UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

# UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

**STRUCTURAL ANALYSIS OF PROTEINS FROM THE RADICAL SAM SUPERFAMILY**

by

**Pedro Cleto Esteves Guerreiro Dinis**

Thesis for the degree of Doctor of Philosophy

November 2015

**UNIVERSITY OF SOUTHAMPTON**

# ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Thesis for the degree of Doctor of Philosophy

**STRUCTURAL ANALYSIS OF PROTEINS FROM THE RADICAL SAM SUPERFAMILY**

Pedro Cleto Esteves Guerreiro Dinis

The Radical SAM superfamily is a large group of enzymes which use an iron-sulfur cluster to catalyse the reductive cleavage of S-adenosylmethionine (SAM), resulting in the formation of a highly reactive intermediate. This potent oxidant is used to functionalise relatively inert substrates to catalyse an extensive role of reactions: cofactor biosynthesis; anaerobic metabolism; methylation and post-translational modifications.

Most members of this family share some structural similarities, most notably a [4Fe-4S] cluster, coordinated by a cysteine triad motif; some conserved motifs for the binding of SAM and an active site with a full or partial triose-phosphate isomerase (TIM) barrel fold. However, despite a quarter of a million predicted protein sequences, the number of structures from which to draw conclusions is severely limited, not reaching 50 available X-ray structures in the Protein Data Bank.

This project aims to solve the structures of three proteins of this superfamily. HydG, an enzyme involved in the biosynthesis of the complex active-site cofactor of the [Fe-Fe] HydA hydrogenase, is proposed to use L-tyrosine and an auxiliary cluster to form a distinctive $Fe(CO)_2CN$ synthon. The structure presented in this work allows a deeper understanding of the mechanism behind the complex product formation, by use of a never before reported [5Fe-5S] cluster.

LipA is an enzyme responsible for the last step on the production of the lipoyl cofactor, the insertion of two sulfur atoms, at C6 and C8 of an octanoyl subunit. The currently accepted mechanism of catalysis has LipA serving as both a catalyst and a substrate, by sacrificing its auxiliary [4Fe-4S] cluster as the sulfur source. With a crystallographic structure recently published, it was possible to identify a residue (serine) coordinating the cluster. Previous mutagenesis studies showed the mutation of this serine has a strong impact on activity, and the crystallization

of a serine to cysteine mutant in this project strengthens the mechanistic proposal being developed.

The third enzyme, Cfr is responsible for the methylation of A2503 in the 23S ribosomal subunit, conferring resistance to several antibiotics, making it a major health concern. Co-crystallization of the protein with an RNA fragment may help direct the efforts to design effective inhibitors, and the first steps in the expression and purification of a small rRNA fragment suitable for crystallization studies are expressed herein.

# Table of Contents

**Table of Contents** ..... **i**

**List of Tables** ..... **xi**

**List of Figures** ..... **xiii**

**DECLARATION OF AUTHORSHIP** ..... **xxiii**

**Acknowledgements** ..... **xxv**

**Definitions and Abbreviations** ..... **xxvii**

**Chapter 1: INTRODUCTION** ..... **1**

1.1 The Radical SAM Superfamily ..... 1

1.1.1 [4Fe-4S] Cluster Interaction With SAM in Radical SAM Enzymes ..... 3

1.1.1.1 Chemical Nature of the [4Fe-4S] Cluster ..... 4

1.1.1.2 SAM:[4Fe-4S] Cluster Binding ..... 5

1.1.1.3 Reduction Potential in the Interaction of SAM with the [4Fe-4S] Cluster ..... 6

1.1.1.4 Propagation of the Radical in Radical SAM Reactions ..... 8

1.1.2 Structure of Radical SAM Enzymes ..... 9

1.1.2.1 SAM Binding Site ..... 12

1.1.2.1.1 SAM Methionyl Moiety ..... 13

1.1.2.1.2 SAM Adenine Moiety ..... 13

1.1.2.1.3 SAM Ribose Moiety ..... 13

1.1.2.2 Substrate Binding Site ..... 14

1.1.2.3 Reductant Binding Site ..... 14

1.1.2.4 Exceptions in Cluster Binding ..... 14

1.1.3 The [4Fe-4S] Cluster ..... 15

1.1.3.1 Auxiliary Clusters in RSAM Enzymes ..... 16

1.2 Hydrogenases and Maturation Pathways ..... 20

1.2.1 The Hydrogenase Family of Enzymes ..... 20

1.2.2 Hydrogenase Subclasses ..... 21

1.2.2.1 [FeFe]-Hydrogenase ..... 24

1.2.2.1.1 The [FeFe]-Hydrogenase Mechanism of Action ..... 25

1.2.2.2 H-Cluster Maturation ..... 26
1.2.2.2.1 Scaffold Protein HydF ..... 27
1.2.2.2.2 Putative Small Ligand Dithiolate Synthase HydE ..... 30
1.2.2.2.3 Tyrosine Lyase HydG ..... 31
1.2.2.3 Understanding the Mechanism of HydG ..... 34
1.2.2.3.1 Mechanistic Relationship Between ThiH and HydG ..... 34
1.2.2.3.2 The Mechanism of Action of HydG ..... 35
1.3 Lipoic Acid and Lipoyl Synthase ..... 40
1.3.1 2—Oxoacid Dehydrogenases ..... 41
1.3.2 Glycine Cleavage System ..... 43
1.3.3 Pathways of Synthesis/Incorporation of Lipoic Acid ..... 45
1.3.3.1 Exogenous Uptake of α-Lipoic Acid ..... 45
1.3.3.2 Endogenous Production of α-Lipoic Acid ..... 45
1.3.4 Sulfur Insertion Enzymes ..... 47
1.3.4.1 Biotin Synthase (BioB) ..... 47
1.3.4.2 Lipoyl Synthase (LipA) and its Mechanism of Action ..... 48
1.4 Antibiotic Resistance Conferring Methyltransferase Cfr ..... 52
1.4.1 Cfr and Antibiotic Multiresistance ..... 54
1.4.2 Mechanism of Action of the RSAM Methyltransferases RlmN and Cfr 56
1.4.3 Structure of RlmN and Cfr ..... 60
**Chapter 2: Structural Characterization of HydG by X-Ray Crystallography ..... 63**
2.1 Introduction ..... 63
2.2 Thermophilic HydGs as a Source of Optimised Homogeneous Protein ..... 65
2.3 HydG Protein Expression ..... 66
2.3.1 Plasmid Design and Expression System ..... 66
2.3.2 Recombinant Protein Expression and Purification ..... 67
2.3.3 Chemical Reconstitution of [4Fe-4S] Clusters in HydG ..... 75
2.4 Preliminary Activity Characterization of Thermophilic HydGs ..... 78
2.5 Initial Crystallization Studies ..... 82
2.5.1 Reproducibility of *Ti*HydG Crystal Formation in 24-Well Format ..... 85

2.5.2 Reproducibility of *Ti*HydG Crystal Formation in 96-Well Format ..... 86
2.5.3 Microseeding to Increase the Reproducibility of *Ti*HydG Crystal Formation in 96-Well Format ..... 87
2.6 Data Collection and Structural Model Building ..... 89
2.6.1 X-Ray Data Collection ..... 89
2.6.2 Model Building and Structural Refinement ..... 95
2.7 HydG Structural Analysis ..... 99
2.7.1 Overall Fold ..... 99
2.7.2 Substrate and Co-Substrate Access Sites ..... 103
2.7.3 Auxiliary Cluster ..... 105
2.7.4 Putative Interaction Site With HydF ..... 107
2.7.5 L-Tyrosine Binding Site ..... 110
2.7.5.1 Molecular Modelling of L-Tyrosine in its Putative Active Site ..... 116
2.8 Structural Comparison With Enzymes From the Amino Acid Lyase Family ..... 117
2.8.1 NosL ..... 117
2.8.2 HydE ..... 123
2.9 Co-Crystallization Studies of *Ti*HydG With L-Tyrosine ..... 128
2.9.1 Attempts to Obtain *Ti*HydG:Tyrosine Co-Crystals by Increasing *Ti*HydG Concentration ..... 128
2.9.2 Crystallization With Low Concentration of *Ti*HydG ..... 133
2.10 EPR Analysis ..... 137
2.11 Mechanistic Update ..... 139
2.12 Conclusions ..... 142

**Chapter 3: Site Directed Mutagenesis of *Ti*HydG ..... 145**

3.1 Introduction ..... 145
3.2 TiHydG His265 Mutagenesis by PCR ..... 148
3.3 TiHydG His265 Mutant Expression and Purification ..... 150
3.3.1 UV-Vis Characterization and Iron Content Analysis of *Ti*HydG Mutants ..... 153
3.3.2 *Ti*HydG His265 Mutant EPR Characterization ..... 154

3.4 TiHydG His265 Mutants' Activity ..... 155
3.5 TiHydG His265Glu Mutant Crystallization and Data Analysis ..... 160
3.5.1 *Ti*HydG His265Glu Structural Analysis ..... 162
3.6 *Ti*HydG His265 Mutants Secondary Structure Estimation by Circular Dichroism165
3.6.1 Principles of Circular Dichroism for Proteins ..... 165
3.6.2 Circular Dichroism Analysis of *Ti*HydG WT and Mutants ..... 165
3.7 Conclusions ..... 171

**Chapter 4: *Ti*HydG Activity With Substrate Analogues ..... 173**

4.1 Introduction ..... 173
4.2 *Ti*HydG Activity With Substrate Analogues ..... 175
4.3 *Ti*HydG Crystallization With Tyrosine Analogues ..... 185
4.3.1 Structural Data and Density Analysis of *Ti*HydG Crystals Grown in the Presence of THIQCA and Tyramine ..... 187
4.3.2 *Ti*HydG Crystallization in the Presence of Different Amino Acids .... 189
4.3.3 Further Studies of *Ti*HydG Crystallization in the Presence of L-Cysteine ..... 190
4.3.4 TiHydG Crystal Soaking in the Presence of Diatomic Ligands ..... 194
4.4 Inorganic Synthesis of $\alpha$-Iminocarboxylic Acid ..... 197
4.4.1 Formation and UV-Spectroscopy of the Fe-DHG Complex ..... 198
4.4.2 HydG Auxiliary Cluster Activity in the Presence of Fe-DHG Complex200
4.5 Summary and conclusions ..... 201

**Chapter 5: Further Insights Into the Structure Characterization of Lipoyl Synthase LipA 205**

5.1 Introduction ..... 205
5.2 Expression and Purification of *Te*LipA2 ..... 208
5.2.1 Chemical Reconstitution of *Te*LipA2 ..... 210
5.3 Crystallization of *TeLipA2* in Higher Glycerol Content ..... 211
5.4 Chemical Reconstitution of *TeLipA*2 Under Different Reducing Conditions .. 214
5.4.1 Crystallization of *TeLipA2* in Different Reconstitution Conditions ... 216

5.5 Chemical Reconstitution of *Te*LipA2 With Sodium Selenide ..... 220

5.5.1 Crystallization of *Te*LipA2 Reconstituted With Selenium ..... 222

5.6 Crystallization of *Te*LipA2 in the Presence of Substrate Analogues ..... 223

5.6.1 Crystallization of *Te*LipA2 in the Presence of MTAN ..... 227

5.7 Mutation of the $[4Fe\text{-}4S]_{Aux}$ Cluster Coordinating Serine in *Te*LipA2 ..... 229

5.7.1 Purification of Expressed *Te*LipA2 Ser283Cys and Ser283Ala Mutants230

5.7.2 Crystallization of *Te*LipA2 Ser283 Mutants..... 234

5.8 Conclusions ..... 238

**Chapter 6: Transcription of an RNA Fragment for Co-Crystallization Studies With 23S Methyltransferase Cfr ..... 239**

6.1 Introduction ..... 239

6.2 RNA Fragment *in vivo* Expression and Purification by Straptividin Affinity Chromatography ..... 241

6.3 RNA Purification by RP-HPLC ..... 245

6.4 RNA Analysis by Mass Spectrometry ..... 246

6.5 RNA purification Using a Commercial Kit ..... 247

6.6 *Ss*Cfr Expression and Purification ..... 248

6.6.1 Chemical Reconstitution of *Ss*Cfr..... 251

6.7 Preliminary *Ss*Cfr Crystallization Studies ..... 252

6.8 Design of a co-Expression System for *Ss*Cfr and its rRNA Substrate ..... 253

6.8.1 Confirmation of the Correct pCDFDuet-1 Constructs..... 255

6.8.2 Co-Expression of Cfr With its rRNA Substrate ..... 258

6.9 Expression of an RNA Fragment Using *in vitro* Transcription..... 259

6.10 Conclusions ..... 261

**Chapter 7: Conclusions and future work ..... 263**

7.1 Hydrogenase Maturase HydG..... 263

7.1.1 Structure of *Ti*HydG ..... 263

7.1.2 Future Work in HydG Crystallography..... 264

7.1.3 Mutagenesis Studies of *Ti*HydG His265 ..... 265

7.1.4 Future Work in *Ti*HydG Mutagenesis ..... 266
7.1.5 Activity Studies With Substrate Analogues ..... 267
7.1.6 Future Work in *Ti*HydG Mutagenesis ..... 267
7.2 Lipoyl Synthase LipA ..... 268
7.2.1 Studies on *Te*LipA2 ..... 269
7.2.2 Future Work in the Mechanism of LipA ..... 270
7.3 Methyltransferase Cfr ..... 270
7.3.1 Studies on *Ss*Cfr and RNA Transcription ..... 270
7.3.2 Future Work in Cfr ..... 271
**Chapter 8: Experimental section ..... 273**
8.1 Principles of Protein Crystallization ..... 273
8.2 The Crystallization Process ..... 273
8.2.1 Improving the Crystallization Conditions ..... 275
8.2.1.1 Salt ..... 275
8.2.1.2 Polymers ..... 276
8.2.1.3 Organic Solvents ..... 277
8.2.1.4 Other Additives ..... 277
8.2.1.5 Temperature and pH ..... 277
8.2.2 Vapour Diffusion Method ..... 278
8.2.3 Cryo-cooling of Crystals ..... 279
8.2.4 Crystal Seeding ..... 279
8.2.5 Matthews Coefficient ..... 281
8.3 X-Ray Diffraction ..... 282
8.3.1 The Structure Factor ..... 284
8.3.2 Solving the Phase Problem ..... 287
8.3.2.1 Molecular Replacement ..... 287
8.3.2.2 Multiple Isomorphous Replacement ..... 287
8.3.2.3 Anomalous Diffraction ..... 287
8.3.2.4 Data Analysis ..... 289
8.3.2.5 Data Refinement ..... 290
8.4 Materials ..... 292

8.5 Equipment ..... 295
8.5.1 DNA/RNA amplification and analysis ..... 295
8.5.2 Bacterial growth ..... 295
8.5.3 pH determination ..... 295
8.5.4 Anaerobic protein purification ..... 296
8.5.5 SDS-PAGE analysis ..... 296
8.5.6 Centrifugation ..... 296
8.5.7 UV-Vis spectroscopy ..... 296
8.5.8 CD spectroscopy ..... 297
8.5.9 HPLC Analysis ..... 297
8.5.10 Crystallography ..... 297
8.6 Media, buffers and solutions ..... 297
8.7 General Microbiological Techniques ..... 309
8.7.1 Small scale culture ..... 310
8.7.2 *E.coli* cell competence by treatment with $CaCl_2$ ..... 310
8.7.3 Competent cell transformation ..... 310
8.7.4 Isolation of plasmid DNA (miniprep) ..... 311
8.7.5 Analytical digestion of purified DNA ..... 311
8.7.6 DNA-vector ligation ..... 311
8.7.7 DNA agarose gel electrophoresis ..... 312
8.7.7.1 RNA Bleach Agarose Gel Electrophoresis ..... 312
8.7.7.2 RNA Formaldehyde Agarose Gel Electrophoresis ..... 312
8.8 Expression and purification of RSAM enzymes ..... 312
8.8.1 Large scale expression ..... 312
8.8.2 Auto-induction large scale expression ..... 313
8.8.3 General purification ..... 313
8.8.4 Cfr purification of RSAM proteins ..... 314
8.8.5 Protein estimation by Bradford quantification ..... 315
8.8.6 15% SDS-PAGE denaturing gel ..... 316
8.9 Chemical reconstitution and iron content analysis of RSAM enzymes ..... 316
8.9.1 Chemical reconstitution of dual [4Fe-4S] cluster containing proteins 316

8.9.2 Chemical reconstitution of single [4Fe-4S] cluster containing proteins317
8.9.3 Iron quantification by the modified method of Fish ..... 317
8.10 Activity assay of HydG proteins ..... 318
8.10.1 Cyanide fluorescence conversion ..... 319
8.11 Screening of crystallization conditions with Hampton PEGRx HT and/or Molecular Dimensions PACT Premier commercial broad screens ..... 320
8.11.1 Manual reproducibility of the identified crystallization conditions. 321
8.12 Reproducibility of crystallization conditions in the 24-well format and 96-well format ..... 322
8.13 Microseeding using seed beads ..... 322
8.14 Crystallization of concentrated *Ti*HydG with L-tyrosine ..... 322
8.15 Crystallization of *Ti*HydG, under different concentrations of protein and SAM323
8.16 Crystallization of 0.3 mM *Ti*HydG with 12 mM L-tyrosine ..... 324
8.17 Primer extension PCR ..... 325
8.18 Circular dichroism analysis of *Ti*HydG mutants ..... 326
8.19 *Ti*HydG crystallization with L-amino acids and reconstitution additives ..... 327
8.20 Soaking of pre-grown crystals with potential ligands ..... 328
8.21 Synthesis of iron-bound $\alpha$-iminocarboxylic acids ..... 328
8.21.1 Synthesis of iron-bound $\alpha$-iminocarboxylic acids, with lower concentration of substrates ..... 329
8.21.2 *Ti*HydG Activity Assay in the Presence of $\alpha$-iminocarboxylic acid ... 329
8.22 *Te*LipA2 Crystallization With High Glycerol Content ..... 330
8.23 *Te*LipA2 Chemical Reconstitution With Different Reducing Agents ..... 331
8.24 *Te*LipA2 Crystallization in the Presence of MTAN ..... 331
8.25 RNA Extraction by Guanidine thiocyanate Chloroform:Phenol ..... 331
8.26 RNA purification by Streptactin Affinity Chromatography ..... 332
8.27 RNA Analysis by HPLC ..... 332

**Appendix A ..... 333**

A.1 Plasmid maps for studies in HydG ..... 333
A.2 Standard curves for the quantification of products in *Ti*/*Tl* HydG activity assays ..... 338

A.3 Crystallization data for *Ti*HydG WT crystals ..... 339
A.4 Crystallization data for *Ti*HydG WT crystals with 6 mM L-tyrosine ..... 346
A.5 Crystallization data for low protein (0.3 mM) *Ti*HydG WT crystals with 12 mM L-tyrosine ..... 348
A.6 *Ti*HydG mutant activity ..... 350
A.7 *Ti*HydG His265Glu crystallization ..... 351
A.8 *Ti*HydG activity with substrate analogues ..... 354
A.9 *Ti*HydG crystallization with THIQCA ..... 355
A.10 *Ti*HydG crystallization with tyramine ..... 358
A.11 *Ti*HydG crystallization with L-cysteine ..... 361
A.12 *Ti*HydG crystallization with CO/CN ..... 364
A.13 Quantification of cyanide formation by *Ti*HydG in the presence of an a-iminocarboxylic acid complex ..... 368
A.14 Plasmid map of *Te*LipA2 ..... 369
A.15 *Te*LipA 2 crystallization with 2.5 molar equivalents of $FeCl_3$/$Na_2S$ ..... 371
A.16 Crystallization of *Te*LipA 2 reconstituted with sodium selenide ..... 376
A.17 Crystallization of *Te*LipA 2 with substrate analogues ..... 379
A.18 Crystallization of *Te*LipA 2 with substrate analogue ESDK(o)AD and MTAN ..... 383
A.19 Crystallization of *Te*LipA 2 Ser283Cys ..... 387
A.20 Gene encoding the Cfr-substrate rRNA Fragment ..... 390
A.21 Plasmid map of *cfr* containing pCDFD1 vector ..... 392

**List of References ..... 395**

# List of Tables

**Table 1 Molecular weights of ISC proteins** ..... 74

**Table 2 Protein yield estimation from different HydG purifications.** ..... 75

**Table 3 Iron quantification of HydG samples before and after each reconstitution step.** ..... 76

**Table 4 Quantification of product formation by *Tl*HydG and *Ti*HydG.** ..... 81

**Table 5 Conditions corresponding to the wells shown in Figure 45** ..... 84

**Table 6 EDNA strategies for the collection of complete datasets of the *Ti*HydG crystals tested.** 90

**Table 7 Data collection statistics. Highest resolution shell shown in parenthesis** ..... 93

**Table 8 Refinement statistics for the crystallographic dataset of *Ti*HydG, following refinement.** ..... 98

**Table 9 Score values for molecular replacement using *Phaser*.** ..... 130

**Table 10 Refinement statistics for the crystallographic datasets of *Ti*HydG with6.0 mM L-tyrosine, following refinement.** ..... 130

**Table 11 Score values for molecular replacement using *Phaser*.** ..... 136

**Table 12 Protein yields and iron quantification from *Ti*HydG His265Glu and *Ti*HydG His265Ala.** ..... 153

**Table 13 Quantification of product formation by *Ti*HydG WT and mutants.** ..... 156

**Table 14 Predicted secondary structure values obtained by the deconvolution of the circular dichroism data for *Ti*HydG wild-type and mutants.** ..... 170

**Table 15 Quantification of product formation by *Ti*HydG with different substrate analogues.** 179

**Table 16 Protein and iron content estimation at different stages of *TeLipA*2 purification.** ... 210

**Table 17 Glycerol concentration in crystallization experiment with *Te*LipA2.** ..... 212

**Table 18 Protein and iron content estimation at different stages of *TeLipA*2 mutants' purification.** ..... 233

**Table 19 Protein estimation from *Ss*Cfr purification** ..... 249

**Table 20 Score values for molecular replacement using *Phaser*.** ..... 389

# List of Figures

**Figure 1 [4Fe-4S]: SAM interaction and reductive cleavage.** ..... 4

**Figure 2 Radical formation and stabilization from 3',4'-anhydroadenosyl-l-methionine.** ..... 5

**Figure 3 Energetics of reductive SAM cleavage at the active site of lysine-2,3-aminomutase (LAM).** ..... 7

**Figure 4 TIM barrel structural comparison.** ..... 10

**Figure 5 Conserved core-fold of members of the RSAM superfamily of enzymes.** ..... 11

**Figure 6 Conserved SAM binding motifs of the Radical SAM superfamily.** ..... 12

**Figure 7 Biosynthetic pathway for the introduction of the [4Fe-4S] cluster into a generic apo-enzyme by the ISC machinery.** ..... 16

**Figure 8 SPASM/Twitch RSAM structure comparison.** ..... 17

**Figure 9 Sequence alignment of different RSAM that possess auxiliary iron-sulfur clusters.** ..... 19

**Figure 10 Organometallic cofactor and protein structure of the three different main classes of hydrogenases.** ..... 23

**Figure 11 Proposed mechanism for $H_2$ production in the active site of [FeFe]-hydrogenases.** 26

**Figure 12 Structure of HydF.** ..... 28

**Figure 13 Proposed biosynthesis pathway for HydA H Cluster.** ..... 29

**Figure 14 Proposed HydG catalysed mechanism of L-tyrosine cleavage and diatomic product formation.** ..... 33

**Figure 15 General mechanism of biosynthesis of thiamine phosphate.** ..... 34

**Figure 16 Alternative pathways of tyrosine reductive cleavage.** ..... 36

**Figure 17 Structure of NosL with its substrate, tryptophan, bound.** ..... 37

**Figure 18 Mechanism of CO/$^{-}$CN formation from DHG through cluster-independent deprotonation.** ..... 38

**Figure 19 SF-FTIR and EPR spectroscopies following the formation of iron-bound diatomic products of HydG.** ..... 39

**Figure 20 α-Lipoic acid, in its reduced (dihydro-α-lipoic acid) and oxidized forms.** ..... 40

**Figure 21 E1 subunit catalyses the decarboxylation of a α-ketoacid.** ..... 42

**Figure 22 Reaction mechanism of the glycine cleavage system.** ..... 44

**Figure 23 Pathways of lipoic acid synthesis and uptake.** ..... 46

**Figure 24 Proposed mechanism of formation of the thiophene ring in biotin by BioB.** ..... 48

**Figure 25 Formation of a deuterium-labelled monothiolated intermediate by LipA.** ..... 49

**Figure 26 Proposed mechanism of action by lipoyl synthase (LipA) for lipoic acid synthesis.** 51

**Figure 27 Antibiotics and their interaction with the 50S bacterial ribosomal subunit.** ..... 53

**Figure 28 Classes of RSAM methyltransferases and methylation targets.** ..... 55

**Figure 29 Proposed mechanism of Cfr methylation reaction, with aromatization and hydride shift.** ..... 57

**Figure 30 Proposed cysteinyl radical based methylation of RNA by Cfr, with different potential mechanisms for adduct resolution.** ..... 59

**Figure 31 Structure of RlmN and modelled structure of Cfr.** ..... 61

**Figure 32 Two step mechanism of tyrosine cleavage by HydG/ThiH.** ..... 64

**Figure 33 Phylogenetic tree of different thermophilic organisms, based on HydG primary protein sequences present.** ..... 66

**Figure 34 Co-expression vector (pRD003) for *Thermoanaerobacter italicus* HydG with the *E.coli isc* operon.** ..... 67

**Figure 35 Schematic representation of the purification steps for RSAM enzymes.** ..... 68

**Figure 36 HydG purification chromatography.** ..... 70

**Figure 37 15% SDS-PAGE analysis of the purification stages of *Tl*HydG.** ..... 71

**Figure 38 15% SDS-PAGE analysis of the purification stages of *Tm*HydG.** ..... 72

**Figure 39 15% SDS-PAGE analysis of the purification of *Ti*HydG.** ..... 73

**Figure 40 UV-vis spectra of HydG proteins, before and after reconstitution steps.** ..... 77

**Figure 41 Time dependent SAM and tyrosine cleavage by *Ca*HydG wild type enzyme.** .....78

**Figure 42 Identification of reaction products of HydG catalysed tyrosine cleavage by RP-HPLC.** .....79

**Figure 43 RT-HPLC chromatogram of the activity assays for *Tl*HydG and *Ti*HydG.** .....80

**Figure 44 Examples of the crystallization wells from the screening of HydG from *Thermotoga lettingae*** .....83

**Figure 45 Crystallization wells of *Ti*HydG** .....84

**Figure 46 24-Well replication of previous crystal forming conditions.** .....85

**Figure 47 Screening wells that showed crystal formation of *Ti*HydG with crystallization solutions mixed in house.** .....86

**Figure 48 Microseeding results.** .....88

**Figure 49 X-ray diffraction of *Ti*HydG crystal F1.** .....92

**Figure 50 Partial structure solution obtained by AutoSol.** .....96

**Figure 51 Structural model of *Ti*HydG as the result of automated model building and refinement with Autobuild.** .....97

**Figure 52 Polygon analysis of *Ti*HydG F1 crystal structure.** .....99

**Figure 53 Dimeric structure of HydG from *Thermoanerobacter italicus*.** .....100

**Figure 54 2-dimensional topology of HydG, revealing the secondary structure.** .....101

**Figure 55 Conserved SAM binding motifs of the RAM superfamily, as present in *Ti*HydG.** ...102

**Figure 56 *Ti*HydG conserved [4Fe-4S] radical SAM cluster and ligands.** .....103

**Figure 57 Potential substrates' entrance sites.** .....104

**Figure 58 Octahedral coordination of the labile iron as observed in the crystal structure of *Ti*HydG.** .....106

**Figure 59 *Ti*HydG auxiliary cluster and ligands.** .....107

**Figure 60 B-factor changes in $\alpha$-helix 391-398 near *Ti*HydG auxiliary cluster, caused by the change in interactions between three residue partners.** .....110

**Figure 61 Ligand binding sites for *Ti*HydG structure, identified by Metapocket.** ..... 111

**Figure 62 Model of *Ti*HydG first putative substrate binding site.** ..... 112

**Figure 63 Overlay of the[4Fe-4S]$_{RS}$ cluster with SAM and the substrate of different RSAM enzymes.** ..... 113

**Figure 64 Model of *Ti*HydG second putative substrate binding site.** ..... 114

**Figure 65 Model of *Ti*HydG TIM barrel-derived tunnel.** ..... 115

**Figure 66 Composite model of tyrosine and SAM binding to HydG.** ..... 117

**Figure 67 Sequence alignment between *Ti*HydG and *Sa*NosL.** ..... 118

**Figure 68 Structural comparison between *Ti*HydG and *Sa*NosL** ..... 120

**Figure 69 Proposed mechanism for NosL.** ..... 121

**Figure 70 Structural comparison between the internal TIM barrel of *Ti*HydG and *Sa*NosL along the TIM barrel axis.** ..... 122

**Figure 71 *Tm*NosL active site.** ..... 123

**Figure 72 Sequence alignment between *Ti*HydG and *Tm*HydE.** ..... 124

**Figure 73 Structural comparison between *Ti*HydG and *Tm*HydE.** ..... 126

**Figure 74 Putative active site for *Tm*HydE.** ..... 127

**Figure 75 Co-crystallization of *Ti*HydG with 6 mM tyrosine** ..... 129

**Figure 76 POLYGON analysis of *Ti*HydG crystal formed in the presence of 6.0 mM L-tyrosine.** 131

**Figure 77 Structure of *Ti*HydG grown in the presence of 6 mM L-tyrosine.** ..... 132

**Figure 78 Microseeding experiment with lower concentration of *Ti*HydG** ..... 133

**Figure 79 Occurrence of crystal formation in lower *Ti*HydG and SAM concentration.** ..... 134

**Figure 80 *Ti*HydG crystal formed in high L-tyrosine content** ..... 135

**Figure 81 Structural information obtained for the low concentration crystal of *Ti*HydG (0.3 mM), grown with 3.0 mM SAM and 12.0 mM L-tyrosine.** ..... 137

**Figure 82 EPR spectroscopic studies of *So*HydG.** ..... 138

**Figure 83 Proposed mechanism of action of HydG, divided onto two cycles.** .....141

**Figure 84 Conservation of the labile iron coordinating histidine, and potential target for mutagenesis.** .....147

**Figure 85 Primer extension PCR.** .....148

**Figure 86 PCR fragment amplification for the creation of *Ti*HydG mutants.** .....149

**Figure 87 15% SDS-PAGE gels of samples obtained in the purification steps of *Ti*HydG mutant *Ti*HydG His265Glu.** .....151

**Figure 88 15% SDS-PAGE gels of samples obtained in the purification steps of *Ti*HydG mutant *Ti*HydG His265Ala.** .....152

**Figure 89 UV spectra of *Ti*HydG mutants together with the calculated 410/280 values.** .....154

**Figure 90 EPR spectra of *Ti*HydG His265Glu.** .....155

**Figure 91 Product quantification of the enzymatic reaction of HydG WT and His mutants.** .156

**Figure 92 HydG catalysed L-tyrosine cleavage, with futile cycle.** .....158

**Figure 93 Examples of crystallization wells showing the formation of *Ti*HydG His265Glu crystals.** .....161

**Figure 94 POLYGON analysis of the refinement data for *Ti*HydG His265Glu crystal.** .....162

**Figure 95 Structural elements of *Ti*HydG His265Glu.** .....164

**Figure 96 Potential model for *Ti*HydG His265Ala aggregation, by stacking of C-terminal β-sheets.** .....168

**Figure 97 Circular dichroism spectroscopic data (green) and predicted modelling (blue) using the K2d algorithm [611].** .....169

**Figure 98 Tyrosine analogues previously used in ThiH and HydG activity assays.** .....174

**Figure 99 New tyrosine analogues tested in HydG activity assays**.....175

**Figure 100 Mechanism for product formation by HydG, together with potential by-product pathways**. .....177

**Figure 101 Quantification of products in *TiHydG* activity assay in the presence of L-tyrosine analogues.** .....178

**Figure 102 Proposed mechanism for tyramine radical mediated cleavage.** ..... 182

**Figure 103 Proposed mechanism for L-DOPA radical mediated cleavage.** ..... 183

**Figure 104 Potential mechanisms of L-tyrosinamide radical cleavage.** ..... 184

**Figure 105 *Ti*HydG crystal formation in the presence of substrate analogues** ..... 186

**Figure 106 Model and density map of *Ti*HydG obtained from crystals grown in the presence of substrate analogues.** ..... 188

**Figure 107 Crystal formation with amino acids and additives, and respective diffraction pattern.** ..... 190

**Figure 108 Potential mechanism of cysteine mediated transfer for the $Fe(CO)_2(CN)$ synthon.** 191

**Figure 109 Model and density map of *Ti*HydG crystal soaked with L-cysteine.** ..... 193

**Figure 110 Soaking experiment of *Ti*HydG crystals with the diatomic products cyanide and carbon monoxide.** ..... 194

**Figure 111 Anomalous signal surrounding both clusters of different crystals of *Ti*HydG soaked with increasing concentrations of KCN.** ..... 196

**Figure 112 Ferrous $\alpha$-iminocarboxylic complexes** ..... 197

**Figure 113 Dehydroglycine:Fe complex** ..... 198

**Figure 114 UV Spectra following the formation of Fe-DHG** ..... 199

**Figure 115 UV spectra of the formation of Fe-DHG complex, with lower concentration of substrates.** ..... 200

**Figure 116 Formation of cyanide, by *Ti*HydG in the presence of the Fe complex.** ..... 201

**Figure 117 Structure superposition of LipA (green) and BioB (red).** ..... 206

**Figure 118 Structure comparison *Te*LipA2 crystals.** ..... 207

**Figure 119 SDS-PAGE gels of samples obtained in the purification of *TeLipA*2.** ..... 209

**Figure 120 UV-vis spectra of *Te*LipA2.** ..... 211

**Figure 121 Crystals obtained for *TeLipA2* in the presence of increasing concentrations of glycerol** ..... 212

**Figure 122 Crystal diffraction from *TeLipA2* grown in the presence of increasing concentrations of glycerol** ..... 213

**Figure 123 UV spectra of *TeLipA2* under different reducing conditions, together with the calculated 410/280 values.** ..... 215

**Figure 124 Crystallization of *Te*LipA2 $Fe_{low}$.** ..... 217

**Figure 125 POLYGON analysis of the *Te*LipA2 $Fe_{low}$ crystal** ..... 218

**Figure 126 Model and density map of *Te*LipA2 $Fe_{low}$.** ..... 219

**Figure 127 UV spectra of *TeLipA2* reconstituted with iron(III) chloride and sodium selenide** 221

**Figure 128 Crystallization of *Te*LipA2, where the reconstitution was carried out with $FeCl_3$ and $Na_2Se$.** ..... 222

**Figure 129 Two small peptide analogues of the native substrate of LipA.** ..... 224

**Figure 130 Crystals of *TeLipA2* grown in the presence of SAM and a synthetic substrate** ..... 225

**Figure 131 Model and density maps of *TeLipA2* co-crystallized with substrate peptide analogues** ..... 226

**Figure 132 Model and density map of *TeLipA2* obtained from a crystal grown in the presence of MTAN and ESKD(o)AD.** ..... 228

**Figure 133 SDS-PAGE gels of samples obtained in the purification of *TeLipA2* Ser283Ala** ..... 231

**Figure 134 SDS-PAGE gels of samples obtained in the purification of *TeLipA2* Ser283Cys** ..... 232

**Figure 135 UV spectra foe *Te*LipA2 Ser283 mutants** ..... 234

**Figure 136 Crystallization of *Te*LipA Ser283 mutants.** ..... 235

**Figure 137 POLYGON analysis of the *Te*LipA2 Ser283Cys crystal** ..... 236

**Figure 138 Structure comparison of native *TeLipA2* with the Ser283Cys mutant.** ..... 237

**Figure 139 23S rRNA secondary structure (3' half)** ..... 240

**Figure 140 Gene for the transcription of an RNA fragment.** ..... 241

**Figure 141 Chromatographic run of a total RNA sample on a Streptactin column.** ..... 242

**Figure 142 Bleach Agarose RNA gel** ..... 243

**Figure 143 Formaldehyde agarose RNA gel.** ..... 244

**Figure 144 Chromatographic profile obtained for the total RNA sample in a C4 reverse column.** ..... 245

**Figure 145 Chromatographic profile and mass analysis obtained for the Streptactin purified RNA sample.** ..... 246

**Figure 146 Formaldehyde Agarose RNA gel for the samples obtained using a commercial RNA kit** ..... 247

**Figure 147 15% SDS-PAGE analysis of the purification fractions of *Ss*Cfr** ..... 250

**Figure 148 UV-vis spectra of *Ss* Cfr** ..... 251

**Figure 149 Needle-like crystals obtained for *Ss*Cfr.** ..... 252

**Figure 150 Testing of *Ss*Cfr needle-like crystal formation.** ..... 253

**Figure 151 Mutagenesis primers.** ..... 254

**Figure 152 Agarose analysis of the PCR amplification of the RNA fragment encoding gene.** 254

**Figure 153 Agarose gel analysis of the PCR amplification for the RNA encoding gene.** ..... 255

**Figure 154 Restriction enzyme digestion analysis by agarose gel.** ..... 257

**Figure 155 15% SDS-PAGE gels of samples obtained in the purification of *Ss* Cfr co-expressed with 23S rRNA fragment.** ..... 259

**Figure 156 *in vitro* transcription** ..... 260

**Figure 157 2D-Phase Diagram.** ..... 274

**Figure 158 Protein solubility as a function of salt concentration.** ..... 276

**Figure 159 Structure of commonly used polymers in macromolecular crystallization** ..... 277

**Figure 160 Most common crystallization setups for vapour diffusion.** ..... 278

**Figure 161 Example of a successful micro seeding experiment.** ..... 280

**Figure 162 2D-phase model with and without the addition of crystal seeds.** ..... 281

**Figure 163 Bragg's law and consequent diffraction** ..... 283

**Figure 164 Representation of unit cell, crystal lattice and a plane with miller indices.** ..... 284

**Figure 165 General process of data processing.**.....289

# DECLARATION OF AUTHORSHIP

I, Pedro Cleto Dinis declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

Structural analysis of proteins from the Radical SAM superfamily

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Part of this work has been published as:

Harmer, J.; Hiscox, M.; Dinis, P.; Fox, S.; Iliopoulos, A.; Hussey, J.; Sandy, J.; Van Beek, F.; Roach, P. (2014) Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *Biochem J*, *464*(1), 123-133

Dinis, P., Suess, D.; Fox, S.; Harmer, J.; Driesener, R.; De La Paz, L.; Roach, P. (2015). X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *PNAS*, *112*(5), 1362-1367.

Signed: .....

Date:.....

# Acknowledgements

Thank you to everyone that has accompanied me during this doctoral work. This is a much mine as it is yours.

I would particularly like to thank Prof. Peter Roach, my supervisor, for his expert advice and encouragement throughout the development of this work. I am truly grateful for all his guidance and patience. I am particularly humbled by the freedom I encountered in pursuing my own set of experiments, even when doom was looming. His door was always opened, and his excitement over new ideas is truly contagious.

A big thank you, of course, to the entire Roach group for their camaraderie. I must give special mention to Dr. Jenny Harmer, Dr. Martin Hiscox and Dr. Rebecca Driesener, who took me under their wing and gave me guidance, without which, none of this project could ever be possible.

I would also like to thank Prof. Jon Essex and Dr. Stephen Fox, our collaborators in creating the substrate model of HydG. Their insights regarding the possible mechanism of the enzyme were truly appreciated and have helped frame the chemical context within the active site of the enzyme.

Dr. Jim Spencer and Dr. Martin Challand have been most generous in helping setting up the studies in Cfr and without them no work on that project would have been possible.

I must also mention Prof. David Britt and Dr. Daniel Suess. The fortuitous meeting at the Iron-Sulfur Cluster conference has proven more than beneficial, and the knowledge they share in both EPR and the HydG system made for amazing discussions.

I am incredibly grateful to Rachel Leahy for her endless help with the English revision. Thank you so much for reviewing the text in your own personal time and for caring.

To Emanuel Dutra, an insightful meteorologist, as well as Carina Santos for all the pep talks.

To António Guerreiro, for being a model of what one can accomplish in life.

To my parents: muito obrigado por tudo o que fizeram por mim para eu poder chegar aqui.

This work is dedicated to my wife, Mariana. All in life, the good and bad, is meant to be shared; for no memories are lost when I am near you and every effort is emboldened by your presence.

# Definitions and Abbreviations

| | |
|---|---|
| $[4Fe\text{-}4S]_{RS}$ | SAM binding [4Fe-4S] cluster |
| $[4Fe\text{-}4S]_{Aux}$ | Auxiliary [4Fe-4S] cluster |
| $\varepsilon_{xxx}$ | Extinction coefficient at XXX nm |
| ΔCTD | C-terminal deletion mutant |
| Aa | amino acids |
| 4-HPHPA | p-hydroxyphenyl-α- hydroxypropionic acid |
| 4-HPPA | p-hydroxyphenylpropionic acid |
| ACP | Acyl carrier protein |
| AD | Anomalous diffraction |
| ADH | Acetoin dehydrogenase complex |
| AMP | Adenosyl monophosphate |
| anSME | Anaerobic sulfatase maturating enzyme |
| ATP | Adenosine triphosphate |
| BCDH | Branched chain oxoacid dehydrogenase complex |
| BioB | Biotin synthase |
| BME | β-mercaptoethanol |
| BtrN | Butirosin biosynthetic enzyme |
| BSA | Bovine serum albumin |
| *Ca* | *Clostridium acetobutylicum* |
| CBI | Fluorescent 1-cyanobenz[f]isoindole |
| CD | Circular dichroism |
| Cfr | Chloramphenicol-florfenicol resistance enzyme |
| CofH | Cofactor F420 biosynthesis enzyme |

| | |
|---|---|
| CRW | Comparative RNA Website and project |
| DFT | Density functional theory |
| DHG | Dehydroglycine |
| DLS | Diamond Light Source |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DOA | 5'-methyl of 5'-deoxyadenosine |
| $DOA^{\bullet}$ | SAM-derived 5'-deoxyadenosyl radical |
| DOIA | 2-deoxy-*scyllo*-inosamine |
| DE3 | BL21-Star cells |
| DTB | Dethiobiotin |
| DTH | Dithionite |
| DTT | Dithiothreitol |
| E1 | 2-Oxoacid dehydrogenase |
| E2 | Dihydrolipoyl acyltransferase |
| E3 | Dihydrolipoyl dehydrogenase |
| EDNA | Characterization and strategy software |
| EDTA | Ethylenediaminetetraacetic acid |
| ENDOR | Electron-nuclear double resonance spectroscopy |
| EPR | Electron paramagnetic resonance spectroscopy |
| ESRF | European Synchrotron Radiation Facilities |
| $FAD/FADH_2$ | Flavin adenine dinucleotide |
| FTIR | Fourier transform infrared spectroscopy |
| GDP | Guanosine-5'-diphosphate |

| | |
|---|---|
| GCV | Glycine cleavage system |
| GGE | Glycine rich motif |
| GNTC | Guanidine thiocyanate |
| GTP | Guanosine-5’-triphosphate |
| HCNAbc | Cyanide synthase |
| HPLC-MS | High-Performance Liquid chromatography–mass spectrometry |
| H-protein | Lipoamide-containing carrier protein |
| HTHIQCA | L-7-Hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid |
| HydA | [Fe-Fe] hydrogenase |
| HydE | [FeFe] hydrogenase H-cluster radical SAM maturase |
| HydF | [FeFe] hydrogenase H-cluster GTPase scaffold maturase |
| HydG | [FeFe] hydrogenase H-cluster radical SAM maturase/tyrosine lyase |
| HYSCORE | Hyperfine sublevel correlation |
| IP-RP-HPLC | Ion Pair Reverse Phase High Performance Liquid Chromatography |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| ISC | Iron Sulfur Cluster |
| ITC | Isothermal calorimetry |
| $k_{cat}^{app}$ | Apparent turnover number |
| LA | α-lipoic acid |
| LAM | Lysine 2,3-aminomutase |
| LC-MS | Affinity purification with streptavidin |
| LCP | Lipoyl-carrier protein |
| LD | Lipoyl domain |
| L-DOPA | L-3,4-dihydroxyphenylalanine |

| | |
|---|---|
| LLG | Log likelihood gain |
| LipA | Lipoyl synthase |
| LipB | Octanoyltransferase |
| LplA | Lipoate protein ligase A |
| LplB | Lipoate protein ligase B |
| L-protein | dihydrolipoamide dehydrogenase |
| (L-) Tyr | (L-)Tyrosine |
| MAD | Multiple wavelength anomalous diffraction |
| MCS | Multiple cloning site |
| MiaB | tRNA methylthiotransferase |
| MoaA | Mo cofactor biosynthesis protein A |
| MPD | 2-methyl-2,4-pentanediol |
| MR | Molecular replacement |
| mRNA | Messenger RNA |
| MRSA | Methicillin-resistant *Staphylococcus aureus* |
| MTA | Methylthioadenosine |
| MTAN | 5'-methylthioadenosine/S-adenosyl-L-homocysteine nucleosidase |
| MTases | Methyltransferases |
| $NAD^+$/NADH | Nicotinamide adenine dinucleotide |
| NMR | Nuclear magnetic resonance |
| NosL | Tryptophan lyase |
| ODH | 2-oxoglutarate dehydrogenase complex |
| pCD003 | Alanine mutant plasmids |
| pCD004 | Glutamic acid mutant plasmids |

| | |
|---|---|
| PCR | Polymerase chain reaction |
| PDB | Protein Data Bank |
| PDH | Pyruvate dehydrogenase complex |
| PEG | Polyethyelene glycol |
| PFL-AE | Pyruvate formate-lyase-activating enzyme |
| PLP | Pyridoxal 5'-phosphate |
| P-protein | glycine dehydrogenase |
| PTC | Peptidyl transferase centre |
| RimO | Ribosomal protein S12 methylthiotransferase |
| RlmN | 23S rRNA m2A2503 methyltransferase |
| RMSD | Root-mean-square deviation |
| RNA | Ribonucleic acid |
| RNAses | Ribonucleases |
| RNR | Ribonucleotide reductase |
| *So* | *Solfalobus solfataricus* |
| RP-HPLC | Reversed phase High-Performance Liquid chromatography |
| rrn | Ribosomal RNA alleles |
| rRNA | Ribosomal RNA |
| RSAM | Radical SAM |
| RT | Retention time |
| SAD | Single wavelength anomalous diffraction |
| SAH | S-adenosyl-L-homocysteine |
| SAM | S-adenosyl-L-methionine |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |

| | |
|---|---|
| SHMT | Serine hydroxymethyltransferase |
| SF-FTIR | Stopped-Flow FTIR |
| *Ss* | *Staphylococcus sciuri* |
| TCA | Citric acid cycle |
| TCEP | tris(2-carboxyethyl)phosphine |
| TF Z-score | Translation function Z-score |
| THF | Tetrahydrofolate |
| ThiH | Tyrosine lyase |
| *Ti* | *Thermoanerobacter italicus* |
| TIM | Triosephosphate isomerase |
| *Tl* | *Thermotoga lettingae* |
| *Tm* | *Thermotoga maritima* |
| TPP | Thiamine pyrophosphate |
| T-protein | Aminomethyltransferase |
| tRNA | Transfer RNA |
| UV-Vis | Ultraviolet-Visible |
| VRSA | Vancomycin-resistant *Staphylococcus aureus* |
| XDS | X-ray Detector Software |
| WT | Wild type |

# Chapter 1: INTRODUCTION

## 1.1 The Radical SAM Superfamily

In 1897 Buchner introduced the concept of enzymes [1], which provided the basis for understanding how the majority of chemical reactions are catalysed in all domains of life. Less than a century later, the first enzymatic mechanism that used radical based chemistry was identified for ribonucleotide reductase (RNR) [2]. Radical reactions are an important concept in chemistry and the term radical refers to the presence of at least one unpaired valence electron: in many situations, the free radical shows a high degree of reactivity towards other substances [3]. By using the high reactivity of radicals in a controlled environment, namely the active sites of enzymes, nature has found in radical based enzymes a suitable mechanistic pathway to catalyse chemically challenging reactions [4].

Many of the enzymatic reactions which involve radical intermediates require a metal and/or an organic cofactor to facilitate the radical-based transformation of substrates. Three enzymatic subclasses can be readily defined within this large group [4]:

i) Enzymes where the radical is formed in the substrate, such as the radical-ion intermediates in DNA photolyase [5]
ii) Enzymes where the radical is formed at a residue of the protein. The mechanism commonly involves glycine [6-8], but also tyrosine [9-11] and cysteine [12-15]
iii) Enzymes which use cofactors for the formation of radicals, of which coenzyme B12 dependent enzymes are one of the most well understood examples [16-19]

An additional family, the radical SAM (RSAM) family, contains examples that could appear in each of these subclasses [20], making it a fascinating family of enzymes. All members of the RSAM family possess a [4Fe-4S] cluster that reductively cleaves S-adenosyl-L-methionine (SAM) for the formation of a deoxyadenosyl radical [21-25]. In some instances, SAM can be used as a true cofactor, being regenerated at the end of the catalytic cycle [26, 27]. Most often, however, SAM is used as a co-substrate, being consumed in the reaction [28]. The radical attack from the SAM-derived 5'-deoxyadenosyl radical (DOA$^{\bullet}$) usually occurs directly onto the substrate by hydrogen atom abstraction, though the formation of a protein-residue based radical has also been reported [29-32].

There is a clear relationship between the RSAM family and the coenzyme B12 enzyme family [33] as both make use of the 5'-deoxyadenosyl radical as a highly reactive intermediate. However, with upwards of 250,000 predicted sequences being currently available in the SUPERFAMILY database [34], the RSAM superfamily has proved to be much larger and widespread than the coenzyme B12 enzymes (20,000 sequences in the SUPERFAMILY database).

Unsurprisingly, given the sheer number of potential enzyme activities that could result from these sequences, only a small subsection of enzymes from the RSAM superfamily have been studied, with less than 100 biosynthetic pathways subjected to experimental studies of RSAM enzyme activity [20, 35]. Members of the RSAM superfamily have been shown to catalyse a diverse and interesting list of reactions: methylation [36], methylthiolation [37], thioether bond formation [38], sulfur insertion [39], ring formation [40], radical-mediated ring contraction [41], oxidative decarboxylation [42], adenosine addition [43], enzymatic activation [44, 45], fragmentation [46, 47] and other complex rearrangement mechanisms which have been widely reviewed [21, 48-53].

The first member of the RSAM superfamily was identified in 1970 [54] as an enzyme which contained iron (Fe) and required SAM and anaerobic incubation with a reducing agent to catalyse its reaction (Figure 1). Soon, other enzymes which shared these characteristics were discovered [55-58], and some structural similarities were identified, namely a cysteine motif CxxxCx$\phi$C ($\phi$ being an aromatic residue), responsible for binding of an iron-sulfur cluster ([Fe-S]) [59, 60]. However, a key step in the recognition of the RSAM superfamily came with a rigorous bioinformatics analysis in 2001 [61], with two conserved regions being clearly identified: the aforementioned cysteine-conserved motif and an additional conserved glycine-rich SAM binding site, the latter also being observed in SAM dependent methyltransferases [62].

The incorporation of [Fe-S] clusters in these proteins, bound to the cysteine-conserved motif, has an extraordinary impact on these proteins, both structurally (see chapter 1.1.2), as well as being responsible for their extraordinary reactivity (see chapter 1.1.1).[Fe-S] clusters are one of the most versatile prosthetic groups in nature [63], with involvement in gene expression [64], Fe metabolism [65-67], oxygen sensing [68], oxidative stress response [69], redox coupled proton transfer reactions [70, 71] and disulfide reduction [72], among many others [63, 73, 74].This already extensive list of reactions has grown considerably larger thanks to the radical SAM superfamily. As will be discussed in this Chapter, the RSAM family utilizes the [4Fe-4S] cubane cluster present in its active site to initiate radical-based chemistry [23] in such diverse reactions as amine shifts [54], oxidative decarboxylation [75], sulfur donation reactions [58, 76-79], metallo-

cofactor biosynthesis [80-82], nucleic acid methylation [83]; cofactor biosynthesis [78, 84, 85] amongst numerous others [21, 22, 25, 35, 86].

### 1.1.1 [4Fe-4S] Cluster Interaction With SAM in Radical SAM Enzymes

In the RSAM superfamily, the [4Fe-4S] **1** cluster is involved in the initiation of radical catalysis by inducing the reductive cleavage of SAM **2**, leading to the formation of a highly reactive primary carbon radical intermediate, predominantly DOA$^{\bullet}$ **3** (Figure 1) [87-89]. Although the involvement of an [Fe-S] cluster in the reduction of a sulfonium ion to the corresponding alkyl radical and thioether was almost unprecedented when these systems were first studied, an important function of the RSAM enzyme was to modify the active site thermodynamics of SAM cleavage to yield a more energetically favourable process [90-94].

The role of the cluster in this reaction is now known to be two-fold: to anchor SAM via the interaction of the amino acid moiety with the unique iron (the only iron not cysteine-bound) [95, 96] and to catalyse the electron transfer step which leads to the cleavage of the molecule and the formation of the radical intermediate [97-99].

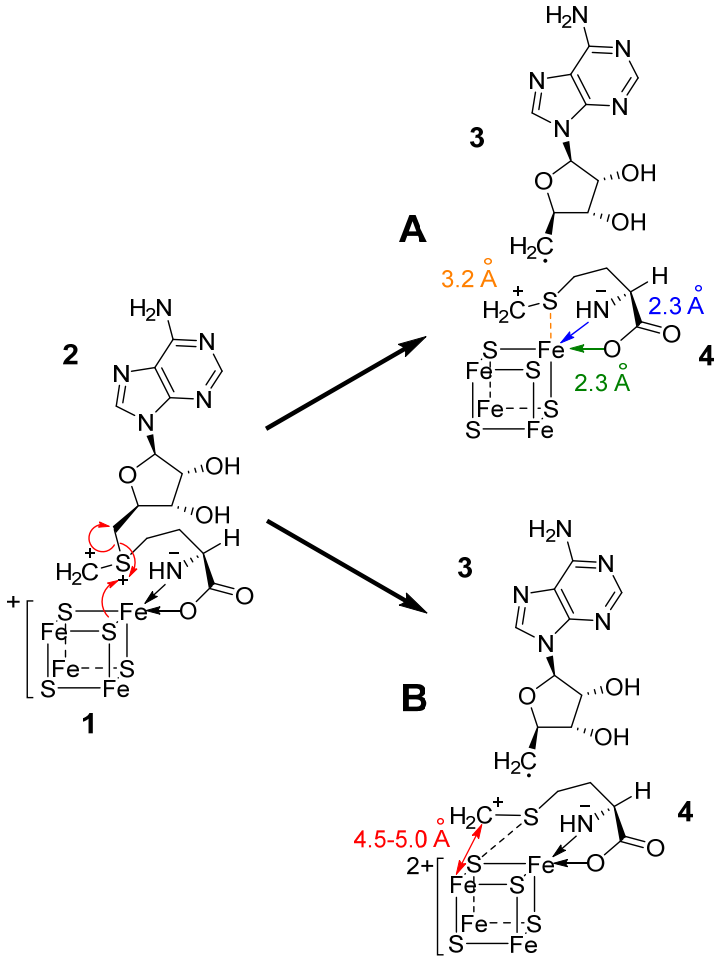



**Figure 1 [4Fe-4S]: SAM interaction and reductive cleavage.**

Binding of SAM **2** to the [4Fe-4S] cluster **1** present in RSAM enzymes and its cleavage to form DOA$^{\bullet}$ **3** and methionine **4** (bound to the cluster). **A:** Proposed binding model for SAM in reactions where this molecule acts as a cofactor. Model adapted from Nicolet *et al.* and Dey *et al.* [87, 100]; **B:** Proposed binding model for SAM in reactions where it acts as a co-substrate, adapted from Walsby *et al.* and Cosper *et al.* [89, 101] . Evidence from later studies [87] has led to A being widely accepted as the mechanistic model for SAM cleavage in RSAM enzymes.

### 1.1.1.1 Chemical Nature of the [4Fe-4S] Cluster

During the purification steps of RSAM enzymes a mixture of different cluster states is commonly observed, hinting at a mechanism of formation and ready interconversion that mimics the assembly of the cluster *in vivo* (section 1.1.3) [59, 102, 103]. For many RSAM enzymes, mutations of the cysteine triad have been shown to result in the inactivation of the enzyme [60, 83, 104-118], providing support for the importance of the cluster in the activity of the enzyme.

The necessity of a[4Fe-4S] cluster 1 in the +1 oxidation state (formally $[Fe^{III}Fe^{II}_3S_4]$), as shown in Figure 1, was initially deduced by the requirement of both an anaerobic environment as well as a strong reducing agent (commonly dithionite) for optimal activity [54]. A better indicator for the reductive activity of the $[4Fe\text{-}4S]^+$ cluster came from the stoichiometric correlation of the loss of the EPR signal corresponding to the (EPR active) $[4Fe\text{-}4S]^+$ with the formation of the radical product [119, 120]. Unfortunately, the reactivity of the $DOA^{\bullet}$ radical that arises from SAM cleavage makes it very unstable and short-lived and therefore it is not practically feasible to detect directly. This difficulty was circumvented using the SAM analogue, 3',4'-anhydroadenosyl-l-methionine **5** to form the stabilised allylic 3',4'-anhydro-5'-deoxyadenosyl radical analogue **6** (Figure 2). The analogue radical is spin delocalized and highly stabilised, making it possible to observe directly by EPR [121, 122], and the strong correlation between its increase of signal and the loss of the $[4Fe\text{-}4S]^+$ signal provided strong evidence for the radical-based mechanism shown in Figure 1.

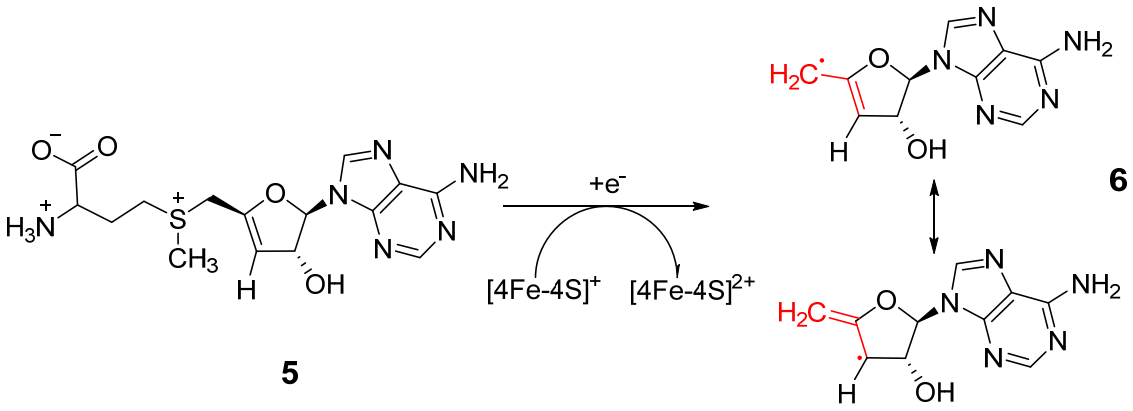



**Figure 2 Radical formation and stabilization from 3',4'-anhydroadenosyl-l-methionine**.

After proton abstraction, the resulting allylic anhydroadenosyl 6 radical is stabilized.

### 1.1.1.2 SAM:[4Fe-4S] Cluster Binding

The interactions of SAM with the [4Fe-4S] cluster were also actively investigated: for example, Mössbauer studies with labelling the unique iron with $Fe^{57}$ showed spectroscopic changes in the presence of SAM [95-97], strongly hinting at an interaction between SAM and the cluster. A similar conclusion was reached by the changes in the Raman [59, 97], EPR [59, 60, 97] and ENDOR [21, 50, 89] signals upon SAM binding. However, despite providing strong evidence for the binding

of SAM to the [4Fe-4S] cluster, these studies did not precisely define the spatial relationship of the cluster to SAM. Spectroscopic studies of lysine 2,3-aminomutase (LAM) [98, 99, 123] and pyruvate formate-lyase activating enzyme (PFL-AE) [89, 101, 124] led to the proposal of two distinct models for SAM binding, shown in Figure 1.

In both models, the carboxyl and the amino groups of SAM establish a bi-dentate interaction with the unique iron in the [4Fe-4S] cluster **1**, with the differences lying in the interaction of the cluster with the sulfonium ion of SAM. At this early stage of the RSAM studies the lack of any x-ray structural data, together with valid results arising from the different enzymes, led to an initial conclusion that the sulfonium of SAM **2** would interact with the unique iron or with a sulfur of the cluster, depending on whether SAM was acting as cofactor (LAM; Figure 1 A) or co-substrate (PFL-AE, Figure 1 B) respectively [89, 98].Subsequent studies, based on the available radical SAM enzyme structures with products, substrates or analogues bound at the active site and analysed using density functional theory quantum modelling [87] and S K-edge X-ray absorption spectroscopy [100] seem to indicate that the interaction between SAM's sulfonium and the unique iron (Figure 1 A) is shared across the RSAM superfamily.

#### 1.1.1.3 Reduction Potential in the Interaction of SAM with the [4Fe-4S] Cluster

The formation of the five-membered chelate ring shown in Figure 1A facilitates inner-sphere electron transfer [87, 99], with the electron transfer occurring directly from the cluster to the SAM sulfonium ion. The sulfur of SAM is held in close proximity (close to van der Waals radii) to the unique iron of the cluster and the S-C bond is homolytically cleaved, forming the 5'-deoxyadenosyl radical [21, 87, 99, 125, 126] and methionine.

The C-S bond of SAM is too strong to be homolytically cleaved simply by the binding of the molecule to the enzyme's active site [127]. Comparison of the energetics required for reductive cleavage of sulfonium ions in the enzyme and in free solution with model compounds is informative: the values for the reduction potential (2+/1+) of the [4Fe-4S] clusters present in the RSAM enzymes range from -400 mV to -500 mV (Figure 3, [4Fe-4S]:LAM:PLP complex) [92, 128], a value considerably smaller than the -1600 to -1800 mV calculated for the chemical reductive cleavage of a trialkylsulfonium compound [129, 130].

The binding of SAM to the cluster has two significant results that allow this 'electrode potential gap' to be bridged. Firstly, the binding of SAM increases the potential of the [4Fe-4S] cluster to around -430 mV [93], allowing most biological reducing systems to reduce the cluster from +2 to the +1 state [131, 132]. Additionally, the reduction potential of SAM is drastically increased when bound to the cluster; this has been calculated at -990 mV [93], a value significantly different from that of the free-solution generic trialkylsulfonium (Figure 3, used to estimate the free solution reduction potential of SAM, -1600 to -1800 mV).

The binding of the (non-SAM) substrate to the active site further lowers the cluster reduction potential by 170 mV [93], which, together with small changes in the conformation of the enzyme facilitates the electron transfer and subsequent reaction.

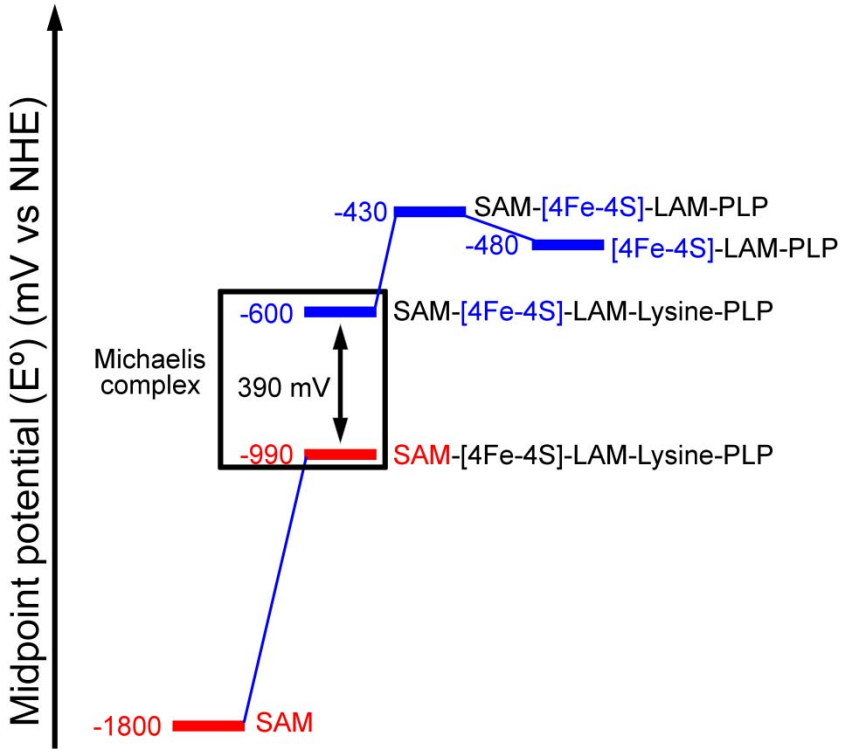



**Figure 3 Energetics of reductive SAM cleavage at the active site of lysine-2,3-aminomutase (LAM).**

SAM related midpoint redox potentials are shown in red and [4Fe-4S] cluster associated potentials in blue. Adapted from Wang *et al*. [93]

These results indicate that LAM exerts considerable control of the cleavage of SAM for the generation of the radical and this occurs most easily in the Michaelis complex (LAM:SAM:Lys:PLP complex. This is not unexpected as the generation of radicals within a cell needs to be highly regulated, to prevent radical mediated damage that *in vivo* can lead to cell death [133, 134]. Without the correct positioning of the substrate, SAM is too stable and its cleavage is generally a slow process [87, 135]. SAM is also a fairly scarce resource, requiring important primary metabolites methionine and ATP to be formed by SAM synthetase [136, 137], making it is one of the most important cofactors in the cell [138] and the unwanted cleavage of the compound an energetically expensive effort.

However*, in vitro* studies which use a significantly stronger reducing agent (usually dithionite, E° = -660mV) [139] than the ones present within the cell (flavodoxin, E° = -360-490 mV, depending on the organism) [140] can often result in the uncoupled formation of the 5'-deoxyadenosyl radical [88, 141], with either a buffer component, such as DTT, or the protein presumed to be acting as the source of hydrogen atoms to quench the 5'-deoxyadenosyl radical [99].

#### 1.1.1.4 Propagation of the Radical in Radical SAM Reactions

The homolytic cleavage of SAM **2** leads to the formation of a DOA$^{\bullet}$ species **3** (Figure 1). This radical is a strongly oxidising species capable of abstracting a hydrogen atom from a wide range of RSAM substrates, leading to the formation of a radical intermediate, the nature of which is dependent on the precise structure of the second (non-SAM) substrate [142].

In the majority of the RSAM reactions, the hydrogen is abstracted from the substrate molecule [33, 37-43, 46, 47, 56-58, 78, 79, 83, 85, 143]. In some instances, however, the abstraction occurs from an individual residue of the enzyme. This ordinarily occurs with the so-called activating enzymes, most commonly leading to the formation of a glycyl radical that is essential to the catalytic activity of the activated enzyme [44, 45, 88, 144, 145].

Interestingly a small subset of the RSAM enzymes is able to re-form the SAM molecule at the end of the reaction cycle after the hydrogen atom abstraction [27, 146], with the product radical able to re-abstract a hydrogen atom from the 5'-methyl of 5'-deoxyadenosine (DOA). Following the reduction, the resultant DOA$^{\bullet}$ **3** recombines with the methionine **4** and returns an electron to the [4Fe-4S] cluster, reducing it to the initial state of the reaction demonstrated in Figure 1.

Studies with ENDOR spectroscopy have provided an insightful view of this mechanism of abstraction [135, 147], which together with crystallographic data [22, 148] show the structure of the enzymatic active site, in particular the location for hydrogen abstraction, very near the S-C5' bond of SAM. When SAM cleavage occurs, van der Waals constraints ensure that the radical position only needs to move slightly. This active site topology enforces the proximity of the incipient 5'-deoxyadenosyl radical and the C-H bond that will be cleaved, thereby leading to highly selective hydrogen atom abstraction from the substrate and minimizing undesirable abstraction side reactions (for example, from active site residues [148].

### 1.1.2 Structure of Radical SAM Enzymes

As RSAM enzymes have a wide diversity of substrates, the structures and sequences amongst members of this family generally show a high degree of variability in the substrate binding regions. An important exception can be found in what is known as the "core fold" (Figure 5) of the RSAM family, where the secondary structural elements are highly conserved within members of this family [21, 22, 50, 75, 101, 149-151].

This core fold is composed of a partial triosephosphate isomerase (TIM) barrel structure. A TIM-barrel fold is typically characterized by an eight-stranded, all parallel β-sheet made up of repeating β/α motifs curved in such a manner that the eight strands form a barrel which is surrounded by the eight $\alpha$-helices. Most Radical SAM enzymes contain what is known as a partial TIM barrel as it contains only six of the eight β/α units [21].

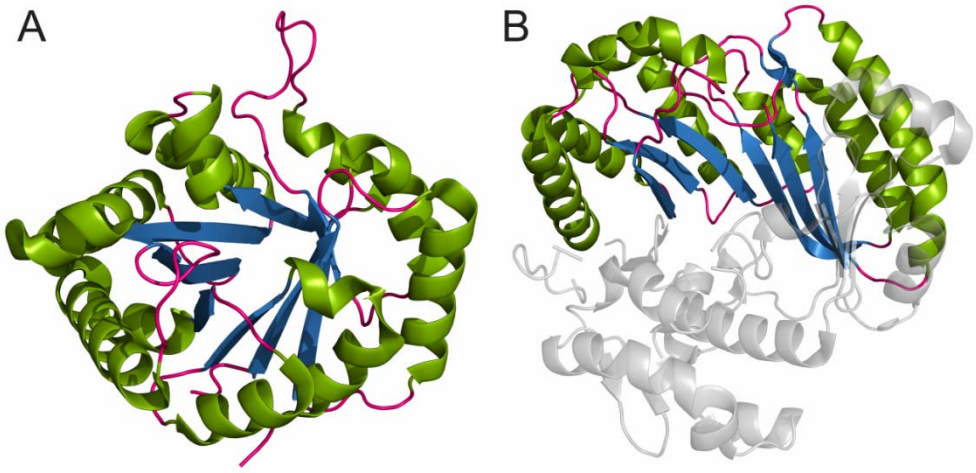



**Figure 4 TIM barrel structural comparison.**

Secondary structure that composes the TIM barrel structure is colour coded, with alpha helices in green, beta strands in blue and random loops in pink. **A:** Complete TIM barrel from triose phosphate isomerase (PDBID: 1TIM). **B:** Partial TIM barrel of the RSAM coproporphyrinogen III oxidase (HemN) (PDBID: 1OLT). N-terminal and C-terminal extensions are shown in grey.

The two missing β/α units lead to the formation of a lateral opening where the active site of the enzyme is often located (Figure 4 B). The degree of openness of the barrel is thought to correlate with the substrate size, as larger barrel openings allow larger substrates to bind in the active site, while a smaller barrel opening allows smaller substrates to bind while shielding the active site from surrounding solvent [152]. The [4Fe-4S] cluster is located within the active site, together with the binding sites for SAM and the substrate [21, 75, 92, 129, 151].

The RSAM family possesses a characteristic cysteine triad containing motif, CxxxCxϕC, (Figure 5, Figure 6) responsible for the coordination the [4Fe-4S] cluster [20].The three cysteines residues are bound to three iron atoms present in a [4Fe-4S] cluster, leaving a unique iron unbound. It is with this unique iron that SAM interacts (Figure 1) [21, 50, 61, 75, 92, 127-129, 131, 135]. The cysteine motif is located in a loop found at the top (C-terminal end) of the first β- strand of the core structure. This cluster bearing loop is responsible for the positioning the [4Fe-4S] cluster within the active site, as well as forming hydrophobic interactions to the adenine moiety in SAM through the aromatic ϕ residue of the cysteine motif, typically a tyrosine of phenylalanine [50, 61, 149, 150].

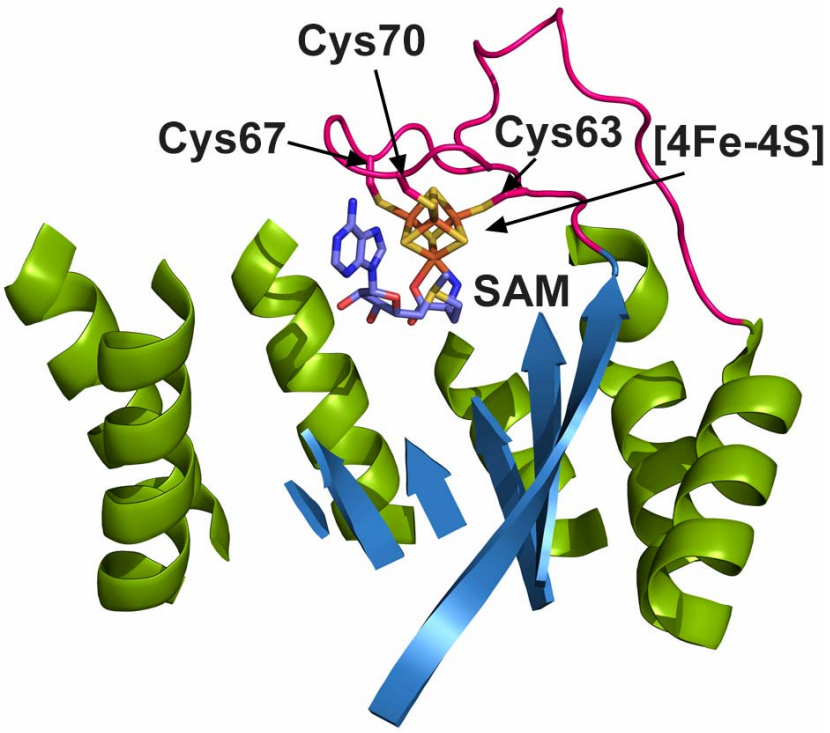



**Figure 5 Conserved core-fold of members of the RSAM superfamily of enzymes.**

Core of the RSAM enzyme HydE, with loops omitted for clarity (PDBID: 4JXC). The CxxxCxxC motif is found at the top (pink), with the three cysteines shown as stick models and the cubane [4Fe-4S] cluster shown in yellow/orange. SAM, in lilac, is shown binding to the [4Fe-4S] cluster. The partial TIM-barrel is revealed, with $\alpha$-helices shown in green and $\beta$-sheets shown in blue.

The combined interactions between the protein, cofactors and substrate result in a high proportion of conserved structure for the active site, consistent with the similarities in the general mechanism of the RSAM superfamily. Outside of this core fold, little to no similarity seems to exist between the RSAM enzymes of the different subfamilies; indeed, significant sequence and structural variation has evolved to permit the application of the common radical generating module to diverse substrates and metabolic requirements [61, 149, 150].

#### 1.1.2.1 SAM Binding Site

The SAM binding residues are commonly found dispersed throughout the protein sequence, comprising mostly hydrophilic residues from each of the core β-strands. The web of connections formed between SAM and the enzyme forces the correct positioning and conformation of the co-substrate for radical formation, through electron transfer, and hydrogen abstraction, preventing any unwanted abstraction, a side reaction that could lead to enzyme inactivation [22, 91].

Although the residues that interact with SAM are not always tightly conserved within the superfamily, a high level of conformational similarity within the different RSAM subfamilies commonly occurs [22, 92, 128, 135].

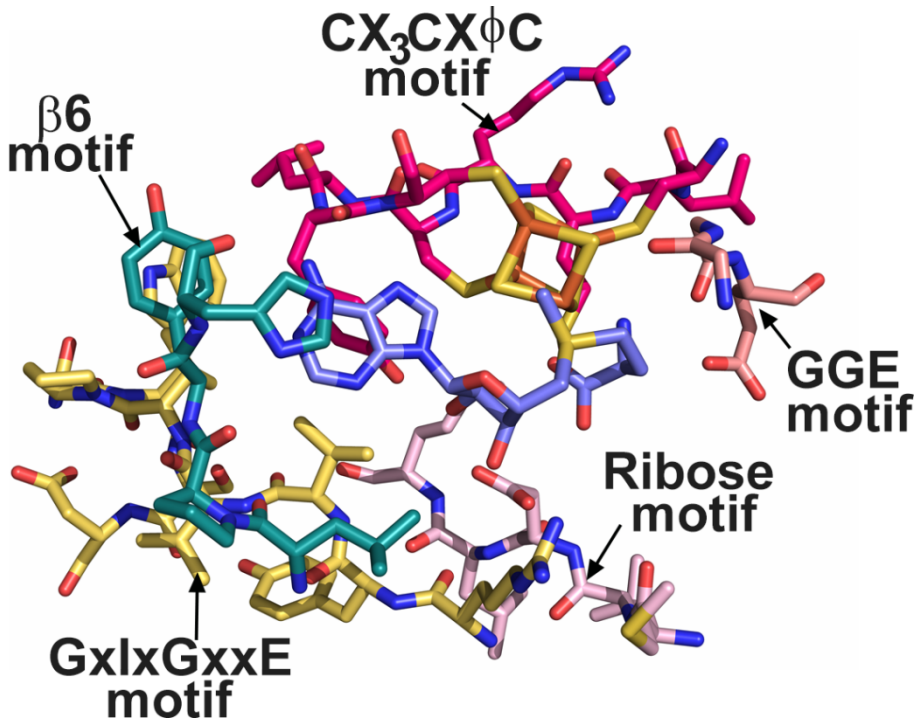



**Figure 6 Conserved SAM binding motifs of the Radical SAM superfamily.**

Colour key: magenta, $CX_3CX\phi C$-containing loop; salmon, "GGE" motif; light pink, ribose motif; golden, "GxIxGxxE" motif; teal, β6 structural motif; green, SAM. Structure of PFL-AE (PDB ID: 3CB8). Adapted from Dowling *et al.* [149].

#### 1.1.2.1.1 SAM Methionyl Moiety

As described previously in section 1.1.1.2, SAM is bound directly to a unique iron atom of the [4Fe-4S] cluster through its amino and methionyl carboxylate groups. The amino group is locked in the correct positioning for electron transfer by the interactions with what is known as the glycine rich motif (GGE) (Figure 6, salmon) [22, 50, 149]. In contrast with the conserved interactions of the amino group with the protein backbone, the interactions with the carboxylate are not conserved within the superfamily, allowing the conformation of SAM's methionine to differ between the enzymes of this superfamily [22, 149].

#### 1.1.2.1.2 SAM Adenine Moiety

The amino acid interactions with the adenine moiety show a high degree of structural conservation across the superfamily. Hydrophobic interactions occur in three specific regions: the conserved aromatic residue of the $CX_3CX\phi C$ motif (Figure 6, magenta); one/two residues located at the end of the β6 structural motif (Figure 6, pale cyan) and the GxIxGxxE motif [22, 50, 149, 150] (Figure 6, golden). The GxIxGxxE motif, located on the β5 sheet and following loop, stabilizes the structure of the SAM adenine binding site (Figure 6, pale cyan) [22, 149].

#### 1.1.2.1.3 SAM Ribose Moiety

The ribose 2' and 3' hydroxyl groups may form hydrogen bonds with charged or polar side chains of the β4 sheet, a group of residues termed the ribose motif (Figure 6, light pink), and a region which is less well conserved than the others within this superfamily. One of the conserved charged or polar residues is thought to have an important role during catalysis through the control of the orientation of either the DOA moiety or the substrate [153].

#### 1.1.2.2 Substrate Binding Site

Owing to the widely diverse range of reactions catalysed by the RSAM superfamily, the binding of the substrate is highly tailored to each specific reaction. As a consequence of this diversity, it is difficult to identify areas of sequence similarity that interact with the substrate across the superfamily as a whole [22, 61, 149].

Nonetheless, it has been shown that most substrates bind in the lateral opening of the partial TIM barrel with the protein-substrate interactions occurring predominantly in the C-terminus of the enzyme. It does seem that substrate orientation, with regard to the abstraction point is also fairly consistent within the RSAM superfamily, as the general mechanism of action is conserved across all members of the family [22, 149].

#### 1.1.2.3 Reductant Binding Site

As previously mentioned in section 1.1.1.3, the formation of the intermediate radical by members of the RSAM superfamily requires the reduction of the [4Fe-4S] cluster to the +1 redox state. *In vivo* this appears to occur by the action of a biological reductase, which binds in a region with the lowest conservation of all the core fold [22, 149, 150]. The reductant binding site is located on the surface near the cluster-binding loop and opposite the opening of the partial TIM barrel, meaning the substrate can bind at the same time as the reductant. As the reduction step is a vital component of the activity, the high degree of variability in the reductase binding region most likely indicates a diversity in the proteins responsible for electron donation, such as (and not limited to) flavodoxin [131, 154], ferredoxin [32] or adrenodoxin [155].

#### 1.1.2.4 Exceptions in Cluster Binding

Despite the cysteine-coordinating motif being highly conserved within the superfamily, and a powerful tool for identification of enzymes capable of coordinating the [4Fe-4S] clusters, two enzymes have been discovered that do not possess this motif, but still possess a mechanistic resemblance to the RSAM family of the enzymes. In the first case, the enzyme QueE possesses an eleven amino acid insertion between the first two cysteines ($CX_{14}CX\Phi C$) [41], while in ThiC the cluster-binding region is not part of the SAM radical core, but is instead found in a separate

domain [116]. These uncharacteristic enzymes demonstrate the diversity available to the wider RSAM family and indicate the potential for many more members of the RSAM superfamily that might exist away from those identified from sequence alignments.

### 1.1.3 The [4Fe-4S] Cluster

Iron-sulfur clusters in biology are traditionally seen as stable redox cofactors, commonly used as electron transfer units [73], with a modular nature that allows for several different, and sometimes interconvertible, structures [63]. In the context of the RSAM superfamily early Mössbauer studies identified a mixture of [2Fe-2S], [3Fe-4S] and [4Fe-4S] clusters [103], with the presence of a reducing agent, such as dithionite, resulting in predominantly [4Fe-4S] clusters [102, 156].

Conversely, the exposure of RSAM enzymes to $O_2$, NO, oxidizing agents, pH changes and mutations to the cluster binding residues result in a destabilization of the [4Fe-4S] cluster even in reducing conditions. The degradation of the [4Fe-4S] cluster can lead to the reversible cleavage into [2Fe-2S] units or to the formation of a [3Fe-4S] intermediate during the breakdown [157-159].

The interconversion of different iron-sulfur clusters is also seen during [4Fe-4S] cluster biosynthesis [160, 161] (Figure 7). Attesting to the importance of iron-sulfur cluster regulation, different networks have been identified for cluster assembly: the NIF system, which assembles the [Fe-S] clusters of nitrogenases in azotrophs [162, 163]; the ISC assembly system in mitochondria and bacteria [164-166]; the SUF system in bacteria, Archaea and plant chloroplasts [165, 167-169] and the CIA system in eukaryotic cytosol [170].

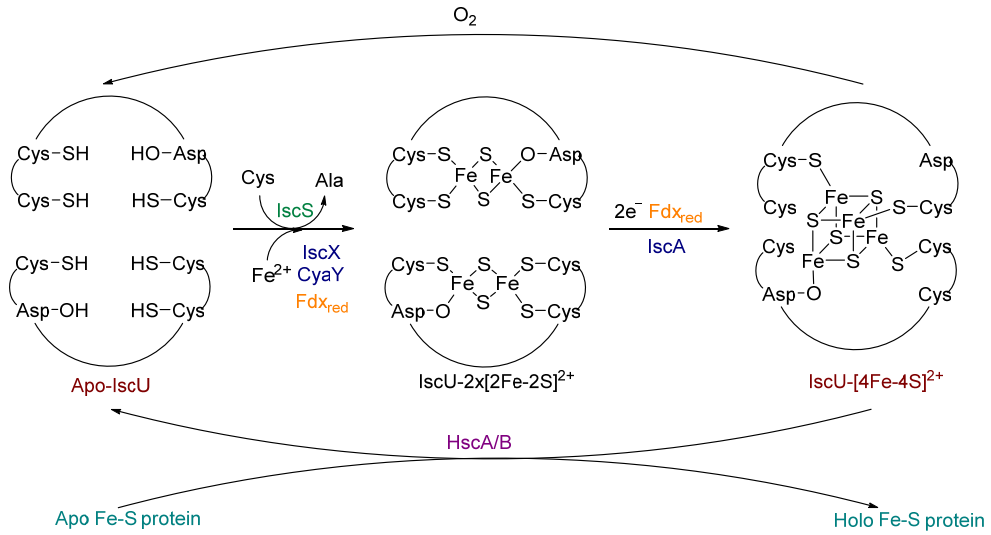



**Figure 7 Biosynthetic pathway for the introduction of the [4Fe-4S] cluster into a generic apo-enzyme by the ISC machinery.**

Cysteine desulfurase IscS and iron transferases IscA/IscX/CyaY provide the sulfide and iron atoms, respectively to the scaffold enzyme IscU. Ferredoxin catalyses the reductive coupling of to [2Fe-2S] clusters into one [4Fe-4S] cluster, which can then be transferred to the apo-enzyme by HsCA/B. Adapted from Blanc *et al.* [164].

### 1.1.3.1 Auxiliary Clusters in RSAM Enzymes

A significant subset of RSAM proteins are able to coordinate additional [Fe-S] cluster(s) in their structure [171, 172], with most of the roles for the auxiliary cluster(s) being currently poorly defined [171]. In the past few years however, two different subgroups of enzymes which contain at least one auxiliary cluster have been identified: the SPASM/Twitch group [173] and the sulfur insertion subgroup [172].

The SPASM subfamily of enzymes, predicted to have over 1,000 members, is postulated to be involved in post-translational modification of peptides [174]. The enzymes within this group share a conserved seven cysteine motif $CX_{9\text{-}15}GX_4C–X_n–CX_2CX_5CX_3C–X_n–C$ [175-178], which bind two auxiliary clusters. So far, only one crystallographic structure of a SPASM enzyme is available [177], revealing that both auxiliary clusters have their four irons fully coordinated by four cysteine residues from within and outside the seven cysteine motif (Figure 8). Interestingly, the topology around the first auxiliary cluster, known as the Twitch domain [178], can also be found in enzymes

which do not belong to the SPASM subfamily. In the Twitch containing enzymes the first auxiliary cluster can be observed, but no coordination for a second cluster is possible [178, 179].

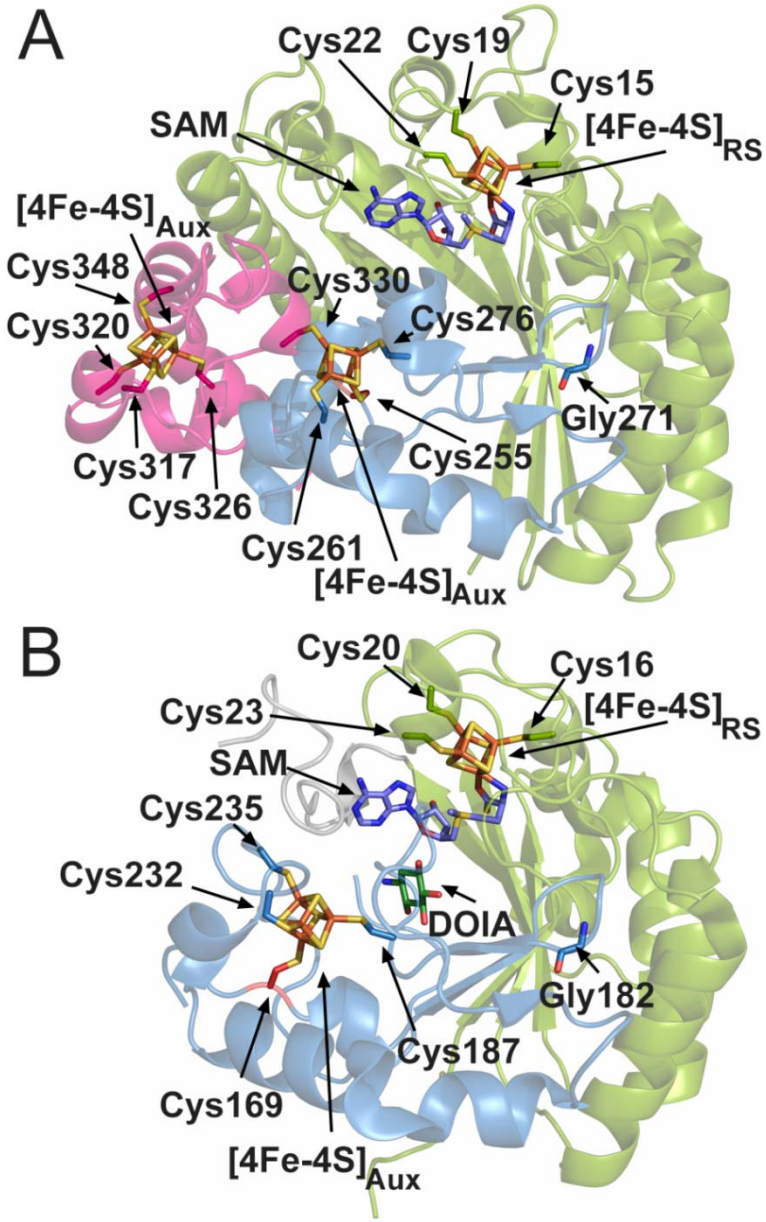



**Figure 8 SPASM/Twitch RSAM structure comparison.**

The RSAM fold is shown in green, the conserved Twitch fold is shown in blue and the additional elements that form the SPASM fold are shown in pink. Iron-sulfur clusters (orange and yellow) and SAM (lilac) are shown as stick models.

**A:** Structure of AnSMEcpe (PDBID: 4K39), a member of the SPASM subfamily of enzymes. The seven cysteines and glycine that form the motif $C_{261}$-$X_9$-$G_{271}$- $X_4$-$C_{276}$-$X_{40}$-$C_{317}$- $X_2$-$C_{320}$- $X_5$-$C_{333}$- $X_3$-$C_{330}$ as well as the non-conserved cysteine (red), coordinating two auxiliary [4Fe-4S] clusters, are represented by sticks. **B:** Structure of BtrN (PDBID: 4M7T), with the small Twitch fold, that coordinates one auxiliary cluster, shown in blue. The cluster binding cysteines of the conserved Twitch motif $C_{169}$-$X_{12}$-$G_{182}$-$X_4$-$C_{187}$-$X_{44}$-$C_{232}$-$X_2$-$C_{235}$ and the non-conserved cysteine $C_{255}$ (red) are represented by stick models.

Members of the sulfur insertion sub family are the most well studied cases of auxiliary cluster whose function is known. The enzymes that belong to this group seem to catalyse the formation of C-S bonds at non-activated carbon atoms, with the auxiliary cluster proposed to serve as either the source of the sulfur for thiolation in a self-sacrificing reaction [76, 180-182], or to provide the binding location for a polysulfide moiety which provides the sulfur atom for the reaction and leads ultimately to methylthiolation reactions [37, 79, 183].

Alongside the enzymes from these two groups, several other RSAM enzymes with an auxiliary cluster have been discovered [171], such as the [Fe-Fe] hydrogenase maturation enzymes HydG/HydE [184], benzylsuccinate synthase [185], tioether bond formation SkfB [186] and wybutosine biosynthetic enzyme TYW1 [117]. Even though this second cluster is also commonly coordinated by three cysteines, the cluster's position in the sequence is much less conserved. The overall lack of a conservation of the number of residues between the cysteines $CX_aCX_bC$ that coordinate the second cluster makes it harder to accurately predict the presence of a secondary cluster by simple alignment studies (Figure 9).

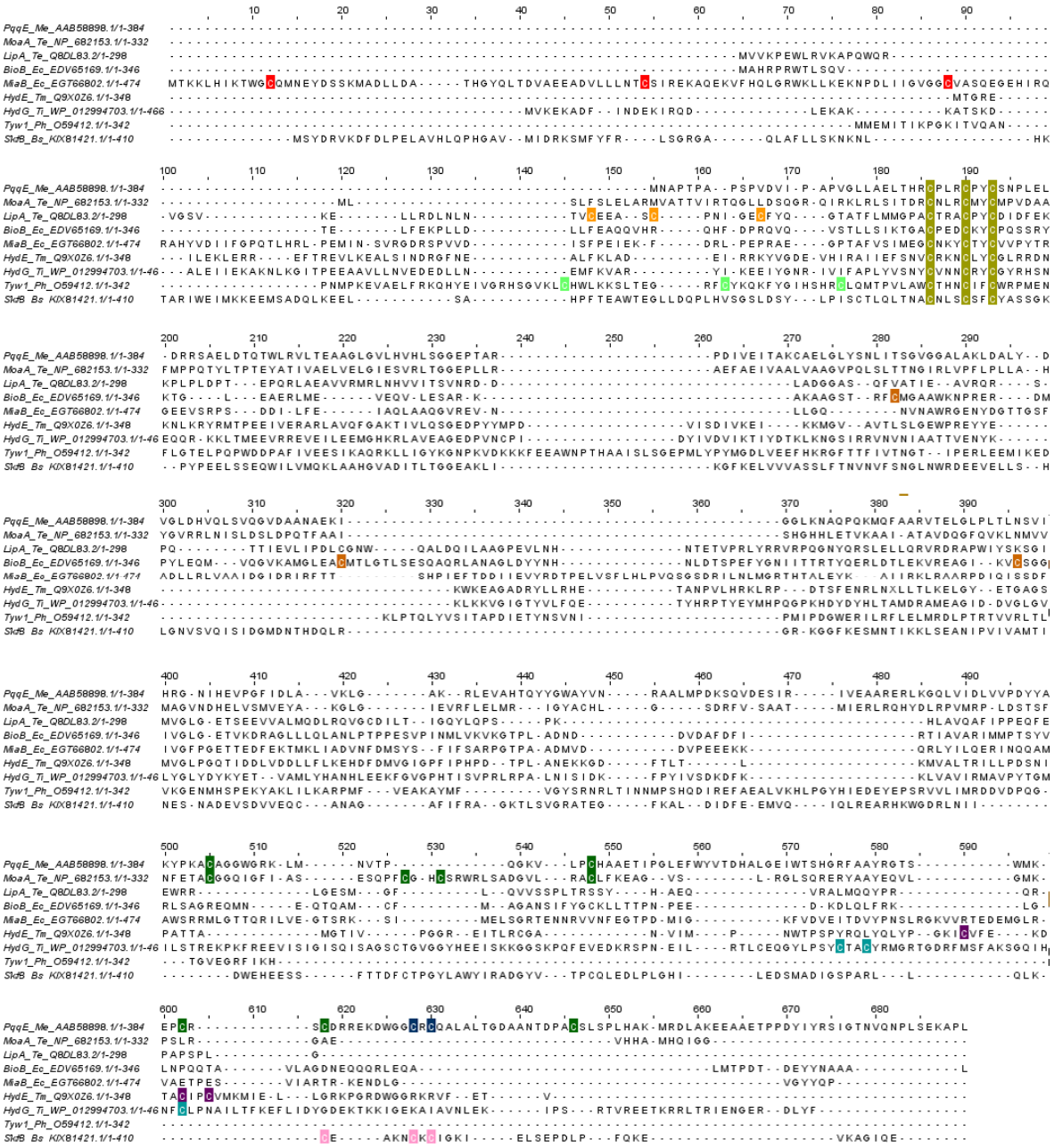



**Figure 9 Sequence alignment of different RSAM that possess auxiliary iron-sulfur clusters.**

In yellow it is shown the highly conserved CxxxCxxC motif characteristic of this superfamily. The cysteines that coordinate an auxiliary cluster are shown in colour, respective to their origin. Twitch and SPASM are shown in green and blue. In orange the cysteines that coordinate the sacrificial cluster in sulfur donating enzymes LipA (light orange) and BioB (dark orange). Cysteines of the methylthioadenosine subgroup are shown in red. Cysteines of the highly related [Fe-Fe] hydrogenase maturases HydG and HydE are shown in purple and teal. Tyw1 cysteines are shown in light green and SkfB in pink.

## 1.2 Hydrogenases and Maturation Pathways

The need for viable energy sources to replace the use of fossil fuels is one of the greatest challenges of modern times [187]. Renewable energy sources such as solar and wind power represent a potential replacement, however they are limited by the challenges associated with storing they energy they generate [188, 189]. In this regard, the use of hydrogen in fuel cells possesses several advantages, as it can theoretically be produced from renewable energy sources such as sunlight [190] and can be used as energy storage, both due to the high energy density of the H-H bond [191] as well as the ability to be easily transported in liquid form.

One of the greatest disadvantages of using hydrogen as a fuel source lies with the high costs associated with its production: molecular hydrogen is generated from protons and electrons, a process that requires expensive catalysts [192]. It has, however been demonstrated that species of bacteria and green algae can perform the reversible reduction of protons, thanks to the presence of hydrogenase metallo-enzymes [193, 194]. So far the ability to introduce a large scale production of hydrogen through use of these biological systems has been limited, in no small part due to the hydrogenase enzymes' sensitivity to inactivation in the presence of oxygen [195]. A deeper understanding of the steps that lead to the formation of an active hydrogenase may hold the key to circumventing this problem and to bio-engineer a more robust system of bio-hydrogen formation [196-198].

### 1.2.1 The Hydrogenase Family of Enzymes

The early atmosphere of Earth was highly reducing, with a significantly higher level of atmospheric hydrogen present [199-201]. This may have led to the use of hydrogen as a major fuel source for the earliest life forms, which could have utilised hydrogenase enzymes as part of their metabolism [200]. This class of enzymes were first discovered over 80 years ago [202] and are primarily present in prokaryotes [203, 204]; though their presence in eukaryotes has also been reported, however, to a much lesser extent [205, 206].

$$2\,H^+ + 2\,e^- \rightleftharpoons H_2$$

**Equation 1 Interconversion of molecular hydrogen and protons catalysed by enzymes of the hydrogenase superfamily.**

Hydrogenases are a very important class of enzymes for various metabolic processes, as hydrogen is often either the dominant product or a reactant in redox metabolism [202, 206, 207].The metalloenzymes hydrogenases use transition metal cofactors to catalyse the interconversion of protons and electrons to dihydrogen [208] (Equation 1). In general terms the equilibrium shift towards reduction of protons or oxidation of dihydrogen is dependent on the oxygen content of the environment.

In a completely anaerobic environment, organisms that possess hydrogenases are known to use protons as electron acceptors, producing $H_2$ upon the fermentation of organic material [209, 210]. Conversely, in aerobic environments cyanobacteria have used the oxidation of $H_2$ to yield $H_2O$ in order to generate a proton gradient for ATP synthase [211, 212], or as a simple mechanism to remove excess reducing equivalents from the cell [213]. Various exceptions, however, are known in nature. In oxygen rich environments, dinitrogen-fixing bacteria reduce protons in their nitrogen metabolism [214-216], while in anoxic environments, methanogenic Archaea as well as acetogenic and sulfur reducing bacteria are known to oxidize $H_2$ as their sole energy source [217].

### 1.2.2 Hydrogenase Subclasses

Hydrogenases are a heterogeneous class of enzymes and an interesting example of convergent evolution, where unrelated ancestral enzymes have evolved chemically related mechanisms [203, 207, 218]. The different subclasses have been widely studied and reviewed [208, 219-223], differing in sequence, structure, specific activities, oxygen tolerance, nature of electron donors and acceptors, reactants, and, most importantly, cofactor content [203, 224]. The nature of the metallic cofactor has been widely used to classify hydrogenases:

1. The [Fe]-only hydrogenase, also known as methylene-tetrahydromethanopterin dehydrogenase (Hmd) [225], is the only subclass not to contain any Fe-S clusters [226, 227], having instead a mononuclear iron site hexacoordinated to a guanylyl pyridinol cofactor [228-230]. The unique cofactor structure present in this enzyme catalyses the reversible hydride transfer from dihydrogen to methenyl-tetrahydromethanopterin as an intermediate step in the reduction of carbon dioxide to methane by methanogens [224, 231, 232].

2. [NiFe]-hydrogenases have a hetero-dinuclear cofactor with a nickel and an iron atom together with a [4Fe-4S] cluster [233, 234]). This subclass of hydrogenases seem to favour the hydrogen oxidation direction of reaction [233, 235-237] , whist also being less sensitive to oxygen inhibition [233, 238]. Within the [NiFe]-hydrogenase subclass, two sub-groups show a significantly higher tolerance for oxygen. One such subgroup includes the membrane-bound hydrogenases, where the most significant change occurs in the proximal FeS cluster responsible for the electron transfer, due to both the presence of six potential binding histidines and its unusual [3Fe-4S] coordination [239-241]. The other oxygen tolerant hydrogenase is composed of seleno-hydrogenases, where one of the cysteine thiolates bound to the cofactor's nickel is replaced by a selenocysteine [241, 242].

3. [FeFe]-hydrogenase (HydA) has a six iron cofactor, termed the H-cluster, comprising of a [4Fe-4S] cluster bound by a cysteinyl thiolate to a [2Fe-2S] cluster [243-246]. Although this enzyme is the most susceptible to irreversible oxygen inactivation [247, 248], HydA has a very high level of hydrogen forming activity [208, 249].

The differences in activities between [NiFe]- and [FeFe]-hydrogenases can partly be attributed to their reducing partners [250]. [FeFe]-hydrogenases are typically coupled with ferredoxin, which has a reduction potential of -453 mV [251, 252], a value close to the midpoint potential of the $H_2/H^+$ pair (-420 mV) at neutral pH [253]. Most [NiFe]-hydrogenases, on the other hand, are coupled with a weaker reducing partner, such as NAD(P)H (-320 mV), which disfavours hydrogen production over the [FeFe]-hydrogenase/ferredoxin partners [254]. An exception has been shown in the [NiFe]-hydrogenase from *Synechocystis sp.*, an enzyme which can be reduced by ferredoxin. The lower reduction potential of the reducing partner means this [NiFe]-hydrogenase is able to efficiently form hydrogen [255], being known as a bidirectional hydrogenase.

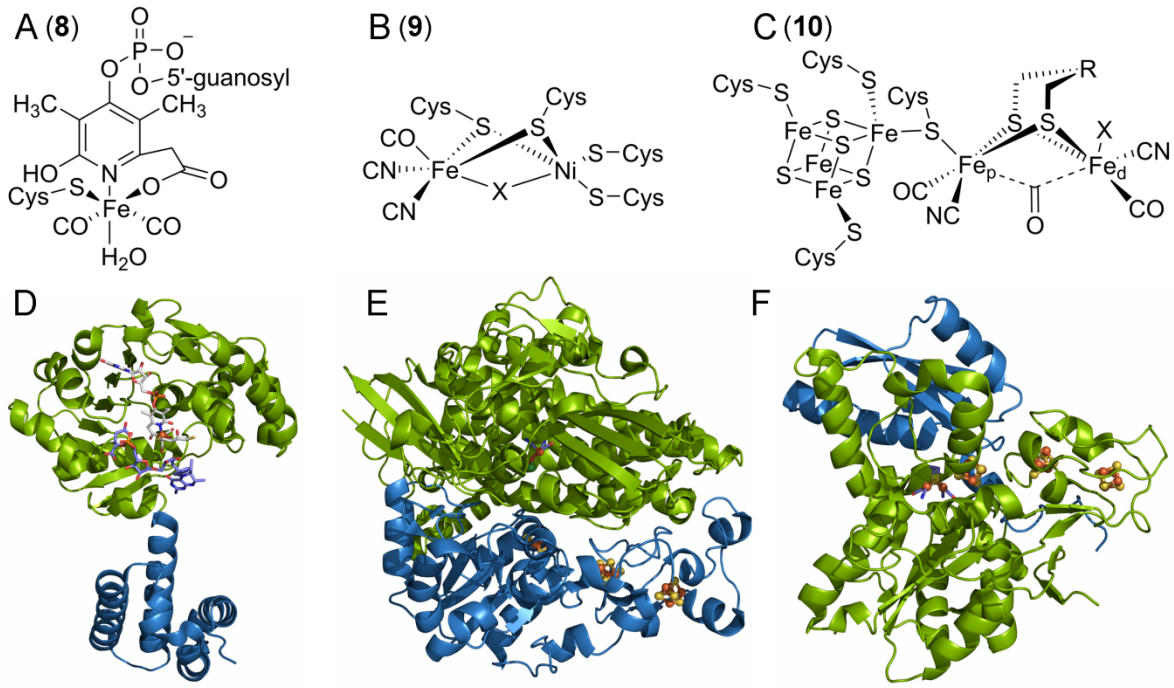



**Figure 10 Organometallic cofactor and protein structure of the three different main classes of hydrogenases.**

**A:** Fe-guanylyl-pyridinol cofactor **8** [256, 257]. **B:** NiFe cofactor **9** [258]. **C:** H-cluster **10**, where the dithiolate bridgehead (R) has been shown to be NH [259]. **D:** Structure of Fe-only hydrogenase from *Methylobacterium extorquens* (PDBID: 3H65). The Rossmann-fold N-terminus is shown in green, binding the cofactor in white and the substrate in blue. The C-terminus, in blue, serves as a dimerization unit. **E:** Structure of NiFe-hydrogenase from *Desulfovibrio gigas* (PDBID: 1YQ9). The small subunit, containing the Fe-S clusters responsible for electron transfer is shown in blue. The large subunit, where the Ni-Fe cofactor is observed is shown in green **F:** Structure of FeFe-hydrogenase from *Desulfovibrio desulfuricans* (PDBID: 1HFE). The secondary structure is shown in green and blue (C-terminus).

Despite their differences, in all three types of hydrogenases there is an redox-inactive, low spin Fe(II) which is ligated by at least one sulfur atom and three unusual π-accepting ligands such as CO, CN and/or a pyridinol that are arranged similarly in a square-pyramidal or octahedral geometry [247, 260].

The cofactor **8** in [Fe]-only hydrogenase can be considered the most unusual structure, as it features an iron-acyl carbon bond [256, 257], while the metallic cofactors of both [NiFe]- **9** and [FeFe]-hydrogenases **10** share some similarities. In both [NiFe]- and [FeFe]- cofactors, the two metals of the cofactor are linked through thiolate bridges (cysteines in [NiFe]-hydrogenases and

azapropanedithiolate in [FeFe]-hydrogenases) and possess iron-bound CO and CN ligands [261, 262].

Furthermore, both [NiFe]- and [FeFe]-hydrogenases appear to have a fast electron transfer system. This system comprises of [4Fe-4S] clusters within the protein, forming an electron transfer chain, together with networks of hydrophobic channels that form small pathways from the surface of the protein into the active site, allowing for the passage of gas molecules [242, 246, 247, 258, 263, 264]. In the [FeFe]-hydrogenase's case the channels can also allow for the reversible inhibitor CO [265] or the irreversible inhibitor $O_2$ [195] to permeate onto the active site, which is why organisms that possess [FeFe]-hydrogenases are usually obligate anaerobes.

#### 1.2.2.1 [FeFe]-Hydrogenase

The cofactor responsible for the activity of the [FeFe]-hydrogenase (HydA), known as the H-cluster **10** (Figure 10 C), consists of a [4Fe-4S] cluster electron transfer unit, connected via a single cysteine thiol to the binuclear Fe-Fe site, known as the 2Fe subcluster. The 2Fe subcluster possesses a proximal ($Fe_p$) and a distal ($Fe_d$) iron atoms (relative to the cubane cluster), with ligands of both irons bound in a square pyramidal structure, inverted in respect to each other [266]. Both iron atoms are coordinated by CO, CN and a dithiolate bridge [259, 267] which is proposed to keep the metal centres in a low oxidation and spin states, making the iron more similar to heavier transition metals which would perhaps be unavailable in nature [261, 268].

A dithiolate bridge connects both irons, being positioned above the vacant coordination site (X in Figure 10 C) allowing it to act as a shuttle for protons. The identity of the bridgehead atom was subject of intense discussion in the scientific community, being proposed as either a carbon, oxygen or nitrogen atoms [246, 259, 269]. The initial evidence for an amine cofactor was supported by hyperfine sub-level correlation (HYSCORE) and electron nuclear double resonance (ENDOR) spectroscopies as well as theoretical modelling and quantum refinement of existing X-ray data [259, 270-273]. Further evidence for the amine bridgehead was found by the incorporation of synthetic Fe2 subcluster mimics into HydA and analysing their activity. These studies showed that full activation of the apo-hydrogenase could only be obtained when an azadithiolate bridge was in place [274, 275].

Despite there only being one covalent bond between the cofactor and the protein (the coordination of the cysteine to $Fe_p$), a network of other interactions has been shown to be critical for efficient hydrogen formation activity of the hydrogenase [276]. Mutagenesis studies in HydA from *Clostridium pasteurianum* showed the CO that bridges $Fe_p$ with $Fe_d$ is involved in an

electrostatic interaction with of methionine ($Met_{353}$); the CN bound to $Fe_d$ forms a strong hydrogen bond to a neighbouring lysine ($Lys_{358}$), whilst a different methionine ($Met_{497}$) together with a cysteine ($Cys_{299}$) participate in an electrostatic interaction with the amine from the azadithiolate bridge [269, 276].

#### 1.2.2.1.1 The [FeFe]-Hydrogenase Mechanism of Action

The structures of HydA from *Clostridium pasteurianum* [249] and *Desulfovibrio desulfuricans* [246] reveal two different intermediate states of the enzyme during the reaction, known as $H_{ox}$ **12** and $H_{red}$ **13**, which differ in the coordinating solvent molecule as well as the bridging of the carbon monoxide ligand. These two states are the foundation of the proposed model for the mechanism of HydA (Figure 11), while many of the intermediate stages proposed have only been partially characterized [277]. Recently, one additional state, considered a "super" reduced state ($H_{sred}$ **14**), has been described [278], leading to an update of the mechanism of action.

In the proposed model of catalysis [279] one electron from the biological reducing donor ferredoxin [118, 251] travels through the [4Fe-4S] cluster and reductively activates the [2Fe-2S] cluster, with the loss of a bound water molecule. This leaves the cluster in a [Fe(II)Fe(I)] state ($H_{ox}$) **12**, which is stabilized by the CO and $^{-}$CN ligands surrounding it [244, 280]. A proton is thought to enter through a very small channel that leads from the surface of the protein to the 2Fe-subcluster [244, 245, 247], together with an additional electron that leads to the $H_{red}$ **13** state of the cluster.

A further reducing step leads to the newly identified super reduced state, $H_{sred}$ **14**, where the proton is believed to be associated with a neighbouring protein residue. The binding of another proton to the amino group of the azadithiolate bridge, together with the re-abstraction of the hydrogen from the neighbouring amino acid, reduced to its hydride form, allows for the re-oxidation of the subcluster to a loaded form of $H_{ox}$ **15**, which catalyses the formation of molecular hydrogen.

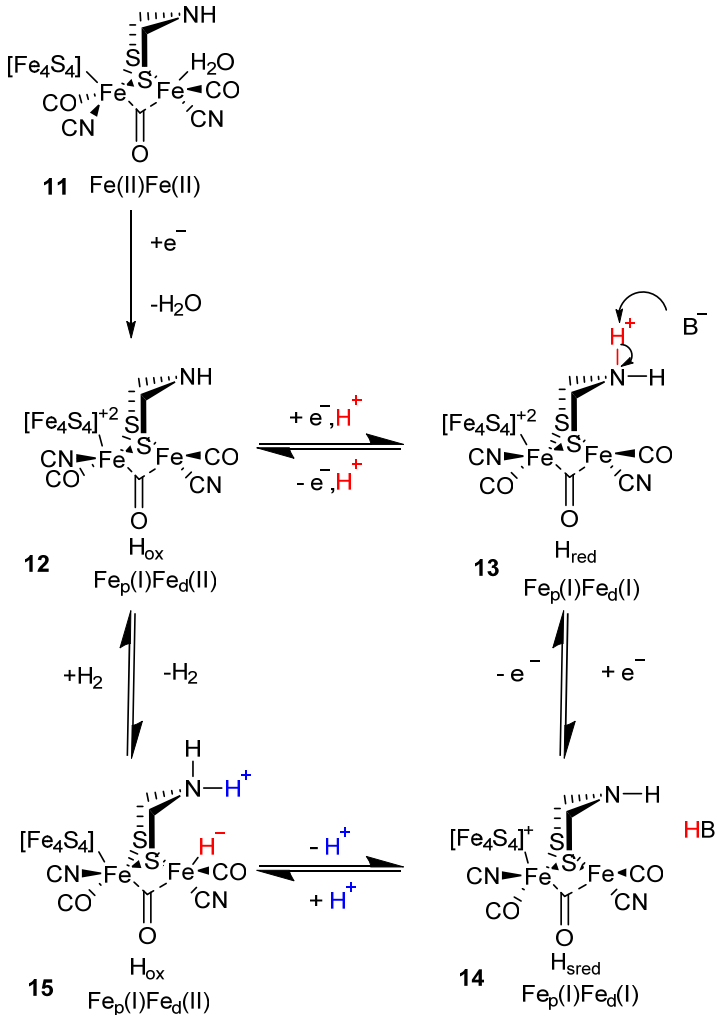



**Figure 11 Proposed mechanism for $H_2$ production in the active site of [FeFe]-hydrogenases.**

The oxidation states of the irons ($Fe_p$ and $Fe_d$) are shown below the scheme. The mechanism of reaction is proposed to begin with the reductive binding of a proton to the azadithiolate bridge **13**, reducing the H-cluster from its $H_{ox}$ state **12** to $H_{red}$. An additional reduction, with the proton associating with a nearby amino acid residue leads to the super reduced state $H_{sred}$ **14**. A second external proton will bind to the $H_{sred}$ together with the hydrogen anion, oxidizing the H cluster and forming hydrogen. Adapted from Adamska *et al.* [278].

### 1.2.2.2 H-Cluster Maturation

During a random mutagenesis trial to test which genes were necessary for $H_2$ production in *Chlamydomonas reinhardtii*, Posewitz and colleagues identified a particular mutant (*hydEF-1*) which could express the hydrogenase protein but did not show any hydrogenase specific activity [81]. The *hydEF* gene, which in prokaryotes is divided into two independent domains, *hydE* and *hydF*, was found adjacent to the *hydG* gene and all organisms that contain [FeFe]-hydrogenases possess homologues of *hydE, hydF* and *hydG* [81].

Heterologous expression of HydE, HydF and HydG triad together with HydA in *E. coli* showed a complete activation of the hydrogenase activity [118, 281]. Additionally, an inactive HydA, expressed in *ΔhydEFG* was successfully activated *in vitro* if incubated with cell extracts that co-expressed HydE, HydF and HydG [282]. $HydA^{\Delta EFG}$ was shown to require the assembly of the [4Fe-4S] subcluster region of the H-cluster assembled prior to the activation of the maturases [283], indicating the three *hyd* genes were responsible for the assembly of the 2Fe-subcluster.

#### 1.2.2.2.1 Scaffold Protein HydF

Initial observation of the sequence of HydF showed the presence of motifs characteristic of the nTPase family of enzymes in the N-terminal region [81, 118], as well as a [4Fe-4S] binding $CX_aCX_bC$ sequence at the C-terminal end [284, 285].The GTPase activity present in the enzyme [286] seems unrelated to the interactions with the hydrogenase [287], but rather with the interaction with the other maturases, as the rate of GTP hydrolysis increased around 50% in the presence of HydE and HydG [282, 287].

Spectroscopic data revealed the cluster assembly by HydF mimicked that of HydA with both [4Fe-4S] cluster and CO/CN bound [2Fe-2S] cluster spectroscopic signals being observed [284, 286, 287]. However, the mode of coordination of the clusters to the HydF protein is still poorly understood. An available apo-structure of HydF [133] showed the spatial positioning of a previously identified iron-sulfur cluster binding motif ($CxHx_{45-53}HCGGC$) [81, 118] Figure 12. Additionally, mutagenesis studies have demonstrated the importance of the residues of the motif in cluster binding [285, 288] and in hydrogenase activation [118]. As the six residues are insufficient to bind both a [4Fe-4S] cluster as well as a [2Fe-2S] cluster, either other necessary residues have yet to be identified, or the clusters are coordinated across subunit interfaces [285].

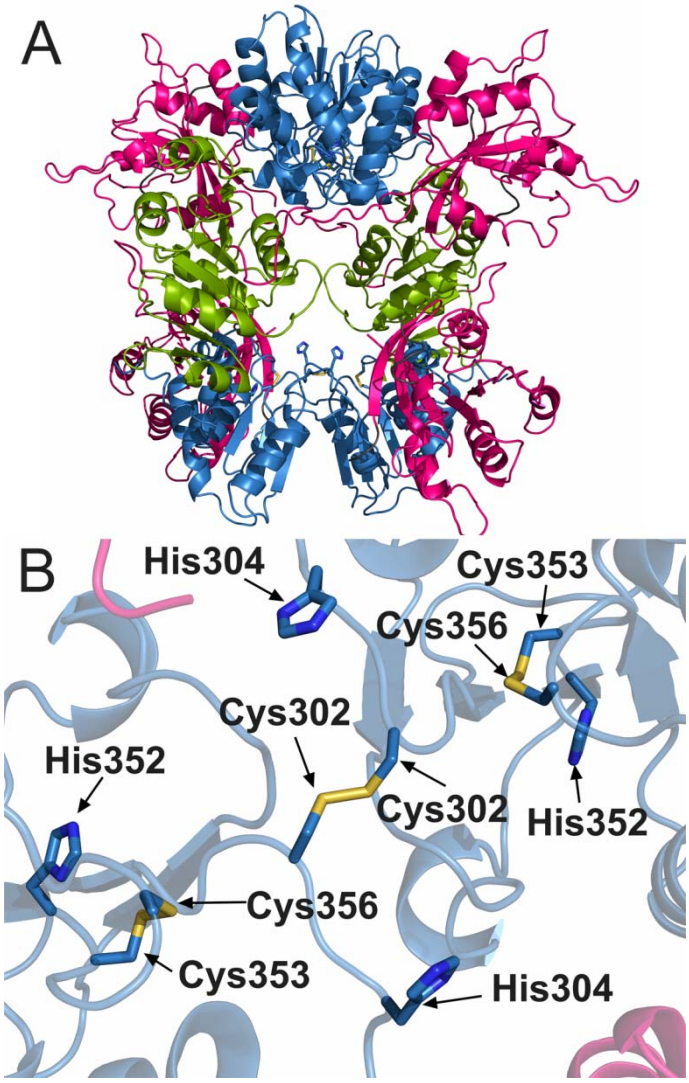



**Figure 12 Structure of HydF.**

**A:** Homotetrameric form of apo-HydF (PDBID: 3QQ5). GTP-binding domains (N-terminal Domain I) are shown in pink, with residues of the disordered loop proposed to bind GTP shown in dark grey. The domains responsible for the tetramerization of HydF (Domain II) are shown in green. The iron-sulfur cluster binding domains (Domain III) are shown in blue. **B:** Dimeric active site. Cysteines and histidines proposed to be responsible for cluster coordination are shown as stick models. Multimerization achieved using PDBe PISA interface [289], with a calculated interface area of 1479.4 Å and $\Delta G$ = -65.7 kJ $mol^{-1}$.

Expression studies of all three maturases in the absence of HydA showed an accumulation of HydF containing a binuclear [FeFe]-cluster with the inorganic ligands (CO and CN) bound to it [287, 290]. When this "loaded" HydF form alone was incubated with an apo-form of HydA (expressed in the absence of any of the maturases), an activation of HydA was observed, generating a functional hydrogenase [287, 290]. Additionally, completely synthetic H-cluster mimics can be

loaded onto HydF and are successfully transferred to an apo-hydrogenase [274, 275], providing mechanistic evidence of the stepwise nature of the maturation process.

These results led to the current model of cofactor formation (Figure 13). In this model, HydF is functioning as a scaffold where HydE and HydG perform chemical modifications to the H-cluster precursor [287, 291, 292]. The interaction with HydG provides the CO and CN ligands [47, 293], while the bridging dithiolate has been proposed to be formed in a reaction requiring HydE [294, 295]. The final step in such a biosynthetic pathway would be the transfer of the complete 2Fe subcluster from HydF **22** into the hydrogenase, as HydA already carries a [4Fe-4S] cluster that can be inserted by the housekeeping iron sulfur cluster biosynthesis machinery [283].

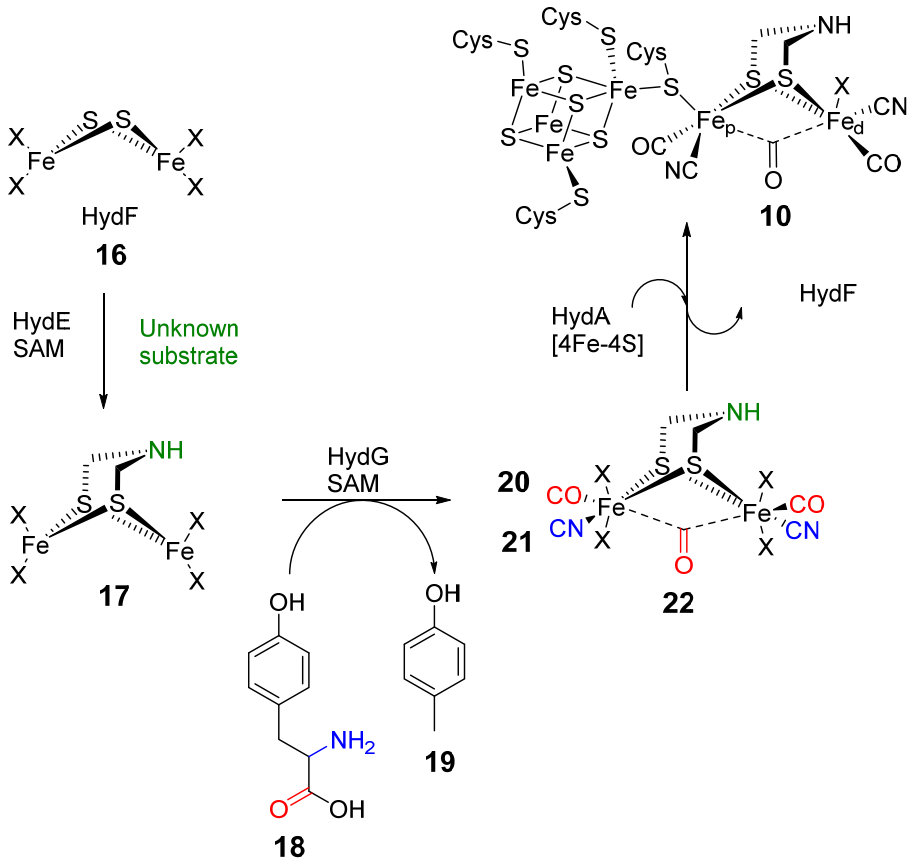



**Figure 13 Proposed biosynthesis pathway for HydA H Cluster.**

HydE uses a yet unknown substrate to form the dithiolate bridge on the [2Fe-2S] cluster **17** of HydF. HydG uses tyrosine **18** to produce *p*-cresol **19**, CO **20** and ¯CN **21**, the last two of which are added to the [2Fe-2S] cluster **19** as well. The matured [2Fe-2S] cluster is then transferred from HydF to HydA, activating the hydrogenase **10**. X - Residues coordinating to protein (cysteines). R-NH Dithiolate bridge. Adapted from Shepard *et al* [287].

Subsequent studies have, however, started calling this model into question. Recent studies of the mechanism of HydG show this enzyme may have as its product an organometallic cluster, the so called $Fe(CN)(CO)_2$ ligand which may be transferred intact to the subsequent maturation steps [296]. If this is proven to be the case, an as yet uncharacterised form of the cofactor is likely initially pre-synthetized by HydG, prior to the transfer into HydF. The formation of the azadithiolate bridge by HydE may occur using either HydG or HydF as substrates, prior to the activation of the hydrogenase.

A recent *in vitro* study [184] has gone even further by proposing that HydG may supply the [2Fe-2S] cluster with the CO and CN bound onto the hydrogenase. In this case, HydF would function solely as a transferase, stabilizing HydG in the activation of apo-HydA. These results seem, however, to completely contradict the ability to pre-load HydF with a synthetic mimic. As such, the hypothesis requires further studies to identify the actual biosynthetic pathway.

#### 1.2.2.2.2 Putative Small Ligand Dithiolate Synthase HydE

HydE is the only member of the maturases whose precise role in the maturation process, except that it is necessary [276], is still unknown [208, 297, 298]. Nonetheless it has been postulated that this enzyme is involved in the formation of the dithiolate bridge. The substrate that HydE uses is also yet to be identified, however it is thought to be a common bacterial metabolite, as heterologous expression of the HydEFG maturases in *E. coli* together with HydA leads to an active hydrogenase [282, 299].

Sequence alignments [150] confirmed that HydE was a RSAM enzyme. A second $CX_aCX_bC$ motif can be found in its sequence, which, in the light of spectroscopic studies, was proposed to bind to another [4Fe-4S] cluster [80]. The presence of an auxiliary cluster was further confirmed when the structure was solved, though only a $[2Fe\text{-}2S]_{Aux}$ cluster was observed, possibly due to some oxidative degradation of the cluster [297].

Due to the presence of an auxiliary cluster and a high degree of sequence similarity, one hypothesis could be that HydE and biotin synthase (BioB) [300] insert sulfur atoms by related mechanisms, sacrificing the auxiliary cluster as a sulfur donor, for the formation of the small dithiolate bridge in the case of HydE [287]. However, the presence of the auxiliary cluster is not completely conserved across all HydE's, and mutagenesis studies showed that HydA was still active even after the removal of the auxiliary cluster, discrediting this sacrificial cluster hypothesis [297]. An alternative proposal suggests this auxiliary cluster function as a protective agent against external oxidants, being oxidized before the catalytic RSAM cluster suffers degradation [301].

The relatively high sequence homology between HydG and HydE led to the suggestion that the dithiolate bridge could be derived from a free tyrosine molecule [302]. However a spectroscopic study of matured HydA studied in the presence of radiolabelled $^{15}$N-L-tyrosine, determined that the source of the azadithiolate is not likely to be L-tyrosine [303].

A more recent study tried to identify potential substrates for HydE, showing that thiols promote SAM cleavage. The formation of the DOA$^{\bullet}$ radical (see section 1.1.1) was confirmed by measuring deuterium incorporation into the resulting deutero-DOA. In particular, compounds with a functional group that have a negative charge at a distance of approximately 5 Å from the thiol, could be likely candidates as the substrate for the enzyme [304]. These results are still preliminary as the expected by-products of the cleavage of the tested compounds (D/L-cysteine, L-homocysteine, mercaptopyruvate, DTT, 3-mercaptopropionic acid and Coenzyme M) were not observed.

While it is still broadly accepted that HydE is responsible for the formation of the azadithiolate bridge following the assembly of the Fe(CO)$_2$CN-[Fe$_3$-S$_4$] complex [305], a recent *in vitro* study achieved HydA maturation in the absence of HydE [184]. The authors proposed a different role for this maturase in assisting with the translocation of the H-cluster precursor from HydG and HydF into apo-HydA.

#### 1.2.2.2.3 Tyrosine Lyase HydG

The *hydG* gene's sequence, was initially identified as part of the RSAM superfamily of enzymes [81], a result which was confirmed by the presence of a characteristic [4Fe-4S] EPR spectra and reductive cleavage of SAM [80]. Analysis of HydG sequences from different organisms revealed the presence of N-terminal and C-terminal extensions, the latter three conserved cysteines which were postulated to bind to a second [4Fe-4S] cluster [80, 302].

The high sequence similarity between HydG and the enzyme tyrosine lyase (ThiH) led to the identification of a common substrate, L-tyrosine, for HydG [302] and to the newly classification of a RSAM subfamily of amino acid lyases [306], of which tryptophan lyase NosL [307] and CofH [308] are also members.

The enzymatic mechanism for HydG has been widely studied over the past decade. From the initial observation that *p*-cresol **19** was released after tyrosine cleavage [302], further studies identified the products of the overall reaction as cyanide 20 [47] and carbon monoxide 19 [293] (Figure 14). Kinetic analysis of HydG catalysed reactions revealed a cyanide $k_{cat}$ of 20 x $10^{-4}$ $s^{-1}$,

with close to a 1:1 formation of *p*-cresol [47]. The formation of carbon monoxide arising from tyrosine cleavage was later shown to have a similar $k_{cat}$ of $11x10^{-4}$ $s^{-1}$ [293]. Using labelled $^{13}C$ and $^{15}N$ L-tyrosine samples it was also revealed that all five diatomic ligands in the H-cluster (Figure 10 C) derive from the carboxylate and amino substituents of tyrosine [309].

The proposed auxiliary cluster coordinating cysteines were initially shown necessary for the activation of HydA [118], indicating the auxiliary cluster had a role in the activity of HydG. When the entire C-terminal stretch of HydG was removed by mutagenesis (ΔCTD) only small amounts of *p*-cresol were observed from the incubation with L-tyrosine (less than 2% compared with the wild type activity [310]). An in depth kinetic analysis of the same mutant revealed no change on the rate of tyrosine cleavage (similar $k_{cat}$ observed for the WT and ΔCTD) but instead, a remarkable effect on tyrosine affinity, with a much higher $K_M$ observed for ΔCTD [311]. The lack of the C-terminal domain was also shown to lead to virtually no formation of the diatomic products [311, 312]. Instead, a significant accumulation of glyoxylate was observed [311]. Glyoxylate was shown to form during turnover of ThiH [313], presumed to be the product of dehydroglycine hydrolysis, indicating a strong similarity between the truncated HydG and ThiH [311].

Intriguingly, the dependence of the overall reaction on the C-terminus is not indicative of the necessity for the auxiliary cluster to be present. Mutations of two auxiliary cluster coordinating cysteines to serines (SxxS) show complete abolishment of carbon monoxide formation, but with some cyanide formation still being observed [311, 312]. However, it is yet unclear whether the two diatomic ligands are formed by different reaction pathways, like was initially proposed [312], as the timescale of cyanide formation by the cysteine mutant is a hundred-fold slower than the wild type enzyme, which may be indicative of a secondary path for cyanide formation that is not enzymatically relevant [314].

These results seemed to suggest that the N-terminal RSAM [4Fe-4S] cluster in HydG was responsible for cleavage of SAM and L-tyrosine, while the C-terminus was necessary for the subsequent formation of the diatomic products [310-312]. The mechanism of action of the enzyme could then be divided into two halves: the RSAM conserved reductive cleavage of the tyrosine $C_\alpha$-$C_\beta$ bond and the breakdown of a glycine-related intermediate into the diatomic products.

Following CO/CN formation, the diatomic ligands, and potentially iron, need to be moved through the maturation pathway of HydA. It was initially thought that the inorganic components were brought together on the HydF scaffold protein [35, 208, 291, 293]. More recent studies have proposed the formation of an $Fe(CO)_2CN$-[Fe3-S4] complex, where the unique iron of the auxiliary cluster is loaded with the inorganic ligands (Figure 14). The iron-ligand complex would then be

transferred onto HydF in what can be effectively considered an iron-centred self-sacrifice of the auxiliary cluster [296, 314]. These results were also confirmed by maturation of HydA with a $^{57}$Fe enriched HydG, which displayed the radioactively labelled irons in the [2Fe] H subcluster [296]. Together with the studies that propose a transfer of the loaded iron directly to the apo-hydrogenase [184], it appears that the model for hydrogenase maturation is far from being completely understood, although the role of HydG is undoubtedly a vital component of the pathway.

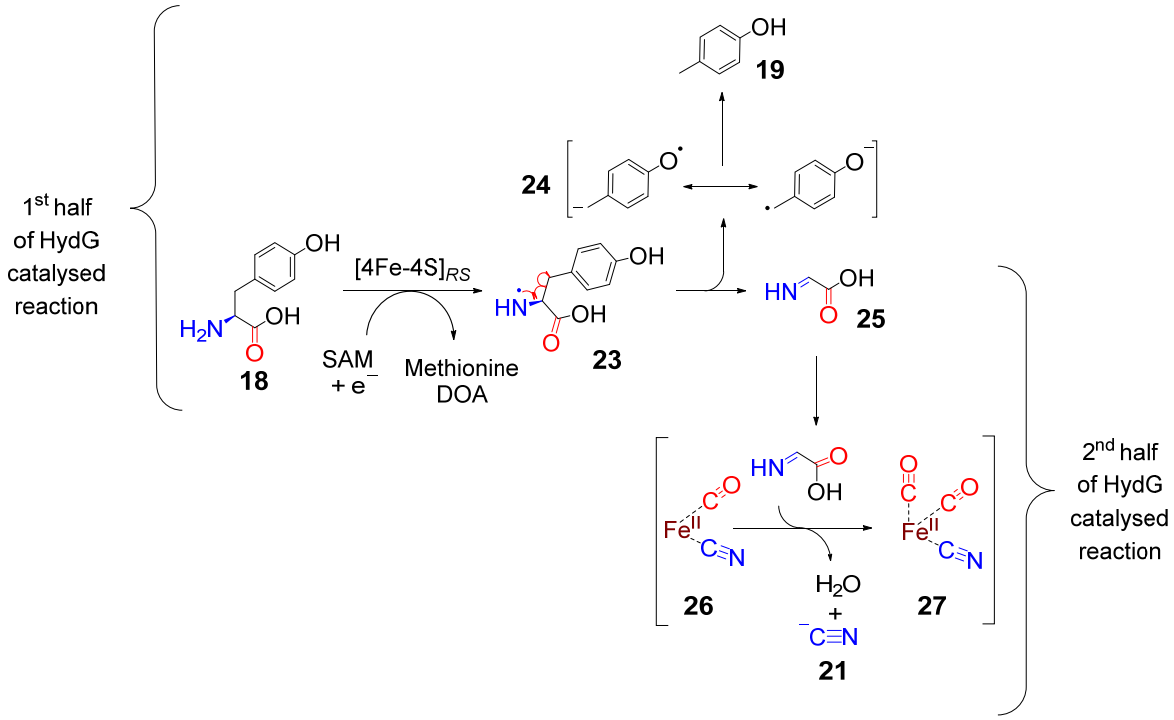



**Figure 14 Proposed HydG catalysed mechanism of L-tyrosine cleavage and diatomic product formation.**

The abstraction of an hydrogen from the amino group of L-tyrosine leads to its heterolytic cleavage into DHG **25** and an oxidobenzyl radical **24** that can be followed spectroscopically by the formation of *p*-cresol. The second half of the reaction is less understood, but leads to the formation of iron-bound CO and $^{-}$CN molecules. Adapted from Kuchenreuther *et al* [314].

### 1.2.2.3 Understanding the Mechanism of HydG

#### 1.2.2.3.1 Mechanistic Relationship Between ThiH and HydG

As briefly mentioned previously, the discovery that HydG and ThiH share a high sequence homology (27%), being both members of the RSAM superfamily, led to the identification of a common substrate and mechanism of action [302].

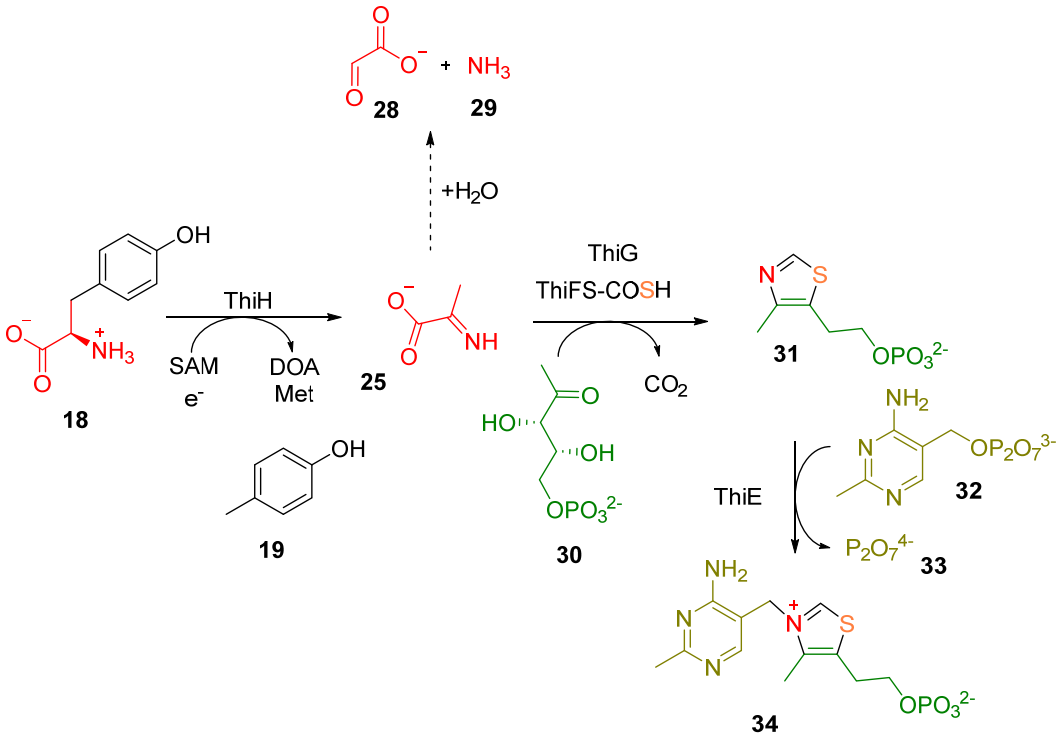



**Figure 15 General mechanism of biosynthesis of thiamine phosphate.**

ThiH catalyses the first step of the reaction, cleaving tyrosine **18** to yield DHG **25**. The *in vivo* relevant reaction begins with a ThiG catalysed condensation of dehydroglycine **25** (from ThiH), thiocarboxylate (from ThiFS) and 1-deoxy-D-xylulose 5-phosphate **30**, to yield thiazole-phosphate carboxylate tautomer **31**. ThiE then facilitates the aromatization of **31** with hydroxymethyl pyrimidine pyrophosphate **32**, yielding thiamine **34**. In the absence of its enzymatic partners, ThiH releases DHG **25**, which is hydrolytically cleaved into glyoxylate **28** and ammonia **29**. Adapted from Jurgenson *et al.* [315].

ThiH was shown to be fully active in the presence of SAM and L-tyrosine under reducing conditions, catalysing the $C_\alpha$-$C_\beta$ bond cleavage of tyrosine as part of a complex that forms the thiamine thiazole moiety [313, 316].The resulting product of the reaction was proposed to be a C-centred glycine radical and quinone methide or dehydroglycine and a phenyl radical, depending on whether the cleavage of the substrate was homolytic or heterolytic, respectively. Density functional theory (DFT) calculations predicted the homolytic pathway of $C_\alpha$-$C_\beta$ bond cleavage in tyrosine to be slightly thermodynamically more favourable, but not significantly so, allowing the prediction of different mechanisms of tyrosine hydrogen abstraction and cleavage (Figure 16) [316]. If the associated component ThiG is not present to uptake dehydroglycine or the glycyl radical for thiazole formation, ThiH cleaves tyrosine to release *p*-cresol and glyoxylate, the latter being the product of dehydroglycine hydrolysis *in vitro* [313].

Due to the similarities in both sequence and substrate, it has been suggested that the mechanism of tyrosine fragmentation is conserved between HydG and ThiH [302]. However, unlike in the *in vitro* reaction catalysed by ThiH, little amounts of glyoxylate are observed in L-tyrosine cleavage by HydG [47]. These results suggest DHG (or the related glycyl radical) functions as an intermediate on the pathway to the formation of CO and cyanide [47, 293, 302, 310, 311].

#### 1.2.2.3.2 The Mechanism of Action of HydG

There is a general acceptance that the reaction catalysed by HydG starts with the reductive cleavage of SAM to create DOA$^\bullet$ **3** (Figure 1) which will then abstract an H atom from L-tyrosine, leading to $C_\alpha$-$C_\beta$ bond cleavage (Figure 14). Different mechanisms for the hydrogen abstraction and $C_\alpha$-$C_\beta$ bond cleavage have been put forward over the years (Figure 16).

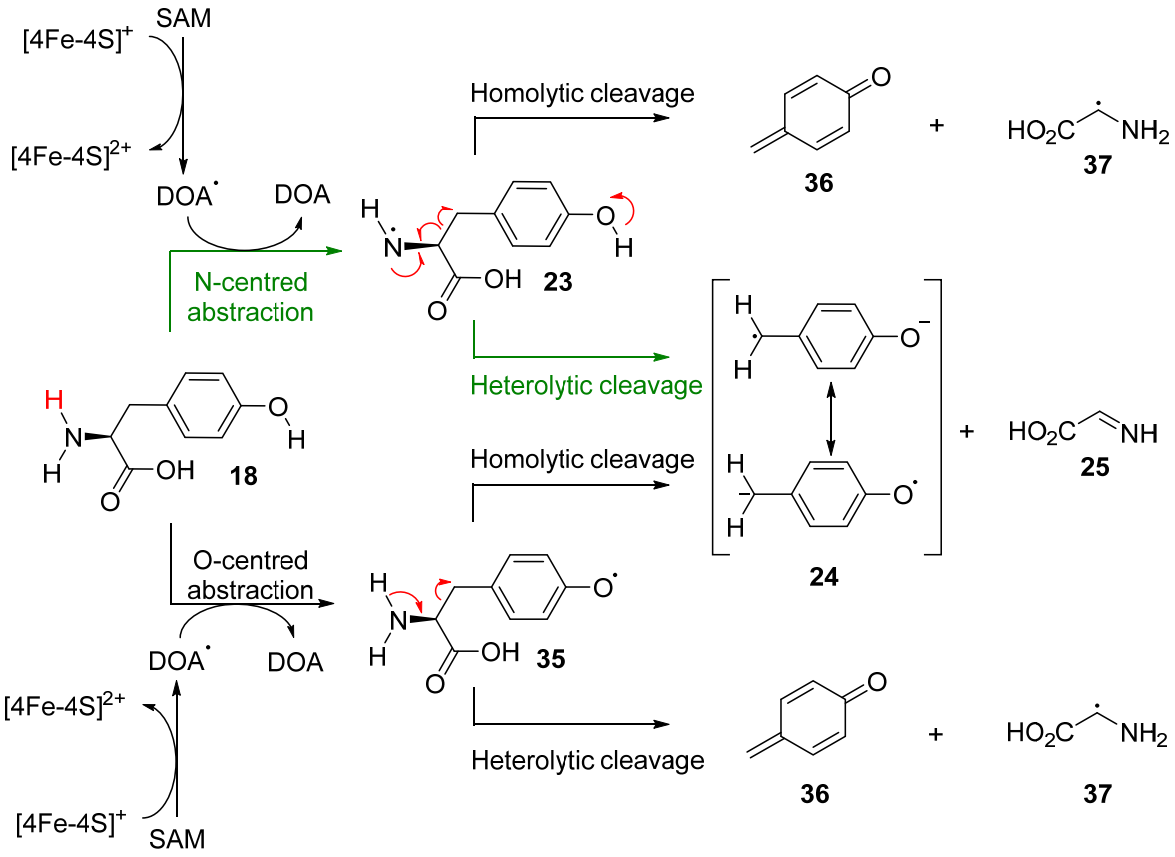



**Figure 16 Alternative pathways of tyrosine reductive cleavage.**

The position of hydrogen abstraction by the DOA$^{\bullet}$ occurs either in the phenolic region of L-tyrosine **35** [312]; or in the amino group **23** [317]. The heterolytic cleavage of the O-based tyrosyl radical or homolytic cleavage of the N-based tyrosyl radical are predicted to yield a glycyl radical **37** and *p*-quinone methide **36**. Conversely the homolytic cleavage of the O-based tyrosyl radical or heterolytic cleavage of the N-based tyrosyl radical are predicted to yield a oxidobenzyl radical **24** and dehydroglycine **25**. The currently most accepted pathway is presented in green. Adapted from Suess *et al.* [318].

The initial abstraction step was shown to occur in a reversible manner in a solvent-exchangeable position [319], but there was no unambiguous determination of the abstraction location. Two possibilities arise from the results: an abstraction in the phenolic –OH **35** or in the amino –NH **23** of tyrosine. Additionally, the characterization of the intermediates of the reaction by EPR did not help to distinguish between either abstraction locations as both were considered equally favourable and lead to a common intermediate (shown as the radical anion in Figure 16) [314].

Initially most studies favoured the initial hydrogen abstraction step to occur in the hydroxyl group of the phenol of tyrosine [293, 302, 310, 311, 314, 320] (Figure 16, **35**). More recently, structural information obtained from the highly related enzyme NosL seems to indicate hydrogen abstraction occurring from the amino group of tryptophan [317, 321] (Figure 17). As NosL also promotes $C_\alpha$-$C_\beta$ bond cleavage in its substrate, the possibility of a conserved mechanism of substrate binding and abstraction could be predicted (Figure 16, green).

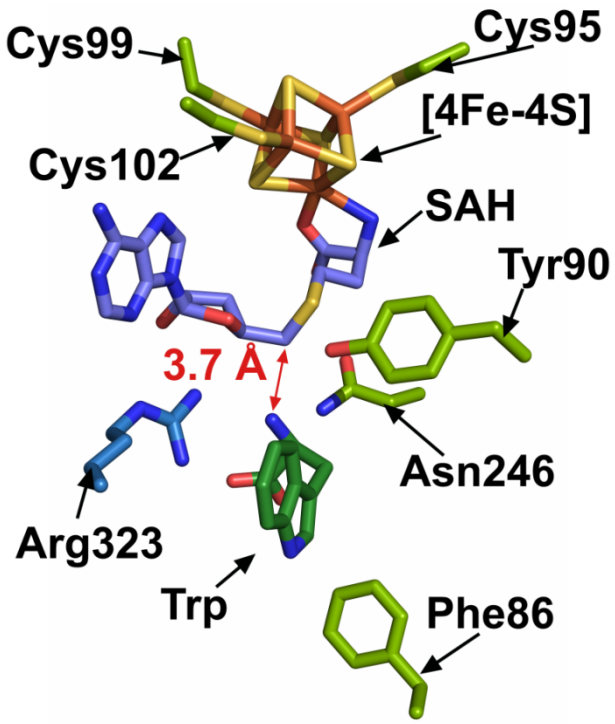



**Figure 17 Structure of NosL with its substrate, tryptophan, bound.**

The structure of NosL (PDBID: 4R33) shows SAH bound to the [4Fe-4S] and with the SAH C5 carbon at a 3.7 Å from the proposed abstraction group, the amino substituent of tryptophan. Adapted from Nicolet *et al.* [321].

Despite initial contention over whether tyrosine cleavage was homolytic and yielding a glycyl radical [312], or heterolytic and forming dehydroglycine [47, 293, 310, 311], a recent rapid freeze-quench EPR spectroscopy study has provided evidence for an intermediate *p*-cresol radical (oxidobenzyl radical, Figure 16) [314], indicating dehydroglycine is an intermediate to cyanide and CO formation.

The mechanism by which CO/$^{-}$CN are formed from DHG is much less well understood. There is precedence for cyanide formation from glycine, using a DHG intermediate, in cyanide synthase HCNAbc [322], but the reaction yields carbon dioxide. Alternatively, the decarbonylation of DHG

could yield both cyanide and CO in a single step, but decarbonylation only has precedence in α-tertiary amino acids heated in the presence of (activating) phosphorous oxychloride in inert solvents [323].

A different proposal, which attempts to justify the measurements of free cyanide in ΔCTD [311, 312], predicts two deprotonation events occurring to DHG, resulting in the cleavage of the C-C bond and release of $^{-}CN$ and a $^{-}CO_2H$ species [317]. This latter product of DHG scission could then bind to an Fe d orbital in the auxiliary cluster using its lone pair, to generate the CO ligand in a similar reaction to anaerobic CO dehydrogenase (Figure 18) [317, 324].

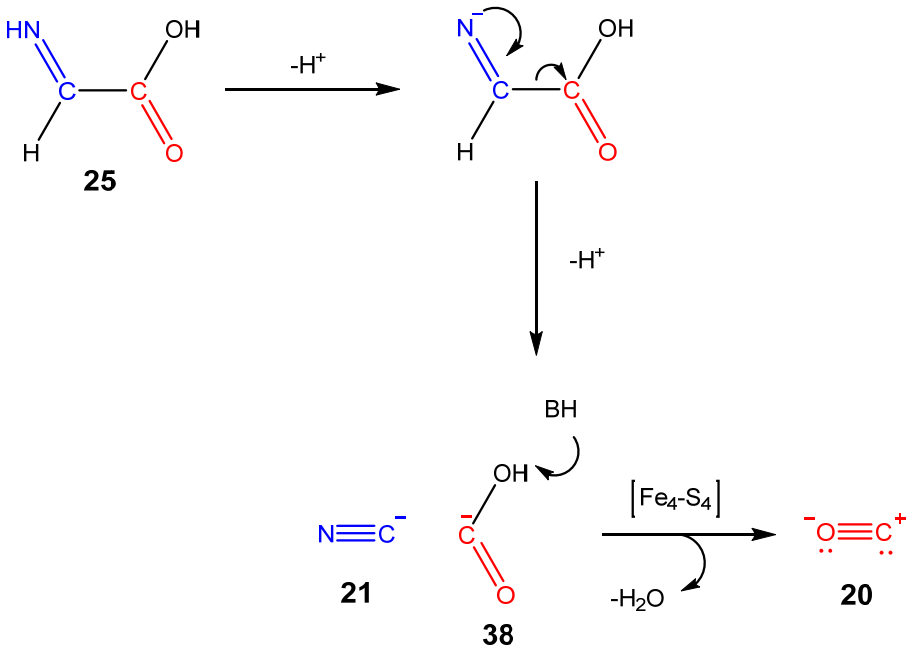
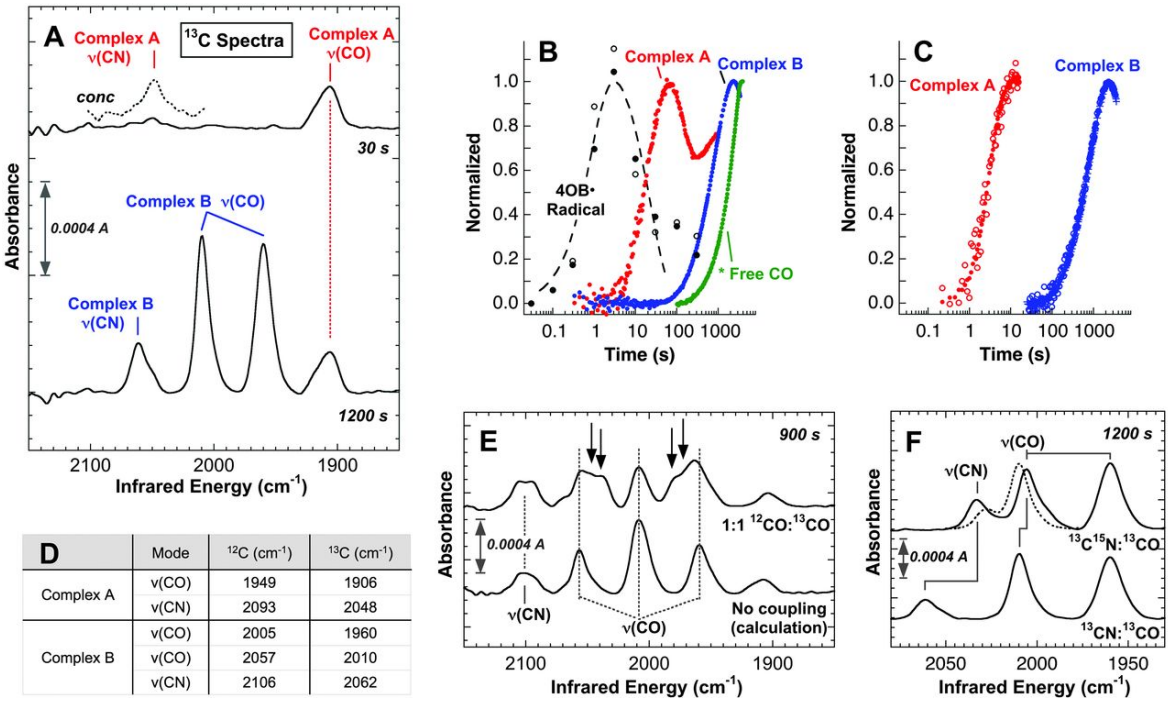



**Figure 18 Mechanism of CO/$^{-}$CN formation from DHG through cluster-independent deprotonation.**

Two initial deprotonation events of DHG **25** lead to a cluster independent release of cyanide **21**. The subsequent protonation of the carboxylate **38**, in a reaction dependent on the [4Fe-4S] as a cofactor, leads to the release of carbon monoxide **20** as the products of DHG degradation.

Very recently, an EPR/FTIR study followed the reaction of HydG and revealed that shortly after the formation of oxidobenzyl radical (Figure 16), signal for a [Fe(CO)$^{-}$CN] species was observed (Figure 19, Complex A). Over time, however, the signal suffered significant changes (Figure 19, Complex B) and density functional theory (DFT) models agreed with the proposal of a $Fe(CO)_2(CN)$ synthon being formed. These studies allowed the update of the mechanism of HydG, leading onto the current understanding of HydG's reaction mechanism, shown in Figure 14.




**D**

| | Mode | $^{12}$C (cm$^{-1}$) | $^{13}$C (cm$^{-1}$) |
|---|---|---|---|
| Complex A | ν(CO) | 1949 | 1906 |
| | ν(CN) | 2093 | 2048 |
| Complex B | ν(CO) | 2005 | 1960 |
| | ν(CO) | 2057 | 2010 |
| | ν(CN) | 2106 | 2062 |

**Figure 19 SF-FTIR and EPR spectroscopies following the formation of iron-bound diatomic products of HydG.**

**A:** SF-FTIR spectra showing the stretch vibrations of CO and ¯CN ligands, changing over time from a signal of one cyanide and one carbon monoxide (red), to a complex signal with two carbon monoxide and one cyanide (blue). **B:** Time dependency of formation and decay of mechanistic relevant species: 4OB•, determined by EPR spectroscopy, consistent with the Cα-Cβ cleavage of tyrosine [314]; and SF-FTIR data of the peak heights for complex A (red), complex B (blue); free CO captured by myoglobin (green). **C:** Comparison of the SF-FTIR peak heights for Complex A (red) and Complex B (blue), with ν($^{13}$CN) indicated by ° and ν($^{13}$CN) indicated by •. **D:** Table of frequencies for observed IR bands of complexes A and B prepared by using Tyr (middle column) or ($^{13}$C)$_9$-Tyr (right-most column). **E:** SF-FTIR spectrum of complex B, at 900 s, using a 1:1 mixture of Tyr and $^{13}$COO-Tyr (Top). Arrows indicate new bands not present in measurements with fully Tyr or fully $^{13}$COO-Tyr measurements, as shown in the average of the $^{12}$CO and $^{13}$CO product B spectra (bottom). The appearance of new bands can be ascribed as the mixing of the two fundamental ν(CO) modes. **F:** SF-FTIR spectrum of complex B, at 1200 s, using ($^{13}$C)$_9$$^{15}$N-Tyr (top) and ($^{13}$C)$_9$Tyr (bottom). Predicted CN and CO bands for $^{13}$C$^{15}$N-containing complex B in the absence of ν(CN)/ν(CO) vibrational mixing is shown as a dotted line. The differences shown indicate that both the ν(CN) and the ν(CO) modes are also vibrationally mixed. Taken together, these coupling data suggest that all three ligands (two CO and one CN$^-$) in complex B are bound to a single Fe-centre. Figure from the studies by Kuchenreuther *et al.* [296], reprinted with permission.

## 1.3 Lipoic Acid and Lipoyl Synthase

Over half a century ago, several research groups, working in parallel, made the key discovery that bacterial growth can be enhanced in the presence of the enzymatic cofactor 6,8-dithiooctanoic acid, or α-lipoic acid (LA) [325-328]. The oxidation of lipoic acid leads to a disulfide linkage, with the formation of a five-membered ring including three backbone carbons (Figure 20).

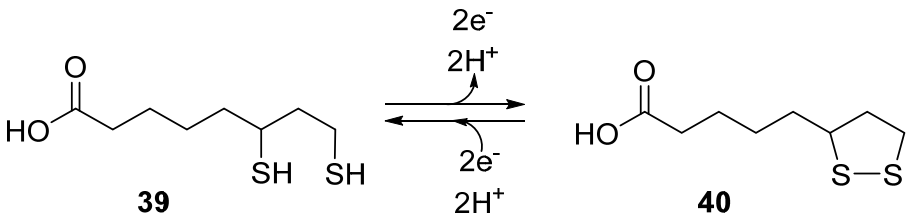



**Figure 20 α-Lipoic acid, in its reduced (dihydro-α-lipoic acid) and oxidized forms.**

Lipoic acid (LA) was identified as a good medical supplement, showing positive effects in the treatment of side effects of Alzheimer's [329, 330] and diabetes [331-333] among others [334-336]. It was initially proposed these results arose from the ability of lipoic acid for radical scavenging as well as metal coordination [337, 338], though later studies have shown that lipoic acid can also have some pro-oxidant activity *in vivo* (see reviews [339, 340]), indicating there are multiple mechanisms by which the advantageous effects of lipoic acid should occur.

Apart from the potential medical uses of LA, it is known that LA is a crucial cofactor of several α-ketoacid dehydrogenase and glycine cleavage system multienzyme complexes [341-343]. In its bioactive form LA is bound covalently by an amide linkage to the Ɛ-amino group of a target lysine residue on the lipoyl bearing protein domain [344]. The lipoyl-lysine functions like a flexible arm enabling the delivery of intermediates between different subunits of the complex [345].

When attached to its appropriate acceptor protein, the lipoyl cofactor plays two important roles, functioning as both an electron acceptor [346, 347] and a shuttle for the reaction intermediates [348] amongst the active sites of the subunits that form the complex. The most widely known multienzyme complexes [171, 349] that use lipoic acid as a cofactor are the pyruvate dehydrogenase complex (PDH) [350], 2-oxoglutarate dehydrogenase complex (ODH) [351], branched chain oxoacid dehydrogenase complex (BCDH) [352] and acetoin dehydrogenase complex (ADH) [353], collectively known as the 2-oxoacid dehydrogenase complexes. Additionally, the glycine cleavage system (GCV) also uses lipoic acid as a cofactor [354] and possesses many

similarities with the 2-oxoacid dehydrogenase multienzyme complexes, despite all enzymes being independent in this case.

### 1.3.1 2—Oxoacid Dehydrogenases

The 2-oxoacid dehydrogenases are multimeric enzyme systems that carry out the oxidative decarboxylation of different substrates, following the same general mechanism (Figure 21).

Both pyruvate dehydrogenase (PDH) complex and 2-oxoglutarate dehydrogenase complex (ODH) (also known as $\alpha$-ketoglutarate dehydrogenase complex), catalyse critical steps within the citric acid cycle (TCA), making LA an important cofactor in aerobic respiration [341, 355, 356]. Lipoic acid is also required for catabolism of acetoin by the acetoin dehydrogenase (ADH) complex. ADH is mostly homologous to PDH, producing acetyl CoA and acetaldehyde in organisms that produce acetoin as a fermentation product [357]. The final member of this group of enzymes, branched-chain amino acid dehydrogenase (BCDH), is involved in the degradation of all branched-chain amino acids (valine, leucine and isoleucine) [358, 359] to their respective branched-chain acyl-CoAs which are used to make branched-chain fatty acids [358, 359].

Each complex consists of three enzyme components, two of them ($E_1$ and $E_2$) unique to each enzyme, though functionally similar, and a third ($E_3$), which is encoded by the same gene across the family [360, 361]. The central core of the complexes are made up of multimers of the lipoic acid-containing subunits ($E_2$ subunits) arranged in octahedral or dodecahedral formation depending on the complex and the organism [341].

The E1 component is a specific diphosphate-dependent decarboxylase using thiamine pyrophosphate (TPP) **41** as a cofactor, producing an acyl-TPP conjugate **43** [362]. The thiazole ring of TPP initiates the reaction by the formation of a carbanion which attacks the carbonyl carbon of the substrate oxoacid (**42**). The carbonyl-carboxyl bond in the substrate is cleaved, with the release of $CO_2$ (or acetaldehyde in the case of α-ketoacid dehydrogenase), and reductive acetylation of the lipoyl group attaches the product to the E2 subunit [363, 364].

Lipoic acid is covalently attached to the ε-amino group of a conserved lysine of the lipoyl domain (LD), which protrude from the E2 subunits on flexible proline-rich linkers and carries reaction intermediates between the various subunits [345].The disulfide bond is reduced by attacking the TPP-carbonyl double bond, resulting in a thioester **45** which is moved from the E1 subunit, to interact with the thiol of CoA **47** in the E2 subunit, while lipoic acid is reoxidized by E3 [341].

The E2 component is a dihydrolipoyl acyl-transferase, and is the key to the structure and function of the multienzyme complex. It is composed is composed of a catalytic C-terminus core domain, which contains the acyl-transferase site, a peripheral small subunit binding domain, where the E1 and E3 components bind, and one to three lipoyl domain region at the N-terminus, where the lipoic acid is attached to the specific lysine residue [365-368].

The E3 component is a dihydrolipoyl dehydrogenase that serves to regenerate the disulfide bond of the lipoic acid cofactor, readying it for a new cycle of catalysis [171, 345] and transfers the electrons to $NAD^+$ **51** via a flavin cofactor **49**. The energy from decarboxylation is conserved by thioester formation and through formation of NADH **52** [345, 347, 352].

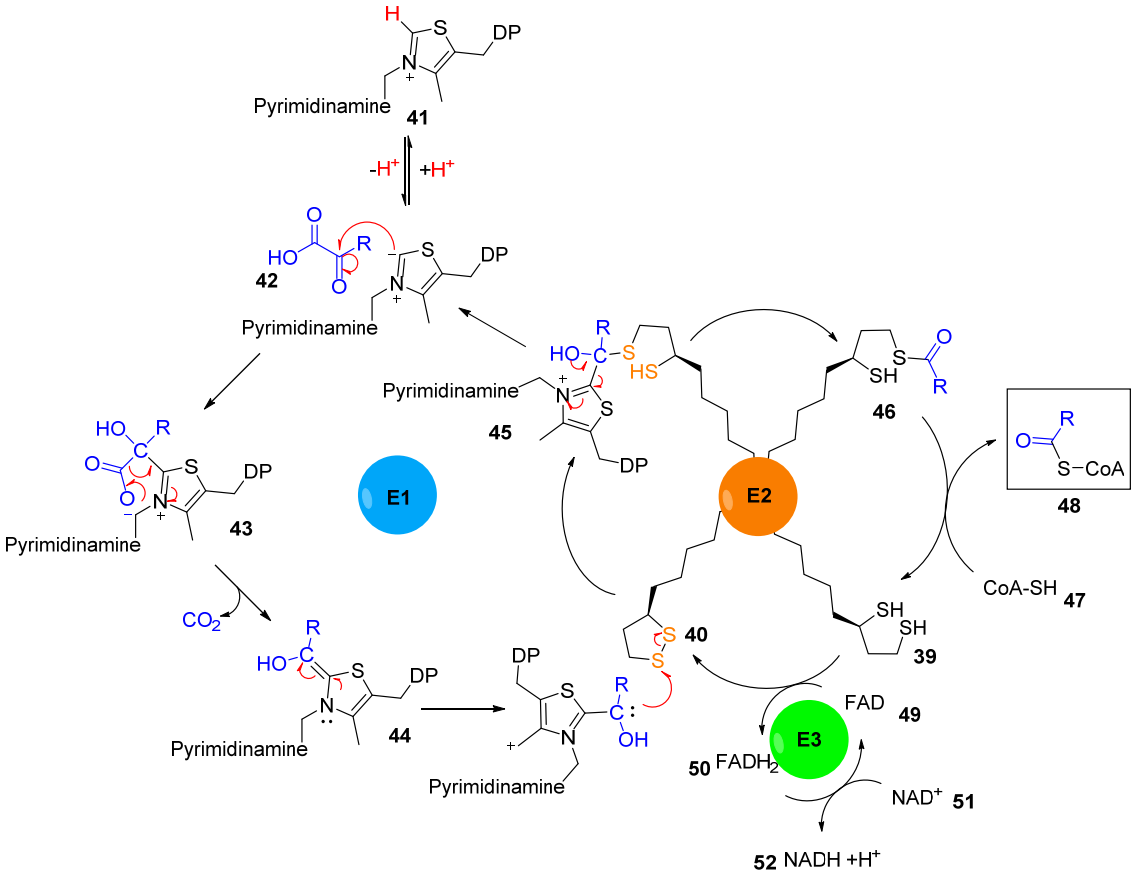



**Figure 21 E1 subunit catalyses the decarboxylation of a α-ketoacid.**

An acyl group is transferred to the lipoyl E2-subunit, with reduction of the lipoyl cofactor **40**. The acyl group is transferred to Co-Enzyme A **48**, releasing a reduced dihidrolipoyl cofactor which can be oxidized by the FAD dependent E3 subunit.E1 (2-Oxoacid dehydrogenase) and E2 (Dihydrolipoyl acyltransferase) are unique subunits for each system. The E3 subunit (Dihydrolipoyl dehydrogenase) is analogous across these complexes. Adapted from Mooney [369].

### 1.3.2 Glycine Cleavage System

The glycine cleavage system (GCV), also known as glycine decarboxylase system and glycine dehydrogenase system, is ubiquitous in all domains of life and was first identified in 1961 by Sagers and Gunsalus [370, 371]. Despite having similarities in the mechanism with 2-oxoacid dehydrogenases, GCV is technically not a complex, as all the component proteins dissociate easily and behave as independent proteins.

The GCV enzymatic system catalyses the oxidative decarboxylation and deamination of glycine **53**, with formation of carbon dioxide **59**, ammonia **29** and the concomitant reduction of $NAD^+$ **51** to NADH **52** [370]. The remaining carbon, the methylene carbon of glycine, is then transferred to 5,6,7,8-tetrahydropteroylpolyglutamate (tetrahydrofolate, THF) **57** to form $N^5$,$N^{10}$-methylene-5,6,7,8-tetrahydropteroylpolyglutamate ($CH_2$–THF) **58** [370-372] (Figure 22).

The system is composed of four enzymes, three analogous to the components of the 2-oxoacid dehydrogenase complexes and a transferase protein. With a similar role to that of $E_2$, the H-protein is a small heat-stable protein of about 14 kDa [343], containing a lipoyl-lysine arm that shuttles the reaction intermediates between the active sites of the components of the glycine cleavage system [345, 373, 374]. The P-protein (glycine dehydrogenase) is similar to the $E_1$ unit but uses pyridoxal phosphate **54** instead of TPP **41** as the cofactor for the decarboxylation of glycine **53** (Figure 22) [375].

The enzyme which does not have an homologue in 2-ODH multienzyme complexes is T-protein, an aminomethyltransferase that that transfers methylamine from lipoate complex **56** to tetrahydrofolate (THF) **57**, yielding methylene-THF **58** and ammonia **29** [376, 377]. Lastly the L-protein (dihydrolipoamide dehydrogenase) re-oxidizes the lipoate **40** from the H-protein with reduction of its FAD **49** cofactor [171, 378]. The L-protein is in all regards similar to E3, and can be effectively replaced by the lipoamide dehydrogenase in some organisms [379, 380].

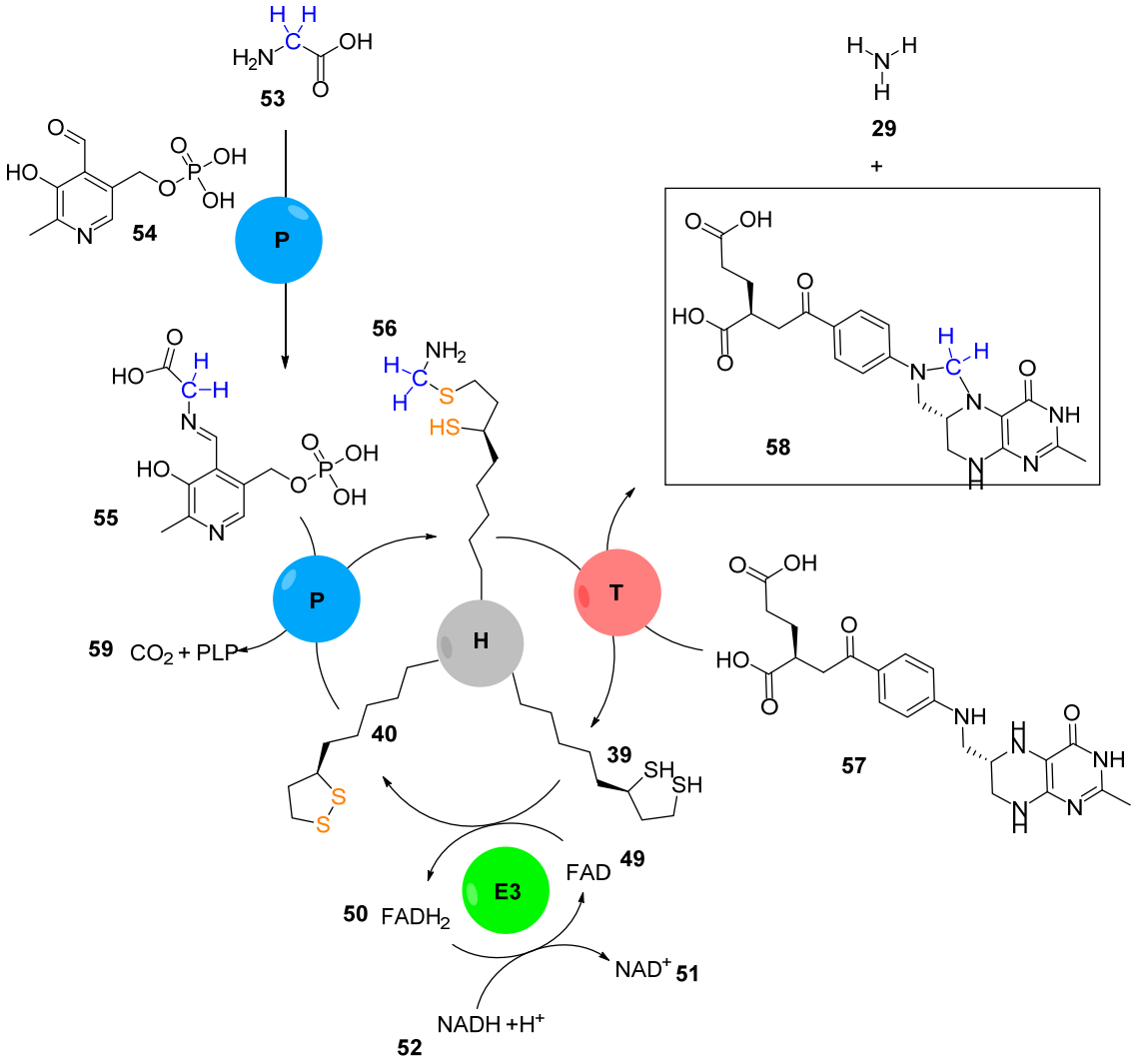



**Figure 22 Reaction mechanism of the glycine cleavage system.**

The P-protein initially catalyses the formation of a Schiff base between glycine **53** and PLP **54**, as well as its decarboxylation, transferring methylamine into the lipoyl cofactor of the H-protein. Protein T methylates THF **57** with production of ammonia and reduction of the lipoic acid. L-protein oxidizes the dihydrolipoyl group for a new cycle of the reaction. Adapted from Perham *et al.* [345].

### 1.3.3 Pathways of Synthesis/Incorporation of Lipoic Acid

#### 1.3.3.1 Exogenous Uptake of α-Lipoic Acid

Due to its importance in central metabolism most organisms possess more than one pathway to incorporate lipoic acid [381, 382]. A mechanism of exogenous uptake of the cofactor shows remarkable variation in different organisms. In *Escherichia coli*, the most well researched pathway, free lipoic acid that is obtained from the environment is attached into a lipoyl-carrier protein (LCP) **60** by the enzyme LplA, in a reaction which requires ATP and magnesium for lipoic acid activation [383, 384]. LplA has also been shown to transfer octanoic acid into the LCP **61** albeit at a lower efficiency when compared to the attachment of lipoic acid [385] (Figure 24).

The Gram-positive *Bacillus subtilis* encodes three LplA homologues [386] where one acts in a manner similar to LplA, another functions as an octanoyltransferase [386, 387] and the third is an amidotransferase, which catalyses the last transfer onto the lipoyl domains of the multienzyme complexes [386, 388]. The archaea *Thermoplasma acidophilum* contains LplA ligase that has two separately encoded subunits, LplA and LplB, which are required for the formation of the acyl-AMP intermediate [389, 390]. Additionally, the eukaryotic pathway for the uptake of lipoic acid is not well understood and may increase the diversity of mechanisms present in nature.

#### 1.3.3.2 Endogenous Production of α-Lipoic Acid

Neither mutations in the *lplA* gene nor growth of cells in the absence of external lipoic acid abolish the presence of lipoylated E2 domains (21). These results, together with the identification of the role of the *lipB* gene in the formation of octanoyl- and lipoyl- E2 domains, led to the suggestion of an endogenous pathway, where the product of *lipB* would be involved [381]. LipB was shown to be a lipoyl (octanoyl)-acyl carrier protein:protein transferase, catalyses the transfer of both octanoyl- and lipoyl- groups from the acyl-acyl carrier protein to the LD [385], with high affinity for octanoyl-ACP [391].

Octanoate **62** was shown to serve as a precursor to lipoic acid **40** in *E. coli*. The sulfur atoms were exchanged for hydrogen atoms at carbons 6 and 8 without loss of protons at carbons 5 and 7, suggesting a direct replacement without an unsaturated intermediate [392]. Introduction of the sulfur at carbon 6 proceeded with inversion of configuration [393, 394].

In 1993, another enzyme relevant to the endogenous synthesis of the lipoyl group was purified [395] as the product of the *lipA* gene. It was shown to have iron-sulfur clusters present [396, 397] and reported lipoyl synthase activity [395, 397, 398]. Later studies showed the preferred

substrate for LipA was an octanoyl-lipoyl domain over octanoyl-ACP **64** [349], strongly indicating LipA catalysis occurs after octanoyl transfer from LipB to an E2 domain.

The LipB/LipA mechanism of lipoyl synthesis can thus be divided into two steps. First, an octanoyltransferase (LipB) transfers an octanoyl moiety from the acyl carrier protein (ACP) **64** of fatty acid synthesis to a lipoyl carrier protein (LCP) **61** [385, 399]. A thioester-bound acyl-enzyme intermediate is formed in the process [399]. LipA is then responsible for the last step of the production of the lipoyl cofactor in all of the lipoyl-accepting domains, inserting two sulfur atoms at carbons 6 and 8 of an n-octanoyl chain appended to a specific lysine residue of a LCP **60** [58, 382]. The insertion of sulfur into these non-activated carbons is done in a stepwise manner, with stereo selectivity towards C6 and regio selectivity towards C8 [400]. The observation of a second cysteine-containing motif ($CX_4CX_5C$) led to the discovery of another [4Fe-4S] cluster, which is proposed to function as a self-sacrificing sulfur donor [39, 401].

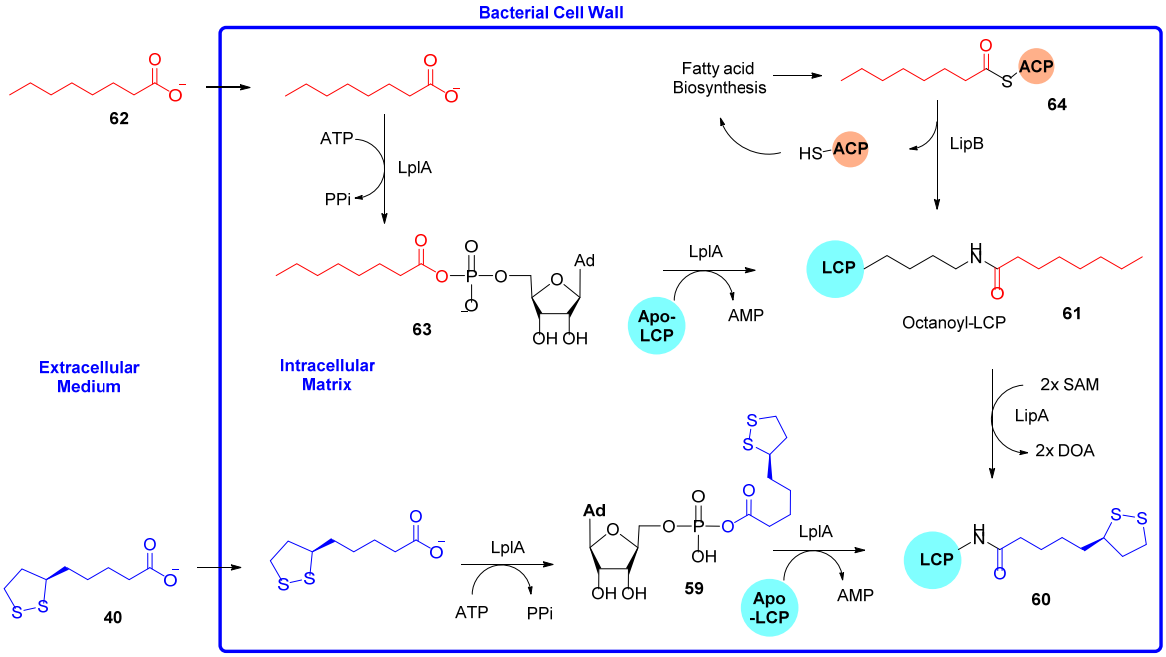



**Figure 23 Pathways of lipoic acid synthesis and uptake.**

The organism may synthesise the lipoyl cofactor by sulfur insertion into octanoyl groups, originating from fatty acid biosynthesis or uptake of exogenous octanoate (octanoyl shown in red). Alternatively, exogenous lipoate can be activated and attached to the LCP (lipoyl shown in blue). The key sulfur insertion step requires lipoyl synthase (LipA). Adapted from Spalding *et al.* [402].

### 1.3.4 Sulfur Insertion Enzymes

#### 1.3.4.1 Biotin Synthase (BioB)

The final step of biotin **67** biosynthesis is the insertion of a sulfur atom between C6 and C9 of dethiobiotin **65** [403]. The enzyme BioB itself was later identified as the responsible for the sulfur insertion [404], but initially the source of the sulfur was unknown [403]. The activity of BioB was shown to be dependent on SAM, even though it was shown not to be the sulfur donor [405]. Results with radiolabelled $^{35}$S-SAM [77, 405] as well as $^{35}$S-cysteine [77] showed none of these compounds were likely to be the source of the sulfur atom.

The necessity of SAM, as well as a reducing agent together with the substrate led to the identification of BioB as part of the RSAM enzymes [406], which was confirmed by identification of a [4Fe-4S] cluster present [59, 407]. Further spectroscopy and electrochemistry studies showed that reconstituted BioB contained both a [4Fe-4S] cluster as well as a distinct [2Fe-2S] cluster, with both cluster being relevant for the activity of the enzyme [128, 408]. Mutagenesis of conserved cysteines that did not coordinate the RSAM [4Fe-4S] cluster showed less iron content and a change in the spectroscopic characteristics of the enzyme [104, 105]. Following the [2Fe-2S] cluster Mössbauer signal showed the signal was stable in the absence of substrate, but decreased during enzymatic turnover, indicating the reaction played a role in cluster degradation [409].

A particularly ambitious study obtained almost one full turnover in the absence of exogenous iron, sulfide and dithiothreitol (DTT), and UV-spectroscopic data observed was consistent with a [2Fe-2S] cluster degradation [410]. Further evidence for the mechanism that involves a self-sacrificial role of the [2Fe-2S] auxiliary cluster was reported in a BioB sample reconstituted with selenide. By reconstituting BioB with selenide, a [2Fe-2Se] cluster was observed in Raman spectroscopy and, under turnover conditions, selenobiotin was generated. If exogenous sulfur was added to the reaction a mixture of biotin and selenobiotin was obtained, again supporting the role of the auxiliary cluster as the source of sulfur (and selenide) [411].

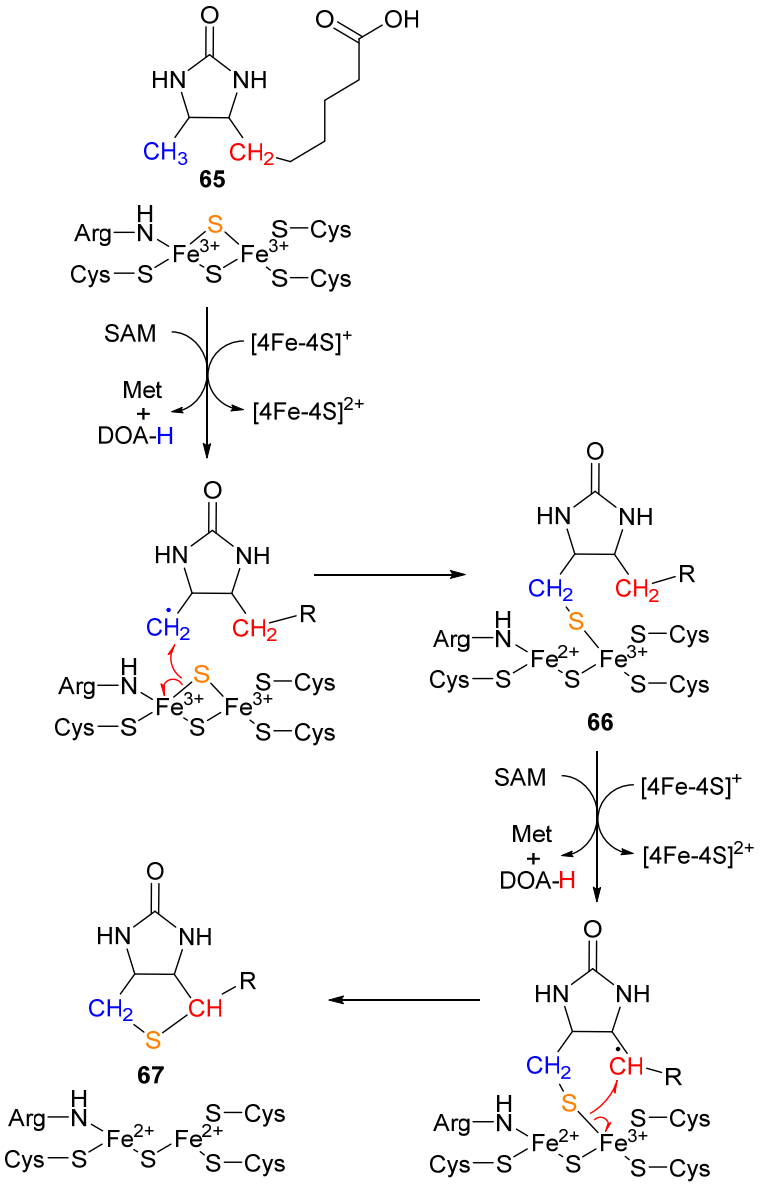



**Figure 24 Proposed mechanism of formation of the thiophene ring in biotin by BioB.**

Two SAM cleavage events abstract protons from C9 and C6 of dethiobiotin, respectively. After the first abstraction, the C9-based radical reduces the $[2Fe\text{-}2S]^{2+}$ cluster, leading to a C9-S bond **66**. The second SAM mediated hydrogen abstraction, in position C6 of the intermediate leads to the formation of a second C-S bond, with sulfur abstraction from the cluster and release of biotin.

### 1.3.4.2 Lipoyl Synthase (LipA) and its Mechanism of Action

Lipoyl synthase (LipA) and biotin synthase (BioB) share a high degree of sequence similarity [396], as well as mechanistic parity, as both catalyse sulfur insertion reactions [39, 412] and require two molecules of SAM per molecule of product produced [413, 414].

Unlike BioB, which possesses a [2Fe-2S] auxiliary cluster [409, 410], consistent with its role as a single sulfur donor, studies on LipA showed higher Fe content which, together with EPR and Mössbauer spectroscopic data, indicated the presence of an additional [4Fe-4S] cluster instead [415]. This difference in the auxiliary cluster was proposed to reflect the necessity of LipA to donate two sulfur atoms for the conversion of the octanoyl group **62** to the lipoyl group **39** [58].

Analysis of a LipA reaction that included a mixture of LipA enzymes reconstituted either with natural $^{32}S$ or isotope labelled $^{34}S$ showed that the product predominantly possessed either fully natural $^{32}S$ or fully isotope labelled $^{34}S$. This seemed to indicate that both sulfur atoms in the synthesized lipoyl groups are derived from the same enzyme molecule [401]. A later study led to the proposal that the sulfur insertions occurred in a sequential manner, as the monothiolated intermediate **70** was successfully isolated, and by use of a tri-deuterated substrate, the difference in mass led to the observation that selectivity for initial C6 abstraction is significant [400].

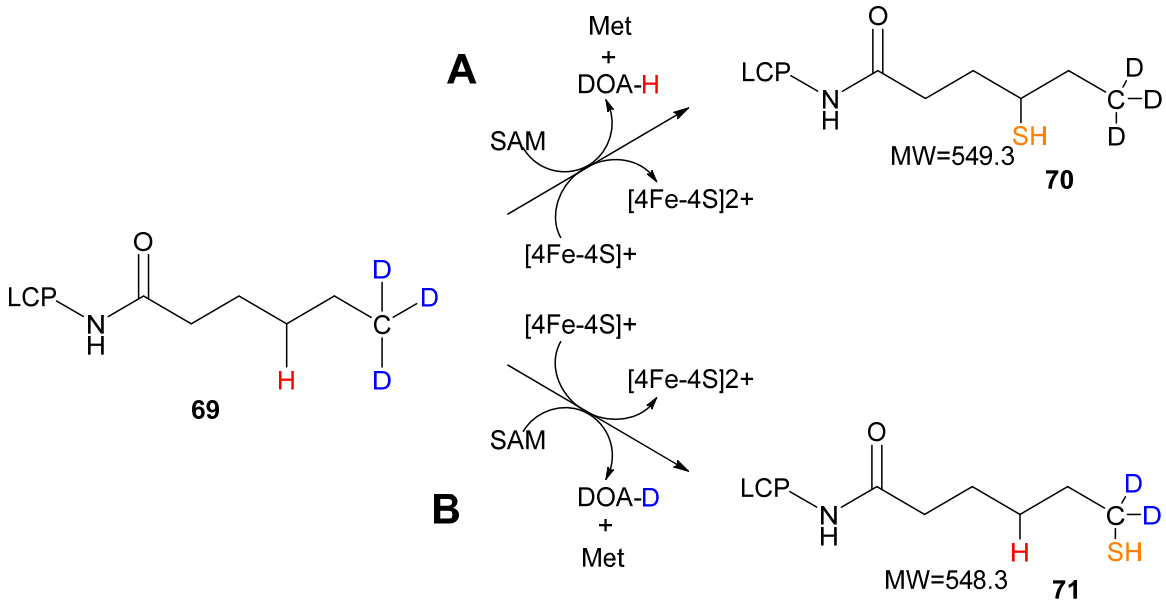



**Figure 25 Formation of a deuterium-labelled monothiolated intermediate by LipA.**

If the radical abstraction on the octanoyl substrate in LipA first occurred in C6 (Pathway A), the thiolation would give rise to a molecular weight that would account for three deuterium present at C8 **70**. Conversely, If the initial abstraction were to occur in C8 (Pathway B), one deuterium would be abstracted by DOA$^{\bullet}$, leading to a molecular weight one unit lower **71** than the product in pathway A. Adapted from Douglas *et al.* [400].

LipA and H-protein have been co-purified as a complex by anion exchange chromatography, due to the presence of a covalent cross-link between a monothiolated H-protein **68** and the auxiliary cluster of LipA [416]. Mössbauer characterization of the cross-linked species did not show the presence two $[4Fe\text{-}4S]^{2+}$ cluster signals, as had been observed in free monomeric LipA. Instead, a presence of one paramagnetic $[3Fe\text{-}3S]^{+}$ cluster was observed, and further showed the formation of this signal was kinetically consistent with the formation of the 6-thiooctanoyl-peptide [416]. Moreover, as the 6-thiooctanoyl-peptide turned over to a lipoyl peptide (i.e. after the second sulfur insertion), there was concomitant loss of the paramagnetic $[3Fe\text{-}3S]^{+}$ signal, with the release of free $Fe^{2+}$ [416]. Additionally, a more recent study directly probed the interactions of a substrate analogue radical (2,4-hexadienoyl-H protein) and a $^{57}Fe$ labelled $[4Fe\text{-}4S]_{Aux}$ cluster, revealing the substrate analogue is located within bonding distance of the auxiliary cluster [417].

These results allowed for the proposal of the mechanism for LipA shown in Figure 26. This enzyme uses two SAM molecules to sequentially abstract a hydrogen atom from two different non-activated carbons, C6 and C8, from its substrate [414]. Initially the first 5'-dA$^{\bullet}$ abstracts a hydrogen atom from C6 of the octanoyl group. The resulting C6 secondary radical is then quenched by reacting with one of the sulfide ions of the second [4Fe-4S] cluster.

During this step, the auxiliary cluster becomes unstable and one $Fe^{2+}$ ion is released, forming the $[3Fe\text{-}3S]^{+}$ cluster observed by Mössbauer characterisation. A second SAM molecule is able to bind to the RSAM cluster, and another 5’-dA$^{\bullet}$ is formed, which abstracts a hydrogen atom from C8 of the octanoyl group. The resulting C8 primary radical is quenched by reacting with an adjacent sulfide of the auxiliary cluster. The addition of two protons facilitates dissociation of the product, a reduced form of the lipoyl cofactor. This is presumed to coincide with the release of some components of the sulfur donor cluster, which, as a result, requires rebuilding for a subsequent reaction cycle [39, 58, 416].

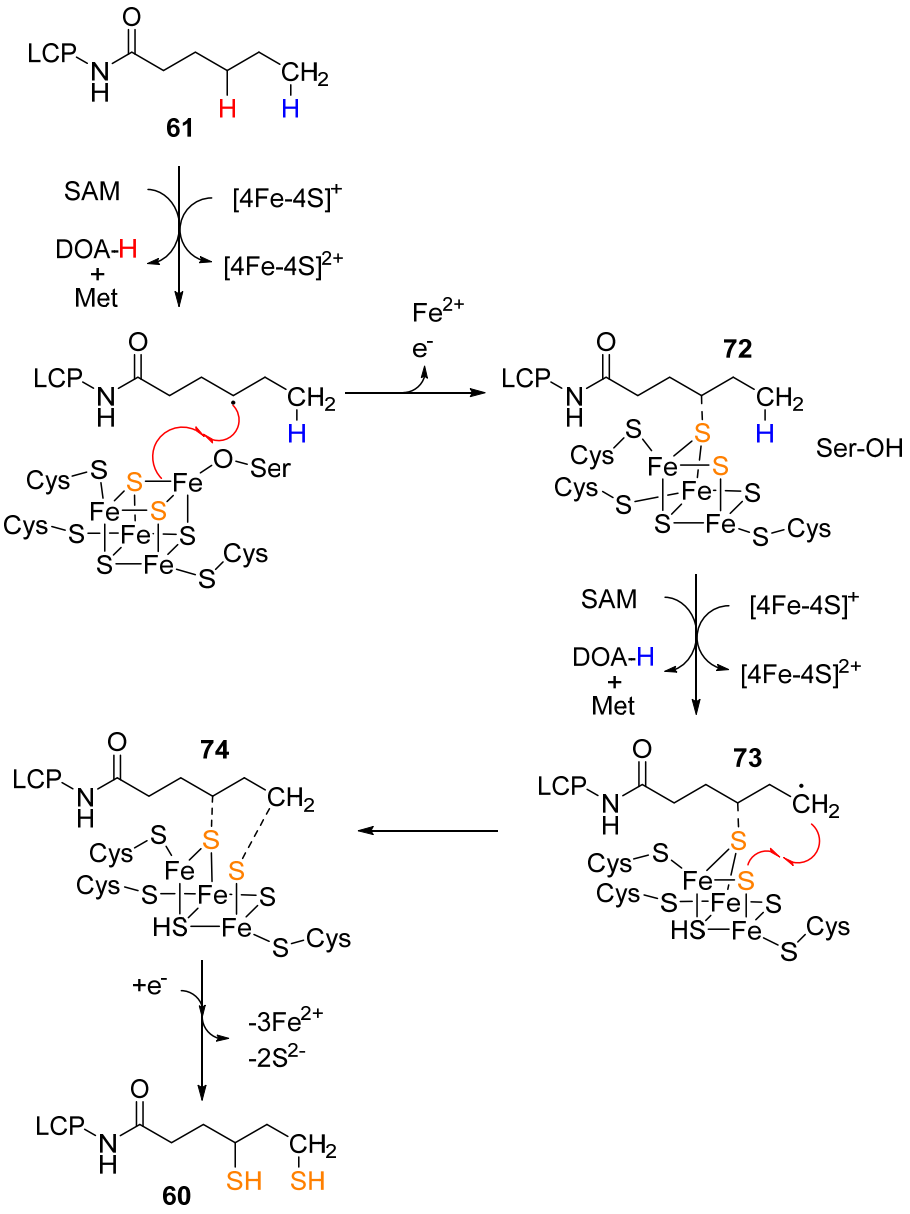



**Figure 26 Proposed mechanism of action by lipoyl synthase (LipA) for lipoic acid synthesis.**

SAM cleavage yields the highly reactive DOA$^{\bullet}$, which abstracts an hydrogen from C6 of the octanoyl substrate. An electron transfer from the cluster yields the formation of a covalent bound monothiolated intermediate. A second DOA$^{\bullet}$ catalysed H abstraction, in C8, followed by an additional electron provided by the auxiliary cluster, yields the two sulfur insertions. Adapted from Booker *et al.* [39].

## 1.4 Antibiotic Resistance Conferring Methyltransferase Cfr

The Gram-positive bacteria *Staphylococcus aureus* (*S. aureus*) was first discovered in the 1880s and is considered to be a major pathogen [418, 419]. The appearance of a methicillin-resistant *S. aureus* (MRSA) [420] as well as a vancomycin-resistant *S. aureus* (VRSA) strains, has led to a widespread concern over drug-resistant pathogens [421-423].

At the turn of the 21st century, the pharmaceutical company Pharmacia released a new antibiotic, linezolid, in an effort to combat drug resistance. Linezolid is part of the oxazolidinone family of antibiotics, the only class of antimicrobial agent with a novel structure to be made available in over two decades [424, 425]. Linezolid (Figure 27 **75**) includes three rings, with an acetamidomethyl tail attached to the oxazolidinone ring.

The ribosome is an important target for antimicrobial therapies due to the complexity of its structure and function [426-428]. Linezolid is one such antimicrobial that targets the ribosome to disrupt protein synthesis leading to bacterial cell death (see Figure 27A for some common ribosome interacting antibiotics) [429-431].

In bacteria the ribosome is divided into the large 50S subunit and the small 30S subunit [432, 433]. The large subunit is composed of 23S rRNA, 5S rRNA and approximately 30 proteins, while the 30S is further divided into 16S ribosomal RNA (rRNA) and approximately 20 proteins [426, 433, 434].

The small subunit is responsible for decoding the messenger RNA (mRNA) and ensuring the correct binding of the transfer RNA (tRNA), which carries the amino acids necessary for protein assembly. The covalent linking of amino acids via peptide bonds occurs in the peptidyl transferase centre (PTC) [435, 436], located in the central loop of domain V of 23S RNA [437] in the 50S subunit. The PTC facilitates the nucleophilic attack [436, 438] of an aminoacyl-tRNA tRNA on the carbonyl carbon of the nascent peptide, binding to it and effectively extending the peptide by one amino acid [437, 439].The newly synthesized polypeptide is then able to leave the large subunit through a channel called the exit tunnel [434, 440].

Several crystal structures show that the binding of different classes of antibiotics, such as linezolid **75**, chloramphenicol **76**, lincosamides and macrolides, occurs in the 50S subunit [431, 441-443]. The structure with linezolid bound, for example, shows the antibiotic binding to the 23S rRNA, with the distal ring pointing towards the PTC and the oxazolidinone ring and acetamidomethyl tail pointing towards the nascent peptide exit tunnel. The drug overlaps with the position of the aminoacyl moiety of a A-site bound tRNA, leading to a prevention of the translation [431].

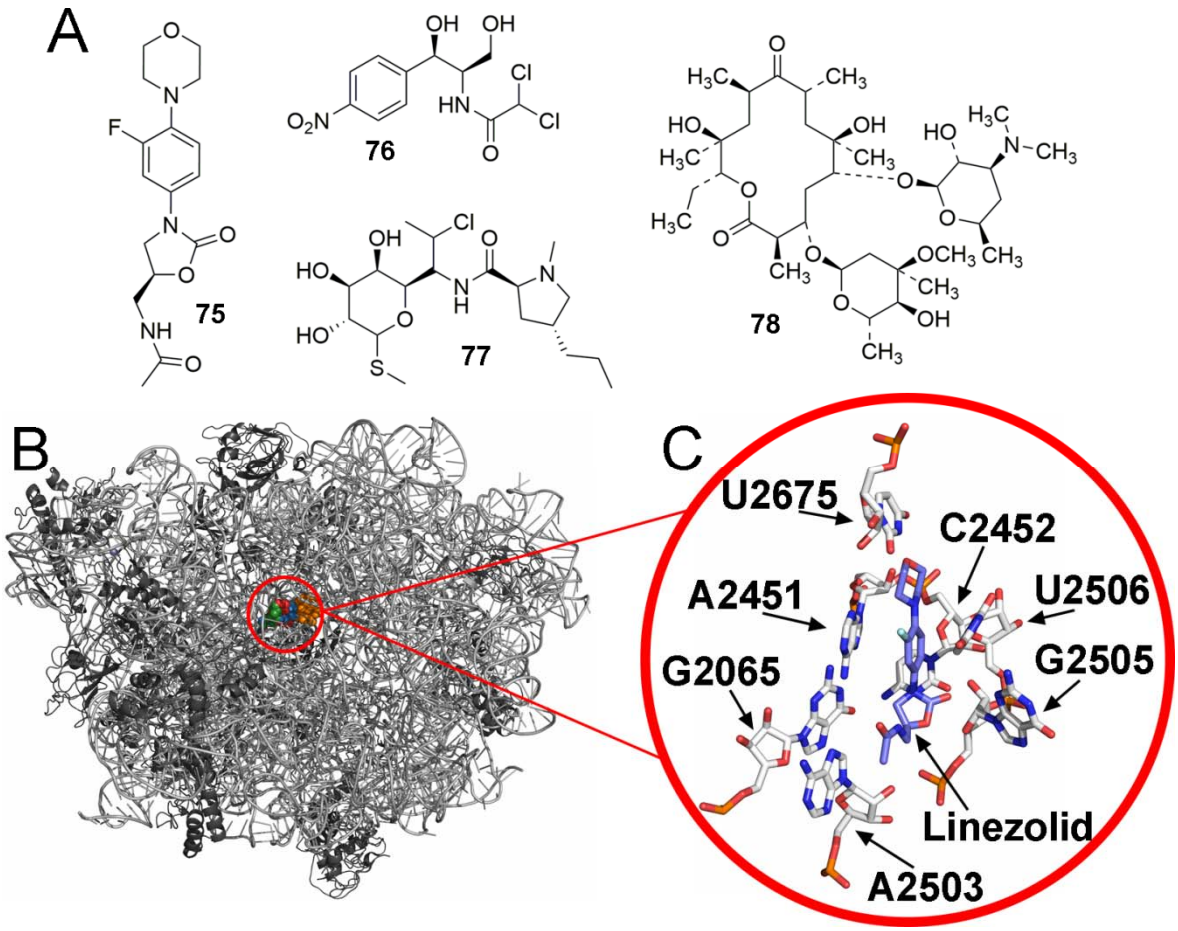



**Figure 27 Antibiotics and their interaction with the 50S bacterial ribosomal subunit.**

**A:** Common antibiotic structures: Linezolid **75**; Chloramphenicol **76**; Clindamycin **77**, of the lincosamide class of antibiotics; Erythromycin **78**, of the macrolide class of antibiotics. **B:** Overall view of the *D. sulfuricans* 50S ribosomal unit. The nucleotide structures are shown in white, amino acids in grey. Linezolid is shown in red, chloramphenicol in green, clindamycin in blue and erythromycin in orange. **C:** View of linezolid (lilac) within the binding pocket formed by eight universally conserved nucleotides (white, *E.coli* numbering) of the 23S rRNA, including A2503. Adapted from Wilson *et al.* and Brown *et al.* [431, 441].

There have been reports of mutations to the nucleotides of the 50S ribosomal unit that lead to increased antibiotic resistance, presumably by interfering with the drug's binding site [430, 444-447]. However, as bacteria possess multiple copies of the ribosomal RNA (*rrn*) alleles, a fortuitous mutation of one copy alone has a modest effect in antibiotic resistance [446, 448]. The accumulation of an antibiotic resistant mutation in additional alleles can lead to an increased resistance which is known as the gene dosage effect [448-450]. These same mutations could however negatively affect ribosome function and, in the absence of antibiotic selection present in

the environment, the mutants could suffer selective pressure to reverse the mutation [451, 452]. Therefore, a more common resistance mechanism is based on the chemical modification of rRNA by enzymes.

### 1.4.1 Cfr and Antibiotic Multiresistance

In 2000, Schwartz and colleagues discovered a multi-resistance plasmid in *Staphylococcus sciuri* isolates of animal origin [453], identifying the *cfr* gene as responsible for the antibiotic resistance. This gene was initially shown to give resistance to chloramphenicol and florfenicol, and was thus named *cfr* for *c*hloramphenicol florfenicol resistance. Furthermore, the gene has been identified in different organisms [454], including Gram-negative bacteria [455, 456], making the understanding of the mechanism behind *cfr*-mediated antibiotic resistance crucial in developing effective treatments for multi-resistant bacterial infections.

It was shown the *cfr* gene encodes a methyltransferase (Cfr) that methylates adenosine 2503 (A2503) in 23S rRNA [457, 458]. A2503 is located at the entrance of the nascent peptide exit tunnel of the PTC [459, 460] and, due to the functional importance of the site, several classes of antibiotics have been developed to bind to that area (see Figure 27). It was unsurprising then that the modification catalysed by Cfr was shown to prevent several classes of antibiotics from binding to the PTC [457, 460-462] including linezolid [463]. The presence of the methyl group is proposed to either cause a shift in the position of A2503 decreasing drug binding [462], or directly obstruct the binding of the drugs [458].

The antibiotic resistance gene *cfr* shows a significant degree of similarity (32% sequence identity) to the endogenous bacterial gene *rlmN* [83, 458, 464]. RlmN is a methyltransferase that catalyses methylation of A2503 at position C2 of the adenine base [465, 466], as well as the methylation of A37 in tRNAs [467]. Unlike Cfr, however, RlmN's modification seems to be an endogenous post-translational modification to increase translation fidelity and does not confer any relevant antibiotic resistance [83, 466, 468].

Amongst the naturally occurring RNA modifications, methylation of the nucleoside constitutes about 2/3 of all changes [469, 470]. These methylations are formed by RNA methyltransferases (MTases), belonging to four unrelated superfamilies [471, 472]. MTases from the Rossmann-fold superfamily [473], the SPOUT superfamily [474] and the RSAM superfamily [307] use SAM as a methyl donor; whereas enzymes from the FAD/NAD(P)-binding protein superfamily use 5,10-methylenetetrahydrofolate (5,10-$CH_2$-THF) as the methyl donor [475].

Both Cfr and RlmN belong to the RSAM family of enzymes [476] and are part of the methylase group of RSAMs, which are further divided into three classes [307]: Class A, to which RlmN and Cfr belong, contain only the canonical RSAM domain (see section 1.1.2), together with two additional conserved cysteines essential for catalysis [31, 477, 478] and methylate $sp^2$-hybridized carbon atoms [31, 478]. Class B enzymes are thought to possess both a RSAM domain as well as a cobalamin (B12) binding domain [479, 480], being able to methylate a highly diverse set of substrates, including inert $sp^3$ carbons [481-483], aromatic heterocycles [114] and phosphinates [484]. Lastly, class C contains enzymes with a separate C-terminal domain, similar to the HemN domain, which are thought to assist in substrate binding [75]. Similar to class A MTases, the enzymes in class C catalyse the methylation of $sp^2$-hybridized carbon centres [485, 486], but do not contain the two conserved cysteines required for methylation via a Class A mechanism [307].

Recently a new type of MTase has been discovered. It contains the RSAM cysteine motif, but uses 5,10-$CH_2$-THF as the methyl donor [487], bridging the RSAM and the FAD/NAD(P)-binding MTases and potentially giving rise to a new class (D) of RSAM methyltransferases [488].

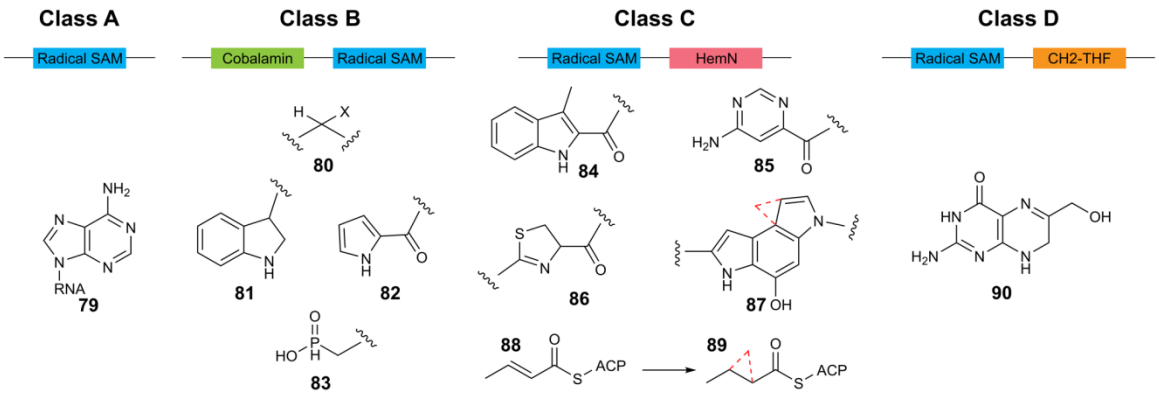



**Figure 28 Classes of RSAM methyltransferases and methylation targets.**

**Class A**: **79**, C2/C8 methylation of rRNA [476] or tRNA [467]. **Class B:** Methylation of *sp3*-hybridized carbon centres **80** [481, 483]; $sp^2$-hybrdized carbon methylation in aromatic heterocycles: indole rings **81** and pyrrole rings **82** [114, 489]; **83** Phosphinate methylation [484]. **Class C: 84** 3-methyl-2-indolic acid (MIA) methylation [485, 490]; **85** Thiazole heterocycle methylation [486]; **86**, Pyrimidine ring methylation [491, 492] ; Cyclopropane ring formation in yatakemycin **87** [493] and Jaw5 cyclopropane formation **89** [494]; Pterin **90** methylation [488]. Adapted from *Zhang et al.* [495].

### 1.4.2 Mechanism of Action of the RSAM Methyltransferases RlmN and Cfr

The stoichiometry of the reactions catalysed by Cfr/RlmN indicated the use of two molecules of SAM per methylation [83]. The first molecule of SAM is proposed to be the source of DOA$^{\bullet}$ **3** while the second provides the methyl group , releasing S-adenosyl-L-homocysteine (SAH) **91** as a by-product of the reaction [83]. As Cfr/RlmN appears to possess only one [4Fe-4S] cluster [476, 496], this cluster was proposed to have dual functionality, acting as both the source of DOA$^{\bullet}$ and as a methylating agent [476].

Until recently, methyl transfer reactions using SAM as a co-substrate were reported to happen solely via $SN_2$ displacement [497], where SAM acts as an electrophilic donor of methyl groups to more nucleophilic O or N atoms [498].The target carbons for Cfr/RlmN were, however, shown to be electrophilic [499]. Additionally, the direct abstraction from an $sp^2$ hybridized carbon would require high activation energy [83], also indicative of an alternative mechanism of action.

The first evidence for the unique mechanism of action of Class A MTases was obtained by deuterium labelling of a RNA fragment at A2503 C2 position. The results showed no formation of a DOA-$^2$H product, indicating that the direct abstraction of the product did not occur [500]. When the same authors used deuterium labelled SAM ([*methyl*-$^2H_3$]SAM, they observed incorporation of two deuterium atoms into the RNA fragment and one final deuterium present in DOA-$^2$H [500]. These results seemed to suggest that one SAM molecule is reductively cleaved to form DOA$^{\bullet}$, as part of the regular RSAM mechanism (see section 1.1.1), which would abstract a hydrogen atom from a second molecule of SAM.

The uniqueness of these steps led to the proposal of a mechanism of action in which the abstraction of an hydrogen from the second SAM, would create a methylene radical **92** which would attack a hydrogen in the target position of the nucleotide, followed by a hydride shift caused by the aromatization of the product **93** and concomitant expulsion of SAH [500].

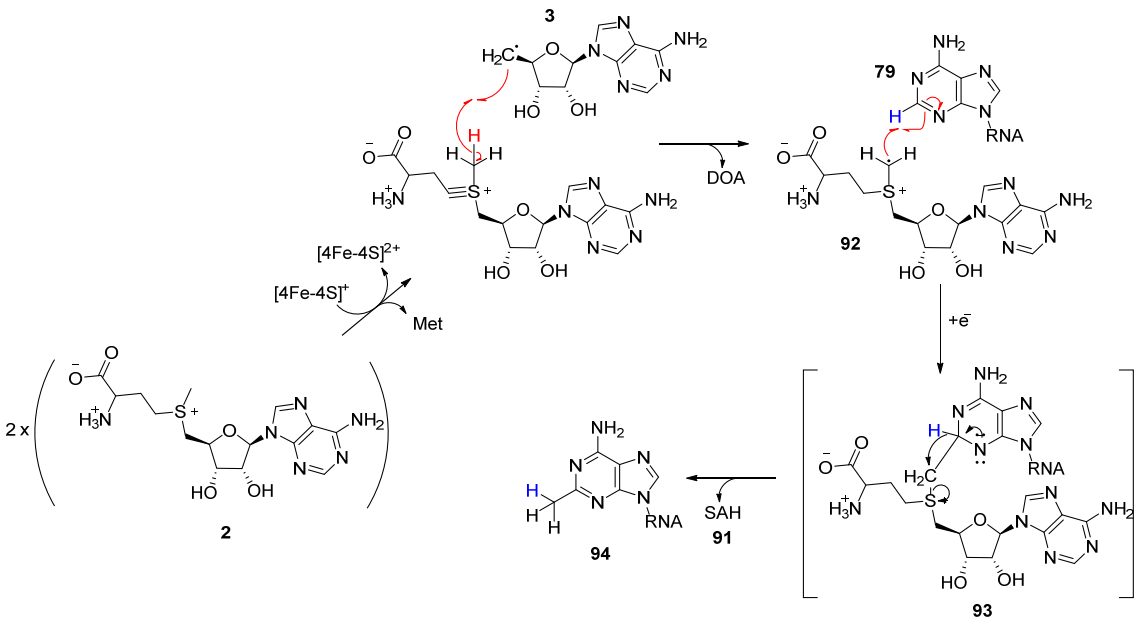



**Figure 29 Proposed mechanism of Cfr methylation reaction, with aromatization and hydride shift.**

In this mechanism two molecules of SAM bind to the active site, with one yielding the mechanistically conserved DOA$^{\bullet}$ which abstract a hydrogen from the methyl group of the second SAM molecule. The methyl based radical **92** formed can then attack the substrate **79,** adding to the amidine carbon of the substrate. One electron reduction releases the product and SAH. Adapted from Yan and Fujimori [500].

This mechanism came into question, however, when new SAM-labelling studies failed to replicate the direct transfer of a methyl group onto the substrate [477]. The same group was able to show, by careful purification of fully reconstituted RlmN and Cfr, that the first turnover reaction did not require exogenous SAM as a source of the methyl group [477]. Instead the DOA$^{\bullet}$ was shown to abstract a hydrogen atom from a methylated cysteine of the protein.

The importance of the residue for activity was confirmed by enzyme mutagenesis studies in RlmN [31], which managed to trap a covalent intermediate of the reaction by mutating one of the conserved cysteines. Additionally the structure of RlmN [496] was obtained, which strongly suggested that the active site could not bind two SAM molecules and the RNA substrate simultaneously, together with structural evidence of a methyl group attached to one of the conserved cysteines (Cys355) [496].

In both Cfr and RlmN, it was shown that SAM can bind with high affinity to the enzyme in the absence of the RNA substrate or the methylated cysteine [501]. However, only in the presence of

a methylated cysteine is the [4Fe-4S] cluster re-oxidised with formation of DOA$^{\bullet}$ radical [501]. A new mechanism of action was proposed (Figure 30), in accordance with the mounting evidence for a methylcysteine radical-based mechanism [31, 477].

It was proposed the methyl group of one molecule of SAM is transferred to one of the conserved cysteines via a polar $S_N2$ mechanism, resulting in formation of an *S*-methylcysteine 89 (mCys338 in Cfr) residue and SAH **91** [31]. This reaction is followed by the homolytical cleavage of the second SAM **2** molecule, with DOA$^{\bullet}$ **3** radical generation, which will abstract a hydrogen atom from the methylated cysteine [477]. The newly formed thiomethylene radical **96** attacks the C2 or C8 amidine carbons, and after one-electron oxidation and loss of a proton a transient rRNA–protein adduct is formed **97**, which was observed by EPR studies [502]. The covalent adduct is resolved by attack of the second cysteine (Cys105 in Cfr) to form a disulfide bridge and release the methylated rRNA **94** [31, 477, 503] (Figure 30).

While the steps described above have become widely accepted, the precise mechanism of product dissociation is not well understood as of yet, and the second cysteine could have a more significant role than to simply create the persulfide bridge (Figure 30, black), by stabilizing a thiyl radical (Figure 30; green), abstract the target proton (Figure 30, yellow) or both (Figure 30, purple) [478].

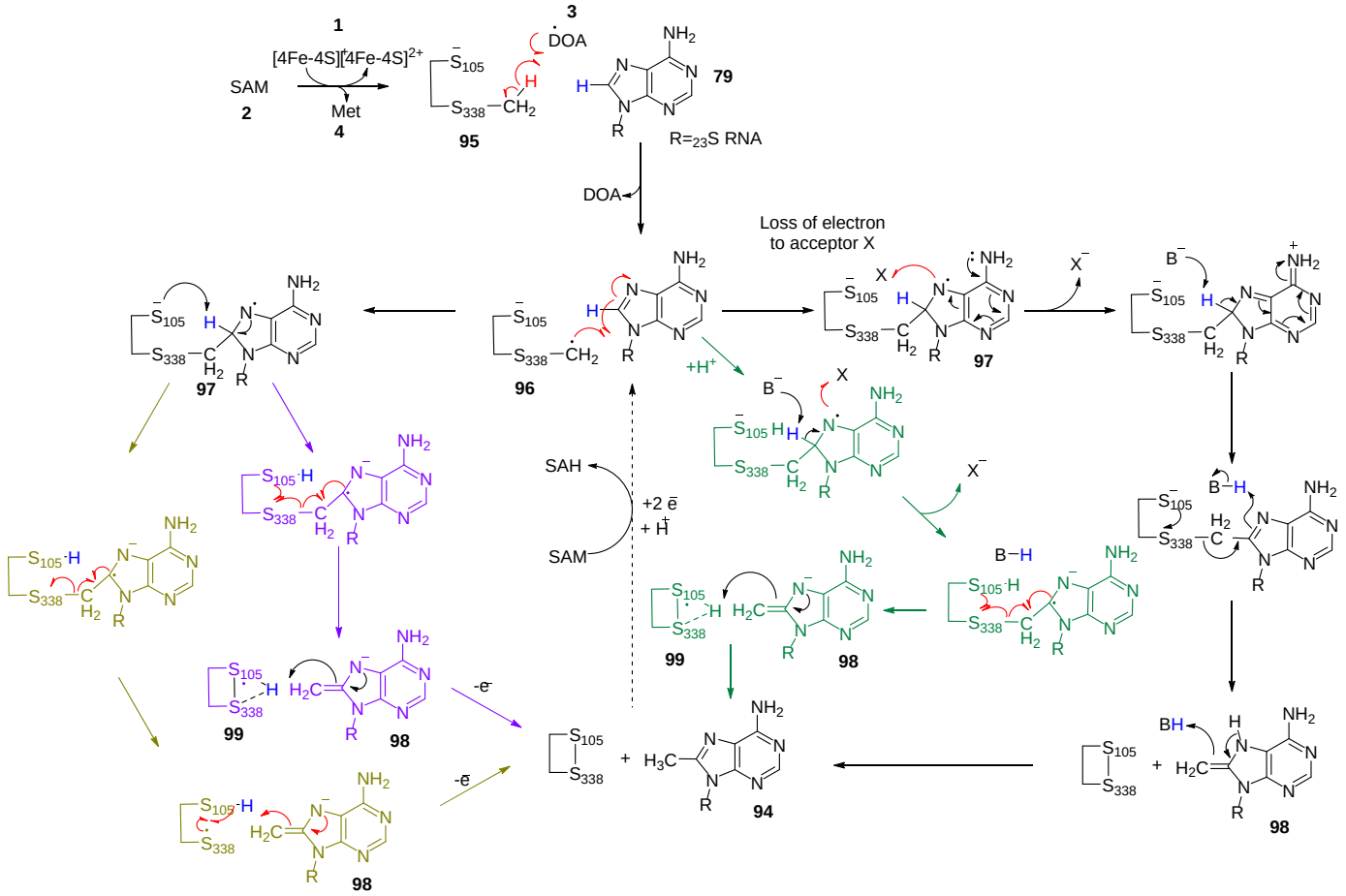



**Figure 30 Proposed cysteinyl radical based methylation of RNA by Cfr, with different potential mechanisms for adduct resolution.**

In black, reduction of the adduct, with Cys105 forming a disulfide bond. In green, Cys105 stabilizes the thiyl radical formed **99**. In yellow Cys105 is responsible for the deprotonation of the product. In purple Cys105 both abstracts the proton and forms the disulphide radical species **99**. BH and $B^-$ are general acid and base residues in the active site. X is a general electron acceptor in the active site. Adapted from McCusker *et al.* and Silakov *et al.* [31, 478].

### 1.4.3 Structure of RlmN and Cfr

As discussed previously, a crystal structure of RlmN was recently obtained [496], showing the common structural elements of the RSAM superfamily (see section 1.1.2): its overall fold forms a partial TIM barrel, and the cysteine motif containing loop, which coordinates the [4Fe-4S] cluster, can be observed at the carboxyl edge of the barrel.

Apart from these elements, RlmN possesses in the three additional β-strands of the N-terminus that extend the barrel laterally (Figure 31). The final 60 residues in the N-terminal form four short α-helices, and are a variant of the HhH2 fold, a sub-structure that mediates substrate recognition in proteins that target nucleic acids [504]. The C-terminus possesses an extension, comprising residues 351 to 360, where the mechanistically relevant Cys355 [31, 477] is located. This extension is disordered in the absence of SAM, however when SAM is bound, a small loop is formed that places the cysteine in the active site, at 3.5 Å from the methionine moiety of SAM and in a favourable position for $SN_2$ displacement [496].

Due to their sequence and mechanistic similarities, it was proposed that Cfr has evolved from RlmN [464]. In fact, Cfr appears to have retained the ability to methylate C2 [83], although at lower efficiency levels, suggesting a flexible binding pocket capable of fitting two different conformations of A2503 [477].

Modelling of Cfr, using RlmN as a template showed conservation of the catalytic residues and most surrounding residues of the core fold. Several key differences could be observed, with a lack of an extended loop between α1 and β2 and the lack of the c-terminal helix, connected to the critical 351-360 linker [496].

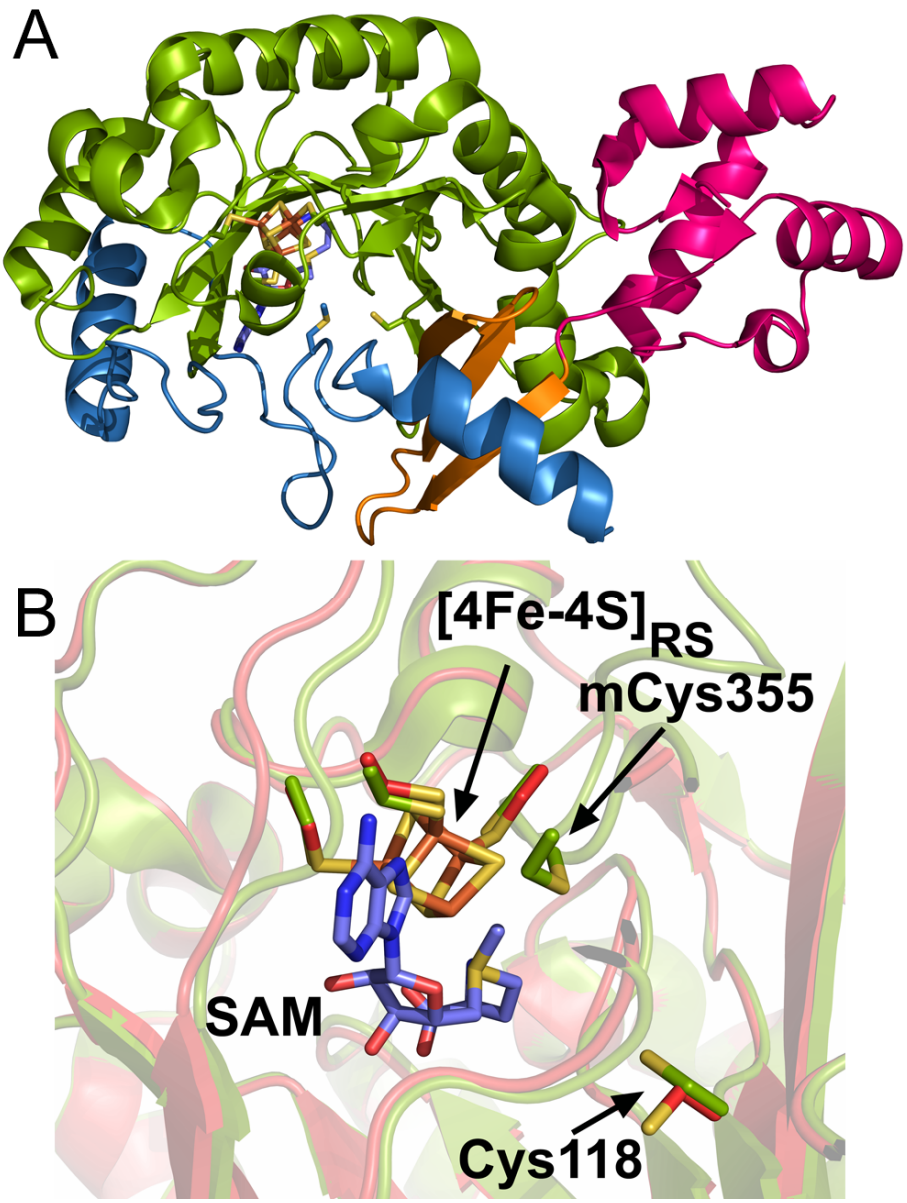



**Figure 31 Structure of RlmN and modelled structure of Cfr.**

**A:** Structure of RlmN(PDB ID: 3RFA) in the presence of SAM. Pink – N terminal domain; Orange – β' 1-3 extension; Green– Partial TIM barrel β6/α6 core fold; Blue – C terminal domain. The iron sulfur cluster, in yellow and orange, and the activity relevant cysteines are shown as stick models. **B:** Comparison of the RlmN structure in the presence (green) and absence (red) of SAM. When SAM is bound to the [4Fe-4S] cluster a loop can be seen on top of the active site, where the mCys355 can be observed, which is disordered in the absence of the co-substrate. Both RlmN structures have, as expected, a high degree of structural similarity, with an average distance between the main carbon atoms (RMSD) of only 1.08 Å. Adapted from Boal *et al.* [496].

# Chapter 2: Structural Characterization of HydG by X-Ray Crystallography

## 2.1 Introduction

Mechanistic studies into the role of the maturase HydG, discussed at length in section 1.2.2.3, led to the identification of the substrate L-tyrosine [302] as well as the diatomic products carbon monoxide [293] and cyanide[47]. These results cemented the role of HydG as a provider of diatomic ligands [309, 310, 505, 506] during the maturation of [FeFe]-hydrogenase HydA [184, 507, 508].

In the absence of structural data for HydG, the currently accepted mechanism of hydrogen abstraction derives from the data obtained in the structure of the related enzyme NosL, an amino acid lyase that also catalyses $C_\alpha$-$C_\beta$ bond cleavage from tryptophan. The structure showed tryptophan in the active site, in a position where the hydrogen abstraction is likely to occur in the α-amino hydrogen atom [321] (see section 1.2.2.3.2, Figure 17).

A recent EPR study with HydG, despite not detecting a tyrosyl radical nor a glycyl radical, did observe the tyrosine derived 4-oxidobenzyl radical (4OB$^\bullet$ 24) (section 1.2.2.3.2, Figure 16), consistent with heterolytic cleavage of $C_\alpha$-$C_\beta$ bond of tyrosine following amino hydrogen abstraction, to form dehydroglycine (DHG).

Meanwhile, spectroscopic evidence supported the hypothesis that HydG possessed not only the RSAM [4Fe-4S] cluster but also an additional auxiliary [4Fe-4S] cluster [80, 293, 310], which mutagenesis studies showed was required for enzymatic activity [118]. The auxiliary cluster is a site differentiated group that does not seem to bind SAM [312], nor be directly related with tyrosine cleavage [311]. Bearing those results in mind, the role of the secondary cluster is likely linked to the second part of the reaction, which yields CO/$^-$CN from the easily hydrolysed DHG or a close equivalent. This idea was strongly supported by an FTIR study which showed the formation of a $Fe(CO)_2CN$ synthon in the auxiliary cluster of HydG [296] (see section 1.2.2.3.2, Figure 19).

A significant gap in the understanding of the enzyme HydG is the lack of an available structure that reveals the relationship between the two clusters and their active sites. A structural homology model was created for the related pair ThiH and HydG using HydE and BioB as model structures [295].

HydG and ThiH possess a high degree of sequence similarity, apart from the 80 amino acid C-terminus extension that contains the $[4Fe\text{-}4S]_{Aux}$ cluster binding $CX_2CX_{22}$ motif, which is absent in ThiH [118]. The comparison between the hypothetical models for ThiH and a mutant HydG lacking the C-terminal end showed both enzyme sequences could be divided into three equivalent regions. A common N-terminal region, responsible for SAM and [4Fe-4S] binding and the site of $DOA^{\bullet}$ radical formation; a central region required for tyrosine binding and cleavage; and C-terminal region that is highly divergent in sequence between the enzymes, and that ultimately determines the fate of the amino acid derived fragment(s) [295]. This outcome presumably includes the transfer of the products into different downstream proteins, presumed to be HydF in the hydrogenase maturation pathway and ThiG in the thiamine biosynthesis pathway, respectively.

The division of HydG into separate regions resonates well with the possible division of the activity of HydG into two halves (see section 1.2.2.2.3, Figure 14). As shown by the mutagenesis activity assays removal of the auxiliary cluster coordination does not interfere significantly with the ThiH/HydG shared tyrosine Cα-Cβ cleavage [311], but it has a direct effect on the outcome of dehydroglycine or related glycyl intermediate

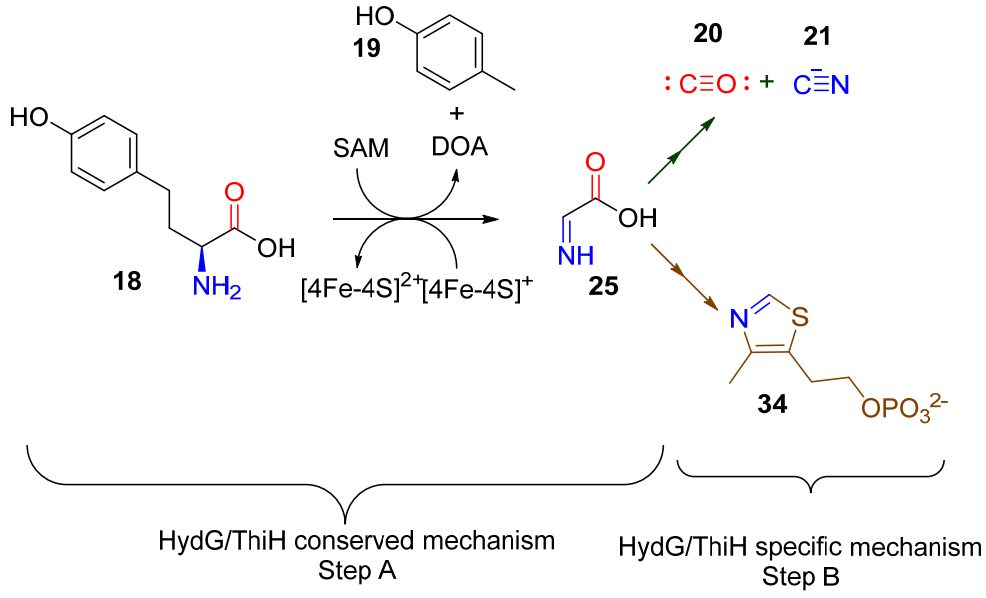



**Figure 32 Two step mechanism of tyrosine cleavage by HydG/ThiH.**

The first half of the reaction, concerning the cleavage of L-tyrosine with the observed release of *p*-cresol and formation of dehydroglycine, is conserved in both ThiH [313] and HydG [302]. The second half of the reaction, however, is specific to each metabolic pathway.

The general information obtained in the homology model, allows the prediction of a structural framework for the different stages of the mechanism of HydG, but does not provide any indication on how each step of the mechanism can occur. In an effort to increase the understanding of this complex enzyme, HydG crystallization studies were carried out.

The main results of this chapter, which include the structural data of *Ti*HydG (PDBID: 4WCX), together with EPR spectroscopic analysis of its clusters was recently published in Dinis *et al.* [509].

## 2.2 Thermophilic HydGs as a Source of Optimised Homogeneous Protein

Expression and purification of sufficient quantities of homogenous protein solution is one of the bottlenecks of protein crystallography, as a reasonable amount of protein is necessary for each crystallization experiment, and many cycles of screening for the appropriate crystallization conditions may be required [510]. In addition, and as a general rule, most natural sequence variation (neutral mutations that do not greatly affect activity) mostly occur in residues unrelated to the active site [511]. These sequence differences can accumulate in surface residues on an evolutionary timescale [512-514].

Furthermore, protein crystallization heavily depends on protein-protein contacts between the surfaces of the different molecules for the formation of the crystal lattice [515]. As such, the expression of proteins with the same catalytic activities but divergent sequences arising in different organisms (and therefore having potential for differences in the surface residues) can be used to increase the possibilities of a successful crystallization experiment [516].

Another key point to bear in mind is that proteins of extremophiles, in particular thermophiles and halophiles, tend to be more stable as they contain more stabilizing salt bridges together with more hydrogen bonds [517]. As a broad generalization, this stability often facilitates purification, leading to higher yields and a more homogeneous protein sample that improves the possibility of a successful crystallization [518].

Early research in the Roach group had made use of HydG from *Clostridium acetobutylicum* [47], an expression system originally developed and kindly donated by Broderick and Peters' group [282]. In an effort to identify a subset of HydGs suitable for crystallization screening, a phylogenetic study was carried out by R. Driesener [519] on *hydG* genes from thermophilic organisms. From the thirteen sequences analysed (Figure 33), three genes were selected from different clades to

maximize the degree of sequence diversity within the HydG family. The selected organisms and their respective sequence identification numbers were: *Thermotoga lettingae* TMO (GENID: Tlet_1748), *Thermoanerobacter italicus* Ab 9 (GENID: Thit_0582) and *Thermotoga maritima* MSB8 (GENID: TM_1267).

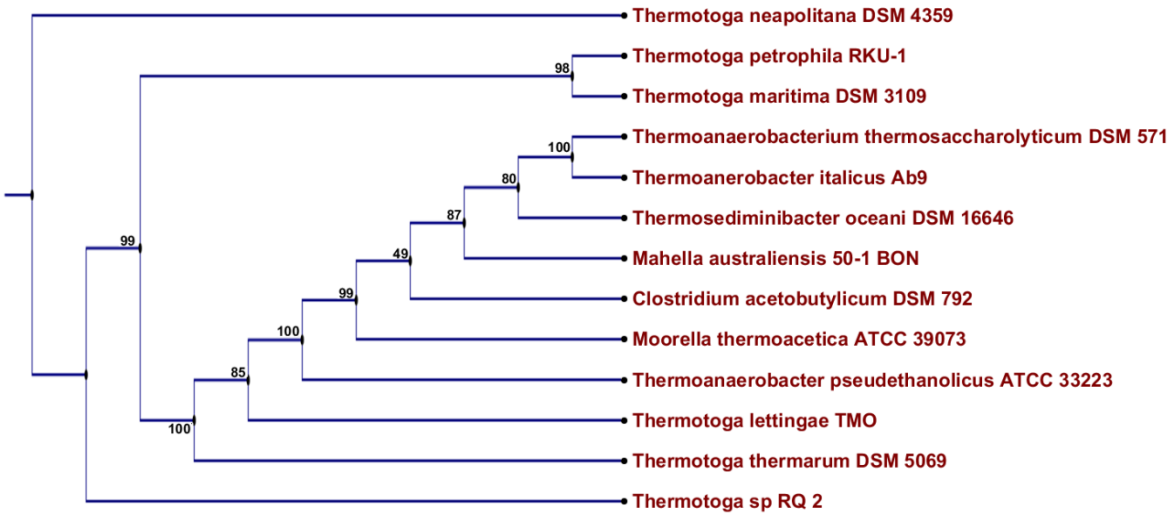



**Figure 33 Phylogenetic tree of different thermophilic organisms, based on HydG primary protein sequences present.**

The software CLC Sequence Viewer 6.9.1 was used to create the tree using the Neighbor Joining algorithm. The confidence values for each node are shown by the bootstrap numerical value of 100 replicates. Adapted from [519].

## 2.3 HydG Protein Expression

### 2.3.1 Plasmid Design and Expression System

Synthetic codon optimised *hydG* genes were acquired commercially (Geneart Inc.) and subcloned by R. Driesener into a pFM024 vector [520] together with an N-terminal hexahistidine-tag to allow affinity purification [521] (Figure 34). The pFM024 vector (Appendix 10.1, Figure A1) is a pBADHis derived vector [522], which allows the co-expression of a gene of choice, inserted between NcoI and XhoI sites, with the *E.coli isc* operon, derived from vector pMK400 [523]. The *isc* operon encodes the proteins responsible for the assembly of iron-sulfur clusters in bacteria. The co-expression of the iron-sulfur cluster assembly proteins together with the RSAM enzymes is thought to improve the [4Fe-4S] content of the latter [523].

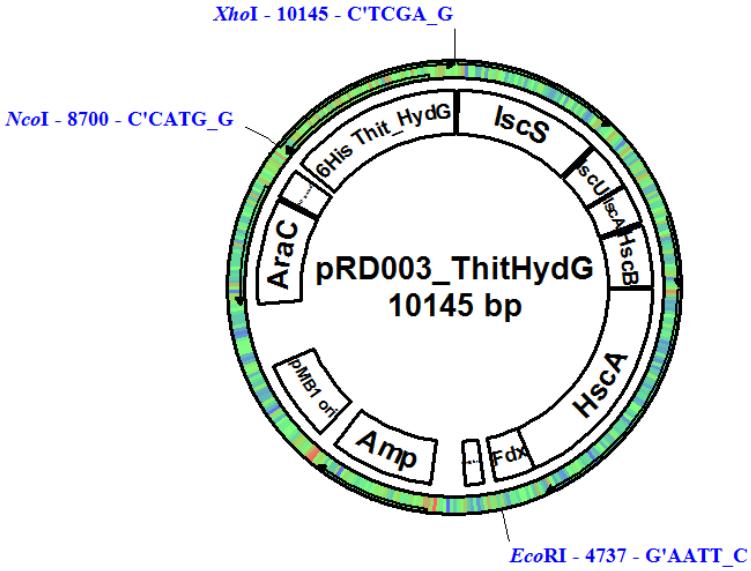



**Figure 34 Co-expression vector (pRD003) for *Thermoanaerobacter italicus* HydG with the *E.coli isc* operon.**

The constructs for the three thermophilic *hydG*s (Appendix A1) had been successfully transformed (Method 9.4.3) into *E. coli* BL21 (DE3) cells by R. Driesener [519], a strain developed for expression of recombinant proteins [524].

### 2.3.2 Recombinant Protein Expression and Purification

The large scale expression of the HydG (typically a 5L culture) from each genetic background yielded similar amounts of a dark-brown cell pellet, between 5 and 8 g/L of cell culture (Method 9.5.1). The culture of HydG from *Thermotoga lettingae* (*Tl*HydG) showed the highest cell density, 40g of cell pellet for each 5L growth, while *Thermotoga maritima* (*Tm*HydG) produced the lowest yield, never higher than 25g per 5L growth. HydG from *Thermoanerobacter italicus* (*Ti*HydG) showed values closer to *Thermotoga lettingae*, with 35 g of cells per 5L growth on average.

The purification method was slightly modified from a previously reported procedure [47] (Method 9.5.3) and is summarized in the scheme below (Figure 35).

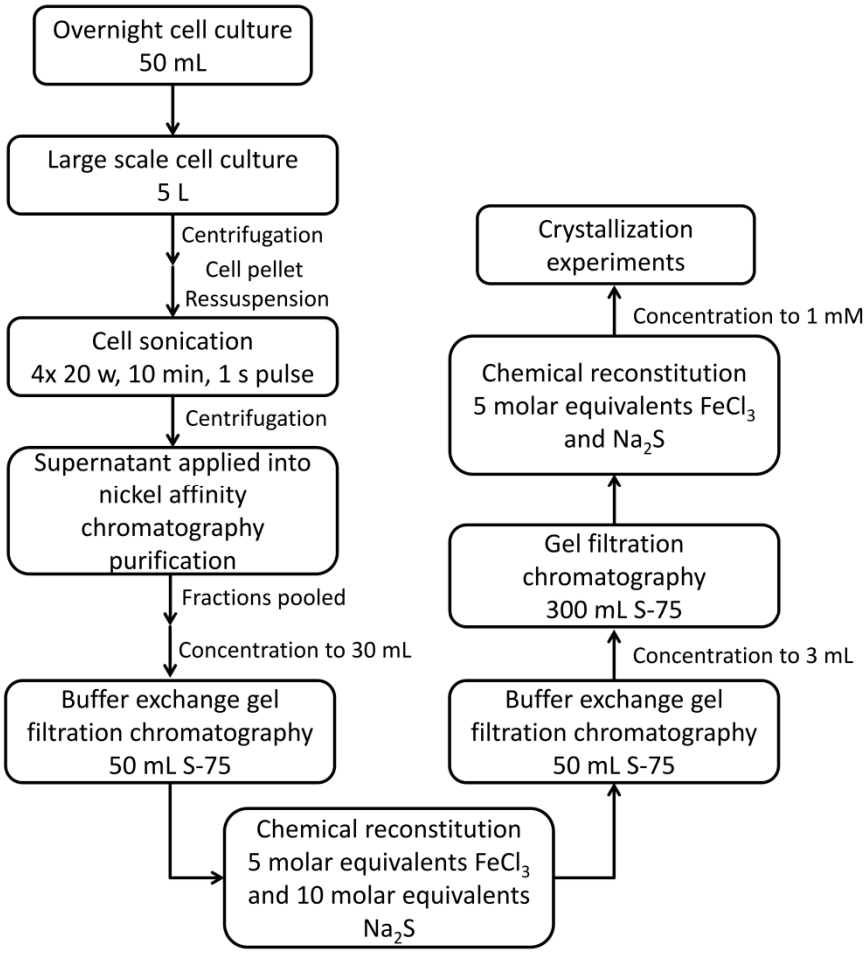



**Figure 35 Schematic representation of the purification steps for RSAM enzymes.**

After growth of a large scale cell culture, the cells are collected by centrifugation and the cell pellet is sonicated to lyse the cells. The cell debris is removed by centrifugation and the cleared lysate is applied to a nickel-sepharose column for purification of the his-tagged proteins. The fractions collected from the nickel affinity purification are pooled and buffer exchanged to a buffer without imidazole and the iron sulfur clusters of the protein are reconstituted. The resultant protein solution is then concentrated and applied to a S-75 sepharose gel filtration column and the purified golden-brown fractions are collected, pooled, reconstituted again and concentrated to ~1 mM. All samples are then stored as aliquots (100 $\mu$L) at -80ºC.

All purification stages were carried out in an anaerobic glove box (Belle Technology, Portesham, UK), kept at $O_2$ < 2 ppm. Following cell lysis a centrifugation step removed the cell debris, leaving the proteins in a clear lysate which is then applied to a nickel charged column. The hexa-histidine tag binds to the metal ions and is retained in the column, while the non-tagged proteins from the bacterial host cell are mostly eluted directly from the column.

The tagged protein can then be released by applying a gradient of imidazole, a ligand that competes with the $(His)_6$-tag for the nickel.

The metal-chelating properties of imidazole [525] usually disrupt the [4Fe-4S] cluster(s) of RSAM enzymes, resulting in at least partial loss of the cluster. Practical experience suggests that loss of the cluster(s) can have a direct impact on protein stability and may lead to its aggregation and precipitation. A chemical reconstitution (Method 9.6.1) is thereby introduced during the purification stages as an attempt to increase protein stability.

Prior to the addition of exogenous iron and sulfur for chemical reconstitution of the samples the imidazole must be removed from the eluted fractions through the prompt use of a small gel filtration column for buffer exchange. The column separates molecules by size, with larger molecules (such as proteins) being excluded from the chromatographic matrix and being eluted first, while small molecules (such as salts and imidazole) taking longer to be eluted [526].

Very often the single purification step of nickel-affinity is insufficient with several cell-derived proteins co-purifying with the target protein. As such, a secondary purification step, based on size exclusion chromatography with a larger column (300 mL, column XK-16, 100 cm), was chosen to improve the purity of the samples [526]. Using this method, an initial dark peak is routinely observed prior to protein elution, corresponding to an aggregated mixture of target protein, bacterial proteins and iron-sulfide [527] which chromatographs as a very large size molecular unit, potentially indicating the formation of colloidal iron sulfide and can be discarded.

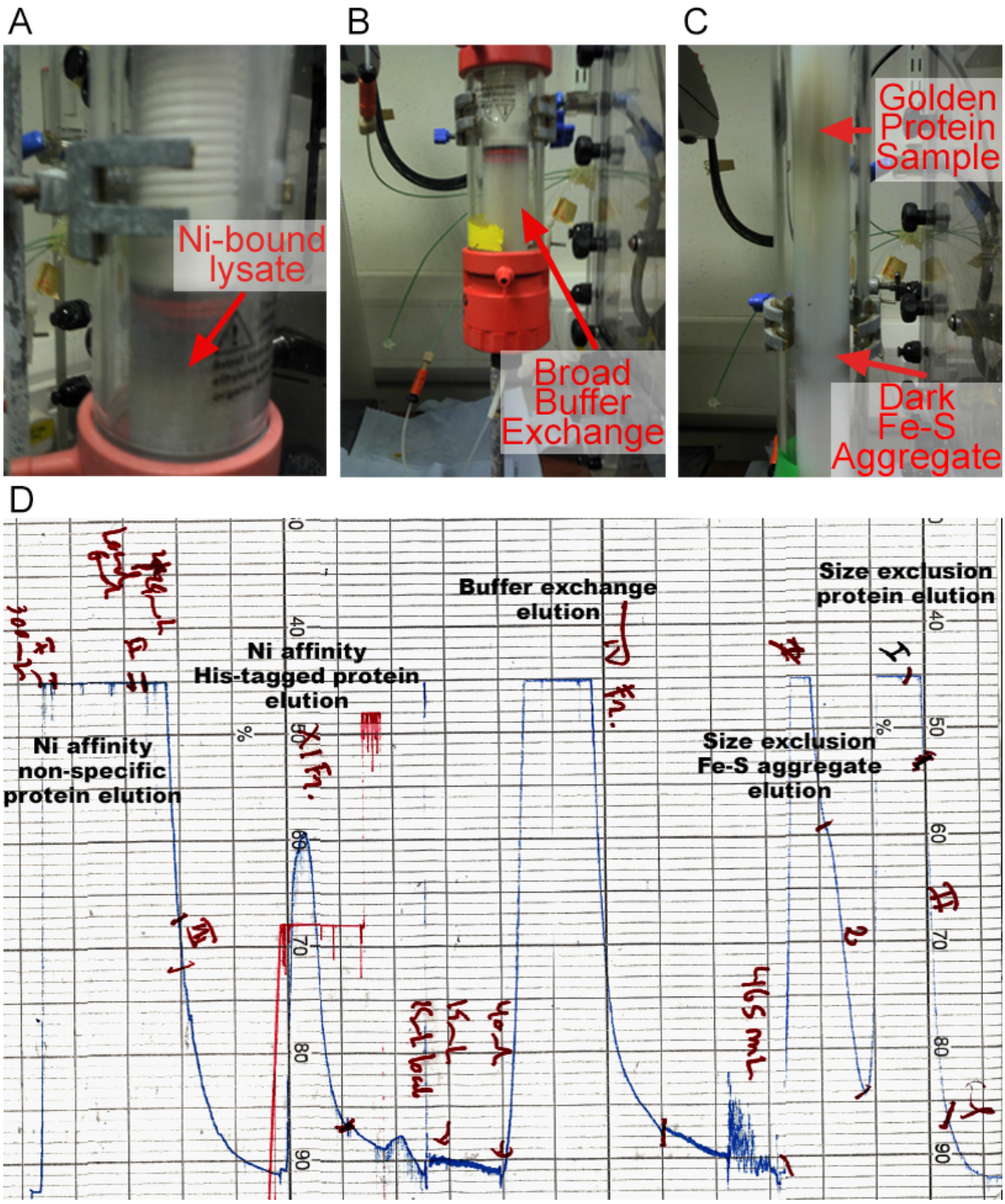



**Figure 36 HydG purification chromatography.**

**A:** Picture of a loaded Ni-NTA column. **B:** Picture of a loaded, small scale, S-75 gel filtration column, used for buffer exchange into a system without imidazole. **C:** Picture of a size exclusion chromatography S-75 column where a reconstituted sample of HydG was applied into. The larger Fe-S-protein aggregate is shown eluting faster (dark brown band), while the pure HydG sample is seen at the top (golden brown band); **D:** Example of an elution chromatogram during each step of the purification of HydG.

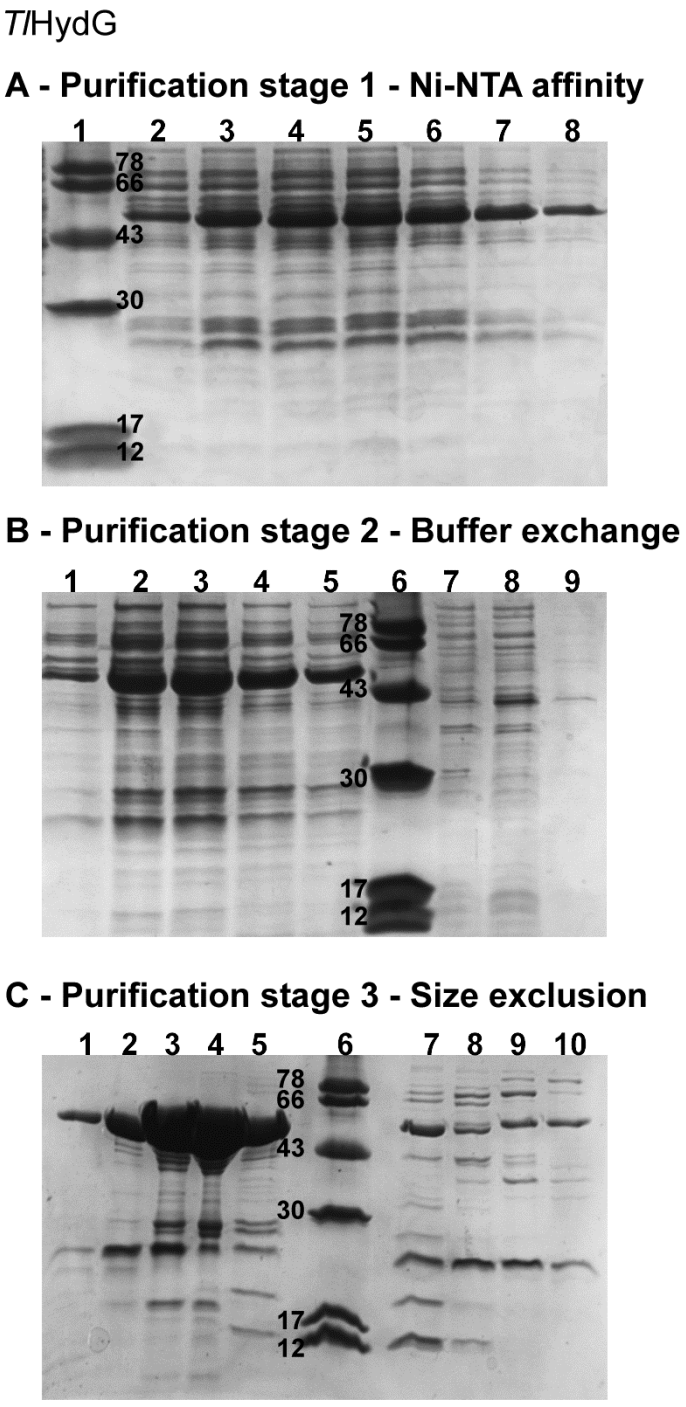



**Figure 37 15% SDS-PAGE analysis of the purification stages of *Tl*HydG.**

**A:** Ni-NTA affinity purification: 1 – Protein molecular weight markers; 2-8 – Samples collected at 60% Buffer B, corresponding to the desired protein at 55 kDa. **B:** Buffer exchange chromatography: 1-5 – Fractions collected from the small S-75 column used to buffer exchange into Buffer C; 6 – Protein molecular weight markers; 7 – Sample of cell lysate; 8 – Insoluble fraction; 9 – Clear lysate loaded onto the Ni-NTA column. **C:** Size exclusion chromatography: 1-5 – Soluble *Tl*HydG, with a golden-brown colour; 6 – Protein molecular weight markers; 7-10 – Dark brown/black [Fe-S] excess and protein aggregate.

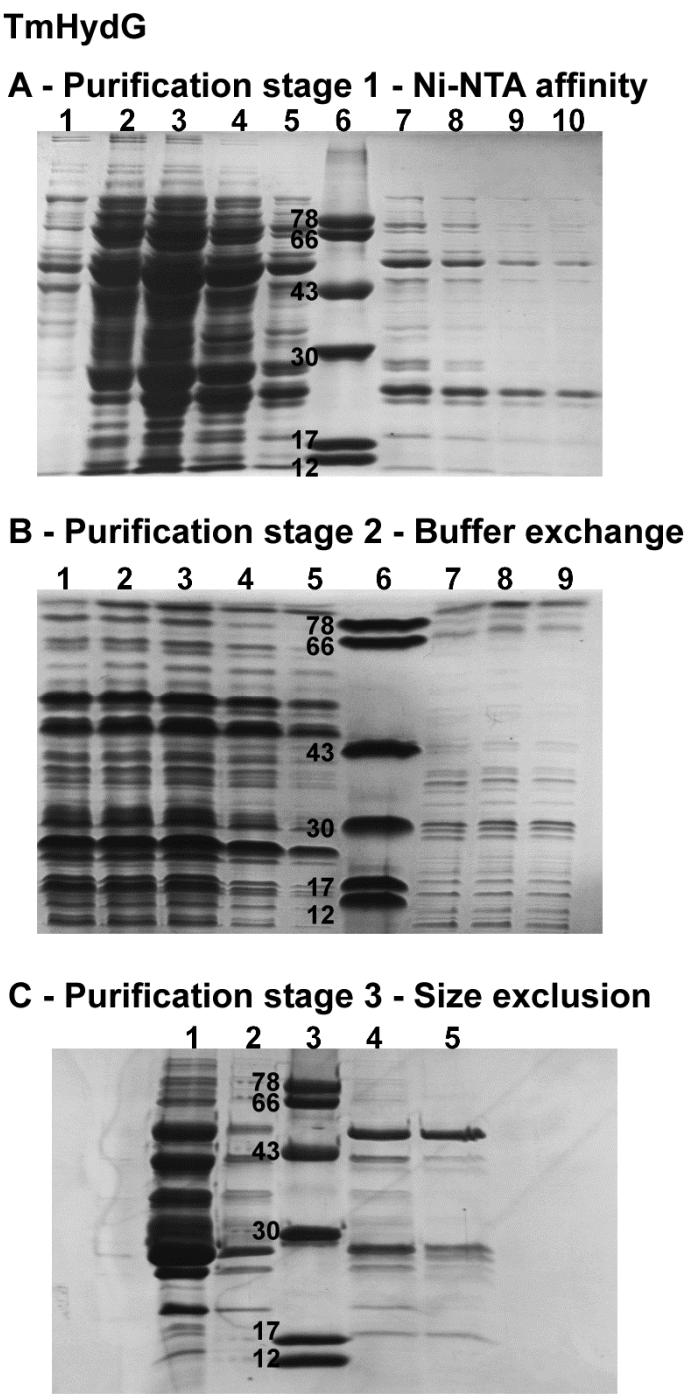



**Figure 38 15% SDS-PAGE analysis of the purification stages of *Tm*HydG.**

**A:** Ni-NTA affinity purification: 1-5 – Samples collected at 60% Buffer B, corresponding to the desired protein at 55 kDa; 6 – Protein molecular weight markers. **B:** Buffer exchange chromatography: 1-5 – Fractions collected from the small S-75 column used to buffer exchange into Buffer C; 6 – Protein molecular weight markers; 7 – Sample of cell lysate; 8 – Insoluble fraction; 9 – Clear lysate loaded onto the Ni-NTA column. **C:** Size exclusion chromatography: 1-2 – Dark brown/black [Fe-S] excess and protein aggregate; 3 – Protein molecular weight markers; 4-5 – Soluble *Tm*HydG, with a golden-brown colour.

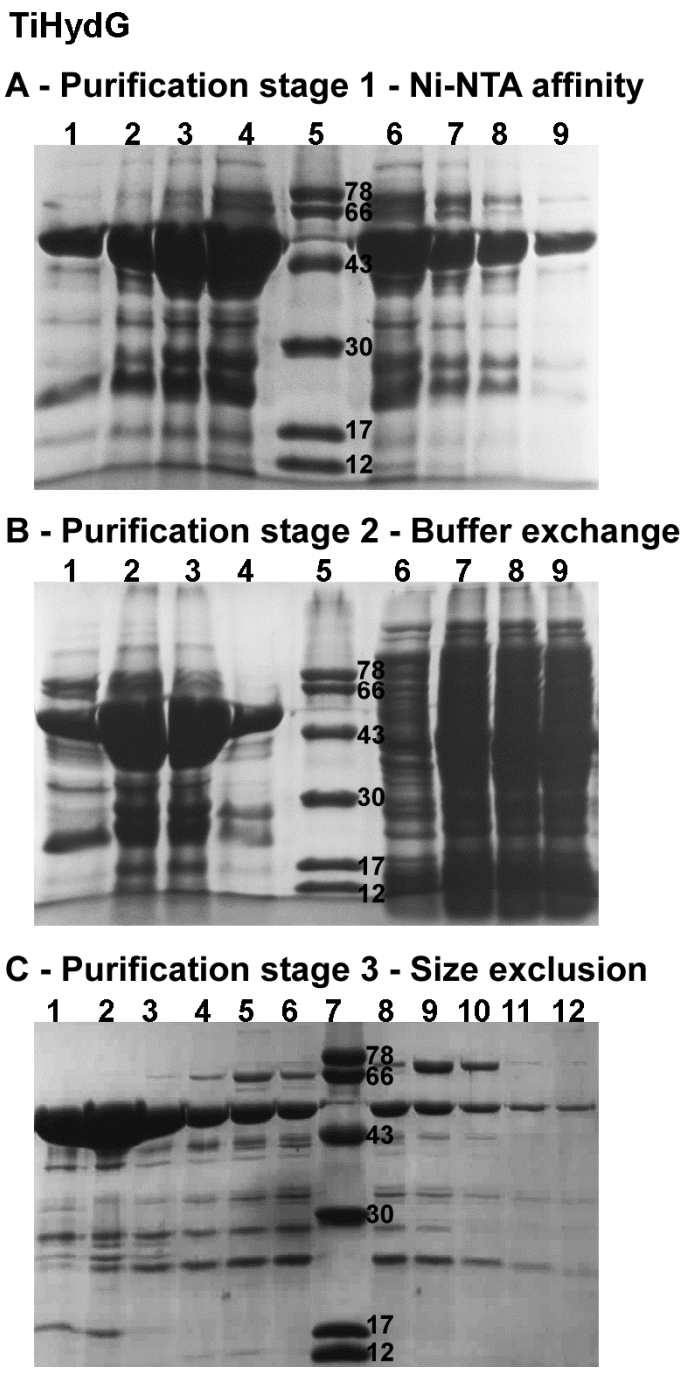



**Figure 39 15% SDS-PAGE analysis of the purification of *Ti*HydG.**

**A:** Ni-NTA affinity purification: 1 – Protein molecular weight markers; 2-8 – Samples collected at 60% Buffer B, corresponding to the desired protein at 55 kDa. **B:** Buffer exchange chromatography: 1-5 – Fractions collected from the small S-75 column used to buffer exchange into Buffer C; 6 – Protein molecular weight markers; 7 – Sample of cell lysate; 8 – Insoluble fraction; 9 – Clear lysate loaded onto the Ni-NTA column. **C:** Size exclusion chromatography: 1-5 – Soluble *Ti*HydG, with a golden-brown colour; 6 – Protein molecular weight markers; 7-10 – Dark brown/black [Fe-S] excess and protein aggregate.

Analysis by denaturing SDS-PAGE has shown that this expression method leads to the co-elution of the protein with some unexpected bands: 66 kDa; 43 kDa; 27 kDa; 25 kDa and 17 kDa, which are present in the samples even following the secondary size exclusion purification step. Their levels of co-purification seem to remain mostly stable across the purification of the different HydG enzymes (Figure 37; Figure 38; Figure 39). The chosen expression method co-expresses several proteins of the ISC machinery which correlate with most of the low molecular bands observed (Table 1). The highest molecular bands observed (78 kDa and 100 kDa) may arise from complexes formed between these enzymes [164].

**Table 1 Molecular weights of ISC proteins**

Adapted from Kriek *et al.* [523].

| Protein name | Protein molecular weight (kDa) |
|---|---|
| HscA | 65 |
| IscS | 46 |
| HscB | 20 |
| IscU | 14 |
| IscA | 12 |
| Fdx | 12 |

The levels of expression of the ISC-related proteins are quite low when compared with the yields of HydG, in particular *Tl*HydG and *Ti*HydG, whose purity can still be estimated at approximately 80% using the imaging software ImageJ [528]. Consistent with the observation of a lower amount of cell pellet, the overall yield of *Tm*HydG was significantly lower than the other two recombinant enzymes, leading to a much lower purity (50% purity estimated with ImageJ). Total protein concentration after purification was estimated using the Bradford method (Method 9.5.5) [529].

**Table 2 Protein yield estimation from different HydG purifications.**

The table summarizes the amount of cell paste that was obtained from a 5L cell culture and the total amount of protein after the main purification steps.

| Protein | Initial Pellet (g) | Total Protein Buffer exchange (mg) | Total Protein Size Exclusion (mg) |
|---|---|---|---|
| *Tl*HydG | 55 | 165 | 113 |
| *Tm*HydG | 38 | 50 | 29 |
| *Ti*HydG | 45 | 158 | 160 |

The protein yield estimates for each of the three HydG enzymes purified shows a clear distinction between *Tm*HydG, which does not express as successfully, and *Tl*HydG and *Ti*HydG. While several approaches could be followed to improve the yield of *Tm*HydG [530], all subsequent studies focused on the two proteins that showed a immediate higher level of protein expression (*Tl*HydG and *Ti*HydG).

### 2.3.3 Chemical Reconstitution of [4Fe-4S] Clusters in HydG

The [4Fe-4S] clusters in RSAM proteins are very labile and reversibly decompose under oxidizing conditions [531]. Additionally, the elution from the nickel affinity column by use of imidazole, which can not only chelate nickel but other transition metals like iron [532], may lead to degradation of the composition of the cluster. To restore the [4Fe-4S] cluster contents of a protein, a chemical reconstitution can be carried out (Method 9.6.1), by the slow addition of $FeCl_3$ and $Na_2S$ under agitation [533-535]. Employing the Fish method [536], the levels of iron were quantified in the enzymes before and after a chemical reconstitution, to estimate the degree of cluster assembly in the protein (Method 9.6.3).

**Table 3 Iron quantification of HydG samples before and after each reconstitution step.**

| Protein | Before 1st reconstitution (Meq) | After 1st reconstitution (Meq) | Before 2nd reconstitution (Meq) | After 2nd reconstitution (Meq) |
|---|---|---|---|---|
| *Tl*HydG | 1.82 ± 0.6 | 12.84 ± 3.6 | 4.31 ± 0.9 | 10.87 ± 2.2 |
| *Ti*HydG | 3.98 ± 0.2 | 13.45 ± 4.0 | 6.03 ± 0.4 | 10.46 ± 3.5 |

The chemical reconstitution of the proteins can also be followed by UV-visible measurement, as the [4Fe-4S] clusters have a characteristic 410nm peak, with an extinction coefficient of about 4000 $cm^{-1}$ $M^{-1}$ per iron atom, though these values are highly dependent on the oxidation state of the cluster [534]. Furthermore, the ratio of the cluster characteristic peak (410 nm) to that of the protein characteristic peak (280 nm) is a good indicator of the relative proportion of cluster chromophore to peptide, and values of 0.3-0.4 range are indicatives of an intact [4Fe-4S] cluster [537] per equivalent of protein.

The protein peaks at 280 nm consistently shows values above the theoretical extinction coefficient (*Ti*HydG $\varepsilon_{280}$ =41,720 $M^{-1}$ $cm^{-1}$ and *Tl*HydG $\varepsilon_{280}$ =41,260 $M^{-1}$ $cm^{-1}$) and two reasons may help explaining these results. First, it is well documented that Bradford estimations of protein concentration in RSAM enzymes are often imprecise [79, 128]. It does, however provide a quick estimation within an acceptable range, making it ideal for repeated testing across all purification and crystallization stages. Furthermore, the addition of an excess of $FeCl_3$ and $Na_2S$ often leads to some colloidal aggregation of iron-sulfide, which may increase the background signal of the samples, as was previously shown for *Ca*HydG [519]. This increase can artificially increase the characteristic $[4Fe\text{-}4S]^{2+}$ absorption peak at 410 nm, limiting the estimation of $[4Fe\text{-}4S]^{2+}$ present in the samples.

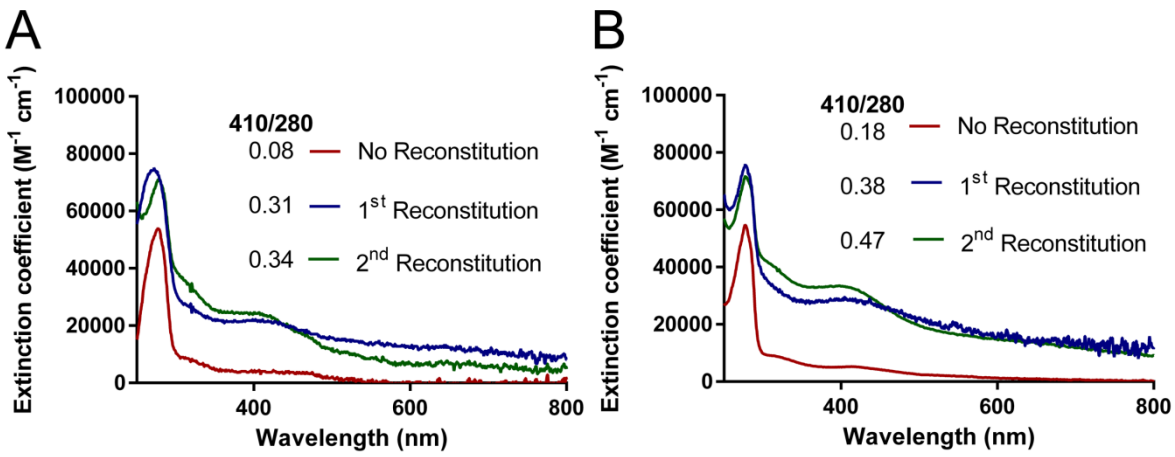



**Figure 40 UV-vis spectra of HydG proteins, before and after reconstitution steps.**

Red trace indicates samples prior to any reconstitution. Blue trace is the UV spectra of samples following a 10 Meq $FeCl_3/Na_2S$ reconstitution. Green trace is the spectra after the final reconstitution step with 5 Meq. $FeCl_3/Na_2S$. The increase in the signal intensity at 410 nm is indicative of the presence of [4Fe-4S] clusters. **A:** *Tl*HydG reconstitution. Protein quantification by Bradford method estimated 2.5 μM before reconstitution (red trace), 2.4 μM after the first reconstitution (blue trace) and 2.8 μM after the second reconstitution (green trace).B: *Ti*HydG reconstitution. Protein quantification by Bradford method estimated 17.4 μM (red trace), 15.4 μM (blue trace) and 11.8 μM (green trace).

Following the initial chemical reconstitution with 10 molar equivalents (Meq) of $FeCl_3$ and $Na_2S$, the background signal of both samples increases significantly (Figure 40, 1st reconstitution), due to the excess iron-sulfide aggregates in the sample. These results explain why the calculated iron values (Table 3) are significantly higher than the predicted eight irons arising from two fully reconstituted [4Fe-4S] clusters. The presence of these colloidal Fe-S aggregates makes the size exclusion chromatography step in the purification critical in obtaining pure protein fractions (section 2.3.2).

The second reconstitution, carried out after gel filtration, increases the 410 nm signal, with lower background interference (Figure 40, 1st reconstitution). The calculated extinction coefficients are similar for both proteins ([*Ti*HydG]= 11.8 μM, $\varepsilon_{410}$= 32,900 $M^{-1}$ $cm^{-1}$; [*Tl*HydG]= 6.0 μM; $\varepsilon_{410}$= 28,900 $M^{-1}$ $cm^{-1}$) and are indicative with two $[4Fe\text{-}4S]^{2+}$ clusters present in the protein [534]. In agreement with the UV results, the iron content was estimated at around 10 iron atoms irons per protein (Table 3), with very small differences between the *Ti* and *Tl* HydG enzymes. The excess iron calculated may be related to a small fraction of the iron-sulfur aggregate that is mostly, but

perhaps not completely, removed in the size exclusion chromatography step. However, the presence of this small amount of protein aggregates did not disturb the downstream applications of the protein.

## 2.4 Preliminary Activity Characterization of Thermophilic HydGs

Following a successful purification, where a homogeneous, soluble and stable enzyme is purified, assaying the enzyme's activity enables the assessment of whether the enzyme has maintained its correct folding and native activity or not. In the case of RSAM enzymes, it is a particularly important step as [4Fe-4S] cluster mediated SAM cleavage provides a good measure of the success of the chemical reconstitution.

Experiments by R. Driesener have laid the foundation for the kinetic analysis of the enzymatic activity of HydG [519]. An end-point HydG reaction was carried out with an incubation of 60 minutes (Method 9.7), which, following the first order reaction profile described previously (Figure 41), would guarantee a near-maximum amount of product [47, 311].

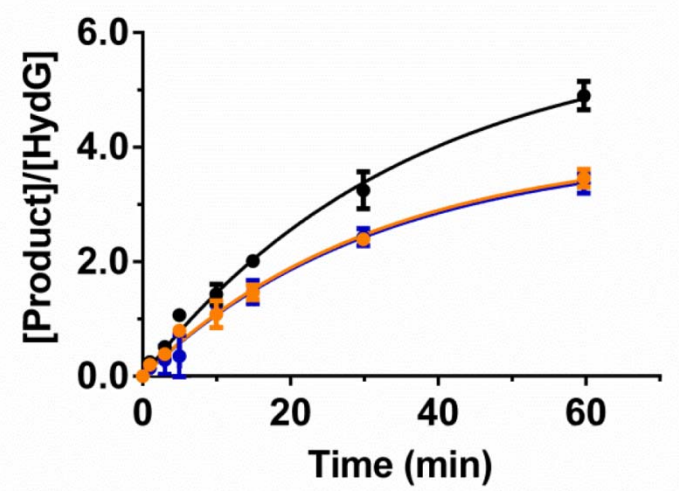



**Figure 41 Time dependent SAM and tyrosine cleavage by *Ca*HydG wild type enzyme.**

Formation of DOA (-●-), *p*-cresol (-●-) and cyanide (-●-). Reprinted with permission, from R. Driesener [519].

Formation of the *p*-cresol, DOA and cyanide products (Figure 42) was followed by RP-HPLC. The RSAM conserved reaction, where SAM is reductively cleaved after hydrogen abstraction, was followed by the formation of DOA, measurable by its absorbance at 254 nm and a retention time (RT) of 18.8 min by comparison with standards. The HydG specific cleavage of L-tyrosine **18** was followed by *p*-cresol **19** formation, which was measured by both its absorbance at 280 nm as well as its native fluorescence with $\lambda_{ex}$= 274 nm, $\lambda_{em}$= 312 nm, and a RT 25.7 min.

The instability of dehydroglycine **25** does not allow for its direct measurement [47, 311], but the product of its cleavage, cyanide **21**, was observed after conversion to a fluorescent derivative following a modified method of Tracqui [538]. The derivatisation reaction allows the observation of the fluorescent 1-cyanobenz[f]isoindole (CBI) **104**, with $\lambda_{ex}$= 418 nm, $\lambda_{em}$= 454 nm, and an RT of 12.8 min.

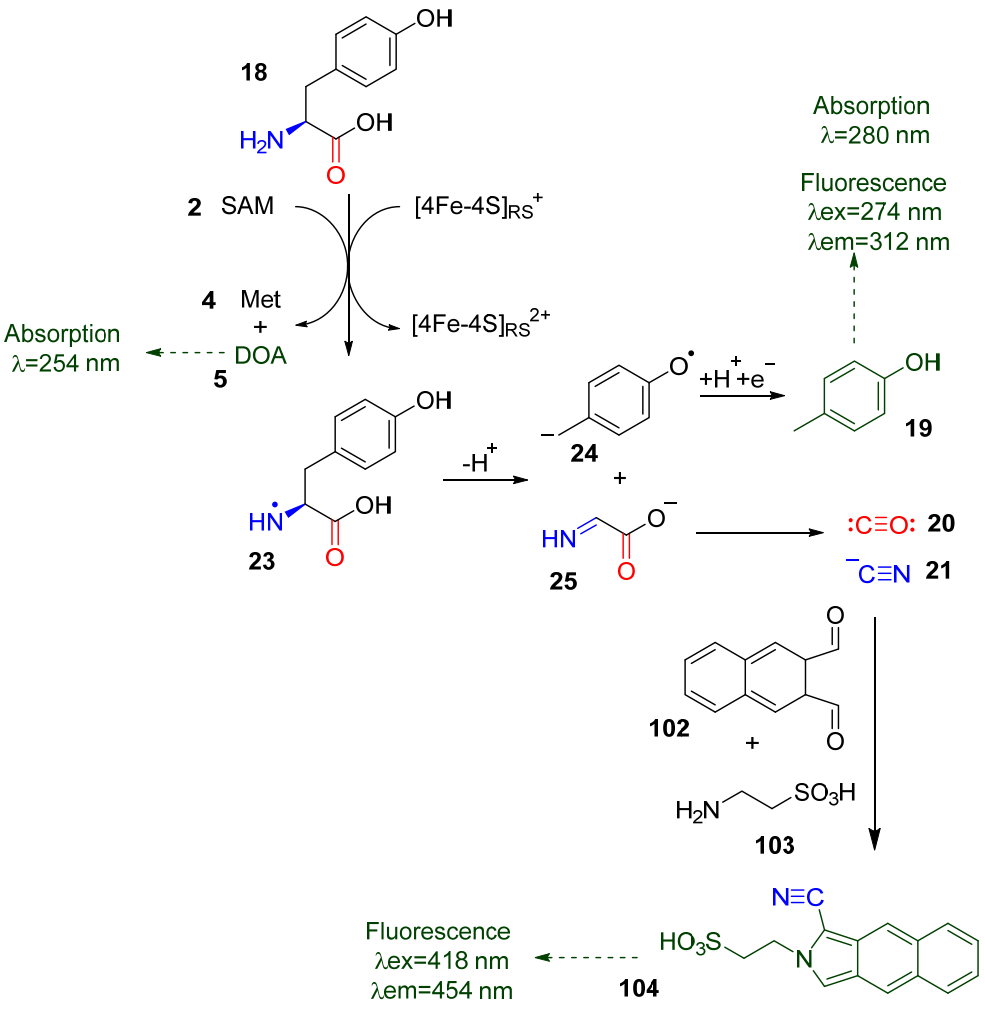



**Figure 42 Identification of reaction products of HydG catalysed tyrosine cleavage by RP-HPLC.**

SAM **2** turnover is followed by the formation of DOA **5**; $C_\alpha$-$C_\beta$ cleavage of tyrosine **18** is measured with the formation of *p*-cresol **19** and formation of the diatomic products is followed by the derivatisation of cyanide **21**.

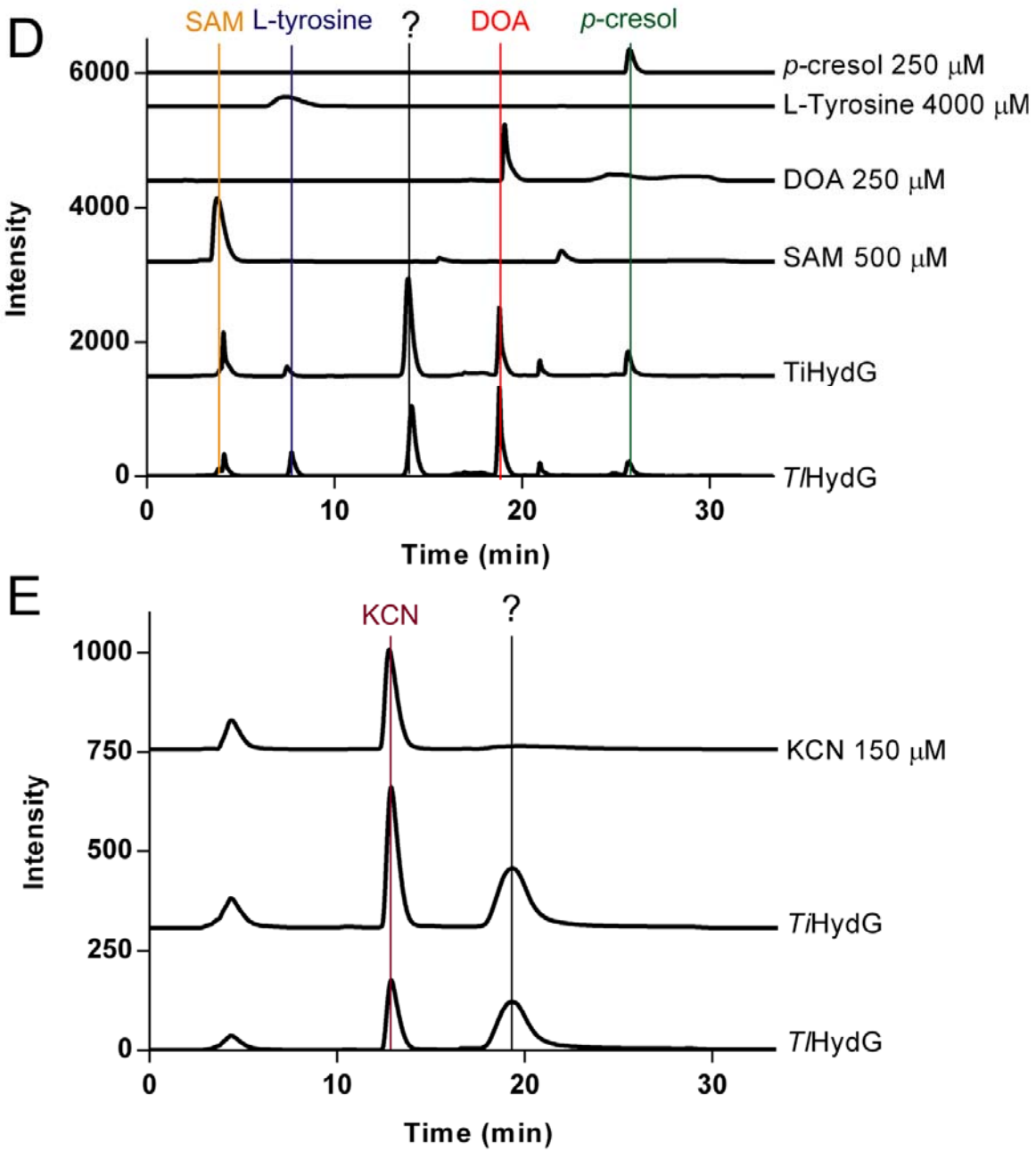



**Figure 43 RT-HPLC chromatogram of the activity assays for *Tl*HydG and *Ti*HydG.**

**A:** Analysis of SAM and L-tyrosine cleavage, together with SAM (RT=4.1 min), DOA (RT=18.8 min), L-tyrosine (RT= 7.8 min) and *p*-cresol standards (RT=25.7 min). A peak can be observed at RT=13.9 min for both protein samples, but absent in all standards, which is currently unidentified. **B:** Analysis of cyanide formation, together with potassium cyanide standards (RT=12.8 min). The protein samples show a strong, broad peak around RT=19.7 min which is unidentified.

**Table 4 Quantification of product formation by *Tl*HydG and *Ti*HydG.**

The apparent rate does not account for variation of rate with time, but was calculated as a final end point assuming a linear process as follows: $\text{Rate} = \frac{\Delta C}{\text{time}}$. Apparent turnover number was determined from the apparent linear rates.

| Enzyme | Product | ΔC (μM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number (x$10^{-4}$ $s^{-1}$) |
|---|---|---|---|---|
| *Tl*HydG | DOA | 172.3 | 47.8 | 9.6 |
| | *p*-cresol | 47.6 | 13.2 | 2.6 |
| | Cyanide | 34.0 | 9.5 | 1.9 |
| *Ti*HydG | DOA | 128.7 | 35.8 | 7.1 |
| | *p*-cresol | 116.2 | 32.3 | 6.5 |
| | Cyanide | 68.1 | 18.9 | 3.8 |

A single activity measurement was carried out for each enzyme, as the goal was simply to observe full activity following reconstitution. Subsequent activity studies in chapters 4 and 5 were carried out in triplicate. The quantification of product formation was calculated using standard concentrations of DOA, *p*-cresol and cyanide (Appendix A2).

Both enzymes show both L-tyrosine cleavage as well as cyanide formation, indicating, that the RSAM cluster and the auxiliary cluster, respectively, are intact [296]. An unidentified peak can be observed in both activity assays of DOA/*p*-cresol at 13.9 minutes, which does not represent any of the controls. Further activity studies, in chapters 4 and 5, do not show this peak present in the assay chromatograms, thus it is assumed the nature of the peak is an impurity present in the column. Additionally, the analysis of the cyanide derivatisation assay also shows a large peak at 19.7 minutes, which could not be identified.

Some differences can be observed in the turnover of *Tl*HydG and *Ti*HydG, with the first showing a large gap between the apparent turnover of DOA and *p*-cresol. The increased cleavage of SAM is well known in RSAM enzymes in the presence of strong reducing agents, such as dithionite, which allows the uncoupled turnover of SAM even in the absence of substrate [406, 539].

Despite only single point estimations, the calculation of an apparent turnover number does allow the comparison of values with those presented in a full kinetic analysis of the activity of *Ca*HydG [311]. The results show a three-fold decrease in the $k_{cat}^{app}$ in the activity of *Ti*HydG (DOA $k_{cat}^{app}$ = 7.1 x$10^{-4}$ $s^{-1}$; *p*-cresol $k_{cat}^{app}$ = 6.5 x$10^{-4}$ $s^{-1}$ and $^{-}$CN $k_{cat}^{app}$ = 3.8 x$10^{-4}$ $s^{-1}$) when compared with *Ca*HydG (DOA $k_{cat}^{app}$ = 23.0 ± 3.0 x$10^{-4}$ $s^{-1}$; *p*-cresol $k_{cat}^{app}$ = 18.0 ± 2.0 x$10^{-4}$ $s^{-1}$ and $^{-}$CN $k_{cat}^{app}$ = 6.0 ± 0.0 x$10^{-4}$ $s^{-1}$), and close to 7-fold decrease in $k_{cat}^{app}$ for *p*-cresol when *Ca*HydG activity is compared with *Tl*HydG (*p*-cresol $k_{cat}^{app}$ = 1.9 x$10^{-4}$ $s^{-1}$).

The lower activity at 37ºC for both thermophilic enzymes in comparison with the mesophilic *Ca*HydG was expected, as there is a well-documented difference when comparing mesophilic and thermophilic isoforms of enzymes [540-542].

## 2.5 Initial Crystallization Studies

All crystallization studies were carried out within an anaerobic chamber ($O_2$ < 2 ppm) and at a constant temperature of 18 °C, where the crystallization screens as well as all crystallization trays were kept.

Initial screening for crystallization conditions for *Tm*HydG had been attempted by Dr. Jenny E. Harmer, using a commercial Hampton PEGRx HT 96-well screen, yielding strong phase separation, but no formation of crystals. These initial results, together with the work also carried out in the crystallization of LipA [543], established a framework for potential RSAM crystallization: high protein concentration (approximately 1 mM), together with several equivalents of SAM (5-10 mM) being mixed with high concentrations of PEG precipitant in the crystallization conditions (Method 9.8).

Using the Hampton Research PEGRx HT broad screen did not yield any successful crystallization of *Tl*HydG. Most conditions led to heavy and almost immediate (within 8 h) precipitation of the protein. This could be interpreted as indicating an excessive amount of protein in the drop which may have led to a quick change into the precipitate phase without reaching a mesostable nucleation (see section 8.2).

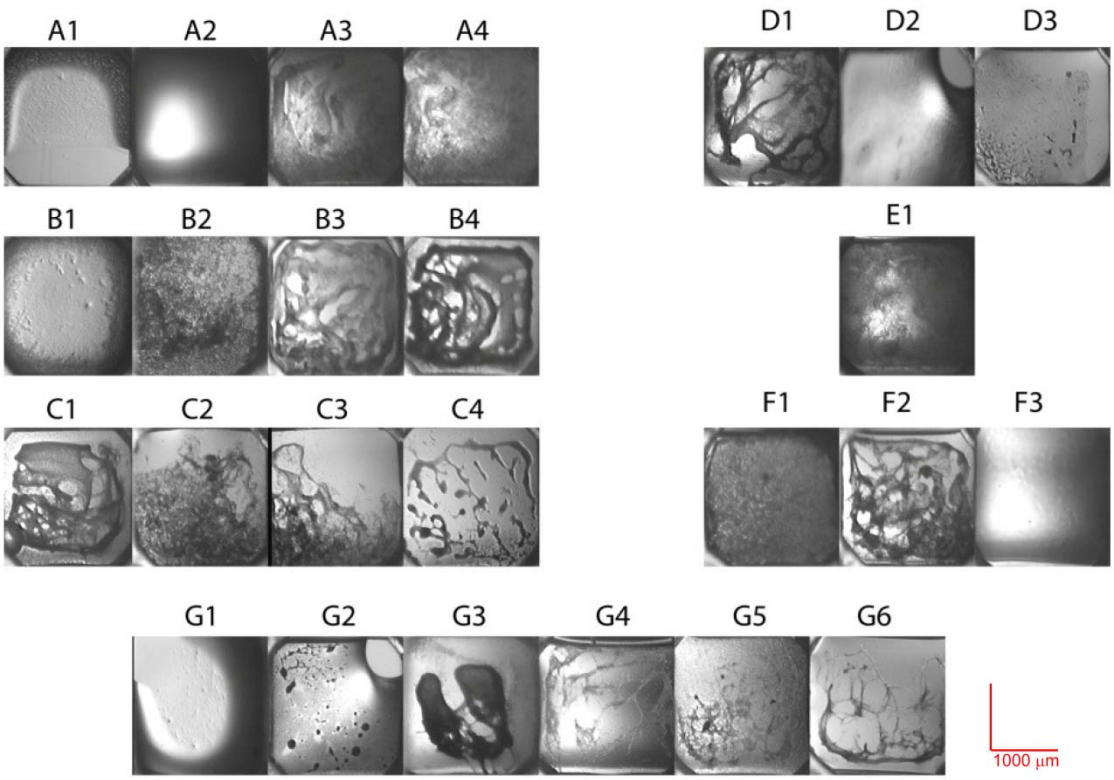



**Figure 44 Examples of the crystallization wells from the screening of HydG from *Thermotoga lettingae***

The crystallization was carried out using Hampton PEGRx HT commercial broad screen. Extensive aggregation and precipitation can be seen, but no clear crystal formation was observed in any of the 96-conditions screened.

Using the same approach for *Ti*HydG, this time by means of two different broad PEG screens, Hampton Research PEGRx HT and Molecular Dimensions PACT Premier, led to an unambiguous identification of four different conditions where the enzyme (0.8 mM) crystallizes, in the presence of 10 mM SAM and either no tyrosine (conditions A8 from PACT Premier and F1 from PEGRx HT) or 3 mM L-tyrosine (conditions E12 from PACT Premier and C11 from PEGRx HT). Crystallization was observed after three weeks of incubation and the four crystals were collected from the respective drops, using a crystal loop, directly frozen in liquid $N_2$ and stored until they could be tested at the Diamond synchrotron facilities (section 2.6).

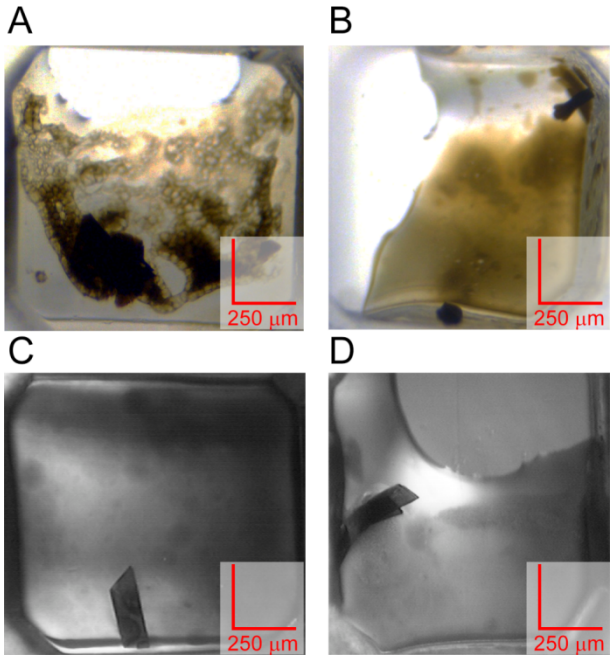



**Figure 45 Crystallization wells of *Ti*HydG**

Successful crystallization conditions for HydG from *Thermoanaerobacter italicus*. **A:** Condition A8 from PACT Premier; and **B:** condition F1 from Hampton PEGRx, both grown in the presence of SAM (10 mM). **C:** Condition E12 from PACT Premier and **D:** condition C11 from Hampton PEGRx, both grown in the presence of SAM (10 mM) and tyrosine (3 mM).

**Table 5 Conditions corresponding to the wells shown in Figure 45**

| Figure | Screen | Condition | Formulation | |
|---|---|---|---|---|
| A | Hampton PEGRx | A8 | 0.1 M MES, pH 6.0 | 22% v/v polyethylene glycol 400 |
| B | PACT Premier | F1 | 0.1 M Bis-Tris Propane, pH 6.5 | 20% w/v polyethylene glycol 3,350 | 0.2M sodium fluoride |
| C | PACT Premier | E12 | | 20% w/v polyethylene glycol 3,350 | 0.2M sodium malonate |
| D | Hampton PEGRx | C11 | 0.1 M Bis-Tris, pH 6.5 | 20% w/v polyethylene glycol monomethyl ether 5,000 |

### 2.5.1 Reproducibility of *Ti*HydG Crystal Formation in 24-Well Format

The success of a crystallization study lies in the ability to replicate the conditions identified in the commercial crystallization broad screens discussed in section 2.5. Further manipulations to either the protein mixture or the formulation identified, can lead to an improvement in crystal size and diffraction, crystallization in the presence of substrates, products and inhibitors and crystallization of mutants [544]. Furthermore, solving the structures from isomorphous crystals is often straightforward, even if identifying the crystallization condition is not.

From an experimental perspective, it is often more desirable to have crystal growth occurring by the hanging drop method, using a 24-well format, rather than the sitting drop method in a 96-well format. The main advantage lies in the ability to increase the volume of the crystallization drop, which is no longer constrained by the size of the small well, allowing a higher variability in the composition of protein:precipitant and potentially leading to larger crystals. A crystal grown in the hanging drop method is also easier to remove into the crystal loop to be tested.

Unfortunately changing from 24-well to 96-well does not necessarily guarantee crystal formation. The volume of precipitant in the reservoir and the volume of air between the reservoir and the drop lead to different equilibrium constants that may disturb the crystallization events. These small variations are the most likely explanation into why any attempt to change into 24-well format was unsuccessful (Method 9.8.1) (Figure 46).

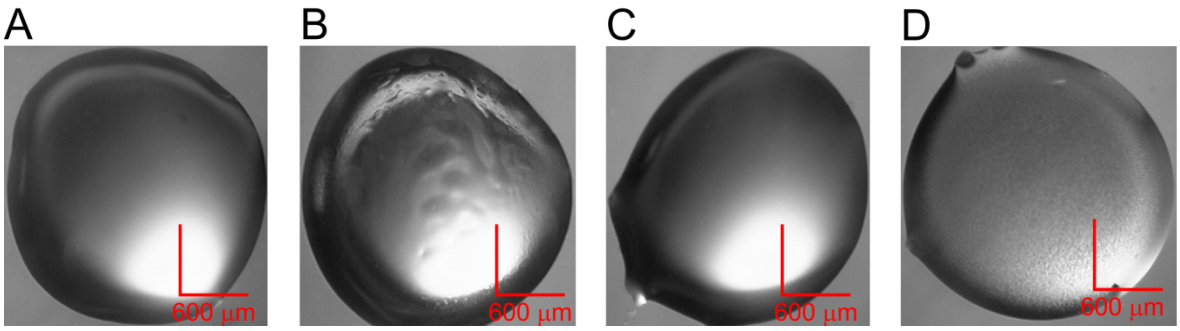



**Figure 46 24-Well replication of previous crystal forming conditions.**

Images of 24-well hanging drops for *Ti*HydG crystallization under the conditions where crystals were obtained in sitting drop 96-well plates. A – Condition A8 from PACT Premier screen; B – Condition F1 from PEGRx Hampton screen; C – Condition E12 from PACT Premier screen; D – Condition C11 from PEGRx Hampton screen.

### 2.5.2 Reproducibility of *Ti*HydG Crystal Formation in 96-Well Format

With the limitation of growing the crystals under 96-well format, all initial formulations were replicated in house (see Method, Table M18) and crystal growth was attempted. Of the four conditions described in Table 5 (section 2.5), only two of them showed formation of crystals: condition C11 from, Hampton PEGRx HT screen and condition F1 from PACT Premier screen.

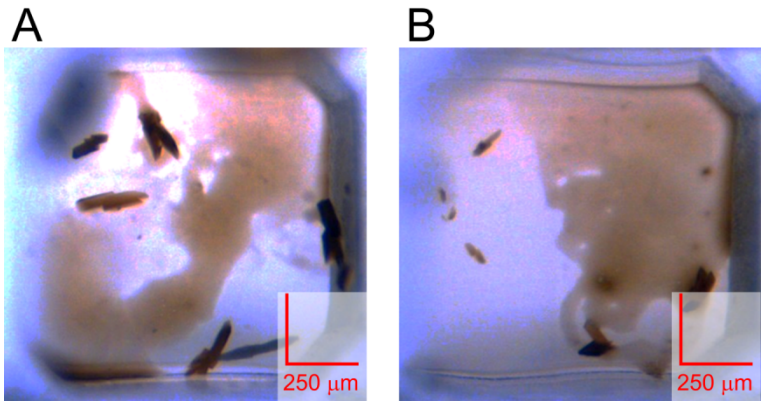



**Figure 47 Screening wells that showed crystal formation of *Ti*HydG with crystallization solutions mixed in house.**

**A:** Crystals grown in condition C11 of PEGRx HT (0.1 M Bis-Tris pH 6.5, 20% PEGME 5,000). **B:** Crystals grown in condition F1 of PACT Premier (0.1 M Bis-Tris pH 6.5, 20% w/v PEGME 5,000).

The reproducibility of crystal formation was, however, far from consistent. Replicating the same condition 288 times (3 sitting wells in a 96-well plate), led to only 48 drops showing crystals for both conditions, translating into a 16.7% chance of obtaining crystal growth. Several reasons can be presumed to be the cause of such a low level frequency of crystal formation. The replication was made with stock solutions prepared in house, which may not accurately mimic the stocks commercially available. Additionally, small unavoidable experimental variations during the set-up of the crystallization conditions may lead to differences in the amount of precipitant/protein added into the drop. Going further, there are unavoidable timing differences when pipetting the samples which can allow some evaporation to occur. Evaporation variation, however, seems unlikely as an explanation for the results observed, bearing in mind there was no clear pattern of crystal formation in the plate crystal.

### 2.5.3 Microseeding to Increase the Reproducibility of *Ti*HydG Crystal Formation in 96-Well Format

In a crystallization experiment, the initial solution conditions may be unsaturated. If that is the case, nucleation cannot occur [545]. Seeding (section 8.2.4) is a technique where the optimization of growth conditions is independent from those required for nucleation, improving the probability of ordered crystal growth. In the case of microseeding, a crystal is placed in a stabilizing solution and broken into microscopic particles which can then serve as tiny nucleation points. This seed stock usually needs to be diluted, as a high number of nucleation points in the crystallization solution leads to the formation of a shower of microcrystals. On the other hand, if the seed stock is too dilute, there is the possibility of not transferring any nuclei, and no crystal is formed.

A previously grown crystal in condition C11 from Hampton PEGRx screen was transferred into an Eppendorf tube with 100 μL of condition C11. Using a Hampton Research seed bead, the crystal was crushed and serial dilutions were made from $1x10^{-1}$ to $1x10^{-9}$ (Method 9.10).

The timing regarding the addition of seeds is of critical importance, as the formation of crystals from seeds only occurs if seeds are added when the drop is in the metastable zone of crystal formation (see section 8.2.4, Figure 162). Though it is possible to create a crystallographic profile of a protein, obtaining a phase diagram for the protein studied, this is usually labour intensive and not widely attempted [546]. Seeds were added to the drops after 8h incubation, as visually the drops showed little formation of precipitate by that stage.

The crystal growth obtained was far from regular. A typical result in a seeding experiment reveals a shower of small crystals that would grow larger in size and smaller in number as the degree of dilution increased (see section 8.2.4, Figure 161). The results obtained in this experiment (Figure 48), however, show two crystalline formations at $1x10^{-2}$ (Figure 48 B) and one crystal at $1x10^{-9}$ (Figure 48 I), with no crystals being observed at the other dilutions. These results seem to only reproduce the low consistency of these crystals' replication in the absence of seeding.

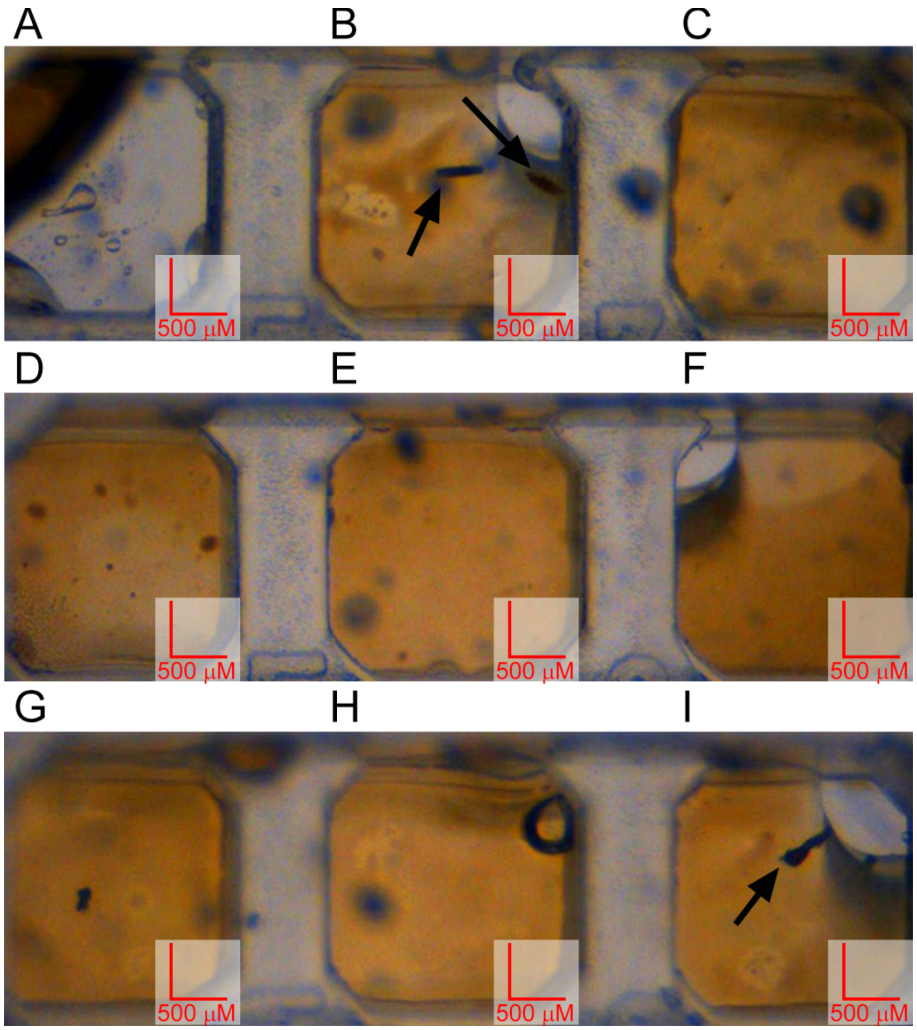



**Figure 48 Microseeding results.**

Microseeding experiment with crystals from *Ti*HydG, using the F1 condition from Hampton PEGRx screen. Serial dilutions on a 1/10 basis, from A – $1x10^{-1}$ to I – $1x10^{-9}$.

With the failure of microseeding to improve crystallization reproducibility, macroseeding techniques could prove invaluable. In macroseeding a single crystal is moved to fresh drops, in an attempt to enlarge the crystal. It is a time consuming process, and a lot of care must be taken, as moving the single crystal from drop to drop may lead to fractures in the crystal, or, in extreme cases, loss of the crystal [547].

Microseeding of HydG was, however, proven more successful when the protein concentration was decreased (see section 2.9.2), allowing for the expected control in crystal formation.

## 2.6 Data Collection and Structural Model Building

### 2.6.1 X-Ray Data Collection

The four initial crystals obtained for *Ti*HydG (see section 2.5, Figure 45) were flash frozen in liquid $N_2$ in the absence of external cryoprotectants. The addition of a cryoprotectant to the drop solution is usually an important parameter in obtaining high quality diffraction data from a frozen crystal. Rapid freezing can lead to ice crystal formation which causes damage to the protein crystal, leading to a decrease the overall quality of the data. However, adding a crystal to a cryoprotectant is also not without risks, as further crystal manipulations may also cause damage [548, 549]. Fortunately, in the case of *Ti*HydG, the high levels of polyethyelene glycols (PEGs) present in the crystallization condition can function as suitable cryoprotectants, as their ability to displace water molecules and formation of an amorphous glass when frozen prevents the formation of ice crystals[548, 549].

The four crystals, named A8, F1, E12 and C11 due to the different screen conditions where each developed (Figure 45 and Table 5) were exposed to X-ray radiation at Diamond Light Source, at a constant temperature of 100K, on beamline I04. This beamline possesses the capability of a tuneable wavelength, which is advantageous to obtain anomalous data near the κ-iron edge, at 1.75 Å [550]. A series of diffraction images (dataset) was obtained by rotation of the crystal in the X-ray beam, sweeping the incident beam through the reciprocal space of the crystal. The collection of two initial images, 90 ° apart, allows the automatic pipeline EDNA [551] to predict the probable space group of the crystal and put forward a collection strategy that would allow a complete dataset to be collected. An important aspect of EDNA is the prediction of radiation damage, which would negatively affect the quality of the data, in the calculations for the collection strategy.

EDNA strategies are usually divided in: normal collection; a soft collection, for crystals which are particularly sensitive to radiation damage; and high multiplicity, where significantly more images are collected. The latter strategy is particularly important for anomalous data because both Friedel's pairs need to be collected for each reflection (effectively doubling the amount of data needed to be collected), to observe where the equality (Friedel's law) breaks down. High multiplicity ensures reliable data, but requires the crystal to be presented to the beamline for longer, and may lead to crystal damage before a full dataset is collected [552]. The initial information obtained from EDNA for the four crystals is shown in Table 6.

**Table 6 EDNA strategies for the collection of complete datasets of the *Ti*HydG crystals tested.**

| Crystal | Approximate Resolution (Å) | Space group | Oscillation (°) | No. images for full dataset Standard | Exposure time (s) Standard | No. images for full dataset Anomalous | Exposure time (s) Anomalous |
|---|---|---|---|---|---|---|---|
| A8 | 3.0 | P1 | 0.2 | 1780 | 0.067 | 2400 | 0.067 |
| C11 | 2.5 | P1 | 0.2 | 1330 | 0.204 | 1800 | 0.314 |
| E12 | 2.4 | P1 | 0.2 | 1345 | 0.067 | 1800 | 0.067 |
| F1 | 2.0 | P1 | 0.2 | 1834 | 0.067 | 2400 | 0.067 |

The initial estimations by EDNA showed a lot of similarities between the four crystals tested. Crystal C11 seems to have a worse crystal quality, as a much higher exposure time was predicted in order to collect the information. All crystals were predicted to diffract to similar resolution ranges, and all samples shared the same space group, P1.

A space group identifies which symmetry elements are present in the crystal lattice. Its determination is an essential step in structural analysis as understanding the symmetry behind the crystal lattice ensures minimum information is needed for the complete description of the contents of the unit cell. HydG's space group, P1, is characterized by an absence of any symmetry operators (rotation, rotation-inversion, screw axes,), other than lattice translations along the: a, b and c unit cell directions. A consequence of this space group is that a 180º crystal rotation was necessary to permit the collection of a full data set. 360º of data would be considered a full data set, but it is possible to collect only half of the data (180º) due to the diffraction intensities showing centrosymmetric symmetry even if the crystal does not, a phenomenon known as the Friedel's Law [553]. The speed of state-of-the-art synchrotron facilities, such as the ones used from this project, greatly minimizes the problem of lack of symmetry, as the highly sensitive detector allows for a very fast collection of several sets worth of data.

With the appropriate strategy calculated, a fluorescence scan was also carried out to confirm the presence of iron in the samples. The presence of iron in RSAM can be exploited to collect anomalous data in the Fe κ-edge, opening the possibility of solving the structure by SAD (see section8.3.2.3). As expected, all crystals showed a strong signal near the fluorescent edge of iron, at approximately 7,120 eV (Figure 49 C and Appendix A3).

Two sets of data for all four crystals were collected from the exact same position in the crystals, both standard data and high multiplicity (anomalous) data. The diffraction images for the crystals tested were processed automatically by X-ray Detector Software (XDS) [554], with data reduction/scaling completed automatically using Pointless and SCALA [555]. Results for each data collected are summarized below. For a deeper discussion regarding each of the processing stages (see section 8.3.2.4).

An example of a diffraction image for the crystal which diffracted to highest resolution (crystal F1) is seen in Figure 49 A. The automatic data processing presents a set of data, shown in Table 5, which assesses the quality of data. Data relating to the three other data sets collected (crystals A8, C11 and E12) is presented in Appendix section A3.

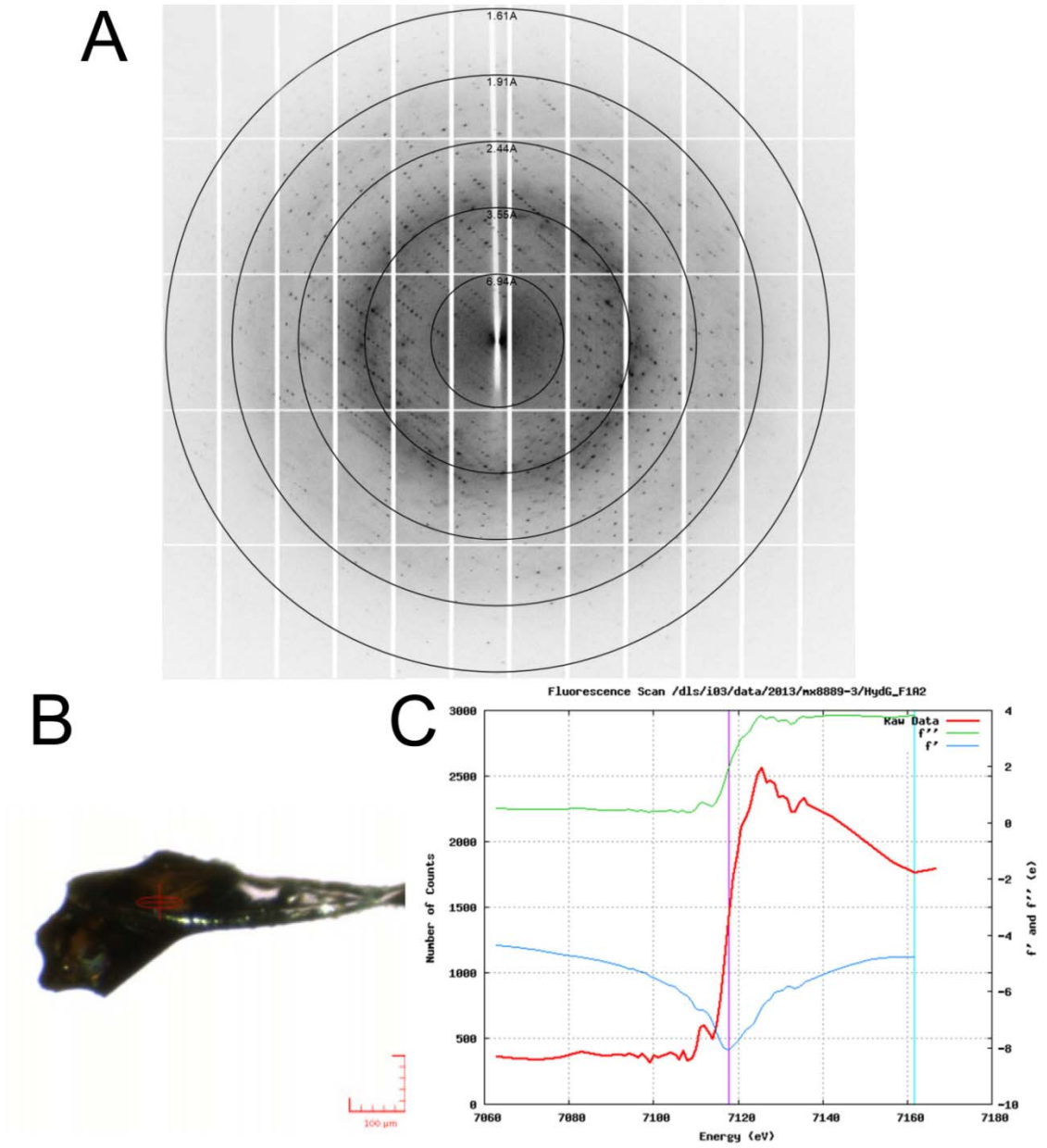



**Figure 49 X-ray diffraction of *Ti*HydG crystal F1.**

**A:** Diffraction pattern from 12 keV synchrotron radiation to 1.6 Å resolution, from the crystal of *Ti*HydG grown in the presence of SAM (0.8 mM HydG, 10 mM SAM), directly frozen after formation in condition F1 of PACT Premier screen (0.1M Bis-Tris propane, pH 6.5;0.2M sodium fluoride; 20% PEG 3350). **B:** Crystal placed in the loop during testing. Loop size – 0.3–0.4 mm, Beam size – 80 x 20 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescence count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table 7 Data collection statistics. Highest resolution shell shown in parenthesis**

| | **HydG** | **HydG Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | F1 Hampton PEGRx | |
| **Data collection** | | |
| Space group | P1 | |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 54.13; 56.19; 84.92 | |
| α, β, γ (°) | 89.59; 83.62; 66.84 | |
| | | *Peak* |
| Wavelength (Å) | 0.97625 | 1.73891 |
| No. of reflections | 181098 (9804-2389) | 173565 (3610-2406) |
| No. of unique reflections | 74941 (4283-876) | 53419 (1242-704) |
| Resolution (Å) | 84.33-1.59 (7.11-1.59) | 84.41-1.84 (8.23-1.84) |
| $R_{sym}$ or $R_{merge}$ | 0.039 (0.033-0.349) | 0.049 (0.043-0.484) |
| $I$ / σ$I$ | 12.3 (30.6-2.2) | 14.6 (26.9-2.1) |
| Completeness (%) | 96.2 (96.4-94.5) | 63.3 (77.6-4.8) |
| Redundancy | 2.7 (2.7-2.7) | 1.6 (1.5-1.1) |
| Twinning score | 2.37, no twinning | 2.25, no twinning |

In Table 7, a collection of values are shown that evaluate the quality of the data collected. The most common measurement of crystal data quality relates to the resolution of the data collected, which indicates the distance corresponding to the smallest observable feature. If two objects are closer than the given resolution value, it will not be possible to distinguish between the electron densities of the two [553]. Resolution values below 2.5 Å are considered good resolution, as the backbone of the protein, together with many side chains, can be clearly observed [556]. It is interesting to note the significant discrepancy between the prediction by the EDNA pipeline and the final quality of the data collected. Crystal F1 diffracted to 1.59 Å, a significant improvement over the 2.5 Å predicted. As previously discussed, EDNA attempts to predict loss of data quality due to radiation damage which, in some cases, leads to a very conservative estimation of the final resolution.

The number of reflections (spots) depends on the size of the unit cell (the bigger the cell, the higher the number of reflections) and on the resolution (the higher the resolution, the more reflections observed in a diffraction experiment). When the redundant individual measurements of reflections are combined, it is possible to obtain the number of unique reflections. The division of the number of reflections by the number of unique reflections is called redundancy. The higher the symmetry of a crystal, the higher the number of repeated reflections and the lower the redundancy. Values of redundancy above 3.0 are generally considered acceptable, although the TiHydG data does not achieve this as a consequence of the low symmetry P1 space group [553].

$R_{merge}$ is a measurement of the agreement between multiple redundant measurements of the same reflections with lower values usually meaning better data quality. Since $R_{merge}$ measures the difference between the same reflections collected at different time points, the redundancy value is related to the $R_{merge}$, with redundancy showing the number of times the reflection has been collected. A higher redundancy can sometimes denote a higher $R_{merge}$, though it will usually be accompanied by a more precise averaged intensity of the spots/reflections [553]. Crystal damage during data collection will lead to a higher $R_{merge}$.

$I/\sigma I$ is the measurement of the signal-to-noise ratio, calculated by average intensity of the reflections divided by the average standard deviation of the intensities of the same reflections. The higher this value, the more precise is the measurement of the intensity of the data [553].

The completeness of the data set relates to how many unique reflections have been collected. Since most data are ultimately used in the calculation of electron density or Patterson maps via Fourier transforms, the completeness of data is of vital importance. The intensity of a single reflection is a factor to a single term in the Fourier summation. If the data is not collected, the missing data cannot contribute to the summation (i.e. the term is implicitly zero) [553]. The

completeness of the anomalous data sets was particularly low in all datasets collected, a consequence of compounding the difficulties in collecting data on a tricyclic (P1) system with the loss of symmetry arising from the breakdown of Friedel's law. Diamond light source I04 beamline now possesses a Mini Kappa goniometer head, which allows the crystal position to be rotated, aligning a symmetry axis of the crystal with the spindle axis, resulting in more a more accurate estimation of anomalous differences, but it was not used to test the crystals discussed in this project.

Twinning is a phenomenon whereby two parts of the crystal have distinct orientations and their reciprocal lattices overlap significantly. The most common form of twinning is twinning by merohedry, where the two diffraction patterns from different crystal orientations overlap extensively in diffraction space. In this case, the recorded intensities are a mixture (sum) of intensities from all the contributing lattices, and the overlapping reflections do not have the same Miller index (see section 8.3). The data obtained is effectively a mixture of intensities that are formally unrelated to each other, and, consequently, crystals with a high lattice overlap cannot be used for structure resolution [553].

### 2.6.2 Model Building and Structural Refinement

As no structure of a HydG was available to solve the phase problem with molecular replacement [557], SAD was used near the absorption edge of iron [550], to obtain the information relative to the iron present in the [4Fe-4S] clusters. AutoSol from *Phenix* program suite [558, 559] was used to automatically solve the phase problem and determine the initial positions of iron atoms in the structure. The program is also able to predict the position of several residues of the protein (Figure 50), but the model building ability is quite limited. It does however seem to show a TIM barrel-like formation, with two clusters positioned close to the modelled residues (shown in green, Figure 50). The iron anomalous signal repeats itself beyond the defined protein structure, (Figure 50, $[5Fe\text{-}XS]_{sym}$) and are likely to arise due to symmetry elements of the crystal.

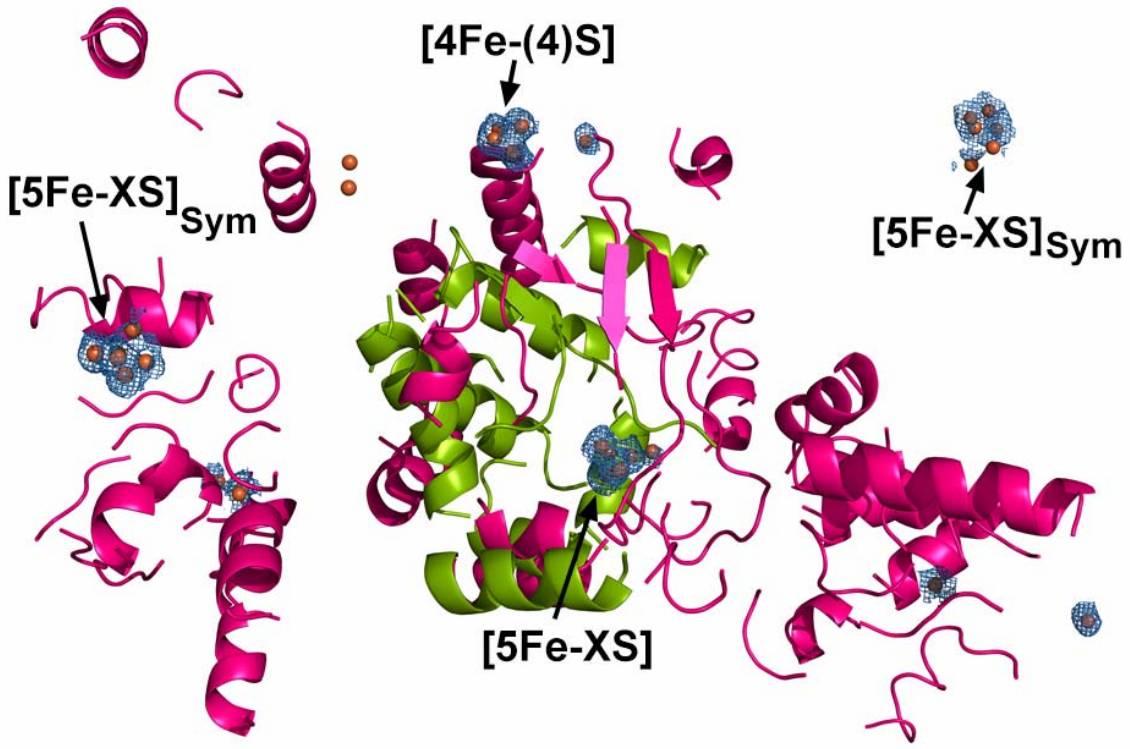



**Figure 50 Partial structure solution obtained by AutoSol.**

The anomalous data set of *Ti*HydG F1 crystals was used to identify the position of the iron atoms (red spheres, and anomalous density in blue, cutoff $\sigma$= 3.0), as well as create an initial partial model of the protein in green. In pink some $\alpha$-helices elements, built entirely on alanine and glycine are shown, as part of the structure building process. The iron signals are observed clustered together in a cubane-like structure, likely to be indicative of the [4Fe-4S] clusters of the protein. Several of these cubane elements show an additional iron signal.

The program AutoBuild from Phenix program suite [560] uses the phase information obtained in AutoSol, together with high resolution reflection information collected at higher energy (standard data) to improve the initial 118 residue model of the protein. The model calculated by AutoBuild is, as can be seen in Figure 51, a significantly more complete structure, with two monomers built (477 residues built out of 865 residues present in two monomers), attesting to the ease of the solution of structures with the available program suites. Interestingly, the iron placement in AutoSol showed several clusters containing an additional iron near the cubane structure (Figure 50), while the number of irons in AutoBuild seems much more conservative, with one of the models presenting two potential [5Fe] cubane clusters, while the second monomer showing just two independent iron atoms. The iron content and overall structure of the clusters will be discussed at length in section 2.7.3.

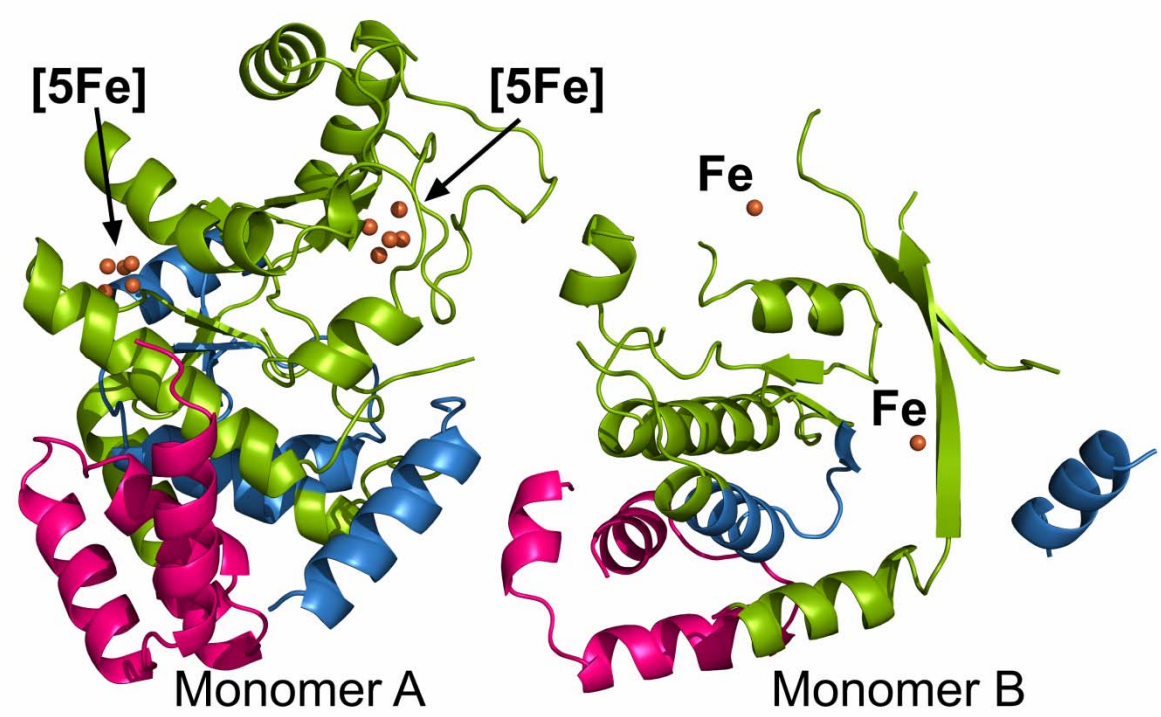



**Figure 51 Structural model of *Ti*HydG as the result of automated model building and refinement with Autobuild.**

Two monomers are present, with a significant asymmetry in the building process in each, with 310 residues in monomer A and 167 in monomer B. N-terminus is shown in pink, RSAM fold in green and C-terminus in blue.

Manual manipulation of the models was carried out in WinCoot [561], allowing the addition of ligands and residues for a better fit of the structure into the density. The successive manual modifications were followed by an automatic refinement cycle in Phenix Refine [559], which optimises rotamers, bond lengths and angles, torsion angles, as well as automatically adding water molecules and potentially ligands.

Several cycles of manual and automatic refinement were completed until all electron density was populated with structural information, and no more improvements could be achieved. The progress of the refinement cycles was followed by calculation of the R-work and R-free values, as a means of validation of the process ( see section 8.3.2.5). For high resolution structures a R-work value below 0.2, and an R-free value not too distant from the R-work (a difference of 0.05 units or less) are usually indicative of a successful structure assignment [562, 563].

**Table 8 Refinement statistics for the crystallographic dataset of *Ti*HydG, following refinement.**

| | **F1 Hampton PEG Rx** |
|---|---|
| $R_{work}$ / $R_{free}$ | 0.18/0.22 |
| No. atoms: | |
| Chain A / Chain B | 7331/7247 |
| Ligand/ion | 48/63 |
| Water | 815 |
| *B-factor*s: | |
| Chain A / Chain B | 32.0/28.4 |
| Ligand/ion | 28.5/22.2 |
| Water | 37.8 |
| R.m.s deviations: | |
| Bond lengths (Å) | 0.027 |
| Bond angles (°) | 2.083 |

Within the *Phenix* software, there is a simple but powerful validation tool, Polygon [564], that directly compares the quality of the structural data, with all the structures in the Protein DataBank (PDB) of similar resolution. It reveals, rather intuitively, in an one-dimensional histogram, much of the same data as Table 8, making it easy to identify any weaknesses the model may have from overfitting (worse than average $R_{free}/R_{work}$ values) to unusual conformations in the residues (RMSD of bonds and angles).

Following refinement of the *Ti*HydG structure, POLYGON analysis shows the B-factor in the structure is much higher than average. A high B-factor is usually indicative of disorder in the structure translating into an increased coordinate error. However, a protein with many flexible

regions may have higher than average B-factor and still have an accurate model of its structure. A discussion of flexible and disordered regions that may be influencing this value is presented in section 2.7.2.

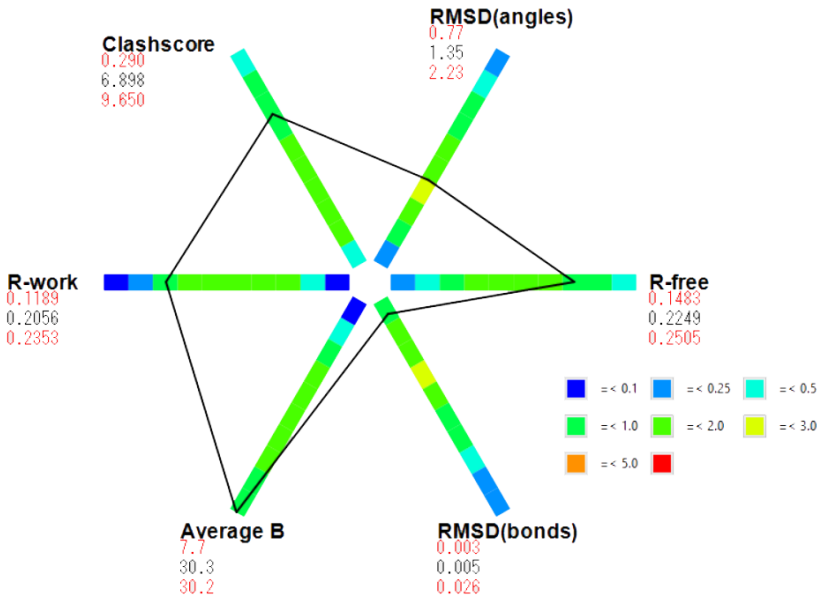



**Figure 52 Polygon analysis of *Ti*HydG F1 crystal structure.**

All values apart from B-factor fall within the average for structures close to 1.6 Å resolution. Each histogram represents a weighted average of structures with similar resolution range.

## 2.7 HydG Structural Analysis

### 2.7.1 Overall Fold

The models obtained from the four crystals grown (section 2.5) reveal the same slightly atypical TIM Barrel formation, as the C-terminus provides two extra β/α secondary structures, to form a full TIM Barrel. In all structures the same homodimerization of the protein is observed, with a 123° rotation between monomers (Figure 53). Several interactions (Glu12-Arg453; Gln16-Phe466; Asn10-Tyr465; Glu225-Lys183; Glu258-Arg392) bind both monomers together, with the dimerization occurring in such a way that the C-terminus of one of the monomers is very close to the second monomer. This would make the binding of an acceptor protein such as HydF and transfer of CO and $^{-}$CN [291, 296] very difficult within that monomer, strongly supporting the idea that the dimerization is most likely related to the crystallization process and not the active form of the enzyme. Analysis using the web service PISA confirmed that no specific interactions exist

between both monomers and the structures probably should not form a complex in solution [289], similar to the previous reports that analytical gel filtration chromatography eluted HydG with a calculated molecular weight of 53 kDa [80, 519], the monomeric form of the protein.

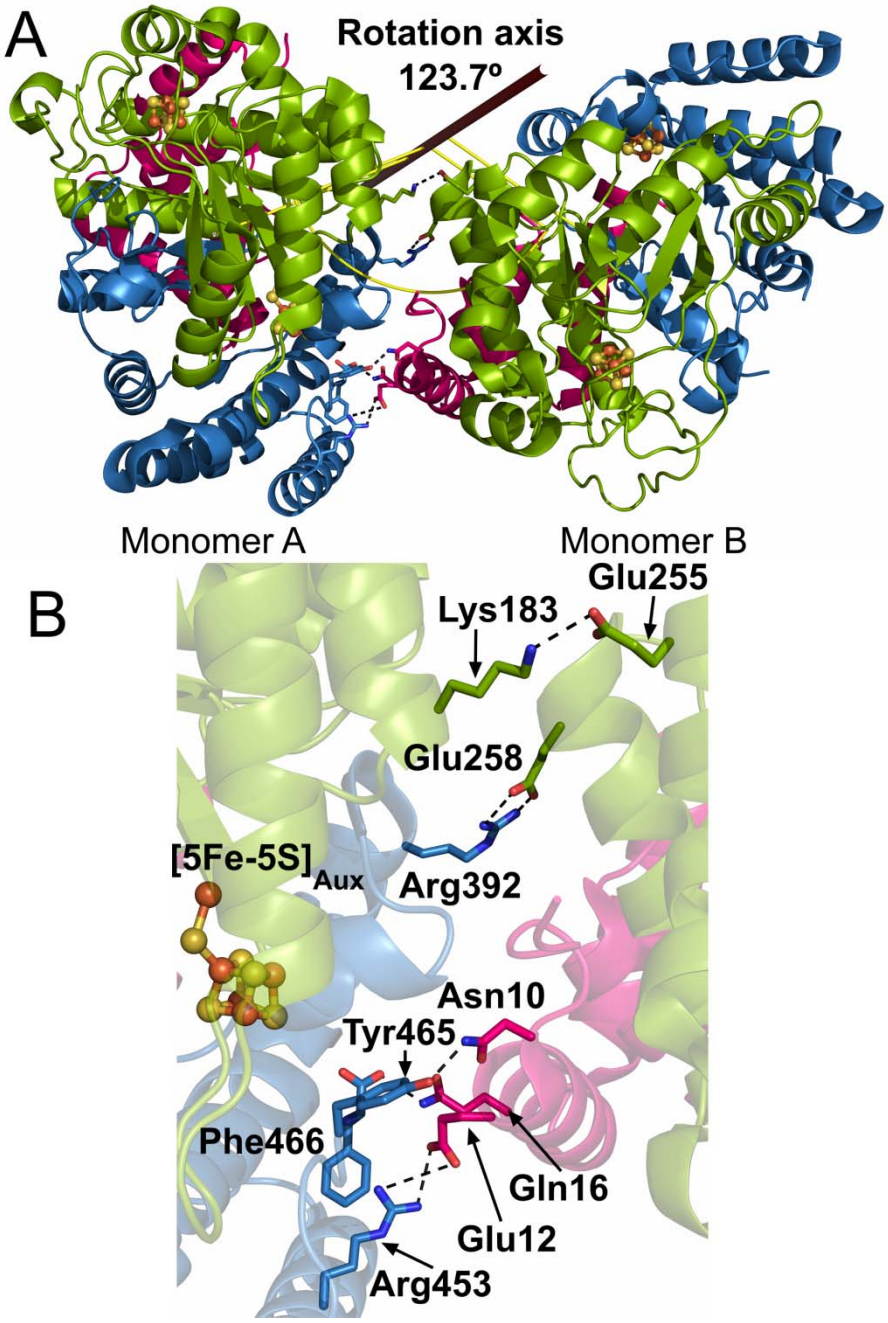



**Figure 53 Dimeric structure of HydG from *Thermoanerobacter italicus*.**

The partial TIM barrel, common to the RSAM superfamily is shown in green. The N-terminus is shown in magenta, and the C-terminus in blue. Each monomer possesses both clusters present, shown in yellow and orange. The TIM barrel forms a tunnel on the bottom of which the auxiliary [4Fe-4S] cluster is found, bound to the C-terminus. **A:** The rotation axis between the two monomers is shown, with the angle of rotation in yellow. Rotation axis calculated using draw_rotation_axis.py script, designed by Dr. Pablo Calvo [565]. **B:** Inter-monomer interactions.

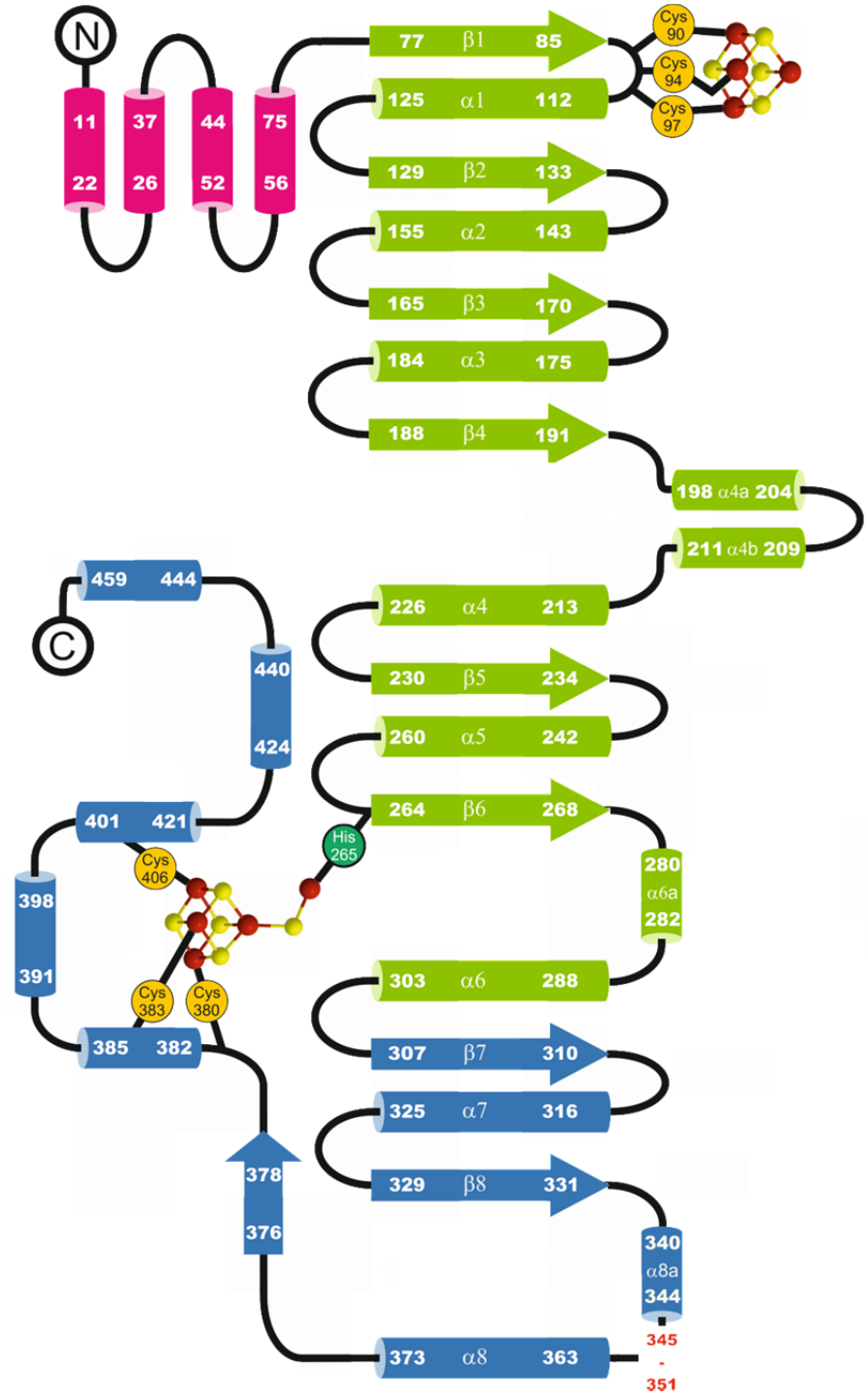



**Figure 54 2-dimensional topology of HydG, revealing the secondary structure.**

The N-terminus is composed of 5 α-helices (purple); the RSAM domain (green) shows the cysteine-containing loop following the first β-strand, together with three extra alpha helices, a pair following β4 and additional α-helix after β6. The C-terminus is composed of two α-helix/β-strand pairs to form the full TIM barrel. Before the last α-helix of the TIM barrel, a disordered loop can be observed, with a small α-helix being formed prior to the mobile residues. The C-terminus still possesses an additional β-strand, and 5 α-helices, three of them responsible for binding the auxiliary cluster.

The SAM binding cluster is, as previously discussed, highly conserved in RSAM enzymes (section 1.1.2, Figure 5), being bound by a loop that stems from the first β-strand. The characteristic CxxxCxϕC is found in residues 90-97 and Tyr96 is the conserved aromatic residue which interacts with SAM (Figure 55, magenta). The structural motifs that have previously been identified in RSAM and reviewed in detail [22, 149] are also observed in this structure (and highlighted in Figure 55).

The main chain of the first residue in the GGE motif (actually an AGE sequence, residues 134-136 in *Ti*HydG) interacts with the amino group of SAM, in a manner not dissimilar to that proposed for the more common glycine [22, 150] (Figure 55, salmon). The sequentially variable GxIxGxxE motif (in *Ti*HydG, GVLYGLYD, residues 234-241) shows a significant amount of sequence similarity to BioB where the motif was firstly identified [22, 150, 180]. In *Ti*HydG's structure the hydrophobic interactions of the motif's conserved isoleucine have been replaced by those of leucine 236, while the motif's glutamic acid is replaced by Asp241 (Figure 55, tan). Glutamic acid 194, present in the loop following β4 strand in the structure of TiHydG, is the residue which interacts with the hydroxyl groups of SAM's ribose, in a motif which has a reportedly high degree of variability [22, 150] (Figure 55, light pink). Lastly, the β6 motif (PRLR, residues 270-273 in *Ti*HydG) provides interactions to the adenine base of SAM (Figure 55, teal).

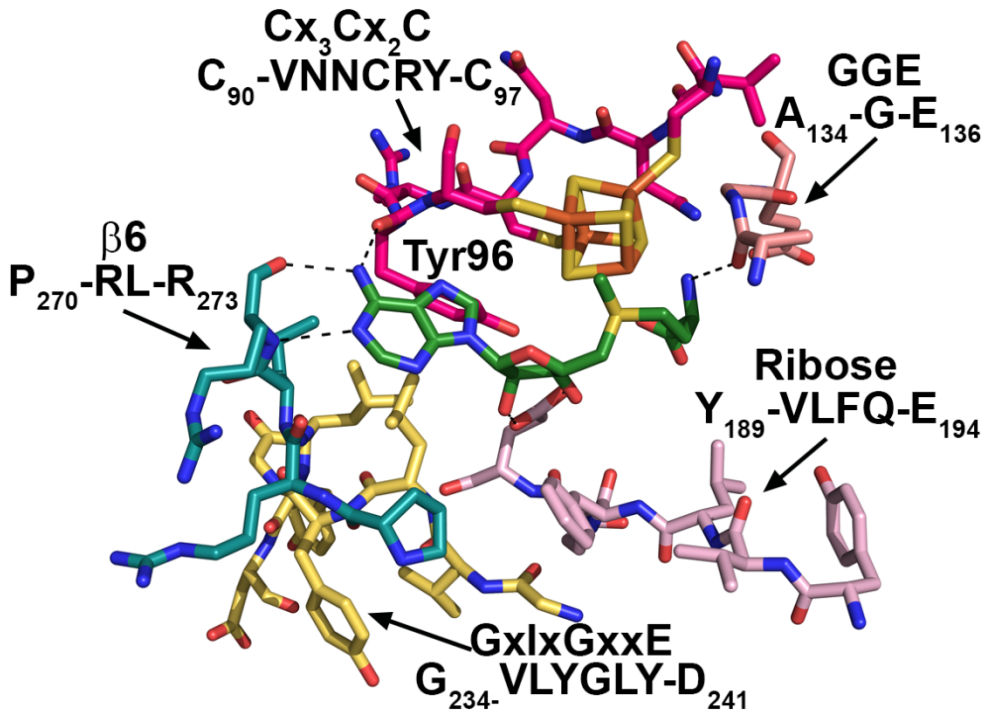



**Figure 55 Conserved SAM binding motifs of the RAM superfamily, as present in *Ti*HydG.**

Magenta: CX3CXϕC-containing loop; Salmon: "GGE" motif; Light pink: ribose motif; Tan: "GxIxGxxE" motif; Teal: β6 structural motif; Green: SAM. Adapted from [149].

### 2.7.2 Substrate and Co-Substrate Access Sites

Comparison of the two monomers A and B shows the presence of a cluster-bound SAM in monomer B, while the co-substrate seems absent in monomer A. Near the $[4Fe\text{-}4S]_{RSAM}$ cluster of monomer A, only density corresponding to methionine can be seen coordinating the cluster, with no electron density available for the adenosyl moiety (Figure 56). This may be explained if turnover of the enzyme has occurred, either in solution or in crystal, with the methionine binding to the active site and the 5'-deoxyadenosine being either disordered or in the bulk medium.

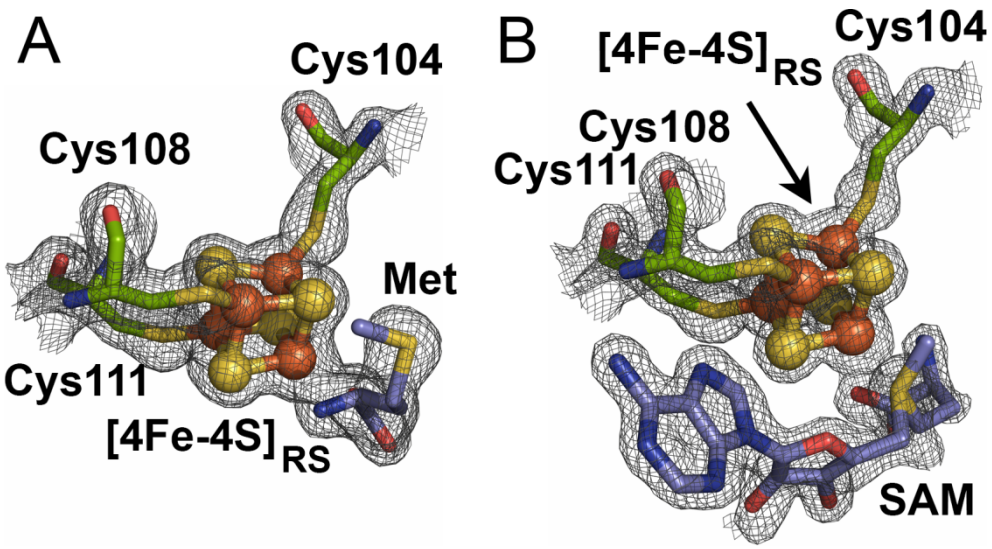



**Figure 56 *Ti*HydG conserved [4Fe-4S] radical SAM cluster and ligands.**

Residues from the N-terminal domain are shown in pink, from the radical SAM core are shown in green and residues from the C-terminal domain are shown in blue. **A:** The $[4Fe\text{-}4S]_{RS}$ cluster from monomer A, methionine, from Sam degradation, bound. **B:** The $[4Fe\text{-}4S]_{Aux}$ cluster from monomer B, SAM bound. Structure obtained with Hampton PEGRx HT F1. This 2Fo-Fc map is contoured at 1.5 σ around the clusters and ligands.

Additional differences can be seen close to the RSAM active site, in particular within a rather disordered loop, which forms part of the C-terminal domain and is found on the top of the active site (in the orientation shown in Figure 57 A). In particular, Tyr 341 has a very different spatial location in both monomers. When SAM is present, it intrudes near the methionine of the co-substrate (Figure 57 A, orange) but, in the absence of SAM the residue moves 14 Å away from the co-substrate, and the vacated space is occupied by Gln354 (Figure 57 A, blue).

It seems likely the movements in the loop correspond to an entrance point of either the substrate or SAM, the two actors in the mechanism associated with the SAM cluster. The possibility of SAM-

entrance seems less likely than the alternative, as the loop belongs to the non-RSAM conserved C-terminus of the enzyme. The conserved GGE motif, together with residues 169-172 belonging to β3 of the TIM barrel are mostly responsible for the co-substrate tunnel leading into the active site (Figure 57 B) [149]. This leaves the role of the loop as the entrance location of L-tyrosine, which may explain its disordered state, as the substrate is not present in either monomer.

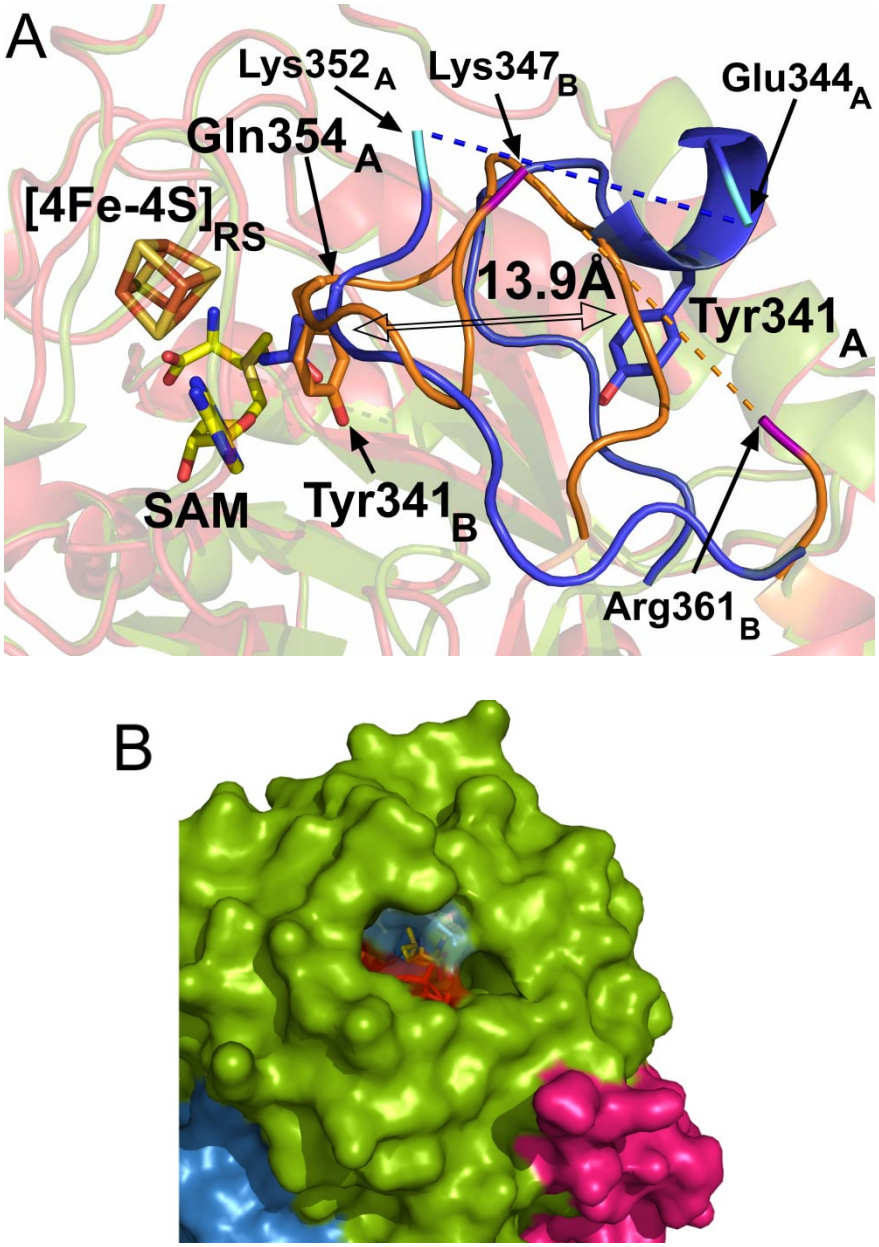



**Figure 57 Potential substrates' entrance sites.**

**A:** Disordered loop in *Ti*HydG structure. Monomer A's overall fold is shown in green, with the partial disordered loop in blue. Residues 344-352 were not built in the structure due to a lack of appropriate density. The open conformation in the absence of SAM is most easily understood with the position of Tyr341, which in monomer A is part of a small α-helical structure, while Gln354, is located near the SAM binding site.

Monomer B's overall fold is shown in red, with the partial disordered loop in orange. Residues 347-361 are absent in the structure. Tyr341, when SAM is present, is located in the Gln354 site of monomer A, representing a 13.9 Å change in position of Tyr354 towards the active site when SAM is present. **B:** Surface view of proposed location for SAM access to the active site. Residues AGE (134-136) in light blue, residues NIAAT (169-173) in red; SAM in yellow.

### 2.7.3 Auxiliary Cluster

At the bottom of the tunnel running the length of the TIM barrel, an auxiliary [4Fe-4S] cluster can be found, enclosed and isolated from the external environment by the C-terminus of the enzyme. Much like the SAM binding cluster, it is bound by three cysteines ($CX_2CX_{22}C$, residues 380-406).

The distance between clusters is unprecedentedly larger than any other reported structure of an auxiliary cluster containing RSAM enzyme, with a 25 Å separation between the two cubane units, significantly higher than: RimO, 7.3 Å [37]; LipA, 12.3 Å [182]; BtrN, 15.8 Å [566]; anSME and MoaA, 16.4 Å [178, 179]. The C-terminus of the enzyme creates a secondary pocket which can putatively function as an auxiliary active site and consideration of the activity shown by the ΔCTD mutant suggest this would be for the formation of the inorganic ligand(s) [296].

Of the four datasets collected, only the highest resolution F1 shows a divergence on the structural nature of the auxiliary cluster. In this structure the unique iron of the auxiliary cubane core shows coordination to a non-protein ligand that was unanticipated and chemically unidentified.

When the iron anomalous data was analysed for monomer A in this particular structure, a strong fifth iron signal was visible in the auxiliary cluster, as had been initially predicted in the AutoSol stage while using the iron κ-edge anomalous data to solve the phase problem (see section 2.6.2, Figure 50). Additional building and refinement of the structure confirmed the presence of a [5Fe-5S] auxiliary cluster (density map in Figure 59 C), composed of a typical [4Fe-4S] cubane cluster, a bridging (μ2) sulfide and the additional iron (Figure 58). This iron was refined to occupancy lower than unity, at 0.73, so the additional iron is present in only 73 out of 100 unit cells within the crystal. Due to the apparent ease by which this iron appeared to be removed from its location led it to be termed the 'labile iron', a terminology that also reflects its proposed function in HydG turnover, as described more fully in section 2.11.

The labile iron is hexacoordinate, showing an approximately octahedral geometry, binding the aforementioned μ2 sulfide, a single protein residue (His265) positioned *trans* to the sulfide ion,

two water molecules and bidentate non-protein amino acid residue. There is an asymmetry in the distances between the bridging μ2 sulfide and both irons it binds. The bond length towards the unique iron of the cubane [4Fe-4S] is 2.2(5) Å, a value well within other measured lengths of Fe-S clusters [87, 301]. The bond length between the μ2 sulfide and the labile iron is higher at 2.5 Å, a long, but not unprecedented interatomic distance [567].

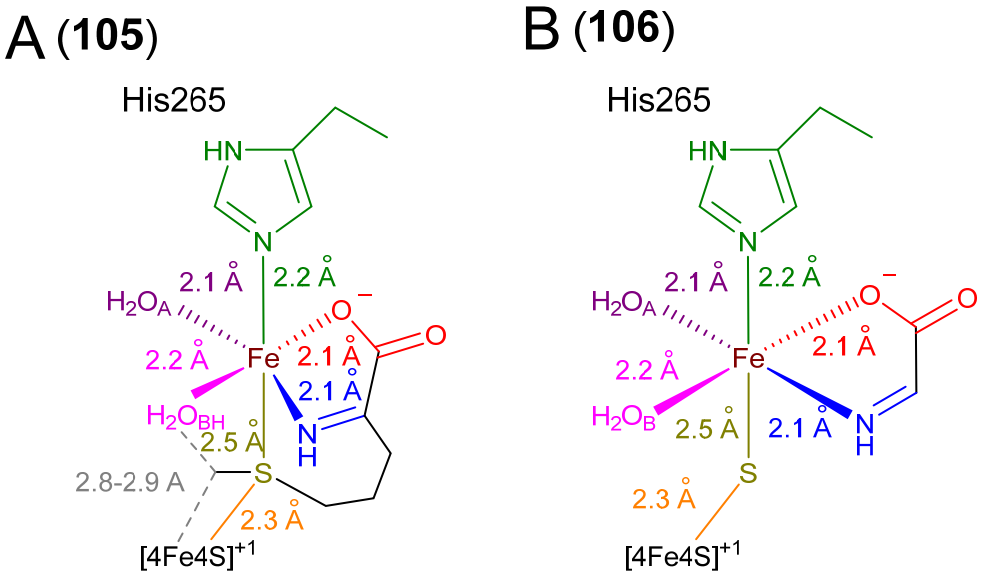



**Figure 58 Octahedral coordination of the labile iron as observed in the crystal structure of *Ti*HydG.**

Octahedral coordination of the labile iron as observed in the crystal structure of *Ti*HydG. Inter-atom distances measured in the structure are shown. **A:** Potential methionine binding model **105**. **B:** Alanine binding model **106**.

The free amino acid shown to coordinate the labile iron is likely to be methionine, derived from SAM cleavage during crystallisation. An initial model was built, where the bridging sulfide would arise from the sulfur in the methyl tioether group of methionine (Figure 58 A). However, the side chain does not show enough well resolved density to permit accurate refinement, possibly derived from the flexibility of the side chain. An alanine residue was modelled at the site instead (Figure 58 B and Figure 59 A/C). The fortuitous presence of this bound amino acid has contributed to development of the proposed mechanism for CO/CN formation from DHG, which could potentially coordinate the labile iron in a similar bidentate manner (see section 2.11).

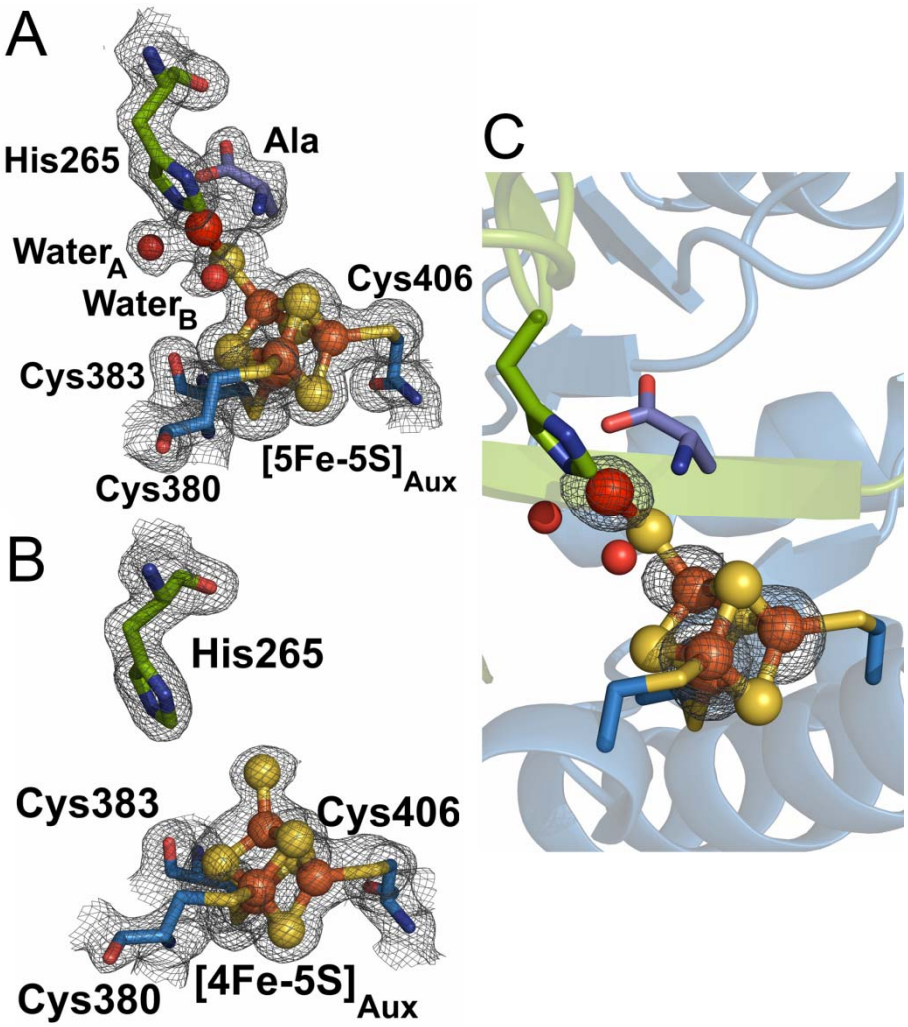



**Figure 59 *Ti*HydG auxiliary cluster and ligands.**

Residues from the radical SAM core are shown in green, and residues from the C-terminal domain are shown in blue. The labile iron is shown in bright red. **A:** The $[5Fe\text{-}5S]_{Aux}$ cluster from monomer A, structure obtained with PEG Rx F1, showing the coordinating water molecules, amino acid ligand, and His265. **B:** The $[4Fe\text{-}5S]_{Aux}$ cluster. The 2Fo-Fc map is contoured at 1.0 σ cutoff around the cluster and ligands. **C:** The $[5Fe\text{-}5S]_{Aux}$ cluster anomalous data, with iron atoms located using the anomalous difference Fourier map shown in grey and contoured in *A–D* at 5.0σ cutoff.

### 2.7.4 Putative Interaction Site With HydF

When looking into the topology of HydG, it was tempting to see if the structure of the helix 391-398, which connects two auxiliary cluster helices (see Figure 54), could be interpreted as to give clues about protein:protein interactions with other maturases. This helix is located towards the external solvent surrounding the protein, making it well positioned to undergo a conformational

change which would easily reveal the auxiliary cluster of HydG, in order to transfer the potential HydG product, a $Fe(CO)_2CN$ synthon to the 2Fe-subcluster template domain of HydF.

When comparing the two monomers we observe a significant change in the Debeye-Waller factor, or B-factor, of this helix, calculated as:

$$\mathbf{B = 8\pi^2 \times u_i}$$

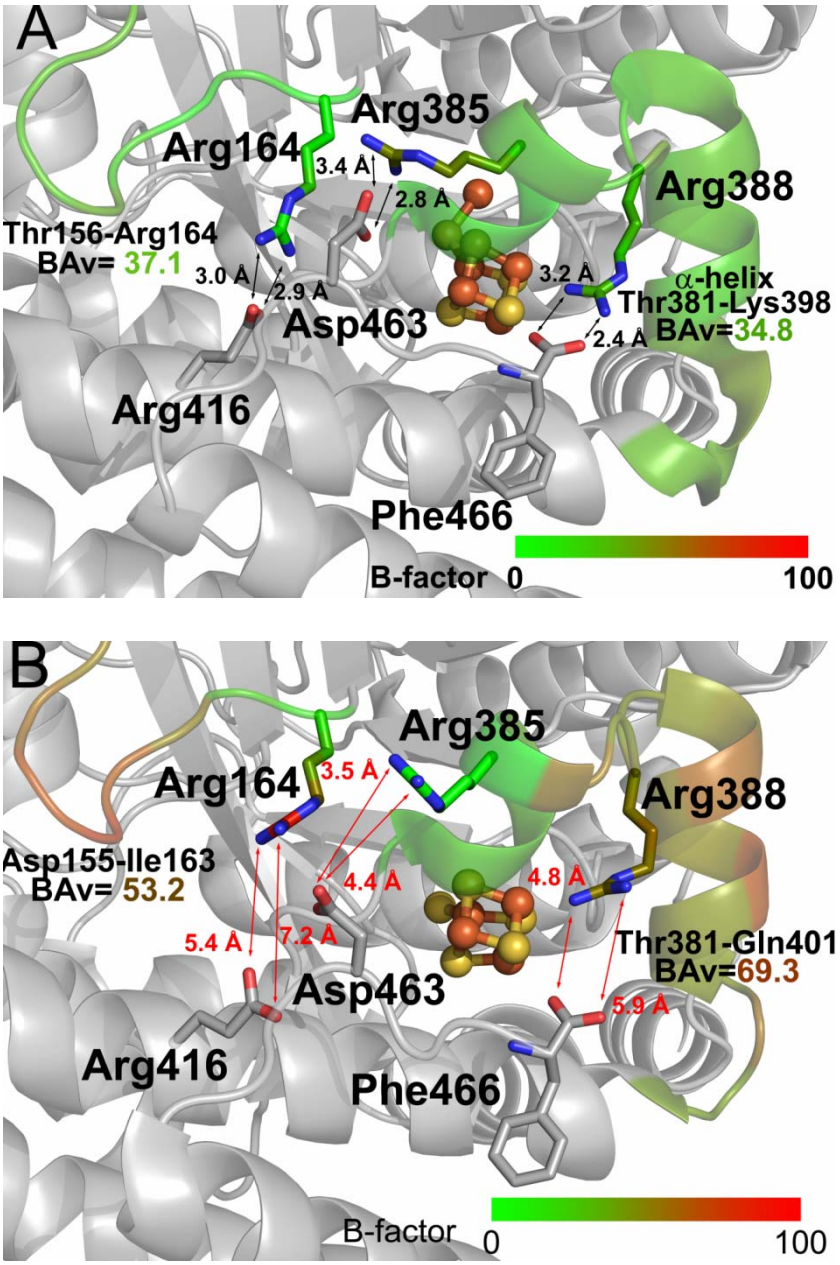
**Equation 2 Calculation of the Debye-Waller factor (B-factor).**

$u_i^2$ is the mean square displacement of atom *i*, and the B-factor is measured in $Å^2$.

Atoms which show low B-factors in the structure are usually well ordered and stationary, while higher B-factors can be interpreted as more flexible positions in the protein [568].

When the labile iron is present, in monomer A, the helix seems to have much lower B factors (Figure 60 A), suggesting a better ordered and less flexible helix; in contrast, the higher B factors in monomer B (Figure 60 B) may indicate a higher degree of flexibility. Additionally, the position of the auxiliary clusters in each monomer do not perfectly overlap with each other, due to a small displacement of the auxillary cubanes (1.9 Å measured between bridging μ2 sulfides of each monomer) (Figure 60 C).

When analysing the network of interactions that provide structure to the C-terminal domain in monomer A, three ionic interactions stand out: Arg399 of the solvent exposed helix interacting with Asp477, Arg402 with the C-terminal carboxylate in Phe480 and Glu430 with Arg178 (Figure 60 A). In monomer B this network is fully disrupted (Figure 60 B), contributing to an to an overall increased B-factor in α-helix 391-398, which could move as a gateway permitting access for a downstream partner enzyme such as HydF. It is interesting to note that none of the amino acids which participate in binding the labile iron required interactions are actually present in this α -helix, but instead in the preceding and subsequent α -helices.


A
Arg385
Arg164
3.4 Å
2.8 Å
Arg388
Thr156-Arg164
BAv= 37.1
3.0 Å
2.9 Å
3.2 Å
α-helix
Thr381-Lys398
BAv=34.8
Asp463
2.4 Å
Arg416
Phe466
B-factor 0
100
B
Arg385
Arg164
3.5 Å
Arg388
Asp155-Ile163
BAv= 53.2
4.4 Å
4.8 Å
5.4 Å
7.2 Å
Thr381-Gln401
BAv=69.3
Asp463
5.9 Å
Arg416
Phe466
B-factor 0
100

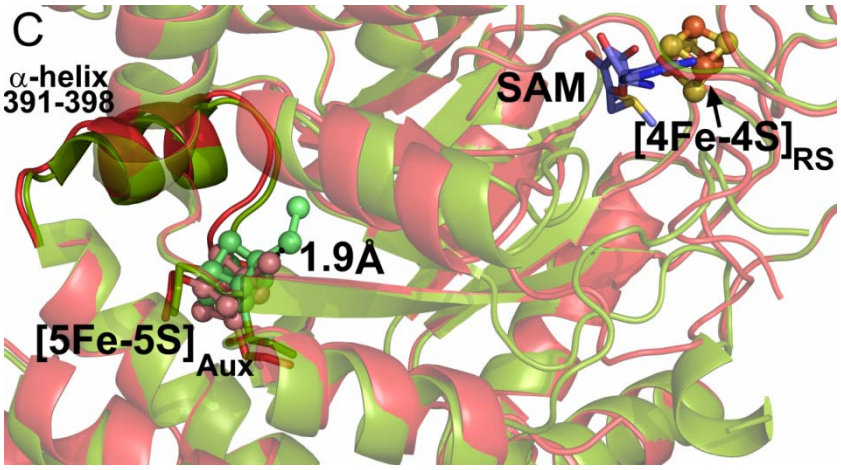



**Figure 60 B-factor changes in α-helix 391-398 near *Ti*HydG auxiliary cluster, caused by the change in interactions between three residue partners.**

B-factor changes in α-helix 405-412 near *Ti*HydG auxiliary cluster, caused by the change in interactions between three residue partners. **A:** Monomer A, with labile iron in position, where three interactions are present (Arg164-Glu416; Arg385-Asp463; Arg388-Phe466), decreasing the overall B-factor of Asp155-Ile163 loop and Thr381-Gln401 α-helix, the latter of which may be responsible for allowing the labile iron to be transferred to the scaffold maturase HydF. **B:** Monomer B, with labile iron absent and consequently the three interactions are no longer occurring, increasing the B-factor of Asp155-Ile163 loop and Thr381-Gln401 α-helix. **C:** Superposition of monomers A and B, showing an overview of the protein. Monomer A, together with its $[5Fe\text{-}5S]_{Aux}$ cluster are shown in green, while monomer B and the $[4Fe\text{-}5S]_{Aux}$ are presented in red. Unlike the B-factor values, the structural changes occurring in α-helix 391-398 are quite small. However, a 1.9 Å difference in the position of the bridging μ2 sulfide can be observed between monomer A and B.

### 2.7.5 L-Tyrosine Binding Site

Despite having obtained two crystals in the presence of the natural substrate L-tyrosine, no density could be attributed to this amino acid in any of the structures. Pocket analysis, using MetaPocket [569], indicated an extended active-site cavity stretching between the iron-sulfur clusters of HydG and running approximately down the axis of the TIM barrel, allowing the prediction of two potential binding locations, and connected by a funnel-shaped passage within the TIM barrel.

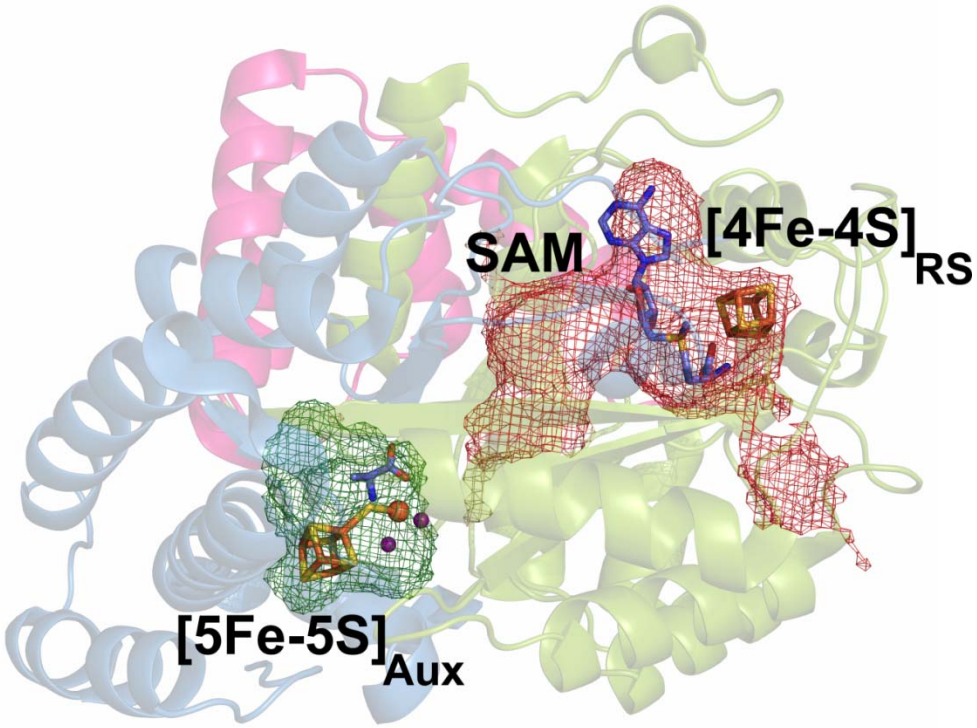



**Figure 61 Ligand binding sites for *Ti*HydG structure, identified by Metapocket.**

The first site (red) covers the RSAM active site, where the $[4Fe\text{-}4S]_{RS}$ cluster and SAM occupy a large portion of the available pocket, with an additional empty location directly below SAM near the TIM barrel formation. The second site (green) contains he $[5Fe\text{-}5S]_{AUX}$ cluster with the available ligands also shown, occupying a significant area of the pocket.

In principle, there are various putative sites into which tyrosine could be docked, but comparison of the two extreme possibilities (close to SAM or close to the auxillary cluster) are described below and allowed the development of a tentative tyrosine binding model.

The first possible tyrosine binding site is located near the SAM cluster (Figure 62). There is a pocket right beneath SAM where tyrosine could bind, being stabilized by hydrophobic residues such as Phe81, Pro83, Phe192, Pro270 and Phe355. Several polar residues can interact with tyrosine in this pocket stabilizing it in the correct orientation for hydrogen abstraction in the amino group of tyrosine, for example: Tyr85, Asn169, Ser268, Thr312, Ser331 and Ser334, with further interactions also arising from Glu133 and Arg313.

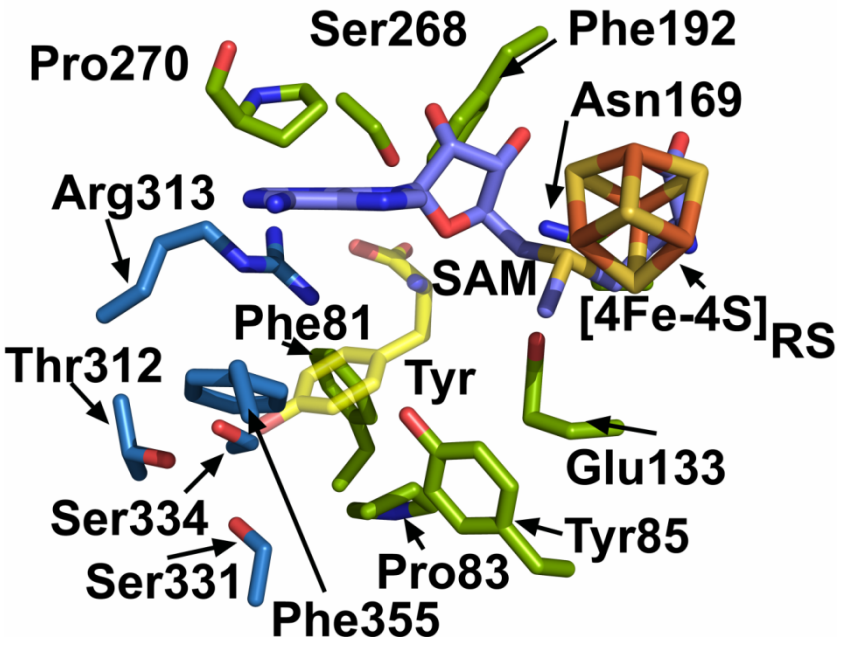



**Figure 62 Model of *Ti*HydG first putative substrate binding site.**

Residues from the RSAM core fold are shown in green and from N-terminus in blue. The $[4Fe\text{-}4S]_{RS}$ cluster is shown in yellow and orange, SAM in lilac and a potential tyrosine bound is shown in yellow.

Several structures of RSAM enzymes are available with substrate and SAM bound in the active site (79;88;92;321;322;323;324), which show a significant conservation of the substrate position [22, 149]. This is consistent with the conserved mechanism of action of RSAM enzymes, with the hydrogen abstraction location being well positioned to be targeted by $DOA^{\bullet}$ to initiate the reaction (see section 1.1.1). By overlaying known structures with the substrate present (Figure 63), one can predict a similar binding location for tyrosine near the SAM binding site.

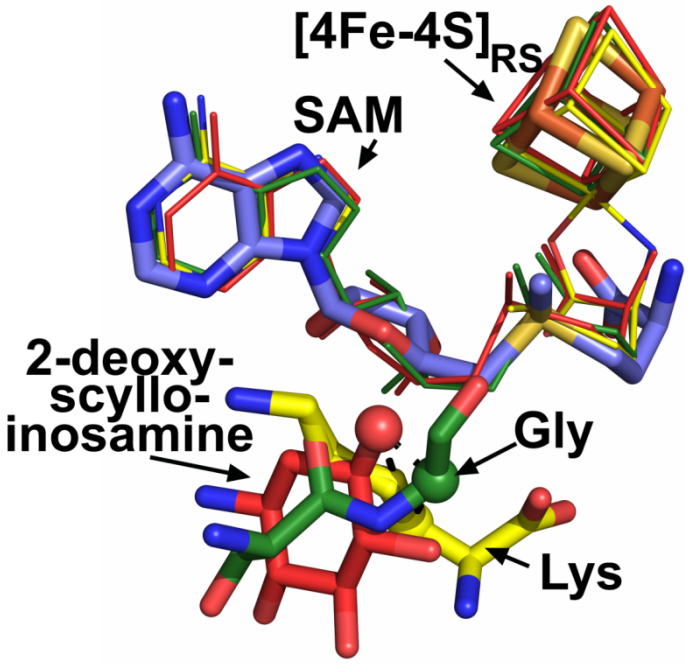



**Figure 63 Overlay of the[4Fe-4S]$_{RS}$ cluster with SAM and the substrate of different RSAM enzymes.**

The structure of BtrN is shown in red (PDBID: 4M7T); in yellow the structure of LAM (PDBID: 25AH) and in green the structure of PFL-AE (PDBID: 3CB8). The substrate atoms where the radical-driven abstraction is proposed to occur are shown as a sphere, and the location of the abstraction is tightly conserved across the three enzymes (less than 3 Å difference). [4Fe-4S]$_{RS}$ (yellow and orange) and SAM (lilac) of *Ti*HydG are shown in stick models, and in a conserved position with the three other structures.

The second possibility for L-tyrosine binding would involve directly-binding to the auxiliary cluster (Figure 64), as proposed previously by Kuchenreuther *et al.* [314]. Tyrosine binding could be anticipated to bind to the labile iron in a bidentate manner analogous to the non-protein amino acid observed in the *Ti*HydG structure (see section 2.7.3). The site is formed by three arginines (Arg77, Arg129 and Arg165), the labile iron coordinating His265, Val79 and Thr266, all arising from the partial TIM barrel; Gly307, Ser328, Gln329 and Phe393 from the C-terminus and interestingly one residue, Tyr74 from the N-terminus. The coordination of the residue would be accomplished mostly by the labile iron, as well as a potential interaction of the carboxylic acid moiety with Ser328. Only one interaction to the phenolic OH, by Arg77, can be predicted in this model.

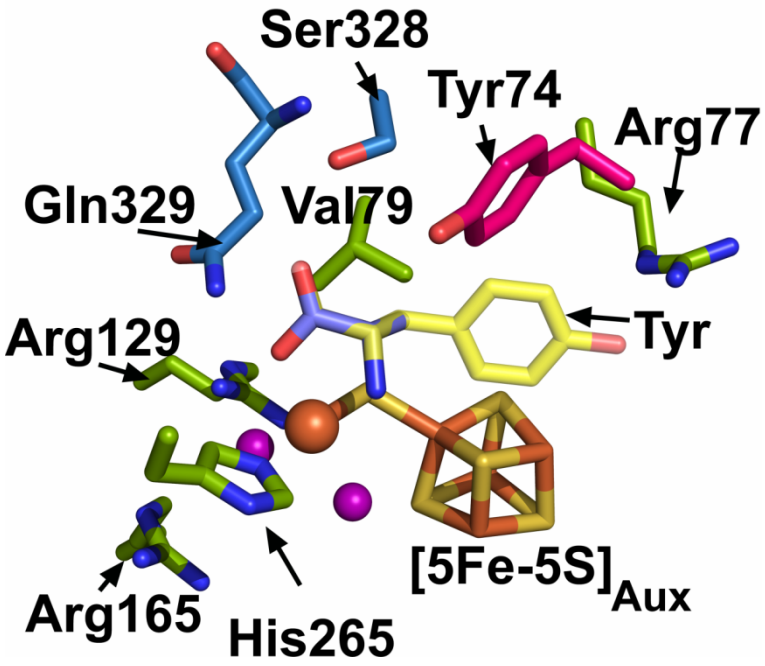



**Figure 64 Model of *Ti*HydG second putative substrate binding site.**

Residues from the RSAM core fold are shown in green and from N-terminus in blue. The [5Fe-5S]$_{Aux}$ cluster is shown in yellow and orange and the potential tyrosine bound is shown in yellow, with the main chain modelled after the alanine residue coordinating the labile iron.

The two [4Fe-4S] and their respective active sites are separated in the structure by the TIM barrel (Figure 65), which forms an internal cavity with a small 'free' volume presumed to contain disordered waters (Figure 61). The top of the tunnel has three charged or polar amino acids, namely Glu133, Asn 169, Ser268 and Ser331. The middle of the barrel is predominantly hydrophobic, with the following residues: Phe81, Ala131, Val190 and Ile309. The bottom of the barrel, near the C-terminal auxiliary cluster has the labile iron coordinating His265. Additionally, two arginines, Arg129 and Arg 165, are present, together with two other polar residues, Asp230 and Gln329, which may have roles in the formation of CO and cyanide.

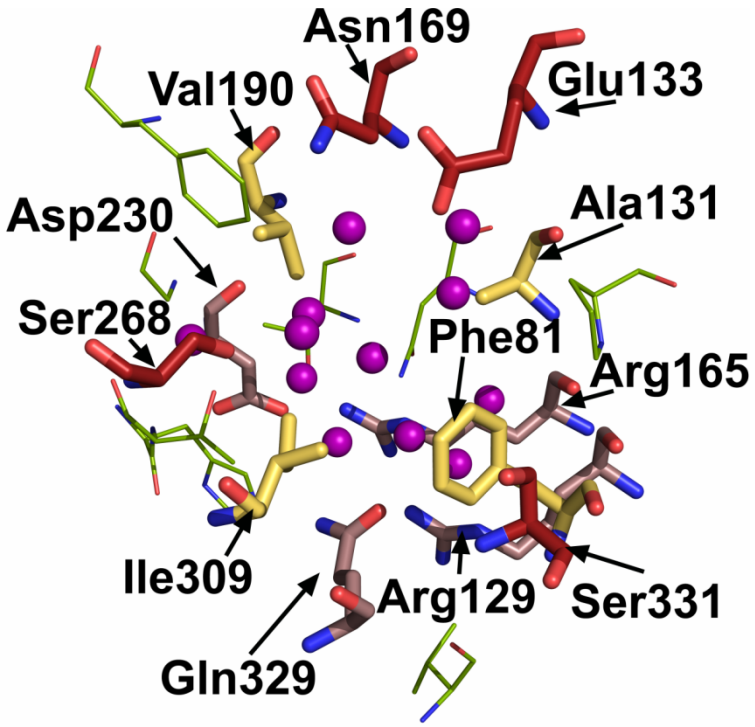



**Figure 65 Model of *Ti*HydG TIM barrel-derived tunnel.**

Residues arising from the conserved. Acidic/polar residues near the RSAM active site are shown in red, hydrophobic residues that form the centre of the tunnel are shown in pale yellow and charged/polar residues near the auxiliary cluster are shown in mauve. RSAM partial TIM barrel is shown in green, and C-terminus in blue. Several water molecules permeate the tunnel formation and are shown in purple.

For the mechanism of HydG to involve binding of tyrosine to the auxiliary cluster while the initiation of catalysis by radical formation occurs at the RSAM cluster, the radical would need to be propagated from the RSAM active site, through the internal cavity formed by the TIM barrel until it could abstract the hydrogen from tyrosine to drive forward the reaction. Although the propagation of a radical through such large distances (over 12 Å in the case of HydG) is not unheard of [570], the necessary network of residues which would allow such a transfer does not seem to be in place within the TIM barrel, in particular since the central region of the cavity is lined with mostly hydrophobic residues.

In contrast, if the breakdown of tyrosine occurs in the SAM cluster active site , the proposed cleavage product dehydroglycine [47, 293, 314], instead of the radical, would have to be moved through the tunnel to reach the second cluster, where it would then be presumably cleaved into $^{-}$CN and CO.

Due to the hydrophobic nature of the centre of the tunnel, it has been proposed that DHG, were it to travel through the channel, would be present in its neutral form [317]. However, the hydrophobic region is relatively small, and both the putative location of the formation of DHG as well as the end of the tunnel nearing the auxiliary cluster, possess charged residues, as described above. The presence of the acidic residue (Glu133) near the RSAM active site and basic residues (Arg129, Arg 165, His 265) near the auxiliary active site may form an electrostatic field responsible for directing the hydrolytically unstable DHG towards the labile iron of the auxiliary cluster. This field and the movement of DHG may explain why a DHG hydrogen abstraction from the 4-oxidobenzyl radical ($4OB^{\bullet}$), resulting in a tyrosine rearrangement similar to the mechanism of the highly related enzyme NosL [571], does not occur in HydG.

However, preliminary modelling carried out by Dr. Stephen Fox (Prof. Jonathan Essex's group), does not show a sufficiently strong electric field between each of the active sites of HydG to solely drive translation of dehydroglycine anion ($NH_2CHCOO^-$) through the tunnel, and further experimental evidence will be needed to fully understand how the movement of DHG occurs.

#### 2.7.5.1 Molecular Modelling of L-Tyrosine in its Putative Active Site

Structural information of the related enzyme NosL [321] (see section 1.2.2.3.2) confirmed the idea that the substrate in amino acid lyases is bound near the SAM binding site. To visualize a potential binding mode for tyrosine, a modelling study was carried out by Dr Stephen Fox, a member of Professor Jonathan Essex's group in the Computational Chemistry section of the University of Southampton.

The structure of NosL was used for superimposition, and L-tyrosine manually docked into the pocket is a pose analogous to that of tryptophan that that had been experimentally observed in NosL (PDBID: 4R33). The HydG:Tyr model complex was then energy minimized for both the ligand and neighbouring residues with Merck molecular force field MMFF94 [572] to a gradient of 0.1 kcal/mol/Å.

The resulting model places the amino group of tyrosine, the proposed target for abstraction [321],5.1 Å away from the 5'-carbon of SAM, a value slightly longer than those observed for the SAM to substrate distance in substrate-bound structures of other RSAM enzymes (3.7-4.1 Å) [151, 180, 573, 574].

The tyrosine amino-acid moiety is held in place by the side chain of Arg313, while the aromatic ring of Tyr is surrounded by the side chains of Phe81, Pro83 and Phe355. The phenolic O-H is positioned to make three hydrogen bonds to Ser331, Ser334 and Thr312 (Figure 66).

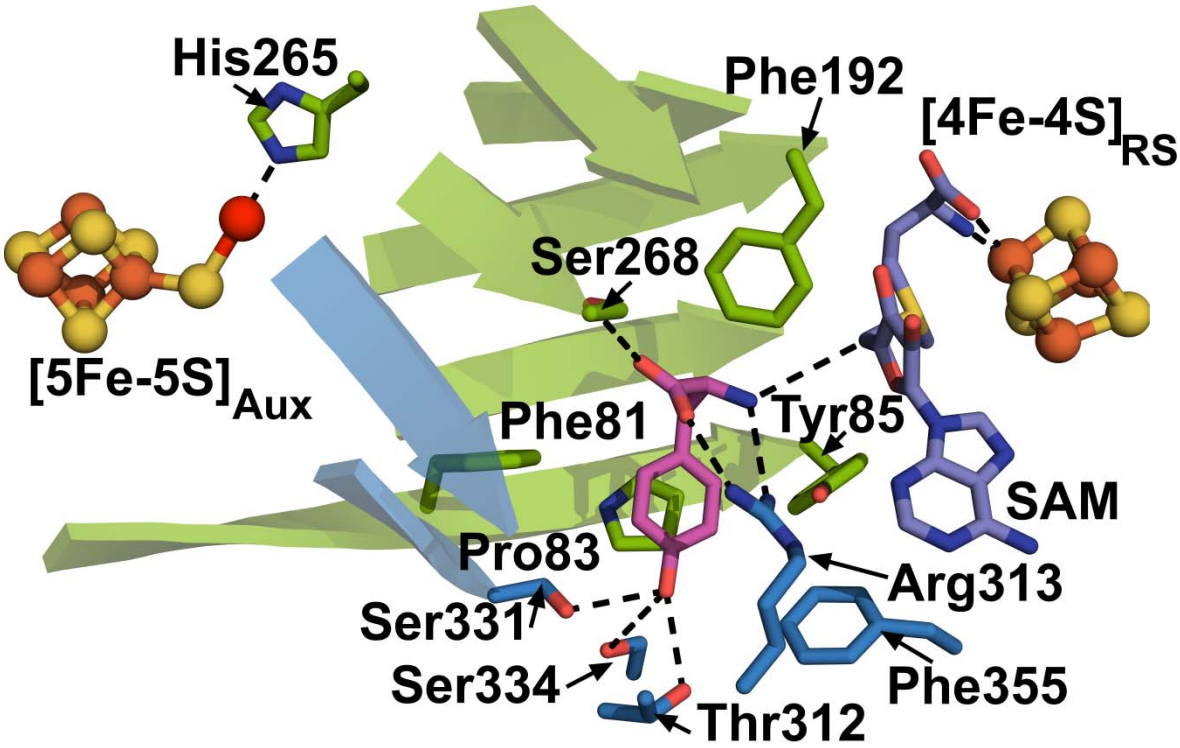



**Figure 66 Composite model of tyrosine and SAM binding to HydG.**

Residues from the radical SAM core are shown in green, and residues from the C-terminal domain are shown in blue. The labile iron is shown in bright red, tyrosine is in purple, and SAM is in purple.

## 2.8 Structural Comparison With Enzymes From the Amino Acid Lyase Family

### 2.8.1 NosL

Both HydG [302] and NosL [307, 321] are part of a sub-group of the RSAM family of enzymes, termed the amino acid lyases, which also comprise ThiH [313] and CofH [308] (for which there are no structures yet available). The close relationship between HydG and NosL, with a 28% sequence identity calculated with ClustalOmega [575] (Figure 67), translates into remarkably similar structures, with an a root mean square deviation of 6.34 Å. The root-mean-square deviation

(RMSD) [576] was calculated by the *align* package in Pymol [577], and is a measurement of the difference of the distances between the atoms of both structures:

$$RMSD = \frac{\sqrt{SUM(d_{ii})^2}}{N}$$

**Equation 3 Calculation of the Root-Mean Square Deviation.**

$d_{ii}$ is the distance between the $i^{th}$ atom of structure 1 and the $i^{th}$ atom of structure 2 and N is the number of atoms matched in each structure.

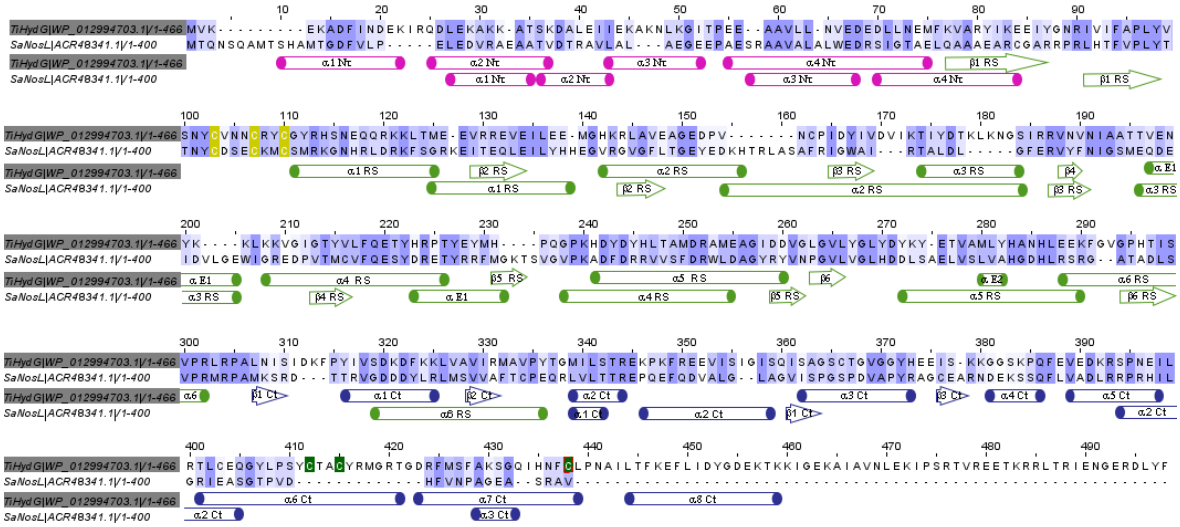
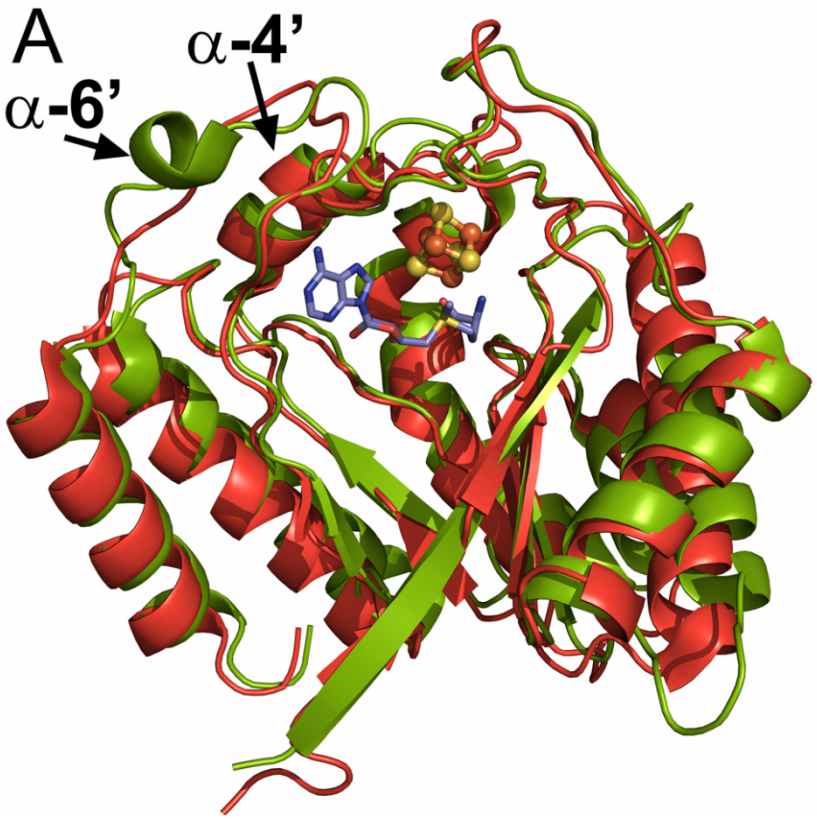


**Figure 67 Sequence alignment between *Ti*HydG and *Sa*NosL.**

Sequence is colour coded, with high conservation shown in darker blue and no conservation in white. The cysteine motif that binds the [4Fe-4S]RS cluster are shown in yellow and cysteines that bind the auxiliary cluster in *Ti*HydG in green. Secondary structures of both enzymes shown as cylinders (α-helix) or arrows (β-strand), and colour coded as follows: pink, N-terminus; green, RSAM conserved fold; blue, C-terminus. Alignment results obtained from ClustalOmega [575] and visualized using JalView. Gene UNIPROTID: D3T7F1 (*Ti*HydG) and C6FX51 (*Sa*NosL).

The similarities between both HydG and NosL structures start in the N-terminal region, which both consist of 4 alpha-helices (Figure 68 B). The general orientation of the helices are similar, but some spatial differences in the positioning of the secondary structure can be observed (RMSD $C_\alpha$ for N-terminal residues= 4.72 Å) The core 6 β/α partial TIM barrel fold, conserved across the RSAM family of enzymes, is less well conserved between HydG and NosL than the N-terminus

(RMSD for $C_a$ residues between the RSAM conserved fold= 6.90 Å). The differences observed between the two RSAM conserved secondary structures are predominant in several loops that surround the active site, together with two atypical additional alpha helices present in HydG's structure, NosL shares the one present after β4, but not the small alpha helix after β6 (Figure 68 A, green and red arrows, respectively). The initial secondary structures present in the C-terminus, which provide the two β strands and two α-helices necessary to close the TIM barrel, are well conserved between both enzymes. In the case of NosL, however, shortly following the closure of the barrel the C-terminal end of the protein is reached, and the large C-terminal extension of five alpha helices, which contains the auxiliary cluster in HydG, is absent (Figure 68 C). However the present structure in NosL is still remarkably well conserved in HydG, with the RMSD difference between the two C-terminal residues being calculated at 4.0 Å.



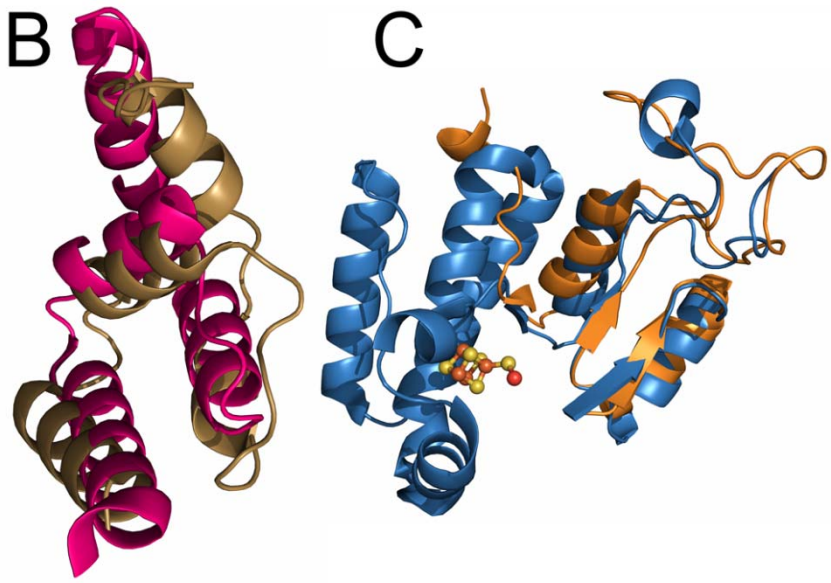



**Figure 68 Structural comparison between *Ti*HydG and *Sa*NosL**

**A:** TIM barrel of HydG is shown in green and NosL (PDBID: 4RR3) in red. **B:** HydG's N-terminus shown in magenta and NosL in brown. **C:** C-terminus from HydG shown in blue and NosL in orange. The auxiliary [5Fe-5S] cluster present in HydG I s absent NosL. The structures show a high degree of similarity both in the overall fold, as well as in the N-terminal extension, but vary significant in the C-terminal end, where *Ti*HydG contains a 5-fold helical extension that coordinates the auxiliary cluster, absent in *Sa*NosL. *Ti*HydG PDBID: 4WCX and *Sa*NosL PDBID: 4RR3.

Aspects of the mechanism of reaction of NosL are shared with ThiH and HydG, in particular the cleavage of the Cα-Cβ bond of an amino acid (tryptophan in the case of NosL, tyrosine for ThiH and HydG), leading to the formation of either dehydroglycine or a glycyl radical, which is used in the subsequent reactions [307] (Figure 69).

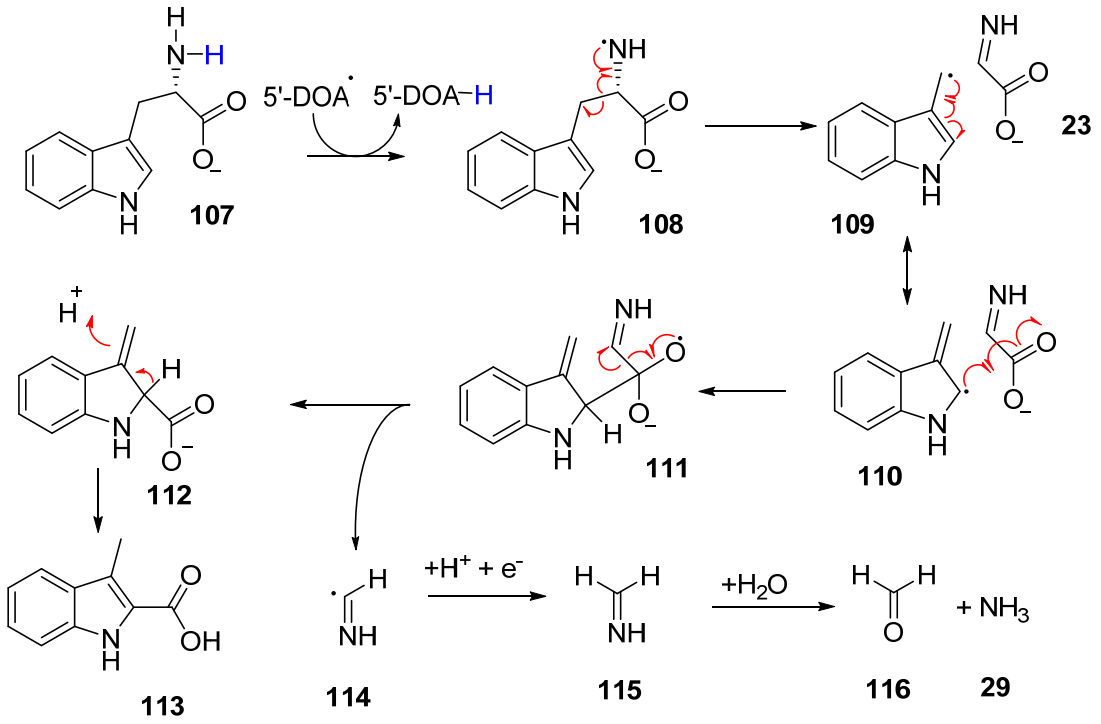



**Figure 69 Proposed mechanism for NosL.**

Following hydrogen abstraction from the amino group of tryptophan **107**, catalysed by DOA•, the scission of $C_\alpha$-$C_\beta$ bond generates a 3-methyl-indole radical **109** and DHG **25**. The addition of **25** to **109** leads to the formation of the intermediary radical **111**, that leads to a second scission, with the release of a cyano radical **114**, and formation of **112** which isomerises into the final product 3-methyl-2-indolic acid **113**. Adapted from Bhandari *et al.* [578].

Unlike HydG, however, after cleavage of the amino acid substrate, neither of the fragments leaves the active site in NosL. The indole radical is thus formed by tryptophan cleavage and recombination of the methyl indole radical with dehydroglycine, leading to the formation of the product 3-methyl-2-indolic acid [307, 321]. A closer comparison of the TIM barrel tunnel reveals the structural basis behind the differentiation of the two mechanisms.

Four residues: Phe86, Tyr173, Asn246 and Val319 (*Sa*NosL numbering) in NosL show an interesting structural contrast to their HydG equivalents: Phe81, Asn167, Gly232 and Ile309 in HydG (*Ti*HydG numbering) (Figure 70). In HydG, as previously discussed, the tunnel permits access between the RSAM and the auxiliary active sites(Figure 70 B); in contrast the NosL 'tunnel' is completely sealed by bulkier residues (Figure 70 C), preventing the escape of intermediates formed in the RSAM active site towards the C-terminus. The closure of the tunnel guarantees DHG is in position for the subsequent rearrangement steps.

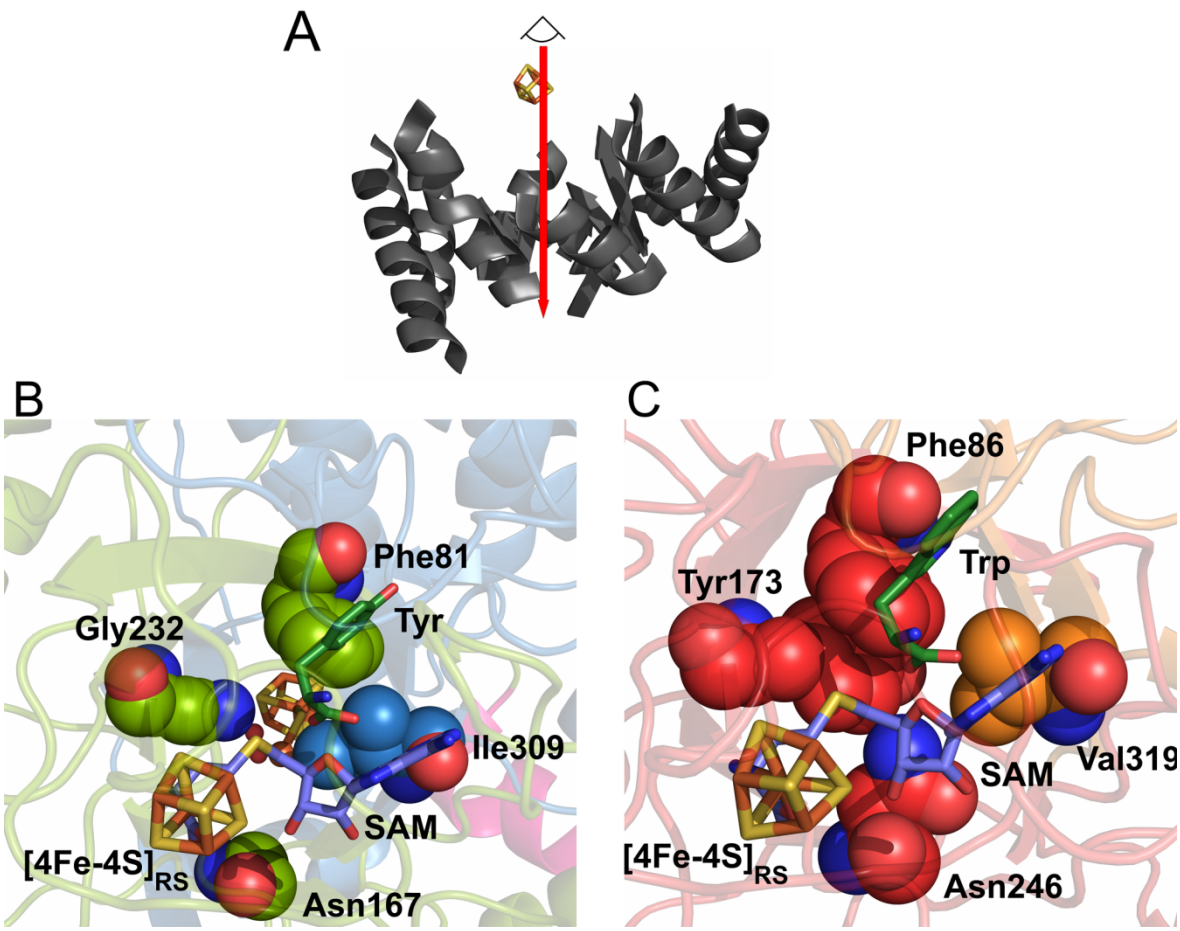



**Figure 70 Structural comparison between the internal TIM barrel of *Ti*HydG and *Sa*NosL along the TIM barrel axis.**

**A:** Visualization axis. **B:** Open tunnel conformation of HydG, showing enough space between the residues for DHG to pass, moving towards the auxiliary cluster. **C:** Closed tunnel conformation of NosL, where the side chains of Phe86 and Tyr173 interact closing any passage toward the C-terminus. *Ti*HydG PDBID: (4WCX and *Sa*NosL PDBID: 4RR3.

With both structure and initial mechanism being potentially conserved across the amino acid lyase sub-group of RSAM enzymes, the crystal structure of NosL provides invaluable information regarding the positioning of the amino acid substrate [321].

The active site (Figure 71) clearly defines the interactions present in NosL, and many of the same residues can be seen conserved in HydG, strongly supporting an analogous placement of the substrate and the same radical abstraction mechanism through the amino group [321, 578].

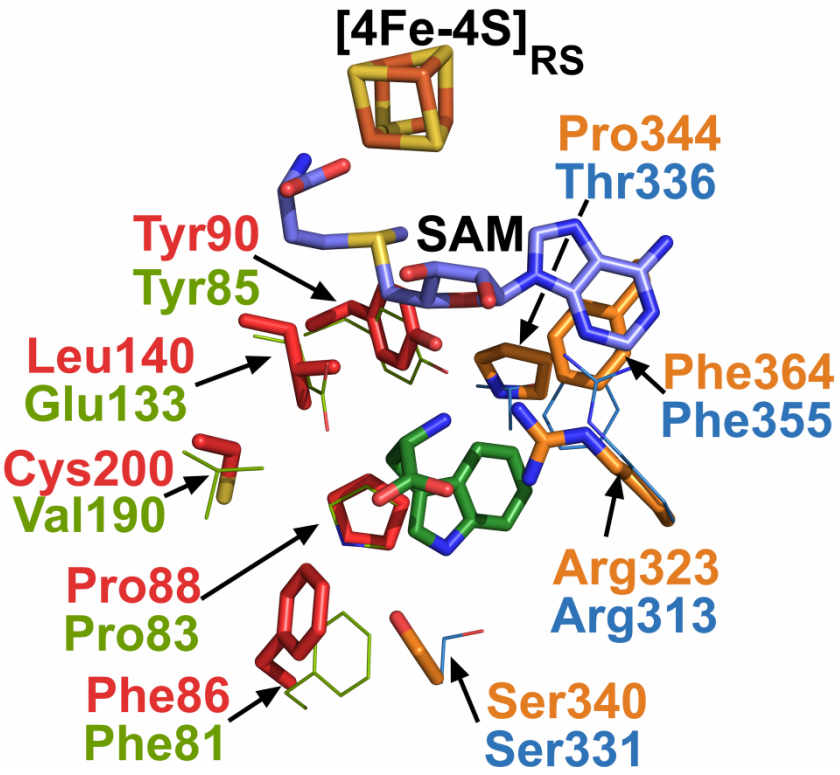



**Figure 71 *Tm*NosL active site.**

Substrate tryptophan is shown in green, SAM in blue, $[4Fe\text{-}4S]_{RSAM}$ in orange and yellow and interacting residues are shown as sticks coloured by their origin in the structure (red: partial TIM barrel fold; orange: C-terminus). Respective residues in HydG are shown as lines (green: partial TIM barrel fold; blue: C-terminus). Phe86, Pro88, Tyr90, Phe202, Arg323 and Phe364 all have exact correspondents in HydG (Phe81, Pro83, Tyr85, Phe192, Arg313 and Phe355). Three residues are not conserved across both structures, Leu 140 (Glu133 in *Ti*HydG), Thr321 (Ser311 in *Ti*HydG) Pro344 (Thr336 in *Ti*HydG), owing perhaps to the different nature of the substrate. *Ti*HydG PDBID: 4WCX and *Sa*NosL structure PDBID:4R33.

### 2.8.2 HydE

The maturases HydG and HydE are both members of the RSAM superfamily of enzymes, both contain an additional auxiliary cluster and are necessary for H-cluster biosynthesis. They share approximately a 26% sequence homology (Figure 72) and various structural similarities seem to be present (Figure 72 and Figure 73), with the superposition of both enzymes giving and RMSD of 5.10 Å.

Despite not being formally considered an amino acid lyase, HydE possesses a higher degree of similarity with HydG than NosL. The fact that two very similar enzymes are present within the same complex mechanism of co-factor assembly is remarkable. These similarities have led McGlynn and colleagues [579] to propose that both HydG and HydE arise from a recent common ancestor.

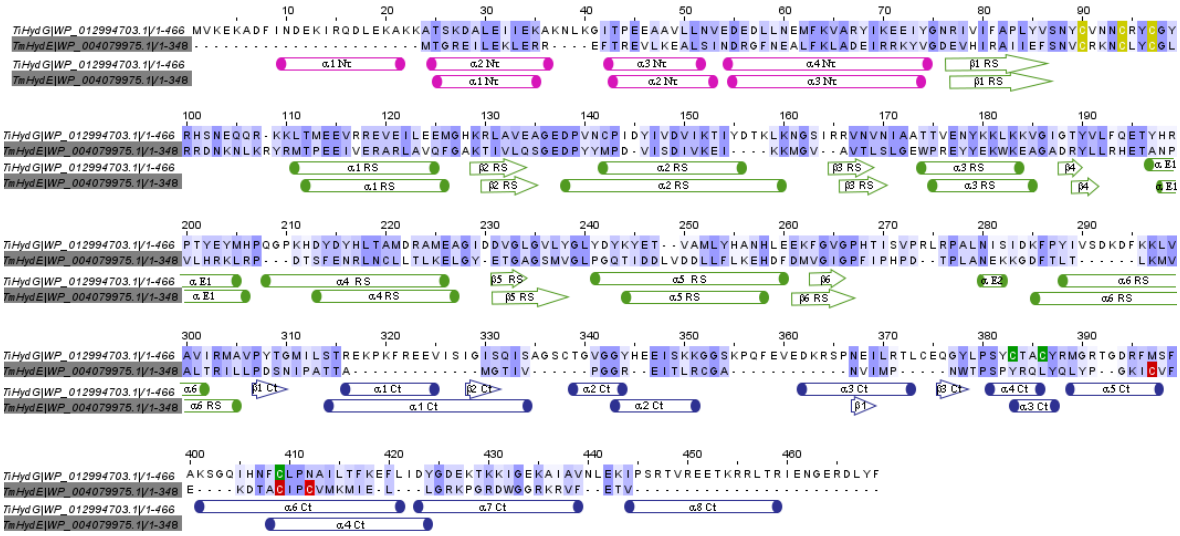


**Figure 72 Sequence alignment between *Ti*HydG and *Tm*HydE.**

Sequence is colour coded, with high conservation shown in darker blue and no conservation in white. The cysteine motif that binds the $[4Fe\text{-}4S]_{RS}$ cluster are shown in yellow and cysteines that bind the auxiliary cluster are shown in green (*Ti*HydG) or red (*Tm*HydE). Secondary structures of both enzymes shown as cylinders ($\alpha$-helix) or arrows ($\beta$-strand), and colour coded as follows: pink, N-terminus; green, RSAM conserved fold; blue, C-terminus. Alignment results obtained from ClustalOmega [575] and visualized using JalView. Gene UNIPROTID: D3T7F1 (*Ti*HydG) and Q9X0Z6 (*Tm*NosL).

The N-termini of these enzymes share high structural similarity (RMSD= 4.4 Å) being formed solely by $\alpha$-helices (Figure 73 B), with one additional $\alpha$-helix being present in HydG to extend the N-terminal region at the beginning of the peptide (Figure 73 A). The core partial TIM barrel fold, highly conserved across the RSAM family, shows, as in the comparison with NosL, a modest degree of similarity between the two enzymes (RMSD= 5.3 Å), with cluster and SAM molecules being positioned in a highly conserved manner (Figure 73 A). As with NosL (section 3.4.2.7.1) only the additional $\alpha$-helix after $\beta$4 can be seen in HydE, an additional secondary feature that the amino acid lyase sub-family of enzymes seem to share (Figure 73 A, red arrow). The small $\alpha$-helix following $\beta$6 is also absent in NosL and seems specific towards HydG (Figure 73 A, green arrow).

The TIM barrel, made complete by the C-terminal extension in each enzyme, creates a tunnel starting from the RSAM active site through the entirety of the structure, approximately down the axis of the TIM barrel. Volume calculations for the available pockets in the protein, using a 1.0 Å probe with CastP [580], reveal a pocket of 822 $Å^3$ for HydE (previously reported as 991 $Å^3$ [297]) and a pocket of 1033 $Å^3$ in the case of HydG (see section 2.7.5, Figure 61). The increased size of the pocket arises from the C-terminus extension HydG possesses.

It is in the C-terminal region that the more significant differences can be observed between the structures of HydG and HydE (RMSD= 6.6 Å). There is an initial similarity between both structures arising from the two β-strands and α-helices that complete the full TIM barrel present in both enzymes. For example, the disordered loop discussed in section 2.7.2 is also present in HydE, and has also been implicated in allowing the substrate enter the active site [297]. However, the auxiliary cluster of HydE, present in its [2Fe-2S] state in the available structures [297] (despite spectroscopic its studies hinting at a [4Fe-4S] cluster [80]), is coordinated by cysteines arising from α8 of the TIM barrel and preceding loop (Figure 73 C, orange arrow). It should be noted that very little structural elements can be found following this cluster coordinating α-helix in HydE. In sharp contrast with the structure of HydE, in HydG the position of the three auxiliary cluster-coordinating cysteines are found in α-helices 9 and 10 (Figure 73 C, blue arrow), part of a larger 5 α-helices extension absent in HydE.

These changes lead to an incredible spatial variation in the auxiliary cluster positioning, with the [5Fe-5S] cluster of HydG being positioned directly below the TIM-barrel tunnel, whilst both clusters in HydE are, despite their separation of 20 Å, still located within the same single active site. Additionally, the presence of the auxiliary cluster in *Ti*HydG is highly conserved within all members of the HydG enzymes and its absence directly affects the enzyme's native activity [311]. No direct correlation between the presence of HydE's auxiliary cluster and activity could be found [297] with HydE from some organisms not even containing the three cysteines necessary for cluster coordination, indicating a strict requirement for the presence solely of the RSAM cluster [581].

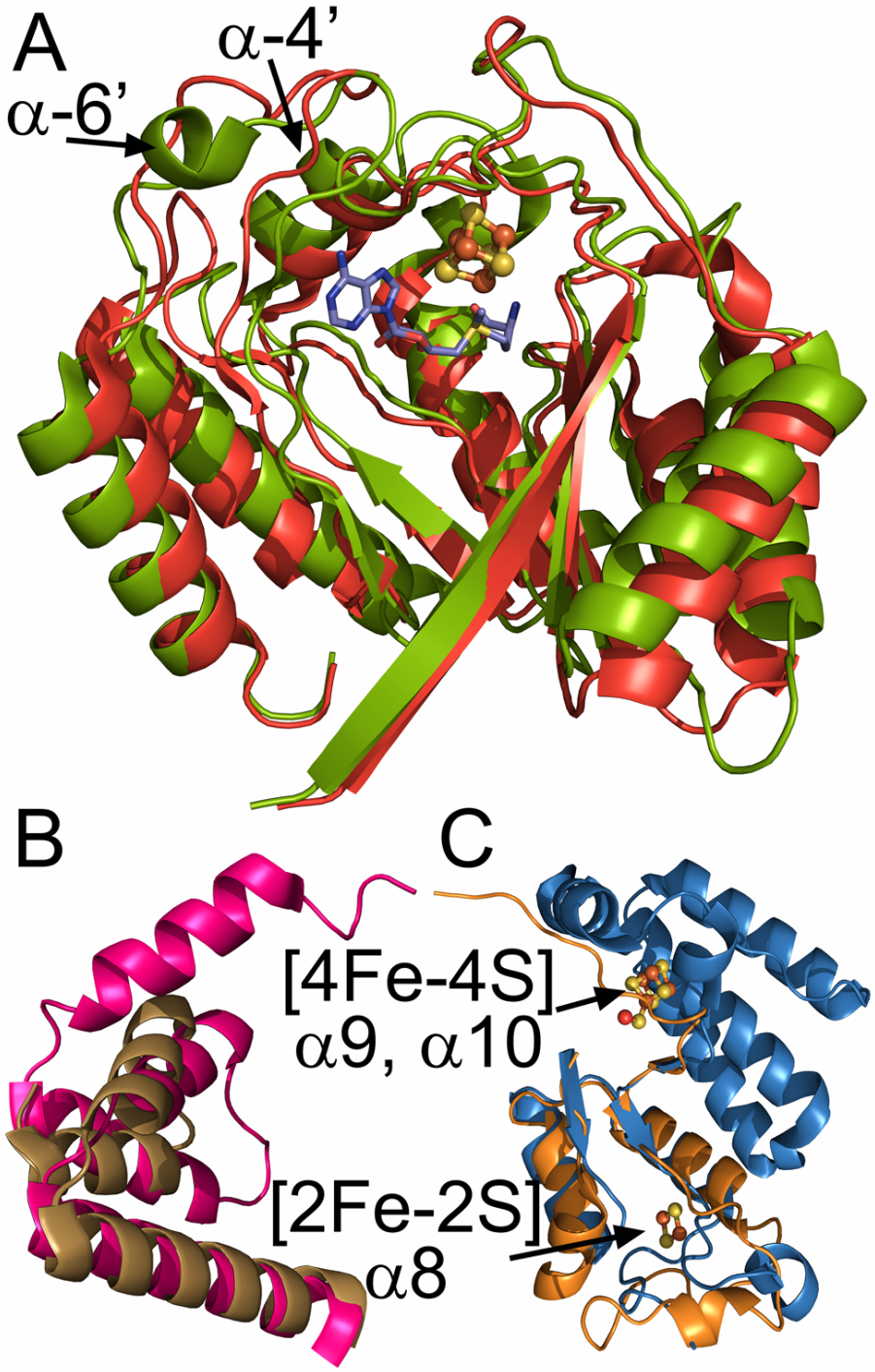



**Figure 73 Structural comparison between *Ti*HydG and *Tm*HydE.**

**A:** Partial TIM barrel comparison. HydG is shown in green and HydE in red. The location of the [4Fe-4S] cluster and SAM is well conserved, while the auxiliary [4Fe-4S] are in very different spatial locations. **B:** N-terminus comparison, with HydG's being shown in pink and HydE in brown. C: C-terminus comparison, with HydG's being shown in blue and HydE in orange. In the C-terminus, where the auxiliary clusters are found, there are little similarities between the proteins. *Ti*HydG PDBID: 4WCX and *Tm*HydE PDBID: 3IIZ.

Neither of the available structures of HydE shows the presence of the substrate, despite an extensive co-crystallization library experiment with various small molecules [581]. The structure does show, however, the presence of several chloride ions near highly conserved residues in two potential active sites. The first of these sites showed two disordered chloride ions near Arg159, Tyr306 and a PATTA HydE conserved motif; while the second active site, located at the opposite end of TIM barrel in relation to the RSAM site, showed another chloride ion unusually bound to Thr134, with Arg54 and Arg155 compensating for the negative charge of the ion. Interestingly, despite the residues not being conserved in relation to HydG, the position of the first putative active site in HydE is shared with NosL and potentially HydG as well (Figure 74).

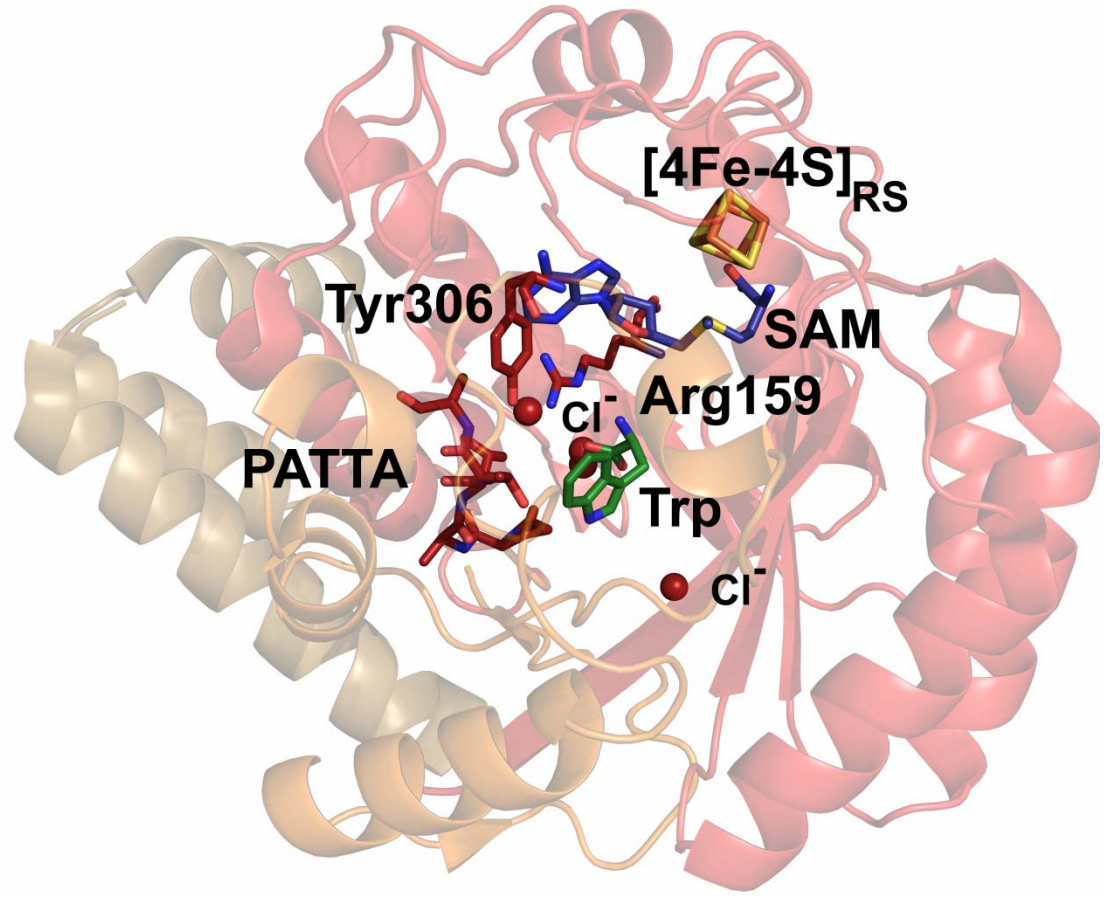



**Figure 74 Putative active site for *Tm*HydE.**

The structure of HydE is shown as a semi-transparent cartoon. The highly conserved Arg159, Tyr306 and 266PATTA motif, together with chloride ions are shown in red. The SAM cluster, conserved in all RSAM enzymes is shown in orange and yellow, together with SAM in blue. Tryptophan from NosL is shown in green, located fairly near the active site formed by the conserved residues and where the chloride ions are located. *Tm*HydE structure PDBID:3IIX [297]; *Sa*NosL structure PDBID:4R33 [321].

## 2.9 Co-Crystallization Studies of *Ti*HydG With L-Tyrosine

The initial commercial broad screens (Section 2.5) provided no crystals of the HydG:tyrosine complex, despite co-crystallization attempts in the presence of L-tyrosine. It has to be noted that L-tyrosine shows very low solubility at the biological pH range (5.5-9.0) [582] where the HydG protein is stable, severely limiting the final concentration of substrate possible for crystallography studies. The highest concentration stock solution possible was 20.0 mM, leading to only 3.0 mM being present in the crystallization mixture.

As the value of $K_M$ for tyrosine, calculated in studies with *Ca*HydG, was shown to be relatively high ($K_{MTyr}$ = 278.9 ± 3.4 mM) [311], the binding of the substrate to the protein is not very efficient. A substrate concentration of 10 x $K_M$ is often used in enzymatic studies as a saturating condition (90% of enzyme molecules are substrate bound) [583], and has similar applications in crystallography studies, as it increases the chances of the formation of an enzyme:substrate complex. The concentration of 3.0 mM of L-tyrosine is slightly above the 10x $K_M$ value, but no crystals with the HydG:tyrosine complex were obtained, perhaps due to the thermophilic nature of *Ti*HydG showing even less affinity for the substrate when compared with *Ca*HydG (see section 2.4).

Different modifications to the crystallization method were attempted to increase the concentration of tyrosine. The addition of methanol or 2-propanol to improve the solubility of the compound did not vary significantly the final concentration of the stock solution and completely inhibited crystal formation. Additionally, despite an increased solubility of tyrosine in the organic solvent DMSO, the addition of the DMSO:tyrosine mix into a buffered aqueous solution led to the immediate precipitation of L-tyrosine, making this solution unsuitable for crystallization experiments.

### 2.9.1 Attempts to Obtain *Ti*HydG:Tyrosine Co-Crystals by Increasing *Ti*HydG Concentration

As a means of overcoming the concentration restriction for the substrate, whilst keeping the crystallization conditions the same (Method 9.8), *Ti*HydG stock solution was concentrated to 3 mM, allowing for the tyrosine in the drop to be raised from 3 mM up to 9 mM, whilst maintaining the same final concentration of protein. Using the Hampton PEG RX commercial screen, a novel condition showed formation of crystals after 4-week incubation, in the presence of 6 mM L-tyrosine.

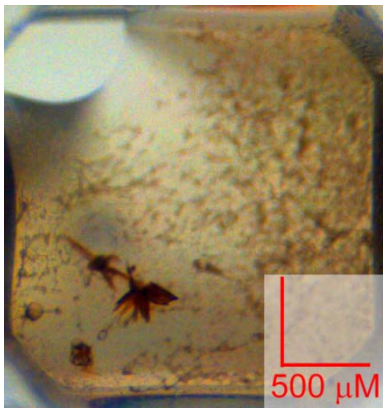



**Figure 75 Co-crystallization of *Ti*HydG with 6 mM tyrosine**

Crystals formed in the presence of 6 mM L-tyrosine in condition C2 of Hampton PEGRx HT commercial screen (0.1 M Imidazole, pH 7.0; 20% Jeffamine® ED-2001, pH 7.0).

The crystals were flash frozen in liquid $N_2$, in the absence of any external cryoprotectants and tested in Diamond Light Source facilities, beamline I03, at a constant temperature of 100 K and the high energy 0.975 Å wavelength. The complete dataset was processed automatically by X-ray Detector Software (XDS) [554], with data reduction and scaling done automatically using Pointless and SCALA [555]. Data collection results are summarized in Appendix A.4.

Being in possession of a complete structure of the enzyme, even in the absence of substrates, greatly simplifies obtaining the structural data, as molecular replacement can be used [557] (see section 8.3.2.1). *Phaser* uses rotation and translation functions to fit the search model in the correct orientation into the electron density data. After each rotation/translation functions, the log likelihood gain (LLG) is calculated, corresponding to the data which explains the observation values with the highest probability. The functions stop when the LLG score can be improved no longer. The higher the translation function Z-score (TF Z-score), the more likely the solution is to be correct, with values of above 5 standard deviations being considered a correct assignment [584]. As the crystals arise from the same enzyme (*Ti*HydG), molecular replacement easily finds a solution with very high scores, creating a suitable initial model.

**Table 9 Score values for molecular replacement using *Phaser*.**

| | **0.8 mM *Ti*HydG + 10 mM SAM + 6 mM L-tyr** |
|---|---|
| | **Hampton PEGRx C2** |
| LL G-score (best) | 4531.052 |
| TF Z-score (best) | 7.2 |

The use of molecular replacement allows an initial look into the available structure. Despite this, the low resolution data for the sample (3.0 Å) limited the degree to which the structure could be refined [585]. Using Polygon from the Phenix software package [559], the data obtained was compared with all the structures within the Protein DataBank database with the same resolution, as a means to validate the results (Figure 76). All parameters fell well within the average values for the already accepted structures, allowing us to confidently use the resulting model and respective electron density map to observe whether density for the substrate is present.

**Table 10 Refinement statistics for the crystallographic datasets of *Ti*HydG with6.0 mM L-tyrosine, following refinement.**

| | **C2 Hampton PEG Rx HT** |
|---|---|
| $R_{work}$ / $R_{free}$ | 0.22/0.27 |
| No. atoms: | |
| Chain A / Chain B | 6894/6600 |
| Ligand/ion | 27/32 |
| Water | 35 |

| | C2 Hampton PEG Rx HT |
|---|---|
| *B-factors*: | |
| Chain A / Chain B | 40.5/29.0 |
| Ligand/ion | 16.8/28.2 |
| Water | 30.2 |
| R.m.s deviations: | |
| Bond lengths (Å) | 0.006 |
| Bond angles (°) | 0.86 |

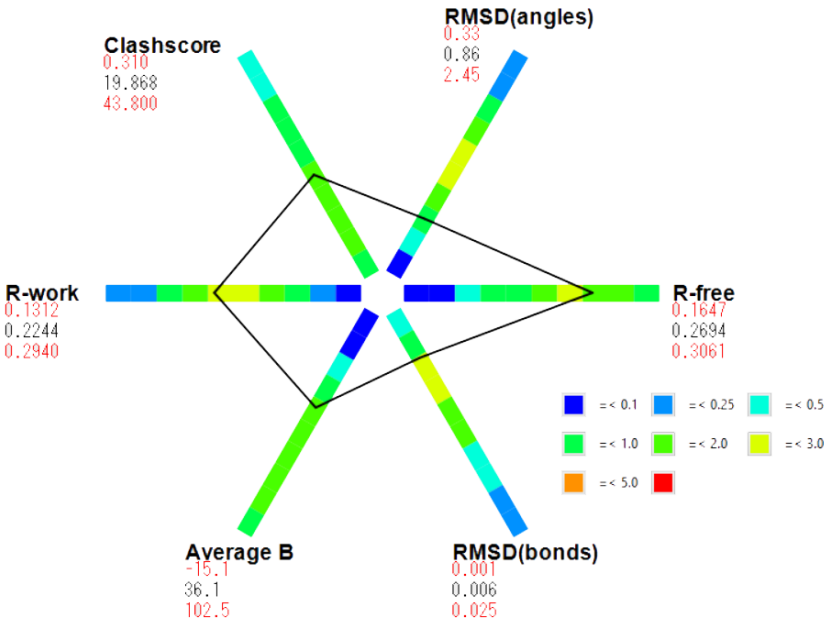



**Figure 76 POLYGON analysis of *Ti*HydG crystal formed in the presence of 6.0 mM L-tyrosine.**

Comparison of several quality indicators to structures with similar resolution within the PDB database. Histograms are coloured based on weighted number of structures with similar resolution.

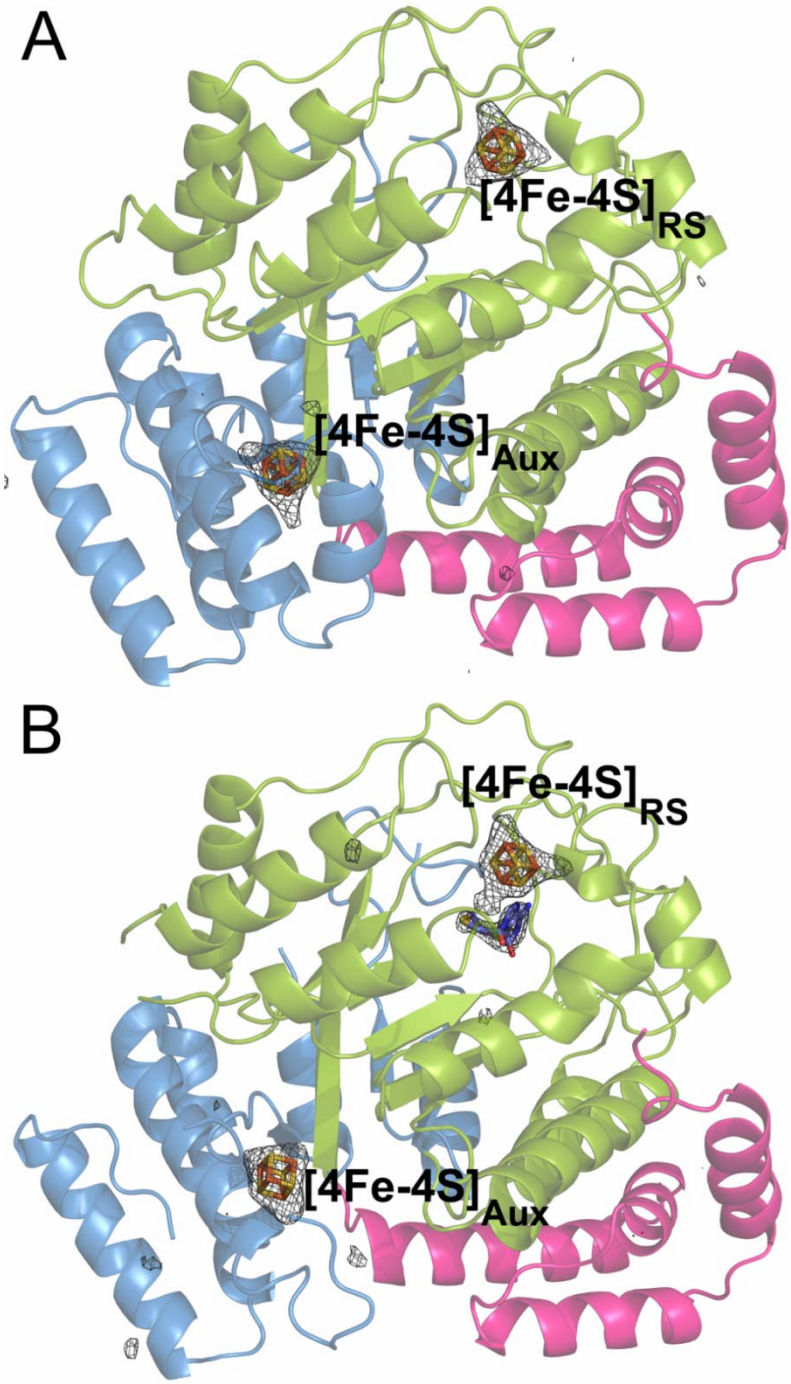



**Figure 77 Structure of *Ti*HydG grown in the presence of 6 mM L-tyrosine.**

The mf0-Dfc electron density difference map is shown (σ cutoff=4.0), indicating non-residue density. The map does not reveal any density for any ligand apart from both clusters and SAM in monomer B. **A:** Data from monomer A. **B:** Data from monomer B.

The structure shows no well resolved density available for L-tyrosine in either potential active site (see section 2.7.5), nor the presence of electron density for the labile iron in the auxiliary cluster.

### 2.9.2 Crystallization With Low Concentration of *Ti*HydG

Using the opposite principle that guided section 3.9.1, if crystal formation can be reliably obtained with lower concentrations of protein, the overall amount of L-tyrosine introduced in the drops can be increased. A seeding experiment was prepared, using condition F1 of Hampton PEGRx HT, the only crystallization condition where crystals are obtained with the labile iron present. As previously discussed (section 2.5.3), formation of crystals previous micro-seeding experiments were mostly unsuccessful. Lowering the protein and SAM concentration in the drops has led to more reliable seeding experiments, as shown in Figure 80. The results show a shower of small crystals when the dilution of the seed mother is low and many microscopic crystals are available to form nucleation points, and as the dilutions increase, the number of potential crystals decrease and can grow larger due to more protein being available for each crystal.

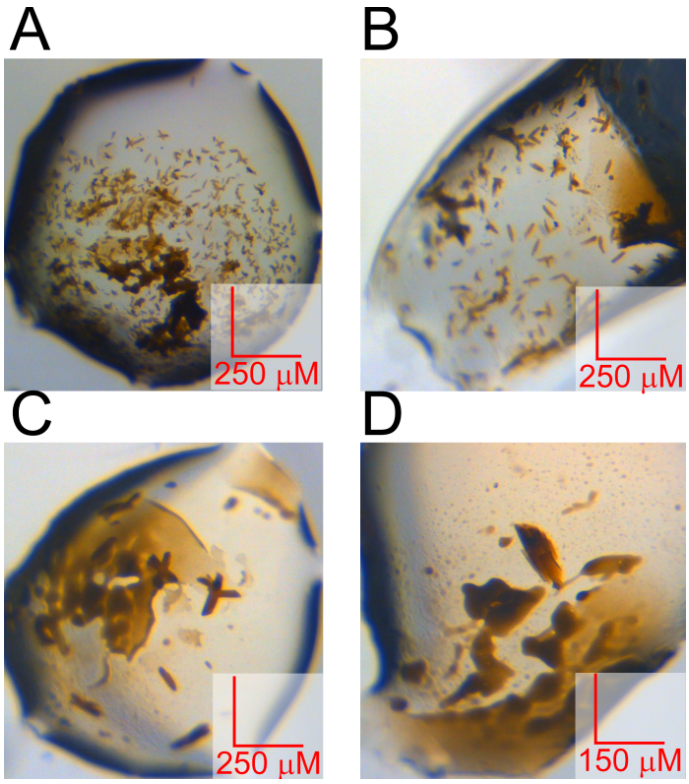



**Figure 78 Microseeding experiment with lower concentration of *Ti*HydG**

The seeding stock was diluted by serial dilutions on a 1/10 basis, from **A:** $1x10^{-1}$ to **D:** $1x10^{-4}$. The number of crystals decreases with the dilution factor, but increase in size. Crystals grown in condition F1 from PACT Premier (0.1 M Bis-Tris propane pH 6.5, 0.2M sodium fluoride, 20% PEG 3,350).

Three different final concentrations of *Ti*HydG were tested to verify what the lowest concentration of protein was where crystallization was still observed: 0.29 mM, 0.44 mM and 0.88 mM *Ti*HydG. Additionally, three concentrations of SAM were also tested, to provide a wide range scan of protein:SAM ratio: 0.5 mM; 1.0 mM and 3.0 mM (Method 9.11). As shown in Figure 79, lowering the protein concentration to 0.3 mM did not inhibit crystal formation and, when allied with a decrease in SAM to 3.0 mM, the crystal formation became significantly more reliable.

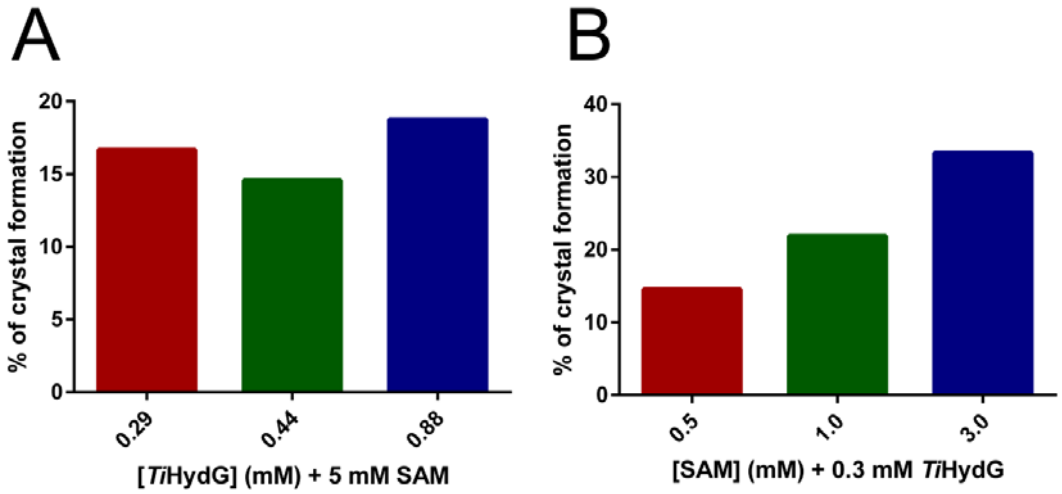



**Figure 79 Occurrence of crystal formation in lower *Ti*HydG and SAM concentration.**

All crystallization attempts used F1 condition from Hampton PEGRx (0.1 M Bis-Tris propane pH 6.5; 0.2 M sodium fluoride; 20% w/v PEG 3350). Percentage of crystal formation calculated by the number of wells in a 96-well plate that showed crystals. **A:** Changes in crystal formation with different HydG concentrations. **B:** Changes in crystal formation with different SAM concentrations and low [*Ti*HydG].

Having identified the most likely condition for crystal formation with the least amount of protein, 0.3 mM *Ti*HydG and 3.0 mM SAM, a new set of seeding experiments were carried out, increasing L-tyrosine concentration to 12 mM, forty times the concentration of protein (Method 9.12). The introduction of L-tyrosine in the crystallization system decreased significantly the formation of crystals, and only 1/288 macrocrystals formed. Such a significant decrease in the number of crystallization events may indicate the formation of an heterogeneous mixture of free protein and protein complexed with L-tyrosine, which may negatively affect the crystallization process [544]. Optimization of the conditions is required to improve reliability of the reproduction of these results.

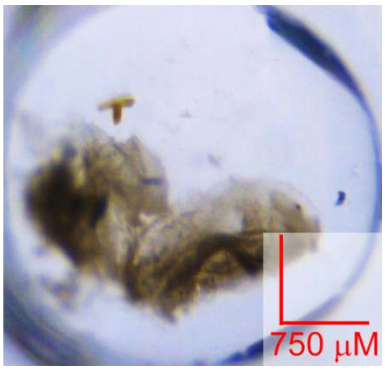



**Figure 80 *Ti*HydG crystal formed in high L-tyrosine content**

Crystal formed in the presence of 12 mM L-tyrosine, 0.3 mM *Ti*HydG and 3.0 mM SAM. Condition F1 from the commercial screen Hampton PEGRx HT.

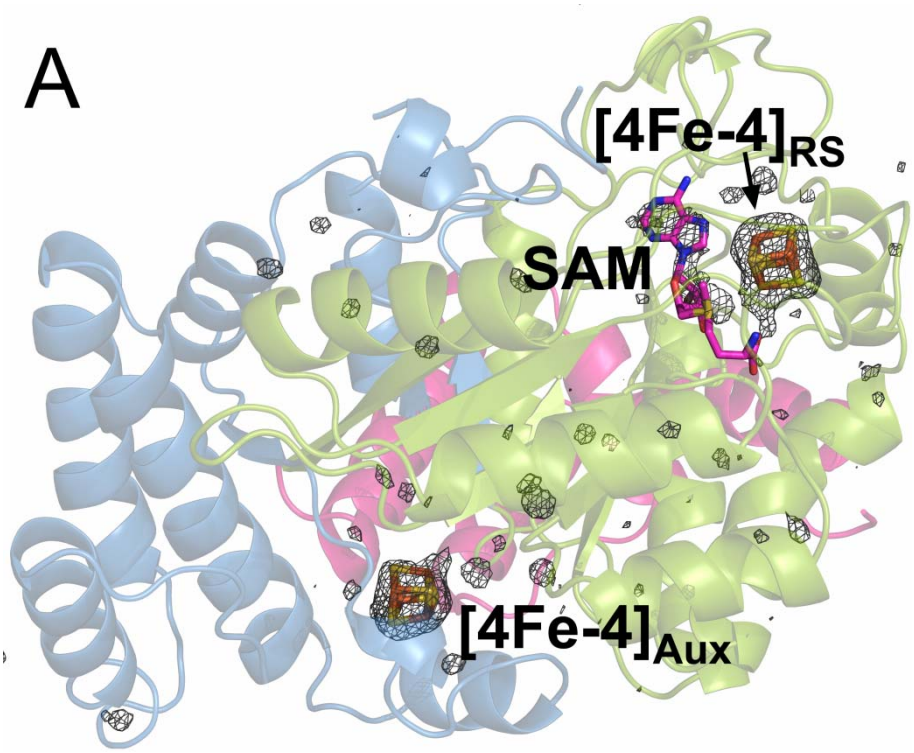
The crystal was flash frozen in liquid $N_2$, in the absence of any external cryoprotectants and tested in the European Synchrotron Radiation Facilities (ESRF), beamline BM14U, at a constant temperature of 100 K and 1.7396 Å wavelength, for collection of anomalous data. The complete dataset was processed by GrenADeS, the automatic pipeline available at ESRF [586].Data processing results and refinement are summarized in Appendix A.5.

The high-resolution structure of *Ti*HydG was once again used for molecular replacement and a high-scoring solution was obtained (Table 11). Unfortunately, the electron density maps fail to show any density that could be attributed to the substrate (Figure 81), despite SAM being present in both monomers. The presence of SAM in monomer A may have arisen from a higher effective concentration of the co-substrate comparing with the amount of protein used for crystallization.

Interestingly, the density of the auxiliary cluster seems to have the labile iron absent, despite the crystal being formed in condition F1 (Table 5) which reliably crystallizes *Ti*HydG showing the complete $[5Fe5S]_{Aux}$ cluster. The absence of the labile iron can possibly be attributed to either changes in protein concentration when setting up the crystallization condition (which are much less saturating and the labile iron may have been released onto solution), or to the excess of L-tyrosine causing a pH shift in the drop which reduces the instability of the labile iron in its native position. The latter could also explain why a single crystal was formed in this crystallization experiment.

**Table 11 Score values for molecular replacement using *Phaser*.**

| | Hampton PEGRx F1 |
|---|---|
| LL G-score (best) | 9367.654 |
| TF Z-score (best) | 8.0 |



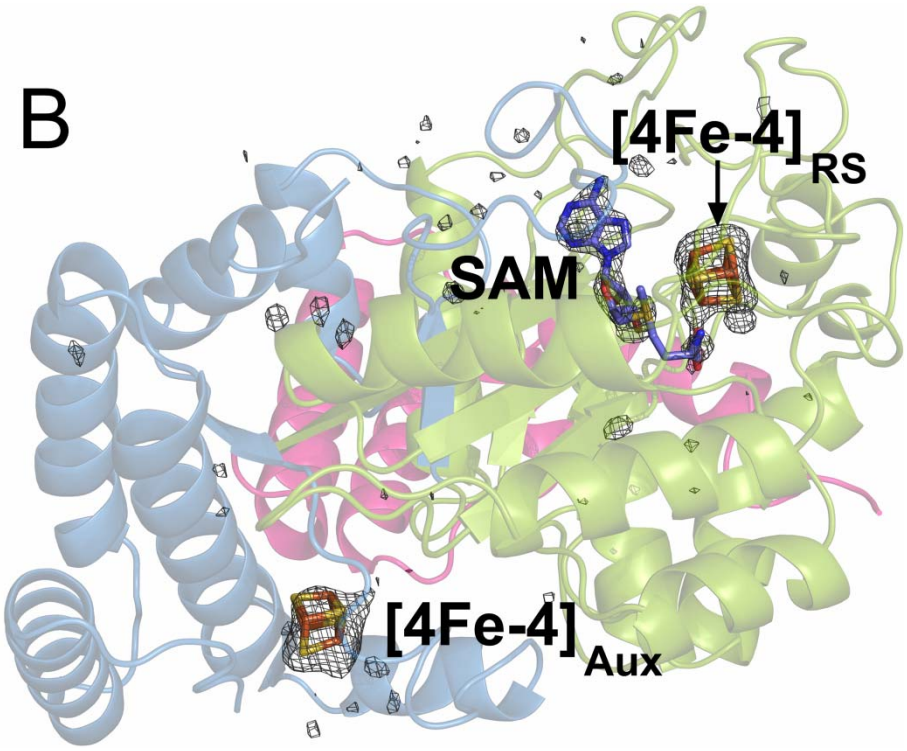



**Figure 81 Structural information obtained for the low concentration crystal of *Ti*HydG (0.3 mM), grown with 3.0 mM SAM and 12.0 mM L-tyrosine.**

The mf0-Dfc electron density difference map is shown (σ=2.0), indicating non-residue density. The electron density corresponding to both clusters together with SAM in both monomer A (magenta) as well as in monomer B (lilac), with no additional density observable indicative of any other ligands apart from some water molecules. **A:** Monomer A. **B:** Monomer B.

## 2.10 EPR Analysis

Electron paramagnetic resonance (EPR) is a commonly used technique to probe the nature of Fe-S clusters [73], as only paramagnetic species are EPR active and variations in g-values can be indicative of different forms present. To provide further experimental evidence for the existence of the [5Fe-5S] auxiliary cluster in HydG, collaborative work was initiated with Dr. Daniel Suess from Prof. David Britt's group in UC Davis (CA, USA). The signal of the [5Fe-5S] cluster was observed. This group prepared, characterized and analysed the EPR spectra of samples of HydG from *Solfalobus solfataricus* (*So*), validating the results obtained in the structure.

In the presence of a reducing agent, *So*HydG displayed two distinct signals, one corresponding to the SAM-bound [4Fe-4S] cluster S=1/2 system with g = [2.01; 1.88; 1.85], as had been previously reported [311, 314]; and a new signal with S=5/2 and g = [9.5; 4.7; 4.1; 3.7]. This spin system is the

proposed result of the ferromagnetic coupling between an S=1/2 $[4Fe\text{-}4S]^{+}$ cluster and an S=2 $Fe^{2+}$ centre, which was confirmed by spectral simulation [587]. Additionally, as iron lability is proposed to be a key feature of the maturation process [296], loss of the iron should translate into a change from S= 5/2 [5Fe-5S] cluster into S =1/2 [4Fe-4S] cluster. The addition of $K^{13}CN$ into a reduced HydG sample with SAM results into two S= 1/2 EPR signals, g= [2.09; 1.94; 1.93] and [2.06; 1.95; 1.93] and, as expected, a loss of the S= 5/2 signal.

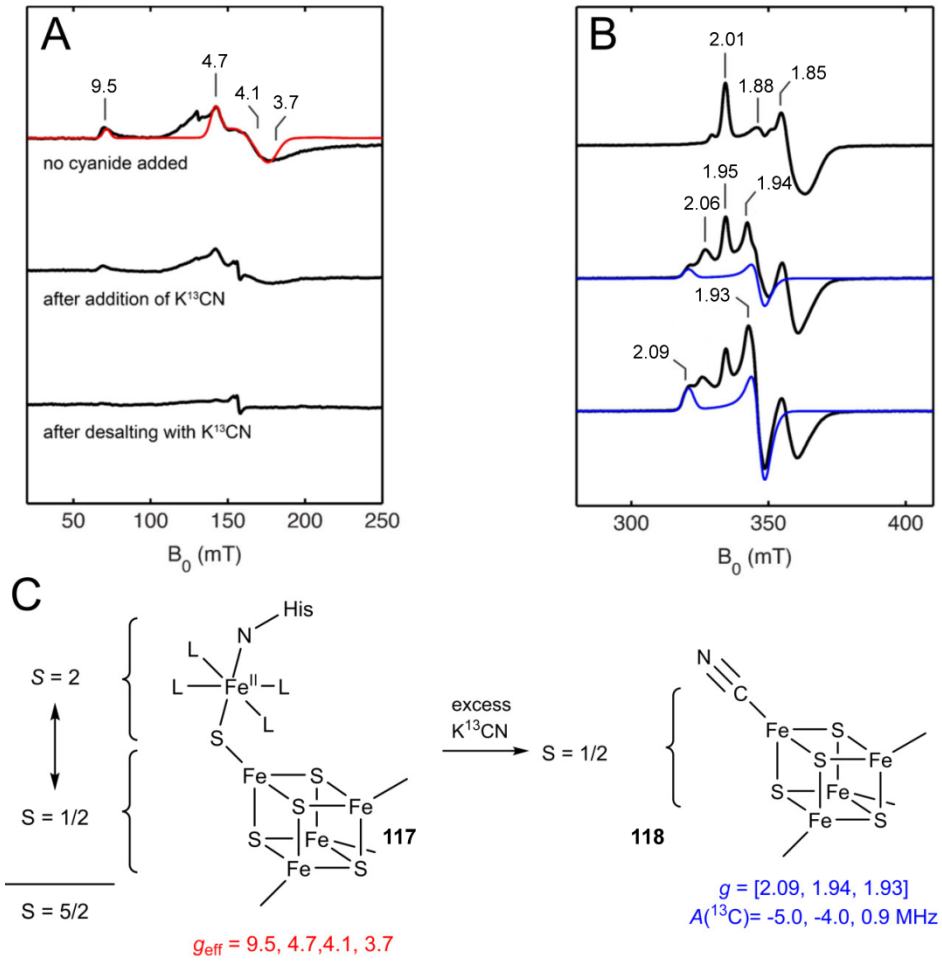



**Figure 82 EPR spectroscopic studies of *So*HydG.**

X-band EPR spectra of So HydG recorded under the following conditions: 9.38 GHz, 10 K, and **A:** 5.00 or **B:** 0.126 mW. In both **A** and **B** the top spectra represents dithionite-reduced HydG in the presence of 3 mM SAM; middle spectra represents dithionite-reduced HydG in the presence of 3 mM SAM and 20 mM $K^{13}CN$ and. bottom spectra represents dithionite-reduced HydG after desalting three times with buffer containing 3 mM SAM, 15 mM $K^{13}CN$, and 10 mM dithionite. Red trace is the simulation of the S = 5/ 2 species **108** assigned to the [5Fe-5S] cluster with D =+ 4.5 cm − 1 and E / D = 0.255, whose model is shown in **C**. Blue traces are the simulations of the species assigned to a $^{13}CN$-bound [4Fe- 4S] cluster **109** with g = [2.09, 1.94, 1.93] and σ = (0.03, 0.02, 0.04), whose model is shown in **C**.

## 2.11 Mechanistic Update

From the structural information obtained, a significant leap in understanding the mechanism of action of HydG was achieved. Assuming the placement of L-tyrosine near the SAM binding site occurs *in vivo*, SAM-derived 5'-deoxyadenosyl radical abstracts an α-amino hydrogen atom [321], followed by the predicted homolysis of L-tyrosine Cα-Cβ bond to yield 4-oxidobenzyl radical anion and DHG [314]. The hydrophobic nature of the active site may help to exclude water, protecting DHG from hydrolysis and allowing the molecule to proceed unmodified to the second half of the reaction.

The idea that the first half of the reaction (L-tyrosine cleavage) is independent from the second half (CO/CN formation), had been previously hinted at by activity assays of a C-terminal deleted HydG from *Clostridium acetobutylicum* which was shown to still cleave L-tyrosine, to yield p-cresol [311, 312]. The structural data shows the means of division of the half-reactions into two active sites each with its cluster. Separating them is the tunnel formed by the full TIM Barrel, which shows only water molecules in the structures, but which the modelling studies show a DHG molecule could be allowed to pass unobstructed. The structure of the [5Fe-5S] auxiliary cluster places the labile iron in a unique position to bind DHG that is able to pass through the tunnel Ligand coordination shows an octahedral geometry in the structure, which may well replicate the *in vivo* binding of two DHG. The labile iron also opens up possibilities regarding cleavage of DHG, explaining how two electron requirements for breaking the organic bond can be surpassed, with one electron arising from the labile iron and the second from the $[4Fe\text{-}4S]^{2+}$ cluster.

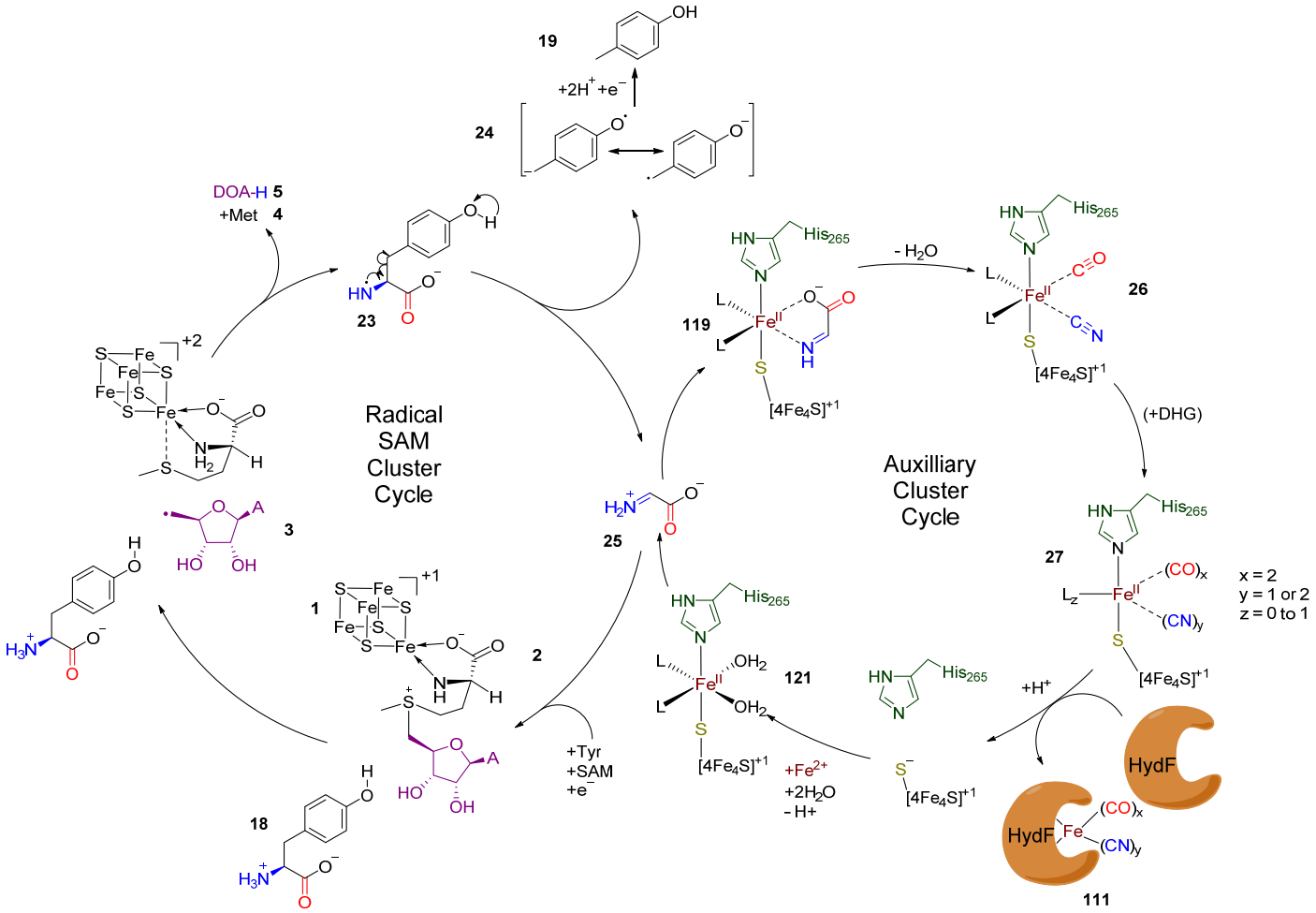



**Figure 83 Proposed mechanism of action of HydG, divided onto two cycles.**

The RSAM cluster cycle is responsible for the radical hydrogen abstraction of the amino group of tyrosine **18**, leading to its heterolytic cleavage. The subsequent product of this reaction, dehydroglycine (DHG), **25**, becomes the substrate of the auxiliary cluster cycle, binding to the labile iron. The precise mechanism for DHG cleavage is yet unknown, but two DHG may be necessary to bind in order to obtain the $Fe(CO)_2(CN)$ **25** species reported by Kuchenreuther *et al.* [296].

## 2.12 Conclusions

From an initial set of three HydG enzymes from different thermophilic organisms, only two were expressed and purified in a suitable manner for crystallization trials. Enzymes from *Thermotoga lettingae* and *Thermoanaerobacter italicus* showed expression levels upwards of 100 mg per 5L culture. The UV-vis spectroscopic data indicates two [4Fe-4S] clusters present in the protein following chemical reconstitution, with additional iron being present, quantified at 10 molecules of iron per protein molecule. End-point activity analysis confirmed the correct reconstitution and folding of the proteins, with values of tyrosine and SAM cleavage similar to the previously reported.

Commercial crystallization broad screens identified several conditions for the crystallization of HydG from one of the organisms (*Thermoanaerobacter italicus*), with the light-brown crystals being tested at the Diamond Light Source synchrotron facilities to obtain high resolution structural data. Replication of the crystals using both the commercial conditions as well as in house replicated conditions is possible, but highly unreliable.

A structural model for the protein was created using the iron anomalous data to initiate a *de novo* assembly of the molecule. The radical SAM core fold is, as expected, present in HydG, together with 2 $\beta/\alpha$ secondary structures arising from C-terminus which complete the enzyme core into a full TIM barrel. Together with the RSAM [4Fe-4S] cluster, an auxiliary cubane cluster is also present in the structure, both being separated by a 25 Å tunnel formed from the TIM barrel interior. The auxiliary cluster, coordinated by three cysteines, shows a unique composition as the unique iron has a $\mu$2 sulfide bridging to an additional iron, termed labile iron due to its low occupancy in the structure, which is hexacoordinated in a octahedral-like geometry to His265 and the aforementioned sulfide in the principal axis, and to two water molecules and one free amino acid residue in the plane. The nature of the free amino acid ligand could not be fully determined, but it's likely to be methionine arising from SAM degradation, an event visible in one monomer of the structure.

Despite extensive attempts to co-crystallize the protein in the presence of the substrate L-tyrosine, the relatively low concentration of substrate achievable, even following optimization of the crystallization conditions, may be the main deterrent responsible for the lack of structural information obtained regarding its binding site location. The available structure does not fully disprove the proposed positioning of L-tyrosine bound to the auxiliary cluster [296], however the conservation of the substrate near the RSAM cluster in all enzymes of the RSAM superfamily, and

in particular the structural information available for another amino acid lyase, NosL [321], strongly suggests L-tyrosine binds to an available site beneath SAM. Further optimisation of crystallization conditions in the presence of L-tyrosine may provide the definitive answer to its binding.

The cleavage of L-tyrosine in the RSAM active site would mean that the product DHG would have to transverse 13 Å towards the labile iron. A central hydrophobic region in the centre of the TIM barrel tunnel has led to the proposal of DHG traversing in its neutral state, although both edges of the tunnel seem to have charged residues, which could initiate an electromotive force if DHG was charged. Mutagenesis studies in Glu133, Arg129 and Arg165 could provide support for the latter proposal, bearing in mind the C-terminus arginines are positioned near the auxiliary cluster and could be necessary for the formation of CO/$^{-}$CN from DHG.

Another important consideration lies in the necessary dimerization with HydF for product transfer. It seems sensible to assume such dimerization occurs in the C-terminus in regards to HydG, where the auxiliary cluster, with products bound, is located. An early estimate of the binding location for HydF was proposed based on the differences in the disorder levels in a uniquely positioned $\alpha$-helix when the labile iron is present or absent. Mutagenesis in the residues of that $\alpha$-helix could have profound consequences downstream in the activation of the hydrogenase, if the proposal is true, while co-crystallization of the HydG:HydF complex could provide a broader picture of the different stages of the maturation process.

# Chapter 3: Site Directed Mutagenesis of *Ti*HydG

## 3.1 Introduction

While knowledge of the three-dimensional structure of the protein can provide clues regarding its mechanism, mutagenesis is a powerful tool that allows the mechanistic hypothesis to be tested, with many examples of such studies having been carried out in RSAM enzymes [22, 178, 179, 588].

Using the polymerase chain reaction method to introduce specific mutations in the gene is a well-established technique first introduced in the 1980's [589, 590]. It has been used efficiently in RSAM enzymes to investigate the presence of [4Fe-4S] clusters by changing the coordinating cysteines to serines/alanines [83, 104-118]. Creating these mutants, unable to coordinate the $[4Fe\text{-}4S]_{RS}$, is of particular relevance if the protein is thought to possess additional clusters, as it means the non-RSAM clusters can be more readily identified by the remaining iron content of the protein and analysis of UV and/or EPR spectra [37, 38, 118, 171, 176, 177, 186, 591].

In HydG, the relevance of the auxiliary cluster for activity was probed by use of site directed mutagenesis, changing the coordinating cysteines to serines, as well as complete removal of the C-terminus by the introduction of a stop codon prior to an 88 amino acid extension which was absent in ThiH [311, 312]. The observed changes in activity [311, 312], as well as EPR spectra[311], provided strong support of the conserved RSAM cluster being responsible for tyrosine cleavage by a radical abstraction mechanism [312].

The formation of the diatomic ligands from dehydroglycine seems to be, at least in part, related to the auxiliary cluster, with carbon monoxide formation being abolished in the mutants. However, the detection of cyanide in the absence of the auxiliary cluster through cysteine to serine mutations [311, 312], but not in the absence of the C-terminus [311], led to two conflicting mechanisms being presented (section 1.2.2.3.2).

The first, in agreement with the mutagenesis results, predicts an enzymatic elimination mechanism, where two hydrogen abstractions leads to formation of free cyanide and formic acid, the latter of which binds to the auxiliary cluster for the elimination of water to occur, and leading to an auxiliary cluster with CO bound and a free cyanide molecule [317] (section 1.2.2.3.2, Figure 18).

New studies have however have come to question this proposal, in particular the formation of a free cyanide, with the spectroscopic characterisation of a $Fe(CO)_2CN$ synthon [296] (section 1.2.2.3.2, Figure 19). An alternative pathway was put forward, where in which the auxiliary cluster is responsible for the formation of both diatomic products.

This new '$Fe(CO)_2CN$ synthon' hypothesis shows the distinct advantage of be easily extended to include transfer of both diatomic ligands into the scaffold enzyme HydF, in particular if one considers the unique position of the labile iron as an integral player in that transfer step. Testing the necessity of the labile iron in the activity of the enzyme may provide evidence for or against a mechanism involving synthon transfer. As such, the single residue coordination of the labile iron was targeted for mutagenesis studies.

The unique coordination of the labile iron to a single residue of the protein makes His265 a remarkable target for mutagenesis. Sequence alignment studies reveal this residue is highly conserved in all HydG sequences, but absent in related amino acid lyase enzymes (ThiH and CofH, Figure 84 C), which may indicate a potential role of the residue in the HydG-specific DHG cleavage reaction.

Altering the labile iron coordination goes one step further in the modification of the auxiliary cluster coordination [312], as the three coordinating cysteines are still in place for the cubane [4Fe-4S] cluster to be intact, minimizing the alterations to the C-terminus of the enzyme. If the mechanism of diatomic product formation is truly independent of the auxiliary cluster, changes to His265 could result to near WT-levels of cyanide formation. Alternatively, if the mechanism of diatomic formation is common for both $CO/^{-}CN$ and dependent on the $[5Fe\text{-}5S]_{Aux}$ cluster (section 2.11), alteration of the labile iron coordination could lead to the abolishment of diatomic ligand formation.

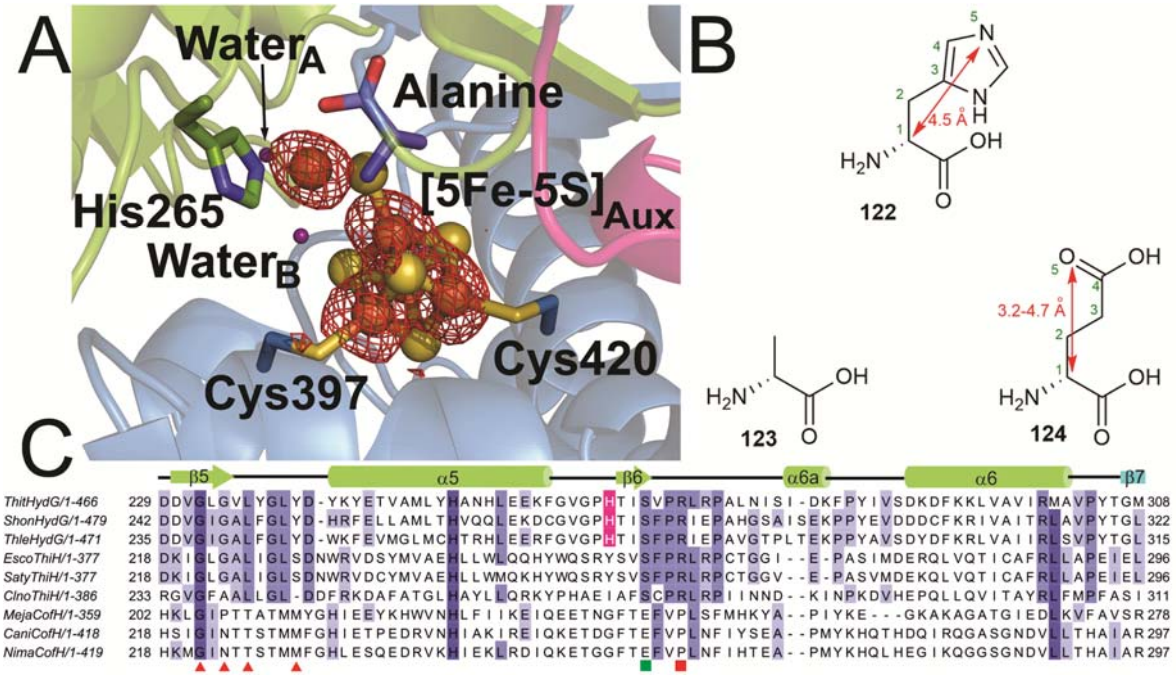



**Figure 84 Conservation of the labile iron coordinating histidine, and potential target for mutagenesis.**

**A:** *Ti*HydG model of auxiliary cluster, with anomalous density map (cutoff 5.0 sigma). **B:** Structure of histidine**122**, alanine **123** and glutamic acid **124**. Number of atoms from the alpha carbon to the atom that can coordinate the iron are shown in green. Average distances between $C_\alpha$ and the potential iron-coordinating atom are shown in red. **C:** Sequence alignment of aromatic amino acid lyases encompassing His265 (*Ti*HydG numbering), in purple. The amino acid is conserved in all HydG's but absent in ThiH, NosL and CofH enzymes. *Thermoanaerobacter italicus* HydG, UniProt accession code D3T7F1; *Shewanella oneidensis* HydG, Q8EAH9; *Clostridium acetobutylicum* HydG, F7ZVC7 ; *Escherichia coli* ThiH, P30140; *Salmonella typhimurium* ThiH, Q9S498; *Streptomyces actuosus* NosL, C6FX5; *Nocardia sp*. ATCC 202099 NocL, E5DUI3 ; *Methanocaldococcus jannaschii* CofH, Q58826 and *Candidatus Nitrosoarchaeum* limnia BG20 CofH , S2E5F8. Sequences were aligned with Multalin [592] and the figure was prepared with Jalview [593].

To test this hypothesis, two mutants were designed, in which His265 was changed into either an alanine **123** or a glutamic acid **124**. The alanine mutant should completely abolish the proteinacious coordination of the labile iron, potentially leading to an impaired activity due to either removal of the labile iron, or to a repositioning of the iron away from its substrate capture point. Glutamic acid was chosen as a second mutant as it can potentially place a coordinating oxygen atom at the same distance as a histidine nitrogen donor, possibly leading to an active enzyme.

## 3.2 TiHydG His265 Mutagenesis by PCR

Primer extension PCR [594] (Method 9.14) was used to introduce the desired mutations in the sequence that encodes HydG. Two primers A and D are complementary to each of the ends of the protein gene, and can be used to insert any mutation. Two additional primers, B and C, are complementary to each other and partially complementary to the location where the mutation is to occur, losing that complementarity in the position where the mutation is to be inserted.

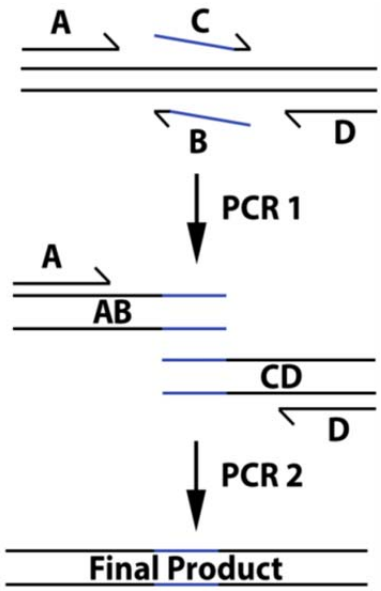



**Figure 85 Primer extension PCR.**

Insertion of target mutagenesis by using two pairs of primers. A and D are primers which are completely complementary to each end of the target gene, while B and C are the mutation specific pairs with complementary ends.

Two sets of PCR occur, the first one creating fragments that go from each end of the gene to the mutagenesis site (AB and CD in Figure 85), while the second PCR is accomplished using the resulting fragments from the first PCR as the template and the conserved primers A and B as the extension primers to yield the final product.

The initial amplification gave rise to two DNA fragments of 853 bp (AB) and 615 bp (CD). They were mixed in a 1:1 ratio as templates for the second PCR amplification, to unite each fragment and amplify the full sequence of *Ti*HydG, now with the mutation inserted. The constructs were confirmed by sequencing, using the MRC PPU DNA sequencing services of the University of Dundee.

The mutated *Ti*HydG genes were inserted into pRD003 vector (section 2.3.1, Figure 34), using the NcoI/XhoI site to remove the wild type *Ti*HydG gene present there. The newly constructed plasmids pCD003 (alanine mutant) and pCD004 (glutamic acid mutant) were transformed into competent *E.coli* BL21 (DE3) cells (Method 9.4.3) for use in protein expression studies.

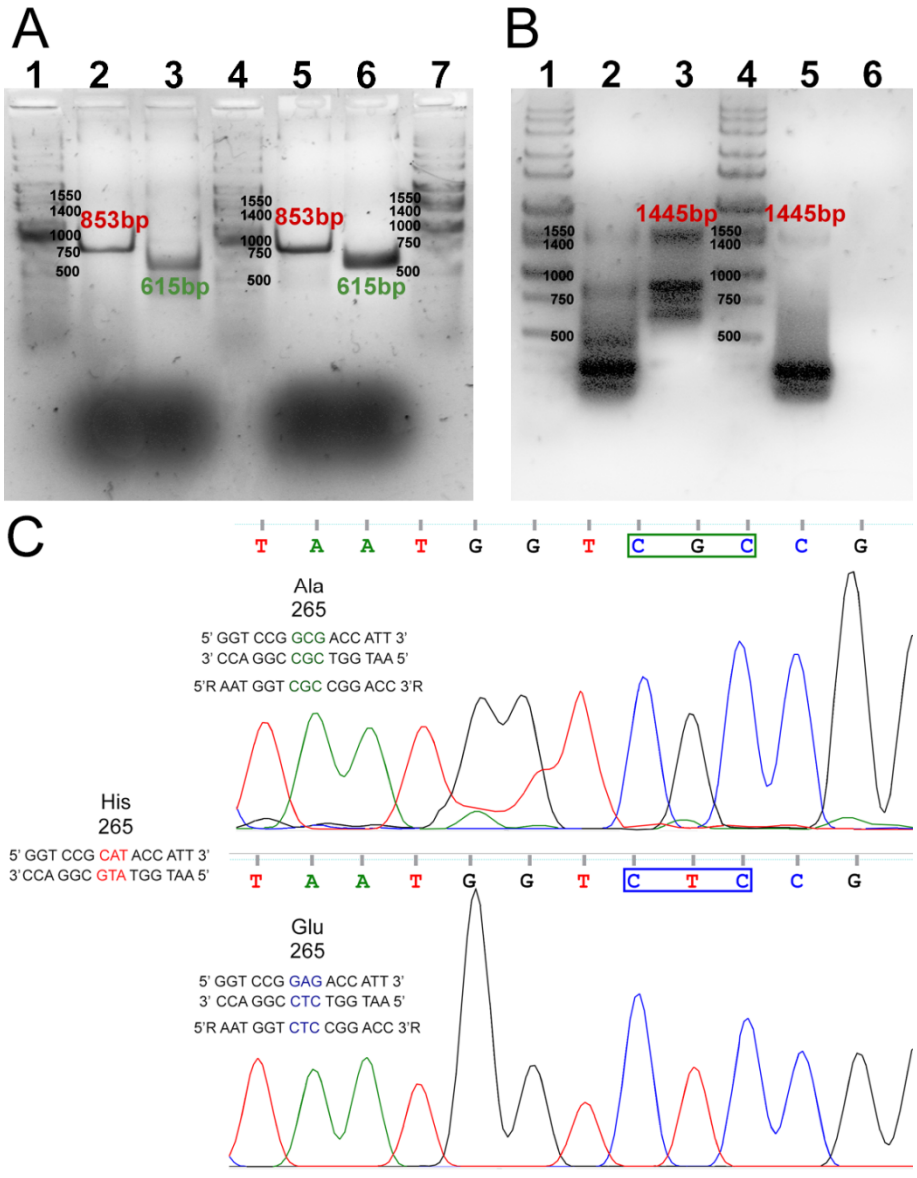



**Figure 86 PCR fragment amplification for the creation of *Ti*HydG mutants.**

**A:** 1st PCR amplification with primer extension PCR: 1;7 –DNA molecular weight marker. 2 – AB amplification for *Ti*HydG His265Glu, with expected 853 bp band present. 3 – CD amplification for *Ti*HydG His265Glu, with expected 615 bp band present. 4 – AB amplification for *Ti*HydG His265Ala, with expected 853 bp band present. 5 – CD amplification for *Ti*HydG His265Ala, with expected 615 bp band present. Reactions done in duplicate.

**B:** The products of the 1st PCR amplification were used for 2nd PCR amplification of the primer extension PCR: 1;4 - DNA molecular weight marker. 2, 3 – *Ti*HydG His265Glu. Several non-specific amplification bands are observed, together with the correct band at 1445 bp. 5;6 – *Ti*HydG His265Ala. A faint band of the correct size is observed in lane 5. Second amplification failed (6), and no bands observed. **C:** Sequencing analysis for both constructs. The use of a reverse primer leads to the sequence presented being the reverse complement of the 5'→3' sequence. For both alanine (green) and glutamic acid (blue) the reverse complement sequence is shown (5'R→3'R).

## 3.3 TiHydG His265 Mutant Expression and Purification

The recombinant expression of each of the *Ti*Purification of both *Ti*HydG mutants routinely showed the presence of the same impurities present in the purification of the WT enzyme. The estimation of the impurity levels by ImageJ still reveals over 80% purity in *Ti*HydG His265Glu purification. In contrast, the calculated purity levels in the case of *Ti*HydG His265Ala are of only 50%, a consequence of the lower yield of the protein of interest (Table 12). In that regard, *Ti*HydG His265Ala more closely resembles *Tm*HydG.

Following the buffer exchange step, the iron sulfur clusters were chemically reconstituted by adding 10 molar equivalents of $FeCl_3$ and $Na_2S$. While the results observed for the His265Glu mutant were similar to those previously recorded for the WT enzyme (section 2.3.3), the addition of $FeCl_3$/$Na_2S$ to *Ti*HydG His265Ala led to the sample quickly becoming very dark, likely to be the result of the accumulation of precipitation and colloidal iron sulfide. This event suggested the alanine mutant was not able to coordinate such an extensive amount of iron and sulfide as the wild type protein. Any attempts to extend the period of incubation and/or to slow the addition of the reconstitution constituents, allowing a more extensive incorporation of iron and sulfide did nothing to prevent the formation of insoluble Fe-S aggregates which, when pelleted by centrifugation, left close to no enzyme present in the supernatant. To prevent the extensive precipitation of *Ti*HydG His265Ala, this mutant was reconstituted using half the molar equivalents (5 M eq.) of $FeCl_3$ and $Na_2S$, levels, which was observed to increase the stability of the protein (Method 9.6.2).

After the gel filtration purification step, the overall purity of both mutant enzymes improved significantly, albeit the overall yield of *Ti*HydG His265Ala suffered a significant decrease, with an increased amount of protein lost during the reconstitution step.

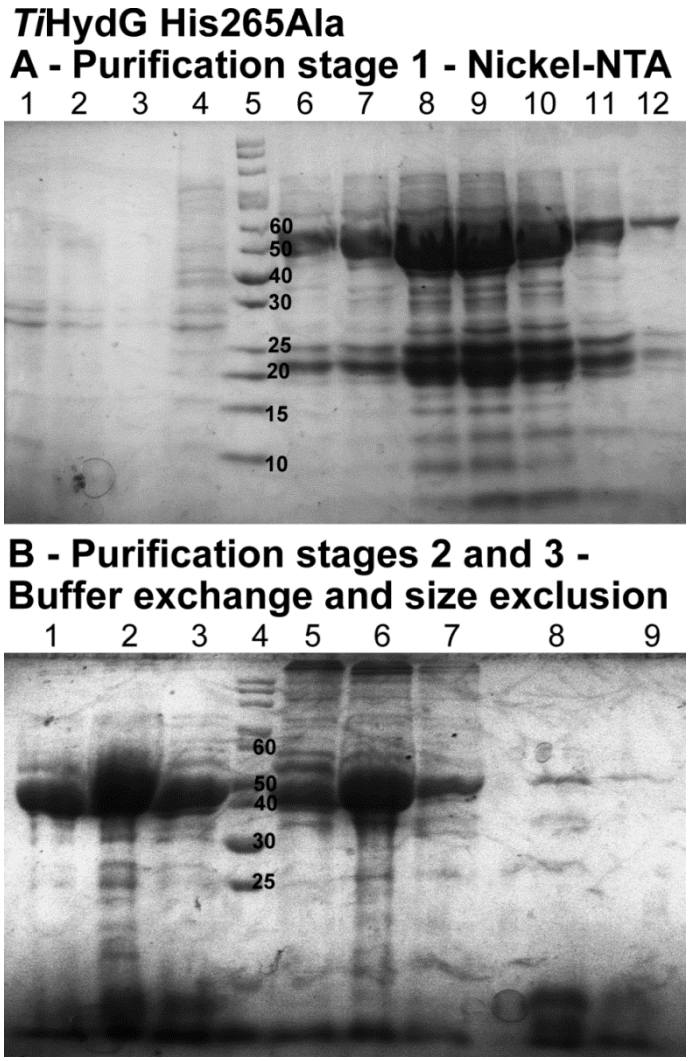



**Figure 87 15% SDS-PAGE gels of samples obtained in the purification steps of *Ti*HydG mutant *Ti*HydG His265Glu.**

**A:** Nickel affinity purification: 1 –Sample of cell lysate; 2 – Insoluble fraction; 3 – Clear lysate loaded onto the Ni-NTA column; 4 – Sample from the wash peak. 5 - Protein molecular weight markers; 6-11 – Samples collected at 60% Buffer B, corresponding to the desired protein at 55 kDa. **B:** Buffer exchange and size exclusion chromatographies: 1-4 – Fractions buffer exchanged into Buffer C; 5 – Protein molecular weight markers; 6-8 – [Fe-S] excess and protein aggregate; 9-11 – Soluble *Ti*HydG His265Glu, with the correct 55 kDa size.

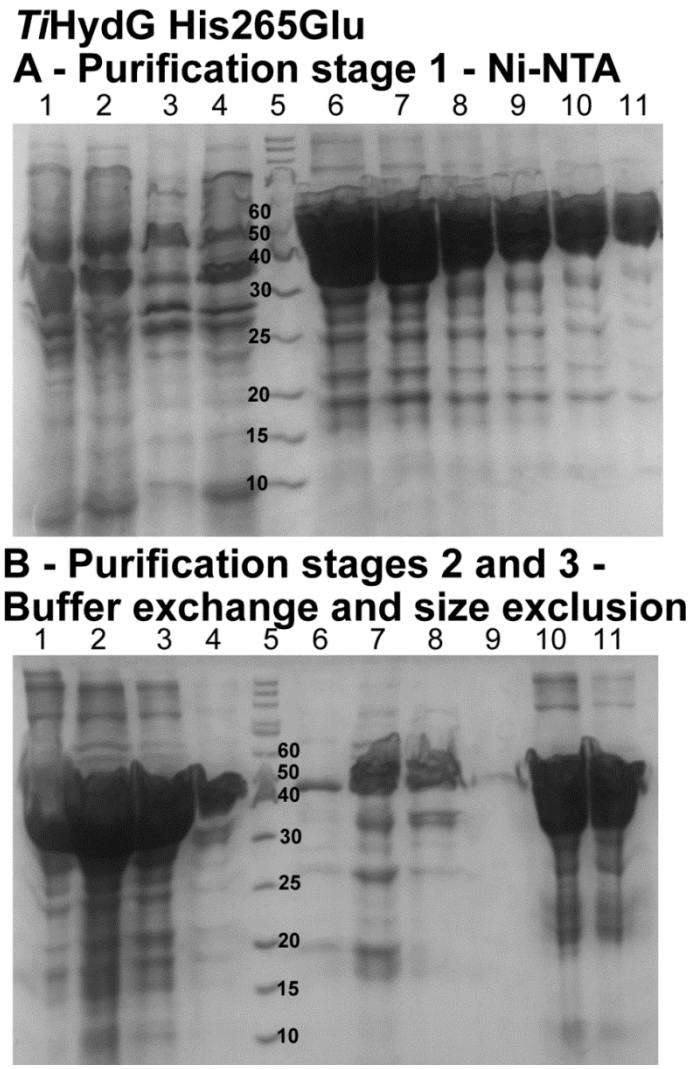



**Figure 88 15% SDS-PAGE gels of samples obtained in the purification steps of *Ti*HydG mutant *Ti*HydG His265Ala.**

**A:** Nickel affinity purification: 1 –Sample of cell lysate; 2 – Insoluble fraction; 3 – Clear lysate loaded onto the Ni-NTA column; 4 – Sample from the wash peak. 5 - Protein molecular weight markers; 6-12 – Samples collected at 60% Buffer B, corresponding to the desired protein at 55 kDa. **B:** Buffer exchange and size exclusion chromatographies: 1-3 – Fractions buffer exchanged into Buffer C; 4 – Protein molecular weight markers; 5-7 – [Fe-S] excess and protein aggregate; 8-9 – Soluble *Ti*HydG His265Ala, with the correct 55 kDa size.

The Bradford estimation (Method 9.5.5) (Table 12) of the protein yield reflects how much more unstable *Ti*HydG His265Ala is when compared with the other two enzymes, possibly owing to the fact that it cannot fully create the auxiliary [5Fe-5S] auxiliary cluster. It is during the second reconstitution step that the yield of *Ti*HydG His265Ala suffers a significant dip, even when milder conditions were used (3.5 molar equivalents of FeCl3 and Na2S instead of the usual 5 molar equivalents).

### 3.3.1 UV-Vis Characterization and Iron Content Analysis of *Ti*HydG Mutants

The UV-vis analysis of the two mutants shows a clear distinction between *Ti*HydG His265Glu ([*Ti*HydG His265Glu]= 23.8 μM, $\varepsilon_{410}$= 34,000 $cm^{-1}$ $M^{-1}$), closely resembling the *Ti*HydG WT enzyme (section 2.3.3), which contrasts with *Ti*HydG His265Ala ([*Ti*HydG His265Ala]= 15.7 μM, $\varepsilon_{410}$= 22,300 $cm^{-1}$ $M^{-1}$) (Figure 89). The spectra of *Ti*HydG His265Ala shows a much less prominent peak at 410 nm and the dramatic changes in the [4Fe-4S] characteristic spectra may indicate that the His to Ala mutation can affect the stability of the entire cubane auxiliary cluster and not solely the labile iron. Iron quantification provides additional support to the notion that the *Ti*HydG His265Ala possesses a significantly impaired auxiliary cluster, with the iron content being calculated at approximately 3.3 ± 1.5 irons per protein.

**Table 12 Protein yields and iron quantification from *Ti*HydG His265Glu and *Ti*HydG His265Ala.**

| Protein | Total protein Ni-NTA (mg) | Iron content before 1st reconstitution (molar equivalents) | Iron content after 1st reconstitution (molar equivalents) | Total Protein gel filtration (mg) | Iron content before 2nd reconstitution (molar equivalents) | Iron content after 2nd reconstitution (molar equivalents) |
|---|---|---|---|---|---|---|
| *Ti*HydG H265E | 210 | 6.44 ± 1.8 | 14.55 ± 5.8 | 182 | 5.60 ± 1.6 | 12.11 ± 2.4 |
| *Ti*HydG H265A | 85 | 0.75 ± 3.0 | 3.19 ± 1.0 | 22 | 2.04 ± 0.2 | 3.33 ± 1.5 |

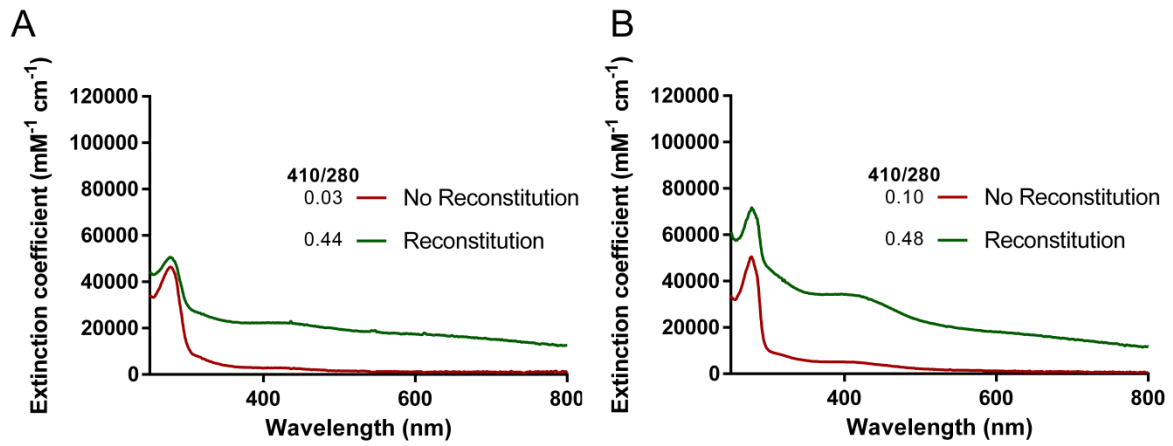



**Figure 89 UV spectra of *Ti*HydG mutants together with the calculated 410/280 values.**

**A:** UV-Vis spectra obtained with samples of *Ti*HydG His265Ala mutant, before chemical reconstitution (red, 12.0 μM) and after the final chemical reconstitution (green, 15.7 μM). **B:** UV-vis spectra of the *Ti*HydG His265Glu mutant before any chemical reconstitution (25.3 μM) and after the final chemical reconstitution (23.8 μM).

### 3.3.2 *Ti*HydG His265 Mutant EPR Characterization

Extending our collaborative effort with Prof. David Britt's group at UCDavis (CA, USA) has revealed initial EPR data concerning both mutants. *Ti*HydG His265Glu shows an EPR signal consistent with the presence of two [4Fe-4S] clusters, both the radical SAM and complete auxiliary cluster, as seen by the change in the low-power signal in the presence of DTH and SAM. The high powered S=5/2 signal, corresponding to the labile iron, can also be observed in the samples, showing an overall EPR spectra very similar to that of *So*HydG (see section 2.10).

The changes in EPR signal for *Ti*HydG His265Ala are much more dramatic, with the sample being completely EPR silent (data not shown). It is currently uncertain whether the RSAM conserved [4Fe-4S] cluster signal is absent due to the degradation of the protein during physical transport of the samples to the United States, or if it has some mechanistic significance. Some unknown structural changes could have occurred to prevent cluster reduction, but so far no proposed model for this event could be put forward. Such changes would definitely attribute a central role of His265 for the correct folding of the entirety of HydG's structure.

The UV spectra and iron quantification show results consistent with the presence of at least one $[4Fe\text{-}4S]^{2+}$ cluster (section 3.3.1), but the oxidative degradation of clusters can give rise to EPR silent $[2Fe\text{-}2S]^{2+}$ clusters in the protein [595]. However, the presence of dithionite in the system

would still reduce the two clusters to a paramagnetic $[2Fe\text{-}2S]^{+}$ observable by EPR [596]. Another degradation pathway could lead to the formation of an intermediate $[3Fe\text{-}4S]^{+}$ cluster, as observed in RSAM enzyme PFL-AE [96], which under strong reducing conditions could give rise to the EPR silent $[3Fe\text{-}4S]^{0}$ state [596]. The loss of the bridging sulfide together with the unique iron of the cluster would be necessary to give rise to these results and further studies are needed to clarify the lack of EPR signal of *Ti*HydG His265Ala.

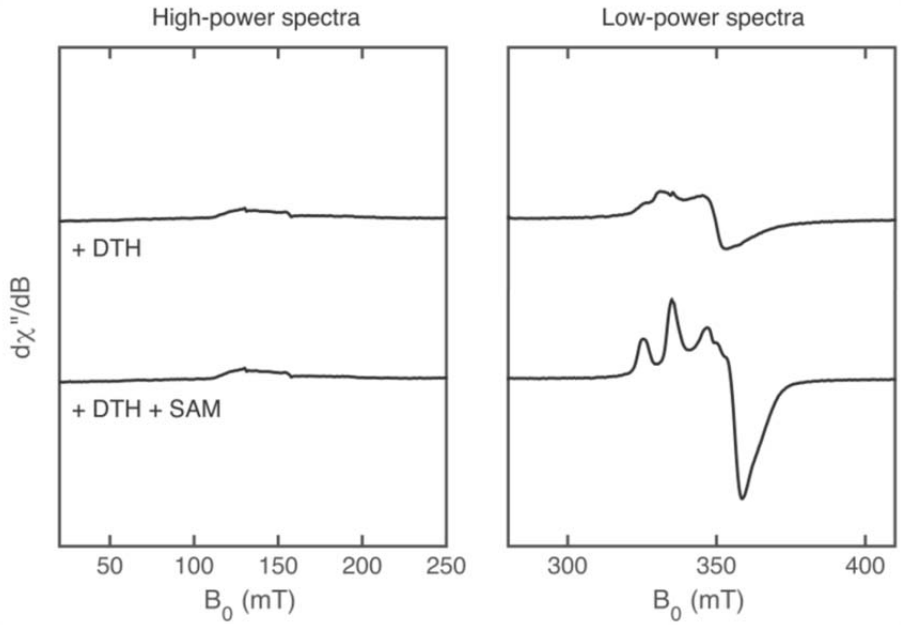



**Figure 90 EPR spectra of *Ti*HydG His265Glu.**

The low-power spectra is indicative of the presence of [4Fe-4S] clusters which, upon addition of SAM, increases complexity due to the interaction of SAM with the RSAM conserved cluster. At high power, the signal is similar to the S=5/2 proposed to arise from the $[5Fe\text{-}5S]_{Aux}$ cluster (Figure 82 C).

## 3.4 TiHydG His265 Mutants' Activity

To compare the activity of *Ti*HydG WT, *Ti*HydG His265Glu and *Ti*HydGHis265Ala, each enzyme (50 μM) was incubated with SAM (500 μM) and L-tyrosine (4000μ) in 50 mM Hepes buffer, pH 7.4 with 500 mM KCl at 37 °C for 60 minutes after initiating the reaction by addition of dithionite (1000 μM) (Method 9.7) [311]. An additional set of reactions were prepared in the absence of L-tyrosine to account for the uncoupled turnover of SAM [47]. The formation of products was quantified by use of DOA, *p*-cresol and cyanide standards (standard curves and HPLC traces are shown in Appendix A.6).

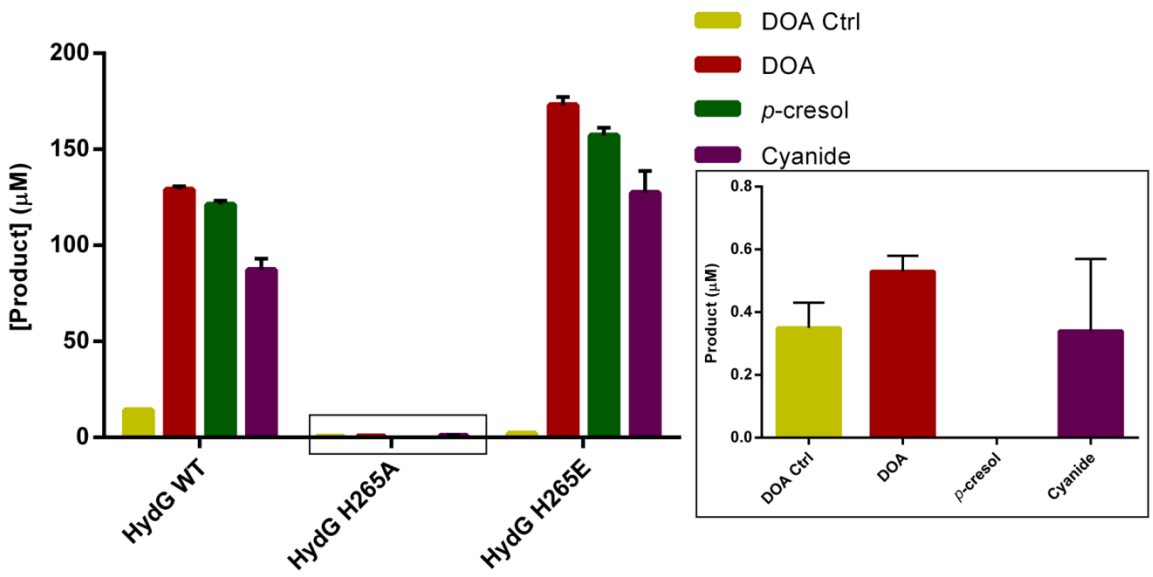



**Figure 91 Product quantification of the enzymatic reaction of HydG WT and His mutants.**

Quantification of the formation of DOA, *p*-cresol and cyanide by *Ti*HydG WT, *Ti*HydG His265Ala and *Ti*HydG His265Glu in the presence of SAM (DOA Ctrl) and SAM + L-tyrosine. A clear difference exists between *Ti*HydG His265Ala and the other two enzymes, with the prior having virtually no activity.

**Table 13 Quantification of product formation by *Ti*HydG WT and mutants.**

The apparent rate is not a kinetic measurement, but an estimation from the end-point result after 1h reaction. $K_{cat}^{app}$ was estimated directly from the linear rate.

| Enzyme | Product | ΔC (µM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number $k_{cat}^{app}$ (x$10^{-4}$ $s^{-1}$) |
|---|---|---|---|---|
| ***Ti*HydG WT** | DOA | 129.0 ± 1.7 | 35.8 ± 0.47 | 7.2 ± 0.1 |
| | *p*-cresol | 121.2 ± 2.1 | 33.7 ± 0.58 | 6.8 ± 0.1 |
| | cyanide | 87.2 ± 5.9 | 24.2 ± 1.64 | 4.8 ± 0.3 |
| | DOA (uncoupled turnover) | 14.3 ± 0.3 | 4.0 ± 0.1 | 0.8 ± 0.0(2) |

| Enzyme | Product | ΔC (μM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number $k_{cat}^{app}$ ($x10^{-4}$ $s^{-1}$) |
|---|---|---|---|---|
| ***Ti*HydG His265Ala** | DOA | 0.5(3) ± 0.0(5) | 0.1(5) ± 0.0(1) | 0.0(3) ± 0.0(03) |
| | *p*-cresol | 0 | 0 | 0 |
| | cyanide | 0.3(4) ± 0.2(3) | 0.0(9) ± 0.0(6) | 0.0(2) ± 0.0(1) |
| | DOA (uncoupled turnover) | 0.35 ± 0.1 | 0.0(9) ± 0.0(3) | 0.0(2) ± 0.0(06) |
| ***Ti*HydG His265Glu** | DOA | 173.0 ± 4.2 | 48.1 ± 1.2 | 9.6 ± 0.2 |
| | *p*-cresol | 157.2 ± 4.1 | 43.7 ± 1.1 | 8.7 ± 0.2 |
| | cyanide | 127.4 ± 11.4 | 35.4 ± 3.2 | 7.1 ± 0.6 |
| | DOA (uncoupled turnover) | 2.1 ± 0.2 | 0.6 ± 0.1 | 0.12 ± 0.0(2) |

The rates of reaction for *Ti*HydG are, as expected, quite similar to the previously estimated (section 2.4). As had been reported for *Tm*HydG [47], *Ti*HydG WT showed the formation of DOA, *p*-cresol and cyanide in an approximate 1:1:1 ratio. However, *Ti*HydG seems to show higher uncoupled turnover even in the absence of the substrate tyrosine (11% of the catalytic reaction), values higher than reported in both *Ca*HydG [47] and ThiH [46].

This high value for uncoupled SAM cleavage is unexpected in radical based enzymes under biological conditions, where multiple controls exist to reduce unwanted reactivity (see section 1.1.1.3) [94]. However, the use of a strong reducing agent to initiate the reaction, such as sodium dithionite, may cause some uncoupled turnover due to its reducing power being greater than the common biological partners ($E°_{Dithionite}$=-696 mV measured with *Av*Flavodoxin) [139]. Some additional differences between the results of *Ca*HydG and *Ti*HydG can possibly be explained due to the latter being a thermophilic enzyme, which in general show lower activity than their

mesophilic counterparts, in particular as the reaction was incubated at relatively low temperatures [540-542].

These differences between *Ca*HydG and *Ti*HydG enzymes could also explain why the $k_{cat}^{app}$ of DOA and *p*-cresol are much closer to one another in *Ti*HydG, with DOA $k_{cat}^{app}$= 7.2 ± 0.1 x$10^{-4}$ $s^{-1}$ and *p*-cresol $k_{cat}^{app}$= 6.8 ± 0.1 x$10^{-4}$ $s^{-1}$ (Table 13) than in the reactions with *Ca*HydG DOA $k_{cat}^{app}$= 23.0 ± 0.3 x$10^{-4}$ $s^{-1}$ and *p*-cresol $k_{cat}^{app}$= 18.0 ± 0.2 x$10^{-4}$ $s^{-1}$ [47]. These results suggest that unlike the reductive uncoupled turnover of SAM, which is higher in *Ti*HydG, the potential futile cycle with the formation of a tyrosyl radical (Figure 92) is seemingly less prevalent in HydG from *Thermoanaerobacter italicus* comparing with *Ca*HydG.

The rate difference between L-tyrosine cleavage and DHG breakdown is quite significant (Table 13), and can arise both due to a small, but non-zero, formation of glyoxylate, as previously reported [47, 311] as well as a non-complete derivatisation of cyanide.

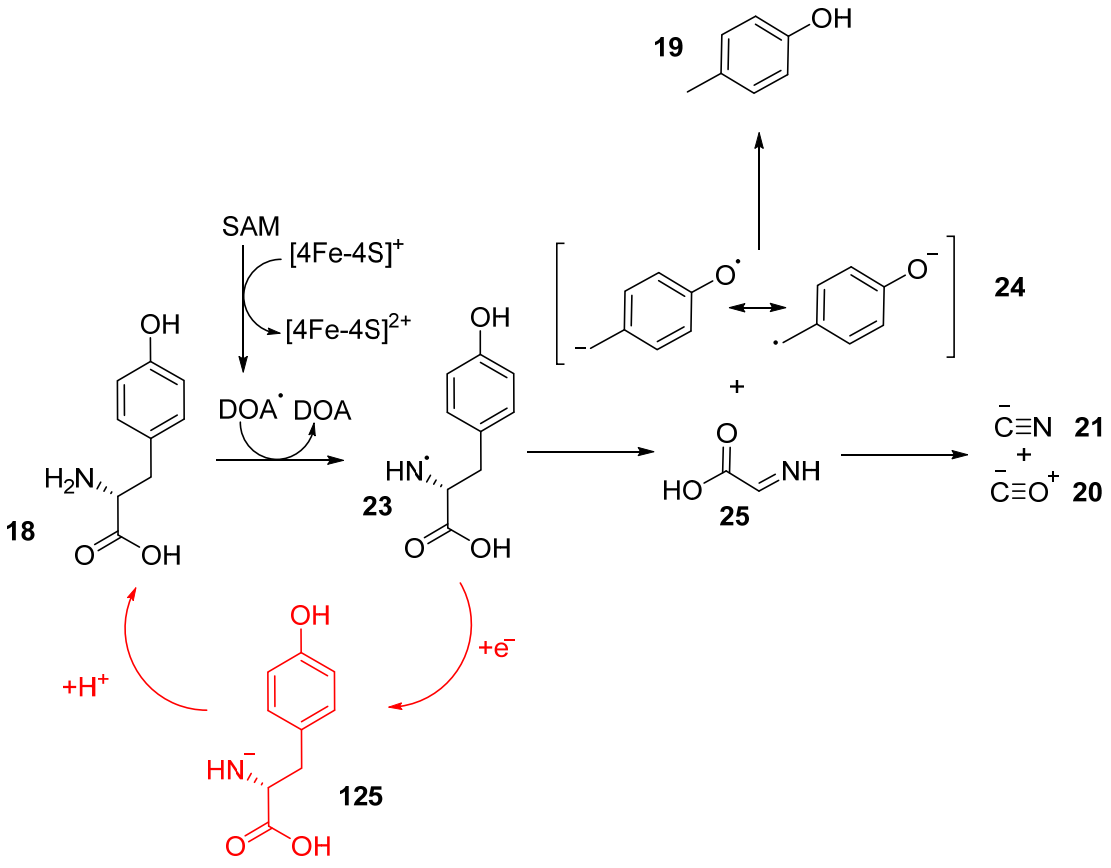



**Figure 92 HydG catalysed L-tyrosine cleavage, with futile cycle.**

Mechanism of HydG catalysed L-tyrosine cleavage, with the futile cycle by reduction and hydrogen re-abstraction shown in red. This pathway may occur in the presence of strong reducing agents.

The alanine mutant, with a presumed impaired Fe coordination (see sections 3.3.1and 3.3.2), shows virtually no turnover of substrates, with turnover numbers for the three measured products of the reaction close to zero (Table 13). The non-zero detection of cyanide is likely to have arisen from some background noise, as no *p*-cresol was detected, and the levels of cyanide presumably formed are below the method's quantification limit of 5 μM [47].

Such dramatic changes in both steps of the activity, shows the impaired coordination of the auxiliary cluster has a profound impact in product formation. The loss of RSAM related activity is unexpected, as even mutants where the C-terminus was removed by insertion of a stop codon before the 80 amino acid extension (ΔCTD) showed tyrosine cleavage at levels close to that of the WT enzyme in the presence of saturating substrates [311, 312]. Those results led Nicolet and colleagues [312] to propose the hypothesis that the auxiliary cluster was not required for the radical cleavage of L-tyrosine (see introduction 1.2.2.3.2). In spite of the ΔCTD mutation not affecting His265, it is bewildering that the mutation of one residue which coordinates an iron of the auxiliary cluster, proposed to be labile, has a more severe consequence for activity than the removal of the entire cluster. A possible explanation for this discrepancy arises from the general instability of the *Ti*HydG His265Ala enzyme (see section 3.3), which may have become partially unfolded and precipitated during the incubation period of the reaction.

Surprisingly the glutamic acid mutant seems to show higher apparent turnover of substrates than the wild-type enzyme (around 140% of native activity), as well as having a significantly smaller amount of uncoupled turnover in the absence of substrate ($DOA_U$ $K_{cat}^{app}$= $0.12x10^{-4}$ $s^{-1}$ for *Ti*HydG His265Glu and $0.8 x10^{-4}$ $s^{-1}$ for *Ti*HydG WT.

The low values of SAM cleavage without L-tyrosine estimated for the mutant enzyme are similar to the previous reported results for *Tm*HydG and ThiH [46, 47]. However, to the decreased uncoupled turnover in the absence of substrate, *Ti*HydG His265Glu also possesses similarly low uncoupled turnover numbers in the presence of L-tyrosine, with a SAM:tyrosine cleavage rate of 1.00 : 0.91, closer to the values of *Ti*HydG WT (1.0 : 0.94) than *Ca*HydG (1.00 : 0.77) [311]. The lower uncoupled turnover activity in the glutamic acid mutant enzyme strongly suggests the structural modification which occurred in the auxiliary cluster active site, directly affects the RSAM conserved reaction. A deeper kinetic analysis of this mutant enzyme in the future, together with structural comparisons in the presence of the substrates, could help understand in what manner the tighter control on the reaction occurs.

The increased cyanide formation when compared with the wild type enzyme (see Table 13) most likely hints at a direct DHG cleavage by the labile iron. It is, however, impossible to discard the possibility that the structural changes arising from the difference in labile iron coordination, which

lead to a higher L-tyrosine cleavage in *Ti*HydG His265Glu and complete loss of activity in *Ti*HydGHis265Ala, may have had a positive effect in the elimination reaction of DHG to form cyanide (see section 1.2.2.3.2, Figure 18).

Assuming DHG is cleaved by a mechanism dependent on the labile iron (as shown in section 2.11), two competing possibilities arise. The mutation can lead to a stronger coordination of the labile iron which allows it to withstand more turnovers before removing itself from the adequate position to capture DHG. Alternatively the labile iron can be bound more weakly, being more easily removed, allowing for a new cycle of the half reaction to occur more readily. In either case, it is possible that the maturation process will be impaired if attempted with HydG His265Glu. If the iron is too tightly bound, the rate of transfer of the labile iron to HydF will be decreased; whereas if the iron is bound too loosely, the correct coordination of the iron may not be achieved, and an incomplete subcluster may be formed in HydF.

Bond energies between metal atoms and an imidazole group are reportedly weaker (283 kJ $mol^{-1}$) than those with the carboxylic group of a benzoate (709 kJ $mol^{-1}$) [597]. These results are consistent with a stronger ionic interaction of the $COO^-$ with $Fe^2$ and correlate well with the iron quantification results (section 3.3.1) which showed a higher iron content for the *Ti*HydG His265Glu mutant when compared with the WT enzyme.

## 3.5 TiHydG His265Glu Mutant Crystallization and Data Analysis

Crystallographic studies using *Ti*HydG His265Ala were uniformly unsuccessful, as the yield of the protein is low, the ability to concentrate is very limited due to extensive protein aggregation and even when a crystallization tray was set up using 0.3 mM of enzyme (together with 3 mM SAM), almost immediate (2h) precipitation of the protein occurred in nearly all wells. From the observations during the purification of *Ti*HydG His265Ala during purification and during the limited crystallisation trials, is seems most likely that this mutant is significantly destabilized and unlikely to yield useful crystals.

The more stable *Ti*HydG His265Glu proved easier to crystallize. Crystals grew following a three-day incubation (in 2 out of 288 tested wells) in condition F1 from Molecular Dimensions PACT premier (0.1M Bis-Tris propane, pH 6.5, 0.2 M NaF, 20% w/v PEG 3,350) (Method 9.8.1), the same condition used for Ti*HydG* WT that allowed for the observation of a complete $[5Fe\text{-}5S]_{Aux}$ cluster. Growing the crystals of the mutant in a condition where the labile iron was identified was vital to probe the potential coordination of glutamate 265 to the iron.

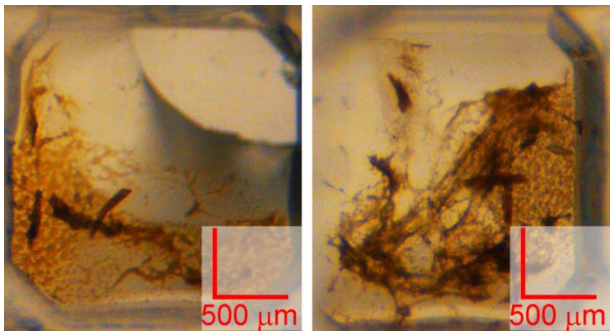



**Figure 93 Examples of crystallization wells showing the formation of *Ti*HydG His265Glu crystals.**

*Ti*HydG His265Glu crystals formed in the presence of 10 mM SAM, under condition F1 from the commercial screen PACT premier (0.1M Bis-Tris propane, pH 65.; 0.2M sodium fluoride; 20% PEG 3,350).

The crystals were collected and frozen in a manner similar to the wild type crystals. They were tested in Diamond Light Source facilities, beamline I02 and. both a high energy dataset, at 0.979 Å, as well as a dataset at energy values close to the measured κ-edge of iron were collected. The latter was used to observe the iron-only anomalous signal to unambiguously confirm the presence of the labile iron in the structure. Both datasets were processed automatically by X-ray Detector Software (XDS) [554], with data reduction/scaling done automatically using Pointless and SCALA [555]. All data quality and data refinement statistics are summarized in Annex 10.8, and show much of the same strengths of high resolution (1.75 Å), good I/sigma (8.5) and $R_{merge}$ (0.063) together with the same weaknesses of lower completeness (84%) and multiplicity (2.6).

Despite collecting an anomalous dataset to confirm the presence of the labile iron molecular replacement was the method used to create the model of *Ti*HydG His265Glu, with the mutation introduced in the search model prior to using the molecular replacement software *Phaser* [584]. The high solution scores (Appendix A.7) are indicative of highly isomorphous crystals of the mutant relative to the wild type (Figure 95), indicating that there was no extensive change in the structure caused by the mutation of His265.

Refinement of the data was carried out as discussed previously and POLYGON analysis reveals a structure that falls within the average values of Protein DataBase accepted structures (Figure 94).

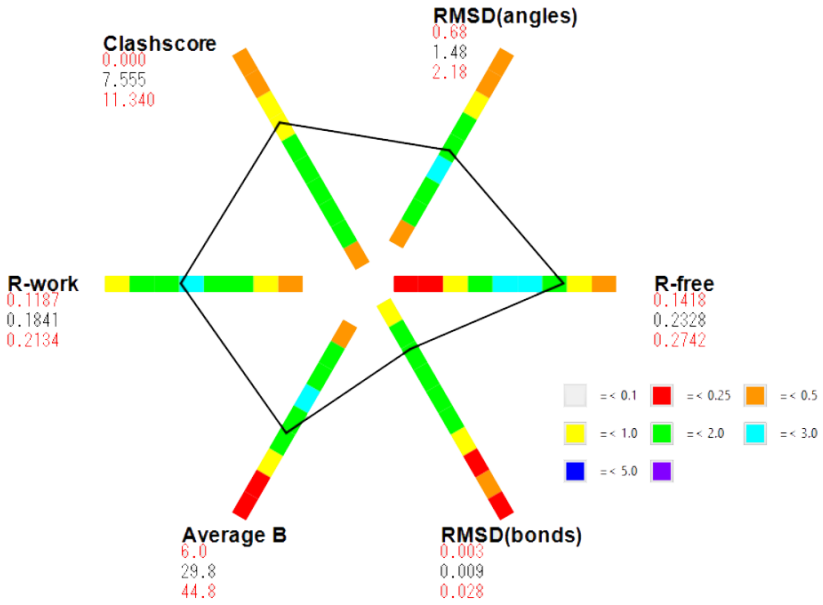



**Figure 94 POLYGON analysis of the refinement data for *Ti*HydG His265Glu crystal.**

Comparison of several quality indicators to structures with similar resolution within the PDB database. Histograms are coloured based on weighted number of structures with similar resolution.

### 3.5.1 *Ti*HydG His265Glu Structural Analysis

There is an extraordinarily high degree of structural similarity between HydG WT and the His265Glu mutant (average RMSD= 0.61 Å), but small changes can still be observed within both structures, mostly restricted to the disordered loop in monomer A (see section 2.7.2). These differences are likely to exist due to SAM being present in the active site of both monomers in contrast to the degradation of SAM observed in the wild type crystal (see section 2.7.2). It is interesting to note that the presence of (unreacted) SAM in both monomers in the crystal structure of the mutant may reflect the lower rate of uncoupled turnover observed for *Ti*HydG His265Glu (section 3.4). The structure does reveal the presence of the labile iron in the auxiliary cluster, being potentially coordinated by water molecules. $Water_A$ in the structure of *Ti*HydG His265Glu is positioned in a similar location to that of the coordinating $water_A$ in the *Ti*HydG WT structure (section 2.7.3), though in the His265Glu structure the $water_B$ binding site is presumably vacant (Figure 95 C). In place of the amino acid binding in a bidentate manner shown in the structure of *Ti*HydG WT, three waters were modelled, two in a position close to that of the amino and carboxylate groups. When observing the maximum likelihood map 2mFo-dFC density at a lower σ cutoff (Figure 95 C, red) it raises the possibility that another non-protein residue could be

bound to the labile iron, but the electron density is far too poorly resolved for an unambiguous allocation of the ligand, and thus the water molecules were put in place.

The anomalous data confirms (Figure 95 B) the position of the fifth iron of the auxiliary cluster, while the occupancy calculation reveals only 52% of the molecules within the crystal contain the elusive labile iron. The occupancy values correspond to a 30% loss of signal intensity when compared with the *Ti*HydG WT structure (see section 2.7.3).This decrease in observed occupancy may indicate an increased lability of the glutamate mutant. Moreover, this may provide a possible reason for the increased activity, as described in section 3.4. The faster dissociation of the labile iron from the protein after DHG cleavage could explain the higher levels of cyanide formed by *Ti*HydG His265Glu. Detailed kinetic analysis of turnover and transfer onwards to complete HydA cluster assembly will be required to determine if the His265Glu mutant is activated for just some, or all of the steps it needs to carry out in the maturation pathway.

Assuming the evolutionary pathway would have selected the most efficient residue for hydrogenase maturation to be present in HydG (His265) which is reasonable given its conservation, and assuming also that two events of DHG cleavage are to occur to form the $Fe(CO)_2CN$ (or similar) synthon [296], the higher lability of the iron when Glu265 is present may lead to an incomplete loading of the iron with enough diatomic ligands and/or premature release of the partially assembled synthon. The maturation process might then stall, or might create an H-subcluster without the full complement of inorganic ligands required to activate HydA.

It would be unwise, however, to presume that a single set of crystallographic data is solid enough evidence for an accurate estimation of the interaction between the labile iron and the protein. As an X-ray structure is modelled from an average of the electron density of all the molecules that constitute the crystal, small changes in the crystallization method can have a strong effect in the iron content. Discrepancies in the quality of the crystal data can also be responsible for differences in the calculated occupancies. In fact, structures with worse than atomic resolution are often insufficient to obtain precise information on atom occupancy and location [598], compounding even more the issue. Additionally the more ordered hexa-coordination of the labile iron in the wild type structure, with a clear amino acid together with two water molecules bound (see section 2.7.3) may introduce an increased stability on the position of the atom, which would also justify the differences in occupancy levels.

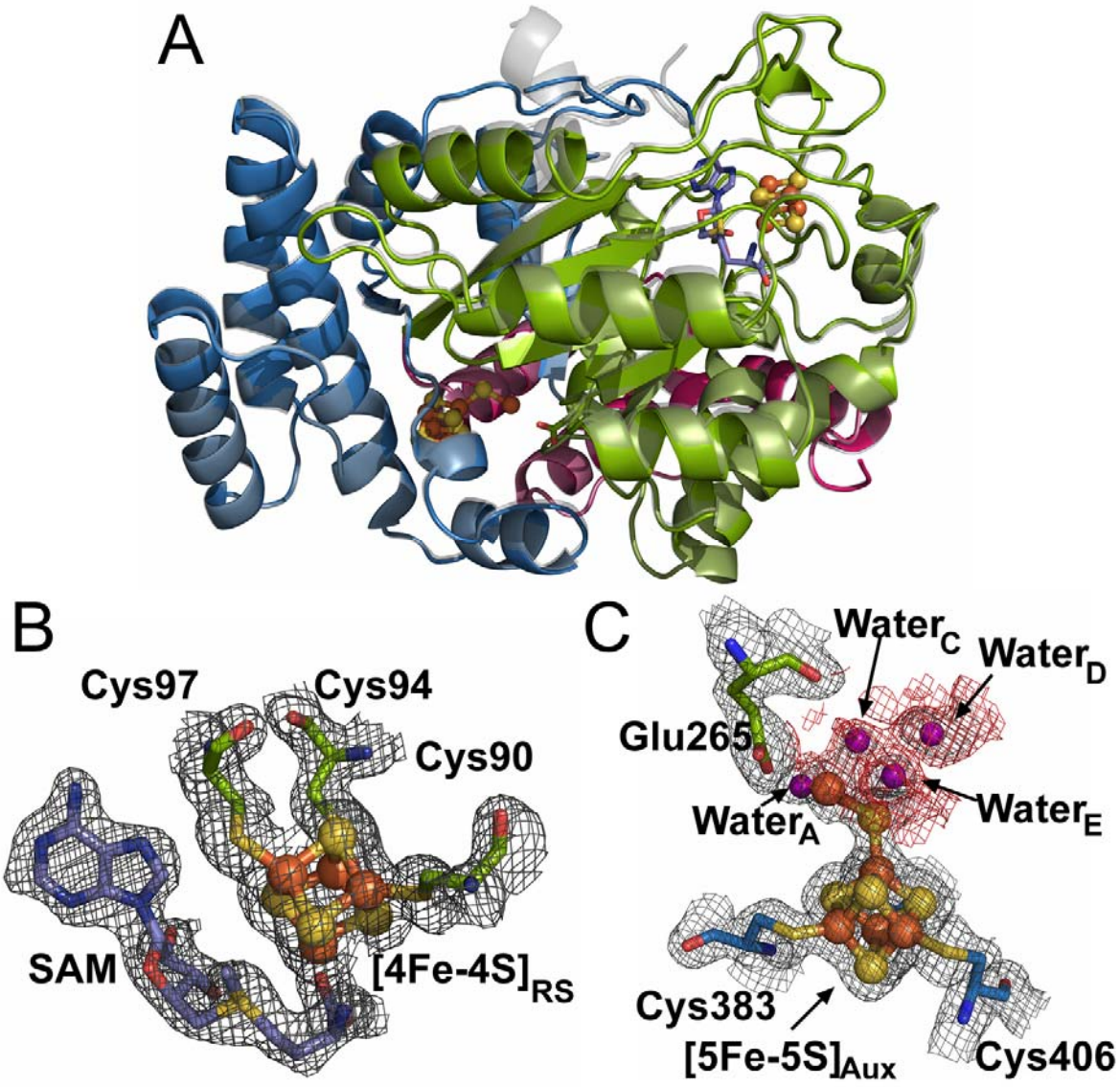



**Figure 95 Structural elements of *Ti*HydG His265Glu.**

**A:** Overall fold of *Ti*HydG His265Glu, with *Ti*HydG WT overlaid in grey. The structures show a near-identical level of similarity, apart from the partial disordered C-terminus loop. **B:** Iron anomalous Fourier difference map, contoured at 5.0 σ, for the auxiliary cluster of *Ti*HydG His265Glu. **C:** The 2Fo-Fc map, contoured at 1.0 σ, around the RSAM cluster and coordinating cysteines, SAM bound. **D:** In grey the 2Fo-Fc map, contoured at 1.0 σ, around the $[5Fe\text{-}5S]_{Aux}$ cluster, the three coordinating cysteines, Glu265 and four potential water molecules coordinating the labile iron. In red the 2Fo-Fc map surrounding waters C, D, E, contoured at 0.5 σ.

## 3.6 *Ti*HydG His265 Mutants Secondary Structure Estimation by Circular Dichroism

### 3.6.1 Principles of Circular Dichroism for Proteins

A chiral molecule absorbs left- and right-circularly polarized light to different extents and have different indices of refraction for the two waves. Circular dichroism (CD) is a light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light by the asymmetric (chiral) molecule (for full reviews on the technique see [599-602]).

For proteins, the absorption occurs mostly in the ultraviolet region of the spectrum due to the nature of the chromophores present in any given protein: the peptide bond, the amino acid side chains and any prosthetic groups. Different secondary structures of peptide bonds have different spectra at far UV regions (190-250 nm), while the spectra at near UV (250-300 nm) is sensitive to disulfide bonds as well as several aromatic side chains, in particular phenylalanine (signals from 250-270 nm); tyrosine (270-290 nm); and tryptophan (280-300 nm). However, the contributions to the CD spectra by the aromatic side chains in the far UV regions (where secondary structural information is located) in proteins is usually negligible [603, 604].

If any number of chromophores from the same type, such as particular folds in the protein backbone, are located in close proximity to one another, they can behave as a single absorbing unit, leading to the formation of a characteristic CD signal [605]. For example, α-helical proteins have negative bands at 222 nm and 208 nm and a positive band at 193 nm [606]; β-strands have a negative band at 218 nm and a positive band at 195 nm [607]; while random coils have a low signal above 210 nm and a negative band near 195 nm [608]. The aim of a protein CD experiment is to be able to deconvolute the sum of these signals in order to obtain a percentage of the structural features. In the region of 230-178 nm, one expects to observe effects of backbone conformational changes while CD effects at longer wavelengths (>230 nm) should isolate aromatic chromophore contributions and being environment dependent should reflect more global, three-dimensional properties of the protein [603-605].

### 3.6.2 Circular Dichroism Analysis of *Ti*HydG WT and Mutants

The difficulties encountered during the purification of *Ti*HydG His265Ala makes it challenging to explain its absolute lack of activity (section 3.4) and EPR signal (section 3.3.2). To observe whether there were significant changes to the enzyme's secondary structure when a mutation is

introduced, circular dichroism was used in *Ti*HydG WT, *Ti*HydG His265Ala and *Ti*HydG His265Glu samples. Samples were collected immediately following their purification with nickel affinity chromatography, which are termed herein 'not reconstituted' samples, as well as samples which were chemically reconstituted after size exclusion chromatography, termed 'reconstituted' samples. Analysis by CD of not reconstituted and reconstituted samples was anticipated to provide some insight on the structural changes caused by the formation (or lack thereof) of the clusters in these HydG proteins.

Circular dichroism is a very sensitive technique, with most of the common biochemical buffers usually used for protein storage being incompatible due to interference of the buffer molecules, in particular NaCl in the far UV region where protein absorbance is also observed [604]. On the other hand, the instability of the *Ti*HydG His265Ala mutant meant a buffer exchange to a system with phosphate buffer, in the absence of glycerol or high salts, would be unsuitable and likely lead to immediate protein precipitation.

Taking advantage of the high sensitivity of the method and the relatively high concentration of the proteins under study (up to 1.0 mM in the case of *Ti*HydG WT and His265Glu; and 0.3 mM for *Ti*HydG His265Ala), the protein samples were diluted into 50 mM phosphate buffer, pH 7.4, to a final concentration of 1.5 μM prior to measuring the CD spectra (see Method 9.15). The data was obtained in a Jasco J-720 spectrophotometer, together with a standard without protein as a blank, which was subsequently subtracted from the samples' chromatograms using JASCO 1.10 software. Prior to data analysis, the results were standardized from degrees of ellipticity values into mean residue ellipticity [609], which allows the direct comparison between the different mutants. The conversion of the results followed equation 1:

$$[\theta]_{mrw} = \frac{100 \times \theta}{c \times l \times n}$$

**Equation 4 Calculation of mean residue ellipticity.**

θ is the ellipticity obtained from the spectrophotometer (degrees); c is the protein concentration (mM); l is the path length (1 cm) and n is the number of residues.

Data analysis was carried out using the online deconvolution system Dichroweb [610], using the available algorithm $K_2D$ [611]. The choice of algorithm is an important step in CD data analysis, and $K_2D$ was chosen to analyse HydG's data mainly due to experimental constraints. The spectrophotometer used did not allow for a scan into the far UV region of the spectra (175-200 nm) and the $K_2D$ algorithm is the most suitable for the smaller spectral window where the HydG data was collected.

The structural information collected in the crystal structures of *Ti*HydG WT and *Ti*HydG His265Glu, which show an equivalent fold and overall structure (see section 2.7.1, Figure 54) was used to calculate the percentage of secondary structures alpha-helix, beta-strand and random coil (Table 14, crystallographic values). The values calculated in Dichroweb for the reconstituted samples of *Ti*HydG WT and *Ti*HydG His265Glu agree to a large extent with the X-ray structure, and the spectroscopic profile between these two proteins is also similar (Figure 97 A2 and B2).

The *Ti*HydG His256Ala mutant is very unstable without a chemical reconstitution, and the data obtained possessed high background noise (Figure 97 C1), leading to a poor fit of the data with the calculated error above the values accepted by the $K_2$d algorithm (Table 14). The high error associated with the non-reconstituted sample of *Ti*HydG His265Ala does not allow for any meaningful conclusions to be taken from the spectra. The chemical reconstitution improves the stability of this mutant, though the secondary structure predicted is different from that obtained for the wild type and glutamic acid mutant, as can be seen both by a large increase of the β-strand present (Table 14), as well as the morphology of the spectra (Figure 97 C2). Residues in the C-terminal region account for 35% of the total structure of the protein, thus it is not unreasonable to suppose the increase in 10-15% of β-strand content may be limited to the cluster-free C-terminus.

Interactions between β-strands are known to cause protein aggregation by the creation of a network of hydrogen bonds between different sheets [612], and can lead to the formation of Aβ amyloid plaques which have been implicated in neurodegenerative diseases [613-615]. The sharp rise in β-strand content in *Ti*HydG His265Ala may be causing the formation of a similar stacking unit (Figure 96), leading to the extensive aggregation observed in the purification stages of the enzyme (section 3.3).

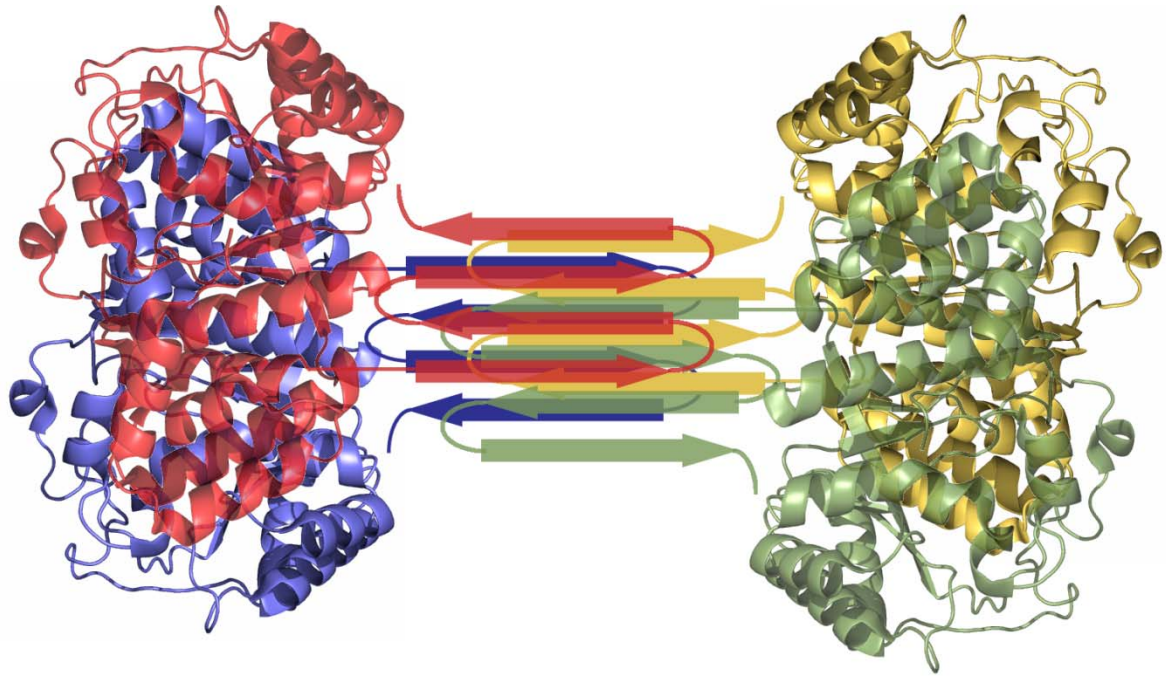


**Figure 96 Potential model for *Ti*HydG His265Ala aggregation, by stacking of C-terminal β-sheets.**

N-terminus and RSAM core fold were modelled from the structure of wild type *Ti*HydG (PDBID: 4WCX), with the C-terminal end being represented as a drawing of β-strands (arrows). Four anti-parallel beta strands were stacked on top of one another to mimic protein aggregation units.

It is possible that without the anchoring of the labile iron to the end of the RSAM domain, the entire C-terminus in *Ti*HydG His265Ala changed into a different conformation, leading into a fold where L-tyrosine cleavage cannot occur, which could explain the abolishment of activity (section 3.4). In the future, a comparative analysis of *Ti*HydG His265Ala CD spectra with the CD spectra of HydG unable to coordinate the auxiliary cluster (cysteine to serine mutations in the triad that coordinates the auxiliary cluster have been reported [312] could give credence to this significant change of secondary structure.

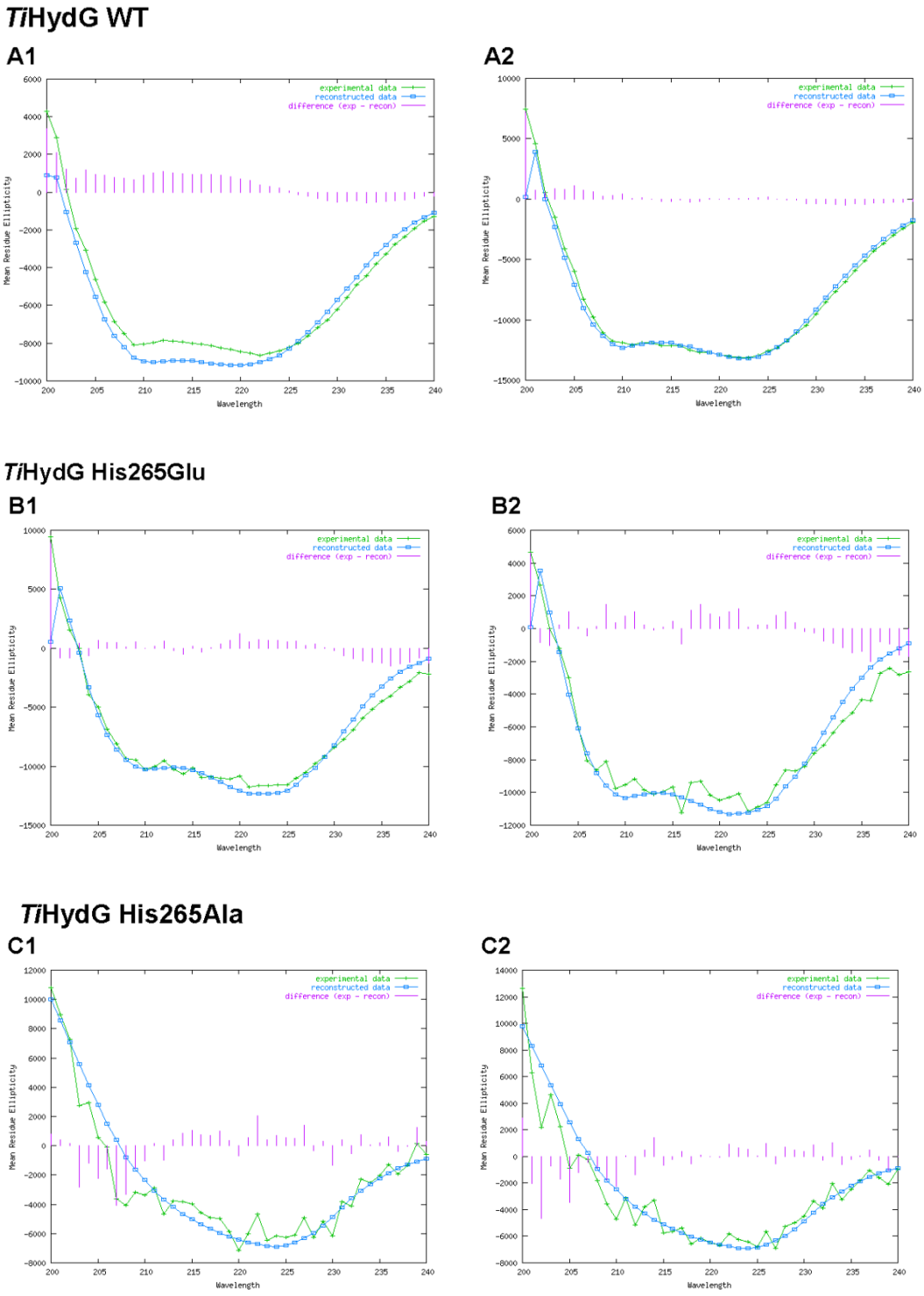



**Figure 97 Circular dichroism spectroscopic data (green) and predicted modelling (blue) using the K2d algorithm [611].**

**A:** *Ti*HydG WT samples. A1 – Sample which was not chemical reconstituted. A2 – Chemically reconstituted sample. **B:** *Ti*HydG His265Glu samples. B1 – Sample which was not chemical reconstituted. B2 – Chemically reconstituted sample. **C:** *Ti*HydG His265Ala samples. C1 – Sample which was not chemical reconstituted. C2 – Chemically reconstituted sample.

**Table 14 Predicted secondary structure values obtained by the deconvolution of the circular dichroism data for *Ti*HydG wild-type and mutants.**

The predicted secondary structure contents were calculated following the structure available for *Ti*HydG (section 2.7.1). The $K_2D$ algorithm fits predicted data with different content of alpha-helix, beta-strands and random coil, and the least square distance fitting is used. A calculated error is presented, with values above 0.22 being considered too high for a correct fitting.

| | | α-helix (%) | β-strand | Random coil | Least Square Distance | Maximum Error |
|---|---|---|---|---|---|---|
| Crystallographic values | | 50 | 9 | 41 | | |
| *Ti*HydG WT | Not reconstituted | 32 | 14 | 53 | 75.3 | 0.18 |
| | Reconstituted | 40 | 10 | 50 | 7.61 | 0.08 |
| *Ti*HydG His265Glu | Not reconstituted | 40 | 4 | 56 | 31.00 | 0.08 |
| | Reconstituted | 37 | 9 | 55 | 65.6 | 0.122 |
| *Ti*HydG His265Ala | Not reconstituted | 9 | 43 | 47 | 110.4 | >0.22 |
| | Reconstituted | 45 | 23 | 31 | 89.1 | 0.18 |

## 3.7 Conclusions

To assess the importance of the coordination of the labile iron of the auxiliary cluster, two mutants were designed, expressed and purified. , Changes focused on the auxillary cluster ligand histidine 265, with site directed mutants to either alanine or glutamate being prepared. From the purification stage onwards, it was clear whether the mutation had high impact in the stability of the protein, with *Ti*HydG His265Glu showing expression levels similar to the *wild type* and *Ti*HydG His265Ala appeared unable to fold correctly, leading to a predisposition to precipitation as well as lower iron coordination. The precipitation of the alanine mutant rendered any crystallographic studies practically impossible, but the CD spectra of its samples confirmed a significant change in the secondary structure content when compared with the *wild type* and glutamic acid mutant, stressing the importance of the anchoring of the auxiliary cluster by the histidine residue.

The activity results with *Ti*HydG His265Glu did not provide a definitive answer regarding whether the auxiliary cluster is directly responsible for the breakdown of DHG into both cyanide and carbon monoxide. The change introduced is still able to coordinate the labile iron, as shown by the crystallographic results, and it does result in an increased ability to form cyanide. However, the levels of L-tyrosine cleavage were also increased significantly by the mutation of His265Glu and it is possible that the overall influence of the glutamate leads to more DHG being degraded through the cluster independent elimination mechanism in section 1.2.2.3.2 Figure 18.

The activity of *Ti*HydG His265Ala would theoretically help in answering this problem, as the lack of cyanide formation would certainly imply the necessity of the labile iron. Regrettably the alanine mutant showed a complete lack of activity for SAM cleavage, tyrosine cleavage or cyanide formation. These results are slightly surprising, as UV-Vis spectra still appears to detect the presence of [4Fe-4S] clusters, and previous mutagenesis studies had shown L-tyrosine cleavage is independent of the auxiliary cluster [311, 312].

Nonetheless, such an impact in the enzyme when the iron coordinating residue is changed gives at least direct evidence for the central role of the labile iron which seems to serve as an indirect support for the mechanism of loaded iron transfer (section 2.11).

# Chapter 4: *Ti*HydG Activity With Substrate Analogues

## 4.1 Introduction

The solution of *Ti*HydG's structure with both clusters intact has updated HydG's trifunctional enzymatic mechanism [317] (section 2.11), leading to a greater insight into the second half of the reaction, where DHG is cleaved onto the diatomic products. While the elimination mechanism, with formation of free cyanide independent from the auxiliary cluster (see section 1.2.2.3.2 Figure 18) [317] is difficult to rigorously disprove, the proposed model of DHG coordination to the labile iron, mimicking the amino acid coordination in the structure, has some advantages.

Firstly, and due to its octahedral coordination, the presence of the labile iron can justify the binding of up to two DHG molecules, which in turn, may help to explain the observation of a $Fe(CO)_2CN$ species [296]. Secondly, if the cleavage mechanism of DHG occurs in the [5Fe-5S] auxiliary cluster, leaving the $CO/^{-}CN$ diatomic products Fe-bound, the completed synthon, currently proposed as '$Fe(CO)_2CN$ ' is ready for the subsequent steps of hydrogenase maturation. This may include transfer to HydF and HydG is known to be the source of the iron present in HydF during the maturation step [296]; whereas if free cyanide is formed, a mechanism of cyanide capture by Fe is still necessary, a step for which no proposal or evidence has arisen.

Probing the reaction with substrate analogues could help further elucidate the understanding of the entire mechanistic model of reaction. Despite this, so far, modifications in both the phenolic –OH (such as phenylalanine and derivatives) or in the amino/carboxylate functional groups have led to a complete loss of activity [319], indicating a high degree of substrate specificity on part of HydG.

Additionally, the identification of substrate analogues is an often used methodology to circumvent difficulties in co-crystallizing with the native substrate. Furthermore, understanding which modifications within the substrate change or maintain the native reaction can give important clues regarding the active site and the entire enzymatic mechanism of action.

Studies in substrate analogue activity with the related tyrosine lyase ThiH tested several compounds (Figure 98), showing that both the carboxylic acid moiety as well as the phenolic hydroxyl groups are necessary for substrate binding. Both *p*-hydroxyphenyl-α-hydroxypropionic acid (4-HPHPA) **127** and *p*-hydroxyphenylpropionic acid (4-HPPA) **128** significantly increased SAM cleavage when compared with a control without substrate (from 6% to 50% increase) [46]. Interestingly, *p*-hydroxyphenylcinnamic acid was not successful in promoting SAM cleavage **129**,

possibly indicative of lack of significant binding by this analogue, and perhaps indicating the $C_\alpha$-$C_\beta$ bond requires some flexibility to reach the active site [46]. Furthermore, 4-substitutred phenylalanine analogues **132-134** and phenylalanine itself **131** led to no activity reported by the enzyme, which was interpreted as the abstraction point occurring in the –OH of the phenol ring of tyrosine [46]. The authors noted a lack of activity even in the presence of 4-fluorophenylalanine **133**, giving some support to their mechanism of action, as fluorine has 97% of the van der Wals radius of oxygen and is also electronegative enough to create hydrogen bonds [46]. These studies, however, did not fully disprove the possibility of an amino hydrogen abstraction, and the structure of the related tryptophan lyase NosL, with tryptophan bound, seems to indicate the latter to be the mechanism of substrate binding in amino acid lyases [321].

Following these studies in ThiH [46], the same tyrosine analogues were tested by Dr. R. Driesener in HydG [519], together with some ortho-substituents on the tyrosine phenolic ring **135, 136**. Unlike ThiH, no analogues tested revealed a significantly increased DOA formation [519], showing that HydG's active site is even more selective than that of tyrosine lyase ThiH, including having a stronger stereoselectivity, as SAM cleavage in the presence of D-tyrosine was the same as the negative control in HydG [519], while double in ThiH [46].

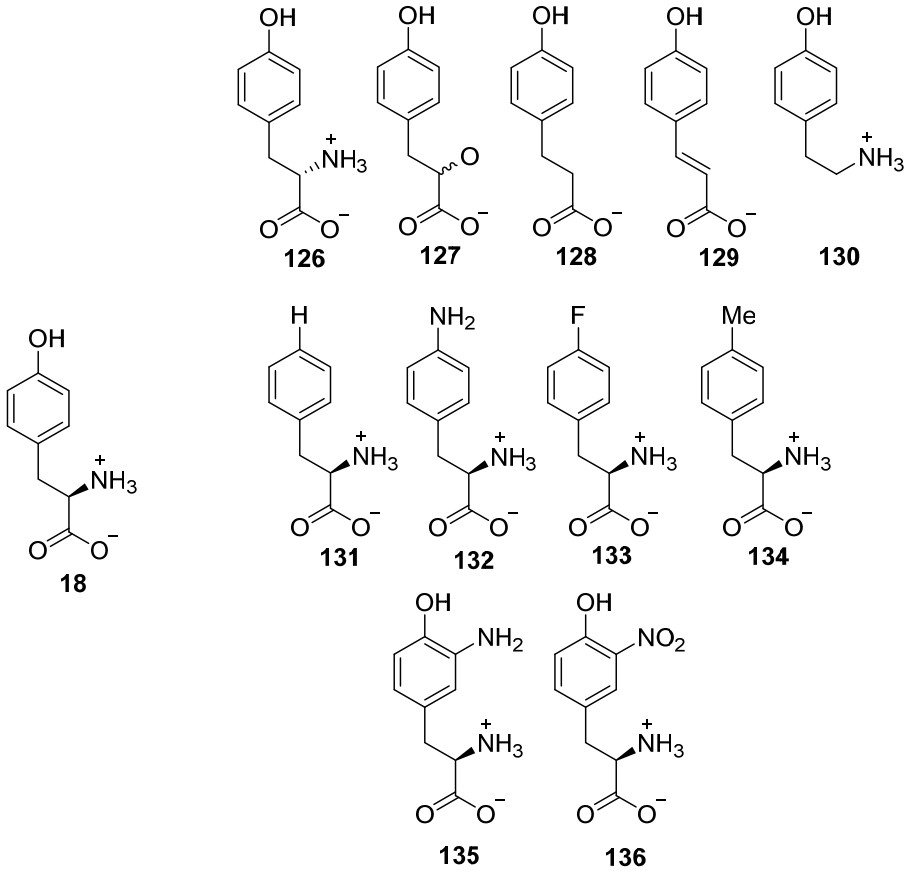



**Figure 98 Tyrosine analogues previously used in ThiH and HydG activity assays.**

## 4.2 *Ti*HydG Activity With Substrate Analogues

The subtle sequence variations naturally occurring in enzymes from organism to organism can sometimes give rise to differences in substrate specificity [542, 616]. While many of the most relevant residues for activity are conserved, there is enough variation within the active sites that justifies sometimes surprising changes between source organisms. The reactivity of *Ti*HydG was investigated with the most promising analogues derived from the studies of both ThiH and *Tm*HydG: tyramine **130**, 4-HPHPA **127** and 4-HPPA **128**, together with four new analogues: L-DOPA **137**, L-tyrosinamide **138**, N-methyl-L-tyrosine **139** and L-7-Hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid **140** (HTHIQCA) (Figure 99).

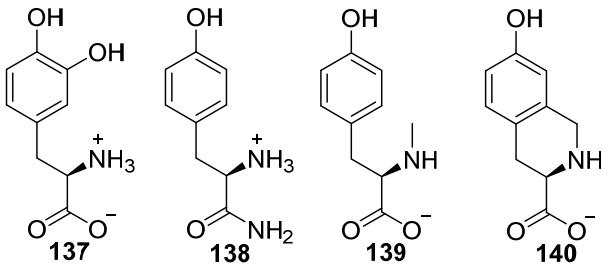



**Figure 99 New tyrosine analogues tested in HydG activity assays**

L-DOPA **137**, L-tyrosinamide **138**, N-methyl-L-tyrosine **139** and L-7-Hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid **140** (HTHIQCA).

Regardless of the substrate analogue used, four general results can be predicted from probing HydG's activity with L-tyrosine analogues: no activity is observed, indicating that the modification interfered with the ability of the compound to reach the active site; only DOA formation is observed, showing the compound interacts with the enzyme in such a way that it increases SAM turnover, but no Cα-Cβ scission is observed, perhaps due to a futile cycle being in place (Figure 100) or because the hydrogen atom abstraction step fails (and the substrate analogue can be considered a competitive inhibitor); DOA and *p*-cresol (or related compound) are observed as products of the reaction, showing the hydrogen abstraction by the SAM derived radical leads to a $C_\alpha$-$C_\beta$ bond cleavage of the substrate analogue in a manner similar to that of the native substrate; and finally, DOA, *p*-cresol (or related compound) and cyanide are observed, indicating an effective substrate analogue that most likely mimics the main pathway of reaction.

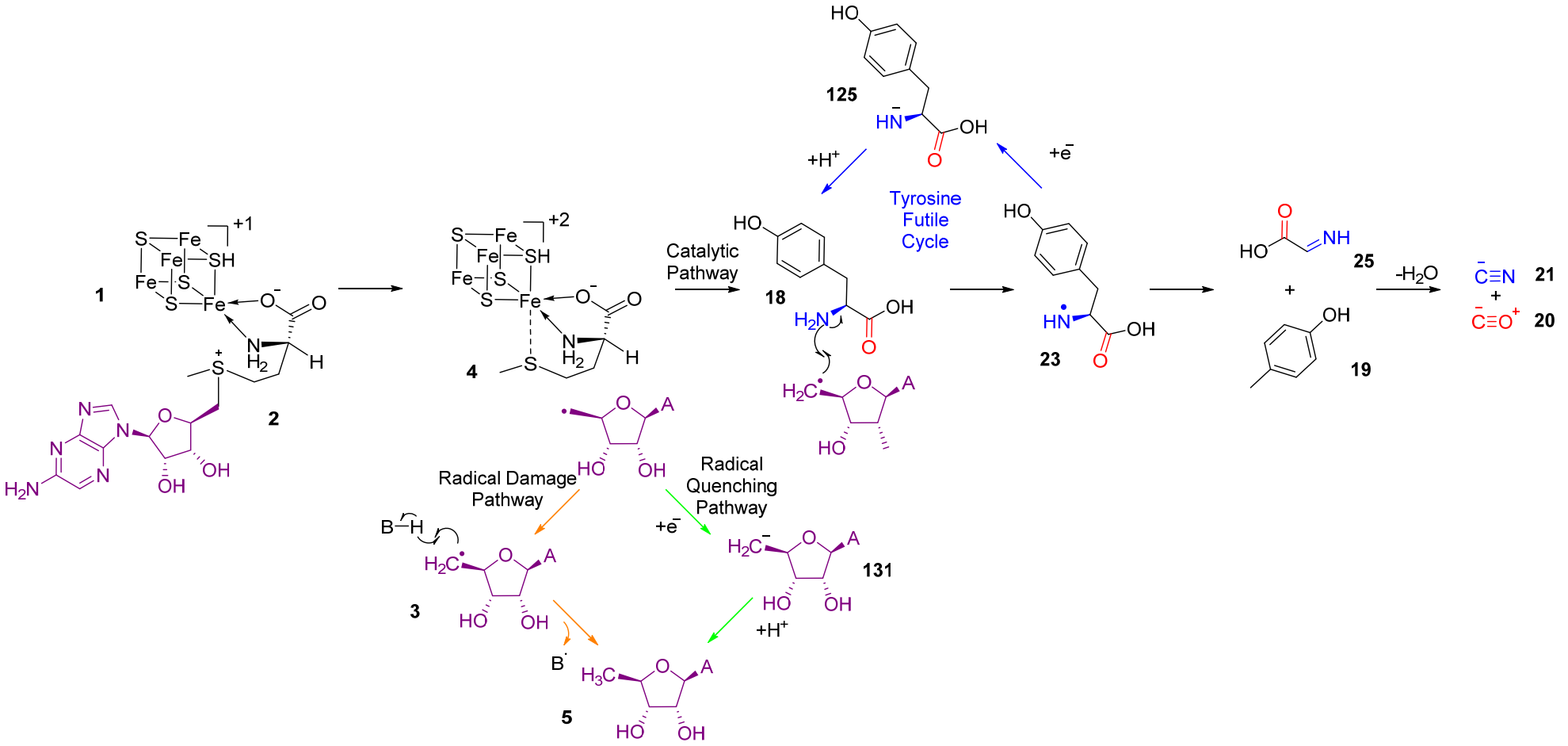



**Figure 100 Mechanism for product formation by HydG, together with potential by-product pathways**.

Black arrows show the mechanism of product formation. In orange, the DOA radical damage pathway. The newly formed, amino acid radical is either reduced and protonated, to quench the radical or the radical is propagated, potentially inactivating the enzyme. The DOA radical quenching pathway is shown with green arrows. Finally the substrate-based radical can be reduced and protonated (tyrosine futile cycle) shown in blue.

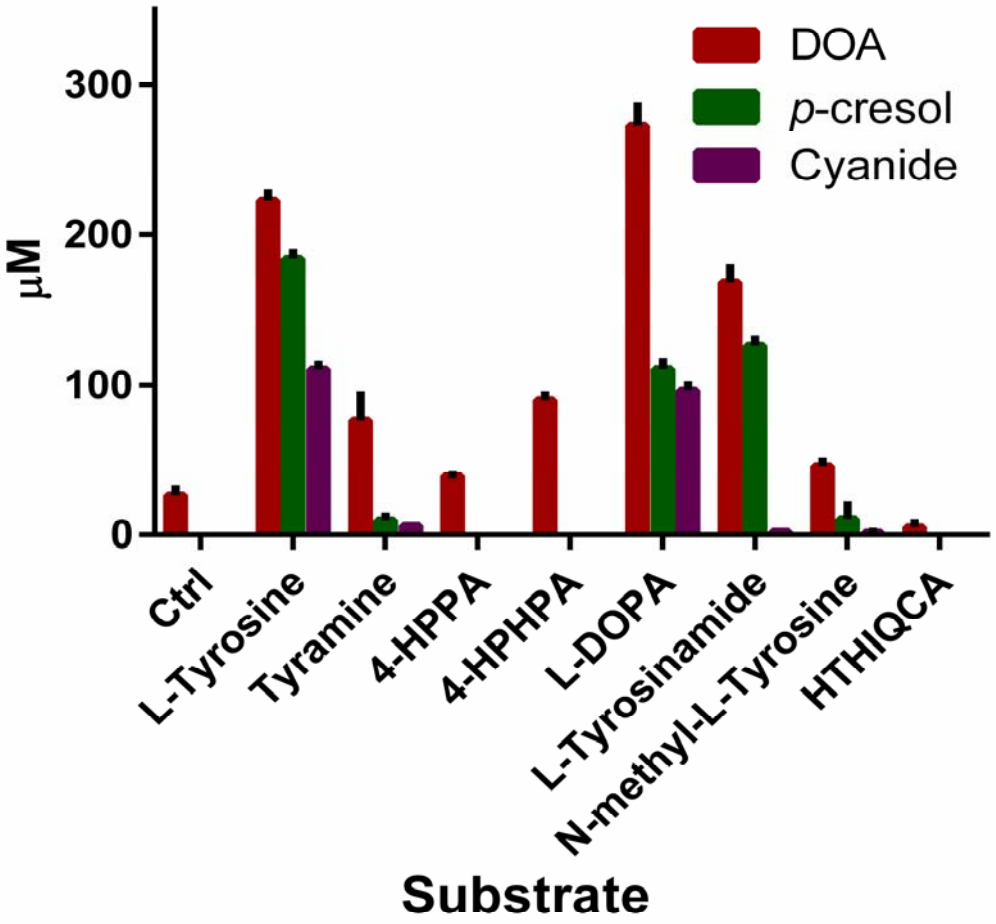



**Figure 101 Quantification of products in *TiHydG* activity assay in the presence of L-tyrosine analogues.**

**Table 15 Quantification of product formation by *Ti*HydG with different substrate analogues.**

The apparent rate is not a kinetic measurement, but an estimation from the end-point result after 1h reaction, assuming linearity of the reaction.

| Substrate | Product | ΔC (μM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number $k_{cat}^{app}$ ($x10^{-4}$ $s^{-1}$) | Reaction Ratio (SAM : Substrate : Intermediate) |
|---|---|---|---|---|---|
| L-tyrosine **18** | DOA | 222.5 ± 4.8 | 61.8 ± 1.3 | 12.4 ± 0.3 | 1.00 : 0.83 : 0.50 (: 0.03) |
| | *p*-cresol | 184.2 ± 3.4 | 51.2 ± 0.9 | 10.2 ± 0.2 | |
| | Cyanide | 110.8 ± 3.0 | 30.8 ± 0.8 | 6.2 ±0.2 | |
| | DOA (Uncoupled turnover) | 26.1 ± 4.0 | 7.3 ± 1.1 | 1.5 ± 0.2 | |
| Tyramine **130** | DOA | 76.8 ± 4.8 | 21.3 ± 1.3 | 4.3 ± 0.3 | 1.00 : 0.13 : 0.01 |
| | *p*-cresol | 9.9 ± 2.6 | 2.8 ± 0.7 | 0.6 ± 0.1 | |
| | Cyanide | 0.6 ± 0.1 | 0.2 ± 0.0(3) | 0.03 ± 0.0(1) | |
| 4-HPPA **128** | DOA | 39.3 ± 0.9 | 10.9 ± 0.3 | 2.2 ± 0.1 | 1.00 : 0.00 : 0.00 |
| | *p*-cresol | 0.0 | 0.0 | 0.0 | |
| | Cyanide | 0.0 | 0.0 | 0.0 | |

| Substrate | Product | ΔC (μM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number $k_{cat}^{app}$ ($x10^{-4}$ $s^{-1}$) | Reaction Ratio (SAM : Substrate : Intermediate) |
|---|---|---|---|---|---|
| 4-HPHPA **127** | DOA | 90.0 ± 3.3 | 25.0 ± 0.9 | 5 ± 0.2 | 1.00 : 0.00 : 0.00 |
| | *p*-cresol | 0.0 | 0.0 | 0.0 | |
| | Cyanide | 0.0 | 0.0 | 0.0 | |
| L-DOPA **137** | DOA | 273.1 ± 12.7 | 75.9 ± 3.5 | 15.2 ± 0.7 | 1.00 : 0.41 : 0.35 |
| | Methylcatechol | 110.9 ± 4.6 | 30.8 ± 1.3 | 6.2 ± 0.3 | |
| | Cyanide | 96.4 ± 0.1 | 26.8 ± 0.0(3) | 5.4 ± 0.0(1) | |
| L-Tyrosinamide **138** | DOA | 168.2 ± 10.0 | 46.7 ± 2.8 | 9.3 ± 0.6 | 1.00 : 0.75 : 0.01 |
| | *p*-cresol | 126.1 ± 4.2 | 35.0 ± 1.2 | 7 ± 0.2 | |
| | Cyanide | 2.2 ± 0.1 | 0.61 ± 0.0(3) | 0.12 ± 0.0(1) | |
| N-methyl-L-tyrosine **139** | DOA | 46.0 ± 3.2 | 12.8 ± 0.9 | 2.6 ± 0.2 | 1.00 : 0.23 : 0.03 |
| | *p*-cresol | 10.5 ± 9.5 | 2.9 ± 2.6 | 0.6 ± 0.5 | |
| | Cyanide | 1.5 ± 0.8 | 0.4 ± 0.2 | 0.1 ± 0.0(4) | |

| Substrate | Product | ΔC (μM) | Apparent linear rate (nM $s^{-1}$) | Apparent turnover number $k_{cat}^{app}$ (x$10^{-4}$ $s^{-1}$) | Reaction Ratio (SAM : Substrate : Intermediate) |
|---|---|---|---|---|---|
| HTHIQCA **140** | DOA | 5.2 ± 2.8 | 1.4 ± 0.8 | 0.3 ± 0.2 | 1.00 : 0.00 : 0.00 |
| | *p*-cresol | 0.0 | 0.0 | | |
| | Cyanide | 0.0 | 0.0 | | |

As expected, by using the native substrate L-tyrosine **18**, the formation of both DOA and *p*-cresol is observed at the highest rate and in an approximate 1:1 ratio, serving as a positive control for the assay. The reaction of reduced *Ti*HydG with SAM, in the absence of substrate, shows the uncoupled turnover of the enzyme (Figure 100, green arrows) to produce DOA.

Showing results similar to the ones reported for ThiH [46] and largely dissimilar to the results reported for *Ca*HydG [519], both 4-HPPA **118** and 4-HPHPA **119** show a rate of SAM cleavage significantly higher than the negative control (5-fold higher in the case of 4-HPPA and 12 fold higher in the reaction with 4-HPHPA). However, SAM turnover was not accompanied by $C_\alpha$-$C_\beta$ cleavage of the products, with no *p*-cresol being observed after 1h incubation. The simplest explanation for these observations is the creation of a futile cycle (see Figure 100, blue) due to substrate binding to the enzyme and the abstraction occurring in another position, either directly in an enzyme amino acid, or in the hydroxyl group of HPHPA, perhaps explaining the increase of uncoupled turnover when compared with HPPA. The HPHPA radical may then be quenched back to starting material or create a radical damage pathway that can inhibit enzymatic activity. Compounds which show an increase in the turnover of SAM but that are not cleaved might be a valuable co-crystallization partner for *Ti*HydG to obtain a structure with a substrate present, as they could potentially enter the active site and remain in position due to a lack of effective turnover.

Another result that appears to contradict previous reports arises from the reaction with tyramine **120**. Neither tyrosine lyase [46] nor *Ca*HydG [519] showed any significant increase of SAM cleavage when incubated with this compound. However, when the reaction is carried out with *Ti*HydG some DOA formation was observed (a 3-fold increase in turnover over the negative

control). The results also show a non-zero formation of *p*-cresol **18**, potentially indicating some $C_\alpha$-$C_\beta$ cleavage, though the results are within the detection limit for *p*-cresol [519] and may constitute a false positive. If cleavage does occur, it is possible the lack of cyanide measurement arises from the hydrolysis of hydrogen cyanide into formaldehyde and ammonia (Figure 102). In future assays, the formation of formaldehyde could be tested using, for example, 3-methyl-2-benzothiazolone hydrazone [617], which would confirm the pathway of tyramine degradation by HydG.

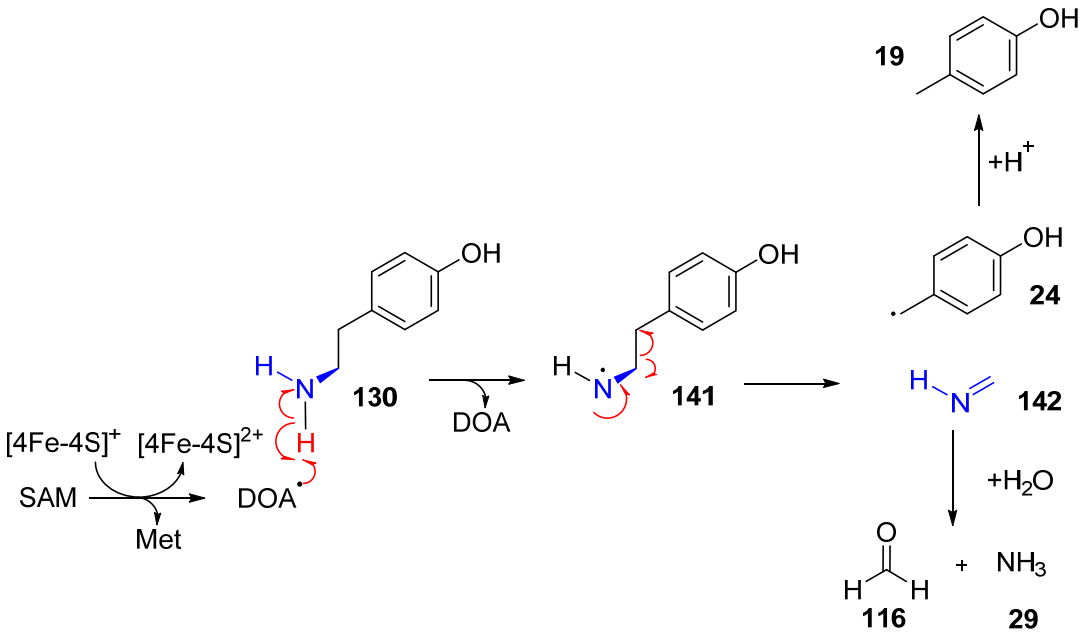



**Figure 102 Proposed mechanism for tyramine radical mediated cleavage.**

The scission of the $C_\alpha$-$C_\beta$ bond of tyramine could yield an oxidobenzyl radical **24**, similar to the reaction with tyrosine, together with hydrogen cyanide **142**. Hydrolysis of the latter could potentially yield formaldehyde **116** and ammonia **29**.

L-DOPA **127** (Figure 103) can be considered a full agonist for the enzyme *Ti*HydG, as its HydG-catalysed cleavage leads to the formation of DOA, methylcatechol and cyanide, indicating complete turnover. The calculation of methylcatechol **136** formation is, however, an estimation as the quantification was carried using *p*-cresol as a standard.

DOA $k_{cat}^{app}$ with L-DOPA is higher (15.2 ± 0.7 x $10^{-4}$ $s^{-1}$) than when L-tyrosine is used as a substrate (12.4 ± 0.3 x $10^{-4}$ $s^{-1}$), but methylcatechol **136** $k_{cat}^{app}$ (6.2 ± 0.3 x $10^{-4}$ $s^{-1}$) is lower than its *p*-cresol counterpart (10.2 ± 0.3 x $10^{-4}$ $s^{-1}$). With a higher difference between DOA and $C_\alpha$-$C_\beta$ cleavage product, L-DOPA seems to induce a higher rate of uncoupled turnover by the enzyme. The

additional hydroxyl group may change the stereoelectronics, slowing down the $C_\alpha$-$C_\beta$ bond cleavage, allowing the quenching pathway (see Figure 100, blue) to occur with higher efficiency relative to the natural tyrosine substrate.

It is worth noting that when the reaction does proceed through the main pathway of the reaction, the apparent turnover numbers of formation for methylcatechol and cyanide are much closer (methylcatechol $k_{cat}^{app}$= 6.2 ± 0.3 x $10^{-4}$ $s^{-1}$ and cyanide $k_{cat}^{app}$= 5.4 ± 0.0 x $10^{-4}$ $s^{-11}$ respectively) than those of *p*-cresol and cyanide with the native substrate (*p*-cresol $k_{cat}^{app}$= 51.2 ± 0.9 x $10^{-4}$ $s^{-1}$ and cyanide $k_{cat}^{app}$= 30.8 ± 0.8 x $10^{-4}$ $s^{-11}$ respectively). This indicates that, following $C_\alpha$-$C_\beta$ bond cleavage of L-DOPA, less DHG is lost to the glyoxylate pathway, or other potential degradation mechanisms. The reasons behind these differences are not completely understood, but two possible explanations should be considered. The first can be related to small conformational changes of L-DOPA in the active site, perhaps being positioned closer to the tunnel linking the RSAM and auxiliary cluster active sites, protecting DHG from hydrolysis. The second is related to the enzyme kinetics. The higher rate of tyrosine cleavage with L-tyrosine in relation to L-DOPA (51 nM $s^{-1}$ in comparison with 30 nM $s^{-1}$) may be closer to the saturation point of the auxillary cluster active site, leading to more DHG to go through a secondary degradation pathway when the native substrate is present in the *in vitro* conditions.

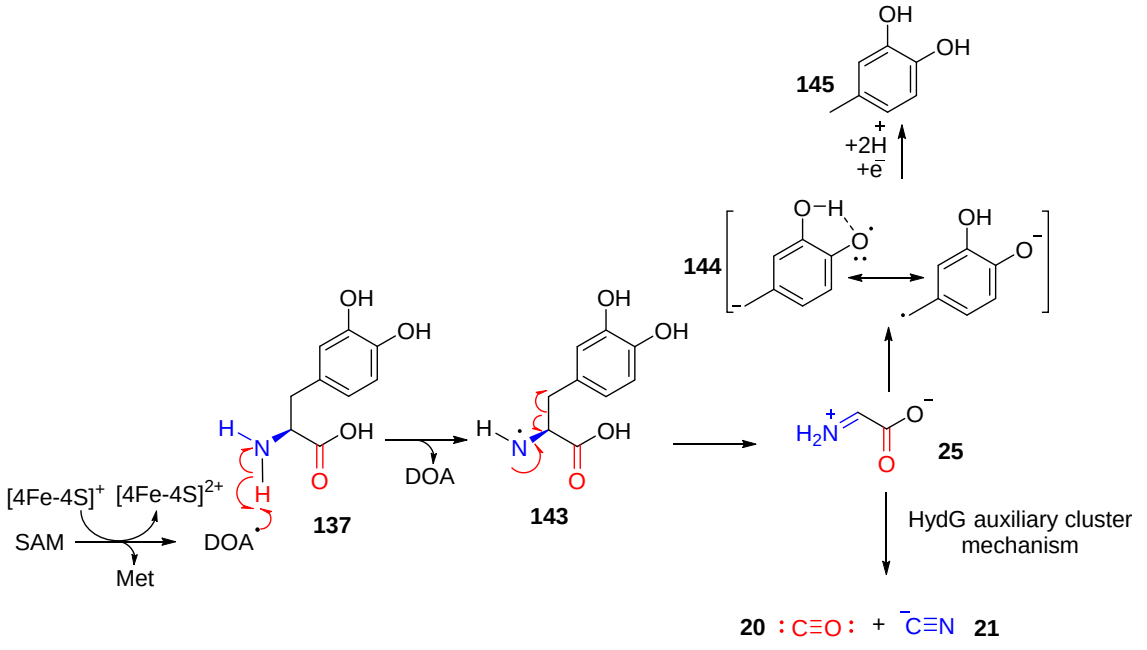



**Figure 103 Proposed mechanism for L-DOPA radical mediated cleavage.**

The mechanism of L-DOPA cleavage closely follows that of L-tyrosine, leading to the measurable formation of cyanide and methylcatechol.

In contrast with L-DOPA, tyrosinamide shows similar rates of turnover for both tyrosinamide and SAM cleavages, indicating a low incidence of uncoupled turnover of SAM. When observing the binding interactions predicted in section 2.7.5.1, changing the carboxylate to an amide [618] will still allow the hydrogen bond pairing with Ser268, the only interacting residue, which may explain the observed minimal impact in compound scission. Considering the mechanism of substrate cleavage is conserved, it is then possible that either 2-oxoacetamide **148** or amino(hydroxy)acetic acid **149** are released as products of the reaction (Figure 104 blue). If the breakdown into cyanide were to be independent of the auxiliary cluster (see section 1.2.2.3.2 Figure 18), one could additionally predict a similar elimination mechanism for the formation of cyanide **21** from 2-imino-acetamide **150** (Figure 104 red).

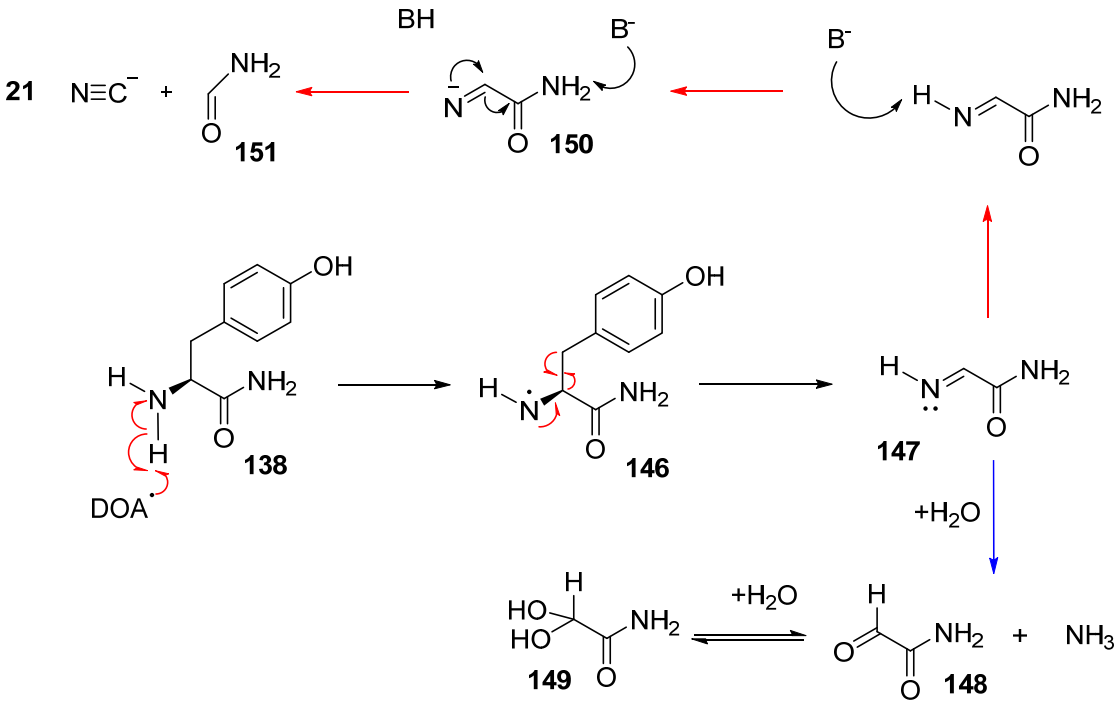



**Figure 104 Potential mechanisms of L-tyrosinamide radical cleavage.**

Following $C_\alpha$-$C_\beta$ bond cleavage in L-tyrosinamide, dehydroglycinamide **147** is expected to form. Hydrolysis of dehydroglycinamide (blue) can be predicted to yield 2-oxoacetamide **148** or amino(hydroxy)acetic acid **149**. Alternatively, two deprotonation evens could lead to the formation of cyanide **21** and formamide **151**.

HTHIQCA **130** is a conformationally locked tyrosine analogue in which the torsion angles around the $C_\alpha$-$C_\beta$ bond are constrained by the cyclic nature of the molecule. The lack of flexibility in the $C_\alpha$-$C_\beta$ could lead to similar lack of activity as reported for *p*-hydroxyphenylcinnamic acid **119** [46].

The results indeed show that the substrate analogue does not show any appreciable SAM cleavage, which could indicate lack of binding to the active site, but the levels of DOA formation are significantly lower than the negative control (20% of DOA formation compared with the uncoupled turnover without any substrate). These alterations could indicate THIQCA does bind to the enzyme, but the mechanism behind lower DOA formation is as yet unclear. It is possible the compound binds to the active site in a different conformation to tyrosine that affects SAM binding, nudging the SAM molecule away from its ideal position. In the future modelling with energy minimization should be used to confirm whether THIQCA could bind to the active site. If the position of the substrate analogue is conserved in regards to L-tyrosine, another possibility is that the lower SAM turnover could arise from the resulting THIQCA radical being stable enough so that cleavage does not occur. This would lead to a re-abstraction of hydrogen from DOA and reforming of SAM, similar to the mechanism reported in LAM [118, 619].

N-methyl-L-tyrosine **129** induced a small rate of SAM turnover (Table 15), and possibly some limited formation of *p*-cresol. However, much like tyramine, the values of *p*-cresol are within the detection limit [519], and SAM uncoupled turnover may be in effect (Figure 100, blue). As the orientation and conformation of the analogue when binding to the active site is unknown, it is impossible to predict whether the abstraction can readily occur at the amino group, where the N-H bond is predicted at -388 kJ/mol; or from the pendant methyl group, where the C-H bond is predicted at -412 kJ/mol [620].

Presuming both the substrate binding as well as the abstraction mechanism are conserved between N-methyl-L-tyrosine and the modelled L-tyrosine (section 2.7.5.1), the addition of a methyl group could occur within the active site, as the large modelled distance (5.1 Å) between C5 of SAM and the amino abstraction point, shows the possibility of a methyl group to be placed in the cavity.

## 4.3 *Ti*HydG Crystallization With Tyrosine Analogues

Despite only preliminary studies, the activity assays with several analogues have already raised interesting questions regarding the potential mechanism of radical-based scission. 4-HPHPA significantly increased SAM cleavage without subsequent $C_\alpha$-$C_\beta$ scission; tyramine and L-tyrosinamide may have suffered limited amount of $C_\alpha$-$C_\beta$ bond cleavage; while reactions with L-DOPA showed measurable cyanide formation. In the other end of the spectrum, HTHIQCA not only does not lead to an increased rate of SAM cleavage, it effectively decreases it in relation to

the control in the absence of any substrate. The knowledge that these analogues interact with the enzyme, potentially in the same binding location as L-tyrosine, makes them exceptional targets for co-crystallization with HydG, in particular as a crystal with the native substrate present has yet to be obtained.

All substrate analogues used in these experiments (4-HPHPA, L-DOPA, tyramine, L-tyrosinamide and THIQCA) share with L-tyrosine a low solubility in water at near neutral pH, leading to only 3 mM being present in the crystallization drop. Direct co-crystallization using the commercial broad screens Molecular Dynamics PACT Premier and Hampton PEGRx HT (Method 9.8) showed no crystals formed even following a three month crystallisation experiment.

The addition of seeds (Method 9.10) changed the outcome for both Tyramine and HTHIQCA, with few small crystals being observed after 72 hour period; while the co-crystallization experiments with 4-HPHPA, L-DOPA and L-tyrosinamide showed no crystal formation even after the addition of crystal seeds.

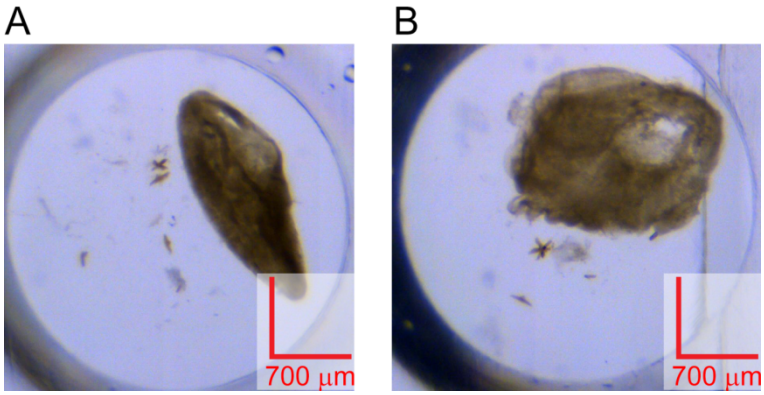



**Figure 105 *Ti*HydG crystal formation in the presence of substrate analogues**

*Ti*HydG (0.8 mM) crystals grown with the addition of microseeds ($10^{-2}$ dilution) in the presence of 10 mM SAM and **A:** 3 mM THIQCA; **B:** Tyramine. All crystals grown in condition F1 of Hampton PEGRx HT (0.1M Bis-Tris proptane, pH 65.;0.2M sodium fluoride; 20% PEG 3350).

### 4.3.1 Structural Data and Density Analysis of *Ti*HydG Crystals Grown in the Presence of THIQCA and Tyramine

Crystals grown in the presence of THIQCA were tested at ESRF, beamline BM14U, at a constant temperature of 100 K and using a 1.7396 Å wavelength, corresponding to the iron κ-edge. One of the crystals diffracted to a level where a complete dataset could be collected, being processed by the automatic pipeline GrenADeS [586].

The tyramine crystals was tested at Diamond Light Source facilities, beamline I04 where a full set of data was collected at 0.9375 Å wavelength, with the data processed automatically using Diamond pipeline [554, 555, 621] (see Appendix A.9 and Appendix A.10). The data, while relatively low in resolution (3 Å), was still assessed for any density corresponding to the substrate analogue.

Both ESRF and DLS have integrated *EDNA* framework for data analysis, which includes DIMPLE (DIfference Map PipeLinE) [622]. If a substrate-free structure PDB file of the protein being studied is present when the data is being processed, it will create a difference map to allow the user to screen non-proteinacious derived density, quickly assessing whether a potential substrate is bound, screening if further analysis of data is required.

Using the same principle, the data collected was used in molecular replacement with the available structure of HydG (PDBID: 4WCX), with the ligands deleted, showing a very likely solution (8.0 TFZ score) for the structure. Molecular replacement statistics are presented in Appendix A.9 (THIQCA) and Appendix A.10 (tyramine), together with the data collection statistics.

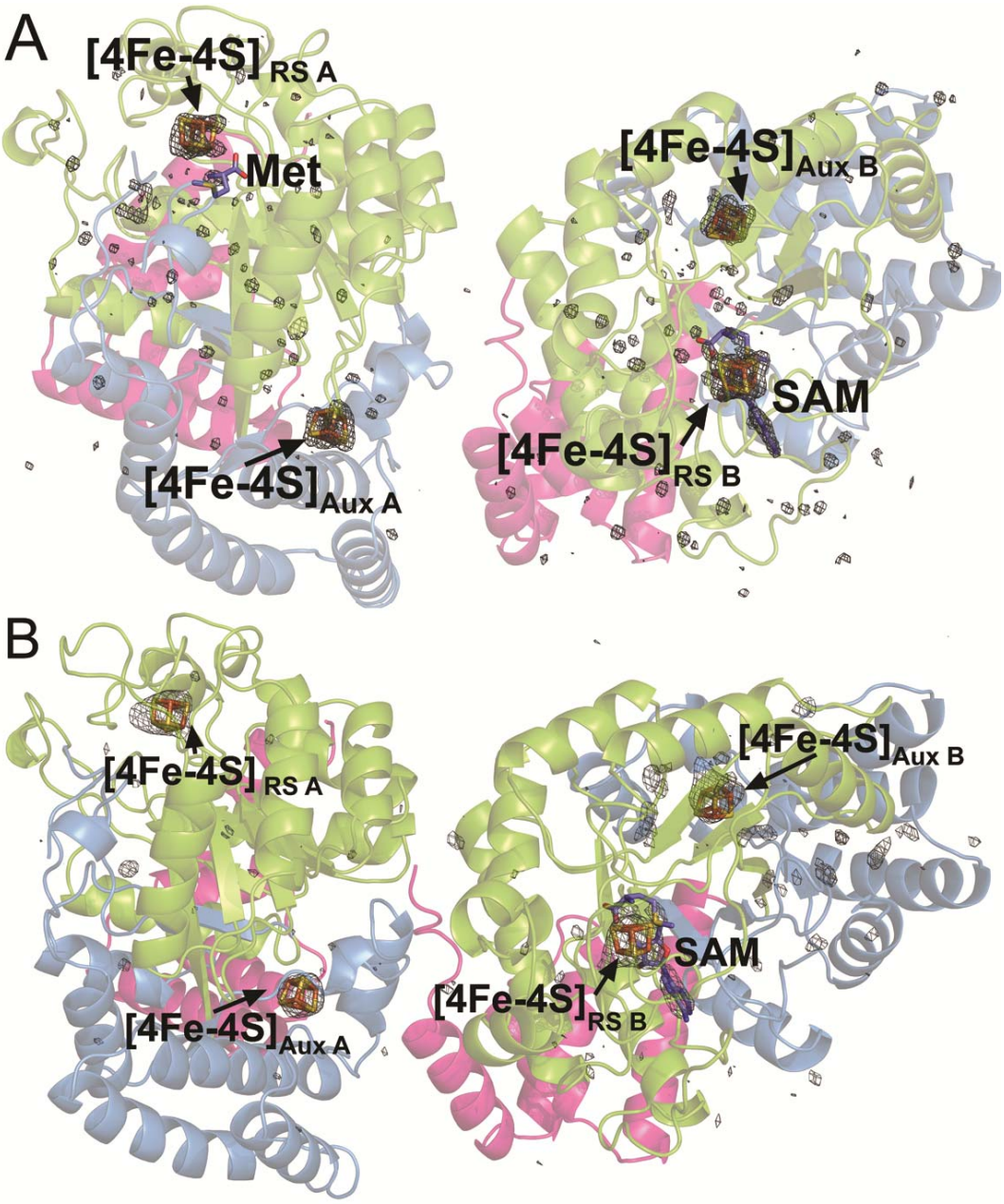



**Figure 106 Model and density map of *Ti*HydG obtained from crystals grown in the presence of substrate analogues.**

The N-terminus is shown in pink, the partial TIM barrel in green and the C-terminus in blue. Both [4Fe-4S] clusters are shown in yellow and orange, SAM and SAM-derived methionine in lilac. The mFo-DFc difference map (cutoff $\sigma$= 3.0) is observed as a black grid, revealing density for both clusters and SAM derived compounds, but no additional density observed. **A:** 2.2 Å data from crystal grown with THIQCA. No density for the labile iron can be observed (an Fe atom is modelled in the structure, but no density can be observed near its site. **B:** Structural data obtained from crystal grown with tyramine (3.0 Å). In monomer A there is no density for SAM nor methionine. In both monomers, only the [4Fe-4S] form of the auxiliary cluster is observed. The lower resolution shows less defined densities for the clusters.

No density for either tyrosine analogue (THIQCA or tyramine) can be observed in the models obtained from Phaser (Figure 106), nor is there clear density for the labile iron. The latter can be explained, in the case of THIQCA, due to experimental time limitations, as the shipping of the crystals for analysis in a dry shipper forced the samples to be exposed to oxygen for several days prior to testing. The low data resolution of tyramine may be unsuitable to clearly observe density for a low occupancy atom such as the labile iron.

### 4.3.2 *Ti*HydG Crystallization in the Presence of Different Amino Acids

The initial structure of *Ti*HydG obtained with a full auxiliary [5Fe-5S] cluster showed the presence of an amino acid coordinating the labile iron through its amino and carboxylate moieties (see section 2.7.3). The presence of the amino acid was fortuitous, as structures can be obtained where the labile iron does not show any coordinating ligands apart from His265, the bridging sulphide and partially disordered water molecules (see section 3.5.1).

A large co-crystallization experiment was carried out (Method 9.16), with the aid of seeding, where different amino acids were tested in different concentrations (0.5; 1 and 5 mM), together with SAM and additional additives. Glycine was tested as the simplest amino acid and the one that closest resembles DHG; L-methionine was chosen as the most likely candidate for the residue present in the available structure, as consequence of SAM degradation near the $[4Fe\text{-}4S]_{RS}$ cluster, while L-cysteine was used for its ability to coordinate iron with its thiol sidechain, together with its proposed role in labile iron placement [623]. Additionally, to improve the occupancy of labile iron, external $FeCl_2$ (5.0 mM) and/or $Na_2S$ (3.0 mM) were introduced into the crystallization drop.

Together with the addition of the non-substrate amino acids and additives, the presence and absence of 3.0 mM L-tyrosine was also tested, as small conformational changes in the coordination of the auxiliary cluster may alter the ability of the protein to bind L-tyrosine.

After 24-hour incubation the samples which had had $Na_2S$ added showed the formation of heavy black precipitate even in the absence of $FeCl_2$. All other samples had no morphological differences between them and the seed stock dilutions were added.

The additives severely hampered the growth of crystals, which were only visible after a three month incubation period, in sharp contrast to the three day growth observed when crystals are co-crystallized with SAM alone. Crystals obtained were frozen and tested at Diamond Light Source facilities, in the microfocus line ID-24, at a constant wavelength of 0.95375 Å. Despite an extensive amount of crystals tested, from various conditions, their diffraction patterns were much

too weak to obtain any structural data. The disparities in crystal formation time and quality of data obtained indicate that considerable optimization is required to change the parameters in the crystallization conditions, in order to improve diffraction and ultimately obtain any structural information.

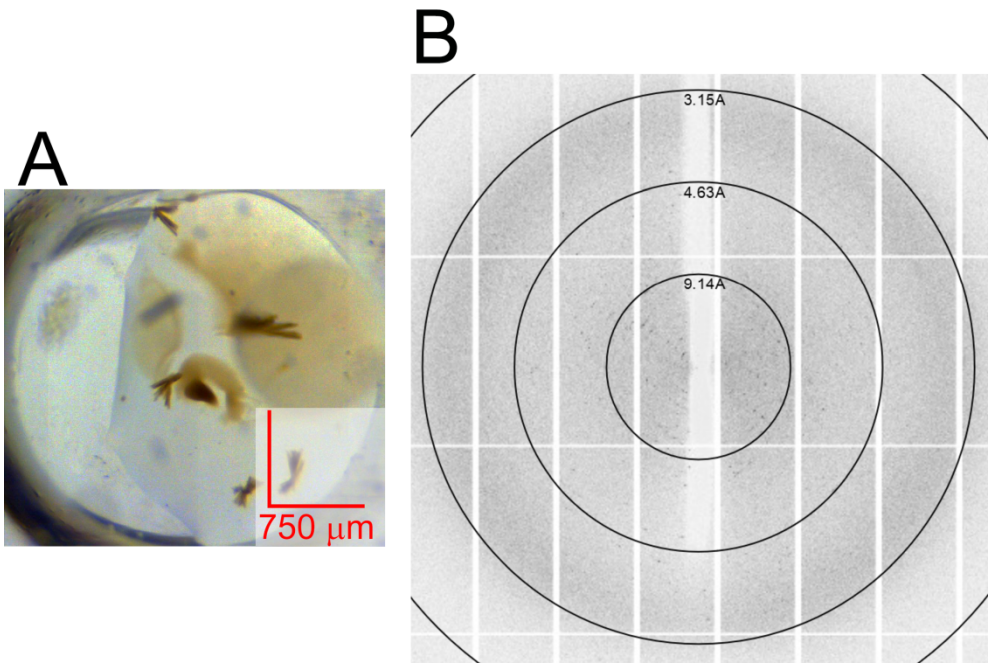



**Figure 107 Crystal formation with amino acids and additives, and respective diffraction pattern.**

**A:** Example of *Ti*HydG crystals grown in the presence of 10 mM SAM, 3 mM Fe(III)$Cl_3$ , 3 mM $Na_2S$, 3 mM glycine. Crystallization condition F1 of Hampton PEGRx HT (0.1M Bis-Tris proptane, pH 65.;0.2M sodium fluoride; 20% PEG 3350). **B:** Diffraction pattern of the same crystals, obtained in the synchrotron facilities at Diamond Light Source, beamline I24.

### 4.3.3 Further Studies of *Ti*HydG Crystallization in the Presence of L-Cysteine

The presence of L-cysteine during the maturation stages of the H-cluster has been known to improve the rate of formation of the active hydrogenase, despite a lack of understanding on the precise mechanism by which this occurs [184]. L-cysteine has been proposed as a possible substrate for HydE, albeit the latest studies show only a modest increase in SAM turnover when presented with this amino acid in relation to other thiols [304]. The coordination of an amino acid in the *Ti*HydG structure with the labile iron [509] has led to the possibility that the μ2 sulfide actually arises from the side chain of cysteine **147**, which would provide a role for cysteine in the process of iron transfer [623], explaining why the addition of this amino acid seemed to improve the maturation of HydA [624].

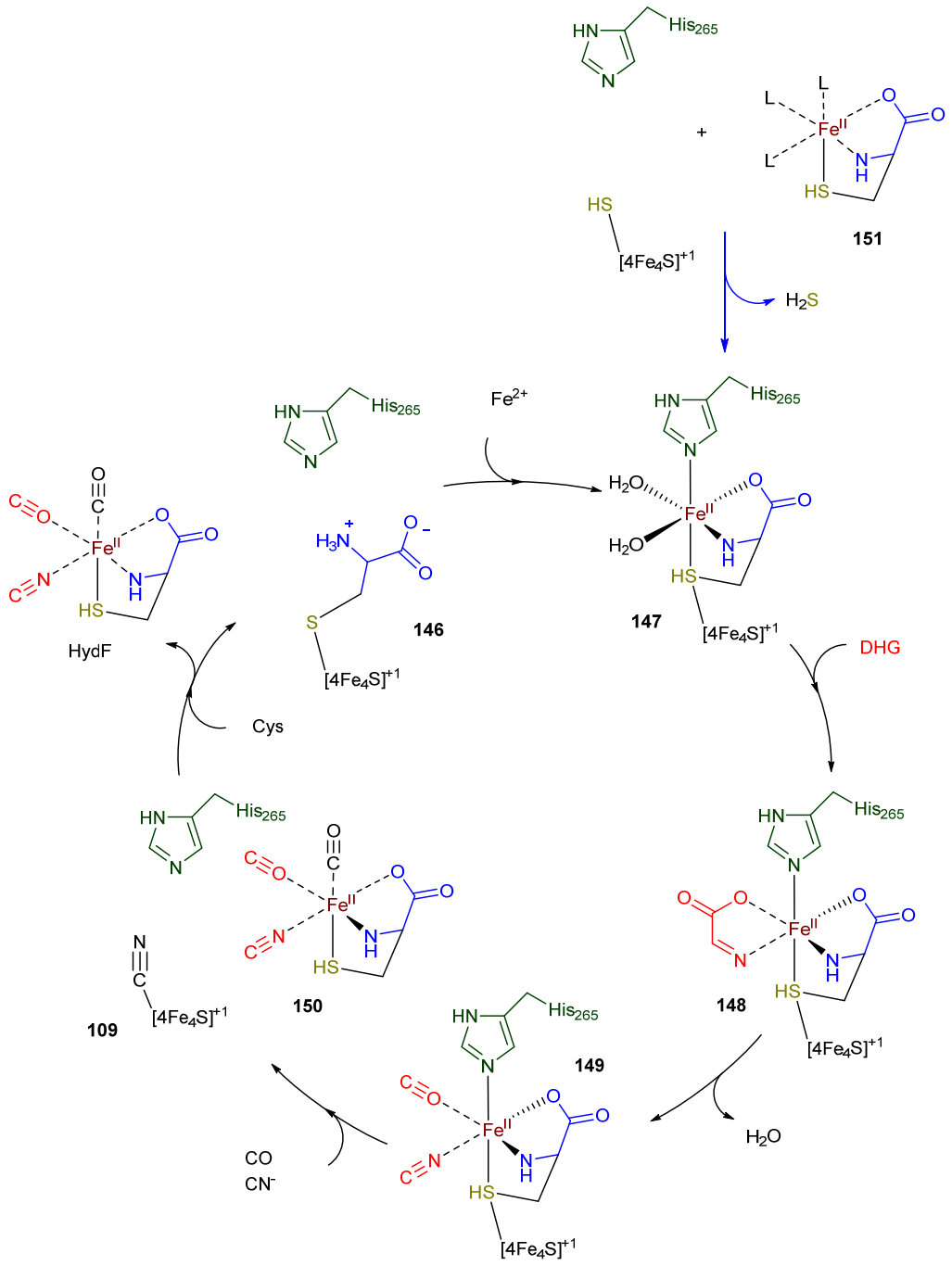



**Figure 108 Potential mechanism of cysteine mediated transfer for the $Fe(CO)_2(CN)$ synthon.**

The mechanism proposed shows the resting state of HydG with an exogenous cysteine bound to the [4Fe-4S] auxiliary cluster **146**. The introduction of $Fe^{2+}$ in the system, leads to the octahedral coordination around the labile iron **148** where DHG cleavage into CO and cyanide occurs **149**. The cysteine coordinating the loaded iron can then dissociate from the cluster and His265 coordination **150**, presumably to be transferred to the maturase scaffold HydF. An alternative mechanism (blue), where a cysteine coordinating an iron **151** binds to a [4Fe-5S]Aux cluster seems more consistent with the structural information available for *Ti*HydG. Adapted from Suess *et al.* [623].

To test whether this theory had any validity, crystal formation in the presence of L-cysteine was attempted, in an effort to replicate the formation of the amino acid-labile iron bonding. To the same crystallization conditions reported previously (section 2.5) 1.0 mM, 5.0 mM and 10.0 mM of L-cysteine were added to the crystal drop. Unfortunately, it was observed that the presence of L-cysteine in the crystallization has a severe impact in the morphology of the crystal formation, with very small crystalline formations observed.

To attempt to remedy this situation, pre-grown crystals in condition F1 (section 2.5, Table 5) were transferred onto fresh drops where L-cysteine was mixed with the crystallization condition (5.0 mM L-cysteine) (Method 9.16). The crystals were soaked in the cysteine-rich formulation for four hours or twenty-four hours and subsequently frozen and tested at Diamond Light Source Synchrotron facilities, beamline I02. Crystals that were allowed to incubate in the fresh drop for 24h diffracted to low resolution (>5.0 Å), and no relevant information was collected from the structure. Of the crystals which were incubated for the lowest amount of time (4h), one diffracted to high resolution (1.8 Å). As it was important to confirm the potential coordination of cysteine to the labile iron, two data sets were obtained, one at high energy (12 kEV) and one at an energy near the iron k-edge (7.2 kEV), for observation of the iron anomalous data. (Appendix A.11). The same molecular replacement method to search for substrate analogue density (section 4.3.1) was used to verify if cysteine was present near the auxiliary cluster.

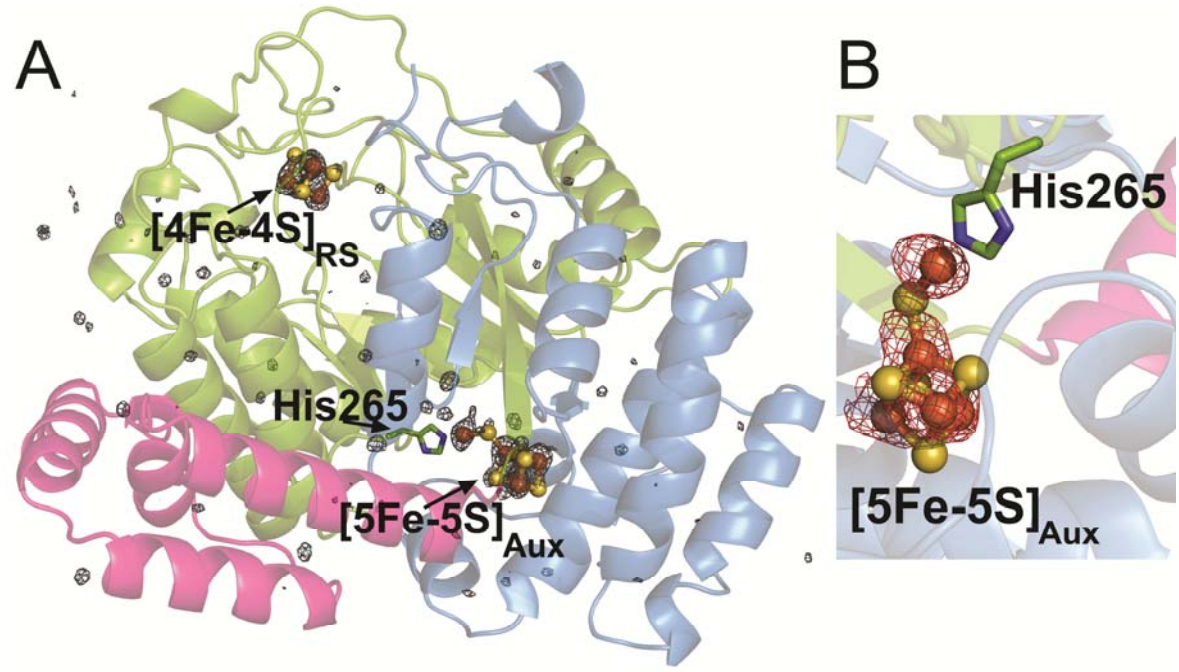



**Figure 109 Model and density map of *Ti*HydG crystal soaked with L-cysteine.**

**A:** Overall fold of the protein, with the mFo-dFc difference map present, correlating with the ligands $[4Fe\text{-}4S]_{RS}$, $[5Fe\text{-}5S]_{Aux}$, SAM and methionine. **B:** Anomalous signal around the $[5Fe\text{-}5S]_{Aux}$ cluster, with iron atoms located using the anomalous difference Fourier map shown in grey and contoured in at 5.0 σ.

The electron density near the auxiliary cluster corresponding to the anomalous signal shows the labile iron still present (Figure 109 B), but the electron density shown in the difference map is not well defined enough to model any ligand apart from some potential water molecules (Figure 109 A).

As it currently stands it is difficult to confirm the validity of the role attributed to cysteine by Suess and colleagues [623]. The high resolution structure discussed in Chapter 3, shows monomer B coordinating a bridging sulfide and not a cysteine, conflicting with the labile iron free state of the auxiliary cluster **150** in Figure 108. It is possible a cysteine already coordinating an iron could displace the bridging sulfide and the general mechanism being conserved (alternative pathway with **151** in Figure 108, blue), but further studies will be required to assess if L-cysteine does play a part *in vivo* in the transfer of the labile iron.

### 4.3.4 TiHydG Crystal Soaking in the Presence of Diatomic Ligands

The proposed mechanism of action (see section 2.11) places the labile iron in central stage, not only in the formation of the diatomic ligands, but also in transferring the bound CO/$^{-}$CN to the scaffold enzyme HydF as part of the maturation process. Despite structural information placing the labile iron in a good position to capture the product(s) of L-tyrosine cleavage, obtaining a structure with the diatomic ligands bound would certainly strengthen the proposal for the role of labile iron in DHG breakdown, while allowing the observation of the labile iron coordination after DHG cleavage.

Pre-grown crystals in condition F1 of Hampton PEGRx were allowed to soak in F1 drops with increasing amounts of cyanide (0.1-1 mM) for 2 hours. Additionally, samples were soaked for 48h in either 2 mM cyanide or in a carbon monoxide saturated solution. After the soaking period, the crystals were directly frozen in liquid $N_2$ and tested at either DLS facilities, beamline I02 (2h soaking crystals) or at ESRF facilities, beamline ID29 (48h soaking crystals). The data obtained for the *Ti*HydG crystal soaked with CO had in general very poor quality, owing to a quick degradation of the crystal while testing. The signal to noise ratio is very low (1.5), which together with a very low completeness (58%) leads to an unreliable set of data, and, as such, the dataset was discarded (see Appendix A.12 for the data collection statistics).

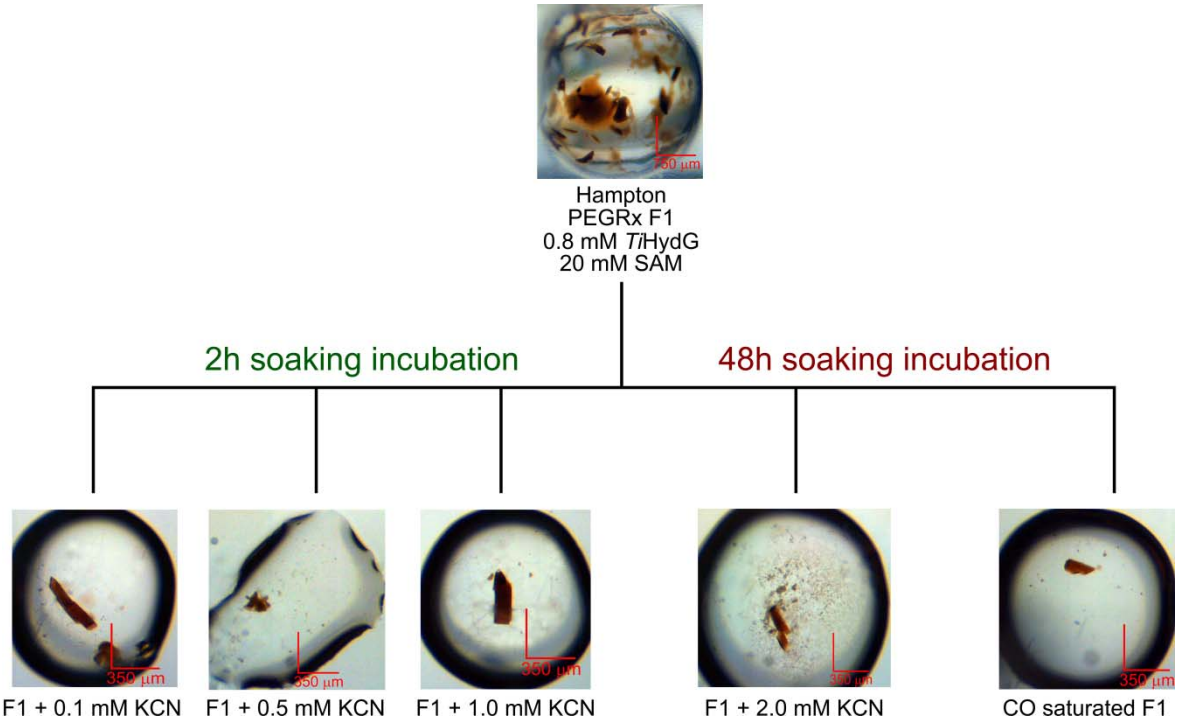



**Figure 110 Soaking experiment of *Ti*HydG crystals with the diatomic products cyanide and carbon monoxide.**

Regardless of the amount of cyanide added and the time incubated, no diatomic molecule can be observed binding to the labile iron. However, there appears to be an inverse relation between the amount of cyanide added and the occupancy of the iron, with highest values of KCN leading to lower amounts of labile iron present. Furthermore, for the most extreme case, where a high content (2.0 mM) of cyanide was incubated for 48h, the labile iron, together with the unique iron of the $[4Fe\text{-}4S]_{Aux}$ cluster, are completely absent, together with the unique iron of the $[4Fe\text{-}4S]_{Aux}$ cluster, effectively showing that cyanide can degrade the auxiliary cluster with the formation of a $[3Fe\text{-}4S]_{Aux}$ cluster (Figure 111).

The cluster degradation with cyanide was also observed by Dr. Daniel Suess, as samples treated with the diatomic compound lost the characteristic signal of the auxiliary cluster [509].The progressive loss of the label iron signal with increasing cyanide concentrations could be considered a confirmation of the mechanism by which the iron is released, upon diatomic ligand formation. However, the interaction of iron with cyanide anions is well documented, creating extraordinarily stable compounds [625], including the pigment Prussian blue [626], meaning that the removal of the iron by cyanide could also have occurred from a purely inorganic process.

In the structures (Figure 111), the $[4Fe\text{-}4S]_{RS}$ cluster is still intact, showing four distinct iron signals when observing the anomalous data collected near the iron κ-edge (Figure 111). This would make the four iron signals of the RSAM conserved cluster an internal control over cyanide degradation, and reveals a very strong selectivity of the cyanide towards the labile and unique irons of the $[5Fe\text{-}5S]_{Aux}$ cluster.

An explanation for these results could arise from the coordination of SAM protecting the unique iron of the $[4Fe\text{-}4S]_{RS}$ cluster, reducing its availability towards the formation of a cyanide complex. This would contrast with the auxiliary cluster whereas, after removal of the labile iron and subsequent release of the bridging μ2 sulfide, the unique iron of the now $[4Fe\text{-}4S]_{Aux}$ cluster would be potentially unprotected. Despite not being fully equivalent, a study characterizing the ability of SAM to reduce autoxidation of $Fe^{2+}$ in the presence of other iron chelators, showed a very weak effect of SAM to prevent the oxidation, indicating the interaction of SAM with an iron does not prevent interactions with other molecules [627]. It is possible, then, that SAM is not responsible for the unreactivity of the unique iron of the $[4Fe\text{-}4S]_{RS}$ cluster to excess cyanide in solution. The resistance of the $[4Fe\text{-}4S]_{RS}$ to cyanide degradation, be it by SAM coordination or by design of the protein structure, could also serve as a protection against RS cluster degradation from any unwanted release of cyanide upon DHG degradation.

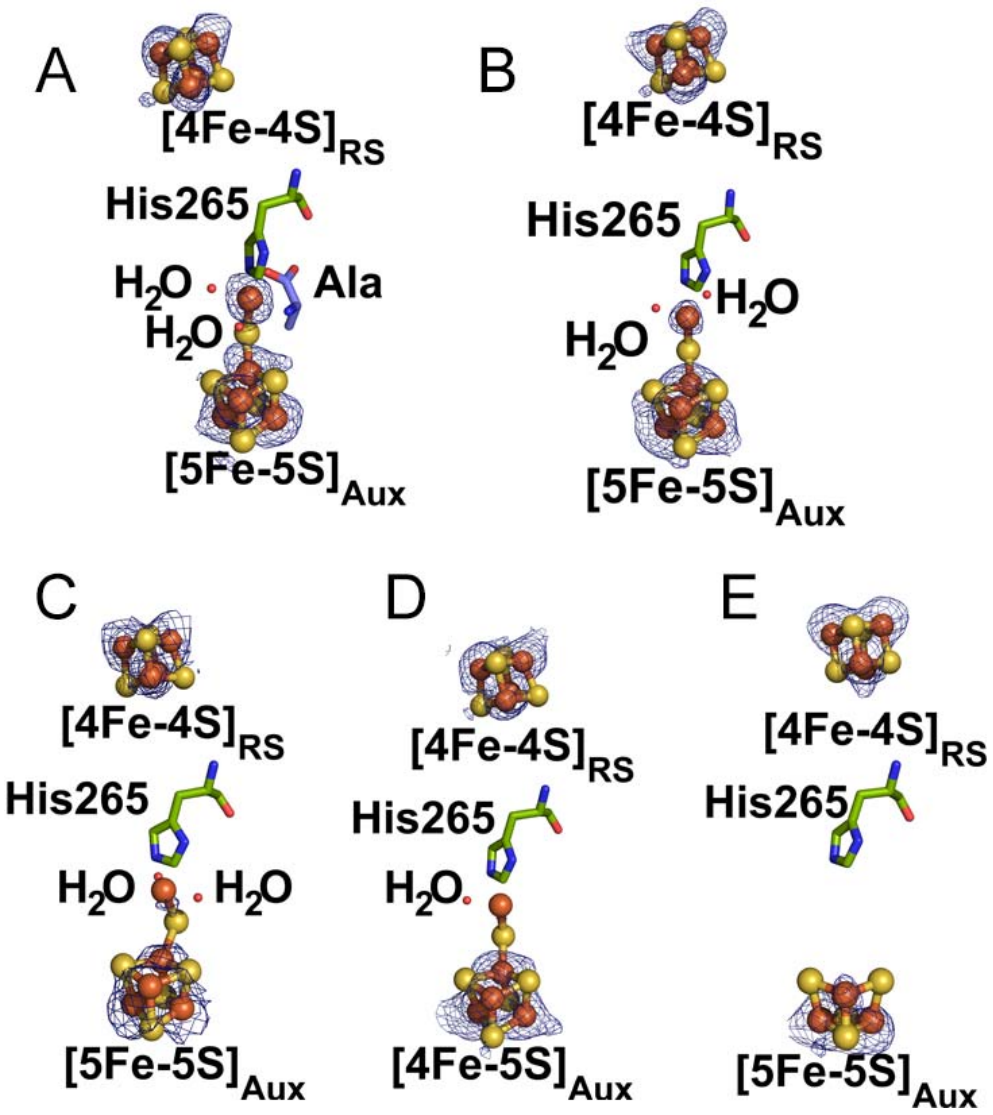
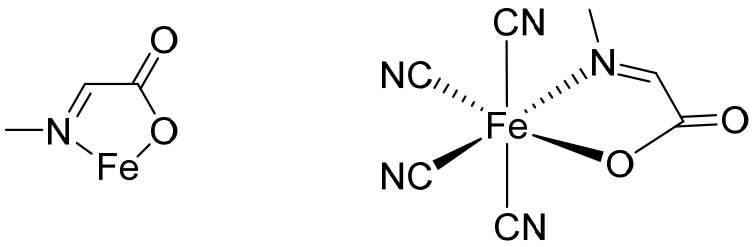



**Figure 111 Anomalous signal surrounding both clusters of different crystals of *Ti*HydG soaked with increasing concentrations of KCN.**

The anomalous difference Fourier map of the irons that constitute both the $[4Fe\text{-}4S]_{RS}$ and the $[4Fe\text{-}5S]_{Aux}$ cluster are shown in blue, contoured at 3.5 σ. **A:** Control data with 0 mM KCN (PDBID: 4WCX). **B**: Data obtained for the crystal soaked for 2h with 0.1 mM KCN. **C:** Data relative to the crystal incubated with 0.5 mM KCN for 2h. **D**: Data obtained for the crystal soaked for 2h with 1.0 mM KCN. The signal for the labile iron is practically absent **E:** Data for the crystal incubated with 2.0 mM KCN for 24h. The degradation of the auxiliary cluster was severe, and only a [3Fe-4S] can be built in the available density.

## 4.4 Inorganic Synthesis of α-Iminocarboxylic Acid

The identification of the labile iron has introduced the possibility of an inorganic centre in which the formation of diatomic ligands takes place. Interestingly, a related non-enzymatic reaction has been described previously, where the incubation of an amine, pyruvate and Fe(II) gives rise to the formation of an α-iminocarboxylic acid [628]. When cyanide was also present in the system, the complex became more stable towards acid hydrolysis as well as becoming oxygen-tolerant, with tetracyano-α-iminocarboxylato-ferrate(II) compounds (Figure 112) being proposed as the products of the reaction, indicative of an octahedral coordination not dissimilar from structure **119** proposed in the mechanism of HydG (section 2.11).




**158** **159**

**Figure 112 Ferrous α-iminocarboxylic complexes**

Iron(II)-α-diimine chromophoric chelate ring **158** and tetracyano(-α-iminocarboxylato)ferrate(II) compound **159** as products of α-iminocarboxylic acid formation in the presence of ferrous iron. Adapted from [628].

Following the same mechanism with ammonia, pyruvate and Fe(II) to yield dehydroglycine would create a framework to study the mechanism of DHG cleavage, effectively bypassing the first half of HydG's reaction (section 2.11, Figure 83). Additionally, the formation of a stable complex could serve as a potential target for crystallization studies to reliably observe a coordinated labile iron with its substrate (DHG) bound.

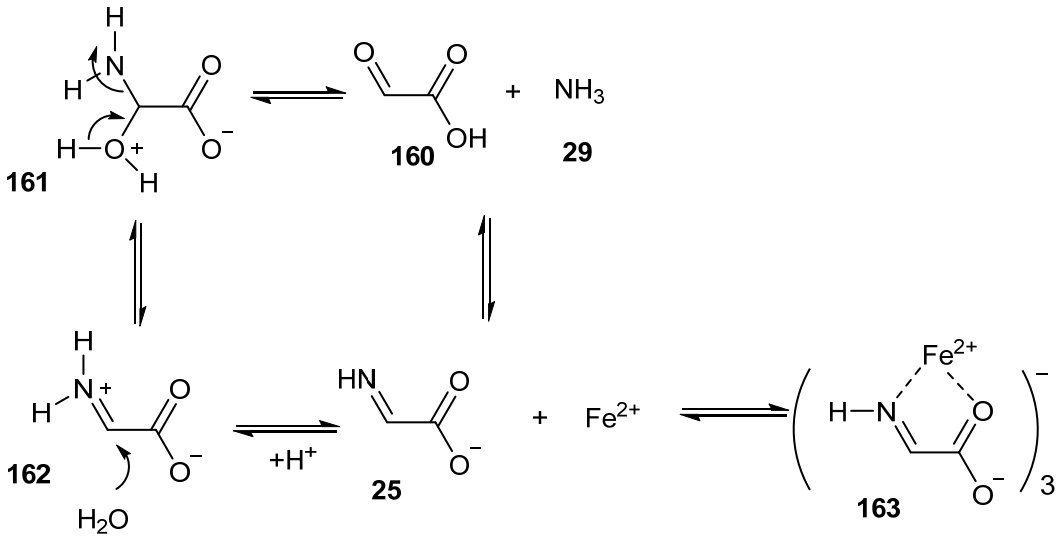



**Figure 113 Dehydroglycine:Fe complex**

Possible mechanism for the coordination of ferrous iron by DHG **163** by the formation of α-aminocarboxylic acids with glyoxylate **160** and ammonia **29**.

### 4.4.1 Formation and UV-Spectroscopy of the Fe-DHG Complex

An initial replication of the conditions previously reported by Krumholz [628] (Method 9.18), mixing glyoxylate with varying concentration of $FeCl_2$ and ammonia led to the formation of dim, purple coloured solutions after 1h incubation. The reactions are identified by their iron rate, calculated as shown in Equation 5.

$$X_{Fe} = \frac{\mathrm{Fe(II)}}{(\mathrm{Fe(II)} + \mathrm{NH_3})}$$

**Equation 5 Iron rate calculation.**

The samples were spun-down, removing a white precipitate, however UV-vis spectra still showed a strong interference making the results difficult to analyse, as the peak broadens throughout the entire visible spectrum, despite a broad peak at the previously reported wavelength 475-520 nm [628] being present in the samples. The white precipitate was replicated in the control where $FeCl_2$ was mixed with $NH_3$ in the absence of glyoxylate **160**, leading to a strong interference of the absorbance (Figure 114 B).

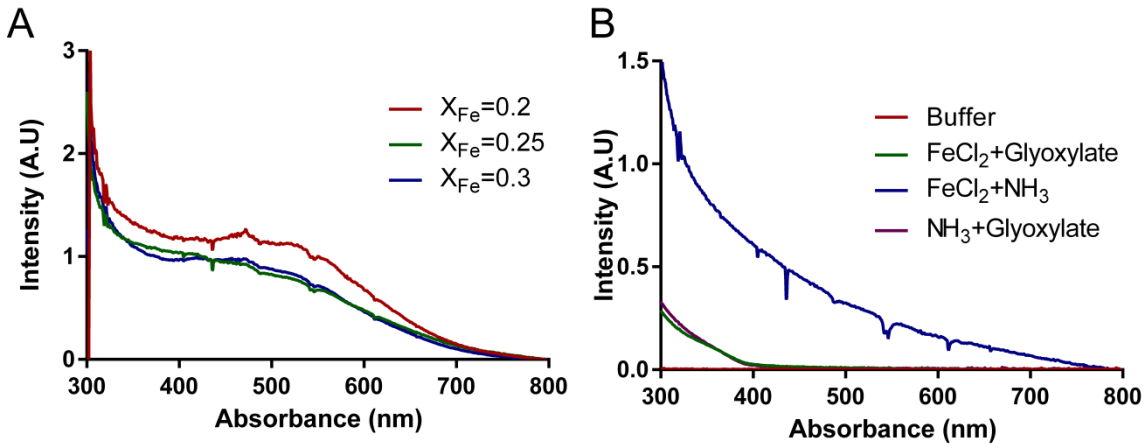



**Figure 114 UV Spectra following the formation of Fe-DHG**

**A:** Results obtained for the reaction made with different $X_{Fe}$ rates **B:** Controls of just the reaction buffer (red); $FeCl_2$ and glyoxylate (green); $FeCl_2$ and $NH_3$ (blue); glyoxylate and $NH_3$ (purple) with no relevant peak at 530 nm corresponding to the product. The mixture of $FeCl_2$ and $NH_3$ led to the formation of a white precipitate, resulting in a broad peak.

To improve the formation of the chromophore compound while reducing the amount of precipitate formed, the conditions of the experiment were altered (Method 9.18.1), by diluting the reactants 10-fold, while keeping the rates of Fe to $NH_3$ ($X_{Fe}$) between 0.1-0.3. A clear purple solution was formed after incubation, without a white precipitate observed, leading to a clear UV-Vis spectra (Figure 115 A). The reported observations by Krumholz [628], showed that at a constant iron(II) concentration and increasing methylamine values resulted in an increase in the absorption maximum, from 440 nm until 477 nm, until a 3:1 amine:iron ratio was reached ($X_{Fe}$=0.25). Further increases in the amine content resulted in a decrease of the absorption maximum, corresponding to formation of species with less than a 3:1 amine:iron ratio when the amine is in a large excess [628].

Changing methylamine for ammonia in the system led to the absorption maximum being at 530 nm. Additionally, while a similar relation between absorbance and $X_{Fe}$ could be traced, the signal at low $X_{Fe}$ values remains almost unchanged up until $X_{Fe}$= 0.2, following which a sharp increase was observed. The maximum was reached when $X_{Fe}$=0.25, in accordance with the results reported for methylamine and ethylamine [628], followed by a signal decrease when $X_{Fe}$=0.3, albeit the spectra with the highest $X_{Fe}$ rate also reveals a significant increase in background noise, with a clear peak no longer being observable.

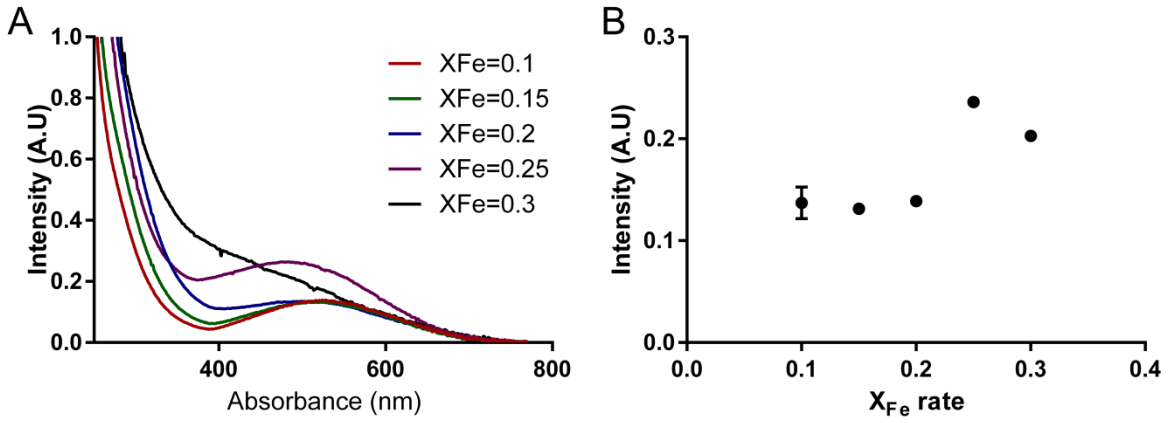



**Figure 115 UV spectra of the formation of Fe-DHG complex, with lower concentration of substrates.**

**A:** Results obtained for the reaction made with 10-fold dilutions of all relevant compounds, using increasing $X_{Fe}$ rates. **B:** Diagram of $A_{530nm}$ peak intensity, showing a dependency of the peak with the $X_{Fe}$ rate, reaching a maximum at $X_{Fe}$=0.25.

### 4.4.2 HydG Auxiliary Cluster Activity in the Presence of Fe-DHG Complex

With the successful formation of a dehydroglycine:iron(II) complex, the hypothesis of bypassing the radical SAM chemistry in HydG could be tested, by measuring the formation of cyanide in the absence of L-tyrosine (Method 9.18.2). HydG was incubated with the product of the reaction of $FeCl_2$, glyoxylate and ammonia, together with the strong reducing agent dithionite, at 37 °C for one hour. Two controls were done, one without the addition of the enzyme, to measure any DHG cleavage that can occur in highly reducing conditions in the absence of the enzymatic environment; and another with the addition of the complex but without dithionite present. DHG cleavage reaction is proposed to make use of the reducing power of the $[4Fe\text{-}4S]_{Aux}$ and with no reducing agent present, close to no reaction is expected.

The cyanide quantification results for both controls are similar, though non-zero, potentially due to some background interference, in particular since the values are low enough to be close to the limit of the methodology [311]. When the enzyme is incubated with the Fe-DHG complex and dithionite, the formation of cyanide is twice that of the controls, indicating successful, albeit very low, turnover by the enzyme.

These results are very preliminary, as analysis of the Fe:DHG complex by LC-MS and NMR have not yet been carried out and are necessary as a confirmation of the nature of what compound is

coordinating iron. Provided the formation of the complex is as predicted (Figure 113), optimization of the concentration of protein:complex, to increase the amount of cyanide formed, would allow for the first time studies on the reaction with the elusive DHG. Furthermore, as the formation of the complexes can be achieved with different aliphatic amines [628], different complexes can be tested for both scission activity on HydG as well as transfer onto HydF.

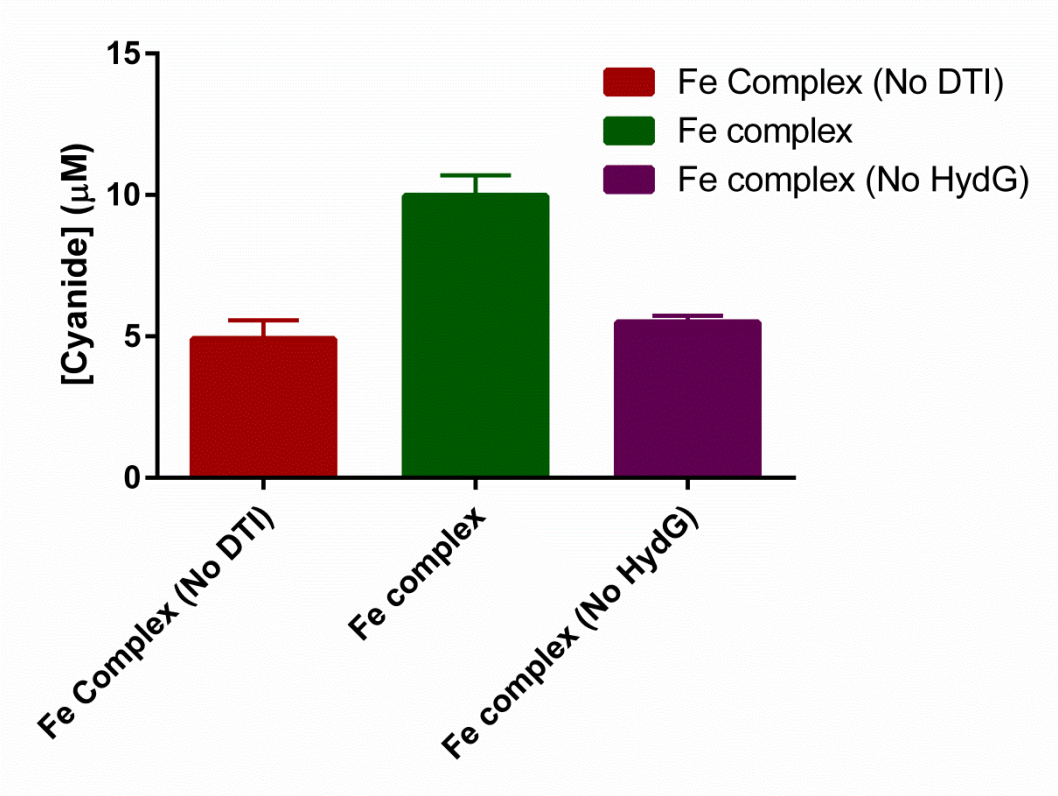



**Figure 116 Formation of cyanide, by *Ti*HydG in the presence of the Fe complex.**

Controls where the reactions were equivalent except for the absence of a reducing agent (No DTH) or enzyme (No HydG) are also shown.

## 4.5 Summary and conclusions

In the absence of a structure with the native substrate to guide the proposed mechanism of action, analogue studies test which moieties are absolutely necessary for interaction with the enzyme and can answer many questions regarding the enzyme's activity. Activity studies with ThiH emphasized the importance of the phenolic O-H bond [46], results which were shared in HydG [519].

Probing the most promising substrates for L-tyrosine analogues of the previous studies, together with new compounds, brought some interesting revelations. Firstly, the removal of the amino group leads to a high rate of uncoupled turnover, but no formation of *p*-cresol (HPHPA **127**; HPPA **128**). These results are in accordance with the latest proposal of the abstraction occurring by the amino group of L-tyrosine [317, 321].

Addition of other functional groups into the aromatic ring leads to a higher rate of uncoupled turnover (L-DOPA **137**), most likely due to steric changes of the side chain, which decrease the ability of the Cα-Cβ bond to be cleaved. The carboxylic acid unit can be successfully changed into an amide (L-tyrosinamide **138**), leading to a significant release of *p*-cresol. Additionally, and unlike previous observations [46], the absence of the carboxylic acid moiety does not completely impede the interaction with the active site, as tyramine **130** led to DOA formation. Interestingly, the cyclization of L-tyrosine in the amino moiety, forcing the molecule into a more planar geometry, not only completely abolishes activity, but decreases the uncoupled turnover of SAM (HTHIQCA **140**).

The identification of a small library of molecules which interact in different ways with the enzyme, paved way for a wider screen of crystallization attempts. Unfortunately, most compounds inhibited severely the ability to obtain crystals. HTHIQCA and tyramine alone were able to be added onto the system and show some crystalline formation, after the use of microseeding. The structural data obtained, however, did not show any density for the substrate analogues.

Building upon the structural information already collected (see chapter 3), a few amino acids were chosen for co-crystallization, to help in the identification of the residue coordinating the labile iron, though, so far, the crystals obtained have had too weak diffraction patterns to obtain reliable structural information. The addition of small molecules, namely iron and sulphide to increase the occupancy of the complete auxiliary cluster, also lead to a significant drop in crystal quality, which is perhaps related to the asymmetry between monomer A and monomer B, where only one is ever shown to contain the $[5Fe\text{-}5S]_{Aux}$.

It seems likely that the crystal form obtained for *Thermoanaerobacter italicus* HydG in the presence of SAM, with space group P1 and unit cell approximately 55.0; 55.0; 90.0; 90.0; 90.0; 65.0 (a; b; c; α; β; γ), is too stable in the absence of additional substrates. This leads to a lack of crystals with other elements bound regardless of them being the native substrate, substrate analogues or indeed other small molecules. Returning to an initial stage where commercial broad screens are to be used and focusing on either expressing a different HydG (*Tl*HydG seems a likely candidate) or by doing broad screens altering the ratio of protein:SAM:substrate may hold the key

for a successful co-crystallization. In order to limit the scope of such studies, estimating the $K_d$ of the protein in relation to different substrate analogues would provide a good selection method for the most relevant substrates, in addition to activity assays. Techniques such as isothermal calorimetry (ITC) have become common tools to assess binding affinities [629] and could guide future experiments, in particular as binding constants are a good guideline for the concentration of substrate to add in a co-crystallization experiment.

Attempting to obtain a crystallographic snapshot of the proposed final product, where the labile iron coordinates CO/CN, was also mildly unsuccessful. Allowing a pre-grown crystal to soak in different concentrations of the diatomic ligands at different timings has not, at least so far, yielded a structure of the loaded labile iron. Instead, the different concentrations of cyanide seem to indicate an inverse relation with the occupancy of the labile iron, with a more extreme event, at 2.0 M of cyanide for 48h, degrading the auxiliary cluster to what is mostly a $[3Fe\text{-}4S]_{Aux}$ species. Both the ability of the labile iron to be abstracted by the presence of cyanide as well as the selective degradation of the auxiliary cluster versus the RSAM conserved cluster are at least indicative of the potential $Fe(CO)_2(CN)$ donation mechanism proposed previously [296].

To further the ability to study the events related to the auxiliary cluster, a chemical process that leads to the formation of iron-bound $\alpha$-iminocarboxylic acids was investigated. Following the protocol established by Krumholz [628], the formation of a Fe-DHG complex was initiated by mixing Fe(II) with glyoxylate and ammonia. The use of the smallest aliphatic amine led to the formation of a purple complex with an absorption maximum at 520 nm, a higher wavelength than those reported for methyl- and ethyl- amines. The compound reaches its highest level of formation when there is a three-fold excess of amine to iron (Xfe=0.25). UV-vis was the only method used for following the reaction, and in the future, confirmation by mass spectrometry and H- and C- NMR are to be carried out.

The ability of HydG to uptake a loaded iron into the correct position was tested by incubating a fully reconstituted enzyme with the complex for one hour. Not unexpectedly, the results show only modest activity by the enzyme as with the labile iron in place, the exchange rate from free labile iron onto a loaded iron complex may be very slow. Additionally, one of the DHG bound to the complex needs to be removed, to release the hexacoordination of the iron so that His265 and the bridging $\mu$2 sulfide can bind.

To maximize the binding potential of HydG, the labile iron should be removed. EDTA is often used to obtain the apo-form of RSAM enzymes, as it is a strong iron chelator [37, 627, 630]. A careful experiment must be designed to understand the amount of EDTA necessary and the incubation

time required, so that the EPR signal for both cluster is still present, the L-tyrosine cleavage activity is not impaired (which would indicate $[4Fe\text{-}4S]_{RS}$ cluster destruction), but no cyanide is formed. Such an enzymatic state would, theoretically, be more susceptible to iron uptake, which could increase the cyanide rate of formation from the complex.

These preliminary results should also be verified with HydG mutants without the $[4Fe\text{-}4S]_{RS}$ cluster. Whilst it is not possible to predict whether the removal of the RSAM conserved cluster would have a general effect on the folding of the protein, which could irreversibly inhibit cyanide formation, the removal of the unnecessary cluster for the second half of the reaction could help in subsequent studies of the DHG cleavage steps by EPR.

# Chapter 5: Further Insights Into the Structure Characterization of Lipoyl Synthase LipA

## 5.1 Introduction

Lipoyl synthase (LipA) and biotin synthase (BioB) belong to a unique subgroup of RSAM enzymes which are responsible for the insertion of sulfur into inactivated carbons of octanoic acid and dethiobiotin, respectively [39, 58, 76]. A common theme shared by these enzymes is the presence of an auxiliary cluster, a [2Fe-2S] cluster in the case of the one sulfur donation by BioB [128], and a [4Fe-4S] cluster in LipA ,which catalyses a two-sulfur donation mechanism [415]. Interestingly, kinetic data on both enzymes revealed only one turnover per enzyme molecule, even in the presence of exogenous sulfur [58, 410, 414, 631], leading to the suggestion that the enzyme would work as both a catalyst and reagent, with the inserted sulfur atoms deriving from the enzyme itself, likely from auxiliary cluster degradation [39, 171]. This theory gained a strong support when Raman spectroscopy with a selenium-reconstituted BioB showed the formation of selenobiotin [411]. In an analogous manner, in lipoyl synthase, the two sulfur atoms were shown to arise from the same source [401]), with the further observation of a monothiolated cross-linked intermediate, bound to a partially degraded cluster clearly supporting the self-sacrifice mechanism [172, 416, 632].

The structural information obtained by X-ray crystallography of BioB was one of the first of the RSAM superfamily [180], showing both the native substrate, dethiobiotin, as well as SAM in the active site, helping to consolidate the understanding of the RSAM conserved activity and structure [50]. Until recently, no structure of LipA had been obtained, but thanks to the efforts of Dr. J. E. Harmer, two high resolution structures are now available: one of the structures was obtained in the presence of S-adenosyl-homocysteine (SAH), of which only the degradation product L-cysteine can be observed, and the other with the methylthioadenosine (MTA) a product of SAM degradation [182].

Despite mechanistic similarities, LipA and BioB are significantly different in their structure, with LipA only possessing a partial 6 $\beta/\alpha$ TIM barrel [182], being similar to enzymes which contain the common RSAM fold [22], whilst BioB possesses a full 8 $\beta/\alpha$ barrel [180]. In both enzymes' structures the $[4Fe\text{-}4S]_{RS}$ cluster is conserved, but the location of the proposed sulfur donor

clusters, [2Fe-2S] of BioB and second [4Fe-4S] of LipA, are located in different regions of the active site, demonstrating that even within subfamilies significant differences can be found.

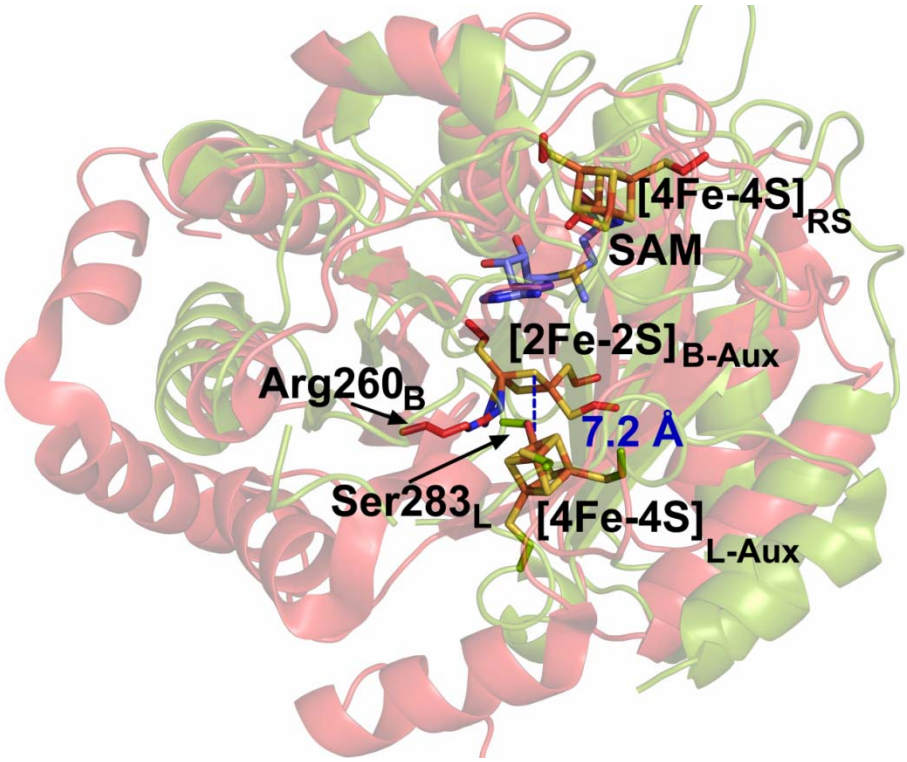



**Figure 117 Structure superposition of LipA (green) and BioB (red).**

The position of the RSAM [4Fe-4S] cluster, together with the partial 6 β/α TIM barrel fold is conserved across the enzymes, however other structural components are largely dissimilar, leading to a 7.2 Å difference in the placement of the auxiliary cluster.

Apart from the core fold, where the $[4Fe\text{-}4S]_{RS}$ is located, LipA possesses both N- and C-terminal extensions, which close the active site. In the N-terminus the $CX_4CX_5C$ cysteine triad is located, in order to bind the $[4Fe\text{-}4S]_{Aux}$ [414], while the C-terminus extension provides the highly conserved Ser283 in *TeLipA*2 which interacts with the unique iron of the $[4Fe\text{-}4S]_{Aux}$. The presence of such unique residue interaction is unprecedented in any RSAM structure and raises questions regarding its mechanistic importance, as the weaker serine-Fe coordination is often used in mutagenesis studies to remove [4Fe-4S] clusters [110, 112, 115, 118, 312, 633]. The serine is part of a conserved RXSX motif, where Arg281 is responsible for a cation-π interaction with SAM, as shown by the interaction observed with the adenosyl ring of MTA (Figure 118). The arginine

interaction works as a signal to reduce intercluster distances, from 13.98 Å in the SAH structure to 12.69 Å in the structure containing MTA [182].

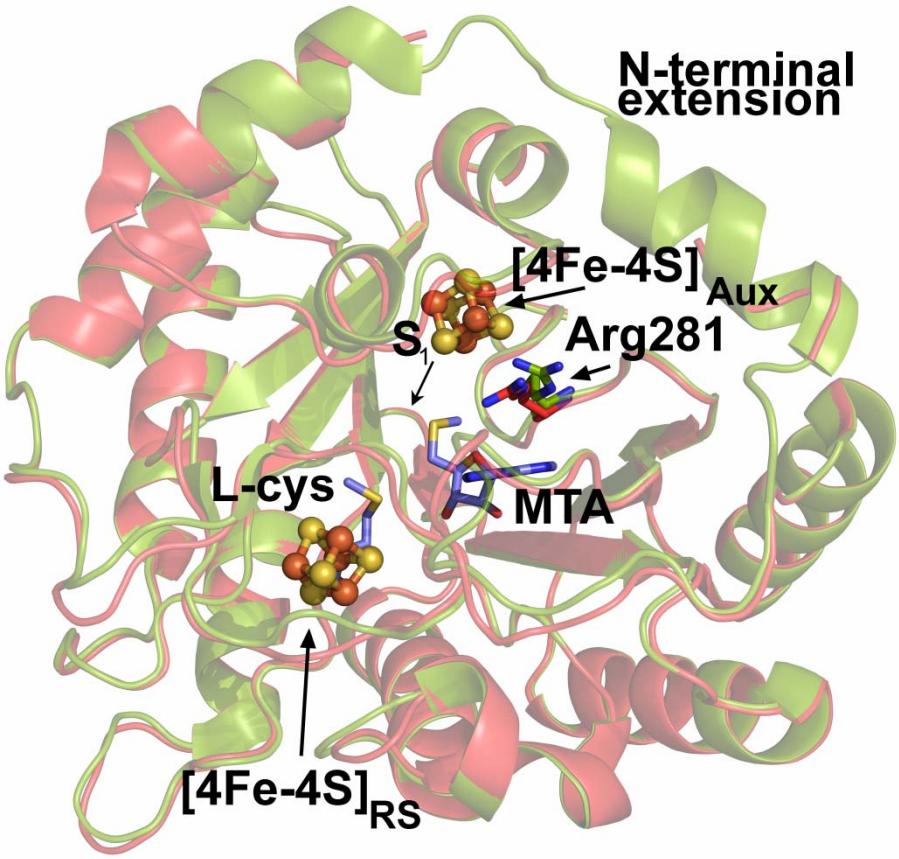



**Figure 118 Structure comparison *Te*LipA2 crystals.**

Crystal structure of *Te*LipA2 crystallized with SAM, degraded to MTA, (PDBID: 4U0O) (red), or crystallized with SAH (PDBID: 4U0P) (green), with cysteine coordination to the $[4Fe\text{-}4S]_{RS}$ cluster. In the structures a 1.3 Å change in the location of Arg281 is observed, shown interacting with MTA in the SAM crystal, but shifting away from that position in the absence of an adenosyl ring. This movement of Arg281 has been associated with sensing the presence of SAM to reduce the intercluster distances, necessary for the sulfur donation mechanism of LipA by the $[4Fe\text{-}4S]_{Aux}$ cluster.

Under reducing conditions, DTT is seen in the active site together with the SAM degradation product MTA. The presence of these molecules in the active site leads to some flexibility in the N-terminal loop (residues 13-21) with undefined density, while the structure of SAH shows an

additional $\alpha$-helix shielding the active site. The position of MTA resembles that of adenosine in SAM, being adjacent to the GxIxGxxE motif (G208-LMLGLG-E215 in *TeLipA*2) and conserved in many RSAM enzymes. The YNHNLD motif found in BioB, proposed to interact both the ribose motif of SAM as well as the substrate in that structure [180], is also present in *TeLipA*2 (Y166-NHNL-E171). However, in LipA's structure, the YNHNLD residues are located too far to bind SAM and are not predicted to interact with the octanoyl substrate.

Two channels into the bulk medium are reported to have direct access into the active site of LipA, the smallest leading to the $[4Fe\text{-}4S]_{RS}$ cluster and likely to be responsible for the interaction with the *in vivo* reductive mechanism. The larger channel is thought to be positioned between the $[4Fe\text{-}4S]_{RS}$ and the $[4Fe\text{-}4S]_{Aux}$, allowing the ocatanoyl-lysine residue from a lipoyl domain to enter [342]. Only one sulfide of the $[4Fe\text{-}4S]_{Aux}$ is positioned directly towards the active site, most likely making it the first sulfur to be incorporated into the substrate (Figure 118, $S_1$). All three other sulfide atoms that constitute the cluster are located at too great a distance away from the substrate binding site. This led to the proposal of a rearrangement of the cluster in the presence of the monothiolated intermediate taking precedence to the second sulfur transfer. Despite a lack of substrate being present in the structure, docking studies placed a flexible octanoyl-lysine in the active site, with the C6 and C8 position of the octanoyl side chain falling between SAM and the $[4Fe\text{-}4S]_{Aux}$, in the same orientation to that of DTB in BioB [180].

## 5.2 Expression and Purification of *Te*LipA2

The nature of the work described herein attempts to build upon the structural studies initiated by Dr. M. Hiscox and Dr J. E. Harmer, who successfully transformed, expressed, purified and crystallized one of the two lipoyl synthases present (*lipa*2) from *Thermosynechococcus enlongatus* (*Te*) [182, 634]. The gene coding for this lipoyl synthase (UNIPROT: Q8DLC2) had been previously inserted into the pFM024 plasmid [523] by Dr. M. Hiscox (Appendix A.14), under the control of a pBAD promoter, which is inducible with l-arabinose [522]. The protein gene was co-expressed with the *E. coli isc* operon, as this operon increases the iron and sulfide content of the protein [523]. In order to purify the protein, a hexa-histidine tag coding region was added to the C-terminal end of the gene, as a means to purify the protein by nickel affinity chromatography [521].

The purification of *Te*LipA2 followed the same general protocol used for the purification of *Ti*HydG in section 2.3.2, and as discussed, a small amount of ISC proteins were co-eluted from the different purification stages, an event which did not affect the crystallization studies of *Te*LipA2. The overall yield of purification was approximately 16 mg protein/L culture.

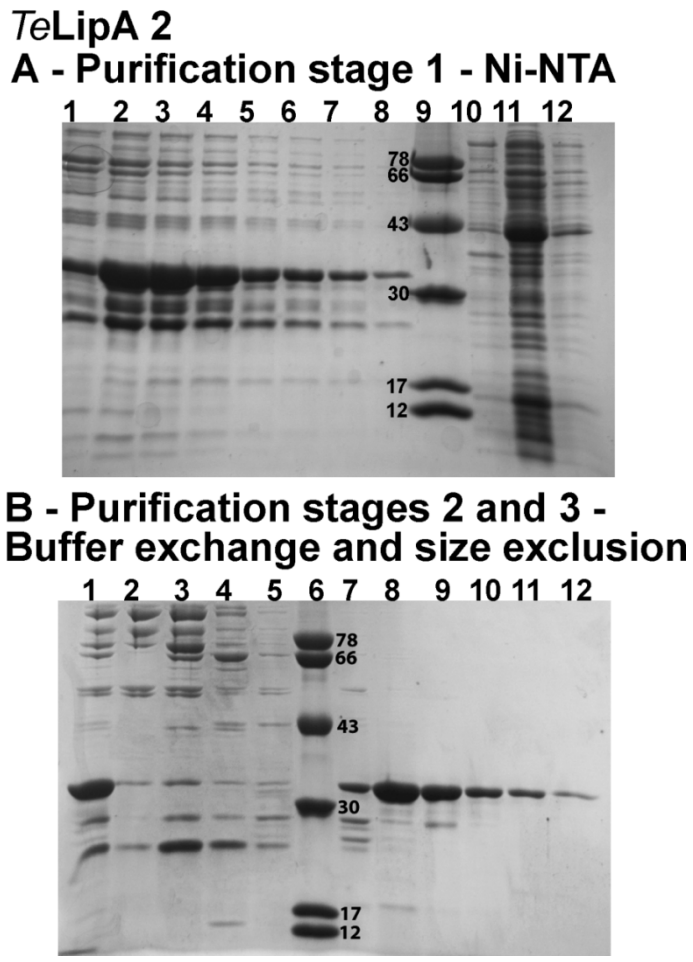



**Figure 119 SDS-PAGE gels of samples obtained in the purification of *TeLipA*2.**

**A:** Nickel-NTA affinity purification results: 1-8 – Eluted fractions containing *TeLipA*2 (32 kDa); 9 – Protein molecular weight markers; 10 – Sonicated lysate; 11 – Insoluble fraction; 12 – Clear lysate applied onto the column. **B:** Buffer exchange chromatography and S-75 gel filtration purification results: 1-3 – Fractions collected from the 50 mL S-75 used for buffer exchange; 4-5 - Initial fractions eluted from the 300 mL S75 gel filtration column, corresponding to an aggregate of protein and excess [Fe-S]; 6 – Protein molecular weight markers; 7-12 – Golden-brown band, corresponding to the soluble *TeLipA2*, with the correct 32 kDa size.

### 5.2.1 Chemical Reconstitution of *Te*LipA2

Iron content and UV-vis studies on *TeLipA*2 prior and following reconstitution, yielded results similar to the previously reported by Cicchillo and colleagues [415]. The UV-spectra following reconstitution has both the proteinacious 280 nm peak as well as a clear defined 410 nm peak, indicating the presence of [4Fe-4S] cluster(s) [534].

**Table 16 Protein and iron content estimation at different stages of *TeLipA*2 purification.**

| Protein | Total protein Ni-NTA (mg) | Iron content before 1st reconstitution (molar equivalents) | Iron content after 1st reconstitution (molar equivalents) | Total Protein gel filtration (mg) | Iron content before 2nd reconstitution (molar equivalents) | Iron content after 2nd reconstitution (molar equivalents) |
|---|---|---|---|---|---|---|
| ***Te*LipA2** | 79 | 4.43 ± 0.7 | 7.24 ± 1.8 | 33 | 4.02 ± 2.3 | 6.22 ± 0.8 |

Following chemical reconstitution, the extinction coefficient values at 410 nm rose from $\varepsilon_{410}$= 16,900 $cm^{-1}$ $M^{-1}$ obtained in the non-reconstituted sample (4.0 μM) to $\varepsilon_{410}$=27,700 $cm^{-1}$ $M^{-1}$ for the reconstituted sample (3.5 μM), indicative of close to two $[4Fe\text{-}4S]^{2+}$ clusters per protein molecule [534, 635]. Comparing these results to those of *Ti*HydG (section 2.3.3), another enzyme containing two $[4Fe\text{-}4S]^{2+}$ clusters, shows that despite an initial higher cluster content, indicative of less degradation by imidazole, *Te*LipA 2 does not reconstitute as readily, leading to a lower than the expected 410 nm peak. Consistent with these results is the presence of an additional small shoulder at 320 nm, a peak which has been associated with the presence of $[2Fe\text{-}2S]^{2+}$ clusters [128], and which has previously been reported in purified lipoyl synthase by Kriek and colleagues [523]. The UV-Vis results correlate well with the iron quantification: the number of irons per molecule of *TeLipA*2 increased from 4.43 ± 1.5 in samples prior to reconstitution to 6.22 ± 0.8 following chemical reconstitution.

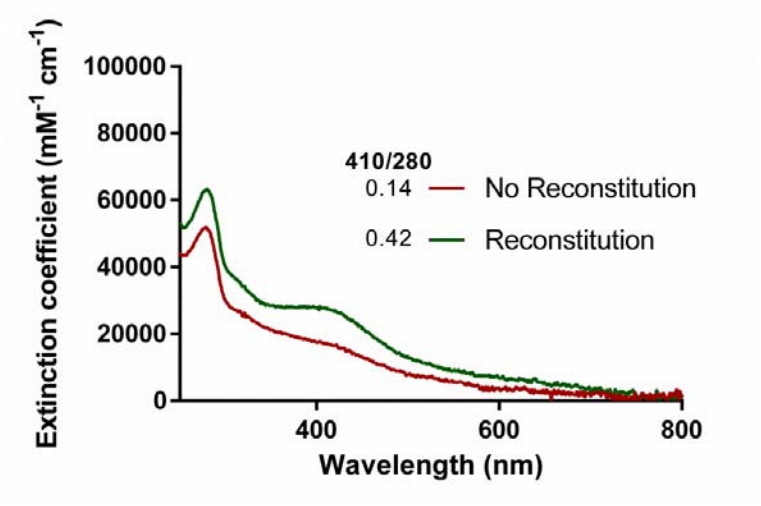



**Figure 120 UV-vis spectra of *Te*LipA2.**

Red trace indicates samples prior to any reconstitution ([*Te*LipA2]= 4.0 μM) and green trace shows the UV-vis spectra of the reconstituted sample ([*Te*LipA2]= 3.5 μM). The increase in the peak at 410 nm is indicative of $[4Fe\text{-}4S]^{2+}$ assembly, while the shoulder at 320 nm indicates the presence of $[2Fe\text{-}2S]^{2+}$ clusters.

## 5.3 Crystallization of *TeLipA2* in Higher Glycerol Content

The purification of *TeLipA2* in buffers with high glycerol content (10% w/v glycerol), together with the presence of PEG molecules in the crystallization condition (15% w/v 20,000 PEG), allows for direct freezing of the crystals [636] without significant macroscopic damage. Although small amounts (around 1%) of a given cryoprotectant, such as glycerol, are often shown to be sufficient for efficient cryoprotection [637], the levels of glycerol introduced into the system during the experiment may reach as high as 50% [638]. To verify whether any changes could be observed in the diffraction quality of crystals when the glycerol content was increased, a set of experiments was carried out where glycerol was introduced into the drop prior to incubation for crystal formation (Method 9.19).

**Table 17 Glycerol concentration in crystallization experiment with *Te*LipA2.**

Table summarizing the glycerol concentrations in the protein and well solutions (left) as well as the final concentrations present in the crystallization drop (right)

| Initial % of Glycerol | | % of glycerol in the final crystallization condition | | |
|---|---|---|---|---|
| **Protein solution** | **Well solution** | **1:1 (protein/well solution)** | **1:2 (protein/well solution)** | **2:1 (protein/well solution)** |
| 10 | 0 | 5 | 3.33 | 6.67 |
| 13.5 | 4.8 | 9.1 | 7.7 | 10.6 |
| 18 | 9.1 | 13.5 | 12.1 | 15.0 |
| 22.5 | 13.0 | 17.8 | 16.2 | 19.3 |
| 27 | 16.7 | 21.8 | 20.1 | 23.6 |

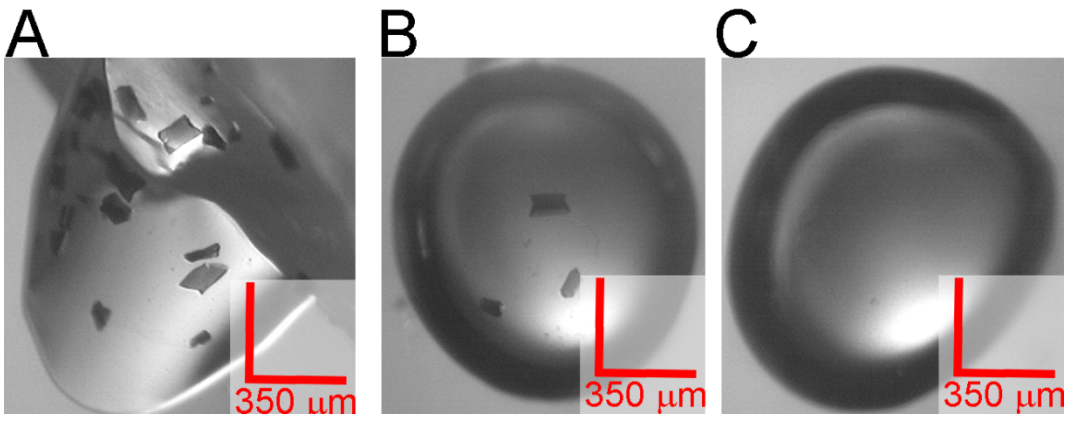



**Figure 121 Crystals obtained for *TeLipA2* in the presence of increasing concentrations of glycerol**

A: 7.7% glycerol content in the drop. B: 10.6% of glycerol content in the drop. C: 17.8% glycerol content in the drop.

As shown in Figure 121 the increasing concentration of glycerol decreased the amount of crystals per drop, with concentrations above 15% being devoid of crystals. The crystals also suffered slight morphological changes, becoming more compact as the glycerol concentration increased. Crystal samples grown in three different drops (7.7% glycerol; 10.6% glycerol and 15% glycerol) were directly frozen and tested in Diamond Synchrotron Light Source I04, at 100K and 12 kEv.

Despite the widespread use of glycerol as a cryoprotectant [548, 638], the growth of crystals in presence of a higher concentration of this substance showed a dramatic reduction in the quality of the diffraction data obtained (Figure 15), even at the smallest increase (from 5% to 9%). The results for higher glycerol contents consistently show that any increase in the cryoprotectant concentration during the crystallization stage yields very low diffraction crystals.

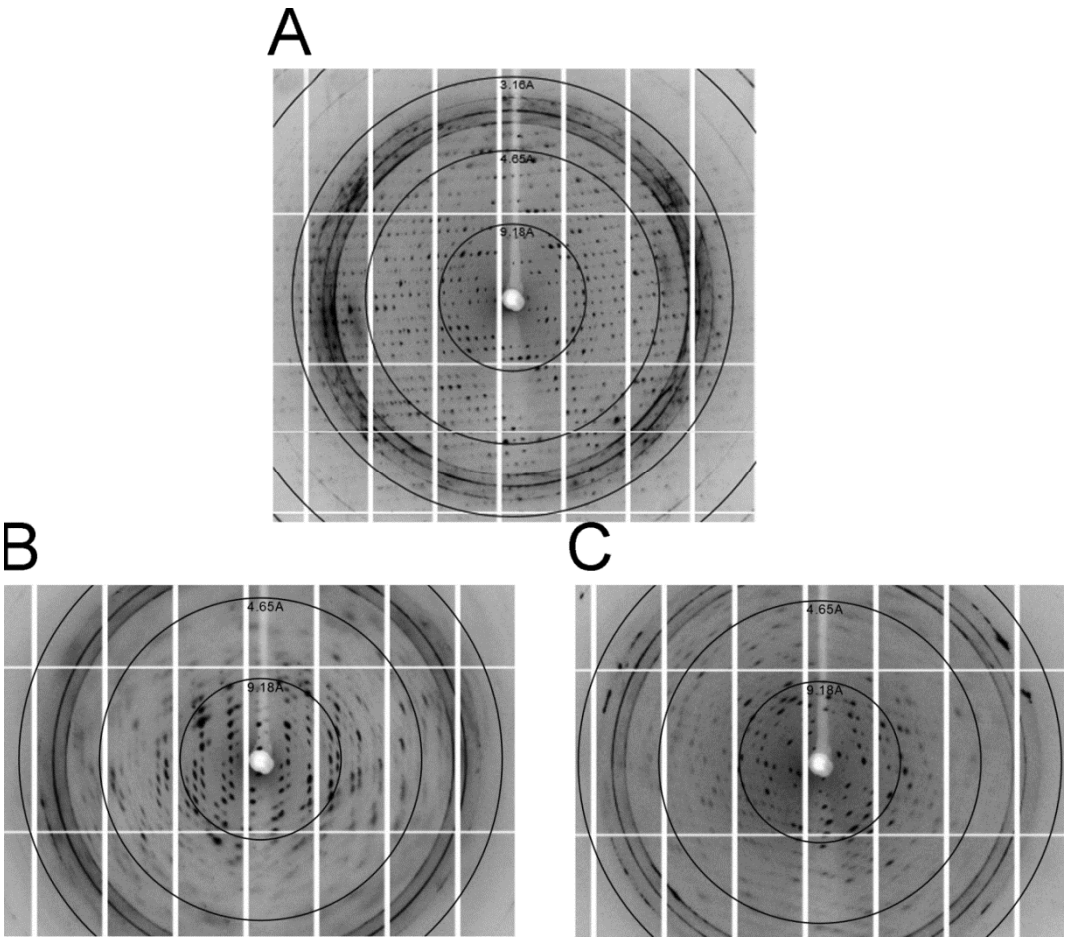



**Figure 122 Crystal diffraction from *TeLipA2* grown in the presence of increasing concentrations of glycerol**

**A:** 7.7% glycerol. **B:** 10.6% glycerol. **C:** 15.0% glycerol. The increase of glycerol led to overall lower resolutions in the data collection.

## 5.4 Chemical Reconstitution of *TeLipA*2 Under Different Reducing Conditions

DTT is the most common reducing agent used during the chemical reconstitution step [38, 79, 408, 410, 639-642]. However its presence at high concentration (5.0 mM) in the crystallization system of *Te*LipA2 leads to DTT being present in the active site of the structure grown in the presence of SAM [182]. This may influence the ability to co-crystallize with an intact SAM molecule, leading to structural changes that could reducing the ability of the system to bind an octanoyl-lysine substrate.

An experiment was designed to reduce the total amount of DTT present in the protein sample prior to crystallization, by testing the extent of chemical reconstitutions in the presence of lower (1.0 mM) quantities of DTT (termed $DTT_{low}$). Other reductants, namely β-mercaptoethanol (BME) and tris(2-carboxyethyl)phosphine) (TCEP), were also tested at the same concentration of 1.0 mM, prior to observing the UV-vis differences that accompanied these changes. Additionally a reconstitution was carried out where only half the amount of $FeCl_3$ and $Na_2S$ was added (2.5 protein molar equivalents of each). The resulting sample was termed $Fe_{low}$ and aimed to provide a control for a presumably more incomplete reconstitution.

Despite its reduction potential being unknown, TCEP is often considered the strongest reducing agent of the three tested [643] and is reportedly [644] able to reduce both BME, with an $E_0$=-0.26 V measured at pH 7.0 and 25ºC against a standard hydrogen electrode [645], and DTT, with an $E_0$= -0.33 V measured at pH 7.0 against a $DPN^+$-DPNH system [646]. However, the similarities in the UV-spectra when the reconstitution is carried out with either TCEP or BME strengthen the idea that the role of DTT during the reconstitution step is greater than simply providing reducing equivalents. The lower extinction coefficients at $\varepsilon_{410}$= 14,000 $cm^{-1}$ $mM^{-1}$ for TCEP (2.7 μM) and $\varepsilon_{410}$= 21,800 $cm^{-1}$ $mM^{-1}$ for BME (2.6 μM), together with the high background signal at higher wavelengths are indicative of an incomplete reconstitution and the formation of unspecific iron-sulfide aggregates [534].

In contrast with these results, when the reconstitution is carried out with DTT present, even with a 5-fold decrease in the concentration of DTT to 1.0 mM or a two-fold decrease in the molar equivalents of $FeCl_3$ and $Na_2S$ added, the reconstitution occurs to a more significant extent. The extinction coefficients at $\varepsilon_{410}$= 20,700 $cm^{-1}$ $mM^{-1}$ and $\varepsilon_{410}$= 23,000 $cm^{-1}$ $mM^{-1}$ for $DTT_{low}$ (3.9 μM)

and $Fe_{low}$ (4.3 μM), respectively, seem to suggest the presence of more than one cluster reconstituted on average [523].

These results seem to show that the iron chelating ability of DTT [647, 648], with its two thiol groups, plays an important role in the reconstitution step. This role of DTT is also alluded to in the increased background noise at high wavelength in samples without DTT, most likely arising from non-specific iron sulfide formation (compare results in Figure 123 with those of Figure 120).

Additionally, protein quantification after reconstitution shows that the loss of protein is much less extensive when DTT is present in the system following reconstitution, with 4.0 μM of *Te*LipA2 being estimated for both reconstitutions with DTT, while BME/TCEP predicts less than 3.0 μM. This decrease in the concentration of protein is also indicative of a more extensive protein aggregation with iron sulfide when DTT is not present in the system.

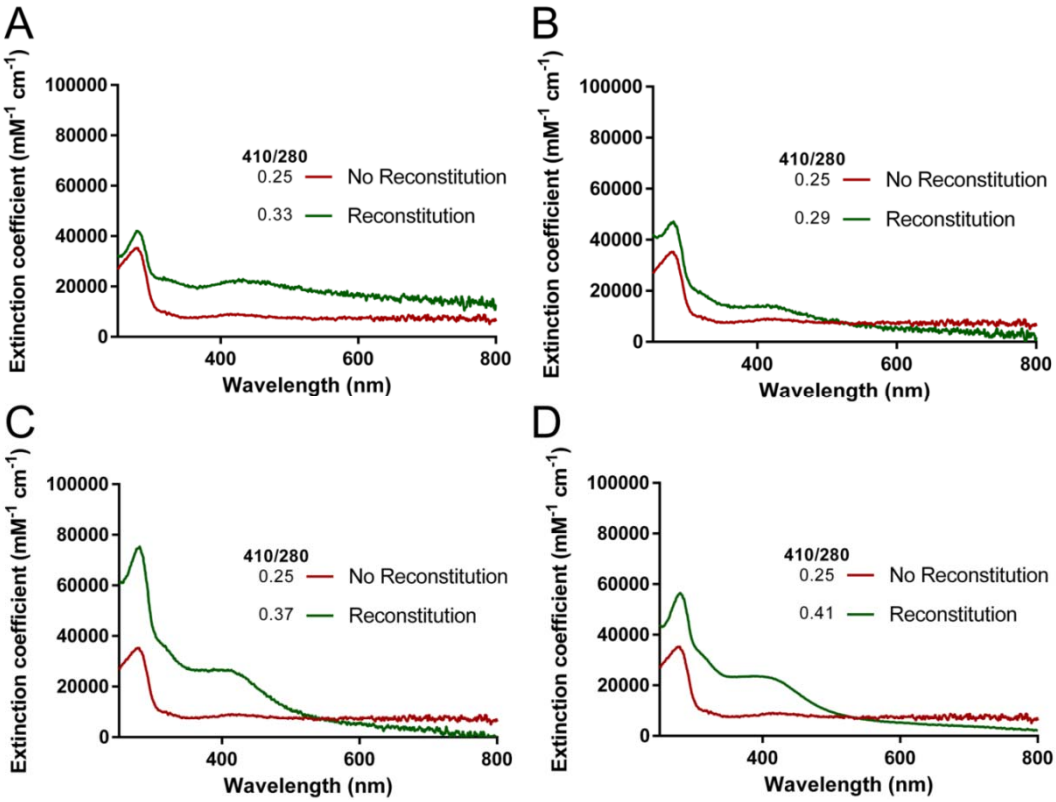



**Figure 123 UV spectra of *TeLipA2* under different reducing conditions, together with the calculated 410/280 values.**

There is a clear contrast between *Te*LipA2 samples reconstituted in the absence of DTT, **A:** BME (2.6 μM) and **B:** TCEP (2.7 μM), showing a less prominent 410 nm peak following reconstitution (green); and samples with DTT present, **C:** 1 $DTT_{low}$ (3.9 μM) and **D:** $Fe_{low}$(4.3 μM), which after reconstitution (red trace) show a more well defined characteristic $[4Fe\text{-}4S]^{2+}$ signal.

### 5.4.1 Crystallization of *TeLipA2* in Different Reconstitution Conditions

A crystallization experiment was carried out using *Te*LipA2 samples reconstituted in the different reducing conditions discussed in the previous section, together with 10 mM SAM. Reducing the overall DTT concentration in the sample would have hopefully led to an absence of the reducing agent in the active site, accommodating an intact molecule of SAM instead. Unfortunately, the changes to the reconstitution method interfered with the ability to form crystals in both a 96-well format, using Hampton PEGRx HT commercial broad screen, as well as in 24-well format under the condition where crystals had previously been obtained (100 mM Bicine, pH 8.5, 3% dextran sulfate sodium salt, 15% PEG 20,000).

In order to increase the chances of crystal formation, a micro seeding experiment was carried out in 24-well format (Method 9.10). A pre-grown crystal of *Te*LipA 2 reconstituted in the presence of 5 mM DTT was shattered into small fragments in a small volume of well mixture. The seed stock was used for 10-fold dilution series and 0.1 μL of these seeding solutions were then added to fresh drops with protein reconstituted with the different reducing agents. Despite the addition of seeds to help with the nucleation stage, only the sample reconstituted with 2.5 molar equivalents of $FeCl_3/Na_2S$ showed any crystal formation (Figure 124). The crystallization of *Te*LipA2 with reduced reconstitution equivalents could lead to an incomplete cluster formation, as hinted by its UV-Vis spectra with a lower $\varepsilon_{410}$, and obtaining the crystal structure of the enzyme with the auxiliary cluster in a partially degraded stage provide some information regarding the fate of the cluster after sulfur donation [632].

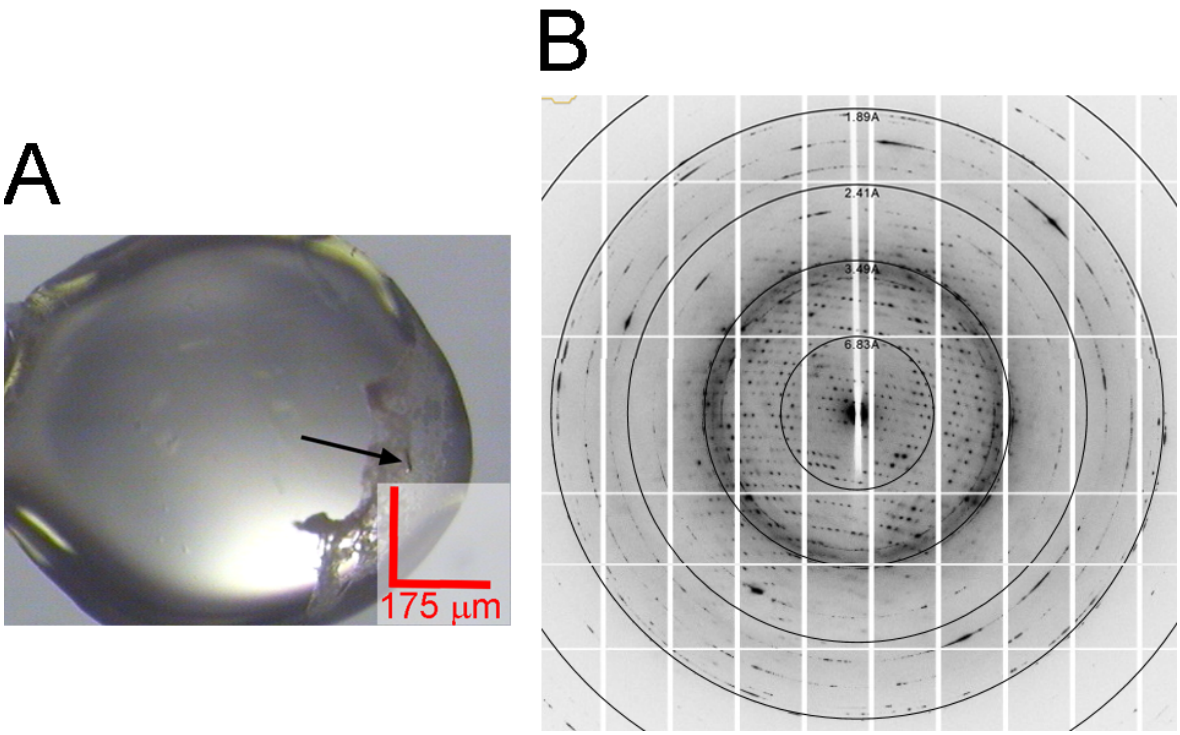



**Figure 124 Crystallization of *Te*LipA2 $Fe_{low}$.**

**A:** Small crystal obtained for *Te*LipA2, where the reconstitution step was achieved with 2.5 molar equivalents of $FeCl_3$ and $Na_2S$. 0.8 mM of enzyme were mixed with SAM and allowed to equilibrate for 24h prior to addition of seeds. A small crystal was observed at 1.0 x $10^{-4}$ dilution alone. Crystallization condition H12 from Hampton PEGRx HT (0.1 M Bicine pH 8.5, 3% w/v Dextran sulfate sodium salt, 15% PEG 20,000). **B:** Crystal diffraction at 12 keV, to 2.4 Å maximum resolution, from the *Te*LipA2 $Fe_{low}$.

The crystal was tested at Diamond Light Source, beamline I03 under a stream of nitrogen to ensure a constant temperature of 100 K. Both data collection and data refinement statistics are shown in Appendix A.15. The medium resolution (2.4 Å) data possesses good values for completeness (97%), multiplicity (3.7), signal-to-noise (10.3) and $R_{merge}$ (0.056), and molecular replacement identified a very high score model (TFZ 48.2), indicating the structure is most likely conserved between the crystal data and the model (PDBID: 4U0O).

Despite all data pointing towards a simple case of structure refinement, manual improvement of the model in WinCoot followed by refinement in *Phenix.refine* never lowered the $R_{work}/R_{free}$ values

below 0.225/0.277, slightly above average for data at this resolution range. Additionally refinement analysis by POLYGON showed a significantly higher B-factor (Figure 125). The reasons behind this much higher temperature factor are two-fold. First, in the protein backbone two loops that were previously stable are now highly disordered and no residues could be built (42-48 and 242-248). The first loop (42-48) is responsible for the binding of the auxiliary cluster, and despite density for all three cysteines being present the increased disorder may have arisen from a more incomplete reconstitution (Figure 126). The second reason for such a high B-factor arises from the molecule of DTT bound to the $[4Fe\text{-}4S]_{RS}$ cluster, with 94.3 $Å^2$, indicating the molecule is moving rather freely. This is not unexpected, as the binding of DTT to the cluster is an artefact of the crystallization condition, and not mechanistically relevant.

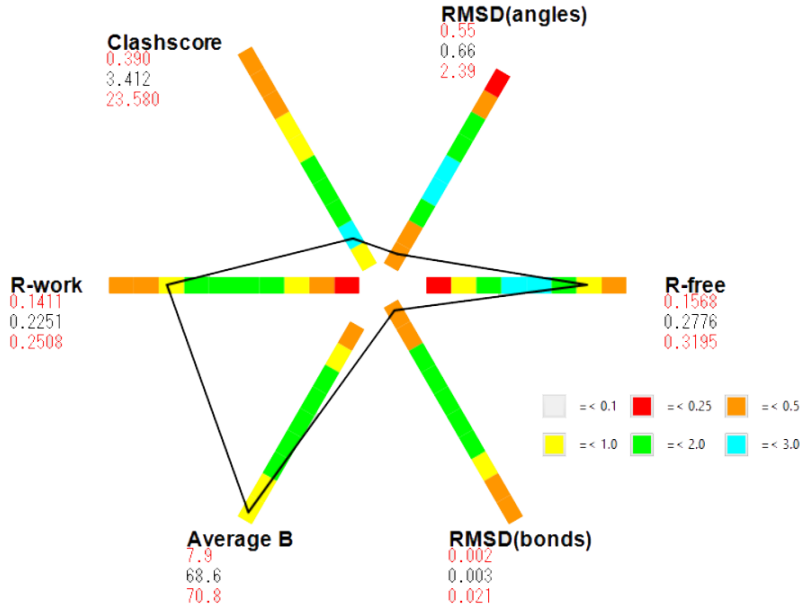



**Figure 125 POLYGON analysis of the *Te*LipA2 $Fe_{low}$ crystal**

Comparison of several quality indicators to structures with similar resolution within the PDB database. Histograms are coloured based on weighted number of structures with similar resolution.

The structural information obtained reveals the presence of both clusters present (Figure 126), perhaps unexpected when considering the UV-vis and iron quantification results (section 5.4). A likely explanation to these results arises from the increased random protein aggregation observed in the crystallization stage (Figure 124), indicating that a significant portion of the protein present

in the drop did not crystallize. It is possible that protein molecules that did not contain both clusters were less stable and formed random aggregation, whilst the few protein molecules in solution which were able to be fully reconstituted despite the less favourable conditions, were able to successfully crystallized.

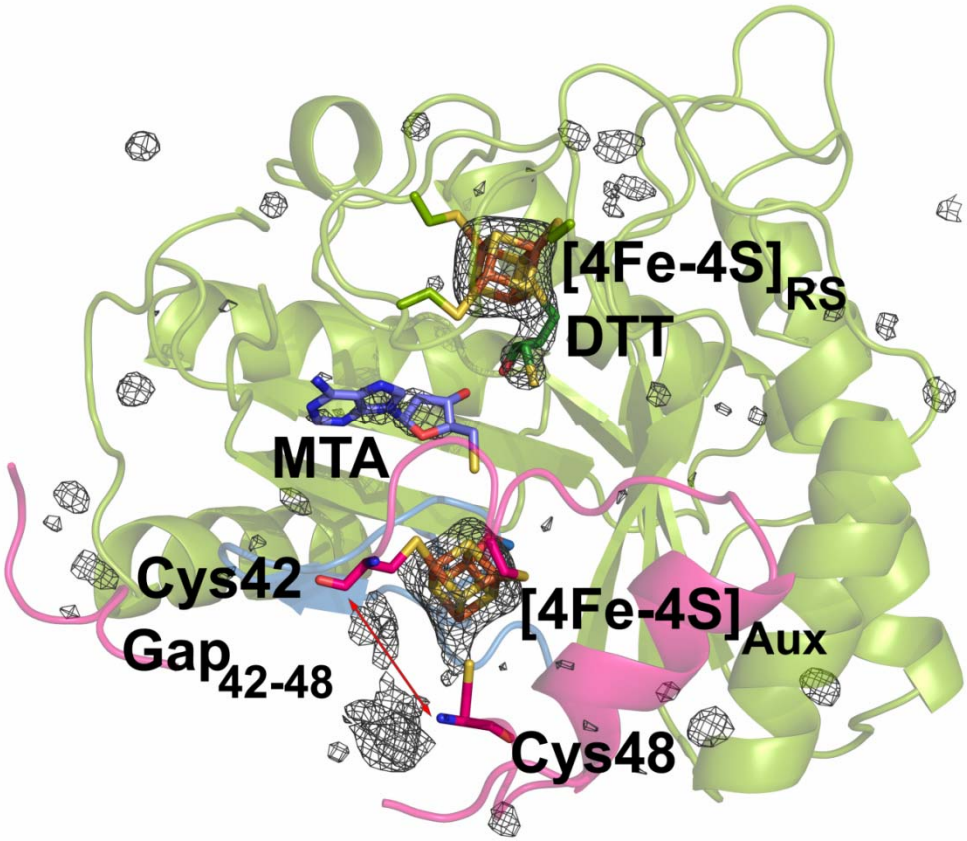



**Figure 126 Model and density map of *Te*LipA2 $Fe_{low}$.**

N-terminus is shown in pink, central RSAM domain in green and C-terminus in blue. mFo-DFc difference map shown in the absence of the ligands (σ= 3.0), revealing density that can be attributed to each cluster, MTA and DTT, together with some density belonging to the residues that constitute loop 42-48. Both [4Fe-4S] clusters are shown in yellow and orange, 5'-MTA in lilac and DTT in green.

## 5.5 Chemical Reconstitution of *Te*LipA2 With Sodium Selenide

When BioB was first described as a catalyst of sulfur donation, two possible donating groups were proposed [649]: an Fe-S centre of the protein [76, 650] or a substrate sulfur species [651] possibly DTT from the buffer. It had been previously demonstrated that SAM consumed in the reaction was not to the source of the sulfur [405], and over the years evidence accumulated that the $[2Fe\text{-}2S]_{Aux}$ cluster was the most likely source of the sulfur [76, 77, 408, 410, 652]. In a study by Bui and colleagues, BioB was reconstituted with iron and selenide to follow the formation of selenobiotin [411]. It showed the ability to selectively form a $[2Fe\text{-}2Se]_{Aux}$ cluster and that selenobiotin was formed upon turnover, strengthening the case for self-sacrifice in BioB [39, 171, 653].

The incorporation of selenium into proteins is a staple method of protein crystallography [654, 655], as a means to obtain anomalous diffraction, leading to the solution of the phase problem (see section 8.3.2.3). Combining the ability to identify the location of selenium atoms in a crystal structure, with the formation of an selenium intermediate of the reaction for lipoyl synthase could strengthen the theory that the mechanism of sulfur insertion is conserved in BioB and LipA, arising from a self-sacrifice mechanism of the auxiliary cluster [401].

The reconstitution steps were replicated as before (see section 7.2), however $Na_2S$ was exchanged for $Na_2Se$. Two methodologies were tested: with one method carrying out both reconstitution steps $Na_2Se$ and the second using $Na_2S$ for the first reconstitution, and $Na_2Se$ for the second reconstitution following size exclusion chromatography (Method 9.6.1).

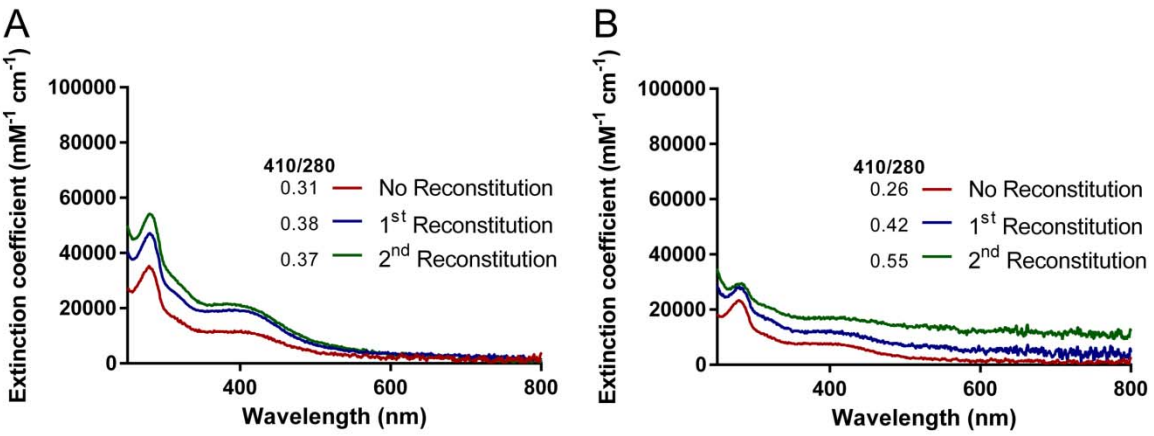



**Figure 127 UV spectra of *TeLipA2* reconstituted with iron(III) chloride and sodium selenide**

The reconstitution of TeLipA2 with sodium selenide leads to a less defined 410 nm peak. **A:** First reconstitution (blue) carried out with $FeCl_3$/$Na_2S$ and the second with $FeCl_3$/$Na_2Se$ (green). Following the second reconstitution the peak near 410 nm did not increase significantly. **B:** Both reconstitutions carried out with $FeCl_3$/$Na_2Se$, leading to an increase in protein and/or iron selenide aggregation, as shown by the background increase at high wavelength.

Samples reconstituted with sodium selenide had a much higher tendency to aggregate, with a much more significant portion of the protein lost at each purification stage. As a result there is a higher than average background signal in the UV spectra (Figure 127), indicative of non-specific Fe-Se interactions. The signal of the cluster specific 410 nm peak did not display very significant changes with the chemical reconstitution step with both reconstitutions being carried out with $Na_2Se$, a result similar to the reported for BioB, where the characteristic peaks were found to be less well defined when $Na_2Se$ was added to the proteins [411]. The sample prior to being reconstituted (4.4 μM) shows $\varepsilon_{410}$= 7,250 $cm^{-1}$ $M^{-1}$, and increases to $\varepsilon_{410}$= 12,000 $cm^{-1}$ $M^{-1}$ after the first $Na_2Se$ reconstitution (3.2 μM), raising further slightly after the second reconstitution (2.9 μM) $\varepsilon_{410}$= 16,900 $cm^{-1}$ $M^{-1}$. The lack of defined features that follow the addition of sodium selenide meant the sample which was initially reconstituted with $Na_2S$ shows a chromatographic profile closer to the usual reconstitution (compare Figure 127 A and Figure 120).

The addition of $Na_2S$ dramatically changes the extinction coefficient at 410 nm from $\varepsilon_{410}$= 10,900 $cm^{-1}$ $M^{-1}$ prior to reconstitution (4.6 μM) to $\varepsilon_{410}$= 18,600 $cm^{-1}$ $M^{-1}$ after reconstitution (4.3 μM).

The addition of 5 Meq of $FeCl_3/Na_2Se$ for the second reconstitution changes the spectroscopic profile slightly, with $\varepsilon_{410}$= 21,000 $cm^{-1}$ $M^{-1}$ after both reconstitution processes (2.9 μM).

### 5.5.1 Crystallization of *Te*LipA2 Reconstituted With Selenium

The successful incorporation of selenium into the auxiliary cluster of *TeLipA2* was to be confirmed X-ray crystallography [656]. Unfortunately, the use of sodium selenide as a reconstitution partner led to a severe inhibition of crystal formation. As mentioned previously, samples reconstituted twice with $Na_2Se$ have a high tendency for aggregation. This tendency led to the rapid formation of an amorphous dark precipitate following the mixture with the crystallization condition, and no crystalline features were apparent Figure 128 A.

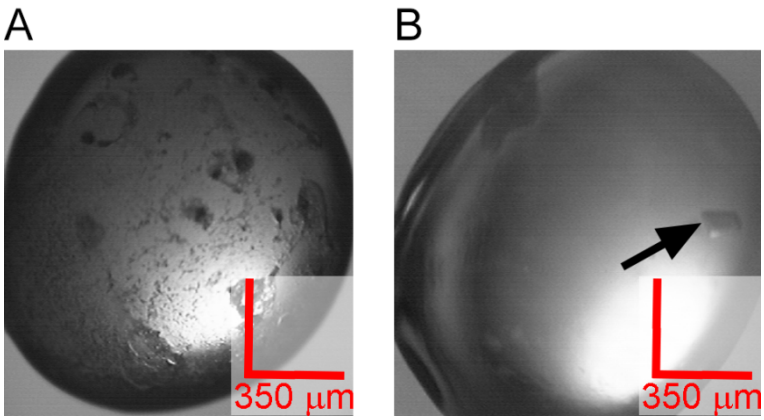



**Figure 128 Crystallization of *Te*LipA2, where the reconstitution was carried out with $FeCl_3$ and $Na_2Se$.**

When both reconstitution steps carried out with $Na_2Se$ (**A**), no crystalline formation was observed and dark aggregates were seen forming. If *Te*LipA2 was initially reconstituted with $FeCl_3/Na_2S$, and only during the second reconstitution was $Na_2Se$ added to the sample, crystallization was not as severely impeded (**B**) and a single crystal was observed. Crystallization condition H12 from Hampton PEGRx HT (0.1 M Bicine pH 8.5, 3% w/v Dextran sulfate sodium salt, 15% PEG 20,000).

The sample which was reconstituted once with $Na_2Se$, however, showed the formation of a single crystal (B), with much lower random aggregation occurring in the crystallization condition. The crystal was collected and directly frozen in liquid nitrogen prior to testing in Diamond Light Source facilities, beamline i03. The data was collected at high energy (12 kEv), near the selenium κ-edge. The fluorescence scan confirmed the presence of selenium in the samples. Unfortunately, the quality of the diffraction data is too low for refinement and analysis (Appendix A.16). The high $R_{merge}$, together with a very low signal to noise ratio indicate either very low quality data, or wrong attribution of symmetry. Manual integration of the diffraction data was attempted using MOSFLM [657], but so far indexation with different predicted unit cell/space groups has not improved the subsequent integration and no structural information has been derived from this experiment.

## 5.6 Crystallization of *Te*LipA2 in the Presence of Substrate Analogues

The crystal structure of *Te*LipA2 [182] allowed for the creation of a potential model of substrate binding, placing the octanoyl moiety between SAM bound to the [4Fe-4S]RS cluster and the $[4Fe\text{-}4S]_{Aux}$ cluster. The presumed distance between the first accepting carbon of the substrate (C6) and the proximal sulfur atom of $[4Fe\text{-}4S]_{Aux}$ is 6.4 Å, an intramolecular distance consistent with the self-sacrificial model. The distances between C8 and any of the additional sulfurs is, however, over 8 Å [182], making it plausible the second donation reaction only occurs after some cluster rearrangement, potentially with iron loss [416].

To try and obtain a substrate-bound structure, substrate analogues were initially designed, synthesized and tested by Dr. M. Hiscox, who showed that both molecules can be dithiolated by the enzyme [634]. The substrate analogues are hexa-peptides with a lysine containing an octanoyl chain appended, mimicking the natural substrate of the enzyme.

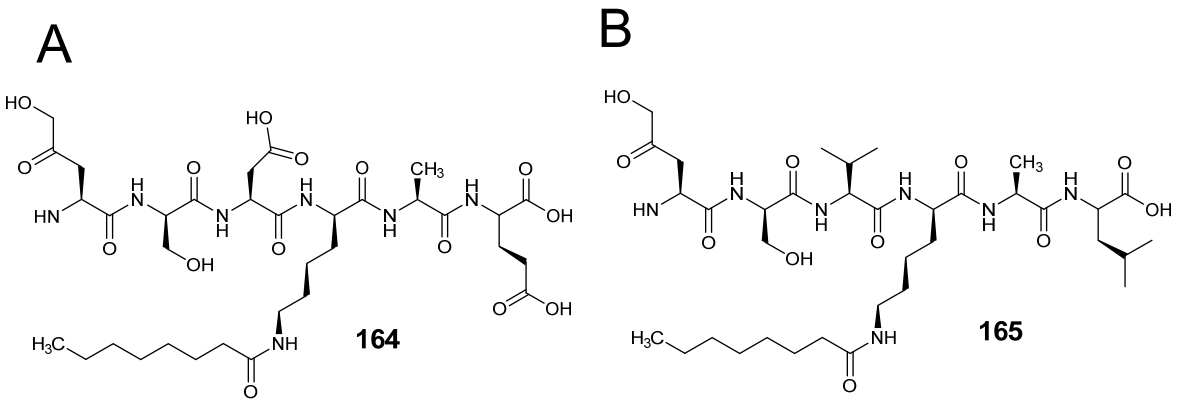



**Figure 129 Two small peptide analogues of the native substrate of LipA.**

**A:** ESDK(o)AD **164**; **B:** ESVK(o)AV **165**. Each letter corresponds to an amino acid, with lysine having an appended octanoyl unit (o).

The low solubility of the substrate analogues at biological pH (25 mM maximum solubility) hampered any crystallographic studies with substrate concentration above 5 mM. A broad crystallization screen, using Hampton PEGRx HT was initially attempted (Method 9.8), using 0.75 mM *Te*LipA2 together with 10 mM SAM, 5 mM of either substrate and either the presence or absence of 0.1 mM sodium dithionite. The addition of the reducing agent served as an attempt to collect a crystal of either the product or of the intermediate-bound compound.

The crystallization screening identified one condition where crystal growth was achievable: 0.1 M bicine pH 8.5, 3% (w/v) dextran sulfate sodium salt, 15% (w/v) PEG 20,000; the same condition used to crystallize *Te*LipA2 [182]. The crystallization conditions were replicated in a 24-well format, where crystals were grown in the presence of both substrates (Figure 130).

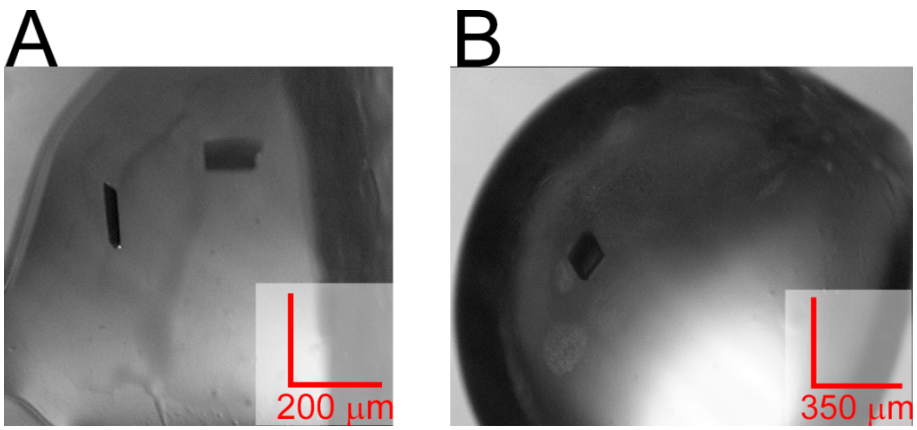



**Figure 130 Crystals of *TeLipA2* grown in the presence of SAM and a synthetic substrate**

**A:** Co-crystallization with 5 mM ESDK(o)AD **164**. **B:** Co-crystallization with 5 mM ESVK(o)AV **165**.

The crystals obtained for *TeLipA2* with substrate analogues were exposed to X-ray radiation at Diamond Light Source, at a constant temperature of 100 K, on beamline I24. An anomalous data set was not required, as the structure of the native enzyme LipA was available. That allowed the structure to be solved, using only a high resolution dataset obtained at 0.975 Å and molecular replacement with the native structure [557]. The diffraction images for the crystals tested were processed automatically by X-ray Detector Software (XDS) [554]. The data reduction and scaling were done automatically using Pointless and SCALA [555] and data processing statistics are summarized in Appendix A.17. Using the previously reported structure of *TeLipA*2 (PDBID: 4U0O) for molecular replacement, a high-scoring solution was obtained with *Phenix.phaser* [557]. Immediately following molecular replacement, the difference map was visualized, to foresee if the substrate analogues would be present in the structure, prior to further refinement. Unfortunately neither of the two substrate analogues were proven suitable for co-crystallization, as Figure 131 A and B show, as in neither case the peptide is seen interacting with LipA. With the exception of the amino acid backbone of the enzyme, the difference map only shows density that can be allocated to both [4Fe-4S] clusters, together with MTA and a RS cluster bound DTT. As no relevant information could be extracted from either of the data sets, further refinement was not pursued.

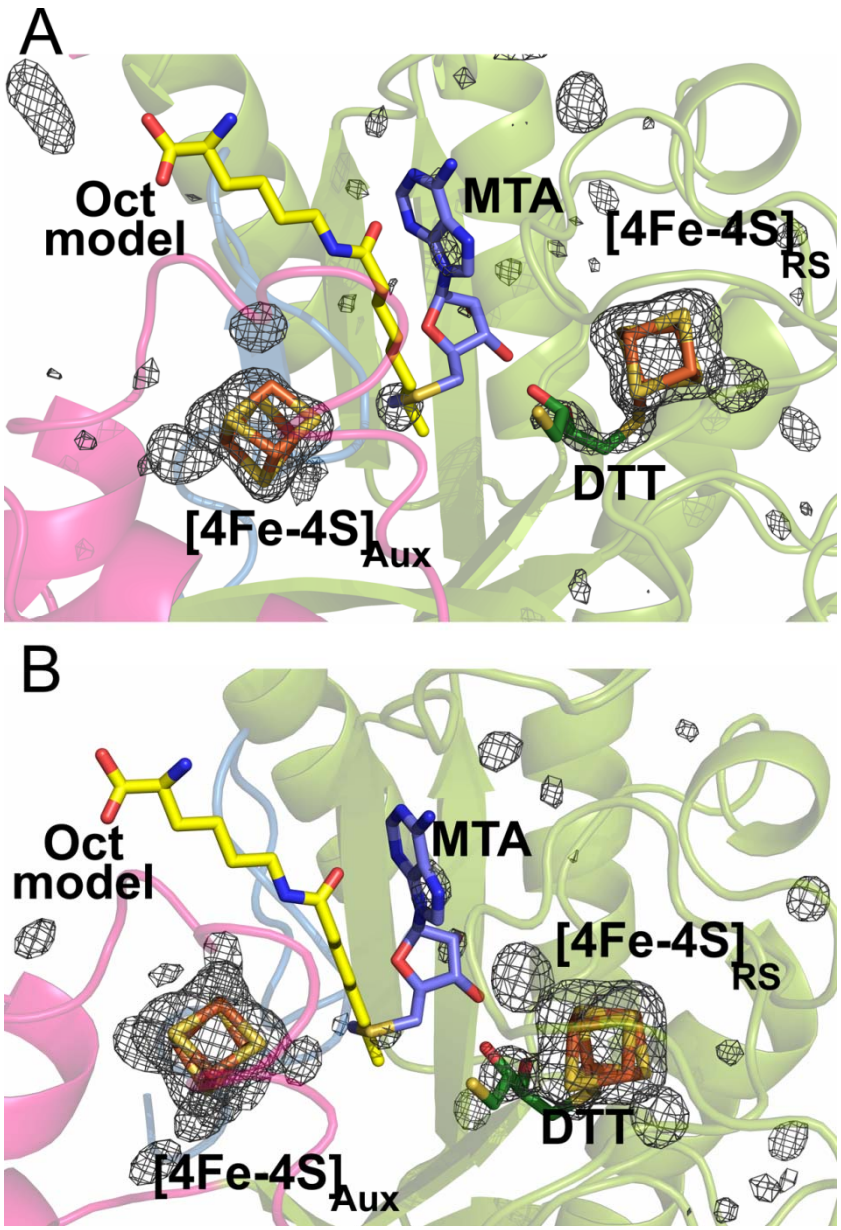



**Figure 131 Model and density maps of *TeLipA2* co-crystallized with substrate peptide analogues**

N-terminus is shown in pink, central RSAM domain in green and C-terminus in blue. mFo-DFc difference map shown in the absence of the ligands ($\sigma$= 3.0). Ligands are modelled from PDBID: 4U0O. [4Fe-4S] clusters are shown in yellow and orange, 5'-MTA in lilac and DTT in green. A: Crystallization in the presence of ESDK(o)AD peptide **158**. The difference map does not show density for 5'-MTA nor the substrate analogue. B: Crystallization in the presence of ESVK(o)AV peptide **159**. No synthetic peptide density can be observed. Both structures are shown with the predicted model of binding of an octanoyl substrate (yellow). The modelling was carried out by Andreas Illiopoulos and Dr. Stephen Fox, both supervised by Prof. Jon Essex, University of Southampton. In both crystallization experiments, no density for the predicted, or any alternative ligation mode of an octanoyl substrate, is observed.

### 5.6.1 Crystallization of *Te*LipA2 in the Presence of MTAN

The two crystal structures available for LipA (PDBID: 4U0O, 4U0P) showed degradation of the co-substrate SAM or related compound SAH [182]. The absence of the adenosyl moiety in the structure obtained with SAH led to very small changes in the structure (see section 5.1, Figure 118) and is currently unknown whether an intact SAM molecule could show further changes in the protein fold. Those changes could lead to the identification of further key residues, or be dramatic enough to allow the presence of a substrate analogue into the active site.

It has been reported [658] that 5'-methylthioadenosine/*S*-adenosyl-L-homocysteine nucleosidase (MTAN) is capable of cleaving the adenosyl products of SAM degradation. This activity has been explored in kinetic assays of RSAM enzymes [47, 484, 659, 660] as a means to decrease potential product inhibition [658, 661]. A crystallization experiment was set-up to determine whether the ability of MTAN to remove the accumulation of SAM degradation products, in this case MTA, could be exploited. By incubating *Te*LipA2 with SAM and MTAN (see Method 9.23), the levels of the degradation by-product in the system were expected to be low enough for the formation of the crystals to occur with an intact SAM molecule instead.

The presence of MTAN did not completely inhibit the formation of crystals, though, remarkably, crystals only seemed to form with one of the two available substrate analogues, ESDK(o)AD **164**. The crystal obtained for *Te*LipA2, grown in the presence of MTAN and the substrate analogue ESKD(o)AD was exposed to X-ray radiation at Diamond Light Source, at a constant temperature of 100 K, on beamline I04-1. The structure of the native enzyme (PDBID: 4U0O) was used for molecular replacement[557] and data collection and Phaser statistics are summarized in Appendix A.18.

Despite some selectivity to form crystals in regards to the different substrate analogues present in the system, the active site shows no density corresponding to a small oligonucleotide bound to the protein. Additionally, the structural information obtained from the crystal formed in the presence of MTAN is analogous to the previous reported structures of *TeLipA*2 grown in the presence of SAM [182], as the addition of MTAN did not significantly alter the presence of MTA in the active site of the enzyme (Figure 132, red density). It is possible that the reductive cleavage of SAM occurs following formation of the crystal, making the release of MTA from the active site, as well as the binding of a new SAM molecule more limited. Assuming the levels of MTA in the drop are relatively low due to the action of MTAN, the binding of SAM to the active site of the enzyme

with release of MTA could be considered a case of replacement soaking, where one bound molecule is exchanged in the active site with a molecule that exists in high concentration in the medium [662, 663], and the correct ratio of both ligands will need to be further explored for a successful crystallization with SAM.

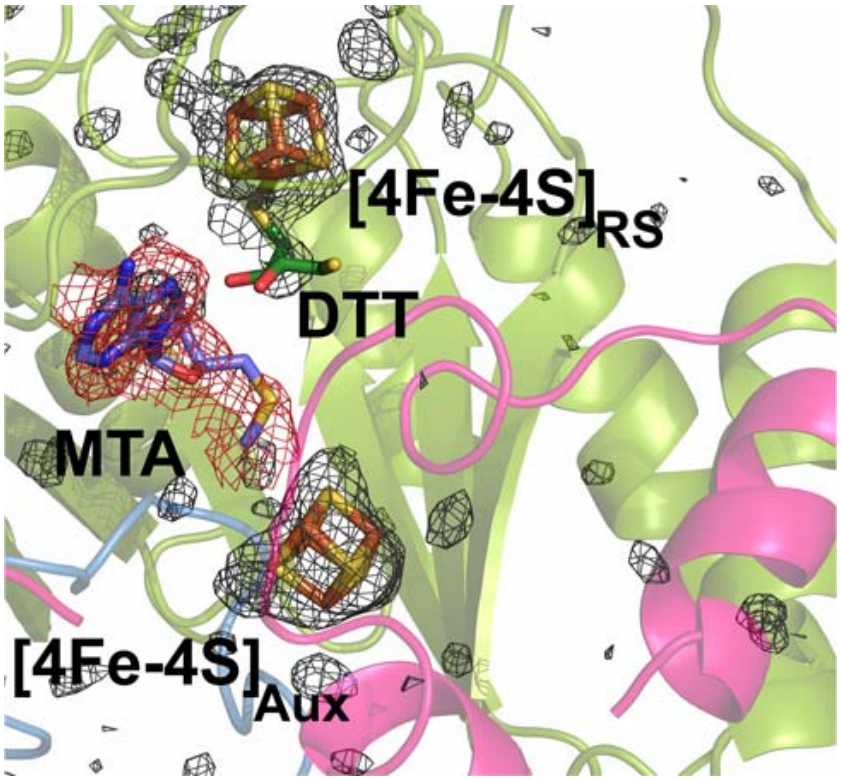



**Figure 132 Model and density map of *TeLipA2* obtained from a crystal grown in the presence of MTAN and ESKD(o)AD.**

The N-terminus is shown in pink, the partial TIM barrel in green and the C-terminus in blue. Both [4Fe-4S] clusters are shown in yellow and orange, 5'-MTA in lilac and DTT in green. The mFo-DFc difference map (cutoff $\sigma$= 3.0) is observed as a black grid, revealing density for both clusters, DTT and 5'-MTA. An additional mFo-DFc difference map (cutoff $\sigma$= 1.5), in red, makes clearer the presence of 5'-MTA in the active site.

## 5.7 Mutation of the $[4Fe\text{-}4S]_{Aux}$ Cluster Coordinating Serine in *Te*LipA2

*Te*LipA2's crystal structure revealed an intriguing coordination of the $[4Fe\text{-}4S]_{Aux}$ cluster by Ser283, a highly conserved residue in most lipoyl synthases [182]. The precise role of the residue in the activity of the enzyme has not yet been fully established, though the mutation of the cluster binding serine ( Ser283) into cysteine or alanine revealed, for both mutants, no formation of lipoyl product and a reduced turnover of SAM when incubated with an octanoyl substrate analogue [182]. The ability of these mutants to turnover SAM contrasts with results for an *E.coli* LipA mutant which lacked the $[4Fe\text{-}4S]_{Aux}$ cluster. In the absence of the auxiliary cluster, no SAM formation was observed [415], indicating that the presence of the secondary cluster is both necessary for SAM turnover and unaffected to some extent by the serine mutation. Similarly, biotin synthase also possesses a unique residue, in BioB's case arginine, bound to its $[2Fe\text{-}2S]_{Aux}$ cluster [180] and mutation of the arginine into cysteine led to a two-third decrease in the activity of the enzyme [133]. The combined results seem to indicate that the sulfur donating mechanism for both enzymes requires non-traditional ligands to the auxiliary cluster as an important aspect of their activity.

More intriguingly, activity studies by Dr. M. Hiscox showed a significant difference between the alanine and cysteine mutants. While both alanine and cysteine mutants have a significant effect in the formation of the lipoyl product and in SAM turnover, the cysteine mutant was shown to be able to form measurable amounts of the 6-thio-octanoyl intermediate, which were not observed in the alanine mutant [634]. These results, together with the possibility of cluster disassembly [416], led to the proposal that Ser283 may need to dissociate from the unique iron after the formation of a substrate intermediate to allow the re-adjustment of the auxiliary cluster for the second sulfur abstraction (see 1.3.4.2 Figure 26).

The mutagenesis of Ser283 into alanine or cysteine was achieved by synthesizing two DNA fragments with the desired mutations that could be inserted into the BglII/XhoI restriction sites available on the pFM024 plasmid (see pMH003 plasmid map, appendix A.14). The synthesis of the fragments and ligation into the appropriate location were carried out by Eurofins MWG.

### 5.7.1 Purification of Expressed *Te*LipA2 Ser283Cys and Ser283Ala Mutants

The purification method was as described for *Ti*HydG in section 2.3.2, and similarly to the results obtained in the two mutants of *Ti*HydG (section 3.3), there is a sharp contrast in the levels of protein aggregation of *Te*LipA2 Ser283Cys when compared with *Te*LipA2 Ser283Ala. During the purification stage, the *Te*LipA2 Ser283Cys mutant shows an equivalent level of stability to that of the wild-type and an overall higher yield of protein is obtained (approximately 25 mg protein/L culture). That is not the case for the *TeLipA2* Ser283Ala mutant, where the ability to be kept in solution without aggregation suffers significantly, leading to a much lower yield of protein during the purification stages (less than 4 mg protein/L culture as the final yield of an expression/purification experiment). The significant differences observed between the two mutants seem to indicate that the coordination of the unique iron of the $[4Fe\text{-}4S]_{Aux}$ has a strong role in the overall structure of the protein. When a cysteine is introduced it is likely that a coordination of the auxiliary cluster is still possible, perhaps even more strongly so than the wild type enzyme, leading to a stable protein. When the mutation introduces an alanine, however, coordination of the auxiliary cluster becomes impaired and the overall stability of the protein suffers significantly.

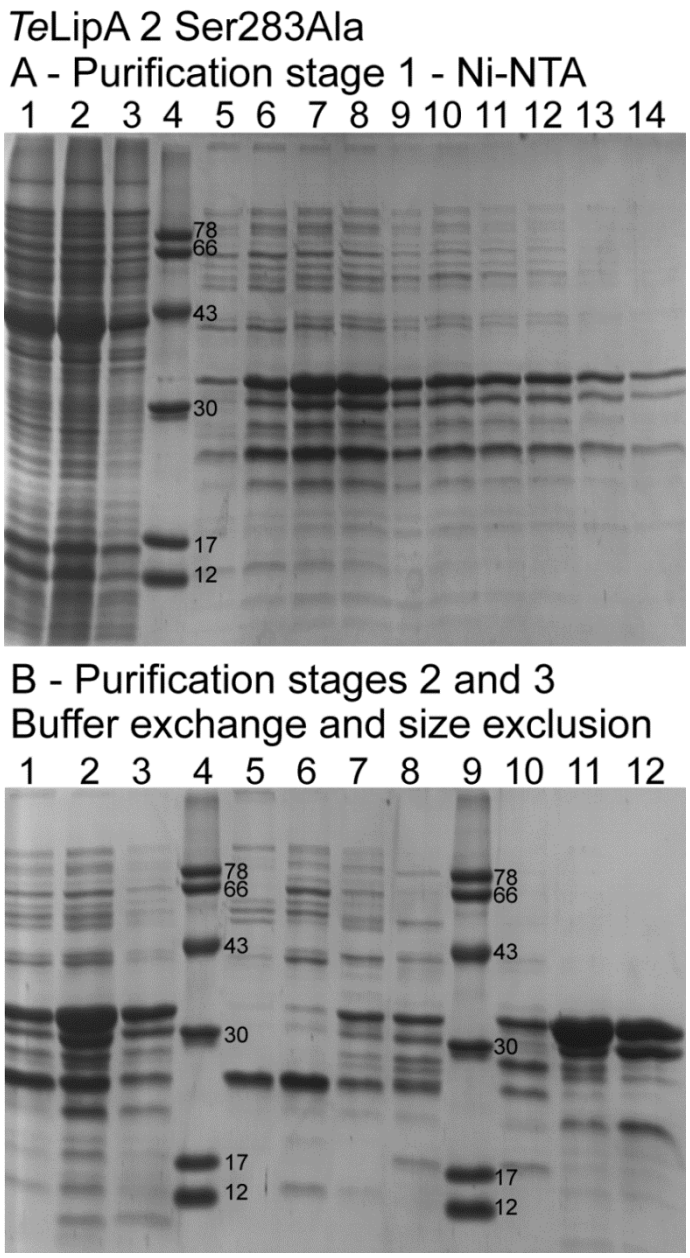



**Figure 133 SDS-PAGE gels of samples obtained in the purification of *TeLipA2* Ser283Ala**

**A:** Nickel affinity purification results: 1 – Ins - Insoluble fraction; 2 – Sonicated lysate; 3 – Flow through of the nickel affinity column; 4 – Protein molecular weight markers; 5-14 - Eluted fractions containing *TeLipA*2 Ser283Ala (32 kDa). **B:** Buffer exchange and S-75 size exclusion results: 1-3 – Fractions collected from the 50 mL S75 used for buffer exchange; 4,9 – Protein molecular weight markers; 5-8 – Dark aggregate of protein and excess [Fe-S]; 10-12 – Golden-brown band, corresponding to the soluble *TeLipA2* Ser283Ala at 32 kDa.

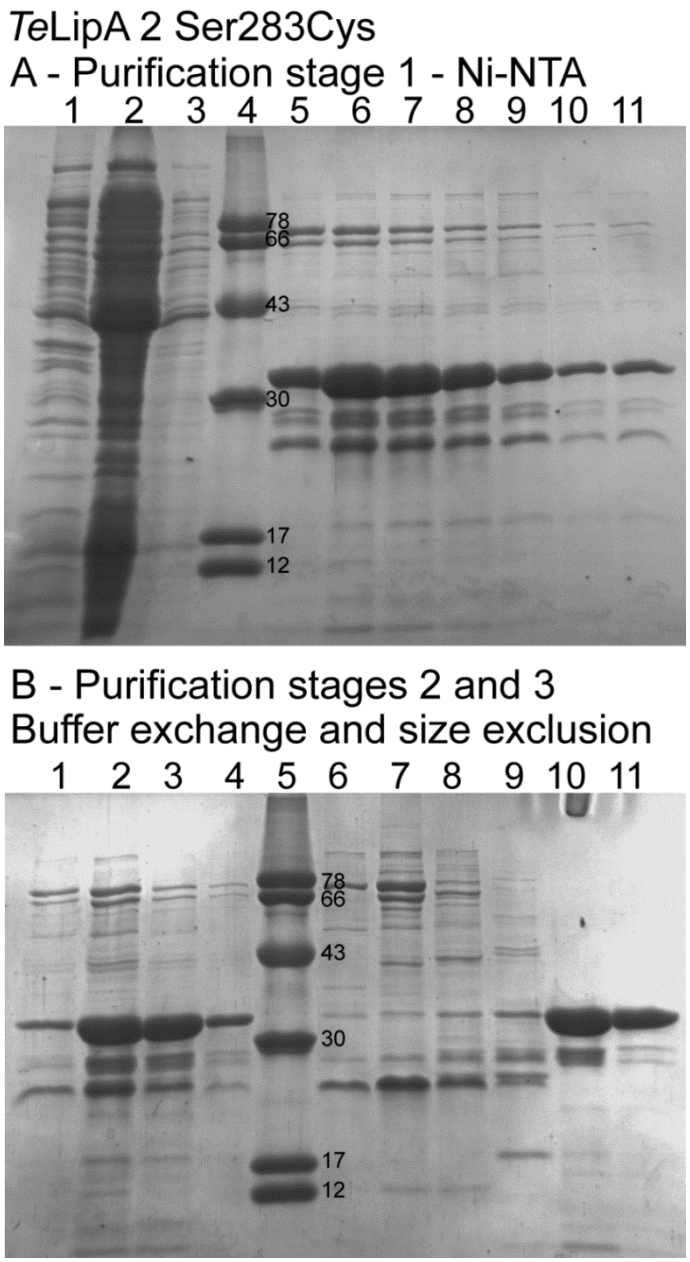



**Figure 134 SDS-PAGE gels of samples obtained in the purification of *TeLipA2* Ser283Cys**

**A:** Nickel affinity purification results: 1 – Ins - Insoluble fraction; 2 – Sonicated lysate; 3 – Flow through of the nickel affinity column; 4 – Protein molecular weight markers; 5-14 - Eluted fractions containing *TeLipA*2 Ser283Cys (32 kDa). **B:** Buffer exchange and S-75 size exclusion results: 1-3 – Fractions collected from the 50 mL S75 used for buffer exchange; 4,9 – Protein molecular weight markers; 5-8 – Dark aggregate of protein and excess [Fe-S]; 10-12 – Golden-brown band, corresponding to the soluble *TeLipA2* Ser283Cys at 32 kDa.

**Table 18 Protein and iron content estimation at different stages of *TeLipA*2 mutants' purification.**

| Protein | Total protein Ni-NTA (mg) | Iron content before 1st reconstitution (molar equivalents) | Iron content after 1st reconstitution (molar equivalents) | Total Protein gel filtration (mg) | Iron content before 2nd reconstitution (molar equivalents) | Iron content after 2nd reconstitution (molar equivalents) |
|---|---|---|---|---|---|---|
| TeLipA2 Ser283Ala | 18 | 1.06 ± 0.4 | 3.81 ± 1.9 | 11 | 3.22 ± 1.6 | 5.64 ± 1.8 |
| TeLipA2 Ser283Cys | 128 | 5.35 ± 0.2 | 8.54 ± 2.0 | 43 | 5.52 ± 2.8 | 7.09 ± 1.3 |

The mutation of Ser283Ala does not seem to greatly impair the chemical reconstitution of both clusters, as can be seen in the UV spectra (Figure 135 B). Despite some background noise at higher wavelength, the extinction coefficient of the 410 nm peak ($\varepsilon_{410}$= 23,100 $cm^{-1}$ $M^{-1}$, [*Te*LipA2 Ser283Ala]= 3.3 μM) is a value relatively close to the reported for *Te*LipA2 wild type (see section 5.2.1), and significantly above the 15,000 $cm^{-1}$ $M^{-1}$ usually attributed to one $[4Fe\text{-}4S]^{2+}$ present [534, 635]. Additionally, protein stability seems to increase following chemical reconstitution, making this step vital following purification to prevent extensive protein loss. Over time, however, the protein starts to aggregate, possibly indicative of cluster degradation due to lack of correct coordination.

UV-Vis spectrum analysis of *Te*LipA2 Ser283Cys shows, in accordance with the iron quantification (Table 18 ), a higher than average cluster presence in the protein prior to any reconstitution. Following the chemical addition of $FeCl_3$/$Na_2S$ the extinction coefficient of the 410 nm peak rises from 21,400 $cm^{-1}$ $mM^{-1}$ ([*Te*LipA2 Ser283Cys]= 2.4 μM) to 30,200 $cm^{-1}mM^{-1}$ ([*Te*LipA2 Ser283Cys]= 2.1 μM), a value consistent with two $[4Fe\text{-}4S]^{2+}$ clusters present, and higher than the value observed for the wild type enzyme [534, 635].

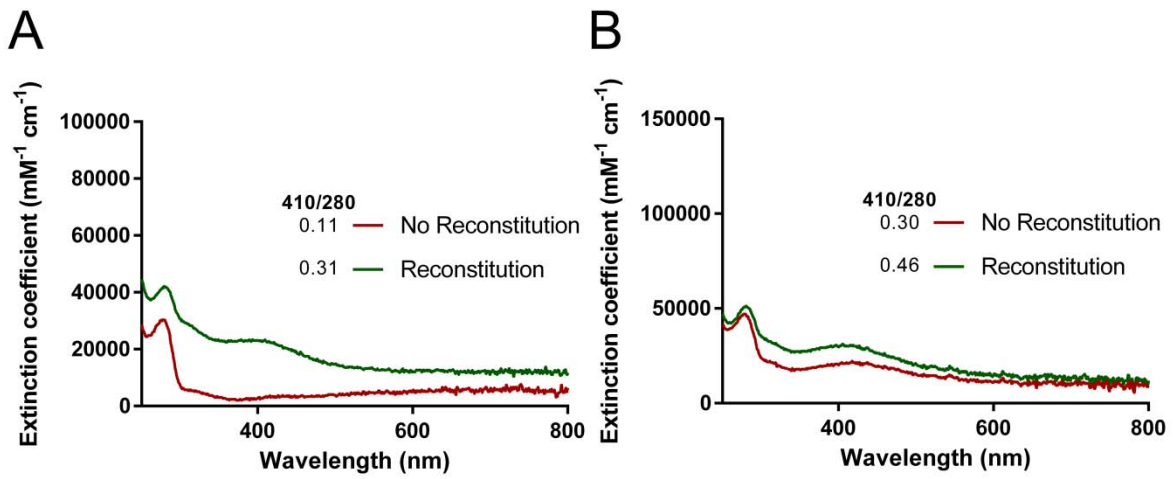



**Figure 135 UV spectra foe *Te*LipA2 Ser283 mutants**

Following purification the spectra of *Te*LipA2 Ser283Ala shows no presence of the $[4Fe\text{-}4S]^{2+}$ characteristic peak at 410 nm (**A**, red), and chemical reconstitution restores the spectra (**A**, green) to levels close to that of the wild type enzyme. Conversely, *Te*LipA2 Ser283Cys seems to be purified with a high content of both [4Fe-4S] clusters in place (**B**, red) and chemical reconstitution only increases the values at 410 nm slightly (**B**, green).

### 5.7.2 Crystallization of *Te*LipA2 Ser283 Mutants

Both mutant proteins from *Te*LipA2 were successfully crystallized under the same condition that yielded the crystals of the wild-type enzyme [182], with small variations. Crystals of *Te*LipA2 Ser283Cys were obtained in the exact same conditions, at 0.8 mM protein with 10 mM SAM, after 24h incubation in condition H12 of Hampton PEGRx screen (0.1 M Bicine pH 8.5, 3% dextran sulfate sodium salt, 15% PEG 20,000 (Method 9.9). *Te*LipA2 Ser283Ala, however, proved much harder to crystallize, as the protein precipitated rapidly following any freezing/thawing cycle. This created a difficult situation where crystallization could only be attempted immediately after a purification cycle had been carried out.

With a fresh batch of protein, *TeLipA2* Ser283Ala was shown to form spherulites [664] at 0.4 mM, a lower concentration than the native enzyme. Unfortunately, these results were hard to replicate and the use of the spherulites as either microseeds, by crushing them and doing serial dilutions (Method 9.10), or as macroseeds, where one spherulite was transferred to a freshly prepared

mixture of *Te*LipA2 Cys283Ala and precipitant, never yielded additional crystals, nor changed the morphology of the crystal during the macroseeding experiment.

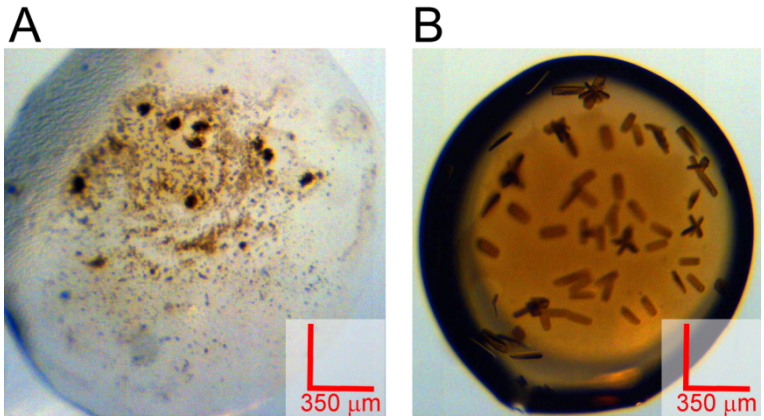



**Figure 136 Crystallization of *Te*LipA Ser283 mutants.**

**A:** Spherulite formation obtained with *Te*LipA2 Ser283Ala. **B:** Crystals obtained with *Te*LipA2 Ser283Ala mutant. Condition H12 from Hampton PEGRx (0.1 M Bicine pH 8.5, 3% w/v dextrane sulfate sodium salt, 15% w/v PEG 20,000).

Several crystals obtained for the Ser283Cys mutant of *Te*LipA2 were exposed to X-ray radiation at Diamond Light Source, at a constant temperature of 100 K, on beamline I03. Despite growing easily and being morphologically similar to the native *Te*LipA2 crystals, all but one of the crystals of the mutated enzyme diffracted to 8 Å or worse, and that single full dataset was collected at medium resolution (3.5 Å) (Data collection statistics in Appendix A.19).

The structure was solved using molecular replacement unit Phaser-MR present in Phenix framework [557]. This model was manually improved by use of WinCoot [561], followed by successive refinement cycles in phenix.refine [665] until the R-values could not be improved further. However, the relatively low resolution of the dataset severely limited the ability to refine it to a high standard, and a high $R_{work}/R_{free}$ was obtained that could not be decreased (0.31/0.36).

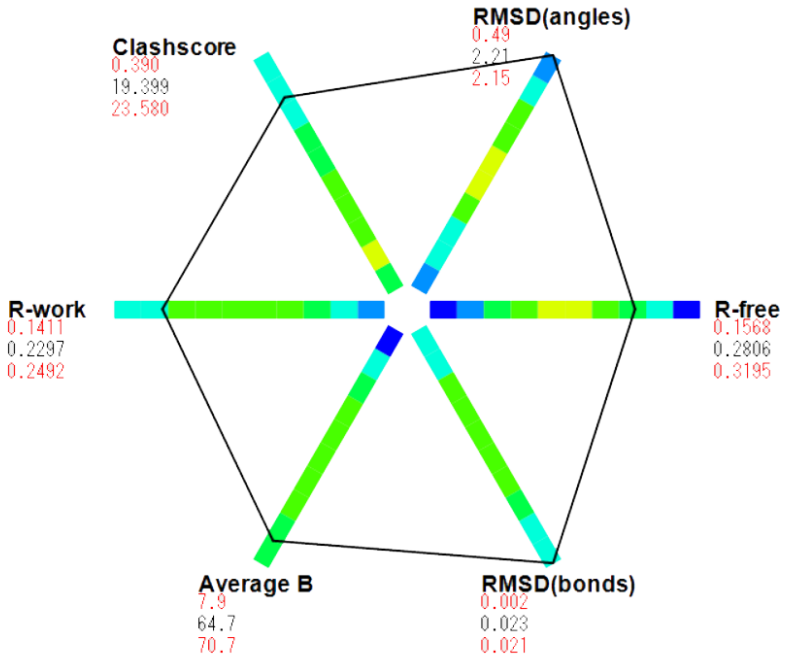



**Figure 137 POLYGON analysis of the *Te*LipA2 Ser283Cys crystal**

Comparison of several quality indicators to structures with similar resolution within the PDB database. Histograms are coloured based on weighted number of structures with similar resolution.

The structure of Ser283Cys is very similar to that of the native enzyme (RMSD= 2.0 Å), as can be seen in Figure 138, with Cys283 forming a bond to the auxiliary cluster of the protein. The general conservation of the structure of the auxiliary cluster goes in accordance with the idea that the inhibition of product following Ser283Cys mutation does not arise from inability to coordinate the $[4Fe\text{-}4S]_{Aux}$ cluster. Instead, cysteine coordination may prevent the release of the unique iron following the formation of the mono-thiolated intermediate, preventing the proposed cluster re-arrangement [416]. The rearrangement of the cluster seems necessary for positioning the second sulfur atom from the cluster nearer the C8 radical intermediate.

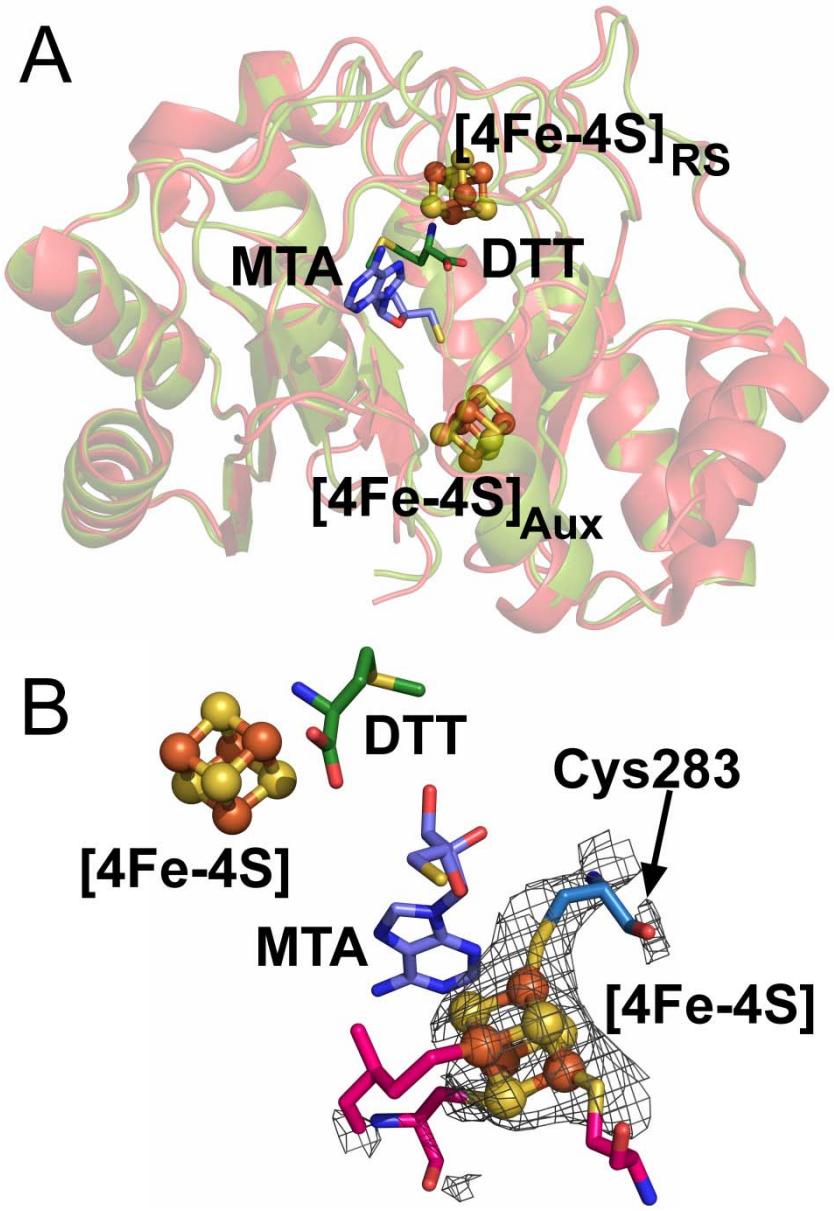



**Figure 138 Structure comparison of native *TeLipA2* with the Ser283Cys mutant.**

**A:** Superposition of the structure of native *Te*LipA2 (blue) and the mutant *Te*LipA2 Ser283Cys (green). **B:** Auxiliary [4Fe-4S] cluster bound by four cysteines. The bottom cysteine (Cys283) in the *TeLipA2* Ser283Cys (green) is seen coordinating an iron of the $[4Fe\text{-}4S]_{Aux}$ cluster and is superimposed with the coordinating Ser283 of the native *TeLipA2* (blue).

## 5.8 Conclusions

With the methodology for *TeLipA2* expression, purification and crystallization well established [182], a diverse array of experiments were carried out to increase the available information present in the protein's structure.

The testing of different reducing conditions in an attempt to obtain a crystal structure with no DTT bound in the active site was not successful. It did, however consolidate the knowledge that the two thiol functional groups provide more than simply reducing equivalents, to keep the [4Fe-4S] clusters from oxidizing and degrading. DTT may indeed coordinate free iron, directly competing with sulfide in solution to prevent iron-sulfide aggregation, and allow a more controlled release of iron for cluster reconstitution, while preventing the excessive formation of colloidal and protein aggregates.

Thanks to the efforts of Dr. M. Hiscox, two synthetic substrate analogues, shown to form a lipoyl product [634], were available for co-crystallization experiments. The low solubility of the oligopeptides restricted the concentration in the crystallization drop, leading to no crystal formation where substrate binding could be observed. The addition of MTAN, an enzyme known to use methylthioadenosine as substrate, into the crystallization conditions showed some selectivity between the two available peptides and, with further investigation, may provide the basis for a successful co-crystallization experiment. Very recently the group of Prof. Squire Booker has published an article revealing the structure of LipA during an intermediate step of sulfur insertion [666] which, together with spectroscopic data of a cross-linked substrate-LipA species containing three iron and three sulfide atoms [632] provides extraordinary evidence for the cluster degradation mechanism proposed [39]. The crystallographic data is yet to be available in PDB [667] thus an extensive discussion of the information contained was not possible.

By successfully crystallizing *TeLipA*2 Ser283Cys mutant, it was shown that Cys283 can coordinate the unique iron of the [4Fe-4S]$A_{ux}$ cluster, explaining the mutant's ability to form a mono-thiolated intermediate of the reaction [634]. As *TeLipA2* Ser283Cys's ability to form the dithiolated product is mostly absent, the re-arrangement of the [4Fe-4S]$_{Aux}$ to allow for a closer donating sulfur to the C8 carbon to the intermediate is likely to require the dissociation of the Ser-Fe bond, which is prevented by the mutation. The co-crystallization of this mutant with the substrate would provide a greater insight into the mechanism of action, as a trapped intermediate would be presumably easy to obtain.

# Chapter 6: Transcription of an RNA Fragment for Co-Crystallization Studies With 23S Methyltransferase Cfr

## 6.1 Introduction

As previously discussed (section 1.4) Cfr is an enzyme which catalyses the methylation of A2503, a nucleotide present in the 23S component of the prokaryotic large ribosomal subunit (50S) [457, 668]. Subsequent studies in the activity of Cfr showed that the segment necessary for 100% Cfr activity is limited to an 178 base oligonucleotide corresponding to the helix system H90-H92 and the adjacent single-stranded stretch that includes A2503 [83].

The current accepted mechanism (section 1.4.2) proposes vital roles for two conserved cysteines, Cys105 and Cys338. Mutation in any of the two residues leads to a loss of activity [464, 501] and the parallels with the related enzyme RlmN suggests Cys338 is likely to suffer methylation by a molecule of SAM, as seen in RlmN structure [496]; while Cys105 is responsible for releasing the methylated nucleic acid from its covalent bond with the protein. In that regard, mutagenesis of Cys105 should allow the creation of a stable covalent RNA:protein adduct, as was shown for RlmN [31], which could be an invaluable target for co-crystallization studies.

The crystallization of RNA and RNA complexes is notoriously harder than obtaining crystals of proteins alone [670]. In that regard, high quantities of RNA are required for crystallization screening and optimization, making the yield of commercial *in* vitro transcription kits often insufficient. Some studies have reported the *in vivo* expression RNA tagged with a streptavidin specific aptamer (S1) in order to purify RNA and RNA:protein complexes [671, 672]. Modifications on this purification mechanism revealed that the addition of RNA aptamers, small RNA fragments that can form specific secondary structures, can be used to protect RNA degradation *in vivo* [673], to allow the purification of large amounts of RNA.

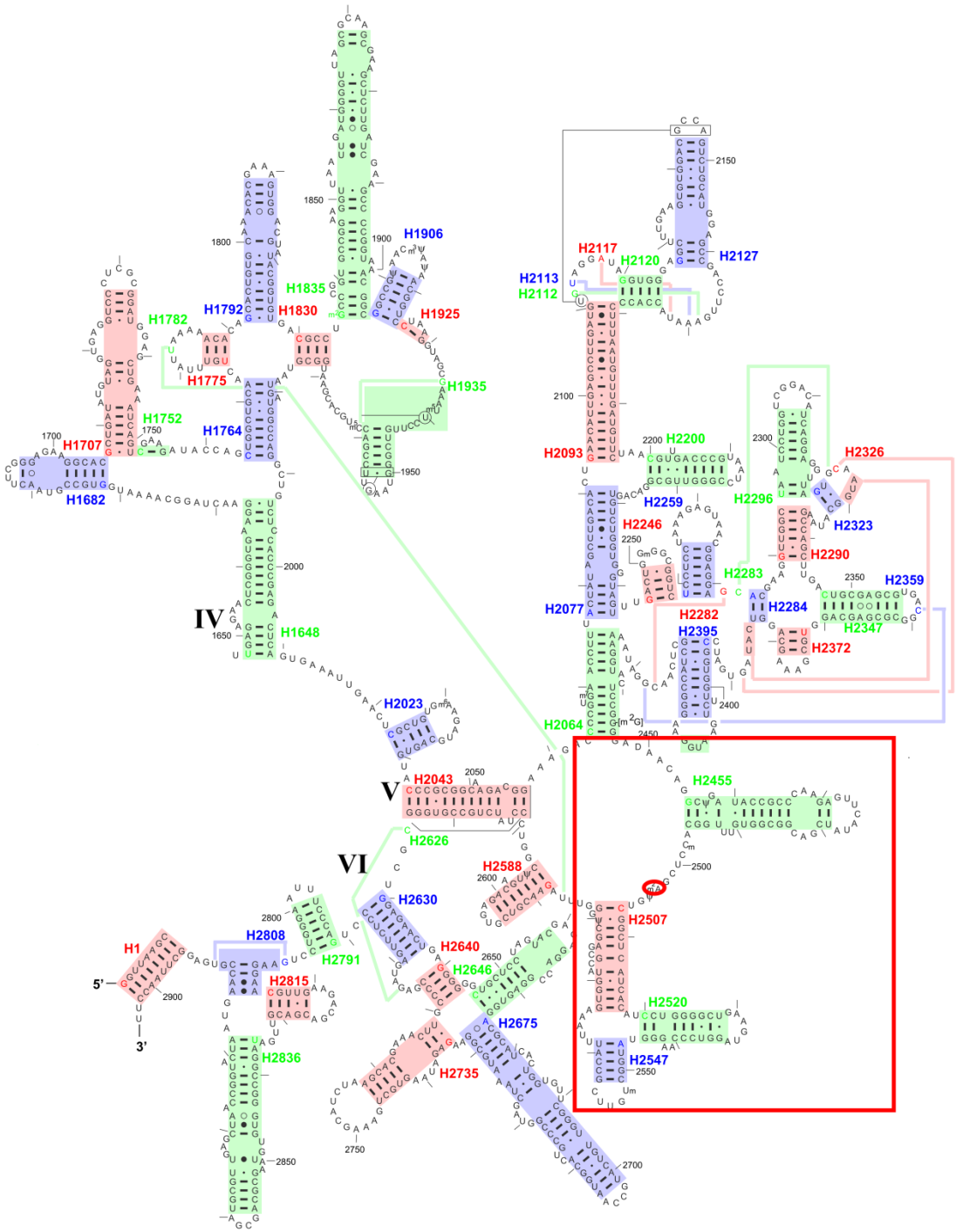



**Figure 139 23S rRNA secondary structure (3' half)**

Red circle corresponds to the position of A2503, where Cfr methylation is proposed to occur. Red square indicates the region necessary for 100% Cfr activity, as described by Yan and colleagues [83]. Figure obtained in the Comparative RNA Website and project (CRW) [669].

## 6.2 RNA Fragment *in vivo* Expression and Purification by Straptividin Affinity Chromatography

A gene was designed following the method of Ponchon and Dardel [674], surrounding the DNA coding the RNA fragment of interest by the genes of two self-cleaving ribozymes. These small RNA structures are able to catalyse the cleavage of the RNA sequence by an attack of a 2'-O of a ribose to the adjacent 3' phosphate [675]. In the case of RNA expression, the ribozymes (Hammerhead and HDV ribozymes) are able to protect the gene from nuclease activity and T7 transcription heterogeneity [674, 676].

The purification of the RNA occurs by the presence of an RNA streptavidin tag [671]. This tag binds to the Strep-Tactin column in a similar manner to the usual eight amino acid tag used in protein purification [677]. After purification, the addition of $Mg^{2+}$ would induce the self-cleavage process of the ribozymes releasing the final product, tag-free.

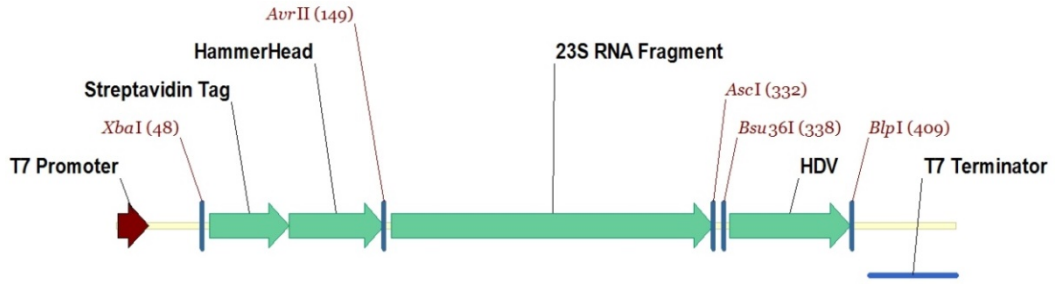



**Figure 140 Gene for the transcription of an RNA fragment.**

The gene coding the substrate fragment for Cfr (23S RNA Fragment) is surrounded by genes of two self-cleaving ribozymes, together with a Streptavidin tag. Several restriction sites are present for manipulation of the gene: XbaI/BlpI are used for introduction into the pET-16b vector; AvrII/AscI are present to change the gene of interest; AscI/Bsu36I can be used to add other tags, if necessary.

The gene, shown in Figure 140, was ordered from EuroFins MWG and placed in an IPTG inducible pET-16b plasmid (Appendix A.20). The plasmid was transformed into *E.coli* BL21 (DE3) STAR cells (from ThermoFisher®), a strain that contains a mutation in the *rne* gene, leading to the expression of a truncated RNAse E, which lacks the ability to degrade RNA [678].

Despite elaborate procedures having been developed for RNA isolation, working with RNA remains a practical challenge [679]. RNA is highly susceptible to degradation due to its chemical

structure, which becomes particularly problematic because of the ease of ribonuclease (RNAse) contamination from the environment [679].

A typical method for RNA isolation from cells, and used in this project, is based on a mono-phasic solution of chloroform:phenol and guanidine isothiocyanate [680, 681] (Method 9.22). The method takes advantage of the fact that, under acidic conditions, RNA is highly soluble in aqueous solution, while most of DNA and proteins are moved towards an organic chloroform:phenol mixture. Phenol, which is slightly miscible in water, helps to denature proteins that may remain in the aqueous solution, while chloroform, which has a significant higher density ($\rho$=1.49 g $cm^{-3}$) than water ($\rho$=0.99 g $cm^{-3}$), helps to form a clear phase separation. The addition of guanidine thiocyanate helps in the inactivation of many RNAses and hydroxyquinoline is an often used additive to prevent phenol oxidation, chelate metals which may over time cause the degradation of RNA and give a clear yellow colour to the organic phase [679], making the separation of both phases easier. The precipitation of RNA from the aqueous phase occurs by the addition of isopropanol and centrifugation. The total RNA obtained from a 1 L growth was approximately 1500 μg of total RNA, significantly lower than the reported 10-50 mg of recombinant RNA purified by Ponchon *et al.* [682].

1 mg of the extracted RNA was applied onto a Streptactin Sepharose HighPerformance column for affinity purification (Method 9.23). The samples were eluted with buffer containing $MgCl_2$ to cause the ribozymes to self-cleave and release the desired fragment [674, 676].

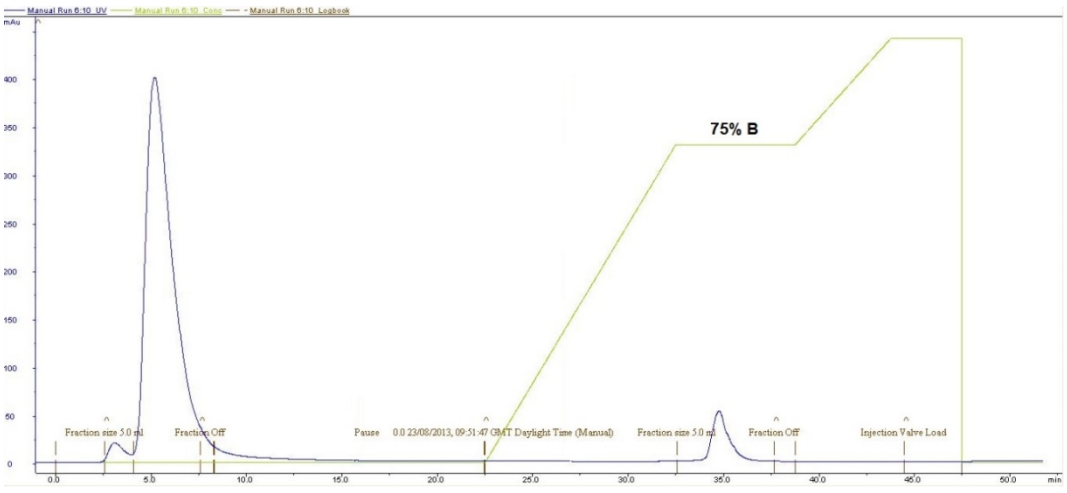



**Figure 141 Chromatographic run of a total RNA sample on a Streptactin column.**

An initial peak, corresponding to the flow through of RNA that does not effectively bind to the column can be observed after 5 minutes (1 mL/min). At 22.5 min a gradient was started to 75% buffer B (100 mM $MgCl_2$), which caused the elution of a single peak, presumably the target RNA. No further peak eluted at higher concentrations of $MgCl_2$.

The analysis of the fractions collected from the streptavidin affinity purification of RNA was carried out using a bleach agarose gel (Method 9.4.7.1) (**Figure 142**). The addition of 1% bleach to a high percentage agarose gel (2%) was reported to reduce RNA degradation, with little sacrifice to the resolution of the RNA bands [683].

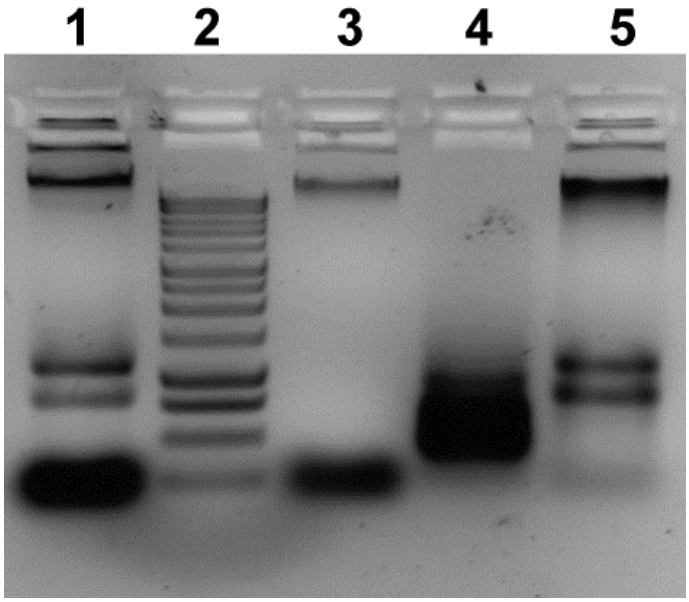



**Figure 142 Bleach Agarose RNA gel**

**1:** Total RNA before purification. **2:** DNA ladder (10,000 bp), used as control of the migration in the gel. **3:** Flow-through peak. **4:** RNA ladder (1,000 nt), with no clear separation visible. **5:** Purified RNA peak.

Some separation of the components loaded into the column can be observed, with the flow through possessing mostly a dark low molecular band, while the purified peak seems to show two higher bands present. All samples seem to also contain a band near the top of the gel, possibly corresponding to RNA that did not migrate correctly in the gel. Regrettably, the separation of the RNA ladder is far from clear, making it difficult to analyse if the purified sample contains the necessary fragment.

Due to the lack of suitable separation of the RNA ladder using the bleach agarose gel, a different separation method was attempted, by using a formaldehyde agarose gel (Method 9.4.7.2), first described in 1971 [684].

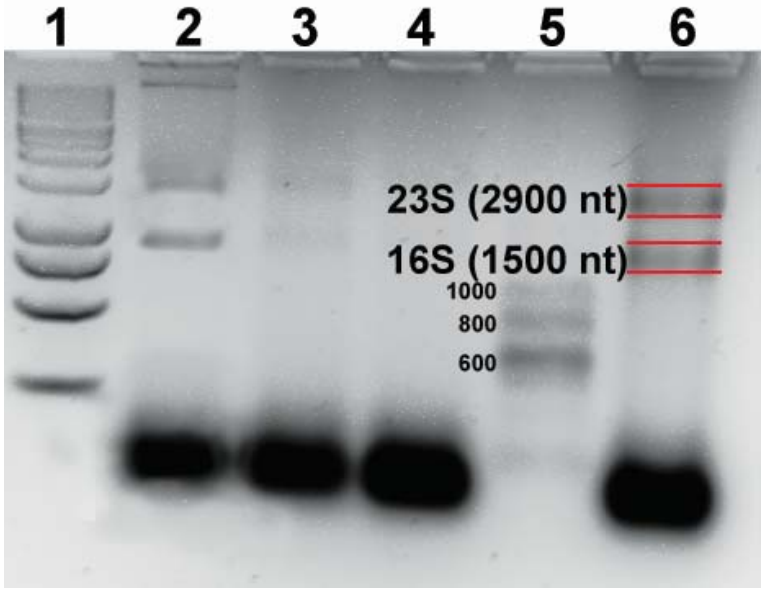



**Figure 143 Formaldehyde agarose RNA gel.**

**1:** DNA ladder (10,000 bp), used as control of the migration in the gel. **2:** Total RNA before purification. **3,4:** RNA samples from the flow-through peak. **5:** RNA ladder (1,000 nt) **6:** Purified RNA peak. The purified bands migrate above the 1000 nt band of the RNA ladder, a likely indication they do not correspond to the desired RNA fragment, but to the 16S and 23S ribosomal units from *E. coli.*

The use of formaldehyde agarose gel allows a better separation of the RNA ladder than the bleach agarose gel method. The purified RNA peak shows a significant band of small fragments of RNA, likely to correspond to RNA degraded by RNAses. Additionally, two bands above 1000 nt can be observed, which are likely to correspond to the 16S and 23S ribosomes of *E. coli*, which possess a size of 1500 nt and 2900 nt, respectively [685, 686]. There is, however, no bands corresponding to either the 178 nt 23S RNA fragment, nor to the 361 nt RNA fragment with ribozymes in any of the lanes. These results strongly indicate the *in vivo* transcription has not been successful, or if it does occur, only small amounts of the fragment are present.

## 6.3 RNA Purification by RP-HPLC

Over the last decade, IP-RP-HPLC (Ion Pair Reverse Phase High Performance Liquid Chromatography) has become a more commonplace strategy to analyse and purify nucleotides [687-689]. Despite most studies describing the use of a nonporous alkylated poly(styrene-divinylbenzene) as the stationary phase, an available system using a (non-polar) C4 column with butyl-silica as the stationary phase was used to test the separation of RNA. A small sample (10 μg) of total RNA, purified as discussed in section 6.2, was loaded onto a Vydac® 4.6 x250 mm, 5 μm C4 column (Method 9.24), using the pairing agents TEA and HFIP with a methanol gradient [690].

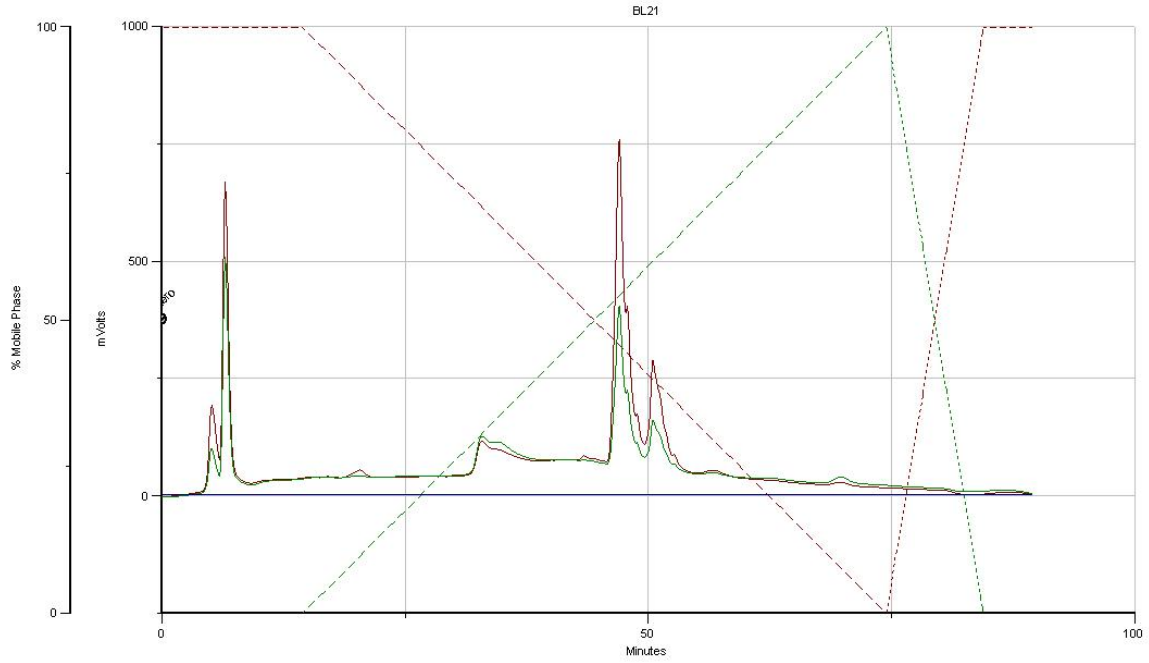



**Figure 144 Chromatographic profile obtained for the total RNA sample in a C4 reverse column.**

Prior to the beginning of the gradient two peaks can be seen eluting. As the gradient begins to rise, an initial broad peak can be observed (RT= 32.0 min) followed by two additional peaks (RT = 45.0 min and RT= 50.0 min). Red trace: UV at 260 nm and green trace: UV at 280 nm.

As can be observed in Figure 144, two main fragments are immediately eluted with 0% of organic phase, with at least 3 other peaks eluting at around 50% organic phase, indicating some separation of molecules within the sample is occurring.

## 6.4 RNA Analysis by Mass Spectrometry

Samples of streptactin purified RNA were analysed by Liquid chromatography–mass spectrometry (HPLC-MS) by J. Herniman, of the Mass Spectrometry Group of University of Southampton. The method used for separation is analogous to the one used for RP-HPLC reported in section 6.3.

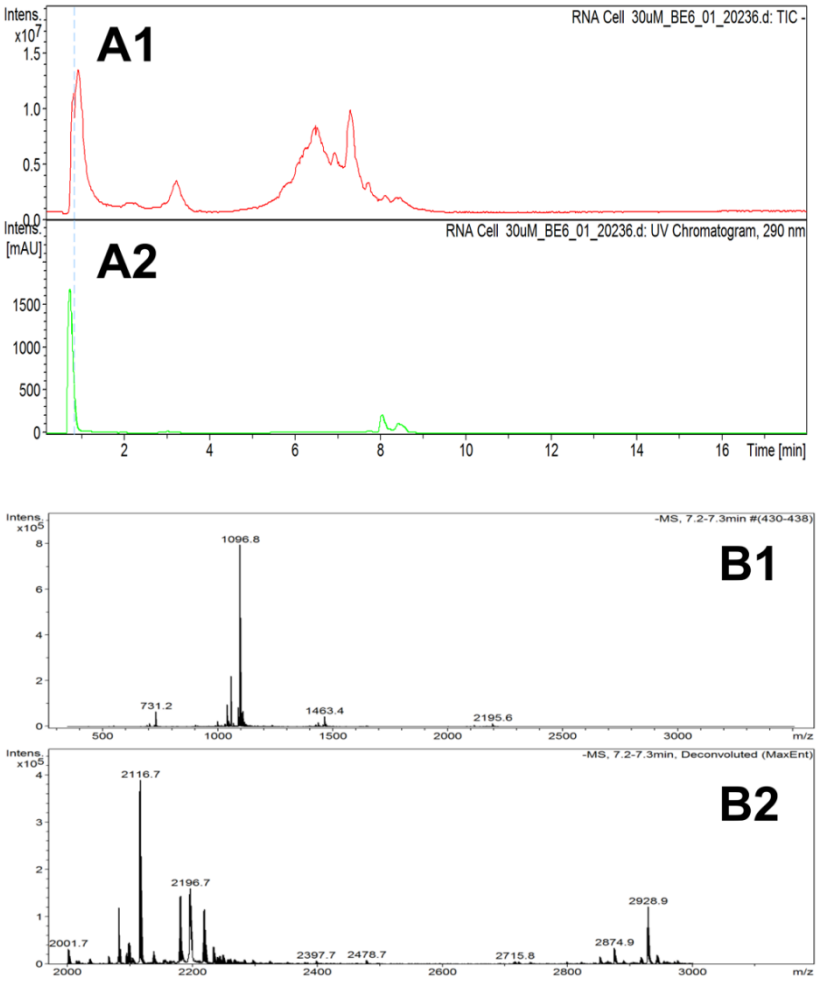



**Figure 145 Chromatographic profile and mass analysis obtained for the Streptactin purified RNA sample.**

The sample of streptactin purified RNA was separated by RP-HPLC prior to injection into the MS unit for analysis. **A1** and **A2** show the chromatogram obtained for the sample. Analysis of the mass/charge **B1**, and deconvolution of the signal for a more appropriate analysis of data **B2** shows only low mass/charge signals, indicative of nucleotide fragments no bigger than 10 bases [691]. It is possible that the samples suffered degradation, through contamination of RNAses, in an analogous result to that observed in the agarose gels, and further efforts need to be made in order to minimize this occurrence.

## 6.5 RNA purification Using a Commercial Kit

To guarantee that the lack of desired fragment during *in vivo* RNA purification (section 6.2) wasn't due to improper RNA purification, RNA was purified from a 10 mL culture using ThermoScientific's GeneJET RNA purification kit, following the manufacturer's instructions. A small yield of RNA was obtained in the end of the purification (108 ng/μL, 50 μL total volume) which was analysed in a formaldehyde agarose gel (Figure 146).

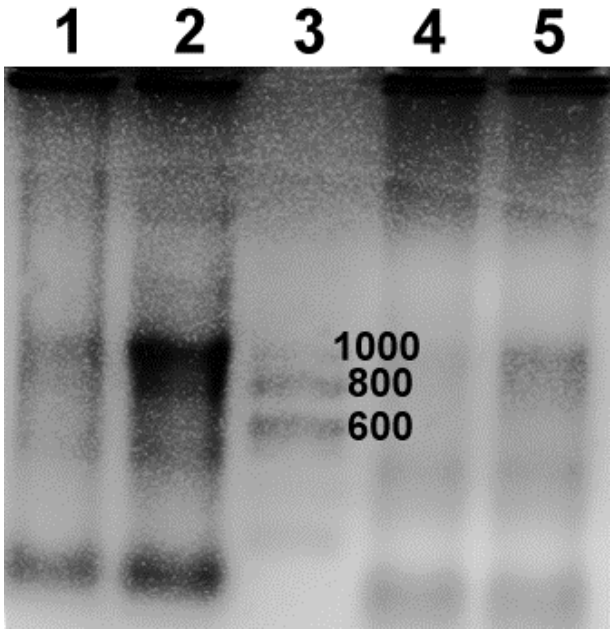



**Figure 146 Formaldehyde Agarose RNA gel for the samples obtained using a commercial RNA kit**

**1, 2**: RNA purified from *E. coli* BL21 (DE3) STAR cell line with an empty pET-16b inserted, used as a control; **3:** RNA ladder (1,000 nt); **4,5:** RNA purified from *E. coli* BL21 (DE3) STAR with pET-16b with the gene encoding the RNA fragment present.

The results do not differ significantly from the ones obtained with the guanidine thiocyanate (GNTC) purification method. Despite the presence of a broad and faint band below 600 bases in the samples with the synthetic gene, the same results can be visible in the control (pET-16b only), a strong indication that the desired fragment is not present in the samples analysed. These results also suggest the absence of the desired RNA fragment does not arise from degradation of the RNA through the action of exogenous RNases. It seems that this cell line is either not expressing the RNA in a suitable amount, or the RNA fragment is being endogenously degraded.

## 6.6 *Ss*Cfr Expression and Purification

The expression and purification of Cfr from *Staphylococcus sciuri* (Uniprot Q9FBG4) was achieved by using two plasmids: pDB182, is a pBAD-related plasmid (l-arabinose inducible), which codes for the *isc operon* and has an ampicillin resistance gene. The second, pMC115 [501], is a pCDFDuet-1 related plasmid, possessing the *cfr* gene together with a streptomycin resistance gene and is under control of the *lac operon* which is induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) (See Appendix A.21). Both plasmids were graciously provided by Dr. M Challand and Dr. Jim Spencer from the University of Bristol. The plasmids were sequentially inserted into *E.coli* BL21-Star (DE3) cells (Method 9.4.3) to use in routine protein expression.

The use of the protein expression methodology shown in section 3.2 was proven to be inefficient when expressing Cfr, as the growth yield for cfr producing cells in the fermenter was remarkably low (0.5 g of wet cell paste for a 5 L culture). The low yield raised the possibility that recombinant Cfr is toxic at high levels in the *E.coli* cell system, leading to a very anaemic cell growth [692].

Growing cells in auto-induction media is a methodology commonly used to address this issue [693], as auto-induction media has the particular property of possessing both glucose and lactose as their energy sources. The bacteria preferentially use glucose, which leads to a growth of the cell volume without induction of the gene regulated by the *lacI* gene. When all the glucose of the medium is used up the cells slowly adapt to the presence of lactose, which not only serves as an energy source, but one of its metabolites (allolactose) induces the *lac operon* (in an analogous manner to IPTG), leading to a more controlled expression of the protein [693].

The methodology for Cfr protein purification was therefore slightly modified (Method 9.5.4). The auto-induction media shows a very high yield of wet cell paste (~38 g/5 L culture), and much of the protein is found in the soluble fraction (Figure 147). Additionally, the use of separate induction Method for the ISC machinery and the relevant protein seems to decrease the yield of unnecessary proteins co-eluting with the RSAM enzyme, as the purity of *Ss*Cfr is higher (92% estimated by ImageJ) than both HydG and LipA.

Unfortunately the protein is also rather unstable in the presence of high imidazole (Buffer B), which has led to protein aggregation, and subsequent degradation, resulting in an overall low protein yield after purification. In an attempt to improve the stability of the protein after elution from the nickel affinity chromatography column, the samples were eluted from the nickel affinity column into a larger volume of buffer with no imidazole (5 mL elution in Buffer B into 10 mL of Buffer C). The stability of the protein increased dramatically, being able to stay in the solution with no observable amount of precipitation for over an hour. Regrettably, this method increased

the total volume collected from the nickel affinity chromatography threefold and, consequently, the necessary time for the concentration step increased significantly as well. This, in turn, led to eventual protein aggregation during the concentration step, and a similar yield of protein was obtained in the end. A decrease in glycerol concentration in Buffer C was likewise attempted, as the presence of a high concentration of glycerol leads to slower concentration times in the Amicon system, but the stability of the protein was severely affected and this change was proven unsuitable.

**Table 19 Protein estimation from *Ss*Cfr purification**

It shows the amount of cellular paste that was obtained in the protein expression as well as the values total protein content after each Bradford estimation.

| Protein | Initial Cell Pellet (g) | Total Protein Ni-NTA chromatography (mg) | Total Protein Size exclusion chromatography (mg) |
|---|---|---|---|
| *Ss*Cfr | 58 | 89 | 10 |

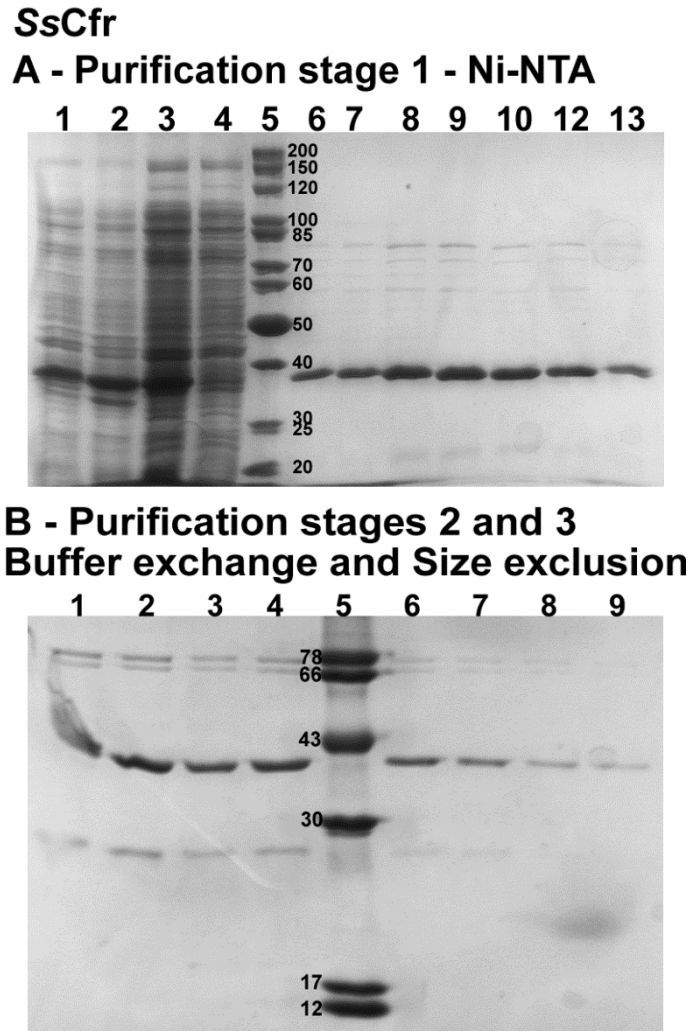



**Figure 147 15% SDS-PAGE analysis of the purification fractions of *Ss*Cfr**

**A:** Ni-NTA purification analysis: 1 – Cell lysate; 2 – Insoluble fraction; 3 – Clear lysate applied onto the column; 4 – Flow through; 5 –Protein molecular weight markers; 6-13 – Samples collected at 60% buffer B, corresponding to *Ss*Cfr (40 kDa). **B:** Buffer exchange and size exclusion purification analysis: 1-4 – Fractions collected from the small S-75 column used to buffer exchange into Buffer C; 5 – Protein molecular weight markers; 6 – Dark brown/black [Fe-S] excess and protein aggregate; 7-9 – Soluble *SsiCfr*, with a golden-brown colour

### 6.6.1 Chemical Reconstitution of *Ss*Cfr

The reconstituted sample of *Ss*Cfr (2 μM) shows a lower extinction coefficient at 410 nm ($\varepsilon_{410}$= 16,400 $cm^{-1}$ $M^{-1}$) when compared with *Ti*Hydg (section 2.3.3) and *Te*LipA 2 (section 5.2.1), which is to be expected as Cfr possesses solely one [4Fe-4S] cluster [476]. The instability of the protein is clearly visible in the UV-Vis spectra, as a very high background can be observed following reconstitution.

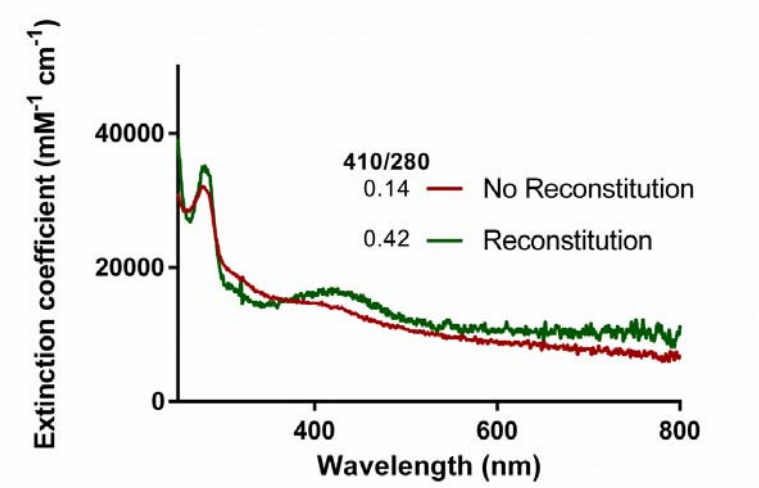



**Figure 148 UV-vis spectra of *Ss* Cfr**

The UV-vis spectrum of Cfr after reconstitution (green trace, [*Ss*Cfr]= 4.8 μM) shows a more prominent peak at 410 nm when compared with the sample before the reconstitution step (red trace , [*Ss*Cfr]= 2.0 μM), indicative of a successful reconstitution. The significant signal noise at higher wavelengths likely arises from iron-sulfide colloids and protein aggregate.

Furthermore, and unlike the other RSAM enzymes, the extinction coefficient at 410 nm of the sample prior to reconstitution (4.8 μM) is already close to the expected values for a $[4Fe\text{-}4S]^{2+}$ cluster ($\varepsilon_{410}$= 14,500 $cm^{-1}$ $M^{-1}$), possibly indicative that the two vector system (section 6.6) may allow for a better *in* vivo reconstitution.

## 6.7 Preliminary *Ss*Cfr Crystallization Studies

An initial screening for crystallization of *Ss*Cfr was carried out using Hampton PEGRx HT broad commercial screen in the presence of SAM (Method 9.10). Three different concentrations of protein were tested: 0.1 mM; 0.6 mM and 1 mM. The protein rapidly precipitates in the higher concentrations, but at 0.1 mM one condition was identified that leads to the formation of needle-like crystals (Figure 149).

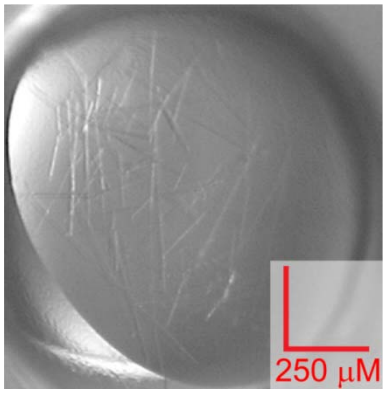



**Figure 149 Needle-like crystals obtained for *Ss*Cfr.**

The crystal like formations of *Ss*Cfr (0.1M) in the presence of SAM (3.0 mM) formed after 5-day incubation. Condition A2 from Hampton PEGRx commercial broad screen (0.1 M sodium citrate dihydrate pH 5.5, 38% PEG 200)

The crystals were collected into a cryo-loop and frozen directly in liquid nitrogen, being kept in it until the diffraction test at Diamond Light Source, beamline I24 (microfocus line). The collection of the sample was particularly challenging, as the needle-like crystals are hard to observe, very fragile and difficult to collect individual fragments. Additionally the direct freezing method was proven ineffective for these samples as the sample in the loop was completely dark (Figure 150 A). Advantage was taken of the grid scan function at DLS, where several locations in the loop are tested in quick succession to locate the protein crystal. The diffraction pattern shown in (Figure 150 B) was obtained after identifying the position of the crystal.

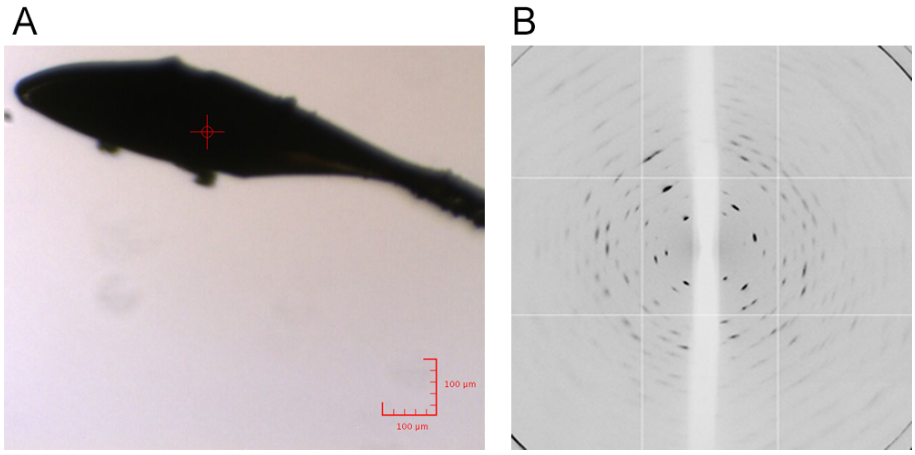



**Figure 150 Testing of *Ss*Cfr needle-like crystal formation.**

**A:** Loop bearing the *Ss*Cfr needle-like crystal. **B:** Crystal diffraction from a needle-like crystal from Cfr.

The resolution of the data is too low for any downstream processing to obtain any information regarding the protein's structure. Nonetheless the diffraction pattern is consistent with that of a macromolecule [694], and further optimization of both the conditions to obtain better crystals is necessary in order to solve the protein's structure.

## 6.8 Design of a co-Expression System for *Ss*Cfr and its rRNA Substrate

The low efficiency of RNA purification (section 6.2), as well as the instability of *Ss*Cfr when expressed in the absence of its substrate(section 6.6) could, potentially, both be circumvented by the co-expression of a protein:RNA complex. There have been various reports of ribonucleotide protein:RNA complexes co-purified for crystallization (see reviews [695-697].

To attempt to co-express *Ss*Cfr with its related rRNA substrate fragment, the gene designed to encode the RNA fragment (section 6.2, Figure 140) was inserted onto the available multiple cloning site (MCS1) of the pCDFDuet-1 vector, containing the cfr gene in MCS2.

As no compatible ends could be obtained for the RNA fragment encoding gene with the vector, primers were designed to mutate the XbaI/BlpI sites surrounding the 23S DNA into compatible NcoI/NdeI sites (Figure 151). The primers were used in primer extension PCR (Method 9.14) to amplify the target gene with the modified ends.


**Figure 151 Mutagenesis primers.**

Primers, designed to mutate XbaI and BlpI sites to NcoI and SacI sites, respectively. This last pair of restriction enzymes are compatible with the free multiple cloning site (MCS) in pCDFDuet-1 plasmid (MCS1).

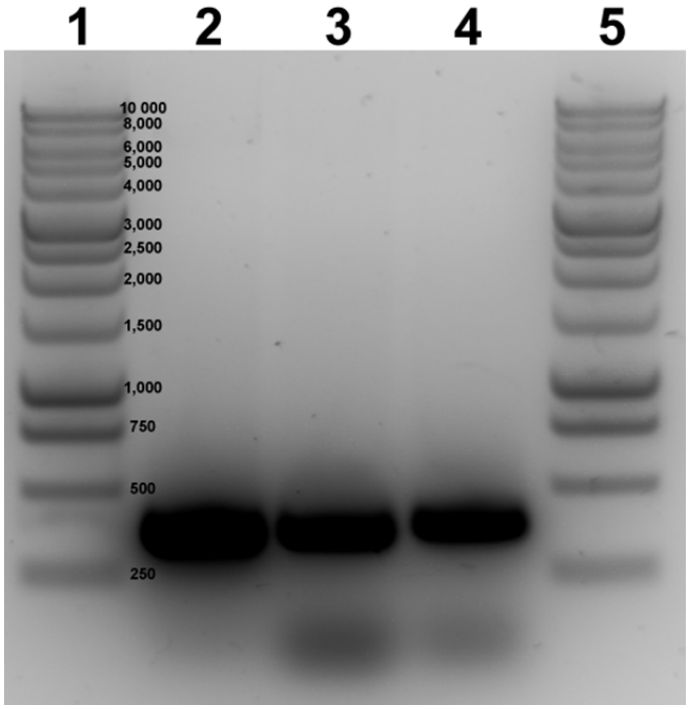



**Figure 152 Agarose analysis of the PCR amplification of the RNA fragment encoding gene.**

The amplification of the RNA encoding gene with the primers that introduce a mutation in each end, to obtain compatible restriction sites was successful. **1,5**: DNA molecular weight marker. **2-4:** Amplified DNA, with the expected size (372 bp).

Following amplification and purification of the relevant DNA fragment (Figure 152), the fragment and the vector were digested, ligated and transformed successfully (Method 9.4.6). The insertion

of the DNA fragment that encodes the rRNA fragment was carried out, not only onto a pCDFDuet-1 vector containing the *cfr* gene, but also onto a different pCDFDuet-1 vector which contained a mutated *cfr* gene encoding a Cys105Ala mutation on *Ss*Cfr. The new pCDFDuet-1 vector was also generously provided by our collaborators Dr. M. Challand and Dr. J. Spencer from the University of Bristol. This Cfr Cys105Ala mutation occurs in an amino acid proposed to be required for the mechanism of action [464], and a cysteine to alanine mutagenesis in the related enzyme in RlmN led to the formation of a stable covalent protein:RNA adduct as reported by McCusker *et al.* [698].

### 6.8.1 Confirmation of the Correct pCDFDuet-1 Constructs

In order to confirm whether the 23S DNA amplified fragment (Figure 152) had been successfully inserted to the MCS1 of the pCDFDuet-1 vector plasmids, a PCR amplification of the gene fragment was carried out using the same primers (section 6.8, Figure 151), but with the pCDFDuet-1 vector as template. Only if the fragment was successfully introduced into MCS1 of the pCDFDuet-1 vector should PCR amplified copies of the 372 bp DNA fragment be visible by agarose gel. As multiple colonies had been selected (Method 9.4.3), plasmid preparation of three colonies for both constructs was carried out (Method 9.4.4) to test by PCR amplification.

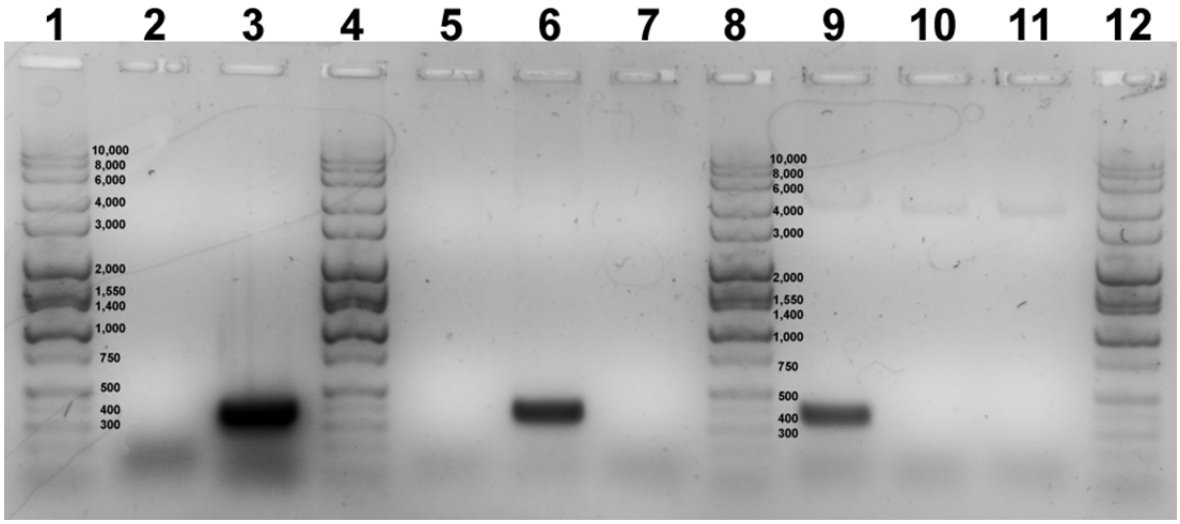



**Figure 153 Agarose gel analysis of the PCR amplification for the RNA encoding gene.**

**1, 4, 8, 12:** DNA molecular weight marker (10,000 bp). **2:** Negative control, where the PCR amplification was carried out in pCDFDuet-1, containing the *cfr wt* gene in MCS2 but an empty MCS1. **3**: Positive control, using the pET16b vector containing the RNA encoding gene. **5, 6, 7:** PCR amplifications of PCDFDuet-1 vector with *cfr wt*. **9, 10, 11**: PCR amplifications of PCDFDuet-1 vector with *cfr* Cys105Ala. Expected size of the amplification is of 372 bp.

The results indicate that of the 6 colonies, only two possessed the correct transformation, one from the vector containing the *cfr wt* gene and one from the vector containing the *cfr* Cys105Ala gene. In order to validate these results, vector DNA that showed positive PCR amplification for the RNA fragment encoding gene of interest were tested by restriction enzyme digestion (Method 9.4.5). The vectors were digested with either restriction enzymes flanking the RNA fragment encoding DNA in MCS1 (NcoI/SacI) or flanking the *cfr* gene in MCS2 (NdeI/XhoI) (see Appendix A.21 for the pCDFDuet-1 plasmid map).

Additionally, the *cfr* gene contains the specific sequence recognized by the restriction enzyme BamHI. BamHI is also one of the available sites within MCS-1, and a successful introduction of the 23S DNA fragment would remove the BamHI site, preventing a double digestion from occurring. This fortuitous presence of the BamHI sequence in both the *cfr* gene and the MCS1 allows an additional test for the construct. Following BamHI digestion, only one band should be present, corresponding to the size of the plasmid (5125 bp). However, if the ligation wasn't successful, the BamHI site in MCS1 is still present, and a double digestion is observed, with two bands expected, at 4068bp and 745 bp.

NcoI/SacI digestions should show a 4760 bp corresponding to the pCDFDuet-1 vector plasmid and a 365 bp band for the RNA encoding fragment inserted. NdeI/XhoI digestions serve as an internal control, as they excise the band corresponding to the Cfr protein, leading to a 1085 bp band in the agarose gel.

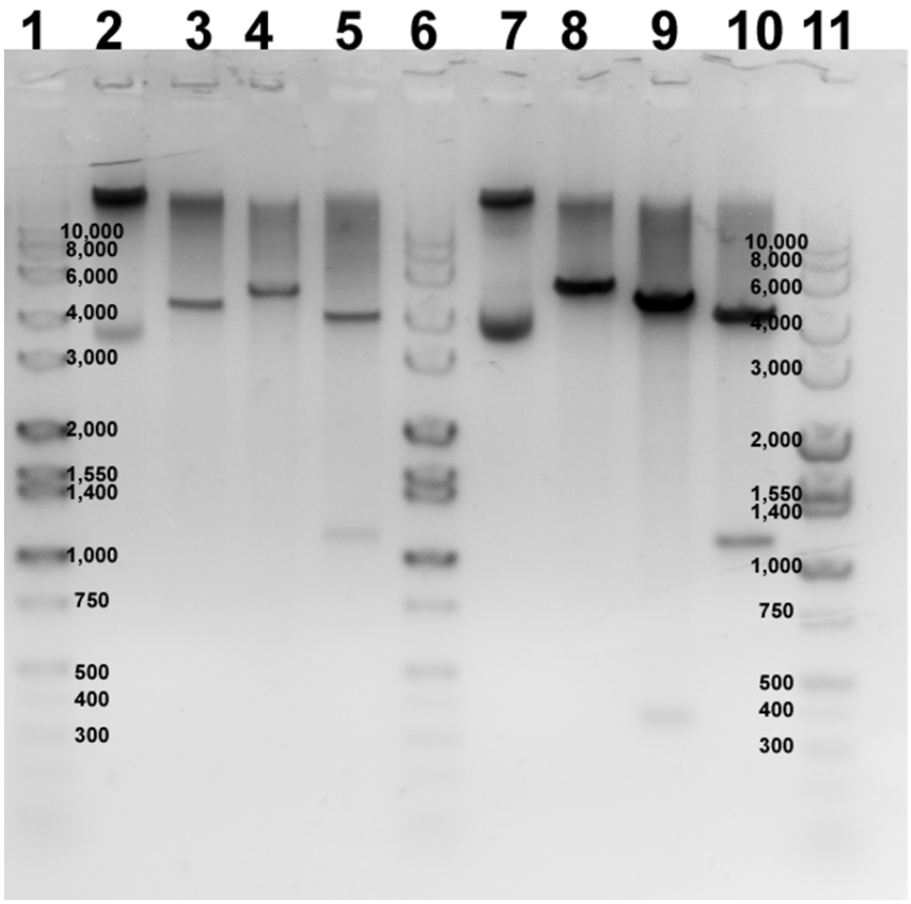



**Figure 154 Restriction enzyme digestion analysis by agarose gel.**

**1, 6, 11:** DNA molecular weight marker; **2-5:** pCDFDuet-1 plasmid containing the *cfr* wt gene: **2:** Non-digested control; **3:** BamHI digestion. The absence of a 745 bp fragment is indicative the 23S DNA fragment was successfully inserted in MCS-1; **4:** NcoI/XhoI double digestion. If the introduction of the 23S DNA fragment was successful, a band at 372 bp should be observed, which seems to be absent. **5:** NdeI/XhoI double digest. A band at 1085 bp is expected, corresponding to the *cfr* gene. **7-10:** pCDFDuet-1 plasmid containing the *cfr* Cys105Ala gene: **7:** Non-digested control; **8:** BamHI digestion. The absence of a 745 bp fragment is indicative the 23S DNA fragment was successfully inserted in MCS-1; **9:** NcoI/XhoI double digestion. The presence of the 372 bp is another indication of the successful assembly of the vector. **10:** NdeI/XhoI double digest. A band at 1085 bp is expected, corresponding to the *cfr* gene.

The results generally confirm the correct insertion of the RNA encoding fragment in both pCDFDuet-1 vectors. In both the wt *cfr* gene and in the mutant BamHI only gives rise to a single band of approximately 5000 bp. The digestion with NcoI/SacI, expected to yield the RNA encoding

fragment and the pCDFDuet-1 vector, seem to have failed in *cfr wt* samples. However with two of the controls testing positive for the insertion, the success of the construct was assumed.

### 6.8.2 Co-Expression of Cfr With its rRNA Substrate

With the successful creation of Duet constructs that can express simultaneously *Ss*Cfr (WT and Cys105Ala) with the RNA fragment, a co-expression study was carried out. Firstly the pCDFDuet-1 vectors, together with a pDB1282 vector were sequentially inserted into BL21(DE3) STAR cells (Method 9.4.3). The new cultures were used for protein expression and purification, as described in section 6.6 (Method 9.5.4).

The collected fractions for both Cfr *wt* and Cfr Cys105Ala were analysed by SDS-PAGE (Figure 155). Unfortunately, there does not seem to be any significant differences in the migration of the proteins putatively co-expressed with the rRNA fragment in comparison with Cfr *wt* expressed from the pCDFDuet-1 vector without the 23S rRNA fragment gene present (section 6.6). The lack of an apparent covalent adduct, as described by McCusker and colleagues [31], indicates that either the RNA fragment is not being expressed in sufficient amounts , or is being rapidly degraded in the cell.

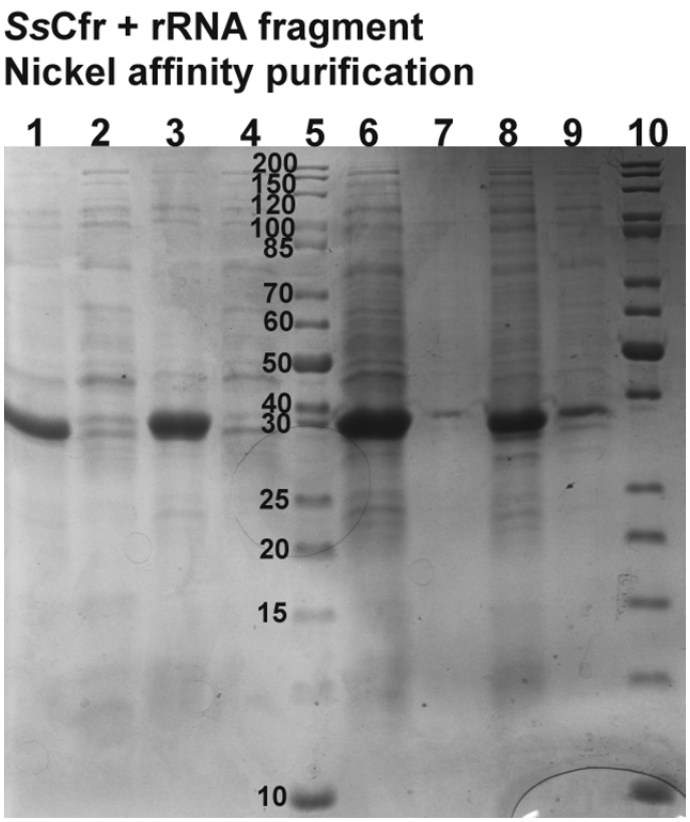



**Figure 155 15% SDS-PAGE gels of samples obtained in the purification of *Ss* Cfr co-expressed with 23S rRNA fragment.**

**1:** *Ss*Cfr WT expressed in the absence of the RNA fragment coding gene (empty MSC1). **2:** Insoluble fraction collected from the same expression as **1**. **3:** *Ss*Cfr WT co-expressed with the gene encoding an RNA fragment in MCS-1 of pCDFDuet-1. **4:** Insoluble fraction of the expression study **3**. **5,10:** Protein molecular weight marker. **6:** *Ss*Cfr Cys105Ala expressed in the absence of the RNA encoding gene. **7:** Insoluble fraction of the same expression study as **6**. **8**: *Ss*Cfr Cys105Ala co-expressed with the gene encoding an RNA fragment. **9**: Insoluble fraction of the same expression study as **8.** The lack of significant differences in the migration of the samples heavily implies there was no measurable formation of protein:RNA adduct.

## 6.9 Expression of an RNA Fragment Using *in vitro* Transcription

As the *in vivo* co-expression studies did not yield the expected results, which would facilitate the co- purification of the enzyme with the substrate bound, an attempt was made at using T7 RNA polymerase commercial kit to amplify the correct RNA fragment *in vitro*, to be subsequently be used as a substrate, as reported previously [83, 477, 500]. The protocol provided by the supplier was rigorously followed, with the digestion of the purified pET16-b vector, containing the gene of

interest, after the sequence to be transcribed. This will allow T7 RNA polymerase to extend the RNA fragment until there are is no further complementary DNA, and it will dissociate from the DNA, starting a new reaction.

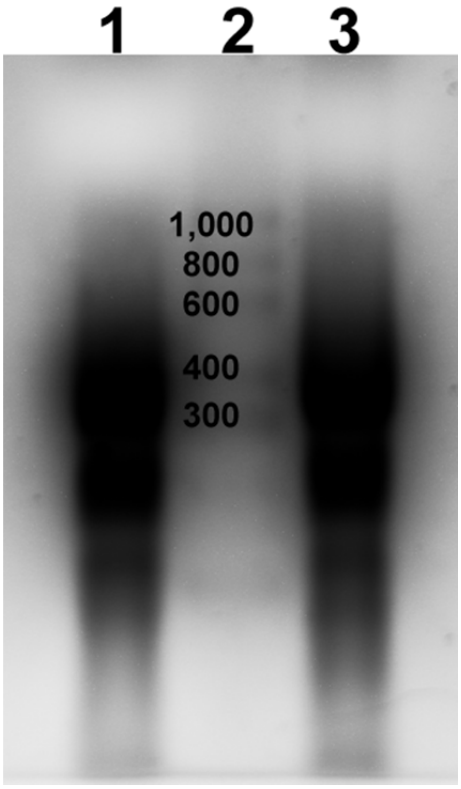



**Figure 156 *in vitro* transcription**

Formaldehyde agarose analysis of the RNA transcription results. **1, 3:** Transcription results. **2:** RNA molecular weight marker (1,000 nt).

The results, though preliminary, do show a very large band between 300 bases and 400 bases, which is consistent with the size of the DNA fragment amplified (382 bp). A second large band can be seen below 200 bp, which may correspond to the 23S fragment without the self-cleaving ribozyme ends (183 bases). The exact content of the reaction buffer provided by the kit is not divulged by Ambion, but should it contain magnesium or another divalent cation, as reported for *in vitro* transcription [699], it may explain how the fragment was released, by inducing the self-cleaving activity of the ribozyme ends [700]. These initial results show how powerful *in vitro* transcription can be in obtaining pure RNA, which will certainly be useful for co-crystallization studies of Cfr.

## 6.10 Conclusions

*In vivo* expression has been described as a valuable alternative to produce high amounts of RNA [674], a goal which would be attractive for co-crystallization of the RNA:Cfr complex. Despite the different methodologies used for total RNA purification (phenol:chloroform extraction and commercial purification kits) and analysis (affinity purification with streptavidin; LC-MS), there was no unambiguous expression of the relevant RNA fragment, despite using protective scaffolds to ensure the decrease of RNA degradation [673].

An *in vivo* co-expression of the RNA:protein complex was also attempted, taking advantage of a mutant (Cfr Cys105Ala) which is predicted to form a stable, covalent, intermediate of the reaction, but with little success. A major concern regarding the co-purification of RNA-protein complexes is their breakdown, which can easily occur due to changes in ionic strength, solute concentration and subcellular localization [701]. UV radiation has been shown to foment covalent bonds between protein and nucleic acids [701], and several developments have been devised to use UV-crosslinking, in an attempt to stabilize the complexes prior to cell lysis [702, 703]. It is possible that the use of this methodologies would tip the scale in the purification of the Cfr:RNA complex, and could be attempted in the near future.

With the *in vivo* RNA expression showing less than ideal yields, the use of an *in vitro* transcription kit was attempted. Despite only preliminary results, bands that may correspond to the desired transcription product could be observed. The possibility of expressing T7 RNA polymerase could allow the accumulation of RNA products *in vitro*, with a more controlled risk of RNA degradation.

Cfr purification required changes into the protocol, as the expression of the protein was shown to be toxic to the host organism. The use of auto-induction media was proven to be an effective measure against the toxicity and, despite a relatively low yield, the protein was purified and crystallized successfully in the presence of SAM. The crystals, needle-like structures, were very fragile had had a poor diffraction pattern, but shine a ray of hope for a successful co-crystallization with RNA, provided the substrate can be expressed in sufficient levels.

# Chapter 7: Conclusions and future work

## 7.1 Hydrogenase Maturase HydG

Within the last five years the understanding of the mechanism of HydG has deepened considerably. It is now generally accepted HydG is responsible for transferring the iron [296], carbon monoxide [293] and cyanide [47] into the scaffold enzyme HydF [287]. Additionally previous theoretical models [310] together with a crystallographic structure of a $\Delta$-CTD HydG missing the auxiliary cluster [317] revealed a full TIM barrel formation and three distinct regions: the conserved RSAM active site, an internal tunnel formed by the TIM barrel and a distinct region where the auxiliary cluster is present.

The identification of a $Fe(CO)_2CN$ entity in the $[4Fe\text{-}4S]_{Aux}$ cluster that could be correlated with enzymatic function [296], was indicative of the way by which the transfer of products onto the maturation partner(s) could take place. At the time, however, criticism was raised over this proposal, as the coordination of the unique iron of a [4Fe-4S] cluster with both amino and carboxylate groups of tyrosine as well as with three diatomic ligands had no precedence in literature [704, 705].This apparent implausibility was answered by the structural data of HydG reported in chapter 3.

### 7.1.1 Structure of *Ti*HydG

A fully reconstituted and active HydG from *Thermoanaerobacter italicus* (10.8 ± 2.2 irons/protein molecule and DOA/*p*-cresol/cyanide $k_{cat}^{app} \simeq$ 7.0/6.5/4.0 $x10^{-4}$ $s^{-1}$, respectively) was successfully crystallized and the structure solved. A crystallographic dimer was observed in the structure and certain key differences in each monomer gave a fortuitous further insight into the mechanism of HydG.

One of the monomers (labelled monomer A) showed a coordinating cysteine bound to the $[4Fe\text{-}4S]_{RS}$ cluster without the adenosyl moiety, representative of SAM degradation. Monomer B instead revealed an intact SAM being coordinated. The accompanying structural changes observed when comparing both models allowed the proposal of the disordered loop 345-360 as the likely entrance for the substrate L-tyrosine. Careful analysis of the anomalous data collected near the iron $\kappa$-edge of the *Ti*HydG crystal showed an unmistakable fifth iron signal near the

auxiliary cluster of monomer A of the enzyme. As the structure was manually updated and refined, a final model was put forward.

It shows this fifth iron, termed 'labile iron' due to both its low occupancy as well as its proposed role as carrier of diatomic molecules, bound to the $[4Fe\text{-}4]_{Aux}$ cluster by a bridging μ2 sulfide. The iron in the structure was coordinated in a apparent octahedral conformation to the sulfide bridge, a single residue of the protein, His265, two water molecules and a single non-proteinacious amino acid bound in a bidentate manner through the α-amino and α-carboxyl groups. The nature of this unprecedented cluster was confirmed by Dr. Daniel Suess and Prof. David Britt, which obtained EPR data consistent with a S=5/2 iron signal present in the auxiliary cluster. It is this labile iron that is now proposed to be responsible for the $Fe(CO)_2CN$ synthon observed [296].

When analysing the iron anomalous data on the auxiliary cluster of monomer B, the absence of the labile iron was apparent, though the bridging sulfide was still in place. Once more the asymmetries between both monomers were proven advantageous. Firstly, this structure credence to the possibility of labile iron transfer, as the structural elements necessary to bind the iron are still in place. Secondly, by directly comparing both structures, it was observed an increase in the B-factor of a specific α-helix (405-412) in monomer B. This helix is positioned as a gateway between the auxiliary cluster and the solvent, leading to the proposal that it will play a role in the interaction with a partner enzyme, likely to be the scaffold maturase HydF.

The collected structural information for *Ti*HydG did not show the substrate tyrosine bound, even when co-crystallized with increasing amounts of the compound. Different strategies were employed to circumvent the low solubility of L-tyrosine at near biological pH (section 3.5), though the results remained unchanged. In the absence of a co-crystal with the substrate, L-tyrosine was modelled into the active site, using the information for the binding of tryptophan in the related amino acid lyase NosL.

### 7.1.2 Future Work in HydG Crystallography

As the extensive efforts in co-crystallization of *Ti*HydG with L-tyrosine (section 3.5) and synthetic analogues (section 5.X) did not yield a relevant structure with ligand density present, one must re-evaluate the crystallization process. In particular, the purification of *Tl*HydG showed similar yields to that of *Ti*HydG (22 mg protein/L culture for *Tl*HydG and 32 mg protein/L culture for *Ti*HydG), and crystallization efforts carried out by Dr. Rebecca Driesener showed the formation of crystalline features which diffracted to medium resolution (2.7 Å), in the absence of either SAM or

tyrosine. Optimization of the conditions to allow crystallization in the presence of the substrates may give rise to a different crystal conformation which will allow tyrosine binding.

A more ambitious project would be the crystallization of a HydG:HydF complex. HydF was shown to crystallize aerobically [706], which limited the information in regards to the nature of the preassembled cluster. The authors described the reconstituted form of the enzyme to be more unstable, possibly due to the lack of a coordinating partner. The co-crystallization of HydG with HydF, in different reconstitution stages, would provide an incredible understanding on the interactivity of the different maturases and perhaps create a closed system where the diatomic products could be observed binding to either protein. Preliminary co-purification studies have already been reported by Vallese *et al.* allowing them to conclude HydF interacts with both HydG and HydE [707], and could be used as models for co-purification.

### 7.1.3 Mutagenesis Studies of *Ti*HydG His265

As a means to address the pivotal role of the labile iron in the newly proposed mechanism of action, the site directed mutagenesis of the iron coordinating His265 was carried out (Chapter 4). Two changes were introduced: His to Ala, to abolish the coordination mechanism of the labile iron and His to Glu, as the most likely mutation to maintain the labile iron coordination.

Spectroscopic analysis revealed how different the cluster content was between the two mutants. *Ti*HydG His265Glu was shown to fully reconstitute both [4Fe-4S] clusters by UV-spectroscopy ($\varepsilon_{410}$= 34,000 $cm^{-1}$ $M^{-1}$) and EPR analysis, showing similar results to that of the wild type enzyme.

Introduction of an alanine in the position of His265, on the other hand, had a dramatic impact on the protein. The protein had a tendency to aggregate in solution and chemical reconstitution was only successfully accomplished by halving the Fe-S content added to the sample. As a consequence, the UV-spectra for *Ti*HydGHis265Ala was indicative of an incomplete reconstitution ($\varepsilon_{410}$= 22,300 $cm^{-1}$ $M^{-1}$). Additionally, EPR analysis by our colleagues at UC Davis (Dr. Daniel Suess and Prof. David Britt) revealed an EPR silent sample. This surprising result is hard to explain, with either the clusters being absent, possibly due to sample degradation, or by being found in an EPR silent state: $[2Fe\text{-}2S]^{2+}$, $[3Fe\text{-}4S]^{+}$ and $[4Fe\text{-}4S]^{2+}$ are all diamagnetic [708]. However, these states should not be present in the presence of strong reductants, such as dithionite and so far no explanation can be put forward for these results.

Activity assays carried out with both mutants and the wild type HydG as control consolidated the asymmetry between His265Ala and the other two enzymes, showing virtually no turnover for any of the substrates. Unlike the inactive alanine mutant, His265Glu was observed to efficiently form

cyanide ($K_{cat}^{app}$= 7.1 x $10^{-4}$ His 265Glu vs $K_{cat}^{app}$ =4.8 x $10^{-4}$ WT). The activity differences observed in His265Glu when compared with the wild type are likely to arise from the stronger coordination of carboxylate to the labile iron.

To verify the nature of the coordination of the labile iron in the mutants, a crystallization experiment was carried out. However, the instability of *Ti*HydG His265Ala limited the ability to obtain any crystal formation with this enzyme. The crystal structure of *Ti*HydG His265Glu was tested and its structure analysed, showing a high degree of similarity between the wild type and the His265Glu mutant, with the labile iron being coordinated by the carboxylate moiety of glutamic acid, but no amino acid residue is seen coordinating the iron. Instead four water molecules are clearly visible, and other ligands may be present, though the electron density analysis did not allow for an unambiguous determination of their nature. These results confirmed the correct placement of the labile iron in *Ti*HydG His265Glu.

In order to probe the reason behind such dramatic results across all studies with *Ti*HydG His265Ala, the secondary structure content was estimated by CD before and after reconstitution for all three HydG enzymes (wild type and mutants). Both wild type and *Ti*HydG His265Glu show similar predicted ratios of $\alpha$-helix (40% and 37%), $\beta$-strand (10% and 9%) and random coil (50% and 55%); values close to the crystallographic structure obtained for each (50% $\alpha$-helix, 9% $\beta$-strand, 41% random coil). The calculated values for the alanine mutant (45% $\alpha$-helix, 23% $\beta$-strand, 31% random coil) are indicative of a different conformational fold taken by the enzyme, with a significant increase in $\beta$-strands. It is possible the change in C-terminal fold leads to an increased aggregation of the protein, explaining the instability verified.

### 7.1.4 Future Work in *Ti*HydG Mutagenesis

The structural differences hinted at in *Ti*HydG His265Ala could be further verified by analysing the CD spectra of HydG mutants unable to coordinate the auxiliary cluster (Cys to Ala/Ser), as a means of validating the preliminary results discussed in section 4.6.2. It is to be expected that the removal of the coordination of the auxiliary cluster would have a similar effect on the overall fold of the protein.

An important question regarding the activity of *Ti*HydG His265Glu is how efficiently would the transfer of the $Fe(CO)_2CN$ synthon be with this enzyme. As the coordination of the labile iron is proposed to be stronger with glutamic acid in comparison with histidine, due to the ionic nature of the interaction, one could predict a more inefficient transfer of the H-cluster precursors. *In*

*vitro* maturation of [Fe-Fe] hydrogenase is a well-established method [282], which would provide a better understanding on the importance of the coordinating histidine residue.

### 7.1.5 Activity Studies With Substrate Analogues

Due to the lack of structural information regarding tyrosine binding in HydG, the use of synthetic analogues was probed in an attempt to circumvent the issue. The list of potential tyrosine analogues is rather extensive, with a good range of compound having been tested previously [709, 710]. Previous studies with HydG had failed to identify any substrate analogue which would suffer $C_\alpha$-$C_\beta$ cleavage, but the activity assays carried out with *Ti*HydG (chapter 4) revealed key compounds that interact with the enzyme.

The reaction with L-DOPA, despite increasing the amount of uncoupled turnover of SAM, was shown to release significant amounts of methylcatechol and cyanide ($k_{cat}^{app}$= 6.2 x $10^{-4}$ $s^{-1}$ and 5.4 x $10^{-4}$ $s^{-1}$, respectively). Tyrosinamide was another compound shown to be reductively cleaved by HydG (*p*-cresol $k_{cat}^{app}$= 7.0 x $10^{-4}$ $s^{-1}$), though as expected no cyanide was observed. Additionally, 4-HPHPA and tyramine showed an increase in SAM cleavage (SAM $k_{cat}^{app}$= 5.0 x $10^{-4}$ $s^{-1}$ and 4.3 x $10^{-4}$ $s^{-1}$, respectively, but little to no *p*-cresol formed. Unlike all other compounds tested, the cyclic THIQCA was shown to possibly decrease the rate of SAM turnover, potentially indicating a binding mode that interfered with SAM binding. These five compounds were used in co-crystallization experiments, but as their solubility levels were not dramatically different from the native substrate, the concentration of ligands in the crystallization mix was limited. Only crystallization in presence of tyramine and THIQCA showed the formation of crystals, but the structural analysis did not show any substrate analogue in the active site.

Influenced by the report of Krumholz [711] for the formation of $\alpha$-iminocarboxylic acids coordinating iron, the protocol described was modified to use ammonia as the amine, to possibly lead to the formation of Fe-bound DHG. Preliminary results seem indicative of the successful formation of the complex, followed by UV-Vis; and an activity assay where the complex was incubated with HydG in the absence of SAM or L-tyrosine, showed a greater than average cyanide formation, bypassing the first half of the mechanism of HydG.

### 7.1.6 Future Work in *Ti*HydG Mutagenesis

$C_\alpha$-$C_\beta$ cleavage of L-tyrosinamide and possibly tyramine should be confirmed by the analysis of the presence of predicted products (2-oxoacetamide and formaldehyde), by coupled assay reactions

or mass-spectrometry and/or NMR. Note that these predicted products assume there is no coordination to the labile iron, which in the case of L-tyrosinamide cleavage product dehydroglycinamide, may be an incorrect assumption.

The ability of the labile iron to coordinate different compounds should also be explored, to gain further insights into the mechanism of DHG cleavage. Glycine analogues could be used to test their reactivity in the auxiliary cluster active site, in a manner similar to the tyrosine analogues tested in Chapter 5, to better understand, together with the available structure of the auxiliary cluster active site, what drives the degradation of the glycine derivative into the diatomic components. A potential method for testing glycine analogues would be the formation of iron-bound $\alpha$-iminocarboxylic acids using different amines. The preliminary results show some promising developments, but there are significant follow-up studies necessary.

Firstly, confirmation of the formation of DHG by mass spectrometry and NMR would provide direct evidence the complex has been successfully formed. Additionally, treatment of HydG with EDTA, in low enough concentration so that both clusters are intact but the labile iron is removed, could increase the cyanide formation rate in the absence of substrate tyrosine, from the initial low values reported (cyanide $k_{cat}^{app}$= 0.3 x $10^{-4}$ $s^{-1}$, approximately). Developing a reliable system for iron:glycine complex analogue incorporation will undoubtedly create many opportunities to study the less understood half of the mechanism, the formation of diatomic ligands.

## 7.2 Lipoyl Synthase LipA

The work carried out by the groups of Prof. P. L. Roach [400, 523, 543, 642] and Prof. S. J. Booker [391, 401, 414-416, 712] have provided overwhelming evidence that LipA is both the catalyst of sulfuration as well as the substrate of the reaction, by donating a sulfur atom from its auxiliary cluster with concomitant degradation.

The structural data of LipA (PDBID: 4U0P/4U0O) [543] allowed the modelling of the potential active site location, located between both clusters, within a close distance to one of the sulfur atoms of the $[4Fe\text{-}4S]_{Aux}$ cluster. This proposed binding mode has also been confirmed by hyperfine coupling of the paramagnetic signal of a substrate analogue (2,4-hexadeinoyl-H pro) radical and the $^{57}Fe$ signal from the auxiliary cluster, indicating that both elements are within bonding distance [712]. However, the lack of a co-crystal with the substrate bound limits the conclusions that can be inferred for the mechanism of the reaction.

### 7.2.1 Studies on *Te*LipA2

To remedy the lack of an available co-crystal with the octanoyl unit bound to the active site of *Te*LipA2, several co-crystallization experiments were carried (chapter 7) out with two octanoylated synthetic peptide analogues, previously shown by Dr. M Hiscox to lead to the formation of the lipoyl form of the peptides (Hiscox, thesis). The low solubility of the compounds limited the concentration range that was used in the studies, and no structure with a substrate present was obtained (section 7.5).

As the structural information within these data sets consistently revealed the coordination of DTT to the $[4Fe\text{-}4S]_{RS}$ as well as the presence of the SAM degradation product 5'-MTA, an effort was made to minimize the presence of each within the crystallization condition, in the hopes that a new crystal form could be obtained where the substrate would be bound. Reconstitution of the enzyme in the presence of lower DTT, or in the absence of DTT but with different reducing agents, revealed the indispensability of the thiol in a successfully reconstitution (section 7.3). The stability of lipoyl synthase in the absence of a fully formed auxiliary cluster is greatly diminished and no crystals grew under these conditions. Likewise the removal of 5'-MTA, using MTAN, did not produce the desired effect, as no structure with an intact SAM was obtained.

The crystal structure of LipA (PDBID: 4U0P/4U0O) also revealed a unique coordination of the $[4Fe\text{-}4S]_{Aux}$ cluster by a serine (Ser283 in the structure) and kinetic analysis shows target-site mutagenesis in that residue leads to either cysteine or alanine lead abolishment of the natural reaction cycle [543]. These results led the authors to propose a unique role for Ser283: following the formation of the monothiolated intermediate, the serine would be released from binding the $[4Fe\text{-}3S]_{Aux}$ cluster, possibly by protonation of the hydroxyl side chain, to allow the re-arrangement of the cluster for the second abstraction to occur.

The structure reported in this work (section 7.6) shows the cysteine mutation is still able to coordinate the fourth iron in the $[4Fe\text{-}3S]_{Aux}$ cluster, leading to the proposal that the mutation prevents the release and subsequent rearrangement of the auxiliary cluster, justifying the lack of lipoylation activity observed in the *Te*LipA Ser283Cys mutant. Conversely, if no coordination is present to the fourth iron of the auxiliary cluster, by Ser283Ala mutation, the reaction was also shown to be abortive. The crystallization of this mutant was not successful, however, its high instability could be indicative of cluster disassembly in the absence of the substrate, revealing how extraordinary the control exerted by Ser283 is in the mechanism of reaction.

### 7.2.2 Future Work in the Mechanism of LipA

As it currently stands, the first insertion of a sulfur atom in the octanoyl substrate seems to be fairly understood: spectroscopic data is able to observe the cross-linked intermediate [713] and the structure based model of octanoic acid in the active site reveals an initial target sulfur of the $[4Fe\text{-}4S]_{Aux}$ for the donation to the substrate [543]. The focus should now turn to understand the second sulfur insertion, in particular what changes occur in the cluster for a second sulfur to be in position to be transferred to the cluster-bound intermediate.

An X-ray structure of *Te*LipA2 Ser283Cys with the intermediate bound could provide the structural insight into the complex prior to the second sulfur insertion. Complementary information could be obtained using co-crystals of LipA with the substrate, which depending on the tight control of the stoichiometry of the reaction, could reveal the structure of the enzyme with either the intermediate or the product bound. Within these studies, crystallization with a selenium-reconstituted sample (section 7.4) could be invaluable to more accurately identify the position of the selenium atoms by anomalous data.

An important consideration to bear in mind in future crystallization studies of LipA is the ability to successfully express the small H protein (10 kDa) and synthesize the octanoyl form of the protein [414], to use as a co-crystallization partner. The protein:protein complex, unlike a small synthetic analogue, can prove much more stable and will successfully incorporate the octanoic group in the correct position of the active site.

## 7.3 Methyltransferase Cfr

With the danger of widespread multi-resistance bacteria being present, understanding the mechanisms of resistance and how to counteract them has become the focus of intense studies. Within the last decade a significant body of work has led to a better understanding of the mechanism of action of Cfr. However, the lack of an available structure, in particularly one with the RNA substrate, limits the ability to accurately predict how a potential inhibitor would interact [714].

### 7.3.1 Studies on *Ss*Cfr and RNA Transcription

The expression and purification of soluble *Ss*Cfr was generally successful, despite the enzyme being more unstable than the other wild type enzymes in this project. Crystal-like formations in

the presence of SAM were achieved, but further optimization is required in the future to obtain relevant structural data.

Obtaining significant amounts of RNA proved more challenging. In an effort to produce milligram amounts of RNA for co-crystallization studies, *in vivo* transcription using *E.coli* followed by total RNA purification was attempted. A gene was designed encoding an RNA fragment previously shown to lead to measurable Cfr activity [715]. To reduce the endogenous degradation of the RNA by RNAses the gene was flanked by two protective ribozymes, and a gene encoding a streptavidin RNA tag was also inserted to aid in the purification. Unfortunately, the levels of transcription of the desired RNA never reached a level that could be quantified, even when the RNA was transcribed *in vivo* with concomitant expression of Cfr.

### 7.3.2 Future Work in Cfr

The results with *in vitro* transcription of the desired RNA fragment showed some promising results, but they require validation to confirm the fragment necessary for Cfr activity is being produced. Furthermore, the reported yields of RNA per kit reaction (3-5 mg) effectively mean less than 50 nmol of RNA produced each time, when the necessary quantities for co-crystallization should be in the millimolar range. A solution could be achieved by the expression and purification of T7 RNA polymerase. The protocol for T7 RNA polymerase protein expression [716] and effective transcription of target genes [717] are available and could be explored as a means to produce large amounts of RNA that would not be subject to endogenous degradation.

If substantial amounts of pure, stable RNA are transcribed, its incubation with Cfr may increase the stability of the enzyme, leading to better yields and potentially higher chance of success in co-crystallization.

# Chapter 8: Experimental section

## 8.1 Principles of Protein Crystallization

The main focus of this project lies in obtaining structural information on different RSAM enzymes, and how to apply that knowledge to deepen the understanding of what guides the mechanism of action within each enzymatic process. X-ray crystallography is the best tool available to obtain structural data, despite some shortcomings. The most common problem in macromolecular X-ray crystallography studies lies with the formation of stable crystals that diffract to high enough resolution to obtain the structural information. This chapter aims to introduce common elements, theoretical as well as experimental, specific to this technique.

## 8.2 The Crystallization Process

Crystallization can summarily be described as the controlled precipitation of the macromolecule through super saturation of the solvent [544]. The formation of the crystal is a phase transition phenomenon that involves: nucleation, growth of the nuclei and cessation of growth [718, 719]. The formation of the crystallization nucleus is the rate-limiting step in the crystallization process [720], where the sample transitions from a disordered state into a more ordered one, theoretically by formation of semi-ordered intermediates [721]. These semi-ordered intermediates can either be the precursors to amorphous precipitate or microcrystals. Amorphous precipitate tends to be predominant either when the protein concentration is solution is well above the saturation point, or when it is brought into the saturation point too quickly [722].

After formation of nucleation points (microcrystals), they will continue to grow, by addition of new molecules to the surface of the crystal, due to either defects in the crystalline order or by the random formation of small nucleation spots at the surface of the crystal [723]. As long as the metastable protein concentration is maintained, the crystals will continue to grow until the saturation zone is reached and cessation of growth occurs.

A 2D phase diagram (Figure 157) is a common visualization aid for the crystallization process described above. It shows the concentration of protein as a function of one parameter of the crystallization condition, with all other parameters held constant [724]. The diagram is divided in four regions: In the under saturated region, no crystal growth is possible, and crystals that are added are quickly dissolved. Above the saturation curve, there is no dissolution of crystals. The

first region of the super saturation zone, metastable region, has no spontaneous nucleation. In the labile region is where nucleation can occur spontaneously. Finally, in the precipitation zone, the concentration of protein is so high that only amorphous precipitate is formed [724].

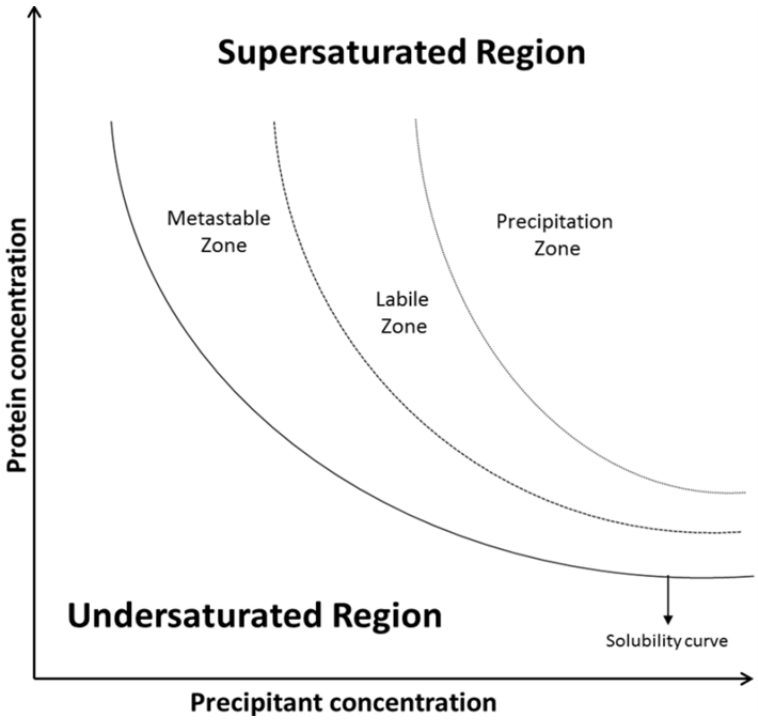



**Figure 157 2D-Phase Diagram.**

Below the solubility curve, there is the under saturated region where crystals do not form. The supersaturated region is divided into three zones: the metastable zone, where spontaneous nucleation does not occur; the labile zone, where new crystal formation may happen; and the precipitation zone, where due to the high concentration of protein, only amorphous precipitate is formed. Adapted from Asherie, N. [724].

Under supersaturation the system is driven towards equilibrium where the molecule is partitioned between entering solid phase and staying in solution. In order to enter the solid phase the individual macromolecule loses both the ability to freely rotate and translate, leading to an overall loss of entropy of the system. To circumvent this unfavourable energetic barrier, the stable bonds formed between molecules reduce the free energy of the system, driving forward the ordering process [725].

In order to minimize the formation of amorphous precipitate and promote stable, controlled nucleation, it is necessary to bring the system slowly to a limited degree of super saturation. This is traditionally achieved by adjusting the conditions to promote crystallization [544, 718, 726].

### 8.2.1 Improving the Crystallization Conditions

The formation of a crystal is the bottleneck of macromolecular X-ray structure resolution. There are various parameters and Method to improve the crystallization process, but the finding of the best conditions for crystal growth remains largely empirical [726]. The single most important factor in obtaining a protein crystal is the properties of the protein sample itself. The sample needs to have a high degree of purity, homogeneity and stability [727].

Provided a pure, stable, homogeneous sample of protein is obtained, the biggest variable in the crystallization process : is the addition of precipitant agents, external additives which can modify the protein precipitation events (Figure 157) [728]. Testing different precipitating agents to determine preliminary conditions for crystal growth is a common methodology, known as a broad screen procedure, and various commercial solutions are readily available. If an initial condition where crystalline-like formations is observed, that condition can be further optimised by varying the precipitant type, precipitant/protein concentrations and other parameters [728, 729].

#### 8.2.1.1 Salt

Many proteins are not soluble in pure water, and the addition of salts in the solution is mandatory to increase the samples' stability. The increase of the solubility of proteins in the presence of salts can be explained through interactions between charged residues at the surface of the protein and the ions in solution [730].

The presence of salt in a crystallization system is, however, a complex interaction. As the ionic strength of the solution decreases, the solubility of ionic compounds in the solution, such as proteins, will decrease as well [731]. This will cause the protein to aggregate by Coulombic attraction between opposite charges on different protein molecules, which may lead to the formation of crystals. This process of decreasing protein solubility by decreasing the ionic network surrounding the molecule is termed "salting-in" [730, 732].

On the other hand, proteins and the salt compete directly to bind water molecules. If the concentration of salt in solution is high enough, the macromolecules become deprived of solvent

and associate between themselves to satisfy their electrostatic requirements. This is a process called “salting-out” [730, 732].

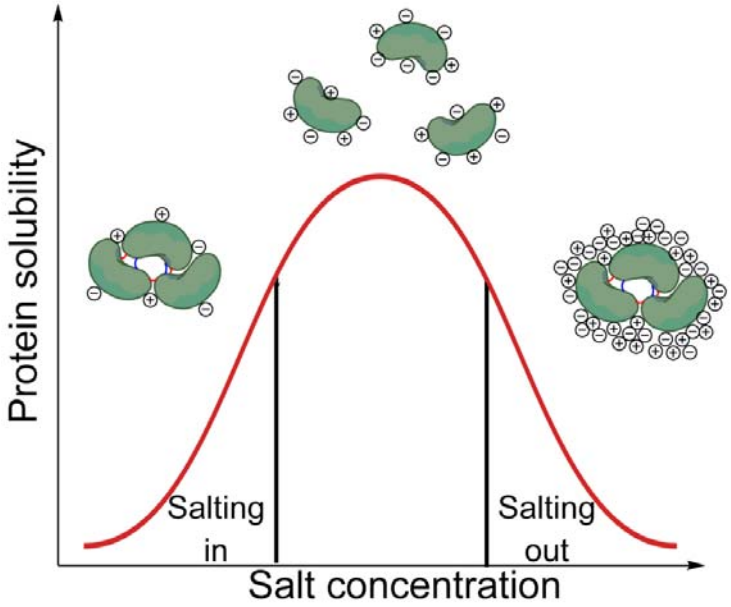



**Figure 158 Protein solubility as a function of salt concentration.**

At low salt concentrations the proteins ionic requirements are satisfied by protein:protein interactions (salting-in). Decreasing the salt concentration below a certain point will lead to protein aggregation to guarantee the proteic ionic requirements (salting-in). At high salt concentrations, the salt ions sequester the solvent, forcing the proteins to aggregate to maintain their hydration layers (salting-out).

#### 8.2.1.2 Polymers

Polymers, such as polyethylene glycols (PEG) (Figure 159, **94**) are molecules which due to their flexible nature are able to sequester a great amount of solvent and space [733]. This, in turn, forces the macromolecules in solution to aggregate to keep the hydration layers. The larger the size of the polymer, the more effective it usually is at forcing the macromolecule to aggregate [732]. Another advantage of the use of PEG in the crystallization is their ability to suffer vitrification when frozen [734], decreasing the formation of ice crystals that could damage the protein. Additionally due to the size of the chains commonly used in crystallography (200-20,000 ethylene-glycol units), it is improbable that they can directly contact the interior of the protein and be present in the active site during the structural analysis stage.

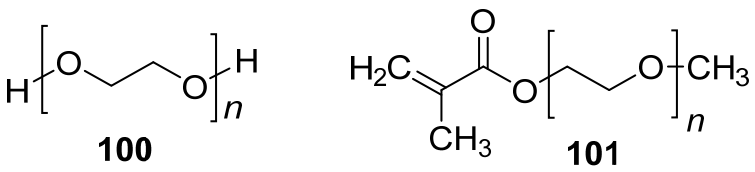



**Figure 159 Structure of commonly used polymers in macromolecular crystallization**

Both polyethylene glycol (PEG) **100** and polyethyelene glycol methyl ether **101** are long chain polymers that are available in many commercial crystallization broad screens.

#### 8.2.1.3 Organic Solvents

Organic solvents are also able to sequester water molecules and can also reduce the dielectric constant of the medium. This latter effect reduces the effective electrostatic shielding between individual macromolecules [544, 732]. However, most organic solvents are highly volatile, making them difficult to accurately handle during the crystallization stage. Regardless, certain compounds, such as 2-methyl-2,4-pentanediol (MPD) has been shown to be particularly useful. MPD is a small polyalcohol, which can compete for water, lower the solution dielectric constant as well as working as a mild detergent [735].

#### 8.2.1.4 Other Additives

In additive screening, a small molecule ligand is added into the crystallization solution, in the hopes its presence will either stabilize the protein or assist in the formation of crystal contacts. Frequently used additives include substrates and products of the protein, metal ions, or any ligand or inhibitor, making a foreknowledge of the mechanism of the reaction (in case of enzymes) a valuable tool in the crystallization process[544, 736].

#### 8.2.1.5 Temperature and pH

Two other important considerations for protein crystallization are the pH and temperature. Proteins have a solubility minimum which corresponds to the isoelectric point of the protein [737]. Additionally the dipole moment of the protein, its conformation and its aggregation state drastically depend on the pH, which can influence the process of salting-in/salting-out [730]. Most crystallization experiments are carried out either at room temperature (25ºC) or 4ºC. The

solubility of proteins in salt solutions tends to increase at low temperatures, whereas in PEG and MPD the solubility decreases with decreasing temperature, making the ideal temperature dependent on the system being studied [728].

### 8.2.2 Vapour Diffusion Method

The most widely used method for crystal formation is the vapour diffusion method, which can be divided into two slightly different techniques: the hanging drop and the sitting drop [544, 720] Method. Both are largely equivalent, using the evaporation and diffusion between solutions of different concentrations as a means of approaching and achieving super saturation of protein macromolecules [544, 720].

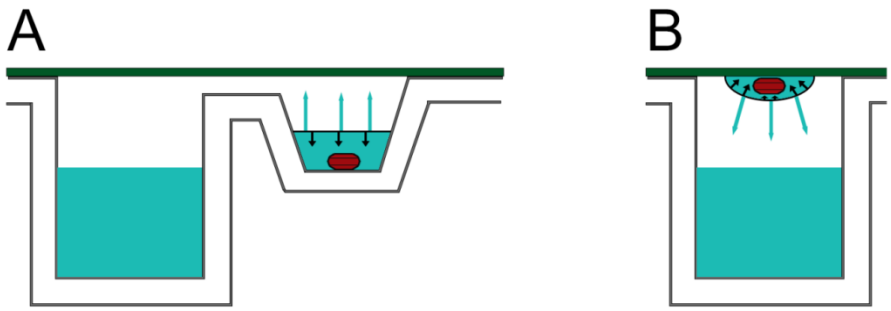



**Figure 160 Most common crystallization setups for vapour diffusion.**

**A:** Sitting drop technique, where a mixture of protein and precipitant is placed in an individual well that is connected to the main well for vapour diffusion to occur. **B:** Hanging drop technique, where the protein:precipitant mixture is placed in a cover slip, closing the reservoir.

After mixing a small volume of protein solution and precipitant agent to form a small drop (from 50 nL to 4 µL), the drop is placed next to a reservoir containing a larger volume of precipitant (from 25 µL to 1 mL). The system is sealed from the outside, and the difference in precipitant concentration between the drop and the well solution causes water to evaporate from the drop to an equilibrium state. This evolution of concentrations the crystallization drop in precipitant and in protein leads to super saturation state and hence to the nucleation and crystal growth [544, 729].

### 8.2.3 Cryo-cooling of Crystals

Even though X-ray diffraction could be considered a non-destructive scattering technique, macromolecular crystals presented to a highly powered X-ray can suffer extensive radiation damage [738, 739]. Collecting the X-ray diffraction information at lower temperature was shown by Hope [740] to significantly reduce the radiation damage of samples. Additionally, at lower temperatures, the atomic movements are reduced, which was proposed to contribute to a better quality of the data, as the estimation of atom location became more accurate [740, 741].

As a result, freezing the crystal samples has become a widespread technique prior to the collection of protein X-ray diffraction data. However, as a significant percentage of volume in protein crystals is usually occupied by solvent [742], ice formation of water molecules within the crystal structure could lead to volume expansion, damaging the crystalline arrangement.

As discussed previously the addition of external agents such as PEG (8.2.1.2), MPD (8.2.1.3) and very commonly, glycerol [638], are usually added as cryo-protectants [637]. Cryo-protectants can, however, alter the protein solubility causing crystal cracking or dissolution [743] and identification of the optimum cryo-protecting solution is necessary to solve a protein structure at cryogenic temperature [744].

### 8.2.4 Crystal Seeding

Seeding is a technique where the optimization of growth conditions is independent from those required for nucleation [546, 745]. It can be divided into two categories, based on the size of the seeds: in macro seeding [746], a single grown crystal is transferred through multiple new protein:precipitant mix; and micro seeding [747] where sub-microscopic crystals are added to the protein: precipitant mix.

In the case of micro seeding, a crystal is placed in a stabilizing solution, sometimes the precipitant solution itself, and broken into microscopic particles which will then serve as small nucleation points. This seed stock usually needs to be diluted, as a high concentration of nucleation points can lead to the formation of a shower of microcrystals (Figure 161, dilution $10^{-1}$). On the other hand, if the seed stock is too dilute, there is the possibility of not transferring any nuclei, and no crystal is formed [747, 748].

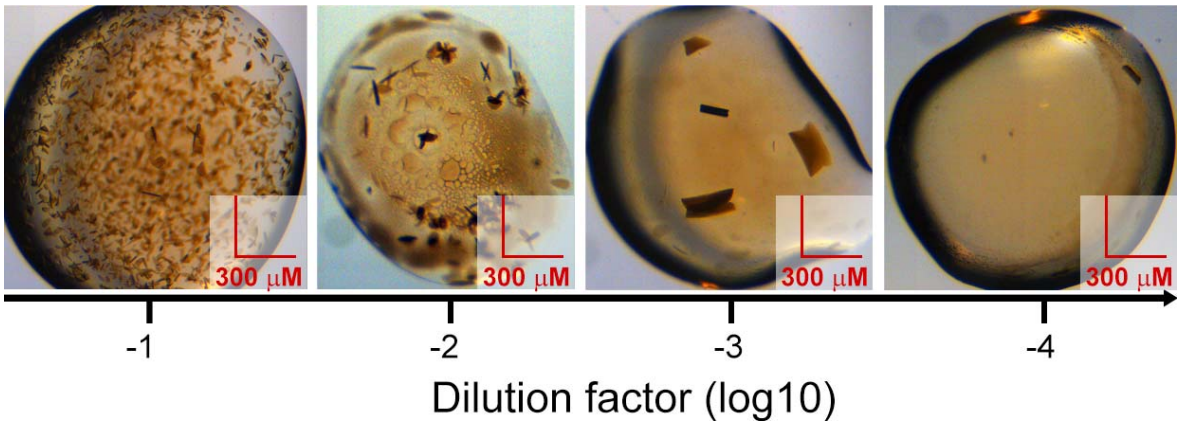



**Figure 161 Example of a successful micro seeding experiment.**

A pre-grown crystal of the hydrogenase maturase protein HydG was crushed into small fragments in a stabilizing solution, followed by sequential dilution of the seed stock. A small volume of the seed stock was injected in crystal drops that did not show protein crystallization.

An alternative method of micro-seeding that does not require serial dilutions was first described by Enrico Stura [749]. In this streak-seeding technique, a cat whisker is used to dislodge small fragments from a pre-grown crystal. The whisker then passed through the drop from one side to the opposite side in one smooth motion, leaving behind small nucleation points.

In a crystallization experiment it is necessary to guarantee the system is saturated in order for a seeding experiment to work. If a saturation condition was not reached, nucleation cannot occur and transferring of a formed crystal into the solution will lead to its dissolution [750].

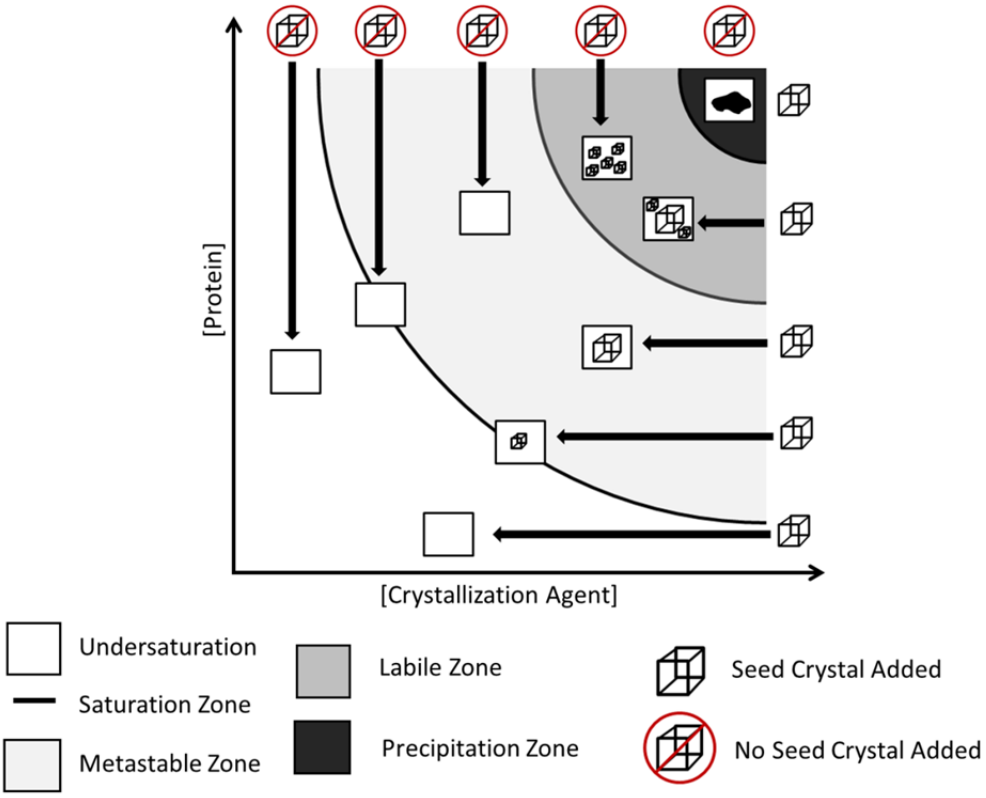



**Figure 162 2D-phase model with and without the addition of crystal seeds.**

Two dimensional representation of a crystallization experiment with and without the addition of a crystal seed. The white region corresponds to an under saturated condition, where the addition of a seed will dissolve into solution. At the interface between under saturation and saturation the nucleation seed will neither dissolve nor increase in size. The next region, super saturation, can be divided into three sub-regions: in the first, metastable super saturation, spontaneous nucleation does not occur but any crystal added to the solution will grow. This is the ideal zone for seeding experiments. The second, named labile super saturation, has spontaneous nucleation occurring, and any seed added will also grow. The final sub-region is the precipitation zone where no crystal growth occurs and amorphous precipitate is formed. Adapted from Luft, J [751].

### 8.2.5 Matthews Coefficient

Protein molecules in crystals are held together by weaker forces, primarily hydrogen bonds between hydrated protein surfaces. However, as proteins have very specific 3-D structures, it would be impossible to fulfil all spaces perfectly in the crystal grid. The presence of solvent molecules within a protein crystal structure is, in that regard, an important component of the crystal. Solvent content of protein crystals was first systematically introduced by Brian W. Matthews in 1968 [742], which defined:

$$V_M = \frac{V}{M \times n \times X}$$

**Equation 6 Matthews coefficient.**

Calculation of the specific volume occupied by a protein within a unit cell ($V_M$), also known as Matthews coefficient. V, volume of the unit cell; M, molecular weight of the protein; n, number of asymmetric units in the unit cell (determined by the number of symmetry operators in the space group), and X is a variable that relates unique properties on the protein sample. All these components reflect the number of molecules in the asymmetric unit.

Most importantly, the Matthews coefficient also helps to evaluate the overall packing features of protein molecules in the crystal. Higher solvent content may be indicative that protein packing may not be very compact, which can negatively affect the quality of the data obtained. In those cases dehydration of the sample, by air drying the crystal, crystallizing in vapour diffusion using higher salt content, and soaking with dehydrating compounds may improve the diffraction data obtained from the crystals [736, 752].

## 8.3 X-Ray Diffraction

In 1912 Bragg deduced there was a relation between the incident wavelength and scattering angle of the beam and the distance between atomic planes of a crystal lattice in what is now known as Bragg's law:

$$n \times \lambda = 2 \times d \times \sin\theta$$

**Equation 7 Bragg's Law.**

n, positive integer; $\lambda$, wavelength of the incident beam; d, distance between the crystal planes; $\theta$, angle of diffraction.

The reflected photons can interfere constructively or destructively. As constructive interference occurs only at a small number of discrete scattering angles, determined by the crystal lattice, the results are the spots (named Bragg spots) observed on diffraction patterns (Figure 163) [515, 736].

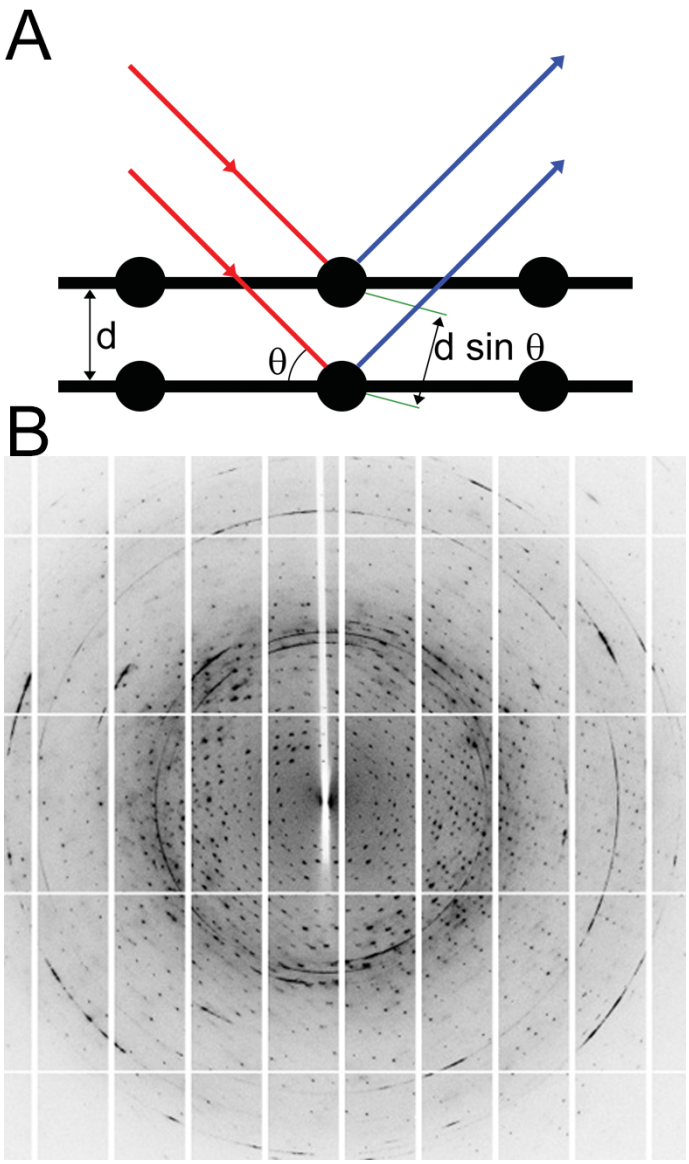



**Figure 163 Bragg's law and consequent diffraction**

The constructive interference of reflected waves in **A** obeys Bragg's law (Equation 7) and experimentally leads to discrete Bragg spots which can be observed in the diffraction pattern of *Ti*HydG crystal **B**.

A crystal contains many identical copies of a molecule arranged into a three dimensional periodic structure, named the lattice. One can think of the lattice as a train of delta functions in 3-dimensions, and the crystal as a single molecule convolved with these delta functions. The basic repeating unit that forms the lattice is called the unit cell, characterized by the parameters (a, b, c, α, β, γ), corresponding to the lengths (a, b, c) and angles (α, β, γ) of the repeating unit.

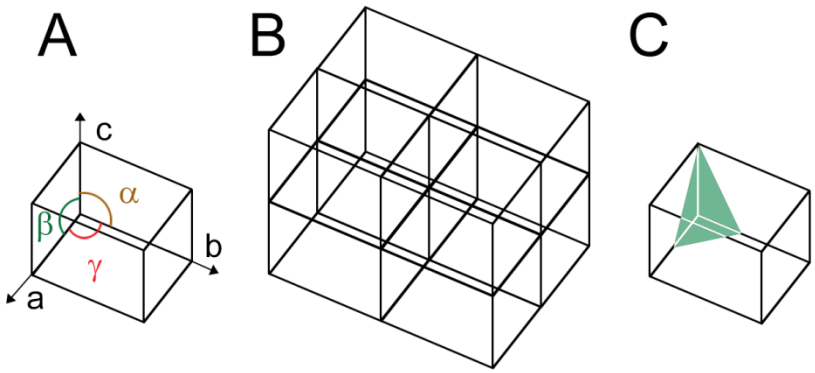



**Figure 164 Representation of unit cell, crystal lattice and a plane with miller indices.**

**A:** Unit cell with length a, b, c and angles $\alpha$, $\beta$, $\gamma$. **B:** A crystal lattice is an infinite array of repeating unit cells. **C:** A plane that intersects the a, b, c edges in a unit cell, is said to possess Miller indices (h, k, l). The represented plane is (2,2,1), as the intercept in the a direction is ½ the length of a and b and one full length of c.

In a diffraction experiment, the intensities and the position of reflections are measured. From the spots in the diffraction pattern, the diffraction angle $\theta$ and $d_{(hkl)}$ can be calculated based on Bragg's law. After a series of $d_{(hkl)}$ have been calculated, the values are investigated to both assign the index h,k,l to each reflection as well as calculate the unit cell parameters (a, b, c; $\alpha$, $\beta$, $\gamma$) [515, 736].

### 8.3.1 The Structure Factor

The way in which the separate diffracted rays combine to form an image depends on three factors within each ray: the direction, the amplitude and the phase. In X-ray crystallography, the scattered rays are separately observed and their intensities measured as the intensity of the spots in the pattern (Figure 163).

By determining the intensity of the signal, one can calculate the amplitude of the scattered rays, through Equation 8.

$$I_{h,k,l} = A^2$$

**Equation 8 Intensity of h,k,l related to the amplitude (A) of the diffracted ray.**

Furthermore, by identifying the Miller indices (h,k,l) of the crystal plane giving rise to each scattered ray, the direction of the ray is specified. Only the phase information of the diffracted ray cannot be directly measured, but it is related to the structure factor $F_{h,k,l}$. [515]. The structure factor $F_{(h,k,l)}$ of any x-ray reflection h,k,l represents a quantity related to the intensity of the reflection which is independent of the method for observation and being related solely to the structure that gives rise to that reflection.

$$F_{h,k,l} = A_{h,k,l}\exp(i\alpha_{h,k,l})$$

**Equation 9 Structure factor $F_{h,k,l}$.**

The structure factor $F_{h,k,l}$ describes the amplitude A and phase $\alpha$ of wave diffracted from a crystal plane h,k,l.

If the coordinates of the structure are known, the structure factor $F_{h,k,l}$ can be easily determined by Equation 10.

$$F_{h,k,l} = \sum_{j=1}^{atoms} f_{(j)} e^{[2\pi\, i(hx_j + hy_j + hz_j)]}$$

**Equation 10 Structure factor $F_{h,k,l}$,**

The structure factor $F_{h,k,l}$, can be determined by the sum of the coordinates of all atoms in the unit cell. $x_j$,$y_j$,$z_j$ are the positional coordinates of the $j^{th}$ atom and $f_j$ is the scattering factor of the $j^{th}$ atom

However, in a crystal experiment, the coordinates of the structure are what is to be determined and the amplitudes are measured. However, the exponential function in Equation 10 is periodic, and a Fourier Transform can be applied.

The Fourier transform of the convolution of two functions is equal to the product of their separate Fourier transforms. As such, the diffraction pattern from the crystal is related to the Fourier transform of the lattice, multiplied by the Fourier transform of a single molecule [515]. As only the electrons of the atoms in the molecule interact appreciably with the X-ray beam, the Fourier transform of the molecule is the Fourier transform of the electron distribution in the molecule, often called the electron density map [515]:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{h,k,l} e^{[-2\pi i(hx+ky+lz)]}$$

**Equation 11 Electron density calculation.**

ρ (xyz) is the electron density at point (xyz) in the unit cell; V is the volume of the unit cell; $F_{h,k,l}$ is structure factor. x, y, z is the position of the atoms in the normalized unit cell.

Unfortunately, as previously mentioned, a crystallographic experiment only measures the amplitudes (A) of the diffracted beams and not the phases (α), and consequently only the magnitude of the structure factor ($|F_{h,k,l}|$) is known (Equation 9). Thus Equation 11 can be re-written as Equation 12.

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} |F_{obs}| e^{[-2\pi i(hx+ky+lz)-\alpha_{h,k,l}]}$$

**Equation 12 Electron density calculation.**

ρ (xyz) = electron density at point (xyz) in the unit cell; V = the volume of the unit cell; $|F_{obs}|$ structure factor amplitude for the reflection and $\alpha_{h,k,l}$ is the phase angle of $F_{obs}$.

The need to determine the phase component which cannot be directly measured in an experiment is known as the “phase problem”. For small molecules, the loss of information associated with the phase problem can be solved using direct phasing methods. These methods are, however, unavailable for large molecules, as the complexity of the calculations rises significantly with the complexity of the molecule [753].

### 8.3.2 Solving the Phase Problem

#### 8.3.2.1 Molecular Replacement

If previous knowledge of a similar protein structure is available, molecular replacement (MR) is the easiest method to solve the “phase problem” [754, 755]. Using the known model as a template, the phase angles of the known model are applied to the indexed diffraction data of the unknown crystal structure. Automatic programs have been devised that create rotation and translation algorithms to find the highest correlation between the experimental diffraction measurements and those calculated from the model. As the number of determined structures available continues to rise, so does the potential template models for molecular replacement.

#### 8.3.2.2 Multiple Isomorphous Replacement

When no similar structure exists, it is necessary to use isomorphous replacement and related Method. From Bragg’s law, the crystal lattice can be determined by the diffraction pattern alone, and two crystals with the same lattice will have the same index h, k, l. If, theoretically, a salt crystal could be found which shared the same crystal grid as the protein direct method could be used to solve the structure of the salt to obtain both the intensity and the phase information. The phase from the salt could then be used to help determine the phase of the protein crystal This is the basic principle of isomorphous replacement [515].

In most proteins, when heavy atoms are not available in the structure, they are added to the protein, so that the native protein structure and the heavy atom derivative crystals are essential the same, or isomorphous [756]. Then, both metal-bound and native crystals are tested, comparing the phases of two heavy atoms in different positions, and using one as a reference. The differences of each reflection between the two datasets are attributed to the heavy atoms [515, 757].

#### 8.3.2.3 Anomalous Diffraction

If the protein already possesses a heavy atom, there is no need to neither modify the structure nor obtain three isomorphous crystal structures. The scattering of the heavy atom can be modified and its position may be inferred. This method is known as anomalous diffraction (AD) [758].

In crystal grids, under normal diffraction conditions, structure factors of the reflections from opposite sides of the same plane ($_{h,k,l}$) are identical in amplitudes, but opposite in phases. This is called Friedel's law. Additionally, as experimentally the incident energy is much greater than the binding energies within the atom, the electrons are considered free. The scattering of X-rays by these free electrons is known as Thompson scattering [759].

The approximation that the electrons in a macromolecule within a crystal obey the Thompson scattering stops being true when the incident X-ray energy is close to the binding energies of electrons. These energy values are known as the absorption edges of the atom and, in the case of heavy atoms, their absorption edges coincide with X-ray energies used for crystallographic diffraction [760].

When the incident energy is near the absorption edge, some photons are scattered normally, and some photons are absorbed and re-emitted with altered phase. Thus, the scattered photon gains an imaginary component to its phase, and Friedel's law is broken. The scattering becomes a complex quantity and is written as:

$$f = f_0 + f' + if''$$

**Equation 13 Scattering factor**

Scattering factor, f. $f_0$ is the Thompson scattering factor; f' is the real anomalous correcting term; if'' is the imaginary anomalous scattering correction term.

In this case, as the anomalous scattering is caused by heavy atoms the differences between $F_{h,k,l}$ and $F_{-h,-k,-l}$ arise from the heavy atoms' anomalous scatterings. These differences could be used to extract the position of heavy atoms, and further calculate the phases of protein's structure factors [515].

Two methods exist that take advantage of anomalous diffraction: single wavelength anomalous diffraction (SAD) and multiple wavelength anomalous diffraction (MAD) [736, 761]. In a MAD experiment several datasets are collected, not only peak dataset, but also inflection dataset, or even remote dataset and the differences between diffraction amplitudes at different wavelengths are used to obtain a unique phase choice [762, 763]. In SAD, only one single crystal anomalous scattering is used. The knowledge that the imaginary anomalous scattering correction term breaks the symmetry of Friedel pairs makes it possible to determine the heavy atom substructure using

direct Method, which are usually only available for small molecules, which in turn allows to solve the phases of the protein diffraction data [550].

#### 8.3.2.4 Data Analysis

After a set of diffraction images is obtained analysis and reduction of the data needs to be carried out.

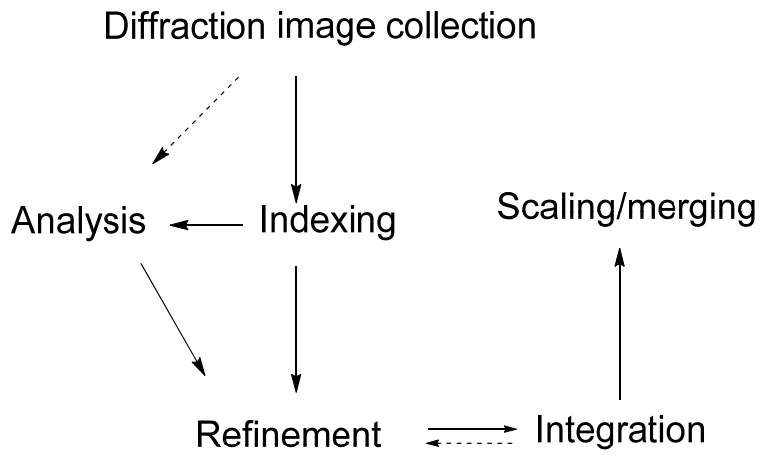



**Figure 165 General process of data processing.**

The diffraction images are indexed and then reflections are integrated. All reflections from the data collection are merged and scaled. Adapted from Incardona *et al.* [551].

Synchrotron beamlines now possess pipelines that automatically carry out the majority of the data processing. In Diamond Light Source (DLS) the pipeline begins using X-ray Detector Software (XDS) [554]. Auto indexing algorithms are able to calculate the unit cell dimensions and orientation of the crystal, as long as beam position, crystal-detector distance and radiation wavelength are known [764]. After an initial orientation has been found, the differences between the observed diffraction spot positions and the predicted positions calculated are minimized.

Once refinement of the crystal orientation for one image is complete all spots in the images are integrated. Integration predicts the position of each Bragg spot, estimates their intensities and associates an error to that estimation [748]. This outputs a list of reflections (hkl) and their unscaled intensities (XDS).

Following integration, Pointless and Scala [555, 621] are used to determine the space group and scale the intensities obtained from XDS. Placing every image in a relative scale (scaling) removes the effects of both the variations in the intensity of the X-ray source and in the intensity of the diffraction caused by changes in the volume of the crystal as it rotates.

The processing by Scala also includes the use of Truncate [765] which estimates the amplitudes of the reflections from the averaged intensities. This process forces all negative observations to become positive, and the weakest reflections are inflated, in order not to underestimate those observations [133].

The amplitudes of the reflections are then determined from their scaled intensities, but the phase, as described previously, cannot be directly determined without solving the 'phase problem' ([743, 766]. After scaling, a scaled structure factor file is obtained for data refinement and structure determination.

Qualitative factors should be considered during data collection process to get a good data set. Several variables can be experimentally changed, such as exposure time for each collection to maximize completeness and minimize radiation damage, crystal: detector distance to maximize data signal and minimize overloads . The data quality after scaling is usually judged by the global $R_{merge}$ factor [767], though other factors such as redundancy, $\sigma$I/noise ratio, mosaicity factor are also important parameters to judge the data quality [555].

$$R_{merge} = \frac{\sum_{h,k,l} \sum_{i} |I_{h,k,l} - < I_{h,k,l} >}{\sum_{h,k,l} \sum_{i} < I_{h,k,l} >}$$

**Equation 14 $R_{merge}$**

Calculation of $R_{merge}$ of the diffraction data. $<I_{h,k,l}>$ is the of Friedel related observations in a unique reflection.

#### 8.3.2.5 Data Refinement

Structure refinement aims to optimise the agreement between an atomic model and both the observed diffraction data and chemical restraints [768]. In other words, the target of refinement is to minimize the difference between structure factors calculated from the model and the observed structure factors.

WinCoot is a program that allows for the observation of protein coordinates and associated electron density, together with manual manipulation of the model to improve the fit of the density [561]. Together with manual improvements of the fit, Phenix.refine [558] is a program which automatically provides several adjustments, such as fixing rotamers, introducing water molecules as well as building ligands into the positive density.

This process can be repeated iteratively through cycles of refinement which is reflected in the R-factor ($R_{work}$) defined as:

$$R = \frac{\sum|F_{obs}| - |F_{calc}|}{\sum|F_{obs}|}$$

**Equation 15 R-factor calculation.**

$F_{obs}$ is structure factors from measured diffraction intensity; $F_{calc}$ is the structure factors calculated from the model.

During the refinement process, R-factor should become smaller and smaller, with a value of below 0.2 in high resolution structures usually representing a well-defined protein structure [769, 770].

As a method to reduce over-refinement and to function as a quality control process the calculated and observed structure factors, the free R-factor (or $R_{free}$) was introduced [770]. $R_{free}$ is computed in the same way as the conventional $R_{work}$, but applied to a small set of randomly chosen intensities which are not used during the refinement of the structural model [770, 771], becoming a less biased measure of the refinement process. High differences between the values of $R_{free}$ and $R_{work}$ are usually indicative that too many parameters were introduced in the structure model artificially reducing the $R_{work}$ value.

## 8.4 Materials

**Table M1 Chemicals used in this study and their suppliers**

| Chemical | Supplier |
|---|---|
| SAM dihydrochloride | Sigma Aldrich |
| SAH | |
| Iron (III) chloride | |
| Iron (II) chloride | |
| Sodium sulphite nonahydrate | |
| L-tyrosine | |
| Tyramine | |
| L-DOPA | |
| L-Tyrosinamide | |
| N-methyl-L-tyrosine | |
| HTHIQCA | |
| Trifluoroacetic acid | |
| Triethylamine | |
| Bacto-tryptone | Oxoid |
| Yeast extract | |
| DTT | Melford |
| IPTG | |

| Chemical | Supplier |
|---|---|
| Antibiotics | Melford |
| Imidazole<br>HEPES<br>Glycerol<br>NaCl<br>30% polyacrylamide-bis solution<br>Formic acid<br>Sodium dodecyl sulfate<br>Sodium carbonate<br>Bacto-agar<br>HPLC-grade solvents for HPLC | Fisher Scientific |
| 4-HPPA<br>4-HPHPA | Astatech, inc. |

**Table M2 Consumables used in this study and respective suppliers**

| Consumable | Supplier |
|---|---|
| 24-well crystallization plates<br>96-well crystallization plates | Greiner Bio-One |
| 0.22 μm filters | Millipore |

| Consumable | Supplier |
|---|---|
| 50 mL Falcon Tubes<br>15 mL Falcon tubes<br>PAGE-Ruler protein ladder<br>EZ-run protein ladder<br>GeneRuler 1Kb DNA ladder<br>RiboRuler low range RNA ladder | Fisher Scientific |
| Nancy-520 nucleotide dye<br>Bradford reagent<br>Antifoam 204 | Sigma-Aldrich |
| E.Z.N.A. miniprep kit II<br>E.Z.N.A. agarose gel extraction kit | VWR International |
| EDTA-free cOmplete protease inhibitor tablets | Roche |
| Pre-packed Sephadex G-25 PD-10 columns<br>Superdex 75<br>Chelating sepharose Fast Flow<br>Streptavidin Sepharose High Performance | GE Healthcare |
| PES based spin filters, 10 kDa cut-off | Sartorius Stedim Biotech |
| Far UV quartz cuvettes, 1 cm path | Starna Scientific Ltd. |
| Analytical HPLC columns | Phenomenex |

**Table M3 Enzymes used in this study and respective suppliers**

| Enzyme | Supplier |
|---|---|
| BSA | Melford |
| Restriction enzymes (and buffers) | New England Biolabs |
| Phusion polymerase | Fisher Scientific |

## 8.5 Equipment

### 8.5.1 DNA/RNA amplification and analysis

PCR DNA amplifications and RNA *in vitro* transcription were carried out in a Eppendorf Mastercycler gradient. Agarose gels were run in an electrophoresis apparatus (Bio-Rad). Gel images were collected in SynGene GeminyBV UV transilluminator, and analysed with GeneTools. DNA and RNA quantification was carried out in a Nanodrop ND-1000 spectrophotometer.

### 8.5.2 Bacterial growth

Media for bacterial growth was sterilized in a PriorClave steam autoclave (PriorClave ltd). Small scale (100 mL) bacterial cultures were incubated in an Innova 4400 incubator shaker, 37ºC with 180 rpm agitation (New Brunswick, Scientific). 5L cultures for protein expression were grown in a BioFlo ® 110 fermenter unit (Eppendorf UK Ltd). Bacterial plate cultures were grown in an Economy Incubator Size 2 (Gallenkamp).

### 8.5.3 pH determination

pH measurements were performed using a Mettler Delta 340 pH meter connected to a Mettler Toledo Inlab 413 Combination Electrode calibrated at pH 4.0, pH 7.0 and pH 10.0 before use and stored in 3 M potassium chloride.

### 8.5.4 Anaerobic protein purification

All experiments involving oxygen sensitive proteins were carried out at constant temperature (20ºC) inside an anaerobic glove box (Belle Technology, Portesham, UK, $O_2$ < 2 ppm), maintained at less than 2 ppm $O_2$. The anaerobic glove box was equipped with a Pharmacia Acta FPLC (GE Healthcare, Buckinghamshire, UK) and a VC 130 Sonicator (Sonics and Materials, Newtown, Connecticut, USA). All containers (bottles, jars, tubes, beakers etc.) and equipment required during the experiments were degassed overnight inside the glove box. Buffers and solutions were bubbled with N for 10 min before being introduced into the glovebox and were allowed to deoxygenate overnight before use. Purification was carried out with Pharmacia XK columns and protein solutions were concentrated by ultrafiltration using an Amicon pressure-cell or spin filters.

### 8.5.5 SDS-PAGE analysis

The migration of protein samples through a polyacrylamide gel was carried out in an electrophoresis apparatus (Bio-Rad). Gel images were collected in SynGene GeminyBV, using the available white light. Densitometric analysis of SDS-PAGE gels was carried out with ImageJ.

### 8.5.6 Centrifugation

Samples were centrifuged at 4 ºC using a Sorval centrifuge fitted with a SLC6000 or a JA-14 rotor. For small volumes (< 1.5 ml) a bench top microcentrifuge (Centrifuge 5415 D) was used at room temperature.

### 8.5.7 UV-Vis spectroscopy

General absorbance readings and UV-vis spectra were recorded on a Thermoscientific BioMate 3 spectrophotometer. Anaerobic UV-vis spectra of RSAM enzymes was collected in a Lamba 2 spectrophotometer (Perkin-Elmer), or on a USB2000 spectrophotometer using a light source Mini-D2-GS (Ocean Optics (Duiven, The Netherlands)).

### 8.5.8 CD spectroscopy

CD spectroscopy analysis was carried out with a Jasco J-710CD, under a constant flux of nitrogen (20 psi). Lamps were cooled by a chilled water loop.

### 8.5.9 HPLC Analysis

Chromatographic separation of compounds was achieved on a workstation (Gilson 321 Pump with H1 Heads and a Gilson 234 Autoinjector) equipped with a dual wavelength Gilson-UV-Vis-155 detector and a Shimadzu RF-10AXL fluorimeter. Chromatograms were analysed using Gilson Unipoint Software (version 5.11).

### 8.5.10 Crystallography

Crystal trays were manually prepared inside the glove box (Belle Technology, Portesham, UK), maintained at less than 2 ppm $O_2$. Crystal growth was monitored using a Meiji EMZ-13TR microscope with an Infinity 1 camera situated inside the glove box and connected via a USB cable to a laptop outside. Images were taken using the associated Infinity Capture or Infinity Analyse software.

## 8.6 Media, buffers and solutions

**Table M4 2xTY media.**

| Component | Quantity |
|---|---|
| Bacto-Tryptone | 80g |
| YeastExtract | 50g |
| NaCl | 25g |
| Water | 5000mL |

**Table M5 Autoinduction media.**

| Component | Quantity |
|---|---|
| Bacto-Tryptone | 100g |
| YeastExtract | 25g |
| NaCl | 25g |
| Na2HPO4 | 6 |
| NaH2PO4 | 3 |
| Water | 5000mL |

**Table M6 Purification Buffers (LipA and HydG).**

| Component | Quantity | | | |
|---|---|---|---|---|
| | Buffer A | Buffer B | Buffer C | Buffer D |
| HEPES | 25 mM – 4.77g | 25 mM – 4.77g | 25 mM – 4.77g | 25 mM – 4.77g |
| NaCl | 500 mM – 29.22g | 500 mM – 29.22g | 500 mM – 29.22g | 0 |
| MgCl2 | 0 | 0 | 0 | 500 mM – 47.61g |
| Glycerol | 10% - 100g | 10% - 100g | 10% - 100g | 10% - 100g |
| Imidazole | 20 mM – 1.36g | 500mM – 34.05g | 0 | 0 |
| Dithiothreitol | 0 | 0 | 5 mM – 0.77g | 5 mM – 0.77g |
| Water | Adjust to 1L | Adjust to 1L | Adjust to 1L | Adjust to 1L |

pH adjusted to 7.4

**Table M7 Purification Buffers (Cfr).**

| Component | Quantity | | |
|---|---|---|---|
| | Buffer A | Buffer B | Buffer C |
| Bis-Tris Propane | 50 mM – 14.12g | 50 mM – 14.12g | 50 mM – 14.12g |
| NaCl | 500 mM – 29.22g | 500 mM – 29.22g | 0 |
| KCl | 0 | 0 | 500 mM – 37.28g |
| Glycerol | 20% - 100mL | 20% - 100mL | 10% - 100mL |
| Imidazole | 20 mM – 1.36g | 500mM – 34.05g | 0 |
| Dithiothreitol | 0 | 0 | 5 mM – 0.77g |
| Water | Adjust to 1L | Adjust to 1L | Adjust to 1L |

pH adjusted to 8.4

**Table M8 GNTC Buffer for RNA extraction**

| Component | Quantity |
|---|---|
| Guanidine thiocyanate | 4M (70.9 g) |
| Sodium citrate | 1.1g |
| Sodium citrate | 25 mM (1.1g) |
| 0.5% Sodium lauryl succinate | 2.5 mL |
| Water | 150 mL |

pH adjusted to 7.0

**Table M9 5X MOPS buffer**

| Component | Quantity |
|---|---|
| MOPS | 0.2 M (83.72 g) |
| Sodium acetate | 0.05 M (8.23 g) |
| EDTA | 0.005 M (20 mL of a 0.5 M solution) |
| Water | 2L |

pH adjusted to 7.0

**SDS-PAGE:**

**Table M10 SDS-PAGE resolving gel (4 gels)**

| Component | Volume/mL |
|---|---|
| Water | 4.6 |
| 30% Acrylamide Mix | 10 |
| 1.5M Tris pH 8.8 | 5 |
| 10% SDS | 0.2 |
| 10% Ammonium persulfate | 0.2 |
| TEMED | 0.008 |

**Table M11 SDS-PAGE stacking gel (4 gels)**

| Component | Volume/mL |
|---|---|
| Water | 5.4 |
| Components | Volume/mL |
| 30% Acrylamide Mix | 1.34 |
| 1.5M Tris pH 8.8 | 1 |
| 10% SDS | 0.08 |
| 10% Ammonium persulfate | 0.08 |
| TEMED | 0.008 |

**Table M12 SDS-PAGE loading Buffer**

| Component | Quantity |
|---|---|
| 0.2M Tris-HCl pH 6.8 | 2.5mL |
| DTT | 154mg |
| SDS | 200mg |
| Bromophenol Blue | 10mg |
| Glycerol | 1mL |
| Water | Adjust to 5mL |

**Table M13 SDS-PAGE running buffer**

| Component | Quantity |
|---|---|
| Tris Base | 15.1g |
| Glycine | 94g |
| SDS | 5g |
| Water | 5L |

**Table M14 SDS-PAGE Gel staining solution**

| Component | Quantity |
|---|---|
| Coomasie Brilliant Blue | 25g |
| Methanol | 450mL |
| Glacial Acetic Acid | 100mL |
| Water | 450mL |

**Table M15 SDS-PAGE Gel destaining solution**

| Component | Volume/mL |
|---|---|
| Methanol | 375 |
| Glacial Acetic Acid | 250 |
| Water | 4375 |

**Iron Quantification by the Fish Method:**

**Table M16 Reagent A:**

| Component | Quantity |
|---|---|
| $KMnO_4$ | 0.22g |
| Water | 5mL |
| HCl (1.2M) | 5mL |

**Table M17 Reagent B:**

| Component | Quantity |
|---|---|
| Ascorbic Acid | 4.4g |
| Ammonium Acetate | 4.85g |
| Ferrozine | 40mg |
| Neocuproine | 40mg |
| Water | 12.5mL |

**Table M18 *Ti*HydG identified crystallization conditions**

| Screening kit | Hampton Condition | Components | | |
|---|---|---|---|---|
| Pact premier | E12 | | 20% w/v polyethylene glycol 3,350 | 0.2M Sodium Malonate |
| Pact premier | F1 | 0.1 M Bis-Tris Propane, pH 6.5 | 20% w/v polyethylene glycol 3,350 | 0.2M Sodium Fluoride |
| PEGRx HT | A8 | 0.1 M MES monohydrate, pH 6.0 | 22% v/v Polyethylene glycol 400 | |
| PEGRx HT | C2 | 0.1 M Imidazole, pH 7.0 | 20% w/v Jeffamine ED-2001, pH 7.0 | |
| PEGRx HT | C9 | 0.1 M Tris, pH 8.0 | 28% w/v polyethylene glycol 4000 | |
| PEGRx HT | C11 | 0.1 M Bis-Tris, pH 6.5 | 20% w/v polyethylene glycol monomethyl ether 5,000 | |
| PEGRx HT | D2 | 0.1 M imidazole, pH 7.0 | 20% w/v polyethylene glycol 6,000 | |
| PEGRx HT | F6 | 0.1 M Bicine, pH 8.5 | 30% w/v Polyethylene glycol 1500 | 10% Isopropanol |
| PEGRx HT | G7 | 0.1 M Imidazole, pH 7.0 | 2% v/v polyethylene glycol 400 | 24% polyethylene glycol monomethyl ether 5000 |

**Table M19 LipA Crystal Condition**

| Hampton Condition | Components | | |
|---|---|---|---|
| PEG RX2 48 | 0.1M Bicine, pH 8.5 | 15% w/v Polyethylene glycol 20.000 | 3% w/v Dextran sulphate sodium salt |

**Table M20 PCR reaction mix**

| Component | Quantity |
|---|---|
| Phusion HF Buffer | 10μL |
| 2 mM dNTP | 5 μL |
| Primer | 1.0 μM |
| Template DNA | 10 ng |
| Phusion DNA Polymerase | 0.5 μL |
| Water | Up to 50 μL |

**Table M21 DNA analytical digestion reaction mix**

| Component | Quantity |
|---|---|
| Plasmid DNA | 50-250 ng/μL |
| Buffer (10X) | 1μL |
| Restriction enzyme (each) | 0.5 μL |
| Water | Up to 50 μL |

**Table M22 DNA ligation mix**

| Component | Quantity |
|---|---|
| Insert DNA | 75 ng |
| Vector DNA | 25 ng |
| T4 DNA ligase buffer (10X) | 2 μL |
| T4 DNA Ligase | 1 μL |
| Water | Up to 20 μL |

**Table M23 Agarose Gel**

| Component | Quantity |
|---|---|
| Water | 50 mL |
| Agarose | 0.6 g |
| Nancy-520 Stain | 2.5 μL |

**Table M24 Buffer TAE (50X)**

| Component | Quantity |
|---|---|
| Water | 1000 mL |
| Tris Base | 242 g |
| Acetic Acid (Glacial) | 57.1 mL |
| 0.5M EDTA, pH 8.0 | 100 mL |

**RNA Extraction:**

**Table M25** Buffer L

| Component | Quantity |
|---|---|
| Water | 100 mL |
| Tris-HCl | 121.1 mg |
| Magnesium Acetate | 142.3 mg |

pH adjusted to 7.4

**Table M26 NaCl solution**

| Component | Quantity |
|---|---|
| Water | 100 mL |
| NaCl | 2.92 mg |

**Table M27 40 mM Phosphate Buffer**

| Component | Quantity |
|---|---|
| Water | 100mL |
| Monosodium phosphate monohydrate | 23.34mg |
| Disodium phosphate, heptahydrate | 61.86mg |

pH adjusted to 7.0

**Bleach Agarose:**

**Table M28 1% agarose gel**

| Component | Quantity |
|---|---|
| Water | 50 mL |
| Agarose | 0.5 g |
| Sodium hypochlorite (6%) | 500 □L |

**Formaldehyde Agarose:**

**Table M29 2% agarose gel**

| Component | Quantity |
|---|---|
| Water | 37 mL |
| Agarose | 1 g |
| 10X MOPS | 5 mL |
| Formaldehyde (37%) | 8 mL |

**Table M30 Buffer MOPS (10X)**

| Component | Quantity |
|---|---|
| Water | 1000 mL |
| MOPS | 41.9 g |
| Sodium Acetate.3H2O | 8.2 g |
| EDTA | 3.72 g |

## 8.7 General Microbiological Techniques

Standard sterile techniques were applied throughout microbiological experiments. Growth media (Table M4 2xTY media.) and heat stable solutions were sterilized in a PriorClave autoclave (Priorclave Ltd) at 121 °C for 25 min. Solutions that could not be sterilized, due to their heat lability, were filter sterilised through 0.22 μm filters (Millipore), for e.g., the appropriate antibiotic for growth. Liquid bacterial cultures were incubated in an Innova 4400 Incubator Shaker (New Brunswick Scientific) at 180 rpm. Bacterial plate cultures were grown at 37 °C overnight in an Economy Incubator Size 2 (Gallenkamp).

**Table M31 Stock solutions and amount of antibiotics added to growth media.**

| Antibiotic | Stock concentration | Final concentration |
|---|---|---|
| Ampicillin | 100 mg/mL in $H_2O$ | 100 μg/mL |
| 0Streptomycin | 50 mg/mL in $H_2O$ | 50 μg/mL |

### 8.7.1 Small scale culture

The appropriate antibiotic was added to 2xTY media (100 mL) (Table M4 2xTY media.), followed by the introduction of a small inoculum from a glycerol stock (section 9.4.2). The culture was left overnight, at 37° C, under agitation (180 rpm).

### 8.7.2 *E.coli* cell competence by treatment with $CaCl_2$

2xTY media (Table M4 2xTY media.) was inoculated with 1 mL overnight cultures of *E. coli* (TOP10: analytical digests and DNA isolation, BL21(DE3): protein expression studies, BL21(DE3) STAR: RNA expression studies). The resulting culture was incubated with shaking (37 °C, 180 rpm) until an $OD_{600}$ of 0.4 of reached. The culture was immediately placed on ice for 20 min, to guarantee a constant cold temperature, and then harvested by centrifugation (50 mL, 4 °C, 4,000 rpm, 10 min). The supernatant was decanted and the pellet resuspended in 10 mL ice cold 0.1 M $CaCl_2$. The culture was then incubated under agitation at 4 °C (20 min), followed by harvesting once more by centrifugation (4 °C, 4,000 rpm, 10 min). The supernatant was discarded and the cells gently resuspended in 5 mL ice cold 0.1 M $CaCl_2$ +15% glycerol. The cultures were aliquoted (150 μL) in microtubes and frozen at -80 °C.

### 8.7.3 Competent cell transformation

150 μL of competent *E. coli* cells were defrosted on ice (10 min) and gently resuspended. The plasmid DNA (10μl, 10 ng/μL) was added and the mixture was incubated on ice (30 min). The cells were heat shocked (42 °C, 30 seconds) in a water bath. Cells were then returned to ice and 250μl 2x TY medium added. Cells were then incubated under agitation (200rpm, 37◦C, 1hr), after which

time they were plated out in agar plates supplemented with the corresponding antibiotic, and incubated at 37°C overnight. A control without the addition of plasmid DNA was also grown and plated in the same conditions.

Single colonies were then picked from the plasmid DNA plates when the control plate was shown empty or with negligible colony growth when compared to the plasmid plates. The colonies were grown in 10mL 2xTY with the appropriate antibiotic overnight (180 rpm, 37°C). Glycerol stocks were made from the overnight (500µl culture, 500µl autoclave sterilised 75% glycerol) and stored at -80ºC.

### 8.7.4 Isolation of plasmid DNA (miniprep)

*E. coli* cells were harvested from overnight cultures (10 mL, 8,000 rpm, 4 ˚C, 10 min) and plasmid DNA isolated/purified using the E.Z.N.A. miniprep kit II as per the manufacturer's instructions. The relevant plasmid was eluted from the purification column with sterile deionised water (50 µL) and stored at –20 °C until required.

### 8.7.5 Analytical digestion of purified DNA

Purified DNA was digested with appropriate restriction enzymes and relevant NEB buffer, as discussed by the manufacturer. See Table M21for the assay mixture.

### 8.7.6 DNA-vector ligation

Purified vector and insert DNA were digested (see Method 8.7.5) with two restriction enzymes giving rise to compatible ends (see Method 8.7.5). The ligation mix as described in Table M22 and allowed to incubate for 15 minutes at room temperature. The sample can then be used directly in cell transformation (see Method 8.7.3).

### 8.7.7 DNA agarose gel electrophoresis

1% w/v agarose was added to 1X TAE buffer (Table M24). The solution was incubated under agitation for 5 min followed by heating in a microwave to melt the agarose. The solution was allowed to cool and 0.05% (v/v) Nancy-520 nucleotide dye was added to the agarose solution, which was then cast in a gel mould with a comb and allowed to solidify. 50 μL of DNA sample and 5 μL DNA ladder were loaded onto the gel. The gel was run in an electrophoresis apparatus in 1X TAE buffer at 100v for 45 min. The gel was visualized in Syngene GeminyBV under the transilluminator UV light.

#### 8.7.7.1 RNA Bleach Agarose Gel Electrophoresis

The same methodology described in 9.4.6 applies to the analysis of RNA by agarose gel, with slight modifications. 2 % w/v agarose was added to TAE followed by heating in the microwave. After the solution was cooled, 0.5 mL (1%) sodium hypochlorite was added to the gel, together with 0.05% (v/v) Nancy-520 nucleotide dye. 250 ng of RNA (quantified by Nanodrop) were loaded in each lane. The gel was run at 60 V for 120 minutes.

#### 8.7.7.2 RNA Formaldehyde Agarose Gel Electrophoresis

1 g of agarose was mixed with 20 mL 5X MOPS buffer (Table M30) and 72 mL water before dissolving by heating. In a fume hood, 17.6 mL of formaldehyde was added to the cooled solution and mixed thoroughly. and 0.05% (v/v) Nancy-520 nucleotide dye was added to the agarose solution, which was then cast in a gel mould with a comb and allowed to solidify. 250 ng of RNA (quantified by Nanodrop) were loaded in each lane. The gel was run at 60 V for 120 minutes.

## 8.8 Expression and purification of RSAM enzymes

### 8.8.1 Large scale expression

2xTY (Table M4) media was sterilized in a fermenter unit (New Brunswick BioFlo® 310). Ampicillin (Table M31) and antifoam (0.1 mg/L) (Antifoam 204) were added, followed by inoculation with 50 mL of a small scale culture (see Method 8.7.1). The culture was incubated at 37° C, with variable agitation (50-250 rpm). Agitation was controlled by the level of $O_2$ present in the media, with the

optimal $O_2$ values being set at 40% of the available at full aeration. Cell growth was followed by optical density at 600 nm. When $OD_{600}$=0.8, expression of the relevant protein then induced by addition of filter sterilized 20% L-arabinose (50mL/L). The expression of the protein was carried out at 27° C for six hours. Cells were harvested by centrifugation (6000 rpm, 4° C, 20 min) and cell paste was stored at -80° C.

### 8.8.2 Auto-induction large scale expression

Autoinduction media (Table M5) (4x 1.25L) was sterilized by autoclaving in 4L flasks. A small scale culture (50 mL) of the desired strain was incubated at 37ºC overnight. 12.5 mL of a filter-sterilized 60% (v/v) glycerol solution, 6.25 mL of a filter sterilized 10% (w/v) glucose solution and 31.25 mL of a 8% (w/v) lactose solution were added to each of the four flasks. Additionally, streptomycin (Table M31) was also added. A small scale culture (see Method 8.7.1) was inoculated into the four flasks and the culture was grown at 27ºC overnight under constant agitation (180 rpm). The following morning the cells were harvested by centrifugation (8000 rpm, 4ºC, 15 min) and the cell paste was stored at -80ºC until its use for purification.

### 8.8.3 General purification

The purification steps were carried out in an anaerobic glovebox (Belle Technologies MR1 anaerobic glove box). Cell paste was thawed and resuspended in buffer A (Table M6) together with lysozyme (10 mg) and protease inhibitor tablets as per the manufacturer's instructions (cOmplete™ EDTA-free protease inhibitor cocktail tablets by Roche®, 1 tablet/50 mL). The mixture was incubated under agitation until completely resuspended. The cells were lysed by sonication (4 x 10 min, 1 second pulse, 20 W). The cell debris was removed by centrifugation (18,000 rpm, 4° C, 30 min) and the clear supernatant applied to a nickel charged column (Chelating sepharose FF, pre-equilibrated with buffer A). The column was washed with five column volumes of buffer A, to elute proteins from the host organism. A step gradient (100 mL to 60% buffer B) was introduced to elute the samples bound to the affinity column. Samples (100 μL) from collected fractions (15 mL) were taken for SDS-PAGE analysis. The darkest fractions, likely to indicate the presence of [4Fe-4S] containing proteins, were pooled together and concentrated using an Amicon concentrator (Merck Millipore®) to 20 mL. The concentrated sample was applied to a small (50 mL) S-75 gel filtration column, pre-equilibrated with buffer C, for buffer exchange. The fractions collected were chemically reconstituted (see Method 8.9.1) and concentrated once more using an

Amicon, to a final volume of 3 mL. The 3 mL reconstituted sample was applied onto a size exclusion S-75 column (300 mL) where the separation between iron-sulfide:protein aggregates and pure protein was achieved. The collected fractions of golden brown protein were pooled together, reconstituted and concentrated to 1 mM of protein concentration. The samples were then aliquoted and frozen at -80 °C.

### 8.8.4 Cfr purification of RSAM proteins

The purification steps were carried out inside an anaerobic glove box. The cell paste was resuspended in Buffer A (3mL/g) with the addition of lysozyme (10mg/100mL) and EDTA-free protease tablets (1tablet/50mL), and allowed to incubate in a cold bath for 1h, under agitation. The cells were then lysed by sonication (10 min, 1 second pulse, 20W). Cell debris was removed by centrifugation (14000 rpm, 4 °C, 30 min) and the supernatant collected was then loaded onto a Ni affinity column pre-equilibrated with buffer A (Table M7). The column was then further washed with buffer A to remove unbound protein. After the complete removal of the unbound proteins, a gradient step with buffer B (Table M7 Purification Buffers (Cfr).) was used to elute the desired protein (50mL; 60% Buffer B), and the protein was collected in fractions from the AKTA purifier.

The protein was eluted in 5 mL fractions into 15 mL flasks, where 10 mL of buffer C (Table M7) without DTT were already present, in an attempt to dilute the amount of imidazole present. A small sample of each fraction was taken for SDS-PAGE analysis and protein quantification with the method of Bradford (see Method 8.8.5). The fractions were then pooled together and concentrated to 30 mL.

During the concentration step, protein precipitation was observed. The protein was transferred into 50 mL Falcon tubes and spun-down (6000 rpm, 10 min) to remove the insoluble protein. The soluble fraction was subsequently loaded to a small S-75 gel filtration column for buffer exchange into buffer C, for storage.

Samples were taken again from each fraction and the purest fractions (as detected by SDS-PAGE) were chemically reconstituted (see Method 8.9.2) and concentrated once more using an Amicon, to a final volume of 3 mL. The 3 mL reconstituted sample was applied onto a size exclusion S-75 column (300 mL) where the separation between iron-sulfide:protein aggregates and pure protein was achieved. The collected fractions of golden brown protein were pooled together, reconstituted and concentrated to 1 mM of protein concentration. The samples were then aliquoted and frozen at -80 °C.

### 8.8.5 Protein estimation by Bradford quantification

5 mg of BSA were dissolved in 1 mL water. This stock solution was used to make the dilutions shown in Table M32.

**Table M32 BSA dilutions for protein quantification by the Bradford method.**

| Final concentration (mg/mL) | Stock BSA (μL) | Water (μL) |
|---|---|---|
| 1 | 200 | 800 |
| 0.5 | 100 | 900 |
| 0.25 | 50 | 950 |
| 0.125 | 25 | 975 |
| 0.05 | 10 | 990 |
| 0.025 | 5 | 995 |
| 0.0125 | 2.5 | 997.5 |

Protein samples were diluted 1/100 in water. 20 μL of the protein sample or a standard were added to a 1 mL cuvette, to which was then added 1mL Bradford reagent. The absorbance was measured at 595 nm and the standard values were used to plot a calibration curve which was used to calculate the protein concentration.

### 8.8.6 15% SDS-PAGE denaturing gel

The resolving gel mix (Table M10) was poured to plates and the surface covered in isopropanol (50%) to prevent the formation of bubbles. Once the gel had solidified, the isopropanol was removed and stacking gel (

**Table M11**) was applied, together with a Teflon comb. After the stacking gel was set, the comb was removed and the gel washed with distilled water. A 1:1 mix of protein sample and loading buffer (Table M12) (total volume 40 μL) was mixed and heated at 95 °C for 5 min. 10uL of each sample was applied to a well, along with a protein ladder. The separation was carried out in a BioRad electrophoresis apparatus, running at 180V, 45 min in SDS-PAGE running buffer (

**Table M13**). After the run was complete, the gels were stained using the SDS-staining solution (

Table **M14**) and heated in the microwave for ~1min. The gels were destained overnight with the destaining solution (

Table **M15**) prior to visualization in the SynGene GeminyBV.

## 8.9 Chemical reconstitution and iron content analysis of RSAM enzymes

### 8.9.1 Chemical reconstitution of dual [4Fe-4S] cluster containing proteins

The pooled protein fraction was allowed to incubate under agitation for 90 min. For the first chemical reconstitution, following Ni-NTA affinity chromatography, 10 equivalents of $FeCl_3$, in respect to the protein concentration, were added dropwise over 30 min. The solution was then stirred for a further 60 min before the dropwise addition of 10 equivalents of $Na_2S.9H_2O$ over 60 min. The solution is further stirred for 120 min before being centrifuged (6,000 rpm, 10 min) to remove excess iron sulfide. For the reconstitution that follows the size exclusion chromatography,

only 5 molar equivalents of $FeCl_3$ and $Na_2S$ are added, with the general protocol remaining equivalent.

### 8.9.2 Chemical reconstitution of single [4Fe-4S] cluster containing proteins

The methodology for chemical reconstitution is analogous to the reported in section 9.6.1, with the only change being the molar equivalent amounts of $FeCl_3$/$Na_2S$ added in each of the reconstitution steps: 5 molar equivalents of each are added in the first reconstitution step, following Ni-NTA affinity purification and buffer exchange into an imidazole free buffer; and 3 Meq of $FeCl_3$/$Na_2S$ are added following size exclusion chromatography.

### 8.9.3 Iron quantification by the modified method of Fish

$FeSO_4.9H_2O$ (5mg) was dissolved in 1 mL water and diluted 1/100. The standards shown in Table M33 were made with the dilute solution, to calculate a standard curve.

**Table M33 $Fe^{2+}$ dilutions for protein quantification by the Fish method.**

| $Fe^{2=}$stock solution (180 μM)/μL | $H_2O$ /μL | Nmol $Fe^{2=}$ |
|---|---|---|
| 400 | 600 | 72 |
| 350 | 650 | 63 |
| 300 | 700 | 54 |
| 250 | 750 | 45 |
| 200 | 800 | 36 |
| 150 | 850 | 27 |
| 100 | 900 | 18 |

| $Fe^{2=}$stock solution (180 μM)/μL | $H_2O$ /μL | Nmol $Fe^{2=}$ |
|---|---|---|
| 50 | 950 | 9 |
| 30 | 970 | 5 |
| 10 | 990 | 2 |
| 0 | 1000 | 0 |

Protein samples were diluted 1/100. To each standard and sample, 500 μL of Reagent A (Table M16) was added. The samples were then heated at 60'C for 2h before cooling to room temperature. 100 μL of reagent B (Table M17) was added and the samples were allowed to stand for 10 min before measuring the absorbance at 562 nm. The standard values were used to plot a calibration curve which was used to calculate the iron content of the proteins analysed.

## 8.10 Activity assay of HydG proteins

All manipulations of HydG occurred inside an anaerobic glove box to ensure anaerobic conditions. Previously reconstituted HydG samples (50 μM final concentration) were incubated with SAM (500 μM) and L-tyrosine or a substrate analogue (4000 μM). Dithionite was added (1000 μM) to initiate the reaction. One control was made without the addition of L-tyrosine, to account for uncoupled turnover of SAM. The reaction mixtures were incubated at 37 °C, 1h and the reaction stopped by the addition of 20% perchloric acid (15 μL). Protein aggregate formed due to the addition of perchloric acid was removed by pelleting with centrifugation (13,000 rpm, 10 min).

The supernatant was collected and 40 μL were injected into a C18 Kinetex 5U 100A reverse phase column (4.6 x 150 mm, 5 μm, 130 Å, 0.8 mL/min) and analysed by HPLC. DOA formation was followed by measuring the changes in $Abs_{254nm}$ and to measure *p*-cresol formation fluorescence was used, $\lambda_{em}$=274 nm; $\lambda_{ex}$=312 nm. The fluorescence scanner gain was set at 3, and the sensitivity at 3, to give a total of 16x magnification of signal. Standards of known concentrations (100 μM, 250 μM, 500 μM) of DOA and *p*-cresol were created to obtain a standard curve with which to calculate the amount of product formed.

The method used for elution of the samples was as follows. (A – 2mM ammonium formate, pH3.0; B – methanol):

20% B – 0 min

20% B – 5 min

65% B – 25 min

100% B – 27 min

100% B – 37 min

The data was processed using Unipoint software.

### 8.10.1 Cyanide fluorescence conversion

From the supernatant used for HPLC analysis of the formation of DOA and *p*-cresol, 10 μL was mixed with 45 μL methanol, 5 μL 20% $NH_3$, 15 μL taurine (83 mM) and 25 μL NDA (10 mM in methanol). The mixture was incubated for 30 minutes, to allow extensive formation of the cyanide-bound fluorescent product in solution. The reaction was stopped by addition of 20% perchloric acid (15 μL), immediately followed by neutralization of the acid by addition of 30 μL 1.0 M phosphate buffer, pH 7.4 and 20 μL 1.0 M NaOH. 40 μL of this sample were injected into a C18 Kinetex 5U 100A reverse phase column (4.6 x 150 mm, 5 μm, 130 Å, 0.8 mL/min) and analysed by HPLC. Presence of cyanide was detected by fluorescence $\lambda_{em}$=418 nm; $\lambda_{ex}$=454 nm, with the gain of the fluorimeter being set at 1 and the sensitivity at 2, for a total of 32x magnification of signal.

A set of standards of potassium cyanide were done to use as a calibration curve of the amount of cyanide formed.

**Table M34 KCN dilutions of the standard curve used for cyanide quantification**

| Cyanide stock | | Cyanide final | | Vbuffer (μL) |
|---|---|---|---|---|
| ci (μM) | vi (μL) | cf (μM) | vf (μL) | |
| 50 | 4.5 | **1.25** | 180 | 138.5 |
| 50 | 9 | **2.5** | 180 | 134 |
| 50 | 18 | **5** | 180 | 125 |
| 50 | 36 | **10** | 180 | 107 |
| 50 | 90 | **25** | 180 | 53 |

| Cyanide stock | | Cyanide final | | Vbuffer (μL) |
|---|---|---|---|---|
| ci (μM) | vi (μL) | cf (μM) | vf (μL) | |
| 500 | 18 | **50** | 180 | 125 |
| 500 | 36 | **100** | 180 | 107 |
| 500 | 72 | **200** | 180 | 71 |

The method used for elution of the samples was as follows. (A – 2mM ammonium formate, pH3.0; B – methanol):

40% B – 0 min

40% B – 5 min

100% B – 15 min

100% B – 25 min

60% B - 26 min

60% B – 35 min

The data was processed using Unipoint software.

## 8.11 Screening of crystallization conditions with Hampton PEGRx HT and/or Molecular Dimensions PACT Premier commercial broad screens

All manipulations of the proteins occurred inside an anaerobic glove box to ensure anaerobic conditions ($O_2$ < 2 ppm). A reconstituted and concentrated (1.0 mM) sample of the RSAM enzyme was thawed and mixtures with SAM and SAM + substrate were created as shown in Table M35 and Table M36.

Using a 96-well plate, 100 μL of the screening solution were loaded into the reservoir from which 1 μL was pipetted to the wells of a 96-well plate. 1 μL of the protein mixes was then pipetted to the same well, forming a 1:1 mixture of protein and precipitant solution. The 96-well plate was sealed with sealing film and the plate was incubated at constant temperature (20 °C). The higher concentration of precipitant in the reservoir leads to water loss inside the drop, due to osmosis, which increases localized concentration of both the sample and precipitant in the drops, encouraging crystal growth.

**Table M35 Protein mix for co-crystallization with SAM.**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| Protein | 1.0 | 0.875 | 131.25 |
| SAM | 80 | 10 | 18.75 |
| Total | | | 150 |

**Table M36** Protein mix for co-crystallization with SAM and the enzyme-specific substrate

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| Protein | 1.0 | 0.725 | 108.75 |
| SAM | 80 | 10 | 18.75 |
| Substrate | 20 | 3 | 22.5 |
| Total | | | 150 |

### 8.11.1 Manual reproducibility of the identified crystallization conditions

After the identification of crystallization conditions by using the commercial screen, the same conditions were replicated with purchased materials and kept inside an anaerobic glove box, to use in a manner analogous to the described in Method 8.11.

## 8.12 Reproducibility of crystallization conditions in the 24-well format and 96-well format

Crystallization in 24-well format is similar to the methodology described for crystallization in the 96-well format (Method 8.11). The crystallization stock solutions did not vary (Table M35 and

Table **M36**), but the crystallization conditions were manually created following the manufacturer's guidelines. All crystallization conditions were kept inside the anaerobic glovebox. 500 μL of crystallization solution was added to the reservoir and a mixture of 1 μL protein and 1μL well solution were used to form the crystal drops in the cover slips.

## 8.13 Microseeding using seed beads

To prepare the seed stock, a freshly grown crystal was removed from its drop and mixed with fresh crystallization condition (100 μL). A small bead, purchased from Molecular Dynamics, was added to the Eppendorf, which aids in the crushing and homogeneous mixing of crystal particles across the seed stock. Serial dilutions of the seed stock are created, and 0.1 μL of the different dilutions are added to previously prepared drops of protein:precipitant.

## 8.14 Crystallization of concentrated *Ti*HydG with L-tyrosine

The crystallization experiment was carried out following the same general procedure detailed in Method 8.11, but changing the overall concentration of the constituents in the protein mix, as shown in Table M37.

**Table M37 Protein mix for co-crystallization of high concentration of *Ti*HydG with SAM and L-tyrosine**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Ti*HydG | 3.0 | 0.88 | 44 |
| SAM | 80 | 10 | 18.8 |

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| L-tyrosine | 20 | 6/9 | 45/67.5 |
| Buffer | | | 42.25/19.75 |
| Total | | | 150 |

## 8.15 Crystallization of *Ti*HydG, under different concentrations of protein and SAM

The co-crystallization studies with lower concentrations of *Ti*HydG or SAM were carried out following the same general procedure depicted in Method 8.11, but changing the overall concentration of the protein mix, as shown in Table M38. For the crystallization, the condition F1 from PEGRx HT commercial screen (see Table M18) was used.

**Table M38 Protein mix for co-crystallization with lowered concentration of SAM.**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Ti*HydG | 1.0 | 0.875 | 131.25 |
| SAM A | 80 | 3 | 5.63 |
| Buffer | | | 13.12 |
| SAM B | 80 | 1 | 1.88 |
| Buffer | | | 16.82 |
| SAM C | 80 | 0.5 | 0.94 |
| Buffer | | | 17.76 |
| Total Volume (μL) | | | 150 |

**Table M39 Protein mix for co-crystallization with lower concentration of *Ti*HydG**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Ti*HydG (A) | 1.0 | 0.875 | 131.25 |
| SAM | 80 | 3 | 5.63 |
| Buffer | | | 0 |
| *Ti*HydG B | 1.0 | 0.44 | 66 |
| Buffer | | | 65.25 |
| *Ti*HydG C | 1.0 | 0.29 | 43.5 |
| Buffer | | | 87.75 |
| Total Volume (μL) | | | 150 |

## 8.16 Crystallization of 0.3 mM *Ti*HydG with 12 mM L-tyrosine

The co-crystallization studies with forty-fold excess of L-tyrosine followed the same general procedure depicted in section Method 8.11, but changing the overall concentration of the protein mix Table M40. The condition F1 from PEGRx HT commercial screen (see Table M18) was used for the crystallization.

**Table M40 Protein mix for co-crystallization of *Ti*HydG with SAM and L-tyrosine**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Ti*HydG | 1.0 | 0.29 | 43.5 |
| SAM | 80 | 3 | 5.63 |
| L-tyrosine | 20 | 12 | 90.0 |
| Buffer | | | 10.87 |
| Total | | | 150 |

## 8.17 Primer extension PCR

4 sets of primers are required for this PCR technique, two of them (A and D) complementary to the terminal sequence of the protein and two (B and C) with a non-complementary branch where the mutation is located. In order to design the two mutations (H285A and H285E) the following primers were ordered:

A – 5'- AGG AAT TAA CCA TGG GCA GCA GCC -3'

D – 5'- CTC CTG TTA CTC GAG TTA AAA GTA -3'

His265Ala mutant:

B – 5'- AAT GGT GCG CGG ACC-3'

C – 5'– GGT CCG GCG ACC ATT –3'

His265Glu mutant:

B – 5'- AAT GGT GAG CGG ACC-3'

C – 5'– GGT CCG GAG ACC ATT –3'

**PCR reaction:**

The PCR reaction mix was placed in an Eppendorf Mastercycler under the following protocol:

Initial denaturation – 98ºC, 30 sec

30x cycles:

Denaturation – 98ºC, 10 sec

Annealing – 57ºC, 30 sec

Extension – 72ºC, 45 sec

Final extension – 72ºC, 10 min

After the protocol finished, the samples were analysed by agarose gel.

## 8.18 Circular dichroism analysis of *Ti*HydG mutants

*Ti*HydG His265Glu, *Ti*HydG His265Ala and *Ti*HydG WT were purified and samples were concentrated to 1 mM and frozen at -80 °C before chemical reconstitution. Different samples were expressed, purified and reconstituted as discussed in section 8.8. Both reconstituted and non-reconstituted samples for all three enzymes were desalted using a disposable PD-10 column (GE Healthcare), into 10 mM phosphate buffer, pH 7.4. Samples were diluted in the same buffer to a concentration of 0.2 mg/mL and placed in high quality 1 mm quartz cuvette. To minimize potential errors, the same cuvette was used to obtain the spectra of both the samples and the buffer for substraction.

The quartz cuvette was sealed with a cap and parafilm to minimize oxygen contamination that could degrade the [4Fe-4S] clusters. Only one sample was removed from the anaerobic glove box at a time for measurement. The samples were analysed in a Jasco J-710CD spectrophotometer, under a constant flux of nitrogen (20 psi).

## 8.19 *Ti*HydG crystallization with L-amino acids and reconstitution additives

A similar experiment to that described in Method 8.14 was carried out, with a concentrated sample of *Ti*HydG being used to allow the addition of L-amino acids without diluting the overall mix of the reaction.

**Table M41 Protein mix for co-crystallization of high concentration of *Ti*HydG with SAM and L-tyrosine**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Ti*HydG | 3.0 | 0.88 | 44 |
| SAM | 80 | 10 | 18.8 |
| Additive | Variable | Variable | Variable |
| Buffer | | | Variable |
| Total | | | 150 |

The different amino acids and additives were mixed using the general table, leading to 40 different conditions tested. Stock solutions of $FeCl_2.SO_4$ (50 mM), $Na_2S.9H_2O$ (30 mM), glycine (40 mM), L-methionine (40 mM) and L-cysteine (40 mM) were made in Buffer C and mixed to a final concentration shown in Table M42.

**Table M42 Table of additives for co-crystallization with *Ti*HydG. Concentration values are final concentration of samples**

| **Additives** | **Amino acid ligands** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Glycine | | | L-methionine | | | L-cysteine | | | |
| | 0.5 mM | 1.0 mM | 5.0 mM | 0.5 mM | 1.0 mM | 5.0 mM | 0.5 mM | 1.0 mM | 5.0 mM | None |
| $FeCl_2$ (5.0 mM) | | | | | | | | | | |
| $Na_2S$ (3.0 mM) | | | | | | | | | | |
| $FeCl_2$/$Na_2S$ (5.0 mM/3.0 mM) | | | | | | | | | | |
| None | | | | | | | | | | Positive control |

## 8.20 Soaking of pre-grown crystals with potential ligands

Pre-grown crystals using the methodologies described in Method 8.11 and 8.13 were transferred into a fresh drop with the crystallization condition mixed with the soaking ligand at relevant concentrations. The well was sealed and the crystal was allowed to incubate in a saturating mixture of the ligand for a set amount of time (between 2h and 24h), before being flash frozen in liquid nitrogen and stored for further analysis.

## 8.21 Synthesis of iron-bound α-iminocarboxylic acids

Stock solutions of glyoxylate (1000 mM), $FeCl_2$ (200 mM) and ammonia (2000 mM) were made in water, in strictly anaerobic conditions, using an anaerobic glove box ($O_2$ < 2 ppm). 450 μL of glyoxylate were mixed with varying concentrations of $FeCl_2$ (150, 187.5, 250 μL, to a final concentration of 30, 37.5 and 45 mM) and ammonia (60, 56.25, 52.50 μL, to a final concentration of 120, 112.5 and 105 mM), to obtain the $X_{fe}$ ratio of 0.2, 0.25 and 0.3, respectively. Samples were

incubated, under agitation, for 1h, and any precipitate formed was removed by centrifugation (6,000 rpm, 10 min). The UV- spectra of the samples was measured using the Lamba 2 spectrophotometer. Controls where only two of the components were mixed were carried out in parallel.

### 8.21.1 Synthesis of iron-bound α-iminocarboxylic acids, with lower concentration of substrates

The same protocol as described in 8.21 was carried out using diluted substrates as shown in Table M43.

**Table M43 Concentrations used for the formation of Fe-bound dehydroglycine. $X_{Fe}$ is calculated as [Fe]/([Fe]+[$NH_3$]).**

| [Glyoxylate] mM | [$FeCl_2$] mM | [$NH_3$] mM | $X_{Fe}$ |
|---|---|---|---|
| 45.00 | 1.50 | 13.50 | 0.10 |
| 45.00 | 2.25 | 12.75 | 0.15 |
| 45.00 | 3.00 | 12.00 | 0.20 |
| 45.00 | 3.75 | 11.25 | 0.25 |
| 45.00 | 4.5 | 10.50 | 0.30 |

### 8.21.2 *Ti*HydG Activity Assay in the Presence of α-iminocarboxylic acid

50 μM of reconstituted *Ti*HydG was incubated with a diluted sample of the Fe-DHG complex, formed as described in section 9.18.1, using the concentration of iron present in the mixture to estimate a 10-fold (500 μM) of iron present in relation with the enzyme. The reaction was initiated by the addition of dithionite, 1000 μM, as described in section 9.7. The samples were incubated for 1h, at 37°C, and cyanide formation was estimated following the general cyanide fluorescence protocol, Method 8.10.1.

## 8.22 *Te*LipA2 Crystallization With High Glycerol Content

To increase the concentration of glycerol in the crystallization condition, two glycerol stock solutions were made: one with Buffer C, to mix with the protein sample; and one with the crystallization solution, to add to the well solution.

As the protein solution already contains 10% glycerol in the purification Buffer C, only a small volume of glycerol stock was needed to increase the concentration in the protein mix:

**Table M44 Protein mix to increase the glycerol concentration**

| Solution | Stock Conc | Final Conc | Volume Added/μL |
|---|---|---|---|
| LipA | 1.0 mM | 0.65 mM | 97.2 |
| SAM | 80 mM | 10 mM | 18.75 |
| Glycerol | 75% | (10)+3% | 6 |
| | | (10)+8% | 16 |
| | | (10)+12.5% | 25 |
| | | (10)+17.0% | 34 |
| Final volume (mL) | | | 150 |

To create the desired crystallization solutions enriched with glycerol, the values in Table M44 were used.

**Table M45 Protein mix for co-crystallization *Te*LipA2 with SAM and MTAN**

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/μL |
|---|---|---|---|
| *Te*LipA 2 | 1.0 | 0.78 | 116.3 |
| SAM | 80 | 10 | 18.8 |
| MTAN | 0.5 | 0.05 | 15 |
| Total | | | 150 |

## 8.23 *Te*LipA2 Chemical Reconstitution With Different Reducing Agents

The protein fraction was initially reconstituted as described in section 9.6.1 with 10 Meq. $FeCl_3/Na_2S$, followed by size exclusion chromatography, with a modified Buffer C* (25 mM Hepes, pH 7.4, 500 mM NaCl, 10% glycerol, 0 mM DTT), to clear the excess Fe-S aggregate formed. The golden fractions corresponding to pure *Te*LipA2 were pooled together and divided equally into five tubes to test the extent of the reconstitution with different reducing agents.

1.0 mM of DTT (5.0 mM in the control), 1.0 mM BME or 1.0 mM TCEP were added slowly, allowing the protein to incorporate the reducing agent under agitation for 30 minutes, prior to the addition of $FeCl_3$. To all samples but one, the control, 5 Meq of $FeCl_3$ and $Na_2S$ were added slowly under agitation. Only 2.5 Meq of $FeCl_3$ and $Na_2S$ were added to the control.

## 8.24 *Te*LipA2 Crystallization in the Presence of MTAN

The general protocol, described in section 9.9, for crystallization of LipA was followed, with the only modification being the addition of MTAN to the protein mix prior to setting up the crystal tray. The protein mix solution was prepared as described in Table M46.

**Table M46** Protein mix for co-crystallization *Te*LipA2 with SAM and MTAN

| Solution | Stock Conc/mM | Final Conc/mM | Volume Added/µL |
|---|---|---|---|
| *Te*LipA 2 | 1.0 | 0.78 | 116.3 |
| SAM | 80 | 10 | 18.8 |
| MTAN | 0.5 | 0.05 | 15 |
| Total | | | 150 |

## 8.25 RNA Extraction by Guanidine thiocyanate Chloroform:Phenol

*E.coli* cells (BL21 and BL21 Star) transformed with the pET16b plasmid containing the rRNA gene of interest were grown overnight at 37ºC in 1L flasks with 250mL 2X TY with ampicillin (100 µg/mL). The cells were then centrifuged at 6000 g, for 30 minutes. The supernatant was discarded and the cells resuspended in GNTC buffer (Table M1), 10 mL/g cell tissue. To this solution 0.1 V 2.0

M sodium acetate, pH 4.0; 1 V acid phenol (equilibrated in water); 0.3 V chloroform are added stepwise, with vortexing between each step. After incubation on ice, two forms should be visible: organic and aqueous. The samples were centrifuged 10,000 rpm, for 25-minutes at 4° C, to guarantee good phase separation. The aqueous layer (on top) was carefully removed onto a new tube, where 1.1 V isopropanol were added. The samples were incubated overnight at -20° C. Next day the samples were centrifuged at 10,000 rpm, 25 min 4° C and the supernatant discarded. 1.1 V of 70% ethanol were added and the RNA pellet resuspended carefully. Samples were once again centrifuged at 10,000 rpm, 25 min 4° C and the supernatant discarded. The RNA sample was resuspended in 500-1000 μL nuclease free water and stored at -20° C.

## 8.26 RNA purification by Streptactin Affinity Chromatography

A small volume (50 μL) of RNA sample, purified as described in 8.25, was injected into a Streptavidin affinity column, previously equilibrated with 10 mM phosphate buffer, pH 7.0. The column was washed with 3 CV of 10 mM phosphate buffer, pH 7.0 and the elution of non-binding RNA occurs. A gradient (50 minutes, 100%, 0.5 mL/min) of 10 mM phosphate buffer, pH 7.0 with 100 mM $MgCl_2$ was used to release the bound RNA. The column was washed with 10 mM phosphate buffer, pH 7.0, 2.5 mM dethiobiotin to remove all RNA fragments still bound to the column.

## 8.27 RNA Analysis by HPLC

50 μL of RNA sample (20 ng/μL) was injected into a Vydac® 4.6 x250 mm, 5 μm C4 column for separation by RP-HPLC. The buffer system used was a HFIP (200 mM)/TEA (8 mM) mixture in 98% $H_2O$ for the aqueous buffer and 98% MeOH for the organic solvent.

The method used for elution of the samples was as follows:

0% B – 0 min

0% B – 15 min

100% B – 75 min

0% B – 90 min

# Appendix A

## A.1 Plasmid maps for studies in HydG

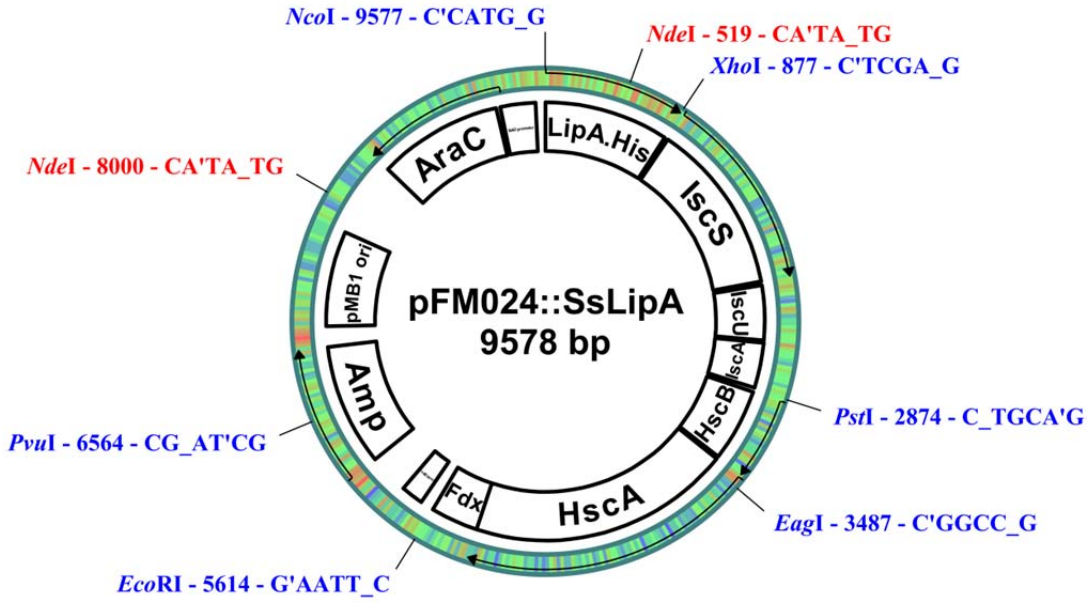



**Figure A1** pBAd/His derived pFM024 plasmid, containing the *isc* operon as well as the gene coding *Solfulobus solfataricus* LipA, tagged C-terminally by a 6His fragment for affinity purification,

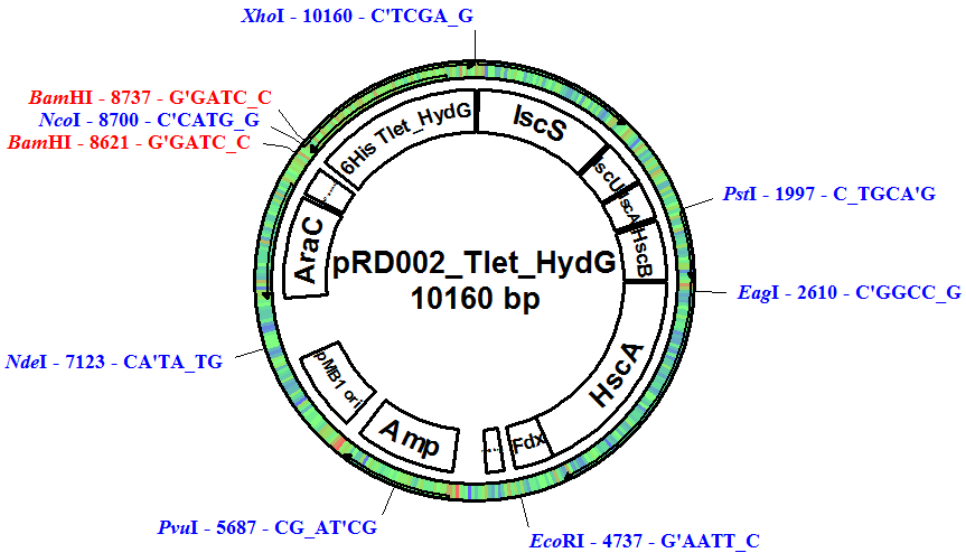



**Figure A2** Modified pFM024 plasmid (pRD002), containing the *Thermotoga lettingae* (*Tl*) HydG gene, tagged N-terminally by a 6His fragment for affinity purification

Appendix A

**Nucleotide sequence for the*TlhydG* gene present in plasmid pRD002**

CACTATAGGGCGAATTGGCGGAAGGCCGTCAAGGCCGCATCCATGGGCAGCAGCCATCACCATCATCACC
ACAGCCAGGATCCGATGTATACCTTTGTGAAAAAAATCCTGGATAGCGAAAGCTTTATCCCGCATGAAGAA
ATTAATCGTCTGCTGAGCCTGACCAAAAACCCGGACAAAAAATACATTGAAGAGATCCTGGAAAAAAGCCT
GAATAAAGAACGTCTGGATCTGAGCGAAACCGCAGCACTGCTGAATGCAGAAAATCCGGAACTGGTTGAA
ATGATTTTTACCGCAGCAAAATCCCTGAAAGAGAAAATTTACGGTAATCGCATTGTTCTGTTTGCACCGCTG
TATATTGGTAACGAATGCATTAACAATTGCAGCTATTGTGGTTTTCGCAGCGAGAATAAACAGGTGATTCG
TAAAACCCTGACCGATGAAGAACTGACCAATGAACTGTTTGCCCTGACCAGCAAAGGTCATAAACGTCTGA
TTGTTGTGTTTGGTGAACATCCGGTTTATAGCCCTGAATATATTGCCAAAACCATCGAAAAAATCTATAGCT
ATCGCAATGGCAATGGTGAAATTCGTCGTGTTAATGTTAATGCCGCACCGCAGACCGTTGAAGGTTATAAA
ACCATTCGTAATGCAGGCATTGGCACCTTTCAGATTTTTCAAGAAACCTATCATCTGCCGACCTACAAAAAA
TACCATAAAGCAGGTCCGAAAAGCAGCTATACCTATCGTCTGTTTGGTCTGGATCGTGCAATGATTGCAGG
TATTGATGATGTTGGTATTGGTGCACTGTTTGGCCTGTATGATTGGAAATTTGAAGTTATGGGTCTGATGTG
TCATACCCGTCATCTGGAAGAACGTTTTGGTGTTGGTCCGCATACCATTAGCTTTCCGCGTATTGAACCGGC
AGTTGGTACACCGCTGACAGAAAAACCGCCTTATGCAGTTAGCGATTATGATTTTAAACGCCTGGTTGCCAT
TATTCGTCTGAGCGTTCCGTATACCGGTCTGATTCTGACCGCACGTGAAAAAGTTCAGCTGCGTGATGAAG
TTATTAAACTGGGTGTTAGCCAGATTGATGCAGGTAGCAGCATTGGTGTGGGTAGCTATGCACAGAAAGA
TCAAGAAATTGTTCGCAAAAGCCAGTTTATTCTGGGTGATACCCGTACCCTGGATGAAGTGATCAAAGAAC
TGGCAATGGAAGGCTATATTCCGAGCTTTTGTACCGCATGTTATCGTGCAGGTCGTACCGGTCAGCATTTTA
TGGAATTTGCAATTCCGGGTTTCGTGAAAGAATTTTGTACCCCGAATGCCCTGCTGACCTTTAAAGAATATC
TGATTGATTACGCCACCAACGAAACCAAACTGATTGGTGAAAAACTGATCGAAAAAGAACTGCTGAAAATT
CCGGATAGCCGTCGTGAAACCGTGCGCAAAATGCTGGTTAATATTGAAAAAGGTGAACGTGATGTGCGCC
TGTAACTCGAGCTGGGCCTCATGGGCCTTCCGCTCACTGCCCGCTTTCCAG

***Tl*HydG amino acid sequence, with the [4Fe-4S] cluster-coordinating cysteines in red**

MGSSHHHHHHSQDPMYTFVKKILDSESFIPHEEINRLLSLTKNPDKKYIEEILEKSLNKERLDLSETAALLNAENPE
LVEMIFTAAKSLKEKIYGNRIVLFAPLYIGNECINNCSYCGFRSENKQVIRKTLTDEELTNELFALTSKGHKRLIVVFG
EHPVYSPEYIAKTIEKIYSYRNGNGEIRRVNVNAAPQTVEGYKTIRNAGIGTFQIFQETYHLPTYKKYHKAGPKSSY
TYRLFGLDRAMIAGIDDVGIGALFGLYDWKFEVMGLMCHTRHLEERFGVGPHTISFPRIEPAVGTPLTEKPPYAV
SDYDFKRLVAIIRLSVPYTGLILTAREKVQLRDEVIKLGVSQIDAGSSIGVGSYAQKDQEIVRKSQFILGDTRTLDEVI
KELAMEGYIPSFCTACYRAGRTGQHFMEFAIPGFVKEFCTPNALLTFKEYLIDYATNETKLIGEKLIEKELLKIPDSR
RETVRKMLVNIEKGERDVRL

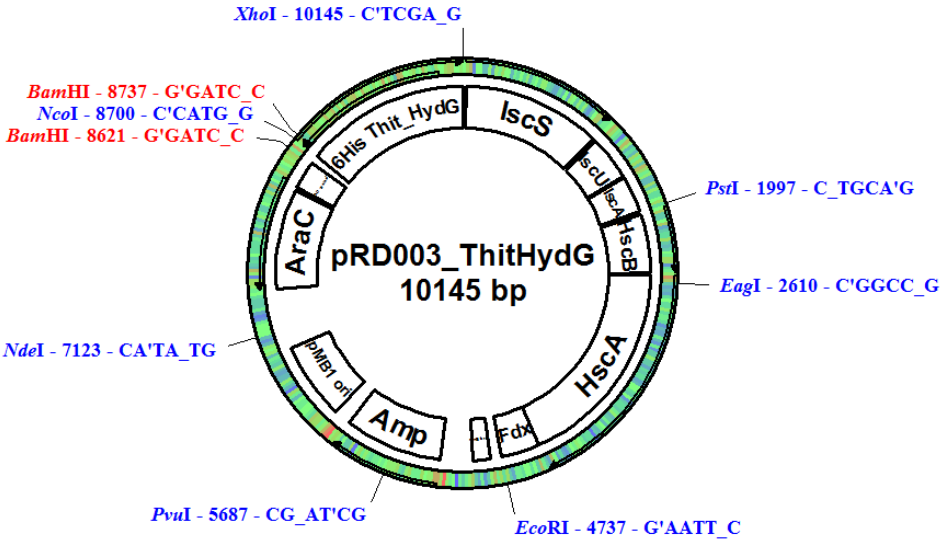



**Figure A3** Plasmid pRD003, containing the *Thermoanaerobacter italicus* (*Ti*) HydG gene, tagged N-terminally by a 6His fragment for affinity purification

**Nucleotide sequence for the *TihydG* gene in plasmid pRD003**

CACTATAGGGCGAATTGGCGGAAGGCCGTCAAGGCCGCATCCATGGGCAGCAGCCATCACCATCATCACC
ACAGCCAGGATCCGATGGTTAAAGAAAAAGCCGATTTCATCAACGACGAAAAAATTCGTCAGGATCTGGA
AAAAGCCAAAAAAGCAACCAGCAAAGATGCCCTGGAAATTATCGAGAAAGCGAAAAATCTGAAAGGCATC
ACACCGGAAGAAGCAGCAGTTCTGCTGAATGTTGAAGATGAAGATCTGCTGAACGAGATGTTTAAAGTTG
CCCGTTATATCAAAGAAGAGATCTACGGTAATCGCATCGTTATTTTTGCACCGCTGTATGTGAGCAATTATT
GCGTGAATAATTGCCGCTATTGCGGTTATCGTCATAGCAATGAACAGCAGCGTAAAAAACTGACAATGGAA
GAAGTTCGTCGCGAAGTTGAAATTCTGGAAGAAATGGGTCATAAACGTCTGGCAGTTGAAGCCGGTGAAG
ATCCGGTTAATTGTCCGATTGATTATATCGTGGATGTGATCAAAACCATCTACGATACCAAACTGAAAAATG
GTAGCATTCGTCGCGTGAATGTTAATATTGCAGCAACCACCGTGGAAAACTACAAAAAACTGAAAAAAGTG
GGCATCGGCACCTATGTTCTGTTTCAAGAAACCTATCATCGTCCGACCTATGAATATATGCATCCGCAGGGT
CCGAAACACGATTATGATTATCATCTGACCGCAATGGATCGTGCAATGGAAGCAGGTATTGATGATGTTGG
TCTGGGTGTTCTGTATGGTCTGTATGATTACAAATATGAAACCGTGGCCATGCTGTATCATGCAAATCATCT
GGAAGAGAAATTTGGTGTTGGTCCGCATACCATTAGCGTTCCGCGTCTGCGTCCGGCACTGAATATTAGCA
TTGATAAATTTCCGTACATCGTGAGCGATAAAGATTTCAAAAAACTGGTTGCCGTTATTCGTATGGCAGTTC
CGTATACCGGCATGATTCTGAGCACCCGTGAAAAACCTAAATTTCGCGAAGAAGTGATTAGCATCGGTATT
AGCCAGATTAGCGCAGGTAGCTGTACCGGTGTTGGTGGTTATCATGAAGAAATTAGCAAAAAAGGTGGTA
GCAAACCGCAGTTTGAAGTGGAAGATAAACGTAGCCCGAACGAAATTCTGCGTACCCTGTGTGAACAGGG
TTATCTGCCGAGCTATTGTACCGCATGTTATCGTATGGGTCGTACCGGTGATCGTTTTATGAGCTTTGCAAA
AAGTGGCCAGATCCATAACTTTTGTCTGCCGAATGCAATCCTGACCTTCAAAGAATTTCTGATCGATTATGG

TGATGAGAAAACCAAAAAAATCGGCGAAAAAGCGATTGCCGTTAACCTGGAAAAAATCCCGAGCCGTACC
GTTCGTGAAGAAACCAAACGTCGTCTGACCCGTATTGAAAATGGTGAACGTGATCTGTACTTTTAACTCGA
GCTGGGCCTCATGGGCCTTCCGCTCACTGCCCGCTTTCCAG

**_Ti_HydG amino acid sequence, with the [4Fe-4S] cluster-coordinating cysteines in red**

MGSSHHHHHHSQDPMVKEKADFINDEKIRQDLEKAKKATSKDALEIIEKAKNLKGITPEEAAVLLNVEDEDLLNE
MFKVARYIKEEIYGNRIVIFAPLYVSNYCVNNCRYCGYRHSNEQQRKKLTMEEVRREVEILEEMGHKRLAVEAGE
DPVNCPIDYIVDVIKTIYDTKLKNGSIRRVNVNIAATTVENYKKLKKVGIGTYVLFQETYHRPTYEYMHPQGPKHD
YDYHLTAMDRAMEAGIDDVGLGVLYGLYDYKYETVAMLYHANHLEEKFGVGPHTISVPRLRPALNISIDKFPYIV
SDKDFKKLVAVIRMAVPYTGMILSTREKPKFREEVISIGISQISAGSCTGVGGYHEEISKKGGSKPQFEVEDKRSPN
EILRTLCEQGYLPSYCTACYRMGRTGDRFMSFAKSGQIHNFCLPNAILTFKEFLIDYGDEKTKKIGEKAIAVNLEKIP
SRTVREETKRRLTRIENGERDLYF

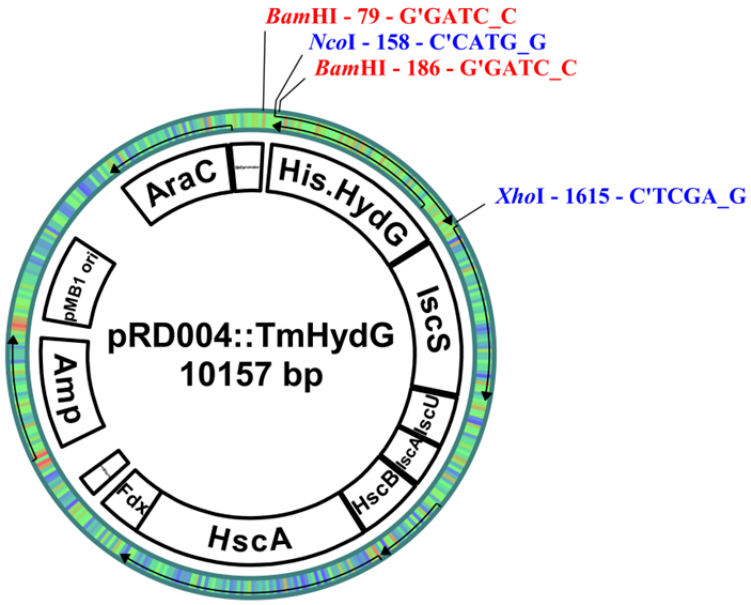



**Figure A4 Plasmid pRD004, containing the _Thermotoga maritima_ (_Tm_) HydG gene, tagged N-terminally by a 6His fragment for affinity purification**

**Nucleotide sequence for the _TmhydG_ gene in plasmid pRD004**

CACTATAGGGCGAATTGAAGGAAGGCCGTCAAGGCCGCATCCATGGGGCATCATCATCACCATCATCAGG
ATCCGATGTGCATGTATGTTTTTGTTAAAGAACGTGTGGAAAGCCGCAGCTTTATTCCGGAAGAAAAAATC
TTTGAGCTGCTGGAAAAAACCAAAAATCCTGATCCGGCACGTGTGCGTGAAATTATTCAGAAAAGCCTGGA

TAAAAATCGCCTGGAACCTGAAGAAACCGCAACCCTGCTGAATGTTGAAGATCCGGAACTGCTGGAAGAA
ATTTTTGAAGCAGCACGTACCCTGAAAGAACGCATTTATGGTAATCGTATTGTTCTGTTTGCACCGCTGTAT
ATTGGCAACGATTGTATTAATGATTGCGTGTATTGCGGTTTTCGCGTGAGCAATAAAGTTGTTGAACGTCG
CACCCTGACCGAAGAACAGCTGAAAGAAGAAGTTAAAGCACTGGTTAGCCAGGGTCATAAACGTCTGATT
GTTGTTTATGGTGAACACCCGAATTACAGCCCTGAATTTATTGCACGTACCATCGATATTGTGTATAACACC
AAATATGGCAACGGTGAAATTCGTCGTGTTAATGTTAATGCAGCACCGCAGACCATTGAAGGCTATAAAAT
CATTAAAAGCGTGGGCATTGGCACCTTTCAGATTTTTCAAGAAACCTATCACCGCGAAACGTACCTGAAACT
GCATCCGCGTGGTCCGAAAAGCAATTATAACTGGCGTCTGTATGGTCTGGATCGTGCAATGATGGCAGGT
ATTGATGATGTTGGTATTGGTGCACTGTTTGGTCTGTATGATTGGAAATTTGAAGTTATGGGCCTGCTGTAT
CATACCATTCATCTGGAAGAACGTTTTGGTGTTGGTCCGCATACCATTAGCTTTCCGCGTATTAAACCGGCA
ATTAATACCCCGTATAGCCAGAAACCGGAACATGTTGTTAGTGATGAGGACTTCAAAAAACTGGTGGCAAT
TATTCGTCTGAGCGTTCCGTATACCGGCATGATTCTGACCGCACGTGAACCGGCAAAACTGCGTGATGAAG
TTATTAAACTGGGTGTGAGCCAGATTGATGCAGGTAGCCGTATTGGCATTGGTGCCTATAGCCATAAAGAA
GATGACGAAGATCGTAAACGTCAGTTTACCCTGGAAGATCCGCGTCCGCTGGATCAGGTTATGCGTAGCCT
GCTGAAAGAGGGTTTTGTTCCGAGCTTTTGTACCGCATGTTATCGTGCAGGTCGTACCGGTGAACATTTTAT
GGAATTTGCAATTCCGGGTTTTGTGAAAAACTTTTGCACCCCGAATGCCCTGTTTACACTGCAAGAATATCT
GTGTGATTATGCCACCGAAGAAACCCGTAAAGTTGGTGAAGAAGTGATTGAACGCGAACTGCAAAAAATG
AATCCGAAAATTCGTGAACGTGTTCGTGAAGGTCTGGAAAAAATCAAACGTGGTGAACGTGATGTGCGCT
TTTAACTCGAGCTGGGCCTCATGGGCCTTCCTTTCACTGCCCGCTTTCCAG

***Tm*HydG amino acid sequence, with the [4Fe-4S] cluster-coordinating cysteines in red**

MGHHHHHHQDPMCMYVFVKERVESRSFIPEEKIFELLEKTKNPDPARVREIIQKSLDKNRLEPEETATLLNVEDP
ELLEEIFEAARTLKERIYGNRIVLFAPLYIGNDCINDCVYCGFRVSNKVVERRTLTEEQLKEEVKALVSQGHKRLIVV
YGEHPNYSPEFIARTIDIVYNTKYGNGEIRRVNVNAAPQTIEGYKIIKSVGIGTFQIFQETYHRETYLKLHPRGPKSN
YNWRLYGLDRAMMAGIDDVGIGALFGLYDWKFEVMGLLYHTIHLEERFGVGPHTISFPRIKPAINTPYSQKPEH
VVSDEDFKKLVAIIRLSVPYTGMILTAREPAKLRDEVIKLGVSQIDAGSRIGIGAYSHKEDDEDRKRQFTLEDPRPL
DQVMRSLLKEGFVPSFCTACYRAGRTGEHFMEFAIPGFVKNFCTPNALFTLQEYLCDYATEETRKVGEEVIEREL
QKMNPKIRERVREGLEKIKRGERDVRF

## A.2 Standard curves for the quantification of products in *Ti*/*Tl* HydG activity assays

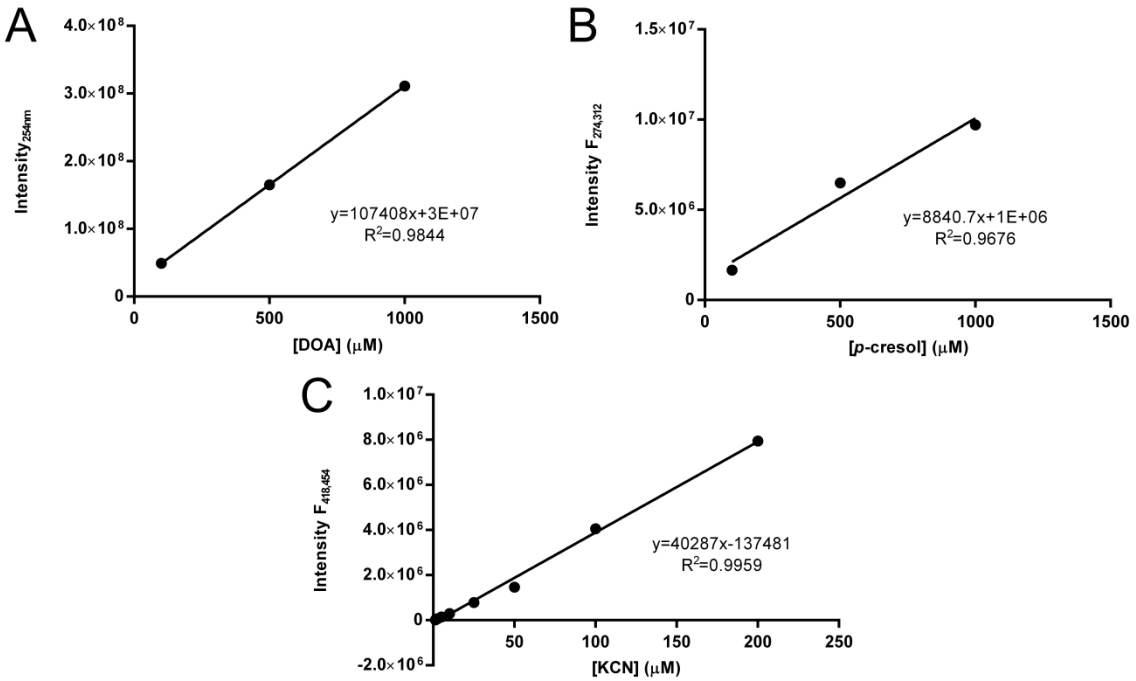



**Figure A5** Standard curves for deoxyadenosine (A), *p*-cresol (B) and cyanide (C).

## A.3 Crystallization data for *Ti*HydG WT crystals

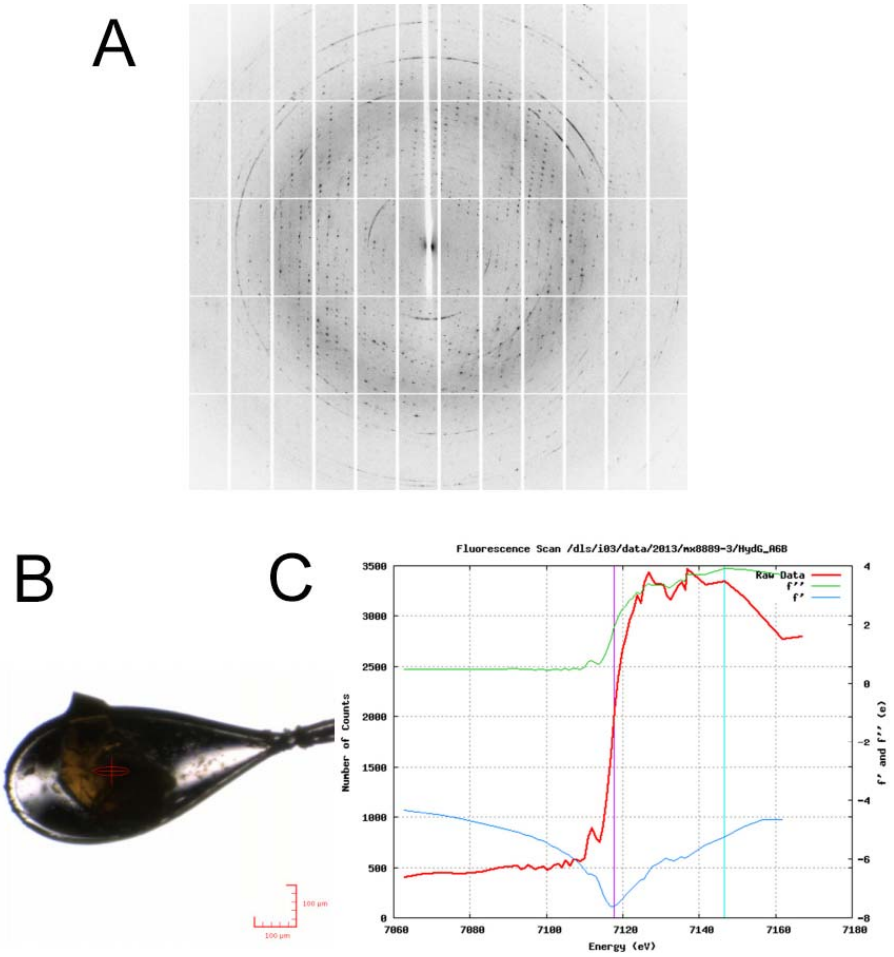



**Figure A6** X-ray diffraction of *Ti*HydG crystal A8. **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 1.9 Å resolution. **B:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM (0.8 mM HydG, 10 mM SAM), directly frozen after formation in condition A8 of Pact Premier screen (0.1M MES monohydrate, pH 6.5; 22% v/v PEG 400). Loop size – 0.5–0.7 mm, Beam size – 80 x 20 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

Appendix A

**Table A1** Data collection statistics for *Ti*HydG crystal A8. Highest resolution shell shown in parenthesis

| | **HydG** | **HydG Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | A8 Pact Premier | |
| **Data collection** | | |
| Space group | P1 | |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 54.10; 55.77; 85.13 | |
| α, β, γ (°) | 90.20; 96.17; 112.61 | |
| | | *Peak* |
| Wavelength (Å) | 0.97625 | 1.73891 |
| No of reflections | 152031 (4487-2072) | 182547 (2928-1044) |
| No of unique reflections | 63203 (2193-796) | 55964 (848-440) |
| Resolution (Å) | 51.42-1.88 (8.41-1.88) | 51.41-1.82 (8.16-1.82) |
| $R_{sym}$ or $R_{merge}$ | 0.033 (0.028-0.396) | 0.040 (0.045-0.323) |
| $I$ / σ$I$ | 14.5 (34.3-2.1) | 18.7 (28.6-2.7) |
| Completeness (%) | 86.1 (98.4-40.3) | 64.5 (81.3-4.1) |
| Redundancy | 2.4 (2.6-2.1) | 1.6 (1.5-1.0) |
| Twinning score | 2.35, no twinning | 2.26, no twinning |

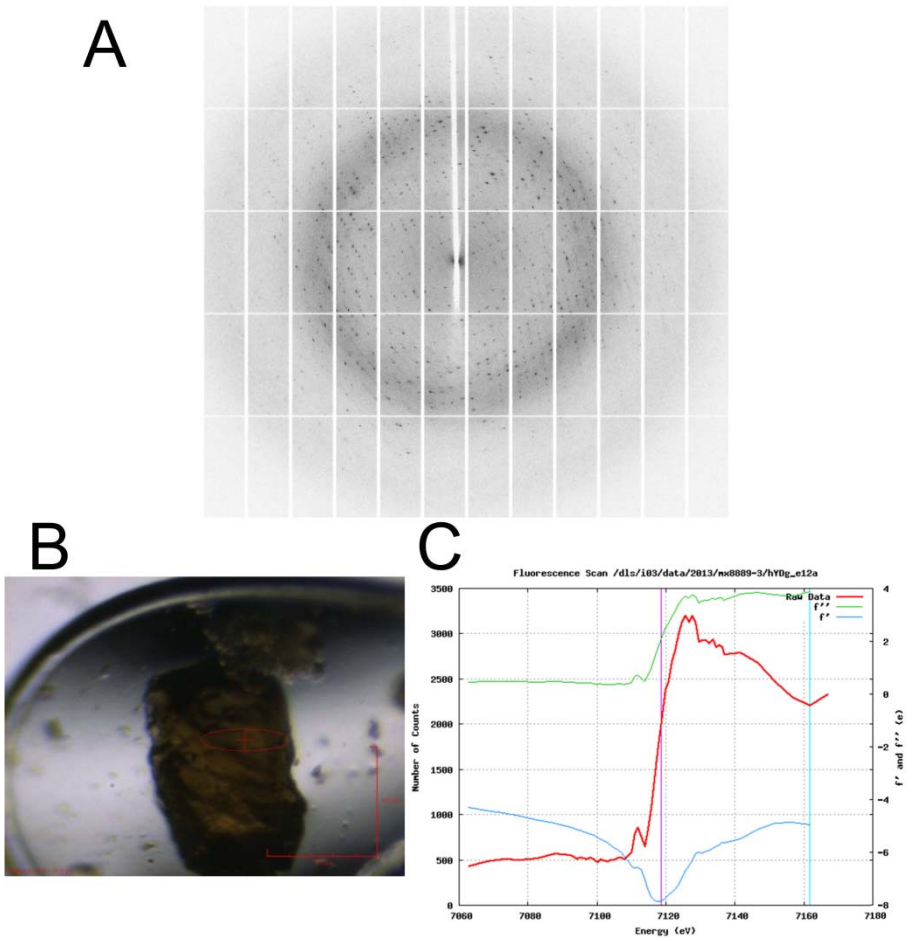



**Figure A7** X-ray diffraction of *Ti*HydG crystal E12. **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 1.8 Å resolution. **B:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM and L-tyrosine (0.8 mM HydG, 10 mM SAM, 3 mM tyrosine), directly frozen after formation in condition E12 of Pact Premier screen (0.2M sodium malonate; 20% w/v PEG 3350). Loop size – 0.3–0.4 mm, Beam size – 80 x 20 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A 2** Data collection statistics for *Ti*HydG crystal E12. Highest resolution shell shown in parenthesis

| | **HydG** | **HydG Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | E12 Pact Premier | |
| **Data collection** | | |
| Space group | P1 | |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 53.86; 55.57; 85.27 | |
| α, β, γ (°) | 90.33; 95.79; 112.33 | |
| | | *Peak* |
| Wavelength (Å) | 0.97625 | 1.7921 |
| No of reflections | 200142 (6910-2418) | 123660 (1615-1311) |
| No of unique reflections | 78324 (3043-933) | 40747 (756-509) |
| Resolution (Å) | 51.34-1.83 (8.18-1.83) | 84.65-1.73 (7.74-1.73) |
| $R_{sym}$ or $R_{merge}$ | 0.036 (0.025-0.403) | 0.034 (0.033-0.472) |
| $I$ / σ$I$ | 14.2 (37.2-1.9) | 18.4 (37.3-1.8) |
| Completeness (%) | 94.2 (98.0-66.2) | 66.7 (88.4-4.7) |
| Redundancy | 2.6 (2.6-2.3) | 1.6 (1.7-1.2) |
| Twinning score | 2.37, no twinning | 2.42, no twinning |

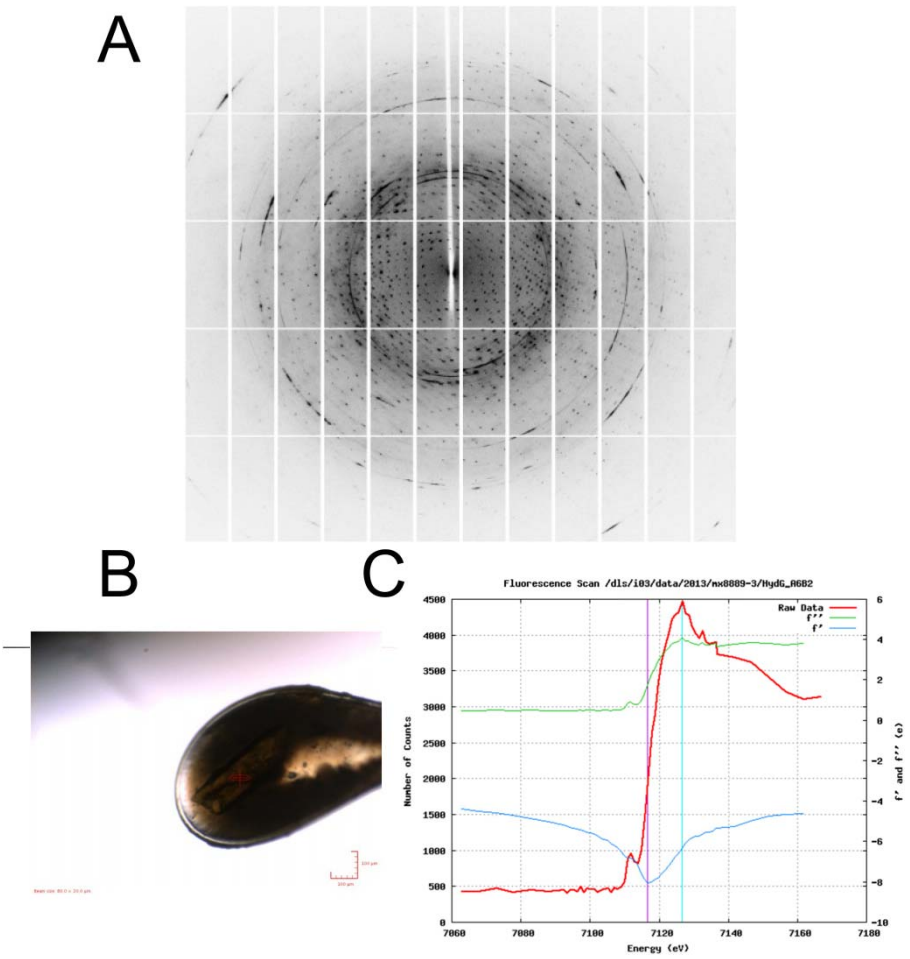



**Figure A8** X-ray diffraction of *Ti*HydG crystal C11. **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 2.0 Å resolution. **B:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM and L-tyrosine (0.8 mM HydG, 10 mM SAM, 3 mM tyrosine), directly frozen after formation in condition C11 of Hampton PEGRx HT screen (0.1M Bis-Tris, pH 6.5; 20% w/v PEG-ME 5000). Loop size – 0.5–0.7 mm, Beam size – 80 x 20 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A3** Data collection statistics for *Ti*HydG crystal C11. Highest resolution shell shown in parenthesis

| | **HydG** | **HydG Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | C11 Hampton PEGRx | |
| **Data collection** | | |
| Space group | P1 | |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 54.34; 56.9; 85.14 | |
| α, β, γ (°) | 88.69; 83.46; 66.33 | |
| | | *Peak* |
| Wavelength (Å) | 0.97625 | 1.72921 |
| No of reflections | 57392 (3010-794) | 174961 (6319-2596) |
| No of unique reflections | 26361 (1721-297) | 58842 (2682-726) |
| Resolution (Å) | 84.56-1.99 (8.90-1.99) | 52.14-1.73 (7.75-1.73) |
| $R_{sym}$ or $R_{merge}$ | 0.125 (0.089-0.922) | 0.033 (0.034-0.230) |
| $I$ / σ$I$ | 5.7 (11.6-2.0) | 19.8 (38.2-3.0) |
| Completeness (%) | 93.7 (97.8-94.1) | 60.5 (80.7-5.2) |
| Redundancy | 2.3 (2.6-2.2) | 1.5 (1.5-1.1) |
| Twinning score | 2.77, no twinning | 2.37, no twinning |

**Table A4** Refinement statistics for the crystallographic datasets of *Ti*HydG, following refinement. Due to the lack of novel structural information when comparing with the higher resolution structure discussed in Chapter 03, the refinement of these datasets was not carried out to completion.

| | **A8 Pact Premier** | **E12 Pact Premier** | **C11 PEG Rx** |
|---|---|---|---|
| $R_{work}$ / $R_{free}$ | 0.21/0.26 | 0.24/0.28 | 0.22/0.24 |
| No. atoms: | | | |
| Chain A / Chain B | 7215/7059 | 7200/7098 | 7126/6984 |
| Ligand/ion | 100 | 102 | 88 |
| Water | 289 | 362 | 486 |
| *B*-factors: | | | |
| Chain A / Chain B | 50.5/45.1 | 53.2/41.8 | 25.9/23.1 |
| Ligand/ion | 51.5 | 47.92 | 25.4/20.3 |
| Water | 41.7 | 39.2 | 39.5 |
| R.m.s deviations: | | | |
| Bond lengths (Å) | 0.030 | 0.037 | 0.017 |
| Bond angles (°) | 0.814 | 1.36 | 1.774 |

## A.4 Crystallization data for *Ti*HydG WT crystals with 6 mM L-tyrosine

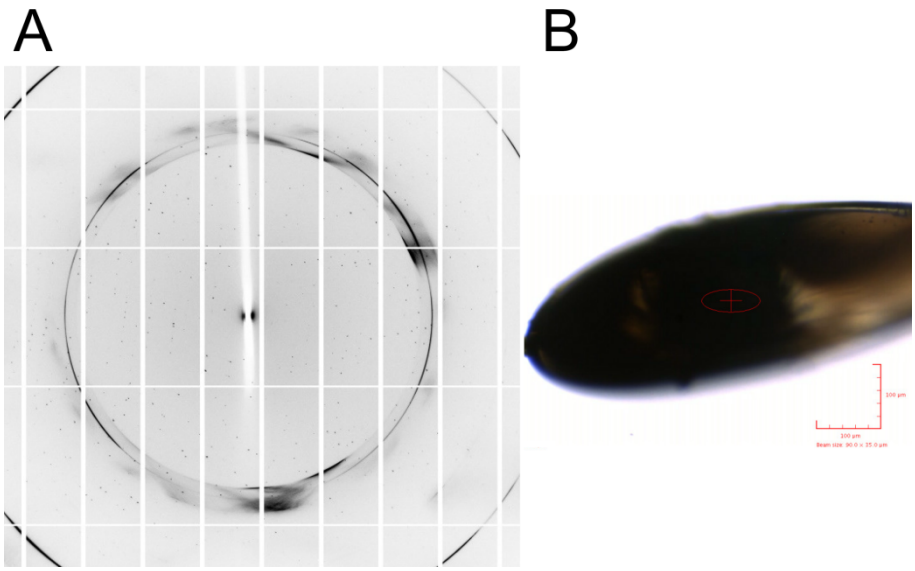



**Figure A9** X-ray diffraction of *Ti*HydG crystal grown with 6 mM L-tyrosine. **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 3.0 Å resolution. **B:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM and L-tyrosine (0.8 mM HydG, 10 mM SAM, 6 mM L-tyrosine), directly frozen after formation in condition C2 of Hampton PEGRx HT commercial screen (0.1 M Imidazole, pH 7.0; 20% Jeffamine® ED-2001, pH 7.0). Loop size – 0.4–0.5 mm, Beam size – 90 x 35 µm.

**Table A5** Data collection statistics. Highest resolution shell shown in parenthesis

| | **HydG** |
|---|---|
| **Crystallization condition** | C2 Hampton PEGRx |
| **Data collection** | |
| Space group | P1 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 54.18; 56.36; 85.22 |
| α, β, γ (°) | 89.66; 83.75; 66.82 |
| Wavelength (Å) | 0.97625 |
| No of reflections | 77277 (5900-909) |
| No of unique reflections | 18100 (1342-199) |
| Resolution (Å) | 84.64-3.00 (13.42-3.00) |
| $R_{sym}$ or $R_{merge}$ | 0.169 (0.041-0.537) |
| $I$ / σ$I$ | 6.3 (20.9-2.1) |
| Completeness (%) | 98.2 (97.1-98.7) |
| Redundancy | 4.3 (4.6-4.4) |
| Twinning score | 2.34, no twinning |

## A.5 Crystallization data for low protein (0.3 mM) *Ti*HydG WT crystals with 12 mM L-tyrosine

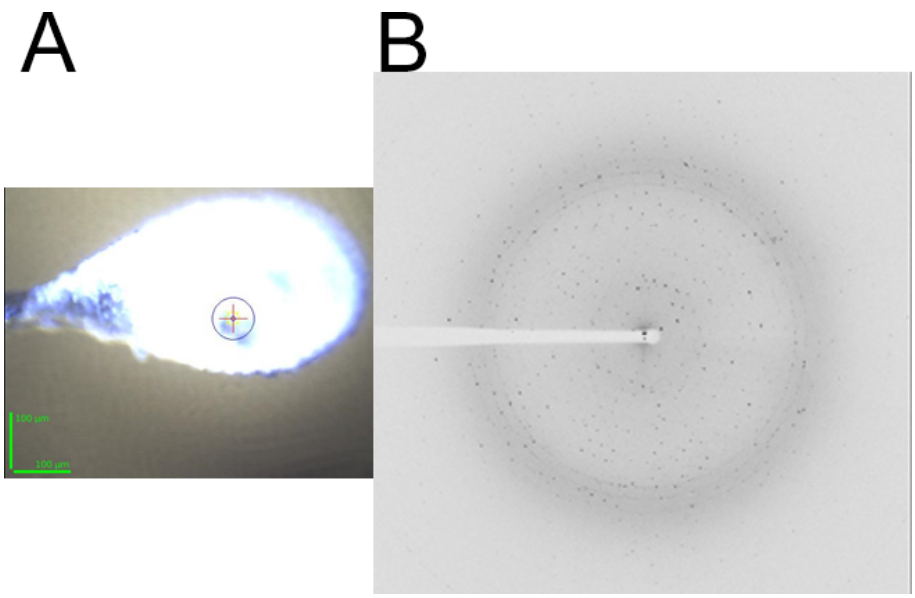



**Figure A10 A:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM and L-tyrosine (0.3 mM HydG, 3.0 mM SAM, 12.0 mM L-tyrosine), directly frozen after formation in condition F1 of Hampton PEGRx HT screen (0.1M Bis-Tris proptane, pH 65.;0.2M sodium fluoride; 20% PEG 3350). The loop showed a large formation of ice surrounding the sample. Loop size – 0.4–0.5 mm, Beam size – 50 x 50 µm. **B:** Associated diffraction pattern from 7.1 keV synchrotron radiation to 2.3 Å resolution.

**Table A6** Data collection statistics. Highest resolution shell shown in parenthesis

| | **HydG** |
|---|---|
| **Crystallization condition** | F1 Hampton PEGRx |
| **Data collection** | |
| Space group | P1 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 53.9; 56.0; 84.9 |
| α, β, γ (°) | 90.0; 83.9; 67. 2 |
| Wavelength (Å) | 1.73745 |
| No of reflections | 272295 |
| No of unique reflections | 13095 |
| Resolution (Å) | 100.0-2.31 (2.39-2.31) |
| $R_{sym}$ or $R_{merge}$ | 0.086 (0.485) |
| $I$ / σ$I$ | 19.24 (3.98) |
| Completeness (%) | 87.6 (44.5) |
| Redundancy | 6.81 (3.38) |
| Twinning Z-score | 1.071, no twinning |

## A.6 *Ti*HydG mutant activity

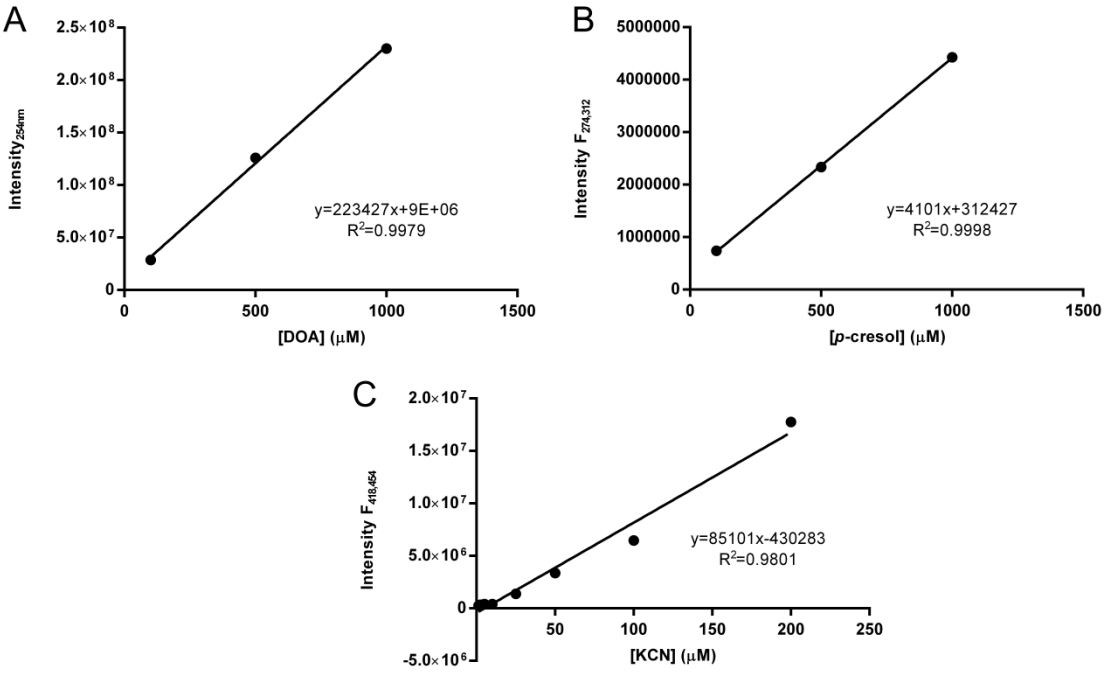



**Figure A11** Standard curves for deoxyadenosine (A), *p*-cresol (B) and cyanide (C)

## A.7 *Ti*HydG His265Glu crystallization

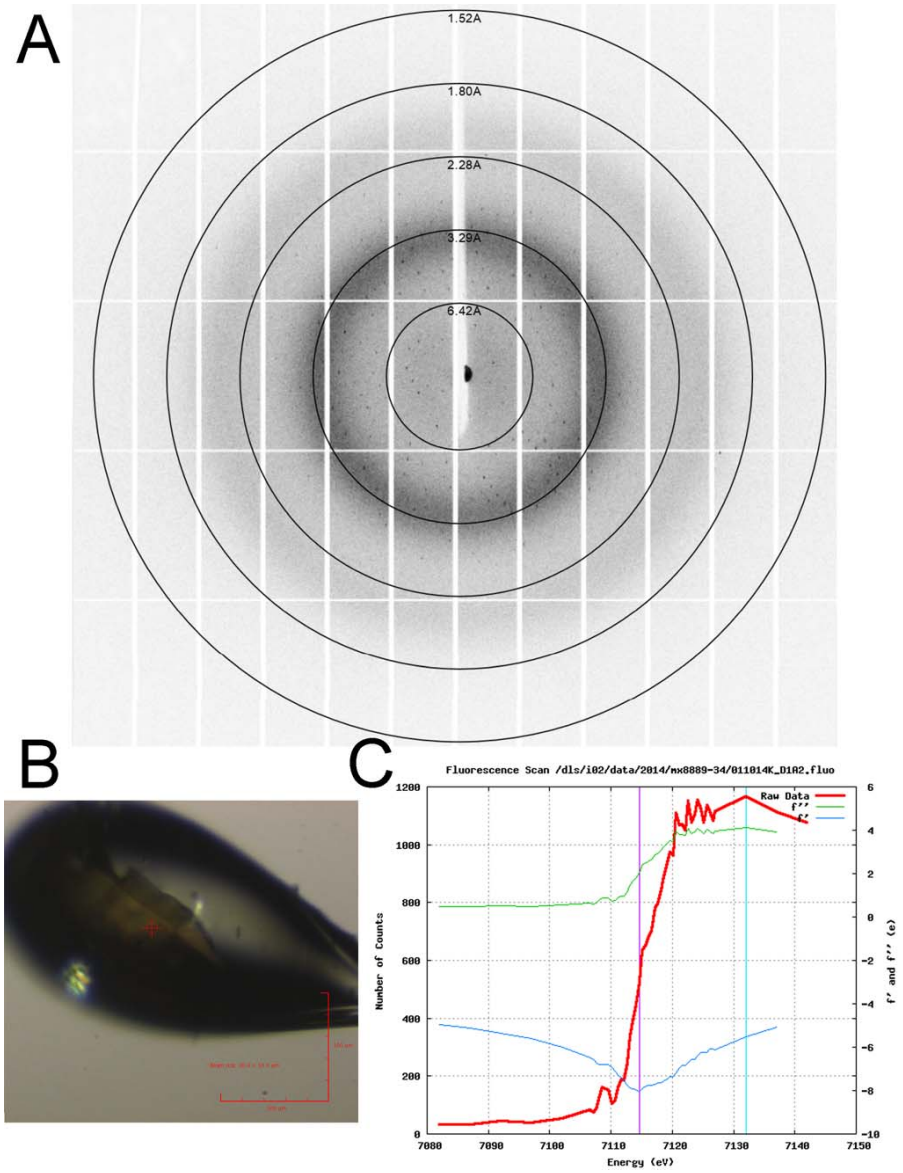



**Figure A12 A:** Associated diffraction pattern from 12 keV synchrotron radiation to, 2.2 Å resolution. **B:** Loop containing a crystal of *Ti*HydG His265Glu, grown in the presence of SAM (0.8 mM HydG, 10 mM SAM, directly frozen after formation in condition F1 of Hampton PEGRx HT (0.1M Bis-Tris proptane, pH 65.; 0.2M sodium fluoride; 20% PEG 3350). Loop size – 0.3–0.4 mm, Beam size – 10 x 10 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A7** Data collection statistics. Highest resolution shell shown in parenthesis

| | **HydG H265E** | **HydG H265E Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | F1 Hampton PEGRx | |
| | **Data collection** | |
| Space group | P1 | |
| | Cell dimensions | |
| *a*, *b*, *c* (Å) | 53.74; 55.03; 84.71 | |
| α, β, γ (°) | 90.75; 95.87; 111.91 | |
| | | *Peak* |
| Wavelength (Å) | 0.97944 | 1.73767 |
| No of reflections | 198522 (6168-2475) | 375676 (26924-4596) |
| No of unique reflections | 75876 (2434-974) | 55077 (3859-636) |
| Resolution (Å) | 44.98- 1.75 (7.83-1.75) | 51.02-2.02 (9.03-2.02) |
| $R_{sym}$ or $R_{merge}$ | 0.063 (0.047-0.309) | 0.103 |
| $.I / \sigma I$ | 8.5 (18.2-2.1) | 11.3 (28.1-3.0) |
| Completeness (%) | 84.1 (97.5-36.0) | 90.0 (95.5-86.3) |
| Redundancy | 2.6 (2.5-2.5) | 6.8 (7.2-7.0) |
| Twinning score | 2.26, no twinning | 2.20, no twinning |

**Table A8** Score values for molecular replacement using *Phaser*.

| | **Hampton PEGRx F1** |
|---|---|
| LL G-score (best) | 9513.579 |
| TF Z-score (best) | 13.6 |

**Table A9** Refinement statistics for *Ti*HydG His265Glu dataset, following refinement

| **Refinement** | ***Ti*HydG His265Glu F1 PEG Rx HT** |
|---|---|
| Resolution (Å) | 44.98- 1.75 (7.83-1.75) |
| No. reflections (free set) | 75876 (2434-974) |
| $R_{work}$ / $R_{free}$ | 0.18/0.23 |
| No. atoms: | |
| Chain A / Chain B | 3596/3540 |
| Ligand/ion | 89 |
| Water | 731 |
| *B*-factors: | |
| Chain A / Chain B | 32.0/26.3 |
| Ligand/ion | 30.2 |
| Water | 35.7 |
| R.m.s deviations: | |
| Bond lengths (Å) | 0.009 |
| Bond angles (°) | 1.481 |

## A.8 *Ti*HydG activity with substrate analogues

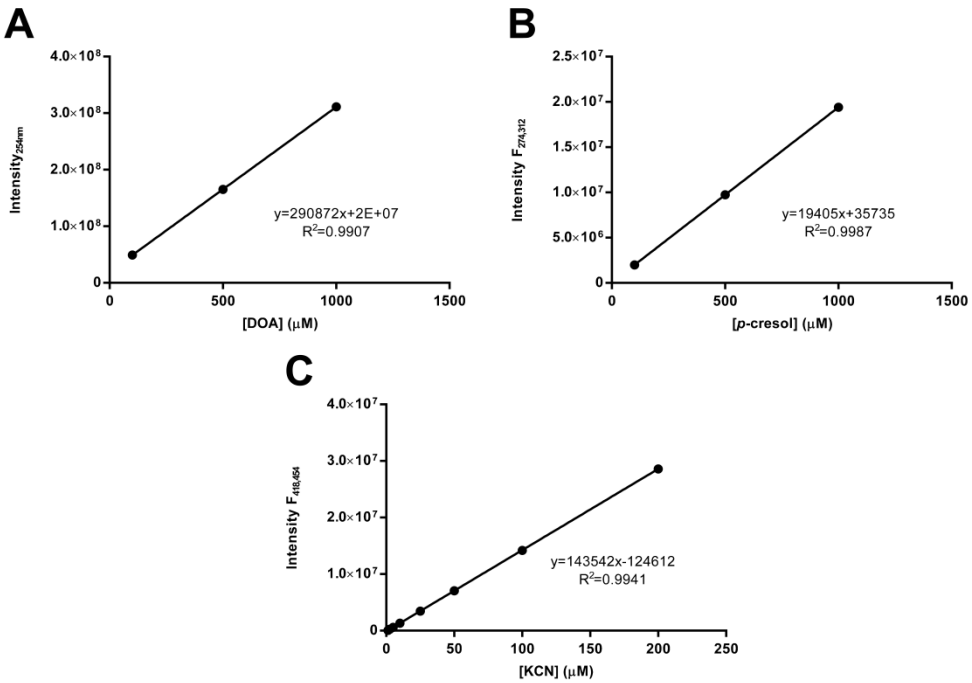



**Figure A13** Standard curves for deoxyadenosine (A) p-cresol (B) and cyanide (C).

## A.9 *Ti*HydG crystallization with THIQCA

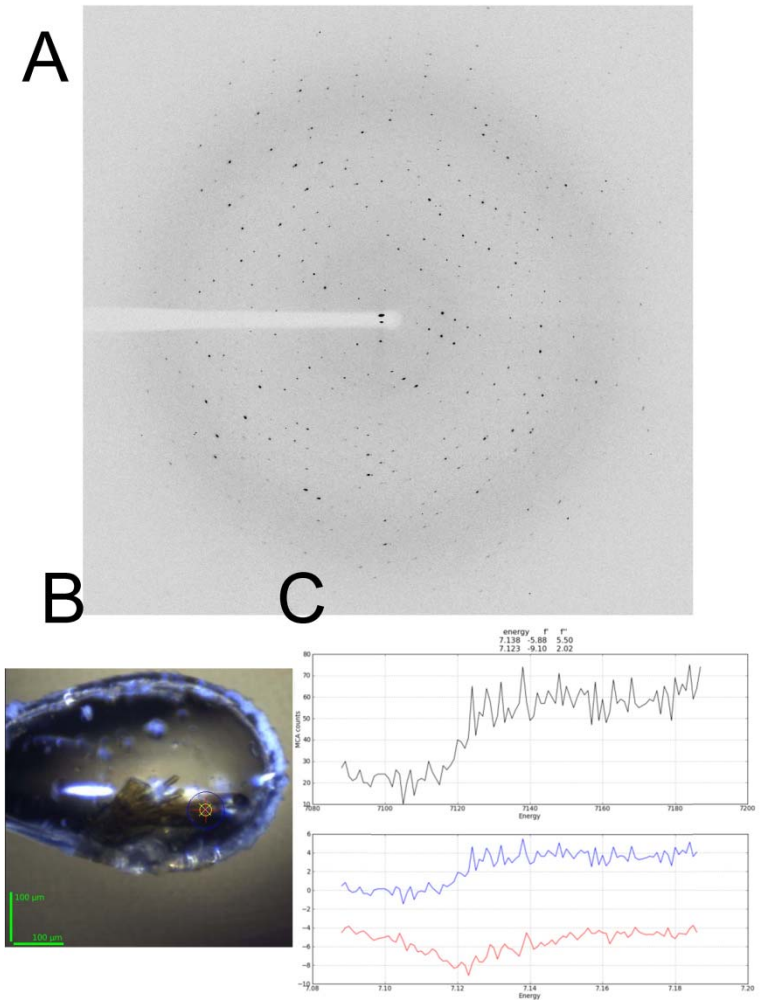



**Figure A14 A:** Associated diffraction pattern from 12 keV synchrotron radiation to, 2.2 Å resolution. **B:** Loop containing a crystal of *Ti*HydG, grown in the presence of SAM and THIQCA (0.8 mM HydG, 10 mM SAM, 30 mM THIQCA), directly frozen after formation in condition F1 of Hampton PEGRx HT (0.1M Bis-Tris proptane, pH 65.;0.2M sodium fluoride; 20% PEG 3350). The loops showed a large formation of ice surrounding the sample. Loop size – 0.4–0.5 mm, Beam size – 50 x 50 µm. **C:** Fluorescence scan near the iron κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A 10** Data collection statistics. Highest resolution shell shown in parenthesis

| | ***Ti*HydG THIQCA** |
|---|---|
| **Crystallization condition** | F1 Hampton PEGRx |
| **Data collection** | |
| Space group | P1 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 54.0; 55.9; 84.8 |
| α, β, γ (°) | 89.8; 83.8; 66.9 |
| Wavelength (Å) | 1.74037 |
| No of reflections | 346982 |
| No of unique reflections | 24812 |
| Resolution (Å) | 100-2.17 (2.25-2.17) |
| $R_{sym}$ or $R_{merge}$ | 0.124 (1.17) |
| $I$ / σ$I$ | 15.47 (1.75) |
| Completeness (%) | 91.9 (64.8) |
| Redundancy | 7.23 (5.01) |

**Table A 11 Score values for molecular replacement using *Phaser***

| | ***Ti*HydG + THIQCA** |
|---|---|
| LL G-score (best) | 12068.483 |
| TF Z-score (best) | 8.0 |

## A.10 *Ti*HydG crystallization with tyramine

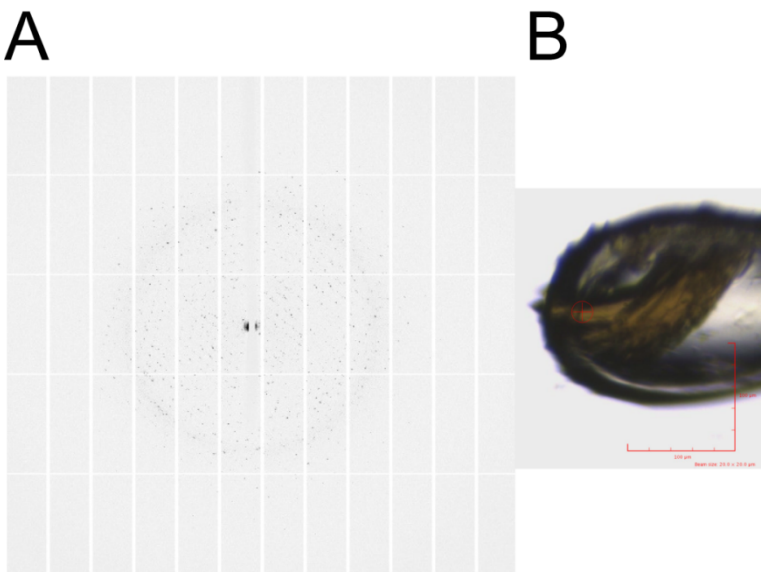



**Figure A15 A:** Associated diffraction pattern from 12 keV synchrotron radiation to 3.0 Å resolution. **B:** Loop containing a *Ti*HydG crystal, grown with SAM and L-tyramine (0.8 mM HydG, 10 mM SAM, 3 mM L-tyramine), directly frozen after 72h formation in condition F1 of Hampton PEGRx HT screen (0.1M Bis-Tris proptane, pH 65.;0.2M sodium fluoride; 20% PEG 3350) Loop size – 0.2–0.3 mm, Beam size – 20 x 20 µm.

**Table A 12** Data collection statistics. Highest resolution shell shown in parenthesis

| | ***Ti*HydG L-tyramine** |
|---|---|
| **Crystallization condition** | F1 Hampton PEGRx |
| **Data collection** | |
| Space group | P1 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 53.99; 56.14; 85.42 |
| α, β, γ (°) | 90.08; 84.02; 67.23 |
| Wavelength (Å) | 0.97949 |
| No of reflections | 153984 (5267-1850) |
| No of unique reflections | 17364 (692-202) |
| Resolution (Å) | 45.72-2.96 (13.24-2.96) |
| $R_{sym}$ or $R_{merge}$ | 0.136 (0.658-0.055) |
| $I$ / σ$I$ | 12.5 (31.5-2.8) |
| Completeness (%) | 90.7 (93.5-48.7) |
| Redundancy | 8.9 (9.2-7.6) |
| Twinning Z-score | 2.12 |

**Table A13** Score values for molecular replacement using *Phaser*.

| | ***Ti*HydG + L-tyramine** |
|---|---|
| LL G-score (best) | 4387.405 |
| TF Z-score (best) | 7.6 |

## A.11 *Ti*HydG crystallization with L-cysteine

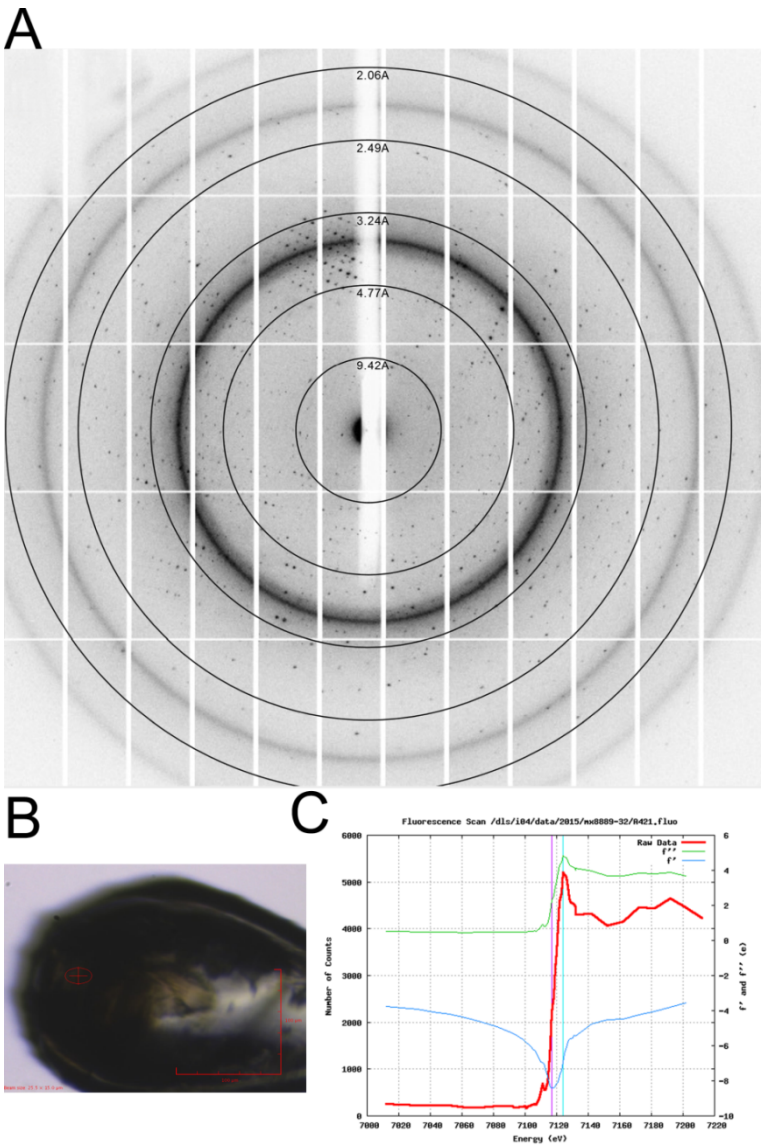



**Figure A16** Testing of *Ti*HydG crystal soaked with L-Cys **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 1.8 Å resolution. **B:** Loop containing a crystal of *Ti*HydG soaked with 5.0 mM of L-cysteine for 4 hours. Crystal pre-grown in the presence of SAM (0.8 mM HydG, 20 mM SAM), directly frozen after formation in condition F1 of PEGRx HT Hampton screen (0.1M Bis-Tris proptane, pH 65.; 0.2M sodium fluoride; 20% PEG 3350). Loop size – 0.3–0.4 mm, Beam size – 25.5 x 15 µm. **C:** Fluorescence scan near the selenium κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A14** Data collection information for *Ti*HydG crystal soaked 4h with L-cysteine (5.0 mM).

| | **HydG + L-cys** | **Fe (SAD)** |
|---|---|---|
| **Crystallization condition** | F1 Hampton PEGRx | |
| | **Data collection** | |
| Space group | P1 | |
| | Cell dimensions | |
| *a*, *b*, *c* (Å) | 60.64; 85.18; 91.95 | |
| α, β, γ (°) | 93.97; 92.12; 95.23 | |
| | | *Peak* |
| Wavelength (Å) | 0.97949 | 1.73842 |
| No of reflections | 577632 (15918-7622) | 536659 (27037-6528) |
| No of unique reflections | 148387 (4750-1908) | 111838 (6586-1321) |
| Resolution (Å) | 51.46-1.77 (7.92-1.77) | 42.36-2.01 (8.99-2.01) |
| $R_{sym}$ or $R_{merge}$ | 0.081 (0.846-0.043) | 0.112 (1.041-0.055) |
| $I$ / σ$I$ | 6.7 (17.6-1.1) | 7.2 (17.7-1.2) |
| Completeness (%) | 83.3 (95.3-36.1) | 86.4 (93.6-63.0) |
| Redundancy | 3.9 (4.0-3.4) | 4.8 (4.9-4.1) |
| Twinning score | 3.46, no twinning | 3.30, no twinning |

**Table A15** Score values for molecular replacement using *Phaser*.

| | *Ti*HydG + 5.0 mM Cysteine + 10 mM SAM |
|---|---|
| LL G-score (best) | 15807.34 |
| TF Z-score (best) | 13.1 |

## A.12 *Ti*HydG crystallization with CO/CN

**Table A16** Data collection statistics. Highest resolution shell shown in parenthesis. Refinement of the datasets is still ongoing and the current $R_{work}/R_{free}$ is shown at the end of the table.

| | ***Ti*HydG + 0.1 mM KCN** | ***Ti*HydG + 0.5 mM KCN** | ***Ti*HydG + 1.0 mM KCN** | ***Ti*HydG + 2.0 mM KCN** | ***Ti*HydG + CO** |
|---|---|---|---|---|---|
| **Soaking time** | 2h | 2h | 2h | 48h | 48h |
| **Crystallization condition** | F1 Hampton PEGRx | | | | |
| **Data collection** | | | | | |
| Space group | P1 | P1 | P1 | P1 | P1 |
| Cell dimensions | | | | | |
| *a*, *b*, *c* (Å) | 53.62; 54.89; 84.79 | 53.20; 54.54; 83.73 | 53.66; 55.09; 84.79 | 54.0; 55.8; 85.2 | 53.8; 55.1; 84.9 |
| α, β, γ (°) | 90.76; 95.76; 11.96 | 90.82; 95.88; 111.76 | 90.65; 95.91; 112.21 | 89.4; 83.2; 67.3 | 90.1; 96.4; 112.45 |
| Wavelength (Å) | 097946 | 0.97945 | 0.97941 | | 0.9998 |
| No of reflections | 508014 (7996-7004) | 129167 (9044-1633) | 385905 (24653-3992) | 278648 | 176981 |
| No of unique reflections | 77937 (1229-1152) | 35883 (2618-3949 | 94685 (6462-1027) | 29054 | 10425 |
| Resolution (Å) | 50.83-1.63 (7.29-1.63) | 44.59-2.32 (10.38-2.32) | 50.93-1.70 (7.60-1.70) | 100.0-1.82 (1.89-1.82) | 100.0-2.06 (2.13-2.06) |

| | ***Ti*HydG + 0.1 mM KCN** | ***Ti*HydG + 0.5 mM KCN** | ***Ti*HydG + 1.0 mM KCN** | ***Ti*HydG + 2.0 mM KCN** | ***Ti*HydG + CO** |
|---|---|---|---|---|---|
| Rsym or Rmerge | 0.068 (0.051-0.596) | 0.181 (0.633-0.083) | 0.055 (0.532-0.034) | 0.073 (0.575) | 0.092 |
| I / σI | 15.1 (31.8-2.2) | 5.5 (14.3-1.6) | 11.8 (32.7-2.1) | 9.0 (1.86) | 1.5 (0.2) |
| Completeness (%) | 70.1 | 96.6 (96.0-94.7) | 96.7 (95.9-90.5) | 94.1 (90.8) | 58.3 (50.7) |
| Redundancy | 6.5 (6.5-6.1) | 3.8 | 4.1 (3.9-3.8) | 3.4 (3.31) | 3.19 (3.02) |
| Twinning Z-score | 2.23, no twinning | 2.27, no twinning | 2.20, no twinning | 2.36, no twinning | 1.19, no twinning |
| Current Rwork | 0.195 | 0.264 | 0.255 | 0.234 | 0.203 |
| Current Rfree | 0.236 | 0.328 | 0.301 | 0.298 | 0.264 |

**Table A17** Anomalous data collection statistics. Highest resolution shell shown in parenthesis.

| | *Ti*HydG + 0.1 mM KCN | *Ti*HydG + 0.5 mM KCN | *Ti*HydG + 1.0 mM KCN | *Ti*HydG + 2.0 mM KCN |
|---|---|---|---|---|
| **Soaking time** | 2h | 2h | 2h | 48h |
| **Crystallization condition** | F1 Hampton PEGRx | | | |
| **Data collection** | | | | |
| Space group | P1 | P1 | P1 | P1 |
| Cell dimensions | | | | |
| *a*, *b*, *c* (Å) | 53.64; 55.03; 85.25 | 53.78; 55.18; 84.77 | 53.70; 55.13; 85.19 | 53.90; 55.19; 84.94 |
| α, β, γ (°) | 91.16; 95.50; 112.06 | 90.61; 95.88; 111.67 | 90.73; 95.72; 112.25 | 90.103; 96.423; 112.108 |
| Wavelength (Å) | 1.736479 | 1.73889 | 1.73524 | 1.73775 |
| No of reflections | 448754 (17501-5812() | 317390 (21096-2998) | 342006 (23960-4323) | 62104 |
| No of unique reflections | 64942 (2706-808) | 46707 (3334-410) | 50082 (3558-578) | 4921 |
| Resolution (Å) | 50.91-1.88 (8.41-1.88) | 15.0-2.14 (9.57-2.14) | 45.09-2.09 (9.35-2.09) | 49.57-2.96 (3.06-2.96) |
| $R_{sym}$ or $R_{merge}$ | 0.067 (0.049-0.719) | 0.107 (0.693-0.067) | 0.088 (0.641-0.051) | 0.004 (0.188) |
| $I$ / σ$I$ | 15.6 | 13.3 (32.4-2.8) | 14.2 (38.6-2.8) | 2.2 (0.2) |

| | *Ti*HydG + 0.1 mM KCN | *Ti*HydG + 0.5 mM KCN | *Ti*HydG + 1.0 mM KCN | *Ti*HydG + 2.0 mM KCN |
|---|---|---|---|---|
| Completeness (%) | 86.1 | 90.6 (85.6-72.7) | 91.3 (96.1-87.7) | 94.4 (92.5) |
| Redundancy | 6.9 (7.2-6.5) | 6.8 (7.3-6.3) | 6.8 (7.5-6.7) | 3.5 (2.9) |
| Twinning Z-score | 2.1, no twinning | 2.03, no twinning | 2.07, no twinning | 2.26; no twinning |

## A.13 Quantification of cyanide formation by *Ti*HydG in the presence of an a-iminocarboxylic acid complex

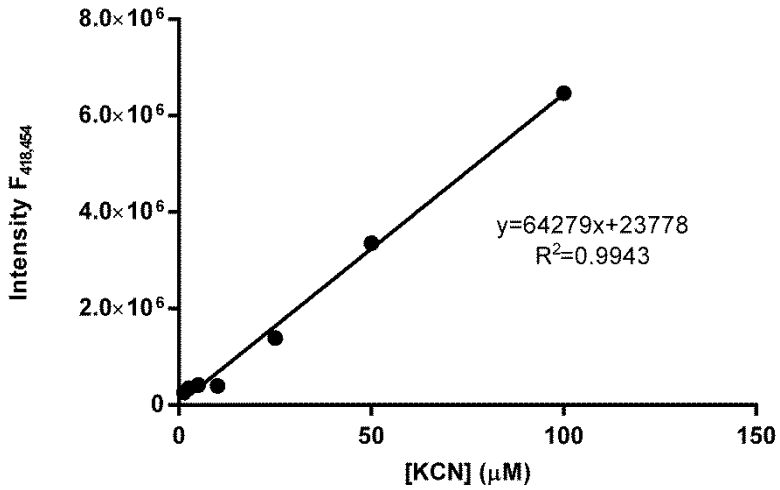



**Figure A17** Standard curve used in cyanide quantification.

## A.14 Plasmid map of *Te*LipA2

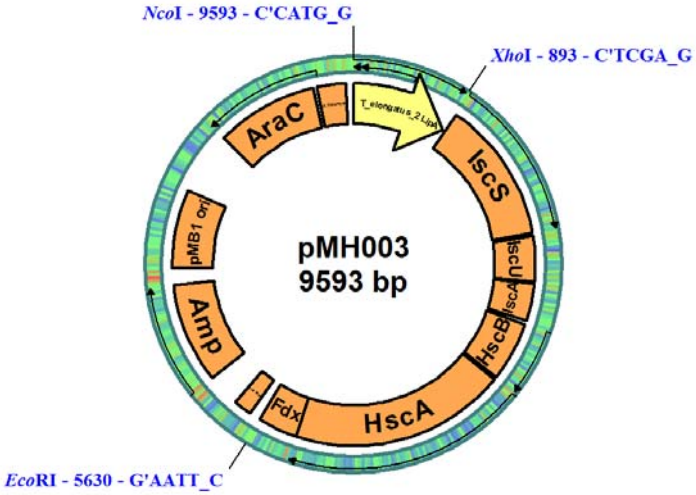



**Figure A18** pFM025 (section 10.1 Figure A1) derived pMH003 plasmid, containing the *isc* operon as well as the gene coding *Thermosynechococcus enlongatus* LipA 2, tagged C-terminally by a 6His fragment for affinity purification,

***Telipa2* gene sequence introduced in pFM024, to form pMH003**

ATGGCACTGAGTCGTCCGCTGCCGAGCTGGCTGCGTAAACCGCTGGGTAAAGCAAGCGAAATTAGCACCG
TTCAGCGTCTGGTTCGTCAGTATGGTATTCATACCATTTGTGAAGAAGGTCGTTGTCCGAATCGTGGTGAAT
GTTATGGTCAGAAAACCGCAACCTTTCTGCTGCTGGGTCCGACCTGTACCCGTGCATGTGCATTTTGTCAGG
TTGAAAAAGGTCATGCACCGGCAGCAGTTGATCCGGAAGAACCGACCAAAATTGCAGCAGCCGTTGCAAC
CCTGGGTCTGCGTTATGTTGTTCTGACCAGCGTTGCACGTGATGATCTGCCGGATCAGGGTGCAGGTCAGT
TTGTTGCCACAATGGCAGCAATTCGTCAGCGTTGTCCGGGTACAGAAATTGAAGTTCTGAGTCCGGATTTT
CGTATGGATCGTGGTCGTCTGAGCCAGCGTGATTGTATTGCACAGATTGTTGCAGCACAGCCTGCATGCTA
TAATCATAATCTGGAAACCGTTCGTCGTCTGCAAGGTCCGGTTCGTCGTGGTGCAACCTATGAAAGCAGCC
TGCGTGTTCTGGCAACCGTTAAAGAACTGAATCCGGATATTCCGACCAAATCAGGTCTGATGCTGGGCCTG
GGTGAAACCGAAGCAGAAATTATTGAAACCCTGAAAGATCTGCGTCGTGTTGGTTGTGATCGTCTGACCCT
GGGCCAGTATCTGCCTCCGAGCCTGAGCCATCTGCCGGTTGTGAAATATTGGACACCGGAAGAATTTAACA
CCCTGGGTAATATTGCCCGTGAACTGGGTTTTAGCCATGTTCGTAGCGGTCCGCTGGTTCGTAGCAGCTATC
ATGCAGCAGAAGGTGGTCATCATCATCACCATCATTAACTCGA

Appendix A

**Translated sequence of *Te*LipA2 protein, with the [4Fe-4S] coordinating cysteines and serine in red**

MALSRPLPSWLRKPLGKASEISTVQRLVRQYGIHTICEEGRCPNRGECYGQKTATFLLLGPTCTRACAFCQVEKG
HAPAAVDPEEPTKIAAAVATLGLRYVVLTSVARDDLPDQGAGQFVATMAAIRQRCPGTEIEVLSPDFRMDRGR
LSQRDCIAQIVAAQPACYNHNLETVRRLQGPVRRGATYESSLRVLATVKELNPDIPTKSGLMLGLGETEAEIIETLK
DLRRVGCDRLTLGQYLPPSLSHLPVVKYWTPEEFNTLGNIARELGFSHVRSGPLVRSSYHAAEGGHHHHHH

## A.15 *Te*LipA 2 crystallization with 2.5 molar equivalents of $FeCl_3/Na_2S$

**Table A18** Data collection information for the *TeLipA2*crystal reconstituted with 2.5 molar equivalent of $FeCl_3$ and $Na_2S$. Highest resolution shell shown in parenthesis

| | ***Te*LipA*2*** **2.5 molar equivalent reconstitution** |
|---|---|
| **Crystallization condition** | H12 Hampton PEGRx |
| **Data collection** | |
| Space group | $C222_1$ |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 73.36; 162.00; 59.12 |
| α, β, γ (°) | 90.00; 90.00; 90.00 |
| Wavelength (Å) | 0.97625 |
| No of reflections | 52351 (3609) |
| No of unique reflections | 13984 (981) |

| | ***Te*LipA*2*<br>**2.5 molar equivalent reconstitution** |
|---|---|
| Resolution (Å) | 43.49-2.38 (10.65-2.38) |
| $R_{sym}$ or $R_{merge}$ | 0.056 (0.034-0.305) |
| $I$ / σ$I$ | 10.3 (23.4-2.0) |
| Completeness (%) | 96.5 (95.7-93.5) |
| Redundancy | 3.7 (3.1-3.7) |
| Twinning Z-score | 2.74, no twinning |

**Table A19** Refinement statistics for the crystal structure of *Te*LipA 2 reconstituted with 2.5 Meq. $FeCl_3/Na_2S$

| **Refinement** | ***Te*LipA2 2.5 molar equivalent reconstitution** |
|---|---|
| Resolution (Å) | 43.49-2.38 |
| $R_{work}$ / $R_{free}$ | 0.225/0.278 |
| No. atoms: | |
| Chain A / Chain B | 7285/6971 |
| Ligand/ion | 134 |
| Water | 470 |
| B-factors: | |
| Chain A | 74.92 |
| Ligand/ion | 78.97 |
| Water | 67.16 |

| **Refinement** | ***Te*LipA2 2.5 molar equivalent reconstitution** |
|---|---|
| R.m.s deviations: | |
| Bond lengths (Å) | 0.007 |
| Bond angles (°) | 0.649 |

**Table A20** Score values for molecular replacement using *Phaser*.

| | ***Te*LipA2 + ESVK(o)AV** |
|---|---|
| LL G-score (best) | 2533.572 |
| TF Z-score (best) | 48.2 |

## A.16 Crystallization of *Te*LipA 2 reconstituted with sodium selenide

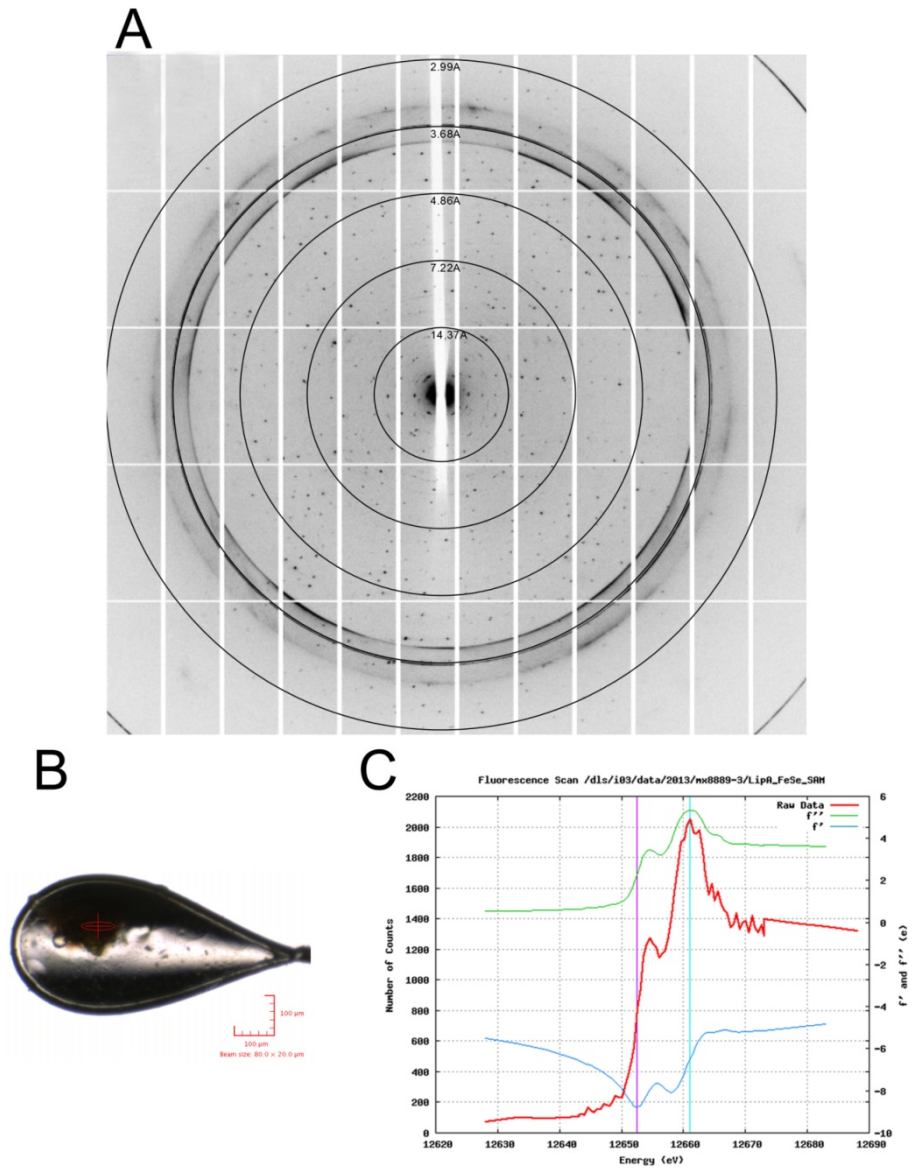



**Figure A19** Testing of *Te*LipA2 crystal reconstituted with $Na_2Se$ **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 4.0 Å resolution. **B:** Loop containing a crystal of *TeLipA2* reconstituted with 5 molar equivalents of $FeCl_3/Na_2Se$, grown in the presence of SAM (0.8 mM LipA, 10 mM SAM), directly frozen after formation in condition H12 of PEGRx HT Hampton screen (0.1M Bicine, pH 8.5; 3% w/v dextran sulfate sodium salt, 15% w/v PEG 20,000). Loop size – 0.4–0.5 mm, Beam size – 80 x 20 µm. **C:** Fluorescence scan near the selenium κ-edge. The red line shows the fluorescent count related to the presence of iron atoms in the structure. f' and f'' correspond to, respectively, the real and imaginary correction components of the atomic structure factor.

**Table A21** Data collection information for the *TeLipA2*crystal reconstituted with $FeCl_3/Na_2Se$. Highest resolution shell shown in parenthesis

| | ***Te*LipA2**<br>**$FeCl_3/Na_2Se$ reconstitution** |
|---|---|
| **Crystallization condition** | H12 Hampton PEGRx |
| **Data collection** | |
| Space group | $C222_1$ |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 137.32; 219.02; 84.08 |
| α, β, γ (°) | 90.00; 114.17; 90.00 |
| Wavelength (Å) | 0.97779 |
| No of reflections | 122,345 (8884) |
| No of unique reflections | 19,007 (1408) |
| Resolution (Å) | 63.08-4.00 (17.89-4.00) |
| $R_{sym}$ or $R_{merge}$ | 1.287 (0.220-3.922) |

| | ***Te*LipA2**<br>**$FeCl_3/Na_2Se$ reconstitution** |
|---|---|
| $I$ / σ$I$ | 0.4 (1.5-0.0) |
| Completeness (%) | 96.5 (95.7-93.5) |
| Anomalous completeness (%) | |
| Redundancy | 6.4 (6.2-6.3) |
| Twinning Z-score | 2.74, no twinning |

## A.17 Crystallization of *Te*LipA 2 with substrate analogues

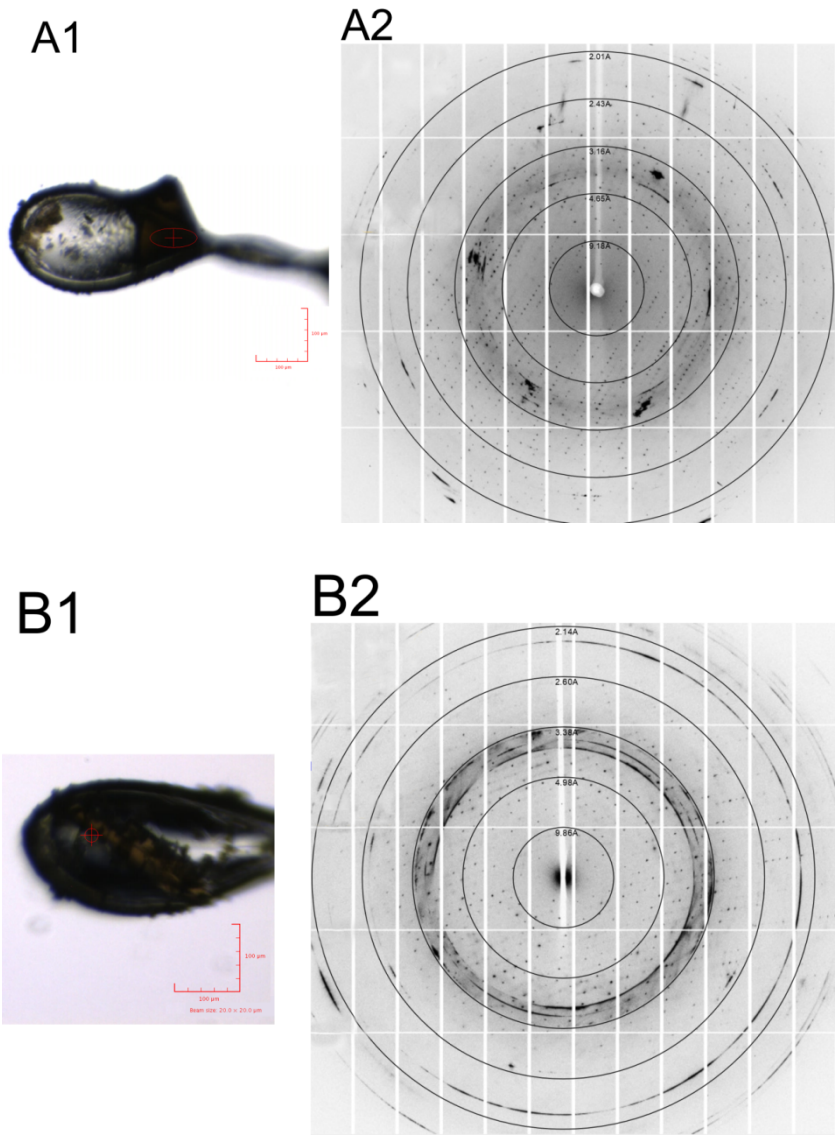



**Figure A20** Testing of *Te*LipA2 crystals reconstituted with peptide analogues. **A:** Crystal grown in the presence of ESDK(o)AD peptide. **A1:** Associated diffraction pattern from 12 keV synchrotron radiation to 1.9 Å resolution. **A2:** Loop containing a crystal of *TeLipA2* reconstituted with 5 molar equivalents of $FeCl_3$/$Na_2Se$, grown in the presence of SAM (0.8 mM LipA, 10 mM SAM), directly frozen after formation in condition H12 of PEGRx HT Hampton screen (0.1M Bicine, pH 8.5; 3% w/v dextran sulfate sodium salt, 15% w/v PEG 20,000). Loop size – 0.4–0.5 mm, Beam size – 80 x 20 µm. **B:** Crystal grown in the presence of ESVK(o)AV peptide. **B1:** Associated diffraction pattern from 12 keV synchrotron radiation to 2.1 Å resolution. **B2:** Loop containing a crystal of *TeLipA2* reconstituted with 5 molar equivalents of $FeCl_3$/$Na_2Se$, grown in the presence of SAM (0.8 mM LipA, 10 mM SAM), directly frozen after formation in condition H12 of PEGRx HT Hampton screen (0.1M Bicine, pH 8.5; 3% w/v dextran sulfate sodium salt, 15% w/v PEG 20,000). Loop size – 0.4–0.5 mm, Beam size – 80 x 20 µm.

**Table A22** Data collection information for the *Te*LipA2 crystals grown with peptide substrate analogues.

| | ***Te*LipA2** | |
|---|---|---|
| **Substrate** | ESDK(o)AD | ESVK(o)AV |
| **Crystallization condition** | H12 Hampton PEGRx | |
| **Data collection** | | |
| **Space group** | $C222_1$ | $C222_1$ |
| **Cell dimensions** | | |
| **a, b, c (Å)** | 70.74; 161.36; 59.15 | 72.86; 162.39; 58.84 |
| **α, β, γ (°)** | 90.0; 90.0; 90.0 | 90.0; 90.0; 90.0 |
| **Wavelength (Å)** | 0.97625 | 0.97625 |
| **No of reflections** | 234519 (17297-2246) | 89491 (6347-1085) |
| **No of unique reflections** | 26281 (1926-327) | 20583 (1519-277) |
| **Resolution (Å)** | 59.15-1.91 (8.54-1.91) | 44.06-2.08 (9.30-2.08) |
| **$R_{sym}$ or $R_{merge}$** | 0.071 (0.028-0.478) | 0.041 (0.445-0.024) |
| **I / σI** | 17.5 (41.6-5.5) | 14.1 (36.5-2.5) |
| **Completeness (%)** | 98.3 (98.4-93.6) | 96.1 (98.1-97.9) |
| **Redundancy** | 8.9 (9.0-6.9) | 4.3 (4.2-3.9) |
| **Twinning Z-score** | 2.92, no twinning | 3.92, no twinning |

| | | ***Te*LipA2** |
|---|---|---|
| **Current R-work** | 0.294 | 0.283 |
| **Current R-free** | 0.337 | 0.342 |

**Table A23** Score values for molecular replacement using *Phaser*.

| | ***Te*LipA2 + ESDK(o)AD** | ***Te*LipA2 + ESVK(o)AV** |
|---|---|---|
| LL G-score (best) | 3598.52 | 3469.96 |
| TF Z-score (best) | 58.4 | 55.1 |

## A.18 Crystallization of *Te*LipA 2 with substrate analogue ESDK(o)AD and MTAN

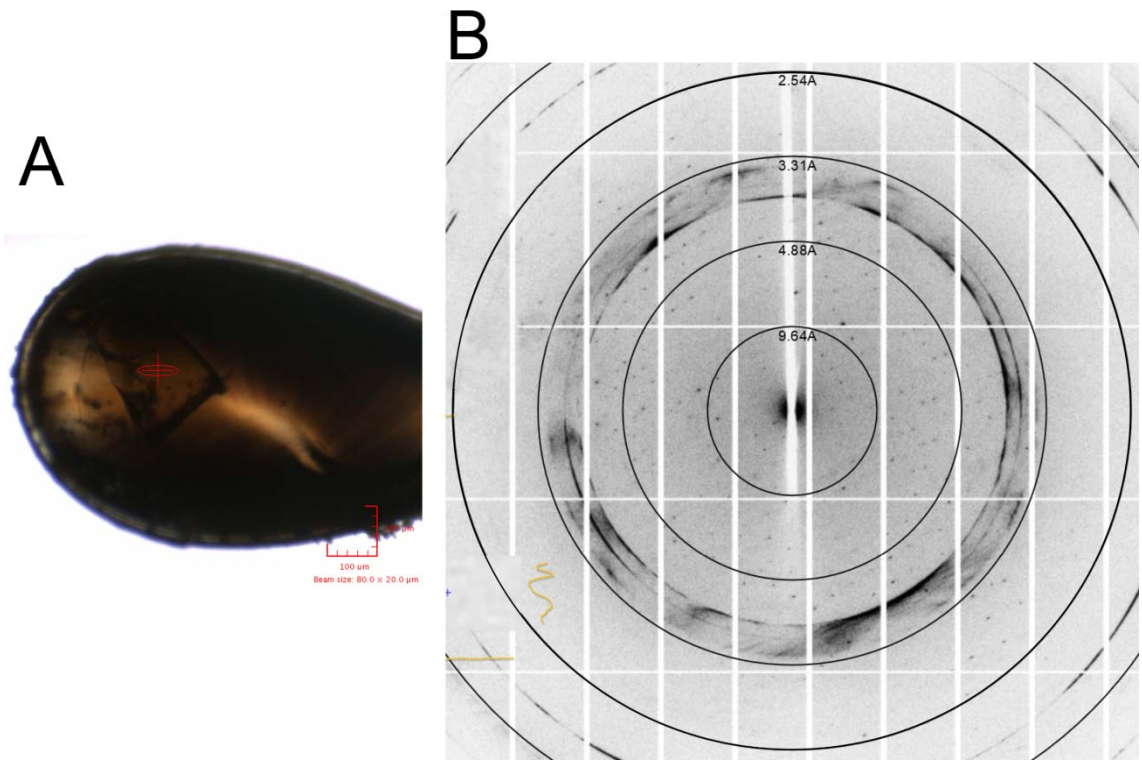



**Figure A21** Testing of *Te*LipA2 crystals reconstituted with ESDK(o)AD and MTAN. **A:** Associated diffraction pattern from 12 keV synchrotron radiation to 1.9 Å resolution. **B:** Loop containing a crystal of *TeLipA2* reconstituted with 5 molar equivalents of $FeCl_3$/$Na_2Se$, grown in the presence of SAM (0.8 mM LipA, 10 mM SAM, 5.0 mM ESDK(o)AD, 100 µM MTAN), directly frozen after formation in condition H12 of PEGRx HT Hampton screen (0.1M Bicine, pH 8.5; 3% w/v dextran sulfate sodium salt, 15% w/v PEG 20,000). Loop size – 0.4–0.5 mm, Beam size – 80 x 20 µm.

**Table A24** Data collection information for the *Te*LipA2 crystals grown with peptide substrate analogue ESDK(o)AD as well as MTAN.

| | ***Te*LipA 2** |
|---|---|
| **Substrate** | SAM + ESDK(o)AD + MTAN |
| **Crystallization condition** | Hampton PEGRx H12 |
| **Data collection** | |
| **Space group** | C 2 2 2 |
| **Cell dimensions** | |
| **a, b, c (Å)** | 73.73; 164.76; 59.58 |
| **α, β, γ (°)** | 90.0; 90.0; 90.0 |
| **Wavelength (Å)** | 0.97625 |
| **No of reflections** | 14787 (1140-154) |
| **No of unique reflections** | 4528 (343-61) |
| **Resolution (Å)** | 59.58-3.55 (15.88-3.55) |
| **$R_{sym}$ or $R_{merge}$** | 0.121 (0.342-0.025) |
| **I / σI** | 6.8 (22.0-2.9) |
| **Completeness (%)** | 97.3 (99.5-90.6) |
| **Redundancy** | 3.3 (3.3-2.5) |
| **Twinning Z-score** | 2.93, no twinning |

| | ***Te*LipA2** |
|---|---|
| **Current R-work** | 0.278 |
| **Current R-free** | 0.322 |

**Table A25** Score values for molecular replacement using *Phaser*.

| | ***Te*LipA2 + ESDK(o)AD + MTAN** |
|---|---|
| LL G-score (best) | 3173.10 |
| TF Z-score (best) | 50.9 |

## A.19 Crystallization of *Te*LipA 2 Ser283Cys

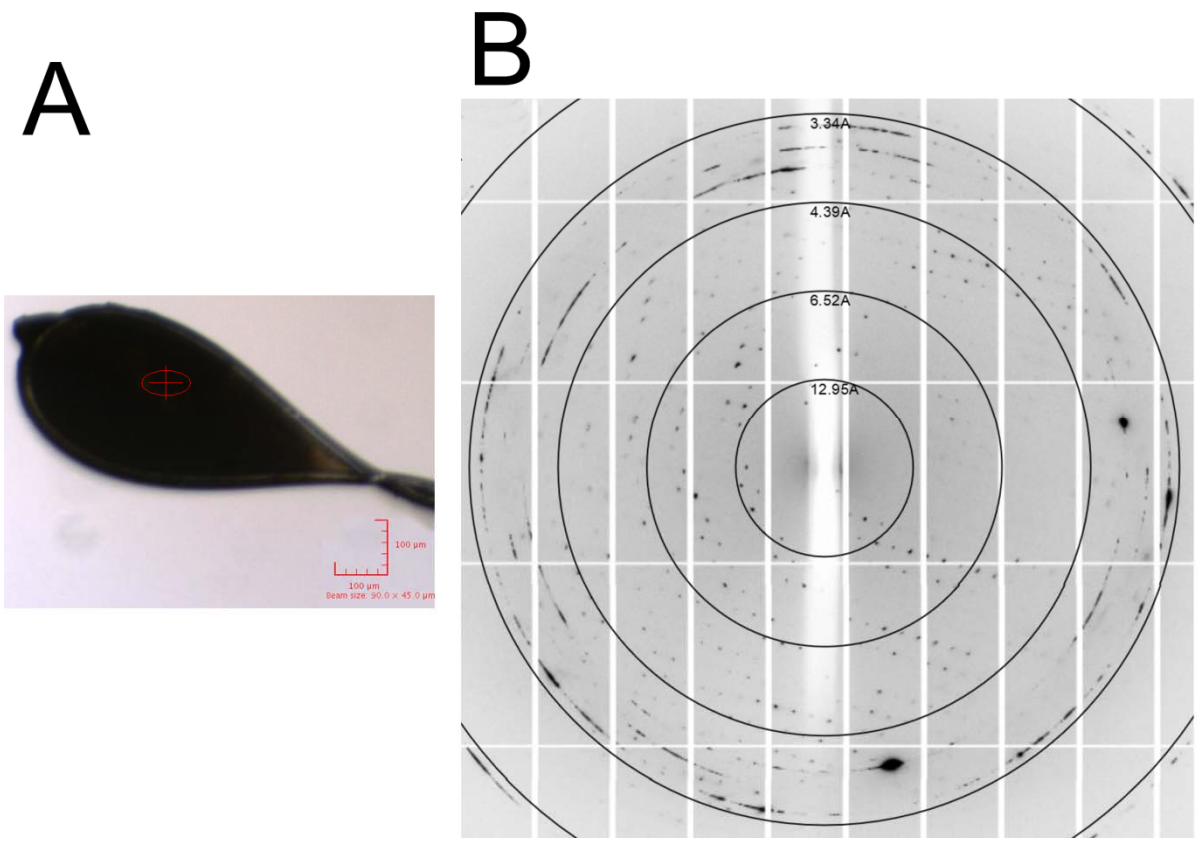



**Figure A22** Testing of *Te*LipA2 Ser283Cys crystal. **A:** Loop containing a crystal of *TeLipA2* Ser283Cys (0.8 mM LipA, 10 mM SAM), directly frozen after formation in condition H12 of PEGRx HT Hampton screen (0.1M Bicine, pH 8.5; 3% w/v dextran sulfate sodium salt, 15% w/v PEG 20,000). Loop size – 0.4–0.5 mm, Beam size – 90 x 45 µm. **b:** Associated diffraction pattern from 12 keV synchrotron radiation to 3.0 Å resolution.

**Table A26** Data collection information for the *Te*LipA2 Ser283Cys crystals. The refinement process is ongoing and the current $R_{work}$ and $R_{free}$ values are presented at the bottom of the table.

| | ***Te*LipA2 Ser283Cys** |
|---|---|
| **Substrate** | SAM |
| **Crystallization condition** | Hampton PEGRx H12 |
| **Data collection** | |
| **Space group** | C 2 2 $2_1$ |
| **Cell dimensions** | |
| **a, b, c (Å)** | 72.51; 162.48; 58.59 |
| **α, β, γ (°)** | 90.0; 90.0; 90.0 |
| **Wavelength (Å)** | 0.97625 |
| **No of reflections** | 24849 (2109-300) |
| **No of unique reflections** | 6549 (497-89) |
| **Resolution (Å)** | 43.39-3.02 (13.51-3.02) |
| **$R_{sym}$ or $R_{merge}$** | 0.059 (0.603-0.028) |
| **I / σI** | 12.3 (29.7-2.4) |
| **Completeness (%)** | 92.6 (98.1-91.8) |
| **Redundancy** | 3.8 (4.2-3.4) |
| **Twinning Z-score** | 3.34, no twinning |

| | ***Te*LipA2 Ser283Cys** |
|---|---|
| **$R_{work}$ current** | 0.313 |
| **$R_{free}$ current** | 0.364 |

**Table 20 Score values for molecular replacement using *Phaser*.**

| | ***Te*LipA2 + ESDK(o)AD + MTAN** |
|---|---|
| LL G-score (best) | 3173.10 |
| TF Z-score (best) | 50.9 |

## A.20 Gene encoding the Cfr-substrate rRNA Fragment

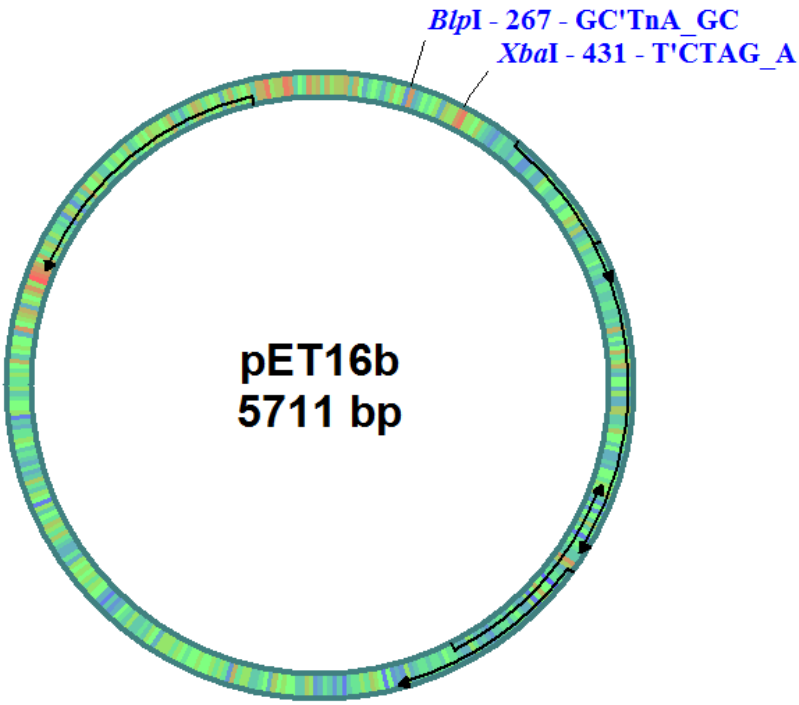



**Figure A23** Pet-16b plasmid map, showing BlpI and XbaI sites, where the gene of interest was inserted.

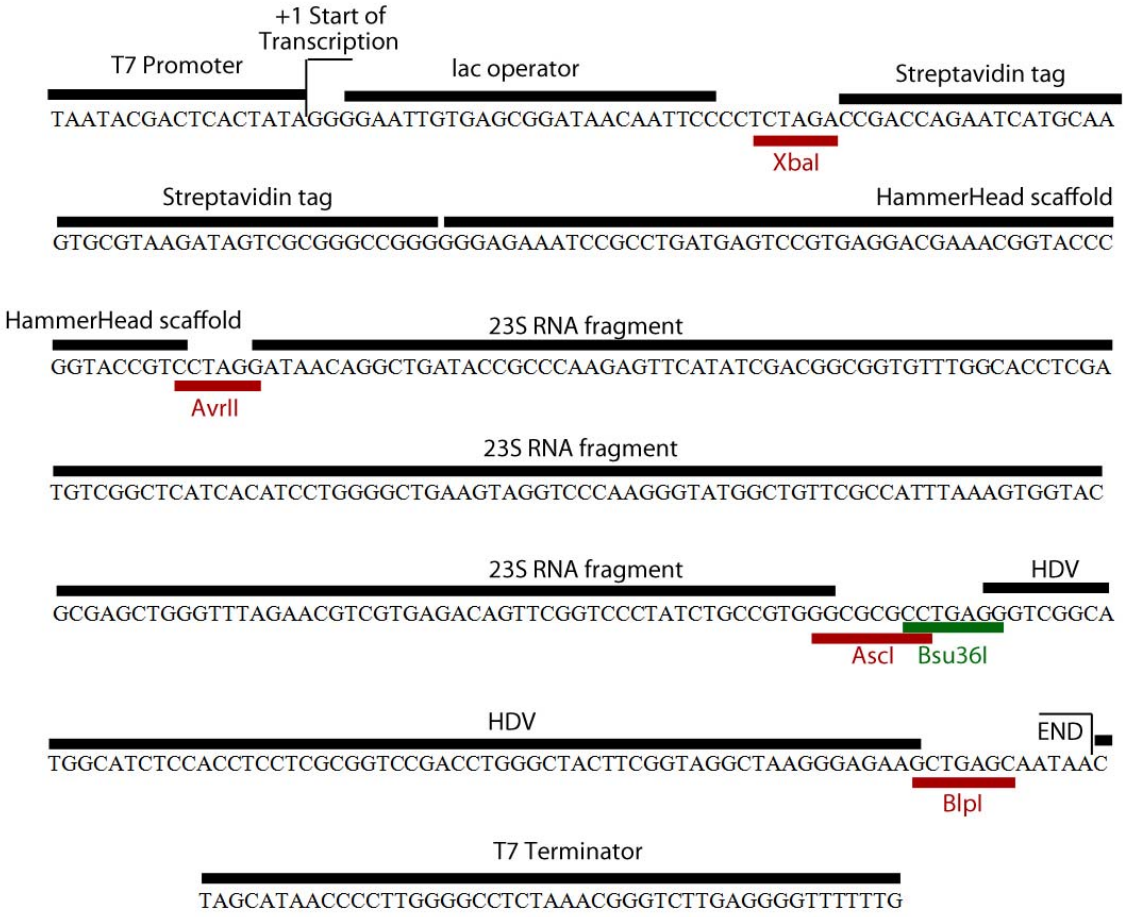



**Figure A24** Pet-16b Sequence for Transcription of an rRNA Fragment

## A.21 Plasmid map of *cfr* containing pCDFD1 vector

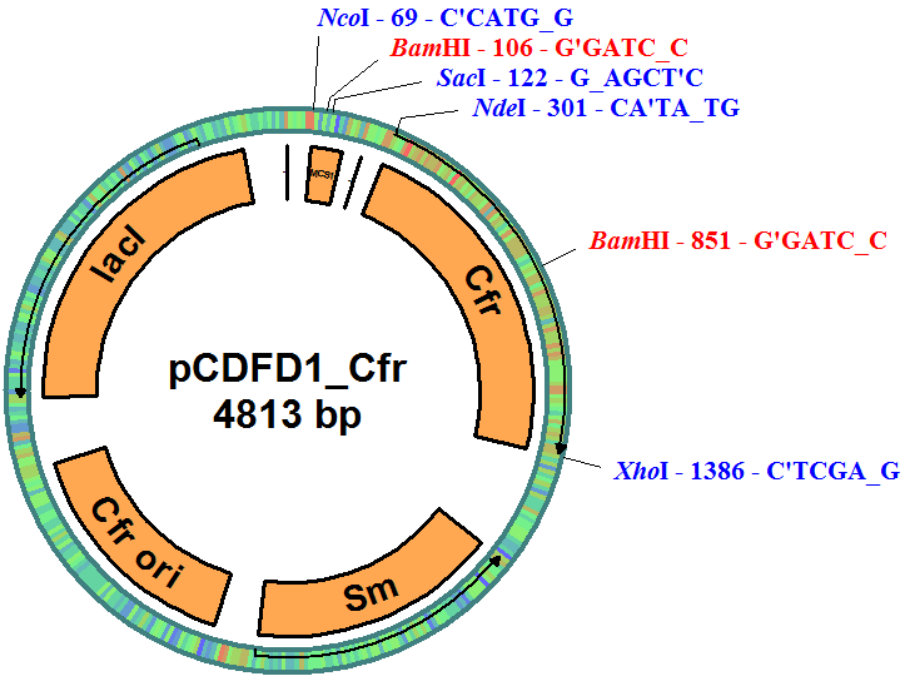



**Figure A25** pCDFDuet-1 vector, containing the *cfr* gene from *Staphylococcus sciuri*, tagged N-terminally by a 6His fragment for affinity purification.

**Nucleotide sequence for the *Sscfr* gene present in plasmid pCDFD1_Cfr**

ATGCATCACCATCACCATCACAAAGAAATGAATTTTAATAATAAAACAAAGTATGGTAAAATACAGGAATTT
TTAAGAAGTAATAATGAGCCTGATTATAGAATAAAACAAATAACCAATGCGATTTTTAAACAAAGAATTAG
TCGATTTGAGGATATGAAGGTTCTTCCAAAATTACTTAGGGAGGATTTAATAAATAATTTTGGAGAAACAG
TTTTGAATATCAAGCTCTTAGCAGAGCAAAATTCAGAGCAAGTTACGAAAGTGCTTTTTGAAGTATCAAAG
AATGAGAGAGTAGAAACGGTAAACATGAAGTATAAAGCAGGTTGGGAGTCATTTTGTATATCATCACAAT
GCGGATGTAATTTTGGGTGTAAATTTTGTGCTACAGGCGACATTGGATTGAAAAAAAACCTAACTGTAGAT
GAGATAACAGATCAAGTTTTATACTTCCATTTATTAGGTCATCAAATTGATAGCATTTCTTTTATGGGAATG
GGTGAAGCTCTAGCCAACCGTCAAGTATTTGATGCTCTTGATTCGTTTACGGATCCTAATTTATTTGCATTAA
GTCCTCGTAGACTTTCTATATCAACGATTGGTATTATACCTAGTATCAAAAAAATAACCCAGGAATATCCTC
AAGTAAATCTTACATTTTCATTACACTCACCTTATAGTGAGGAACGCAGCAAATTGATGCCAATAAATGATA
GATACCCAATAGATGAGGTAATGAATATACTCGATGAACATATAAGATTAACTTCAAGGAAAGTATATATA

GATTATATCATGTTGCCTGGTGTAAATGATTCTCTTGAGCATGCAAACGAAGTTGTTAGCCTTCTTAAAAGT
CGCTATAAATCAGGGAAGTTATATCATGTAAATTTGATACGATACAATCCTACAATAAGTGCACCTGAGAT
GTATGGAGAAGCAAACGAAGGGCAGGTAGAAGCCTTTTACAAAGTTTTGAAGTCTGCTGGTATCCATGTC
ACAATTAGAAGTCAATTTGGGATTGATATTGACGCTGCTTGTGGTCAATTATATGGTAATTATCAAAATAGC
CAATAGCTCTAGC

***Ss*Cfr amino acid sequence, with [4Fe-4S] coordinating cysteines in red**

MHHHHHHKEMNFNNKTKYGKIQEFLRSNNEPDYRIKQITNAIFKQRISRFEDMKVLPKLLREDLINNFGETVLNI
KLLAEQNSEQVTKVLFEVSKNERVETVNMKYKAGWESFCISSQCGCNFGCKFCATGDIGLKKNLTVDEITDQVL
YFHLLGHQIDSISFMGMGEALANRQVFDALDSFTDPNLFALSPRRLSISTIGIIPSIKKITQEYPQVNLTFSLHSPYSE
ERSKLMPINDRYPIDEVMNILDEHIRLTSRKVYIDYIMLPGVNDSLEHANEVVSLLKSRYKSGKLYHVNLIRYNPTIS
APEMYGEANEGQVEAFYKVLKSAGIHVTIRSQFGIDIDAACGQLYGNYQNSQ

# List of References


1. Buchner, E., Rapp, R., *Alkoholische gährung ohne hefezellen.* Ber Deut Chem Ges, 1897. **30**(3): p. 2668-2678.

2. Reichard, P., Ehrenberg, A. , *Ribonucleotide reductase--a radical enzyme.* Science, 1983. **221**: p. 514-519.

3. Gowenlock, B., Snelling, D., *Free radicals in organic chemistry.* Adv. Chem. Serv., 1962. **36**: p. 150.

4. Buckel, W., Golding, B., *Radical enzymes in anaerobes.* Annu Rev Microbiol, 2006. **60**: p. 27-49.

5. Sancar, A., *Structure and function of DNA photolyase.* Biochemistry, 1994. **33**(1): p. 2-9.

6. Eklund, H., Fontecave, M., *Glycyl radical enzymes: a conservative structural basis for radicals.* Structure, 1999. **7**(11): p. R257-R262.

7. Logan, D., Andersson, J., Sjöberg, B., Nordlund, P., *A glycyl radical site in the crystal structure of a class III ribonucleotide reductase.* Science, 1999. **283**(5407): p. 1499-1504.

8. Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S., Wagner, A., *Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase.* Nat Struct Biol, 1999. ***6***(10): p. 969-975.

9. Hsi, L., Hoganson, C., Babcock, G., Smith, W., *Characterization of a tyrosyl radical in prostaglandin endoperoxide synthase-2.* Biochem Bioph Res Co, 1994. **202**(3): p. 1592-1598.

10. Heinecke, J., Li, W., Francis, G., Goldstein, J., *Tyrosyl radical generated by myeloperoxidase catalyzes the oxidative cross-linking of proteins.* J Clin Invest, 1993. **91**(6): p. 2866-2872.

11. Stubbe, J., Riggs-Gelasco, P., *Harnessing free radicals: formation and function of the tyrosyl radical in ribonucleotide reductase.* Trends Biochem Sci, 1998. **23**(11): p. 438-443.

12. Persson, A., Sahlin, M., Sjöberg, B., *Cysteinyl and substrate radical formation in active site mutant E441Q of Escherichia coli class I ribonucleotide reductase.* J Biol Chem, 1998. **273**(47): p. 31016-31020.

13. Licht, S., Gerfen, G., Stubbe, J., *Thiyl radicals in ribonucleotide reductases.* Science, 1996. **271**(5248): p. 477-481.

14. Cordes, M., Giese, B., *Electron transfer in peptides and proteins.* Chem Soc Rev, 2009. **38**(4): p. 892-901.

15. Fontecave, M., *Ribonucleotide reductases and radical reactions.* Cell Mol Life Sci, 1998. **54**(7): p. 684-695.

16. Blackburn, R., Kyaw, M., Phillips, G., Swallow, A., *Free radical reactions in the coenzyme B12 system.* J Chem Soc, Faraday transactions 1: Physical chemistry in condensed phases, 1975. **71**: p. 2277-2287.

17. Halpern, J., *Mechanisms of coenzyme B12-dependent rearrangements.* Science, 1985. **227**(4689): p. 869-875.

18. Neil, E., Marsh, G., *Coenzyme B12 (cobalamin)-dependent enzymes.* Essays Biochem, 1999. **34**: p. 139-154.

19. Bucher, D., Sandala, G., Durbeej, B., Radom, L., Smith, D., *The elusive 5'-deoxyadenosyl radical in coenzyme-B12-mediated reactions.* J Am Chem Soc, 2012. **134**(3): p. 1591-1599.

20. Sofia, H., Chen, G., Hetzler, B., Reyes-Spindola, J., Miller, N., *Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods.* Nucleic Acids Res, 2001. **29**(5): p. 1097-1106.

21. Frey, P., Hegeman, A., Ruzicka, F., *The radical SAM superfamily.* Crit Rev Biochem Mol Biol, 2008. **43**(1): p. 63-88.

22. Vey, J., Drennan, C., *Structural insights into radical generation by the radical SAM superfamily.* Chem Rev, 2011. **111**(4): p. 2487-2506.

23. Wang, S.,Frey, P., *S-adenosylmethionine as an oxidant: the radical SAM superfamily.* Trends Biochem Sci, 2007. **32**(3): p. 101-110.

24. Booker, S., *Radical SAM enzymes and radical enzymology.* Biochim Biophys Acta, 2012. **1824**(11): p. 1151-1153.

25. Roach, P., *Radicals from S-adenosylmethionine and their application to biosynthesis.* Curr Opin Chem Biol, 2011. **15**(2): p. 267-275.

26. Frey, P., *Radical mechanisms of enzymatic catalysis.* Annu Rev Biochem, 2001. **70**: p. 121-148.

27. Cheek, J., Broderick, J., *Direct H atom abstraction from spore photoproduct C-6 initiates DNA repair in the reaction catalyzed by spore photoproduct lyase: evidence for a reversibly generated adenosyl radical intermediate.* J Am Chem Soc, 2002. **124**(12): p. 2860-2861.

28. Marsh, E., Patwardhan, A., Huhta, M., *S-adenosylmethionine radical enzymes.* Bioorg Chem, 2004. **32**(5): p. 326-340.

29. Tamarit, J., Mulliez, E., Meier, C., Trautwein, A.,Fontecave, M, *The anaerobic ribonucleotide reductase from Escherichia coli.* J Biol Chem, 1999. **274**(44): p. 31291-31296.

30. Henshaw, T., Cheek, J., Broderick, J., *The [4Fe-4S] 1+ cluster of pyruvate formate-lyase activating enzyme generates the glycyl radical on pyruvate formate-lyase: EPR-detected single turnover.* J Am Chem Soc, 2000. **122**(34): p. 8331-8332.

31. McCusker, K., Medzihradszky, K., Shiver, A., Nichols, R., Yan, F., Maltby, D., Gross, C., Fujimori, D., *Covalent intermediate in the catalytic mechanism of the radical S-adenosyl-L-methionine methyl synthase RlmN trapped by mutagenesis.* J Am Chem Soc, 2012. **134**(43): p. 18074-18081.

32. Demick, J., Lanzilotta,W., *Radical SAM activation of the B(12)-independent glycerol dehydratase results infFormation of 5'-deoxy-5'-(methylthio)adenosine and not 5'-deoxyadenosine.* Biochemistry, 2011. **50**(4): p. 440-442.

33. Frey, P., *Lysine 2,3-aminomutase: is adenosylmethionine a poor man's adenosylcobalamin?* FASEB J, 1993. **7**(8): p. 662-670.

34. Wilson, D., Madera, M., Vogel, C., Chothia, C., Gough, J., *The SUPERFAMILY database in 2007: families and functions.* Nucleic Acids Res, 2007. **35**(suppl 1): p. D308-D313.

35. Shisler, K., Broderick, J., *Emerging themes in radical SAM chemistry.* Curr Opin Struc Biol, 2012. **22**(6): p. 701-710.

36. Bauerle, M., Schwalm, E., Booker, S., *MechanisticdDiversity of radical SAM-dependent methylation.* J Biol Chem, 2014: p. jbc-R114.

37. Forouhar, F., Arragain, S., Atta, M., Gambarelli, S., Mouesca, J., Hussain, M., Xiao, R., Kieffer-Jaquinod, S., Seetharaman, J., Acton, T., Montelione, G., *Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases.* Nat Chem Biol, 2013. **9**(5): p. 333-338.

38. Flühe, L., Knappe, T., Gattner, M., Schäfer, A., Burghaus, O., Linne, U., Marahiel, M. A., *The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A.* Nat Chem Biol, 2012. **8**(4): p. 350-357.

39. Booker, S., Cicchillo, R., Grove, T., *Self-sacrifice in radical S-adenosylmethionine proteins.* Curr Opin Chem Biol, 2007. **11**(5): p. 543-552.

40. Perche-Letuvée, P., Kathirvelu, V., Berggren, G., Clemancey, M., Latour, J., Maurel, V., Garcia-Serres, R., *4-Demethylwyosine synthase from Pyrococcus abyssi is a radical-S-adenosyl-L-methionine enzyme with an additional [4Fe-4S](+2) cluster that interacts with the pyruvate co-substrate.* J Biol Chem, 2012. **287**(49): p. 41174-41185.

41. Dowling, D., Bruender, N., Young, A., McCarty, R., Bandarian, V., Drennan, C., *Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism.* Nat Chem Biol, 2014. **10**(2): p. 106-112.

42. Layer, G., Pierik, A., Trost, M., Rigby, S., Leech, H., Grage, K., Jahn, D., *The substrate radical of Escherichia coli oxygen-independent coproporphyrinogen III oxidase HemN.* J Biol Chem, 2006. **281**(23): p. 15727-15734.

43. Mahanta, N., Fedoseyenko, D., Dairi, T., Begley, T., *Menaquinone biosynthesis: Formation of aminofutalosine requires a unique radical SAM enzyme.* J Am Chem Soc, 2013. **135**(41): p. 15318-15321.

44. Frey, M., Rothe, M., Wagner, A., Knappe, J., *Adenosylmethionine-dependent synthesis of the glycyl radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom. Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site.* J Biol Chem, 1994. **269**(17): p. 12432-12437.

45. Gambarelli, S., Luttringer, F., Padovani, D., Mulliez, E., Fontecave, M., *Activation of the Anaerobic Ribonucleotide Reductase by S-Adenosylmethionine.* Chem Bio Chem, 2005. **6**(11): p. 1960-1962.

46. Challand, M., Martins,F., Roach, P., *Catalytic activity of the anaerobic tyrosine lyase required for thiamine biosynthesis in Escherichia coli.* J Biol Chem, 2010. **285**(8): p. 5240-5248.

47. Driesener, R., Challand, M., McGlynn, S., Shepard, E., Boyd, E., Broderick, J., Roach, P., *[FeFe]-hydrogenase cyanide ligands derived from S-adenosylmethionine-dependent cleavage of tyrosine.* Angew Chem Int Ed Engl, 2010. **49**(9): p. 1687-1690.

48. Hiscox, M., Driesener,R., Roach,P., *Enzyme catalyzed formation of radicals from S-adenosylmethionine and inhibition of enzyme activity by the cleavage products.* Biochim Biophys Acta, 2012. **1824**(11): p. 1165-1177.

49. Atta, M., Arragain, S., Forouhar, F., Hunt, J., Fontecave, M., *S-Adenosylmethionine-dependent radical-based modification of biological macromolecules.* Curr Opin Struc Biol, 2010. **20**(6): p. 684–692.

50. Layer, G., Heinz, D., Jahn, D., Schubert, W., *Structure and function of radical SAM enzymes.* Curr Opin Chem Biol, 2004. **8**(5): p. 468-476.

51. Challand, M., Driesener, R., Roach, P., *Radical S-adenosylmethionine enzymes: Mechanism, control and function.* Nat Prod Rep, 2011. **28**(10): p. 1696-1721.

52. Mehta, A., Abdelwahed, S., Mahanta, N., Fedoseyenko, D., Philmus, B., Cooper, L., Begley, T., *Radical SAMeEnzymes in cofactor biosynthesis: a treasure trove of complex organic radical rearrangement reactions.* J Biol Chem, 2014: p. jbc-R114.

53. Wang, J., Woldring, R., Román-Meléndez, G., McClain, A., Alzua, B., Marsh, E., *Recent advances in radical SAM enzymology: new structures and mechanisms.* ACS Chem Biol, 2014. **9**(9): p. 1929-1938.

54. Chirpich, T., Zappia, V., Costilow, R., Barker, H., *Lysine 2,3-aminomutase. Purification and properties of a pyridoxal phosphate and S-adenosylmethionine-activated enzyme.* J Biol Chem, 1970. **245**(7): p. 1778-1789.

55. Knappe, J., Schmitt,T., *A novel reaction of S-adenosyl-L-methionine correlated with the activation of pyruvate formate-lyase.* Biochem Bioph Res Co, 1976. **71**(4): p. 1110-1117.

56. Eliasson, R., Fontecave, M., Jörnvall, H., Krook, M., Pontis, E., Reichard, P., *The anaerobic ribonucleoside triphosphate reductase from Escherichia coli requires S-adenosylmethionine as a cofactor.* P Natl Acad Sci, 1990. **87**(9): p. 3314-3318.

57. Sanyal, I., Cohen, G., Flint, D., *Biotin synthase: purification, characterization as a [2Fe-2S] cluster protein, and in vitro activity of the Escherichia coli bioB gene product.* Biochemistry, 1994. **33**(12): p. 3625-3631.

58. Miller, J., Busby, R., Jordan, S., Cheek, J., Henshaw, T., Ashley, G., Marletta, M., *Escherichia coli LipA is a lipoyl synthase: in vitro biosynthesis of lipoylated pyruvate dehydrogenase complex from octanoyl-acyl carrier protein.* Biochemistry, 2000. **39**(49): p. 15166-15178.

59. Duin, E., Lafferty, M., Crouse, B., Allen, R., Sanyal, I., Flint, D. H., Johnson, M., *[2Fe-2S] to [4Fe-4S] cluster conversion in Escherichia coli biotin synthase.* Biochemistry, 1997. **36**(39): p. 11811-11820.

60. Külzer, R.P., T., Kappl, R., Hüttermann, J., Knappe J., *Reconstitution and characterization of the polynuclear iron-sulfur cluster in pyruvate formate-lyase-activating enzyme. Molecular properties of the holoenzyme form.* J Biol Chem, 1998. **273**(9): p. 4897–4903.

61. Sofia, H., Hetzler, B., Reyes-Spindola, J., Miler, N., *Self-sacrifice in radical S-adenosylmethionine proteins.* Nucleic Acids Res, 2001. **29**(5): p. 1097-1106.

62. Martin, J., McMillan, F., *SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold.* Curr Opin Struc Biol, 2002. **12**(6): p. 783-793.

63. Johnson, D., Dean, D., Smith, A., Johnson, M., *Structure, function, and formation of biological iron-sulfur clusters.* Annu Rev Biochem, 2005. **74**: p. 247-281.

64. Kiley, P., Beinert, H., *The role of Fe–S proteins in sensing and regulation in bacteria.* Curr Opin Microbiol, 2003. **6**(2): p. 181–185.

65. Haile, D., Tang, C., Chin, J., Harford, J., Klausner, R., *Reciprocal control of RNA-binding and aconitase activity in the regulation of the iron-responsive element binding protein: role of the iron-sulfur cluster.* P Natl Acad Sci, 1992. **89**(16): p. 7536-7540.

66. Harrison, P., Arosio, P., *The ferritins: molecular properties, iron storage function and cellular regulation.* Biochim Biophys Acta, 1996. **1275**(3): p. 161-203.

67. Rouault, T.K., R., *Iron-sulfur clusters as biosensors of oxidants and iron.* Trends Biochem Sci, 1996. **21**(5): p. 174-177.

68. Crack, J., Green, J., Thomson, A., *Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR).* J Biol Chem, 2004. **279**(10): p. 9278-9286.

69. Hidalgo, E., Demple, B., *An iron-sulfur center essential for transcriptional activation by the redox-sensing SoxR protein.* EMBO J, 1994. **13**(1): p. 138-146.

70. Camba, R., Hunsicker-Wang, L., Burgess, B., Stout, C., Hirst, J., Armstrong, F., *Mechanisms of redox-coupled proton transfer in proteins: role of the proximal proline in reactions of the [3Fe-4S] cluster in Azotobacter vinelandii ferredoxin I.* Biochemistry, 2003. **42**(36): p. 10589-10599.

71. Henderson, R., *Proton transfer to synthetic [Fe-S] based clusters.* Coordin Chem Rev, 2005. **249**(17): p. 1841-1856.

72. Walters, E., Johnson, M., *Ferredoxin:thioredoxin reductase: disulfide reduction catalyzed via novel site-specific [4Fe-4S]cluster.* Chemistry Photosynth Res., 2004. **79**(3): p. 249-264.

73. Beinert, H.H., R., Münck, E., *Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures.* Science, 1997. **277**(5326): p. 653-659.

74. Fontecave, M., *Iron-sulfur clusters: ever-expanding roles.* Nat Chem Biol, 2006. **2**(4): p. 171-174.

75. Layer, G., Moser, J., Heinz, D., Jahn, D., Schubert, W., *Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes.* EMBO J, 2003. **22**(23): p. 6214-6224.

76. Tse Sum Bui, B., D., Fournier, F., Ploux, O., Méjean, A., Marquet, A., *Biotin synthase mechanism: on the origin of sulfur.* FEBS letters, 1998. **440**(1): p. 226-230.

77. Gibson, K., Pelletier, D.. Turner, I., *Transfer of sulfur to biotin from biotin synthase (BioB protein).* Biochem Bioph Res Co, 1999. **254**(3): p. 632-635.

78. Pierrel, F., Douki, T., Fontecave, M., Atta, M., *MiaB protein is a bifunctional radical-S-adenosylmethionine enzyme involved in thiolation and methylation of tRNA.* J Biol Chem, 2004. **279**(46): p. 47555-47563.

79. Lee, K., Saleh, L., Anton, B., Madinger, C., Benner, J., Iwig, D., Booker, S., *Characterization of RimO, a new member of the methylthiotransferase subclass of the radical SAM superfamily.* Biochemistry, 2009. **48**(42): p. 10162-10174.

80. Rubach, J., Brazzolotto, X., Gaillard, J., Fontecave, M., *Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from Thermatoga maritima.* FEBS Letters, 2005. **579**(22): p. 5055-5060.

81. Posewitz, M., King, P., Smolinski, S., Zhang, L., Seibert, M., hirardi, M., *Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase.* J Biol Chem, 2004. **279**(24): p. 25711-25720.

82. Allen, R., Chatterjee, R., Ludden, P., Shah, V., *Incorporation of iron and sulfur from NifB cofactor into the iron-molybdenum cofactor of dinitrogenase.* J Biol Chem, 1995. **270**(45): p. 26890-26896.

83. Yan, F., LaMarre, J., Röhrich, R., Wiesner, J., Jomaa, H., Mankin, A., Fujimori, D., *RlmN and Cfr are radical SAM enzymes involved in methylation of ribosomal RNA.* J Am Chem Soc, 2010. **132**(11): p. 3953-3964.

84. Leonardi, R. Roach, P., *Thiamine biosynthesis in Escherichia coli: in vitro reconstitution of the thiazole synthase activity.* J Biol Chem, 2004. **279**(17): p. 17054-62.

85. Martinez-Gomez, N., Downs, D., *ThiC is an [Fe-S] cluster protein that requires AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in thiamin synthesis.* Biochemistry, 2008. **47**(35): p. 9054-9056.

86. Booker, S., Grove, T., *Mechanistic and functional versatility of radical SAM enzymes.* F1000 biology reports, 2010: p. B2-B52.

87. Nicolet, Y., Amara, P., Mouesca, J., Fontecilla-Camps, J., *Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins.* Proc Natl Acad Sci, 2009. **106**(35): p. 14867-14871.

88. Padovani, D., Thomas, F., Trautwein, A., Mulliez, E., Fontecave, M, *Activation of class III ribonucleotide reductase from E. coli. The electron transfer from the iron-sulfur center to S-adenosylmethionine.* Biochemistry, 2001. **40**(23): p. 6713-6719.

89. Walsby, C., Hong, W., Broderick, W., Cheek, J., Ortillo, D., Broderick, J., Hoffman, B., *Electron-nuclear double resonance spectroscopic evidence that S-adenosylmethionine binds in contact with the catalytically active [4Fe-4S](+) cluster of pyruvate formate-lyase activating enzyme.* J Am Chem Soc, 2002. **124**(12): p. 3143-3151.

90. Daley, C., Holm, R., *Reactivity of [Fe4S4 (SR) 4] 2-, 3-clusters with sulfonium cations: Analogue reaction systems for the initial step in biotin synthase catalysis.* Inorg Chem, 2001. **40**(12): p. 2785-2793.

91. Ugulava, N., Frederick, K., Jarrett, J., *Control of adenosylmethionine-dependent radical generation in biotin synthase: a kinetic and thermodynamic analysis of substrate binding to active and inactive forms of BioB.* Biochemistry, 2003. **42**(9): p. 2708-2719.

92. Hinckley, G., Frey,P., *Cofactor dependence of reduction potentials for [4Fe-4S]2+/1+ in lysine 2,3-aminomutase.* Biochemistry, 2006. **45**(10): p. 3219-3225.

93. Wang, S., Frey,P., *Binding energy in the one-electron reductive cleavage of S-adenosylmethionine in lysine 2,3-aminomutase, a radical SAM enzyme.* Biochemistry, 2007. **46**(45): p. 12889-12895.

94. Duschene, K., Veneziano, S., Silver, S., Broderick, J., *Control of radical chemistry in the AdoMet radical enzymes.* Curr Opin Chem Biol, 2009. **13**(1): p. 74-83.

95. Cosper, M., Jameson, G., Davydov, R., Eidsness, M., Hoffman, B., Huynh, B., Johnson, M., *The [4Fe-4S](2+) cluster in reconstituted biotin synthase binds S-adenosyl-L-methionine.* J Am Chem Soc, 2002. **124**(47): p. 14006-14007.

96. Krebs, C., Broderick, W., Henshaw, T., Broderick, J., Huynh, B., *Coordination of adenosylmethionine to a unique iron site of the [4Fe-4S] of pyruvate formate-lyase activating enzyme: a Mossbauer spectroscopic study.* J Am Chem Soc, 2002. **124**(6): p. 912-913.

97. Ollagnier, S., Mulliez, E., Schmidt, P., Eliasson, R., Gaillard, J., Deronzier, C., Fontecave, M., *Activation of the anaerobic ribonucleotide reductase from Escherichia coli. The essential role of the iron-sulfur center for S-adenosylmethionine reduction.* J Biol Chem, 1997. **272**(39): p. 24216-24223.

98. Cosper, N., Booker, S., Ruzicka, F., Frey, P., Scott, R. , *Direct FeS cluster involvement in generation of a radical in lysine 2,3-aminomutase.* Biochemistry, 2000. **39**(51): p. 15668-15673.

99. Jarrett, J., *The generation of 5'-deoxyadenosyl radicals by adenosylmethionine-dependent radical enzymes.* Curr Opin Chem Biol, 2003. **7**(2): p. 174-182.

100. Dey, A., Peng, Y., Broderick, W. E., Hedman, B., Hodgson, K., Broderick, J., Solomon, E., *S K-edge XAS and DFT calculations on SAM dependent pyruvate formate-lyase activating enzyme: nature of interaction between the Fe4S4 cluster and SAM and its role in reactivity.* J Am Chem Soc, 2011. **133**(46): p. 18656-18662.

101. Cosper, M., Cosper, N., Hong, W., Shokes, J., Broderick, W., Broderick, J., Scott, R., *Structural studies of the interaction of S-adenosylmethionine with the [4Fe-4S] clusters in biotin synthase and pyruvate formate-lyase activating enzyme.* Protein Sci, 2003. **12**(7): p. 1573-1577.

102. Liu, A. and A. Graslund, *Electron paramagnetic resonance evidence for a novel interconversion of [3Fe-4S](+) and [4Fe-4S](+) clusters with endogenous iron and sulfide in anaerobic ribonucleotide reductase activase in vitro.* J Biol Chem, 2000. **275**(17): p. 12367-12373.

103. Krebs, C., Henshaw, T., Cheek, J., Huynh, B., Broderick, J., *Conversion of 3Fe-4S to 4Fe-4S clusters in native pyruvate formate-lyase activating enzyme: Mössbauer characterization and implications for mechanism.* J Am Chem Soc, 2000. **122**(50): p. 12497-12506.

104. Hewitson, K. , Ollagnier-de Choudens, S., Sanakis, Y., Shaw, N., Baldwin, J., Münck, E., Fontecave, M., *The iron-sulfur center of biotin synthase: site-directed mutants.* J Biol Inorg Chem, 2002. **7**(1-2): p. 83-93.

105. Hewitson, K., Baldwin, J., Shaw, N., Roach, P., *Mutagenesis of the proposed iron-sulfur cluster binding ligands in Escherichia coli biotin synthase.* FEBS Lett, 2000. **466**(2-3): p. 372-376.

106. Boutigny, S., Saini, A., Baidoo, E., Yeung, N., Keasling, J., Butland, G., *Physical and functional interactions of a monothiol glutaredoxin and an iron sulfur cluster carrier protein with the sulfur-donating radical S-adenosyl-L-methionine enzyme MiaB.* J Biol Chem, 2013. **288**(20): p. 14200-14211.

107. Jiang, D., Guo, H., Xu, C., Chang, J., Gu, B., Wang, L., Guo, J., *Identification of three interferon-inducible cellular enzymes that inhibit the replication of hepatitis C virus.* J Virol, 2008. **82**(4): p. 1665-1678.

108. Kamat, S., Williams, H., Dangott, L., Chakrabarti, M., Raushel, F., *The catalytic mechanism for aerobic formation of methane by bacteria.* Nature, 2013. **497**(7447): p. 132-136.

109. Fajardo-Cavazos, P., Rebeil, R., Nicholson, W., *Essential cysteine residues in Bacillus subtilis spore photoproduct lyase identified by alanine scanning mutagenesis.* Curr Microbiol, 2005. **51**(5): p. 331-335.

110. DeFraia, C., Wang, Y., Yao, J., Mou, Z., *Elongator subunit 3 positively regulates plant immunity through its histone acetyltransferase and radical S-adenosylmethionine domains.* BMC Plant Biol, 2013. **13**(1): p. 102-115.

111. Zhu, X., Dzikovski, B., Su, X., Torelli, A., Zhang, Y., Ealick, S., Lin, H., *Mechanistic understanding of Pyrococcus horikoshii Dph2, a [4Fe-4S] enzyme required for diphthamide biosynthesis.* Mol Biosyst, 2011. **7**(1): p. 74-81.

112. Layer, G., Verfürth, K., Mahlitz, E., Jahn, D, *Oxygen-independent coproporphyrinogen-III oxidase HemN from Escherichia coli.* J Biol Chem, 2002. **277**(37): p. 34136-34142.

113. Fang, Q., Peng, J., Dierks, T., *Post-translational formylglycine modification of bacterial sulfatases by the radical S-adenosylmethionine protein AtsB.* J Biol Chem, 2004. **279**(15): p. 14570-14578.

114. Pierre, S., Guillot, A., Benjdia, A., Sandström, C., Langella, P., Berteau, O., *Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes.* Nat Chem Biol, 2012. **8**(12): p. 957-959.

115. Paraskevopoulou, C., Fairhurst, S., Lowe, D., Brick, P., Onesti, S., *The elongator subunit Elp3 contains a Fe4S4 cluster and binds S-adenosylmethionine.* Mol Microbiol, 2006. **59**(3): p. 795-806.

116. Chatterjee, A., Li, Y., Zhang, Y., Grove, T., Lee, M., Krebs, C., Ealick, S., *Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily.* Nat Chem Biol, 2008. **4**(12): p. 758-765.

117. Suzuki, Y., Noma, A., Suzuki, T., Senda, M., Senda, T., Ishitani, R.,Nureki, O., *Crystal structure of the radical SAM enzyme catalyzing tricyclic modified base formation in tRNA.* J Mol Biol, 2007. **372**(5): p. 1204-14.

118. King, P., Posewitz, M., Ghirardi, M., Seibert, M., *Functional studies of [FeFe] hydrogenase maturation in an Escherichia coli biosynthetic system.* J Bacteriol, 2006. **188**(6): p. 2163-2172.

119. Petrovich, R., Ruzicka, F., Reed, G., Frey, P., *Characterization of iron-sulfur clusters in lysine 2,3-aminomutase by electron paramagnetic resonance spectroscopy.* Biochemistry, 1992. **31**(44): p. 10774-10781.

120. Lieder, K., Booker, S., Ruzicka, F., Beinert, H., Reed, G., Frey, P., *S-Adenosylmethionine-dependent reduction of lysine 2,3-aminomutase and observation of the catalytically functional iron-sulfur centers by electron paramagnetic resonance.* Biochemistry, 1998. **37**(8): p. 2578-2585.

121. Magnusson, O., Frey, P., *Spectroscopic evidence for the participation of an allylic analogue of the 5'-deoxyadenosyl radical in the reaction of lysine 2, 3-aminomutase.* J Am Chem Soc, 1999. **121**(41): p. 9764-9765.

122. Magnusson, O., Reed, G., Frey, P., *Characterization of an allylic analogue of the 5'-deoxyadenosyl radical: an intermediate in the reaction of lysine 2,3-aminomutase.* Biochemistry, 2001. **40**(26): p. 7773-7782.

123. Chen, D., Walsby, C., Hoffman, B., Frey, P, *Coordination and mechanism of reversible cleavage of S-adenosylmethionine by the [4Fe-4S] center in lysine 2,3-aminomutase.* J Am Chem Soc, 2003. **125**(39): p. 11788-11789.

124. Walsby, C., Ortillo, D., Broderick, W., Broderick, J., Hoffman, B., *An anchoring role for FeS clusters: chelation of the amino acid moiety of S-adenosylmethionine to the unique iron site of the [4Fe-4S] cluster of pyruvate formate-lyase activating enzyme.* J Am Chem Soc, 2002. **124**(38): p. 11270-11271.

125. Knappe, J., Blaschkowski, H., Gänzler, M., *Post-translational activation introduces a free radical into pyruvate formate-lyase.* P Natl Acad Sci, 1984. **81**(5): p. 1332-1335.

126. Mulliez, E., Gaillard, J., Reichard, P., *An iron-sulfur center and a free radical in the active anaerobic ribonucleotide reductase of Escherichia coli.* J Biol Chem, 1993. **268**(4): p. 2296-2299.

127. Kamachi, T., Kouno, T., Doitomi, K., Yoshizawa, K, *Generation of adenosyl radical from S-adenosylmethionine (SAM) in biotin synthase.* Journal of Inorganic Biochemistry, 2011. **105**(6): p. 850-857.

128. Ugulava, N., Gibney, B., Jarrett, J., *Biotin synthase contains two distinct iron-sulfur cluster binding sites: chemical and spectroelectrochemical analysis of iron-sulfur cluster interconversions.* Biochemistry, 2001. **40**(28): p. 8343-8351.

129. Saeva, F., Morgan, B, *Mechanism of one-electron electrochemical reductive cleavage reactions of sulfonium salts.* J. Am. Chem. Soc., 1984. **106**(15): p. 4121-4125.

130. Grimshaw, J., *Electrochemistry of the sulphonium group.* The Sulphonium Group, 1981. **Volume 1**: p. 141-155.

131. Mulliez, E., Padovani, D., Atta, M., Alcouffe, C., Fontecave, M., *Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center.* Biochemistry, 2001. **40**(12): p. 3730-3736.

132. Brazeau, B., Gort, S., Jessen, H., Andrew, A, Liao, H., *Enzymatic activation of lysine 2, 3-aminomutase from Porphyromonas gingivalis.* Appl Environ Microb, 2006. **72**(9): p. 6402-6404.

133. Imlay, J., *Pathways of oxidative damage,* Annu Rev Microbiol, 2003. **57**(1), p.395-418

134. Chandra, J., Samali, A., Orrenius, S., *Triggering and modulation of apoptosis by oxidative stress.* Free Radical Biol Med, 2000. **29**(3): p. 323-333.

135. Lees, N., Chen, D., Walsby, C., Behshad, E., Frey, P., Hoffman, B., *How an enzyme tames reactive intermediates: positioning of the active-site components of lysine 2,3-aminomutase during enzymatic turnover as determined by ENDOR spectroscopy.* J Am Chem Soc, 2006. **128**(31): p. 10145-10154.

136. Cantoni, G., *Biological methylation: selected aspects.* Annu Rev Biochem, 1975. **44**(1): p. 435-451.

137. Markham, G., Hafner, E., Tabor, C., Tabor, H., *S-Adenosylmethionine synthetase from Escherichia coli.* J Biol Chem, 1980. **255**(19): p. 9082-9092.

138. Pajares, M., Corrales, F., Durán, C., Mato, J., Alvarez, L., *How is rat liver S-adenosylmethionine synthetase regulated?* FEBS letters, 1992. **309**(1): p. 1-4.

139. Mayhew, S., *The redox potential of dithionite and SO− 2 from equilibrium reactions with flavodoxins, methyl viologen and hydrogen plus hydrogenase.* Eur J Biochem, 1978. **85**(2): p. 535-547.

140. Yoch, D.C., Valentine, R. C., *Ferredoxins and flavodoxins of bacteria.* Annu Rev Microbiol, 1972. **26**(1): p. 139-162.

141. Ollagnier-de Choudens, S., Sanakis, Y., Hewitson, K., Roach, P., Münck, E., Fontecave, M., *Reductive cleavage of S-adenosylmethionine by biotin synthase from Escherichia coli.* J Biol Chem, 2002. **277**(16): p. 13449-13454.

142. Frey, P., Booker, S., *Radical mechanisms of S-adenosylmethionine-dependent enzymes.* Adv Protein Chem, 2001. **58**: p. 1-45.

143. Pierrel, F., Hernandez, H., Johnson, M., Fontecave, M., Atta, M., *MiaB protein from Thermotoga maritima: Characterization of an extremely thermophilic tRNA-methylthiotransferase.* J Biol Chem, 2003. **278**(32): p. 29515-29524.

144. O'Brien, J., Raynaud, C., Croux, C., Girbal, L., Soucaille, P., Lanzilotta, W., *Insight into the mechanism of the B12-independent glycerol dehydratase from Clostridium butyricum: preliminary biochemical and structural characterization.* Biochemistry, 2004. **43**(16): p. 4635-4645.

145. Krieger, C., Roseboom, W., Albracht, S., Spormann, A., *A stable organic free radical in anaerobic benzylsuccinate synthase of Azoarcus sp. strain T.* J Biol Chem, 2001. **276**(16): p. 12924-12927.

146. Moss, M., Frey, P., *Activation of lysine 2,3-aminomutase by S-adenosylmethionine.* J Biol Chem, 1990. **265**(30): p. 18112-18115.

147. Horitani, M., Byer, A., Shisler, K., Chandra, T., Broderick, J., Hoffman, B., *Why nature uses radical SAM enzymes so widely: electron nuclear double resonance studies of lysine 2, 3-aminomutase show the 5'-dAdo•" free radical" Is neverfFree.* J Am Chem Soc, 2015. **137**(22): p. 7111-7121

148. Berkovitch, F., Behshad, E., Tang, K., Enns, E., Frey, P., Drennan, C., *A locking mechanism preventing radical damage in the absence of substrate, as revealed by the x-ray structure of lysine 5,6-aminomutase.* Proc Natl Acad Sci USA, 2004. **101**(45): p. 15870-15875.

149. Dowling, D., Vey, J., Croft, A., Drennan, C, *Structural diversity in the AdoMet radical enzyme superfamily.* Biochim Biophys Acta, 2012. **1824**(11): p. 1178-1195.

150. Nicolet, Y., Drennan, C., *AdoMet radical proteins--from structure to evolution--alignment of divergent protein sequences reveals strong secondary structure element conservation.* Nucleic Acids Res, 2004. **32**(13): p. 4015-4025.

151. Vey, J., Yang, J., Li, M., Broderick, W., Broderick, J., Drennan, C., *Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme.* Proc Natl Acad Sci USA, 2008. **105**(42): p. 16137-16141.

152. Broderick, J., Duffus, B., Duschene, K., Shepard, E., *Radical S-adenosylmethionine enzymes.* Chem Rev, 2014. **114**(8): p. 4229-4317.

153. Farrar, C., Jarrett, J., *Protein residues that control the reaction trajectory in S-adenosylmethionine radical enzymes: mutagenesis of asparagine 153 and aspartate 155 in Escherichia coli biotin synthase.* Biochemistry, 2009. **48**(11): p. 2448-2458.

154. Bianchi, V., Fontecave, M., Mulliez, E., Hoover, D., Matthews, R., Reichard, P., *Flavodoxin is required for the activation of the anaerobic ribonucleotide reductase.* Biochem Bioph Res Co, 1993. **197**(2): p. 792-797.

155. Picciocchi, A., Douce, R., Alban, C., *The plant biotin synthase reaction Identification and characterization of essential mitochondrial accessory protein components.* J Biol Chem, 2003. **278**(27): p. 24966-24975.

156. Ugulava, N., Gibney, B., Jarrett, J., *Iron-sulfur cluster interconversions in biotin synthase: dissociation and reassociation of iron during conversion of [2Fe-2S] to [4Fe-4S] clusters.* Biochemistry, 2000. **39**(17): p. 5206-5214.

157. Broderick, J.,Henshaw, T., Cheek, J., Wojtuszewski, K., Smith, S., Trojan, M., McGhan, R., Kopf, A., Kibbey, M., Broderick, W., *Pyruvate formate-lyase-activating enzyme: strictly anaerobic isolation yields active enzyme containing a [3Fe–4S]+ cluster.* Biochem Bioph Res Co, 2000. **269**: p. 451–456.

158. Cosper, M., Jameson, G., Hernández, H., Krebs, C., Huynh, B., Johsnon, M., *Characterization of the cofactor composition of Escherichia coli biotin synthase.* Biochemistry, 2004. **43**: p. 2007-2021.

159. Zhang, B., Crack, J., Subramanian, S., Green, J., Thomson, A., Le Brun, N., Johnson, M., *Reversible cycling between cysteine persulfide-ligated [2Fe-2S] and cysteine-ligated [4Fe-4S] clusters in the FNR regulatory protein.* P Natl Acad Sci, 2012. **109**(39): p. 15734-15739.

160. Bandyopadhyay, S., Chandramouli, K., Johnson, M., *Iron-Sulfur Cluster Biosynthesis.* BiochemSoc T, 2008. **36**(6): p. 1112-1119.

161. Raulfs, E., O'Carrol., I., Dos Santos, P., Unciuleac, M., Dean, D., *In vivo iron–sulfur cluster formation.* P Natl Acad Sci, 2008. **105**(25): p. 8591-8596.

162. Takahashi, Y., Tokumoto, U., *A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids.* J Biol Chem, 2002. **277**(32): p. 28380-28383.

163. Johnson, D., Dos Santos, P., Dean, D., *NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in Azotobacter vinelandii.* Biochem Soc T, 2005. **33**(1): p. 90-93.

164. Blanc, B., Gerez, C., de Choudens, S., *Assembly of Fe/S proteins in bacterial systems: Biochemistry of the bacterial ISC system.* Biochim Biophys Acta- Mol Cel Res, 2015. **1853**(6): p. 1436-1447.

165. Roche, B., Aussel, L., Ezraty, B., Mandin, P., Py, B., Barras, F., *Iron/sulfur proteins biogenesis in prokaryotes: formation, regulation and diversity.* Biochim Biophys Acta-Bioenergetics, 2013. **1827**(8): p. 923-937.

166. Lill, R., Srinivasan, V., Mühlenhoff, U., *The role of mitochondria in cytosolic-nuclear iron–sulfur protein biogenesis and in cellular iron regulation.* Curr Opin Microbiol, 2014(22): p. 111-119.

167. Iwasaki, T., *Iron-sulfur world in aerobic and hyperthermoacidophilic archaea Sulfolobus.* Archaea, 2010. **2010**: p. 1-14.

168. Balk, J., Schaedler, T., *Iron cofactor assembly in plants.* Annu Rev Plant Biol, 2014. **65**: p. 125-153.

169. Boyd, E., Thomas, K., Dai, Y., Boyd, J., Outten, F., *Interplay between oxygen and Fe–S cluster biogenesis: insights from the Suf pathway.* Biochemistry, 2014. **53**(37): p. 5834-5847.

170. Sharma, A., Pallesen, L., Spang, R., Walden, W., *Cytosolic iron-sulfur cluster assembly (CIA) system: factors, mechanism, and relevance to cellular iron regulation.* J Biol Chem, 2010. **285**(35): p. 26745-26751.

171. Lanz, N., Booker, S., *Identification and function of auxiliary iron–sulfur clusters in radical SAM enzymes.* Biochim Biophys Acta- Prot Proteomics, 2012. **1824**(11): p. 1196-1212.

172. Lanz, N., Booker, S., *Auxiliary iron–sulfur cofactors in radical SAM enzymes.* Biochim Biophys Acta- Mol Cell Res, 2015. **1853**(6): p. 1316-1334.

173. Grell, T., Goldman, P., Drennan, C., *SPASM and twitch domains in AdoMet radical enzyme structures.* J Biol Chem, 2014: p. jbc-R114.

174. Haft, D., *Bioinformatic evidence for a widely distributed, ribosomally produced electron carrier precursor, its maturation proteins, and its nicotinoprotein redox partners.* BMC genomics, 2011. **12**(1): p. 21-35.

175. Haft, D., Basu, M., *Biological systems discovery in silico: radical S-sdenosylmethionine protein families and their target peptides for posttranslational modification.* J. Bacteriol., 2011. **193**(11): p. 2745-2755.

176. Benjdia, A., Subramanian, S., Leprince, J., Vaudry, H., Johnson, M., Berteau, O., *Anaerobic sulfatase-maturating enzyme--a mechanistic link with glycyl radical-activating enzymes?* FEBS J, 2010. **277**(8): p. 1906-1920.

177. Grove, T., Lee, K., St. Clair, J., Krebs, C., Booker, S., *In vitro characterization of AtsB, a radical SAM formylglycine-generating enzyme that contains three [4Fe-4S] clusters.* Biochemistry, 2008. **47**(28): p. 7523-7538.

178. Goldman, P., Grove, T., Sites, L., McLaughlin, M., Booker, S., Drennan, C., *X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification.* Proc Natl Acad Sci USA, 2013. **110**(21): p. 8519-8524.

179. Hanzelmann, P., Schindelin, H., *Binding of 5'-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoaA provides insights into its mechanism.* Proc Natl Acad Sci USA, 2006. **103**(18): p. 6829-6834.

180. Berkovitch, F., Nicolet, Y., Wan, J., Jarrett, J., Drennan, C, *Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme.* Science, 2004. **303**(5654): p. 76-79.

181. Cicchillo, R., Baker, M., Schnitzer, E, Newman, E., Krebs, C., Booker, S., *Escherichia coli L-serine deaminase requires a [4Fe-4S] cluster in catalysis.* J Biol Chem, 2004. **279**(31): p. 32418-32425.

182. Harmer, J., Hiscox, M., Dinis, P., Fox, S., Iliopoulos, A., Hussey, J., Sandy, J., Van Beek, F., Essex, J., Roach, P., *Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions.* Biochem J, 2014. **464**(1): p. 123-133.

183. Hernández, H., Pierrel, F., Elleingand, E., García-Serres, R., Huynh, B., Johnson, M., Atta, M., *MiaB, a bifunctional radical-S-adenosylmethionine enzyme involved in the thiolation and methylation of tRNA, contains two essential [4Fe-4S] clusters.* Biochemistry, 2007. **46**(17): p. 5140-5147.

184. Kuchenreuther, J., Britt, D., Swartz, J., *New insights into [FeFe] hydrogenase activation and maturase function.* Plos One, 2012. **7**(9): p. e45850.

185. Leuthner, B., Leutwein, C., Schulz, H., Hörth, P., Haehnel, W., Schiltz, E., Heider, J., *Biochemical and genetic characterization of benzylsuccinate synthase from Thauera aromatica: a new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism.* Mol Microbiol, 1998. **28**(3): p. 615-628.

186. Flühe, L., Burghaus, O., Wieckowski, B., Giessen, T., Linne, U., Marahiel, M., *Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor.* J Am Chem Soc, 2013. **135**(3): p. 959-962.

187. Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., *Climate change 2013: the physical science basis.* Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, 2013.

188. Ibrahim, H., Ilinca, A., Perron, J., *Energy storage systems—characteristics and comparisons.* Renew Sust Energ Rev, 2008. **12**(5): p. 1221-1250.

189. Lindley, D., *Smart grids: the energy storage problem.* Nature News, 2010. **463**(7277): p. 18-20.

190. Webster, L., Ibrahim, S., Wright, J., Pickett, C., *Solar fuels: visible-light-driven generation of dihydrogen at p-type silicon electrocatalysed by molybdenum hydrides.* Chem-Eur J, 2012. **18**(37): p. 11798-11803.

191. Cook, T., Dogutan, D., Reece, S., Surendranath, Y., Teets, T., Nocera, D., *Solar energy supply and storage for the legacy and nonlegacy worlds.* Chem Rev, 2010. **110**(11): p. 6474-6502.

192. Lu, G., da Costa, J., Duke, M., Giessler, S., Socolow, R., Williams, R., Kreutz, T., *Inorganic membranes for hydrogen production and purification: a critical review and perspective.* J Colloid Interf Sci, 2007. **314**(2): p. 589-603.

193. Melis, A., Happe, T., *Hydrogen production. Green algae as a source of energy.* Plant Physiol, 2001. **127**(3): p. 740-748.

194. Happe, T., Kaminski, A., *Differential regulation of the Fe-hydrogenase during anaerobic adaptation in the green alga Chlamydomonas reinhardtii.* Eur J Biochem, 2002. **269**(3): p. 1022-1032.

195. Stripp, S., Goldet, G., Brandmayr, C., Sanganas, O., Vincent, K., Haumann, M., Happe, T., *How oxygen attacks [FeFe] hydrogenases from photosynthetic organisms.* P Natl Acad Sci, 2009. **106**(41): p. 17331-17336.

196. Cammack, R., Frey, M., Robson, R. (Eds.), *Hydrogen as a fuel: learning from nature.* CRC Press, 2002.

197. Karyakin, A., Morozov, S., Karyakina, E., Zorin, N., Perelygin, V., Cosnier, S., *Hydrogenase electrodes for fuel cells.* Biochem Soc T, 2005. **33**(1): p. 73-75.

198. Mertens, R., Liese, A., *Biotechnological applications of hydrogenases.* Curr Opin Biotech, 2004. **15**(4): p. 343-348.

199. Kasting, J., Brown, L., *The early atmosphere as a source of biogenic compounds.* The Molecular Origins of Life, Cambridge University Press, 1998.

200. Tian, F., Toon, O., Pavlov, A. , De Sterck, H., *A hydrogen-rich early Earth atmosphere.* Science, 2005. **308**(5724): p. 1014-1017.

201. Van Kranendonk, M., *Earth's early atmosphere and surface environments: A review.* Geol Soc Am Sp Paper, 2014. **504**: p. 105-130.

202. Stephenson, M., Stickland, L. , *Hydrogenase: a bacterial enzyme activating molecular hydrogen: The properties of the enzyme.* Biochem J, 1931. **25**(1): p. 205-214.

203. Vignais, P., Colbeau, A., *Molecular biology of microbial hydrogenases.* Curr Issues Mol Biol, 2004. **6**(2): p. 159-188.

204. Madigan, M., Martinko, J., Stahl, D., and Clark, D., *Brock Biology of Microorganisms.* Benjamin Cummings, 2010.

205. Vignais, P., Billoud, B., Meyer, J., *Classification and phylogeny of hydrogenases.* FEMS Microbiol Rev, 2001. **25**(4): p. 455-501.

206. Vignais, P, Billoud, B., *Occurrence, classification, and biological function of hydrogenases: an overview.* Chem Rev, 2007. **107**(10): p. 4206-4272.

207. Tamagnini, P., Axelsson, R., Lindberg, P., Oxelfelt, F., Wünschiers, R., Lindblad, P., *Hydrogenases and hydrogen metabolism of cyanobacteria.* Microbiol Mol Biol R, 2002. **66**(1): p. 1-20.

208. Frey, M., *Hydrogenases: hydrogen-activating enzymes.* Chem Bio Chem, 2002. **3**(2-3): p. 153-160.

209. Adams, M., Mortenson, L., Chen, J., *Hydrogenase.* Biochim Biophys Acta, 1981. **594**(2-3): p. 105-175.

210. Sapra, R., Bagramyan, K., Adams, M., *A simple energy-conserving system: proton reduction coupled to proton translocation.* P Natl Acad Sci, 2003. **100**(13): p. 7545-7550.

211. Graf, M., Bokranz, M., Böcher, R., Friedl, P., Kröger, A., *Electron transport driven phosphorylation catalyzed by proteoliposomes containing hydrogenase, fumarate reductase and ATP synthase.* FEBS letters, 1985. **184**(1): p. 100-103.

212. Trchounian, K., Pinske, C., Sawers, R., Trchounian, A., *Dependence on the F0F1-ATP synthase for the activities of the hydrogen-oxidizing hydrogenases 1 and 2 during glucose and glycerol fermentation at high and low pH in Escherichia coli.* J Bioenerg Biomembr, 2011. **43**(6): p. 645-650.

213. Stojanowic, A., Mander, G., Duin, E., Hedderich, R., *Physiological role of the F420-non-reducing hydrogenase (Mvh) from Methanothermobacter marburgensis.* Arch Microbiol, 2003. **180**(3): p. 194-203.

214. Burgess, B., Lowe, D., *Mechanism of molybdenum nitrogenase.* Chem Rev, 1996. **96**(7): p. 2983-3012.

215. Xie, H., Cao, Z., *Enzymatic reduction of nitrate to nitrite: insight from density functional calculations.* Organometallics, 2009. **29**(2): p. 436-441.

216. Schwartz, E., Fritsch, J., Friedrich, B., *4 H2-Metabolizing Prokaryotes.* Springer Berlin Heidelberg, 2013.

217. Bélaich, J., Bruschi, M., Garcia, J., *Microbiology and biochemistry of strict anaerobes involved in interspecies hydrogen transfer.* Springer Science Business Media, 2012**.**

218. Meyer, J., *[FeFe] hydrogenases and their evolution: a genomic perspective.* Cell Mol Life Sci, 2007. **64**(9): p. 1063-1084.

219. Spek, T.A., A., Happe, R., Yun, S., Bagley, K., Stufkens, D., Albracht, S., *Similarities in the Architecture of the Active Sites of Ni-Hydrogenases and Fe-Hydrogenases Detected by Means of Infrared Spectroscopy.* Eur J Biochem, 1996. **237**(3): p. 629-634.

220. Evans, D., Picker, C., *Chemistry and the hydrogenases.* Chem Soc Rev, 2003. **32**(5): p. 268-275.

221. Armstrong, F., *Hydrogenases: active site puzzles and progress.* Curr Opin Chem Biol, 2004. **8**(2): p. 133-140.

222. Lubitz, W., Tumas, W., *Hydrogen: an overview.* Chem Rev, 2007. **107**(10): p. 3900-3903.

223. Tard, C., Picket, C., *Structural and Functional Analogues of the Active Sites of the [Fe]-,[NiFe]-, and [FeFe]-Hydrogenases.* Chem Rev, 2009. **109**(6): p. 2245-2274.

224. Shima, S., Thauer, R., *A third type of hydrogenase catalyzing H2 activation.* Chem Rec, 2007. **7**(1): p. 37-46.

225. Pilak, O., Mamat, B., Vogt, S., Hagemeier, C., Thauer, R., Shima, S.,Ermler, U., *The crystal structure of the apoenzyme of the iron–sulphur cluster-free hydrogenase.* J Mol Biol, 2006. **358**(3): p. 798-809.

226. Thauer, R., Klein, A., Hartmann, G., *Reactions with molecular hydrogen in microorganisms: evidence for a purely organic hydrogenation catalyst.* Chem Rev, 1996. **96**(7): p. 3031-3042.

227. Zirngibl, C., Hedderich, R., Thauer, R., *N5, N10-Methylenetetrahydromethanopterin dehydrogenase from Methanobacterium thermoautotrophicum has hydrogenase activity.* FEBS Letters, 1990. **261**(1): p. 112-116.

228. Hartmann, G., Klein, A., Linder, M., Thauer, R., *Purification, properties and primary structure of H2-formingN 5, N10-methylenetetrahydromethanopterin dehydrogenase fromMethanococcus thermolithotrophicus.* Arch Microbiol, 1996. **165**(3): p. 187-193.

229. Buurman, G., Shima, S., Thauer, R., *The metal-free hydrogenase from methanogenic archaea: evidence for a bound cofactor.* FEBS letters, 2000. **485**(2): p. 200-204.

230. Lyon, E., Shima, S., Boecher, R., Thauer, R., Grevels, F., Bill, E., Albracht, S., *Carbon monoxide as an intrinsic ligand to iron in the active site of the iron-sulfur-cluster-free hydrogenase H2-forming methylenetetrahydromethanopterin dehydrogenase as revealed by infrared spectroscopy.* J Am Chem Soc, 2004. **126**(43): p. 14239-14248.

231. Korbas, M., Vogt, S., Meyer-Klaucke, W., Bill, E., Lyon, E., Thauer, R., Shima, S., *The iron-sulfur cluster-free hydrogenase (Hmd) is a metalloenzyme with a novel iron binding motif.* J Biol Chem, 2006. **281**(41): p. 30804-30813.

232. Hiromoto, T., Warkentin, E., Moll, J., Ermler, U., Shima, S., *The Crystal Structure of an [Fe]-Hydrogenase–Substrate Complex Reveals the Framework for H2 Activation.* Angew Chem Int Edit, 2009. **48**(35): p. 6457-6460.

233. Przybyla, A., Robbins, J., Menon, N., Peck Jr, H., *Structure-function relationships among the nickel-containing hydrogenases.* FEMS microbiology reviews, 1992. **88**(2): p. 109-136.

234. Volbeda, A., Garcin, E., Piras, C., de Lacey, A., Fernandez, V., Hatchikian, E., Fontecilla-Camps, J., *Structure of the [NiFe] hydrogenase active site: evidence for biologically uncommon Fe ligands.* J Am Chem Soc, 1996. **118**(51): p. 12989-12996.

235. Vignais, P., Touissant, B., Colbeau, A., *Regulation of hydrogenase gene expression.* Anoxygenic Photosynthetic Bacteria, Springer Netherlands, 2004.

236. Friedrich, B., Schwartz, E., *Molecular biology of hydrogen utilization in aerobic chemolithotrophs.* Annu Rev Microbiol, 1993. **47**(1): p. 351-383.

237. Albracht, S., *Nickel hydrogenases: in search of the active site.* Biochim Biophys Acta-Bioenergetics, 1994. **1188**(3): p. 167-204.

238. Vincent, K., Parkin, A., Armstrong, F., *Investigating and exploiting the electrocatalytic properties of hydrogenases.* Chem Rev, 2007. **107**(10): p. 4366-4413.

239. Burgdorf, T., Lenz, O., Buhrke, T., Van Der Linden, E., Jones, A., Albracht, S., Friedrich, B., *[NiFe]-hydrogenases of Ralstonia eutropha H16: modular enzymes for oxygen-tolerant biological hydrogen oxidation.* J Mol Microbiol Biotech, 2005. **10**(2-4): p. 181-196.

240. Goris, T., Wait, A., Saggu, M., Fritsch, J., Heidary, N., Stein, M., Lenz, O., *A unique iron-sulfur cluster is crucial for oxygen tolerance of a [NiFe]-hydrogenase.* Nat Chem Biol, 2011. **7**(5): p. 310-318.

241. Shafaat, H., Rüdiger, O., Ogata, H., Lubitz, W., *[NiFe] hydrogenases: a common active site for hydrogen metabolism under diverse conditions.* Biochim Biophys Acta- Bioenergetics, 2013. **1827**(8): p. 986-1002.

242. Garcin, E., Vernede, X., Hatchikian, E., Volbeda, A., Frey, M., Fontecilla-Camps, J., *The crystal structure of a reduced [NiFeSe] hydrogenase provides an image of the activated catalytic center.* Structure, 1999. **7**(5): p. 557-566.

243. Adams, M., *The structure and mechanism of iron-hydrogenases.* Biochim Biophys Acta-Bioenergetics, 1990. **1020**(2): p. 115-145.

244. Popescu, C., Münck, E., *Electronic structure of the H cluster in [Fe]-hydrogenases.* J Am Chem Soc, 1999. **121**(34): p. 7877-7884.

245. Peters, J., *Structure and mechanism of iron-only hydrogenases.* Curr Opin Struc Biol, 1999. **9**(6): p. 670-676.

246. Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C., Fontecilla-Camps, J., *Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center.* Structure, 1999. **7**(1): p. 13-23.

247. Fontecilla-Camps, J., Volbeda, A., Cavazza, C., Nicolet, Y., *Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases.* Chem Rev, 2007. **107**(10): p. 4273-303.

248. Vincent, K., Parkin, A., Lenz, O., Albracht, S., Fontecilla-Camps, J., Cammack, R., Armstrong, F., *Electrochemical definitions of O2 sensitivity and oxidative inactivation in hydrogenases.* J Am Chem Soc, 2005. **127**(51): p. 18179-18189.

249. Peters, J., Lanzilotta, W., Lemon, B., Seefeldt, L., *X-ray crystal structure of the Fe-only hydrogenase (CpI) from Clostridium pasteurianum to 1.8 angstrom resolution.* Science, 1998. **282**(5395): p. 1853-1858.

250. Ghirardi, M., Posewitz, M., Maness, P., Dubini, A., Yu, J., Seibert, M., *Hydrogenases and hydrogen photoproduction in oxygenic photosynthetic organisms.* Annu. Rev. Plant Biol., 2007. **58**: p. 71-91.

251. Schut, G., Adams, M., *The iron-hydrogenase of Thermotoga maritima utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production.* J Bacteriol, 2009. **191**(13): p. 4451-4457.

252. Sterner, R., *Ferredoxin from Thermotoga maritima.* Method Enzymol, 2001. **334**: p. 23-30.

253. Wood, E., Sutton, R., Rockett, B., Swindells, P., *Chemistry for the life sciences.* Taylor Francis London, 2001.

254. Ducat, D., Sachdeva, G., Silver, P., *Rewiring hydrogenase-dependent redox circuits in cyanobacteria.* P Natl Acad Sci, 2011. **108**(10): p. 3941-3946.

255. Gutekunst, K., Chen, X., Schreiber, K., Kaspar, U., Makam, S., Appel, J., *The bidirectional NiFe-hydrogenase in Synechocystis sp. PCC 6803 is reduced by flavodoxin and ferredoxin and is essential under mixotrophic, nitrate-limiting conditions.* J Biol Chem, 2014. **289**(4): p. 1930-1937.

256. Yang, X., Hall, M., *Trigger mechanism for the catalytic hydrogen activation by monoiron (iron−sulfur cluster-free) hydrogenase.* J Am Chem Soc, 2008. **130**(43): p. 14036-14037.

257. Fujishiro, T., Kahnt, J., Ermler, U., Shima, S., *Protein-pyridinol thioester precursor for biosynthesis of the organometallic acyl-iron ligand in [lsqb] Fe [rsqb]-hydrogenase cofactor.* Nature Comms, 2015. **6**.

258. Volbeda, A., Charon, M., Piras, C., Hatchikian, E., Frey, M., Fontecilla-Camps, J., *Crystal structure of the nickel–iron hydrogenase from Desulfovibrio gigas.*Nature, 1995. **373**: p. 580-587

259. Silakov, A., Wenk, B., Reijerse, E., Lubitz, W., *14N HYSCORE investigation of the H-cluster of [FeFe] hydrogenase: evidence for a nitrogen in the dithiol bridge.* Phys Chem Ch Ph, 2009. **11**(31): p. 6592-6599.

260. Shima, S., Pilak, O., Vogt, S., Schick, M., Stagni, M., Meyer-Klaucke, W., Ermler, U., *The crystal structure of [Fe]-hydrogenase reveals the geometry of the active site.* Science, 2008. **321**(5888): p. 572-575.

261. Pierik, A., Hulstein, M., Hagen, W., Albracht, S., *A low-spin iron with CN and CO as intrinsic ligands forms the core of the active site in [Fe]-hydrogenases.* Eur J Biochem, 1998. **258**(2): p. 572-578.

262. Pierik, A., Roseboom, W., Happe, R., Bagley, K., Albracht, S., *Carbon monoxide and cyanide as intrinsic ligands to iron in the active site of [NiFe]-Hydrogenases NiFe (CN) 2CO, Biology's way to activate H2.* J Biol Chem, 1999. **274**(6): p. 3331-3337.

263. Higuchi, Y., Yagi, T., Yasuoka, N., *Unusual ligand structure in Ni–Fe active center and an additional Mg site in hydrogenase revealed by high resolution X-ray structure analysis.* Structure, 1997. **5**(12): p. 1671-1680.

264. Montet, Y., Amara, P., Volbeda, A., Vernede, X., Hatchikian, E., Field, M., Fontecilla-Camps, J., *Gas access to the active site of Ni-Fe hydrogenases probed by X-ray crystallography and molecular dynamics.* Nat Struct Biol, 1997. **4**(7): p. 523-526.

265. Bennett, B., Lemon, B., Peters, J., *Reversible carbon monoxide binding and inhibition at the active site of the Fe-only hydrogenase.* Biochemistry, 2000. **39**(25): p. 7455-7460.

266. Singleton, M., Bhuvanesh, N., Reibenspies, J., Darensbourg, M., *Synthetic support of de novo design: sterically bulky [FeFe]-hydrogenase models.* Angewandte Chemie, 2008. **120**(49): p. 9634-9637.

267. Nicolet, Y., de Lacey, A., Vernede, X., Fernandez, V., Hatchikian, E. . Fontecilla-Camps, J., *Crystallographic and FTIR spectroscopic evidence of changes in Fe coordination upon reduction of the active site of the Fe-only hydrogenase from Desulfovibrio desulfuricans.* J Am Chem Soc, 2001. **123**(8): p. 1596-601.

268. Nicolet, Y., Cavazza, C., Fontecilla-Camps, J., *Fe-only hydrogenases: structure, function and evolution.* J Inorg Biochem, 2002. **91**(1): p. 1-8.

269. Pandey, A., Harris, T., Giles, L., Peters, J., Szilagyi, R., *Dithiomethylether as a ligand in the hydrogenase H-cluster.* J Am Chem Soc, 2008. **130**(13): p. 4533-4540.

270. Olsen, M., Barton, B., Rauchfuss, T., *Hydrogen activation by biomimetic diiron dithiolates.* Inorg Chem, 2009. **48**(16): p. 7507-7509.

271. Ryde, U., Greco, C., De Gioia, L., *Quantum refinement of [FeFe] hydrogenase indicates a dithiomethylamine ligand.* J Am Chem Soc, 2010. **132**(13): p. 4512-4513.

272. Erdem, Ö., Schwartz, L., Stein, M., Silakov, A., Kaur-Ghumaan, S., Huang, P., Lubitz, W., *A model of the [FeFe] hydrogenase active site with a biologically relevant azadithiolate bridge: a spectroscopic and theoretical investigation.* Angew Chem Int Edit, 2011. **50**(6): p. 1439-1443.

273. Carroll, M., Barton, B., Rauchfuss, T.,Carroll, P., *Synthetic models for the active site of the [FeFe]-hydrogenase: catalytic proton reduction and the structure of the doubly protonated intermediate.* J Am Chem Soc, 2012. **134**(45): p. 18843-18852.

274. Berggren, G., Adamska, A., Lambertz, C., Simmons, T., Esselborn, J., Atta, M., Happe, T., *Biomimetic assembly and activation of [FeFe]-hydrogenases.* Nature, 2013. **499**(7456): p. 66-69.

275. Esselborn, J., Lambertz, C., Adamska-Venkatesh, A., Simmons, T., Berggren, G., Noth, J., Fontecave, M, *Spontaneous activation of [FeFe]-hydrogenases by an inorganic [2Fe] active site mimic.* Nat Chem Biol, 2013. **9**(10): p. 607-609.

276. Knörzer, P., Silakov, A., Foster, C., Armstrong, F., Lubitz, W., Happe, T., *Importance of the protein framework for catalytic activity of [FeFe]-hydrogenases.* J Biol Chem, 2012. **287**(2): p. 1489-1499.

277. Winkler, M., Esselborn, J., Happe, T., *Molecular basis of [FeFe]-hydrogenase function: An insight into the complex interplay between protein and catalytic cofactor.* Biochim Biophys Acta- Bioenergetics, 2013. **1827**(8): p. 974-985.

278. Adamska, A., Silakov, A., Lambertz, C., Rüdiger, O., Happe, T., Reijerse, E., Lubitz, W., *Identification and Characterization of the "Super-Reduced" State of the H-Cluster in [FeFe] Hydrogenase: A New Building Block for the Catalytic Cycle?* Angew Chem Int Edit, 2012. **51**(46): p. 11458-11462.

279. Greco, C.. Bruschi, M., De Gioia, L., Ryde, U., *A QM/MM investigation of the activation and catalytic mechanism of Fe-only hydrogenases.* Inorg Chem, 2007. **46**(15): p. 5911-5921.

280. Darensbourg, M., Lyon, E., Zhao, X., Georgakaki, I., *The organometallic active site of [Fe] hydrogenase: models and entatic states.* P Natl Acad Sci, 2003. **100**(7): p. 3683-3688.

281. Girbal, L., von Abendroth, G., Winkler, M., Benton, P., Meynial-Salles, I., Croux, C., Soucaille, P., *Homologous and heterologous overexpression in Clostridium acetobutylicum and characterization of purified clostridial and algal Fe-only hydrogenases with high specific activities.* Appl Environ Microb, 2005. **71**(5): p. 2777-2781.

282. McGlynn, S., Ruebush, S., Naumov, A., Nagy, L., Dubini, A., King, P., Peters, J., *In vitro activation of [FeFe] hydrogenase: new insights into hydrogenase maturation.* J Biol Inorg Chem, 2007. **12**(4): p. 443-447.

283. Mulder, D., Ortillo, D., Gardenghi, D., Naumov, A., Ruebush, S., Szilagyi, R., Peters, J., *Activation of HydA(ΔEFG) requires a preformed [4Fe-4S] cluster.* Biochemistry, 2009. **48**(26): p. 6240-8.

284. Czech, I., Stripp, S., Sanganas, O., Leidel, N., Happe, T., Haumann, M., *The [FeFe]-hydrogenase maturation protein HydF contains a H-cluster like [4Fe4S]–2Fe site.* FEBS letters, 2011. **585**(1): p. 225-230.

285. Joshi, N., Shepard, E., Byer, A., Swanson, K., Broderick, J., Peters, J., *Iron–sulfur cluster coordination in the [FeFe]-hydrogenase H cluster biosynthetic factor HydF.* FEBS letters, 2012. **586**(22): p. 3939-3943.

286. Brazzolotto, X., Rubach, J., Gaillard, J., Gambarelli, S., Atta, M., Fontecave, M., *The [Fe-Fe]-hydrogenase maturation protein HydF from Thermotoga maritima is a GTPase with an iron-sulfur cluster.* J Biol Chem, 2006. **281**(2): p. 769-774.

287. Shepard, E.., McGlynn, S., Bueling, A., Grady-Smith, C., George, S., Winslow, M., Broderick, J., *Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold.* Proc Natl Acad Sci USA, 2010. **107**(23): p. 10448-10453.

288. Berto, P., Di Valentin, M., Cendron, L., Vallese, F., Albertini, M., Salvadori, E., Costantini, P., *The [4Fe-4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies.* Biochim Biophys Acta-Bioenergetics, 2012. **1817**(12): p. 2149-2157.

289. Krissinel, E., Henrick, K., *Inference of macromolecular assemblies from crystalline state.* J Mol Biol, 2007. **372**(3): p. 774-797.

290. Czech, I., Silakov, A., Lubitz, W., Happe, T., *The [FeFe]-hydrogenase maturase HydF from Clostridium acetobutylicum contains a CO and CN- ligated iron cofactor.* FEBS letters, 2010. **584**(3): p. 638-642.

291. McGlynn, S., Shepard, E., Winslow, M., Naumov, A., Duschene, K., Posewitz, M., Peters, J., *HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis.* FEBS Lett, 2008. **582**(15): p. 2183-2187.

292. Vallese, F., Berto, P., Ruzzene, M., Cendron, L., Sarno, S., Rosa, E., Giacometti, G., Costantini, P., *Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process.* The J Biol Chem, 2012. **287**(43): p. 36544-36555.

293. Shepard, E., Duffus, B., George, S., McGlynn, S., Challand, M., Swanson, K., Broderick, J., *[FeFe]-hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide.* J Am Chem Soc, 2010. **132**(27): p. 9247-9249.

294. Broderick, J., Byer, A., Duschene, K., Duffus, B., Betz, J., Shepard, E., Peters, J., *H-Cluster assembly during maturation of the [FeFe]-hydrogenase.* J Biol Inorg Chem, 2014. **19**(6): p. 747-757.

295. Nicolet, Y. Fontecilla-Camps, J., *Structure-function relationships in [FeFe]-hydrogenase active site maturation.* J Biol Chem, 2012. **287**(17): p. 13532-13540.

296. Kuchenreuther, J., Myers, W., Suess, D., Stich, T., Pelmenschikov, V., Shiigi, S., George, S., *The HydG enzyme generates an Fe(CO)(2)(CN) synthon in assembly of the FeFe hydrogenase H-cluster.* Science, 2014. **343**(6169): p. 424-427.

297. Nicolet, Y., Rubach, J., Posewitz, M., Amara, P., Mathevon, C., Atta, M., Fontecilla-Camps, J., *X-ray structure of the [FeFe]-hydrogenase maturase HydE from Thermotoga maritima.* J Biol Chem, 2008. **283**(27): p. 18861-18872.

298. Shepard, E., Boyd, E., Broderick, J., Peters, J., *Biosynthesis of complex iron-sulfur enzymes.* Curr Opin Chem Biol, 2011. **15**(2): p. 319-327.

299. Peters, J., Szilagyi, R., Naumov, A., Douglas, T., *A radical solution for the biosynthesis of the H-cluster of hydrogenase.* FEBS letters, 2006. **580**(2): p. 363-367.

300. Jarrett, J., *Biotin synthase: enzyme or reactant?* Chem Biol, 2005. **12**(4): p. 409-410.

301. Nicolet, Y., Rohac, R., Martin, L., Fontecilla-Camps, J., *X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe4S4 clusters.* Proc Natl Acad Sci USA, 2013. **110**(18): p. 7188-7192.

302. Pilet, E., Nicolet, Y., Mathevon, C., Douki, T., Fontecilla-Camps, J., Fontecave, M., *The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly.* FEBS Lett, 2009. **583**(3): p. 506-~511.

303. Myers, W., Stich, T., Suess, D., Kuchenreuther, J., Swartz, J., Britt, R., *The cyanide ligands of [FeFe] hydrogenase: pulse EPR studies of 13C and 15N-labeled H-cluster.* J Am Chem Soc, 2014. **136**(35): p. 12237-12240.

304. Betz, J., Boswell, N., Fugate, C., Holliday, G., Akiva, E., Scott, A., Broderick, J., *[FeFe]-hydrogenase maturation: insights into the role HydE plays in dithiomethylamine biosynthesis.* Biochemistry, 2015. **54**(9): p. 1807-1818.

305. Lubitz, W., Ogata, H., Rüdiger, O., Reijerse, E., *Hydrogenases.* Chem Rev, 2014. **114**(8): p. 4081-4148.

306. Jarrett, J., Farrar, C., *Radical S-adenosylmethionine (SAM) superfamily.* eLS, 2007.

307. Zhang, Q., Li, Y., Chen, D., Yu, Y., Duan, L., Shen, B., Liu, W., *Radical-mediated enzymatic carbon chain fragmentation-recombination.* Nat Chem Biol, 2011. **7**(3): p. 154-160.

308. Decamps, L., Philmus, B., Benjdia, A., White, R., Begley, T., Berteau, O., *Biosynthesis of F0, precursor of the F420 cofactor, requires a unique two radical-SAM domain enzyme and tyrosine as substrate.* J Am Chem Soc, 2012. **134**(44): p. 18173-18176.

309. Kuchenreuther, J., George, S., Grady-Smith, C., Cramer, S., Swartz, J., *Cell-free H-cluster synthesis and [FeFe] hydrogenase activation: all Five CO and CN- Ligands derive from tyrosine.* Plos One, 2011. **6**(5).

310. Tron, C., Cherrier, M., Amara, P., Martin, L., Fauth, F., Fraga, E., Fontecilla-Camps, J., *Further characterization of the [FeFe]-hydrogenase maturase HydG.* Eur J Inorg Chem, 2011. **7**: p. 1121-1127.

311. Driesener, R., Duffus, B., Shepard, E., Bruzas, I., Duschene, K., Coleman, N., Broderick, J., *Biochemical and kinetic characterization of radical S-Adenosyl-l-methionine enzyme HydG.* Biochemistry, 2013. **52**(48): p. 8696-8707

312. Nicolet, Y., Martin, L., Tron, C., Fontecilla-Camps, J., *A glycyl free radical as the precursor in the synthesis of carbon monoxide and cyanide by the [FeFe]-hydrogenase maturase HydG.* FEBS Lett, 2010. **584**(19): p. 4197-4202.

313. Kriek, M., Martins, F., Challand, M., Croft, A., Roach, P., *Thiamine biosynthesis in Escherichia coli: identification of the intermediate and by-product derived from tyrosine.* Angew Chem Int Ed Engl, 2007. **46**(48): p. 9223-9226.

314. Kuchenreuther, J., Myers, W., Stich, T., George, S., NejatyJahromy, Y., Swartz, J., Britt, R., *A radical Intermediate in tyrosine scission to the CO and CN- ligands of FeFe hydrogenase.* Science, 2013. **342**(6157): p. 472-475.

315. Jurgenson, C., Begley, T., Ealick, S., *The structural and biochemical foundations of thiamin biosynthesis.* Annu Rev Biochem, 2009. **78**: p. 569-603.

316. Kriek, M., Martins, F., Leonardi, R., Fairhurst, S., Lowe, D., Roach, P., *Thiazole synthase from Escherichia coli: an investigation of the substrates and purified proteins required for activity in vitro.* J Biol Chem, 2007. **282**(24): p. 17413-23.

317. Nicolet, Y., Pagnier, A., Zeppieri, L., Martin, L., Amara, P., Fontecilla-Camps, J., *Crystal structure of HydG from Carboxydothermus hydrogenoformans: a trifunctional [FeFe]-hydrogenase maturase.* Chem Bio Chem, 2015. **16**(3): p. 397-402.

318. Suess, D., Britt, R., *EPR spectroscopic studies of [FeFe]-hydrogenase ,aturation.* Top Cat, 2015. **58**(12-13): p. 699-707.

319. Duffus, B., Ghose, S., Peters, J., Broderick, J., *Reversible H atom abstraction catalyzed by the radical S-adenosylmethionine enzyme HydG.* J Am Chem Soc, 2014. **136**(38): p. 13086-13089.

320. Nicolet, Y., Fontecilla-Camps, J., Fontecave, M., *Maturation of [FeFe]-hydrogenases: Structures and mechanisms.* Int J Hydrogen Energ, 2010. **35**(19): p. 10750-10760.

321. Nicolet, Y., Zeppieri, L., Amara, P., Fontecilla-Camps, J. C., *Crystal structure of tryptophan lyase (NosL): Evidence for radical formation at the amino group of tryptophan.* Angewandte Chemie, 2014. **126**(44): p. 12034-12038.

322. Laville, J., Blumer, C., Von Schroetter, C., Gaia, V., Défago, G., Keel, C., Haas, D., *Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent Pseudomonas fluorescens CHA0.* J Bacteriol, 1998. **180**(12): p. 3187-3196.

323. Dean, R., Padgett, H., Rapoport, H., *A high yield regiospecific preparation of iminium salts.* J Am Chem Soc, 1976. **98**(23): p. 7448-7449.

324. Jeoung, J., Dobbek, H., *Carbon dioxide activation at the Ni, Fe-cluster of anaerobic carbon monoxide dehydrogenase.* Science, 2007. **318**(5855): p. 1461-1464.

325. Reed, L., *From lipoic acid to multi-enzyme complexes.* Protein Sci, 1998. **7**(1): p. 220-224.

326. O'kane, D., Gunsalus, I., *Accessory factor requirement for pyruvate oxidation.* J Bacteriol, 1947. **54**(1): p. 20-20.

327. Kline, L., Barker, H., *A new growth factor required by Butyribacterium rettgeri.* J Bacteriol, 1950. **60**(3): p. 349-363.

328. Guirard, B., Snell, E., Williams, R., *The nutritional role of acetate for lactic acid bacteria, the response to substances related to acetate.* Archives of biochemistry, 1946. **9**: p. 361-379.

329. Holmquist, L., Stuchbury, G., Berbaum, K., Muscat, S., Young, S., Hager, K., Münch, G., *Lipoic acid as a novel treatment for Alzheimer's disease and related dementias.* Pharmacol Therapeut, 2007. **113**(1): p. 154-164.

330. Maczurek, A., Hager, K., Kenklies, M., Sharman, M., Martins, R., Engel, J., Münch, G., *Lipoic acid as an anti-inflammatory and neuroprotective treatment for Alzheimer's disease.* Adv Drug Delivery Rev, 2008. **60**(13): p. 1463-1470.

331. Borcea, V., Nourooz-Zadeh, J., Wolff, S., Klevesath, M., Hofmann, M., Urich, H., Nawroth, P., *α-Lipoic acid decreases oxidative stress even in diabetic patients with poor glycemic control and albuminuria.* Free Radical Bio Med, 1999. **26**(11): p. 1495-1500.

332. Packer, L., Kraemer, K., Rimbach, G., *Molecular aspects of lipoic acid in the prevention of diabetes complications.* Nutrition, 2001. **17**(10): p. 888-895.

333. Ziegler, D., Nowak, H., Kempler, P., Vargha, P., Low, P., *Treatment of symptomatic diabetic polyneuropathy with the antioxidant α-lipoic acid: a meta-analysis.* Diabetic Med, 2004. **21**(2): p. 114-121.

334. Smith, A., Shenvi, S., Widlansky, M., Suh, J., Hagen, T., *Lipoic acid as a potential therapy for chronic diseases associated with oxidative stress.* Curr Med Chem, 2004. **11**(9): p. 1135-1146.

335. Wollin, S., Jones, P., *α-Lipoic acid and cardiovascular disease.* J Nutr, 2003. **133**(11): p. 3327-3330.

336. Bilska, A., Wlodek, L., *Lipoic acid-the drug of the future.* Pharmacol Rep, 2005. **57**(5): p. 570-577.

337. Packer, L., Witt, E., Tritschler, H., *Alpha-lipoic acid as a biological antioxidant.* Free Radical Bio Med, 1995. **19**(2): p. 227-250.

338. Biewenga, G., Haenen, G., Bast, A., *The pharmacology of the antioxidant lipoic acid.* Gen Pharmacol-Vasc S, 1997. **29**(3): p. 315-331.

339. Moini, H., Packer, L., Saris, N., *Antioxidant and prooxidant activities of α-lipoic acid and dihydrolipoic acid.* Toxicol Appl Pharm, 2002. **182**(1): p. 84-90.

340. Çakatay, U., *Pro-oxidant actions of α-lipoic acid and dihydrolipoic acid.* Med hypotheses, 2006. **66**(1): p. 110-117.

341. Cronan, J., Zhao, X., Jiang, Y., *Function, attachment and synthesis of lipoic acid in Escherichia coli.* Adv Microb Physiol, 2005. **50**: p. 103-146.

342. Reed, L., *Multienzyme complexes.* Accounts Chem Res, 1974. **7**(2): p. 40-46.

343. Fujiwara, K., Okamura, K., Motokawa, Y., *Hydrogen carrier protein from chicken liver: purification, characterization, and role of its prosthetic group, lipoic acid, in the glycine cleavage reaction.* Arch Biochem Biophys, 1979. **197**(2): p. 454-462.

344. Koike, M., Reed, L., *α-keto acid dehydrogenation complexes II. the role of protein-bound lipoic acid and flavin adenine dinucleotide.* J Biol Chem, 1960. **235**(7): p. 1931-1938.

345. Perham, R., *Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions.* Annu Rev Biochem, 2000. **69**: p. 961-1004.

346. Howie, J., Houts, J., Sawyer, D., *Oxidation-reduction chemistry of DL-. alpha.-lipoic acid, propanedithiol, and trimethylene disulfide in aprotic and in aqueous media.* J Am Chem Soc, 1977. **99**(19): p. 6323-6326.

347. Reed, L., *A trail of research from lipoic acid to α-keto acid dehydrogenase complexes.* J Biol Chem, 2001. **276**(42): p. 38329-38336.

348. Reed, L., Hackert, M., *Structure-function relationships in dihydrolipoamide acyltransferases.* J Biol Chem, 1990. **265**: p. 8971-8974.

349. Zhao, X., Miller, J., Jiang, Y., Marletta, M., Cronan, J., *Assembly of the covalent linkage between lipoic acid and its cognate enzymes.* Chem Biol, 2003. **10**(12): p. 1293-1302.

350. Patel, M., Roche, T., *Molecular biology and biochemistry of pyruvate dehydrogenase complexes.* FASEB J, 1990. **4**(14): p. 3224-3233.

351. Rex Sheu, K., Blass, J., *The α-ketoglutarate dehydrogenase complex.* Ann NY Acad Sci, 1999. **893**(1): p. 61-78.

352. Perham, R., *Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein.* Biochemistry, 1991. **30**(35): p. 8501-8512.

353. Oppermann, F., Schmidt, B., Steinbüchel, A., *Purification and characterization of acetoin: 2, 6-dichlorophenolindophenol oxidoreductase, dihydrolipoamide dehydrogenase, and dihydrolipoamide acetyltransferase of the Pelobacter carbinolicus acetoin dehydrogenase enzyme system.* J Bacteriol, 1991. **173**(2): p. 757-767.

354. Kikuchi, G., Hiraga, K., *The mitochondrial glycine cleavage system.* Mol Cell Biol, 1982. **45**(3): p. 137-149.

355. Reed, L., Willms, C., *Purification and resolution of the pyruvate dehydrogenase complex (Escherichia coli).* Method Enzymol, 1966. **9**: p. 247-265.

356. Wieland, O., *The mammalian pyruvate dehydrogenase complex: structure and regulation.* Rev Physiol Bioch P, 1983. **96**: p. 123-170.

357. Xiao, Z., Xu, P., *Acetoin metabolism in bacteria.* Crit Rev Microbiol, 2007. **33**(2): p. 127-140.

358. Pettit, F.Y., S., Reed, L., *Purification and characterization of branched chain α-keto acid dehydrogenase complex of bovine kidney.* P Natl Acad Sci, 1978. **75**(10): p. 4881-4885.

359. Harper, A., Miller, R., Block, K., *Branched-chain amino acid metabolism.* Annu Rev Nutr, 1984. **4**(1): p. 409-454.

360. Smith, M., Neidhardt, F., *2-Oxoacid dehydrogenase complexes of Escherichia coli: cellular amounts and patterns of synthesis.* J Bacteriol, 1983. **156**(1): p. 81-88.

361. Burns, G., Brown, T., Hatter, K., Sokatch, J., *Comparison of the amino acid sequences of the transacylase components of branched chain oxoacid dehydrogenase of Pseudomonas*

*putida, and the pyruvate and 2-oxoglutarate dehydrogenases of Escherichia coli.* Eur J Biochem, 1988. **176**(1): p. 165-169.

362. Schellenberger, A., *Sixty years of thiamin diphosphate biochemistry.* Biochim Biophys Acta-Protein Struct M, 1998. **1385**(2): p. 177-186.

363. Danson, M., Ferscht, A., Perham, R., *Rapid intramolecular coupling of active sites in the pyruvate dehydrogenase complex of Escherichia coli: mechanism for rate enhancement in a multimeric structure.* P Natl Acad Sci, 1978. **75**(11): p. 5386-5390.

364. Berg, A., Westpahl, A., Bosma, H., De Kok, A., *Kinetics and specificity of reductive acylation of wild-type and mutated lipoyl domains of 2-oxo-acid dehydrogenase complexes from Azotobacter vinelandii.* Eur J Biochem, 1998. **252**(1): p. 45-50.

365. Griffin, T., Lau, K., Chuang, D., *Characterization and conservation of the inner E2 core domain structure of branched-chain alpha-keto acid dehydrogenase complex from bovine liver. Construction of a cDNA encoding the entire transacylase (E2b) precursor.* J Biol Chem, 1988. **263**(28): p. 14008-14014.

366. Lau, K., Lee, J., Fisher, C., Cox, R., Chuang, D., *Premature termination of transcription and alternative splicing in the human transacylase (E2) gene of the branched-chain α-ketoacid dehydrogenase complex.* FEBS letters, 1991. **279**(2): p. 229-232.

367. Packman, L., Hale, G., Perham, R., *Repeating functional domains in the pyruvate dehydrogenase multienzyme complex of Escherichia coli.* EMBO J, 1984. **3**(6): p. 1315-1318.

368. Milne, J., Wu, X., Borgnia, M., Lengyel, J., Brooks, B., Shi, D., Subramaniam, S., *Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy.* J Biol Chem, 2006. **281**(7): p. 4364-4370.

369. Mooney, B., Miernyk, Jan, A., Randall, D., *The complex fate of alpha-ketoacids.* Annu Rev Plant Biol, 2002. **53**: p. 357-375.

370. Sagers, R., Gunsalus, I., *Intermediary metabolism of Diplococcus glycinophilus I: glycine cleavage and one-carbon interconversions.* J Bacteriol, 1961. **81**(4): p. 541-549.

371. Kikuchi, G., *The glycine cleavage system: composition, reaction mechanism, and physiological significance.* Mol Cell Biol, 1973. **1**(2): p. 169-187.

372. Richert, D., Aamberg, R., Wilson, M., *Metabolism of glycine by avian liver.* J Biol Chem, 1962. **237**(1): p. 99-103.

373. Motokawa, Y., Kikuchi, G., *Glycine metabolism in rat liver mitochondria.* Arch Biochem Biophys, 1971. **146**(2): p. 461-466.

374. Pares, S., Cohen-Aaddad, C., Sieker, L.,Neuburger, M., Douce, R., *X-ray structure determination at 2.6-A resolution of a lipoate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves.* P Natl Acad Sci, 1994. **91**(11): p. 4850-4853.

375. Fujiwara, K., Motokawa, Y., *Mechanism of the glycine cleavage reaction. Steady state kinetic studies of the P-protein-catalysed reaction.* J Biol Chem, 1983. **258**(13): p. 8156-8162.

376. Guilhaudis, L., Simorre, J., Blackledge, M., Marion, D., Gans, P., Neuburger, M., Douce, R., *Combined structural and biochemical analysis of the HT complex in the glycine*

*decarboxylase cycle: evidence for a destabilization mechanism of the H-protein.* Biochemistry, 2000. **39**(15): p. 4259-4266.

377. Fujiwara, K., Okamura-Ikeda, K., Motokawa, Y., *Mechanism of the glycine cleavage reaction. Further characterization of the intermediate attached to H-protein and of the reaction catalysed by T-protein.* J Biol Chem, 1984. **259**(17): p. 10664-10668.

378. Faure, M., Bourguignon, J., Neuburger, M., Macherel, D., Sieker, L., Ober, R., Douce, R., *Interaction between the lipoamide-containing H-protein and the lipoamide dehydrogenase (L-protein) of the glycine decarboxylase multienzyme system.* Eur J Biochem, 2000. **267**(10): p. 2890-2898.

379. Turner, S., Ireland, R., Rawsthorne, S., *Purification and primary amino acid sequence of the L subunit of glycine decarboxylase. Evidence for a single lipoamide dehydrogenase in plant mitochondria.* J Biol Chem, 1992. **267**(11): p. 7745-7750.

380. Steiert, P., Stauffer, L., Stauffer, G., *The lpd gene product functions as the L protein in the Escherichia coli glycine cleavage enzyme system.* J Bacteriol, 1990. **172**(10): p. 6142-6144.

381. Morris, T., Reed, K.,. Cronan, J., *Lipoic acid metabolism in Escherichia coli: the lplA and lipB genes define redundant pathways for ligation of lipoyl groups to apoprotein.* J Bacteriol, 1995. **177**(1): p. 1-10.

382. Booker, S., *Unraveling the pathway of lipoic acid biosynthesis.* Chem Biol, 2004. **11**(1): p. 10-12.

383. Morris, T., Reed, K., Cronan, J., *Identification of the gene encoding lipoate-protein ligase A of Escherichia coli. Molecular cloning and characterization of the lplA gene and gene product.* J Biol Chem, 1994. **269**(23): p. 16091-16100.

384. Fujiwara, K., Toma, S., Okamura-Ikeda, K., Motokawa, Y., Nakagawa, A., Taniguchi, H., *Crystal structure of lipoate-protein ligase a from Escherichia coli Determination of the lipoic acid-binding site.* J Biol Chem, 2005. **280**(39): p. 33645-33651.

385. Jordan, S., Cronan, J., *The Escherichia coli lipB gene encodes lipoyl (octanoyl)-acyl carrier protein:protein transferase.* J Bacteriol, 2003. **185**(5): p. 1582-1589.

386. Martin, N., Christensen, Q., Mansilla, M., Cronan, J., de Mendoza, D., *A novel two-gene requirement for the octanoyltransfer reaction of Bacillus subtilis lipoic acid biosynthesis.* Mol Microbiol, 2011. **80**(2): p. 335-349.

387. Christensen, Q., Cronan, J. *Lipoic acid synthesis: a new family of octanoyltransferases generally annotated as lipoate protein ligases.* Biochemistry, 2010. **49**(46): p. 10024-10036.

388. C Christensen, Q., Martin, N., Mansilla, M., de Mendoza, D., Cronan, J., *A novel amidotransferase required for lipoic acid cofactor assembly in Bacillus subtilis.* Mol Microbiol, 2011. **80**(2): p. 350-363.

389. Posner, M., Upadhyay, A., Bagby, S., Hough, D., Danson, M., *A unique lipoylation system in the Archaea.* FEBS Journal, 2009. **276**(15): p. 4012-4022.

390. Christensen, Q.,Cronan, J., *The Thermoplasma acidophilum LplA-LplB complex defines a new class of bipartite lipoate-protein ligases.* J Biol Chem, 2009. **284**(32): p. 21317-21326.

391. Nesbitt, N., Baleanu-Gogonea, C., Cicchillo, R., Goodson, K., Iwig, D., Broadwater, J., Booker, S., *Expression, purification, and physical characterization of Escherichia coli lipoyl(octanoyl)transferase.* Protein Expr Purif, 2005. **39**(2): p. 269-282.

392. Parry, R, *Biosynthesis of lipoic acid. 1. Incorporation of specifically tritiated octanoic acid into lipoic acid.* J Am Chem Soc, 1977. **99**(19): p. 6464-6466.

393. Parry, R. Trainor, D., *Biosynthesis of lipoic acid .2. Stereochemistry of sulfur introduction at C-6 of octanoic-acid.* J Am Chem Soc, 1978. **100**(16): p. 5243-5244.

394. White, R., *Stable isotope studies on the biosynthesis of lipoic acid in Escherichia coli.* Biochemistry, 1980. **19**(1): p. 15-19.

395. Reed, K. Cronan, J., *Lipoic acid metabolism in Escherichia coli: sequencing and functional characterization of the lipA and lipB genes.* J Bacteriol, 1993. **175**(5): p. 1325-1336.

396. Busby, R., Schelvis, J., Yu, D., Babcock, G., Marletta, M., *Lipoic acid biosynthesis: LipA is an iron-sulfur protein.* J Am Chem Soc, 1999. **121**(19): p. 4706-4707.

397. Ollagnier-de Choudens, S., Fontecave, M., *The lipoate synthase from Escherichia coli is an iron-sulfur protein.* FEBS letters, 1999. **453**(1): p. 25-28.

398. Hayden, M., Huang, I., Iliopoulos, G., Orozco, M., Ashley, G., *Biosynthesis of lipoic acid: characterization of the lipoic acid auxotrophs Escherichia coli W1485-lip2 and JRG33-lip9.* Biochemistry, 1993. **32**(14): p. 3778-3782.

399. Zhao, X., Miller, J., Cronan, J., *The reaction of LipB, the octanoyl-[acyl carrier protein]:protein N-octanoyltransferase of lipoic acid synthesis, proceeds through an acyl-enzyme intermediate.* Biochemistry, 2005. **44**(50): p. 16737-16746.

400. Douglas, P., Kriek, M., Bryant, P., Roach, P., *Lipoyl synthase inserts sulfur atoms into an octanoyl substrate in a stepwise manner.* Angew Chem Int Ed Engl, 2006. **45**(31): p. 5197-5199.

401. Cicchillo, R. Booker, S., *Mechanistic investigations of lipoic acid biosynthesis in Escherichia coli: both sulfur atoms in lipoic acid are contributed by the same lipoyl synthase polypeptide.* J Am Chem Soc, 2005. **127**(9): p. 2860-2861.

402. Spalding, M., Prigge, S., *Lipoic acid metabolism in microbial pathogens.* Microbiol Mol Biol R, 2010. **74**(2): p. 200-228.

403. Ifuku, O., Kishimoto, J., Haze, S., Yanagi, M., Fukushima, S., *Conversion of dethiobiotin to biotin in cell-free extracts of Escherichia coli.* Biosci Biotech Bioch, 1992. **56**(11): p. 1780-1785.

404. Jarrett, J.T., *7 Biotin synthase: a role for iron-sulfur clusters in the radical-mediated generation of carbon-sulfur bonds.* Iron-sulfur clusters in chemistry and biology, 2014.

405. Florentin, D., Bui, B., Marquet, A., Ohshiro, T., Izumi, Y., *On the mechanism of biotin synthase of Bacillus sphaericus.* C R Acad Sci III, 1994. **317**(6): p. 485-488.

406. Guianvarc'h, D., Florentin, D., Bui, B., Nunzi, F., Marquet, A., *Biotin synthase, a new member of the family of enzymes which uses S-adenosylmethionine as a source of deoxyadenosyl radical.* Biochem Bioph Res Co, 1997. **236**(2): p. 402-406.

407. Bui, B., Florentin, D., Marquet, A., Benda, R., Trautwein, A., *Mössbauer studies of Escherichia coli biotin synthase: evidence for reversible interconversion between [2Fe-2S] 2+ and [4Fe-4S] 2+ clusters.* FEBS letters, 1999. **459**(3): p. 411-414.

408. Ugulava, N., Surerus, K., Jarrett, J., *Evidence from Mossbauer spectroscopy for distinct [2Fe-2S](2+) and [4Fe-4S](2+) cluster binding sites in biotin synthase from Escherichia coli.* J Am Chem Soc, 2002. **124**(31): p. 9050-9051.

409. Jameson, G., Cosper, M., Hernández, H., Johnson, M., Huynh, B., *Role of the [2Fe-2S] cluster in recombinant Escherichia coli biotin synthase.* Biochemistry, 2004. **43**(7): p. 2022-2031.

410. Ugulava, N., Sacanell, C., Jarrett, J., *Spectroscopic changes during a single turnover of biotin synthase: destruction of a [2Fe-2S] cluster accompanies sulfur insertion.* Biochemistry, 2001. **40**(28): p. 8352-8358.

411. Tse Sum Bui, B., Mattioli, T., Florentin, D., Bolbach, G., Marquet, A., *Escherichia coli biotin synthase produces selenobiotin. Further evidence of the involvement of the [2Fe-2S]2+ cluster in the sulfur insertion step.* Biochemistry, 2006. **45**(11): p. 3824-3834.

412. Fontecave, M., Ollagnier-de-Choudens, S., Mulliez, E., *Biological radical sulfur insertion reactions.* Chem Rev, 2003. **103**(6): p. 2149-2166.

413. Begley, T., Xi, J., Kinsland, C., Taylor, S., McLafferty, F., *The enzymology of sulfur activation during thiamin and biotin biosynthesis.* Curr Opin Chem Biol, 1999. **3**(5): p. 623-629.

414. Cicchillo, R., Iwig, D., Jones, A., Nesbitt, N., Baleanu-Gogonea, C., Souder, M., Booker, S., *Lipoyl synthase requires two equivalents of S-adenosyl-L-methionine to synthesize one equivalent of lipoic acid.* Biochemistry, 2004. **43**(21): p. 6378-6386.

415. Cicchillo, R., Lee, K., Baleanu-Gogonea, C., Nesbitt, N., Krebs, C., Booker, S., *Escherichia coli lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide.* Biochemistry, 2004. **43**(37): p. 11770-11781.

416. Lanz, N., Pandelia, M., Kakar, E., Lee, K., Krebs, C., Booker, S., *Evidence for a Catalytically and Kinetically Competent Enzyme-Substrate Cross-linked Intermediate in Catalysis by Lipoyl Synthase.* Biochemistry, 2014. **53**(28): p. 4557-4572.

417. Lanz, N., Rectenwald, J., Wang, B., Kakar, E., Laremore, T., Booker, S., Silakov, A., *Characterization of a radical intermediate in lipoyl cofactor biosynthesis.* J Am Chem Soc, 2015. **137**(41): p. 13216-13219.

418. Deurenberg, R., Stobberingh, E., *The evolution of Staphylococcus aureus.* Infection, genetics and evolution, 2008. **8**(6): p. 747-763.

419. Lindsay, J., *Staphylococcus: molecular genetics.* Horizon Scientific Press, 2008.

420. Grundmann, H., Aires-de-Sousa, M., Boyce, J., Tiemersma, E., *Emergence and resurgence of meticillin-resistant Staphylococcus aureus as a public-health threat.* Lancet, 2006. **368**(9538): p. 874-885.

421. Livermore, D., *Antibiotic resistance in staphylococci.* Int J Antimicrob Ag, 2000. **16**: p. 3-10.

422. Reacher, M., Shah, A., Livermore, D., Wale, M., Graham, C., Johnson, A., George, R., *Bacteraemia and antibiotic resistance of its pathogens reported in England and Wales between 1990 and 1998: trend analysis.* Bmj, 2000. **320**(7229): p. 213-216.

423. Henwood, C., Livermore, D., Johnson, A., James, D., Warner, M., Gardiner, A., *Susceptibility of Gram-positive cocci from 25 UK hospitals to antimicrobial agents including linezolid.* J Antimicrob Chemoth, 2000. **46**(6): p. 931-940.

424. Hutchinson, D., *Oxazolidinone antibacterial agents: a critical review.* Curr Top Med Chem, 2003. **3**(9): p. 1021-1042.

425. Livermore, D., *Linezolid in vitro: mechanism and antibacterial spectrum.* J Antimicrob Chemoth, 2003. **51**(suppl 2): p. ii9-ii16.

426. Polaceck, N., Mankin, A., *The ribosomal peptidyltransferase center: Structure, function, evolution, inhibition.* Crit Rev Biochem Mol, 2005. **40**(5): p. 285-311.

427. Dunkle, J., Xiong, L., Mankin, A., Cate, J., *Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyltransferase center explain spectra of drug action.* P Natl Acad Sci, 2010. **107**(40): p. 17152-17157.

428. Poehlsgaard, J., Douthwaite, S., *The bacterial ribosome as a target for antibiotics.* Nature Rev Microbiol, 2005. **3**(11): p. 870-881.

429. Matassova, N., Rodnina, M., Endermann, R., Kroll, H., Pleiss, U., Wild, H., Wintermeyer, W., *Ribosomal RNA is the target for oxazolidinones, a novel class of translational inhibitors.* RNA, 1999. **5**(07): p. 939-946.

430. Xiong, L., Kloss, P., Douthwaite, S., Andersen, N., Swaney, S., Shinabarger, D., Mankin, A. S., *Oxazolidinone resistance mutations in 23S rRNA ofEscherichia coli reveal the central region of domain V as the primary site of drug action.* J Bacteriol, 2000. **182**(19): p. 5325-5331.

431. Wilson, D., Schluenzen, F., Harms, J., Starosta, A., Connell, S., Fucini, P., *The oxazolidinone antibiotics perturb the ribosomal peptidyl-transferase center and effect tRNA positioning.* P Natl Acad Sci, 2008. **105**(36): p. 13339-13344.

432. Tissieres, A., Watson, J., Schlessinger, D., Hollingworth, B., *Ribonucleoprotein particles from Escherichia coli.* J Mol Biol, 1959. **1**(3): p. 221-233.

433. Kurland, C., *Structure and function of the bacterial ribosome.* Annu Rev Biochem, 1972. **41**(1): p. 377-408.

434. Martinko, J., Madigan, M., *Brock biology of microorganisms.* 2005.

435. Maden, B., Monro, R., *Ribosome-catalyzed peptidyl transfer.* Eur J Biochem, 1968. **6**(2): p. 309-316.

436. Trobro, S., Å., J., *Mechanism of peptide bond synthesis on the ribosome.* P Natl Acad Sci USA, 2005. **102**(35): p. 12395-12400.

437. Nissen, P., Hansen., J., Ban, N., Moore, P., Steitz, T., *The structural basis of ribosome activity in peptide bon synthesis.* Science, 2000. **289**(5481): p. 920-930.

438. Sievers, A., Beringer, M., Rodnina, M., Wolfenden, R., *The ribosome as an entropy trap.* P Natl Acad Sci USA, 2004. **101**(21): p. 7897-7901.

439. Trobro, S., Åqvist., J., *A model for how ribosomal release factors induce peptidyl-tRNA cleavage in termination of protein synthesis.* Mol Cell, 2007. **27**(5): p. 758-766.

440. Brown, C., Stockwell, P., Trotman, C. , Tate, W., *The signal for the termination of protein synthesis in procaryotes.* Nucleic Acids Res, 1990. **18**(8): p. 2079-2086.

441. Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., Franceschi, F., *Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria.* Nature, 2001. **413**(6858): p. 814-821.

442. Hansen, J., Moore, P., Steitz, T., *Structures of five antibiotics bound at the peptydiltransferase centre of the large ribosomal unit.* J Mol Biol, 2003. **330**(5): p. 1061-1075.

443. Ippolito, J., Kanyo, Z., Wang, D., Franceschi, F., Moore, P., Steitz, T., Duffy, E., *Crystal Structure of the oxazolidinone antibiotic linezolid bound to the 50S ribosomal subunit.* J Med Chem, 2008. **51**(12): p. 3353-3356.

444. Weisblum, B., *Erythromycin resistance by ribosome modification.* Antimicrob Agents Ch, 1995. **39**(3): p. 577-585.

445. Prunier, A., Malbruny, B., Tandé, D., Picard, B., Leclercq, R., *Clinical isolates of Staphylococcus aureus with ribosomal mutations conferring resistance to macrolides.* Antimicrob Agents Ch, 2002. **46**(9): p. 3054-3056.

446. Canu, A., Malbruny, B., Coquemont, M., Davies, T. A., Appelbaum, P. C., Leclercq, R., *Diversity of ribosomal mutations conferring resistance to macrolides, clindamycin, streptogramin, and telithromycin in Streptococcus pneumoniae.* Antimicrob Agents Ch, 2002. **46**(1): p. 125-131.

447. Tsiodras, S., Gold, H., Sakoulas, G., Eliopoulos, G., Wennersten, C., Venkataraman, L., Ferraro, M. J., *Linezolid resistance in a clinical isolate of Staphylococcus aureus.* Lancet, 2001. **358**(9277): p. 207-208.

448. Pantosti, A., Sanchini, A., Monaco, M., *Mechanisms of antibiotic resistance in Staphylococcus aureus.* Future Microbiol, 2007. **2**(3) p.323-334.

449. Marshall, S., Donskey, C., Hutton-Thomas, R., Salata, R., Rice, L., *Gene dosage and linezolid resistance in Enterococcus faecium and Enterococcus faecalis.* Antimicrob Agents Ch, 2002. **46**(10): p. 3334-3336.

450. Lobritz, M., Hutton-Thomas, R., Marshall, S., Rice, L., *Recombination proficiency influences frequency and locus of mutational resistance to linezolid in Enterococcus faecalis.* Antimicrob Agents Ch, 2003. **47**(10): p. 3318-3320.

451. Wolter, N., Smith, A., Farrell, D., Klugman, K., *Heterogeneous macrolide resistance and gene conversion in the pneumococcus.* Antimicrob Agents Ch, 2006. **50**(1): p. 359-361.

452. Andersson, D., Levin, B., *The biological cost of antibiotic resistance.* Curr Opin Microbiol, 1999. **2**(5): p. 489-493.

453. Schwarz, S., Werckenthin, C., Kehrenberg, C., *Identification of a plasmid-borne chloramphenicol-florfenicol resistance gene in Staphylococcus sciuri.* Antimicrob Agents Ch, 2000. **44**(9): p. 2530-2533.

454. Dai, L., Wu, C., Wang, M., Wang, Y., Wang, Y., Huang, S., Shen, J., *First report of the multidrug resistance gene cfr and the phenicol resistance gene fexA in a Bacillus strain from swine feces.* Antimicrob Agents Ch, 2010. **54**(9): p. 3953-3955.

455. Wang, Y., Wang, Y., Wu, C., Schwarz, S., Shen, Z., Zhang, W., Shen, J., *Detection of the staphylococcal multiresistance gene cfr in Proteus vulgaris of food animal origin.* J Antimicrob Chemoth, 2011. **66**(11): p. 2521-2526.

456. Shen, J., Wang, Y., Schwarz, S., *Presence and dissemination of the multiresistance gene cfr in Gram-positive and Gram-negative bacteria.* J Antimicrob Chemoth, 2013. **dkt092**.

457. Kehrenberg, C., Schwarz, S., Jacobsen, L., Hansen, L., Vester, B., *A new mechanism for chloramphenicol, florfenicol and clindamycin resistance: methylation of 23S ribosomal RNA at A2503.* Mol Microbiol, 2005. **57**(4): p. 1064-1073.

458. Giessing, A., Jensen, S., Rasmussen, A., Hansen, L., Gondela, A., Long, K., Kirpekar, F., *Identification of 8-methyladenosine as the modification catalyzed by the radical SAM methyltransferase Cfr that confers antibiotic resistance in bacteria.* RNA, 2009. **15**(2): p. 327-336.

459. Ban, N., Nissen, P., Hansen, J., Moore, P., Steitz, T., *The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution.* Science, 2000. **289**(5481): p. 905-920.

460. Smith, L., Mankin, A., *Transcriptional and translational control of the mlr operon, which confers resistance to seven classes of protein synthesis inhibitors.* Antimicrob Agents Ch, 2008. **52**(5): p. 1703-1712.

461. Booth, M., Challand, M., Emery, D., Roach, P., Spencer, J., *High-level expression and reconstitution of active Cfr, a radical-SAM rRNA methyltransferase that confers resistance to ribosome-acting antibiotics.* Protein Expr Purif, 2010. **74**(2): p. 204-210.

462. Long, K., Poehlsgaard, J., Kehrenbeg, C., Schwarz, S., Vester, B., *The Cfr rRNA methyltransferase confers resistance to phenicols, lincosamides, oxazolidinones, pleuromutilins, and streptogramin A antibiotics.* Antimicrob Agents Ch, 2006. **50**(7): p. 2500-2505.

463. Kloss, P., Xiong, L., Shinabarger, D., Mankin, A., *Resistance mutations in 23 S rRNA identify the site of action of the protein synthesis inhibitor linezolid in the ribosomal peptidyl transferase center.* J Mol Biol, 1999. **294**(1): p. 93-101.

464. Kaminska, K., Purta, E., Hansen, L., Bujnicki, J., Vester, B., Long, K., *Insights into the structure, function and evolution of the radical-SAM 23S rRNA methyltransferase Cfr that confers antibiotic resistance in bacteria.* Nucleic Acids Res, 2010. **38**(5): p. 1652-1663.

465. Toh, S., Xiong, L., Bae, T., Mankin, A., *The methyltransferase YfgB/RlmN is responsible for modification of adenosine 2503 in 23S rRNA.* RNA, 2008. **14**(1): p. 98-106.

466. Vázquez-Laslop, N., Ramu, H., Klepacki, D., Kannan, K., Mankin, A., *The key function of a conserved and modified rRNA residue in the ribosomal response to the nascent peptide.* EMBO J, 2010. **29**(18): p. 3108-3117.

467. Benítez-Páez, A., Villarroya, M., Armengod, M., *The Escherichia coli RlmN methyltransferase is a dual-specificity enzyme that modifies both rRNA and tRNA and controls translational accuracy.* RNA, 2012. **18**(10): p. 1783-1795.

468. LaMarre, J., Howden, B., Mankin, A., *Inactivation of the indigenous methyltransferase RlmN in Staphylococcus aureus increases linezolid resistance.* Antimicrob Agents Ch, 2011. **55**(6): p. 2989-2991.

469. Czerwoniec, A., Dunin-Horkawicz, S., Purta, E., Kaminska, K., Kasprzak, J., Bujnicki, J., Grosjean, H., Rother, K., *MODOMICS: a database of RNA modification pathways 2008 update.* Nucleic Acid Res, 2009. **37**: p. D118-D121.

470. Cantara, W., Crain, P., Rozenski, J., McCloskey, J., Harris, K., Zhang, X., Vendeix, F., Fabris, D., Agris, P., *The RNA modification database, RNAMDB: 2011 update.* Nucleic Acids Res, 2010: p. D195-D201.

471. Czerwoniec, A., Kasprzak, J., Kaminska, K., Rother, K., Purta, E., Bujnicki, J., *Folds and functions of domains in RNA modification enzymes. In DNA and RNA modification enzymes: structure, mechanism, functions, and evolution.* Landes, Bioscience, Austin, TX. 2003.

472. Schubert, H., Blumenthal, R., Cheng, X., *Many paths to methyltransfer: a chronicle of convergence.* Trends Biochem Sci, 2003. **28**(6): p. 329-335.

473. Fauman, E., Blumenthal, R., Cheng, X., *Structure and evolution of AdoMet-dependent methyltransferases. S-adenosylmethionine-dependent methyltransferases: structures and functions.* World Scientific Publishing, Singapore, 1999.

474. Anantharaman, V., Koonin, E., Aravind, L., *SPOUT: a class of methyltransferases that includes spoU and trmD RNA methylase superfamilies, and novel superfamilies of predicted prokaryotic RNA methylases.* J Mol Microbiol Biotechn, 2002. **4**(1): p. 71-76.

475. Urbonavicius, J., Brochier-Armanet, C., Skouloubris, S., Myllykallio, H., Grosjean, H., *In vitro detection of the enzymatic activity of folate-dependent tRNA (Uracil-54,-C5)-methyltransferase: evolutionary implications.* Method Enzymol, 2007. **425**: p.103-119.

476. Grove, T., Radle, M., Krebs, C., Booker, S., *Cfr and RlmN Contain a Single [4Fe-4S] Cluster, which Directs Two Distinct Reactivities for S-Adenosylmethionine: Methyl Transfer by SN2 Displacement and Radical Generation.* J Am Chem Soc, 2011. **133**(49), 19586-19589

477. Grove, T., Benner, J., Radle, M., Ahlum, J., Landgraf, B., Krebs, C., Booker, S., *A radically different mechanism for S-adenosylmethionine-dependent methyltransferases.* Science, 2011. **332**(6029): p. 604-607.

478. Silakov, A., Grove, T., Radle, M., Bauerle, M., Green, M., Rosenzweig, A., Booker, S., *Characterization of a cross linked protein–nucleic acid substrate radical in the reaction catalyzed by RlmN.* J Am Chem Soc, 2014. **136**(23): p. 8221-8228.

479. Matthews, R., *Cobalamin-dependent methyltransferases.* Accounts Chem Res, 2001. **34**(8): p. 681-689.

480. Chan, K., Thompson, S., O'Hagan, D., *The mechanisms of radical SAM/cobalamin methylations: an evolving working hypothesis.* Chem Bio Chem, 2013. **14**(6): p. 675-677.

481. Ávan der Donk, W., *New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin.* Chem Commun, 2007. **4**: p. 359-361.

482. Allen, K., Wang, S., *Initial characterization of Fom3 from Streptomyces wedmorensis: the methyltransferase in fosfomycin biosynthesis.* Arch Biochem Biophys, 2014. **543**: p. 67-73.

483. Kim, H., McCarty, R., Ogasawara, Y., Liu, Y., Mansoorabadi, S., LeVieux, J., Liu, H., *GenK-catalyzed C-6' methylation in the biosynthesis of gentamicin: isolation and characterization of a cobalamin-dependent radical SAM enzyme.* J Am Chem Soc, 2013. **135**(22): p. 8093-8096.

484. Werner, W., Allen, K., Hu, K., Helms, G., Chen, B., Wang, S., *In vitro phosphinate methylation by PhpK from Kitasatospora phosalacinea.* Biochemistry, 2011. **50**(42): p. 8986-8988.

485. Yu, Y., Duan, L., Zhang, Q., Liao, R., Ding, Y., Pan, H., Wendt-Pienkowski, E., Tang, G., Shen , B., Liu, W., *Nosiheptide biosynthesis featuring a unique indole side ring formation on the characteristic thiopeptide framework.* ACS Chem Biol, 2009. **4**(10): p. 855-864.

486. Morris, R., Leeds, J., Naegeli, H., Oberer, L., Memmert, K., Weber, E., LaMarche, M., Parker, C., Burrer, N., Esterow, S., Hein, A., Schmitt, E., Krastel, P., *Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu.* J Am Chem Soc, 2009. **131**(16): p. 5946-5955.

487. Allen, K., Xu, H., White, R., *Identification of a Unique Radical S-Adenosylmethionine Methylase likely involved in methanopterin biosynthesis in methanocaldococcus jannaschii.* J Bacteriol, 2014. **196**(18): p. 3315-3323.

488. Bauerle, M., Schwalm, E., Booker, S., *Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation.* J Biol Chem, 2015. **290**(7): p. 3995-4002.

489. Westrich, L., Heide, L., Li, S., *CloN6, a novel methyltransferase catalysing the methylation of the pyrrole-2-carboxyl moiety of clorobiocin.* Chem Bio Chem, 2003. **4**(8): p. 768-773.

490. Ding, Y., Yu, Y., Pan, H., Guo, H., Li, Y., Liu, W., *Moving posttranslational modifications forward to biosynthesize the glycosylated thiopeptide nocathiacin I in Nocardia sp. ATCC202099.* Mol BioSystems, 2010. **6**(7): p. 1180-1185.

491. Du, L., Sánchez, C., Chen, M., Edwards, D. J., Shen, B., *The biosynthetic gene cluster for the antitumor drug bleomycin from Streptomyces verticillus ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase.* Chem Biol, 2000. **7**(8): p. 623-642.

492. Galm, U., Wendt-Pienkowski, E., Wang, L., George, N., Oh, T., Yi, F., Tao, M., Coughlin, J., Shen, B., *The biosynthetic gene cluster of zorbamycin, a member of the bleomycin family of antitumor antibiotics, from Streptomyces flavoviridis ATCC 21892.* Mol BioSystems, 2009. **5**(1): p. 77-90.

493. Huang, W., Xu, H., Li, Y., Zhang, F., Chen, X. Y., He, Q. L., Tang, G. L., *Characterization of yatakemycin gene cluster revealing a radical S-adenosylmethionine dependent methyltransferase and highlighting spirocyclopropane biosynthesis.* J Am Chem Soc, 2012. **134**(21): p. 8831-8840.

494. Hiratsuka, T., Suzuki, H., Kariya, R., Seo, T., Minami, A., Oikawa, H., *Biosynthesis of the Structurally Unique Polycyclopropanated Polyketide–Nucleoside Hybrid Jawsamycin (FR-900848).* Angew Chem Int Edit, 2014. **53**(21): p. 5423-5426.

495. Zhang, Q., Van Der Donk, W., Liu, W., *Radical-mediated enzymatic methylation: a tale of two SAMS.* Accounts Chem Res, 2011. **45**(4): p. 555-564.

496. Boal, A., Grove, T., McLaughlin, M., Yennawar, N., Booker, S., Rosenzweig, A., *Structural basis for methyl transfer by a radical SAM enzyme.* Science, 2011. **332**(6033): p. 1089-1092.

497. Frey, P., *Enzymatic Reaction Mechanisms.* Oxford University Press, New York, 2007.

498. Coward, J., Sweet, W., *Kinetics and mechanism of methyl transfer from sulfonium compounds to various nucleophiles.* J Org Chem, 1971. **36**(16): p. 2337-2346.

499. Lin, H., *S-Adenosylmethionine-dependent alkylation reactions: When are radical reactions used?* Bioorgan Chem, 2011. **39**(5): p. 161-170.

500. Yan, F., Fujimori, D., *RNA methylation by radical SAM enzymes RlmN and Cfr proceeds via methylene transfer and hydride shift.* Proc Natl Acad Sci USA, 2011. **108**(10): p. 3930-3934.

501. Challand, M., Salvadori, E., Driesener, R., Kay, C., Roach, P., Spencer, J., *Cysteine Methylation Controls Radical Generation in the Cfr Radical AdoMet rRNA Methyltransferase.* PLoS One, 2013. **8**(7): p. e67979.

502. Grove, T., Livada, J., Schwalm, E., Green, M., Booker, S., Silakov, A., *A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr.* Nat Chem Biol, 2013. **9**(7): p. 422-427.

503. Hioe, J., Zipse, H., *Hydrogen Transfer in SAM-Mediated Enzymatic Radical Reactions.* Chem-Eur J, 2012. **18**(51): p. 16463-16472.

504. Doherty, A., Serpell, L., Ponting, C., *The helix-hairpin-helix DNA-Binding Motif: A structural basis for non-sequence specific recognition of DNA.* Nucleic Acids Res, 1996. **24**(13): p. 2488-2497.

505. Mulder, D., Shepard, E., Meuser, J., Joshi, N., King, P., Posewitz, M., Broderick, J., Peters, J., *Insights into [FeFe]-hydrogenase structure, mechanism, and maturation.* Structure, 2011. **19**(8): p. 1038-1052.

506. Swanson, K., Duffus, B., Beard, T., Peters, J., Broderick, J., *Cyanide and carbon monoxide ligand formation in hydrogenase biosynthesis.* Eur J Inorg Chem, 2011. **2011**(7): p. 935-947.

507. Mulder, D., Boyd, E., Sarma, R., Lange, R., Endrizzi, J., Broderick, J., Peters, J., *Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA(DeltaEFG).* Nature, 2010. **465**(7295): p. 248-251.

508. Kuchenreuther, J., Grady-Smith, C., Bingham, A., George, S., Cramer, S., Swartz, J., *High-Yield Expression of Heterologous [FeFe] Hydrogenases in Escherichia coli.* Plos One, 2010. **5**(11): p. e15491.

509. Dinis, P., Suess, D., Fox, S., Harmer, J., Driesener, R., De La Paz, L., Essex, J., Britt, D., Roach, P., *X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly.* P Natl Acad Sci, 2015. **112**(5): p. 1362–1367.

510. Dale, G., Oefner, C., D'Arcy, A., *The protein as a variable in protein crystallization.* J Struct Biol, 2003. **142**(1): p. 88-97.

511. Guharoy, M., Chakrabarti, P., *Conserved residue clusters at protein-protein interfaces and their use in binding site identification.* BMC Bioinformatics, 2010. **11**(1): p. 286.

512. Ng, P., Henikoff, S., *Predicting the effects of amino acid substitutions on protein function.* Annu. Rev. Genomics Hum. Genet., 2006. **7**: p. 61-80.

513. Guo, H. Choe, J., Loeb, L., *Protein tolerance to random amino acid change.* P Natl Acad Sci USA, 2004. **101**(25): p. 9205-9210.

514. Tokuriki, N., Tawfik, D., *Stability effects of mutations and protein evolvability.* Curr Opin Struc Biol, 2009. **19**(5): p. 596-604.

515. Rhodes, G., *Crystallography made crystal clear: a guide for users of macromolecular models.* Academic Press, 2010.

516. Savchenko, A., Yee, A., Khachatryan, A., Skarina, T., Evdokimova, E., Pavlova, M., Edwards, A. , *Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches.* Proteins: Structure, Function, and Bioinformatics, 2003. **50**(3): p. 392-399.

517. Kumar, S., Nussinov, R., *Salt bridge stability in monomeric proteins.* J Mol Biol, 1999. **293**(5): p. 1241-1255.

518. Jenney Jr, F., Adams, M., *The impact of extremophiles on structural genomics (and vice versa).* Extremophiles, 2008. **12**(1): p. 39-50.

519. Driesener, R., *Mechanistic and structural characterisation of HydG catalysed L-tyrosine cleavage.* Doctoral dissertation, University of Southampton, 2014.

520. Martins, F., *Studies in thiamine biosynthesis*, Doctoral dissertation, University of Southampton, 2010.

521. Hochuli, E., Döbeli, H., Schacher, A., *New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues.* J Chromatography A, 1987. **411**: p. 177-184.

522. Guzman, D., Carson, M., Beckwith, J., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-4130.

523. Kriek, M., Petersa, L., Takahashib, Y., Roach, P., *Effect of iron-sulfur cluster assembly proteins on the expression of Escherichia coli lipoic acid synthase.* Protein Expr Purif, 2003. **28**(2): p. 241-245.

524. Sørensen, H., Mortensen, K., *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* J Biotechn, 2005. **115**(2): p. 113-128.

525. Glusker, J., *Structural aspects of metal liganding to functional groups in proteins.* Adv. Protein Chem, 1991. **42**(1): p. 76.

526. Barth, H., Jackson, C., Boyes, B., *Size exclusion chromatography.* Anal Chem, 1994. **66**(12): p. 595R-620R.

527. Lanz, N. Grove, T., Gogonea, C., Lee, K., Krebs, C., Booker, S., *RlmN and AtsB as models for the overproduction and characterization of radical SAM proteins.* Method Enzymol, 2011. **516**: p. 125-152.

528. Abràmoff, M., Magalhães, P., Ram, S., *Image processing with ImageJ.* Biophotonics Int, 2004. **11**(7): p. 36-42.

529. Bradford, M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**(1): p. 248-254.

530. Baneyx, F., *Recombinant protein expression in Escherichia coli.* Curr Opin Biotech, 1999. **10**(5): p. 411-421.

531. Nicolet, Y., Fontecilla-Camps, J., *10 Iron-sulfur clusters and molecular oxygen: function, adaptation, degradation, and repair.* Iron-Sulfur Clusters in Chemistry and Biology, 2014.

532. Sundberg, R., Martin, R., *Interactions of histidine and other imidazole derivatives with transition metal ions in chemical and biological systems.* Chem Rev, 1974. **74**(4): p. 471-517.

533. Malkin, R., Rabinowitz, J., *The reconstitution of clostridial ferredoxin.* Biochem Bioph Res Co, 1966. **23**(6): p. 822-827.

534. Sweeney, W., Rabinowitz, J., *Proteins containing 4Fe-4S clusters: an overview.* Annu Rev Biochem, 1980. **49**(1): p. 139-161.

535. Lill, R., *Function and biogenesis of iron–sulphur proteins.* Nature, 2009. **460**(7257): p. 831-838.

536. Fish, W., *Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples.* Methods Enzymol, 1988. **158**: p. 357-64.

537. Thayer, M., Ahern, H., Xing, D., Cunningham, R., Tainer, J., *Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure.* EMBO J, 1995. **14**(16): p. 4108.

538. Tracqui, A., Raul, J., Geraut, A., Berthelon, L., Ludes, B., *Determination of blood cyanide by HPLC-MS.* J Anal Toxicol, 2002. **26**(3): p. 144-148.

539. Layer, G., Grage, K., Teschner, T., Schünemann, V., Breckau, D., Masoumi, A., Jahn, M., Heathcote, P., Trautwein, A., Jahn, D., *Radical S-adenosylmethionine enzyme coproporphyrinogen III oxidase HemN: functional features of the [4Fe-4S] cluster and the two bound S-adenosyl-L-methionines.* J Biol Chem, 2005. **280**(32): p. 29038-29046.

540. Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E., Kern, D., *Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair.* Nature Struct Mol Biol, 2004. **11**(10): p. 945-949.

541. Lam, S., Yeung, R., Yu, T., Sze, K., Wong, K., *A rigidifying salt-bridge favors the activity of thermophilic enzyme at high temperatures at the expense of low-temperature activity.* PLoS Biol, 2011. **9**(3): p. e1001027.

542. Siddiqui, K., Thomas, T., *Protein adaptation in extremophiles.* Nova Publishers, 2008.

543. Harmer, J., Hiscox, M., Dinis, P., Fox, S., Iliopoulos, A., Hussey, J., Sandy, J., Van Beek, F., Essex, J., Roach, P., *Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions.* Biochem J, 2014. **464**(1): p. 123-133.

544. McPherson, A., *Crystallization of biological macromolecules.* Cold Spring Harbor Laboratory Press, 1999.

545. Ducruix, A., Giegé, R., *Crystallization of nucleic acids and proteins.* IRL Press at Oxford University Press, 1992.

546. Bergfors, T., *Seeds to crystals.* J Strut Biol, 2003. **142**(1): p. 66-76.

547. Stura, E.A., *Seeding.* Protein crystallization: techniques, strategies, and tips: a laboratory manual, 1999. International University Line, La Jolla, Calif: p. 139-153.

548. Mitchell, E., Garman, E., *Flash freezing of protein crystals: investigation of mosaic spread and diffraction limit with variation of cryoprotectant concentration.* J Appl Crystallogr, 1994. **27**(6): p. 1070-1074.

549. Henderson, R., *Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction.* P Biol Sci, 1990. **241**(1300): p. 6-8.

550. Dauter, Z., Dauter, M., Dodson, E., *Jolly sad.* Acta Crystallogr D, 2002. **58**(3): p. 494-506.

551. Incardona, M., Bourenkov, G., Levik, K., Pieritz, R., Popov, A., Svensson, O., *EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis.* J Synchrotron Radiat, 2009. **16**(6): p. 872-879.

552. Ruso, J., Piñeiro, Á., *Proteins in solution and at interfaces: methods and applications in biotechnology and materials science.* John Wiley Sons, Inc., 2013.

553. Ooi, L., *Principles of X-ray Crystallography.* Oxford University Press, 2010.

554. Kabsch, W., *Xds.* Acta Crystallogr D, 2010. **66**(2): p. 125-132.

555. Evans, P., *Scaling and assessment of data quality.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 72-82.

556. Messerschmidt, A., *X-ray crystallography of biomacromolecules: a practical guide.* John Wiley Sons, Inc., 2007.

557. McCoy, A., *Solving structures of protein complexes by molecular replacement with Phaser.* Acta Crystallogr D, 2007. **63**(1): p. 32-41.

558. Adams, P. Grosse-Kunstleve, R., Hung, L., Ioerger, T., McCoy, A., Moriarty, N., Terwilliger, T., *PHENIX: building new software for automated crystallographic structure determination.* Acta Crystallogr D, 2002. **58**(11): p. 1948-1954.

559. Adams, P. Afonine, P., Bunkóczi, G., Chen, V., Davis, I., Echols, N., Headd, J., Hung, L., Kapral, G., Grosse-Kunstleve, R., McCoy, A., Moriarty, N., Oeffner, R., Read, R., Richardson, D., Richardson, J., Terwilliger T., Zwart, P., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-221.

560. Terwilliger, T., Grosse-Kunstleve, R., Afonine, P., Moriarty, N., Zwart, P., Hung, L., Read, R., Adams, P., *Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 1): p. 61-69.

561. Emsley, P., Cowtan, K., *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-2132.

562. Laskowski, R., MacArthur, M., Thornton, J., *Validation of protein models derived from experiment.* Curr Opin Struc Biol, 1998. **8**(5): p. 631-639.

563. Brunger, A., Kuriyan, J., Karplus, M., *Crystallographic R factor refinement by molecular dynamics.* Science, 1987. **235**(4787): p. 458-460.

564. Urzhumtseva, L., Afonine, P., Adams, P., Urzhumtsev, A., *Crystallographic model quality at a glance.* Acta Crystallogr D, 2009. **65**(3): p. 297-300.

565. http://www.pymolwiki.org/index.php/RotationAxis.

566. Goldman, P., Groveb, T., Bookerb, S., Drennan, C., *X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry.* Proc Natl Acad Sci USA, 2013. **110**(40): p. 15949-15954.

567. Macedo, S., Mitchell, E., Romão, C., Cooper, S., Coelho, R., Liu, M., Xavier, A., LeGall, J., Bailey, S., Garner, D., Hagen, W., Teixeira, M., Carrondo, M., Lindley, P., *Hybrid cluster proteins (HCPs) from Desulfovibrio desulfuricans ATCC 27774 and Desulfovibrio vulgaris (Hildenborough): X-ray structures at 1.25 Å resolution using synchrotron radiation.* J Biol Inorg Chem, 2002. **7**(4-5): p. 514-525.

568. Parthasarathy, S., Murthy, M., *Analysis of temperature factor distribution in high-resolution protein structures.* Protein Sci, 1997. **6**(12): p. 2561-2567.

569. Huang, B., *MetaPocket: a meta approach to improve protein ligand binding site prediction.* OMICS A Journal of Integrative Biology, 2009. **13**(4): p. 325-330.

570. Minnihan, E., Nocera, D., Stubbe, J., *Reversible, long-range radical transfer in E. coli class Ia ribonucleotide reductase.* Acc Chem Res, 2013. **46**(11): p. 2524-2535.

571. Ji, X., Li, Y., Ding, W., Zhang, Q., *Substrate-tuned catalysis of the radical S-adenosyl-L-methionine enzyme NosL involved in nosiheptide niosynthesis.* Angewandte Chemie, 2015. **127**(31): p. 9149-9152.

572. Halgren, T., *The Merck Molecular Force Field: Application and validation for ''real world'' problems.* Abstracts of Papers of the American Chemical Society, 1996. **212**: p. 128.

573. Lepore, B., Ruzicka, F., Frey, P., Ringe, D., *The x-ray crystal structure of lysine-2,3-aminomutase from Clostridium subterminale.* Proc Natl Acad Sci USA, 2005. **102**(39): p. 13819-13824.

574. Goldman, P., Grove, T., Booker, S., Drennan, C. , *X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry.* P Natl Acad Sci, 2013. **110**(40): p. 15949-15954.

575. Thompson, J., Higgins, D., Gibson, T., *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Res, 1994. **22**(22): p. 4673-4680.

576. Maiorov, V., Crippen, G., *Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins.* J Mol Biol, 1994. **235**(2): p. 625-634.

577. DeLano, W., *The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.* 2002.

578. Bhandari, D., Xu, H., Nicolet, Y., Fontecilla-Camps, J., Begley, T., *Tryptophan lyase (NosL): mechanistic insights from substrate analogues and mutagenesis.* Biochemistry, 2015. **54**(31): p. 4767-4769.

579. McGlynn, S., Boyd, E., Shepard, E., Lange, R., Gerlach, R., Broderick, J., Peters, J., *Identification and characterization of a novel member of the radical AdoMet enzyme superfamily and implications for the biosynthesis of the Hmd hydrogenase active site cofactor.* J Bacteriol, 2010. **192**(2): p. 595-598.

580. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., Liang, J., *CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated resiudes.* Nucleic Acids Res, 2006. **34**: p. W116-W118.

581. Böck, A., King, P., Blokesch, M., Posewitz, M., *Maturation of hydrogenases.* Adv Microb Physiol, 2006. **51**: p. 1-225.

582. Hitchcock, D., *The solubility of tyrosine in acid and in alkali.* The Journal of general physiology, 1924. **6**(6): p. 747-575.

583. Blackstock, J.C., *Guide to biochemistry.* Butterworth-Heinemann, 2014.

584. McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L., Read, R., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(4): p. 658-674.

585. Tickle, I., Laskowski, R., Moss, D., *Rfree and the Rfree ratio. I. Derivation of expected values of cross-validation residuals used in macromolecular least-squares refinement.* Acta Crystallogr D, 1998. **54**(4): p. 547-557.

586. Monaco, S., Gordon, E., Bowler, M., Delageniere, S., Guijarro, M., Spruce, D., Nanao, M. H., *Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF.* J Appl Crystallogr, 2013. **46**(3): p. 804-810.

587. Stoll, S., Schweiger, A., *EasySpin, a comprehensive software package for spectral simulation and analysis in EPR.* Journal of Magnetic Resonance, 2006. **178**(1): p. 42-55.

588. Kneuttinger, A., Heil, K., Kashiwazaki, G., Carell, T., *The radical SAM enzyme spore photoproduct lyase employs a tyrosyl radical for DNA repair.* Chem Commun, 2013. **49**(7): p. 722-724.

589. Willems, H., Jäger, C., *Polymerase chain reaction.* Ullmann's Encyclopedia of Industrial Chemistry, 1984.

590. Dulau, L., Cheyrou, A., Aigle, M., *Directed mutagenesis using PCR.* Nucleic Acids Res, 1989. **17**(7): p. 2873-2873.

591. Lees, N., Hänzelmann, P., Hernandez, H., Subramanian, S., Schindelin, H., Johnson, M., Hoffman, B., *ENDOR spectroscopy shows that guanine N1 binds to [4Fe− 4S] cluster II of the S-adenosylmethionine-dependent enzyme MoaA: mechanistic implications.* J Am Chem Soc, 2009. **131**(26): p. 9184-9185.

592. Corpet, F., *Multiple sequence alignment with hierarchical clustering.* Nucleic Acids Res, 1988. **16**(22): p. 10881-10890.

593. Clamp, M., Cuff, J., Searle, S., Barton, G., *The jalview java alignment editor.* Bioinformatics, 2004. **20**(3): p. 426-427.

594. Ho, S., Hunt, H., Horton, R., Pullen, J., Pease, L., *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene, 1989. **77**(1): p. 51-59.

595. Khoroshilova, N., Popescu, C., Münck, E., Beinert, H., Kiley, P. J., *Iron-sulfur cluster disassembly in the FNR protein of Escherichia coli by O2:[4Fe-4S] to [2Fe-2S] conversion with loss of biological activity.* P Natl Acad Sci, 1997. **94**(12): p. 6087-6092.

596. Hagen, W., *EPR spectroscopy of iron-sulfur proteins.* Adv. Inorg. Chem, 1992. **38**: p. 165-222.

597. Nimmermark, A., Öhrström, L., Reedijk, J., *Metal-ligand bond lengths and strengths: are they correlated? A detailed CSD analysis.* Z Kristallogr-Cryst Mat, 2013. **228**(7): p. 311-317.

598. Burger, E., Andrade, S., Einsle, O., *Active sites without restraints: high-resolution analysis of metal cofactors.* Curr Opin Struc Biol, 2015. **35**: p. 32-40.

599. Gillard, R., *Circular dichroism. A review.* Analyst, 1963. **88**(1052): p. 825-828.

600. Johnson Jr, W., *Analysis of circular dichroism spectra.* Method Enzymol, 1992. **210**: p. 426-447.

601. Fasman, G., *Circular dichroism and the conformational analysis of biomolecules.* Springer Science Business Media, 2013.

602. Berova, N., Nakanishi, K., *Circular dichroism: principles and applications.* John Wiley Sons, 2000.

603. Johnson Jr, W., *Secondary structure of proteins through circular dichroism spectroscopy.* Annu Rev Biophys Bio, 1988. **17**(1): p. 145-166.

604. Kelly, S., Jess, T., Price, N., *How to study proteins by circular dichroism.* Biochim Biophys Acta-Proteins and Proteomics, 2005. **1751**(2): p. 119-139.

605. Sreerama, N., Woody, R., *Computation and analysis of protein circular dichroism spectra.* Method Enzymol, 2004. **383**: p. 318-351.

606. Holzwarth, G., Doty, P., *The ultraviolet circular dichroism of polypeptides1.* J Am Chem Soc, 1965. **87**(2): p. 218-228.

607. Greenfield, N., Fasman, G., *Computed circular dichroism spectra for the evaluation of protein conformation.* Biochemistry, 1969. **8**(10): p. 4108-4116.

608. Venyaminov, S., Baikalov, I., Shen, Z., Wu, C., Yang, J., *Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set.* Anal Biochem, 1993. **214**(1): p. 17-24.

609. Myers, J., Pace, C., Scholtz, J., *Helix propensities are identical in proteins and peptides.* Biochemistry, 1997. **36**(36): p. 10923-10929.

610. Whitmore, L., Wallace, B., *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data.* Nucleic Acids Res, 2004. **32**(2): p. W668-W673.

611. Andrade, M., Chacon, P., Merelo, J., Moran, F., *Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network.* Protein Engineering, 1993. **6**(4): p. 383-390.

612. Brockwell, D., Paci, E., Zinober, R., Beddard, G., Olmsted, P., Smith, D., Radford, S., *Pulling geometry defines the mechanical resistance of a β-sheet protein.* Nature Struct Mol Biol, 2003. **10**(9): p. 731-737.

613. Von Bergen, M., Barghorn, S., Biernat, J., Mandelkow, E., Mandelkow, E., *Tau aggregation is driven by a transition from random coil to beta sheet structure.* Biochim Biophys Acta-Mol Basis Dis, 2005. **1739**(2): p. 158-166.

614. Ma, B., Nussinov, R., *Simulations as analytical tools to understand protein aggregation and predict amyloid conformation.* Curr Opin Chem Biol, 2006. **10**(5): p. 445-452.

615. Colletier, J., Laganowsky, A., Landau, M., Zhao, M., Soriaga, A., Goldschmidt, L., Eisenberg, D., *Molecular basis for amyloid-β polymorphism.* P Natl Acad Sci, 2011. **108**(41): p. 16938-16943.

616. Tomaiuolo, M., Bertram, R., Houle, D., *Enzyme isoforms may increase phenotypic robustness.* Evolution, 2008. **62**(11): p. 2884-2893.

617. Oliveira, F., Leite, B., Andrade, M., Korn, M., *Determination of total aldehydes in fuel ethanol by MBTH method: sequential injection analysis.* J Brazil Chem Soc, 2005. **16**(1): p. 87-92.

618. Dixon, D., Dobbs, K., Valentini, J., *Amide-water and amide-amide hydrogen bond strengths.* The Journal of Physical Chemistry, 1994. **98**(51): p. 13435-13439.

619. Frey, P., Ballinger, M., Reed, G., *S-adenosylmethionine: a'poor man's coenzyme B12'in the reaction of lysine 2, 3-aminomutase.* Biochem Soc T, 1998. **26**(3): p. 304-310

620. Sanderson, R., *Chemical Bonds and Bonds Energy.* Elsevier, 2012.

621. Winter, G., *xia2: an expert system for macromolecular crystallography data reduction.* J Appl Crystallogr, 2009. **43**(1): p. 186-190.

622. Fisher, S., Levik, K., Williams, M., Ashton, A., McAuley, K., *SynchWeb: a modern interface for ISPyB.* J Appl Crystallogr, 2015. **48**(3): p. 927-932.

623. Suess, D., Bürstel, I., De La Paz, L., Kuchenreuther, J., Pham, C., Cramer, S., Britt, R., *Cysteine as a ligand platform in the biosynthesis of the FeFe hydrogenase H cluster.* P Natl Acad Sci, 2015. **112**(37): p. 11455-11460.

624. Kuchenreuther, J., Stapleton, J., Swartz, J., *Tyrosine, cysteine, and S-adenosyl methionine stimulate in vitro [FeFe] hydrogenase activation.* PLoS One, 2009. **4**(10): p. e7565.

625. Sharpe, A., *Chemistry of cyano complexes of the transition metals.* Academic Press, 1976.

626. Buser, H., Schwarzenbach, D., Petter, W., Ludi, A., *The crystal structure of Prussian blue: Fe4 [Fe (CN) 6] 3. xH2O.* Inorg Chem, 1977. **16**(11): p. 2704-2710.

627. Caro, A., Cederbaum, A., *Antioxidant properties of S-adenosyl-L-methionine in Fe(2+)-initiated oxidations.* Free Radic Biol Med, 2004. **36**(10): p. 1303-1316.

628. Krumholz, P., *Studies on the coordinate bond. VII. Ferrous complexes of α-iminocarboxylic acids.* Inorganica Chimica Acta, 1967. **1**: p. 27-33.

629. Jayanthi, S., Morris, J., Kachel, B., Al-Ameer, M., Henderson, R., *The Versatility of Isothermal Titration Calorimetry in Modern Biology.* J Anal Bioanal Tech, 2015. **6**: p. e121.

630. Ollagnier-de Choudens, S., Sanakis, Y., Hewitson, K., Roach, P., Baldwin, J., Münck, E., Fontecave, M., *Iron-sulfur center of biotin synthase and lipoate synthase.* Biochemistry, 2000. **39**(14): p. 4165-4173.

631. Ollagnier-de-Choudens, S., Mulliez, E., Fontecave, M., *The PLP-dependent biotin synthase from Escherichia coli: mechanistic studies.* FEBS Lett, 2002. **532**(3): p. 465-468.

632. Lanz, N., Pandelia, M., McLaughlin, M., Rectenwald, J., Silakov, A., Drennan, C., Booker, S., *Evidence for the sacrificial role of the auxiliary [4Fe-4S] cluster of lipoyl synthase.* FASEB J, 2015. **29**(1 Supplement): p. 572-574.

633. McIver, L., Baxter, R. Campopiano, D., *Identification of the [Fe-S] cluster-binding residues of Escherichia coli biotin synthase.* J Biol Chem, 2000. **275**(18): p. 13888-13894.

634. Hiscox, M., *Investigations into lipoic acid biosynthesis.* Doctoral dissertation, University of Southampton, 2014.

635. Collins, D. Zhou, H., *Iron–sulfur models of protein active sites.* Encyclopedia Inorg BioInorg Chem, 2011.

636. Garman, E., Owen, R., *Cryocooling and radiation damage in macromolecular crystallography.* Acta Crystallogr D, 2006. **62**(1): p. 32-47.

637. Kuwamoto, S., Akiyama, S., Fujisawa, T., *Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants.* J Synchrotron Radiat, 2004. **11**(6): p. 462-468.

638. Garman, E., *'Cool'crystals: macromolecular cryocrystallography and radiation damage.* Curr Opin Struc Biol, 2003. **13**(5): p. 545-551.

639. Wecksler, S., Stoll, S., Tran, H., Magnusson, O., Wu, S., King, D., Klinman, J., *Pyrroloquinoline quinone biogenesis: demonstration that PqqE from Klebsiella pneumoniae is a radical S-adenosyl-L-methionine enzyme.* Biochemistry, 2009. **48**(42): p. 10151-10161.

640. Feng, J., Wu, J., Dai, N., Lin, S., Xu, H. H., Deng, Z., He, X., *Discovery and characterization of BlsE, a radical S-adenosyl-L-methionine decarboxylase involved in the blasticidin S biosynthetic pathway.* PloS one, 2013. **8**: p. e68545.

641. Yokoyama, K., Numakura, M., Kudo, F., Ohmori, D., Eguchi, T., *Characterization and mechanistic study of a radical SAM dehydrogenase in the biosynthesis of butirosin.* J Am Chem Soc, 2007. **129**(49): p. 15147-15155.

642. Bryant, P., Kriek, M., Wood, R., Roach, P., *The activity of a thermostable lipoyl synthase from Sulfolobus solfataricus with a synthetic octanoyl substrate.* Anal Biochem, 2006. **351**(1): p. 44-49.

643. Getz, E., Xiao, M., Chakrabarty, T., Cooke, R., Selvin, P., *A comparison between the sulfhydryl reductants tris (2-carboxyethyl) phosphine and dithiothreitol for use in protein biochemistry.* Anal Biochem, 1999. **273**(1): p. 73-80.

644. Burns, J., Butler, J., Moran, J., Whitesides, G., *Selective reduction of disulfides by tris (2-carboxyethyl) phosphine.* J Org Chem, 1991. **56**(8): p. 2648-2650.

645. Lees, W., Whitesides, G., *Equilibrium constants for thiol-disulfide interchange reactions: a coherent, corrected set.* J Org Chem, 1993. **58**(3): p. 642-647.

646. Cleland, W., *Dithiothreitol, a new protective reagent for SH groups.* Biochemistry, 1964. **3**(4): p. 480-482.

647. Krężel, A., Leśniak, W., Jeżowska-Bojczuk, M., Młynarz, P., Brasuñ, J., Kozłowski, H., Bal, W., *Coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant.* Journal of inorganic biochemistry, 2001. **84**(1): p. 77-88.

648. Lees, W., Singh, R., Whitesides, G., *meso-2, 5-Dimercapto-N, N, N', N'-tetramethyladipamide: a readily available, kinetically rapid reagent for the reduction of disulfides in aqueous solution.* J Org Chem, 1991. **56**(26): p. 7328-7331.

649. Mejean, A., Bui, B., Florentin, D., Ploux, O., Izumi, Y., Marquet, A., *Highly purified biotin synthase can transform dethiobiotin into biotin in the absence of any other protein, in the presence of photoreduced deazaflavin.* Biochem Bioph Res Co, 1995. **217**(3): p. 1231-1237.

650. Tse Sum Bui, B., Escalettes, F., Chottard, G., Florentin, D., Marquet, A., *Enzyme-mediated sulfide production for the reconstitution of [2Fe− 2S] clusters into apo-biotin synthase of Escherichia coli.* Eur J Biochem, 2000. **267**(9): p. 2688-2694.

651. Ollagnier-de-Choudens, S., Mulliez, E., Hewitson, K., Fontecave, M., *Biotin synthase is a pyridoxal phosphate-dependent cysteine desulfurase.* Biochemistry, 2002. **41**(29): p. 9145-9152.

652. Jarrett, J., *The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase.* Arch Biochem Biophys, 2005. **433**(1): p. 312-321.

653. Lin, S., Cronan, J., *Closing in on complete pathways of biotin biosynthesis.* Mol Biosyst, 2011. **7**(6): p. 1811-1821.

654. Hendrickson, W., Horton, J., LeMaster, D., *Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure.* EMBO J, 1990. **9**(5): p. 1665-1672.

655. Ealick, S., *Advances in multiple wavelength anomalous diffraction crystallography.* Curr Opin Chem Biol, 2000. **4**(5): p. 495-499.

656. Minor, W., Tomchick, D., Otwinowski, Z., *Strategies for macromolecular synchrotron crystallography.* Structure, 2000. **8**(5): p. R105-R110.

657. Leslie, A., Powell, H., *Processing diffraction data with MOSFLM.* , in *Evolving methods for macromolecular crystallography*. 2007, Springer Netherlands. p. 41-51.

658. Choi-Rhee, E., Cronan, J., *A nucleosidase required for in vivo function of the S-adenosyl-L-methionine radical enzyme, biotin synthase.* Chem Biol, 2005. **12**(5): p. 589-593.

659. Farrar, C., Siu, K., Howell, P., Jarrett, J., *Biotin synthase exhibits burst kinetics and multiple turnovers in the absence of inhibition by products and product-related biomolecules.* Biochemistry, 2010. **49**(46): p. 9985-9996.

660. Palmer, L., Downs, D., *The thiamine biosynthetic enzyme ThiC catalyzes multiple turnovers and is inhibited by S-adenosylmethionine (AdoMet) metabolites.* J Biol Chem, 2013. **288**(42): p. 30693-30699.

661. Challand, M.R., et al., *Product inhibition in the radical S-adenosylmethionine family.* FEBS Lett, 2009. **583**(8): p. 1358-62.

662. Skarzynski, T., Thorpe, J., *Industrial perspective on X-ray data collection and analysis.* Acta Crystallogr D, 2006. **62**(1): p. 102-107.

663. Carvalho, A., Trincão, J., Romão, M., *X-ray crystallography in drug discovery.* Ligand-Macromolecular Interactions in Drug Discovery. Humana Press, 2010.

664. Benvenuti, M., Mangani, S., *Crystallization of soluble proteins in vapor diffusion for x-ray crystallography.* Nature Protocols, 2007. **2**(7): p. 1633-1651.

665. Afonine, P., Grosse-Kunstleve, R., Echols, N., Headd, J., Moriarty, N., Mustyakimov, M., Adams, P., *Towards automated crystallographic structure refinement with phenix.refine.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 4): p. 352-367.

666. McLaughlin, M., Lanz, N., Goldman, P., Lee, K. H., Booker, S., Drennan, C., *Caught in the Act: Snapshots of Sulfur Insertion by Lipoyl Synthase.* FASEB J, 2015. **29**(1 Supplement): p. 895-4.

667. Berman, H., Henrick, K., Nakamura, H., *Announcing the worldwide protein data bank.* Nature Struct Mol Biol, 2003. **10**(12): p. 980-980.

668. Kowalak, J., Bruenger, E., McCloskey, J., *Posttranscriptional modification of the central loop of domain V in Escherichia coli 23 S ribosomal RNA.* J Biol Chem, 1995. **270**(30): p. 17758-17764.

669. http://www.rna.icmb.utexas.edu/.

670. Ke, A., Doudna, J., *Crystallization of RNA and RNA–protein complexes.* Methods, 2004. **34**(3): p. 408-414.

671. Srisawat, C., Goldstein, I., Engelke, D., *Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures.* Nucleic Acids Res, 2001. **29**(2): p. e4-e4.

672. Srisawat, C., Engelke, D., *Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins.* RNA, 2001. **7**(4): p. 632-641.

673. Ponchon, L., Dardel, F., *Recombinant RNA technology: the tRNA scaffold.* Nature methods, 2007. **4**(7): p. 571-576.

674. Ponchon, L., Dardel, F., *Large scale expression and purification of recombinant RNA in Escherichia coli.* Methods, 2011. **54**(2): p. 267-273.

675. Ferré-D'Amaré, A., Scott, W., *Small self-cleaving ribozymes.* Cold Spring Harb Perspect Biol, 2010. **2**: p. 10.

676. Hartmann, R., Bindereif, A., Schön, A., Westhof, E., *Handbook of RNA biochemistry.* Weinheim: Wiley-VCH, 2005.

677. Srisawat, C., Engelke, D., *RNA affinity tags for purification of RNAs and ribonucleoprotein complexes.* Methods, 2002. **26**(2): p. 156-161.

678. Lopez, P., Marchand, I., Joyce, S., Dreyfus, M., *The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo.* Mol Microbiol, 1999. **33**(1): p. 188-199.

679. Farrell, R.E., *RNA methodologies.* Wiley-VCH Verlag GmbH Co. KGaA, 1998.

680. Chomczynski, P., Sacchi, N., *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Anal Biochem, 1987. **162**(1): p. 156-159.

681. Chomczynski, P., Sacchi, N., *The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on.* Nature Protocols, 2006. **1**(2): p. 581-585.

682. Ponchon, L., Beauvais, G., Nonin-Lecomte, S., Dardel, F., *A generic protocol for the expression and purification of recombinant RNA in Escherichia coli using a tRNA scaffold. .* Nature Protocols, 2009. **4**(6): p. 947-959.

683. Aranda, P., LaJoie, D., Jorcyk, C., *Bleach gel: A simple agarose gel for analysing RNA quality.* Electrophoresis, 2011. **33**(2): p. 366-369.

684. Boedtker, H., *Conformation independent molecular weight determinations of RNA by gel electrophoresis.* Biochim Biophys Acta, Nucleic Acids and Protein Synthesis, 1971. **240**(3): p. 448-453.

685. Muto, A., *Nucleotide distribution of Escherichia coli 16S ribosomal ribonucleic acid.* Biochemistry, 1970. **9**(19): p. 3683-3694.

686. Moore, P., *The ribosome at atomic resolution.* Biochemistry, 2001. **40**(11): p. 3243-3250.

687. Azarani, A., Hecker, K., *RNA analysis by ion-pair reverse-phase high performance liquid chromatography.* Nucleic Acids Res, 2001. **29**(2): p. 1-9.

688. Dickman, M.J., *Effects of sequence and structure in the separation of nucleic acids using ion pair reverse phase high performance liquid chromatography.* J Chromatography A, 2005. **1076**(1): p. 83-89.

689. Noll, B., Seiffert, S., Vornlocher, H., Roehl, I., *Characterization of small interfering RNA by non-denaturing ion-pair reversed-phase liquid chromatography.* J Chromatography A, 2011. **1218**(33): p. 5609-5617.

690. Beverly, M., Hartsough, K., Machemer, L., Pavco, P., Lockridge, J., *Liquid chromatography electrospray ionization mass spectrometry analysis of the ocular metabolites from a short interfering RNA duplex.* Journal of Chromatography B, 2006. **835**(1): p. 62-70.

691. Curtis, M., Minier, M., Chitranshi, P., Sparkman, O., Jones, P., Xue, L., *Direct analysis in real time (DART) mass spectrometry of nucleotides and nucleosides: elucidation of a novel fragment [C 5 H 5 O]+ and its in-source adducts.* J Am Soc Mass Spectr, 2010. **21**(8): p. 1371-1381.

692. Welch, M., Villalobos, A., Gustafsson, C., Minshull, J. Methods of Enzymology, 2011.

693. Studier, F., *Protein production by autoinduction in high-density shaking cultures.* Protein Express Purif, 2005. **41**(1): p. 207-234.

694. Drenth, J., *X-Ray Crystallography.* John Wiley Sons, Inc., 2007.

695. Jones, S., Daley, D., Luscombe, N., Berman, H., Thornton, J., *Protein–RNA interactions: a structural analysis.* Nucleic Acids Res, 2001. **29**(4): p. 943-954.

696. Ellis, J., Broom, M., Jones, S., *Protein–RNA interactions: structural analysis and functional classes.* Proteins: Structure, Function, and Bioinformatics, 2007. **66**(4): p. 903-911.

697. Bahadur, R., Zacharias, M., Janin, J., Nucleic Acids Res, 2008.

698. McCusker, K., Medzihradszky, K., Shiver, A., Nichols, R., Yan, F., Maltby, D., Galonić Fujimori, D., *Covalent intermediate in the catalytic mechanism of the radical S-adenosyl-L-methionine methyl synthase RlmN trapped by mutagenesis.* J Am Chem Soc, 2012. **134**(43): p. 18074-18081.

699. Losick, R., *In vitro transcription.* Annu Rev Biochem, 1972. **41**(1): p. 409-446.

700. Pyle, A., *Ribozymes: a distinct class of metalloenzymes.* Science, 1993. **261**(5122): p. 709-714.

701. Greenberg, J., *Ultraviolet light-induced crosslinking of mRNA to proteins.* Nucleic Acids Res, 1979. **6**(2): p. 715-732.

702. Ule, J., Jensen, K., Mele, A., Darnell, R., *CLIP: a method for identifying protein–RNA interaction sites in living cells.* Methods, 2005. **37**(4): p. 376-386.

703. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Tuschl, T., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.* Cell, 2010. **141**(1): p. 129-141.

704. Shepard, E., Mus, F., Betz, J., Byer, A., Duffus, B., Peters, J. , Broderick, J., *[FeFe]-hydrogenase maturation.* Biochemistry, 2014. **53**(25): p. 4090-4104.

705. Pickett, C., *Making the H-cluster from scratch.* Science, 2014. **343**(6169): p. 378-379.

706. Cendron, L., Berto, P., D'Adamo, S., Vallese, F., Govoni, C., Posewitz, M. C., Zanotti, G. , *Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation.* J Biol Chem, 2011. **286**(51): p. 43944-43950.

707. Vallese, F., Berto, P., Ruzzene, M., Cendron, L., Sarno, S., De Rosa, E., Costantini, P., *Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process.* J Biol Chem, 2012. **287**(43): p. 36544-36555.

708. Beinert, H., Holm, R. H., Münck, E., *Iron-sulfur clusters: nature's modular, multipurpose structures.* Science, 1997. **277**(5326): p. 653-659.

709. Challand, M.R., Martins, F. T., Roach, P. L., *Catalytic activity of the anaerobic tyrosine lyase required for thiamine biosynthesis in Escherichia coli.* J Biol Chem, 2010. **285**(8): p. 5240-5248.

710. Driesener, R., *Mechanistic and structural characterisation of HydG catalysed L-tyrosine cleavage* Doctoral dissertation, University of Southampton, 2014.

711. Krumholz, P., *Studies on the coordinate bond. VII. Ferrous complexes of α-iminocarboxylic acids.* Inorganica Chimica Acta, 1967. **1**(27-33).

712. Lanz, N., Rectenwald, J., Wang, B., Kakar, E., Laremore, T., Booker, S., Silakov, A. , *Characterization of a radical intermediate in lipoyl cofactor biosynthesis.* J Am Chem Soc, 2015. **137**(41): p. 13216–13219.

713. Lanz, N., Pandelia, M., Kakar, E., Lee, K., Krebs, C., Booker, S., *Evidence for a catalytically and kinetically competent enzyme–substrate cross-linked intermediate in catalysis by lipoyl synthase.* Biochemistry, 2014. **53**(28): p. 4557-4572.

714. Kalyaanamoorthy, S., Chen, Y. P. P., *Structure-based drug design to augment hit discovery.* Drug Discov Today, 2011. **16**(17): p. 831-839.

715. Yan, F., LaMarre, J., Röhrich, R., Wiesner, J., Jomaa, H., Mankin, A.,Fujimori, D. , *RlmN and Cfr are radical SAM enzymes involved in methylation of ribosomal RNA.* J Am Chem Soc, 2010. **132**(11): p. 3953-3964.

716. Tabor, S., Richardson, C. , *A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes.* P Natl Acad Sci, 1985. **82**(4): p. 1074-1078.

717. Studier, F., Moffatt, B., *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.* J Mol Biol, 1986. **189**(1): p. 113-130.

718. Kam, Z., Shore, H., Feher, G., *On the crystallization of proteins.* J Mol Biol, 1978. **123**(4): p. 539-555.

719. Veesler, S., Boistelle, R., Crystallization of Nucleic Acids and Proteins: a Practical Approach, 1999.

720. McPherson, A., *Preparation and analysis of protein crystals.* Wiley, New York, 1982.

721. Haas, C., Drenth, J., *Understanding protein crystallization on the basis of the phase diagram.* J Crystal Growth, 1999. **196**(2): p. 388-394.

722. Blow, D., Chayen, N., Lloyd, L., Saridakis, E., *Control of nucleation of protein crystals.* Protein Sci, 1994. **3**(10): p. 1638-1643.

723. Durbin, S., Carlson, W., Saros, M., *In situ studies of protein crystal growth by atomic force microscopy.* Journal of Physics D: Applied Physics, 1993. **26**(8B): p. B128-B132.

724. Asherie, N., *Protein crystallization and phase diagrams.* Methods, 2004. **34**(3): p. 266-272.

725. Patro, S., Prybycien, T., *Simulations of kinetically irreversible protein aggregate structure.* Biphys J, 1994. **66**(5): p. 1274-1289.

726. Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y., McPherson, A., *Screening and optimization strategies for macromolecular crystal growth.* Acta Crystallogr D, 1994. **50**(4): p. 414-423.

727. Durbin, S., Feher, G., *Protein crystallization.* Annu Rev Phys Chem, 1996. **47**(1): p. 171-204.

728. McPherson, A., *Introduction to protein crystallization.* Methods, 2004. **34**(3): p. 254-265.

729. Drenth, J., *X-Ray Crystallography.* John Wiley Sons, Inc.. 2007.

730. Arakawa, T., Timasheff, S., *Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding.* Biochemistry, 1984. **23**(25): p. 5912-5923.

731. Pitzer, K., *Electrolyte theory-improvements since Debye and Hückel.* Accounts Chem Res, 1977. **10**(10): p. 371-377.

732. Wiencek, J., *New strategies for protein crystal growth. Annual review of biomedical engineering.* 1, 1999. **1**(505-534).

733. Ilett, S.O., A., Poon, W., Pusey, P., *Phase behavior of a model colloid-polymer mixture.* Physical Review E, 1995. **51**(2): p. 1344.

734. Berejnov, V., Husseini, N. S., Alsaied, O. A., Thorne, R. E., *Effects of cryoprotectant concentration and cooling rate on vitrification of aqueous solutions.* J Appl Crystallogr, 2006. **39**(2): p. 244-251.

735. Anand, K., Pal, D., Hilgenfeld, R., *An overview on 2-methyl-2, 4-pentanediol in crystallization and in crystals of biological macromolecules.* Acta Crystallogr D, 2002. **58**(10): p. 1722-1728.

736. Bergfors, T., *Protein crystallization.* Internat'l University Line, 2009.

737. Kantardjieff, K., Rupp, B., *Protein isoelectric point as a predictor for increased crystallization screening efficiency.* Bioinformatics, 2004. **20**(14): p. 2162-2168.

738. Nave, C., *Radiation damage in protein crystallography.* Radiat Phys Chem, 1995. **45**(3): p. 483-490.

739. Garman, E., *Radiation damage in macromolecular crystallography: what is it and why should we care?* Acta Crystallogr D, 2010. **66**(4): p. 339-351.

740. Hope, H., *Cryocrystallography of biological macromolecules: a generally applicable method.* Acta Crystallogr B, 1988. **44**(1): p. 22-26.

741. Garman, E., Schneider, T., *Macromolecular cryocrystallography.* J Appl Crystallogr, 1997. **30**(3): p. 211-237.

742. Matthews, B., *Solvent content of protein crystals.* J Mol Biol, 1968. **33**(2): p. 491-497.

743. Warkentin, M., Berejnov, V., Husseini, N., Thorne, R., *Hyperquenching for protein cryocrystallography.* J Appl Crystallogr, 2006. **39**(6): p. 805-811.

744. Rodgers, D., *Practical cryocrystallography.* Method Enzymol, 1997. **276**: p. 183-203.

745. Bollag, D.E., S., Rozycki, M., *Protein methods.* New York: Wiley-Liss, 1996. **2nd ed**: p. 231-269.

746. Thaller, C., Weaver, L., Eichele, G., Wilson, E., Karlsson, R., Jansonius, J., *Repeated seeding technique for growing large single crystals of proteins.* J Mol Biol, 1981. **147**(3): p. 465-469.

747. Stura, E., Wilson, I., *Analytical and production seeding techniques.* Methods, 1990. **1**(1): p. 38-49.

748. Stura, E., Wilson, A., *Crystallization of nucleic acids and proteins: a practical approach, by A. Ducruix R. Giegé.* New York: Oxford University Press, 1992.

749. Stura, E., Wilson, I., *Applications of the streak seeding technique in protein crystallization.* J Crystal Growth, 1991. **110**(1): p. 270-282.

750. Riès-Kautt, M., Ducruix, A., Giegé, R., *Crystallization of nucleic acids and proteins: a practical approach.* Oxford University Press, 1992.

751. Luft, J., DeTitta, G., *A method to produce microseed stock for use in the crystallization of biological macromolecules.* Acta Crystallogr D, 1999. **55**(5): p. 988-993.

752. Heras, B., Martin, J., *Post-crystallization treatments for improving diffraction quality of protein crystals.* Acta Crystallogr D, 2005. **61**(9): p. 1173-1180.

753. Hauptman, H., *Phasing methods for protein crystallography.* Curr Opin Struc Biol, 1997. **7**(5): p. 672-680.

754. Evans, P., McCoy, A., *An introduction to molecular replacement.* Acta Crystallogr D, 2007. **64**(1): p. 1-10.

755. Scapin, G., *Molecular replacement then and now.* Acta Crystallogr D, 2013. **69**(11): p. 2266-2275.

756. Blundell, T., Johnson, L., *Protein crystallography.* Academic Press Inc., London, 1976.

757. Baker, P., Farrants, G., Stillman, T., Britton, K., Helliwell, J., Rice, D., *Isomorphous replacement with optimized anomalous scattering applied to protein crystallography.* Acta Crystallogr A, 1990. **46**(9): p. 721-725.

758. Hendrickson, W., *Determination of macromolecular structures from anomalous diffraction of synchrotron radiation.* Science, 1991. **254**(5028): p. 51-58.

759. Warren, B., *X-ray Diffraction.* Addison-Wesley Publishing Co, 1969.

760. Matthews, B., *Anomalous dispersion. International Tables for Crystallography* 2001, 1970. **Volume F**: p. 293-298.

761. Burla, M., Carrozzini, B., Cascarano, G., Giacovazzo, C., Polidori, G., *SAD or MAD phasing: location of the anomalous scatterers.* Acta Crystallogr D, 2003. **59**(4): p. 662-669.

762. Schmidt, A., Gonzalez, A., Morris, R., Costabel, M., Alzari, P., Lamzin, V., *Advantages of high-resolution phasing: MAD to atomic resolution.* Acta Crystallogr D, 2002. **58**(9): p. 1433-1441.

763. Hendrickson, W., Ogata, C., *Phase determination from multiwavelength anomalous diffraction measurements.* Method Enzymol, 1997. **276**: p. 494-523.

764. Leslie, A., Powell, H., *Methods in Macromolecular Crystallography.* IOS Press, Amsterdam, 2001(65-71).

765. *Collaborative Computation Project, Number 4.* Acta Crystallographica, 1994. **50**: p. 760-763.

766. Hauptman, H., *The phase problem of x-ray crystallography.* Rep Prog Phys, 1991. **54**(11): p. 1427.

767. Weiss, M., Hilgenfeld, R., *On the use of the merging R factor as a quality indicator for X-ray data.* J Appl Crystallogr, 1997. **30**(2): p. 203-205.

768. Hendrickson, W., *Stereochemically restrained refinement of macromolecular structures.* Method Enzymol, 1984. **115**: p. 252-270.

769. Laskowski, R., MacArthur, M., Moss, D., Thornton, J., *PROCHECK: a program to check the stereochemical quality of protein structures.* J Appl Crystallogr, 1993. **26**(2): p. 283-291.

770. Brünger, A., *Free R value: a novel statistical quantity for assessing the accuracy of crystal structures.* Nature, 1992. **355**: p. 472-475.

771. Brunger, A., *Assessment of phase accuracy by cross validation: the free R value. Methods and applications.* Acta Crystallogr D, 1993. **49**(1): p. 24-36.