UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS

School of Chemistry

Applications of Intein Mediated Ligation

by

Robert James Wood

Thesis for the degree of Doctor of Philosophy

September 2005

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF CHEMISTRY

Doctor of Philosophy

APPLICATIONS OF INTEIN MEDIATED LIGATION

by Robert James Wood

Intein mediated ligation provides a site-specific method for the attachment of a wide range of molecular probes to proteins. The method has some inherent practical difficulties which have limited its application. The use of this valuable technique for the site specific fluorescent labelling and controlled electrode immobilisation of proteins was investigated, with intein mediated ligation chosen as a method due to its site specificity. Plasmids for the expression of target-intein fusion proteins were assembled and fusion protein expression optimised in a range of six *E. coli* strains, four target-intein fusion proteins were purified by Ni affinity chromatography. The reaction of a model target-intein fusion protein with a fluorescent cysteine derivative was investigated under a range of conditions including anaerobicity of the reaction and the addition of low molecular weight thiols. It was found that the ligation of the fluorophore to the target protein is critically dependant upon the degree of oxidation of the fluorescent cysteine label. Unusually it was shown that efficient direct reaction of the fluorescent label with the target-intein thioester was possible, but only under rigorously anaerobic conditions. Optimal, more reliable conditions for efficient reaction include anaerobic reaction with the addition of thiophenol (10 mM) or aerobic reaction with the further addition of the water soluble reducing agent, tricarboxyethylphosphine.

The glassy carbon electrode surface was oxidatively modified with diaminodecane, which was coupled with a protected cysteine derivative. Each electrode derivatisation step was characterised electrochemically and by XPS. Cysteine deprotection was followed by reaction with BioB-intein, although protein was present at the electrode surface (as observed by XPS), it was unclear whether the protein had been attached via intein mediated ligation or had non-specifically adsorbed.

Contents

List of Figures

List of Tables

List of Equations

DECLARATION OF AUTHORSHIP

I, Robert James Wood, declare that the thesis entitled 'Applications of Intein Mediated Ligation' and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- 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;
- where I have consulted the published work of others, this is always clearly attributed;
- 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;
- I have acknowledged all main sources of help;
- 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;
- parts of this work have been published as:

Wood, R.J., Pascoe, D. D., Brown, Z. K., Medlicott, E. M., Kriek, M., Neylon, C. and Roach, P.L., *Bioconjugate Chemistry*, (2004), **15**, 366-372.

Signed: [redacted]

Date: March 2006

Acknowledgements

I thank those here mentioned for their contributions made to this work; towards the synthesis of FTEC by project students David Pascoe, Zöe Brown and Emma Medlicott, Mike Wang for his work on the expression and purification of BFP-intein and Veronique Calleja for pVC100. John Langley and Julie Herniman (Mass spectrometry), Joan Street and Neil Wells (NMR) and Graham Beamson (XPS) for their expert help.

I would like to warmly thank Peter Roach and Phil Bartlett for their supervision throughout this work and their great help in the process of writing this thesis. I thank Phil for his steady advice and calm thought and focus through times when it was difficult, working on such a broad project, to remain focused upon the goals which I wanted to achieve. I particularly thank Peter for letting me have a go at such an interesting and exiting project, which allowed me to thoroughly explore every area I was interested in, I do not think that this would have been as possible on any other project! His overwhelming enthusiasm for his work and interest in each aspect were a great encouragement during my time in his lab. I also thank my advisor, Tom Brown, for his encouragement and advice during this study.

I thank Marco, Penny, Paul, Hannah and Rohan for their great help in proof reading my thesis, and Lyn for helping me tie it all together at the end, a job I don't think I could have managed alone!

I would particularly like to thank my parents Jane and Tim, my sister Becky, Martin, Marco, Filipa, my group mates and my girlfriend Lyn for their support and help during my work towards this thesis. Thanks especially to Lyn for her patience during a tough year, hopefully we can rest and have fun now!

Thanks also to Tillmann, Roberta and all other Roach / Bartlett group members past and present!!

Abbreviations

BioB:	*Escherichia coli* Biotin synthase
BFP:	*Aequorea victoria* green fluorescent protein, blue variant
BSA:	Bovine serum albumin
C:	Cysteine
CDI:	Carbonyldiimidazole
CoXAsH:	4',5'-bis(1,3,2-dithioarsolan-2-yl)-2,8-dichloro-3,6-dihydroxy-9H-xanthen-9-one
Cys:	Cysteine
DAD:	Decanediamine
DCM:	Dichloromethane
DIPEA:	Diisopropylethylamine
DMF:	Dimethylformamide
DNA:	deoxyribonucleic acid
DTT:	Dithiothretol
E. coli:	*Escherichia coli*
EDC:	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EPL:	Expressed protein ligation
FAD:	Flavin adenine dinucleotide
FcAA:	Ferroceneacetic acid
FlAsH:	4',5'-bis(1,3,2-dithioarsolan-2-yl)fluorescein
FITC:	Fluorescein isothiocyanate
FPLC:	Fast protein liquid chromatography
Fpr:	*Escherichia coli* Flavodoxin NADP+ reductase
FRET:	Fluorescence resonance energy transfer
FTEC:	N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide
GFP:	*Aequorea victoria* green fluorescent protein
GshA:	*Escherichia coli* Glutamate-cysteine ligase
GshB:	*Escherichia coli* Glutathione synthase
hAGT:	*Homo sapiens* O^6-alkylguanine DNA alkyl transferase
His:	Histidine
His_6-tag:	Six Histidine tag

HPLC:	High pressure liquid chromatography
IML:	Intein mediated ligation
I_{par}:	Parallel fluorescence intensity
I_{per}:	Perpendicular fluorescence intensity
Intein:	*Saccharomyces cerevisiae VMA1* intein
kDa:	Kilo Dalton
L:	Leucine
LipA:	*Sulfolobus solfataricus* Lipoyl synthase
MESA:	Mercaptoethanesulphonic acid
NADH:	Nicotinamide adenine dinucleotide
NMR:	Nuclear magnetic resonance
p:	Plasmid
PBS:	Phosphate buffered saline
PCR:	polymerase chain reaction
Pfu:	*Pyrococcus furiosus*
PGE:	Pyrolytic edge planar graphite
PKBα:	*Homo sapiens* Protein kinase Bα
PTEN:	phosphate and tensin homologue deleted on *Homo sapiens* chromosome ten
PyBop:	Benzotriazol-1-yl-oxytripyrrolidinophosphoniumhexafluorophosphate
ReASH:	4',5'-bis(1,3,2-dithioarsolan-2-yl)-resorufin
RNA:	Ribonucleic acid
S:	Serine
SAM:	S-adenosyl-L-methionine
SCE:	Saturated calomel electrode
SDS:	Sodium dodecyl sulphate
SDS-PAGE:	SDS polyacrylamide gel electrophoresis
SolA:	*Escherichia coli* N-methyltryptophan oxidase
Taq:	*Thermus aquaticus*
TCEP:	Tricarboxyethylphosphine
TEMED:	N,N,N',N'-Tetramethylethylenediamine
TFA:	Trifluoroacetic Acid
Tris:	Tris(hydroxymethylethyl)aminomethane

UV:	Ultraviolet
XPS:	X-ray photoelectron spectroscopy

Notations

In this report, gene names are quoted in italics with no capitalization of the first letter, e.g. *bioB*.

Gene products are named in normal type, with first letter capitalized, e.g. BioB.

Chapter 1:- Introduction

1.1 Applications of fluorescently labelled proteins

Amongst the methods available for the study of protein structure and their interaction with other biomolecules are X-ray diffraction, nuclear magnetic resonance, electron microscopy and scanning probe techniques. These analytical methods require large quantities of the purified biological components (proteins, DNA etc), are often conducted under non physiological conditions and are seldom suitable for the observation of biomolecular interactions or processes in real time or at high throughput (*1*).

The use of fluorescence techniques for the study of biomolecules has allowed many of these problems to be circumvented. Recent developments have been made which permit the selective, real time analysis of molecular interactions both *in vitro* (*2*) and *in vivo* (*3, 4*). It has even been shown to be possible to monitor interactions between two single molecules under physiological conditions (*5, 6*).

A prerequisite for the application of such fluorescence techniques is that the molecule of interest must exhibit strong fluorescence and be distinguishable above the background intrinsic fluorescence of the other biochemical components present. It is rare that the protein or ligand of interest is fluorescent itself, and most proteins are indistinguishable above the background. For this reason it is necessary to label molecules of interest with suitable fluorescent probes. Proteins suitably modified with fluorescent labels have been used in fluorescence polarisation (*7*), fluorescence resonance energy transfer (FRET) (*8*) and fluorescence microscopy studies (*9, 10*). Our interest in methods for protein modification stems from a desire to specifically fluorescently label a range of proteins so that these fluorescence techniques could be used to study protein–protein

and protein–substrate interactions. Some of the most frequently used applications of fluorescent labelling of proteins are discussed below.

1.1.1 Fluorescence polarization

Fluorescence polarization (*7*) is a very useful technique for the detection of binding between two molecules. It is a measurement which provides information about the molecular orientation and the mobility of a fluorophore conjugated to another molecule such as a protein (*11*). In fluorescence polarization, a sample is excited with linearly polarised light and the emitted fluorescence intensity measured at a fixed angle relative to the excitation beam, usually 90°. The emitted light passes through an emission polarizer which is orientated to pass light which is polarised either parallel or perpendicular with respect to the exiting light (by rotation of the polarising filter). This results in two measurements, a parallel fluorescence intensity (I_{par}) and a perpendicular intensity (I_{per}), where the difference between these two values carries the information regarding the mobility of the excited fluorophore. Fluorescence polarization, or P is therefore calculated from equation 1.1.

$$P = (I_{par} - I_{per}) / (I_{par} + I_{per})$$

Equation 1.1 Fluorescence polarization

The orientational distribution of fluorophores attached to small rapidly rotating molecules becomes randomised as the molecules rotate quickly over the time required for emission (for example fluorescein fluorescence lifetime = 4 ns (*12*)) resulting in low polarisation of the emitted photons, and low fluorescence polarization (*13*). Dyes attached to large, slowly rotating molecules will exhibit high fluorescence polarization as they will rotate relatively little prior to emission. This method has been widely applied *in vitro* and examples include the binding of a fluorescein labelled extendin with a glucagon like peptide-1 receptor (*14*) and the interaction of citrate synthase with malate dehydrogenase (*15*). This technique has also been developed in microplate reader format, to allow high throughput screening (*16-18*).

1.1.2 Fluorescence resonance energy transfer (FRET)

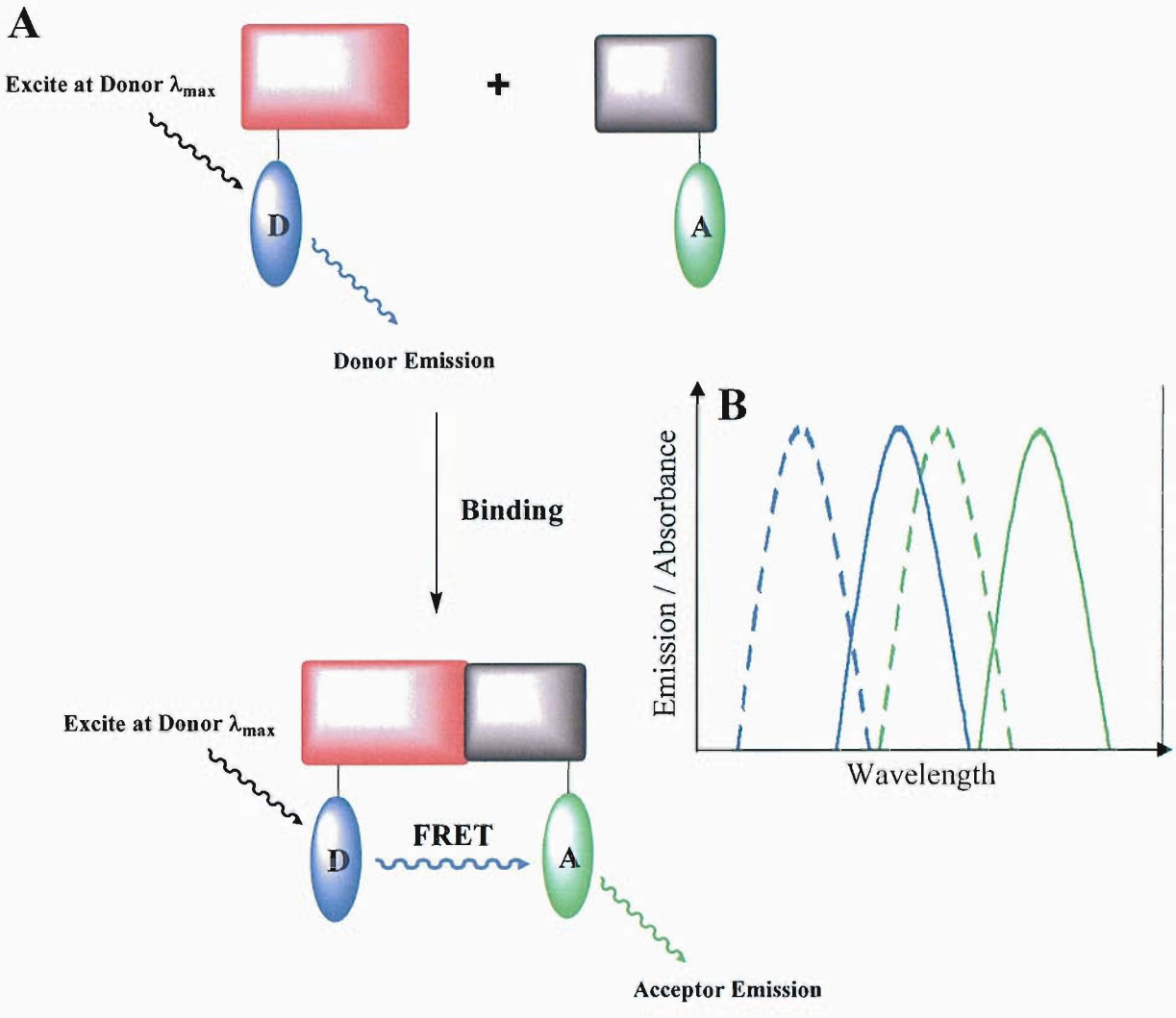


Figure 1.1 Fluorescence resonance energy transfer: Two proteins, coloured red and black (A), are labelled with a donor (blue) and acceptor (green) fluorophore respectively. Excitation of the donor fluorophore in the absence of the other protein leads to emission at the donor wavelength. When the black protein is added and binding occurs, the donor and acceptor fluorophores become close in space. The emission spectrum ((B) solid lines) of the blue fluorophore is chosen such that it overlaps well with the excitation spectrum ((B) dotted lines) of the green fluorophore. In this case, excitation of the donor leads to radiationless energy transfer to the acceptor fluorophore, and emission at the acceptor wavelength.

FRET is used to detect whether two fluorescently labelled molecules are in very close proximity, which typically occurs when they are bound to one another (*19*). It requires

the labelling of the molecules whose interaction is of interest with two different fluorophores, a donor (D) and an acceptor (A) (Fig. 1.1A). The donor and acceptor molecules are selected such that the emission spectrum of the donor overlaps strongly with the excitation spectrum of the acceptor (Fig. 1.1B). If this is the case, and the two molecules are in close proximity (usually separation of less than 10 nm is required) when the donor molecule is excited rather than emitting the energy at its characteristic emission frequency, the energy is transferred non-radiatively to the acceptor molecule resulting in emission at the acceptors characteristic emission wavelength (*20*). The efficiency of FRET is highly dependent upon the distance between the donor and acceptor, and is proportional to the inverse sixth power with respect to the distance between the two fluorophores (*21*). Typically FRET is used to measure the binding of a ligand and receptor, as upon binding the fluorophores are brought close together in space and FRET will occur.

FRET has been a very widely applied technique (*8, 22*), used in many examples in which the interaction is between a protein and another molecule. The interaction of many proteins may be studied *in vivo* using FRET based measurements (*23*), in this case proteins are often expressed as a fusion protein with the *Aequorea victoria* green fluorescent protein and its variants (*24*), potentially allowing real time *in vivo* observation of cellular events using this powerful technique.

1.1.3 Fluorescence microscopy

Fluorescence microscopy (*9, 10*) is a method through which fluorescently labelled biomolecules are visualised *in situ* within the host cells (*25*). This type of live or fixed cell imaging requires that the cells are loaded with a quantity of labelled protein which has a minimal effect upon the normal cellular processes whilst maintaining a favourable signal to noise ratio. Microscopy allows the cellular localisation of proteins to be assessed under different conditions (*26*) and when proteins are labelled with different fluorophores, it permits the interaction between proteins within their natural environment to be investigated by simultaneous microscopy and FRET measurements (*10*). Observation of the cellular location and interactions of proteins within the host cell under various stimuli is an exciting and developing application of fluorescent labelling in the field of systems biology. Various routes to the modification of proteins

with fluorophores have been used, and some of the general methods are discussed in section 1.2.

1.2 Protein modification

Protein modification chemistry (*27*) has played an important role in the study of many proteins, its role is to tailor the physical properties of the protein of interest without altering its activity or mode of action. The properties are altered to suit the application of interest, but usually to facilitate the characterisation of the protein and to widen the range of techniques available for its study. Proteins may be covalently modified using a range of methods (*28*) to suit the required purpose, for example, reagents and procedures have been designed to preserve electrostatic charge (*29*), to increase hydrophobicity (*30*), to alter susceptibility to proteolysis (*31*), to introduce fluorescent labels (*32*), radiolabels (*33*), metal ions (*34*), carbohydrate moieties (*35*) and affinity tags such as biotin (*36*). Proteins are also modified to alter their specificity (*37*), and to cross link proteins thus stabilising complexes for analysis (*38*).

The most commonly used routes for the modification of proteins involve covalent attachment of the label to reactive functionalities available within the protein structure, usually one of primary amines, cysteine sulphydryl groups and carboxylic acids (*39*).

1.2.1 Reaction with primary amines (*40*)

The labelling of primary amines is the most widely used method of protein labelling due to their availability within the protein structure. The modifying reagent reacts with primary amine functionalities which are found at the protein N-terminus and lysine residues contained within the protein structure, these are abundant and easily modified due to their reactivity.

The main factors which affect the amine reactivity are the environment of the amine and its basicity. Aliphatic amines such as the ε-amino group of lysine are of moderate basicity and will react with most acylating agents. The concentration of the free base form of these amines below pH 8.0 is low (pKa lysine ε-amino group = 10.5 (*41*)) and

so the modification reaction is pH dependent, such that it is usual for a more basic pH to be used for reaction with these residues. The α-amine at the protein N-terminus has a pKa of ~ 8 (*41*), and so may be modified closer to neutral pH. Amine reactive labels are often acylating agents, which react to form amides, thioureas and sulfonamides upon reaction.

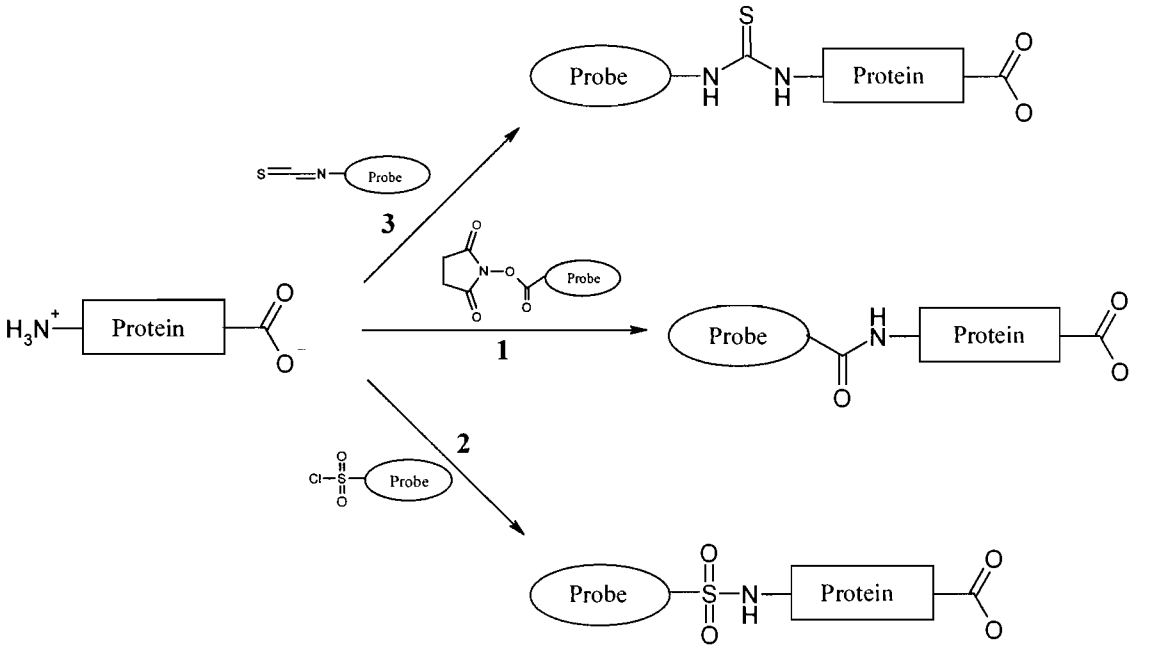


Figure 1.2 Protein labelling via reaction with primary amines

Primary amines may be modified by reaction with an activated carboxylic acid (Fig. 1.2, **1**), examples include the use of N-hydroxysuccinimidyl- (*42*) and tetrafluorophenyl-esters (*43*). Primary amines may be reacted with widely available isothiocyanates (Fig 1.2, **3**) which react to form thioureas, but the thiourea formed may be unstable over time, leading to degradation (*44*). Examples of their application include fluorescent labelling using highly water soluble labels such as fluorescein isothiocyanate and tetramethylrhodamine isothiocyanate forming a labelled protein which can act as a FRET donor / acceptor pair for binding studies (*45*). Sulphonyl chlorides (Fig. 1.2, **2**) are highly reactive with amines, but are also fairly unstable in water at higher pH (*46*). Once conjugated to the protein however, the sulphonamide product is very stable, even able to withstand complete protein hydrolysis (*47*). The reactivity of sulphonyl chlorides however may also result in non-specific reaction with

other functional groups such as phenols (tyrosine), thiols (cysteine) and imidazoles (histidine) (*48*).

The general reactivity of primary amines towards the modification reagents allows the labelling procedure to be very straightforward, but the specificity of the modification is poor.

1.2.2 Modification via reaction with thiols (*49*)

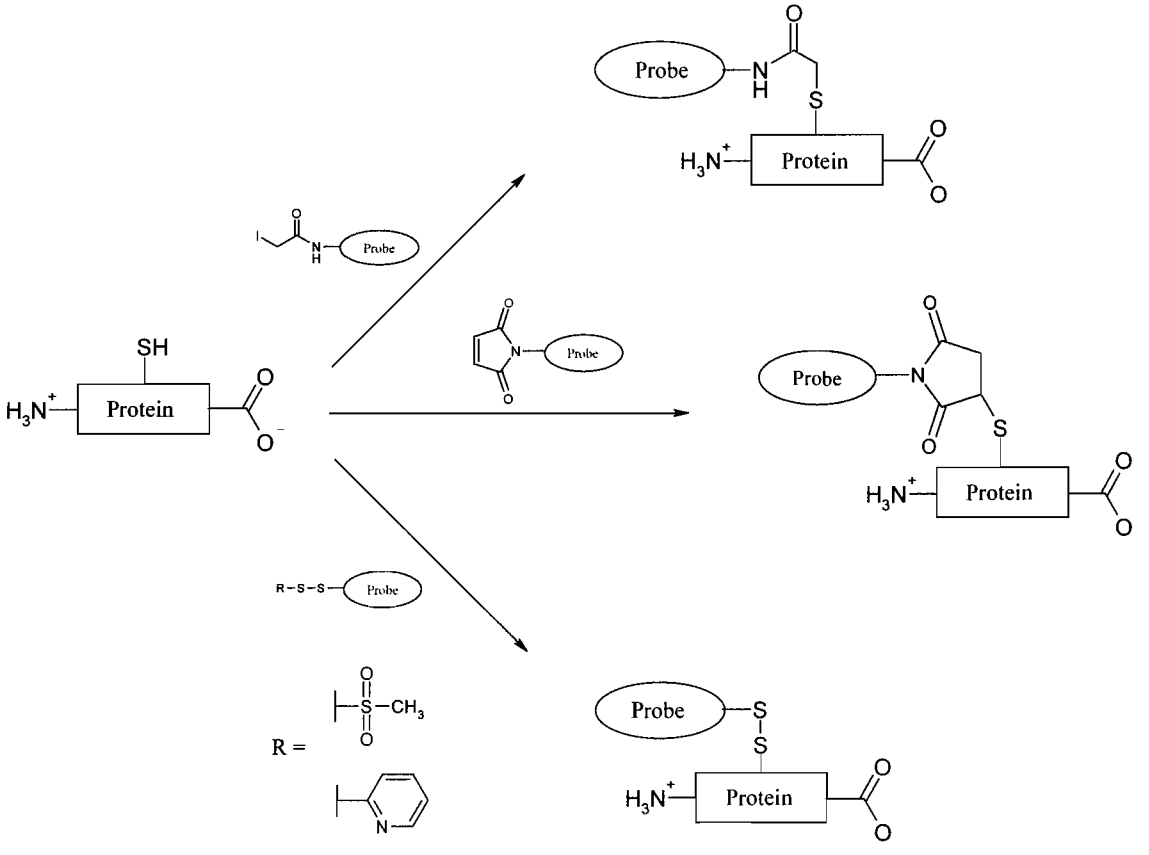


Figure 1.3 Protein labelling via modification of thiols

In contrast to amine modification, thiol modification can be relatively selective. Site specific labelling may be achieved by mutagenesis of the target protein so that a single cysteine residue is left at the required site in the protein. Thiols are usually modified by alkylating reagents, for example iodoacetamides efficiently react with thiols (Fig. 1.3, **1**) forming a thioether product (*50*); however, a potential problem with their application is the side reaction with histidine residues (*51*). Reaction with maleimides (*52*) is more

selective (*53*), and involves a Michael type addition of the thiol across the double bond (Fig. 1.3, **2**) forming a thioether. Thiols may also be derivatised by disulphide exchange reagents (*54*), which are compounds containing a disulphide group which is able to participate in a thiol exchange reaction with a protein cysteine thiol yielding a reversibly labelled protein (Fig. 1.3, **3**). Reagents used for this type of modification include methanethiosulphonate (*55*) and thiopyridyl (*56*) mixed disulphides. A disadvantage of cysteine labelling is the lack of specificity when the proteins contain more than one cysteine residue. A second disadvantage is that often cysteine residues are required for protein activity (*57*), so cysteine labelling may result in a reduction of catalytic or biological activity.

1.2.3 Derivatisation by reaction with carboxylic acids

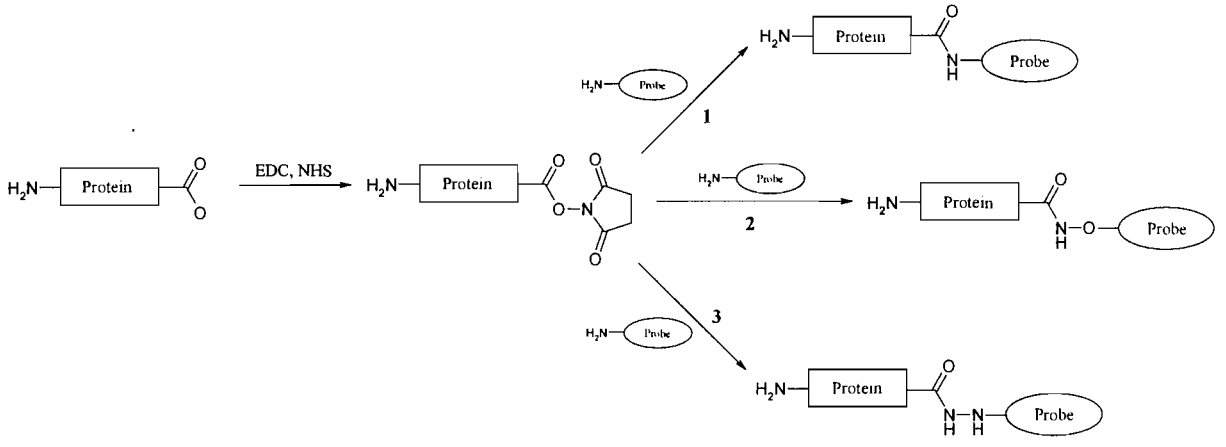


Figure 1.4 Protein labelling via reaction with carboxylic acids

Carboxylic acids, found at the protein C-terminus, and in glutamic acid and aspartic acid residue side chains, may react in aqueous medium to form amides, esters, hydrazides and hydroxamic acids (Fig. 1.4). Reaction of the protein with a water soluble carbodiimide such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) in the presence of N-hydroxysuccinimide (Fig. 1.4, **1**) activates the carboxylic acid towards reaction with amine modified probes, to form an amide bond between the protein and the label (*58*). Inter-protein reaction between lysine residues is a distinct problem with this type of modification, but this may be avoided by reaction at a relatively low pH, in the presence of a high concentration of nucleophile. Reaction with hydrazines and hydroxylamines (Fig. 1.4, **2** and **3**) to form hydrazides and hydroxamic

acids is often preferable if reaction with carboxylic acids is necessary, as they generally are more reactive at lower pH than aliphatic amines, greatly reducing undesired side reactions.

1.2.4 Fluorescent labelling via expression as a GFP fusion protein

A method of protein modification which has been specifically adopted for fluorescent labelling of target proteins *in vivo* is expression as a fusion protein with *Aequorea victoria* green fluorescent protein (GFP) (*3, 59*). For *in vivo* application, conventional covalent protein labelling requires the protein to be fluorescently labelled prior to its introduction into the cellular environment. This however is not ideal, as proteins are often post-translationally modified *in vivo* for targeting to their required location within the cell and protein labelled *in vitro* may not contain the required modification. As GFP becomes fluorescent *in vivo* after oxidative modification (*3*), GFP fusions which can be expressed *in situ* within the host cells offer the great advantage of avoiding the use of *in vitro* labelling and microinjection. The expression of these proteins within the host cells also means that the GFP fusions can be localised to specific sites by the appropriate targeting signals (*60, 61*). Site directed mutagenesis has provided a range of colour variants of GFP (*26*), which allow the simultaneous observation of cellular location and interaction of several proteins of interest simultaneously by fluorescence microscopy and FRET (*62*).

Although this is by far the most widely used method for *in vivo* protein labelling (*63-65*), a drawback is that both GFP must fold correctly as a fusion protein to allow maturation, and expression of proteins as GFP fusions may interfere with target folding (*66*) which might have an adverse affect on the biological activity. A number of recent developments have been made with respect to the specific chemical labelling of target proteins directly *in vivo* (*4, 67*). These methods for the incorporation of small fluorescent labels into protein domains include FlAsH (1.2.5, Fig. 1.5) (*68*) and covalent labelling of *Homo sapiens* O^6-alkylguanine DNA alkyl transferase (1.2.6, Fig. 1.6) (*69*).

1.2.5 *In vivo* labelling with fluorescent dyes using FlAsH (*68*)

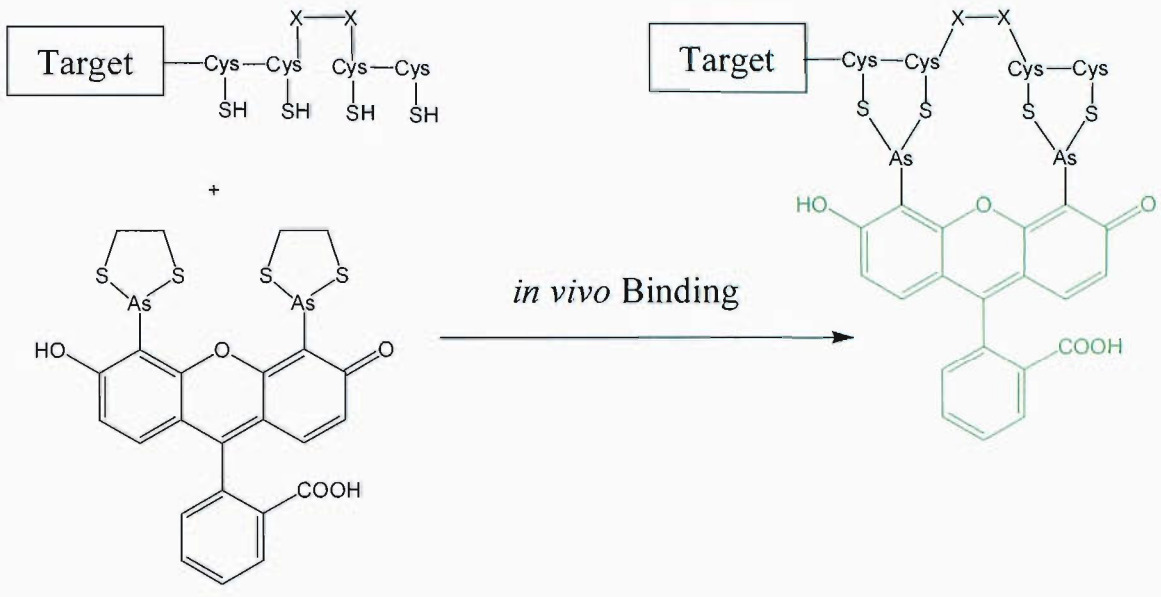


Figure 1.5 FlAsH: biarsenical fluorescein protein labelling via binding to a tetracysteine motif which is fused to the target protein. The label fluorescence greatly increases upon binding.

A technique developed for site specific protein labelling within cells with a low molecular weight fluorophore is the site specific labelling of a tetracysteine motif (Fig. 1.5), introduced into the target protein domain at the required position by mutagenesis. Arsenoxides have a well known affinity for closely spaced cysteine pairs, so two arsenoxide groups were incorporated into the fluorescein structure, forming FlAsH (*70*), which binds strongly to peptides containing the CCXXCC motif, where X is any amino acid other than cysteine, which, as this is rare in natural protein sequences results in site specific labelling. Expression of the CCXXCC motif as a fusion with the target protein and the addition of the membrane permeable, non-fluorescent FlAsH to the cells results in incorporation of the label into the binding motif. Upon formation of the label:protein complex, the label becomes dramatically more fluorescent (*71*). Several variants of FlAsH are available, a red biarsenic ligand ReAsH (*71*) and blue label, ChoXAsH may also be used. This is an excellent method due to its specificity and low background fluorescence. However, the label and tetracysteine motif are toxic to the host cells and the tetracysteine motif must be in its reduced form for labelling to occur.

1.2.6 *In vivo* fluorescent labelling of *Homo sapiens* O^6-alkylguanine DNA alkyl transferase

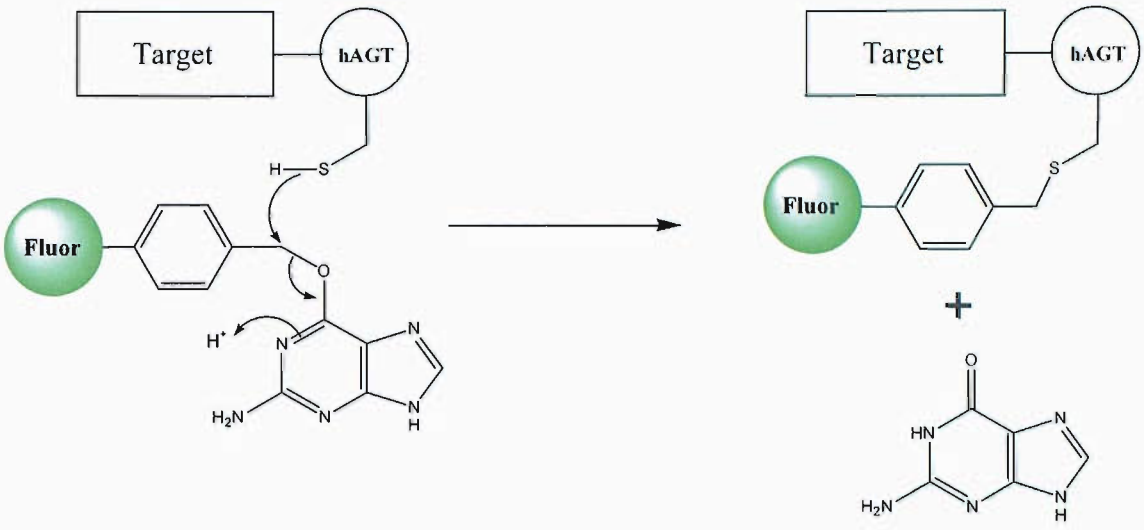


Figure 1.6 Covalent labelling of *Homo sapiens* O^6-alkylguanine DNA alkyl transferase with benzyl guanine derivatives.

A second method of *in vivo* labelling involves the target protein being expressed as a fusion with *Homo sapiens* O^6-alkylguanine DNA alkyl transferase (hAGT). Fluorescently labelled O^6-benzylguanine derivatives irreversibly and specifically label the hAGT which is fused to the target proteins and expressed within mammalian cells (*72*). This reaction involves specific reaction of the O^6-benzylguanine derivative with a reactive cysteine residue (Fig. 1.6) HAGT has a size of 25 kDa, which is similar to that of GFP, therefore the effect of these types of fusion upon target protein activity should be similar, and generally non interfering. A range of fluorophores have been prepared as O^6-benzylguanine derivatives, allowing FRET based studies between hAGT labelled fusions and GFP variant fusion proteins (*4*). The rate of reaction of label with hAGT is not significantly affected by the nature of the label incorporated, which allows the incorporation of a wide range of probes to hAGT-target fusion proteins (*73*). One drawback of this method is that it may only be applied for protein labelling in hAGT deficient cell lines, as the benzyl guanine substrates will react with the naturally occurring hAGT.

1.3 Intein mediated ligation (*74*)

The methodology for protein modification *in vitro* as described in section 1.2 has remained more or less unchanged over recent years, the most commonly used methods being non-specific (amine, thiol or carboxylate modification) except in the case of proteins which contain only one cysteine residue either naturally or through genetic manipulation (*27*). An inherent problem with these methods is the inability for them to be used directly in the *in vivo* environment. Studying proteins in their natural environment is essential for true understanding of their function in relation to other proteins in the living cell.

Protein splicing and its recent application via 'expressed protein ligation' (EPL) (*75, 76*), otherwise known as 'intein mediated ligation' (IML) (*77, 78*) have the potential to have a dramatic impact upon the field of protein science, aiding the synthesis, modification and isolation of proteins.

1.3.1 Inteins and protein splicing (*79*)

Protein splicing is a naturally occurring process in which an intervening protein domain, or 'intein' (*80*) catalyses its own excision from two flanking peptide sequences, or 'exteins,' joining the exteins with a native peptide bond, forming a new active protein domain. The intein mediated splicing reaction has been shown to be independent of the flanking extein sequences, allowing an intein to ligate almost any protein sequence.

The first intein was discovered in 1990 (*81, 82*) and since then the reaction mechanism and its products have been well characterised. The first intein identified was the *Saccharomyces cerevisiae* VMA1 intein, and since this over 100 new protein splicing elements have been identified (*83*). Alignment and examination of the known intein sequences has indicated the importance of the N- and C-terminal intein residues and the first downstream amino acid (*84*), the mechanistic roles of which were further confirmed by site directed mutagenesis (*85*).

The critical residues of the intein domain for protein splicing were found to be a cysteine or serine at the intein N-terminus and a histidine-asparagine sequence at the C-terminus followed by a cysteine, serine or threonine residue. The discovery of a thermophilic intein in *Pyrococcus*, whose splicing reaction was found be sufficiently slowed or stopped by lowering the temperature allowed the demonstration of the splicing reaction of a purified precursor and demonstrated the existence of a branched intermediate (*86*). From this it was found that the first step of protein splicing was linear thioester formation at the intein N-terminus (*87*). Cleavage of the C-terminal peptide bond was found to occur via succinimide formation by the final asparagine residue (*88*).

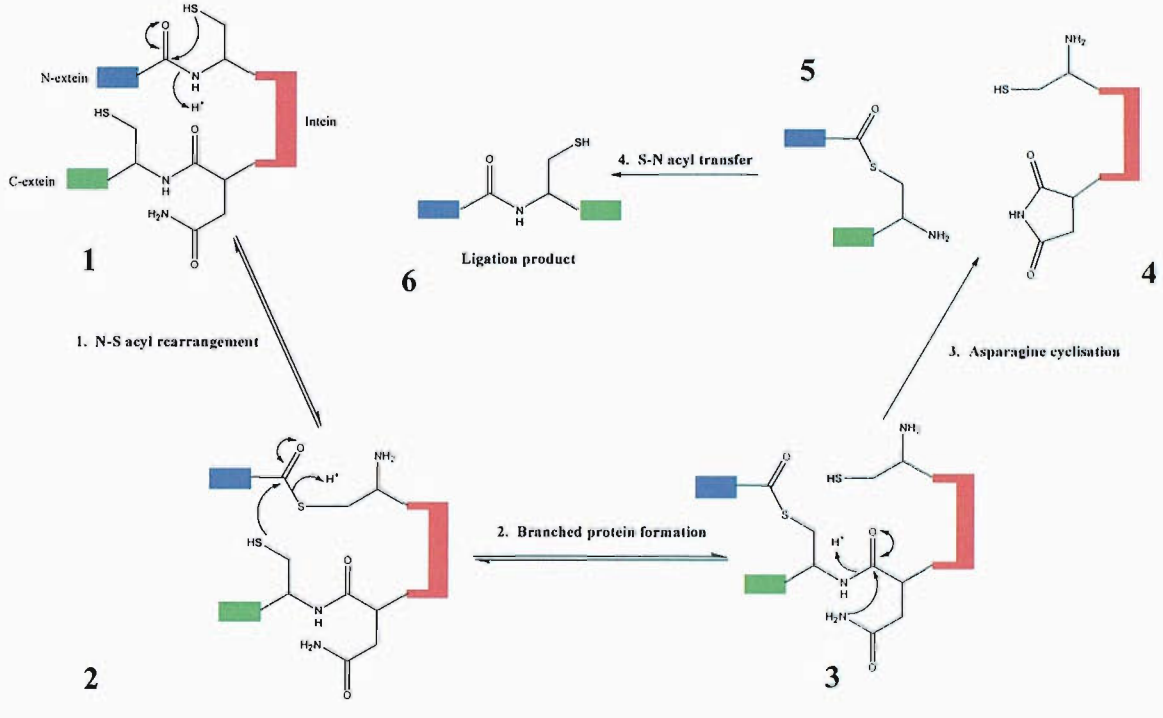


Figure 1.7 Intein mediated protein splicing mechanism (*79*)

These results can be explained by a four step mechanism for protein splicing (Fig. 1.7) (*89*). The splicing reaction begins with an N-S or N-O acyl rearrangement (step 1) which generates a (thio)ester bond between the intein N-terminal residue and the upstream extein residue (**2**). A trans(thio)esterification reaction (step 2) then occurs via intramolecular attack by the hydroxyl or sulphydryl group of the first residue downstream of the intein domain. The resulting protein (**3**) then has two N-termini and is referred to as the branched intermediate. Succinimide formation (**4**) by the intein C-

terminal asparagine releases it from the two extein sequences (step 3), which are now linked by a (thio)ester bond (**5**). The splicing reaction is completed by rearrangement of the (thio)ester bond via a spontaneous O-N or S-N acyl rearrangement (step 4) forming the amide product (**6**).

The first step, N-S or N-O rearrangement at the N-terminal intein splicing junction is surprising as the amide-thioester equilibrium usually strongly favours amide bond formation. The crystal structure of an intein precursor (prior to the first step) shows the peptide bond at the splice junction to be in an energetically disfavoured conformation, and the strain is released upon formation of the thioester (*90*).

The modification of intein domains via site directed mutagenesis has enabled the development of novel techniques for the study of protein *in vitro*. The autoprocessing ability of inteins was harnessed by formation of constructs for protein purification (*91, 92*). The mutation of residues at the active site of the intein allowed the splicing reaction to be stopped at an intermediate point in which a thioester is formed between the N-extein and intein domains, but further reaction is blocked.

In one system of protein purification (the IMPACT™ system), the C-extein domain of the *Saccharomyces cerevisae* VMAI intein was removed and replaced with the *Bacillus circulans* chitin binding domain for affinity purification (Fig. 1.8) (*93*). The protein of interest, or 'target' domain is expressed as a fusion protein, replacing the N-extein. Rearrangement of the bond between the first residue of the intein domain and the C-terminal residue of the target results in the formation of a thioester between the two domains. In the purification system, the fusion protein is bound to a chitin resin and the target domain released by thiol induced cleavage, allowing elution of the target protein from the affinity resin (*91*).

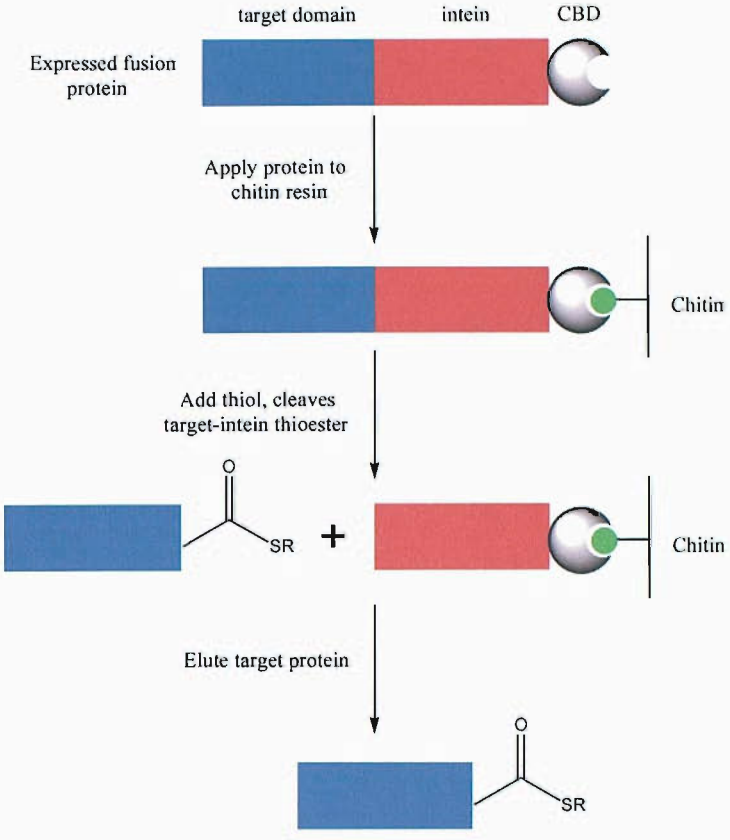


Figure 1.8 Protein purification using intein domains

If the thiol chosen to cleave the target-intein thioester is chosen appropriately the cleaved target protein will form a stable α-thioester (*94*) that is relatively stable to hydrolysis. However, the thioester will undergo attack by stronger nucleophiles such as an N-terminal cysteine. This intein mediated activation of the C-terminus of target protein, followed by ligation with a peptide containing an N-terminal cysteine residue is known as intein mediated ligation (IML) (*77*).

1.3.2 The intein mediated ligation reaction

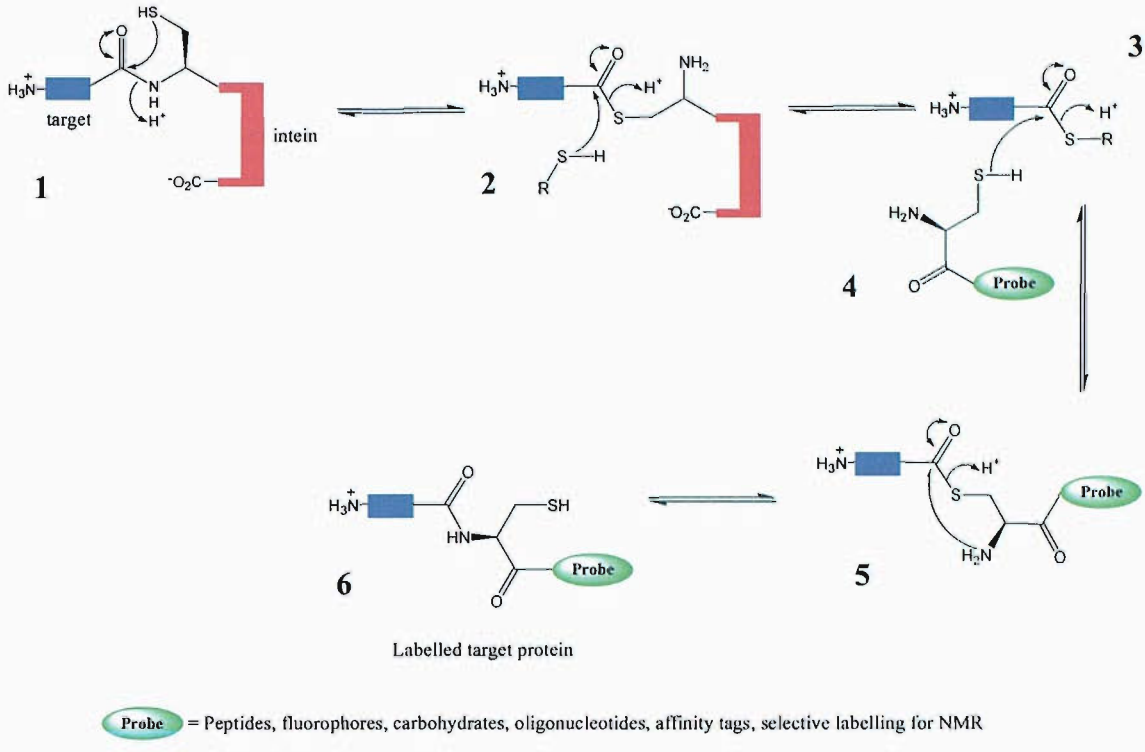


Figure 1.9 Intein mediated ligation. Acyl transfer of the target-intein fusion protein (**1**) from the alpha amino group to the thiol of the reactive cysteine yields thioester (**2**), which is attacked by a low molecular weight thiol to give the more reactive thioester (**3**). Attack of a cysteine based label (**4**) yields thioester (**5**), which rearranges to give the thermodynamic amide product (**6**).

Intein mediated ligation (Fig. 1.9) allows the chemical ligation of two polypeptides, at least one of which is recombinant in nature. This allows the semisynthesis of protein domains with the appropriate introduction of the diverse range of functionalities possible via synthetic organic chemistry. IML has been used to introduce functionalities as broad as carbohydrates (*95*), fluorescent labels (*96*), oligonucleotides (*97*), peptides (*75*) and protein domains (*98*), and isotopic labels for NMR studies into large recombinant proteins (*99*). Simple intein mediated activation of a protein domain C-terminus can therefore theoretically allow specific modification of the protein with the probe of choice by synthesising a cysteine based label conjugate.

The distinct advantage of this methodology over the previously described approaches is the specificity of the labelling reaction. A possible future development is that if the label may enter the cellular environment, protein domains expressed within cells may be fluorescently labelled *in vivo*. Our aim is to apply intein mediated ligation to the fields of fluorescent labelling and electrode immobilisation of proteins (as described in chapter 5).

1.3.3 General methodology for intein mediated ligation

Intein mediated ligation relies upon the reaction of a target protein with an intein formed C-terminal thioester with a synthetic molecule containing the required modification, and an N-terminal cysteine. The synthetic peptide may be modified as required to contain unnatural amino acids and chemical or biochemical labels (*100*). Intein mediated ligation utilises the IMPACT™ system, as is described in section 1.3. The target protein is cloned into an expression vector, removing the stop codon allowing expression as a fusion protein in which the target is attached to the intein N-terminus and the fusion protein purified from the cell lysate by chitin affinity (*91*).

Direct reaction of the target-intein thioester with the N-terminal cysteine containing label has been observed to be inefficient (*75*). Therefore the target-intein thioester is first reacted with a low molecular weight thiol, forming a new more reactive thioester (*101*). Generally the thiol chosen is small and strongly nucleophillic to maximise the probability of accessing the target-intein thioester and reacting with it. Two different practical methods have been investigated at this point. In one method the target protein may be eluted from the column and stored, followed by ligation with the peptide (*102*). In this case the thiol chosen should form a thioester which is less reactive and therefore less prone to hydrolyse (*103*), and typically β-mercaptoethanol is used in this case (*94, 103*). In the other method, IML can be performed directly on the chitin resin and both transesterification and ligation with the synthetic peptide may occur simultaneously (*75*). Using this method, as the isolation of the intermediate thioester is not required, a more reactive intermediate thioester may be formed to increase the efficiency of the ligation reaction, typically mercaptoethanesulphonic acid (*98*) or thiophenol (*75*) are applied in this case. This method requires initial binding of the target-intein thioester to

the chitin resin, followed by exposure typically with a mixture of up to 2 % thiophenol (*104, 105*) and the peptide containing an N-terminal cysteine residue. The reaction is then left sealed overnight to allow the ligation reaction to occur prior to elution of the ligated product.

1.3.4 Limitations of intein mediated ligation (*77*)

There are a number of critical factors which have been found to interfere with the viability of the intein mediated ligation reaction for protein labelling.

i. The ability to cleave the target-intein thioester both *in vivo* and *in vitro* may be affected by the target protein sequence (*77*).

ii. *In vivo* cleavage of the target-intein thioester lowers the yield of intact fusion protein isolated, with the observation of 10 – 50 % cleaved fusion protein typically observed (*77*). Although this may hinder the isolation of intact fusion protein for intein mediated ligation, it indicates that the target-intein thioester is reactive. On the other hand, the nature of the target protein may cause the target-intein thioester to be virtually unreactive towards intein mediated ligation.

iii. Direct reaction of the cysteine based label with the target-intein thioester is inefficient (*75*), and the reaction requires the addition of high concentrations of a low molecular weight thiol (up to 3 % β-mercaptoethanol (*97*), 2 % thiophenol (*105*), or 2 % MESA (*95*)) to form a more reactive intermediate thioester with which the label may more efficiently react. Even with the addition of the low molecular weight thiol, reaction yields of < 60 % over long reaction times (3 days) have been observed (*94*). It would be preferable to avoid the addition of this thiol.

iv. The low molecular weight thiols are often added at high concentration and are toxic towards the user and potentially the cells that it may be added to, and highly reactive. This may cause reaction with less robust proteins especially

those for instance containing delicate cofactors such as iron-sulphur clusters (*57*) which may react with the aggressive thiols added.

v. The hydrolysis of the intermediate α-thioester product may occur during the reaction, reducing the yield of labelled protein and complicating the purification procedure (*103*).

vi. Genetic modification of the target gene is necessary as the stop codon must be removed to allow expression of the full length fusion protein (*91*).

1.3.5 Fluorescent labelling via intein mediated ligation

There are few reported examples of the practical application of intein mediated ligation for the simple C-terminal labelling of a protein domain with a fluorescent cysteine derivative. Tolbert and Wong investigated reaction of a cysteine derivative with the *E. coli* maltose binding protein (*106*). The synthesis of a dansyl cysteine derivative was successful and ligation with the target protein was achieved but the ligation reaction did not proceed cleanly. Under the reaction conditions (30 mM mercaptoethanesulphonic acid and 2 mM cysteine derivative) some of the target protein was observed to be unstable and precipitated. Iakovenko labelled the Rab7 protein with a short peptide (*107*) (CL[dansyl]SCSC) in the presence of 500 mM mercaptoethanesulphonic acid over 24 hours at room temperature. Ebright reported the labelling of the *E. coli* RNA polymerase using a fluorescein derivative of cysteine (*96*), which involved the coupling of fluorescein iodoacetamide with a trityl protected cysteine. Deprotection was followed by reaction with the RNA polymerase-intein fusion protein on a chitin column in buffer containing 0.5 % thiophenol and 0.5 mM tricarboxyethyl phosphine, a water soluble reducing agent which has been shown to be efficient at reducing disulphides. After eight hours of reaction the labelled protein was eluted and the labelled protein utilised in FRET based DNA interactions with Cy5 labelled DNA substrates. The yield of fluorescently labelled protein was 90 %.

1.4 Aims of the project

The aim of this project was to investigate new methods for the fluorescent labelling of proteins and their immobilisation on electrode surfaces using intein mediated ligation, including experiments addressing some of the limitations of this methodology.

1.5 References

1. Hovius, R., Vallotton, P., Wohland, T., and Vogel, H. (2000) *Trends in Pharmacological Science 21*, 266-273.

2. Parola, A. L., Lin, S., and Kobilka, B. K. (1997) *Analytical Biochemistry 254*, 88-95.

3. Tsien, R. Y. (1998) *Annual Review of Biochemistry 67*, 509-544.

4. Miller, L. W., and Cornish, V. W. (2005) *Current Opinion in Chemical Biology 9*, 56-61.

5. Weiss, S. (1999) *Science 283*, 1676-1683.

6. Cherry, R. J., Morrison, I. E., Karakikes, I., Barber, R. E., Silkstone, G., and Fernandez, N. (2003) *Biochemistry Society Transactions 31*, 1028-1031.

7. Jameson, D. M., and Seifried, S. E. (1999) *Methods 19*, 222-233.

8. Selvin, P. R. (2000) *Nature Structural Biology 7*, 730-734.

9. Yeung, T., Touret, N., and Grinstein, S. (2005) *Current Opinion in Microbiology 8*, 350-358.

10. Wallrabe, H., and Periasamy, A. (2005) *Current Opinion in Biotechnology 16*, 19-27.

11. Heyduk, T., Ma, Y., Tang, H., and Ebright, R. H. (1996) *Methods in Enzymology 274*, 492-503.

12. Hanley, Q. S., and Clayton, A. H. (2005) *Journal of Microscopy 218*, 62-67.

13. Owicki, J. C. (2000) *Journal of Biomolecular Screening 5*, 297-306.

14. Chicchi, G. G., Cascieri, M. A., Graziano, M. P., Calahan, T., and Tota, M. R. (1997) *Peptides 18*, 319-321.

15. Tompa, P., Batke, J., Ovadi, J., Welch, G. R., and Srere, P. A. (1987) *Journal of Biological Chemistry 262*, 6089-6092.

16. Ason, B., and Reznikoff, W. S. (2004) *Nucleic Acids Research 32*, E83.

17. Sittampalam, G. S., Kahl, S. D., and Janzen, W. P. (1997) *Current Opinion in Chemical Biology 1*, 384-391.

18. Levine, L. M., Michener, M. L., Toth, M. V., and Holwerda, B. C. (1997) *Analytical Biochemistry 247*, 83-88.

19. Jantz, D., and Berg, J. M. (2002) *Nature Biotechnology 20*, 126-127.

20. Gordon, G. W., Berry, G., Liang, X. H., Levine, B., and Herman, B. (1998) *Biophysical Journal 74*, 2702-2713.

21. Lilley, D. M. J., and Wilson, T. J. (2000) *Current Opinion in Biotechnology 4*, 507-517.

22. Selvin, P. R. (1995) in *Biochemical Spectroscopy*, 300-334.

23. Piehler, J. (2005) *Current Opinion in Structural Biology 15*, 4-14.

24. Van Roessel, P., and Brand, A. H. (2002) *Nature Cell Biology 4*, E15-20.

25. Michalet, X., Kapanidis, A. N., Laurence, T., Pinaud, F., Doose, S., Pflughoefft, M., and Weiss, S. (2003) *Annual Review of Biophysics and Biomolecular Structure 32*, 161-182.

26. Phillips, G. J. (2001) *Fems Microbiology Letters 204*, 9-18.

27. Means, G. E. and Feeney, G. E. (1990) *Bioconjugate Chemistry 1*, 2-12.

28. Davis, B. G. (2003) *Current Opinion in Biotechnology 14*, 379-386.

29. Means, G. E., and Feeney, G. E. (1968) *Biochemistry 7*, 1366-1371.

30. Nishikawa, A. H., Morita, R. Y., and Becker, R. R. (1968) *Biochemistry 7*, 1506-1513.

31. Rice, R. H., Means, G. E., Brown, W. D. (1977) *Biochimica and Biophysica Acta. 492*, 316-321.

32. Kapanidis, A. N., and Weiss, S. (2002) *Journal of Chemical Physics 117*, 10953-10964.

33. Rice, R. H., and Means, G. E., (1971) *Journal of Biological Chemistry 246*, 831-832.

34. Mears, C. F., McCall, M. J., Deshpande, S. V., DeNArdo, S. J., and Goodwin, D. A. (1988) *International Journal of Cancer 2*, 99-102.

35. Chen, V. J., and Wold, F. (1984) *Biochemistry 23*, 3306-3311.

36. Bogusiewicz, A., Mock, N. I., and Mock, D. M. (2005) *Analytical Biochemistry 337*, 98-102.

37. Kaiser, E. T., Laurence, D. S., and Rokita, S. E. (1985) *Annual Review of Biochemistry 54*, 565-595.

38. Kalkhof, S., Ihling, C., Mechtler, K., and Sinz, A. (2005) *Analytical Chemistry 77*, 495-503.

39. Sinz, A. (2003) *Journal of Mass Spectrometry 38*, 1225-1237.

40. Parola, A. L., Lin, S. S., and Kobilka, B. K. (1997) *Analytical Biochemistry 254*, 88-95.

41. Forsyth, W. R., Gilson, M. K., Antosiewicz, J., Jaren, O. R., and Robertson, A. D. (1998) *Biochemistry 37*, 8643-8652.

42. Lomant, A. J., and Fairbanks, G. (1976) *Journal of Molecular Biology 104*, 243-261.

43. Wilbur, D. S., Pathare, P. M., Hamlin, D. K., Rothenberg, S. P., and Quadros, E. V. (1999) *Bioconjugate Chemistry 10*, 912-920.

44. Banks, P. R., and Paquette, D. M. (1995) *Bioconjugate Chemistry 6*, 447-458.

45. Zhu, K., Gonzalez-Pedrajo, B., and Macnab, R. M. (2002) *Biochemistry 41*, 9516-9524.

46. Lefevre, C., Kang, H. C., Haugland, R. P., Malekzadeh, N., Arttamangkul, S., and Haugland, R. P. (1996) *Bioconjugate Chemistry 7*, 482-489.

47. Seiler, N. (1970) *Methods of Biochemical analysis 18*, 259-337.

48. Valeur, B., and Keh, E. (1979) *Journal of Physical Chemistry 83*, 3305-3307.

49. Cha, A., and Bezanilla, F. (1997) *Neuron 19*, 1127-1140.

50. Brake, J. M., and Wold, F. (1962) *Biochemistry 1*, 386-391.

51. Musci, G., and Berliner, L. J. (1986) *Biochemistry 25*, 4887-4891.

52. Weltman, J. K., Szaro, R. P., Frackelton, A. R., Dowben, R. M., Bunting, J. R., and Calthow, R. E.,. (1973) *Journal of Biological Chemistry 248*, 3173-3177.

53. Heitz, J. R., Anderson, C. D. and Anderson, B. M., (1968) *Archives of Biochemistry and Biophysics 127*, 627.

54. Apenten, R. K. O. (1998) *International Journal of Biological Macromolecules 23*, 19-25.

55. Matsumoto, K., Davis, B. G., and Jones, J. B. (2002) *Chemistry 8*, 4129-4137.

56. Haring, D., Lees, M. R., Banaszak, L. J., and Distefano, M. D. (2002) *Protein Engineering 15*, 603-610.

57. Kriek, M., Peters, L., Takahashi, Y., and Roach, P. L. (2003) *Protein Expression and Purification 28*, 241-245.

58. Hoare, D. G. and Koshland, D. E. (1966) *Journal of the American Chemical Society 88*, 2057.

59. Shimomura, O. (2005) *Journal of Microscopy 217*, 1-15.

60. Rook, M. S., Lu, M., and Kosik, K. S. (2000) *Journal of Neuroscience 20*, 6385-6393.

61. Gonzalez, C., and Bejarano, L. A. (2000) *Trends in Cell Biology 10*, 162-165.

62. Gordon, G. W., Berry, G., Liang, X. H., Levine, B., and Herman, B. (1998) *Biophysical Journal 74*, 2702-2713.

63. Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002) *Nature Reviews in Molecular Cell Biology 3*, 906-918.

64. O'Rourke, N. A., Meyer, T., and Chandy, G. (2005) *Current Opinion in Chemical Biology 9*, 82-87.

65. March, J. C., Rao, G., and Bentley, W. E. (2003) *Applied Microbiology and Biotechnology 62*, 303-315.

66. Chang, H. C., Kaiser, C. M., Hartl, F. U., and Barral, J. M. (2005) *Journal of Molecular Biology, 353*, 397-409.

67. Johnsson, N., and Johnsson, K. (2003) *Chembiochem 4*, 803-810.

68. Griffin, B. A., Adams, S. R., and Tsien, R. Y. (1998) *Science 281*, 269-272.

69. Keppler, A., Pick, H., Arrivoli, C., Vogel, H., and Johnsson, K. (2004) *Proceedings of the National Academy of Sciences of the United States of America 101*, 9955-9959.

70. Griffin, B. A., Adams, S. R., Jones, J., and Tsien, R. Y. (2000) *Methods in Enzymology 327,* 565-578.

71. Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J., and Tsien, R. Y. (2002) *Journal of the American Chemical Society 124*, 6063-6076.

72. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003) *Nature Biotechnology 21*, 86-89.

73. Kindermann, M., Sielaff, I., and Johnsson, K. (2004) *Bioorganic and Medicinal Chemistry Letters 14*, 2725-2728.

74. David, R., Richter, M. P. O., and Beck-Sickinger, A. G. (2004) *European Journal of Biochemistry 271*, 663-677.

75. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

76. Muir, T. W. (2001) *Synlett*, 733-740.

77. Xu, M. Q., and Evans, T. C. (2001) *Methods 24*, 257-277.

78. Wood, R. J., Pascoe, D. D., Brown, Z. K., Medlicott, E. M., Kriek, M., Neylon, C., and Roach, P. L. (2004) *Bioconjugate Chemistry 15*, 366-372.

79. Evans, T. C., and Xu, M. Q. (2002) *Chemical Reviews 102*, 4869-4883.

80. Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J., and Belfort, M. (1994) *Nucleic Acids Research 22*, 1125-1127.

81. Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., and Stevens, T. H. (1990) *Science 250*, 651-657.

82. Hirata, R., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990) *Journal of Biological Chemistry 265*, 6726-6733.

83. Perler, F. B. (2000) *Nucleic Acids Research 28*, 344-345.

84. Pietrokovski, S. (1994) *Protein Science 3*, 2340-2350.

85. Hirata, R., and Anraku, Y. (1992) *Biochemistry and Biophysics Research Communications 188*, 40-47.

86. Xu, M. Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993) *Cell 75*, 1371-1377.

87. Chong, S. R., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M. Q. (1996) *Journal of Biological Chemistry 271*, 22159-22168.

88. Shao, Y., Xu, M. Q., and Paulus, H. (1995) *Biochemistry 34*, 10844-10850.

89. Paulus, H. (2001) *Bioorganic Chemistry 29*, 119-129.

90. Poland, B. W., Xu, M. Q., and Quiocho, F. A. (2000) *Journal of Biological Chemistry 275*, 16408-16413.

91. Chong, S. R., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. (1997) *Gene 192*, 271-281.

92. Zhang, A. H., Gonzalez, S. M., Cantor, E. J., and Chong, S. R. (2001) *Gene 275*, 241-252.

93. Chong, S. R., Montello, G. E., Zhang, A. H., Cantor, E. J., Liao, W., Xu, M. Q., and Benner, J. (1998) *Nucleic Acids Research 26*, 5109-5115.

94. Dawson, P. E., Churchill, M. J., Ghadiri, M. R., and Kent, S. B. H. (1997) *Journal of the American Chemical Society 119*, 4325-4329.

95. Macmillan, D., and Bertozzi, C. R. (2000) *Tetrahedron 56*, 9515-9525.

96. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001) *Cell 106*, 453-463.

97. Lovrinovic, M., Seidel, R., Wacker, R., Schroeder, H., Seitz, O., Engelhard, M., Goody, R. S., and Niemeyer, C. M. (2003) *Chemical Communications*, 822-823.

98. Evans, T. C., Benner, J., and Xu, M. Q. (1999) *Journal of Biological Chemistry 274*, 3923-3926.

99. Xu, R., Ayers, B., Cowburn, D., and Muir, T. W. (1999) *Proceedings of the National Academy of Sciences of the United States of America 96*, 388-393.

100. David, R., Richter, M. P., and Beck-Sickinger, A. G. (2004) *European Journal of Biochemistry 271*, 663-677.

101. Hofmann, R. M., and Muir, T. W. (2002) *Current Opinion in Biotechnology 13*, 297-303.

102. Evans, T. C., Benner, J., and Xu, M. Q. (1998) *Protein Science 7*, 2256-2264.

103. Ayers, B., Blaschke, U. K., Camarero, J. A., Cotton, G. J., Holford, M., and Muir, T. W. (1999) *Biopolymers 51*, 343-354.

104. Severinov, K., and Muir, T. W. (1998) *Journal of Biological Chemistry 273*, 16205-16209.

105. Zhang, Z. S., Shen, K., Lu, W., and Cole, P. A. (2003) *Journal of Biological Chemistry 278*, 4668-4674.

106. Tolbert, T. J., and Wong, C. H. (2000) *Journal of the American Chemical Society 122*, 5421-5428.

107. Iakovenko, A., Rostkova, E., Merzlyak, E., Hillebrand, A. M., Thoma, N. H., Goody, R. S., and Alexandrov, K. (2000) *Febs Letters 468*, 155-158.

Chapter 2:- Cloning, expression and purification of intein fusion proteins

2.1 Introduction

Intein mediated activation of the C-terminus of a protein domain has provided a general site specific method for the attachment of a wide range of molecules to proteins (*1, 2*). The aims are to apply this methodology to develop a straightforward method for site specific fluorescent labelling of proteins, and to provide an orientation specific method for the immobilisation of potentially delicate proteins on electrode surfaces for biophysical and other diverse studies.

The target proteins selected for C-terminal intein fusion were BioB, LipA, Fpr, SolA and PKBα, and are discussed in detail below.

2.1.1 Biotin synthase (BioB) (*3*)

E. coli Biotin synthase inserts a sulphur atom into dethiobiotin (Fig. 2.1). BioB contains a reactive $[4Fe\text{-}4S]^{+}$ (*4*) cofactor and is a member of the radical SAM family of proteins (*5*). The hydrogen atoms are abstracted by adenosyl radicals, that are derived from SAM which is homolytically cleaved in a one electron reduction by the BioB iron-sulphur cluster. This protein provides an interesting electrochemical and fluorescent labelling target. FRET based binding assays and electrochemical study (*6, 7*) may provide an insight into the nature and interaction of the iron sulphur cluster with electron transfer cofactors required for activity. The BioB [4Fe-4S] cluster is known to be labile (*8*). Aggressive labelling conditions may affect the stability of the protein, providing a test of the mildness of intein mediated ligation conditions for application to unstable proteins.

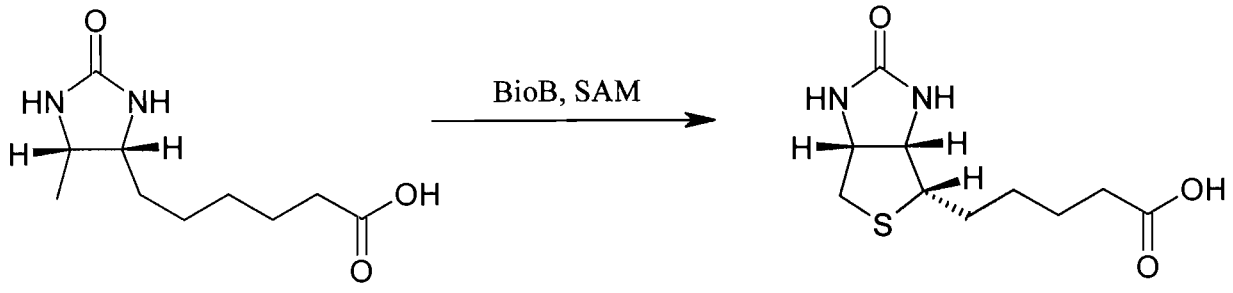


Figure 2.1 Biotin synthase sulphur insertion reaction

2.1.2 Lipoyl synthase (LipA) (*9*)

Sulfolobus solfataricus Lipoyl synthase is required for the insertion of sulphur into octanoyl groups which are attached to the E2 subunit of Pyruvate dehydrogenase complex (*10*) (Fig. 2.2). LipA is a member of the radical SAM family and has recently been shown to contain two $[4Fe\text{-}4S]^{+}$ clusters (*11*) and to transfer sulphur from one of these to the substrate (*12*). In essence, LipA is a very similar protein to BioB (*13*).

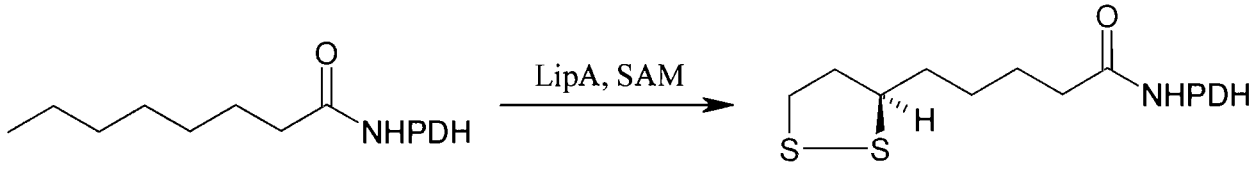


Figure 2.2 LipA reaction scheme

2.1.3 Flavodoxin reductase (*14*)

This *E. coli* flavoprotein (containing an flavin adenine dinucleotide (FAD) cofactor) is an electron transfer protein required for the reduction of Flavodoxin. These electron transfer proteins are required for reductive activation of many proteins such as anaerobic ribonucleotide reductase (*14*), biotin synthase (*15*), pyruvate formate lyase (*16*), and cobalamin-dependent methionine synthase (*17*). The flexible C-terminus of Fpr has previously been identified (*18*), which, combined with the relatively small size of this protein (27.6 kDa) may favour efficient formation and labelling of the target–intein thioester. Its electrochemistry has also been well characterised (*19-21*), making this an

ideal reference protein for both fluorescent labelling and surface immobilisation applications.

2.1.4 N-methyl-L-tryptophan oxidase (SolA) (*22*)

E. coli N-methyl-L-tryptophan oxidase oxidatively cleaves the methyl group from N-methyl-L-tryptophan forming hydrogen peroxide from dioxygen (*23*), the electrochemical detection of hydrogen peroxide has been well documented (*24-27*) so this protein has been selected for surface immobilisation. It is anticipated that the detection of hydrogen peroxide produced by the active protein will allow the stability of proteins immobilised in this way to be examined as H_2O_2 production will cease upon protein denaturation.

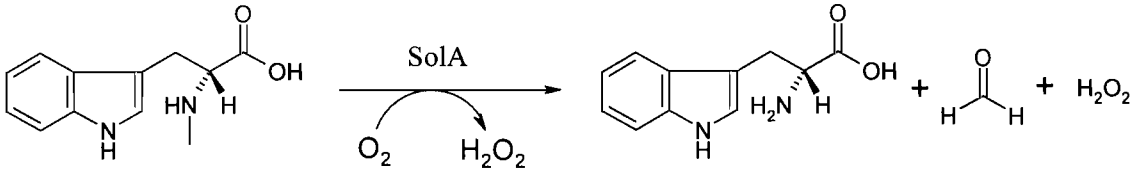


Figure 2.3 SolA reaction scheme

2.1.5 Protein kinase B α (PKBα) (*28*)

Homo sapiens PKBα has been found to phosphorylate serine and threonine residues that lie in Arg-Xaa-Arg-Xaa-Xaa-Ser/Thr- motifs (*29*). PKBα has been identified to play an essential role in tumour suppressor gene PTEN (phosphotase and tensin homologue deleted on chromosome 10) controlled tumorigenisis (*30*). Fluorescent labelling of this protein may allow its role in tumorigenisis to be further investigated.

This work was undertaken in collaboration with Veronique Calleja, Cancer Research UK.

2.1.6 Design of intein fusion proteins

The utilisation of intein mediated ligation requires the expression of target proteins as C-terminal intein fusions (*1*). Rearrangement of the target-intein amide has been shown to result in a unique thioester between C-terminus of the target protein, and N-terminal cysteine residue of the intein domain which may then undergo selective attack by a range of nucleophiles (*31*). The *Saccharomyces cerevisiae* VMAI intein (*32*) subcloned from vector pTYB1, IMPACT CN system (*33*) was used in these studies.

When considering plasmids for the expression of target-intein fusion proteins, the following points were considered:

- The stop codon of the target gene was removed to allow expression of the full length fusion protein.

- A glycine residue was added between the target gene and reactive cysteine residue at the intein N-terminus; this has been shown to be the optimal amino acid for thioester formation (*34*).

- A C-terminal His_6-tag has been added to the intein domain, to allow metal affinity purification of the expressed fusion protein.

- The domain labelled 'intein' includes DNA encoding intein and chitin binding domains from pTYB1, as an alternative purification method.

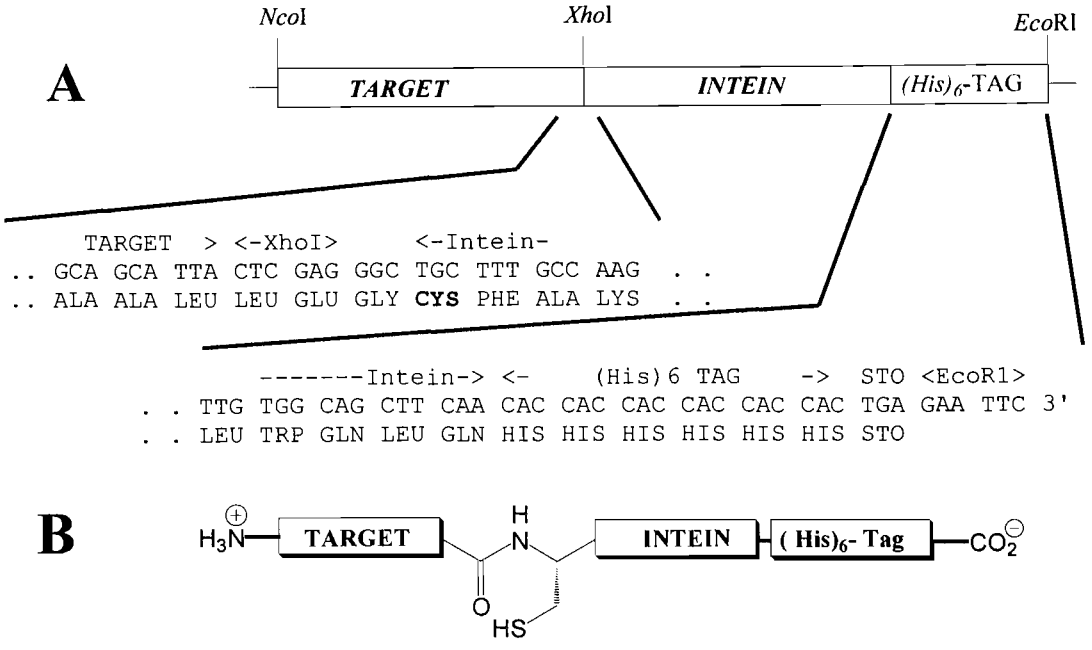


Figure 2.4 Construct design for the expression of intein fusion proteins; A = Primer design, B = Expressed protein.

2.2 Expression plasmid construction methodology

The strategy for the assembly of the required vectors is outlined in Fig. 2.5.

i. The correctly modified intein sequence was amplified from pTYBI and inserted into pBAD-TOPO.

ii. The intein encoding fragment was transferred to the *Xho*I – *Eco*RI sites of pBAD-HisA.

iii. Amplification of the target gene with *Nco*I - *Xho*I sites and insertion into pBAD-TOPO.

iv. Transfer of the modified target sequence into pRJW/2960/82 using unique *Nco*I – *Xho*I sites.

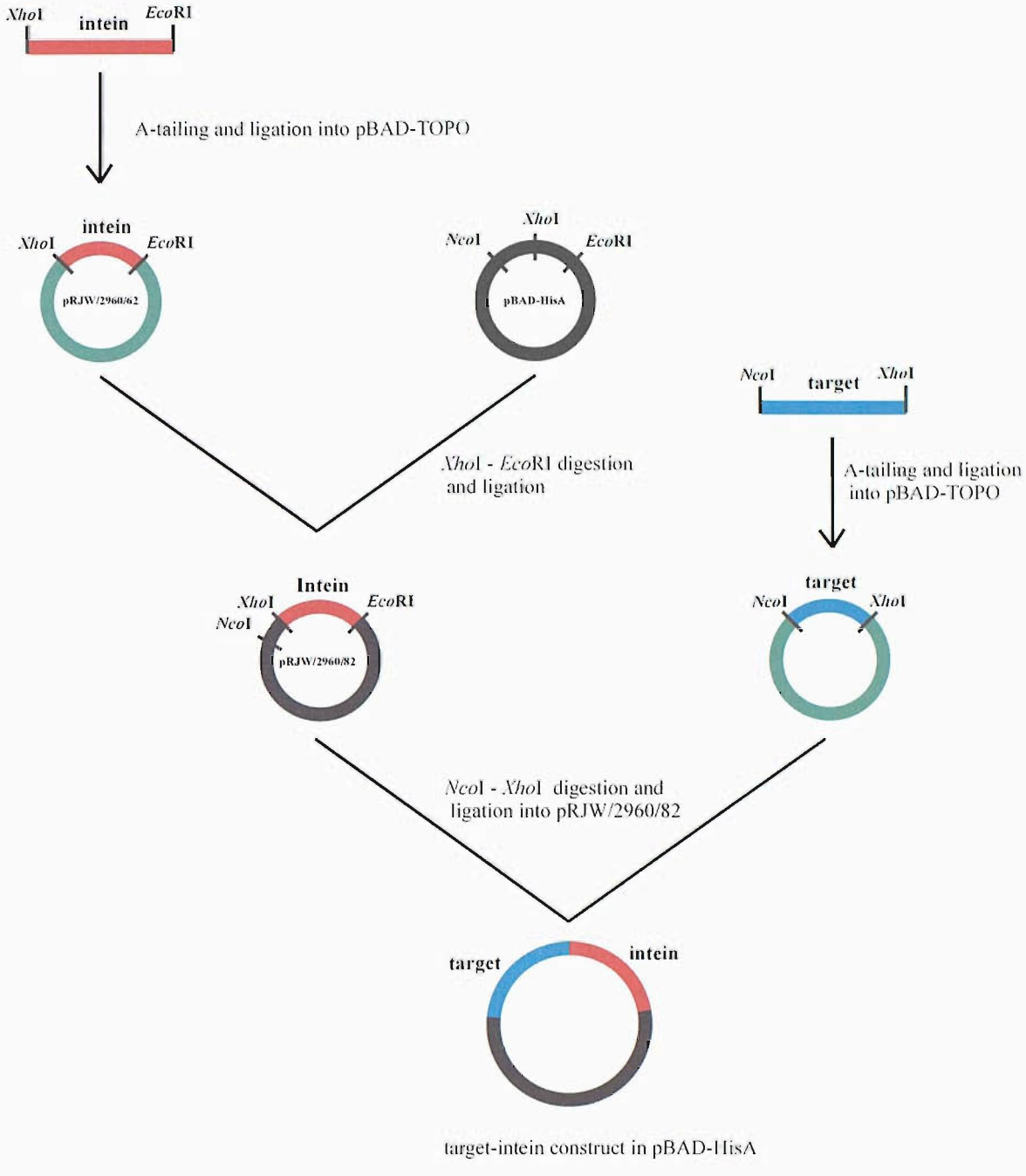


Figure 2.5 Schematic representation of assembly of expression plasmids, colours show the fate of DNA fragments.

Target genes were amplified by PCR using *Pyrococcus furiosus* (*Pfu*) DNA polymerase which has 5' – 3' proofreading activity and therefore a low error rate; 1.5 x 10^{-6} errors per base pair, compared with 1 x 10^{-4} for *Taq* polymerase (*35*). PCR products were A-tailed using *Thermus aquaticus* (*Taq*) DNA polymerase which adds a single deoxyadenosine in a template independent fashion to the 3' ends of the amplified

fragments, allowing ligation into the pBAD-TOPO TA subcloning vector which contains a 5' deoxythymidine overhang.

Plasmid pRJW/2960/82 containing the intein domain (see appendix A for plasmid maps) was constructed to allow the straightforward assembly of plasmids for the expression of target-intein fusion proteins under the control of an arabinose inducible promoter. The plasmid contains an *Nco*I site at the start codon of the target gene and a unique *Xho*I site allowing insertion of the target sequence in frame with the downstream gene encoding the intein domain.

Plasmids pRJW/2960/54 (Fpr), pRJW/2960/72 (BioB), pRJW/3104/35 (LipA), pRJW/3104/55 (SolA) and VC100 (PKBα) were constructed for subcloning of target genes into pRJW/2960/82 to form the required target-intein expression vectors. Table 2.1 lists target / intein vectors assembled.

Plasmid Name	Parent Plasmid	Insert	Resistence	Inducer
pRJW/2960/54A	pBAD-TOPO	*Nco*I-*fpr*-*Xho*I-*Bam*HI	Ampicillin	-
pRJW/2960/62	pBAD-TOPO	*Xho*I-*intein*-*Eco*RI	Ampicillin	-
pRJW/2960/72A	pBAD-TOPO	*Nco*I-*bioB*-*Xho*I-*Bam*HI	Ampicillin	-
pRJW/2960/82	pBAD-HisA	*Xho*I- *intein*-*Eco*RI	Ampicillin	Arabinose
pRJW/2960/88	pBAD-HisA	*Nco*I-*fpr*-*intein*-*Eco*RI	Ampicillin	Arabinose
pRJW/2960/89	pBAD-HisA	*Nco*I-*bioB*-*intein*-*Eco*RI	Ampicillin	Arabinose
pRJW/3104/35	pBAD-TOPO	*Nco*I-*lipA*-*Xho*I-*Bam*HI	Ampicillin	-
pRJW/3104/40	pBAD-HisA	*Nco*I-lipA-*intein*-*Eco*RI	Ampicillin	Arabinose
pRJW/3104/55	pBAD-TOPO	*Nco*I-*solA*-XhoI-*Pst*I	Ampicillin	-
pRJW/3104/100	pBAD-HisA	*Nco*I-*solA*-*intein*-*Eco*RI	Ampicillin	Arabinose
pVC100	pBAD-TOPO	*Nco*I-*PKBα*-*Xho*I	Ampicillin	
pRJW/3953/77	pBAD-HisA	*Nco*I-*PKBα*-*intein*-*Eco*RI	Ampicillin	Arabinose

Table 2.1 Assembled plasmids

2.2.1 Amplification and subcloning of the *Saccharomyces cerevisiae VMAI* intein gene

The Sce *VMAI* intein gene was amplified by PCR and inserted into the pBAD-TOPO subcloning vector. The *Xho*I-*intein*-*Eco*RI fragment was then subcloned into pBAD-HisA, forming pRJW/2960/82.

2.2.2 Amplification and cloning of the *Homo Sapiens* PKBα gene

The *PKBα* gene was amplified and cloned into the blunt end pBAD-TOPO vector by Veronique Calleja (VC100, Appendix A).

Subcloning of the *Nco*I-*PKBα*-*Xho*I fragment from VC100 was complicated by the presence of 3 *Nco*I and 2 *Xho*I restriction sites. The strategy employed was to restrict pVC100 in two stages. Firstly *Xho*I restriction and purification of the 1526 bp fragment, which was restricted with *Nco*I and the 1481 bp fragment gel purified yielding the required *Nco*I-*PKBα*-*Xho*I insert. This was ligated into pRJW/2960/82 forming pRJW/3953/77, *Nco*I-*PKBα*-*Xho*I-intein-*Eco*RI in pBAD-HisA (Appendix A).

2.3 Target-intein expression studies

For each of the BioB-, LipA-, Fpr- and SolA-intein fusion proteins, six *E. coli* strains were investigated for both optimal expression of the fusion protein and the extent of *in vivo* cleavage. Recombinant proteins are usually expressed in *E. coli* strain BL21(DE3) (*36*), which has in our hands been found to be very reliable for the expression of small globular proteins. The soluble expression of target-intein proteins in this strain was initially found to be at poor levels (< 1 % total protein), with significant proportions of the fusion protein found to have reacted with an intracellular nucleophile, further reducing the yield of fusion protein. Target-intein genes were then expressed in six *E. coli* strains to optimise yields of soluble protein (Table 2.2).

Strain	Genotype	Comments	References
BL21(DE3)	*E. coli* B, F-, *dcm*, o*mpT*, *hsdS*(*r*B-*mB*-), *gal*λ(DE3)	Generally used as an expression host. Deficient in Omptin and Lon proteases.	(*36*)
LMG194	F-, Δ*lacX74*, *galE*, *thi*, *rpsL*, Δ*phoA*, (PvuII), Δ*ara714*, *leu*::Tn10	Generally used as an expression host for pBAD derived plasmids. Δara714 blocks background arabinose catabolism.	(*37*)
AN1459	*K12, supE44, thi-1, leuB6, thr-1, ilvC, hsdR, recA*	*recA* mutant - involved in DNA repair and recombination, reduces homologous recombination of plasmid with host DNA, giving more stable inserts.	(*38*)
C43(DE3)	*E. coli* B, F-, *dcm*, o*mpT*, *hsdS*(*r*B-*mB*-), *gal*λ(DE3) (and two uncharacterised mutations)	C43 is a BL21(DE3) double mutant, selected for overexpression of toxic proteins; can tolerate inclusion bodies at a high level without toxic effect observed for BL21(DE3).	(*39*)
NM554	*recA13*, *araD139*, Δ*(ara-leu)7696*, Δ*(lac)17A*, *galU*, *galK*, *hsdR, rpsl, mcrA, mcrB*	*recA* mutant, usually used for plasmid host for cloning.	(*40*)
821	*F-, thr-1, araC14, leuB6(Am), DE(gpt-proA)62, lacY1, tsx-33, qsr'-0, glnV44(AS), galK2(Oc), LAM-, Rac-0, hisG4(Oc), rfbD1, mg1-51, gshA2, rpsL31, kdgK51, xy1A5, mtl-1, argE3(Oc), thi-1.*	*gshA2* mutation blocks glutathione biosynthesis	(*41*)

Table 2.2 *E. coli* strain genotypes

E. coli strain 821

A significant problem encountered in intein fusion protein expression was intracellular cleavage of the fusion protein prior to isolation. A likely explanation of this cleavage is reaction with an intracellular thiol which reacts with the protein thioester during cell growth. Glutathione (*42*), a tripeptide containing a free sulfhydryl group, is the principle free thiol within many cell types. It is responsible for maintaining the balance of the intracellular electrochemical potential and functions as the reducing agent for glutaredoxin during deoxyribonucleotide biosynthesis.

Glutathione is present at high levels in cells (up to 5 mM), and cycles between the reduced and oxidised forms (in which two tripeptides are linked through a disulphide bond), with the disulphide being reduced by Glutathione reductase. The glutathione biosynthetic pathway (*43*) (Fig. 2.6) involves two steps, a condensation reaction between cysteine and glutamate catalysed by Glutamate-cysteine ligase (*gshA*) followed by condensation of the dipeptide with glycine catalysed by Glutathione synthase (*gshB*).

As *E. coli* strain 821 has a *gshA*$^-$ mutation that blocks the biosynthesis of glutathione, expression of target-intein fusion proteins in this strain may reduce the unwanted thioester cleavage, increasing the efficiency with which the intact target-intein protein can be isolated.

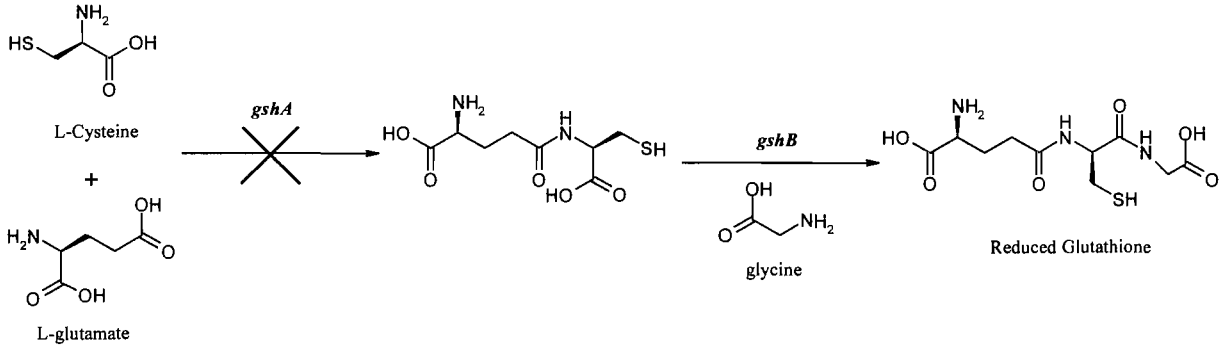


Figure 2.6 The glutathione biosynthetic pathway

2.3.1 SDS-PAGE gel analysis

An example of a typical expression study result is shown in Fig. 2.7 with Fpr-intein, small scale nickel affinity purification shown (crude lysate not shown). Often the target-intein thioester has been cleaved by an intracellular thiol prior to isolation. In this case, target-intein fusion protein cleavage is significant, with two major bands being observed. Ni affinity purified sample contains both the target-intein and intein proteins as the His_6-tag is located at the intein C-terminus. The proportion of cleavage of the target-intein thioester was judged by manual band integration using Syngene Genetools software unless otherwise stated.

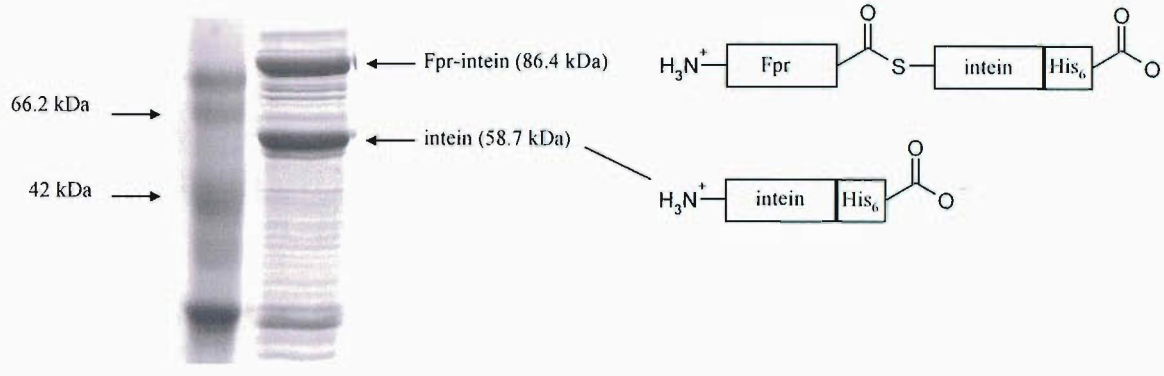


Figure 2.7 Identification of target-intein expression. Example is small scale Ni affinity purified Fpr-intein.

2.3.2 Expression of BioB-intein

The expression of BioB-intein was investigated in six strains of *E. coli*. Cultures were grown in 100 mL of 2-YT medium and the expression compared in cleared lysates and after small scale Ni affinity purification, with analysis by SDS-PAGE (10 %, Fig. 2.8). Expression was on a 100 mL culture scale. BioB-intein has been overexpressed in all six strains (Fig. 2.8 A), the purified samples (B) show significant quantities of cleaved intein. It is clear that strain 821 yields the largest quantity of intact fusion protein.

BioB-intein was overexpressed on a large scale in BL21(DE3).

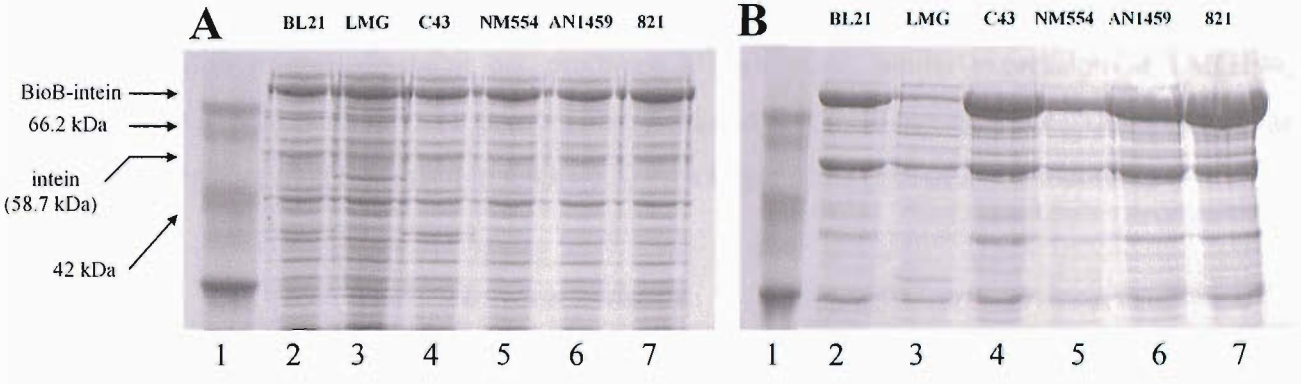


Figure 2.8 BioB-intein expression study, (A) cleared lysates. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = expression in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression in strain 821. (B) BioB-intein expression study, Ni affinity purified protein. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = expression in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression in strain 821

2.3.3 Expression of LipA-intein

LipA-intein displays quite different properties to the other fusion proteins, as it undergoes relatively little cleavage prior to isolation. Somewhat surprisingly, strain 821 shows a poor yield of protein and a relatively high proportion of cleavage, ~30 %. This elevated cleavage may be caused by the overproduction of the dithiol lipoic acid. In the absence of glutathione, lipoic acid has been shown to take on the role of the principle intracellular redox component (*44-46*). Lipoic acid biosynthesis may have been increased by strain 821 to compensate for the lack of glutathione, this coupled with the overexpression of lipoic acid synthase might support high levels within the cells.

Expression in AN1459 resulted in 60 % thioester cleavage, BL21(DE3) showed little fusion protein cleavage but produced a poor yield, whilst expression in LMG194, C43(DE3), and NM554 gave a good yield and low (<20 %) cleavage. LipA-intein was therefore overexpressed in C43(DE3) on a larger scale (5 L of 2-YT medium).

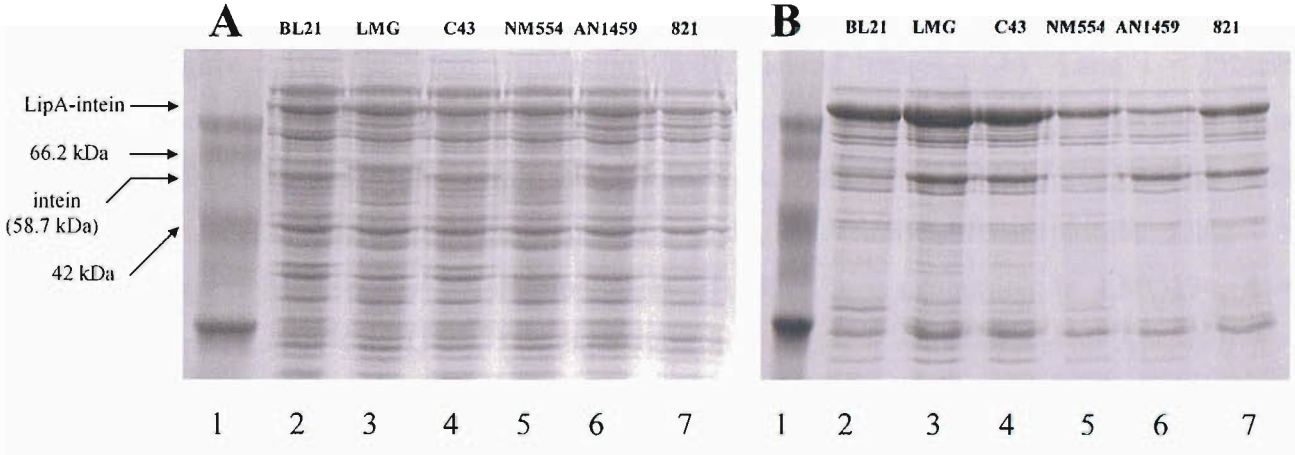


Figure 2.9 LipA-intein expression study, (A) cleared lysates. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = LipA-intein in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression in strain 821. (B) LipA-intein expression study, Ni affinity purified protein. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = expression in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression in strain 821.

2.3.4 Expression of Fpr-intein

Expression in strain 821 (Fig. 2.10) resulted in the highest yield of intact fusion protein and lowest fraction of cleaved Fpr-intein. In strains BL21(DE3), LMG194, C43(DE3), NM554 and AN1459 expression levels were comparable, but the fraction of cleaved protein was 40 - 60 % in comparison with 30 % in strain 821. Fpr-intein was therefore overexpressed on a large scale in strain 821.

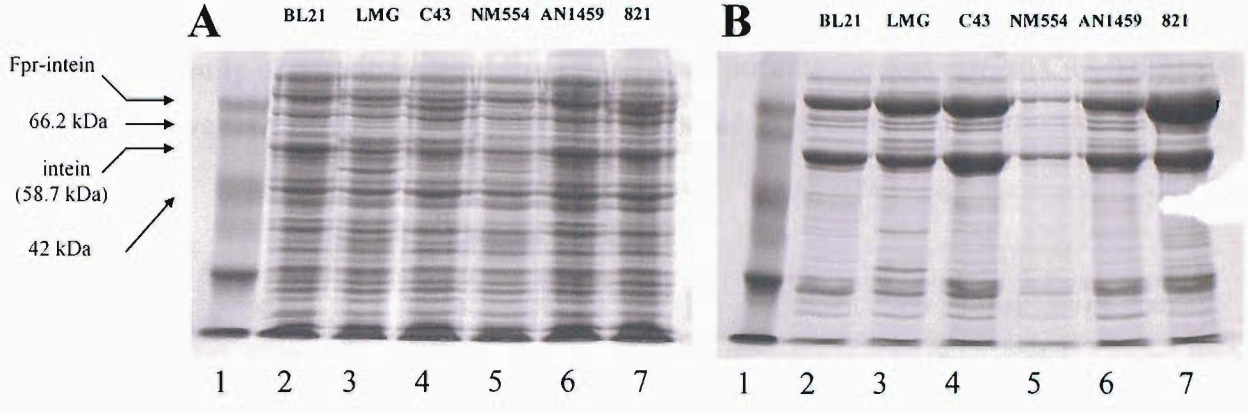


Figure 2.10 Fpr-intein Expression Study, cleared lysates. (A) Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = expression in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression in strain 821. (B) Fpr-intein expression study, Ni affinity purified protein Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in LMG194, 4 = expression in C43(DE3), 5 = expression in NM554, 6 = expression in AN1459, 7 = expression strain 821.

2.3.5 Expression of SolA-intein

SolA-intein expression was investigated in *E. coli* strains BL21(DE3), C43(DE3), NM554, AN1459 and 821. SolA-intein is expressed (Fig. 2.11) in all five strains, with high levels of cleaved protein isolated after small scale Ni affinity purification in all except for NM554. A high proportion of cleavage is observed in BL21(DE3) (70 %), AN1459 (60 %) and strain 821 (60 %).

Although the lowest proportion of cleavage resulted from expression in NM554 (25 %), the quantity of protein isolated from small scale purification was low, so C43 (DE3) (20 % cleavage) was used as the host for large scale expression and purification of SolA-intein.

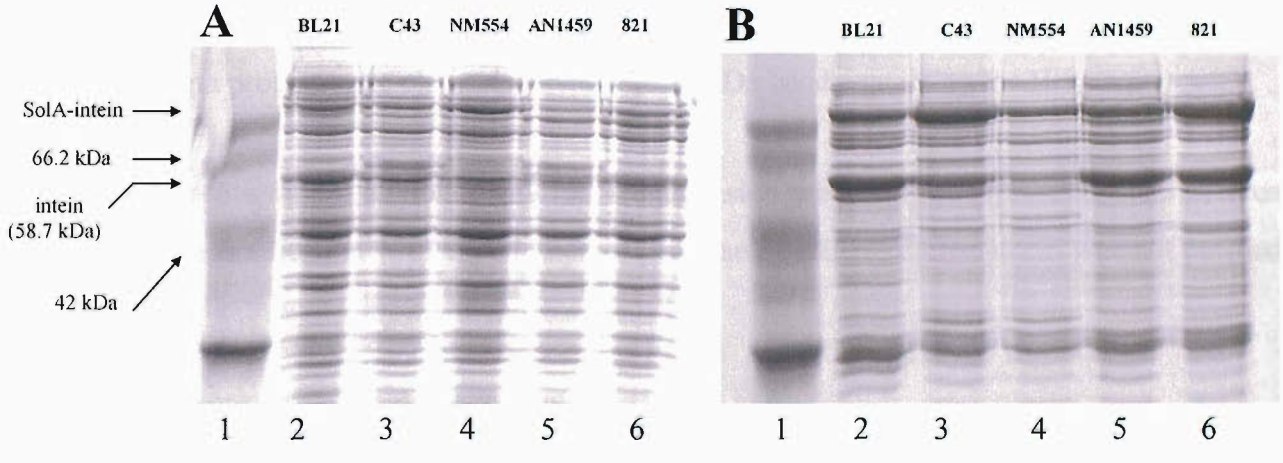


Figure 2.11 SolA-intein Expression Study. (A) cleared lysates. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in C43(DE3), 4 = expression in NM554, 5 = expression in AN1459, 6 = expression in strain 821. (B) SolA-intein expression study, Ni affinity purified protein. Lane 1 = Protein Marker, 2 = expression in BL21(DE3), 3 = expression in C43(DE3), 4 = expression in NM554, 5 = expression in AN1459, 6 = expression in strain 821.

2.3.6 Expression of PKBα-intein

Plasmid pRJW/3953/77 (PKBα-intein in pBADHisA) was transformed into *E. coli* strains BL21(DE3), C43(DE3), LMG194 and strain 821. Small scale expression of the PKBα-intein fusion protein was investigated. Expression in BL21(DE3) is representative of the results observed. The expected mass of PKBα-intein is 115.9 kDa. Analysis of the cleared cell lysates and small scale nickel affinity purified protein revealed no intact fusion protein. In both the supernatant and Ni purified samples (Fig. 2.12A) an intense band that may be the cleaved intein is present. His-staining (a fluorescent dye which strongly binds to the His_6-tag) showed this band to strongly fluoresce, supporting this hypothesis. A strong band of size ~ 62 kDa is also seen in both supernatant and Ni purified samples (Fig. 2.12A,B,C) which also binds to the His-stain, suggesting that this band is also His_6-tagged. It seems possible that the PKBα-intein contains a sequence that may be cleaved by an *E. coli* protease, causing the full length protein to be cleaved, leaving the 62 kDa band containing the His_6-tag.

Due to a failure to isolate any full length fusion protein from expression studies, even upon purification with the addition of protease inhibitors, PKBα-intein was not purified on a large scale.

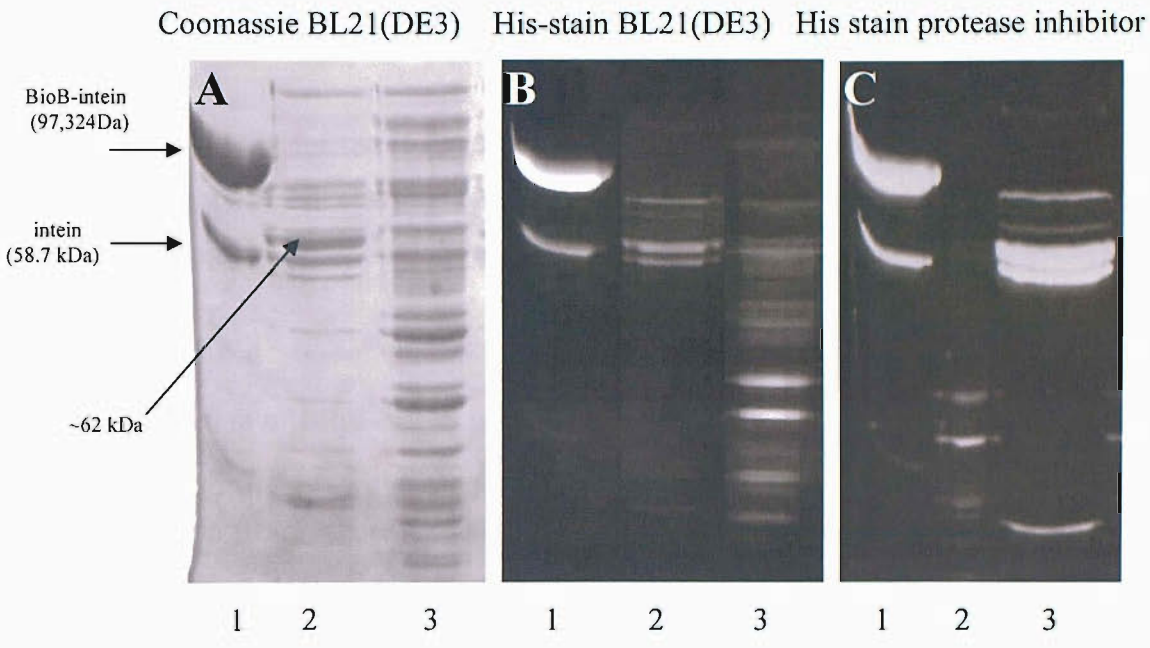


Figure 2.12 10 % SDS-PAGE analysis of small scale expression of PKBα-intein in BL21 (DE3). (A) Lane 1 = BioB-intein reference sample, 2 = Small scale Ni affinity purification sample, 3 = cleared lysate. (B) His stained gel visualised under UV. Lane 1 = BioB-intein reference sample, 2 = cleared lysate, 3 = Small scale Ni affinity purification sample. (C) 10 % SDS-PAGE analysis of small scale expression of PKBα-intein in BL21 (DE3), lysis buffer conditions contained protease inhibitors. His stained gel visualised under UV. Lane 1 = BioB-intein reference sample, 2 = cleared lysate, 3 = Small scale Ni affinity purification sample.

2.4 Purification of target-intein fusion proteins

The general strategy for the isolation of target-intein fusion proteins was to apply the cleared lysate to a Ni charged chelating sepharose column, elute with imidazole, concentrate the most pure fractions and to further purify the concentrate by gel filtration chromatography.

2.4.1 Purification of BioB-intein

The BioB-intein fusion protein was eluted early in the imidazole gradient from the Ni affinity column, at 100 mM imidazole concentration. This indicates a relatively low affinity, which may be due to the large size of the fusion protein hindering the affinity interaction. This step yielded 140 mg of >70 % pure protein as approximated by eye. Fractions 4 – 20 were analysed by SDS-PAGE (Fig. 2.13). Fractions 10 – 20 were pooled and concentrated. Fractions 4-9 contained large quantities of cellular proteins (data not shown).

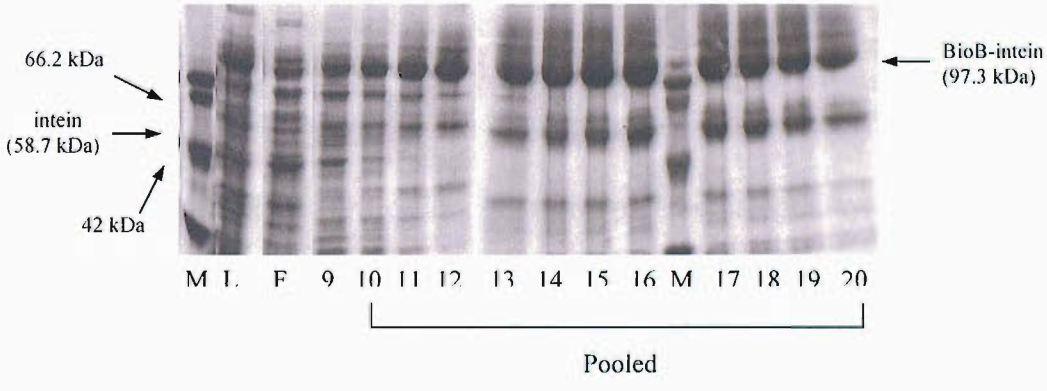


Figure 2.13 BioB-intein Purification, Ni affinity chromatography; Lane M = protein marker, L = cleared lysate, F = flow through. Numbered lanes represent fraction numbers.

Early attempts at the purification of BioB-intein were unsuccessful due to problems encountered during the gel filtration step. The buffer initially used for gel filtration contained no salt and resulted in protein precipitation, but 25 mg of fusion protein was recovered for use to test the stability of the protein. A simple pH and salt concentration

screen indicated that the protein did not precipitate from buffer at pH 7.0 containing salt (500 mM).

Concentrated protein (6 mL, 25 mg/mL) from the Ni affinity purification step was applied to an S-200 gel filtration column and the protein eluted without precipitation, fractions 33 – 48 were light brown in colour. Fractions were analysed by SDS-PAGE (Fig. 2.14) and the purest stored at -80 °C, without concentration or pooling. This yielded 60 mg of >80 % pure BioB-intein fusion protein (as estimated by eye), the major impurities being cleavage products. The retention volume of BioB-intein was 300 mL.

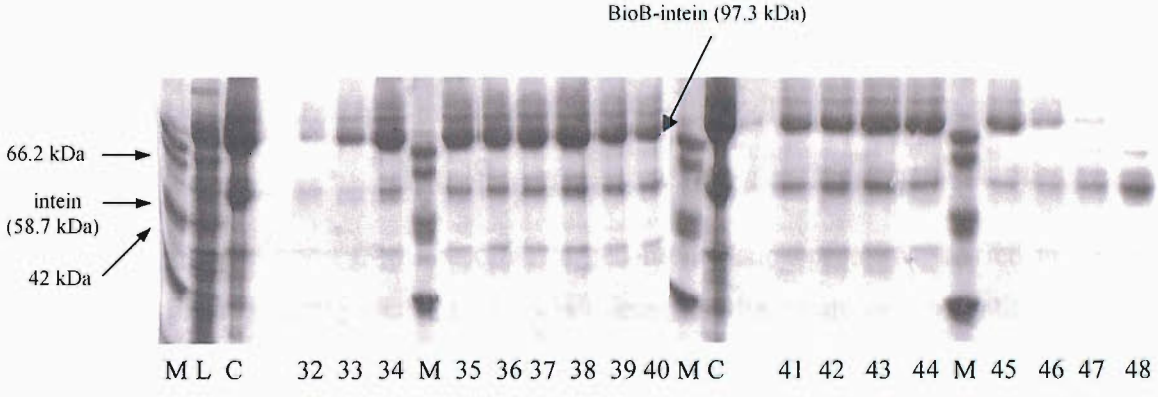


Figure 2.14 BioB-intein Purification, gel filtration chromatography; Lane M = protein marker, L = cleared lysate, C = concentrate post Ni affinity. Numbered lanes represent fraction numbers.

	Total protein (mg)	Purity (%)	BioB-intein (mg)	Yield (%)
Crude	1800	10	180	100
Ni Eluted	200	70	140	78
S-200 Eluted	60	95	57	32

Table 2.3 BioB-intein purification summary

2.4.2 Purification of LipA-intein

The LipA-intein fusion protein was eluted at an imidazole concentration of 125 mM. This purification step was efficient, yielding 240 mg of highly purified LipA-intein. Fractions 13 – 23 were analysed by SDS-PAGE (10 %, Fig. 2.15A), fractions 16 – 22 were pooled and concentrated. The protein could not be concentrated beyond 12 mg / mL without protein loss through precipitation.

Concentrated protein (20 mL, 12 mg / mL) was applied to a pre-equilibrated S-200 gel filtration column and the protein eluted. Fractions 34 – 49 were light brown in colour and were analysed by SDS-PAGE (Fig. 2.15B). Most pure fractions were not pooled but were directly stored at -80 °C. This purification yielded 140 mg of >95 % (as estimated by eye) pure LipA-intein. Retention volume of LipA-intein = 360 mL.

LipA-intein is set apart from the other target-intein fusion proteins purified in terms of its stability. Very little cleaved intein is present in the lysate or is co-eluted with the fusion protein at the Ni affinity step. This stability suggests that the reactivity of the fusion protein thioester may be dependant upon the nature of the target protein. The folding of the target may hinder both the initial formation of the thioester and the attack of a nucleophile after its formation.

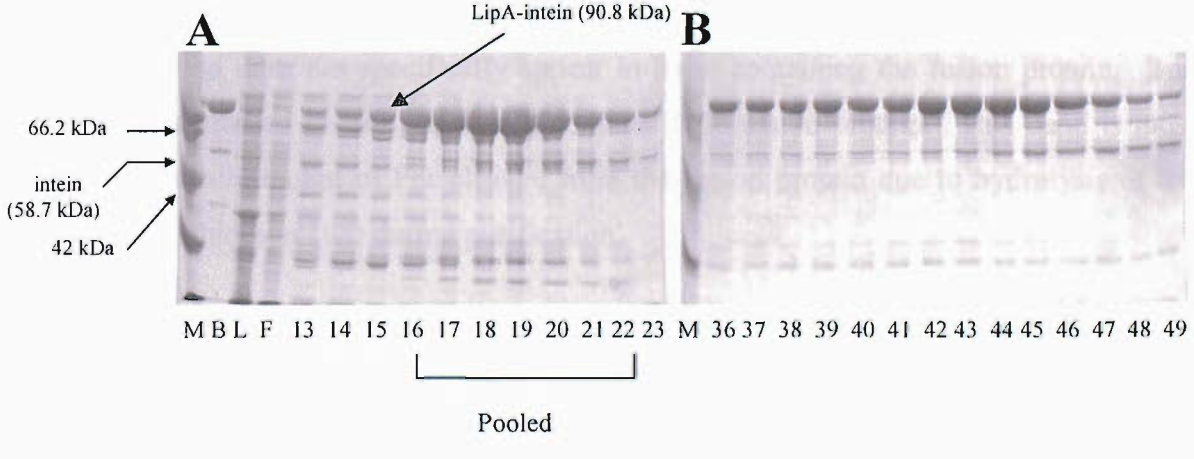


Figure 2.15 LipA-intein Purification; (A) Ni affinity chromatography, (B) gel filtration chromatography. M = protein marker, B = BioB-intein reference, L = cleared lysate, F = flow through. Numbered lanes indicate fraction numbers.

	Total protein (mg)	Purity (%)	LipA-intein (mg)	Yield (%)
Crude	5000	5	250	100
Ni Eluted	240	95	230	90
S-200 Eluted	140	95	130	54

Table 2.4 LipA-intein purification summary

2.4.3 Purification of Fpr-intein

The Fpr-intein fusion protein was eluted from the Ni affinity column by imidazole at a concentration of 100 mM. This purification step yielded 240 mg >50 % pure protein as estimated by eye. Once again, a large quantity of intein co-eluted with the fusion protein, indicating a significant degree of cleavage prior to purification. Fractions 8 – 19 (Fig. 2.16 A) were pooled and concentrated to 3.5 mL. This sample of concentrated protein (3.5 mL, 68 mg/mL) from the Ni affinity purification step was applied to an S-75 gel filtration column and the protein eluted at a retention volume of 360 mL. Fractions 30 – 33 (Fig. 2.16B) were deep yellow in colour and were immediately stored at -80 °C.

Considering Fig. 2.16B, fractions 31 and 37, fraction 31 contains Fpr-intein, intein and two bands at ~ 30 kDa. The loading buffer used to run these gels contains DTT, which is known to efficiently cleave target-intein fusion proteins (*47*), so this band, which appears only in lanes containing fusion protein is DTT cleaved Fpr. The slightly smaller band does not specifically appear in lanes containing the fusion protein. It is most concentrated in lane 12, which contains little fusion protein. This band is most likely a small amount of Fpr cleaved from the fusion protein due to hydrolysis of the Fpr-intein thioester (*48*) during purification.

This preparation yielded 120 mg of >70 % pure Fpr-intein.

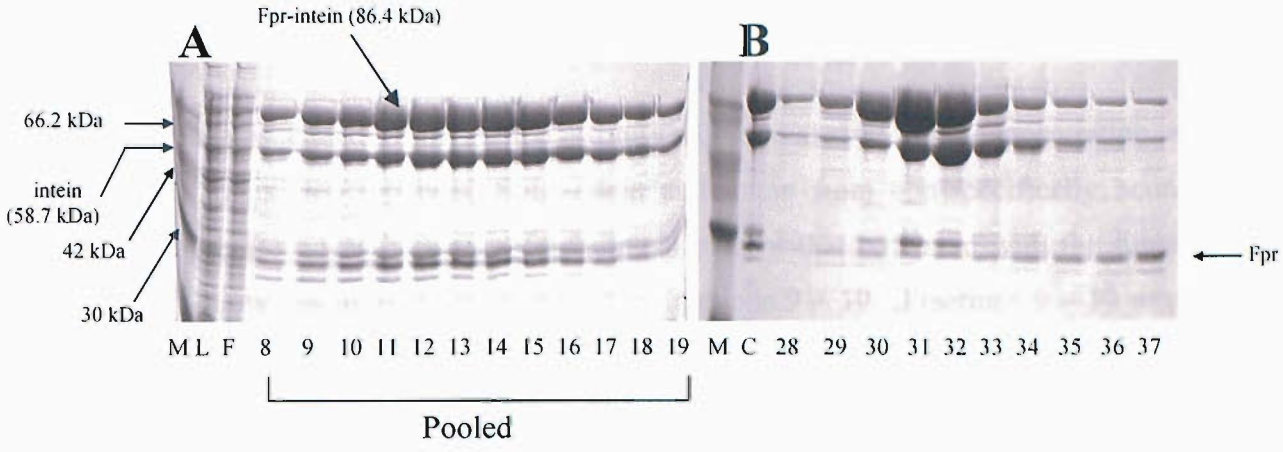


Figure 2.16 Fpr-intein Purification. (A) Ni affinity chromatography, (B) = gel filtration chromatography. M = protein marker, L = cleared lysate, F = flow through, C = concentrated protein from Ni affinity, numbered lanes indicate fraction numbers.

	Total protein (mg)	Purity (%)	Fpr-intein (mg)	Yield (%)
Crude	5000	3	150	100
Ni Eluted	240	~60	140	90
S-200 Eluted	180	~70	120	85

Table 2.5 Fpr-intein purification summary

2.4.4 Purification of SolA-intein

SolA-intein was eluted from the Ni affinity column at 140 mM concentration of imidazole. This step separated SolA-intein and intein from non-specifically bound proteins well, there was however >50 % intein present that co eluted with the fusion protein. Yellow colouration was observed in fractions 9 – 19. Fractions 9 – 19 were analysed by SDS-PAGE (10 %, Fig 2.17A), fractions 12-16 were combined and concentrated. Yield ~ 60 mg 50 % pure SolA-intein.

An aliquot of concentrated protein (3 mL, 20 mg / mL) from the Ni affinity purification step was applied to an S-200 gel filtration column and the protein eluted at retention volume 400 mL. Fractions 43 – 56 (10 % SDS-PAGE, Fig. 2.17B) were light yellow in colour and were immediately stored at -80 °C. This preparation yielded 55 mg >60 % pure protein.

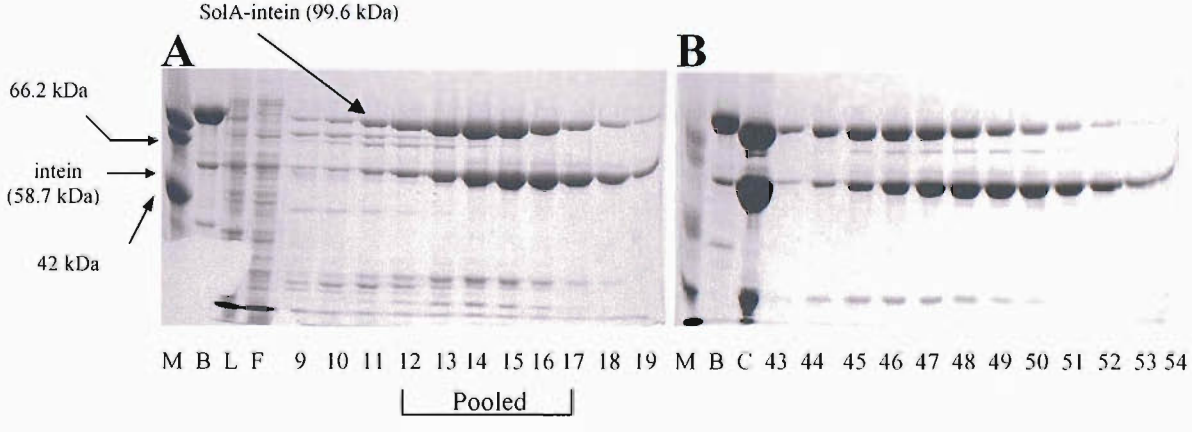


Figure 2.17 SolA-intein Purification, (A) Ni affinity purification, (B) gel filtration chromatography. M = protein marker, B = BioB-intein reference, L = cleared lysate, F = flow through, C = concentrate post Ni affinity. Numbered lanes indicate fraction numbers.

	Total protein (mg)	Purity (%)	SolA-intein (mg)	Yield (%)
Crude	1800	5	90	100
Ni Eluted	60	50	30	33
S-200 Eluted	55	60	30	33

Table 2.6 SolA-intein purification summary

2.5 Summary and Conclusions

Plasmid pRJW/2960/82 has been constructed for the simple assembly of plasmids for the expression of target-intein fusion proteins under the control of an arabinose inducible promoter. Plasmids pRJW/2960/88, pRJW/2960/89, pRJW/3104/40, pRJW/3104/100 and pRJW/3953/77 were constructed from pRJW/2960/82 for the expression of *E. coli* Fpr-, BioB- and SolA-, and *Sulfolobus solfataricus* LipA-intein and *Homo sapiens* PKBα-intein fusion proteins respectively.

Several fusion proteins were expressed at high levels from the expression plasmids in various *E. coli* strains. Protein isolated from the Ni affinity step in each case contains a mixture of the full length fusion protein and intein domain, possibly due to *in vivo* reaction with abundant intracellular thiols such as glutathione. The selection of *E. coli* strain for overexpression was empirical, but often expression of fusion proteins in *E. coli* strain 821, which is unable to produce glutathione, yielded the highest proportion of full length protein. The expression of full length PKBα-intein was unsuccessful, this protein may have been cleaved either by an intracellular thiol or an *E. coli* protease.

2.6 References

1. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

2. Perler, F. B., and Adam, E. (2000) *Current Opinion in Biotechnology 11*, 377-383.

3. Cammack, R. (1992) *Advances in Inorganic Chemistry 38*, 281-322.

4. Marsh, E. N. G., Patwardhan, A., and Huhta, M. S. (2004) *Bioorganic Chemistry 32*, 326-340.

5. Layer, G., Heinz, D. W., Jahn, D., and Schubert, W. D. (2004) *Current Opinion in Chemical Biology 8*, 468-476.

6. Heering, H. A., Weiner, J. H., and Armstrong, F. A. (1997) *Journal of the American Chemical Society 119*, 11628-11638.

7. Jin, W., Wollenberger, U., Bernhardt, R., Stocklein, W. F. M., and Scheller, F. W. (1998) *Bioelectrochemistry and Bioenergetics 47*, 75-79.

8. Sanyal, I., Cohen, G., and Flint, D. H. (1994) *Biochemistry 33*, 3625-3631.

9. Miller, J. R., Busby, R. W., Jordan, S. W., Cheek, J., Henshaw, T. F., Ashley, G. W., Broderick, J. B., Cronan, J. E., and Marletta, M. A. (2000) *Biochemistry 39*, 15166-15178.

10. Vandenboom, T. J., Reed, K. E., and Cronan, J. E. (1991) *Journal of Bacteriology 173*, 6411-6420.

11. Cicchillo, R. M., Lee, K. H., Baleanu-Gogonea, C., Nesbitt, N. M., Krebs, C., and Booker, S. J. (2004) *Biochemistry 43*, 11770-11781.

12. Cicchillo, R. M., and Booker, S. J. (2005) *Journal of the American Chemical Society 127*, 2860-2861.

13. Ollagnier-de Choudens, S., Sanakis, Y., Hewitson, K. S., Roach, P., Baldwin, J. E., Munck, E., and Fontecave, M. (2000) *Biochemistry 39*, 4165-4173.

14. Bianchi, V., Reichard, P., Eliasson, R., Pontis, E., Krook, M., Jornvall, H., and Haggardljungouist, E. (1993) *Journal of Bacteriology 175*, 1590-1595.

15. Sanyal, I., Gibson, K. J., and Flint, D. H. (1996) *Archives of Biochemistry and Biophysics 326*, 48-56.

16. Blaschkowski, H. P., Neuer, G., Ludwig-Festl, M., and Knappe, J. (1982) *European Journal of Biochemistry 123*, 563-569.

17. Fujii, K., and Huennkens, F. M. (1974) *Journal of Biological Chemistry 249*, 6745-6753.

18. Ingelman, M., Bianchi, V., and Eklund, H. (1997) *Journal of Molecular Biology 268*, 147-157.

19. Pueyo, J. J., Gomezmoreno, C., and Mayhew, S. G. (1991) *European Journal of Biochemistry 202*, 1065-1071.

20. McIver, L., Leadbeater, C., Campopiano, D. J., Baxter, R. L., Daff, S. N., Chapman, S. K., and Munro, A. W. (1998) *European Journal of Biochemistry 257*, 577-585.

21. Madoz-Gurpide, J., Abad, J. M., Fernandez-Recio, J., Velez, M., Vazquez, L., Gomez-Moreno, C., and Fernandez, V. M. (2000) *Journal of the American Chemical Society 122*, 9808-9817.

22. Wagner, M. A., Khanna, P., and Jorns, M. S. (1999) *Biochemistry 38*, 5588-5595.

23. Khanna, P., and Jorns, M. S. (2001) *Biochemistry 40*, 1441-1450.

24. Wang, J., Ciszewski, A., and Naser, N. (1992) *Electroanalysis 4*, 777-782.

25. Kacaniklic, V., Johansson, K., Markovarga, G., Gorton, L., Jonssonpettersson, G., and Csoregi, E. (1994) *Electroanalysis 6*, 381-390.

26. Sakslund, H., Wang, J., Lu, F., and Hammerich, O. (1995) *Journal of Electroanalytical Chemistry 397*, 149-155.

27. Evans, S. A. G., Elliott, J. M., Andrews, L. M., Bartlett, P. N., Doyle, P. J., and Denuault, G. (2002) *Analytical Chemistry 74*, 1322-1326.

28. Brazil, D. P., and Hemmings, B. A. (2001) *Trends in Biochemical Sciences 26*, 657-664.

29. Alessi, D. R., Caudwell, F. B., Andjelkovic, M., Hemmings, B. A. and Cohen, P. (1996) *Febs Letters 399*, 333–338.

30. Stiles, B., Gilman, V., Khanzenzon, N., Lesche, R., Li, A., Qiao, R., Liu, X., and Wu, H. (2002) *Molecular and Cell Biology 22*, 3842-3851.

31. Evans, T. C., and Xu, M. Q. (1999) *Biopolymers 51*, 333-342.

32. Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., and Stevens, T. H. (1990) *Science 250*, 651-657.

33. Chong, S. R., Montello, G. E., Zhang, A. H., Cantor, E. J., Liao, W., Xu, M. Q., and Benner, J. (1998) *Nucleic Acids Research 26*, 5109-5115.

34. Chong, S. R., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998) *Journal of Biological Chemistry 273*, 10567-10577.

35. Cline, J., Braman, J. C., and Hogrefe, H. H. (1996) *Nucleic Acids Research 24*, 3546-3551.

36. Hyman, H. C., and Honigman, A. (1986) *Journal of Molecular Biology 189*, 131-141.

37. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) *Journal of Bacteriology 177*, 4121-4130.

38. Elvin, C. M., Dixon, N. E., and Rosenberg, H. (1986) *Molecular & General Genetics 204*, 477-484.

39. Miroux, B., and Walker, J. E. (1996) *Journal of Molecular Biology 260*, 289-298.

40. Raleigh, E. A., Murray, N. E., Revel, H., Blumenthal, R. M., Westaway, D., Reith, A. D., Rigby, P. W. J., Elhai, J., and Hanahan, D. (1988) *Nucleic Acids Research 16*, 1563-1575.

41. Apontoweil, P., and Behrends, W. (1975) *Molecular & General Genetics 141*, 91-95.

42. Berg, J. M., Tymoczco, J. L., Stryer, L. (2002) *Biochemistry*, 5 ed., W. H. Freeman & co.

43. Apontoweil P, B. W. (1975) *Biochimica and Biophysica Acta. 1*, 1-9.

44. Mirandavizuete, A., Martinezgalisteo, E., Aslund, F., Lopezbarea, J., Pueyo, C., and Holmgren, A. (1994) *Journal of Biological Chemistry 269*, 16631-16637.

45. MirandaVizuete, A., RodriguezAriza, A., Toribio, F., Holmgren, A., LopezBarea, J., and Pueyo, C. (1996) *Journal of Biological Chemistry 271*, 19099-19103.

46. Prieto-Alamo, M. J., Jurado, J., Gallardo-Madueno, R., Monje-Casas, F., Holmgren, A., and Pueyo, C. (2000) *Journal of Biological Chemistry 275*, 13398-13405.

47. Wu, C., Seitz, P. K., and Falzon, M. (2000) *Molecular and Cellular Endocrinology 170*, 163-174.

48. Nichols, N. M., Benner, J. S., Martin, D. D., and Evans, T. C. (2003) *Biochemistry 42*, 5301-5311.

Chapter 3:- Reactivity of the BioB-intein thioester towards nucleophiles and fluorescent label synthesis

3.1 Introduction

Fluorescent labelling via intein mediated ligation requires a fluorescent molecule containing a nucleophilic moiety, which will react with the target-intein thioester. This reaction results in covalent fluorescent labelling of the target protein domains.

With the objective of developing a highly efficient labelling method, we set out to measure the extent to which a series of nucleophiles reacted with a model system, BioB-intein fusion protein. The second phase of the project would apply the most efficient nucleophile, as a conjugate with a fluorescent group for site specific labelling of target proteins. The label should be small (300 – 700 Da) (*1*), minimising the effect upon the folding (*2*), localisation (*3*) and biochemical properties (*4*) of the target protein, and specifically react with the target-intein thioester.

Literature examples of intein mediated ligation utilise an N-terminal cysteine based peptide derivatised with the required modification. Examples include conjugation with a fluorescent label (*5*), carbohydrates (*6*) or an oligonulcleotide (*7*) for specific cleavage of the target-intein thioester. Several low molecular weight thiols (i.e. β-mercaptoethanol (*8*), DTT (*9*) and MESA (*10*)) have been utilised to react with the target-intein thioester, but a wider study of alternative nucleophiles has not been undertaken. One of the first steps in developing a labelling method was therefore to study the reactivity of the BioB-intein thioester with a wide range of nucleophiles.

3.2.1 Reaction of α-effect N-nucleophiles with BioB-intein

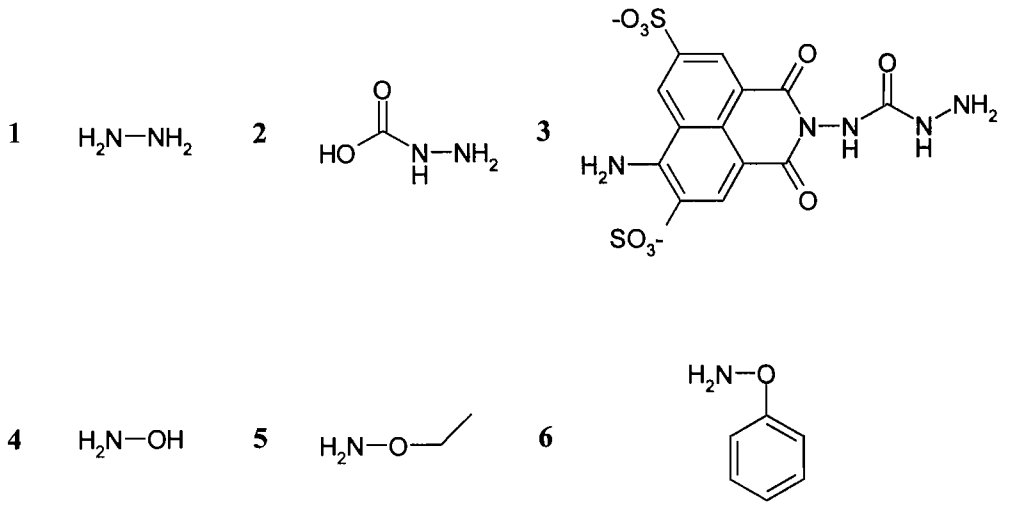


Figure 3.1 α-effect amine nucleophile structures.

The reactivity of the model target-intein thioester towards a range of nucleophiles was investigated. The amine nucleophiles (Fig. 3.1, numbering refers to table 3.1) were reacted with the protein in the absence of any thiol catalyst. The amines used in this experiment were all α-effect nucleophiles (*11, 12*). The α-effect is the enhanced reactivity of nucleophiles that have an unshared pair of electrons on the atom adjacent to the nucleophillic centre, relative to a normal nucleophile of the same approximate basicity.

In these experiments, the nucleophile was added in solution directly to the fusion protein sample. The negative control samples contained BioB-intein to which the appropriate quantity of 50 mM Tris/HCl buffer pH 7.0 had been added. Samples were incubated at room temperature. The concentration of BioB-intein was 0.4 mg / mL, 5 μM, nucleophiles were present in >1000 fold excess.

At pH 7.0, hydrazine failed to react over the full concentration range tested, leaving the fusion protein intact. However, hydroxylamine (Fig. 3.1, lanes 2) reacted even at low concentration (10 mM), unfortunately this reactivity was not observed for substituted hydroxylamines.

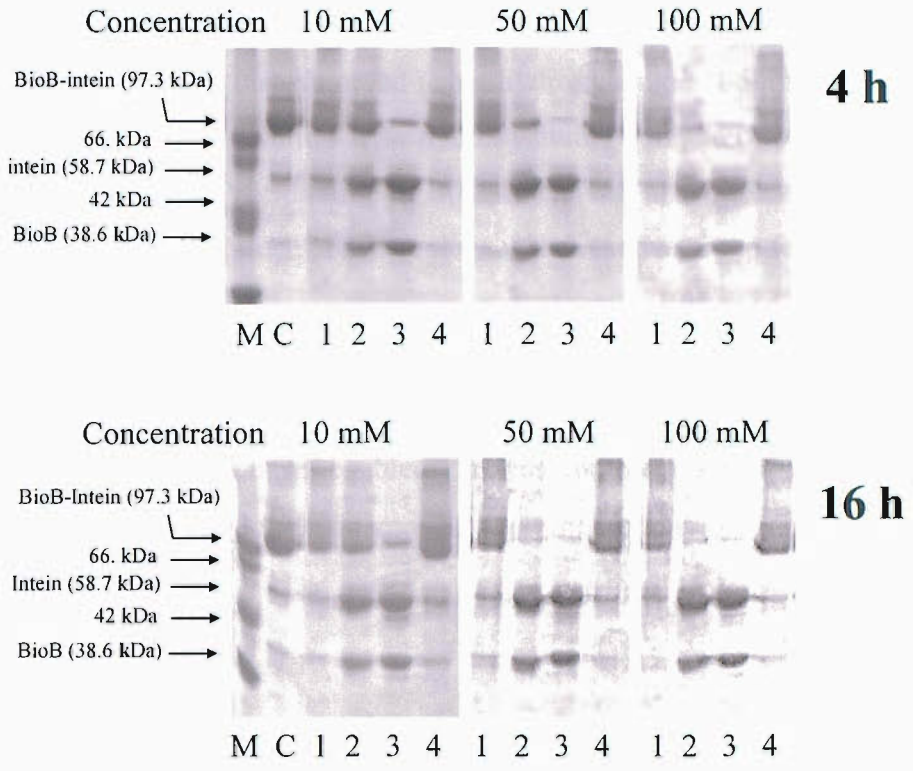


Figure 3.2 Reaction of α-effect N-nucleophiles with BioB-intein, in 50 mM tris buffer pH 7.0, room temperature. 10 % SDS-PAGE analysis, loading buffer contains no DTT. Time 4 or 16 h as indicated, concentration is written above panels. M = protein marker, C = negative control, 1 = Hydrazine, 2 = Hydroxylamine, 3 = DTT, 4 = NaI.

Reagent	pH 7.0 Cleavage	pH 9.0 Cleavage
Hydrazine (**1**)	N	Y
Acetic Hydrazide (**2**)	N	-
Lucifer Yellow (**3**)	N	-
Hydroxylamine (**4**)	Y	-
Ethylhydroxylamine (**5**)	N	-
Benzylhydroxylamine (**6**)	N	-

Table 3.1 Nitrogen Nucleophiles used to cleave the target-intein thioester.

The pKa[1] of hydrazine = 7.9 and therefore at neutral pH hydrazine is predominantly singly protonated, the non reactivity may be due to the loss of the donating lone pair necessary for the α-effect. Repeating the hydrazine cleavage reaction (Fig 3.3) at pH

7.0 (lanes 1), pH 7.5 (lanes 2), pH 8.0 (lanes 3), pH 8.5 (lanes 4) and pH 9.0 (lanes 5) showed hydrazine reaction with the BioB-intein thioester at pH ≥8.0, although this was still less efficient than that observed using hydroxylamine at 50 mM, reaction time 4 h with >90 % complete reaction with hydroxylamine compared to <70 % reaction with hydrazine at pH 8.0. Precipitation of BioB-intein was also observed at pH >8.0, further limiting the potential for use of this moiety as the reactive group for a fluorescent label.

Hydrazine derivatives (data not shown) acetic hydrazide and Lucifer yellow also showed no reaction towards the BioB-intein thioester at pH 7.0 – 9.0. Reaction with Lucifer yellow would have enabled direct fluorescent labelling. In this case the delocalisation of the α-lone pair into the adjacent carbonyl bond is a possible cause of the loss of reactivity.

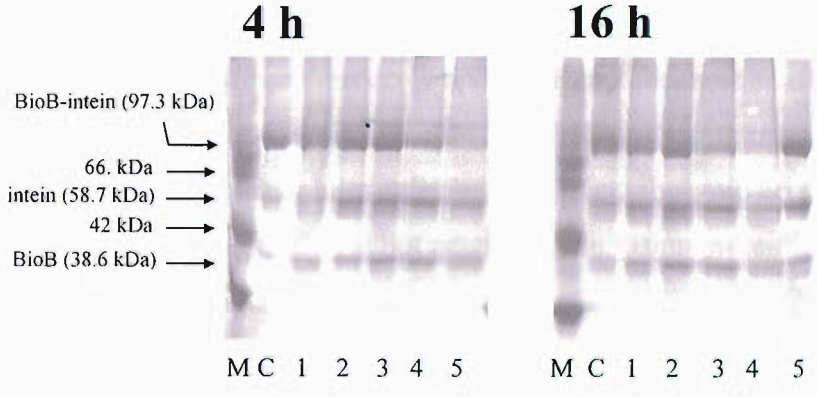


Figure 3.3 Effect of pH upon hydrazine reaction with BioB-intein. Reaction with 50 mM hydrazine, in 50 mM Tris/HCl buffer, room temperature. Analysis by 10 % SDS-PAGE, loading buffer contained no DTT. M = protein marker, C = negative control, pH indicated as follows; 1 = pH 7.0, 2 = pH 7.5, 3 = 8.0, 4 = pH 8.5, 5 = pH 9.0.

3.2.2 Thiol nucleophile BioB-intein cleavage experiments

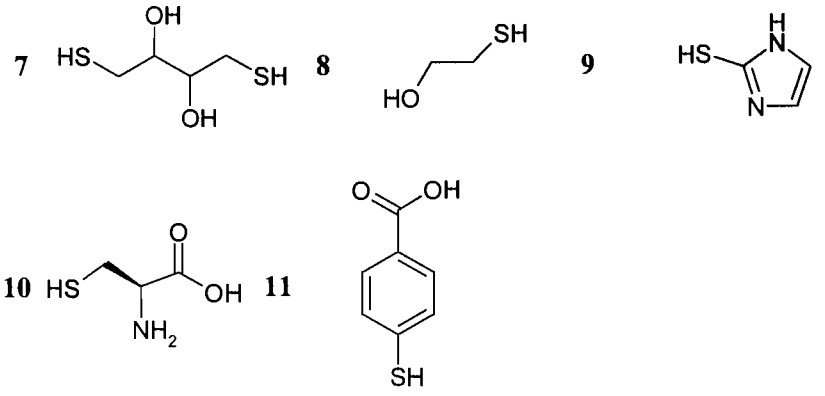


Figure 3.4 Thiol nucleophile structures.

The reactivity of a range of thiol nucleophiles towards the BioB-intein thioester was investigated (Fig. 3.4) and the nucleophiles selected were DTT, L-cysteine, 2-mercapoethanol, 2-mercaptoimidazole and 4-mercaptobenzoic acid (Table 3.2, Fig. 3.4). Reactions were carried out in 50 mM Tris/HCl, pH 7.0 over 4 / 16 h at room temperature. Of the nucleophiles selected, 4-mercaptobenzoic acid was insoluble at pH 7.0, and 2-mercaptoimidazole did not react with the fusion protein.

Over 4 h, 10 mM concentration, cysteine reacted most efficiently with the BioB-intein thioester (Fig. 3.5, lane 4), resulting in a >99 % complete reaction. DTT and 2-mercaptoethanol also reacted well, reacting to >90 % completion. Reaction over 16 h at 10 mM concentration resulted in complete reaction with these compounds. It is clear that thiol nucleophiles react more efficiently than the α-effect amine nucleophiles at neutral pH. Cysteine was observed to react most efficiently with the BioB-intein thioester, so cysteine, the natural substrate was selected as the reactive group to be conjugated to the fluorophore in the design of the fluorescent label.

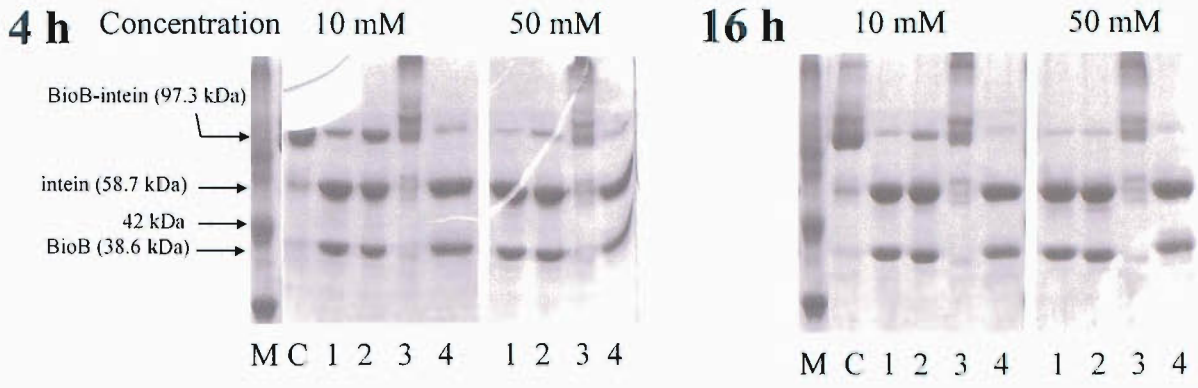


Figure 3.5 Reaction of thiol nucleophiles with the BioB-intein thioester. Reaction in 50 mM Tris/HCl buffer pH 7.0, room temperature. 10 % SDS-PAGE analysis, loading buffer contains no DTT. Lanes M = protein marker, C = negative control, 1 = DTT, 2 = 2-mercaptoethanol, 3 = 2-mercaptoimidazole, 4 = L-cysteine.

Nucleophile	pH 7.0 Cleavage
DTT (**7**)	Y
Mercaptoethanol (**8**)	Y
Mercaptoimidazole (**9**)	N
L-cysteine (**10**)	Y
4-mercaptobenzoic acid (**11**)	Insoluble

Table 3.2 Summary of thiol nucleophile reactivity.

3.2.3 Inorganic nucleophile BioB-intein cleavage experiments

The reaction of a model target-intein fusion protein with iodide and phosphate was investigated (Fig. 3.6). Reaction with iodide was in 50 mM Tris/HCl, pH 7.0 over 16 h. It was observed that iodide did not react with the BioB-intein thioester. Gill ***et al.*** reported that if thioester hydrolysis was conducted in phosphate buffer, that the rate of reaction was greatly increased by the formation of a phosphate intermediate (*13*). It was thought that this may be an alternative to the thiol activated intermediate required for intein mediated ligation. Reaction of the phosphate dianion with the model fusion protein was investigated by exchanging the protein into potassium phosphate buffer at concentrations of 10, 25, 50, 100 and 250 mM, pH 7.0 (Fig. 3.6). No reaction was observed at any phosphate concentration, but further investigation of the formation of intermediates as a route to efficient labelling is discussed in chapter 4.

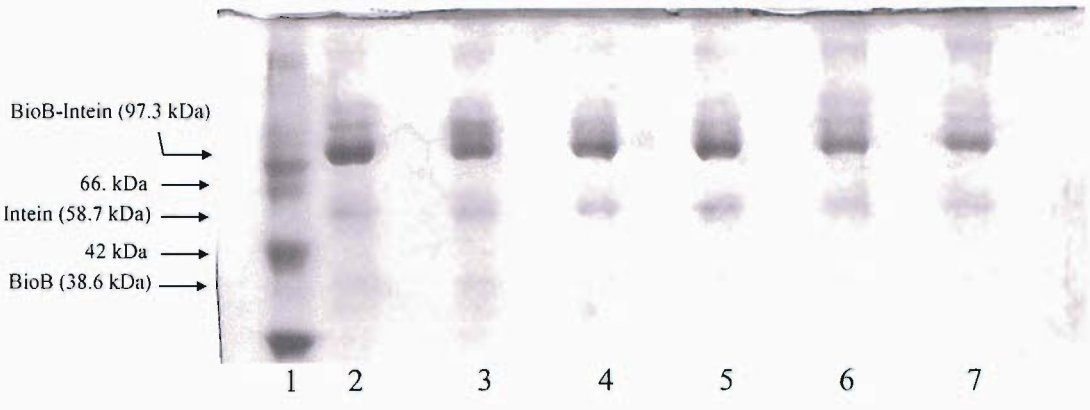


Figure 3.6 Reaction of phosphate buffer with the BioB-intein thioester. Reaction in potassium phosphate buffer, concentration as specified, pH 7.0, room temperature, 16 h reaction time. Loading buffer contained no DTT. Lane 1 = Protein marker, 2 = BioB-intein reference, 3 = 10 mM, 4 = 25 mM, 5 = 50 mM, 6 = 100 mM, 7 = 250 mM.

Nucleophile	pH 7 Cleavage
NaI	N
$KHPO_4$	N

Table 3.3 Summary of the reaction of inorganic nucleophiles with BioB-intein

3.3 Preparation of fluorescent label, N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide (FTEC)

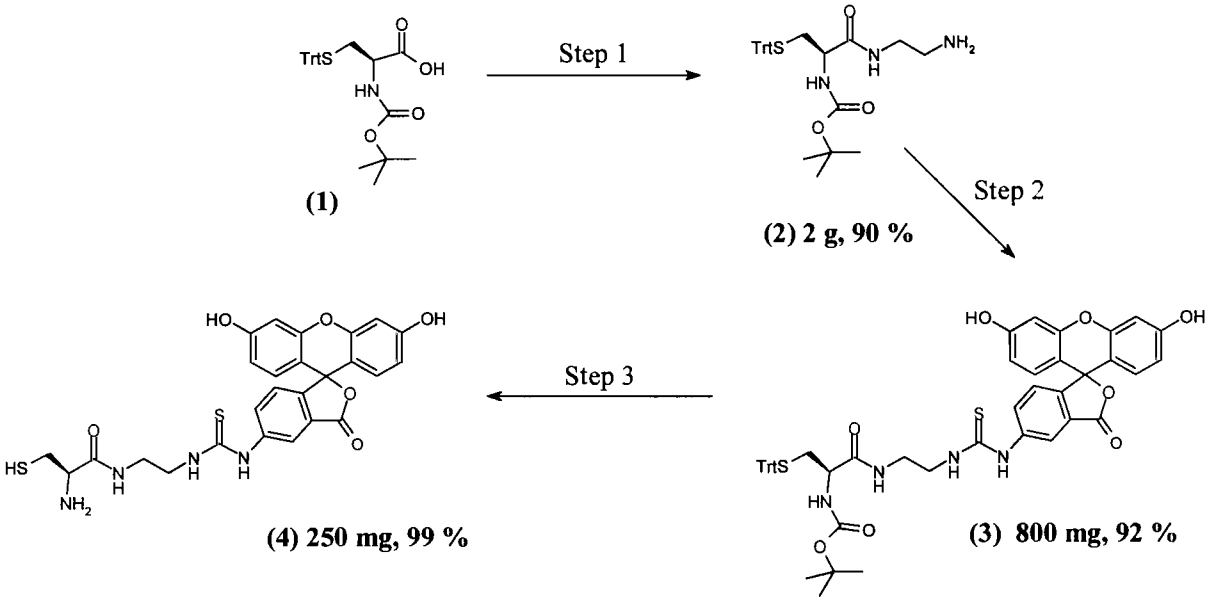


Figure 3.7 FTEC synthesis reaction scheme Step 1 = ethylenediamine, CDI, DMF, 2 h. Step 2 = FITC, methanol, 16 h. Step 3 = 95 % TFA, water, triisopropylsilane.

The investigation of the cleavage of a model target-intein fusion protein showed cysteine to be the most efficient molecule at reacting with the fusion protein thioester. The fluorescent label was therefore based on a conjugate consisting of cysteine coupled to a short ethylenediamine linker, to which fluorescein was conjugated, completing the fluorescent label (Fig. 3.7 **(4)**). A short relatively rigid linker was chosen as they have been shown to create a less flexible label structure, limiting the movement of the fluorophore when attached to the protein. If the fluorophore is free to change its position in relation to the protein, significant noise may be observed in fluorescence measurements such as FRET, caused by constant changes in the distances separating fluorophore and donor molecules (*14*). Fluorescein was selected as the fluorophore as it is a widely used, affordable molecule for protein labelling. Its favourable properties include: high absorptivity, excellent fluorescence quantum yield and good water solubility.

Conditions for the synthesis of FTEC are summarised in Fig. 3.7. In the first step, ethylenediamine was coupled to N-Boc-S-Trityl-L-cysteine (*15*). The resuting amine was then reacted with fluorescein isothiocyanate to form the fully protected thiourea **(3)** which was found to be unstable to silica flash chromatography. This intermediate was prepared on a large scale and stored in this form, small aliquots were deprotected as required. In total, 250 mg of FTEC have been prepared in a near quantitative yield.

3.3.1 HPLC analysis of FTEC

The possibility of FTEC oxidation to its symmetric disulphide was investigated by reverse phase HPLC. Oxidation to the disulphide would leave FTEC unreactive towards the target-intein thioester. Samples were analysed using a reverse phase C18 column. Three samples were analysed; Fig. 3.8A shows analysis of a sample of freshly prepared FTEC which was directly dissolved in 10 % ACN in water containing 0.1 % TFA, the cysteine thiol was stable to oxidation under acidic conditions. Detection at 480 / 260 nm revealed a single peak at retention time 19.7 minutes, thought to be fully reduced FTEC.

The second sample (Fig. 3.8B) was of freshly prepared FTEC which was dissolved in 10 % ACN in water. The sample was left open to air for 30 minutes before acidification with TFA and analysis. This revealed a single peak at retention 21.3 minutes. This is fully oxidised FTEC, which as would be expected for the more hydrophobic disulphide has greater retention.

The third sample (Fig. 3.8C) is freshly prepared FTEC which was degassed in a glovebox under anaerobic conditions (<0.2 ppm O_2) and dissolved in deoxygenated 10 % ACN in water. The sample was stored under anaerobic conditions for 16 h, before acidification with TFA and analysis. The result is a 60:40 mixture of two peaks, one with retention time 19.7 minutes and one 21.3 minutes, corresponding to a 40 % oxidised sample of FTEC, even at low oxygen retention inside the glovebox.

This analysis has shown FTEC to be very sensitive to oxidation to the symmetric disulphide in solution. In only 30 minutes under aerobic conditions complete oxidation takes place, even under anaerobic conditions 40 % oxidation was observed. This may

have been due to some residual oxygen contained within plastic tubes or trapped within the FTEC solid when it was introduced into the glovebox.

It is clear that direct labelling of proteins using FTEC must be carried out under anaerobic conditions, as oxidised FTEC will be non-reactive. N.B. The maximum solubility of FTEC in 50 mM Tris/HCl buffer pH 8.1 is 1 mg / mL.

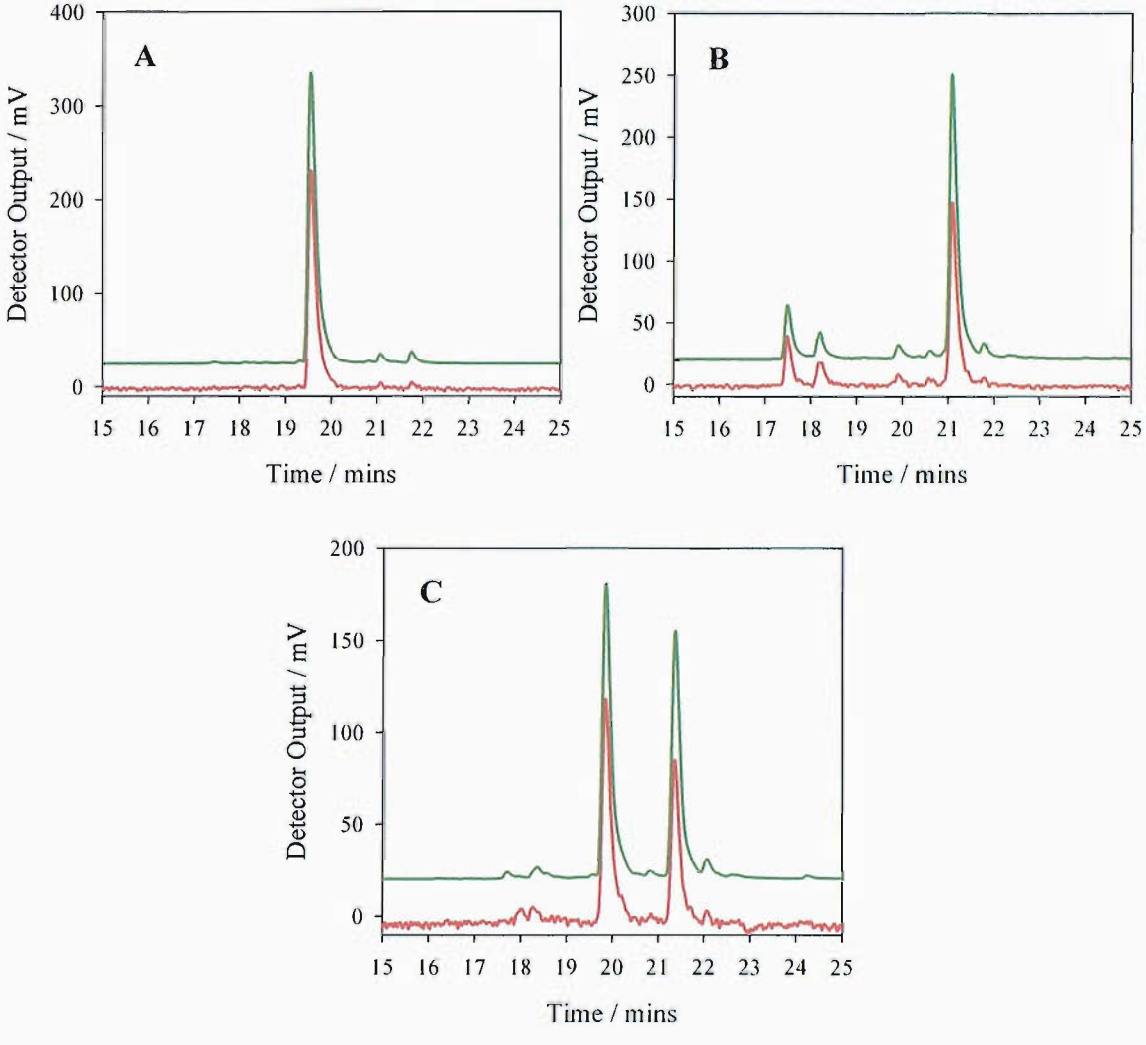


Figure 3.8 HPLC analysis of FTEC, red; λ = 480 nm, raw data × 10 , green; λ = 260 nm, raw data offset by 20. (A) Acidified stored sample of FTEC (reduced), (B) Oxidised sample of FTEC, (C) FTEC post overnight storage in 50 mM pH 8.1 Tris / HCl buffer under anaerobic conditions (<0.2 ppm O_2).

3.4 Optimising conditions for SDS-PAGE analysis of intein mediated fluorescent labelling reactions

SDS-PAGE was anticipated to be a useful technique for the analysis of intein mediated labelling reactions, however, the sample loading buffer contained DTT, with the potential for side reaction with the target-intein thioester. Therefore, the conditions for SDS-PAGE sample preparation were investigated in detail. Ideally, samples would be prepared without the addition of DTT to avoid the side reaction, so initially samples lacked DTT, which resulted (Fig. 3.9) in non-specific fluorescent labelling (*16*), which presumably arises from disulphide bond formation between FTEC and cysteine residues in the proteins.

Conditions used to prepare samples (5 minutes at 80 °C) for SDS-PAGE were selected to allow efficient denaturation and reduction of disulphide bonds, whilst minimising DTT dependent cleavage of the target-intein thioester.

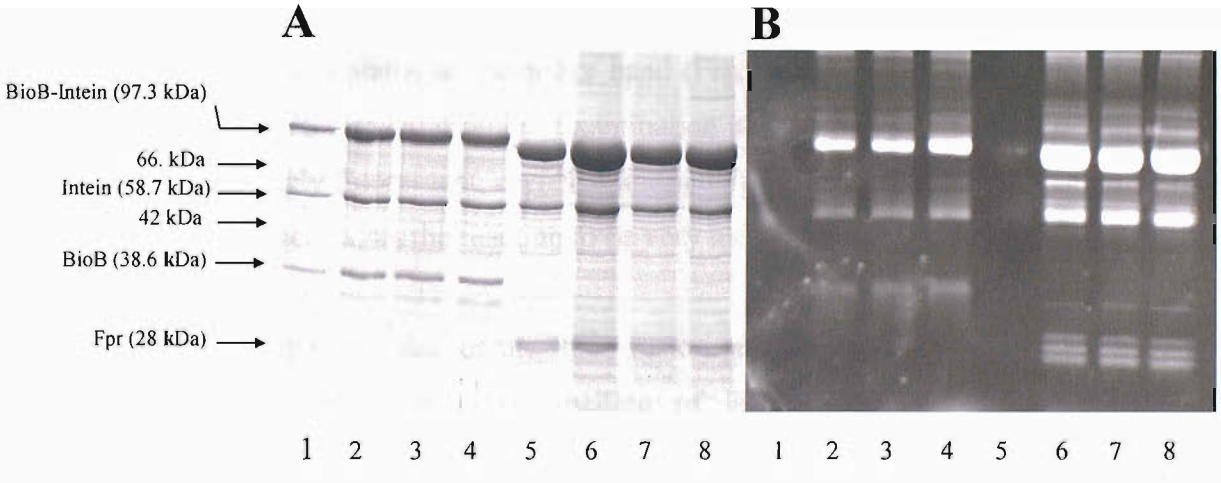


Figure 3.9 10 % SDS-PAGE analysis of fluorescent labelling of BioB, Fpr and LipA using FTEC (2.3 mM) under anaerobic conditions, loading buffer does not contain DTT. (A) Coomassie blue stained gel viewed under white light, (B) gel viewed on a UV light box. Lane 1 = BioB-intein negative control, 2 = BioB labelling reaction 4 h, 3 = BioB labelling reaction 24 h, 4 = BioB labelling reaction 48 h, 5 = Fpr-intein negative control, 6 = Fpr labelling reaction 4 h, 7 = Fpr labelling reaction 24 h, 8 = Fpr labelling reaction 48 h.

3.5 Investigating the direct fluorescent labelling of target proteins with FTEC

The long term objective is to use intein mediated ligation for efficient fluorescent labelling. To obtain some initial data for this reaction, the direct fluorescent labelling of BioB, Fpr and LipA (Fig. 2.10) using FTEC was attempted. Reaction conditions were carefully selected to provide the maximum extent of reaction, for example, tubes and buffers for the reactions were thoroughly degassed, and reactions were carried out under anaerobic conditions (glovebox, <0.2 ppm oxygen). Tris/HCl, pH 8.1 was the reaction buffer and FTEC was added at a >250 fold molar excess, the reaction time was 24 hours.

Samples were analysed by SDS-PAGE. Firstly considering the Coomassie stained gels, (Fig. 2.10 A,C) each of the proteins show negative controls to have reacted to a small extent with the DTT within the loading buffer (lanes 1, 3, 5). The BioB and LipA negative control and labelling reaction contain only one target protein band, which corresponds to the DTT cleaved protein. In the case of Fpr however, the fluorescent labelling reaction contains an extra Fpr band (Fig. 2.10, lane 4) at a higher molecular weight to the DTT cleaved protein. Examination under UV (Fig 2.10B) shows the new band to be strongly fluorescent, corresponding to Fpr fluorescently labelled with FTEC. The coomassie gel shows the reaction to be very inefficient, resulting in a <5 % yield.

Fpr-intein was the only one of the three fusion proteins to react with FTEC directly, resulting in specific C-terminal labelling of Fpr. This lack of reactivity was disappointing, and may be related to steric hindrance (*17*) towards attack of FTEC upon the target-intein thioester.

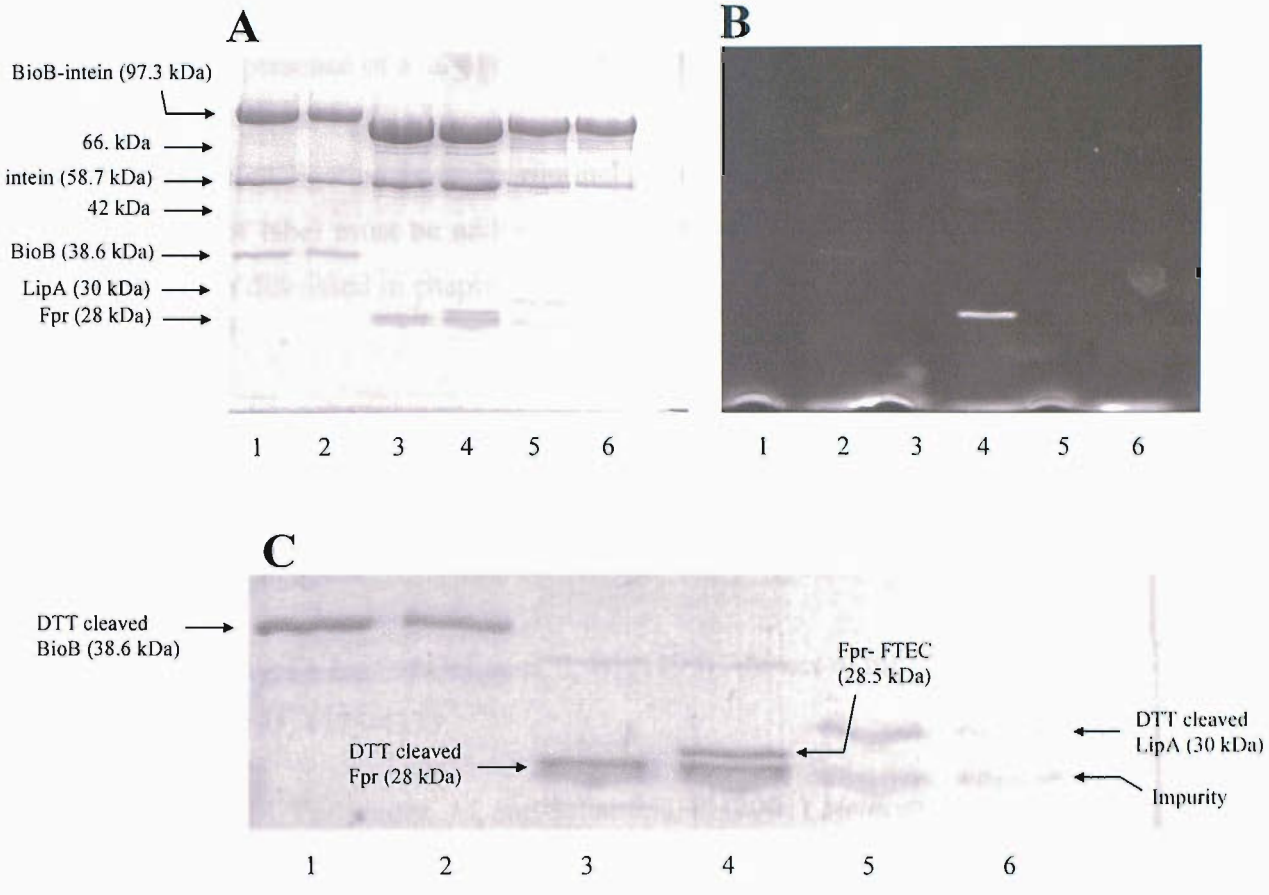


Figure 3.10 10 % SDS-PAGE analysis of fluorescent labelling of BioB, Fpr and LipA using FTEC (2.3mM) under anaerobic conditions, reaction time = 24 h. Loading buffer contains DTT. (A) Coomassie blue stained gel; Lane 1 = BioB-intein negative control, 2 = BioB-intein labelling reaction, 3 = Fpr-intein negative control, 4 = Fpr-intein labelling reaction, 5 = LipA-intein negative control, 6 = LipA-intein labelling reaction. (B) Fig. 2.9A viewed on a UV light box. (C) Fig. 2.9A expanded image of cleaved target proteins.

3.6 Summary and Conclusions

Cysteine, the natural substrate in protein splicing was found to be the most efficient molecule for reaction with the BioB-intein thioester. A cysteine based fluorescent label, FTEC, was synthesised over a simple 3 step route in 74 % yield. FTEC was found to be sensitive to oxidation to its symmetrical disulphide in solution, so fluorescent labelling was conducted under anaerobic conditions. Preliminary labelling experiments with

BioB-, Fpr-, and LipA-intein resulted in poor incorporation of the fluorescent label onto Fpr only, in the presence of a large excess of label.

If intein mediated ligation is to be routinely used for fluorescent labelling, the poor incorporation of label must be addressed. Optimisation of intein mediated ligation is essential, and is discussed in chapter 4.

3.7 References

1. Kapanidis, A. N., and Weiss, S. (2002) *Journal of Chemical Physics 117*, 10953-10964.

2. Daugherty, D. L., and Gellman, S. H. (1999) *Journal of the American Chemical Society 121*, 4325-4333.

3. Day, R. N., Periasamy, A., and Schaufele, F. (2001) *Methods 25*, 4-18.

4. Meyer, E., and Fromherz, P. (1999) *European Journal of Neuroscience 11*, 1105-1108.

5. Iakovenko, A., Rostkova, E., Merzlyak, E., Hillebrand, A. M., Thoma, N. H., Goody, R. S., and Alexandrov, K. (2000) *Febs Letters 468*, 155-158.

6. Macmillan, D., and Bertozzi, C. R. (2000) *Tetrahedron 56*, 9515-9525.

7. Lovrinovic, M., Seidel, R., Wacker, R., Schroeder, H., Seitz, O., Engelhard, M., Goody, R. S., and Niemeyer, C. M. (2003) *Chemical Communications*, 822-823.

8. Ayers, B., Blaschke, U. K., Camarero, J. A., Cotton, G. J., Holford, M., and Muir, T. W. (1999) *Biopolymers 51*, 343-354.

9. Chong, S. R., Montello, G. E., Zhang, A. H., Cantor, E. J., Liao, W., Xu, M. Q., and Benner, J. (1998) *Nucleic Acids Research 26*, 5109-5115.

10. Evans, T. C., Benner, J., and Xu, M. Q. (1999) *Journal of Biological Chemistry 274*, 3923-3926.

11. Edwards, J. O., Pearson, R.G. (1962) *Journal of the American Chemical Society 84*, 16-24.

12. Um, I. H., and Buncel, E. (2000) *Journal of Organic Chemistry 65*, 577-582.

13. Gill, M. S., Neverov, A. A., and Brown, R. S. (1997) *Journal of Organic Chemistry 62*, 7351-7357.

14. Weiss, S. (1999) *Science 283*, 1676-1683.

15. Tolbert, T. J., and Wong, C. H. (2000) *Journal of the American Chemical Society 122*, 5421-5428.

16. Krull, I. S., Strong, R., Sosic, Z., Cho, B. Y., Beale, S. C., Wang, C. C., and Cohen, S. (1997) *Journal of Chromatography B 699*, 173-208.

17. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

Chapter 4:- Fluorescent labelling of proteins via intein mediated ligation (*1*)

4.1 Introduction

Initial attempts at direct fluorescent labelling of BioB-, Fpr- and LipA-intein with FTEC resulted in poor incorporation of the fluorescent label, despite the use of a large excess of FTEC. It has been previously reported that the direct reaction of molecules containing an N-terminal cysteine residue with a target-intein fusion protein via intein mediated ligation was unsuccessful (*2*) and has been rationalised in terms of steric hindrance from the bulky intein domain (*3*). This problem was circumvented by literature concerning 'native chemical ligation', a method which utilises the synthesis of large peptides and their ligation to form synthetic protein domains (*4*). This occurs via reaction of an N-terminal cysteine residue with a C-terminal thioester, the reactivity of which was regulated by controlling the nature the α-thioester group (*5*). Typically, a thiol such as thiophenol is added to the reaction mixture to regulate the reactivity of the protein thioester, improving the yield of the ligation reaction (*6*).

This strategy was later applied to intein fusion proteins to enhance the yield of intein mediated ligation reactions, mechanistically it is assumed that this is via trans-thoiesterification to the more reactive intermediate (*7*) (Fig. 4.1 (**3**)). The most widely utilised thiols for the activation of the target-intein thioester are thiophenol, typically used at 195 mM (*3, 8*), mercaptoethanesulphonic acid (MESA), 30 mM (*9*) and ethanethiol, 0.49 M (*10, 11*). Under these conditions, cysteine derivatives are thought to form the kinetic product (**5**) and then to rearrange to the thermodynamically more stable amide (**6**) (*3, 5*).

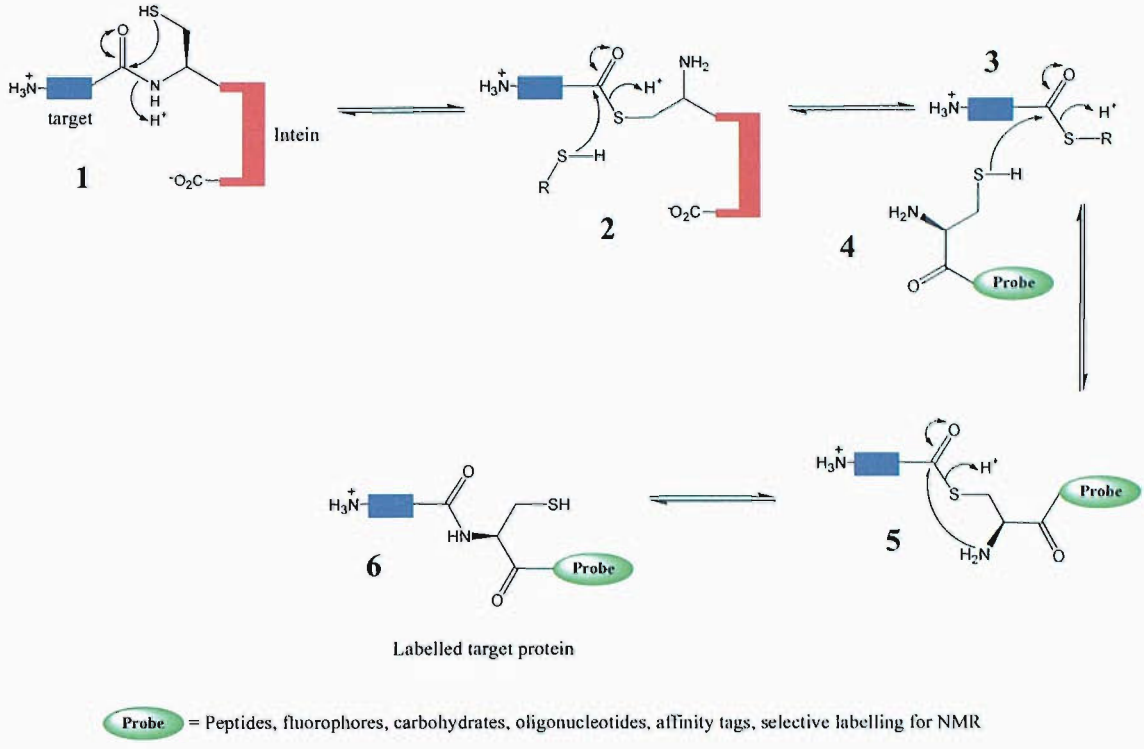


Figure 4.1 The mechanism of intein-mediated protein labelling. Acyl transfer of the target-intein fusion protein (**1**) from the alpha amino group to the thiol of the reactive cysteine yields thioester (**2**), which is attacked by a low molecular weight thiol to give the more reactive thioester (**3**). Attack of a cysteine based labelling reagent (**4**) yields thioester (**5**), which rearranges to give the thermodynamic amide product (**6**).

It was anticipated that the formation of thioester intermediates could lead to more efficient ligation of the FTEC label to the target protein domains. Therefore the factors that were important for the achievement of efficient and consistent labelling yields were investigated.

4.2 Comparison of aerobic and anaerobic thioester formation

The reaction of a range of low molecular weight thiols towards the BioB-intein thioester was investigated, with the objective of identifying the most efficient molecule for use in forming the reactive intermediate thioester (**3**). With this in mind, thiophenol derivatives were selected on the basis of substituents that might modifiy the electronic properties or water solubility. Thiols selected were thiophenol, 2-methoxythiophenol, 4-fluorothiophenol, 2-aminothiophenol and mercaptoethanesulphonic acid.

BioB-intein was selected as a model target-intein fusion protein. Experiments were carried out in parallel either under aerobic or anaerobic conditions, thiol concentration was 25 mM, and the results of these reactions are shown in Fig. 4.2. The general trend in reactivity was for thiols to be less effective under aerobic conditions than anaerobic. 2-aminothiophenol was found to be most reactive towards the BioB-intein thioester (Fig 4.2 lane 3), but is difficult to handle as it has poor solubility and forms an emulsion. In the case of thiophenol and 2-methoxythiophenol, a white precipitate was observed to form within 5 minutes, which was presumed to be due to disulphide formation upon exposure to air. 2-methoxythiophenol precipitation was seen to cause co-precipitation of BioB-intein (Fig 4.2, lane 4). The general trend in reactivity of the thiols towards the BioB-intein thioester under aerobic conditions was poor, suspected to be due to the formation of the unreactive symmetric disulphide of the reactive thiols. As a result, the effect of anaerobicity upon the efficiency of reaction was therefore investigated.

Under anaerobic conditions, the extent of reaction was greatly improved and no precipitation of thiols was observed. Thiophenol, 2-aminothiophenol, 4-fluorothiophenol and MESA were efficient at reacting with the BioB-intein thioester, forming the reactive α-thioester intermediate.

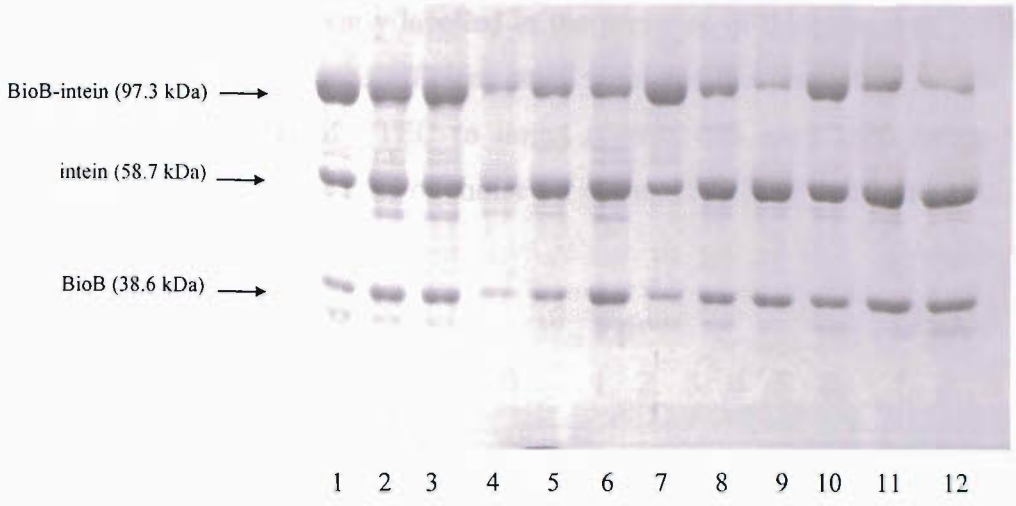


Figure 4.2 10 % SDS-PAGE analysis of the reaction of BioB-intein fusion protein with various thiols (25 mM), loading buffer contains DTT (50 mM). Lanes 1-6, aerobic; lanes 7-12, anaerobic reaction. Lanes 1 and 7 negative control (no thiol), lanes 2 and 8, thiophenol; lanes 3 and 9, 2-aminothiophenol; lanes 4 and 10, 2-methoxythiophenol lanes 5 and 11, mercaptoethanesulphonic acid; lanes 6 and 12, 4-fluorothiophenol.

4.3 Assay of thiol activated fluorescent labelling

Having investigated the intermediate thioester formation, the effect of a range of thiols upon the reactivity of BioB-, LipA- and Fpr-intein towards FTEC under anaerobic conditions was investigated. FTEC was reacted in the presence of thiophenol, 2-methoxythiophenol, 4-fluorothiophenol and MESA over 16 hours. In control experiments, the low molecular weight thiols were omitted. In the absence of the low molecular weight thiol, a poor / variable extent of labelling resulted, a maximum observed was ~10 % labelling of Fpr as estimated from Coomassie stained gel, and no labelling was observed of BioB or LipA.

The addition of low molecular weight thiols greatly increased the quantity of cleaved fusion protein, as observed in Fig. 4.3, lower panels, and the yield of fluorescently labelled protein was seen to increase in all cases. A broad correlation between the efficiency of cleavage, and the efficiency of labelling was observed, which suggests that under these conditions once an intermediate thioester is formed, reaction with FTEC is fast, leading to efficient fluorescent labelling. BioB-intein and LipA-intein are most

efficiently fluorescently labelled in the presence of thiophenol or MESA. Thiophenol was the slightly more efficient small thiol for the promotion of the intein mediated ligation reaction of FTEC to target protein and even considering its toxicity was considered to be the thiol of choice for optimal labelling under these conditions.

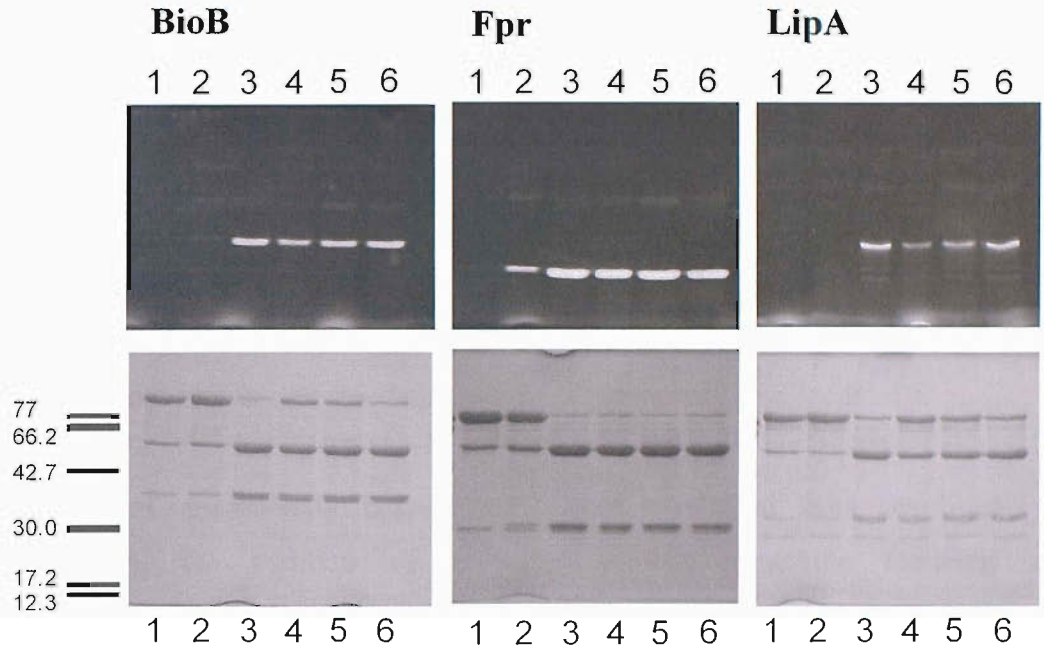


Figure 4.3 10 % SDS-PAGE analysis of fluorescent labelling reaction with BioB- , Fpr- and LipA-intein fusion proteins in the presence of various thiols (each 25 mM) anaerobically, over 16 hours. Analysis via 10 % SDS-PAGE, loading buffer contains DTT (50 mM). Top panels are fluorescence images (illuminated by transilluminator, λ = 302 nm), lower panels are Coomassie stained. In each case: lane 1 = no thiol, no FTEC, 2 = FTEC only, 3 = FTEC plus thiophenol, 4 = FTEC plus 2-methoxythiophenol, 5 = FTEC plus 4-fluorothiophenol, 6 = FTEC plus MESA.

4.4 Optimisation of intein mediated labelling reaction conditions

Having identified thiophenol as the most efficient thiol for promotion of fluorescent labelling via intein mediated ligation, the kinetics of fluorescent labelling of BioB under different conditions was compared, to further optimise the reaction. A timecourse of fluorescent labelling of BioB over 24 h was conducted. The extent of fluorescent labelling could be estimated by the integrating the fluorescent band intensity using Genetools software from gels viewed on a UV light box using a Genegenius imager.

Aliquots were withdrawn from the labelling reaction at different timepoints and the reaction quenched by freezing at -80 °C until analysis. The progress of the reaction was followed by SDS-PAGE. The extent of fluorescent labelling was quantified normalised with respect to a fluorescent standard (sample of previously labelled BioB).

The effect of anaerobicity (experiments were conducted both aerobically and anaerobically), the addition of a water soluble phosphine reducing agent tricarboxyethylphosphine (TCEP) to aerobic reactions and the addition of thiophenol on the efficiency of the labelling reaction were investigated. TCEP (*12*) was added to attempt to solve the problem of aerobic oxidation of FTEC / thiophenol, as it has previously been shown to be efficient for the reduction of disulphides (*13*).

A plot of fluorescence against time (Fig. 4.4) revealed that the labelling reaction proceeded most efficiently under anaerobic conditions in the presence of thiophenol. This data was fitted to a first order exponential, to give a pseudo first order rate constant of 0.30 ± 0.04 h^{-1}. Labelling under aerobic conditions with the addition of both thiophenol and reducing agent TCEP was also fitted, giving a rate constant of 0.23 ± 0.04 h^{-1}. Over multiple experiments, these rate constants consistently fell within 0.1 – 0.3 h^{-1}, which is comparable to that observed by Muir (*10*).

The aerobic experiments with thiophenol added, but TCEP omitted gave very interesting results. The initial observed rate is very similar to the previous examples, but the reaction is seen to quickly reach a maximum at ~0.4, after one hour of reaction and then almost stops. The most likely explanation is that thiophenol quickly oxidises,

forming a white precipitate under aerobic conditions. Prior to oxidation, thiophenol may both maintain FTEC in a reduced state, and react with BioB-intein, allowing efficient labelling. Once thiophenol oxidises and aggregates, the reaction comes to a halt. Reactions under other conditions showed little fluorescent labelling (< 10 %). Without thiophenol, even aerobically with TCEP, FTEC was unable to react with BioB-intein, only in the presence of thiophenol was efficient reaction observed.

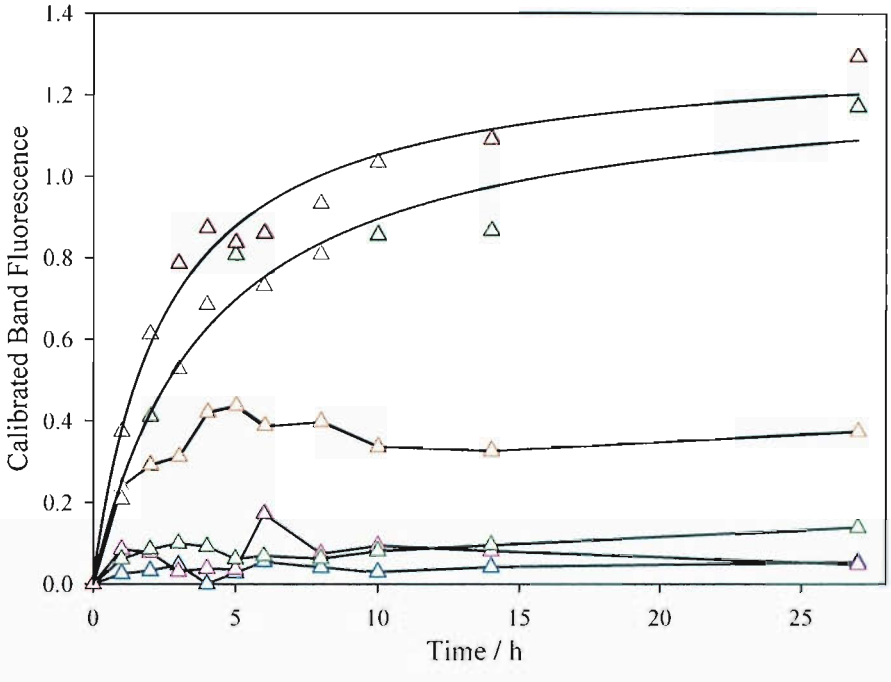


Figure 4.4 Timecourse of fluorescent labelling reaction with BioB-intein fusion protein. Graph showing timecourse of fluorescent labelling under various conditions, timepoints were 1, 2, 3, 4, 5, 6, 8, 10, 14, 26 hours. The following conditions were compared: (Δ) aerobic; (Δ) aerobic plus phosphine; (Δ) aerobic with thiophenol; (Δ) aerobic with phosphine and thiophenol; (Δ) anaerobic; (Δ) anaerobic with thiophenol. Data for anaerobic reactions with thiophenol and for aerobic reaction in the presence of thiophenol and TCEP were fitted to a first order exponential. For the remaining slow reactions data cannot be fitted and the points are connected to indicate the general trend.

4.5 Effect of thiophenol concentration on the labelling reaction

Although thiophenol has been shown to be the most efficient thiol at promoting intein mediated ligation, it is a toxic and highly reactive compound. It may limit the application of intein mediated ligation to more stable proteins, as for instance, an aggressive thiol such as thiophenol may react with labile Fe-S clusters (*14*) present in proteins such as BioB and LipA. Ideally, the addition of the low molecular weight thiol would be unnecessary for reaction with delicate proteins or in biological systems. It was therefore of interest to determine the minimum concentration of thiophenol required for efficient target protein labelling. The labelling of BioB was investigated anaerobically, using FTEC and a range of thiol concentrations from 0 – 25 mM over 16 h (Fig. 4.5).

Samples to which no thiophenol was added showed a variable yield of labelled BioB, from 0 – 80 % (in most cases 0 – 10 %, very occasionally higher proportions of labelling were observed). The underlying reason for this variability is as yet undetermined, but is likely to be simply due to oxidation of FTEC to its symmetric disulphide, as discussed in section 3.2.5. Variability of the oxygen content in samples used in reaction or in anaerobicity of the glovebox may contribute to this, as the FTEC thiol was observed to be highly oxygen sensitive. A concentration of 10 mM thiophenol consistently was observed to ensure efficient fluorescent labelling of BioB with FTEC under anaerobic conditions.

These results suggest that the small molecule thiol, in this case thiophenol, may have a twofold function in labelling reactions. Firstly, to form the more reactive thioester intermediate and secondly to maintain the FTEC thiol in a reduced and hence reactive state. Under these reaction conditions it is not possible to deconvolute the relative importance of these roles, as other parameters such as anaerobicity and steric constraints of the target-intein junction may also contribute (*15*).

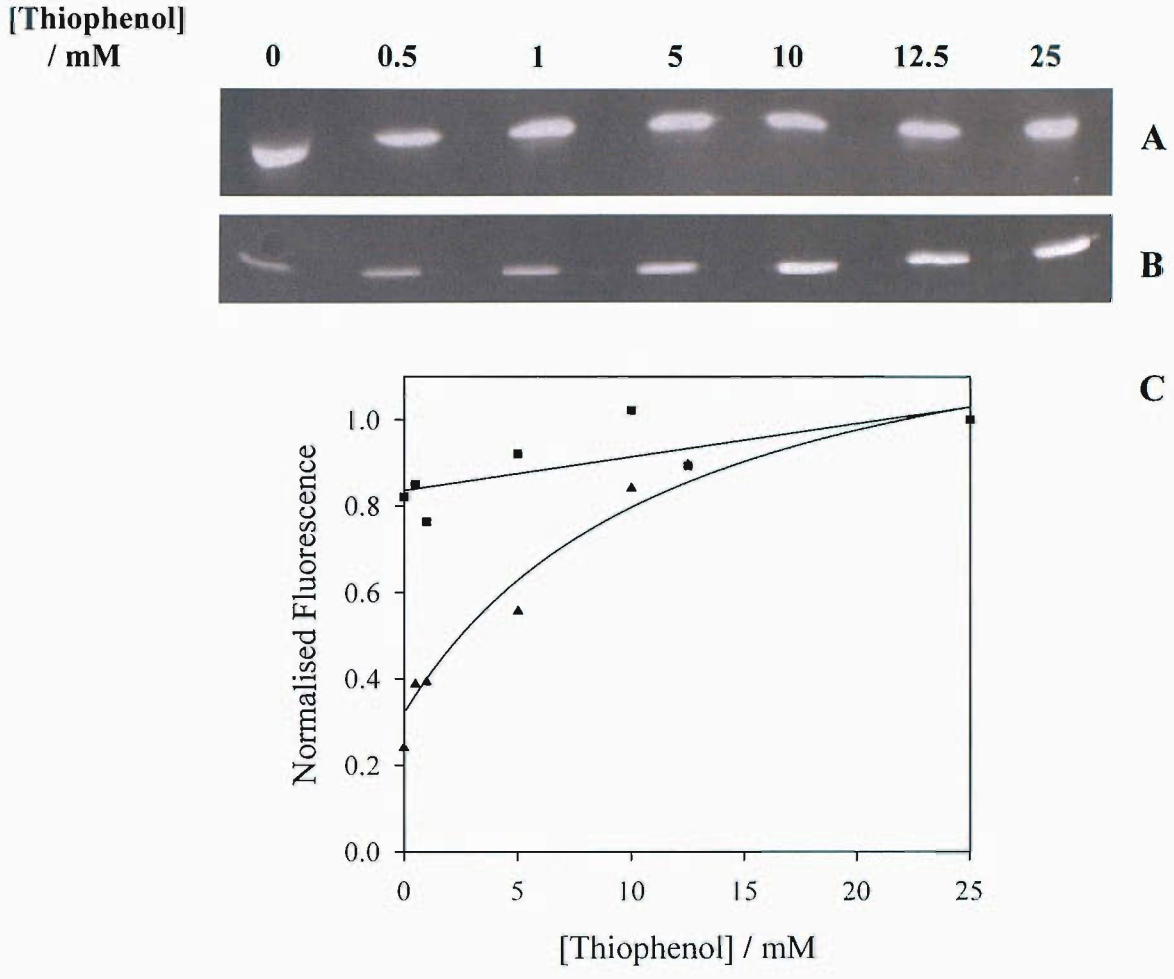


Figure 4.5 Examples of two fluorescent labelling reactions of BioB as analysed by SDS-PAGE, loading buffer contained DTT (50 mM), gels were viewed on a transilluminator, λ = 302 nm. (A) and (B) contain two different samples of FTEC (2.3 mM) in the presence of increasing concentrations of thiophenol as shown. (C) Graph showing yield of fluorescent protein at various thiophenol concentrations: (▲) A; (■) B. Data is fit to a straight line (A) and exponential curve (B).

4.6 Effect of the nature of the target protein upon reaction kinetics

The variable extent of labelling of different target proteins in section 4.3 suggested that the nature of the target protein may have an effect upon the rate of fluorescent labelling via intein mediated ligation. The rate of fluorescent labelling of BioB, Fpr, LipA, SolA and blue fluorescent protein (BFP, expressed and purified in collaboration with Mike Wang, Roach Group) was compared directly, by a timecourse of fluorescent labelling over 24 h, using optimal labelling conditions for BioB, which were anaerobic containing 10 mM thiophenol and 2.3 mM FTEC (Fig. 4.6). The rate of fluorescent labelling was judged as in section 4.4, with samples withdrawn and stopped by freezing at -80 °C, and analysis by SDS-PAGE.

In a plot of normalised fluorescence against time, data for BioB could be fitted to a first order exponential with a pseudo first order rate constant of 0.48 ± 0.06 h^{-1}. For the other proteins, reaction is too fast, and fitting is inappropriate. There is clearly a difference in the rate of fluorescent labelling of the proteins. The rate of labelling of LipA, Fpr and SolA is very similar, each reacting to completion in 2 h. The rate of BioB labelling is considerably slower, but BFP was observed to undergo very slow reaction, with <5 % labelling complete in 24 h.

This is the first comparison of reactivity of a range of target-intein fusion proteins. The rate of labelling of Fpr was predicted to be fast from its crystal structure (section 2.1.3) and the comparatively very slow rate of labelling of BFP may be caused by its rigid β-barrel structure (*16*), which may have an effect upon the ability of the BFP-intein junction to fold correctly to form the essential thioester. The lack of flexibility in the junction region may also sterically hinder the attack of FTEC.

These results suggest that intein mediated ligation should be a successful route for the site specific, mild labelling of proteins in many cases. Additionally, information regarding optimal sequences at the linker region between the target and intein domains for fluorescent labelling may be gained from these types of experiments. It may be possible to alter the nature of the BFP-intein junction for instance to increase reactivity

by the addition of a linking sequence into the BFP-intein junction, which may for instance correspond in sequence with the Fpr C-terminus.

The rate of labelling of Fpr, LipA and SolA was difficult to compare directly due to their fast kinetics. This method for measuring the course / rate of fluorescent labelling is suitable for gaining approximate kinetic data, but for more accurate information (i.e. a timecourse over 1 h, with a timepoint every 5 minutes for accurate rate constant determination) cannot be gained due to the difficulty of quenching the labelling reaction and error in timepoints due to rate of sample removal from the glovebox / rate of freezing and analysis. An alternative method is required for this type of experiment.

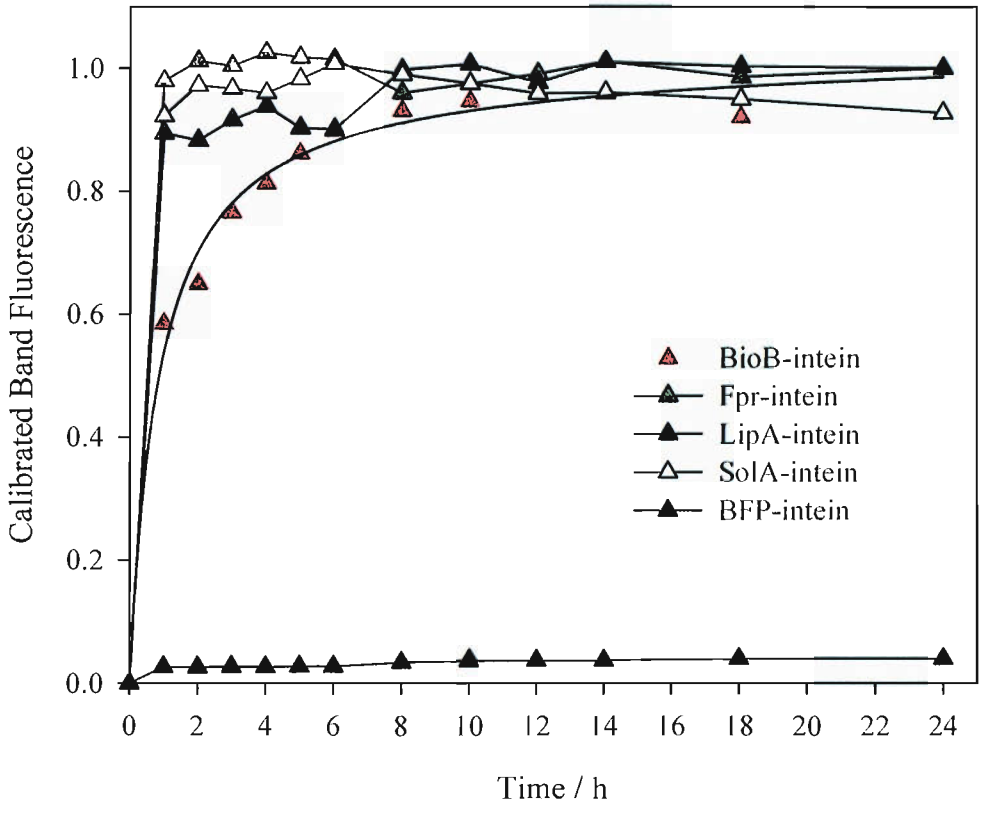


Figure 4.6 Timecourse of fluorescent labelling of BioB-, Fpr-, LipA-, SolA- and BFP-intein. Labelling was under anaerobic conditions in the presence of thiophenol (25 mM) using FTEC (2.3 mM). Samples were withdrawn from the reaction mixtures at times 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 18 and 24 hours and were analysed by SDS-PAGE. Gels were viewed on a transilluminator, λ = 302 nm, and bands quantified using genetools software. Data for BioB-intein was fitted to a first order exponential, remaining data could not be fitted and points are connected to show the general trend.

4.7 Fluorescent Labelling of BFP

Due to the slow rate and yield of fluorescent labelling of BFP, the effect of the concentration of thiophenol on the yield of labelled BFP under anaerobic conditions was investigated. The rate of formation of the activated intermediate thioester may be faster at higher concentrations of thiophenol.

The yield of fluorescently labelled BFP was greatest at concentrations of 100 and 200 mM thiophenol. Even under these aggressive conditions, a yield of <10 % (assessed from Coomassie stained gel) was observed. It can be concluded that the slow rate of fluorescent labelling is due to one of two factors, either the BFP-intein thioester forms but is inaccessible, or the tight folding of the BFP domain affects the folding of the intein junction in a way which causes the formation of the thioester to become less favourable, limiting the rate of reaction. The BFP-intein thioester is seen to form as partial reaction forming cleaved intein and fluorescently labelled BFP is observed, however, which of these factors is the major cause of this slow rate requires further investigation.

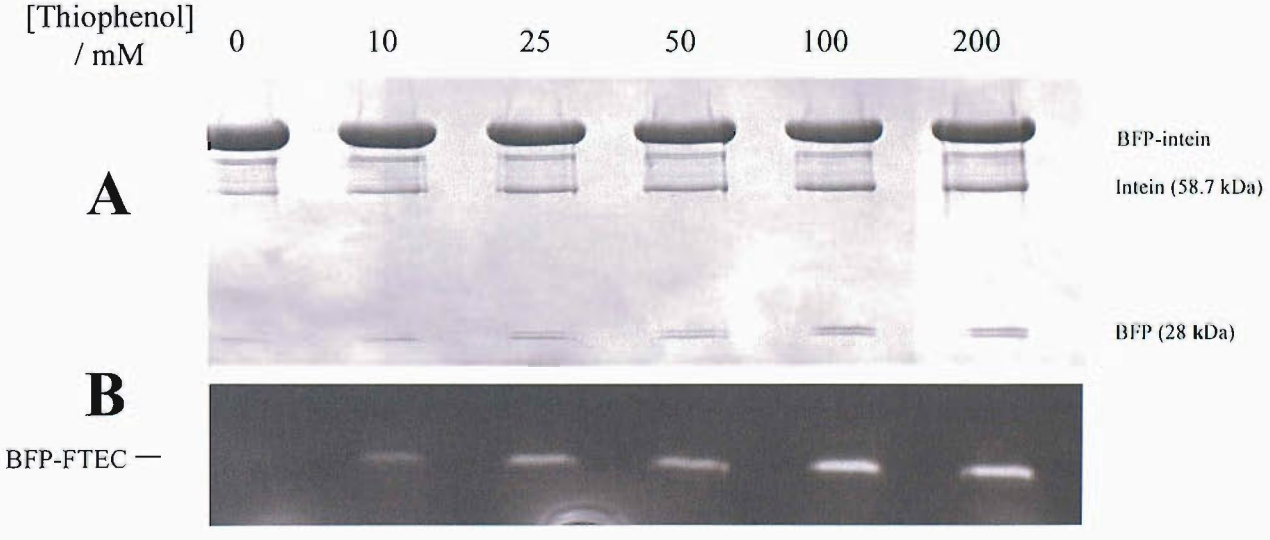


Figure 4.7 Fluorescent labelling of BFP anaerobically in the presence of thiophenol, 16 h reaction, sample loading buffer contains DTT (50 mM). (A) Coomassie stained, (B) gels viewed on a UV light box (λ = 302 nm). Lanes correspond to thiophenol concentrations as stated.

4.8 *in vivo* fluorescent labelling

In vivo fluorescent labelling is a powerful technique for the study of proteins in their natural environment, information such as cellular localisation (*17*) and in situ protein interaction (*18*) could be gained from this type of experiment. Currently relatively few techniques are available for *in vivo* fluorescent labelling (*19*), often proteins must be expressed in host cells as a GFP fusion (*20*). There is a great demand for new techniques to expand this field.

If the fluorescent label FTEC can cross the cell wall and accumulate inside the cell, under reducing intracellular conditions, intein mediated ligation to a target protein expressed as an intein fusion within host cells may be possible. The accumulation of FTEC within the *E. coli* cytoplasm was investigated by growing *E. coli* cells to an OD_{600} of 0.6, and incubating the cells in 0.5 mg / mL FTEC in PBS for 3 h.

After careful washing, comparison of cell lysates shows the FTEC treated cell lysate (Fig. 4.9) to be strongly fluorescent, indicating that FTEC may have accumulated in *E. coli*, opening the possibility for this technique to be applied to an *in vivo* system.

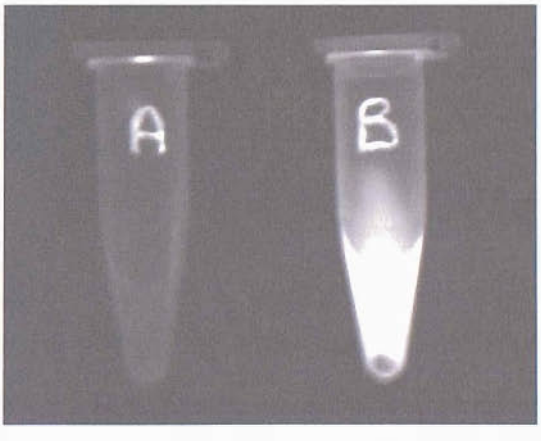


Figure 4.8 Accumulation of FTEC in the *E. coli* cell lysate. (A) Cleared cell lysate of cells incubated with no FTEC. (B) Cleared cell lysate of cells incubated with 0.5 mg/ml FTEC for 3 hours. Samples viewed on a transilluminator, (λ = 302 nm).

4.9 Conclusions

For most proteins, practical optimal conditions for intein mediated fluorescent labelling with FTEC have been identified. It is possible to use only FTEC under rigorously anaerobic conditions to label target proteins, but variable yields are obtained. Optimal labelling conditions were identified as occurring under anaerobic conditions with the addition of thiophenol (10 mM). Labelling under aerobic conditions with the further addition of TCEP, a water soluble phosphine reducing agent also produced a high yield after 24 h reaction.

The nature of the target protein domain was shown to be an important factor in the rate of fluorescent labelling, as shown by the slow labelling reaction of BFP. Further adjustment of the target-intein junction may be required to increase thioester reactivity, enabling efficient labelling in all cases.

FTEC was observed to accumulate in *E. coli* cytoplasm, suggesting that it may be possible to apply intein mediated fluorescent labelling to an *in vivo* system for the study of proteins in their natural environment.

4.10 References

1. Wood, R. J., Pascoe, D. D., Brown, Z. K., Medlicott, E. M., Kriek, M., Neylon, C., and Roach, P. L. (2004) *Bioconjugate Chemistry 15*, 366-372.

2. Evans, T. C., and Xu, M. Q. (1999) *Biopolymers 51*, 333-342.

3. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

4. Schnolzer, M., and Kent, S. B. H. (1992) *Science 256*, 221-225.

5. Dawson, P. E., Muir, T. W., Clarklewis, I., and Kent, S. B. H. (1994) *Science 266*, 776-779.

6. Evans, T. C., Benner, J., and Xu, M. Q. (1998) *Protein Science 7*, 2256-2264.

7. Chong, S. R., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. (1997) *Gene 192*, 271-281.

8. Severinov, K., and Muir, T. W. (1998) *Journal of Biological Chemistry 273*, 16205-16209.

9. Tolbert, T. J., and Wong, C. H. (2000) *Journal of the American Chemical Society 122*, 5421-5428.

10. Ayers, B., Blaschke, U. K., Camarero, J. A., Cotton, G. J., Holford, M., and Muir, T. W. (1999) *Biopolymers 51*, 343-354.

11. Evans, T. C., Benner, J., and Xu, M. Q. (1999) *Journal of Biological Chemistry 274*, 3923-3926.

12. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001) *Cell 106*, 453-463.

13. Burns, J. A., Butler, J. C., Moran, J., and Whitesides, G. M. (1991) *Journal of Organic Chemistry 56*, 2648-2650.

14. Sanyal, I., Cohen, G., and Flint, D. H. (1994) *Biochemistry 33*, 3625-3631.

15. Xu, M. Q., and Evans, T. C. (2001) *Methods 24*, 257-277.

16. Wachter, R. M., King, B. A., Heim, R., Kallio, K., Tsien, R. Y., Boxer, S. G., and Remington, S. J. (1997) *Biochemistry 36*, 9759-9765.

17. Pouli, A. E., Emmanouilidou, E., Zhao, C., Wasmeier, C., Hutton, J. C., and Rutter, G. A. (1998) *Biochemical Journal 333*, 193-199.

18. Keppler, A., Pick, H., Arrivoli, C., Vogel, H., and Johnsson, K. (2004) *Proceedings of the National Academy of Sciences of the United States of America 101*, 9955-9959.

19. Chen, I., and Ting, A. Y. (2005) *Current Opinion in Biotechnology 16*, 35-40.

20. Festenstein, R., Pagakis, S. N., Hiragami, K., Lyon, D., Verreault, A., Sekkali, B., and Kioussis, D. (2003) *Science 299*, 719-721.

Chapter 5:- Electrode surface immobilisation of proteins by intein mediated ligation

5.1 Introduction

5.1.1 Applications of protein immobilisation

Through evolution, a wide range of proteins for electron transport and the catalysis of oxidative and reductive reactions have been created. The motivation behind the immobilisation of proteins on electrode surfaces is in the harnessing of these enzymes allowing their application in a wide range of bioelectronic technology (*1*). Protein modified electrodes have been applied as biosensors (*2*), diagnostic protein arrays (*3*), biofuel cells (*4*), as biocatalysts for organic synthesis (*5*), logic gates, and optical memories (*6*).

The most basic feature of bioelectronics is the requirement to immobilise a biomaterial onto a conductive support. Biomaterials that have been immobilised include proteins (*7*), receptors (*8*), antibodies or antigens (*9*), oligonucleotides (*10*) and small molecules such as biotin (*11*) which exhibit a specific interaction with the material which is to be assembled. Different electronic responses have been employed to detect the functions occurring at the biological molecule, including the flow of current (*12*), changes in potential (*13*), capacitance (*14*) and impedance (*15*).

The most widely studied is the immobilisation of proteins, as their highly specific enzyme-substrate interactions and often high turnover rates have enabled the development of highly specific biocatalysts. The beauty of the use of redox enzymes in any application requiring the oxidation / reduction of a material is their specificity and the ability to alter the rate of the chemical transformation by controlling the

electrochemical potential. The potential applied across the electrode / solution interface provides the driving force for the reaction, with the resulting current giving a direct measurement of the rate of catalyst turnover.

For application for example in biosensing, a level of organisation must be introduced which allows integration of the biomolecule with the conductive support such that biocatalytic transformations are electrochemically detected (Fig. 5.1). Amperometric detection of enzyme generated products such as H_2O_2 (i.e. glucose oxidase (*16, 17*) or other oxidases (*18*)) has commonly been utilised in the development of biosensor devices.

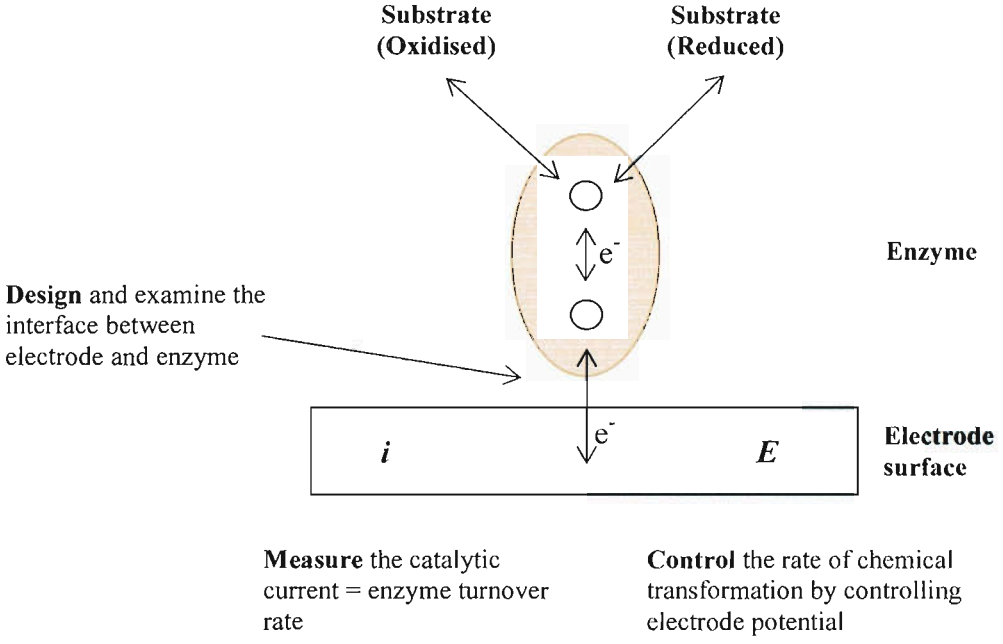


Figure 5.1 Depiction of the study of an electrode immobilised enzyme.

There are several challenges confronting the development of a bioelectronic device using proteins (*19*). The general requirement of redox proteins for a cofactor (such as NADH or a flavin cofactor such as FAD) which will require regeneration post electron transfer is a major challenge. Often these cofactors are expensive and tend to be easily adsorbed on electrode surfaces, these immobilised cofactors then may degrade at the surface (*20*).

The stability of the enzyme is crucial to its application, the maximum activity is often observed to only be maintained for a short period of time (*21*) due to protein denaturation. The protein then must be immobilised onto the electrode surface in a manner which achieves reproducible and fast electron transfer. Technical challenges include the even distribution of the enzyme on the electrode surface and the modification of the surface to provide a stable linkage.

For all applications, the common element is the requirement for methodologies for reproducible, controlled protein immobilisation on electrode surfaces. Methods should allow control over the immobilisation process such that the enzyme will be stable and orientated such that efficient electron transfer with the electrode surface may occur. Enzyme orientation is a crucial factor in all applications, as the three dimensional structure of proteins is complex, and immobilisation must not interfere with the access of the substrate with the protein active site, and efficient electron transfer must occur. Below are discussed some of the methods commonly used for protein immobilisation on electrode surfaces.

5.1.2 Electrostatic Immobilisation

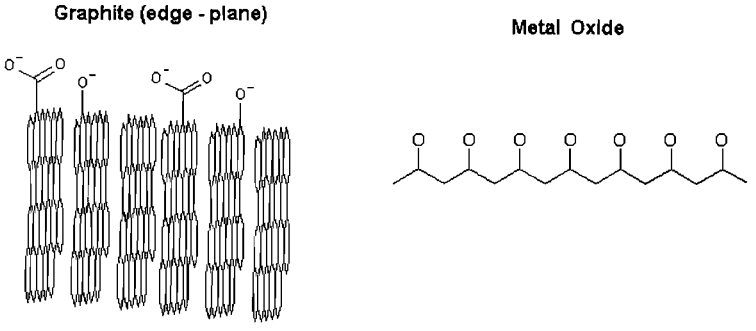


Figure 5.2 Electrostatic immobilisation

Proteins are known to spontaneously adsorb onto a range of electrode surfaces, an interaction that may be strong and irreversible (*22*). The physical adsorption of the protein directly onto the electrode surface offers the advantages of simplicity, wide application and minimal requirement for surface tailoring.

This simple method requires optimisation of the electrostatic interaction between surface and protein. The most common surfaces used in this context are pyrolytic edge-planar graphite (PGE) which has acidic C-O functionalities and gold modified with self assembled monolayers (SAMs). Cytochrome c (*23*) has been found to strongly adsorb at gold electrodes modified with ω-carboxylate alkane thiol SAMs. This is due to surface interaction with lysine residues located around the exposed heme edge. Another example of a surface used is a metal oxide, which is a stable hydrophilic material. Metal oxides are often also transparent, which allows spectrochemical changes to be studied as the immobilised protein is reduced and oxidised (*24*). Protein stability in the case of the adsorbed system has been improved by cross-linking the protein molecules using glutaraldehyde (*25*).

The *Desulfovibrio* hydrogenase catalyses the reversible oxidation:

$$H_2 \leftrightarrow 2H^+ + 2e^-$$

so that hydrogenases have the ability to either evolve or oxidise H_2 (*26*). At conventional electrodes, hydrogenases do not generally exhibit reversible electrochemistry. An approach to their immobilisation was via attraction of the negatively charged domain at which electron transfer occurs towards the electrode surface using a pyrolytic graphite electrode coated with a cationic poly(L-lysine) layer (*27, 28*). This example shows that it is essential to orientate the protein correctly at the electrode surface to allow efficient electron transfer to occur. These adsorbed electrodes have however shown to have poor stability, with enzyme activity observed to reduce quickly during the first day of storage, to inactive after a week (*29*).

The drawbacks of this method are in the lack of control over the process. There is scope for proteins to be immobilised in random distribution and orientation and there is little ability to control the separation between electrode surface and protein molecule. Due to the nature of the interaction with the surface, proteins may be denatured at the protein - electrode interface. This has however been proven to be a reliable and relatively simple method for the study of many protein redox systems.

5.1.3 Covalent bonding

Covalent attachment of enzymes to an electrode surface provides a more stable protein layer. Covalent modification is typically a more complex procedure than adsorption as it involves three steps, firstly the electrode surface has to be activated, the electrode must then be modified with a reactive group towards the protein and the enzyme must then be coupled to the surface. This method however often provides a more stable immobilised enzyme layer (*30*). Methods include activation of a carbon surface by oxidation, providing carboxylic acid functionalities, which, through carbodiimide coupling may be activated and the protein directly coupled to the surface via an amide bond (*31*). Glucose oxidase has been successfully immobilised via this route resulting in a stable catalytically active enzyme which remains active for up to ten days of storage (*32*).

Another approach which has been widely applied to protein immobilisation involves the reaction of a metal oxide (for example SnO_2, In_2O_3, TiO_2) or glassy carbon / pyrolytic graphite with the versatile cyanuryl chloride (2,4,6-trichlorotriazine). This molecule reacts with the hydroxyl groups on the electrode surface to form oxy-substituted mono- and dichlorotriazine layers (*33*). These functionalities are reactive towards lysine and tyrosine side chains, and the resulting protein monolayer is attached to the electrode by one or more reactive amino acid residues exposed on the protein surface. Examples of immobilised proteins via this route are *Aspergillus niger* glucose oxidase (*34*) for glucose sensing, *Crotalus atrox* L-aminoacid oxidase (*35*) for phenylalanine detection and *E. coli* alkaline phosphotase (*36*) for catechol phosphate analysis.

Other methods for covalent protein immobilisation include functionalisation of the surface with an amine for coupling with the protein C-terminus and the modification of gold electrodes with thiol self-assembled monolayers for further covalent linkage with proteins.

5.1.4 Immobilisation in surfactant films

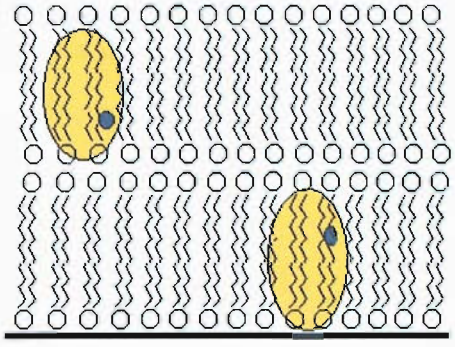


Figure 5.3 Schematic representation of the immobilisation of proteins in membranes.

This method provides an environment for membrane proteins which models that of the cell membrane phospholipid bilayer. The surfactants are water-soluble and have a polar head with two long alkyl chains. Immobilisation of proteins in these layers allows them to be free to diffuse throughout the layer, but they cannot escape into the bulk aqueous phase. These layers are cast onto electrode surfaces from solutions of surfactants in an organic solvent or by aqueous vesicle dispersions, examples include the immobilisation of cytochrome c_3 (*37*) and myoglobin (*38*). It was observed that the protein was strongly adsorbed in a stable film, and showed a strong electrochemical response when compared to adsorption of protein at bare / SAM modified electrodes. It is suggested that this method has the advantage that it avoids the problems of the formation of an insulating layer of denatured protein or impurities onto the surface which inhibit the passage of electrons. The disadvantages are that it is not suitable for unstable enzymes and it is unlikely that all will give reversible voltammetry in these films.

5.1.5 Mediated protein electron transfer:- 'wiring' the protein to the electrode surface

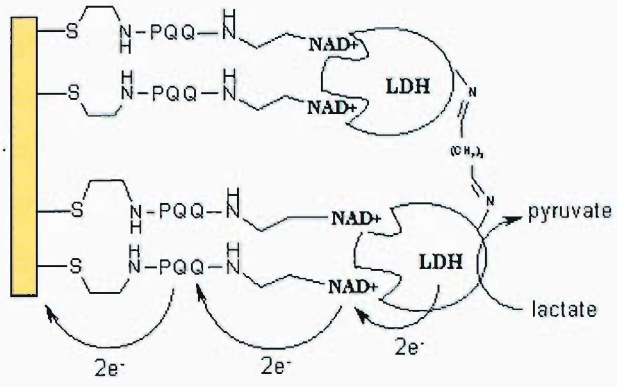


Figure 5.4 An integrated lactate dehydrogenase electrode and the bioelectrocatalytical oxidation of lactate at the electrode surface.

A class of proteins known as oxidoreductases (e.g flavodoxin $NADP^+$ reductase) utilise small molecule cofactors and in contrast to the large membrane proteins (whose redox active centres are mostly accessible to the outer surface of the protein) their electroactive component is often buried deep within the protein structure, and is therefore often in poor contact with the electrode giving little or no response to cycling potential.

One method used is to 'wire' the protein cofactor directly onto the electrode surface, an example is the study of bonding the lactate dehydrodgenase cofactor NAD^+ to a gold surface. This was done by immobilising the NAD^+ onto a tether containing PQQ (mediator). To this was then added the apo-protein which bound the cofactor after which the proteins were cross linked on the surface with glutaraldehyde to stabilise the enzyme layer formed (*39*) (Fig. 5.4). The stability of the layer was found to be moderate (after 30 minutes, only 25 % of the biocatalyst had dissociated from the surface), the resulting integrated protein layer was able to biocatalyse the oxidation / reduction of the cofactor whilst catalytically turning over lactate $\rightarrow$ pyruvate. This biocatalysis was made possible by the recycling of the reduced NADH $\rightarrow$ NAD^+ by the oxidation of the NADH by PQQ, as the electrons are passed down the wire from lactate $\rightarrow$ NAD^+ $\rightarrow$ PQQ $\rightarrow$ electrode, producing a novel self contained biocatalyst (*40*).

This method for the study of oxidoreductases has proved successful. Another example is the immobilisation of glucose oxidase (*17*) via tethering of its flavin cofactor. Horseradish peroxidase reconstitution was investigated (*41*) using a similar method. This showed that this process is highly sensitive to the chain length of the tether. It was found that a C12 chain length was optimal - a shorter chain showed no successful reconstitution, resulting in slow electron transfer kinetics, it was suggested that this might be improved by the addition of mediator to the tether. This highlights a significant problem with this method, as the cofactors are often bound deep within the protein structure, the tethers will often have to be long to allow reconstitution which may lead to difficulties in protein reconstitution as well as causing the kinetics of electron transfer to be slow.

5.1.6 Proteins engineered for surface immobilisation:- Histidine tags

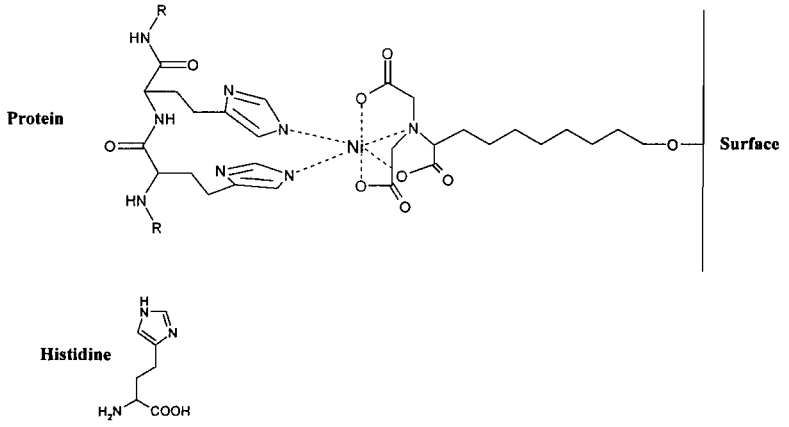


Figure 5.5 Protein immobilisation via Ni^{2+} coordination to a His_6-tag.

Histidine is a naturally occurring amino acid, whose side chain contains an imidazole moiety. This has a strong affinity for transition metal ions such as Ni^{2+} and Cu^{2+}. The addition of a sequence of 6-histidines (Fig. 5.5) to the N- or C-terminus of a protein sequence by PCR provides a site which can bind to a metal ion attached to an electrode surface. Often the metal ion is coordinated to a nitrilotriacetic acid group (Fig. 5.5). The metal ion tends to coordinate to two of the histidines present within the tag, the length ensures that the histidines protrude from the protein surface, optimising the binding process. Alternatively, if the 3D structure of the protein is known, two

histidines only have been shown to be required for coordination, for example, two histidines were engineered into the outer surface of the *Anabaena* ferrodoxin $NADP^+$ reductase (*42*). The protein was modified with a histidine pair His-X_3-His motif, where X = a randomised amino acid sequence. This was coordinated (on a gold electrode) to a Cu^{2+} ion bound to a SAM terminating in a nitrilotriacetic acid group. The histidine pair was engineered in two different positions on surface exposed sites of the protein, which was shown to be immobilised at the surface in two different orientations, predicted by the position of the mutations.

An alternative is immobilisation using a cysteine residue specifically engineered into the surface of the protein, which will covalently bond with the gold electrode surface, for instance, azaurin (*43*), the 'blue copper protein' was immobilised by the addition of an engineered cysteine residue into the structure. The Au-S bond formed tethers the protein to the surface, restricting its lateral movement unlike in the electrostatic case, where the protein is free to migrate over the surface. This is an advantage over the electrostatic method as control is gained over the molecular orientation. The disadvantages are that modification of the protein sequence is required by mutagenesis pre immobilisation and there is no real control over the manner in which the proteins are arranged in terms of density at the electrode surface.

5.1.7 Surface immobilisation via intein mediated ligation

Intein mediated ligation has been shown to provide a mild method for the activation of the C-terminus of a protein domain towards reaction with a cysteine based molecular probe (*44*). The application of this technology to protein immobilisation would potentially provide a highly specific method for the attachment of proteins to electrode surfaces.

The potential benefits of this method are that the orientational control over the immobilisation process is high, as the protein is reactive only in a single position, the protein will be immobilised in its native state – no further modification of the protein is required for instance removal of reactive residues at the protein surface. It looks to provide a clean method for surface derivatisation, and the separation between surface

and protein may simply be controlled by varying the distance between electrode surface and reactive cysteine. Particularly in cases such as immobilisation on vireous carbon for high surface area biocatalytic / bio fuelcell applications and for techniques such as screen printing this method may provide good control over the derivatisation process and some control over the density of protein loading on the electrode surface may also be gained. The use of a range of protecting groups might allow the systematic immobilisation of different proteins on the same electrode.

To immobilise proteins via this technique, two surfaces which may be derivatised are gold and carbon. Gold as an electrode surface however was not investigated in this work because there is a possibility that the reactive cysteine residue required for reaction with the intein fusion protein may itself react with the gold surface, so the inert carbon surface was investigated as the target for protein immobilisation.

5.1.8 Surface characterisation by X-ray photoelectron spectroscopy (*45*)

A powerful technique for the study of surface modification is X-ray photoelectron spectroscopy. It is a method which essentially involves measuring the energy spectrum of electrons ejected from a sample upon bombardment with monoenergetic X-rays. The energies of ejected electrons differ according to their orbitals of origin, and may be related to different ionization potentials for the same atom or molecule.

The recorded spectra therefore serve to identify the elements present and provide information about the chemical structure of the sample. For example, the K-shell ionisation potentials for the first eighteen elements in the periodic table vary from 14 – 3203 eV. For example, elemental C1s = 280 eV, N1s = 400 eV and O1s = 530 eV. An experimental observation of an ionisation potential within a few volts of 400 eV could therefore be assigned to the presence of nitrogen in the sample. The 'chemical shift' or extent to which the ionisation potential or binding energy varies from the elemental value gives provides information as to the molecular environment of the atom, for instance the presence of two peaks of equal area close to 400 eV would indicate the presence of two nitrogen atoms in a molecule in non-equivalent positions.

This technique is well suited to surface analysis of chemically modified electrodes, as X-rays can only eject electrons from the uppermost 10 nm of a solid. It has been used to analyse many of the samples in this chapter and has proved invaluable for analysis of surface modifications which could not simply be electrochemically characterised.

5.2 Electrochemical modification of a glassy carbon electrode by the reduction of aromatic diazonium salts (*46*)

The single electron reduction of aryl diazonium salts at a glassy carbon electrode results in covalent modification of the surface with aryl groups (Fig 5.6). The reduction is usually effected in acetonitrile (ie non-aqueous solution) and is controlled by cyclic voltammetry.

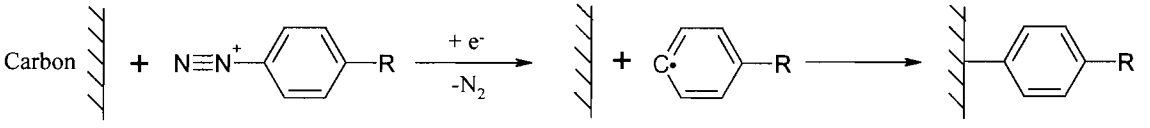


Figure 5.6 Modification of glassy carbon with aromatic diazonium salts.

This method is ideal for the derivatisation of a carbon electrode in our case, as it has been well characterised, and the electrodes have been shown to be stable over a long period to storage in air and to sonication in aggressive organic solvents.

The diazonium salt of 4-nitrobenzene has been selected for this step, as once covalently attached, the reversible one electron reduction of the nitro group will allow the surface coverage of nitrobenzene to be assessed. This can then be reduced in protic medium to the amine for coupling chemistry (*47*). This type of modification was appealing as it may allow monitoring of the coupling efficiency at each step, as for instance, each modification might introduce an electroactive moiety, allowing the coupling yield after each step to be calculated.

5.2.1 Derivatisation of a glassy carbon electrode with nitrobenzene

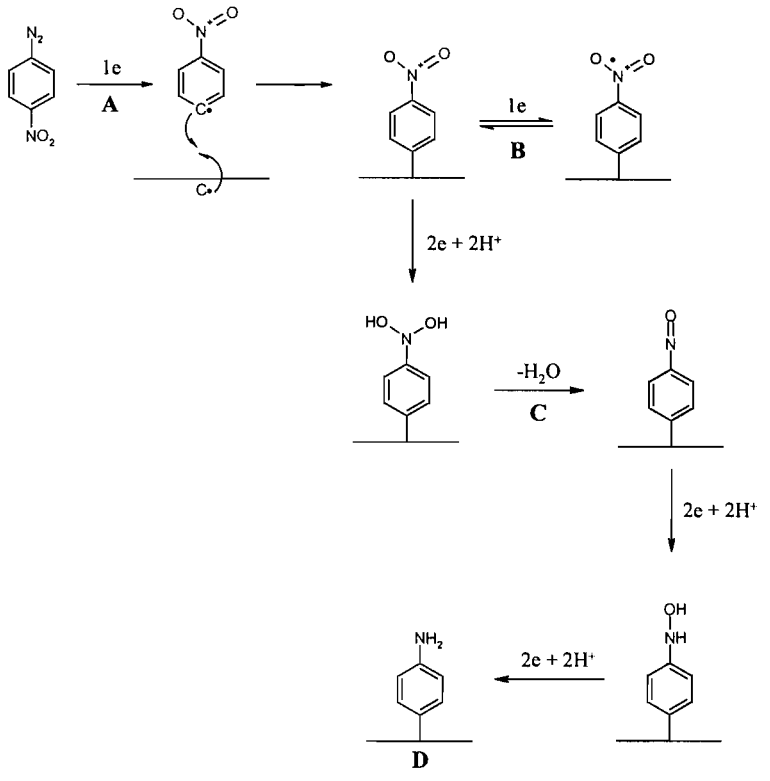


Figure 5.7 The derivatisation of a glassy carbon electrode with aniline. Reference to voltammograms is in bold type, as shown in Fig. 5.8 and 5.9,

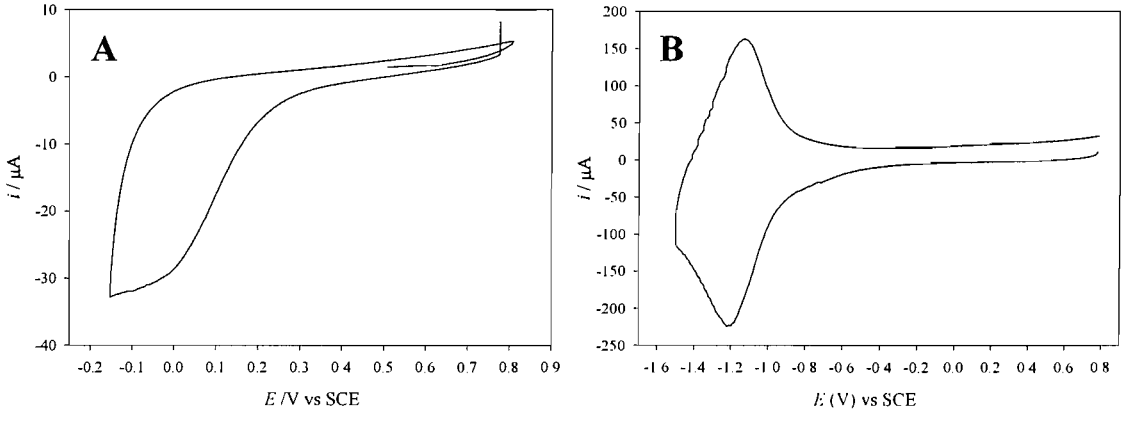


Figure 5.8 (**A**) Derivatisation of a 7 mm diameter glassy carbon electrode via voltammetry in 1 mM 4-nitrobenzene diazonium salt / 0.1 M NBu_4BF_4 in acetonitrile, ν = 50 mV / s, Ar, RT. (**B**) 7 mm diameter glassy carbon electrode derivatised with 4-nitrobenzene, voltammetry in 0.1 M NBu_4BF_4 in acetonitrile, ν = 50 mV / s, Ar, RT.

The first derivatisation step involves a one electron reduction of nitrobenzene diazonium salt at the glassy carbon electrode surface (Fig 5.7, 5.8 A). The voltammetry involves a single potential cycle from +0.8 V to –0.15 V vs SCE. As the potential is cycled, a broad reduction peak is observed which is representative of the one electron diazonium reduction.

The resulting coverage of the electrode with nitrobenzene was determined by voltammetry of the derivatised electrode in pure electrolyte (Fig. 5.7, 5.8 B). The cyclic voltammogram shows the potential to be swept from 0.8 to -1.4 V VS SCE. A reduction current is observed sweeping to -1.4 V vs SCE and on the return scan, an oxidative current is observed, showing the reversible one electron reduction of nitrobenzene as in Fig. 5.7 B. It is possible to determine the coverage (Γ_R) of nitro groups on the surface from the area under either the single electron reduction or oxidation peak of nitrobenzene (Fig. 5.8 B). The area under the peak is representative of the charge passed i.e. the number of single electron reduction / oxidations occurring at the electrode surface. The electrode coverage is calculated by integrating the charge under the peak and dividing this by the scan rate, giving us the charge passed, Q. To convert this to a surface coverage we use:

$$\Gamma_R = Q / nFA$$

Equation 1. Surface coverage calculation

Where n is the number of electrons involved the reaction, F is Faraday's constant and A is the area of the electrode in cm^2, this will give a coverage (Γ_R) in moles / cm^2. In this case, the experimental values range from $8 - 14 \times 10^{-10}$ moles / cm^2, compared to a theoretical coverage of 12.5×10^{-10} moles / cm^2 based upon a close packed arrangement of nitrobenzene on the electrode surface, with the aromatic rings orientated perpendicular to the surface. The close agreement between theoretical and experimentally observed results suggests that the modification reaction results in a monolayer coverage of nitrobenzene.

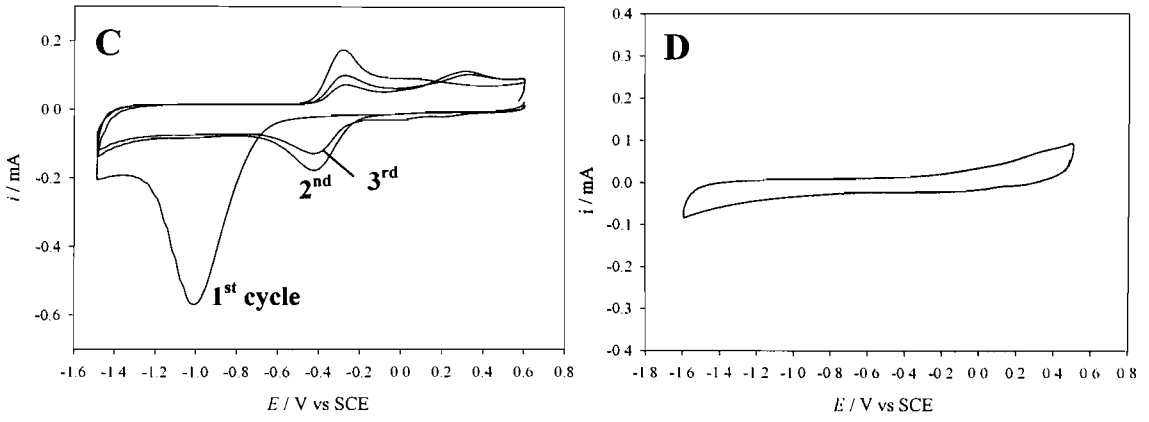


Figure 5.9 (**C**) Voltammetry of nitrobenzene derivatised 7 mm diameter glassy carbon electrode in 0.1M KCl solution in 9:1 water : ethanol, 3 scans, ν = 50 mV / s, Ar, RT. (**D**) Voltammetry of aniline modified 7 mm glassy carbon electrode in 0.1M KCl solution in 9:1 water / ethanol post electrolysis at -1.4 V vs SCE for 100 minutes, ν = 50 mV / s, Ar, RT.

The reduction of nitrobenzene to aniline on the electrode surface was by electrolysis at -1.4 V vs SCE for 110 minutes. Investigation of the reduction by cyclic voltammetry between + 0.7 and -1.5 V vs SCE (Fig. 5.9 C which shows the first three cycles) shows a large, irreversible reduction peak at -1.0 V vs SCE, the area under which, assuming the molecule is orientated at the surface as before corresponded to a nitrobenzene coverage of 12×10^{-10} moles / cm^2 with the charge under the aqueous nitrobenzene reduction peak corresponding on average, to a five electron reduction per mole of immobilised nitrobenzene, when the complete reduction of nitrobenzene requires a six electron reduction. On the return scan in the first cycle, an oxidation peak is observed at – 0.4 V vs SCE which can be assigned to the formation of an intermediate species which accounts for the remaining incompletely reduced nitrobenzene. Further potential cycling shows the reduction of the size of the peak with each cycle.

This behaviour has been observed in the literature for the reduction of nitrobenzene to aniline, it has been reported to be the NHOH / NO intermediate (*48*), full reduction is achieved by electrolysis (Fig 5.9 D). This indicated that in the first cycle, the majority of the nitrobenzene is completely reduced, but a proportion of the molecules undergo only a two electron reduction, forming the NO intermediate, whose reduction is completed over a long electrolysis at -1.4 V vs SCE.

5.2.2 Derivatisation of an aniline modified glassy carbon electrode via conventional coupling chemistry

The reactivity of the aniline modified surface towards 4-nitrobenzoic acid using conventional amide coupling chemistry was investigated. This molecule was chosen as again, its coupling to the aniline modified surface could be directly monitored by electrochemistry of the nitro group. Coupling with either dicyclohexylcarbodiimide or carbonyldiimidazole in DCM, DMF or acetonitrile failed to produce any coupling of nitrobenzene at the surface which could be detected electrochemically. This unreactivity of the aniline modified surface along with the lengthy modification procedure led to the investigation of an alternative method for the derivatisation of glassy carbon.

5.3 Covalent modification of glassy carbon via the electrochemical oxidation of amines

Primary and secondary amines may be covalently attached to a glassy carbon surface via oxidation in anhydrous ethanol. Cyclic voltammetry past the irreversible amine oxidation peak leads to modification (*49*) (Fig. 5.10).

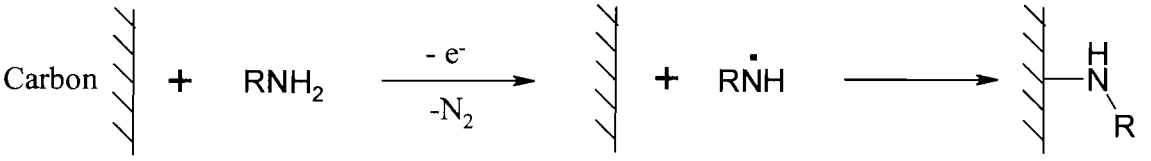


Figure 5.10 Modification of glassy carbon with primary amines via oxidation.

This method has been widely used for the modification of carbon surfaces with a wide range of amines. The grafting of diaminoalkanes onto the carbon surface using this methodology has led to their use as a handle for further coupling chemistry, so these were applied for glassy carbon modification.

Two different chain lengths were investigated; ethylenediamine and decanediamine modification was compared. These chain lengths were selected as it has been shown that for optimal control over the efficient transfer of electrons to the protein active site,

the cofactor must make efficient contact with the surface. For this efficient transfer, there are two possibilities; either tethering the protein very close to the surface may reduce the distance between the electrode surface and cofactor to a minimum giving the optimal electron transfer rate (*50*) (hence the use of ethylenediamine), or a longer tether (i.e. decanediamine) may lead to a weak interaction between the protein and the electrode surface, but this in turn allows the protein freedom to explore many different orientations without being desorbed into solution (*51*). Both of these approaches have led to fast electron transfer for cytochrome c, either by immobilisation directly on a gold surface via a surface cysteine residue, or by immobilisation onto 6-mercaptohexan-1-ol.

5.3.1 Functionalisation of glassy carbon via oxidation of ethylenediamine

The glassy carbon surface was derivatised with ethylenediamine by electrolysis at 2 V vs SCE. Cyclic voltammetry between 0 and 2 V vs SCE (Fig. 5.11), shows the large, broad, irreversible oxidation peak at +1.3 V vs SCE, which is consistent with the literature (*49*).

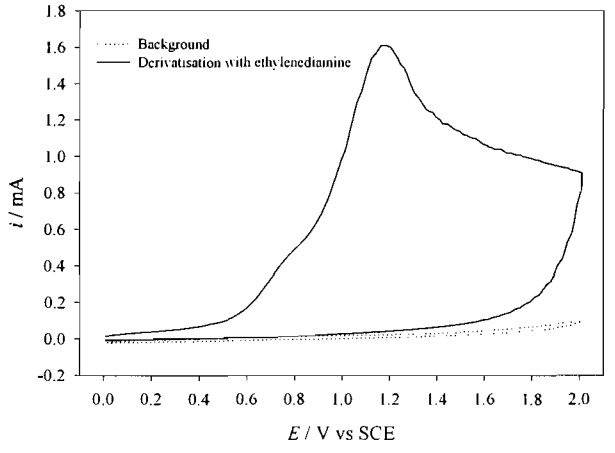


Figure 5.11 Derivatisation of a 7 mm diameter glassy carbon electrode with ethylenediamine by cyclic voltammetry in 0.2 M ethylene diamine in 0.1 M NBu_4BF_4 / acetonitrile, $\nu = 20$ mV / s, Ar, RT.

This monoelectronic oxidation of both amines has been shown to lead to surface reaction in three ways; ethylenediamine can attach via one amine, by both amines or can further react to polymerise (Fig 5.12).

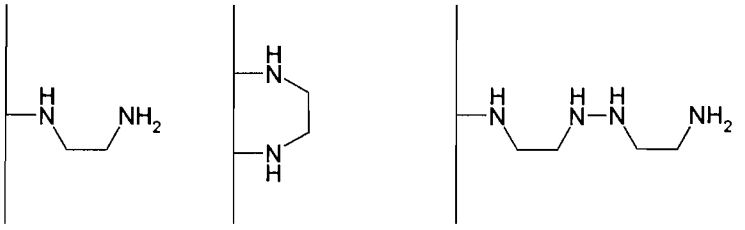


Fig 5.12 Modes of attachment of ethylenediamine to a glassy carbon surface.

Due to this mixed derivatisation, a relatively low proportion of the surface might be expected to be reactive towards further derivatisation. Attempts at conventional amide coupling as described in section 5.2.2 were unsuccessful also, no derivatisation of the surface with electroactive groups was observed. In the literature, only 1 % of the amine groups on the electrode surface were reported to be free for further chemistry (*52*), therefore this would be the maximum coupling yield expected, this result is therefore not very surprising.

Derivatisation with long chain alkanediamines has been shown to result only in attachment via one amine group, highly improving the loading or the surface with reactive primary amine moeties, improving the chance of further reaction.

5.3.2 The nature of the glassy carbon surface

Glassy carbon is manufactured in a two step process, first, the bulk material is moulded from dry or liquid phenol formaldehyde resin, followed by carbonisation by thermal treatment for a period of time that may be as long as 3-4 months. This results in a form of carbon which is amorphous, and unlike graphite the carbon layers are not ordered but are 'crumpled' (*53*) as in Fig. 5.13.

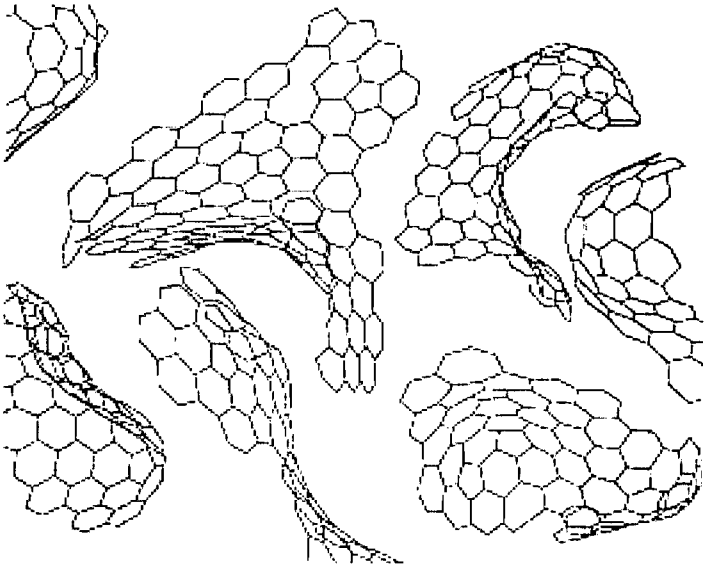


Figure 5.13 Representation of the glassy carbon surface

The bare, polished glassy carbon surface will typically contain a 7 – 20 % O/C ratio, with various functionalities being introduced through the polishing process (*54*).

5.3.3 Surface characterisation by X-ray photoelectron spectroscopy (XPS)

The composition of the glassy carbon surface was analysed by XPS using a Scienta ESCA300 photoelectron spectrometer (Daresbury, UK). For this technique, samples were prepared on glassy carbon plates (10 × 10 × 1 mm), to which an electrical contact was made using silver paint / insulated electrical wire, coated with epoxy resin to hold it in position, this could then easily be removed prior to analysis. Glassy carbon plates were prepared for derivatisation by polishing in alumina slurries to 0.3 μm particle size and prior to derivatisation, the surfaces were further cleaned by sonication in water.

The bare surface was characterised by XPS (Fig 5.14, 5.15) and the surface composition is given in table 5.1.

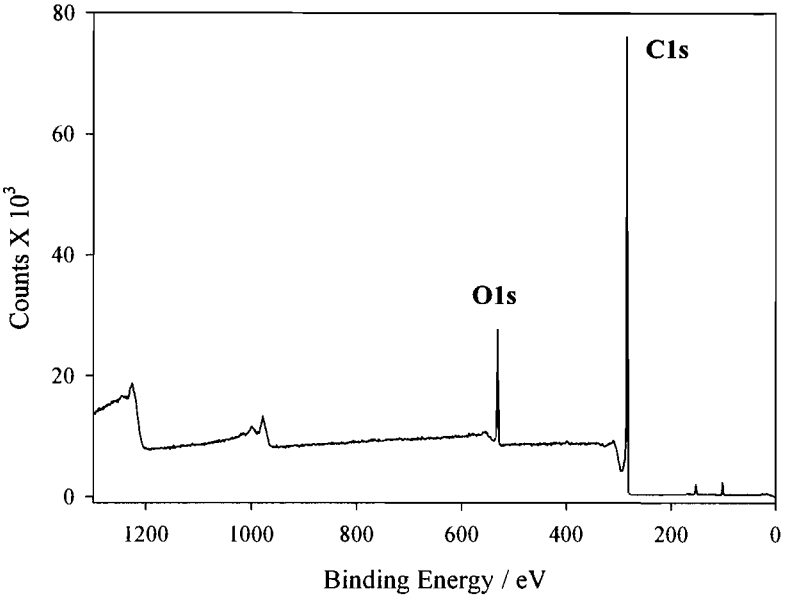


Figure 5.14 Bare glassy carbon. XPS Survey; slit = 1.9 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

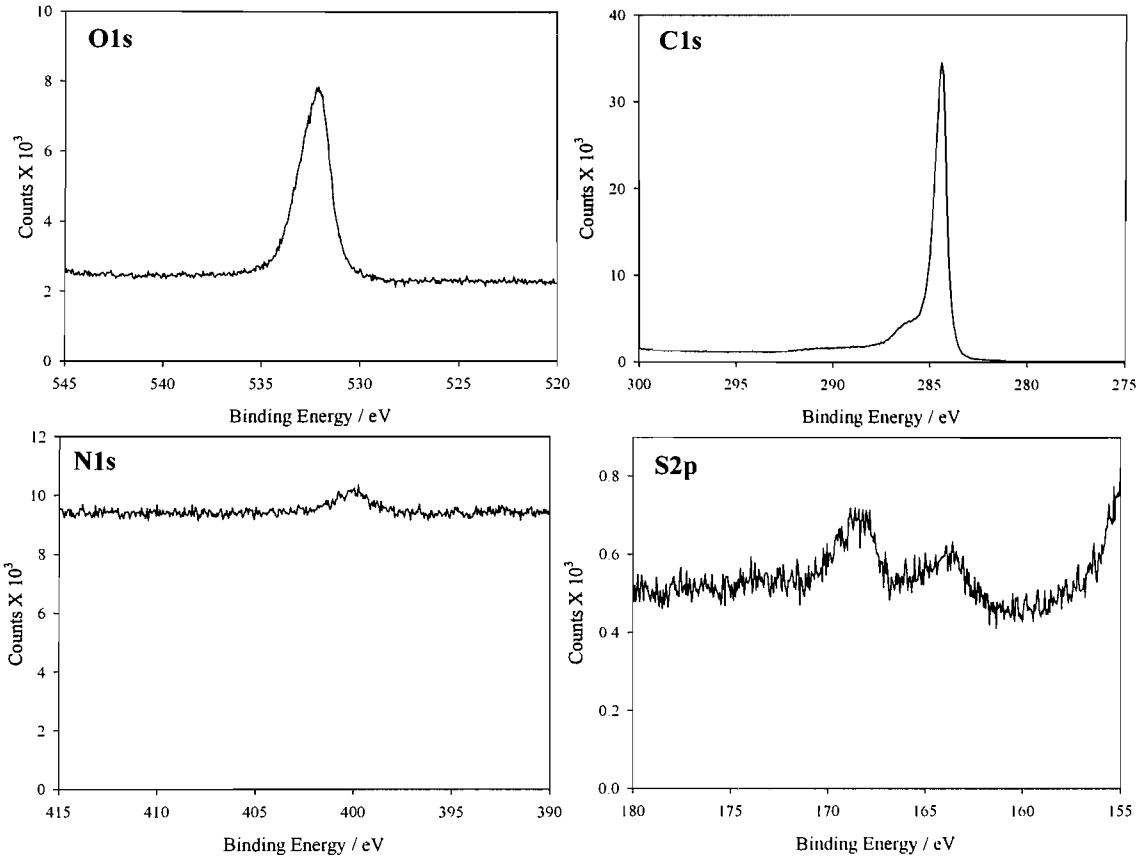


Figure 5.15 Bare glassy carbon. XPS core level region scans; slit = 1.9 mm, Take off angle = 90 °, step energy = 0.050 eV, time per step = 0.5 s for O1s and C1s, 1 s for N1s and S2p. X-ray energy = 1487 eV. Number of scans; O1s = 1, C1s = 1, N1s = 2, S2p = 2.

Atom	Conc / atom %
C	90
O	9.2
N	0.6
S	0.2

Table 5.1 Surface concentration of atoms on the bare glassy carbon surface

Fig. 5.14 is a broad survey spectrum of the bare glassy carbon surface at low resolution, this identifies the elements present within the surface whilst Fig. 5.15 shows high resolution spectra of peaks corresponding to each element, so that fine structure resulting from bonding may be observed. The survey shows the main constituents of the bare glassy carbon surface to be carbon and oxygen, with the relative quantity of

oxygen (9.2 %) within the surface layer in agreement with reported values. The high resolution spectra show a small quantity of sulphur present. The position of the peak at 165 eV represents a sulphur environment in which the atom is bonded to an element of low electronegativity in contrast with the larger peak at 168 eV, the shift is indicative of bonding to an element such as oxygen, representing a small quantity of oxidised sulphur on the surface, most probably as SO_3^{2-}. This background sulphate impurity is observed in all of the XPS spectra shown in this chapter, and is most likely residual from the polymer used as a starting material for the glassy carbon manufacturing process. The C1s spectrum shows a major peak at 284.4 eV a peak which is indicative of graphitic carbon (*53*), the broad peak at 286.5 eV arises from covalent bonding with oxygen in the form of alcohols and phenolic moieties introduced during the polishing process (*54*).

5.3.4 Derivatisation of glassy carbon by the oxidation of decanediamine

Diaminodecane (DAD) is covalently attached to the glassy carbon surface via oxidation of the primary amine to its cation radical followed by reaction with edge-plane sites at the GC surface (*55*). This oxidation is observed in Fig. 5.16 as a broad peak from +0.4 - +1.3 V vs Ag / Ag^+. The charge under the broad peak (over the three cycles) is equivalent to a 100 fold excess (for a 1 e^- oxidation, $2 \times 10^{-9} \pm 5 \times 10^{-10}$ moles / cm^2 over >100 experiments) to that required to form a close packed monolayer of DAD on the GC surface.

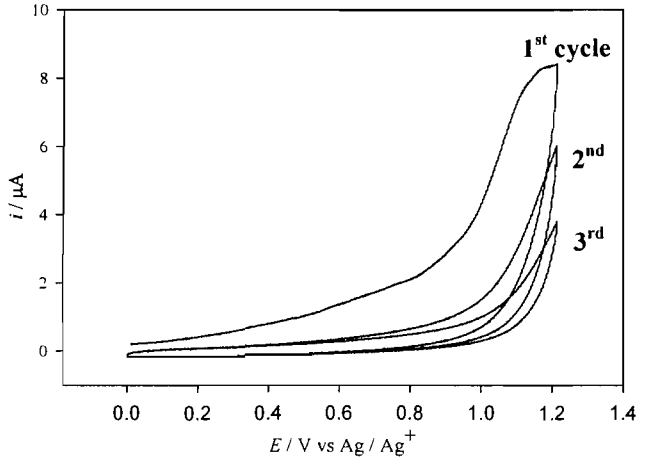


Figure 5.16 Voltammetry (first 3 cycles as indicated) of a bare 3 mm diameter glassy carbon electrode in an ethanolic solution of 2 mM diaminodecane / 0.1 M $LiClO_4$, RT, Ar, Pt gauze CE, vs Ag / Ag^+, scan rate = 20 mV / s.

This derivatisation step was characterised by voltammetry in ruthenium hexamine. A redox molecule such a ruthenium hexamine is often utilised as a probe of surface derivatisation. When the surface is non-modified, the rate of electron transfer is fast, with reversible behaviour and so should obey the Randles-Sevičk equation:

$$I_p = 2.69 \times 10^5 n^{3/2} A C D^{1/2} \nu^{1/2}$$

Equation 5.2 The Randles-Sevičk equation for T = 298 K.

Where I_p is the peak current density, n is the number of electrons in the process, C is the concentration of the redox species, D is the diffusion coefficient and ν is the scan rate. The peak current density (I_p) should therefore be proportional to the square root of the scan rate ($\nu^{1/2}$) for a reversible system. If the surface has been derivatised with a long chain alkane, the separation between redox molecule and surface will be significantly increased assuming a high coverage of a compact film. The relationship between the rate of electron transfer and distance between electron donor and acceptor is described by the 'superexchange model' (*56*), in which the rate of electron transfer decreases exponentially with increasing separation. If the rate of electron transfer is decreased, it is expected that as the scan rate is increased that the rate of mass transport increases and becomes comparable to the rate of electron transfer. The most noticeable effect of which is that the peak separation increases. An example of this is shown in Fig. 5.17, which compares voltammetry at the bare and derivatised glassy carbon electrode in ruthenium hexamine solution. The potential is cycled from +0.5 to -0.6 V vs SCE in aqueous solution, on the forward cycle a sharp peak is observed at -0.2 V vs SCE, which represents the single electron reduction of Ru^{3+} to Ru^{2+}, on the return cycle, a symmetrical oxidation peak is observed with peak separation of 60 mV which represents the oxidation of Ru^{2+} to Ru^{3+}.

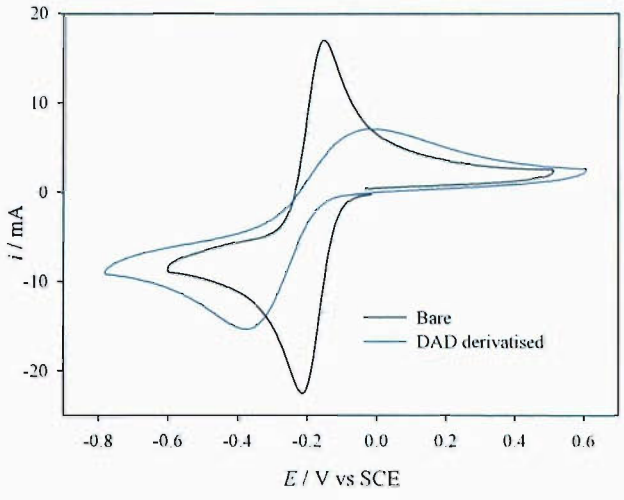


Figure 5.17 Characterisation of the glassy carbon electrode by voltammetry in 2 mM $Ru(NH_3)_6.Cl_2$ in 0.1 M KCl pH 5.0 (Bare / Derivatised 7 mm diameter glassy carbon working electrode), limits -0.6 to +0.5 V vs SCE, scan rate = 50 mV/s, Pt gauze CE, Ar, RT.

Once derivatised with diaminodecane, voltammetry in ruthenium hexamine shows a significant change in the voltammogram shape. The reduction and oxidation peaks are beginning to lose their symmetry and the peak separation is greatly increased to ~ 300 mV, suggesting that the rate of electron transfer is slower due to separation of the redox couple from the surface.

Voltammetry at a range of scan rates shows that in both cases the data (Fig. 5.18, 5.19), is in agreement with the Randles-Sevičk equation, $I_p \alpha \nu^{1/2}$. The peak separation at the bare glassy carbon electrode is essentially independent of scan rate (60 mV at 2 mV / s to 80 mV at 500 mV / s) but in the derivatised case, the voltammetry shows quasi reversible behaviour, as E_p varies with scan rate.

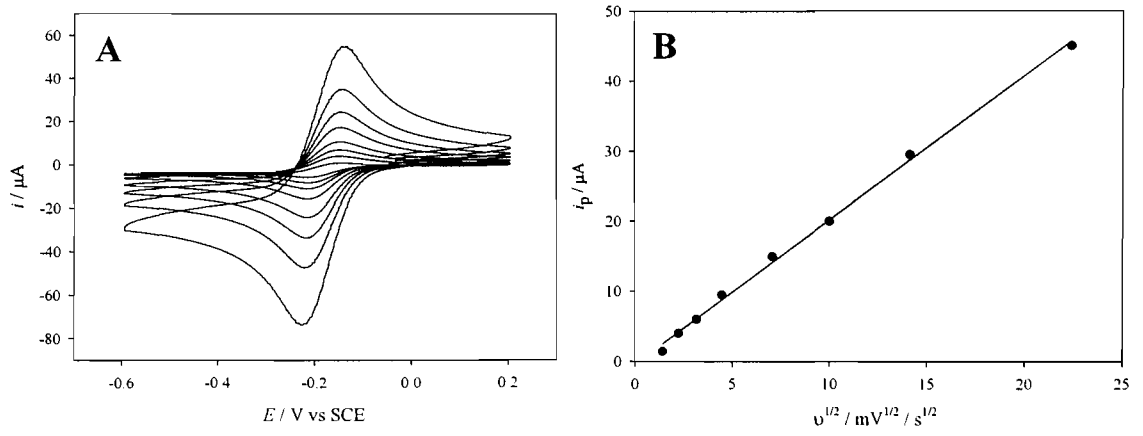


Figure 5.18 (A) Voltammetry (first cycle at each scan rate) of bare 3 mm diameter glassy carbon electrode in 2 mM $Ru(NH_3)_6.Cl_3$ / 0.1 M KCl, RT, Ar, Pt gauze CE, vs. SCE, varying scan rate; 2, 5, 10, 20, 50, 100, 200, 500 mV / s (A). (B) Plot of I_p vs. $\nu^{1/2}$.

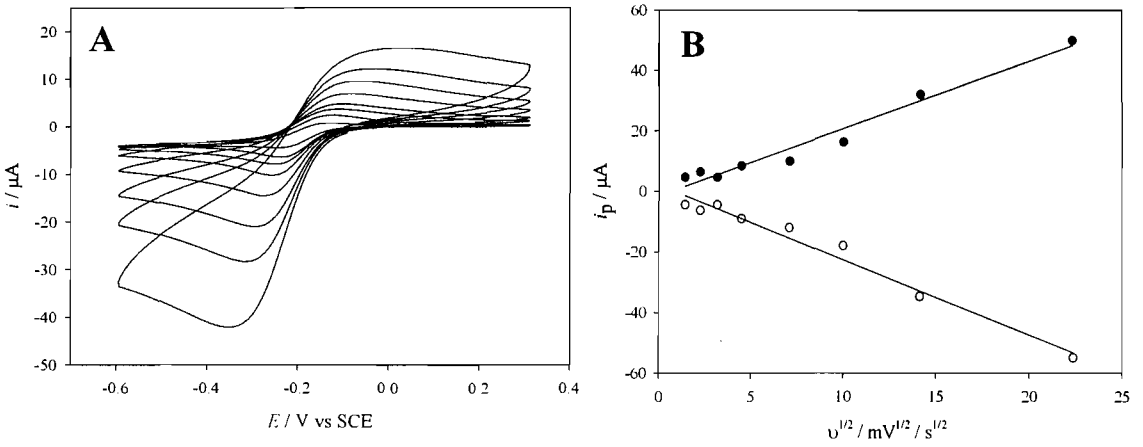


Figure 5.19 (A) Voltammetry (1st scan at each scan rate) of DAD derivatised 3mm diameter Glassy Carbon electrode in 2 mM $Ru(NH_3)_6.Cl_3$ / 0.1 M KCl, RT, Ar, Pt gauze CE, vs. SCE, varying scan rate; 2, 5, 10, 20, 50, 100, 200, 500 mV / s. (B) Plot of I_p vs. $\nu^{1/2}$.

From the slope of the plot of (Fig. 5.18 B) the experimental value of D = 3.39 × 10^{-6} cm^2 / s as calculated using the Randles-Sevičk equation (equation 5.2) using an electrode area of 0.0707 cm^2, n = 1 and $[Ru(NH_3)_6.Cl_3]$ = 2 × 10^{-6} moles / cm^3. Data from the ruthenium hexamine voltammetry at the bare and DAD derivatised glassy

carbon surface at scan rates of 500, 200 mV / s could be simulated using DIGISIM (Fig. 5.20) giving values of D of and k as given in tables 5.2 and 5.3.

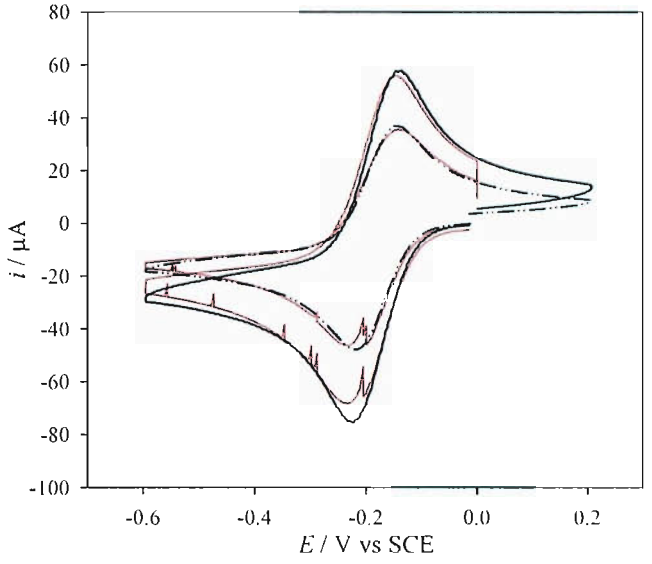


Figure 5.20 DIGISIM simulated voltammograms for the $Ru^{2+/3+}$ redox couple at the bare glassy carbon electrode.

ν / V / s	E_o / V	α	k / s × 10^{-3}	D cm^2 / s × 10^{-6}
0.5	-0.176 ± 0.005	0.47 ± 0.09	6.33 ± 0.12	8.82 ± 0.6
0.2	-0.183 ± 0.002	0.547 ± 0.024	4.38 ± 0.24	10. 1 ± 0.2

Table 5.2 Parameters as simulated by DIGISIM for bare glassy carbon electrode.

ν / V / s	E_o / V	α	k / s × 10^{-3}	D cm^2 / s × 10^{-6}
0.5	-0.175 ± 0.012	0.69 ± 0.05	0. 448 ± 0.15	4.36 ± 0.36
0.2	-0.198 ± 0.01	0.676 ± 0.046	0. 41 ± 0.11	5.87 ± 0.41

Table 5.3 Parameters as simulated by DIGISIM for DAD derivatised glassy carbon electrode.

The value of D in the case of the bare and DAD derivatised electrode is close to the experimental and reported vale of 6 × 10^{-6} cm^2 / s (*57*). The electron transfer rate constant, *k*, is approximately ten times less at the diaminodecane modified glassy carbon electrode. The increased distance between redox group and electrode surface

was expected to decrease the rate of electron transfer as is observed. The rate of electron transfer is expected to decrease exponentially with increasing distance from the surface as described in equation 5.3 (*58*).

$$k^0_{app} = k^0 \exp(-\beta d)$$

Equation 5.3

Where β is the tunnelling parameter (for alkyl chains, β = 0.79 $Å^{-1}$ (*59*)), k^0_{app} is the apparent rate constant at zero overpotential, *d* is the barrier thickness (number of bonds) and k^0 is the rate constant at the bare electrode at zero overpotential. The theoretical value of $k^0_{app} = 3.82 \times 10^{-7}$ cm^2 / s, assuming that the separation of ten C-C bonds and two C-N bonds approximately corresponds to a vale of $d = 12$ and that the alkyl chain is in a close packed monolayer. This value is three orders of magnitude slower than that simulated, this may be expected however, as the oxidative derivatisation of the glassy carbon surface results only in the derivatisation of the edge-planar regions of the glassy carbon surface, accounting for 18.6 % of the surface composition (*60*). This will result in a surface with a mixture of derivatised / non derivatised regions, the rate of electron transfer will be faster in the non-derivatised regions therefore increasing the observed rate.

5.3.5 Stability of the decanediamine derivatised surface

The intention to use the decanediamine modified electrodes for conventional coupling chemistry led to the investigation of their stability in a variety of organic solvents. Derivatised electrodes were left to soak overnight in a range of solvents, before drying and analysis of the surface by cyclic voltammetry in ruthenium hexamine (Fig. 5.21). The voltammetry shows that in all cases the blocking properties of the surface were decreased. The effect was least in water and dichloromethane and most pronounced in dimethylformamide. This loss of blocking can be attributed to the removal of the diamine from the electrode surface. It was however reported that diaminodecane should only react with the edge-planar regions of the electrode surface, which accounts for 18.6 % of the electrode surface (*60*). The blocking would therefore be expected to be poor at

such a coverage. The loss of blocking observed may then be due to the desorption of adsorbed diaminodecane on the electrode surface.

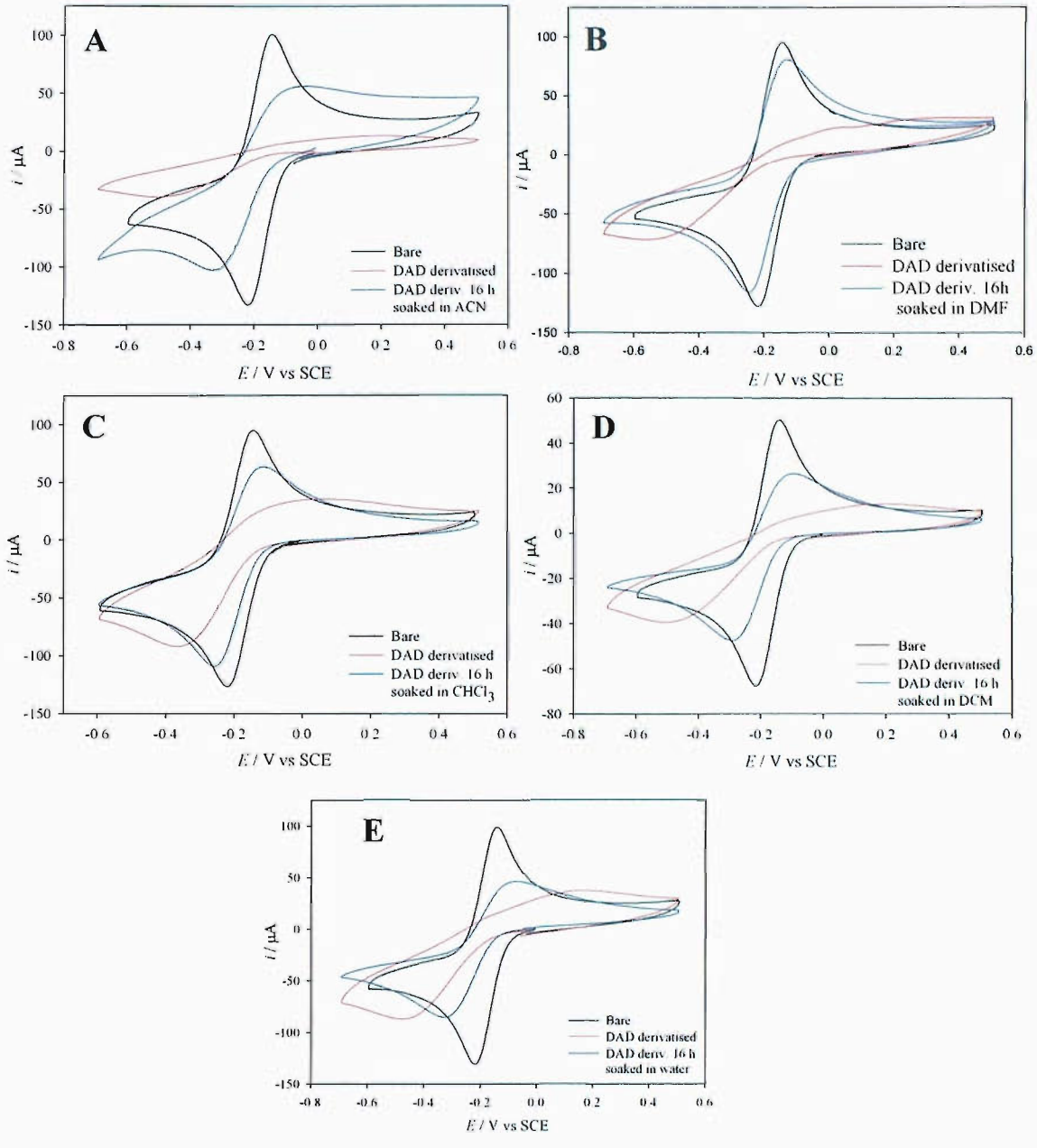


Figure 5.21 3 mm diameter glassy carbon derivatised by voltammetry in an ethanolic 2 mM diaminodecane / 0.1 M $LiClO_4$ solution, 3 cycles, limits 0 - + 1.3 V vs Ag / Ag^+, Pt gauze CE, Ar, RT, ν = 20 mV / s. Post voltammetry electrodes were incubated in (A) DMF, (B) DCM, (C) $CHCl_3$, (D) ACN, (E) H_2O for 16 h, RT. Characterisation in $Ru(NH_3)_6.Cl_2$ 2 mM in 0.1 M KCl pH 5.0 (Bare / Derivatised 3 mm diameter GC WE), limits -0.6 - +0.5 V vs SCE, scan rate = 50 mV/s, Pt gauze CE, Ar, RT.

The diaminodecane modified glassy carbon electrode was characterised by XPS.

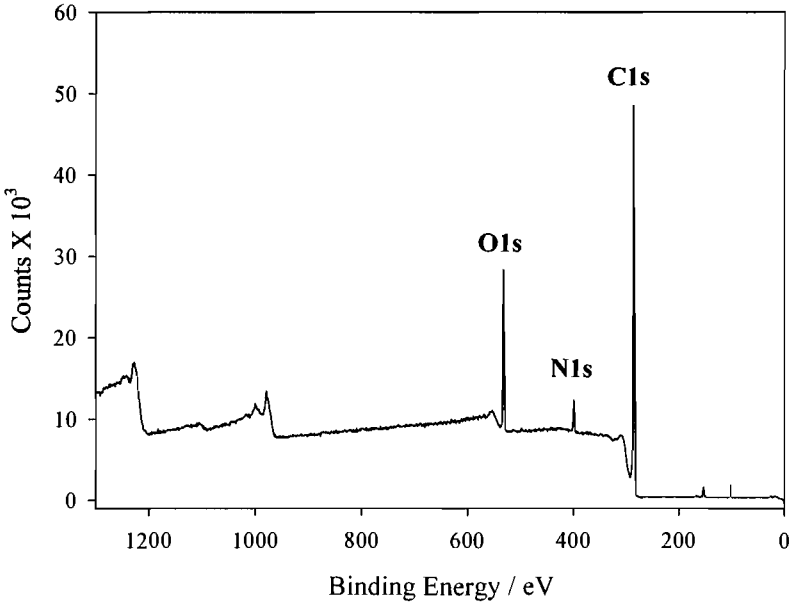


Figure 5.22 Diaminodecane modified glassy carbon. XPS Survey; slit = 1.9 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

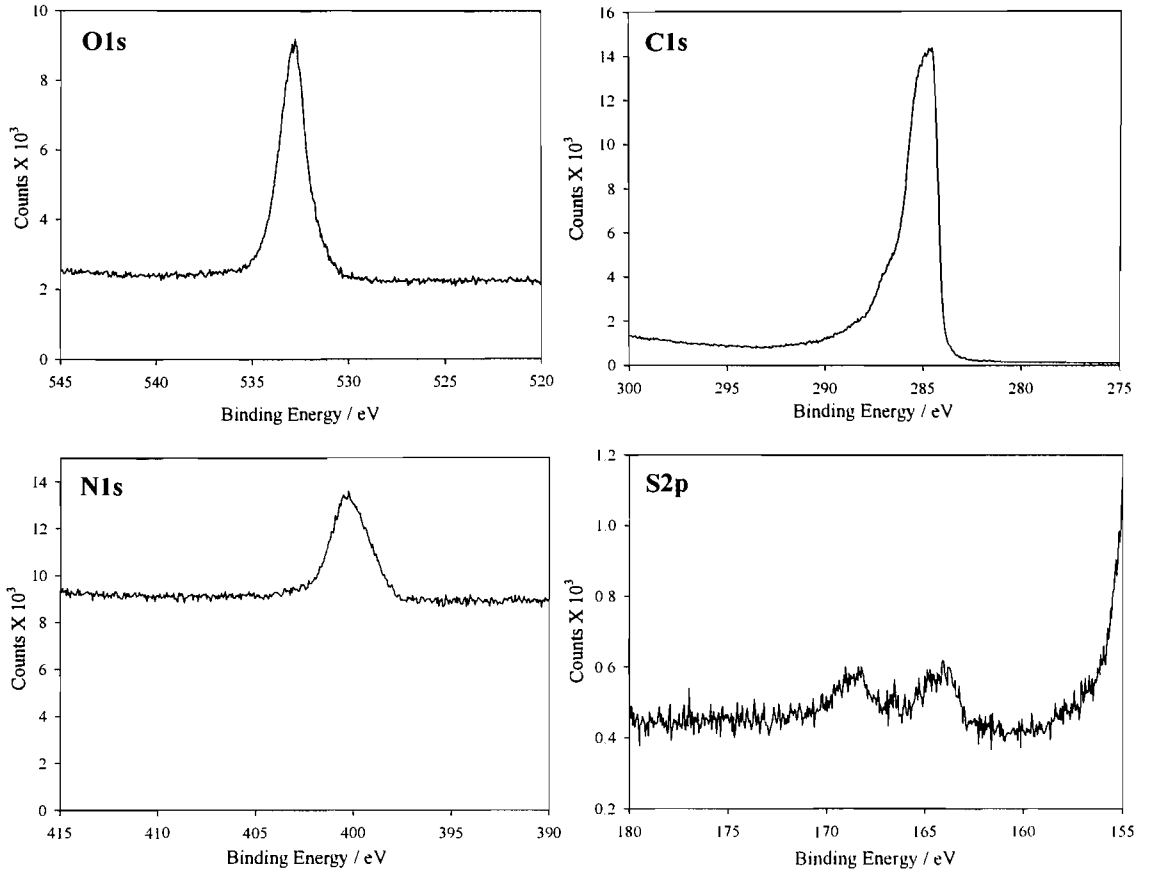


Figure 5.23 Diaminodecane modified glassy carbon. XPS core level region scans; slit = 1.9 mm, TOA = 90°, step energy = 0.050 eV, time per step = 0.5 s for O1s and C1s, 1 s for N1s and S2p. X-ray energy = 1487 eV. Number of scans; O1s = 1, C1s = 1, N1s = 2, S2p = 2.

Atom	Conc / atom %
C	83.5
O	12.1
N	4.2
S	0.2

Table 5.4 Surface concentration of atoms on the diaminodecane modified glassy carbon surface.

The first change in comparison with that observed for bare glassy carbon in the XPS scan (Fig. 5.22) is that there is the appearance of a peak at 400 eV, which corresponds to the N1s peak. This shows that the surface has been derivatised with diaminodecane.

The surface composition shows the nitrogen content to have increased from 0.6 to 4.2 % of the atom content of the surface.

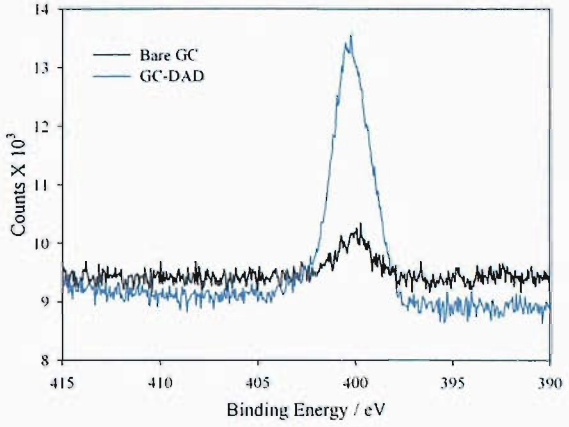


Figure 5.24 Comparison of the N1s region of the bare and diaminodecane modified glassy carbon surface.

Comparison of the bare vs derivatised N1s region scan show the significant increase in the intensity of the N1s peak (Fig. 5.24). The O1s and S2p regions (Fig. 5.23) have remained as observed in the bare glassy carbon sample, the sulphur content of the surface has remained the same and the oxygen content is within that expected for a mechanically polished surface.

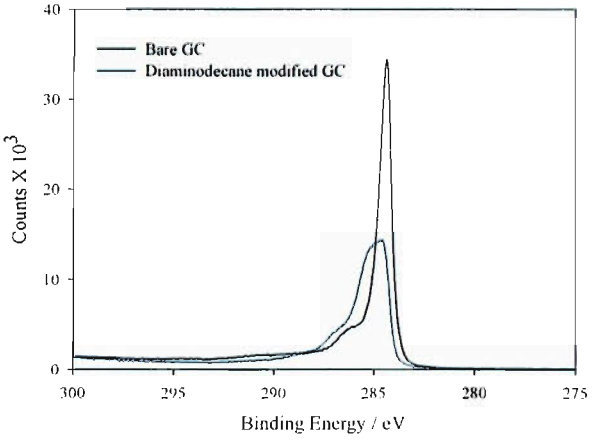


Figure 5.25 Comparison of bare and diaminodecane modified C1s regions.

The C1s region exhibits (Fig. 5.25) significantly changed properties to the bare glassy carbon surface. The peak at 284.5 eV has significantly reduced in size, this peak represents the glassy carbon aromatic C-C bonds, its reduction is size is due to the

diaminodecane blocking the surface from the X-rays which can only penetrate 10 nm into the sample surface. The broadness of the major peak is due to a contribution from a peak at 285.0 eV from the long chain alkane C-C bond atoms. The shoulder at 286.6 eV observed in the bare glassy carbon spectrum remains, and there is a contribution from the C-N carbon atoms under this shoulder to the main peak at 285.8 eV.

5.3.6 Coupling of decanediamine with ferroceneacetic acid

The ability to attach carboxylic acids to the surface using conventional amide coupling chemistry was first investigated using ferroceneacetic acid, a molecule which has been coupled to long chain diamine modified surfaces in the literature (Fig. 5.26). The coupling was achieved using the water soluble coupling reagent EDC.

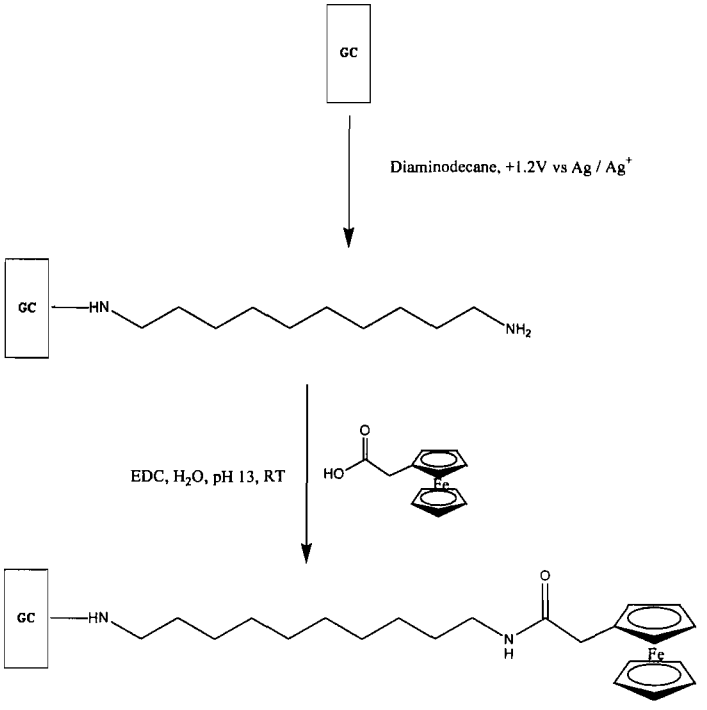


Figure 5.26 Coupling of ferroceneacetic acid to the diaminodecane modified glassy carbon electrode.

The success of the coupling reaction was judged by cyclic voltammetry in aqueous 0.1 M $LiClO_4$ solution (Fig. 5.28). This voltammetry shows the ferrocene reversible redox couple ($E_{1/2}$ = +300 mV, peak separation remains independent of scan rate at ~90 mV). Coupling resulted in a coverage of $6.1 \times 10^{-11} \pm 5 \times 10^{-12}$ moles / cm^2, over > 25

experiments, (25 %) compared to a theoretical maximum coverage of 2.4×10^{-10} moles / cm^2 (Fig. 5.27).

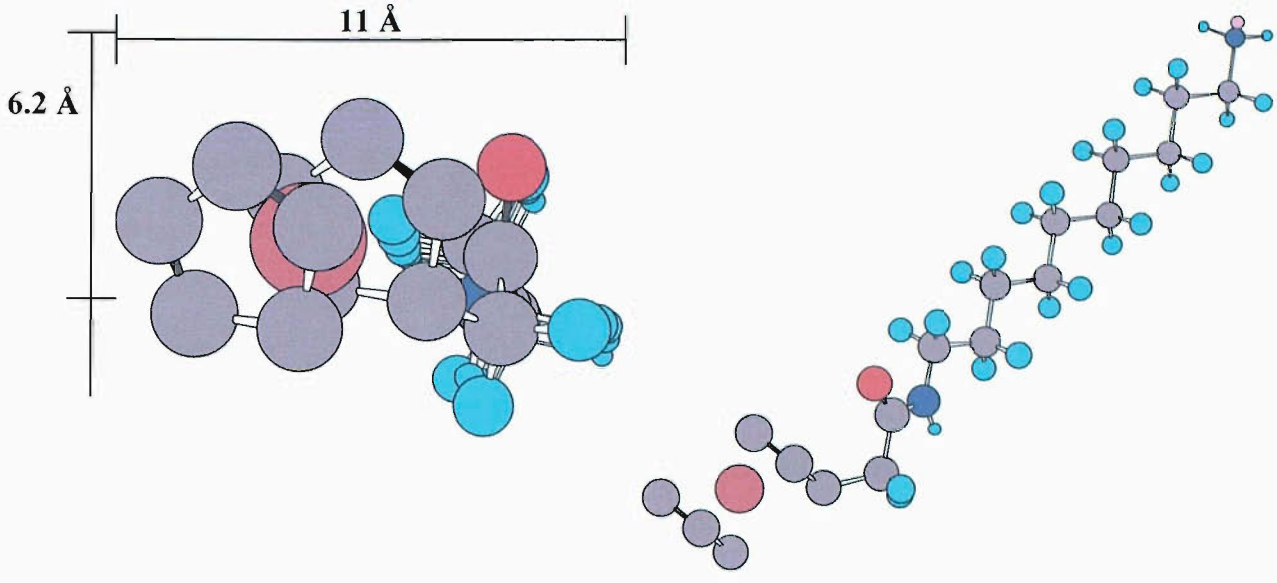


Figure 5.27 Modelled dimensions of the diaminodecane-ferroceneacetic acid modification for surface coverage calculations.

The glassy carbon electrode surface is composed of a mixture of basal planes and edge planes (Fig. 5.13), the relative amounts of edge plane orientations on the surface has recently been elucidated (*60*) as 18.6 % of the surface area. If covalent bonding of diaminodecane is only possible at edge plane sites, then the maximum coverage of DAD can be estimated to be 18.6 % of the maximum coverage of the surface area of the electrode = 4.1×10^{-10} moles / cm^2.

This then leads to the maximum coverage of FcAA on the electrode surface in the orientation as modelled in Fig. 5.27 (perpendicular to the surface) of 4.5×10^{-11} moles / cm^2. The observed coverage of $6.1 \times 10^{-11} \pm 5 \times 10^{-12}$ moles / cm^2 is close to that predicted for covalent attachment to decanediamine attached to the edge-plane sites at the electrode surface.

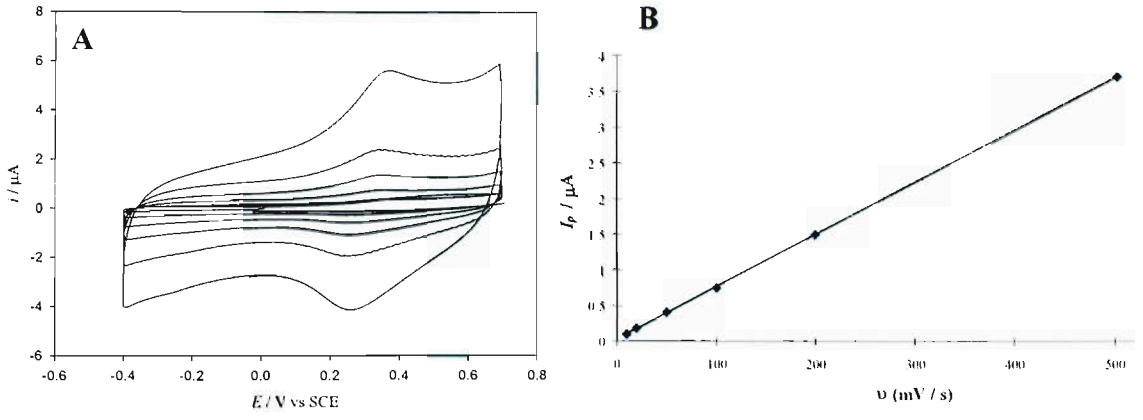


Figure 5.28 (A) Glassy carbon 3 mm diameter electrode derivatised by (i) oxidation of decanedimine (ii) couping with FcAA (10 mM FcAA, 10 mM EDC, aqueous pH 14), voltammetry at scan rates 2, 5, 10, 20, 50, 10, 200, 500 mV / s in 0.1 M $LiClO_4$, sweep limits -0.2 - + 0.7 V vs SCE, Ar, RT, Pt gauze CE. (B) Plot of I_p vs. ν.

In the case of ferroceneacetic acid where the electroactive species has been attached to a chemically modified electrode, different behaviour is observed to that of a species in solution (as observed for ruthenium hexamine voltammetry). If the species is directly attached to the electrode surface, the reduction / oxidation peaks of the voltammogram will be sharp and symmetrical, with charges associated with oxidation / reduction equal. The analysis of this type of system is simplified as mass transport effects are not considered. In this case the peak current density is given by:

$$I_p = (n^2 F^2 \Gamma_R / 4RT).\nu$$

Equation 5.4 I_p = peak current, n = number of electrons involved in the electron transfer, Γ_R = surface coverage of the molecule as calculated from reduction charge, ν = scan rate.

This shows $I_p \alpha \nu$ as opposed to $I_p \alpha \nu^{1/2}$ as for the reversible reaction in solution. This relationship is observed for the ferroceneacetic acid coupled electrode (Fig. 5.28 B), $I_p \alpha \nu$ with the line passing through the origin. The peaks however do not exhibit typical directly adsorbed behaviour (i.e. symmetrical peak shape and no peak separation). This is due to the large separation between redox molecule and electrode surface and hence

slowed electron transfer kinetics are observed, similar behaviour to that of ruthenium hexamine at the diaminodecane modified electrode.

Determination of the FcAA electrochemical rate constant

It is possible to estimate the experimental electron transfer rate constant by plotting the positions of the oxidation and reduction peaks of the ferrocene voltammograms shown in Figure 5.29 (A) as a function of the log of the scan rate, in a plot referred to as a 'trumpet plot', as in Fig. 5.29 (B) below.

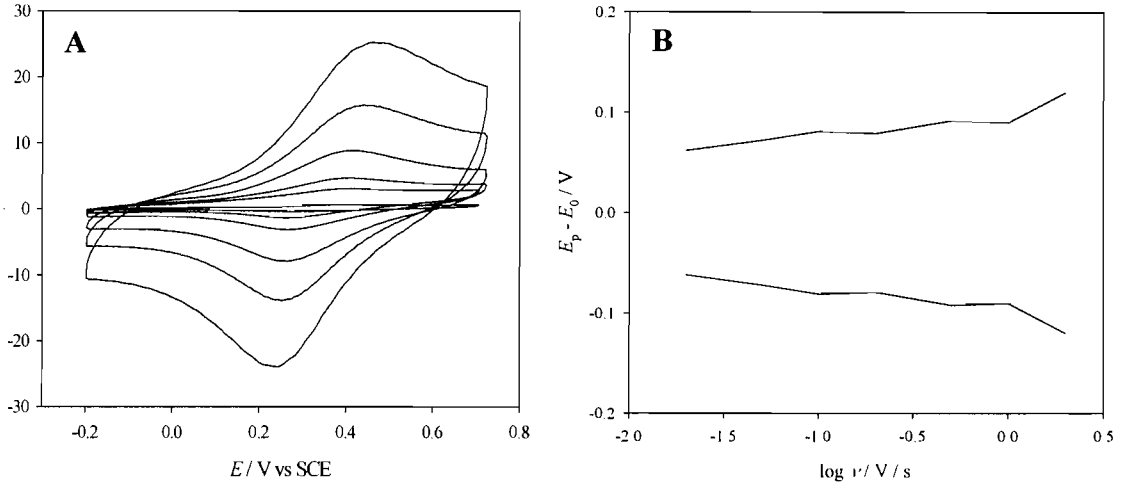


Figure 5.29 (A) Glassy carbon 3 mm diameter electrode derivatised by (i) oxidation of decanediamine (ii) coupling with FcAA (10 mM FcAA, 10 mM EDC, aqueous pH 14), voltammetry at scan rates 20, 50, 100, 200, 500, 1000, 2000 mV / s in 0.1 M $LiClO_4$, sweep limits -0.2 - + 0.7 V vs SCE, Ar, RT, Pt gauze CE. (B) Plot of $E_p - E_0$ vs log ν.

At the scan rates used in this experiment, the peak separation was not sufficient to apply the method of Laviron (*61*) in which the rate of electron transfer may be calculated directly from a plot of E_p vs logν. The rate of electron transfer was therefore estimated by comparing the trumpet plot above to theoretical curves created using ideal calculated positions of oxidation and reduction peaks as a function of scan rate (log scale), for a simple electron transfer process with various rate constants, where k_0 = 0, 1, 2, 3, 4. Positions were predicted using the Butler-Volmer model, with n = 1, a = 0.5 and an $E_{1/2}$ of 0 V (*62*), and a rate constant of 10 s^{-1} extracted.

The ferroceneacetic acid derivatised surface was analysed by XPS.

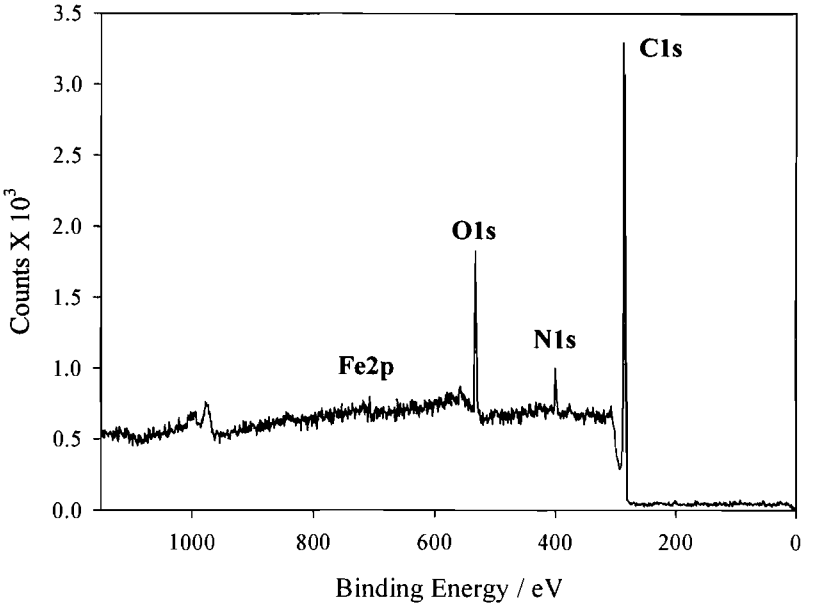


Figure 5.30 Glassy carbon electrode modified with diaminodecane-ferroceneacetic acid. XPS Survey; slit = 2.8 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

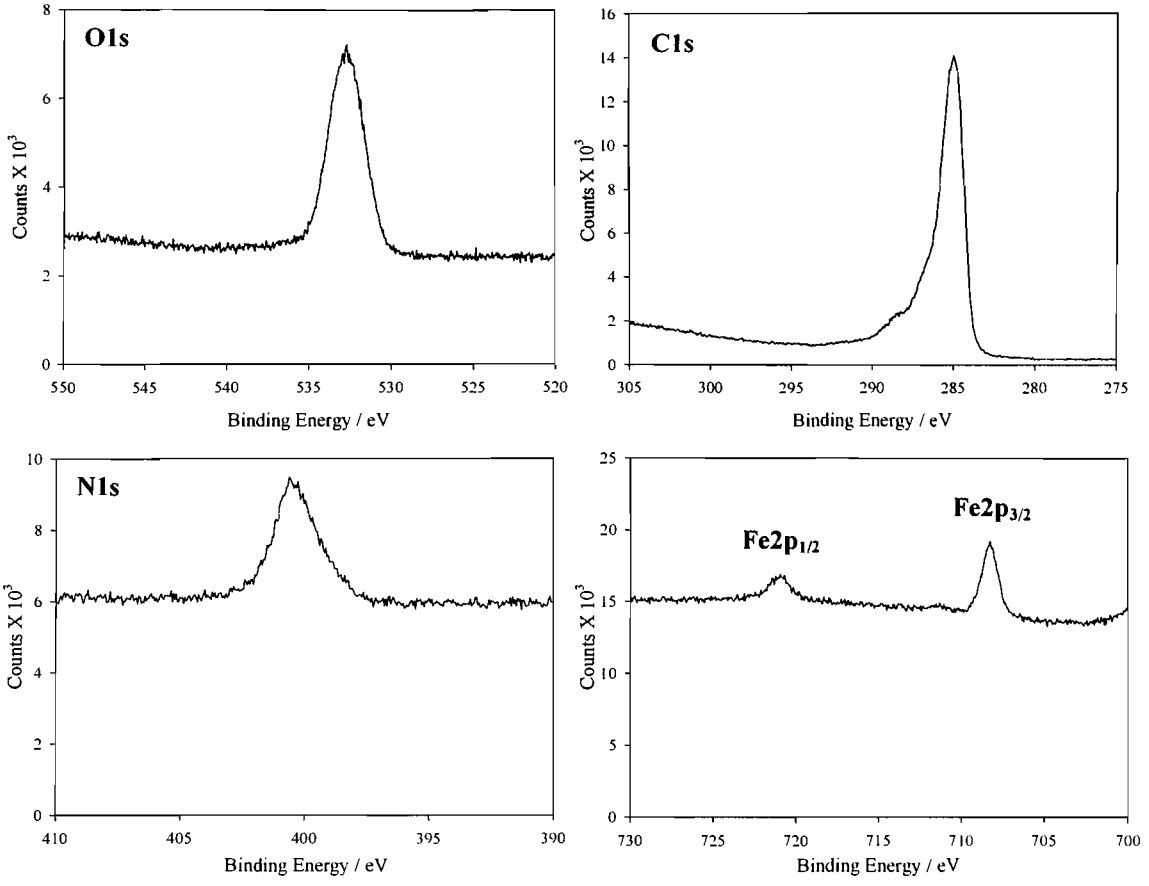


Figure 5.31 Diaminodecane-ferroceneacetic acid modified glassy carbon. XPS core level region scans; slit = 1.9 mm, TOA = 90°, step energy = 0.050 eV, time per step = 0.1 s, X-ray energy = 1487 eV. Number of scans; O1s, C1s = 2, N1s = 5, Fe2p = 10.

Atom	Conc / atom %
C	82.1
O	12.3
N	5.3
Fe	0.3

Table 5.5 Atomic composition of the diaminodecane-ferroceneacetic acid modified glassy carbon electrode.

The 1325 – 0 eV survey of the surface is very similar to that observed for the diaminodecane modified surface (Fig. 5.22). The atomic composition of the surface is given in table 5.5, as calculated for weighted areas under the relevant peaks in the

region scans (Fig. 5.31). The composition similarity with the diaminodecane modified surface demonstrates the reproducibility of this method of surface modification. The appearance of Fe on the electrode surface confirms the immobilisation of ferrocene, the proportion of Fe vs N at the electrode surface suggests a coupling efficiency of 11 % of the available amines at the surface.

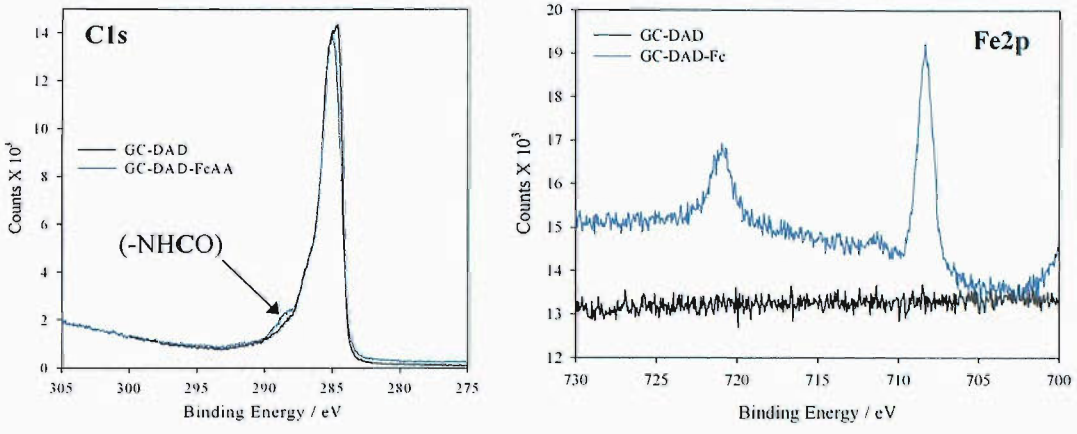


Figure 5.32 Comparison of C1s and Fe2p regions of the diaminodecane (GC-DAD) and diaminodecane-ferroceneacetic acid (GC-DAD-FcAA) modified surfaces.

The major differences in the XPS data are in the comparison of the C1s and Fe2p spectra between the diaminodecane derivatised and ferroceneacetic acid coupled electrodes. The C1s spectrum shows a shoulder appearing in the spectrum at 288.5 eV, which is consistent with the formation of amide bonds (-NHCO-), the Fe2p spectrum clearly demonstrates the appearance of $Fe2p_{1/2}$ and $Fe2p_{3/2}$ peaks. These changes indicate the immobilisation of an iron containing molecule via an amide bond at the electrode surface, confirming the derivatisation.

5.3.7 Coupling and deprotection of *N*-Boc-*S*-trityl-L-cysteine to a diaminodecane derivatised glassy carbon electrode

As stated earlier, the electrode surface must be derivatised with cysteine for protein immobilisation via intein mediated ligation (Fig. 5.33). After the successful carbodiimide coupling of ferroceneacetic acid to the diaminodecane modified glassy carbon surface, the coupling of an *N*-boc *S*-trityl protected cysteine derivative to the diaminodecane modified surface was attempted via PyBop coupling in DMF. Glassy carbon electrodes were derivatised with diaminodecane and directly placed into a DMF/ PyBop / diisopropylethylamine (DIPEA) mixture and left to react overnight before careful washing and sonication to remove adsorbed cysteine from the surface.

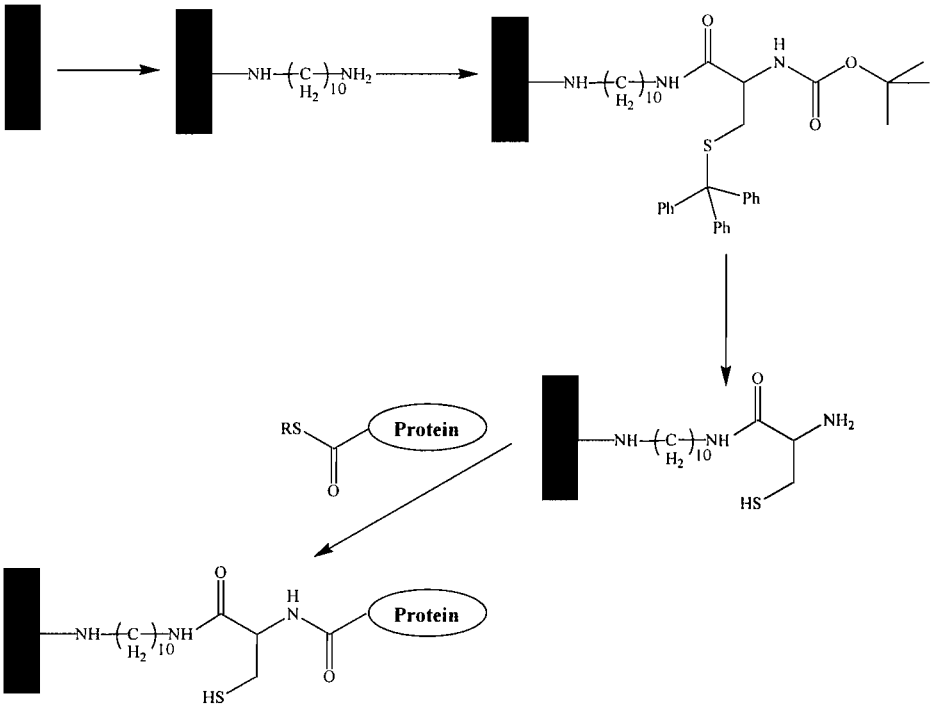


Figure 5.33 Proposed route to protein immobilisation via intein mediated ligation.

The resulting surface derivatisation was analysed by XPS.

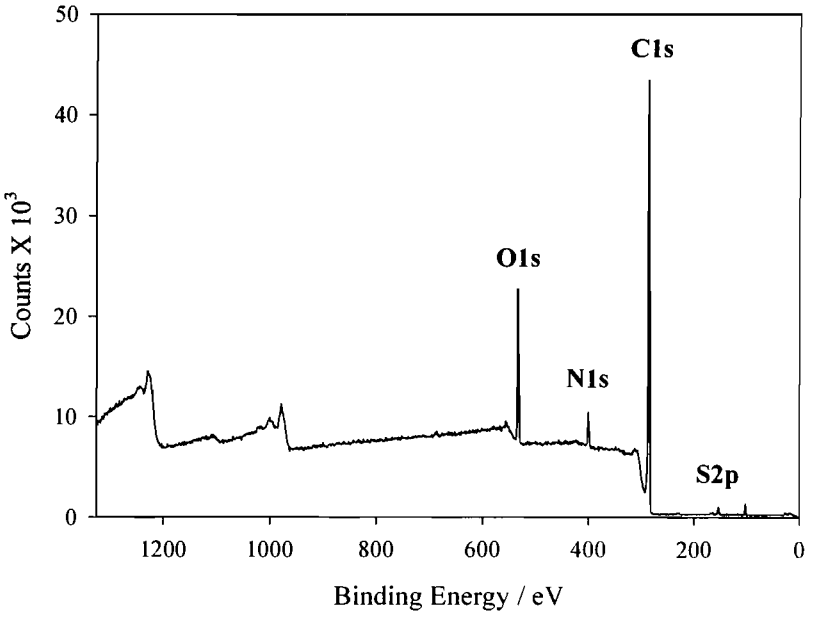


Figure 5.34 Glassy carbon electrode modified with diaminodecane-*N*-Boc-*S*-trityl-L-cysteine. XPS Survey; slit = 2.8 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

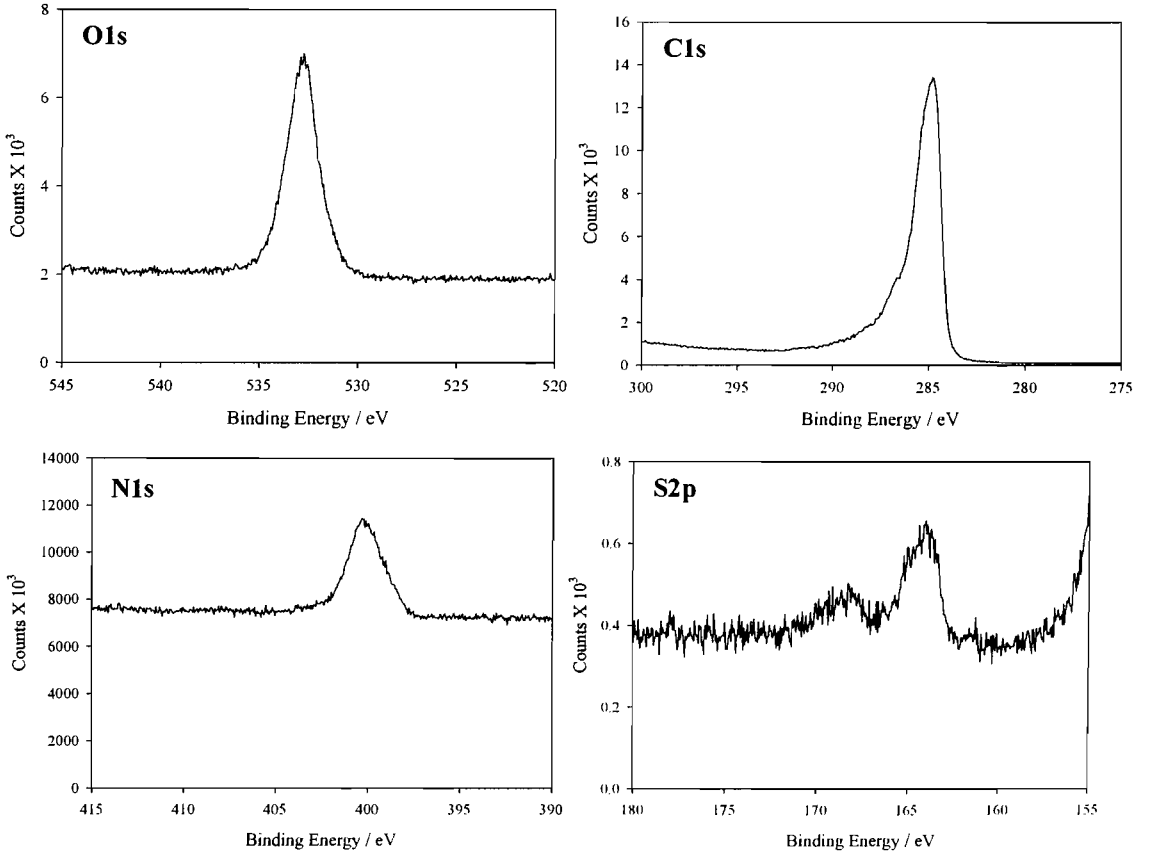


Figure 5.35 Diaminodecane-*N*-Boc-*S*-trityl-L-cysteine modified glassy carbon. XPS core level region scans; slit = 1.9 mm, TOA = 90°, step energy = 0.050 eV, step time = 0.1 s. X-ray energy = 1487 eV. Number of scans; O1s, C1s = 1, N1s and S2p = 2.

Atom	Conc / atom %
C	83.7
O	11.5
N	4.4
S	0.4

Table 5.6 Atomic composition of the diaminodecane-*N*-Boc-*S*-trityl-L-cysteine modified glassy carbon electrode.

The surface atomic content is again very similar to that of the diaminodecane modified surface (Fig. 5.22). The sulphur content has risen by 0.2 %, indicating a coupling efficiency with available amine groups of 9 %. An interesting observation was that although the reaction solvent for amide coupling is dimethylformamide, which was shown to alter the electrochemical blocking properties of the electrode (Fig. 5.21) thought to be due to stripping of the adsorbed diamine from the surface, this was not observed by XPS, suggesting that the covalently bonded decanediamine is stable to DMF exposure.

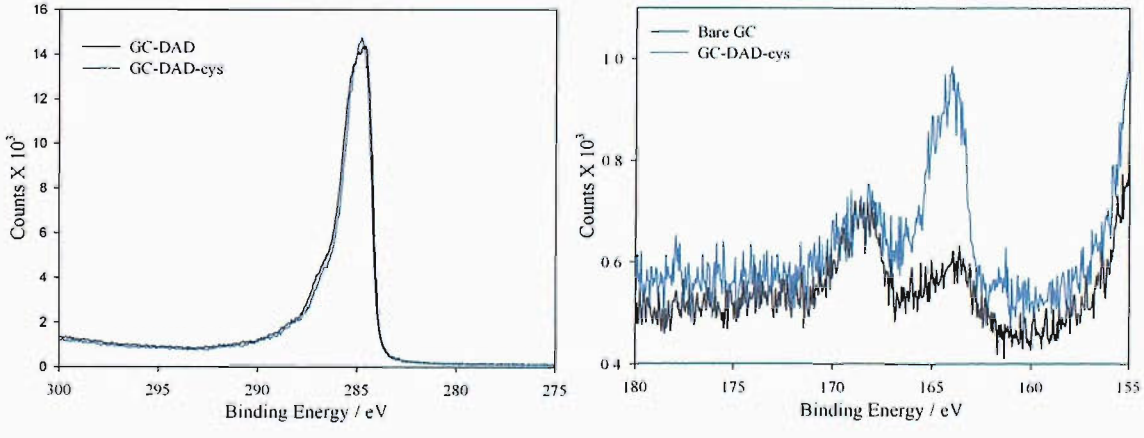


Figure 5.36 Comparison of C1s and S2p regions of the diaminodecane (GC-DAD) and proected diaminodecane-L-cysteine (GC-DAD-cys) modified glassy carbon electrode.

Proof of cysteine attachment at the electrode surface is shown in the S2p spectrum, which shows a significant increase in the non-oxidised sulphur content of the surface when compared with the S2p spectrum of the diaminodecane modified glassy carbon (Fig 5.36 B). Comparison of the C1s spectra (Fig 5.36 A) shows no change in the

spectrum, unlike the ferrocene acetic acid coupled case, where an amide carbon peak was observed. This leads to the possibility that although cysteine is present at the electrode surface, it may have been adsorbed as opposed to covalently attached, there is no evidence of a shoulder at 399.6 eV in the N1s spectrum either, which is also indicative of the absence of amide bonds. The low efficiency of surface attachment may cause the change in the spectrum to be too small to easily observe.

5.3.8 Deprotection of *N*-Boc-*S*-trityl-L-cysteine coupled to a diaminodecane derivatised glassy carbon electrode

The diaminodecane-*N*-Boc-*S*-trityl-L-cysteine modified glassy carbon electrode was deprotected using 95 % trifluoroacetic acid, electrode modifications were analysed by XPS.

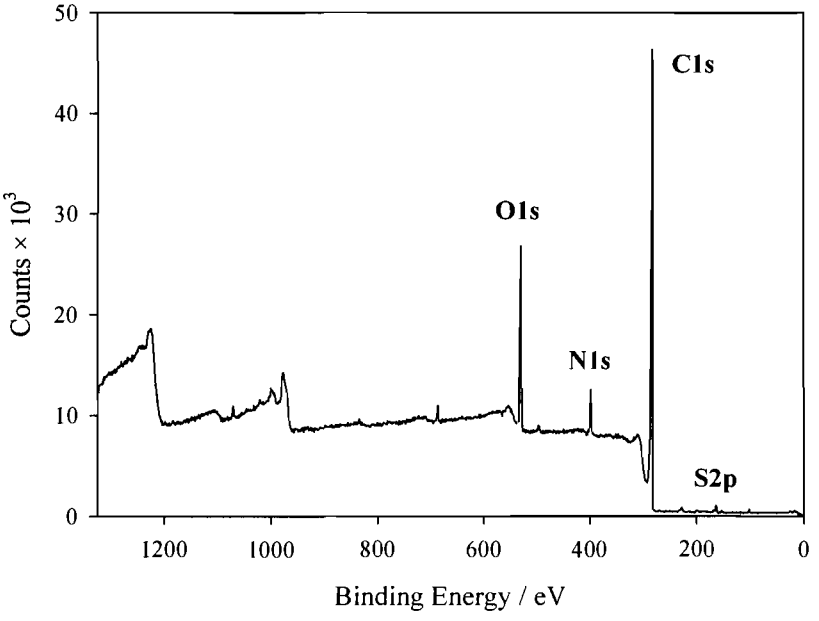


Figure 5.37 Glassy carbon electrode modified with diaminodecane-L-cysteine. XPS Survey; slit = 1.9 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

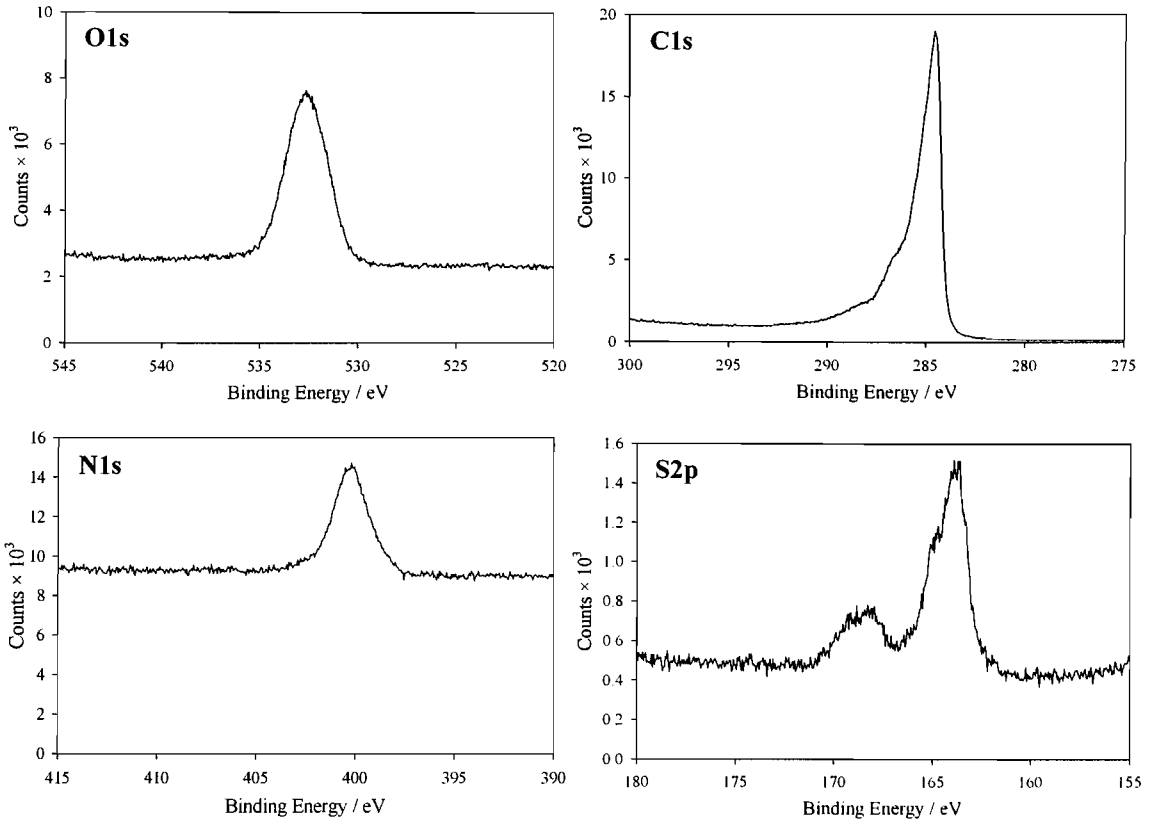


Figure 5.38 Diaminodecane-L-cysteine modified glassy carbon. XPS core level region scans; slit = 1.9 mm, TOA = 90°, step energy = 0.050 eV, time per step = 0.1 s 1. X-ray energy = 1487 eV. Number of scans; O1s and C1s = 1, N1s and S2p = 2.

Atom	Conc / atom %
C	82.6
O	12.1
N	4.4
S	0.9

Table 5.7 Atomic composition of the diaminodecane-L-cysteine modified glassy carbon electrode.

The relatively harsh treatment of the surface with a strong acid seems not to have etched the diaminodecane from the surface, indicating that this method results in a stable electrode modification. The atomic composition of the deprotected electrode again shows very similar proportions of C, N and O atoms to that of the diamonodecane

modified surface. There is however a significantly higher quantity of sulphur on the surface at 0.9 %. This indicates that potentially a coupling efficiency of 40 % has been achieved (separate electrodes were used for each set of XPS data).

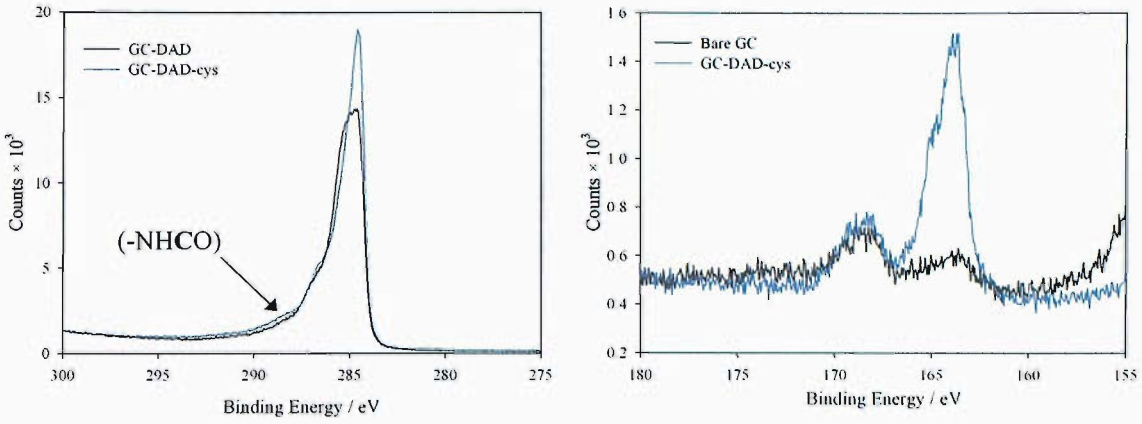


Figure 5.39 Comparison of C1s and S2p regions of the diaminodecane (GC-DAD) and diaminodecane-L-cysteine (GC-DAD-cys) modified glassy carbon electrode.

Comparison of the diaminodecane and diaminodecane-L-cysteine modified electrode C1s and S2p regions show significant differences (Fig. 5.39). In this case, the C1s contains a shoulder at 288.5 eV as observed in the ferroceneacetic acid coupling reaction, indicative of the formation of amide bonds. The S2p spectrum shows a large increase in the non oxidised sulphur, consistent with the presence of cysteine at the electrode surface. These XPS results are more promising, as the large increase in the quantity of sulphur at the electrode surface combined with the observation of amide carbons in the C1s spectrum support the covalent coupling of cysteine to the electrode surface via formation of an amide bond.

5.4 Immobilisation of BioB on a decanediamine-L-cysteine modified glassy carbon electrode

The immobilisation of Biotin synthase (BioB) at the diaminodecane-L-cysteine modified electrode was investigated using XPS. The immobilisation of the protein may be followed by the increase in the amide bond peak in the C1s spectrum and in the surface atomic composition.

The deprotected electrodes were transferred into an anaerobic atmosphere to prevent cysteine thiol disulphide formation and were treated in one of two ways. To immobilise the protein (Fig 5.33), activation via intein mediated ligation is required as discussed in Chapter 4. The protein was activated using thiophenol and the electrode placed directly into a protein labelling reaction where the electrode surface replaces the fluorophore as the nucleophile for covalent C-terminal attachment to the protein. The negative control was exposure to BioB which had been cleaved with cysteine and purified in a mixture containing thiophenol anaerobically. Thiophenol will prevent oxidation of the surface / cysteine thiol at the C-terminus and prevent non-specific attachment via disulphide formation (as observed in Chapter 3).

5.4.1 Negative control BioB immobilisation

Two sets of electrodes (two immobilisation reactions and two negative controls) were examined by XPS.

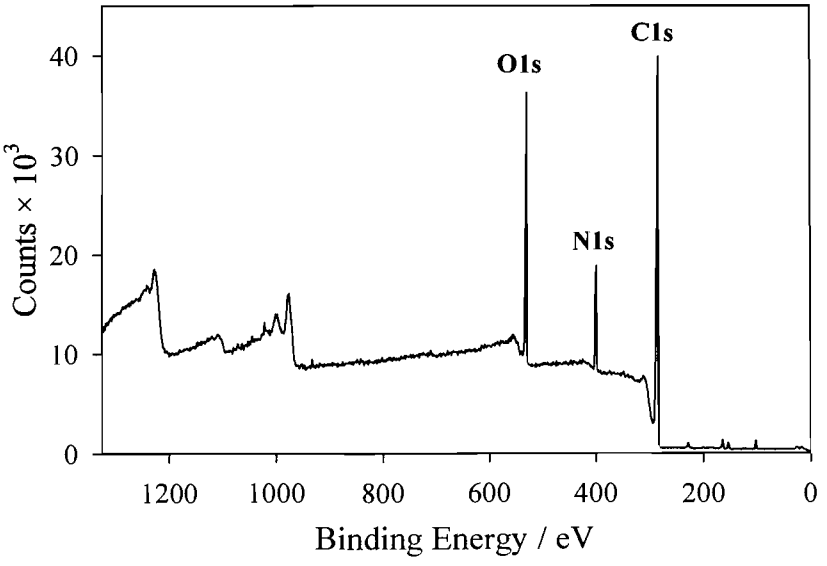


Figure 5.40 Glassy carbon electrode modified with diaminodecane-L-cysteine, negative control BioB immobilisation reaction. XPS Survey; slit = 1.9 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

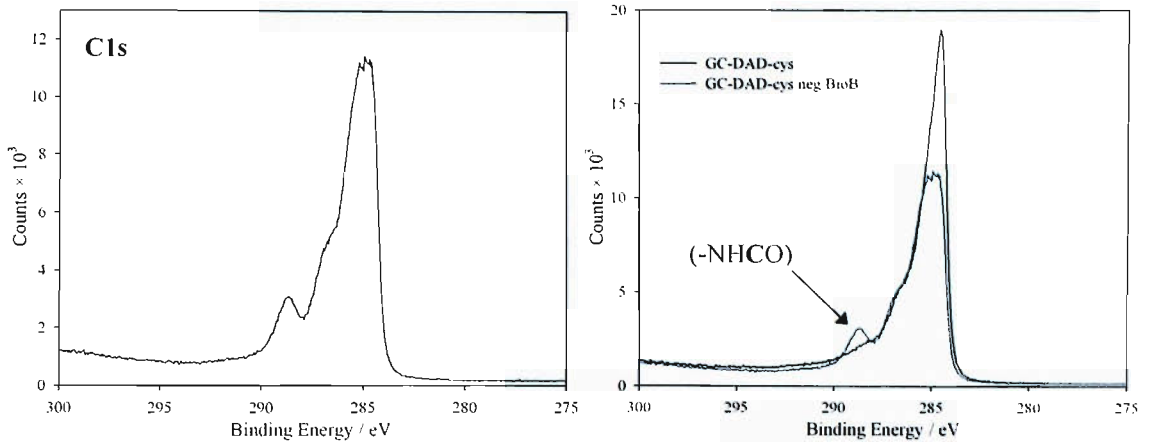


Figure 5.41 BioB immobilisation reaction negative control on diaminodecane-L-cysteine modified glassy carbon. (A) XPS core level region scan; slit = 0.8 mm, TOA = 90°, step energy = 0.050 eV, time per step = 0.1 s 1. X-ray energy = 1487 eV. Number of scans; C1s = 1. (B) Comparison of C1s region of the diaminodecane-L-cysteine (GC-DAD-cys) modified glassy carbon electrode and BioB immobilisation negative control.

Atom	Conc / atom %
C	75.2
O	14.8
N	9
S	1

Table 5.8 Atomic composition of the BioB-intein immobilisation negative control at a diaminodecane-cysteine modified glassy carbon electrode.

It is clear from the survey (Fig. 5.40) that there have been significant changes in the nature of the surface post treatment of the diaminodecane-L-cysteine modified electrode with cysteine cleaved BioB, the negative control. The relationship between the three main elements, carbon, oxygen and nitrogen has significantly altered. Analysis of the atomic composition of the electrode surface (Table 5.8) shows that the proportion of oxygen and nitrogen at the surface has significantly increased, as would be expected upon protein adsorption on an electrode surface.

The C1s region (Fig 5.41 A) shows a strong amide peak at 288.5 eV, indicating the presence of a protein layer. The shape of the peak has changed (Fig. 5.41 B) such that the contribution of the surface aromatic C-C bonds is now barely noticeable; the overriding types of carbon environments are now from the diaminodecane alkane C-C bonds and protein C-O and amide carbons. The XPS results prove that the control electrode has been coated with an adsorbed layer of BioB.

5.4.2 BioB immobilisation reaction

The specific reaction of intein activated BioB with the cysteine modified electrode surface was investigated and analysed by XPS.

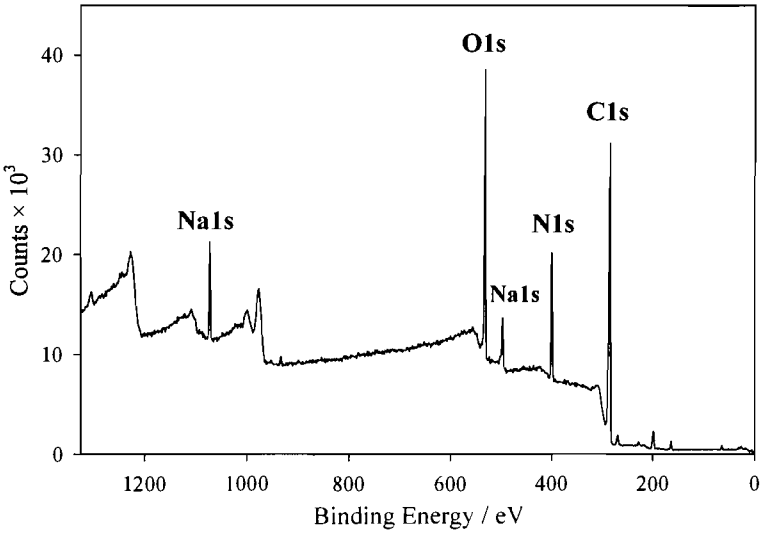


Figure 5.42 Glassy carbon electrode modified with diaminodecane-L-cysteine, BioB immobilisation reaction. XPS Survey; slit = 1.9 mm, TOA = 90°, start energy = 0 eV, end energy = 1325 eV, step energy = 1.0 eV, time per step = 0.5 s, X-ray energy = 1487 eV.

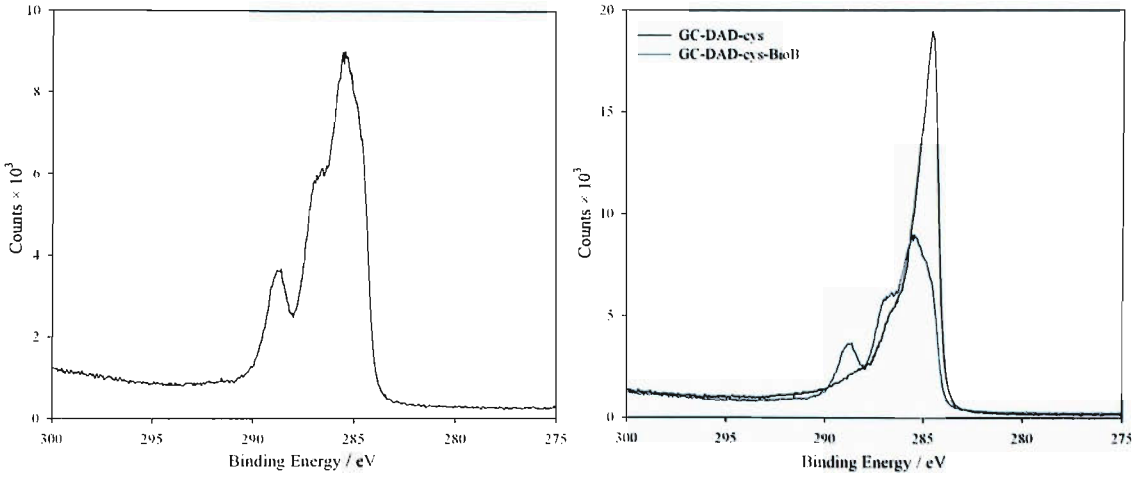


Figure 5.43 BioB immobilisation reaction on diaminodecane-L-cysteine modified glassy carbon. (A) XPS core level region scan; slit = 0.8 mm, TOA = 90°, step energy = 0.050 eV, time per step = 0.1 s 1. X-ray energy = 1487 eV. Number of scans; C1s = 1. (B) Comparison of C1s region of the diaminodecane-L-cysteine (GC-DAD-cys) modified glassy carbon electrode and BioB immobilisation negative control.

Atom	Conc / atom %
C	70.9
O	15.9
N	12.3
S	0.9

Table 5.9 Atomic composition of the BioB immobilisation reaction at a diaminodecane-cysteine modified glassy carbon electrode.

The survey (Fig. 5.42) again clearly shows the increase in intensity of the N1s and O1s peaks in comparison with the C1s peak, indicating the presence of protein at the electrode surface. The survey also shows peaks associated with Na1s electrons, the presence of sodium in the sample was due to its storage in buffer containing sodium chloride prior to analysis. The atomic composition of the electrode surface (Table 5.9) shows the presence of a high proportion of nitrogen in comparison with the diaminodecane-cysteine modified electrode (4.4 vs 12.3 %), the electrode must have a layer of protein present at the surface, but from these results it is not possible to gain information regarding the mode of immobilisation.

The C1s region again shows dramatic change from the cysteine modified electrode (Fig. 5.43) and is further modified in comparison to the negative control (Fig. 5.44). There is now virtually no contribution towards the C1s region from the bare glassy carbon surface, the dominating features are the strong amide peak at 288.5 eV, the protein C-O bonds at 287 eV and the decanediamine alkane C-C bonds at 285.0 eV.

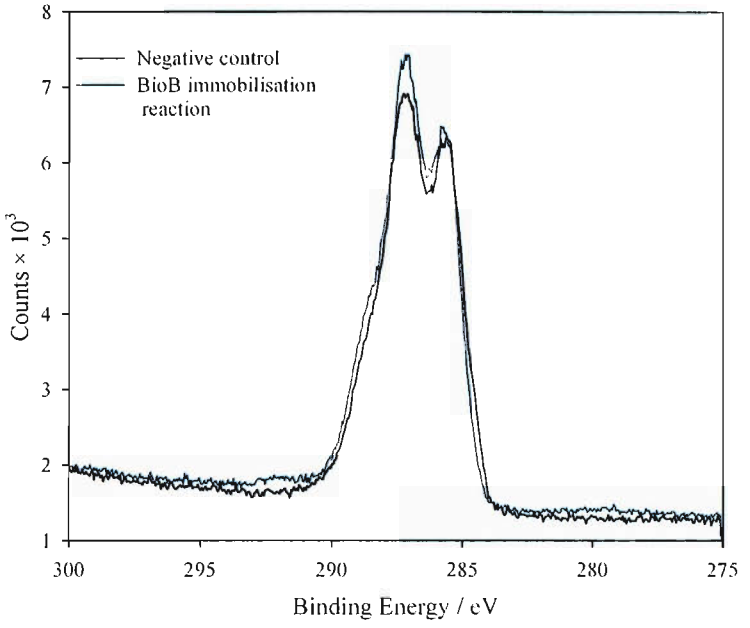


Figure 5.44 Comparison of averaged C1s region of the diaminodecane-L-cysteine modified glassy carbon electrode BioB negative control and immobilisation reaction.

Comparison of the averaged (over two experiments) negative control vs BioB immobilisation reactions (Fig. 5.44) shows the immobilisation reactions electrode surfaces to contain a higher proportion of protein, as highlighted in the previous data. The difference however is small, and there is a very significant amount of adsorbed protein at the electrode in the negative control reactions. These initial attempts at immobilising proteins on electrode surfaces by intein mediated ligation were inconclusive, the major problem being the adsorption of protein onto the electrode surface in negative control reactions.

5.5 Conclusions

The derivatisation of the glassy carbon electrode towards functionalisation with cysteine and then site specific protein immobilisation via intein mediated ligation was attempted.

Derivatisation via reduction of the diazonium salt of nitrobenzene and its reduction led to the modification of the electrode surface with aniline. This was however unreactive towards conventional coupling chemistry.

Glassy carbon was next derivatised via the oxidation of ethylenediamine, which was also found to be unreactive towards amide coupling. Derivatisation with decanediamine resulted in an amine modified surface which was reacted firstly as a trial with ferroceneacetic acid and then with a protected cysteine derivative. The protected cysteine derivative was deprotected to reveal the required reactive cysteine moiety at the electrode surface.

The results of reaction with Biotin synthase via intein mediated ligation were not conclusive, the negative control reactions showed BioB to adsorb onto the electrode surface. A higher concentration of BioB was present at the electrode which had been exposed to an intein mediated ligation reaction, but from the XPS data alone it is not possible to tell the mode of attachment of the protein to the electrode surface.

5.6 References

1. Willner, I. I., and Katz, E. (2000) *Angewandte Chemie International Edition 39*, 1180-1218.

2. Scheller, F. W., Wollenberger, U., Warsinke, A., and Lisdat, F. (2001) *Current Opinion in Biotechnology 12*, 35-40.

3. Predki, P. F. (2004) *Current Opinion in Chemical Biology 8*, 8-13.

4. Mano, N., Mao, F., and Heller, A. (2003) *Journal of the American Chemical Society 125*, 6588-6594.

5. Munge, B., Estavillo, C., Schenkman, J. B., and Rusling, J. F. (2003) *Chembiochem 4*, 82-89.

6. Willner, I. (1997) *Accounts of Chemical Research 30*, 347-356.

7. Mrksich, M., and Whitesides, G. M. (1996) *Annual Review Biophysics and Biomolecular Structure 25*, 55-78.

8. Anicet, N., Anne, A., Moiroux, J., and Saveant, J. M. (1998) *Journal of the American Chemical Society 120*, 7115-7116.

9. Bourdillon, C., Demaille, C., Moiroux, J. and Saveant, J-M. (1996) *Accounts of Chemical Research 29*, 529-535.

10. Charles, P. T., Vora, G. J., Andreadis, J. D., Fortney, A. J., Meador, C. E., Dulcey, C. S., and Stenger, D. A. (2003) *Langmuir 19*, 1586-1591.

11. Anzai, J., Takeshita, H., Kobayashi, Y., Osa, T., and Hoshi, T. (1998) *Analytical Chemistry 70*, 811-817.

12. Sethi, R. S. (1994) *Biosensors and Bioelectronics 9*, 243-264.

13. Ghindilis, A. L., and Kurochkin, I. N. (1994) *Biosensors and Bioelectronics 9*, 353-359.

14. Mirsky, V. M., Riepl, M., and Wolfbeis, O. S. (1997) *Biosensors and Bioelectronics 12*, 977-989.

15. Athey, D., Ball, M., McNeil, C. J. and Armstrong, R. D. (1995) *Electroanalysis 7*, 270-273.

16. Sakslund, H., Wang, J., Lu, F., and Hammerich, O. (1995) *Journal of Electroanalytical Chemistry 397*, 149-155.

17. Katz, E., Riklin, A., Heleg-Shabtai, V., Willner, I., and Buckmann, A. F. (1999) *Analytica Chimica Acta 385*, 45-58.

18. Kacaniklic, V., Johansson, K., Markovarga, G., Gorton, L., Jonssonpettersson, G., and Csoregi, E. (1994) *Electroanalysis 6*, 381-390.

19. Wong, T. S., and Schwaneberg, U. (2003) *Current Opinion in Biotechnology 14*, 590-596.

20. Bartlett, P. N., Simon, E., and Toh, C. S. (2002) *Bioelectrochemistry 56*, 117-122.

21. Zhang, S., Wright, G., and Yang, Y. (2000) *Biosensors & Bioelectronics 15*, 273-282.

22. Bianco, P. (2002) *Journal of Biotechnology 82*, 393-409.

23. Song, S., Clark, R.A., Bowden, E. F. and Tarlov, M.J. (1993) *Journal of Physical Chemistry 97*, 6564-6572.

24. Topoglidis, E., Campbell, C. J., Palomares, E., and Durrant, J. R. (2002) *Chemical Communications (Camb)*, 1518-1519.

25. Patolsky, F., Katz, E., Heleg-Shabtai, V. and Willner, I. (1998) *Chemistry - a European Journal 4*, 1068-1073.

26. Odom, J. M., and Peck, H. D., Jr. (1984) *Annual Review of Microbiology 38*, 551-592.

27. Armstrong, F. A. (1990) *Structure and Bonding 72*, 137-221.

28. Lojou, E., Pieulle, L., Guerlesquin, F. and Bianco, P. (2002) *Journal of Electroanalytical Chemistry 523*, 150-159.

29. Draoui, K., Bianco, P., Haladjian, J., Guerlesquin, F. and Bruschi, M. (1991) *Journal of Electroanalytical Chemistry 313*, 201-214.

30. Razumas, V., Jasaitis, J. and Kulys, J. (1984) *Bioelectrochemistry and Bioenergetics 12*, 297-322.

31. Bourdillon, C., Bourgeois, J.-P., and Thomas, D. (1980) *Journal of the American Chemical Society 102*, 4231-4235.

32. Bianco, P. and Haladjian, J. (1990) *Journal of Electroanalytical Chemistry 293*, 151-163.

33. Yacynych, A. and Kuwana, T. (1978) *Analytical Chemistry 50*, 640-645.

34. Ianniello, R. M. and Yacynych, A. M., (1981) *Analytical Chemistry 52*, 2090-2095.

35. Ianniello, R. M. and Yacynych, A. M. (1981) *Analytica Chimica Acta 131*, 123-132.

36. Razumas, V., Jasaitis, J. and Kulys, J. (1983) *Bioelectrochemistry and Bioenergetics 10*, 427-439.

37. Bianco, P. and Haladjian, J. (1994) *Journal of Electroanalytical Chemistry 367*, 79-84.

38. Nassar, A. E., Willis, W. S., and Rusling, J. F., (1995) *Analytical Chemistry 67*, 2386-92.

39. Katz, E., Heleg-Shabtai, V., Bardea, A., Willner, I., Rau, H. K., and Haehnel, W. (1998) *Biosensors and Bioelectronics 13*, 741-756.

40. Willner, I., Katz, E. and Willner, B. (1997) *Electroanalysis 9*, 965-977.

41. Zimmermann, H., Lindgren, A., Schuhmann, W., and Gorton, L. (2000) *Chemistry-a European Journal 6*, 592-599.

42. Madoz-Gurpide, J., Abad, J. M., Fernandez-Recio, J., Velez, M., Vazquez, L., Gomez-Moreno, C., and Fernandez, V. M. (2000) *Journal of the American Chemical Society 122*, 9808-9817.

43. Jeuken, L. J. C., and Armstrong, F. A. (2001) *Journal of Physical Chemistry B 105*, 5271-5282.

44. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

45. Baker, A. D. and Betteridge, D. (1972) *Photoelectron Spectroscopy*, Pergamon Press.

46. Delamar, M., Hitimi, R., Pinson, J. and Savéant, J.M. (1992) *Journal of the American Chemical Society 114*, 5883.

47. Allongue, P., Delamar, M., Desbat, B., Fagebaume, O., Hitmi, R., Pinson, J., and Saveant, J. M. (1997) *Journal of the American Chemical Society 119*, 201-207.

48. Ortiz, B., Saby, C., Champagne, G. Y., and Belanger, D. (1998) *Journal of Electroanalytical Chemistry 455*, 75-81.

49. Barbier, B., Pinson, J., Desarmot, G., and Sanchez, M. (1990) *Journal of the Electrochemical Society 137*, 1757-1764.

50. Heering, H. A., Wiertz, F. G., Dekker, C., and de Vries, S. (2004) *J American Chemical Society 126*, 11103-11112.

51. Baymann, F., Barlow, N. L., Aubert, C., Schoepp-Cothenet, B., Leroy, G., and Armstrong, F. A. (2003) *FEBS Letters 539*, 91-94.

52. Downard, A. J. (2000) *Electroanalysis 12*, 1085-1096.

53. McCreery, R. L. (1991) *Electroanalytical Chemistry 17*, 221-374.

54. Chen, P. and McCreery, R.L. (1996) *Analytical Chemistry 68*, 3958-3965.

55. Deinhammer, R. S., Ho, M., Anderegg, J. W., and Porter, M. D. (1994) *Langmuir 10*, 1306-1313.

56. McConnell, H. M. (1961) *Journal of Chemical Physics 35*, 508-515.

57. Guillorn, M. A., McKnight, T. E., Melechko, Z., Merkulov, V. I., Britt, P. F., Austin, D. W., Lowndes, D. Z., and Simpson, M. L. (2001) *Journal of Applied Physics 91*, 3824-3828.

58. Finklea, H. O., and Hanshew, D. D. (1992) *Journal of the American Chemical Society 114*, 3173-3181.

59. Tomfohr, J. K., and Sankley, O. F. (2002) *Physical Review B 65*, 245105-245117.

60. Kim, C. H., Pyun, S., and Kim, J. H. (2003) *Electrochimica Acta 48*, 3455-3463.

61. Laviron, E. (1978) *Journal of Electroanalytical Chemistry 101*, 19-28.

62. Armstrong, F. A., Camba, R., Heering, H. A., Hirst, J., Jeuken, L. J., Jones, A. K., Leger, C., and McEvoy, J. P. (2000) *Faraday Discussions*, 191-203; discussion 257-68.

Chapter 6:- Overall conclusions and future work

6.1 Cloning, expression and purification of target-intein fusion proteins

Plasmid pRJW/2960/82 was constructed to allow the simple assembly of plasmids for the expression of target-intein fusion proteins. The target-intein fusion protein expressed from this plasmid contained a C-terminal His_6-tag which allowed simple metal affinity purification, as well as the *Bacillus circulans* chitin binding domain (*1*) as an alternative purification method.

A common limitation of intein mediated ligation is the observation of *in vivo* cleavage of the target-intein thioester, which lowers the yield of intact fusion protein isolated, with typically 10 – 50 % cleavage observed (*2*). This cleavage is likely to be caused by intracellular thiols, such as glutathione. The expression of Fpr-, BioB-, LipA- and SolA-intein was investigated in six different *E. coli* strains, including strain 821 (*3*), a mutant which is unable to biosynthesise glutathione, the principle source of reduced thiol within *E. coli*. It was found that the amount of cleavage of each protein was variable, and the selection of an *E. coli* strain was empirical, but often expression in *E. coli* strain 821 yielded the highest proportion of intact fusion protein. The amount of *in vivo* cleavage of each protein could be optimised, but not avoided.

Future work might include the investigation of the replacement of the *Saccaromyces cerevisae* intein gene with the theromophillic *Pyrococcus* GB-D intein gene (*4*), as the formation of the intermediate thioester has been shown to be temperature dependent. Expression of this modified intein as a C-terminal fusion with the target, and induction at low temperatures (12 – 20 °C) may stop thioester formation until warming of the fusion protein occurs (to 37 °C), therefore avoiding undesired intracellular cleavage. This method may prevent the complete cleavage observed of the PKBα-intein fusion

protein, allowing its isolation. Fpr-, BioB-, LipA- and SolA-intein were expressed and purified by simple Ni affinity chromatography.

6.2 Optimisation of intein mediated fluorescent labelling reactions

The reactivity of the BioB-intein thioester towards various nucleophiles was investigated and the most efficient reaction was observed using cysteine, so a cysteine based fluorescent label was produced. The fluorescent label (FTEC) was synthesised via reaction of a protected cysteine-ethylenediamine conjugate with an activated fluorescein derivative. The synthesis of FTEC was simple and completed in an overall yield of 74 %. The simplicity of the synthesis and potential for reaction of the cysteine-ethylenediamine intermediate with the wide range of available activated fluorophores should allow the future production of a wide range of available fluorophores, for use in FRET based assays (*5*).

Under aerobic conditions, FTEC was found to be very sensitive to oxidation to its symmetric disulphide, so the direct labelling of Fpr-, BioB-, LipA-intein was attempted anaerobically. Only Fpr was labelled directly, in low yield, so the optimisation of the labelling reaction was investigated. It was found that the addition of a low molecular weight thiol, of which thiophenol was found to be most reactive, and anaerobicity were essential for efficient reaction. Under anaerobic conditions, it was possible in most cases to label the target protein using a thiophenol concentration of only 10 mM. In comparison with the usual concentration of low molecular weight thiol used for intein mediated ligation, this is a significant improvement, as thiophenol concentrations as high as 195 mM are often used (*6*). Under rigorously anaerobic conditions it was even found that unusually, direct reaction with the target protein was possible without the need for the addition of thiophenol, but to ensure efficient labelling, 10 mM thiophenol should be added. It should also be noted that for the most part, labelling was complete after only 2-5 hours, this combination of improved conditions should assist in the fluorescent labelling of more delicate proteins which may not be stable to the harsh conditions often used for intein mediated ligation.

The comparison of the rate of labelling of Fpr-, BioB-, LipA-, SolA- and BFP-intein showed that the nature of the target protein may have a significant impact upon the rate of the fluorescent labelling reaction. In most cases, the rate of fluorescent labelling was fast, complete within 2-5 hours, the rate of labelling of BFP was very slow, with only 5 % labelling observed in 24 hours. The reason for this may be due to either its rigid β-barrel structure (*7*), which may have an effect upon the ability of the BFP-intein junction to fold correctly forming the essential thioester, or the lack of flexibility in the junction region may also sterically hinder the attack of FTEC.

Future work may investigate the effect of a series of linker sequences which might be introduced between the target and intein domains, aiming to improve the folding of both by allowing a greater separation and flexibility in the connecting sequence, the addition of a linker may also allow the target-intein thioester to be less sterically hindered, improving the accessibility for reaction. A good starting point for this may be to add the C-terminal region of Fpr for instance, a protein which is readily labelled to the C-terminus of BFP, and to examine the effect this has upon the rate of reaction. It will also be necessary to develop new assays to study the rate of the intein mediated fluorescent labelling reaction, so as the affect of these modifications may be studied more accurately, in real time with a more reliable assay than SDS-PAGE analysis. The fluorescence techniques described in Chapter 1.1 may be useful for this; for instance BFP may be labelled with FTEC and the labelling reaction followed by FRET measurements, or the rate of cleavage of BFP-intein followed by fluorescence polarisation, either of these methods may allow a more detailed study of reaction with the target-intein thioester.

6.3 Electrode modification towards the immobilisation of proteins via intein mediated ligation

The modification of a glassy carbon electrode with a reactive cysteine residue, followed by reaction with a target-intein fusion protein was attempted. The glassy carbon surface could be reproducibly modified by the electrochemical oxidation of diaminodecane, which is attached to the surface via only one of its amine groups. The free primary amine was then coupled with a protected cysteine derivative. Deprotection and reaction

with BioB via intein mediated ligation, under optimal conditions as found for fluorescent labelling were inconclusive, as the negative control reactions showed BioB to adsorb onto the electrode surface. A higher concentration of BioB was present at the electrode which had been exposed to an intein mediated ligation reaction, but from the XPS data alone it was not possible to identify the mode of attachment of the protein to the electrode surface.

Further work on the electrode immobilisation of proteins via intein mediated ligation might include optimisation of the surface immobilisation reaction and the use of BSA to block some of the non-specific surface-protein interactions. Protecting groups that might be electrochemically removed may allow better characterisation of the cysteine derivatised surface. The XPS experiments provide valuable characterisation of the electrode surface and should be widely used in future work, the cysteine coupling reaction may be optimised using this technique. The use of a fluorescent target protein such as GFP may allow more simple analysis of the immobilisation reaction via fluorescence microscopy.

6.4 References

1. Chong, S. R., Montello, G. E., Zhang, A. H., Cantor, E. J., Liao, W., Xu, M. Q., and Benner, J. (1998) *Nucleic Acids Research 26*, 5109-5115.

2. Xu, M. Q., and Evans, T. C. (2001) *Methods 24*, 257-277.

3. Apontoweil, P., and Behrends, W. (1975) *Molecular & General Genetics 141*, 91-95.

4. Xu, M. Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993) *Cell 75*, 1371-1377.

5. Selvin, P. R. (2000) *Nature Structural Biology 7*, 730-734.

6. Zhang, Z. S., Shen, K., Lu, W., and Cole, P. A. (2003) *Journal of Biological Chemistry 278*, 4668-4674.

7. Wachter, R. M., King, B. A., Heim, R., Kallio, K., Tsien, R. Y., Boxer, S. G., and Remington, S. J. (1997) *Biochemistry 36*, 9759-9765.

Chapter 7:- Experimental

7.11 Materials

Restriction enzymes, *Taq* DNA Polymerase, T4 DNA Ligase, Wizard® *Plus* Minipreps DNA Purification System and Wizard® PCR Preps DNA Purification System were purchased from Promega (Southampton). *Pfu Turbo*® DNA polymerase was purchased from Stratagene, dNTP's from MBI Fermentas, primers were synthesised by Oswel. Agarose was purchased from Bioline and low melting point agarose from Gibco BRL.

IPTG and DTT were obtained from Melford Laboratories Ltd. and Lancaster Chemical Company Ltd. respectively. Bacto Tryptone, Yeast Extract and Bacto Agar, for culture media, were purchase from Oxoid and Difco. Plasmids were obtained from Promega or Novagen.

Polyacrylamide-bis polyacrylamide (30% w/v, 37:5:1) was purchased from Amresco, Superdex 75 (S-75), Superdex 200 (S-200) and Chelating Fast Flow resins were purchased from Pharmacia. All other chemicals used were of the highest quality available and purchased from Aldrich, Sigma or Fluka.

PCR

A Progene Thermal Cycler was used for PCR reactions.

Centrifugation

Samples were centrifuged at 4 °C in a Avanti™ J-25 centrifuge (Beckman) centrifuge. For small volumes (sample less than 1.5 mL), a microCentrifugette 4214 (ALC) was used at room temperature (11,000 g).

pH determination

pH measurements were performed using a Mettler Delta 340 pH meter connected to a Mettler Toledo Inlab 413 Combination Electrode. This was calibrated at pH 4.0 to 7.0 or pH 7.0 to 10.0 before use and stored in 3 M potassium chloride.

Fermentation

Fermentation was carried out in an Incubator Shaker innova™4400 (New Brunswick Scientific).

Incubations

Bacterial plate cultures were grown at 37 °C overnight in a Economy Incubator Size 2 (Gallenkamp). Liquid cultures were incubated in an Incubator Shaker innova™4400 (New Brunswick Scientific) at 37 °C, 180 rpm unless otherwise stated.

Protein concentration

Protein samples were concentrated by ultrafiltration in an Amicon pressure cell.

7.12 General experimental methods

Standard sterile techniques were applied during microbiological experiments. Growth media / heat stable solutions were autoclaved, whilst heat labile solutions (antibiotics, IPTG and arabinose) were filter sterilised through 0.22 μm filters.

Growth media were supplemented with the appropriate antibiotic at the following concentrations; ampicillin, 100 μg / mL, kanamycin, 30 μg / mL, streptomycin, 10 μg / mL. Protein expression was induced by the addition of arabinose (0.2 % w/v) for pBAD derived plasmids and IPTG (1 mM) for pET derived plasmids.

Plasmids and their resistance markers are listed in table 2.1.

Method 1. PCR

PCR reactions (50 μL) were processed in a thermocycler using the following method, unless otherwise reported.

	Volume	Concentration	Total Quantity
Sterile Water	31.5 - 25.5 μL		
DMSO	5 μL		
dNTPs	1 μL	10 mM	10 nmoles
DNA template	1.2 μL	100 ng / μL , 10 ng / μL	120 ng
Forward / Reverse primers	2.5 μL Each	20 μM	0.05 nmoles
***PfU* Turbo DNA polymerase**	1 μL	2.5 U / mL	2.5 U

Table 7.1 PCR general reaction mixture

The reaction program was:-

1 cycle
95 °C 10 min
80 °C 3 min

30 cycles
95 °C 1 min
55 °C 1 min
72 °C 2 min

1 cycle
72 °C 10 min
4 °C storage

PCR products were analysed by agarose gel electrophoresis and purified using QIAquick® PCR purification kit, following the manufacturers guidelines.

Method 2. A-tailing

Purified PCR products were A-tailed using the following conditions.

	Volume	Concentration	Total Quantity
PCR Product	7 μL	20-100 ng / μL	140-700 ng
Taq polymerase buffer	1 μL	10X	0.9X
$MgCl_2$	0.8 μL	25 mM	20 nmoles
ATP	1 μL	2 mM	2 nmoles
Taq DNA polymerase	1 μL	5 U / μL	5 U

Total volume = 10.8 μL. Reactions were maintained at 72 °C for 30 minutes.

Table 7.2 A-tailing reaction mixture

Method 3. Ligation into pBAD-TOPO® vector

A-tailed PCR product (method 2) was ligated into pBAD-TOPO® following manufacturers instructions. The resulting ligation reaction mixture was transformed into TOP-10 competent cells (method 4).

Method 4. Transformation

Competent cells were purchased (TOP-10) or prepared by either the rubidium chloride (*1*) or One Step method (*2*).

Competent cell aliquots (50-100 μL) were thawed on ice for 10 minutes before addition of a ligation reaction (10 μL, method 3) or purified plasmid DNA (1 μL) with gentle mixing. Reactions were maintained on ice for 30 minutes before being heat shocked in a temperature controlled waterbath (42 °C, 45 seconds). The cultures were then placed on ice and SOC medium added (250 μL). Cultures were then incubated in a shaker (1 hour, 37 °C, 180 rpm), and spread onto agar plates containing the appropriate antibiotic and incubated overnight (37 °C). Colonies were characterised by PCR screening (method 5), minipreps (method 6) and analytical digestion (method 7).

Method 5. PCR screening of colonies

Colonies were screened for inserts by direct colony PCR, using vector-specific primers (pBAD-TOPO® reverse and insert forward, as specified). Colonies were picked from plates and used as a template in PCR reactions.

	Volume	Concentration	Total Quantity
Sterile Water	7 µl	20-100 ng/µl	140-700 ng
Taq Polymerase buffer	1 µl	10 X	0.9 X
$MgCl_2$	0.8 µl	25 mM	20 nmoles
ATP	1 µl	2 mM	2 nmoles
Forward and reverse primers	2.5 µl	20 µM	0.05 nmoles
Taq DNA polymerase	1µl	5U / µl	5U

Table 7.3 PCR screening reaction mixture

The residual cells from a colony were used to inoculate 10 mL 2-YT medium, cultured overnight (37 °C, 180 rpm) for storage. PCR reaction conditions:-

1 cycle
95 °C 10 min
80 °C 3 min (Taq polymerase added as a 1:10 diluted solution in water)

30 cycles
95 °C 1 min
55 °C 1 min
72 °C 1min / kb

1 cycle
72 °C 10 min
4 °C storage

Method 6. Minipreps

Plasmid DNA was isolated using Wizard® *Plus* Minipreps DNA Purification System, used as stated in manufacturers instructions. Sterile water was used to elute the isolated plasmid DNA.

Method 7. Restriction enzyme digestion (analytical and preparative)

Analytical digestion of plasmid DNA (45 μL, 50-75 ng / μL) isolated from bacterial culture (5 mL, method 6) was carried out using the following conditions.

	Volume	Concentration	Total Quantity
Plasmid	5 μL	50 - 75 ng / μL	250 - 375 ng
Buffer	1 μL	10X	1X
BSA	1 μL	1 mg / mL	1 μg
Restriction Enzyme (each)	0.5 μL	10 U / μL	5 U
Sterile Water	2 μL		

Table 7.4 Analytical digestion reaction mixture

Preparative digestion of plasmid DNA (25 μL, 100 - 300 ng / μL):

	Volume	Concentration	Total Quantity
Plasmid	20 μL	100 - 300 ng / μL	2 - 6 μg
Buffer	3 μL	10X	0.9X
BSA	3 μL	1 mg / mL	3 μg
Restriction Enzyme (each)	3 μL	10 U / μL	30 U

Total volume = 32 μL.

Table 7.5 Preparative digestion reaction mixture

Reactions were incubated at 37 °C for 1.5 hours.

Method 8. Purification of digested fragments

Completed preparative digestion reactions (30 μL) were loaded onto a 1 % low melting point agarose gel. The desired product was excised and purified using the QIAquick® Gel Extraction Kit following the manufacturers instructions.

Method 9. Ligation into an expression vector

Purified fragments from preparative digestions (vector and fragment) were ligated in a 10 μL reaction mixture.

	Volume	Concentration	Total Quantity
T4 DNA Ligase Buffer	1 μL	10X	1X
Backbone	2 μL	50 ng / μL	100 ng
Insert	*x* μL*	25 - 100 ng / μL	
T4 DNA ligase	1 μL	3 U / μL	3 U
Sterile Water	To a final volume of 10 μL		

* *x* = varied according to the concentration of vector and insert, to give a 3:1 insert:vector molar ratio.

Table 7.6 Ligation reaction mixture

Ligation reactions were incubated overnight at 4 °C and transformed into TOP-10 competent cells (method 4). Colonies were characterised by PCR screen (method 5) and analytical digestions of minipreps (methods 6,7).

Method 10 Glycerol freeze preperation

Permanent stocks of plasmid bearing strains were prepared by adding sterilised glycerol (75 % v/v, 125 μL) to bacterial culture (500 μL). These were stored at -80 °C.

Method 11 Protein concentration determination

Protein concentration was assayed using the method of Bradford (*3*).

Method 12. 10 % SDS-PAGE denaturing gel

For 10 mL of resolving gel solution (5 mL per plate) the following components were mixed in the order as shown in Table 7.7:-

Reagent	Quantity
H_2O	4.0 mL
30 % Acrylamide / bis acrylamide	3.3 mL
1.5 M Tris/HCl (pH 8.8)	2.5 mL
10 % SDS	0.1 mL
10 % Ammonium Persulphate	0.1 mL
TEMED	4 mL

Table 7.7 Resolving gel mixture

4 mL of this solution was then poured into each plate and allowed to set for 45 minutes. The surface of the gel was covered with a thin layer of water, which was then removed carefully and the following stacking gel was then prepared as in Table 7.8.

Reagent	Quantity
H_2O	3.4 mL
30 % Acrylamide / bis acrylamide	0.83 mL
1.5 M Tris/HCl (pH 8.8)	0.63 mL
10 % SDS	0.05 mL
10 % Ammonium Persulphate	0.05 mL
TEMED	5 mL

Table 7.8 Stacking gel mixture

This mixture was then directly poured onto the top of the resolving gel and a Teflon comb inserted into the gel solution. Teflon combs were removed after one hour and gels immediately used or stored at 4 °C.

Samples were prepared by mixing 20 μL protein sample with 20 μL sample loading buffer (Table 7.9), denaturing at 95 °C for 5 minutes. Samples were then applied to the gel.

Reagent	Quantity
0.2 M Tris/HCl (pH 6.8)	2.5 mL
DTT	154 mg
SDS	200 mg
Bromophenol Blue	10 mg
Glycerol	1 mL
Deionised water	Adjust volume to 10 mL

Table 7.9 Sample loading buffer stock solution

Reagent	Quantity
Tris Base	15.1 g
Glycine	94 g
10% SDS solution	50 mL
Deionised water	Adjust volume to 1000 mL

Table 7.10 SDS-PAGE running buffer (×5 stock solution)

Electrophoretic separation was at 180 V (~15 V / cm) (in SDS-PAGE running buffer, 1 ×) and analysed. Gels were stained using Coomassie brilliant blue stain (Table 7.11) and destained (Table 7.12).

Reagent	Quantity
Coomassie brilliant blue	2.5 g
Methanol : water (1 : 1)	90 mL
Glacial acetic acid	10 mL

Table 7.11 Coomassie brilliant blue protein stain

Reagent	Quantity
Deionised water	4375 mL
Methanol : water (1 : 1)	375 mL
Glacial acetic acid	250 mL

Table 7.12 Destain solution

Method 13 Small scale expression experiments

An overnight starter culture (10 mL) (inoculated from glycerol freeze stock) was used to inoculate 2-YT medium (100 mL) containing appropriate antibiotics (Table 7.13). Culture was incubated in a shaker (37 °C, 180 rpm), and growth monitored by OD_{600}. Cells were induced at OD_{600} = 0.6 by the addition of arabinose for pBAD expression vectors (final concentration 0.2 %) or IPTG (final concentration 1 mM) for the pET expression vector. The incubation temperature was reduced to 27 °C and cells grown for a further 4 h before harvesting by centrifugation (8 min, 10,000 rpm). Cell pellets were resuspended in lysis buffer (750 μL) and lysed by sonication (6 × 3 s on / off). Cell debris was separated by centrifugation (13,000 rpm, 10 min) and the supernatant decanted. Cell debris was resuspended in lysis buffer (100 × pellet volume). Protein concentration was then estimated by Bradford assay (method 11). The protein content of both supernatant and pellet were analysed by SDS-PAGE (method 12).

Antibiotic	Final Concentration	Stock Solution
Kanamycin	30 μg / mL	30 mg / mL in H_2O
Ampicillin	100 μg / mL	100 mg / mL in H_2O
Chloramphenicol	34 μg / mL	34 mg / mL in EtOH
Streptomycin	10 μg / mL	10 mg / mL in H_2O

Table 7.13 Antibiotic solutions

Method 14 Large Scale Expression Experiments

An overnight starter culture was used as a 1 % innoculum for 4 × 1.25 L 2-YT medium (containing appropriate antibiotic). Cultures were incubated in an orbital shaker (37 °C, 180 rpm) and cell growth monitored by OD_{600}. At OD_{600} 0.6 cells were induced (method 13) and the incubation temperature lowered to 27 °C. Cells were harvested five hours post induction by centrifugation (JA14 rotor, 10,000 rpm, 8 min, 4 °C) and stored at -80 °C.

Method 15. Cell Lysis

The frozen cell pellet was resuspended in 5 × w / v lysis buffer (50 mM Tris/HCl pH 8.1, 0.5 M NaCl, 10 % glycerol). Lysosyme (final concentration of 0.1 mg/mL) and benzonaze (final concentration 10 U / mL) were added, and the mixture stirred for 30 minutes. The mixture was then sonicated (10 × 30 s bursts) whilst cooled on ice. Cell debris was separated by centrifugation (JA14 rotor, 10,000 rpm, 30 min, 4 °C). The supernatant was then carefully decanted.

Method 16. Protein purification: Fast Protein Liquid Chromatography (FPLC)

All enzyme purifications were performed using a Pharmacia FPLC System. Buffers were prepared in Milli-Q water. Resin specifications are given in Table 7.14.

Resin	Particle Size / μm	Matrix	Use
Superdex 75	13	Spherical composite of cross-linked agarose and dextran.	Size exclusion seperation of globular proteins mw 3 - 70 kDa
Superdex 200	13	Spherical composite of cross-linked agarose and dextran.	Size exclusion seperation of globular proteins mw 10 - 600 kDa
Chelating Sepharose Fast Flow	45 - 165	Highly cross linked 6 % agarose, derivatised with chelating iminodiacetic acid groups.	Medium for immobilised metal ion affinity chromatography

Table 7.14 Resin types used for protein purification

Method 17 Small scale expression study / nickel spin column purification

Cells from a cell culture (25 mL of a 100 mL growth) were harvested by centrifugation, resuspended in 1 mL lysis buffer and lysed by sonication (5 × 5s bursts). Cell debris was separated by centrifugation (13,000 rpm, 15 mins).

Nickel spin columns were prepared by adding chelating sepherose (150 μL) (Pharmacia chelating sepherose^{TM} fast flow) to a Quiagen QIAquick spin column, excess buffer was removed by centrifugation (3,000 rpm, 1 min). N.B all subsequent elutions were by centrifugation (3,000 rpm, 1 min, 4 °C). The following protocol was then followed:-

- Resin was charged with 0.2 M nickel sulphate (300 μL) and eluted by centrifugation.

- The resin was washed with 50 mM Tris/HCl pH 8.1 (700 μL).

- To bind the His tagged protein, cleared lysate (700 μL) was applied to the column, mixed, left to stand for 1 minute followed by centrifugation. The flow through was then reapplied, left to stand for one minute and eluted (repeated once).

- The column was washed twice with 700 μL of 50 mM Tris/HCl buffer pH8.1 containing 50 mM imidazole and 0.5 M NaCl. After the final wash, the column was centrifuged once to remove excess buffer.

- His tagged protein was eluted with Tris/HCl buffer pH 8.1 (50 μL, 50 mM) containing imidazole (0.5 M) and NaCl (0.5 M). This buffer was well mixed with the resin and left to stand for five minutes before elution.

The eluted proteins were analysed by 10 % SDS-PAGE (method 12).

Method 18 Anaerobic experiments

Where stated, reactions were maintained under anaerobic conditions using a glovebox, maintained at 20 °C under an atmosphere of nitrogen at <0.2 ppm O_2. Buffers for these experiments were introduced into the glovebox and were allowed to deoxygenate overnight before use in preparation of reagent solutions. Aliquots of protein (100 μL) were allowed to deoxygenate for 1 h before use.

Method 19 Conditions for SDS-PAGE analysis of intein mediated fluorescent labelling reactions

Protein samples (10 μL) were denatured in Gel Loading Buffer (5 μL) for five minutes at 80 °C (*1*) and analysed by SDS-PAGE on 10 % gels. Gels were visualised under UV illumination and after Coomassie staining using a Syngene GeneGenius imager. The images were analysed using GeneTools software (Syngene, Cambridge, UK).

7.2 Experimental for Chapter 2

Assembly of target-intein expression vectors

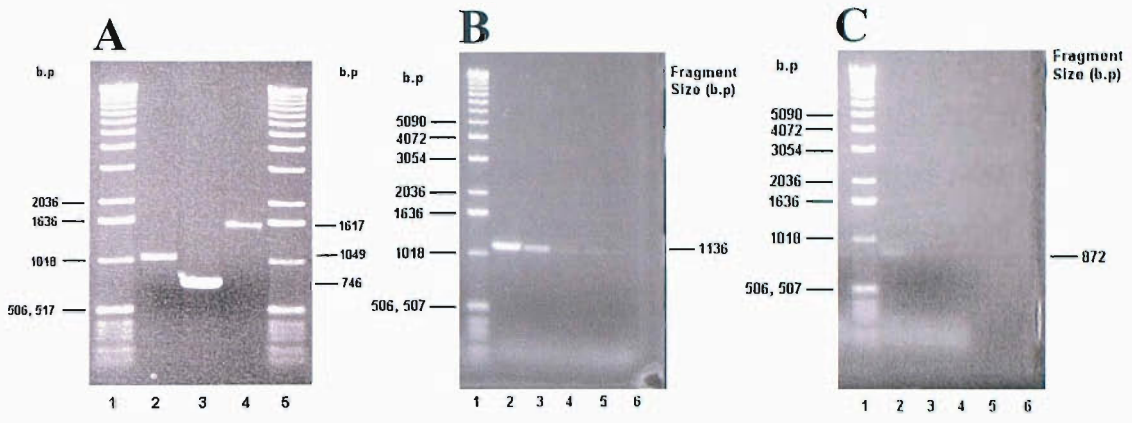


Figure 7.1 PCR products; (A) Lanes 1, 5 = 1 kb DNA ladder, 2 = Amplified *bioB* (1049 bp), 3 = *fpr* (746 bp), 4 = *intein* (1617 bp) gene. (B) *solA* (1136 bp) PCR products; Lane 1 = 1 kb DNA ladder, 2 – 6 = 1:1, 1:10, 1:50, 1:100, 1:1000 genomic DNA dilution. (C) *lipA* (872 bp) PCR products; reactions as per 7.1 (B).

PCR primers

pfbioB 5'ccatggctcaccgcccacgc 3'

prbioB 5'-ggatccttactcgagtaatgctgccgcgttgtaatattcg 3'

pffpr 5'ccatggctgattgggtaacaggc 3'

prfpr 5'ttagccctcgagccagtaatgctccgctgt 3'

pflipA 5'ccatggagaaaaaagttcaaattgc 3'

prlipA 5'ggatcctcactcgagactatgattcttagcattttcaatcgcc 3'

pfint 5'ctcgagggctgctttgccaagggt 3'

print 5'gaattctcagtggtggtggtggtggtgttgaagctgccacaaggcaggaacgt 3'.

pfsolA 5' ctgcagttactcgagttggaagcgggaaagcctgaatgg 3'

prsolA 5' ccatgggaaaatacgatctcatcattatt 3'

Primers for *bioB* (pfbioB and prbioB), *fpr* (pffpr and prfpr), *lipA* (pflipA and prlipA) and *solA* (pfsolA and prsolA) were designed such that the PCR products contained 5' *Nco*I and 3' *Xho*I restriction sites. Primers for *SCE VMAI intein* were designed so that

the PCR product contained 3' *Xho*I - gly - cys extension (*4*) and 5'*Eco*RI site, with a C-terminal His_6-tag in the expressed protein sequence.

7.2.1 Assembly of pRJW/2960/82: intein in pBAD-HisA

The Sce *VMAI* intein gene was amplified by PCR (method 1) from pTYB1 from the NEB IMPACT™ kit as a template, *Pfu* DNA polymerase and primers pfint and print. The annealing temperature was altered from standard to 50 °C, and elongation time was extended to 5 min. The product (Fig. 7.1A, lane 4) was ligated into pBAD-TOPO (method 2, 3) and (pRJW/2960/62) transformed into competent TOP-10 cells (method 4). Plasmid DNA was isolated (method 6) and positive colonies identified by analytical digestion (method 7).

The *XhoI-intein-Eco*RI fragment was subcloned into pBAD-HisA (method 7, 8, 9) to form pRJW/2960/82. Colonies containing the correct plasmid were identified by analytical digestion (method 7). All pDNA isolated was found to be positive for the *intein* insert (Fig. 7.2).

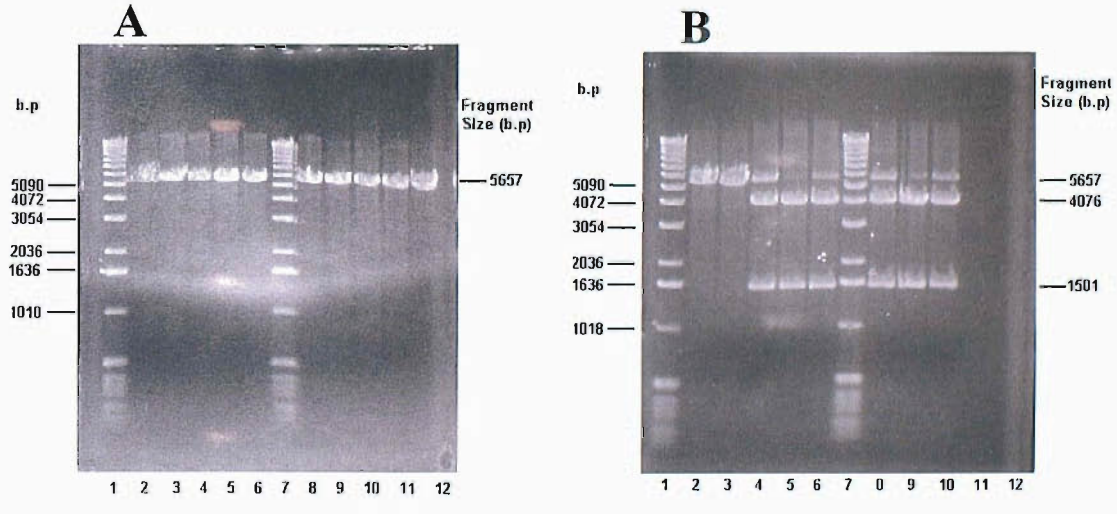


Figure 7.2 Analytical digestion of *Xho*I-*intein-Eco*RI ligation into pBAD-HisA. (A) Lanes 1, 7, 1 Kb ladder, lanes 2 – 6, 7 restriction with *Nco*I (expected fragment 5657 bp), lanes 9 – 12 restriction with *Nde*I (expected fragment 5657 bp). (B) Lanes 1, 7, 1 Kb ladder, lanes 2,3 restriction with *Nde*I (expected fragment 5657 bp), lanes 4 – 6 and 8 – 10, restriction with *Xho*I / *Eco*RI (expected fragments 4076, 1581 bp).

7.2.2 Amplification of *fpr* and subcloning into pRJW/2960/82

The *fpr* gene was amplified from *E. coli* genomic DNA by PCR (Method 1) using primers pffpr and prfpr (Annealing temperature adjusted to 60 °C). The PCR product (Fig. 7.3, lane 3) was ligated into pBAD-TOPO as described in 7.2.1. The *Nco*I-*fpr*-*Xho*I fragment was subcloned (method 7, 8, 9) into pRJW/2960/82 to assemble the *fpr-intein* expression vector pRJW/2960/88.

Correct plasmids were identified by analytical digestion (method 7). All pDNA isolated was positive for the *fpr-intein* insert (Fig. 7.4).

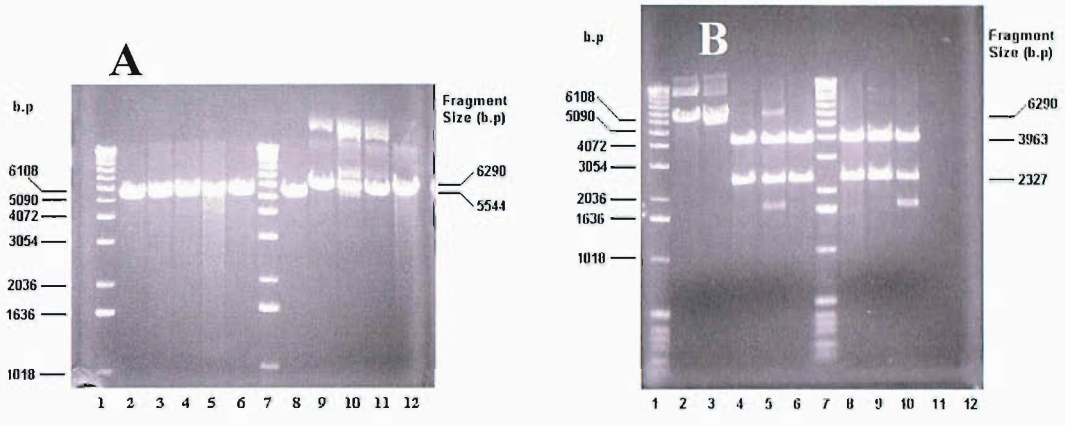


Figure 7.4 Analytical digestion of *Nco*I-*fpr*-*Xho*I ligation into pRJW/2960/82. (A) Lanes 1, 7, 1 Kb ladder, lanes 2 – 6, 8 restriction with *Nco*I / *Xho*I (expected fragments 746, 5544 bp), lanes 9 – 12 restriction with *Xho*I (expected fragment 6290 bp). (B) Lanes 1, 7, 1 Kb ladder, lanes 2 – 3, restriction with *Xho*I (expected fragment 6290 bp), lanes 4 – 6, 8 - 10 restriction with *Eco*RI / *Xho*I (expected fragments 2327, 3963 bp).

7.2.3 Amplification of *bioB*, *solA* and *lipA* genes, subcloning into pRJW/2960/82

The *E. coli bioB* (primers pfbioB and prbioB), *solA* (primers pfsolA and prsolA) and *Sulfolobus solfataricus lipA* (primers pflipA and prlipA) genes were amplified from genomic DNA using the standard protocol (Method 1). Subsequent cloning was as described in section 7.2.2, assembling pRJW/2960/89 (*bioB-intein*), pRJW/3104/100 (*solA-intein*) and pRJW/3104/40 (*lipA-intein*).

Positive inserts were identified by analytical digestion (method 7). In the case of the assembly of pRJW/2960/89, three out of the six colonies were found to be positive for the *bioB-intein* insert (Fig. 7.5).

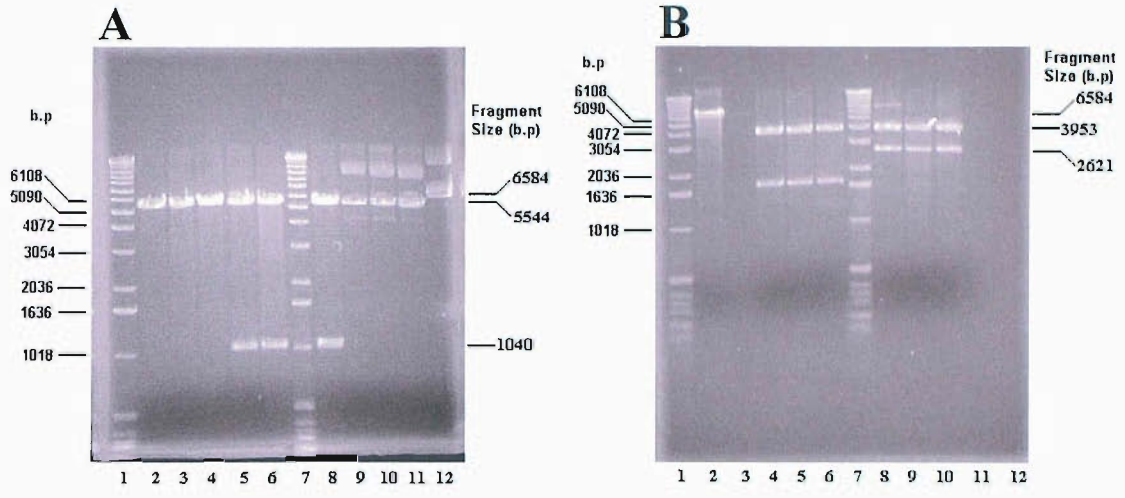


Figure 7.5 Analytical digestion of *Nco*I-*bioB*-*Xho*I ligation into pRJW/2960/82. (A) Lanes 1, 7, 1 Kb ladder, lanes 2 – 6, 8 restriction with *Nco*I / *Xho*I (expected fragments 1040, 5544 bp), lanes 9 – 12 restriction with *Xho*I (expected fragment 6584 bp). (B) Lanes 1, 7, 1 Kb ladder, lanes 2 – 3, restriction with *XhoI*(expected fragment 6584 bp), lanes 4 – 6, 8 - 10 restriction with Eco*RI* / *XhoI* (expected fragments 2621, 3963 bp).

One colony was found to be positive in the assembly of the *solA-intein* construct (Fig. 7.6)

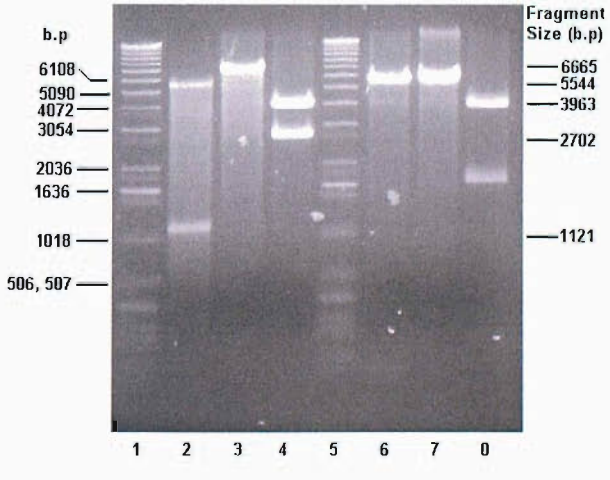


Figure 7.6 Analytical digestion of *Nco*I-*solA*-*Xho*I ligation into pRJW/2960/82. Lanes 1, 5, 1 Kb ladder, lanes 2, 6 restriction with *Nco*I / *Xho*I (expected fragments 1121, 5544

bp), lanes 3, 7 restriction with *Xho*I (expected fragment 6665 bp), lanes 4, 8 restriction with *Nco*I / *Eco*RI (Expected fragments 2702, 3963 bp).

In the case of the assembly of pRJW/3104/40, both colonies were found to be positive for the *lipA-intein* insert (Fig. 7.7).

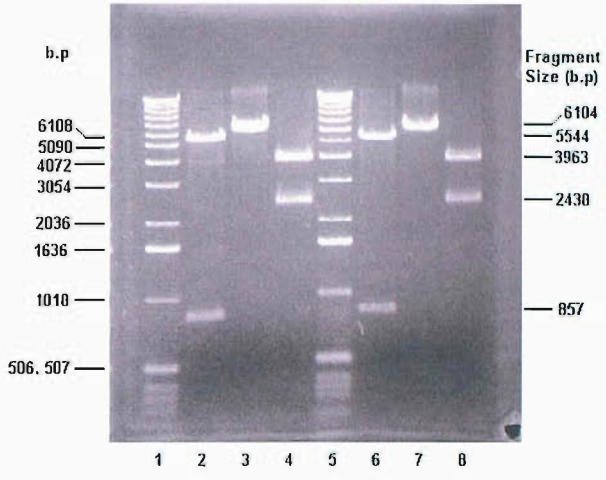


Figure 7.7 Analytical digestion of *Nco*I-*lipA-Xho*I ligation into pRJW/2960/82. Lanes 1, 5, 1 Kb ladder, lanes 2, 6 restriction with *Nco*I / *Xho*I (expected fragments 857, 5544 bp), lanes 3, 7 restriction with *Xho*I (expected fragment 6401 bp), lanes 4, 8 restriction with *Nco*I / *Eco*RI (Expected fragments 2438, 3963 bp).

7.2.3 Amplification of *PKBα* and subcloning into pRJW/2960/82

The *Homo sapiens PKBα* gene was amplified and ligated into pBAD-TOPO by Dr. Veronique Calleja forming pVC100 (Appendix A). The *Nco*I-*PKBa-Xho*I fragment was restricted in two steps. Firstly pVC100 was restricted on a large scale with *Xho*I only (method 7), the fragment gel purified (method 8) and the resulting product restricted preparitively with *Nco*I. The product was gel purified and ligated into pRJW/2960/82 assembling pRJW/3953/77.

Positive inserts were identified by analytical digestion (method 7). In the case of the assembly of pRJW/3953/77, one colony was found to be positive for the *PKBα-intein* insert (Fig. 7.8).

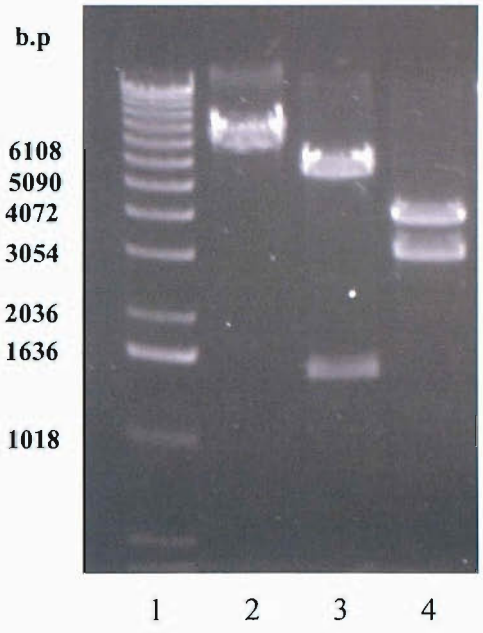


Fig. 7.8 Analytical digestion of pRJW/3953/77. Lane 1 = 1 kb DNA ladder, 2 = restriction with *Xho*I (expected fragment = 7025 bp), 3 = restriction with *Nco*I / *Xho*I (expected fragments = 1482, 5543 bp), 4 = restriction with *Nco*I / *Eco*RI (expected fragments = 3063, 3962bp).

7.2.5 Expression of target-intein fusion proteins

pRJW/2960/88, pRJW/2960/88, pRJW/3104/100, pRJW/3104/40 and pRJW/3953/77 were transformed (method 4) into competent BL21(DE3), C43(DE3), LMG194, AN1459, NM554 and strain 821 cells. Small scale expression experiments (method 14) / Ni affinity column (method 17) purification of target-intein fusion proteins from these strains were analysed by 10 % SDS-PAGE (method 12). Where identification of the expressed fusion protein was unclear (PKBα-intein), gels were first stained using His-Stain-532 (Molecular probes) and visualised under UV light before conventional staining.

Large scale expression and purification of target-intein fusion proteins

7.2.6 Purification of Fpr-intein

Buffer A = 50 mM Tris/HCl pH 8.1 + 500 mM NaCl + 50 mM imidazole
Buffer B = 50 mM Tris/HCl pH 8.1 + 500 mM NaCl + 500 mM imidazole
Buffer C = 50 mM Tris/HCl pH 8.1 + 500 mM NaCl

Ni affinity chromatography

Fpr-intein was overexpressed on a large scale (method 14, 15 L culture) from pRJW/2960/88 in strain 821. A 95 g cell pellet was lysed (method 15) and the cleared lysate applied to a pre-equilibrated Ni affinity column (50 mL column volume, flow rate = 5 mL / min, washed with 150 mL buffer A, 150 mL buffer B and 200 mL buffer A). Once loading was complete, the column was washed with 200 mL buffer A. Protein was eluted at 5 mL / min, over a gradient of buffer B from 0 – 100 % over 600 mL. 12 mL fractions were collected, light yellow colouration was observed in fractions 9 – 18 (protein concentration fractions 9 - 18 = 0.3 – 3 mg / mL). Protein purity of fractions 8 – 19 was assessed by SDS-PAGE (method 12, 10 % gel). Fractions 8 - 19 were combined and concentrated by ultrafiltration (50 kDa MW cut off filter) to a volume of 3.5 mL.

Gel filtration chromatography

To a pre-equilibrated S-75 gel filtration column (column volume = 600 mL, washed with 2 L buffer C) was applied concentrated supernatant (3.5 mL, 68 mg / mL) from the Ni affinity purification. Protein was eluted at 3 mL / min, collecting 9 mL fractions. Deep yellow colouration was observed in fractions 30 – 33 (retention time of Fpr-intein = 360 mL). Protein concentration was assessed by Bradford Assay (method 11), fractions 29 – 34 were found to contain 1.5 – 10 mg / mL protein. These fractions were immediately stored at -80 °C. Content of fractions 28 – 37 was assessed by SDS-PAGE (method 12, 10 % gel).

7.2.7 Purification of BioB-intein

Ni affinity chromatography

BioB-intein was overexpressed from pRJW/2960/89 in BL21(DE3) (method 14), yielding 25 g cell paste from 5 L culture. These cells were lysed and protein purified by Ni affinity as described in 7.2.6, except that pH buffers A and B was 7.0. Light brown colouration was observed in fractions 4 – 20 (protein concentration fractions 7-18 = 2.3 – 8.2 mg / mL). Fractions 10 – 20 were combined and concentrated to a volume of 7 mL.

Gel filtration chromatography

Method as described in 7.2.6, except that pH buffer C = 7.0. Protein was eluted from an S-200 gel filtration column (800 mL volume) at a retention time of 300 mL. Most pure fractions were directly frozen at -80 °C.

7.2.8 Purification of SolA-intein

Ni affinity chromatography

SolA-intein was overexpressed from pRJW/3104/100 in C43(DE3) (method 14), yielding 32 g cell paste from 10 L cell culture. Cells were lysed and Ni affinity purified as described in 7.2.6. Light yellow colouration was observed in fractions 9 – 19, most pure fractions were pooled and concentrated.

Gel filtration chromatography

Concentrated SolA-intein was purified by gel filtration as described in 7.2.6 using an S-200 column (800 mL volume). Retention time of SolA-intein = 400 mL. Most pure fractions were directly frozen at -80 °C.

7.2.9 Purification of LipA-intein

Ni affinity chromatography

LipA-intein was overexpressed from pRJW/3104/40 in C43(DE3), yielding 50 g cell paste from 10 L culture. LipA-intein was purified following the protocol in 7.2.6. Light brown colouration was observed in fractions 13 – 23, most pure fractions were pooled and concentrated.

Gel filtration chromatography

LipA-intein was purified as described in 7.2.6 using an S-200 column (800 mL volume). Retention time of LipA-intein = 360 mL. Most pure fractions were directly frozen at -80 °C.

7.3 Experimental for Chapter 3

7.3.1 BioB-intein aerobic nucleophile dependant cleavage reactions

Reaction buffer was 50 mM Tris/HCl pH 7.0, reactions were incubated at room temperature, samples were removed at 4 and 24 h time points for analysis. Reactions were stopped by quick freezing at -80 °C. Reagent solutions were prepared in 50 mM Tris/HCl buffer, pH 7.0, pH was adjusted to 7.0 using 1 M NaOH / HCl as necessary.

To BioB-intein (60 μL of a 0.7 mg / mL stock) was added one of the following reagents hydrazine, acetic hydrazide, Lucifer yellow, hydroxylamine, ethyl hydroxylamine, benzyl hydroxylamine, DTT, 2-mercaptoethanol, 2-mercaptoimidazole, 4-mercaptobenzoic acid, cysteine or NaI (500 mM stock solutions) for a final concentration 10, 50 or 100 mM, to complete the reaction volume (100 μL) make 50 mM Tris/HCl pH 7.0 was added as required. Negative control reactions contained only BioB-intein and Tris/HCl buffer.

7.3.2 Reaction of BioB-intein with potassium phosphate

BioB intein was exchanged into potassium phosphate buffer pH 7.0 at concentrations 10, 25, 50, 100 and 250 mM, using a PD-10 gel filtration column. Protein was diluted to 0.4 mg / mL and 100 μL samples were incubated at room temperature for 4 / 24 h and samples stored at -80 °C.

Preperation of Fluorescent Label, FTEC

7.3.3 Synthesis of N-Boc-S-Trityl-N-2-aminoethyl-L-cysteinamide(*5*)

N-Boc-*S*-trityl-L-cysteine [1.0 g, 2.16 mmol] was dissolved in DMF (10 mL) and solid carbonyl diimidazole (420 mg, 2.6 mmol) was added. After 30 minutes, ethylene diamine (1.4 mL, 21.6 mmol) was added in a single portion and the reaction stirred at room temperature for a further two hours. The DMF was then removed *in vacuo* and the residue dissolved in dichloromethane (50 mL). The solution was washed with water (5 x 50 mL) and saturated brine (50 mL) and dried with anhydrous magnesium sulfate. The solvent was reduced to ~10 mL *in vacuo*, yielding N-Boc-S-Trityl-N-2-aminoethyl-L-cysteinamide as a white crystalline solid which was collected by filtration and dried *in vacuo* (980 mg, 90 %). M.p. 76 – 80 °C; ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.29 (6H, d, J = 8.1 Hz), 7.21 (9H, m), 6.65 (1H, t, J = 5.2 Hz), 4.89 (1H, d, J = 7.4 Hz), 3.85 (1H, dd, J = 12.5, 6.6 Hz), 3.25 (2H, ddd, J = 11.4, 4.7, 1.5 Hz), 2.78 (2H, t, J = 5.88 Hz), 2.71 (1H, m), 2.54 (1H, dd, J = 12.5, 5.15 Hz), 1.42 (9H, s). ESMS 506.3 (MH^+).

7.3.4 Synthesis of N-Boc-S-Trityl-N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide

N-Boc-S-Trityl-N-2-aminoethyl-L-cysteinamide (500 mg, 0.99 mmol) was dissolved in methanol (20 mL) and a solution of fluorescein isothiocyanate (424 mg, 1.09 mmol) in methanol (30 mL) added drop-wise. The reaction was stirred under nitrogen for 20 hours. The solvent was removed *in vacuo* to yield an orange residue which was redissolved in 80 % MeOH/20 % water (20 mL). The solution was applied to a Supelclean LC-18 reverse phase column (60 mL) that had been pre-equilibrated in 80 % MeOH/20 % water and the compound eluted in the same solvent. The purest fractions were combined and the solvent removed *in vacuo* to afford the slightly unstable N-Boc-S-Trityl-N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide as an orange solid (815 mg, 92 % yield). M.p. 184 – 190 °C; IR (neat) 3256, 2918, 2849, 1667 , 1592, 1487, 1443, 1249, 1158, 840 cm^{-1}; ^{1}H NMR (400 MHz, CD_3OD), δ 8.0 (1H, s), 7.58 (1H, d, J = 8.0 Hz), 7.3-7.0 (15H, m), 6.97 (1H, d, J = 8.5 Hz), 6.66 (2H,

d, J = 9 Hz), 6.6 (2H, d, J = 2 Hz), 6.44 (2H, dd, J = 8.8, 2.0 Hz), 3.88 (1H, m), 3.6-3.8 (2H, m), 3.37 (2H, t, J = 5.15 Hz), 2.48 (2H, d, J = 6.6 Hz), 1.31 (9H, s); ^{13}C NMR (100 MHz, CD_3OD) δ 182.8, 174.0, 173.3, 171.2, 157.0, 154.7, 146.0, 145.6, 141.5, 130.4, 129.5, 128.7, 128.6, 127.7, 127.6, 127.4, 126.4, 120.9, 114.8, 112.0, 103.3, 80.7, 67.7, 64.1, 55.0, 45.0, 39.9, 35.0, 28.4; HRES-MS calculated for $C_{50}H_{47}N_4O_8S_2$ [M + H]: 895.2830; found: 895.2840.

7.3.5 Synthesis of N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide (FTEC)

Water (100 μL), triisopropylsilane (92 μL, 0.448 mmol) and trifluoroacetic acid (4 mL) were added to N-Boc-S-Trityl-N-[2-[[[(fluorescein-5-yl)amino]thioxomethyl]amino]ethyl]-L-cysteinamide (100 mg, 0.112 mmol) under nitrogen. The reaction mixture was stirred at room temperature for 30 minutes. The solvent was removed *in vacuo* to afford a crude oily orange residue that was triturated with diethyl ether (5 mL). The resultant orange solid was collected by filtration and washed with a further portion of diethylether (10 mL) to afford FTEC as an orange solid (55 mg, 89 % yield). M.p. 190 – 195 °C; IR (neat) 3263, 3061, 2933, 1665, 1585, 1535, 1175, 1115 cm^{-1}; ^{1}H NMR (400 MHz, CD_3OD) δ 8.21 (1H, d, J = 2.0 Hz), 7.79 (1H, dd, J = 8.1, 2.0 Hz), 7.20 (1H, d, J = 8.1 Hz), 6.80 (2H, d, J = 8.1 Hz), 6.77 (2H, d, J = 2.0 Hz), 6.62 (2H, dd, J = 8.1, 2.0 Hz), 4.07 (1H, dd, J = 6.6, 5.1 Hz), 3.95-3.80 (2H, m), 3.62-3.49 (2H, m), 3.11-3.03 (2H, m); ^{13}C NMR (100 MHz, CD_3OD) δ 182.4, 169.9, 167.8, 161.5, 153.8, 147.0, 141.2, 130.9, 129.7, 128.3, 125.3, 120.0, 113.3, 111.0, 102.5, 65.9, 55.2, 43.8, 39.3, 25.3; HRES-MS calculated for $C_{26}H_{25}N_4O_6S_2$ [M + H]: 553.1210; found: 553.1201.

7.3.6 HPLC analysis of FTEC

The extent to which FTEC had oxidised to its disulphide was assessed by reverse phase HPLC using a Gilson System Workcenter and a Phenomenex Prodigy 5 μm ODS-2 (150 × 4.6 mm) column. HPLC Buffer A was 0.1 % TFA in 10 % acetonitrile / water, HPLC buffer B was 0.1 % TFA in acetonitrile. The following conditions were used during the separation with linear gradients between time points: flow rate 0.8 mL / min;

detector wavelength 1 = 480 nm; detector wavelength 2, 280 nm; t = 0 min, 95 % A, 5 % B; t = 5.0 min, 95 % A, 5 % B; t = 40.0 min, 10 % A, 90 % B.

Three samples were analysed by HPLC (in duplicate). Firstly a sample of freshly prepared FTEC (1 mg, 3.2.4) was dissolved in 1 mL of HPLC buffer A and immediately analysed. The second sample of freshly prepared FTEC from the same batch (1 mg) was dissolved in 1 mL HPLC buffer A not containing TFA, was left to stand in air for 30 minutes before 0.1 % TFA was added, and the sample analysed. The third sample was a freshly prepared sample of FTEC (1 mg), which was introduced into an anaerobic atmosphere, dissolved in deoxygenated HPLC buffer A (1 mL) not containing TFA and stood in the glovebox for 16 hours (Method 18) before acidification with 0.1 % TFA and immediate analysis.

The following retention times were recorded: FTEC, 19.7 minutes; FTEC disulphide, 21.3 minutes.

7.3.7 Fluorescent labelling of target proteins using FTEC only

Reactions were maintained under anaerobic conditions (Method 18) and contained the following reagents at final concentrations: BioB-intein (0.4 mg / mL), Fpr-intein (0.7 mg / mL), LipA-intein (0.4 mg / mL), FTEC (2.3 mM) and 50 mM Tris/HCl buffer pH 8.1 in a total volume of 100 μL. Negative control reactions contained only protein in Tris/HCl buffer, labelling reactions contained intein fusion proteins, Tris/HCl buffer and FTEC (2.3 mM). Reactions were stopped by freezing at – 80 °C after 24 h and samples were analysed by SDS-PAGE (method 19).

7.4 Experimental for Chapter 4

7.4.1 Comparison of aerobic and anaerobic thiol dependant cleavage reactions

Reaction buffer was 50 mM Tris/HCl pH 8.1, reactions were maintained aerobically at room temperature and at 20 °C anaerobically (method 18). Reactions were incubated for 16 h and stopped by quick freezing at -80 °C. Reaction volume was typically 100 μL. Reagent solutions were prepared in 50 mM Tris/HCl buffer, pH 8.1, pH was adjusted to 8.1 using 1 M NaOH / HCl as required.

To BioB-intein (60 μL of a 0.7 mg / mL stock) was added a thiol solution, thiols were thiophenol, 2-aminothiophenol, 2-methoxythiophenol, mercaptoethanesulphonic acid or 4-fluorothiophenol (500 mM stocks), for final concentration 25 mM. To complete the reaction volume (100 μL), 50 mM Tris/HCl pH 8.1 was added as necessary. Negative control reactions contained only BioB-intein and Tris/HCl buffer. Reactions were analysed by SDS-PAGE (10 %) (method 12, 19), using loading buffer containing DTT.

7.4.2 Assay of thiol activated fluorescent labelling

The anaerobic reaction (method 18) of BioB-, Fpr- and LipA-intein with fluorescent label FTEC was investigated in the presence of thiols (25 mM) thiophenol, 2-methoxythiophenol, MESA or 4-fluorothiophenol over 24 h, reaction volume was typically 100 μL.

Reactions contained 50 mM Tris/HCl buffer, pH 8.1 containing 500 mM NaCl, BioB-intein (60 μL of a 0.7 mg / mL stock), FTEC (2.3 mM), thiol (from a 500 mM stock solution), as required. Reagent solutions were prepared in 50 mM Tris/HCl buffer, pH 8.1, pH was adjusted using 1 M NaOH / HCl as needed. Control reactions (containing BioB-intein only) were compared with reaction using FTEC only and FTEC with additional small thiol. Reaction volume was adjusted to 100 μL using reaction buffer.

Reactions were analysed by SDS-PAGE (10 %) (method 12, 19), using loading buffer containing DTT and were visualised on a UV light box prior to Coomassie staining.

7.4.3 Comparing the progress of labelling reactions under various conditions

The reaction of BioB-intein with FTEC in parallel under a range of conditions was investigated.

All reactions were prepared in 50 mM Tris/HCl buffer pH 8.1 containing 500 mM NaCl. Total reaction volume was 750 μL, reactions contained BioB-intein (0.42 mg/mL) and FTEC (2.3 mM) and to these were added thiophenol (25 mM final concentration from a 500 mM stock in reaction buffer) and TCEP (10 mM final concentration from a 100 mM stock in reaction buffer) as neccesary. Reactions were incubated either aerobically or anaerobically (method 18) and aliquots (40 μL) withdrawn and stored at -80 °C after 1, 2, 3, 4, 5, 6, 8, 10, 14 and 27 hours. The aliquots were thawed and immediately analysed by SDS-PAGE (10 %) (method 12, 19), using loading buffer containing DTT and were visualised on a UV light box prior to Coomassie staining. Images were analysed using GeneTools software (Syngene, Cambridge, UK) to gain individual band fluorescence intensity data for the assessment of reaction kinetics.

7.4.4 Effect of thiophenol concentration on labelling reaction

Labelling reactions (100 μL) contained BioB-intein (0.4 mg / mL), 50 mM Tris/HCl buffer pH 8.1 and thiophenol at concentrations 0, 0.5, 1, 5, 10, 12.5, and 25 mM. Reactions were maintained under anaerobic conditions for 16 hours and analysed by SDS-PAGE (10 %).

7.4.5 Effect of the nature of the target protein upon reaction kinetics

The reaction of BioB-, LipA-, Fpr-, SolA- and BFP-intein with FTEC was investigated. Labelling reactions contained (final concentrations) fusion protein (0.42 mg / mL), thiophenol (10 mM) and FTEC (2.3 mM), total reaction volume was 1 mL. Reactions were maintained under anaerobic conditions (method 18) and samples were withdrawn and stored at -80 °C at timepoints 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 18 and 24 h. Samples were analysed by SDS-PAGE (10 %) (method 12, 19) as described in 7.4.3.

7.4.6 *In vivo* fluorescent labelling

A 10 mL overnight culture (2YT medium, 37 °C, 180 rpm) of pRJW/2960/88 in BL21(DE3) was inoculated from glycerol stock (method 10) and was incubated overnight. This was used as a 1:100 inoculum for 100 mL 2YT. Cell growth was monitored by OD_{600}, and once it had reached 0.6 cells were induced by the addition of arabinose (1 mL of a 20 % solution). The growth temperature was lowered to 27 °C and cells were incubated for 3 h, after which they were harvested by centrifugation (8 min, 10,000 rpm). The cell pellet was washed with phosphate buffered saline (25 mL, three washes, recovery by centrifugation) followed by resuspension in PBS (8 mL). The cells were split and to one half was added FTEC in PBS (to a total volume of 40 mL, [FTEC] = 0.5 mg / mL, containing ampicillin) and to the other PBS (36 mL, containing ampicillin). Cells were incubated for a further 2 h, before harvesting by centrifugation (8 min, 10,000 rpm). Cells were washed with PBS (resuspension in 25 mL, five washes, recovery by centrifugation) and stored at – 80 °C.

Cells were resuspended in lysis buffer (50 mM Tris/HCl containing 10 % glycerol, 1 mL) and lysed by sonication (5 × 3 s bursts), cell debris was separated by centrifugation (13, 000 rpm, 15 mins) and lysate fluorescence analysed on a UV light box.

7.5 Experimental for Chapter 5

Electrochemistry general experimental

All solutions were prepared using deionised water (Whatman still and purification system, resistivity > 10 MΩ / cm) and compounds purchased from Sigma-Aldrich and used as received (unless otherwise stated): alumina powder and polishing cloths (Buehler), *N*-Boc-*S*-trityl-L-cysteine, PyBop and DIPEA (Novabiochem), peptide synthesis grade DMF (Rathburn), HPLC grade acetonitrile and ethanol (Riedel de Haen), glassy carbon rod (3 mm diameter, 100 mm length) for carbon disk electrode construction (Tokai Carbon), glassy carbon plates (10 × 10 × 1 mm) for XPS experiments (Goodfellow). For electrochemical experiments, an in-house built potentiostat was used, with National Instruments Labview 5.1 software. Electrochemical simulations were conducted using DIGISIM® 3.1.

7.5.1 Electrode construction

The 7 mm diameter glassy carbon disk electrodes were manufactured (source unknown) and consisted of a glassy carbon rod sealed into a KEL-F polymer rod (10 mm diameter), with a contact made using a brass rod.

The 3 mm diameter glassy carbon disk electrodes were hand made and were prepared by sealing a short section (5 mm length) of glassy carbon rod into a length of KEL-F polymer (30 mm length, 10 mm diameter) using a slow setting epoxy resin. An electrical contact was made by connecting the carbon with a brass rod using silver epoxy. The brass rod was sealed into the KEL-F polymer using a quick setting epoxy resin.

The glassy carbon plate electrodes (10 × 10 × 1 mm), which were used to prepare samples for XPS measurements, were prepared by contacting one side of the plate with insulated electrical wire using silver paint. Once the paint was dry, this side was coated with epoxy resin, sealing the contact. Once the electrodes had been derivatised, the

epoxy coating and contact could easily be lifted away from the surface, allowing simple mounting of samples on stubs as required for analysis.

7.5.2 Electrochemical modification of a glassy carbon electrode by the reduction of 4-nitrobenzenediazonium salt and its reduction (*6, 7*)

A glassy carbon electrode (7 mm diameter), prepared by polishing in 3, 1 and 0.3 μm alumina slurries was derivatised by electrolysis at -0.7 V vs SCE for 20 minutes in a 1 mM solution of 4-nitrobenzenediazonium salt in a solution of 0.1 M NBu_4BF_4 in freshly distilled acetonitrile, at room temperature. Solutions were degassed with Argon for 10 minutes prior to each experiment, CE = Pt gauze, reference = SCE. The efficiency of coupling was judged by transferring the derivatised electrode into a solution of pure electrolyte and cycling the potential between 0.8 and -1.5 V vs SCE, observing the characteristic reversible 1e couple at -1.2 V vs SCE. The electrode was then placed into a 0.1 M KCl solution in 9:1 water / ethanol and the nitro groups reduced by electrolysis at -1.4 V for 100 minutes.

7.5.3 Derivatisation and characterisation of an aniline modified glassy carbon electrode via conventional coupling chemistry

The coupling of 4-nitrobenzeneacetic acid to the aniline modified glassy carbon surface was attempted by the immersion of the electrode into one of the following mixtures: 4-nitrobenzeneacetic acid (0.18 g, 1 mmole) and either carbonyldiimidazole (0.16 g, 1 mmole) or DCC (0.12 g, 0.6 mmoles) in DMF, acetonitrile or DCM (10 mL in each case). Analysis of coupling was as described in section 7.5.2 for nitro group coverage analysis.

7.5.4 Functionalisation of glassy carbon via oxidation of ethylenediamine (*8*)

A freshly polished (to 0.3 μm alumina) glassy carbon electrode (7 mm diameter) was derivatised with ethylene diamine by electrolysis at a potential of –2 V (vs SCE) in a solution of 0.2 M ethylenediamine in NBu_4BF_4 / acetonitrile for 90 seconds at room

temperature under argon. The counter electrode was a Pt gauze and reference electrode was SCE.

7.5.5 Derivatisation of glassy carbon by the oxidation of decanediamine (DAD) (*9*)

A freshly polished (to 0.3 μm alumina powder) GC 3mm diameter electrode was derivatised by cyclic voltammetry (3 cycles) in the range 0 to +1.25 V vs Ag / Ag^+ in an ethanolic 2 mM solution of DAD / 0.1 M $LiClO_4$. The solution was degassed with argon prior to derivatisation (10 mins), CE = Pt gauze, reference = Ag wire (cleaned in 10 % HNO_3), experiments were run at room temperature. After cycling, the electrode was rinsed with ethanol, followed by sonication for 5 minutes in ethanol and water.

7.5.6 Characterisation of the diaminodecane modified surface

Electrodes (3 mm diameter glassy carbon disk electrodes) were characterised by cyclic voltammetry at varying scan rates (as indicated) in an aqueous solution of 2 mM $Ru(NH_3)_6.Cl_3$ / 0.1 M KCl (pH 5), under argon at room temperature. The counter electrode was Pt gauze, the reference electrode was SCE.

7.5.7 Stability of the decanediamine derivatised surface

The diaminodecane modified glassy carbon surfaces were characterised before and after soaking in one of the following solvents overnight; water, DCM, DMF, acetonitrile or chloroform. Electrodes were dried under a stream of argon prior to analysis.

7.5.8 Coupling of decanediamine with ferroceneacetic acid

Ferroceneacetic acid (FcAA) was coupled to the DAD monolayer by overnight immersion of the GC-DAD electrodes in a 5 mM FcAA / 10 mM EDC aqueous solution, pH 14, adjusted using 5M NaOH, at room temperature. The electrodes were thoroughly washed with water and ethanol prior to use. Films were characterised by cyclic voltammetry in an aqueous 0.1 M solution of $LiClO_4$ at varying scan rate (as

indicated). The solution was degassed with argon prior to analysis (10 mins), CE = Pt gauze, reference = SCE, experiments were run at room temperature.

7.5.9 Coupling and deprotection of an N-Boc, S-trityl protected cysteine derivative to a DAD derivatised electrode

Diaminodecane coupled glassy carbon electrodes were coupled with a protected cysteine derivative by immersion in a mixture of *N*-Boc-*S*-trityl-L-cysteine (0.046 g, 0.1 mmoles), PyBop (*10*) (0.1 g, 0.2 mmoles) and DIPEA (0.065 g, 0.5 mmoles) in DMF (20 mL) overnight. Electrodes were washed thoroughly with DMF (5 × 10 mL) and dried in a stream of nitrogen before deprotection.

7.5.10 Deprotection of the protected cysteine modified electrode

Glassy carbon electrodes modified with a protected cysteine derivative as described in 7.5.9 were deprotected anaerobically by immersion in a mixture of TFA (20 mL), water (0.5 mL) and triisopropyl silane (0.5 mL, 3.2 mmoles) for two hours, followed by washing with thoroughly degassed water (100 mL).

7.5.11 Immobilisation of BioB

Samples of C-terminal MESA (for immobilisation) and L-cysteine (negative control) modified BioB were prepared by reaction of BioB-intein (7.2 μM) with either L-cysteine (50 mM) or MESA (50 mM) anaerobically in buffer (50 mM Tris/HCl buffer pH 8.1 containing 0.5 M NaCl), in a reaction volume of 10 mL overnight. Reactions were concentrated (to a volume of 100 μL) and derivatised BioB purified by gel filtration (high resolution superdex S-75, column volume 50 mL, flow rate = 0.5 mL / min). Most pure fractions were pooled, and added anaerobically to previously deprotected cysteine modified surfaces and allowed to react overnight before washing in buffer (50 mL).

7.5.2 Characterisation of electrode surfaces by XPS

Glassy carbon samples were modified as described above and for transport were secured with a small piece of plasticine in a glass vial (10 mL volume) and were stored in a saturated atmosphere of water (for non-enzyme modified electrodes) or 50 mM Tris/HCl buffer, pH 8.1 containing 0.5 M NaCl for enzyme modified electrodes. For analysis, samples were carefully mounted on an aluminium sample stub and dried in a stream of compressed air immediately prior to introduction into the sample chamber. Samples were analysed with a Scienta ESCALAB300 photoelectron spectrometer (Daresbury, UK), samples were impacted with Al Kα radiation at 1486.7eV, and analysed either by a broad spectrum survey at low resolution (1 eV steps) or core level regions investigated at high resolution (0.05 eV steps).

1. Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual 2nd Ed.*

2. Chung, C. T., Niemela, S. L., Miller, R. H. (1989) *Proceedings of the National Academy of Sciences of the United States of America 86*, 2172-2175.

3. Bradford, M. M. (1976) *Analytical Biochemistry 72*, 248-254.

4. Chong, S. R., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. (1997) *Gene 192*, 271-281.

5. Tolbert, T. J., and Wong, C. H. (2000) *Journal of the American Chemical Society 122*, 5421-5428.

6. Delamar, M., Hitimi, R., Pinson, J. and Savéant, J.M. (1992) *Journal of the American Chemical Society 114*, 5883.

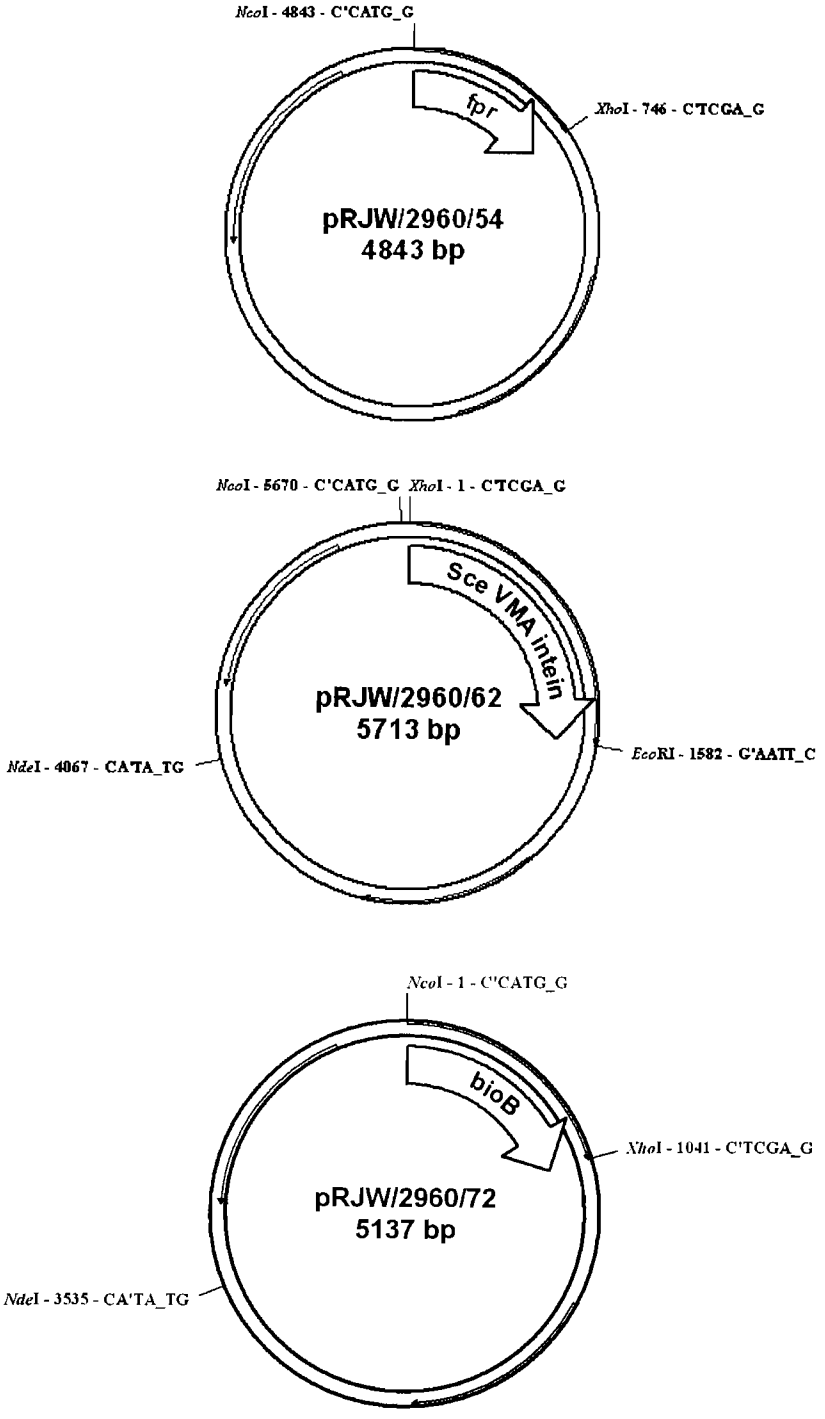
7. Allongue, P., Delamar, M., Desbat, B., Fagebaume, O., Hitmi, R., Pinson, J., and Saveant, J. M. (1997) *Journal of the American Chemical Society 119*, 201-207.

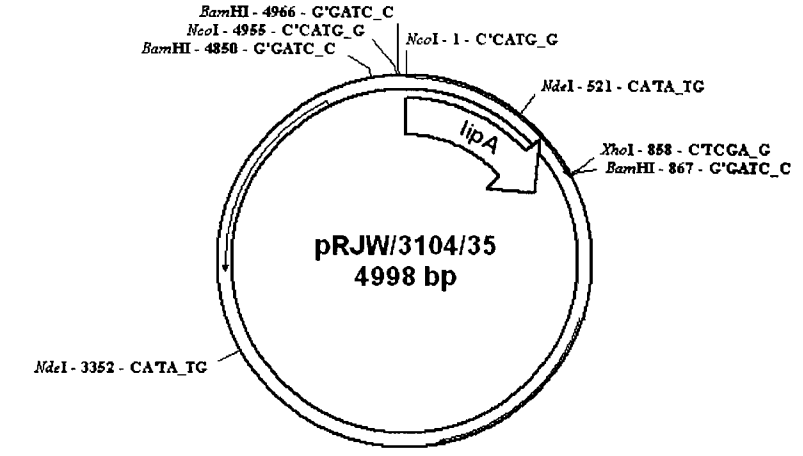
8. Barbier, B., Pinson, J., Desarmot, G., and Sanchez, M. (1990) *Journal of the Electrochemical Society 137*, 1757-1764.

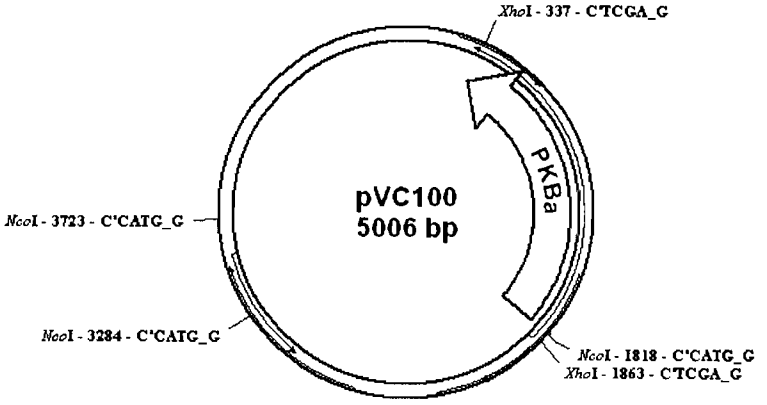
9. Liu, J. a. D., S.,. (2000) *Electrochemistry Communications 2*, 707-712.

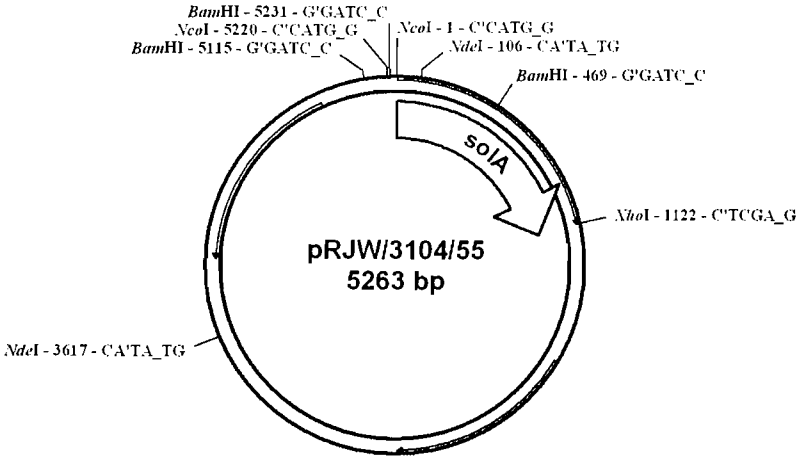
10. Coste, J., Lenguyen, D., and Castro, B. (1990) *Tetrahedron Letters 31*, 205-208.

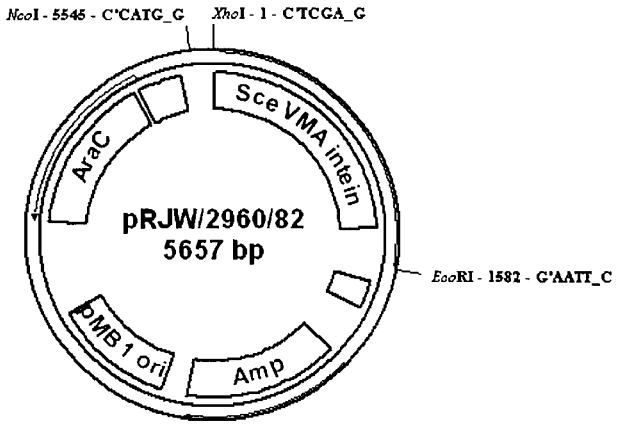
Appendix A:- Plasmid maps

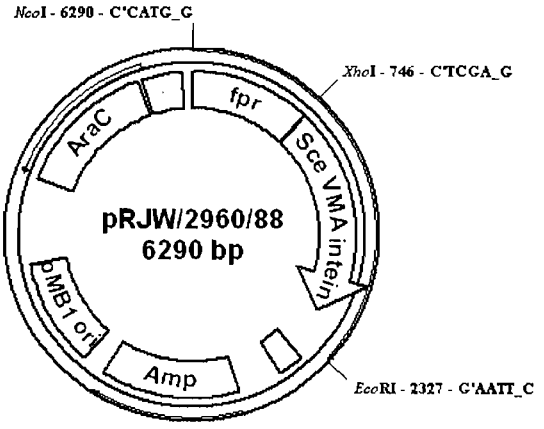


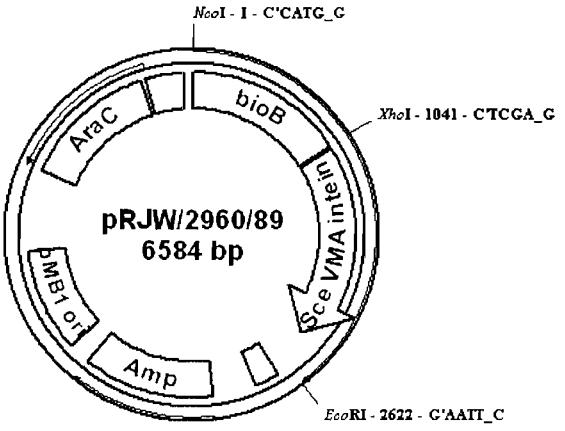

BamHI - 4966 - G'GATC_C
NcoI - 4955 - C'CATG_G
BamHI - 4850 - G'GATC_C
NcoI - 1 - C'CATG_G
NdeI - 521 - CA'TA_TG
lipA
XhoI - 858 - C'TCGA_G
BamHI - 867 - G'GATC_C
pRJW/3104/35
4998 bp
NdeI - 3352 - CA'TA_TG

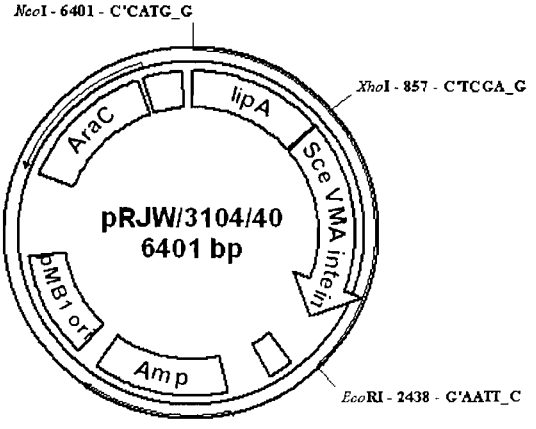

XhoI - 337 - C'TCGA_G
PKBa
pVC100
5006 bp
NcoI - 3723 - C'CATG_G
NcoI - 3284 - C'CATG_G
NcoI - 1818 - C'CATG_G
XhoI - 1863 - C'TCGA_G

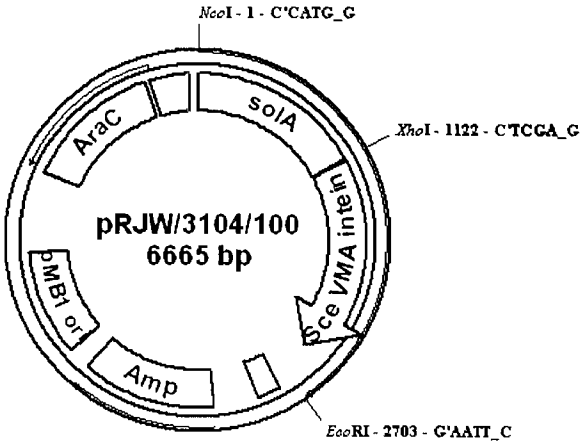

BamHI - 5231 - G'GATC_C
NcoI - 5220 - C'CATG_G
BamHI - 5115 - G'GATC_C
NcoI - 1 - C'CATG_G
NdeI - 106 - CA'TA_TG
BamHI - 469 - G'GATC_C
solA
XhoI - 1122 - C'TCGA_G
pRJW/3104/55
5263 bp
NdeI - 3617 - CA'TA_TG

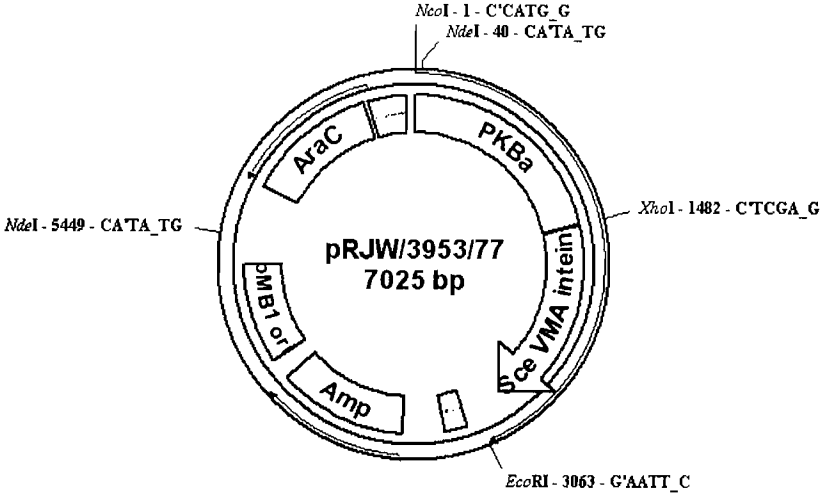

NcoI - 5545 - C'CATG_G
XhoI - 1 - C'TCGA_G
AraC
Sce VMA intein
pRJW/2960/82
5657 bp
EcoRI - 1582 - G'AATT_C
pMB1 ori
Amp


NcoI - 6290 - C'CATG_G
XhoI - 746 - C'TCGA_G
fpr
AraC
Sce VMA intein
pRJW/2960/88
6290 bp
pMB1 ori
Amp
EcoRI - 2327 - C'AATT_C


NcoI - 1 - C'CATG_G
bioB
XhoI - 1041 - C'TCGA_G
AraC
Sce VMA intein
pRJW/2960/89
6584 bp
pMB1 ori
Amp
EcoRI - 2622 - C'AATT_C


NcoI - 6401 - C'CATG_G
lipA
XhoI - 857 - C'TCGA_G
AraC
Sce VMA intein
pRJW/3104/40
6401 bp
pMB1 ori
Amp
EcoRI - 2438 - G'AATT_C


NcoI - 1 - C'CATG_G
AraC
solA
XhoI - 1122 - C'TCGA_G
pRJW/3104/100
6665 bp
pMB1 ori
Sce VMA intein
Amp
EcoRI - 2703 - G'AATT_C


NcoI - 1 - C'CATG_G
NdeI - 40 - CA'TA_TG
AraC
PKBa
NdeI - 5449 - CA'TA_TG
XhoI - 1482 - C'TCGA_G
pRJW/3953/77
7025 bp
pMB1 ori
Sce VMA intein
Amp
EcoRI - 3063 - G'AATT_C