University of Southampton

Faculty of Environmental and Life Sciences

School of Biological Sciences

Structural Basis of Specific IgG Deactivation by Streptococcal Immune Evasion Enzymes

by

Abigail Sophie Louise Sudol

ORCID ID 0000-0002-9420-9758

Thesis for the degree of *Doctor of Philosophy*

March 2024

University of Southampton

Abstract

Faculty of Environmental and Life Sciences
School of Biological Sciences

Thesis for the degree of Doctor of Philosophy

Structural Basis of Specific IgG Deactivation by Streptococcal Immune Evasion Enzymes

by
Abigail Sophie Louise Sudol

Enzymatic cleavage of IgG antibodies is a common strategy used by pathogens for the ablation of immune effector function. The bacterium *Streptococcus pyogenes*, for example, is a common human pathogen and is highly adapted for human infection. The bacterium secretes several enzymes which inactivate human IgG, including the protease IdeS and endoglycosidases EndoS/EndoS2, which catalyse cleavage and deglycosylation of IgG antibodies, respectively. This activity, along with their fine specificity for IgG, has enabled the development of these enzymes with diverse applications. IdeS has received clinical approval for use as a pre-treatment for hypersensitised patients undergoing kidney transplantation, and, along with EndoS, has been tested in several pre-clinical models of autoimmune disease. The endoglycosidases EndoS and EndoS2 have additional application in engineering antibody glycosylation. Despite these extensive therapeutic and biotechnological applications, the structural basis for how these enzymes specifically deactivate IgG is unclear.

This thesis presents crystal structures of IdeS, EndoS and EndoS2 in complex with their IgG1 Fc substrate, which were generated using "less-crystallisable" Fc variants, to evade the common bottleneck of preferential Fc self-crystallisation in the crystallisation of Fc-protein complexes. These structures reveal the molecular basis behind the exquisite specificity of each enzyme for IgG. The IdeS protease displays extensive Fc recognition and encases the antibody hinge region. Conversely, the glycan hydrolase domain in EndoS/EndoS2 traps the Fc glycan in a previously unobserved "flipped-out" conformation, while additional recognition of the Fc peptide surface is driven by the so-called carbohydrate-binding module. This work reveals the molecular basis of specific IgG deactivation by these streptococcal immune evasion enzymes, which will aid in the development of next-generation enzymes for a wide range of applications.

Table of Contents

Table of Tables

Table of Figures

Table of Figures

List of Accompanying Materials

Crystal structures presented in this thesis deposited within the Protein Data Bank (PDB):

1. “Less-crystallisable” IgG1 Fc^{E382S} (PDB 8A48)
2. IgG1 Fc in complex with IdeS (PDB 8A47)
3. EndoS in complex with IgG1 Fc (PDB 8A49)
4. EndoS2 in complex with IgG1 Fc (PDB 8Q5U)

Research Thesis: Declaration of Authorship

Print name: Abigail Sophie Louise Sudol

Title of thesis: Structural Basis of Specific IgG Deactivation by Streptococcal Immune Evasion Enzymes

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

I confirm that:

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

Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. *Nat Commun.* 2022;13:7801

Bauer-Smith H, Sudol ASL, Beers SA, Crispin M. Serum immunoglobulin and the threshold of Fc receptor-mediated immune activation. *Biochim Biophys Acta Gen Subj.* 2023:130448

Pre-print:

Sudol ASL, Tews I, Crispin M. Bespoke conformation and antibody recognition distinguishes the streptococcal immune evasion factors EndoS and EndoS2. *bioRIV.* 2023:553389

Signature: .....Date: 24/03/24

Acknowledgements

I feel incredibly lucky to have been given this opportunity to complete a PhD, definitely the most academically and mentally challenging thing I have done in my life. Writing up the work completed over these past four years almost negates how long and difficult some of these experiments were at the time, and it feels surreal to finally be finishing with all the work presented here. I am monumentally grateful to everyone who has helped along the way, without whom I surely would not have been able to finish.

Firstly, none of this work would have been possible without the constant help and guidance from my supervisors Max Crispin and Ivo Tews. Throughout this project I have discovered a passion for structural biology and this is entirely down to them. Max has always been very enthusiastic and supportive about my area of research which has been a great source of encouragement over the years. Ivo has taught me the vast majority of my knowledge of X-ray crystallography, and has always been very helpful with any tough problems I have encountered. Despite their quite different supervisory styles, being their first joint student has been very successful in my opinion! This work would also not have been possible without the support of Against Breast Cancer and the University of Southampton, who funded this project.

I am extremely fortunate to have been a part of two very kind and supportive lab groups. I would like to thank all the past and present members from the Crispin and Tews labs, all of whom have made the lab environment enjoyable to work in: Himanshi, Lydia, Joel, Maddy, Wenwen, Jacob, Al, Hannah, Grace, Dylan, John, Yasa, Charlie, Charlotte, Hayden, Rachel, Jack, Izzy, Danny, Anjala, Martin, Chris O and Chris H. In particular, I must thank John Butler for being an amazing mentor to me in the lab, and always being around to help with all the problems I have ran into/mistakes I have made. I would also like to thank Dylan Ivory for his help with the various IdeS/EndoS projects and lab work behind the scenes, which was a great help to me particularly during the tough paper writing periods, and Chris Holes for his help setting up extensive numbers of crystal trays for me. Further thanks goes to the beamline staff at Diamond Light Source and the European Synchrotron Radiation Facility, for their assistance with countless beamline trips and unattended data collection sessions. I must also thank Ben Clark for always being around to help with the centrifuge, a particular enemy of mine in the lab!

This PhD has been a very rewarding, but also very tough challenge for me, and I certainly would not have reached this point without the support of everyone around me. In particular, I would like to thank Charlotte, Rosalba and Peter for being amazing housemates throughout my time in Southampton. It has been so nice to live with such great friends going through the same PhD experience as me. I must give a shout-out to Aimee, James, Sandra and Nick (who, despite their insistence otherwise, will likely never read this), for being very supportive friends from afar. Lastly,

Acknowledgements

I would like to thank all my other friends, my boyfriend and particularly my family for their constant support and encouragement. I would not have been able to do this without you all.

Definitions and Abbreviations

3HB ..... 3-helix bundle

AAV ..... Adeno-associated virus

ADCC ..... Antibody-dependent cellular cytotoxicity

ADCP ..... Antibody-dependent cellular phagocytosis

AU ..... Asymmetric unit

C-3HB ..... C-terminal 3-helix bundle

CBM ..... Carbohydrate-binding module

CDC ..... Complement-dependent cytotoxicity

CEF ..... Complement evasion factor

C_H domain ..... Constant heavy domain

C_L domain ..... Constant light domain

cPRA ..... Calculated panel-reactive antibody

Cryo-EM ..... Cryo-electron microscopy

CV ..... Column volumes

DC ..... Dendritic cell

DLS ..... Diamond Light Source

EM ..... Electron microscopy

EndoS ..... Endoglycosidase S

EndoS2 ..... Endoglycosidase S2

ER ..... Endoplasmic reticulum

ESRF ..... European Synchrotron Radiation Facility

Fab ..... Fragment antigen binding

Fc ..... Fragment crystallisable

FcRn ..... Neonatal Fc receptor

FcγR..... Fc gamma receptor

GAS ..... Group A *Streptococcus*

GH ..... Glycosidase

GlcNAC..... *N*-acetylglucosamine

GPI..... Glycosylphosphatidylinositiol

HDX ..... Hydrogen-deuterium exchange

HEPES..... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

hIg..... Hybrid immunoglobulin

HLA..... Human leukocyte antigen

IdeS..... IgG-degrading enzyme of *Streptococcus pyogenes*

IFN-γ..... Interferon-γ

Ig..... Immunoglobulin

IgG..... Immunoglobulin G

IL ..... Interleukin

IPTG ..... Isopropyl β-D-1-thiogalactopyranoside

ITAM..... Immunoreceptor tyrosine-based activation motif

ITIM ..... Immunoreceptor tyrosine-based inhibition motif

JCSG..... Joint Center for Structural Genomics

LDS ..... Lithium dodecyl sulfate

LRR ..... Leucine-rich repeat

MES..... 2-ethanesulfonic acid

MHC..... Major histocompatibility complex

Mincle..... Macrophage inducible C-type lectin

MME ..... Monomethyl ether

MPD 2-methyl-2,4-pentanediol

MWCO ..... Molecular weight cut-off

NCS ..... Non-crystallographic symmetry

NETs ..... Neutrophil extracellular traps

NK cell ..... Natural killer cell

NMR ..... Nuclear magnetic resonance

OD_{600} ..... Optical density at 600 nm

PBS ..... Phosphate-buffered saline

PDB ..... Protein Data Bank

PEG ..... Polyethylene glycol

PepO ..... Endopeptidase O

PI3K ..... Phosphoinositide 3 kinase

PLCγ ..... Phospholipase C gamma

PRL ..... Proline-rich loop

PTM ..... Post-translational modification

RBC ..... Red blood cell

RMSD ..... Root-mean-squared deviation

ROS ..... Reactive oxygen species

Rpm ..... Rotations per minute

RSRZ ..... Real Space R-value normalised by residue type and resolution

SDS-PAGE ..... Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SEC ..... Size exclusion chromatography

SH2 ..... SRC homology 2

SIC ..... Streptococcal inhibitor of complement-mediated lysis

SLO ..... Streptolysin O

SLS ..... Streptolysin S

SpeB ..... Streptococcal pyrogenic exotoxin B

SpyCEP ..... *S. pyogenes* cell envelope proteinase

SRC ..... Src family kinases

TLS ..... Translation-libration-screw

TMAO ..... Trimethylamine-*N*-oxide

TRIM21 ..... Tripartite motif-containing protein 21

V_H domain ..... Variable heavy domain

V_L domain..... Variable light domain

WT ..... Wild-type

γIMCs ..... Interferon-γ-producing immature myeloid cells

Chapter 1 Introduction

During infection, the human body works to destroy the invading pathogen using a series of immune defences [1-3]. Unfortunately, this results in a strong selective pressure on the pathogen to find ways to survive, and has thus resulted in the evolution of sophisticated strategies for evasion of the human immune system. During instances in which immune activation becomes aberrant, such as in autoimmune disease, or is unwanted, such as following organ transplantation, strategies for dampening the immune response are an attractive therapeutic avenue.

The bacterium *Streptococcus pyogenes* is highly adapted for human infection and thus possesses a diverse range of immune evasion mechanisms, including enzymes which directly modulate components of the immune system [4, 5]. This bacterium secretes several enzymes which target immunoglobulin G (IgG) antibodies: IgG-degrading enzyme of *S. pyogenes* (IdeS) [6], endoglycosidase S (EndoS) [7] and endoglycosidase S2 (EndoS2) [8]. The ability of these bacterial agents to specifically deactivate IgG has prompted their investigation as biologics for various therapeutic and biotechnological applications, as well as extensive research into their structure and function, which will be discussed in the following sections. This thesis describes work towards elucidating the structural basis behind the exquisite substrate specificity of the aforementioned enzymes for IgG antibodies, in order to facilitate their continued development.

Some of the figures included in this chapter and prepared by the author (specifically, **Figure 4**, **Figure 5**, **Figure 7** and **Figure 12**) have been adapted or reproduced from the publication: Bauer-Smith H, Sudol ASL, Beers SA, Crispin M. Serum immunoglobulin and the threshold of Fc receptor-mediated immune activation. *Biochim Biophys Acta Gen Subj.* 2023:130448.

1.1 The human immune system

The immune response against an incoming threat can be classified into two main branches: innate and adaptive immunity. In brief, the innate immune response constitutes the first line of defence and is non-specific, i.e., not tailored towards the specific foreign antigen recognised [1]. Components of innate immunity in humans include: physical barriers to infection (e.g., skin and mucosal membranes), physiological barriers (e.g., acidic pH in the stomach), non-specific phagocytic/endocytic cells and inflammatory responses to infection (causing influx of phagocytic cells to the infection site) [1, 2]. In contrast, adaptive immunity is dependent upon recognition and destruction of a specific antigen, and retains “memory” of that antigen, so that a targeted immune response can be launched upon future exposure to the same antigen. Cells involved in adaptive immunity include T and B lymphocytes [1, 3].

T cells are activated by antigen-presenting cells, which present foreign peptides using a complex of cell surface proteins called the major histocompatibility complex (MHC), or also known as human leukocyte antigens (HLAs) [1]. These proteins are integral in distinguishing "self" antigens from foreign antigens, and thus are the cause of rejection during organ transplantation (discussed further in section 1.4.1). MHC-peptide complexes bind immature T cells with compatible T cell receptors, causing the cells to differentiate into more specialised cells types. One such subtype, termed T helper cells, subsequently secrete cytokines which cause further activation of B cells [1, 3].

B cells are stimulated directly by foreign antigens, which bind compatible proteins called antibodies present on their cell surface. This causes their differentiation into several specialised cells: plasma cells, which produce and secrete antibodies specific towards the target antigen, and memory cells, whose function is to produce large amounts of these antibodies upon antigen re-exposure. These antibodies, also called immunoglobulins, bind cells displaying the target foreign antigen, marking them for immune destruction [1, 9]. The structure and function of these proteins are discussed in the following section.

1.1.1 Immunoglobulins

Immunoglobulins (Igs) are Y-shaped proteins produced by B cells which target specific "foreign" antigens present on pathogen cell surfaces, following initial recognition of said antigen by the immune system. Ig binding to a foreign antigen has the function of marking the pathogen for immune destruction (termed opsonisation) by various immune effector pathways [1, 9], some of which will be discussed in the following sections. Igs are comprised of two identical heavy and two identical light chains linked together with disulphide bonds (**Figure 1**) [10]. In humans, there are five types of heavy chain (α, δ, ε, γ and μ), which define the Ig isotype as either IgA, IgD, IgE, IgG or IgM, respectively. There are also two forms of light chain: κ and λ. IgA and IgG antibodies can be further classified into two and four distinct subclasses (IgA1–2 and IgG1–4), respectively [9, 11, 12] (**Figure 1**).

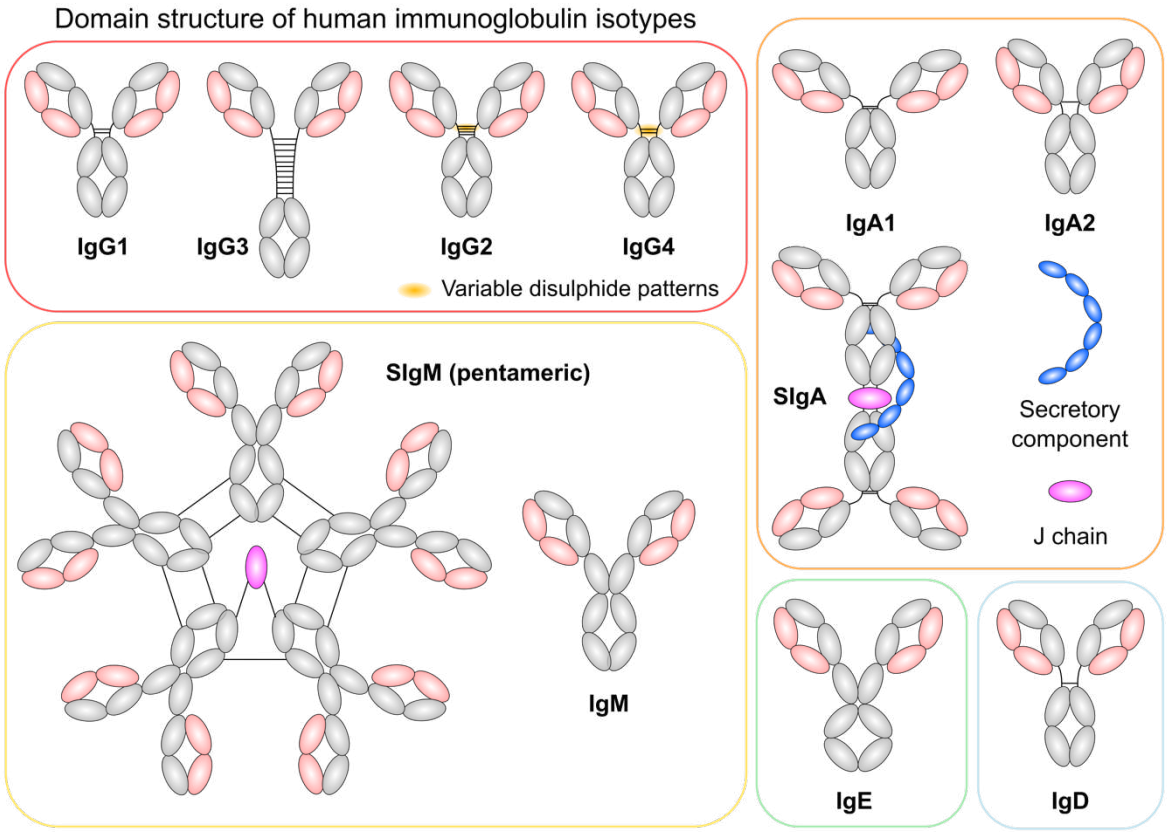


Figure 1: Domain structure of human immunoglobulins. There are five distinct Ig isotypes in humans: IgA, IgD, IgE, IgG and IgM. IgA and IgG can be classified into further subclasses: IgA1–2 and IgG1–4. Igs are homodimers, comprised of two identical heavy chains (depicted in silver), which are each linked to two identical light chains (depicted in pink). Heavy chains comprise an N-terminal variable (V_H) domain, followed by three constant (C_H) domains (exceptions are IgE and IgM which contain four C_H domains). The first and second C_H domains in IgA, IgD and IgG are linked by a flexible hinge region. IgA antibodies can form dimers by associating with the J-chain (dark pink) and secretory component (blue). IgM can form larger oligomers, such as pentamers, by association with the J-chain [9]. *N*- and *O*-linked glycosylation is omitted for clarity.

Within each Ig heavy and light chain, there exists an N-terminal, highly variable (V) region, which binds target antigens. The remaining chain is comprised of varying numbers of constant (C) domains, whose sequence is largely conserved (for each Ig isotype) [10]. During B cell maturation, genes encoding these V and C regions are assembled in tightly-regulated recombination pathways [12-14], resulting in Ig production. In brief, Ig variable regions, encoded by V_H, D_H and J_H genes (V_L and J_L in the light chain gene locus) are shuffled in a process called V(D)J recombination, which produces randomly-altered variable region sequences [12, 13]. Upon activation by a specific antigen which can bind the presented Ig variable regions, B cells undergo further genetic alteration to fine-tune Ig production, without affecting the antigen specificity [12]. Affinity maturation of antibodies is achieved by a process called somatic hypermutation, which introduces further mutations into the Ig variable region gene loci [14, 15]. Moreover, a process called class-switching modifies the Ig heavy chain locus (C_H), which alters the Ig isotype produced. This mechanism functions by deletional recombination (and is therefore irreversible): the isotypes which can be "switched" to thus depends

on the order of C_H genes within the locus [9, 11, 16]. Such antibody production is regulated by cytokine signals, which direct the B cell to produce a particular Ig isotype [12, 16].

The various Ig isotypes possess different structure, and thus perform different functions as part of the adaptive immune response. IgM is the first isotype secreted by B cells, and is therefore important in initial pathogen neutralisation and clearance. The multimeric structure of IgM (typically pentameric, as shown in **Figure 1**, but it can also be hexameric) provides high avidity towards antigens and so, upon binding, IgM is a strong inducer of the complement pathway (detailed in section 1.1.3) [9, 17, 18]. IgD plays a role in mucosal immunity [19], while IgE mediates allergic responses, immune responses against parasites, and growing evidence suggests a role in protection against tumours [9, 20]. IgA is the most abundant antibody within humans and plays a central role in the neutralisation of a diverse range of pathogens, particularly at mucosal surfaces [21]. Both IgA isotypes can form dimers, stabilised by association with the secretory component (**Figure 1**); so-called "secretory IgA" (SIgA in **Figure 1**) is transported to the gut lumen and has a role in coating microbes for maintenance of gut homeostasis [9, 22, 23]. In-depth discussions of these antibody isotypes are beyond the scope of this thesis, which is focussed on IgG antibodies and their deactivation by bacterial immune evasion factors.

1.1.2 Immunoglobulin G

IgG is the most abundant Ig isotype found within human serum and is thus involved in the majority of humoral antibody responses. It can be classified into four subtypes (IgG1–4; numbered in order of decreasing abundance, with IgG1 being most abundant), which are highly similar in amino acid sequence but possess different functional properties [11, 24]. IgG1 and IgG3 antibodies induce strong immune responses against protein antigens, by activation of Fcγ receptors (FcγRs) and the complement pathway (discussed in section 1.1.3). IgG4 also responds to protein antigens, but is typically produced following repeated antigen exposure [11, 24, 25]. In contrast, IgG2 responses are induced following exposure to polysaccharide antigens [11, 24-26].

Owing to their sequence similarity, IgG antibodies have a similar domain architecture, comprising two heavy and two light chains, which are linked by disulphide bonds. Each heavy chain comprises an N-terminal variable (V_H) domain, followed by three constant domains (C_H1, C_H2 and C_H3; also called Cγ1, Cγ2 and Cγ3): similarly, the light chains possess and N-terminal variable (V_L) domain followed by a constant region (C_L). As depicted in **Figure 2**, these chains assemble to form the characteristic "Y" shape of antibodies, comprised of two "fragment antigen binding" (Fab) regions, which bind target antigens, and a "fragment crystallisable" (Fc) region, which mediates immune activation. A flexible hinge region links the Cγ1 and Cγ2 domains, which permits conformational flexibility of the molecule. Furthermore, IgG antibodies all have a conserved *N*-linked glycan site

within each Cγ2 chain, at amino acid N297 [11]. This moiety is implicit in Fc-mediated effector functions of IgG and will be discussed further in section 1.1.4.

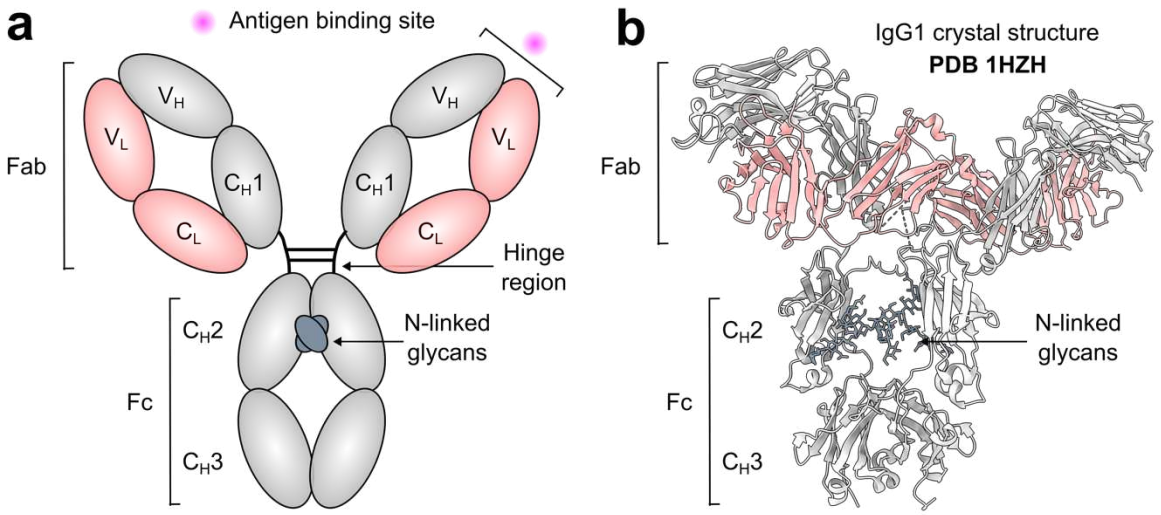


Figure 2: Structure of IgG antibodies. a Schematic representation of IgG1. **b** Crystal structure of full-length IgG1 (PDB 1HZH [27]). **a, b** Heavy and light chains are depicted in light grey and light pink, respectively. Fc *N*-linked glycans at N297 are shown in dark grey. Antigen binding sites are located at the tip of each Fab, as indicated with a pink circle.

There are a few notable differences between the structures of IgG subclasses. The greatest diversity exists within the hinge region (**Figure 3**), particularly in the case of IgG3, which contains up to 62 amino acids (dependent upon the allotype) [24]. This confers an elongated structure of IgG3 compared to the other IgG subclasses [28]. This extended hinge has been suggested to enable improved antigen binding [29] and improved ability to induce phagocytosis [30], and has also been implicated in FcγRI binding, although several studies show that the ability of IgG3 to bind complement with high affinity is not due to its longer hinge region [31-33]. Recent cryo-EM structures of antigen-bound IgG3 alone and in complex with complement factors have provided further insight: the extended hinge in IgG3 allows tighter packing of the Fab domains on the antigenic cell surface, which may enhance neutralisation, and the increased distance from the cell surface to the Fc may enable improved receptor binding when compared to complexes with IgG1/IgM) [34].

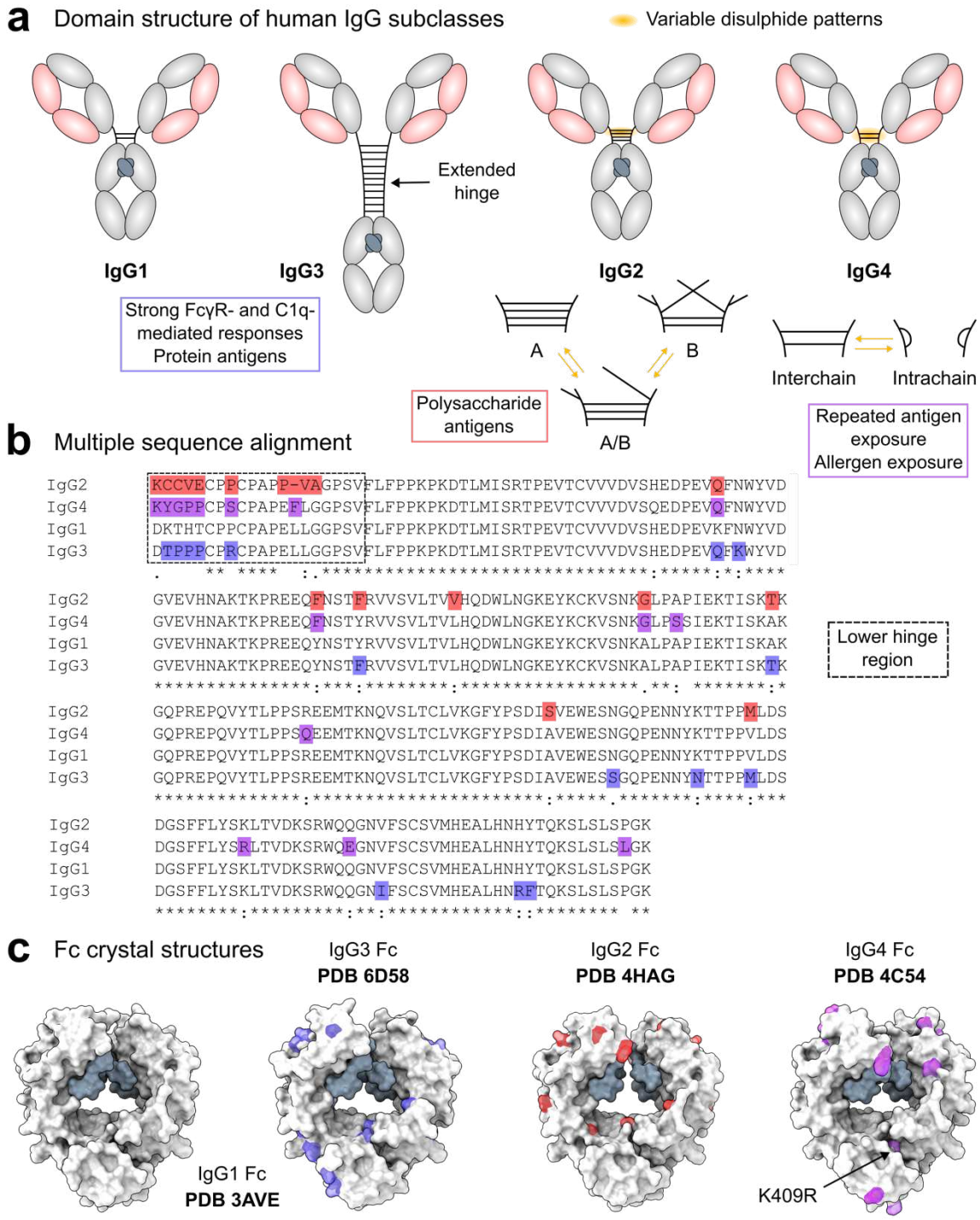


Figure 3: Structural comparison of IgG isotypes. a Domain structure of IgG subclasses, as depicted in **Figure 2**. Varying disulphide bridges within the hinge regions of IgG2 and IgG4 are indicated. **b** Multiple sequence alignment of IgG1–4 lower hinge (from amino acid 221, as numbered in IgG1) and Fc regions, generated using Clustal Omega [35]. Sequence corresponding to sections of the lower hinge region is indicated with a dashed box. Amino acid substitutions in IgG2, IgG3 and IgG4 (relative to IgG1) are coloured red, blue and purple, respectively. **c** Crystal structures of IgG1, IgG2, IgG3 and IgG4 Fcs (PDB 3AVE [36], 4HAG [37], 6D58 (deposited but not yet published) and 4C54 [38], respectively). Amino acid substitutions relative to IgG1 are mapped onto the corresponding structure and coloured as in **b**.

IgG2 and IgG4 antibodies have slightly shorter hinge regions (of 12 amino acids) compared to IgG1, whose hinge length is 15 amino acids [24]. In addition, IgG2 can be classified into further subclasses (IgG2A, A/B and B; **Figure 3a**) owing to different disulphide formations within its hinge region and

Fab domains [39]. This confers increased rigidity of the IgG2 hinge and has been shown to result in "super agonistic" antibodies [40-42]. Within IgG4, the exchange of P228 to a serine (compared to IgG1) allows for easier hinge disulphide breakage. This modification, along with a K409R exchange (relative to IgG1; **Figure 3c**) which impedes interactions between Cγ3 domains, promotes the "Fab arm exchange" observed in IgG4 antibodies, which enables formation of bi-specific IgG4 antibodies [43-45].

1.1.3 Immune activation by IgG

IgG antibodies display a wide range of effector functions in response to detected pathogens, which is summarised in a recent review by Lu *et al.* [46]. Firstly, IgG can exhibit neutralisation responses, which simply involves binding of the antibody Fab regions to antigens, in order to inhibit the pathogen in some way. For example, IgG can neutralise toxins produced by bacteria and/or prevent pathogen entry into cells. Antibodies have also been shown to inhibit other virulence factors which enable invasion or pathogenesis, such as by inhibiting biofilm formation [47] or preventing bacterial adherence to epithelial cells [48]. Although immune activation by the Fc domain is integral in IgG-mediated immune responses, some studies have shown that neutralisation with antibody Fab fragments can provide immune protection [48-50].

The Fc region of antibodies is responsible for binding immune effector molecules, and thus mediates immune activation against foreign antigens. Immune activation is induced by activation of signalling pathways, which are discussed in detail below.

Fcγ receptor-mediated immune activation

IgG Fc binds a series of receptors present on immune effector cells, termed Fcγ receptors (FcγRs). These receptors are classified according to affinity for monovalent IgG: FcγRI is termed the "high affinity" receptor, while the remaining receptors are classed as "low affinity" [51] (**Figure 4**). FcγRs have different affinities for the various IgG subclasses [52] and are populated differently across different cell types [53] (**Table 1**). Fc-mediated effector functions induced by binding FcγRs include antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP). ADCC is largely mediated by Fc binding to FcγRIIIA on natural killer (NK) cells [54]; this mechanism plays a role in clearing viral infections, and is increasingly being utilised for tumour cell killing by therapeutic antibodies [46, 55, 56]. In contrast, ADCP is mediated by a range of phagocytes and granulocytes, and is a commonly-utilised mechanism against bacterial pathogens [46].

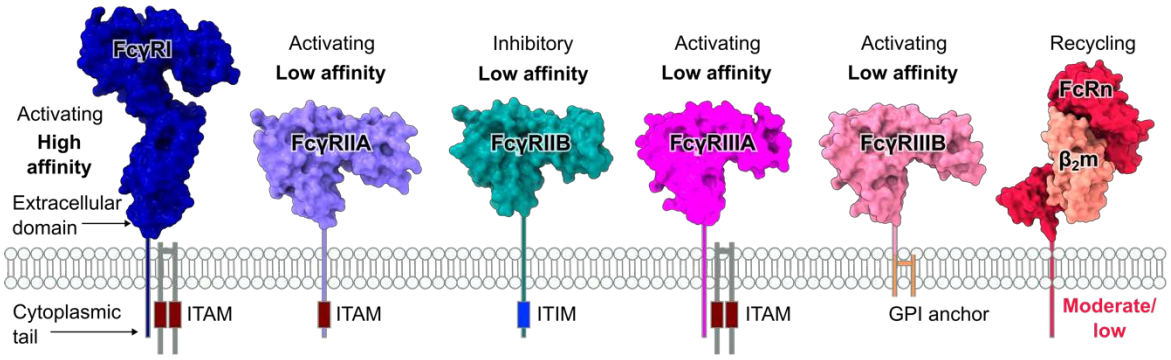


Figure 4: Structure of human FcγRs. Crystal structures for soluble FcγR domains are shown (PDB codes 4W4O [57], 1H9V [58], 3WJJ [59], 5XJE [60] and 6EAQ [61] for FcγRs I, IIA, IIB, IIIA and IIIB, respectively), showing overall domain architecture. FcγRI and FcγRIIIA cytoplasmic tails associate with a common γ-chain (shown in grey) for downstream signalling. FcγRIIIB is associated to the membrane via a GPI anchor. ITAM, immunoreceptor tyrosine-based activation motif; ITIM, immunoreceptor tyrosine-based inhibition motif; GPI, glycosylphosphatidylinositiol. Adapted from Figure 1 in Bauer-Smith *et al.* [62].

Table 1: Properties of human FcγRs.

Receptor	FcγRI	FcγRIIA		FcγRIIB	FcγRIIIA		FcγRIIIB
Allele	-	H131	R131	-	F158	V158	-
K_D (nM)[a]	15.4	192	286	8,330	855	500	5,000[b]
Subclass preference	1>3>4	1>3>2>4	1>3>4>1	4>3>1>2	3>1>4>2	3>1	3>1
Expression profile[c]	Monocytes/ macrophages, DCs, (neutrophils, mast cells)[d]	Monocytes/ macrophages, neutrophils, DCs, basophils, mast cells, eosinophils		B cells, DCs, basophils	NK cells, monocytes/ macrophages		Neutrophils

[a]Affinity constants are reported for IgG1 (to 3 significant figures), and calculated from K_A values reported in Bruhns *et al.* [52]. [b]Data is reported for the NA1 isoform of FcγRIIIB. [c]Expression profiles are reported from a review by Bruhns [53]. [d]Parentheses indicate inducible expression.

Upon Fc binding and subsequent FcγR clustering, downstream signalling pathways are induced to modulate activity of the immune effector cell. The majority of FcγRs possess an immunoreceptor tyrosine-based activation motif (ITAM), either within their cytoplasmic tail directly or by association with a common γ-chain, which mediate immune cell activation (**Figure 5**). In NK cells, activation of this pathway results in the release of granzymes and perforins which lyse the target cell [63, 64]. In phagocytic cells, it regulates microtubule growth to promote phagosome formation [46]. FcγRIIB conversely contains an immunoreceptor-based inhibition motif (ITIM), which inhibits downstream signalling [11]. The interplay between activating and inhibitory FcγRs mediates whether an immune cell becomes activated or not, which is affected by the expression profile of receptors on a particular cell type [65]. In addition, some antibodies show preferential binding to the inhibitory FcγRIIB, such as IgG4, despite its lower binding affinity to the other FcγRs. IgG4 has therefore been proposed to

act as a dampener of immune activity, by blocking prolonged activity of stronger binders IgG1 and IgG3, or preventing sustained responses against allergens by IgE [43].

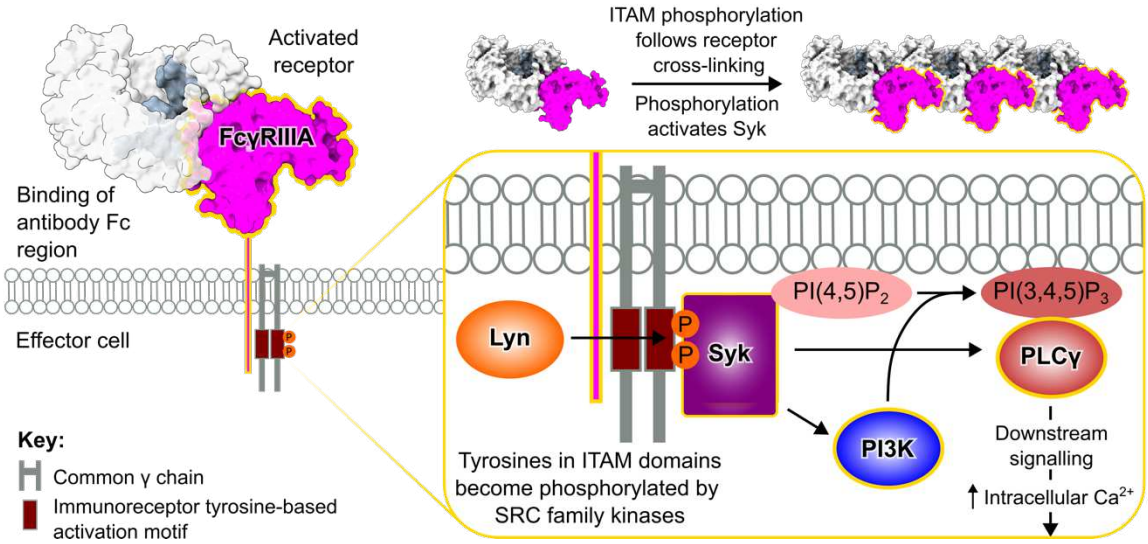


Figure 5: Signalling by ITAM domains within FcγRs. Tyrosine residues in ITAM motifs become phosphorylated by SRC family kinases, such as Lyn, following cross-linking of cell surface FcγRs by IgG-immune complexes. This leads to recruitment and activation of SRC homology 2 (SH2)-containing kinases, such as Syk. This results in downstream activation of phospholipase C gamma 1 (PLCγ) and phosphoinositide 3 kinase (PI3K). PI3K catalyses phosphorylation of $PI(4,5)P_2$ into $PI(3,4,5)P_3$ present in the plasma membrane, which serves as a docking site for PLCγ, thus recruiting it to the membrane and promoting further phosphorylation and activation. PLCγ generates inositol triphosphate and diacylglycerol second messenger molecules, which increases intracellular Ca^{2+} and activates protein kinase C, which induces various immune effector functions [66]. Adapted from Figure 1 in Bauer-Smith *et al.* [62].

IgG Fc binds asymmetrically to FcγRs in a 1:1 stoichiometric ratio, which involves residues in the IgG hinge region, the Cγ2 domain and the Fc *N*-linked glycan [11, 67, 68]. IgG1 and IgG3 bind to all FcγRs, while IgG2 and IgG4 generally show much weaker binding [52]. Amino acid variations have been implicated in the distinct FcγR binding profiles of IgG isotypes. For example, IgG2 contains several exchanges within the lower hinge region (such as E233P, L235A and deletion of G236, relative to IgG1) which contribute to its reduced binding [67, 69-71]. Similarly, IgG4 antibodies have been shown to induce lower levels of immune activation when compared to other subclasses [43, 52, 72, 73], which is likely due to substitution of amino acids implicated in IgG1 binding FcγRs and complement (including L234F, A327G and P331S, relative to IgG1) [38, 43, 67, 74, 75].

Complement-mediated immune activation

The IgG Fc region also mediates immune effector function by activation of the complement cascade, termed complement-dependent cytotoxicity (CDC). Upon Fc binding to the C1q subunit within the C1 complex (comprising $C1qr_2s_2$), the C1r and C1s proteases are activated, which initiates a cascade involving cleavage of other complement proteins, and culminates in formation of a 'membrane attack complex' (comprised of C5b, C6, C7, C8 and C9 subunits) which induces pore formation and lysis

of the target cell. Activation of the complement cascade also results in opsonisation of target cells with cleavage fragments (C3b, C4b and C1q), thereby marking the cells for immune destruction, and recruitment of immune cells to the infection site (by production of C3a and C5a peptides) [76]. C3a and C5a peptides can also induce pro-inflammatory cytokine release and generation of reactive oxygen species to aid in pathogen clearance [77].

IgG Fc binds C1q across a similar interface to the FcγRs (**Figure 6**). IgG1 and IgG3 are strong activators of complement, with reported affinity constants (K_A) of ~1.5×10^8 M^{-1}, while IgG2 and IgG4 display little to no complement binding and activation [24, 75, 78]. As elucidated with binding and mutagenesis studies, amino acids involved in the C1q binding interaction include L235 in the hinge region [79], along with D270, K322, P329 and P331 in the Cγ2 domains [75, 80, 81]. A cryo-electron microscopy (cryo-EM) structure of the Fc-C1q complex also found Fc residues 266–272, 294–300 and 325–331 to be implicit within the binding interface [82]. The reduced ability of IgG2 and IgG4 to activate complement is largely attributed to L235A [79] and P331S [75] substitutions, respectively (relative to IgG1). As elucidated by the cryo-EM structure [82], IgG Fc can bind complement as a hexameric ring, in which one Fc interacts with one C1q subunit (**Figure 6**). Complement activity can thus also be modulated by amino acid substitutions which promote formation this hexameric assembly [83, 84], such as E345K and E430G exchanges (within IgG1) [84]. Fc *N*-glycosylation also affects this binding interaction, as discussed in section 1.1.4.

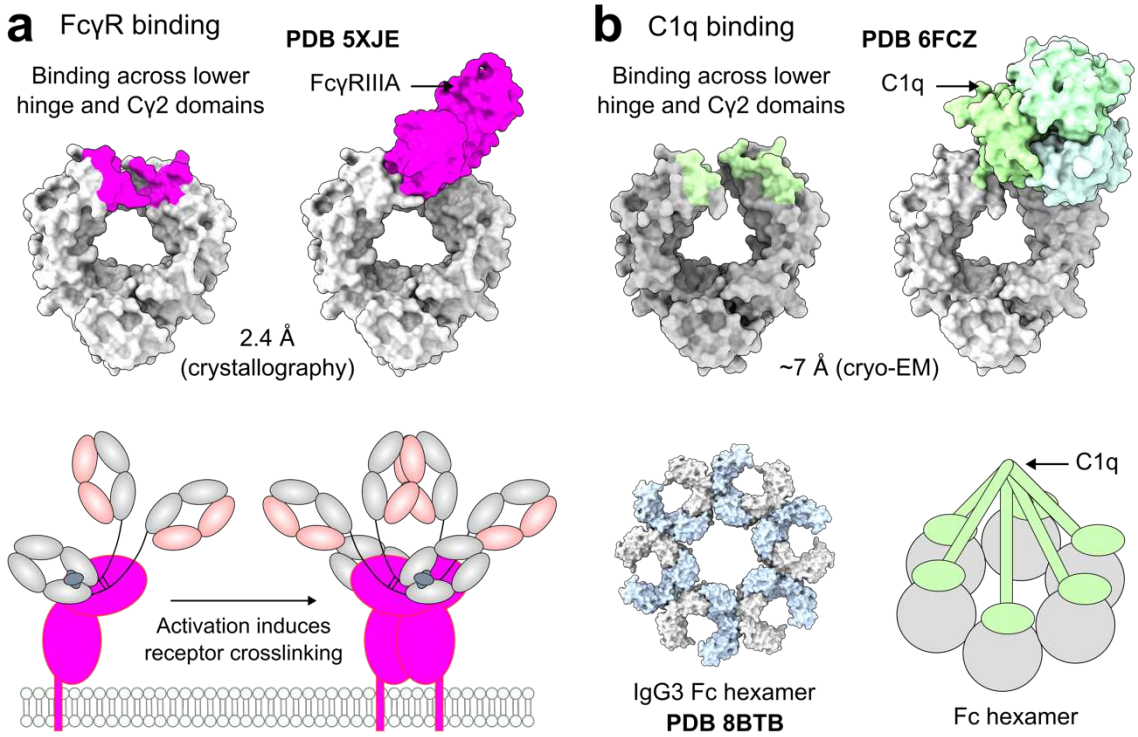


Figure 6: Comparison of FcγR and complement activation by IgG Fc. IgG Fc binds FcγRs (**a**) and complement factor C1q (**b**) across similar regions of its protein surface (lower hinge region and Cγ2 domains). Experimentally-determined structures of IgG Fc in complex with FcγRIIIA (crystal structure; PDB 5XJE [60]) and C1q (cryo-EM; PDB 6FCZ [82]) are depicted. Interfaces of each protein with IgG Fc are mapped onto the Fc protein surface, as calculated using PDBePISA [85]. Schematic models of binding are depicted. **a** FcγR binding induces formation of immune complexes, which activates downstream signalling. **b** A hexameric ring of Fcs (e.g., as observed in PDB 8BTB [34]) binds C1q, with one Fc homodimer binding one C1q subunit. *N*-linked glycans are omitted for clarity.

A recent study by So *et al.* demonstrated that C1q can directly inhibit Fc:FcγR interactions and prevent FcγR-mediated functional effects, particularly for larger immune complexes [86]. This is likely due to the overlapping Fc binding sites of C1q and FcγR, and suggests an additional role of Fc:C1q binding as a buffering mechanism against prolonged immune complex activation of FcγRs.

Other ligands for IgG Fc

IgG Fc has several other binding partners which modulate its immune activation. A notable example is the neonatal Fc receptor (FcRn), which is expressed across a wide range of tissues and cell types, particularly myeloid cells [87, 88]. Binding to IgG occurs at the Fc Cγ2–Cγ3 interface [89], at the low pH of 5–6 within early and recycling endosomes (due to the protonation of H310 and H435 side chains), which enables IgG salvage from lysosomal degradation; the Fc subsequently dissociates at physiological pH of ~7.4 [90-92] (**Figure 7**). A single amino acid exchange of H345R, present in some allotypes of IgG3 (relative to IgG1), has been implicated with the reduced half-life of this subclass compared to IgG1,2 and 4 [93].

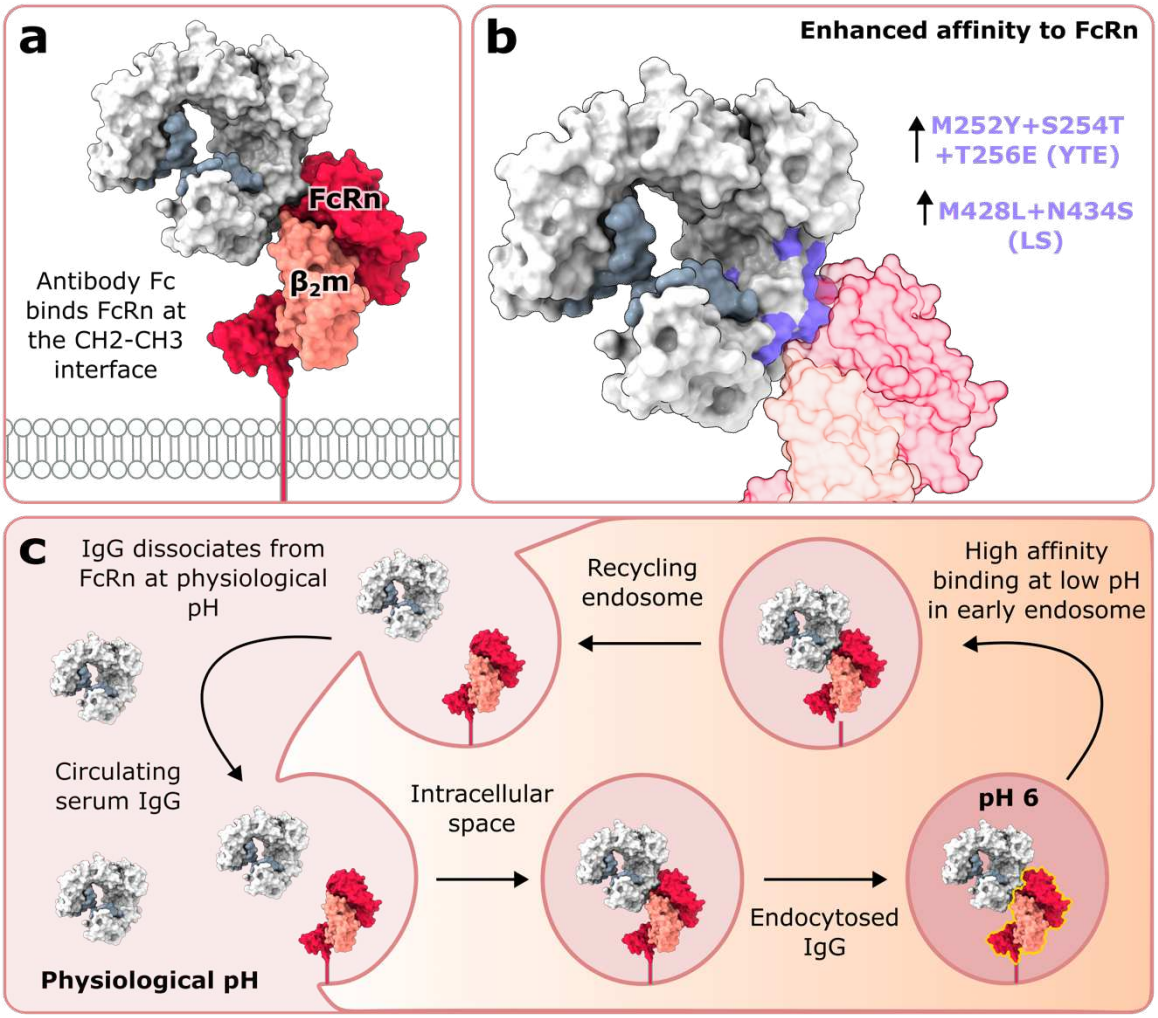


Figure 7: FcRn and biodistribution of IgG. a Crystal structure of IgG Fc bound to human FcRn (PDB 7Q15 [94]) shows how the receptor binds IgG Fc at the interface between its Cγ2 and Cγ3 domains. **b** Amino acid variations in IgG Fc which affect binding affinity to FcRn, which are currently being utilised in the clinic [95, 96]. IgG Fc residues implicated in FcRn binding are coloured purple. **c** FcRn regulates the *in vivo* persistence of IgG and enables biodistribution of IgG within tissues. IgG can be salvaged in early or sorting endosomes from lysosomal degradation, due to increased binding affinity for FcRn at acidic, endosomal pH (pH 6). Figure is taken from Bauer-Smith *et al.* [62].

Tripartite motif-containing protein 21 (TRIM21) is an additional IgG Fc binding partner [11]. In brief, this molecule is a cytosolic receptor expressed highly in immune cells and binds Fc at a similar position to FcRn (at the Cγ2–Cγ3 interface) [97]. TRIM21 binds to intracellular pathogens opsonized by IgA, IgG or IgM, which triggers their degradation by the proteasome, along with transcription of cytokines which regulate immune function [98, 99].

1.1.4 IgG Fc glycosylation

The Fc region of IgG has a conserved *N*-linked glycosylation site at N297 on each Cγ2 heavy chain (**Figure 8**). This post-translational modification is implicit in maintaining thermal stability and structural integrity of the Fc [100-105], and is involved in Fc-mediated effector functions [11, 103, 106]. The core glycan structure comprises *N*-acetylglucosamine (GlcNAc) and mannose

monosaccharides; further processing occurs to add various amounts of galactose, core fucose, sialic acid and/or bisecting GlcNac. The precise composition of *N*-linked glycans is modulated by glycosyltransferase and glycosidase enzymes present in the endoplasmic reticulum (ER) and Golgi apparatus [107].

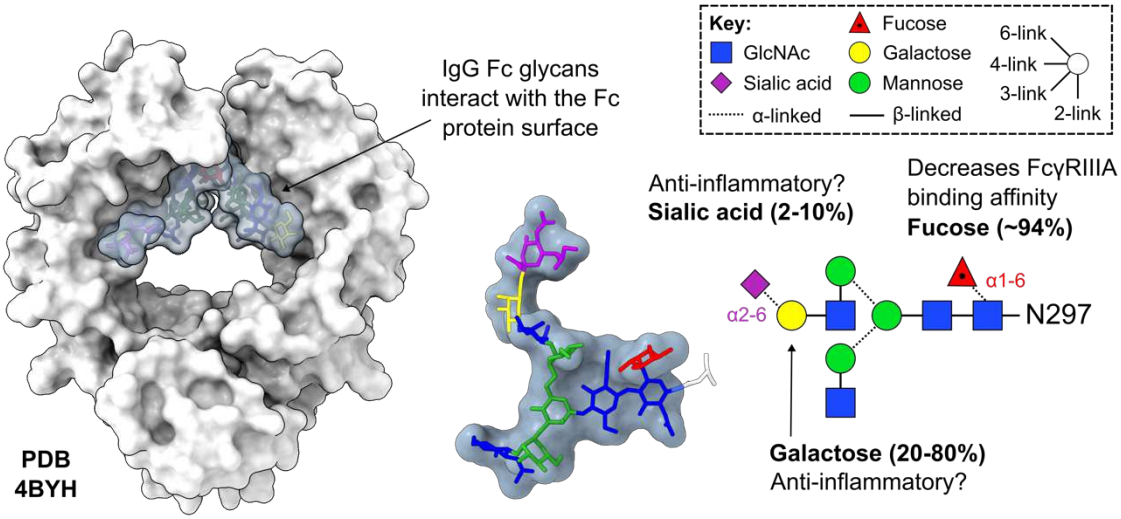


Figure 8: IgG Fc glycosylation at N297. Crystal structure of IgG1 Fc bearing sialylated glycans is shown (PDB 4BYH [108]). Monosaccharides within the crystal structure are coloured to reflect the Symbol Nomenclature For Glycans [109]. Monosaccharide prevalence in serum IgG is indicated [110].

The Fc region of IgG displays around 20 different common glycoforms [111, 112]. Fc glycans are largely biantennary, complex-type, with ~94 % fucosylation, variable galactosylation (20–80 %) and low levels of bisecting GlcNAc (10–15%) and sialic acid (2–10 %) [110] (**Figure 8**). Due to their positions between the Fc Cγ2 domains, these glycans are relatively sterically-restricted and are thus relatively well-resolved in crystal structures [113]. Aromatic residues within the Cγ2 domain interact with monosaccharide residues in order to stabilise the protein [103, 114]. Despite this, Fc glycans still possess conformational flexibility, evidenced by incomplete resolution in crystal structures and nuclear magnetic resonance/molecular dynamics studies [115-119].

The Fc glycans help mediate interactions with FcγRs [74, 120, 121], and their removal impedes Fc-mediated effector functions [103, 105, 106, 122-124]. The structural basis for this has been heavily researched. Within the plethora of IgG Fc crystal structures, the Cγ2 domains are observed in a broad range of conformational states relative to each other, while the Cγ3 homodimer remains unchanged [113]; a molecular dynamics study found that *N*-glycan truncation increases the conformational flexibility of the Cγ2 domains [125], indicating that their presence stabilises the Fc. The glycans interact with amino acids along the Fc protein surface (including L235, F241, F243, K248, V262, V264 and Y300, amongst others [103]), and truncation of monosaccharides from the glycan has been reported to increase disorder within the Fc protein [100, 103]. NMR studies have elucidated that the presence of these glycans stabilises the Fc C′E loop, which contains the glycosylation site at N297,

thereby priming the Fc for FcγIIIA binding [121, 126]. NMR and crystallographic studies have also shown that interactions with the Fc protein surface increase as glycans undergo processing from immature, oligomannose-type to mature, complex-type [127, 128].

Fc glycoengineering

Modulation of the Fc *N*-linked glycans has been utilised to optimise antibody effector function. IgG Fc glycans lacking core fucose have up to ~50-fold greater affinity for FcγRIIIA, which increases ADCC levels [112, 129-132]. This increased affinity is thought to arise from improved interactions with the N162 glycan on FcγRIIIA: Falconer *et al.* proposed that the presence of fucose in Fc glycans limits the conformations which can be samples by the receptor N162 glycan, thereby impeding high affinity binding [133]. Afucosylated antibodies are therefore being used clinically for their improved immune effector function [56].

Galactosylation can further increase FcγRIIIA binding affinity of IgG Fc lacking core fucose [112, 132]. It has also been shown to induce stronger CDC responses in IgG1 and IgG3 antibodies [112, 134], by promoting hexameric Fc formation which drives C1q binding [135, 136]. Sialylation can further promote this increased binding effect [112, 134]; however, sialylation is most commonly associated with anti-inflammatory activity [110, 137-140], although the mechanism for this has not been fully elucidated.

Given the importance of the Fc *N*-linked glycans to antibody immune effector function, deglycosylated antibodies can be used in therapeutic contexts where immune activation is unwanted [141]. Deglycosylation at N297 perturbs the Fc quaternary structure, as indicated by crystallographic analysis [105] and small-angle X-ray scattering [142]. The therapeutic uses of enzymatic deglycosylation of IgG Fc are discussed in section 1.4.

1.2 *Streptococcus pyogenes*

Streptococcus pyogenes, also known as group A *Streptococcus* (GAS), is a Gram-positive, β-hemolytic bacterium, comprised of cocci (round cells) which commonly link in long chains [143]. Its cell structure comprises, from the innermost layer: a cell membrane, a peptidoglycan-containing cell wall and, a capsule of hyaluronic acid (not always present). This bacterium is characterised by possessing the group A carbohydrate as a cell surface antigen, according to the Lancefield classification [144, 145]. GAS is further classified by its surface M protein, encoded by the *emm* gene, which gives rise to over 220 serotypes [146]. The M protein is embedded within the cell wall via its C-terminus; the N-terminal region is hypervariable and thus exhibits the antigenic diversity which dictates serotype classification [146, 147]. Other GAS surface antigens include lipoteichoic acid and protein F, which mediate adhesion to host cells [148, 149].

1.2.1 Disease caused by GAS

GAS causes a wide variety of disease in humans. Infection can be mild; for example, superficial infections such as pharyngitis, scarlet fever and impetigo. In more severe cases, infection can penetrate into deeper tissues. Examples of more invasive infection include streptococcal toxic shock syndrome and necrotising fasciitis. Moreover, GAS infection can lead to immune-mediated sequelae, including acute rheumatic fever, rheumatic heart disease and acute post-streptococcal glomerulonephritis [150]. Recent estimates predicted the occurrence of over half a million deaths from GAS infection each year, the majority of which occur in developing countries [151].

1.2.2 Treatment of GAS

GAS infections are currently treated with antibiotics; however, the bacterium has evolved a multitude of resistance mechanisms against these (reviewed by Brouwer *et al.* [150]). Generally, β-lactam antibiotics such as penicillin are still effective, although subclinical resistance has been reported in some strains [152, 153]. Efforts to manufacture a vaccine for GAS are challenging, in part due to the extensive genetic diversity within the GAS genome and the serotype-specific responses to GAS infection. Other contributing factors include the difficulty in producing effective animal models of a human-adapted pathogen, and concerns about possible autoimmune complications arising from the vaccine [150, 154].

Nevertheless, GAS vaccine research is still ongoing [155]. All current vaccines currently being clinically tested target the M protein, which target either a mixture of the hypervariable N-terminal peptides from clinically-relevant GAS serotypes, or the conserved C-terminal region of the protein (which should confer protection against all GAS serotypes) [150]. StreptAnova is the most clinically-advanced multivalent N-terminal epitope-based vaccine, which was found to be well-tolerated and immunogenic in a Phase I trial [156]. Vaccines targeting M protein C-terminal epitopes are also in development [157-160]. In addition, there are several vaccine candidates targeting non-M protein antigens which have been efficacious in animal models of GAS infection [161-164].

1.2.3 Host immune response against GAS

Upon GAS infection, both innate and adaptive immunity contribute in bacterial clearance. Immune responses against GAS isolates causing superficial versus invasive infection can vary, due to differential expression of various virulence factors which modulate the host immune response [165] (discussed in section 1.3).

Innate immunity

Numerous cell types are involved in innate responses against GAS, which can determine the pathogenicity of the infection. Neutrophils, for example, produce antimicrobial peptides, reactive

oxygen species (ROS) and neutrophil extracellular traps (NETs) for clearance of bacteria [165]. The response of neutrophils is thought to be implicit in initial responses: impaired neutrophil activity has been shown to promote the severity of GAS infection in animal models [166, 167]. Activity of macrophages [168, 169] and dendritic cells (DCs) [170, 171] is also important for controlling initial GAS infection. These cells produce a variety of cytokines and chemokines which mediate downstream immune responses, such as interleukin (IL)-1β and CXCL1, which promote neutrophil recruitment [172, 173]. A recent study by Imai *et al.* demonstrated the role of macrophage inducible C-type lectin (Mincle) in recognising monoglucosyldiacylglycerol, which comprises part of the lipoteichoic acid anchor in GAS, and subsequently inducing secretion of ROS, nitric oxide and proinflammatory cytokines in DCs [174]. Moreover, another subset of myeloid cells termed interferon-γ (IFN-γ)-producing immature myeloid cells (γIMCs), have been isolated from mouse models of severe, invasive GAS infection [175, 176]. Matsumura *et al.* demonstrated that, upon GAS recognition, such γIMCs secrete IL-6, which in turn upregulates Mincle [175].

Adaptive immunity

Accumulation of anti-GAS IgG antibodies are thought to confer increased immunity to GAS over time [177, 178]. Antibodies produced in response to both serotype-specific GAS infection (i.e., targeted against the M protein) and conserved GAS antigens have been reported [179-181]. Subsequent activation of the complement pathway by these antibodies is an important immune mechanism against GAS [4].

In order to prolong infection, GAS has evolved a myriad of mechanisms for targeting various components involved in the host immune response, which will be discussed in the following section.

1.3 Immune evasion mechanisms of GAS

GAS is highly adapted for human infection and thus has developed an extensive range of immune evasion mechanisms, which are summarised below (**Figure 9**).

Inhibition of the complement pathway

The surface M protein (along with M-like proteins, such as protein H) is a major virulence factor in GAS. As well as promoting adhesion to host epithelial cells, these proteins also bind complement factors and immunoglobulins [182, 183]. M proteins also help GAS evade phagocytic killing by the complement system by binding other host proteins and [184, 185] complement inhibitors [186]. For example, binding of fibrinogen has been proposed to prevent recognition of opsonised GAS by steric hindrance [184]. In contrast, recruitment of plasmin(ogen) results in degradation of opsonised complement factor C3b [185]. Protein F1 (also known as SfbI), also promotes GAS adhesion and entry into host epithelial cells, and provides similar immune evasion activity, by preventing C3b deposition and thus preventing complement activation [187]. The hyaluronic acid capsule, although

not present in all GAS strains, also confers resistance to phagocytic killing, likely due to prevention of immune cells from recognising opsonising complement factors [188]. Additional GAS cell surface proteins, such as FbaA and Scl1, aid in immune evasion by binding regulators of complement activation [189, 190].

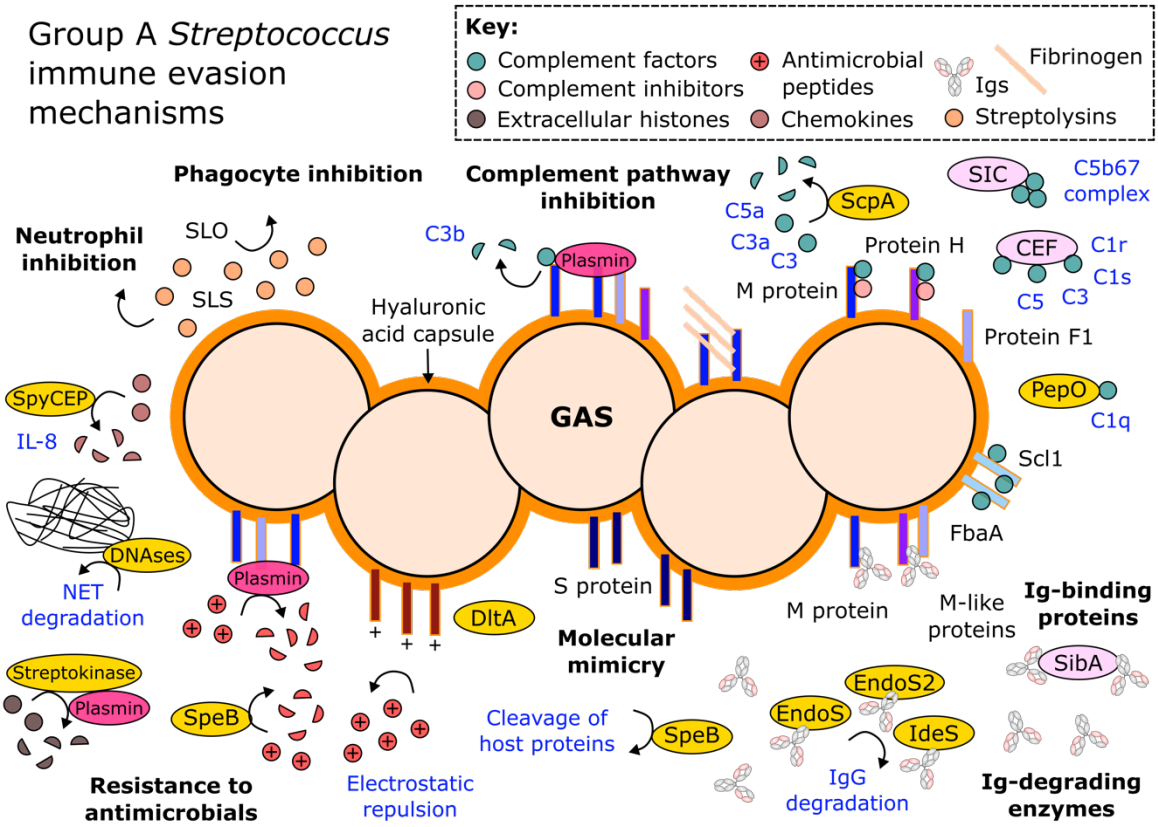


Figure 9: Immune evasion strategies of Group A *Streptococcus*. Schematic representation of strategies utilised by GAS for immune evasion, including: complement pathway inhibition, phatocyte/neutrophil inhibition, resistance to antimicrobials, molecular mimicry, Ig-binding proteins and Ig-degrading enzymes. Partially adapted from Figure 1 in Walker *et al.* [5].

GAS also produces a peptidase (ScpA) which targets the complement factor C5a for cleavage [191], and has recently been reported to cleave C3 and C3a factors [192]. Moreover, streptococcal inhibitor of complement-mediated lysis (SIC) prevents formation of the membrane attack complex by binding the C5b67 complex [193].

Recent studies have identified additional GAS immune evasion factors. A study by Honda-Ogawa *et al.* identified an endopeptidase (PepO) secreted by GAS, which binds the complement activator C1q. Deletion of *pepO* rendered GAS more susceptible to killing by human serum, indicating its involvement in immune evasion [194]. The mechanism of this protein has not been fully elucidated; however, it has been shown to regulate the activity of GAS protease SpeB (discussed below) [195]. Okumura *et al.* also identified a novel GAS protein which binds various complement factors (C1r, C1s, C3 and C5), which was named complement evasion factor (CEF). Again, its role in GAS infection has not been fully determined, although a *Δcef* deletion mutant showed decreased virulence compared to wild-type GAS [196].

Phagocyte inhibition

GAS secretes cytolytic toxins, termed streptolysins, which aid in virulence by causing pore formation in various host cells, including immune cells [197]. Streptolysin O (SLO) has the ability to impair neutrophil function by suppressing ROS production by these cells, along with production of NETs [198]. Along with streptolysin S (SLS), SLO also induces killing of macrophages by inducing programmed cell death [199].

Neutrophil inhibition

GAS secretes a protease called *S. pyogenes* cell envelope proteinase (SpyCEP) which cleaves IL-8, a chemokine which recruits and activates neutrophils [200]. Moreover, secreted DNAses degrade the DNA component of NETs, and thereby release any bacteria trapped within [201, 202]. DNAses also provide the function of autodegrading bacterial DNA, which prevents recognition of this foreign DNA by immune cells [203].

Resistance to antimicrobials

Extracellular histone proteins can function as potent antimicrobial agents [204]. GAS utilises a protein called streptokinase to protect itself from this type of killing: streptokinase binds and activates the host zymogen plasminogen, which mediates cleavage of extracellular histones as the recruited plasminogen acquires plasmin-like (i.e. protease) activity [205]. Plasmin recruitment to the GAS membrane determines the pathogenicity of GAS infection [206], and can assist with cleavage of antimicrobial peptides such as human cathelicidin LL-37 [207]. The broad-spectrum protease SpeB (discussed in further detail below) also degrades this antimicrobial peptide [208].

Antimicrobial peptides are typically positively-charged, and thus are attracted to bacterial negative surface charges. Modification of the cell surface to increase the net positive surface charge therefore helps to repel these peptides. GAS achieves this using a protein called DltA, which catalyses D-alanylation of lipoteichoic acid on the bacterial cell surface [209].

Molecular mimicry

Wierzbicki *et al.* have recently described the activity of an additional cell surface protein, named S protein, which is highly conserved in GAS and helps the bacterium evade immune detection by binding red blood cell membranes [210]; this observation also provides a rationale for the haemolytic properties of this bacterium.

Ig-binding proteins

Several GAS proteins have the ability to bind and sequester immunoglobulins by their Fc region, which prevents their downstream Fc-mediated immune activity. Along with binding complement factors, M and M-like proteins on the cell surface also bind IgG and IgA antibodies [183, 211, 212].

The secreted immunoglobulin binding protein from GAS (SibA) also binds the Fc and Fab regions of IgG, along with IgA and IgM antibodies [213].

Ig-degrading enzymes

In addition to all the immune evasion mechanisms outlined above, GAS utilises several enzymes for the inactivation of immunoglobulins, which directly inhibits the adaptive immune response against the bacterium.

Streptococcal pyrogenic exotoxin B (SpeB) is a broad-spectrum cysteine protease and major virulence factor of GAS [214]. Substrates for SpeB include epithelial junction proteins [215], extracellular matrix proteins [216], autophagy proteins [217], chemokines [218] and the antimicrobial peptide LL-37 [219]. It promotes a pro-inflammatory response by cleaving and activating pro-inflammatory cytokines IL-1β and IL-36γ [220, 221], which has been implicated in severe GAS infection [222].

SpeB also degrades molecules involved in the immune response against GAS, including various complement factors [223-225]. More controversially, it has been reported to cleave IgA, IgD, IgE, IgG and IgM antibodies [226, 227]. A study by Persson *et al.* reported that SpeB only cleaves immunoglobulins under reducing conditions, with no activity under physiological conditions [228]. In contrast, recent work by Blöchl *et al.* reported that SpeB does cleave all IgG subclasses, under both reducing and non-reducing conditions. Cleavage within the IgG hinge region was observed under reducing conditions only; however, the major IgG cleavage site observed under non-reducing conditions was at the junction between the Fc Cγ2 and Cγ3 domains [229]. SpeB may therefore contribute to GAS immune evasion by IgG degradation.

The focus of this thesis is enzymes secreted by GAS which specifically target IgG antibodies, namely IgG-degrading enzyme of *Streptococcus pyogenes* (IdeS) and endoglycosidases S and S2 (EndoS/EndoS2). These enzymes are discussed in detail in the following sections.

1.3.1 IdeS

IdeS is a 35 kDa protease secreted by GAS, which cleaves within the lower hinge region of IgG (specifically, between amino acids G236 and G237) [6]. Cleavage at this site separates the immune-modulating Fc region from the Fabs which recognise antigens, yielding $F(ab')_2$ and Fc fragments [6, 230, 231], thereby rendering the antibody inactive. IdeS activity inhibits the killing of GAS by opsonophagocytosis or neutrophil ROS production [6, 232], and so helps GAS evade the adaptive immune response. As discussed in section 1.1.2, the IgG hinge region is important for conformational flexibility of the molecule, and is also involved in the interaction of IgG with FcγRs and C1q [11, 67, 68]. This region is thus an attractive target for inhibiting antibody activity, and is targeted by

several other proteases, including papain, pepsin, and some matrix metalloproteases [230, 231, 233]. Within the literature, IdeS has also been referred to as Mac1/2 [232, 234] and Sib35 [235].

Different GAS strains have been found to possess one of two alleles of the *ideS* gene [236]; the two resulting variants of IdeS (often distinguished by naming the first and second variants IdeS/Mac-1 and Mac-2, respectively) display less than 50 % sequence identity between residues 111–205 [236], but possess largely indistinguishable endopeptidase activity [237]. An exception to this is Mac-2 from the M28 serotype of GAS, whose impaired enzymatic activity is attributed to a cysteine side chain at position 257, which can form a disulphide bond with the catalytic cysteine at position 94 [237]. Full cleavage of IgG is mediated in two steps in which each Fc chain is cleaved separately, with a much slower rate of proteolysis for the second chain [230, 231, 238, 239]. Deveuve *et al.* reported how cleavage of one antibody chain by IdeS resulted in much weaker binding to FcγRIIIA and C1q, and rendered the antibody unable to induce FcγRIIIA-dependent NK cell functional responses [238]. Similarly-impeded Fc-mediated effector functions have been reported for single-cleaved antibodies by other lower hinge-targeting proteases [231, 238, 240-242].

Enzymatic characterization of IdeS previously reported non Michaelis-Menten kinetics for IgG cleavage [243-245], which provided evidence for IdeS functioning as a dimer, a suggestion corroborated by crystallographic analyses [246]. A more recent study by Vindebro *et al*., however, investigated the kinetics of each chain cleavage separately, reporting K_M values of 7.2 μM and 28 μM, along with k_{cat} values of 10.1 s^{-1} and 0.10 s^{-1} for the first and second IgG cleavage steps, respectively. Both reaction steps were found to follow typical Michaelis-Menten curves, indicating that IdeS is active as a monomer [239]. Dynamic light scattering and analytical size exclusion chromatography (SEC) also provided evidence for IdeS behaving as a monomer [239].

Prior to the start of this project, several crystal structures of IdeS had been determined [246, 247]. The structure of wild-type IdeS from GAS serotype M1 (strain MGAS5005; PDB 2AU1 [246]) is depicted in **Figure 10**. IdeS displays an α/β fold, similar to other proteases within the papain superfamily [246, 247]. Adopting the nomenclature for this papain superfamily, the protein is comprised of two domains: the L-domain, containing mostly α-helices, and the R-domain, comprising a twisted antiparallel β-sheet. The L and R domains mostly comprise the N- and C-terminal halves of the protein sequence, respectively (**Figure 10a**). The catalytic activity of IdeS is attributed to its single cysteine residue (C94) [6, 236] and is therefore classified as a cysteine protease. This residue, along with H262 and D284 (as numbered in Mac-1) form the catalytic triad, which is present at the junction between the L and R domains [247, 248] (**Figure 10a, b**).

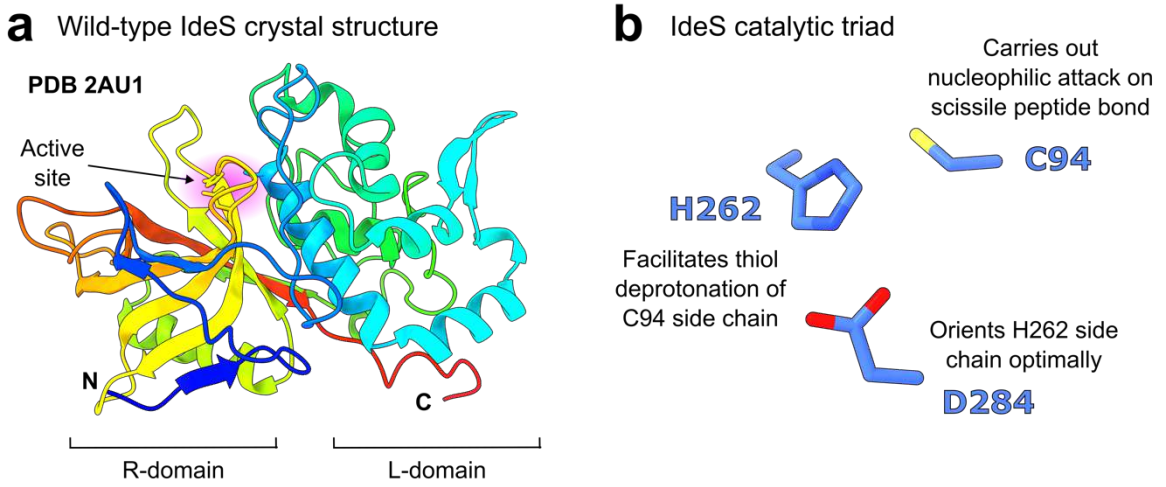


Figure 10: Crystal structure of wild-type IdeS from GAS serotype M1. a Overall structure of IdeS (PDB 2AU1 [246]), coloured as a rainbow from the N-(blue) to the C-terminus (red). **b** Catalytic triad within IdeS, comprising C94, H262 and D284 amino acids. Side chains are coloured by heteroatom (oxygen, red; nitrogen, dark blue; sulphur, yellow).

Agniswamy *et al.* measured the binding affinities of IdeS H262A and D284A variants for IgG1 to be reduced ~20-fold relative to that of the wild-type enzyme (K_D of 2.6 μM), along with undetectable catalytic activity [246]. Amino acid K84 was also proposed to be important for catalysis by comprising the oxyanion hole, a positively-charged pocket within the active site which helps to stabilise a negatively-charged intermediate formed during catalysis. This residue was identified by an overlay of the IdeS active site to papain and cathepsin B (other papain superfamily cysteine proteases) [247]. Unsurprisingly, a K84A substitution displayed no detectable binding or endopeptidase activity for IgG1 [246].

Other amino acids found to be involved in enzyme activity include D286 and Y255. Alanine substitution of each of these two amino acids reduced IdeS binding affinity for IgG1 by ~10-fold, and also resulted in reduced catalytic activity [246, 248]. D286 was proposed to be important for orienting K84, as it was observed interacting with this side chain within an IdeS crystal structure [247]. Similarly, the role of Y255 was attributed to its interaction with D286, stabilising D286 and, in turn, K84 [246].

IdeS is distinct from the family of papain superfamily cysteine proteases due to its strict specificity for IgG (its only known substrate) [6], an observation which is in stark contrast to other bacterial cysteine proteases displaying broad substrate specificities [249-251]. IdeS cleaves all four subclasses of IgG, although IgG2 is cleaved more slowly [6]. This is likely due to IgG2 lacking the double glycine motif within the lower hinge region that is cleaved by IdeS, with an amino acid sequence of P–VAGPS for residues 233–239 (distinct to ELLGGPS for the equivalent section in IgG1; see **Figure 3**) [6]. Vincents *et al*. demonstrated that IdeS can cleave neither proteins with similar amino acid sequences to IgG within the hinge region, nor synthetic peptides designed to mimic the IgG1 hinge [244], indicating that the enzyme recognises IgG using exosite binding.

The protein sequence of IdeS additionally contains an RGD motif (amino acids 214–216), which is typically associated with integrin binding [236]. Indeed, Lei *et al.* reported IdeS binding to human integrins αvβ3 and αIIbβ3 [236]. Interestingly, SpeB also contains an RGD motif [252]. The importance of this motif in IdeS is unknown; however, it is possible that integrin binding may serve the function of IdeS binding to immune cells, in order to intercept Fc-mediated immune activation.

1.3.2 Endoglycosidases

GAS additionally secretes two endoglycosidases which specifically remove the *N*-linked glycans present at N297 on all IgG Fc subclasses: endoglycosidase S (EndoS) [7] and endoglycosidase S2 (EndoS2) [8]. Specifically, these enzymes cleave the β1–4 glycosidic linkage between the two *N*-acetylglucosamine (GlcNAc) residues within the glycan core, releasing the majority of the glycan and leaving a single GlcNAc, variably modified with fucose. As discussed in section 1.1.4, this glycan modification is important for the structural integrity of the Fc [100-105], and mediates binding to FcγRs and complement [11, 103, 106]. Therefore, IgG Fc glycan removal by EndoS/EndoS2 inhibits Fc-mediated effector functions [105, 122, 253]. EndoS was demonstrated to improve survival of GAS in an opsonophagocytic assay, due to reduced FcγR- and complement-mediated immune activation [253]. Moreover, IgG glycan hydrolysis by EndoS has been shown to promote virulence and survival of the bacterium *in vivo* [254], thus demonstrating its role in immune evasion.

EndoS and EndoS2 are encoded by the *ndoS* and *ndoS2* genes, respectively, which are 53 % identical in their nucleotide sequences and only 37 % similar in amino acid sequence [8]. The majority of GAS serotypes secrete EndoS [7], while serotype M49 secretes EndoS2 [8]. Both enzymes cleave Fc *N*-linked glycans from all four IgG Fc subclasses [8, 122]; however, they show distinct Fc glycoform specificities. EndoS is specific towards complex-type, biantennary glycans on IgG Fc [7, 105, 255], whereas EndoS2 exhibits broader glycan specificity, also cleaving other classes of *N*-linked glycans such as oligomannose-type and hybrid-type [8, 256]. Moreover, both enzymes are highly specific for IgG [8, 227], although EndoS2 was reported to also have some activity against α1-acid glycoprotein [8]. The enzymes lose activity against denatured IgG, indicating that specific recognition of the IgG protein surface is required [8, 227].

Prior to the start of this project, crystal structures of EndoS and EndoS2 had been elucidated, in both apo forms and in complex with their isolated glycan substrates [257-259] (**Figure 11**). The enzymes have similar domain architectures, both containing a glycosylhydrolase (GH) domain, a leucine-rich repeat (LRR) domain, a hybrid immunoglobulin (hIg) domain and a carbohydrate-binding module (CBM). EndoS additionally contains N- and C-terminal 3-helix bundles (3HB) (**Figure 11a**). The GH domain contains the glycan-binding groove, and thus contains the active site dyad residues: D233/E235 for EndoS [258, 260], and D184/E186 for EndoS2 [8, 259]. The broader glycoform

specificity exhibited by EndoS2 appears to be possible due to a larger glycan-binding groove within the active site which can accommodate bulkier glycans [259], in comparison to that of EndoS [258].

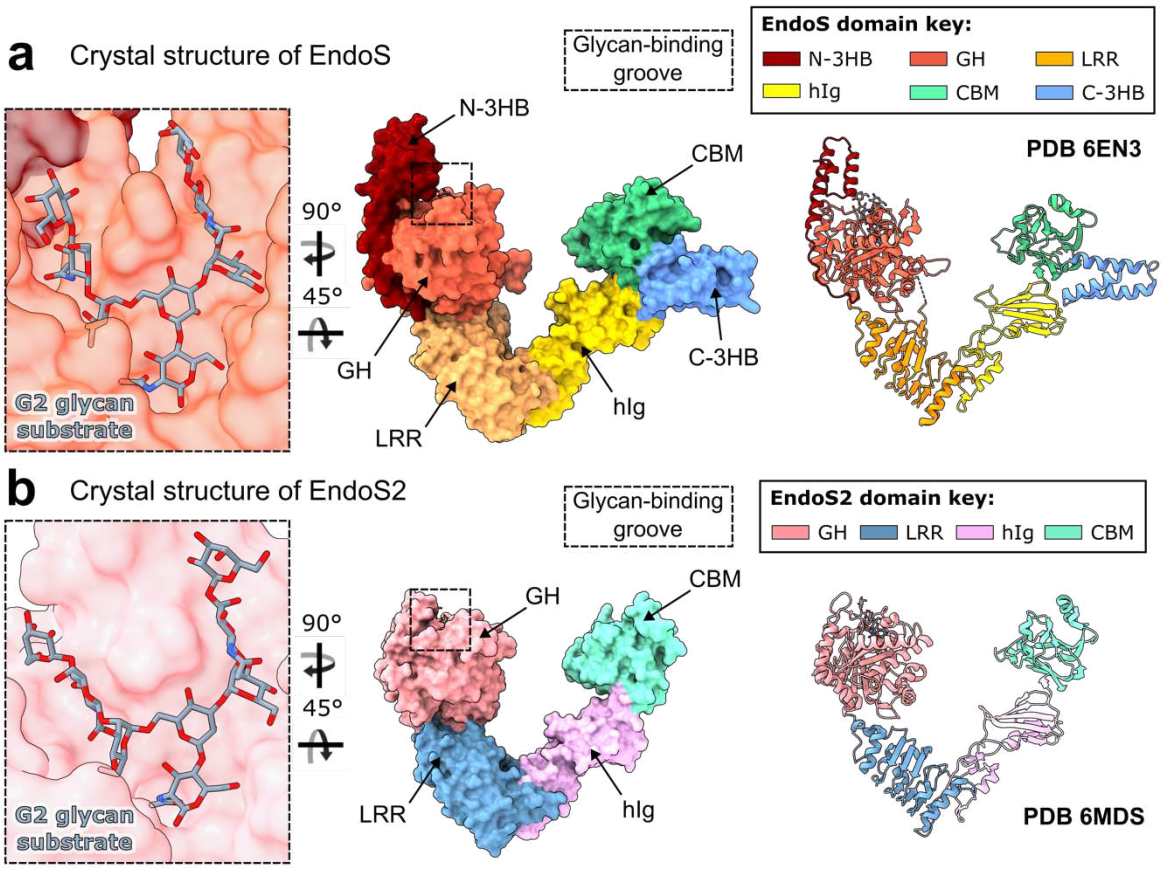


Figure 11: Crystal structures of EndoS and EndoS2 in complex with G2 oligosaccharide substrates. Structures of EndoS (**a**; PDB 6EN3 [258]) and EndoS2 (**b**; PDB 6MDS [259]), from GAS serotypes M1 and M49, respectively. Both enzymes comprise a glycosylhydrolase (GH) domain, leucine-rich repeat (LRR) domain, hybrid immunoglobulin (hIg) domain and a carbohydrate-binding module (CBM). EndoS has additional N- and C-terminal 3-helix bundles (N-3HB and C-3HB, respectively). G2 oligosaccharides are bound within the GH domain of each enzyme, which are depicted as sticks and coloured by heteroatom (oxygen, red; nitrogen, blue).

However, it was unclear from these crystal structures how exactly how EndoS and EndoS2 are able to recognise their IgG substrate with such high specificity. Dixon *et al.* showed that the smallest fragment of IgG that EndoS can hydrolyse is the isolated Cγ2 domain, suggesting that the antibody Fab regions and Cγ3 domains are dispensable for EndoS activity [261]. Within EndoS, the CBM was named according to its homology to carbohydrate-binding modules from *Clostridium thermocellum* proteins (e.g., *Ct*CBM62, which binds galactose-containing polysaccharides). Several studies have shown how removal of the CBM, in both EndoS and EndoS2, severely hinders enzyme activity [257, 259, 261]. It was thus postulated that the CBM is somehow involved in Fc glycan recognition, although the molecular basis behind this had not been defined.

Moreover, despite the EndoS and EndoS2 crystal structures showing a similar overall domain architecture [257, 259], domain-swap experiments showed that substitution of the GH domain from EndoS2 onto an EndoS scaffold produces a non-functional enzyme. However, subsequent

substitution of the EndoS2 CBM onto this scaffold results in a chimeric enzyme with EndoS2-like activity [259]. These experiments indicate that the GH domain and CBM from EndoS2 have co-evolved to produce optimum enzymatic activity, although the structural basis for this is unknown.

1.4 Applications of IgG-degrading enzymes

It is clear that GAS utilises IdeS and EndoS/EndoS2 enzymes as part of its sophisticated immune evasion strategy [6-8]. However, these enzymes are unique in that they specifically target IgG [8, 227, 244]. There is thus significant interest in repurposing these enzymes as biologics for use in a wide range of therapeutic and biotechnological applications, which are summarised below.

1.4.1 Organ transplantation

Organ transplantation is currently the best treatment for patients suffering from end-stage renal disease. The availability of transplants is limited, however, by the need to match a patient with a suitable donor. Human leukocyte antigens (HLAs), which are present on almost all nucleated cells, identify cells as "self" and are thus not targeted by the immune system [1]. A patient receiving a transplant must therefore be screened for anti-HLA antibodies which would recognise HLA molecules on a donated organ as "foreign", ultimately leading to organ rejection [262]. Patient sensitisation is a significant barrier to successful organ transplantation: in 2017, ~30 % of the patients on the waiting list for a kidney transplant in the US were sensitised to HLA, with ~50 % of those considered highly sensitised (calculated panel-reactive antibody (cPRA) of at least 80 %) [263].

IdeS has been employed as a means of desensitisation in such patients, due to its ability to effectively cleave any donor-specific antibodies circulating within a patient prior to transplantation. Under the name Imlifidase (brand name Idefirix®), it has recently been clinically-approved for use as a pre-treatment for kidney transplantation in hypersensitised patients [262, 264, 265]. An international, phase II clinical trial showed how Imlifidase treatment resulted in an ~88.9 % graft survival rate (at 6 months post-transplantation) in highly sensitised patients (median of 99.83 % cPRA) [264], and is therefore highly effective as a pre-treatment.

Moreover, EndoS has been tested as a desensitization treatment in a mouse model of bone marrow transplantation. In combination with Imlifidase, it was shown to be successful at inactivating donor-specific antibodies and improved the survival of transplanted bone marrow cells [266].

1.4.2 Autoimmune disease

IgG-degrading activity has clear applications to autoimmune disease, in which pathogenic auto-antibodies cause aberrant activation of the immune system. IdeS has thus been tested in a variety of pre-clinical models of autoimmune disease, such as small vessel vasculitis [267], arthritis [268],

thrombocytopenic purpura [269], antibody-mediated skin injury [270], anti-glomerular basement membrane disease [271, 272], heparin-induced thrombocytopenia [273], and Guillain-Barré syndrome [274]. IdeS has also been tested in a phase 2a clinical trial for anti-glomerular basement membrane disease, after which 67% of patients were no longer dependent on dialysis (the current best treatment) [275], although these results have yet to be confirmed with a randomised trial.

A recent paper by Lynch *et al.* also describes the anchoring of IdeS to FcγRIIA for treatment of immune thrombocytopenia (ITP: a bleeding disorder characterised by aberrant platelet activation, which is caused by IgG-mediated FcγRIIA activation), which resulted in reduced thrombocytopenia in an ITP mouse model [276]. Being a bacterial enzyme, IdeS itself is immunogenic and can be inhibited by anti-IdeS antibodies generated during infection [277]. Therefore, a more targeted approach for IdeS administration, as described in Lynch *et al.* [276], could be useful in reducing the required dose and minimising harmful immune responses against IdeS.

EndoS has also been successfully utilised as a treatment in several pre-clinical models of autoimmune disease, including anti-glomerular basement membrane disease [271], systemic lupus erythematosus [278, 279], arthritis [279-281], small vessel vasculitis [267, 282], thrombocytopenic purpura [283], antibody-mediated skin injury [270], multiple sclerosis [284] and autoimmunity against type VII collagen [285].

1.4.3 *In vivo* gene therapy

Gene therapy is being utilised for the treatment of several diseases, such as hemophilia B [286], which involves the administration of an adeno-associated virus (AAV) vector for the delivery of a missing/defective gene. However, such treatment can be inhibited by the presence of neutralising antibodies against the viral capsid [287]. A recent study by Leborgne *et al.* demonstrated how IdeS administration can reduce anti-AAV antibodies in human plasma samples *in vitro*, as well as murine and non-human primate models *in vivo* [288], indicating its potential use for *in vivo* gene therapy.

1.4.4 Inactivation of competing serum IgG

IgG circulates in human serum at concentrations significantly greater than the K_D of IgG Fc:FcγR binding interactions, which results in the majority of FcγRs being bound with IgG under physiological conditions (reviewed in [62]). Therefore, an administered therapeutic antibody must compete with serum IgG for binding FcγRs, and the presence of serum IgG weakens the Fc-mediated effector functions of a therapeutic antibody. A study by Kelton *et al.* in 1985 demonstrated this using radiolabelled red blood cells; clearance rates of these cells were measured to be quicker when lower levels of serum IgG was present. They also showed how the abnormally fast RBC clearance in a patient with hypogammaglobulaemia could be returned to "normal" levels by gamma globulin treatment [289] (**Figure 12a**). Subsequent work by Preithner *et al.* showed how the inhibitory effect

of serum IgG is dependent upon the density of target antigen: in instances of high target density, the inhibitory effect of serum IgG can be overcome using higher concentrations of therapeutic antibody. In contrast, when target density is low, the presence of serum IgG limits the maximum response elicited, even at high therapeutic antibody concentrations [290] (**Figure 12b)**.

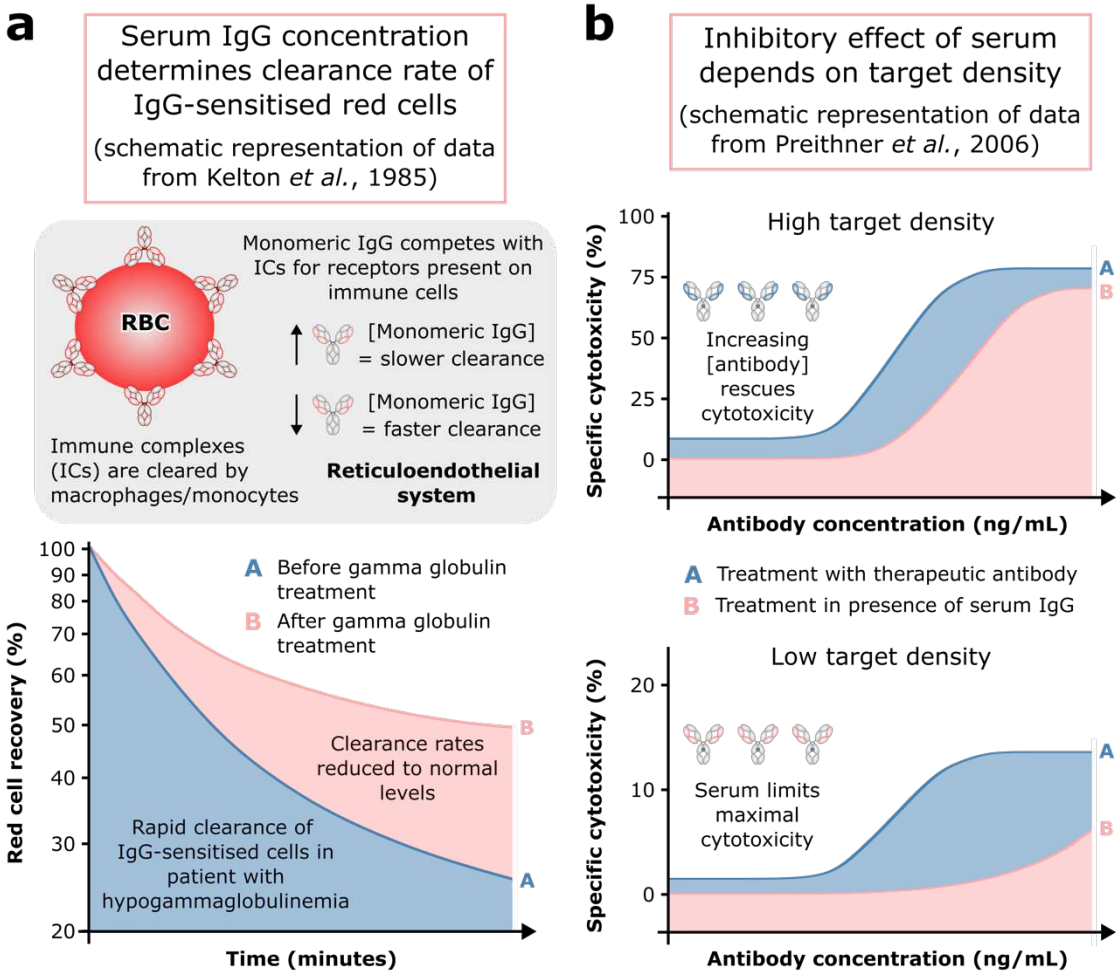


Figure 12: Effect of serum IgG concentration on antibody function. a Schematic representation of data from Kelton *et al.* (1985) [289], showing effect of serum IgG concentration on clearance rate of IgG-sensitised red blood cells (RBCs). Unusually fast clearance rate in a patient with hypogammaglobulaemia was reduced to a normal level following gamma globulin treatment. **b** Schematic representation of data from Preithner *et al.* (2006) [290], showing how the inhibitory effect of serum IgG depends on target antigen density. The inhibitory effect of competing serum IgG can be overcome using higher concentrations of therapeutic antibody when target density is high; however, the maximum antibody response is limited where target density is low, and cannot be overcome with increasing therapeutic antibody concentration. Figure taken from Bauer-Smith *et al.* [62].

The strategy of serum IgG removal is therefore being investigated for the potentiation of therapeutic antibodies. IdeS can cleave IgG present on B cells, thereby preventing downstream responses to B cell antigen recognition [291], and directly prevents FcγR activation by IgG [292]. Järnum *et al.* showed how IdeS treatment can potentiate the efficacy of rituximab in a tumour mouse model. An *ex vivo* experiment was also carried out, in which sera from two human participants (from a phase II clinical trial of IdeS utilisation as a pre-treatment for kidney transplantation [293]) before and after

IdeS treatment was tested for rituximab blocking ability. Sera from both patients prior to IdeS treatment significantly inhibited rituximab-mediated ADCP, while sera collected 24 hours after IdeS treatment had no inhibitory effect [292]. These results indicate that, in humans, a single dose of IdeS can reduce serum IgG levels enough to prevent its inhibitory effect on therapeutic antibody efficacy.

Similarly, EndoS has been researched for its ability to block the inhibitory effect of serum IgG. Baruah *et al.* demonstrated that, *in vitro*, serum IgG significantly reduced the ability of a monoclonal antibody to bind to FcγRIIIA [105]. This antibody was designed to be resistant to EndoS cleavage, by engineering the Fc glycosylation to be oligomannose-type, a glycoform not cleaved by EndoS [256]. Addition of EndoS significantly increased binding of the antibody to FcγRIIIA, demonstrating its potential use in "receptor refocussing", as termed by the authors [105].

1.4.5 Engineering IgG Fc glycosylation

As discussed in section 1.1.4, IgG Fc glycoforms influence antibody effector functions by mediating binding to FcγRs and complement [103, 105, 106, 122-124]. The precise control of antibody glycosylation has thus been applied in several clinically-used antibodies for improved immune effector function [56, 294-297]. The ability of EndoS and EndoS2 to specifically remove IgG Fc *N*-linked glycans has led to biotechnological use of these enzymes as tools for antibody glycoengineering [107, 298]. In brief, wild-type EndoS and EndoS2 can be used to cleave off existing Fc *N*-linked glycans, and transglycosylation reactions can subsequently be performed on this antibody scaffold to generate desired glycoforms [299-303] (**Figure 13**). Crystal structures of EndoS/EndoS2 in complex with their isolated glycan substrates revealed the carbohydrate groups bound within a groove in the GH domain, bringing them into close proximity with the catalytic dyad (D233/E235 in EndoS; D184/E186 in EndoS2; **Figure 13a**).

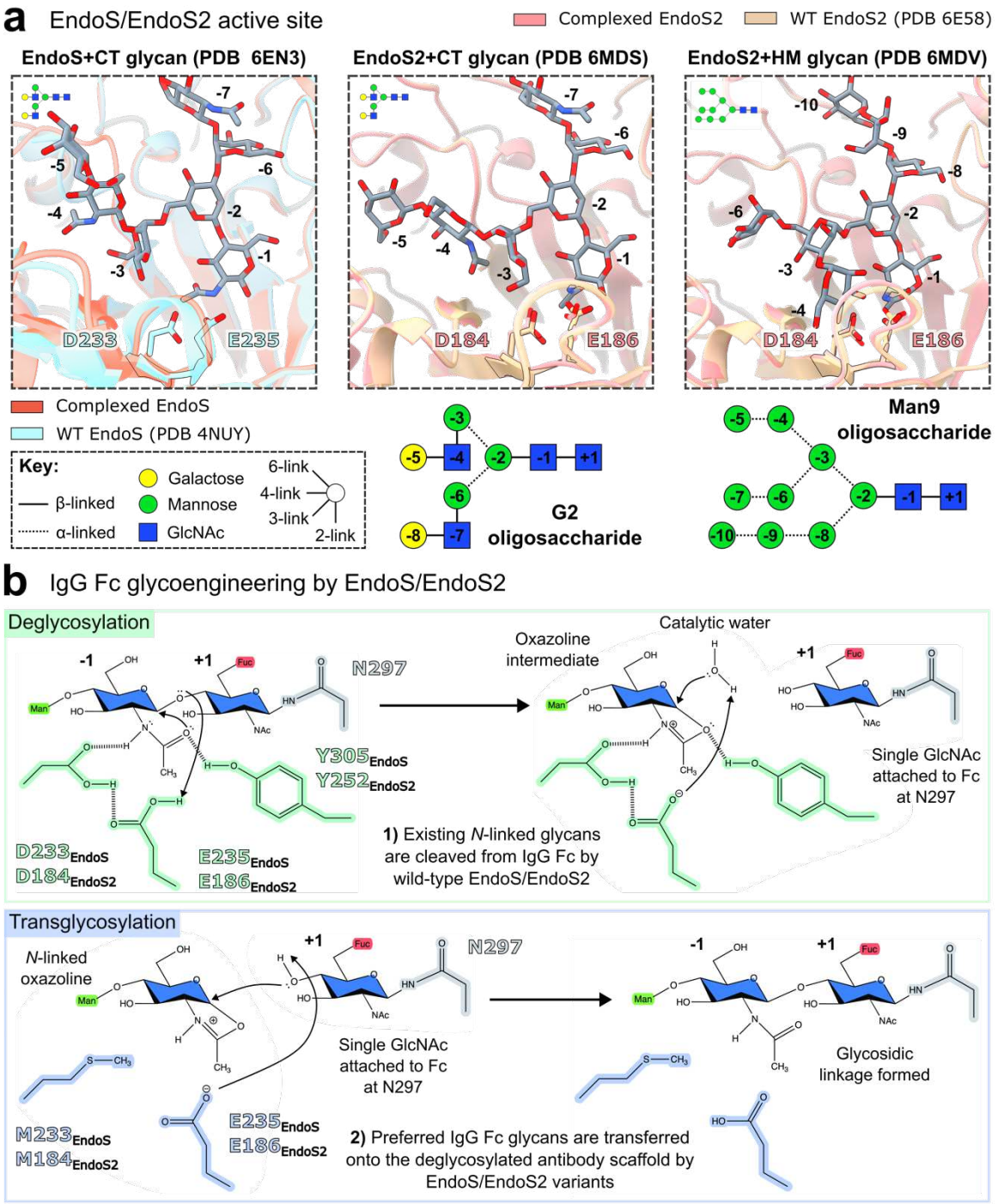


Figure 13: IgG Fc *N*-linked glycan remodelling by EndoS/EndoS2. a Crystal structures of EndoS (PDB 6EN3 [258]) and EndoS2 (complex-type glycan, PDB 6MDS; high-mannose glycan PDB 6MDV [259]) in complex with their isolated glycan substrates. Glycans are bound within the glycosidase domain of the endoglycosidase, in the vicinity of the catalytic dyad (D233/E235 for EndoS; D184/E186 for EndoS2). **b** IgG Fc glycoengineering by EndoS/EndoS2. Existing glycans are cleaved from IgG Fc using a wild-type endoglycosidase. Preferred Fc glycoforms are generated in a transglycosylation reaction, in which an *N*-linked oxazoline is used as a substrate and forms a glycosidic bond with the GlcNAc (+1) on the deglycosylated antibody scaffold [304].

EndoS and EndoS2 cleave between the two core GlcNAcs of IgG Fc glycans (marked as +1 and –1 in **Figure 13**). Cleavage occurs in a hydrolysis reaction, in which the catalytic glutamate (E235 in EndoS; E186 in EndoS2) transfers a proton to the glycosidic oxygen, cleaving the glycosidic bond. The carbonyl oxygen within the GlcNAc *N*-acetyl group attacks the anomeric carbon to generate an

oxazoline intermediate, which is oriented by interaction with D233/D184. This intermediate breaks down by nucleophilic attack of the glutamate side chain of an incoming water molecule, which subsequently attacks the anomeric carbon within the intermediate to regenerate the hemiacetal GlcNAc (-1) group (**Figure 13b**) [304]. Aside from the catalytic dyad, additional residues are thought to aid in catalysis, such as a conserved tyrosine (Y305/255 in EndoS/EndoS2) which helps to optimally position the *N*-acetyl group in GlcNAc (-1) [304, 305].

Replacement of the catalytic water with a sugar molecule (such as an *N*-linked oxazoline, whose structure mimics the intermediate generated in hydrolysis) allows a glycosylation reaction to occur, in which the GlcNAc (+1), activated by the catalytic glutamate, attacks an incoming donor carbohydrate to form a glycosidic bond (**Figure 13b**) [304]. Both wild-type EndoS [255, 306] and EndoS2 [301] have been shown to possess some transglycosylation activity, although enzyme variants have been developed which optimise this process. Huang *et al.* reported two EndoS variants, D233A and D233Q, which are more efficient glycosynthases than wild-type EndoS and generated precise, homogenous glycoforms of rituximab (such as the fully-sialylated S2G2F) for improved effector functions [299]. These EndoS variants can also be used to produce azido-tagged IgG glycoforms, which have application as antibody-drug conjugates [299, 307, 308].

Within EndoS2, the equivalent residue to D233 is D184; Li *et al.* demonstrated how a D184M exchange in EndoS2 confers potent transglycosylation activity (with yields exceeding 90%) of oligomannose, hybrid-type and complex-type glycans to deglycosylated IgG [301]. Subsequently, Tong *et al.* showed that the equivalent EndoS variant, D233M, showed an improved transglycosylation/hydrolysis ratio compared to the previously-described D233A and D233Q variants. The greater transglycosylation efficiency of EndoS^{D233M} was found to be due to a ~4-fold increase in turnover rate (measured by k_{cat}), and improved binding affinity for deglycosylated antibody (K_M of ~30 μM, compared to >200 μM for EndoS^{D233A}); similar observations were made in a comparison of EndoS2^{D184A} and EndoS2^{D184M} variants [300]. An EndoS2 T138Q variant has also been described, which displays potent transglycosylation activity [302]. Moreover, wild-type EndoS2 has been utilised in “one-step” reactions for synthesis of antibody-drug conjugates [309, 310].

1.5 Aims

Despite the extensive therapeutic and biotechnological applications of the GAS immune evasion enzymes IdeS and EndoS/EndoS2, it was largely unknown how these enzymes are able to specifically target and deactivate IgG antibodies. Therefore, the aim of this project was to elucidate the structural basis behind the unique specificities of these enzymes for IgG, by generating X-ray crystal structures of the inactive enzymes in complex with the IgG Fc region. Previous attempts to yield these co-

crystals are thought to have been limited by the strong ability of IgG Fc to crystallise on its own, thereby inhibiting co-crystal growth.

An initial aim was to elucidate the structural basis behind the strong crystallisation ability of IgG Fc, in order to understand how this may be overcome for co-crystallisation experiments. Chapter 3 describes analysis of "typical" IgG Fc crystallisation, and the subsequent identification of Fc variants which display reduced self-crystallisation, by exchanging an amino acid implicit in the Fc crystal packing interactions. These Fc variants were used for subsequent co-crystallisation experiments with the aforementioned GAS immune evasion enzymes.

Within chapters 4 and 5, crystal structures for IdeS, EndoS and EndoS2 in complex with their IgG1 Fc substrate are described. For each of the enzyme-Fc complexes, the structural basis behind enzyme activity and specificity towards IgG are discussed. In addition, chapter 5 describes the differences in IgG recognition between related enzymes EndoS and EndoS2. Final conclusions and perspectives are outlined in chapter 6.

Chapter 2 Materials and Methods

2.1 Cloning and mutagenesis

All synthesised constructs were optimised for their respective expression systems using the codon optimization tool from GENWIZ (https://clims4.genewiz.com/Toolbox/CodonOptimization).

2.1.1 Immunoglobulins

2.1.1.1 IgG1 Fcs

Immunoglobulins were sourced as encoded within a pFUSE-hIgG1-Fc vector (plasmid purchased from InvivoGen), which results in expression of residues 221–447 in IgG1. Wild-type IgG1 Fc was expressed exactly as encoded within this plasmid; E382R, E382S and E382A exchange variants (collectively called "E382X") were introduced to this construct by site-directed mutagenesis, using a QuikChange II kit (Agilent). The method was used as described in the manufacturer's protocol. Primers for site-directed mutagenesis of E382X variants were synthesised by Eurofins Genomics and are shown in **Table 2**.

Table 2: Primers used for introduction of E382R, E382S and E382A exchanges into IgG1 Fc.

Construct	Mutagenic primer sequence (written 5′→3′)	
	Forward primer	**Reverse primer**
IgG1 Fc^{E382R}	GACATCGCCGTGGAGTGGAGTGGAG GAGCAATGGGCAGCCGGAGAACAAC	TTGTTCTCCGGCTGCCCATTGCTCCT CCACTCCACGGCGATGTCG
IgG1 Fc^{E382S}	GACATCGCCGTGGAGTGGAGTGGAG CAGCAATGGGCAGCCGGAGAACAAC	TTGTTCTCCGGCTGCCCATTGCTGCTC CACTCCACGGCGATGTCG
IgG1 Fc^{E382A}	GACATCGCCGTGGAGTGGAGTGGGC GAGCAATGGGCAGCCGGAGAACAAC	TTGTTCTCCGGCTGCCCATTGCTCGCC CACTCCACGGCGATGTCG

Sequences of the wild-type IgG1 Fc construct and the E382X Fc variants are displayed as a multiple sequence alignment in **Figure 14**.

Figure 14: IgG1 Fc constructs expressed from pFUSE-hIgG1-Fc vectors. pFUSE-hIgG1-Fc vector expresses IgG1 Fc residues 221–447. E382R/S/A exchanges introduced with site-directed mutagenesis are indicated with a red asterisk. Sequences were aligned with Clustal Omega [35] and the resulting alignment depicted using ESPript 3.0 [311].

Genes for IgG1 Fc E382X variants containing an additional L234C exchange (to introduce an additional disulphide bridge within the lower hinge) were cloned into a pFUSE-hIgG1-Fc vector by NBS Biologicals.

2.1.1.2 "Asymmetric" IgG1 Fc^{E382R}

A construct for a "short" IgG1 Fc γ-heavy chain was designed to mimic one that had been cleaved by IdeS (therefore, starting from G237). This chain would be combined with an equivalent, "long" IgG1 Fc γ-chain (starting from D221, as encoded within the pFUSE-hIgG1-Fc vector), thereby mimicking an Fc that has been cleaved by IdeS on one chain only. Knob-into-hole amino acid exchanges were included to encourage the formation of heterodimer Fcs (i.e., not homodimers of two "short" chains or two "long" chains, as described by Merchant *et al.* [312]). The "knob" exchanges (S354C, T366W and K409A) were included within the "long" Fc chain, while the "hole" exchanges (Y349C, T366S, L368A, F405K and Y407V) were included within the "short" chain. An E382R exchange was included with both "long" and "short" chains, in order to discourage preferential Fc self-crystallisation. The two constructs designed for expression of this "asymmetric" IgG1 Fc variant are depicted in **Figure 15**.

a "Long" IgG1 Fc^{E382R} chain (residues 221-447)

PAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN
AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ
PREPQVYTLPPCREEMTKNQVSLWCLVKGFYPSDIAVEWRSNGQPENNYKTTPPVL
DSDGSFFLYSALTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

b "Short" IgG1 Fc^{E382R} chain (residues 237-447)

GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR
EEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQV
CTLPPSREEMTKNQVSLSCAVKGFYPSDIAVEWRSNGQPENNYKTTPPVLDSDGSF
KLVSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Figure 15: Protein sequences of "long" and "short" IgG1 Fc constructs used for "asymmetric" Fc formation. a "Long" IgG1 Fc containing "knob" amino acid exchanges (S354C, T366W and K409A; highlighted in purple). **b** "Short" IgG1 Fc containing "hole" amino acid exchanges (Y349C, T366S, L368A, F405K and Y407V; highlighted in pink). **a, b** E382R exchanges are included to discourage Fc self-crystallisation (highlighted in orange).

2.1.1.3 IgG2/3/4 Fc^{E382R}

Protein constructs for IgG2, IgG3 and IgG4 Fc (GenBank accession numbers AAB59393.1, CAA27268.1 and AAB59394.1, respectively) were designed as shown in **Figure 16**. The equivalent glutamate to E382 in IgG1 Fc was exchanged to an arginine, in order to discourage preferential Fc self-crystallisation.

```
IgG4      KYGPPCPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVD 60
IgG2      KCCVECPPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVD 59
IgG3      DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 60
          .    ** ****  :.********************************:*******:****

IgG4      GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK 120
IgG2      GVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK 119
IgG3      GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 120
          ***************:***:********:*****************.**: *******:*

IgG4      GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWRSNGQPENNYKTTPPVLDS 180
IgG2      GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDISVEWRSNGQPENNYKTTPPMLDS 179
IgG3      GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWRSSGQPENNYNTTPPMLDS 180
          **************:**********************:*****.*******:****:***

IgG4      DGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLGK 227
IgG2      DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 226
IgG3      DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK 227
          ********:*********:**:************::******** **
```

Figure 16: Multiple sequence alignment of IgG2/3/4 Fc^{E382R} constructs. The majority of sequence diversity occurs within the lower hinge region (indicated with the dashed box). E382R exchanges were included within the constructs in order to discourage preferential Fc self-crystallisation (highlighted in red). Sequence alignment was generated using Clustal Omega [35].

2.1.2 IdeS and related constructs

All constructs were designed and cloned into a pET21a(+) vector by NBS Biologicals. Constructs were designed to contain a C-terminal linker and His-tag (sequence LEHHHHHH) to enable protein purification by nickel affinity chromatography.

2.1.2.1 Truncated IdeS^{C94A}

A gene construct encoding *ideS* from serotype M59 of *S. pyogenes* (GenBank accession number AFC66043.1) was designed. The construct was optimised for crystallisation by encoding only amino acids 41–339 of IdeS (based off the resolved amino acids within elucidated crystal structures [246, 247]). A C94A exchange was included to abolish catalytic activity, as previously elucidated [247]. The resulting protein construct is shown in **Figure 17**.

Truncated IdeS^{C94A}

MSEVTPYHVTSVWTKGVTPPAKFTQGEDVFHAPYVANQGWYDITKTFNGKDDLLA
GAATAGNMLHWWFDQNNEKIEAYLKKHPDKQKIMFGDQELLDVRKVINTKGDQT
NSELFNYFRDKAFPGLSARRIGVMPDLVLDMFINGYYLNVYKTQTTDVNRTYQEKDR
RGGIFDAVFTRGDQSKLLTSRHDFKEKTLKEISDLIKKELTEGKALGLSHTYANVRIN
HVINLWGADFDSNGNLEAIYVTDSDSNASIGMKKYFVGVNSAGKVAISAKEIKEDN
IGAQVLGLFTLSTGQDSWNQLEHHHHHH

Figure 17: Protein sequence of truncated, inactive IdeS variant utilised for co-crystallisation experiments. Truncated IdeS (amino acids 41–339) with a C94A exchange to abolish catalytic activity (amino acid A94 is highlighted in orange). C-terminal linker and His-tag are highlighted in blue.

2.1.2.2 Truncated Imlifidase^{C94A}

A gene for a truncated form of Imlifidase (isolated from *S. pyogenes* serotype M1; GenBank accession number AAK33786.1) was designed to encode amino acids 41–337. A C94A amino acid exchange was included to abolish catalytic activity. The protein sequence of this Imlifidase construct is shown in **Figure 18**.

Truncated Imlifidase^{C94A}

MSEVTPYHVTSVWTKGVTPPANFTQGEDVFHAPYVANQGWYDITKTFNGKDDLLA
GAATAGNMLHWWFDQNKDQIKRYLEEHPEKQKINFNGEQMFDVKEAIDTKNHQL
DSKLFEYFKEKAFPYLSTKHLGVFPDHVIDMFINGYRLSLTNHGPTPVKEGSKDPRG
GIFDAVFTRGDQSKLLTSRHDFKEKNLKEISDLIKKELTEGKALGLSHTYANVRINHV
INLWGADFDSNGNLKAIYVTDSDSNASIGMKKYFVGVNSAGKVAISAKEIKEDNIG
AQVLGLFTLSTGQDSWNQLEHHHHHH

Figure 18: Protein sequence of truncated, inactive Imlifidase variant utilised for co-crystallisation experiments. Truncated Imlifidase (amino acids 41–338) with a C94A exchange to abolish catalytic activity. Amino acid A94 is highlighted in orange; C-terminal linker and His-tag are highlighted in blue.

2.1.3 Endoglycosidases

All constructs were designed and cloned into a pET21a(+) vector by NBS Biologicals. Constructs were designed to contain a C-terminal linker and His-tag (sequence LEHHHHHH).

2.1.3.1 Truncated EndoS$^{D233A/E235L}$

A gene construct of *ndoS* from serotype M1 of *S. pyogenes* (GenBank accession number AP012491.2) was designed to encode amino acids 98–995 of EndoS. D233A and E235L exchanges were included to abolish catalytic activity, as described previously [258]. The protein construct is shown in **Figure 19**.

Truncated EndoS$^{D233A/E235L}$

MIPEKIPMKPLHGPLYGGYFRTWHDKTSDPTEKDKVNSMGELPKEVDLAFIFHDWT
KDYSLFWKELATKHVPKLNKQGTRVIRTIPWRFLAGGDNSGIAEDTSKYPNTPEGN
KALAKAIVDEYVYKYNLDGLDVAVLHDSIPKVDKKEDTAGVERSIQVFEEIGKLIGP
KGVDKSRLFIMDSTYMADKNPLIERGAPYINLLLVQVYGSQGEKGGWEPVSNRPEK
TMEERWQGYSKYIRPEQYMIGFSFYEENAQEGNLWYDINSRKDEDKANGINTDITG
TRAERYARWQPKTGGVKGGIFSYAIDRDGVAHQPKKYAKQKEFKDATDNIFHSDYS
VSKALKTVMLKDKSYDLIDEKDFPDKALREAVMAQVGTRKGDLERFNGTLRLDNPA
IQSLEGLNKFKKLAQLDLIGLSRITKLDRSVLPANMKPGKDTLETVLETYKKDNKEEP
ATIPPVSLKVSGLTGLKELDLSGFDRETLAGLDAATLTSLEKVDISGNKLDLAPGTEN
RQIFDTMLSTISNHVGSNEQTVKFDKQKPTGHYPDTYGKTSLRLPVANEKVDLQSQ
LLFGTVTNQGTLINSEADYKAYQNHKIAGRSFVDSNYHYNNFKVSYENYTVKVTDS
TLGTTTDKTLATDKEETYKVDFFSPADKTKAVHTAKVIVGDEKTMMVNLAEGATVIG
GSADPVNARKVFDGQLGSETDNISLGWDSKQSIIFKLKEDGLIKHWRFFNDSARN
PETTNKPIQEASLQIFNIKDYNLDNLLENPNKFDDEKYWITVDTYSAQGERATAFSN
TLNNITSKYWRVVFDTKGDRYSSPVVPELQILGYPLPNADTIMKTVTTAKELSQQKD
KFSQKMLDELKIKEMALETSLNSKIFDVTAINANAGVLKDCIEKRQLLKKLLEHHHH
HH

Figure 19: Protein sequence of truncated, inactive EndoS variant used for crystallography. Construct includes amino acids 98–995, and contains D233A/E235L exchanges to abolish catalytic activity (highlighted in orange). C-terminal linker and His-tag are highlighted in blue.

2.1.3.2 Full-length EndoS$^{D233A/E235L}$

The *ndoS* gene from serotype M1 of *S. pyogenes* (GenBank accession number AP012491.2) was also utilised to design a construct encoding full-length EndoS (amino acids 37–995, as described previously [258]). D233A/E235L exchanges were included to render the enzyme inactive. The protein sequence of this construct is included within **Figure 20**.

Full length EndoS$^{D233A/E235L}$

MEEKTVQVQKGLPSIDSLHYLSENSKKEFKEELSKAGQESQKVKEILAKAQQADKQ
AQELAKMKIPEKIPMKPLHGPLYGGYFRTWHDKTSDPTEKDKVNSMGELPKEVDLA
FIFHDWTKDYSLFWKELATKHVPKLNKQGTRVIRTIPWRFLAGGDNSGIAEDTSKY
PNTPEGNKALAKAIVDEYVYKYNLDGLDVAVLHDSIPKVDKKEDTAGVERSIQVFEE
IGKLIGPKGVDKSRLFIMDSTYMADKNPLIERGAPYINLLLVQVYGSQGEKGGWEPV
SNRPEKTMEERWQGYSKYIRPEQYMIGFSFYEENAQEGNLWYDINSRKDEDKANG
INTDITGTRAERYARWQPKTGGVKGGIFSYAIDRDGVAHQPKKYAKQKEFKDATDN
IFHSDYSVSKALKTVMLKDKSYDLIDEKDFPDKALREAVMAQVGTRKGDLERFNGT
LRLDNPAIQSLEGLNKFKKLAQLDLIGLSRITKLDRSVLPANMKPGKDTLETVLETYK
KDNKEEPATIPPVSLKVSGLTGLKELDLSGFDRETLAGLDAATLTSLEKVDISGNKLD
LAPGTENRQIFDTMLSTISNHVGSNEQTVKFDKQKPTGHYPDTYGKTSLRLPVANE
KVDLQSQLLFGTVTNQGTLINSEADYKAYQNHKIAGRSFVDSNYHYNNFKVSYENY
TVKVTDSTLGTTTDKTLATDKEETYKVDFFSPADKTKAVHTAKVIVGDEKTMMVNLA
EGATVIGGSADPVNARKVFDGQLGSETDNISLGWDSKQSIIFKLKEDGLIKHWRFF
NDSARNPETTNKPIQEASLQIFNIKDYNLDNLLENPNKFDDEKYWITVDTYSAQGER
ATAFSNTLNNITSKYWRVVFDTKGDRYSSPVVPELQILGYPLPNADTIMKTVTTAKEL
SQQKDKFSQKMLDELKIKEMALETSLNSKIFDVTAINANAGVLKDCIEKRQLLKKLL
EHHHHHH

Figure 20: Protein sequence of full-length, inactive EndoS variant used for crystallography. Construct includes amino acids 37–995, and contains D233A/E235L exchanges to abolish catalytic activity (highlighted in orange). C-terminal linker and His-tag are highlighted in blue.

2.1.3.3 EndoS2$^{D184A/E186L}$

A gene for *ndoS2* from serotype M49 of *S. pyogenes* (GenBank accession number KC155346.1) was designed to encode amino acids 38–843 of EndoS2, as described previously [259]. The construct includes D184A/E186 exchanges to abolish catalytic activity (**Figure 21**).

EndoS2$^{D184A/E186L}$ (residues 38-843)

MEKTVQTGKTDQQVGAKLVQEIREGKRGPLYAGYFRTWHDRASTGIDGKQQHPEN
TMAEVPKEVDILFVFHDHTASDSPFWSELKDSYVHKLHQQGTALVQTIGVNELNGR
TGLSKDYPDTPEGNKALAAAIVKAFVTDRGVDGLDIAILHEFTNKRTPEEDARALNV
FKEIAQLIGKNGSDKSKLLIMDTTLSVENNPIFKGIAEDLDYLLRQYYGSQGGEAEV
DTINSDWNQYQNYIDASQFMIGFSFFEESASKGNLWFDVNEYDPNNPEKGKDIEG
TRAKKYAEWQPSTGGLKAGIFSYAIDRDGVAHVPSTYKNRTSTNLQRHEVDNISHT
DYTVSRKLKTLMTEDKRYDVIDQKDIPDPALREQIIQQVGQYKGDLERYNKTLVLTG
DKIQNLKGLEKLSKLQKLELRQLSNVKEITPELLPESMKKDAELVMVGMTGLEKLNL
SGLNRQTLDGIDVNSITHLTSFDISHNSLDLSEKSEDRKLLMTLMEQVSNHQKITVK
NTAFENQKPKGYYPQTYDTKEGHYDVDNAEHDILTDFVFGTVTKRNTFIGDEEAFAI
YKEGAVDGRQYVSKDYTYEAFRKDYKGYKVHLTASNLGETVTSKVTATTDETYLVDV
SDGEKVVHHMKLNIGSGAIMMENLAKGAKVIGTSGDFEQAKKIFDGEKSDRFFTW
GQTNWIAFDLGEINLAKEWRLFNAETNTEIKTDSSLNVAKGRLQILKDTTIDLEKMD
IKNRKEYLSNDENWTDVAQMDDAKAIFNSKLSNVLSRYWRFCVDGGASSYYPQYT
ELQILGQRLSNDVANTLKDLLEHHHHHH

Figure 21: Protein sequence of inactive EndoS2 variant used for crystallography. Construct encodes amino acids 38–843 and includes D184A/E186L exchanges (highlighted in orange) to abolish catalytic activity. A C-terminal linker and His-tag were also included (sequence LEHHHHHH; highlighted in blue).

2.2 Protein expression

2.2.1 Immunoglobulins

Immunoglobulin constructs were transformed into XL10-Gold® Ultracompetent cells (Agilent) using the manufacturer's protocol. DNA was extracted prior to transfection using a QIAGEN® Plasmid Maxi Kit (Agilent) using the manufacturer's protocol. Constructs were transiently expressed in FreeStyle293F cells (ThermoFisher) using FreeStyle™ MAX Reagent (ThermoFisher), as described in the manufacturer's protocol. Following transfection, cells were incubated at 37 °C and 8 % CO_2 for seven days, shaking at 125 rpm (New Brunswick S41i incubator).

2.2.2 IdeS, EndoS and EndoS2 variants

Proteins originating from the *S. pyogenes* bacterium were expressed in *E. coli* BL21 (DE3)pLysS cells (ThermoFisher), all using the same method.

Constructs were transformed into 30 µL of expression cells (using 1 µL of DNA at a concentration of 50 ng/µL), using the manufacturer's protocol. For large-scale expressions, 1 L cultures of cells were grown in 2 L baffled conical flasks (Fisher Scientific). Cells were grown in Terrific Broth (Melford) with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Cells were grown at 37 °C, shaking at 200 rpm (in an Innova 43 R incubator, New Brunswick Scientific), until an optical density

at 600 nm (OD_{600}) of 0.8 was reached. Protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), after which the incubation temperature was lowered to 25 °C and cultures were left to shake overnight at 200 rpm. Cells were harvested by centrifugation at 6220 × *g* for 20 minutes, after which supernatant was discarded. Cell pellets were stored at −20 °C prior to purification.

2.3 Protein purification

2.3.1 Immunoglobulins

Cells were harvested by centrifugation at 3100 × *g* for 30 minutes, following which cell pellets were discarded. Supernatants were filtered through a Nalgene™ Rapid-Flow™ Sterile Disposable Filter Unit (0.2 μm; ThermoFisher) to remove any remaining cell debris. All immunoglobulins were purified by affinity purification with a HiTrap Protein A HP column (Cytiva), with the exception of IgG3 Fc^{E382R} which was purified with a HiTrap Protein G HP column (Cytiva). Buffers used for antibody purification are detailed in **Table 3**.

Table 3: Buffers used for antibody purification.

Buffer	**Composition**	
Binding	A: 20 mM NaH_2PO_4 B: 20 mM Na_2HPO_4	Mix A and B to reach pH 7.0
Elution	A: 0.1 M sodium citrate tribasic dihydrate B: 0.1 M citric acid monohydrate	Mix A and B to reach pH 3.0
Neutralisation	1 M Tris-HCl, pH 9.0	
SEC	PBS	

For affinity purification, the column was first equilibrated with 10 column volumes (CV) binding buffer, prior to sample application. The column was washed with 10 CV binding buffer and antibodies subsequently eluted with 30 mL elution buffer (into 10 mL neutralisation buffer). Wash steps were carried out at a flow rate of 5 mL/min; sample application and elution were carried out at flow rates of 2.5 mL/min.

Antibodies were exchanged into phosphate-buffered saline (PBS) buffer using a Vivaspin 20 centrifugal concentrator (MWCO 30 kDa, Sigma) and concentrated to under 1 mL. Further purification was carried out by size exclusion chromatography (SEC), using a Superdex 200 16/600 column (Cytiva) equilibrated in PBS. The sample was applied to the column at 1 mL/min.

2.3.2 IdeS, EndoS and EndoS2 variants

All purification steps up to and including the nickel affinity purification were carried out at 4 °C, with buffers and cells kept on ice throughout the purification process. Buffers used in purification are detailed in **Table 4**. Cell pellets were resuspended in PBS containing 2 μg/mL DNAse1 (Sigma) and a pinch of lysozyme (Sigma), and homogenised with a glass homogeniser (Fisher) and broken apart using a cell disruptor (Constant Cell Disruption Systems). Cell lysates were centrifuged at 3100 × *g* for 20 minutes to remove any remaining cell debris, and then centrifuged at 100,000 × *g* for 1 hour, to remove cell membranes.

Table 4: Buffers used for purification of bacterial proteins expressed in *E. coli.*

Buffer	**Composition**
Resuspension	PBS with 2 μg/mL DNAse1 and a pinch of lysozyme
Dilution	100 mM HEPES, 150 mM NaCl, pH 7.5 (NaOH), 10 mM imidazole
Binding	100 mM HEPES, 150 mM NaCl, pH 7.5 (NaOH)
Elution	100 mM HEPES, 150 mM NaCl, pH 7.5 (NaOH), 500 mM imidazole
SEC	10 mM HEPES, 150 mM NaCl, pH 8.0 (NaOH)

The remaining supernatant was diluted in an equal volume of dilution buffer, filtered through a Nalgene™ Rapid-Flow™ Sterile Disposable Filter Unit (0.2 μm; ThermoFisher) and subsequently applied to a HisTrap HP column (Cytiva) equilibrated with 10 CV binding buffer. Following sample application, the column was washed with 10 CV binding buffer with 10 mM imidazole (achieved with a mix of 98 % binding buffer and 2 % elution buffer). Proteins were eluted with a stepwise gradient of increasing elution buffer %, collecting 25 mL fractions at increasing concentrations of imidazole (specific concentrations used for each protein are detailed in their respective results chapters). Wash steps were carried out with a flow rate of 5 mL/min; sample application and elution steps were carried out with a flow rate of 2.5 mL/min.

Fractions were concentrated separately using Vivaspin 20 centrifugal concentrators (MWCO 10 kDa) and checked for purity using SDS-PAGE (see section 2.5.2). Fractions containing the most and purest protein were then pooled and applied (at 1 mL/min) to a Superdex 75 16/600 column (Cytiva) equilibrated in 10 mM HEPES, 150 mM NaCl, pH 8.0.

2.4 Enzyme-substrate complex formation

2.4.1 IdeS-IgG1 Fc complexes

Purified IdeS and IgG1 Fc (E382X variants) were combined in a 1:1 molar ratio. The sample was concentrated to under 1 mL and then applied to a Superdex 200 16/600 column (Cytiva) at 1 mL/min. The column was equilibrated in 10 mM HEPES, 150 mM NaCl, pH 8.0. Peak fractions were subsequently pooled and checked for purity with SDS-PAGE prior to crystallisation.

2.4.2 EndoS-IgG1 Fc complexes

Purified EndoS and IgG1 Fc E382X variants were combined in a 1:1 molar ratio. The sample was concentrated to under 1 mL and then applied to a Superdex 200 16/600 column (Cytiva) at 1 mL/min. The column was equilibrated in 10 mM HEPES, 150 mM NaCl, pH 8.0. Peak fractions were subsequently pooled and checked for purity with SDS-PAGE prior to crystallisation.

N.B. following initial crystallisation attempts, the stoichiometry of EndoS to IgG1 Fc in the crystal structure was observed to be 2:1 (i.e., two EndoS molecules binding one IgG1 Fc homodimer). Subsequent purifications were adjusted so that EndoS and IgG1 Fc were combined in a 2:1 molar ratio.

Full-length EndoS and EndoS2 were combined with IgG1 Fcs in a 2:1 molar ratio and purified using an identical method.

2.5 Biophysical methods

2.5.1 Determination of protein concentration

Protein concentrations were determined using a DS-11+ Spectrophotometer (DeNovix). Molecular weight and extinction coefficient parameters for each construct were calculated using the ProtParam tool [313] from the ExPASY server.

2.5.2 SDS-PAGE

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out using NuPAGE™ 4 to 12% Bis-Tris protein gels. 5 μL of NuPAGE™ lithium dodecyl sulfate (LDS) Sample Buffer was added to 15 μg of protein sample prior to loading. The buffer used for SDS-PAGE was NuPAGE™ MES SDS Running Buffer. Gels were run at 160 V for 45 minutes, stained using Coomassie Brilliant Blue (ThermoFisher) for 20 minutes (or longer if needed) and destained using a solution comprised of 10 % acetic acid, 40 % methanol and 50 % milliQ water. 10 μL of

Precision Plus Protien™ Kaleidoscope™ Prestained Protein Standards (Bio-Rad) was run alongside protein samples, as molecular weight standards.

2.6 Protein crystallisation

Proteins were exchanged into a buffer containing 50 mM HEPES, 150 mM KCl, pH 7.5 (KOH) using Vivaspin 20 centrifugal concentrators (Sigma) prior to crystallisation.

Crystallisation experiments were kindly set up by Christopher Holes (Technician, Macromolecular Crystallography, University of Southampton), as detailed below. Sitting-drop crystallisation experiments were set up into Intelli-Plate 96-3 LVR plates (Alphabiotech), using an Oryx4 robot (Douglas Instruments). Experiments for each well were set up in triplicate, with the top, middle and bottom droplets for each condition comprising 200 nL protein + 100 nL condition (2:1), 100 nL protein + 100 nL condition (1:1) and 100 nL protein + 200 nL condition (1:2), respectively. Droplets were set up next to a reservoir containing 40 μL of crystallisation condition. Experiments using seed stock to promote crystal growth were comprised of 100 nL protein + 100 nL condition + 100 nL seed. Crystallisation experiments were attempted with JCSG-plus™ (a screen which uses results from a crystallisation study by the Joint Center for Structural Genomics), Morpheus and ProPlex™ HT-96 screens (Molecular Dimensions). Crystals were left to grow at 21 °C.

Crystals were cryoprotected with a solution comprising the mother liquor (i.e., solution the crystals were grown in) with 20 % glycerol added. Crystals were harvested using LithoLoops (Molecular Dimensions) and vitrified in liquid nitrogen prior to data collection.

Crystalline seed stock for various projects was kindly made by Christopher Holes (Technician, Macromolecular Crystallography, University of Southampton), using MicroSeed Beads (Molecular Dimensions). Crystals to be made into seed stock were pipetted into an Eppendorf tube containing the bead, along with the reservoir of the crystallisation condition. Crystals were crushed by vortexing the seed bead tube, keeping the tube cold by transferring to ice every ~30 seconds, and then stored at −20 °C prior to use.

A hanging-drop optimisation screen (detailed in section 4.2.3, Appendix D) was set up manually using a 24-well plate (Molecular Dimensions). The screen was designed by the 'Make Tray' server on the Hampton Research website (https://hamptonresearch.com/make-tray.php). Volumes of reagents to add were calculated by this server, by inputting the desired crystallisation condition to be optimised, in order to comprise a final reservoir volume of 500 μL. Droplets were set up as 1 μL protein + 1 μL reservoir, and suspended over the reservoir using 22 mm^2 Siliconized Cover Slips (Molecular Dimensions). Wells were sealed using Molykote High Vacuum Grease (Sigma), and the plate was left to incubate at 21 °C.

2.6.1 Quantification of IgG1 Fc crystallisation

For comparison of crystallisation ability of various IgG1 Fc constructs, identical sitting-drop crystallisation experiments were set up as detailed above. Crystallisation of wild-type IgG1 Fc and each Fc^{E382X} variant was trialled with JCSG-plus™ and Morpheus screens, at Fc concentrations of 10 mg/mL and 15 mg/mL. Crystals were left to grow at 21 °C for eight days, after which the number of conditions producing crystals were counted.

2.6.2 IgG1 Fcs

Crystals of IgG1 Fc^{E382S} were grown in 0.2 M ammonium sulphate, 0.1 M Tris (pH 7.5), 25 % w/v PEG 8000. Crystals of IgG1 Fc^{E382A} were grown in 0.17 M ammonium sulphate, 25.5 % w/v PEG 4000, 15 % v/v glycerol.

2.6.3 Truncated $IdeS^{C94A}$-IgG1 Fc^{E382A} complex

Crystal trays were set up with truncated, inactive IdeS in complex with each of the Fc E382X variants. Crystals of the $IdeS^{C94A}$-IgG1 Fc^{E382A} complex were grown in condition F9 from the Morpheus crystallisation screen, comprising 0.12 M monosaccharides mix, 0.1 M buffer system 3 (pH 8.5), 30 % v/v precipitant mix 1.

2.6.4 Truncated $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} complex

Crystal trays for truncated, inactive EndoS were set up in complex with all the IgG1 Fc^{E382X} variants. Subsequent crystallisation experiments were set up using seed stock of EndoS-IgG1 Fc crystals harvested from initial trials, diluted to 1 in 1000. Crystals of the $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} complex were grown in condition B8 from the Morpheus crystallisation screen [314], comprising 0.09 M halogens, 0.1 M buffer system 2 (pH 7.5), 37.5 % v/v precipitant mix 4.

2.6.5 $EndoS2^{D184A/E186L}$-IgG1 $Fc^{L234C/E382A}$ complex

Crystal trays of inactive EndoS2 were set up in complex with all the IgG1 Fc^{E382X} variants, truncated versions of these variants, and variants containing an additional disulphide bridge within the hinge region (detailed in section **5.2.3**). Crystals of the protein complex were grown in condition G12 from the Morpheus crystallisation screen [314], comprising 0.1 M carboxylic acids, 0.1 M buffer system 3, pH 8.5, precipitant mix 4.

2.7 X-ray crystallography

Protein crystals were taken to either Diamond Light Source (DLS) in Oxford, UK, or the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, for X-ray diffraction studies.

2.7.1 Data collection

Vitrified crystals were stored in liquid nitrogen prior to data collection, and were placed in a 100 K cryo-stream during data collection. Crystallographic data collection parameters for proteins investigated in this thesis are detailed in the tables below.

Table 5: Data collection parameters for X-ray diffraction of IgG1 Fc crystals.

Parameter	**Protein**	
	IgG1 Fc^{E382S}	**IgG1 Fc^{E382A}**
Beamline	ID30A-3 (ESRF)	ID23-2 (ESRF)
Wavelength (Å)	0.968	0.873
Flux (photons/s)	6.50e+11	5.45e+11
Transmission (%)	10.2	10.3
Exposure (s)	0.024	0.031
Beam size (μm)	15 × 15	4 × 4
Oscillation angle (°)	0.25	0.1
Number of images	3600	3600

Table 6: Data collection parameters for X-ray diffraction of IdeS^{C94A}-IgG1 Fc^{E382A} crystals.

Parameter	IdeS^{C94A}-IgG1 Fc^{E382A}
Beamline	ID30A-3 (ESRF)
Wavelength (Å)	0.9677
Flux (photons/s)	7.32e+11
Transmission (%)	10.2
Exposure (s)	0.005
Beam size (μm)	15 × 15
Oscillation angle (°)	0.1
Number of images	1200

Table 7: Data collection parameters for crystals of EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} and EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystals used for structure solution.

Parameter	Protein	
	EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}	EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$
Beamline	I03 (DLS)	I03 (DLS)
Wavelength (Å)	0.9763	0.9763
Flux (photons/s)	9.07e+11	8.52e+11
Transmission (%)	10.00	10.00
Exposure (s)	0.0091	0.0182
Beam size (μm)	80 × 20	80 × 20
Oscillation angle (°)	0.10	0.1
Number of images	3600	3600

2.7.2 Data processing

Any specific data processing steps are described in the relevant results chapters. In brief, diffraction data images were processed using DIALS [315], xia2 [316] and/or XDS [317]. Merging of multiple data sets was achieved using the xia2.multiplex pipeline [318].

2.7.3 Structure determination and analysis

Specific methods for structure determination and refinement are outlined in the relevant results chapters. In brief, software within the ccp4i2 interface [319] was used for crystal structure solution and refinement. Structures were solved by molecular replacement, using the program Molrep [320]. Successive rounds of model building and refinement were carried out for iterative model improvement, using the programs Coot [321] and Refmac5 [322], respectively. Carbohydrates within protein structures were modelled in Coot [323] and validated using Privateer [324]. MolProbity [325, 326] and the Protein Data Bank (PDB) validation server [327] were utilised for validation of models. PDBePISA [85] was used for analysis of protein-protein interfaces in crystal structures. Figures depicting protein structures were prepared using UCSF ChimeraX [328].

Chapter 3 Engineering IgG1 Fc for protein complex crystallisation

IgG is the most abundant antibody present in human serum, and is potent against protein and carbohydrate-based antigens [11]. As discussed in chapter 1 (section 1.1.3), the "fragment crystallisable", or Fc, region binds receptors present on immune cells, as well as other proteins involved in immune activation, in order to activate signalling cascades which ultimately induce immune responses against foreign antigens [11]. Due to this strong immune-activating ability, IgG can be targeted for degradation in order to hinder immune response, such as during bacterial infection, or within the tumour microenvironment [230, 231, 233]. The majority of this thesis describes attempts to elucidate the structural basis behind IgG degradation as an immune evasion mechanism of the bacterium *S. pyogenes*.

X-ray crystallography is a high-resolution structural biology technique which can resolve atomic-level resolution of protein structure. The major drawback of this technique is that it often requires milligrams of highly purified protein, and the subsequent growth of protein crystals (often serendipitous) which diffract to high resolution. The use of multiple crystallisation screens allows hundreds of different conditions to be tested for crystallisation of a particular protein, in the hope that this will yield at least one "hit". The ability of a protein to crystallise is affected by a multitude of factors, and many methods exist for encouraging crystal growth, which are discussed in an excellent review by Deller *et al.* [329] and are summarised below.

By nature, proteins are dynamic and thus forcing them into a conformationally-restrained lattice of molecules, which is required for crystallography, is not always favourable. Therefore, in order to improve the chances of successful crystallisation, the smallest, most compact form of a protein should be used. Methods such as ESPRIT (Expression of Soluble Proteins by Random Incremental Truncation [330, 331]) and Combinatorial Domain Hunting [332, 333] can identify soluble protein fragments ideal for structure determination. Techniques such as hydrogen-deuterium exchange (HDX) coupled with mass spectrometry can be used to identify flexible regions within a protein [334-337], which can inform protein construct design for successful crystallisation [338-340]. *In situ* proteolysis can also be used to remove accessible, flexible regions of proteins and identify stable peptide fragments for crystallisation [341-351]. In addition, proteins can often be stabilised by the addition of ligands or additives such as polyethylene glycol [329, 352].

Modification of protein surface amino acids is a common technique for enhancing crystallisation [353, 354]. Reductive alkylation, such as methylation of the ε-amino groups in lysine side chains [355], can promote crystallisation by introducing new surface contacts for crystal packing interactions [356, 357]. A similar technique is the exchange of large, flexible amino acids on the

protein surface which rarely form crystal contacts (such as lysines, glutamates and glutamines). Such modifications reduce the surface entropy of the molecule, which, according to thermodynamic principles, should aid in the crystallisation process [353, 354, 358, 359] (**Figure 22**). A recent study by Gao *et al.* describes the use of this approach to generate crystals of several botulinum neurotoxins with improved diffraction quality [360], while Diaz *et al.* reported the use of surface entropy reduction to grow crystals of human spermine oxidase, prior to which no crystals were obtained [361].

$$\Delta G°_{cryst} = \Delta H°_{cryst} - T\Delta S°_{cryst}$$

Figure 22: Simplified equation for the thermodynamics of crystallisation. As discussed in Derewenda *et al.* [354], the enthalpic contribution towards crystallisation ($\Delta H°_{cryst}$) is governed by formation of weak intermolecular bonds within the crystal lattice and is therefore minimal. The entropic contribution ($\Delta S°_{cryst}$) is thus the main determinant of whether a protein will crystallise or not.

In a similar tactic, post-translational modifications (PTMs) present on protein surfaces can often hinder crystallisation. *N*-linked glycosylation, for example, is a very common PTM within eukaryotic systems. *N*-linked glycans possess both large conformational and compositional heterogeneity: IgG Fc glycans, which are relatively homogenous and sterically-restricted [36, 113] compared to those in other proteins, still comprise around 20 different glycoforms [111, 112] and possess significant conformational flexibility [115-119]. Strategies to remove glycans which hinder crystallisation include generation of protein variants lacking consensus sequences for glycosylation (Asn-X-Ser/Thr; [362]), use of glycosylation-deficient cell lines for expression [363-366] and use of glycosylation inhibitors to stall *N*-glycan processing at a particular stage [367-369] followed by deglycosylation with EndoH [370].

Moreover, crystallisation can be encouraged by stabilisation of the quaternary structure within a protein crystal. For example, introduction of cysteine side chains can induce formation of disulphide bridges within the crystal packing interface [371], which can stabilise novel oligomeric assemblies with improved crystallisation [372-375]. Engineered metal-binding sites [375, 376] and leucine zipper-like interfaces [377] have also been utilised for this purpose. An example of improved crystallisation by engineered crystal packing is a structure of the neonatal Fc receptor (FcRn) in complex with IgG Fc [89]: crystallisation of the native complex resulted in dimers of FcRn packing in-between Fc homodimers, which formed a so-called “oligomeric ribbon” structure diffracting anisotropically to very low resolution (**Figure 23a**; [378]). In order to overcome this, Martin *et al.* designed a heterodimeric Fc which could bind FcRn on one γ-chain only, which altered the crystal packing and allowed structure solution to 2.8 Å resolution (**Figure 23b**; [89]). Furthermore, crystal packing can sometimes be improved by co-crystallisation of a protein with a binding partner, thereby altering the surface area available to form packing interactions [379, 380]; examples include co-crystallisation with DNA [381] and antibody fragments [382].

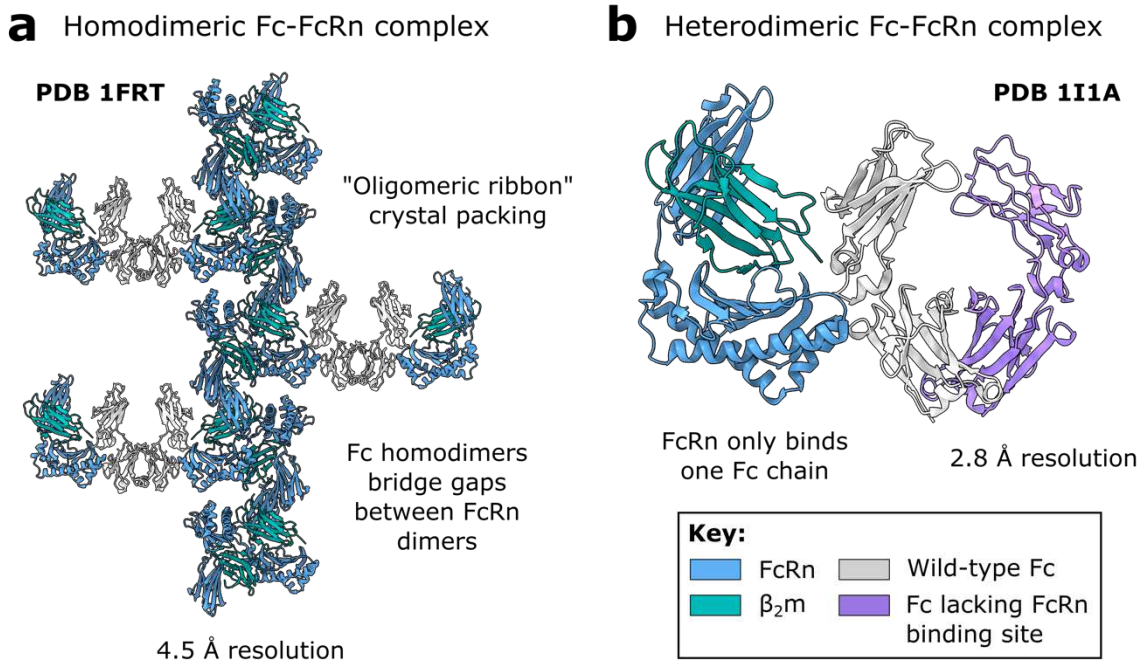


Figure 23: Improved resolution of FcRn-Fc crystal structure by altered crystal packing. a Crystal structure of homodimeric Fc-FcRn complex (PDB 1FRT [378]) diffracting to 4.5 Å. Resolution is limited by "oligomeric ribbon" structure formed during crystallisation. **b** Crystal structure of heterodimeric Fc-FcRn complex (PDB 1I1A [89]) diffracting to 2.8 Å. Crystal packing is altered by engineering one Fc γ-chain to lack the FcRn binding site. Adapted from Figure 1 in Martin *et al.* [89].

There are several examples of crystal lattice engineering carried out for crystallisation of proteins involved in immune recognition [89, 383, 384]. However, the crystallisation of IgG-protein complexes presents a unique challenge. As its name indicates, the Fc region of IgG is highly crystallisable, a property which can outcompete the growth of Fc-protein complex crystals. This is highlighted by the number of crystal structures deposited within the PDB: 107 structures of apo, human IgG Fc had been deposited as of August 2023. In contrast, 45 structures of human IgG Fc in complex with another protein had been solved (not including structures presented within this thesis). The solution to this problem was hypothesised to be the use of IgG Fc constructs with reduced self-crystallisation ability, in order to encourage growth of Fc-protein complex crystals. This was attempted by substitution of an amino acid found to be important for Fc crystal packing (described in section 3.1). This strategy, in which a highly self-crystallisable binding component has been engineered to promote protein complex crystallisation, has been used elsewhere: for example, by Kirchdoerfer *et al.* for crystallisation of a *Bacillus anthracis* glycoprotein in complex with a variable lymphocyte receptor [384]. In addition, Davies *et al.* applied mutagenesis of an omalizumab Fab to prevent its self-crystallisation and generate a crystal structure of the IgE Fc-Fab complex [385, 386].

The following chapter describes analysis of typical IgG1 Fc crystallisation, and the design of "less-crystallisable" IgG1 Fc variants for Fc-protein complex crystallisation. These Fc variants were subsequently used for structural characterisation of the interactions of immune evasion enzymes from *S. pyogenes* with IgG antibodies (discussed in subsequent chapters). Some of the results presented in

this chapter (specifically, **Figure 26** and **Figure 27**, along with the IgG1 Fc^{E382S} crystal structure with PDB 8A48) appear in the following publication: Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. *Nat Commun.* 2022;13:7801. This publication is included within Appendix A for reference.

3.1 Analysis of typical IgG1 Fc crystallisation

The co-crystallisation of IgG Fc with enzymes is notoriously difficult, due to the inherent ability of the Fc fragment to crystallise on its own. The aim of this work was to identify favourable contacts present in typical Fc crystals, in order to devise a strategy to counteract its selective self-crystallisation. Human IgG Fc structures deposited within the PDB have commonly crystallised in the $P2_12_12_1$ space group (~61 % of apo IgG Fc structures, as of August 2023; **Table 8**).

Table 8: Space groups reported for apo, human, IgG Fc crystal structures deposited within the PDB.

Space group[a]	Number of IgG Fc structures	Average resolution (Å)	Average solvent content (%)	Average Matthews coefficient
$P2_12_12_1$	65	2.25	55.1	2.75
$C222_1$	14	2.35	56.6	2.76
$P12_11$	7	2.30	50.0	2.55
$P6_122$	6	2.65	65.6	3.63
$C121$	5	2.40	55.0	2.78

[a]The five most common space groups are listed (as of August 2023).

For all the deposited IgG Fc crystal structures, the reported space group has little influence on resolution, solvent content or Matthews coefficient (**Table 8**). The exception to this is structures crystallising in $P6_122$, which diffracted to lower resolution and had less-tightly packed crystal lattices (as inferred from greater solvent content and Matthews coefficient), although this may be biased by the small sample size of structures. The range of space groups reported reflects the vast diversity of Fc structures crystallised, with a range of amino acid exchanges included and differing glycosylation which affect crystallisation. It is clear, however, that the preferred space group for IgG Fc crystals is $P2_12_12_1$.

The crystal lattice contacts present in a typical, wild-type Fc structure crystallised in $P2_12_12_1$ (PDB 3AVE [36]) were studied, in order to identify amino acids which are important in this favourable

packing arrangement. This structure of wild-type IgG1 Fc reveals a tightly packed crystal lattice, with largely conserved contacts across both Fc chains (**Figure 24**). Of note was amino acid E382, which forms salt bridges with R255 in a neighbouring Fc molecule (and vice versa), in both chains of the Fc homodimer (**Figure 24b**). It was therefore hypothesised that replacement of this side chain would hinder the self-association of the Fc into this preferred crystal lattice.

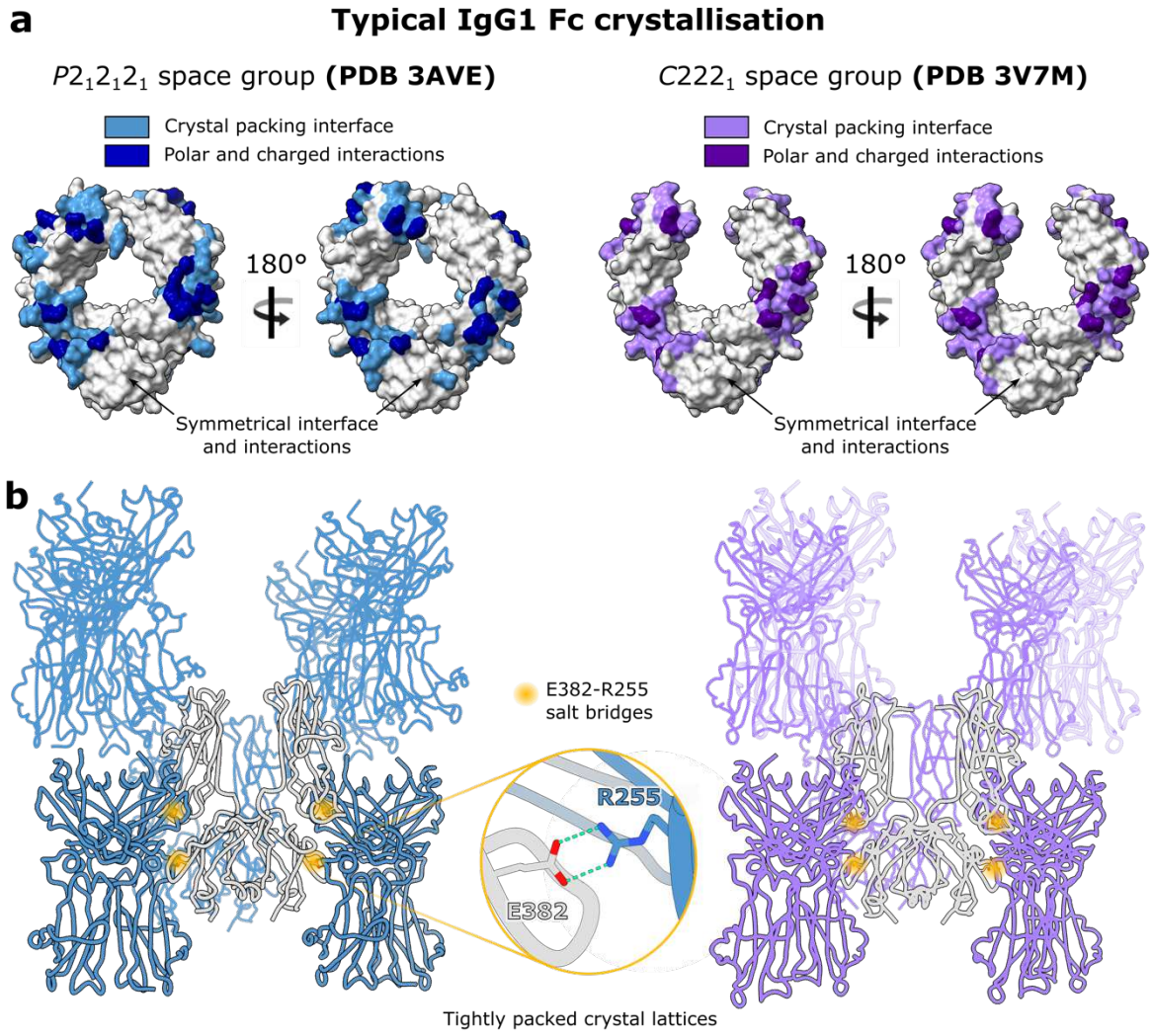


Figure 24: Observed crystal packing in "typical" IgG Fc crystal structures. a Analysis of crystal packing interface and interactions present in a typical, wild-type IgG1 Fc structure, as calculated using PDBePISA [85]. **b** Crystal packing resulting from IgG Fc crystallisation. E382–R255 salt bridges between symmetry-related Fcs in the crystal lattice are highlighted with a yellow circle. Neighbouring Fcs in the crystal lattice contacting the origin Fc within a 3.2 Å contact distance are shown. **a, b** Analysis relating to IgG Fc structures crystallising in typical space groups $P2_12_12_1$ (PDB 3AVE [36]) and $C222_1$ (PDB 3V7M [387]) is depicted in blue and purple, respectively.

While the majority of IgG Fc crystal structures have been solved in space group $P2_12_12_1$, the second most common space group reported for IgG Fc structures is $C222_1$ (**Table 8**). Upon inspection, the majority of structures solved in this space group also have the conserved E382–R255 salt bridges and conserved crystal packing when compared to those solved in $P2_12_12_1$ (**Figure 24**); this results in ~70% of deposited IgG Fc crystal structures containing this salt bridge. Due to the homodimeric nature of the Fc, a pseudosymmetry axis can be placed between the two Cγ2–Cγ3 heavy chains,

which allows the structure to be solved in $C222_1$, if a single Cγ2–Cγ3 unit is utilised as the asymmetric unit. The simple conclusion from this analysis is that Fc structures solved in $P2_12_12_1$ and $C222_1$ space groups (as reported in the PDB) utilise the E382–R255 salt bridges in their favourable crystal packing arrangements.

Having identified the E382–R255 salt bridge commonly present in IgG Fc crystal lattices, IgG1 Fc variants containing amino acid exchanges at E382 were designed with the aim of hindering the self-crystallisation of the Fc. Three distinct variants were designed: IgG1 Fc^{E382R}, IgG1 Fc^{E382S} and Fc^{E382A}. The substitution to an arginine was hypothesised to disrupt the typical crystal lattice formation with the introduction of a large, oppositely charged amino acid at position 382. Substitutions to serine and alanine were also utilised: these amino acids are small and therefore commonly used for removal of residues without disrupting overall protein structure (e.g., generation of inactive IdeS and EndoS variants [246, 258]).

3.2 Expression and purification of IgG1 Fc^{E382X} variants

As described in section 2.1.1, E382R/S/A amino acid exchanges were introduced into wild-type IgG1 Fc using site-directed mutagenesis. The resulting constructs were expressed in FreeStyle293F cells and purified with affinity chromatography followed by SEC. The purity of Fcs was checked using SDS-PAGE, which is displayed in **Figure 25**. A equivalent, wild-type (WT) IgG1 Fc was expressed and purified to homogeneity using an identical protocol (**Figure 25**).

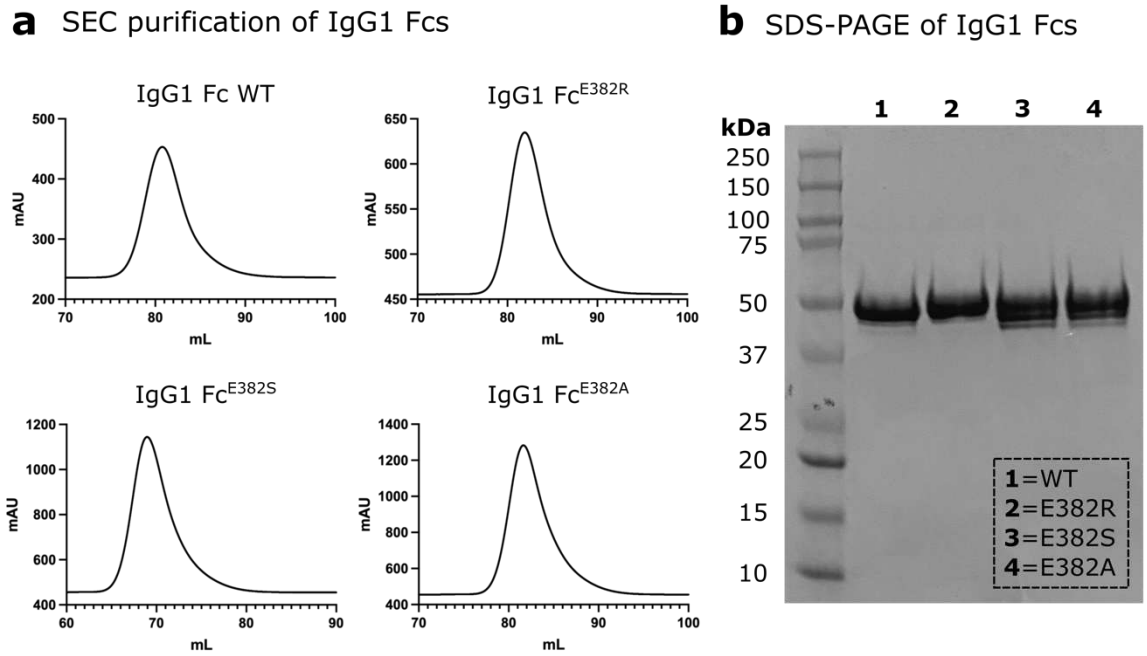


Figure 25: Purification of IgG1 Fcs. a SEC purification of wild-type (WT) IgG1 Fc, and Fc E382R/S/A variants. **b** SDS-PAGE of purified IgG1 Fcs. Within each sample, there is a single band consistent with the molecular weight of IgG1 Fc (~50 kDa).

Each Fc eluted as a single peak after SEC purification (**Figure 25a**). N.B. the baseline for each SEC run is not at zero due to a software fault. Fcs eluted at ~81–82 mL, with the exception of Fc^{E382S} which eluted at ~69 mL (the reason for this observation is unknown and was possibly a fault with the software/an unknown error in the protocol. Other SEC purifications for this protein construct eluted as expected; see Appendix B for an example). Subsequent SDS-PAGE analysis revealed, for each sample, a single band present at ~50 kDa, which is consistent with the molecular weight of IgG1 Fc (**Figure 25b**). A small amount of smearing is present within each of the bands, which likely reflects the heterogenous glycosylation present within each Fc sample.

Following successful purification of WT IgG1 Fc and IgG1 Fc^{E382X} variants, samples of each Fc were taken forward for crystallisation experiments.

3.3 Quantification of IgG1 Fc crystallisation

In order to compare the crystallisation abilities of the IgG1 Fc^{E382X} variants versus a wild-type IgG1 Fc, identical crystallisation experiments were set up for each Fc, using two different crystallisation screens (JCSG-plus™ and Morpheus; both from Molecular Dimensions). For each Fc, plates were set up at 10 mg/mL and 15 mg/mL. Crystals were left to grow at 21 °C and, after eight days, the number of conditions in each screen producing crystals were counted. The results of this analysis are depicted in **Figure 26**.

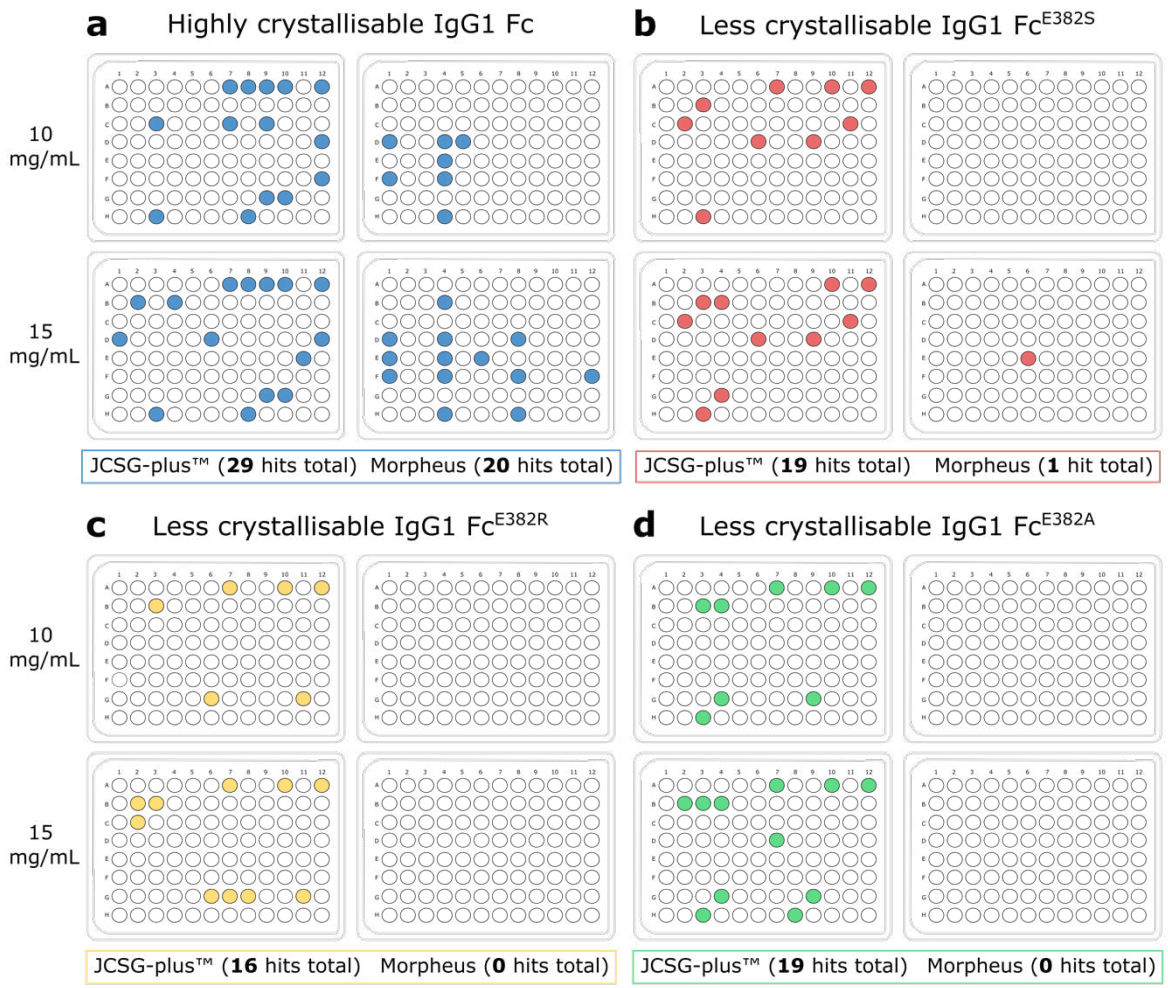


Figure 26: Crystallisation of wild-type IgG1 Fc and IgG1 Fc^{E382X} variants. Fcs were crystallised in JCSG-plus™ and Morpheus screens, at 10 mg/mL and 15 mg/mL, and at 21 °C. Crystal "hits" are indicated with a coloured circle. Results for each Fc are shown in panels **a**, **b**, **c** and **d** for wild-type IgG1 Fc, IgG1 Fc^{E382S}, IgG1 Fc^{E382R} and IgG1 Fc^{E382A} IgG1 Fc variants, respectively.

Across the four crystallisation experiments set up with the wild-type Fc, a total of 49 crystal "hits" were observed (29 in JCSG-plus™ and 20 in Morpheus; **Figure 26a**). In contrast, all the Fc^{E382X} variants displayed reduced crystallisation abilities with respect to the wild-type. Fc^{E382S} and Fc^{E382A} variants displayed similar levels of crystallisation, with 20 and 19 hits, respectively, resulting in a reduction in crystallisation ability compared to the wild-type of 59 % and 61 %, respectively (**Figure 26b, d**). Fc^{E382R} displayed further reduced crystallisation with 16 hits (**Figure 26c**), resulting in a 67 % reduction compared to the wild-type Fc. For all three Fc^{E382X} variants, crystallisation was essentially abolished in the Morpheus screen.

This analysis showed a clear reduction in crystallisation of the IgG1 Fc^{E382X} variants; however, crystallisation was not abolished completely. We therefore sought to determine the crystal structures of these Fc variants, in order to elucidate how the E382X exchanges affect the crystal packing.

3.4 Structural characterisation of IgG1 Fc^{E382X} variants

3.4.1 IgG1 Fc^{E382S}

Crystals of IgG1 Fc^{E382S} were grown in 0.2 M ammonium sulphate, 0.1 M Tris (pH 7.5), 25 % w/v PEG 8000. Crystals were cryo-protected in mother liquor with 20 % glycerol added and flash-frozen in liquid nitrogen prior to data collection, which was carried out on beamline ID30A-3 at the European Synchrotron Radiation Facility (ESRF). Parameters for data collection are detailed in **Table 5** (section 2.7.1).

These crystals diffracted to ~3 Å. Diffraction images were processed using XDS [317] and scaled in space group $P3_221$. The data were truncated to a final resolution of 3.04, to yield a CC_{half} of >0.3 for the highest resolution shell [388, 389]. Data collection statistics for this crystal are detailed in **Table 9** below. As of August 2023, space group $P3_221$ had not previously been reported for a human apo IgG Fc structure; however, there were examples of IgG Fc crystallised in complex with a small peptide (for example, PDB code 5DVK [390]) where the reported space group was $P3_221$. Interestingly, this peptide binds at the Cγ2–Cγ3 interface within the Fc, around the same area as E382, which likely affected the ability of these Fcs to crystallise in their "typical" lattice.

Table 9: Crystallographic data collection statistics for less-crystallisable IgG1 Fc^{E382S} variant.

Data Collection	**IgG1 Fc^{E382S}**
Beamline	ID30A-3 (ESRF)
Resolution range (Å)	47.5–3.04 (3.25–3.04)[a]
Space group	$P3_221$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 106.78, 106.78, 104.01 90, 90, 120
Wavelength (Å)	0.968
Unique reflections	13476 (2358)
Completeness (%)	99.7 (98.4)
R_{merge}	0.034 (1.117)
R_{meas}	0.048 (1.580)
R_{pim}	0.034 (1.117)
$I/\sigma(I)$	15.2 (0.6)
Multiplicity	1.9 (1.9)
CC_{half}	0.999 (0.321)
Wilson *B* factor (Å^2)	113.7

[a]Values for the highest resolution shell are shown in parentheses.

The structure of IgG1 Fc^{E382S} was solved by molecular replacement, using a wild-type IgG1 Fc structure as a search model (PDB 3AVE). When refining the model, non-crystallographic symmetry (NCS) restraints were used, in addition to restraints provided by the PDB-REDO server [391]. Refinement statistics for this structure are included within **Table 10** below. Coordinates for this crystal structure have been deposited within the PDB, with accession code 8A48.

Table 10: Crystallographic refinement statistics for less-crystallisable IgG1 Fc^{E382S} variant.

Refinement	**IgG1 Fc^{E382S}**
Number of reflections (all/free)	13476/654
R_{work} (%)	19.7
R_{free} (%)	21.8
RMSD[1]: Bonds (Å) Angles (°)	 0.0025 0.797
Molecules per ASU[2]	1
Atoms per ASU[2]	3,549
Average *B* factors (Å^2) (protein/ligand/water)	 (141.27/176.11/105.09)
Model quality (Ramachandran plot): Most favoured region (%) Allowed region (%) Outliers (%)	 95.88 3.63 0.48, 0.0*

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
*according to wwPDB X-ray Structure Validation Report

This crystal structure of the IgG1 Fc^{E328S} variant was subsequently compared to that of a wild-type IgG1 Fc, which crystallised in the "typical" space group $P2_12_12_1$ (as discussed in section 3.1). Interfaces within the crystal packing arrangement were calculated using PDBePISA [85]. This analysis is depicted in **Figure 27**.

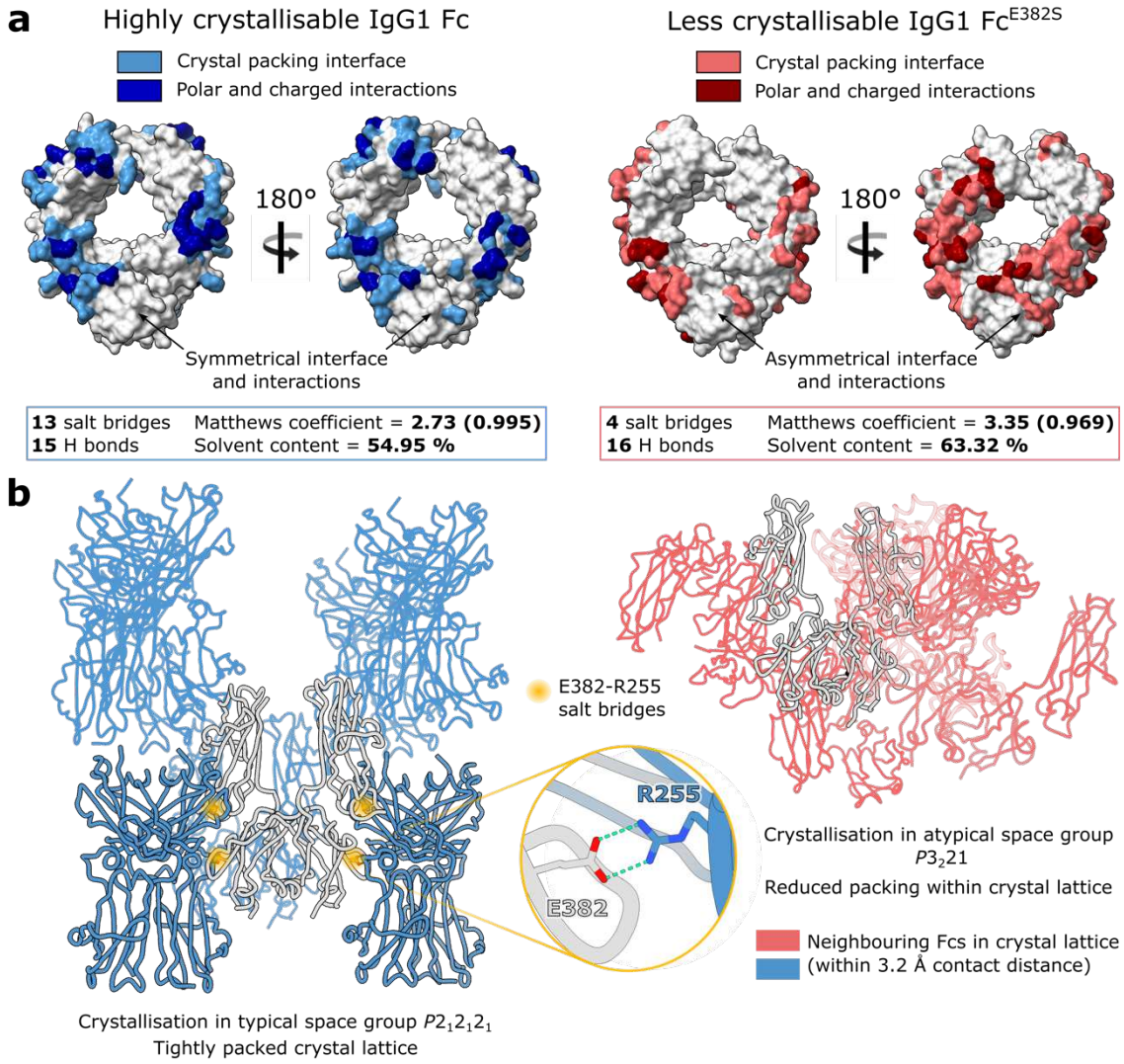


Figure 27: Observed crystal packing in wild-type and "less crystallisable" IgG1 Fc^{E382S} fragment. a Analysis of crystal packing interface and interactions present in a typical, wild-type IgG1 Fc crystal structure (PDB code 3AVE [36]) and IgG1 Fc^{E382S} variant, as calculated by PDBePISA [85]. **b** Crystal packing resulting from crystallisation in typical space group $P2_12_12_1$ and atypical space group $P3_221$, for the wild-type IgG1 Fc and IgG1 Fc^{E382S} variant, respectively. E382-R255 salt bridges between symmetry-related Fcs in the $P2_12_12_1$ crystal lattice are highlighted with a yellow circle. Neighbouring Fcs in the crystal lattice contacting the origin Fc within a 3.2 Å contact distance are shown. **a, b** Analysis relating to the wild-type IgG1 Fc and IgG1 Fc^{E382S} variant is depicted in blue and red, respectively.

Analysis of the crystal contacts revealed that the IgG1 Fc^{E328S} variant makes fewer interactions with symmetry-related molecules in the crystal (four salt bridges and sixteen hydrogen bonds; **Figure 27a**), which are asymmetrical across the two Fc chains, resulting in altered crystal packing (**Figure 27b**). Furthermore, as calculated within the ccp4i2 interface [319], the E382S variant had a higher solvent content and Matthews coefficient compared to the wild-type Fc (**Figure 27a**), and additionally when compared to the average values for Fc crystals with the $P2_12_12_1$ space group (**Table 8**). The Matthews coefficient of a crystal is defined as the crystal volume per unit of protein molecular weight [392], therefore the smaller value for the wild-type structure indicates a more tightly packed crystal. Crystallisation of this Fc variant therefore appears to have been rendered less favourable.

3.4.2 IgG1 Fc^{E382A}

Crystals of the IgG1 Fc^{E382A} variant were grown in 0.17 M ammonium sulphate, 25.5 % w/v PEG 4000. These crystals were cryoprotected (as described in section 2.6) and taken for data collection on beamline ID23-2 at the ESRF. Parameters for data collection are detailed in **Table 5** (section 2.7.1). IgG1 Fc^{E382A} crystals from this condition diffracted to ~2.6 Å; diffraction images from this dataset were subsequently processed using xia2 and DIALS [315, 316]. The final resolution of the dataset was cut to 2.7 Å, to yield a completeness of >95 % and a CC_{half} of >0.3 for the highest resolution shell [388, 389]. Data collection statistics are detailed below in **Table 11**.

Table 11: Crystallographic data collection statistics for less-crystallisable IgG1 Fc^{E382A} variant.

Data Collection	**IgG1 Fc^{E382A}**
Beamline	ID23-2 (ESRF)
Resolution range (Å)	46.14–2.70 (2.74–2.70)[a]
Space group	$P3_221$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 106.6, 106.6, 103.3 90, 90, 120
Wavelength (Å)	0.873
Unique reflections	19,105 (1046)
Completeness (%)	99.9 (95.8)
R_{merge}	0.189 (3.691)
R_{meas}	0.194 (3.799)
R_{pim}	0.045 (0.887)
$I/\sigma(I)$	11.3 (0.98)
Multiplicity	18.6 (18.2)
CC_{half}	0.998 (0.540)
Wilson *B* factor (Å^2)	59.1

[a]Values for the highest resolution shell are shown in parentheses.

The structure of the IgG1 Fc^{E382A} fragment was solved with molecular replacement (using PDB 3AVE as a search model), in the same "atypical" space group of $P3_221$ as observed for the IgG1 Fc^{E382S} variant. Automatic translation-libration-screw (TLS) parameters [393] and automatic NCS restraints improved the *R* factors and thus were applied during refinement. The model is of good quality, with no Ramachandran outliers as calculated by the wwPDB X-ray Structure Validation Report (see Appendix C). Refinement statistics for the final structure are listed in **Table 12** below.

Table 12: Crystallographic refinement statistics for less-crystallisable IgG1 Fc^{E382A} variant.

Refinement	**IgG1 Fc^{E382A}**
Number of reflections (all/free)	19,063/912
R_{work} (%)	19.4
R_{free} (%)	25.3
RMSD[1]: Bonds (Å) Angles (°)	 0.0085 1.57
Molecules per ASU[2]	1
Atoms per ASU[2]	3,698
Average *B* factors (Å^2) (protein/ligand/water)	71.6/97.0/59.2
Model quality (Ramachandran plot): Most favoured region (%) Allowed region (%) Outliers (%)	 97.57 2.19 0.24, 0*

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
*according to wwPDB X-ray Structure Validation Report

Analysis of the crystal packing was carried out using PDBePISA [85], as described above for the IgG1 Fc^{E382S} model, and is depicted in **Figure 28**. The IgG1 Fc^{E382A} crystal structure was found to display similarly altered crystallisation, as observed for the Fc^{E382S} crystal structure. When compared to the example of a "typical" IgG Fc crystal (PDB 3AVE, as in **Figure 24**, **Figure 27**), IgG1 Fc^{E382A} makes fewer interactions within the crystal lattice (three salt bridges and thirteen hydrogen bonds). The calculated solvent content (of 62.9%) and Matthews coefficient (of 3.32; **Figure 28**) is also higher than those values from the wild-type crystal (as well as the average values of Fc structures crystallising in $P2_12_12_1$; **Table 8**), indicating a less-tightly packed crystal lattice.

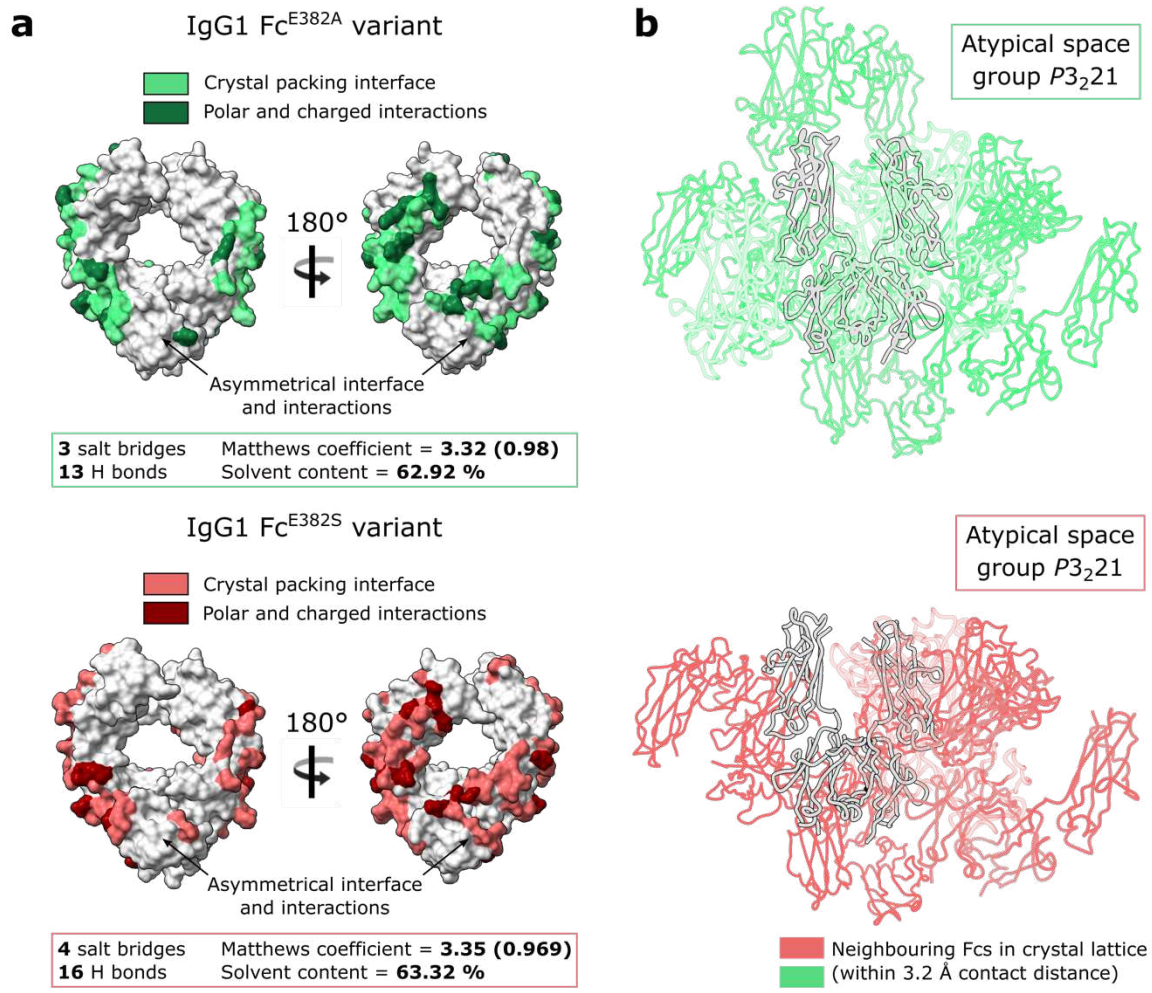


Figure 28: Comparison of crystal packing in "less crystallisable" IgG1 Fc^{E382A} and IgG1 Fc^{E382S} fragments. a Analysis of crystal packing interface and interactions present within the crystal of each Fc variant, as calculated by PDBePISA [85]. **b** Crystal packing resulting from crystallisation in atypical space group $P3_221$ for each Fc variant. Neighbouring Fcs in the crystal lattice contacting the origin Fc within a 3.2 Å contact distance are shown. **a, b** Analysis relating to the IgG1 Fc^{E382A} and IgG1 Fc^{E382S} variants is depicted in green and red, respectively.

Figure 28 shows how the IgG1 Fc^{E382A} and Fc^{E382S} variants crystallised in the same "atypical" space group, and this display very similar crystal packing interfaces. This is perhaps unsurprising, given the similar crystallisation profiles observed (**Figure 26** in section 3.3) by these two Fc variants. It therefore appears that crystallisation of these two Fc variants has been rendered less favourable.

3.4.3 IgG1 Fc^{E382R}

Although the IgG1 Fc^{E382R} variant was still able to crystallise at reduced capacity (see **Figure 26**, section 3.3), no diffracting crystals were obtained for analysis of the structure. It can therefore be concluded that the crystallisation ability of this Fc variant has also been severely impeded.

3.5 Discussion

The Fc fragment of IgG is known to crystallise readily, indicated inherently with its name "fragment crystallisable", as first identified in 1959 [394]. This poses a significant barrier to the generation of Fc-protein complexes. The aim of this chapter was therefore to elucidate the amino acid residues involved in this favourable Fc crystal packing arrangement, and from this deduce a strategy to impair Fc self-crystallisation, in order to promote protein complex crystallisation.

Analysis of the presently deposited human IgG Fc crystal structures revealed the strong preference for crystallisation in the $P2_12_12_1$ space group. Upon inspection, this favourable lattice arrangement is driven by the formation of salt bridge interactions between E382 and R255 side chains between adjacent molecules within the crystal lattice. This salt bridge was also found to be present in the majority of Fc crystal structures crystallising in space group $C222_1$, further highlighting its importance. It was therefore hypothesised that disruption of this salt bridge would hinder Fc crystallisation.

It is interesting to note that this salt bridge occurs in close proximity to the Fc Cγ2–Cγ3 interface. Inspection of the Fc-protein complexes currently deposited within the PDB shows that around half the complexed proteins bind Fc around this interface (which includes Protein A from *Staphylococcus aureus*, and the neonatal Fc receptor). Binding at this interface may therefore help prevent favourable Fc crystal lattice interactions.

Three IgG1 Fc variants were designed: E382R, E382S and E382A. The crystallisation ability of each variant was tested against a wild-type Fc using a simple experiment, in which identical crystallisation experiments were set up for each Fc and the number of crystals produced were counted. Each Fc variant displayed significantly reduced crystallisation compared to the wild-type, with a ~60% reduction in the E382A and E382S variants, and a near 70% reduction in the E382R variant. This may reflect the introduction of the oppositely-charged arginine residue in this variant, while the other two variants are more conservative and only remove the negative charge of the glutamate. It can be concluded that all three Fc variants are "less-crystallisable", compared to Fcs able to form the intermolecular E382R–R255 salt bridges.

Crystal structures were determined for the IgG1 Fc^{E382S} and Fc^{E382A} variants, which both crystallised in the $P3_221$ space group. It was interesting to note that this space group has not previously been reported for a human, apo IgG Fc crystal structure, although it is reported for a group of Fc structures crystallised in complex with a peptide binding around the Fc Cγ2–Cγ3 interface (e.g., PDB 5DVK [390]). Inspection of the crystal lattice packing for these two Fc variants revealed smaller packing interfaces and fewer packing interactions when compared to a typical, wild-type IgG1 Fc structure (PDB 3AVE). Larger calculated solvent content and Matthews coefficients for these variants also indicate less tightly-packed crystals when compared to the wild-type. Crystals of the IgG1 Fc^{E382R}

variant did not diffract sufficiently to enable structure solution, also suggesting severely reduced crystallisation, although this does not mean well-diffracting crystals of this variant are impossible to obtain.

In general, it can be concluded that this panel of Fc variants have been rendered “less-crystallisable” by exchange of the amino acid E382. It can be envisaged that the exchange of any amino acid similarly impacting crystal lattice formation could be employed for generating protein complex crystals, as evidenced in the solution of other immune complexes using this strategy [384, 386]. The IgG1 Fc^{E382R}, Fc^{E382S} and Fc^{E382A} variants were therefore utilised for the attempted co-crystallisation with IgG-degrading enzymes from *S. pyogenes*, which is described in the following chapters.

Chapter 4 Extensive substrate recognition of IgG Fc by the streptococcal protease IdeS

The bacterium *Streptococcus pyogenes* has evolved a diverse range of mechanisms for evading the human adaptive immune system [4]. This pathogen secretes multiple enzymes that target components of the immune system [4, 5]: one such example is IgG-degrading enzyme of *Streptococcus pyogenes* (IdeS), which directly targets and cleaves IgG antibodies [6]. Specifically, the enzyme cleaves between glycines 236 and 237 in the lower hinge region, yielding $F(ab')_2$ and Fc fragments [6, 230, 231]. The enzyme cleaves each Fc chain in sequential steps [230, 231]; however, a single cleavage is sufficient to render the antibody unable to recruit the immune system [231, 238]. The enzyme therefore has a role in immune evasion, by impeding cellular responses through immune recruitment mediated by the antibody Fc domain [6, 232, 238].

The interest in IdeS is two-fold: firstly, the enzyme is extremely efficient at cleaving human IgG [238]. In addition, IdeS is highly specific for IgG, and in fact no other substrates for the enzyme have been identified [6], in stark contrast to other cysteine proteases within the papain superfamily [7, 249-251]. Interestingly, IdeS is unable to hydrolyse synthetic peptides mimicking the IgG hinge region [244], which indicates that exosite binding to other regions of the Fc is required for IgG recognition by IdeS. This unique activity has prompted extensive research into the structure and function of this enzyme, in order to aid in its development for a wide range of therapeutic applications.

Under the name Imlifidase (brand name Idefirix®), IdeS has been clinically-approved as a pre-treatment for transplantation in hypersensitized individuals with chronic kidney disease [262, 264, 265]. The ability of IdeS to inactivate IgG has also led this enzyme to be investigated for use in the deactivation of pathogenic antibodies, such as in treatment of autoimmune disorders [267, 271, 273-275, 278, 280], deactivation of neutralising antibodies for *in vivo* gene therapy [288], and the potentiation of therapeutic antibodies by deactivation of competing serum IgG [105, 292]. Imlifidase has also been used in combination with endoglycosidase S (another IgG-inactivating enzyme secreted by *S. pyogenes*, which is discussed in chapter 5) for inactivation of donor-specific antibodies in murine allogeneic bone marrow transplantation [266].

Prior to the start of this project, it was largely unknown exactly how IdeS is able to cleave human IgG with such high specificity. This chapter describes work towards determining the crystal structure of IdeS in complex with the Fc region of IgG, in order to elucidate the structural basis behind its action. The panel of "less-crystallisable" IgG1 Fc variants described in chapter 3 were utilised for this work, in order to prevent selective self-crystallisation of the Fc, which had been observed in previous attempts to crystallise this protein-protein complex (see Dixon, E.V. (2014). Mechanisms

of immunoglobulin deactivation by *Streptococcus pyogenes* [PhD thesis]. University of Oxford. [395]).

Analysis of the IdeS-IgG1 Fc crystal structure (PDB 8A47) is included in the following publication: Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. *Nat. Commun.* 2022;13:7801. This publication is included within Appendix A for reference.

4.1 Expression and purification of IdeS^{C94A}-IgG1 Fc complexes

IdeS and related bacterial constructs were expressed in BL21 (DE3)pLysS cells and purified as detailed in sections 2.2.2 and 2.3.2.

4.1.1 IdeS^{C94A}-IgG1 Fc complex

Truncated IdeS^{C94A} was initially purified using nickel affinity purification, using step-wise elutions at increasing concentrations of imidazole: 20, 40 80, 100, 200 and 300 mM. SDS-PAGE analysis of each eluted fraction revealed a dominant band consistent with the molecular weight for truncated IdeS^{C94A} (~34 kDa) within the 80 mM, 100 mM and 200 mM imidazole fractions (**Figure 29a**). These fractions were therefore taken forward for further purification. IgG1 Fcs containing E382R/S/A exchanges were expressed and purified as detailed in sections 2.2.1 and 2.3.1. **Figure 29b** shows SDS-PAGE analysis of the Fcs following protein A purification; each of the eluted Fc samples contained a dominant band at a size consistent with IgG1 Fc molecular weight (~50 kDa). The E382S and E382A variants were homogenous following this purification; the E382R variant contained additional contaminating bands at ~150 and ~300 kDa. This Fc variant was therefore purified further using SEC (over a HiLoad® 16/600 Superdex 200 pg column), in order to remove these impurities.

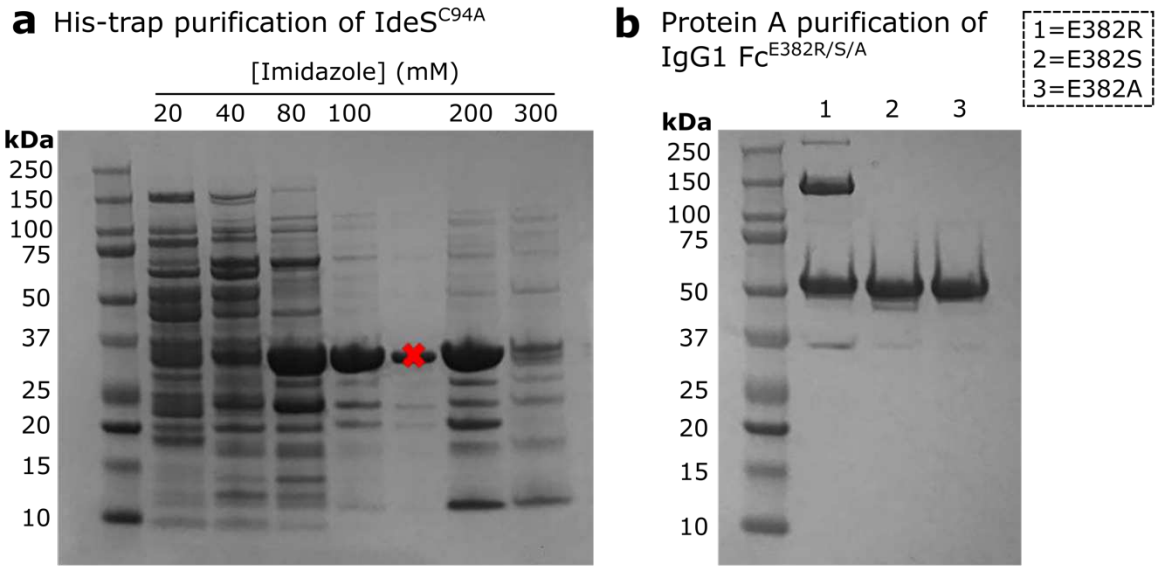


Figure 29: Purification of truncated IdeS^{C94A} and IgG1 Fc$^{E382R/S/A}$. a SDS-PAGE gel of fractions eluted from nickel affinity purification of IdeS^{C94A}. Truncated IdeS^{C94A} is overexpressed and present as a dominant band at ~35 kDa. Fractions were collected at 20 mM, 40 mM, 80 mM, 100 mM, 200 mM and 300 mM imidazole. The 80 mM, 100 mM and 200 mM fractions only were pooled for further purification. The lane marked by a red cross had overspilled from another lane and should be ignored. **b** SDS-PAGE of eluted IgG1 Fc$^{E382R/S/A}$ antibodies from Protein A purification. Each Fc is present as a dominant band at ~50 kDa. All Fcs were taken forward for further purification.

As judged by the SDS-PAGE in **Figure 29a**, the 80 mM, 100 mM and 200 mM fractions from truncated IdeS^{C94A} nickel affinity purification were taken forward for SEC purification (using a HiLoad® 16/600 Superdex 75 pg column). Each run produced a main peak eluting at ~96 mL (**Figure 30a**). T21 and T22 fractions (corresponding to 93.2–96.7 mL and 96.7–100.2 mL, respectively) were concentrated separately and subjected to SDS-PAGE analysis, which revealed that the majority of impurities had been removed (**Figure 30b**). The 100 mM and 200 mM fractions were pooled prior to subsequent purification.

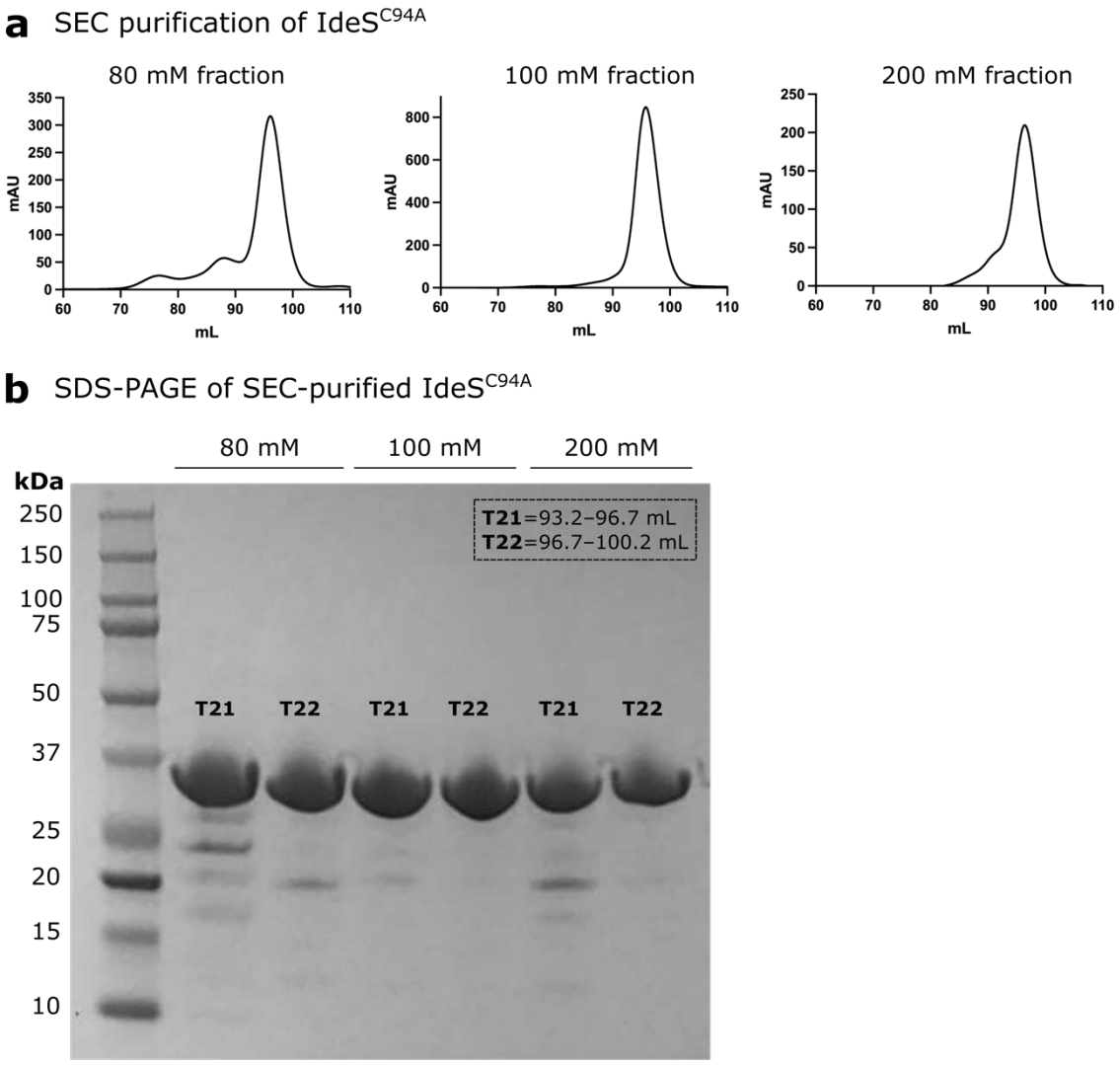


Figure 30: SEC purification of truncated IdeS^{C94A}. a Individual SEC purifications of IdeS^{C94A} fractions from nickel affinity purification (80, 100 and 200 mM). Fractions corresponding to the main peak (T21 and T22 for each SEC run) were concentrated separately. **b** SDS-PAGE analysis of main peak fractions from each SEC run in **a**. Each lane contains a dominant band consistent with the molecular weight for truncated IdeS^{C94A} (~35 kDa).

The ability of IdeS to form stable complexes with IgG1 Fc was subsequently investigated using analytical SEC (**Figure 31a**). Purified IdeS, IgG1 Fcs and their complexes were passed through the same HiLoad® 16/600 Superdex 200 pg column, at 1 mL/min, equilibrated in 10 mM HEPES, 150 mM KCl (pH 8.0). As expected, IdeS eluted at a later time point (~99 mL) compared to the three IgG1 Fcs (95.95 mL, 96.53 mL and 95.87 mL for E382R, E382S and E382A variants, respectively). IdeS was subsequently combined with each of the IgG1 Fc variants in a 1:1 molar ratio and applied to the column again. Each of the IdeS-IgG1 Fc complexes eluted earlier than the Fc on its own (at 92.87 mL, 92.53 mL and 92.42 mL for complexes with Fc variants E382R, E382S and E382A, respectively; **Figure 31a**), indicating the presence of a heavier molecular weight species. Each of the complexes eluted as a single peak with a slight shoulder on the right-hand side, indicating either lower molecular weight impurities or some complex dissociation.

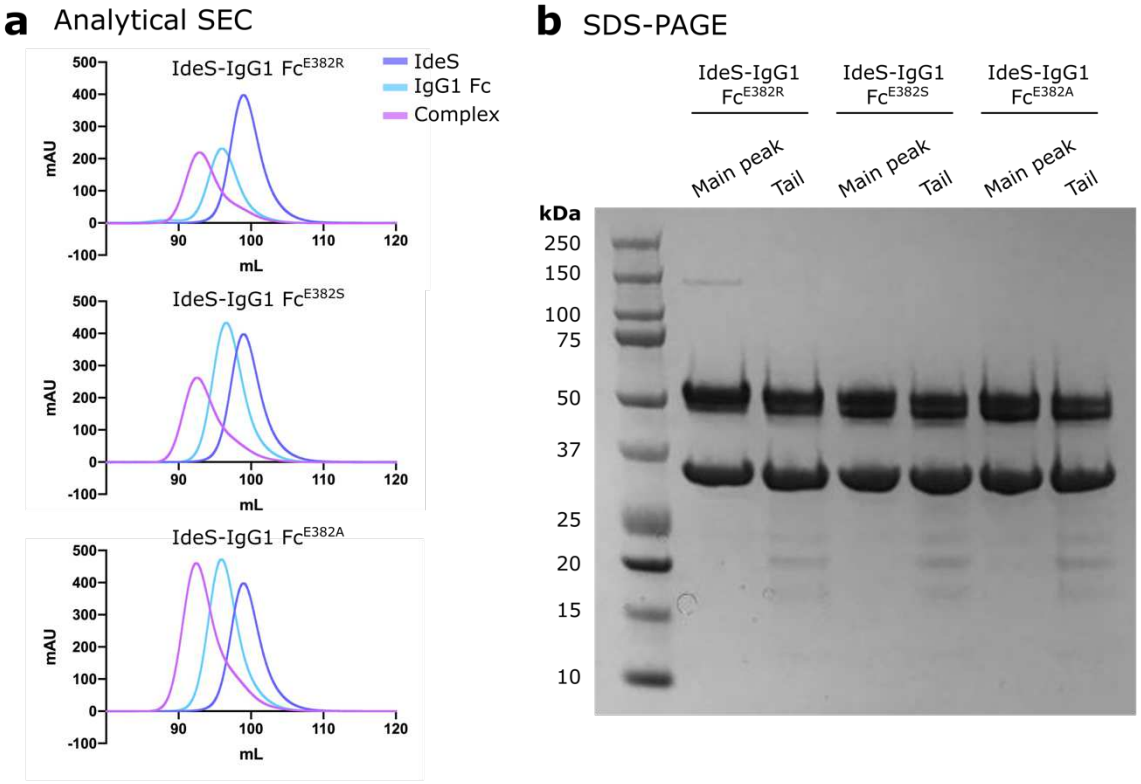


Figure 31: Purification of truncated IdeS^{C94A}-IgG1 Fc$^{E382R/S/A}$ complexes. **a** Analytical SEC of IdeS^{C94A}-IgG1 Fc$^{E382R/S/A}$ complexes. IdeS^{C94A} and Fcs containing E382R, E382S and E382A substitutions were first applied to a HiLoad® 16/600 Superdex 200 pg column, then complexes were combined in a 1:1 molar ratio and applied to the column again. **b** SDS-PAGE of the main peak and tail fractions following purification of each complex. Main peaks for each complex contain dominant bands for IgG1 Fc$^{E382R/S/A}$ (~50 kDa) and IdeS^{C94A} (~35 kDa). Main peak fractions only were taken forward for crystallisation.

Fractions corresponding to the main and "tail" peaks of each complex were concentrated separately and analysed using SDS-PAGE. The tail peaks each contained dominant bands consistent with the molecular weights of IdeS and IgG1 Fc, with small amounts of lower molecular weight impurities. Main peaks of each complex were judged pure as no contaminating bands could be observed, with the exception of a faint band at ~150 kDa in the IdeS-IgG1 Fc^{E382R} complex (**Figure 31b**). Fractions corresponding to the main peak of each complex only were taken forward for crystallisation.

4.1.2 Imlifidase^{C94A}-IgG1 Fc complexes

Following successful expression, purification and crystallisation of IdeS in complex with IgG1 Fc, additional related targets were identified. One such target was the clinically-approved variant of IdeS, named Imlifidase (brand name Idefirix®). This variant originates from the Mac-1 isoform of IdeS, which partially differs in sequence from the IdeS construct purified in the previous section (see chapter 1, section 1.3.1 and **Figure 45** within this chapter) [236]. The two variants have identical catalytic activity against IgG [237], thus the mechanism for binding IgG Fc is likely to be conserved.

However, a crystal structure of the Imlifidase-IgG1 Fc complex was pursued, in order to obtain structural information about the clinically-relevant IdeS variant.

A truncated variant of Imlifidase was expressed and purified as described in sections 2.2.2 and 2.3.2, with the inclusion of a C94A exchange to abolish catalytic activity. SDS-PAGE analysis of the eluted fractions from nickel affinity purification is shown in **Figure 32a**. A band consistent with the molecular weight of truncated Imlifidase^{C94A} (~34 kDa) was present in all fractions; however, fewer contaminating bands were observed in the 80 mM 100 mM and 200 mM fractions (**Figure 32a**). These fractions were therefore taken forward for further purification; the 100 mM and 200 mM fractions were combined prior to SEC purification, due to low protein concentrations present within these fractions.

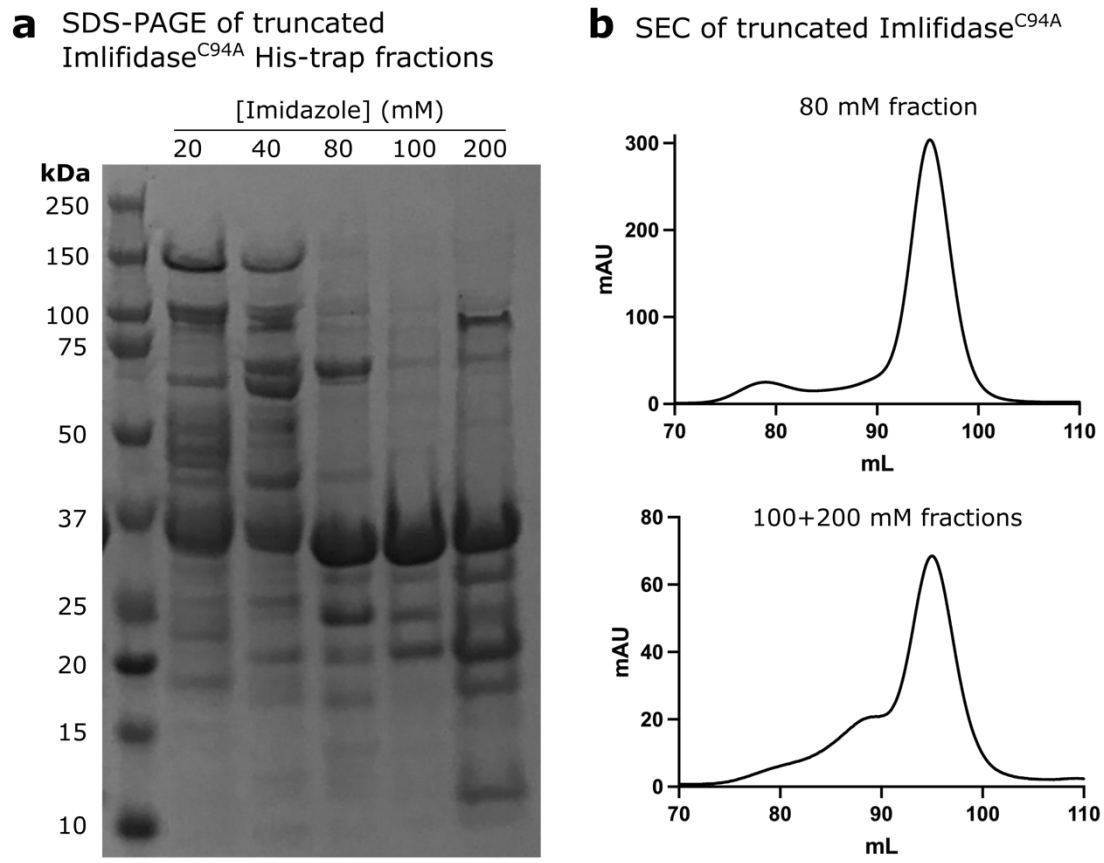


Figure 32: Purification of truncated Imlifidase^{C94A}. a SDS-PAGE of fractions from nickel affinity purification of Imlifidase, corresponding to elution at 20 mM, 40 mM, 80 mM, 100 mM and 200 mM imidazole. A dominant band consistent with the molecular weight of truncated Imlifidase^{C94A} (~34 kDa) is present within the 80 mM, 100 mM and 200 mM fractions. **b** SEC purification of the 80 mM fraction, and 100 + 200 mM fractions, of truncated Imlifidase^{C94A}, using a HiLoad® 16/600 Superdex 75 pg column.

Figure 32b shows SEC purification of the 80 mM truncated Imlifidase^{C94A}, and that of the pooled 100 mM and 200 mM fractions. Both samples resulted in a main peak eluting at ~95 mL. SEC of the 100 + 200 mM fraction produced a peak with relatively low mAU (~70), which additionally contained heavier molecular weight impurities or aggregation, as evidenced by a shoulder on the left-

hand side of the peak. In contrast, SEC of the 80 mM fraction produced a single peak only, with an approximately 4-fold increase in mAU (**Figure 32b**). Therefore, protein from this SEC run only was taken forward for further experiments. Fractions corresponding to the main peak only were pooled.

Truncated Imlifidase^{C94A} was subsequently combined in a 1:1 molar ratio with the panel of "less-crystallisable" IgG1 Fc variants (purified as described in section 2.3.1). Complexes were purified using the same method as the truncated IdeS^{C94A}-IgG1 Fc complexes (described in section 2.4.1). Each complex was applied to a HiLoad® 16/600 Superdex 200 pg column for additional purification, and eluted as a single peak at ~95 mL (**Figure 33a**). Elution volumes cannot be compared to that of apo truncated Imlifidase^{C94A} (detailed in **Figure 32**), however, due to use of different types of filtration columns utilised during these purifications.

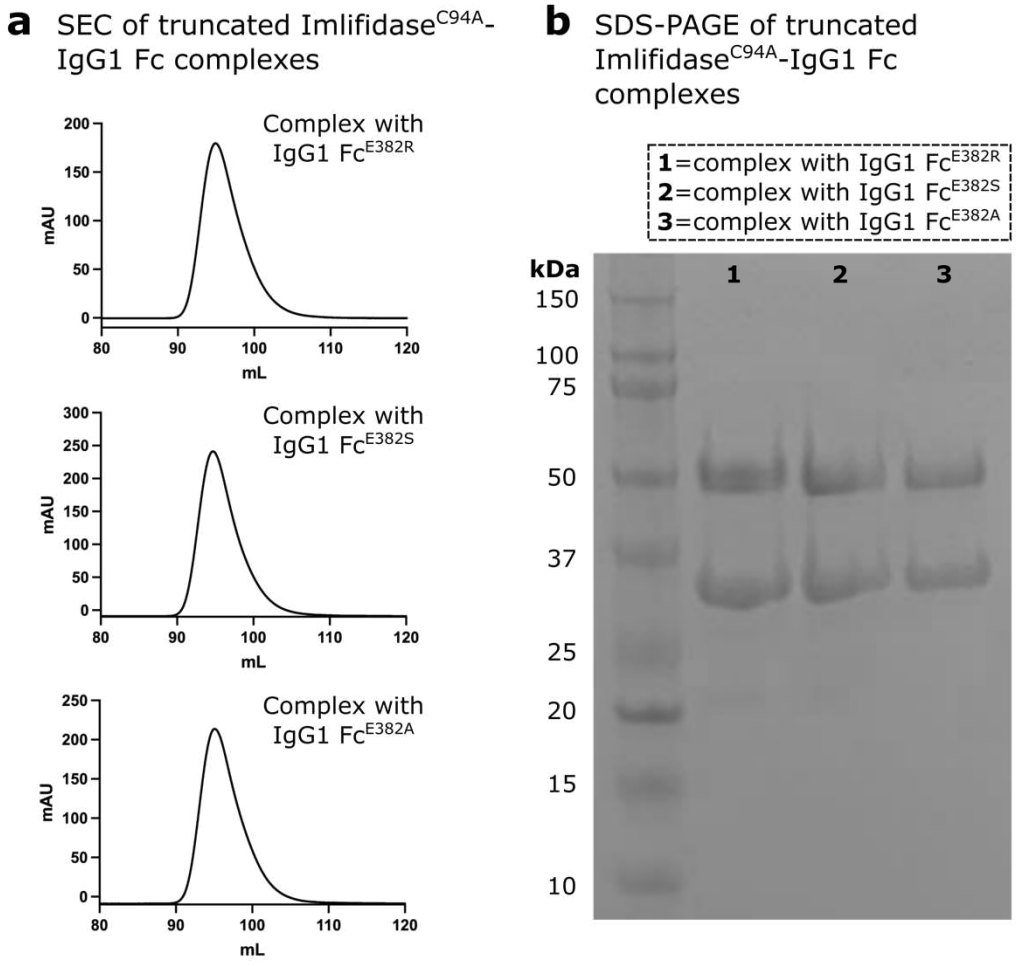


Figure 33: Purification of truncated Imlifidase^{C94A}-IgG1 Fc complexes. a SEC purification of complexes over a HiLoad® 16/600 Superdex 200 pg column. Fractions corresponding to the main peak of each complex were pooled for crystallisation. **b** SDS-PAGE of the purified complexes. Within the sample of each complex, dominant bands consistent with the molecular weights of truncated Imlifidase (~34 kDa) and IgG1 Fc (~50 kDa) are present. N.B. gel picture is missing the 250 kDa band in the protein ladder due to an error; no additional bands were present on the gel above this point.

SDS-PAGE analysis revealed that each of the complexes had been purified to homogeneity: within each sample, dominant bands are present which are consistent with the molecular weights of truncated Imlifidase (~34 kDa) and IgG1 Fc (~50 kDa) (**Figure 33b**). Samples for each of the Imlifidase-IgG1 Fc complexes were therefore taken forward for crystallisation experiments.

4.1.3 IdeS^{C94A}-asymmetric IgG1 Fc complex

IdeS cleaves each IgG Fc chain in two distinct reactions, in which the second chain is cleaved ~100-fold more slowly, as reported by Vindebro *et al.* [239]. In order to investigate the structural basis of how IdeS cleaves the second Fc chain, an "asymmetric IgG1 Fc" construct was designed to mimic an Fc that has been cleaved by IdeS on a single chain. As detailed in section 2.1.1.2, this asymmetric Fc was comprised of one "long" strand containing "knob" amino acid exchanges, and one "short" strand (starting with amino acid G237) containing "hole" amino acid exchanges [312]. Such knob and hole variants were included in order to encourage a stable complex forming between these distinct Fc chains, and discourage the formation of "knob-knob" or "hole-hole" Fc homodimers. E382R substitutions were also included to reduce Fc self-crystallisation.

The asymmetric Fc was expressed in HEK293-F cells and purified using Protein A affinity purification, followed by SEC purification (sections 2.2.1 and 2.3.1). Asymmetric Fc eluted from the SEC column as a single peak at ~77 mL, indicating that a stable species had formed. An "uncleaved" control "knob-hole" construct (i.e., both Fc chains start with amino acid 221, see section 2.1.1.2 for details) eluted slightly earlier, at ~76 mL, and as a single peak, indicating that this Fc heterodimer had also formed correctly (**Figure 34a**). SDS-PAGE analysis showed that, compared to the "uncleaved" Fc construct, the asymmetric Fc ran slightly further on the gel, consistent with a molecular weight reduction of ~0.7 kDa (achieved from IdeS cleavage of one chain of construct **1**; **Figure 34b**).

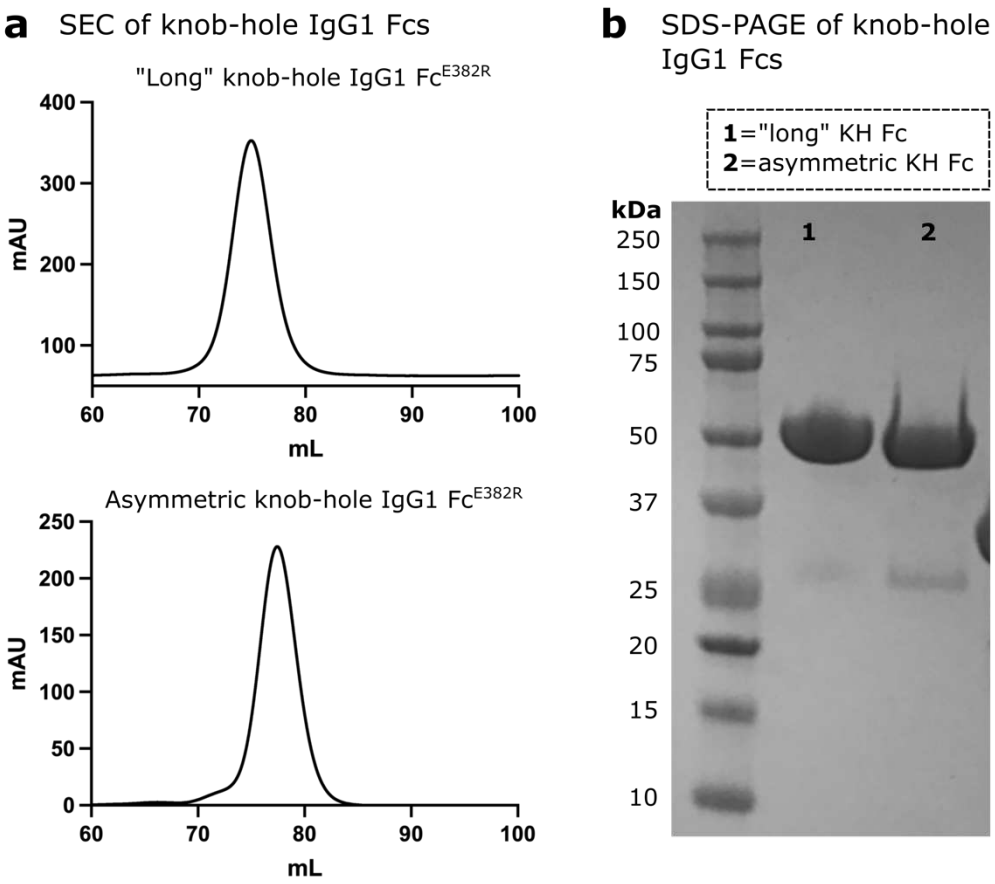


Figure 34: Purification of knob-hole (KH) IgG1 Fc^{E382R} constructs. a SEC purification of Fc constructs using a HiLoad® 16/600 Superdex 200 pg column. N.B. mAU was not zeroed correctly prior to purification of the "long" knob-hole Fc, therefore the baseline is lower than zero. **b** SDS-PAGE analysis of purified Fc constructs. Each Fc contains a dominant band at the approximate molecular weight (~49 kDa for the "long" Fc; ~48 kDa for the asymmetric Fc).

Following this purification, the asymmetric Fc was combined with IdeS^{C94A} and Imlifidase^{C94A} (purified using identical methods as described in section 2.3.2) in a 1:1 molar ratio and purified again using SEC (**Figure 35**). Each complex eluted as a single peak only, at ~99 mL (**Figure 35a**). SDS-PAGE analysis showed that each complex contained a dominant band consistent with the molecular weight of the asymmetric Fc (~48 kDa) and that of truncated IdeS/Imlifidase (~35 kDa and ~34 kDa, respectively). The asymmetric Fc-Imlifidase complex still contained a small amount of lower molecular weight impurities, however (**Figure 35b**).

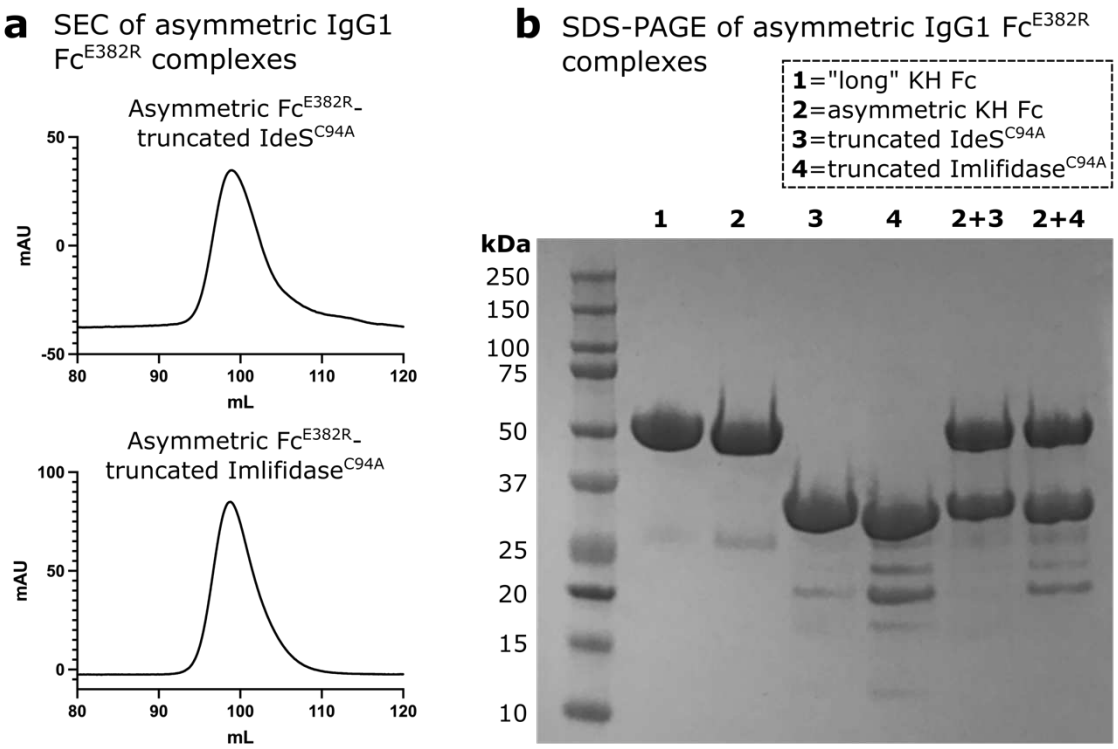


Figure 35: Purification of asymmetric IgG1 Fc^{E382R} complexes. a Asymmetric IgG1 Fc^{E382R} and truncated IdeS^{C94A}/imlifidase^{C94A} were combined in a 1:1 molar ratio and purified using SEC, yielding a single peak for each complex. N.B. mAU was not zeroed correctly prior to purification of the asymmetric Fc-IdeS complex, therefore the baseline is lower than zero. **b** SDS-PAGE. Lanes comprise, from left to right: "long" knob-hole IgG1 Fc^{E382R}, asymmetric IgG1 Fc^{E382R}, truncated IdeS^{C94A}, truncated Imlifidase^{C94A}, asymmetric Fc-IdeS complex, asymmetric Fc-Imlifidase complex. N.B. lanes **1** and **2** in **Figure 34** show the same SDS-PAGE gel as included here.

For each complex, fractions corresponding to the main peak were pooled and taken forward for crystallisation.

4.1.4 IdeS^{C94A} in complex with IgG2, IgG3 and IgG4 Fc

Along with IgG1, IdeS additionally cleaves IgG2, IgG3 and IgG4 antibodies [6]. The various IgG subclasses are similar in sequence (as discussed in section 1.1.2), but contain some notable differences, such as the length and disulphide patterning in the hinge region [24]. Compared to the other IgG subclasses, IgG2 antibodies contain a distinct sequence of P–VAGPS within the lower hinge (compared to ELLGGPS in IgG1; see **Figure 3**), and thus does not possess the G236–G237 motif cleaved by IdeS, which likely explains the decreased binding affinity of IdeS to this isotype [6]. In order to elucidate the structural basis behind the ability of IdeS to cleave IgG2, IgG3 and IgG4 antibodies, variants of these antibody Fc regions were designed (detailed in section 2.1.1.3), with the inclusion of E382R substitutions to reduce preferential Fc self-crystallisation.

IgG2 Fc^{E382R}, IgG3 Fc^{E382R} and IgG4 Fc^{E382R} fragments were expressed and purified as described in section 2.2.1 and 2.3.1. Unfortunately, expression yields for these constructs were relatively poor, and so SEC purification was not carried out, in order to retain enough protein for crystallisation

experiments. Truncated IdeS^{C94A} was expressed and purified (as described in sections 2.2.2 and 2.3.2) and then combined with each of the IgG Fc fragments in a 1:1 molar ratio. SDS-PAGE analysis of these complexes revealed the presence of dominant bands consistent with the molecular weights of IgG Fc (~50 kDa) and truncated IdeS^{C94A} (~35 kDa) within each sample (**Figure 36**). However, the IgG2 Fc-IdeS complex contained some faint bands for heavier molecular weight impurities or aggregation, and the IgG4 Fc-IdeS contained an additional band at approximately 30 kDa (**Figure 36**).

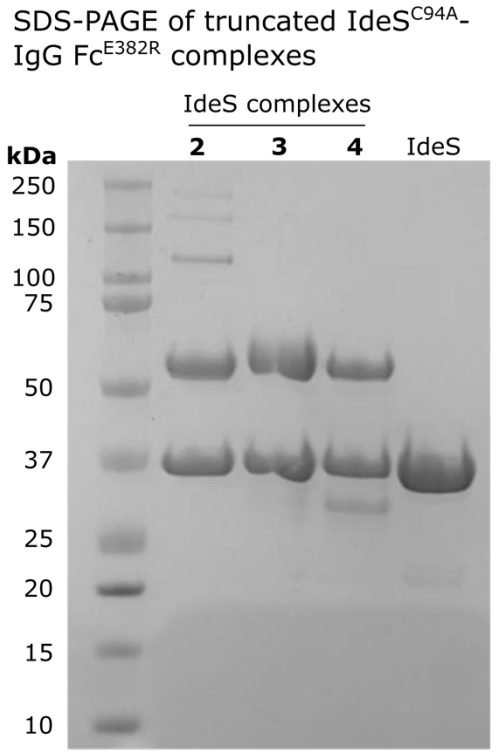


Figure 36: SDS-PAGE of truncated IdeS^{C94A}-IgG Fc^{E382R} complexes. Truncated IdeS^{C94A} was combined in a 1:1 molar ratio with IgG2 Fc^{E382R} (**2**), IgG3 Fc^{E382R} (**3**) and IgG4 Fc^{E382R} (**4**) and subsequently analysed by SDS-PAGE. Truncated IdeS^{C94A} alone was analysed also as a control. Complexes contain dominant bands consistent with the molecular weights of IgG Fc (~50 kDa) and truncated IdeS (~35 kDa).

Despite small amounts of contamination in these samples, each of the complexes was deemed largely pure, and so samples were taken forward for crystallisation experiments.

4.2 Crystallisation of IdeS-IgG1 Fc complexes

The panel of "less-crystallisable" IgG1 Fcs described in chapter 3 were used to attempt crystallisation of the truncated IdeS^{C94A}-IgG Fc complex. Following successful crystallisation of this complex, the strategy of rendering IgG Fc less crystallisable with substitution of amino acid E382 was applied to a broader range of IgG antibodies also cleaved by IdeS, and crystallisation of these complexes was attempted also. All crystals were grown at 21 °C.

4.2.1 IdeS^{C94A}-IgG1 Fc complex

Truncated IdeS^{C94A}-IgG1 Fc complexes were purified as described in section 2.4.1. Initial crystallisation trials yielded crystal hits in several conditions across the JCSG-plus™ and Morpheus screens (**Figure 37**). The crystals were vitrified and stored in liquid nitrogen prior to data collection, which was carried out on beamline ID30A-3 at the ESRF (Grenoble, France). Parameters used for data collection are detailed in **Table 6** (section 2.7.1).

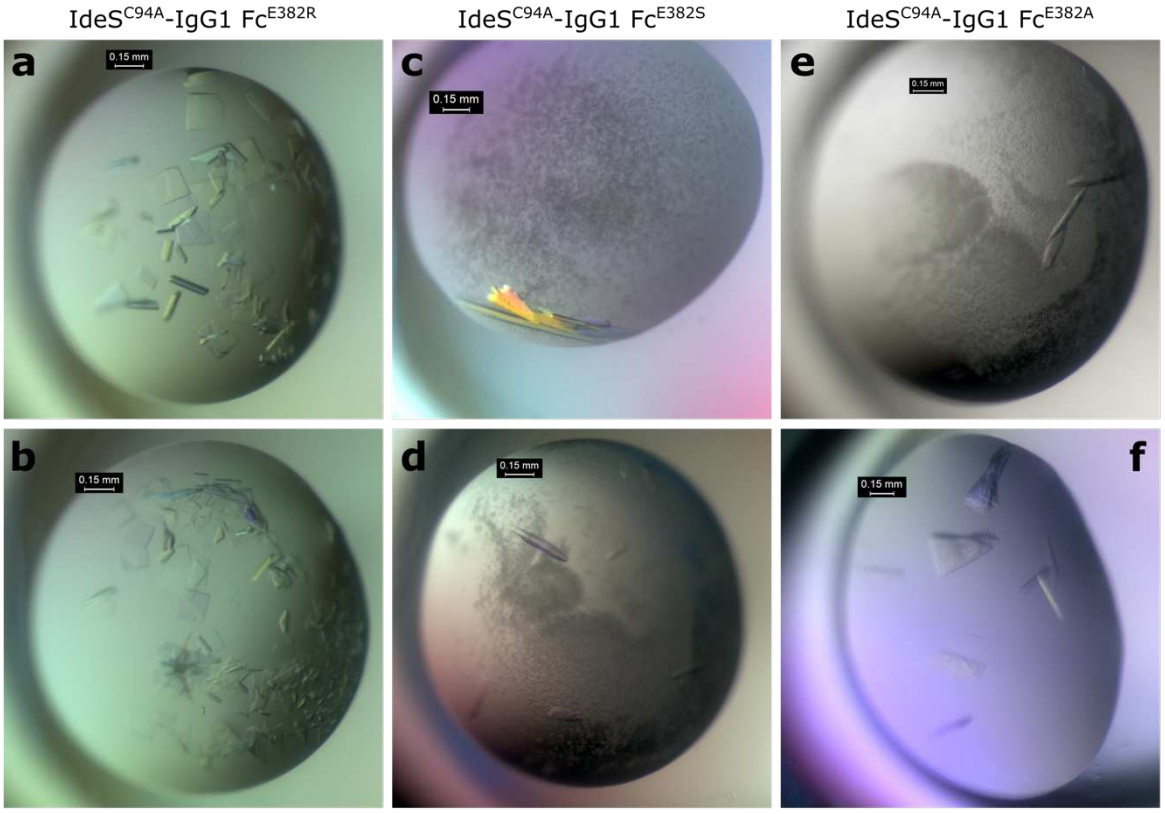


Figure 37: Crystallisation of truncated IdeS^{C94A}-IgG1 Fc complexes. Examples of crystals grown for each complex are pictured: IdeS^{C94A}-IgG1 Fc^{E382R} (**a**, Morpheus F9; **b**, Morpheus D9); IdeS^{C94A}-IgG1 Fc^{E382S} (**c**, JCSG-plus™ C2; **d**, Morpheus H9); IdeS^{C94A}-IgG1 Fc^{E382A} (**e**, Morpheus F9; **f**, Morpheus G8). All crystals were grown by sitting-drop vapour diffusion, at 21 °C.

Crystals grown in condition F9 from the Morpheus crystallisation screen (**Figure 37e**) diffracted to ~2.3 Å. Two datasets were collected from the same crystal and merged together using xia2.multiplex [318] in order to improve the completeness in the highest resolution shell. This merged dataset was scaled in space group *C*121 and truncated to a final resolution of 2.34 Å, to yield a CC_{half} value of >0.3 [388, 389] for the highest resolution shell. Data collection statistics for this complex are detailed in **Table 13** below.

Table 13: Crystallographic data collection statistics for truncated IdeS^{C94A}-IgG1 Fc^{E382A} complex.

Data Collection	**Truncated IdeS^{C94A}-IgG1 Fc^{E382A}**
Beamline	ID30A-3 (ESRF)
Resolution range (Å)	48.53–2.34 (2.38–2.34)[a]
Space group	*C* 1 2 1
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 217.56, 108.45, 63.05 90.00, 90.06, 90.00
Wavelength (Å)	0.9677
Unique reflections	60351 (3005)
Completeness (%)	97.7 (98.8)
R_{merge}	0.229 (0.877)
R_{meas}	0.225 (0.978)
R_{pim}	0.112 (0.426)
$I/\sigma(I)$	5.3 (1.0)
Multiplicity	4.96 (4.98)
CC_{half}	0.954 (0.335)
Wilson *B* factor (Å^2)	34.8

[a]Values for the highest resolution shell are shown in parentheses.

The structure of IdeS^{C94A}-IgG1 Fc^{E382A} was solved by molecular replacement. Two rounds of molecular replacement were required: firstly, the enzyme was fitted into the density using a search model for wild-type IdeS (PDB 1Y08). This solution was used as a fixed model for a second round of molecular replacement, in which a model of wild-type IgG1 Fc (PDB 3AVE) was searched for. The presence of twinning was detected in the data by an improved R_{free} upon refinement in Refmac5 with the option for twinning enabled (compared to an equivalent refinement job with default

parameters) [322]. Subsequent analysis of the data intensity statistics using *phenix.xtriage* within the *PHENIX* software suite [396] also indicated the presence of twinning (see Appendix E).

The twin fraction converged to 0.493 for operator *-h, -k, l* during refinement, which shows an almost perfect twin. Therefore, the potential for model bias is much greater, and care needs to be taken in analysing fine details of the structure. Automatically-defined NCS restraints were used during refinement, which also resulted in lower *R* factors. Final refinement statistics are detailed in **Table 14**.

Table 14: Crystallographic refinement statistics for IdeS^{C94A}-IgG1 Fc^{E382A} complex.

Refinement	**IdeS^{C94A}-IgG1 Fc^{E382A}**
Number of reflections (all/free)	60350/3049
R_{work} (%)	18.2
R_{free} (%)	20.6
Twinning fraction	0.507 for *h,k,l* 0.493 for *-h,-k,l*
RMSD[1]: Bonds (Å) Angles (°)	 0.0079 1.36
Molecules per ASU[2]	2
Atoms per ASU[2]	6,185
Average *B* factors (Å^2) (protein/ligand/water)	(38.48/45.71/38.81)
Model quality (Ramachandran plot): Most favoured region (%) Allowed region (%) Outliers (%)	 96.95 2.77 0.28, 0.1*

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
*as calculated within the full wwPDB X-ray Structure Validation Report

The final model is of good quality, with an R_{free} of 20.6 % and only a single Ramachandran outlier, as calculated within the wwPDB X-ray Structure Validation Report. As discussed in section 4.3, this outlier is believed to be "true", as it arises from a distortion in the IgG hinge backbone at the scissile peptide bond (*i.e.,* where the enzyme would cleave IgG if it were active). A *B* factor distribution reveals that the structure is generally well-ordered, with relatively higher *B* factors in the Fc Cγ2 domain bound by the enzyme (**Figure 38**), which is also discussed in section 4.3. The crystal structure reveals how the enzyme binds over this chain of the Fc, in order to encase the lower hinge region in a cavity containing the active site residues. The presence of the IgG hinge at this site is validated with clear electron density in this region (**Figure 38**), shown by the final density map and a polder omit map (calculated using *phenix.polder* module in the *PHENIX* software suite [396, 397]). This structure has been deposited within the PDB with code 8A47.

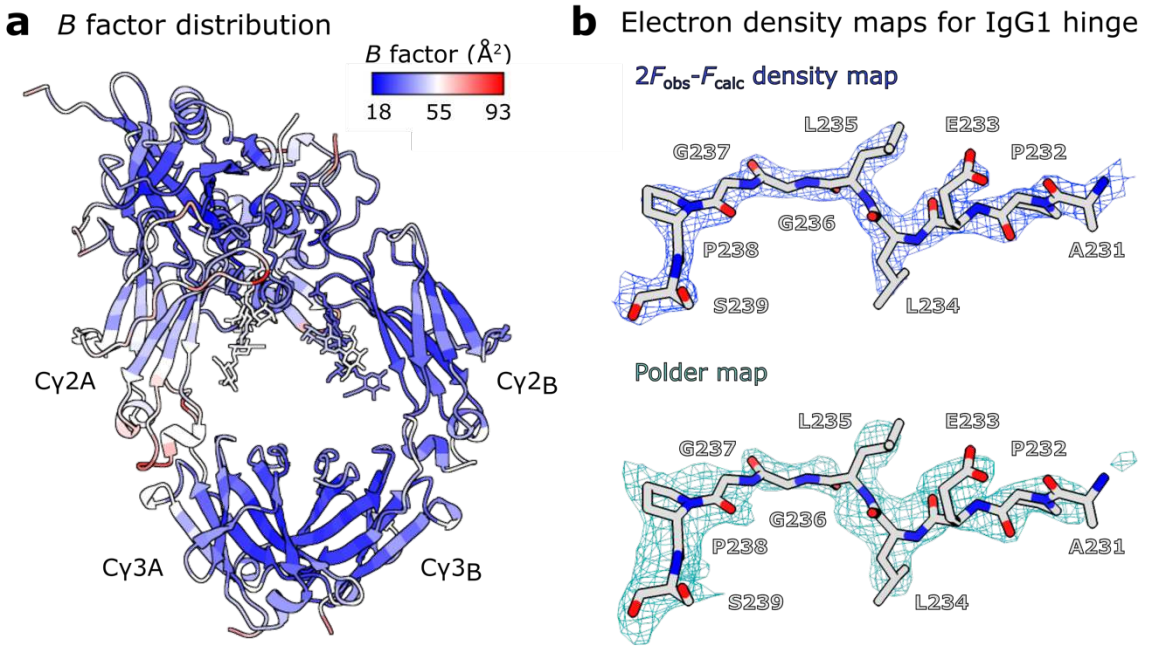


Figure 38: Analysis of IdeS^{C94A}-IgG1 Fc^{E382A} structure. a Distribution of average *B* factors per residue for IdeS^{C94A}-IgG1 Fc^{E382A} complex. **b** Final $2F_{obs}$–F_{calc} electron density map (blue, contoured at 1.5 σ), and polder map calculated in *PHENIX* [396, 397] (teal, contoured at 3 σ) for the IgG1 Fc^{E382A} hinge region (residues 231–239 in chain A) bound within the IdeS^{C94A} active site. The IgG1 Fc^{E382A} hinge peptide is depicted as sticks and coloured silver, with oxygen atoms coloured red and nitrogen atoms coloured blue.

A full analysis of this crystal structure is included within section 4.3.

4.2.2 Imlifidase^{C94A}-IgG1 Fc complexes

Complexes of truncated Imlifidase^{C94A}-IgG1 Fc were expressed and purified as detailed in section 4.1.2. Unfortunately, no crystals large enough for data collection were obtained from crystallisation experiments using JCSG-plus™ or Morpheus crystallisation screens.

4.2.3 IdeS^{C94A}-asymmetric IgG1 Fc^{E382R} complex

Complexes of the asymmetric IgG1 Fc^{E382R} (i.e., cleaved by IdeS on one chain) with truncated IdeS^{C94A} and truncated Imlifidase^{C94A} were expressed and purified as described in section 4.1.3. Some small, fragile crystals of the IdeS-asymmetric Fc complex grew in a few different conditions (**Figure 39**) from the JCSG-plus™ and Morpheus crystallisation screens. Most of these crystals were unsuitable for data collection due to their small size. An optimisation screen was also designed based on condition **b** from **Figure 39** (see Appendix D), in order to try and improve the crystals grown in this condition; however, the crystals grown in the initial screening could not be reproduced.

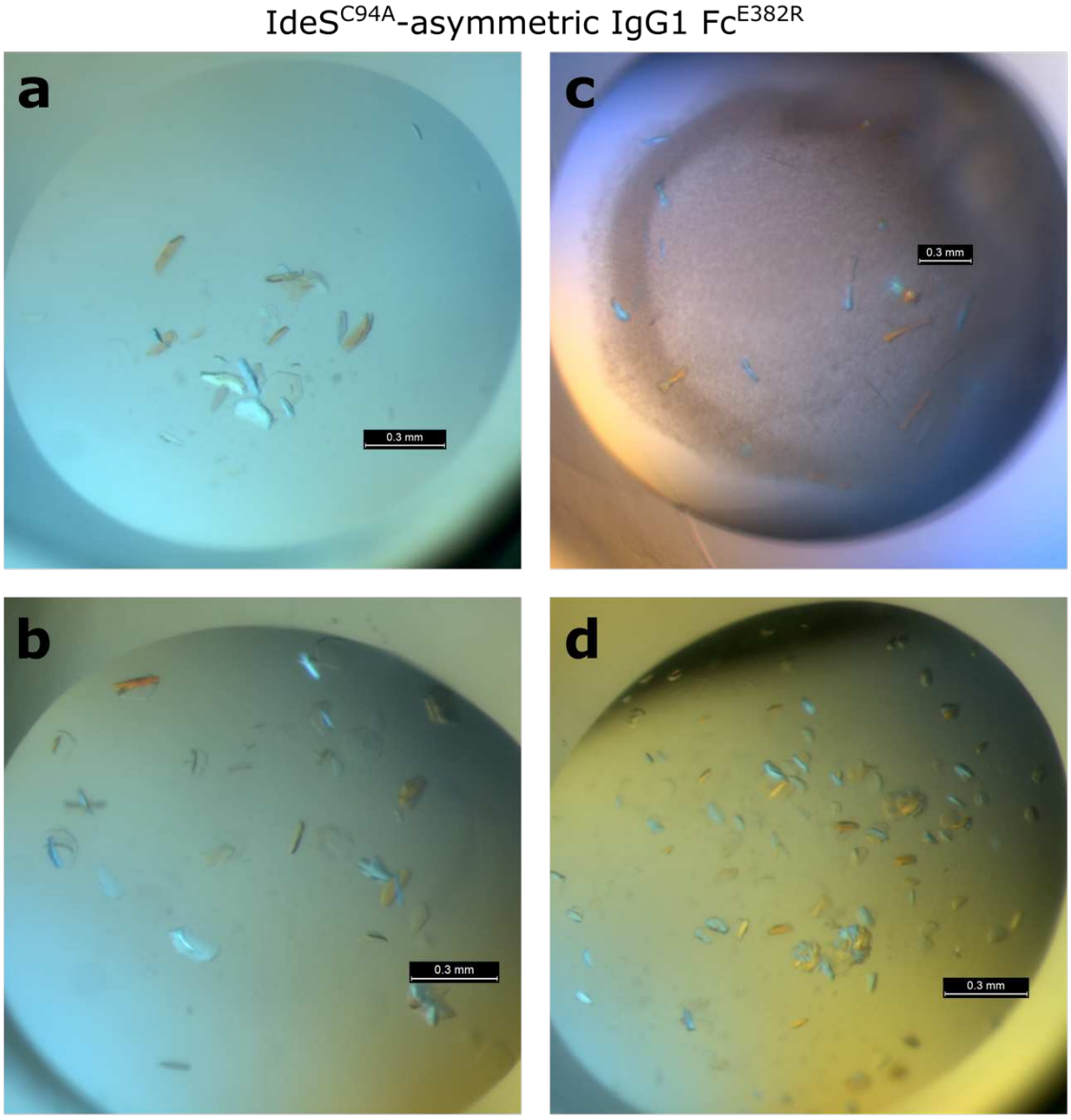


Figure 39: Crystallisation of truncated IdeS^{C94A} in complex with asymmetric IgG1 Fc^{E382R}. Crystals were grown in condition A5 from the Morpheus crystallisation screen (**a**) and conditions G4, G5 and A9 from the JCSG-plus™ screen (**b**, **c** and **d**, respectively).

Crystals grown from condition **c** in **Figure 39** (from subsequent purifications) were harvested and taken for data collection on beamline I04 at Diamond Light Source. Data collection was carried out using parameters detailed in **Table 15**.

Table 15: Data collection parameters for X-ray diffraction of truncated IdeS^{C94A}-asymmetric IgG1 Fc^{E382R} complex.

Parameter	**Truncated IdeS^{C94A}-asymmetric IgG1 Fc^{E382R} complex**
Beamline	I04 (DLS)
Wavelength (Å)	0.9795
Flux	2.82e+11
Transmission (%)	100
Exposure (s)	0.0137
Beam size (μm)	30×20
Oscillation angle (°)	0.1
Number of images	3600

A low-resolution data set (diffraction to ~3.5 Å) was obtained from a single crystal. Diffraction images were processed using DIALS and cut to a final resolution of 3.45 Å. The data were scaled in space group $P6_1$. Data collection statistics for this complex are detailed in **Table 16** below.

Table 16: Crystallographic data collection statistics for truncated IdeS^{C94A}-asymmetric IgG1 Fc^{E382R} complex.

Data Collection	**Truncated IdeS^{C94A}-asymmetric IgG1 Fc^{E382R}**
Beamline	I04 (DLS)
Resolution range (Å)	48.91–3.45 (3.51–3.45)[a]
Space group	$P6_1$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 97.8, 97.8, 103.2 90, 90, 120
Wavelength (Å)	0.9795
Unique reflections	7,459 (358)
Completeness (%)	100 (100)
R_{merge}	0.145 (3.661)
R_{meas}	0.148 (3.749)
R_{pim}	0.033 (0.806)
$I/\sigma(I)$	13.5 (0.4)
Multiplicity	20.9 (21.6)
CC_{half}	1.000 (0.400)
Wilson *B* factor (Å^2)	108.4

[a]Values for the highest resolution shell are shown in parentheses.

The structure was solved by molecular replacement, using PDB entry 3AVE (wild-type IgG1 Fc) as a search model. However, it was discovered that the asymmetric Fc has crystallised on its own, and no density consistent with an additional protein bound was observed. Therefore, no further structure solution of this dataset was pursued.

4.2.4 IdeS^{C94A} in complex with IgG2, IgG3 and IgG4 Fc

Complexes of truncated IdeS^{C94A} with IgG2 Fc^{E382R}, IgG3 Fc^{E382R} and IgG4 Fc^{E382R} were expressed and purified as detailed in section 4.1.4. Initial crystallisation experiments were carried out using Morpheus, JCSG-plus™ and Proplex™ screens. Crystals grew in several conditions across these screens, for the complexes with IgG2 Fc^{E382R} and IgG4 Fc^{E382R}, as shown in **Figure 40**. No crystals were obtained for the complex with IgG3 Fc^{E382R}. Crystals large enough for data collection were harvested and taken to Diamond Light Source.

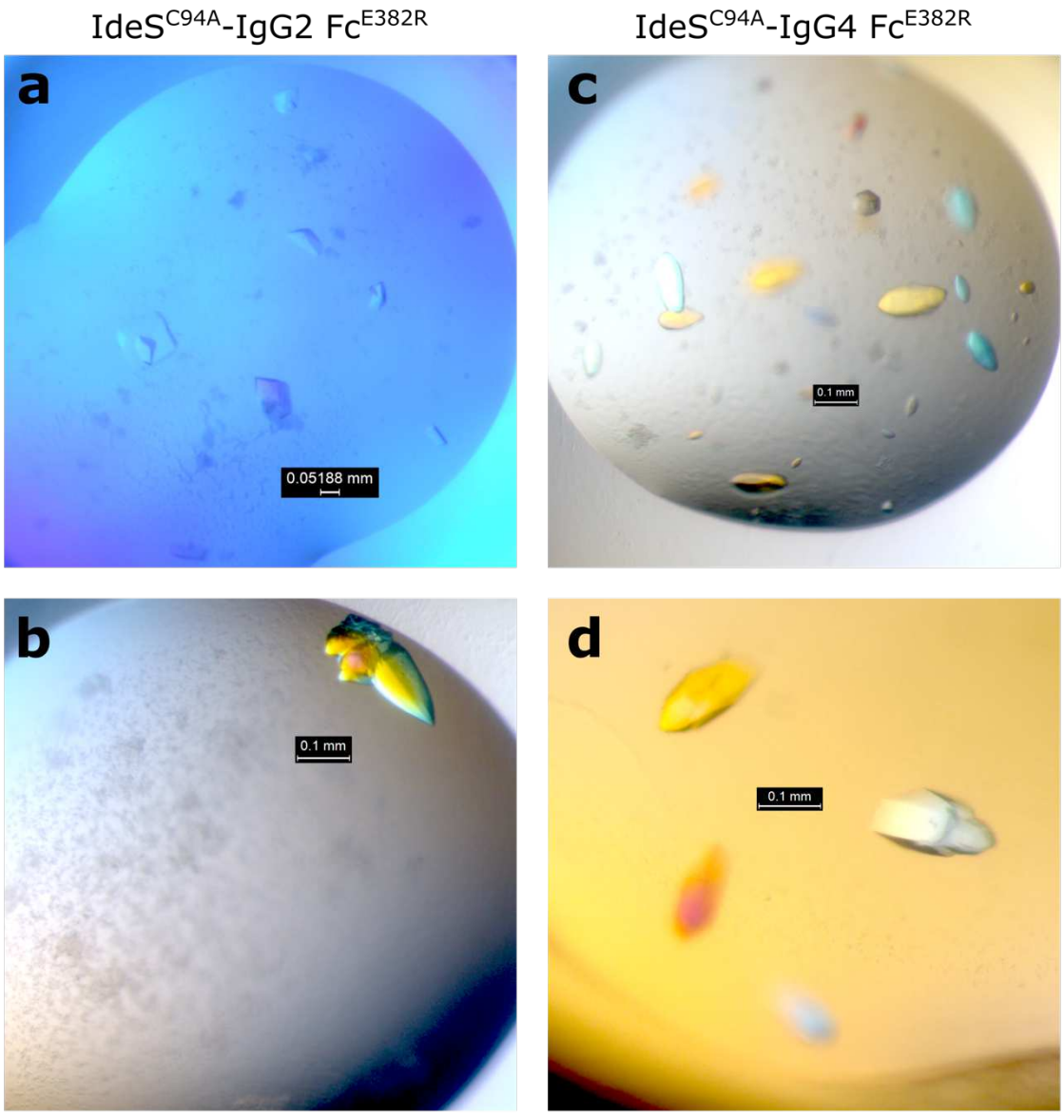


Figure 40: Crystals of truncated IdeS^{C94A} in complex with IgG2 Fc^{E382R} and IgG4 Fc^{E382R}. Pictured crystals grew in the following conditions: Proplex™ D11 (**a**); JCSG-plus™ B2 (**b**); JCSG-plus™ A5 (**c**); Morpheus B12 (**d**).

Crystals for the IdeS^{C94A}-IgG4 Fc^{E382R} complex typically diffracted to 3.5 Å or lower. However, a 2.6 Å dataset was obtained from a crystal grown in condition H9 from the Morpheus crystallisation screen. Additionally, a ~2 Å dataset was obtained for the IdeS^{C94A}-IgG2 Fc^{E382R} crystals, which were

grown in condition A5 from the JCSG-plus™ screen (**Figure 40b**). These data were collected on beamline I03 at Diamond Light Source, using parameters detailed in **Table 17**.

Table 17: Data collection parameters for X-ray diffraction of truncated IdeS^{C94A} in complex with IgG4 Fc^{E382R} and IgG2 Fc^{E382R}.

Parameter	**Truncated IdeS^{C94A}-IgG4 Fc^{E382R} complex**	**Truncated IdeS^{C94A}-IgG2 Fc^{E382R} complex**
Beamline	I03 (DLS)	I03 (DLS)
Wavelength (Å)	0.9763	0.9763
Flux	6.51e+11	8.84e+11
Transmission (%)	10	10
Exposure (s)	0.0091	0.0070
Beam size (μm)	80×20	80×20
Oscillation angle (°)	0.1	0.1
Number of images	3600	3600

For the IdeS^{C94A}-IgG4 Fc^{E382R} complex, merged reflections were downloaded from the xia2/DIALS automatic processing pipeline [315, 316] were downloaded and used for structure solution. The data were scaled in space group $P6_1$ and truncated to a resolution of 2.6 Å. Diffraction images from the IdeS^{C94A}-IgG2 Fc^{E382R} complex were processed using DIALS and also scaled in space group $P6_1$. This dataset was truncated to a resolution of 1.95 Å. Data collection statistics for both datasets are detailed in **Table 18** below.

Table 18: Crystallographic data collection statistics for truncated IdeS^{C94A}-IgG4 Fc^{E382R} and -IgG2 Fc^{E382R} complexes.

Data Collection	**Truncated IdeS^{C94A}-IgG4 Fc^{E382R}**	**Truncated IdeS^{C94A}-IgG2 Fc^{E382R}**
Beamline	I03 (DLS)	I03 (DLS)
Resolution range (Å)	69.53–2.60 (2.65–2.60)	49.12–1.95 (1.98–1.95)[a]
Space group	$P6_1$	$P6_1$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 109.8, 109.8, 101.9 90, 90, 120	 98.2, 98.2, 100.6 90, 90, 120
Wavelength (Å)	0.9763	0.9763
Unique reflections	21,510 (1049)	40,041 (1922)
Completeness (%)	100 (100)	99.7 (96.7)
R_{merge}	0.080 (4.402)	0.136 (1.707)
R_{meas}	0.082 (4.507)	0.140 (1.796)
R_{pim}	0.018 (0.966)	0.031 (0.545)
$I/\sigma(I)$	22.0 (0.3)	10.8 (0.4)
Multiplicity	21.0 (21.6)	19.0 (10.5)
CC_{half}	1.000 (0.252)	0.998 (0.801)
Wilson *B* factor (Å^2)	87.05	31.58

[a]Values for the highest resolution shell are shown in parentheses.

The structures were solved by molecular replacement, using PDB entries 4C54 (human IgG4 Fc) and 4HAG (human IgG2 Fc) as search models. However, the structures revealed that each of the antibodies had crystallised on their own (i.e., not in complex with the enzyme). Structure solution was therefore not pursued to completion; however, crystal packing interactions were analysed in order to determine amino acids involved in Fc self-crystallisation.

IgG2 Fc^{E382R}

Interface analysis revealed that the majority of crystal packing interactions within the *P*6$_1$ cell occur between one Fc molecule and a symmetry-related molecule at (x−y+1, x+1, z+1/6). The interface has a total surface area of 1071.8 Å^2, as identified by PDBePISA [85]. This analysis revealed that the exchanged R382 in IgG2 Fc^{E382R} is involved in the crystal contacts at this interface, in both Fc chains. **Figure 41** shows the packing interactions present at this interface: in one Fc molecule, R382 forms interactions with the backbone and side chain of H433. R382 in the second Fc molecule interacts with the backbone of Y436 and N434. The nearby residue E380 also interacts with N434 (**Figure 41b**).

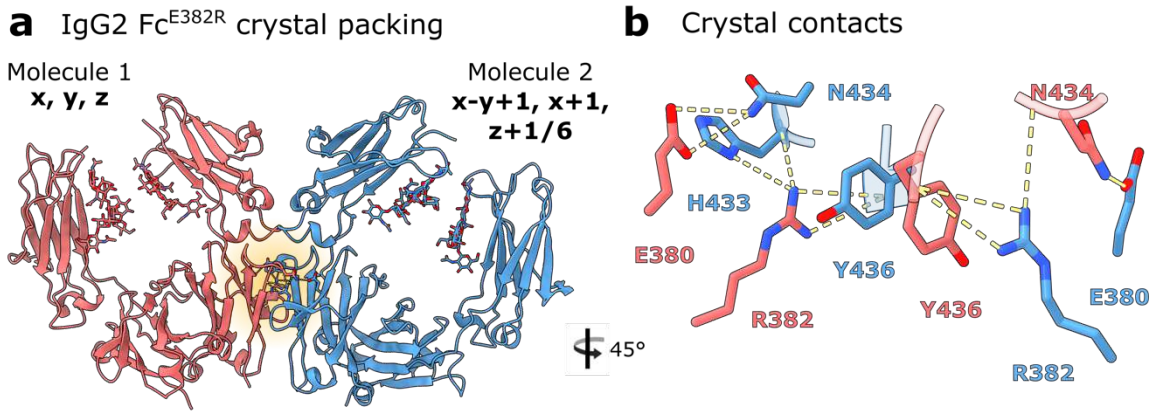


Figure 41: Analysis of crystal packing interactions within IgG2 Fc^{E382R} structure solved in *P*6$_1$. **a** Crystal packing between molecule 1 (at x, y, z) and a symmetry-related molecule (at x−y+1, x+1, z+1/6) within the crystal. **b** Interactions present within the packing interface between these two molecules. Exchanged amino acid R382 (E382 in the wild-type protein) forms several interactions at this interface. Hydrogen bonds are depicted as yellow dashed lines.

This analysis indicates that R382 and E380 side chains could be modified, such as to an alanine, in order to render the Fc less crystallisable, which could be utilised in future attempts to crystallise the IgG2 Fc complexes.

IgG4 Fc^{E382R}

Interface analysis of the IgG4 Fc^{E382R} crystal in *P*6$_1$ revealed the largest crystal packing interface to be between a molecule at (x, y, z) and a symmetry-related molecule at (x−y, x, z+1/6). This interface has an interface of 995.9 Å^2, as identified by PDBePISA [85]. As also observed for the IgG2 Fc^{E382R} crystal, R382 and E380 side chains are involved in the crystal packing interactions at this interface. As depicted in **Figure 42**, R382 on one chain forms hydrogen bonds with the backbone of N434 and Y436, while that of the other chain forms hydrogen bonds with the H433 and Y436 backbone. E380 on both Fc chains forms a hydrogen bond with N434 side chain and a salt bridge with the H433 side chain.

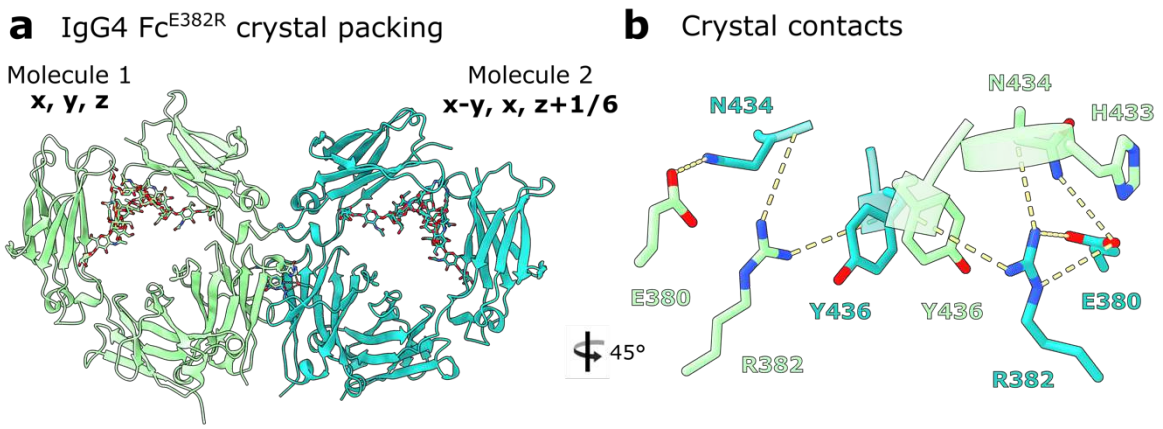


Figure 42: Analysis of crystal packing interactions within IgG4 Fc^{E382R} structure solved in *P*6$_1$. **a** Crystal packing between molecule 1 (at x, y, z) and a symmetry-related molecule (at x−y, x, z+1/6) within the crystal. **b** Interactions formed by E380 and exchanged side chain R382 within the packing interface between these two molecules. Hydrogen bonds are depicted as yellow dashed lines.

As for the IgG2 Fc^{E382R} crystal, the exchange of amino acid E382 to an arginine has likely not significantly hindered crystal packing. Therefore, future attempts to crystallise IgG4 Fc in complex with another protein could instead be attempted with a construct utilising E380A and/or E382A exchanges, in order to prevent the crystal packing interactions observed here.

4.3 Structural basis of IgG1 Fc recognition by IdeS

Prior to the start of this project, several crystal structures of wild-type and inactive IdeS had been solved [246, 247]. The enzyme was known to be highly efficient and highly specific for IgG [6, 244]; however, it was largely unknown exactly how IdeS is able to recognise its sole substrate. This section describes the analysis of the inactive IdeS-IgG1 Fc crystal structure.

Electron density resolves amino acids 43–339 in IdeS, as well as 229–445 and 230–444 for chains A and B in IgG1 Fc, respectively. Density is also observed for seven/eight monosaccharide residues at the *N*-linked glycosylation site (at N297) on Fc chains A and B, comprising a fucosylated biantennary glycan with a single β-1,2-linked GlcNAc on the mannose 6-arm (chain A) and the equivalent glycan with terminal β-1,2-linked GlcNAc on both arms (chain B). The final structure was refined to 2.34 Å (**Table 14**, section 4.2.1) and is depicted in **Figure 43**.

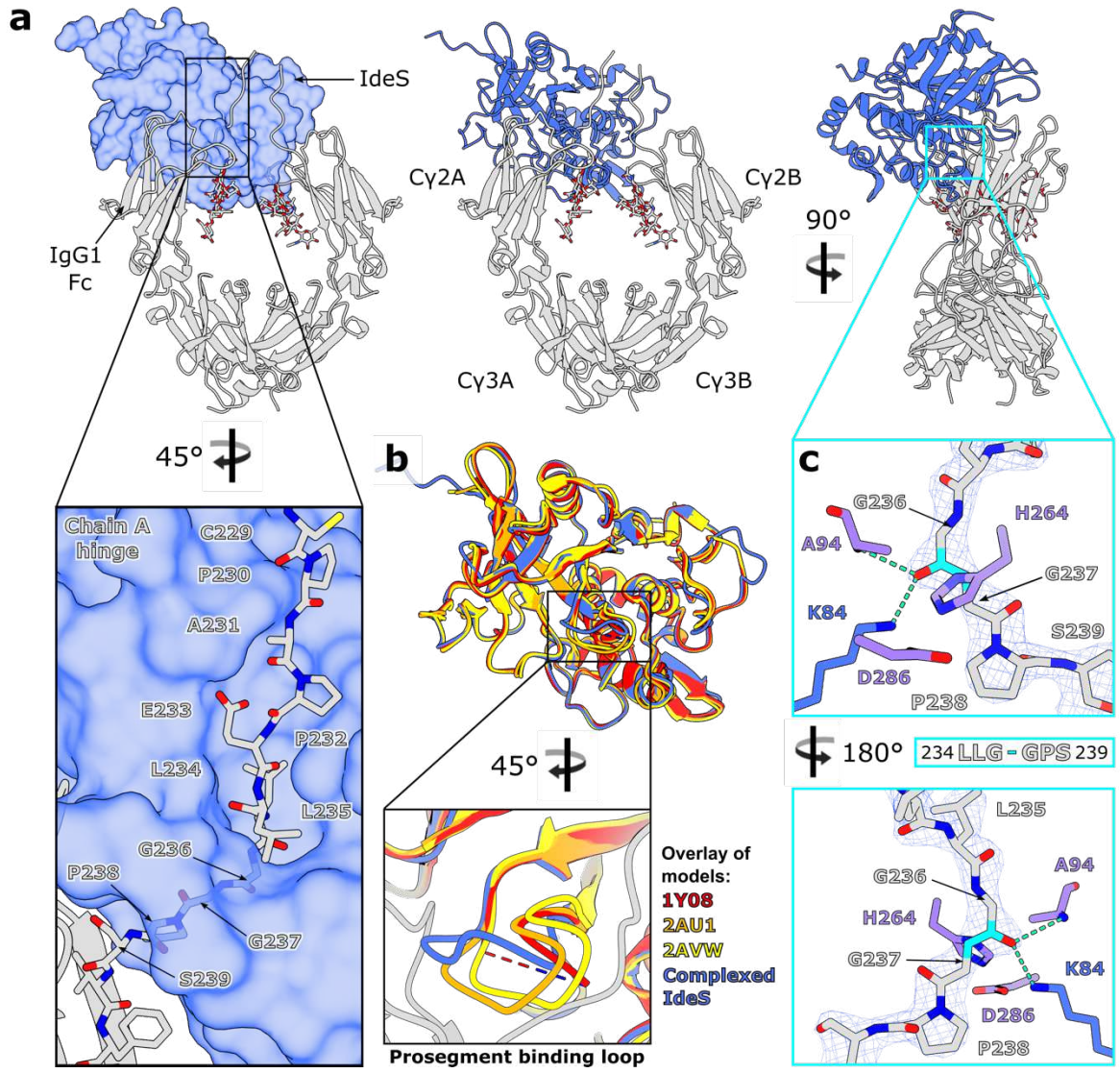


Figure 43: Crystal structure of IdeS^{C94A}-IgG1 Fc^{E382A} complex. a Overall structure of complex, with IdeS^{C94A} shown as a surface and IgG1 Fc^{E382A} shown as a cartoon. *N*-linked glycans and the IgG1 Fc^{E382A} hinge peptide in the focused panel are shown as sticks and coloured by heteroatom (oxygen, red; nitrogen, blue; sulphur, yellow). **b** Superposition of complexed IdeS^{C94A} with three published apo structures of IdeS (PDB codes 1Y08, 2AU1 and 2AVW, coloured in red, orange and yellow, respectively) and focused view of the prosegment binding loop. **c** Binding of IgG1 Fc^{E382A} hinge peptide within the IdeS^{C94A} active site. Fc^{E382A} peptide and IdeS^{C94A} active site residues are depicted as sticks and coloured by heteroatom; catalytic triad residues are coloured purple. The scissile peptide bond is coloured in cyan; hydrogen bonds are depicted as green dashes. The final $2F_{obs}–F_{calc}$ electron density map corresponding to the Fc^{E382A} peptide is shown (contoured at 1.5 σ). **a, b, c** IdeS^{C94A} is coloured blue; IgG1 Fc^{E382A} is coloured in silver.

The crystal structure shows asymmetric binding of IdeS across the Cγ2 domains of the Fc and its lower hinge region (**Figure 43a**). It appears that the upper hinge region of IgG and its Fab regions do not contribute significantly to complex formation, as indicated with the lack of electron density for the hinge above residues 229/230 (in Fc chains A and B, respectively), and reported cleavage of both full-length IgG and its Fc fragment by IdeS [244]. However, it cannot be formally excluded from this crystal structure alone that there is some interaction of IdeS with the IgG Fab regions.

The 1:1 stoichiometry observed in the crystal structure is consistent with previous kinetic analyses [239] showing that IdeS functions predominantly in a monomeric form. The enzyme appears to clamp down over the lower hinge region of one Fc chain (**Figure 43a**), creating a cavity in which the catalytic residues are brought into close proximity with the cleavage site. Binding of the enzyme to the Fc appears to displace the two Cγ2 domains slightly, as shown by superposition with a structure of wild-type IgG1 Fc (PDB code 3AVE) (**Figure 44b**). Residues within the Cγ2 domain in chain A have higher *B* factors compared to the rest of the complex (**Figure 44a**), which could indicate that binding of IdeS pulls this domain away slightly from the rest of the antibody.

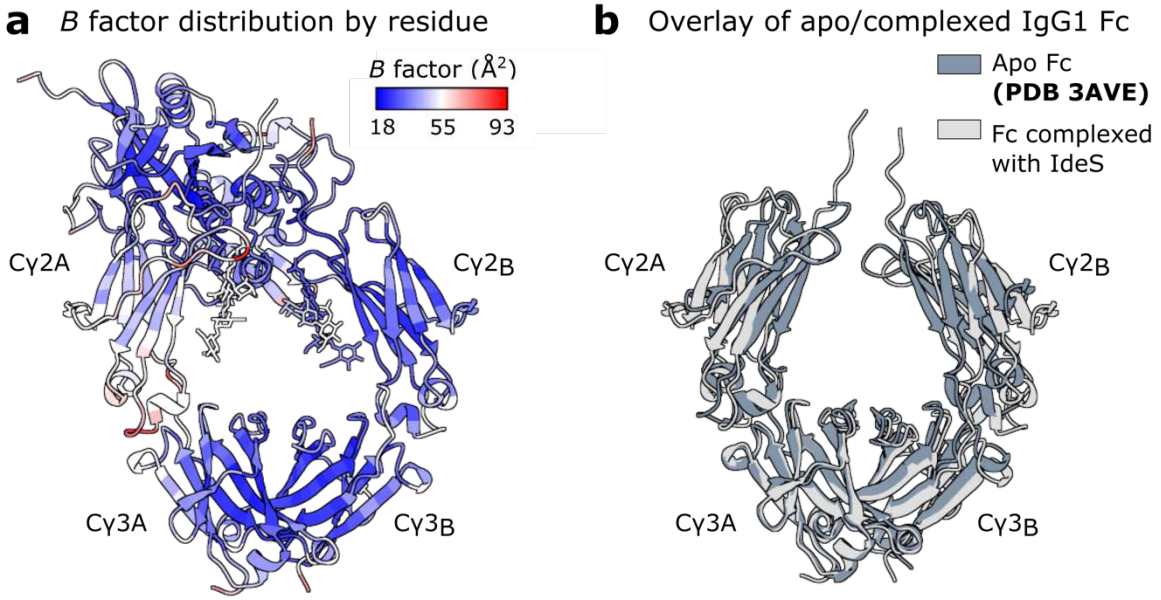


Figure 44: Analysis of IdeS^{C94A}-IgG1 Fc^{E382A} crystal structure. a Distribution of average *B* factors per residue for IdeS^{C94A}-IgG1 Fc^{E382A} complex. **b** Superposition of complexed IgG1 Fc^{E382A} (depicted in silver) with wild-type, apo IgG1 Fc (PDB code 3AVE; depicted in dark grey).

Role of prosegment binding loop in IdeS-Fc complex

IdeS crystallised in complex with IgG1 Fc here is the Mac-2 variant, and thus deviates in sequence from the three published apo structures of IdeS (all of which are the Mac-1 variant; **Figure 45**). Despite this, a structural alignment shows very few deviations (**Figure 43b**).

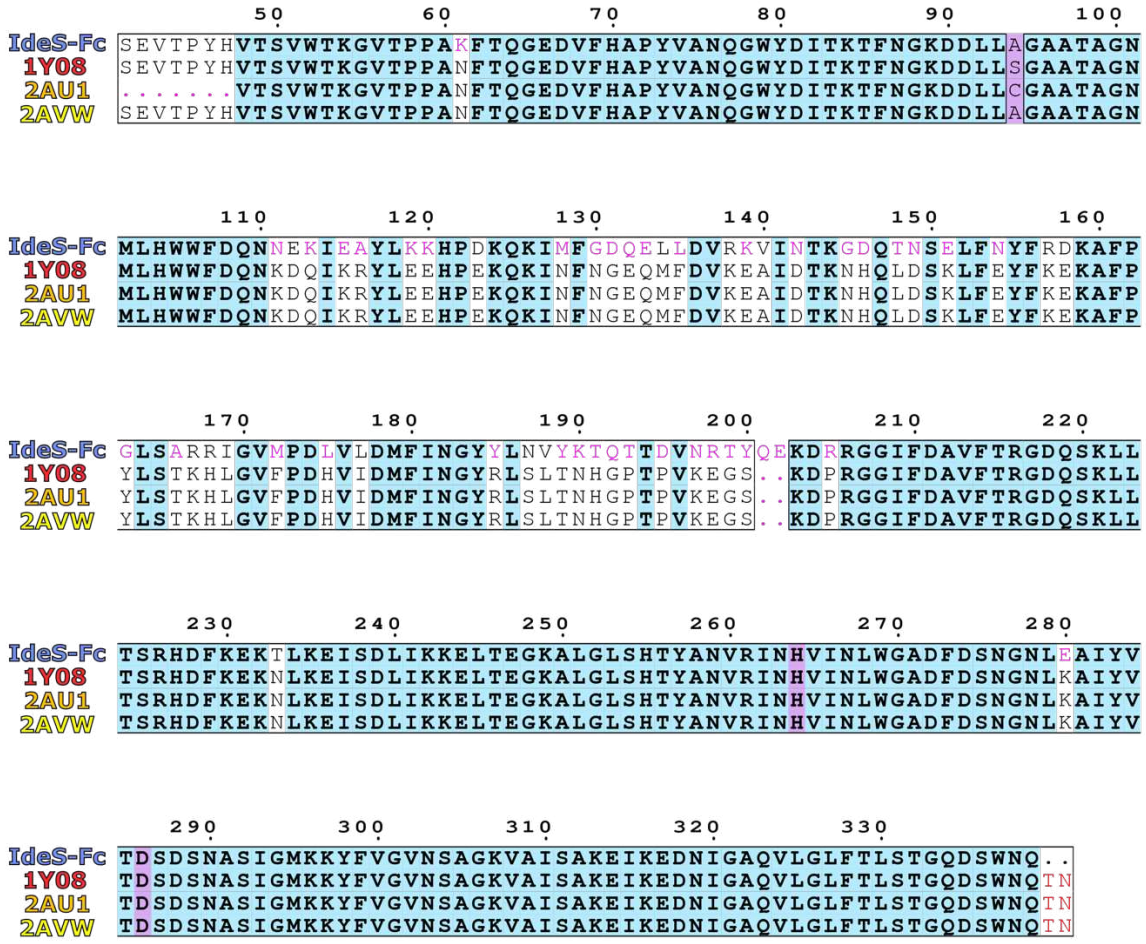


Figure 45: Multiple sequence alignment of complexed IdeS^{C94A} vs published apo IdeS structures (PDB codes 1Y08, 2AU1 and 2AVW; all Mac-1 variants). Alignment was generated in Clustal Omega [35] and depicted using ESPript 3.0 [311]. Conserved residues are coloured blue; catalytic triad residues are highlighted in purple; similar residues are shown in black font; residues with differing chemical properties are coloured in pink font. Numbering is shown for the complexed form of IdeS^{C94A} (Mac-2).

Complexed IdeS contains ten α-helices and twelve β-strands, as calculated by DSSP [398, 399] (**Figure 46**). Inspection of the structural superposition reveals that the loop located between β-strands seven and eight is modelled in distinct conformations for each of the apo structures [246] and is not included within 1Y08 [247] (**Figure 43b**), signifying its inherent flexibility in the apo form. This loop is equivalent to the "prosegment binding loop" present in other papain superfamily cysteine proteases; in these enzymes, which are synthesised as inactive zymogens, this loop packs against the prosegment as a mode of inhibition [400-402]. In complexed IdeS, the loop curls upwards to accommodate the Fc hinge within the active site cavity (**Figure 43a**).

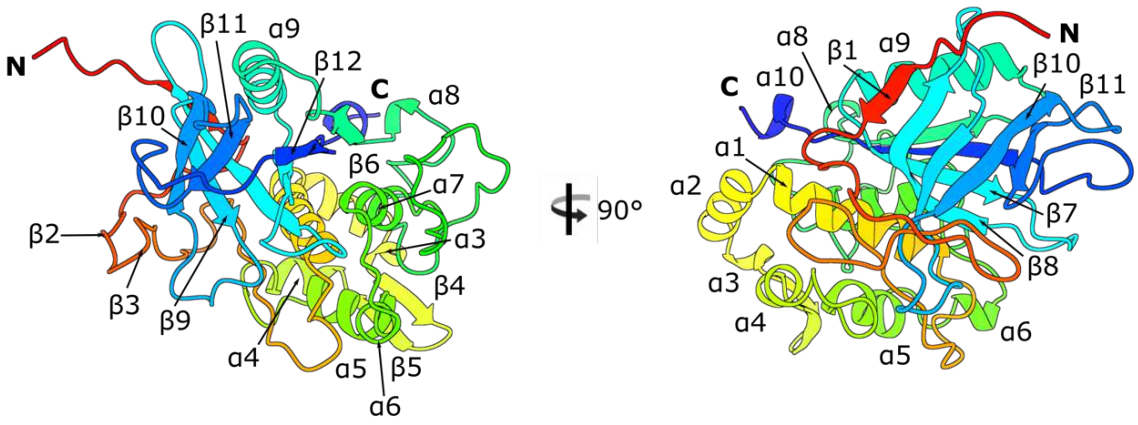


Figure 46: Secondary structure of complexed IdeS^{C94A}. IdeS is coloured as a rainbow from the N- (red) to C-terminus (blue) and labelled with secondary structure, as calculated by DSSP [398, 399].

Alanine substitution variations within this loop were previously found to have little effect on both IdeS binding to IgG, and its catalytic activity [246]. The IdeS-Fc crystal structure shows, however, that the majority of interactions present here involve the IdeS backbone, whose conformation won't be significantly altered by alanine substitutions. The inability of IdeS to cleave IgG hinge-mimicking peptides [244] also indicates an occlusion of the active site in the absence of substrate, especially given the strong potential of hydrogen bonding and hydrophobic interactions observed at the Fc hinge (discussed in the following section). It is possible that a conformational change in the active site could occur upon Fc binding; however, superposition of wild-type IdeS (PDB code 2AU1) with the complexed enzyme shows the catalytic triad residues in very similar conformations, although this cannot be ruled out given the sequence diversity present (**Figure 47**). Moreover, IdeS^{C94A} has been shown to retain antibody binding and inhibit IgG-mediated phagocytosis at levels comparable to the wild-type enzyme [236], suggesting that the inactive enzyme retains antibody binding. Therefore, the likely role of this loop is mediation of substrate access to the active site.

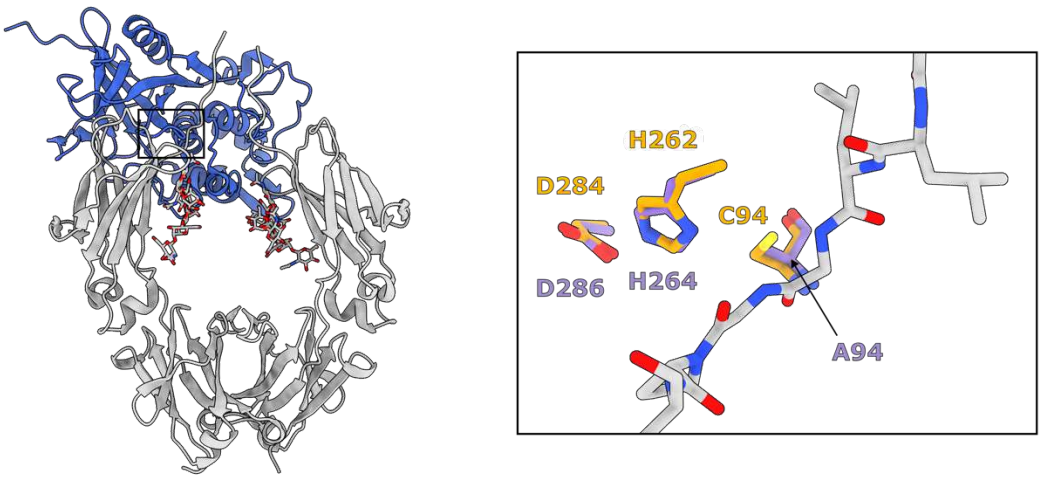


Figure 47: Superposition of catalytic residues of complexed IdeS^{C94A} vs apo wild-type IdeS. IgG1 Fc^{E382A} is coloured silver; *N*-linked glycans and residues in zoom panel are coloured by heteroatom (oxygen, red; nitrogen, blue). IdeS^{C94A} is coloured blue, catalytic residues in zoom panel are coloured purple. Wild-type IdeS (PDB 2AU1) is coloured orange.

Interaction of IgG1 Fc hinge at IdeS active site

Clear density is observed for the IgG1 Fc hinge region bound within the IdeS active site cavity, in both the final electron density map (**Figure 43c**) and a polder map as calculated in *PHENIX* [396, 397] (**Figure 38b**). The carbonyl oxygen of G236 forms a hydrogen bond with the amide nitrogen of the catalytic cysteine (exchanged to alanine here) and the side chain of K84, which collectively form the oxyanion hole, as predicted [246, 247]. Binding of the hinge distorts the Fc peptide backbone at G236 in order to promote scissile bond cleavage (**Figure 43c**); this residue is thus identified in Molprobity [325] as a Ramachandran outlier. Superposition of wild-type IdeS (PDB code 2AU1) with the complexed enzyme gives an indication for placement of the catalytic cysteine side chain (**Figure 47**): in this conformation, the cysteine sulphur is ideally poised for nucleophilic attack on the carbonyl carbon within the scissile peptide bond.

Extended exosite binding of IdeS to the Fc Cγ2 domains

It has long been suspected that IdeS must recognise its sole substrate IgG with exosite binding [244, 246, 247]. The IdeS-Fc crystal structure now reveals that IdeS binds across both chains of the Fc region (**Figure 48a**). Unsurprisingly, the most extensive interface is formed with the Fc chain being cleaved (annotated as chain A in the crystal structure) (**Figure 48b**), with an interface area of 1392 Å^2 and a solvation free energy gain upon interface formation of -16.2 kcal/mol, as calculated by PDBePISA [85]. The interface extends across the entire hinge region (C229–S239; **Figure 48b**), with hydrogen bonds formed with the backbone at A231, L234, G236 and G237 and the E233 side chain, and favourable hydrophobic interactions predicted here (inferred by positive solvation energies of hinge residues). Within the Fc Cγ2 domain, IdeS interacts with residues in proximity of

the BC loop, which aids in stabilising an "open" conformation of the prosegment binding loop (as discussed above), and additionally the FG loop (**Figure 48b**).

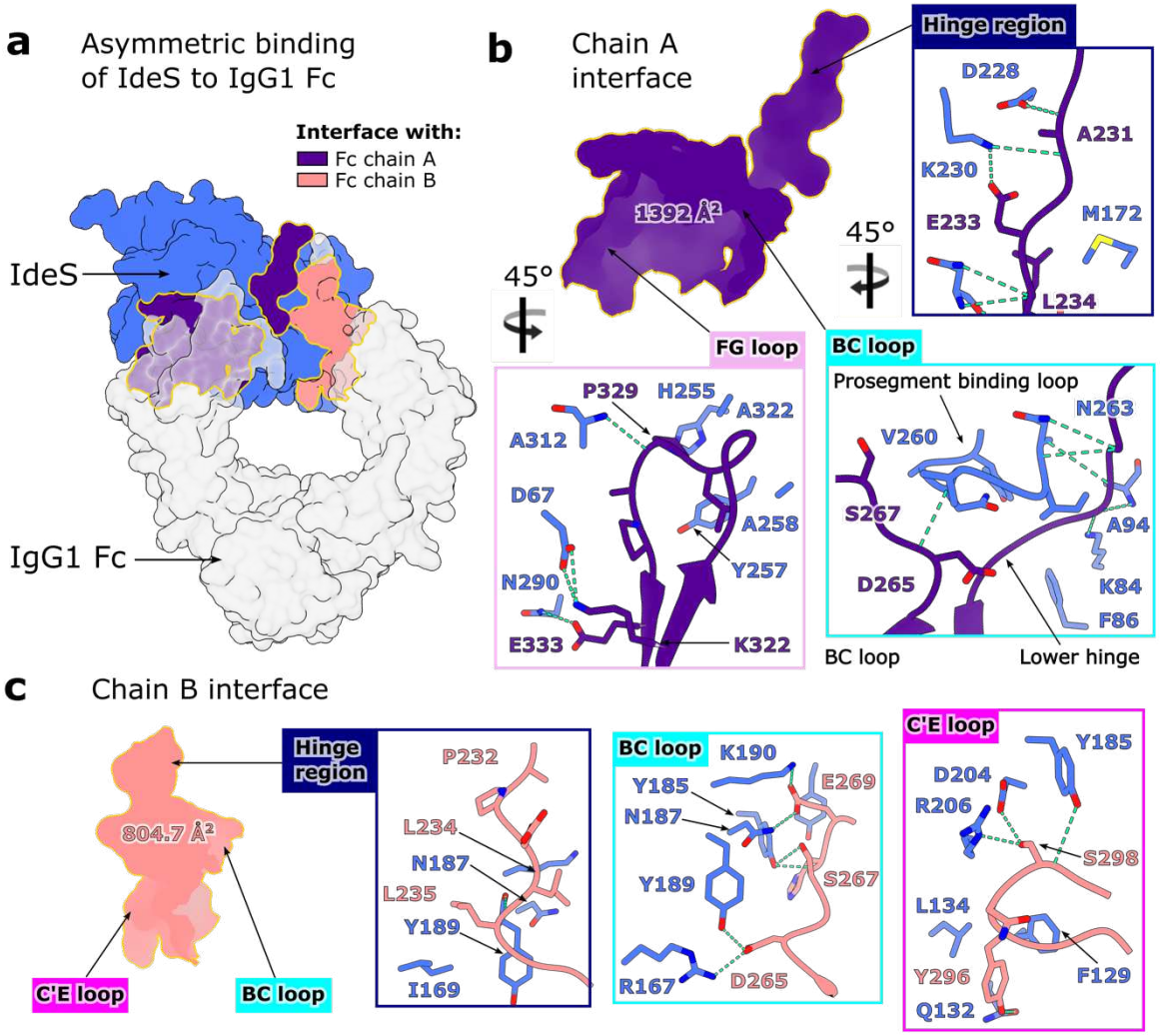


Figure 48: Asymmetric binding interface of IdeS^{C94A}-IgG1 Fc^{E382A} complex. a Overall view of complex depicted as a surface, with IdeS^{C94A} coloured blue and IgG1 Fc^{E382A} coloured silver. Interfaces of IdeS^{C94A} with chains A and B of the Fc are coloured indigo and coral, respectively. Glycans within the Fc have been omitted for clarity. **b** Interface between IdeS^{C94A} and IgG1 Fc^{E382A} chain A, involving the Fc hinge region, BC loop and FG loop. **c** Interface between IdeS^{C94A} and IgG1 Fc^{E382A} chain B, involving the hinge region, BC loop and C′E loop. **b, c** Residues involved in binding are depicted as sticks and coloured by heteroatom (oxygen, red; nitrogen, blue), with hydrogen bonds depicted as green dashes.

A secondary interface is formed across the second Fc chain (annotated as chain B in the crystal structure; **Figure 48c**), with an interface area of 804.7 Å^2 and a solvation free energy gain of -7.6 kcal/mol. A smaller proportion of the Fc hinge contributes (A231–G237), but PDBePISA predicts favourable hydrophobic interactions here, albeit not to the same extent as chain A. Subsequent recognition of this Fc chain is driven by interactions with the BC loop, and, in contrast to chain A, the C′E loop containing the *N*-linked glycan (**Figure 48c**). PDBePISA additionally predicts a small number of interactions between the enzyme and the Fc *N*-linked glycans; the lack of electron density for any monosaccharides past β-1,2-linked GlcNAc suggests that any further glycan processing

doesn't affect complex formation, and that IdeS can accommodate IgG with heterogenous glycosylation.

Although IdeS interacts with both chains in the Fc hinge simultaneously, following cleavage of the first chain, the complex would need to dissociate before the second cleavage could occur. This observation is also evidenced by detection of single-cleaved Fc in enzymatic assays and in clinical studies [230, 239, 293, 403]. The binding interface is likely altered for single-cleaved Fc and would likely explain the slower rate of cleavage [230, 231, 239]. It is also interesting to note that, aside from the hinge region, IdeS binds Fc regions implicit in Fcγ-receptor binding, an observation also inferred by its ability to counteract Fc-mediated effector functions by competitive binding inhibition [236]. Moreover, IdeS residues interacting with the Fc are largely conserved across the two IdeS isoforms, and any substitutions are mostly to similar amino acids, which aids in explaining their near identical activity [237].

4.4 Discussion

The mechanism of action of the streptococcal protease IdeS on human IgG was largely unknown prior to the start of this project. This chapter details work towards the goal of solving the structure of this enzyme in complex with IgG, using the panel of "less-crystallisable" IgG1 Fcs described in chapter 3 to encourage co-crystal growth. A truncated version of the enzyme was designed (lacking the signal sequence, which is not resolved in IdeS crystal structures), in order to remove any flexible regions which may inhibit crystallisation. Catalytic activity was abolished by the inclusion of a C94A amino acid exchange. Following successful crystallisation of one complex, a series of other constructs were tested, including the clinically-used variant of IdeS (Imlifidase), and additional IgG substrates cleaved by the enzyme.

All complexes were successfully expressed and purified, as detailed in section 4.1. However, only the truncated IdeS^{C94A}-IgG1 Fc complex was successfully crystallised. Purification of this complex using SEC resulted in a peak shift relative to the apo enzyme and the Fc, indicating stable complex formation, and subsequent SDS-PAGE analysis showed how this additional purification removed some remaining lower molecular weight impurities. As a result, complexes with each of the "less-crystallisable" Fcs produced crystals in several different conditions.

The Imlifidase-IgG1 Fc complexes were purified to homogeneity as judged by SDS-PAGE. Although direct analytical SEC was not performed, complexes eluted as a single peak, indicating stable complex formation. Initial crystallisation attempts did not produce crystals large enough for data collection, however. Further attempts could utilise crystalline seed stock of the IdeS-IgG1 Fc complex, now that structure solution has confirmed these crystals are of the protein-protein complex.

Crystallisation trials of the asymmetric IgG1 Fc in complex with IdeS did yield some initial hits, although some of these could not be reproduced in subsequent purifications. It was later elucidated that the crystals forming in condition G5 from the JCSG-plus™ screen were of the antibody only, and diffracted to relatively low resolution (~3.5 Å). With more time, crystals from some of the other conditions could be optimised further: condition A5 from the Morpheus screen, in particular, produced some sharper-looking crystals (**Figure 39a**), although the quality of a crystal cannot be judged based on its appearance alone. In addition, in further purifications, seed stock of the IdeS-IgG1 Fc complex could be utilised to try and encourage complex crystallisation.

Expression yields for IgG2 Fc^{E382R}, IgG3 Fc^{E382R} and IgG4 Fc^{E382R} were low, and thus limited purification steps were carried out in order to retain sufficient protein for crystallisation experiments. Although SDS-PAGE analysis showed complexes were relatively pure, particularly the IgG3 Fc-IdeS complex, some faint contaminating bands were present. SEC purification of the complexes would thus remove some of these unwanted contaminants, and likely also remove any leftover apo proteins, which would improve the chances of successful crystallisation. The lower affinity of IgG2 Fc for IdeS might make this complex unsuitable for SEC purification, however. As with the other complexes described in this chapter, crystallisation could also be attempted using IdeS-IgG1 Fc seed stock to encourage complex crystallisation, and complexes with Imlifidase could also be tested.

Crystals obtained for the IgG2 Fc and IgG4 Fc-IdeS complexes were found to be the respective antibodies alone. IgG4 Fc^{E382R} appeared to crystallise poorly, with the majority of crystals diffracting to fairly low resolution (>3 Å). A single crystal diffracted to ~2.6 Å, however. Analysis of the crystal packing for this construct, along with that of IgG2 Fc^{E382R} which diffracted to <2 Å, revealed that the exchanged arginine at position 382 was involved in the crystal packing interactions, along with nearby amino acid E380. IgG2/4 Fc could therefore be rendered "less-crystallisable" by utilising amino acid exchanges of these residues, such as to an alanine side chain, which may encourage co-crystal growth.

The crystal structure of truncated $IdeS^{C94A}$ in complex with IgG1 Fc^{E382A} was solved to a resolution of 2.34 Å. This confirmed previous suggestions that IdeS recognises IgG Fc using exosite binding [244, 246, 247], by revealing how the enzyme not only binds along a large section of the IgG hinge region (specifically, residues 229–239), but also across both Fc Cγ2 domains. The enzyme clamps down over the Fc lower hinge region, which brings the scissile peptide bond (between glycines 236 and 237) into close proximity with the active site. Substrate access to the active site appears to be mediated by the so-called "prosegment binding loop", which is observed in a distinct conformation in the complexed enzyme, with respect to previously-elucidated apo structures [246, 247], and has a role in enzyme inhibition in related cysteine proteases [400-402].

The IdeS-IgG1 Fc crystal structure has enabled mapping of the extensive binding interface recognised by the enzyme. The Fc chain being "cleaved" within the structure is contacted most

extensively, while a secondary interface is formed with the other chain within the Fc homodimer. The hinge region of this second Fc chain is too far away from the IdeS active site to be cleaved within a single binding event, which is also evidenced by detection of single-cleaved Fc in enzymatic assays and in clinical studies [230, 239, 293, 403]. The binding interface for cleavage of the second Fc chain is likely altered, due to the destabilisation caused by a single cleavage of the hinge region, which in itself renders the antibody unable to activate the immune system [238]. This altered binding interface is likely less extensive, which would explain the slower cleavage rate for this chain [230, 231, 239]; crystallisation of the single-cleaved Fc could be used to confirm this.

The structural knowledge presented here will aid in further expanding the therapeutic uses of IdeS. For example, the deactivation of serum IgG using IdeS can strengthen the potency of therapeutic antibodies [292]; this strategy could be applied to potentiate any therapeutic antibody, in theory, if the antibody were designed to be resistant to cleavage by these enzymes, a venture which will be aided greatly by structural information. A current limitation of using bacterial enzymes as therapeutics is that they are inherently immunogenic, and so, during infection with *S. pyogenes*, anti-IdeS antibodies are produced [277]. Design of immunologically-distinct enzyme variants will therefore be imperative for the long-term therapeutic use of IdeS, and would enable repeated dosing. The *de novo* design of enzymes is only recently emerging [404] and its success is improved by utilising structure prediction methods [405]. The IdeS-Fc crystal structure can thus be used as a template to inform future structure-guided design of new enzyme variants.

To conclude, the crystal structure of IdeS in complex with its IgG1 Fc substrate provides a structural rationale for the unique properties of this enzyme, particularly its exquisite substrate specificity towards human IgG [6].

Chapter 5 Distinct substrate recognition by IgG-specific endoglycosidases EndoS and EndoS2

Streptococcus pyogenes secretes two endoglycosidases, EndoS [7] and EndoS2 [8], which specifically cleave the *N*-linked glycans from IgG Fc. The two enzymes differ slightly in their substrate specificities: EndoS is highly specific towards IgG bearing biantennary, complex-type Fc glycans [7, 255], whereas EndoS2 can also cleave oligomannose-type, hybrid-type and glycans containing bisecting GlcNAc [8, 256]. These carbohydrate moieties, covalently linked to residue N297 on each heavy chain of the Fc, are important for the structural integrity of the molecule [100-105] and are implicit in Fc-mediated immune activation, by mediating FcγR and complement binding [11, 103, 106]. Removal of these *N*-linked glycans is thus detrimental to Fc-mediated immune activation [105, 122]. Along with the protease IdeS (discussed in chapter 4), these enzymes are thought to help *S. pyogenes* to evade the immune system [253, 254].

The ablation of IgG Fc-mediated immune function by EndoS and EndoS2 has potential applications for dampening unwanted immune activation, such as during organ transplantation, or in an autoimmune disease context. EndoS has been successfully utilised as a treatment in several pre-clinical models of autoimmune disease [267, 270, 271, 278-281, 283-285], and, along with IdeS, was shown to be successful at inactivating donor-specific antibodies in a murine model of bone marrow transplantation [266]. The specific deactivation of competing serum IgG by EndoS and/or IdeS is also being investigated for the potentiation of therapeutic antibodies [105, 292]. Furthermore, EndoS and EndoS2 have additional biotechnological applications as tools for engineering antibody glycosylation [107, 298]: firstly, native glycans can be cleaved using wild-type EndoS or EndoS2, and then the desired glycoforms can be generated using this deglycosylated antibody scaffold. Such transglycosylation reactions have been optimised using variants of EndoS and EndoS2 [299-303]. The precise control of antibody glycosylation has been applied in several clinically-used antibodies for improved immune effector function [56, 294-297], thus demonstrating the potential utility of these enzymes in antibody engineering.

The potential therapeutic and biotechnological uses of EndoS and EndoS2 have prompted extensive research into their structure and function. Prior to the start of this project, crystal structures for both EndoS [257, 258] and EndoS2 [259] had been solved in complex with their respective glycan substrates, which identified a glycan-binding groove within their respective catalytic domains. This groove was observed to be larger in EndoS2, which allows the enzyme to accommodate a broader range of Fc glycans, including the bulkier oligomannose-type [259]. In addition, both enzymes were found to contain a so-called “carbohydrate binding module” (CBM). This domain was found to be essential for activity of both EndoS and EndoS2 [257, 259, 261], although its precise function was unknown.

This chapter describes work towards the goal of determining the structural basis behind the extreme specificity of EndoS and EndoS2 for their IgG Fc substrate. The panel of "less-crystallisable" IgG1 Fc variants described in chapter 3 were utilised with the aim of obtaining crystal structures of EndoS and EndoS2 in complex with IgG1 Fc.

Analysis of the EndoS-IgG1 Fc crystal structure (PDB 8A49) is included in the following publication: Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. *Nat Commun.* 2022;13:7801. This publication is included within Appendix A for reference. Furthermore, the EndoS2-IgG1 Fc crystal structure (PDB 8Q5U), along with analysis comparing EndoS and EndoS2 recognition of IgG, is included within the following pre-print: Sudol ASL, Tews I, Crispin M. Bespoke conformation and antibody recognition distinguishes the streptococcal immune evasion factors EndoS and EndoS2. *bioRIV*. 2023:553389 (see Appendix F).

5.1 Expression and purification of EndoS$^{D233A/E235L}$ and EndoS2$^{D184A/E186L}$ complexes with IgG1 Fc

In order to obtain crystal structures of the endoglycosidases in complex with uncleaved IgG1 Fc, inactive versions of the enzymes were required. Previous structural work had described constructs of EndoS and EndoS2 which could be successfully expressed, purified and crystallised [257-259, 406]; therefore, sequences of these constructs were utilised for enzyme production in this work. Constructs for full-length EndoS [258] and a truncated form of EndoS lacking the N-terminal 3-helix bundle (residues 98–995) [257] were both tested, since both versions of the enzyme hydrolyse IgG Fc glycans [257], and a truncated version may be more suitable for crystallographic studies. Specific mutations of catalytic residues were included to abolish enzyme activity: D233A/E235L exchanges were included within EndoS as previously described [258], while D184A/E186L exchanges were included as the equivalent substitutions within EndoS2. Constructs were synthesised into pET21a(+) vectors and designed to contain a C-terminal linker and His-tag (sequence LEHHHHHH) by Eurofins Genomics, as detailed in section 2.1.3. EndoS$^{D233A/E235L}$ and EndoS2$^{D184A/E186L}$ constructs were subsequently expressed in *E. coli* BL21 (DE3)pLysS cells and purified as detailed in sections 2.2.2 and 2.3.2, respectively. EndoS/EndoS2 was subsequently combined with IgG1 Fcs and purified further as described in section 2.4.2.

5.1.1 EndoS$^{D233A/E235L}$-IgG1 Fc

The truncated form of EndoS$^{D233A/E235L}$ was purified using nickel affinity purification. Fractions corresponding to 80 mM, 100 mM and 200 mM elutions were collected and analysed by SDS-PAGE (**Figure 49a**). A dominant band with a molecular weight consistent with that of truncated EndoS (~101 kDa) was present in each of the fractions. Fractions were concentrated separately (in 50 kDa

MWCO concentrators) and purified further using SEC. The trace for the 80 mM fraction produced a very broad peak suggesting lots of impurities or aggregation present; traces for 100 mM and 200 mM fractions, however, showed a single peak (eluting at ~103 mL and ~104 mL, respectively) with minimal impurities either side of the curve (**Figure 49b**). Fractions corresponding to the main peak for the 100 mM and 200 mM SEC runs were pooled for further investigation.

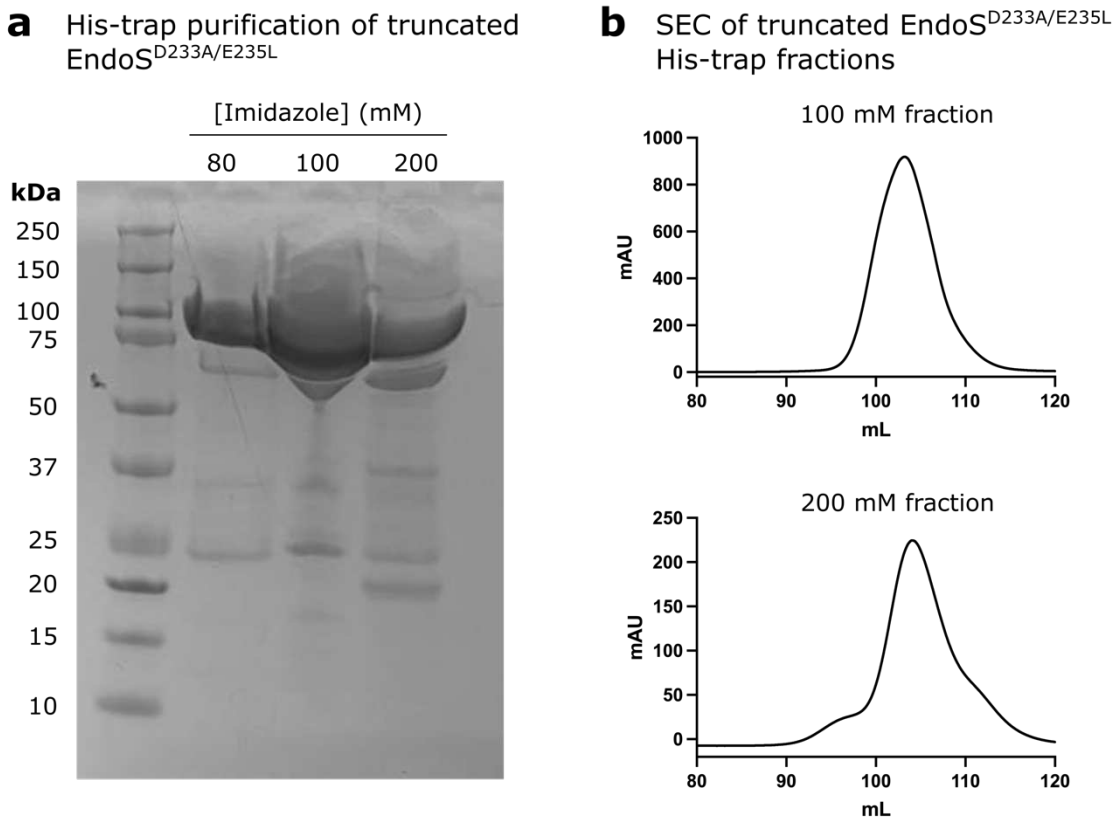


Figure 49: Purification of truncated EndoS$^{D233A/E235L}$. a SDS-PAGE of 80 mM, 100 mM and 200 mM fractions collected from nickel affinity purification of truncated EndoS$^{D233A/E235L}$. The main band in each fraction at ~101 kDa is consistent with overexpression of truncated EndoS. **b** SEC of collected 100 mM and 200 mM fractions from nickel affinity chromatography. Fractions corresponding to the main peak only were pooled for further investigation.

EndoS was combined with IgG1 Fcs in a 1:1 molar ratio for complex formation, due to previous work indicating that the enzyme behaves as a monomer in solution [257, 258]. Truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} and EndoS$^{D233A/E235L}$-IgG1 Fc^{E382A} complexes were subsequently subjected to SEC purification over a HiLoad® 16/600 Superdex 200 pg column, at a flow rate of 1 mL/min, which resulted in two separate species eluting off the column (**Figure 50a**). The main peak for each complex eluted at ~103 mL, while the second peak eluted at ~109 mL. SDS-PAGE analysis of the two peaks for each complex revealed that, in each peak, there was a band consistent with the molecular weights of both truncated EndoS (~101 kDa) and IgG1 Fc (~50 kDa). Each peak also contained a contaminating band at ~65 kDa (**Figure 50b**).

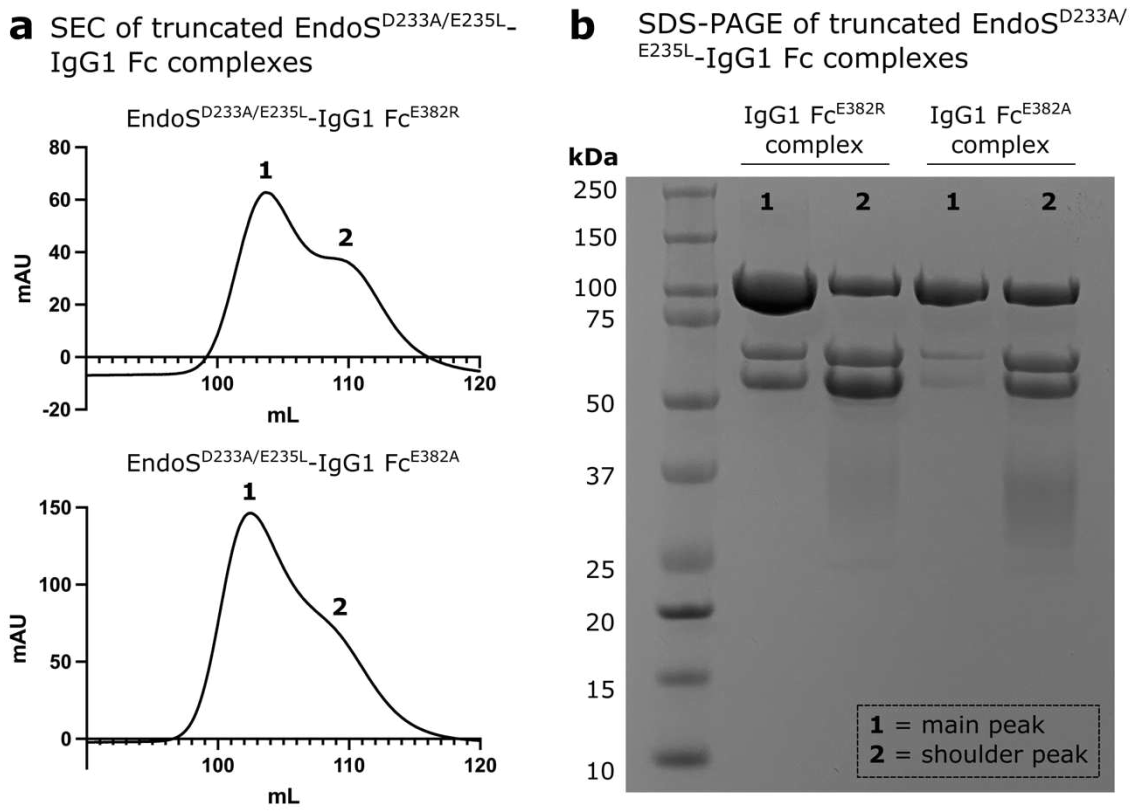


Figure 50: Purification of truncated EndoS$^{D233A/E235L}$-IgG1 Fc complexes. **a** SEC of complexes combined in a 1:1 molar ratio and applied to a HiLoad® 16/600 Superdex 200 pg column at 1 mL/min. The samples eluted as two separate species, annotated as peaks **1** and **2**. Fractions corresponding to these peaks were pooled and concentrated separately. **b** SDS-PAGE of peaks **1** and **2** following SEC of truncated EndoS$^{D233A/E235L}$-IgG1 Fc complexes. For peak **1** of each complex, there are bands present consistent with the molecular weights of truncated EndoS (~101 kDa) and IgG1 Fc (~50 kDa). Fractions corresponding to peak **1** for each complex only were pooled and taken forward for crystallisation.

In peak **1** for each complex, the 101 kDa band is much more prominent compared to the 50 kDa band, suggesting a greater proportion of EndoS compared to Fc. Conversely, in peak **2**, the 101 kDa band is less prominent, and the 50 kDa band is relatively stronger (**Figure 50b**). Although it would be expected for the heavier molecular weight protein to absorb more dye and therefore produce a more prominent band, the relative intensities of the bands did not seem consistent with a 1:1 complex. However, the presence of a ~50 kDa band in peak **1** (more visible for the Fc^{E382R} complex; an error likely resulted in less protein being loaded for the Fc^{E382A} complex and therefore fainter bands) suggested some complex formation. Fractions corresponding to peak **1** of each complex were pooled and taken forward for crystallisation experiments.

Initial crystallisation trials of the truncated EndoS$^{D233A/E235L}$-IgG1 Fc$^{E382R/E382A}$ complexes yielded crystals in many different conditions across the JCSG-plus™ and Morpheus crystallisation screens. Crystals from condition A8 from JCSG-plus™ diffracted to ~4 Å and yielded an initial solution of the structure, which revealed that the EndoS-IgG1 Fc complex had crystallised in a 2:1 stoichiometry (*i.e.*, two copies of EndoS in complex with one copy of the Fc homodimer). Therefore, for subsequent purifications, EndoS was combined with IgG1 Fc in a 2:1 molar ratio, as shown in **Figure 51** below.

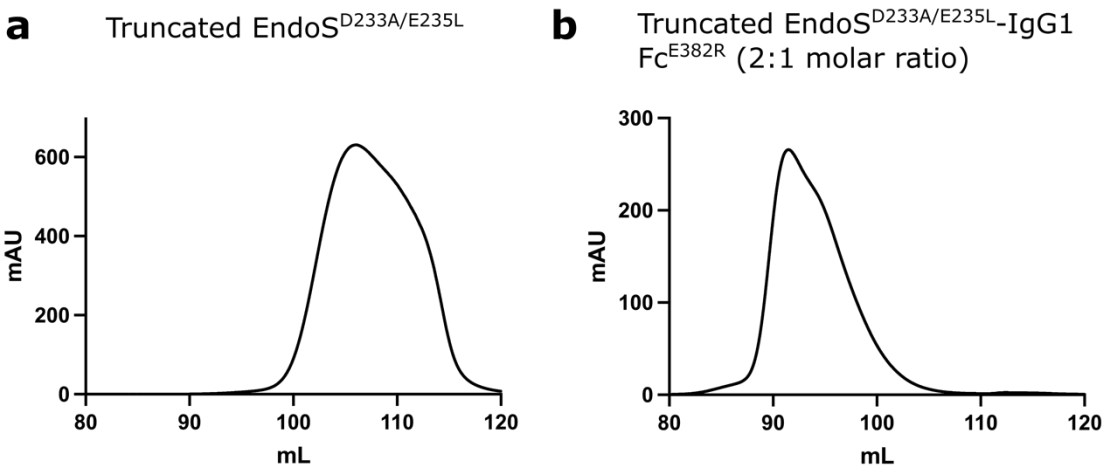


Figure 51: Purification of truncated EndoS$^{D233A/E235L}$-IgG1 Fc 2:1 complex. Comparison of SEC of truncated EndoS alone (**a**) versus truncated EndoS combined in a 2:1 molar ratio with IgG1 Fc^{E382R} (**b**), over a HiLoad® 16/600 Superdex 200 pg column, using the same batch of EndoS protein. The enzyme-Fc complex eluted as a single peak (with a slight shoulder), suggesting formation of a stable complex. N.B. curves cannot be compared directly due to different flow rates of the SEC purifications (1 mL/min in **a**; 0.5 mL/min in **b**).

When combined in a 2:1 molar ratio, truncated EndoS-IgG1 Fc eluted as a single peak, with a slight shoulder on the right (**Figure 51b**), either from complex dissociation or lower molecular weight impurities. The complex also eluted more quickly (at 91.53 mL) compared to apo EndoS (~106 mL) from the same batch of protein (**Figure 51a**), which typically indicates elution of a heavier molecular weight species, (*i.e.,* a complex). However, due to these SEC purifications being performed using different flow rates (1 mL/min in **a**; 0.5 mL/min in **b**), the curves cannot be compared directly. SEC of the complex was carried out at a slower flow rate in order to maximise the separation of impurities/aggregation from the sample, and thus obtain a purer batch of protein immediately prior to crystallisation. Fractions corresponding to the main peak of the complex only were taken forward for subsequent crystallisation experiments.

5.1.2 Full-length EndoS$^{D233A/E235L}$-IgG1 Fc

Following successful crystallisation of the truncated form of EndoS in complex with IgG1 Fc, crystallisation of the complexed full-length, inactive enzyme was attempted (residues 37–995 as described previously [258]; residues 1–36 correspond to the signal peptide). Full-length EndoS$^{D233A/E235L}$ (EndoSFL) and IgG1 Fc^{E382A} were purified as described in sections 2.3.1 and 2.3.2. SDS-PAGE analysis of the fractions from nickel affinity purification of EndoSFL revealed a dominant band consistent with the EndoSFL molecular weight of 108 kDa within the 100 mM and 200 mM fractions (**Figure 52a**); these fractions were pooled and taken forward for further purification.

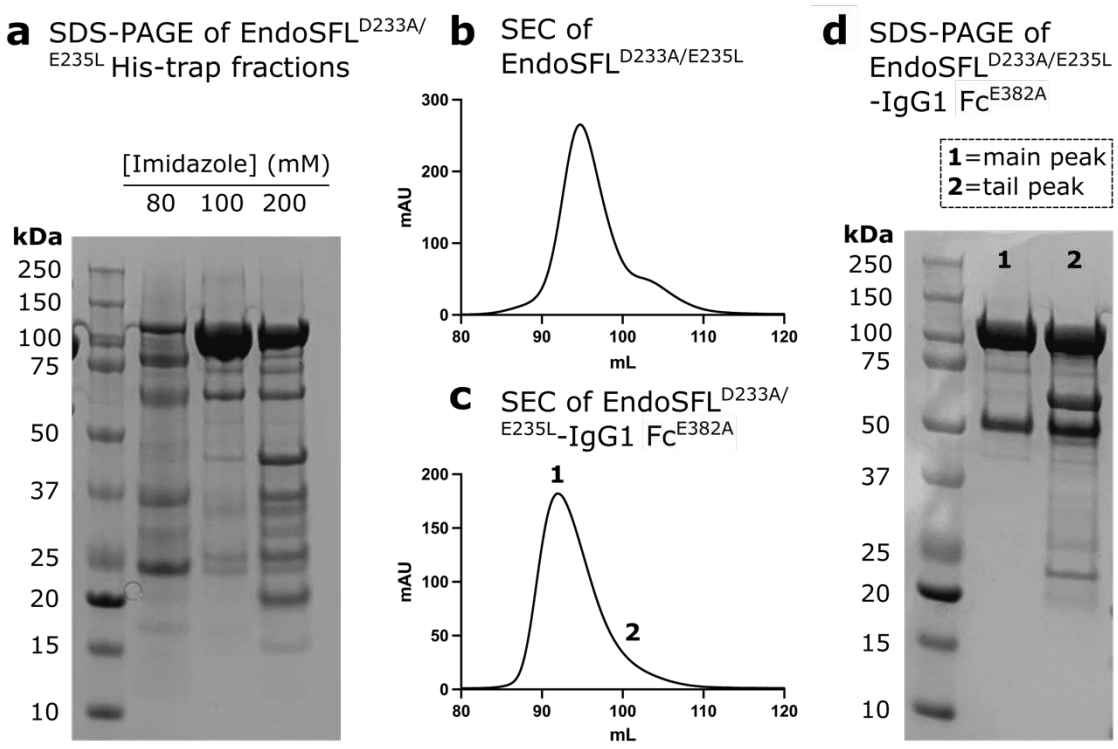


Figure 52: Purification of full-length EndoS$^{D233A/E235L}$ and its complex with IgG1 Fc^{E382A}. **a** SDS-PAGE of 80 mM, 100 mM and 200 mM fractions from nickel affinity purification of full-length EndoS$^{D233A/E235L}$ (EndoSFL). The 100 mM and 200 mM fractions contained a dominant band just above 100 kDa, which is consistent with the molecular weight of EndoSFL (108 kDa). **b** SEC purification of pooled 100 mM and 200 mM fractions from nickel affinity purification. **c** SEC purification of EndoSFL$^{D233A/E235L}$-IgG1 Fc^{E382A} (2:1 molar ratio of enzyme to Fc). **d** SDS-PAGE analysis of the main peak of the purified complex in **c**, versus the tail of the peak. The main peak (lane **1**) contained bands consistent with molecular weights of EndoSFL$^{D233A/E235L}$ (108 kDa) and IgG1 Fc^{E382A} (50 kDa). Fraction **2**, (corresponding to the tail of the peak) contained the same bands, along with additional contaminating bands. Fractions corresponding to the main peak only were pooled and taken forward for crystallisation.

SEC purification of the pooled 100 + 200 mM EndoSFL fractions produced a single peak, which eluted at ~95 mL (**Figure 52b**). The purified protein was subsequently combined with IgG1 Fc^{E382A} in a 2:1 molar ratio (of enzyme to Fc) and applied to the column again, after which the complex eluted as a single peak at ~92 mL (**Figure 52c**). The earlier elution of this species with respect to EndoSFL alone suggested formation of a stable 2:1 complex. SDS-PAGE analysis of the main peak revealed dominant bands consistent with the molecular weights of EndoSFL (108 kDa) and IgG1 Fc^{E382R} (50 kDa). The tail of the peak contained additional contaminating bands (**Figure 52d**); therefore, fractions corresponding to the main peak only were taken forward for crystallisation.

5.1.3 EndoS2$^{D184A/E186L}$-IgG1 Fc

Inactive EndoS2 (EndoS2$^{D184A/E186L}$) was expressed and purified as detailed in sections 2.2.2 and 2.3.2. SDS-PAGE analysis of the 80 mM, 100 mM and 200 mM fractions eluted from nickel affinity purification showed that only the 80 mM fraction contained a dominant band consistent with the expected molecular weight of EndoS2 (~91 kDa) (**Figure 53a**), therefore protein from this fraction

only was taken forward for further purification. The following SEC purification resulted in the elution of a single peak at ~90 mL (**Figure 53b**).

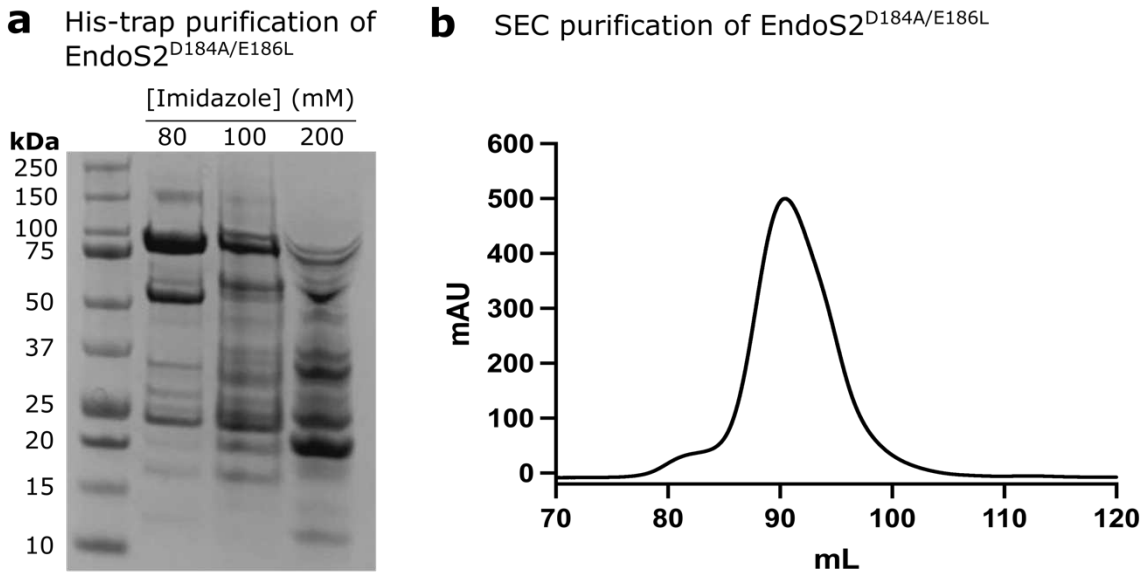


Figure 53: Purification of EndoS2$^{D184A/E186L}$. **a** SDS-PAGE of EndoS2$^{D184A/E186L}$ 80 mM, 100 mM and 200 mM fractions from nickel affinity purification. A dominant band consistent with the molecular weight of EndoS2 (~91 kDa) is present in the 80 mM fraction. **b** SEC purification of EndoS2$^{D184A/E186L}$ collected from the 80 mM fraction.

Due to the similar domain architectures of EndoS and EndoS2 [257-259], it was hypothesised that EndoS2 may also bind IgG1 Fc in a 2:1 stoichiometry (of enzyme to Fc). Therefore, EndoS2 and "less-crystallisable" IgG1 Fcs were combined in a 2:1 molar ratio and purified again with SEC (**Figure 54a**). Each of the complexes eluted as a single peak with a slight shoulder to the right, and eluted earlier than inactive EndoS2 alone, indicating formation of stable complexes (complexes with Fc E382R, E382S and E382A variants eluted at ~87 mL, ~88 mL and ~89 mL, respectively). SDS-PAGE of the main and tail peaks for each complex revealed that the purified complexes were in the main peak, and the tail peaks contained additional lower molecular weight impurities (**Figure 54b**). Fractions corresponding to the main peak only were taken forward for crystallisation.

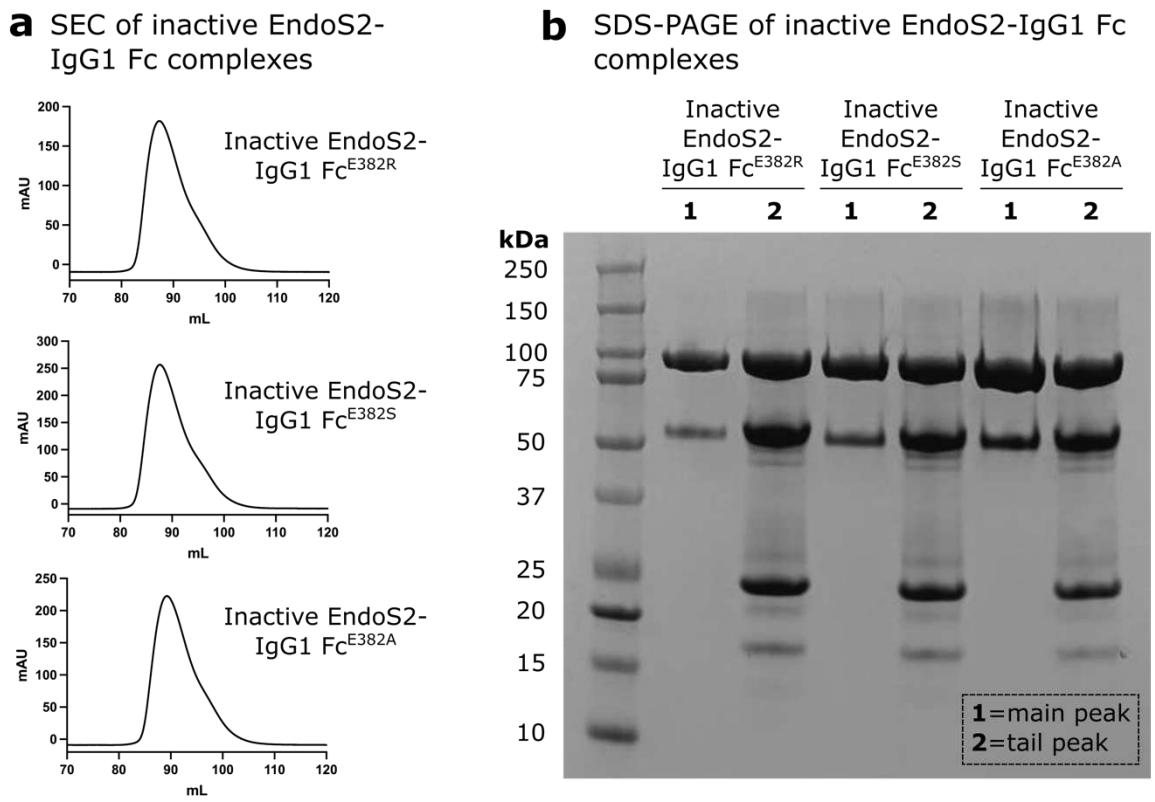


Figure 54: Purification of EndoS2$^{D184A/E186L}$-IgG1 Fc complexes. **a** SEC of inactive EndoS-IgG1 Fc complexes (combined in a 2:1 molar ratio of enzyme to Fc). Each complex eluted as a single peak with a slight tail peak on the right-hand side; fractions corresponding to the main and tail peaks were pooled separately. **b** SDS-PAGE of main and tail peaks for inactive EndoS2-IgG1 Fc complexes. Main peaks (**1**) for each complex contain bands consistent with the molecular weights of EndoS2 (~91 kDa) and IgG1 Fc (~50 kDa). Tail peaks (**2**) contain additional contaminating bands. Fractions corresponding to the main peak only for each complex were taken forward for crystallisation.

Initial crystallisation of each of the three EndoS2$^{D184A/E186L}$-IgG1 Fc complexes produced crystals in many different conditions across the JCSG-plus™, Morpheus and Proplex™ crystallisation screens; however, all these crystals diffracted to very low resolution (~6 Å or lower). The flexible nature of the Fc was hypothesised to possibly be inhibiting crystallisation; therefore, an IgG1 Fc construct with L234C exchanges was designed, in an attempt to engineer in an additional disulphide bridge within the hinge region and thereby reduce flexibility in this region. IgG1 Fcs containing both E382R/A exchanges and the L234C exchanges were expressed and purified as detailed in sections 2.2.1 and 2.3.1.

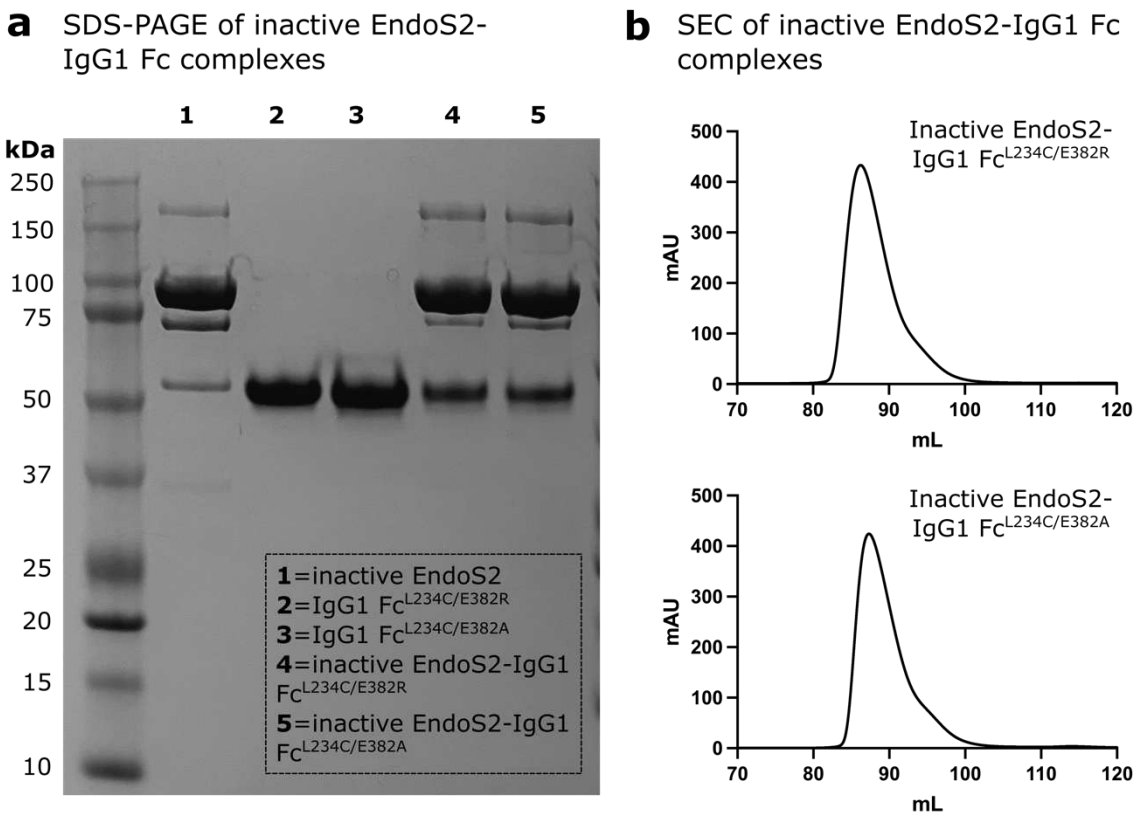


Figure 55: Purification of inactive EndoS2-clamped hinge IgG1 Fc complexes. **a** SDS-PAGE of inactive EndoS2 and IgG1 Fcs containing L234C mutations (in addition to E382R/A exchanges) before and after complex formation. Dominant bands are present at molecular weights consistent with that of inactive EndoS2 (~91 kDa) and IgG1 Fcs (~50 kDa). **b** SEC purification of inactive EndoS2-IgG1 Fc complexes, each yielding a single peak with a slight tail on the right-hand side. Complexes were applied to a HiLoad® 16/600 Superdex 200 pg column at 0.5 mL/min.

SDS-PAGE analysis revealed that, prior to complex formation, IgG1 $Fc^{L234C/E382R}$ (lane **2**) and IgG1 $Fc^{L234C/E382A}$ (lane **3**) had been purified to homogeneity, while inactive EndoS2 still contained some impurities (lane **1**; **Figure 55a**). EndoS2 and the clamped hinge IgG1 Fcs were subsequently combined in a 2:1 molar ratio and purified again by SEC, which yielded single peaks for each complex, eluting at ~86 mL and ~87 mL for the complexes with Fcs from lanes **2** and **3**, respectively. Small amounts of lower-molecular weight impurities were observed as a small shoulder on the right-hand side of each peak (**Figure 55b**); these "tail" sections were discarded, and fractions corresponding to the main peak only were taken forward for crystallisation experiments.

SDS-PAGE of the resulting complexes revealed dominant bands consistent with the molecular weight of inactive EndoS2 (~91 kDa), along with less-prominent bands consistent with the molecular weight of IgG1 Fc (~50 kDa), as observed for previous EndoS/EndoS2-Fc 2:1 complexes (**Figure 50b**, **Figure 52c**, **Figure 54b**). Small amounts of contaminating bands at just above 150 kDa and ~75 kDa were also observed; however, crystals of the inactive EndoS2-IgG1 Fc complex were still obtained from these protein samples.

5.2 Crystallisation of EndoS/EndoS2-IgG1 Fc complexes

The expressed and purified endoglycosidase-IgG1 Fc complexes were next used for crystallisation experiments (full methods of which are detailed in section 2.6.4). All crystals were grown at 21 °C.

5.2.1 EndoS$^{D233A/E235L}$-IgG1 Fc complex

Initial crystallisation trials for the truncated (residues 98–995), inactive form of EndoS in complex with IgG1 Fc^{E382R} and Fc^{E382A} yielded many different hits across the JCSG-plus™ and Morpheus crystallisation screens (**Figure 56**). The majority of these crystals diffracted to very low resolution (between 5 and 7 Å); however, crystals from condition A8 from the JCSG-plus™ screen produced crystals diffracting to ~4 Å. The diffraction images from one of these datasets were processed to a resolution of 3.88 Å and used to solve an initial structure of the complex.

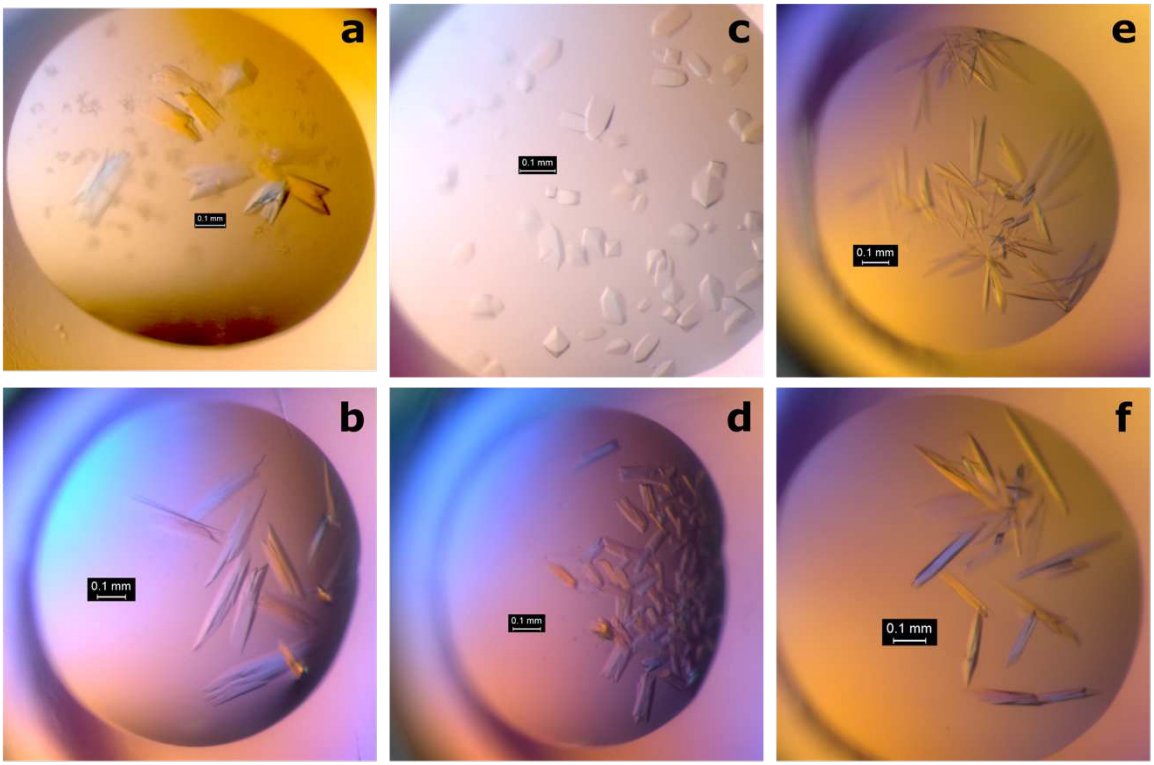


Figure 56: Crystals of EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex. Crystals grew in many different conditions, including: JCSG-plus™ A8, B4, D1 and G4 (**a**, **b**, **c** and **d**, respectively) and Morpheus D10 and F10 (**e** and **f**, respectively).

Subsequent crystallisation trials were set up using freshly purified protein complexes. Crystals from the initial trials were harvested to make seed stock (as detailed in section 2.6), and utilised for seeding these new plates, in an attempt to find novel crystal hits and yield better quality crystals. These subsequent trials also utilised the full panel of "less-crystallisable" IgG1 Fc variants (E382R, E382S and E382A). In total, two rounds of seeding yielded crystals diffracting to sufficient resolution for

structure solution. The full list of crystals taken for data collection are detailed in **Table 19**; crystals diffracting well enough for structure solution are detailed in **Table 20**.

Table 19: Crystals of truncated EndoS$^{D233A/E235L}$-IgG1 Fc complexes taken for crystallographic data collections.

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}	JCSG-plus™ A8	I03	10/04/2022	4.03
	JCSG-plus™ A12	I03	13/05/2022	6.17
	JCSG-plus™ B2	I03	30/05/2022	3.73
	JCSG-plus™ B4	I03	10/04/2022	5.30
	JCSG-plus™ B12	I03	30/05/2022	5.76
	JCSG-plus™ C4	I03	13/05/2022	6.69
	JCSG-plus™ D1	I03	10/04/2022	6.82
	JCSG-plus™ G4	I03	13/05/2022	3.79
	JCSG-plus™ G6	I03	10/04/2022	N/A[b]
	JCSG-plus™ G7	I03	13/05/2022	4.00
	JCSG-plus™ H6	I03	30/05/2022	N/A
	Morpheus A2	I03	30/05/2022	6.65
	Morpheus A6	I03	30/05/2022	5.70
	Morpheus A10	I03	30/05/2022	4.15
	Morpheus B8	I03	30/05/2022	3.25
	Morpheus D2	I03	30/05/2022	5.56
	Morpheus D6	I03	30/05/2022	N/A
	Morpheus D10	I03	13/05/2022	N/A
	Morpheus D11	I03	30/05/2022	5.34

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
	Morpheus E3	I03	30/05/2022	7.00
	Morpheus E10	I03	13/05/2022	6.69
	Morpheus F10	I03	13/05/2022	6.29
	Morpheus H10	I03	13/05/2022	5.52
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382S}	JCSG-plus™ A9	I03	30/05/2022	6.01
	JCSG-plus™ A10	I03	30/05/2022	3.94
	JCSG-plus™ B4	I03	30/05/2022	5.93
	JCSG-plus™ C4	I03	30/05/2022	N/A
	JCSG-plus™ C10	I03	30/05/2022	N/A
	JCSG-plus™ E12	I03	30/05/2022	N/A
	JCSG-plus™ G4	I03	30/05/2022	3.49
	JCSG-plus™ G6	I03	30/05/2022	6.72
	JCSG-plus™ G7	I03	30/05/2022	3.85
	JCSG-plus™ H6	I03	30/05/2022	N/A
	Morpheus A10	I03	30/05/2022	3.72
	Morpheus F6	I03	30/05/2022	5.79
	Morpheus H10	I03	30/05/2022	N/A
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382A}	JCSG-plus™ B4	I03	30/05/2022	N/A
	JCSG-plus™ B12	I03	10/04/2022	N/A
	JCSG-plus™ C4	I03	30/05/2022	N/A
	JCSG-plus™ E12	I03	30/05/2022	N/A
	JCSG-plus™ G4	I03	30/05/2022	3.57

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
	JCSG-plus™ G7	I03	30/05/2022	N/A
	JCSG-plus™ H6	I03	30/05/2022	N/A
	Morpheus A2	I03	30/05/2022	6.22
	Morpheus B2	I03	30/05/2022	4.95
	Morpheus E2	I03	30/05/2022	6.09
	Morpheus F2	I03	30/05/2022	5.98
	Morpheus F10	I03	10/04/2022	6.21
	Morpheus H10	I03	30/05/2022	7.71

[a]Best resolution is given as determined by the xia2/DIALS automatic processing pipeline. [b]N/A refers to either no diffraction or processing failure by xia2/DIALS.

Table 20: Crystals of the EndoS$^{D233A/E235L}$-IgG1 Fc complex used for structure solution.

Complex	Crystallisation condition	Date acquired	Resolution (Å)[a]
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}	0.2 M ammonium formate, 20 % w/v PEG[b] 3350 (JCSG-plus™ A8)	10/04/2022	3.88
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}	0.2 M TMAO[c], 0.1 M Tris (pH 8.5), 20 % w/v PEG 2000 MME (JCSG-plus™ G4)	13/05/2022	3.71
EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}	0.09 M halogens (contains sodium fluoride; sodium bromide; sodium iodide), 0.1 M buffer system 2 (contains sodium HEPES; MOPS (acid)), pH 7.5, 37.5 % v/v precipitant mix 4 (contains 25 % v/v MPD[d]; 25 % w/v PEG 1000, 25 % w/v PEG 3350) (Morpheus B8)	30/05/2022	3.45

[a]Diffraction images were manually processed to the given resolution. [b]PEG, polyethylene glycol; [c]TMAO, trimethylamine *N*-oxide; [d]MPD, 2-methyl-2,4-pentanediol.

The best dataset obtained was produced by crystals of the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex, which were grown in condition B8 of the Morpheus crystallisation screen. Data collection was carried out on beamline I03 at Diamond Light Source (Oxfordshire, UK), using parameters detailed in **Table 7** (section 2.7.1). The complex crystallised in an orthorhombic space group $P2_12_12_1$, with

unit cell dimensions a = 96.529, b = 174.294, c = 193.059 and $\alpha = \beta = \gamma = 90°$. The data were cut to a final resolution of 3.45 Å, yielding a CC_{half} of 0.599 for the highest resolution shell. Final data collection statistics for the complex are detailed in **Table 21** below.

Table 21: Crystallographic data collection statistics for truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex.

Data Collection	**Truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}**
Beamline	I03 (DLS)
Resolution range (Å)	49.78–3.45 (3.51–3.45)[a]
Space group	$P2_12_12_1$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 96.529, 174.294, 193.059 90, 90, 90
Wavelength (Å)	0.9763
Unique reflections	43678 (2124)
Completeness (%)	100 (100)
R_{merge}	0.144 (1.951)
R_{meas}	0.150 (2.023)
R_{pim}	0.040 (0.532)
$I/\sigma(I)$	11.2 (0.6)
Multiplicity	13.7 (14.3)
CC_{half}	1.000 (0.599)
Wilson *B* factor (Å^2)	116.4

[a]Values for the highest resolution shell are shown in parentheses.

The structure of the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex was solved by molecular replacement, using the program Molrep [320] within the ccp4i2 software suite [319]. The structure was solved using several rounds of molecular replacement, in which existing models for EndoS (PDB 4NUY) and IgG1 Fc (PDB 3AVE) were searched for separately within the density, yielding a model

containing two copies of truncated EndoS$^{D233A/E235L}$ and one IgG1 Fc^{E382R} homodimer. Section 2.7.3 details the general methods used for the structure solution process. Due to the relatively low resolution of the data, electron density maps were calculated using map sharpening in Refmac5 [322]. In addition, automatically-calculated non-crystallographic symmetry (NCS) restraints were used in refinement, which resulted in lower *R* factors. Final refinement statistics are included within **Table 22**.

Table 22: Crystallographic refinement statistics for truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex.

Refinement	**Truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R}**
Number of reflections (all/free)	43613/2116
R_{work} (%)	25.3
R_{free} (%)	31.1
RMSD[1]: Bonds (Å) Angles (°)	 0.0018 0.604
Molecules per ASU[2]	3
Atoms per ASU[2]	17087
Average *B* factors (Å^2) (protein/carbohydrate/water)	157, 156, 90
Model quality (Ramachandran plot): Most favoured region (%) Allowed region (%) Outliers (%)	 92.65 6.33 1.02, 0.4*

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
*as reported in the full wwPDB X-ray Structure Validation Report

This model of the truncated EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex has been deposited within the PDB (with accession number 8A49). The model has a relatively high R_{free} of 31.3 %, along with relatively high *B* factors, which can in part be expected due to the low resolution of the data. Validation of the structure in the PDB additionally computes an RSRZ score (Real Space R-value, normalised against residue type and resolution), which gives an indication of how well the model fits

against the experimental data. The EndoS-IgG1 Fc model also has a relatively high RSRZ score compared to structures of a similar resolution. Within the structure, one copy of truncated EndoS is well-resolved in the electron density, along with the Fc homodimer; however, the second copy of EndoS exhibits much more disorder and less well-resolved electron density. This is visualised in higher *B* factors and RSRZ scores for residues within this copy of the enzyme (annotated as chain D in the structure), as shown in **Figure 57**. The well-ordered copy of the complex was therefore used for structure depiction and analysis (discussed in section 5.3.1).

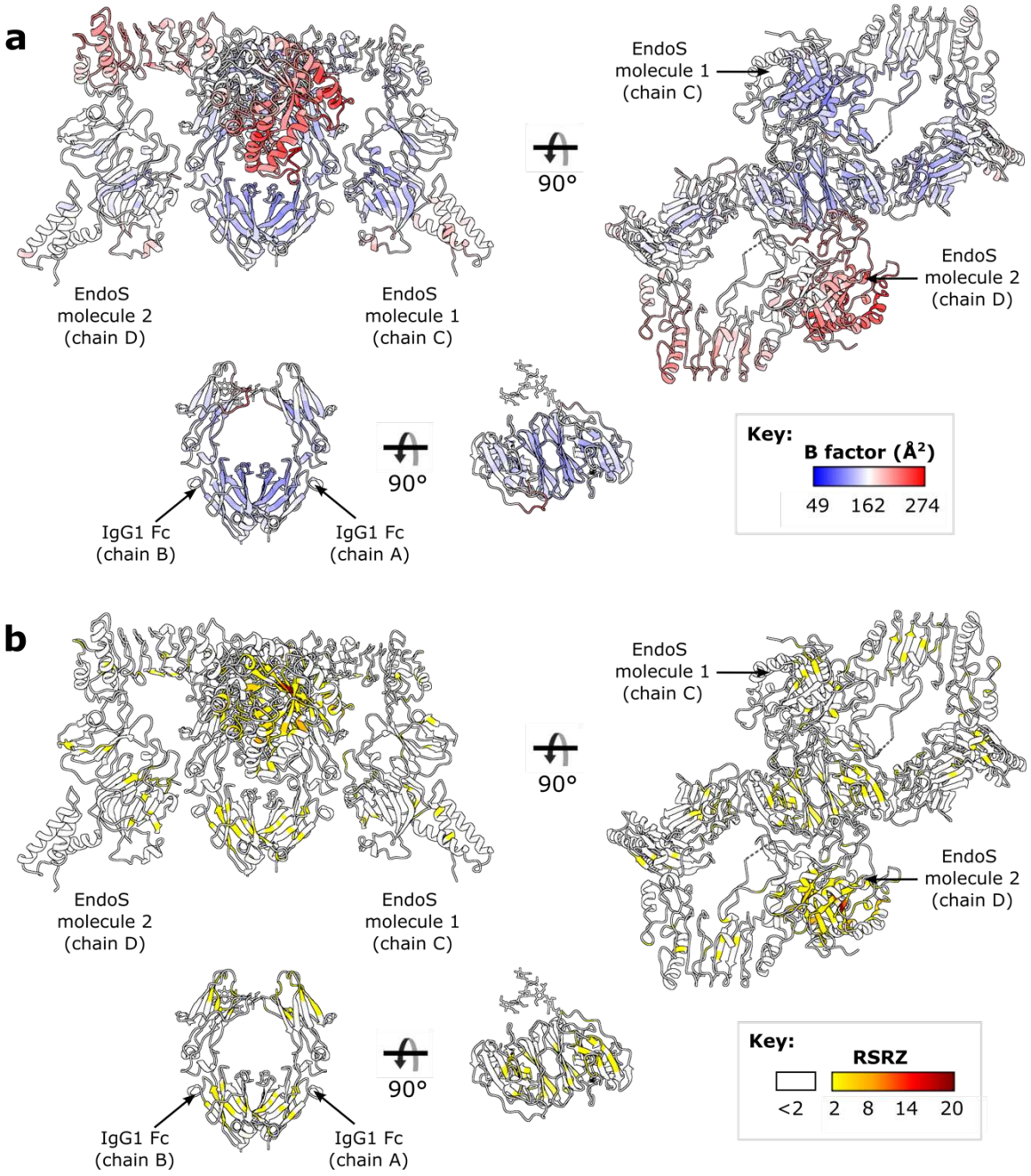


Figure 57: Analysis of order within the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} crystal structure. a *B* factor distribution (depicted as average *B* factor per residue); **b** RSRZ distribution as reported in the wwPDB X-ray Structure Validation Report for this structure (PDB 8A49). Residues identified as outliers (RSRZ of 2.0 or greater) are coloured as shown in the scale; residues with RSRZ lower than 2.0 are coloured white.

As discussed in section 5.3.1, EndoS captures the Fc *N*-linked glycan in a "flipped-out conformation, which has not been observed previously in a crystal structure of IgG Fc. The presence of this carbohydrate group within the crystal structure was validated using a polder map [397], as included within **Figure 58**.

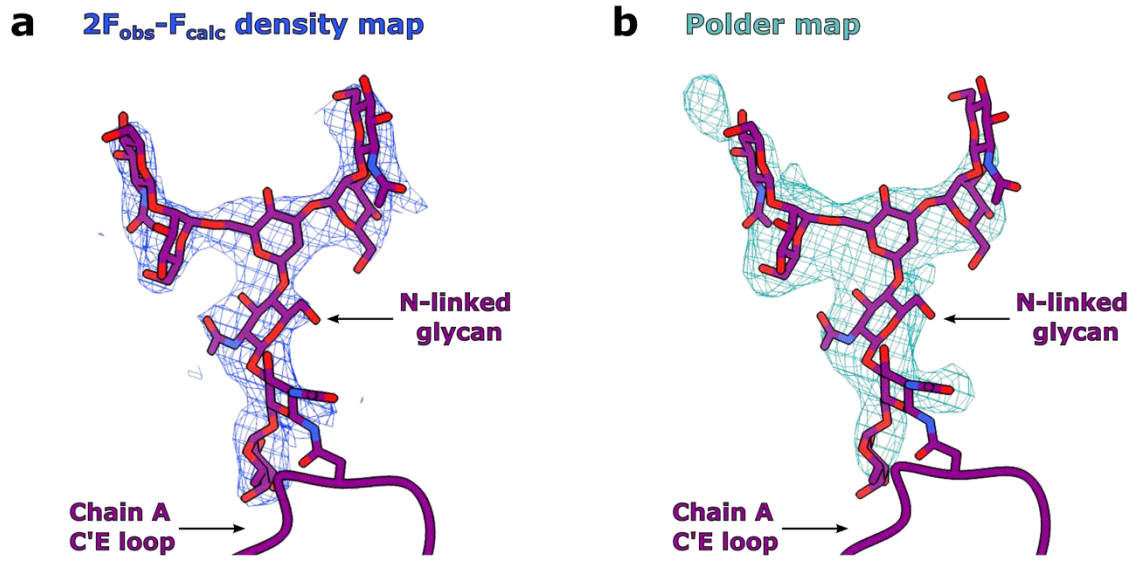


Figure 58: Electron density maps for the *N*-linked glycan on IgG1 Fc^{E382R} in complex with EndoS$^{D233A/E235L}$. **a** Final $2F_{obs}–F_{calc}$ electron density map (blue, contoured at 1.1 σ) for the *N*-linked glycan within chain A of IgG1 Fc^{E382R}. **b** Polder map for the *N*-linked glycan (teal, contoured at 3.5 σ), as calculated with the *phenix.polder* tool [397] within the *PHENIX* software suite [396]. A model of the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex was generated by molecular replacement in Molrep [320], using PDB models devoid of a glycan in this conformation (4NUY for EndoS; 3AVE for IgG1 Fc), and subsequently used in calculation of the polder map. **a, b** The glycan and the N297 side chain are depicted as sticks and coloured by heteroatom (oxygen in red and nitrogen in blue); the C′E loop is depicted as a cartoon.

A full discussion of the EndoS-IgG1 Fc complex is included within section 5.3.1.

5.2.2 Full-length EndoS$^{D233A/E235L}$-IgG1 Fc complex

Crystallisation trials for the full-length EndoS$^{D233A/E235L}$-IgG1 Fc^{E382A} complex were set up in JCSG-plus™, Morpheus and Proplex™ screens, using seed stock of EndoS-Fc crystals in a 1:100 dilution. Despite SDS-PAGE indicating that the complex had been purified to homogeneity (**Figure 52c**), crystallisation trials were unsuccessful and did not yield any hits that could be further optimised. However, only a single purification of this complex was attempted due to time limitations, and with only one IgG1 Fc construct.

5.2.3 EndoS2$^{D184A/E186L}$-IgG1 Fc complex

Initial crystallisation attempts for the EndoS2$^{D184A/E186L}$-IgG1 Fc complex produced lots of crystals across the JCSG-plus™, Morpheus and Proplex™ crystallisation screens. These initial trays were

seeded with stock made from crystals of the truncated EndoS$^{D233A/E186L}$-IgG1 complexes (detailed in section 5.2.1) at a 1:100 dilution. These initial crystals diffracted very poorly, if at all (between 6 and 9 Å). Subsequent crystallisation trials were set up using seed stock made from crystals of EndoS2$^{D184A/E186L}$-IgG1 Fc; multiple rounds of seeding were attempted but did not improve the diffraction capabilities of the crystals. A strategy to reduce the flexibility of the Fc within this complex, by inclusion of L234C exchanges within the lower hinge (in an attempt to engineer an additional stabilising disulphide bridge, detailed in section 5.1.3), was also utilised. A selection of EndoS2-Fc crystals grown throughout these experiments is pictured in **Figure 59** below.

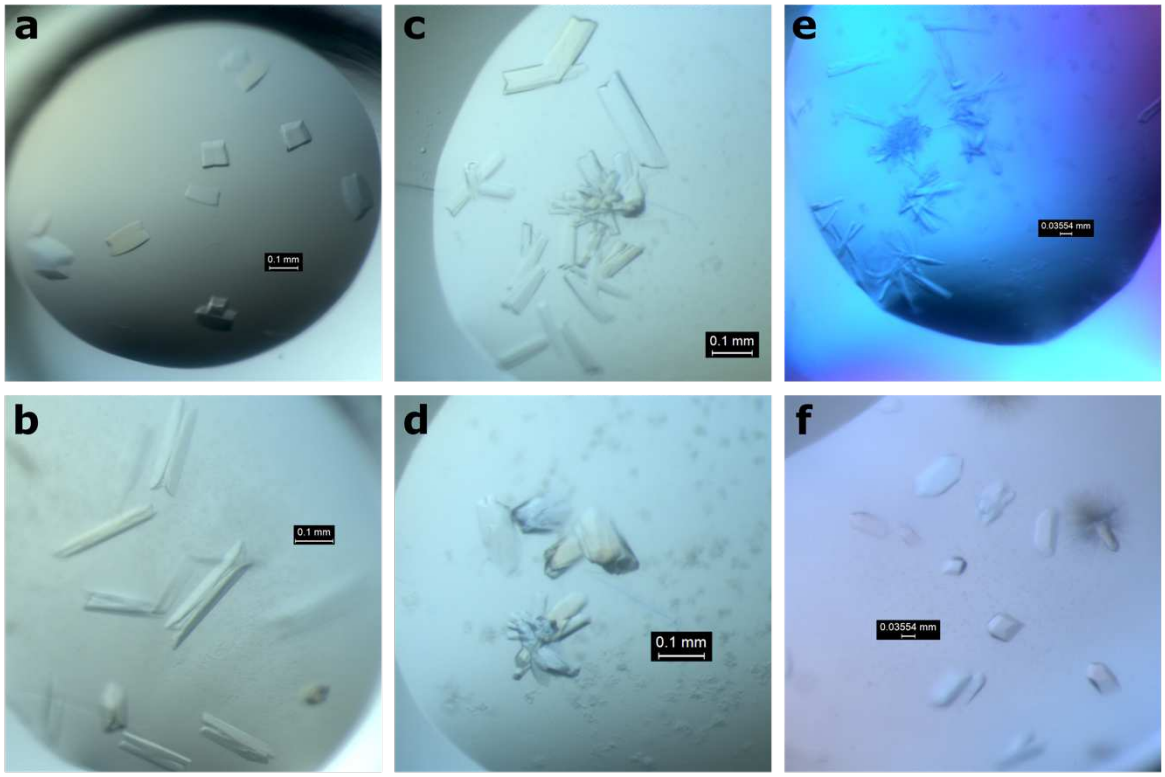


Figure 59: Crystals of EndoS2$^{D184A/E186L}$-IgG1 Fc complexes. a, b, c Crystals of EndoS2$^{D184A/E186L}$-IgG1 Fc^{E382R} complex, grown in: JCSG-plus™ G1 (**a**), Morpheus A6 (**b**), Proplex™ D7 (**c**) and Proplex™ E9 (**d**). **e** Crystals of EndoS2$^{D184A/E186L}$- IgG1 Fc$^{L234C/E382R}$ complex grown in JCSG-plus™ B2. **f** Crystals of EndoS2$^{D184A/E186L}$- IgG1 Fc$^{L234C/E382A}$ complex, grown in Morpheus G12.

The majority of all these crystals still diffracted very poorly (see **Table 23**). However, crystals harvested from a single well (condition G12 from the Morpheus crystallisation screen) diffracted to ~3 Å and were used to solve the structure of the complex (**Figure 59f**).

Table 23: Crystals of EndoS2$^{D184A/E186L}$-IgG1 Fc complexes taken for crystallographic data collections.

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
EndoS2$^{D184A/E186L}$-IgG1 Fc^{E382R}	JCSG-plus™ A2	I03	22/07/2022	7.37
	JCSG-plus™ A3	I03	22/07/2022	N/A[b]
	JCSG-plus™ A6	I03	22/07/2022	6.17
	JCSG-plus™ A11	I03	22/07/2022	8.19
	JCSG-plus™ B7	I03	22/07/2022	N/A
	JCSG-plus™ B12	I03	24/05/2023	6.03
	JCSG-plus™ C7	I03	22/07/2022	N/A
	JCSG-plus™ D2	I03	22/07/2022	N/A
	JCSG-plus™ E11	I04	17/09/2022	N/A
	JCSG-plus™ F2	I03	22/07/2022	7.61
	JCSG-plus™ G1	I03	22/07/2022	6.80
	JCSG-plus™ H2	I03	22/07/2022	N/A
	JCSG-plus™ H11	I03	22/07/2022	6.94
	Morpheus A2	I04	17/09/2022	7.73
	Morpheus A3	I03	22/07/2022	N/A
	Morpheus A6	I04	17/09/2022	N/A
	Morpheus B2	I04	17/09/2022	N/A
	Morpheus B8	I03	24/05/2023	8.44
	Morpheus D5	I04	17/09/2022	N/A
	Morpheus E7	I04	17/09/2022	N/A

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
	Morpheus F1	I04	17/09/2022	8.55
	Morpheus F10	I03	24/05/2023	N/A
	Morpheus H2	I04	17/09/2022	N/A
	Morpheus H5	I04	17/09/2022	N/A
	Proplex™ B12	I03	24/05/2023	N/A
	Proplex™ C7	I04	17/09/2022	N/A
	Proplex™ C8	I04	17/09/2022	N/A
	Proplex™ E12	I04	17/09/2022	N/A
	Proplex™ F2	I04	17/09/2022	N/A
EndoS2$^{D184A/E186L}$-IgG1 Fc^{E382S}	JCSG-plus™ D1	I04	17/09/2022	4.46
	JCSG-plus™ G1	I04	17/09/2022	N/A
	JCSG-plus™ G6	I04	17/09/2022	7.00
	Morpheus A10	I04	17/09/2022	N/A
	Morpheus B12	I04	17/09/2022	N/A
	Morpheus D7	I04	17/09/2022	N/A
	Morpheus D9	I04	17/09/2022	N/A
	Morpheus F12	I04	17/09/2022	N/A
	Proplex™ C1	I04	17/09/2022	N/A
	Proplex™ D3	I04	17/09/2022	N/A
	Proplex™ D9	I04	17/09/2022	7.92
	Proplex™ E3	I04	17/09/2022	N/A
	Proplex™ E4	I04	17/09/2022	N/A

Complex	Condition crystals harvested from	Beamline	Date	Best resolution achieved (Å)[a]
	Proplex™ E6	I04	17/09/2022	N/A
	Proplex™ E8	I04	17/09/2022	N/A
	Proplex™ E10	I04	17/09/2022	N/A
EndoS2$^{D184A/E186L}$-IgG1 Fc^{E382A}	JCSG-plus™ A2	I04	17/09/2022	N/A
	JCSG-plus™ B2	I04	17/09/2022	5.72
	JCSG-plus™ G1	I04	17/09/2022	7.60
	Morpheus A2	I04	17/09/2022	N/A
	Morpheus A11	I04	17/09/2022	N/A
	Proplex™ C8	I04	17/09/2022	N/A
	Proplex™ E6	I04	17/09/2022	N/A
	Proplex™ E10	I04	17/09/2022	N/A
	Proplex™ F2	I04	17/09/2022	7.40
EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382R}$	Proplex™ B10	I03	24/05/2023	N/A
	Proplex™ D3	I03	24/05/2023	N/A
	Proplex™ D4	I03	24/05/2023	7.06
	Proplex™ D8	I03	24/05/2023	N/A
EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$	Morpheus D9	I03	24/05/2023	4.39
	Morpheus G12	I03	24/05/2023	3.07
	Morpheus H8	I03	24/05/2023	8.73

[a]Best resolution is given as determined by the xia2/DIALS automatic processing pipeline. [b]N/A indicates either no diffraction or processing failure by the xia2/DIALS pipeline.

The crystals diffracting to ~3 Å were grown in condition G12 from the Morpheus crystallisation screen, and comprised inactive EndoS2 in complex with IgG1 Fc$^{L234C/E382A}$. Data collection was carried out on beamline I03 at Diamond Light Source, using parameters detailed in **Table 7** (section 2.7.1). Diffraction images were processed and data scaled in space group $P4_32_12$, with unit cell

dimensions of $a = b = 228.775$, $c = 161.629$ and $\alpha = \beta = \gamma = 90°$. The data were truncated to a resolution of 3.00 Å, to yield a CC_{half} of >0.6 for the highest resolution shell. Data collection statistics for this complex are detailed in **Table 24** below.

Table 24: Crystallographic data collection statistics for the EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex.

Data Collection	**EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$**
Beamline	I03 (DLS)
Resolution range (Å)	49.91-3.00 (3.05-3.00)[a]
Space group	$P4_32_12$
Unit cell dimensions: *a, b, c* (Å) α, β, γ (°)	 228.775, 228.775, 161.629 90, 90, 90
Wavelength (Å)	0.9763
Unique reflections	85,926 (4250)
Completeness (%)	100 (100)
R_{merge}	0.187 (2.614)
R_{meas}	0.191 (2.668)
R_{pim}	0.036 (0.527)
$I/\sigma(I)$	12.6 (0.5)
Multiplicity	27.5 (25.4)
CC_{half}	1.000 (0.640)
Wilson *B* factor (Å^2)	76.994

[a]Values for the highest resolution shell are reported in parentheses.

The structure of inactive EndoS2-IgG1 Fc was solved by molecular replacement, using search models of apo EndoS2 (PDB 6E58) and IgG1 Fc (PDB 3AVE). Within the complex, each EndoS2 monomer binds one chain within an Fc homodimer. Successive rounds of molecular replacement were required for modelling the whole asymmetric unit, which contains three copies of the EndoS2-

half Fc complex (depicted in **Figure 60a**). Automatic NCS restraints were used for refinement of the complex, along with defined translation-libration-screw (TLS) groups (detailed in **Table 25**) and restraints generated by the PDB-REDO server [391]. TLS groups were defined for the various domains in EndoS2 (detailed in section 5.3.2), and the Cγ2/Cγ3 domains within IgG Fc. Electron density maps were calculated using map sharpening in Refmac5 [322]. Final refinement statistics are shown in **Table 26** below.

Table 25: Defined translation-libration-screw (TLS) groups used in refinement of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex.

EndoS2$^{D184A/E186L}$	**IgG1 Fc$^{L234C/E382A}$**
TLS For Peptide chain D RANGE ′D 45. ′ ′D 386. ′ RANGE ′D 387. ′ ′D 547. ′ RANGE ′D 548. ′ ′D 680. ′ RANGE ′D 681. ′ ′D 832. ′	TLS For Peptide chain A RANGE ′A 237. ′ ′A 340. ′ RANGE ′A 341. ′ ′A 444. ′
TLS For Peptide chain E RANGE ′E 264. ′ ′E 386. ′ RANGE ′E 387. ′ ′E 547. ′ RANGE ′E 548. ′ ′E 680. ′ RANGE ′E 681. ′ ′E 832. ′	TLS For Peptide chain B RANGE ′B 237. ′ ′B 340. ′ RANGE ′B 341. ′ ′B 444. ′
TLS For Peptide chain F RANGE ′F 46. ′ ′F 386. ′ RANGE ′F 387. ′ ′F 547. ′ RANGE ′F 548. ′ ′F 680. ′ RANGE ′F 681. ′ ′F 832. ′	TLS For Peptide chain C RANGE ′C 236. ′ ′A 340. ′ RANGE ′C 341. ′ ′C 443. ′

Table 26: Crystallographic refinement statistics for EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex.

Refinement	**EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$**
Number of reflections (all/free)	85,849/4376
R_{work} (%)	21.6
R_{free} (%)	25.1
RMSD[1]: Bonds (Å) Angles (°)	 0.0066 1.18
Molecules per ASU[2]	6* (3 copies of EndoS2-half Fc complex)
Atoms per ASU[2]	21,168
Average *B* factors (Å^2) (protein/ligand/water)	 (105.75/105.07/73.20)
Model quality (Ramachandran plot)[3]: Most favoured region (%) Allowed region (%) Outliers (%)	 95.98 3.82 0.19, 0.0**

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
[3]calculated in Molprobity
*each half Fc (*i.e.*, one chain within the homodimer) is counted as a separate molecule
**according to wwPDB X-ray Structure Validation Report

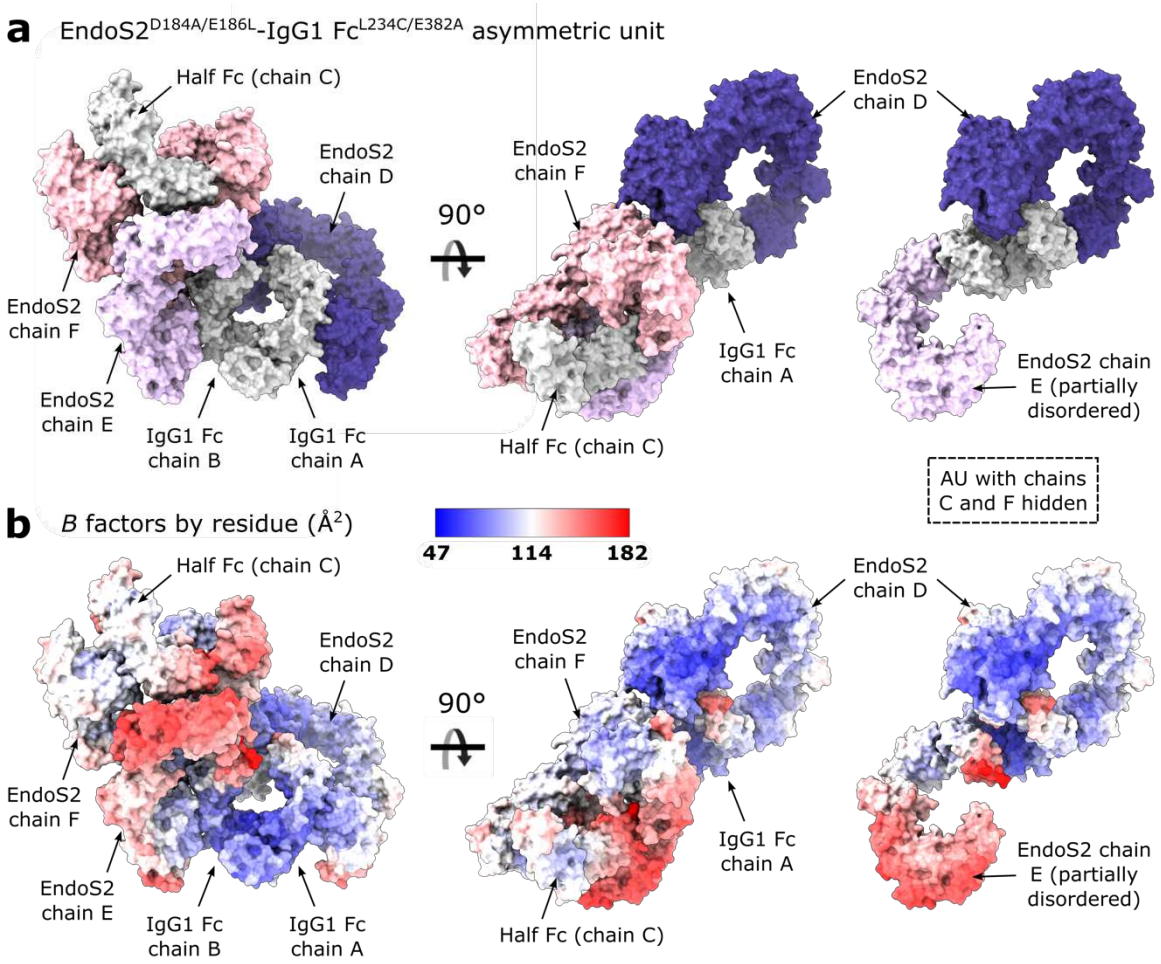


Figure 60: Crystal structure of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. a Arrangement of molecules within the crystal asymmetric unit (AU), containing three copies of the half Fc-EndoS2 complex. IgG1 Fc$^{L234C/E382A}$ is coloured silver; EndoS2$^{D184A/E186L}$ molecules are coloured purple, lilac and pink. **b** *B* factors by residue. **a, b** Far right panels show the complex AU with chains C and F hidden for clarity.

As also observed for the EndoS-Fc complex, one copy of the EndoS2-Fc complex (annotated as chains D and A for the enzyme and Fc monomer, respectively) is well-ordered within the structure, and is used for the discussion and depiction of the structure in the following sections. In contrast, the second copy of EndoS2 (annotated as chain E), which binds the second monomer within the Fc chain A/B homodimer, is more disordered, which can be visualised by higher *B* factors compared to the rest of the complex (**Figure 60b**). This copy of the enzyme was difficult to place in the density with molecular replacement, and as such the whole enzyme could not be built reliably into the density. Despite this, the model is of good quality for a resolution of 3 Å, with no Ramachandran outliers and a final R_{free} of 25.1 (**Table 26**). The structure has been pre-deposited within the PDB with code 8Q5U (at the time of writing, the model is being held for publication, i.e., not yet available to the public). The wwPDB X-ray Structure Validation Report for this structure is included within Appendix G.

The Fc *N*-linked glycan is captured by the enzyme in a "flipped-out" conformation, as seen within the EndoS-Fc complex and discussed in the following sections. The presence of the glycan moiety in this conformation was validated using a polder map [396, 397] (shown in **Figure 61b**). The glycan

residues also have relatively low *B* factors (**Figure 61a**), indicating it is bound in a stable conformation.

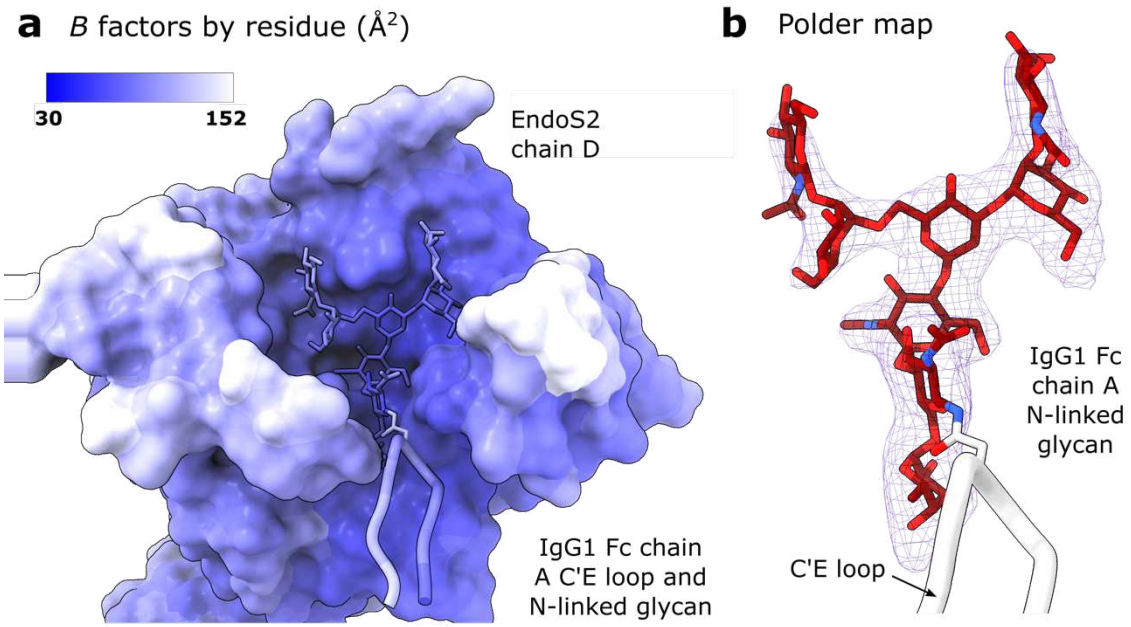


Figure 61: Validation of observed *N*-linked glycan conformation in EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystal structure. **a** *B* factors per residue. **b** Polder map for the *N*-linked glycan modelled for chain A of IgG1 Fc$^{L234C/E382A}$, calculated in *PHENIX* [396, 397] and contoured at 3 σ.

Full details of this crystal structure are discussed in section 5.3.2.

5.3 Structural analysis of EndoS/EndoS2 complexes with IgG1 Fc

Prior to the start of this project, there were several known structures of endoglycosidases in complex with their glycan substrate, including EndoS and EndoS2 [258, 259, 407-409]. This section details the crystal structures of inactive EndoS and EndoS2 in complex with their full IgG1 Fc substrates. Crystal structures of each complex are described separately and subsequently compared in section 5.4.

5.3.1 EndoS$^{D233A/E235L}$-IgG1 Fc complex

The structure of truncated EndoS$^{D233A/E235L}$ (residues 98-995) in complex with IgG1 Fc^{E382R} was solved in space group $P2_12_12_1$ and refined to a resolution of 3.45 Å, as detailed in section 5.2.1. The AU contains two copies of EndoS, each of which binds one Cγ2–Cγ3 monomer within the Fc homodimer. The best-ordered copy of the complex (annotated as chains A and C for the Fc monomer and enzyme, respectively) is used for structure depiction and analysis. The overall structure of the complex is depicted in **Figure 62**.

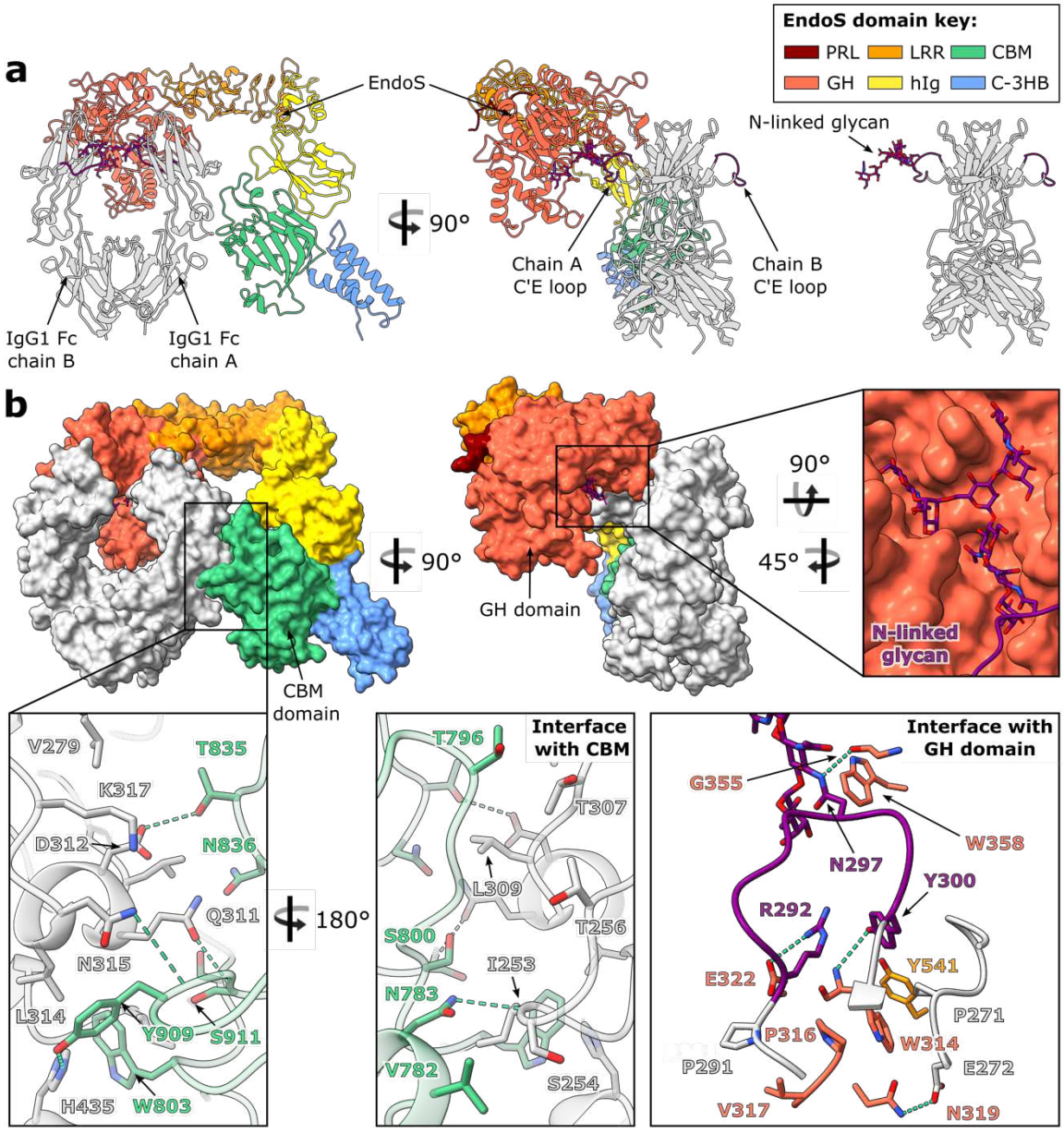


Figure 62: Crystal structure of EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex. a Overall structure of complex depicted as a cartoon. IgG1 Fc^{E382R} is coloured silver, with C′E loops coloured purple; the *N*-linked glycan is shown as sticks and coloured by heteroatom (oxygen, red; nitrogen, blue). EndoS domains are coloured as follows: proline-rich loop (PRL), maroon; glycosidase domain (GH), red; leucine-rich repeat domain (LRR), orange; hybrid Ig domain (hIg), yellow; carbohydrate-binding module (CBM), green; C-terminal 3-helix bundle (C-3HB), blue. **b** EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex depicted as a surface, highlighting binding to IgG1 Fc^{E382R} by the CBM and GH domains. Residues involved in binding are depicted as sticks and coloured by heteroatom. Hydrogen bonds are depicted as green dashes.

This structure of EndoS shows the same “V” shape as observed in its previously-solved structures [257, 258]. Following the previously-described domain classification [257], the structure comprises, from the N- to the C terminus: a proline-rich loop (PRL; residues 98–112), a glycosidase domain (GH; residues 113–445), a leucine-rich repeat domain (LRR; residues 446–631), a hybrid Ig domain (hIg; residues 632–764), a carbohydrate-binding module (CBM; residues 765–923) and a C-terminal three-helix bundle domain (C-3HB; residues 924–995) (**Figure 62a**). One Cγ2 domain in IgG1 Fc

(annotated as chain A in our structure) binds across the termini of the "V", in-between the GH domain and CBM, with the rest of the antibody remaining exposed to the surrounding solvent. Previous work investigating the ability of EndoS to cleave the *N*-linked glycans from various Fc fragments in comparison to full-length IgG showed EndoS was able to cleave the majority of glycans in all instances, indicating that the IgG Fab regions are unimportant in complex formation [261]. However, the possibility that EndoS interacts with the IgG Fab regions cannot be excluded based on this crystal structure alone.

The *N*-linked glycan on this chain is "flipped-out" from the usually-observed position between the two IgG Fc Cγ2 domains [36] and is bound within the previously-identified glycosidase domain cavity [258] (**Figure 62b**). A structural overlay with full-length EndoS in complex with its G2 oligosaccharide substrate (PDB code 6EN3 [258]) shows that the overall morphology and domain organisation of EndoS is approximately maintained (**Figure 63a**), apart from a slight shift of the CBM and C-3HB, likely due to a pinching of the CBM around the Fc as it binds.

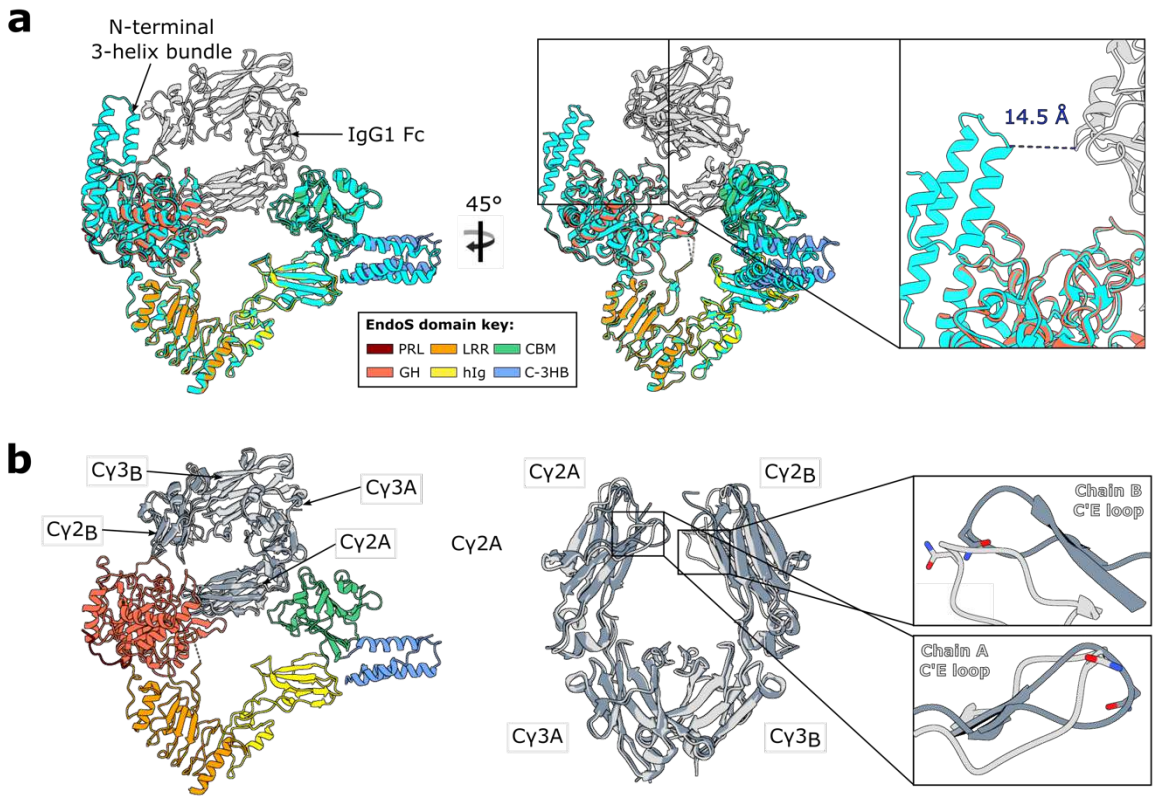


Figure 63: Superposition of EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex with respective apo structures. a Superposition of full-length EndoS (PDB code 6EN3, coloured in cyan) with complexed EndoS$^{D233A/E235L}$. Focused view of N-terminal 3-helix bundle (not present in the complexed EndoS$^{D233A/E235L}$ construct) indicates that it would not bind IgG1 Fc. **b** Superposition of wild-type IgG1 Fc (PDB 3AVE; coloured in dark grey) with complexed IgG1 Fc^{E382R}, which reveals conformational changes in the C′E loops containing the *N*-linked glycans at N297 (whose side chain is shown as sticks and coloured by heteroatom). **a, b** Complexed EndoS$^{D233A/E235L}$ and IgG1 Fc^{E382R} are coloured as in **Figure 62a**. *N*-linked glycans have been omitted for clarity.

Role of CBM in governing specificity of EndoS for IgG

The structure of the EndoS-Fc complex reveals how one Cγ2 domain of the Fc binds across the GH domain and CBM (**Figure 62**). As calculated by PDBePISA [85], the interface between chain A of the Fc and EndoS comprises an area of 1323.5 Å^2 and yields a solvation free energy gain of -9.1 kcal/mol. The GH domain of EndoS is observed forming contacts with the glycan-containing C′E loop, while the CBM forms additional interactions at the Fc Cγ2–Cγ3 interface (**Figure 62b**). Residue W803 within the CBM of EndoS, whose substitution to an alanine has previously been shown to abolish hydrolytic activity against all human IgG subclasses [257], appears to act as a hydrophobic "plug"; it binds within a cavity at the Fc Cγ2–Cγ3 interface containing residues I253, H310, L314 and H435 (**Figure 62b**), and has the highest solvation energy (of 2.06 kcal/M) of all EndoS residues calculated by PDBePISA, indicating that strong hydrophobic interactions are present here. A small number of contacts are also predicted between EndoS and the second Fc Cγ2 domain, although these are unlikely to be necessary for complex formation, given that EndoS can cleave the Fc Cγ2 lacking the hinge region, which is likely to be monomeric [261].

The complex structure presented here corroborates previous findings that both the GH domain and the CBM are important for IgG Fc binding to EndoS [257] and resulting glycan hydrolysis [261], and that EndoS can cleave the Cγ2 homodimer fragment of IgG Fc [261]. While substitution of residues within the glycan binding site of EndoS (and the related enzyme EndoS2) abolishes hydrolytic activity [258, 259], EndoS lacking the CBM can still hydrolyse IgG, albeit at greatly reduced capacity [257, 261]. Therefore, the CBM appears to drive additional specificity of EndoS for the Fc peptide surface.

Interestingly, although the name "carbohydrate-binding module" was assigned to the EndoS CBM based on homology to a legitimate carbohydrate-binding domain [257] and previous work has indicated that it has the capacity to bind galactose (albeit with low affinity) [261], here the CBM is not observed to bind the Fc *N*-linked glycan. Furthermore, the N- and C-terminal 3 helix bundles, which are homologous to IgG-binding protein A from *Staphylococcus aureus* [258, 410], are not interacting with the substrate polypeptide within this complex. A structural overlay of complexed EndoS with full-length EndoS (PDB code 6EN3) indicates that the N-terminal bundle would not contact the Fc (**Figure 63a**), thus its contribution to EndoS-IgG binding and glycan hydrolysis is likely solely due to stabilisation of the GH domain, as suggested previously [258]. Indeed, existence of the crystal structure is evidence in itself that EndoS forms a stable complex with IgG in the absence of the N-terminal bundle.

Stoichiometry of the EndoS-IgG Fc complex

Within the crystal, each of the two IgG Fc chains binds a distinct EndoS molecule, resulting in a complex with 2:2 stoichiometry (**Figure 64a**). The first EndoS molecule (chain C) is binding chain

A of the Fc, and clear electron density for the Fc N297 glycan binding within the EndoS GH domain cavity is observed, which was identified with a previous crystal structure [258] (**Figure 64b**). The polder map for this carbohydrate group, calculated in *PHENIX* [396, 397], supports the presence of an uncleaved *N*-linked glycan in the substrate binding pocket of EndoS, as discussed in section 5.2.1 (**Figure 58b**). The observation of an Fc glycan in this "flipped-out" conformation is in strong contrast to typical crystal structures of IgG Fc, whose *N*-linked glycans are interspersed between the Cγ2 domains [36] (**Figure 64b**).

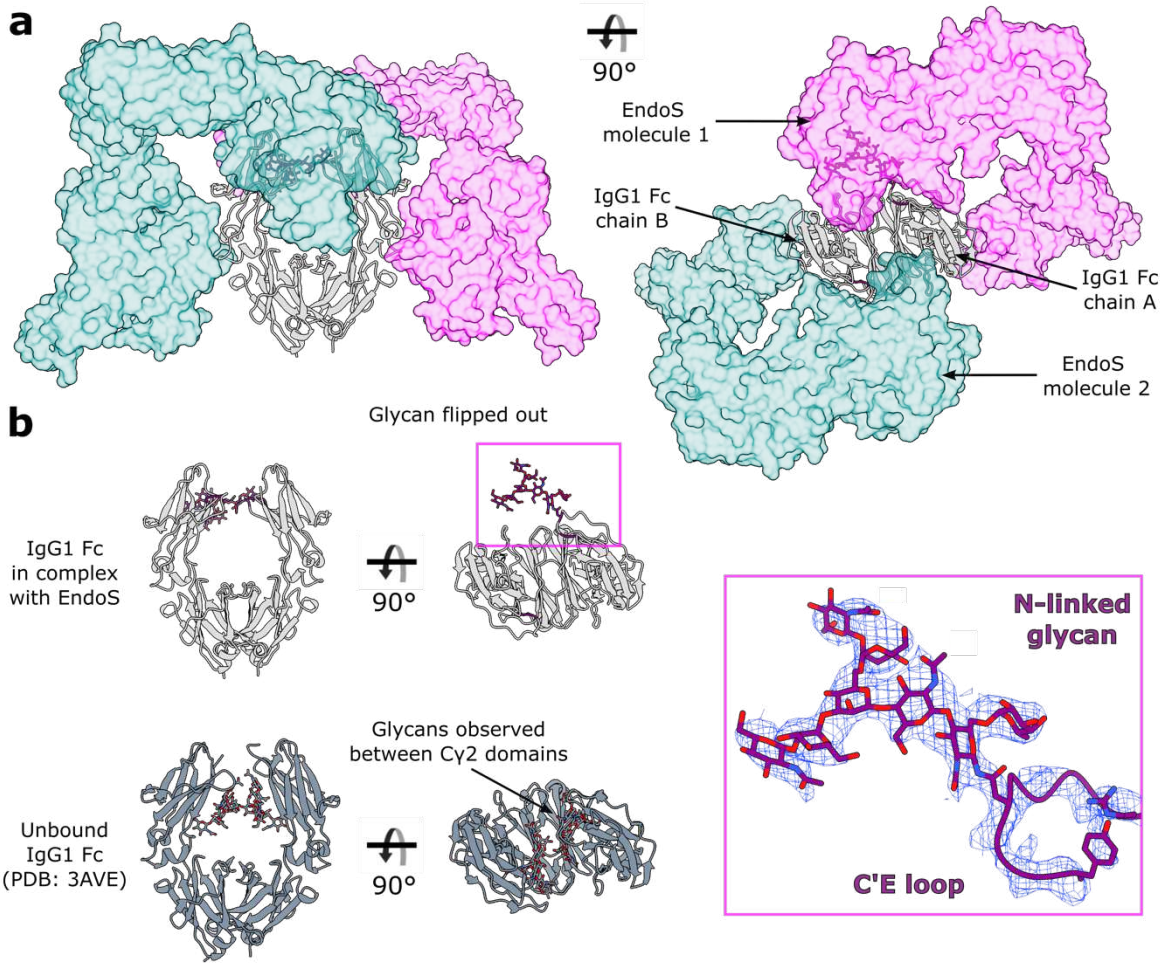


Figure 64: Stoichiometry of the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex. a One EndoS$^{D233A/E235L}$ molecule interacts with one chain within IgG1 Fc^{E382R}, resulting in an overall 2:2 binding stoichiometry. The two EndoS molecules are coloured teal and magenta and depicted as a surface at 50 % transparency, while the Fc is coloured silver and depicted as a cartoon. C′E loops within the Fc are coloured purple; the *N*-linked glycan is depicted as sticks and coloured by heteroatom (oxygen, red; nitrogen, blue). **b** Comparison of *N*-linked glycan positions observed in IgG1 Fc^{E382R} bound to EndoS$^{D233A/E235L}$, and a wild-type IgG1 Fc structure (PDB code 3AVE, coloured in dark grey). *N*-linked glycan is observed in a "flipped-out" structure in the complexed Fc^{E382R}, while *N*-linked glycans in typical Fc structures are observed between the Fc Cγ2 domains. Electron density from the final $2F_{obs}$-F_{calc} map, corresponding to the glycan and C′E loop, is shown (contoured at 1.1 σ).

Chain B of the Fc appears to be binding a second EndoS molecule (chain D; **Figure 64a**); however, this EndoS molecule is poorly resolved in the electron density, with higher *B* factors and a greater proportion of residues identified as RSRZ outliers (**Figure 57**; section 5.2.1). Moreover, the Fc N297

glycan in chain B is not fully visible in the electron density, although the second EndoS molecule appears to bind this Fc chain in the same manner as its more ordered counterpart (**Figure 64a**). Electron density for the Fc C′E loop in chain B (albeit less clear than chain A) also indicates that the glycan is in close proximity to the GH domain in the second EndoS molecule.

It is fascinating to observe the glycan trapped in this "flipped-out" conformation, and this substantiates several recent studies documenting the existence of IgG Fc glycan conformational heterogeneity [115-119]. Superposition of this complexed IgG with a wild-type Fc (PDB 3AVE) illustrates that movement of the glycan into this position is governed by movement of the C′E loop only (**Figure 64b**), although it is possible that the lower resolution of the data is masking small chain shifts. The observation of a "flipped-out" glycan conformation may also provide a structural explanation for the ability of cellular glycosidases and glycosyltransferases to process this otherwise sterically-restricted substrate. Moreover, it appears that the capture of Fc *N*-linked glycans in this state allows space for two enzymes to bind the Fc homodimer simultaneously (**Figure 64**); however, there is no evidence to suggest that this 2:2 assembly is required for activity, especially given previous work showing that EndoS is largely monomeric in solution [257, 258]. Although EndoS crystallised here is lacking the N-terminal 3-helix bundle, a structural superposition with full-length EndoS (**Figure 63a**) suggests 2:2 binding would be able to occur in its presence.

5.3.2 EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex

The EndoS2-IgG1 Fc complex was solved in space group $P4_32_12$, with three copies of the complex modelled within the AU, as detailed in section 5.2.3. As observed for the EndoS-IgG1 Fc structure (**Figure 62**), one copy of EndoS2 binds a Cγ2–Cγ3 monomer within the Fc homodimer. In addition, one copy of the complex is not as well-resolved in the electron density (annotated as chains E and B, for the copies of EndoS2 and half Fc, respectively) and thus not fully built into the density. This is also inferred from a *B* factor distribution of the AU content (**Figure 60b**). The best-resolved copy of the complex (annotated as chains D and A, for the EndoS2 and Fc copies, respectively) is therefore used for depiction and analysis of the structure. The final model was refined to 3.0 Å and is displayed in **Figure 65**.

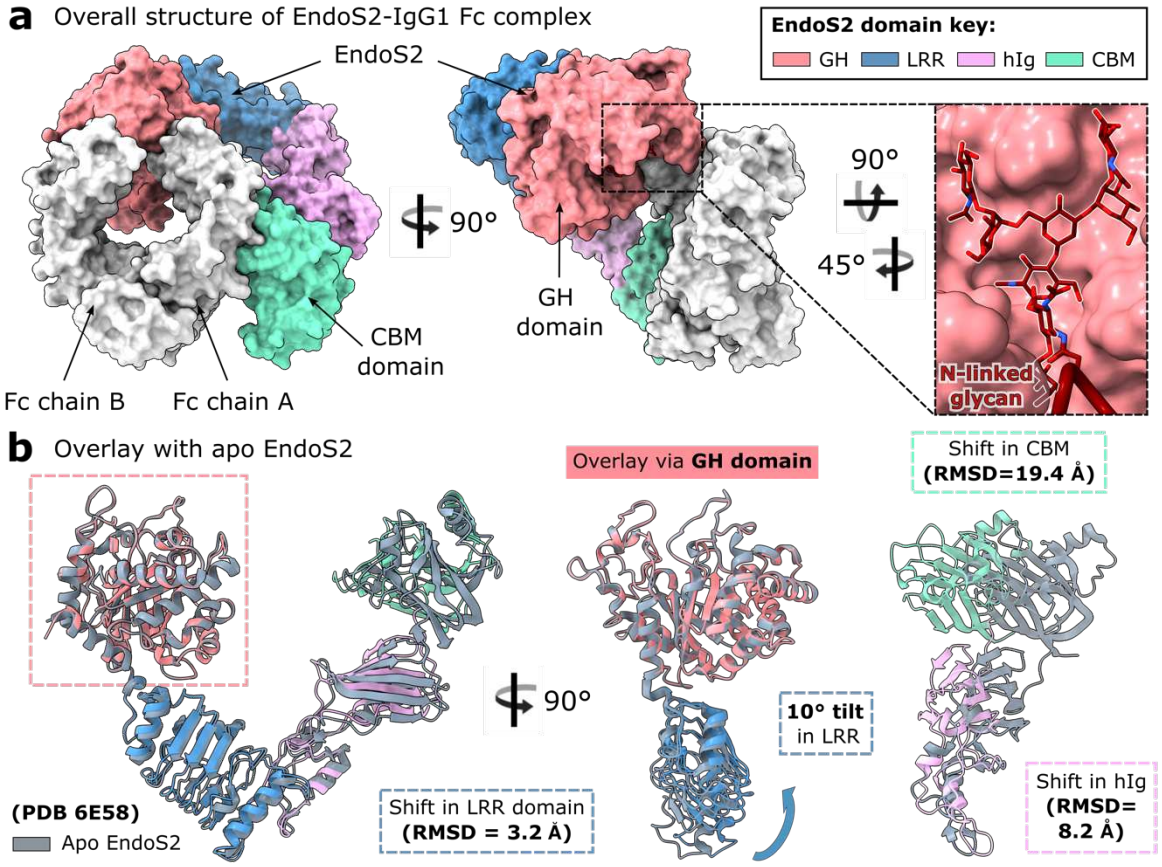


Figure 65: Overall structure of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. a Surface representation of the EndoS-IgG1 Fc complex. Two EndoS2 molecules bind each Fc homodimer in the crystal (see **Figure 66**); for clarity, only one EndoS2 molecule is shown here. IgG1 Fc is depicted in silver, with the N297 glycan depicted as maroon sticks and coloured by heteroatom (oxygen, red; nitrogen, blue). EndoS2 domains are coloured as follows: glycoside hydrolase (GH), coral; leucine-rich repeat (LRR), blue; hybrid immunoglobulin (hIg), lilac; carbohydrate-binding module (CBM), green. **b** Superposition of apo EndoS2 (PDB 6E58) and EndoS2 in complex with IgG1 Fc, with respect to the GH domain (aligned Cαs for EndoS2 residues 46–386). IgG1 Fc in complexed EndoS2 is omitted here for clarity. The figure reveals the shift of GH and CBM relative to each other, with the entire GH–LRR–hIg–CBM scaffold in different orientation. RMSDs and plane angle differences for EndoS2 domain shifts after these superpositions are indicated, as calculated in UCSF ChimeraX [328].

EndoS2 comprises a glycosylhydrolase domain (GH, coral, residues 43–386), a leucine-rich repeat domain (LRR, blue, residues 387–547), a hybrid immunoglobulin domain (hIg, lilac, residues 548–680) and a carbohydrate binding module (CBM, green, residues 681–843) [259]. The enzyme is shaped like a "V", and captures one Cγ2 domain of the Fc through binding by the GH and CBM domains (**Figure 65**), similar to that observed in the structure of the EndoS-Fc complex. Each Fc γ-chain is seen to interact with one EndoS2 molecule, similar to what was observed with the EndoS-Fc crystal lattice (**Figure 66**).

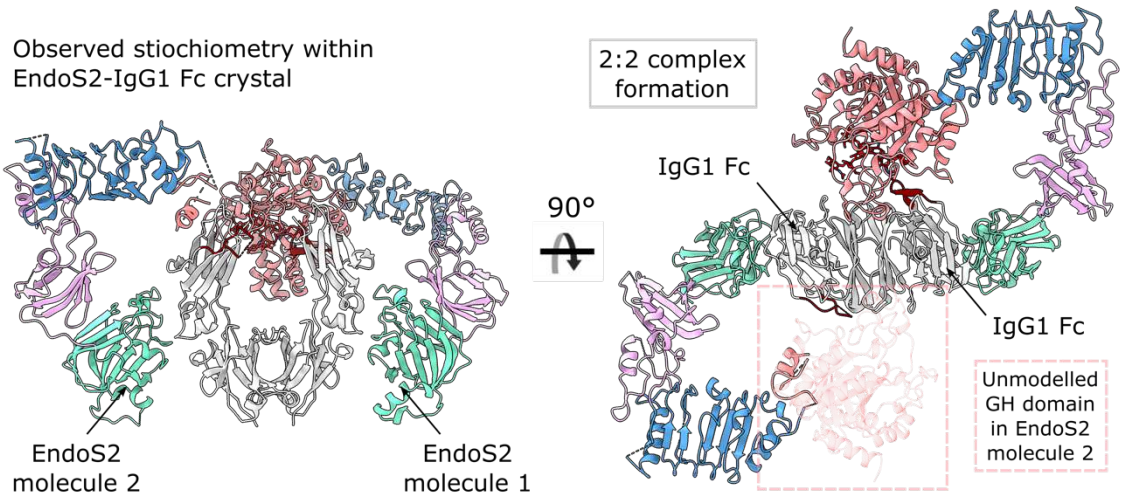


Figure 66: Stoichiometry within EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystal structure. One EndoS2 molecule binds a single chain of the IgG1 Fc homodimer within the EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystal structure, resulting in a 2:2 stoichiometry as also observed for the EndoS-IgG1 Fc complex. EndoS2 and IgG1 Fc are coloured as in **Figure 65**.

In the EndoS2-IgG1 Fc complex, the uncleaved Fc *N*-linked glycan is captured in a "flipped-out" conformation and bound within the GH domain groove (**Figure 65**a), which was identified by co-crystals of EndoS2 in complex with glycan substrates [259]. This observation is in stark contrast to the typically-observed position of Fc glycans within crystal structures, in which they appear to sit in-between the Cγ2 domains and interact with Fc surface residues [36, 113, 114, 128]. The carbohydrate is well-ordered, displaying low *B* factors with respect to the average *B* factor of the complex (105.8 Å^2, **Table 26**, section 5.2.3). A polder map additionally displays clear density consistent with an uncleaved glycan in this conformation (**Figure 61**, section 5.2.3). Overall, this demonstrates the same mode of glycan capture as observed in the EndoS-IgG1 Fc complex, and further corroborates literature reporting heterogeneity of Fc *N*-linked glycan conformations [115-119].

Superposition of complexed and apo EndoS2 (PDB 6E58) reveals a shifting of multiple domains upon Fc binding (**Figure 65b**). An overlay with respect to the GH domain (calculated by aligning Cαs for residues 46–386) results in a 10° tilt of the LRR (with respect to its position in the apo structure). Consequently, the hIg and CBM are also displaced (with RMSDs of 8.2 Å and 19.4 Å, respectively). An overlay with respect to the CBM (calculated by aligning Cαs for residues 681–843) results in smaller domain shifts: the adjacent hIg domain is tilted by 4.2°, while the LRR and GH are displaced by 5.1 Å and 6.3 Å, respectively. Therefore, both the GH and CBM domains are rearranged with respect to the rest of the enzyme upon Fc binding. The conformation of the LRR–hIg scaffold is also affected: an overlay with respect to the LRR domain (calculated by aligning Cαs for residues 387–547) results in a 11.6° tilt in the hIg. Thus, the relative domain positioning within EndoS2 is shifted upon binding the Fc substrate, although discrete domains can be fully superimposed and so do not undergo individual conformational changes. Analysis of apo and glycan-bound EndoS2 crystal structures has revealed no conformational changes upon glycan binding [259],

which indicates that EndoS2 domain shifts observed here are solely due to binding the Fc peptide surface.

The binding interface of the EndoS2-IgG1 Fc complex is analysed in **Figure 67**. The GH and CBM domains within EndoS2 both contact the same Fc Cγ2 domain. This corroborates hydrogen-deuterium exchange data of the EndoS2-rituximab complex, which suggested a role for both domains in complex formation [259]. Crystal structures of EndoS2 in complex with oligomannose- and complex-type glycan substrates have identified GH amino acids involved in glycan binding [259]; the EndoS2-Fc complex structure additionally reveals interactions with the first GlcNAc and fucose moieties within the glycan (**Figure 67b**). The aromatic side chains of F189, Y251, Y252 and W297 provide hydrophobic stacking interactions, while the Y252 backbone and Q255 side chain form hydrogen bonds with the fucose (**Figure 67b**).

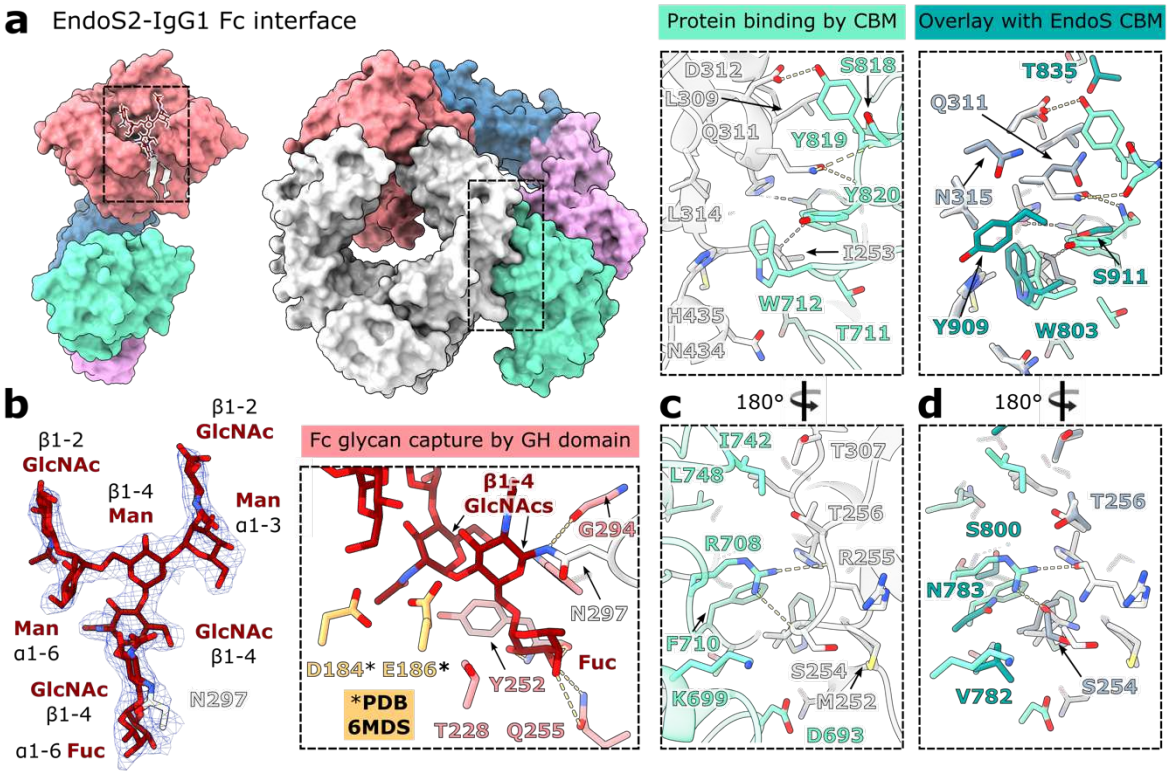


Figure 67: Binding interface of the EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. **a** Overall structure of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex, with interfacing areas highlighted. **b** *N*-linked glycan capture by the EndoS2 GH domain. Final $2F_{obs}–F_{calc}$ electron density map for the *N*-linked glycan and residue N297 is shown (contoured at 1.5 σ). Approximate positions of catalytic residues D184 and E186 at the active site are indicated (from an overlay with wild-type EndoS2 (PDB 6MDS) via the GH domain: aligned Cαs of residues 43–386). **c** Fc protein surface binding by EndoS2 CBM. EndoS2 and IgG1 Fc are coloured as in Figure 1; residues involved in binding are shown as sticks and coloured by heteroatom. Hydrogen bonds are depicted as yellow dashes. **d** Overlay of EndoS2 CBM interface with EndoS CBM (PDB 8A49). EndoS residues are coloured teal; IgG1 Fc complexed with EndoS is coloured dark grey.

Atypical ϕ and ψ torsion angles are observed for the β1–4 linkage between the two core GlcNAcs of 5.9° and –121.2°, respectively, using C1–O–C(x)′–C(x–1)′ as the definition of ψ. These values lie outside the range of average glycosidic linkages reported in crystal structures (–73.7° ± 8.4 for ϕ;

116.8° ± 15.6 for ψ angles, calculated from 163 structures) [411]. In a catalytic context, the distortion of the torsion angles at the site of hydrolysis is consistent with promoting cleavage of the β1–4 linkage. Such glycan distortion is also observed within the EndoS-Fc complex, with equivalent ϕ and ψ torsion angles of –58.6° and –121.2°, respectively. In addition, several crystal structures of endoglycosidases in complex with their cleaved glycan substrate show the second GlcNAc adopting a higher-energy, skew-boat conformation, and so glycan distortion has been suggested to comprise part of the catalytic cycle in such enzymes [259, 304, 412]. As observed for the structure of EndoS2 bound to an isolated, complex-type glycan (PDB 6MDS [259]), the second GlcNAc is bound close to the catalytic dyad residues D184 and E186 (exchanged to alanine and leucine in the structure of the EndoS2-Fc complex, respectively; **Figure 67b**). An overlay of PDB 6MDS with the EndoS2-Fc complex via the GH domain indicates that, in this approximate position, E186 is oriented in close proximity to the distorted β1–4 linkage between the two GlcNAcs, facilitating nucleophilic attack and subsequent cleavage of the glycan.

The CBM interface in the EndoS2-IgG1 Fc complex is analysed in **Figure 67c**. Amino acid W712 in EndoS2 has been shown to be important for Fc binding, as substitution to alanine abolished hydrolytic activity towards IgG Fc bearing complex-type or oligomannose-type glycans [259]. The EndoS2-Fc structure reveals how W712 binds within a hydrophobic pocket at the Fc Cγ2–Cγ3 interface, formed by I253, H310, L314, N434 and H435. Thus, W712 makes similar interactions as W803 in EndoS, similarly shown to be essential for enzyme activity [257] and observed binding in the same cavity within the EndoS-Fc crystal structure (**Figure 67d**). Another important amino acid in EndoS2 is Y820: Klontz *et al*. found that a serine exchange variant of this side chain severely reduced hydrolytic activity towards IgG bearing both complex- and oligomannose-type glycans [259]. This side chain makes not only a main chain contact to Fc residue Q311, but also forms stacking interactions with the Q311 side chain (**Figure 67c**). The equivalent amino acid in EndoS is a serine, similarly providing a main chain interaction with Q311 but lacking the hydrophobic interaction. A further parallel between the two structures is observed with the side chain of EndoS2 Y819 that hydrogen bonds to the D312 side chain from the Fc, similar to the equivalent interactions of the EndoS T835 side chain.

5.4 Comparison of EndoS/EndoS2 recognition of IgG1 Fc

Comparison of the EndoS2 and EndoS CBM interfaces with IgG1 Fc reveals some unique interactions formed by EndoS2, such as the side chain of R708 that hydrogen bonds to the Fc backbone (**Figure 67c**, **d**). PISA analysis [85] also predicts distinct Fc protein surface interfaces across the two enzymes: the EndoS2-IgG1 Fc interface has a surface area of 977.9 $Å^2$ and a solvation free energy gain of –8.6 kcal/mol. Conversely, EndoS forms a 1323.5 $Å^2$ interface with the Fc, with a –9.1 kcal/mol solvation free energy gain. A sequence alignment of EndoS and EndoS2 reveals that

side chains interfacing with the Fc protein are not particularly conserved (**Figure 68**). Moreover, aside from residue W712 (W803 in EndoS), other CBM residues found to be important for IgG Fc binding (R908 and E833 in EndoS [257]; Y820 in EndoS2 [259]) are not conserved (**Figure 68**). These observations indicate distinct modes of IgG recognition by the two endoglycosidases.

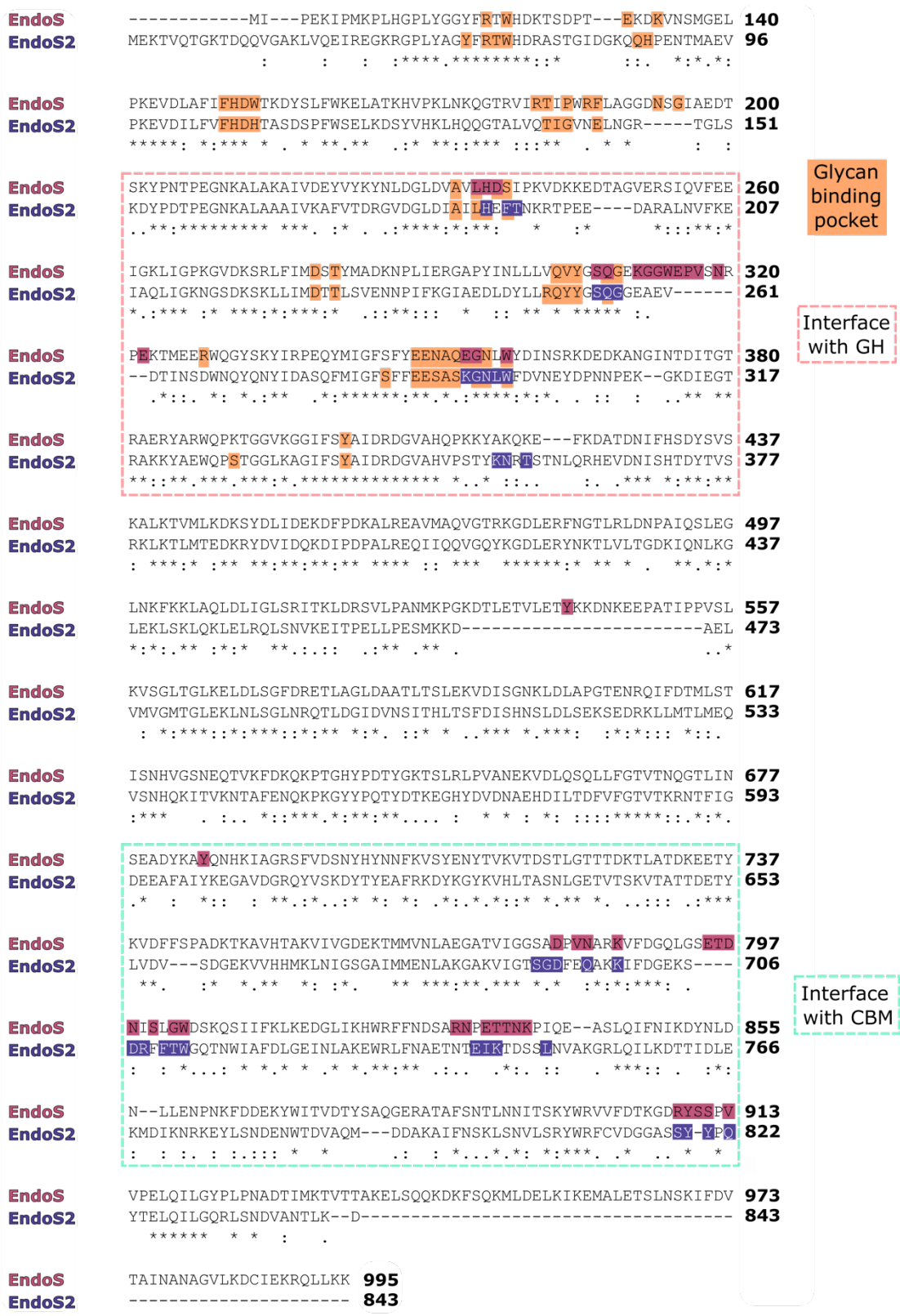


Figure 68: Multiple sequence alignment of EndoS2 and EndoS protein sequences. Sequences are shown as present in crystal structures of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ (PDB 8Q5U) and EndoS-IgG1 Fc (PDB 8A49). Interfaces of EndoS2 and EndoS with the IgG1 Fc protein surface are coloured purple and maroon, respectively, with interfaces formed by the GH domain and CBM of each enzyme indicated with dashed lines. Interfaces of each enzyme with their glycan substrate are coloured in orange. Sequence alignment was generated by Clustal Omega [35].

Although the two enzymes bind IgG1 Fc in a similar way, a comparison of the EndoS2-Fc and EndoS-Fc crystal structures reveals differing angles of recognition between the two enzymes. Superposition of the two complexes with respect to the interfacing Fc Cγ2 domain (calculated by aligning Cαs of amino acids 237–340) shows how different GH domain and CBM orientations allow the enzymes to discover different surfaces on the Fc peptide (**Figure 69a**). This is accompanied with distinct Fc surface interfaces recognised by the two enzymes (**Figure 69b**). Differential capture of the Fc by the two enzymes is also observed: superposition of EndoS2- and EndoS-bound Fc by their Cγ3 domains (alignment of Cαs for amino acids 341–444) reveals more "open" Cγ2 domain placements in the former (**Figure 69c**). However, an overlay of the interfacing Cγ2 (alignment of Cαs for amino acids 238–340) within the complexed Fcs and the equivalent domain from a wild-type IgG1 Fc structure (PDB 3AVE) shows that conformational changes only occur in the C′E loop, allowing capture of the Fc glycan in a "flipped-out" state (**Figure 69b**). A superposition of the GH domains from EndoS- and EndoS2-Fc complexes shows that the interfacing Fc Cγ2 domains are captured in different orientations (**Figure 70**), which may reflect the altered Cγ2 domain placements observed here.

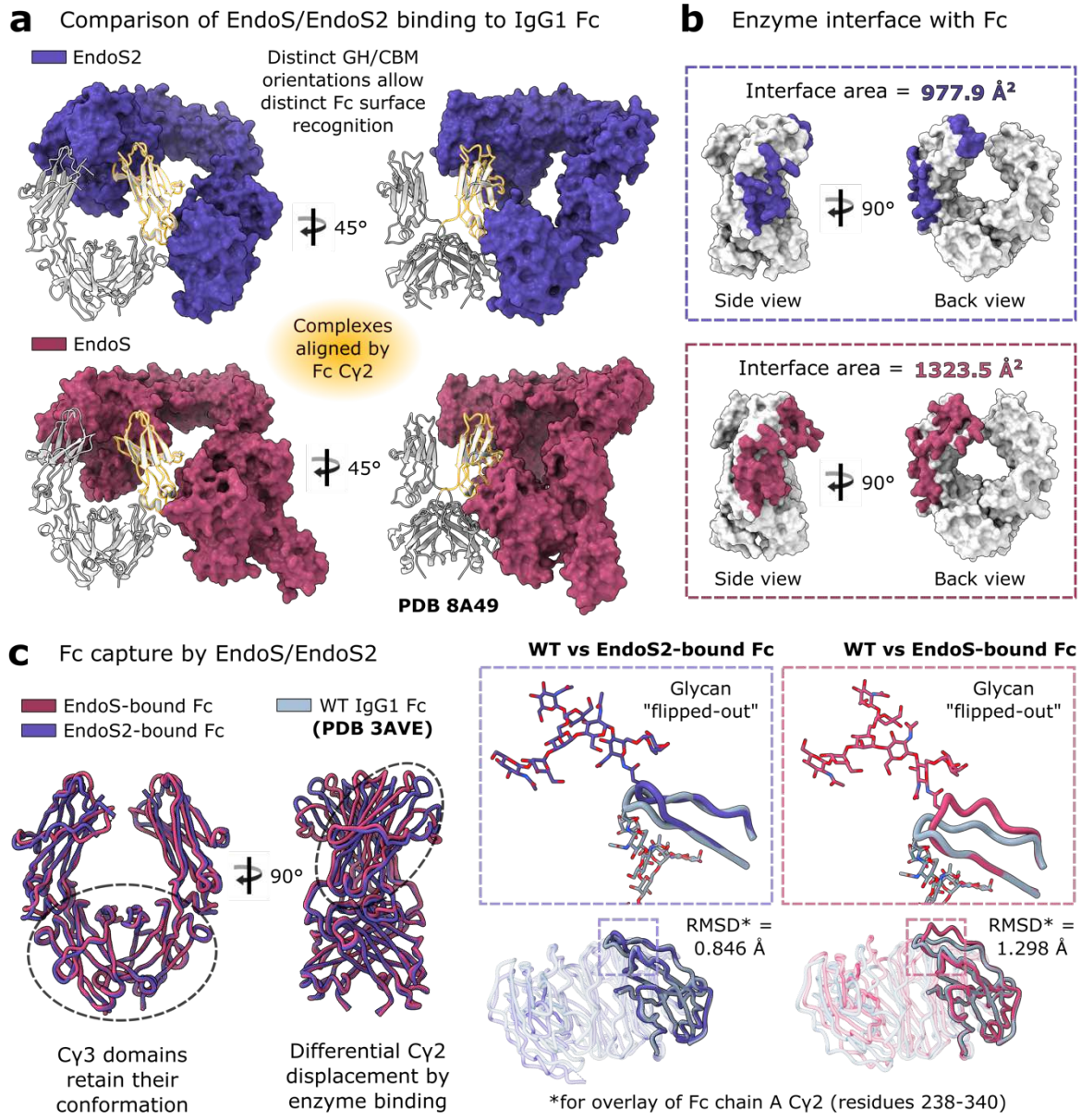


Figure 69: Comparison of EndoS and EndoS2 binding their IgG1 Fc substrate. **a** EndoS2-Fc and EndoS-Fc crystal structures, superimposed with respect to the interfacing Fc Cγ2 domain (Cαs for residues 237–340). Complexed EndoS and EndoS2 are depicted in maroon and purple, respectively; complexed IgG1 Fc is depicted in silver. **b** Binding interfaces of each enzyme mapped onto IgG1 Fc, as calculated by PDBePISA [85]. **c** Comparison of Fc capture by EndoS and EndoS2. Superposition of EndoS-bound and EndoS2-bound Fc relative to the Cγ3 homodimer (Cαs for residues 341–444) reveals differential shifting of Cγ2 domains upon enzyme binding. Fc *N*-linked glycans (shown as sticks and coloured by heteroatom) are captured by both enzymes in a "flipped-out" conformation, occurring with shifts in the C′E loop which is covalently linked to the glycan (relative to a wild-type IgG1 Fc structure: PDB 3AVE, depicted in dark grey). EndoS-bound and EndoS2-bound Fc are coloured maroon and purple, respectively.

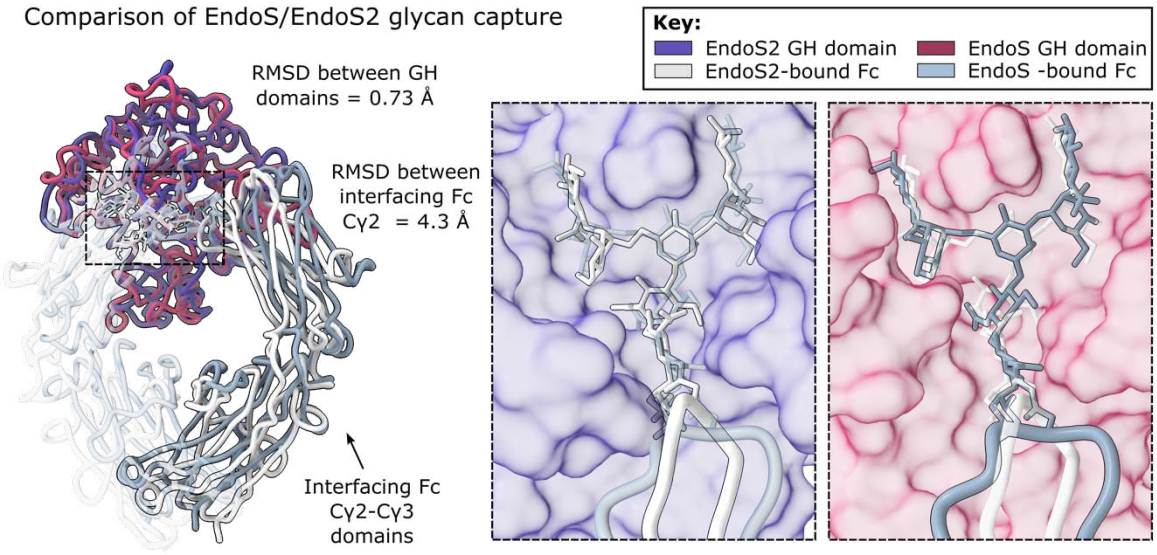


Figure 70: Superposition of GH domains from complexes of EndoS2 and EndoS (PDB 8A49) with IgG1 Fc. Complexes were superimposed in ChimeraX [328] in relation to GH domains (amino acids 46–386 for EndoS2; amino acids 113–445 for EndoS), with a calculated RMSD of 0.73 Å between 273 pruned atom pairs. EndoS and EndoS2 GH domains are coloured as in **Figure 69**; EndoS- and EndoS2-bound IgG1 Fc are coloured dark grey and silver, respectively. Fc *N*-linked glycans are shown as sticks.

The observed differences in EndoS and EndoS2 binding to IgG1 Fc, in distinct protein-protein interfaces and differing angles of recognition for the GH and CBM domains (**Figure 69**), provide a structural rationale for experiments with chimeric EndoS/EndoS2 enzymes, which showed that substitution of the EndoS2 GH onto an EndoS scaffold was not enough to confer EndoS2-like activity, but rather required substitution of both EndoS2 GH and CBM domains [259]. This chimeric enzyme was still not fully functional with respect to wild-type EndoS2 [259], thus the LRR–hIgG scaffold must also play a role in orienting the GH and CBM domains optimally. The EndoS2-Fc crystal structure can help guide future efforts for synthesis of chimeric enzymes, which have application in engineering antibody glycosylation. Enzymes using EndoS/EndoS2 scaffolds are of particular interest due to their exquisite specificity for IgG antibodies: for example, a recent paper by Fan *et al*. shows how the glycosidase domains in EndoS/EndoS2 can be replaced with an α-L-fucosidase from *Lactobacillus casei* BL23 for significantly enhanced antibody defucosylation, when compared to the native fucosidase [413].

5.5 Discussion

This chapter describes work towards determining the structural basis behind the unique specificity of two endoglycosidases secreted by the *S. pyogenes* bacterium: EndoS and EndoS2. The "less-crystallisable" IgG1 Fc variants described in chapter 3 were used in an attempt to co-crystallise IgG1 Fc in complex with both EndoS and EndoS2. Crystallisation using two different constructs of EndoS were attempted: full-length, inactive EndoS (EndoSFL$^{D233A/E235L}$) and truncated, inactive EndoS

(EndoS$^{D233A/E235L}$). The equivalent catalytic residues were exchanged in EndoS2 to produce an inactive construct of this enzyme (EndoS2$^{D184A/E186L}$).

All three proteins were successfully expressed and purified from *E.coli* BL21 (DE3)pLysS cells. Each of the expressed EndoS/EndoS2 constructs were subsequently combined with IgG1 Fc variants, in order to produce complexes for crystallisation experiments. Despite previous work indicating that EndoS behaves as a monomer in solution [257, 258], combining EndoS and IgG1 Fcs in a 1:1 molar ratio did not produce a single protein species as judged by SEC. Successful crystallisation of the EndoS-IgG1 Fc complex confirmed that, within the crystal, two copies of EndoS were binding a single Fc homodimer. Therefore, in subsequent purifications of EndoS and EndoS2, the enzyme was combined with IgG1 Fcs in a 2:1 molar ratio. When purified by SEC, these 2:1 complexes eluted as a single peak, indicating stable complex formation. The solved structure of EndoS2 in complex with IgG1 Fc later also confirmed its 2:1 binding stoichiometry in the crystal lattice. However, for both EndoS and EndoS2, binding events of one enzyme molecule interacting with one Fc Cγ2-Cγ3 monomer appear to be independent (i.e., EndoS/EndoS2 are not dimeric).

The truncated EndoS-IgG1 Fc and EndoS2-IgG1 Fc complexes both crystallised readily, with hits in many conditions across multiple crystallisation screens. However, upon sample irradiation, it was found that the majority of these crystals diffracted very poorly. Two rounds of seeding further crystallisation plates produced EndoS-Fc crystals diffracting to 3.45 Å. The EndoS2-Fc complex was much more difficult to solve; seeding techniques did not improve the resolution and neither did most attempts to reduce flexibility of the Fc. A single condition was found for inactive EndoS2 in complex with a Fc engineered to contain an additional disulphide bridge in the lower hinge region (conferred by L234C exchanges), which produced crystals diffracting to ~3 Å and enabled structure solution. Crystallisation of the full-length EndoS-Fc complex was unsuccessful and produced no initial hits; further attempts to crystallise this complex could be attempted with the full range of "less-crystallisable" Fcs.

Prior to the start of this project, crystal structures of EndoS and EndoS2 in complex with their isolated glycan substrates had been elucidated, allowing identification of the glycan-binding grooves present within the glycosidase (GH) domain [258, 259]. This domain, along with a putative carbohydrate-binding module (CBM), was known to be important for enzymatic activity in both EndoS and EndoS2 [257-259, 261]. The crystal structures of EndoS-IgG1 Fc and EndoS2-IgG1 Fc presented within this chapter corroborate these findings, by showing that the GH and CBM domains, which form the termini of these V-shaped enzymes [257-259], bind across one Cγ2 domain of IgG Fc. The role of the CBM is therefore to drive additional specificity for the enzyme towards its IgG substrate. This provides a structural rationale for the observation that EndoS unable to hydrolyse the glycans from denatured IgG Fc [227]. A recent cryo-EM structure of EndoS in complex with IgG Fc also corroborates these findings [414].

Examination of both crystal structures revealed that the Fc *N*-linked glycans have been captured in a "flipped-out" conformation, due to conformational changes within the Fc C′E loops. This is an exciting observation and has not been previously seen within an IgG Fc crystal structure: Fc *N*-linked glycans are typically observed between the Cγ2 domains, where they interact with particular Fc residues [36, 113, 114, 128]. However, the flexible nature of IgG Fc glycans has been studied using other techniques [115-119]; the crystal structures presented here therefore corroborate the existence of Fc glycan conformational heterogeneity. This also provides a structural rationale for the ability of these enzymes, and other glycosidases, to access the (apparently) sterically-restricted glycan substrate.

A comparison of the interfaces of EndoS/EndoS2 binding to IgG1 Fc reveals, despite the overall similarity in binding modes, distinct modes of IgG recognition. The LRR–hIgG scaffold of each enzyme, although not involved in substrate recognition or catalysis, appear to be important in orienting the GH and CBM domains for optimal binding to their respective interfaces on IgG Fc. An EndoS scaffold with the GH domain from EndoS2 was found to be non-functional, but subsequent substitution of the CBM from EndoS2 restored partial EndoS2-like activity [259]. The crystal structures presented here provide a structural rationale for these experiments, by revealing that the interfaces formed by the enzyme binding to the Fc protein surface are not as highly conserved as those formed with the Fc *N*-linked glycan. Moreover, these structures can be utilised as templates for future structure-guided design of novel IgG-specific chimeric enzymes, as recently demonstrated by Fan *et al.* [413], in which a fucosidase enzyme was fused onto EndoS scaffold (lacking its catalytic domain), to yield an IgG-specific fucosidase. Given the co-dependence of the GH domain and CBM within these endoglycosidases [259], and the role of the LRR–hIgG scaffold in orienting these domains for optimal IgG binding, simple substitution of a catalytic domain onto an IgG-specific scaffold may not yield a functional enzyme. Structural information for two distinct IgG-specific enzyme scaffolds (from EndoS and EndoS2) provides two possible templates for chimeric enzyme design, which should greatly aid such ventures.

Overall, the crystal structures of EndoS and EndoS2 in complex with IgG1 Fc presented here reveal the molecular basis behind their extensive substrate specificity. This structural knowledge will aid in the continued development of enzyme variants, and chimeric EndoS/EndoS2-based enzymes, as tools for improved antibody glycoengineering and as biologics for use in immune modulation.

Chapter 6 Conclusions

The aim of this thesis was to elucidate the structural basis behind the unique specificity of IdeS and EndoS/EndoS2 for IgG antibodies, which provides the bacterium *Streptococcus pyogenes* with potent evasion mechanisms against the human immune system. The successful expression, purification, crystallisation and structure solution of the inactive enzymes in complex with IgG1 Fc is described.

Chapter 3 presents a panel of IgG1 Fc variants designed to be "less-crystallisable", by substitution of amino acid E382, which forms an important salt bridge with R255 in typical Fc crystal lattice interactions. A crystal counting experiment was utilised to confirm that the three Fc variants (containing E382R, E382S and E382A exchanges) crystallise less readily than a wild-type Fc. Structural analysis of IgG1 Fc^{E382S} and Fc^{E382A} showed similarly altered crystallisation, with both variants crystallising in space group $P3_221$, not previously reported for a human, apo Fc structure. These Fc variants were hypothesised to be of use for co-crystallisation experiments of enzyme-Fc complexes; successful crystallisation of these complexes demonstrates their suitability for this purpose, and highlights their potential use for other Fc-protein complexes, such as complexes with FcγRs.

The IdeS-IgG1 Fc crystal structure confirms the long-suspected belief that IdeS must recognise IgG using exosite binding, as it shows extensive binding across both chains of the hinge region and both Fc Cγ2 domains. The enzyme is observed encasing the IgG hinge region to be cleaved within the active site cavity, mediated by the so-called "prosegment binding loop". Binding of the hinge peptide at the active site distorts the scissile peptide bond to promote cleavage. IdeS cleaves both chains of the hinge sequentially and the structure indicates an altered binding interface for the second chain cleavage. Further work needs to be undertaken to understand the structural basis of this, along with the ability of IdeS to additionally cleave the IgG2, IgG3 and IgG4 subclasses.

Similarly, the EndoS-IgG1 Fc and EndoS2-IgG1 Fc crystal structures confirm the importance of not only the catalytic glycosylhydrolase domain, but also the carbohydrate-binding module for EndoS/EndoS2 activity. The role of the CBM was previously unclear and proposed to be in glycan recognition; however, the structures presented here elucidate its role in Fc protein surface recognition, and thus in driving specificity towards IgG. This mode of binding was corroborated shortly after publication of the EndoS-Fc crystal structure (PDB 8A49) with a cryo-EM structure of the EndoS-IgG1 Fc structure. Moreover, these crystal structures observe the Fc *N*-linked glycan in a "flipped-out" conformation, corroborating previous molecular dynamics simulations and providing a structural rationale for the ability of these enzymes, and likely other endoglycosidases, to process this relatively sterically-restricted substrate.

Comparison of the EndoS and EndoS2-Fc structures additionally reveals differences in Fc recognition, providing a structural rationale for recent experiments showing how the catalytic domain from EndoS2 cannot be readily substituted with that from EndoS. This is important to note for the design of chimeric enzymes for use in antibody glycan remodelling. The high specificity of EndoS and EndoS2 for IgG makes them ideal scaffolds for IgG glycoengineering, and the structural information presented herein will aid in their development as chimeric enzymes.

In conclusion, the crystal structures presented within this thesis provide a structural rationale for the unique properties of these immune evasion enzymes, particularly their exquisite substrate specificity towards human IgG. Furthermore, these structures can serve as templates for structure-guided design of novel enzymes, such as immunologically-distinct IdeS/endoglycosidase variants for long-term therapeutic use, and development of novel IgG-specific endoglycosidases with varying glycosyltransferase activities for precise antibody glycoengineering.

Appendix A Publication 1

Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. *Nat Commun.* 2022;13:7801.

Publication is included overleaf for reference:

nature communications

Article

https://doi.org/10.1038/s41467-022-35340-z

Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS

Received: 16 August 2022

Accepted: 25 November 2022

Published online: 17 December 2022

Check for updates

Abigail S. L. Sudol[1], John Butler[1], Dylan P. Ivory[1], Ivo Tews[1] & Max Crispin[1]

Enzymatic cleavage of IgG antibodies is a common strategy used by pathogenic bacteria to ablate immune effector function. The *Streptococcus pyogenes* bacterium secretes the protease IdeS and the glycosidase EndoS, which specifically catalyse cleavage and deglycosylation of human IgG, respectively. IdeS has received clinical approval for kidney transplantation in hypersensitised individuals, while EndoS has found application in engineering antibody glycosylation. We present crystal structures of both enzymes in complex with their IgG1 Fc substrate, which was achieved using Fc engineering to disfavour preferential Fc crystallisation. The IdeS protease displays extensive Fc recognition and encases the antibody hinge. Conversely, the glycan hydrolase domain in EndoS traps the Fc glycan in a "flipped-out" conformation, while additional recognition of the Fc peptide is driven by the so-called carbohydrate binding module. In this work, we reveal the molecular basis of antibody recognition by bacterial enzymes, providing a template for the development of next-generation enzymes.

The bacterium *Streptococcus pyogenes* has evolved a diverse range of mechanisms for evading the human adaptive immune system[1]. Infection with *S. pyogenes* can be mild, causing for example throat infections, but at the other extreme can cause terminal necrotising fasciitis[2]. Two enzymes secreted by this bacterium, IdeS[3] and EndoS[4], directly target and cleave IgG antibodies, and thereby impede cellular responses through immune recruitment mediated by the antibody Fc domain. The specificity of these enzymes for IgG has led to the development of a wide range of clinical and biotechnology applications[5] and has warranted extensive studies of their enzymology.

Of the two immune evasion factors, IdeS is most advanced in clinical development[6–8]. *S. pyogenes* expresses two variants of this enzyme (often distinguished by naming the first and second variants IdeS/Mac-1 and Mac-2, respectively), which display less than 50 % sequence identity within the middle third of the protein[9], but nonetheless exhibit largely indistinguishable endopeptidase activity[10]. The enzyme targets IgG by cleaving within the lower hinge region, yielding $F(ab')_2$ and Fc fragments[3,11,12], an activity which has enabled its development (specifically, the Mac-1 enzyme variant) as a pre-treatment for transplantation in hypersensitised individuals with chronic kidney disease (Imlifidase, brand name Idefirix®)[6–8]. Along with EndoS, it has further potential use in the deactivation of pathogenic antibodies in autoimmune disorders[13–19], deactivation of neutralising antibodies for in vivo gene therapy[20], and for the potentiation of therapeutic antibodies by deactivation of competing serum IgG[21,22]. Imlifidase has also been used in combination with EndoS for inactivation of donor-specific antibodies in murine allogeneic bone marrow transplantation[23].

The endoglycosidase EndoS has additional biotechnological applications in engineering antibody glycosylation:[24] it hydrolyses the β−1,4 linkage between the first two N-acetylglucosamine (GlcNAc) residues within biantennary complex-type N-linked glycans on IgG Fc, thereby removing the majority of the glycan[4]. The related enzyme EndoS2 from serotype M49 of *S. pyogenes* also targets IgG[25] but exhibits broader glycan specificity[26]. Variants of both enzymes have been utilised in transglycosylation of various glycoforms to intact IgG to enable precise antibody glycan remodelling[24,27–29].

[1]School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK. e-mail: Ivo.Tews@soton.ac.uk; Max.Crispin@soton.ac.uk

It is still unclear, however, how exactly these enzymes specifically target and degrade IgG. Full cleavage of an antibody by IdeS occurs in two distinct steps, in which the second chain is cleaved more slowly;[11,12] this observation, along with the finding that IdeS exhibits low activity towards synthetic hinge peptides[30], suggests a more extensive recognition interface with the target IgG. Similarly, multiple domains within EndoS contribute to substrate recognition and catalysis[31–33], but the molecular details of substrate recognition remain undefined.

Here, we illustrate the molecular basis behind the unique substrate specificity of these enzymes using X-ray crystallography. We show that mutation of IgG Fc residue E382, which consistently forms salt bridge interactions in Fc crystal structures, can be used as a strategy to discourage Fc self-crystallisation and thus promote crystallisation of protein complexes. We present crystal structures of IdeS/IgG1 and EndoS/IgG1 complexes, to a resolution of 2.34 Å and 3.45 Å, respectively, and map the extensive interfaces that are formed in these complexes. Understanding substrate interaction and recognition of these enzymes facilitates their further clinical development and their application as highly specific biotechnological tools.

Results and discussion

Analysis of Fc crystal structures for Fc engineering

The co-crystallisation of IgG Fc with enzymes is notoriously difficult, due to the inherent ability of the Fc fragment to crystallise on its own. We therefore sought to identify favourable contacts present in typical Fc crystals, in order to devise a strategy to counteract its selective self-crystallisation.

We have observed, from looking at structures currently present in the PDB, that human IgG Fc commonly crystallises in the $P2_12_12_1$ space group (60.2 % of 108 apo IgG Fc structures, as of August 2022). We studied the crystal lattice contacts present in a typical, wild-type Fc structure (PDB 3AVE[34]), in order to identify amino acid residues which are important in this favourable packing arrangement. As calculated in PDBePISA[35], model 3AVE forms thirteen salt bridges and fifteen hydrogen bonds with neighbouring molecules within its crystal lattice (Fig. 1b). In addition, contacts are largely conserved across both Fc chains, resulting in a tight packing arrangement (Fig. 1b, c). We identified residue E382, which forms salt bridges with R255 in a neighbouring Fc molecule (and vice versa), in both Fc chains (Fig. 1b). We hypothesised that replacement of this residue would hinder the self-association of the Fc into this preferred crystal lattice, and therefore designed three IgG1 Fc variants: E382R, E382S and E382A, which we collectively term as "Fx" variants.

In order to compare the crystallisation abilities of our "Fx" variants versus a wild-type IgG1 Fc, we set up identical crystallisation experiments in JCSG-plus™ and Morpheus screens (both from Molecular Dimensions) for each Fc at 10 mg/mL. Crystals were left to grow at 21 °C and, after eight days, the number of conditions in each screen producing crystals were counted. Of the three Fx variants, we illustrate results here for the E382S variant. Wild-type IgG1 Fc produced a total of 21 crystal "hits" (14 in JCSG-plus™ and 7 in Morpheus). In contrast, the E382S variant produced 9 hits in total, all of which were in JCSG-plus™ (Fig. 1a); this variant therefore displayed a ~57 % reduction in crystallisation compared to the wild-type Fc. The E382A and E382R variants similarly produced no hits in the Morpheus screen, and yielded 8 and 6 hits in JCSG-plus™, respectively.

Crystals of the Fc E382S variant were found to have grown in an atypical space group $P3_221$. The structure was determined by molecular replacement using 3AVE as a search model and refined to a resolution of 3.04 Å (Supplementary Table 1, Supplementary Fig. 1). As of August 2022, this space group has not previously been reported for a human apo IgG Fc structure; we did, however, find examples of IgG Fc crystallised in complex with a small peptide (for example, PDB 5DVK[36]) where the reported space group was $P3_221$. Interestingly, this peptide binds at the Cγ2-Cγ3 interface within the Fc, around the same area as E382.

Analysis of the crystal contacts revealed that this variant makes fewer interactions with symmetry-related molecules in the crystal (four salt bridges and sixteen hydrogen bonds; Fig. 1b), which are asymmetrical across the two Fc chains, resulting in altered crystal packing (Fig. 1c). Furthermore, as calculated within the ccp4i2 interface[37], the E382S variant had a higher solvent content and Matthews coefficient compared to the wild-type Fc (Fig. 1b), indicating that the molecules are less tightly packed in this crystal form. We conclude that crystallisation of this Fc variant has been rendered less favourable; this indicates that other substitutions at E382, as in the E382R and E382A variants, would have similarly altered crystallisation. In general, we envisage that any mutation impacting lattice formation could be similarly employed. We subsequently used these E382 variants for screening of enzyme-Fc complexes, and believe that they would similarly be ideal for crystallisation of other Fc complexes, such as those of IgG with Fc γ-receptors.

The IdeS-IgG1 Fc complex

IdeS from *Streptococcus pyogenes* (strain MGAS15252), containing a C94A mutation to abolish catalytic activity, was combined in a 1:1 molar ratio with our panel of IgG1 Fc variants, and the resulting complexes were purified by size exclusion chromatography (Supplementary Fig. 2). We obtained crystals of IdeS in complex with the Fc^{E382A} variant, which crystallised in space group C121 (Supplementary Table 2). The structure was determined by molecular replacement with 1Y08 and 3AVE search models. The data for this crystal is twinned, with a refined twin fraction of 0.493 for operator *-h,-k,l*, as determined using twin refinement in Refmac5[38]. Such a high twin fraction means that significant model bias is to be expected, and care must be taken in model analysis (see Supplementary Table 2 and Supplementary Fig. 4).

Electron density resolves amino acids 43-339 in IdeS, as well as 229-445 and 230-444 for chains A and B in IgG1 Fc, respectively. We additionally observe density for seven/eight monosaccharide residues at the N-linked glycosylation site (at N297) on Fc chains A and B, comprising a fucosylated biantennary glycan with a single β−1,2-linked GlcNAc on the mannose 6-arm (chain A) and the equivalent glycan with terminal β−1,2-linked GlcNAc on both arms (chain B). The final structure was refined to 2.34 Å (Supplementary Table 2, Supplementary Fig. 3) and is depicted in Fig. 2.

The crystal structure shows asymmetric binding of IdeS across the Cγ2 domains of the Fc and its lower hinge region (Fig. 2a). We envisage that the upper hinge region of IgG and its Fab regions do not contribute significantly to complex formation, as indicated with the lack of electron density for the hinge above residues 229/230 (in Fc chains A and B, respectively), and reported cleavage of both full-length IgG and its Fc fragment by IdeS[30]. However, we cannot formally exclude the possibility that there is some interaction of IdeS with the IgG Fab regions.

The 1:1 stoichiometry observed in the crystal structure is consistent with previous kinetic analyses[39] showing that IdeS functions predominantly in a monomeric form. The enzyme appears to clamp down over the lower hinge region of one Fc chain (Fig. 2a), creating a cavity in which the catalytic residues are brought into close proximity with the cleavage site. Binding of the enzyme to the Fc appears to displace the two Cγ2 domains slightly, as shown by superposition with a structure of wild-type IgG1 Fc (PDB 3AVE) (Supplementary Fig. 4b). Residues within the Cγ2 domain in chain A have higher *B* factors compared to the rest of the complex (Supplementary Fig. 4a), which could indicate that binding of IdeS pulls this domain away slightly from the rest of the antibody.

Role of prosegment binding loop in IdeS-Fc complex

IdeS crystallised in complex with IgG1 Fc here is the Mac-2 variant, and thus deviates in sequence from the three published apo structures of IdeS (all of which are the Mac-1 variant; Supplementary Fig. 5a). Despite this, a structural alignment shows very few deviations (Fig. 2b). Complexed IdeS contains ten α-helices and twelve β-strands, as calculated

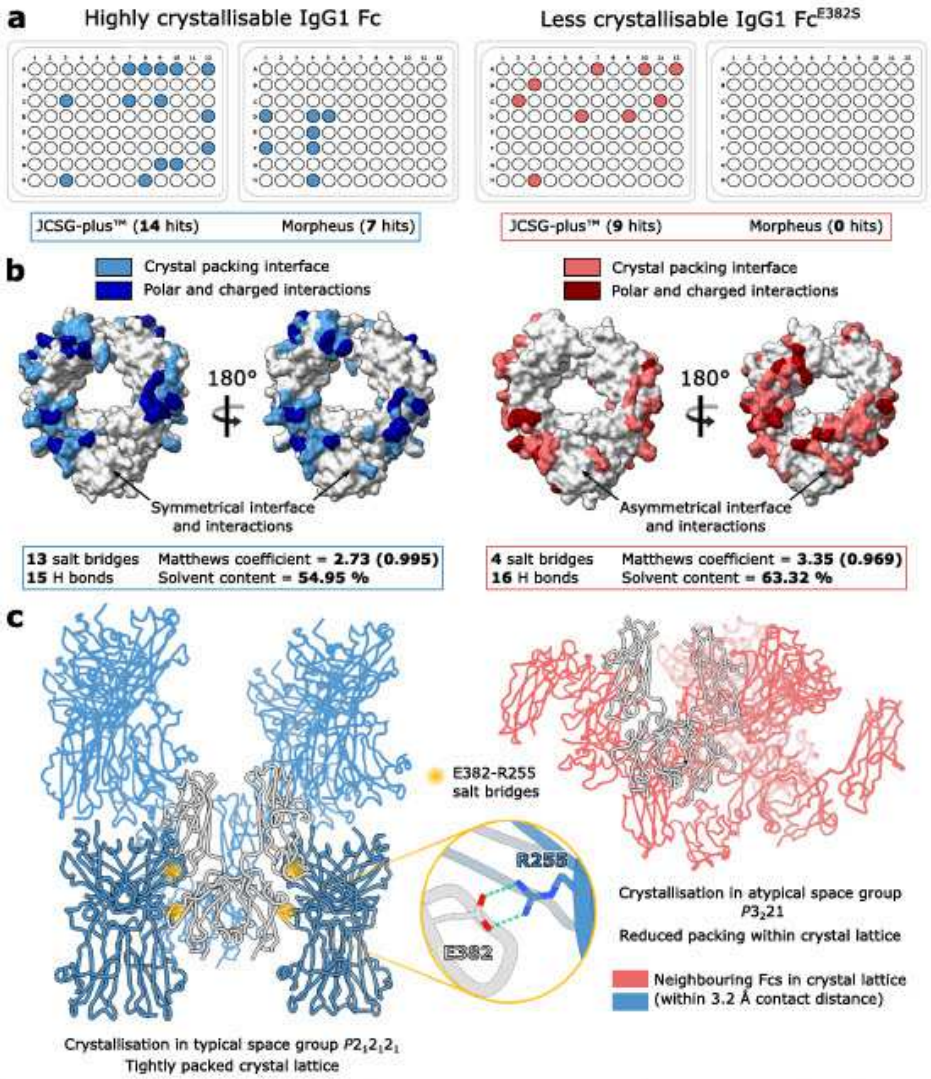


Fig. 1 | Observed crystal packing in wild-type and "less crystallisable" IgG1 Fc fragments. **a** Crystallisation of wild-type IgG1 Fc and IgG1 Fc^{E382S} variant in JCSG-plus™ and Morpheus screens, at 10 mg/mL and 21 °C. Crystal "hits" are indicated with a coloured circle. **b** Analysis of crystal packing interface and interactions present in a typical, wild-type IgG1 Fc crystal structure (PDB ID 3AVE) and IgG1 Fc^{E382S} variant, as calculated by PDBePISA[35]. **c** Crystal packing resulting from crystallisation in typical space group $P2_12_12_1$ and atypical space group $P3_221$, for the wild-type IgG1 Fc and IgG1 Fc^{E382S} variant, respectively. E382-R255 salt bridges between symmetry-related Fcs in the $P2_12_12_1$ crystal lattice are highlighted with a yellow circle. Neighbouring Fcs in the crystal lattice contacting the origin Fc within a 3.2 Å contact distance are shown. **a**–**c** Analysis relating to the wild-type IgG1 Fc and IgG1 Fc^{E382S} variant is depicted in blue and red, respectively.

by DSSP[40,41] (Supplementary Fig. 4c); we note that the loop located between β-strands seven and eight is modelled in distinct conformations for each of the apo structures[42] and is not included within 1YO8[43] (Fig. 2b), signifying its inherent flexibility in the apo form. This loop is equivalent to the "prosegment binding loop" present in other papain superfamily cysteine proteases; in these enzymes, which are synthesised as inactive zymogens, this loop packs against the prosegment as a mode of inhibition[44–46]. In complexed IdeS, the loop curls upwards to accommodate the Fc hinge within the active site cavity (Fig. 2a).

Alanine substitution mutations within this loop were previously found to have little effect on neither IdeS binding to IgG, nor its catalytic activity[42]. Our structure shows, however, that the majority of interactions present here involve the IdeS backbone, whose conformation won't be significantly altered by alanine mutations. The inability of IdeS to cleave IgG hinge-mimicking peptides[30] also indicates an occlusion of the active site in the absence of substrate, especially given the strong potential of hydrogen bonding and hydrophobic interactions observed at the Fc hinge (discussed in the following section). It is possible that there may be a conformational change in the active site upon binding; however, superposition of wild-type IdeS (PDB 2AU1) with the complexed enzyme shows the catalytic triad residues in very similar conformations, although we cannot rule this out given the sequence diversity present (Supplementary Fig. 5). Moreover, IdeS^{C94A} has been shown to retain antibody binding and

https://doi.org/10.1038/s41467-022-35340-z

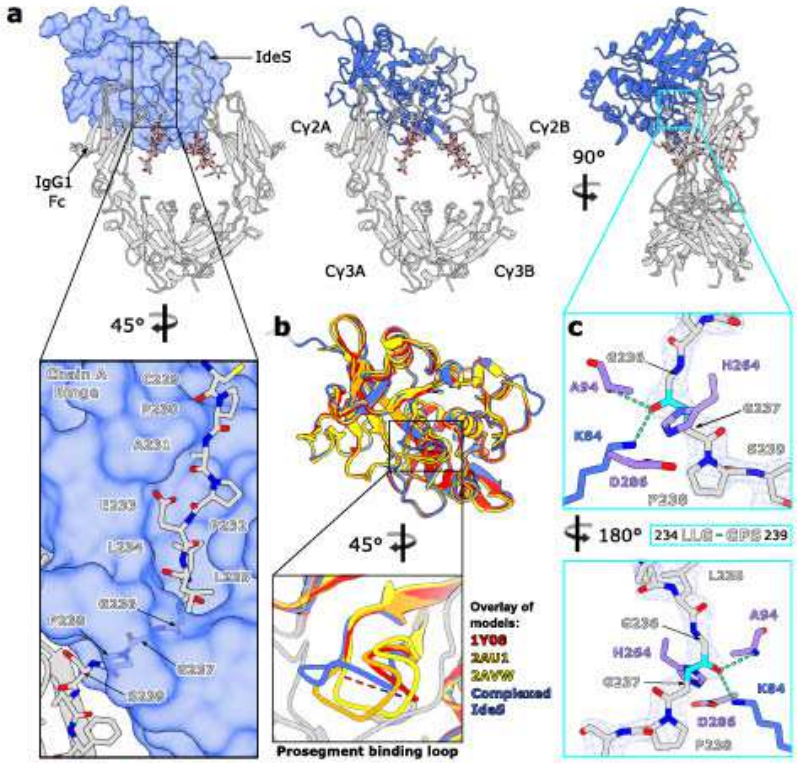


Fig. 2 | Crystal structure of IgG1 Fc^{EM2A}-$IdeS^{C94A}$ complex. **a** Overall structure of complex, with $IdeS^{C94A}$ shown as a surface and IgG1 Fc^{EM2A} shown as a cartoon. N-linked glycans and the IgG1 Fc^{EM2A} hinge peptide in the focused panel are shown as sticks and coloured with oxygen, nitrogen and sulphur atoms in red, blue and yellow, respectively. **b** Superposition of complexed $IdeS^{C94A}$ with three published apo structures of IdeS (PDB IDs 1Y08, 2AU1 and 2AVW, coloured in red, orange and yellow, respectively) and focused view of the prosegment binding loop. **c** Binding of IgG1 Fc^{EM2A} hinge peptide within the $IdeS^{C94A}$ active site. Fc^{EM2A} peptide and $IdeS^{C94A}$ active site residues are depicted as sticks and coloured by heteroatom; catalytic triad residues are coloured purple. The scissile peptide bond is coloured in cyan; hydrogen bonds are depicted as green dashes. The final $2F_{obs}$-F_{calc} electron density map corresponding to the Fc^{EM2A} peptide is shown (weighted at 1.5 σ). **a**–**c** $IdeS^{C94A}$ is coloured blue; IgG1 Fc^{EM2A} is coloured in silver.

inhibit IgG-mediated phagocytosis at levels comparable to the wild-type enzyme[9], suggesting that the inactive enzyme retains antibody binding. We therefore conclude that the likely role of this loop is mediation of substrate access to the active site.

Interaction of IgG1 Fc hinge at IdeS active site

We observe clear density for the IgG1 Fc hinge region bound within the IdeS active site cavity, in both the final electron density map (Fig. 2c) and a polder map as calculated in *PHENIX*[47,48] (Supplementary Fig. 4d). The carbonyl oxygen of G236 is hydrogen bonded to the amide nitrogen of the catalytic cysteine (mutated to alanine here) and the side chain of K84, which collectively form the oxyanion hole, as predicted[42,43]. Binding of the hinge distorts the Fc peptide backbone at G236 in order to promote scissile bond cleavage (Fig. 2c); this residue is thus identified in Molprobity[49] as a Ramachandran outlier. Superposition of wild-type IdeS (PDB 2AU1) with the complexed enzyme gives an indication for placement of the catalytic cysteine side chain (Supplementary Fig. 5b): in this conformation, the cysteine sulphur is ideally poised for nucleophilic attack on the carbonyl carbon within the scissile peptide bond.

Extended exosite binding of IdeS to the Fc Cγ2 domains

It has long been suspected that IdeS must recognise its sole substrate IgG with exosite binding[30,42,43]. Our structure now reveals that IdeS binds across both chains of the Fc region (Fig. 3a). Unsurprisingly, the most extensive interface is formed with the Fc chain being cleaved (annotated as chain A in our structure) (Fig. 3b), with an interface area of 1392 Å^2 and a solvation free energy gain upon interface formation of −16.2 kcal/mol, as calculated by PDBePISA[35]. The interface extends across the entire hinge region (C229-S239; Fig. 3b), with hydrogen bonds formed with the backbone at A231, L234, G236 and G237 and the E233 side chain, and favourable hydrophobic interactions predicted here (inferred by positive solvation energies of hinge residues). Within the Fc Cγ2 domain, IdeS interacts with residues in proximity of the BC loop, which aids in stabilising an "open" conformation of the prosegment binding loop (as discussed above), and additionally the FG loop (Fig. 3b).

A secondary interface is formed across the second Fc chain (annotated as chain B in our structure; Fig. 3c), with an interface area of 804.7 Å^2 and a solvation free energy gain of −7.6 kcal/mol. A smaller proportion of the Fc hinge contributes (A231-G237), but PDBePISA predicts favourable hydrophobic interactions here, albeit not to the same extent as chain A. Subsequent recognition of this Fc chain is driven by interactions with the BC loop, and, in contrast to chain A, the C′E loop containing the N-linked glycan (Fig. 3c). PDBePISA additionally predicts a small number of interactions between the enzyme and the Fc N-linked glycans; the lack of electron density for any monosaccharides past β−1,2-linked GlcNAc suggests that any further glycan

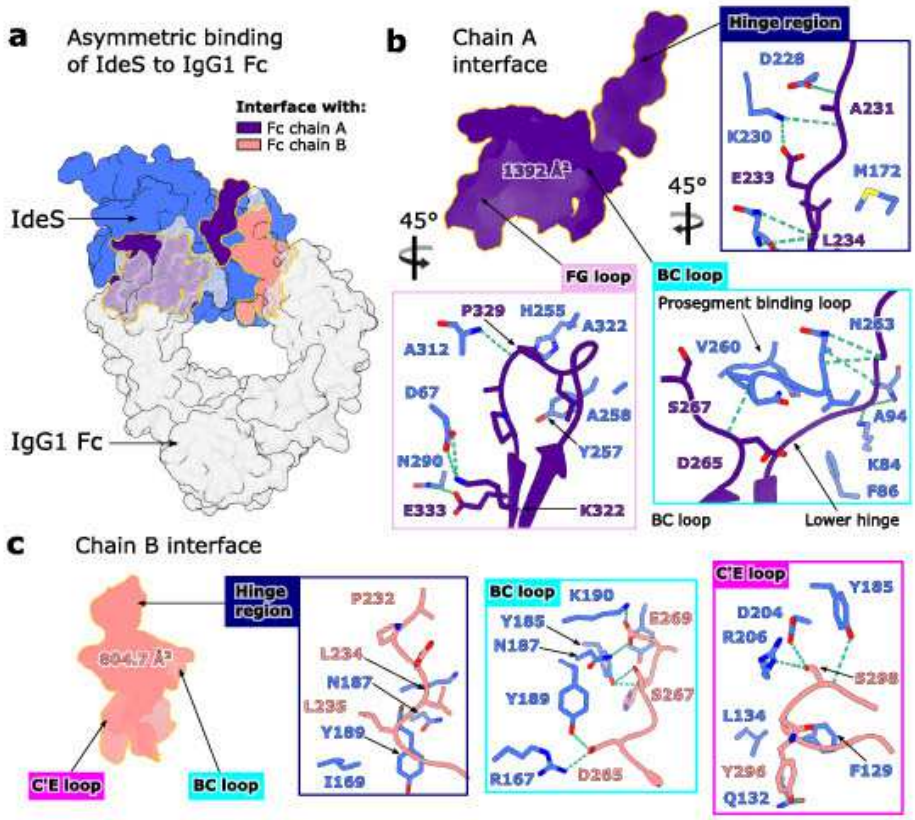


Fig. 3 | Asymmetric binding interface of IdeS^{C94A}-IgG1 Fc^{E382R} complex. a Overall view of complex depicted as a surface, with IdeS^{C94A} coloured blue and IgG1 Fc^{E382A} coloured silver. Interfaces of IdeS^{C94A} with chains A and B of the Fc are coloured indigo and coral, respectively. Glycans within the Fc have been omitted for clarity. **b** Interface between IdeS^{C94A} and IgG1 Fc^{E382A} chain A, involving the Fc hinge region, BC loop and FG loop. **c** Interface between IdeS^{C94A} and IgG1 Fc^{E382A} chain B, involving the hinge region, BC loop and C'E loop. **b**, **c** Residues involved in binding are depicted as sticks and coloured by heteroatom (oxygen in red and nitrogen in blue), with hydrogen bonds depicted as green dashes.

processing doesn't affect complex formation, and that IdeS can accommodate IgG with heterogenous glycosylation.

Although IdeS interacts with both chains in the Fc hinge simultaneously, following cleavage of the first chain, the complex would need to dissociate before the second cleavage could occur. This observation is also evidenced by detection of single-cleaved Fc in enzymatic assays and in clinical studies[11,39,50,51]. We suspect that the binding interface is altered for single-cleaved Fc and that this explains its slower rate of cleavage[11,12,39]. It is also interesting to note that, aside from the hinge region, IdeS binds Fc regions implicit in Fcγ-receptor binding, an observation also inferred by its ability to counteract Fc-mediated effector functions by competitive binding inhibition[9]. Moreover, we observe that IdeS residues interacting with the Fc are largely conserved across the two IdeS isoforms, and any substitutions are mostly to similar amino acids, which aids in explaining their near identical activity[10].

The EndoS-IgG1 Fc complex

To date, there are several known structures of endoglycosidases in complex with their glycan substrates[26,33,52–54]. Here, we present the structure of truncated EndoS (residues 98-995, as described previously[31]) in complex with its IgG1 Fc substrate (E382R variant). A catalytically inactive version of EndoS was generated by the inclusion of D233A/E235L substitutions, as described previously[33]. The complex, containing two copies of EndoS and one IgG1 Fc molecule, crystallised in space group $P2_12_12_1$ and was refined to a resolution of 3.45 Å (Supplementary Table 3, Supplementary Fig. 6). The second copy of EndoS (annotated as chain D) is much more poorly resolved in the electron density compared to the rest of the structure (discussed below), and thus we have used the more ordered copy of EndoS (annotated as chain C) for structure depiction and analysis. In addition, the density resolves only the N-linked glycan on chain A of the Fc (not its counterpart on chain B), which is bound within the more ordered copy of EndoS. We therefore elucidate the mode of glycan binding and IgG recognition by inspecting the interaction of EndoS with this Fc chain. The final structure is depicted in Fig. 4.

Our structure of EndoS shows the same "V" shape as observed in its previously solved structures[31,33]. Following the previously-described domain classification[31], the structure comprises, from the N- to the C-terminus: a proline-rich loop (residues 98-112), a glycosidase domain (residues 113-445), a leucine-rich repeat domain (residues 446-631), a hybrid Ig domain (residues 632-764), a carbohydrate-binding module (CBM; residues 765-923) and a C-terminal three-helix bundle domain (C-3HB; residues 924-995) (Fig. 4a). One Cγ2 domain in IgG1 Fc (annotated as chain A in our structure) binds across the termini of the "V", in-between the glycosidase domain and CBM, with the rest of the

https://doi.org/10.1038/s41467-022-35340-z

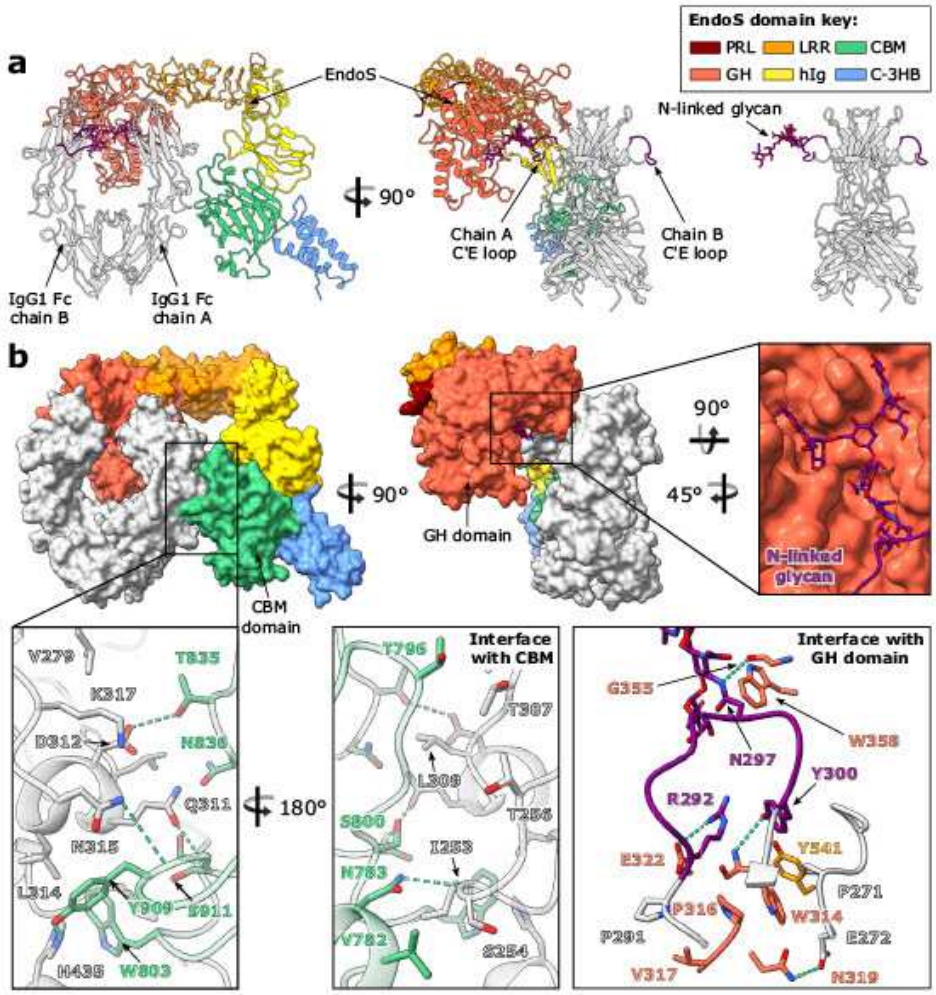


Fig. 4 | Crystal structure of EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex. a Overall structure of complex depicted as a cartoon. IgG1 Fc^{E382R} is coloured silver, with its C'E loops coloured purple; the N-linked glycan is shown as sticks and coloured by heteroatom (oxygen in red and nitrogen in blue). EndoS domains are coloured as follows: proline-rich loop (PRL), maroon; glycosidase domain (GH), red; leucine-rich repeat domain (LRR), orange; hybrid Ig domain (hIg), yellow; carbohydrate-binding module (CBM), green; C-terminal 3-helix bundle (C-3HB), blue. **b** EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex depicted as a surface, highlighting binding to IgG1 Fc^{E382R} by the CBM and GH domains. Residues involved in binding are depicted as sticks and coloured by heteroatom. Hydrogen bonds are depicted as green dashes.

antibody remaining exposed to the surrounding solvent. Previous work investigating the ability of EndoS to cleave the N-linked glycans from various Fc fragments in comparison to full-length IgG showed EndoS was able to cleave the majority of glycans in all instances, indicating that the IgG Fab regions are unimportant in complex formation[32]. However, as with the IdeS-IgG1 Fc structure, we are unable to exclude the possibility that EndoS interacts with the IgG Fab regions from this crystal structure alone.

The N-linked glycan on this chain is "flipped-out" from its usually-observed position between the two IgG Fc Cγ2 domains[34] and is bound within the previously-identified glycosidase domain cavity[33] (Fig. 4b). A structural overlay with full-length EndoS in complex with its G2 oligosaccharide substrate (PDB 6EN3[33]) shows that the overall morphology and domain organisation of EndoS is approximately maintained (Supplementary Fig. 7a), apart from a slight shift of the CBM and C-3HB, likely due to a pinching of the CBM around the Fc as it binds.

Role of CBM in governing specificity of EndoS for IgG

Our structure of the EndoS-Fc complex reveals how one Cγ2 domain of the Fc binds across the glycosidase domain and CBM (Fig. 4). As calculated by PDBePISA[35], the interface between chain A of the Fc and EndoS comprises an area of 1323.5 Å^2 and yields a solvation free energy gain of −9.1 kcal/mol. The glycosidase domain of EndoS is observed forming contacts with the glycan-containing C'E loop, while the CBM forms additional interactions at the Fc Cγ2-Cγ3 interface (Fig. 4b). We note that residue W803 within the CBM, whose substitution to an alanine has previously been shown to abolish hydrolytic activity against all human IgG subclasses[31], appears to act as a hydrophobic "plug"; it binds within a cavity at the Cγ2-Cγ3 interface containing Fc

Appendix A

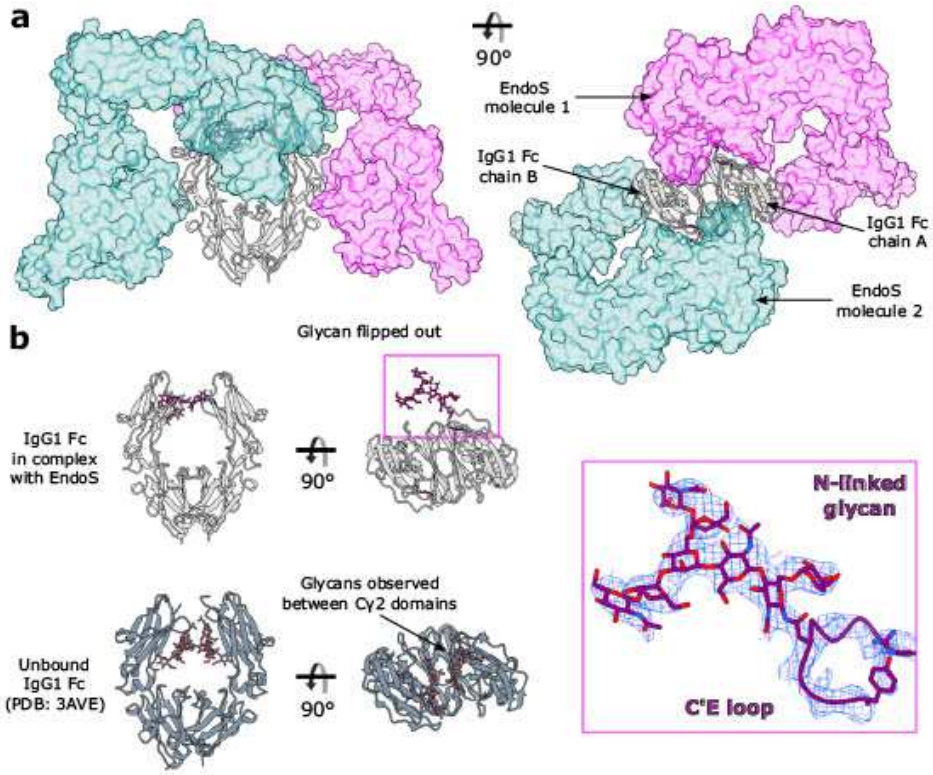


Fig. 5 | Stoichiometry of the EndoS$^{D233A/E235L}$-IgG1 Fc^{E382R} complex. a One EndoS$^{D233A/E235L}$ molecule interacts with one chain within IgG1 Fc^{E382R}, resulting in an overall 2:2 binding stoichiometry. The two EndoS molecules are coloured teal and magenta and depicted as a surface at 50% transparency, while the Fc is coloured silver and depicted as a cartoon. C'E loops within the Fc are coloured purple; the N-linked glycan is depicted as sticks and coloured by heteroatom (oxygen in red and nitrogen in blue). **b** Comparison of N-linked glycan positions observed in IgG1 Fc^{E382R} bound to EndoS$^{D233A/E235L}$, and a wild-type IgG1 Fc structure (PDB ID 3AVE, coloured in dark grey). N-linked glycan is observed in a "flipped-out" structure in the complexed Fc^{E382R}, while N-linked glycans in typical Fc structures are observed between the Fc Cγ2 domains. Electron density from the final $2F_{obs}$-F_{calc} map, corresponding to the glycan and C'E loop, is shown (weighted at 1.1 σ).

residues I253, H310, L314 and H435 (Fig. 4b), and has the highest solvation energy (of 2.06 kcal/M) of all EndoS residues calculated by PDBePISA, indicating that strong hydrophobic interactions are present here. A small number of contacts are also predicted between EndoS and the second Fc Cγ2 domain, although these are unlikely to be necessary for complex formation, given that EndoS can cleave the Fc Cγ2 lacking the hinge region (likely monomeric)[32].

The complex structure presented here corroborates previous findings that both the glycosidase domain and the CBM are important for IgG Fc binding[31] and glycan hydrolysis[32], and that EndoS can cleave the Cγ2 homodimer fragment of IgG Fc[32]. The related enzyme EndoS2 likely binds IgG in a similar manner;[25] hydrogen-deuterium exchange mass spectrometry on this complex has similarly indicated strong binding of IgG to the glycosidase domain and the CBM[26]. While mutation of residues within the glycan binding site of both enzymes completely abolishes their hydrolytic activity[26,33], EndoS lacking the CBM can still hydrolyse IgG, albeit at greatly reduced capacity[31,32]. Therefore, the CBM appears to drive additional specificity of EndoS for the Fc peptide surface.

Interestingly, although the CBM was assigned based on homology to a legitimate carbohydrate-binding domain[31] and previous work has indicated that it has the capacity to bind galactose (albeit with low affinity)[32], here the CBM is not observed to bind the Fc N-linked glycan. Furthermore, the N- and C-terminal 3 helix bundles, which are homologous to IgG-binding protein A from *Staphylococcus aureus*[33,35], are not interacting with the substrate polypeptide within this complex. A structural overlay of complexed EndoS with full-length EndoS (PDB 6EN3) indicates that the N-terminal bundle would not contact the Fc (Supplementary Fig. 7a), thus its contribution to EndoS-IgG binding and glycan hydrolysis is likely solely due to stabilisation of the glycosidase domain, as suggested previously[33]. Indeed, existence of the crystal structure is evidence in itself that EndoS forms a stable complex with IgG in the absence of the N-terminal bundle.

Stoichiometry of the EndoS-IgG Fc complex

Within the crystal, each of the two IgG Fc chains binds a distinct EndoS molecule, resulting in a complex with 2:2 stoichiometry (Fig. 5a). The first EndoS molecule (chain C) is binding chain A of the Fc, and we observe clear electron density for the N297 glycan binding within the EndoS glycosidase domain cavity previously identified[33] (Fig. 5b). The polder map for this carbohydrate group, calculated in *PHENIX*[47,48], supports the presence of an uncleaved N-linked glycan in the substrate binding pocket of EndoS (Supplementary Fig. 8). This observation of an Fc glycan in this "flipped-out" conformation is in strong contrast to typical crystal structures of IgG Fc, whose N-linked glycans are interspersed between the Cγ2 domains[34] (Fig. 5b).

Chain B of the Fc appears to be binding a second EndoS molecule (chain D; Fig. 5a); however, this EndoS molecule is poorly resolved in

the electron density, with higher *B* factors and a greater proportion of residues identified as RSRZ outliers (Supplementary Fig. 9). Moreover, the Fc N297 glycan in chain B is not visible in the electron density, although the second EndoS molecule appears to bind this Fc chain in the same manner as its more ordered counterpart (Fig. 5a). Electron density for the Fc C'E loop in chain B (albeit less clear than chain A) also indicates that the glycan is in close proximity to the glycosidase domain in the second EndoS molecule.

It is fascinating to observe the glycan trapped in this "flipped-out" conformation, and this substantiates several recent studies documenting the existence of IgG Fc glycan conformational heterogeneity[56–60]. Superposition of this complexed IgG with a wild-type Fc (PDB 3AVE) illustrates that movement of the glycan into this position is governed by movement of the C'E loop only (Supplementary Fig. 7b), although it is possible that the lower resolution of the data is masking small chain shifts. The observation of a "flipped-out" glycan conformation may also provide a structural explanation for the ability of cellular glycosidases and glycosyltransferases to process this otherwise sterically-restricted substrate. Moreover, it appears that the capture of Fc N-linked glycans in this state allows space for two enzymes to bind simultaneously; however, there is no evidence to suggest that this 2:2 assembly is required for activity, especially given previous work showing that EndoS is largely monomeric in solution[31,33]. Although EndoS crystallised here is lacking the N-terminal 3-helix bundle, a structural superposition with full-length EndoS (Supplementary Fig. 7a) suggests 2:2 binding would be able to occur in its presence.

The crystal structures presented here provide a structural rationale for the unique properties of these two enzymes, particularly their exquisite substrate specificity towards human IgG. Understanding the molecular basis of this activity is critical for expanding their clinical and biotechnological use. For example, the deactivation of serum IgG using both IdeS and EndoS can strengthen the potency of therapeutic antibodies;[21,22] this strategy could be applied to potentiate any therapeutic antibody, in theory, if the antibody were designed to be resistant to cleavage by these enzymes, a venture which can be aided greatly with structural information. This will also be invaluable in the synthesis of immunologically-distinct enzyme variants which retain identical activity, for their long-term therapeutic use. While EndoS variants have already been designed to expand the ability to engineer antibody glycosylation[27–29], the structural information presented here will allow this to be extended further. We also envisage that this structural information will help in the development of anti-streptococcal biologics resistant to enzyme-mediated degradation. To conclude, this work will assist in the continued development of IdeS and EndoS as enzymatic tools with wide clinical and biotechnological applications.

Methods

Cloning, expression and purification of IdeS/EndoS

Gene fragments encoding IdeS^{C94A} (amino acids 41-339, gene accession number AFC66043.1) and EndoS$^{D233A/E235Q}$ (amino acids 98-995, as described previously[31,33], genome accession number AP012491.2) were synthesised to contain a C-terminal linker and His tag (sequence LEHHHHHH), and cloned into pET21a(+) vectors by NBS Biologicals. Constructs for truncated, inactive IdeS and EndoS used for IgG complex crystallography are shown in Supplementary Fig. 10. Constructs were expressed in *E. coli* BL21 (DE3)pLysS cells (Thermo Fisher). Cells were grown at 37 °C in Terrific Broth (Melford) in the presence of 100 µg/mL ampicillin and 34 µg/mL chloramphenicol, until an OD_{600} of 0.8 was reached, when protein expression was induced by addition of 1 mM IPTG. Cells were left to shake overnight at 25 °C, 200 rpm (Innova 43 R incubator; New Brunswick Scientific). Cells were collected by centrifugation at 6220 × *g* for 20 min, resuspended in PBS containing 2 µg/mL DNAse1 (Sigma) and a pinch of lysozyme (Sigma), homogenised using a glass homogeniser and broken apart using a cell disruptor (Constant Cell Disruption Systems). The remaining sample was centrifuged first at 3100 × *g* for 20 min, then again at 100,000 × *g* for one hour, to remove remaining cell debris and cell membranes. The resulting supernatant was subsequently filtered through a 0.2 µm membrane. Proteins were purified from the supernatant using Ni affinity chromatography with a HisTrap HP column (Cytiva) followed by size exclusion chromatography with a Superdex 75 16/600 column (Cytiva), equilibrated in 10 mM HEPES, 150 mM NaCl, pH 8.0.

Cloning, expression and purification of IgG1 Fcs

IgG1 Fcs were expressed from a pFUSE-hIgG1-Fc vector (encoding residues 221-447 of IgG1; plasmid purchased from InvivoGen). Wild-type IgG1 Fc was expressed exactly as encoded within this plasmid; mutations for the E382A/S/R constructs were introduced by site-directed mutagenesis (QuikChange II kit; Agilent), using mutagenic primers (Supplementary Table 4) synthesised by Eurofins Genomics. Sequences of resulting IgG1 Fc constructs are shown in Supplementary Fig. 11. Fcs were transiently expressed in FreeStyle293F cells (ThermoFisher), using FreeStyle™ MAX Reagent (ThermoFisher), as described in the manufacturer's protocol. Cells were left to incubate at 37 °C, 8% CO_2, shaking at 125 rpm (New Brunswick S41i incubator), and harvested after seven days by centrifugation at 3100 x *g* for 30 minutes. Supernatants were filtered through a 0.2 µm membrane and antibodies purified by affinity purification with a HiTrap Protein A HP column (Cytiva), followed by size exclusion chromatography with a Superdex 200 16/600 column (Cytiva) in 10 mM HEPES, 150 mM NaCl (pH 8.0).

Quantification of IgG1 Fc crystallisation

Wild-type IgG1 Fc and Fc E382A/S/R variants were expressed and purified as detailed above. Identical sitting drop vapour diffusion crystallisation trays were set up using an Oryx4 robot (Douglas Instruments), in JCSG-plus™ and Morpheus crystallisation screens (Molecular Dimensions), using Fcs concentrated to 10 mg/mL. Crystals were left to grow at 21 °C for eight days, after which the number of crystal "hits" were counted (a "hit" constitutes crystals observed growing in a particular condition).

Protein crystallisation

Protein concentrations were determined with a DS-11+ Spectrophotometer (DeNovix), using molecular weight and extinction coefficients calculated by the ProtParam tool[61]. IdeS/EndoS were combined with IgG1 Fcs in a 1:1 molar ratio and applied to a Superdex 200 16/600 column (Cytiva) equilibrated in 10 mM HEPES, 150 mM NaCl (pH 8.0). Fractions corresponding to the main peak only were pooled for crystallisation. Purified complexes were exchanged into 50 mM HEPES, 150 mM KCl (pH 7.5) prior to crystallisation, using a Vivaspin 20 centrifugal concentrator (MWCO 30 kDa; Sigma). Sitting drop vapour diffusion crystallisation trays were set up using an Oryx4 robot (Douglas Instruments). Crystals of the IdeS^{C94A}-IgG1 Fc^{E382A} complex were grown in 0.12 M monosaccharides mix, 0.1 M buffer system 3 (pH 8.5), 30 % v/v precipitant mix 1 (Morpheus crystallisation screen; Molecular Dimensions). Crystals of the EndoS$^{D233A/E235Q}$-IgG1 Fc^{E382R} complex were grown in 0.09 M halogens, 0.1 M buffer system 2 (pH 7.5), 37.5 % v/v precipitant mix 4 (Morpheus crystallisation screen; Molecular Dimensions). Crystals of IgG1 Fc^{E382S} were grown in 0.2 M ammonium sulphate, 0.1 M Tris (pH 7.5), 25 % w/v PEG 8000. Crystals were cryoprotected in mother liquor with 20 % glycerol added and flash-frozen in liquid nitrogen prior to data collection.

Data collection and structure determination

Data collection for the IdeS^{C94A}-IgG1 Fc^{E382A} complex was carried out at the European Synchrotron Radiation Facility (Grenoble, France) on beamline ID30A-3, under a 100 K cryostream (λ = 0.9677 Å). Data

collection for the IgG1 Fc^{E382S} variant was carried out at the European Synchrotron Radiation Facility on beamline ID30A-3, also at 100 K (λ = 0.968 Å). Data collection for the $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} complex was carried out at Diamond Light Source (Oxford, UK) on beamline I03, at 100 K (λ = 0.9763 Å). Data processing of diffraction images was carried out using DIALS[62] and XDS[63]. Structures were solved by molecular replacement with the program Molrep[64]. 3AVE was used as a search model to solve the IgG1 Fc^{E382S} structure; $IdeS^{C94A}$-IgG1 Fc^{E382A} and $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} were solved using initial search models for the enzyme (PDB ID 1Y08 for IdeS; 6EN3 for EndoS), after which the resulting solution was used as a fixed model for a second round of molecular replacement, using 3AVE as the search model. Models were improved with successive rounds of model building and refinement, using Coot[65] and Refmac5[38], respectively, within the ccp4i2 suite[37]. Due to the presence of twinning in the $IdeS^{C94A}$-IgG1 Fc^{E382A} data, this structure was refined with the option for twinning ticked in Refmac5 and converged to a twin fraction of 0.493 for operator *-h,-k,l*. All structures were refined using local non-crystallographic symmetry restraints. Electron density maps for the $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} model were calculated using map sharpening in Refmac5. The PDB-REDO[66] server was used to generate restraints for the IgG1 Fc^{E382S} model for use in refinement. MolProbity[49] and the PDB validation server[67] were used for model validation prior to deposition. Carbohydrates were modelled in Coot[68] and validated using Privateer[69]. Data collection and refinement statistics for IgG1 Fc^{E382S}, $IdeS^{C94A}$-IgG1 Fc^{E382A} and $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} models are presented in Supplementary Tables 1–3, respectively. Polder maps were calculated with the *phenix.polder* tool[48] within the *PHENIX* software suite[47] using default settings. Protein complex interfaces were analysed using PDBePISA[35]. UCSF ChimeraX[70] was used to prepare figures depicting protein structure.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon request. Structure factor files and atomic coordinates for each of the crystallographic models presented in this study have been deposited in the PDB, with accession codes 8A47, 8A48, and 8A49, for the $IdeS^{C94A}$-IgG1 Fc^{E382A} complex, IgG1 Fc^{E382S} and the $EndoS^{D233A/E235L}$-IgG1 Fc^{E382R} complex, respectively. Previously-elucidated crystal structures analysed in this study are deposited in the PDB, with accession codes 3AVE, 1Y08, 2AU1, 2AVW and 6EN3, which correspond to wild-type IgG1 Fc, $IdeS^{C94S}$, wild-type IdeS, $IdeS^{C94A}$ and $EndoS^{D233A/E235L}$, respectively.

References

1. Laabei, M. & Ermert, D. Catch me if you can: *Streptococcus pyogenes* complement evasion strategies. *J. Innate Immun.* **11**, 3–12 (2019).
2. Cunningham, M. W. Pathogenesis of group A streptococcal infections. *Clin. Microbiol Rev.* **13**, 470–511 (2000).
3. von Pawel-Rammingen, U., Johansson, B. P. & Björck, L. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J.* **21**, 1607–1615 (2002).
4. Collin, M. & Olsén, A. EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J.* **20**, 3046–3055 (2001).
5. Collin, M. & Björck, L. Toward clinical use of the IgG specific enzymes IdeS and EndoS against antibody-mediated diseases. *Methods Mol. Biol.* **1535**, 339–351 (2017).
6. Huang, E., Maldonado, A. Q., Kjellman, C. & Jordan, S. C. Imlifidase for the treatment of anti-HLA antibody-mediated processes in kidney transplantation. *Am. J. Transpl.* **22**, 691–697 (2022).
7. Jordan, S. C. et al. Imlifidase desensitizationin crossmatch-positive, highly sensitized kidney transplant recipients: resultsof an international phase 2 trial (Highdes). *Transplantation* **105**, 1808–1817 (2021).
8. Kjellman, C. et al. Outcomes at 3 years posttransplant in imlifidase-desensitized kidney transplant patients. *Am. J. Transpl.* **21**, 3907–3918 (2021).
9. Lei, B. et al. Opsonophagocytosis-inhibiting mac protein of group A *Streptococcus*: identification and characteristics of two genetic complexes. *Infect. Immun.* **70**, 6880–6890 (2002).
10. Söderberg, J. J., Engström, P. & von Pawel-Rammingen, U. The intrinsic immunoglobulin G endopeptidase activity of streptococcal Mac-2 proteins implies a unique role for the enzymatically impaired Mac-2 protein of M28 serotype strains. *Infect. Immun.* **76**, 2183–2188 (2008).
11. Ryan, M. H. et al. Proteolysis of purified IgGs by human and bacterial enzymes in vitro and the detection of specific proteolytic fragments of endogenous IgG in rheumatoid synovial fluid. *Mol. Immunol.* **45**, 1837–1846 (2008).
12. Brezski, R. J. et al. Tumor-associated and microbial proteases compromise host IgG effector functions by a single cleavage proximal to the hinge. *Proc. Natl. Acad. Sci. USA* **106**, 17864–17869 (2009).
13. Yang, R. et al. Successful treatment of experimental glomerulonephritis with IdeS and EndoS, IgG-degrading streptococcal enzymes. *Nephrol. Dial. Transpl.* **25**, 2479–2486 (2010).
14. Lood, C. et al. IgG glycan hydrolysis by endoglycosidase S diminishes the proinflammatory properties of immune complexes from patients with systemic lupus erythematosus: a possible new treatment. *Arthritis Rheum.* **64**, 2698–2706 (2012).
15. Segelmark, M. & Björck, L. Streptococcal enzymes as precision tools against pathogenic IgG autoantibodies in small vessel vasculitis. *Front Immunol.* **10**, 2165 (2019).
16. Kizlik-Masson, C. et al. Cleavage of anti-PF4/heparin IgG by a bacterial protease and potential benefit in heparin-induced thrombocytopenia. *Blood* **133**, 2427–2435 (2019).
17. Takahashi, R. & Yuki, N. Streptococcal IdeS: therapeutic potential for Guillain-Barré syndrome. *Sci. Rep.* **5**, 10809 (2015).
18. Uhlin, F. et al. Endopeptidase cleavage of anti-glomerular basement membrane antibodies in vivo in severe kidney disease: an open-label phase 2a study. *J. Am. Soc. Nephrol.* **33**, 829–838 (2022).
19. Nandakumar, K. S. et al. Streptococcal Endo-β-*N*-acetylglucosaminidase suppresses antibody-mediated inflammation in vivo. *Front Immunol.* **9**, 1623 (2018).
20. Leborgne, C. et al. IgG-cleaving endopeptidase enables in vivo gene therapy in the presence of anti-AAV neutralizing antibodies. *Nat. Med* **26**, 1096–1101 (2020).
21. Järnum, S. et al. Enzymatic inactivation of endogenous IgG by IdeS enhances therapeutic antibody efficacy. *Mol. Cancer Ther.* **16**, 1887–1897 (2017).
22. Baruah, K. et al. Selective deactivation of serum IgG: a general strategy for the enhancement of monoclonal antibody receptor interactions. *J. Mol. Biol.* **420**, 1–7 (2012).
23. Lin, J. et al. Desensitization using imlifidase and EndoS enables chimerism induction in allosensitized recipient mice. *Am. J. Transpl.* **20**, 2356–2365 (2020).
24. Goodfellow, J. J. et al. An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. *J. Am. Chem. Soc.* **134**, 8030–8033 (2012).
25. Sjögren, J. et al. EndoS2 is a unique and conserved enzyme of serotype M49 group A *Streptococcus* that hydrolyses N-linked glycans on IgG and α1-acid glycoprotein. *Biochem J.* **455**, 107–118 (2013).
26. Klontz, E. H. et al. Molecular basis of broad spectrum N-glycan specificity and processing of therapeutic IgG monoclonal antibodies by endoglycosidase S2. *ACS Cent. Sci.* **5**, 524–538 (2019).

27. Huang, W., Giddens, J., Fan, S. Q., Toonstra, C. & Wang, L. X. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *J. Am. Chem. Soc.* **134**, 12308–12318 (2012).
28. Tong, X., Li, T., Li, C. & Wang, L. X. Generation and comparative kinetic analysis of new glycosynthase mutants from *Streptococcus pyogenes* endoglycosidases for antibody glycoengineering. *Biochemistry* **57**, 5239–5246 (2018).
29. Du, J. J., Klontz, E. H., Guerin, M. E., Trastoy, B. & Sundberg, E. J. Structural insights into the mechanisms and specificities of IgG-active endoglycosidases. *Glycobiology* **30**, 268–279 (2020).
30. Vincents, B., von Pawel-Rammingen, U., Björck, L. & Abrahamson, M. Enzymatic characterization of the streptococcal endopeptidase, IdeS, reveals that it is a cysteine protease with strict specificity for IgG cleavage due to exosite binding. *Biochemistry* **43**, 15540–15549 (2004).
31. Trastoy, B. et al. Crystal structure of *Streptococcus pyogenes* EndoS, an immunomodulatory endoglycosidase specific for human IgG antibodies. *Proc. Natl Acad. Sci. USA* **111**, 6714–6719 (2014).
32. Dixon, E. V. et al. Fragments of bacterial endoglycosidase S and immunoglobulin G reveal subdomains of each that contribute to deglycosylation. *J. Biol. Chem.* **289**, 13876–13889 (2014).
33. Trastoy, B. et al. Structural basis for the recognition of complex-type N-glycans by endoglycosidase S. *Nat. Commun.* **9**, 1874 (2018).
34. Matsumiya, S. et al. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J. Mol. Biol.* **368**, 767–779 (2007).
35. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
36. Leaver-Fay, A. et al. Computationally designed bispecific antibodies using negative state repertoires. *Structure* **24**, 641–651 (2016).
37. Potterton, L. et al. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D. Struct. Biol.* **74**, 68–84 (2018).
38. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D. Biol. Crystallogr* **67**, 355–367 (2011).
39. Vindebro, R., Spoerry, C. & von Pawel-Rammingen, U. Rapid IgG heavy chain cleavage by the streptococcal IgG endopeptidase IdeS is mediated by IdeS monomers and is not due to enzyme dimerization. *FEBS Lett.* **587**, 1818–1822 (2013).
40. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
41. Touw, W. G. et al. A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **43**, D364–D368 (2015).
42. Agniswamy, J. et al. Crystal structure of group A *Streptococcus* Mac-1: insight into dimer-mediated specificity for recognition of human IgG. *Structure* **14**, 225–235 (2006).
43. Wenig, K. et al. Structure of the streptococcal endopeptidase IdeS, a cysteine proteinase with strict specificity for IgG. *Proc. Natl Acad. Sci. USA* **101**, 17371–17376 (2004).
44. Coulombe, R. et al. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.* **15**, 5492–5503 (1996).
45. Cygler, M. et al. Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure* **4**, 405–416 (1996).
46. Sivaraman, J., Lalumière, M., Ménard, R. & Cygler, M. Crystal structure of wild-type human procathepsin K. *Protein Sci.* **8**, 283–290 (1999).
47. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D. Struct. Biol.* **75**, 861–877 (2019).
48. Liebschner, D. et al. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D. Struct. Biol.* **73**, 148–157 (2017).
49. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D. Biol. Crystallogr* **66**, 12–21 (2010).
50. Bockermann, R. et al. Imlifidase-generated single-cleaved IgG: implications for transplantation. *Transplantation* https://doi.org/10.1097/TP.0000000000004031 (2021).
51. Lorant, T. et al. Safety, immunogenicity, pharmacokinetics, and efficacy of degradation of anti-HLA antibodies by IdeS (imlifidase) in chronic kidney disease patients. *Am. J. Transpl.* **18**, 2752–2762 (2018).
52. Seki, H. et al. Structural basis for the specific cleavage of core-fucosylated *N*-glycans by endo-β-*N*-acetylglucosaminidase from the fungus *Cordyceps militaris*. *J. Biol. Chem.* **294**, 17143–17154 (2019).
53. Waddling, C. A., Plummer, T. H., Tarentino, A. L. & Van Roey, P. Structural basis for the substrate specificity of endo-β-*N*-acetylglucosaminidase F3. *Biochemistry* **39**, 7878–7885 (2000).
54. García-Alija, M. et al. Mechanism of cooperative N-glycan processing by the multi-modular endoglycosidase EndoE. *Nat. Commun.* **13**, 1137 (2022).
55. Deis, L. N. et al. Suppression of conformational heterogeneity at a protein-protein interface. *Proc. Natl Acad. Sci. USA* **112**, 9028–9033 (2015).
56. Frank, M., Walker, R. C., Lanzilotta, W. N., Prestegard, J. H. & Barb, A. W. Immunoglobulin G1 Fc domain motions: implications for Fc engineering. *J. Mol. Biol.* **426**, 1799–1811 (2014).
57. Barb, A. W. et al. NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation. *Biochemistry* **51**, 4618–4626 (2012).
58. Barb, A. W. & Prestegard, J. H. NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. *Nat. Chem. Biol.* **7**, 147–153 (2011).
59. Lee, H. S. & Im, W. Effects of N-glycan composition on structure and dynamics of IgG1 Fc and their implications for antibody engineering. *Sci. Rep.* **7**, 12659 (2017).
60. Harbison, A. M., Brosnan, L. P., Fenlon, K. & Fadda, E. Sequence-to-structure dependence of isolated IgG Fc complex biantennary N-glycans: a molecular dynamics study. *Glycobiology* **29**, 94–103 (2019).
61. Wilkins, M. R. et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **112**, 531–552 (1999).
62. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D. Struct. Biol.* **74**, 85–97 (2018).
63. Kabsch, W. XDS. *Acta Crystallogr D. Biol. Crystallogr* **66**, 125–132 (2010).
64. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D. Biol. Crystallogr* **66**, 22–25 (2010).
65. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D. Biol. Crystallogr* **66**, 486–501 (2010).
66. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220 (2014).
67. Gore, S. et al. Validation of structures in the Protein Data Bank. *Structure* **25**, 1916–1927 (2017).
68. Emsley, P. & Crispin, M. Structural analysis of glycoproteins: building N-linked glycans with Coot. *Acta Crystallogr D. Struct. Biol.* **74**, 256–263 (2018).
69. Agirre, J. et al. Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **22**, 833–834 (2015).

70. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Acknowledgements
We are grateful to the beamline scientists on I03 at Diamond Light Source (DLS) and ID30A-3 at the European Synchrotron Radiation Facility (ESRF), and to Chris Holes for his support with the macromolecular crystallisation facility at the University of Southampton. This work was supported by DLS, ESRF (MX2373) and the School of Biological Sciences, University of Southampton.

Author contributions
M.C. conceived the project. I.T. and M.C. supervised the project. A.S.L.S., J.B., D.P.I. and I.T. performed experimental work. A.S.L.S., I.T. and M.C. performed computational work. A.S.L.S. prepared the original draft together with initial data presentation. I.T. and M.C. reviewed and edited the manuscript. All authors have read and approve the final manuscript.

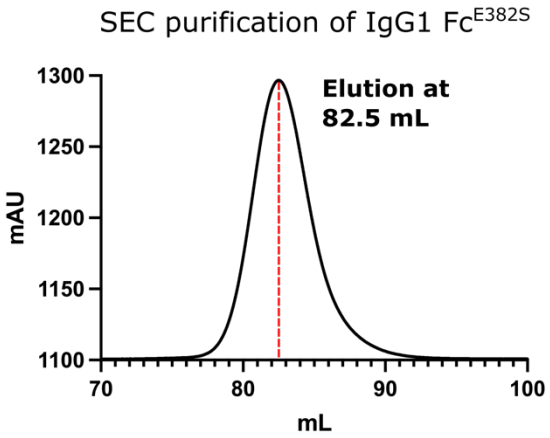
Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-35340-z.

Correspondence and requests for materials should be addressed to Ivo Tews or Max Crispin.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Appendix B Example IgG1 Fc^{E382S} SEC purification



Example IgG1 Fc^{E382S} SEC purification over a HiLoad® 16/600 Superdex 200 pg column (method detailed in chapter 2, section 2.3.1). Peak elutes at ~82.5 mL, which is similar to the values for the other IgG1 Fcs reported in **Figure 25**. N.B. baseline is not at zero due to a fault with the software.

Appendix C wwPDB X-ray Structure Validation Report for IgG1 Fc^{E382A}

Preliminary wwPDB X-ray Structure Validation Report for IgG1 Fc^{E382A} crystal structure (not yet deposited). Report is included overleaf:

WORLDWIDE PDB
PROTEIN DATA BANK

Preliminary Full wwPDB X-ray Structure Validation Report (i)

Jan 10, 2023 – 11:04 AM EST

This wwPDB validation report is NOT for manuscript review

This is a Preliminary Full wwPDB X-ray Structure Validation Report.

This report is produced by the standalone wwPDB validation server.
The structure in question has not been deposited to the wwPDB.
This report should not be submitted to journals.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

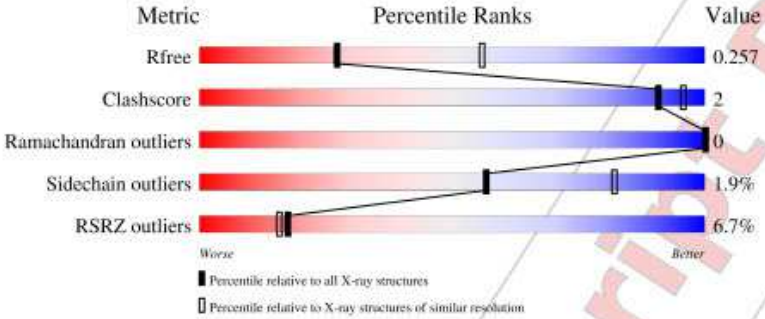
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	227	12% 86% 5% 9%
1	B	227	85% 6% 8%
2	C	9	78% 22%
2	D	9	89% 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FUC	C	9	-	-	-	X

2 Entry composition

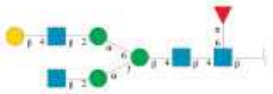
There are 6 unique types of molecules in this entry. The entry contains 7134 atoms, of which 3436 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein.

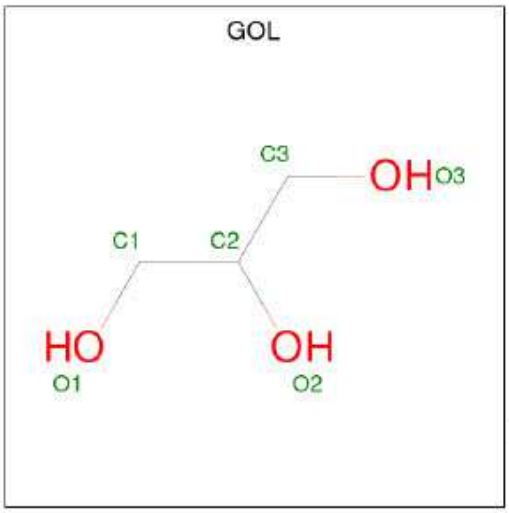
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	207	Total	C	H	N	O	S	63	0	0
			3206	1032	1589	271	307	7			
1	B	208	Total	C	H	N	O	S	56	0	0
			3287	1051	1633	280	316	7			

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



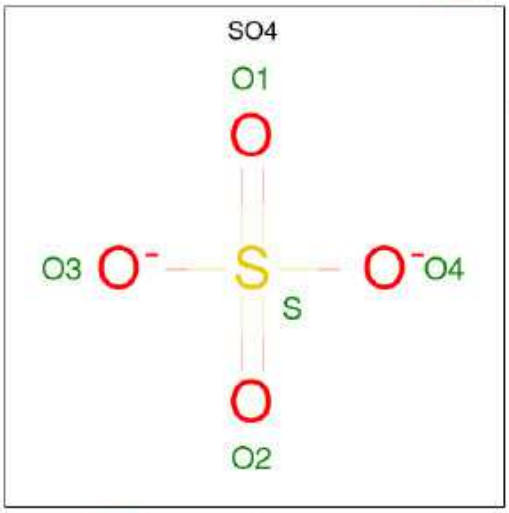
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	9	Total	C	H	N	O	23	0	0
			213	62	103	4	44			
2	D	9	Total	C	H	N	O	23	0	0
			213	62	103	4	44			

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 14 3 8 3	2	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	Cl	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	185	Total	O	0	0
			185	185		

3 Residue-property plots

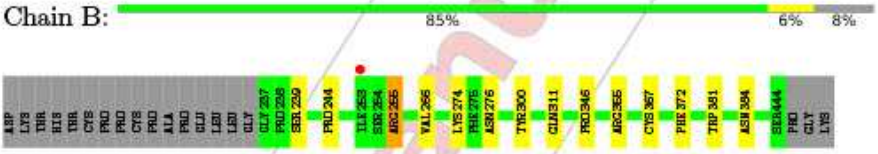
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1:

Chain A:

• Molecule 1:

Chain B:



• Molecule 2: beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

• Molecule 2: beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	106.55Å 106.55Å 103.27Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.18 – 2.70 46.14 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.18-2.70) 99.7 (46.14-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0403	Depositor
R, R_{free}	0.194 , 0.253 0.203 , 0.257	Depositor DCC
R_{free} test set	912 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	67.1	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 53.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.010 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7134	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, FUC, CL, BMA, NAG, MAN, SO4, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.46	0/1662	0.72	0/2269
1	B	0.47	0/1699	0.75	0/2314
All	All	0.46	0/3361	0.73	0/4583

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	255	ARG	Sidechain

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1617	1589	1566	6	0
1	B	1654	1633	1622	5	0
2	C	110	103	94	1	0
2	D	110	103	94	1	0
3	A	6	8	8	2	0
4	E	15	0	0	1	0
5	G	1	0	0	0	0
6	F	185	0	0	1	0
All	All	3698	3436	3384	12	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (12) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:7:MAN:O4	2:D:7:MAN:O6	2.13	0.66
1:B:346:PRO:HB3	1:B:372:PHE:HB3	1.95	0.47
1:A:379:VAL:H	3:A:501:GOL:H2	1.79	0.47
1:A:346:PRO:HB3	1:A:372:PHE:HB3	1.99	0.45
1:B:367:CYS:HB2	1:B:381:TRP:CZ2	2.53	0.44
1:A:286:ASN:HA	6:F:118:HOH:O	2.18	0.43
1:B:266:VAL:HB	1:B:300:TYR:HB2	2.00	0.43
1:B:255:ARG:NH2	4:E:2:SO4:O1	2.46	0.42
1:A:367:CYS:HB2	1:A:381:TRP:CZ2	2.55	0.42
1:A:266:VAL:HB	1:A:300:TYR:HB2	2.02	0.42
1:A:393:THR:HG21	3:A:501:GOL:H32	2.03	0.41
1:B:274:LYS:HE3	1:B:276:ASN:OD1	2.20	0.41

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	205/227 (90%)	200 (98%)	5 (2%)	0	100	100
1	B	206/227 (91%)	201 (98%)	5 (2%)	0	100	100
All	All	411/454 (90%)	401 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	184/209 (88%)	182 (99%)	2 (1%)	73	90
1	B	192/209 (92%)	187 (97%)	5 (3%)	46	75
All	All	376/418 (90%)	369 (98%)	7 (2%)	57	82

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	239	SER
1	A	284	VAL
1	B	239	SER
1	B	244	PRO
1	B	311	GLN
1	B	355	ARG
1	B	384	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	310	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

18 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	C	1	1,2	14,14,15	0.71	0	17,19,21	0.36	0
2	NAG	C	2	2	14,14,15	0.82	0	17,19,21	0.52	0
2	BMA	C	3	2	11,11,12	0.65	0	15,15,17	1.10	0
2	MAN	C	4	2	11,11,12	0.61	0	15,15,17	0.76	0
2	NAG	C	5	2	14,14,15	0.62	0	17,19,21	0.57	0
2	GAL	C	6	2	11,11,12	0.95	0	15,15,17	0.92	0
2	MAN	C	7	2	11,11,12	0.77	0	15,15,17	0.99	0
2	NAG	C	8	2	14,14,15	0.62	0	17,19,21	1.00	1 (5%)
2	FUC	C	9	2	10,10,11	0.58	0	14,14,16	0.85	0
2	NAG	D	1	1,2	14,14,15	0.63	0	17,19,21	0.38	0
2	NAG	D	2	2	14,14,15	0.52	0	17,19,21	0.64	0
2	BMA	D	3	2	11,11,12	0.89	0	15,15,17	1.29	0
2	MAN	D	4	2	11,11,12	0.84	0	15,15,17	0.74	0
2	NAG	D	5	2	14,14,15	0.64	0	17,19,21	0.90	0
2	GAL	D	6	2	11,11,12	1.08	0	15,15,17	0.82	0
2	MAN	D	7	2	11,11,12	0.83	1 (9%)	15,15,17	0.93	0
2	NAG	D	8	2	14,14,15	0.55	0	17,19,21	0.47	0
2	FUC	D	9	2	10,10,11	0.64	0	14,14,16	1.16	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
2	BMA	C	3	2	-	0/2/19/22	0/1/1/1
2	MAN	C	4	2	-	2/2/19/22	0/1/1/1
2	NAG	C	5	2	-	4/6/23/26	0/1/1/1
2	GAL	C	6	2	-	0/2/19/22	0/1/1/1
2	MAN	C	7	2	-	2/2/19/22	0/1/1/1
2	NAG	C	8	2	-	3/6/23/26	0/1/1/1
2	FUC	C	9	2	-	-	0/1/1/1
2	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1
2	BMA	D	3	2	-	0/2/19/22	0/1/1/1
2	MAN	D	4	2	-	0/2/19/22	0/1/1/1
2	NAG	D	5	2	-	4/6/23/26	0/1/1/1
2	GAL	D	6	2	-	2/2/19/22	0/1/1/1
2	MAN	D	7	2	-	2/2/19/22	0/1/1/1
2	NAG	D	8	2	-	1/6/23/26	0/1/1/1
2	FUC	D	9	2	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	7	MAN	C2-C3	-2.08	1.49	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	8	NAG	C2-N2-C7	2.04	125.81	122.90

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	7	MAN	O5-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6
2	C	8	NAG	C4-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6
2	C	7	MAN	C4-C5-C6-O6
2	C	8	NAG	O5-C5-C6-O6
2	D	7	MAN	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	7	MAN	O5-C5-C6-O6
2	C	5	NAG	C8-C7-N2-C2
2	C	5	NAG	O7-C7-N2-C2
2	D	5	NAG	C8-C7-N2-C2
2	D	5	NAG	O7-C7-N2-C2
2	C	5	NAG	C4-C5-C6-O6
2	C	5	NAG	O5-C5-C6-O6
2	C	4	MAN	O5-C5-C6-O6
2	D	5	NAG	O5-C5-C6-O6
2	D	6	GAL	O5-C5-C6-O6
2	D	5	NAG	C4-C5-C6-O6
2	C	4	MAN	C4-C5-C6-O6
2	C	8	NAG	C1-C2-N2-C7
2	D	8	NAG	C4-C5-C6-O6
2	D	6	GAL	C4-C5-C6-O6

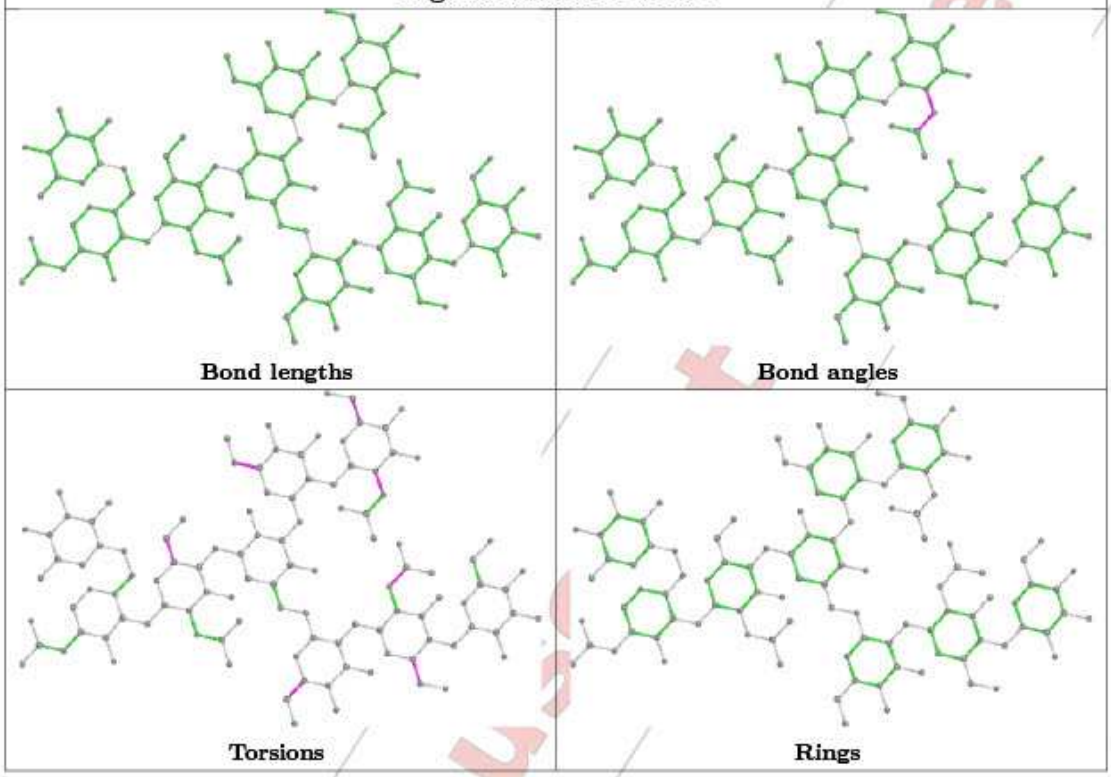
There are no ring outliers.

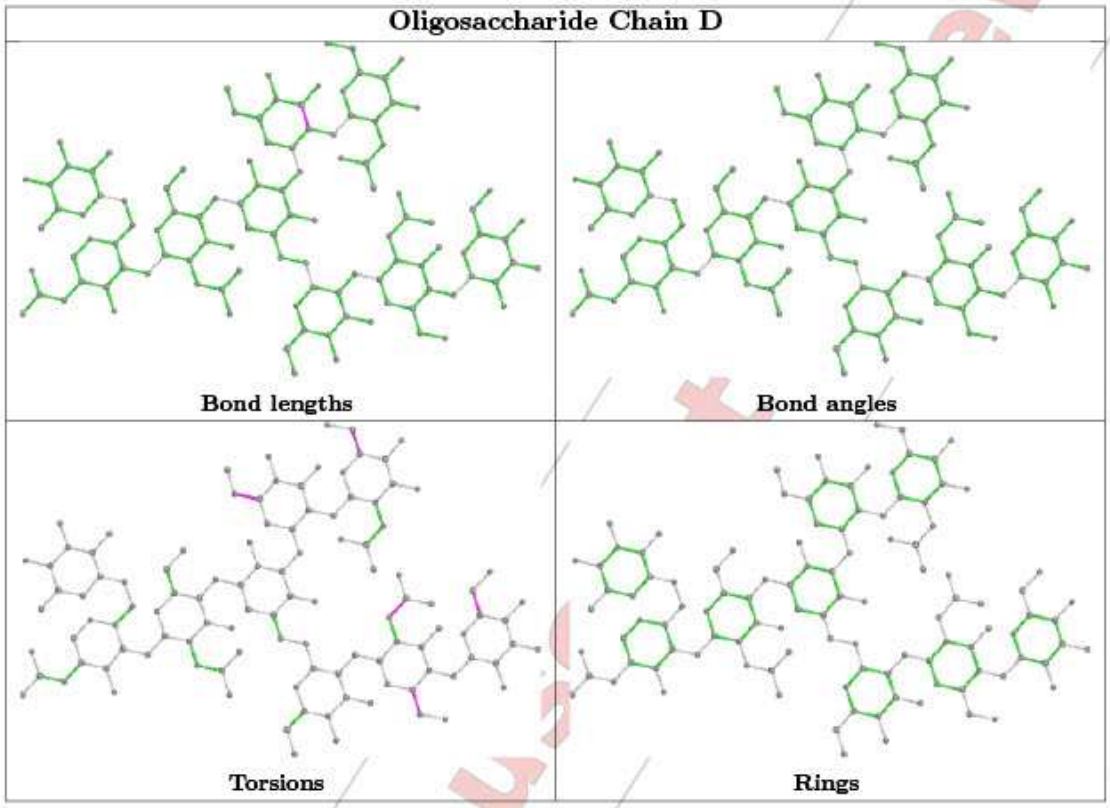
2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	7	MAN	1	0
2	D	7	MAN	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain C





5.6 Ligand geometry

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SO4	E	3	-	4,4,4	0.29	0	6,6,6	0.26	0
4	SO4	E	2	-	4,4,4	0.37	0	6,6,6	0.14	0
4	SO4	E	1	-	4,4,4	0.20	0	6,6,6	0.45	0
3	GOL	A	501	-	5,5,5	0.22	0	5,5,5	0.65	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	501	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	GOL	O1-C1-C2-C3
3	A	501	GOL	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	2	SO4	1	0
3	A	501	GOL	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	207/227 (91%)	0.84	27 (13%) 3 2	39, 70, 126, 156	0
1	B	208/227 (91%)	0.50	1 (0%) 91 92	42, 65, 101, 119	0
All	All	415/454 (91%)	0.67	28 (6%) 17 16	39, 67, 114, 156	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300	TYR	5.8
1	A	266	VAL	5.2
1	A	268	HIS	4.4
1	A	332	ILE	4.3
1	A	298	SER	4.1
1	A	291	PRO	4.1
1	A	267	SER	3.9
1	A	295	GLN	3.7
1	A	264	VAL	3.7
1	A	299	THR	3.6
1	A	270	ASP	3.5
1	A	296	TYR	3.4
1	A	302	VAL	3.3
1	A	269	GLU	3.1
1	A	271	PRO	2.9
1	A	294	GLU	2.9
1	A	239	SER	2.9
1	A	241	PHE	2.9
1	A	323	VAL	2.7
1	A	290	LYS	2.7
1	A	265	ASP	2.5
1	A	305	VAL	2.5
1	A	292	ARG	2.5
1	B	253	ILE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	303	VAL	2.5
1	A	293	GLU	2.2
1	A	328	LEU	2.1
1	A	297	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

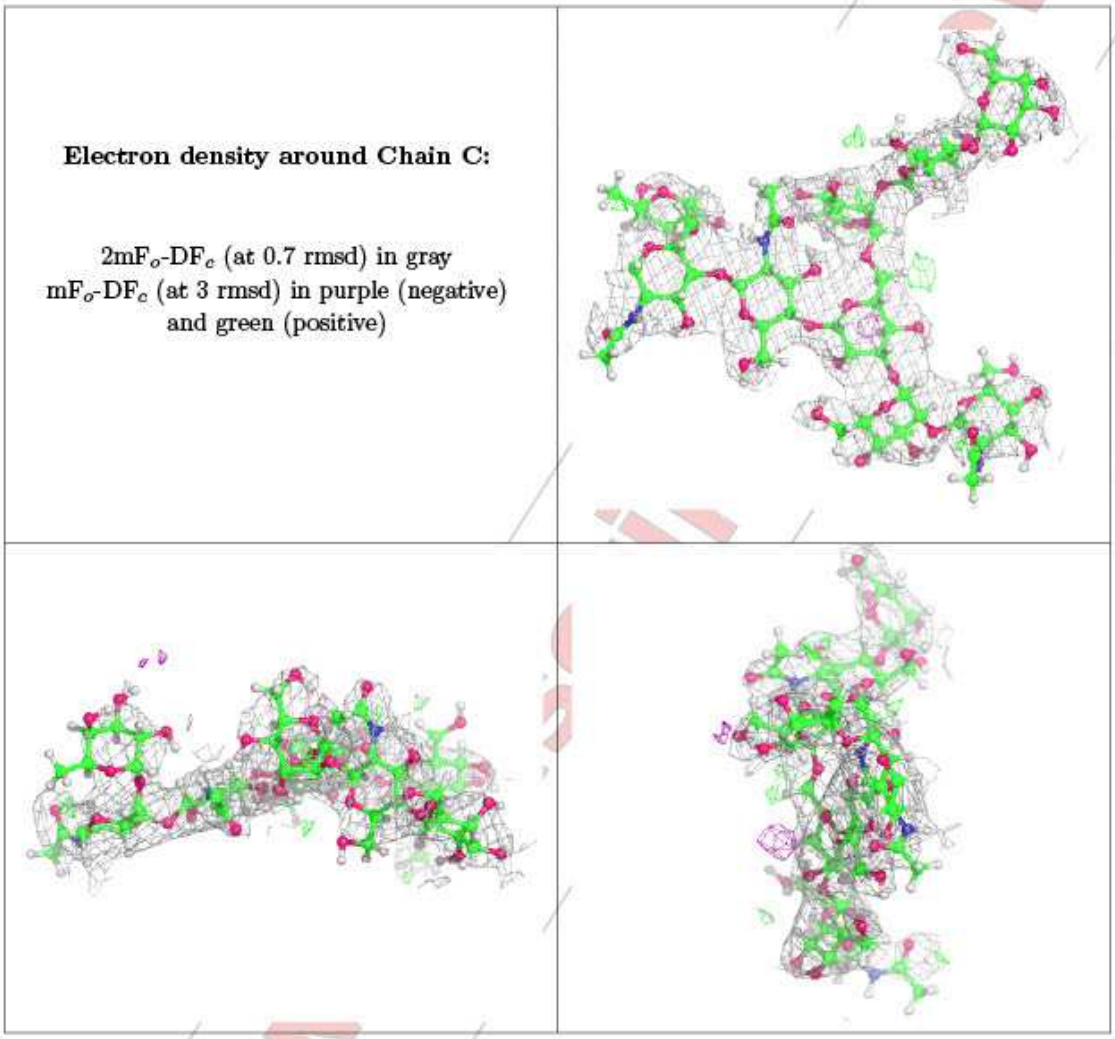
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

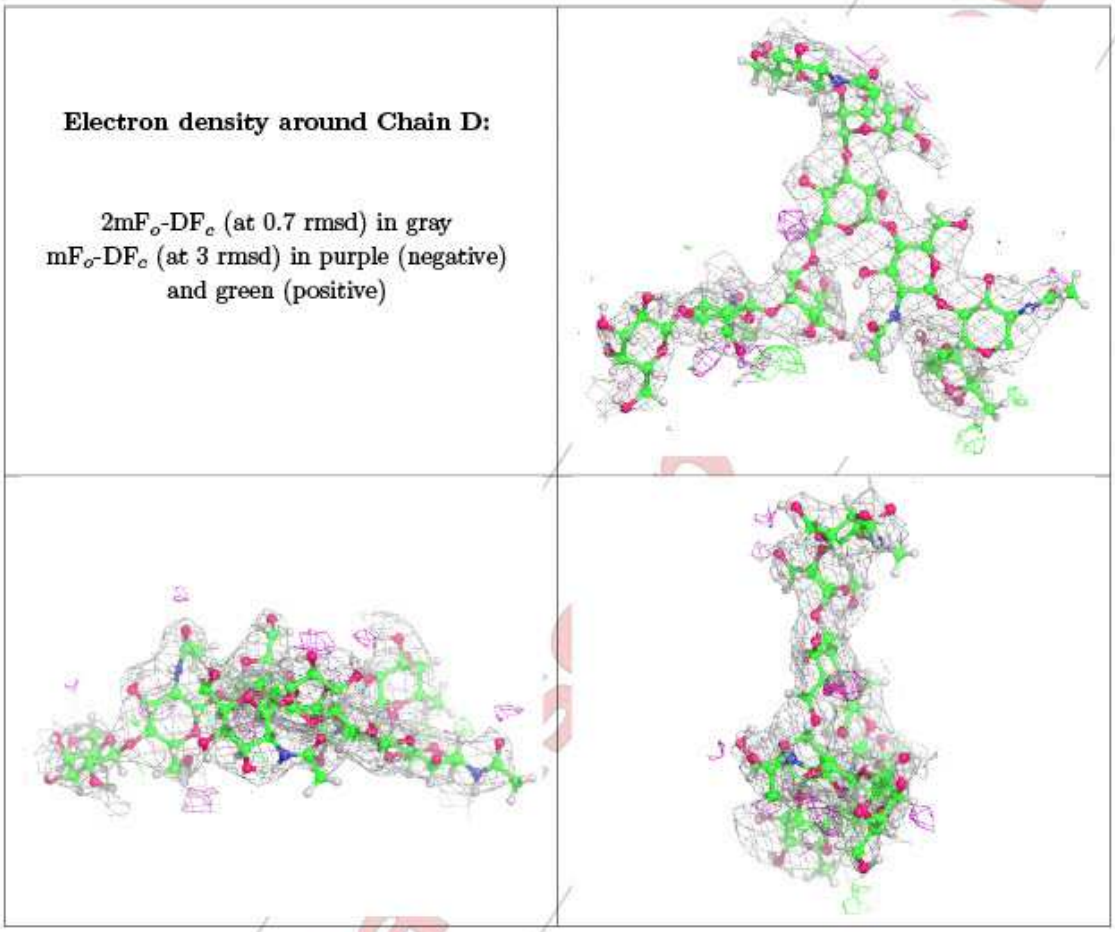
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	C	8	14/?	0.71	0.31	52,131,152,158	3
2	FUC	C	9	10/?	0.72	0.54	52,140,150,167	3
2	GAL	C	6	11/?	0.85	0.36	52,99,110,113	4
2	MAN	C	7	11/?	0.85	0.27	52,98,114,122	3
2	NAG	D	8	14/?	0.85	0.21	52,121,131,146	3
2	GAL	D	6	11/?	0.88	0.26	52,105,114,121	4
2	NAG	C	1	14/?	0.88	0.27	52,106,114,125	1
2	NAG	C	5	14/?	0.90	0.27	52,101,113,114	2
2	MAN	C	4	11/?	0.91	0.21	52,98,109,114	3
2	NAG	C	2	14/?	0.91	0.21	52,87,105,110	2
2	FUC	D	9	10/?	0.91	0.15	52,89,92,96	3
2	NAG	D	5	14/?	0.93	0.20	52,71,76,83	2
2	BMA	D	3	11/?	0.93	0.13	52,61,73,74	2
2	BMA	C	3	11/?	0.94	0.10	52,75,79,80	2
2	MAN	D	7	11/?	0.95	0.18	52,87,94,106	3
2	NAG	D	2	14/?	0.95	0.14	52,73,76,77	2
2	NAG	D	1	14/?	0.95	0.20	52,84,89,93	1
2	MAN	D	4	11/?	0.96	0.16	52,75,79,87	3

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain C:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	E	3	5/?	0.82	0.33	84,104,134,156	0
3	GOL	A	501	6/?	0.87	0.26	30,73,80,86	2
4	SO4	E	2	5/?	0.88	0.32	116,120,125,155	0
4	SO4	E	1	5/?	0.95	0.26	67,79,91,94	0
5	CL	G	1	1/?	0.99	0.30	41,41,41,41	1

6.5 Other polymers

There are no such residues in this entry.

Appendix D IdeS^{C94A}-asymmetric IgG1 Fc^{E382R} optimisation screen

Optimisation screen set up for an initial hit of the asymmetric IgG1 Fc^{E382R} in complex with truncated IdeS^{C94A}, obtained from condition G4 (from the JCSG-plus™ crystallisation screens; see **Figure 39**).

Table 27: Optimisation screen set up for crystallisation of truncated IdeS^{C94A}-asymmetric IgG1 Fc^{E382R}, from an initial hit in condition G4 (JCSG-plus™).

A1	A2	A3	A4	A5	A6	pH 7.5
B1	B2	B3	B4	B5	B6	pH 8.0
C1	C2	C3	C4	C5	C6	pH 8.5
D1	D2	D3	D4	D5	D6	pH 9.0
18 %	19 %	20 %	21 %	22 %	23 %	

This condition comprised 0.2 M TMAO, 0.1 M Tris (pH 8.5), 20 % w/v PEG 2000 MME; the optimisation was designed to mimic this condition, with varying PEG 2000 MME (from 18 % to 23 %; x-axis) and pH (7.5 to 9.0; y-axis). TMAO, trimethylamine-*N*-oxide; PEG, polyethylene glycol.

Appendix E Diagnostic tests for twinning: IdeS-IgG1 Fc crystal structure

Twinning is a crystallographic pathology in which a crystal contains several distinct domains with distinct orientations. Upon irradiation with X-rays, this results in a diffraction pattern which comprises a sum of the diffraction by these distinct domains. These diffraction patterns overlap depending on how the domains are oriented in relation to each other. The presence of twinning therefore affects the distribution of observed reflection intensities [415]. Detection of this data pathology is important for successful structure solution and proper refinement of the data.

Analysis of the truncated IdeS^{C94A}-IgG1 Fc^{E382A} data set in *phenix.xtriage* [396] suggested the presence of twinning within the crystal (see **Figure 71** and **Figure 72** below). In brief, these analyses study the distribution of reflection intensities throughout the data set. **Figure 71a** shows the *N(Z)* plot, which is a cumulative intensity distribution. Within a twinned data set, the overlapping diffraction patterns comprising the distinct twin domains results in more "average" reflection intensities, and fewer apparent "extreme" intensities (i.e., very strong or very weak reflections). Therefore, a twinned data set tends to produce a cumulative intensity distribution with a more sigmoidal shape.

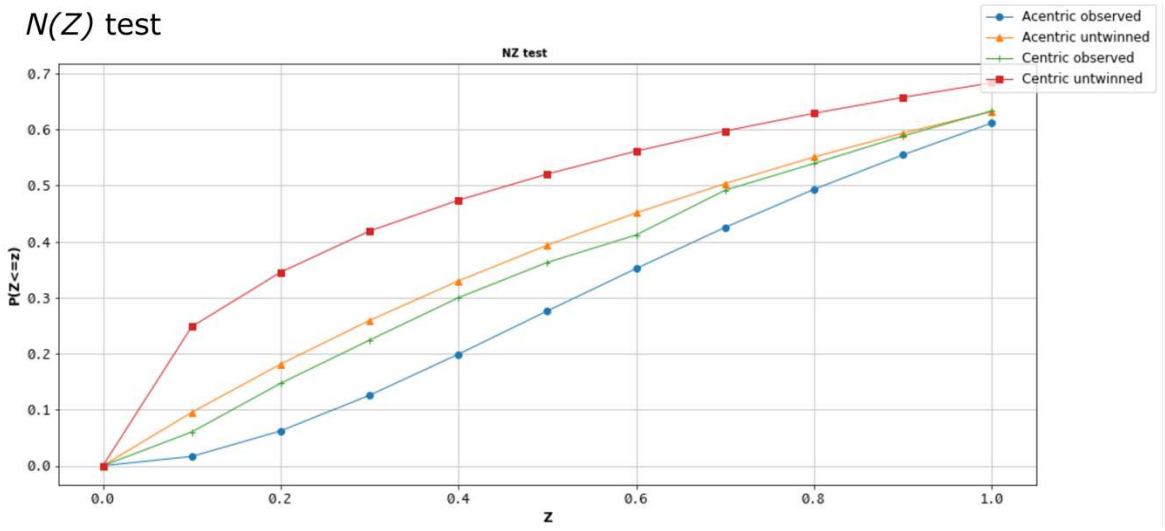


Figure 71: *N(Z)* test for detection of twinning within the truncated IdeS^{C94A}-IgG1 Fc^{E382A} crystallographic data set. The cumulative intensity distribution of the data produces a more sigmoidal-shaped curve (blue for acentric reflections; green for centric reflections) compared to theoretical distributions for "normal" data. Analysis was performed by *phenix.xtriage* [396].

A more robust test for twinning is the *L* test, which is also a cumulative distribution of observed reflection intensities. This test examines differences in intensities of pairs of spatially-related reflections (i.e., close together within the diffraction pattern), therefore this test is less affected by the presence of other possible crystal pathologies (e.g., anisotropy or non-crystallographic symmetry)

[416]. *L* is derived below, where h_1 and h_2 are two spatially-related reflections, and *I* is the intensity value of each reflection:

$$L = \frac{I(h_1) - I(h_2)}{I(h_1) + I(h_2)}$$

A cumulative distribution of *L*, as shown in **Figure 72**, is expected to be linear for "normal" data. Analysis of the data from the truncated IdeS^{C94A}-IgG1 Fc^{E382A} crystal shows that the *L* distribution is significantly different from that expected for normal data (orange) and closely follows that expected for perfectly-twinned data (green).

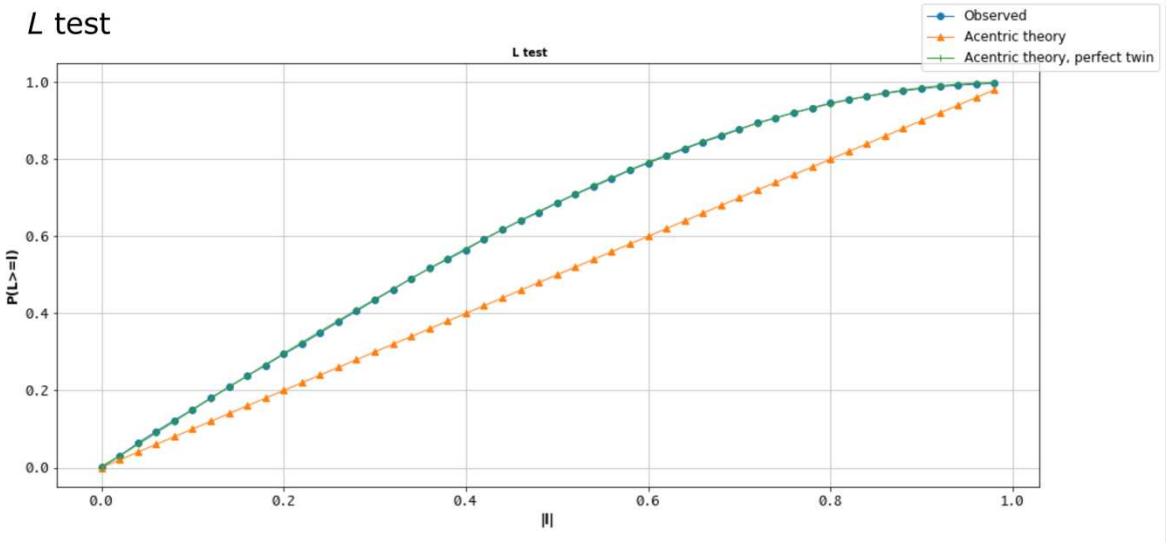


Figure 72: *L* test for detection of twinning within the truncated IdeS^{C94A}-IgG1 Fc^{E382A} crystallographic data set. The *L* distribution differs significantly from that expected for "normal" data (orange) and closely resembles that predicted for perfectly-twinned data (green), which indicates the presence of twinning within this data set. Analysis was performed by *phenix.xtriage* [396].

Appendix F Publication 2 (pre-print)

Sudol ASL, Tews I, Crispin M. Bespoke conformation and antibody recognition distinguishes the streptococcal immune evasion factors EndoS and EndoS2. *bioRIV*. 2023:553389.

This publication pre-print is included overleaf for reference:

Bespoke conformation and antibody recognition distinguishes the streptococcal immune evasion factors EndoS and EndoS2

Abigail S. L. Sudol[1], Ivo Tews[1]*, Max Crispin[1]*.

[1]School of Biological Sciences, University of Southampton, Southampton, SO17 1BJ, UK

*Correspondence to Ivo Tews (Ivo.Tews@soton.ac.uk) and Max Crispin (Max.Crispin@soton.ac.uk)

Abstract

The IgG-specific endoglycosidases EndoS and EndoS2 from *Streptococcus pyogenes* ablate IgG function by removing the conserved N-linked glycans present on the Fc region. Their role in immune evasion, by inactivation of IgG antibodies, has led these enzymes to be investigated as therapeutics for suppressing unwanted immune activation. Their activity and precise substrate specificity has also prompted the development of these enzymes as tools for engineering IgG glycosylation. Recent structural studies have revealed how EndoS drives specificity for IgG by binding the Fc peptide surface with a domain that has homology for a carbohydrate-binding module (CBM). Here, we present the crystal structure of the EndoS2-IgG1 Fc complex at 3.0 Å resolution. Comparison with the EndoS-IgG1 Fc structure reveals a similar mode of interaction, but slightly different orientations resulting from different interfaces with glycosidase and CBM domains, leading to recognition of distinct Fc surfaces. These findings rationalise previous observations that non-catalytic domains cannot readily be substituted. The structural information presented here will guide the continued development of IgG-specific endoglycosidases in antibody glycoengineering and immunotherapy.

Key words: antibody; glycosylation; crystal; endoglycosidase; Fc; structure

Main text

The pathogen *Streptococcus pyogenes* is highly adapted for human infection and has thus developed a range of immune evasion mechanisms for prolonging infection [1]. This pathogenic bacterial species possesses a toolkit of enzymes which specifically target and degrade immunoglobulin G (IgG) antibodies, the most abundant antibody found within human serum [2]. For example, the protease IgG-degrading enzyme of *S. pyogenes* (IdeS) deactivates IgG by cleaving within the lower hinge region, yielding $F(ab')_2$ and Fc fragments [3, 4]. Another protease, SpeB, also cleaves a broader range of immunoglobulins and other components of the immune system [5, 6].

Streptococcal EndoS [7] and EndoS2 [8] enzymes are additionally utilised by the bacterium for the removal of N-linked glycans from IgG Fc. Specifically, these enzymes cleave the β1–4 glycosidic linkage between the two *N*-acetylglucosamine (GlcNAc) residues within the glycan core, releasing the majority of the glycan and leaving a single GlcNAc, variably modified with fucose. EndoS is specific towards complex-type, biantennary glycans on IgG Fc [7, 9], whereas EndoS2 exhibits broader glycan specificity, also cleaving other classes of N-linked glycans such as oligomannose-type and hybrid-type [8, 10]. The broader substrate specificity exhibited by EndoS2 appears to be possible due to a larger glycan-binding groove within the active site which can accommodate such bulkier glycans [11], in comparison to that of EndoS [12]. The glycosylation of IgG occurs at residue N297 within the Fc Cγ2 domains and is conserved across all IgG subtypes [13]. Glycosylation of the Fc domain is implicated in maintaining Fc structural integrity [14-19], Fc gamma receptor (FcγR)- and complement-mediated immune activation [2, 17, 20]. Removal of the glycan therefore impedes Fc-mediated effector functions [19, 21], although activity can sometimes be detected for particular IgG subclasses when deglycosylated Fcs are presented in immune complexes [22].

EndoS was demonstrated to improve survival of *S. pyogenes* in an opsonophagocytic assay, due to reduced FcγR- and complement-mediated immune activation [23]. Moreover, IgG glycan hydrolysis by EndoS has been shown to promote virulence and survival of the bacterium *in vivo* [24], thus demonstrating its role in immune evasion. This highlights the potential of EndoS, and EndoS2, as therapeutic agents for dampening unwanted immune activation, such as during organ transplantation or in an autoimmune disease context. Indeed, EndoS has been successfully utilised as a treatment in several pre-clinical models of autoimmune disease [25-

34], and, in combination with IdeS, was shown to be successful at inactivating donor-specific antibodies in a murine model of bone marrow transplantation [35]. The specific deactivation of competing serum IgG by EndoS and/or IdeS is also being investigated for the potentiation of therapeutic antibodies [19, 36].

Additional interest in these enzymes stems from potential use in biotechnological application, providing tools for antibody glycoengineering [37, 38]. Wild-type enzymes can be used to cleave off unwanted glycoforms, and transglycosylation reactions can be performed on this antibody scaffold to generate desired glycoforms [39-43]. Such transglycosylation reactions have been optimised using variants of these enzymes, although both wild-type EndoS [9, 44] and EndoS2 [41] have been shown to possess some transglycosylation activity. A further development has been the use of wild-type EndoS2 in "one-step" reactions for synthesis of antibody-drug conjugates [45, 46]. The precise control of antibody glycosylation has been applied in several clinically-used antibodies for improved immune effector function [47-51], thus demonstrating their utility in antibody engineering.

The potential therapeutic and biotechnological uses of EndoS and EndoS2 have prompted extensive research into the structure and function of these enzymes. These endoglycosidases are multi-domain enzymes, with the catalytic glycosidase domain and the so-called carbohydrate binding module (CBM) being of particular interest [12, 52-54]. Recent structural studies revealed that the functional role of the CBM in EndoS is not to bind carbohydrate, but rather to specify peptide binding on the Fc surface [55, 56], which provided a structural rationale for the abrogated enzymatic activity of EndoS lacking the CBM [52, 53], further explaining the inability of EndoS to cleave glycans from denatured IgG [6]. Similarly, it is known for that the CBM of EndoS2 is essential for activity. Furthermore, the CBM must work in tandem with the catalytic domain, as a chimeric enzyme displacing the EndoS2 catalytic domain onto the EndoS scaffold leads to a non-functional enzyme [11].

We therefore sought to investigate the structural basis behind the specificity of EndoS2 for IgG, and how it differs from EndoS in substrate recognition. We used the previously established "less-crystallisable" Fc to obtain crystals of the EndoS2-IgG1 Fc complex [55]. We further introduced a L234C exchange, in an attempt stabilise the flexible hinge region through disulphide formation (see Supplementary Information). The enzyme $EndoS2_{38–843}$ was expressed as an inactive exchange variant D184A/E186L with C-terminal His-tag (see SI, [11]). Purified $EndoS2^{D184A/E186L}$ and IgG1 $Fc^{L234C/E382A}$ were combined in a 2:1 molar ratio and the

resulting complex purified by size exclusion chromatography and submitted to crystallisation for structure determination (see SI).

Crystallographic studies of enzymes in complex with their glycan substrate have provided insight into the mechanism of action of endoglycosidases [11, 12, 57-59]. Our recent structure of EndoS in complex with its IgG1 Fc substrate [55] and a cryo-electron microscopy study [56] provided further insight. Here, we present the crystal structure of EndoS2 in complex with IgG1 Fc substrate at a resolution of 3.0 Å (Table 1), with N-linked glycans modelled using the carbohydrate module in Coot [60] and validated using Privateer [61], as shown in Figure 1. The complex crystallised in space group $P4_32_12$, with three copies of the EndoS2-half Fc complex within the asymmetric unit (Table 1).

Table 1: Crystallographic data collection and refinement statistics for EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. Values for the highest resolution shell are reported in parentheses.

Data Collection	
Beamline	I03 (Diamond Light Source)
Resolution range (Å)	49.91-3.00 (3.05-3.00)
Space group	$P4_32_12$
Unit cell dimensions: *a*, *b*, *c* (Å) α, β, γ (°)	 228.8, 228.8, 161.6 90, 90, 90
Wavelength (Å)	0.9763
Unique reflections	85,926 (4250)
Completeness (%)	100 (100)
R_{merge}	0.187 (2.614)
R_{meas}	0.191 (2.668)
R_{pim}	0.036 (0.527)
$I/\sigma(I)$	12.6 (0.5)
Multiplicity	27.5 (25.4)
CC half	1.000 (0.640)
Wilson *B* factor (Å^2)	77
Refinement	
Number of reflections (all/free)	85,849/4376
R_{work} (%)	21.6
R_{free} (%)	25.2
RMSD[1]: Bonds (Å) Angles (°)	 0.0066 1.18
Molecules per ASU[2]	6* (3 copies of EndoS2-half Fc complex)
Atoms per ASU[2]	21,168
Average *B* factors (Å^2) (protein/ligand/water)	(105.2/105.0/73.1)

Model quality (Ramachandran plot)[3]:	
Most favoured region (%)	96.06
Allowed region (%)	3.75
Outliers (%)	0.19, 0.0**

[1] RMSD, root-mean-squared deviation
[2] ASU, asymmetric unit
[3]calculated in Molprobity
*each half Fc (*i.e.*, one chain within the homodimer) is counted as a separate molecule
**according to wwPDB Structure Validation Report

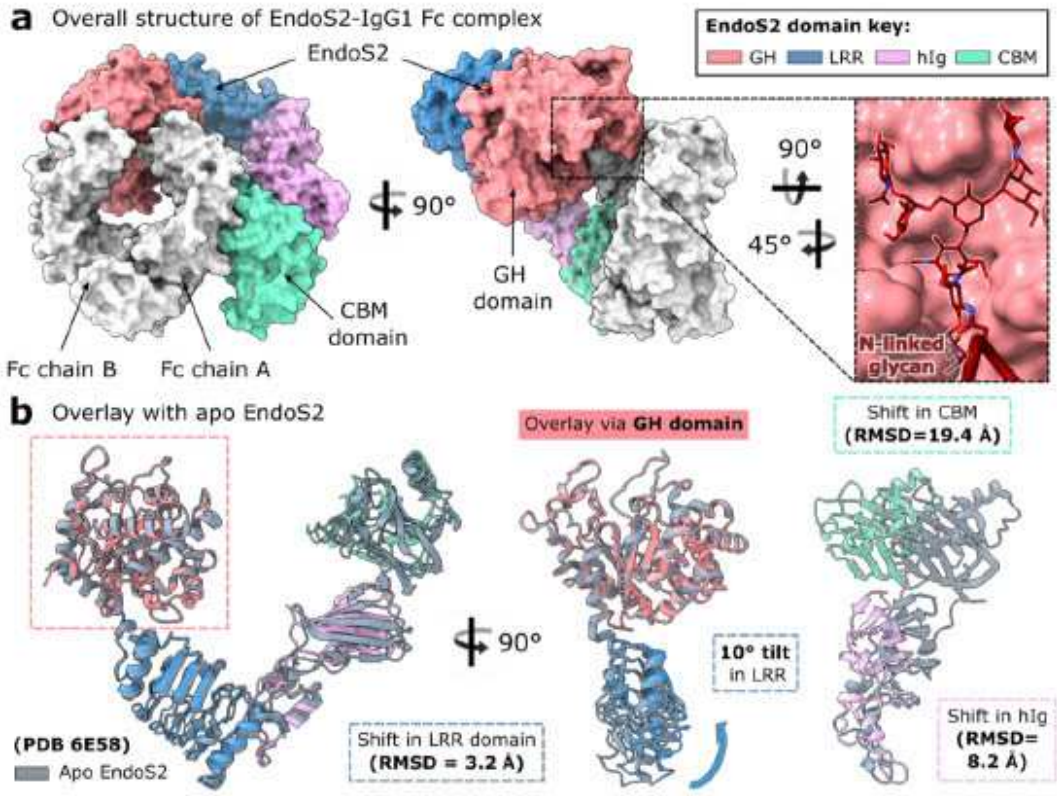


Figure 1: Overall structure of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. **a** Surface representation of the EndoS-IgG1 Fc complex. Two EndoS2 molecules bind each Fc homodimer in the crystal (see Supplementary Figure 2); for clarity, only one EndoS2 molecule is shown here. IgG1 Fc is depicted in silver, with the N297 glycan depicted as maroon sticks and coloured by heteroatom (oxygen, red; nitrogen, blue). EndoS2 domains are coloured as follows: glycoside hydrolase (GH), coral; leucine-rich repeat (LRR), blue; hybrid immunoglobulin (hIg), lilac; carbohydrate-binding module (CBM), green. **b** Superposition of apo EndoS2 (PDB 6E58) and EndoS2 in complex with IgG1 Fc, with respect to the GH domain (aligned Cαs for EndoS2 residues 46–386). IgG1 Fc in complexed EndoS2 is omitted here for clarity. The figure reveals the shift of GH and CBM relative to each other, with the entire GH-LRR-hIg-CBM scaffold in different orientation. RMSDs and plane angle differences for EndoS2 domain shifts after these superpositions are indicated, as calculated in UCSF ChimeraX. Images depicting protein structure were prepared using UCSF ChimeraX [62].

EndoS2 comprises a glycosylhydrolase domain (GH, coral, residues 43–386), a leucine-rich repeat domain (LRR, blue, residues 387–547), a hybrid immunoglobulin domain (hIg, lilac, residues 548–680) and a carbohydrate binding module (CBM, green, residues 681–843) [11]. The enzyme is shaped like a "V", and captures one Cγ2 domain of the Fc through binding by the GH and CBM domains (Figure 1), similar to that observed in the structure of the EndoS-Fc complex [55, 56]. Each Fc γ-chain is seen to interact with one EndoS2 molecule, similar to what was observed with the EndoS-Fc crystal lattice [55].

In the EndoS2-IgG1 Fc complex, the uncleaved Fc N-linked glycan is captured in a "flipped-out" conformation and bound within the GH domain groove (Figure 1a), which was identified by co-crystals of EndoS2 in complex with glycan substrates [11]. This observation is in stark contrast to the typically-observed position of Fc glycans within crystal structures, in which they appear to sit in-between the Cγ2 domains and interact with Fc surface residues [13, 63-65]. The carbohydrate is well-ordered, displaying low *B* factors with respect to the average *B* factor of the complex (105.8 Å^2, Table 1). A polder map additionally displays clear density consistent with an uncleaved glycan in this conformation (Supplementary Figure 3). Overall, this demonstrates the same mode of glycan capture as observed in the EndoS-IgG1 Fc complex [55], and further corroborates literature reporting heterogeneity of Fc N-linked glycan conformations [66-70].

Superposition of complexed and apo EndoS2 (PDB 6E58) reveals a shifting of multiple domains upon Fc binding (Figure 1b). An overlay with respect to the GH domain (calculated by aligning Cαs for residues 46–386) results in a 10° tilt of the LRR (with respect to its position in the apo structure). Consequently, the hIg and CBM are also displaced (with RMSDs of 8.2 Å and 19.4 Å, respectively). An overlay with respect to the CBM (calculated by aligning Cαs for residues 681–843) results in smaller domain shifts: the adjacent hIg domain is tilted by 4.2°, while the LRR and GH are displaced by 5.1 Å and 6.3 Å, respectively. Therefore, both the GH and CBM domains are rearranged with respect to the rest of the enzyme upon Fc binding. The conformation of the LRR-hIg scaffold is also affected: an overlay with respect to the LRR domain (calculated by aligning Cαs for residues 387–547) results in a 11.6° tilt in the hIg. Thus, the relative domain positioning within EndoS2 is shifted upon binding the Fc substrate, although discrete domains can be fully superimposed and so do not undergo individual conformational changes. Analysis of apo and glycan-bound EndoS2 crystal structures has

revealed no conformational changes upon glycan binding [11], which indicates that EndoS2 domain shifts observed here are solely due to binding the Fc peptide surface.

The binding interface of the EndoS2-IgG1 Fc complex is analysed in Figure 2. The GH and CBM domains within EndoS2 both contact the same Fc Cγ2 domain. This corroborates hydrogen-deuterium exchange data of the EndoS2-rituximab complex, which suggested a role for both domains in complex formation [11]. Crystal structures of EndoS2 in complex with oligomannose- and complex-type glycan substrates have identified GH amino acids involved in glycan binding [11]; the EndoS2-Fc complex structure additionally reveals interactions with the first GlcNAc and fucose moieties within the glycan (Figure 2b). The aromatic side chains of F189, Y251, Y252 and W297 provide hydrophobic stacking interactions, while the Y252 backbone and Q255 side chain form hydrogen bonds with the fucose (Figure 2b).

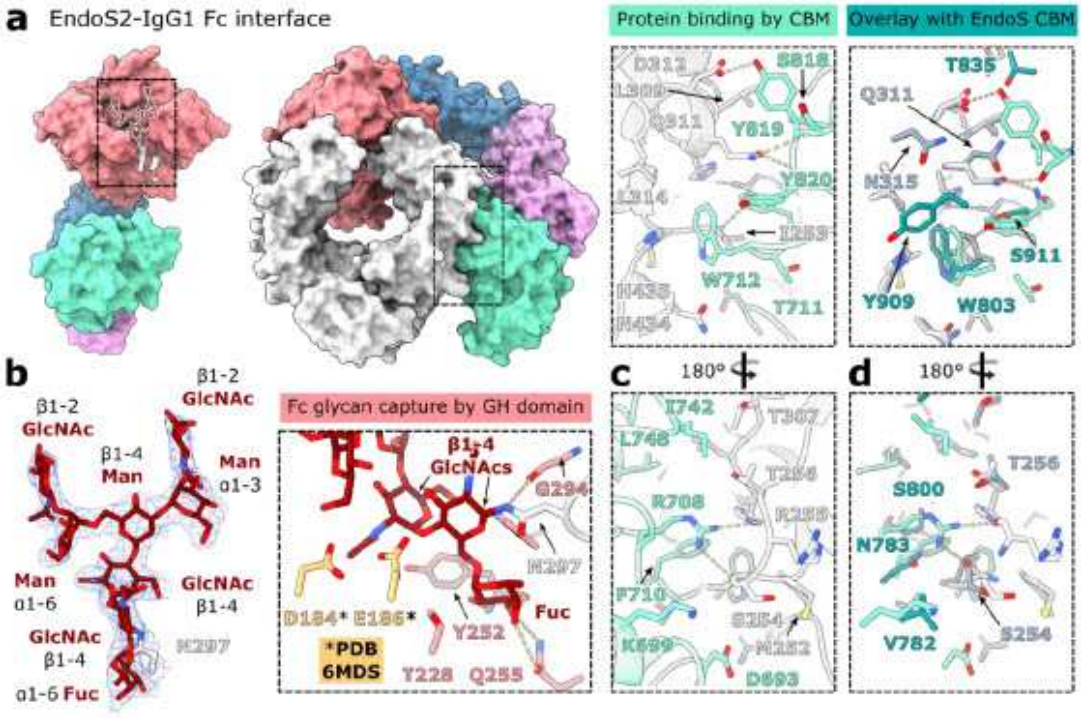


Figure 2: Binding interface of the EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex. a Overall structure of EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ complex, with interfacing areas highlighted. **b** N-linked glycan capture by the EndoS2 GH domain. Final $2F_{obs}$–F_{calc} electron density map for the N-linked glycan and residue N297 is shown (weighted at 1.5 σ). Approximate positions of catalytic residues D184 and E186 (from an overlay with PDB 6MDS *via* the GH domain: aligned Cαs of residues 43–386) at the active site are indicated. **c** Fc protein surface binding by EndoS2 CBM. EndoS2 and IgG1 Fc are coloured as in Figure 1; residues involved in binding are shown as sticks and coloured by heteroatom. Hydrogen bonds are depicted as yellow dashes. **d** Overlay of EndoS2 CBM interface with EndoS CBM (PDB code 8A49). EndoS residues are coloured teal; IgG1 Fc complexed with EndoS is coloured dark grey. Residues involved in interface formation were identified using PDBePISA [71].

We observe atypical ϕ and ψ torsion angles for the β1–4 linkage between the two core GlcNAcs of 5.9° and –121.2°, respectively, using C1–O–C(x)′–C(x–1)′ as the definition of ψ. These values lie outside the range of average glycosidic linkages reported in crystal structures (–73.7° ± 8.4 for ϕ; 116.8° ± 15.6 for ψ angles, calculated from 163 structures) [72]. In a catalytic context, the distortion of the torsion angles at the site of hydrolysis is consistent with promoting cleavage of the β1–4 linkage. Such glycan distortion is also observed within the EndoS-Fc complex, with equivalent ϕ and ψ torsion angles of –58.6° and –121.2°, respectively [55]. In addition, several crystal structures of endoglycosidases in complex with their cleaved glycan substrate show the second GlcNAc adopting a higher-energy, skew-boat conformation, and so

glycan distortion been suggested to comprise part of the catalytic cycle in such enzymes [11, 73, 74]. As observed for the structure of EndoS2 in complex with a complex-type glycan (PDB 6MDS [11]), the second GlcNAc is bound close to the catalytic dyad residues D184 and E186 (exchanged to alanine and leucine in our structure, respectively; Figure 2b), as indicated with an overlay of wild-type, apo EndoS2 (PDB 6MDS) via the GH domain. In this approximate position, E186 is oriented in close proximity to the distorted β1–4 linkage between the two GlcNAcs, facilitating nucleophilic attack and subsequent cleavage of the glycan.

The CBM interface in the EndoS2-IgG1 Fc complex is analysed in Figure 2c. Amino acid W712 in EndoS2 has been shown to be important for Fc binding, as substitution to alanine abolished hydrolytic activity towards IgG Fc bearing complex-type or oligomannose-type glycans [11]. The EndoS2-Fc structure reveals how W712 binds within a hydrophobic pocket at the Fc Cγ2-Cγ3 interface, formed by I253, H310, L314, N434 and H435. Thus, W712 makes similar interactions as W803 in EndoS, similarly shown to be essential for enzyme activity [53] and observed binding in the same cavity within the EndoS-Fc crystal structure [55] (Figure 2d). Another important amino acid in EndoS2 is Y820: Klontz *et al.* found that a serine exchange variant of this side chain severely reduced hydrolytic activity towards IgG bearing both complex- and oligomannose-type glycans [11]. This side chain makes not only a main chain contact to Fc residue Q311, but also forms stacking interactions with the Q311 side chain (Figure 2c). The equivalent amino acid in EndoS is a serine, similarly providing a main chain interaction with Q311 but lacking the hydrophobic interaction [55]. A further parallel between the two structures is observed with the side chain of EndoS2 Y819 that hydrogen bonds to the D312 side chain from the Fc, similar to the equivalent interactions of the EndoS T835 side chain [55].

Comparison of the EndoS2 and EndoS CBM interfaces with IgG1 Fc reveals some unique interactions formed by EndoS2, such as the side chain of R708 that hydrogen bonds to the Fc backbone (Figure 2c and 2d). PISA analysis [71] also predicts distinct Fc protein surface interfaces across the two enzymes: the EndoS2-IgG1 Fc interface has a surface area of 977.9 Å^2 and a solvation free energy gain of –8.6 kcal/mol. Conversely, EndoS forms a 1323.5 Å^2 interface with the Fc, with a –9.1 kcal/mol solvation free energy gain. A sequence alignment of EndoS and EndoS2 reveals that side chains interfacing with the Fc protein are not particularly conserved (Supplementary Figure 4). Moreover, aside from residue W712 (W803 in EndoS), other CBM residues found to be important for IgG Fc binding (R908 and E833 in EndoS [53];

Y820 in EndoS2 [11]) are not conserved (Supplementary Figure 4). These observations indicate distinct modes of IgG recognition by the two endoglycosidases.

Although the two enzymes bind IgG1 Fc in a similar way, a comparison of the EndoS2-Fc and EndoS-Fc crystal structures reveals differing angles of recognition between the two enzymes. Superposition of the two complexes with respect to the interfacing Fc Cγ2 domain (calculated by aligning Cαs of residues 237–340) shows how different GH domain and CBM orientations allow the enzymes to discover different surfaces on the Fc peptide (Figure 3a). This is accompanied with distinct Fc surface interfaces recognised by the two enzymes (Figure 3b). We additionally observe differential capture of the Fc by the two enzymes: superposition of EndoS2- and EndoS-bound Fc by their Cγ3 domains (alignment of Cαs, residues 341–444) reveals more "open" Cγ2 domain placements in the former (Figure 3c). However, an overlay of the interfacing Cγ2 (alignment of Cαs, residues 238–340) within the complexed Fcs and the equivalent domain from a wild-type IgG1 Fc structure (PDB 3AVE) shows that conformational changes only occur in the C′E loop, allowing capture of the Fc glycan in a "flipped-out" state (Figure 3b) [55]. A superposition of the GH domains from EndoS- and EndoS2-Fc complexes shows that the interfacing Fc Cγ2 domains are captured in different orientations (Supplementary Figure 5), which may reflect the altered Cγ2 domain placements observed here.

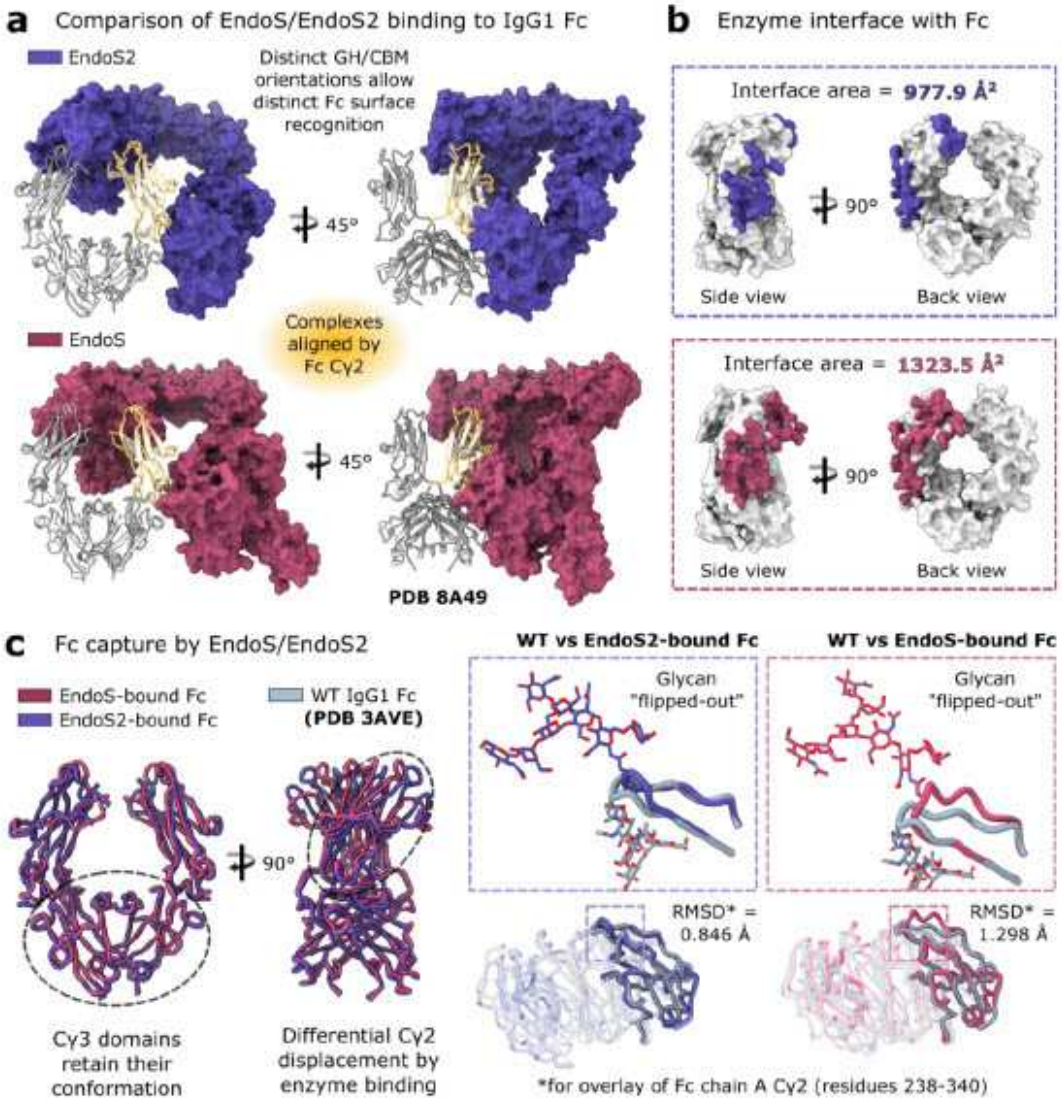


Figure 3: Comparison of EndoS and EndoS2 binding their IgG1 Fc substrate. a EndoS2-Fc and EndoS-Fc crystal structures, superimposed with respect to the interfacing Fc Cγ2 domain (Cαs for residues 237–340). Complexed EndoS (PDB 8A49; [55]) and EndoS2 are depicted in maroon and purple, respectively; complexed IgG1 Fc is depicted in silver. **b** Binding interfaces of each enzyme mapped onto IgG1 Fc, as calculated by PDBePISA [71]. **c** Comparison of Fc capture by EndoS and EndoS2. Superposition of EndoS-bound and EndoS2-bound Fc relative to the Cγ3 homodimer (Cαs for residues 341–444) reveals differential shifting of Cγ2 domains upon enzyme binding. Fc N-linked glycans (shown as sticks and coloured by heteroatom) are captured by both enzymes in a "flipped-out" conformation, occurring with shifts in the C′E loop which is covalently linked to the glycan (relative to a wild-type IgG1 Fc structure: PDB code 3AVE, depicted in dark grey). EndoS-bound and EndoS2-bound Fc are coloured maroon and purple, respectively.

The observed differences in EndoS and EndoS2 binding to IgG1 Fc, in distinct protein-protein interfaces and differing angles of recognition for the GH and CBM domains (Figure 3), provide a structural rationale for experiments with chimeric EndoS/EndoS2 enzymes, which showed that substitution of the EndoS2 GH onto an EndoS scaffold was not enough to confer EndoS2-like activity, but rather required substitution of both EndoS2 GH and CBM domains [11]. This chimeric enzyme was still not fully functional with respect to wild-type EndoS2 [11], thus the LRR–hIgG scaffold must also play a role in orienting the GH and CBM domains optimally. We envisage that our crystal structure can guide future efforts for synthesis of chimeric enzymes, which have application in engineering antibody glycosylation. Enzymes using EndoS/EndoS2 scaffolds are of particular interest due to their exquisite specificity for IgG antibodies: for example, a recent paper by Fan *et al*. shows how the glycosidase domains in EndoS/EndoS2 can be replaced with an α-L-fucosidase from *Lactobacillus casei* BL23 for significantly enhanced antibody defucosylation, when compared to the native fucosidase [75].

Taken together, our crystal structure of EndoS2 in complex with IgG1 Fc reveals the molecular basis behind its extensive substrate specificity. This structural knowledge will aid in the continued development of EndoS2, and chimeric EndoS/EndoS2-based enzymes, as tools for improved antibody glycoengineering and as biologics for use in immune modulation.

Protein Data Bank accession number

Model coordinates and structure factors for the EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystal structure have been deposited in the Protein Data Bank with accession number 8Q5U.

Acknowledgements

We are grateful to the beamline scientists on I03 at Diamond Light Source (DLS), and to Chris Holes for his support with the macromolecular crystallisation facility at the University of Southampton. We are also grateful to John Butler for his continued support and encouragement. This work was supported by DLS (MX29835) and the School of Biological Sciences, University of Southampton.

CRediT authorship contribution statement

Abigail Sudol: Conceptualization, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Ivo Tews:** Investigation, Formal analysis, Supervision, Writing - review & editing, Funding acquisition. **Max Crispin:** Conceptualization, Investigation, Formal analysis, Supervision, Writing - original draft, Writing - review & editing, Funding acquisition.

References

[1] Laabei M, Ermert D. Catch me if you can: *Streptococcus pyogenes* complement evasion strategies. J Innate Immun. 2019;11:3-12.

[2] de Taeye SW, Rispens T, Vidarsson G. The Ligands for Human IgG and Their Effector Functions. Antibodies (Basel). 2019;8:30.

[3] von Pawel-Rammingen U, Johansson BP, Björck L. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. EMBO J. 2002;21:1607-15.

[4] Ryan MH, Petrone D, Nemeth JF, Barnathan E, Björck L, Jordan RE. Proteolysis of purified IgGs by human and bacterial enzymes *in vitro* and the detection of specific proteolytic fragments of endogenous IgG in rheumatoid synovial fluid. Mol Immunol. 2008;45:1837-46.

[5] Nelson DC, Garbe J, Collin M. Cysteine proteinase SpeB from Streptococcus pyogenes - a potent modifier of immunologically important host and bacterial proteins. Biol Chem. 2011;392:1077-88.

[6] Collin M, Olsén A. Effect of SpeB and EndoS from Streptococcus pyogenes on human immunoglobulins. Infect Immun. 2001;69:7187-9.

[7] Collin M, Olsén A. EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. EMBO J. 2001;20:3046-55.

[8] Sjögren J, Struwe WB, Cosgrave EF, Rudd PM, Stervander M, Allhorn M, et al. EndoS2 is a unique and conserved enzyme of serotype M49 group A *Streptococcus* that hydrolyses N-linked glycans on IgG and α1-acid glycoprotein. Biochem J. 2013;455:107-18.

[9] Goodfellow JJ, Baruah K, Yamamoto K, Bonomelli C, Krishna B, Harvey DJ, et al. An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. J Am Chem Soc. 2012;134:8030-3.

[10] Sjögren J, Cosgrave EF, Allhorn M, Nordgren M, Björk S, Olsson F, et al. EndoS and EndoS2 hydrolyze Fc-glycans on therapeutic antibodies with different glycoform selectivity

and can be used for rapid quantification of high-mannose glycans. Glycobiology. 2015;25:1053-63.

[11] Klontz EH, Trastoy B, Deredge D, Fields JK, Li C, Orwenyo J, et al. Molecular basis of broad spectrum N-glycan specificity and processing of therapeutic IgG monoclonal antibodies by endoglycosidase S2. ACS Cent Sci. 2019;5:524-38.

[12] Trastoy B, Klontz E, Orwenyo J, Marina A, Wang LX, Sundberg EJ, et al. Structural basis for the recognition of complex-type N-glycans by endoglycosidase S. Nat Commun. 2018;9:1874.

[13] Kiyoshi M, Tsumoto K, Ishii-Watabe A, Caaveiro JMM. Glycosylation of IgG-Fc: a molecular perspective. Int Immunol. 2017;29:311-7.

[14] Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. J Mol Biol. 2003;325:979-89.

[15] Zheng K, Bantog C, Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. MAbs. 2011;3:568-76.

[16] Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, et al. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. Mol Immunol. 2000;37:697-706.

[17] Yamaguchi Y, Nishimura M, Nagano M, Yagi H, Sasakawa H, Uchida K, et al. Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. Biochim Biophys Acta. 2006;1760:693-700.

[18] Zheng K, Yarmarkovich M, Bantog C, Bayer R, Patapoff TW. Influence of glycosylation pattern on the molecular properties of monoclonal antibodies. MAbs. 2014;6:649-58.

[19] Baruah K, Bowden TA, Krishna BA, Dwek RA, Crispin M, Scanlan CN. Selective deactivation of serum IgG: a general strategy for the enhancement of monoclonal antibody receptor interactions. J Mol Biol. 2012;420:1-7.

[20] Wada R, Matsui M, Kawasaki N. Influence of N-glycosylation on effector functions and thermal stability of glycoengineered IgG1 monoclonal antibody with homogeneous glycoforms. MAbs. 2019;11:350-72.

[21] Allhorn M, Olin AI, Nimmerjahn F, Collin M. Human IgG/Fc gamma R interactions are modulated by streptococcal IgG glycan hydrolysis. PLoS One. 2008;3:e1413.

[22] Lux A, Yu X, Scanlan CN, Nimmerjahn F. Impact of immune complex size and glycosylation on IgG binding to human FcγRs. J Immunol. 2013;190:4315-23.

[23] Collin M, Svensson MD, Sjöholm AG, Jensenius JC, Sjöbring U, Olsén A. EndoS and SpeB from Streptococcus pyogenes inhibit immunoglobulin-mediated opsonophagocytosis. Infect Immun. 2002;70:6646-51.

[24] Naegeli A, Bratanis E, Karlsson C, Shannon O, Kalluru R, Linder A, et al. Streptococcus pyogenes evades adaptive immunity through specific IgG glycan hydrolysis. J Exp Med. 2019;216:1615-29.

[25] Yang R, Otten MA, Hellmark T, Collin M, Björck L, Zhao MH, et al. Successful treatment of experimental glomerulonephritis with IdeS and EndoS, IgG-degrading streptococcal enzymes. Nephrol Dial Transplant. 2010;25:2479-86.

[26] Lood C, Allhorn M, Lood R, Gullstrand B, Olin AI, Rönnblom L, et al. IgG glycan hydrolysis by endoglycosidase S diminishes the proinflammatory properties of immune complexes from patients with systemic lupus erythematosus: a possible new treatment? Arthritis Rheum. 2012;64:2698-706.

[27] Nandakumar KS, Collin M, Happonen KE, Lundström SL, Croxford AM, Xu B, et al. Streptococcal Endo-β-*N*-acetylglucosaminidase suppresses antibody-mediated inflammation *in vivo*. Front Immunol. 2018;9:1623.

[28] Segelmark M, Björck L. Streptococcal enzymes as precision tools against pathogenic IgG autoantibodies in small vessel vasculitis. Front Immunol. 2019;10:2165.

[29] Collin M, Shannon O, Björck L. IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions. Proc Natl Acad Sci U S A. 2008;105:4265-70.

[30] Mihai S, Albert H, Ludwig RJ, Iwata H, Björck L, Collin M, et al. In vivo enzymatic modulation of IgG antibodies prevents immune complex-dependent skin injury. Exp Dermatol. 2017;26:691-6.

[31] Albert H, Collin M, Dudziak D, Ravetch JV, Nimmerjahn F. In vivo enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. Proc Natl Acad Sci U S A. 2008;105:15005-9.

[32] Benkhoucha M, Molnarfi N, Santiago-Raber ML, Weber MS, Merkler D, Collin M, et al. IgG glycan hydrolysis by EndoS inhibits experimental autoimmune encephalomyelitis. J Neuroinflammation. 2012;9:209.

[33] Hirose M, Vafia K, Kalies K, Groth S, Westermann J, Zillikens D, et al. Enzymatic autoantibody glycan hydrolysis alleviates autoimmunity against type VII collagen. J Autoimmun. 2012;39:304-14.

[34] Nandakumar KS, Collin M, Olsén A, Nimmerjahn F, Blom AM, Ravetch JV, et al. Endoglycosidase treatment abrogates IgG arthritogenicity: importance of IgG glycosylation in arthritis. Eur J Immunol. 2007;37:2973-82.

[35] Lin J, Boon L, Bockermann R, Robertson AK, Kjellman C, Anderson CC. Desensitization using imlifidase and EndoS enables chimerism induction in allosensitized recipient mice. Am J Transplant. 2020;20:2356-65.

[36] Järnum S, Runström A, Bockermann R, Winstedt L, Crispin M, Kjellman C. Enzymatic inactivation of endogenous IgG by IdeS enhances therapeutic antibody efficacy. Mol Cancer Ther. 2017;16:1887-97.

[37] Le NP, Bowden TA, Struwe WB, Crispin M. Immune recruitment or suppression by glycan engineering of endogenous and therapeutic antibodies. Biochim Biophys Acta. 2016;1860:1655-68.

[38] Wang LX, Tong X, Li C, Giddens JP, Li T. Glycoengineering of Antibodies for Modulating Functions. Annu Rev Biochem. 2019;88:433-59.

[39] Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. J Am Chem Soc. 2012;134:12308-18.

[40] Tong X, Li T, Li C, Wang LX. Generation and comparative kinetic analysis of new glycosynthase mutants from *Streptococcus pyogenes* endoglycosidases for antibody glycoengineering. Biochemistry. 2018;57:5239-46.

[41] Li T, Tong X, Yang Q, Giddens JP, Wang LX. Glycosynthase Mutants of Endoglycosidase S2 Show Potent Transglycosylation Activity and Remarkably Relaxed Substrate Specificity for Antibody Glycosylation Remodeling. J Biol Chem. 2016;291:16508-18.

[42] Shivatare SS, Huang LY, Zeng YF, Liao JY, You TH, Wang SY, et al. Development of glycosynthases with broad glycan specificity for the efficient glyco-remodeling of antibodies. Chem Commun (Camb). 2018;54:6161-4.

[43] Trastoy B, Du JJ, García-Alija M, Li C, Klontz EH, Wang LX, et al. Sculpting therapeutic monoclonal antibody N-glycans using endoglycosidases. Curr Opin Struct Biol. 2022;72:248-59.

[44] Tong X, Li T, Orwenyo J, Toonstra C, Wang LX. One-pot enzymatic glycan remodeling of a therapeutic monoclonal antibody by endoglycosidase S (Endo-S) from Streptococcus pyogenes. Bioorg Med Chem. 2018;26:1347-55.

[45] Shi W, Li W, Zhang J, Li T, Song Y, Zeng Y, et al. One-step synthesis of site-specific antibody-drug conjugates by reprograming IgG glycoengineering with LacNAc-based substrates. Acta Pharm Sin B. 2022;12:2417-28.

[46] Zhang X, Ou C, Liu H, Prabhu SK, Li C, Yang Q, et al. General and Robust Chemoenzymatic Method for Glycan-Mediated Site-Specific Labeling and Conjugation of Antibodies: Facile Synthesis of Homogeneous Antibody-Drug Conjugates. ACS Chem Biol. 2021;16:2502-14.

[47] Golay J, Da Roit F, Bologna L, Ferrara C, Leusen JH, Rambaldi A, et al. Glycoengineered CD20 antibody obinutuzumab activates neutrophils and mediates phagocytosis through CD16B more efficiently than rituximab. Blood. 2013;122:3482-91.

[48] Kaneko Y, Nimmerjahn F, Ravetch JV. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. Science. 2006;313:670-3.

[49] Fiedler W, Stoeger H, Perotti A, Gastl G, Weidmann J, Dietrich B, et al. Phase I study of TrasGEX, a glyco-optimised anti-HER2 monoclonal antibody, in patients with HER2-positive solid tumours. ESMO Open. 2018;3:e000381.

[50] Fiedler W, DeDosso S, Cresta S, Weidmann J, Tessari A, Salzberg M, et al. A phase I study of PankoMab-GEX, a humanised glyco-optimised monoclonal antibody to a novel tumour-specific MUC1 glycopeptide epitope in patients with advanced carcinomas. Eur J Cancer. 2016;63:55-63.

[51] Pereira NA, Chan KF, Lin PC, Song Z. The "less-is-more" in therapeutic antibodies: Afucosylated anti-cancer antibodies with enhanced antibody-dependent cellular cytotoxicity. MAbs. 2018;10:693-711.

[52] Dixon EV, Claridge JK, Harvey DJ, Baruah K, Yu X, Vesiljevic S, et al. Fragments of bacterial endoglycosidase S and immunoglobulin G reveal subdomains of each that contribute to deglycosylation. J Biol Chem. 2014;289:13876-89.

[53] Trastoy B, Lomino JV, Pierce BG, Carter LG, Günther S, Giddens JP, et al. Crystal structure of *Streptococcus pyogenes* EndoS, an immunomodulatory endoglycosidase specific for human IgG antibodies. Proc Natl Acad Sci U S A. 2014;111:6714-9.

[54] Aytenfisu AH, Deredge D, Klontz EH, Du J, Sundberg EJ, MacKerell AD. Insights into substrate recognition and specificity for IgG by Endoglycosidase S2. PLoS Comput Biol. 2021;17:e1009103.

[55] Sudol ASL, Butler J, Ivory DP, Tews I, Crispin M. Extensive substrate recognition by the streptococcal antibody-degrading enzymes IdeS and EndoS. Nat Commun. 2022;13:7801.

[56] Trastoy B, Du JJ, Cifuente JO, Rudolph L, García-Alija M, Klontz EH, et al. Mechanism of antibody-specific deglycosylation and immune evasion by Streptococcal IgG-specific endoglycosidases. Nat Commun. 2023;14:1705.

[57] Seki H, Huang Y, Arakawa T, Yamada C, Kinoshita T, Iwamoto S, et al. Structural basis for the specific cleavage of core-fucosylated *N*-glycans by endo-β-*N*-acetylglucosaminidase from the fungus *Cordyceps militaris*. J Biol Chem. 2019;294:17143-54.

[58] Waddling CA, Plummer TH, Tarentino AL, Van Roey P. Structural basis for the substrate specificity of endo-β-*N*-acetylglucosaminidase F3. Biochemistry. 2000;39:7878-85.

[59] García-Alija M, Du JJ, Ordóñez I, Diz-Vallenilla A, Moraleda-Montoya A, Sultana N, et al. Mechanism of cooperative N-glycan processing by the multi-modular endoglycosidase EndoE. Nat Commun. 2022;13:1137.

[60] Emsley P, Crispin M. Structural analysis of glycoproteins: building N-linked glycans with Coot. Acta Crystallogr D Struct Biol. 2018;74:256-63.

[61] Agirre J, Iglesias-Fernández J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. Privateer: software for the conformational validation of carbohydrate structures. Nat Struct Mol Biol. 2015;22:833-4.

[62] Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 2021;30:70-82.

[63] Bowden TA, Baruah K, Coles CH, Harvey DJ, Yu X, Song BD, et al. Chemical and structural analysis of an antibody folding intermediate trapped during glycan biosynthesis. J Am Chem Soc. 2012;134:17554-63.

[64] Matsumiya S, Yamaguchi Y, Saito J, Nagano M, Sasakawa H, Otaki S, et al. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. J Mol Biol. 2007;368:767-79.

[65] Yu X, Baruah K, Harvey DJ, Vasiljevic S, Alonzi DS, Song BD, et al. Engineering hydrophobic protein-carbohydrate interactions to fine-tune monoclonal antibodies. J Am Chem Soc. 2013;135:9723-32.

[66] Frank M, Walker RC, Lanzilotta WN, Prestegard JH, Barb AW. Immunoglobulin G1 Fc domain motions: implications for Fc engineering. J Mol Biol. 2014;426:1799-811.

[67] Barb AW, Meng L, Gao Z, Johnson RW, Moremen KW, Prestegard JH. NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation. Biochemistry. 2012;51:4618-26.

[68] Barb AW, Prestegard JH. NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. Nat Chem Biol. 2011;7:147-53.

[69] Lee HS, Im W. Effects of N-glycan composition on structure and dynamics of IgG1 Fc and their implications for antibody engineering. Sci Rep. 2017;7:12659.

[70] Harbison AM, Brosnan LP, Fenlon K, Fadda E. Sequence-to-structure dependence of isolated IgG Fc complex biantennary N-glycans: a molecular dynamics study. Glycobiology. 2019;29:94-103.

[71] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774-97.

[72] Petrescu AJ, Petrescu SM, Dwek RA, Wormald MR. A statistical analysis of N- and O-glycan linkage conformations from crystallographic data. Glycobiology. 1999;9:343-52.

[73] Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE. Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nat Struct Biol. 1996;3:638-48.

[74] Du JJ, Klontz EH, Guerin ME, Trastoy B, Sundberg EJ. Structural insights into the mechanisms and specificities of IgG-active endoglycosidases. Glycobiology. 2020;30:268-79.

[75] Fan S, Li W, Zhang K, Zou X, Shi W, Liu Z, et al. Enhanced antibody-defucosylation capability of α-L-fucosidase by proximity-based protein fusion. Biochem Biophys Res Commun. 2023;645:40-6.

Appendix G wwPDB X-ray Structure Validation Report for EndoS2-IgG1 Fc

wwPDB X-ray Structure Validation Report for EndoS2$^{D184A/E186L}$-IgG1 Fc$^{L234C/E382A}$ crystal structure (deposited with code 8Q5U but being held for publication at time of writing). Report is included overleaf:

Full wwPDB X-ray Structure Validation Report (i)

Aug 10, 2023 – 09:22 am BST

PDB ID	:	8Q5U
Title	:	Endoglycosidase S2 in complex with IgG1 Fc
Deposited on	:	2023-08-09
Resolution	:	3.00 Å(reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

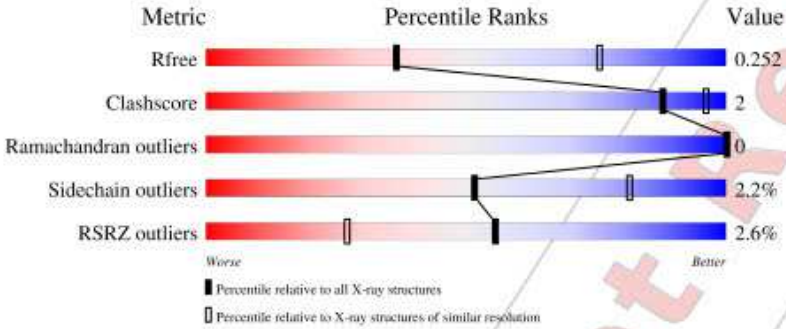
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



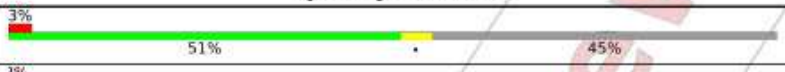
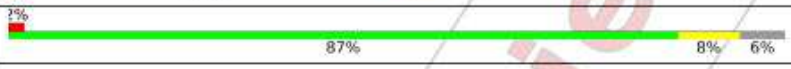
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	227	88% · · 8%
1	B	227	% 88% · 9%
1	C	227	4% 86% 5% 9%
2	D	816	% 89% 7% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	816	
2	F	816	
3	G	8	
4	H	7	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	PGE	D	903	-	-	-	X
6	PGE	F	901	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 41613 atoms, of which 20445 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Uncharacterized protein DKFZp686C11235.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	208	Total	C	H	N	O	S	62	0	0
			3268	1049	1621	276	315	7			
1	C	206	Total	C	H	N	O	S	66	0	0
			3188	1025	1581	271	304	7			
1	B	207	Total	C	H	N	O	S	62	0	0
			3224	1029	1606	275	307	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	234	CYS	LEU	engineered mutation	UNP Q6MZV7
A	382	ALA	GLU	engineered mutation	UNP Q6MZV7
C	234	CYS	LEU	engineered mutation	UNP Q6MZV7
C	382	ALA	GLU	engineered mutation	UNP Q6MZV7
B	234	CYS	LEU	engineered mutation	UNP Q6MZV7
B	382	ALA	GLU	engineered mutation	UNP Q6MZV7

- Molecule 2 is a protein called Endo-beta-N-acetylglucosaminidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	788	Total	C	H	N	O	S	183	0	0
			12398	3936	6137	1069	1241	15			
2	E	447	Total	C	H	N	O	S	133	0	0
			6793	2166	3346	585	687	9			
2	F	769	Total	C	H	N	O	S	192	0	0
			11928	3801	5879	1029	1204	15			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	37	MET	-	initiating methionine	UNP A0A8H2N1T2
D	184	ALA	ASP	engineered mutation	UNP A0A8H2N1T2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	186	LEU	GLU	engineered mutation	UNP A0A8H2N1T2
D	844	LEU	-	expression tag	UNP A0A8H2N1T2
D	845	LEU	-	expression tag	UNP A0A8H2N1T2
D	846	GLU	-	expression tag	UNP A0A8H2N1T2
D	847	HIS	-	expression tag	UNP A0A8H2N1T2
D	848	HIS	-	expression tag	UNP A0A8H2N1T2
D	849	HIS	-	expression tag	UNP A0A8H2N1T2
D	850	HIS	-	expression tag	UNP A0A8H2N1T2
D	851	HIS	-	expression tag	UNP A0A8H2N1T2
D	852	HIS	-	expression tag	UNP A0A8H2N1T2
E	37	MET	-	initiating methionine	UNP A0A8H2N1T2
E	184	ALA	ASP	engineered mutation	UNP A0A8H2N1T2
E	186	LEU	GLU	engineered mutation	UNP A0A8H2N1T2
E	844	LEU	-	expression tag	UNP A0A8H2N1T2
E	845	LEU	-	expression tag	UNP A0A8H2N1T2
E	846	GLU	-	expression tag	UNP A0A8H2N1T2
E	847	HIS	-	expression tag	UNP A0A8H2N1T2
E	848	HIS	-	expression tag	UNP A0A8H2N1T2
E	849	HIS	-	expression tag	UNP A0A8H2N1T2
E	850	HIS	-	expression tag	UNP A0A8H2N1T2
E	851	HIS	-	expression tag	UNP A0A8H2N1T2
E	852	HIS	-	expression tag	UNP A0A8H2N1T2
F	37	MET	-	initiating methionine	UNP A0A8H2N1T2
F	184	ALA	ASP	engineered mutation	UNP A0A8H2N1T2
F	186	LEU	GLU	engineered mutation	UNP A0A8H2N1T2
F	844	LEU	-	expression tag	UNP A0A8H2N1T2
F	845	LEU	-	expression tag	UNP A0A8H2N1T2
F	846	GLU	-	expression tag	UNP A0A8H2N1T2
F	847	HIS	-	expression tag	UNP A0A8H2N1T2
F	848	HIS	-	expression tag	UNP A0A8H2N1T2
F	849	HIS	-	expression tag	UNP A0A8H2N1T2
F	850	HIS	-	expression tag	UNP A0A8H2N1T2
F	851	HIS	-	expression tag	UNP A0A8H2N1T2
F	852	HIS	-	expression tag	UNP A0A8H2N1T2

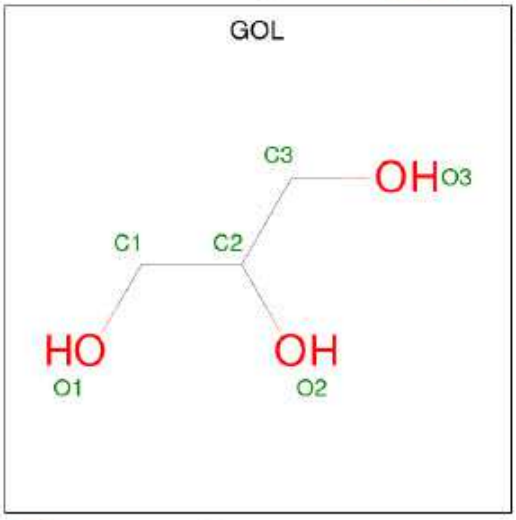
- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	8	Total	C	H	N	O	18	0	0
			195	56	96	4	39			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	7	Total	C	H	N	O	15	0	0
			167	48	82	3	34			

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



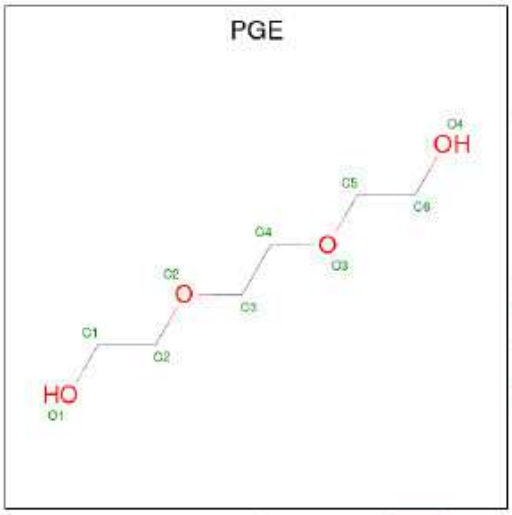
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	2	0
			14	3	8	3		
5	D	1	Total	C	H	O	2	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C H O 14 3 8 3	2	0
5	B	1	Total C H O 14 3 8 3	2	0

- Molecule 6 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total C H O 24 6 14 4	1	0
6	D	1	Total C H O 24 6 14 4	1	0
6	D	1	Total C H O 24 6 14 4	1	0
6	F	1	Total C H O 16 4 9 3	1	0
6	F	1	Total C H O 24 6 14 4	1	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total Ca 1 1	0	0
7	E	1	Total Ca 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	1	Total Ca 1 1	0	0

- Molecule 8 is water.

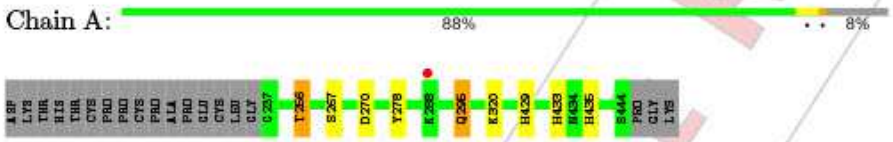
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	43	Total O 43 43	0	0
8	D	105	Total O 105 105	0	0
8	C	1	Total O 1 1	0	0
8	E	19	Total O 19 19	0	0
8	F	51	Total O 51 51	0	0
8	B	62	Total O 62 62	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

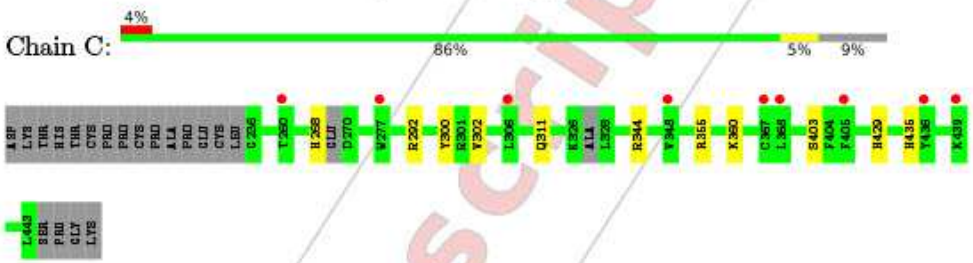
- Molecule 1: Uncharacterized protein DKFZp686C11235

Chain A:



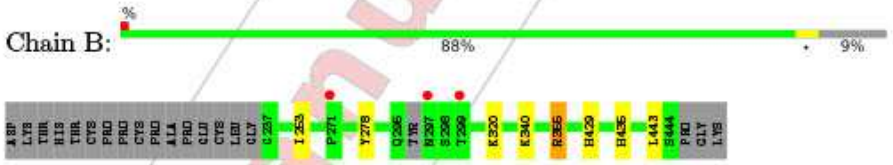
- Molecule 1: Uncharacterized protein DKFZp686C11235

Chain C:



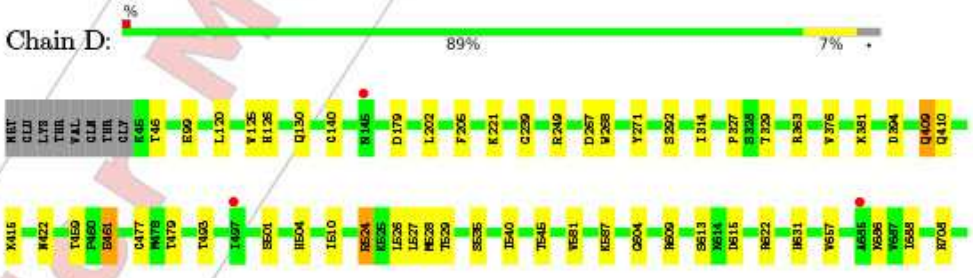
- Molecule 1: Uncharacterized protein DKFZp686C11235

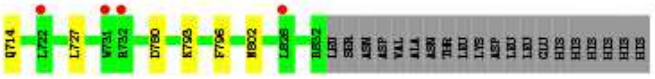
Chain B:



- Molecule 2: Endo-beta-N-acetylglucosaminidase

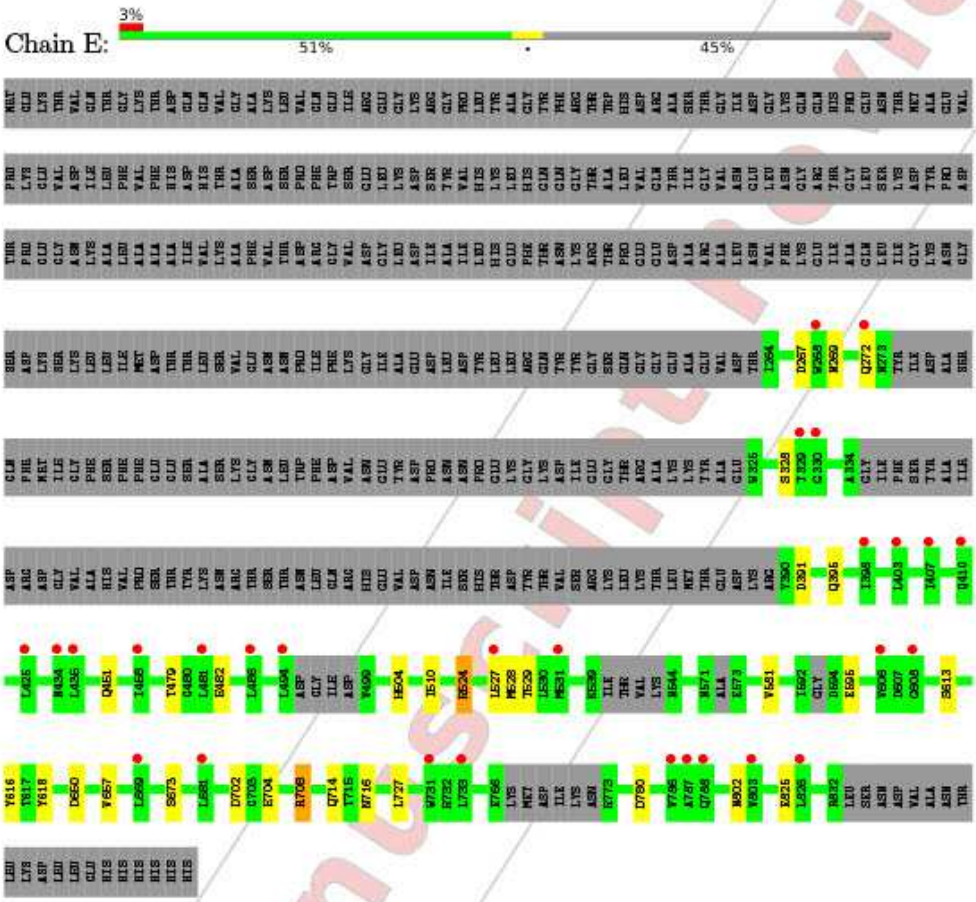
Chain D:





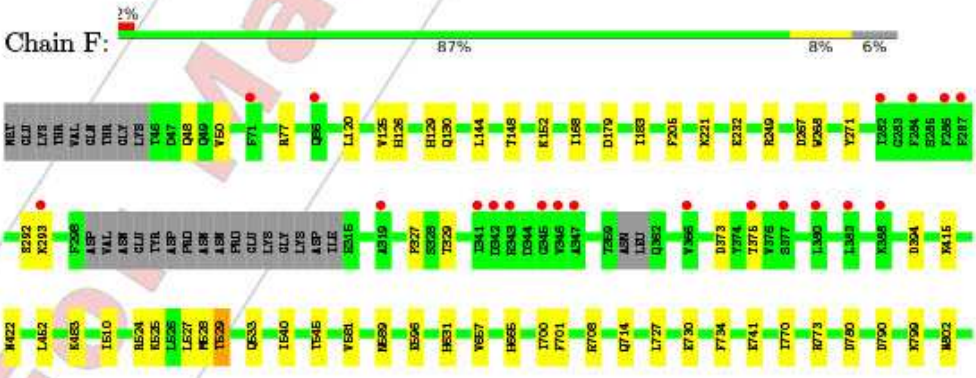
• Molecule 2: Endo-beta-N-acetylglucosaminidase

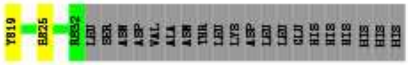
Chain E:



• Molecule 2: Endo-beta-N-acetylglucosaminidase

Chain F:





• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	228.78Å 228.78Å 161.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.91 – 3.00 49.91 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.91-3.00) 93.7 (49.91-3.00)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.52 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.216 , 0.252 0.221 , 0.252	Depositor DCC
R_{free} test set	4385 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	79.4	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 56.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	41613	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MAN, FUC, GOL, PGE, CA, BMA, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.30	0/1693	0.55	0/2309
1	B	0.30	0/1660	0.54	0/2262
1	C	0.30	0/1650	0.54	0/2249
2	D	0.28	0/6379	0.53	0/8619
2	E	0.28	0/3500	0.52	0/4732
2	F	0.28	0/6161	0.51	0/8334
All	All	0.28	0/21043	0.53	0/28505

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	4
2	E	0	1
2	F	0	1
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	363	ARG	Sidechain
2	D	524	ARG	Sidechain
2	D	609	ARG	Sidechain
2	D	622	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	E	524	ARG	Sidechain
2	F	524	ARG	Sidechain

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1647	1621	1603	6	0
1	B	1618	1606	1581	6	0
1	C	1607	1581	1548	6	0
2	D	6261	6137	6102	29	0
2	E	3447	3346	3262	13	0
2	F	6049	5879	5814	24	0
3	G	99	96	85	1	0
4	H	85	82	73	1	0
5	A	6	8	8	1	0
5	B	12	16	16	0	0
5	D	6	8	8	0	0
6	D	30	42	42	4	0
6	F	17	23	23	0	0
7	D	1	0	0	0	0
7	E	1	0	0	0	0
7	F	1	0	0	0	0
8	A	43	0	0	2	0
8	B	62	0	0	2	0
8	C	1	0	0	0	0
8	D	105	0	0	1	0
8	E	19	0	0	1	0
8	F	51	0	0	2	0
All	All	21168	20445	20165	83	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (83) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:144:LEU:HD11	2:F:183:ILE:HD13	1.67	0.76
2:F:452:LEU:O	8:F:1001:HOH:O	2.11	0.68
2:E:595:GLU:HG3	2:E:618:TYR:CD2	2.29	0.68
2:F:148:THR:HA	2:F:152:LYS:HE2	1.75	0.68
1:A:295:GLN:HG2	8:A:635:HOH:O	1.95	0.67
1:B:443:LEU:O	8:B:601:HOH:O	2.14	0.66
2:F:665:HIS:ND1	8:F:1002:HOH:O	2.32	0.60
2:D:587:LYS:HE2	6:D:901:PGE:H4	1.84	0.59
1:A:256:THR:OG1	8:A:601:HOH:O	2.15	0.59
2:F:529:THR:O	2:F:533:GLN:HG2	2.02	0.58
2:D:540:ILE:HG23	2:D:545:THR:HG21	1.86	0.57
2:E:702:ASP:O	2:E:825:GLU:OE2	2.25	0.55
2:F:373:ASP:OD2	2:F:375:THR:OG1	2.25	0.54
2:F:540:ILE:HG23	2:F:545:THR:HG21	1.90	0.54
2:D:524:ARG:O	2:D:528:MET:HG2	2.08	0.53
4:H:1:NAG:O3	4:H:2:NAG:H2	2.09	0.53
1:C:311:GLN:OE1	2:F:819:TYR:HA	2.10	0.52
2:D:510:ILE:HD11	2:D:527:LEU:HD13	1.93	0.51
2:D:604:GLY:O	8:D:1001:HOH:O	2.18	0.51
2:F:394:ASP:OD1	2:F:415:LYS:NZ	2.43	0.51
2:D:99:GLU:O	2:D:381:LYS:HG3	2.09	0.51
1:B:355:ARG:NH2	8:B:604:HOH:O	2.36	0.51
2:D:329:THR:HB	2:D:422:ASN:ND2	2.25	0.51
2:E:595:GLU:HG3	2:E:618:TYR:CE2	2.45	0.51
2:D:394:ASP:OD1	2:D:415:LYS:NZ	2.44	0.50
2:E:524:ARG:O	2:E:528:MET:HG2	2.11	0.50
2:D:409:GLN:HG3	2:D:410:GLN:HG3	1.93	0.50
2:D:459:THR:OG1	2:D:461:GLU:HG2	2.11	0.49
2:F:329:THR:HB	2:F:422:ASN:ND2	2.27	0.49
1:C:355:ARG:H	1:C:355:ARG:CD	2.24	0.49
2:F:510:ILE:HD11	2:F:527:LEU:HD13	1.95	0.49
2:F:179:ASP:O	2:F:221:LYS:HA	2.13	0.48
1:A:267:SER:OG	1:A:270:ASP:OD1	2.32	0.48
2:D:179:ASP:O	2:D:221:LYS:HA	2.14	0.48
2:D:126:HIS:O	2:D:130:GLN:HG2	2.14	0.48
2:D:581:VAL:HG21	2:D:657:VAL:HG21	1.96	0.48
2:E:510:ILE:HD11	2:E:527:LEU:HD13	1.96	0.48
2:F:581:VAL:HG21	2:F:657:VAL:HG21	1.96	0.48
2:E:581:VAL:HG21	2:E:657:VAL:HG21	1.95	0.47
2:F:734:PHE:CD2	2:F:825:GLU:HB3	2.50	0.47
2:F:144:LEU:CD2	2:F:168:ILE:HD13	2.45	0.46
2:D:314:ILE:HG22	2:D:376:VAL:HG13	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:278:TYR:HB2	1:B:320:LYS:HB3	1.98	0.46
2:D:140:GLY:HA3	3:G:7:NAG:H81	1.98	0.46
2:E:650:ASP:OD2	2:E:673:SER:OG	2.22	0.46
2:D:535:SER:HB3	2:D:540:ILE:HD11	1.98	0.45
2:F:126:HIS:O	2:F:130:GLN:HG3	2.16	0.45
2:D:268:TRP:CD1	2:D:327:PRO:HB3	2.52	0.45
2:E:269:ASN:O	2:E:272:GLN:HG2	2.17	0.45
1:B:355:ARG:NE	1:B:355:ARG:H	2.14	0.45
2:F:700:ILE:HG13	2:F:701:PHE:CE2	2.52	0.44
2:F:734:PHE:HD2	2:F:825:GLU:HB3	1.83	0.44
2:D:477:GLY:HA2	2:D:501:SER:HB2	1.99	0.44
2:E:613:SER:OG	2:E:616:TYR:HB2	2.18	0.44
2:F:268:TRP:CD1	2:F:327:PRO:HB3	2.53	0.44
2:D:686:LYS:HD2	2:D:688:ILE:HG22	2.00	0.44
1:C:344:ARG:NH1	1:C:403:SER:HB3	2.33	0.44
2:D:120:LEU:O	2:D:125:VAL:HG23	2.19	0.43
2:F:727:LEU:HD11	2:F:802:ASN:HB3	2.00	0.43
2:D:793:LYS:HE3	2:D:796:PHE:CD2	2.54	0.43
1:A:278:TYR:HB2	1:A:320:LYS:HB3	2.01	0.43
2:D:479:THR:O	2:D:504:HIS:ND1	2.47	0.43
2:E:727:LEU:HD11	2:E:802:ASN:HB3	2.01	0.43
1:B:429:HIS:O	1:B:435:HIS:HA	2.19	0.42
2:F:120:LEU:O	2:F:125:VAL:HG23	2.19	0.42
2:E:716:ASN:OD1	8:E:1001:HOH:O	2.21	0.42
1:A:433:HIS:HA	5:A:501:GOL:H32	2.01	0.42
2:E:708:ARG:NH1	1:B:253:ILE:O	2.53	0.42
2:D:267:ASP:HB3	2:D:271:TYR:CE2	2.55	0.42
2:F:267:ASP:HB3	2:F:271:TYR:CE2	2.55	0.42
2:D:727:LEU:HD11	2:D:802:ASN:HB3	2.01	0.41
2:D:493:THR:HG22	6:D:901:PGE:H52	2.03	0.41
2:D:587:LYS:HE2	6:D:901:PGE:H5	2.02	0.41
1:C:429:HIS:O	1:C:435:HIS:HA	2.21	0.41
1:C:268:HIS:CD2	1:C:300:TYR:HH	2.38	0.41
1:C:292:ARG:CG	1:C:302:VAL:HG22	2.51	0.41
2:D:202:LEU:HD11	2:D:239:GLY:HA3	2.02	0.41
2:D:613:SER:OG	2:D:615:ASP:OD1	2.36	0.41
2:E:479:THR:O	2:E:504:HIS:ND1	2.45	0.41
1:A:429:HIS:O	1:A:435:HIS:HA	2.21	0.40
2:D:587:LYS:HE2	6:D:901:PGE:C4	2.50	0.40
2:F:50:VAL:HG11	2:F:129:HIS:HB3	2.02	0.40
2:F:770:ILE:HD13	2:F:773:ARG:NH2	2.37	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	206/227 (91%)	203 (98%)	3 (2%)	0	100	100
1	B	203/227 (89%)	201 (99%)	2 (1%)	0	100	100
1	C	200/227 (88%)	196 (98%)	4 (2%)	0	100	100
2	D	786/816 (96%)	756 (96%)	30 (4%)	0	100	100
2	E	431/816 (53%)	411 (95%)	20 (5%)	0	100	100
2	F	763/816 (94%)	732 (96%)	31 (4%)	0	100	100
All	All	2589/3129 (83%)	2499 (96%)	90 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	190/209 (91%)	188 (99%)	2 (1%)	73	90
1	B	185/209 (88%)	183 (99%)	2 (1%)	73	90
1	C	181/209 (87%)	180 (99%)	1 (1%)	86	95
2	D	678/712 (95%)	666 (98%)	12 (2%)	59	85
2	E	360/712 (51%)	349 (97%)	11 (3%)	40	75
2	F	645/712 (91%)	624 (97%)	21 (3%)	38	73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	2239/2763 (81%)	2190 (98%)	49 (2%)	52	81

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	256	THR
1	A	295	GLN
2	D	46	THR
2	D	205	PHE
2	D	249	ARG
2	D	292	SER
2	D	409	GLN
2	D	461	GLU
2	D	526	LEU
2	D	529	THR
2	D	631	HIS
2	D	708	ARG
2	D	714	GLN
2	D	780	ASP
1	C	360	LYS
2	E	267	ASP
2	E	328	SER
2	E	391	ASP
2	E	395	GLN
2	E	451	GLN
2	E	482	GLU
2	E	529	THR
2	E	704	GLU
2	E	708	ARG
2	E	714	GLN
2	E	780	ASP
2	F	48	GLN
2	F	77	ARG
2	F	205	PHE
2	F	232	GLU
2	F	249	ARG
2	F	292	SER
2	F	293	LYS
2	F	483	LYS
2	F	525	LYS
2	F	528	MET
2	F	529	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	589	ASN
2	F	596	GLU
2	F	631	HIS
2	F	708	ARG
2	F	714	GLN
2	F	730	GLU
2	F	741	GLU
2	F	780	ASP
2	F	790	ASP
2	F	799	LYS
1	B	340	LYS
1	B	355	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

15 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NAG	G	1	1,3	14,14,15	0.59	0	17,19,21	0.59	0
3	NAG	G	2	3	14,14,15	0.51	0	17,19,21	1.33	1 (5%)
3	BMA	G	3	3	11,11,12	0.64	0	15,15,17	0.94	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	MAN	G	4	3	11,11,12	0.43	0	15,15,17	0.64	0
3	NAG	G	5	3	14,14,15	0.41	0	17,19,21	0.89	1 (5%)
3	MAN	G	6	3	11,11,12	0.53	0	15,15,17	0.70	0
3	NAG	G	7	3	14,14,15	0.29	0	17,19,21	0.77	1 (5%)
3	FUC	G	8	3	10,10,11	0.46	0	14,14,16	0.45	0
4	NAG	H	1	4,1	14,14,15	0.51	0	17,19,21	0.44	0
4	NAG	H	2	4	14,14,15	0.42	0	17,19,21	1.04	1 (5%)
4	BMA	H	3	4	11,11,12	0.39	0	15,15,17	0.55	0
4	MAN	H	4	4	11,11,12	0.38	0	15,15,17	0.68	0
4	NAG	H	5	4	14,14,15	0.55	0	17,19,21	0.59	0
4	MAN	H	6	4	11,11,12	0.43	0	15,15,17	0.59	0
4	FUC	H	7	4	10,10,11	0.44	0	14,14,16	0.61	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	G	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	4/6/23/26	0/1/1/1
3	BMA	G	3	3	-	1/2/19/22	0/1/1/1
3	MAN	G	4	3	-	0/2/19/22	0/1/1/1
3	NAG	G	5	3	-	1/6/23/26	0/1/1/1
3	MAN	G	6	3	-	2/2/19/22	0/1/1/1
3	NAG	G	7	3	-	2/6/23/26	0/1/1/1
3	FUC	G	8	3	-	-	0/1/1/1
4	NAG	H	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	H	2	4	-	3/6/23/26	0/1/1/1
4	BMA	H	3	4	-	0/2/19/22	0/1/1/1
4	MAN	H	4	4	-	0/2/19/22	0/1/1/1
4	NAG	H	5	4	-	1/6/23/26	0/1/1/1
4	MAN	H	6	4	-	1/2/19/22	0/1/1/1
4	FUC	H	7	4	-	-	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	2	NAG	O5-C1-C2	4.60	118.56	111.29

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	2	NAG	O5-C1-C2	3.28	116.46	111.29
3	G	7	NAG	O5-C1-C2	2.29	114.91	111.29
3	G	5	NAG	O5-C1-C2	2.06	114.54	111.29

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	2	NAG	O5-C5-C6-O6
4	H	2	NAG	C4-C5-C6-O6
3	G	6	MAN	O5-C5-C6-O6
3	G	6	MAN	C4-C5-C6-O6
3	G	7	NAG	C8-C7-N2-C2
3	G	7	NAG	O7-C7-N2-C2
3	G	2	NAG	O5-C5-C6-O6
3	G	5	NAG	C8-C7-N2-C2
4	H	2	NAG	C1-C2-N2-C7
3	G	2	NAG	C4-C5-C6-O6
4	H	5	NAG	C4-C5-C6-O6
4	H	6	MAN	O5-C5-C6-O6
3	G	2	NAG	C1-C2-N2-C7
3	G	2	NAG	C3-C2-N2-C7
3	G	3	BMA	O5-C5-C6-O6

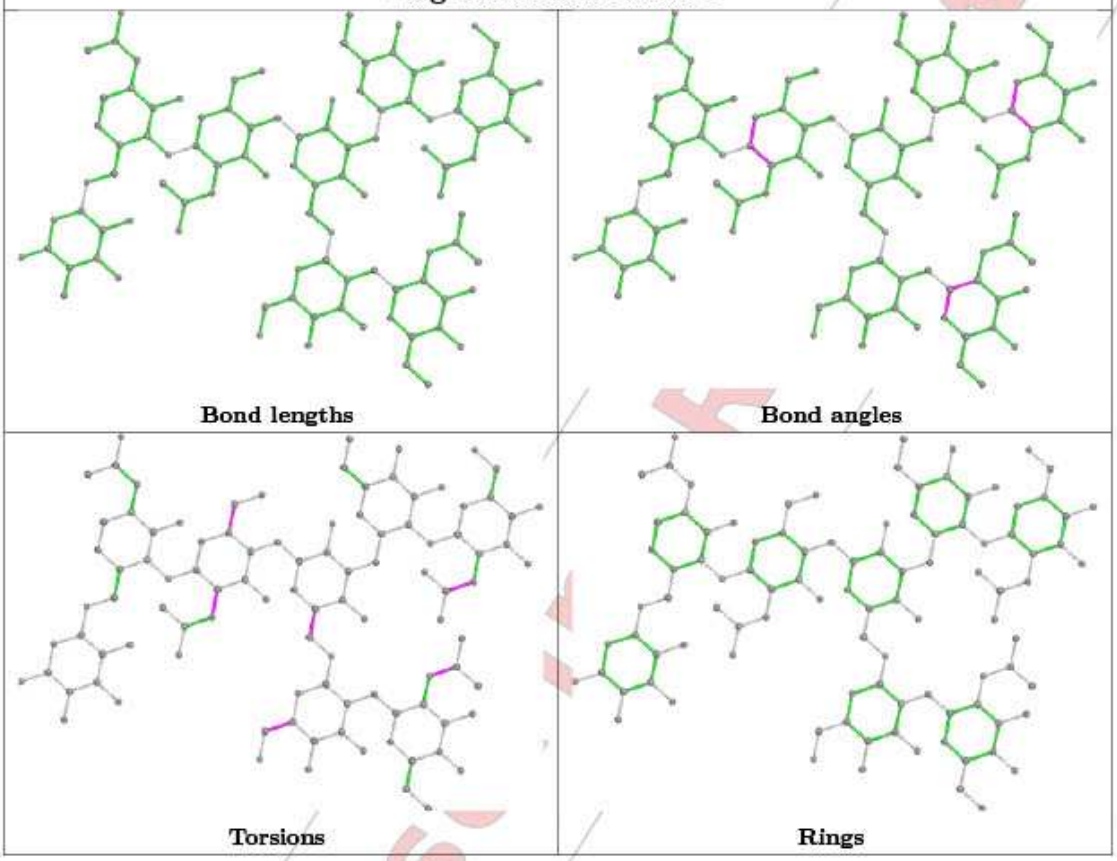
There are no ring outliers.

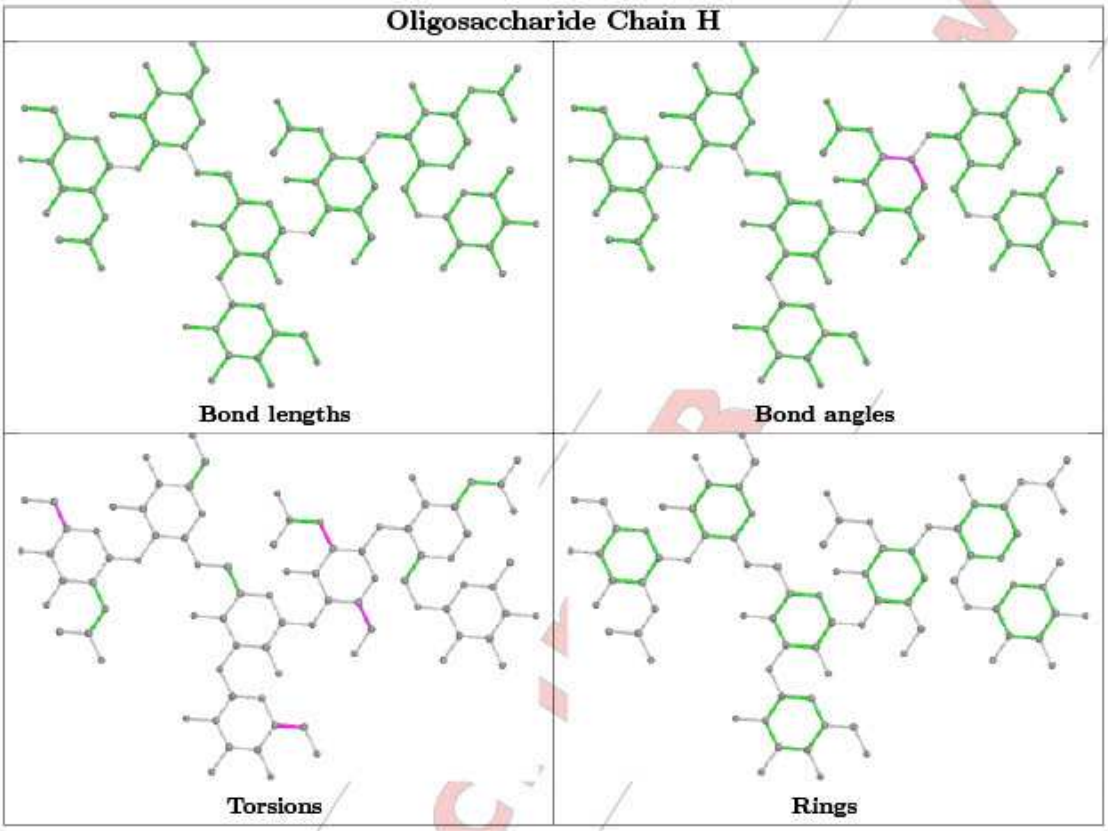
3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1	NAG	1	0
3	G	7	NAG	1	0
4	H	2	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain G





5.6 Ligand geometry

Of 12 ligands modelled in this entry, 3 are monoatomic - leaving 9 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	GOL	B	501	-	5,5,5	0.10	0	5,5,5	0.36	0
6	PGE	F	902	-	9,9,9	0.25	0	8,8,8	0.11	0
6	PGE	D	901	-	9,9,9	0.21	0	8,8,8	0.11	0
6	PGE	D	903	-	9,9,9	0.22	0	8,8,8	0.17	0
5	GOL	D	904	-	5,5,5	0.10	0	5,5,5	0.28	0
5	GOL	B	502	-	5,5,5	0.10	0	5,5,5	0.29	0
6	PGE	D	902	-	9,9,9	0.20	0	8,8,8	0.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	A	501	-	5,5,5	0.09	0	5,5,5	0.30	0
6	PGE	F	901	-	6,6,9	0.23	0	5,5,8	0.25	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	B	501	-	-	2/4/4/4	-
6	PGE	F	902	-	-	4/7/7/7	-
6	PGE	D	901	-	-	6/7/7/7	-
6	PGE	D	903	-	-	5/7/7/7	-
5	GOL	D	904	-	-	0/4/4/4	-
5	GOL	B	502	-	-	4/4/4/4	-
6	PGE	D	902	-	-	3/7/7/7	-
5	GOL	A	501	-	-	0/4/4/4	-
6	PGE	F	901	-	-	1/4/4/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	501	GOL	C1-C2-C3-O3
5	B	502	GOL	C1-C2-C3-O3
5	B	502	GOL	O2-C2-C3-O3
6	D	901	PGE	O1-C1-C2-O2
6	F	902	PGE	O1-C1-C2-O2
6	D	901	PGE	O2-C3-C4-O3
6	D	903	PGE	O3-C5-C6-O4
6	D	903	PGE	O2-C3-C4-O3
5	B	501	GOL	O2-C2-C3-O3
6	F	902	PGE	O2-C3-C4-O3
6	D	903	PGE	O1-C1-C2-O2
6	F	902	PGE	O3-C5-C6-O4
6	D	902	PGE	O3-C5-C6-O4
6	D	901	PGE	C6-C5-O3-C4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	D	903	PGE	C4-C3-O2-C2
6	D	903	PGE	C3-C4-O3-C5
5	B	502	GOL	O1-C1-C2-O2
6	F	902	PGE	C6-C5-O3-C4
6	F	901	PGE	O1-C1-C2-O2
6	D	901	PGE	O3-C5-C6-O4
6	D	901	PGE	C1-C2-O2-C3
6	D	902	PGE	C1-C2-O2-C3
6	D	901	PGE	C3-C4-O3-C5
5	B	502	GOL	O1-C1-C2-C3
6	D	902	PGE	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	901	PGE	4	0
5	A	501	GOL	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	208/227 (91%)	0.13	1 (0%)	91	75	61, 85, 125, 152	0
1	B	207/227 (91%)	0.28	3 (1%)	75	49	59, 86, 146, 175	0
1	C	206/227 (90%)	0.41	9 (4%)	34	13	94, 119, 160, 183	0
2	D	788/816 (96%)	0.22	7 (0%)	84	63	57, 91, 120, 153	0
2	E	447/816 (54%)	0.28	28 (6%)	20	6	83, 134, 159, 166	0
2	F	769/816 (94%)	0.18	20 (2%)	56	27	84, 107, 140, 157	0
All	All	2625/3129 (83%)	0.23	68 (2%)	56	27	57, 104, 149, 183	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	277	TRP	4.9
2	F	346	VAL	4.8
2	E	435	LEU	4.7
2	D	826	LEU	4.3
2	E	330	GLY	4.1
2	F	341	ILE	3.8
1	C	348	VAL	3.5
2	F	345	GLY	3.5
2	E	403	LEU	3.5
2	E	398	ILE	3.4
2	F	284	PHE	3.4
2	F	319	ALA	3.3
2	F	286	PHE	3.3
2	E	458	ILE	3.2
2	E	329	THR	3.0
2	D	722	LEU	2.9
2	E	826	LEU	2.9
2	E	788	GLN	2.9
2	E	787	ALA	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	527	LEU	2.8
2	F	377	SER	2.8
1	C	367	CYS	2.8
2	F	347	ALA	2.8
2	E	407	ILE	2.7
1	C	368	LEU	2.7
2	E	486	LEU	2.7
2	F	71	PHE	2.6
2	F	366	VAL	2.6
2	E	733	LEU	2.6
2	E	268	TRP	2.6
2	E	410	GLN	2.6
2	F	287	PHE	2.5
2	D	731	TRP	2.5
1	C	260	THR	2.5
2	E	531	MET	2.5
2	D	145	ASN	2.4
2	F	388	LYS	2.4
2	E	606	VAL	2.4
2	E	434	ASN	2.4
2	E	681	LEU	2.4
2	D	685	ALA	2.4
2	E	481	LEU	2.4
2	E	669	LEU	2.4
2	F	375	THR	2.3
1	B	297	ASN	2.3
1	B	299	THR	2.3
2	E	272	GLN	2.3
2	D	732	ARG	2.3
1	C	306	LEU	2.3
2	F	293	LYS	2.2
2	E	786	VAL	2.2
2	E	803	VAL	2.2
2	E	494	LEU	2.2
1	A	288	LYS	2.2
2	F	343	ARG	2.2
2	E	608	GLY	2.2
2	F	380	LEU	2.2
2	F	383	LEU	2.1
2	F	342	ASP	2.1
2	E	425	LEU	2.1
1	C	405	PHE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	497	ILE	2.1
1	B	271	PRO	2.1
1	C	436	TYR	2.1
1	C	439	LYS	2.1
2	F	282	ILE	2.1
2	E	731	TRP	2.0
2	F	86	GLN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

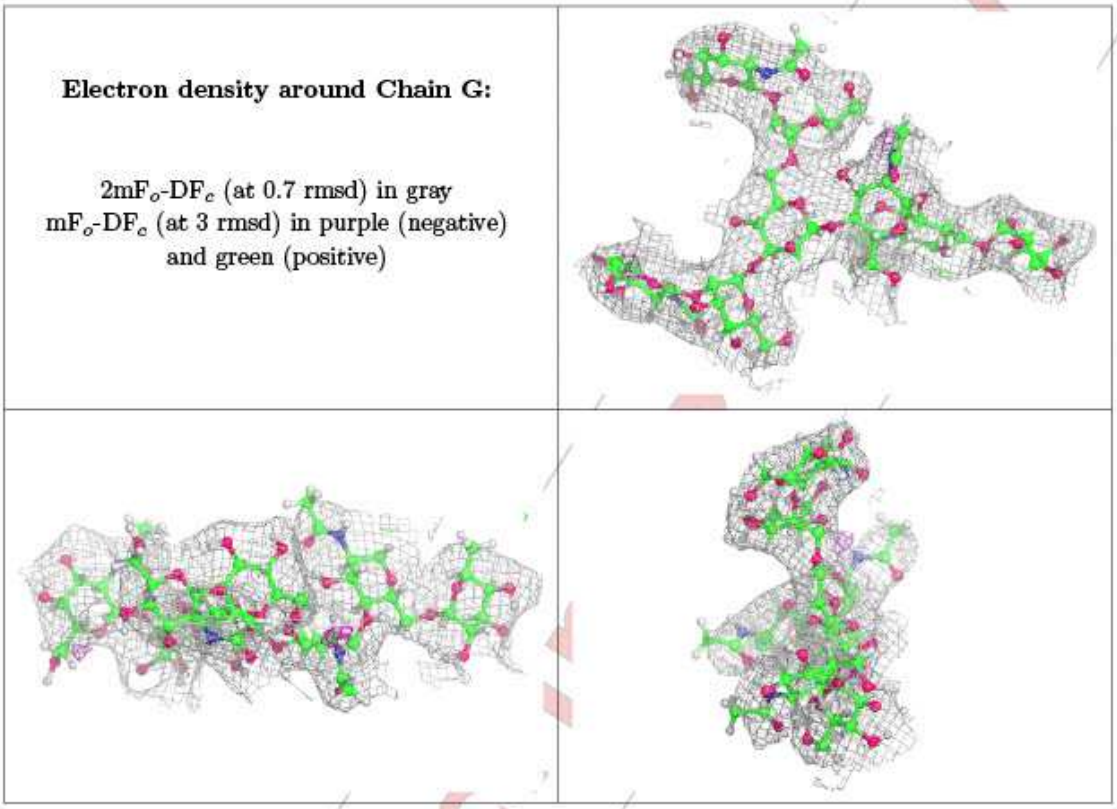
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MAN	H	6	11/12	0.82	0.19	69,123,129,130	3
4	NAG	H	5	14/15	0.86	0.27	69,144,147,153	3
4	NAG	H	2	14/15	0.88	0.22	69,128,134,138	2
4	MAN	H	4	11/12	0.90	0.19	69,124,128,137	2
3	NAG	G	7	14/15	0.91	0.27	69,91,95,97	3
4	NAG	H	1	14/15	0.91	0.19	69,118,127,127	1
4	BMA	H	3	11/12	0.92	0.15	69,123,127,129	1
3	NAG	G	5	14/15	0.92	0.30	69,88,92,95	3
3	MAN	G	6	11/12	0.93	0.19	69,85,88,88	2
4	FUC	H	7	10/11	0.93	0.15	69,106,109,110	3
3	NAG	G	1	14/15	0.95	0.21	63,72,90,96	1
3	NAG	G	2	14/15	0.96	0.27	65,71,76,80	2
3	BMA	G	3	11/12	0.97	0.20	69,71,73,75	2
3	FUC	G	8	10/11	0.97	0.21	69,74,74,77	3
3	MAN	G	4	11/12	0.97	0.24	69,71,74,81	2

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

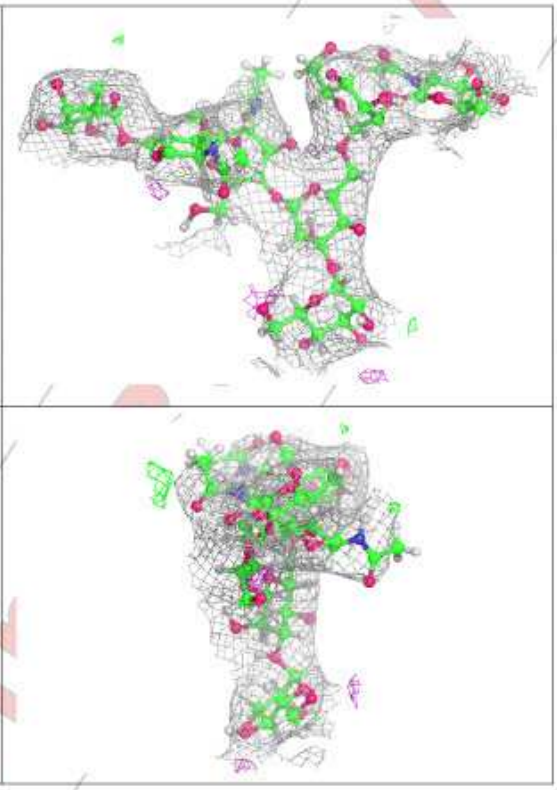
Electron density around Chain G:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around Chain H:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	PGE	D	903	10/10	0.51	1.23	69,142,148,150	1
6	PGE	F	902	10/10	0.65	0.34	69,125,132,133	1
6	PGE	F	901	7/10	0.67	0.52	69,104,108,109	1
5	GOL	D	904	6/6	0.76	0.37	69,105,108,110	2
5	GOL	A	501	6/6	0.77	0.33	69,96,100,100	2
5	GOL	B	502	6/6	0.79	0.31	69,112,116,116	2
7	CA	D	905	1/1	0.80	0.18	135,135,135,135	0
6	PGE	D	901	10/10	0.82	0.48	69,105,110,111	1
5	GOL	B	501	6/6	0.85	0.36	69,104,107,109	2
6	PGE	D	902	10/10	0.85	0.52	69,116,118,119	1
7	CA	F	903	1/1	0.90	0.13	123,123,123,123	0
7	CA	E	901	1/1	0.96	0.19	124,124,124,124	0

6.5 Other polymers

There are no such residues in this entry.

Glossary of Crystallographic Terms

Acentric reflections..... Any reflection that does not have centrosymmetry. These reflections are much more abundant than centric reflections, and are therefore more important to consider when assessing cumulative intensity distributions for twinning analyses, for example.

Anisotropy ..... A crystal is said to be anisotropic when it possesses different diffraction limits in different directions, which results in a more oval-shaped diffraction pattern (rather than a circle).

Asymmetric unit..... The smallest region within a crystal which can generate the full unit cell of the crystal, upon application of the space group symmetry operations.

***B* factor** ..... Describes the displacement of an atom relative to its mean position within an atomic model (higher *B* factors indicate greater displacement). Derived as:

$$B = 8\pi^2\langle u^2\rangle$$

where u is the isotropic displacement of the atom.

***CC*$_{\mathbf{half}}$**..... A correlation coefficient calculated from two randomly-generated halves of a crystallographic set of reflections. The (unmerged) reflections are divided into two equal parts, and a correlation is calculated between the intensity estimates of these two groups, which ranges from a value of 0 (no correlation) to 1 (perfect correlation). The CC_{half} value therefore indicates how well one half of the data predicts the other half. CC_{half} can be calculated for different resolution bins within the data, and is useful for determining the high-resolution cut-off which should be applied (typically at a value of 0.3).

Centric reflections..... Any reflection that does possess centrosymmetry (also called inversion, i.e., an atom at coordinate vector **x** is identical to an atom with coordinate vector $-\mathbf{x}$).

Contour level ..... Usually stated for representation of an electron density map. Level of density is represented in terms of standard deviations (σ) above the mean density; therefore, a map contoured at greater values of σ show areas of the map where the density is stronger (greater signal to noise).

Cryoprotectant..... A substance added to the mother liquor prior to crystal harvesting, which prevents ice formation upon exposure of the crystal to liquid nitrogen. Common examples are glycerol and polyethylene glycol.

Enthalpy, *H* ..... In the context of crystallisation thermodynamics, ΔH describes the heat transfer during crystal formation, i.e., formation of contacts between symmetry-related molecules within the crystal lattice.

Entropy, *S*..... In the context of crystallisation thermodynamics, ΔS describes the change in order during crystal formation, and has a negative value when order is created. The formation of protein molecules within an ordered crystal lattice has inherent negative entropy; this can be overcome by release of ordered solvent molecules around the protein surface upon crystal formation. Replacement of bulky side chains on the protein surface with smaller amino acids, such as alanine, is a strategy for improving chances of successful crystallisation, and is termed surface entropy reduction.

***I*/σ(*I*)**..... Signal-to-noise ratio.

Matthews coefficient..... Crystal volume per unit of protein molecular weight, therefore a smaller value indicates a more tightly-packed crystal lattice.

Molecular replacement ..... A method for estimating the (unknown) phases of a set of structure factors, enabling structure solution. Phases from an existing model (which has a similar structure to the unknown protein structure) are used as initial estimates to phase the unknown structure.

Multiplicity..... Average number of independent measurements for each unique reflection.

NCS... ..... Non-crystallographic symmetry. Arises when more than one copy of a structural unit is present within the asymmetric unit of a crystal, e.g., when a protein crystallises as a dimer.

Refinement ..... Iterative adjustment of parameters defining a crystallographic model, in order to optimise the agreement/fit between the model and observed (raw) data.

$R_{\mathbf{free}}$ ..... Calculated in the same method as R_{work}, but only uses a small percentage of reflections (usually 5%) which aren't used in refinement and therefore aren't biased by the model. R_{free} should be similar to R_{work} for a model that has not been over-refined.

$R_{\mathbf{meas}}$..... Multiplicity-independent *R* factor. Derived from R_{merge}, in which each reflection is corrected by a factor of $\sqrt{\frac{n_{hkl}}{n_{hkl}-1}}$, which removes the dependence on multiplicity. Derived as:

$$\frac{\sum_{hkl}\sqrt{\frac{n_{hkl}}{n_{hkl}-1}}\sum_{i=1}^{n}|I_i(hkl)-\bar{I}(hkl)|}{\sum_{hkl}\sum_{i=1}^{n}I_i(hkl)}$$

where n_{hkl} describes the multiplicity of reflection *hkl*.

$R_{\mathbf{merge}}$..... Merged *R* factor. Indicates the precision of the multiple measurements for each unique reflection. Derived as:

$$\frac{\sum_{hkl}\sum_{i=1}^{n}|I_i(hkl)-\bar{I}(hkl)|}{\sum_{hkl}\sum_{i=1}^{n}I_i(hkl)}$$

where $I_i(hkl)$ is the intensity of reflection *hkl*, $\bar{I}(hkl)$ is the average intensity of the unique reflection, summed over *i* distinct measurements. The value of R_{merge} is affected by multiplicity, in that its value increases with more measurements of the unique reflection (despite the increased precision).

***R*pim** ..... Precision-indicating merged *R* factor. Derived from R_{merge} and accounts for the improvement in precision which is actually present with an increased number of measurements for a unique reflection, by correcting each reflection by a factor of $\sqrt{\frac{1}{n_{hkl}-1}}$. R_{pim} therefore decreases with increasing multiplicity, and is derived as:

$$\frac{\sum_{hkl}\sqrt{\frac{1}{n_{hkl}-1}}\sum_{i=1}^{n}|I_i(hkl)-\bar{I}(hkl)|}{\sum_{hkl}\sum_{i=1}^{n}I_i(hkl)}$$

where n_{hkl} describes the multiplicity of reflection *hkl*.

***R*work** ..... Indicates the fit between raw crystallographic data and the current atomic model. Derived as:

$$R=\frac{\sum_{hkl}|F_{obs}-F_{calc}|}{\sum_{hkl}F_{obs}}$$

where F_{obs} represents observed structure factor amplitudes (calculated from observed reflection intensities) and F_{calc} represents structure factor amplitudes calculated from the current atomic model. A smaller *R* factor therefore indicates a better fit between the raw data and the model.

RSRZ..... Real Space R-value normalised against residue type and resolution. Provides a measure for how well a particular residue fits into the electron density map.

Polder map ..... A type of omit map which is calculated to validate the presence of a particular section of a crystallographic model. Omit maps are typically calculated by omission of these atoms in calculation of the electron density map; if present, the omission of these atoms should result in a positive difference density peak. Observation of this difference density can be masked by bulk solvent modelling, however. Calculation of a polder map accounts for this by preventing bulk solvent entering the omitted region of the model.

Space group ..... A group of symmetry operators describing the arrangement of molecules within a crystal lattice.

Structure factor..... A mathematical description (i.e., a complex wave function) of an observed reflection within a diffraction pattern. Each atom in the unit cell contributes to the scattering of any given reflection, and so a structure factor can be described as the summation of scattered waves from each atom, which interfere constructively and destructively to produce a complex wave. The scattering of an X-ray photon by a single atom is defined by three factors: amplitude, direction and phase. A structure factor describing a particular reflection *hkl* can be defined as:

$$F_{hkl} = \sum_{j=1}^{n} f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

where structure factor F_{hkl} is calculated from n atoms in the unit cell, starting from atom j. The contribution of an atom j to the scattering of reflection hkl is determined by the amplitude f_j, position of the atom within the unit cell (expressed by fractional coordinates hx_j, ky_j and lz_j) and phase α, which is defined implicitly within the exponent $2\pi i(hx_j + ky_j + lz_j)$. i is a complex number used in the definition of a structure factor as a complex vector.

TLS ..... Translation-libration-screw. A refinement protocol in which the positional adjustment of a group of atoms (e.g., a domain within a protein) is parametrised by translation, libration ("rocking" motion) and screw motions.

Twin fraction..... The fractional volume of crystal occupied by the smaller of twin domains present within a crystal.

Twin operator ..... Symmetry operator which relates the orientation of distinct twin domains within a crystal.

Twinning..... The presence of two or more domains within a crystal, containing identically-arranged protein molecules but arranged in distinct orientations (described by the twin operator). Irradiation of such a crystal results in a diffraction pattern comprising contributions from both domains, which may overlap and therefore influence the reflection intensities observed.

Unit cell..... The smallest unit within a crystal which describes the complete space group symmetry repeating across the entire crystal.

Vapour diffusion..... A method used for protein crystallisation, in which a protein is gradually forced into a supersaturated state, by diffusion of water from a protein droplet to a reservoir containing various precipitants (in an enclosed system).

Wilson *B* factor ..... Overall *B*-factor of a crystallographic data set, obtained from a Wilson plot (average reflection intensity vs $\frac{sin^2\theta}{\lambda^2}$, where θ represents the scattering angle and λ represents the X-ray wavelength). Estimates the average *B* factor /degree of order within the crystal.

Bibliography

[1] Marshall JS, Warrington R, Watson W, Kim HL. An introduction to immunology and immunopathology. *Allergy Asthma Clin Immunol*. 2018;14:49.

[2] Turvey SE, Broide DH. Innate immunity. *J Allergy Clin Immunol*. 2010;125:S24-32.

[3] Bonilla FA, Oettgen HC. Adaptive immunity. *J Allergy Clin Immunol*. 2010;125:S33-40.

[4] Laabei M, Ermert D. Catch me if you can: *Streptococcus pyogenes* complement evasion strategies. *J Innate Immun*. 2019;11:3-12.

[5] Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, et al. Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. *Clin Microbiol Rev*. 2014;27:264-301.

[6] von Pawel-Rammingen U, Johansson BP, Björck L. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J*. 2002;21:1607-15.

[7] Collin M, Olsén A. EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J*. 2001;20:3046-55.

[8] Sjögren J, Struwe WB, Cosgrave EF, Rudd PM, Stervander M, Allhorn M, et al. EndoS2 is a unique and conserved enzyme of serotype M49 group A *Streptococcus* that hydrolyses *N*-linked glycans on IgG and α1-acid glycoprotein. *Biochem J*. 2013;455:107-18.

[9] James LK. B cells defined by immunoglobulin isotypes. *Clin Exp Immunol*. 2022;210:230-9.

[10] Schroeder HW, Cavacini L. Structure and function of immunoglobulins. *J Allergy Clin Immunol*. 2010;125:S41-52.

[11] de Taeye SW, Rispens T, Vidarsson G. The ligands for human IgG and their effector functions. *Antibodies (Basel)*. 2019;8:30.

[12] Chi X, Li Y, Qiu X. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. *Immunology*. 2020;160:233-47.

[13] Roth DB. V(D)J recombination: mechanism, errors, and fidelity. *Microbiol Spectr*. 2014;2.

[14] Methot SP, Di Noia JM. Molecular mechanisms of somatic hypermutation and class switch recombination. *Adv Immunol*. 2017;133:37-87.

[15] Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem*. 2007;76:1-22.

[16] Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol*. 2012;12:517-31.

[17] Sharp TH, Boyle AL, Diebolder CA, Kros A, Koster AJ, Gros P. Insights into IgM-mediated complement activation based on *in situ* structures of IgM-C1-C4b. *Proc Natl Acad Sci U S A*. 2019;116:11900-5.

[18] Keyt BA, Baliga R, Sinclair AM, Carroll SF, Peterson MS. Structure, function, and therapeutic use of IgM antibodies. *Antibodies (Basel)*. 2020;9.

[19] Gutzeit C, Chen K, Cerutti A. The enigmatic function of IgD: some answers at last. *Eur J Immunol*. 2018;48:1101-13.

[20] Sutton BJ, Davies AM, Bax HJ, Karagiannis SN. IgE antibodies: from structure to function and clinical translation. *Antibodies (Basel)*. 2019;8.

[21] de Sousa-Pereira P, Woof JM. IgA: structure, function, and developability. *Antibodies (Basel)*. 2019;8.

[22] Sterlin D, Fadlallah J, Adams O, Fieschi C, Parizot C, Dorgham K, et al. Human IgA binds a diverse array of commensal bacteria. *J Exp Med*. 2020;217.

[23] Bunker JJ, Erickson SA, Flynn TM, Henry C, Koval JC, Meisel M, et al. Natural polyreactive IgA antibodies coat the intestinal microbiota. *Science*. 2017;358.

[24] Vidarsson G, Dekkers G, Rispens T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol*. 2014;5:520.

[25] Hjelholt A, Christiansen G, Sørensen US, Birkelund S. IgG subclass profiles in normal human sera of antibodies specific to five kinds of microbial antigens. *Pathog Dis*. 2013;67:206-13.

[26] Ferrante A, Beard LJ, Feldman RG. IgG subclass distribution of antibodies to bacterial and viral antigens. *Pediatr Infect Dis J*. 1990;9:S16-24.

[27] Harris LJ, Skaletsky E, McPherson A. Crystallographic structure of an intact IgG1 monoclonal antibody. *J Mol Biol*. 1998;275:861-72.

[28] Spiteri VA, Goodall M, Doutch J, Rambo RP, Gor J, Perkins SJ. Solution structures of human myeloma IgG3 antibody reveal extended Fab and Fc regions relative to the other IgG subclasses. *J Biol Chem*. 2021;297:100995.

[29] Scharf O, Golding H, King LR, Eller N, Frazier D, Golding B, et al. Immunoglobulin G3 from polyclonal human immunodeficiency virus (HIV) immune globulin is more potent than other subclasses in neutralizing HIV type 1. *J Virol*. 2001;75:6558-65.

[30] Chu TH, Crowley AR, Backes I, Chang C, Tay M, Broge T, et al. Hinge length contributes to the phagocytic activity of HIV-specific IgG1 and IgG3 antibodies. *PLoS Pathog*. 2020;16:e1008083.

[31] Giuntini S, Granoff DM, Beernink PT, Ihle O, Bratlie D, Michaelsen TE. Human IgG1, IgG3, and IgG3 hinge-truncated mutants show different protection capabilities against *meningococci* depending on the target antigen and epitope specificity. *Clin Vaccine Immunol*. 2016;23:698-706.

[32] Norderhaug L, Brekke OH, Bremnes B, Sandin R, Aase A, Michaelsen TE, et al. Chimeric mouse human IgG3 antibodies with an IgG4-like hinge region induce complement-mediated lysis more efficiently than IgG3 with normal hinge. *Eur J Immunol*. 1991;21:2379-84.

[33] Redpath S, Michaelsen T, Sandlie I, Clark MR. Activation of complement by human IgG1 and human IgG3 antibodies against the human leucocyte antigen CD52. *Immunology*. 1998;93:595-600.

[34] Abendstein L, Dijkstra DJ, Tjokrodirijo RTN, van Veelen PA, Trouw LA, Hensbergen PJ, et al. Complement is activated by elevated IgG3 hexameric platforms and deposits C4b onto distinct antibody domains. *Nat Commun*. 2023;14:4027.

[35] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7:539.

[36] Matsumiya S, Yamaguchi Y, Saito J, Nagano M, Sasakawa H, Otaki S, et al. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J Mol Biol*. 2007;368:767-79.

[37] Teplyakov A, Zhao Y, Malia TJ, Obmolova G, Gilliland GL. IgG2 Fc structure and the dynamic features of the IgG CH2-CH3 interface. *Mol Immunol*. 2013;56:131-9.

[38] Davies AM, Rispens T, Ooijevaar-de Heer P, Gould HJ, Jefferis R, Aalberse RC, et al. Structural determinants of unique properties of human IgG4-Fc. *J Mol Biol*. 2014;426:630-44.

[39] Wypych J, Li M, Guo A, Zhang Z, Martinez T, Allen MJ, et al. Human IgG2 antibodies display disulfide-mediated structural isoforms. *J Biol Chem*. 2008;283:16194-205.

[40] White AL, Chan HT, French RR, Willoughby J, Mockridge CI, Roghanian A, et al. Conformation of the human immunoglobulin G2 hinge imparts superagonistic properties to immunostimulatory anticancer antibodies. *Cancer Cell*. 2015;27:138-48.

[41] Liu X, Zhao Y, Shi H, Zhang Y, Yin X, Liu M, et al. Human immunoglobulin G hinge regulates agonistic anti-CD40 immunostimulatory and antitumour activities through biophysical flexibility. *Nat Commun*. 2019;10:4206.

[42] Orr CM, Fisher H, Yu X, Chan CH, Gao Y, Duriez PJ, et al. Hinge disulfides in human IgG2 CD40 antibodies modulate receptor signaling by regulation of conformation and flexibility. *Sci Immunol*. 2022;7:eabm3723.

[43] Rispens T, Huijbers MG. The unique properties of IgG4 and its roles in health and disease. *Nat Rev Immunol*. 2023:1-16.

[44] Labrijn AF, Rispens T, Meesters J, Rose RJ, den Bleker TH, Loverix S, et al. Species-specific determinants in the IgG CH3 domain enable Fab-arm exchange by affecting the noncovalent CH3-CH3 interaction strength. *J Immunol*. 2011;187:3238-46.

[45] Rispens T, Davies AM, Ooijevaar-de Heer P, Absalah S, Bende O, Sutton BJ, et al. Dynamics of inter-heavy chain interactions in human immunoglobulin G (IgG) subclasses studied by kinetic Fab arm exchange. *J Biol Chem*. 2014;289:6098-109.

[46] Lu LL, Suscovich TJ, Fortune SM, Alter G. Beyond binding: antibody effector functions in infectious diseases. *Nat Rev Immunol*. 2018;18:46-61.

[47] Lam H, Kesselly A, Stegalkina S, Kleanthous H, Yethon JA. Antibodies to PhnD inhibit staphylococcal biofilms. *Infect Immun*. 2014;82:3764-74.

[48] Kaur R, Surendran N, Ochs M, Pichichero ME. Human antibodies to PhtD, PcpA, and Ply reduce adherence to human lung epithelial cells and murine nasopharyngeal colonization by *Streptococcus pneumoniae*. *Infect Immun*. 2014;82:5069-75.

[49] Mabry R, Rani M, Geiger R, Hubbard GB, Carrion R, Brasky K, et al. Passive protection against anthrax by using a high-affinity antitoxin antibody fragment lacking an Fc region. *Infect Immun*. 2005;73:8362-8.

[50] Harvill ET, Osorio M, Loving CL, Lee GM, Kelly VK, Merkel TJ. Anamnestic protective immunity to *Bacillus anthracis* is antibody mediated but independent of complement and Fc receptors. *Infect Immun*. 2008;76:2177-82.

[51] Daëron M. Fc receptor biology. *Annu Rev Immunol*. 1997;15:203-34.

[52] Bruhns P, Iannascoli B, England P, Mancardi DA, Fernandez N, Jorieux S, et al. Specificity and affinity of human Fc gamma receptors and their polymorphic variants for human IgG subclasses. *Blood*. 2009;113:3716-25.

[53] Bruhns P. Properties of mouse and human IgG receptors and their contribution to disease models. *Blood*. 2012;119:5640-9.

[54] de Taeye SW, Bentlage AEH, Mebius MM, Meesters JI, Lissenberg-Thunnissen S, Falck D, et al. FcγR binding and ADCC activity of human IgG allotypes. *Frontiers in Immunology*. 2020;11.

[55] Quast I, Peschke B, Lünemann JD. Regulation of antibody effector functions through IgG Fc *N*-glycosylation. *Cell Mol Life Sci*. 2017;74:837-47.

[56] Pereira NA, Chan KF, Lin PC, Song Z. The "less-is-more" in therapeutic antibodies: afucosylated anti-cancer antibodies with enhanced antibody-dependent cellular cytotoxicity. *MAbs*. 2018;10:693-711.

[57] Kiyoshi M, Caaveiro JM, Kawai T, Tashiro S, Ide T, Asaoka Y, et al. Structural basis for binding of human IgG1 to its high-affinity human receptor FcγRI. *Nat Commun*. 2015;6:6866.

[58] Sondermann P, Kaiser J, Jacob U. Molecular basis for immune complex recognition: a comparison of Fc-receptor structures. *J Mol Biol*. 2001;309:737-49.

[59] Mimoto F, Katada H, Kadono S, Igawa T, Kuramochi T, Muraoka M, et al. Engineered antibody Fc variant with selectively enhanced FcγRIIb binding over both FcγRIIa(R131) and FcγRIIa(H131). *Protein Eng Des Sel*. 2013;26:589-98.

[60] Sakae Y, Satoh T, Yagi H, Yanaka S, Yamaguchi T, Isoda Y, et al. Conformational effects of *N*-glycan core fucosylation of immunoglobulin G Fc region on its interaction with Fcγ receptor IIIa. *Sci Rep*. 2017;7.

[61] Roberts JT, Barb AW. A single amino acid distorts the Fc γ receptor IIIb/CD16b structure upon binding immunoglobulin G1 and reduces affinity relative to CD16a. *J Biol Chem*. 2018;293:19899-908.

[62] Bauer-Smith H, Sudol ASL, Beers SA, Crispin M. Serum immunoglobulin and the threshold of Fc receptor-mediated immune activation. *Biochim Biophys Acta Gen Subj*. 2023:130448.

[63] Smyth MJ, Cretney E, Kelly JM, Westwood JA, Street SE, Yagita H, et al. Activation of NK cell cytotoxicity. *Mol Immunol*. 2005;42:501-10.

[64] Wang W, Erbe AK, Hank JA, Morris ZS, Sondel PM. NK cell-mediated antibody-dependent cellular cytotoxicity in cancer immunotherapy. *Front Immunol*. 2015;6:368.

[65] Nimmerjahn F, Ravetch JV. Divergent immunoglobulin G subclass activity through selective Fc receptor binding. *Science*. 2005;310:1510-2.

[66] Junker F, Gordon J, Qureshi O. Fc gamma receptors and their role in antigen uptake, presentation, and T cell activation. *Front Immunol*. 2020;11:1393.

[67] Shields RL, Namenuk AK, Hong K, Meng YG, Rae J, Briggs J, et al. High resolution mapping of the binding site on human IgG1 for Fc gamma RI, Fc gamma RII, Fc gamma RIII, and FcRn and design of IgG1 variants with improved binding to the Fc gamma R. *J Biol Chem*. 2001;276:6591-604.

[68] Caaveiro JM, Kiyoshi M, Tsumoto K. Structural analysis of Fc/FcγR complexes: a blueprint for antibody design. *Immunol Rev*. 2015;268:201-21.

[69] Canfield SM, Morrison SL. The binding affinity of human IgG for its high affinity Fc receptor is determined by multiple amino acids in the CH2 domain and is modulated by the hinge region. *J Exp Med*. 1991;173:1483-91.

[70] Chappel MS, Isenman DE, Everett M, Xu YY, Dorrington KJ, Klein MH. Identification of the Fc gamma receptor class I binding site in human IgG through the use of recombinant IgG1/IgG2 hybrid and point-mutated antibodies. *Proc Natl Acad Sci U S A*. 1991;88:9036-40.

[71] Armour KL, Clark MR, Hadley AG, Williamson LM. Recombinant human IgG molecules lacking Fc gamma receptor I binding and monocyte triggering activities. *Eur J Immunol*. 1999;29:2613-24.

[72] Brüggemann M, Williams GT, Bindon CI, Clark MR, Walker MR, Jefferis R, et al. Comparison of the effector functions of human immunoglobulins using a matched set of chimeric antibodies. *J Exp Med*. 1987;166:1351-61.

[73] Steplewski Z, Sun LK, Shearman CW, Ghrayeb J, Daddona P, Koprowski H. Biological activity of human-mouse IgG1, IgG2, IgG3, and IgG4 chimeric monoclonal antibodies with antitumor specificity. *Proc Natl Acad Sci U S A*. 1988;85:4852-6.

[74] Sondermann P, Huber R, Oosthuizen V, Jacob U. The 3.2-A crystal structure of the human IgG1 Fc fragment-Fc gammaRIII complex. *Nature*. 2000;406:267-73.

[75] Tao MH, Smith RI, Morrison SL. Structural features of human immunoglobulin G that determine isotype-specific differences in complement activation. *J Exp Med*. 1993;178:661-7.

[76] Schartz ND, Tenner AJ. The good, the bad, and the opportunities of the complement system in neurodegenerative disease. *J Neuroinflammation*. 2020;17:354.

[77] Reis ES, Mastellos DC, Hajishengallis G, Lambris JD. New insights into the immune functions of complement. *Nat Rev Immunol*. 2019;19:503-16.

[78] Bindon CI, Hale G, Brüggemann M, Waldmann H. Human monoclonal IgG isotypes differ in complement activating function at the level of C4 as well as C1q. *J Exp Med*. 1988;168:127-42.

[79] Morgan A, Jones ND, Nesbitt AM, Chaplin L, Bodmer MW, Emtage JS. The N-terminal end of the CH2 domain of chimeric human IgG1 anti-HLA-DR is necessary for C1q, Fc gamma RI and Fc gamma RIII binding. *Immunology*. 1995;86:319-24.

[80] Idusogie EE, Presta LG, Gazzano-Santoro H, Totpal K, Wong PY, Ultsch M, et al. Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc. *J Immunol*. 2000;164:4178-84.

[81] Thommesen JE, Michaelsen TE, Løset G, Sandlie I, Brekke OH. Lysine 322 in the human IgG3 C(H)2 domain is crucial for antibody dependent complement activation. *Mol Immunol*. 2000;37:995-1004.

[82] Ugurlar D, Howes SC, de Kreuk BJ, Koning RI, de Jong RN, Beurskens FJ, et al. Structures of C1-IgG1 provide insights into how danger pattern recognition activates complement. *Science*. 2018;359:794-7.

[83] Diebolder CA, Beurskens FJ, de Jong RN, Koning RI, Strumane K, Lindorfer MA, et al. Complement is activated by IgG hexamers assembled at the cell surface. *Science*. 2014;343:1260-3.

[84] de Jong RN, Beurskens FJ, Verploegen S, Strumane K, van Kampen MD, Voorhorst M, et al. A novel platform for the potentiation of therapeutic antibodies based on antigen-dependent formation of IgG hexamers at the cell surface. *PLoS Biol*. 2016;14:e1002344.

[85] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372:774-97.

[86] So EC, Zhou H, Greenwell A, Burch EE, Ji Y, Mérigeon EY, et al. Complement component C1q is an immunological rheostat that regulates Fc:FcγR interactions. *Immunogenetics*. 2023;75:369-83.

[87] Pyzik M, Kozicky LK, Gandhi AK, Blumberg RS. The therapeutic age of the neonatal Fc receptor. *Nat Rev Immunol*. 2023;23:415-32.

[88] Vidarsson G, Stemerding AM, Stapleton NM, Spliethoff SE, Janssen H, Rebers FE, et al. FcRn: an IgG receptor on phagocytes with a novel role in phagocytosis. *Blood*. 2006;108:3573-9.

[89] Martin WL, West AP, Gan L, Bjorkman PJ. Crystal structure at 2.8 A of an FcRn/heterodimeric Fc complex: mechanism of pH-dependent binding. *Mol Cell*. 2001;7:867-77.

[90] Oganesyan V, Damschroder MM, Cook KE, Li Q, Gao C, Wu H, et al. Structural insights into neonatal Fc receptor-based recycling mechanisms. *J Biol Chem*. 2014;289:7812-24.

[91] Kim JK, Firan M, Radu CG, Kim CH, Ghetie V, Ward ES. Mapping the site on human IgG for binding of the MHC class I-related receptor, FcRn. *Eur J Immunol*. 1999;29:2819-25.

[92] Raghavan M, Bonagura VR, Morrison SL, Bjorkman PJ. Analysis of the pH dependence of the neonatal Fc receptor/immunoglobulin G interaction using antibody and receptor variants. *Biochemistry*. 1995;34:14649-57.

[93] Stapleton NM, Andersen JT, Stemerding AM, Bjarnarson SP, Verheul RC, Gerritsen J, et al. Competition for FcRn-mediated transport gives rise to short half-life of human IgG3 and offers therapeutic potential. *Nat Commun*. 2011;2:599.

[94] Brinkhaus M, Pannecoucke E, van der Kooi EJ, Bentlage AEH, Derksen NIL, Andries J, et al. The Fab region of IgG impairs the internalization pathway of FcRn upon Fc engagement. *Nat Commun*. 2022;13:6073.

[95] Zalevsky J, Chamberlain AK, Horton HM, Karki S, Leung IW, Sproule TJ, et al. Enhanced antibody half-life improves *in vivo* activity. *Nat Biotechnol*. 2010;28:157-9.

[96] Dall'Acqua WF, Kiener PA, Wu H. Properties of human IgG1s engineered for enhanced binding to the neonatal Fc receptor (FcRn). *J Biol Chem*. 2006;281:23514-24.

[97] James LC, Keeble AH, Khan Z, Rhodes DA, Trowsdale J. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proc Natl Acad Sci U S A*. 2007;104:6200-5.

[98] Fletcher AJ, Mallery DL, Watkinson RE, Dickson CF, James LC. Sequential ubiquitination and deubiquitination enzymes synchronize the dual sensor and effector functions of TRIM21. *Proc Natl Acad Sci U S A*. 2015;112:10014-9.

[99] Foss S, Watkinson R, Sandlie I, James LC, Andersen JT. TRIM21: a cytosolic Fc receptor with broad antibody isotype specificity. *Immunol Rev*. 2015;268:328-39.

[100] Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol*. 2003;325:979-89.

[101] Zheng K, Bantog C, Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs*. 2011;3:568-76.

[102] Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, et al. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol*. 2000;37:697-706.

[103] Yamaguchi Y, Nishimura M, Nagano M, Yagi H, Sasakawa H, Uchida K, et al. Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. *Biochim Biophys Acta*. 2006;1760:693-700.

[104] Zheng K, Yarmarkovich M, Bantog C, Bayer R, Patapoff TW. Influence of glycosylation pattern on the molecular properties of monoclonal antibodies. *MAbs*. 2014;6:649-58.

[105] Baruah K, Bowden TA, Krishna BA, Dwek RA, Crispin M, Scanlan CN. Selective deactivation of serum IgG: a general strategy for the enhancement of monoclonal antibody receptor interactions. *J Mol Biol*. 2012;420:1-7.

[106] Wada R, Matsui M, Kawasaki N. Influence of *N*-glycosylation on effector functions and thermal stability of glycoengineered IgG1 monoclonal antibody with homogeneous glycoforms. *MAbs*. 2019;11:350-72.

[107] Le NP, Bowden TA, Struwe WB, Crispin M. Immune recruitment or suppression by glycan engineering of endogenous and therapeutic antibodies. *Biochim Biophys Acta*. 2016;1860:1655-68.

[108] Crispin M, Yu X, Bowden TA. Crystal structure of sialylated IgG Fc: implications for the mechanism of intravenous immunoglobulin therapy. *Proc Natl Acad Sci U S A*. 2013;110:E3544-6.

[109] Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology*. 2015;25:1323-4.

[110] Nimmerjahn F, Vidarsson G, Cragg MS. Effect of posttranslational modifications and subclass on IgG activity: from immunity to immunotherapy. *Nat Immunol*. 2023;24:1244-55.

[111] Baković MP, Selman MH, Hoffmann M, Rudan I, Campbell H, Deelder AM, et al. High-throughput IgG Fc *N*-glycosylation profiling by mass spectrometry of glycopeptides. *J Proteome Res*. 2013;12:821-31.

[112] Dekkers G, Treffers L, Plomp R, Bentlage AEH, de Boer M, Koeleman CAM, et al. Decoding the human immunoglobulin G-glycan repertoire reveals a spectrum of Fc-receptor- and complement-mediated-effector activities. *Front Immunol*. 2017;8:877.

[113] Kiyoshi M, Tsumoto K, Ishii-Watabe A, Caaveiro JMM. Glycosylation of IgG-Fc: a molecular perspective. *Int Immunol*. 2017;29:311-7.

[114] Yu X, Baruah K, Harvey DJ, Vasiljevic S, Alonzi DS, Song BD, et al. Engineering hydrophobic protein-carbohydrate interactions to fine-tune monoclonal antibodies. *J Am Chem Soc*. 2013;135:9723-32.

[115] Frank M, Walker RC, Lanzilotta WN, Prestegard JH, Barb AW. Immunoglobulin G1 Fc domain motions: implications for Fc engineering. *J Mol Biol*. 2014;426:1799-811.

[116] Barb AW, Meng L, Gao Z, Johnson RW, Moremen KW, Prestegard JH. NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation. *Biochemistry*. 2012;51:4618-26.

[117] Barb AW, Prestegard JH. NMR analysis demonstrates immunoglobulin G *N*-glycans are accessible and dynamic. *Nat Chem Biol*. 2011;7:147-53.

[118] Lee HS, Im W. Effects of *N*-glycan composition on structure and dynamics of IgG1 Fc and their implications for antibody engineering. *Sci Rep*. 2017;7:12659.

[119] Harbison AM, Brosnan LP, Fenlon K, Fadda E. Sequence-to-structure dependence of isolated IgG Fc complex biantennary *N*-glycans: a molecular dynamics study. *Glycobiology*. 2019;29:94-103.

[120] Lu J, Chu J, Zou Z, Hamacher NB, Rixon MW, Sun PD. Structure of FcγRI in complex with Fc reveals the importance of glycan recognition for high-affinity IgG binding. *Proc Natl Acad Sci U S A*. 2015;112:833-8.

[121] Subedi GP, Barb AW. The structural role of antibody *N*-glycosylation in receptor interactions. *Structure*. 2015;23:1573-83.

[122] Allhorn M, Olin AI, Nimmerjahn F, Collin M. Human IgG/Fc gamma R interactions are modulated by streptococcal IgG glycan hydrolysis. *PLoS One*. 2008;3:e1413.

[123] Nose M, Wigzell H. Biological significance of carbohydrate chains on monoclonal antibodies. *Proc Natl Acad Sci U S A*. 1983;80:6632-6.

[124] Hayes JM, Frostell A, Cosgrave EF, Struwe WB, Potter O, Davey GP, et al. Fc gamma receptor glycosylation modulates the binding of IgG glycoforms: a requirement for stable antibody interactions. *J Proteome Res*. 2014;13:5471-85.

[125] Buck PM, Kumar S, Singh SK. Consequences of glycan truncation on Fc structural integrity. *MAbs*. 2013;5:904-16.

[126] Subedi GP, Hanson QM, Barb AW. Restricted motion of the conserved immunoglobulin G1 *N*-glycan is essential for efficient FcγRIIIa binding. *Structure*. 2014;22:1478-88.

[127] Barb AW. Intramolecular *N*-glycan/polypeptide interactions observed at multiple *N*-glycan remodeling steps through [(13)C,(15)N]-*N*-acetylglucosamine labeling of immunoglobulin G1. *Biochemistry*. 2015;54:313-22.

[128] Bowden TA, Baruah K, Coles CH, Harvey DJ, Yu X, Song BD, et al. Chemical and structural analysis of an antibody folding intermediate trapped during glycan biosynthesis. *J Am Chem Soc*. 2012;134:17554-63.

[129] Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, et al. Lack of fucose on human IgG1 *N*-linked oligosaccharide improves binding to human Fc gamma RIII and antibody-dependent cellular toxicity. *J Biol Chem*. 2002;277:26733-40.

[130] Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, et al. The absence of fucose but not the presence of galactose or bisecting *N*-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J Biol Chem*. 2003;278:3466-73.

[131] Niwa R, Natsume A, Uehara A, Wakitani M, Iida S, Uchida K, et al. IgG subclass-independent improvement of antibody-dependent cellular cytotoxicity by fucose removal from Asn297-linked oligosaccharides. *J Immunol Methods*. 2005;306:151-60.

[132] Subedi GP, Barb AW. The immunoglobulin G1 *N*-glycan composition affects binding to each low affinity Fc γ receptor. *MAbs*. 2016;8:1512-24.

[133] Falconer DJ, Subedi GP, Marcella AM, Barb AW. Antibody fucosylation lowers FcγRIIIa/CD16a affinity by limiting the conformations sampled by the N162-glycan. *ACS Chem Biol*. 2018;13:2179-89.

[134] Peschke B, Keller CW, Weber P, Quast I, Lünemann JD. Fc-galactosylation of human immunoglobulin gamma isotypes improves C1q binding and enhances complement-dependent cytotoxicity. *Front Immunol*. 2017;8:646.

[135] van Osch TLJ, Nouta J, Derksen NIL, van Mierlo G, van der Schoot CE, Wuhrer M, et al. Fc galactosylation promotes hexamerization of human IgG1, leading to enhanced classical complement activation. *J Immunol*. 2021;207:1545-54.

[136] Wei B, Gao X, Cadang L, Izadi S, Liu P, Zhang HM, et al. Fc galactosylation follows consecutive reaction kinetics and enhances immunoglobulin G hexamerization for complement activation. *MAbs*. 2021;13:1893427.

[137] Anthony RM, Nimmerjahn F, Ashline DJ, Reinhold VN, Paulson JC, Ravetch JV. Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science*. 2008;320:373-6.

[138] Bartsch YC, Rahmöller J, Mertes MMM, Eiglmeier S, Lorenz FKM, Stoehr AD, et al. Sialylated autoantigen-reactive IgG antibodies attenuate disease development in autoimmune mouse models of lupus nephritis and rheumatoid arthritis. *Front Immunol*. 2018;9:1183.

[139] Epp A, Hobusch J, Bartsch YC, Petry J, Lilienthal GM, Koeleman CAM, et al. Sialylation of IgG antibodies inhibits IgG-mediated allergic reactions. *J Allergy Clin Immunol*. 2018;141:399-402.e8.

[140] Washburn N, Schwab I, Ortiz D, Bhatnagar N, Lansing JC, Medeiros A, et al. Controlled tetra-Fc sialylation of IVIg results in a drug candidate with consistent enhanced anti-inflammatory activity. *Proc Natl Acad Sci U S A*. 2015;112:E1297-306.

[141] Crispin M. Therapeutic potential of deglycosylated antibodies. *Proc Natl Acad Sci U S A*. 2013;110:10059-60.

[142] Borrok MJ, Jung ST, Kang TH, Monzingo AF, Georgiou G. Revisiting the role of glycosylation in the structure of human IgG Fc. *ACS Chem Biol*. 2012;7:1596-602.

[143] Harvey RA, Champe PC, Fisher BD. Microbiology. *2nd edition* ed: Lippincott Williams & Williams; 2007.

[144] van Sorge NM, Cole JN, Kuipers K, Henningham A, Aziz RK, Kasirer-Friede A, et al. The classical lancefield antigen of group A *Streptococcus* is a virulence determinant with implications for vaccine design. *Cell Host Microbe*. 2014;15:729-40.

[145] Lancefield RC. The antigenic complex of *Streptococcus haemolyticus*: I. Demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. *J Exp Med*. 1928;47:91-103.

[146] McMillan DJ, Drèze PA, Vu T, Bessen DE, Guglielmini J, Steer AC, et al. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. *Clin Microbiol Infect*. 2013;19:E222-9.

[147] Phillips GN, Flicker PF, Cohen C, Manjula BN, Fischetti VA. Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. *Proc Natl Acad Sci U S A*. 1981;78:4689-93.

[148] Hanski E, Caparon M. Protein F, a fibronectin-binding protein, is an adhesin of the group A *Streptococcus Streptococcus pyogenes*. *Proc Natl Acad Sci U S A*. 1992;89:6172-6.

[149] Ferretti JJ, Stevens DL, Fischetti VA, University of Oklahoma. Health Sciences Center. *Streptococcus pyogenes*: basic biology to clinical manifestations. Oklahoma City (OK): University of Oklahoma Health Sciences Center,; 2016. p. 1 online resource (PDF file (x, 727 pages)).

[150] Brouwer S, Rivera-Hernandez T, Curren BF, Harbison-Price N, De Oliveira DMP, Jespersen MG, et al. Pathogenesis, epidemiology and control of group A *Streptococcus* infection. *Nat Rev Microbiol*. 2023;21:431-47.

[151] Hand R, Snelling T, Carapetis J. Group A *Streptococcus*. *Hunter's Tropical Medicine and Emerging Infectious Diseases*. 2019;2020:429-38.

[152] Musser JM, Beres SB, Zhu L, Olsen RJ, Vuopio J, Hyyryläinen HL, et al. Reduced *in vitro* susceptibility of *Streptococcus pyogenes* to β-lactam antibiotics associated with mutations in the *pbp2x* gene is geographically widespread. *J Clin Microbiol*. 2020;58.

[153] Hayes A, Lacey JA, Morris JM, Davies MR, Tong SYC. Restricted sequence variation in *Streptococcus pyogenes* penicillin binding proteins. *mSphere*. 2020;5.

[154] Dale JB, Walker MJ. Update on group A streptococcal vaccine development. *Curr Opin Infect Dis*. 2020;33:244-50.

[155] Walkinshaw DR, Wright MEE, Mullin AE, Excler JL, Kim JH, Steer AC. The *Streptococcus pyogenes* vaccine landscape. *NPJ Vaccines*. 2023;8:16.

[156] Pastural É, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, et al. Safety and immunogenicity of a 30-valent M protein-based group A streptococcal vaccine in healthy adult volunteers: a randomized, controlled phase I study. *Vaccine*. 2020;38:1384-92.

[157] Sekuloski S, Batzloff MR, Griffin P, Parsonage W, Elliott S, Hartas J, et al. Evaluation of safety and immunogenicity of a group A *Streptococcus* vaccine candidate (MJ8VAX) in a randomized clinical trial. *PLoS One*. 2018;13:e0198658.

[158] Ozberk V, Reynolds S, Huo Y, Calcutt A, Eskandari S, Dooley J, et al. Prime-pull immunization with a bivalent M-protein and Spy-CEP peptide vaccine adjuvanted with CAF®01 liposomes induces both mucosal and peripheral protection from *covR/S* mutant *Streptococcus pyogenes*. *mBio*. 2021;12.

[159] Reynolds S, Pandey M, Dooley J, Calcutt A, Batzloff M, Ozberk V, et al. Preclinical safety and immunogenicity of *Streptococcus pyogenes* (Strep A) peptide vaccines. *Sci Rep*. 2021;11:127.

[160] Reynolds S, Rafeek RAM, Hamlin A, Lepletier A, Pandey M, Ketheesan N, et al. *Streptococcus pyogenes* vaccine candidates do not induce autoimmune responses in a rheumatic heart disease model. *NPJ Vaccines*. 2023;8:9.

[161] Di Benedetto R, Mancini F, Carducci M, Gasperini G, Moriel DG, Saul A, et al. Rational design of a glycoconjugate vaccine against group A *Streptococcus*. *Int J Mol Sci*. 2020;21.

[162] Gao NJ, Uchiyama S, Pill L, Dahesh S, Olson J, Bautista L, et al. Site-specific conjugation of cell wall polyrhamnose to protein SpyAD envisioning a safe universal group A streptococcal vaccine. *Infectious Microbes & Diseases*. 2021;3:87-100.

[163] Loh JMS, Rivera-Hernandez T, McGregor R, Khemlani AHJ, Tay ML, Cork AJ, et al. A multivalent T-antigen-based vaccine for group A *Streptococcus*. *Sci Rep*. 2021;11:4353.

[164] Rivera-Hernandez T, Rhyme MS, Cork AJ, Jones S, Segui-Perez C, Brunner L, et al. Vaccine-induced Th1-type response protects against invasive group A *Streptococcus* infection in the absence of opsonizing antibodies. *mBio*. 2020;11.

[165] Matsumura T, Takahashi Y. The role of myeloid cells in prevention and control of group A streptococcal infections. *Biosafety and Health*. 2020;2:130-4.

[166] Navarini AA, Lang KS, Verschoor A, Recher M, Zinkernagel AS, Nizet V, et al. Innate immune-induced depletion of bone marrow neutrophils aggravates systemic bacterial infections. *Proc Natl Acad Sci U S A*. 2009;106:7107-12.

[167] Hidalgo-Grass C, Mishalian I, Dan-Goor M, Belotserkovsky I, Eran Y, Nizet V, et al. A streptococcal protease that degrades CXC chemokines and impairs bacterial clearance from infected tissues. *EMBO J*. 2006;25:4628-37.

[168] Goldmann O, Rohde M, Chhatwal GS, Medina E. Role of macrophages in host resistance to group A *streptococci*. *Infect Immun*. 2004;72:2956-63.

[169] Mishalian I, Ordan M, Peled A, Maly A, Eichenbaum MB, Ravins M, et al. Recruited macrophages control dissemination of group A *Streptococcus* from infected soft tissues. *J Immunol*. 2011;187:6022-31.

[170] Loof TG, Rohde M, Chhatwal GS, Jung S, Medina E. The contribution of dendritic cells to host defenses against *Streptococcus pyogenes*. *J Infect Dis*. 2007;196:1794-803.

[171] Loof TG, Goldmann O, Medina E. Immune recognition of *Streptococcus pyogenes* by dendritic cells. *Infect Immun*. 2008;76:2785-92.

[172] Harder J, Franchi L, Muñoz-Planillo R, Park JH, Reimer T, Núñez G. Activation of the Nlrp3 inflammasome by *Streptococcus pyogenes* requires streptolysin O and NF-kappa B activation but proceeds independently of TLR signaling and P2X7 receptor. *J Immunol*. 2009;183:5823-9.

[173] Fieber C, Kovarik P. Responses of innate immune cells to group A *Streptococcus*. *Front Cell Infect Microbiol*. 2014;4:140.

[174] Imai T, Matsumura T, Mayer-Lambertz S, Wells CA, Ishikawa E, Butcher SK, et al. Lipoteichoic acid anchor triggers Mincle to drive protective immunity against invasive group A *Streptococcus* infection. *Proc Natl Acad Sci U S A*. 2018;115:E10662-E71.

[175] Matsumura T, Ikebe T, Arikawa K, Hosokawa M, Aiko M, Iguchi A, et al. Sequential sensing by TLR2 and Mincle directs immature myeloid cells to protect against invasive group A streptococcal infection in mice. *Cell Rep*. 2019;27:561-71.e6.

[176] Matsumura T, Ato M, Ikebe T, Ohnishi M, Watanabe H, Kobayashi K. Interferon-γ-producing immature myeloid cells confer protection against severe invasive group A *Streptococcus* infections. *Nat Commun*. 2012;3:678.

[177] Sabharwal H, Michon F, Nelson D, Dong W, Fuchs K, Manjarrez RC, et al. Group A *Streptococcus* (GAS) carbohydrate as an immunogen for protection against GAS infection. *J Infect Dis*. 2006;193:129-35.

[178] O'Connor SP, Darip D, Fraley K, Nelson CM, Kaplan EL, Cleary PP. The human antibody response to streptococcal C5a peptidase. *J Infect Dis*. 1991;163:109-16.

[179] Frost HR, Laho D, Sanderson-Smith ML, Licciardi P, Donath S, Curtis N, et al. Immune cross-opsonization within *emm* clusters following group A *Streptococcus* skin infection: broadening the scope of type-specific immunity. *Clin Infect Dis*. 2017;65:1523-31.

[180] Frost H, Excler JL, Sriskandan S, Fulurija A. Correlates of immunity to group A *Streptococcus*: a pathway to vaccine development. *NPJ Vaccines*. 2023;8:1.

[181] Mortensen R, Nissen TN, Blauenfeldt T, Christensen JP, Andersen P, Dietrich J. Adaptive immunity against *Streptococcus pyogenes* in adults involves increased IFN-γ and IgG3 responses compared with children. *J Immunol*. 2015;195:1657-64.

[182] Ermert D, Weckel A, Agarwal V, Frick IM, Björck L, Blom AM. Binding of complement inhibitor C4b-binding protein to a highly virulent *Streptococcus pyogenes* M1 strain is mediated by protein H and enhances adhesion to and invasion of endothelial cells. *J Biol Chem*. 2013;288:32172-83.

[183] Carlsson F, Berggård K, Stålhammar-Carlemalm M, Lindahl G. Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. *J Exp Med*. 2003;198:1057-68.

[184] Whitnack E, Beachey EH. Inhibition of complement-mediated opsonization and phagocytosis of *Streptococcus pyogenes* by D fragments of fibrinogen and fibrin bound to cell surface M protein. *J Exp Med*. 1985;162:1983-97.

[185] Ly D, Taylor JM, Tsatsaronis JA, Monteleone MM, Skora AS, Donald CA, et al. Plasmin(ogen) acquisition by group A *Streptococcus* protects against C3b-mediated neutrophil killing. *J Innate Immun*. 2014;6:240-50.

[186] Ermert D, Shaughnessy J, Joeris T, Kaplan J, Pang CJ, Kurt-Jones EA, et al. Virulence of group A *Streptococci* is enhanced by human complement inhibitors. *PLoS Pathog*. 2015;11:e1005043.

[187] Hyland KA, Wang B, Cleary PP. Protein F1 and *Streptococcus pyogenes* resistance to phagocytosis. *Infect Immun*. 2007;75:3188-91.

[188] Dale JB, Washburn RG, Marques MB, Wessels MR. Hyaluronate capsule and surface M protein in resistance to opsonization of group A *streptococci*. *Infect Immun*. 1996;64:1495-501.

[189] Caswell CC, Han R, Hovis KM, Ciborowski P, Keene DR, Marconi RT, et al. The Scl1 protein of M6-type group A *Streptococcus* binds the human complement regulatory protein, factor H, and inhibits the alternative pathway of complement. *Mol Microbiol*. 2008;67:584-96.

[190] Pandiripally V, Gregory E, Cue D. Acquisition of regulators of complement activation by *Streptococcus pyogenes* serotype M1. *Infect Immun*. 2002;70:6206-14.

[191] Cleary PP, Prahbu U, Dale JB, Wexler DE, Handley J. Streptococcal C5a peptidase is a highly specific endopeptidase. *Infect Immun*. 1992;60:5219-23.

[192] Lynskey NN, Reglinski M, Calay D, Siggins MK, Mason JC, Botto M, et al. Multi-functional mechanisms of immune evasion by the streptococcal complement inhibitor C5a peptidase. *PLoS Pathog*. 2017;13:e1006493.

[193] Fernie-King BA, Seilly DJ, Willers C, Würzner R, Davies A, Lachmann PJ. Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes. *Immunology*. 2001;103:390-8.

[194] Honda-Ogawa M, Sumitomo T, Mori Y, Hamd DT, Ogawa T, Yamaguchi M, et al. Endopeptidase O contributes to evasion from complement-mediated bacteriolysis via binding to human complement factor C1q. *J Biol Chem*. 2017;292:4244-54.

[195] Brouwer S, Cork AJ, Ong CY, Barnett TC, West NP, McIver KS, et al. Endopeptidase PepO regulates the SpeB cysteine protease and is essential for the virulence of invasive M1T1 *Streptococcus pyogenes*. *J Bacteriol*. 2018;200.

[196] Aghababa H, Ting YT, Pilapitiya D, Loh JMS, Young PG, Proft T. Complement evasion factor (CEF), a novel immune evasion factor of *Streptococcus pyogenes*. *Virulence*. 2022;13:225-40.

[197] Sierig G, Cywes C, Wessels MR, Ashbaugh CD. Cytotoxic effects of streptolysin O and streptolysin S enhance the virulence of poorly encapsulated group A *streptococci*. *Infect Immun*. 2003;71:446-55.

[198] Uchiyama S, Döhrmann S, Timmer AM, Dixit N, Ghochani M, Bhandari T, et al. Streptolysin O rapidly impairs neutrophil oxidative burst and antibacterial responses to group A *Streptococcus*. *Front Immunol*. 2015;6:581.

[199] Goldmann O, Sastalla I, Wos-Oxley M, Rohde M, Medina E. *Streptococcus pyogenes* induces oncosis in macrophages through the activation of an inflammatory programmed cell death pathway. *Cell Microbiol*. 2009;11:138-55.

[200] Edwards RJ, Taylor GW, Ferguson M, Murray S, Rendell N, Wrigley A, et al. Specific C-terminal cleavage and inactivation of interleukin-8 by invasive disease isolates of *Streptococcus pyogenes*. *J Infect Dis*. 2005;192:783-90.

[201] Tanaka M, Kinoshita-Daitoku R, Kiga K, Sanada T, Zhu B, Okano T, et al. Group A *Streptococcus* establishes pharynx infection by degrading the deoxyribonucleic acid of neutrophil extracellular traps. *Sci Rep*. 2020;10:3251.

[202] Buchanan JT, Simpson AJ, Aziz RK, Liu GY, Kristian SA, Kotb M, et al. DNase expression allows the pathogen group A *Streptococcus* to escape killing in neutrophil extracellular traps. *Curr Biol*. 2006;16:396-400.

[203] Uchiyama S, Andreoni F, Schuepbach RA, Nizet V, Zinkernagel AS. DNase Sda1 allows invasive M1T1 group A *Streptococcus* to prevent TLR9-dependent recognition. *PLoS Pathog*. 2012;8:e1002736.

[204] Doolin T, Amir HM, Duong L, Rosenzweig R, Urban LA, Bosch M, et al. Mammalian histones facilitate antimicrobial synergy by disrupting the bacterial proton gradient and chromosome organization. *Nat Commun*. 2020;11:3888.

[205] Nitzsche R, Köhler J, Kreikemeyer B, Oehmcke-Hecht S. *Streptococcus pyogenes* escapes killing from extracellular histones through plasminogen binding and activation by streptokinase. *J Innate Immun*. 2016;8:589-600.

[206] Sun H, Ringdahl U, Homeister JW, Fay WP, Engleberg NC, Yang AY, et al. Plasminogen is a critical host pathogenicity factor for group A streptococcal infection. *Science*. 2004;305:1283-6.

[207] Hollands A, Gonzalez D, Leire E, Donald C, Gallo RL, Sanderson-Smith M, et al. A bacterial pathogen co-opts host plasmin to resist killing by cathelicidin antimicrobial peptides. *J Biol Chem*. 2012;287:40891-7.

[208] Nyberg P, Rasmussen M, Björck L. Alpha2-macroglobulin-proteinase complexes protect *Streptococcus pyogenes* from killing by the antimicrobial peptide LL-37. *J Biol Chem*. 2004;279:52820-3.

[209] Kristian SA, Datta V, Weidenmaier C, Kansal R, Fedtke I, Peschel A, et al. D-alanylation of teichoic acids promotes group A *Streptococcus* antimicrobial peptide resistance, neutrophil survival, and epithelial cell invasion. *J Bacteriol*. 2005;187:6719-25.

[210] Wierzbicki IH, Campeau A, Dehaini D, Holay M, Wei X, Greene T, et al. Group A streptococcal S protein utilizes red blood cells as immune camouflage and is a critical determinant for immune evasion. *Cell Rep*. 2019;29:2979-89.e15.

[211] Pleass RJ, Areschoug T, Lindahl G, Woof JM. Streptococcal IgA-binding proteins bind in the Calpha 2-Calpha 3 interdomain region and inhibit binding of IgA to human CD89. *J Biol Chem*. 2001;276:8197-204.

[212] Stenberg L, O'Toole PW, Mestecky J, Lindahl G. Molecular characterization of protein Sir, a streptococcal cell surface protein that binds both immunoglobulin A and immunoglobulin G. *J Biol Chem*. 1994;269:13458-64.

[213] Fagan PK, Reinscheid D, Gottschalk B, Chhatwal GS. Identification and characterization of a novel secreted immunoglobulin binding protein from group A *streptococcus*. *Infect Immun*. 2001;69:4851-7.

[214] Nelson DC, Garbe J, Collin M. Cysteine proteinase SpeB from *Streptococcus pyogenes* - a potent modifier of immunologically important host and bacterial proteins. *Biol Chem*. 2011;392:1077-88.

[215] Sumitomo T, Nakata M, Higashino M, Terao Y, Kawabata S. Group A streptococcal cysteine protease cleaves epithelial junctions and contributes to bacterial translocation. *J Biol Chem*. 2013;288:13317-24.

[216] Kapur V, Topouzis S, Majesky MW, Li LL, Hamrick MR, Hamill RJ, et al. A conserved *Streptococcus pyogenes* extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. *Microb Pathog*. 1993;15:327-46.

[217] Barnett TC, Liebl D, Seymour LM, Gillen CM, Lim JY, Larock CN, et al. The globally disseminated M1T1 clone of group A *Streptococcus* evades autophagy for intracellular replication. *Cell Host Microbe*. 2013;14:675-82.

[218] Egesten A, Olin AI, Linge HM, Yadav M, Mörgelin M, Karlsson A, et al. SpeB of *Streptococcus pyogenes* differentially modulates antibacterial and receptor activating properties of human chemokines. *PLoS One*. 2009;4:e4769.

[219] Schmidtchen A, Frick IM, Andersson E, Tapper H, Björck L. Proteinases of common pathogenic bacteria degrade and inactivate the antibacterial peptide LL-37. *Mol Microbiol*. 2002;46:157-68.

[220] Kapur V, Majesky MW, Li LL, Black RA, Musser JM. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A*. 1993;90:7676-80.

[221] Macleod T, Ainscough JS, Hesse C, Konzok S, Braun A, Buhl AL, et al. The proinflammatory cytokine IL-36γ is a global discriminator of harmless microbes and invasive pathogens within epithelial tissues. *Cell Rep*. 2020;33:108515.

[222] Richter J, Brouwer S, Schroder K, Walker MJ. Inflammasome activation and IL-1β signalling in group A *Streptococcus* disease. *Cell Microbiol*. 2021;23:e13373.

[223] Kuo CF, Lin YS, Chuang WJ, Wu JJ, Tsao N. Degradation of complement 3 by streptococcal pyrogenic exotoxin B inhibits complement activation and neutrophil opsonophagocytosis. *Infect Immun*. 2008;76:1163-9.

[224] Terao Y, Mori Y, Yamaguchi M, Shimizu Y, Ooe K, Hamada S, et al. Group A streptococcal cysteine protease degrades C3 (C3b) and contributes to evasion of innate immunity. *J Biol Chem*. 2008;283:6253-60.

[225] Honda-Ogawa M, Ogawa T, Terao Y, Sumitomo T, Nakata M, Ikebe K, et al. Cysteine proteinase from *Streptococcus pyogenes* enables evasion of innate immunity via degradation of complement factors. *J Biol Chem*. 2013;288:15854-64.

[226] von Pawel-Rammingen U, Björck L. IdeS and SpeB: immunoglobulin-degrading cysteine proteinases of *Streptococcus pyogenes*. *Curr Opin Microbiol*. 2003;6:50-5.

[227] Collin M, Olsén A. Effect of SpeB and EndoS from *Streptococcus pyogenes* on human immunoglobulins. *Infect Immun*. 2001;69:7187-9.

[228] Persson H, Vindebro R, von Pawel-Rammingen U. The streptococcal cysteine protease SpeB is not a natural immunoglobulin-cleaving enzyme. *Infect Immun*. 2013;81:2236-41.

[229] Blöchl C, Holzner C, Luciano M, Bauer R, Horejs-Hoeck J, Eckhard U, et al. Proteolytic profiling of streptococcal pyrogenic exotoxin B (SpeB) by complementary HPLC-MS approaches. *Int J Mol Sci*. 2021;23.

[230] Ryan MH, Petrone D, Nemeth JF, Barnathan E, Björck L, Jordan RE. Proteolysis of purified IgGs by human and bacterial enzymes *in vitro* and the detection of specific proteolytic fragments of endogenous IgG in rheumatoid synovial fluid. *Mol Immunol*. 2008;45:1837-46.

[231] Brezski RJ, Vafa O, Petrone D, Tam SH, Powers G, Ryan MH, et al. Tumor-associated and microbial proteases compromise host IgG effector functions by a single cleavage proximal to the hinge. *Proc Natl Acad Sci U S A*. 2009;106:17864-9.

[232] Lei B, DeLeo FR, Hoe NP, Graham MR, Mackie SM, Cole RL, et al. Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. *Nat Med*. 2001;7:1298-305.

[233] Brezski RJ, Jordan RE. Cleavage of IgGs by proteases associated with invasive diseases: an evasion tactic against host immunity? *MAbs*. 2010;2:212-20.

[234] Lei B, Mackie S, Lukomski S, Musser JM. Identification and immunogenicity of group A *Streptococcus* culture supernatant proteins. *Infect Immun*. 2000;68:6807-18.

[235] Kawabata S, Tamura Y, Murakami J, Terao Y, Nakagawa I, Hamada S. A novel, anchorless streptococcal surface protein that binds to human immunoglobulins. *Biochem Biophys Res Commun*. 2002;296:1329-33.

[236] Lei B, DeLeo FR, Reid SD, Voyich JM, Magoun L, Liu M, et al. Opsonophagocytosis-inhibiting mac protein of group A *Streptococcus*: identification and characteristics of two genetic complexes. *Infect Immun*. 2002;70:6880-90.

[237] Söderberg JJ, Engström P, von Pawel-Rammingen U. The intrinsic immunoglobulin G endopeptidase activity of streptococcal Mac-2 proteins implies a unique role for the enzymatically impaired Mac-2 protein of M28 serotype strains. *Infect Immun*. 2008;76:2183-8.

[238] Deveuve Q, Lajoie L, Barrault B, Thibault G. The proteolytic cleavage of therapeutic monoclonal antibody hinge region: more than a matter of subclass. *Front Immunol*. 2020;11:168.

[239] Vindebro R, Spoerry C, von Pawel-Rammingen U. Rapid IgG heavy chain cleavage by the streptococcal IgG endopeptidase IdeS is mediated by IdeS monomers and is not due to enzyme dimerization. *FEBS Lett*. 2013;587:1818-22.

[240] Jordan RE, Fan X, Salazar G, Zhang N, An Z. Proteinase-nicked IgGs: an unanticipated target for tumor immunotherapy. *Oncoimmunology*. 2018;7:e1480300.

[241] Fan X, Brezski RJ, Fa M, Deng H, Oberholtzer A, Gonzalez A, et al. A single proteolytic cleavage within the lower hinge of trastuzumab reduces immune effector function and *in vivo* efficacy. *Breast Cancer Res*. 2012;14:R116.

[242] Hsiao HC, Fan X, Jordan RE, Zhang N, An Z. Proteolytic single hinge cleavage of pertuzumab impairs its Fc effector function and antitumor activity *in vitro* and *in vivo*. *Breast Cancer Res*. 2018;20:43.

[243] Agniswamy J, Lei B, Musser JM, Sun PD. Insight of host immune evasion mediated by two variants of group A *Streptococcus* Mac protein. *J Biol Chem*. 2004;279:52789-96.

[244] Vincents B, von Pawel-Rammingen U, Björck L, Abrahamson M. Enzymatic characterization of the streptococcal endopeptidase, IdeS, reveals that it is a cysteine protease with strict specificity for IgG cleavage due to exosite binding. *Biochemistry*. 2004;43:15540-9.

[245] Vincents B, Vindebro R, Abrahamson M, von Pawel-Rammingen U. The human protease inhibitor cystatin C is an activating cofactor for the streptococcal cysteine protease IdeS. *Chem Biol*. 2008;15:960-8.

[246] Agniswamy J, Nagiec MJ, Liu M, Schuck P, Musser JM, Sun PD. Crystal structure of group A *Streptococcus* Mac-1: insight into dimer-mediated specificity for recognition of human IgG. *Structure*. 2006;14:225-35.

[247] Wenig K, Chatwell L, von Pawel-Rammingen U, Björck L, Huber R, Sondermann P. Structure of the streptococcal endopeptidase IdeS, a cysteine proteinase with strict specificity for IgG. *Proc Natl Acad Sci U S A*. 2004;101:17371-6.

[248] Lei B, Liu M, Meyers EG, Manning HM, Nagiec MJ, Musser JM. Histidine and aspartic acid residues important for immunoglobulin G endopeptidase activity of the group A *Streptococcus* opsonophagocytosis-inhibiting Mac protein. *Infect Immun*. 2003;71:2881-4.

[249] Kalińska M, Kantyka T, Greenbaum DC, Larsen KS, Władyka B, Jabaiah A, et al. Substrate specificity of *Staphylococcus aureus* cysteine proteases--Staphopains A, B and C. *Biochimie*. 2012;94:318-27.

[250] Potempa J, Sroka A, Imamura T, Travis J. Gingipains, the major cysteine proteinases and virulence factors of *Porphyromonas gingivalis*: structure, function and assembly of multidomain protein complexes. *Curr Protein Pept Sci*. 2003;4:397-407.

[251] Herwald H, Collin M, Müller-Esterl W, Björck L. Streptococcal cysteine proteinase releases kinins: a virulence mechanism. *J Exp Med*. 1996;184:665-73.

[252] Stockbauer KE, Magoun L, Liu M, Burns EH, Gubba S, Renish S, et al. A natural variant of the cysteine protease virulence factor of group A *Streptococcus* with an arginine-glycine-aspartic acid (RGD) motif preferentially binds human integrins alphavbeta3 and alphaIIbbeta3. *Proc Natl Acad Sci U S A*. 1999;96:242-7.

[253] Collin M, Svensson MD, Sjöholm AG, Jensenius JC, Sjöbring U, Olsén A. EndoS and SpeB from *Streptococcus pyogenes* inhibit immunoglobulin-mediated opsonophagocytosis. *Infect Immun*. 2002;70:6646-51.

[254] Naegeli A, Bratanis E, Karlsson C, Shannon O, Kalluru R, Linder A, et al. *Streptococcus pyogenes* evades adaptive immunity through specific IgG glycan hydrolysis. *J Exp Med*. 2019;216:1615-29.

[255] Goodfellow JJ, Baruah K, Yamamoto K, Bonomelli C, Krishna B, Harvey DJ, et al. An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. *J Am Chem Soc*. 2012;134:8030-3.

[256] Sjögren J, Cosgrave EF, Allhorn M, Nordgren M, Björk S, Olsson F, et al. EndoS and EndoS2 hydrolyze Fc-glycans on therapeutic antibodies with different glycoform selectivity and can be used for rapid quantification of high-mannose glycans. *Glycobiology*. 2015;25:1053-63.

[257] Trastoy B, Lomino JV, Pierce BG, Carter LG, Günther S, Giddens JP, et al. Crystal structure of *Streptococcus pyogenes* EndoS, an immunomodulatory endoglycosidase specific for human IgG antibodies. *Proc Natl Acad Sci U S A*. 2014;111:6714-9.

[258] Trastoy B, Klontz E, Orwenyo J, Marina A, Wang LX, Sundberg EJ, et al. Structural basis for the recognition of complex-type *N*-glycans by endoglycosidase S. *Nat Commun*. 2018;9:1874.

[259] Klontz EH, Trastoy B, Deredge D, Fields JK, Li C, Orwenyo J, et al. Molecular basis of broad spectrum *N*-glycan specificity and processing of therapeutic IgG monoclonal antibodies by endoglycosidase S2. *ACS Cent Sci*. 2019;5:524-38.

[260] Allhorn M, Olsén A, Collin M. EndoS from *Streptococcus pyogenes* is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity. *BMC Microbiol*. 2008;8:3.

[261] Dixon EV, Claridge JK, Harvey DJ, Baruah K, Yu X, Vesiljevic S, et al. Fragments of bacterial endoglycosidase S and immunoglobulin G reveal subdomains of each that contribute to deglycosylation. *J Biol Chem*. 2014;289:13876-89.

[262] Huang E, Maldonado AQ, Kjellman C, Jordan SC. Imlifidase for the treatment of anti-HLA antibody-mediated processes in kidney transplantation. *Am J Transplant*. 2022;22:691-7.

[263] Hart A, Smith JM, Skeans MA, Gustafson SK, Wilk AR, Castro S, et al. OPTN/SRTR 2017 annual data report: kidney. *American Journal of Transplantation*. 2019;19:19-123.

[264] Jordan SC, Legendre C, Desai NM, Lorant T, Bengtsson M, Lonze BE, et al. Imlifidase desensitization in crossmatch-positive, highly sensitized kidney transplant recipients: results of an international phase 2 trial (Highdes). *Transplantation*. 2021;105:1808-17.

[265] Kjellman C, Maldonado AQ, Sjöholm K, Lonze BE, Montgomery RA, Runström A, et al. Outcomes at 3 years posttransplant in imlifidase-desensitized kidney transplant patients. *Am J Transplant*. 2021;21:3907-18.

[266] Lin J, Boon L, Bockermann R, Robertson AK, Kjellman C, Anderson CC. Desensitization using imlifidase and EndoS enables chimerism induction in allosensitized recipient mice. *Am J Transplant*. 2020;20:2356-65.

[267] Segelmark M, Björck L. Streptococcal enzymes as precision tools against pathogenic IgG autoantibodies in small vessel vasculitis. *Front Immunol*. 2019;10:2165.

[268] Nandakumar KS, Johansson BP, Björck L, Holmdahl R. Blocking of experimental arthritis by cleavage of IgG antibodies *in vivo*. *Arthritis Rheum*. 2007;56:3253-60.

[269] Johansson BP, Shannon O, Björck L. IdeS: a bacterial proteolytic enzyme with therapeutic potential. *PLoS One*. 2008;3:e1692.

[270] Mihai S, Albert H, Ludwig RJ, Iwata H, Björck L, Collin M, et al. *In vivo* enzymatic modulation of IgG antibodies prevents immune complex-dependent skin injury. *Exp Dermatol*. 2017;26:691-6.

[271] Yang R, Otten MA, Hellmark T, Collin M, Björck L, Zhao MH, et al. Successful treatment of experimental glomerulonephritis with IdeS and EndoS, IgG-degrading streptococcal enzymes. *Nephrol Dial Transplant*. 2010;25:2479-86.

[272] Soveri I, Mölne J, Uhlin F, Nilsson T, Kjellman C, Sonesson E, et al. The IgG-degrading enzyme of *Streptococcus pyogenes* causes rapid clearance of anti-glomerular basement membrane antibodies in patients with refractory anti-glomerular basement membrane disease. *Kidney Int*. 2019;96:1234-8.

[273] Kizlik-Masson C, Deveuve Q, Zhou Y, Vayne C, Thibault G, McKenzie SE, et al. Cleavage of anti-PF4/heparin IgG by a bacterial protease and potential benefit in heparin-induced thrombocytopenia. *Blood*. 2019;133:2427-35.

[274] Takahashi R, Yuki N. Streptococcal IdeS: therapeutic potential for Guillain-Barré syndrome. *Sci Rep*. 2015;5:10809.

[275] Uhlin F, Szpirt W, Kronbichler A, Bruchfeld A, Soveri I, Rostaing L, et al. Endopeptidase cleavage of anti-glomerular basement membrane antibodies *in vivo* in severe kidney disease: an open-label phase 2a study. *J Am Soc Nephrol*. 2022;33:829-38.

[276] Lynch DR, Stringham EN, Zhang B, Balbin-Cuesta G, Curtis BR, Palumbo JS, et al. Anchoring IgG-degrading enzymes to the surface of platelets selectively neutralizes antiplatelet antibodies. *Blood Adv*. 2022;6:4645-56.

[277] Akesson P, Moritz L, Truedsson M, Christensson B, von Pawel-Rammingen U. IdeS, a highly specific immunoglobulin G (IgG)-cleaving enzyme from *Streptococcus pyogenes*, is inhibited by specific IgG antibodies generated during infection. *Infect Immun*. 2006;74:497-503.

[278] Lood C, Allhorn M, Lood R, Gullstrand B, Olin AI, Rönnblom L, et al. IgG glycan hydrolysis by endoglycosidase S diminishes the proinflammatory properties of immune complexes from patients with systemic lupus erythematosus: a possible new treatment? *Arthritis Rheum*. 2012;64:2698-706.

[279] Albert H, Collin M, Dudziak D, Ravetch JV, Nimmerjahn F. *In vivo* enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. *Proc Natl Acad Sci U S A*. 2008;105:15005-9.

[280] Nandakumar KS, Collin M, Happonen KE, Lundström SL, Croxford AM, Xu B, et al. Streptococcal endo-β-*N*-acetylglucosaminidase suppresses antibody-mediated inflammation *in vivo*. *Front Immunol*. 2018;9:1623.

[281] Nandakumar KS, Collin M, Olsén A, Nimmerjahn F, Blom AM, Ravetch JV, et al. Endoglycosidase treatment abrogates IgG arthritogenicity: importance of IgG glycosylation in arthritis. *Eur J Immunol*. 2007;37:2973-82.

[282] van Timmeren MM, van der Veen BS, Stegeman CA, Petersen AH, Hellmark T, Collin M, et al. IgG glycan hydrolysis attenuates ANCA-mediated glomerulonephritis. *J Am Soc Nephrol*. 2010;21:1103-14.

[283] Collin M, Shannon O, Björck L. IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions. *Proc Natl Acad Sci U S A*. 2008;105:4265-70.

[284] Benkhoucha M, Molnarfi N, Santiago-Raber ML, Weber MS, Merkler D, Collin M, et al. IgG glycan hydrolysis by EndoS inhibits experimental autoimmune encephalomyelitis. *J Neuroinflammation*. 2012;9:209.

[285] Hirose M, Vafia K, Kalies K, Groth S, Westermann J, Zillikens D, et al. Enzymatic autoantibody glycan hydrolysis alleviates autoimmunity against type VII collagen. *J Autoimmun*. 2012;39:304-14.

[286] Miesbach W, Meijer K, Coppens M, Kampmann P, Klamroth R, Schutgens R, et al. Gene therapy with adeno-associated virus vector 5-human factor IX in adults with hemophilia B. *Blood*. 2018;131:1022-31.

[287] Mingozzi F, High KA. Overcoming the host immune response to adeno-associated virus gene delivery vectors: the race between clearance, tolerance, neutralization, and escape. *Annu Rev Virol*. 2017;4:511-34.

[288] Leborgne C, Barbon E, Alexander JM, Hanby H, Delignat S, Cohen DM, et al. IgG-cleaving endopeptidase enables *in vivo* gene therapy in the presence of anti-AAV neutralizing antibodies. *Nat Med*. 2020;26:1096-101.

[289] Kelton JG, Singer J, Rodger C, Gauldie J, Horsewood P, Dent P. The concentration of IgG in the serum is a major determinant of Fc-dependent reticuloendothelial function. *Blood*. 1985;66:490-5.

[290] Preithner S, Elm S, Lippold S, Locher M, Wolf A, da Silva AJ, et al. High concentrations of therapeutic IgG1 antibodies are needed to compensate for inhibition of antibody-dependent cellular cytotoxicity by excess endogenous immunoglobulin G. *Molecular immunology*. 2006;43:1183-93.

[291] Järnum S, Bockermann R, Runström A, Winstedt L, Kjellman C. The bacterial enzyme IdeS cleaves the IgG-type of B cell receptor (BCR), abolishes BCR-mediated cell signaling, and inhibits memory B cell activation. *J Immunol*. 2015;195:5592-601.

[292] Järnum S, Runström A, Bockermann R, Winstedt L, Crispin M, Kjellman C. Enzymatic inactivation of endogenous IgG by IdeS enhances therapeutic antibody efficacy. *Mol Cancer Ther*. 2017;16:1887-97.

[293] Lorant T, Bengtsson M, Eich T, Eriksson BM, Winstedt L, Järnum S, et al. Safety, immunogenicity, pharmacokinetics, and efficacy of degradation of anti-HLA antibodies by IdeS (imlifidase) in chronic kidney disease patients. *Am J Transplant*. 2018;18:2752-62.

[294] Golay J, Da Roit F, Bologna L, Ferrara C, Leusen JH, Rambaldi A, et al. Glycoengineered CD20 antibody obinutuzumab activates neutrophils and mediates phagocytosis through CD16B more efficiently than rituximab. *Blood*. 2013;122:3482-91.

[295] Kaneko Y, Nimmerjahn F, Ravetch JV. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science*. 2006;313:670-3.

[296] Fiedler W, Stoeger H, Perotti A, Gastl G, Weidmann J, Dietrich B, et al. Phase I study of TrasGEX, a glyco-optimised anti-HER2 monoclonal antibody, in patients with HER2-positive solid tumours. *ESMO Open*. 2018;3:e000381.

[297] Fiedler W, DeDosso S, Cresta S, Weidmann J, Tessari A, Salzberg M, et al. A phase I study of PankoMab-GEX, a humanised glyco-optimised monoclonal antibody to a novel tumour-specific MUC1 glycopeptide epitope in patients with advanced carcinomas. *Eur J Cancer*. 2016;63:55-63.

[298] Wang LX, Tong X, Li C, Giddens JP, Li T. Glycoengineering of antibodies for modulating functions. *Annu Rev Biochem*. 2019;88:433-59.

[299] Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *J Am Chem Soc*. 2012;134:12308-18.

[300] Tong X, Li T, Li C, Wang LX. Generation and comparative kinetic analysis of new glycosynthase mutants from *Streptococcus pyogenes* endoglycosidases for antibody glycoengineering. *Biochemistry*. 2018;57:5239-46.

[301] Li T, Tong X, Yang Q, Giddens JP, Wang LX. Glycosynthase mutants of endoglycosidase S2 show potent transglycosylation activity and remarkably relaxed substrate specificity for antibody glycosylation remodeling. *J Biol Chem*. 2016;291:16508-18.

[302] Shivatare SS, Huang LY, Zeng YF, Liao JY, You TH, Wang SY, et al. Development of glycosynthases with broad glycan specificity for the efficient glyco-remodeling of antibodies. *Chem Commun (Camb)*. 2018;54:6161-4.

[303] Trastoy B, Du JJ, García-Alija M, Li C, Klontz EH, Wang LX, et al. Sculpting therapeutic monoclonal antibody *N*-glycans using endoglycosidases. *Curr Opin Struct Biol*. 2022;72:248-59.

[304] Du JJ, Klontz EH, Guerin ME, Trastoy B, Sundberg EJ. Structural insights into the mechanisms and specificities of IgG-active endoglycosidases. *Glycobiology*. 2020;30:268-79.

[305] Synstad B, Gåseidnes S, Van Aalten DM, Vriend G, Nielsen JE, Eijsink VG. Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase. *Eur J Biochem*. 2004;271:253-62.

[306] Tong X, Li T, Orwenyo J, Toonstra C, Wang LX. One-pot enzymatic glycan remodeling of a therapeutic monoclonal antibody by endoglycosidase S (Endo-S) from *Streptococcus pyogenes*. *Bioorg Med Chem*. 2018;26:1347-55.

[307] Iwamoto M, Yamaguchi T, Sekiguchi Y, Oishi S, Shiiki T, Soma M, et al. Pharmacokinetic and pharmacodynamic profiles of glyco-modified atrial natriuretic peptide derivatives synthesized using chemo-enzymatic synthesis approaches. *Bioconjug Chem*. 2018;29:2829-37.

[308] Li C, Wang LX. Chemoenzymatic methods for the synthesis of glycoproteins. *Chem Rev*. 2018;118:8359-413.

[309] Shi W, Li W, Zhang J, Li T, Song Y, Zeng Y, et al. One-step synthesis of site-specific antibody-drug conjugates by reprograming IgG glycoengineering with LacNAc-based substrates. *Acta Pharm Sin B*. 2022;12:2417-28.

[310] Zhang X, Ou C, Liu H, Prabhu SK, Li C, Yang Q, et al. General and robust chemoenzymatic method for glycan-mediated site-specific labeling and conjugation of antibodies: facile synthesis of homogeneous antibody-drug conjugates. *ACS Chem Biol*. 2021;16:2502-14.

[311] Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42:W320-4.

[312] Merchant AM, Zhu Z, Yuan JQ, Goddard A, Adams CW, Presta LG, et al. An efficient route to human bispecific IgG. *Nat Biotechnol*. 1998;16:677-81.

[313] Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 1999;112:531-52.

[314] Gorrec F. The MORPHEUS protein crystallization screen. *J Appl Crystallogr*. 2009;42:1035-42.

[315] Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol*. 2018;74:85-97.

[316] Winter G, Lobley CM, Prince SM. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr*. 2013;69:1260-73.

[317] Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010;66:125-32.

[318] Gildea RJ, Beilsten-Edmands J, Axford D, Horrell S, Aller P, Sandy J, et al. xia2.multiplex: a multi-crystal data-analysis pipeline. *Acta Crystallogr D Struct Biol*. 2022;78:752-69.

[319] Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, et al. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol*. 2018;74:68-84.

[320] Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*. 2010;66:22-5.

[321] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010;66:486-501.

[322] Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*. 2011;67:355-67.

[323] Emsley P, Crispin M. Structural analysis of glycoproteins: building *N*-linked glycans with Coot. *Acta Crystallogr D Struct Biol*. 2018;74:256-63.

[324] Agirre J, Iglesias-Fernández J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. Privateer: software for the conformational validation of carbohydrate structures. *Nat Struct Mol Biol*. 2015;22:833-4.

[325] Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010;66:12-21.

[326] Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci*. 2018;27:293-315.

[327] Gore S, Sanz García E, Hendrickx PMS, Gutmanas A, Westbrook JD, Yang H, et al. Validation of structures in the Protein Data Bank. *Structure*. 2017;25:1916-27.

[328] Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci*. 2021;30:70-82.

[329] Deller MC, Kong L, Rupp B. Protein stability: a crystallographer's perspective. *Acta Crystallogr F Struct Biol Commun*. 2016;72:72-95.

[330] Mas PJ, Hart DJ. ESPRIT: a method for defining soluble expression constructs in poorly understood gene sequences. *Methods Mol Biol*. 2017;1586:45-63.

[331] Yumerefendi H, Tarendeau F, Mas PJ, Hart DJ. ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J Struct Biol*. 2010;172:66-74.

[332] Reich S, Puckey LH, Cheetham CL, Harris R, Ali AA, Bhattacharyya U, et al. Combinatorial domain hunting: an effective approach for the identification of soluble protein domains adaptable to high-throughput applications. *Protein Sci*. 2006;15:2356-65.

[333] Littler E. Combinatorial domain hunting: solving problems in protein expression. *Drug Discov Today*. 2010;15:461-7.

[334] Englander SW. Hydrogen exchange and mass spectrometry: a historical perspective. *J Am Soc Mass Spectrom*. 2006;17:1481-9.

[335] Konermann L, Pan J, Liu YH. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev*. 2011;40:1224-34.

[336] Zheng J, Strutzenberg T, Pascal BD, Griffin PR. Protein dynamics and conformational changes explored by hydrogen/deuterium exchange mass spectrometry. *Curr Opin Struct Biol*. 2019;58:305-13.

[337] James EI, Murphree TA, Vorauer C, Engen JR, Guttman M. Advances in hydrogen/deuterium exchange mass spectrometry and the pursuit of challenging biological systems. *Chem Rev*. 2022;122:7562-623.

[338] Spraggon G, Pantazatos D, Klock HE, Wilson IA, Woods VL, Lesley SA. On the use of DXMS to produce more crystallizable proteins: structures of the *T. maritima* proteins TM0160 and TM1171. *Protein Sci*. 2004;13:3187-99.

[339] Pantazatos D, Kim JS, Klock HE, Stevens RC, Wilson IA, Lesley SA, et al. Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. *Proc Natl Acad Sci U S A*. 2004;101:751-6.

[340] Fowler ML, McPhail JA, Jenkins ML, Masson GR, Rutaganira FU, Shokat KM, et al. Using hydrogen deuterium exchange mass spectrometry to engineer optimized constructs for crystallization of protein complexes: case study of PI4KIIIβ with Rab11. *Protein Sci*. 2016;25:826-39.

[341] Bai Y, Auperin TC, Tong L. The use of *in situ* proteolysis in the crystallization of murine CstF-77. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2007;63:135-8.

[342] Mandel CR, Gebauer D, Zhang H, Tong L. A serendipitous discovery that *in situ* proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p). *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2006;62:1041-5.

[343] Dong A, Xu X, Edwards AM, Chang C, Chruszcz M, Cuff M, et al. *In situ* proteolysis for protein crystallization and structure determination. *Nat Methods*. 2007;4:1019-21.

[344] Wernimont A, Edwards A. *In situ* proteolysis to generate crystals for structure determination: an update. *PLoS One*. 2009;4:e5094.

[345] Bou-Nader C, Zhang J. Rational engineering enables co-crystallization and structural determination of the HIV-1 matrix-tRNA complex. *STAR Protoc*. 2022;3:101056.

[346] Gheyi T, Rodgers L, Romero R, Sauder JM, Burley SK. Mass spectrometry guided *in situ* proteolysis to obtain crystals for X-ray structure determination. *J Am Soc Mass Spectrom*. 2010;21:1795-801.

[347] Little DJ, Whitney JC, Robinson H, Yip P, Nitz M, Howell PL. Combining *in situ* proteolysis and mass spectrometry to crystallize *Escherichia coli* PgaB. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2012;68:842-5.

[348] Lei L, Egli M. *In situ* proteolysis for crystallization of membrane bound cytochrome P450 17A1 and 17A2 proteins from zebrafish. *Curr Protoc Protein Sci*. 2016;84:29.16.1-29.16.9.

[349] Zhang Y, Wang S, Jia Z. *In situ* proteolysis condition-induced crystallization of the XcpVWX complex in different lattices. *Int J Mol Sci*. 2020;21.

[350] Civril F, Hopfner KP. Crystallization of mouse RIG-I ATPase domain: *in situ* proteolysis. *Methods Mol Biol*. 2014;1169:27-35.

[351] Remaut H, Tang C, Henderson NS, Pinkner JS, Wang T, Hultgren SJ, et al. Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell*. 2008;133:640-52.

[352] Hassell AM, An G, Bledsoe RK, Bynum JM, Carter HL, Deng SJ, et al. Crystallization of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2007;63:72-9.

[353] Derewenda ZS. Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D Biol Crystallogr*. 2010;66:604-15.

[354] Derewenda ZS, Vekilov PG. Entropy and surface engineering in protein crystallization. *Acta Crystallogr D Biol Crystallogr*. 2006;62:116-24.

[355] Walter TS, Meier C, Assenberg R, Au KF, Ren J, Verma A, et al. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*. 2006;14:1617-22.

[356] Sledz P, Zheng H, Murzyn K, Chruszcz M, Zimmerman MD, Chordia MD, et al. New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Sci*. 2010;19:1395-404.

[357] Tan K, Kim Y, Hatzos-Skintges C, Chang C, Cuff M, Chhor G, et al. Salvage of failed protein targets by reductive alkylation. *Methods Mol Biol*. 2014;1140:189-200.

[358] Derewenda ZS. Rational protein crystallization by mutational surface engineering. *Structure*. 2004;12:529-35.

[359] Cooper DR, Boczek T, Grelewska K, Pinkowska M, Sikorska M, Zawadzki M, et al. Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr D Biol Crystallogr*. 2007;63:636-45.

[360] Gao L, Lam KH, Liu S, Przykopanski A, Lübke J, Qi R, et al. Crystal structures of OrfX1, OrfX2 and the OrfX1-OrfX3 complex from the *orfX* gene cluster of botulinum neurotoxin E1. *FEBS Lett*. 2023;597:524-37.

[361] Diaz E, Adhikary S, Tepper AWJW, Riley D, Ortiz-Meoz R, Krosky D, et al. Structure of human spermine oxidase in complex with a highly selective allosteric inhibitor. *Commun Biol*. 2022;5:787.

[362] Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH. The amino acid following an Asn-X-Ser/Thr sequon is an important determinant of *N*-linked core glycosylation efficiency. *Biochemistry*. 1998;37:6833-7.

[363] Aricescu AR, Owens RJ. Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology. *Curr Opin Struct Biol*. 2013;23:345-56.

[364] Kozak S, Bloch Y, De Munck S, Mikula A, Bento I, Savvides SN, et al. Homogeneously *N*-glycosylated proteins derived from the GlycoDelete HEK293 cell line enable diffraction-quality crystallogenesis. *Acta Crystallogr D Struct Biol*. 2020;76:1244-55.

[365] Reeves PJ, Callewaert N, Contreras R, Khorana HG. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous *N*-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc Natl Acad Sci U S A*. 2002;99:13419-24.

[366] Puthalakath H, Burke J, Gleeson PA. Glycosylation defect in Lec1 Chinese hamster ovary mutant is due to a point mutation in *N*-acetylglucosaminyltransferase I gene. *J Biol Chem*. 1996;271:27818-22.

[367] Elbein AD, Tropea JE, Mitchell M, Kaushal GP. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. *J Biol Chem*. 1990;265:15599-605.

[368] Gross V, Tran-Thi TA, Vosbeck K, Heinrich PC. Effect of swainsonine on the processing of the asparagine-linked carbohydrate chains of alpha 1-antitrypsin in rat hepatocytes. Evidence for the formation of hybrid oligosaccharides. *J Biol Chem*. 1983;258:4032-6.

[369] Almahayni K, Spiekermann M, Fiore A, Yu G, Pedram K, Möckl L. Small molecule inhibitors of mammalian glycosylation. *Matrix Biol Plus*. 2022;16:100108.

[370] Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, et al. Glycoprotein structural genomics: solving the glycosylation problem. *Structure*. 2007;15:267-73.

[371] Quistgaard EM. A disulfide polymerized protein crystal. *Chem Commun (Camb)*. 2014;50:14995-7.

[372] Banatao DR, Cascio D, Crowley CS, Fleissner MR, Tienson HL, Yeates TO. An approach to crystallizing proteins by synthetic symmetrization. *Proc Natl Acad Sci U S A*. 2006;103:16230-5.

[373] Forse GJ, Ram N, Banatao DR, Cascio D, Sawaya MR, Klock HE, et al. Synthetic symmetrization in the crystallization and structure determination of CelA from *Thermotoga maritima*. *Protein Sci*. 2011;20:168-78.

[374] Heinz DW, Matthews BW. Rapid crystallization of T4 lysozyme by intermolecular disulfide cross-linking. *Protein Eng*. 1994;7:301-7.

[375] Leibly DJ, Arbing MA, Pashkov I, DeVore N, Waldo GS, Terwilliger TC, et al. A suite of engineered GFP molecules for oligomeric scaffolding. *Structure*. 2015;23:1754-68.

[376] Laganowsky A, Zhao M, Soriaga AB, Sawaya MR, Cascio D, Yeates TO. An approach to crystallizing proteins by metal-mediated synthetic symmetrization. *Protein Sci*. 2011;20:1876-90.

[377] Yamada H, Tamada T, Kosaka M, Miyata K, Fujiki S, Tano M, et al. 'Crystal lattice engineering,' an approach to engineer protein crystal contacts by creating intermolecular symmetry: crystallization and structure determination of a mutant human RNase 1 with a hydrophobic interface of leucines. *Protein Sci*. 2007;16:1389-97.

[378] Burmeister WP, Huber AH, Bjorkman PJ. Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature*. 1994;372:379-83.

[379] Warke A, Momany C. Addressing the protein crystallization bottleneck by cocrystallization. *Crystal Growth & Design*. 2007;7:2219-25.

[380] Bukowska MA, Grütter MG. New concepts and aids to facilitate crystallization. *Curr Opin Struct Biol*. 2013;23:409-16.

[381] Partridge BE, Winegar PH, Han Z, Mirkin CA. Redefining protein interfaces within protein single crystals with DNA. *J Am Chem Soc*. 2021;143:8925-34.

[382] Chesterman C, Arnold E. Co-crystallization with diabodies: a case study for the introduction of synthetic symmetry. *Structure*. 2021;29:598-605.e3.

[383] Ramsland PA, Bradford TM, Farrugia W, Hogarth PM. Engineering immune system glycoproteins to form uniform crystalline lattices. *Acta Crystallographica Section A*. 2005;61:c230.

[384] Kirchdoerfer RN, Herrin BR, Han BW, Turnbough CL, Cooper MD, Wilson IA. Variable lymphocyte receptor recognition of the immunodominant glycoprotein of *Bacillus anthracis* spores. *Structure*. 2012;20:479-86.

[385] Davies AM, Allan EG, Keeble AH, Delgado J, Cossins BP, Mitropoulou AN, et al. Allosteric mechanism of action of the therapeutic anti-IgE antibody omalizumab. *J Biol Chem*. 2017;292:9975-87.

[386] Mitropoulou AN, Ceska T, Heads JT, Beavil AJ, Henry AJ, McDonnell JM, et al. Engineering the Fab fragment of the anti-IgE omalizumab to prevent Fab crystallization and permit IgE-Fc complex crystallization. *Acta Crystallogr F Struct Biol Commun*. 2020;76:116-29.

[387] Sibéril S, Ménez R, Jorieux S, de Romeuf C, Bourel D, Fridman WH, et al. Effect of zinc on human IgG1 and its FcγR interactions. *Immunol Lett*. 2012;143:60-9.

[388] Karplus PA, Diederichs K. Assessing and maximizing data quality in macromolecular crystallography. *Curr Opin Struct Biol*. 2015;34:60-8.

[389] Karplus PA, Diederichs K. Linking crystallographic model and data quality. *Science*. 2012;336:1030-3.

[390] Leaver-Fay A, Froning KJ, Atwell S, Aldaz H, Pustilnik A, Lu F, et al. Computationally designed bispecific antibodies using negative state repertoires. *Structure*. 2016;24:641-51.

[391] Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ*. 2014;1:213-20.

[392] Matthews BW. Solvent content of protein crystals. *J Mol Biol*. 1968;33:491-7.

[393] Urzhumtsev A, Afonine PV, Adams PD. TLS from fundamentals to practice. *Crystallogr Rev*. 2013;19:230-70.

[394] Porter RR. The hydrolysis of rabbit y-globulin and antibodies with crystalline papain. *Biochem J*. 1959;73:119-26.

[395] Dixon EV. Mechanisms of immunoglobulin deactivation by *Streptococcus pyogenes*: University of Oxford; 2014.

[396] Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75:861-77.

[397] Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, et al. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol*. 2017;73:148-57.

[398] Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983;22:2577-637.

[399] Touw WG, Baakman C, Black J, te Beek TA, Krieger E, Joosten RP, et al. A series of PDB-related databanks for everyday needs. *Nucleic Acids Res*. 2015;43:D364-8.

[400] Coulombe R, Grochulski P, Sivaraman J, Ménard R, Mort JS, Cygler M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J*. 1996;15:5492-503.

[401] Cygler M, Sivaraman J, Grochulski P, Coulombe R, Storer AC, Mort JS. Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure*. 1996;4:405-16.

[402] Sivaraman J, Lalumière M, Ménard R, Cygler M. Crystal structure of wild-type human procathepsin K. *Protein Sci*. 1999;8:283-90.

[403] Bockermann R, Järnum S, Runström A, Lorant T, Winstedt L, Palmqvist N, et al. Imlifidase-generated single-cleaved IgG: implications for transplantation. *Transplantation*. 2021.

[404] Yeh AH, Norn C, Kipnis Y, Tischer D, Pellock SJ, Evans D, et al. *De novo* design of luciferases using deep learning. *Nature*. 2023;614:774-80.

[405] Bennett NR, Coventry B, Goreshnik I, Huang B, Allen A, Vafeados D, et al. Improving *de novo* protein binder design with deep learning. *Nat Commun*. 2023;14:2625.

[406] Trastoy B, Lomino JV, Wang LX, Sundberg EJ. Liquid-liquid diffusion crystallization improves the X-ray diffraction of EndoS, an endo-β-*N*-acetylglucosaminidase from *Streptococcus pyogenes* with activity on human IgG. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2013;69:1405-10.

[407] Seki H, Huang Y, Arakawa T, Yamada C, Kinoshita T, Iwamoto S, et al. Structural basis for the specific cleavage of core-fucosylated *N*-glycans by endo-β-*N*-acetylglucosaminidase from the fungus *Cordyceps militaris*. *J Biol Chem*. 2019;294:17143-54.

[408] Waddling CA, Plummer TH, Tarentino AL, Van Roey P. Structural basis for the substrate specificity of endo-β-*N*-acetylglucosaminidase F3. *Biochemistry*. 2000;39:7878-85.

[409] García-Alija M, Du JJ, Ordóñez I, Diz-Vallenilla A, Moraleda-Montoya A, Sultana N, et al. Mechanism of cooperative *N*-glycan processing by the multi-modular endoglycosidase EndoE. *Nat Commun*. 2022;13:1137.

[410] Deis LN, Wu Q, Wang Y, Qi Y, Daniels KG, Zhou P, et al. Suppression of conformational heterogeneity at a protein-protein interface. *Proc Natl Acad Sci U S A*. 2015;112:9028-33.

[411] Petrescu AJ, Petrescu SM, Dwek RA, Wormald MR. A statistical analysis of *N*- and *O*-glycan linkage conformations from crystallographic data. *Glycobiology*. 1999;9:343-52.

[412] Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE. Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat Struct Biol*. 1996;3:638-48.

[413] Fan S, Li W, Zhang K, Zou X, Shi W, Liu Z, et al. Enhanced antibody-defucosylation capability of α-L-fucosidase by proximity-based protein fusion. *Biochem Biophys Res Commun*. 2023;645:40-6.

[414] Trastoy B, Du JJ, Cifuente JO, Rudolph L, García-Alija M, Klontz EH, et al. Mechanism of antibody-specific deglycosylation and immune evasion by streptococcal IgG-specific endoglycosidases. *Nat Commun*. 2023;14:1705.

[415] Parsons S. Introduction to twinning. *Acta Crystallogr D Biol Crystallogr*. 2003;59:1995-2003.

[416] Padilla JE, Yeates TO. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr D Biol Crystallogr*. 2003;59:1124-30.