

Arabinofuranoside–Antibody Interactions: A Case Study of Furanoside–Protein Binding

by

Parnian Lak

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

© Parnian Lak, 2014

Abstract

Carbohydrate–protein interactions play key roles in a range of biological processes. Noticeable advancements have been made in understanding the features associated with the recognition of pyranosides (six-membered ring sugars) by proteins. In contrast, similar detailed studies about the interaction of furanosides (five-membered ring sugars) with proteins are scarce. Such interactions are increasingly recognized as important, particularly in host–pathogen recognition.

In this thesis, we provide a detailed investigation of a furanoside–protein interaction system, by studying the recognition of a cell-wall arabinofuranoside oligosaccharide fragment of mycobacteria by antibodies. The first binding system addressed is the interaction of a hexaarabinofuranoside antigen (Ara6) of mycobacterial lipoarabinomannan with the CS-35 antibody. Our approach was centered on developing CS-35 single chain variable fragment (scFv) as a useful tool to engineer the binding site. This enabled us to create a library of its mutants. Structural and binding aspects of various fragments interacting with Ara6 were studied from different aspects by various methods such as surface plasmon resonance (SPR), saturation transfer difference NMR spectroscopy, and circular dichroism spectroscopy. Ultimately, we determined the specificity motifs of this binding system including key amino acids, the epitope of the furanoside ligand, and the affinities and kinetics of the binding. In addition to the molecular-level study of Ara6–CS-35 system, we also investigated Ara6-CS-40 and Ara6-906.4321 systems from several aspects. ScFv technology, SPR spectroscopy, STD NMR spectroscopy, CD spectroscopy, and homology modelling assisted us in developing a more detailed picture of these interactions. Finally, we compared the three systems in search for the

common features in the recognition of Ara6 by these antibodies. Overall, this research serves to provide a more detailed understanding of the molecular recognition of furanosides by proteins.

Preface

This is an original document. The Ara6 antigen has been synthesized by Dr. Maju Joe in the Department of Chemistry at the University of Alberta. The crystal structure in Chapter 3 of this thesis (906.4321 antibody in combination with oligosaccharide antigen) was obtained by Dr. Kenneth Ng and Dr. Tomohiko Murase in University of Calgary. The sequences of the antibody variable fragments in Chapter 3 were received from Mr. Gareth Lambkin in the Biological Services of the Chemistry Department. The Saturation Transfer Difference of the CS-35 F_{ab} with the hexasaccharide antigen was obtained with the assistance of Dr. Margaret Johnson.

“The secret of life is molecular recognition; the ability of one molecule to recognize another through weak bonding interactions.”

- Linus Pauling

Acknowledgements

First and foremost, I sincerely thank my supervisor Prof. Todd L. Lowary for his support during the past five years. I acknowledge his constant efforts and guidance through this project, as well as the freedom and opportunity he provided to me to explore on my own.

I would like to acknowledge the Alberta Glycomics Center for providing the funding for this research.

I have been fortunate to know the wonderful people in Lowary's group; I appreciate their help, especially Dr. Maju Joe for providing the synthetic antigen, and Mr. Blake Zhang in the Biology Lab. Also, I thank Mr. Gareth Lambkin in Biological Services.

I warmly thank Prof. Robert Campbell and his group for the insightful discussions.

With sincere appreciations to our collaborators Drs. Kenneth Ng and Tomohiko Murase in University of Calgary for the crystallographic data, and Dr. John Klassen's group for the mass spectrometry data.

I am also indebted to the people in the Chemistry Department, with whom I have interacted in the course of my graduate studies. I am grateful to the NMR facility, in particular Mr. Mark Miskolzie, Mass Spectrometry facility, in particular Dr. Randy Whittal and Mr. Bela Reiz, and Analytical Facility, in particular Dr. Wayne Moffat.

I acknowledge the thesis committee Dr. David Bundle, Dr. Harry Brumer, Dr. Christopher Cairo, Dr. Robert Campbell, and Dr. Julianne Gibbs-Davis.

I appreciate my friends, whose friendship helped me stand through hard times.

I would like to express my deep gratitude to my family; to my parents and my sister, for their constant love, encouragement, and belief in me; and to my fiancé, Siavash, for proofreading my thesis, and for always being a source of love, support, and patience.

Table of Contents

Chapter 1 : Introduction.....	1
1.1 Overview and Premise	1
1.2 Carbohydrate-Binding Proteins.....	2
1.3 Antibody Fragments.....	3
1.4 Single Chain Variable Fragments (scFv)	5
1.5 Kinetics and Thermodynamics of Carbohydrate–Protein Interactions.....	6
1.6 Molecular Forces in Carbohydrate–Protein Interactions	7
1.6.1 Hydrogen Bonds	7
1.6.2 Interactions Involving Charged Groups.....	9
1.6.3 Hydrophobic Interactions	10
1.6.4 Role of Water.....	13
1.7 Carbohydrate–Antibody Interactions: General Insights	14
1.7.1 Interaction of the Se155-4 Antibody with the <i>Salmonella</i> Serogroup B O- Antigen.....	15
1.7.2 Interaction of S-20-4 Antibody with <i>Vibrio cholera</i> O1 Polysaccharide Antigen 20	
1.7.3 Interaction of Antibodies with <i>Shigella flexneri</i> Cell-Wall LPS Antigens.....	24
1.7.4 Summary.....	30
1.8 The Interaction of CS-35 Antibody with Mycobacterial Arabinofuranoside Antigen	31
1.9 Thesis Research Overview	37
References	37
Chapter 2 : Specificity of the CS-35 Antibody to the Ara6 Antigen.....	44
2.1 Introduction	44
2.2 Materials and Methods	46
2.2.1 Construction of CS-35 scFv Genes.....	47
2.2.2 Transformation, Expression, and Refolding of scFv Fragments.....	47
2.2.3 Protein Purification.....	48
2.2.4 Generation of the CS-35 ScFv Mutants Library.....	48

2.2.5 Circular Dichroism (CD) Spectroscopic Study of the CS-35 Antibody Fragments.....	49
2.2.6 Confirmation of the Presence of Disulfide Bonds in the scFv	49
2.2.7 Binding of the CS-35 Antibody Fragments to Ara6 by Surface Plasmon Resonance (SPR) Spectroscopy	49
2.2.8 Studies of the CS-35 scFv by ESI-MS	50
2.2.9 Binding of the CS-35 Fragments and Mutants to 2-4 by Saturation Transfer Difference NMR Spectroscopy.....	50
2.3 Results.....	51
2.3.1 Generation of the CS-35 scFv, Structure.....	51
2.3.2 Intra-Domain Disulfide Bonds	55
2.3.3 Generation of CS-35 scFv Mutants	57
2.3.4 SPR Binding of Wild-Type CS-35 Fragments to Ara6	58
2.3.5 CS-35 scFv Mutants Binding.....	61
2.3.6 Structural Studies by Mass Spectrometry.....	64
2.3.7 STD NMR Spectroscopy of Wild-Type and Mutant Proteins with 2-3	68
2.4 Discussion.....	74
2.4.1 CS-35 scFv Structure and Stability.....	74
2.4.2 Discussion of Binding	76
2.4.3 Inter-Molecular Forces and Specificity Motifs in the CS-35 scFv–Ara6 System.....	77
2.4.4 Comparison of the Ara6–CS-35 System with Previous Pyranoside–Antibody Interaction Systems.....	83
2.5 Conclusions	85
References	85

Chapter 3 : Characterization of the Binding of the CS-40 and 906.4321

Antibodies to the Ara6 Antigen	88
3.1 Introduction	88
3.2 Materials and Methods	90
3.2.1 Construction of CS-40 and 906.4321 scFv Genes.....	90
3.2.2 Transformation, Expression, and Refolding of scFv Fragments	91

3.2.3 Protein Purification and Characterization.....	92
3.2.4 Circular Dichroism (CD) Study of CS-40 and 906.4321 Antibody Fragments	92
3.2.5 Binding of CS-40 and 906.4321 scFvs to Ara6 by Surface Plasmon Resonance (SPR) Spectroscopy	93
3.2.6 Saturation Transfer Difference NMR Spectroscopy.....	94
3.2.7 Modeling the Binding Sites of CS-40 and 906.4321 Antibodies	94
3.3 Results.....	95
3.3.1 Generation of CS-40 and 906.4321 scFvs	95
3.3.2 Circular Dichroism Study of Antibody Fragments.....	97
3.3.3 Analysis of the binding of CS-40 and 906.4321 scFv by SPR Spectroscopy	98
3.3.4 STD NMR Analysis of the Ara6 Epitope in Binding to 906.4321 F _{ab}	101
3.3.5 ScFv Binding Site Models.....	102
3.4 Discussion	111
3.5 Conclusions	118
References	119
Chapter 4 : Conclusions and Future Directions	121
4.1 Summary: Generation and Binding Specificity of scFv Mutants	121
4.2 Future Work: Optimization of CS-40 and 906.4321 scFv Production	123
4.3 Future Work: Thermodynamic Binding Parameters	124
4.4 Future Work: Crystal structures of scFvs	124
4.5 Future Work: Mutational Studies on CS-40 and 906.4321 scFvs.....	125
4.6 Future Work: Applying the Present Information in the Design of New Tools for Studying Mycobacterial Disease	125
References	126
Appendices	135
A1. List of Primers	135
A2. FPLC results for the monomer mutants that bind to Ara6.....	137
A3. SDS PAGE results for CS-35 scFv mutants.....	140
A4. CD Spectrum for Phe95Ala in Far UV.....	143
A5. SPR Fittings into 1:1 Binding Model.....	144

A6. Equilibrium Binding Analyses.	148
A7. Homology Results for CS-40 and 906.4321 by Swiss modeling.....	151
A8. Table of chemical shifts assignment of Ara6.....	153
A9. Sequences of scFv fragments and alignments	154
A10. Control experiments for STD NMR	160

List of Tables

Table 1.1 General characteristics of strong hydrogen bonds	8
Table 2.1 Mutants of CS-35 scFv, and the molecular interactions that are affected upon mutation.....	57
Table 2.2 SPR binding affinity of the wild-type CS-35 scFv with 2-3 . Affinity was obtained from association and dissociation rates (k_a and k_d , respectively), and also from equilibrium binding analysis according to Equation (2.1).	59
Table 2.3 Mutants and their binding to ligand 3 in SPR.....	61
Table 2.4 Kinetics and affinity evaluation of the six scFv mutants showing binding to 2-3 from different complementary determining regions (CDR) for ligand 2-3 from SPR experiments.	63
Table 3.1 Binding constants and association and dissociation rates of CS-40 and 906.4321 scFvs to 3-3 at determined by SPR spectroscopy. Affinities were calculated from association and dissociation rates, and also from equilibrium binding analysis	100
Table 3.2 Comparison of the binding profiles of three scFvs (CS-40, 906.4321, and CS-35) to Ara6 antigen from SPR spectroscopic results.	112
Table 3.3 Comparison of key amino acids involved in Ara6 binding in CS-40, 906.4321 and CS-35.....	115

List of Figures

Figure 1.1 a) Schematic structure of antibody. H = heavy; L= light; C = constant; V = variable' b) Complementary determining regions in heavy chain and light chain	4
Figure 1.2 Schematic structure of scFv.....	5
Figure 1.3 A typical intermolecular hydrogen bond.....	8
Figure 1.4 Hydrogen bond formation of a galactopyranose residue with planar side chains of amino acids.....	9
Figure 1.5 Structure of <i>N</i> -acetyl glucosamine	9
Figure 1.6 Structure of neuraminic acid. The carboxylate moiety can be involved in ionic interactions and hydrogen bonds.....	10
Figure 1.7 a) Schematic depiction of hydrophobic (B) and hydrophilic (A) faces of galactopyranose b) Stacking interaction of Trp and Phe residues in Pokeweed Lectin-D2 in combination with a trisaccharide ligand	11
Figure 1.8 Carbon backbone of Neu5Ac in pink	11
Figure 1.9 Examples of topologies for CH- π interactions. ⁵⁹	13
Figure 1.10 Structure of L-arabino pyranoside, and D-galacto pyranoside.....	14
Figure 1.11 Trisaccharide epitope (1-1) of <i>Salmonella</i> serotype B O-chain LPS, which is recognized by the Se1554 antibody.....	16
Figure 1.12 Se155-4 F _{ab} binding site in combination with its trisaccharide ligand (2-1) and a structured water molecule (showed by red x).	17
Figure 1.13 Structure of the synthetic tethered derivatives of the trisaccharide antigen recognized by Se155-4 antibody	18
Figure 1.14 Discrepancy in the Gal-Man conformation of (1-1) between scFv ⁷⁰ and F _{ab} ⁶⁵ crystal structures.....	19
Figure 1.15 Structure of the perosamine-containing monosaccharide (1-2) and disaccharide (1-3) epitopes used in crystallization experiments.....	21
Figure 1.16 S-20-4 F _{ab} hydrophobic binding site in combination with <i>Vibrio cholera</i> cell-wall monosaccharide (1-2) antigen	22
Figure 1.17 S-20-4 F _{ab} in combination with the terminal disaccharide derivative	22

Figure 1.18 A Tyr and a Trp residue in S-20-4, in close proximity to the monosaccharide ligand (1-2).....	23
Figure 1.19 Structure of the synthetic pentasaccharide ABCDA' (1-4) and synthetic trisaccharide BC*D (1-5) antigens used in X-ray crystallographic studies of the SYA/J6 F _{ab}	25
Figure 1.20 SYA/J6 F _{ab} in combination with the ABCDA' pentasaccharide antigen.....	26
Figure 1.21 Structure of a pentasaccharide AB(E)CD (1-6) unit that is recognized by the F22-4 antibody.....	28
Figure 1.22 F22-4 F _{ab} fragment in combination with <i>Shigella flexneri</i> decasaccharide antigen.....	29
Figure 1.23 Pro95 and Met96 in <i>cis</i> peptide bond in F22-4 F _{ab} fragment in combination with <i>Shigella flexneri</i> decasaccharide antigen.....	30
Figure 1.24 Composite structure of mycobacterial LAM.....	32
Figure 1.25 Ara6 antigen present in both mycobacterial AG and LAM.....	33
Figure 1.26 Regions of the Ara6 antigen recognized by CS-35 as determined by STD-NMR spectroscopy. The strongest saturation transfer effects are seen with residues A, B, C, and E.....	34
Figure 1.27 Main CDR domains of CS-35 F _{ab} in combination with Ara6.....	35
Figure 1.28 CS-35 F _{ab} in complex with Ara6.....	35
Figure 1.29 Phe95 and Gly96 of the CDR H3 in the structure of CS-35 F _{ab}	36
Figure 2.1 Structure of Ara6 in LAM (2-1), and binding studies (2-2, 2-3, and 2-4).....	46
Figure 2.2 SDS PAGE (29 KDa) and ESI mass of wild type scFv.....	52
Figure 2.3 FPLC results of the CS-35 scFv. The peak at 18 mL belongs to the monomer; the dimer elutes at 14 mL.....	52
Figure 2.4 The far UV CD spectrum of the wild-type CS-35 scFv before and after adding 4, The near UV CD spectrum of the wild-type CS-35 scFv before (green) and after adding 2-4.....	53
Figure 2.5 CD spectroscopic results in far UV region for CS-35 F _{ab} (a) and mAb (b).....	54
Figure 2.6 Aromatic residues of the CS-35 F _{ab} binding site around 2-2.....	55
Figure 2.7 Schematic depiction of intra-domain disulfide bonds, and the reduction and alkylation reaction.....	56

Figure 2.8 ESI results after treating scFv with IAA in the presence and absence of the reducing agent	56
Figure 2.9 The residues of the CS-35 binding site that were mutated	58
Figure 2.10 SPR binding of the wild-type CS-35 scFv with 2-3	58
Figure 2.11 SPR binding of the CS-35 F _{ab} with 2-3	59
Figure 2.12 Residues around ring A.	60
Figure 2.13 SPR binding of the CS-35 scFv dimer with 2-3	61
Figure 2.14 Six essential residues in the binding site of the CS-35 F _{ab}	62
Figure 2.15 SPR binding of a) Asn34Asp, b) Tyr50Ala, and c) Tyr98Phe to Ara6 2-3	64
Figure 2.16 ESI mass depiction of two conformations of the CS-35 scFv	65
Figure 2.17 His ionization equilibrium	66
Figure 2.18 ESI-MS study of the CS-35 scFv at three pH conditions: pH : 3 (a), pH : 8 (b), and pH : 6.8 (c).....	66
Figure 2.19 Two charge states of CS-35 mutants in ESI mass.....	67
Figure 2.20 ES-MS spectrum of the CS-35 scFv lacking the extra sequence	68
Figure 2.21 ¹ H NMR spectrum of 2-4 (a), STD experiment of the CS-35 F _{ab} and 2-4	69
Figure 2.22 STD NMR spectrum of the CS-35 scFv (a) and its four mutants (b) with 2-4	70
Figure 2.23 Residues that were studied by STD NMR upon mutation to Ala.....	71
Figure 2.24 Protons that are mostly affected by mutations.....	73
Figure 2.25 Hydrogen bond network around ring E	74
Figure 2.26 CH–π interactions of Tyr98, Phe95, and Trp33 with 2-2 from the crystal structure of the CS-35 F _{ab}	78
Figure 2.27 Phe95–Gly96–Asn97–Tyr98 segment of the CS-35 F _{ab} around 2-2 in the X-ray structure of the complex	79
Figure 2.28 Asn97 and ring C of 2 in the X-ray structure of the CS-35 F _{ab}	80
Figure 2.29 a) Hydrogen bond making residues around ring E in the crystal structure of the CS-35 F _{ab} and 2-2 . b) The distances of the interactions.....	81
Figure 2.30 Asn34 close to ring A	82
Figure 2.31 Position of residue Ser31, indirectly critical for binding of CS-35 F _{ab} to 2-2	83
Figure 2.32 Similarity of the Trp and His residues interactions with the oligosaccharide antigens in the CS-35–Ara6 system, and Se155-4–trisaccharide system.....	84

Figure 3.1 Ara6-containing structures found in LAM (3-1) and compounds used in binding studies with the CS-35, CS-40, and 906.4321 antibodies (3-2 to 3-5).....	89
Figure 3.2 SDS PAGE results for the CS-40 scFv (Left), and the 906.4321 scFv (Right).....	95
Figure 3.3 FPLC results for 906.4321 scFv (a), CS-40 oligomer (b), and CS-40 scFv monomer (c).....	96
Figure 3.4 Far UV circular dichroism results with CS-40 (a) and 906.4321 (b) scFv fragments. Near UV CD spectroscopic results with CS-40 mAb (c) and 906.4321 F _{ab} (d) with and without Ara6.	97
Figure 3.5 a) SPR curves for the binding of CS-40 b) Equilibrium binding analysis.....	99
Figure 3.6 a) SPR curves for the binding of 906.4321 b) Equilibrium binding analysis.....	100
Figure 3.7 ¹ H 1D NMR spectrum of 3-4 , and the STD NMR spectrum in the presence of the 906.4321 F _{ab} . An STD NMR experiment with the CS-35 F _{ab} and 3-4	101
Figure 3.8 Homology model of the 906.4321 scFv.....	103
Figure 3.9 Superposition of 906.4321 and CS-35 variable domains.	104
Figure 3.10 Aromatic residues in the 906.4321 binding site.....	105
Figure 3.11 Similar residues in 906.4321 binding site, in comparison to the CS-35.....	106
Figure 3.12 Model of CS-40 scFv.....	107
Figure 3.13 CS-40 superposition with CS-35 variable domains.....	108
Figure 3.14 Aromatic residues in the binding site of CS-40 antibody model.....	109
Figure 3.15 Phe–Pro–Phe–Tyr sequence in CDR H3 domain of CS-40.....	110
Figure 3.16 Similar residues in CS-40 in comparison to the CS-35 antibody.....	110
Figure 3.17 Asn34 of CS-35 is substituted by Tyr 59 in CS-40, extending its OH group in close proximity to ring A, and stacking against ring A.....	114
Figure 3.18 Superposition of similar Asn residues in CS-35 and 906.4321.....	116
Figure 4.1 Ara6 structure.	122

List of Abbreviations

Å	Ångstrom
Abe	Abequose
ABP	Arabinose Binding Protein
Ala	Alanine
Ara6	Hexa Arabinofuranoside
Araf	Arabinofuranoside
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
CD	Circular Dichroism
CDR	Complementary Determining Region
CM5	Carboxy methyl dextran 5
ConA	Concanavalin A
Cys	Cysteine
D-Gal	D-Galactose
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
ESI-MS	Electrospray Ionization Mass Spectrometry
ELISA	Enzyme-Linked Immunosorbent aAssay
F _{ab}	Fragment antigen binding
F _c	Fragment constant

FPLC	Fast Protein Liquid Chromatography
Glc _p	Glucopyranose
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
His	Histidine
IAA	Iodoacetamide
Ig	Immunoglobulin
Ile	Isoleucine
ITC	Isothermal Titration Calorimetry
kDa	Kilodalton
k _a	Association rate
K _A	Association Constant
k _d	Dissociation rate
<i>M. avium</i>	<i>Mycobacterium avium</i>
<i>M. leprae</i>	<i>Mycobacterium leprae</i>
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>
<i>M.tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
MBSU	Molecular Biology Services Unit
Met	Methionine
mg	Milligram
MHz	Mega Hertz

mL	Milliliter
mM	Millimolar
Neu5Ac	<i>N</i> -acetyl Neuraminic acid
Ng	Nanogram
nm	Nanometer
nM	Nanomolar
NMR	Nuclear Magnetic Resonance
NTA	Nitrilotriacetic acid
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
Phe	Phenylalanine
ppm	Parts Per Million
Pro	Proline
Rhap	Rhamnopyranose
R _{eq}	Equilibrium Response
R _{max}	Maximum Response
RU	Resonance Unite (Response Unit)
scFv	Single chain Fragment variable
Ser	Serine
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
STD	Saturation Transfer Differenc
TB	Tuberculosis
Thr	Threonine

Trp	Tryptophan
Tyr	Tyrosine
μg	Microgram
μM	Micromolar
UV	Ultraviolet
Val	Valine
V _H	Variable Heavy chain
V _L	Variable Light chain
WT	wild type

Chapter 1 : Introduction

1.1 Overview and Premise

Carbohydrate–protein interactions play crucial roles in physiological and pathophysiological functions and numerous biological events, including cell adhesion, cellular growth, fertilization, endocytosis of deleterious glycoconjugates, inflammation, and immunological responses.¹ They also play important roles in host–pathogen recognition² and contribute towards the mechanisms by which certain pathogens escape from the immune response.¹ Cell migration, apoptosis, and metastasis are other processes that carbohydrate–protein interactions regulate.³

Proteins that recognize carbohydrates should be able to discriminate among a large number of sugar structures with various stereochemistry of hydroxyl groups and linkages. A considerable amount of research has focused on deciphering the recognition aspects of carbohydrates by lectins (carbohydrate binding proteins) and antibodies, so that there is now a deep understanding of the fundamentals of many such interactions.⁴ Different aspects of carbohydrate–protein binding including affinity, structural features, thermodynamics, and molecular forces have been discussed.⁵⁻⁸ However, most of these studies have focused on the recognition of pyranosides⁵ (six membered ring carbohydrates) by proteins, and similar information about how furanosides (five-membered ring carbohydrates) interact with proteins is more limited than those information for pyranosides.^{9,10} Although furanosides are increasingly recognized as important biological components,¹¹⁻¹⁴ the key elements in their molecular recognition by proteins have stayed enigmatic.

In this thesis I describe our efforts to provide a more detailed picture of furanoside–protein interactions and the molecular motifs that govern this recognition. Our approach centers on the development of single chain variable fragments of monoclonal antibodies that recognize a mycobacterial hexaarabinofuranoside (Ara6) antigen as a tool to study the furanoside–antibody

binding characteristics by a combination of techniques. In the next sections of this chapter, I will briefly overview the essential concepts necessary to understand this challenge.

1.2 Carbohydrate-Binding Proteins

Carbohydrate-binding proteins include enzymes, membrane transport proteins, bacterial periplasmic binding proteins involved in sugar transport, lectins, and antibodies.⁵

The concept of the specific interaction of proteins with glycans goes back to Emil Fischer, who referred to the recognition of enzymes for their glycan substrates as a "lock and key" interaction.¹² Enzymes indeed shaped the early history of glycan-protein interaction studies. Lysozyme, an endoglycosidase that degrades bacterial cell walls, was the first carbohydrate binding protein whose crystal structure was solved¹³ by X-ray crystallography. Since then, X-ray structures of many other carbohydrate-protein complexes have been solved.^{14,15}

Lectins are another class of glycan binding proteins that have been extensively studied. Lectins by definition (from *lectus*, to select or choose) are carbohydrate-binding proteins, and exist in most living organisms, from viruses and bacteria to plants and animals.¹⁶ Typically, enzymes or antibodies that recognize carbohydrates are not classified as lectins. Lectins interact with carbohydrates non-covalently and reversibly with high specificity accompanied by few conformational changes in the protein. Lectins usually bind to di- tri- and tetrasaccharides with significantly higher affinities than monosaccharides. The binding site is usually a shallow depression on the surface of the protein, and in some lectins close to a Ca^{2+} or Mg^{2+} ion. Because of this water-exposed, shallow binding site, the affinity of lectins for monosaccharides is low.¹⁶ A good example is the interaction of influenza haemagglutinin lectin with sialic acid, which is the best studied example of a viral carbohydrate binding protein. This interaction is 1000 times weaker than the affinity of the interaction of this sugar for the enzyme neuraminidase in the same virus.¹⁷ However, lectins show exquisite affinity for oligosaccharides through multivalent recognition.⁵ Therefore, haemagglutinin lectin binding to glycans is enhanced by avidity through binding to multiple carbohydrate units. These interactions involve water-mediated hydrogen bonds, direct hydrogen bonds, and van der Waals forces.⁵ Concanavalin A (ConA) binding to mannose oligosaccharides is an outstanding representative of carbohydrate-binding proteins. This system has been extensively studied and has significantly contributed to our understanding

of carbohydrate–protein interactions.¹⁸ The crystal structure of ConA, as well as the crystal structure of influenza virus hemagglutinin, are of historical significance.¹⁹ The studies of Lemieux and Kabat on the combining sites of lectins and antibodies towards specific blood group antigens have provided critical information about the concepts of such interactions.¹⁹

Among carbohydrate-binding proteins, antibodies and lectins have served as tremendously valuable tools to unravel molecular principles underlying the recognition of carbohydrates by proteins without an enzymatic activity for their ligands⁵. In the following sections I will focus mostly on the general concepts of bacterial carbohydrate–antibody interactions, which are relevant to the subject of this thesis.

1.3 Antibody Fragments

Antibodies or immunoglobulins (IgA, IgD, IgE, IgG, IgM) are glycoproteins with critical roles in the immune system, which act by identifying foreign organisms such as bacteria and viruses.²⁰ Each antibody consists of two heavy (H) chains and two light (L) chains (**Figure 1.1**). Each heavy chain has three constant (C) domains and one variable (V) domain. In turn, each light chain has one constant domain and one variable domain. F_{ab} (Fragment antigen binding) is a proteolytic fragment of an antibody that contains one complete light chain, paired with one heavy chain fragment, which includes a variable domain, and only one constant domain. F_c (Fragment constant) is also a proteolytic fragment of an antibody that contains four constant domains from heavy chains.²⁰ The F_c domain recruits immunological elements through interaction with the complement system and provides long serum half life. Antibodies recognize specific antigens via their binding sites, also called paratopes, located at the upper tips of their “Y” shaped structure. Paratopes recognize particular epitopes displayed on specific antigens, which are complementary to their shape.²¹ Each antibody binding site consists mainly of six hypervariable loops called complementary determining regions (CDRs) within a β -sheet framework.²²⁻²⁵ There are three CDRs in the light chain (L1, L2, L3) and three in the heavy chain (H1, H2, H3). The CDR H3 has the greatest sequence variation among CDRs in antibodies, and often makes more contribution to antigen recognition than the other five CDRs.²⁶

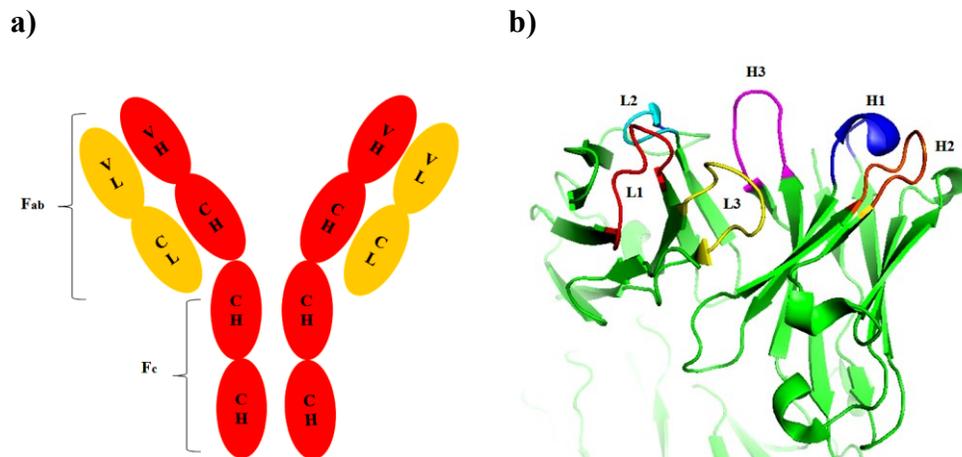


Figure 1.1 a) Schematic structure of antibody. H = heavy; L= light; C = constant; V = variable? F_{ab} = Fragment antigen binding; F_c = Fragment constant.²⁷ **b)** Complementary determining regions in heavy chain and light chain (PDB code: 3HNS)¹⁰ Reproduced from Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *Journal of Molecular Biology* **2009**, 392, 381, permission obtained

Antibodies that bind to carbohydrate antigens have, in general, low to moderate intrinsic affinity. The ability to bind to two antigens significantly increases the functional affinity of an antibody (also called avidity). Since 1975, when Kohler and Milstein introduced the production of antibodies with defined specificity through hybridoma technology, monoclonal antibodies (mAbs) have been increasingly considered as valuable research tools, as they can be produced in high quantities, bind specifically to an antigen, and are easily standardized.^{28,29}

However, monoclonal antibodies have several limitations in structural studies.²¹ First, the technology of monoclonal antibody production is laborious and time consuming. Second, kinetic studies of mAbs are more complicated than those of the corresponding smaller fragments due to having two binding sites and avidity effects. Difficulties in studying mAbs by mutational studies and manipulation of the binding site is among other limitations of using full length antibodies for molecular studies. Moreover, the microheterogeneity of mAbs, which is the result of glycosylation patterns on the F_c domain, complicates crystal structure and mass spectrometric studies. These drawbacks have led to numerous investigations to produce smaller fragments of antibodies²¹ for binding and structural studies.

The F_{ab} , which can be produced by the enzymatic removal of the F_c domain, is a common antibody fragment used in structural studies. IgGs can be digested into F_{ab} domains through

proteolysis with enzymes such as papain and pepsin.²⁰ They can further be engineered into fragments such as single chain variable fragments (scFv), single V_H or V_L domains, and diabodies (bivalent scFv fragments).³⁰ Advances in recombinant antibody technology during the past decades have significantly facilitated the genetic manipulation of antibody fragments.^{31,32}

1.4 Single Chain Variable Fragments (scFv)

The scFv fragment of an antibody consists of the variable regions of heavy (V_H) and light (V_L) chains, connected together by a flexible peptide linker²¹ (**Figure 1.2**). An scFv has one antigen binding site, i.e., it is monovalent. ScFvs are preferred over mAbs and F_{ab} fragments for detailed studies of antibody–antigen interactions due to the possibility they provide to manipulate binding sites by site-directed mutagenesis. Such manipulation also allows for the identification of fragments with improved affinity for the antigen. Another advantage of the scFv technology is the convenient and efficient expression of the protein in *E. coli*.²¹ Moreover, the small size of scFvs allows these fragments to penetrate more rapidly and evenly into tumors, and other tissues, as compared to full length antibodies, with less possibility of developing immune response.^{33,34}

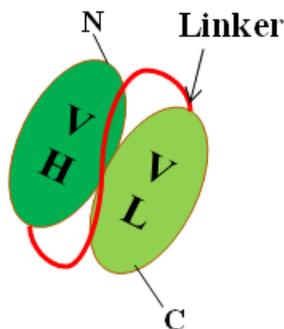


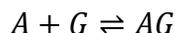
Figure 1.2 Schematic structure of scFv

However, there are some important considerations in generating the scFv of an antibody, such as proper folding, stability, and solubility. The scFv lacks the constant domains that stabilize the structure of the F_{ab} fragment and mAbs. Therefore, the main challenge in the production of a functional scFv is to reduce the effect of deleting the constant domains on the

stability of the scFv to the minimum. Various constructs of V_H-V_L or V_L-V_H domains and linker lengths have been studied by Huston et al.³⁵ The flexible linker used to connect the two variable domains allows the proper folding of the binding site without disturbing the antigen binding ability of the domains.³⁶ The amino acid composition of the linker should provide enough flexibility and polarity to avoid interaction with V_H and V_L domains. The most popular scFv linkers have sequences consisting of stretches of Gly and Ser residues, e.g. V_H-(Gly₄Ser)₃-V_L. Charged residues such as Glu and Lys also have been used to enhance solubility.³⁷

1.5 Kinetics and Thermodynamics of Carbohydrate-Protein Interactions^{19,38}

The monovalent binding of an antibody to a glycan can be defined by a simple equation:



The rate of the forward reaction is called association rate (k_a) or on-rate (k_{on}), and the rate of the backward reaction is called dissociation rate (k_d) or off-rate (k_{off}). The association or affinity constant (K_A) is therefore:

$$K_A = \frac{k_{on}}{k_{off}}$$

The affinity constant, K_A , is related to the concentration of the antibody and glycan:

$$K_A = \frac{[AG]}{[A][G]}$$

The affinity constant is also related to thermodynamic parameters:

$$\Delta G = \Delta H - T \Delta S = -RT \ln K_A$$

Favorable enthalpy changes are considered to be the result of the formation of hydrogen bonds and hydrophobic interactions at the combining site. However, the effect of the entropic component on K_A has been debated. Lemieux considered entropy as a favorable factor in carbohydrate-protein recognition, and postulated it to originate from water removal from the binding surface.³⁹ However, other studies suggest that entropic effects are explained by the restriction of movement of molecules upon binding.^{40,41} Such restriction in movement and flexibility is therefore considered as an unfavorable factor in binding.

There are many different ways to investigate glycan-protein interactions. Each approach has its advantages and disadvantages. Methods such as X-ray diffraction and NMR spectroscopy,

as well as Surface Plasmon Resonance (SPR), together with isothermal titration calorimetry, mass spectrometry and computational analyses, are the most important techniques that have served to unravel the principles of these interactions.¹⁹

1.6 Molecular Forces in Carbohydrate–Protein Interactions

Biomolecular recognition in water is suggested to occur in two steps: 1) encounter, which is the orientation of two molecules by non-specific forces to form a hydrophobic encounter complex; and 2) docking with specific interaction formation.^{42,43}

High affinity antibodies that are highly specific toward their cognate antigens are proposed to have a "lock and key" binding type.⁴⁴ However, more cross-reactive antibodies employ induced fit or conformational changes to optimize complementarity in binding.⁴⁵ Carbohydrates are polar compounds and, as would be expected from their structures, usually make an abundant number of specific hydrogen bonds with proteins. However, nonpolar interactions also play an important role in building a stable carbohydrate–protein complex.⁵ Research aimed at understanding the relative contributions of each of these forces is always desired. Ideal studies should give insights into the participation of these forces in solution, considering the dynamic nature of binding as well as the role of water.

1.6.1 Hydrogen Bonds⁴⁶

Hydrogen bonding is a salient feature of the binding of carbohydrates and proteins, and the energies associated with these interactions lie between covalent bonds and van der Waals interactions. In an antibody–carbohydrate complex, these hydrogen bonds are shielded from the bulk solvent, and exist in a lower dielectric environment. Such hydrogen bonds are enthalpically stronger than those formed between the ligand and water.⁴⁷ Hydrogen bond's strength differs as a function of atomic distance and direction. These characteristics are summarized in **Table 1.1** and **Figure 1.3** for a typical $\text{OH}\cdots\text{O}=\text{C}$ hydrogen bond.

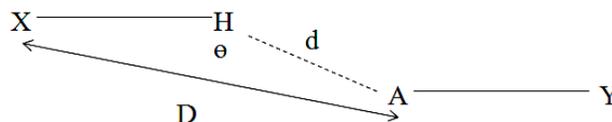


Figure 1.3 A typical intermolecular hydrogen bond. D is the distance between two electronegative atoms, d is the hydrogen bond length. θ is the X–H–A angle.

Table 1.1 General characteristics of strong hydrogen bonds

Bond Energy (–kcal/mol)	4–15
D (Å)	2.5–3.2
d (Å)	1.5–2.2
θ (°)	130–180

Sugar hydroxyl groups can act as the donor of a single hydrogen bond, and as the acceptor of one or two hydrogen bonds. The ring oxygen can act as a hydrogen bond acceptor. Different protein side-chain and main-chain groups participate in forming hydrogen bonds with carbohydrates. The torsional freedom of sugar hydroxyl groups permits optimization of hydrogen bond formation with planar carbonyl groups of proteins, although fixing of hydroxyl groups is accompanied by an entropic cost. Hydrogen bond donors in proteins are primarily reported to be the main-chain amide groups and the side-chain amide group of asparagine and, less frequently, glutamine. The carbonyl or carboxylate groups of the main chain, or aspartic acid and glutamic acid, are frequent hydrogen bond acceptors. Hydroxyl groups from tyrosine, serine, and threonine are less common as either the donors or the acceptors of hydrogen bonds with sugar hydroxyl groups. The fact that the hydroxyl groups of proteins are less frequently participating in hydrogen bonding suggests that the entropic cost of fixing both the sugar and the protein hydroxyl groups is large.⁵

Carbohydrate–protein interactions can also involve cooperative hydrogen bonds. Planar amino acid side chains (e.g., carboxylates and amides) make an effective hydrogen bond geometry for vicinal sugar hydroxyl groups in an equatorial/equatorial or equatorial/axial configuration¹⁶ (**Figure 1.4**).

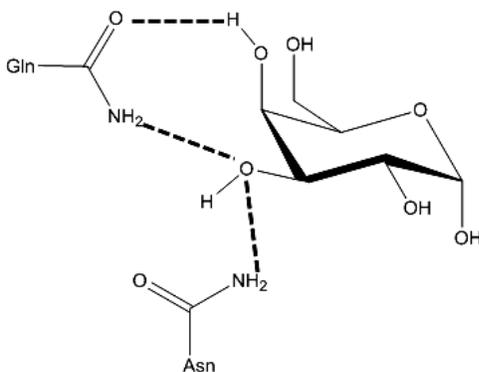


Figure 1.4 Hydrogen bond formation of a galactopyranose residue with planar side chains of amino acids (here glutamine and asparagine).

The N–H group of an amide in a carbohydrate, e.g., *N*-acetyl-glucosamine, (**Figure 1.5**) can act as a hydrogen bond donor to a carbonyl or carboxylate oxygen. Alternatively, the acetamido substituent often accepts hydrogen bonds.

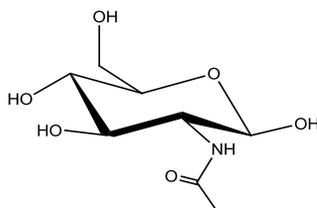


Figure 1.5 Structure of *N*-acetyl glucosamine

Unlike hydroxyl groups, the nitrogen and the carbonyl oxygens of the acetamido substituent have a fixed, planar geometry. Therefore the energetic penalty of fixing hydroxyl groups of amino acids, such as serine, in a hydrogen bond is not seen for amino acids that use their side chain carbonyl or amide groups for hydrogen bonding.⁴⁸

1.6.2 Interactions Involving Charged Groups

N-Acetylneuraminic acid (Neu5Ac) (**Figure 1.6**) and some of its derivatives^{49,50} that bind to lectins⁵⁰ have been extensively studied for the interaction of charged sugars with proteins. The

interaction of the Neu5Ac carboxylate of a trisaccharide receptor with arginine side chains in the polyoma virus is a good example of charge–charge interactions,⁵² which are required for the specificity of the recognition. It was shown that this Neu5Ac residue is critical for polyoma infectivity, by making several interactions with the receptor, including a salt bridge between the carboxylate group of the Neu5Ac and the guanidinium group of Arg77. This key interaction helps polyoma virus be distinguished from other structurally relevant viruses that are not infectious. However, although these charge–charge interactions are important, the carboxylate moiety has been reported to be involved mostly as an acceptor in essential hydrogen bonds with polar side chains, main chain amides, and ordered water molecules in the binding site.⁵

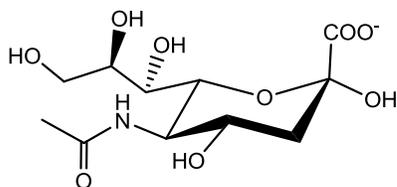


Figure 1.6 Structure of *N*-acetyl-neuraminic acid. The carboxylate moiety can be involved in ionic interactions and hydrogen bonds

1.6.3 Hydrophobic Interactions⁵

Although carbohydrates are highly polar molecules, they can also interact with proteins through hydrophobic interactions. The aliphatic side chains of proteins do not interact frequently with sugars; however, the delocalized electron cloud of an aromatic residue often favorably stacks against the sugar aliphatic protons. The best representative of such interactions is the packing of the galactose hydrophobic face (B-face) against one or more aromatic side chains of the protein such as phenylalanine or tyrosine (**Figure 1.7**).

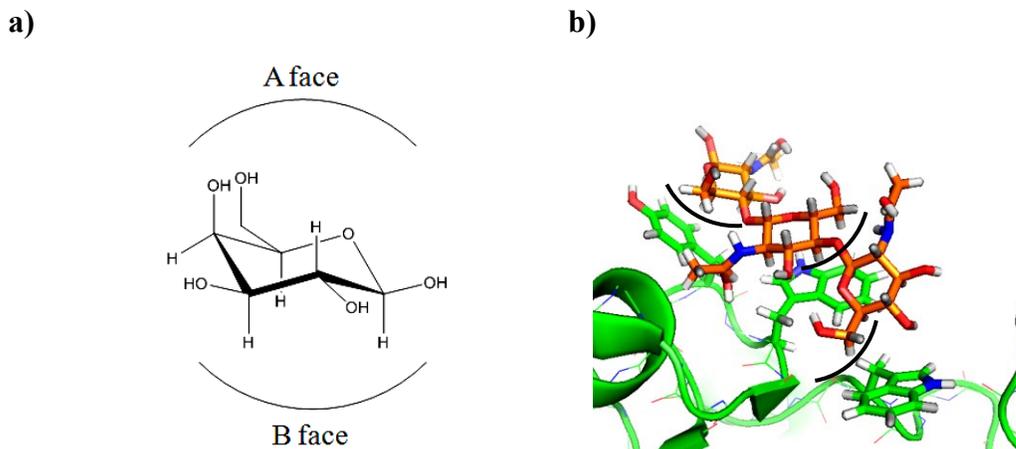


Figure 1.7 a) Schematic depiction of hydrophobic (B) and hydrophilic (A) faces of galactopyranose b) Stacking interaction of Trp and Phe residues in Pokeweed Lectin-D2 in combination with trio-*N*-acetylchitotriose ligand, the hydrophobic interfaces are shown in black (PDB code: 1ULM)⁵³

The methyl group of the acetamido moiety of *N*-acetyl-glucosamine and *N*-acetyl-galactosamine can also interact with an aromatic ring by hydrophobic forces. The carbon backbone of the glycerol moiety of *N*-acetyl-neuraminic acid also serves as a nonpolar surface to protein aromatic side chains (**Figure 1.8**).

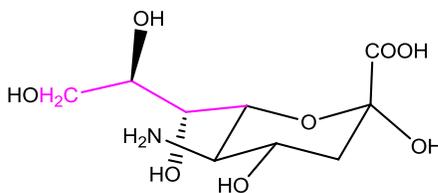


Figure 1.8 Carbon backbone of the glycerol moiety neuraminic acid in pink

CH- π bonds are important carbohydrate-protein interactions that have attracted increasing attention in recent years.⁵⁴ These interactions were first studied in systems other than carbohydrates. In 1952 Tamres, by using calorimetric methods, reported that the interaction of a carbon acid and a π -system is attractive, and thus, chloroform dissolves in benzene exothermically.⁵⁵ For carbohydrates, Quijcho and Vyas first reported that D-galactose and D-

maltose are sandwiched between two aromatic amino acid residues in many carbohydrate–protein interactions.^{47,56} However, the importance of CH– π bonds in carbohydrate–protein interactions was first reported by Haratara and Muraki in the binding of Lysozyme to tri-*N*-acetylchitotriose.^{57,58} It has also been suggested that CH– π interaction is essential for glycosidase activity; removal of the aromatic residues in *Trichoderma reesei* cellulase enzyme reduced the ligand binding free energy and led to the deactivation of the enzyme.⁵⁹ The high number of aromatic residues – Trp, Tyr, Phe, and His – in the binding sites of the carbohydrate binding antibodies has also been frequently reported, which will be discussed in the following sections. Moreover, such interactions have been highlighted in the interaction of carbohydrate binding modules with their substrates.¹⁴ Carbohydrate binding modules are sequences in enzymes that process carbohydrates. Such sequences are in the catalytic site, or close to the catalytic site of the enzyme that can bind to carbohydrate ligands.¹⁴ It has been shown that the carbohydrate binding of such modules is assisted by multiple hydrophobic amino acids that can get involved in CH– π interaction.¹⁴

A single CH– π interaction has an energy of 1.5–2.5 kcal mol⁻¹. However, cooperativity is a salient feature of CH– π recognition. Carbohydrates usually make multiple CH– π interactions with a protein, which results in significant interaction energies when considered in aggregate. Electronegative atoms adjacent to the CH group make the C–H bond sufficiently polar to be attracted by the electron cloud of the aromatic ring. The greater the ability of the CH group to share the hydrogen, the larger the stabilizing effect of a CH– π bond. Therefore, CH groups in carbohydrates, which are close to multiple electronegative oxygen atoms, are well suited for forming such interactions. A noticeable characteristic of CH– π interactions is that they are present in both polar protic solvents such as water as well as in nonpolar solvents. This capability distinguishes the CH– π bonds from conventional hydrogen bonds, which can be disturbed in the presence of water. Directionality is also an important aspect of CH– π interactions, which makes them distinctive from London dispersion forces. The aromatic side chain of amino acids face the CH groups in different topologies, and sometimes one sugar residue becomes sandwiched between two aromatic side chains by multiple CH– π interactions (**Figure 1.9**).

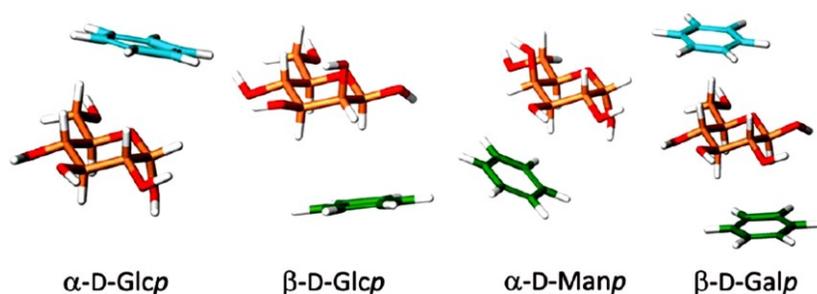


Figure 1.9 Examples of topologies for CH– π interactions.⁶⁰ Reproduced from Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. s. *Accounts of Chemical Research* **2012**, *46*, 946, permission obtained.

1.6.4 Role of Water

The role of water in the carbohydrate–protein binding has been investigated.⁶¹ Because the shape of the ligand is complementary to the binding site of the protein, it can displace small water molecules to form a stable complex. Upon ligand binding, water is released from the surface of the formed complex. Therefore, the protein–ligand binding has been considered as a desolvation phenomenon. Water release contributes favorably to the enthalpy and free energy of binding.⁶² Water molecules act both as hydrogen bond donors and acceptors. Lemieux reviewed his ideas on the role of water in molecular recognition in an interesting paper in 1996.³⁹ He considered water as a catalyst that make activated complexes with the protein and the ligand, which are necessary for molecular association, but is regenerated after the complex formation.

However, some water molecules bound to the free protein remain in the binding site after binding to the ligand, and make critical hydrogen bonds that stabilize the complex. Water-mediated hydrogen bonds between carbohydrates and proteins can be as strong as direct protein–sugar hydrogen bonds.⁶³ These water molecules serve as structural elements of binding site architecture and contribute to the specificity of the binding.⁵ The enthalpy gain that results from such extra water-mediated hydrogen bonds might be greater than the entropic penalty paid for the immobilization of the involved water molecules.⁶⁴ Water molecules can form an intricate network to make the antibody–antigen interface fully complementary. Therefore, a water-mediated hydrogen bond network has a stabilizing effect on ligand–protein interactions.

A good example of the role of water in the binding specificity is the interaction of the L-arabinose binding protein (ABP) with L-arabinose and D-galactose (**Figure 1.10**).⁵⁵ The protein

has a slightly higher affinity for L-arabinose. High-resolution crystal structures of the complexes⁵⁵ revealed that both ligands adopt similar orientations in the ABP binding pocket. However, replacement of D-galactose by L-arabinose results in the CH₂OH group of D-galactose being replaced in the complex by one water molecule.

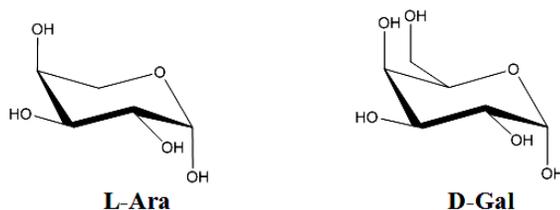


Figure 1.10 Structure of L-arabino pyranoside, and D-galacto pyranoside

This water molecule makes one extra water-mediated hydrogen bond in the L-arabinose–ABP complex. This leads to an increase in the specificity of the protein for L-arabinose in comparison to D-galactose. The thermodynamic impact of this change is small, which means that at least for this protein, the entropic cost of adding a water molecule to the binding site is similar to the cost of fixing the CH₂OH group of D-galactose. Water molecule affects the specificity of this interaction and also slightly increases the affinity of the protein for L-arabinose.

1.7 Carbohydrate–Antibody Interactions: General Insights

Various aspects of antibody–carbohydrate recognition have been studied to date, which I will briefly review. Detailed descriptions of such interactions are ideally achieved by structural analysis and binding assessment of the antibody–carbohydrate complexes from different aspects. A few crystal structures of bacterial oligosaccharide–antibody systems have been reported to date. Site-directed mutagenesis of scFv fragments, molecular modeling, surface plasmon resonance (SPR), NMR spectroscopy, microcalorimetry, and mass spectrometry are among the most common methods that have been used to investigate different aspects of carbohydrate–antibody interactions. In the following sections I will briefly overview the studies on the bacterial carbohydrate–antibody interactions that have contributed the most to our current

knowledge of this biomolecular interaction. These systems are all representatives of pyranoside–antibody interactions. I will end the subject with a summary of a bacterial furanoside–antibody recognition system.

1.7.1 Interaction of the Se155-4 Antibody with the *Salmonella* Serogroup B O-Antigen

Salmonella is a Gram-negative bacterium, responsible for food poisoning.⁶⁵ Like other Gram-negative bacteria, the cell wall contains lipopolysaccharide (LPS). The polysaccharide portion of the serogroup B LPS is recognized by the monoclonal antibody Se1554.⁶⁶ This interaction system represents the most detailed and comprehensive study available for bacterial carbohydrate–antibody interactions. Various structural and energetic aspects of this recognition have been investigated. Indeed, the crystal structure of the Se155-4 F_{ab} fragment in combination with an oligosaccharide antigen, reported by Bundle and coworkers in 1991, was the first crystal structure resolved for a carbohydrate–antibody complex.⁶⁷ This structure revealed a pocket-shaped combining site dominated by Trp and His residues, and a structured water molecule. The discovery of multiple aromatic amino acid residues in the combining site of the antibody, distinguished it from the previously studied carbohydrate-binding proteins (lectins,, enzymes).

The dodecasaccharide antigen consisted of three tetrasaccharide repeating units of the LPS O-chain $[\rightarrow 3)\text{-}\alpha\text{-D-Galp}(1\rightarrow 2)\text{-}[\alpha\text{-D-Abep}(1\rightarrow 3)]\text{-}\alpha\text{-D-Manp}(1\rightarrow 4)\text{-}\alpha\text{-L-Rhap}(1\rightarrow)]_3$, in which Abe stands for abequose, or 3,6-dideoxy-D-xylo-hexose. It was shown that the epitope recognized by Se155-4 is a branched trisaccharide (**1-1**) consisting of the Gal, Abe and Man residues (**Figure 1.11**).

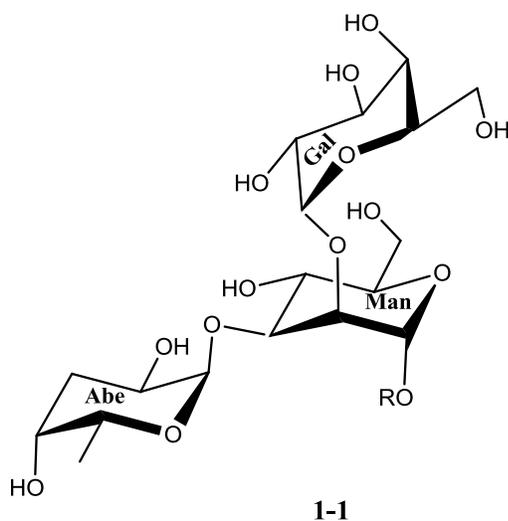


Figure 1.11 Trisaccharide epitope (**1-1**) of *Salmonella* serotype B O-chain LPS, which is recognized by the Se1554 antibody.

In this complex, the Abe residue is buried in a hydrophobic pocket, while the Gal and Man residues are more water exposed and are involved in hydrogen bonds with the protein (**Figure 1.12**). The binding site is complementary to the shape of the trisaccharide. The bound conformation of the ligand was similar to the observed conformations for the free ligand, previously obtained by NMR studies.⁶⁸

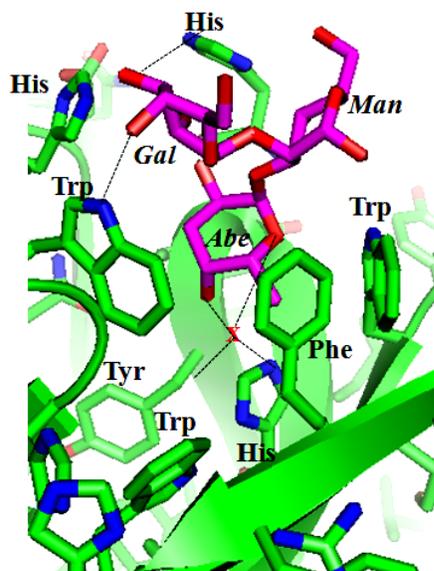


Figure 1.12 Se155-4 F_{ab} binding site in combination with its trisaccharide ligand (**1**) and a structured water molecule (showed by red x). The Abe residue is buried in the binding site facing multiple aromatic residues. (PDB code: 1MFE)⁶⁷

An interesting study on this system was the characterization of the entropic costs resulting from the immobilization of glycosidic linkages in the bound state.⁶⁸ The flexibility of the glycosidic linkages can be considered to result in an unfavorable energetic loss upon binding. To address this, tethered synthetic ligands in which glycosidic bonds of the Abe–Gal–Man trisaccharide were restricted to mimic the bound state (**Figure 1.13**) were synthesized.

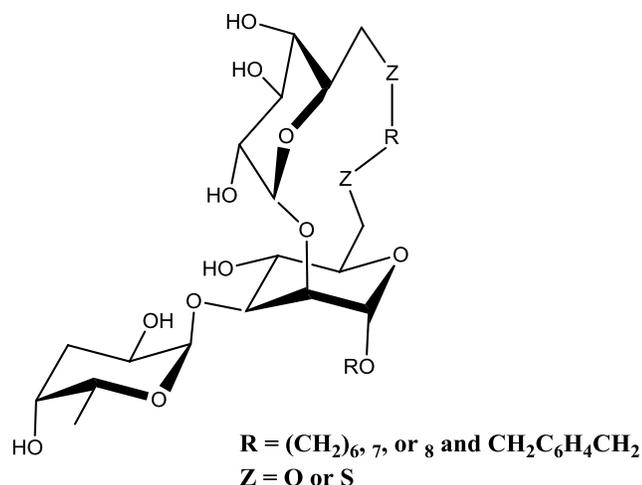


Figure 1.13 Structure of the synthetic tethered derivatives of the trisaccharide antigen recognized by Se155-4 antibody

This preorganized structure was supposed to decrease the energy cost of binding. However, calorimetric studies with the Se155-4 F_{ab} fragment showed that that ΔG did not change substantially after freezing the trisaccharide in the tethered compounds. This modification had little effect on ΔH and ΔS . Indeed, the fact that the entropy gain did not improve in these systems indicated that oligosaccharides may display a restricted range of conformations. This result was consistent with Lemieux's postulate about the role of entropy in the binding energy.

Calorimetric studies also confirmed that there are substantial hydrophobic interactions at the combining site.⁷⁰ This was consistent with the crystal structure, which had shown to be rich in hydrophobic residues. Abequose, being a dideoxy sugar, was shown to be the most important residue in the binding energy. It was shown to be the main contributor to the nonpolar interaction with Se155-4 antibody by stacking against two key Trp residues. This hydrophobic interaction and the displacement of water molecules upon binding provide the main source of binding energy.

Detailed epitope mapping of this recognition was accomplished by using a synthetic panel of monodeoxy oligosaccharide derivatives, to probe the importance of various hydrogen bonds. It was shown that a few direct and water-mediated hydrogen bonds were essential for the recognition.⁷¹

Three years after the aforementioned structure, the scFv fragment of Se155-4 was crystallized (PDB code: 1MFE) in combination with the same Gal–Abe–Man trisaccharide antigen.⁷² This structure was important for comparison with the F_{ab}. It was shown that the positions of all amino acid residues that interact with the trisaccharide in the scFv were almost the same as the positions in the F_{ab}. However, while the Abe–Gal linkage is exactly the same in two complexes, the flexible Gal–Man linkage adopts a different conformation in the scFv binding site compared to the F_{ab}–trisaccharide complex (**Figure 1.14**).

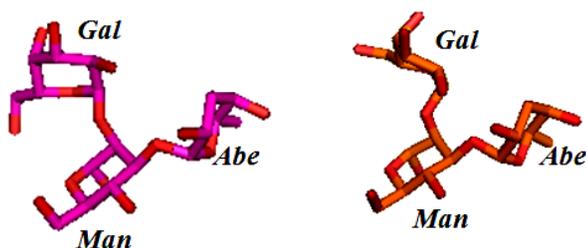


Figure 1.14 Discrepancy in the Gal–Man conformation of (1-1) between scFv⁷² and F_{ab}⁶⁷ crystal structures

This was suggested to be the result of crystal packing forces. The affinity of the interaction for the scFv–trisaccharide complex was shown to be in close agreement with the affinity of the F_{ab}–trisaccharide complex. Both K_D values were in the micromolar range, which is typical for carbohydrate–antibody interactions.⁷³ The behavior of the scFv fragment was consistent in both solution and gas phase. A gas phase (mass spectrometric) study of the scFv association and dissociation to the trisaccharide antigen showed that the conformation of the ligand is retained in both solution and in the gas phase when it is in complex with the scFv.⁷⁴ The hydrogen bonds and hydrophobic interactions were preserved in the gaseous phase to result in the same torsional angles as the crystal structure.

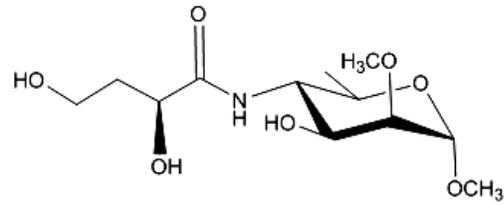
Production of the scFv fragment of Se155-4 was indeed an important step toward the characterization of the antibody paratope and kinetics of the binding. The scFv was used as a tool for generating a library of mutants among which mutants with improved on-rates were generated.⁷³ The Se155-4 scFv was further studied for the influence of valence on the kinetics of binding, by comparing the binding profiles of Se155-4 scFv monomer, scFv dimer, and mAb to the sugar antigen by SPR.⁷³ It was emphasized that the traces of scFv dimer in a monomer

sample could lead to slower off-rates and faster on-rates, which complicate the binding profile. Therefore, the scFv monomer is always desired for the kinetics and affinity studies of the antibody binding to the carbohydrate antigen.

1.7.2 Interaction of S-20-4 Antibody with *Vibrio cholera* O1 Polysaccharide Antigen

Vibrio cholera is a Gram-negative bacterium; the *Ogawa* serotype is responsible for cholera disease.⁷⁶ The S-20-4 antibody recognizes the O-chain of the LPS with an affinity in the micromolar range as determined by SPR.⁷⁷ Binding studies of S-20-4 IgG with synthetic methyl glycoside derivatives of the carbohydrate antigen revealed that the non-reducing terminal *N*-acylated perosamine (4-amino-4,6-dideoxy-2-*O*-methyl-D-mannose **Figure 1.15**) residue (**1-2**) accounts for 90% of the recognition.⁷⁸ The 2-*O*-methyl group of was also shown to be vital for binding. The crystal structure of S-20-4 F_{ab} unliganded, and in combination with the synthetic mono- and disaccharide (**3**) analogs of the O-antigen, were reported in 2000 (**Figures Figure 1.15, Figure 1.16, and Figure 1.17**).⁷⁷

1-2:



1-3:

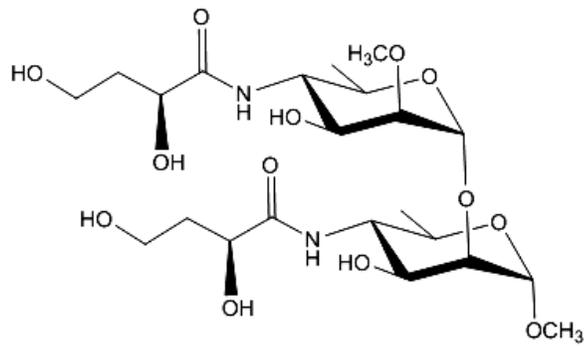


Figure 1.15 Structure of the perosamine-containing monosaccharide (1-2) and disaccharide (1-3) epitopes used in crystallization experiments

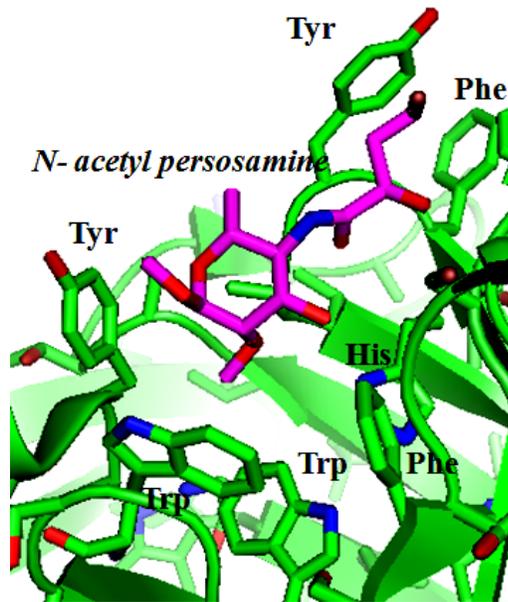


Figure 1.16 S-20-4 F_{ab} hydrophobic binding site in combination with *Vibrio cholera* cell-wall monosaccharide (1-2) antigen (PDB code: IF4X)⁷⁵

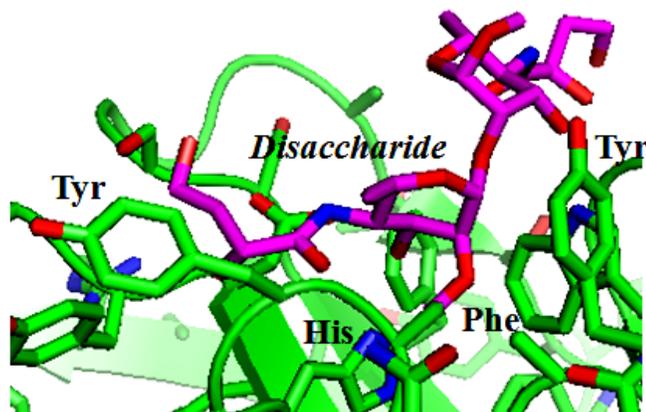


Figure 1.17 S-20-4 F_{ab} in combination with the terminal disaccharide derivative (1-3). The second residue extends out of the binding site. (PDB code: IF4Y)⁷⁵

Comparison of the unliganded structure of S-20-4 with the complex containing the monosaccharide showed a small conformational change for the antibody combining site upon binding. The *N*-acylated perosamine binds in an extended cavity facing central hydrophobic

pocket in proximity to Trp, Tyr, Ala, and His residues. Upon binding, Tyr 32 and Trp 91 move slightly toward the terminal perosamine residue (**Figure 1.18**). This perosamine residue also becomes stabilized by six hydrogen bonds to Asp and His residues of the antibody.

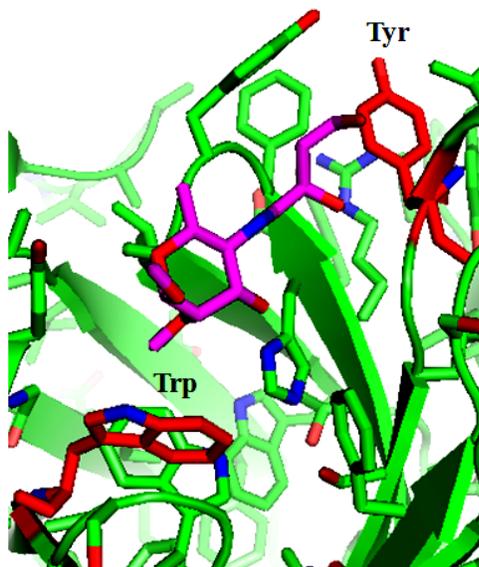


Figure 1.18 A Tyr and a Trp residue in S-20-4, in close proximity to the monosaccharide ligand (1-2) (PDB code: IF4X).⁷⁵

The disaccharide structure (**Figure 1.17**) showed that the second *N*-acylated perosamine residue extends away from the antibody binding site. Therefore, this residue makes fewer contacts with the paratope compared to the first residue. Additional sugar residues extending from the reducing-end perosamine would be positioned out of the binding cavity without making contacts to the antibody binding site.

The structural specificity for the monosaccharide epitope observed in the crystal structure confirmed the previously-mentioned binding studies using the synthetic analogs of the antigen.⁷⁸ This carbohydrate–antibody interaction system was interesting because one sugar residue was responsible for most of the affinity, as for many of the other carbohydrate–antibody complexes more than one sugar is required. In addition, the architecture of the binding site was well tailored for combining hydrophobic and polar interactions to maximize the contact with a monosaccharide antigen.

1.7.3 Interaction of Antibodies with *Shigella flexneri* Cell-Wall LPS Antigens

Shigellosis, or bacillary dysentery, arises from the destruction of the colonic mucosa caused by the Gram-negative bacterium *Shigella*.⁷⁹ Antibodies directed against the bacterial cell wall LPS were shown to induce immunity against this pathogen. In this section, I present carbohydrates from two serotypes of *S. flexneri*, which have been studied for carbohydrate–antibody interaction.⁸⁰

S. flexneri serotype Y is recognized by SYA/J6 antibody, which binds to a pentasaccharide epitope consisting of four Rhap moieties and on GlcpNAc residue (**Figure 1.19**). The epitope, typically abbreviated as ABCDA' (**1-4**), is a linear oligosaccharide and is composed of one repeating unit plus an additional Rhap residue from an adjacent repeating unit in the LPS O-chain. The crystal structure (**Figure 1.20**) of SYA/J6 F_{ab} in combination with a synthetic ABCDA' derivative (**Figure 1.19**), as well as in the unbound state was resolved in 2002.⁸¹ Moreover, the crystal structure of a modified antigen (BC*D) (**1-5**), where the C2-OH of rhamnose C was deoxygenated (**Figure 1.18**), in complex with SYA/J6 F_{ab} was obtained.⁸¹ This modified antigen had been shown to bind to the antibody with typical micromolar affinity, but about 7- and 17-fold tighter than those with the ABCDA' and the natural trisaccharide (BCD) epitopes, respectively.⁸²

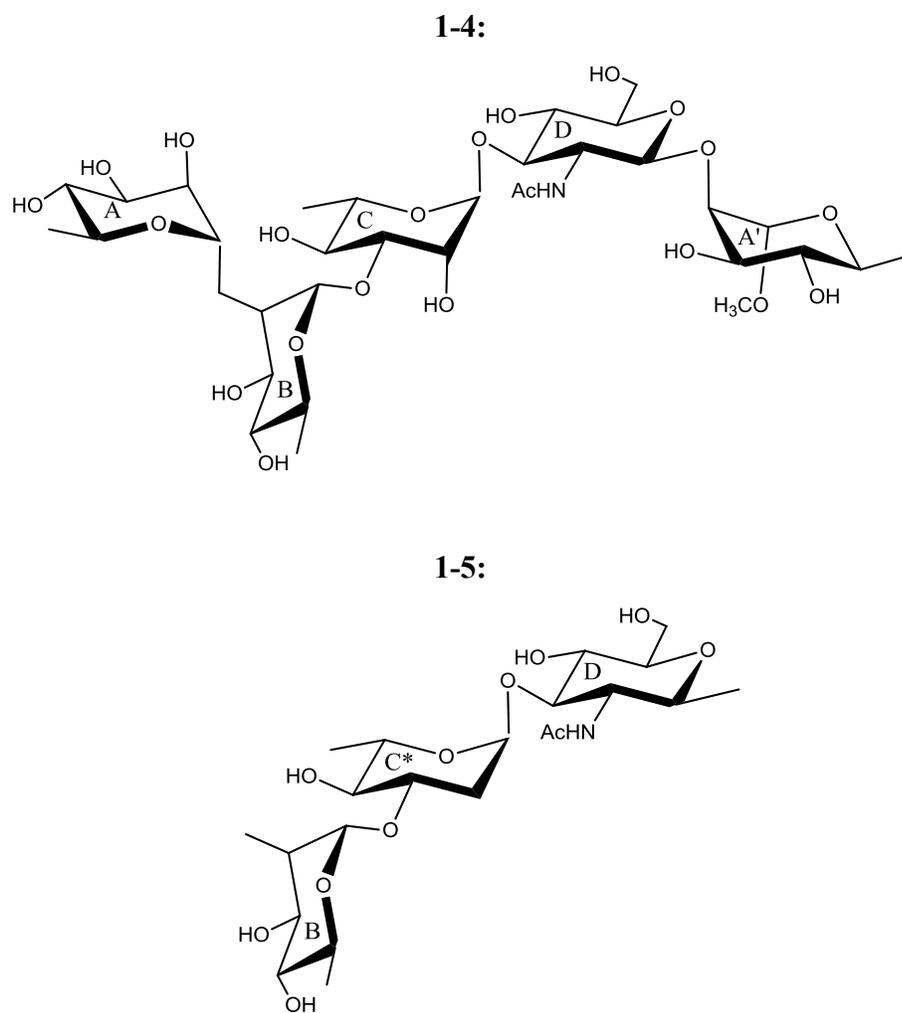


Figure 1.19 Structure of the synthetic pentasaccharide ABCDA' (**1-4**) and synthetic trisaccharide BC*D (**1-5**) antigens used in X-ray crystallographic studies of the SYA/J6 F_{ab}

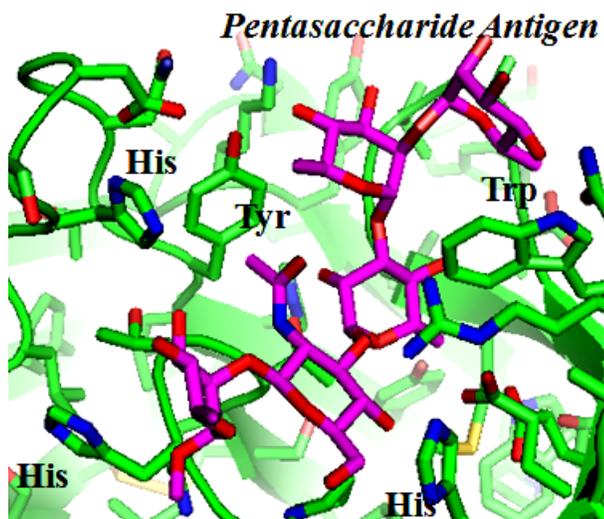


Figure 1.20 SYA/J6 F_{ab} in combination with the ABCDA' pentasaccharide antigen (PDB code: 1M7I)⁷⁹

Both antigens showed good complementarity to the binding site. The CDR H3 domain of the antibody was shown to make the highest number of contacts with both antigens. The CD and C*D rings made the dominant contribution to the binding by forming van der Waals interactions with residues such as Trp, Tyr, and Ala. Residue C and C* were buried at the deepest center of the binding site, completely inaccessible to the solvent and made the most hydrophobic contacts with the antibody, while residues B, D and especially A are solvent accessible. A significant number of the polar residues including Glu were also detected in the binding site. GlcNAc D made the highest number of hydrogen bonds and water-mediated hydrogen bonds among other residues, mainly with Glu and Arg residues. The conformation of the antigen had been previously studied by NMR spectroscopy, and comparison between the crystal structure and NMR data showed that the sugar antigen undergoes conformational changes upon binding, in particular around the C–D and the D–A glycosidic linkages.⁸³

Based on the crystal structure and the obvious contribution of rings C and C* to the formation of the contact surface, it was suggested the higher affinity of the BC*D ligand is because of improved nonpolar interactions resulting from the removal of the 2-OH in residue C*. This enhances the fitting of ring C* in the binding groove, compared to the ring C counterpart in the natural ligand.⁸¹ Moreover, BC*D fits more snugly in the binding site, resulting in improved

hydrogen bonds, including an important-water mediated hydrogen bond that is missing in the ABCDA'-antibody complex.

Calorimetric studies showed that the binding of the BCD and BC*D ligands were enthalpy-driven, and the enthalpy contribution to the binding affinity for BC*D was greater than the natural segment BCD. In contrast, however, the ABCDA' binding is entropy-driven, which was suggested to be because of the displacement of water from the binding site by residues A and A'.⁸¹

S. flexneri serotype 2a is another example of a studied carbohydrate-antibody interaction.⁸¹ A number of oligosaccharide fragments of serotype 2a O-antigen were synthesized⁸⁴⁻⁸⁶ and it was shown that the specificity for this antibody arises from a branched pentasaccharide moiety, where the third rhamnose is modified by a glucose, designated as AB(E)CD (**1-6**) (**Figure 1.21**). Among the antibodies that were tested against serotype 2a antigen, F22-4 showed to be unique in its ability to bind to the different oligosaccharide fragments of serotype 2a. The crystal structure of F22-4 F_{ab} unliganded, and in combination with the synthetic decasaccharide [AB(E)CD]₂ and pentadecasaccharide [AB(E)CD]₃ fragments of the *S. flexneri* 2a O-antigen were reported in 2008.⁷⁹

1-6:

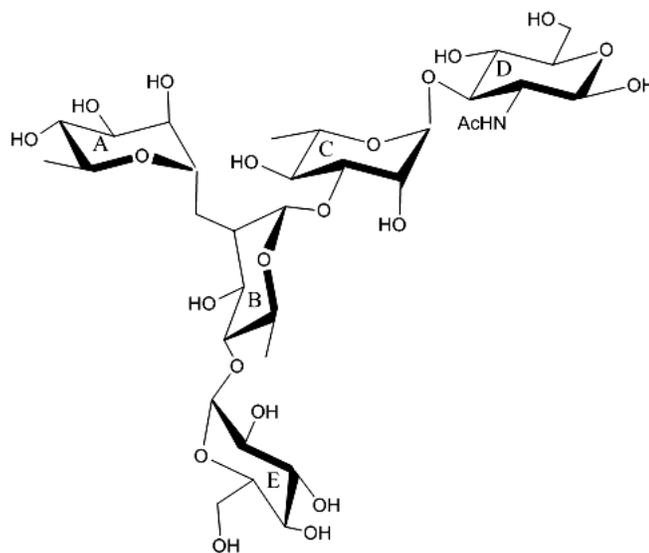


Figure 1.21 Structure of a pentasaccharide AB(E)CD (1-6) unit that is recognized by the F22-4 antibody

The binding site (**Figure 1.22**) consists of two cavities. One of these cavities holds the branching Glc residue, E. In this aspect, it is similar to Se155-4, which accommodates the branching Abe residue of the epitope in a cavity. In both of the liganded structures, the epitope was a nine carbohydrate residue unit [AB(E)CDA'B'(E')C']. This epitope includes six sugar units, the non-reducing rhamnose units and both branching glucose units, making direct contact with the antibody groove-shaped binding site.

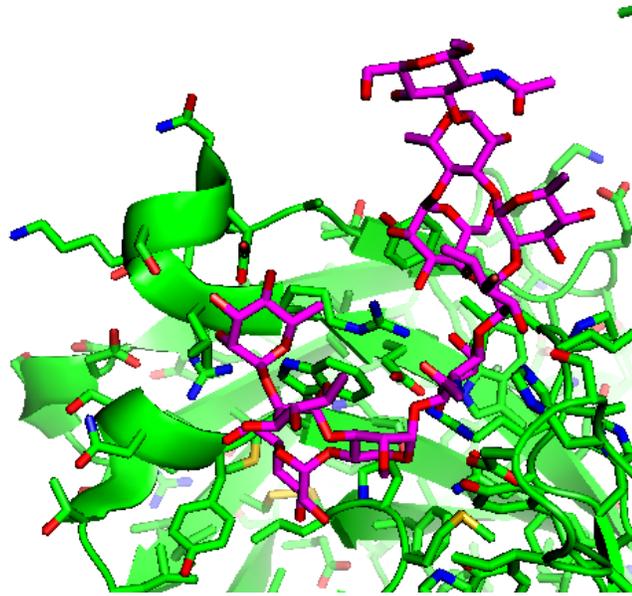


Figure 1.22 F22-4 F_{ab} fragment in combination with *Shigella flexneri* deca-saccharide antigen (PDB code: 3BZ4)⁷⁷

An interesting feature of the oligosaccharide was that it formed a compact helical structure that was accommodated in the antibody-combining site. GlcNAc D and Glc E formed the most of the contact via multiple hydrogen bonds, water mediated hydrogen bonds, and van der Waals interactions. These results were later confirmed by saturation transfer difference (STD) NMR analysis and molecular dynamic calculations.⁸⁷ An interesting structural aspect of the antibody, on the other hand, was a *cis* peptide bond detected between Pro 95 and Met 96 of the heavy chain in both liganded and unliganded structures.

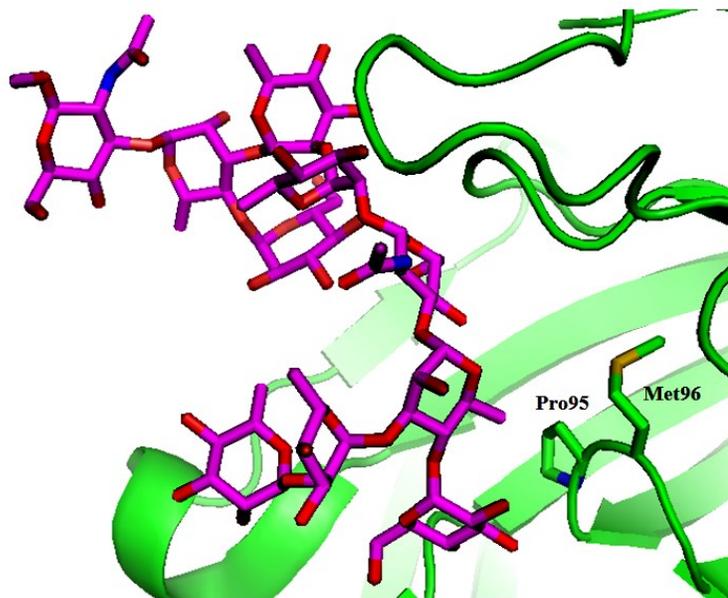


Figure 1.23 Pro95 and Met96 in *cis* peptide bond in F22-4 F_{ab} fragment in combination with *Shigella flexneri* deca-saccharide antigen (PDB code: 3BZ4)⁷⁷

Comparison of the two antibodies recognizing the O-antigen of two serotypes reveals that the BCD conformation in the two complexes is very similar, and CDR H3 was the most important domain for binding to the carbohydrate antigen in both antibodies. Interestingly, the major difference between these two antibodies arises from the CDR H3 domain size that is involved in epitope recognition. This domain is nine amino acid residues in SYA/J6, and four amino acid residues in F22-4. The short CDR H3 domain in F22-4 indeed creates a deep pocket to accommodate the branched Glu (E) in the AB(E)CD epitope.

1.7.4 Summary

The systems presented here are some of the most studied examples of pyranoside–antibody interactions, representing various aspects of carbohydrate recognition by antibodies. Other systems such as the recognition of *Brucella melitensis*^{88,89} or *Francisella tularensis*⁹⁰⁻⁹² cell wall oligosaccharides by different antibodies have revealed characteristics that were in agreement with the aspects revealed by the mentioned systems. Although in each system the antibody recognized a specific antigen with unique characteristics, there are some common aspects in all of the presented systems. In all cases the significant number of aromatic amino acid residues in

the binding site interacting with the carbohydrate antigen is noteworthy, and hydrophobic residues such as Trp are also essential for binding. Trp and His are the most frequently reported residues to be critical for binding in these systems. Small oligosaccharides tend to be accommodated by pocket- or cavity-type binding sites rich in hydrophobic amino acid residues, while longer oligosaccharides bind in groove-type binding sites. Induced fit is a characteristic of these interactions. Both the carbohydrate antigen and the antibody combining site might experience induced fit for optimum binding. The *S. flexneri* oligosaccharide (1-6) is an interesting example as the sugar adopts a compact helical conformation upon binding to be accommodated in the binding site.

1.8 The Interaction of CS-35 Antibody with Mycobacterial Arabinofuranoside Antigen

Although there is a substantial amount of research on the recognition of glycan antigens by antibodies, most of these investigations have been carried out using pyranoside ligands. Having higher energy than their six membered sugar counterparts,⁹³ furanose rings exist in a variety of twist and envelope conformations, separated by low energy barriers. A number of bacteria, fungi, and parasites biosynthesize furanoside-containing polysaccharides.^{94,95} Furanose-containing glycoconjugates are typically found on the surfaces of these organisms and interact with their environment and induce the production of antibodies.

The interactions of furanoses with the transporter proteins in microorganisms have been investigated. For instance, studies on bacterial ribofuranose⁹⁶ and bacterial galactofuranose⁹⁷ transporters have revealed the structural features that distinguish the furanose transporter proteins from pyranose transporter proteins. The biosyntheses of bacterial and fungal furanosides have also been investigated as examples of furanose-enzyme interactions.^{95, 98-100} Antibodies are also an important class of proteins that interact with bacterial carbohydrates. I reviewed a few examples of antibody–pyranose interaction systems. However, the information about such interactions between furanose moieties and antibodies, is limited. Although some interaction motifs are expected to be common between other proteins that interact with furanose moieties, this thesis focuses on the features that antibodies use to recognize their furanose antigen in their

binding sites. I will review the information on a furanose–antibody interaction system, which occurs between the cell wall oligosaccharides of mycobacteria and murine antibodies.

Arabinogalactan (AG) and lipoarabinomannan (LAM), which are found in the cell wall of mycobacterial species, are prominent examples of furanose-containing glycoconjugates.¹⁰¹ *Mycobacterium tuberculosis*, *M. leprae*, and *M. avium*, which are causes of tuberculosis, leprosy, and a tuberculosis-like disease common in HIV-positive individuals, respectively, are important representatives of furanose-containing bacteria. LAM has been implicated in the inhibition of T-cell function, the induction of cytokines¹⁰² and evocation of CD-1 restricted T-cell responses¹⁰³ and it is also a potent immunogen. LAM is composed of an arabinan domain, containing arabinofuranose residues attached to a mannopyranan core (Figure 1.24).¹⁰⁴

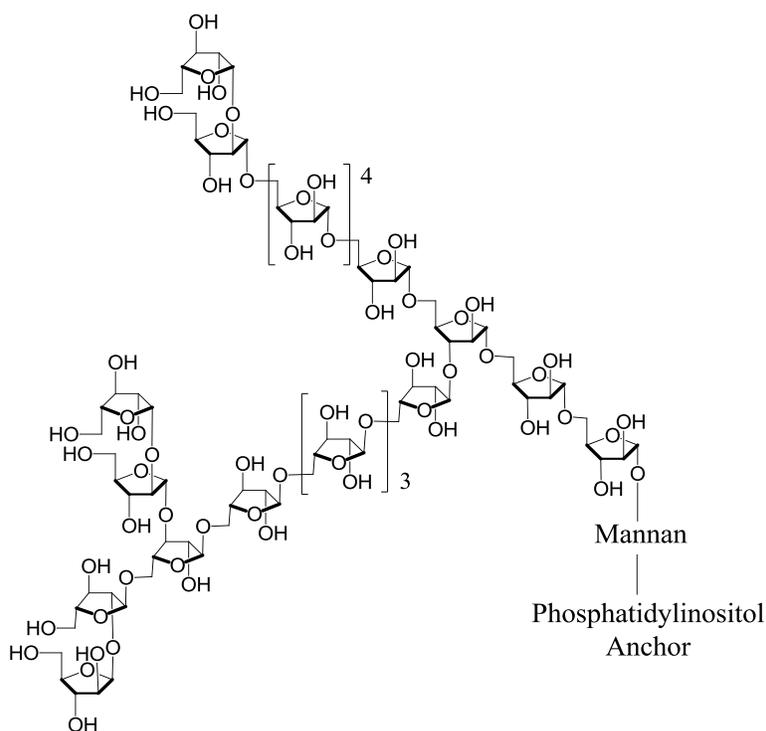


Figure 1.24 Composite structure of mycobacterial LAM

The CS-35 monoclonal antibody was raised against *M. leprae* LAM in 1986.¹⁰⁵ Since then, this antibody has been used as a reference antibody for the characterization of other mAbs against LAM and to investigate the role of LAM in the interaction of mycobacteria with the immune system.¹⁰⁶⁻¹¹¹ In 2002, it was shown that a terminal hexasaccharide moiety in LAM, Ara6 (**Figure 1.21**), is bound by the antibody.¹⁰³ This hexasaccharide (**Figure 1.25**) motif is found in both AG and LAM and its interaction with CS-35 represents a model system to study the interaction of furanosides with proteins.

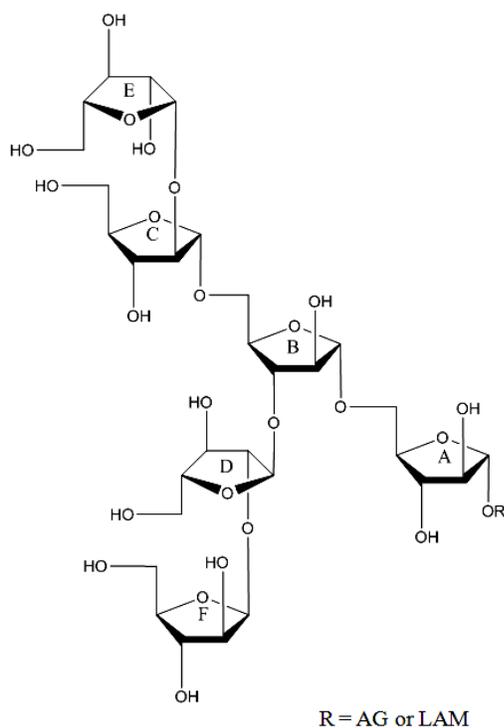


Figure 1.25 Ara6 antigen present in both mycobacterial AG and LAM.

The first study investigating the molecular details of the interaction of Ara6 with CS-35 was reported in 2007.⁹ A panel of substrates was synthesized and their binding to the CS-35 F_{ab} were studied by electrospray mass spectrometry (ES-MS) and STD NMR spectroscopy. It was shown that residues A, B, C, and E were important for binding, and the contribution of each residue of the hexasaccharide structure to the recognition was defined (**Figure 1.26**).

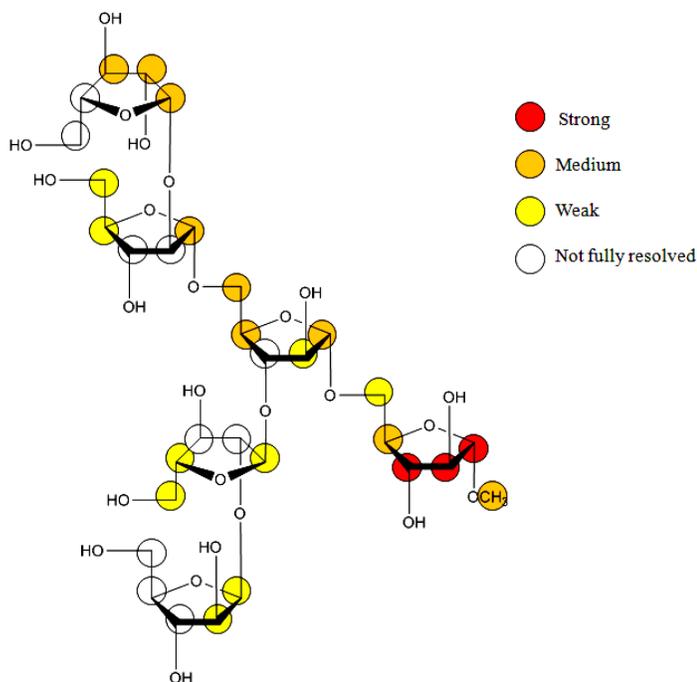


Figure 1.26 Regions of the Ara6 antigen recognized by CS-35 as determined by STD-NMR spectroscopy. The strongest saturation transfer effects are seen with residues A, B, C, and E.⁹

The high resolution crystal structure of CS-35 F_{ab} combined with Ara6 was reported two years later, and this provided valuable information on the Ara6–antibody combining site.¹⁰ The binding site was shown to be a broad groove at one end and a triangular canyon at the other, which was highly complementary to the Y-shaped structure of Ara6. CDRs L1, L2, and H3 domains formed most of the contact area (**Figure 1.27**). The binding epitope that was present in the crystal structure included rings A, B, C and E (**Figure 1.28**).

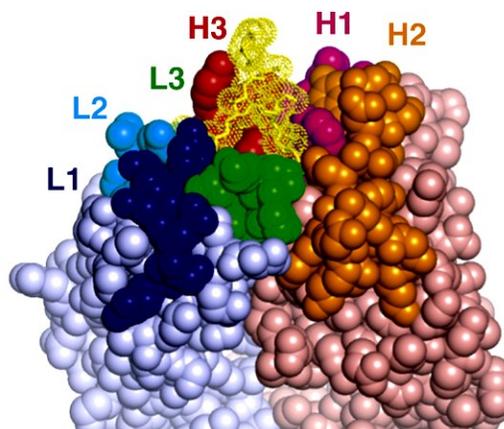


Figure 1.27 Main CDR domains of CS-35 F_{ab} in combination with Ara6.¹⁰

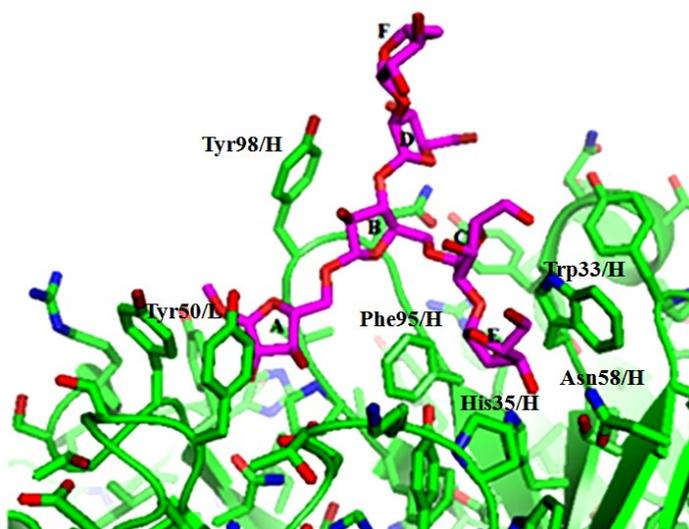


Figure 1.28 CS-35 F_{ab} in complex with Ara6 (PDB code: 3HNS)¹⁰ Some of the residues that are in close contact with Ara6 are highlighted in the structure.

This crystal structure revealed intermolecular interactions involving the Ara6 antigen including hydrogen bonds and water mediated hydrogen bonds as well as a substantial number of aromatic residues in close contact with the antigen. Trp33 and His35 positioning to Ara6 was very similar to the positioning of same residues to the trisaccharide antigen in Se155-4 antibody; these residues were previously reported to form half of the binding specificity of Se155-4. While

rings D and F extended out of the binding site in the crystal structure, rings A, B, C, and E were in close contact with the antibody. Ring B showed significant induced fit in the bound structure with an obvious deviation from its preferred solution conformation. The glycosidic linkages connecting ring A to B, and B to C also adopted extended conformations. However, the reason ring B adopted such an unusual conformation in the binding site could not be explained based on the structure. A noteworthy feature of the antibody structure was an unusual *cis* peptide bond between F95/H and G96/H (**Figure 1.29**).

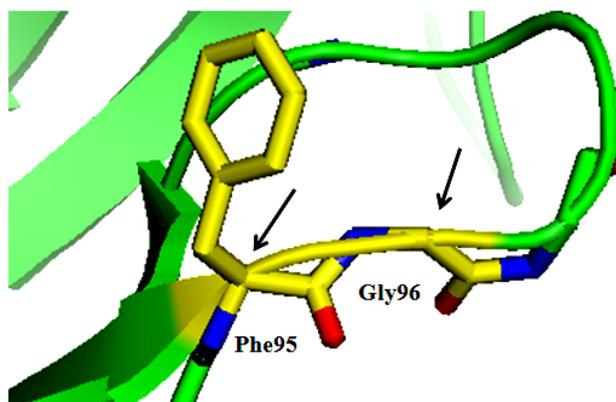


Figure 1.29 Phe95 and Gly96 of the CDR H3 in the structure of CS-35 F_{ab}; the *cis* peptide bond is indicated (PDB code: 3HNS)¹⁰ The two alpha carbon atoms are shown by arrows.

These unexplained structural findings, combined with the scarce amount of information about the nature of furanoside–protein interactions motivated us to study this binding interaction in more detail. This thesis is devoted to developing a more detailed picture of furanoside–antibody interactions, and to clarify some of the ambiguities of the recognition of Ara6 by CS-35. The results will contribute to the understanding of the principles of furanoside–antibody interactions and ultimately lead to a better comprehension of carbohydrate–protein interactions in general.

1.9 Thesis Research Overview

To probe structural and chemical elements in furanosides recognition by proteins, I manipulated the antibody–Ara6 interaction system by engineering three anti-LAM antibodies, CS-35, CS-40,¹⁰⁴ and 906.4321.¹⁰⁵ Our approach is centered on generating scFv fragments to study the binding profiles, and to manipulate the binding pocket by site-directed mutagenesis. A combination of single site mutation, SPR, STD NMR spectroscopy, circular dichroism, and homology modeling were applied to illuminate the recognition motifs from different aspects.

In Chapter 2 of this thesis, the generation and characterization of the CS-35 scFv fragment is discussed, and its binding to Ara6 is investigated. Also the design and generation of a library of scFv mutants and subsequent binding is presented as a means of probing the molecular details of the Ara6–CS-35 interaction.

In Chapter 3, the generation and characterization of scFv fragments of two other monoclonal antibodies that were known to bind to Ara6, CS-40 and 906.4321, is described. Experimental methods and computational approaches are used to investigate the binding of these two antibodies to Ara6.

Finally, Chapter 4 summarizes the contributions and the results of the present thesis and provides guidelines, ideas, and suggestions for the extension of this research in the field of carbohydrate–protein interactions.

References

- (1) Engering, A.; Geijtenbeek, T. B.; van Kooyk, Y. *Trends Immunol* **2002**, *23*, 480.
- (2) Hooper, L. V.; Gordon, J. I. *Glycobiology* **2001**, *11*, 1R.
- (3) Gabius, H.-J.; André, S.; Kaltner, H.; Siebert, H.-C. *Biochim. Biophys. Acta* **2002**, *1572*, 165.
- (4) Gabius, H.-J.; André, S.; Jiménez-Barbero, J.; Romero, A.; Solís, D. *Trends Biochem. Sci.* **2011**, *36*, 298.
- (5) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441.
- (6) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.

- (7) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387.
- (8) Woods, R. J.; Tessier, M. B. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575.
- (9) Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T. *J. Am. Chem. Soc.* **2007**, *129*, 10489.
- (10) Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *J. Mol. Biol.* **2009**, *392*, 381.
- (11) Ryder, M.; Tate, M.; Jones, G. *J. Biol. Chem.* **1984**, *259*, 9704.
- (12) Lemieux, R. U.; Spohr, U. *Adv. Carb. Chem. Biochem.* **1994**, *50*, 1.
- (13) Blake, C.; Koenig, D.; Mair, G.; North, A.; Phillips, D.; Sarma, V. *Nature* **1965**, *206*, 757.
- (14) Boraston, A.; Bolam, D.; Gilbert, H.; Davies, G. *J. Biochem.* **2004**, *382*, 769.
- (15) Nagae, M.; Yamaguchi, Y. *Int. J. Mol. Sci.* **2012**, *13*, 8398.
- (16) Ambrosi, M.; Cameron, N. R.; Davis, B. G. *Org. Biomol. Chem.* **2005**, *3*, 1593.
- (17) Janakiraman, M. N.; White, C. L.; Laver, W. G.; Air, G. M.; Luo, M. *Biochemistry* **1994**, *33*, 8172.
- (18) Goldstein, I.; Hollerman, C.; Smith, E. *Biochemistry* **1965**, *4*, 876.
- (19) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. **1999**.
- (20) Abbas, A. K.; Lichtman, A. H.; Pillai, S. *Cell. Mol. Immunol.*; Elsevier Health Sciences, 1994.
- (21) Ahmad, Z. A.; Yeap, S. K.; Ali, A. M.; Ho, W. Y.; Alitheen, N. B. M.; Hamid, M. *J. Immunol. Res.* **2012**, *2012*, 980250.
- (22) Alzari, P.; Lascombe, M.; Poljak, R. *Annu. Rev. Immunol.* **1988**, *6*, 555.
- (23) Davies, D. R.; Metzger, H. *Annu. Rev. Immunol.* **1983**, *1*, 87.
- (24) Capra, J. D.; Edmundson, A. B. *Sci. Am.* **1977**, *236*, 50.
- (25) Te Wu, T.; Kabat, E. A. *J. Exp. Med.* **1970**, *132*, 211.
- (26) Mian, I. S.; Bradwell, A. R.; Olson, A. J. *J. Mol. Biol.* **1991**, *217*, 133.
- (27) Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland New York, 1991; Vol. 2.

- (28) Klimka, A.; Matthey, B.; Roovers, R.; Barth, S.; Arends, J.; Engert, A.; Hoogenboom, H. *Br. J. Cancer* **2000**, *83*, 252.
- (29) Watkins, N. A.; Ouwehand, W. H. *Vox Sang* **2000**, *78*, 72.
- (30) Holliger, P.; Hudson, P. J. *Nature Biotechnol.* **2005**, *23*, 1126.
- (31) Boss, M. A.; Kenten, J. H.; Wood, C. R.; Emtage, J. S. *Nucleic Acids Res.* **1984**, *12*, 3791.
- (32) Kontermann, R. E.; Müller, R. *J. Immunol. Methods* **1999**, *226*, 179.
- (33) Colcher, D.; Pavlinkova, G.; Beresford, G.; Booth, B.; Choudhury, A.; Batra, S. *Quar. J. Nuc. Med. (AIMN) (IAR)* **1998**, *42*, 225.
- (34) Hudson, P. J.; Kortt, A. A. *J. Immunol. Methods* **1999**, *231*, 177.
- (35) Huston, J. S.; Levinson, D.; Mudgett-Hunter, M.; Tai, M.-S.; Novotný, J.; Margolies, M. N.; Ridge, R. J.; Bruccoleri, R. E.; Haber, E.; Crea, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5879.
- (36) Huston, J. S.; Mudgett-Hunter, M.; Tai, M.-S.; McCartney, J.; Warren, F.; Haber, E.; Oppermann, H. *Methods Enzymol.* **1991**.
- (37) Whitlow, M.; Bell, B. A.; Feng, S.-L.; Filpula, D.; Hardman, K. D.; Hubert, S. L.; Rollence, M. L.; Wood, J. F.; Schott, M. E.; Milenic, D. E. *Protein Eng.* **1993**, *6*, 989.
- (38) Maeda, H.; Schmidt-Kessen, A.; Engel, J.; Jaton, J.C. *Biochemistry* **1997**, *16*, 4086.
- (39) Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373.
- (40) Carver, J. *Pure Appl. Chem.* **1993**, *65*, 763.
- (41) Lindh, I.; Hindsgaul, O. *J. Am. Chem. Soc.* **1991**, *113*, 216.
- (42) Haselkorn, D.; Friedman, S.; Givol, D.; Pecht, I. *Biochemistry* **1974**, *13*, 2210.
- (43) Ross, P. D.; Subramanian, S. *Biochemistry* **1981**, *20*, 3096.
- (44) Wedemayer, G. J.; Patten, P. A.; Wang, L. H.; Schultz, P. G.; Stevens, R. C. *Science* **1997**, *276*, 1665.
- (45) Koshland, D. E. *Angew. Chem. Int. Ed.* **1995**, *33*, 2375.
- (46) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond: in Structural Chemistry and Biology*; Oxford University Press, 2001; Vol. 9.
- (47) Quioco, F. A. *Annu. Rev. Biochem.* **1986**, *55*, 287.

- (48) Vyas, N. K. *Curr. Opin. Struct. Biol.* **1991**, *1*, 732.
- (49) Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.; Park, S. J.; Skehel, J. J.; Wiley, D. C. *Biochemistry* **1992**, *31*, 9609.
- (50) Watowich, S. J.; Skehel, J. J.; Wiley, D. C. *Structure* **1994**, *2*, 719.
- (51) Merritt, E. A.; Sarfaty, S.; Akker, F. V. D.; L'Hoir, C.; Martial, J. A.; Hol, W. G. *Protein Sci.* **1994**, *3*, 166.
- (52) Stehle, T.; Yan, Y.; Benjamin, T. L.; Harrison, S. C. *Nature* **1994**, *369*, 160.
- (53) Hayashida, M.; Fujii, T.; Hamasu, M.; Ishiguro, M.; Hata, Y. *J. Mol. Biol.* **2003**, *334*, 551.
- (54) Nishio, M. *Phys. Chem.* **2011**, *13*, 13873.
- (55) Tamres, M. *J. Am. Chem. Soc.* **1952**, *74*, 3375.
- (56) Quioco, F. A.; Vyas, N. K. *Nature* **1984**, *310*, 381.
- (57) Harata, K.; Muraki, M. *Acta Crystallogr. Section D: Biological Crystallography* **1997**, *53*, 650.
- (58) Muraki, M. *Pro. Pept. Let.* **2002**, *9*, 195.
- (59) Payne, C. M.; Bomble, Y. J.; Taylor, C. B.; McCabe, C.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Biol. Chem.* **2011**, *286*, 41028.
- (60) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. S. *Acc. Chem. Res.* **2012**, *46*, 946.
- (61) Dunitz, J.D. *Science*, **1994**, *264*, 670
- (62) Toone, E. J. *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.
- (63) Watson, K. A.; Mitchell, E. P.; Johnson, L. N.; Bichard, C. J.; Orchard, M. G.; Fleet, G. W.; Oikonomakos, N. G.; Leonidas, D.; Son, J. C. *Biochemistry* **1994**, *33*, 5745.
- (64) Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973.
- (65) Coburn, B.; Grassl, G. A.; Finlay, B. *Immunol. Cell Biol.* **2007**, *85*, 112.
- (66) Sigurskjold, B.; Bundle, D. *J. Biol. Chem.* **1992**, *267*, 8371.
- (67) Cygler, M.; Rose, D. R.; Bundle, D. R. *Science* **1991**, *253*, 442.
- (68) Bundle, D. R.; Baumann, H.; Brisson, J.-R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183.

- (69) Bundle, D. R.; Alibés, R.; Nilar, S.; Otter, A.; Warwas, M.; Zhang, P. *J. Am. Chem. Soc.* **1998**, *120*, 5317.
- (70) SIGURSKJOLD, B. W.; Altman, E.; BUNDLE, D. R. *Eur. J. Biochem.* **1991**, *197*, 239.
- (71) Bundle, D. R.; Eichler, E.; Gidney, M. A. J.; Meldal, M.; Ragauskas, A.; Sigurskjold, B. W.; Sinnott, B.; Watson, D. C.; Yaguchi, M.; Young, N. M. *Biochemistry* **1994**, *33*, 5172.
- (72) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423.
- (73) MacKenzie, C. R.; Hiram, T.; Deng, S.-j.; Bundle, D. R.; Narang, S. A.; Young, N. M. *J. Biol. Chem.* **1996**, *271*, 1527.
- (74) Kitova, E. N.; Wang, W.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 13980.
- (75) Deng, S.; MacKenzie, C. R.; Sadowska, J.; Michniewicz, J.; Young, N. M.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1994**, *269*, 9533.
- (76) Reidl, J.; Klose, K. E. *FEMS Microbiol. Rev.* **2002**, *26*, 125.
- (77) Villeneuve, S.; Souchon, H.; Riottot, M.-M.; Mazié, J.-C.; Lei, P.-S.; Glaudemans, C.; Kováč, P.; Fournier, J.-M.; Alzari, P. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8433.
- (78) Wang, J.; Villeneuve, S.; Zhang, J.; Lei, P.-s.; Miller, C. E.; Lafaye, P.; Nato, F.; Szu, S. C.; Karpas, A.; Bystricky, S. *J. Biol. Chem.* **1998**, *273*, 2777.
- (79) Vulliez-Le Normand, B.; Saul, F.; Phalipon, A.; Belot, F.; Guerreiro, C.; Mulard, L. A.; Bentley, G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9976.
- (80) Jennison, A. V.; Verma, N. K. *FEMS Microbiol. Rev.* **2004**, *28*, 43.
- (81) Vyas, N. K.; Vyas, M. N.; Chervenak, M. C.; Johnson, M. A.; Pinto, B. M.; Bundle, D. R.; Quioco, F. A. *Biochemistry* **2002**, *41*, 13575.
- (82) Bundle, D.; Oxford University Press: New York: 1999, p 370.
- (83) Bock, K.; Josephson, S.; Bundle, D. R. *J. Chem. Soc. Perkin Transaction 2* **1982**, 59.
- (84) Bélot, F.; Wright, K.; Costachel, C.; Phalipon, A.; Mulard, L. A. *J. Org. Chem.* **2004**, *69*, 1060.
- (85) Wright, K.; Guerreiro, C.; Laurent, I.; Baleux, F.; Mulard, L. A. *Org. Biomol. Chem.* **2004**, *2*, 1518.
- (86) Bélot, F.; Guerreiro, C.; Baleux, F.; Mulard, L. A. *Chem. Eur. J.* **2005**, *11*, 1625.

- (87) Theillet, F.-X.; Frank, M.; Vulliez-Le Normand, B.; Simenel, C.; Hoos, S.; Chaffotte, A.; B elot, F.; Guerreiro, C.; Nato, F.; Phalipon, A. *Glycobiology* **2011**, *21*, 1570.
- (88) Rose, D.; Przybylska, M.; To, R.; Kayden, C.; Vorberg, E.; Young, N.; Bundle, D.; Oomen, R. *Prot. Sci.* **1993**, *2*, 1106.
- (89) Young, N.; Gidney, M.; Gudmundsson, B.; MacKenzie, C.; To, R.; Watson, D.; Bundle, D. *Mol. Immunol.* **1999**, *36*, 339.
- (90) Wang, Q.; Shi, X.; Leymarie, N.; Madico, G.; Sharon, J.; Costello, C. E.; Zaia, J. *Biochemistry* **2011**, *50*, 10941.
- (91) Lu, Z.; Madico, G.; Roche, M. I.; Wang, Q.; Hui, J. H.; Perkins, H. M.; Zaia, J.; Costello, C. E.; Sharon, J. *Immunology* **2012**, *136*, 352.
- (92) Lu, Z.; Rynkiewicz, M. J.; Yang, C. Y.; Madico, G.; Perkins, H. M.; Wang, Q.; Costello, C. E.; Zaia, J.; Seaton, B. A.; Sharon, J. *Immunology* **2013**, *140*, 374.
- (93) Angyal, S. J. *Adv. Carb. Chem. Biochem.* **1984**, *42*, 15.
- (94) Lowary, T. L. *Curr. Opin. Chem. Biol.* **2003**, *7*, 749.
- (95) Houseknecht, J. B.; Lowary, T. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 677.
- (96) Bagaria, A.; Bagaria, A.; *Proteins Struct. Funct. Bioinfo.* **2011**, *79*, 1352.
- (97) Horler, R.; M uller, A.; David C. *J. Biol. Chem.* **2009**, *284*, 31156.
- (98) Miyanaga, A.; Koseki, T.; Matsuzawa, H.; Wakagi, T.; Shoun, H.; Fushinobu, S. *J. Biol. Chem.* **2004**, *279*, 44907.
- (99) Rose, N. L.; Completo, G. C.; Shuang-Jun, L.; McNeil, M. R.; Placic, M.; Lowary, T. L. *J. Am. Chem. Soc.* **2006**, *128*, 6721.
- (100) Nassau, P.M.; Martin, S. L.; Brown, R. E.; Weston, A.; Monsey, D.; McNeil, M. R.; Duncan, K. *J. Bacteriol.* **1996**, *178*, 1047.
- (101) Brennan, P. J.; Nikaido, H. *Annu. Rev. Biochem.* **1995**, *64*, 29.
- (102) Yoshida, A.; Koide, Y. *Infec. Immun.* **1997**, *65*, 1953.
- (103) Sieling, P.; Chatterjee, D.; Porcelli, S.; Prigozy, T.; Mazzaccaro, R.; Soriano, T.; Bloom, B.; Brenner, M.; Kronenberg, M.; Brennan, P. *Science* **1995**, *269*, 227.
- (104) Briken, V.; Porcelli, S. A.; Besra, G. S.; Kremer, L. *Mol. Microbiol.* **2004**, *53*, 391.
- (105) Hunter, S. W.; Gaylord, H.; Brennan, P. *J. Biol. Chem.* **1986**, *261*, 12345.
- (106) Hamasur, B.; K allenius, G.; Svenson, S. B. *Vaccine* **1999**, *17*, 2853.

- (107) Means, T. K.; Lien, E.; Yoshimura, A.; Wang, S.; Golenbock, D. T.; Fenton, M. J. *J Immunol.* **1999**, *163*, 6748.
- (108) Bottger, E.; Teske, A.; Kirschner, P.; Bost, S.; Hirschel, B.; Chang, H.; Beer, V. *The Lancet* **1992**, *340*, 76.
- (109) Kaur, D.; Lowary, T. L.; Vissa, V. D.; Crick, D. C.; Brennan, P. J. *Microbiology* **2002**, *148*, 3049.
- (110) Navoa, J. A. D.; Laal, S.; Pirofski, L.-A.; McLean, G. R.; Dai, Z.; Robbins, J. B.; Schneerson, R.; Casadevall, A.; Glatman-Freedman, A. *Clin. Diag. Lab. Immunol.* **2003**, *10*, 88.
- (111) Gaylord, H.; Brennan, P.; Young, D.; Buchanan, T. *Infec. Immun.* **1987**, *55*, 2860.

Chapter 2 : Specificity of the CS-35 Antibody to the Ara6 Antigen

2.1 Introduction

Mycobacteria are the cause of important diseases such as tuberculosis (TB) and leprosy. The obstacles¹ in the treatment of TB, especially in immuno-compromised patients, and the emergence of drug resistant TB cases have necessitated research to seek new mechanisms for controlling TB. An attractive target in mycobacteria is the unique structure of the cell wall, which is recognized by antibodies. Ara6, as I discussed in Chapter 1, is an antigenic part of the cell wall, which binds to the CS-35 antibody.² Moreover, this system provides the opportunity to study the molecular basis of furanoside recognition by proteins. Furanosides, the five membered ring sugars are interesting chemical structures, with several polar functional groups and flexible skeleton, whose recognition by proteins is not as studied as their six membered ring counterparts (pyranosides). In Chapter 1, we highlighted the value of the CS-35–Ara6 interaction system as a model to study the recognition of furanosides by proteins.

The CS-35 F_{ab} had been previously generated by the enzymatic digestion of the CS-35 mAb.² This fragment was used in the initial binding studies with Ara6 (**2-2**, **Figure 2.1**) and its derivatives.² By using the electrospray mass (ES-MS) technique, the affinity of the CS-35 F_{ab} to Ara6 (**2-2**) was shown to have a K_D of 6.2×10^{-6} M. A characteristic of carbohydrate–protein interactions is the low affinity of binding, which is usually in the millimolar range.³ It is interesting that Ara6 carbohydrate is recognized by the antibody with such high affinity. By using saturation transfer difference nuclear magnetic resonance (STD NMR) experiments, the binding epitope of the antigen to the CS-35 F_{ab} was shown to contain mainly rings A, B, C, and E.² This was further corroborated by the X-ray crystal structure of F_{ab} in combination with **2-2**.⁴ The crystal structure revealed some direct hydrogen bonds, and water mediated hydrogen bonds, and a number of aromatic residues that encompassed the antigen in close proximity to the furanoside rings. However, the forces involved in this interaction and the specificity of the

recognition remained enigmatic. To define the molecular motifs that govern this interaction, and to discover the functional groups that tailor the specificity of this recognition, we were interested in manipulating the binding site by mutations and evaluating the binding behavior upon changing the amino acid residues. This would lead to defining the contribution of specific residues and intermolecular interactions to the binding. However, mAbs and the F_{ab} are not suitable for this purpose, for the reasons outlined in Chapter 1. Therefore, we considered applying single chain variable fragment (scFv) technology as a tool to manipulate the CS-35 binding site by point mutations. Use of a scFv has proved to be useful in studying some aspects of the interaction of *Salmonella* serogroup B O-antigen with Se155-4 antibody,⁵ which is a bacterial pyranoside–antibody interaction system.⁵ We generated CS-35 scFv and a panel of its mutants, each representing one amino acid change. This is the first scFv reported to recognize a mycobacterial glycan, and the first scFv that recognizes a bacterial furanoside antigen. A systematic evaluation of the binding details was further achieved by different methods, which will be discussed in the following sections.

This chapter outlines our efforts to define the specificity motifs in CS-35–Ara6 interaction by developing a stable single chain variable fragment (scFv) of CS-35, and its mutants. By using a combination of different methods such as surface plasmon resonance (SPR) spectroscopy, STD NMR spectroscopy, and circular dichroism (CD) we have developed a more detailed picture of the CS-35 antibody binding to Ara6 at the molecular level. The Ara6 antigens used in this chapter were provided by Dr. Maju Joe in the Department of Chemistry at the University of Alberta.

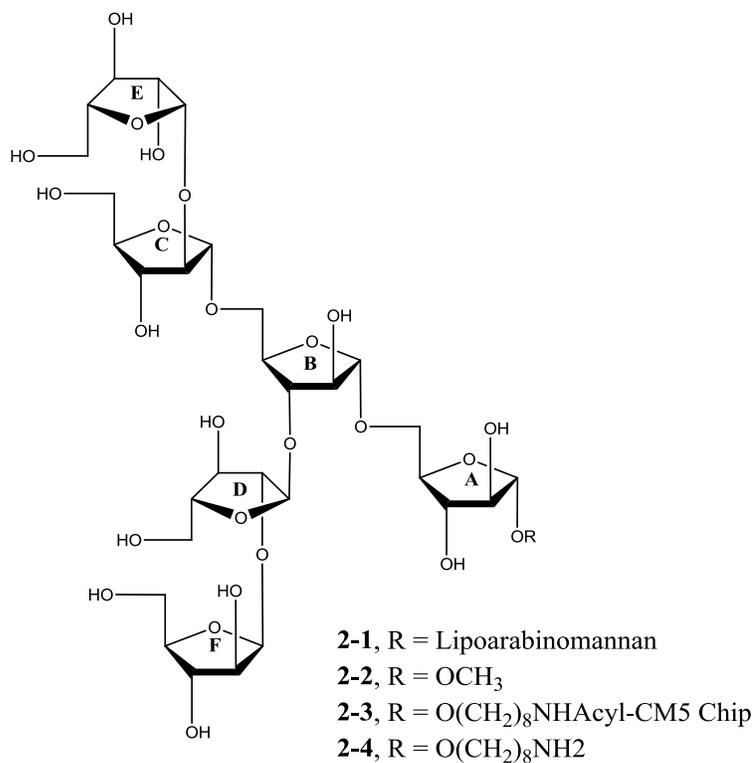


Figure 2.1 Structure of Ara6 in LAM (**2-1**), and binding studies (**2-2**, **2-3**, and **2-4**)

2.2 Materials and Methods

Synthetic DNA oligonucleotides were purchased from Genscript. PCR products and products of restriction digestion were purified by gel electrophoresis and extracted using the GeneJET gel extraction kit (Fermentas, ON). Plasmid DNA was purified using the GeneJET Plasmid Miniprep Kit (Fermentas, ON). Restriction endonucleases were purchased from Invitrogen. Dye terminator cycle sequencing using the BigDye (Applied Biosystems) was used to confirm the complete cDNA sequences. Sequencing reactions were analyzed at the University of Alberta Molecular Biology Service Unit (MBSU). QuikChange II Site-Directed Mutagenesis Kit (Agilent technologies) was used for point mutations. The CS-35 images are derived from the crystal structure PDB code: 3HNS.⁴

2.2.1 Construction of CS-35 scFv Genes

The scFv of the CS-35 antibody was constructed based on the corresponding F_{ab} sequence (PDB: 3HNS) as the template, with V_H-Linker-V_L orientation by splice overlap PCR. The V_H and V_L domains had been previously cloned and sequenced by reverse transcriptase PCR.⁴ The linker separating variable domains was the flexible peptide (Gly₄Ser)₃. The scFv product of the PCR was sequenced and shown to have a mutation at the Ser31 residue in the V_L to Ala. The Ser31^LAla mutant gene was stored for further expression, and was also used as a template to generate the wild type sequence of the scFv by single-site mutation. Ala31 was mutated to Ser, and the correct sequence of the wild type scFv product was confirmed by sequencing. The gene was harbored in pET-28a plasmid between BamHI and HindIII restriction sites for expression purposes, and was also used as the template to create mutants. The gene sequence of the scFv that was inserted in the vector is provided in the appendix. The sequence of the scFv was:

```
EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKQRPGQGLEWIGSIYPGNSD  
TNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFGNYVPFAYWGQGLV  
TVGGGGSGGGGSGGGGSDIQMTQTTSSLSASLGDRVTIGCRASQDIGSYLNWYQKPDG  
AVRLLIYYTSRLHSGVPSRFSGSGSGTHFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTK  
LEIK
```

A scFv gene designed to be harbored in pET-30b plasmid between NdeI and HindIII, and was ordered later through Genscript for some comparison studies of the scFv conformation. I refer to this fragment as scFv for convenience in the following sections. A His tag was added to the sequence at the C terminus for purification purposes.

2.2.2 Transformation, Expression, and Refolding of scFv Fragments

The pET-28a plasmid (200 ng) harboring the wild type scFv gene was transformed into *Escherichia coli* BL21(DE3) host cells by heat shock. Cells were cultured in 50 mL of LB medium, and were then shaken at 37 °C for 8 h. To a 400 mL of LB medium, 10 mL of this culture was transferred and the mixture was shaken in 37 °C for 2 h. This amount of time was sufficient to reach OD 0.8–1, which was optimum for developing inclusion bodies in the cytoplasm of *E. coli* cells. Expression was then induced by adding 1M isopropyl β-D-1-

thiogalactopyranoside and the culture was shaken at 20 °C overnight. The next day, cells were pelleted at 8000 rpm for 30 min, and the pellet was dissolved in a buffer containing 50 mM Tris-HCl, pH 8, 1% Triton X-100, and 150 mM NaCl, followed by cell disruption by French press. Inclusion bodies were pelleted and washed six times successively with a buffer containing 50 mM Tris-HCl, pH 8, and 200 mM NaCl. Inclusion bodies were then dissolved and denatured by shaking at room temperature in a buffer containing 4 mM guanidine and 100 mM Tris-HCl, pH 8 for 2 h. The solution was further injected into a buffer containing 0.5 M arginine and 100 mM Tris-HCl pH 8, and was incubated at 4 °C for 48–72 h. Denatured scFvs were then dialyzed twice against the refolding buffer containing 50 mM Tris-HCl, pH 8, and 300 mM NaCl each time for 10 h.

2.2.3 Protein Purification

The folded scFv fragments were purified by affinity chromatography using Ni-NTA Superflow Qiagen resin. SDS PAGE was used to check the size and purity of fragments. The wild type scFv was also submitted to the mass spectrometry facility at the Department of Chemistry to obtain an ESI-MS spectrum. The scFv fragments were analyzed for monomer, dimer, and higher oligomer content by size-exclusion fast protein liquid chromatography (FPLC) using a Superdex 75 10/30 (GE Healthcare Life Sciences) column. Protein samples were loaded in 1mg/mL concentration, and were eluted in ammonium acetate buffer (50 mM, pH 6.8) with a 0.5mL/min rate.

2.2.4 Generation of the CS-35 ScFv Mutants Library

A library of 12 scFv mutants was generated based on the crystal structure of the CS-35 F_{ab} in combination with **2** for possible contribution of specific functional groups to the binding. The scFv was used as the template for the single site mutations. CS-35 scFv mutants were expressed as inclusion bodies, refolded, and purified under the same conditions as those used for the wild type scFv. SDS PAGE confirmed the size (29 kDa) and the purity of each mutant. The monomer fractions of mutants that could bind to the antigen in the screening were collected using size

exclusion chromatography for the single binding site analysis, as described above for the wild type scFv. Pure monomer fractions of all mutants were stored in 4 °C for further studies.

2.2.5 Circular Dichroism (CD) Spectroscopic Study of the CS-35 Antibody Fragments

CD spectra were obtained in phosphate buffer (50mM, pH 7.2) for wild type scFv, using an Olis DSM 17s spectrometer at room temperature. Samples were run at the Analytical and Instrumentation Lab of Chemistry Department by Dr. Wayne Moffat. The far UV spectra were obtained for 27 μ M of the CS-35 scFv with and without **2-4**. An aliquot containing 200 ng of **2-4** in the same buffer condition was added to the wild type scFv sample, and the far UV spectrum of the complex was recorded. The far UV spectra of the CS-35 mAb and F_{ab} fragment were also collected at 10 μ M and 15 μ M concentrations of the protein samples, respectively, under the same conditions, for comparison. The near UV spectrum was obtained for 27 μ M of the CS-35 scFv in phosphate buffer (50 mM, pH 7.2) with and without **2-4**.

2.2.6 Confirmation of the Presence of Disulfide Bonds in the scFv

A scFv sample containing 200 ng of protein was dissolved in 200 μ L of 50 mM ammonium bicarbonate buffer pH 8. The sample was divided in half. To one half, 10 μ L of 100 mM dithiothreitol (DTT) was added, and the sample was incubated in 50 °C for 30 min. Then to the both samples 10 μ L of 200 mM iodoacetamide (IAA) was added. Both samples were incubated for 20 min, and were then submitted to the Mass Spectrometry Facility in the Department of Chemistry and ES-MS analysis was carried out.

2.2.7 Binding of the CS-35 Antibody Fragments to Ara6 by Surface Plasmon Resonance (SPR) Spectroscopy

Binding affinities and kinetics were investigated by SPR using a BIAcore 3000 (GE Healthcare Life Sciences) instrument at 25 °C. Ligand **2-4** (5 μ M) was covalently immobilized on the surface of the carboxymethylated dextran sensor chip (CM5) chip by amine coupling. A fresh solution of 0.5 M NHS and 10 mM EDC was first injected on the chip surface to activate

the carboxylate groups. Ligand **2-4** was then passed over the flow channel in a buffer containing 10 mM HEPES and 150 mM NaCl, pH 7.2 to generate the surface bound ligand **2-3**. A reference flow channel was prepared by treatment with a solution of 1M aqueous ethanolamine pH 8.5 after the activation of the surface with NHS/EDC. Freshly prepared scFv samples and mutants in concentrations ranging from nanomolar to micromolar (0.26, 0.53, 1.6, 2.6, 4.8 μ M) (extinction coefficient: 51590) were injected to the flow channels in buffer with 10 mM HEPES, pH 7.2, and 150 mM, NaCl, with the rate of 20 μ L/min for 300 seconds. The F_{ab} fragment was also tested and analyzed as a reference under the same conditions. Proteins showing zero response units after the subtraction of the two flow channels were designated as non-binders.

The equilibrium binding of the wild-type scFv, mutants, and F_{ab} fragment were also determined by Scatchard plots using the following equation:

$$\frac{R_{eq}}{C} = K_A R_{max} - K_A R_{eq} \quad (2.1)$$

Where R_{eq} is the equilibrium response units, R_{max} is the resonance signal at saturation, C is the concentration of the protein, and K_A is the association constant. A plot of R_{eq}/C versus R_{eq} has a slope of $-K_A$. Therefore, to obtain K_A and R_{max} from the data, one can do a first-degree (i.e. linear) polynomial regression to these variables (R_{eq}/C versus R_{eq}). We performed this using the “polyfit” command in MATLAB. The K_D value is the reciprocal of the K_A value.

2.2.8 Studies of the CS-35 scFv by ESI-MS

I prepared monomer fragments of scFv and those mutants that retained binding to **2-4** for mass spectrometric studies, in collaboration with Drs. John Klassen and Aneika Leney in the Department of Chemistry. The samples were sprayed in nano ESI in a SYNAPT G2S (Waters 2011) instrument in 50 mM ammonium acetate pH 6.8.

2.2.9 Binding of the CS-35 Fragments and Mutants to 2-4 by Saturation Transfer Difference NMR Spectroscopy

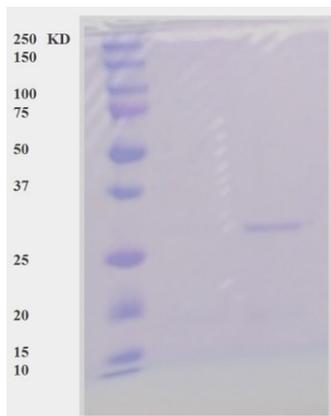
NMR experiments for scFv mutants were recorded on a Varian Inova 600 MHz spectrometer. Lyophilized scFv mutant samples were dissolved in deuterated ammonium acetate (50 mM, pH 6.8), with concentrations in the range of 27–30 μ M. An aliquot of ligand **2-4** was

added to each sample with the final concentration of 0.5 mM. Measurements were done in 3 mm NMR tubes. 1D STD experiments⁶ were performed at 298 K with saturation at 0.5 ppm, 2 ms delay, and with water suppression for mutants. Wild-type scFv and F_{ab} fragment samples were buffer exchanged into the deuterated phosphate buffer (50 mM, pH 7.2). The STD experiments for these two samples were performed at 27 °C on a Varian Inova 700 MHz instrument without water suppression. Saturation at 0.5 ppm with a 3.8 ms delay was applied. A 1D ¹H NMR spectrum of each solution containing protein and ligand **2-4**, as well as STD spectra of scFv and F_{ab} with ligand **2-4**, were recorded.

2.3 Results

2.3.1 Generation of the CS-35 scFv, Structure

The scFv was produced with the yield of 2 mg/L, after expression in inclusion bodies and refolding the denatured proteins. Different buffer and pH conditions were tested in various steps for the generation and stability of the fragments. This method led to high purity of the folded scFv, even before purification by the Ni-NTA column. The scFv fragments are, in fact, protected in the inclusion bodies during six washing steps, which helps removal of impurities. The refolding condition also was optimized for this protein, which could assist the efficacy. The size (29 kDa) and purity of the generated scFv was confirmed by SDS PAGE and ESI mass spectroscopy (**Figure 2.2**).



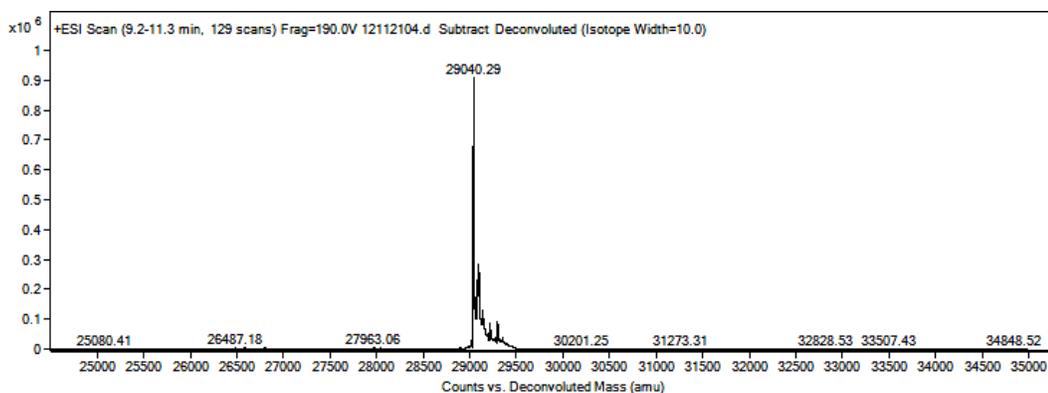


Figure 2.2 SDS PAGE (29 KDa) and ESI mass of wild type scFv

The scFv sample was injected into the size exclusion column, and it was shown to contain monomer scFv peak at 18 mL and dimer scFv peak at 14 mL (**Figure 2.3**). The monomer fragment was collected to study the binding kinetics of a single antigen binding site. The dimer fragment was also collected for comparison. Different buffer and pH conditions were tested for the ratio of monomer to dimer. For the CS-35 scFv, 50 mM ammonium acetate buffer, pH 6.8 was shown to be optimum for the ratio of monomer to dimer. Also, the monomer was stable in this buffer condition at 4 °C for three weeks. Moreover, the more dilute the sample loaded on the size exclusion column, the larger the amount of the monomer scFv was collected. This result was consistent among all the mutants.

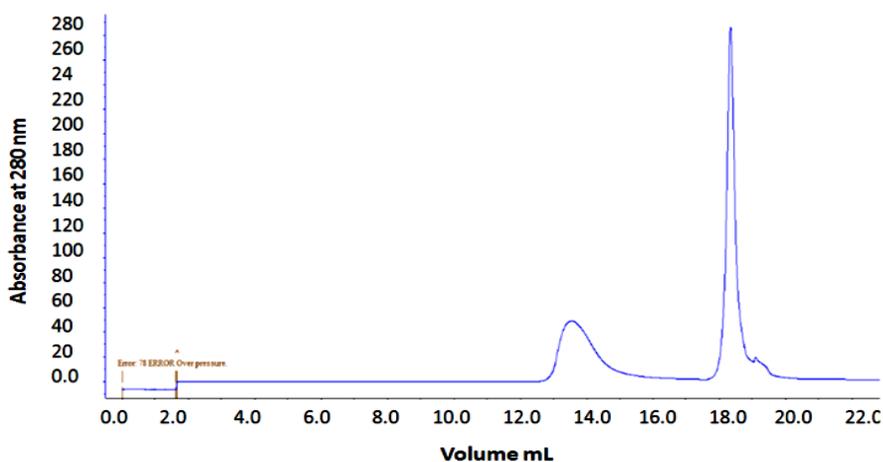


Figure 2.3 FPLC results of the CS-35 scFv. The peak at 18 mL belongs to the monomer; the dimer elutes at 14 mL

CS-35 scFv was generated as inclusion bodies in *E. coli* cytoplasm. The key step in this process is the proper folding of the denatured inclusion bodies. The inclusion bodies were first denatured and then induced to refold in buffer. Therefore, it was important to check the proper folding of this recombinant antibody fragment before using it for binding analysis. The secondary structure of antibodies is composed of β sheets. In the circular dichroism (CD) technique, these β sheets of antibodies show a typical absorption profile in the far UV region. This structure is distinguished by a positive band at 195 nm and a negative band at 218 nm.⁷ If there is misfolding of the protein structure, it will be reflected in the CD spectrum.

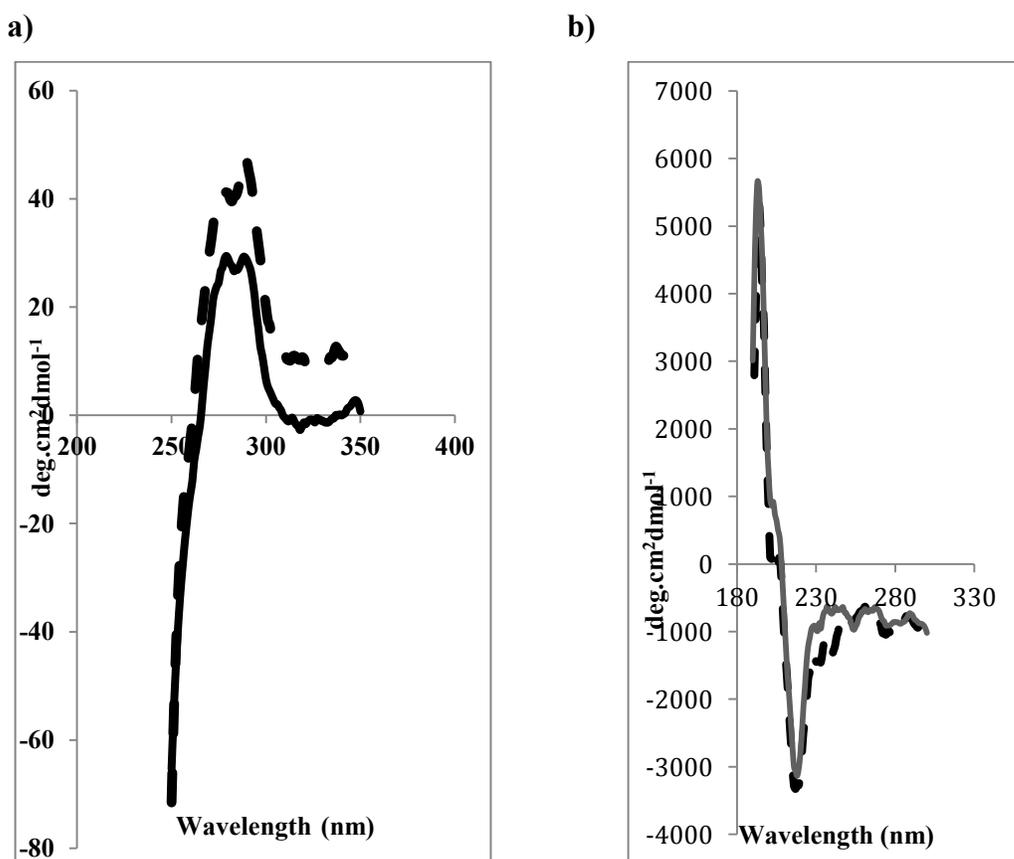


Figure 2.4 The far UV (top) CD spectrum of the wild-type CS-35 scFv before (solid) and after (dashed) adding **2-4** (a), The near UV CD spectrum of the wild-type CS-35 scFv before (solid) and after (dashed) adding **2-4** (b).

The far UV CD spectra of the scFv (**Figure 2.4**), the F_{ab} fragment, and the mAb (**Figure 2.5**) showed a positive band at 195 nm and a negative band at 218 nm, typical of the

secondary structures of antibodies, suggesting that the scFv was properly folded. The far UV spectra of the scFv in complex with Ara6 **2-4** and that of the scFv alone (**Figure 2.4**) were almost identical, implying that binding to **2-4** did not alter the secondary structure of the protein.

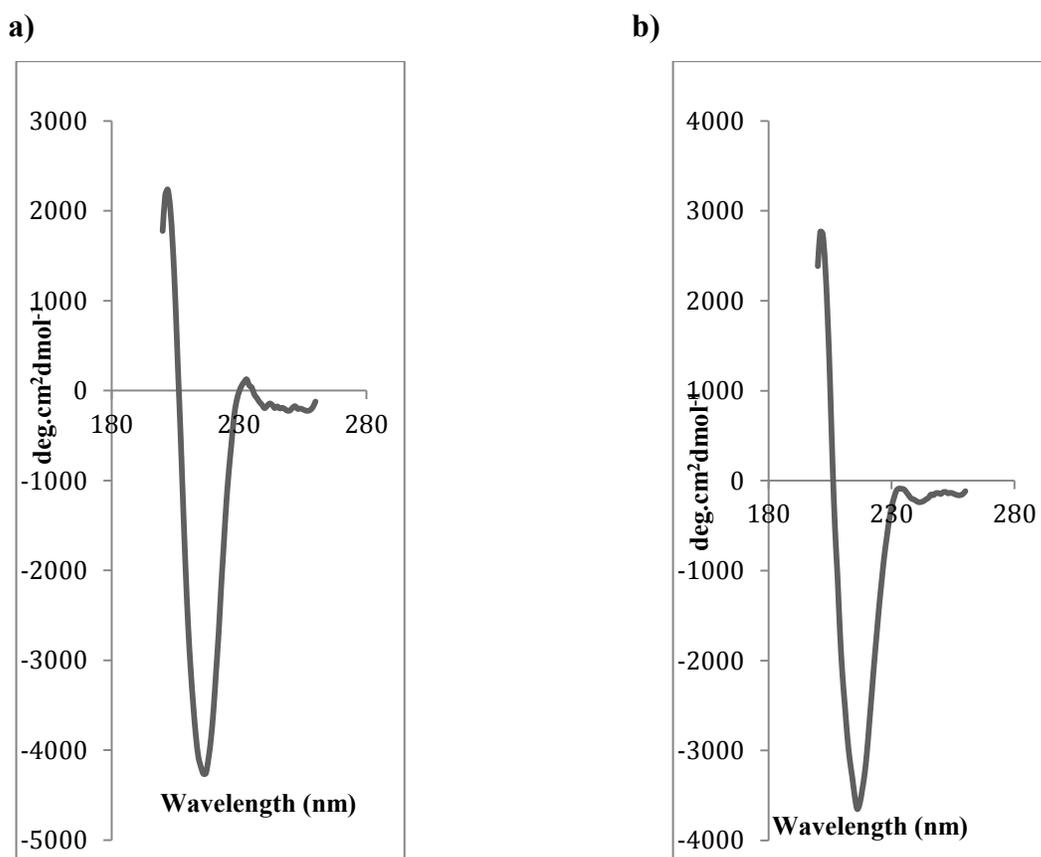


Figure 2.5 CD spectroscopic results in far UV region for CS-35 F_{ab} (a) and mAb (b)

CD spectra in the near UV region (260–320 nm) arise from aromatic amino acid residues, and this absorption is specific for each antibody. Binding of **2-4** to the antibody led to an increase in ellipticity in the near UV (**Figure 2.4**), suggesting that aromatic residues are involved in the binding, and that the tertiary structure of the protein is altered upon binding. The X-ray structure of the CS-35 F_{ab} fragment in complex with **2-2** had revealed some Trp, Tyr and Phe residues in close proximity to the ligand (**Figure 2.6**), supporting this observation.

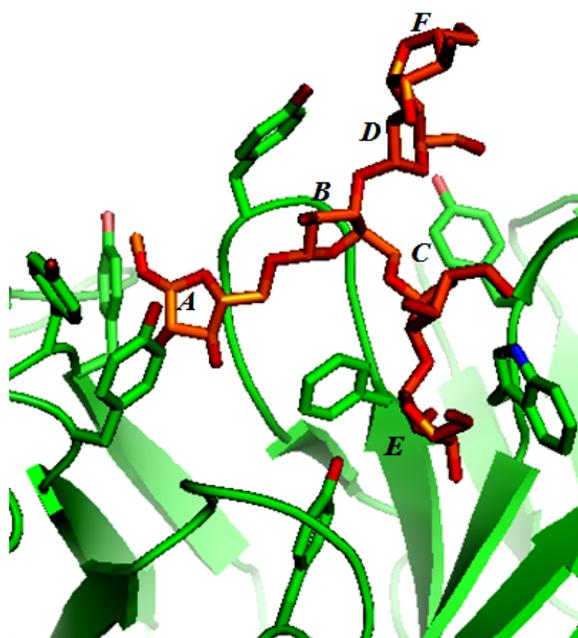


Figure 2.6 Aromatic residues of the CS-35 F_{ab} binding site around 2-2

2.3.2 Intra-Domain Disulfide Bonds

Within both variable domains of an antibody is a disulfide bond. These two intra-domain disulfide bonds are formed between the highly conserved Cys residues in V_H and V_L being H22/H92 and L23/L88, and influence the stability of the domains.⁸ In order to confirm that the two disulfide bonds are formed in the CS-35 scFv during the refolding step, I incubated the wild type scFv sample with DTT, a reducing agent, and then added IAA, an alkylating agent. To a second sample I just added iodoacetamide without the reducing agent. If the cysteines are involved in the disulfide bond, we will not expect to see alkylation by IAA in the second sample, because IAA reacts with free SH groups. The first sample was used as a control; DTT reduction of disulfide bonds leads to thiols that will be further alkylated by IAA. The ESI results showed that the control sample had an increase in mass equal to the mass of four alkylation with acetamide (4×58 amu) corresponding to the four Cys residues in the scFv. In contrast, the other sample was intact, showing no reaction with IAA, implying that the Cys residues are in the oxidized state in disulfide bonds.

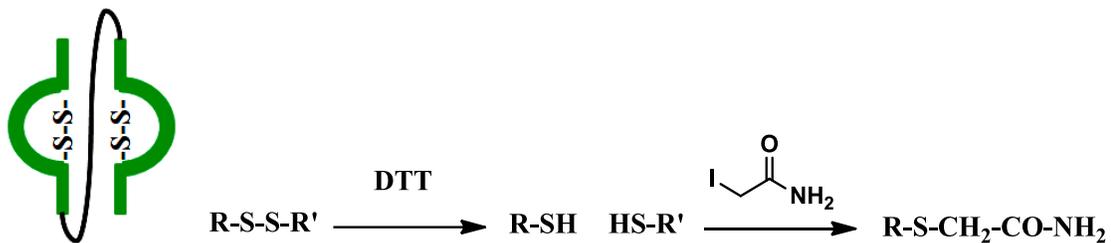


Figure 2.7 Schematic depiction of intra-domain disulfide bonds, and the reduction and alkylation reaction.

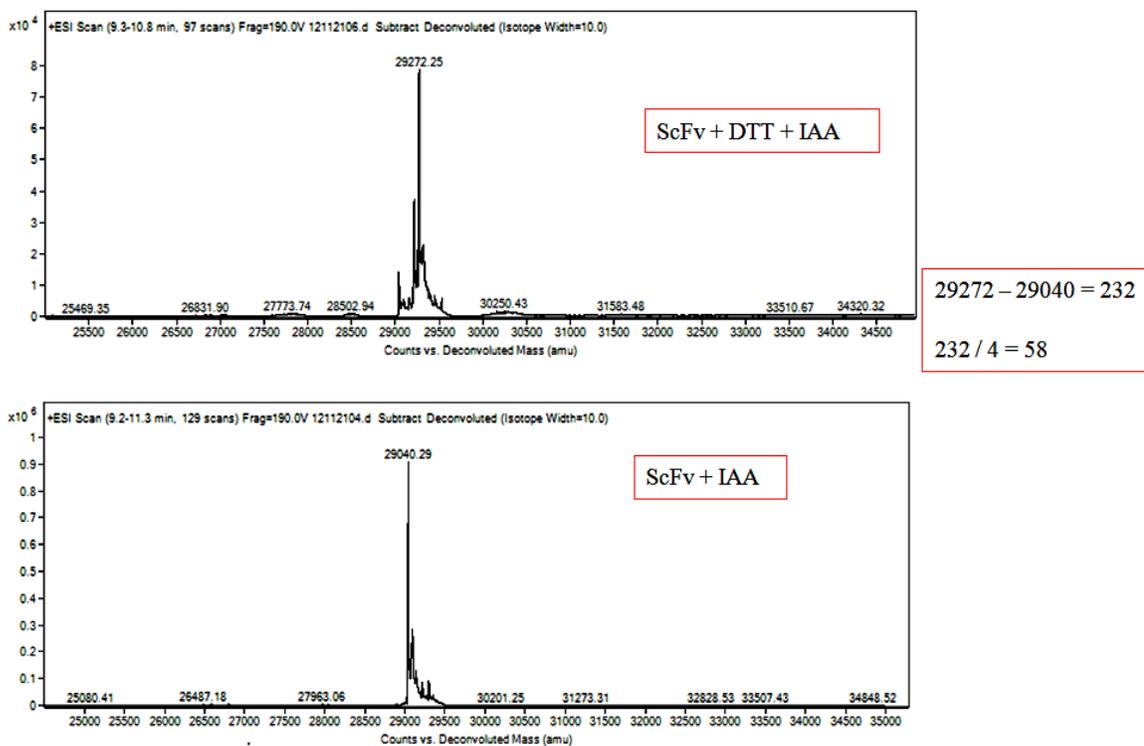


Figure 2.8 ESI results after treating scFv with IAA in the presence (top) and absence (bottom) of the reducing agent

2.3.3 Generation of CS-35 scFv Mutants

Residues were chosen for mutation based on the direct- or water-mediated hydrogen bonds they made with the **2** in the crystal structure, or the possible hydrophobic interactions between the protein and the ligand.

Tyr98 was shown to form a hydrogen bond and also stacking interactions in close proximity to ring B of **2**. Therefore it was mutated to Ala and Phe to evaluate the contribution of hydrogen bond and stacking interaction, respectively, to the binding. Asn34^L was mutated to Ala and Asp to study the necessity of hydrogen bond and its direction in the binding. Asn97^H was also mutated to Asp for the direction of the hydrogen bond. The rest of the mutations were Ala scanning. The mutants prepared were Trp33^HAla, His35^HAla, Asn58^HAla, Ser50^HAla, Phe95^HAla, Asn97^HAsp, Tyr98^HAla, Tyr98^HPhe, Asn34^LAla, Asn34^LAsp, Tyr50^LAla, as well as Ser31^LAla, which was the product of the overlap PCR.

Table 2.1 Mutants of CS-35 scFv, and the molecular interactions that are predicted to be disrupted by mutation

Hydrogen bond	Van der Waals
Trp33 ^H Ala	Trp33 ^H Ala
Tyr98 ^H Phe	Tyr98 ^H Ala
Tyr98 ^H Ala	Tyr50 ^L Ala
His35 ^H Ala	Phe95 ^H Ala
Asn58 ^H Ala	
Ser50 ^H Ala	
Asn34 ^L Ala	
Asn34 ^L Asp	
Asn97 ^H Asp	

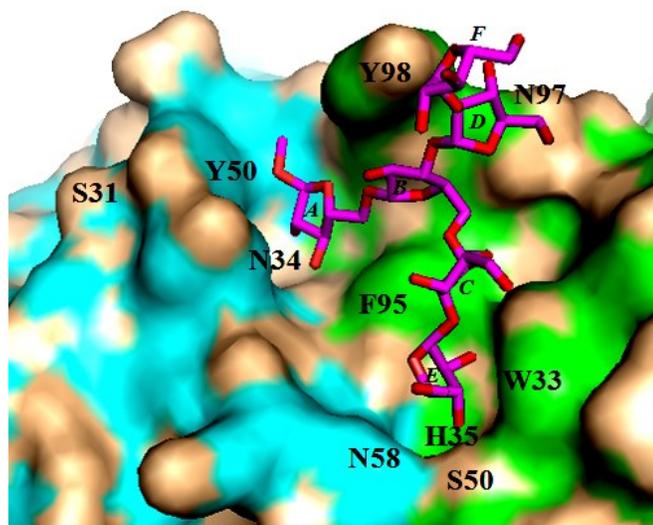


Figure 2.9 The residues of the CS-35 binding site that were mutated.

2.3.4 SPR Binding of Wild-Type CS-35 Fragments to Ara6

We analyzed the binding profiles of the CS-35 scFv, its mutants and the CS-35 F_{ab} fragment to **2-3** by SPR spectroscopy using the BIAevaluation software.⁹ Bovine Serum Albumin (BSA) was used as a negative control. The wild-type scFv showed 1:1 binding with **2-3** (**Figure 2.10**), both association and dissociation phases showed good fitting (Appendix A5) to a 1:1 binding model in the BIAevaluation⁸ software.

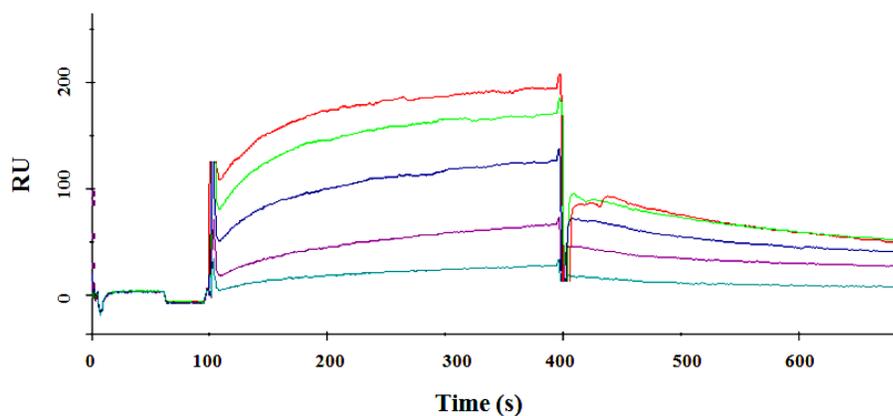


Figure 2.10 SPR binding of the wild-type CS-35 scFv with **2-3**. Concentrations of the scFv injections were 0.26 —, 0.53 —, 1.6 —, 2.6 —, 4.8 — μM .

The dissociation constant (K_D) was calculated to be 0.92 μM . This K_D value was consistent with the value calculated from Equation (2.1) ($K_D = 0.56 \mu\text{M}$). Both values are also in close agreement to the binding constant for the F_{ab} , which was measured (**Figure 2.11**) as a reference by SPR ($K_D = 0.49 \mu\text{M}$). These values are also consistent with the value that was previously reported for the interaction of the CS-35 F_{ab} fragment with ligand **2-2** measured by other ES-MS and isothermal titration calorimetry (ITC) ($K_D = 6.2 \mu\text{M}$ and $5.8 \mu\text{M}$, respectively).²

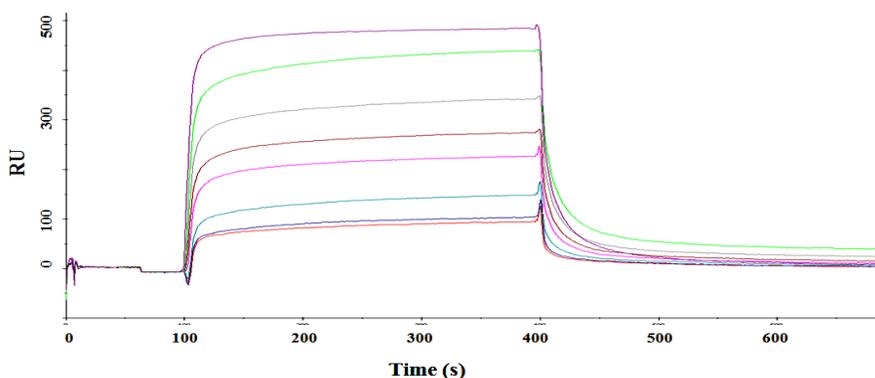


Figure 2.11 SPR binding of the CS-35 F_{ab} with **2-3**. Concentrations were: 0.02 —, 0.05 —, 0.1 —, 0.18 —, 0.31 —, 0.53 —, 1.08 —, 1.51 — μM

Table 2.2 SPR binding affinity of the wild-type CS-35 scFv with **2-3**. Affinity was obtained from association and dissociation rates (k_a and k_d , respectively), and also from equilibrium binding analysis according to Equation (2.1).

Binding	scFv/ 2-3 k_d/k_a	scFv/ 2-3 Equilibrium	F_{ab} / 2-3 k_d/k_a	F_{ab} / 2-2 ⁴
K_D (μM)	0.50	$5.6 \pm 0.05 \times 10^{-7}$	0.49	6.2

The one order of magnitude improvement in the affinity for ligand **2-3** compared to **2-2** presumably results from the improved hydrophobic interactions between the octyl chain of the ligand and the binding site, which is rich in non-polar amino acid side chain residues in that region of the protein. (**Figure 2.12**)

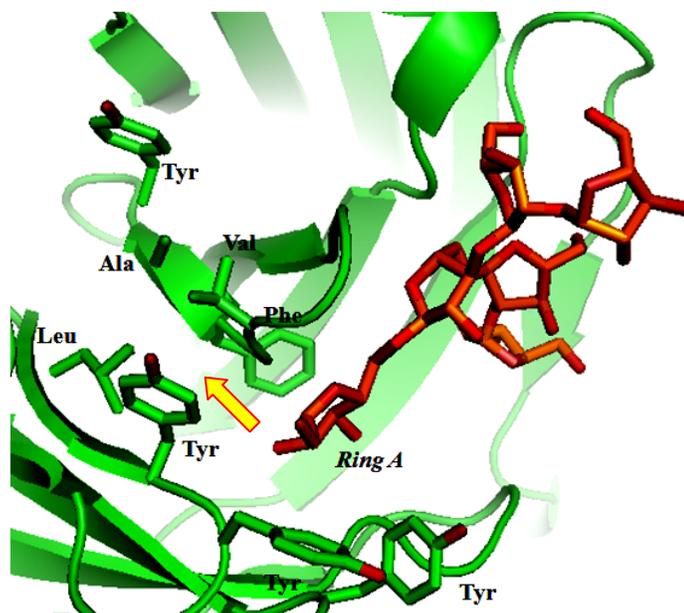


Figure 2.12 Residues around ring A. The room where octyl chain can be accommodated is shown by a yellow arrow.

The fast on-rate ($k_a = 3.86 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$) of the binding is typical of carbohydrate–protein interactions.⁵ However, we noticed a slow off-rate for the scFv–furanoside binding ($k_d = 3.56 \times 10^{-3} \text{ s}^{-1}$). Therefore, the relatively high affinity of this interaction is caused by slow dissociation. This is in agreement with the previous suggestion that the low affinities of carbohydrate–protein interactions are caused by fast dissociation of the protein from carbohydrate.⁵

Bivalent scFv dimers have been increasingly considered as potential tools for affinity enhancements for drug discovery.¹⁰ In order to compare the effect of dimerization on the binding profile of the CS-35 scFv to **2-3**, we tested its binding by SPR spectroscopy. The injection of a sample of dimer scFv showed that it was able to bind to **2-3** (**Figure 2.13**). However, its binding profile deviated from the monomer, by having increased association ($2.69 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) and decreased dissociation ($2.43 \times 10^{-6} \text{ s}^{-1}$) rates. The difference was particularly evident in the slow dissociation phase for the dimer sample which is the result of the double binding of bivalent scFv to two antigens, leading to higher affinity ($K_D (k_d/k_a) = 0.1 \text{ nM}$) due to the avidity effect.¹⁰

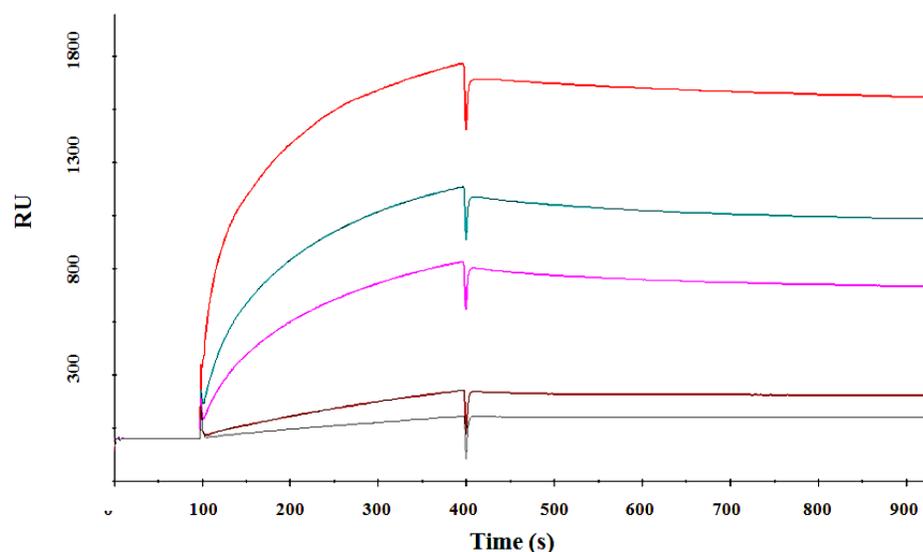


Figure 2.13 SPR binding of the CS-35 scFv dimer with **2-3**. Concentrations: 0.07 —, 0.15 —, 0.46 —, 0.77 —, 1.08 — μM

2.3.5 CS-35 scFv Mutants Binding

After characterizing the binding of the wild-type scFv, the mutants were screened for their ability to bind to **2-3** by SPR spectroscopy. The CS-35 F_{ab} and CS-35 wild type scFv were used as positive controls, and BSA was used as a negative control. Out of the twelve generated mutants, six were unable to recognize the antigen (**Table 2.3**). These included aromatic mutants such as Phe95Ala, Tyr98Ala, and Trp33Ala. Two hydrogen bond-forming residues are in this list as well: Asn34Ala, and Asn97Asp. Ser31Ala, which is not in direct contact with the antigen, is also unable to bind to **2-3** (**Figure 2.14**).

Table 2.3 Mutants and their binding to ligand 3 in SPR

Mutant	SPR Binding*	Interaction of the wild type residue
Trp33 ^H Ala	No	van der Waals and Hydrogen Bonding
Asn34 ^L Ala	No	Hydrogen Bonding
Ser31 ^L Ala	No	—
Phe95 ^H Ala	No	van der Waals

Tyr98 ^H Ala	No	van der Waals and Hydrogen Bonding
Asn97 ^H Asp	No	Hydrogen Bonding
His35 ^H Ala	Yes	Hydrogen Bonding
Asn58 ^H Ala	Yes	Hydrogen Bonding
Ser50 ^H Ala	Yes	Hydrogen Bonding
Tyr50 ^L Ala	Yes	van der Waals
Tyr98 ^H Phe	Yes	van der Waals
Asn34 ^L Asp	Yes	Hydrogen Bonding

* The SPR bindings were analyzed after the subtraction of the response units of the two flow channels. Those with zero response units were considered as non-binders.

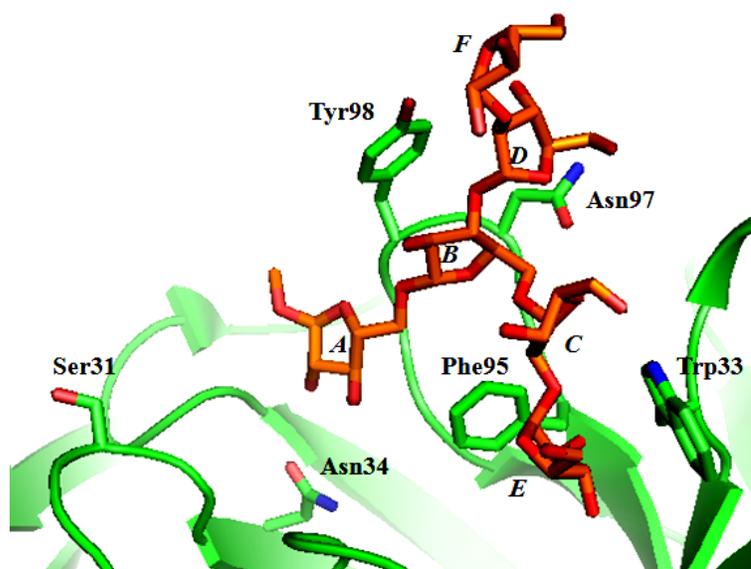


Figure 2.14 Six essential residues in the binding site of the CS-35 F_{ab}

The other six mutants that retained binding were further investigated by SPR spectroscopy to obtain affinities and kinetics. Although mutants His35^HAla, Asn58^HAla, and Ser50^HAla were able to bind to **2-3**, they exhibited substantially decreased binding affinities, in the range of millimolar. These three residues had been shown to be involved in hydrogen bonding with rings C and E of the ligand in the crystal structure.

Mutants that retained binding with greater affinity (Tyr50^LAla, Tyr98^HPhe and Asn34^LAsp), were further analyzed. In all cases, a mixture of monomer and dimer showed biphasic binding. While monomeric Ser50^HAla, His35^HAla, and Asn58^HAla were weakly binding to **2-3**, the samples that contained a mixture of monomer and dimer, and the samples containing pure dimer showed deviations from 1:1 binding with obviously slower dissociations.

The affinities of Tyr50^LAla, Asn34^LAsp, and Tyr98^HPhe remained close to that of the wild-type scFv. Indeed, Tyr50^LAla and Asn34^LAsp mutants showed slower dissociation rates compared to the wild-type fragment.

Table 2.4 Kinetics and affinity evaluation of the six scFv mutants showing binding to **2-3** from different complementary determining regions (CDR) for ligand **3** from SPR experiments.

	CDR	Mutant	K_D (M) k_d/k_a	K_D (M) Equilibrium	k_a (M ⁻¹ s ⁻¹)	k_d (s ⁻¹)
1	H1	His35Ala	-	<10 ³	-	-
2	H1	Asn58Ala	-	<10 ³	-	-
3	H2	Ser50Ala	-	<10 ³	-	-
4	H3	Tyr98Phe	1.85×10^{-6}	$1.8 \pm 0.21 \times 10^{-6}$	2.66×10^3	4.91×10^{-3}
5	L1	Asn34Asp	3.03×10^{-6}	$4.02 \pm 4.2 \times 10^{-6}$	40	1.21×10^{-4}
6	L2	Tyr50Ala	7.4×10^{-6}	$7.7 \pm 0.85 \times 10^{-6}$	56.7	3.35×10^{-4}
7	-	scFv wild	5.0×10^{-7}	$5.6 \pm 0.05 \times 10^{-7}$	3.86×10^3	3.56×10^{-3}

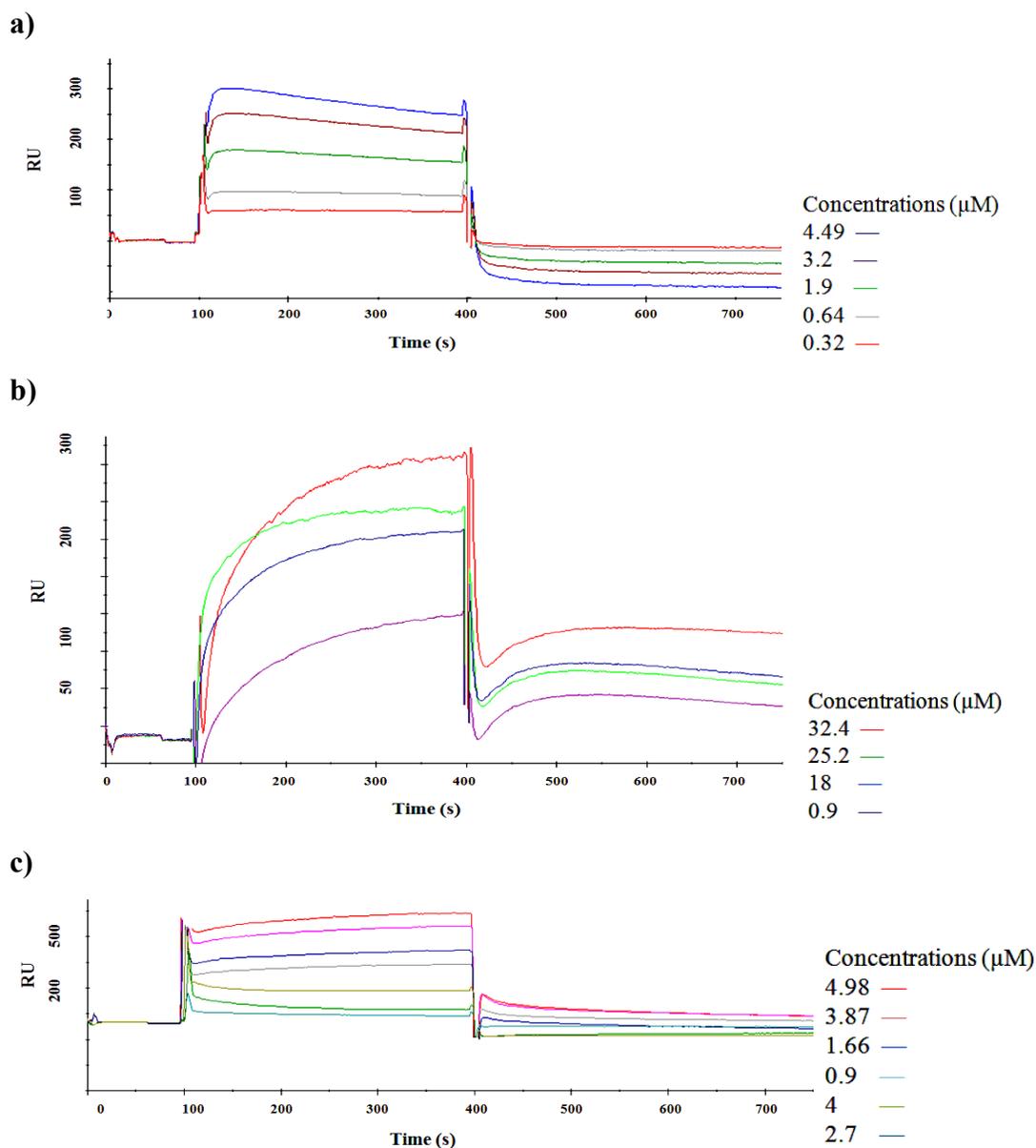


Figure 2.15 SPR binding of **a)** Asn34Asp, **b)** Tyr50Ala, and **c)** Tyr98Phe to Ara6 2-3

2.3.6 Structural Studies by Mass Spectrometry

The initial goal of the mass spectrometric studies was to investigate the binding of the CS-35 scFv and its mutants by mass spectrometry, and to compare the results with those obtained from SPR spectroscopy. However, the results showed two Gaussian distributions at the mass for the wild type scFv. This implied that the scFv existed in two conformations, one compact, the

other extended. The extended conformation is the one which has a larger charge number (z) compared to the compact conformation.¹¹ Therefore, it has smaller m/z value, since the mass is identical in the two species. Wild-type scFv showed binding to **2-4** after being sprayed with the ligand. The two conformations remained after adding **2-4**, which implies that both conformations were able to bind to the ligand. To explain the origin of the second conformation, two hypotheses were tested.

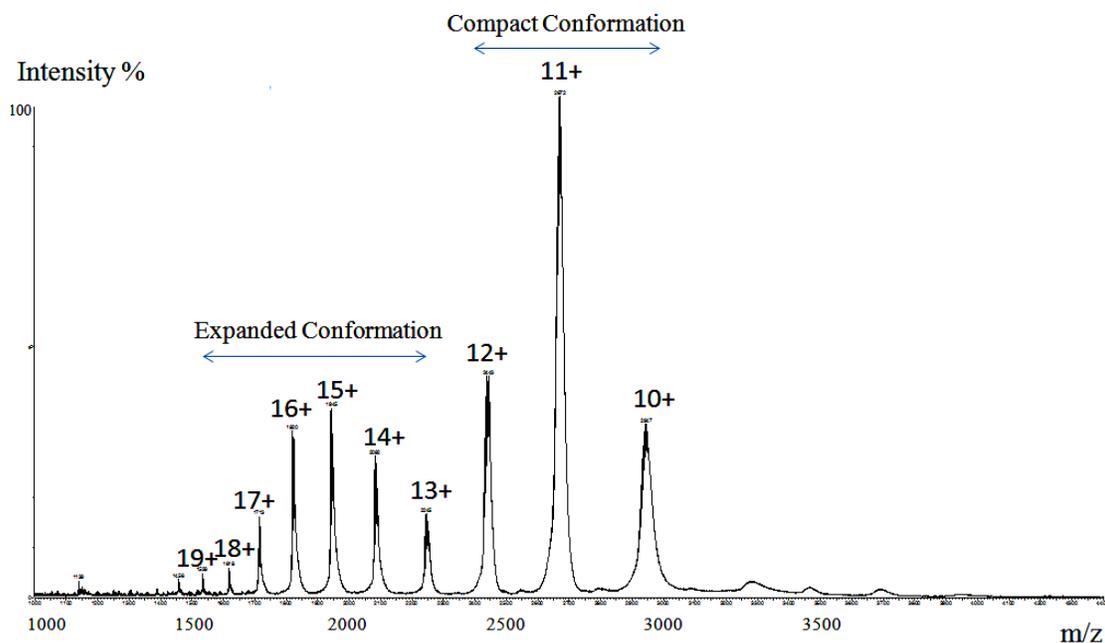


Figure 2.16 ESI mass depiction of two conformations of the CS-35 scFv

First, inspection of the CS-35 sequence in the crystal structure revealed His residues. The pI of a His amino acid is ~ 6.5 , which is close to the pH of the ammonium acetate buffer (6.8). Therefore, the His residue can exist possibly in two protonation states, in equilibrium (**Figure 2.17**). Such an equilibrium might be a source of two conformations.¹² To test the effect of His protonation state on the scFv conformation, the wild type scFv was prepared at three different pH s (3.0, 6.8 and 8.0). The results showed that the change of the pH , and thus the protonation of the protein, does not inhibit the formation of two conformations, although it does alter the relative conformations between the two states.

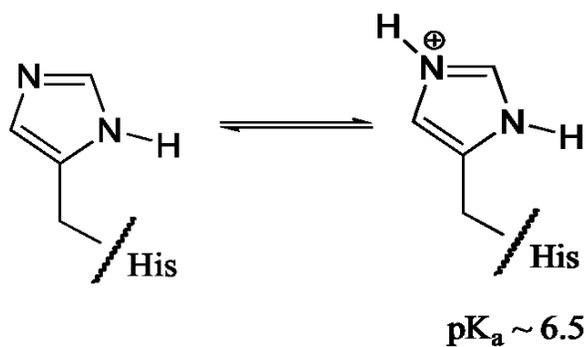


Figure 2.17 His ionization equilibrium

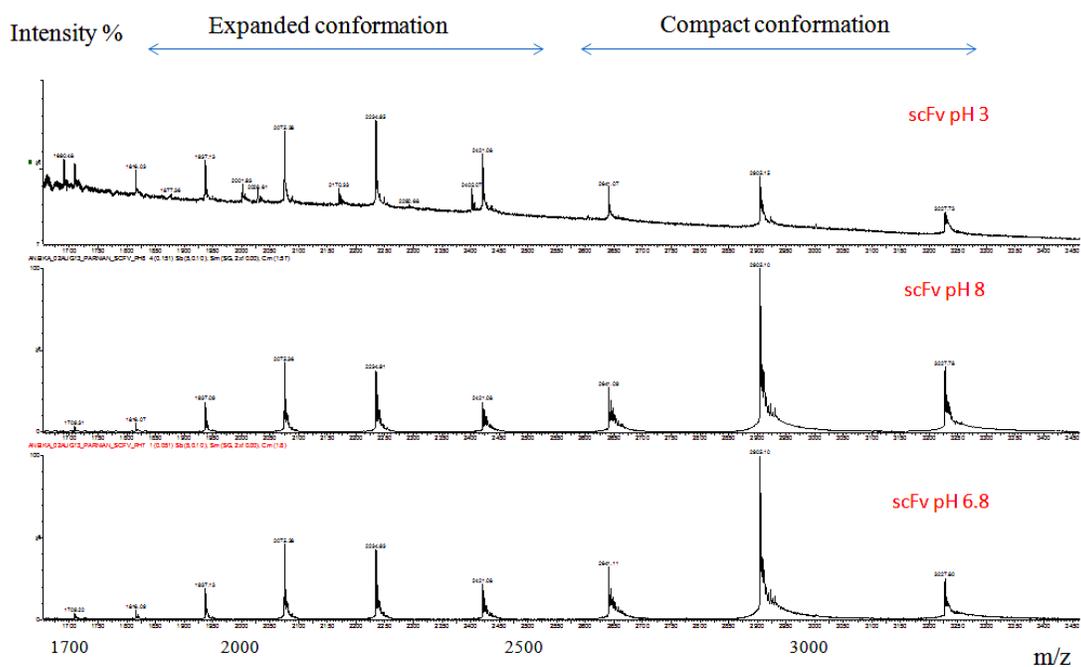


Figure 2.18 ESI-MS study of the CS-35 scFv at three pH conditions: pH : 3 (a), pH : 8 (b), and pH : 6.8 (c)

Subsequently, five scFv mutants that had shown to retain binding to Ara6 **2-4** were sprayed with the ligand under the same buffer conditions at pH 6.8. It was shown that mutants also behaved in the same way as the wild type monomeric scFv, indicating that they also existed in two conformations and that both conformations were able to bind to **2-4**. **Figure 2.19** shows how these mutants revealed two Gaussian shaped distribution of charge states with the same mass and different charges indicating the existence of two conformations.

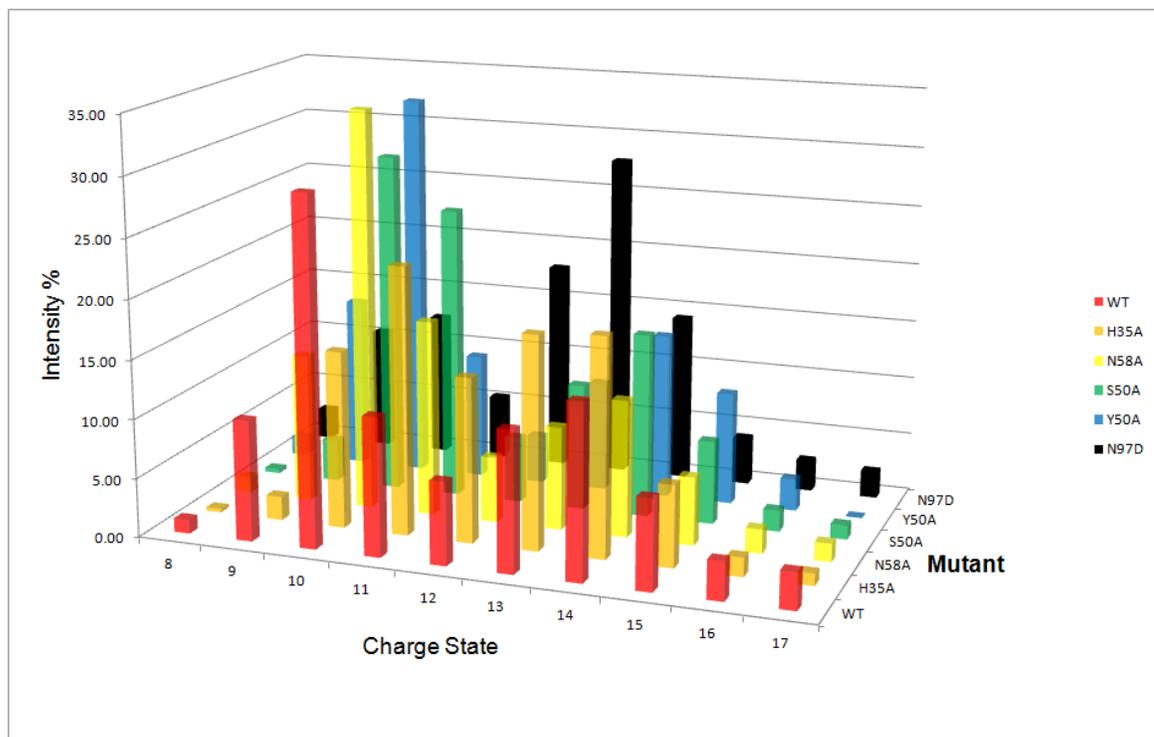


Figure 2.19 Two charge states of CS-35 mutants in ESI mass. For each mutant two distribution sets of charge state are observed that represent the compact and the extended conformations.

Second, I inspected the structure of scFv again. In the construct we had made for the scFv, the N-terminus of the protein was inserted after a BamHI restriction site, which means there was a sequence of 34 amino acids starting from the first Met of the plasmid, and prior to the first residue of the scFv, at the N-terminus of the protein. I hypothesized that the extra sequence could result in the protein existing in more than one conformation. To test this, wild type scFv, which

lacks this extra sequence, was sprayed under the same condition, and showed only one compact conformation. **Figure 2.20** depicts the result of this CS-35 scFv.

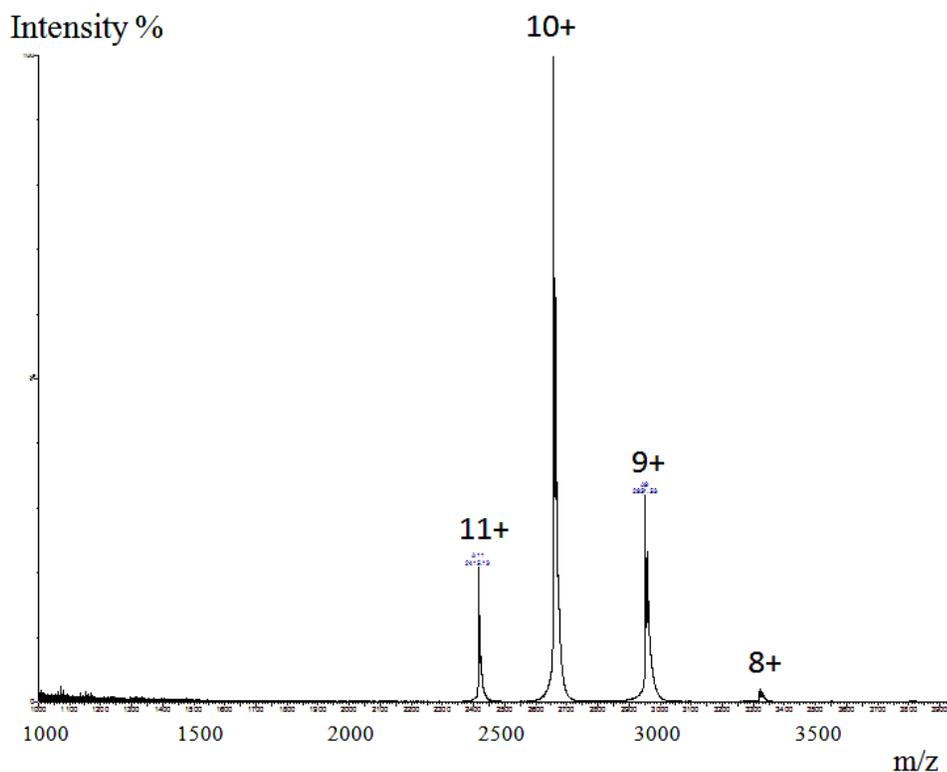


Figure 2.20 ES-MS spectrum of the CS-35 scFv lacking the extra sequence, which precedes the first Met residue at the C-terminus. There is one conformation in corresponding to the compact conformation.

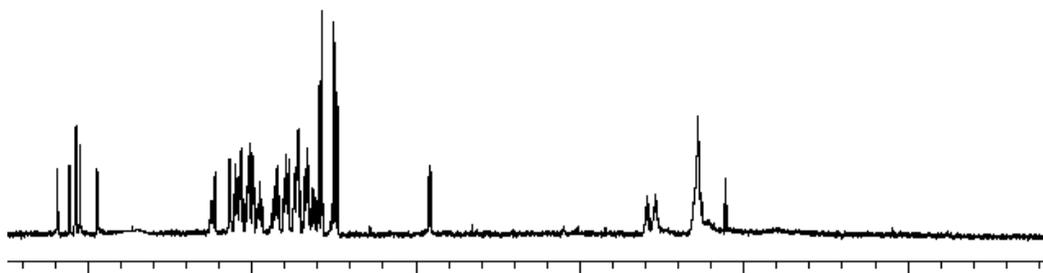
However, according to our collaborators measuring the binding affinity in the presence of two conformation states, required high amounts of scFv sample (at least 50 μ M for each protein sample), which was a limitation in studying affinities by this method, due to the protein yields and behavior in such high concentrations.

2.3.7 STD NMR Spectroscopy of Wild-Type and Mutant Proteins with 2-3

The ^1H NMR spectrum of **2-2** was fully assigned in 2007 and the binding of the CS-35 F_{ab} was then probed using **2-2** as the ligand by STD NMR spectroscopy.² Based on those data, it was

proposed that **2-2** could bind to the CS-35 F_{ab} in one of the two binding modes – one involving the ABCE rings of **2-2**, and the other involving the ABDF rings as the epitope (see **Figure 2.1** for the definition of the ring forms). The favored binding epitope, however, involved rings ABCE and, indeed, this was the only binding mode observed in the crystal structure.⁴ In such a structure, ring B was the hinge at the center of the binding motif. In the crystal structure, this ring adopted a conformation deviating from the low energy conformation in solution, extending ring A, and the C–E arm deep into the binding groove, while the D–F arm extended out of the binding site. I used STD NMR spectroscopy to examine the effect of mutation on epitope recognition by CS-35. The STD NMR spectrum of the F_{ab} with ligand **2-4** were also obtained as the reference (**Figure 2.21**), which were consistent with the previously reported STD NMR results for the interaction of F_{ab} with **2-2**.

a)



b)

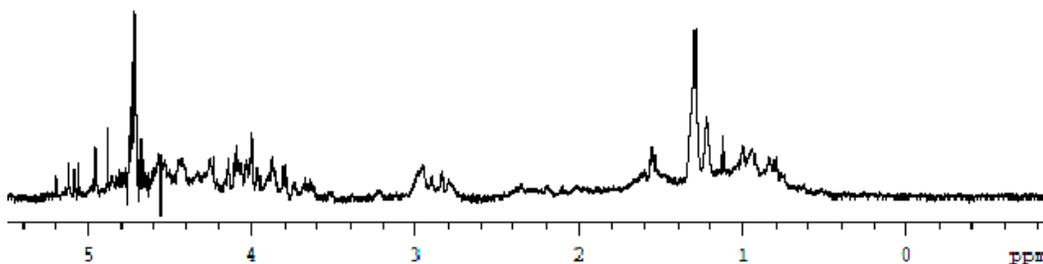


Figure 2.21 ¹H NMR spectrum of **2-4** (a), STD experiment of the CS-35 F_{ab} and **2-4** (b) The extra peaks at 4.3–4.8 ppm belong to water and glycerol traces remaining from the centrifuge filter membrane. The extra peaks below around 1 ppm are residual protein background.

The STD NMR spectra for the four mutants (Tyr50^LAla, His35^HAla, Asn58^HAla, and Ser50^HAla) (**Figure 2.23**) that were shown to recognize the ligand in the SPR spectroscopy experiments, and which could be expressed in good yields, were obtained in combination with ligand 2-4 (**Figure 2.22**).

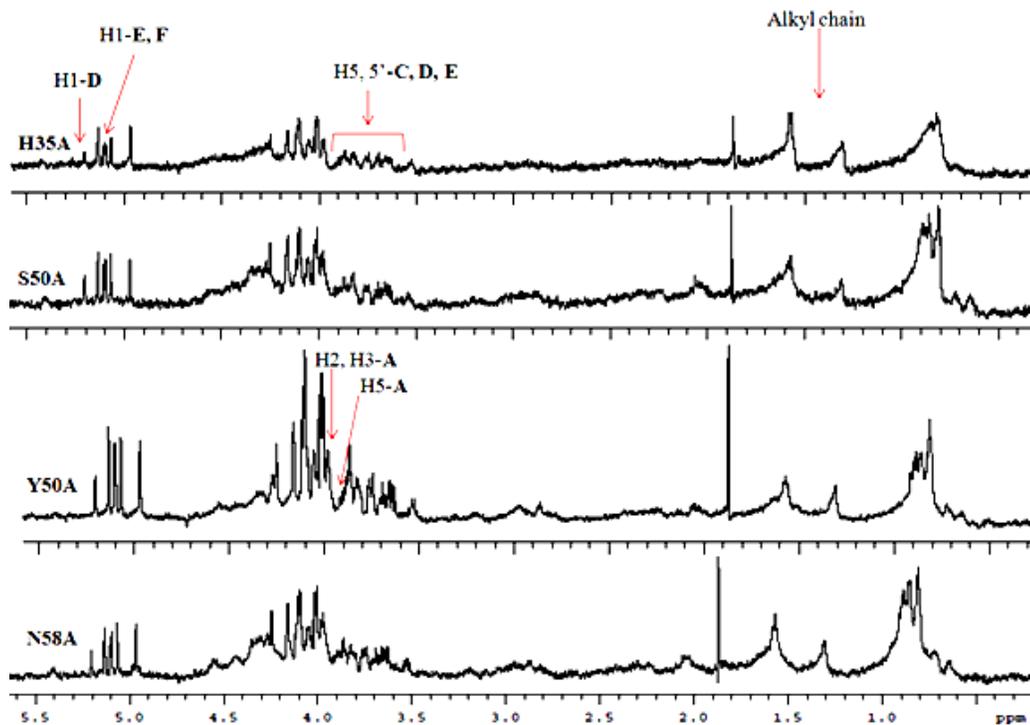


Figure 2.22 STD NMR spectra of four CS-35 scFv mutants (**b**) with 2-4.

Consistent with the SPR spectroscopic results, STD effects confirmed the decreased binding to the Ara6 ligand upon mutation. It is obvious that STD effects are dramatically affected when His35 is mutated to Ala. The decrease is especially evident in the protons belonging to rings C to E, which makes sense with the recognition model, in which His35 is close to these rings, interacting with Ara6 by donating a hydrogen bond to the ring E (**Figure 2.25**). Although the protein tolerates this mutation to some extent, the affinity is considerably reduced ($1.08 (\mu\text{M})^{-1}$ in the wild-type vs $<10^3 (\mu\text{M})^{-1}$ in the His35^HAla mutant).

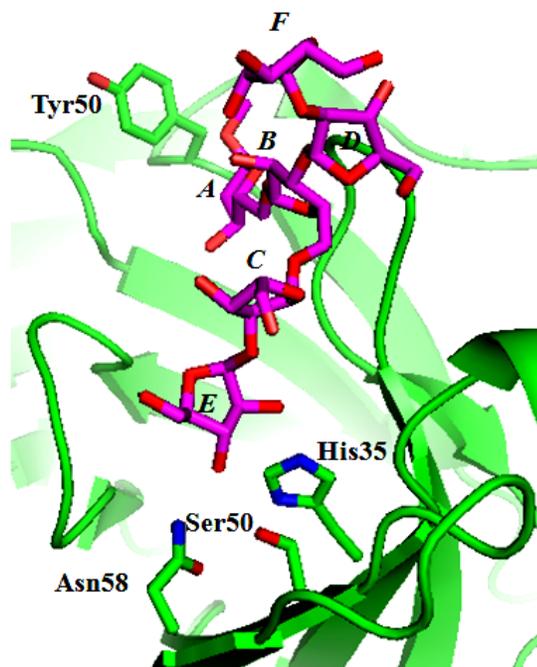
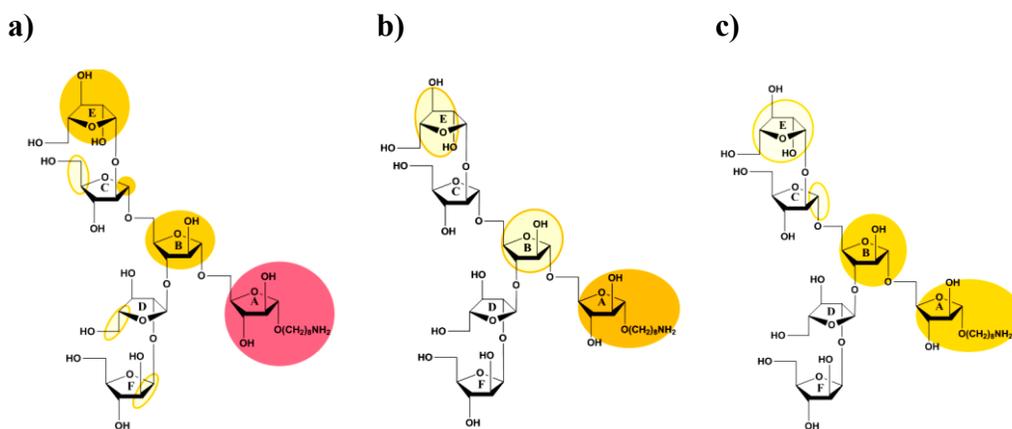


Figure 2.23 Residues that were studied by STD NMR upon mutation to Ala.

The Ser50Ala and Asn58Ala mutants also show decreases in the intensities of rings C, D and E. These three residues (His35, Ser50, and Asn58) form a hydrogen bond network around ring E (**Figure 2.25**). In the STD NMR spectra of all four mutants, however, the essential involvement of rings A and B in the binding epitope was obvious. They also all demonstrated the engagement of the C–E arm as well as the D–F arm in their binding, by showing the peaks corresponding to these rings protons. However, the intensities of the peaks for the D–F arm in all mutants was lower than the those for the C–E arm, which confirms the previous data that the ligand is recognized preferentially by the A, B, C and E rings.² We suggest this is due to the critical CH– π bonds formed between ring E and Phe95 and Trp33 of the antibody, and the hydrogen bond network created between His35, Ser50, and Asn58 from antibody and rings C and E of the antigen. Mutation of Phe95 and Trp33, two CH– π bond forming residues, to Ala leads to the abolition of the binding. The formation of CH– π interactions and hydrogen bonds rely on proper orientation and direction of the involved groups.¹¹ Such critical distance and orientation is only satisfied by the extension of C–E arm from hinge B through the C5 methylene group and the flexibility of this linkage. In contrast, the D–F arm is connected through O3 and thus lacks the methylene group. This results in a less efficient interaction with the two aromatic

residues, and the hydrogen bond network, which is reflected in the decreased intensities of STD NMR spectra for these rings.

The protons that were mostly affected by mutation in the STD NMR spectra are highlighted in **Figure 2.24**. Mutations at His35, Ser50, and Asn58 residues led to decreases in the intensities of the peaks for rings C, D, E, and F, with His35Ala showing the most significant decrease, while protons of ring A and B were less affected. This implies that although the hydrogen bond network by these three amino acid residues is important to interact with ring E, the recognition is tolerant to the alteration of these hydrogen bonds to some extent. Tyr50Ala, which is adjacent to ring A, showed a decrease in the intensities of protons of ring A, but the STD effects for the octyl chain were not affected very much. This result confirms the decreased affinity in SPR data for the Tyr50Ala mutant compared to the wild type, due to the less efficient interaction with ring A.



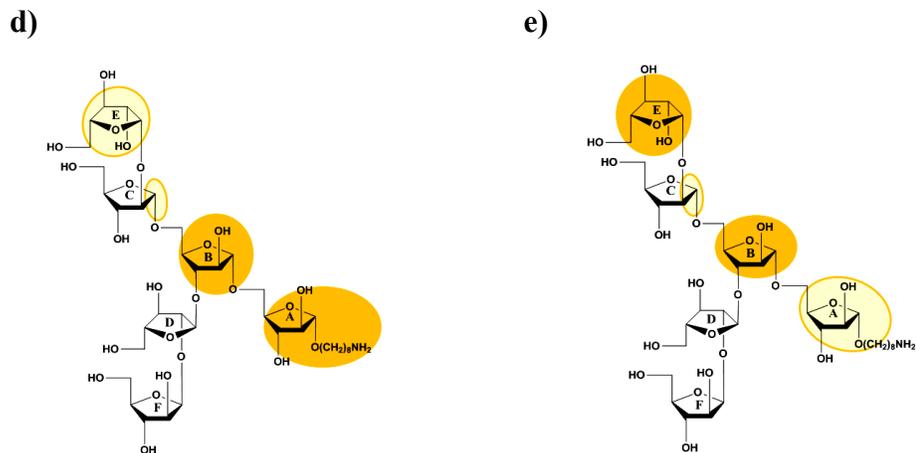


Figure 2.24 Protons with strong (pink) medium (orange) and weak (yellow) STD-NMR signals for **a)** CS-35 wild-type scFv **b)** His35Ala mutant **c)** Ser50Ala mutant **d)** Asn58Ala mutant **e)** Tyr50Ala mutant. The effect of mutating His35 is the most significant. The Ser50 and Asn58 mutants have similar patterns to each other.

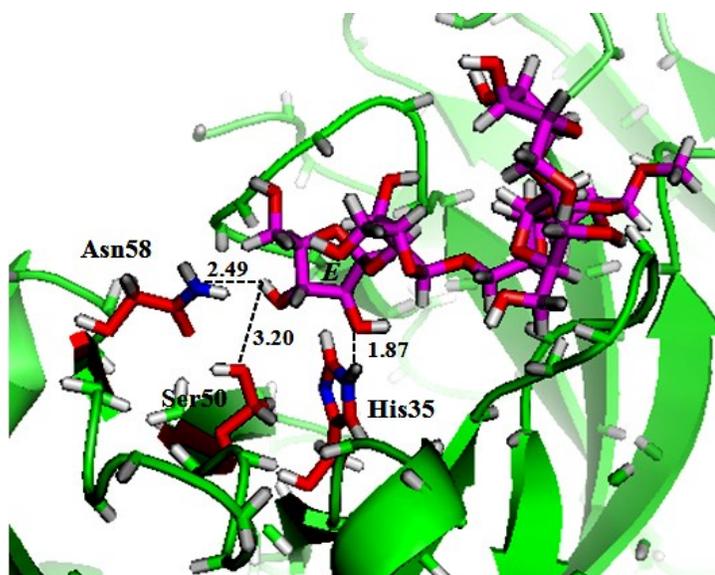


Figure 2.25 Hydrogen bond network around ring E

2.4 Discussion

The specificity motifs of the interaction of arabinofuranosides with proteins have not been investigated in significant detail. A useful approach to define the specificity of an interaction is to manipulate the structure of the binding partners. By generating a stable scFv for the CS-35 antibody, which maintained its ability to bind to the Ara6 ligand, we were able to manipulate the binding site by producing a rational library of CS-35 scFv mutants. Access to these mutants allowed us to investigate the details of the binding to the Ara6 antigen.

2.4.1 CS-35 scFv Structure and Stability

The order of the variable domains in scFvs has been studied and discussed previously. It has been shown that V_H -Linker- V_L is more stable than V_L -Linker- V_H in most of the cases.¹³ The size and amino acid content of the linker have also been studied.¹³ It is known that linkers shorter than 12 amino acids lead to high degrees of dimerization and oligomerization, which is useful for the purpose of generating bivalent or multivalent scFvs. Very long linkers, on the other hand, can lead to problems in folding and stability of the scFv. A size of 15–21 amino acids is considered to be optimum for a stable scFv. Gly residues are incorporated into the linker

to provide enough flexibility and Ser residues provide enough polarity to the structure of the linker, which is important in the stability of the fragment.

The stability of the scFvs is an important requirement for using these fragments in research. Pluckthun and colleagues have pioneered the studies on the stability of various scFvs.^{11,15,16} The overall stability of the two-domain proteins such as a scFvs, depends on the intrinsic stability of the constituent parts, here the V_H and V_L domains, and the stability of the surface between these domains.⁹ A hallmark of the intrinsic stability of protein domains relies on the intra-domain disulfide bonds. Although some scFvs have been shown to tolerate the removal of disulfide bonds by mutating Cys residues, most scFvs cannot refold natively in the absence of these disulfide bonds.¹⁸ I tried various oxidizing and reducing conditions in the refolding of the CS-35 scFv produced in the inclusion bodies. Adding reducing and oxidizing reagents in different steps of refolding complicated the formation of monomer scFv, and led to the formation of aggregates, presumably through the formation of unwanted disulfide bonds. In contrast, in the absence of the redox reagents, the variable domains were refolded naturally from inclusion bodies with the correct intra-molecular disulfide bonds. The CS-35 scFv was properly folded with the typical β -sheet structure of antibodies as determined by UV spectroscopy. By using IAA in the presence and absence of a reducing agent, the scFv was shown to retain its intra-domain disulfide bonds. This fragment was stable up to three weeks at 4 °C, and stable for at least two months at -80 °C.

As we discussed in Chapter 1, the variable domains in F_{ab}s are associated with the constant domains, while scFvs contain only the variable domains. Therefore, the amino acids that form the interface between the variable domains and the constant domain in the F_{ab} become solvent exposed in the scFv. This interface has been shown to be rich in the hydrophobic residues, and can be a source of instability of scFvs in solution, leading to aggregation, including dimer formation.¹⁵ It has been shown that mutations at these hydrophobic patches can lead to improved folding and stability of scFvs.¹⁵ Also, the interface between the two variable domains influences the stability of the structure in the absence of constant domains. This can lead to the formation of dimer scFvs. In order to avoid that, I tested the impact of different factors such as buffer, pH, storage condition (temperature and time), and concentration on the dimerization of the CS-35 scFv. Overall, the equilibrium favored the formation of the monomer scFv in 50 mM ammonium

acetate, pH 6.8. In summary, the CS-35 scFv could be successfully produced in high purity, in the monomeric state, properly folded and containing the conserved intra-domain disulfide bonds.

The scFv mutants were generated under the same conditions. The Phe95Ala mutant was the least stable scFv. Under most of pH (6.8–7.4) and buffer (phosphate, ammonium acetate, Tris sodium hydrochloride) conditions, it precipitated at lower concentrations than other mutants. The stability of Tyr98Ala, and Asn97Ala was also noticeably lower than other mutants. These residues belong to CDR H3 of the antibody and it seems that manipulation at this domain leads to lower stability of the fragment. Overall, among 12 mutants the yield of Asn34Asp was the lowest (0.5 mg/L). In contrast, the Tyr50Ala mutant had the highest yield of production (3 mg/L) and the protein possessed high stability and a lower ratio of dimerization in ammonium acetate buffer. However, it was not stable in HEPES buffer. The fact that single mutations affect the yield of production and stability of the scFv implies how sensitive the interface of the two variable domains is to manipulation.

2.4.2 Discussion of Binding

The affinity of the CS-35 to hexasaccharide **2-2** had previously been probed by ESI mass spectrometry.² In this project, SPR spectroscopy provided the opportunity to study the rates of the association and dissociation of the ligand for the CS-35 scFv and mutants. Among the advantages this method provided to our research was the possibility of using the same immobilized ligand for many measurements. The synthesis of Ara6-based ligands is achieved after many steps; however, once it is covalently attached to the CM5 chip, it is stable for a long time and can be used repeatedly for different proteins at different concentrations.

Hexasaccharide **2-4**, which was used for the binding measurements, contains an amino alkyl chain through which the sugar is fixed to the surface of the SPR chip. The alkyl chain length is assumed to provide enough space from the chip surface to permit the freedom of movement to the sugar chain, to fit properly into the antibody binding site. This chain also **increased** the binding affinity of to the CS-35 F_{ab} presumably due to hydrophobic interactions with several Tyr, Phe, Leu and Ile residues in the binding site.

The F_{ab} and monomer scFv use a single binding site to recognize Ara6-based ligands. The influence of valence on the affinity and kinetics of *Salmonella* antigen and Se155-4 was

investigated by Mackenzie *et. al*, and it was shown that trace amounts of dimer can lead to a biphasic binding in the scFv.⁵ Comparison of monomer and dimer CS-35 scFv also showed that two binding sites dramatically influence the binding by avidity effects.

SPR spectroscopy was also used to screen CS-35 scFv mutants for their ability to bind to an Ara6-containing ligand immobilized on a CM5 chip. The results moved the study one step closer to answer to the question "In this system, what forces govern the recognition of the arabinofuranoside structure by the protein?" The screening results revealed the tolerance of the binding to the change of some hydrogen bonds, and the eradication of the recognition upon manipulating some aromatic residues. A study of the CS-35 scFv with **2-4** in a near UV CD spectroscopic experiment also revealed the involvement of aromatic residues in the binding.

2.4.3 Inter-Molecular Forces and Specificity Motifs in the CS-35 scFv–Ara6 System

Carbohydrate–protein interactions are usually weak when there is a single site of binding. The affinity of such interactions is considerably enhanced by avidity effects through binding with multiple binding sites. The Ara6–CS-35 recognition system, however, possess high intrinsic binding strength with dissociation constants in the sub-micromolar range.

Because there is no site of electrostatic interaction in this system (the ligand is uncharged), hydrogen bonds and hydrophobic interaction are responsible for the affinity of the carbohydrate and protein. Carbohydrates are abundant in the number of hydroxyl groups and ring oxygens that are capable of forming multiple hydrogen bonds. The hydrophobic interactions and hydrogen bonds that are formed between a carbohydrate and a protein should be strong enough to compensate the breaking of hydrogen bonds of both binding partners with water.

The first screen of the 12 mutants for their binding to **2-3** revealed six residues that were critical for the interaction to survive. Among the proteins that were unable to bind to **2-3** were three in which aromatic amino acids (Trp33, Phe95 and Tyr98) had been replaced with Ala. While Tyr98 involves its OH group in a water-mediated hydrogen bond with the 2-OH group of ring B of **2-2**, the results showed that removal of this hydrogen bond (the Tyr98Phe mutant) did not significantly affect the binding. Therefore, the interaction of the aromatic ring of Tyr98 is the main molecular interaction for the recognition of these essential residues. This interaction is a

CH- π bond between the aromatic ring of Tyr98 and H2 of ring B in 2.63 Å distance. Inspection of the CS-35F_{ab}-2-2 crystal structure revealed that the aromatic rings of Trp33 and Phe95 are also in close enough proximity to CH bonds of ring E (C1-H1 and C3-H3) to form CH- π interactions with T-shaped geometries.¹⁸ The distances of these interactions of ring E with Phe95 and Trp33 are 2.85 Å and 2.63 Å, respectively.

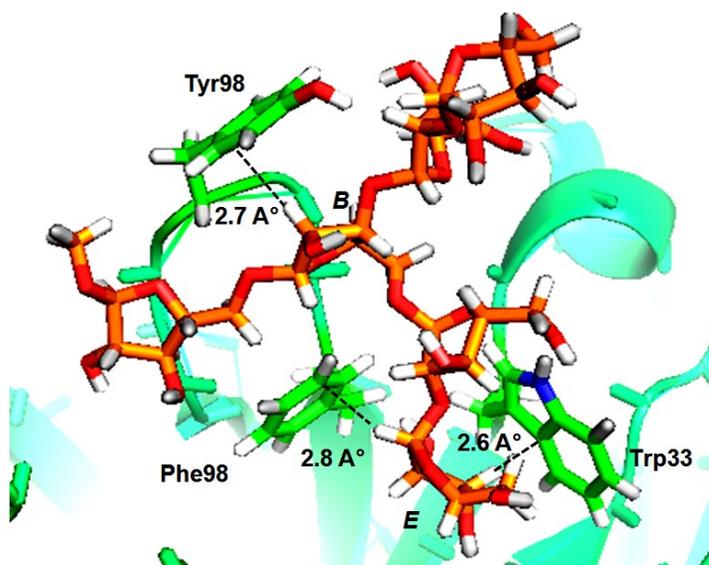


Figure 2.26 CH- π interactions of Tyr98, Phe95, and Trp33 with 2-2 from the crystal structure of the CS-35 F_{ab}. The distances are shown in Ångstroms.

The results from the CD spectroscopic data and mutation studies suggested that there are considerable interactions between the furanoside rings of 2-4 and aromatic amino acid residues of CS-35. Such interactions have been frequently reported for the recognition of pyranosides, which often have discernible hydrophobic faces, by proteins.²¹ In the Ara6-CS-35 interaction system, ring B and E, which make the main CH- π interactions have a N-type ring pucker. The flexibility of ring B assists it in hinging the structure, and making a critical CH- π interaction with Tyr98. The hydrophobic face of Ring E stacks against Trp33; it also makes a CH- π interaction with Phe95. These data exhibit the importance of CH- π bonds in building the affinity and specificity of a furanoside-protein interaction.

Two of the key aromatic residues (Phe95 and Tyr98) occur in the Phe95-Gly96-Asn97-Tyr98 sequence in CDR H3. It is useful to discuss this segment of the CS-35 binding site

individually. This sequence is part of CDR H3, and contains three critical residues necessary for the specificity of the recognition (Phe, Asn, and Tyr). I previously mentioned that these three residues influence the stability of the scFv, as well. This segment also represents an interesting structural example; it has been shown in the crystal structure⁴ that the peptide bond between Phe95 and Gly96 residues is *cis*. This is one of the rare cases where a *cis* peptide bond involves non-proline amino acids. Such *cis* bonds are believed to be functionally important and engaged in regulating biochemical properties.¹¹ In the current recognition system, the *cis* peptide bond positions Phe95 properly to form a critical CH- π contact with H2 of ring E of the ligand (**Figure 2.27**). Asn97 is the third essential residue in this sequence (**Figure 2.28**), which makes a water-mediated hydrogen bond with the ring oxygen of residue C, by its side chain nitrogen. Mutation of this amino acid destroys the binding.

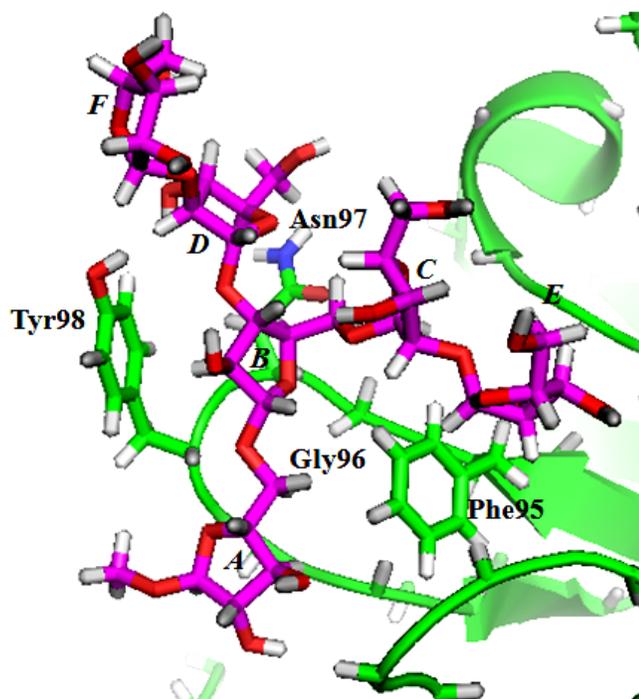


Figure 2.27 Phe95–Gly96–Asn97–Tyr98 segment of the CS-35 F_{ab} around 2-2 in the X-ray structure of the complex

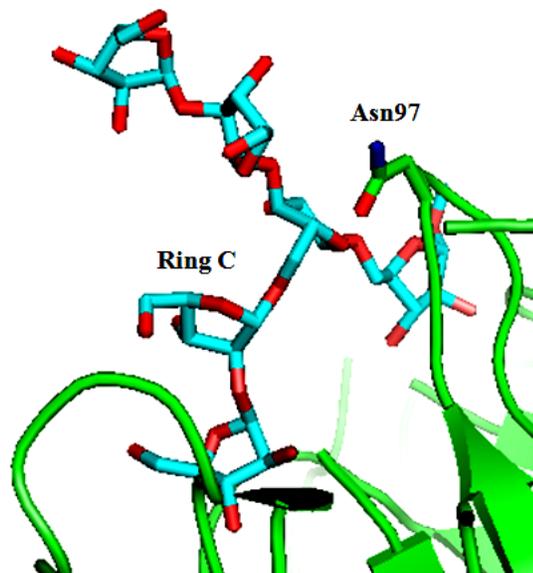
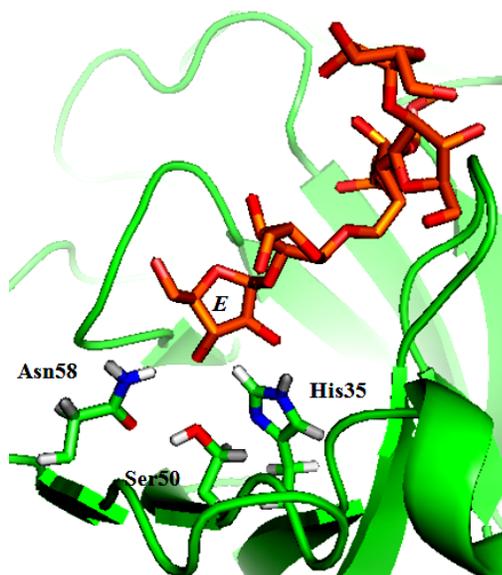


Figure 2.28 Asn97 and ring C of **2** in the X-ray structure of the CS-35 F_{ab}

The binding results showed that the recognition is, to some extent, tolerant to the alteration of the hydrogen bonds formed by His35 (2-OH, ring E), Ser50 (3-OH, ring C) and Asn58 (3-OH, ring C), although the affinities upon mutating these residues are reduced. This hydrogen bond network indeed strengthens the interaction from C–E arm of the ligand. We suggest that the D–F arm, which is shorter than C–E arm by one methylene group, does not extend deeply enough to this network for the effective contact, and is, therefore, the less preferred epitope compared to C–E arm.

a)



b)

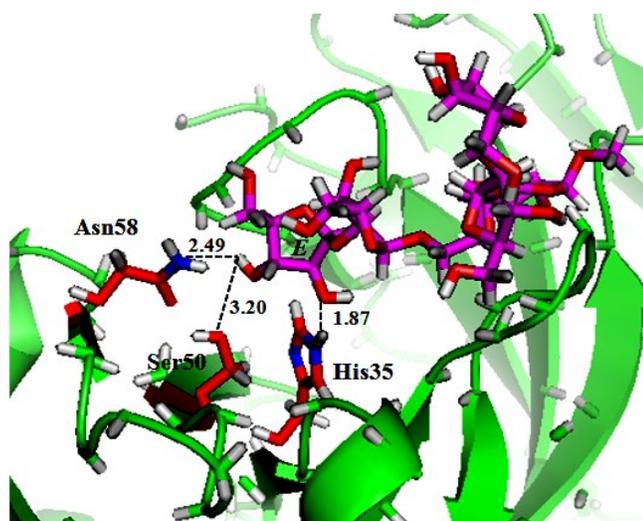


Figure 2.29 a) Hydrogen bond making residues around ring E in the crystal structure of the CS-35 F_{ab} and 2-2. b) The distances of the interactions

Asn34 is another hydrogen-bond-forming residue, mutation of which to Ala, leads to the destruction of binding. However, when it is mutated to Asp, it retains some affinity for the binding.

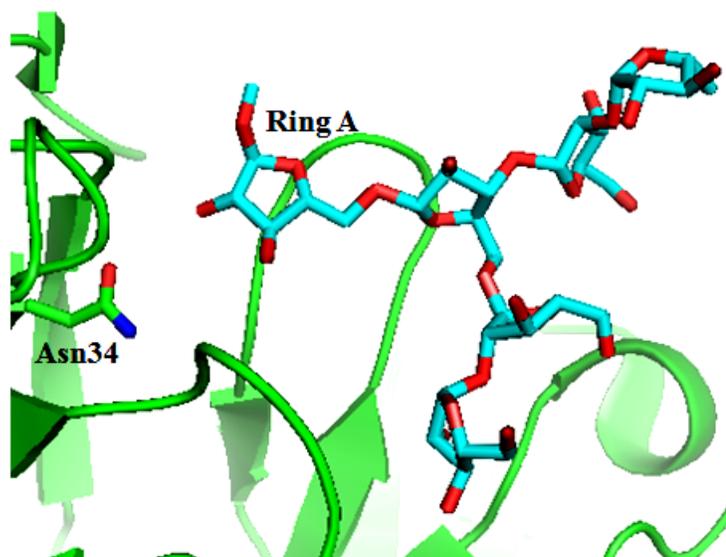


Figure 2.30 Asn34 close to ring A

Ser31 and Tyr50 are residues close to ring A, which bears the alkyl chain. In the crystal structure of the complex, Ser31 is not in direct contact with **2-2**. However, mutating this residue inhibits the recognition, which appears to be because it is in a strategic region close to ring A where key amino acids come together to interact with **2-2** (**Figure 2.31**). The interaction of this residue in the binding site with adjacent amino acids is presumably assisting other residues to align properly to bind to the ligand. In the Tyr50Ala mutant, however, there was not a significant decrease in the STD effects for the octyl chain. This is presumably because of the space it creates in the binding site for accommodating the long alkyl chain upon mutation. This indicates the interaction of ring A with aromatic residue Tyr50 is also somewhat important, although not as significant as other interactions in the combining site.

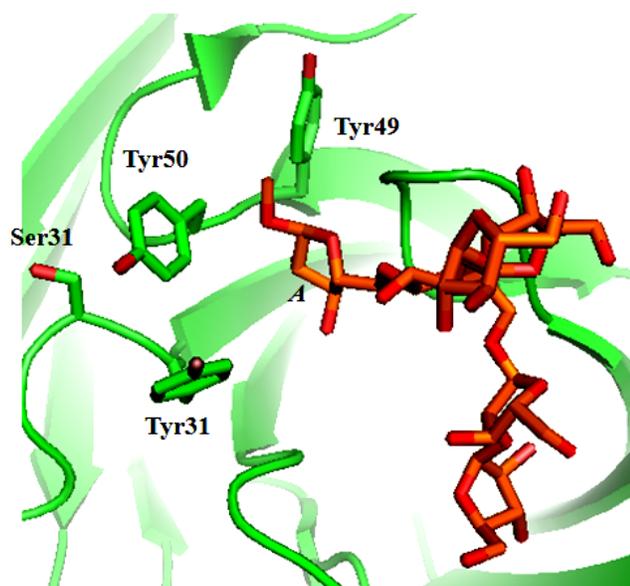


Figure 2.31 Position of residue Ser31, indirectly critical for binding of CS-35 F_{ab} to 2-2

2.4.4 Comparison of the Ara6–CS-35 System with Previous Pyranoside–Antibody Interaction Systems

A brief comparison of the Ara6–CS-35 furanoside–antibody interaction system with the previously reported bacterial carbohydrate–antibody systems can be useful. Most of the detailed studies on the bacterial carbohydrate–antibody interaction have reported the recognition of pyranosides by antibodies. The Se155-4 antibody interaction with the pyranoside antigens of the *Salmonella* cell wall^{5,22–26} is probably the most complete study of pyranoside–antibody binding. Since its crystal structure was reported as the first carbohydrate–antibody complex,²⁷ several pyranoside–antibody systems have highlighted the essence of aromatic residues for the recognition.^{28–32} Among these aromatic residues, Trp and His have been frequently reported as critical residues for binding. The results from our study are consistent with those data. Indeed, the positioning of Trp33 and His35 is surprisingly very similar in both systems,²⁷ shaping the specificity for abequeose and mannose residues in the Se155-4 system, and arabinofuranose ring E in the CS-35 system (**Figure 2.32**). The increased affinity of our system compared to the Se155-4–glycan interaction, is mainly reflected in the enhanced dissociation rate.

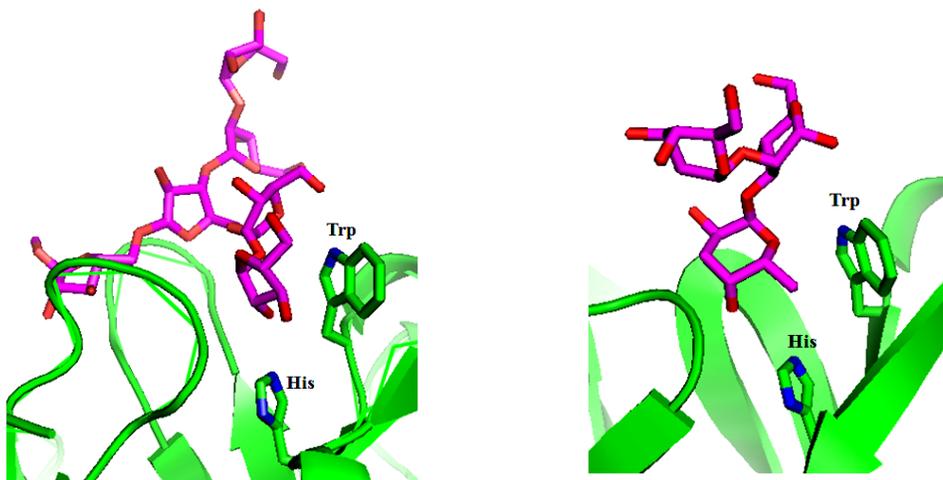


Figure 2.32 Similarity of the Trp and His residues interactions with the oligosaccharide antigens in the CS-35–Ara6 system (left), and Se155-4–trisaccharide system (right).

The role of aromatic–carbohydrate interactions in binding has been discussed for pyranosides.²¹ While aromatic rings stack against the more hydrophobic face of pyranosides, such surfaces are not present in many furanosides, in particular α -arabinofuranoside residues, in their low energy conformations. The binding of **3** to the CS-35 scFv however, revealed multiple critical CH– π bonds whose proper geometries and atomic distances were essential for binding. The importance of hydrogen bonds and water-mediated hydrogen bonds were also in agreement with pyranoside–antibody interactions. Ring E adopts a conformation in the binding site which is optimum to make three essential hydrogen bonds and two critical CH– π interactions. The unusual conformation of ring B is also a consequence of being involved in a vital CH– π bond with Tyr98, and positioning ring A and the C–E arm simultaneously in the binding site. It is interesting to see how the overall architecture of the antigen and the binding site are tailored to make the proper geometry for CH– π interactions. These findings demonstrate how the arabinofuranoside antigen positions each part of the epitope in the right place to interact with CS-35 antibody with a combination of different molecular forces.

2.5 Conclusions

Through the systematic evaluation of the recognition of an arabinofuranoside antigen (**2-3**) by the CS-35 antibody, we provide for the first time, a detailed picture of the motifs involved in a furanoside–protein interaction. To achieve this, we generated a scFv fragment of the CS-35 antibody, and quantitated its interaction with **2-3** by SPR spectroscopy. A library of mutants were rationally designed, generated and evaluated for the impact on the binding affinity. It was shown that the recognition relies heavily on the interactions between multiple aromatic residues of the antibody and the CH groups of the ligand, as well as essential hydrogen bonds, although in general some of the latter interactions are dispensable. Several aromatic residues were shown to dictate the binding and several hydrogen bonds tuned it.

References

- (1) O'Brien, R. J.; Nunn, P. P. *Am. J. Respir. Crit. Care Med.* **2001**, *163*, 1055.
- (2) Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T. *J. Am. Chem. Soc.* **2007**, *129*, 10489.
- (3) Galili, U. *Immunol. Cell Biol.* **2005**, *83*, 674.
- (4) Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *J. Mol. Biol.* **2009**, *392*, 381.
- (5) MacKenzie, C. R.; Hiram, T.; Deng, S.-j.; Bundle, D. R.; Narang, S. A.; Young, N. M. *J. Biol. Chem.* **1996**, *271*, 1527.
- (6) Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784.
- (7) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta -Proteins and Proteomics* **2005**, *1751*, 119
- (8) WoÈrn, A.; Pluckthun, A. *J. Mol. Biol.* **2001**, *305*, 989.
- (9) Biacore Manual, A. *Biacore AB, Uppsala, Sweden* **1997**.
- (10) Hudson, P. J.; Kortt, A. A. *J. Immunol. Methods* **1999**, *231*, 177.
- (11) Plückthun, A.; Pack, P. *Immunotechnology* **1997**, *3*, 83.

- (12) Miranker, A.; Robinson, C.; Radford, S.; Dobson, C. *FASEB J.* **1996**, *10*, 93.
- (13) J de Mel, V. S.; Doscher, M. S.; Martin, P. D.; Edwards, B. F. *FEBS Lett.* **1994**, *349*, 155.
- (14) Ahmad, Z. A.; Yeap, S. K.; Ali, A. M.; Ho, W. Y.; Alitheen, N. B. M.; Hamid, M. J. *Immunol. Res.* **2012**, *2012*, 980250.
- (15) Nieba, L.; Honegger, A.; Krebber, C.; Plückthun, A. *Protein Eng.* **1997**, *10*, 435.
- (16) Rheinhecker, M.; Hardt, C.; Ilag, L. L.; Kufer, P.; Gruber, R.; Hoess, A.; Lupas, A.; Rottenberger, C.; Plückthun, A.; Pack, P. *J. Immunol.* **1996**, *157*, 2989.
- (17) Proba, K.; WoÈrn, A.; Honegger, A.; PluÈckthun, A. *J. of Mol. Biol.* **1998**, *275*, 245.
- (18) Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/ π Interaction: Evidence, Nature, and Consequences*; John Wiley & Sons, 1998; Vol. 21.
- (19) Taha, H. A.; Castillo, N.; Sears, D. N.; Wasylshen, R. E.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theor. Comput.* **2009**, *6*, 212.
- (20) Richards, M. R.; Bai, Y.; Lowary, T. L. *Carb. Res.* **2013**, *374*, 103.
- (21) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. *Acc. Chem. Res.* **2012**, *46*, 946.
- (22) Bundle, D. R.; Baumann, H.; Brisson, J.-R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183.
- (23) Deng, S.; MacKenzie, C. R.; Sadowska, J.; Michniewicz, J.; Young, N. M.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1994**, *269*, 9533.
- (24) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423.
- (25) Sigurskjold, B.; Bundle, D. *J. Biol. Chem.* **1992**, *267*, 8371.
- (26) Pathiaseril, A.; Woods, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 331.
- (27) Cygler, M.; Rose, D. R.; Bundle, D. R. *Science* **1991**, *253*, 442.
- (28) Vulliez-Le Normand, B.; Saul, F.; Phalipon, A.; Belot, F.; Guerreiro, C.; Mulard, L. A.; Bentley, G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9976.
- (29) Vyas, N. K.; Vyas, M. N.; Chervenak, M. C.; Johnson, M. A.; Pinto, B. M.; Bundle, D. R.; Quioco, F. A. *Biochemistry* **2002**, *41*, 13575.
- (30) Rynkiewicz, M. J.; Lu, Z.; Hui, J. H.; Sharon, J.; Seaton, B. A. *Biochemistry* **2012**, *51*, 5684.

- (31) Lu, Z.; Rynkiewicz, M. J.; Yang, C. Y.; Madico, G.; Perkins, H. M.; Wang, Q.; Costello, C. E.; Zaia, J.; Seaton, B. A.; Sharon, J. *Immunology* **2013**, *140*, 374.
- (32) Rose, D.; Przybylska, M.; To, R.; Kayden, C.; Vorberg, E.; Young, N.; Bundle, D.; Oomen, R. *Protein Sci.* **1993**, *2*, 1106.

Chapter 3 : Characterization of the Binding of the CS-40 and 906.4321 Antibodies to the Ara6 Antigen

3.1 Introduction

Carbohydrate epitopes on bacterial, viral, and tumor cells are important disease markers and targets for diagnostic and therapeutic antibodies.¹ Antibody–antigen interaction systems can also serve as valuable models for studying general principles of protein–carbohydrate interactions. In Chapter 1, I explained that the interaction of arabinofuranose moieties in mycobacterial LAM with antibodies is a good model system for studying furanoside recognition by proteins. LAM gives rise to high levels of different antibodies in hosts that are infected with mycobacteria.² One of these antibodies is CS-35, whose interaction with Ara6 (**3-2**, **3-3**, **3-4**) (**Figure 3.1**) was the subject of Chapter 2 of the present thesis. Characterization of the interaction of other antibodies that recognize this antigen (**3-1**) would be beneficial in providing a comparison system for defining the specificity motifs through which antibodies recognize arabinofuranoside antigens in LAM. Two other such antibodies are CS-40³ and 906.4321.⁴ CS-40 was first raised against purified LAM of *Mycobacterium tuberculosis*,⁵ and the 906.4321 antibody was raised against *Mycobacterium leprae*.⁴ By using bovine serum albumin (BSA)-conjugate (**3-5**) (**Figure 3.1**), in ELISA experiments (unpublished data), it was shown that these two antibodies also recognize the Ara6 moiety of LAM.

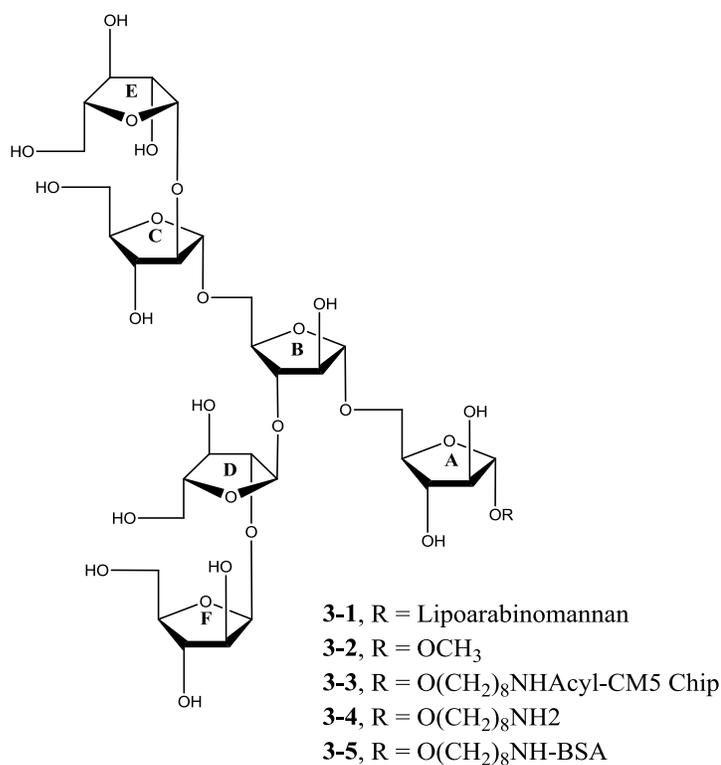


Figure 3.1 Ara6-containing structures found in LAM (**3-1**) and compounds used in binding studies with the CS-35, CS-40, and 906.4321 antibodies (**3-2** to **3-5**).

Except for these data, and some unpublished analysis of the binding of CS-40 and 906.4321 using electrospray mass spectrometry (J. S. Klassen, personal communication), there is no information about how these two antibodies interact with LAM or the Ara6 motif. Studying these interactions will help us not only to define affinity and specificity elements for each of these systems, but also to answer questions such as: How do different antibodies with different protein sequences bind to Ara6? What are the limits of the binding site and epitope in the specificity of such interaction?

Knowing the nature of the binding of these two other antibodies to Ara6 antigen, in combination with the data we have for CS-35–Ara6 interaction, could provide insights into the similarities and differences of the three interaction systems from different aspects, such as affinity of binding, kinetics, key amino acids involved in recognition, and the epitope of the antigen that is shared with each antibody. To date, there has been no crystal structure available for these two antibodies or their fragments. Indeed, efforts to crystallize the 906.4321 F_{ab} have

only recently (in the last 1–2 months, K. Ng personal communication) been successful. Moreover, efforts to generate a pure F_{ab} of CS-40 antibody have failed. Therefore, production of single chain fragment variable (scFv) for these two antibodies would be beneficial for providing the opportunity of single site mutations of the binding sites, and to overcome the difficulties of working with a full size antibody in binding analysis and structural studies. The success I had in studying the interaction of the CS-35 antibody with Ara6 by using its scFv as a tool encouraged us to extend the same strategy to these two other LAM-recognizing antibodies.

In this Chapter, I describe the production of scFv fragments of CS-40 and 906.4321 antibodies. A more detailed picture of the interaction between these antibodies and Ara6 is provided by surface plasmon resonance (SPR) spectroscopy, circular dichroism (CD), molecular modeling, and STD NMR spectroscopy investigations. The results will be compared to those obtained from the CS-35–Ara6 interaction system. The Ara6 antigens used in this chapter were provided by Dr. Maju Joe in the Department of Chemistry at the University of Alberta.

3.2 Materials and Methods

3.2.1 Construction of CS-40 and 906.4321 scFv Genes

The sequences of the CS-40 and 906.4321 variable domains were determined previously by our colleague in the University of Alberta Department of Chemistry, Mr. Gareth Lambkin. Light chain and heavy chain fragments were amplified by PCR from reverse transcription reaction products.⁶⁻⁸ These were then digested with *BfuI* to destroy any aberrant chain DNA that may have been inadvertently amplified. The digested PCR products were then gel purified and subcloned into pJET1.2, and multiple clones were sequenced in both the forward and reverse directions using primers that annealed to the vector backbone. The sequences of the variable fragments were used to construct the scFv gene. A linker containing 15 amino acids (Gly4Ser)₃ connecting the V_H and V_L domains was inserted between the C terminus of the heavy chain and the N terminus of the light chain for each construct. A His tag was added to the C-terminus of the sequence for the purification purposes. The ScFv sequences were harbored between NcoI and

HindIII restriction sites in the pET 28a vector and the synthetic genes were purchased from Genscript.⁹

The sequences of the CS-40 and 906.4321 scFvs are as follows; the linker sequence is shown in red:

CS40 scFv:

GEGSTKGSEVKVVESGGGLVQPGGSLRLSCASSGFTFTEYYMSWVRQPPGKALE
WLG FIRNKPNGYTTTEYSASVKGRFTISRDNQSILYLQMNTLRAEDSATYYCARHFPHY
AMDYWGQGT TTVTVSSAAANGGGGSGGGGSGGGGSGISTMDFQVQIFSLLISASVIMSR
GENVLAQSPAIMSATLGEKVTMSCRASSNVKYIYWYQQKSGASPKLWIYYTSNLAGV
PARFSGSGSGT SYSLTISSVEAEDAATYYCQQFTTSPSIFTFGTGTKLEIKRGGSTSGSHHH
HHH

906.4321 scFv:

GEGSTKGSQVQLQQSGAELVKPGASVKMSCKASGYSFTIYWITWMKQRPGQGLEWIGD
IYPGSVITNYNAKFKSRATLTVDASSSTAYMQLSSLTSEDSAVYYCARSGTNWHWWFD
VWGTGTTTVTVSSAAANGGGGSGGGGSGGGGSGISTMKLPVRLLVLMFWIPASSSDVLM
TQTPLSLPVS LGDQASISCRSSQNIVHSNGNTYLEWYLQKPGQSPKLLIYKVS YRFSGVP
DRFSGSGSGTDFTLKISRVEAEDLG VYYCFQGSHPWTFGGGGTKLEIKRGGSTSGSHHH
HHH

3.2.2 Transformation, Expression, and Refolding of scFv Fragments

Each of the pET-28a plasmids (200 ng) harboring the scFv genes of CS-40, and 906.4321 was transformed into *E. coli* BL21(DE3) host cells by heat shock. The same procedure as that used for the CS-35 scFv was used to express and refold these two scFvs. Cells were cultured in 50 mL of LB medium, followed by shaking at 37 °C for 8 h. To a 400 mL LB medium, 10 mL of this culture was transferred and it was shaken at 37 °C for 2 h. Expression was induced by adding 1M isopropyl β -D-1-thio-galactopyranoside and the culture was then shaken at 20 °C overnight. Cells were then pelleted at 8000 rpm for 30 min, and the pellet was dissolved in a

buffer containing Tris-HCl (50 mM pH 8), Triton X-100 (1%), and NaCl (150 mM), followed by cell disruption using a French press. Inclusion bodies were pelleted and washed six times successively with Tris-HCl (50 mM pH 8) and NaCl (200 mM) buffer. Inclusion bodies were then shaken at room temperature in buffer containing 4 mM guanidine, and 100 mM Tris-HCl pH 8 for 2 h. The solution was further injected into 0.5 M arginine, and 100 mM Tris-HCl buffer and was incubated in 4 °C for 48 h. Denatured proteins were then dialyzed against Tris-HCl (50 mM pH 8), NaCl (300 mM) buffer for 10 h to refold them.

3.2.3 Protein Purification and Characterization

Each sample containing the folded scFv of CS-40 or 906.4321 was purified by affinity chromatography using Ni-NTA Superflow Qiagen resin. The yield of the CS-40 scFv and 906.4321 scFv were 0.5 g/L and 1 g/L, respectively. ScFv fragments were then analyzed for monomer, dimer, and higher oligomer content by fast protein liquid chromatography (FPLC) using a Superdex 75 10/30 (GE Healthcare Life Sciences) column. Protein samples were eluted in ammonium acetate buffer (50 mM, pH 6.8).

3.2.4 Circular Dichroism (CD) Study of CS-40 and 906.4321 Antibody Fragments

CD spectra were obtained in phosphate buffer (50 mM, pH 7.2), using an Olis DSM 17 spectrometer at room temperature. An aliquot containing 200 ng of ligand **3-4** in phosphate buffer (50 mM, pH 7.2) was added to the samples to measure the far-UV and near-UV spectra of the complex in a 1mm and 1cm cell respectively. Samples were run at the Analytical and Instrumentation Lab of Chemistry Department by Dr. Wayne Moffat. CD spectroscopy in the far UV region was used to study the structural features of CS-40 (10 μ M), and 906.4321 (16 μ M) scFv fragments.

To see how the addition of the Ara6 ligand changed the behavior of aromatic amino acid side chains, near UV CD spectra were obtained for the CS-40 mAb (15 μ M) and 906.4321 F_{ab} (18 μ M). 200 ng of ligand **3-4** was added to each sample and the near UV CD spectrum was collected for the complex.

3.2.5 Binding of CS-40 and 906.4321 scFvs to Ara6 by Surface Plasmon Resonance (SPR) Spectroscopy

Binding affinities and kinetics were studied by SPR spectroscopy using a BIAcore 3000 (GE Healthcare Life Sciences) instrument at 25 °C. A CM5 chip was derivatized with Ara6 **3-4** (5 μM) by amine coupling using the following procedure. A freshly prepared aqueous solution of 0.5M NHS and 10 mM EDC was first injected on the chip surface to activate the carboxylate groups. Ligand **3-4** was then passed over the flow channel in a buffer containing 10 mM HEPES and 150 mM NaCl, pH 7.2 to generate the surface bound ligand **3-3**. A reference flow channel was also prepared by injecting 1M aqueous ethanolamine pH 8.5 after the activation of the surface with NHS/EDC. Fresh samples containing monomer scFvs (with the rate of 20 μL/min over 300 seconds of CS-40 and 906.4321) were injected to the flow channels in buffer with 10 mM HEPES containing 150 mM NaCl at pH 7.2. The CS-35 F_{ab} was used as a positive standard prior to measurements to confirm the stability of the chip, and the consistency and reproducibility of the tests done by BIAcore 3000. BSA was used as a negative control. Concentrations were measured several times by measuring the absorbance at 280 nm. Extinction coefficient of 906.4321 and CS-40 scFvs are 74370 and 53080, respectively.

The affinities were obtained from association and dissociation rates of the binding provided by the instrument, and were also calculated from equilibrium binding analysis according to Equation (3.1).

$$\frac{R_{eq}}{C} = K_A R_{max} - K_A R_{eq} \quad (3.1)$$

Where R_{eq} is the equilibrium response units, R_{max} is the resonance signal at saturation, C is the concentration of the protein, and K_A is the association constant. A plot of R_{eq}/C versus R_{eq} has a slope of $-K_A$. To obtain K_A and R_{max} from the data, we performed a first-degree polynomial regression using the “polyfit” command in MATLAB.

3.2.6 Saturation Transfer Difference NMR Spectroscopy

The STD NMR spectroscopic experiment for the 906.4321 F_{ab} was recorded on a Varian Inova 600 MHz spectrometer. The protein was lyophilized from ammonium acetate buffer (50 mM, pH 6.8) and then dissolved in deuterated ammonium acetate (50 mM pH 6.8), at a concentration of 19 μ M. Hexasaccharide **3-4** was added to this sample at a final concentration of 0.5 mM. The measurement was done in a 3 mm NMR tube. 1D STD experiments were performed at 298 K with saturation at 0.5 ppm, with a 2 ms delay, and with water suppression. A ¹H NMR spectrum of **3-4** and was recorded as reference.

3.2.7 Modeling the Binding Sites of CS-40 and 906.4321 Antibodies

A homology model for the variable domains of CS-40 and 906.4321 antibodies (residues 1–377) was generated using the “Automatic Modeling” mode of SWISS-MODEL with default parameters.^{10,11} The crystal structure of CS-35 (PDB ID: 3HNS)¹² was used as the template for the investigation of the common features in the binding sites of the three antibodies, and the difference in the specificity of CS-35, CS-40, and 906.4321 antibodies. Sequence similarities of each variable chain to that of the CS-35 sequence is as follows:

CS-40 V_H: 44%

CS-40 V_L: 54%

906.4321 V_H: 56%

906.4321 V_L: 67%

Ara6 **3-2** from the CS-35 crystal structure (PDB Code: 3HNS) was posed in the binding site of each protein. structures were further inspected as described in more detail in the Results section.

3.3 Results

3.3.1 Generation of CS-40 and 906.4321 scFvs

The scFvs of the CS-40 and 906.4321 antibodies were expressed as inclusion bodies in the cytoplasm of *E. coli* host cells. The purity of the scFv sample obtained from inclusion bodies was high. The size and purity of each fragment were confirmed by SDS PAGE (**Figure 3.2**).

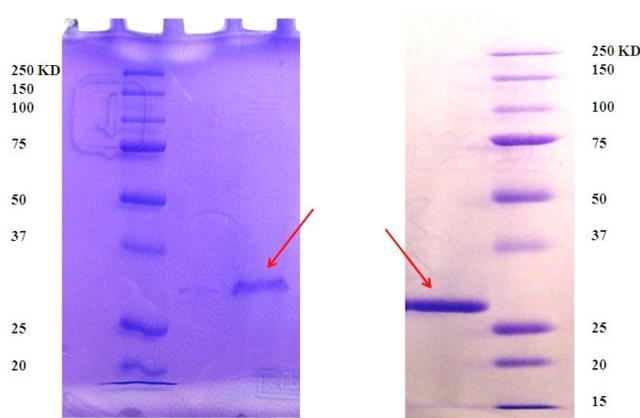


Figure 3.2 SDS PAGE results for the CS-40 scFv (Left), and the 906.4321 scFv (Right)

Size exclusion chromatography was further applied to purify monomer scFv from dimeric and oligomeric species. The monomer scFv peak, which shows up at 18 mL, was collected for each sample. The 906.4321 scFv showed a combination of peaks. However, the ratio of CS-40 monomer to dimer was shown to be highly concentration dependant. At the same concentration of the injection as the 906.4321 scFv (1 mg/mL), the CS-40 scFv was mainly a large oligomer with almost no trace of monomer or dimer. At more dilute concentrations, the ratio of monomer and dimer improved (**Figure 3.3**). This is consistent with the behaviour of the CS-35 scFv, which we described in the previous chapter. In all cases, a lower concentration of the scFv sample shifts the equilibrium from the oligomeric state (dimer or larger oligomers) to the monomer fragment. The purified monomer samples were stable in ammonium acetate buffer for 2–3 weeks at 4 °C.

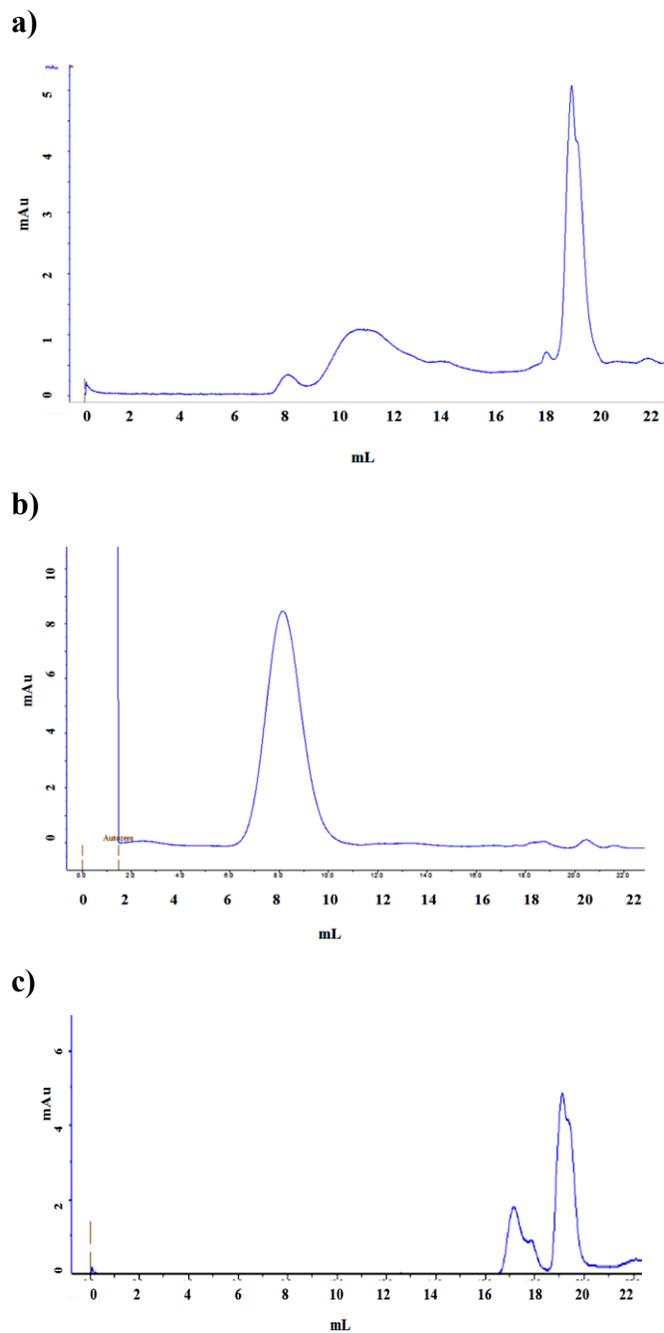


Figure 3.3 FPLC results for 906.4321 scFv **(a)**, CS-40 oligomer **(b)**, and CS-40 scFv monomer **(c)**; the peak starting at 18 mL belongs to the corresponding monomer scFv in each sample. Those at 8 mL correspond to the larger oligomers.

3.3.2 Circular Dichroism Study of Antibody Fragments

Study of the both CS-40 and 906.4321 scFv fragments in far UV region showed positive peaks at 195 nm and negative bands at 218 nm, which is the typical absorption for β sheets in antibody secondary structures.¹³ This confirmed the proper folding of scFvs (

Figure 3.4).¹⁴ Addition of the ligand **3-4** to the proteins did not change the secondary structure of these proteins.

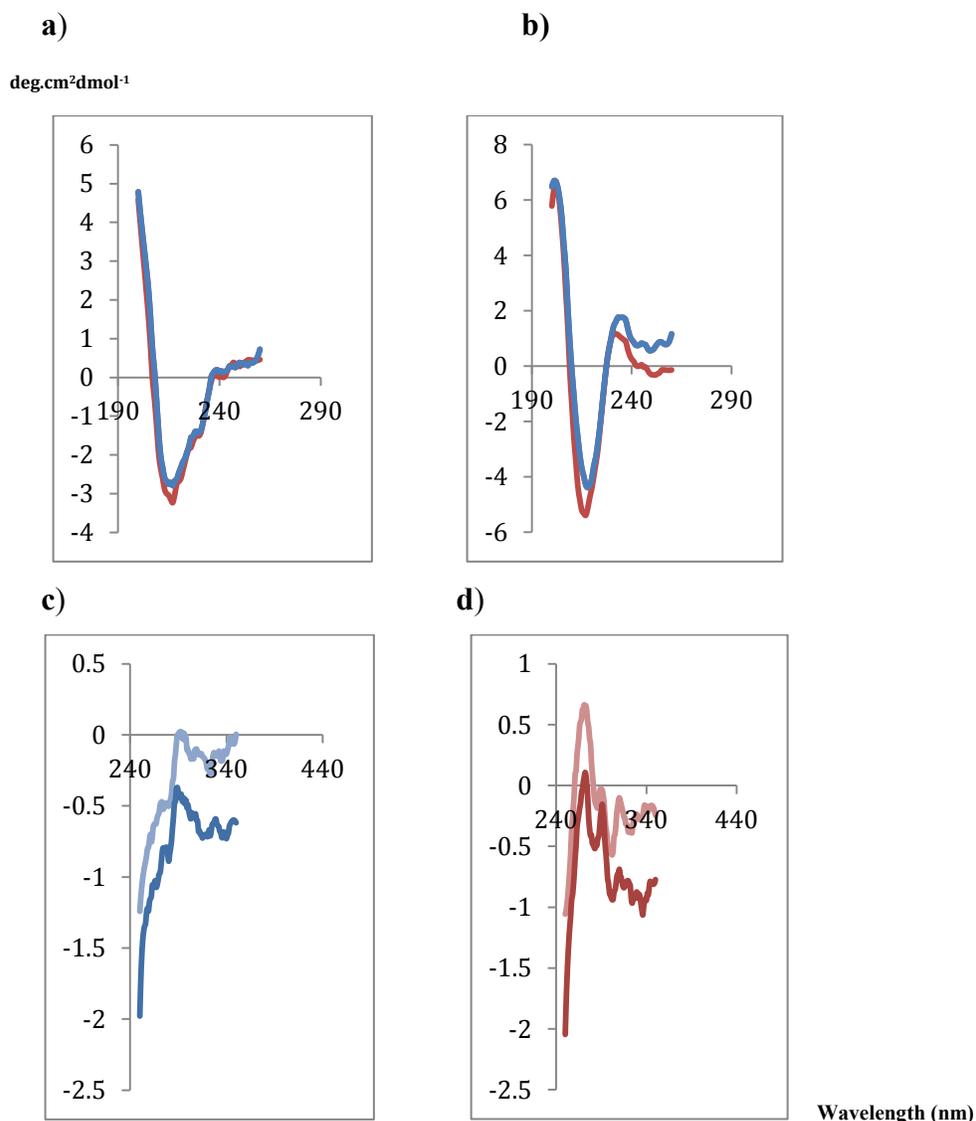


Figure 3.4 Far UV circular dichroism results with CS-40 (a) and 906.4321 (b) scFv fragments. with (blue) and without (red) Ara6. Near UV CD spectroscopic results with CS-40 mAb (c) and 906.4321 F_{ab} (d) with (pale) and without (dark) Ara6. The ScFv fragments were used in the far UV experiments. The F_{ab} and mAb were used for 906.4321 and CS-40, respectively, in the near UV region.

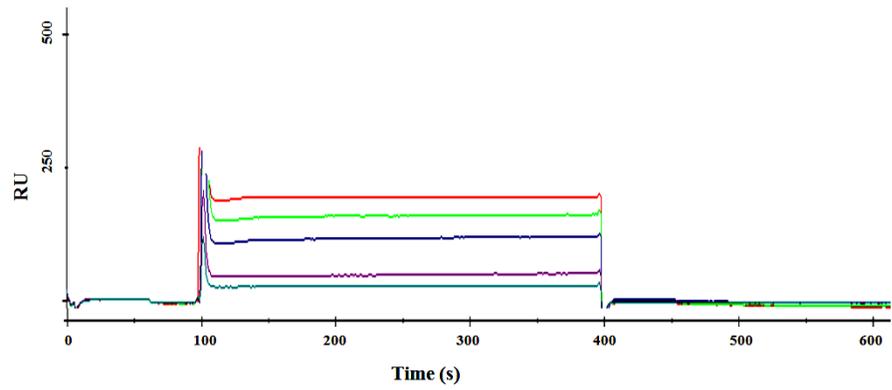
The near UV CD spectrum was used to investigate the possible binding of aromatic side chain residues to Ara6. The comparison of the bound and unbound antibody CD spectra in this region showed that both the CS-40 and 906.4321 antibodies shift to a higher ellipticity upon the addition of **3-4** (

Figure 3.4). This implies the involvement of aromatic residues in the recognition of Ara6 by CS-40 and 906.4321.¹⁴ This is consistent with the results previously observed for the CS-35 antibody, where Trp, Phe, and Tyr residues were contributing to the specificity of the binding. It is also consistent with the previous reports for pyranoside–antibody recognition systems, which were overviewed in the introduction chapter, highlighting the contribution of aromatic residues to these recognition events.

3.3.3 Analysis of the binding of CS-40 and 906.4321 scFv by SPR Spectroscopy

The SPR spectroscopic results for CS-40 scFv and 906.4321 scFv confirmed the previous ELISA data that both fragments were able to bind to Ara6. The binding profiles of both scFvs were then analyzed. Both fragments showed to fit to the 1:1 binding model in BIAevaluation analysis.¹⁵ **Figure 3.5** and **Figure 3.6** summarize the results of the binding of Ara6 immobilized on a CM5 chip to the CS-40 and 906.4321 scFv fragments, respectively. The observed association and dissociation rates, and association constants, are shown in **Table 3.1**.

a)



b)

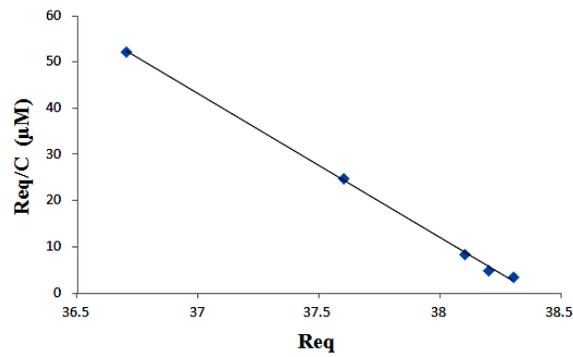


Figure 3.5 a) SPR curves for the binding of CS-40 (Concentrations: 10.7—, 7.6—, 4.5—, 1.5—, and 0.7— μM) to Ara6 immobilized on a CM5 Chip **b)** Equilibrium binding analysis

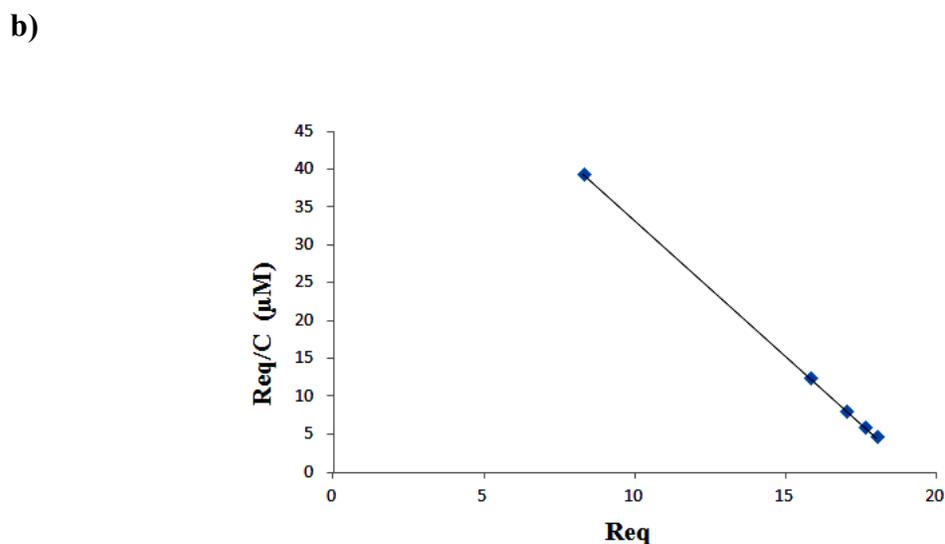
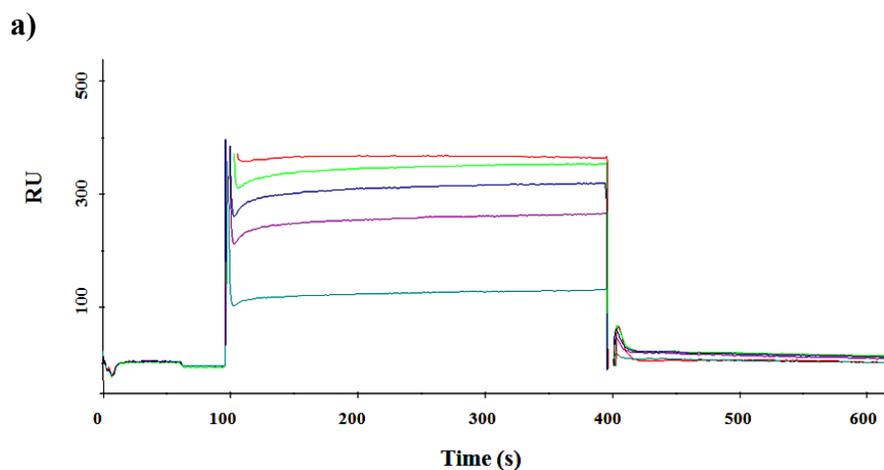


Figure 3.6 a) SPR curves for the binding of 906.4321 (Concentrations: 3.7—, 2.9—, 2.1—, 1.26—, and 0.21— μM) to Ara6 immobilized on a CM5 Chip. **b)** Equilibrium binding analysis

Table 3.1 Binding constants and association and dissociation rates of CS-40 and 906.4321 scFvs to **3** at determined by SPR spectroscopy. Affinities were calculated from association and dissociation rates, and also from equilibrium binding analysis

scFv	K_D (M) k_d/k_a	K_D (M) Equilibrium	k_a ($\text{M}^{-1} \text{s}^{-1}$)	k_d (s^{-1})
CS-40	8.9×10^{-8}	$8.89 \pm 0.12 \times 10^7$	1.04×10^3	3.32×10^{-5}
906.4321	3.57×10^{-7}	$3.58 \pm 0.54 \times 10^6$	4.95×10^3	1.77×10^{-3}

3.3.4 STD NMR Analysis of the Ara6 Epitope in Binding to 906.4321 F_{ab}

In order to investigate the epitope involved in binding to 906.4321 antibody, the STD NMR experiment was carried out with 90.4321 F_{ab} and **3-4**. The STD NMR spectrum of 906.4321 F_{ab} in combination with **3-4** is shown in **Figure 3.7**.

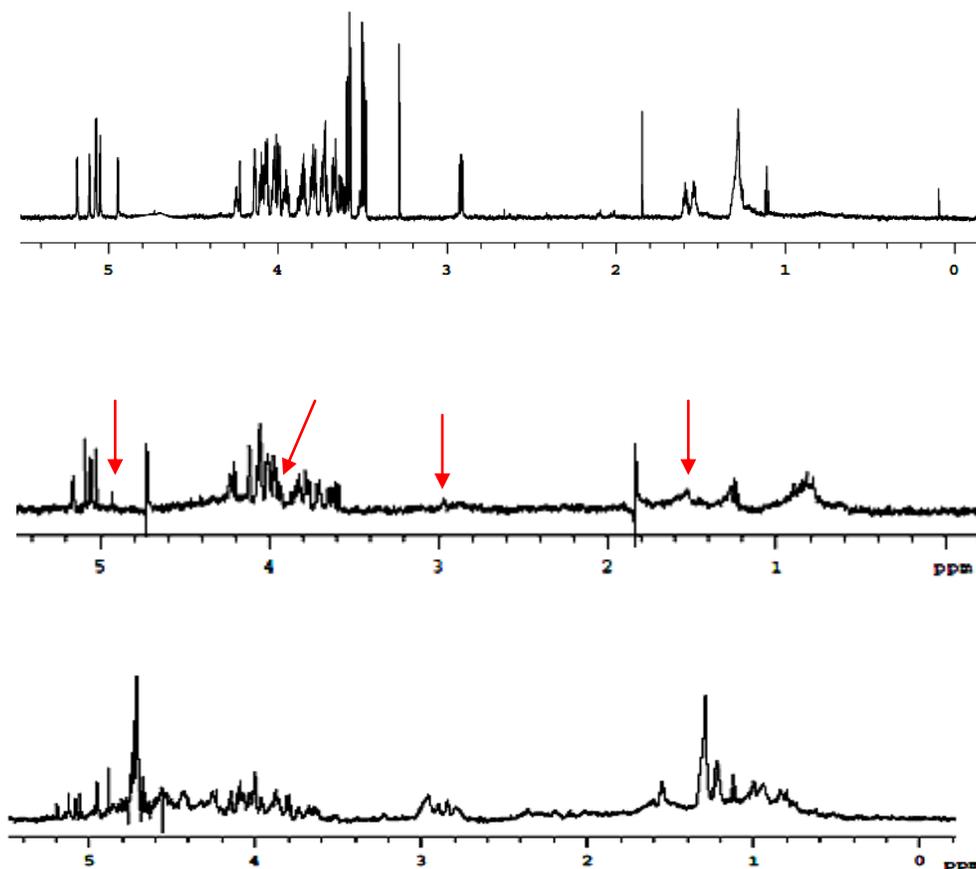


Figure 3.7 ^1H 1D NMR spectrum of **3-4** (top), and the STD NMR spectrum in the presence of the 906.4321 F_{ab} (middle). The diminished signals for ring A and octyl chain are shown with arrows. An STD NMR experiment with the CS-35 F_{ab} and **3-4** is also shown (bottom). The extra peaks below 1 ppm belong to residual protein signals.

Comparison of the STD spectrum with the 1D ^1H NMR spectrum of **3-4** and with STD spectrum of CS-35 F_{ab} in the presence of **3-4** revealed that among all rings, the intensities of the proton signals of ring A in Ara6 **3-4** is decreased the most upon binding to the 906.4321 F_{ab}

(**Figure 3.7**). Also, the decrease in the intensity of the signals for the octyl chain is evident compared to the 1D spectrum of **3-4** and also to the CS-35 STD spectrum with **3-4**. Ring A had shown to be an essential part of the Ara6 epitope to be recognized by CS-35 antibody.¹⁶ The fact that it is almost not involved in the binding of Ara6 to 906.4321 implies that 906.4321 chooses a different epitope of Ara6 for binding. I attempted STD NMR experiment for CS-40 mAb, but the stability of these proteins in the concentrations necessary for the experiment was not sufficient. CS-40 scFv was also produced in low yield far from the concentrations required for STD-NMR experiments.

3.3.5 ScFv Binding Site Models

Homology models of the 906.4321 and CS-40 scFvs provide information about the structural aspects of the binding of Ara6 by these proteins in the absence of crystal structures. Experimental support is always desired to validate the built models. **Figure 3.8** shows the surface of the 906.4321 scFv. Hexasaccharide **3-2**, in the conformation it adopts in the CS-35–**3-2** crystal structure was posed into the binding site. The 906.4321 binding site was shown to be a smaller groove compared to the CS-35 binding site. In particular, at the region where ring B interacts with the protein, the shape of the CDR H3 domain makes the binding site tighter than that of CS-35, so that the CDR H3 would clash with ring B and D in that conformation. Consequently, **3-2** would not be accommodated in the same way in 906.4321 binding site as in the CS-35 binding site. The shape of the groove also deviates from CS-35 mainly in the region of CDR L1 and H3 (**Figure 3.9**). The discrepancies in these two domains have been confirmed in the recent crystal structure of 906.4321 in combination with Ara6 (unpublished data). The Ara6 moiety, therefore, should adopt a different conformation to be accommodated in the smaller binding site of the 906.4321 antibody. This appears to exclude ring A from the binding site based on the STD NMR results and the homology model.

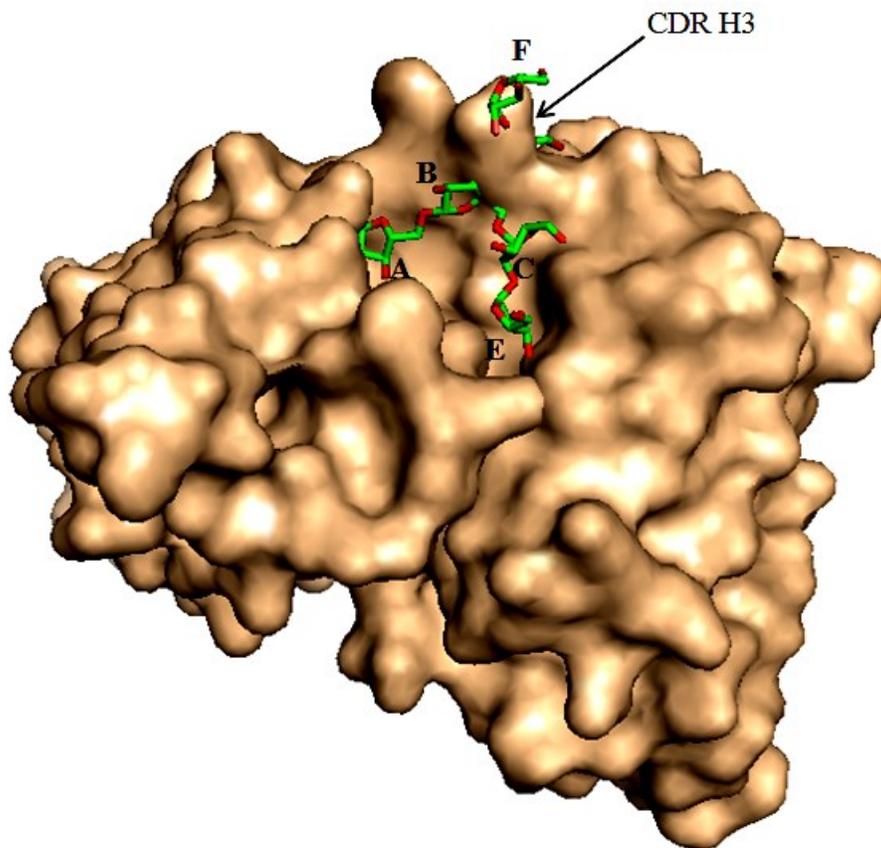


Figure 3.8 Homology model of the 906.4321 scFv. Ara6 was obtained from CS-35 crystal structure (PDB:3HNS) and was posed in the binding site. CDR H3 obviously provides smaller room to Ara6 at the position of ring B. Therefore, the Ara6 should adopt a different bound conformation. Ring A will be in the least contact with the binding site in such a conformation based on the STD NMR results.

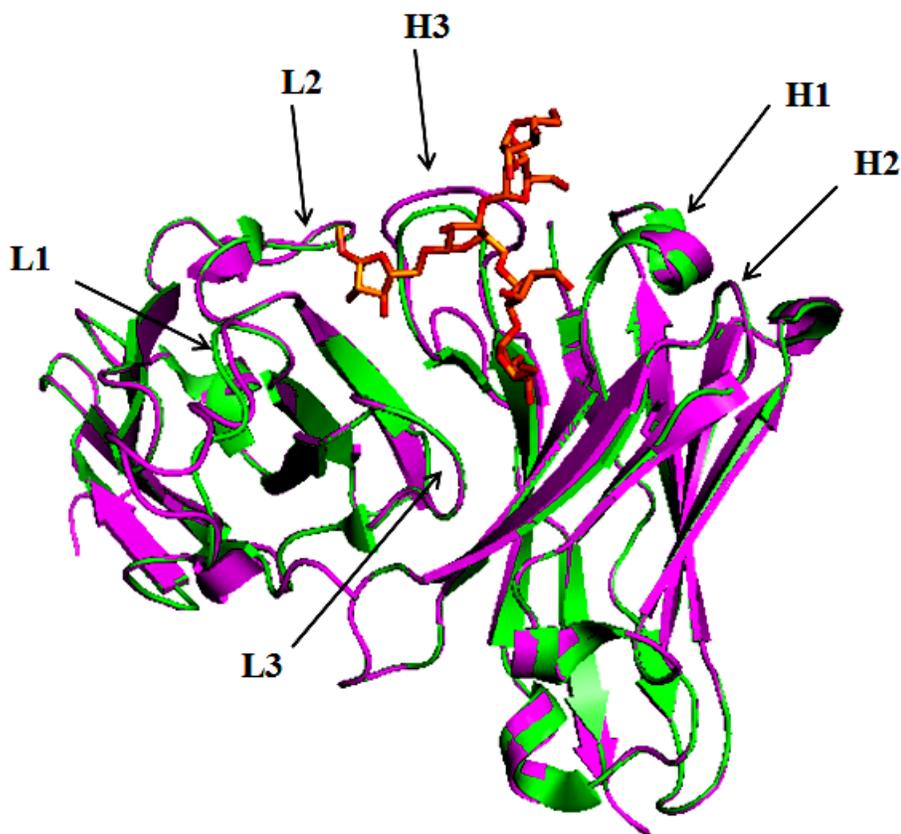


Figure 3.9 Superposition of 906.4321 (magenta, homology model) and CS-35 (green, X-ray structure) variable domains.

The amino acid composition of the 906.4321 variable domains in the model was also shown to be rich in aromatic residues, especially Trp (**Figure 3.10**). We previously (Chapter 1) highlighted the importance of Trp and His in the binding of antibodies to pyranosides. This system contains three Trp residues and one His residue close to the Ara6 ligand.

In many cases, CDR H3, which is at the center of the binding site, is the most important domain in antigen recognition. It often shows the most diversity in length, sequence and structure among the six CDRs of antibodies.^{17,18} This domain is larger in 906.4321, having 11 residues both in the model and crystal structure, compared to nine residues in CS-35 and CS-40. This loop is rich in aromatic amino acids Trp and Phe.

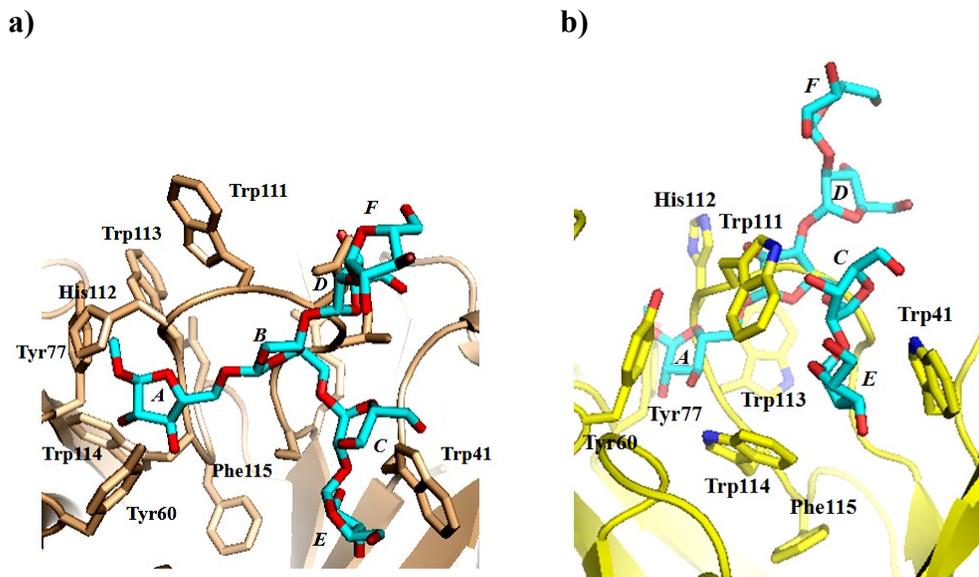


Figure 3.10 Aromatic residues in the 906.4321 binding site model (a) and crystal structure (b) Ara6 was posed from the CS-35 crystal structure (PDB Code: 3HNS)

Superposition of the 906.4321 binding site model and crystal structure on CS-35 crystal structure revealed the residues that are positioned similarly in the two systems (**Figure 3.11**). Among these similar residues, Trp33 and Asn58 of CS-35 have identical counterparts in 906.4321 (Trp41 and Asn67). Moreover, at position Ser31 in CS-35, there is a Thr residue in 906.4321. This Ser residue is not in direct contact with the ligand, but was shown (Chapter 2) to be important for binding to Ara6.

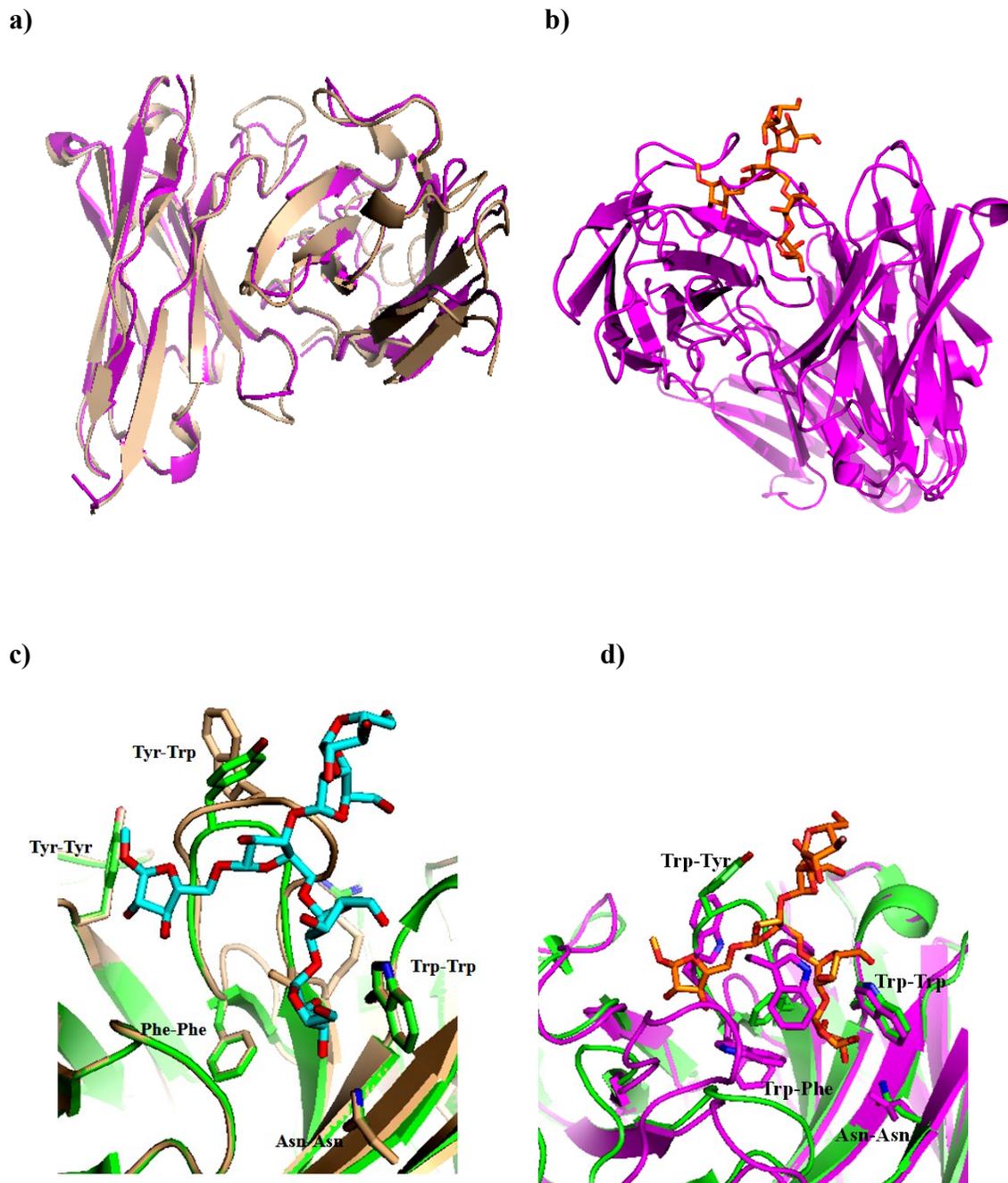


Figure 3.11 **a)** Superposition of 906.4321 (tint) model and crystal structure (magenta). **b)** Ara6 from CS-35 crystal structure posed in the 906.4321 binding site crystal structure. **c)** Similar residues in 906.4321 binding site model (wheat tints) **d)** and 906.4321 crystal structure (magenta), in comparison to the CS-35 (green) antibody

The model was successful in the prediction of loop size, and the position of some residues close to ring B to E. However, the conformation of some residues in the flexible CDR H3 loop

deviates between the model and the crystal structure (**Figure 3.11**). The CDR H3 showed a significant number of aromatic residues, in particular Trp. To investigate the role of these residues in the specificity of 906.4321, single site mutations are suggested.

Inspection of the variable domains of CS-40 in the generated model also revealed a groove in the binding site (**Figure 3.12**). CDR L1, L3 and H2 showed some deviations from those in CS-35. The shape of the CDR H2 creates a small cavity close to rings C and E, in a conformation where rings D and F extend out of the binding site. But there seems to be enough space for rings A and B, similar to CS-35 binding site. Therefore, it is predicted that the Ara6 epitope will interact differently in the region of rings C and E compared to CS-35.

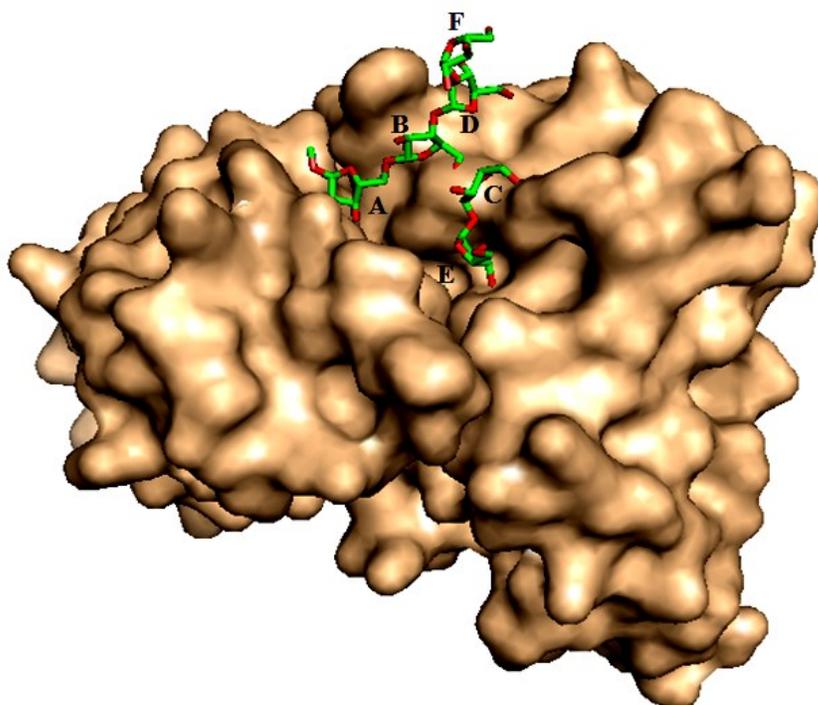


Figure 3.12 Model of CS-40 scFv. The Ara6 was obtained from the CS-35 crystal structure (PDB:3HNS) and was posed in the binding site. Rings A, B, D, and F are in contact with the binding site. The binding site obviously provides less space to the ligand at the connection between rings B and C, this is at the CDR H2 domain. Therefore, the binding from rings C and E arm could deviate from CS-35.

The superposition of CS-40 and CS-35 variable domains showed that CDR H3 has high similarity with regard to size and shape in the two antibodies (**Figure 3.13**). In contrast, CDRs L1, L3, and H2 are different.

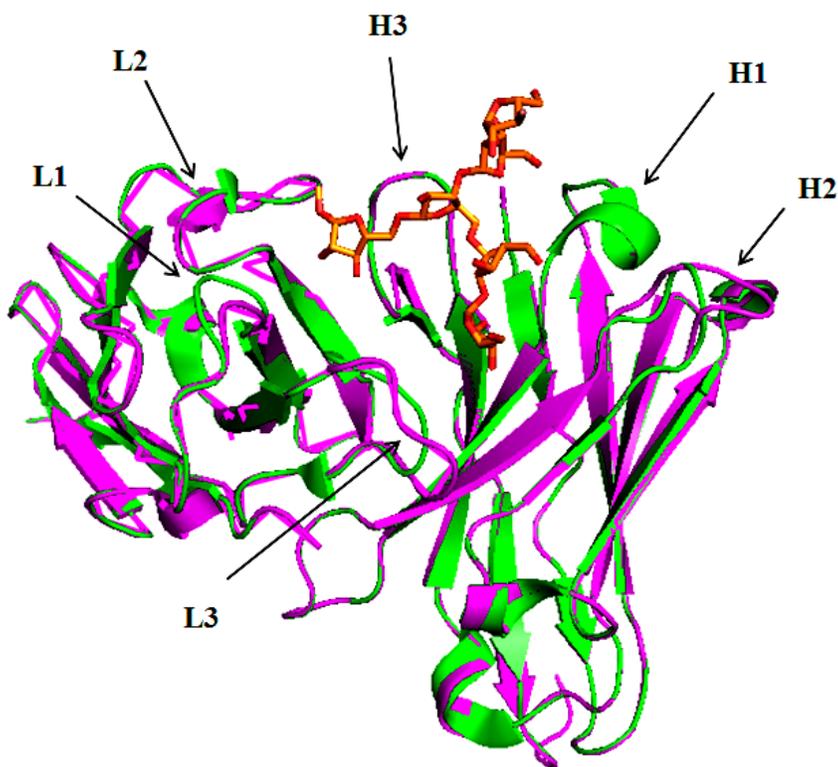


Figure 3.13 CS-40 (magenta) superposition with CS-35 (green) variable domains. Deviations in CDRs L1, L3, and H2 are obvious.

Inspection of the CS-40 binding site structure revealed a binding site with a significant number of aromatic amino acid residues close to the ligand. The binding site contains several aromatic residues in very close contact to the Ara6 ligand (**Figure 3.14**). The CDR H3 loop in CS-40 contains nine residues, which is the same size as this loop in CS-35. This loop has five aromatic residues including a His, two Phe and two Tyr residues. The positioning of Phe112 in this loop corresponds to Tyr98 in CS-35, in close proximity to ring B of Ara6.

The CDR H3 segment of Phe110–Pro111–Phe112–Tyr113 in CS-40 is similar to Phe95–Gly96–Asn97–Tyr98 in CS-35. Actually it was shown in the CS-35 F_{ab} crystal structure, that the

bond of Phe95 with Gly96 is *cis*.¹² It can be predicted that the peptide bond of Phe with Pro is also a *cis* bond in CS-40. Proline residues are observed more frequently than other amino acids in *cis* bonds with adjacent residues.¹⁹

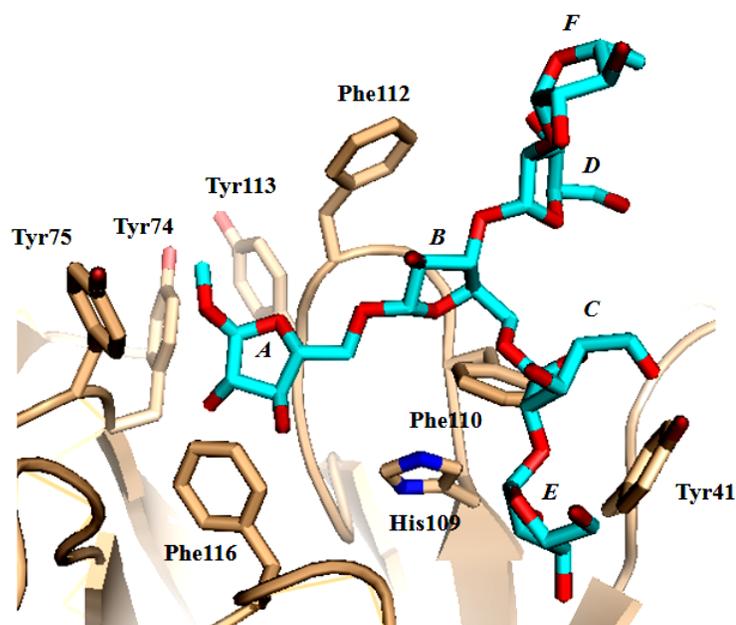


Figure 3.14 Aromatic residues in the binding site of CS-40 antibody model. Ara6 from crystal structure of the complex with CS-35 (PDB:3HNS) is added to the structure.

In addition, several residues appear to be similar in the two structures, including important residues such as Phe95, Tyr98 and Tyr50 in CS-35. In CS-40, these positions correspond to His109, Phe112, and Tyr75, and they seem to have the same type of interactions with the ligand (**Figure 3.16**) as they do in CS-35.

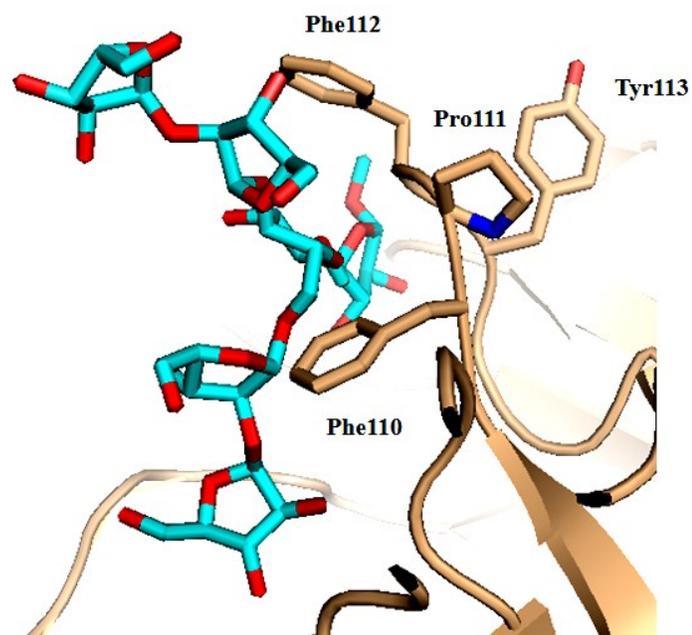


Figure 3.15 Phe–Pro–Phe–Tyr sequence in CDR H3 domain of CS-40

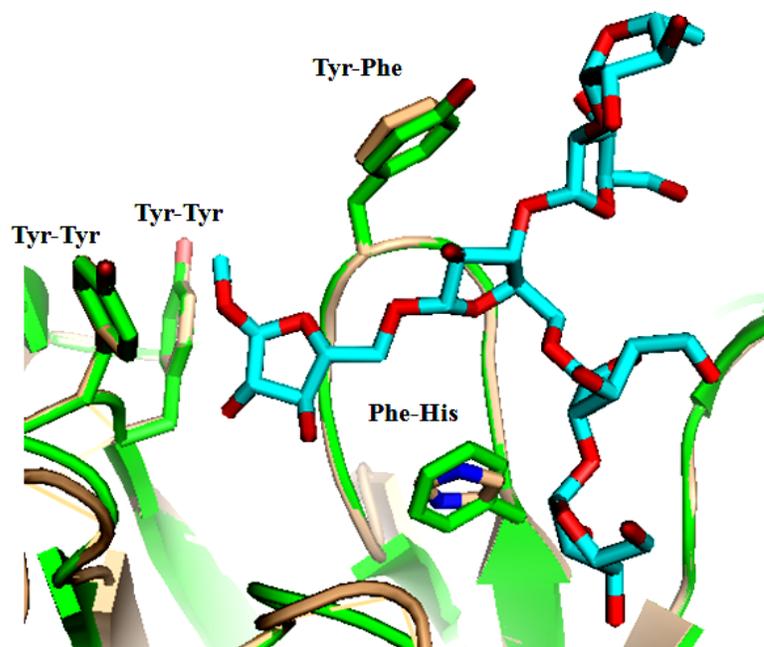


Figure 3.16 Similar residues in CS-40 in comparison to the CS-35 (green) antibody.

3.4 Discussion

The approach pursued in the characterization of the CS-35–Ara6 binding interaction using the scFv of CS-35 antibody was a successful case-study of a furanoside–protein binding system (Chapter 2). The results from that investigation, in addition to the paucity of information about furanoside–protein recognition events, motivated us to extend the same strategy to study the interaction of two other LAM recognizing antibodies, CS-40 and 906.4321, with Ara6. The binding affinity of CS-40 to ManLAM has been reported in to be in the micromolar range ($K_D = 3 \times 10^{-7}$ M).²⁰ However, information about its interaction with arabinofuranose motifs of LAM, which have been shown to be the antigenic epitope, was not available. In this chapter, I described my efforts to build a picture of the interaction of CS-40 and 906.4321 with Ara6.

We considered the scFv technology especially valuable, because attempts to produce a F_{ab} of CS-40 have been unsuccessful. Working with a monoclonal antibody has also its limitations and disadvantages, as was highlighted in Chapter 1. Moreover, the development of scFvs for both CS-40 and 906.4321 antibodies provides the opportunity to manipulate their binding site by mutation, the approach that was useful in characterizing the CS-35–Ara6 interaction system.

As described in the previous chapter for CS-35, we expressed both new fragments by inclusion bodies in *E. coli* system. Refolding of the inclusion bodies was successfully done by affinity column purification of the scFv fragments. The yield of the 906.4321 scFv was close to that of CS-35 scFv, and higher than the CS-40 scFv. The CS-40 scFv had a high tendency to oligomerize and precipitate in concentrations lower than that required for the oligomerization of the CS-35 and 906.4321 scFvs. This implies that the interface of the two variable fragments in CS-40 is not stabilized as efficiently as the two other antibodies. The dimerization of scFv fragments has been discussed extensively.²¹⁻²³ The factors that influence the equilibrium between dimer to monomer include pH, concentration, and temperature. A future direction of CS-40 scFv research could be the optimization of these factors to make the equilibrium shift to more monomer. Optimization of the interface of the two variable fragments by some mutations could also stabilize the interface, leading to higher stability of the monomer fragment, and possibly higher yield of the production. This will result in the production of sufficient quantities of the protein for detailed structural and binding studies.

Both the CS-40 and 906.4321 scFvs were produced with high purity, as determined by SDS PAGE, and were successfully purified as monomer fragments by FPLC. Both fragments were properly folded, based on a far UV CD spectroscopic experiment. The study of the two antibodies in the near UV region demonstrated an increase in ellipticity after adding the ligand, which implies the involvement of aromatic side chain residues in the binding.

The behavior of these two fragments in binding to Ara6 was then analyzed by SPR spectroscopy. The binding of the ligand to both proteins could be fit to a single site binding model. The association rates and dissociation rates, as well as affinities are summarized in comparison with those of CS-35 scFv in **Table 3.2**. A future direction could be to corroborate these data with other binding techniques such as mass spectrometry binding analysis of the scFv fragments to Ara6-based ligands.

Table 3.2 Comparison of the binding profiles of three scFvs (CS-40, 906.4321, and CS-35) to Ara6 antigen from SPR spectroscopic results.

Antibody scFv	K_D (M) k_d/k_a	k_a (M ⁻¹ s ⁻¹)	k_d (s ⁻¹)
CS-40	3.1×10^{-8}	1.04×10^3	3.32×10^{-5}
906.4321	3.5×10^{-7}	4.95×10^3	1.77×10^{-3}
CS-35	0.5×10^{-7}	3.86×10^3	3.56×10^{-3}

Binding analysis of the fragments using SPR spectroscopic data provided a basis for comparative study of the binding interactions. The results showed that the binding profile of 906.4321 scFv is very similar to that of the CS-35 scFv. In contrast, the CS-40 scFv revealed slower dissociation rate, which translates into enhanced affinity by one order of magnitude in comparison to the CS-35 scFv. Other measurement techniques such as ESI mass spectrometry can be useful to assess the binding behavior of the proteins in ammonium acetate buffer in gas phase.

Because there is no crystal structure available for these two antibodies, or their F_{ab} fragments (aside from a very recently obtained one obtained by K. Ng at the University of Calgary), in silico techniques offer an attractive alternative for studying these interactions.¹ Therefore, homology models were generated for the CS-40 and 906.4321 scFvs to define and

explain the similarities and differences in comparison to the CS-35 paratope. Moreover, in the absence of crystal structures, these models serve as a guide to choose particular residues that are close to Ara6 for mutation studies.

To find an explanation for the increased binding in CS-40 compared to CS-35, I checked the CS-35 mutants profile and the CS-40 model. In the mutation studies of the CS-35 system, I showed that the Asn34Asp mutant had a slower dissociation rate. We proposed that the hydrogen bond at this residue is critical for recognition; mutation to Ala destroys the binding. At the position of this residue in CS-40, there is Tyr59 (**Figure 3.13**), extending its OH group in close proximity to the region that ring A of Ara6 occupies, while the same position in CS-35 and 906.4321 has an Asn or a Glu residue, respectively. The Tyr59 in CS-40, which is adjacent to ring A of the ligand, is likely to make more efficient interactions with the ligand both by hydrogen bonding and stacking interactions with ring A. The same position in the CS-35 and 906.4321 antibodies have an Asn or Asp residue, respectively, which are non-aromatic residues, and which are positioned farther from the ligand. In the X-ray structure of CS-35 in complex with **3-2**, the 2-OH group of Ring A of the ligand was shown to be in a water mediated hydrogen bond with Asn34. This OH is actually closer to Tyr59 in CS-40, being 3.07 Å apart. Although the aromatic ring of Tyr 59 might affect the conformation of ring A, it is well positioned to make stacking interaction and direct hydrogen bond in a good distance with the critical OH group of ring A. The contribution of the Tyr59 residue to the improved binding profile of CS-40 could be tested by a single site mutation of this residue to Ala or Asn in the CS-40 scFv as a future direction.

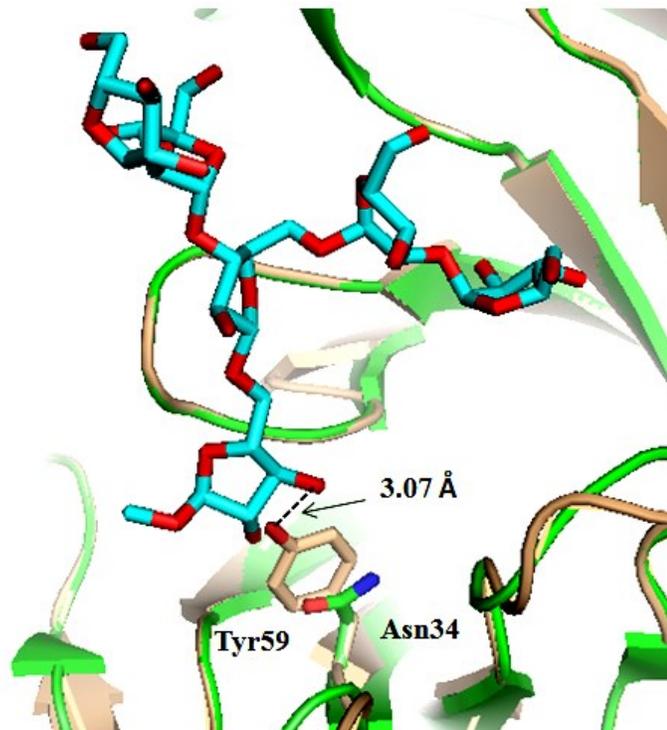


Figure 3.17 Asn34 of CS-35 (green) is substituted by Tyr 59 in CS-40, extending its OH group in close proximity to ring A, and stacking against ring A.

Inspection of the CS-40 and 906.4321 scFv models revealed the topography of the binding sites. There were deviations in some domains compared to CS-35. The paratope was then checked for similarity in the residues that were shown to be critical in the binding of CS-35 antibody to Ara6. Some of the key amino acids were conserved in these antibodies and some other had counterparts that seem to be able to make the same molecular interactions with the ligand. These are summarized in

Table 3.3. An interesting aspect was the observation of a significant number of aromatic residues in the binding sites, some of which are conserved between each of these two antibodies and CS-35. Rings D and F seem to be in more contact with these two antibodies compared to CS-35. However, more experimental evidence is required to describe the epitope.

Table 3.3 Comparison of similar amino acids involved in Ara6 binding in CS-40 model, 906.4321 crystal structure in comparison to CS-35 crystal structure.

Antibody	Rings in Ara6		
	Ring A	Ring B	Ring C/E
906.4321	Tyr77/L*	Trp111/H	Trp41/H, Asn67/H
CS-35	Tyr50/L	Tyr98/H	Trp33/H, Asn58/H
CS-40	Tyr99/L	Phe112/H, His109/H	Tyr41/H
CS-35	Tyr50/L	Tyr98/H, Phe95/H	Trp33/H

* Ring A extends out from the binding site of 906.4321

The 906.4321 binding site is particularly rich in Trp and His. In this aspect it resembles pyranoside–antibody interaction counterparts that were discussed in Chapter 1, more than the other two antibodies, CS-35 and CS-40. It is interesting that all three antibodies recognize Ara6 with a high number of aromatic residues. This implies that a substantial number of CH– π interactions with Ara6 may be conserved in the three systems. Mutation studies to prove this would be a good future direction to complement the picture of these interactions.

Among hydrogen bond-making residues, Asn67 is the most evident residue that is conserved in 906.4321 compared to CS-35. It is likely that this residue forms a hydrogen bond with ring E in the binding site of 906.4321 antibody. Also, His 109 in CS-40 binding site can be predicted to make hydrogen bond with ring C or ring E of the ligand. However, the binding sites of both 906.4321 and CS-40 antibodies are significantly occupied by hydrophobic residues. Also, it is possible that the microarchitecture of other hydrogen bonds are variable in the three systems, considering the large number of OH groups in the ligand, and the freedom of rotation for these functional groups to form different networks of hydrogen bonds with the different antibodies. Such predictions about hydrogen bonds would be supported by performing energy minimization of the ligand, in the absence of a crystal structure that defines the conformation of the ligand in the binding site.

The crystal structure of 906.4321 provided the opportunity to compare the results from modeling. Although the conformational information of Ara6 in the binding site is currently unclear, the results were compared to the model for the protein. The model successfully predicted a small surface area for the binding site of this antibody, which results in a different

epitope conformation. The most obvious deviation between the model and the crystal structure however, is in the CDR H3 loop. As I mentioned before, this loop is the most flexible CDR, with a variable size in different antibodies. The model was able to predict the size of this loop. Furthermore, the model showed that CDR H3 and CDR L1 together provide a smaller binding surface area to the ligand in comparison to that of CS-35. However, the conformation of individual residues especially at CDR H3 loop showed discrepancies between the model and the crystal structure. Noticeably, the model was successful in predicting two conserved Trp and Asn residues that we previously studied in CS-35. Energy minimization of the model and the ligand by molecular dynamic calculations would be a useful approach to increase the prediction property of CS-40 model, for which, there is not a crystal structure.

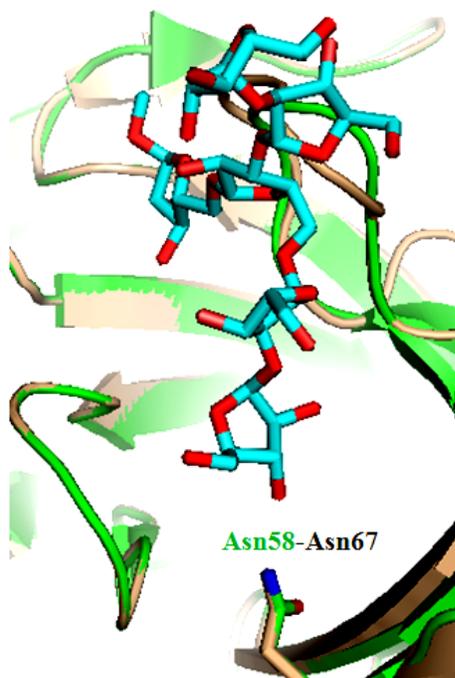


Figure 3.18 Superposition of similar Asn residues in CS-35 and 906.4321 model

The CS-40 scFv showed high similarity to CS-35 in CDR H3. In many antigen–antibody recognition systems, CDR H3 has been shown to make the greatest contribution among the antibody domains to antigen binding,^{24,25} and also has the highest variation among the other domains across different antibodies. For example, the CDR H3 of Se155-4 antibody, which

recognizes a pyranoside antigen on the surface of Salmonella, was investigated as the most important domain of a carbohydrate binding antibody by mutation studies.²⁶ In Chapter 2, we also showed that the CDR H3 of CS-35 makes a substantial contribution to the recognition of Ara6, as a furanoside antigen. Comparison of the created models revealed that the key residue Tyr98 in CDR H3 of CS-35 has counterparts with aromatic residues in CS-40 (Phe112) and 906.4321 (Trp111). The important Phe–Gly–Asn–Tyr segment in CS-35 is similar to the Phe–Pro–Phe–Tyr segment of CDR H3 in CS-40. In contrast, such sequence is absent in the 906.4321 CDR H3. Also, the size of the CDR H3 loop was revealed to be the same in the CS-40 and CS-35 antibodies, containing nine residues. In contrast, the CDR H3 loop in the 906.4321 antibody contains eleven residues.

Except from the CDR H3 size, the topography of the binding site of the 906.4321 antibody differs from CS-35 in the room it provides for accommodating the antigen. The binding site in the 906.4321 antibody is tighter compared to CS-35. The turn in CDR H3 and different shape in CDR L1 are responsible for these differences. This observation is in agreement with the preliminary data from the crystal structure of 906.4321 with Ara6.

To provide some insight into how Ara6 is recognized by the different paratopes in the two antibodies, I investigated the interaction of 906.4321 F_{ab} with **3-4** using STD NMR spectroscopy. The results showed that ring A of Ara6 makes less contact with 906.4321 compared to CS-35. It appears that the limited size of the binding pocket, forces Ara6 to share to the antibody in a different manner than in CS-35. Also, inspection of the model revealed that there is no counterpart for Tyr50 of CS-35 in 906.4321, which we showed to interact with ring A in the CS-35 system. Moreover, Asn34 of CS-35, is replaced by Glu in 906.4321. Manipulation of both of these residues had previously proved to affect the binding of **3-4** by the CS-35 scFv. It is possible that the amino acid substitutions in these positions are also responsible for the decrease in the signals of Ring A. In any case, this deviation is interesting, because in CS-35–Ara6 system, ring A had shown to be a pivotal part of the epitope for binding, whose deletion from the structure of Ara6 significantly reduced the binding affinity, while it makes minimum interaction with 906.4321 antibody. Preliminary analysis of the Ara6 ligand in the binding site of 906-4321 antibody confirmed this observation.

Overall, it seems that the binding of 906.4321 to Ara6 is more similar to CS-35 in rings C–E of the epitope, and deviates in ring A. In contrast, the similar size and shape of CDR H3, and

L2 suggest in CS-35 and CS-40 suggest that both proteins binds rings A and B similarly but deviate for rings C and E. After a comparative analysis of the three antibodies that bind to Ara6, it seems that aromatic residues shape an important part of the specificity. Also key hydrogen bonds seem to be the same for the interaction.

Based on the data provided above and the analysis of it, these results will pave the way for mutation studies on the scFv fragments for structural analysis and possible X-ray crystallizations. This will provide a more detailed picture of the nature of arabinofuranose interactions with antibodies, and furanoside–protein interaction in general. In preliminary data that was provided by our crystallographer collaborators at the University of Calgary, it was shown that the 906.4321 CDR H3 and CDR L1 obviously deviate from those domains in CS-35. Also, the size of the CDR H3 was larger than CS-35 CDR H3, containing eleven residues. The binding site has been shown to be smaller, providing less room to Ara6. All of these data confirm the results from modeling the 906.4321 variable domains.

3.5 Conclusions

In this chapter, I described the preparation of two novel single chain variable fragments of the LAM recognizing antibodies, CS-40 and 906.4321. Both fragments were able to recognize the Ara6 antigen, which is part of the mycobacterial LAM structure. Some aspects of these recognition events such as the affinities, and the generation of homology models and docking of substrates were investigated. Similar to the CS-35–Ara6 system, the participation of aromatic residues in the recognition was shown to be important in the binding from both the CD spectroscopy and modeling results. Common residues were highlighted in the three systems, and affinities were compared. As potential future directions, the binding sites can be studied by single-site mutations and X-ray structures to complement the information about these interactions.

References

- (1) Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P. A.; Yuriev, E. *J. Chem. Info. Model.* **2009**, *49*, 2749.
- (2) Hunter, S. W.; Gaylord, H.; Brennan, P. *J. Biol. Chem.* **1986**, *261*, 12345.
- (3) Navoa, J. A. D.; Laal, S.; Pirofski, L.-A.; McLean, G. R.; Dai, Z.; Robbins, J. B.; Schneerson, R.; Casadevall, A.; Glatman-Freedman, A. *Clin. Diag. Lab. Immunol.* **2003**, *10*, 88.
- (4) Gaylord, H.; Brennan, P.; Young, D.; Buchanan, T. *Infect. and Immun.* **1987**, *55*, 2860.
- (5) Chatterjee, D.; Lowell, K.; Rivoire, B.; McNeil, M. R.; Brennan, P. *J. Biol. Chem.* **1992**, *267*, 6234.
- (6) Yuan, X.; Gubbins, M. J.; Berry, J. D. *J. Immunol. Methods* **2004**, *294*, 199.
- (7) Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* **1997**, *201*, 35.
- (8) Burmester, J.; Plückthun, A. *Antibody Engineering*; Springer: **2001**, p 19.
- (9) GenScript, C. *GenScript Corporation* **2006**.
- (10) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195.
- (11) Guex, N.; Peitsch, M. C.; Schwede, T. *Electrophoresis* **2009**, *30*, S162.
- (12) Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *J. Mol. Biol.* **2009**, *392*, 381.
- (13) Greenfield, N. J. *Nature Protocols* **2007**, *1*, 2876.
- (14) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta (BBA)-Proteins and Proteomics* **2005**, *1751*, 119.
- (15) Biacore, A. *Biacore AB, Uppsala, Sweden* **1997**.
- (16) Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T. *J. Am. Chem. Soc.* **2007**, *129*, 10489.
- (17) Shirai, H.; Kidera, A.; Nakamura, H. *FEBS Lett.* **1996**, *399*, 1.
- (18) Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A. M. *J. Mol. Biol.* **1998**, *275*, 269.
- (19) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291.

- (20) Barenholz, A.; Hovav, A.-H.; Fishman, Y.; Rahav, G.; Gershoni, J. M.; Bercovier, H. *J. Med. Microbiol.* **2007**, *56*, 579.
- (21) Hudson, P. J.; Kortt, A. A. *J. Immunol. Methods* **1999**, *231*, 177.
- (22) Huston, J. S.; Levinson, D.; Mudgett-Hunter, M.; Tai, M.-S.; Novotný, J.; Margolies, M. N.; Ridge, R. J.; Brucoleri, R. E.; Haber, E.; Crea, R. *Proc. Natl. Acad. Sci. U.S.A.* 1988, *85*, 5879.
- (23) Huston, J. S.; Mudgett-Hunter, M.; Tai, M.-S.; McCartney, J.; Warren, F.; Haber, E.; Oppermann, H. *Methods Enzymol.* **1991**.
- (24) Alzari, P.; Lascombe, M.; Poljak, R. *Annu. Rev. Immunol.* **1988**, *6*, 555.
- (25) Mian, I. S.; Bradwell, A. R.; Olson, A. J. *J. Mol. Biol.* **1991**, *217*, 133.
- (26) Brummell, D. A.; Sharma, V. P.; Anand, N. N.; Bilous, D.; Dubuc, G.; Michniewicz, J.; MacKenzie, C. R.; Sadowska, J.; Sigurskjold, B. W. *Biochemistry* **1993**, *32*, 1180.

Chapter 4 : Conclusions and Future Directions

Antibody–carbohydrate interactions are fascinating systems for determining the characteristics of the recognition of carbohydrates by proteins. They have high specificity for their antigens, and the use of scFv technology represents a useful tool for engineering the binding site. Many carbohydrate–antibody interactions occur at the cell wall of bacteria, one of which is mycobacterial species. Our group has had a long standing interest in discovering new mechanisms involved in the pathogenesis, diagnosis, and treatment of mycobacterial infections. The successful synthesis of arabinofuranoside fragments of LAM was an important step that led to the initial binding analysis of mycobacterial arabinofuranosides with the CS-35 antibody. The affinity of the recognition of the Ara6 moiety in LAM was shown to have a low micromolar dissociation constant ($K_D = 6.2 \mu\text{M}$). The good affinity of this interaction, and the crystal structure of the Ara6 in combination with CS-35 F_{ab},^{1,2} raised a number of questions concerning the nature of this recognition. Moreover, the significance of this interaction in mycobacterial species, and the scarce amount of information about the recognition motifs of furanosides by proteins were intriguing motivations to explore this interaction system as an example of furanoside–protein recognition. In this chapter, I summarize my results and will try to give a perspective for the future directions of this research based on the lessons this process taught me.

4.1 Summary: Generation and Binding Specificity of scFv Mutants

One of the most challenging steps was to develop stable scFv derivatives of the three monoclonal antibodies. The generation of an scFv is not always accompanied by success. In the previous chapters I explained the successful approaches that led us to accomplish our goal. However, some of the steps were fulfilled after other approaches had failed.

I created a library of the CS-35 scFv mutants that were probed for their binding to the Ara6 antigen. The parent CS-35 scFv itself was used in various techniques for determining the kinetics of the binding, epitope determination, and structural studies. Overall, we showed that the binding

of Ara6 to the CS-35 antibody is assisted by a combination of significant hydrophobic interactions and hydrogen bonds. The geometry of the combining site was shown to be a determining factor in the specificity, and made the antibody favor ABCE epitope over ABDF epitope¹ (**Figure 4.1**). I discussed the contribution of various residues and functional groups to the binding, and showed how the correct position and distance of each residue assisted the binding. The unusual *cis* peptide bond in the paratope, the unusual conformation of Ring B in the Ara6–CS-35 structure, and the good affinity of the binding were explained based on the mutation results, the molecular forces involved, and the kinetics studies.

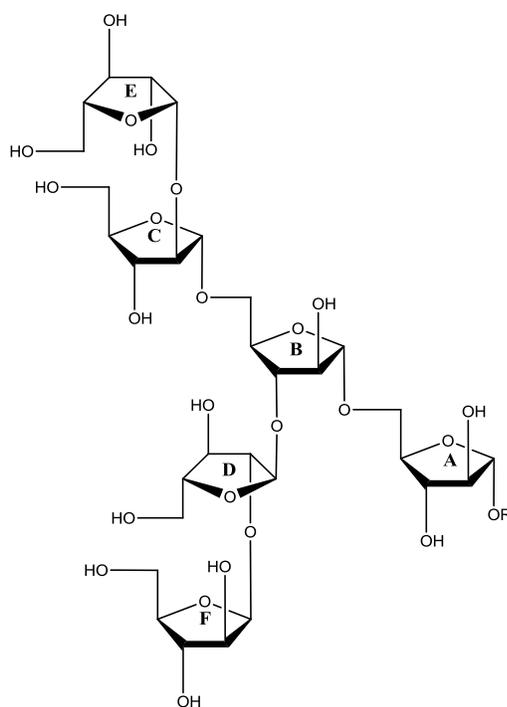


Figure 4.1 Ara6 structure. In nature, R is a complex glycan present in the mycobacterial cell wall.

I also successfully developed scFvs of the CS-40, and 906.4321 monoclonal antibodies. The binding specificity of CS-40 and 906.4321, which had been investigated previously only to a small degree by ELISA experiments, were further studied in more detail, and the paratopes for each was scrutinized by developing homology models. In these systems, the involvement of aromatic residues in the binding was also observed, and the kinetics of binding was probed.

Structural aspects of paratope and epitope were also discussed based on these results. The observed similarities and discrepancies between the CS-35, CS-40, and 906.4321 systems are interesting motivations for further research.

4.2 Future Work: Optimization of CS-40 and 906.4321 scFv Production

My research started by generating the CS-35 scFv. I showed that a functional scFv of this antibody could be produced by inclusion body expression in *E. coli*. However, we had also tried cytoplasmic and periplasmic expression prior to this approach. The inclusion body expression method was successful for the generation of CS-40 and 906.4321 scFvs, as well. To increase the yield of the production for the two latter scFvs in the future, it is worth trying cytoplasmic and periplasmic expression, as well as optimization of the refolding of the proteins after denaturing from inclusion bodies.

The most important step in generating scFvs from inclusion body expression is the proper folding of the denatured proteins. Although the CS-35 scFv was produced effectively in the absence of redox conditions, this could be different for the two other antibodies. Therefore, the expression and refolding conditions should be optimized individually for each scFv to obtain the best yields. Rational mutation at the interface of the V_H and V_L domains is another possible approach to enhance the proper folding and yield of these two scFvs. The hydrophobic patches at such interfaces are usually responsible for the aggregation of the scFvs and their diminished stabilities compared to their corresponding F_{ab} s. Mutating one or two such residues, where the interface of the two variable domains reside based on the model, could improve solubility and stability issues. Buffer conditions (concentration, pH, etc.) also affect the stability of scFv in different steps of production, as we showed for CS-35. These factors could be examined in detail and optimized for the other scFvs. Increasing the yield of the scFv is an important goal, given the range of fascinating applications, provided that they are available in sufficient amounts. I will refer to some such applications in the next sections.

4.3 Future Work: Thermodynamic Binding Parameters

The binding of different fragments to Ara6 was investigated from different aspects, such as eiptope and paratope structural characteristics in the binding, kinetics, and affinity. Valuable information about the nature of the binding can be obtained by investigating the thermodynamic properties of the interactions. Enthalpic contributions to the CS-35 F_{ab}–Ara6 interaction, as well as the determination of the association constants, have been studied by isothermal titration calorimetry (ITC).¹ It would be interesting to measure these data by ITC for the CS-35 scFv. ITC could also provide useful thermodynamic information about the binding of CS-40 and 906.4321 scFvs to Ara6, and the results would be valuable in comparing the three systems. We suggest ITC as a potential direction for these binding measurements. This technique can also provide association constants, which can be used to compare to the SPR spectroscopic results. ITC requires higher amounts of protein, and ligand, compared to SPR spectroscopy, but the thermodynamic information it provide is more detailed, and therefore would be valuable for better understanding of these interactions. Therefore, as outlined above, optimizing the yield of the scFv production for the three systems to give sufficient amount of protein for ITC experiments is an important goal.

4.4 Future Work: Crystal structures of scFvs

The heterogeneous patterns of glycosylation on mAbs make it a challenge to study them by X-ray crystallography. Therefore, F_{abs} are useful antibody fragments in crystallography. As I mentioned in the previous chapters, efforts to obtain a pure CS-40 F_{ab} have failed. In the absence of a CS-40 F_{ab}, the CS-40 scFv provides a valuable opportunity to obtain the X-ray crystal structure in combination with Ara6-based ligands (e.g., **2**). A structure of this type would add a detailed set of data about the paratope and epitope in the combining site, and will be valuable in assessing the precision of the model. The crystal structure of Se155-4 scFv and F_{ab} has been compared to each other before. It would be useful to do the same comparison for the CS-35 scFv and 906.4321 scFvs with their corresponding F_{abs}, both of which are now available.

4.5 Future Work: Mutational Studies on CS-40 and 906.4321 scFvs

Manipulating the binding site of the 906.4321 and CS-40 scFvs is another direction to further study these interactions. In particular, CDR H3, which has been shown to be the most important domain in other antigen–antibody recognition studies, would be a good starting point. The homology models I have developed could be used to pick the residues to mutate based on the possible interactions they make with Ara6. Mutants can be further assessed for their binding to Ara6 by various methods (see above). This direction will experimentally evaluate the importance of different molecular forces in the binding of each system. In particular, mutation of the conserved residues and multiple aromatic residues in the binding site of the 906.4321 antibody is recommended.

4.6 Future Work: Applying the Present Information in the Design of New Tools for Studying Mycobacterial Disease

The ultimate goal in this type of research is to find new approaches to diagnose or treat related diseases. Armed with information about the binding interactions of these proteins to Ara6, scFvs could be optimized by directed evolution to find mutants with higher affinity for the antigen in search for TB diagnostic tools based upon detecting LAM in serum or urine. Moreover, arabinofuranoside antigens with improved interactions could be synthesized based on the present results. Such synthetic approaches can be considered in generating vaccines with high affinity for the related antibodies. Jimenez–Barbero and colleagues have proposed enhancing the stability of the carbohydrate–aromatic complexes by increasing the polarity of the C–H bond via proper substitutions on the pyranose ring, to make more effective CH– π interactions.³ The same strategy can be applied to furanoside systems.

References

- (1) Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T. *Journal of the American Chemical Society* **2007**, *129*, 10489.
- (2) Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *Journal of Molecular Biology* **2009**, *392*, 381.
- (3) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. S. *Accounts of Chemical Research* **2012**, *46*, 946.

Bibliography

- Abbas, A. K.; Lichtman, A. H.; Pillai, S. *Cellular and Molecular Immunology*; Elsevier Health Sciences, **1994**.
- Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P. A.; Yuriev, E. *J. Chem. Info. Model.* **2009**, *49*, 2749.
- Ahmad, Z. A.; Yeap, S. K.; Ali, A. M.; Ho, W. Y.; Alitheen, N. B. M.; Hamid, M. J. *Immunol. Res.* **2012**, *2012*, 980250.
- Alzari, P.; Lascombe, M.; Poljak, R. *Annu. Rev. Immunol.* **1988**, *6*, 555.
- Ambrosi, M.; Cameron, N. R.; Davis, B. G. *Org. Biomol. Chem.* **2005**, *3*, 1593.
- Angyal, S. J. *Adv. Carb. Chem. Biochem.* **1984**, *42*, 15.
- Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195.
- Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. S. *Acc. Chem. Res.* **2012**, *46*, 946.
- Bagaria, A.; *Proteins Struct. Funct. Bioinfo.* **2011**, *79*, 1352.
- Barenholz, A.; Hovav, A.-H.; Fishman, Y.; Rahav, G.; Gershoni, J. M.; Bercovier, H. *J. Med. Microbiol.* **2007**, *56*, 579.
- Bélot, F.; Guerreiro, C.; Baleux, F.; Mulard, L. A. *Chem. Eur. J.* **2005**, *11*, 1625.
- Bélot, F.; Wright, K.; Costachel, C.; Phalipon, A.; Mulard, L. A. *J. Org. Chem.* **2004**, *69*, 1060.
- Biacore Manual, A. *Biacore AB, Uppsala, Sweden* **1997**.
- Blake, C.; Koenig, D.; Mair, G.; North, A.; Phillips, D.; Sarma, V. *Nature* **1965**, *206*, 757.
- Bock, K.; Josephson, S.; Bundle, D. R. *J. Chem. Soc. Perkin Transaction 2* **1982**, 59.
- Boraston, A.; Bolam, D.; Gilbert, H.; Davies, G. *J. Biochem.* **2004**, *382*, 769.
- Boss, M. A.; Kenten, J. H.; Wood, C. R.; Emtage, J. S. *Nucleic Acids Res.* **1984**, *12*, 3791.

- Bottger, E.; Teske, A.; Kirschner, P.; Bost, S.; Hirschel, B.; Chang, H.; Beer, V. *The Lancet* **1992**, *340*, 76.
- Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland New York, 1991; Vol. 2.
- Brennan, P. J.; Nikaido, H. *Annu. Rev. Biochem.* **1995**, *64*, 29.
- Briken, V.; Porcelli, S. A.; Besra, G. S.; Kremer, L. *Mol. Microbiol.* **2004**, *53*, 391.
- Brummell, D. A.; Sharma, V. P.; Anand, N. N.; Bilous, D.; Dubuc, G.; Michniewicz, J.; MacKenzie, C. R.; Sadowska, J.; Sigurskjold, B. W. *Biochemistry* **1993**, *32*, 1180
- Bundle, D. R.; Alibés, R.; Nilar, S.; Otter, A.; Warwas, M.; Zhang, P. *J. Am. Chem. Soc.* **1998**, *120*, 5317.
- Bundle, D. R.; Baumann, H.; Brisson, J.-R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183.
- Bundle, D. R.; Eichler, E.; Gidney, M. A. J.; Meldal, M.; Ragauskas, A.; Sigurskjold, B. W.; Sinnott, B.; Watson, D. C.; Yaguchi, M.; Young, N. M. *Biochemistry* **1994**, *33*, 5172.
- Bundle, D.; Oxford University Press: New York: **1999**, p 370.
- Burmester, J.; Plückthun, A. *Antibody Engineering*; Springer: **2001**, p 19.
- Capra, J. D.; Edmundson, A. B. *Sci. Am.* **1977**, *236*, 50.
- Carver, J. *Pure Appl. Chem.* **1993**, *65*, 763.
- Chatterjee, D.; Lowell, K.; Rivoire, B.; McNeil, M. R.; Brennan, P. *J. Biol. Chem.* **1992**, *267*, 6234.
- Coburn, B.; Grassl, G. A.; Finlay, B. *Immunol. Cell Biol.* **2007**, *85*, 112.
- Colcher, D.; Pavlinkova, G.; Beresford, G.; Booth, B.; Choudhury, A.; Batra, S. *Quar. J. Nuc. Med. (AIMN) (IAR)* **1998**, *42*, 225.
- Cygler, M.; Rose, D. R.; Bundle, D. R. *Science* **1991**, *253*, 442.
- Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387.
- Davies, D. R.; Metzger, H. *Annu. Rev. Immunol.* **1983**, *1*, 87.
- Deng, S.; MacKenzie, C. R.; Sadowska, J.; Michniewicz, J.; Young, N. M.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1994**, *269*, 9533.

- Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond: in Structural Chemistry and Biology*; Oxford University Press, 2001; Vol. 9.
- Dunitz, J.D. *Science*, **1994**, *264*, 670.
- Engering, A.; Geijtenbeek, T. B.; van Kooyk, Y. *Trends Immunol* **2002**, *23*, 480.
- Gabius, H.-J.; André, S.; Jiménez-Barbero, J.; Romero, A.; Solís, D. *Trends Biochem. Sci.* **2011**, *36*, 298.
- Gabius, H.-J.; André, S.; Kaltner, H.; Siebert, H.-C. *Biochim. Biophys. Acta* **2002**, *1572*, 165.
- Galili, U. *Immunol. Cell Biol.* **2005**, *83*, 674.
- Gaylord, H.; Brennan, P.; Young, D.; Buchanan, T. *Infec. Immun.* **1987**, *55*, 2860.
- GenScript, C. *GenScript Corporation* **2006**.
- Goldstein, I.; Hollerman, C.; Smith, E. *Biochemistry* **1965**, *4*, 876.
- Greenfield, N. J. *Nature Protocols* **2007**, *1*, 2876.
- Guex, N.; Peitsch, M. C.; Schwede, T. *Electrophoresis* **2009**, *30*, S162.
- Hamasur, B.; Källenius, G.; Svenson, S. B. *Vaccine* **1999**, *17*, 2853.
- Harata, K.; Muraki, M. *Acta Crystallogr. Section D: Biological Crystallography* **1997**, *53*, 650.
- Haselkorn, D.; Friedman, S.; Givol, D.; Pecht, I. *Biochemistry* **1974**, *13*, 2210.
- Hayashida, M.; Fujii, T.; Hamasu, M.; Ishiguro, M.; Hata, Y. *J. Mol. Biol.* **2003**, *334*, 551.
- Holliger, P.; Hudson, P. J. *Nature Biotechnol.* **2005**, *23*, 1126.
- Hooper, L. V.; Gordon, J. I. *Glycobiology* **2001**, *11*, 1R.
- Horler, R.; Müller, A.; David C. *J. Biol. Chem.* **2009**, *284*, 31156.
- Houseknecht, J. B.; Lowary, T. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 677.
- Hudson, P. J.; Kortt, A. A. *J. Immunol. Methods* **1999**, *231*, 177.
- Hunter, S. W.; Gaylord, H.; Brennan, P. *J. Biol. Chem.* **1986**, *261*, 12345.

- Huston, J. S.; Levinson, D.; Mudgett-Hunter, M.; Tai, M.-S.; Novotný, J.; Margolies, M. N.; Ridge, R. J.; Bruccoleri, R. E.; Haber, E.; Crea, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5879.
- Huston, J. S.; Mudgett-Hunter, M.; Tai, M.-S.; McCartney, J.; Warren, F.; Haber, E.; Oppermann, H. *Methods Enzymol.* **1991**.
- J de Mel, V. S.; Doscher, M. S.; Martin, P. D.; Edwards, B. F. *FEBS Lett.* **1994**, *349*, 155.
- Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291.
- Janakiraman, M. N.; White, C. L.; Laver, W. G.; Air, G. M.; Luo, M. *Biochemistry* **1994**, *33*, 8172.
- Jennison, A. V.; Verma, N. K. *FEMS Microbiol. Rev.* **2004**, *28*, 43.
- Kaur, D.; Lowary, T. L.; Vissa, V. D.; Crick, D. C.; Brennan, P. J. *Microbiology* **2002**, *148*, 3049.
- Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta -Proteins and Proteomics* **2005**, *1751*, 119
- Kitova, E. N.; Wang, W.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 13980.
- Klimka, A.; Matthey, B.; Roovers, R.; Barth, S.; Arends, J.; Engert, A.; Hoogenboom, H. *Br. J. Cancer* **2000**, *83*, 252.
- Kontermann, R. E.; Müller, R. *J. Immunol. Methods* **1999**, *226*, 179.
- Koshland, D. E. *Angew. Chem. Int. Ed.* **1995**, *33*, 2375.
- Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* **1997**, *201*, 35.
- Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973.
- Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373.
- Lemieux, R. U.; Spohr, U. *Adv. Carb. Chem. Biochem.* **1994**, *50*, 1.
- Lindh, I.; Hindsgaul, O. *J. Am. Chem. Soc.* **1991**, *113*, 216.
- Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.
- Lowary, T. L. *Curr. Opin. Chem. Biol.* **2003**, *7*, 749.

- Lu, Z.; Madico, G.; Roche, M. I.; Wang, Q.; Hui, J. H.; Perkins, H. M.; Zaia, J.; Costello, C. E.; Sharon, J. *Immunology* **2012**, *136*, 352.
- Lu, Z.; Rynkiewicz, M. J.; Yang, C. Y.; Madico, G.; Perkins, H. M.; Wang, Q.; Costello, C. E.; Zaia, J.; Seaton, B. A.; Sharon, J. *Immunology* **2013**, *140*, 374.
- MacKenzie, C. R.; Hiramata, T.; Deng, S.-j.; Bundle, D. R.; Narang, S. A.; Young, N. M. *J. Biol. Chem.* **1996**, *271*, 1527.
- Maeda, H.; Schmidt-Kessen, A.; Engel, J.; Jaton, J.C. *Biochemistry* **1997**, *16*, 4086.
- Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784.
- Means, T. K.; Lien, E.; Yoshimura, A.; Wang, S.; Golenbock, D. T.; Fenton, M. J. *J. Immunol.* **1999**, *163*, 6748.
- Merritt, E. A.; Sarfaty, S.; Akker, F. V. D.; L'Hoir, C.; Martial, J. A.; Hol, W. G. *Protein Sci.* **1994**, *3*, 166.
- Mian, I. S.; Bradwell, A. R.; Olson, A. J. *J. Mol. Biol.* **1991**, *217*, 133.
- Miranker, A.; Robinson, C.; Radford, S.; Dobson, C. *FASEB J.* **1996**, *10*, 93.
- Miyanaga, A.; Koseki, T.; Matsuzawa, H.; Wakagi, T.; Shoun, H.; Fushinobu, S. *J. Biol. Chem.* **2004**, *279*, 44907.
- Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A. M. *J. Mol. Biol.* **1998**, *275*, 269.
- Muraki, M. *Pro. Pept. Let.* **2002**, *9*, 195.
- Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. *J. Mol. Biol.* **2009**, *392*, 381.
- Nagae, M.; Yamaguchi, Y. *Int. J. Mol. Sci.* **2012**, *13*, 8398.
- Nassau, P.M.; Martin, S. L.; Brown, R. E.; Weston, A.; Monsey, D.; McNeil, M. R.; Duncan, K. *J. Bacteriol.* **1996**, *178*, 1047.
- Navoa, J. A. D.; Laal, S.; Pirofski, L.-A.; McLean, G. R.; Dai, Z.; Robbins, J. B.; Schneerson, R.; Casadevall, A.; Glatman-Freedman, A. *Clin. Diag. Lab. Immunol.* **2003**, *10*, 88.
- Nieba, L.; Honegger, A.; Krebber, C.; Plückthun, A. *Protein Eng.* **1997**, *10*, 435.
- Nishio, M. *Phys. Chem.* **2011**, *13*, 13873.

- Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/π Interaction: Evidence, Nature, and Consequences*; John Wiley & Sons, 1998; Vol. 21.
- O'Brien, R. J.; Nunn, P. P. *Am. J. Respir. Crit. Care Med.* **2001**, *163*, 1055.
- Pathiaseril, A.; Woods, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 331.
- Payne, C. M.; Bomble, Y. J.; Taylor, C. B.; McCabe, C.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Biol. Chem.* **2011**, *286*, 41028.
- Plückthun, A.; Pack, P. *Immunotechnology* **1997**, *3*, 83.
- Proba, K.; WoÈrn, A.; Honegger, A.; PluÈckthun, A. *J. of Mol. Biol.* **1998**, *275*, 245.
- Quioco, F. A. *Annu. Rev. Biochem.* **1986**, *55*, 287.
- Quioco, F. A.; Vyas, N. K. *Nature* **1984**, *310*, 381.
- Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T. *J. Am. Chem. Soc.* **2007**, *129*, 10489.
- Reidl, J.; Klose, K. E. *FEMS Microbiol. Rev.* **2002**, *26*, 125.
- Rheinnecker, M.; Hardt, C.; Ilag, L. L.; Kufer, P.; Gruber, R.; Hoess, A.; Lupas, A.; Rottenberger, C.; Plückthun, A.; Pack, P. *J. Immunol.* **1996**, *157*, 2989.
- Richards, M. R.; Bai, Y.; Lowary, T. L. *Carb. Res.* **2013**, *374*, 103.
- Rose, D.; Przybylska, M.; To, R.; Kayden, C.; Vorberg, E.; Young, N.; Bundle, D.; Oomen, R. *Prot. Sci.* **1993**, *2*, 1106.
- Rose, N. L.; Completo, G. C.; Shuang-Jun, L.; McNeil, M. R.; Placic, M.; Lowary, T. L. *J. Am. Chem. Soc.* **2006**, *128*, 6721.
- Ross, P. D.; Subramanian, S. *Biochemistry* **1981**, *20*, 3096.
- Ryder, M.; Tate, M.; Jones, G. *J. Biol. Chem.* **1984**, *259*, 9704.
- Rynkiewicz, M. J.; Lu, Z.; Hui, J. H.; Sharon, J.; Seaton, B. A. *Biochemistry* **2012**, *51*, 5684.
- Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.; Park, S. J.; Skehel, J. J.; Wiley, D. C. *Biochemistry* **1992**, *31*, 9609.
- Shirai, H.; Kidera, A.; Nakamura, H. *FEBS Lett.* **1996**, *399*, 1.

- Sieling, P.; Chatterjee, D.; Porcelli, S.; Prigozy, T.; Mazzaccaro, R.; Soriano, T.; Bloom, B.; Brenner, M.; Kronenberg, M.; Brennan, P. *Science* **1995**, *269*, 227.
- Sigurskjold, B. W.; Altman, E.; Bundle, D. R. *Eur. J. Biochem.* **1991**, *197*, 239.
- Sigurskjold, B.; Bundle, D. *J. Biol. Chem.* **1992**, *267*, 8371.
- Stehle, T.; Yan, Y.; Benjamin, T. L.; Harrison, S. C. *Nature* **1994**, *369*, 160.
- Taha, H. A.; Castillo, N.; Sears, D. N.; Wasylshen, R. E.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theor. Comput.* **2009**, *6*, 212.
- Tamres, M. *J. Am. Chem. Soc.* **1952**, *74*, 3375.
- Te Wu, T.; Kabat, E. A. *J. Exp. Med.* **1970**, *132*, 211.
- Theillet, F.-X.; Frank, M.; Vulliez-Le Normand, B.; Simenel, C.; Hoos, S.; Chaffotte, A.; Bélot, F.; Guerreiro, C.; Nato, F.; Phalipon, A. *Glycobiology* **2011**, *21*, 1570.
- Toone, E. J. *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.
- Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. **1999**.
- Villeneuve, S.; Souchon, H.; Riottot, M.-M.; Mazié, J.-C.; Lei, P.-S.; Glaudemans, C.; Kováč, P.; Fournier, J.-M.; Alzari, P. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8433.
- Vulliez-Le Normand, B.; Saul, F.; Phalipon, A.; Belot, F.; Guerreiro, C.; Mulard, L. A.; Bentley, G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9976.
- Vyas, N. K. *Curr. Opin. Struct. Biol.* **1991**, *1*, 732.
- Vyas, N. K.; Vyas, M. N.; Chervenak, M. C.; Johnson, M. A.; Pinto, B. M.; Bundle, D. R.; Quioco, F. A. *Biochemistry* **2002**, *41*, 13575.
- Wang, J.; Villeneuve, S.; Zhang, J.; Lei, P.-s.; Miller, C. E.; Lafaye, P.; Nato, F.; Szu, S. C.; Karpas, A.; Bystricky, S. *J. Biol. Chem.* **1998**, *273*, 2777.
- Wang, Q.; Shi, X.; Leymarie, N.; Madico, G.; Sharon, J.; Costello, C. E.; Zaia, J. *Biochemistry* **2011**, *50*, 10941.
- Watkins, N. A.; Ouwehand, W. H. *Vox Sang* **2000**, *78*, 72.
- Watowich, S. J.; Skehel, J. J.; Wiley, D. C. *Structure* **1994**, *2*, 719.
- Watson, K. A.; Mitchell, E. P.; Johnson, L. N.; Bichard, C. J.; Orchard, M. G.; Fleet, G. W.; Oikonomakos, N. G.; Leonidas, D.; Son, J. C. *Biochemistry* **1994**, *33*, 5745.

- Wedemayer, G. J.; Patten, P. A.; Wang, L. H.; Schultz, P. G.; Stevens, R. C. *Science* **1997**, *276*, 1665.
- Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441.
- Whitlow, M.; Bell, B. A.; Feng, S.-L.; Filpula, D.; Hardman, K. D.; Hubert, S. L.; Rollence, M. L.; Wood, J. F.; Schott, M. E.; Milenic, D. E. *Protein Eng.* **1993**, *6*, 989.
- WoÈrn, A.; Pluckthun, A. *J. Mol. Biol.* **2001**, *305*, 989.
- Woods, R. J.; Tessier, M. B. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575.
- Wright, K.; Guerreiro, C.; Laurent, I.; Baleux, F.; Mulard, L. A. *Org. Biomol. Chem.* **2004**, *2*, 1518.
- Yoshida, A.; Koide, Y. *Infec. Immun.* **1997**, *65*, 1953.
- Young, N.; Gidney, M.; Gudmundsson, B.; MacKenzie, C.; To, R.; Watson, D.; Bundle, D. *Mol. Immunol.* **1999**, *36*, 339.
- Yuan, X.; Gubbins, M. J.; Berry, J. D. *J. Immunol. Methods* **2004**, *294*, 199.
- Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423.

Appendices

A1. List of Primers

List of primers used for splice overlap PCR

ScFv in pET28a

Forward primer: (BamHI)

5'-TAAATA GGATCCGAAGTCCAACCTGCAGCAATCGGGC-3'

Reverse primer: (Middle primer)

5'-CCACCGCCAGAACCGCCACCGCCCACCGTGACCAGGGTACC CTGG-3'

Forward primer: (Middle primer)

5'-CCAGGGTACCCTGGTCACGGTGGGCGGTGGCGGTTCTGGCGG TGG-3'

Reverse primer: (HindIII)

5'-ATTAAAAGCTTTTATTATTTAATTTCCAGTTTCGTGCCGC-3'

Primer used for A31S:

5'- TCGCAGGACATCGGTAGCTATCTGAA-'3

List of the primers used for mutant generation:

S50A:

GGTCTGGAATGGATCGGTGCAATCTACCCGGGCAACTC

W33A:

CATCAGGCTATTCGTTACCAACTATGCGATGCATTGGGTGA

N58A:

ATCCGGGCAACTCTGATACCGCCTACAAACAGAAATTCAAGG

F95A:

CCGCGGTTTATTACTGCACCCGTGCTGGCAATTATGTCC

Y98A:

GTTTATTACTGCACCCGTTTTGGCAATGCTGTCCCGTTCGCC

Y98F:

ACTGCACCCGTTTTGGCAATTTTGTCCCGTTTCG

N34A:

GGACATCGGTTCGTATCTGGCCTGGTACCAGCAAAAAGT

Y50A:

CGGTTCGTCTGTTGATCTATGCCACCAGCCGCC

N34D:

GGACATCGGTTCGTATCTGGACTGGTACCAGCAAAAACCG

N97D:

CTGCACCCGTTTTGGCGACTATGTCCCGTTCGCC

H35A:

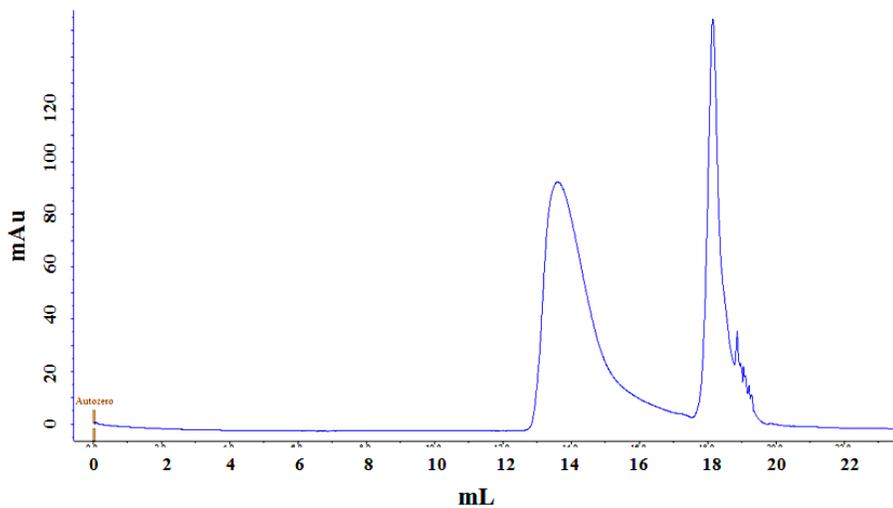
CACCAACTATTGGATGGCTTGGGTGAAACAGCGCCC

A2. FPLC results for the monomer mutants that bind to Ara6

Protein that eluted at 18 mL was collected for binding analysis.

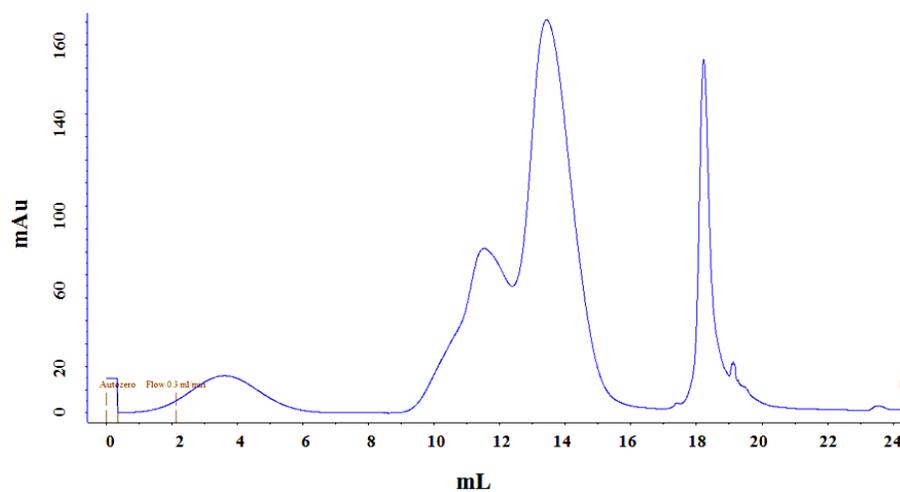
a)

His35Ala:



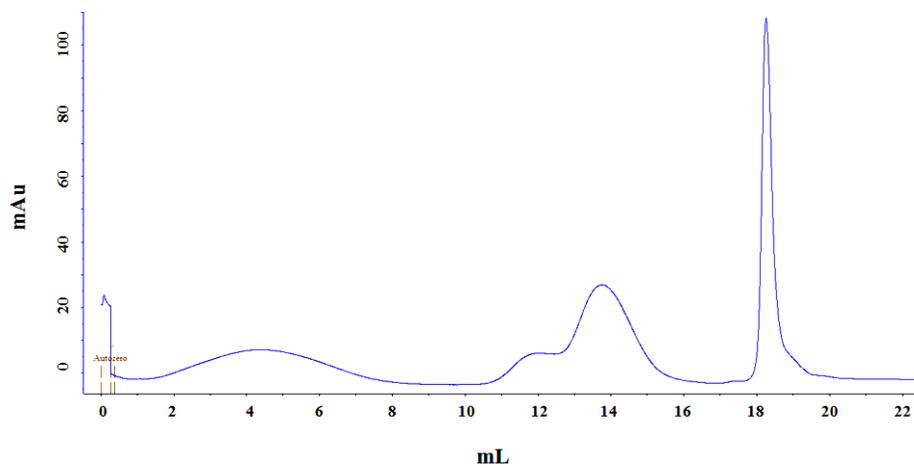
b)

Asn58Ala:



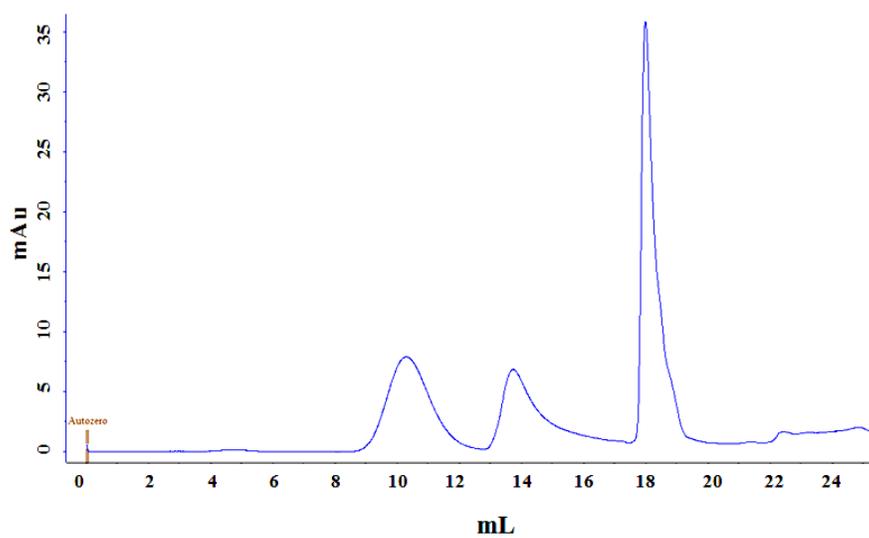
c)

Ser50Ala:



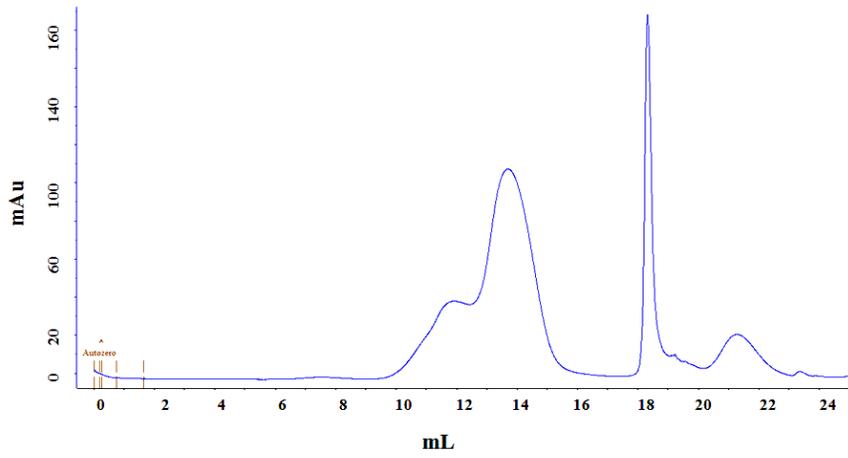
d)

Tyr102Phe:



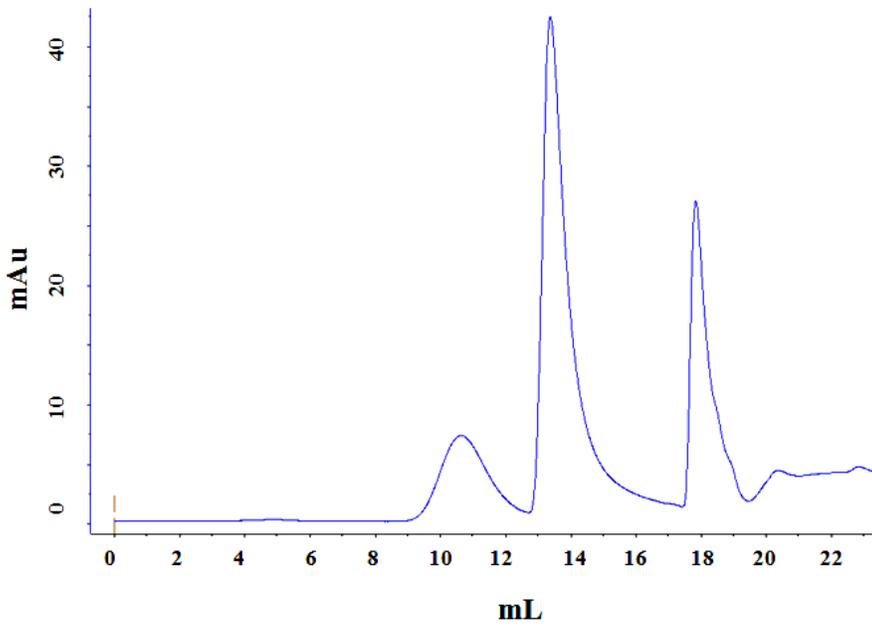
e)

Tyr50Ala:

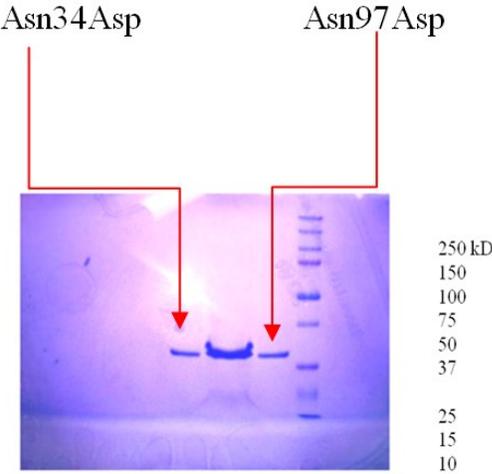
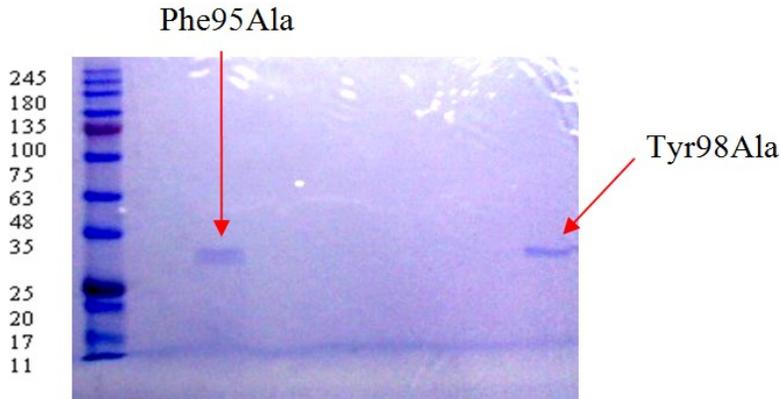


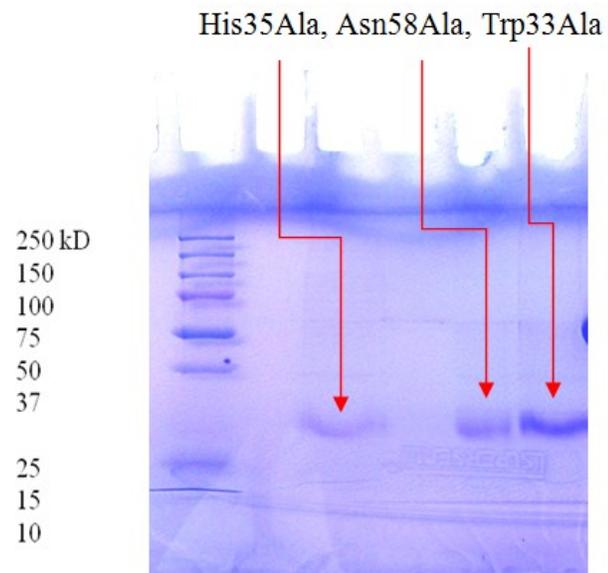
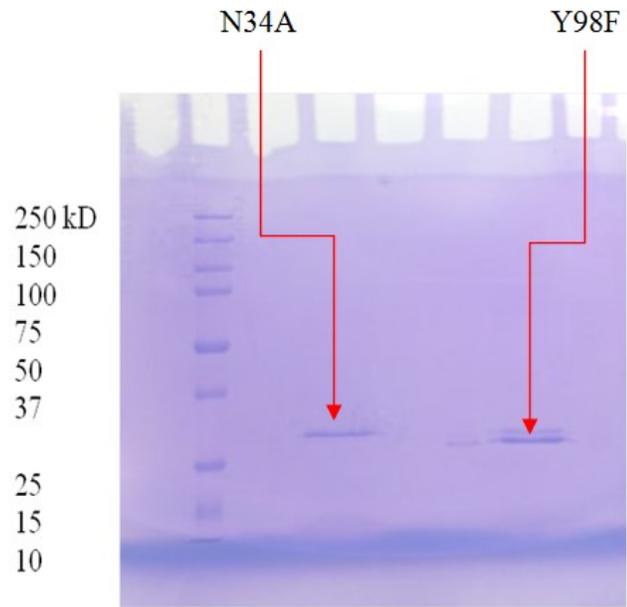
f)

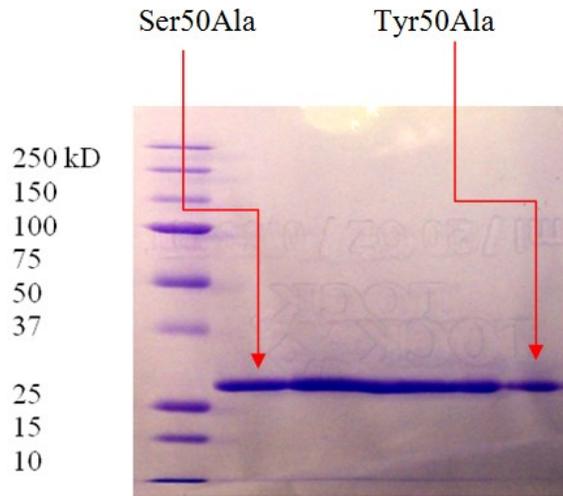
Asn34Asp:



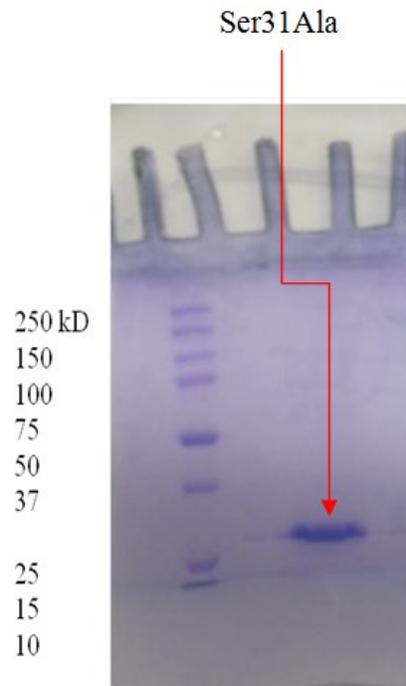
A3. SDS PAGE results for CS-35 scFv mutants





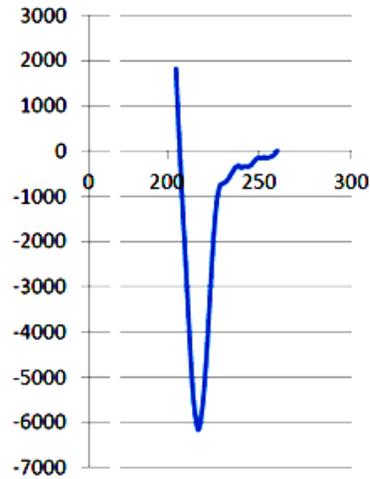


(The peaks in the middle are from CS-35 scFv remaking batches)



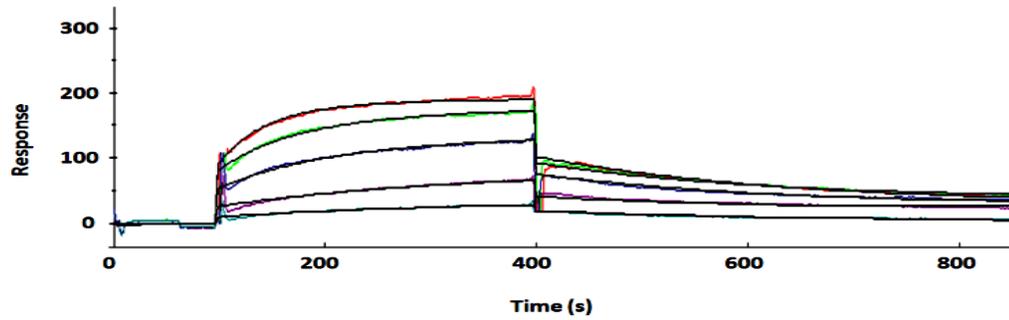
A4. CD Spectrum for Phe95Ala in Far UV

Phe95Ala was tested for its secondary structure in far-UV region by CD spectroscopy. The fact that this mutant showed high oligomerization and aggregation encouraged us to check the secondary structure for any possible impact it might have on the structure. The result showed that the typical secondary structure was conserved.



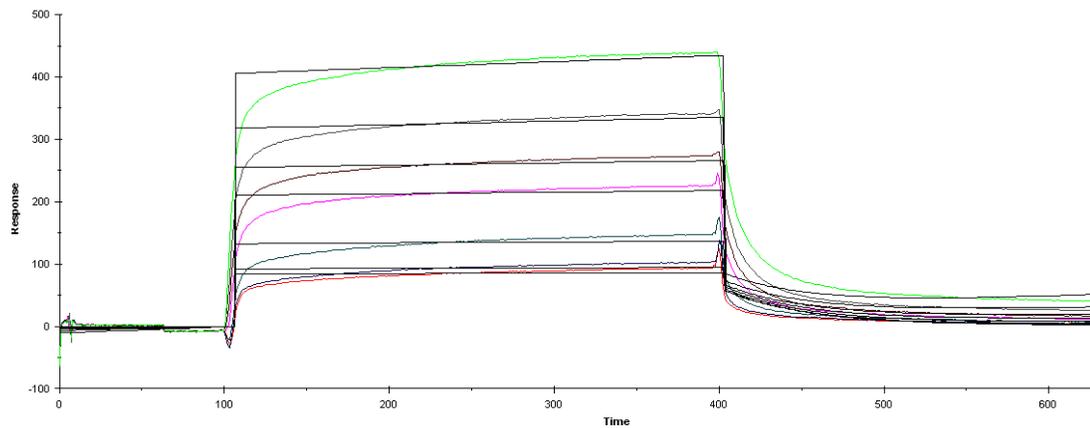
A5. SPR Fittings into 1:1 Binding Model

a) CS-35 scFv



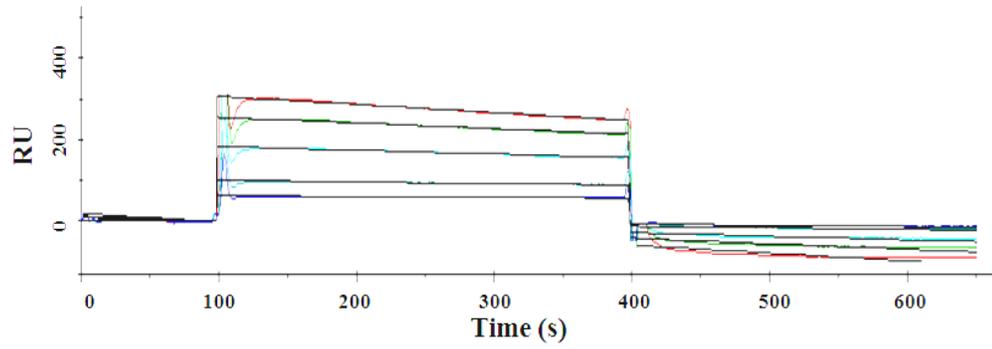
SPR binding of the wild-type CS-35 scFv with **2-3**. Concentrations of the scFv injections were 0.26 —, 0.53 —, 1.6 —, 2.6 —, 4.8 — μM

b) CS-35 Fab

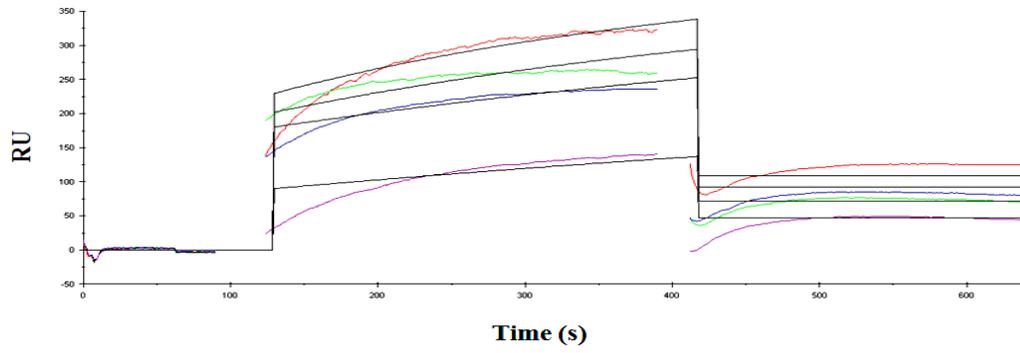


SPR binding of the CS-35 Fab with **2-3**. Concentrations were: 0.02 —, 0.05 —, 0.1 —, 0.18 —, 0.31 —, 0.53 —, 1.08 —, 1.51 — μM

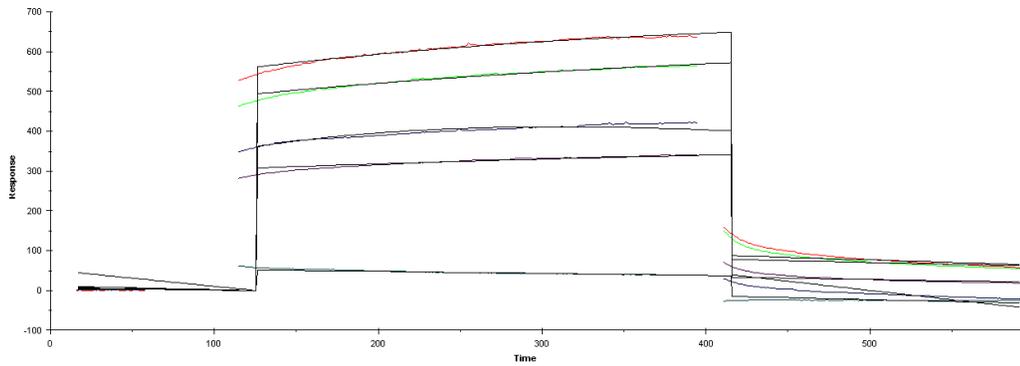
c) Asn34Asp



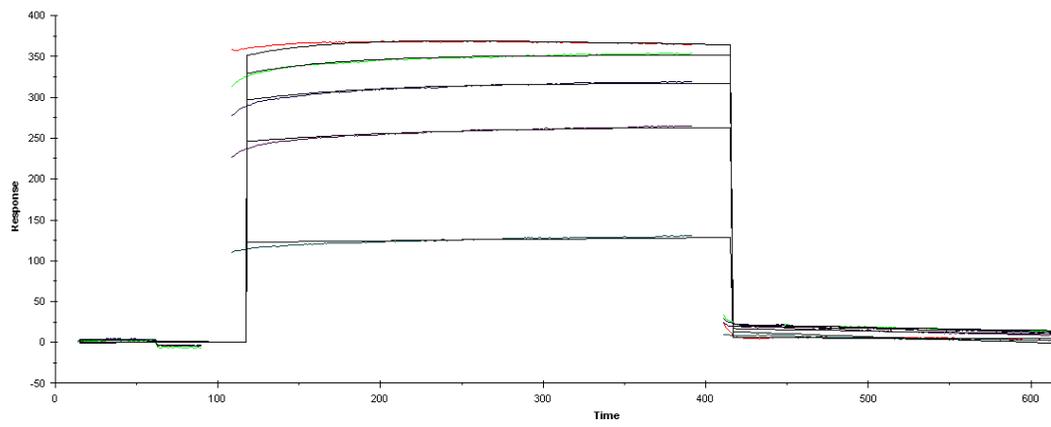
d) Tyr50Ala



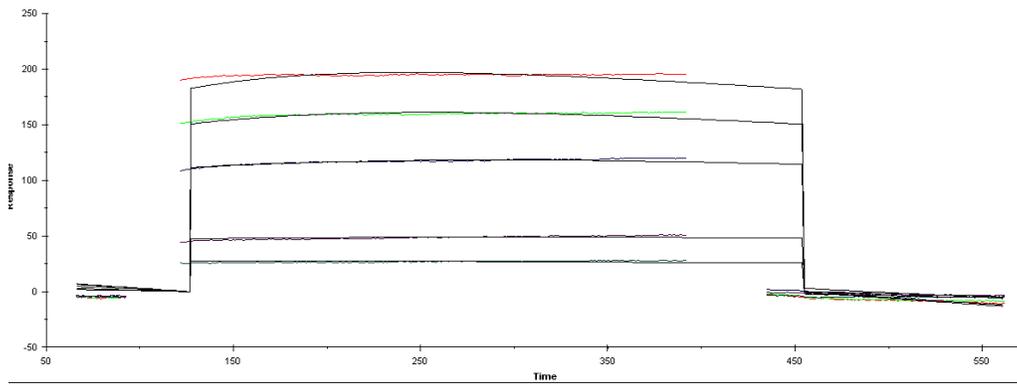
e) Tyr98Phe



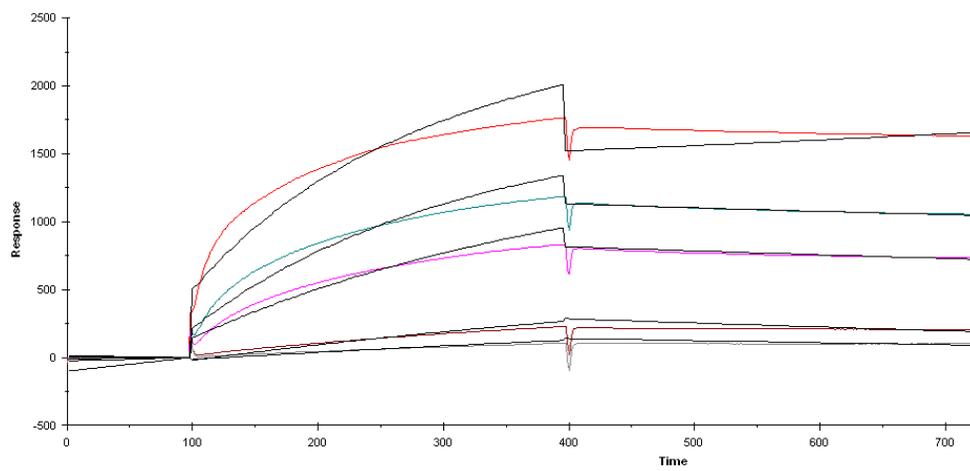
f) 906-4321



g) CS-40

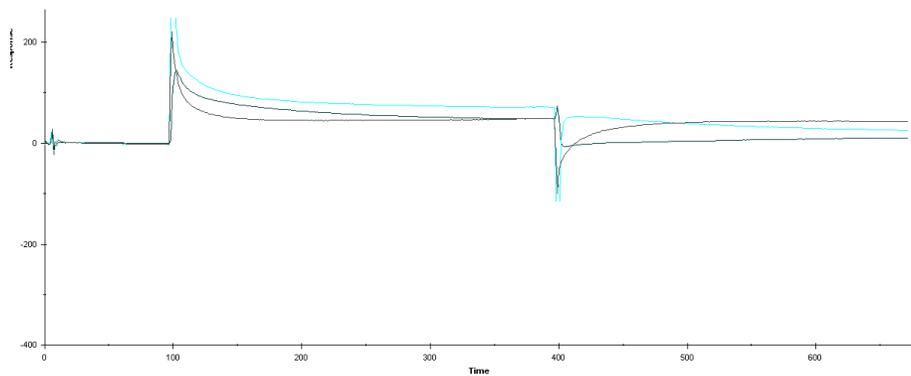


h) CS-35 Dimer

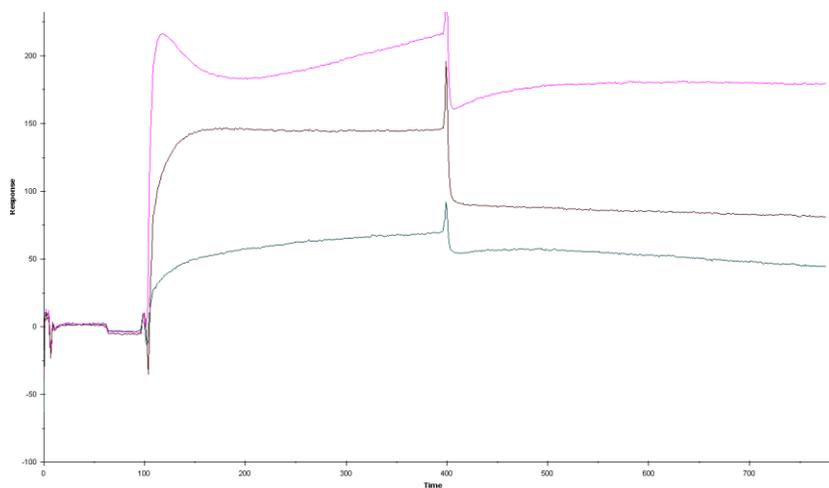


Weak binder mutants: (For these three weak binder mutants the Rmax cannot be obtained at high concentration (about 30 μ M). The protein precipitates above such concentrations)

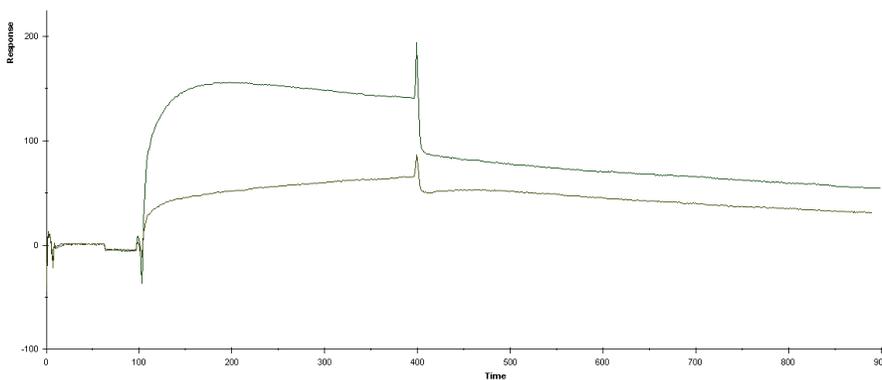
i) Asn58Ala



j) Ser50Ala



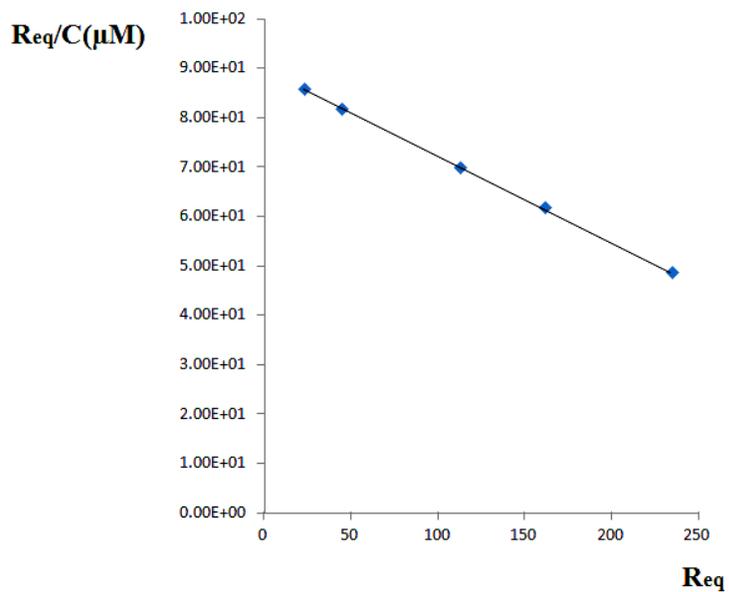
k) His35Ala



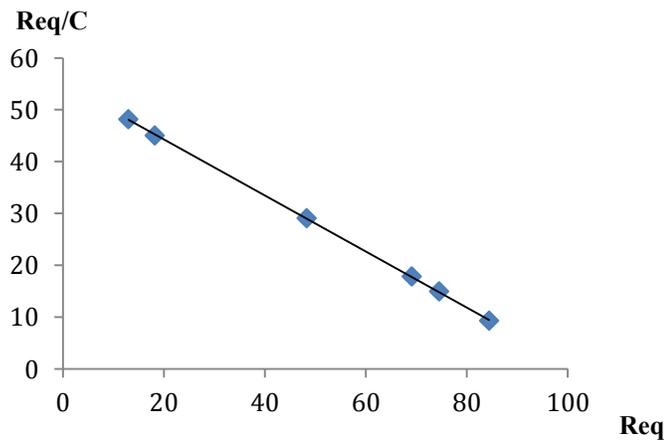
A6. Equilibrium Binding Analyses.

(Concentrations are in μM)

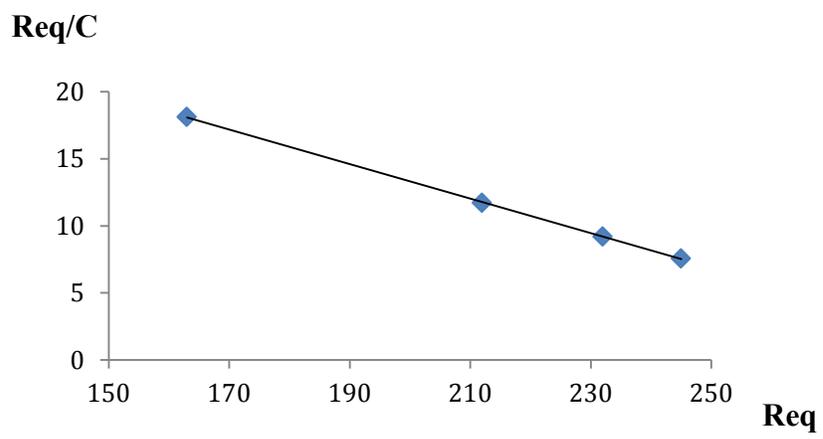
CS-35 scFv



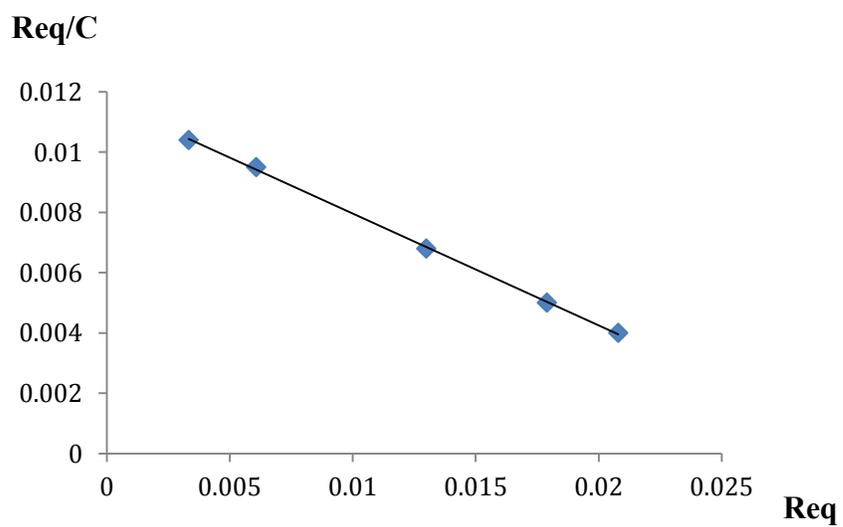
Tyr98Phe



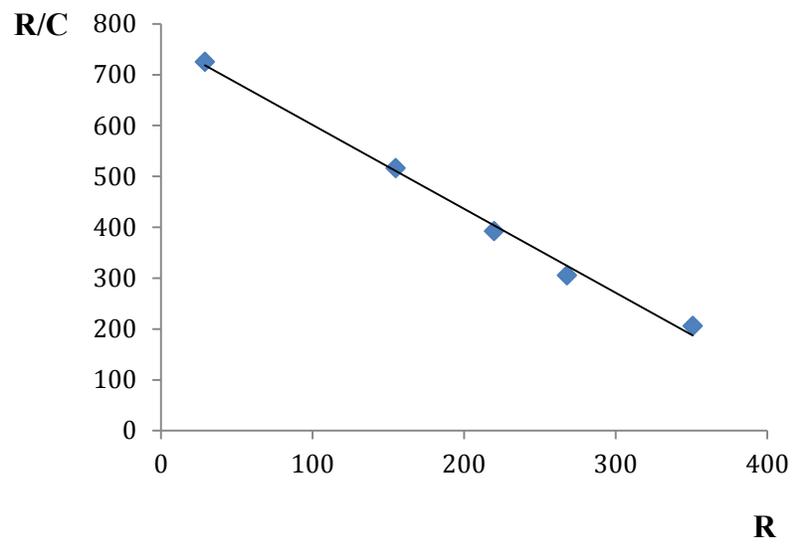
Tyr50Ala



Asn34Asp



CS-35 F_{ab}



A7. Homology Results for CS-40 and 906.4321 by Swiss modeling

CS-40

Workunit: P003616 CS40 - Overview



Print/Save this page as 

Model Summary



Model information:

Modelled residue range: 9 to 131
Based on template: [3hnsH] (2.00 Å)
Remark: No search for template was performed.
Only user specified template was used for modelling.
Sequence Identity [%]: 43.9
Evalue: 3.10e-44

Quaternary structure information: [\[details\]](#)

Template (3hns): HETERO DIMER
Model built :SINGLE CHAIN

Ligand information: [\[details\]](#)

Ligands in the template: BXX: 1, BXY: 1.
Ligands in the model: none.

Quality information: [\[details\]](#)

QMEAN Z-Score: -1.61 

[logs:](#) [\[Templates\]](#) [\[Alignment\]](#) [\[Modelling\]](#)

[display model:](#) as [\[pdb\]](#) - as [\[DeepView project\]](#) - in [\[AstexViewer\]](#)

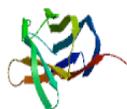
[download model:](#) as [\[pdb\]](#) - as [\[Deepview project\]](#) - as [\[text\]](#)

Workunit: P000002 CS40-L - Overview



Print/Save this page as 

Model Summary



Model information:

Modelled residue range: 27 to 137
Based on template: [3hnsL] (2.00 Å)
Remark: No search for template was performed.
Only user specified template was used for modelling.
Sequence Identity [%]: 54.46
Evalue: 1.50e-21

Quaternary structure information: [\[details\]](#)

Template (3hns): HETERO DIMER
Model built :SINGLE CHAIN

Ligand information: [\[details\]](#)

Ligands in the template: AXR: 1.
Ligands in the model: none.

Quality information: [\[details\]](#)

QMEAN Z-Score: -2.97 

[logs:](#) [\[Templates\]](#) [\[Alignment\]](#) [\[Modelling\]](#)

[display model:](#) as [\[pdb\]](#) - as [\[DeepView project\]](#) - in [\[AstexViewer\]](#)

[download model:](#) as [\[pdb\]](#) - as [\[Deepview project\]](#) - as [\[text\]](#)

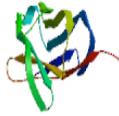
906.4321

Workunit: P000004 906.4321 - Overview



Print/Save this page as 

Model Summary



Model information:

Modelled residue range: 24 to 138
Based on template: [3hnsL] (2.00 Å)
Remark: No search for template was performed.
Only user specified template was used for modelling.
Sequence Identity [%]: 55.65
Evalue: 4.20e-24

Quality information: [\[details\]](#)

QMEAN Z-Score: -2.34 

Quaternary structure information: [\[details\]](#)

Template (3hns): HETERO DIMER
Model built :SINGLE CHAIN

Ligand information: [\[details\]](#)

Ligands in the template: AXR: 1.
Ligands in the model: none.

[logs](#): [\[Templates\]](#) [\[Alignment\]](#) [\[Modelling\]](#)
[display model](#): as [\[pdb\]](#) - as [\[DeepView project\]](#) - in [\[AstexViewer\]](#)
[download model](#): as [\[pdb\]](#) - as [\[Deepview project\]](#) - as [\[text\]](#)

Workunit: P000003 906.4321 - Overview



Print/Save this page as 

Model Summary



Model information:

Modelled residue range: 9 to 129
Based on template: [3hnsH] (2.00 Å)
Remark: No search for template was performed.
Only user specified template was used for modelling.
Sequence Identity [%]: 66.94
Evalue: 1.12e-34

Quality information: [\[details\]](#)

QMEAN Z-Score: -1.72 

Quaternary structure information: [\[details\]](#)

Template (3hns): HETERO DIMER
Model built :SINGLE CHAIN

Ligand information: [\[details\]](#)

Ligands in the template: BXX: 1, BXY: 1.
Ligands in the model: none.

[logs](#): [\[Templates\]](#) [\[Alignment\]](#) [\[Modelling\]](#)
[display model](#): as [\[pdb\]](#) - as [\[DeepView project\]](#) - in [\[AstexViewer\]](#)
[download model](#): as [\[pdb\]](#) - as [\[Deepview project\]](#) - as [\[text\]](#)

A8. Table of chemical shifts assignment of Ara6

Rademacher, C.; *et. al. J. Am. Chem. Soc* **2007**, *129*, 10489.

Ring	H1	H2	H3	H4	H5	H5'	(CH ₂) ₈ NH ₂
A	4.96	4.08	4.04	4.18	3.89	3.81	1.22, 1.56, 1.6, 2.9, 3.45, 3.61
B	5.14	4.32	4.12	4.34	3.95	3.90	
C	5.21	4.23	4.15	4.10	3.88	3.75	
D	5.27	4.23	4.14	4.05	3.87	3.75	
E	5.17	4.18	4.09	3.92	3.81	3.71	
F	5.16	4.17	4.08	3.94	3.82	3.71	

A9. Sequences of scFv fragments and alignments

Alignments from: EMBOSS Needle, www.ebi.ac.uk

CS-35 scFv gene sequence:

```
GAAGTTCAACTGCAACAATCGGGCACCGTTCTGGCTCGTCCGGGCACCTCCGTCAAAAATGTCGTGTAAAGCGTCGG
GTTATTCGTTTACCAACTATTGGATGCATTGGGTCAAACAGCGCCCGGGCCAAGGTCTGGAATGGATTGGTTCTAT
CTATCCGGGCAACAGCGATACCAACTACAAACAGAAATTC AAGGGTAAAGCAAAAACCTGACCGCTGTCACGAGTGC
GTCCACCGCCTACATGGAAGTGAACAGCCTGACGAATGAAGACAGTGCAGTTTATTACTGCACCCGTTTTGGCAAT
TATGTCCCCTCGCTTACTGGGGTCAGGGTACCCTGGTGACGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CGGTGGCGGTTTCAGATATTTCAGATGACCCAAACCACGAGCTCTCTGTCAGCGTCGCTGGGTGATCGTGTACGAT
TGGCTGTCGCGCCTCGCAGGACATCGGTAGCTATCTGAACTGGTACCAGCAAAAACCGGATGGCGCAGTGCCTCT
GCTGATCTATTACACCTCCCGCCTGCATTTCAGGTGTTCCGTCGCGTTTTAGCGGCTCTGGTAGTGGCACCCACTTCA
GCCTGACGATTTCTAATCTGGAACAGGAAGATATCGGCACCTATTTCTGCCACCAAGACACGAAACCGCCGTACAC
CTTCGGTTCGGGCACCAAACTGGAATCAA
```

CS-35 F_{ab} gene sequence

```
GAAGTCCAACGTCAGCAATCGGGCACCGTTCTGGCACGTCCGGGCACCTCCGTCAAAAATGTCTTGTAAGC
ATCAGGCTATTCGTTACCAACTATTGGATGCATTGGGTGAAACAGCGCCCGGGCCAAGGTCTGGAATGG
ATTGGTAGCATCTATCCGGGCAACTCTGATACCAACTACAAACAGAAATTC AAGGGTAAAGCGAAACTGA
CCGCAGTCACGTCAGCATCGACCGCATATATGGAAGTGAACAGTCTGACGAATGAAGACTCCGCGGTTTA
TACTGCACCCGTTTTGGCAATTATGTCCCCTCGCCTACTGGGGCCAGGGTACCCTGGTACGGGTGTC
GCGGCCACCACGACCGCACCGTCCGGTCTACCCGCTGGTGCCGGTTGCTCGGATACCAGCGCAGTCTG
TGACGCTGGGTTGTCTGGTTAAAGGCTATTTTCCGGAACCGGTTACCGTCAAATGGAACACTACGGTGCAC
TACTGAGTTCCGGCGTCCGCACGGTGTTCATCGGTTCTGCAGTCAGGTTTCTATTTCGCTGAGCTCTCTGGT
GACTGTTCCGAGTTCCACGTGGCCGTTCTCAAACCGTTATTTGCAATGTCGCACATCCGGCTAGCAAAACCGAAC
TGATTAACGTCATCGAACCGCGCATTCGGGCGGTGGCGGTTCTGGCGGTGGCGGTAGTGGCGGTGGCGG
TTCCGATATCCAGATGACCCAAACGACCTCATCGCTGAGCGCATCTCTGGGTGATCGTGTACCATTGGC
TGTCGCGCTTACAGGACATCGGTTTCGATCTGAACTGGTACCAGCAAAAACCGGATGGCGCGGTTTCGTC
TGCTGATCTATTACACCAGCCGCTGCATTCTGGGTGTGCCGAGTCGTTTTAGTGGCTCCGGTTCAGGCAC
GCACTTCTCCCTGACCATTTCAAATCTGGAACAGGAAGATATCGGCACCTATTTTGGCCACCAAGACACG
AAACCGCGTACACCTTCGGTAGCGGCACGAAACTGGAATTAACGTCAGATGCAGCTCCGACCGTGA
GCATCTTCCCGCCGAGCTCTGAACAGCTGACGTCGGGCGGTGCTAGCGTGGTTTGTCTTCTGAACAAC
CTACCCGAAAGATATCAACGTTAAATGGAATTCGACGGTCCGAACGCCAGAACGGCGTGTGAACAGC
TGGACCGATCAAGACTCTAAAGATAGTACGTACTCCATGAGTTCACGCTGACCCCTGACGAAAGACGAAT
ATGAACGTCATAATAGTTACACCTGTGAAGCCACGCACAAAACAGCACGTCACCGATTGTCAAAGCTT
CAACCGCAACGAATGTCACCACCATCATCACTAATAA
```

Alignment of CS-35 scFv and mutant sequences

His35Ala

```
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS      50
  |||
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMAWVKRPGQGLEWIGS      50
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
```

Trp33Ala

```
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS      50
  |||
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYAMHWVKRPGQGLEWIGS      50
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
```

Tyr50Ala

```
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS      50
  |||
EMBOSS_001      1 EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS      50
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001     51 IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG     100
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    101 NYVPFAYWGQGLTVTVGGGGSGGGGSGGGGSDIQMTQTSSLSASLGDRV     150
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    151 TIGCRASQDIGSYLNWYQQKPDGAVRLLIYATSRHSGVPSRFSGSGSGT     200
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
EMBOSS_001    201 HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK      238
  |||
```

Tyr98Ala

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFG	100
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
		.	
EMBOSS_001	101	NAVVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

Tyr98Phe

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFG	100
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
		:	
EMBOSS_001	101	NFVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

Phe95Ala

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRFG	100
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTRAG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

Asn97Asp

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

Ser50Ala

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGA	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

Asn58Ala

EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	1	EVQLQQSGTVLARPGTSVKMSCKASGYSFTNYWMHWVKRPGQGLEWIGS	50
EMBOSS_001	51	IYPGNSDTNYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	51	IYPGNSDTAYKQKFKGKAKLTAVTSASTAYMEVNSLTNEDSAVYYCTFRG	100
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	101	NYVPFAYWGQGLTVTVGGGGSGGGSGGGSDIQMTQTSSLSASLGDRV	150
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	151	TIGCRASQDIGSYLNWYQQKPDGAVRLLIYYTSRLHSGVPSRFSGSGSGT	200
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238
EMBOSS_001	201	HFSLTISNLEQEDIGTYFCHQDTKPPYTFGSGTKLEIK	238

A10. Control experiments for STD NMR

STD spectrum of Ara6 (0.5 mM, *d*-ammonium acetate in D₂O) on-resonance: -0.2 ppm shows no signal below 1 ppm.

