Structural investigations into conformational diversity, polyspecificity, and binding mechanisms of near-germline antibodies

by

Ryan J. Blackler
BSc, University of Victoria, 2010

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry and Microbiology

Supervisory Committee

Structural investigations into conformational diversity, polyspecificity, and binding mechanisms of near-germline antibodies

by

Ryan Blackler
B.Sc., University of Victoria, 2010

Supervisory Committee

Dr. Stephen V. Evans (Department of Biochemistry and Microbiology)
Supervisor

Dr. Alisdair Boraston (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Caroline Cameron (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Jeremy Wulff (Department of Chemistry)
Outside Member

Abstract

Supervisory Committee

Dr. Stephen V. Evans (Department of Biochemistry and Microbiology)
Supervisor

Dr. Alisdair Boraston (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Caroline Cameron (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Jeremy Wulff (Department of Chemistry)
Outside Member

The antibody response has evolved under constant pressure to recognize common pathogens and also remain adaptable to novel threats. Given the limited size of the germline antibody repertoire, adaptability requires that some antibodies must be polyspecific for multiple distinct antigens. Despite the profound importance of polyspecificity in the antibody response, the structural features that allow it are not well understood.

Antibodies raised against glycoconjugates of Chlamydiaceae LPS oligosaccharides of the inner-core sugar Kdo (3-deoxy-D-*manno*-oct-2-ulosonic acid) have been shown to cross-react with several inner-core oligosaccharides through conserved recognition of single Kdo residues in a germline-encoded pocket, with additional sugars accommodated by flexible side-chains. Two of these antibodies, S25-2 and S25-39, were observed to bind several Kdo oligosaccharides with an identical binding site conformation, but adopted unique conformations of the heavy chain complementarity determining region loop 3 (CDR H3) in the absence of ligand.

Conformational flexibility of germline antibodies is believed to facilitate polyspecificity by generating multiple unique binding sites in a single antibody. This thesis research further explores the conformational flexibility of the antibodies S25-2 and S25-39 to gain insight into mechanisms of antigen recognition and how this feature may allow polyspecificity. This was achieved first by solving structures of S25-39 from crystals grown in unique conditions to observe alternate CDR H3 conformations, and second by designing synthetic Kdo-based antigens so as both to inhibit interaction with the previously observed liganded conformation of S25-2 and S25-39 and to be accommodated by their observed unliganded conformations.

These structures reveal an unprecedented level of structural diversity of CDR H3, notably including the exact 'liganded' conformation in the absence of ligand. This is the first direct structural evidence that CDR H3 can exist in a conformational equilibrium with antigen binding through a selection mechanism, as opposed to induced fit where antigen causes the observed conformational change. Definitive evidence for binding the synthetic antigens was not obtained, however the resulting structures revealed several additional unique conformations of CDR H3 suggesting that ligands can alter conformational equilibria during crystallization. A unique conformation was also observed with CDR H3 coordinating multiple iodide ions, revealing another potential source of polyspecificity with unique binding paratopes generated by ion coordination.

Finally, the unparalleled level of conformational diversity observed for these antibodies highlights the challenges of antibody structure classification and prediction,

and stresses the need for additional in-depth studies of conformational diversity and binding mechanisms to advance these fields for therapeutic application.

This is the first targeted structural study of flexibility in antibodies and provides insight into their conformational dynamics and antigen-binding mechanisms. These are of fundamental importance in understanding antibody structure and function, a critical consideration in practical applications such as modelling and design of therapeutic or diagnostic antibodies.

Table of Contents

List of Abbreviations

Ab	Antibody
Ag	Antigen
CDR	Complementarity Determining Region
CS	Conformational Selection
ELISA	Enzyme Linked Immunosorbent Assay
F_{AB}	Fragment Antigen-Binding
F_C	Fragment Crystallizable
F_V	Fragment Variable
IF	Induced Fit
Ig	Immunoglobulin
Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
LPS	Lipopolysaccharide
mAb	Monoclonal Antibody
MHC	Major Histocompatibility Complex
MPD	2-Methyl-2,4-pentanediol
PEG	Polyethylene Glycol
rmsd	Root-mean-square deviation
ScFv	Single chain Fragment variable
SPR	Surface Plasmon Resonance
TD	T-Cell Dependent
TI	T-Cell Independent
VH	Heavy chain Variable domain
VL	Light chain Variable domain

List of Tables

List of Figures

Acknowledgments

I am extremely grateful to my supervisor Stephen Evans, who gave me my start in research as an undergraduate and has been endlessly supportive as I pursued this PhD research. Steve is an inspiration in his brilliance and curiosity, and in the care and compassion he shows for his students, which together make him a truly remarkable mentor.

I thank my committee members Caroline Cameron, Alisdair Boraston and Jeremy Wulff for their advice and useful discussions over the years.

Thanks to my collaborators Sven Müller-Loennies, Helmut Brade, Lore Brade, Roger MacKenzie and Paul Kosma, who made this research possible and have always been eager to explore new ideas.

My time in the Evans lab has been enriched by amazing coworkers and friends: Cory Brooks, Brock Schuman, Javier Alfaro, Asha Johal, Kathryn Gomery, Dylan Evans, Matthew Parker, Omid Ghassemi and Susannah Gagnon. I especially thank Cory Brooks, who was my first mentor in the Evans lab and provided me with the foundation in research that helped me to get to where I am today.

Finally, I thank all my friends and family who supported me through the ups and downs of graduate school.

Dedication

To the natural world for its infinite supply of mysteries and wonders, and for providing us with the capacity to explore and admire them.

To my loved ones.

Chapter 1: Introduction

Portions of this chapter are adapted or reproduced from the publication ***Blackler RJ, et al.*** *(2012) Antibody Recognition of Chlamydia LPS: Structural Insights of Inherited Immune Responses. Anticarbohydrate Antibodies, eds Kosma P, Müller-Loennies S (Springer Vienna, Vienna), pp 75–120. See Appendix B for permissions.*

1.1. A history of antibody discovery and theories for their formation

> *"Yet it was with those who had recovered from the disease that the sick and the dying found most compassion. These knew what it was from experience, and had now no fear for themselves; for the same man was never attacked twice – never at least fatally."*
>
> *– Thucydides, Athens approx. 430 BC – Translated in* (1)

Perhaps the first recorded reference to adaptive immunity was in the Greek historian Thucydides' account '*The Peloponnesian war*' where he describes the plague of Athens and the resistance of recovered individuals to re-infection. This phenomenon was observed and exploited across Asia and Europe where inoculation (or 'variolation' – a term derived from the causative agent of smallpox, Variola virus) was practiced since at least the 10th century. However, the foundation of modern immunology is often credited to Edward Jenner's research and insistent promotion of vaccination, the administration of Vaccinia virus (the causative agent of cowpox) for the control of smallpox infection around the year 1800 (2, 3). The term vaccination is now generally used to refer to the controlled presentation of any antigenic material to the immune system to develop immunity to a foreign pathogen or any disease with an identifiable molecular signature (e.g. cancer, Alzheimer's disease, *etc*.).

The role of antibodies in immunity was realised by Emil von Behring and Shibasabura Kitasato in 1890 when they showed that animals given a fatal dose of

diphtheria toxin could be protected through a serum transfer from previously immunized animals by the action of 'anti-toxins' (4). This practise of passive immunization was a major research focus of Paul Ehrlich, who in 1891 coined the term 'antibody' (5) and in 1897 proposed his 'side-chain' theory for their production: that a cellular receptor (the side-chain) will specifically recognize "as a key is to its lock" the foreign substance, and the cell will respond by producing an excess of these receptors, some of which will be released into the serum to function as antibodies (6, 7).

Ehrlich's side-chain theory prevailed for over thirty years, although studies by Karl Landsteiner and others showing that antibodies could be formed against a variety of chemically synthesized haptens cast a growing doubt on the ability of a pre-existing antibody repertoire to match a seemingly infinite antigenic diversity (8). These doubts led to an increasing popularity of 'instructional' theories, implying that antigen played a role in dictating the specificity of antibody, including Linus Pauling's landmark 'direct template' theory in 1940 that proposed antigen at the site of antibody production serves as a folding template to achieve perfect complementary (9, 10). Pauling's theory reconciled the limited size of the antibody repertoire with its ability to match an infinite diversity of antigens but was unable to explain certain phenomena, such as evidence for antibody multivalency through the precipitation of multiply substituted proteins.

Around the same time, Frank Burnet proposed his 'adaptive enzyme' theory that attempted to better explain observed biologic aspects of the immune response compared to Pauling's theory that focused mainly on the chemistry of recognition. Burnet's theory stated that antibody-synthesizing enzymes were modified by contact

with antigen to then produce specific antibodies, and these modified enzymes were replicated and passed on to daughter cells, explaining the heightened immune response upon re-exposure to antigen (11). His theory was modified later in the same decade upon the discovery that proteins are under the control of a 'genome', to suggest it was rather this genome that was modified by antigen to then produce specific antibody (12).

In 1955, largely in attempt to explain pre-existing so-called 'natural antibodies' that could not be accounted for by the instructional theories, Niels Jerne refreshed some of Ehrlich's ideas and presented his 'Natural Selection' theory: that antibodies of all specificities circulated in the blood and upon binding antigen would migrate to specialized cells where they would influence the cell's RNA to cause further production of specific antibody (13).

The critical flaw in theories suggesting an influence of antigen or antibody on DNA or RNA was emphasized in 1958 with Francis Crick's "central dogma" of genetics, that described information flow as a one-way street from DNA to RNA to protein (14). This led Burnet and others back towards concepts first introduced by Ehrlich, for the development of the 'clonal selection' theory: that natural antibodies pre-exist on the surface of lymphocyte clones, where interaction with antigen triggers clone proliferation and somatic mutation to generate "better" antibodies (15–17). Lederberg further suggested that antibody specificity was determined by a unique amino acid sequence encoded in nucleic acid, which was subject to a high rate of mutation to account for antibody diversity.

Although the clonal selection theory was followed by years of debate as to the genetic source of antibody diversity, it accurately predicted the basic tenets by which the antibody response is now known to function. Many copies of unique immunoglobulins (Ig) are present on the surface of individual circulating B-lymphocyte clones as part of the B-cell receptor, and binding to antigen provides a signal to the B-cell to proliferate and differentiate into antibody secreting plasma cells or long-lasting memory B-cells (18–21). The genetic basis for antibody diversity and more detail into the antibody response are given in the following sections.

1.2. Early studies of general antibody structure

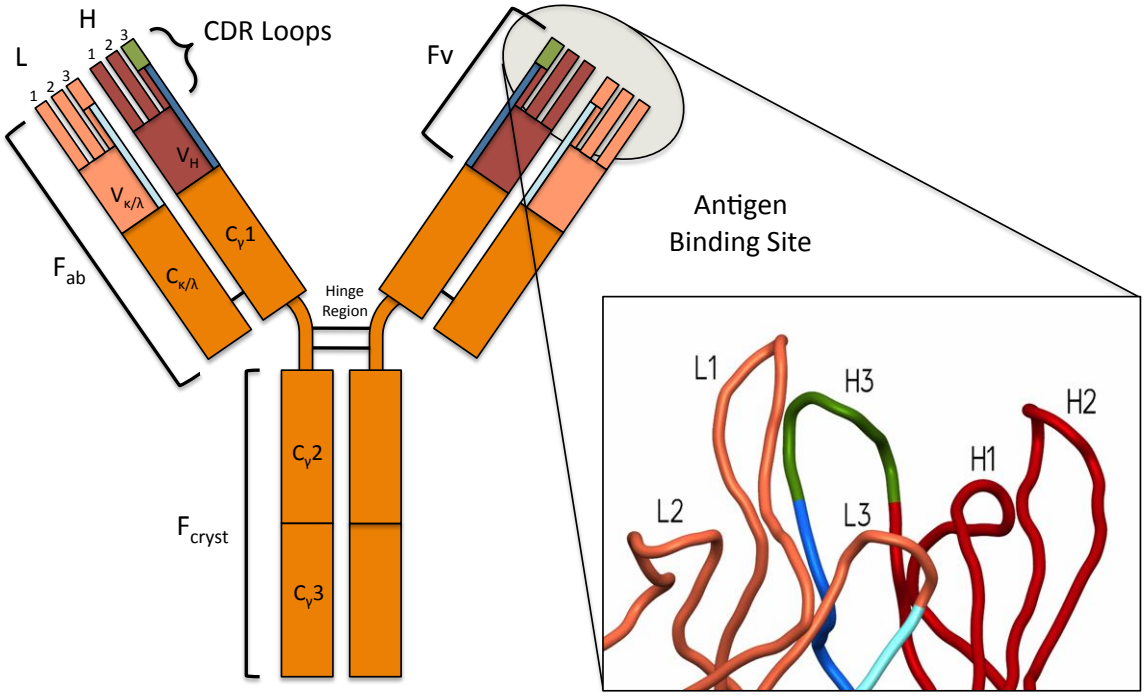
Many fundamentals of antibody structure and function were elucidated in early studies by Rodney Porter and Gerald Edelman, for which they shared the Nobel Prize in medicine in 1972 (22–24). Following in the footsteps of Petermann and Landsteiner, who discovered that an intact immunoglobulin (Ig) is not required for antigen specificity (25–28), Porter's work on the digestion of IgG[1] with papain revealed its multivalent nature and established the existence of fragments that were 'antigen binding' (F_{AB}) and 'crystallizable' (F_C) (Figure 1) (29). Edelman discovered that antibodies were composed of multiple chains cross-linked by disulfide bridges (30, 31). Further research established the 'four peptide chain structure' of IgG with two light chains and two heavy chains (32) and a more precise mapping of disulfide bridges (33). In a landmark paper, Kabat showed that immunoglobulins as a group possessed six regions of hyper-variable

[1] Antibodies can exist in a number of isotypes, of which IgG is the most common. These are explained in section 1.3.2

sequence (three on the light chain and three on the heavy chain), which he hypothesized to lie at the basis of individual antibody specificity (34). The precise domain structure and location of hypervariable regions were then revealed with the first high-resolution crystal structure of an antibody F_{AB}, determined in 1973 by Poljak *et al.* (35), as will be described in section 1.4.

Figure 1: General schematic of IgG structure

The IgG molecule is composed of two heavy chains and two light chains. The heavy chains each possess four domains; V_H, Cγ1, Cγ2 and Cγ3. The heavy chains dimerize through the Cγ2 and Cγ3 domains and through disulfide bonds in the hinge region. The V_H and Cγ1 domains of each heavy chain pair with the Vκ and Cκ or Vλ and Vλ domains depending on whether the light chain is of kappa or lambda type. The dimerized Cγ2 and Cγ3 domains form the 'Fragment Crystallizable' or F_{cryst} or F_C, which is recognized by Fc receptors on various types of immune cells and is responsible for the effector functions of the antibody. The heavy chain V_H and Cγ1 domains paired with light chain Vκ/λ and Cκ/λ domains each form a 'Fragment Antigen-Binding' or F_{AB}. This is further subdivided into the 'Fragment Variable' or F_V, containing only the V_H and Vκ/λ domains. Each F_V contains six loops of hypervariable sequence deemed complementarity determining regions, or CDRs, that form the antigen binding site (inset shows CDR loops from an antibody crystal structure, PDB identifier code 3OKE[2]). The schematic and CDRs are coloured according to which gene segment encodes which region, consistent with Figure 2.



[2] Atomic coordinates available from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (http://www.rcsb.org/pdb/home/home.do)

1.3. The genetic basis for antibody diversity

The generation of antibodies from the recombination of multiple gene segments was first hypothesized by Dreyer and Bennett in 1965 to explain the generation of a vast antibody repertoire from a limited genome size (36). The first experimental evidence for this phenomenon was provided by Hozumi and Tongeawa in 1976, when comparative restriction digests and hybridization revealed that the variable and constant regions of kappa light chain were located on two distant fragments in early embryo DNA versus a single joined region in plasmacytoma DNA (37). Additional DNA digestion, hybridization and sequencing experiments by Tonegawa and others (37–47) led to his seminal publication In 1983, *somatic generation of antibody diversity* (48), where he described the assembly of complete immunoglobulin genes during lymphocyte development from multiple gene segments scattered across the chromosome. In 1987, Tonegawa was awarded the Nobel prize in medicine for this research.

1.3.1. V(D)J recombination: generating the germline repertoire

This assembly involves the recombination of Variable (V), Diversity (D) and Joining (J) gene segments to create the heavy chain variable region, and the recombination of V and J gene segments to create the light chain variable region. The V(D)J-recombination process occurs at the DNA level during lymphocyte development, so that a mature B-cell contains a single contiguous VDJ heavy chain gene segment and a single contiguous VJ light chain gene segment (Figure 2). It is the light and heavy chain variable regions that encode the antibody Fv, which contains the antigen binding site (Figure 1). Mature B-

cells therefore express a single unique antibody as part of their B-cell receptor (BCR) from the rearranged Ig DNA loci.

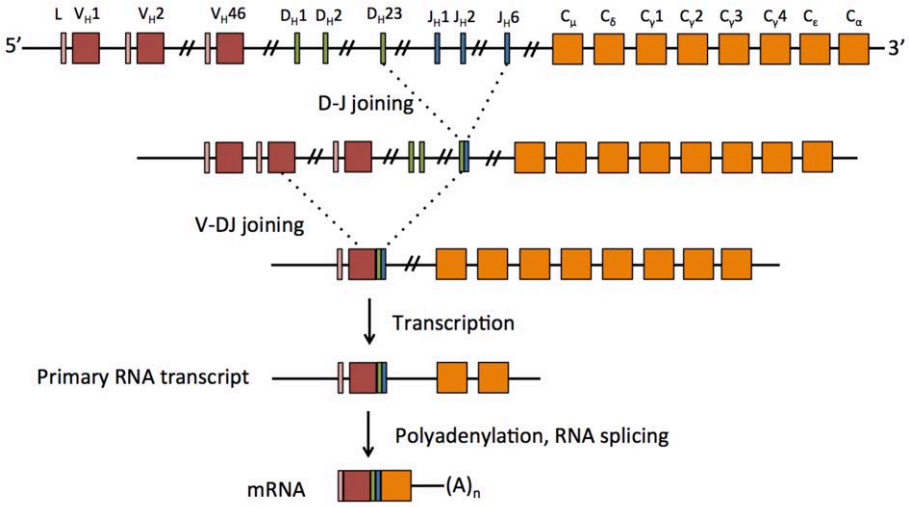
There are a total of 38-46 V, 23 D, 6 J heavy chain gene segments, 29-33 V and 4-5 J lambda light chain gene segments and 31-36 V and 5 J kappa light chain gene segments in humans, and a total of 97 V, 14 D, 4 J heavy chain gene segments, 3 V and 3 J lambda light chain gene segments and 95-96 V and 4 J kappa light chain gene segments in mice (*mus musculus*) (Table 1). V(D)J recombination and differential light and heavy chain pairings can therefore generate a maximum of 1.3×10^7 unique antibody variable fragments in both humans and mice (49–51).

Further diversity is generated during V(D)J recombination through junctional flexibility, where imprecise joining of gene segments can generate alternate amino acids at each coding joint in CDR3, and nucleotide addition or deletion during recombination. Nucleotide addition can occur when DNA cleavage generates single-strand overhangs that are repaired to generate short palindromic sequences at coding joints in a process called P-addition, and also with the addition of up to 15 random non-templated nucleotides at V-DJ or VD-J coding joints by a terminal deoxynucleotidyl transferase (TdT) in a process called N-addition. Nucleotide deletion may also occur at cleavage sites prior to nucleotide addition (18, 52–54). Junctional diversity has been estimated to increase the combinatorial diversity of the germline repertoire to approximately 10^{11} different receptors (49, 55–57).

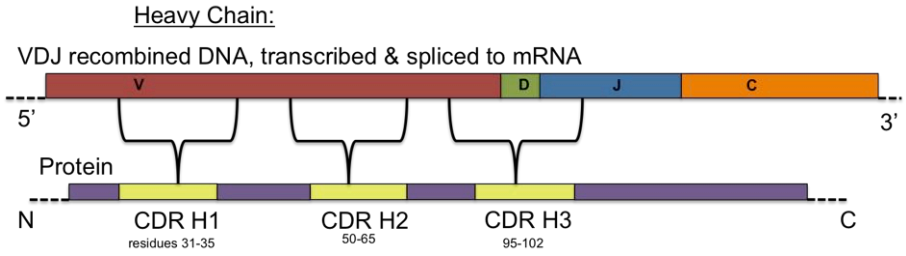
Figure 2: V(D)J germline gene segment recombination

(A) Functional antibody genes are formed by the recombination of germline gene segments during B-cell development, using so-called V, D and J gene segments in the heavy chain and V and J gene segments in the light chain. The primary RNA transcript of recombined VDJ DNA is then spliced with a constant gene segment that determines the antibody isotype, forming a complete heavy chain mRNA. **(B)** and **(C)** show the regions of VDJ gene segments that encode the CDRs. CDR Residue numbers follow Kabat definitions (see section 1.4.1).

A)



B)



C)

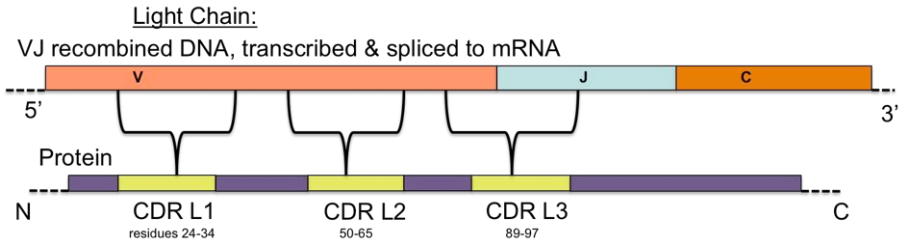


Table 1: Human and mouse Ig gene segment diversity

Gene segment numbers were obtained from the IMGT server[3] locus descriptions of IGH, IGL and IGK, only including functional genes. Ranges given for some gene segments account for allelic polymorphism by insertion/deletion.

	Human	Mouse
Heavy Chain		
V gene segments	38-46	97
D gene segments	23	14
D gene reading frames	6	6
J gene segments	6	4
Combinations	31464-38088	32592
Light Chains		
Kappa		
V gene segments	31-36	94-96
J gene segments	5	4
Combinations	155-180	376-384
Lambda		
V gene segments	29-33	3
J gene segments	4-5	3
Combinations	116-165	9
Combined combinatorial diversity of heavy and light chains	**$8.5 \times 10^6 - 1.3 \times 10^7$**	**1.3×10^7**

[3] (http://www.imgt.org/) (50)

1.3.2. Constant genes dictate antibody isotype

The VDJ or VJ recombined gene segments are further joined with Constant (C) gene segments within their respective Ig DNA loci, through splicing at the RNA level (18, 48). The light chain gene segments are separated in two multi-gene families, lambda (λ) and kappa (κ), located on two separate chromosomes (number 22 and 2 in human, 16 and 6 in mouse, respectively) and each contain different V and J gene segments and a single λ or κ C gene segment. There is a single heavy chain multi-gene family located on a single chromosome (14 in human, 12 in mouse), which, in contrast to light chain, contains multiple unique C gene segments. These are the C_μ, C_δ, C_γ, C_ε and C_α gene segments, with C_γ variants $C_\gamma1$, $C_\gamma2$, $C_\gamma3$ and $C_\gamma4$ in humans, and $C_\gamma1$, $C_\gamma2a$, $C_\gamma2b$ and $C_\gamma3$ in mice (Figure 2).

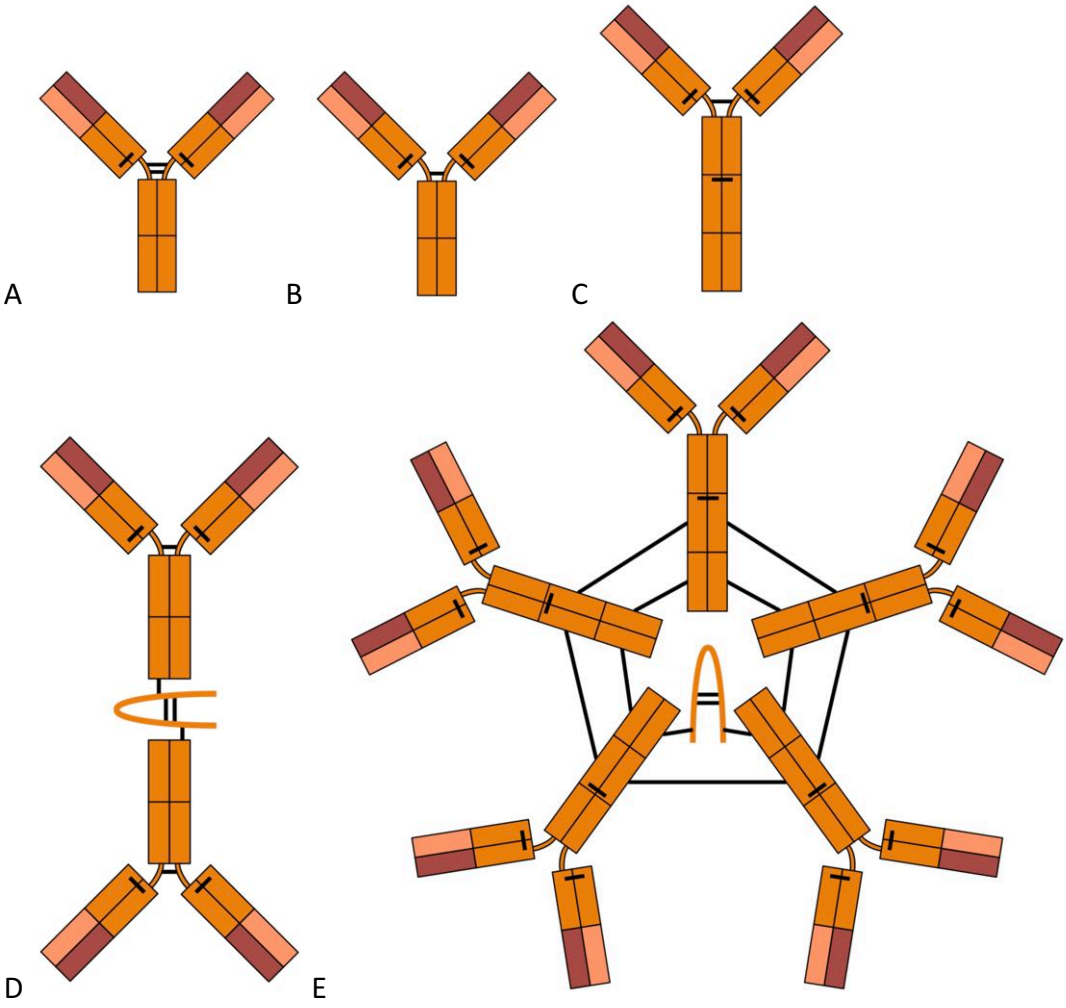
The heavy chain constant gene segment dictates the antibody isotype, for example IgM, IgD, IgG1, IgG2a, IgG2b, IgG3, IgE or IgA in mice. The isotypes have unique sizes and structures (Figure 3), are expressed in different stages of B-cell development or in response to different antigen classes, and have different effector functions. For example, IgG binds to F_C receptor of phagocytes and can cross the placenta, IgA (and IgM to a lesser extent) can be secreted in mucous, and IgE induces mast cell degranulation (18).

Mature, naïve B-cells express membrane bound germline IgM and IgD. Upon activation by antigen and T-helper cell help, a B cell can differentiate and undergo class-switching of the antibody isotype. The exact mechanism of class-switching is still unknown, but the process involves the enzyme 'activation-induced cytidine deaminase'

(AID) and results in the removal of C gene segments from the Immunoglobulin DNA loci so that a new gene segment is immediately downstream of the VDJ recombined DNA. For example, class switching from IgM to IgG involves the removal of C_μ and C_δ gene segments (Figure 2). IgG is the most abundant isotype in serum, and is the most commonly used in molecular biology and medicine because of its small size, stability, diversity of effector functions, and comparative ease of laboratory manipulation and characterization.

Figure 3: General schematics of immunoglobulin isotype structures

Simplified schematics to illustrate overall structures of **(A)** IgG **(B)** IgD **(C)** IgE **(D)** IgA and **(E)** IgM. Schematics are coloured with the V_H domain in red, the V_L domain in pink, and all constant domains in orange. Thick black lines indicate disulphide bonds. The orange loops in (D) and (E) each represent a protein component called the 'joining chain'.



1.3.3. Somatic hypermutation expands the diversity of the germline repertoire

When a mature naïve B-cell displaying membrane-bound IgM and IgD is activated by antigen stimulation and signalling from a T-helper cell, it will undergo somatic hypermutation of the Ig locus in addition to class-switching of the isotype as discussed above (18, 58–62). Somatic hypermutation occurs throughout the VDJ and VJ segments at a rate of approximately one mutation per every one or two cell divisions, which is at least 10^5 times higher than the normal gene mutation rate. After each cell division, B-cells expressing mutant BCRs with higher affinity for the stimulating antigen undergo further proliferation or differentiation into long-lasting Memory cells or antibody-secreting Plasma cells. The iterative process of somatic hypermutation and selection of B-cells displaying higher-affinity BCRs is called *affinity maturation.* Through documentation of mutations during an immune response, it has been observed that mutations are most likely to occur in the CDRs (18, 63).

1.4. Structural studies of antibodies

The first high-resolution structure of an antibody was the 2.8 Å crystal structure of an Fab' fragment of the antibody 'New', determined in 1973 by Poljak *et al.* (35). This structure revealed the four-subunit structure of the Fab, with the V_H and V_L subunits tightly associated to form a V domain, and the C_L and C_H1 subunits tightly associated to form a C domain (Figure 4). Each subunit consists of an immunoglobulin fold: a pair of antiparallel β sheets bridged by a disulfide bond and forming a hydrophobic core. The hypervariable regions of both chains are solvent exposed in loops at one end of the

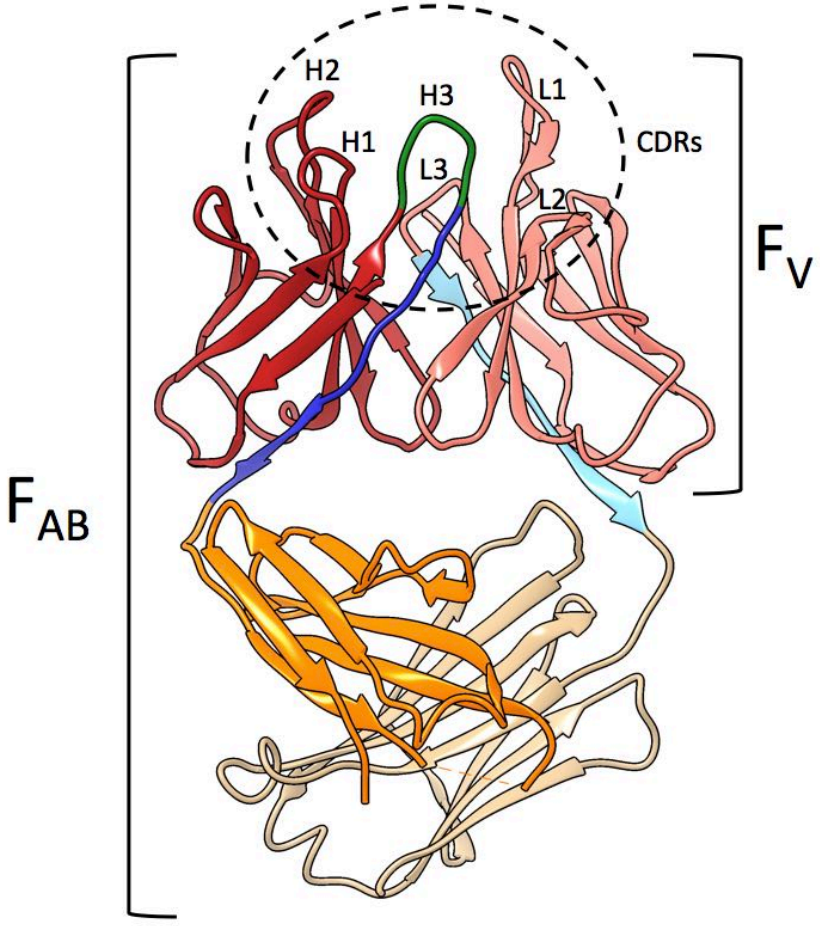
molecule where they are restrained by the framework regions, and together form the antigen binding site. These six hypervariable regions are now referred to as 'complementarity determining regions' or CDRs, and are numbered sequentially in the light and heavy chains as L1, L2, L3, H1, H2 and H3. In the light chain, L1 and L2 are encoded by the germline V_L gene, and L3 by the V_L-J_L junction. In the heavy chain, H1 and H2 are encoded by the V_H gene, whereas H3 is encoded by the V_H-D_H-J_H junction and is therefore the most variable CDR in sequence and structure (Figure 2, Figure 3, Figure 4).

There are now 2238 high-resolution structures[4] of antibodies or their fragments deposited in the protein data bank (www.rcsb.org/pdb) (64), which have confirmed the general structural features described above and have allowed for many highly detailed comparisons and characterizations of antibody structural features (65–74).

[4] According to The Structural Antibody Database (SabDab), accessed Dec 30 2015 <http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/Welcome.php> (107)

Figure 4: High resolution structure of an antibody F_{AB}

Ribbon diagram of a F_{AB} from a 2.4 Å crystal structure (PDB 3OKE). Regions are coloured according to germline gene segment descent as in Figure 1 and Figure 2 with the exception of the light chain constant region, which is coloured tan for better contrast with the heavy chain.



1.4.1. Antibody numbering systems

Given the enormous diversity of antibodies, the significance of antibodies in technology and medicine, and the utility of antibody structural analysis and comparison, it is desirable to have a standardized numbering scheme that allows the comparison of residues in structurally equivalent positions. A standardized numbering scheme allows identification of residues in critical positions to define CDR boundaries, to interact with antigen, to influence CDR conformations, to affect domain association, *etc*. Several such numbering schemes have been developed. The basis of all schemes is to provide consistent numbering of framework regions and to account for different CDR lengths with lettered residue numbers, e.g., 27A, 27B, 27C, *etc*. denote insertions at position 27.

The first scheme was developed by Kabat in 1983, based on sequence alignments before structural data were available (75). This was followed by the Chothia method based on structural data of antibody variable domains, which adjusted the sites of CDR insertions and deletions compared to the Kabat method (76). The Chothia method underwent several adjustments (67, 77), and the additional IMGT (78) and Aho (79) schemes were also introduced to provide unity with T-cell receptor numbering. The Kabat and Chothia schemes are the most commonly used (80), and it is the Kabat scheme used in this thesis to provide consistency with the associated publications and deposited structures.

1.4.2. CDR Canonical conformations

As the first antibody crystal structures were determined, Chothia pioneered attempts to classify CDR conformations based on sequence and structure (76). He suggested that

CDRs may each adopt only a small number of so-called 'canonical structures' based on length and sequence. These classifications were updated by Chothia and others as more structures became available (67, 77, 81), and more recent schemes divide CDR conformations into canonical 'clusters' based on the relative population of each cluster from available structural data (82, 83) (for example, cluster L2-8-1 would be the most common conformation among 8-residue long CDR L2s) and one such scheme is available *via* the webserver PyIgClassify[5] (82, 84).

Efforts to classify the conformations of CDR H3 have been impeded by the hypervariability of that loop in length and amino acid sequence, and there have been a number of studies to focus specifically on H3 classification. Kuroda *et al.* first proposed the "H3 rules" for classifying and predicting its structure in 1996 (85), which have since been revised and expanded by Kuroda and others (85–91).

CDR H3 is divided into two structural regions (Figure 5): the anchor (also called the base or torso), referring to N-terminal residues (-1) to (+1) and C-terminal residues (n – 3) to (n), where residues 1 to n represent H95 to H102 as numbered according to Kabat[6] (80, 92, 93), and the head (also called β-hairpin), referring to those residues in between (87). The base can be classified as either kinked (K) or extended (E) depending on the pseudo-dihedral angle between the Cα atoms from (n – 2) to (n + 1), called θ_{base}, where a kinked base is defined by a dihedral angle within the range of -100° to 100°. These base classifications are further sub-classified into E^P, E^N, K^T, K^G, K^C forms

[5] (http://dunbrack2.fccc.edu/PyIgClassify/)

[6] H in residue number indicates heavy chain, whereas L would indicate a light chain residue.

depending on the specific value of θ_{base} and a second pseudo-dihedral angle θ_{stem} from (n – 4) to (n – 1), and the additional K^+ form designating a second bulge immediately above a K^C base (Table 2). There are four possible structural classes of the β-hairpin region, A through D, which are dictated by CDR length, *n*, and the base subtype (Figure 6) (85–87, 94). The β-hairpin can form an antiparallel β-sheet and potentially a β-turn depending on its sequence. The CDR H3 rules i through iv determine the base type, subtype, β-hairpin structure and β-turn structure, respectively, as described in (85–87) and summarized in Table 3 through Table 6.

More recently, the CDR clustering scheme of North *et al.* (82) has categorized the CDR H3 base/anchor region into eight structural clusters. These clusters are defined by the backbone φ and ψ angles (Figure 7) of the first three and last four residues of CDR H3, and are named H3-anchor-1 through H3-anchor-7 and H3-anchor-*cis*4-1 according to their abundance among structures in the PDB (Table 7). The effect of CDR H3 length on anchor clustering was also investigated (Table 8). The same group clustered full-length CDR H3 conformations by the same methods, and found that while shorter loops cluster well, they are low in population, and longer loops form a large number of small clusters with low predictive value.

A recent analysis of antibody structures found that the majority of CDR H3s possess a kinked anchor, and that this kink is specifically stabilized by the Ig heavy chain fold (88). The authors hypothesize that the preference for a kinked base is an evolved mechanism for loop diversification, as it results in disruption of β-strand pairing and increases structural diversity.

Figure 5: CDR H3 structural regions

CDR H3 is divided into an anchor region (green) and a β-hairpin region (yellow). The anchor is composed of N-terminal residues (-1) to (+1) and C-terminal residues (*n* – 3) to (*n*), where residues 1 to *n* represent H95 to H102 as numbered according to Kabat (80, 92, 93), and the β-hairpin is composed of the residues in between. Shown is the backbone of CDR H3 from PDB 3OKE.

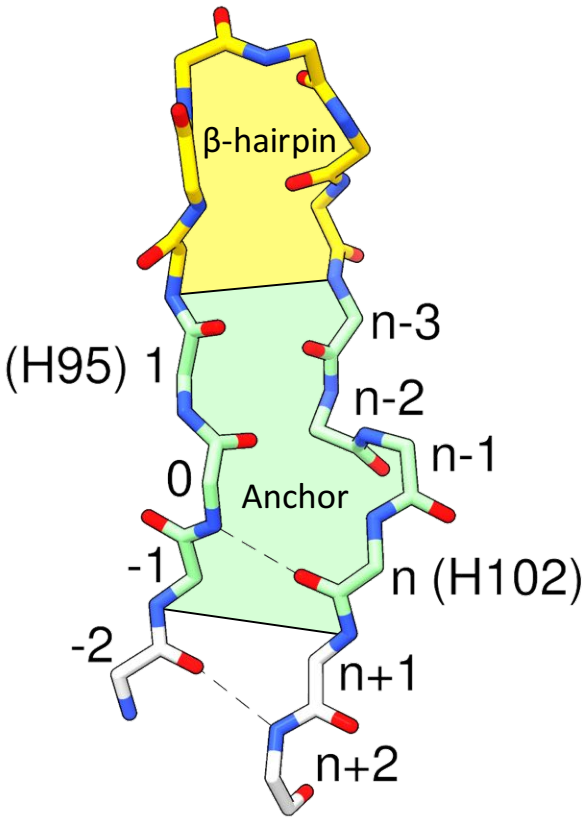


Table 2: CDR H3 structural classifications

Kinked (K) and Extended (E) base types and subtypes are determined by θ_{base} and dictate the resulting β-hairpin length (*m*), which in turn dictates the β-hairpin class. In determining the β-hairpin class, the value 'j' is any non-zero integer. For example, a CDR H3 of length *n* = 9 with a K^G base type will have a β-hairpin of length *m* = 7 that will be of class B. Base type and β-hairpin classification definitions are from reference (87). θ_{base} is the pseudo-dihedral angle between the Cα atoms from (*n* – 2) to (*n* + 1), and θ_{stem} is the pseudo-dihedral angle between the Cα atoms from (*n* – 4) to (*n* – 1).

	Structural definition		Outcome
Base	θ_{base}	θ_{stem}	Resulting β-hairpin length (*m*)
Kinked (K)	-100° to 100°		
K^T		-170° to -100°	*n* – 2
K^G		-100° to -10°	*n* – 2
K^C		-10° to 50°	*n* – 3
K^+		-10° to 50°	*n* – 3
Extended (E)	Outside -100° to 100°		
E^P	Around 160°		*n* – 1
E^N	Around -160°		*n* – 1
β-hairpin	β-hairpin length (*m*)		
Class A	4j-2		
Class B	4j-1		
Class C	4j		
Class D	4j-3		

Figure 6: CDR H3 β-hairpin classes

The β-hairpin region of CDR H3 can adopt four classes depending on its length, *m*, which is determined by the CDR length, *n*, and the CDR base type and subtype as outlined in Table 2. 'j' is any non-zero integer. Blue dashed lines indicate hydrogen bonds, and orange circles indicate key residue positions for CDR H3 rule iv to predict β-turns and outlined in Table 6. The yellow circle in class D indicates the special position of glycine for the case *m* = 5 of rule iv. Adapted from (87).

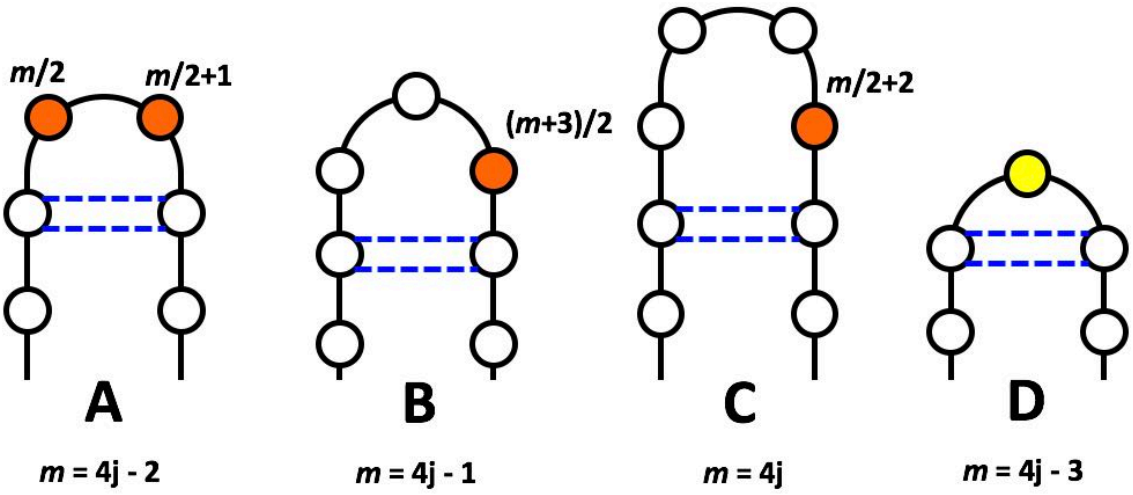


Table 3: CDR H3 rule i: identifying base type

CDR H3 rule i predicts the general base types Kinked (K) or Extended (E) based on the identities of key amino acids in the CDR. For each rule, if the conditions within the blue box are not satisfied, then the top line prediction of the rule is still valid. Adapted from (87).

	Position								
Rule	*n* – 1	*n* – 2	*n* – 3	1	0	-1	Length	Notable signals	Base type
i-a	Not Asp								K
		Hydrophilic							E
		Phe/Met	Ala/Gly	Tyr					E
		Not Phe/Met	Not Ala/Gly				7		E
i-b	Asp				Not Basic				E
		Phe/Met	Ala/Gly						K
								L49: Basic	K
i-c	Asp				Basic	Not Basic			K
								L46: Basic and L36: not Tyr	E

Table 4: CDR H3 rule ii: identifying base sub-type

CDR H3 rule ii predicts the base sub-type from the outcome of rule i and additional key amino acids. Adapted from (87).

		Position			
Rule	Base	*n* – 2	*n* – 3	0	Base subclass
ii-a	K		Gly		K^G
ii-b	K		Trp		
	K	Gly	Large hydrophobic		K^+
ii-c	E			Gly/Arg	E^N
	E			Not Gly/Arg	E^P

Table 5: CDR H3 rule iii: formation of the hydrogen bond ladder

CDR H3 rule iii predicts the structure of the β hairpin region depending on the base type and sub-type from rules I and ii, and additional key amino acids. Adapted from (87).

		Position						
Rule	Base	*2*	*3*	*4*	*m* – 3	*m* – 2	*m* – 1	β-hairpin
iii-a		Pro						Deformed, H-bond ladder broken
				Pro				Deformed, H-bond ladder broken
					Pro			Deformed, H-bond ladder broken
							Pro	Deformed, H-bond ladder broken
iii-b	K	Aromatic					Aromatic	Formed with H-bond ladder
	K		Aromatic			Aromatic		Formed with H-bond ladder

Table 6: CDR H3 rule iv: formation of the β-turn

CDR H3 rule iv predicts the formation of a typical β-turn depending on the β-hairpin class and additional key amino acids. Adapted from (87).

		Position				
Rule	β-hairpin	*m*/2	*m*/2 + 1	(*m*+3)/2	*m*/2 + 2	β-turn
iv	Class A	Gly/Asp/Asn				Typical
	Class A		Gly/Asp/Asn			Typical
	Class B			Gly/Asp/Asn		Typical
	Class C				Gly/Asp/Asn	Typical
	Class D	*When m=5 and the 3rd residue is Gly*				Typical

Figure 7: Regions of the Ramachandran map

Conformational regions of the Ramachandran map as defined in (82): B is the β-sheet region, P is polyproline II, A is the α-helix, D is the δ region (near α-helix but at more negative values of φ), L is the left-handed helix, and G is the γ region (φ>0°, excluding the L and B regions). Figure adapted from (82).

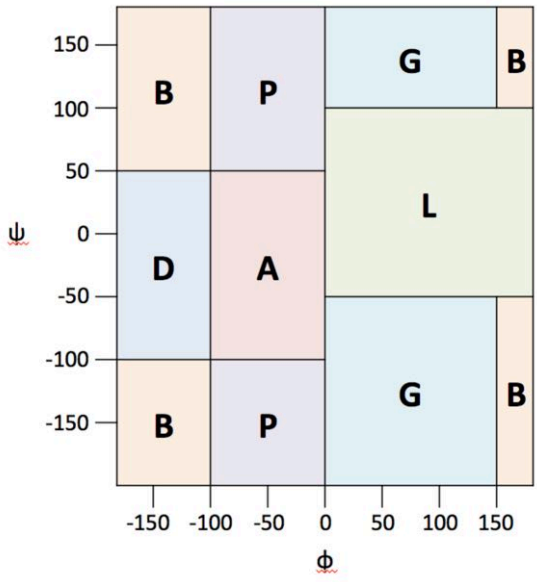


Table 7: CDR H3 anchor clusters

CDR H3 anchor clusters as defined in (82). Loop conformation according to regions defined in Figure 7 are given for the first three and last four residues of CDR H3, defined in (82) as Kabat-numbered residues 94 to 105. Residue conformations that differ from the largest cluster are shown in red, without differentiating between the similar regions B/P, A/D and L/G. Lower case residues indicate cis residues. Median Angle represents the average difference of φ or ψ from the median structure over all the loops in each cluster. Table adapted from (82).

Cluster	Number of unique sequences	Percent unique sequences	Loop conformation	Median Angle (°)	Comments
H3-anchor-1	169	67	BPP BPAB	21	Bulged
H3-anchor-2	32	13	BBB ABAB	30	Non-bulged
H3-anchor-3	23	9	BPP ABAB	23	H3-7 only
H3-anchor-4	10	4	BPA ALAB	7	Non-bulged
H3-anchor-5	12	5	BPB PPBB	33	H3-7 only
H3-anchor-6	5	2	BBL LLAB	16	
H3-anchor-7	4	2	BPP GBBL	26	
H3-anchor-*cis*4-1	2	1	BPA pLAB	38	

Table 8: Abundance of CDR H3-anchor clusters by H3 length

For each loop length of H3, the number of structures ("Count") and the percentage in each cluster are given. The most prevalent cluster percentage for each loop length is highlighted in green. Loop length includes Kabat numbered residues from 94 to 105. Table adapted from (82)

Loop length	Count	Percentage in each cluster (%)							
		1	2	3	4	5	6	7	*cis*4-1
7	34	9	-	-	68	-	18	-	3
8	5	40	20	-	-	40	-	-	-
9	26	35	65	-	-	-	-	-	-
10	29	79	21	-	-	-	-	-	-
11	27	71	11	19	-	-	-	-	-
12	37	74	6	12	-	6	-	2	-
13	40	78	10	5	-	-	-	8	-
14	28	75	4	7	-	14	-	-	-
15	13	92	-	-	-	8	-	-	-
16	24	92	-	-	-	8	-	-	-
>16	34	71	-	29	-	-	-	-	-

1.4.3. Antibody structure prediction

Since the first crystal structures of antibodies were determined and efforts to classify their structures began, the goal has been to develop reliable methods to predict antibody structures from their sequences. Such predictive power enables optimizing antibody affinity, specificity and stability, and is critical to the rapidly advancing field of antibody design for therapy and biotechnology (95–105). There are now many different antibody modeling methods and programs created by academia (105–112) and industry[7], that generally follow the same steps for structure prediction:

1) Selection of template antibody structures based on sequence alignments, including CDRs L1, L2, L3, H1 and H2, and the H3 anchor/base/torso
2) Model building from templates and from canonical CDR structures, and *ab initio* modeling of CDR H3 (i.e. without template or with partial templates)
3) Model evaluation and ranking of conformations

To compare the state-of-the art antibody modeling methods in their abilities to reproduce accurate structures, the "Antibody Modeling Assessment" (AMA) was initiated in 2009, where two homology modeling strategies (independently developed by Accelrys Inc. and the Chemical Computing Group CCG) and two fully automated modeling servers (PIGS and Rosetta Antibody) were compared in a blinded structure prediction study (113). All methods performed similarly with the PIGS, Rosetta, ACC and

[7] For example, Accelrys Antibody Modeling (http://accelrys.com/products/collaborative-science/biovia-discovery-studio/antibody-modeling.html), Chemical Computing Group Protein and Antibody Modeling (https://www.chemcomp.com/MOE-Protein_and_Antibody_Modeling.htm), Schrödinger Biologics Suite (http://www.schrodinger.com/biologics/), Macromoltek SmrtMolAntibody (https://www.macromoltek.com/), Kotai Antibody Builder (http://kotaiab.org/).

CCG models having average root-mean-square deviation (rmsd) of atomic positions[8] from the target structure backbone atoms of 1.3 Å for the whole Fv, 1.0 Å for framework regions, 1.2 Å for CDRs with canonical structures, and 3.1 Å for CDR H3.

Recently, a second antibody modeling assessment (AMA) was undertaken to assess improvements in antibody modeling from method development and the increased number of antibody structures in the PDB available for use as templates (114, 115). In this AMA, the average rmsd for distances of modeled backbone atoms from the target structures were 1.1 Å for the whole Fv, 0.9 Å for framework regions, 1.1 Å for CDRs with canonical structures, and 2.8 Å for CDR H3, showing improvements in all areas.

Given the difficulty in classifying CDR H3 loop conformations as discussed in section 1.4.2, it is not surprising that prediction of CDR H3 structure from sequence is the least reliable. This difficulty has attracted the attention of researchers to develop specialized tools for H3 prediction. For example, the database loop prediction server FREAD (http://opig.stats.ox.ac.uk/webapps/fread/php/) has been repurposed for H3 modeling as FREAD-S, with a focus on sequence similarity, and conFREAD, which additionally includes contact information, neither of which utilize canonical rules but instead rely on a fragment-based database search method (112). conFREAD was reported to achieve an average rmsd of 1.23 Å for CDR H3 prediction, and an rmsd of 1.35 Å when specifically predicting bound structures from the starting point of an

[8] rmsd is a measure of the average distance between atoms of superimposed proteins: $\mathrm{rmsd} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$ where δ is the distance between N pairs of equivalent atoms.

unbound structure. Another recent study presented a method based on an automatic learning technique to select optimal structural templates as starting points for CDR H3 building (also ignoring canonical rules) and claimed the highest accuracy and fastest CDR H3 prediction to date (105).

Significantly, each assessment of antibody structure prediction assumes that the reference structure represents the only possible conformation of the antibody. This may be a fundamental flaw of these assessments, since individual antibodies may adopt a number of unique conformations, as discussed throughout this thesis.

1.5. Expanding the recognition potential of the germline antibody repertoire

Although it has been estimated that VDJ recombination can generate a germline antibody repertoire of approximately 10^{11} unique receptors, at any given time there are only 10^{6} circulating B-cell clones able to respond to antigen (49, 55–57, 116). Based on the finite size of the germline antibody repertoire and the reduced number of available clones, it is widely accepted that some germline antibodies must be able to recognize multiple unique antigens to match their infinite diversity (117–119).

This binding promiscuity is now generally categorized as either cross-reactivity, where chemically-related antigens make highly similar interactions to a single antibody, or polyspecificity, where chemically-distinct antigens interact uniquely with an antibody (120). These behaviours can be achieved through either binding different antigen epitopes[9] with the same antibody paratope[10], binding different epitopes with different

[9] An epitope is the region of antigen specifically interacting with antibody. A single antigen can have many different epitopes.

paratopes of a single antibody conformation, or binding different epitopes through different antibody paratopes generated by conformational flexibility.

One of the most drastic examples of antibody polyspecificity is seen among a large class of antibodies deemed 'natural antibodies' or NAbs, that are produced at a steady state in pre-immune sera and are polyreactive to large numbers of self and non-self antigens (121–123). These are largely IgM isotype but can also be IgA and IgG, and have innate-like functions in pathogen elimination and functions in homeostasis such as removing apoptotic cells. Natural antibodies may represent the extreme of antibody polyspecificity, and in one study a panel of natural antibodies were each observed to react with at least 5 of 13 antigens tested (124), suggesting an astonishing ability to bind diverse antigens. Unfortunately, there have been no structural investigations of natural antibody polyspecificity.

1.5.1. Cross-reactivity

The phenomenon of antibody cross-reactivity has been known for some time, and was quickly identified as a potential source of additional germline antibody recognition diversity (118). Antibody cross-reactivity has significant implications in health and disease and has been identified as the cause of many autoimmune disorders and allergies (125–130). The structural basis of cross-reactivity has therefore been a topic of great interest.

[10] A paratope is the region of antibody specifically interacting with antigen. A single antibody can have multiple paratopes.

Although the term cross-reactivity is sometimes used broadly to define binding by an antibody to anything other than the original epitope, it is generally meant to refer to the strict case where binding to similar antigens is achieved through the same critical antibody contacts to equivalent chemical groups. This is differentiated from the binding of multiple antigens using different paratopes or alternate chemical interactions, which is referred to as polyspecificity (131, 132).

Several structural studies of cross-reactive antibodies in complex with multiple antigens have revealed the molecular basis for the phenomenon, which is the conserved recognition of common features between the multiple antigens and further accommodation of variations or additions that do not drastically impact the core specificity.

For example, the anti-progesterone antibody DB3 bound five conformationally unique steroids in a single binding site pocket (133). Cross-reactivity was achieved through accommodation of the apolar steroid skeleton between two tryptophan side-chains, and additional specific hydrogen bonds to unique substituents among the different steroids.

In another example, the mAb Anti-p24 bound four unique peptides (131). The antigen-derived epitope peptide and an epitope-homologous peptide were similarly bound by the antibody, demonstrating cross-reactivity, whereas two unique peptides were bound by different antibody paratopes within the same antibody conformation, demonstrating polyspecificity, with each interaction through a mixture of polar and apolar contacts.

Cross-reactivity of antibodies to related carbohydrates has also been observed. For example, anti-HIV antibody 2G12 was crystallized in complex with both the carbohydrate backbone of the lipooligosaccharide from *Rhizobium radiobacter* strain Rv3, and with $Man_9GlcNAc_2$ of the high-mannose carbohydrate on HIV-1 gp120, where binding in both cases was achieved through identical recognition of a shared αMan(1→2)αMan(1→2)αMan(1→3)αMan(1→5) epitope.

One of the most thorough investigations of antibody cross-reactivity has been the structural studies of the S25-2 family of antibodies that bind Chlamydiaceae LPS oligosaccharides through conserved recognition of a single Kdo residue and further flexible accommodation of different lengths, linkages and synthetic unnatural modifications (134–142). The results of these studies are discussed in detail in section 1.8.

As cross reactivity has been observed towards proteins, carbohydrates and small molecules with specific hydrogen bonds or charged interactions and/or through less specific hydrophobic interaction, it is clear that there are no particular molecular features that allow cross-reactivity, as it can occur through a variety of interactions.

1.5.2. Polyspecificity

In contrast to cross-reactivity, which is by definition restricted to identical paratope usage to bind nearly identical epitopes, polyspecificity can occur through differential epitope and/or paratope positioning. This can include 'molecular mimicry' where unique antigens can form the same types of interactions with the same residues of the

antibody combining site (125, 143), or can involve binding to unique antigens through different interactions with the same or altered antibody paratopes.

Similar to Anti-p24 above, the germline antibody 36-65 was crystallized in complex with five unique peptides that were observed to bind in unique antibody paratopes (144, 145). Polyspecificity for unique proteins has also been observed, with the anti-hen egg lysozyme (anti-HEL) antibody D1.3 binding both HEL and the anti-idiotypic antibody E5.1 with an almost identical binding site conformation (146).

Polyspecificity toward unique classes of antigen has also been structurally characterized. For example, the antibody SYA/J6 specific for the cell surface O-antigen polysaccharide of the pathogen *Shigella flexneri Y* displayed polyspecific binding to the O-antigen and an octapeptide mimic with identical binding site conformations (147). Likewise, polyspecific binding was observed for the anti-LPS monoclonal antibody F22-4, raised against *Shigella flexneri* serotype 2a, which bound both an LPS decasaccharide and a dodecapeptide mimic in the same paratope of an identical binding site conformation (148). In another example, the antibody 2D10 was observed to bind methyl α-D-mannopyranoside and a dodecapeptide with overlapping paratopes of a single antibody binding site conformation (149).

Conformational flexibility of antibodies has also been well documented, and has been proposed as a mechanism to enable antibody polyspecificity by generating unique antibody paratopes. This concept has its roots in Linus Pauling's 1940 antigen-template theory (10, 150), long before its first kinetic observation (151). Several subsequent

studies have observed limited antibody conformational flexibility both kinetically (152–154) and structurally (130, 136, 140, 155–159).

While conformational flexibility has been observed, structural observations of polyspecificity through conformational changes have been sparse, in contrast to differential epitope and paratope positioning in single antibody binding site conformations. Examples include the germline precursor to antibody 7G12, polyspecific for its cognate immunogen N-methyl-mesoporphyrin and the polyether Jeffamine (160); antibody SPE7, polyspecific for protein antigen and hapten molecules (161, 162); antibody DNA-1, polyspecific for ssDNA and HEPES (130, 158); and mAb BBE6.12H3, polyspecific for four different peptides each with slightly different CDR H3 conformations (163).

A notable example of polyspecificity through structural flexibility is the antibody bH1, which was engineered to have dual specificity for human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor (VEGF) by randomization of the light chain CDRs of the anti-HER2 therapeutic antibody Herceptin, and selection by phage display for developing a second specificity to VEGF (164). Crystal structures of bH1 in complexes with HER2 and VEGF revealed moderate paratope overlap and a significant structural change of CDR L1 that was necessary for binding.

As with the examples of cross-reactivity discussed above, there do not appear to be particular molecular features required for polyspecificity. It has now been observed between peptides, proteins, carbohydrates, DNA, and small molecules, through polar

and apolar interactions, and through conserved or unique epitope/paratope positioning in the same or different antibody binding site conformations.

1.5.3. Conformational flexibility and affinity maturation

Conformational flexibility is intimately related to the process of affinity maturation. If a given germline antibody is flexible, then mutations that stabilize a given conformation can decrease entropic penalties of binding the complimentary ligand and increase affinity. Indeed, many studies of germline and mature antibodies have demonstrated how mutations accrued during affinity maturation reduce flexibility to generate higher affinity for specific antigens and simultaneously reduce cross-reactivity or polyspecificity (144, 165–174).

1.6. The significance of carbohydrates and anti-carbohydrate antibodies

Proteoglycans, glycolipids and glycoproteins are the most prominent types of cell-surface molecules and are integral to cell-signalling, trafficking, adhesion, differentiation, embryogenesis, spermatogenesis, angiogenesis, and fertilization (175, 176). Glycoconjugates are generated by glycosylation pathways of varying complexity involving one or more glycosyltransferases, and the breakdown of these pathways can lead to a disruption of cellular homeostasis. One mechanism credited with the prevention of associated diseases is B-cell surveillance of cell surface oligosaccharides that may be altered as a result of this disruption, and elimination of abnormal cells. Some of these aberrant glycosylations are known tumour-associated-antigens (TAA) and have been well-studied for a variety of cancers (177–179). Human antibodies are

known to be capable of recognizing many of these unusual TAAs, and there is a substantial research focus in the generation of TAA-conjugate vaccines to stimulate immune responses to various cancers.

The prevalence of carbohydrate structures on the surfaces of healthy cells leads to a degree of immune tolerance, and infectious agents can sometimes use these same structures to mask their antigenic surface proteins to evade immune surveillance (127, 180, 181). However, many antibodies to carbohydrate antigens are remarkably specific and a significant protective response against most pathogenic bacteria is still achieved through the generation of antibodies to LPS or capsular polysaccharide (182).

The ability of antibodies to distinguish between closely-related antigens is exemplified in some transfusion mismatches of the human ABO(H) blood group. The antibody response to the foreign blood group antigen is often so severe as to result in fatality, yet the human A and B blood group trisaccharide antigens are nearly identical and differ only in the substitution of a hydroxyl group for an acetamido group on the terminal sugar (183). It is this potential to distinguish between closely related antigens that drives the development of carbohydrate-specific antibodies in diagnostic medicine.

Although many anti-carbohydrate antibodies do display high-specificity, others are known to be highly cross-reactive or polyspecific. This can pose serious problems, as the cross-reactivity of some anti-carbohydrate antibodies is associated with autoimmune disorders that are triggered when infectious organisms display an immunogen with structural similarities to self-antigens (143, 184). This can occur

despite the screening of B-cells during development, when B-cells that recognize self-antigens normally undergo apoptosis (18, 55).

Aside from their sometimes deleterious side-effects, cross-reactivity and polyspecificity are thought to play a key role in expanding the recognition potential of the germline antibody repertoire as discussed above, and are largely responsible for the success of the antibody response to carbohydrates, discussed below.

1.7. The antibody response to carbohydrate antigens

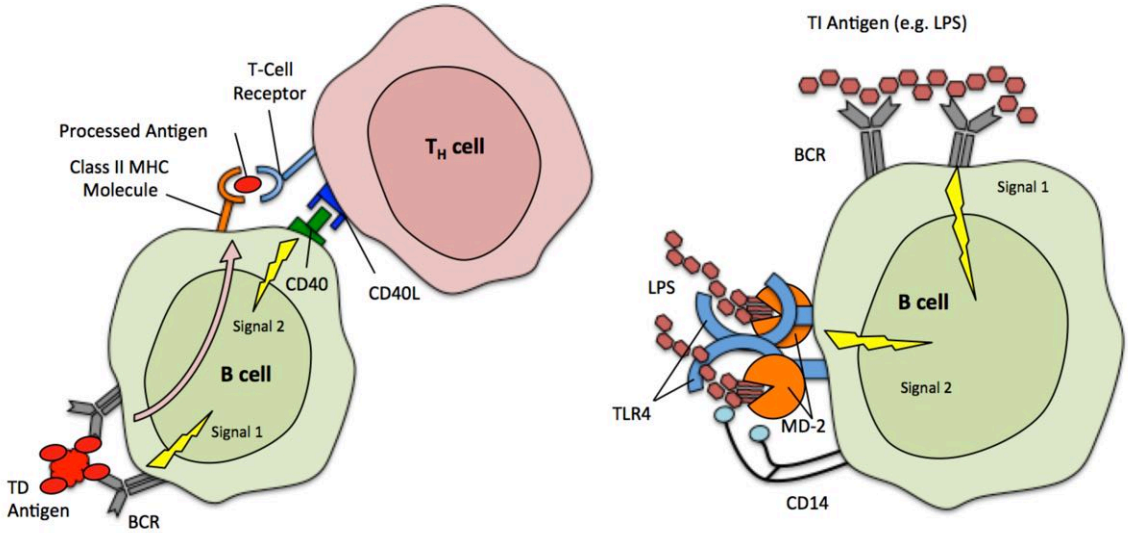
As described in section 1.3, the diversity of the antibody response arises from V(D)J recombination in developing B-cells, and mature circulating B-cells display copies of single BCRs with defined specificity. Circulating B-cells are normally dormant but can be activated upon stimulation by immunogen and co-stimulation by T-helper cells, where they then migrate to peripheral lymphoid organs to undergo somatic hypermutation and affinity maturation.

In special circumstances, B-cell activation can occur without co-stimulation of T-helper cells upon stimulation by thymus-independent (TI) antigens. TI antigens, including the bacterial cell-wall components LPS and capsular polysaccharides or polymeric protein antigens (18, 185, 186), activate B-cells by cross-linking BCRs or through concomitant stimulation of the BCR and the Myeloid differentiation factor-2 (MD-2)/Toll-like Receptor 4 (TLR4) receptor complex (Figure 8). However, the overall humoral response to TI antigens is typically weaker than that to thymus-dependent (TD) antigens, with no generation of memory cells, affinity maturation or class-switching.

The majority of carbohydrate antigens are TI and do not by themselves induce significant affinity maturation (186–188). This suggests that recognition of carbohydrates from common pathogens should be conserved in germline gene segments, as this response has evolved not to rely upon affinity maturation. Furthermore, as there are a limited number of germline gene segments available for the generation of antibodies to all potential antigens, as discussed in section 1.5, evolutionary pressure would select for those gene segments that both protect against common pathogens and remain able to respond to novel threats. Therefore, anti-carbohydrate antibodies provide excellent models for studying the molecular basis of inherited immunity combined with adaptability.

Figure 8: B-cell activation through TD or TI antigens

B-cell activation, proliferation and differentiation occur in response to antigen and may occur in a T-cell independent (TI) or T-cell dependant (TD) manner. TD responses involve the processing and presentation of antigen by B-cells to T-cells and require direct contact between the two. Some antigens such as LPS or capsular polysaccharide may activate B-cells in a TI manner by cross-linking BCRs and interacting with the TLR-4/MD-2 complex.



1.8. Structural studies of antibodies against Chlamydiaceae LPS

Lipopolysaccharide is a highly immunogenic conserved building block of the Gram-negative outer membrane, and can be exploited as an excellent probe of the antibody response to TI carbohydrate antigens (186–189). It is present in the order of 10^6 copies per bacterium, and is crucial for the structural integrity of the membrane and for blocking serum components such as the membrane attack complex (189). Lipopolysaccharides from Enterobacteria are the prototypical example of bacterial endotoxin, and are large molecules generally divided into three components. The lipid A anchor is a β1–6 linked glucosamine disaccharide with each residue acylated on C2 and C3 to embed in the bacterial outer membrane. Attached to lipid A is a short chain of sugars called the core oligosaccharide, which is subdivided into the inner and outer core and can vary significantly between bacterial species. Last is the O-antigen or the O-polysaccharide, which is a repeating oligosaccharide attached to the outer core that varies significantly among bacterial strains.

Under natural conditions, a functional outer membrane in Gram-negative bacteria contains at least Lipid A with an additional (2→4) linked disaccharide of 3-deoxy-α-D-manno-oct-2-ulosonic acid (Kdo), or a single Kdo phosphorylated in position 4 or 5. In some species such as *Burkholderia cepacia* and certain strains of *Acinetobacter*, one of the Kdo residues may be substituted with the isosteric D-glycero-D-talo-oct-2-ulosonic acid (Ko) (190–192).

Chlamydiaceae is a bacterial family containing two genera, *Chlamydia* and *Chlamydophila*, with a total of nine species representing a range of human and animal

pathogens (193–195). The LPS of this family displays an unusual truncated LPS of only Lipid A and Kdo, consisting of the family-specific oligosaccharide Kdo(2→8)Kdo(2→4)Kdo linked (2→6) to the Lipid A GlcN4*P*(1→6)GlcN1*P*, with *Chlamydophila psittaci* also displaying Kdo(2→4)Kdo(2→4)Kdo and the species-specific branched oligosaccharide Kdo(2→4)[Kdo(2→8)]Kdo(2→4)Kdo (Figure 9) (192, 196–201). The antibody response to the family-specific antigen is the basis of a diagnostic test for Chlamydophila pneumoniae infection in humans.

Although most carbohydrates are TI antigens and therefore do not elicit affinity maturation or class-switching responses, these can be induced by conjugating the antigen to a protein or peptide (202–204). This is desirable to obtain IgG for structural studies, as IgM is difficult to work with in the laboratory and enzyme digest will produce the Fv fragment in low yield.

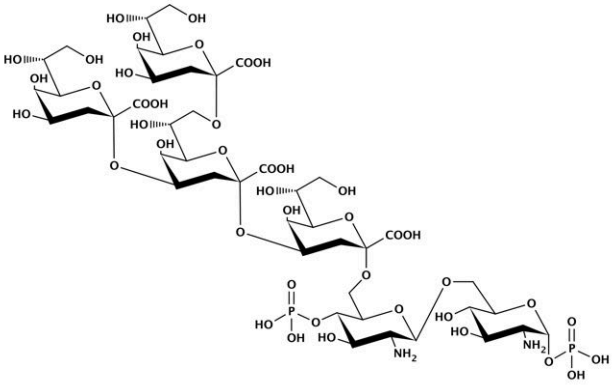
Although inducing an affinity-maturation response to study germline antibodies may seem counterintuitive, much useful information about the germline response can be obtained through this technique. First, the higher affinity of the IgG will better allow co-crystallization with the antigens of interest. Second, the class-switching that accompanies affinity maturation allows the production of large quantities of IgG that can be digested to yield the Fab fragment in high yield. Third, murine germline genes are well defined, and analysis of the antibody sequence of a successful structure determination of an antibody–antigen complex will still reveal the likely germline interactions.

Glycoconjugates of oligosaccharides containing various chlamydial Kdo-based LPS epitopes have proven useful in the immunization of mice for generating a range of antibodies with varying specificities and affinities for Kdo and Ko containing antigens (197, 198, 205–213). An example of the glycoconjugates used to generate these antibodies is shown in Figure 9, and representative lists of the antigens used to raise antibodies and to test their specificities to chlamydial LPS are presented in Table 9 and Table 10. A variety of structures were tested to investigate antibody specificity, including synthetic antigens not found in nature, to probe the limits of cross-reactivity (Figure 10).

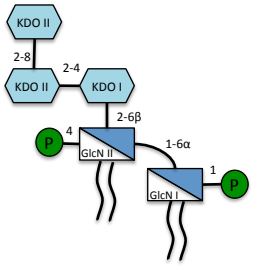
The sizeable panel of antibodies that was developed displayed a range of specificities (Table 10). Some bound the Kdo(2→4)Kdo glycosidic linkage preferentially over the Kdo(2→8)Kdo linkage and vice versa. Some would only bind antigens of one or two carbohydrate residues, while others exclusively bound larger antigens. Some were highly specific for a single epitope, while some were cross-reactive for several distinct Kdo epitopes. Together, these antibodies have provided an unparalleled opportunity to explore germline recognition of carbohydrate antigens and to analyze the effects of specific mutations on binding.

Figure 9: Chlamydiaceae LPS

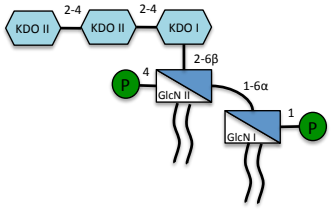
(A) Shows the detailed chemical structure of Chlamydophila psittaci species-specific LPS consisting of the branched oligosaccharide Kdo(2→4)[Kdo(2→8)]Kdo(2→4)-Kdo(2→6)GlcN4P(1→6)GlcN1P. **(B)** – **(D)** are simplified schematics of the Chlamydiaceae family-specific LPS oligosaccharide Kdo(2→8)Kdo(2→4)Kdo and the *Chlamydophila psittaci* species-specific oligosaccharides Kdo(2→4)Kdo(2→4)Kdo and Kdo(2→4)[Kdo(2→8)]Kdo(2→4)Kdo. **(E)** shows an example glycoconjugate used in immunizations to generate the S25-2 type antibodies.



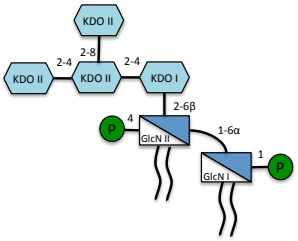
A



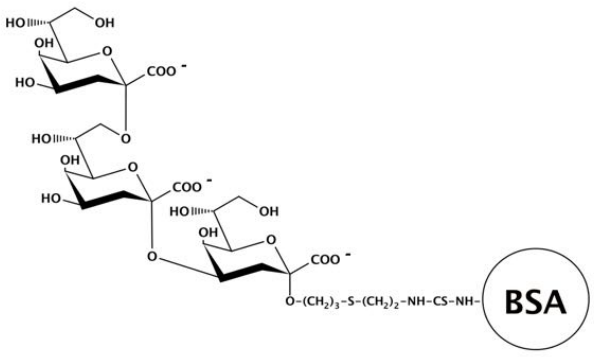
B



C



D



E

Figure 10: Kdo antigens crystallized in complex with S25-2

Various natural and synthetic antigens were used to probe the cross-reactivity of S25-2: **(a)** Kdo(2→4)Kdo(2→4)Kdo, **(b)** Ko, **(c)** Ko(2→4)Kdo, **(d)** 7-epi-Kdo, **(e)** 3,4-dehydro-3,4,5-trideoxy-Kdo(2→8)Kdo, **(f)** 5-deoxy-4-epi-2,3-dehydro-Kdo(4→8)Kdo, and **(g)** Kdo(2→4)KdoC1red

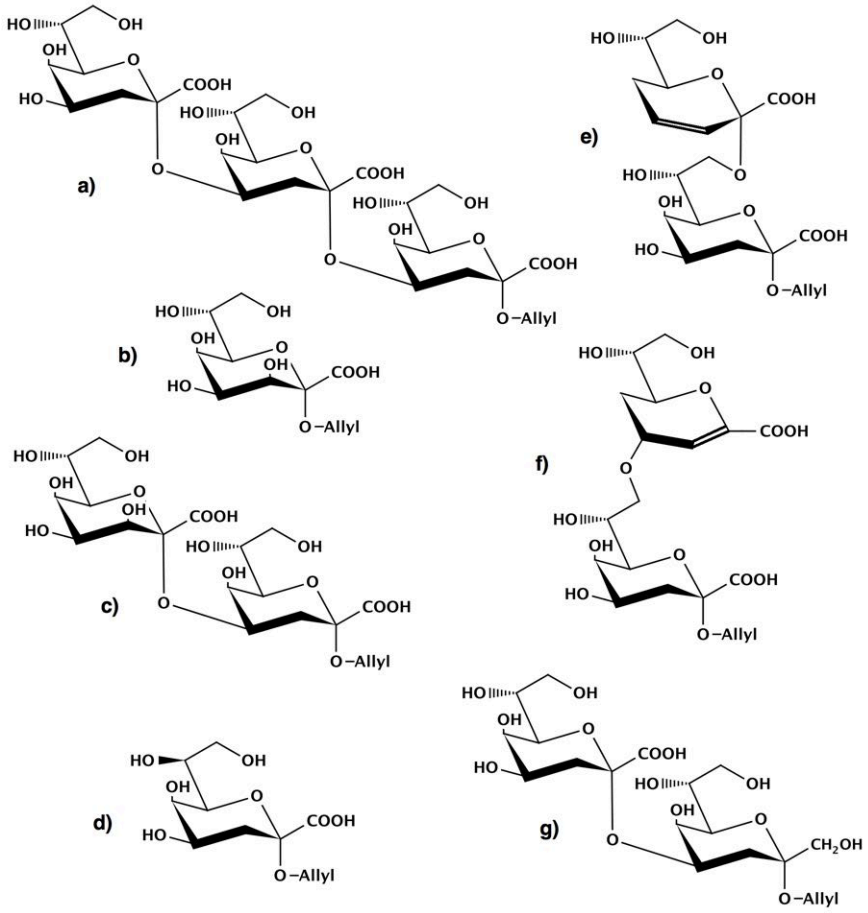


Table 9: Germline gene segment usage of antibodies raised against Kdo oligosaccharide glycoconjugates

Antibodies generated towards Ko- and Kdo-based LPS antigens display a conserved germline gene segment usage despite their generation against distinct antigens, evidencing the presence of a Kdo/Ko binding motif evolutionary conserved in the germline.

		Light Chain (κ)		Heavy Chain		
Clone	**Immunogen (conjugated to BSA)**	**V gene**	**J gene**	**V gene**	**D gene**	**J gene**
S25-2	**Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcNAc**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD2-9*01	IGHJ3*01
S25-39	**Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcNAc**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD2-3*01	IGHJ3*01
S45-18	**Kdo(2→4)Kdo(2→4)Kdo**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD1-1*01	IGHJ4*01
S54-10	**Kdo(2→8)[Kdo(2→4)]Kdo(2→4)Kdo(2→6)βGlcN(1→6)αGlcN**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD2-14*01	IGHJ4*01
S69-4	**Kdo(2→8)[Kdo(2→4)]Kdo(2→4)Kdo(2→6)βGlcN(1→6)αGlcN**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD2-4*01	IGHJ4*01
S67-27	**Ko(2→4)Kdo**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD1-1*02	IGHJ3*01
S73-2	**Kdo(2→8)[Kdo(2→4)]Kdo**	IGKV8-21*01	IGKJ1*01	IGHV7-3*02	IGHD2-3*01	IGHJ4*01

Table 10: Relative binding avidities of S25-2 type antibodies for Chlamydiaceae oligosaccharides determined by ELISA

Antibodies generated by immunization of BALB/c mice with LPS glycoconjugates displayed a variety of binding profiles with varying specificities when tested by ELISA with immobilized glycoconjugates of LPS antigens; some mimics of naturally occurring chlamydial LPS oligosaccharides and other unnatural variations of these epitopes.

	mAb conc. (ng/mL) yielding OD_{405} >0.2 using 2 pmol/well							
Antigen	**S25-2**	**S25-39**	**S45-18**	**S54-10**	**S69-4**	**S73-2**	**S67-27**	**S64-4**
Ko	>1000	>1000	>1000	>1000	>1000	>1000	500	>1000
Ko(2→4)Kdo	>1000	>1000	63	500	>1000	250	63	>1000
Kdo	500	500	16	150	>1000	250	16	>1000
Kdo(2→4)Kdo	500	125	2	63	>1000	8	16	>1000
Kdo(2→4)Kdo(2→4)Kdo	1000	63	0.5	4	16	4	16	>1000
Kdo(2→4)Kdo(2→4)Kdo(2→6)βGlcNAc	1000	32	1	4	16	4	16	>1000
Kdo(2→4)Kdo(2→4)Kdo(2→6)βGlcN4P(1→6)αGlcN	>1000	63	1	8	250	8	16	500
Kdo(2→8)Kdo	250	16	4	250	>1000	250	16	>1000
Kdo(2→8)Kdo(2→4)Kdo	32	8	8	250	>1000	16	16	1000
Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcNAc	63	8	125	1000	>1000	500	63	>1000
Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcN4*P*(1→6)αGlcN1*P*	63	16	1000	>1000	>1000	1000	1000	2
Kdo(2→8)[Kdo(2→4)]Kdo(2→4)Kdo	>1000	32	0.5	4	16	4	16	>1000
Kdo(2→8)[Kdo(2→4)]Kdo(2→4)Kdo(2→6)βGlcN4P(1→6)αGlcN	>1000	125	1	8	1000	16	32	250
Kdo(2→8)[Kdo(2→4)]Kdo(2→4)Kdo(2→6)βGlcN4P(1→6)αGlcN1P	>1000	32	0.5	4	63	4	16	>1000

1.8.1. V region restriction of anti-Chlamydiaceae antibodies

The germline genes from which an affinity-matured antibody has descended can be elucidated by a comparison of the nucleotide sequence of the antibody with those of the germline genes (50). The structure of the antibody in complex with antigen can then show which conserved residues in contact with antigen were likely responsible for the germline interaction, and which residues have been mutated presumably to increase antigen affinity.

The repeated utilization of a particular set of germline genes in response to a particular class of carbohydrate antigen is a common phenomenon known as V-region restriction. It has been observed for both the human and murine antibody response to capsular polysaccharides from *Haemophilus influenzae* (214, 215), *Streptococcus pneumoniae* (216), *Cryptococcus neoformans* (217, 218), *Neisseria meningitidis* (219), and *C. neoformans glucuronoxylomannan* (220). V-region restriction has been hypothesized to be a result of the limited epitope diversity of polysaccharides, being repeating units of short oligosaccharide epitopes (221).

A large proportion of antibodies raised against chlamydial LPS display V-region restriction as demonstrated by their shared germline gene segment usage (Table 9). Remarkably, this is true even for those antibodies raised using chemically and stereochemically distinct immunogens. This redundant usage of germline gene segments in response to Kdo-based immunogens suggests an evolutionary conservation of a combining site or sub-site optimized for Kdo recognition, accentuating the importance of Kdo as an antigenic marker with which the immune system has co-

evolved to provide inherited immunity. Interestingly, several studies have concluded that extensive somatic hypermutation can generate a large measure of paratope diversity even from a limited gene usage (221–224). V-region restriction appears to be an evolved strategy to provide a mechanism of broad recognition of important related carbohydrate antigens that places minimal strain on the size of the germline gene repertoire, and acts as a starting point that can mature towards specific recognition of particular epitopes of these related antigens.

One particular combination of heavy and light chain V genes repeatedly appears in response to different chlamydial LPS immunogens (Table 9), that nevertheless results in antibodies that display a wide range of specificities. As the V genes provide a largely common L1, L2, L3, H1 and H2, the differences in avidity and specificity shown by these antibodies can be attributed first to the different D and J genes that code for H3, and second to mutations arising through affinity maturation (Table 11).

The first of this group of antibodies to be structurally characterized was the archetypical S25-2 (134, 140), which is the closest of the group to germline in sequence with only three V-gene mutations. This antibody was raised against Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcNAc-BSA and displays its highest avidity for Kdo antigens with a (2→8) terminal linkage, and also showed weak cross-reactivity for a range of other antigens (Table 10). The crystal structure of S25-2 was solved in complex with a variety of ligands representing natural chlamydial epitopes as well as modified unnatural Kdo ligands. These structures provided the first glimpse into a recurring theme for the recognition of Kdo-containing antigens by S25-2 type antibodies, which

involves a highly conserved Kdo monosaccharide binding pocket combined with flexible recognition of alternate lengths, linkages, and synthetic modifications.

Table 11: CDR sequence alignment of S25-2 type antibodies

Amino acid alignment of S25-2 type antibody CDRs to germline sequence. Differences from germline sequence are underlined, and the Asn H53 to Lys H53 mutation (discussed in section 1.8.4) is shown in bold. Colour-coded germline gene segments are aligned above the sequences to indicate amino acid origin.

Clone	CDR L1	CDR L2	CDR L3	CDR H1	CDR H2	CDR H3
IGKV8-21*01 and IGHV7-3*02	QSLLNSRTRKNYLA	WASTRES	CKQSYNL	GFTFTDYYMS	FIRNKANGYTTEYSAS	ARD---------
S25-2	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKANGYTTEYSPS	ARDHDGYYE
S25-39	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKA**K**GYTTEYSAS	ARDHDGYYE
S45-11	QSLLNSRTRKSYLA	WAATRES	CKQSYNL---	GFTFTDYYMS	FIRNKP**K**GYTTEYSAS	VRDIYSFGSRD
S45-18	QSLLNSRTRKSYLA	WAATRES	CKQSYNL---	GFTFTDYYMS	FIRNKP**K**GYTTEYSAS	VRDIYSFGSRD
S45-24	QSLLNSRTRKNYLA	WASTRDS	CKQSYTL---	GFTFTDYYMS	FIRNKA**K**GYTTEYSAS	ARDDYDYPYY
S73-2	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKA**K**GYTTEYSAS	ARDINPGSDGYYD
S54-10	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKV**K**GYTIDYSAS	ARDMRRFDDGD
S54-13	QSLLNSRTRKNFLA	WASTRES	CKQFYSL---	GFTFTEYYMS	FIRNKT**K**GYTTEYSTS	ARDKHFGSRD
S54-27	QSLLHSSNQKNYLA	WASTRES	CQQYYRY---	GFTFTDSYMS	FIRDKPNGYTTEYSAS	TRDSRYY
S54-30	QSLLHSSYQKNYLA	WASTRES	CQQYYRY---	GFTFTDYYMS	FIRNKANGYTIEYSAS	ARDTRYY
S69-4	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMG	FIRNKA**K**GYTTEYSAS	ARDLIYFDYDD
S46-5	QSLLNSRTRKNNLA	WASTREF	CKQSSNL---	GFTFTDYYMS	FIRNKANGYTTEYSAS	ARDVDGNYVE
S46-8	QSLLNSRTRKNNLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKPNGYTTEYSVS	TRDVDFNYVE
S61-27	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKA**K**GYTTEYSAS	ARDIITTGVAPHYS
S68-5	QSLFHSRTRKNHLA	WASTRES	CKQSYSL---	GFTFTDYYMS	FIRNRANFYTTEYSAS	ARDSDSYPV
S68-12	QSLLNSRTRKSYLA	WAATRES	CKQSYNL---	GFTFTDYYMS	FIRNKANFYTTEYSAS	ARDSDTYPV
S67-27	QSLLNSRTRKNYLA	WASTRES	CKQSNNL---	GFTFTDYYMS	FIRNKA**K**GYTTEYSAS	ARDISPSYGVYYE
S25-27	QSLLNSRTRKSYLA	WASTRES	CKQSYNL---	GLTFTDYYMS	FIRNKANGYTTEYSAS	ARDHDGYYE
S25-37	QSLLNSRTRKNYLA	WASTRES	CKQSYNL---	GFTFTDYYMS	FIRNKA**K**SYTTEYSAS	TRDHDGYYE
S25-38	QSLLNRRTRKNYLA	WASTRES	CKQSNNL---	GFTFSDFYMS	FIRNRVNGYTTEYSAS	ARDIGYYE
S23-24	QSLLNRRTRKNYLA	WASTRES	CKQSNNL---	GFTFSDFYMS	FIRNRVNGYTTEYSAS	ARDIGYYE

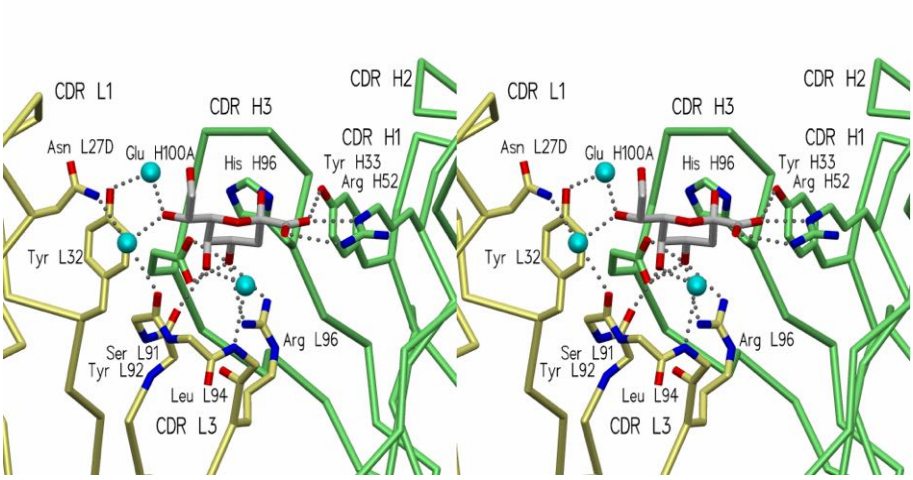
1.8.2. Conserved recognition of Kdo

PDB codes for all crystal structures discussed are given in Table 12

The Kdo monosaccharide binding pocket observed in the S25-2 structures is composed exclusively of germline residues from CDRs H1, H2, L1 and L3, with additional interactions from CDR H3 (Figure 11). The interactions formed with ligand by this pocket consist largely of hydrogen bonds and VDW forces and include a salt bridge from Arg H52 to the Kdo carboxyl group, which was the first observation of a charged residue interaction in anti-carbohydrate antibodies (140, 225). As in other antibody-antigen complexes, a number of coordinated water molecules also play a role in maximizing surface complementarity in the S25-2 structures. Although buried water molecules have been observed in other carbohydrate–antibody complexes (226), none of the waters described above are completely isolated from the solvent shell.

Figure 11: Kdo binding by S25-2

The binding site of S25-2 in complex with Kdo. Hydrogen bonds between S25-2 and Kdo are indicated by dashed black spheres. The S25-2 backbone is shown as an α-carbon trace unless backbone atoms participate in hydrogen bonds, in which case they are shown. The S25-2 heavy chain is coloured green and the light chain yellow, and water molecules are displayed as cyan spheres[11].



[11] Many figures in this thesis are presented in stereoview (all wall-eyed) for better clarity.

Table 12: PDB codes for crystal structures of S25-2 type antibodies

All PDB files are available from http://www.rcsb.org

Antibody	Antigen	PDB Code
S25-2	Unliganded	1Q9K, 1Q9L
	Kdo	1Q9V
	Kdo(2→8)Kdo(2→4)Kdo	1Q9Q
	Kdo(2→4)Kdo	1Q9T
	Kdo(2→8)Kdo	1Q9R
	Kdo(2→4)Kdo(2→4)Kdo	2R2B
	Ko	2R2H
	Ko(2→4)Kdo	2R23
	7-epi-Kdo	2R2E
	KdoC1red(2→4)Kdo	2R1Y
	3,4-Dehydro-3,4,5-trideoxy-Kdo(2→8)Kdo	2R1W
	5-Deoxy-4-epi-2,3-dehydro-Kdo(4→8)Kdo	2R1X, 3BPC
S67-27	Ko	3IJH
	Kdo(2→8)Kdo	3IJY
	Kdo(2→8)-7-O-Me-Kdo	3IKC
	Kdo(2→4)Kdo(2→6)GlcN4*P*(1→6)GlcN1*P*	3IJS
S73-2	Kdo	3HZM
	Kdo(2→4)Kdo	3HZK
	Kdo(2→4)Kdo(2→4)Kdo	3HZY
	Kdo(2→8)Kdo(2→4)Kdo	3HZV
S54-10	Kdo(2→4)Kdo(2→4)Kdo	3I02
S25-39	Unliganded	3OKM
	Kdo	3OKD
	Ko	3OKE
	Kdo(2→4)Kdo	3OKK
	Kdo(2→8)Kdo	3OKL
	Kdo(2→4)Kdo(2→4)Kdo	3OKN
	Kdo(2→8)Kdo(2→4)Kdo	3OKO
S45-18	Unliganded	1Q9O
	Kdo(2→4)Kdo(2→4)Kdo(2→6)GlcN4*P*(1→6)GlcN1*P*	1Q9W
S64-4	Kdo	3PHQ
	Kdo(2→4)Kdo(2→4)Kdo(2→6)GlcN4*P*(1→6)GlcN1*P*	3PHO

1.8.3. Cross-reactivity for alternate lengths and linkages

Antigens of alternate lengths and linkages are accommodated in the S25-2 binding site through highly similar recognition of a terminal Kdo in the conserved binding pocket, with the remainder of each antigen accommodated by interactions with Arg L27F of CDR L1 and Asn H53 of CDR H2, which form a groove above the conserved pocket (Figure 12).

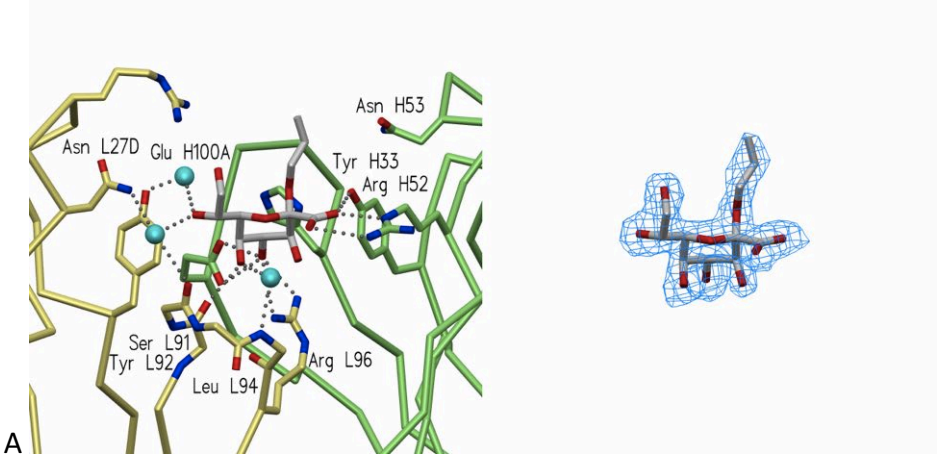
The terminal (2→4) and (2→8) glycosidic linkages of the disaccharide and trisaccharide ligands are accommodated differently by S25-2, with the terminal (2→4) linkage less favourably oriented and showing fewer interactions. Correspondingly, ELISA and SPR data show that S25-2 has higher avidity for sugars with terminal (2→8) linkages than (2→4), and significantly higher avidity for Kdo(2→8)Kdo(2→4)Kdo over Kdo (2→4)Kdo(2→4)Kdo (Table 10 and Table 13). This is also evident in the quality of the electron density maps for the two trisaccharides in the combining site, where the Kdo(2→4)Kdo(2→4)Kdo trisaccharide shows some disorder (Figure 12).

The ability of the near germline antibody S25-2 to cross-react with distinct epitopes primarily through the action of a monosaccharide binding pocket of conserved sequence appears to be an evolved strategy for general recognition of Kdo-containing foreign antigens, which is especially apparent in the recognition of modified and unnatural Kdo ligands. Modified monosaccharide ligands that bind S25-2 include Ko and the 7-epi-Kdo diastereomer. Both of these ligands show slight changes in binding mode albeit with minor (7-epi-Kdo) or significant (Ko) decreases in avidity (Table 10). Disaccharides with modifications of non-terminal Kdo residues such as Kdo

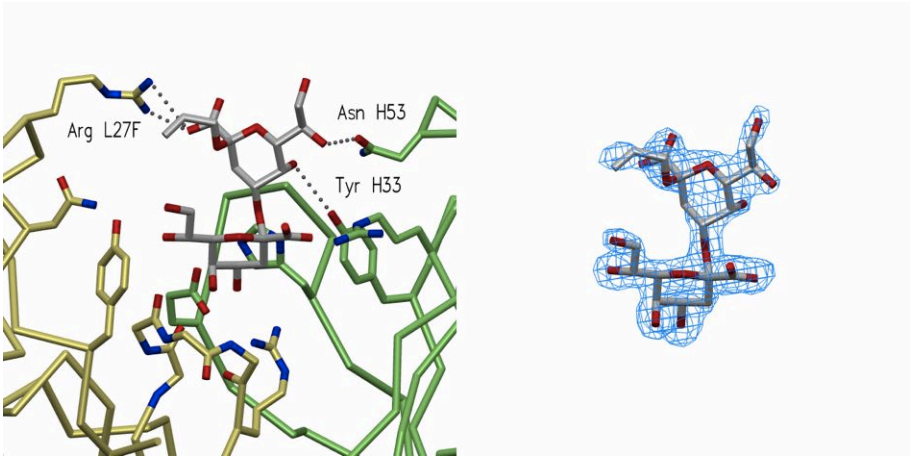
(2→4)KdoC1red are also accommodated by the flexible groove as long as a terminal Kdo is present to bind in the monosaccharide pocket. Compounds with significantly altered terminal sugars were observed to bind in an "upside-down" orientation, so instead of the modified Kdo occupying the monosaccharide binding pocket the second unmodified Kdo is bound. Although the Kdo monosaccharide binding pocket observed in the first structures of S25-2 bound to the natural antigens was initially hypothesized to offer specific recognition by a germline antibody of a single sugar residue on a foreign antigen from which highly-specific antibodies could be developed by affinity maturation (140), the pocket turned out to be surprisingly adaptable and would recognize modified and non-terminal Kdo residues (134).

Figure 12: S25-2 complexes with Kdo antigens

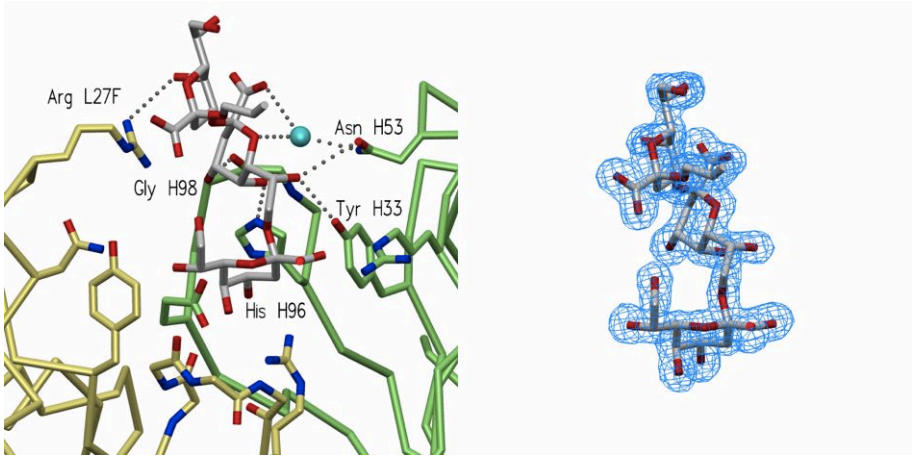
Left panels show binding site interactions from crystal structures of S25-2 in complex with various antigens, and right panels show 2Fo-Fc electron density maps contoured around antigens to 1.0σ. **(A)** Ko, **(B)** Kdo(2→8)Kdo, **(C)** Kdo(2→4)Kdo, **(D)** Kdo(2→8)Kdo(2→4)Kdo, **(E)** Kdo(2→4)Kdo(2→4)Kdo, **(F)** Kdo(2→4)KdoC1red, **(G)** 5-deoxy-4-epi-2,3-dehydro-Kdo(4→8)Kdo and **(H)** 3,4-dehydro-3,4,5-trideoxy-Kdo(2→8)Kdo. Other than in (A), interactions to terminal Kdo residues are identical to those illustrated in Figure 11 and so are not shown.



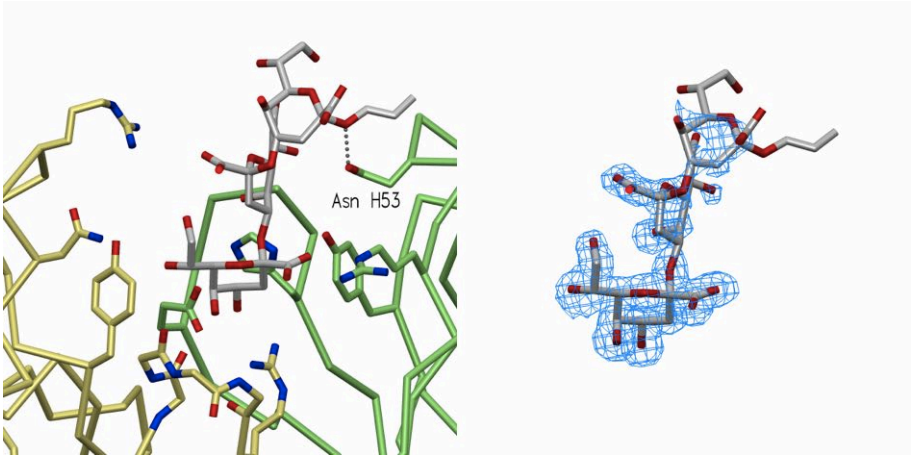
C


Arg L27F
Asn H53
Tyr H33

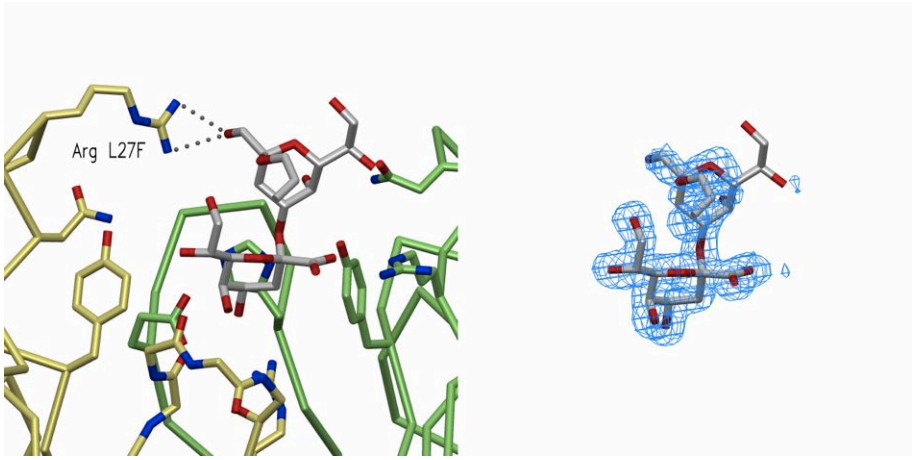
D


Arg L27F
Asn H53
Gly H98
Tyr H33
His H96

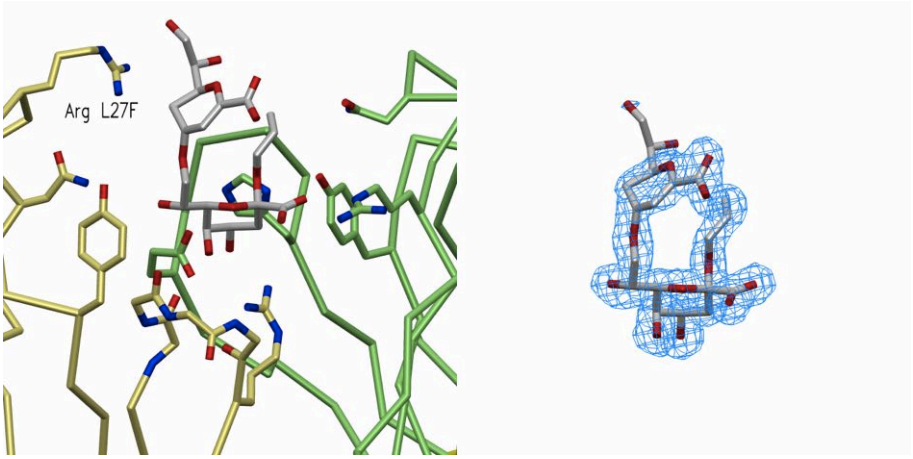
E


Asn H53

F



G



H

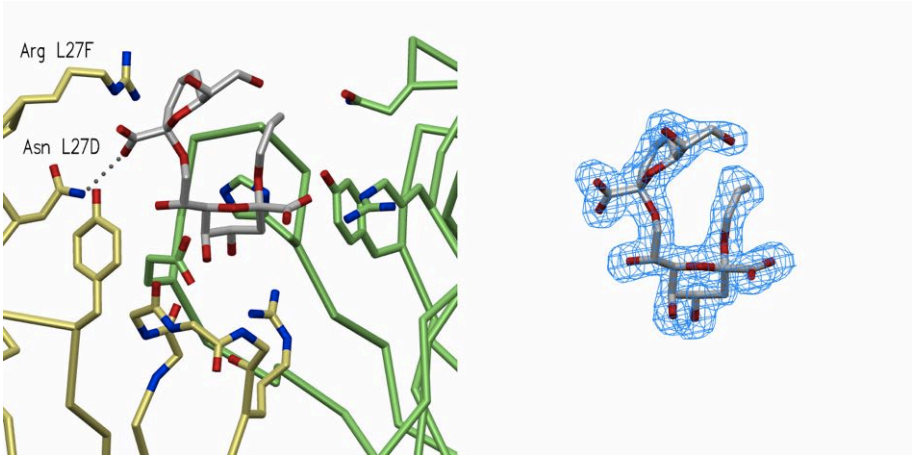


Table 13: Binding affinities of S25-2 and S67-27 to select Kdo oligosaccharides determined by SPR

K_Ds of S25-2 and S67-27 for Kdo antigens were determined by SPR as described in (134, 135).

Antibody/Antigen	K_D ($x10^{-6}$ M)
S25-2	
Kdo	15
Kdo(2→8)Kdo	1.8
Kdo(2→8)Kdo(2→4)Kdo	0.6
Kdo(2→4)Kdo	1.1
Kdo(2→4)Kdo(2→4)Kdo	63
Kdo(2→4)KdoC1red	31
Ko(2→4)Kdo	190
KdoC1red(2→4)Kdo	290
3,4-Dehydro-3,4,5-trideoxy-Kdo(2→8)Kdo	25
5-Deoxy-4-epi-2,3-dehydro-Kdo(4→8)Kdo	16
S67-27	
Kdo	3.5
Kdo(2→8)Kdo	0.91
7-O-Me-Kdo(2→8)Kdo	0.035

1.8.4. Simple mutations significantly improve avidity towards Kdo antigens

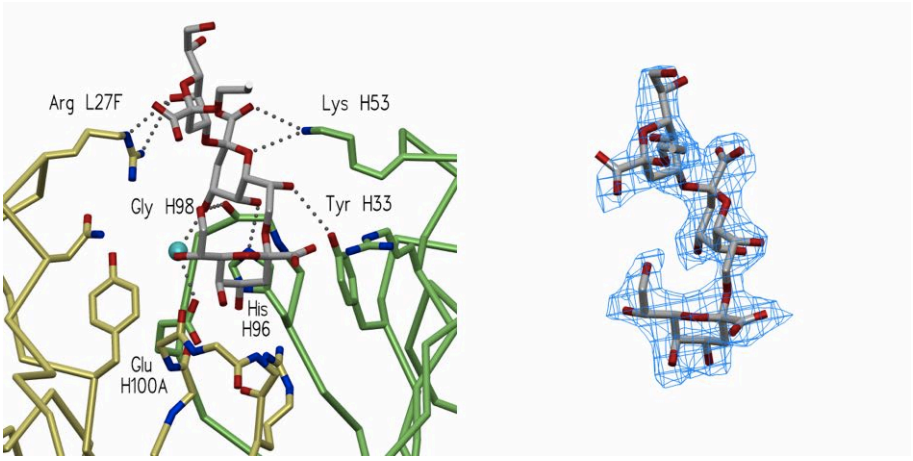
The mAb S25-39 descends from the same heavy and light chain V and J genes as S25-2, and has an identical CDR H3 with the exception of residue H101 that is Ser in S25-2 and Ala in S25-39. H101 is at the base of CDR H3 away from the binding site. There is a single difference in the combining site of these antibodies, where S25-2 possesses Asn H53 in CDR H2 and S25-39 possesses Lys (Table 11). Crystal structures of S25-39 in complex with various Kdo oligosaccharides reveal that Lys H53 forms additional hydrogen bonds and/or charged residue interactions with the second Kdo residue of every oligosaccharide antigen (Figure 13). These observations are consistent with S25-39 having a similar specificity as S25-2 but generally higher avidity for all antigens (Table 10).

The NH53K mutation is the result of a single nucleotide substitution in the germline sequence from T to A that alters the codon from AAU (Asn) to AAA (Lys). This mutation was found in 11 of the 21 S25-2 type antibodies with known sequence (Table 11). The prevalence and profound benefit to binding of this mutation among the S25-2 type antibodies emphasizes the potential of the germline Kdo monosaccharide pocket as a starting point for affinity maturation. However, the preservation of Asn H53 in the germline may be due to its higher bonding versatility (can both donate and accept two hydrogen bonds), which may better enable cross-reactivity or polyspecificity compared to Lys H53 (which can only donate hydrogen bonds and/or form a charged-residue interaction). Furthermore, the accessibility of the AAU (Asn) to AAA (Lys) mutation would not exert a strong evolutionary pressure for its permanent incorporation.

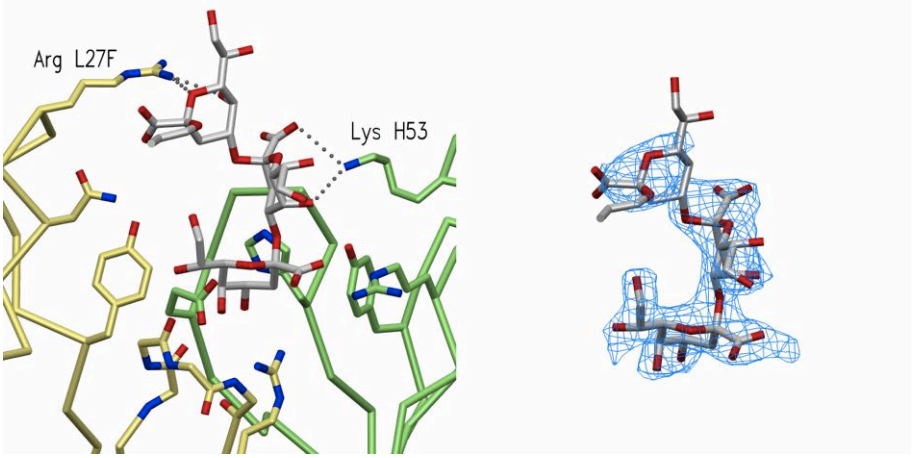
Figure 13: S25-39 in complex with Kdo oligosaccharides

Left panels show binding site interactions from crystal structures of S25-39 in complex with **(A)** Kdo(2→8)Kdo(2→4)Kdo and **(B)** Kdo(2→4)Kdo(2→4)Kdo, and right panels show 2Fo-Fc electron density maps contoured around antigen to 1.0σ. Interactions to terminal Kdo residues are identical to those illustrated in Figure 11 and are not shown.

A



B



1.8.5. Alternate CDR H3 changes the specificity of V-gene pairs

CDR H3 is well known to influence the specificity of antibodies, as many unique loops arise from the hypervariable VDJ interface (174, 227, 228). One landmark study revealed that switching the CDR H3 of otherwise identical IgM molecules can dramatically alter their specificities for haptens and protein antigens (227). The study of the S25-2 type antibodies reveals that not only are the unique interactions via CDR H3 important in most antigen recognition scenarios, but the shape of CDR H3 has a significant influence in directing the antigen binding site toward specificity or cross-reactivity.

The mAb S45-18 shares the same heavy and light chain V genes as S25-2, but possesses different D and J genes to create a unique CDR H3 (Table 11). Whereas S25-2 was raised against an antigen with a (2→8) terminal linkage, S45-18 was raised against the Kdo(2→4)Kdo(2→4)Kdo trisaccharide and shows higher avidity for terminal Kdo(2→4) ligands. The differences in binding specificities between S45-18 and S25-2 (Table 10) can be attributed to the different CDR H3 sequence and conformation.

The structure of S45-18 in complex with Kdo(2→4)Kdo(2→4)-Kdo(2→6)GlcN4*P*(1→6)GlcN1*P* shows that the terminal Kdo residue is recognized by the same monosaccharide binding pocket as in S25-2. However, while S25-2 has a comparatively short CDR H3 and an open combining site, the longer CDR H3 of S45-18 bends inwards to form a more restricted pocket (Figure 14). It possesses a key phenylalanine residue at position H99 that protrudes into the combining site to form favourable stacking interactions with the Kdo(2→4)Kdo(2→4)Kdo antigen. These

interactions are not available with the alternate structure of Kdo(2→8)Kdo(2→4)Kdo, and so binding to this antigen is less favoured (Table 10).

The mAb S54-10 shows strikingly similar specificity to S45-18, but with a different CDR H3 sequence. The structure of S54-10 in complex with Kdo(2→4)Kdo(2→4)Kdo reveals an identical CDR H3 Cα-backbone conformation as S45-18 despite its different sequence, which also contains Phe H99 that forms stacking interactions with terminal Kdo(2→4) ligands and impedes binding of terminal Kdo(2→8) ligands (Figure 15, Table 10). The presence in the germline of multiple DJ combinations that give rise to similar binding motifs suggests a survival advantage conferred by a redundant repertoire specific for pathogens displaying Kdo oligosaccharides.

The mAbs S73-2 and S67-27 possess CDR H3 loops that are longer and of different sequence than S45-18 or S54-10 (Table 11). Structural studies showed that they lean away from the Kdo pocket in both antibodies (Figure 16, Figure 17). Interestingly, these antibodies both adopt a backward leaning CDR H3 despite their being raised with different immunogens; S73-2 with the sterically-challenging Kdo(2→8)[Kdo(2→4)]Kdo trisaccharide, and S67-27 with the comparatively small Ko(2→4)Kdo disaccharide (Table 9).

Crystal structures revealed that S73-2 uses CDR H3 only to interact with the terminal Kdo residue (Figure 16), where Kdo(2→4)Kdo(2→4)Kdo is bound in an orientation similar to that observed with S25-39. Remarkably, S73-2 binds Kdo(2→8)Kdo(2→4)Kdo in a bent conformation with the second Kdo residue in the

conserved pocket, which emphasizes the versatility of the conserved Kdo-binding pocket to bind alternate lengths and linkages.

The mAb S67-27 cross-reacts with all Kdo-based epitopes tested with comparable avidity (Table 10), notably including the synthetic unnatural antigen 7-O-methyl-Kdo(2→8)Kdo, which binds with 30-fold higher avidity than the cognate immunogen Kdo(2→8)Kdo (Table 13). The structure of S67-27 in complex with a number of antigens showed that like S73-2, it possesses a long backward-leaning CDR H3. However, the bend is more pronounced than in S73-2, resulting in a relatively open combining site that allows recognition of many antigens with (2→4) or (2→8) linked terminal Kdo residues. The increased avidity for the unnatural 7-O-Me-Kdo(2→8)Kdo analog is due to hydrophobic contact between the 7-O-Me group and Ile H96, Pro H98 and Tyr H100C, which are exposed by the backward tilt of CDR H3 (Figure 17). S67-27 again demonstrates the adaptability of this V-gene combination to recognize antigens that the immune system is unlikely to have previously encountered.

Together, these structures of S25-2 type antibodies provide many insights into how the germline antibody repertoire has evolved to conserve recognition of important bacterial carbohydrates while also remaining adaptable to unencountered antigens.

Figure 14: S45-18 in complex with Kdo(2→4)Kdo(2→4)KdoGlcN4*P*(1→6)GlcN1*P*

Binding site interactions from crystal structure of S45-18 in complex with Kdo(2→4)Kdo(2→4)KdoGlcN4*P*(1→6)GlcN1*P*. Hydrogen bonds are shown as dashed black lines. Interactions to terminal Kdo residue are identical to those illustrated in Figure 11 and are not shown.

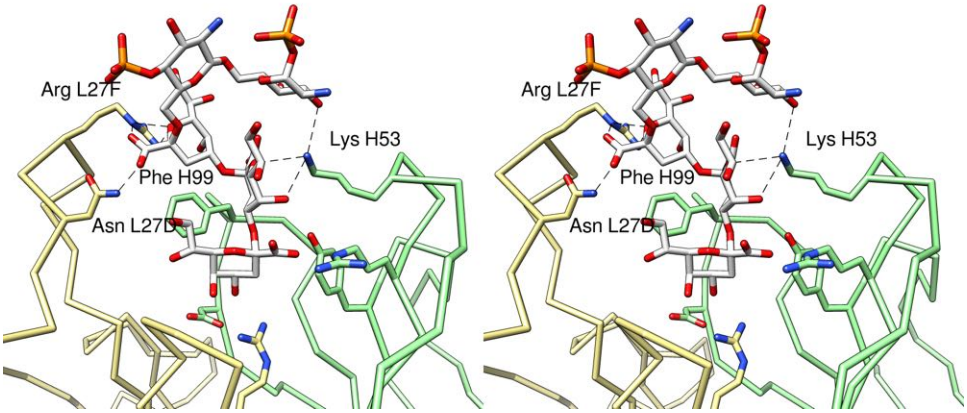


Figure 15: S54-10 in complex with Kdo(2→4)Kdo(2→4)Kdo

Binding site interactions of S54-10 in complex with Kdo(2→4)Kdo(2→4)Kdo. Hydrogen bonds are shown as dashed black lines. Interactions to terminal Kdo residue are identical to those illustrated in Figure 11 and are not shown.

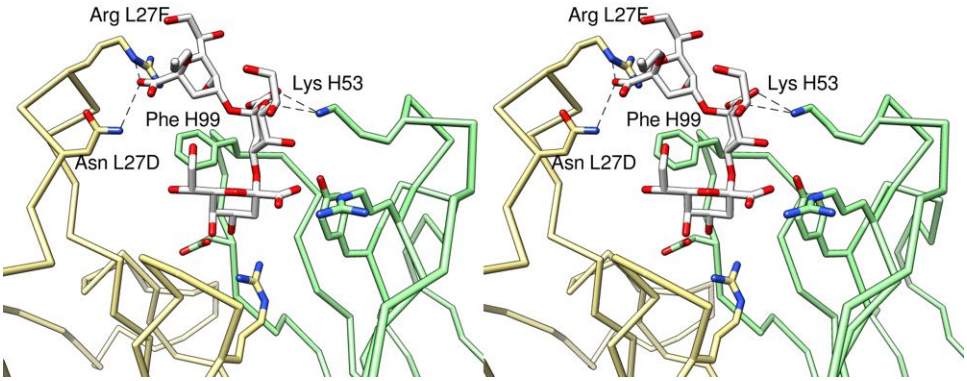


Figure 16: S73-2 in complex with Kdo antigens

Binding site interactions of S73-2 in complex with **(A)** Kdo(2→4)Kdo(2→4)Kdo and **(B)** Kdo(2→8)Kdo(2→4)Kdo. Hydrogen bonds are shown as dashed black lines. Interactions to the Kdo residue in the specificity pocket are identical to those illustrated in Figure 11 and are not shown.

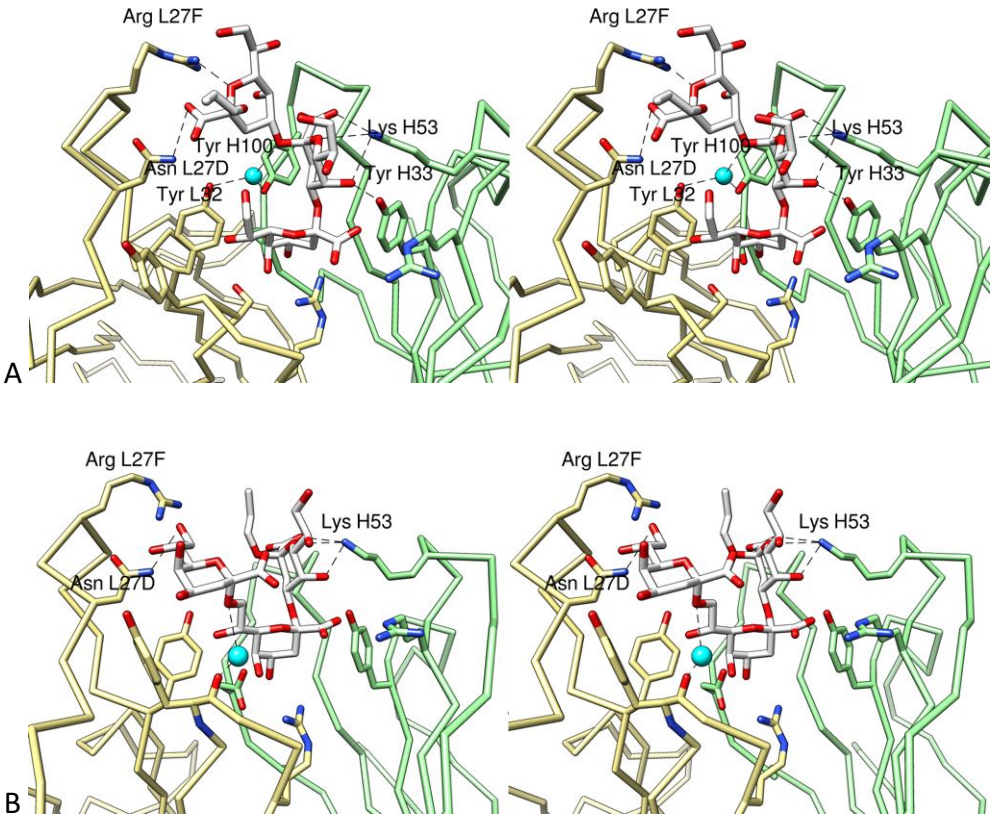
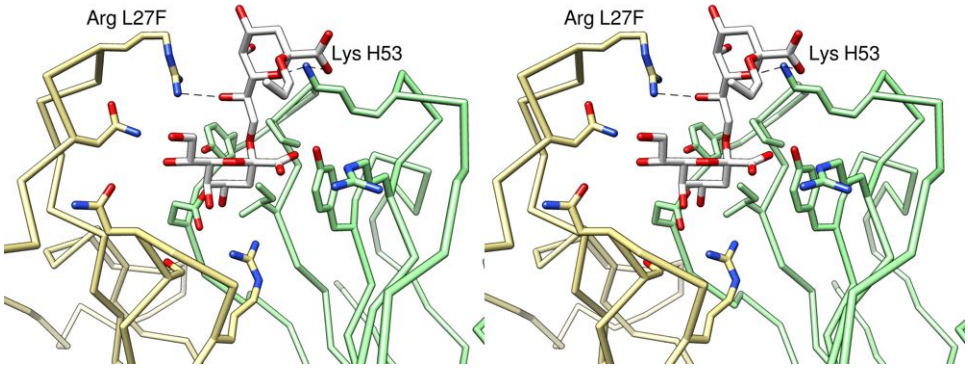


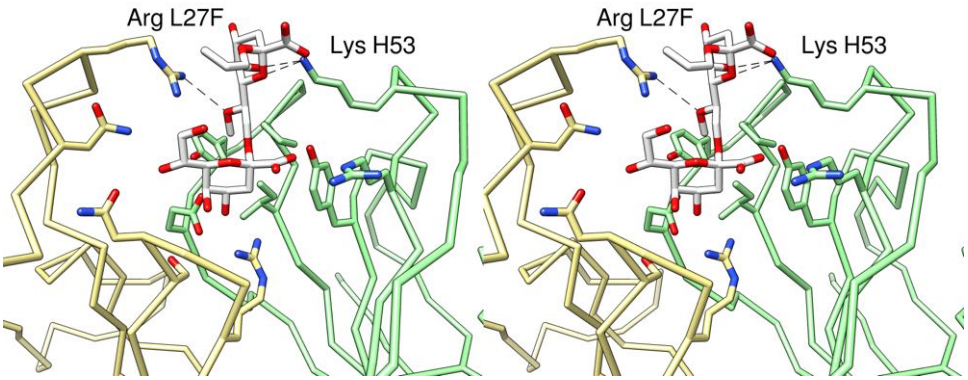
Figure 17: S67-27 in complex with Kdo antigens

Binding site interactions of S67-27 in complex with **(A)** Kdo(2→8)Kdo and **(B)** Kdo(2→8)-7-O-Me-Kdo. Hydrogen bonds are shown as dashed black lines. Interactions to the Kdo residue in the specificity pocket are identical to those illustrated in Figure 11 and are not shown. **(C)** shows a close up of the hydrophobic interactions between Kdo(2→8)-7-O-Me-Kdo and CDR H3.

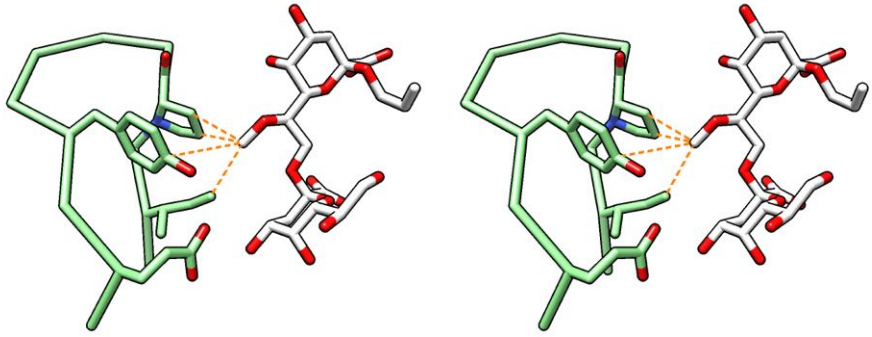
A



B



C



1.8.6. Flexibility of CDR H3 in Kdo antigen binding

In addition to the structures in complex with Kdo oligosaccharides discussed above, the structures of both S25-2 and S25-39 were reported in their unliganded states. Although these antibodies display an identical CDR H3 conformation in all complex structures, there were two unique unliganded conformations observed for S25-2 and a third unique conformation observed for S25-39 (Figure 18).

Although structural differences of CDR H3 between bound and unbound states of antibodies have now been repeatedly observed (229), the differences observed for S25-2 were at the time among a very small number of observations (155, 230–235). The differences observed for S25-2 and S25-39 remain two of the most significant observations in terms of their magnitudes, with rmsd of 3.99, 4.16 and 3.47 Å between the CDR H3s of the shared liganded conformation and of S25-2 unliganded #1, S25-2 unliganded #2 and S25-39 unliganded, respectively[12].

Each unliganded conformation displays complete or nearly complete electron density for backbone atoms (Figure 19), suggesting that the CDR H3 of S25-2 and S25-39 is not completely labile when unliganded, but possesses a number of stable conformers of similar energy. As described in section 1.5.2, such conformational flexibility could enable polyspecific binding by these antibodies.

[12] rmsd calculated using UCSF Chimera (308) between alpha carbon atoms of the 15 residues from Ala H93 to Trp H103 (equal to CDR H3 residues -1 to *n*+1 as described in section 1.4.2), after superimposition of the heavy chain variable domains.

Figure 18: Unliganded conformations of S25-2 and S25-39

The CDR H3 from residue Arg H94 to Trp H103 is shown as ribbon from unliganded structures of S25-2 (pink, green) and S25-39 (tan) and the shared conformation of both S25-2 and S25-39 when bound to Kdo (blue), after alignment of the heavy chain variable domain.

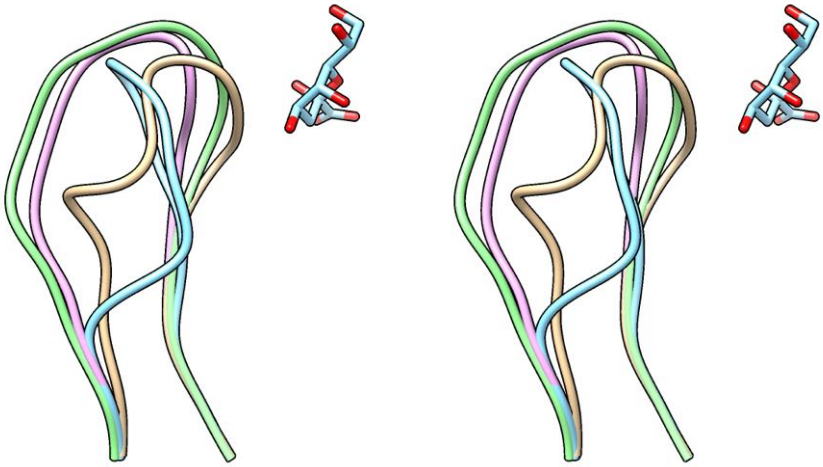
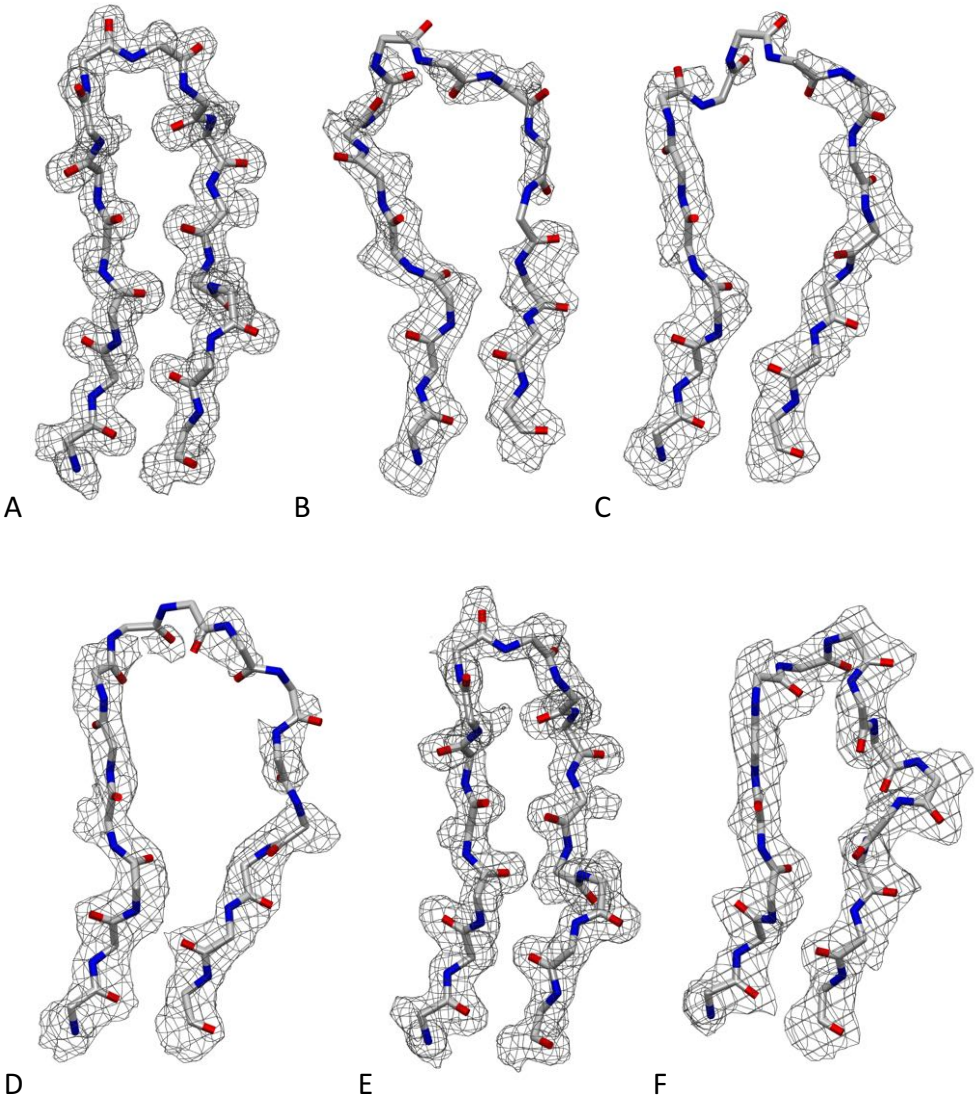


Figure 19: CDR H3 electron density for S25-2 and S25-39 unliganded conformations

2Fo-Fc electron density maps contoured to 1.0σ around CDR H3 backbone atoms of **(A)** S25-2 liganded with Kdo **(B)** S25-2 Unliganded #1 **(C)** S25-2 Unliganded #2 chain B **(D)** S25-2 Unliganded #2 chain D **(E)** S25-39 liganded with Kdo and **(F)** S25-39 unliganded



1.9. Mechanisms of antigen recognition

A large body of research has focused on investigating the more detailed dynamic nature of conformational changes observed with substrate binding. For antibodies, observation of conformational changes with antigen binding have traditionally been categorized as 'induced-fit' (IF), where an initial weak interaction with antigen is followed by structural rearrangement of antibody and/or antigen to achieve better complementarity (158, 159, 231, 236–241). However, beginning in 1965 with the Monod-Wyman-Changeux model to describe cooperative binding phenomena of protein oligomers (242) and accelerating in the last decade, there has been a paradigm shift in the understanding of protein dynamics. Rather than having single lowest-energy conformational states, proteins are now viewed as conformational ensembles with multiple low energy states, where binding events occur by 'conformational selection' (CS) and are followed by population shift (243–252). These binding events can include multiple conformational selection steps or CS followed by IF adjustments (248–250, 253–255).

There has been a recent surge in the literature of kinetic and computational analyses of CS in protein-ligand interactions, including determination of binding affinities for individual conformational states (255, 256), the description of IF and CS on a continuum, shifted by both conformational transition rates and ligand concentration (257–260), and many other observations and discussions of CS (249, 250, 252, 261–266).

Despite this surge in evidence and acknowledgement of the importance of CS in binding, there have been relatively few discussions of CS as a mechanism of antigen

binding by antibodies (122, 123, 162, 267). The few instances of experimental evidence for CS in antibodies include Foote and Milstein's 1994 observation using stopped-flow fluorescence of an equilibrium of antibody conformations with ligand binding preferentially to one form (151), and the 2003 study of James, Roversi and Tawfik, which demonstrated by x-ray crystallography and pre-steady-state kinetics an equilibrium of antibody conformations where hapten binding involved selection of a low-affinity isomer followed by IF to generate the final complex (161).

The relative lack of investigation and observation of CS in antibodies is at odds with the profound importance of flexibility in molecular recognition, as it is crucial to our understanding of polyspecificity at the fundamental level of antibody diversity and off-target effects, and at the applied level of accurate modelling for antibody engineering.

1.10. Objective of this work

The objective of this thesis research is to test the hypothesis that structural flexibility may allow antibody cross-reactivity or polyspecificity and thereby expand the recognition potential of the germline repertoire. Firstly, the flexibility of the anti-carbohydrate antibodies S25-2 and S25-39 is further investigated to gain insight into the conformational diversity possible in near-germline antibodies. This is implemented by crystallizing these antibodies without ligand in numerous conditions to favour alternate packing arrangements likely to be found in different space groups and so selectively stabilize unique conformations, and then using x-ray diffraction to solve and characterize their structures. Secondly, structure-based design is used to modify the

natural antigen of these antibodies, Kdo, to direct cross-reactivity via alternate conformations.

This is the first reported in-depth structural analysis of conformational diversity and its role in antigen binding by near-germline antibodies, and will provide both fundamental and practical insight into the mechanics of antibody-antigen interaction.

Chapter 2: Structural investigations into flexibility and polyspecificity of the anti-carbohydrate antibodies S25-2 and S25-39

The mAbs utilized in this study were produced by my collaborators Sven Müller-Loennies, Lore Brade and Helmut Brade at Research Center Borstel, Leibniz-Center for Medicine and Biosciences, Parkallee 22, Borstel D-23845, Germany. The antigens used were synthesized by Barbara Pokorny and Paul Kosma at the Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria. I performed purification of IgG from hybridoma supernatants, digestion of IgG, purification and crystallization of F_{AB}, and all stages of x-ray diffraction data collection, structure solution and analysis.

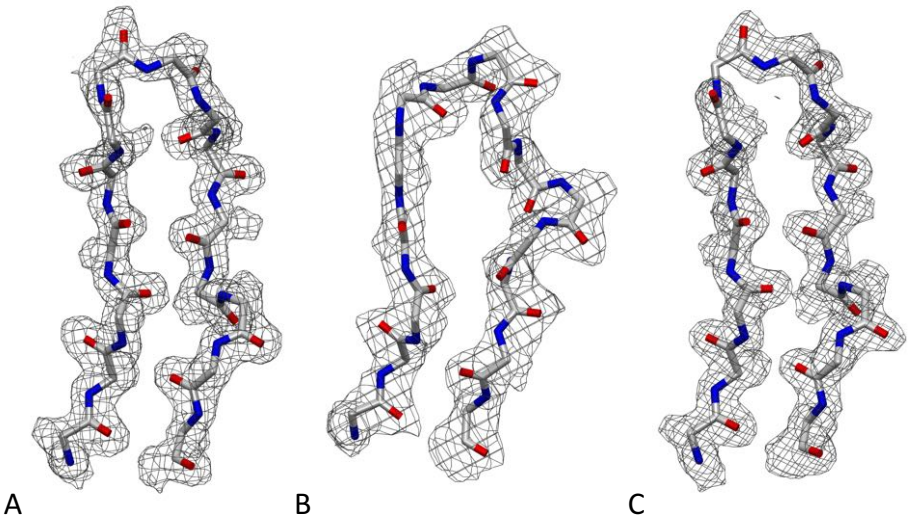
2.1. Results

2.1.1. Crystal Structure of S25-39 Unliganded #2

Two new crystal structures of unliganded S25-39 were solved from crystals grown in conditions different from the original unliganded structure (Table 18, Appendix A). Both of the new structures solved were in the same space group ($P2_1$) as the original Unliganded #1 but with different unit cell dimensions. The first displayed the same CDR H3 conformation as the original Unliganded #1, and the second displayed a CDR H3 conformation identical to the previously observed liganded conformation observed in all complex structures with Kdo antigens (Figure 20). All complex structures with Kdo antigens were solved in space group $P2_12_12_1$ with a unit cell different from this new unliganded structure, called Unliganded #2.

Figure 20: CDR H3 electron density for S25-39 conformations

2Fo-Fc electron density maps contoured to 1.0σ around CDR H3 backbone atoms of **(A)** S25-39 liganded with Kdo **(B)** S25-39 unliganded #1 and **(C)** S25-39 Unliganded #2



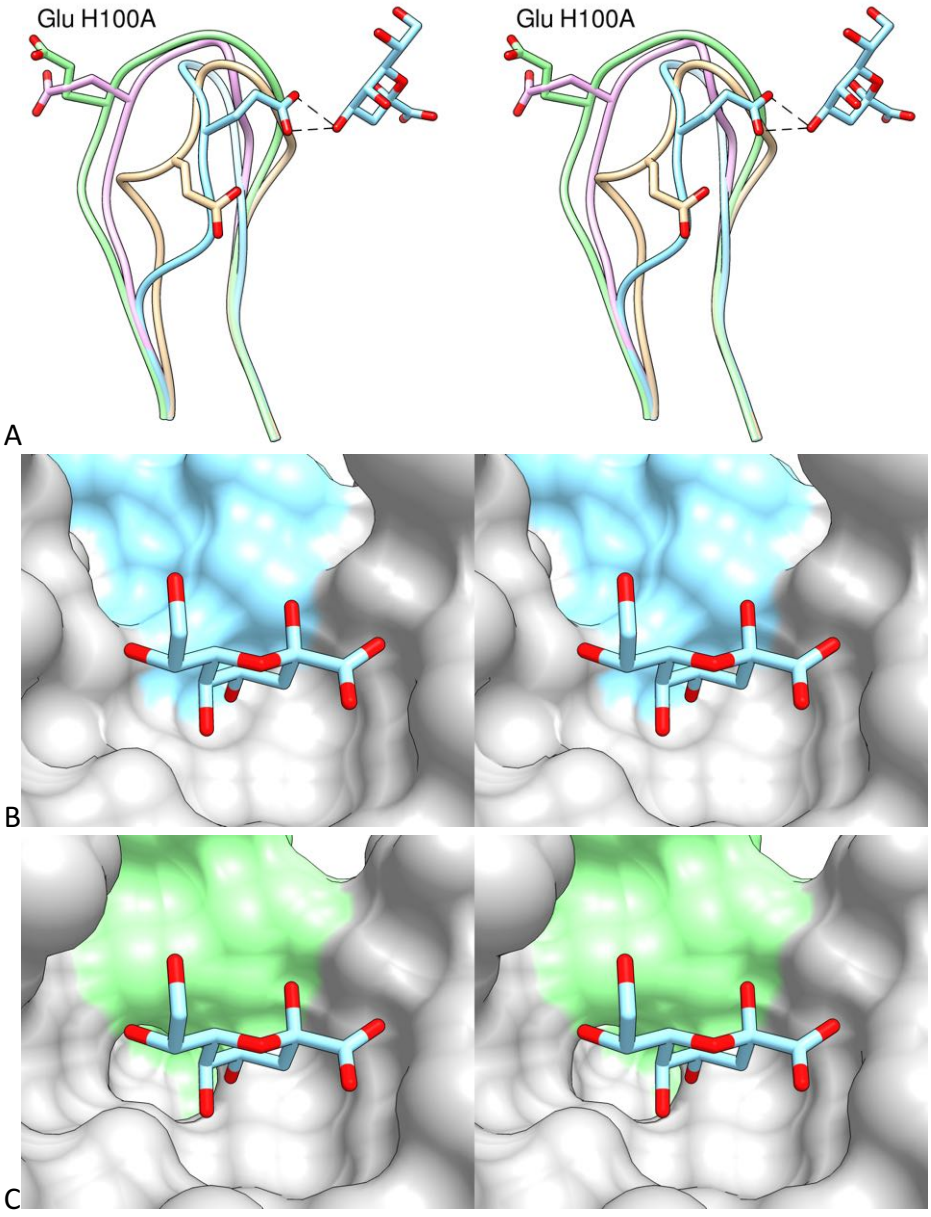
A B C

2.1.2. Design and synthesis of Kdo-based monosaccharide antigens

The ability of S25-39 and S25-2 to bind unique ligands through alternate binding-site conformations was investigated through structure-based ligand design. Based on the binding site topology and ligand interactions observed in previous structures, novel Kdo-based antigens were designed to inhibit interaction of Kdo O4 with Glu H100A of CDR H3 in the liganded conformation (Figure 21 A), but to be compatible with observed unliganded conformations – in particular with S25-2 Unliganded #1 that displays a small hydrophobic pocket in position to accommodate Kdo O4 substitutions (Figure 21 C). The first antigen synthesized and crystallized with S25-39 was 4-MeO-Kdo, which was followed by 4-O-benzyl-KdoOMe, 4-O-Ethoxymethyl-KdoOMe and 4-O-Methoxymethyl-KdoOMe (Figure 22).

Figure 21: Analysis of CDR H3 structures for ligand design

(A) Synthetic Kdo analogues with alternate substituents on O4 were created to inhibit the interaction of Kdo O4 with Glu H100 observed in liganded structures (blue) but remain compatible with observed unliganded structures (S25-2 Unliganded #1, green; S25-2 Unliganded #2, pink; S25-39 Unliganded #1, tan). The binding site surfaces of **(B)** S25-39 liganded with Kdo, **(C)** S25-2 Unliganded #1, **(D)** S25-2 Unliganded #2 and **(E)** S25-39 Unliganded #1 are shown with Kdo overlayed by superimposition with the S25-39 liganded structure, and the surface contributed by CDR H3 coloured as in (A),



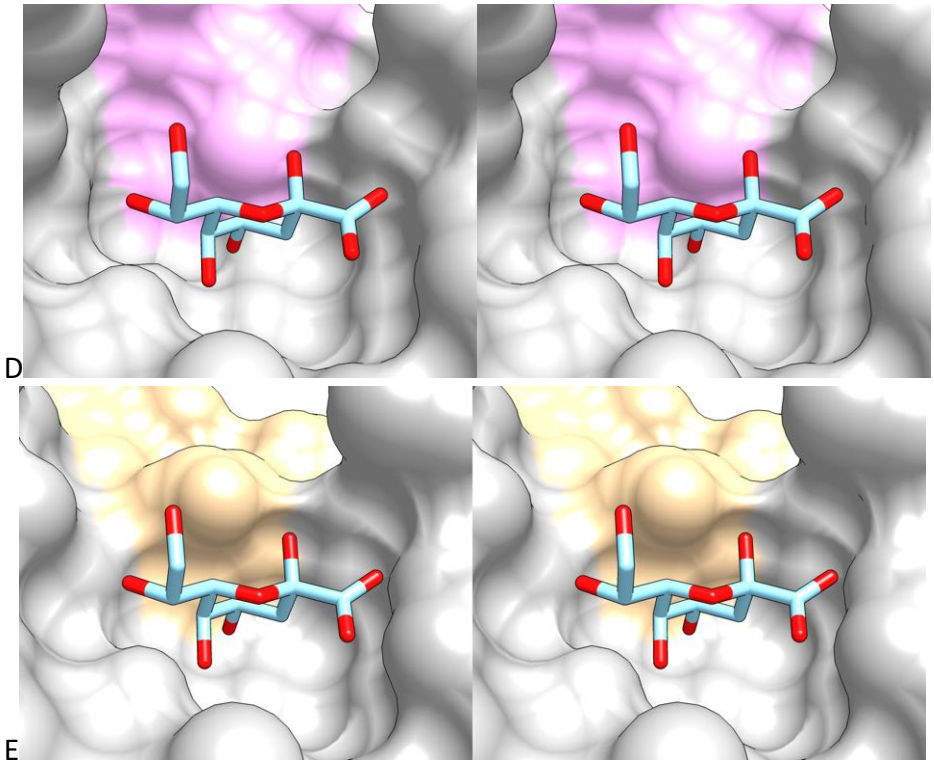
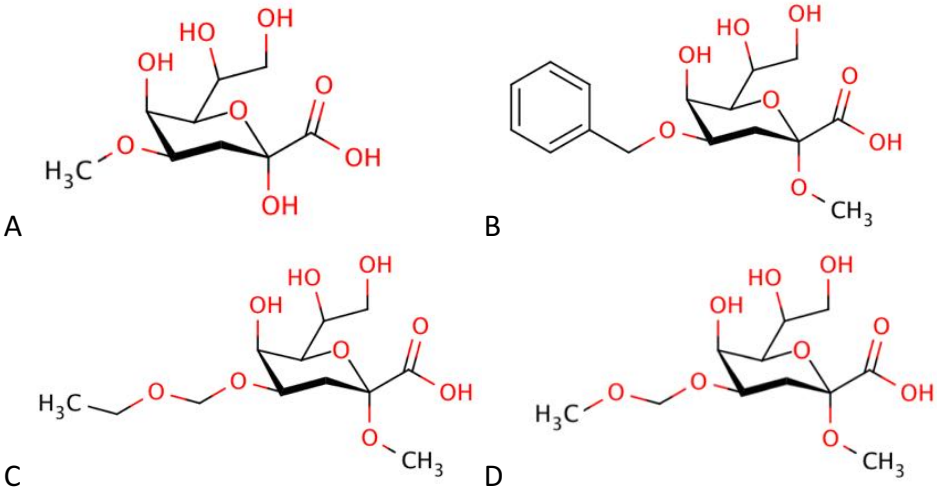


Figure 22: Synthetic Kdo analogues

Chemical structures of synthetic antigens **(A)** 4-MeO-Kdo **(B)** 4-O-benzyl-KdoOMe **(C)** 4-O-Ethoxymethyl-KdoOMe **(D)** 4-O-Methoxymethyl-KdoOMe.



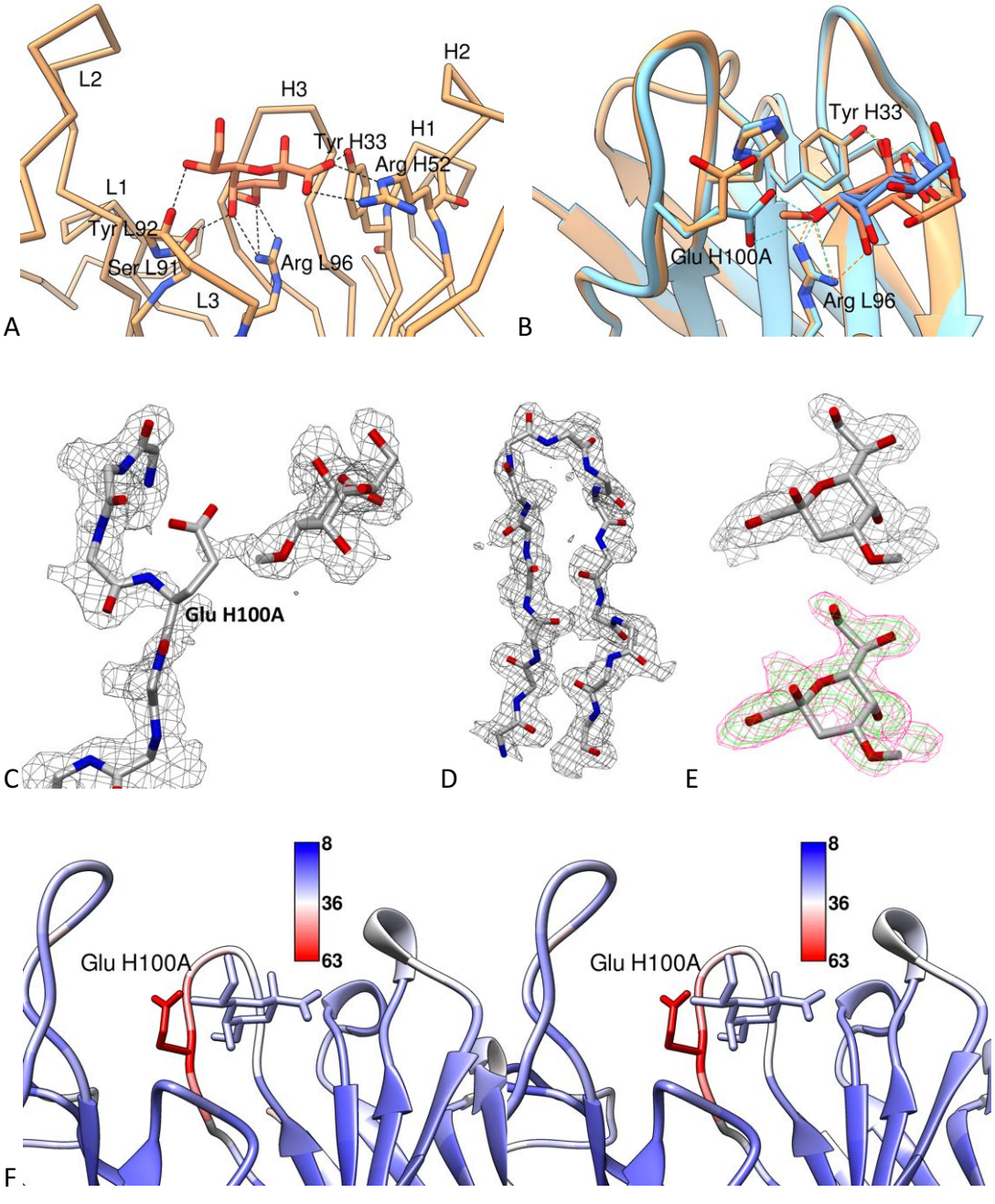
2.1.3. Crystal structures of S25-39 in complex with 4-MeO-Kdo

The crystal structure of S25-39 in complex with 4-MeO-Kdo was solved to 2.00 Å in space group $P2_1$ with 1 molecule in the AU. The structure displays excellent electron density for all main chain atoms with the exception of residues H127 to H133 in the constant region (the opposite end of the F_{AB} from the combining site) that are disordered and not modeled. Data collection and refinement statistics are given in Table 20, Appendix A.

Good electron density for 4-MeO-Kdo is observed in the binding site, with an orientation similar to Kdo in complex with S25-39 (Figure 23). The backbone of CDR H3 adopts the same conformation as previous structures of S25-39 in complex with Kdo antigens. However, the side chain of Glu H100A is completely disordered and the B-factors of neighbouring residues are elevated compared to the average B-factor of the structure, and compared to the CDR H3 B-factors of the Kdo complex structure, indicating destabilization by 4-MeO-Kdo binding (Figure 23, Table 15).

Figure 23: S25-39 binding 4-MeO-Kdo

Specific interactions of 4-MeO-Kdo in complex with S25-39 are shown in **(A)** and a superposition with Kdo-bound S25-39 in **(B)**. 2Fo-Fc electron density maps contoured to 1.0σ around **(C)** 4-MeO-Kdo and residues of CDR H3 surrounding Glu H100A, **(D)** all of CDR H3, and **(E)** 4-MeO-Kdo, also including Fo-Fc omit maps contoured to 2.0σ and 3.0σ. **(F)** shows the binding site coloured by B-factor.



2.1.4. Crystal structures of S25-39 grown with synthetic antigens

After the initial structure of S25-39 in complex with 4-MeO-Kdo was solved, the additional antigens 4-O-benzyl-KdoOMe, 4-O-ethoxymethyl-KdoOMe and 4-O-methoxymethyl-KdoOMe were synthesized and crystallized with both S25-2 and S25-39.

All 14 structures of S25-39 solved from crystals grown in the presence of these antigens were unliganded, and three of these structures were unique (Table 18, Appendix A). Unique structures were solved in space groups $P2_1$ (Unliganded #1 and Unliganded #2, to 1.53 Å and 1.30 Å each with 1 molecule in the AU) and in space group C2 (unliganded #3, to 2.10 Å with 4 molecules in the AU). Of the 14 total structures, there were 7 observations of Unliganded #1, 6 observations of Unliganded #2, and a single observation of Unliganded #3. Data collection and refinement statistics are given in Table 20, Appendix A.

Unliganded #1 displays excellent electron density for all main chain atoms with the exception of residues H127 to H133 in the constant region that are disordered and not modeled. Electron density for a BIS-TRIS molecule from crystallization buffer was observed coordinated to Glu H148, Val H150 and Ala H168 in the constant region. This structure shows an identical CDR H3 conformation to the original S25-39 Unliganded #1 structure but was solved to significantly higher resolution (Figure 24).

Unliganded #2 displays excellent electron density for all main chain atoms with the exception of residues H127 to H133 that are disordered and not modeled. This structure shows an identical CDR H3 conformation to the original S25-39 Unliganded #2 structure but was solved to significantly higher resolution (Figure 24).

Unliganded #3 displays good electron density for all main chain atoms with the exception of residues H127 to H134, B100A to B100B in CDR H3 of chain B, B127 to B132, D128 to D133, F98 to F100B in CDR H3 of chain F, and F127 to F133[13], which are disordered and not modeled. In one out of four molecules in the AU (chains C/D) there was ambiguous electron density in the antigen binding site where 4-O-ethoxymethyl-KdoOMe could be modeled (Figure 25). However, the density could be equally well-modeled by water molecules and an omit difference map is more consistent with water molecules, which are therefore modeled in the final structure. Binding data was not collected to evaluate this interaction. The four molecules in the AU include two unique conformations of CDR H3 in molecules H and D, and two partially disordered conformations in B and F (Figure 24, Figure 26).

[13] When there are multiple F_{AB} molecules in the AU, the light and heavy chains are labelled L/H for one molecule, then A/B, C/D, E/F, *etc*., for others. However, the CDRs of each molecule are still labelled L1, L2, L3, H1, H2 and H3 regardless of chain label.

Figure 24: Electron density of S25-39 CDR H3 conformations

2Fo-Fc electron density maps contoured to 1σ around backbone atoms of CDR H3 of S25-39 **(A)** Unliganded #1, **(B)** Unliganded #2, **(C)** Unliganded #3 chain H, **(D)** Unliganded #3 chain B, **(E)** Unliganded #3 chain D and **(F)** Unliganded #3 chain F.

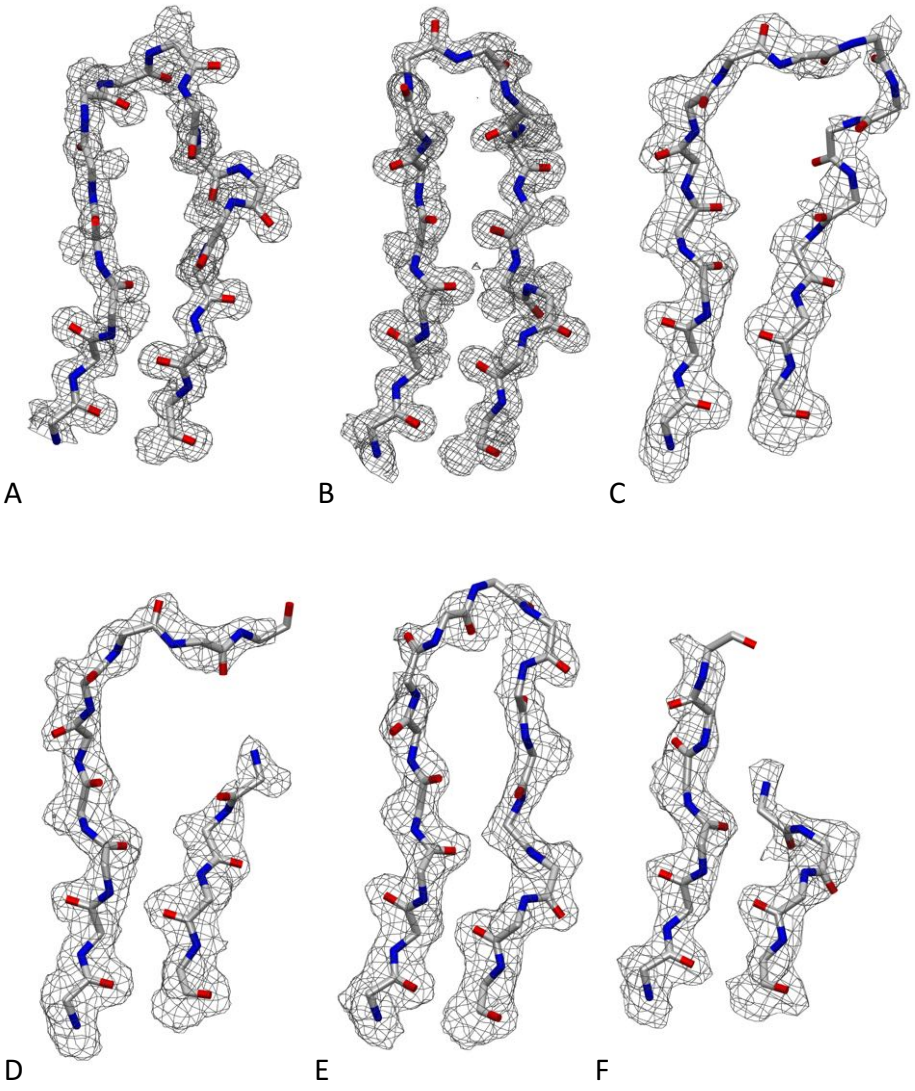
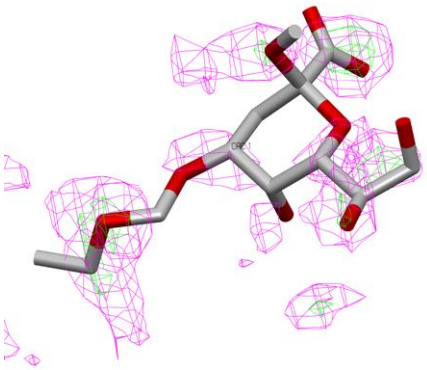
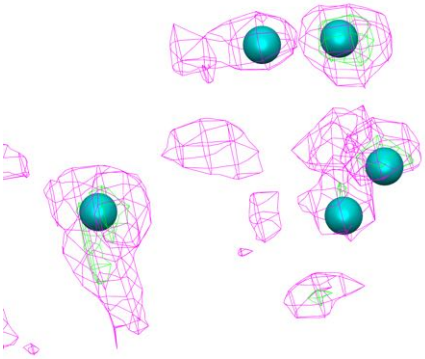


Figure 25: Binding site electron density of S25-39 Unliganded #3

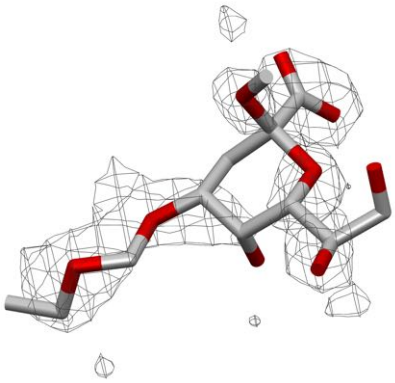
Omit Fo-Fc electron density maps contoured to 2σ (magenta) and 3σ (green) of the S25-39 Unliganded #3 chain C/D binding site with superimposed **(A)** 4-O-Ethoxymethyl-KdoOMe and **(B)** water molecules, and refined 2Fo-Fc electron density maps contoured to 1σ around **(C)** 4-O-Ethoxymethyl-KdoOMe and **(D)** water molecules.



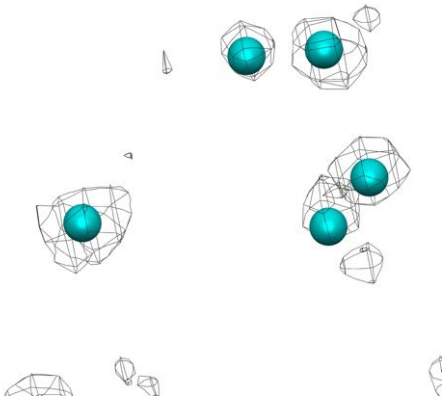
A



B



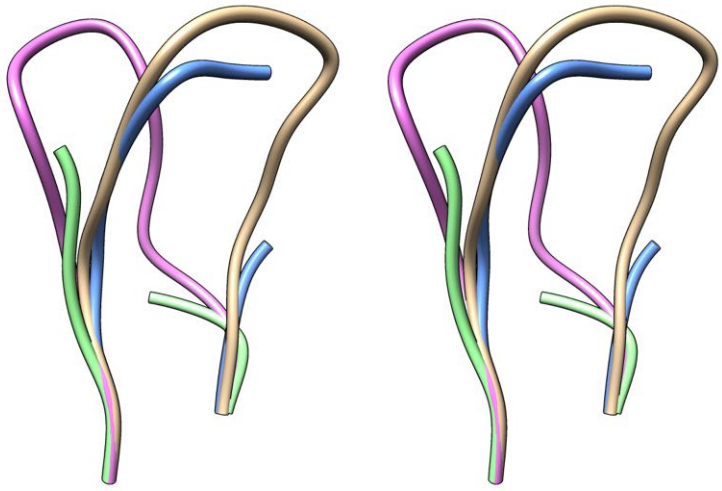
C



D

Figure 26: CDR H3 conformations of S25-39 Unliganded #3

(A) Aligned CDR H3 from Ala 93 to Tyr 102 of each molecule in the S25-39 Unliganded #3 AU. Chain H Tan, Chain B Blue, Chain D pink, Chain F Green.



2.1.5. Crystal structures of S25-2 grown with synthetic antigens

Four crystal structures were solved of S25-2 from crystals grown in the presence of 4-O-benzyl-KdoOMe or 4-O-methoxymethyl-KdoOMe. No well-diffracting crystals were obtained grown with 4-O-ethoxymethyl-KdoOMe. All structures were unliganded and two were unique (Table 19, Appendix A). The new structures of S25-2, Unliganded #3 and Unliganded #4, were both solved in space group $P2_1$ to 2.15 Å and 2.0 Å, respectively, each with 4 molecules in the AU (Unliganded #1 and #2 are the previously published structures 1Q9K and 1Q9L). Data collection and refinement statistics are given in Table 20, Appendix A.

Unliganded #3 displays good electron density for all main chain atoms with the exception of residues B100B and B100C of CDR H3 in molecule B, B128 to B133, and D127 to D133, which are disordered and not modeled. The four molecules in the AU display three unique conformations of CDR H3, with one conformation occurring twice

and one conformation partially disordered (Figure 27, Figure 28), and all conformations are different from the previously observed Unliganded #1 and Unliganded #2. All antigen binding sites are occupied by symmetry-related F_{AB} molecules making several specific interactions, including a salt bridge between the binding site Arg 52 of the heavy chain and Glu 213 of the symmetry-related light chain, mimicking the salt bridge to the carboxylate of Kdo antigens (Figure 29). Each F_{AB} in the AU makes slightly different contacts to the symmetry-related molecules in the binding sites.

Unliganded #4 displays good electron density for all main chain atoms with the exception of H98 to H100A of CDR H3 in molecule H, and B128 to B133, which are disordered and not modeled. This crystal was grown in 0.2 M ammonium iodide and 34 iodide ions are modeled in the final structure. Three of the four molecules in the AU display an identical CDR H3 conformation that is different than previously observed conformations, and the fourth is partially disordered (Figure 30, Figure 31A). Each CDR H3 coordinates three iodide ions in similar positions (Figure 31B). The structure displays the same crystal packing as Unliganded #3 described above, with a symmetry related molecule occupying each binding site (Figure 31C). None of the four F_{AB} molecules of Unliganded #4 show additional structural differences from other S25-2 structures due to iodide coordination.

Figure 27: Electron density of S25-2 Unliganded #3 CDR H3 conformations

2Fo-Fc electron density maps contoured to 1σ around backbone atoms of CDR H3 of S25-2 Unliganded #3 **(A)** chain H, **(B)** chain B, **(C)** chain D and **(D)** chain F.

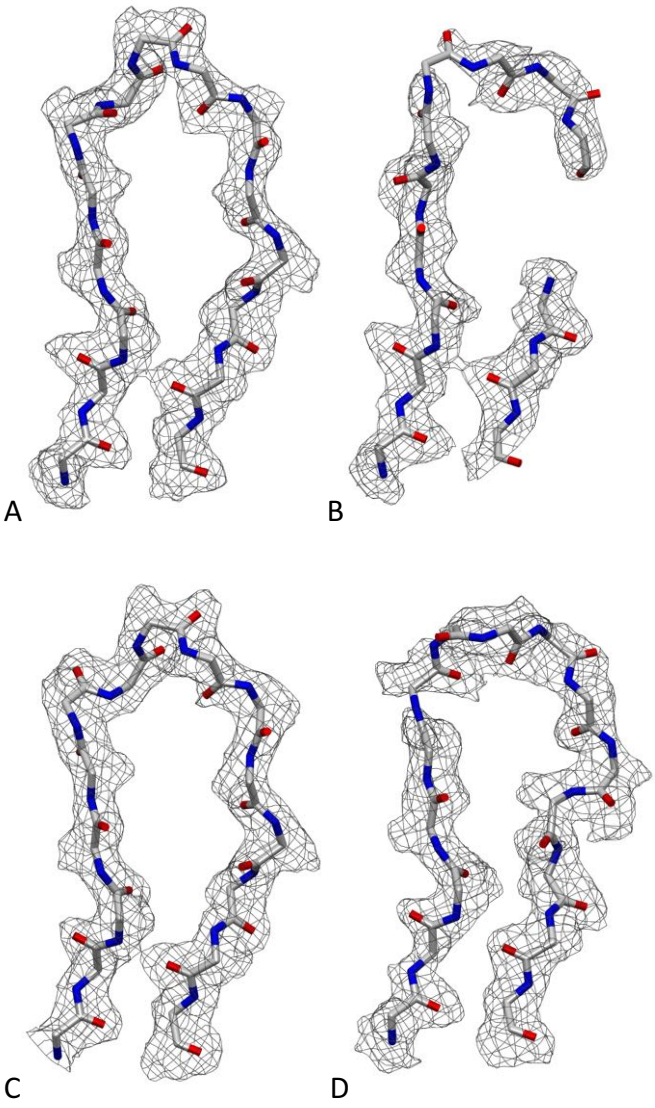


Figure 28: CDR H3 conformations of S25-2 Unliganded #3

(A) Aligned CDR H3 from Ala 93 to Tyr 102 of each molecule in the S25-2 Unliganded #3 AU. Chain H Tan, Chain B Blue, Chain D pink, Chain F Green.

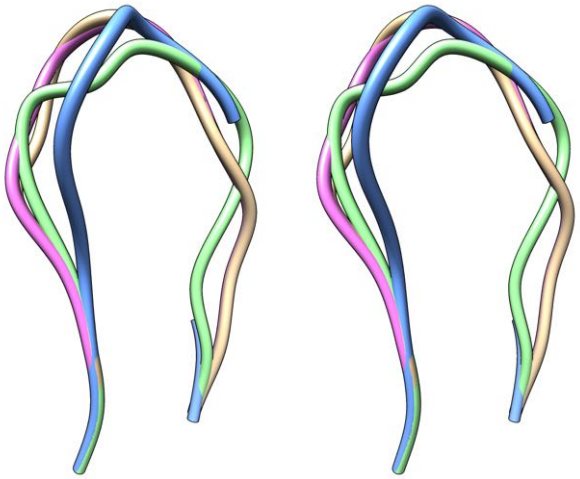
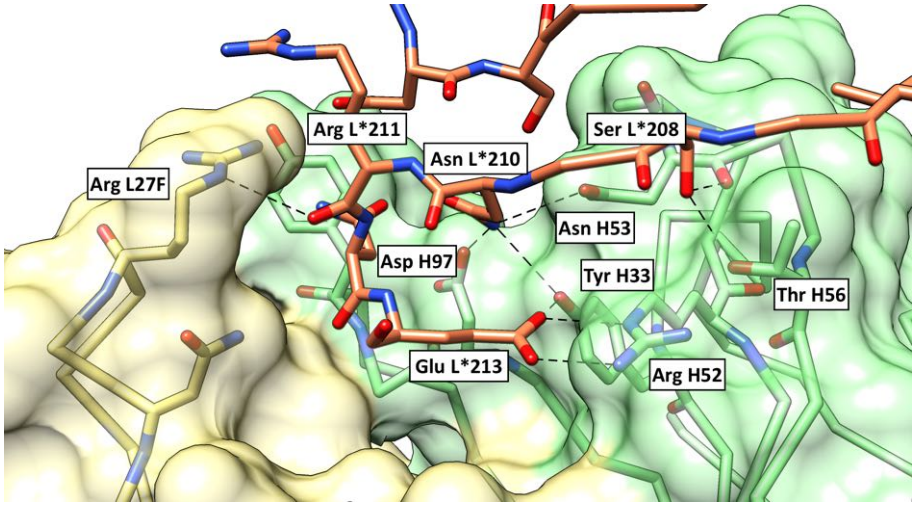
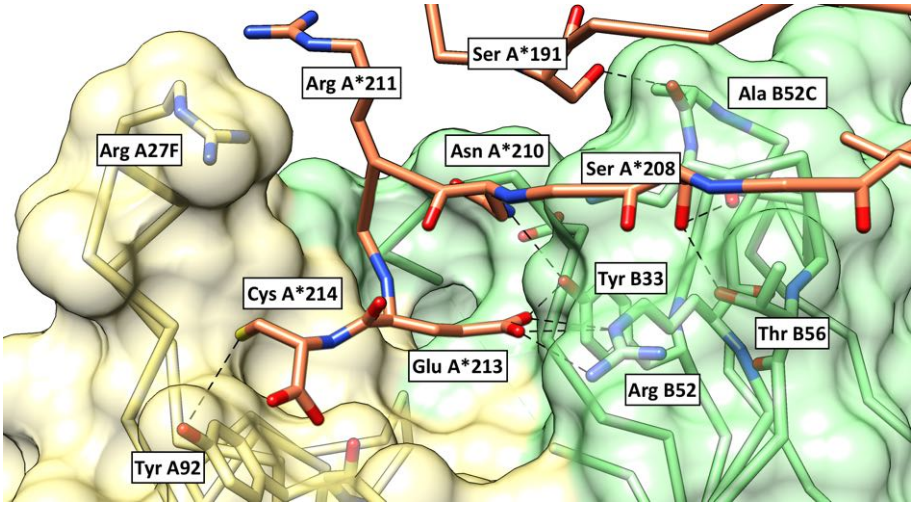


Figure 29: Binding site crystal contacts of S25-2 Unliganded #3

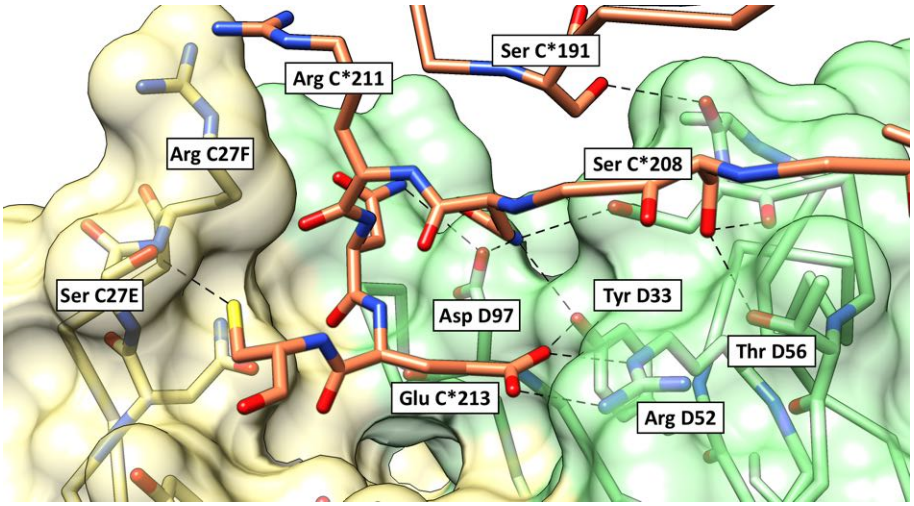
The binding site of each of the four Fab molecules **(A)** L/H **(B)** A/B **(C)** C/D and **(D)** E/F in the AU of the S25-2 Unliganded #3 structure is occupied by the C-terminus of a symmetry-related light chain. The * in residue labels indicates the symmetry-related molecule.

A 

B


Ser A*191
Arg A*211
Ala B52C
Arg A27F
Asn A*210
Ser A*208
Tyr B33
Cys A*214
Thr B56
Glu A*213
Arg B52
Tyr A92

C


Ser C*191
Arg C*211
Arg C27F
Ser C*208
Ser C27E
Tyr D33
Asp D97
Thr D56
Glu C*213
Arg D52

D

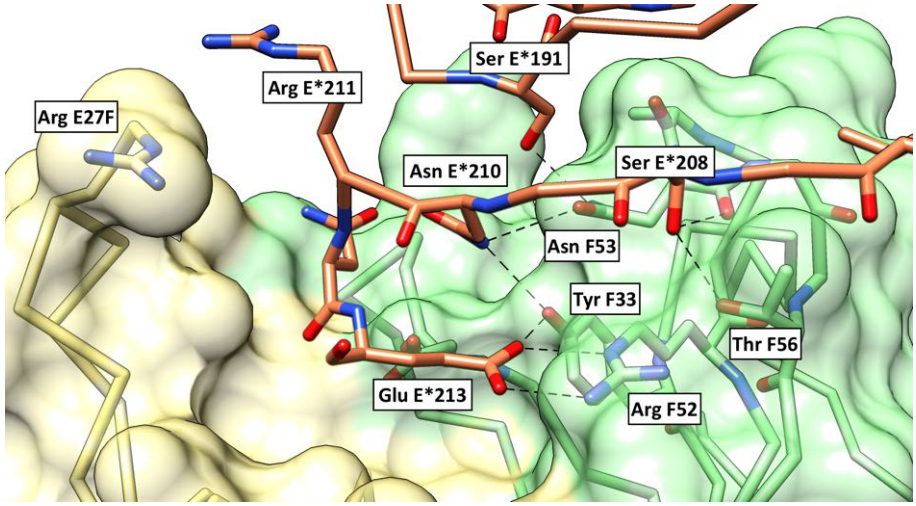

Ser E*191
Arg E*211
Arg E27F
Asn E*210
Ser E*208
Asn F53
Tyr F33
Thr F56
Glu E*213
Arg F52

Figure 30: Electron density of S25-2 Unliganded #4 CDR H3 conformations

2Fo-Fc electron density maps contoured to 1σ around backbone atoms and iodide ions of CDR H3 of S25-2 Unliganded #4 **(A)** chain H, **(B)** chain B, **(C)** chain D and **(D)** chain F.

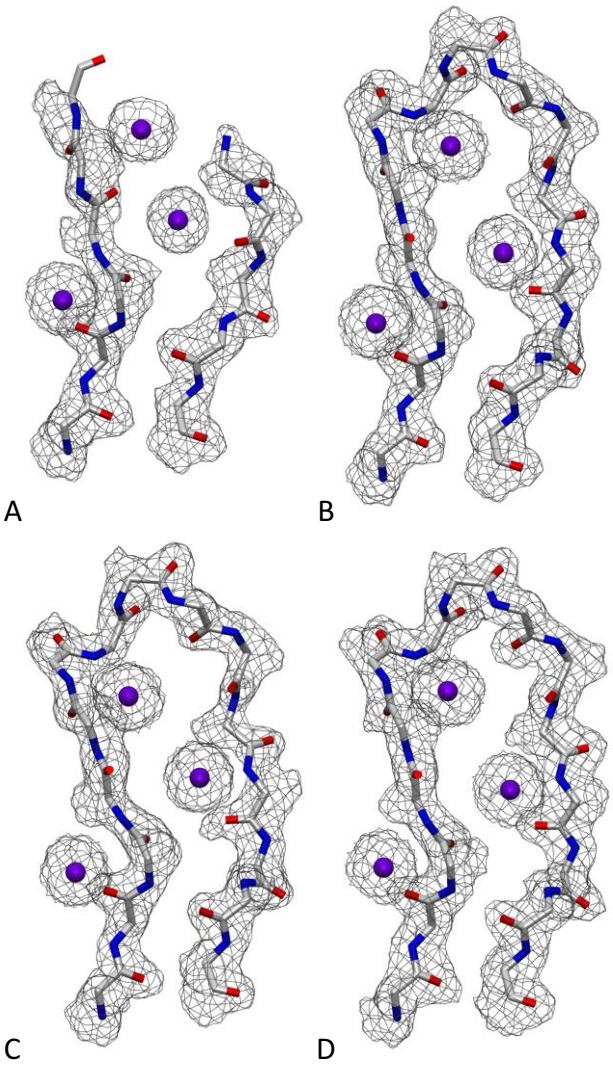
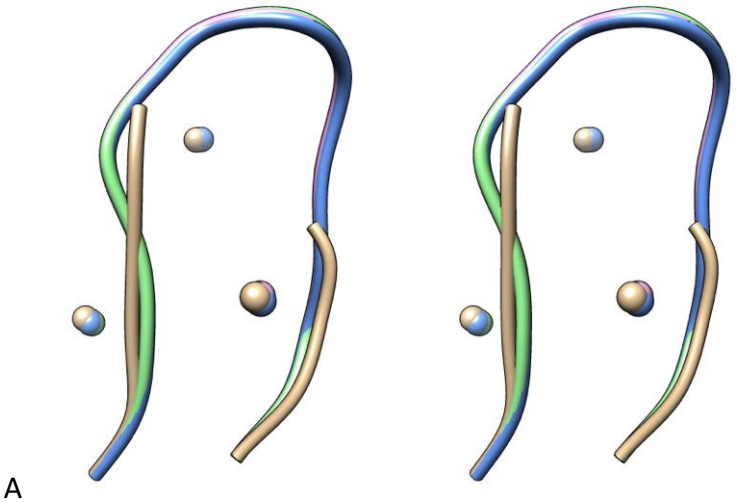
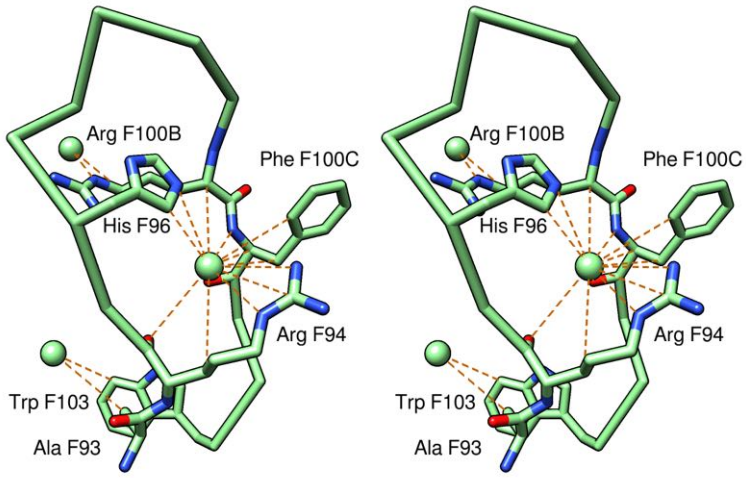


Figure 31: S25-2 Unliganded #4 CDR H3 conformations

(A) Aligned CDR H3 from Ala 93 to Tyr 102 of each molecule in the S25-2 Unliganded #4 AU. Chain H Tan, Chain B Blue, Chain D pink, Chain F Green. Iodide ions are shown as spheres. **(B)** shows contacts[14] to iodide residues from chain F. **(C)** shows an alignment of the binding sites of all four F_{AB} molecules in the AU, each occupied by a symmetry-related light chain. Specific interactions are similar to those shown Figure 31.

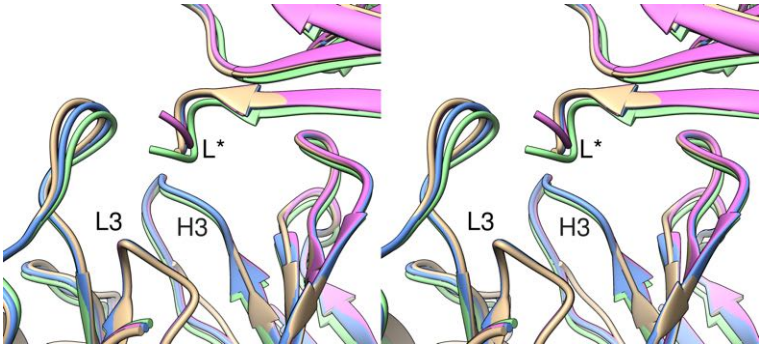


A



B

[14] Determined using UCSF Chimera (308) Find Clashes/Contacts tool with default contact criteria (atoms with ≥ -0.4 Å VDW overlap)

C 

2.1.6. Comparison of S25-39 and S25-2 CDR conformations

Superposition of the Fv regions of all crystal structures of S25-2 and S25-39 reveals conservation in the conformations and positions of most CDRs (Figure 32). CDR L1 shows a nearly identical conformation in all structures but slight shifts in location. When the light chain variable domains of all molecules are aligned there is an overall rmsd[15] of 0.577 Å for CDR L1 Cα atoms and the largest linear displacement of Cα atoms is 3.6 Å, whereas the conformations and positions of CDRs L2 and L3 are essentially identical with overall rmsd of 0.285 Å and 0.305 Å, respectively (Figure 32A). Likewise, when the heavy chain variable domains of all molecules are aligned, the conformations and positions of CDRs H1 and H2 are essentially identical with overall rmsd of 0.281 Å and 0.250 Å, respectively (Figure 32B). The spread of V_H or V_L positions observed upon alignment of the opposite domain reflects a range of possible V_H-V_L orientation angles. V_H-V_L orientation differences in bound and unbound antibody structures are relatively common, and this rearrangement has been proposed as an additional mechanism for generating unique binding sites to expand germline recognition potential (268).

[15] Overall rmsd includes the pairwise deviations of Cα atoms between every pair of structures (only modelled atoms), weighted equally, after superimposition of the relevant variable domain. Calculated with UCSF Chimera (308).

In all S25-2 and S25-39 structures, CDRs H1, H2, L1, L2 and L3 adopt the Chothia canonical conformations H1-1, H2-4, L1-3, L2-1 and L3-6 (76, 77, 269), or the North *et al.* canonical clusters H1-13-1, H2-12-1, L1-17-1, L2-8-1, and L3-8-1 (82, 84), summarized in Table 14. The Chothia canonical conformation L3-6 was an addition to the original Chothia canonical system (272) with Arg in place of Pro at position 96. This conformation was re-classified by North *et al.* (82) as CDR L3-8-1 and further sub-classified by Teplyakov and Gilliland (303) as CDR L3-8-NP.

The most notable difference between these structures is the observation of several unique[16] conformations of CDR H3, with an overall rmsd of 3.224 Å and individual rmsd from the liganded conformation as high as 6 Å (Table 15), and linear displacement of individual Cα atoms as large as 12 Å (Figure 32B, Figure 33). The CDR H3 conformations cover a range from well-ordered with complete electron density for backbone atoms and low B-factors, to mostly disordered with poor or absent electron density (Table 15).

S25-2 and S25-39 do not share any of the same unliganded conformations. There are a larger number of more similar conformations among the S25-2 unliganded structures compared to the S25-39 structures; the overall rmsd of the seven unique S25-2 unliganded conformations is 1.811 Å, whereas the overall rmsd of the four unique S25-39 unliganded conformations is 4.040 Å (Figure 33, Table 16).

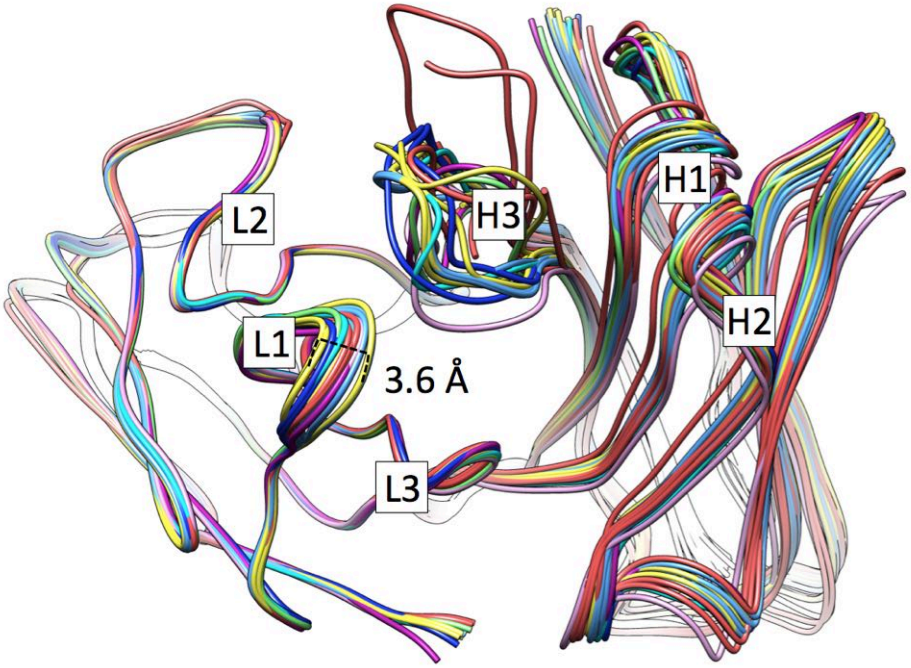
The CDR H3 anchors of all conformations were classified as kinked or extended based on the (n – 2) to (n + 1) Cα dihedral angle, with only the liganded conformation

[16] A unique conformation is defined by a unique Ramachandran code, Table 16.

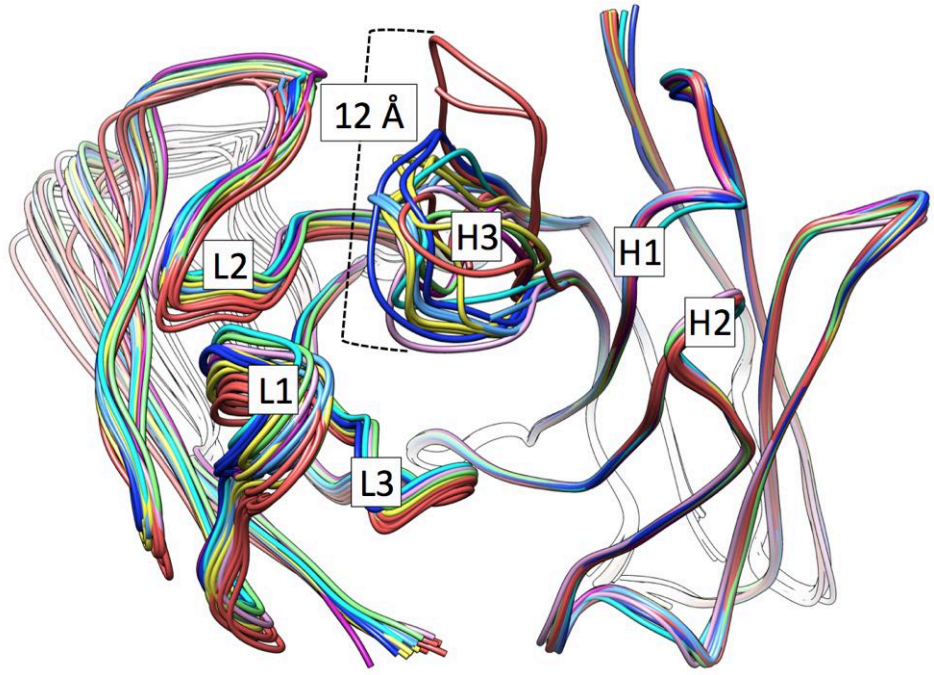
and the S25-39 Unliganded #2 conformation being kinked and all other conformations being extended (Table 16). The Ramachandran code (Figure 7) for each CDR H3 was determined to compare with the North *et al.* anchor clusters (Table 7). No conformation perfectly matches any anchor cluster, but there are some matches when permitting the similar Ramachandran regions B/P, A/D and L/G. The shared liganded conformation and S25-39 Unliganded #2 has a Ramachandran anchor BPB/BBAB that is similar to the H3-anchor-1 cluster BPP/BPAB, and several S25-2 unliganded conformations are similar to the H3-anchor-5 cluster BPB/PPBB. The Ramachandran codes further highlight the similarities of some S25-2 conformations, in particular Unliganded #2, Unliganded #3 chain C, and Unliganded #4. The clusters for entire CDR H3 conformations were also determined using PyIgClassify (Table 16).

Figure 32: Superposition of all S25-2 and S25-39 F_V conformations

Superposition of all S25-2 and S25-39 conformations through alignment of **(A)** the light chain variable domain and **(B)** the heavy chain variable domain. S25-2 liganded dark purple, S25-2 Unliganded #1 cyan, S25-2 Unliganded #2 dark blue, S25-2 Unliganded #3 yellow, S25-2 Unliganded #4 light blue, S25-39 Unliganded #1 light purple, S25-39 Unliganded #2 light green, S25-39 Unliganded #3 medium red. The maximum distance between identical Cα atoms in different conformations is indicated for CDRs L1 and H3.



A



B

Figure 33: Alignment of all S25-2 and S25-39 CDR H3 conformations

Alignment of all S25-2 and S25-39 CDR H3 conformations. **(A)** side view and **(B)** top view of all conformations of both antibodies, **(C)** side view and **(D)** top view of only S25-2, and **(E)** side view and **(F)** top view of only S25-39. S25-2 liganded green, S25-2 Unliganded #1 cyan, S25-2 Unliganded #2 dark blue, S25-2 Unliganded #3 light blue, S25-2 Unliganded #4 medium blue, S25-39 Unliganded #1 red, S25-39 Unliganded #2 yellow, S25-39 Unliganded #3 orange.

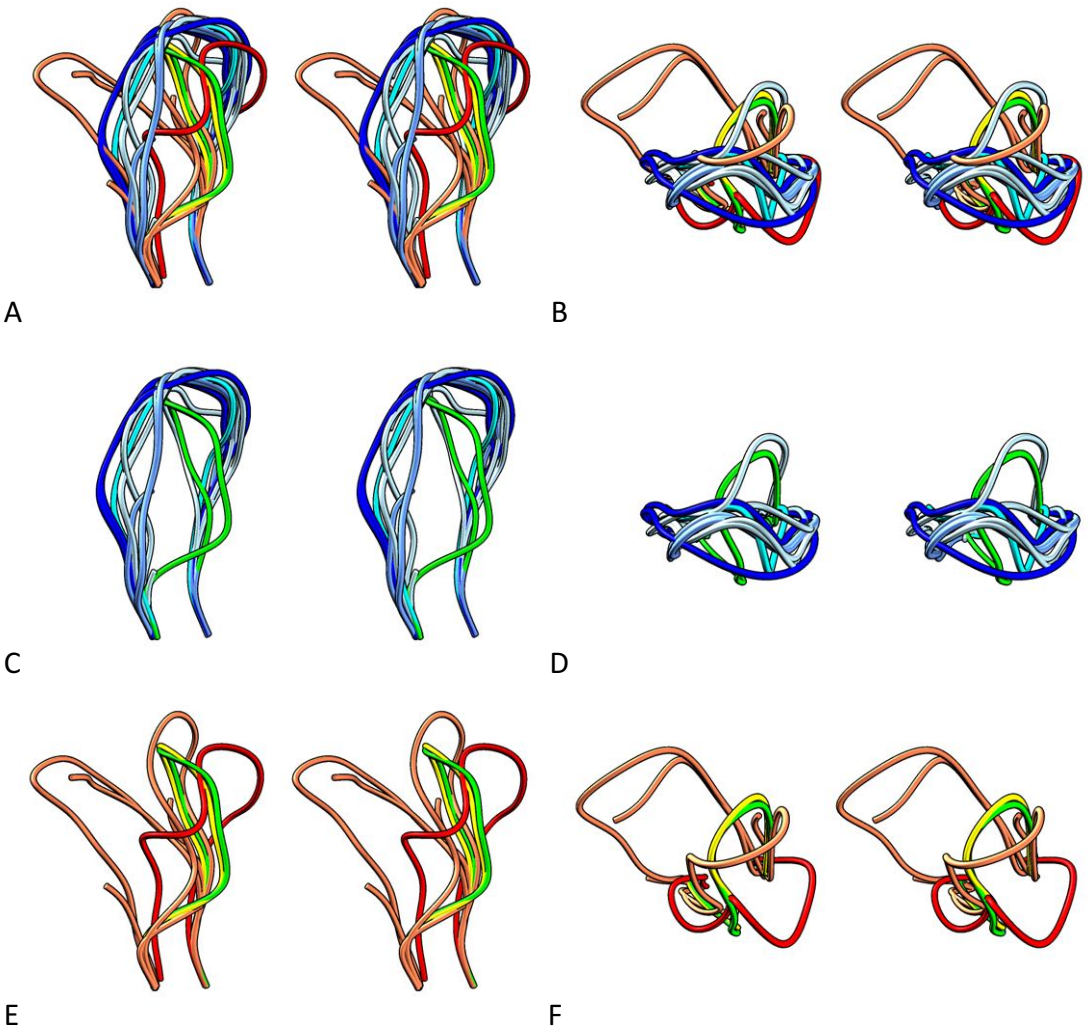


Table 14: Canonical classifications of S25-2 and S25-39 CDRs

Classifications were made with PyIgClassify[17] for North Clusters and the Chothia Canonical Assignment tool from Andrew C.R. Martin's Bioinformatics Group at UCL[18]. *The Canonical class L3-6 was assigned by sequence and structure comparison as described in (270)

CDR	Chothia Canonical	North Cluster
L1	L1-3	L1-17-1
L2	L2-1	L2-8-1
L3	L3-6*	L3-8-1
H1	H1-1	H1-13-1
H2	H2-4	H2-12-1

[17] http://dunbrack2.fccc.edu/PyIgClassify/default.aspx (84)

[18] http://www.bioinf.org.uk/index.html (80)

Table 15: CDR H3 properties of S25-2 and S25-39 structures

For each conformation of CDR H3, several indicators of loop stability are given. The amino acid sequence of H3 is given under residue order, coloured based on whether the backbone atoms are observed with good electron density (green) or poor density (magenta), and residues not modelled due to disorder are replaced by dashes. The average B-factor of backbone atoms for each CDR H3 is given, as well as the difference between this average and the average B-factor of backbone atoms for the whole F_{AB}, coloured magenta if H3 has a higher than F_{AB}-average B-factor and green for lower than F_{AB}-average. The rmsd of each conformation from the liganded conformation was calculated only with Cα atoms.

Crystal	PDB *Chain ID*	Residue order	CDR H3 Avg. B-factor	(CDR H3 Avg. B) - (F_{AB} Avg. B)	RMSD from liganded conformation
S25-2					
Liganded	3T4Y	ARDHDGYYERFSY	16.10	-2.73	0
Unliganded 1	1Q9K	ARDHDGYYERFSY	73.24	34.67	4.081
Unliganded 2	1Q9L				
	B	ARDHDGYYERFSY	70.02	30.05	4.321
	D	ARDHDGYYERFSY	76.10	35.40	3.551
Unliganded 3	XXXX*				
	H	ARDHDGYYERFSY	45.43	5.08	4.124
	B	ARDHDGYYE--SY	53.53	2.83	2.304
	D	ARDHDGYYERFSY	34.89	-4.34	4.080
	F	ARDHDGYYERFSY	52.75	11.09	3.512
Unliganded 4	XXXX				
	H	ARDHD----RFSY	57.11	16.59	3.140
	B	ARDHDGYYERFSY	44.82	1.89	3.628
	D	ARDHDGYYERFSY	46.21	-0.66	3.627
	F	ARDHDGYYERFSY	30.47	-11.49	3.641
S25-39					
Liganded	3OKL	ARDHDGYYERFAY	25.25	-0.47	0
Unliganded 1	3OKM	ARDHDGYYERFAY	31.72	-6.21	3.573
4-MeO-Kdo	XXXX	ARDHDGYYERFAY	37.94	10.8	0.299
Unliganded 1	XXXX	ARDHDGYYERFAY	9.09	-3.98	3.554
Unliganded 2	XXXX	ARDHDGYYERFAY	14.49	-1.21	0.428
Unliganded 3	XXXX				
	H	ARDHDGYYERFAY	38.51	6.36	6.061
	B	ARDHDGYY--FAY	38.76	11.20	4.493
	D	ARDHDGYYERFAY	41.09	13.70	2.195
	F	ARDHD-----FAY	41.58	11.57	0.979

*XXXX indicates unpublished structures that have not yet been deposited in the protein data bank

Table 16: CDR H3 classifications of S25-2 and S25-39 structures

Several characteristics and classifications of CDR H3 structure are given for each observed conformation of S25-2 and S25-39. The PyIgClassify cluster is determined using the PyIgClassify webserver[19]. The H3 Ramachandran sequence is determined according to Figure 7, and the residues composing the H3 anchor (as defined in section 1.4.2) are coloured orange. The similarity to North *et al.* (82) anchor clusters were determined by treating B/P, A/D and L/G as equal when comparing Ramachandran sequences with Table 7.

Crystal	PDB	PyIgClassify cluster	θ_{base}*	Anchor type**	Ramachandran	Similarity to anchor cluster
S25- 2						
Liganded	3T4Y	H3-13-2	19.437	K	BPBPDLDBBBBAB	1
Unliganded 1	1Q9K	H3-13-3	139.626	E	BPBDBAAPBBGLB	NA
Unliganded 2	1Q9L					
	B	H3-13-3	-176.293	E	BPBBALAPPBBPB	5
	D	H3-13-3	-165.203	E	BPBBALAPPBBPB	5
Unliganded 3	XXXX					
	H	H3-13-3	-172.776	E	BPBBPAABPBPPB	5
	B	H3-11-1	NA	E	BPBPPABB----B	NA
	D	H3-13-3	-172.644	E	BPBBALAPPBPPB	5
	F	H3-13-cis7-1	140.757	E	BPBBBPPPBPAPB	NA
Unliganded 4	XXXX					
	H	N/A	163.044	E	BPDG------BBB	NA
	B	H3-13-3	153.742	E	BPBBALAPPBBBB	5
	D	H3-13-3	152.290	E	BPBBALAPPBBBB	5
	F	H3-13-3	153.290	E	BPBBALAPPBBBB	5
S25-39						
Liganded	3OKL	H3-13-2	30.231	K	BPBPDLDPBBBAB	1
Unliganded 1	3OKM	H3-13-2	171.976	E	BPBBPAADAAAPB	NA
4-MeO-Kdo	XXXX	H3-13-2	43.099	K	BPBPDLDPBBBAB	1
Unliganded 1	XXXX	H3-13-2	166.438	E	BPBBPAADAAAPB	NA
Unliganded 2	XXXX	H3-13-2	32.225	K	BPBPDLDPBBBAB	1
Unliganded 3	XXXX					
	H	H3-13-3	-135.295	E	BPABBGPAPDPPB	NA
	B	H3-11-2	NA	E	BPBPBG-----PB	NA
	D	H3-13-1	130.53	E	BPBPBPAAPBPDP	1
	F	NA	NA	E	BPPG------AP	NA

* θ_{base} is the pseudo-dihedral angle between the Cα atoms from (n – 2) to (n + 1)

** The anchor is classified as kinked (K) if θ_{base} is within range of -100° to 100°, otherwise it is classified as extended (E).

[19] http://dunbrack2.fccc.edu/PyIgClassify/default.aspx (84)

2.1.7. Comparison of S25-39 and S25-2 sequence and structures

S25-2 possesses 4 mutations away from the germline sequence, whereas S25-39 possesses 6 (Table 17) (134, 136, 140). The two antibodies do not share any mutations and possess an identical binding site with the exception of residues H53, which is the germline Asn in S25-2 and Lys in S25-39, as discussed in section 1.8.4, and residue H101 that is Ala in S25-39 and Ser in S25-2. H101 is located at the base of CDR H3 on the outside of the binding site and makes no contact to antigen. Ser H101 of S25-2 participates in a hydrogen bond network with Arg H94, Asn H95, Arg H100B, Glu L55 and several coordinated water molecules (Figure 34).

The crystal structures of S25-2 and S25-39 display ten unique[20] unliganded conformations of CDR H3, none that are shared between the two antibodies, and one unique conformation used by both antibodies to bind several Kdo-based antigens (Figure 33) (134, 140).

[20] A unique conformation is defined by a unique Ramachandran code, Table 16. Side-chain positions are not considered.

Table 17: Amino acid sequence alignment of S25-2 and S25-39 variable regions

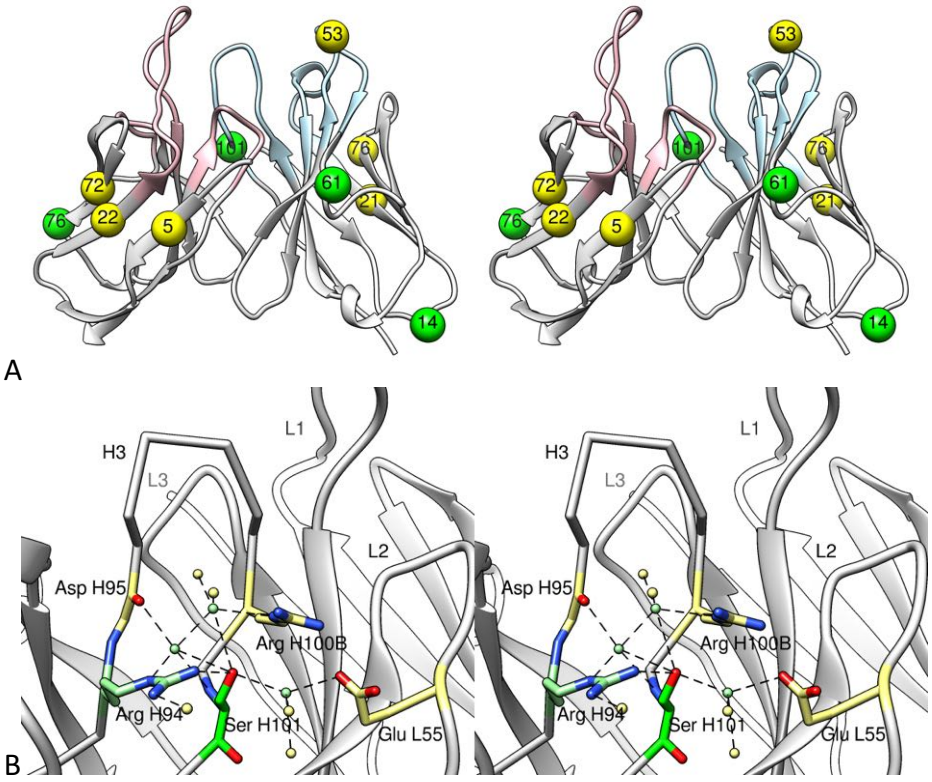
Amino acid sequence alignment of S25-2 and S25-39 with germline, broken into framework and CDR regions. Germline sequence is composed from the mouse gene segments IGKV8-21*01 and IGKJ1*01 for the light chain and IGHV7-3*02, IGHJ3*01 and IGHD2-3*01 for the heavy chain[21]. Amino acids differing from germline sequence are highlighted green for S25-2 and Yellow for S25-39. Kabat numbers of mutated residues are given below the alignment.

	LFR1	CDR L1	LFR2	CDR L2
Germline	DIVMSQSPSSLAVSAGEKVTMSC	KSSQSLLNSRTRKNYLA	WYQQKPGQSPKLLI	YWASTRES
S25-2	DIVMSQSPSSLAVSAGEKVTMSC	KSSQSLLNSRTRKNYLA	WYQQKPGQSPKLLI	YWASTRES
S25-39	DIVMTQSPSSLAVSAGEKVTMNC	KSSQSLLNSRTRKNYLA	WYQQKPGQSPKLLI	YWASTRES
	5 **22**			
	LFR3	**CDR L3**	**LFR4**	
Germline	GVPDRFTGSGSGTDFTLTISSVQAEDLAVYYC	KQSYNLRT	FGGGTKLEIK	
S25-2	GVPDRFTGSGSGTDFTLTITSVQAEDLAVYYC	KQSYNLRT	FGGGTKLEIK	
S25-39	GVPDRFTGSGSGTDFALTISSVQAEDLAVYYC	KQSYNLRT	FGGGTKLEIK	
	72 **76**			
	HFR1	**CDR H1**	**HFR2**	**CDR H2**
Germline	EVKLVESGGGLVQPGGSLRLSC	ATSGFTFTDYYMS	WVRQPPGKALEWLG	FIRNKANGYTTE
S25-2	EVKLVESGGGLVQSGGSLRLSC	ATSGFTFTDYYMS	WVRQPPGKALEWLG	FIRNKANGYTTE
S25-39	EVKLVESGGGLVQPGGSLRLAC	ATSGFTFTDYYMS	WVRQPPGKALEWLG	FIRNKAKGYTTE
	14 **21**			**53**
	HFR3	**CDR H3**	**HFR4**	
Germline	YSASVKGRFTISRDNSQSILYLQMNTLRAEDSATYYC	ARDHDGYYERFAY	WGQGTLVTVS	
S25-2	YSPSVKGRFTISRDNSQSILYLQMNTLRAEDSATYYC	ARDHDGYYERFSY	WGQGTLVTVS	
S25-39	YSASVKGRFTISRDNSQSSLYLQMNTLRAEDSATYYC	ARDHDGYYERFAY	WGQGTLVTVS	
	61 **76**	**101**		

[21] Determined using IMGT/V-QUEST (http://www.imgt.org/) (50)

Figure 34: Amino acid differences of S25-2 and S25-39

(A) Locations of amino acid mutations from the germline of S25-2 (green spheres) and S25-39 (yellow spheres) are shown on a ribbon diagram of the F_V. Mutant residues are numbered, and CDRs are coloured as in Table 17. **(B)** Shows the hydrogen bond network of S25-39 Ser H101, green, with direct hydrogen-bonded residues and waters in light green, and second-shell hydrogen-bonded residues and waters in light yellow.



2.2. Discussion

2.4.1. Affecting antibody conformations with synthetic antigens

To explore the binding potential of previously observed unliganded conformations of S25-2 and S25-39, synthetic antigens were designed that would inhibit interactions with the original antigen-binding conformation but remain compatible with the unliganded conformations. The first compound synthesized and tested was 4-MeO-Kdo, which was successfully crystallized in complex with S25-39. The 4-MeO substituent was intended to abolish the hydrogen bond with Glu H100A seen in the Kdo complex structures, and to be accommodated by a hydrophobic cavity as observed in the S25-2 unliganded #1 structure (Figure 21). This was somewhat successful. In the complex structure with 4-MeO-Kdo, CDR H3 still adopts the same liganded backbone conformation but Glu H100A is disordered and B-factors are elevated for the remainder of the CDR (Figure 23, Table 15). This indicates that the ligand does abolish the interaction with Glu H100A, but binding to this backbone conformation is still more favourable than an alternative.

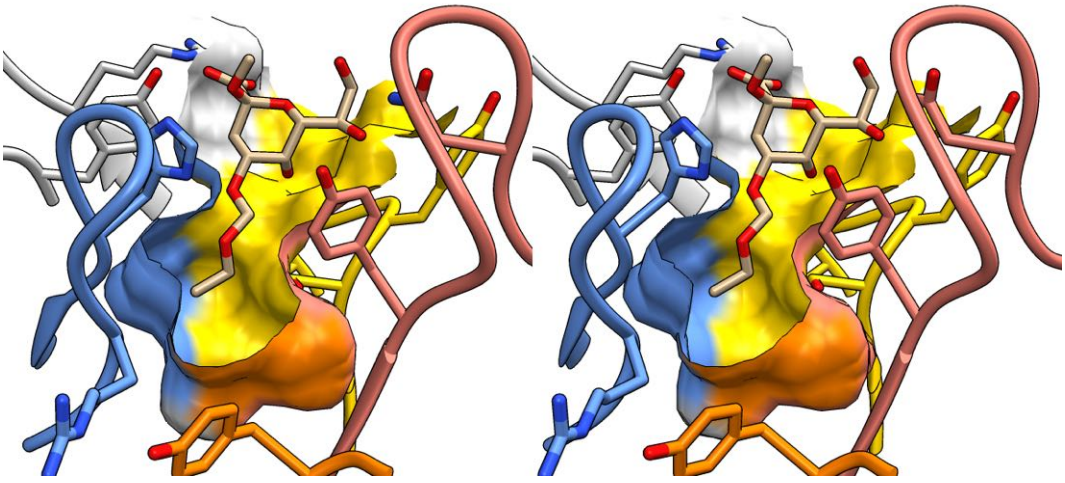
To enhance the observed conflict between 4-MeO-Kdo and the liganded CDR H3 conformation, compounds with larger substituents, 4-O-benzyl-KdoOMe, 4-O-ethoxymethyl-KdoOMe and 4-O-methoxymethyl-KdoOMe, were synthesized and used for co-crystallization with S25-2 and S25-39. Many crystals were obtained from these screens, but while no bound antigens were seen in any of the resulting structures several new unliganded conformations were observed (Figure 32, Figure 33). It is possible that persistent weak interactions with these ligands shift the conformational

equilibrium of antibody during crystal formation to allow stabilization in new conformations, or alternatively cause the disorder observed in some structures.

In one structure from a crystal grown with 4-O-ethoxymethyl-KdoOMe, S25-39 Unliganded #3, there is ambiguous weak density in one molecule out of four in the AU where the ligand can be modeled (Figure 25). Although the electron density for ligand is inadequate to conclude binding, there is complete electron density for the CDR H3 backbone of this molecule in a unique conformation that would accommodate 4-O-ethoxymethyl-KdoOMe, with the 4-O-ethoxymethyl moiety in a cavity formed between CDR H3 and all three light chain CDRs (Figure 35). Regardless of the quality of electron density for ligand, the observation of this conformation that is compatible with ligand binding may suggest that the presence of ligand increased the relative population of this conformation to a level sufficient for crystallization.

Figure 35: Modelling 4-O-ethoxymethyl-KdoOMe into the S25-39 Unliganded #3 chain C/D binding pocket

A partial surface of the binding pocket is shown coloured according to contributions from CDR H3 blue, L1 salmon, L2 orange, and L3 yellow. 4-O-ethoxymethyl-KdoOMe was modelled into the ambiguous electron density of the binding site and refined as shown in Figure 25.



2.4.2. Robust conservation of V-gene encoded carbohydrate recognition

Maintenance of binding to 4-MeO-Kdo despite the disruption of a specific interaction and destabilization of CDR H3 again highlights the amazing versatility of this V-gene conserved Kdo binding pocket. In addition to its ability to accommodate alternate lengths, linkages and unnatural modifications through flexible recognition as described in section 1.8.3, this pocket is also able to accommodate unnatural modifications that actively disrupt specific interactions with CDR H3. CDR H3 is well known to influence the specificity of antibodies (174, 227, 228), such as described above for the S25-2 type antibodies where unique interactions *via* CDR H3 are involved in recognizing most antigens and direct the binding towards specificity or cross-reactivity (section 1.8.5).

Given the importance of CDR H3 in determining antibody specificity it may be surprising at first that 4-MeO-Kdo is still bound by S25-39, but this is consistent with the phenomena of V-gene restriction and conservation of carbohydrate recognition in the germline; It emphasizes the ability of this conserved Kdo-recognition pocket to adapt to novel modifications and expand the recognition potential of the germline antibody repertoire.

2.4.3. Conformational flexibility increases recognition potential in the germline antibody repertoire

The germline antibody repertoire must display a degree of polyspecificity or cross-reactivity to protect against the enormous antigenic diversity it may encounter. This is possible through differential epitope or paratope usage in a rigid antibody combining site, or through a flexible antibody combining site that generates unique paratopes.

The near-germline antibodies S25-2 and S25-39 display both of these mechanisms for generating recognition potential. The shared antigen-binding conformation of S25-2 and S25-39 is cross-reactive for a number of distinct Kdo-based oligosaccharides, where a single Kdo moiety is bound in a germline-encoded pocket and additional lengths, linkages, and unnatural modifications are accommodated in the remainder of the binding site (Figure 12). As shown here with the structures of S25-39 in complex with 4-MeO-Kdo, the germline pocket can even accommodate unnatural additions that directly impede binding contacts.

In addition to this cross-reactivity using a single conformation of CDR H3, the unliganded structures presented here display a range of alternate CDR H3 conformations that generate unique combining sites that may allow recognition of unrelated antigens (Figure 32, Figure 36). Indeed, there is some evidence for binding to 4-O-ethoxymethyl-KdoOMe through an alternate conformation as discussed above.

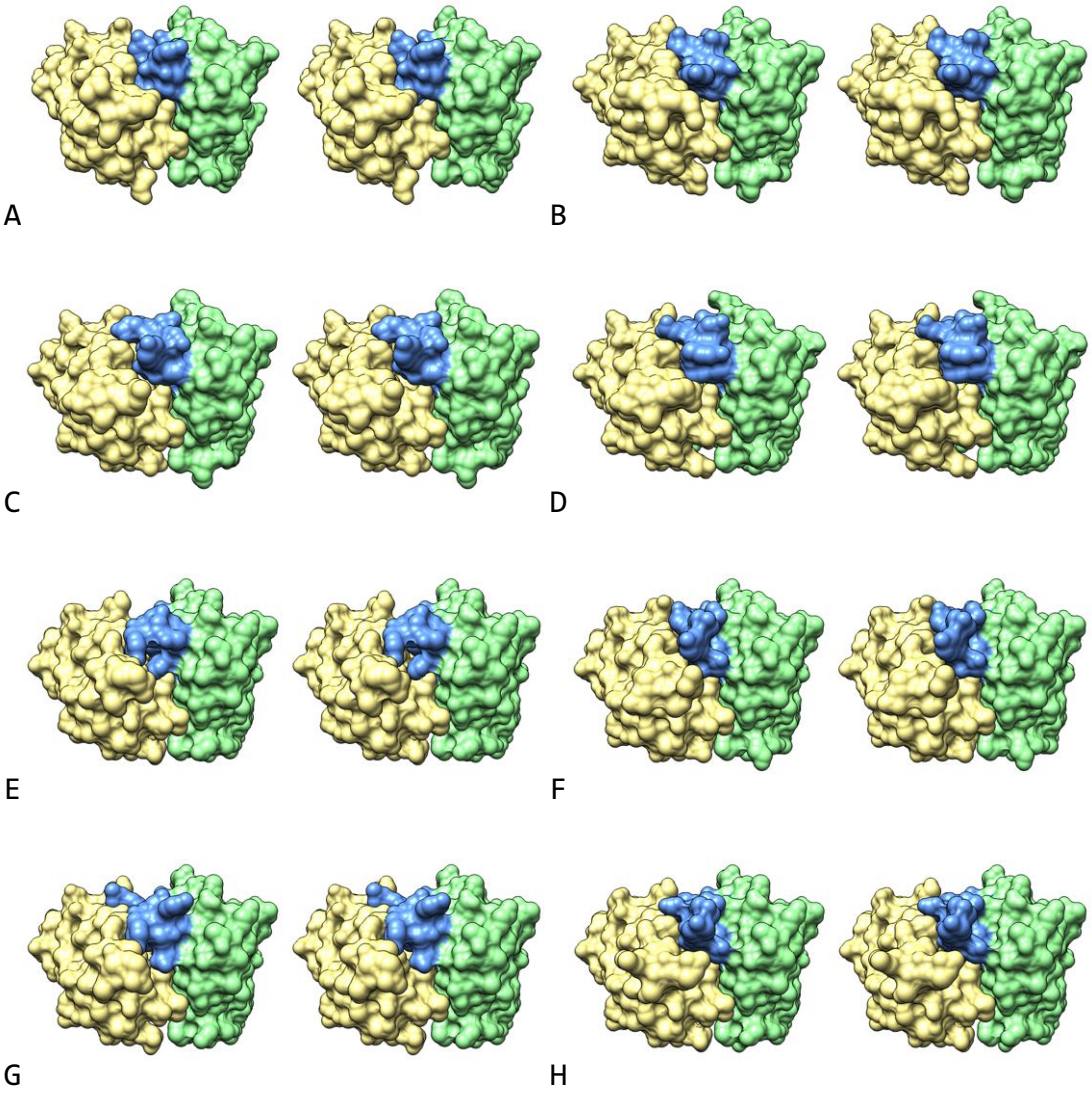
There is a precedent for this type of polyspecificity, as was observed for the antibodies SPE7 (161), 7G12 (160), DNA-1 (130), BBE6.12H3 (163) and bH1 (164, 271) that each recognized unique antigens through different binding site conformations (Figure 37). S25-2 and S25-39 are unique in demonstrating both the largest number of distinct conformations[22], 7 for S25-2 and 4 for S25-39, with the next highest example being 5 for BBE6.12H3, and the largest maximum change in Cα position between conformations, 12 Å for S25-39, compared to the next largest, 10.8 Å for 7G12 (Figure 32, Figure 37).

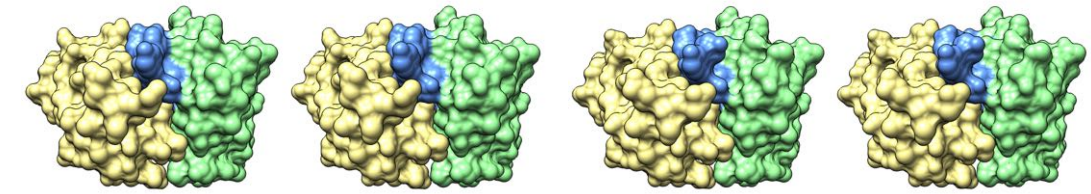
[22] A unique conformation is defined by a unique Ramachandran code, Table 16. Side-chain positions are not considered.

S25-2 and S25-39 have been shown to exhibit cross-reactivity through single binding site conformations and to also exhibit an unprecedented level of structural flexibility, which may allow polyspecificity. These antibodies clearly demonstrate how a single antibody may protect against a large wedge of antigenic space and reconcile the limited size of the germline antibody repertoire with a potentially limitless diversity of antigens.

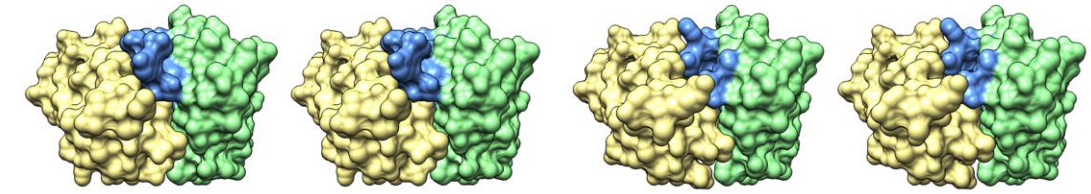
Figure 36: Binding surfaces of alternate S25-2 and S25-39 conformations

Binding surfaces generated by different conformations of CDR H3, with heavy chain green, light chain yellow and CDR H3 blue, are shown for **(A)** S25-2 liganded with Kdo **(B)** S25-2 Unliganded #1 **(C)** S25-2 Unliganded #2 Chains A/B **(D)** S25-2 Unliganded #2 Chains C/D **(E)** S25-2 Unliganded #3 Chains A/B **(F)** S25-2 Unliganded #3 Chains C/D **(G)** S25-2 Unliganded #3 Chains E/F **(H)** S25-2 Unliganded #3 Chains H/L **(I)** S25-2 Unliganded #4 Chains A/B **(J)** S25-2 Unliganded #4 Chains C/D **(K)** S25-2 Unliganded #4 Chains E/F **(L)** S25-2 Unliganded #4 Chains H/L **(M)** S25-39 Unliganded #1 **(N)** S25-39 Unliganded #2 **(O)** S25-39 Unliganded #3 Chains C/D and **(P)** S25-39 Unliganded #3 Chains L/H.

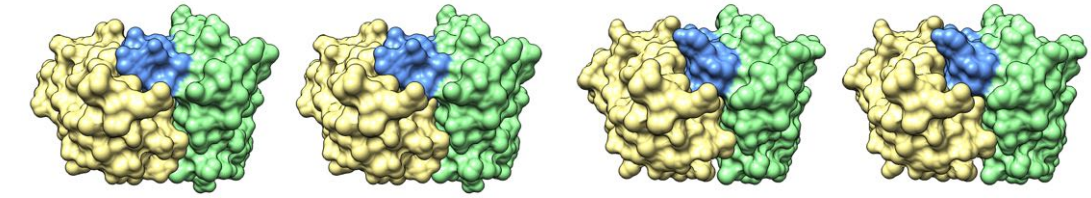




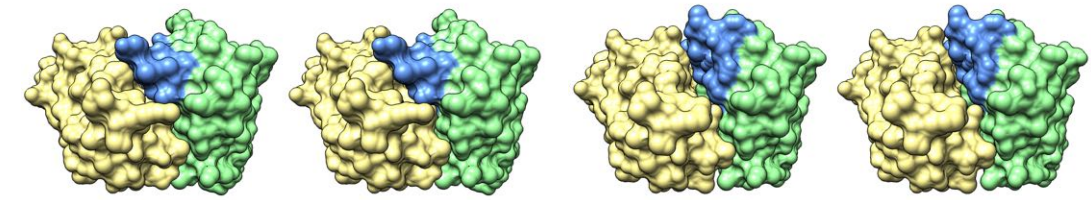
I J



K L



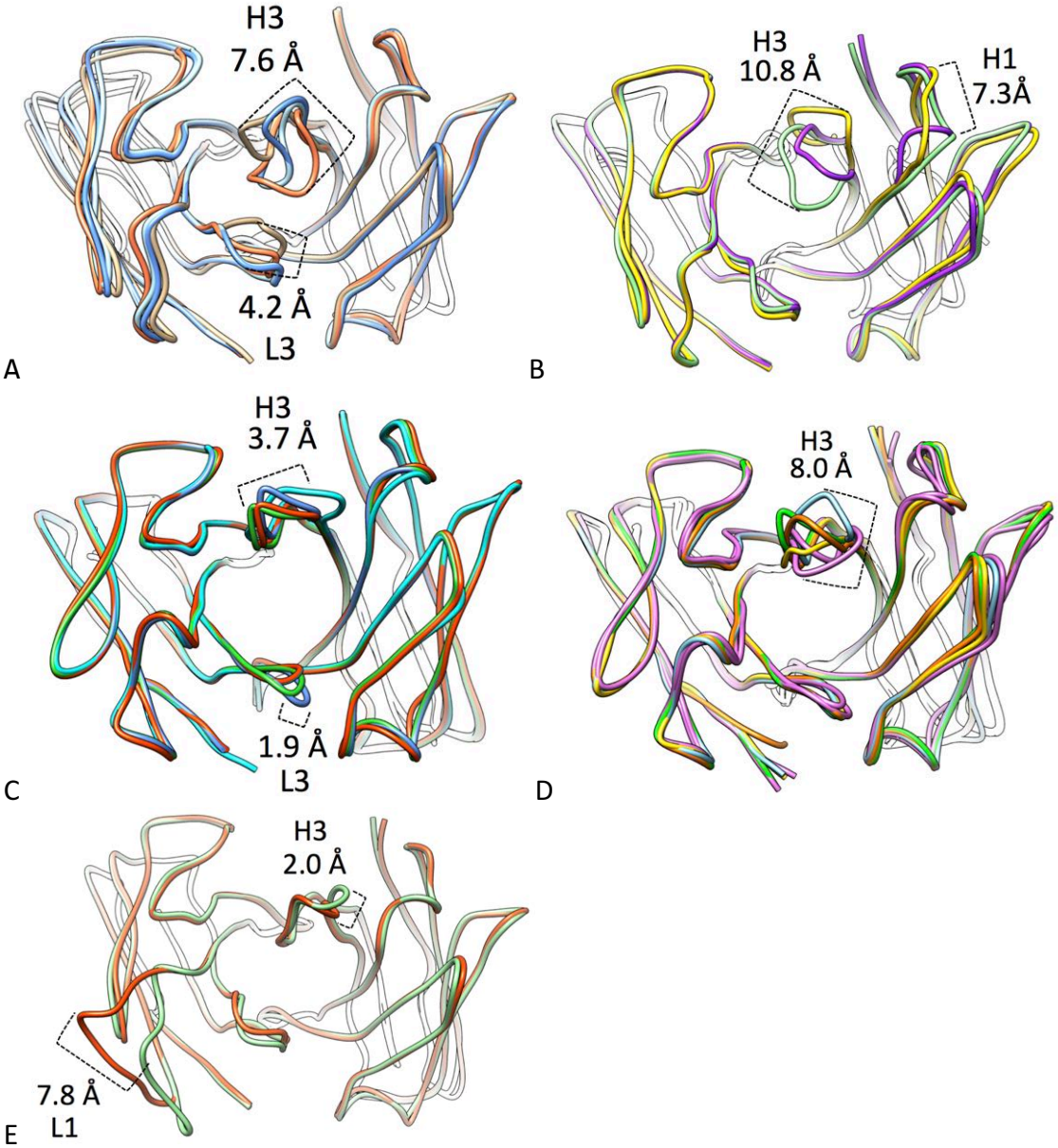
M N



O P

Figure 37: Structural changes of polyspecific antibodies

Maximum Cα displacement of CDR structural changes observed in the polyspecific antibodies **(A)** DNA-1 in complex with ssDNA (1I8M chains AB tan, chains HL orange) and with HEPES (1P7K chains AB light blue, chains HL medium blue), **(B)** 7G12 unliganded (1NGZ green) and in complex with Jeffamine (1NGX yellow), and NMP (1N7M purple), **(C)** SPE7 unliganded (1OAQ cyan) and in complex with Az (1OAU blue), DNP-Ser (1OAR green), and Trx-Shear3 (1OAZ red), **(D)** BBE6.12H3 unliganded (4A6Y purple) and in complex with four different peptides (2Y07 blue, 2Y06 yellow, 2Y36 green, and 2XZQ orange), and **(E)** bH1 in complex with HER2 (3BE1 green) and VEGF (3BDY orange). Light and heavy chain CDR displacements were calculated after alignment of either chain individually, whereas an average alignment of both chains is shown in the figure.



2.4.4. Affinity maturation effects germline conformational repertoire

As discussed in section 1.5.3, conformational flexibility is relevant to the process of affinity maturation, as this process has been shown to reduce antibody flexibility to stabilize a specific antigen-binding conformation, which can increase both affinity and specificity. The effect of amino acid mutations on the conformational repertoire of CDR H3 is demonstrated by S25-2 and S25-39. These antibodies differ by only a single amino acid Ser/Ala H101 at the base of CDR H3, yet were not crystallized in any of the same unliganded conformations. This suggests that the amino acid differences between these antibodies, potentially also including those distant from the binding site, significantly alter their conformational ensembles. Ser H101 of S25-2 interacts with four nearby amino acids either directly or through bridging water molecules (Figure 34), so it is understandable that the Ser/Ala difference may alter the hydrogen bonding network at the base of CDR H3 and influence its conformation.

Interestingly, S25-39 possesses the germline Ala H101 in CDR H3 and displays the largest Cα displacement between conformations (Figure 33), suggesting that the germline sequence may allow a greater magnitude of structural change. In contrast, the conformations displayed by S25-2 are more similar to one another, but a larger number were observed. Both antibodies utilize the same antigen-binding conformation, indicating that affinity maturation can alter the conformational equilibrium to simultaneously preserve or optimize the antigen binding conformation while reducing or altering potentially polyspecific conformations.

2.4.5. S25-39 displays exact conformational selection in antigen binding

The closely related antibodies S25-39 and S25-2 have nearly identical binding sites, with a CDR H3 that has now been observed in ten unique[23] unliganded conformations. One of these well-ordered unliganded conformations is also the active antigen-binding conformation of both antibodies (Figure 32, Figure 33). The several conformations observed with good electron density suggest that these exist as relatively low-energy conformations in equilibrium that are stabilized by specific crystallization conditions. The partially disordered conformations may indicate the stabilization of partial conformations with a number of similar energy variants so that no single full conformation was preferentially stabilized in the crystals. This variety of conformations with different apparent stabilities is consistent with the 'rugged energy landscape' interpretation of proteins (244), where these antibodies exist in an equilibrium of conformations with different energy levels.

Although crystal packing may be a factor affecting the stability and conformation of flexible loops in crystals, the observation of the exact antigen-binding conformation in the absence of ligand with good electron density (S25-39 Unliganded #2, Figure 24) validates the possibility that other unliganded conformations are also biologically relevant and not artefacts of crystal packing. Supporting this conclusion is that the S25-39 Unliganded #1 and #2 structures as well as the structure of S25-39 in complex with 4MeOKdo are in the same space group with very minor differences in unit cell dimensions and packing arrangements (Appendix A), as are the S25-2 Unliganded #3 and

[23] A unique conformation is defined by a unique Ramachandran code, Table 16. Side-chain positions are not considered.

#4 structures. Additionally, S25-39 structures displaying the Unliganded #1 and #2 conformations were solved from crystals grown in identical conditions, and multiple structures displaying each conformation were also solved from crystals grown in unique conditions (Appendix A). These results indicate that a given conformation is not possible *only* in a certain condition or crystal packing context, and likewise, that a certain condition is not capable of stabilizing *only* a single conformation for crystallization. Rather, the crystallization condition may affect the probability of nucleation of a specific conformation into a crystal, and it is the conformation that dictates the crystal packing, not *vice versa*.

This collection of structures illustrates a conformational selection mechanism of complex formation where a Kdo-based antigen selects the appropriate CDR H3 conformation from pre-equilibrium. Significantly, the conformational selection illustrated by these structures is absolute; there is no evidence of conformational selection followed by induced-fit adjustment, as the exact liganded conformation is observed in the absence of ligand. This is in contrast to the only other structural characterization of conformational selection of antibody SPE7 (161). In the analysis of mAb SPE7 structures, James *et al.* presented a case for conformational selection in binding both protein and aromatic hapten through distinct binding-site conformations. This antibody was observed in two unique antigen-free conformations, and the existence of an isomeric pre-equilibrium was supported with pre-steady-state kinetics. James *et al.* argued that the two observed unliganded conformations formed the initial encounter complexes with either hapten or protein antigen, and subsequent induced fit

generated the final antigen bound conformations. The conformational selection between the initial unliganded structures was largely based on a Trp side-chain rotamer of CDR L3, where only one conformation is compatible with hapten binding. However, large structural differences of CDR H3 were observed between both unliganded and liganded structures (Figure 37). Pre-steady-state binding kinetics supported a binding mechanism involving conformational selection of CDR L3 followed by induced-fit isomerization of CDR H3.

Crystal structures of the polyspecific antibody 7G12 also revealed conformation differences between the unliganded and two different liganded states. The authors suggest that each ligand induces significant structural changes of the unliganded antibody binding site (160), but with only a single unliganded structure no distinction between CS and IF can be made.

Antibody BBE3.12H3 was observed to bind four unique peptides through four distinct binding site conformations, each different than an observed unliganded conformation. Again with only a single unliganded structure, no comment on the contributions of conformational selection versus induced fit can be made. Likewise, the antibody DNA-1 was shown to bind ssDNA and HEPES through alternate CDR H3 conformations (130) that were different from two disordered unliganded conformations (158).

In contrast to all of these examples, antibodies S25-2 and S25-39 display many unique antigen-free conformations including the exact antigen-binding conformation in the absence of antigen observed for S25-39, which indicates pure conformational

selection in binding. These structures suggest that the antibodies exist in an equilibrium of conformations that each may be stabilized by specific crystallization conditions or upon antigen binding.

Pre-equilibrium and conformational selection have been suggested previously as a mechanism for Ab-Ag binding and for allowing polyspecificity to expand the recognition potential of the germline repertoire, but structural evidence for these phenomena is generally lacking. As discussed by Sela-Culang *et al*. (229), investigation of unbound antibody structures is less common then the corresponding Ab-Ag complexes that tend to have greater biological impact. Furthermore, those studies that do include unbound structures generally only include a single one, which provides only a snapshot of a potentially dynamic equilibrium of unbound states. Finally, antibodies that do exhibit significant structural flexibility (i.e. germline antibodies) are inherently less amenable to crystallization, which biases the database towards less flexible structures (i.e. affinity-matured antibodies). This study of S25-2 and S25-39 represents the first targeted investigation of unbound antibody conformations and provides unique structural insight into the dynamic conformational equilibrium of unbound Ab, including direct evidence for conformational selection in Ag binding.

The presence of the exact antigen-binding conformation pre-existing in equilibrium may also reflect a significant general feature of anti-carbohydrate antibodies. Carbohydrates are commonly T-cell independent antigens, meaning that they can activate B-cells without T-cell help. As a result, there is no somatic hypermutation of the immunoglobulin locus and no affinity maturation of the antibody.

This suggests that the recognition of important carbohydrates (e.g. from pathogens with which the immune system has co-evolved) has been optimized in the germline to not depend on affinity maturation. In the context of a flexible antibody, this may be achieved through the evolution of a conformational equilibrium that includes the exact antigen-binding conformation, as a compromise between enabling germline polyspecificity through alternate conformations and reducing the energetic penalty of induced-fit binding. In contrast, germline antibodies may recognize T-cell dependent antigens through less favourable multi-step selection-and-adjustment mechanisms since these interactions can be subsequently optimized through affinity maturation.

2.4.6. Ion coordination alters binding site

An interesting case is presented by the Unliganded #4 structure of S25-2 that showed a unique CDR H3 conformation with three coordinated iodide ions (Figure 31). The same conformation and coordination pattern was observed in each of the four molecules of the AU, suggesting an energetic preference for that configuration or its specific stabilization by the crystallization conditions. One of these three ions is coordinated at the base of CDR H3 underneath the antibody binding site (Figure 38), highlighting the extreme flexibility of CDR H3 required for the ion to enter this site. The solvent-excluded location of this iodide further indicates that this CDR H3 conformation cannot be selected by iodide ions, but must be achieved either through induced fit or through a sequence of selection and adjustments steps. None of the additional 22 locations of coordinated iodide ions in the structure showed altered protein conformations compared to structures without iodide.

The effects of ions on protein structure and stability has long been a topic of interest, pioneered in the late 1800s by Franz Hofmeister, who analysed the effects of different cations and anions on the precipitation of egg white proteins to develop what would later be called the Hofmeister series (272). The mechanisms of action of kosmotropes, which give higher-order structure to water, and chaotropes, which disrupt water structure, have since been extensively studied and are still debated (272–290). These two classes of ions "salt-out" or "salt-in" proteins, respectively, meaning to decrease or increase their solubility.

Iodide is a large weakly hydrated anion (a chaotrope) known to interact with weakly hydrated positive portions of proteins like Arg, Lys, His, and backbone amides, as well as well as nonpolar surfaces (278, 284). Iodide increases protein solubility as a chaotrope, but it is also generally denaturing due to its extensive interactions with proteins, notably the peptide backbone. However, given the possibility of specific interactions, it has been observed that bound iodide can stabilize or destabilize proteins depending on the context and on ion concentration (291).

The structure of S25-2 from a crystal grown in ammonium iodide (unliganded #4) shows 34 ordered iodide ions, coordinated in a variety of contexts including charged, uncharged, polar and apolar side chains, and main chain atoms. Specifically in CDR H3, the interactions with iodide leading to the same configuration in all four molecules of the AU is an example of a stabilizing iodide interaction. Similarly, Iodide ions have previously been observed to effect unique conformations of a flexible region in crystal structures of lysozyme (277), and iodide has been shown to increase the thermal

stability of lysozyme at low concentrations and decrease stability at higher concentrations (291). This may reflect a general mechanism for the potential of iodide at certain concentrations to stabilize flexible regions of a protein and by extension the stability of protein in solution, as disordered regions are known to be susceptible to aggregation (292–296).

Only two other antibody fragments have been crystallized in the presence of iodide ions. One, the nanobody[24] 7D12, was crystallized in complex with domain III of the extracellular region of EGFR. A single iodide ion interacts with both Nanobody and ligand, but distant from the major binding interface and with no structural differences from a complex structure lacking iodide (297).

The second antibody fragment, an F_{AB} of the catalytic antibody 15A9, was crystallized with its hapten 5'-phosphopyridoxyl (PPL)-N^{ε}-acetyl-L-lysine as well as pyridoxal-5'-phosphate-L-alanine (PPL-L-alanine) and PPL-D-alanine (298). In all three structures CDR H3 was observed to coordinate an iodide ion. Interestingly, the structures still showed bound ligands despite the direct contact of iodide ions with CDR H3. Furthermore, the transaminase activity of 15A9 in solution was decreased by only 42% with the addition of sodium iodide at the concentration used for crystallization (0.2 M). Unfortunately, a structure of 15A9 without iodide is not available for comparison.

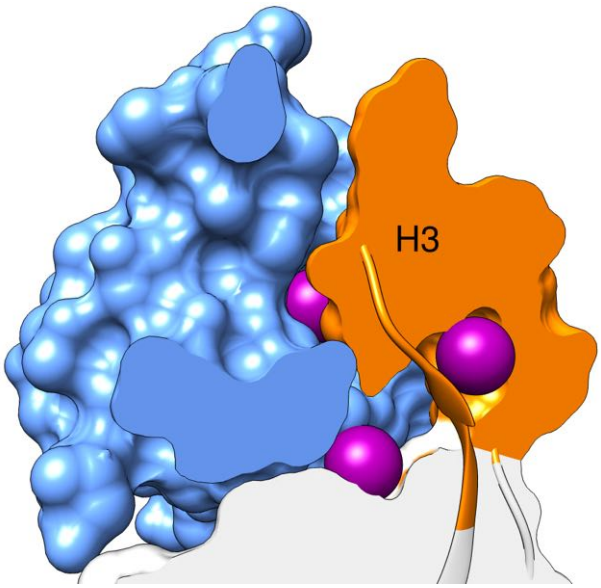
As the CDR H3 of S25-2 and S25-39 can be present in an equilibrium of unique conformations as illustrated by the structures presented in this thesis, it provides a significantly larger number of potential configurations for ion coordination. This may be

[24] A Nanobody, or VHH domain, consists only of the heavy chain variable domain.

extrapolated to all antibodies with flexible binding sites to suggest that interactions with ions may further increase the number of unique paratopes available. As seen with antibody 15A9 described above, iodide-influenced paratopes can be functional in antigen-binding, so these paratopes present an additional avenue for generating antibody polyspecificity. These effects may not be relevant at physiological conditions where the total iodide ion concentration in blood is typically in the nanomolar range (299, 300), but may be significant in molecular biology applications where ion concentrations are much higher.

Figure 38: CDR H3 buried iodides in S25-2 Unliganded #4

A clipped surface of S25-2 Unliganded #4 is shown with the light chain blue, CDR H3 orange and the rest of the heavy chain light grey. Iodide ions are shown as purple spheres. A ribbon representing CDR H3 and flanking residues is also shown.



2.4.7. Conformational classification and prediction of CDRs

Rather than viewing antibodies, enzymes or other proteins as rigid bodies with single conformations, considering their dynamic conformational equilibria has significant implications in structure prediction and structure-based drug design (244). For antibodies, this directly affects the prediction of CDR conformations based on sequence and canonical conformations.

Considering the evolutionary pressure for diversity in the antibody repertoire, it may seem counter-intuitive that five out of six CDR loops often adopt common canonical structures. However, as these loops are almost entirely encoded by single germline gene segments (with the exception of CDR L3) that were generated by gene duplication (20, 21, 185, 301), it is not surprising that they have maintained common

conformations that form functional binding sites and that the vast majority of variability in the antibody response instead is generated by CDR H3 encoded by the hypervariable V-D-J junction.

As discussed in sections 1.4.2 and 1.4.3, the conformations of CDRs L1, L2, L3, H1 and H2 are highly predictable based on sequence, whereas the conformation of CDR H3 is not. The difficulty in CDR H3 classification and prediction is firstly due to the much larger sequence diversity of that loop, and secondly due to the tendency of individual CDR H3 sequences to generate structurally flexible loops (171). The latter feature is clearly demonstrated in the structures of S25-39 and S25-2 (Figure 32, Figure 33).

The CDR H3 loops of S25-2 and S25-39 display a variety of conformations, some of which satisfy H3 rules and are similar to canonical anchor clusters, and some which do not satisfy H3 rules and are not similar to any clusters. The shared liganded conformation and S25-39 Unliganded #2 are kinked and similar to the H3-anchor-1 cluster (Table 16). This conformation satisfies the H3-rule i-a of Kuroda *et al.* (Table 3) with (n – 1) not Asp predicting the kinked base, although all other unliganded conformations display an extended base countering this prediction. Of the remaining unliganded conformations, many of S25-2 are similar to the cluster H3-anchor-5 (Table 16) whereas those of S25-39 are not similar to any defined anchor cluster. The ability of these loops to obey the H3 rules or not and to adopt multiple canonical and non-canonical conformations, while perhaps an extreme case, stresses that flexibility must be taken into account for CDR H3 prediction. This is reflected by the greatest accuracy in CDR H3 prediction so far being achieved by techniques that do not rely on canonical

structures or rules, but instead more adaptive template selection and *ab initio* methods (105, 111, 112).

The lack of shared unliganded conformations observed between S25-2 and S25-39 despite their almost identical CDR H3 indicates another challenge for structure prediction: different environments (in this case the 10 amino acid differences between S25-2 and S25-39) may alter the available CDR H3 conformations or their populations in equilibrium. The observation of different conformations between S25-2 and S25-39 suggests that an identical CDR H3 paired with different V-gene combinations may have vastly different conformational repertoires. Again, this is reflected by the majority of CDR H3 modelling techniques being more successful when taking the protein environment of CDR H3 into account during optimization (114).

It is clear that much more research is required to understand the flexibility and dynamics of CDR H3 before these characteristics can be accurately modelled or predicted. This thesis describes the first targeted structural investigation of antibody conformational diversity, which will hopefully encourage similar investigations of unrelated antibodies that together will provide knowledge of features that allow flexibility and dictate conformation. This knowledge is a critical component alongside improvements in computer processing power and modelling methods to achieve accurate prediction of antibody dynamics and antigen binding.

2.3. Conclusions

Combining cross-reactivity and polyspecificity in a single flexible binding site is an astounding evolved dual-mechanism that allows the immune response to maintain recognition of common pathogens and also adapt to new threats, while minimizing the burden on genome size. The studies of the near-germline anti-carbohydrate antibodies S25-2 and S25-39 presented in this thesis reveal an unprecedented level of conformational diversity in CDR H3, which may enable polyspecificity through alternate conformations in addition to their robust cross-reactivity by conserved recognition of Kdo and Kdo-based synthetic antigens through a single conformation.

These structures also shed light on the modification of conformational equilibria by natural and synthetic antigens and ions, the latter of which were observed to generate unique paratopes that may heighten polyspecificity. Also revealed are the effects of a single mutation in CDR H3 and several mutations distant from the combining site on the conformational equilibria of these similar antibodies that share liganded but not unliganded conformations, echoing studies that suggest affinity maturation can proceed by moderating antibody flexibility.

Significantly, this observation of the exact Kdo-binding conformation in the absence of ligand is the first structural illustration of pure conformational selection in antigen binding. Conformational selection is an ideal implementation of antibody binding-site adaptability as it provides the potential for polyspecificity without the kinetic or entropic disadvantages of induced fit, which may be especially relevant to anti-carbohydrate antibodies that have evolved not to rely upon affinity maturation.

Together, these structures provide insight into the conformational dynamics and antigen-binding mechanisms of near-germline anti-carbohydrate antibodies. These are of fundamental importance to our understanding of antibody structure and function, leading to practical applications such as improved modelling and design of therapeutic or diagnostic antibodies.

2.4. Methods

2.4.1. Antibody F_{AB} preparation

The monoclonal antibodies S25-2 and S25-39 (both IgG1κ) were produced by my collaborators Sven Müller-Loennies, Lore Brade and Helmut Brade at Research Center Borstel, Leibniz-Center for Medicine and Biosciences, Parkallee 22, Borstel D-23845, Germany, by immunization of BALB/c mice with Kdo(2→8)Kdo(2→4)Kdo(2→6)βGlcNAc-BSA as described previously (134, 136). IgG was purified from hybridoma supernatant using 3 mL Pierce protein A/G agarose resin in a 16 mL Kontes Flex-Column gravity flow column from VWR. Hybridoma supernatant was diluted 1:1 in binding buffer (20mM NaH_2PO_4 pH 7.0) and centrifuged for 20 minutes at 13000 rpm before loading onto the column, which was pre-equilibrated with binding buffer. After loading, the column was washed with 15 mL binding buffer and eluted with 0.1 M Glycine pH 2.7 into 1 mL fractions that were immediately neutralized with 100 µl of 1 M Tris-HCL pH 8.5. Fractions were analyzed by SDS-PAGE and absorbance at 280 nm for IgG content. Fractions containing IgG were pooled and dialyzed overnight in 50 mM Hepes pH 7.5. Trial papain digests were performed in 150 µl reactions with 0.1 M Hepes pH 7.5, 2 mM DTT, 2 mM EDTA and 1 mg/ml IgG, with Papain-to-IgG wt:wt ratios of 1:500 and 1:1000, and 30 µl aliquots were removed and quenched with 10 mM Iodoacetamide at 0.5, 1, 2, 3 and 4 hours. Aliquots were run on SDS-PAGE to analyze digestion of IgG into F_{AB} and F_C. A batch digest was then performed on remaining dialyzed IgG at a Papain:IgG ratio of 1:400 for 4 hours, then quenched with 10 mM Iodoacetamide and dialyzed overnight into 20 mM Hepes pH 7.5. The F_{AB} fragments yielded by papain digest were then

purified by HPLC cation-exchange chromatography using a Shodex CM-825 column, with 20 mM Hepes pH 7.5 and a 0 to 1 M NaCl gradient. The F_{AB} peak from multiple runs was pooled and buffer-exchanged into 20 mM Hepes pH 7.5 and concentrated to between 20 and 40 mg/mL using 4 mL Amicon ultra centrifugal filtration devices with a 10 kDa cutoff.

2.4.2. Crystallization of S25-39 and S25-2 Fab

S25-39 and S25-2 F_{AB} were crystallized unliganded or in the presence of the synthetic ligands 4-MeO-Kdo, 4-O-benzyl-KdoOMe, 4-O-Ethoxymethyl-KdoOMe and 4-O-Methoxymethyl-KdoOMe (synthesized by my collaborators Barbara Pokorny and Prof. Paul Kosma at the Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria), each with 12 mg/ml F_{AB} and 5 mM ligand, using Hampton Crystal Screens 1 and 2 in Hampton 24-well hanging-drop vapour diffusion plates set up by hand, and using Qiagen Pegs I and Pegs II screens in Hampton 96-well sitting-drop intelli-plates set up with an ARI Crystal Gryphon robot. Initial hits were optimized in 35 x 10 mm tissue culture plates by hanging-drop vapour diffusion. The conditions yielding crystals from which structures were solved are given in Table 18 and Table 19 of Appendix A.

2.4.3. Data collection, Structure Determination and Refinement

Before freezing, crystals were dipped in reservoir solution with the addition of 30 % MPD for cryoprotection (Table 18 and Table 19, Appendix A). Crystals were flash frozen to -160°C using an Oxford Cryosystems Cryostream 700 crystal cooler. Data were

collected either on a Rigaku R-AXIS 4++ area with X-rays produced by a Rigaku MM-003 generator, or on beamline CMCF-ID at the Canadian Light Source synchrotron (Saskatoon, SK). The data were scaled, averaged and integrated using HKL2000 (302). All structures were solved by molecular replacement using Phaser (303), with a previously solved S25-39 or S25-2 Fab structure as a search model. Model building and refinement were carried out with Coot (304) and Refmac5 (305) through the CCP4 interface (306). Final model and refinement statistics are give in Table 20, Appendix A.

2.4.4. Figures

Figures depicting protein structures and electron density were created using SetoMac, an unpublished development of SETOR (307), and UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311) (308). Chemical drawings were created using ChemAxon MarvinSketch 15.11.30. Other figures were created using Microsot PowerPoint.

Bibliography

1. Silverstein AM (2009) *A History of Immunology* (Academic Press/Elsevier).
2. Gross CP, Sepkowitz KA (1998) The myth of the medical breakthrough: Smallpox, vaccination, and Jenner reconsidered. *Int J Infect Dis* 3(1):54–60.
3. Riedel S (2005) Edward Jenner and the history of smallpox and vaccination. *BUMC Proc* 18:21–25.
4. Behring E von, Kitasato BS (1890) Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Thieren. *Dtsch Medizinische Wochenschrift* 49:1–7.
5. Ehrlich P (1891) Experimentelle untersuchungen über immunität. I. Ueber Ricin. *DMW-Deutsche Medizinische Wochenschrift* 17(32):976–979.
6. Ehrlich P (1897) *Die Wertbemessung des Diphtherieheilserums und deren theoretische Grundlagen* (G. Fischer).
7. Ehrlich P (1901) Die seitenkettentheorie und ihre gegner. *Münch Med Wschr* 18:2123–2124.
8. Fischer, D S (1964) Theories of antibody formation: A review. *Yale J Biol Med* 37(6):1–30.
9. Silverstein AM (1985) A history of theories of antibody formation. *Cell Immunol* 91(1):263–83.
10. Pauling L (1940) A Theory of the Structure and Process of Formation of Antibodies. *J Am Chem Soc* 62:2643–2657.
11. Burnet FM (1941) *The Production of Antibodies. A Review and a Theoretical Discussion.* (Macmillan & Co.).
12. Burnet FM, Fenner F (1949) *The Production of Antibodies*.
13. Jerne NKN (1955) The Natural-Selection Theory of Antibody Formation. *Proc Natl Acad Sci U S A* 41(11):849.
14. Crick FH (1958) On protein synthesis. *Symposia of the Society for Experimental Biology*, p 138.
15. Burnet SFM (1959) *The clonal selection theory of acquired immunity* (University Press Cambridge).
16. Talmage DW (1959) Immunological Specificity Unique combinations of selected natural globulins provide an alternative to the classical concept. *Science (80-)* 129(3364):1643–1648.
17. Lederberg J (1959) Genes and Antibodies: Do antigens bear instructions for antibody specificity or do they select cell lines that arise by mutation? *Science*

(80-) 129:1649–1653.

18. Kindt TJ, Goldsby RA, Osborne BA (2007) *Kuby Immunology* doi:10.1016/S0076-6879(09)04815-0.
19. Chan TD, Brink R (2012) Affinity-based selection and the germinal center response. *Immunol Rev* 247(1):11–23.
20. Litman GW, Rast JP, Fugmann SD (2010) The origins of vertebrate adaptive immunity. *Nat Rev Immunol* 10(8):543–53.
21. Flajnik MF, Kasahara M (2010) Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet* 11(1):47–59.
22. Edelman GM (1991) Antibody structure and molecular immunology. *Scand J Immunol* 34(1):4–22.
23. Porter RR (1991) Structural studies of immunoglobulins. *Scand J Immunol* 34(4):382–388.
24. Steiner LA, Fleishman JB (2008) RR Porter and the structure of antibodies. *Molecular Aspects of Innate and Adaptive Immunity* (Royal Society of Chemistry, Cambridge, UK), pp 3–10.
25. Landsteiner K (1990) *The specificity of serological reactions* (Courier Corporation).
26. Petermann ML (1946) The splitting of human gamma globulin antibodies by papain and bromelin. *J Am Chem Soc* 68(1):106–113.
27. Petermann ML, Pappenheimer Jr AM (1941) The Ultracentrifugal Analysis of Diphtheria Proteins. *J Phys Chem* 45(1):1–9.
28. Rothen A, Landsteiner K (1942) Serological reactions of protein films and denatured proteins. *J Exp Med* 76(5):437–450.
29. Ceppellini R, et al. (1964) Nomenclature for human immunoglobulins. *Immunochemistry* 1(3):145–149.
30. Edelman GM (1959) Dissociation of γ-globulin. *J Am Chem Soc* 81(12):3155–3156.
31. Edelman GM, Poulik MD (1961) Studies on structural units of the γ-globulins. *J Exp Med* 113(5):861–884.
32. Fleischman JB, Porter RR, Press EM (1963) The arrangement of the peptide chains in γ-globulin. *Biochem J* 88(2):220.
33. O'Donnell IJ, Frangione B, Porter RR (1970) The disulphide bonds of the heavy chain of rabbit immunoglobulin G. *Biochem J* 116:261–268.
34. Kabat EA, Wu TT, Bilofsky H (1977) Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J Biol Chem* 252(19):6609–6616.

35. Poljak RJ, et al. (1973) Three-Dimensional Structure of the Fab' Fragment of a Human Immunoglobulin at 2.8-A Resolution. *Proc Natl Acad Sci* 70(12):3305–3310.

36. Dreyer W, Bennett J (1965) The molecular basis of antibody formation: a paradox. *Biochemistry* 54:864–869.

37. Hozumi N, Tonegawa S (1976) Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A* 73(10):3628–3632.

38. Lenhard-Schuller R, Hohn B, Brack C, Hirama M, Tonegawa S (1978) DNA clones containing mouse immunoglobulin kappa chain genes isolated by in vitro packaging into phage lambda coats. *Proc Natl Acad Sci U S A* 75(10):4709–13.

39. Bernard O, Hozumi N, Tonegawa S (1978) Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell* 15(4):1133–44.

40. Brack C, Tonegawa S (1977) Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 nontranslated bases apart. *Proc Natl Acad Sci U S A* 74(12):5652–6.

41. Early P, Huang H, Davis M, Calame K, Hood L (1980) An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell* 19(4):981–992.

42. Max EE, Seidman JG, Leder P (1979) Sequences of Five Potential Recombination Sites Encoded Close to an Immunoglobulin Kappa Constant Region Gene. *Proc Natl Acad Sci U S A* 76(7):3450–4.

43. Brack C, Hirama M, Lenhard-Schuller R, Tonegawa S (1978) A complete immunoglobulin gene is created by somatic recombination. *Cell* 15(1):1–14.

44. Sakano H, Hüppi K, Heinrich G, Tonegawa S (1979) Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* 280(5720):288–294.

45. Sakano H, Maki R, Kurosawa Y, Roeder W, Tonegawa S (1980) Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. *Nature* 286(5774):676–683.

46. Seidman JG, Max EE, Leder P (1979) A kappa-immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature* 280(5721):370–375.

47. Tonegawa S, Maxam a M, Tizard R, Bernard O, Gilbert W (1978) Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain. *Proc Natl Acad Sci U S A* 75(3):1485–9.

48. Tonegawa S (1983) Somatic generation of antibody diversity. *Nature* 302:575–581.

49. Schroeder HW (2006) Similarity and divergence in the development and expression of the mouse and human antibody repertoires. *Dev Comp Immunol* 30(1-2):119–35.

50. Giudicelli V, Chaume D, Lefranc MP (2005) IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic acids Res* 33:D256–61.

51. Lefranc M-P, et al. (2015) IMGT(R), the international ImMunoGeneTics information system(R) 25 years on. *Nucleic Acids Res* 43(D1):D413–D422.

52. Nadel B, Feeney a J (1997) Nucleotide deletion and P addition in V(D)J recombination: a determinant role of the coding-end sequence. *Mol Cell Biol* 17(7):3768–3778.

53. Nadel B, Feeney a J (1995) Influence of coding-end sequence on coding-end processing in V(D)J recombination. *J Immunol* 155(9):4322–9.

54. Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG (2000) The RAG Proteins and V(D)J Recombination: Complexes, Ends, and Transposition. *Annu Rev Immunol* 18(1):495–527.

55. Janeway CA, Travers P, Walport MJ, Shlomchik MJ, others (2001) *Immunobiology: the immune system in health and disease* (Churchill Livingstone London).

56. Glanville J, et al. (2009) Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci U S A*.

57. Georgiou G, et al. (2014) The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* 32(2):158–68.

58. Maul RW, Gearhart PJ (2010) AID and somatic hypermutation. *Adv Immunol* 105(10):159–91.

59. Li Z, Woo CJ, Iglesias-Ussel MD, Ronai D, Scharff MD (2004) The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev* 18(1):1–11.

60. Tomlinson IM, et al. (1996) The imprint of somatic hypermutation on the repertoire of human germline V genes. *J Mol Biol* 256(5):813–817.

61. Eason DD, et al. (2004) Mechanisms of antigen receptor evolution. *Semin Immunol* 16(4):215–26.

62. Maizels N (2005) Immunoglobulin gene diversification. *Annu Rev Genet* 39(June):23–46.

63. Berek C, Milstein C (1987) Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev* 96:23–41.

64. Berman HM, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28(1):235–42.

65. Davies DR, Padlan E a, Sheriff S (1990) Antibody-antigen complexes. *Annu Rev Biochem* 59:439–73.

66. Davies DR, Metzger H (1983) Structural basis of antibody function. *Annu Rev Immunol* 1:87–117.

67. Chothia C, et al. (1989) Conformations of immunoglobulin hypervariable regions. *Nature* 342(6252):877–83.

68. Potter M (1983) Structural correlates of immunoglobulin diversity. *Surv Immunol Res* 2(1):27–42.

69. Elgert KD (1998) Antibody structure and function. *Immunology: Understanding the Immune System*, pp 58–78.

70. Ramsland PA, Farrugia W (2002) Crystal structures of human antibodies: A detailed and unfinished tapestry of immunoglobulin gene products. *J Mol Recognit* 15(L):248–259.

71. Padlan EA (1994) Anatomy of the antibody molecule. *Mol Immunol* 31(3):169–217.

72. Alzari PM, Lascombe MB, Poljak RJ (1988) Three-Dimensional Structure of Antibodies. *Annu Rev Immunol* 6(1):555–580.

73. Braden BC, Poljak RJ (1995) Structural features of the reactions between antibodies and protein antigens. *FASEB J* 9(1):9–16.

74. Sundberg EJ, Mariuzza RA (2002) Molecular recognition in antibody-antigen complexes. *Adv Protein Chem* 61:119–160.

75. Kabat EA, Wu TT, Perry HM, Gottesman KS, Koeler C (1984) Sequences of proteins of immunological interest. *Anal Biochem* 138(1):265.

76. Chothia C, Lesk AM (1987) Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* 196(4):901–917.

77. Al-Lazikani B, Lesk AM, Chothia C (1997) Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol* 273(4):927–948.

78. Lefranc MP, et al. (2005) IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. *Dev Comp Immunol* 29(3):185–203.

79. Honegger A, Plückthun A (2001) Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J Mol Biol* 309(3):657–70.

80. Abhinandan KR, Martin ACR (2008) Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol Immunol* 45(14):3832–3839.

81. Martin AC, Thornton JM (1996) Structural families in loops of homologous

proteins: automatic classification, modelling and application to antibodies. *J Mol Biol* 263(5):800–15.

82. North B, Lehmann A, Dunbrack RL (2011) A new clustering of antibody CDR loop conformations. *J Mol Biol* 406(2):228–56.

83. Nikoloudis D, Pitts JE, Saldanha JW (2014) A complete, multi-level conformational clustering of antibody complementarity-determining regions. *PeerJ* 2:e456.

84. Adolf-Bryfogle J, Xu Q, North B, Lehmann A, Dunbrack RL (2015) PyIgClassify: a database of antibody CDR structural classifications. *Nucleic Acids Res* 43(D1):D432–D438.

85. Shirai H, Kidera A, Nakamura H (1996) Structural classification of CDR-H3 in antibodies. *FEBS Lett* 399(1-2):1–8.

86. Shirai H, Kidera A, Nakamura H (1999) H3-rules : identification of CDR-H3 structures in antibodies. *FEBS Lett* 455:188–197.

87. Kuroda D, Shirai H, Kobori M, Nakamura H (2008) Structural classification of CDR-H3 revisited: a lesson in antibody modeling. *Proteins* 73(3):608–20.

88. Weitzner BDD, Dunbrack RLL, Gray JJJ (2015) The Origin of CDR H3 Structural Diversity. *Structure* 23(2):302–311.

89. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM (1997) Antibody structure, prediction and redesign. *Biophys Chem* 68(1-3):9–16.

90. Oliva B, Bates PA, Querol E, Avilés FX, Sternberg MJ (1998) Automated classification of antibody complementarity determining region 3 of the heavy chain (H3) loops into canonical forms and its application to protein structure prediction. *J Mol Biol* 279(5):1193–210.

91. Koliasnikov O V, Kiral MO, Grigorenko VG, Egorov AM (2006) Antibody CDR H3 modeling rules: extension for the case of absence of Arg H94 and Asp H101. *J Bioinform Comput Biol* 4(2):415–24.

92. Kabat EA, Te Wu T, Perry HM, Gottesman KS, Foeller C (1992) *Sequences of proteins of immunological interest* (DIANE publishing).

93. Martin ACR (1996) Accessing the Kabat antibody sequence database by computer. *PROTEINS Struct Funct Genet* 25(1):130–3.

94. Sibanda BL, Blundell TL, Thornton JM (1989) Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* 206:759–777.

95. Brekke OH, Sandlie I (2003) Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov* 2(1):52–62.

96. Chan AC, Carter PJ (2010) Therapeutic antibodies for autoimmunity and inflammation. *Nat Rev Immunol* 10(5):301–16.

97. Kuroda D, Shirai H, Jacobson MP, Nakamura H (2012) Computer-aided antibody design. *Protein Eng Des Sel* 25(10):507–521.

98. Holliger P, Hudson PJ (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* 23(9):1126–36.

99. Beck A, Wurch T, Bailly C, Corvaia N (2010) Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* 10(5):345–52.

100. Sievers S a., Scharf L, West AP, Bjorkman PJ (2015) Antibody engineering for increased potency, breadth and half-life. *Curr Opin HIV AIDS*:1.

101. Weiner LM, Surana R, Wang S (2010) Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat Rev Immunol* 10(5):317–27.

102. Adams GP, Weiner LM (2005) Monoclonal antibody therapy of cancer. *Nat Biotechnol* 23(9):1147–57.

103. Chames P, Baty D (2009) Bispecific antibodies for cancer therapy: The light at the end of the tunnel? *MAbs* 1(6):539–547.

104. Katsuda T, Sonoda H, Kumada Y, Yamaji H (2012) Antibody Engineering. *Antibody Engineering*, ed Chames P (Humana Press, Totowa, NJ). doi:10.1007/978-1-61779-974-7.

105. Messih MA, Lepore R, Marcatili P, Tramontano A (2014) Improving the accuracy of the structure prediction of the third hypervariable loop of the heavy chains of antibodies. *Bioinformatics* 30(19):2733–2740.

106. Marcatili P, Rosi A, Tramontano A (2008) PIGS: automatic prediction of antibody structures. *Bioinformatics* 24(17):1953–1954.

107. Dunbar J, et al. (2014) SAbDab: The structural antibody database. *Nucleic Acids Res* 42(D1):1–7.

108. Lyskov S, et al. (2013) Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS One* 8(5):5–7.

109. Klausen MS, Anderson MV, Jespersen MC, Nielsen M, Marcatili P (2015) LYRA, a webserver for lymphocyte receptor structural modeling. *Nucleic Acids Res* 43(W1):W349–W355.

110. Shirai H, et al. (2014) High-resolution modeling of antibody structures by a combination of bioinformatics, expert knowledge, and molecular simulations. *Proteins Struct Funct Bioinforma* 82(April):1624–1635.

111. Choi Y, Deane CM (2010) FREAD revisited: Accurate loop structure prediction using a database search algorithm. *Proteins* 78(6):1431–40.

112. Choi Y, Deane CM (2011) Predicting antibody complementarity determining region structures without classification. *Mol Biosyst* 7(12):3327–34.

113. Almagro JC, et al. (2011) Antibody modeling assessment. *Proteins* 79(11):3050–

66.

114. Almagro JC, et al. (2014) Second Antibody Modeling Assessment (AMA-II). *Proteins Struct Funct Bioinforma* 82(8):1553–1562.

115. Teplyakov A, et al. (2014) Antibody modeling assessment II. Structures and models. *Proteins Struct Funct Bioinforma* 82(February):1563–1582.

116. Mak TW, Saunders ME, Jett BD (2013) *Primer to the Immune Response* (Elsevier Science).

117. Berzofsky J, Schechter A (1981) The concepts of crossreactivity and specificity in immunology. *Mol Immunol* 18(8):751–763.

118. Richards FF, Konigsberg WH, Rosenstein RW, Varga JM (1975) On the specificity of antibodies. *Science (80-)* 187(4172):130–7.

119. Rosenstein RW, Musson R a, Armstrong MK, Konigsberg WH, Richards FF (1972) Contact regions for dinitrophenyl and menadione haptens in an immunoglobulin binding more than one antigen. *Proc Natl Acad Sci USA* 69(4):877–81.

120. Van Regenmortel MH V. (2014) Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. *J Mol Recognit* 27(11):627–39.

121. Notkins AL (2004) Polyreactivity of antibody molecules. *Trends Immunol* 25(4):174–179.

122. Dimitrov JD, Pashov AD, Vassilev TL (2012) Antibody polyspecificity: what does it matter? *Adv Exp Med Biol* 750:213–26.

123. Manivel V, Bayiroglu F, Siddiqui Z, Salunke DM, Rao KVS (2002) The Primary Antibody Repertoire Represents a Linked Network of Degenerate Antigen Specificities. *J Immunol* 169(2):888–897.

124. Seigneurin JM, Guilbert B, Bourgeat MJ, Avrameas S (1988) Polyspecific natural antibodies and autoantibodies secreted by human lymphocytes immortalized with Epstein-Barr virus. *Blood* 71(3):581–5.

125. James LC, Tawfik DS (2003) The specificity of cross-reactivity: promiscuous antibody binding involves specific hydrogen bonds rather than nonspecific hydrophobic stickiness. *Protein Sci* 12:2183–2193.

126. Karlsen AE, Dyrberg T (1998) Molecular mimicry between non-self, modified self and self in autoimmunity. *Semin Immunol* 10(1):25–34.

127. Cohen IR (2001) Antigenic mimicry, clonal selection and autoimmunity. *J Autoimmun* 16(3):337–40.

128. Aalberse RC, Kleine Budde I, Stapel SO, van Ree R (2001) Structural aspects of cross-reactivity and its relation to antibody affinity. *Allergy* 56 Suppl 6:27–9.

129. Frank SA (2002) Immunology and Evolution of Infectious Disease. *Heal San Fr* 300(5617):358.

130. Schuermann JP, Henzl MT, Deutscher SL, Tanner JJ (2004) Structure of an anti-DNA fab complexed with a non-DNA ligand provides insights into cross-reactivity and molecular mimicry. *Proteins* 57(2):269–78.

131. Keitel T, et al. (1997) Crystallographic analysis of anti-p24 (HIV-1) monoclonal antibody cross-reactivity and polyspecificity. *Cell* 91(6):811–820.

132. Scheerer P, Kramer A, Otte L, Seifert M, Wessner H (2007) Structure of an anti-cholera toxin antibody Fab in complex with an epitope-derived D -peptide : a case of polyspecific recognition y , z. 263–274.

133. Arevalo J, Taussig M, Wilson I (1993) Molecular basis of crossreactivity and the limits of antibody-antigen complementarity. *Nature* 365:859–863.

134. Brooks CL, et al. (2008) Exploration of specificity in germline monoclonal antibody recognition of a range of natural and synthetic epitopes. *J Mol Biol* 377(2):450–68.

135. Brooks CL, et al. (2010) The role of CDR H3 in antibody recognition of a synthetic analog of a lipopolysaccharide antigen. *Glycobiology* 20(2):138–47.

136. Blackler RJ, et al. (2011) A common NH53K mutation in the combining site of antibodies raised against chlamydial LPS glycoconjugates significantly increases avidity. *Biochemistry* 50(16):3357–68.

137. Blackler RJ, et al. (2012) Antibody Recognition of Chlamydia LPS: Structural Insights of Inherited Immune Responses. *Anticarbohydrate Antibodies*, eds Kosma P, Müller-Loennies S (Springer Vienna, Vienna), pp 75–120.

138. Brooks CL, et al. (2010) Antibodies raised against chlamydial lipopolysaccharide antigens reveal convergence in germline gene usage and differential epitope recognition. *Biochemistry* 49(3):570–81.

139. Gerstenbruch S, et al. (2010) Analysis of Cross-Reactive and Specific Anti-Carbohydrate Antibodies against Lipopolysaccharide from Chlamydophila psittaci. *Glycobiology* 20(4):461–472.

140. Nguyen HP, et al. (2003) Germline antibody recognition of distinct carbohydrate epitopes. *Nat Struct Biol* 10(12):1019–25.

141. Evans DW, et al. (2011) Structural insights into parallel strategies for germline antibody recognition of lipopolysaccharide from Chlamydia. *Glycobiology* 21(8):1049–59.

142. Haji-Ghassemi O, et al. (2014) Groove-type recognition of chlamydiaceae-specific lipopolysaccharide antigen by a family of antibodies possessing an unusual variable heavy chain N-linked glycan. *J Biol Chem* 289(24):16644–16661.

143. Oldstone MB (1998) Molecular mimicry and immune-mediated diseases. *FASEB J* 12(13):1255–1265.

144. Sethi DK, Agarwal A, Manivel V, Rao KVS, Salunke DM (2006) Differential epitope

positioning within the germline antibody paratope enhances promiscuity in the primary immune response. *Immunity* 24(4):429–38.

145. Khan T, Salunke DM (2014) Adjustable Locks and Flexible Keys: Plasticity of Epitope-Paratope Interactions in Germline Antibodies. *J Immunol*. doi:10.4049/jimmunol.1302143.

146. Acqua WD, Goldman ER, Eisenstein E, Mariuzza R a (1996) A Mutational Analysis of the Binding of Two Different Proteins to the Same. *Methods* 2960(96):9667–9676.

147. Vyas NK, et al. (2003) Structural basis of peptide-carbohydrate mimicry in an antibody-combining site. *Proc Natl Acad Sci USA* 100(25):15023–8.

148. Theillet FX, et al. (2009) Structural mimicry of O-antigen by a peptide revealed in a complex with an antibody raised against Shigella flexneri serotype 2a. *J Mol Biol* 388:839–50.

149. Tapryal S, Gaur V, Kaur KJ, Salunke DM (2013) Structural Evaluation of a Mimicry-Recognizing Paratope: Plasticity in Antigen-Antibody Interactions Manifests in Molecular Mimicry. *J Immunol* 191(1):456–63.

150. Haurowitz F (1967) The Evolution of Selective and Instructive Theories of Antibody Formation. *Cold Spring Harb Symp Quant Biol* 32:559–567.

151. Foote J, Milstein C (1994) Conformational isomerism and the diversity of antibodies. *Proc Natl Acad Sci USA* 91(October):10370–10374.

152. Jimenez R, Case D a., Romesberg FE (2002) Flexibility of an antibody binding site measured with photon echo spectroscopy. *J Phys Chem B* 106(5):1090–1103.

153. Kroon GJA, Mo H, Martinez-Yamout MA, Dyson HJ, Wright PE (2003) Changes in structure and dynamics of the Fv fragment of a catalytic antibody upon binding of inhibitor. *Protein Sci* 12(7):1386–94.

154. Komissarov a a, Deutscher SL (1999) Thermodynamics of Fab-ssDNA interactions: contribution of heavy chain complementarity determining region 3. *Biochemistry* 38(44):14631–7.

155. Schulze-Gahmen U, Rini J, Wilson I (1993) Detailed Analysis of the Free and Bound Conformations of an Antibody; X-ray Structures of Fab 17/9 and Three Different Fab-peptide Complexes. *J Mol Biol* 234:1098–1118.

156. Huang C, et al. (2004) Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc Natl Acad Sci U S A* 101(9):2706–2711.

157. Buchbinder JL, Stephenson RC, Scanlan TS, Fletterick RJ (1998) A comparison of the crystallographic structures of two catalytic antibodies with esterase activity. *J Mol Biol* 282(5):1033–41.

158. Schuermann JP, Prewitt SP, Davies C, Deutscher SL, Tanner JJ (2005) Evidence for

structural plasticity of heavy chain complementarity-determining region 3 in antibody-ssDNA recognition. *J Mol Biol* 347(5):965–978.

159. Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA (2007) Maturation of Shark Single-domain (IgNAR) Antibodies: Evidence for Induced-fit Binding. *J Mol Biol* 367(2):358–372.

160. Yin J, Beuscher AE, Andryski SE, Stevens RC, Schultz PG (2003) Structural Plasticity and the Evolution of Antibody Affinity and Specificity. *J Mol Biol* 330(4):651–656.

161. James LC, Roversi P, Tawfik DS (2003) Antibody Multispecificity Mediated by Conformational Diversity. *Science (80-)* 299(February):1362–1367.

162. Wang W, et al. (2013) Conformational selection and induced fit in specific antibody and antigen recognition: SPE7 as a case study. *J Phys Chem B* 117:4912–4923.

163. Khan T, Salunke DM (2012) Structural elucidation of the mechanistic basis of degeneracy in the primary humoral response. *J Immunol* 188(4):1819–27.

164. Bostrom J, et al. (2009) Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site. *Science* 323(5921):1610–1614.

165. Wang F, et al. (2013) Somatic hypermutation maintains antibody thermodynamic stability during affinity maturation. *Proc Natl Acad Sci U S A* 110(11):4261–6.

166. Xu H, et al. (2014) Key mutations stabilize antigen-binding conformation during affinity maturation of a broadly neutralizing influenza antibody lineage. (212):771–780.

167. Adhikary R, et al. (2015) Adaptive Mutations Alter Antibody Structure and Dynamics during Affinity Maturation. *Biochemistry*:150310155001006.

168. Yin J, et al. (2001) A comparative analysis of the immunological evolution of antibody 28B4. *Biochemistry* 40(36):10764–10773.

169. Li T, et al. (2015) Rigidity Emerges during Antibody Evolution in Three Distinct Antibody Systems: Evidence from QSFR Analysis of Fab Fragments. *PLOS Comput Biol* 11(7):e1004327.

170. Schmidt AG, et al. (2013) Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc Natl Acad Sci* 110(1):264–269.

171. Babor M, Kortemme T (2009) Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility. *Proteins* 75(4):846–58.

172. Thorpe IF, Brooks CL (2007) Molecular evolution of affinity and flexibility in the immune system. *Proc Natl Acad Sci U S A* 104(21):8821–8826.

173. Zimmermann J, et al. (2006) Antibody evolution constrains conformational heterogeneity by tailoring protein dynamics. *Proc Natl Acad Sci U S A*

103(37):13722–13727.

174. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC (1997) Structural insights into the evolution of an antibody combining site. *Science* 276(5319):1665–1669.

175. Bucior I, Burger MM (2004) Carbohydrate-carbohydrate interactions in cell recognition. *Curr Opin Struct Biol* 14(5):631–637.

176. Lennarz W (2012) *The biochemistry of glycoproteins and proteoglycans* (Springer Science & Business Media).

177. Hakomori S (1984) Tumor-associated carbohydrate antigens. *Annu Rev Immunol* 2(1):103–126.

178. Hakomori S (1989) Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens. *Adv Cancer Res* 52(25):270.

179. Hakomori S (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *The Molecular Immunology of Complex carbohydrates—2* (Springer), pp 369–402.

180. Lang KS, Burow A, Kurrer M, Lang PA, Recher M (2007) The role of the innate immune response in autoimmune disease. *J Autoimmun* 29(4):206–12.

181. Lutz HU (2007) Homeostatic roles of naturally occurring antibodies: An overview. *J Autoimmun* 29(4):287–294.

182. Astronomo RD, Burton DR (2010) Carbohydrate vaccines: developing sweet solutions to sticky situations? *Nat Rev Drug Discov* 9(4):308–24.

183. Storry, J R, Olsson ML (2009) The ABO blood group system revisited: a review and update. *Immunohematol J Blood Gr Serol Educ* 25(Supplement):48–59.

184. Oldstone MBA (2005) *Molecular Mimicry: Infection-Inducing Autoimmune Disease* ed Oldstone MBA (Springer-Verlag, Berlin/Heidelberg) doi:10.1007/3-540-30791-5.

185. Bonilla F a, Oettgen HC (2010) Adaptive immunity. *J Allergy Clin Immunol* 125(2):S33–40.

186. Vos Q, Lees A, Wu Z, Snapper C, Mond J (2000) B-cell activation by T-cell-independent type 2 antigens as an integral part of the humoral immune response to pathogenic microorganisms. *Immunol Rev* 176:154–170.

187. Mond JJ, Vos Q, Lees A, Snapper CM (1995) T cell independent antigens. *Curr Opin Immunol* 7(3):349–54.

188. Mond J, Lees A, Snapper C (1995) T cell independent antigens type 2. *Annu Rev Immunol* 13:655–692.

189. Raetz CRH, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* 71:635–700.

190. Holst O, Bock K, Brade L, Brade H (1995) The Structures of Oligosaccharide Bisphosphates Isolated from the Lipopolysaccharide of a Recombinant Escherichia coli Strain Expressing the Gene gseA [3-deoxy-d-manno-Octulopyranosonic Acid (Kdo) Transferase] of Chlamydia psittaci 6BC. *Eur J Biochem* 229(1):194–200.

191. Isshiki Y, Kawahara K, Zähringer U (1998) Isolation and characterisation of disodium (4-amino-4-deoxy-β-l-arabinopyranosyl)-(1→8)-(d-glycero-α-d-talo-oct-2-ulopyranosylonate)-(2→4)-(methyl 3-deoxy-d-manno-oct-2-ulopyranosid) onate from the lipopolysaccharide of Burkholderia cepacia. *Carbohydr Res* 313(1):21–27.

192. Süsskind M, Müller-Loennies S, Nimmich W, Brade H, Holst O (1995) Structural investigation on the carbohydrate backbone of the lipopolysaccharide from Klebsiella pneumoniae rough mutant R20/O1−. *Carbohydr Res* 269(1):C1–C7.

193. Everett KDE (2000) Chlamydia and Chlamydiales: more than meets the eye. *Vet Microbiol* 75(2):109–126.

194. Read TD, et al. (2003) Genome sequence of Chlamydophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae. *Nucleic Acids Res* 31(8):2134–2147.

195. Everett KDE, Bush RM, Andersen AA (1999) Emended description of the order Chlamydiales, proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards. *Int J Syst Bacteriol* 49(2):415–440.

196. Brade H, Brade L, Nano FE (1987) Chemical and serological investigations on the genus-specific lipopolysaccharide epitope of Chlamydia. *Proc Natl Acad Sci U S A* 84(8):2508–12.

197. Kosma P, Bahnmüller R, Schulz G, Brade H (1990) Synthesis of a tetrasaccharide of the genus-specific lipopolysaccharide epitope of Chlamydia. *Carbohydr Res* 208:37–50.

198. Müller-Loennies S, et al. (2000) Characterization of high affinity monoclonal antibodies specific for chlamydial lipopolysaccharide. *Glycobiology* 10(2):121–30.

199. Fu Y, Baumann M, Kosma P, Brade L, Brade H (1992) A synthetic glycoconjugate representing the genus-specific epitope of chlamydial lipopolysaccharide exhibits the same specificity as its natural counterpart. *Infect Immun* 60(4):1314–21.

200. Brade L, Nurminen M, Mäkelä PH, Brade H (1985) Antigenic properties of Chlamydia trachomatis lipopolysaccharide. *Infect Immun* 48(2):569–72.

201. Nurminen M, Rietschel ET, Brade H (1985) Chemical characterization of Chlamydia trachomatis lipopolysaccharide. *Infect Immun* 48(2):573–5.

202. Lepenies B, Seeberger PH (2010) The promise of glycomics, glycan arrays and carbohydrate-based vaccines. *Immunopharmacol Immunotoxicol* 32(2):196–207.

203. Goldblatt D (2000) Conjugate vaccines. *Clin Exp Immunol* 119(1):1–3.

204. Huang Y-L, Wu C-Y (2010) Carbohydrate-based vaccines: challenges and opportunities. *Expert Rev Vaccines* 9(11):1257–1274.

205. Brade L, Rozalski A, Kosma P, Brade H (2000) A monoclonal antibody recognizing the 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) trisaccharide αKdo (2→ 4) αKdo (2→ 4) αKdo of Chlamydophila spacepsittaci 6BC lipopolysaccharide. *J Endotoxin Res* 6(5):361–368.

206. Brade L, Gronow S, Wimmer N, Kosma P, Brade H (2002) Monoclonal antibodies against 3-deoxy-α-D-manno-oct-2-ulosonic acid (Kdo) and D-glycero-α-D-talo-oct-2-ulosonic acid (Ko). *J Endotoxin Res* 8(5):357–364.

207. Kosma P, Schulz G, Brade H (1988) Synthesis of a trisaccharide of 3-deoxy-D-manno-2-octulopyranosylonic acid (KDO) residues related to the genus-specific lipopolysaccharide epitope of Chlamydia. *Carbohydr Res* 183(2):183–199.

208. Kosma P, Schulz G, Unger FM, Brade H (1989) Synthesis of trisaccharides containing 3-deoxy-D-manno-2-octulosonic acid residues related to the KDO-region of enterobacterial lipopolysaccharides. *Carbohydr Res* 190(2):191–201.

209. Kosma P, et al. (1999) Synthesis of inner core antigens related to Chlamydia, Pseudomonas and Acinetobacter LPS. *J Endotoxin Res* 5(3):157–163.

210. Kosma P, Reiter A, Hofinger A, Brade L, Brade H (2000) Synthesis of neoglycoproteins containing Kdo epitopes specific for Chlamydophila psittaci lipopolysaccharide. *J Endotoxin Res* 6(1):57–69.

211. Maaheimo H, Kosma P, Brade L, Brade H, Peters T (2000) Mapping the binding of synthetic disaccharides representing epitopes of chlamydial lipopolysaccharide to antibodies with NMR. *Biochemistry* 39(42):12778–12788.

212. Müller-Loennies S, et al. (2002) A novel strategy for the synthesis of neoglycoconjugates from deacylated deep rough lipopolysaccharides. *J Endotoxin Res* 8(4):295–305.

213. Müller-Loennies S, et al. (2006) A monoclonal antibody against a carbohydrate epitope in lipopolysaccharide differentiates Chlamydophila psittaci from Chlamydophila pecorum, Chlamydophila pneumoniae, and Chlamydia trachomatis. *Glycobiology* 16(3):184–196.

214. Adderson EE, Shackelford PG, Quinn A, Carroll WL (1991) Restricted Ig H chain V gene usage in the human antibody response to Haemophilus influenzae type b capsular polysaccharide. *J Immunol* 147(5):1667–1674.

215. Senn BM, et al. (2003) Combinatorial immunoglobulin light chain variability creates sufficient B cell diversity to mount protective antibody responses against pathogen infections. *Eur J Immunol* 33(4):950–961.

216. Shaw DR, Kirkham P, Schroeder HW, Roben P, Silverman GJ (1995) Structure-

function studies of human monoclonal antibodies to pneumococcus type 3 polysaccharide. *Ann N Y Acad Sci* 764(1):370–3.

217. Casadevall A, Scharff MD (1991) The mouse antibody response to infection with Cryptococcus neoformans: VH and VL usage in polysaccharide binding antibodies. *J Exp Med* 174(1):151–160.

218. Pirofski L, Lui R, DeShaw M, Kressel AB, Zhong Z (1995) Analysis of human monoclonal antibodies elicited by vaccination with a Cryptococcus neoformans glucuronoxylomannan capsular polysaccharide vaccine. *Infect Immun* 63(8):3005–3014.

219. Berry JD, Boese DJ, Law DKS, Zollinger WD, Tsang RSW (2005) Molecular analysis of monoclonal antibodies to group variant capsular polysaccharide of Neisseria meningitidis: recurrent heavy chains and alternative light chain partners. *Mol Immunol* 42(3):335–344.

220. Nakouzi A, Casadevall A (2003) The function of conserved amino acids in or near the complementarity determining regions for related antibodies to Cryptococcus neoformans glucuronoxylomannan. *Mol Immunol* 40(6):351–361.

221. Zhou J, Lottenbach KRK, Barenkamp SJ, Lucas AH, Reason DC (2002) Recurrent variable region gene usage and somatic mutation in the human antibody response to the capsular polysaccharide of Streptococcus pneumoniae type 23F. *Infect Immun* 70(8):4083–4091.

222. Zhou J, Lottenbach KR, Barenkamp SJ, Reason DC (2004) Somatic hypermutation and diverse immunoglobulin gene usage in the human antibody response to the capsular polysaccharide of Streptococcus pneumoniae Type 6B. *Infect Immun* 72(6):3505–3514.

223. Reason DC, Zhou J (2004) Correlation of antigenic epitope and antibody gene usage in the human immune response to Streptococcus pneumoniae type 23F capsular polysaccharide. *Clin Immunol* 111(1):132–136.

224. Lucas AH, Moulton KD, Tang VR, Reason DC (2001) Combinatorial library cloning of human antibodies to Streptococcus pneumoniae capsular polysaccharides: variable region primary structures and evidence for somatic mutation of Fab fragments specific for capsular serotypes 6B, 14, and 23F. *Infect Immun* 69(2):853–864.

225. Cygler M, Rose D, Bundle D (1991) Recognition of a Cell-Surface Oligosaccharide of Pathogenic Salmonella by an Antibody Fab Fragment. *Science (80-)* 253(5018):442–445.

226. Bundle DR, Young NM (1992) Carbohydrate-protein interactions in antibodies and lectins. *Curr Opin Struct Biol* 2(5):666–673.

227. Xu JL, Davis MM (2000) Diversity in the CDR3 region of V H is sufficient for most antibody specificities. *Immunity* 13(1):37–45.

228. Li Z, Woo CJ, Iglesias-Ussel MD, Ronai D, Scharff MD (2004) The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev* 18(1):1–11.

229. Sela-Culang I, Alon S, Ofran Y (2012) A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *J Immunol* 189(10):4890–9.

230. Rini JM, Schulze-Gahmen U, Wilson IA (1992) Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science* 255(5047):959–65.

231. Stanfield RL, Takimoto-Kamimura M, Rini JM, Profy AT, Wilson IA (1993) Major antigen-induced domain rearrangements in an antibody. *Structure* 1(2):83–93.

232. Stanfield RL, Wilson I a (1994) Antigen-induced conformational changes in antibodies: a problem for structural prediction and design. *Trends Biotechnol* 12(7):275–9.

233. Wilson I a, Stanfield RL (1994) Antibody-antigen interactions: new structures and new conformational changes. *Curr Opin Struct Biol* 4(6):857–67.

234. Elsen J Van Den, Vandeputte-rutten L, Kroon J, Gros P (1999) Bactericidal Antibody Recognition of Meningococcal PorA by. 274(3):1495–1501.

235. Webster DM, Henry AH, Rees AR (1994) Antibody-antigen interactions. *Curr Opin Struct Biol* 4(2):297.

236. Edmundson AB, et al. (1974) Binding of 2,4-Dinitrophenyl compounds and other small Molecules to a Crystalline A-Type Bence-Jones Dimer. *Biochemistry* 13(18):3816–3827.

237. Verdaguer N, Mateu MG, Bravo J, Domingo E, Fita I (1996) Induced pocket to accommodate the cell attachment Arg-Gly-Asp motif in a neutralizing antibody against foot-and-mouth-disease virus. *J Mol Biol* 256(2):364–376.

238. Herron JN, et al. (1991) An autoantibody to single-stranded DNA: comparison of the three-dimensional structures of the unliganded Fab and a deoxynucleotide-Fab complex. *Proteins* 11(3):159–75.

239. Davies DR, Padlan E a (1992) Twisting into shape. *Curr Biol* 2:254–256.

240. Tormo J, et al. (1994) Crystal structure of a human rhinovirus neutralizing antibody complexed with a peptide derived from viral capsid protein VP2. *EMBO J* 13(10):2247–2256.

241. Wilson A, Stanfield RL (1993) Antibody-antigen interactions. 113–118.

242. Monod J, Wyman J, Changeux J-PJ (1965) On the nature of Allosteric Transitions: A Plausible Model. *J Mol Biol* 12(1):88–118.

243. Changeux J, Edelstein S (2011) Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep* 3(19). doi:10.3410/B3-19.

244. James LC, Tawfik DS (2003) Conformational diversity and protein evolution – a 60-year-old hypothesis revisited. *Trends Biochem Sci* 28(7):361–368.

245. Vogt AD, Pozzi N, Chen Z, Di Cera E (2014) Essential role of conformational selection in ligand binding. *Biophys Chem* 186:13–21.

246. Vogt AD, Di Cera E (2012) Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochemistry* 51(30):5894–5902.

247. Vogt AD, Di Cera E (2013) Conformational selection is a dominant mechanism of ligand binding. *Biochemistry* 52(34):5723–5729.

248. Boehr DD, Nussinov R, Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 5(11):789–96.

249. Csermely P, Palotai R, Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: An extended view of binding events. *Trends Biochem Sci* 35(10):539–546.

250. Nussinov R, Ma B, Tsai CJ (2014) Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophys Chem* 186:22–30.

251. Tsai C, Ma B, Nussinov R (1999) Folding and binding cascades: Shifts in energy landscapes. *Proc Natl …* 96(18):9970–9972.

252. Weikl TR, Paul F (2014) Conformational selection in protein binding and function. *Protein Sci* 23(11):1508–1518.

253. Yu Q, Ye W, Wang W, Chen HF (2013) Global Conformational Selection and Local Induced Fit for the Recognition between Intrinsic Disordered p53 and CBP. *PLoS One* 8(3). doi:10.1371/journal.pone.0059627.

254. Silva D, Bowman GR, Sosa-peinado A, Huang X (2011) A Role for Both Conformational Selection and Induced Fit in Ligand Binding by the LAO Protein. 7(5). doi:10.1371/journal.pcbi.1002054.

255. Plattner N, Noé F (2015) Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat Commun* 6(May):7653.

256. Daniels KG, Suo Y, Oas TG (2015) Conformational kinetics reveals affinities of protein conformational states. *Proc Natl Acad Sci* 112(30):201502084.

257. Greives N, Zhou H (2014) Both protein dynamics and ligand concentration can shift the binding mechanism between conformational selection and induced fit. 111(28):10197–10202.

258. Bosshard HR (2001) Molecular Recognition by Induced Fit: How Fit is the Concept? *News Physiol Sci* 16(4):171–173.

259. Zhou HX (2010) From induced fit to conformational selection: a continuum of binding mechanism controlled by the timescale of conformational transitions. *Biophys J* 98(6):L15–7.

260. Hammes GG, Chang Y, Oas TG (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc Natl Acad Sci USA* 106(33):13737–41.

261. Bohnuud T, Kozakov D, Vajda S (2014) Evidence of Conformational Selection Driving the Formation of Ligand Binding Sites in Protein-Protein Interfaces. *PLoS Comput Biol* 10(10):e1003872.

262. Arai M, Sugase K, Dyson HJ, Wright PE (2015) Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc Natl Acad Sci* 112(31):201512799.

263. Schafer CT, Farrens DL (2015) Conformational Selection and Equilibrium Governs the Ability of Retinals to Bind Opsin. *J Biol Chem* 290(7):4304–4318.

264. Musiani F, et al. (2014) Molecular Dynamics Simulations Identify Time Scale of Conformational Changes Responsible for Conformational Selection in Molecular Recognition of HIV-1 Transactivation Responsive RNA. *J Am Chem Soc* 136(44):15631–7.

265. Abrol R, et al. (2014) Ligand- and mutation-induced conformational selection in the CCR5 chemokine G protein-coupled receptor. *Proc Natl Acad Sci* 111(36):13040–13045.

266. Tobi D, Bahar I (2005) Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc Natl Acad Sci U S A* 102(52):18908–18913.

267. Jimenez R, Salazar G, Baldridge KK, Romesberg FE (2003) Flexibility and molecular recognition in the immune system. *Proc Natl Acad Sci USA* 100(1):92–7.

268. Dunbar J, Fuchs A, Shi J, Deane CM (2013) ABangle: characterising the VH-VL orientation in antibodies. *Protein Eng Des Sel* 26(10):611–620.

269. Chothia C, et al. (1992) Structural repertoire of the human VH segments. *J Mol Biol* 227(3):799–817.

270. Vargas-Madrazo E, Paz-Garcia E (2002) Modifications to canonical structure sequence patterns: Analysis for L1 and L3. *Proteins Struct Funct Genet* 47(2):250–254.

271. Bostrom J, Haber L, Koenig P, Kelley RF, Fuh G (2011) High affinity antigen recognition of the dual specific variants of herceptin is entropy-driven in spite of structural plasticity. *PLoS One* 6(4):e17887.

272. Jungwirth P, Cremer PS (2014) Beyond Hofmeister. *Nat Chem* 6(4):261–263.

273. Lim K, Nadarajah A, Forsytiie EL, Pusey ML (1998) Locations of bromide ions in tetragonal lysozyme crystals. *Acta Crystallogr Sect D Biol Crystallogr* 54(5):899–904.

274. Dauter Z, Li M, Wlodawer A (2001) Practical experience with the use of halides for

phasing macromolecular structures: A powerful tool for structural genomics. *Acta Crystallogr Sect D Biol Crystallogr* 57(2):239–249.

275. Nagem RAP, Polikarpov I, Dauter Z (2003) Phasing on Rapidly Soaked Ions. *Methods Enzymol* 374(2002):120–137.

276. Dauter Z, Dauter M (2001) Entering a New Phase : Using Solvent Halide Ions. *Structure* 9(01):21–26.

277. Steinrauf LK (1998) Structures of monoclinic lysozyme iodide at 1.6 Å and of triclinic lysozyme nitrate at 1.1 Å. *Acta Crystallogr Sect D Biol Crystallogr* 54(5):767–779.

278. Collins KD (2004) Ions from the Hofmeister series and osmolytes: Effects on proteins in solution and in the crystallization process. *Methods* 34(3):300–311.

279. Tadeo X, López-Méndez B, Castaño D, Trigueros T, Millet O (2009) Protein Stabilization and the Hofmeister Effect: The Role of Hydrophobic Solvation. *Biophys J* 97(9):2595–2603.

280. Zhang Y, Cremer PS (2006) Interactions between macromolecules and ions: the Hofmeister series. *Curr Opin Chem Biol* 10(6):658–663.

281. Baldwin RL (1996) How Hofmeister ion interactions affect protein stability. *Biophys J* 71(4):2056–2063.

282. Shortle D, Meeker AK, Gerring SL (1989) Effects of denaturants at low concentrations on the reversible denaturation of staphylococcal nuclease. *Arch Biochem Biophys* 272(1):103–113.

283. Horinek D, Netz RR (2007) Specific ion adsorption at hydrophobic solid surfaces. *Phys Rev Lett* 99(22):48–51.

284. Dzubiella J (2008) Salt-Specific Stability and Denaturation of a Short Salt-Bridge-Forming α-Helix. *J Am Chem Soc* 130(42):14000–14007.

285. Schneider CP, Shukla D, Trout BL (2011) Arginine and the hofmeister series: The role of ion-ion interactions in protein aggregation suppression. *J Phys Chem B* 115(22):7447–7458.

286. Crevenna AH, Naredi-Rainer N, Lamb DC, Wedlich-Söldner R, Dzubiella J (2012) Effects of hofmeister ions on the α-helical structure of proteins. *Biophys J* 102(4):907–915.

287. Dandliker WB, et al. (1967) The effect of chaotropic ions on the dissociation of antigen-antibody complexes. *Biochemistry* 6(5):1460–7.

288. Fox JM, et al. (2015) Interactions between Hofmeister anions and the binding pocket of a protein. *J Am Chem Soc* 137(11):3859–3866.

289. Paterová J, et al. (2013) Reversal of the Hofmeister Series: Specific Ion Effects on Peptides. *J Phys Chem B* 117(27):8150–8158.

290. Lo Nostro P, Ninham BW (2012) Hofmeister phenomena: an update on ion specifLo Nostro, P., & Ninham, B. W. (2012). Hofmeister phenomena: an update on ion specificity in biology. Chemical Reviews, 112(4), 2286–322. doi:10.1021/cr200271jicity in biology. *Chem Rev* 112(4):2286–322.

291. Bye JW, Falconer RJ (2013) Thermal stability of lysozyme as a function of ion concentration: A reappraisal of the relationship between the Hofmeister series and protein stability. *Protein Sci* 22(11):1563–1570.

292. Fink AL (1998) Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold Des* 3(1):9–23.

293. Carvalho SB, et al. (2013) Intrinsically Disordered and Aggregation Prone Regions Underlie β-Aggregation in S100 Proteins. *PLoS One* 8(10):1–11.

294. Uversky VN (2013) A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci* 22(6):693–724.

295. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6(3):197–208.

296. Uversky VN (2011) Intrinsically disordered proteins may escape unwanted interactions via functional misfolding. *Biochim Biophys Acta - Proteins Proteomics* 1814(5):693–712.

297. Schmitz KR, Bagchi A, Roovers RC, van Bergen en Henegouwen PMP, Ferguson KM (2013) Structural Evaluation of EGFR Inhibition Mechanisms for Nanobodies/VHH Domains. *Structure* 21(7):1214–1224.

298. Golinelli-Pimpaneau B, Luthi C, Christen P (2006) Structural Basis for D-Amino Acid Transamination by the Pyridoxal 5'-Phosphate-dependent Catalytic Antibody 15A9. *J Biol Chem* 281(33):23969–23977.

299. Zhang T, Wu Q, Sun HW, Rao J, Kannan K (2010) Perchlorate and iodide in whole blood samples from infants, children, and adults in Nanchang, China. *Environ Sci Technol* 44(18):6947–6953.

300. Blount BC, et al. (2009) Perinatal exposure to perchlorate, thiocyanate, and nitrate in New Jersey mothers and newborns. *Environ Sci Technol* 43(19):7543–7549.

301. Cooper MD, Alder MN (2006) The evolution of adaptive immune systems. *Cell* 124(4):815–22.

302. Otwinowski Z, Minor W (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* 276:307–326.

303. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.

304. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.

305. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–67.

306. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–42.

307. Evans S V (1993) SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J Mol Graph* 11(2):134–8, 127–8.

308. Pettersen EF, et al. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–12.

Appendix A
Crystallographic Data

Table 18: S25-39 crystal conditions

Details of crystallization conditions and resulting structures of S25-39 and each synthetic antigen. Common conditions are highlighted in like colours. Common conformations of CDR H3 observed in the resulting structures are highlighted in like colours (no relation to colours in Condition column). The highest resolution structure of each unique conformation (used for analysis and all figures) is identified with bold outlines in the final column.

S25-39 structures

Ligand	Condition	Drop type	Space group	Unit cell	Crystal form	Resolution	Conformation
None	0.05 M Zinc Acetate 0.05 M Sodium Cacodylate pH 6.5 15 % PEG 3350 10 % MPD	24-well hanging drop	$P2_1$	41.0, 83.2, 69.5 1 in AU	1	2.4	Unliganded #1 (Published – 3OKM)
	0.2 M Ammonium Acetate 0.1 M BIS-TRIS pH 5.5 45 % v/v (+/-)-2-Methyl-2,4,-pentanediol	24-well hanging drop	$P2_1$	55.7, 63.1, 62.4 1 in AU	2	1.9	Unliganded #1
	0.05 M Potassium Chlordide 0.01 M Magnesium Chloride 15 % PEG 6000	24-well hanging drop	$P2_1$	55.7, 63.5, 62.2 1 in AU	2	1.85	Unliganded #2
4-MeO-Kdo	0.1M Hepes pH 7.0 30% Jeffamine ED-2001	10mm dish hanging drop	$P2_1$	52.1, 63.6, 62.6 1 in AU	2	2.0	Liganded
4-O-Methoxymethyl-KdoOMe	0.2M Ammonium Acetate 0.1M Tris pH 8.5 25% PEG 3350	96-well sitting drop	$P2_1$	55.7, 62.3, 61.5 1 in AU	2	1.8	Unliganded #2
	0.15M Potassium Bromide	96-well	$P2_1$	55.6, 63.1, 61.9	2	1.3	Unliganded #2

	30% PEG MME 2000	sitting drop		1 in AU			
	0.15M Potassium Bromide 30% PEG MME 2000	24-well hanging drop	$P2_1$	52.0, 60.6, 63.2 1 in AU	2	2.0	Unliganded #1
	0.1M Sodium acetate trihydrate pH 4.5 25% PEG 3350	96-well sitting drop	$P2_1$	52.1, 61.7, 63.4 1 in AU	2	2.0	Unliganded #1
	0.1M BIS-TRIS pH 6.5 25% PEG 3350	96-well sitting drop	$P2_1$	55.8, 63.4, 62.0 1 in AU	2	1.9	Unliganded #2
	0.1M BIS-TRIS pH 6.5 20% PEG MME 2000	96-well sitting drop	$P2_1$	52.6, 61.2, 63.6 1 in AU	2	1.53	Unliganded #1
	0.2M Sodium chloride 0.1M BIS-TRIS pH 6.5 25% PEG 3350	96-well sitting drop	$P2_1$	52.8, 61.8, 63.6 1 in AU	2	2.0	Unliganded #1
4-O-benzyl-KdoOMe	0.1M Sodium acetate trihydrate pH 4.5 25% PEG 3350	96-well sitting drop	$P2_1$	55.8, 63.2, 61.9 1 in AU	2	1.53	Unliganded #1
	0.2M Sodium chloride 0.1M BIS-TRIS pH 6.5 25% PEG 3350	96-well sitting drop	$P2_1$	55.6, 62.9, 61.1 1 in AU	2	1.9	Unliganded #2
	0.05M Zinc acetate dihydrate 20% PEG 3350	96-well sitting drop	$P2_1$	41.1, 82.8, 137.7 2 in AU	3	1.85	Unliganded #1
	0.1M TRIS-HCL Ph 8.5 25% PEG 3000	24-well hanging drop	$P2_1$	55.7, 63.2, 61.7 1 in AU	2	1.4	Unliganded #2
	0.2M Ammonium fluoride 20% PEG 3350	96-well sitting drop	$P2_1$	55.6, 62.8, 62.0 1 in AU	2	1.9	Unliganded #2
4-O-Ethoxymethyl-KdoOMe	0.1M Sodium acetate trihydrate pH 4.5 25% PEG 3350	96-well sitting drop	$P2_1$	51.7, 61.1, 63.5 1 in AU	2	1.9	Unliganded #1
	0.1M Sodium HEPES pH 7.5 25% PEG MME 2000	24-well hanging drop	C2	101.6, 111.7, 154.7 4 in AU	4	2.1	Unliganded #3

Table 19: S25-2 crystal conditions

Details of crystallization conditions and resulting structures of S25-2 and each synthetic antigen. Common conditions are highlighted in like colours. Common conformations of CDR H3 observed in the resulting structures are highlighted in like colours. The highest resolution structure of each unique conformation (used for analysis and all figures) is identified with bold outlines in the final column.

S25-2 structures

Ligand	Condition	Drop type	Space group	Unit cell	Crystal form	Resolution	Conformation
4-O-Methoxymethyl-KdoOMe	0.2M Ammonium fluoride 20% PEG 3350	10mm dish hanging drop	$P2_1$	73.7, 135.3, 92. 7 4 in AU	5	2.15	Unliganded #3
	0.2M Ammonium iodide 20% PEG 3350	10mm dish hanging drop	$P2_1$	93.6, 135.5, 74.0 4 in AU	5	2.0	Unliganded #4
	0.2M Ammonium iodide 20% PEG 3350	10mm dish hanging drop	$P2_1$	73.9, 135.0, 93.0 4 in AU	5	2.20	Unliganded #4
4-O-benzyl-KdoOMe	0.2M Ammonium fluoride 20% PEG 3350	24-well hanging drop	$P2_1$	74.1, 134.8, 94.1 4 in AU	5	2.37	Unliganded #3

Table 20: Data collection and refinement statistics

Statistics are given only for the highest resolution structure solved of each unique conformation (see Table 18 and Table 19).

	S25-39				S25-2	
	4MeOKdo	Unliganded #1	Unliganded #2	Unliganded #3	Unliganded #3	Unliganded #4
Resolution (Å)	50.00-2.00 (2.07-2.00)	50.00-1.53 (1.56-1.53)	40.00 – 1.30 (1.32-1.30)	40.00-2.10 (2.18-2.10)	40.00-2.15 (2.19-2.15)	40.00-2.00 (2.03-2.00)
Space group	$P2_1$	$P2_1$	$P2_1$	C2	$P2_1$	$P2_1$
a (Å)	52.111	52.610	55.593	101.555	73.733	93.558
b (Å)	63.624	61.218	63.135	111.749	135.329	135.497
c (Å)	62.636	63.573	61.885	154.701	92.667	73.977
Molecules in AU	1	1	1	4	4	4
R_{sym}	0.071 (0.398)	0.109 (0.612)	0.057 (0.547)	0.070 (0.381)	0.083 (0.634)	0.092 (0.690)
R_{pim}	0.042 (0.248)	0.044 (0.323)	0.036 (0.331)	0.046 (0.251)	0.048 (0.364)	0.054 (0.388)
CC1/2	0.853	0.706	0.807	0.777	0.793	0.830
I/σ(I)	16.44 (2.06)	22.75 (2.27)	19.05 (2.23)	16.42 (3.04)	16.07 (1.96)	17.53 (2.61)
Completeness (%)	99.1 (95.1)	99.3 (86.4)	99.4 (99.4)	99.3 (98.5)	99.8 (100.0)	99.8 (100.0)
Redundancy	3.5 (3.0)	6.6 (4.1)	3.7 (3.7)	3.3 (3.2)	4.1 (4.1)	4.1 (4.2)
Unique reflections	27284	59940	103442	99437	98548	122420
Refinement						
R_{work} (%)	21.44	17.60	18.89	17.16	24.78	22.16
R_{free} (%)	26.14	21.17	22.54	20.19	27.80	25.45
no. waters	161	490	599	230	306	436
r.m.s bonds (Å)	0.0103	0.0227	0.0297	0.0197	0.0079	0.0098
r.m.s. angles (º)	1.4945	2.1787	2.6500	1.9593	1.2998	1.4414
Mean B ($Å^2$)	27.783	15.279	17.957	28.998	43.561	44.563

Appendix B
Permissions

The following license was granted by Springer via RightsLink for full-text reproduction from the publication:

Blackler RJ, et al. (2012) Antibody Recognition of Chlamydia LPS: Structural Insights of Inherited Immune Responses. Anticarbohydrate Antibodies, eds Kosma P, Müller-Loennies S (Springer Vienna, Vienna), pp 75–120.

License Number	3814390455897
License date	Feb 22, 2016
Licensed content publisher	Springer
Licensed content publication	Springer eBook
Licensed content title	Antibody Recognition of Chlamydia LPS: Structural Insights of Inherited Immune Responses
Licensed content author	Ryan J. Blackler
Licensed content date	Jan 1, 2012
Type of Use	Book/Textbook
Requestor type	Publisher
Publisher	University of Victoria
Portion	Full text
Format	Electronic
Will you be translating?	No
Print run	1
Author of this Springer article	Yes and you are the sole author of the new work
Title of new book	Thesis: Structural investigations into conformational diversity, polyspecificity, and binding mechanisms of near-germline antibodies
Publisher	University of Victoria
Author of new book	Ryan J. Blackler
Expected publication date of new book	May 2016
Estimated size of new book (pages)	160
Total	0.00 CAD