

**Phage Display Derived Fragment-based Discovery of Antigens for Antibodies
associated with Mycobacterial Infections**

by

Ying Chou

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

Master of Science

Department of Chemistry
University of Alberta

© Ying Chou, 2017

Abstract

Tuberculosis (TB) remains a global health problem. Incorrect or missed serodiagnosis for TB, especially in resource-limited settings, is one of the main reasons for TB epidemics. Clinically, serological tests that use circulating antibodies are attractive because antibodies in serum are present at higher concentration than disease biomarkers, making detection of disease-specific antibodies easier. However, the accuracy of a serological test is highly dependent on the antigens used in the test; existing serological tests for TB provide inconsistent specificities and sensitivities. This issue is further aggravated by the heterogeneity and complexity of the glycolipid antigen across mycobacteria species, which makes rational design of discriminatory antigens difficult.

This thesis approaches these challenges, i.e. the need for simpler and more efficient antigens with high specificity for TB diagnoses, by combining fragment-based discovery with phage display. Through anchoring each genetically-encoded displayed peptide on phage to a small molecule fragment, I constructed glycopeptide libraries that demonstrated their application for the selection of promising antigens for antibodies associated with mycobacterial infections. By using an anti-LAM antibody CS-35 as a model target, I discovered novel hexasaccharide-peptide conjugates with K_D value 10-fold better than that of the hexasaccharide itself (Chapter 2). The selectivity and affinity of these conjugates enhanced modestly when constructed in multivalent presentation and tested in multivalent binding assays. Interestingly, when the hexasaccharide on these conjugates were truncated to monosaccharides, their binding affinity drastically diminished, suggesting that some, if not all, monosaccharides in the hexasaccharide must be present to elicit constructive antigen-antibody interactions (Chapter 3). Consequently,

I demonstrated a site-selective dual modification strategy to create the first phage-displayed macrocyclic glycopeptide library that present two carbohydrate moieties per peptide. The peptide library can be glycosylated first using sodium periodate cleavage on N-terminal serine followed by oxime ligation; the second glycosylation was achieved through alkylation on disulfides using a dichloro-oxime derivative. These libraries show potential for use in fragment-based ligand discovery, and selection for TB-specific antibodies. Finally, I developed a deep-panning system that combines single-round selection with deep sequencing to afford rapid and unbiased selection. The system has been validated on model target ConA and antiFLAG antibody, and will strongly complement the fragment-based ligand discovery to identify disease-specific antigens using serum from diseased patients.

Preface

Portion of **Chapter 1** has been submitted to publication from H. Anany, Y. Chou, S. Cucic, S. Evoy, R. Derda, and M.W. Griffiths, “From Bits and Pieces to Whole Phage to Nanomachines: Pathogen Detection Using Bacteriophage,” *Annual Review of Food Science and Technology*, Vol. 8 (publication date April 2017). I reviewed the current phage display technology that facilitates phage-derived detection system for pathogens. **Figure 1-3 B** is reused with permission from S.A. Goodchild, H. Dooley, R.J. Schoepp, M. Flajnik, and S.G. Lonsdale, “Isolation and characterisation of Ebolavirus-specific recombinant antibody fragments from murine and shark immune libraries,” *Molecular Immunology*, **2011**, 48(15-16), 2027-2037. Copyright © Elsevier; **Figure 1-3 E** is reused with permission from I.B. Sorokulova, E.V. Olsen, I. Chen, B. Fiebor, J.M. Barbaree, V.J. Vodyanoy, B.A. Chin, V.A. Petrenko, “Landscape phage probes for *Salmonella typhimurium*,” *Journal of Microbiological Methods*, **2005**, 63(1), 55-72. Copyright © Elsevier. **Figure 1-4 A** is reused with permission from J. McElhiney, M. Drever, L.A. Lawton, and A.J. Porter, “Rapid isolation of a single-chain antibody against the cyanobacterial toxin microcystin-LR by phage display and its use in the immunoaffinity concentration of microcystins from water,” *Appl Environ Microbiol.*, **2002**, 68(11), 5288-5295. Copyright © American Society for Microbiology. **Figure 1-4 B** is reused with permission from C.E. Chan, B.Z. Zhao, A. Cazenave-Gassiot, S.W. Pang, A.K. Bendt, M.R. Wenk, P.A. MacAry, and B.J. Hanson, “Novel phage display-derived mycolic acid-specific antibodies with potential for tuberculosis diagnosis,” *J. Lipid Res.*, **2013**, 54(10), 2924-2932. Copyright © American Society for Biochemistry and Molecular Biology, Inc. **Figure 1-4 C** is reused with permission from X. Lu, P. Weiss,

and T. Block, “A phage with high affinity for hepatitis B surface antigen for the detection of HBsAg,” *Journal of Virological Methods*, **2004**, 119(1), 51-54. Copyright © Elsevier. **Figure 1-4 D** is reused with permission from M. Mourez, R.S. Kane, J. Mogridge, S. Metallo, P. Deschatelets, B.R. Sellman, G.M. Whitesides, and R.J. Collier, “Designing a polyvalent inhibitor of anthrax toxin,” *Nat Biotechnol.*, **2001**, 19(10), 958-961. Copyright © Nature Publishing Group. **Figure 1-5 C** is reused from G.A. Kouzmitcheva, V.A. Petrenko, and G.P. Smith, “Identifying diagnostic peptides for Lyme Disease through epitope discovery,” *Clin Diagn Lab Immunol.*, **2001**, 8(1), 151-160. Copyright © American Society for Microbiology. **Figure 1-5 D** is reused from G.J. Xu, T.Kula, Q. Xu, M.Z. Li, S.D. Vernon, T. Ndung’u, K. Ruxrungtham, J. Sanchez, C. Brander, R.T. Chung, K.C. O’Connor, B. Walker, H.B Larman, and S.J. Elledge, “Comprehensive serological profiling of human populations using a synthetic human virome,” *Science*, **2015**, 348(6239). Copyright © Science.

Chapter 2 is based on non-published work initiated by R. Derda. I was responsible for optimization of chemical modifications on phage and all the phage selection. B.M. Paschal and C.J. Noren provided the phage library used in the first round of selection. M. Joe synthesized the hexasaccharide aminoxy-derivatives and P. Kitov provided glucose and galactose aminoxy-derivatives. Y. Yang and I synthesized the peptide sequences of the putative hits using SyntArray technology. I developed the synthesis of glycopeptides on paper and fluorescence binding assay on paper. E. Kitova and J. Klassen were involved with the ESI-MS binding assay. I synthesized all glycopeptide hits conjugated to BSA. T. Lowary and P. Kahn were responsible for

printing glycopeptides microarrays. B. Zhang provided F_{ab} and IgG of anti-LAM antibodies. M. Miskolzie was involved in NMR measurements.

Chapter 3 is based on non-published work initiated by R. Derda. The α -arabinose monosaccharide was provided by R. Brunton. I synthesized all glycopeptides and glycopeptide-BSA conjugates. T. Lowary and P. Kahn were involved in printing glycopeptides microarrays. I was responsible for testing glycopeptide microarrays.

Chapter 4 is based on non-published work initiated by R. Derda. P. Kitov provided the mannose derivatives and R. Brunton synthesized the α -arabinose monosaccharide. I was responsible for synthesizing the dichloro-oxime-derivative, optimizing chemical modifications on phage, and selection of phage libraries on antibodies. Blocking phage and monoclonal SWYD phage were provided by N. Bennet. D. Ferrer was involved in phage selection against model target ConA. Several techniques described in this chapter have been published in K.F. Tjhung, F. Deiss, J. Tran, Y. Chou, R. Derda, "Intra-domain phage display (ID-PhD) of peptides and mini-domain proteins censored from canonical pIII phage display", **2015 *Front. Microbiol.* 6:340**. Specifically, I used the published phage displaying FLAG sequence and blue-white screen for the optimization of the selection.

Acknowledgements

I would like to express my sincere gratitude to my colleagues, friends, and my supervisor who have supported me throughout my academic career.

Foremost, I would like to thank my thesis advisor Dr. Derda for his continuous support and guidance to steer me in the right direction. His patience, enthusiasm, mentorship, and knowledge have inspired me immensely and helped me overcome numerous obstacles through my research.

I would also like to thank Prof. Todd Lowary and Prof. John Klassen for their selfless encouragement and insightful comments. As my collaborators, Prof. Lowary and Prof. Klassen have contributed to many aspects of my projects. I am thankful for Prof. Julie Gibbs-Davis and Prof. Jillian Buriak for being my committee members. They have given me valuable advice on my professional career over the past two years. I appreciate Prof. Florence Williams for being my one-arm's length examiner during my defense.

This project would not be possible without my colleagues and friends who were always passionate and approachable. Simon, Reza, Daniel, Frédérique, and Pavel have imparted tremendous input and expertise that greatly assisted my research. Their support on both professional and personal levels has guided me through the hardest times during my studies.

Finally, I wish to acknowledge the fellowship support from the Faculty of Graduate Studies and Research from the University of Alberta and the funding support from Alberta Glycomics Centre. This accomplishment would not have been possible without these supports.

Table of Contents

Abstract.....	ii
Preface.....	iv
Acknowledgements	vii
List of Tables	xii
List of Figures	xiii
List of Abbreviations	xv
Chapter 1: Introduction	1
1.1 Current detection system for pathogens and toxins	1
1.2 Phage display-derived detection system.....	1
1.3 Kinds of phage display system	3
1.3.1 Peptides	4
1.3.2 Antibodies	5
1.4 Selection targets	5
1.4.1 Selections that target whole pathogen organism.....	8
1.4.2 Selections that target fragments of pathogens	14
1.4.3 Selections that target antibodies associated with pathogenic infections of diseases.....	18
1.5 Prospect for phage display-derived detection reagent	22
1.6 Thesis overview	25
Chapter 2: Genetically-encoded fragment-based (GE-FBD) discovery of glycopeptide ligands for antibodies associated with mycobacterial infections.....	27
2.1 Introduction.....	27
2.2 Results and discussion	32
2.2.1 Generation of glycopeptide libraries on phage	32
2.2.2 Selection and sequencing.....	32
2.2.3 Glycopeptides testing using paper-based binding assay	35
2.2.4 Glycopeptides testing using ESI-MS assay	40
2.2.5 Glycopeptides testing in microarray format using monoclonal antibodies	45

2.3	Conclusion	45
2.4	Experimental procedures	47
2.4.1	Materials and general information	47
2.4.2	Chemical modification phage library.....	58
2.4.3	Phage selection.....	49
2.4.3.1	First round of selection	49
2.4.3.2	Second round of selection.....	50
2.4.3.3	Third round of selection.....	50
2.4.4	Phage amplification and preparation for sequencing.....	50
2.4.5	Analysis of the deep-sequencing data.....	52
2.4.6	Synthesis of glycopeptides on paper.....	53
2.4.7	TRITC labeling of mAb CS-35.....	54
2.4.8	Paper-based fluorescence binding assay	55
2.4.9	Synthesis of peptides on solid support.....	55
2.4.10	Synthesis of Ara ₆ -peptide conjugates	57
2.4.10.1	Oxidation of N-terminal serine	57
2.4.10.2	Ligation of Ara ₆ -ONH ₂	57
2.4.11	Synthesis of Glu-peptide conjugates.....	58
2.4.12	NMR to determine product yield	59
2.4.13	ES-MS binding measurements.....	59
2.4.14	Synthesis of glycopeptide-squaramide	61
2.4.15	Synthesis of glycopeptide-BSA conjugates.....	61
2.4.16	Preparation of microarrays.....	62
2.4.17	Microarray testing.....	62
2.4.18	Array scanning and data extraction.....	63
2.4.19	Antigen binding curves to mAbs (CS-35, 906.4321, CS-40)	64
Chapter 3: Structure-activity relationship between truncated glycopeptides and anti-LAM antibodies associated with mycobacterial infections		65
3.1	Introduction.....	65
3.2	Hypothesis and objective.....	67
3.3	Experimental design	68

3.4	Results and discussion	71
3.4.1	Synthesis of glycopeptides.....	71
3.4.2	ESI-MS of glycopeptides binding to monoclonal anti-LAM CS-35	71
3.4.3	Testing the binding of monoclonal to glycopeptides in a microarray format.....	72
3.5	Conclusion	75
3.6	Experimental procedures	76
3.6.1	Materials and general information.....	76
3.6.2	Synthesis of glycopeptide conjugates	77
3.6.3	ESI-MS binding measurements	78
3.6.4	Synthesis of glycopeptide-squaramide	79
3.6.5	Synthesis of glycopeptide-BSA conjugates.....	80
3.6.6	Preparation of microarrays.....	80
3.6.7	Microarray testing.....	81
3.6.8	Array scanning and data extraction.....	81
3.6.9	Antigen binding curves to mAbs (CS-35, 906.4321, CS-40)	82
Chapter 4: Deep-panning of phage-displayed macrocyclic peptide libraries for antibodies associated with mycobacterial infections		83
4.1	Introduction.....	83
4.2	Objective and hypothesis.....	88
4.3	Experimental design	91
4.4	Results and discussion	92
4.4.1	Generation of Ara-M-Ara libraries on phage.....	92
4.4.2	Generation of input libraries for selection	95
4.4.3	Sequential pull-down assays to select binders for ConA, CS-35, 906.4321, and anti-FLAG	96
4.4.4	Analysis of Illumina sequencing.....	99
4.5	Conclusion	105
4.6	Experimental procedures	106
4.6.1	Materials and general information.....	106
4.6.2	Synthesis of dichloro-oxime- α -arabinose (DCO- α Ara).....	108
4.6.3	Reaction of SHCVWWDC with DCO- α Ara.....	109

4.6.3.1 Synthesis of oxidized SHCVWWDC	109
4.6.3.2 Synthesis of α Ara-DCVWWDC	109
4.6.3.3 Synthesis of α Ara-M- α Ara.....	110
4.6.4 Generation of α Ara-M- α Ara libraries on phage	110
4.6.4.1 Bulk amplification of SXCX ₃ C and N- <i>SerX</i> ₇ phage library.....	110
4.6.4.2 Generation of α Ara-XCX ₃ C library	111
4.6.4.3 Modifications of α Ara-XCX ₃ C library with DCO- α Ara	112
4.6.5 Preparation of input mixture for selection	112
4.6.6 Phage selection and extraction of ssDNA.....	113
4.6.7 Preparation for Illumina sequencing.....	115
4.6.8 Analysis of the seep-sequencing data	116
Chapter 5: Conclusion and outlook	117
5.1 Conclusion	117
5.2 Future directions.....	120
References.....	122
Appendix A: Supporting information for chapter 2.....	143
Appendix B: Supporting information for chapter 3.....	211
Appendix C: Supporting information for chapter 4.....	233

List of Tables

Table 1. Targets that have been selected against using phage display, the libraries used for selection, and the potential application of the discovered ligands	9
Table 2. K_D of ligands to CS-35 and 906.4321 F _{ab} measured by ESI-MS.....	42

List of Figures

Figure 1-1. Panning with peptide library displayed on pIII.....	4
Figure 1-2. The available display technology to develop efficient detection systems for specific targets	8
Figure 1-3. Phage libraries targeted for whole pathogen.....	13
Figure 1-4. Phage libraries targeted for fragments of pathogens or toxins.....	16
Figure 1-5. Phage libraries targeted for antibodies associated with disease.....	20
Figure 2-1. The design of glycopeptide antigen Ara ₆ -peptides for anti-LAM antibody CS-35.	31
Figure 2-2. Validation of chemical modifications on peptide, phage clones and phage libraries	33
Figure 2-3. Overview of selection and recovery of phage after multiple rounds of selection	34
Figure 2-4. Detailed selection scheme and post-selection analysis	36
Figure 2-5. Binding of glycopeptides on paper to CS-35	37
Figure 2-6. Optimization of oxidation conditions for peptide immobilized on paper	38
Figure 2-7. Optimization of glycopeptide formation on paper	39
Figure 2-8. ESI-MS measurements of discovered Ara ₆ -peptides	43
Figure 2-9. Ara ₆ -peptides binding in multivalent microarray assay	45
Figure 2-10. Analysis of the deep-sequencing data	53
Figure 3-1. Truncation of hexasaccharide to enhance specificity.....	68
Figure 3-2. Observation of binding properties of Ara ₆ and Ara ₆ -peptides to antibodies on monovalent and multivalent assays	69

Figure 3-3. Synthesis scheme of Ara-peptides for ESI-MS measurements and microarray printing.....	72
Figure 3-4. ESI-MS measurements of Ara-peptides to CS-35 F _{ab}	73
Figure 3-5. Micorarray testing of Ara-peptides on E and H surface.....	74
Figure 3-6. <i>De novo</i> discovery of Ara-X ₇ binders for CS-35	76
Figure 4-1. Overview of the workflow of deep-panning system	84
Figure 4-2. Macrocycle formation by bis-glycosylated peptide macrocycle.....	90
Figure 4-3. Validation of Ara-M-Ara reaction on peptide.....	93
Figure 4-4. Ara-M-Ara reaction on phage.....	95
Figure 4-5. Glycopeptide libraries used for sequential pull-down assays	98
Figure 4-6. Post-selection analysis on model target ConA.....	100
Figure 4-7. Post-selection analysis on model target aFLAG	101
Figure 4-8. Tracing monoclonal phage with control sequences throughout each selection experiments.....	102
Figure 4-9. Post-selection analysis on Ara-XCX ₃ C selected against CS-35	104
Figure 4-10. SDS-PAGE gel of selection supernatant after capturing antibody-phage mixture on protein G-coated beads.....	114

List of Abbreviations

AOB	Aminooxy-biotin
Ara	Arabinose
Ara ₆	<i>Methyl 5-O-<math>\{3,5\text{-di-O-(2-O-β-D-arabinofuranosyl)-α-D-arabinofuranosyl)-α-D-arabinofuranoside}</math></i>
BCG	Bacillus Calmette–Guérin
BIA	Biotin-PEG ₂ -iodoacetamide
BoNT	botulinum neurotoxin
BSA	Bovine serum albumin
ConA	Concanavalin A
CRD	Carbohydrate recognition domain
DCA	1,3-Dchloroacetone
DCO	1,3-Dichloro-oxime derivatives
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)
ELISA	enzyme-linked immunosorbent assay
ESI	Electrospray ionization
F _{ab}	Antigen-binding fragment
FACS	Fluorescence-activated cell sorting

Fmoc	9-Fluorenylmethoxycarbonyl
FTS	Fluorescein-5-thiosemicarbazide
F _v	Fragment variable
Gal	D-Galactose
GE-FBD	Genetically encoded fragment-based discovery
Glu	Glucose
GSH	Glutathione
h	Hour
HBsAg	Hepatitis B surface antigen
HBTU	N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uranium hexafluorophosphate
HEPES	N-(2-hydroxyethyl)piperazine-N'-ethanesulfonic acid
HIV	human immunodeficiency virus
HPAI	Highly pathogenic avian influenza
HRMS	High-resolution mass spectrometry
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IgNAR	Immunoglobulin new antigen receptor antibodies
KD	Dissociation constant
kDa	Kilo Dalton
LAM	Lipoarabinomannan
LC	Liquid chromatography
MACS	Magnetic-activated cell sorting
MALDI	Matrix-assisted laser desorption ionization
Man	D-Mannose

MCW	Mycobacterial cell wall
MHC	Major histocompatibility complex
min	Minute
MOPS	3-(N-morpholino)propanesulfonic acid
MQ	MilliQ
MS	Mass spectrometry
MWCO	Molecular weight cut-off
NMR	Nuclear magnetic resonance
OD	Optical density
PBS	Phosphate buffer saline
PCR	polymerase chain reaction
PFUs	Plaque forming units
PHALISA	Phage-linked immune-absorbent assay
PhIP	Phage immunoprecipitation
PMT	Photomultiplier tube (PMT)
ProP-PD	Proteomic peptide-phage display
PTM	Post-translational modification
Rha	Rhamnose
RiPPs	Ribosomally-made post translationally modified peptides
RNA	ribonucleic acid
RP-HPLC	Reverse phage high-performance liquid chromatography
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
scFv	Single chain fragment variable
sdAb	Single domain antibodies

SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TB	Tuberculosis
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
TRITC	Tetramethylrhodamine
tRNA	transfer RNA
UPLC	Ultra performance liquid chromatography
VSG	Variant surface glycoproteins

Chapter 1: Introduction

1.1 Current detection system for pathogens and toxins

The mainstay of detection of pathogens and toxins can be divided into phenotypic assays, such as culture-based detection or microscopy, and molecular bioassays.¹⁻⁵ Culture- and microscopy-based assays for detection of pathogens have been considered to be the most robust and reliable methods of detection. They are often the core technology of the approved methods for the detection of food- and water-borne pathogens. While phenotypic assays often provide the most robust result, there is a trade-off between the length, complexity, cost of an assay, and the rate at which it yields results. With the increasing demand of faster results, the ultimate challenge in diagnostics research is replacing or upgrading current culture-based and microscopic methods with the help of molecular assays with equivalent sensitivity, specificity, yet with reduced complexity and cost.

1.2 Phage display-derived detection system

Phage-display technology has been one of the work horses for the development of the components such as new peptides, proteins, and antibodies for molecular bioassays. When compared to traditional culture-based methods, which are often time-consuming, due to the slow growth of microorganisms, phage display-derived molecular assays offer rapid diagnostics. In situations where growth-based assays are complicated by the phenotypic heterogeneity of bacteria or the biochemical composition of patient sample,

molecular methods should be consulted. For instance, culture analysis of periodontal samples to detect *Tannerella forsythia* can be highly variable because not all viable bacteria found in periodontal pockets are culturable and the bacteria count does not reflect the number of viable bacteria present *in vivo*.⁶ In addition, the morphology of the colonies of *T. forsythia* varies with growth conditions, making identification difficult.⁷ In contrast, PCR or antigen-antibody method has shown to be more promising for the detection of *T. forsythia*.⁸⁻⁹

Molecular diagnostics can potentially replace culture-based assays or whole animals in assays that quantify the presence of toxins produced by pathogens. For example, mouse-based bioassays that detect 0.03 ng of botulinum toxin require 1 to 2 days.¹⁰ As an alternative, Janda and coworkers¹¹ used peptide display technology to identify a peptide ligand for an ELISA that can detect 1 pg/mL of botulinum neurotoxin (BoNT) within 3 h. This inexpensive capture ELISA system can be used for rapid, sensitive, and highly specific BoNT detection in contaminated foods.

Screening for ligands *via* phage display can be conceptually summarized as a pull-down of target-binding ligands from large libraries of ligands fused to the coat protein of phage. These libraries are generated readily by inserting the genes encoding peptides or proteins into the genes that encode phage coat protein (Figure 1-1). A routine screening strategy employs serial, multi-step enrichment schemes, starting from targets immobilized on a solid support. After an extensive wash to remove weak binders, phage clones retained by the target are eluted off the targets and re-amplified in bacteria to yield a focused library of ligands. This focused library is subjected to another round of selection. Through multiple rounds that alternate panning and amplification processes,

binders with high specificity and affinity can be selected and enriched. Selected phage can be sub-cloned and analyzed by sequencing of phage DNA.

Screening often requires no prior knowledge of the nature of the target nor its interaction with natural ligands; this technology, thus, can uncover ligands for targets that do not have conventional or known binders. For example, selection can incorporate whole pathogen cells as bait and bypass the need for purification or identification of receptors from the target. A similar strategy has been widely employed to develop ligands that bind to specific cell types in multicellular organisms (mouse, human), but the overview of these results is beyond the scope of this chapter and reviewed elsewhere.¹² This chapter introduces briefly the current available display format on bacteriophage. More detailed overview of the display platforms, display methods, and selection techniques can be consulted in previous reviews.¹³⁻¹⁶

1.3 Types of phage display system

Molecular components can be displayed on lytic phage—T7^{17, 18, 19-21}, T4²²⁻²⁴, λ ²⁵²⁶⁻²⁹—or non-lytic filamentous phage or phagemid based on f1, M13, or fd.³⁰ The different life cycles pose constraints on the length and sequences of polypeptide chains that can be displayed on the phage without interfering with life-cycle of phage or assembly of the phage particle. Conceptually, the displayed molecules on phage can be subdivided into two categories: short peptides (7-10 amino acids, usually no more than 20 amino acids), proteins or protein domains among which antibody fragments are most commonly displayed entities.

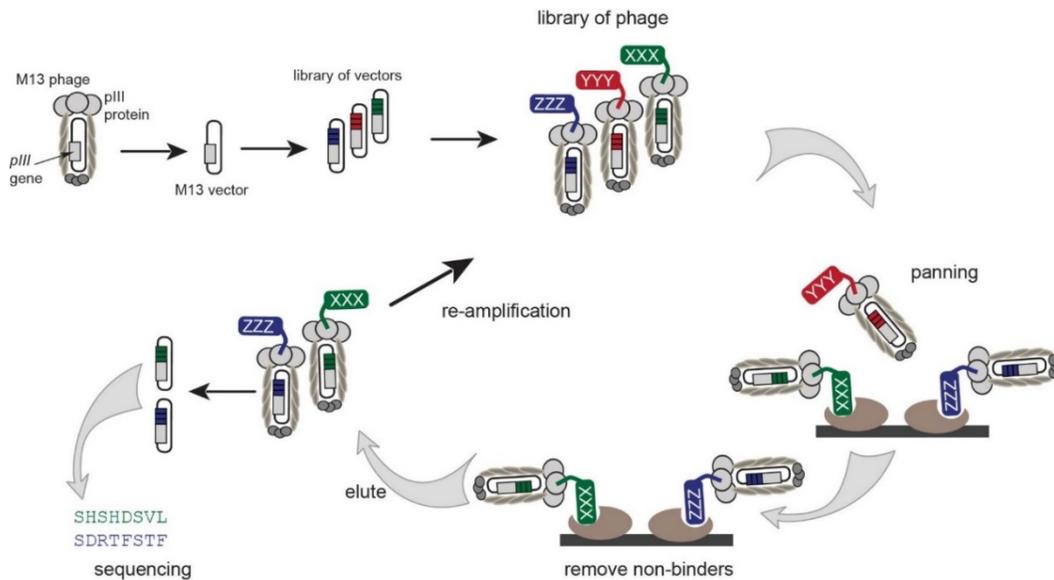


Figure 1-1. Panning with peptide library displayed on pIII. Phage libraries are incubated with target-coated surface or bead. Extensive washing removes the unbound phage. Bound phage are eluted and subsequently amplified in *E. coli*. The whole selection cycle can be repeated 2-3 times. Phage at each round can be processed to extract and sequence DNA to identify putative binders.

1.3.1 Peptides

Display on filamentous phages fd/M13 historically was the first display system and it has been investigated, used and modified the most. Random peptide libraries are most commonly displayed on pIII minor coat protein or pVIII major coat protein.³¹⁻³³ Modification on pVIII benefits from local functionalities of peptides and global functions from entire surface landscape.³⁴ It also allows displaying anywhere from 2 to 100 copies of molecules per phage. In contrast, display on pIII is limited to 2-5 copies. The pIII display peptide libraries are commercially available: Ph.D.-12 (12-mer), Ph.D.-7 (7-mer), Ph.D.-C7C, all of which are expressed at the N-terminus of pIII. Other peptide libraries exist and used for diagnostic discovery include C12C³⁵ and CX9C libraries of peptide disulfides that contain 12 and 9 residues; X6, X8, X15 libraries that contain linear 6-mer,

8-mer and 15-mer or their combination (X15CX). These libraries are not commercial but they can be constructed using readily available cloning vectors (fd-tet³⁶, f88.6³⁷). Other displays using pVII and pIX are reviewed and reported to circumvent protein folding challenging during amplification in bacteria host.³⁸ Larger, longer peptide fragments derived from specific repertoire can also be incorporated. One of the techniques, termed “cDNA display,” has been used to display proteins encoded in a specific genome and has been employed mostly in autoantigen discovery.³⁹

Because the displayed peptides are specified by the corresponding nucleotide inserts, the chemical diversity of polypeptides are restricted to 20 natural amino acids. To widen the chemical and structural space, many groups have sought for strategies that can integrate genetically encoded unnatural amino acids⁴⁰⁻⁴² or perform chemical modifications on displayed peptide libraries.⁴³⁻⁵¹ For example, oxime ligation of hydroxylamine derivative with aldehyde generated from the oxidation of terminal Ser-containing peptide libraries has been employed to discover ligands for carbohydrate-binding proteins,⁵² and it is used in Chapter 2 of my thesis for discovery of glycopeptide antigens.

1.3.2 Antibodies

Different antibody fragments are available in phage display technology, including Fab (antigen-binding fragment), Fv (fragment variable), scFv (single chain fragment variable) and modifications of thereof.⁵³ In some cases sdAb (single domain antibodies), which are devoid of light chains,⁵⁴ are used to discover small-domain binders; these domains can be then assembled into multivalent binders in order to increase avidity.⁵⁵

The sources of these antibodies repertoire can be classified into three: 1) immunized libraries: heavy and light chains are isolated from B cells of immunized animals; 2) synthetic libraries: heavy and light chains gene segments from human B cells are rearranged in vitro with the introduction of artificial complementary determining region of varying loop lengths; or 3) naïve libraries: heavy and light chains are cloned from the naïve IgM and IgG repertoire and randomly assembled to produce scFv.⁵⁶

Most often the antibody fragment is usually fused to the pIII DNA phagemid vector, which requires a wild-type helper phage to generate a fully assembled phage displaying the fused antibody fragment.⁵⁷ Antibody libraries are commonly generated to discover biomarkers, therapeutics, and alternatively as a method to express antibodies due to its much higher efficiency than the traditional hybridoma method.⁵⁸ Using semisynthetic combinatorial and synthetic human antibodies libraries, monoclonal antibodies of high specificity and affinity can be selected and produced for a wide range of protein targets. Strategies for constructing and randomizing various libraries of antibody fragments is an aggressively growing area pursued both in academic and industrial settings; complete overview of the state-of-the-art in this field is beyond the scope of this chapter and can be found elsewhere.⁵⁹⁻⁶⁰

1.4 Selection targets

Using the basic principle of aforementioned selection strategy, detection assays based on phage display technology can be developed for specific targets. The ultimate goal for an effective detection is to probe targets from environmental samples, food or

patients' samples (stool, blood, urine, wound, sputum, saliva, serum). Detection targets can be whole organisms, fragments derived from whole organisms, or serum antibodies that have been produced in response to specific pathogens. Theoretically, phage library screening and counter-screening condition can be optimized to discover binders for the pathogenic target and exhibit no cross-reactivity to closely related non-pathogenic organisms. Putative binders discovered by display technologies can be incorporated in biosensors, ELISA tests, or lateral flow devices. Based on the nature of targets and screening conditions, certain selection targets or phage libraries are more beneficial than others (Figure 1-2).

In this section, we discuss the advantages of detection reagents discovered *de novo* using phage display technology, and address selective cases that can be potentially integrated into detection platforms. A list of examples using phage-derived systems, organized according to target type, is summarized in Table 1. For some pathogens such as HIV, Hepatitis, or bacillus anthracis, numerous efforts have been made to develop diagnostics. A more comprehensive list of phage-display derived system is available at mimoDB 2.0,⁶¹ and the BDB database.⁶² Note that although the literature contains abundant successful examples of selection, the practitioners of selection technology often acknowledge that every target of interest brings new challenges to discovering sensitive and specific binders. As the literature is inherently skewed towards successful examples, it is difficult to poise our discussion with problematic and failed selections. Nevertheless, section 1.5 highlights some problems and emerging solutions in this area.

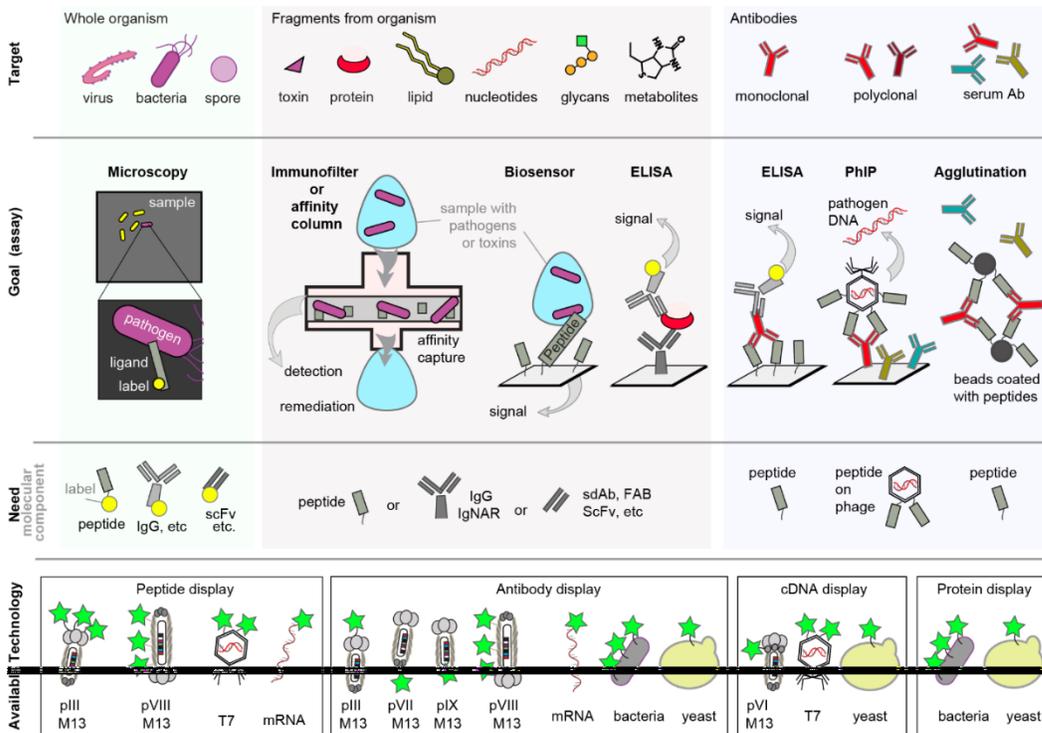


Figure 1-2. The available display technology to develop efficient detection systems for specific targets. scFv: single chain antibodies; sdAb: single domain antibody; IgNAR: immunoglobulin new antigen receptor; ELISA: Enzyme-linked immunosorbent assay; PhIP: Phage immunoprecipitation.

1.4.1 Selections that target whole pathogen organism

Whole organisms include bacteria, viruses, cells, tissues, and spores. Using whole organisms for screening allows selection for antigens that are not stable under isolation or immobilization conditions, and no prior knowledge or antigens is required.

A potentially clinically-useful approach to detect H5N1 virus was demonstrated by Ke and coworkers.⁶³ Three selected clones from human scFv phage libraries showed high affinity to highly pathogenic avian influenza (HPAI) H5N1 but no affinity for the other seasonal epidemic flu strains in ELISA (Figure 1-3A). These scFvs, with further development and experiments on sera, may be used in clinical diagnostic ELISA tests.⁶⁴

Table 1. Targets that have been selected against using phage display, the libraries used for selection, and the potential application of the discovered ligands

Pathogens	Library (target) ^{Ref}	Potential Application	
Cyanobacterial toxin microcystin	scFv (toxin) ⁶⁵⁻⁶⁶	Immunoaffinity column	
	scFv (toxin) ⁶⁷	Diagnostic	
Palytoxin	scFv (toxin) ⁶⁸	Diagnostic	
Tomato spotted wilt virus	scFv (virus) ⁶⁴	Diagnostic	
Human immunodeficiency virus (HIV)	X8, X15CX (Gp120) ⁶⁹ CX12C (monoclonal) ⁷⁰ Ph.D.-12 (Gp41) ⁷¹ X20 (gp120) ⁷² CX9C (polyclonal) ⁷³	Epitope mapping	
	Ph.D.-12 (polyclonal) ⁷⁴ scFv (gp160) ⁷⁵	Diagnostic	
	Ph.D.-12 (capsid) ⁷⁶ Ph.D.-12, Ph.D.-7, Ph.D.-C7C (Gp120) ⁷⁷	Inhibitor for infection	
	f3-15mer, X21 (gp120, gp41) ⁷⁸ Ph.D.-7 (protein) ⁷⁹ scFv (peptide region from protein) ⁸⁰	Therapeutic Candidate	
	<i>Mycobacterium tuberculosis</i>	F _{ab} (fragment) ⁸¹	Diagnostic
	<i>Fusarium verticillioides</i>	scFv (protein) ⁸²	Plant diagnostic
Vaccinia virus	Ph.D.-C7C (monoclonal) ⁸³	Epitope mapping	
	sdAb (virus) ⁸⁴ Ph.D.-12 (virus) ⁸⁵	Diagnostic	
Avian influenza viruses	scFv (virus) ⁶³ Ph.D.-12 (virus) ⁸⁶	Diagnostic	
	Ph.D.-C7C (virus) ⁸⁷ Ph.D.-12 (monoclonal) ⁸⁸	Epitope mapping	
	<i>Xylella fastidiosa</i>	scFv (whole cell) ⁸⁹	Plant diagnostic
Strongyloidiasis	Ph.D.-C7C (purified IgG) ⁹⁰	Therapeutic candidate	
Lyme disease	X6, X15 X8 (polyclonal) ⁹¹	Diagnostic	
Norovirus	Ph.D.-12 (virus) ⁹²	Diagnostic	
Hepatitis C	scFv (peptides from protein) ⁹³ Ph.D.-7 (protein) ⁹⁴ cDNA library (polyclonal) ⁹⁵ Ph.D.-12 (monoclonal) ⁹⁶	Therapeutic candidate	

	CX9C (polyclonal) ⁹⁷ Ph.D.-12 (monoclonal) ⁹⁸	Diagnostic	
	Ph.D.-12, Ph.D.-7 (protein) ⁹⁹	Inhibitor viral infection	
	X30 (monoclonal) ¹⁰⁰	Epitope mapping	
<i>Bacillus anthracis</i>	landscape phage library (cell) ¹⁰¹ Ph.D.-7 (spore) ¹⁰²	Biosorbent	
	Ph.D.-7, Ph.D.-12 (spore) ¹⁰³ Ph.D.-12 (cell) ¹⁰⁴	Detector	
	Ph.D.-12 (protein) ¹⁰⁵	Therapeutic candidate	
	sdAb (virus) ¹⁰⁶	Diagnostic	
	CX9C (virus) ¹⁰⁷	Diagnostic	
Cucumber mosaic virus	Ph.D.-7 (polyclonal) ¹⁰⁸	Epitope mapping	
Mycobacterium	Ph.D.-7, Ph.D.-12 (monoclonal) ¹⁰⁹	Diagnostics, therapeutics	
<i>Mycobacterium avium</i> subspecies	X20 (polyclonal) ¹¹⁰	Diagnostic	
Neurocysticercosis	X15, X30, XCX15, XCX8CX (polyclonal) ¹¹¹ Ph.D.-7 (polyclonal) ¹¹² Ph.D.-12 (polyclonal) ¹¹³	Diagnostic	
	Equine arteritis virus	X6 (polyclonal) ¹¹⁴	Diagnostic
	<i>Phytophthora capsici</i>	Ph.D.-12 (cell) ¹¹⁵	Plant diagnostic
<i>Staphylococcus aureus</i>	Ph.D.-12 (cell) ¹¹⁶ Ph.D.-12 (monoclonal) ¹¹⁷	Diagnostic	
	Equine infectious anemia	Ph.D.-C7C (gp45) ¹¹⁸	Diagnostic
staphylococcal enterotoxin B	Ph.D.-12 (toxin) ¹¹⁹	Detector	
	Ph.D.-7, Ph.D.-12 (monoclonal) ¹²⁰	Epitope mapping	
<i>E. coli</i> cell surface	Ph.D.-12 (cell) ¹²¹ X8 (protein) ¹²²	Therapeutic Detector	
	Ph.D.-12 (protein) ¹²³ X3CX7CX3 (protein) ¹²⁴ Ph.D.-12 (protein) ¹²⁵ f3-15mer (monoclonal) ¹²⁶	Therapeutic	
Hepatitis B	f3-6mer (polyclonal) ¹²⁷ CX9C (polyclonal) ¹²⁸ X30 (monoclonal) ¹²⁹	Epitope mapping	
	cyclic 7-mer (protein) ¹³⁰ X6 (protein) ¹³¹	Inhibitor	
	Ph.D.-C7C (protein) ¹³²	Diagnostic	

Human cytomegalovirus	Fv (virus) ¹³³ F _{ab} (virus lysate) ¹³⁴	Immunofilter Diagnostic
<i>Streptococcus pneumoniae</i>	15-mer (monoclonal) ¹³⁵	Therapeutic
Black currant reversion associated virus	scFv (virus) ¹³⁶	Diagnostic
<i>Trypanosoma brucei gambiense</i>	Ph.D.-12 (polyclonal) ¹³⁷	Diagnostic
Respiratory syncytial virus	F _{ab} (RSV-F protein) ¹³⁸	Epitope mapping
Epstein-Barr virus	X20 (monoclonal) ¹³⁹	Diagnostic
<i>Listeria monocytogenes</i>	Ph.D.-12 (cell) ¹⁴⁰	Diagnostic
Venezuelan equine encephalitis virus	scFv (cell) ¹⁴¹	Diagnostic
Clostridia	Ph.D.-12 (spore) ¹⁴²	Detection
African horsesickness virus	scFv (polyclonal) ¹⁴³	Detection
Ebolavirus	F _{ab} (polyclonal) ¹⁴⁴ scFv (nucleoprotein) ¹⁴⁵	Diagnostic
Bovine rotavirus	Ph.D.-12 (virus) ¹⁴⁶	Diagnostic
Leptospira	7-mer (monoclonal) ¹⁴⁷	Epitope mapping
Environmental contamination	12-mers (2,4,6-trinitrotoluene-BSA) ¹⁴⁸	Flow sensor

Such a test would offer advantages over RT-PCR, the commonly used diagnostic technique for avian flu viruses. While both tests allow virus subtyping within 4 h and can be automated, ELISA with specific antibodies does not require costly equipment and skilled personnel, thereby allowing application of these tests in developing countries where they are most needed.

While most immunoassays utilize IgG/IgM derived from immunization of mammals or single chain variable fragments (scFv), Lonsdale and coworkers¹⁴⁵ has described for the first time the selection of shark-derived IgNAR V antibodies against ebolavirus. IgNARs from the nurse shark are thermally stable and because of their small

antigen binding domain, they can access hindered epitopes otherwise accessible only to small molecules.¹⁴⁹ The selected IgNAR V variants showed affinity to viral nucleoprotein and were able to recognize the whole virus. Although cross-reactivity with other ebolavirus subtypes exists, these IgNARs were more resistant to thermal denaturation than murine scFv and monoclonal IgG (Figure 1-3B). Although no further development of these antibodies were shown in this report, these selected antibodies, when incorporated into sandwich ELISA, may be able to capture viable virus and possible to be integrated in clinical or point-of-care tests.

Single-domain antibodies (sdAb) also possess high thermal stability and are able to refold to active molecules after temperature extremes, which make them more suitable for point-of-care assays when cold-storage is impeded.¹⁵⁰ Reported by Goldman and coworkers,¹⁰⁶ through selection against *B. anthracis* spores, the selected sdAbs showed nanomolar binding affinity to S layer protein EA1, and were able to refold into functional antigen-binding molecules following several rounds of thermal denaturation. When developed into sandwich assay, these immobilized sdAbs had a detection limit of 10⁴ CFU/mL. Along with validation of testing results on patients that are susceptible to infections, these molecules, with unique thermal stability and specific antigen-binding properties may lead to adaptation in point-of-care systems.

One major challenge in phage-display technology is to generate species-specific binding interactions. Khati and coworkers¹⁵¹ used commercially available macrocyclic CX7C peptide libraries to select peptide sequences with affinity to *Mycobacterium tuberculosis* H37Rv. However, the most enriched peptide sequence identified through high-throughput sequencing cross-reacted with other mycobacteria species (*M. bovis*

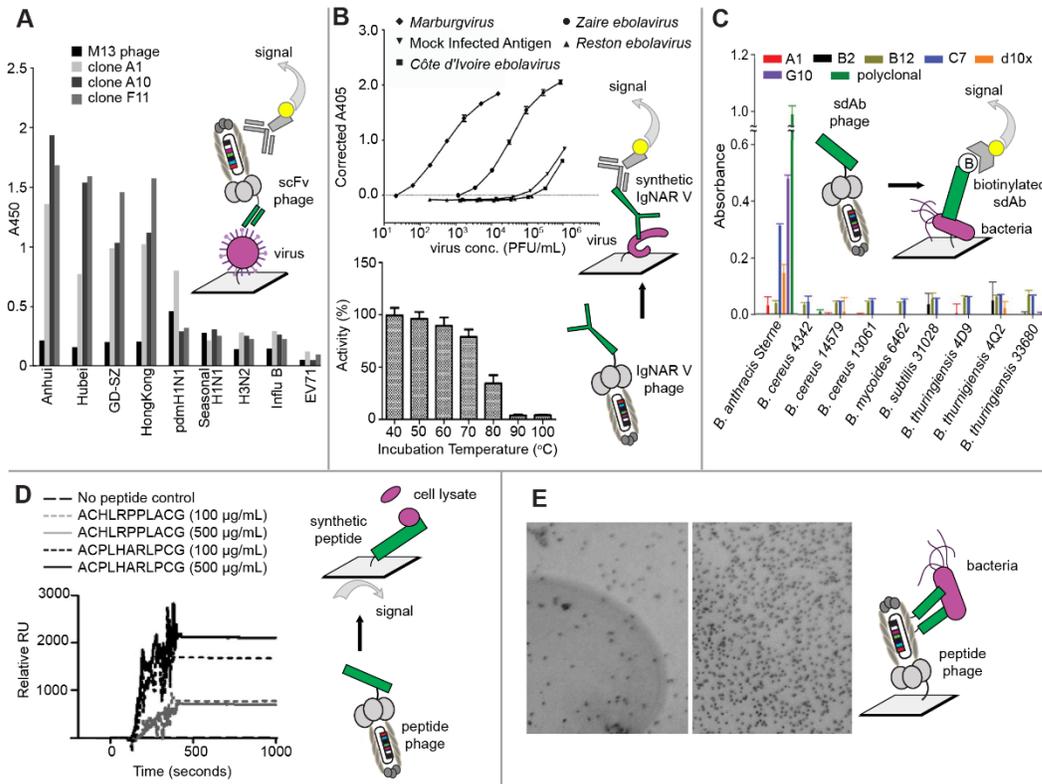


Figure 1-3. Phage libraries targeted for whole pathogen.

(A) ELISA signals showing binding of selected phage clones or the M13 phage library as a whole to H5N1 and other viruses. (B) Binding specificity of a selected IgNAR V clone to Ebolavirus species and the thermal stability of this clone after incubation at different temperatures. (C) Specificity of selected sdAbs to *Bacillus* species. All clones except B12 show preferential binding affinity to *B. anthracis Sterne* by direct binding ELISA. (D) The association of *M. tuberculosis* H37Rv lysate with selected peptides in SPR. High binding was observed for the selected peptide sequence (ACHLRPPLACG), but not the scrambled peptide (ACPLHARLPCG). (E) High power optical microscopy showing binding of *S. Typhimurium* ATCC 13311 to selected phage (right), but not to unrelated phage (left).

BCG, *M. smegmatis*). Surface plasmon resonance (Figure 1-3D) demonstrated this synthetic peptide binding to a protein from *M. tuberculosis* lysate, yet the identity of this protein was not validated nor tested for its presence in other mycobacteria species, which may explain cross-reactivity of this sequence to other mycobacteria. Nonetheless, ELISA tests using these peptides, in contrast with conventional tests that use antibodies, may be

able to detect non-antigenic bacterial cell wall components from patient samples. Using these selected peptides as diagnostics offers great advantage since it is possible that the surface of *M. tuberculosis* also possesses molecules which may not be antigenic.

Interestingly, phage can be used to produce functional materials such as biosorbents to monitor the presence of and to remove pathogenic bacteria (e.g., *Salmonella*).¹⁵² Development of such materials can build on the advantages of landscape phage technology, which allows the production and screening of densely functionalized phage particles with different surface structures and biophysical properties;³⁴ in addition, a high copy number of displayed epitopes can improve binding avidity. Using landscape phage with multiple copies of pVIII modified with peptides against *S. Typhimurium* cells, Otvos and coworkers¹⁵³ discovered five families of bacteria-binding peptides. When immobilized on gold surface, the selected peptide phage clones can be observed under microscopy (Figure 1-3E). This approach may improve current microscopy methods; using labeled probes with high specificity to bacteria may alleviate subjective interpretation. In further studies, selected phage particles were incorporated into biosensors and give sensitivity of 5000 cells/mL from fat-free milk spiked with *S. Typhimurium*.¹⁵⁴ These phage can be beneficial for detection, separation and purification of bacteria prior to their identification with PCR or immunoassays.

1.4.2 Selections that target fragments of pathogens

Proteins, lipids, small molecules, DNA, RNA, and carbohydrates derived from or secreted by the pathogens can be isolated and used as targets for selection. Detection of protein-based toxin is one of the immediate needs in monitoring the quality of food and

water, but canonical methods of detection that use toxicity assays require significant time for detection. To address this issue, Porter and coworkers⁶⁵ selected scFv that can rapidly detect the hepatotoxin microcystin-LR at levels below the World Health Organization limit in drinking water. Although the selected scFv molecules cross-reacted with three other microcystin isomers (microcystin-RR, microcystin-LW, microcystin-LF), the selected scFv could concentrate trace level of microcystin-LR from a large volume of water (Figure 1-4A). With stability in nonphysiological conditions, for example, methanol,¹⁵⁵ which is used routinely to extract microcystins from environment samples, these detecting agents can be used in immunoaffinity applications to clean up and concentrate toxins.

Antibody-based detection is usually limited to protein-based antigens whereas non-polar lipids remain a largely neglected group of antigens, as lipids typically do not exhibit high affinity to antibodies produced by traditional hybridoma methods. However, through *in vitro* selection against mycolic acid extracted from *M. tuberculosis*, Hanson and coworkers⁸¹ identified anti-mycolic F_{ab} antibodies. When converted to IgG, the most sensitive hit can detect *M. tuberculosis*-derived mycolic acids at 4.5 ng in an ELISA assay. The antibodies did not cross react with closely related lipids or other *M. tuberculosis*-derived lipids. Combined with rapid and optimized liquid extraction developed in this study, these antibodies can detect 10⁶ CFU/ml of *M. bovis* BCG, a close relative of *M. tuberculosis* in an ELISA assay that uses lipids extracted from bacteria (Figure 1-4B). Although no data were shown for the detection limit for *M. tuberculosis*, this study demonstrated that lipid biomarkers can be used in diagnostics. This method can be potentially used to target a diverse range of insoluble lipids that previously may have

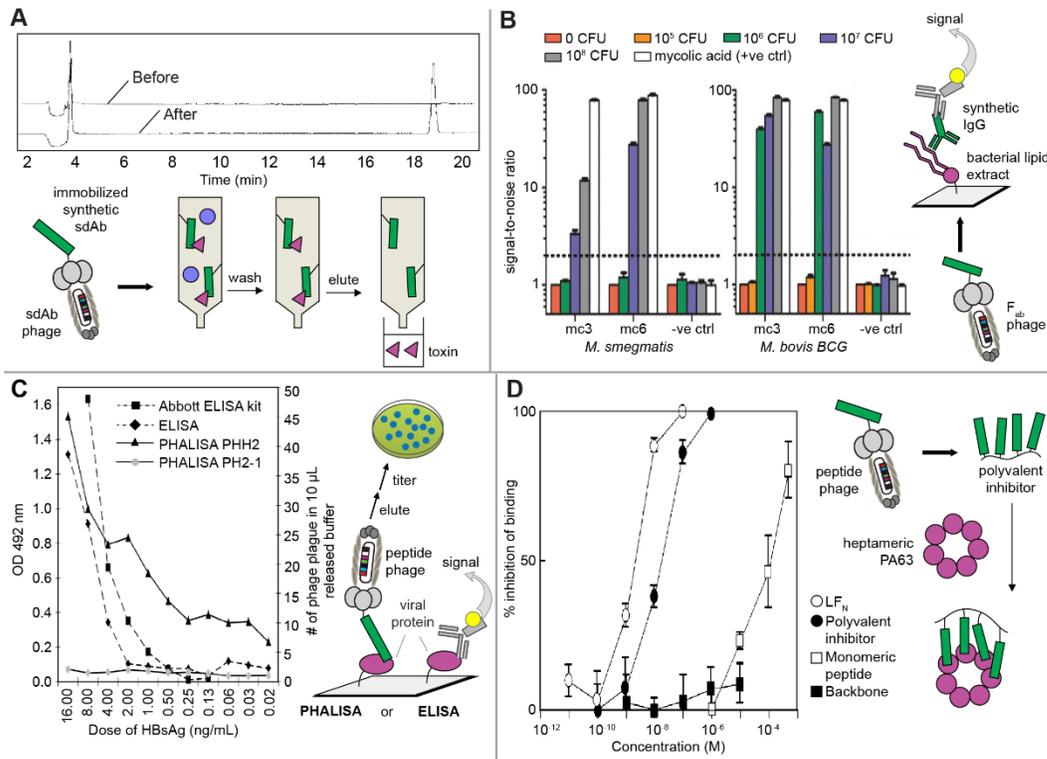


Figure 1-4. Phage libraries targeted for fragments of pathogens or toxins.

(A) HPLC chromatograms of distilled water spiked with microcystin-LR before and after concentration on immunoaffinity column prepared with selected scFv. (B) Detection limit of the two most sensitive IgG mc4 and mc6 (converted from F_{ab}) towards CFU of *M. smegmatis* and *M. bovis* BCG. (C) Comparison between sensitivity of ELISA and PHALISA in detecting HBsAg polypeptide (viral protein). Binding activity is measured by OD (ELISA) or standard phage plaque analysis (PHALISA). PHH2: monoclonal phage obtained after five rounds of panning; PH2-1: monoclonal phage selected after two rounds of panning. (D) Inhibition of the association of anthrax toxin in cell culture. Toxin association was measured as the binding of ³⁵S-labeled LF_N on cells incubated with PA protein.

not been accessible to antibody-based detection. Still further testing on clinical sputum samples will be required to validate the clinical potential of the selected antibodies.

In addition to providing standalone affinity reagents, phage display can give rise to diagnostic tests in which the phage itself serves as the detecting agent and a reporter. For example, Jungkind and coworkers¹⁵⁶ developed a detection platform, PHALISA,

using selected clones carrying random 8-mers. This platform was later used for the detection of HBV surface antigen HbsAg.¹²² Different from ELISA, signal from PHALISA was generated from standard plaque assay, using selected phage clones to bind to target. PHALISA could detect 20-100 pg/mL of HBsAg, while ELISA required 1-10 ng/mL of antigen, showing that PHALISA was 100 times more sensitive in detection of HBsAg per sample than conventional ELISA (Figure 1-4C). Despite high sensitivity, there has been little further work on PHALISA, which makes the development of PHALISA into real-world detection assay preliminary.

In some cases discovered binders also inspire designs of therapeutic agents. For instance, Collier, Whitesides and co-workers used peptides selected from Ph.D.-12 peptide library against a heptameric anthrax toxin to construct a polyvalent potent inhibitor of the toxin.¹⁰⁵ This inhibitor could be useful in therapeutic application against clinical anthrax (Figure 1-4D). Such multivalent inhibitors designed by Whitesides,¹⁵⁷ Bundle,¹⁵⁸ and many others have been particularly useful in blocking oligomeric high-symmetry bacterial toxins.¹⁵⁹⁻¹⁶⁰ However, to date, many multivalent inhibitors failed to enter clinical trials or have been withdrawn from advanced stage clinical trials.¹⁵⁹ It is also not clear whether the success of targeting anthrax toxin with short peptides can be applied to other toxins. Natural ligands for many toxins are cell surface glycans. Targeting such glycan-binding proteins with peptide-based libraries is notoriously challenging and is known to result in ligands that do not bind to the native carbohydrate binding site.¹⁶¹ For example, all attempts to target clinically relevant Shiga Toxin 2 with peptide-based phage-displayed libraries have been unsuccessful (David Bundle, personal communication). On the other hand, antibody-based inhibitors for Stx2 are known and

they have been used in phase 2 clinical trials (e.g., Shigamab™ from BELLUS). We and others have also demonstrated that chemical modification of short peptide libraries with monosaccharide can result in the identification of ligands that target carbohydrate-binding sites in difficult “lectins.”^{49, 52}

Other examples of phage-based discovery of peptides, antibodies and proteins that bind to purified components from pathogenic microorganisms are listed in Table 1.

1.4.3. Selections that target antibodies associated with pathogenic infections or diseases

Antibody profiling can be very useful for pathogen detection. The idea of finding a disease epitope in unpurified sera has persisted for over two decades,^{162 73, 163-164} yet the progress is surprisingly slow. We hypothesize that this is likely due to four factors: i) heterogeneity of serum antibodies;¹⁶⁵ ii) disproportionate display of ligands in genetically-encoded libraries;¹⁶⁶ iii) amplification bias;¹⁶⁷ and iv) incomplete chemical space of genetically-encoded peptide libraries. The latter problem was highlighted by Kodadek and co-workers, who proposed that many peptide-based selections for disease-specific autoantibodies cannot represent a general route to rapid discovery of diagnostic probes because the collections of unmodified peptides are unlikely to contain the primary autoantigens that trigger the most disease-specific autoimmune response.¹⁶⁸ On the other hand, if the monoclonal antibody associated with the disease is known, selection of a purified monoclonal antibody often can give rise to a successful peptide epitope, even in the cases when antigen is not related to peptides. There has been several reports on successful selection of immunogenic peptide mimotopes of carbohydrate antigens.¹⁶⁹⁻¹⁷¹

However, because these studies usually employ non-human antibodies, realization into clinical diagnoses has to be assessed on a case-by-case basis.

One limitation in detecting infectious disease is the difficulty to prepare antigenic extracts from patients' samples. To circumvent this problem, Costa-Cruz and coworkers⁹⁰ proposed a strategy to discover peptides from Ph.D.-C7C libraries that are specific for *stercoralis*-associated antibodies in patients' serum. The most successful phage displayed peptides showed sensitivity >85% and specificity >77.5% using purified IgG (Figure 1-5A). These novel biomarkers can be incorporated into gel agglutination tests, electrochemical sensors, or magnetic hybrid capture ELISA. At this stage, a simple phage-ELISA test may be used as a screening platform to detect human strongyloidiasis.

When the target presents discontinuous epitopes – epitopes with binding residues that are distant in primary sequence but close in proximity in tertiary structure, synthetic peptides antigens with multiple epitope sites may be suitable as detection reagents. For example, Lejon and coworkers¹³⁷ selected peptides from Ph.D.-12 and Ph.D.-C7C libraries against monoclonal antibodies to identify sequences that mimic discontinuous epitopes on the variant surface glycoproteins (VSGs) of *Trypanosoma brucei*. The sequence of some selected peptides matched different parts of the VSG protein, indicating the epitopes of the target VSGs to be discontinuous, whereas others resembled the sequence of the linear region of VSG. While these peptides exhibit diagnostic potential in an inhibition ELISA using positive serum (Figure 1-5B), these discontinuous mimotopes only mimicked the antigen partially and thus were not as efficiently as easily to be recognized as linear mimotopes by human antibodies. To be integrated into clinical applications, the development of these peptides to serological assays are necessary, and

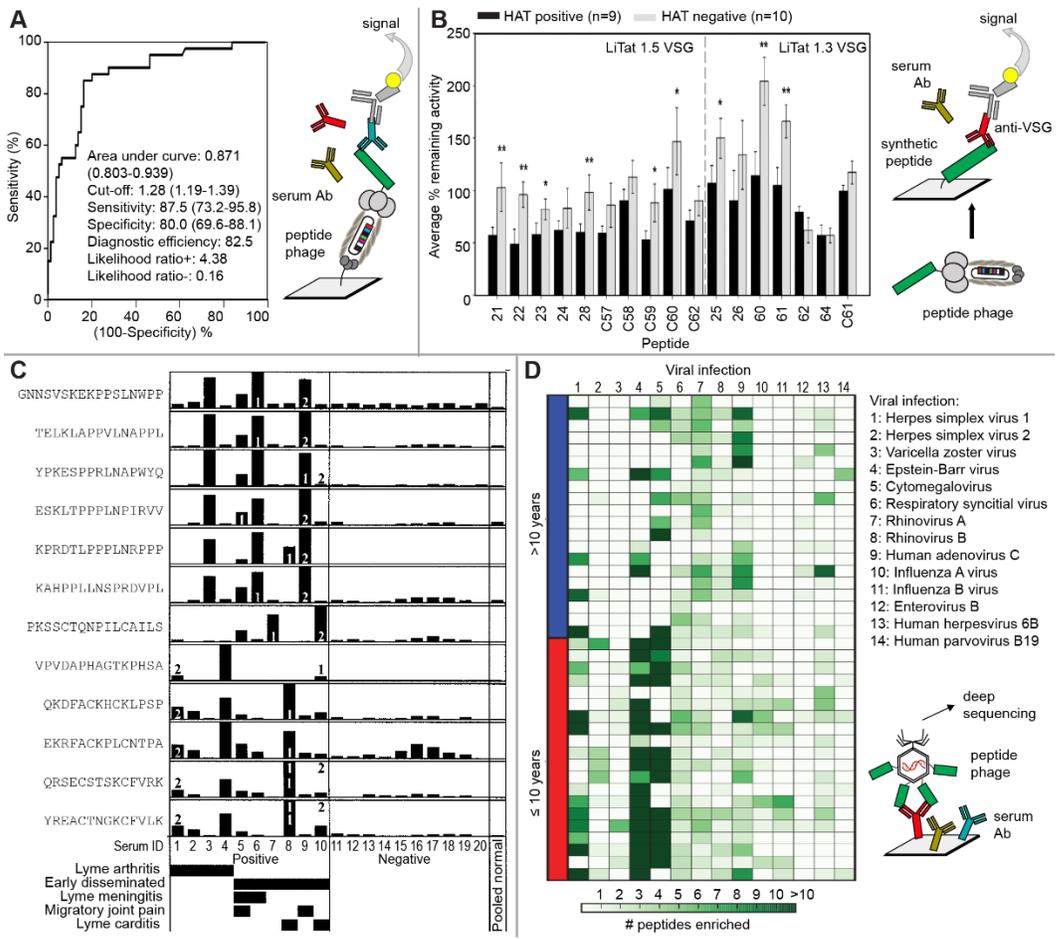


Figure 1-5. Phage libraries targeted for antibodies associated with disease. (A) ROC curve of selected phage clone as detecting agents for human strongyloidiasis using 120 serum samples. (B) Average percent remaining activity of anti-VSG monoclonal antibodies (LiTat 1.5, LiTat 1.3) after binding to mimotopic peptides discovered by phage display using inhibition ELISA. * $p < 0.05$; ** $p < 0.01$. (C) Reactivity of 12 phage-displayed peptides with Lyme disease positive and negative sera as determined by antibody capture ELISA. The ELISA assay is illustrated in (A). (D) A typical Antibody profile from VirScan assay of random donors. Each column is a virus, each row is a sample. The color of intensity of each cell indicates the number of peptides from the virus that were significantly enriched by antibodies in the sample.

serum samples should account for inherent differences in antibody response among individual patients.

In another case, twelve random peptide libraries displayed on fd phage were selected against serum antibodies from patients with Lyme disease.⁹¹ Several diagnostic probes were identified (Figure 1-5C), and the sequences of which were characterized as mimotopes of discontinuous epitopes, as none matched contiguous segments of proteins from the Lyme disease pathogen, *Borrelia burgdorferi*. These several diagnostic probes, or combination of a selected set of these probes, can potentially form the basis of a new diagnostic enzyme-linked immunosorbent assay, and may replace expensive immunoblotting tests that are required for definitive serological diagnosis. However, to be sufficient as diagnostics, the sensitivity and specificity of these probes must be tested on clinical samples.

While randomized peptide libraries represent the major core for targeting antibodies associated with disease, focused peptides libraries derived from the DNA of the pathogens may also improve serological detection of viral infections. Using gene-array DNA synthesis to produce a DNA library that covers 206 viral genomes, Elledge and coworkers¹⁶⁵ generated a VirScan platform using T7 phage library to display peptide fragments from the viral genome. This platform was able to detect past and current viral infections (Figure 1-5D). By combining phage immunoprecipitation with deep sequencing (PhIP-seq), VirScan condensed affinity selection to a single round, avoiding amplification bias. This technology can be easily expanded to include new viruses and human pathogens in order to discover viral epitopes. In another similar study, Sidhu, Kim, and coworkers generated phage libraries that contain human and viral C-terminal peptides through oligonucleotide microarrays.¹⁷² This proteomic peptide-phage display

(ProP-PD) strategy successfully identified short linear motifs for modular protein domains.

1.5 Prospect for phage display-derived detection reagent

Twenty-five years since the advent of phage-display selection, one can recognize the pros and cons of canonical phage display selection procedures that use traditional randomized polypeptide libraries and methods for selection of ligands that employ multiple rounds of panning. Currently, several emerging enabling technologies can further facilitate the discovery of new ligands and diagnostic probes. Firstly, DNA-array based synthesis¹⁷²⁻¹⁷⁵ opens previously impossible opportunities for the production of sparse genetic libraries that contain only desired DNA fragments, such as those derived from a viral genome.

The second enabling technology is the incorporation of post-translational modifications (PTM)^{43, 46, 49, 176-181} to generate libraries with unnatural amino acids¹⁸² or with ribosomally-made post translationally modified peptides (RiPPs),¹⁸³ which are introduced in section 1.3.1. Because the chemical entities cannot be encoded genetically, we believe that such PTM-libraries will be useful when targets recognize non-protein ligands. In addition, these libraries offer advantage to mimic antigens that are critical in immune responses, for instance, citrullination in rheumatoid arthritis and protein cleavage in systemic lupus erythematosus.¹⁸⁴ In my thesis, I demonstrate that such PTM library can be used to discover ligands that mimic antigens for tuberculosis.

Next generation sequencing serves the third facilitator for efficient screening. Historically, multiple rounds of selection have been used to maximize the enrichment of target-binding clones. However, it is apparent now that even a minor proliferation advantage of some phage clones during amplification between rounds can significantly skew the outcome.^{167, 185} This problem is exacerbated when the target is complex, such as with a repertoire of all available extracellular receptors on the cell surface or a repertoire of all antibodies present in serum of an individual patient.¹⁸⁶ The use of high-throughput sequencing minimizes both the number of amplification steps and the amplification bias.^{186-188, 165, 175, 189} These so-called deep-panning strategies that use only one round of selection make identification ligands against such complex targets more feasible than analogous multi-round selections, with identical library and identical target.

While this chapter highlights the use of phage display, many other display platforms exist. Bacterial and yeast display offer at least one very important advantage when compared to phage display: they can be enriched using sequential magnetic-activated cell sorting (MACS) and fluorescence-activated cell sorting (FACS),¹⁹⁰ which allows real-time analysis of target-binding affinity,¹⁹¹⁻¹⁹² specificity, and stability to proteases of the peptides displayed on cell surface.¹⁹³ Using peptide display on a bacterial surface, next generation sequencing (Display-Seq), and computational analysis (IMUNE), Daugherty and coworkers identified celiac disease- specific motifs, QPEQPF[PS]E, a deamidated version of QPQQPF[PS]Q, present in many grain antigens associated with Celiac disease.¹⁹⁴ On the other hand, using yeast display, it is possible to incorporate complex post-translational modifications in displayed proteins. This display system

permits identification of ligands and probes that depend on the presence of the post-translational modifications.¹⁹⁵⁻¹⁹⁶

Bacteriophages have great potential for the development of affinity reagents for capture and detection of pathogens or molecular biomarkers from these pathogens required for molecular diagnostics. Phages are the most convenient carriers for display of synthetic ligand libraries and ligands discovered from such libraries can improve molecular diagnostics. Multivalent display of ligands on phage is reminiscent of the display of multiple recognition elements on multivalent IgG or IgM antibodies. It is thus not surprising to find applications in which a phage is used not only as a discovery tool but also as a detection tool in phage-ELISA, phage-immunoprecipitation, phage-based labeling of cells, and other assays that traditionally employ antibodies as detection reagents. Advances in phage display offer great promise for molecular diagnostics but there are only a few reports that integrate molecular discoveries into functional diagnostic devices. Many fruitful outcomes from phage screening that lead to novel biomarkers and new epitopes (Table 1), are yet to be translated to any end-point assays. Many molecular discoveries have been validated only on a purified target and never tested on “real life samples” (e.g. human serum, native pathogen-containing sample, etc). None have been commercialized yet. Both optimization of phage-based discovery of molecular components and simplification of the production of diagnostic platforms¹⁹⁷⁻²⁰¹ will be necessary to bridge the gap between ligand development and diagnostic kits.

1.6 Thesis overview

In Chapter 1, I described the current trend in the development of detection assays derived from phage display technology. I discussed examples that use different types of targets, including: whole pathogen, fragments of pathogens, or antibodies associated with the pathogens, for the discovery of potential diagnostic probes. Finally, I addressed challenges in this field and enabling technology that can circumvent them.

In Chapter 2, I present my work that employs a genetically encoded fragment-based discovery (GE-FBD) strategy to elaborate a known glycan for anti-lipoarabinomannan (aLAM) antibody into a glycopeptide antigen. These glycopeptide antigens exhibit more specific monovalent and multivalent interactions with subset of anti-LAM antibodies than their parent carbohydrate antigen. These molecules can be the starting point for the design of improved diagnostic devices for serological detection of mycobacterial infections.

In Chapter 3, I investigate the structure-activity relationship of the discovered glycopeptide antigens to anti-LAM antibodies in monovalent and multivalent fashions. I started from truncating glycopeptide antigens that contain either a complex carbohydrate (hexasaccharide) or a simple monosaccharide. I then measured the monovalent affinity and multivalent avidity of the interactions of these truncated antigens with F_{ab} fragments and IgG of anti-LAM antibodies, respectively, and compared them to the performance of the parent glycopeptides (hexasaccharide conjugates). I observed that truncating the hexasaccharide into monosaccharide on the glycopeptides ablated their monovalent affinities to anti-LAM antibodies. Translating these truncated ligands into multivalent assays did not result in binding to anti-LAM antibodies either. These observations

suggest that the monosaccharide with the peptide sequences were not sufficient to elicit specific, constructive interaction, confirming that intact peptide and hexasaccharide are critical for binding.

Chapter 4 describes the optimization of chemical post-translational modifications that convert natural disulfide-bridged peptides libraries displayed on M13 phage to glycosylated macrocyclic peptides libraries. I tailored these libraries to contain two monosaccharides suitable for interactions with anti-LAM antibody—specifically each macrocycle displays both the α - or β -anomers of arabinose. I developed a deep-panning strategy that bypasses the need for multiple rounds of selection and uses only a single round of panning and PCR as the sole amplification step. We believe that the discovery approach that eliminates bias emanating from amplification of phage in bacteria and employs a library with chemical diversity pertinent to mycobacterium derived antigens (Ara) will maximize the success of discovery of TB-antigens directly from serum.

Lastly, Chapter 5 summarizes the advances and limitations I encountered in developing the technology described in Chapters 2-4. I outlined future directions to translate the results of my thesis work into a functional serological test.

Chapter 2: Genetically-encoded fragment-based (GE-FBD) discovery of glycopeptide ligands for antibodies associated with mycobacterial infections

2.1 Introduction

The composition of antibodies in human serum represents a unique signature reflecting the history of infections and vaccinations. From a clinical perspective, antibodies offer greater advantage than other biomarkers for disease diagnosis, as other biomarkers originated or shed from pathogens are often hard to detect due to their presence at low concentration.²⁰² To date, serological tests have been shown successful for many infectious disease diagnoses,²⁰³ such as syphilis²⁰⁴ and Lyme's disease;²⁰⁵ however, incorrect or missed serodiagnoses of TB, especially in resource-limited settings, remains challenging.

The causative agent of TB—*Mycobacterium tuberculosis*—and other Mycobacteria species possess unique cell wall structure that can stimulate host immune response.²⁰⁶ Among the diverse carbohydrates and lipids in the mycobacterial cell wall (MCW), lipoarabinomannan (LAM), or its sub-structures, were of great interest as antigens for diagnostic platforms, particularly because LAM can be recognized by antibodies in infected individuals.²⁰⁷⁻²⁰⁸ Furthermore, many MCW components are species-specific, offering the possibility of their use as biomarkers for various mycobacterial diseases, including tuberculosis, leprosy,²⁰⁹ and *M. avium* infections that frequently plague the immunocompromised.²¹⁰

Neither synthetic nor native antigens can currently act as effective discriminatory agents for TB from real world samples. For example, all Mycobacteria infections give rise to antibodies that can react with many sub-structures of LAM. As a result, current serological tests that use native antigens (e.g., LAM) or substructures of LAM (e.g., Ara₆) demonstrate poor specificity.²¹¹⁻²¹² Even with recombinant antigens, sensitivity and specificity are highly inconsistent.²¹³⁻²¹⁴ Despite these problems, the interest in serodiagnostics remains high due to the low cost, high speed, and simplicity of these tests when compared to PCR- or culture-based detection.²¹⁵ Developing well-defined synthetic antigens with higher discriminatory power will facilitate next-generation serodiagnostics, which, in turn, can be readily accepted by resource-limited communities.

Previously, phage display has been used to discover peptides that bind antisera against neutral polysaccharide of *M. tuberculosis* (PLXGT[L/V]P)²¹⁶ or *M. bovis* BCG (MSPRATI, EQPYLHV, EQPYIEN),¹⁰⁹ as well as monoclonal anti-LAM antibody CS-40 (WEADDKNQHGEG, ISLTEWSMWYRH, EEGPWSTHVGRT, WGNEGGDHLQPV, SLKIRWELKMYQE, AVERWEKHTWSE)²¹⁷ and CS-35 (MSPRATI, SHRLLQTYWSSA).¹⁰⁹ When validated against human TB serum²¹⁷ as an antigen, the discovered peptide ISLTEWSMWYRH exhibited similar specificity (92%) to that of LAM but with compromised sensitivity (50%). In another case, selected 7-mer and 12-mer peptide mimotopes, when tested against positive human serum, demonstrated half of the sensitivity that was exhibited by LAM as an antigen.¹⁰⁹ Although high-throughput screening with random peptide libraries is rapid and can uncover ligands that demonstrate binding properties to targets, without structural evidence, it is unclear whether any of these discovered mimotopes target the antigen binding site of

mycobacteria-associated antibodies. It is uncommon for peptide ligands discovered by technologies like phage display to bind carbohydrate-binding proteins²¹⁸ at the site distal from carbohydrate recognition domain (CRD).²¹⁹ It is also known that IgG specific to a carbohydrate moiety (e.g., methyl- α -D-mannopyranoside, ligand for ConA) does not exhibit cross-reactivity with the carbohydrate-mimicking peptide (e.g., DVFYPYPYASGS, peptide mimetic of methyl- α -D-mannopyranoside),²²⁰ which indicates that selections solely using random peptide libraries and polyclonal antibodies as selection target may elicit ligands that are biased towards only a subset of antibodies. In turn, devices that incorporate such ligands can exhibit compromised sensitivity or specificity.

Systematic bottom-up syntheses of carbohydrate substructures from antigenic glycans is an orthogonal approach that has the potential to discover carbohydrate fragments essential for antigen discovery and synthetic antigens for diagnostic tests.²²¹⁻²²³ Lowary and coworkers undertook such systematic approach to develop a panel of synthetic arabinofuranosyl-containing oligosaccharides and identified that a hexasaccharide (Ara₆), located at the non-reducing termini of the arabinan domains in mycobacterial LAM, exhibit the strongest binding among other fragmented derivatives of LAM to mAb CS-35, an anti-LAM antibody generated against *M. leprae* (Figure 2-1 A).²²⁴ The crystal structure of CS-35 F_{ab} and Ara₆ further identified the key interactions that contribute to strong binding—the two terminal residues in the hexasaccharide to the antigen-binding site of the antibody F_{ab} (Figure 2-1B).²²⁵ However, in a multiplexed serological test developed by Lowary, Kolk, Raap, and coworkers, Ara₆ as a diagnostic antigen showed 100% sensitivity and 39% specificity (positive/negative cutoff point at

3000 of the antigen intensity score).²¹² A possible method to enhance ligand specificity to CS-35 is through further elaboration of native carbohydrate structures in order to incorporate other fragments of LAM; yet this route involves multistep and complex chemical manipulations. Scaling up of such syntheses for validation purposes could be challenging.

To retain ligand binding affinity and simultaneously enhance ligand binding specificity to CS-35 without complex syntheses, I proposed an alternative ligand search strategy that involves elaborating the existing Ara₆ as an anchor for carbohydrate-antibody recognition with random peptides to streamline ligand specificity towards CS-35. To this end, I employed GE-FBD to identify glycopeptide ligands from phage-displayed glycopeptide libraries that contain a constant Ara₆ residue and randomized peptides, which bind to monoclonal anti-LAM antibody CS-35 with higher affinity than Ara₆ alone (Figure 2-1 C-E). Similar libraries have been previously used by Derda group to identify glycopeptides which can bind in the CRD of model lectin ConA.⁵² Although CS-35 is not directly associated with TB, we believe that the selection of glycopeptides, with enhanced affinity and specificity towards specific anti-LAM antibody but not other closely related antibodies, can serve as the first step towards developing technology that yields antigens with increased specificity. Subsequently, these antigens can give rise to TB serodiagnostics with enhanced specificity.

Elucidating the discriminatory potential of the identified Ara₆-peptides for real world serum will have impact beyond immunoprofiling for TB-specific antibodies. For

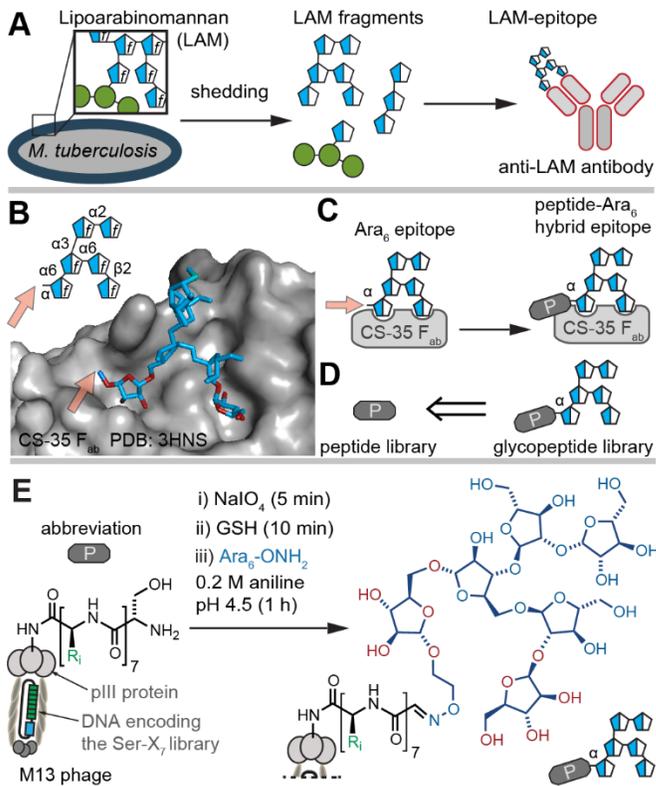


Figure 2-1. The design of glycopeptide antigen Ara₆-peptides for anti-LAM antibody CS-35.

(A) Mycobacteria cell wall contains lipoarabinomannan (LAM) that is constantly shed during infection. Components of LAM can trigger host immune response. (B) Crystal structure of anti-LAM antibody CS-35 F_{ab} to Ara₆, a hexassacharide that is derived from LAM. (C) Starting from previously validated Ara₆ epitope with modest specificity, we propose that peptide-Ara₆ hybrid epitope will increase binding specificity of such epitope to CS-35. (D) These peptide-Ara₆ glycopeptides are generated from peptide libraries. (E) Glycopeptides are discovered using GE-FBD. Starting from Ser-X7 library displayed on M13 phage, oxidation of peptide libraries with NaIO₄ and subsequent oxime ligation with Ara₆-ONH₂ yields glycophage libraries containing constant Ara₆ moiety and randomized heptapeptide moieties.

example, a developed panel of Ara₆-peptides may be used as improved biomarkers for bovine TB, a disease that remains a concern to the cattle industry in Alberta and worldwide.²²⁶ Furthermore, the glycophage display approach can also be used to select discriminatory glycopeptide epitopes in other diseases using privileged but specificity-compromised carbohydrate antigens.

2.2 Results and discussion

2.2.1 Generation of glycopeptide libraries on phage

LC-MS analysis confirmed the formation of glycopeptide conjugate from the peptide sequence SVEKY and Ara₆ hydroxylamine through oxime ligation (Figure 2-2 A). Specifically, treatment of SVEKY with 60 μ M solution of NaIO₄ for 5 min on ice, followed by quenching with 1 mM GSH, led to complete conversion of peptide to glyoxal peptide, as confirmed by LCMS. One-pot addition of 2 mM Ara₆-ONH₂ in 200 mM aniline converted aldehyde to oxime. A minimum concentration of 2 mM of Ara₆-ONH₂ was required for this reaction.

Employing the identical reaction conditions—oxidation and oxime ligation—on a phage that displays sequence SHTHDSVE yielded 69 \pm 12% conversion of peptide to glyoxal and then to oxime, as confirmed by a biotin-streptavidin capture assay previously developed in our group (Figure 2-2 B).^{180, 227} The same optimized conditions converted the N-terminal serine in a random library SX₇ to either Ara₆ or galactose terminated libraries in 24% and 49% yield (Figure 2-2 C). The Gal-X₇ library served as control in the downstream selection process.

2.2.2 Selection and sequencing

In each round of panning, libraries modified with Ara₆ or galactose, and unmodified libraries were incubated in the wells of the 96-well plate coated with CS-35 IgG. Ara₆-X₇ libraries incubated in blank wells served as additional control (Figure 2-3 A). The subset

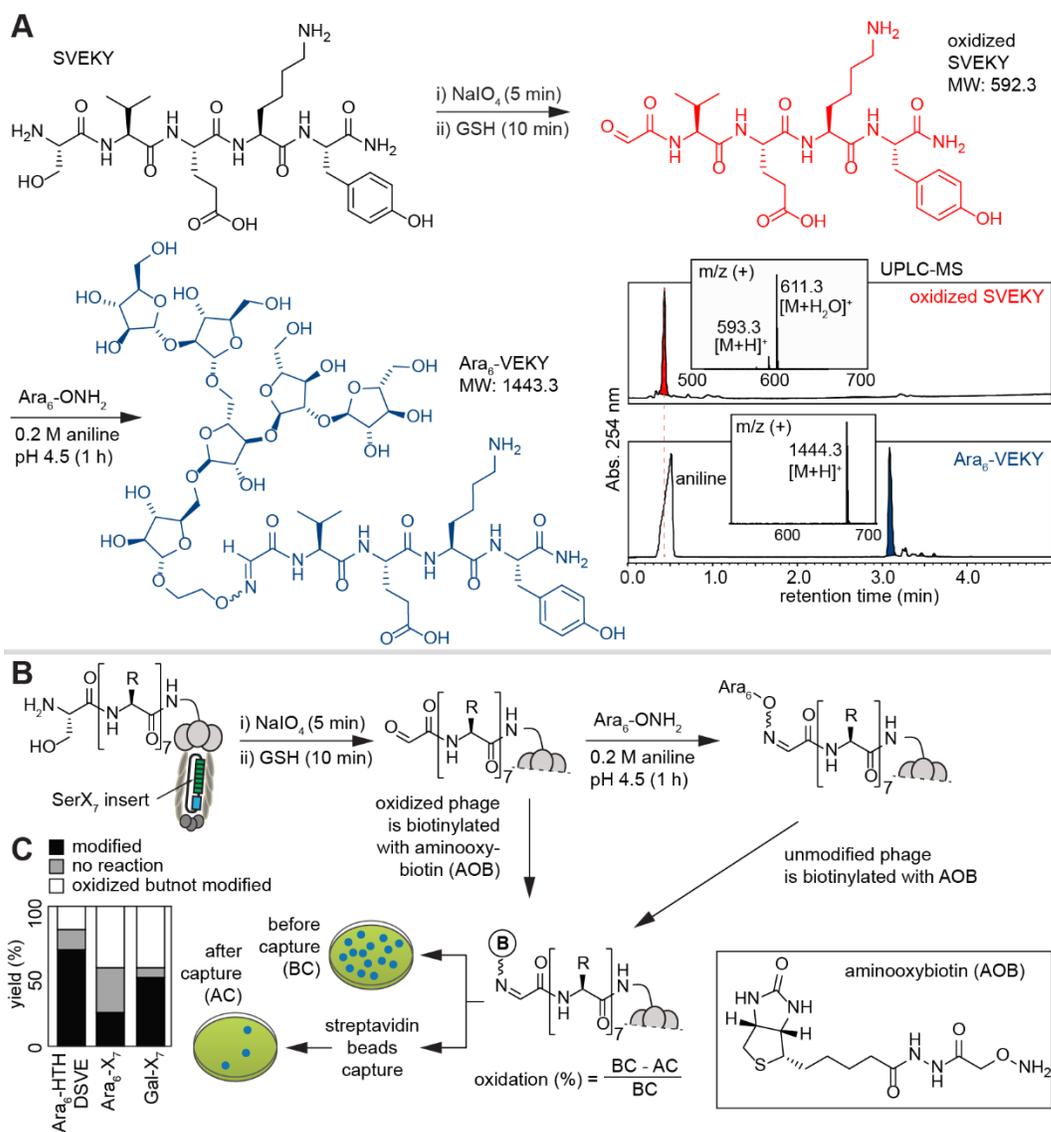


Figure 2-2. Validation of chemical modifications on peptide, phage clones and phage libraries.

(A) LC-MS confirms oxidation and oxime ligation reactions on the peptide SVEKY. (B) Oxidation and oxime ligation reaction conditions and quantification method on phage libraries. (C) The conversion of the oxidation and oxime ligation reactions on monoclonal phage and phage libraries as determined by direct biotin capture (oxidation) and differential biotin capture (ligation).

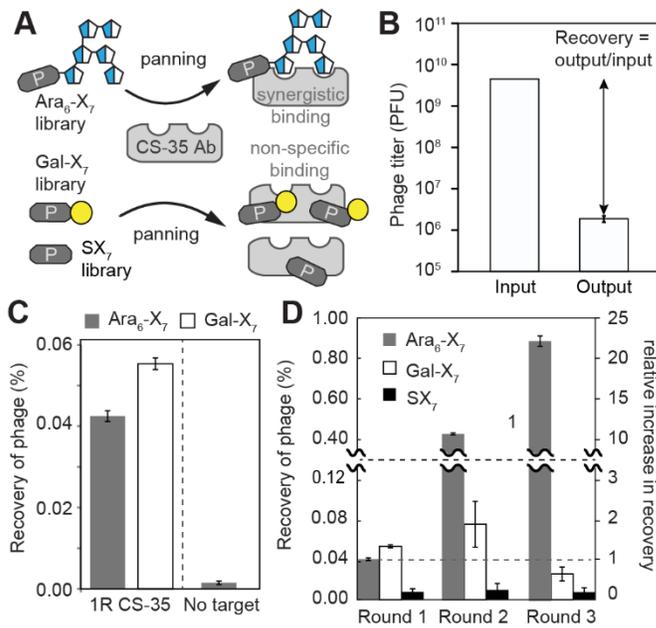


Figure 2-3. Overview of selection and recovery of phage after multiple rounds of selection.

(A) Ara₆-X₇ or control SX₇ and Gal-X₇ libraries were selected against anti-LAM antibody CS-35. (B) Recovery of phage is the mean of eluted phage titers divided by the mean of input phage titers for biopanning ran in triplicate. Error bars represent the propagation of error. (C) Recovery of phage after the first round of biopanning. (D) Recovery of phage after three rounds of biopanning on CS-35. At the third round of selection, library modified with Ara₆ was significantly enriched, whereas the unmodified library or the same library modified with Gal was not.

of the library enriched in wells coated with CS-35 IgG were amplified and further modified (or not modified) and served as inputs for the second round of panning.

The recovery of Ara₆-modified phage throughout three rounds of selection against CS-35 increased by a factor of 22 fold (Figure 2-3 D). The recovery was determined as the ratio of number of phage particles eluted from the target-coated plates to the number of phage particles added to this target prior to each round of panning (Figure 2-3C). Deep sequencing and differential enrichment analysis²²⁸ of the sequences emanating from the first two rounds of selection of a library modified with Ara₆ identified 80 putative glycopeptide binders for CS-35. All 80 hits have been significantly enriched (p<0.05,

R>4) in selection of Ara₆-X₇ library on CS-35 but not in any of the control screens (Figure 2-4).

2.2.3 Glycopeptides testing using paper-based binding assay

I synthesized the peptide portions of the putative glycopeptide hits on patterned Teflon-impregnated paper using previously developed SyntArray technology (Figure 2-5 A).²²⁹ LC-MS of the peptides cleaved from the cellulose through treatment with NH₃ confirmed the peptide sequences on paper (Appendix). The oxidation of N-terminal serine in peptides immobilized on paper was previously optimized such that only the 1,2-aminoalcohol of N-terminal serine in peptides got oxidized while leaving native vicinal 1,2-diols in cellulose intact. The optimization was achieved using a “pulse-chase” method, followed up with LCMS characterization (Figure 2-6 A). Briefly, we treated unmodified cellulose or paper containing immobilized βAla-βAla-Ser peptide with 1 mM of NaIO₄ in carbonate buffer (pH 8.6) at RT for 5 min, followed by the exposure to fluorescein-5-thiosemicarbazide (FTS) in aniline acetate buffer (pH 4.5). We observed near completion of Ser oxidation to form aldehyde, visualized as orange color and fluorescence of the FTS-conjugate (Figure 2-6 B-C). LCMS confirmed the formation of the conjugate after treating paper with wet NH₃ gas overnight (Figure 2-6 D). On the other hand, unmodified cellulose control that was subjected to the same oxidation condition showed only a minor orange color (Figure 2-6 B: the area surrounding the squares). The same residual orange color was observed when untreated cellulose was exposed to FTS, suggesting that minor color was due to non-specific adsorption of FTS and not due to oxidation of cellulose.

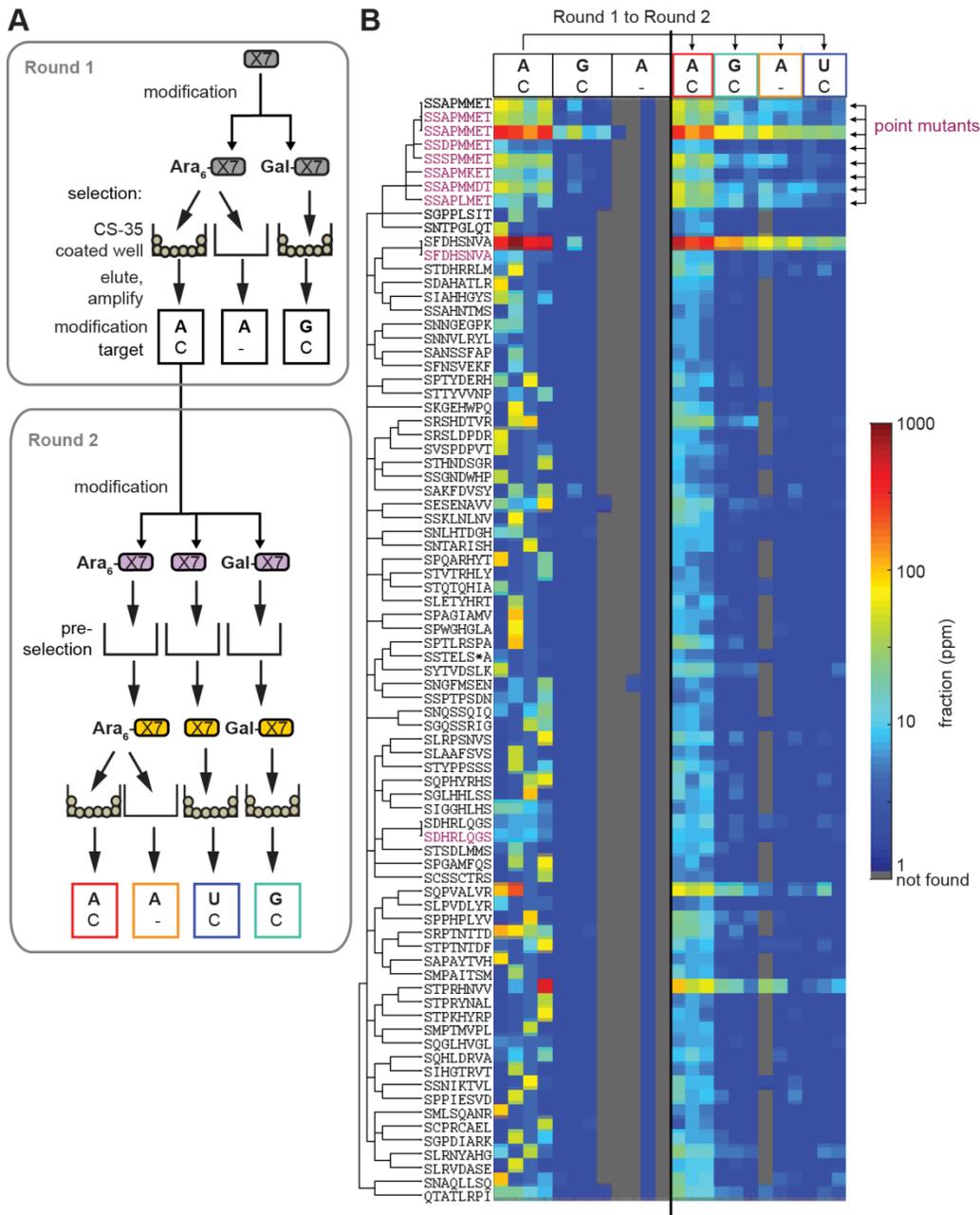


Figure 2-4. Detailed selection scheme and post-selection analysis.

(A) Schemes of panning. (B) Heat map showing 80 putative hits discovered from two rounds of panning. Sequences of putative hits were enriched greater or equal to 4 fold ($R \geq 4$, $p < 0.05$) when compared to control experiments (GC from RI, A- from RI, GC from RII, A- from RII, and UC from RII). Panel B was created by *MakeFigure2_4.m* MatLab script (available in Appendix).

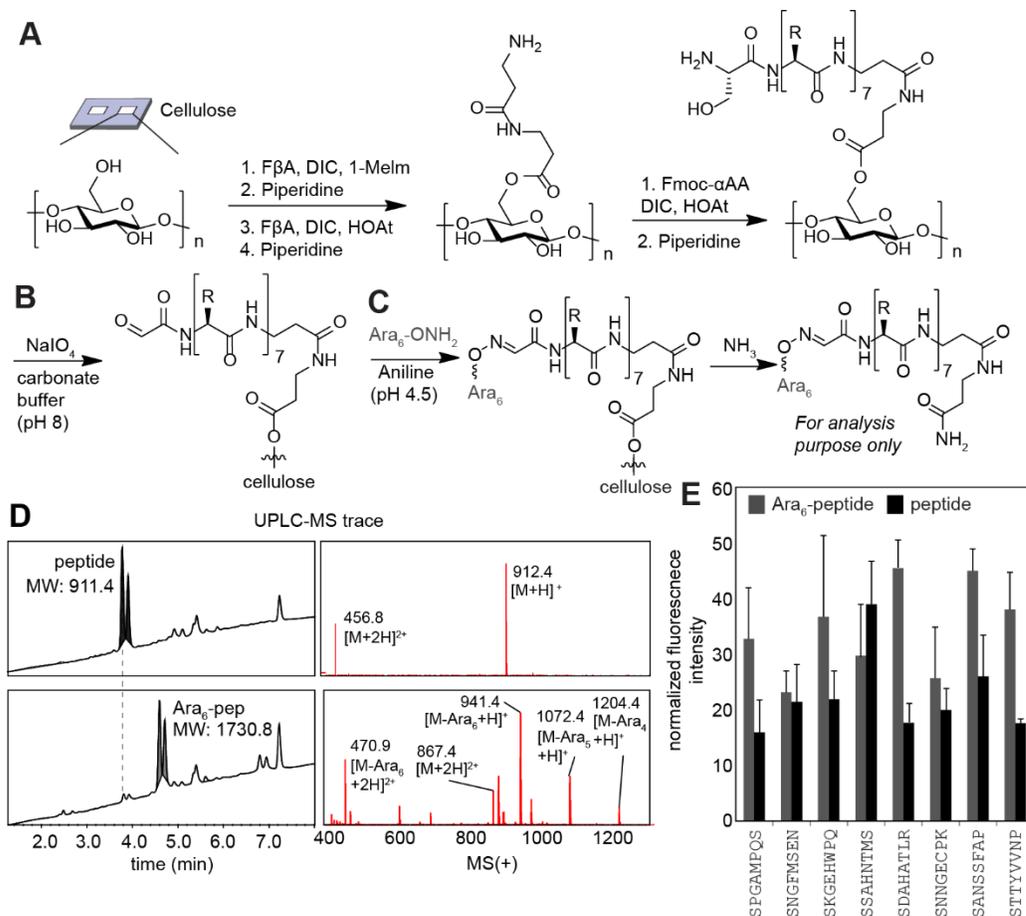


Figure 2-5. Binding of glycopeptides to CS-35 on paper-based assay.

(A) We employed the SyntArray to synthesize peptides on Teflon-patterned paper. (B) Subsequent oxidation and (C) oxime ligation gave rise to glycopeptide arrays. (D) UPLC-MS trace to confirm the formation of Ara₆-peptide (SGPPLSIT-βA-βA) on paper arrays. Only doubly charged species and fragmented singly charged species were observed in MS. (E) The results of paper-based assay showing binding of fluorescently labeled CS-35 to peptides and Ara₆-peptides. The “hits” were defined as Ara₆-peptides that bound significantly more antibody than aglycone-peptide. Fluorescence was obtained using excitation wavelength at 532 nm and long-pass emission filter LPG (≥ 575 nm).

Even when oxidation time was increased to 40 min, cellulose control produced no sign of oxidation (Figure 2-6 B).

To validate and optimize the conversion of peptides to Ara₆-containing glycopeptides on paper, I combined a “pulse-chase” method that can be followed up with

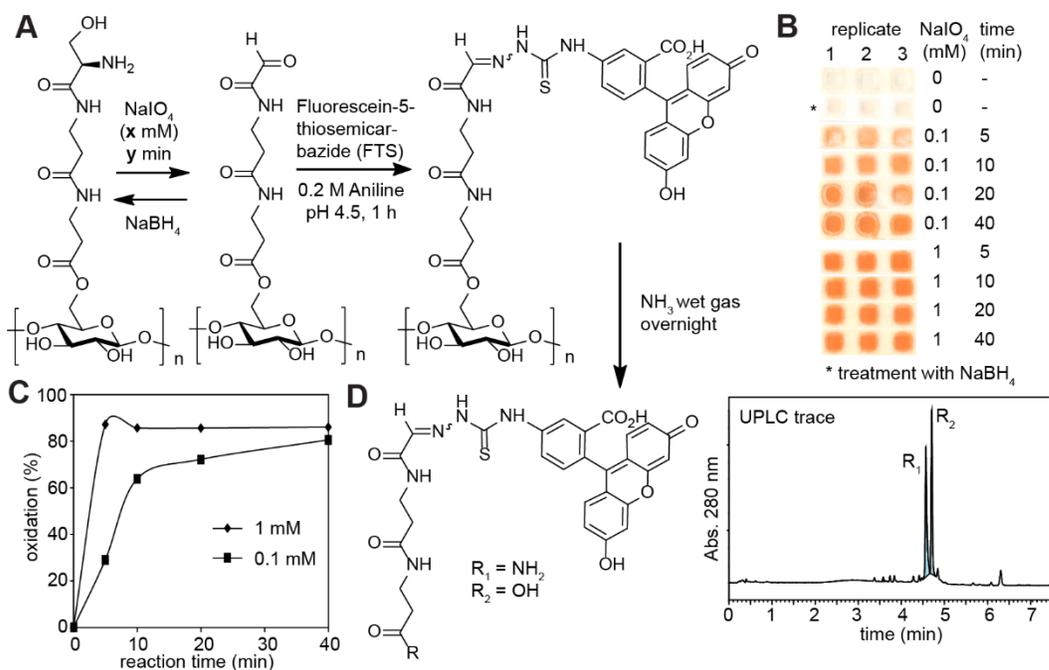


Figure 2-6. Optimization of oxidation conditions for peptides immobilized on paper. (A) Modification of peptides on paper array via aldehyde formation and subsequent ligation to N-terminal aldehyde. (B) Treatment of peptides with different concentrations of NaIO₄ for specified time length, followed by reaction with FTS resulted in the formation of FTS-conjugate (orange). (C) Oxidation went near completion in 5 min using 1 mM NaIO₄. (D) Structures and LCMS validation of FTS-βAla-βAla-conjugates formed in the reaction on paper.

LCMS characterization (Figure 2-7 A). By varying the concentration and amount of aminoxy-derivative (e.g. Rha-ONH₂) spotted onto oxidized peptides on paper (designated as “pulse”) and subsequently chased by the addition of fixed amount of FTS, I observed areas of orange color in different shades (Figure 2-7 B). Little to no orange color after this “pulse-chase” procedure indicated that the oxidized peptide (an aldehyde group) has been consumed in the reaction with Rha- ONH₂ and only a few aldehyde groups were available to react with FTS. Specifically, spotting progressively increasing concentrations of Rha-ONH₂ (0, 1, 2, 5 mM) and followed by FTS-chase produced fainter shades of orange color on paper. These findings were corroborated by the LCMS of the

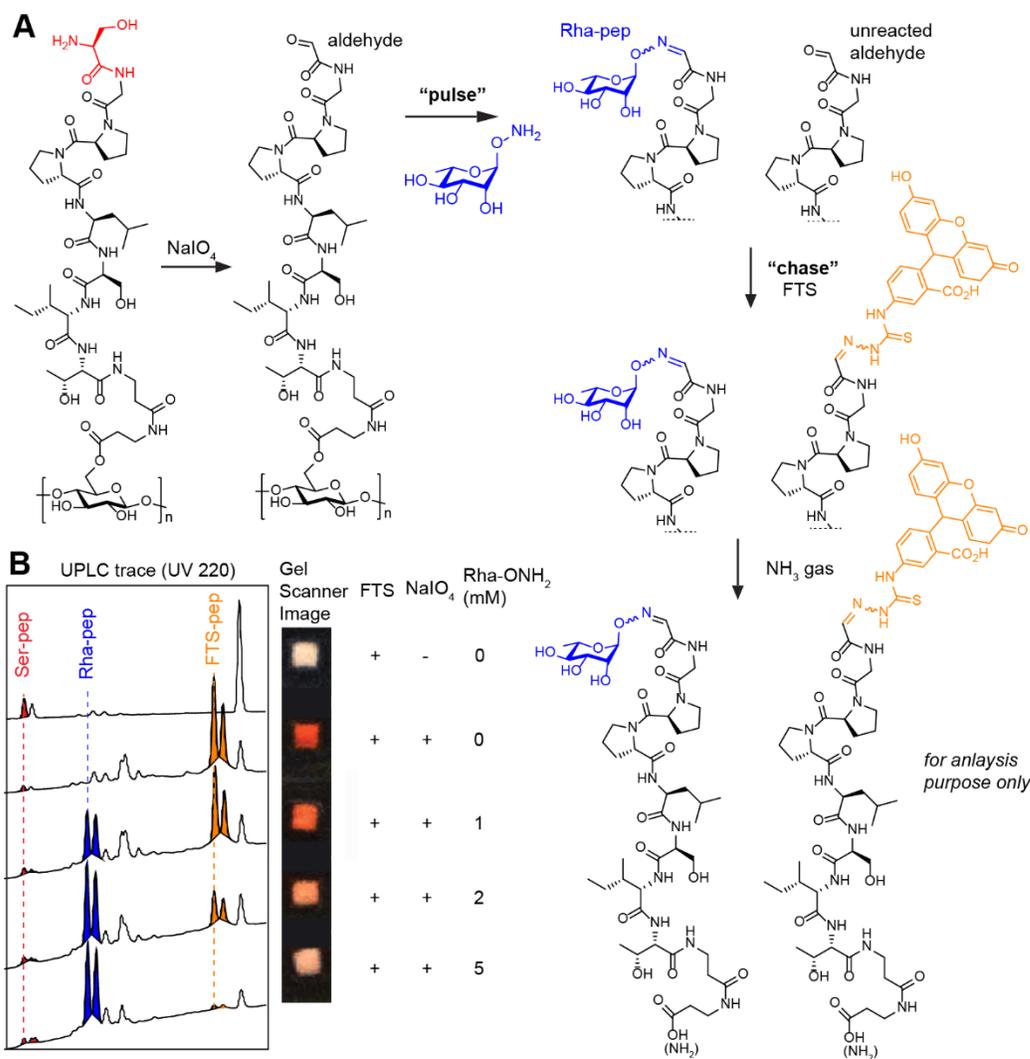


Figure 2-7. Optimization of glycopeptide formation on paper.

(A) N-terminal serine is oxidized, followed by the addition of increasing concentration of aminoxy derivative (Rha-ONH₂). FTS is then added to react with remaining oxidized but un-ligated peptides. (B) FTS readily reacts with oxidized peptides, forming FTS-peptide, and can be visually detected by the orange color on paper. Both colorimetric assay and LCMS confirmed formation of Rha-peptide.

peptides cleaved from paper. This “pulse-chase” method coupled to a colorimetric readout can allow rapid, visual optimization of the reaction conditions.

Both colorimetric assay and LCMS confirmed that treatment with 5 mM Rha-ONH₂ for 1 h quantitatively converted the peptide glyoxal to rhamnose oxime derivative

(Figure 2-7 B). Using the optimized oxidation and oxime ligation conditions, we derivatized eight peptides immobilized on paper with Ara₆-ONH₂ (Figure 2-5 B-C). For a selected peptide SPQARHYT, we confirmed the formation of the Ara₆-PQARHYT conjugate on paper using LCMS (Figure 2-5 D). Exposing these arrays to CS-35 labeled with tetremethylrhodamine isothiocyanate (TRITC) allowed quantifying the binding of TRITC-CS-35 to paper-immobilized glycopeptides or control peptides. Glycopeptides Ara₆-DAHATLR, Ara₆-ANSSFAP, and Ara₆-TTYVVNP exhibited significantly higher binding to TRITC-CS-35 than the aglycon peptide sequences SDAHATLR, SANSSFAP, STTYVVNP (Figure 2-5 E). Interestingly, the peptide SSAHNTMS displayed higher fluorescence intensity than Ara₆-SAHNTMS. These three “positive” and one “negative” glycopeptides were re-synthesized for further validation, using a standard solid-phase peptide synthesis and bioconjugation.

2.2.4 Glycopeptides testing using ESI-MS assay

Although the conditions for the synthesis of oxime glycopeptides have been reported previously, the published conditions resulted in partial oxidation of methionine residue in peptides.¹⁸⁰ To minimize the oxidation of methionine in SSAHNTMS peptide during oxidative cleavage of Ser by NaIO₄ we noted that the rate of Met oxidation is known to be slower than the rate of cleavage of N-terminal Ser.^{180, 230} Armed with this knowledge, we confirmed that kinetically controlled oxidation with 60 μM NaIO₄, for 5 min at 4 °C used for modification of phage libraries did not produce any detectable amounts of oxidized methionine. On the other hand, the use of higher concentration (1 mM or 10 mM), used in previous reports for a preparative synthesis of peptide glyoxals¹⁸⁰

resulted in either partial or complete oxidation (Figure 2-7). LC-MS confirmed the presence or absence of oxidized methionine in these reactions.

With the help of Dr. Elena Kitova from Dr. John Klassen group, I used ESI-MS to measure the binding constants (K_D) of these four Ara₆-containing glycopeptides, the parent glycan Ara₆, parent peptide sequences, and glucose-modified glycopeptides to CS-35 F_{ab} and 906.4321 F_{ab} (Appendix) (Table 2). The binding constants of Ara₆-peptides were measured on two different days with two different stocks of antibody F_{ab} fragments, and the absolute values varied between two experiments; however, the relative values and trends were consistent between these two experiments. The values of ligand **3**, **6**, **9**, **12**, **13** in Table 2 represent the average.

We found that the monovalent binding affinity of Ara₆-ANSSFAP and Ara₆-TTYVVNP was 10 fold lower than the K_D of the glycan Ara₆ to CS-35 F_{ab}. The K_D of the same glycopeptides binding to 906.4321 F_{ab} was only 2 fold lower than that of Ara₆ binding to this F_{ab}. The other ligand, Ara₆-DAHATLR, exhibited K_D to CS-35 F_{ab} 3 fold lower when compared to that of Ara₆. On the other hand, the K_D of Ara₆-DAHATLR for 906.4321 F_{ab} increased 1.2 fold when compared to Ara₆, suggesting that the peptide provided no synergy but rather an interference with binding to 906.4321. This observation can be rephrased as: Ara₆-DAHATLR exhibited 3 fold improved selectivity—determined as the ratio of K_D (antigen-906 Fab) / K_D (antigen-CS-35)—when compared to the Ara₆ antigen. Using the same notation, Ara₆-ANSSFAP, Ara₆-TTYVVNP, and Ara₆-SAHNTMS exhibited 5-6 fold improved selectivity and 3-14 fold of enhanced affinity towards CS-35 than Ara₆.

Table 2. K_D of ligands to CS-35 and 906.4321 F_{ab} measured by ESI-MS

	K_D CS-35 F_{ab} (μ M)	K_D 906.4321 F_{ab} (μ M)
SANSSFAP (1)	372 ± 107	710 ± 344
Glu-ANSSFAP (2)	N/A*	630 ± 852
Ara ₆ -ANSSFAP (3)	1.4 ± 1.0	20 ± 3
STTYVVNP (4)	463 ± 282	2090 ± 1402
Glu-TTYVVNP (5)	N/A*	N/A*
Ara ₆ -TTYVVNP (6)	1.9 ± 1.2	33 ± 4.7
SSAHNTMS (7)	312 ± 114	1082 ± 732
Glu-SAHNTMS (8)	N/A*	83.3*
Ara ₆ -SAHNTMS (9)	3.2 ± 1.0	47 ± 13
SDAHATLR (10)	388 ± 146	1012 ± 310
Glu-DAHATLR (11)	N/A*	N/A*
Ara ₆ -DAHATLR (12)	7.0 ± 6.2	59 ± 8.3
Ara ₆ (13)	19 ± 9.1	52 ± 31

* K_D of ligands could not be measured due to overlap of the mass of F_{ab} -ligand with the mass of F_{ab} .

The peptide aglycones exhibited measurable binding to CS-35 (Figure 2-8), but much higher K_D than Ara₆-containing glycopeptides. Unfortunately, binding constants of Glu-peptides to CS-35 were not attainable through ESI-MS due to the nature of the ligands (see Table 2), but estimated binding of Glu-peptides were significantly weaker than that of Ara₆-peptides. The collective data show that to provide enhanced binding and selectivity towards CS-35, Ara₆, the peptide sequence, and the linker connecting them must work in synergy.

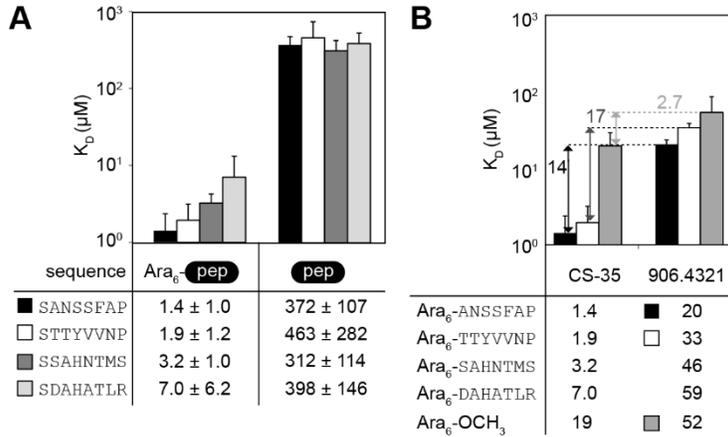


Figure 2-8. ESI-MS measurements of discovered Ara₆-peptides. (A) ES-MS measured K_D for peptides, Ara₆-peptides and glu-peptides to CS-35 F_{ab}. (B) Binding specificity of Ara₆-peptides to CS-35 and 906.4321 F_{ab}.

2.2.5 Glycopeptides testing in microarray format using monoclonal antibodies

To examine the enhancement of selectivity of the discovered glycopeptides in multivalent serological assay, we employed squarate ligation to convert six glycopeptides (2, 3, 4, 5, 11, 12) and Ara₆ (13) to multivalent BSA conjugates that contain an average low (2-4), medium (5-9), or high copy number (more than 9) of glycopeptides per BSA molecule. The average number of glycopeptides conjugated per BSA molecule was determined by MALDI (Appendix). These glycopeptides have been printed and reprinted to investigate the reproducibility of printing. In addition, I re-synthesized and printed Glu-ANSSFAP and Glu-TTYVVNP BSA conjugates in order to examine whether re-synthesis can generate reproducible results.

We screened the arrays using a dilution series of anti-LAM IgGs—CS35, CS-40 and 906.4321, visualized through the addition of fluorescently labelled anti-mouse IgG

secondary antibody (Figure 2-9 A-C). This assay yielded dose-response curves for every ligand-antibody pair. The curves, when fit to a Hill equation

$$F(c) = F_{max} \times \frac{1}{1 + \left(\frac{EC_{50}}{c}\right)^n}$$

yielded the best fit for $n=1$ (i.e. non-cooperative binding) and both EC_{50} and maximum fluorescent intensity F_{max} . Comparing the EC_{50} across the panel of ligands and antibodies, we observed indistinguishable EC_{50} values towards CS-35 and 906.4321 from high-valency Ara_6 -peptide as well as all Ara_6 antigens, regardless of valency. For instance, the EC_{50} of $(Ara_6\text{-TTYVVNP})_{12}\text{-BSA}$, $(Ara_6)_{12}\text{-BSA}$, or $(Ara_6)_4\text{-BSA}$ for CS-35 and 906.4321 were 2-3 nM. On the other hand, the Ara_6 -peptides displayed at low valency exhibited modestly improved selectivity towards CS-35; the EC_{50} of $(Ara_6\text{-TTYVVNP})_2\text{-BSA}$ for CS-35 was 2 nM and for 906.4321 was 25 nM. The differences in EC_{50} translates to 13 fold selectivity. Other low valency Ara_6 -peptides— $(Ara_6\text{-ANSSFAP})_2\text{-BSA}$, $(Ara_6\text{-DAHATLR})_2\text{-BSA}$, yielded similar EC_{50} for CS-35 and exhibited 20 and 13 fold selectivity towards CS-35.

As expected, the control ligands, i.e. glucose-containing peptides, regardless of valency, all showed binding signals to CS-35 and 906.4321 within the range of noise. These observations highlighted that selectivity and binding affinity of multivalent presentation of ligands can only be achieved when the glycopeptide ligands contain the correct carbohydrate moiety (i.e. Ara_6) and only when these Ara_6 -peptides were presented in a low valency fashion.

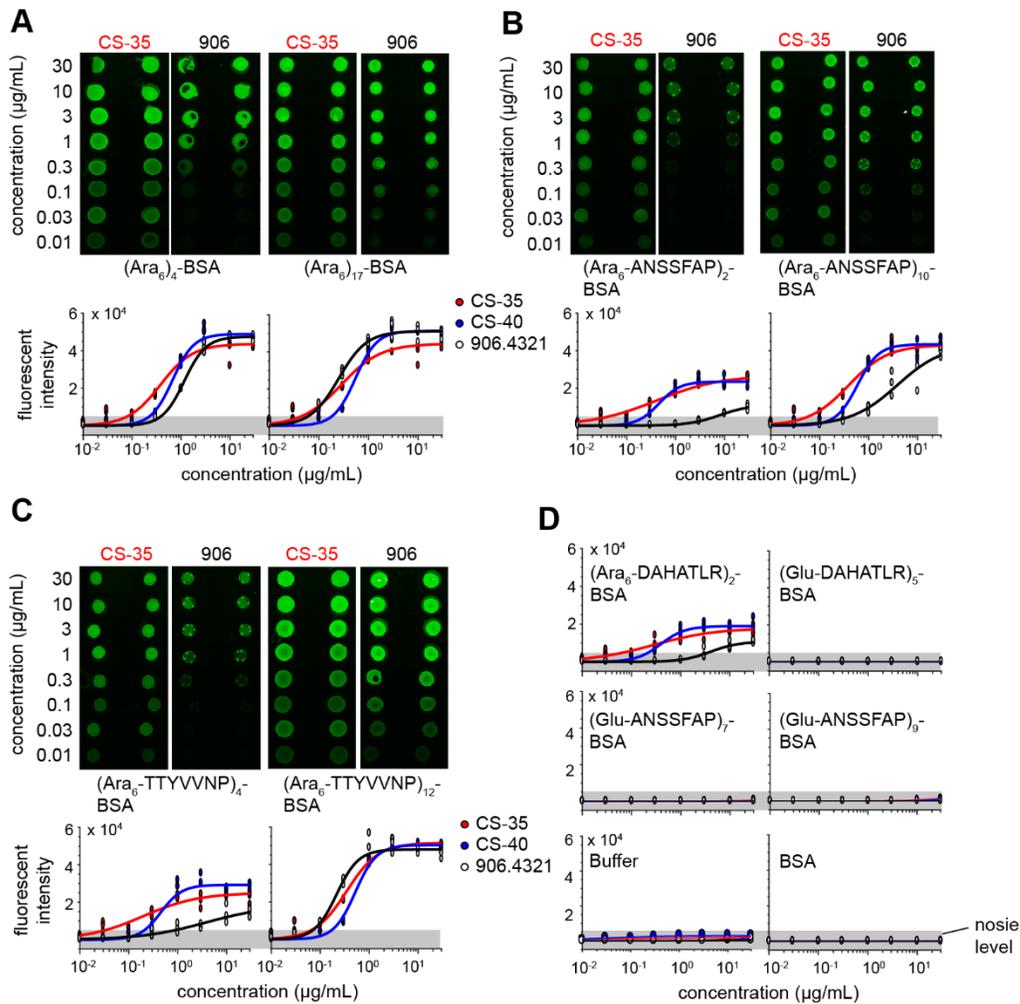


Figure 2-9. Ara₆-peptides binding in multivalent microarray assay. Fluorescence images and dose-response curves of (A) (Ara₆)₄-BSA, (Ara₆)₁₇-BSA, (B) (Ara₆-ANSSFAP)₂-BSA, (Ara₆-ANSSFAP)₁₀-BSA, and (C) (Ara₆-TTYVVNP)₄-BSA, (Ara₆-TTYVVNP)₁₂-BSA binding to CS-35, CS-40, and 906.4321. (D) Dose-response curves of (Ara₆-DAHATLR)₂-BSA and glucose-containing ligands.

2.3 Conclusion

In conclusion, I described the discovery of glycopeptide ligands for anti-LAM antibody CS-35 from phage displayed library of chemically modified peptides. To

facilitate discovery validation of these ligands, I optimized glycopeptide synthesis in three different modalities – on phage, on cellulose, and in solution. In all formats, kinetically controlled oxidation and aniline catalyzed oxime formation effectively converted unprotected peptides to glycopeptides. I show that GE-FBD discovered ligand with higher binding affinity and specificity than the parent glycan in monovalent binding assay. When incorporated into a multivalent binding assay using surface-immobilized antigens and bivalent IgG antibodies, only glycopeptides with Ara₆ presented at low valency (2-4 ligands per BSA) exhibited preferential binding to one anti-LAM antibody over another.

I demonstrated that the multivalent presentations of ligands on the array has improved the selectivity moderately, from 5-6 fold enhancement in monovalent assays to 8-10 fold in multivalent assay. The enhanced effect is prominent only when glycopeptide ligands have the Ara₆ moiety, and only when presented in a low-valency fashion.

While these discovered glycopeptides provided higher selectivity than the carbohydrate Ara₆, the enhancement is rather underwhelming. In the next chapter, I undertook a series of investigations to study the structure-activity relationships of the fragments of discovered glycopeptides to the anti-LAM antibody CS-35, aiming to augment the selectivity further in multivalent assays.

2.4 Experimental procedures

2.4.1 Materials and general information

MOPS buffer contains 50 mM MOPS, 150 mM NaCl, 2 mM CaCl₂ with pH adjusted to 7.4 after preparation. All solutions used for phage work were sterilized by filtration through 0.22 μm filters. Production of mAb IgG of CS-35, CS-40, 906.4321 and F_{ab} of CS-35 and 906.4321 was previously reported.²²⁴ The synthesis of Ara₆-OCH₃ was reported previously.²³¹ RP-HPLC were performed on Waters HPLC system equipped with a Waters 1525 EF binary pump, a Waters FlexInject manual injector (dual mode) and a Waters 2489 tunable UV detector. SymmetryPrepTM C18 semi-preparative column (19 × 50 mm, particle size 5 μm, pore size 100 Å) and XBridge BEH Amide OBD Prep column (19 × 250 mm, particle size 5 μm, pore size 130 Å) were used for the purifications of peptides and N-terminal serine oxidized peptides at a typical flow rate of 12 mL/min. Symmetry C18 analytical column (4.6 × 75 mm, particle size 3.5 μm, pore size 100 Å) was used for all the purifications of Ara₆ glycopeptides glucose glycopeptides at a typical flow rate of 1 mL/min. Aminoxy-biotin (#10009350) was purchased from Cayman Chemical. Characterization of Glu-peptides, Ara₆-peptides and Ara₆-peptide squaramides were performed by UPLC-MS. Characterization of Ara₆-peptide-BSA conjugates were performed by MALDI. Microarrays were printed at Engineering Arts LLC (Phoenix, Arizona) using Schott NEXTERION[®] Slide E and piezoelectric non-contact printing.

2.4.2 Chemical modification of Phage library

N-SerX₇ phage-displayed peptide library (complexity: 3 x 10⁸ PFU) was generated by our collaborators at New England Biolabs according to the referred protocol.^{52, 232} Prior to the chemical modification, the phage library (2 x 10¹² pfu) was dialyzed extensively against 4 L PBS (4 °C, pH 7.4, 4 buffer changes over 24 h, 10K MWCO) to remove the storage buffer which contains 50% (v/v) glycerol. For chemical modification, the phage library (100 μL, 10¹¹ PFU/mL) was oxidized with 0.06 mM sodium periodate (by adding 1 μL of 6 mM NaIO₄ solution in MQ water) on ice for 5 min. The oxidation was quenched by adding 1 μL of 50 mM solution of glutathione in MQ water and incubation at RT for 10 min. The oxidized library was distributed into three separate portions, which were treated with an equal volume of aminoxy-hexasaccharide (2 mM Ara₆-ONH₂ in 200 mM anilinium acetate buffer, pH 4.5), aminoxy-galactose (2 mM Gal-ONH₂ in 200 mM anilinium acetate buffer, pH 4.5), or aminoxy-biotin (2 mM AOB in 200 mM anilinium acetate buffer, pH 4.5) to yield modified libraries designated Ara₆-X₇, Gal-X₇, and Bio-X₇, respectively. The reaction mixtures were incubated for 1 h at RT, after which, the modified libraries were buffer exchanged twice in Amicon® Ultra-4 Centrifugal filter units 10K MWCO using MOPS buffer. Each buffer exchange was performed at 4700 rpm on Heraeus™ Multifuge™ X3R Centrifuge for 15 min. The recovered phage was used for panning experiments. To monitor the oxidation, a small portion of the Bio-X₇ library was diluted and captured with biotin-capture assay as previously described.¹⁸⁰ Typically, up to 80% of monoclonal phage and 40% of phage library were successfully oxidized and biotinylated. To quantify the modification with Gal-ONH₂ and Ara₆-ONH₂, a small portion of the Ara₆-X₇ and Gal-X₇ libraries were

treated with AOB and characterized via biotin-capture assay as previously described.¹⁸⁰ Typically, 60-80% of the oxidized fractions of phage library has been converted to Ara₆-ONH₂ and Gal-ONH₂ conjugates.

2.4.3 Phage selection

2.4.3.1 First round of selection

A 96-well polystyrene plate (Corning #3369) was coated with a solution of CS-35 (100 μ L, 40 μ g/mL) in PBS for overnight at 4 °C. After coating, the plate was rinsed with washing solution (10 \times 300 μ L, 0.1% (v/v) Tween-20 in MOPS buffer) using 405TM Touch Microplate Washer (BioTek). The selection of Ara₆-X₇ library against CS-35 was performed in four separate wells in parallel with the control selections, i.e., Gal-X₇ against CS-35, Ara₆-X₇ against blank wells. Specifically, the modified libraries were added into the corresponding wells (100 μ L/well, 10⁹ PFU/well). After incubating for 1 h at RT, the unbound phage was rinsed with the washing solution (10 \times 300 μ L) using the plate washer. Phage remained in the wells was eluted by adding 200 μ L of glycine elution buffer (0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA). After 9 min of incubation at RT, the elution buffer was transferred into a 1.5-mL microcentrifuge tube containing 30 μ L neutralization buffer of 1 M Tris-HCl (pH 9.1). The recovered phage libraries, 12 samples in total, were processed further as described in subsequent section “Phage amplification and preparation for sequencing.”

2.4.3.2 Second round of selection

After the first round of selection, the amplified phage recovered from the selection of Ara₆-X₇ library against CS-35 was modified with Ara₆-ONH₂ or Gal-ONH₂ as described in previous section “Chemical modification of phage library.” The modified phage was subjected to the second round of selection using the procedure described above, with the following modification: (i) the stringency of washing was increased by using 0.2% (v/v) Tween-20 in MOPS buffer; (ii) all libraries (SX₇, Gal-X₇, Ara₆-X₇) have been preselected on blank wells. Briefly, the libraries (100 μL, 10⁹ PFU/well) were incubated in a single well (not blocked) for 1 h at RT. The supernatants from each well were then used as inputs for the second round of selection.

2.4.3.3 Third round of selection

The third round of selection followed the same procedures as the second round of panning, except for the following modification: the stringency of washing was increased by using 0.5% (v/v) Tween-20 in MOPS buffer.

2.4.4 Phage amplification and preparation for sequencing

The eluted phage was amplified separately by adding the eluates (230 μL) into 6 mL of ER2738 culture (1:100 dilution of overnight culture). The phage and bacterial mixtures were incubated for 4.5 h at 37 °C with vigorous shaking. The cultures were centrifuged (15 min, 4700 rpm) at 4 °C to pellet the bacterial cells. An aliquot (~ 4.5 mL)

of the amplified phage from selection of Ara₆-X₇ library against CS-35 was transferred to a fresh centrifuge tube and precipitated with PEG/NaCl. The combined precipitated library from each replicate served as the input for the subsequent round of selection. The remaining supernatants (~ 1.5 mL) containing the amplified phage were transferred into fresh tubes and the phage ssDNA was isolated using QIAprep spin M13 kit (Qiagen, #27704) according to manufacturer's instructions. The ssDNA was then converted to Illumina-compatible dsDNA amplicon by PCR. Briefly, the ssDNA (~50–100 ng) was combined with 1x Phusion[®] buffer, 200 μM dNTPs (each), 0.5 μM forward and reverse primers, and one unit Phusion[®] High-Fidelity DNA Polymerase in a total volume of 50 μL. Forward (F) and reverse (R) primer sequences, 5' → 3':

F: 5'-CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCCTG

CTGAACCGCTCTTCCGATCTXXXXCCTTTCTATTCTCACTCT-3'

R: 5'-AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACA

CGACGCTCTTCCGATCTXXXXACAGTTTCGGCCGA-3'

The XXXX in the primer sequence denotes four-nucleotide-long barcodes used to trace multiple samples in one Illumina sequencing experiment. The temperature cycling protocol for PCR was as follows: 95 °C for 30 s, followed by 25 cycles of 95 °C for 10 s, 60.5 °C for 15 s and 72 °C for 30 s, and then a final extension at 72 °C for 5 min before holding at 4 °C. The resulting amplicons were characterized by gel electrophoresis using 2% agarose gel, pooled together (20 ng per sample) and purified on E-Gel[®] SizeSelect[™] 2% agarose gel (Invitrogen, #G6610-02). Sequencing was performed using the Illumina

NextSeq platform (Molecular Biology Service Unit, University of Alberta) and analyzed as described in the subsequent section “Analysis of the deep-sequencing data.”

2.4.5 Analysis of the deep-sequencing data

Analysis was adapted from Ng *et al.*⁵² and Matochko *et al.*²²⁸ with minor modifications. Raw FASTQ data were processed using MatLab scripts described in our previous publication.²³³ In short, the scripts use regular expressions to identify sequencing barcodes at the front and back end of the read (see red **XXXX** sequences in the previous section), and isolate the library region (crimson box, in Figure 2-10), which is located between two constant adapter regions (grey in the scheme above). Sequences and their copy numbers in each sequencing experiment were stored as $N \times M$ table, where N is a total number of unique sequences and M is total number of experiments and replicates (Figure 2-10).

Identification of significantly enriched sequences from deep-sequencing data was performed similarly to procedure described in Ng *et al.*⁵² and Tjhung *et al.*²³³ and the entire processing algorithm is bundled in one MatLab script entitled *MakeFigure2_4.m* available in Appendix. In short, for each sequence we calculated the average normalized frequency at which it appeared in each sequencing set and then the ratio as average frequency in test experiment divided by an average frequency in the control experiment. We also calculated the p-value between the replicates of test and replicates of control experiment using a two-sided unequal variance t-test. Hit sequences were defined as those that had $p < 0.05$ and $\text{ratio} \geq 4$ for all control experiments and were presented as heat map,

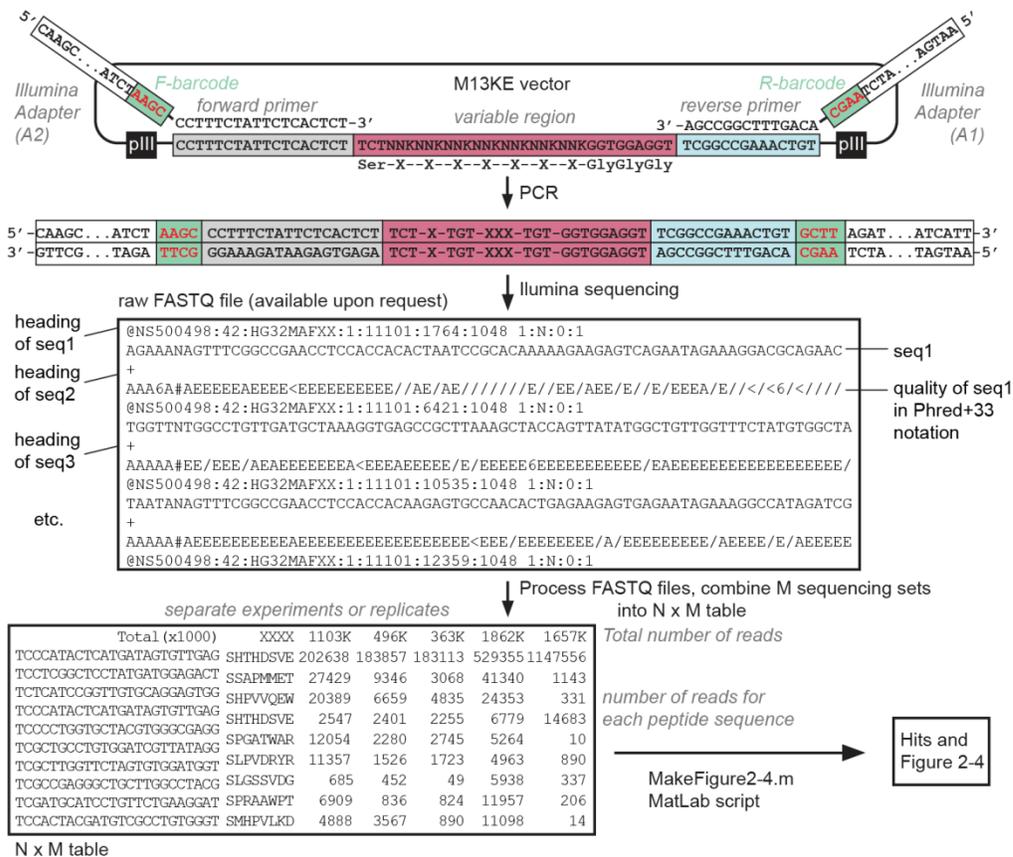


Figure 2-10. Analysis of the deep-sequencing data

as described in Figure 2-3 E and Figure 2-4. Figure 2-4 can be regenerated on demand using *MakeFigure2_4.m* MatLab script, which is available in Appendix.

2.4.6 Synthesis of glycopeptides on paper.

Synthesis of peptides on paper is previously reported by Deiss et al.²²⁹ Prior to ligation of carbohydrates onto peptides established on paper, solid-ink barriers on peptide-containing areas on the arrays were printed for downstream modification of peptides. Five minutes incubation of the array in 120 °C oven generates a wax-patterned paper array that has areas resistant to aqueous solutions. To oxidize the peptides, peptide

arrays were first washed in carbonate buffer, pH 8 to wash away residual TFA. The array was then immersed in 5 mM of sodium periodate (10 mL) in carbonate buffer (pH 8) for 30 min, kept dark at RT. Arrays were then washed with anilinium acetate buffer (pH 4.5) and dried with Kimwipe. Aminoxy-derivative (Ara₆) was spotted at this array (3 μ L, 5 mM) and the array was incubated in a humidity chamber at RT for 1.5 h. After oxime ligation, the array was washed with H₂O, dried in air, and stored at -20 °C until use.

To optimize the oxidation and oxime ligation reactions, we used different concentration of aminoxy rhamnose (1 mM, 2 mM, 5 mM) for oxime ligation. We then submerged the array into solution of fluorescein thiosemicarbazide (FTS, 1 mM in 200 mM anilinium acetate, pH 4.7). This step converted all remaining un-ligated, oxidized peptides to with FTS conjugates, giving rise to orange color on paper. The relative amounts of Rha-peptide and FTS-peptides were also monitored by LC-MS.

2.4.7 TRITC labeling of mAb CS-35.

0.47 mg/mL of CS-35 was dialyzed against 4 L PBS (4 °C, pH 7.4, 4 buffer changes over 24 h, 10K MWCO). 25 mg/mL of TRITC dissolved in DMF was added into recovered CS-35 solution buffered with carbonate to pH 9. Reaction mixture was incubated on a rocker at RT for 2 hrs. The reaction mixture was buffer exchanged with PBS in Amicon Ultra-4 Centrifugal filter units to eliminate the excess of TRITC. The concentration of conjugated antibody was determined from UV-vis spectrum (A_{280} of IgG 1% = 14; ϵ_{280} of IgG = 210,000 M⁻¹cm⁻¹), taking into account that absorbance contribution from the covalently bound TRITC dye at 550 nm and correction factor (CF)

adjusted for the amount of absorbance at 280 nm caused by the dye (CF=0.35) (ThermoScientific TR0031.7):

$$IgG (M) = \frac{A_{280} - 0.34 \times A_{550}}{210000}$$

2.4.8 Paper-based fluorescence binding assay.

Glycopeptide arrays were immersed in blocking buffer (2% BSA, 0.4% Tween in PBS) at RT for 1 h. In a small petri dish, 5 μ g/mL of TRITC-labeled CS35 in blocking buffer were added to glycopeptide arrays. The arrays were incubated at 37 °C for 30 min. The arrays were then washed 2 times with 0.4% Tween in PBS. Fluorescent images were obtained with Typhoon FLA 9500 (GE Healthcare Life Science), using excitation wavelength at 532 nm and long-pass emission filter LPG (\geq 575 nm).

2.4.9 Synthesis of peptides on solid support

Peptides were synthesized using Prelude X Instrument (Protein Technologies, Inc.) Briefly, Rink Amide AM resin (200 mg, 0.91 mmol g^{-1} , 0.18 mmol) was transferred into reaction vessel. CH_2Cl_2 (5 mL) was added to the dried resin for swelling. After 10 min, the solvent was drained. The resin was washed with DMF (2 \times 5 mL) and then deprotected with 20% (v/v) piperidine in DMF (5 mL) for 3 min. The deprotection was repeated for another 3 min using fresh 20% (v/v) piperidine in DMF (5 mL). The resin was washed with DMF (4 mL) for 1 min, followed by CH_2Cl_2 wash (4 \times 4 mL). Fmoc-protected amino acid (1 mmol) dissolved in DMF (5 mL) and HBTU (1 mmol) dissolved in DMF (3 mL) were added to the resin, followed by DIPEA (2 mL, 2 mmol). After 15

min of agitation with nitrogen and heating at 50 °C, the reagents were removed and the resin was washed with DMF (6 × 4 mL). The Fmoc-deprotection, amide coupling, and washing steps were repeated consecutively to elongate the peptide sequence. After final Fmoc-deprotection, the resin was washed with DMF (6 × 4 mL), followed by CH₂Cl₂ (5 × 4 mL). The resin was drained and dried in reaction vessel and then transferred to Poly-Prep® chromatography column. A cleavage cocktail (2 mL) containing TFA/H₂O/phenol/TIPS [85/5/5/5 (v/v/w/v)] was added to the dried resin. The column was left on a rocker for 4 h to cleave the peptide. The flow-through from the column was collected and the resin was rinsed with TFA (1 mL). The combined cleavage mixture was added dropwise to cold diethyl ether (20 mL) in a centrifuge tube. The mixture was incubated on ice for 30 min. The precipitates were centrifuged for 5 min at 3000 rpm. Supernatant was decanted and the precipitates were washed with cold diethyl ether (10 mL). The centrifugation and washing steps were repeated for another two cycles. The precipitates were air-dried and then left under vacuum for overnight. Typical yield: 50-100 mg (crude).

Crude peptide (50 mg) was dissolved in MQ water (0.5 mL); and if peptide did not dissolve, acetonitrile was added in small aliquots (0.05 mL) until the solution appeared clear. The solution was injected into a semi-preparative RP-HPLC system with C18 column following the gradient attached to each HPLC trace of peptides in Appendix using solvent A (MQ water, 0.1% (v/v) TFA) and solvent B (MeCN, 0.1% (v/v) TFA) at a flow rate of 12 mL/min. The fractions corresponding to the main peak were collected. MeCN was removed by evaporation under reduced pressure. The aqueous solution was lyophilized to yield the peptide as white fluffy powder (10–50 mg).

2.4.10 Synthesis of Ara₆-peptide conjugates

2.4.10.1 Oxidation of N-terminal serine

The procedure was adapted from Ng et al.¹⁸⁰ The oxidation of N-terminal serine on peptides SANSSFAP, SDAHATLR, and STTYVVNP followed the same procedure. Example of the procedure for ox-ANSSFAP: Peptide SANSSFAP (10 mg, 13 μ mol, 1 eq.) was dissolved in MOPS buffer (0.5 mL, 200 mM, pH 7.0) in a 1.5-mL microcentrifuge tube. Sodium periodate (3.3 mg, 15.6 μ mol, 1 eq.) was added to the microcentrifuge tube and the reaction mixture was incubated for 1 h at RT.

The oxidation of N-terminal serine on peptide that contained methionine (here: SSAHNTMS) and its purification was carried out by the following: peptide SSAHNTMS (4 mg, 5 μ mol, 1 eq.) was dissolved in 2.25 mL ice-cold PBS. Ice-cold solution of sodium periodate (15 μ L, 400 mM, 1.2 eq.) was added for 7 sec, and the reaction was quenched by the addition of GSH (50 μ L, 400 mM, pH 8, 4 eq.). Reaction was vortexed and immediately injected into semi-reparative RP-HPLC with Amide column. The gradient can be found in Appendix. The fraction containing the product was lyophilized to yield a product as white solid (10%).

2.4.10.2 Ligation of Ara₆-ONH₂: General procedure

The N-terminal oxidized peptides (0.3 μ mol, 1 eq.) were dissolved in 60 μ L water in a 1.5-mL microcentrifuge tube. Ara₆-ONH₂ (0.3 mg, 0.36 μ mol, 1.2 eq.) was dissolved in 18 μ L of water and added to the microcentrifuge tube, followed by the addition of 500

mM anilinium acetate (19 μ L). The oxime ligation was carried out for 1 h at RT. The reaction mixture was injected into a LC/MS system (Agilent 1100 LC/MSD) using symmetry C18 analytical column (WAT066224). Purification was carried out using gradients described in Appendix.

Purification yielded Ara₆-peptide conjugates as a translucent film after lyophilisation. The purity and identity of the product was determined with UPLC-MS (Appendix). The recovered mass of the product and the yields of each conjugate was determined by NMR (Appendix).

2.4.11 Synthesis of Glu-peptide conjugates

The procedure was adapted from Ng *et. al.*⁵² Example of the procedure for Glu-DAHATLR: the peptide SDAHATLR (8 mg, 9 μ mol, 1 eq.) was dissolved in 200 mM MOPS (0.5 mL, pH 7.0). The solution was added to a 1.5-mL microcentrifuge tube containing sodium periodate (2.3 mg, 12 μ mol, 1.2 eq.). The reaction mixture was incubated for 1 h at RT. Glu-ONH₂ (2.7 mg, 13.5 μ mol, 1.5 eq.) dissolved in 200 mM anilinium acetate (0.5 mL, pH 4.7) was added to the reaction mixture. The oxime ligation was carried out for 1 h at RT. The reaction mixture was then injected into a RP-HPLC system using gradients described for each glycopeptide in Appendix. Lyophilization yielded Glu-DAHATLR as white fluffy powder (11%). The purity and identity of the product was confirmed with by UPLC-MS (see Appendix for spectra).

2.4.12 NMR to determine product yield.

The concentration of Ara₆-glycopeptides were measured by NMR, using 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as an internal standard for chemical shift and concentration reference. Lyophilized glycopeptides were dissolved in 150 μL H₂O. 70 μL was taken for ¹H NMR in analysis and combined with 630 μL D₂O. The ¹H spectrum was collected on a 700 MHz, 4 channel Agilent instrument, fitted with a HCN Z-gradient Cold probe at 27 °C. The acquisition software was VNMRJ 4.2A. The relaxation delay was 5 s which included 2 s of presaturation of the residual water peak (HOD). The acquisition time was 3 s and the spectral window was set to 12 ppm, 11 to -1 ppm. A reduced flip angle of 30 degrees was applied to ensure accurate integration. Relaxation delay was set at 5 sec. (See Appendix for the spectra).

2.4.13 ES-MS binding measurements.

The association constants (K_a) for F_{ab} binding to ligands were determined using the direct ESI-MS binding assay. The sample solutions were prepared from aqueous stock solutions of F_{ab} and ligand with known concentrations. Aqueous ammonium acetate was added to the sample solutions to yield a final buffer concentration of ~200 mM. Direct ESI-MS binding assay was performed on a Synapt G2S Q-IMS-TOF mass spectrometer (Waters, Manchester, UK) equipped with nanoflow ESI (nanoESI) source. NanoESI was performed by applying a voltage of ~1 kV to a platinum wire inserted into the nanoESI tip, which was produced from a borosilicate glass capillary (1.0 mm o.d., 0.68 mm i.d.) pulled to ~5 μm o.d. using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). The source temperature and gas flow rates were 60 °C and 2

mL min⁻¹, respectively. The cone, trap and transfer voltages were 30 V, 5 V and 2 V, respectively. MassLynx software (version 4.1) was used for data acquisition and processing. The values of K_a for binding of a monovalent protein (P) to monovalent ligand (L) were calculated from the abundance ratio of L-bound to free P ions (i.e., *R*), measured by ESI-MS:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \quad (1)$$

$$K_a = \frac{[PL]}{[P][L]} = \frac{R}{[L]_0 - \frac{R}{R+1}[P]_0} \quad (2)$$

where [P]₀ and [L]₀ are initial concentrations of protein and ligand, correspondingly.²³⁴

Non-specific protein-ligand binding, if present, was corrected as shown in (2).²³⁵

Two batches of CS-35 F_{ab} generated on different days were used for the binding measurements of all ligands. All ligands (**1-13**) were tested against the first batch, while ligands **3, 6, 9, 12, 13** were tested additionally using the second batch of CS-35 F_{ab}. Previously reported K_a for CS-35 F_{ab} to Ara₆-OCH₃ (ligand **13**, Table 2) was 1.6 × 10⁵ M⁻¹. However, the K_D we obtained drifted from the literature value, possibly due to prolonged storage of the first batch of CS-35 F_{ab}. Mass spectra for aqueous solutions of the first batch of CS-35 F_{ab} were provided in Appendix. The four distinct F_{ab} species, a, b, c, and d, resulted from digestion of the IgG mAb with papain, were detected at charge states ranging +13 to +15. The masses are: (a) 48.25 kDa, (b) 47.90 kDa, (c) 48.92 kDa, and (d) and 49.20 kDa. The major glycoform a was used to calculate binding constants for all ligands. The second batch of CS-35 F_{ab} yielded a different major species at mass of 47.90 kDa. This signal was used to calculate binding constants for ligands **3, 6, 9, 12, 13**.

The binding constants given in Table 2 for ligands **3**, **6**, **9**, **12**, **13** were calculated by averaging the binding constants obtained from two batches of CS-35 F_{ab}.

For 906.4321 F_{ab}, three distinct F_{ab} species, a, b, c, were detected at charge states ranging +13 to +15. The masses are (a) 48.6, (b) 49.7, and (c) 50.0 kDa. The major glycoform a was used for calculating binding constants.

2.4.14 Synthesis of glycopeptide-squaramide

All syntheses of glycopeptide-squaramide followed the same procedure. Example of synthesis of Ara₆-ANSSFAPK-squaramide: Ara₆-ANSSFAPK (0.1 mg, 0.08 μmol, 1 eq.) was dissolved in 40 μL water in a 0.6-mL microcentrifuge tube. Diethyl squarate (0.02 mg, 0.16 μmol, 2 eq.) dissolved in an equal volume of EtOH was added. Saturated NaHCO₃ was added to adjust the pH of the reaction mixture to 8 as measured with universal pH-paper. Reaction was set at RT for 1 h, followed by purification on analytical RP-HPLC system using gradients provided in Appendix for each glycopeptide-squaramide. Yield: 67% after lyophilisation (by weight).

2.4.15 Synthesis of glycopeptide-BSA conjugates

To produce high density conjugates, 15-25 eq. of glycopeptide-squaramide (50-100 μg) was reacted with 1 eq. of BSA (~200 μg) in 10 μL, 500 mM borate buffer, pH 9 for 3 days at RT. The reaction mixture was dialyzed against 2 L of water (3 buffer changes over 8 h) at 4 °C. For medium and low density conjugates, 8-12 eq. or 4-7 eq. of glycopeptide-squaramide was reacted with 1 eq. of BSA (~ 400-500 μg). Conjugates were lyophilized after dialysis, and the conjugation efficiency was determined by MALDI

(Appendix). The quantity of each conjugate was determined by SDS-PAGE gel (Appendix).

2.4.16 Preparation of microarrays

Antigen preparations were aliquoted, lyophilized and sent to Engineering Arts LLC (Phoenix, Arizona) for piezoelectric non-contact printing. Briefly, the antigen spotting solutions were prepared at 0.1 mg/ml in buffer (1:10 PBS+0.005% v/v Triton-X100). Each antigen was micro-arrayed in triplicate spots onto epoxysilane-coated glass slides (Schott, NEXTERION[®] Slide E) with 360 picoliters per spot producing 140-160 μm diameter spots. The spot-to-spot spacing was 0.5 mm (column pitch) and 0.4 mm (row pitch). On each Epoxy slide, 24 replicate arrays were spotted to provide 72 identical spots for each antigen and control. The resulting arrays were sealed in a slide box and stored at -20 °C until use.

2.4.17 Microarray testing

Arrays were stored at -20 °C and, prior to each experiment, the array was equilibrated to room temperature (~ 20 min). The slide was attached to ProPlate[®] Multi-Array Slide System, 24 chambers (Grace Bio-Labs, 246824). Two hundred microliters of washing buffer (0.05% (v/v) tween in PBS) were used to wash each chamber on the slide ($2 \times 200 \mu\text{L}$), followed by one rinse with PBS. Two hundred microliters of blocking buffer was added to the slide, which was then sealed with plastic cover and incubated at 37 °C for 1.5 h. Blocking buffer was then aspirated, and varying concentrations (30, 10, 3,

1, 0.3, 0.1, 0.03, 0.01 $\mu\text{g}/\text{mL}$) of primary monoclonal antibodies (CS-35, CS-40, 906.4321) pre-saturated with blocking buffer (30 min at 37 °C) were added to each chamber of the gasket arrays and incubated at 37 °C for 1 h. After rinsing with washing buffer ($10 \times 200 \mu\text{L}$), followed by two additional washes with PBS ($2 \times 200 \mu\text{L}$), secondary antibodies (goat anti-mouse IgG) pre-saturated with blocking buffer at 1:500 dilution were added and incubated at RT for 1 h. After rinsing with washing buffer ($10 \times 200 \mu\text{L}$), the slide was detached from the gasket and the array was immersed in 50 mL of washing buffer for 5 min. The array was then transferred to 50 mL of milliQ water and immersed for 5 min. Array was dried by spinning in a table-top centrifuge and stored at room temperature in a microarray slide storage box.

2.4.18 Array scanning and data extraction

Arrays were scanned with Molecular Devices Genepix 4000B (Molecular Biology Service Unit, University of Alberta) using 532 nm laser and the photomultiplier tube (PMT) voltage setting at 350 V. The location of each spot on the array was outlined by use of the automatic grid feature of the GenePix Pro 4.1 software and manually adjusted where necessary. The signal intensity from each spot was calculated after subtraction of the local background.

2.4.19 Dose-response curves of antigens binding to mAbs (CS-35, 906.4321, CS-40)

MatLab script *MakeFigure2_9.m* was used for fitting the Hill equation to dose-response curves of each antigen binding to each mAbs. Figure 2-9 E can be generated on demand using MatLab script *MakeFigure2_9.m* available in Appendix.

Chapter 3: Structure-activity relationship between truncated glycopeptides and anti-LAM antibodies associated with mycobacterial infections

3.1 Introduction

In this chapter, I systematically perturbed the structure of the Ara₆-peptides hybrid antigens discovered in Chapter 2 (Figure 3-1 A) in order to investigate the structure-activity relationship of the peptide sequence and arabinose residues in their ability to bind the antigen-binding site of the anti-LAM antibody CS-35 in both monovalent and multivalent binding assays (Figure 3-1 B-C).

Compared to protein-protein interactions, protein-glycan interactions cover substantially lower surface area on the protein, causing small number of direct contacts at the binding interface. These shallow binding sites result in low to moderate binding affinity of a given glycan-binding site to a complementary glycan-motif when studied at monovalent binding level.²³⁶⁻²³⁹ To overcome these problems, in many biological systems, protein-glycan interactions occur in a multivalent or polyvalent fashion.²⁴⁰ Advancing receptor-ligand interactions to polyvalent mode yields several functional advantages, such as: a) increased affinity due to decreased rate of dissociation of two polyvalent entities;²⁴¹ b) signal amplification in biological responses;²⁴² c) efficient communication among cells due to larger surface contact; and d) increased resistance to shear stress during cell-cell recognition processes under severe flow conditions.¹⁵⁷

The benefits of polymerization of ligands have inspired the translation of monovalent ligands to multivalent architecture to improve the design of potent ligands, inhibitors and antagonists of weak protein carbohydrate interactions. Examples of such

weak interactions are interaction of sialic acid-terminated glycans, ubiquitously present on mammalian cells, with various pathological and non-pathological receptors. Whitesides and coworkers developed a polymer displaying multiple copies of sialic acid that were shown to be 10^8 times more effective in blocking the adhesion of influenza viruses to erythrocyte than more monovalent derivatives of sialic acids.²⁴³ Kiessling and coworkers in a similar approach devised a multivalent inhibitor of L-selectin-mediated leukocyte rolling.²⁴⁴ Oligomerization of existing weak ligands has been used extensively in academia to produce potent inhibitors rapidly. Development of a monovalent inhibitor of the same efficacy and inhibitory potency in the case of selectin required additional 20 years.²⁴⁵⁻²⁴⁶ Other areas where multivalent ligands became a commercial success is glycoconjugate vaccines; several of such vaccines have been successful against pathogens in the clinic and several of those have been made commercially available.²⁴⁷

Academic or commercial portable diagnostics that rely on interactions between multivalent antibody and multivalent antigen are inherent beneficiaries of multivalent interactions. To develop new multivalent diagnostic devices that rely on the interactions between IgG or IgM antibodies in sera and the multivalent presentation of immobilized antigens on surface, it is important to understand the binding between putative antigens and antibodies both in mono- and multivalent fashion. More specifically, in the case of mycobacterial infections, CS-35/arabinan interaction serves as a simple model system for probing oligofuranoside-antibody recognition. Lowary and coworkers previously showed that among a panel of oligosaccharides from mycobacterial arabinan, a minimum of tetrasaccharide motif was required for detectable affinities with CS-35 F_{ab}, and pointed out that the residues at the reducing and non-reducing end are essential for tight binding

(Figure 3-1 A).²²⁴ Our goal is to expand this understanding to a putative diagnostic device that contains a multivalent surface display of a peptide-oligoarabinose surrogate antigen for CS-35 and other mycobacteria specific antibodies.

3.2 Hypothesis and Objective

Previously in Chapter 2, I employed GE-FBD to select glycopeptides that consist of hexasaccharide (Ara₆) and heptamer peptides. The selected glycopeptides exhibited enhanced affinity and selectivity in monovalent binding assays towards anti-LAM antibody CS-35 when compared to Ara₆. The closely related 906.4321 antibody did not exhibit such enhanced binding to Ara₆-pep. (Figure 3-1 A). However, in multivalent assays, only low-valency Ara₆-peptides (Ara₆-ANSSFAP, Ara₆-TTYVVNP, Ara₆-DAHATLR) demonstrated mild selectivity. It appears that the multivalent presentation of ligand at specific density abolishes the selectivity of the monovalent glycopeptide construct.

The proximal arabinose attached to the peptide and the arabinose at the non-reducing end of hexasaccharide were previously reported as two of the key residues critical for tight binding to CS-35. We hypothesized that by truncating one of these key residues while retaining the other residue along with the peptide fragment may reduce the monovalent binding affinity of the resulting ligand but yield an optimal truncated antigen with ideal specificity in a multivalent assay. Figure 3-2 B and D justify this hypothesis on a structural basis; truncation of Ara₆ to Ara could maximize the specificity by keeping only the portion of the ligand that provides the highest potential to distinguish between

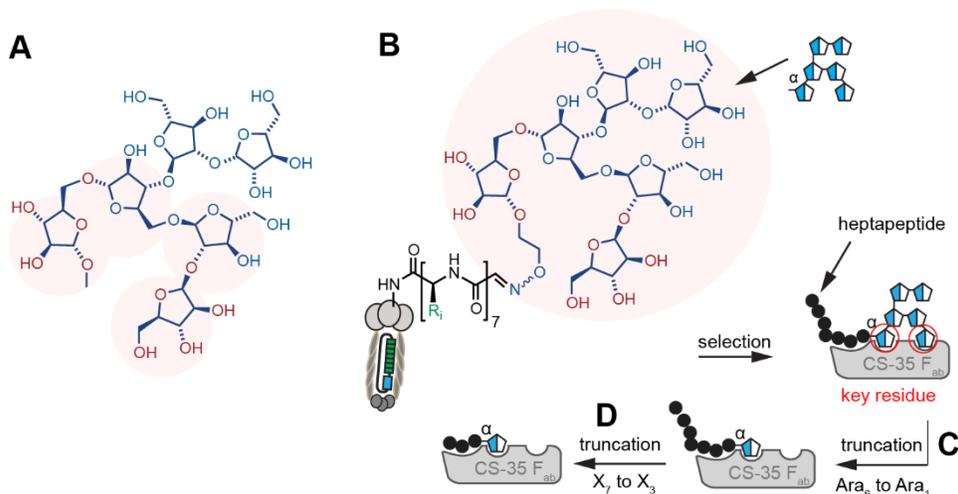


Figure 3-1. Truncation of hexasaccharide to enhance specificity.

(A) The tetrasaccharide motif required for detectable binding to CS-35. The residues at reducing and non-reducing end (red) are critical for tight binding. (B) Chapter 2 employed GE-FBD to identify binding ligands for CS-35 anti-LAM antibody. Ara₆-peptides bind specifically to CS-35 F_{ab} in monovalent fashion. (C) In this chapter, I truncated these discovered glycopeptides to a ligand that contains only a proximal Ara and (D) further truncated the heptapeptide sequence to a tripeptide.

two closely-related anti-LAM antibodies. We note that in the absence of X-ray structure of the complex, we do not understand the exact structural basis of glycopeptide:CS-35 interaction and this proposal is a mere extrapolation from the existing X-ray structure of Ara₆-CS-35 complex (Figure 2-1 B).

3.3 Experimental Design

I downsized the hexasaccharide moiety on the three discovered glycopeptides (Ara₆-ANSSFAP, Ara₆-DAHATLR, Ara₆-TTYVVNP) from Chapter 2 into α -arabinose moiety (Ara-ANSSFAP, Ara-DAHATLR, Ara-TTYVVNP). This proximal α -arabinose

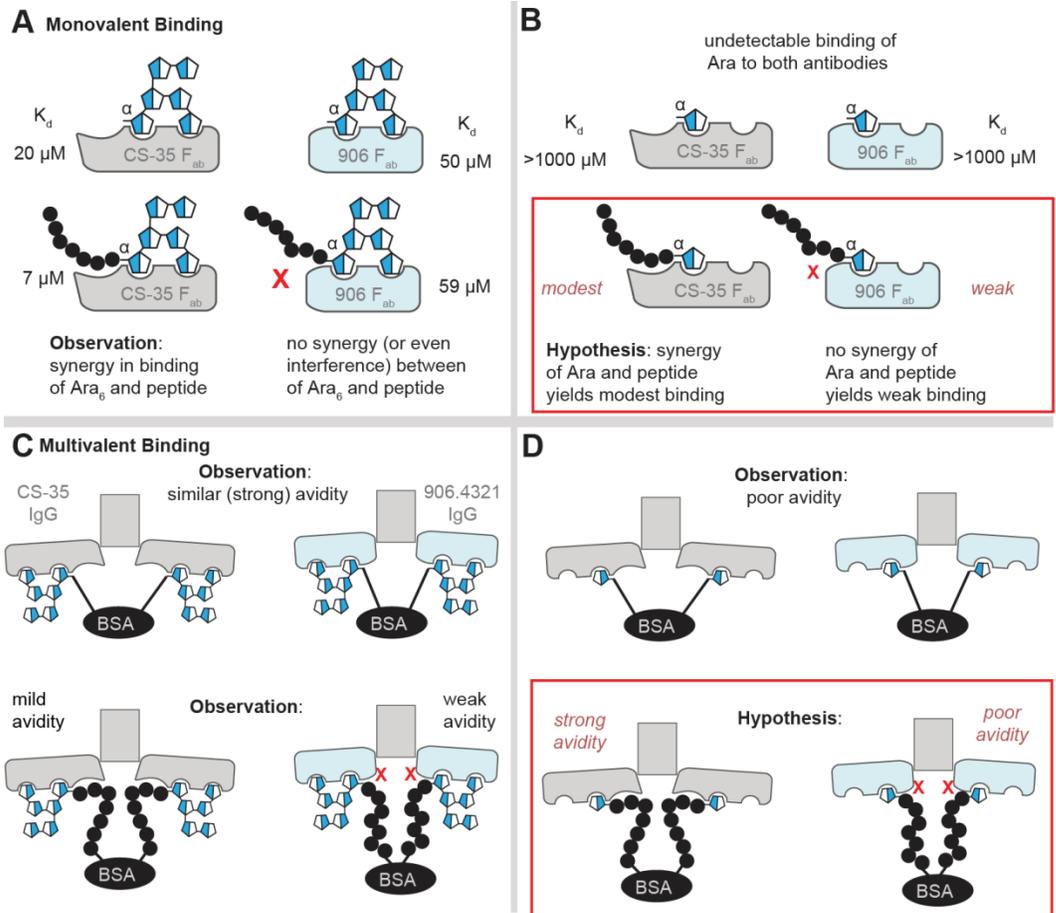


Figure 3-2. Observation of binding properties of Ara₆ and Ara₆-peptides to antibodies on monovalent (A) and multivalent assays (C). We hypothesize that by truncating the hexasaccharide to a monosaccharide (C), the specificity can be retained in multivalent assays (D).

moiety, according to the crystal structure of Ara₆ to CS-35 F_{ab} and monovalent binding assays, is one of the key residues that provides high specificity for Ara₆ (Figure 3-1 A).

Previous study using GE-FBD to select binders for ConA demonstrated that some residues in the discovered heptapeptide contributed minimally to the binding and truncating these residues indeed yielded similar K_D to untruncated sequences.⁵² To explore the impact of truncation on avidity and selectivity, I truncated one heptapeptide

(Ara-TTYVVNP) to tripeptide (Ara-TTY) and compared the binding properties of these two glycopeptides.

For multivalent binding assays in which ligands are immobilized on glass coated slides through amine reactive groups, I printed both glycopeptides (Ara-pep) and their corresponding BSA conjugates in order to examine the difference between direct printing through the amino group on terminal lysine residue in glycopeptides and printing of BSA conjugates through the amino groups on BSA molecules. The advantage of printing glycopeptides directly on slides is the simplification of the production of the potential diagnostic device. It bypasses the need for synthesis of BSA conjugates and their purification. On the other hand, the potential drawback of direct printing is poor control over spacing of the glycopeptides on the surface.

To perform a side-by-side comparison between ligands before truncation and after truncation in a multivalent fashion, I printed the second generation (Ara-pep) and the first generation (Ara₆-pep) glycopeptides BSA conjugates previously synthesized in Chapter 2 side-by-side. In addition, these conjugates were printed on both epoxysilane 2-D (slide E) and thin-film 3-D polymer (slide H) surfaces in order to explore the effect of binding interactions resulting from different type of immobilization strategies on these two slide surfaces.

3.4 Results and Discussion

3.4.1 Synthesis of glycopeptides

Peptides SANSSFAPK, SDAHATLRK, STTYK, SGGGGGGGK, STTYVVNPK, and SANSSFAP were synthesized using solid phase peptide synthesis (SPPS). Oxidation of purified, deprotected peptides by NaIO₄ and one-pot ligation with 1.5 eq. of Ara-OH₂ in the presence of aniline yielded desired glycoconjugates in 27-70% yield (Figure 3-3). These glycopeptides were used for ESI-MS.

3.4.2 ESI-MS of glycopeptides binding to monoclonal anti-LAM CS-35

With the help of Dr. Elena Kitova from Dr. John Klassen group, I employed ESI-MS to measure binding constants (K_D) of these six Ara-containing glycopeptides to CS-35 F_{ab} (Appendix) (Figure 3-4 B). When measurements were performed on different days, standard ligand Ara₆ was used to ensure measurements were reproducible and comparable.

The monovalent binding affinities of all Ara-peptides were significantly weaker than the affinities of Ara₆-peptides of the identical sequence. Both Ara-ANSSFAPK and Ara-ANSSFAP yielded weak to undetectable binding affinity, extrapolated to be in 3.3-3.7 mM range, suggesting that it was not the C-terminal Lys that interfered with binding (Figure 3-4 A). In contrast, the sequence SANSSFAP, when conjugated to Ara₆, yielded K_D 1.4 μ M.

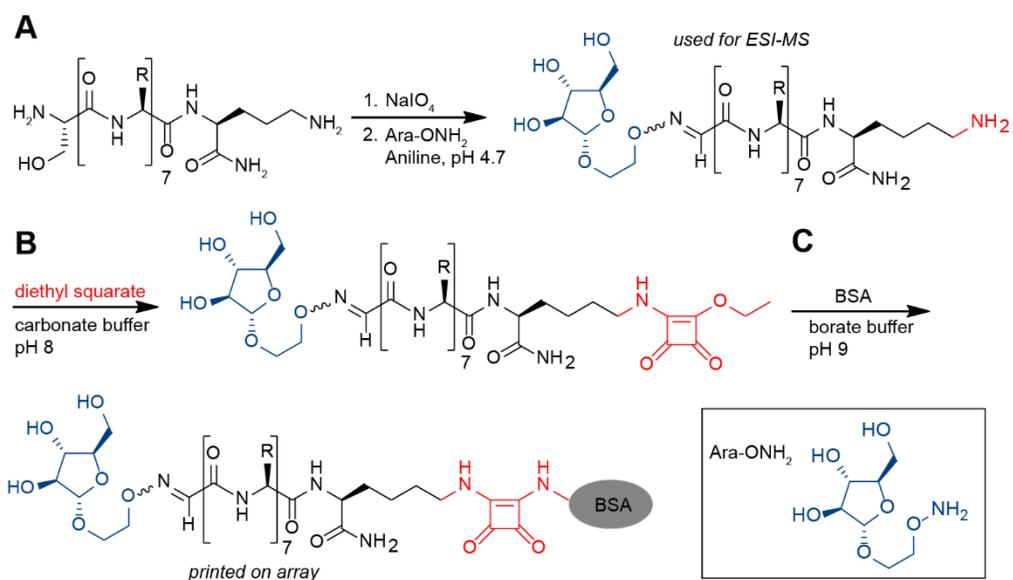


Figure 3-3. Synthesis scheme of Ara-peptides for ESI-MS measurements and microarray printing.

(A) Synthesis of Ara-containing glycopeptides. BSA conjugates of glycopeptides were synthesized through (B) suqaramide chemistry and (C) subsequent conjugation onto BSA.

This observation suggested that the loss of affinity of these second generation glycopeptides is significant and they might not be able to engage in mono or multivalent interaction with CS-35. The original hypothesis described in Figure 3-2 might not be accurate. Still, to examine the performance of these ligands in multivalent assays, I proceeded to synthesize and evaluate multivalent constructs.

3.4.3 Testing the binding of monoclonal antibodies to glycopeptides in a microarray format

All synthesized Ara-peptides were conjugated to BSA at high (>7) and low densities (<4) for printing onto arrays (Figure 3-3 C). Glycopeptides Ara-TTYVVNPK and Ara-GGGGGGK were also conjugated to BSA at medium density (5-6). The

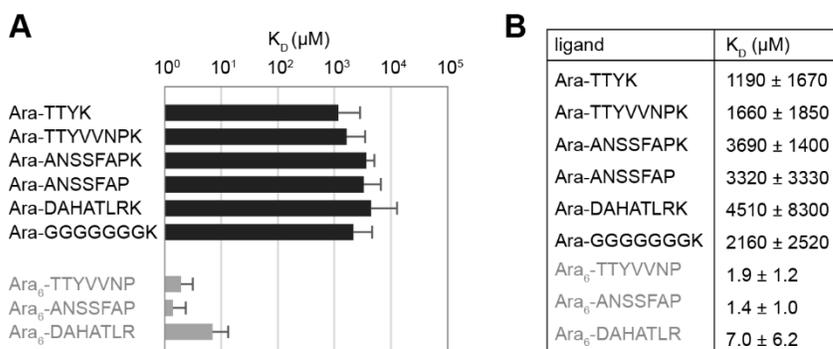


Figure 3-4. ESI-MS measurements of Ara-peptides to CS-35 F_{ab} . ES-MS measured K_D for peptides and Ara-peptides to CS-35 F_{ab} (black) compared to Ara₆-peptides discovered from Chapter 2 (gray) in (A) bar graph and in (B) table.

average number of glycopeptides conjugated per BSA molecule was determined by MALDI (Appendix). I screened arrays using the same procedure as the one described in Chapter 2. In brief, a 3-fold dilution series of anti-LAM IgGs—CS-35, CS-40, and 906.4321—were employed to yield dose-response curves for each printed ligand binding to each antibody. As expected, all resynthesized glucose-containing control ligands (Glu-ANSSFAP and Glu-TTYVVNP) did not show any binding signals at any surface density (data not shown). Unfortunately, none of the Ara-containing glycopeptides, whether conjugated to BSA or directly immobilized on slides, yielded detectable signals on any type of surfaces (Figure 3-5 A). These results show that the combination of monosaccharide and the linker to the peptide sequence used for investigation failed to elicit synergistic binding interactions with CS-35.

Because no Ara-peptides binding interactions with CS-35 were detected, an accurate comparison of performance of ligands printed on different slide surfaces was not possible. Nevertheless, since the first generation (Ara₆-pep) glycopeptides were also printed in parallel, we were able to evaluate a difference in the dose-response curves from

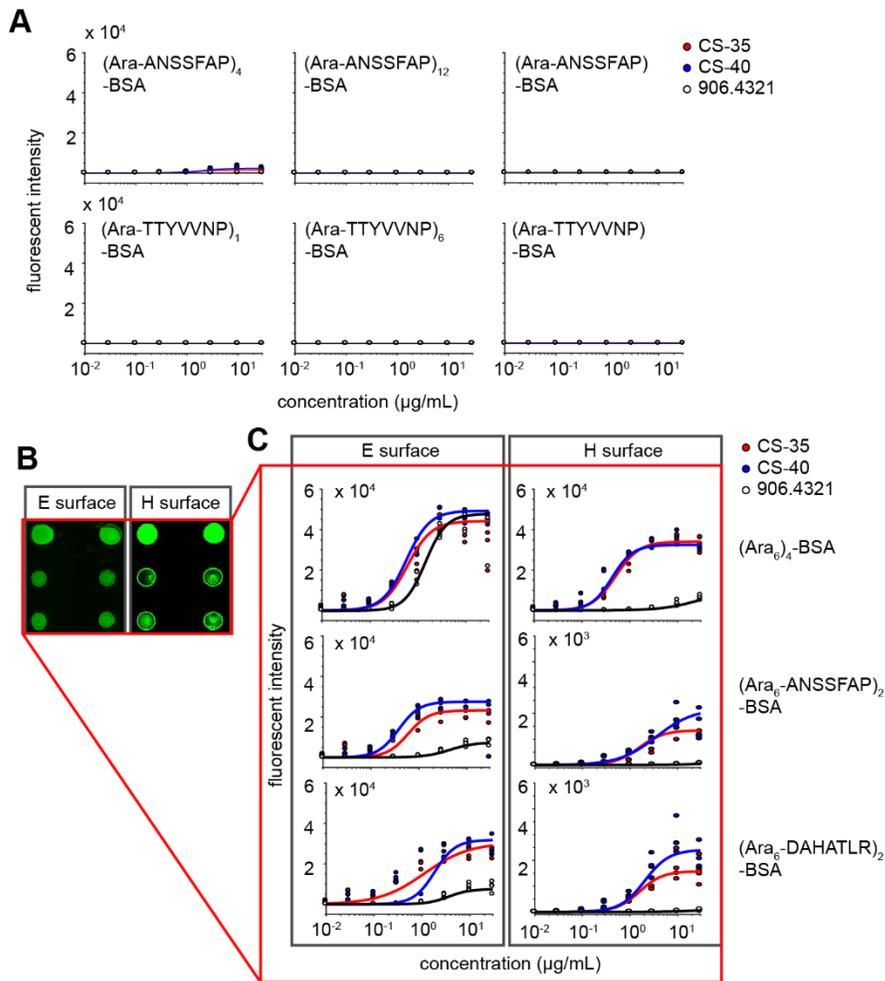


Figure 3-5. Microarray testing of Ara₆- and Ara-peptides on E and H surface. (A) Dose-response curves of Ara-peptides binding to CS-35, CS-40, and 906.4321. (B) Fluorescence images of (Ara₆)₄-BSA, (Ara₆-ANSSFAP)₂-BSA, and (Ara₆-DAHATLR)₂-BSA after binding to 10 $\mu\text{g/mL}$ CS-35 on E surface and H surface, and (C) the corresponding dose-response curves of these three Ara₆-peptides binding to CS-35, CS-40, and 906.4321 on E and H surfaces.

E surface (epoxy-coated glass) to H surface (glass surface presenting N-hydroxysuccinimide reactive groups). The signal intensity of all Ara₆-peptides on H surface at all valencies diminished when compared to E surface. This has caused the low valency ligands to yield even lower or undetectable signals on slide H (Figure 3-5 C). Moreover, we observed, from the fluorescence image, that the morphology of the spots

printed on slide H showed that all spots corresponding to Ara₆-peptides BSA conjugates were of donut shapes, possibly indicating a suboptimal immobilization of ligand on H surface (Figure 3-5 B). Examination of ligand selectivity towards CS-35 showed that all valencies of Ara₆ antigens exhibited high selectivity for CS-35 and CS-40 on slide H. Such selectivity was not observed on slide E (Figure 3-5 C). The selectivity of Ara₆-peptides, according to their individual EC₅₀ towards CS-35 and 906.4321, remained similar from Slide E to slide H. It is not clear whether the discrepancy in selectivities was caused by incomplete immobilization of conjugates or by additional effects present only on polymer film of slide H, which resulted in lower binding affinities of these glycopeptide BSA conjugates.

3.5 Conclusion

To enhance the specificity of discovered Ara₆-peptides from Chapter 2, I proposed to perturb the affinity of the hexasaccharide moiety of these glycopeptides by a truncation that yields monosaccharide-containing glycopeptides. The results showed that these Ara-peptides exhibited a dramatically reduced monovalent binding affinities to CS-35. Consequently, they yielded low to undetectable binding signals in multivalent assays.

I concluded that using this truncation approach for antigen discovery is not a promising path. The sequences that were discovered in synergy with Ara₆ cannot bind in synergy with Ara to CS-35. As a result of this loss of affinity, they cannot be used to distinguish CS-35 from 906.4321. Therefore, to identify Ara-peptides that display high specificity and affinity to CS-35, one has to search for ligands *de novo* using selection of

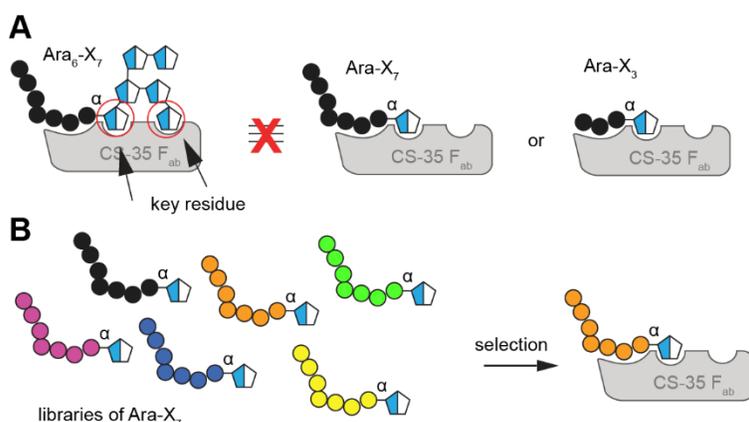


Figure 3-6. *De novo* discovery of Ara-X₇ binders for CS-35.

(A) Truncation of the hexasaccharide to monosaccharide on sequences that were identified in synergy with hexasaccharide did not yield more specific and stronger binders for CS-35 F_{ab}. (B) To discover Ara-peptide ligands for CS-35 F_{ab}, one should search ligands in Ara₁-peptide libraries.

glycopeptide libraries that contain Ara-monosaccharide moiety against the target CS-35 (Figure 3-6). The conclusion leads to the next chapter. In Chapter 4, I will discuss the development and validation of a selection strategy that facilitates the discovery of ligands *de novo*.

3.6 Experimental procedures

3.6.1 Materials and general information

Peptides SANSSFAPK, STTYVVNPK, SDAHATLRK, SGGGGGGGK, and STTYK were synthesized according to standard protocol described in Chapter 2 (section 2.4.9). Production of mAb IgG of CS-35, CS-40, and 906.4321 and F_{ab} of CS-35 was previously reported (ref). The synthesis of Ara-ONH₂ was performed by R. Brunton in

Lowary group using previously developed strategy for synthesis of α -anomeric conjugates of arabinose.

RP-HPLC were performed on Waters HPLC system equipped with a Waters 1525 EF binary pump a Waters FlexInject manual injector (dual mode) and a Waters 2489 tunable UV detector. SymmetryPrepTM C18 semi-preparative column (19 × 50 mm, particle size 5 μ m, pore size 100 Å) and XBridge BEH Amide OBD Prep column (19 × 250 mm, particle size 5 μ m, pore size 130 Å) were used for purifications of peptides and arabinose glycopeptides at a typical flow rate of 12 mL/min. Characterization of Ara-peptide and Ara-peptide squaramides were performed by UPLC-MS. Characterization of Ara-peptide-BSA conjugates were performed by MALDI. Microarrays were printed at Engineering Arts LLC (Phoenix, Arizona) using Schott NEXTERION® slide E, slide H, and piezoelectric non-contact printing. Array scanning and data extraction were described in Chapter 2 section 2.4.17 and 2.4.18.

3.6.2 Syntheses of glycopeptide conjugates

Synthesis of all Ara-containing glycopeptides follow previously published protocol.⁵² An example of synthetic procedure is provided for Ara-GGGGGGGK: the peptide SGGGGGGK (9 mg, 14 μ mol, 1 eq.) was dissolved in 200 mM MOPS (0.3 mL, pH 7.0). The solution was added to a 1.5-mL microcentrifuge tube containing sodium periodate in water (34 μ L, 500 mM, 1.2 eq.). The reaction mixture was incubated for 1 h at RT. Ara-ONH₂ (4.3 mg, 21 μ mol, 1.5 eq.) dissolved in 500 mM anilinium acetate (0.13 mL, pH 4.7) was added to the reaction mixture. The oxime ligation was carried out

for 1 h at RT. The reaction mixture was then injected into a RP-HPLC system using gradients described for each glycopeptide in Appendix. Lyophilization yielded Ara-GGGGGGGK as white fluffy powder (72%). The purity and identity of the product was confirmed with by UPLC-MS (see Appendix for spectra).

3.6.3 ESI-MS binding measurements

The association constants (K_a) for F_{ab} binding to ligands were determined using the direct ESI-MS binding assay. The sample solutions were prepared from aqueous stock solutions of F_{ab} and ligand with known concentrations. Aqueous ammonium acetate was added to the sample solutions to yield a final buffer concentration of ~ 200 mM. Direct ESI-MS binding assay was performed on a Synapt G2S Q-IMS-TOF mass spectrometer (Waters, Manchester, UK) equipped with nanoflow ESI (nanoESI) source. NanoESI was performed by applying a voltage of ~ 1 kV to a platinum wire inserted into the nanoESI tip, which was produced from a borosilicate glass capillary (1.0 mm o.d., 0.68 mm i.d.) pulled to ~ 5 μm o.d. using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). The source temperature and gas flow rates were 60 $^{\circ}\text{C}$ and 2 mL min^{-1} , respectively. The cone, trap and transfer voltages were 30 V, 5 V and 2 V, respectively. MassLynx software (version 4.1) was used for data acquisition and processing. The values of K_a for binding of a monovalent protein (P) to monovalent ligand (L) were calculated from the abundance ratio of L-bound to free P ions (i.e., R), measured by ESI-MS:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \quad (1)$$

$$K_a = \frac{[PL]}{[P][L]} = \frac{R}{[L]_0 - \frac{R}{R+1}[P]_0} \quad (2)$$

where $[P]_0$ and $[L]_0$ are initial concentrations of protein and ligand, correspondingly.²³⁴

Non-specific protein-ligand binding, if present, was corrected as shown in (2).²³⁵

Mass spectra for aqueous solutions of CS-35 F_{ab} were provided in Appendix. The four distinct F_{ab} species, a, b, c, and d, resulted from digestion of the IgG mAb with papain. All four species were detected at charge states ranging +13 to +15. The masses were: (a) 48.25 kDa, (b) 47.90 kDa, (c) 48.92 kDa, and (d) and 49.20 kDa. Because for some ligands, the mass of the first two antibody glycoforms overlapped with their corresponding ligand-antibody complexes, the glycoforms used for calculation of binding constants for each ligand were indicated in the spectra in Appendix.

3.6.4 Synthesis of glycopeptide-squaramide

All syntheses of glycopeptide-squaramide followed the same procedure. Example of synthesis of Ara-ANSSFAPK-squaramide: Ara₆-ANSSFAPK (3 mg, 2.8 μmol, 1 eq.) was dissolved in 50 μL water in a 0.6-mL microcentrifuge tube. Diethyl squarate (1 mg, 5.6 μmol, 2 eq.) dissolved in an equal volume of EtOH was added. Saturated NaHCO₃ was added to adjust the pH of the reaction mixture to 8 as measured with universal pH-paper. Reaction was set at RT for 1 h, followed by purification on semi-preparative RP-HPLC system using gradients provided in Appendix for each glycopeptide-squaramide. Yield: 42% after lyophilisation.

3.6.5 Synthesis of glycopeptide-BSA conjugates

To produce high density conjugates, 15-25 eq. of glycopeptide-squaramide (~1 mg) was reacted with 1 eq. of BSA (~ 3.5 mg) in 60 μ L, 500 mM borate buffer, pH 9 for 2 days at RT. The reaction mixture was dialyzed against 2 L of water (4 buffer changes over 24 h) at RT. For medium and low density conjugates, 8-12 eq. or 4-7 eq. of glycopeptide-squaramide was reacted with 1 eq. of BSA (~ 4 mg). Conjugates were lyophilized after dialysis, and the conjugation efficiency was determined by MALDI (Appendix).

3.6.6 Preparation of microarrays

Antigen preparations were aliquoted, lyophilized and sent to Engineering Arts LLC (Phoenix, Arizona) for piezoelectric non-contact printing. Briefly, the antigen spotting solutions were prepared at 0.1 mg/mL in buffer (1:10 PBS+0.005% v/v Triton-X100). Each antigen was printed with two duplicate spots per array onto epoxysilane-coated glass slides (Schott, NEXTERION® Slide E) or N-hydroxysuccinimide-activated glass slides (Schott, NEXTERION® Slide H) with 360 picoliters per spot producing 140-160 μ m diameter spots. Within each array, center to center spot-to-spot spacing was 0.5 mm (column pitch) and 0.32 mm (row pitch). On each slide, 24 replicate arrays were spotted to provide 48 identical spots for each antigen and control. The resulting arrays were sealed in a slide box and stored at -20 °C until use.

3.6.7 Microarray testing

Arrays were stored at -20 °C and, prior to each experiment, each array was equilibrated to room temperature for ~ 20 min. The slide was attached to ProPlate® Multi-Array Slide System, 24 chambers (Grace Bio-Labs, 246824). Two hundred microliters of washing buffer (0.05% (v/v) tween-20 in PBS) were used to wash each chamber on the slide (2 × 200 µL), followed by one rinse with PBS. Two hundred microliters of blocking buffer was added to the slide, which was then sealed with plastic cover and incubated at 37 °C for 1.5 h. Blocking buffer was then aspirated, and varying concentrations (30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 µg/mL) of primary monoclonal antibodies (CS-35, CS-40, 906.4321) pre-saturated with blocking buffer (30 min at 37 °C) were added to each chamber of the gasket arrays and incubated at 37 °C for 1 h. After rinsing with washing buffer (10 × 200 µL), followed by two additional washes with PBS (2 × 200 µL), secondary antibodies (goat anti-mouse IgG) pre-saturated with blocking buffer at 1:300 dilution were added and incubated at RT for 1 h. After rinsing with washing buffer (10 × 200 µL), the slide was detached from the gasket and the array was immersed in 50 mL of washing buffer for 5 min. The array was then transferred to 50 mL of milliQ water and immersed for 5 min. Array was dried by spinning in a table-top centrifuge and stored at room temperature in a microarray slide storage box.

3.6.8 Array scanning and data extraction

Arrays were scanned with Molecular Devices Genepix 4000B (Molecular Biology Service Unit, University of Alberta) using 532 nm laser and the photomultiplier tube

(PMT) voltage setting at 350 V. The location of each spot on the array was outlined by use of the automatic grid feature of the GenePix Pro 4.1 software and manually adjusted where necessary. The signal intensity from each spot was calculated after subtraction of the local background.

3.6.9 Dose-response curves of Antigens binding to mAbs (CS-35, 906.4321, CS-40)

MatLab script *MakeFigure3_5.m* was used for fitting the Hill equation to dose-response curves of each antigen binding to each mAbs. Figure 3-5 B can be generated on demand using MatLab script *MakeFigure3_5.m* available in Appendix.

Chapter 4: Deep-panning of phage-displayed macrocyclic peptide libraries for antibodies associated with mycobacterial infections

4.1 Introduction

In this chapter, I investigated the synthesis of macrocyclic peptide libraries displayed on phage through two sequential orthogonal chemical modifications of phage-displayed peptides (Figure 4-1A). These libraries serve as potential source for the discovery of antigens against antibodies present in sera of mycobacteria-infected patients. To enable identification of arabinose modified macrocyclic peptides antigens for anti-LAM antibody CS-35, I developed a deep-panning platform that employs one round of selection and Illumina sequencing (Figure 4-1B). I validated this discovery platform on model targets and I believe this technology can be easily transferred to others systems which employ other type of targets or input libraries with different modifications.

The overarching proposal described in this section, i.e. the usage of random libraries to identify peptides or glycopeptides that bind to an antigen-binding site of the antibody, can be justified based on the modern understanding of adaptive immune system. The canonical interpretation of adaptive immunity that triggers antibody responses is that the immune response initiates when antigen-specific B cells that bind to antigens are signaled by helper T cells. When an epitope on a disease-associated protein is recognized by the surface immunoglobulin on a B cell, the protein is internalized and degraded. The peptides derived from the degraded protein are returned to the B-cell surface and bound to MHC class II molecules. Helper T cells that recognize the peptide:MHC complex then

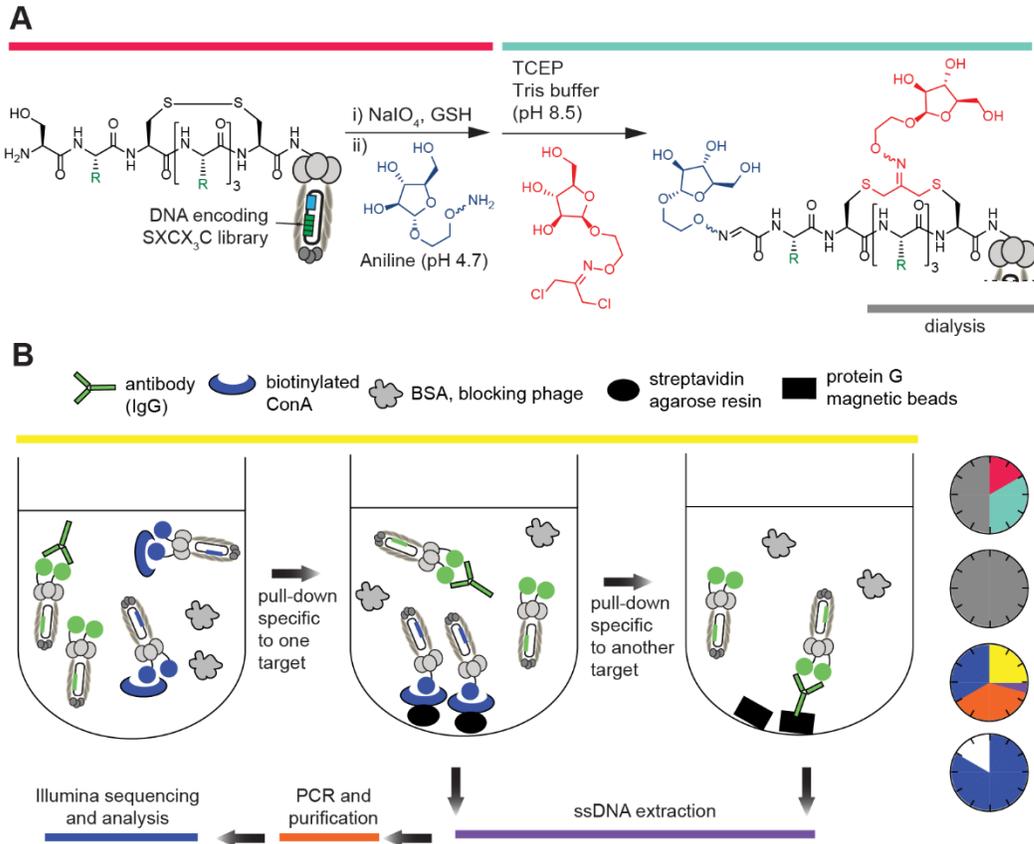


Figure 4-1. Overview of the workflow of deep-panning system.

The entire ligand search can be accomplished within 2 days. (A) Modification of phage library and subsequent purification by dialysis require an additional day. (B-D) Sequential pull-down assays to enrich the ligands that bind to targets being precipitated at each step (3 h). (C) Pull down of biotinylated ConA by streptavidin agarose resin. (D) Pull down of antibody by protein G magnetic beads. (E) Isolation of phage ssDNA (purple, 30 min); PCR of library DNA with primers, E-Gel purification for Illumina sequencing (4.5 h); DNA sequencing with Illumina platform and analysis (14 h).

deliver activating signals to the B cell, which then produce antibody against the peptide antigen in downstream processes.²⁴⁸ MHC class II molecules typically bind to peptides that are 14-20 amino acids long. The peptide segment that elicits close interaction and situates within the groove of the MHC class II molecules consists of 9 amino acids, whereas the remaining peptide residues are outside the groove.²⁴⁹ Due to the unique

interaction and presentation to its associated antibody, each peptide segment can serve as useful disease epitope that corresponds to a specific disease state.

One of the ultimate goals of the work described in this chapter is the understanding of antibody-ligand interactions in the context of intact serum from TB(+) patients in order to design highly specific ligands that bind to TB-associated antibodies present in this serum. Unfortunately, many TB-specific antigens are complex carbohydrates, whose syntheses are usually complicated; moreover, the generation of TB associated antibodies that can recognize carbohydrate epitopes is not well understood. Research efforts in the past 20-30 years suggest that it is possible to identify peptides that bind to antibodies and at the same time recognize carbohydrates; however, these peptides have historically failed to trigger immune response necessary for the production of glycan-binding antibodies. Nevertheless, the mining of peptide or glycopeptide libraries to find a potent and specific binder for carbohydrate-binding antibodies associated with TB can be a productive avenue for development of ligands that bind to antibodies *in vitro*, and this approach can also serve as the basis for specific and sensitive serological tests.

Traditionally, technologies such as peptide arrays and chemically-synthesized peptide libraries were employed to discover disease epitopes.²⁵⁰⁻²⁵² However, these approaches were inherently limited because these synthesized libraries could not cover the diversity of all possible peptides at lengths pertinent to antibody-peptide antigen interactions (e.g. MHC class II molecules bind >9 amino acids, so the number of all possible 9-mers is $20^9 = 5 \times 10^{11}$).

In the 1990's, development of display technology for the production and screening of large peptide libraries displayed on phage made it possible to rapidly

identify a peptide antigen for virtually any purified antibody.^{31, 253-254} For example, through the use of random phage-displayed peptide libraries, Scott²⁵⁵⁻²⁵⁶ and others²⁵⁷⁻²⁵⁸ were able to show that carbohydrate-binding antibodies could be used as bait in selection process and could identify peptides that bind to carbohydrate-binding sites of these antibodies. Another attractive proposal that emerged at the time was the discovery of disease-specific antigens via screening of random libraries against antibody repertoires isolated from disease-bearing patients.^{169, 259} Yet, such approach has only been shown to be successful in a few cases. The existing unsolved challenge can be due to several reasons: the heterogeneity of the serum antibodies,²⁶⁰ the disproportionate display of ligands in genetically-encoded libraries,¹⁶⁶ the amplification bias involved in selection procedures,¹⁶⁷ or mismatched chemical space of genetically-encoded peptide libraries.¹⁶⁵

In the mid 2000's, deep sequencing technology has emerged as a promising tool to accelerate ligand screening through phage display, mRNA display, yeast display, and other related display platforms.^{186, 261-263} Sequencing $>10^5$ clones simultaneously allows ligands to be efficiently identified in a single round of selection. The use of deep sequencing in selection also decreases the number of rounds of iterative binding and amplification, hence reducing secondary biases in selections.^{5, 11-12}

The past 10 years has also brought about several solutions to the limited chemical diversity in display technologies. As most display systems rely on the natural transcription and translation machinery, available ligands are exclusively within the space of peptides made of the 20 natural amino acids. Selections that use these random peptide libraries as the sole source of diversity and serum antibody repertoire as bait may not be able to identify disease epitopes that can be effectively and specifically recognized by

antibodies associated with non-protein antigens from pathogens, because every antibody in the serum antibody repertoire has the potential to recognize a peptide epitope. The abundance of peptide-binding antibodies in any serum may divert the selection and make it difficult to target rare antibodies that originate from disease or pathogens. Note that many antibodies to bacterial pathogens can recognize non-proteins species, such as glycans, lipids, complex glycopeptides and glycolipids. In addition, many diseases are also characterized by post translational modifications (PTM) in proteins, such as citrullination of glucose-regulated protein in type 1 diabetes²⁶⁴ or rheumatoid arthritis,²⁶⁵ phosphorylation of kinases in chronic leukemia,²⁶⁶ and glycosylation in cancer cells.²⁶⁷

Due to the importance of PTMs, considerable progress has been developed to produce genetically-encoded libraries that mimic fragments of modified proteins.²⁶⁸ For example, Schultz and coworkers showcased a peptide library displayed on phage that can be phosphorylated in vitro and used to target protein tyrosine kinases.²⁶⁹ Derda and coworkers⁵² and Mihara and coworkers²⁷⁰ have developed glycopeptide libraries displayed on phage to target carbohydrate-binding proteins. Analogous modification of mRNA-displayed libraries have been pioneered by Roberts,²⁷¹ and developed in groups of Krauss,²⁷² Ito²⁷³ and others. While there exist studies towards expanding chemical diversity for ligand search through post translational modifications, introduction of two distinct chemical modifications in phage-displayed library of peptides has not been reported. Such dual modification is one of the advancements described in this chapter.

Early efforts in the production of libraries with unnatural modifications focus on incorporating tRNA/aminoacyl tRNA synthetase pair in the system of interest. To date, over 100 non-canonical amino acids have been evolved and demonstrated with novel

chemical or biological properties in several biological systems.²⁷⁴⁻²⁸⁰ Unfortunately, only a few reports have used unnatural amino acids for the selection of phage libraries.²⁸¹ Moreover, Schultz, Smider, and coworkers reported that in such system, the yield of library generation favors sequences containing natural amino acids.²⁸¹ As a result, incorporating multiple unnatural amino acids to generate multiple modifications may pose systematic expression biases against sequences containing unnatural amino acids. Suga and coworkers developed technology that employs template-directed ribosomal synthesis of peptides in order to incorporate multiple standard or non-standard amino acids,²⁸² but this technology is rather expensive, and according the inventor himself, is “very difficult to use, and cannot be easily adapted by anyone but his group” (Suga, personal communications). We and others (reviewed by Heinis and Winter²⁸³) recognized that careful design of chemical reactions on natural amino acids can be used to introduce different and multiple chemical modifications simultaneously, in this case, a macrocycles-forming reaction, without the need to use unnatural codons or alter the expression of peptide sequences in any way. I took such approach in this Chapter, and strived to simplify this approach to make it a robust, reproducible modification that can be transferrable to any person with minimal knowledge of chemistry or bioconjugation.

4.2 Objective and Hypothesis

Previously, Chapter 3 suggests that the discovery of ligands with high selectivity towards CS-35 anti-LAM antibody cannot be achieved through truncating the carbohydrate moiety nor truncating the peptide sequence from an existing, identified Ara₆-peptide sequence (discovered in Chapter 2). Rather, ligands should be discovered *de*

novo from libraries that contain the desired, truncated carbohydrate moiety (Ara). As a result, I focused my chemistry efforts on incorporation of one or more Ara-monosaccharide blocks into peptide libraries.

Partial substitution of glycan with simple structural blocks has been reported successful for several carbohydrate-binding proteins.²⁸⁴⁻²⁸⁵ It is possible that the Ara₆ is not the exact core epitope that can be recognized by TB-specific antibodies. Instead, other epitopes that exhibit general features of the Ara₆ epitope—such as simultaneous binding of two Ara residues to the F_v of the antibody—may also be recognized by TB-specific antibodies. I hypothesized that complex glycan can be replaced by a structural core made of non-carbohydrate entity (here peptide) and two monosaccharides that are in contact with the target, forming an overall macrocyclic structure (Ara-M-Ara) (Figure 4-2A).²⁸ Discovery of Ara-peptide antigens has an additional benefit when compared to antigens based on complex carbohydrates (i.e. Ara₆ or Ara₆-peptide), as synthesis of monosaccharides (Ara) are trivial when compared to synthesis of hexasaccharide. Simplicity of synthesis offers advantage when material is needed to be produced at scale for validation purposes.

The chemistry that allows preparation of Ara-M-Ara libraries has been published by the Dawson²⁸⁶ and Derda group.¹⁷⁸ It is conceivable that the attachment of the first α -arabinose residue can generate an α Ara-XCX₃C library through the oxidation of N-terminal Ser and aminoxy-Ara ligation. The disulfides are known to be resistant to NaIO₄ oxidation in controlled conditions. The macrocycles that carry a second arabinose residue thus can be introduced subsequently via stapling of two Cys residues with dichlorooxime arabinose derivative. The ligation strategies leading to such Ara-peptide

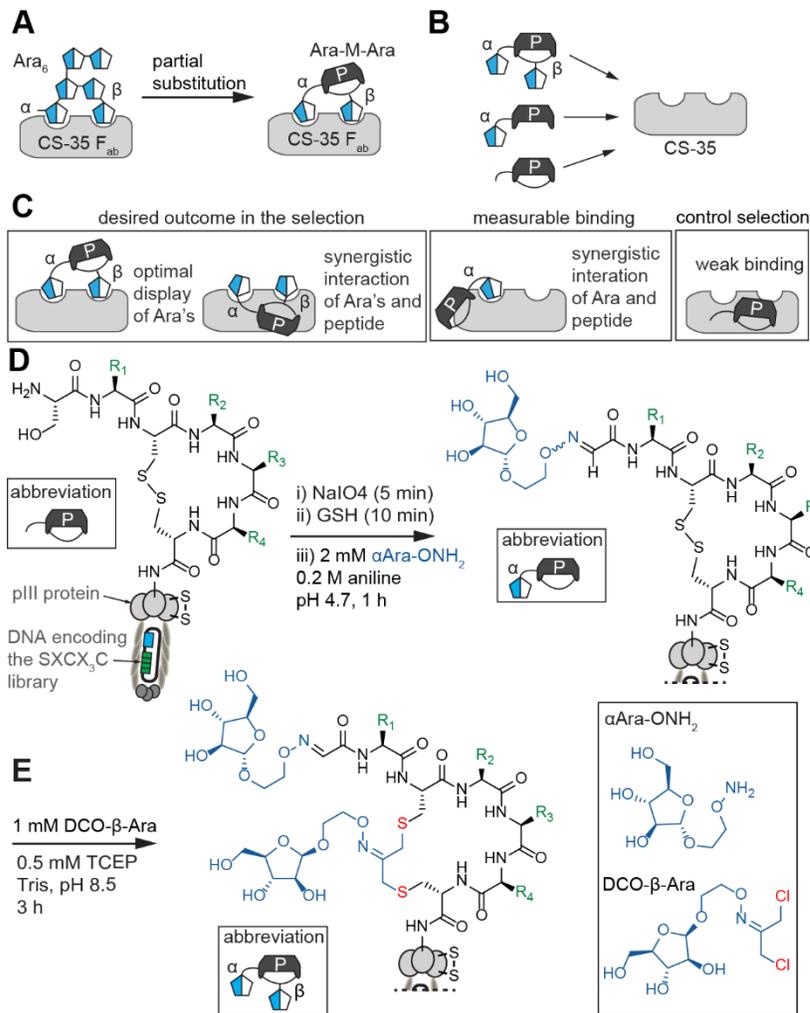


Figure 4-2. Macrocyclic formation by bis-glycosylated peptide macrocycle. (A) Partial substitution of hexasaccharide with bis-glycosylated peptide macrocycle. The α - and β -arabinose residues contribute to the specificity of Ara₆ towards CS-35 F_{ab}. (B)(C) Screening against anti-LAM antibody to yield hybrid Ara-M-Ara epitopes that exhibit similar performance to Ara₆-peptide epitopes. (D)(E) Synthesis of Ara-macrocytic libraries (Ara-M-Ara) through two sequential orthogonal reactions.

libraries are potentially compatible with modification of peptides displayed on phage (Figure 4-2 D-E). The use of such libraries offers advantage over random peptide libraries because selection that uses a peptide library decorated with two strategically positioned Ara epitopes is expected to favor targeting arabinose-interacting antibodies,

which are often not naturally occurring in human serum unless the source of serum has been exposed to exogenous microbial cell walls and plants.

4.3 Experimental Design

Simple glycopeptides antigens can offer advantages over antigens based on complex carbohydrates due to the ease of peptide synthesis when compared to the synthesis of complex glycan. To find such antigens, I chemically attached an α -arabinose on SXCX₃C library to yield α Ara-XCX₃C and Ara-M-Ara libraries. To mimic the interactions between key arabinose residues and CS-35, I optimized the conditions that allow two bridging Ara residues with a semi-rigid linker, forming macrocyclic peptides. Each Ara-macrocycle (Ara-M-Ara) is conformationally restricted, but the overall libraries of Ara-M-Ara will allow the exploration of conformational space in the carbohydrate-binding site of CS-35.

To develop efficient single round deep-panning platform, one must first validate screening procedure using model targets and model libraries. Therefore, I designed a simple in-solution panning process, termed “sequential pull-down assay,” that employs two consecutive pull-down procedures. In this assay, affinity resins were used to capture the target protein, along with the population of phage that were bound to this target. Phage clones precipitated in each “pull-down” step can be further processed to isolate their DNA, which in turn can be analyzed through deep sequencing, generating information about peptide sequences or glycopeptides that were co-precipitated with the associated target.

To validate this pull-down assay, I used antiFLAG (aFLAG) antibody and biotinylated ConA as model targets and incorporated model libraries for the selection procedure. More specifically, I used two different libraries (lib 1 and lib 2) as the inputs for pull-down assays. The composition of these two libraries are: i) lib 1—Ara-modified cyclic peptide libraries (Ara-XCX₃C), linear mannosylated peptide libraries (Man-X₇), monoclonal phage that displays FLAG epitope, and a mannosylated monoclonal phage (Man-WYDKYH). ii) lib 2—unmodified linear peptide libraries (SX₇), unmodified cyclic peptide libraries (SXCX₃C), and monoclonal phage that displays FLAG epitope. In lib 1, Man-WYDKYH served as the internal control clone which was previously identified to show an affinity for ConA;⁵² on the other hand, the monoclonal phage that displayed FLAG epitope served as internal control for the model target aFLAG. In other words, if the selection was successful, the pull-down of biotinylated ConA using streptavidin-coated resins should enrich the population of phage that were mannosylated and contained the motif WYD; at the same time, the pull-down of aFLAG using protein G-coated resins should enrich phage clones with FLAG epitopes.

4.4 Results and Discussion

4.4.1 Generation of Ara-M-Ara libraries on phage

LC-MS analysis confirmed the formation of Ara-M-Ara conjugate in a two-step reaction. Treatment of disulfide bridged SHCVWWDC with NaIO₄, followed by the addition of 8.8 mM α Ara-ONH₂ (1.5 eq.) in 200 mM aniline, converted the peptide to glycopeptide containing N-terminal arabinose, as confirmed by LC-MS (Figure 4-3 B).

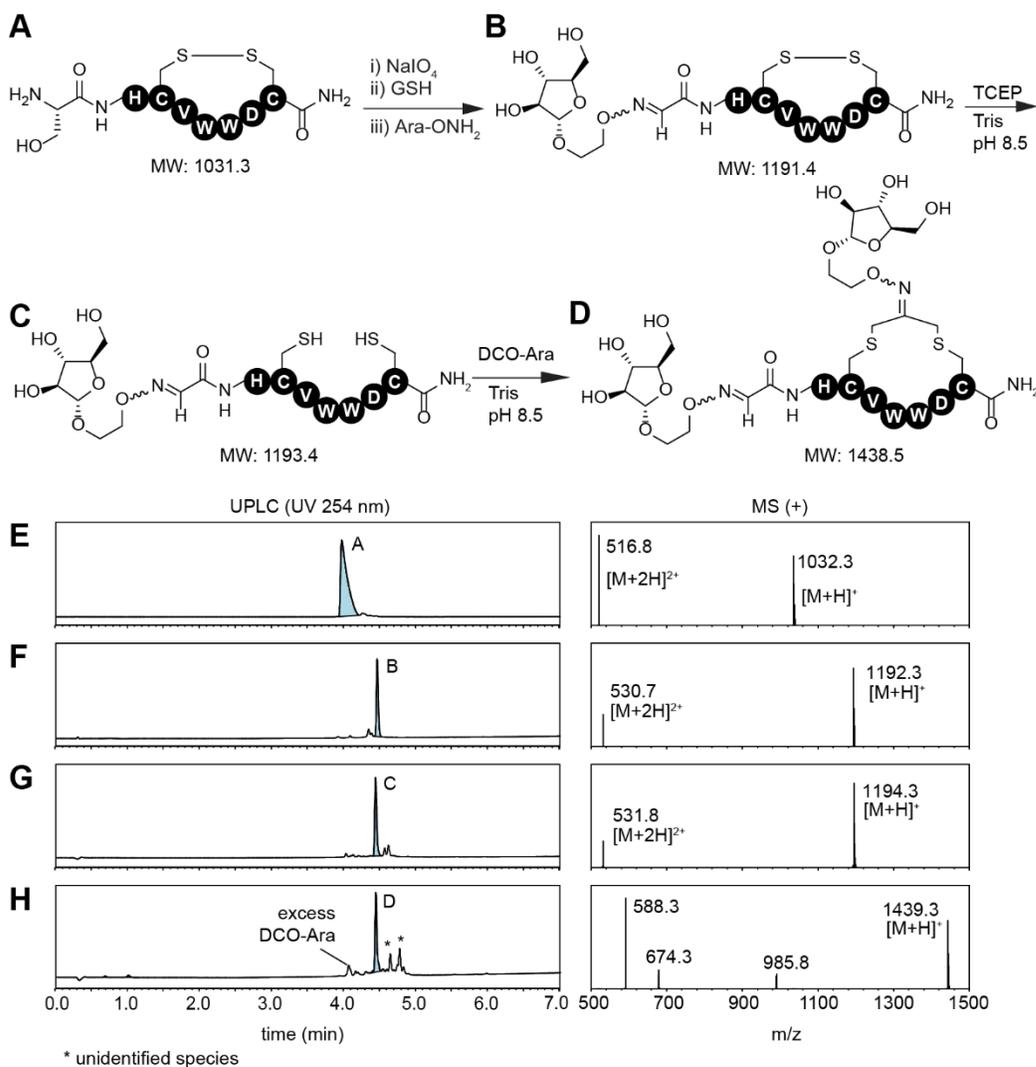


Figure 4-3. Validation of Ara-M-Ara reaction on peptide.

(A) Disulfide bridged peptide SHCVWDC is oxidized and ligated to α Ara-ONH₂ to yield (B) α Ara-HCVWDC. (C) Treatment with TCEP reduced the disulfide bridge. (D) α Ara-HCVWDC reacted with DCO- α Ara to yield Ara-M-Ara. The respective UPLC-MS trace of compound (A) (B) (C) (D) are in (E) (F) (G) (H).

Exposure of HPLC-purified α Ara-HCVWDC peptide disulfide to TCEP in Tris buffer (pH 8.5) reduced disulfide. Subsequent addition of dichloro-oxime- α -arabinose stapled the peptide into a macrocycle through two Cys (Figure 4-3 D). The overall structure carries two α -arabinose residues. LC-MS confirmed the identity of the product (Figure 4-

3 E-H). I then demonstrated that the conditions optimized for a purified peptide can be applied to modify peptides displayed on phage. Employing previously established N-terminal oxime ligation conditions, specifically, 60 μ M of NaIO₄ and 5 mM GSH, followed by 1 mM α Ara-ONH₂, converted a phage clone that displays the sequence SGCSVYC to α Ara-GCSVYC. Differential biotin capture confirmed 64 \pm 15% conversion of this ligation. This phage clone was exposed to TCEP and Tris buffer. Subsequent reaction with 1 mM DCO- α Ara yielded 45% Ara-M-Ara macrocycle. Quantification of reaction was confirmed by a differential biotin-iodoacetamide (BIA) capture before and after the reaction, as previously reported by our group (Figure 4-4 C).²²⁷ The same condition converted SXCX₃C library to Ara-M-Ara library in 7% yield (Figure 4-4 E). Two reasons for such low conversion were: (i) 50% of the library appeared to be unreactive with BIA, indicating either incomplete reduction of the peptide disulfides, or lack of peptide-disulfides on phage altogether, possibly due to proteolytic cleavage of this peptide from phage. (ii) Even after exposure to DCO- α Ara reagent in reaction conditions that completely convert peptide thiol to DCO conjugate, the library appeared to be reactive with BIA. This observation suggests that a large fraction of peptides still contain thiol or potentially other nucleophile that can react with BIA.

Due to the underwhelming performance of the second functionalization step, which introduces a DCO-Ara functionality into the library, I decided to proceed to panning using only the library that contained N-terminal Ara and internal disulfide (Figure 4-4 A). I envision that in the future when the reasons for poor DCO-functionalization are better understood, the screen can be also repeated with bis-modified Ara-M-Ara library (Figure 4-4 C).

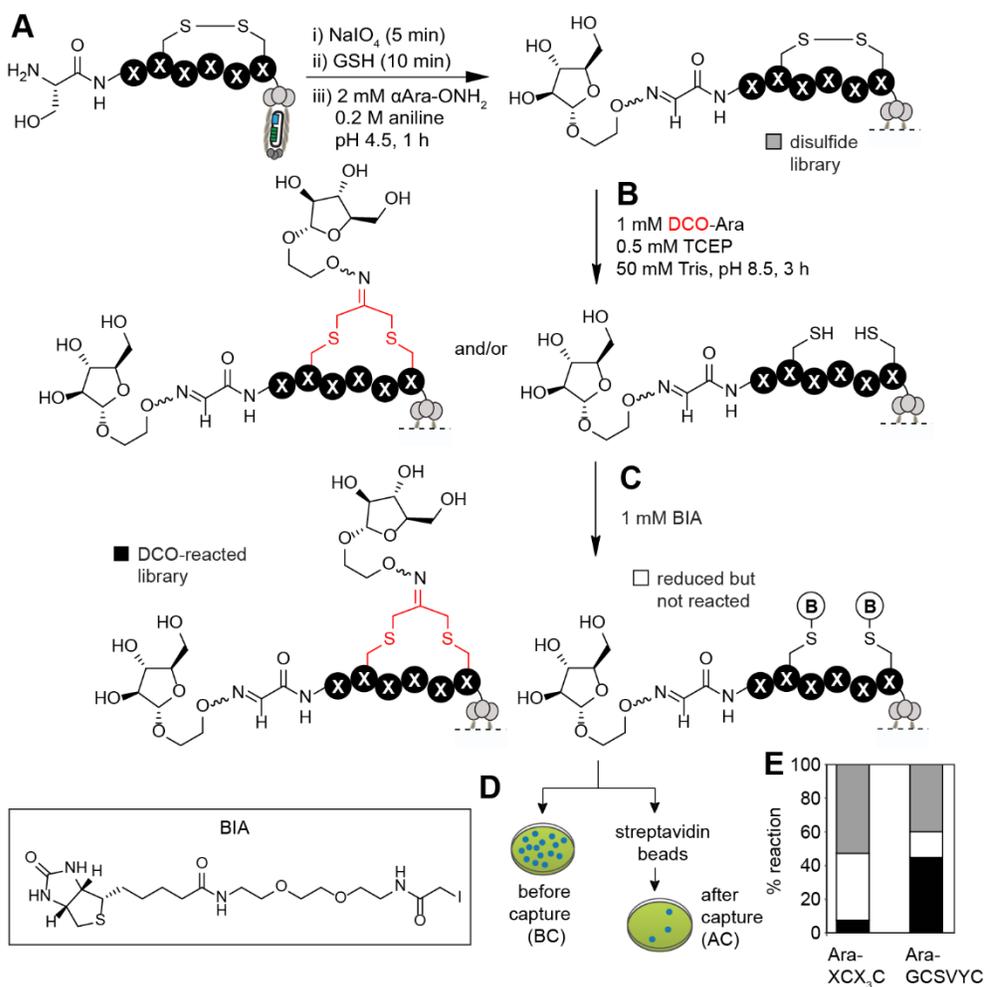


Figure 4-4. Ara-M-Ara reaction on phage.

(A) Phage displaying disulfide bridged peptide SXCX₃C was oxidized and ligated to $\alpha\text{Ara-ONH}_2$ to yield (B) $\alpha\text{Ara-XCX}_3\text{C}$ on phage. (C) Treatment with TCEP, followed by DCO-Ara yielded Ara-M-Ara or reduced but unreacted phage. (D) Exposure to BIA biotinylated unreacted phage, leaving DCO-reacted phage intact. (E) The conversion of the reactions that lead to Ara-M-Ara moiety on SXCX₃C library and SGCSVYC monoclonal phage as determined by the biotin capture technology.

4.4.2 Generation of input libraries for selection

Using the conditions optimized in section 4.4.1, I repeated the modification of the libraries at a scale that yielded sufficient number of phage for multiple replicates of panning (ca. 10^{12} PFU total). Modification of 6×10^{11} PFU of SX₇ library with

aminoxy-mannose yielded 2×10^{11} PFU of library that contained $58 \pm 24\%$ of Man-X₇ library. Modification of SWYDKYH clone with aminoxy-mannose yield a phage that contained $66 \pm 29\%$ Man-WYDKYH. Modification of 10^{12} PFU of SXCX₃C library with α Ara-ONH₂ yielded $\sim 10^{12}$ PFU of library that contained $48 \pm 16\%$ α Ara-XCX₃C library. Procedures for quantification of the chemical modifications were based on previously reported method.¹⁸⁰

4.4.3 Sequential pull-down assays to select binders for ConA, CS-35, 906.4321, and antiFLAG

In this sequential pull-down assay, each library (lib 1 and lib 2) was incubated with two different targets—biotinylated ConA and an antibody—at once, and then the mixture was subjected streptavidin-coated resins. The resins along with the co-precipitated biotinylated ConA and bound phage were removed with centrifugation force, and then the residual supernatant was pull-downed consecutively with protein G-coated resins, which served to capture the antibody. The antibodies used in this proof-of-concept experiment were aFLAG, CS-35, and 906.4321.

Prior to pull-down, resins were blocked with blocking reagents and “blocking phage” to reduce non-specific interactions. All library inputs were also prepared in blocking solutions that contained vast excess of blocking phage. These blocking phage, constructed by the post-doctoral fellow N. Bennett in our group, were designed in such a way that their DNA would not be amplified under the PCR conditions that I used to amplify the pull-downed library populations. This means that even present at high

concentration throughout the selection process, these blocking phage did not interfere with downstream sequencing and analysis.

The amount of each phage species in lib 1 was 3×10^9 PFU Man-X₇: 3×10^9 PFU Ara-XCX₃C: 4×10^4 PFU Man-WYDKYH: 1.8×10^4 PFU DYKDDDDK. The amount of each phage species in lib 2 was 2×10^9 PFU SX₇: 2×10^9 PFU SXCX₃C: 1.8×10^4 PFU DYKDDDDK. The amount of Man-WYDYH and DYKDDDDK used in the selection was based on two assumptions: (i) the fraction of control sequences should be high enough for the sequence to be visible after the selection and (ii) the fraction of the sequences should not be too high to avoid overpowering the recovery of selection libraries at a given depth of Illumina sequencing. More specifically, I assumed that the recovery percentage of the control sequence will be similar to our published report.²⁸⁷ Under this assumption, if the titer of the overall mixture input is 10^{10} PFU and the titer of the recovered mixture output is 10^6 ,²⁸⁷ then adding 6×10^4 PFU of control sequences to the input mixture will make the fraction of the control sequences in the overall input mixture (f_{input}) = $6 \times 10^4 / 10^{10} = 6 \times 10^{-6}$; and the fraction of the control sequences in the recovered output mixture (f_{output}) will be = $0.1 \times 6 \times 10^4 / 10^6 = 6 \times 10^{-3}$. At a typical depth of sequencing of 10^6 reads, these fractions will yield 6 reads of the control sequence in the input sequencing data set and 6000 reads in the output set.

The possible outcomes of the pull-downs using lib 1 or lib 2 against model targets ConA or aFLAG are illustrated in Figure 4-5 G-J. In each scenario, there exists resin-binding phage (non-specific interaction) and target-binding phage (specific interaction).

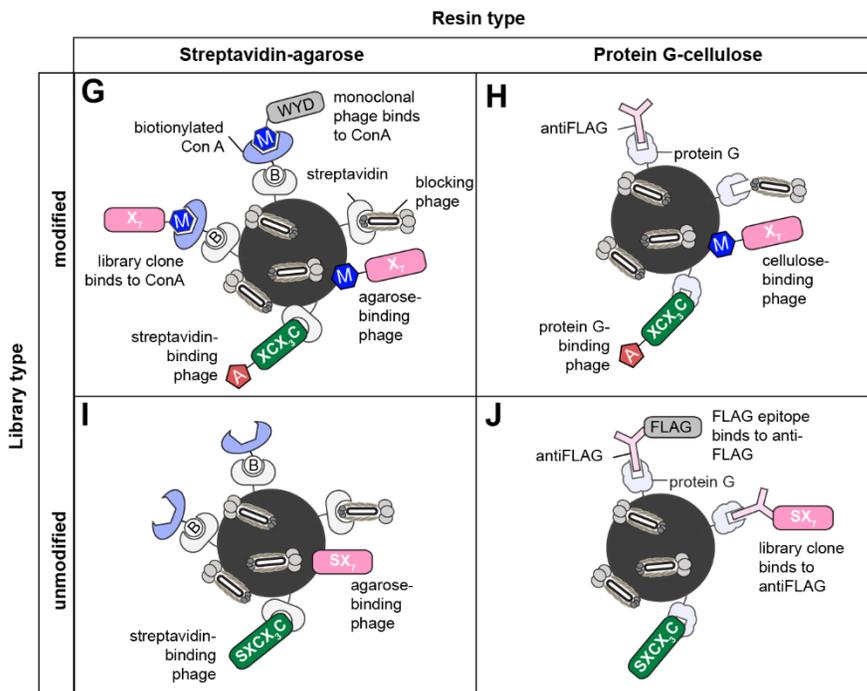
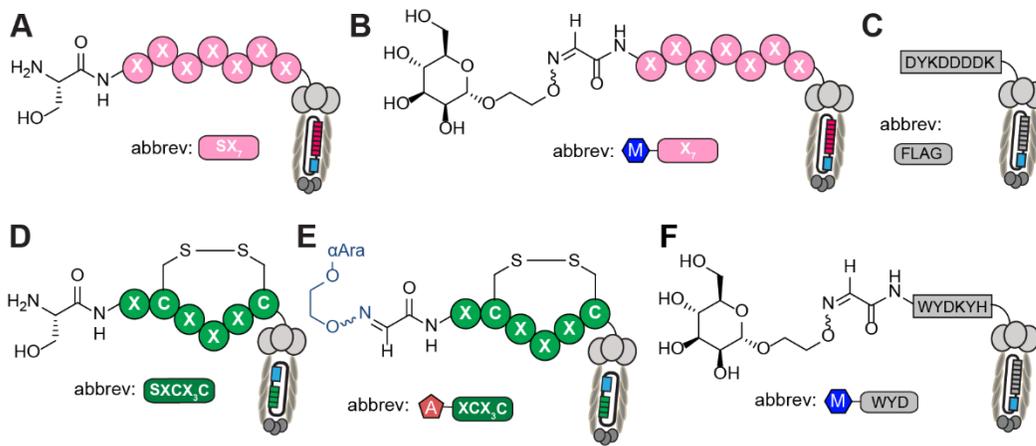


Figure 4-5. Glycopeptide libraries used for sequential pull-down assays and possible outcome of the assay. (A) unmodified SX₇ library; (B) mannosylated SX₇ (Man-X₇); (C) DYKDDDDK clone (FLAG); (D) unmodified SXCX₃C; (E) α Ara modified SXCX₃C library (α Ara-XCX₃C); (F) mannosylated SWYDKYH clone (Man-WYD); (G)-(J) Possible outcome of each pull-down step. In each scenario, there exists resin-binding phage and target-binding phage.

4.4.4 Analysis of Illumina sequencing

To validate the screening procedure against model target ConA, I identified four sets of sequences within the pull-down population against ConA. The four sets of sequences were: i) set **A**—naive lib **1**; ii) set **B**—lib **1** selected on ConA; iii) set **C**—lib **2** selected on ConA; iv) set **D**—lib **1** selected on CS-35. The composition of the population in set **B**, when compared to set **A**, showed significant enrichment of Man-X₇, and this enrichment was not observed in sets **C** and **D** (Figure 4-6 A). The enrichment of Man-X₇ in set **B** but not in set **C** confirmed that affinity was enhanced only when the linear X₇ sequence was modified with mannose (Figure 4-6 D). In addition, the enrichment of Man-X₇ in set **B** but not in set **D** indicates that the mannosylation is only effective when the screening was performed against ConA.

Next, I searched for the enrichment of the N-terminal motif SW[YFW] in sets **A**–**D**, as this motif, identified through GE-FBD, exhibited high affinity for ConA.⁵² As expected, this motif was enriched significantly only in set **B** (Figure 4-6B). In order to identify the sequences that have been enriched in set **B**, I used differential enrichment analysis (volcano analysis) and designated the result as the differences between two of the sets. For example, **B/A** is the difference between set **B** and set **A**. From the results of the difference sets—**B/A**, **B/C**, and **B/D**, I observed that the majority of enriched sequences originated from the Man-X₇ library (Figure 4-6 C), as anticipated. Moreover, a significant fraction of the enriched sequences at the intersection of difference sets **B/A**, **B/D** and **B/C** contain the motif SW[YFW]. These collective data demonstrate that the pull-down assay on model target ConA was functional.

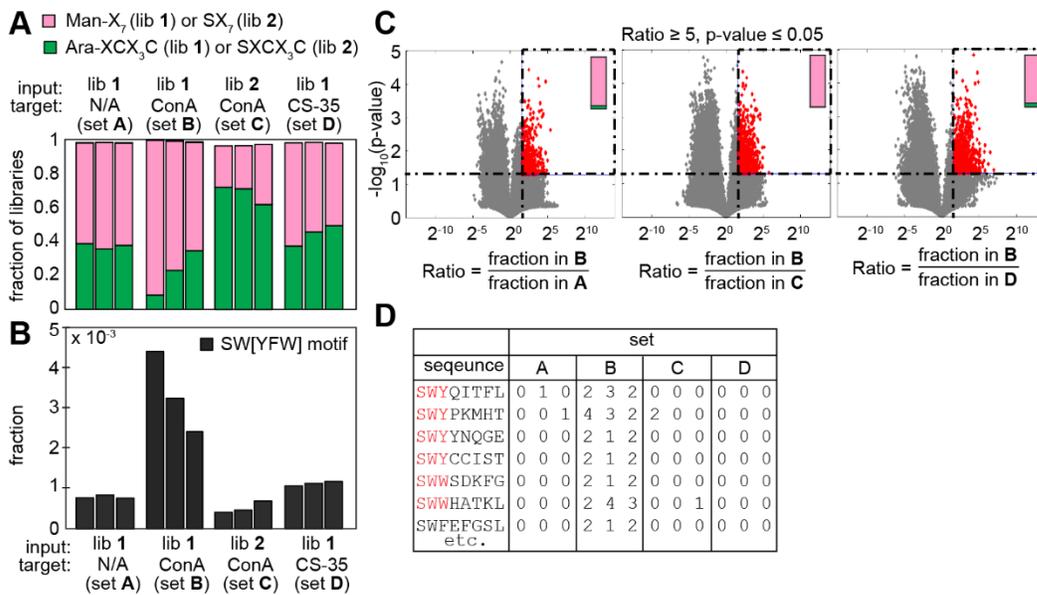


Figure 4-6. Post-selection analysis on model target ConA.

(A) Composition of SX7 and SXCX3C, or Man-X7 and Ara-XCX3C library in each selection output. (B) Fraction of unique sequences that contain SW[YFW] motifs in each selection screen. (C) Volcano plots showing sequences that are significantly enriched ($R \geq 5$, $p < 0.05$) using different set A, C, D as controls. We abbreviate the set of these sequences as B/A, B/C, and B/D. The composition of the enriched sequences were shown on the top right of each volcano plot. (D) Normalized frequency of the sequences that contain SW[YFW] motif, which originate from the enriched sequences in set B over set A, C, and D.

Using the same analysis tools, I validated the pull-down assay on model target aFLAG. I identified four sets of sequences from Illumina sequencing data: i) set **E**—naive lib 2; ii) set **F**—lib 2 selected on aFLAG; iii) set **G**—lib 2 selected on CS-35; iv) set **H**—lib 2 selected on 906.4321. Searching from the sequences within in each set, I observed higher instances of the motif DYK (partial FLAG epitope) within set **F** than set **E**, **G**, and **H** (Figure 4-7 B). These sequences that contain the DYK motif originate from both SXCX₃C and SX₇ libraries, but with a slight higher fraction from SXCX₃C library (Figure 4-7 A, set **F**). Using volcano analysis, I generated three difference sets—**F/E**, **F/G**,

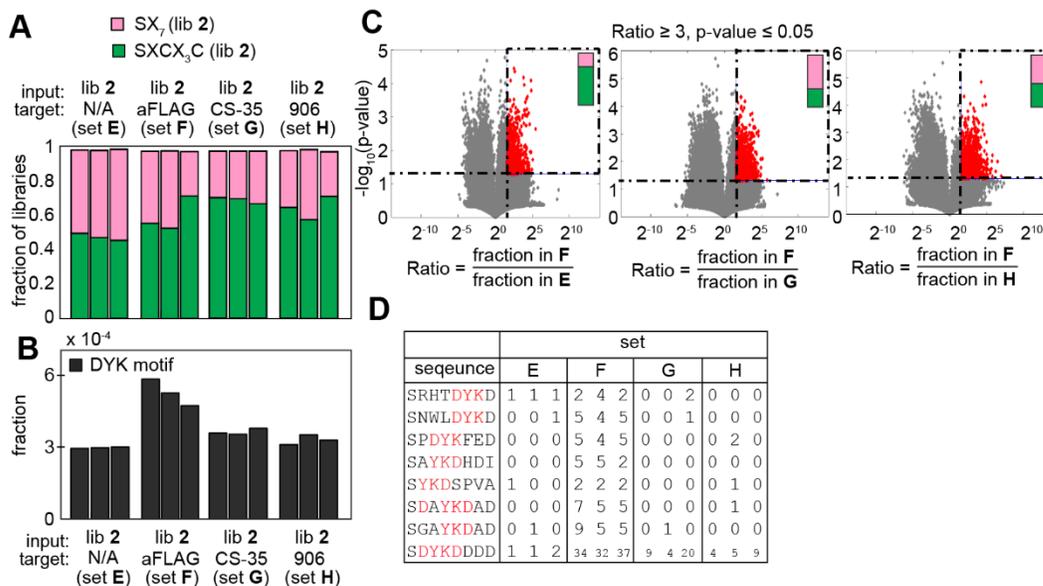


Figure 4-7. Post-selection analysis on model target aFLAG.

(A) Composition of SX_7 and $SXCX_3C$, or Man- X_7 and Ara- XCX_3C library in each selection output. (B) Fraction of unique sequences that contain DYK motifs in each selection screen. (C) Volcano plots showing sequences that are significantly enriched ($R \geq 3$, $p < 0.05$) using different set E, G, H as controls. We abbreviate the set of these sequences as F/E, F/G, and F/H. The composition of the enriched sequences was shown on the top right of each volcano plot. (D) Normalized frequency of the sequences that contain DYK motif, which originate from the enriched sequences in set F over set E, G, and H.

and F/H. The intersection of these three difference sets contained 359 unique sequences, several of which contained the DYK motif.

To examine whether the presences of these two motifs were random events, I searched for the instances of N-terminal SW[YFW] appearing in the sequences pull-downed against aFLAG, as well as the instances of DYK appearing in the population pull-downed with ConA. In the former case, 0.3% of sequence (1 in 359) from selection against aFLAG contained N-terminal SW[YFW] (SWYECCAS). In the latter case, 0.2% of sequence (3 in 1926) from ConA selection carried DYK motifs (SFWNVDYK, SNPWVDYK, SDDYKTRD). These percentages were comparatively lower than the

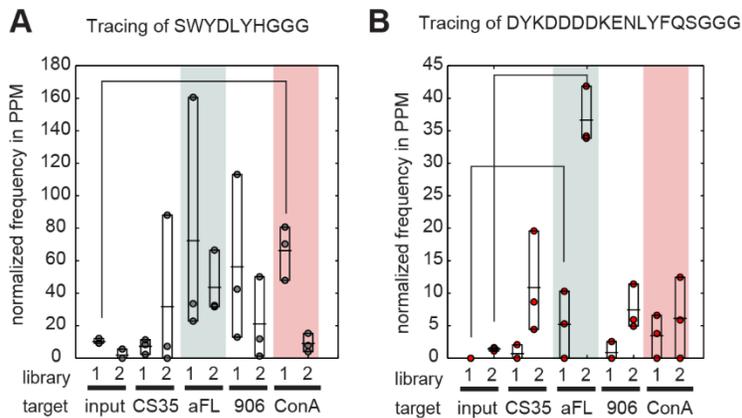


Figure 4-8. Tracing monoclonal phage with control sequences throughout each selection experiment.

(A) The frequency of Man-WYDLYH phage in lib 1 is enriched after selection on ConA. (B) The frequency of DYKDDDDK phage in lib 1 and lib 2 are enriched after selection on aFLAG. However, enrichment of this clone in screens against other targets was also observed.

percentages of DYK appearing in aFLAG screen (2.2%), or SW[YFW] in ConA screen (1.2%), suggesting that the appearances of the DYK (originating from SX₇ and SXCX₃C) and SW[YFW] (originating from Man-X₇) were not random events. However, tracing the fraction of the control monoclonal phage with sequences Man-WYDLYH and SDYKDDDDKENLYFQS, from lib 1, lib 2, to every pull-down screen, I observed only minor enrichment of these clones (Figure 4-8 A-B). For instance, 10⁵ of SDYKDDDDKENLYFQS clones were initially present in 10¹⁰ lib 2 clones, but the pull-down population against aFLAG (~10⁶ clones) contained only 40 SDYKDDDDKENLYFQS clones. The recovery of SDYKDDDDKENLYFQS clone was therefore 40/10⁶ = 0.0004% instead of the anticipated 10%, calculated in section 4.4.3. The reason for this low fraction of recovery remains unclear and the pull-down procedures are subjected to further optimization in order to maximize the recovery.

After the validation of the pull-down assay on model antibody aFLAG, I analyzed the selection outcomes from CS-35 pull-down. I defined one additional set from Illumina sequencing: set **I**— lib **1** selected on aFLAG. Comparing screens against CS-35 using a library containing macrocyclic glycopeptides (set **G**) to a library containing the macrocyclic counterparts without glycosylation (set **D**), I observed that the fraction of SXCX₃C library within the overall pull-down population is similar to that of α Ara-XCX₃C library (Figure 4-9 A), suggesting that the N-terminal arabinose may not confer significant synergy with the cyclic peptides structures within the α Ara-XCX₃C library when binding to CS-35.

When comparing set **D**, **I**, and **H**, I observed no trend of preferential binding of α Ara-XCX₃C to CS-35 (Figure 4-9 A); interestingly, the composition of the pull-down population contained a significant amount of Man-X₇. To identify the enriched sequences in these sets, I used four volcano analyses (Figure 4-9 B) and produced four difference sets: **D/A**, **D/I**, **D/H**, and **D/G**. The intersection of these four difference sets indicated several enriched sequences originated from Man-X₇ library (Figure 4-9 F). I noted that although the structural evidence indicated that CS-35 binds to LAM with naked Ara_f termini (AraLAM), it has also been reported that CS-35 could also recognize Ara_f termini capped with mannose residues (ManLAM).²⁸⁸ This interaction between CS-35 and ManLAM can possibly explain the abundance of mannose-modified glycopeptides in the population that was pull-downed with CS-35.

The collective data indicated that while there are some Ara-glycopeptides identified as putative hits, it is also possible that CS-35 can recognize the macrocyclic library without arabinosylation, and N-terminal arabinosylation on the macrocyclic

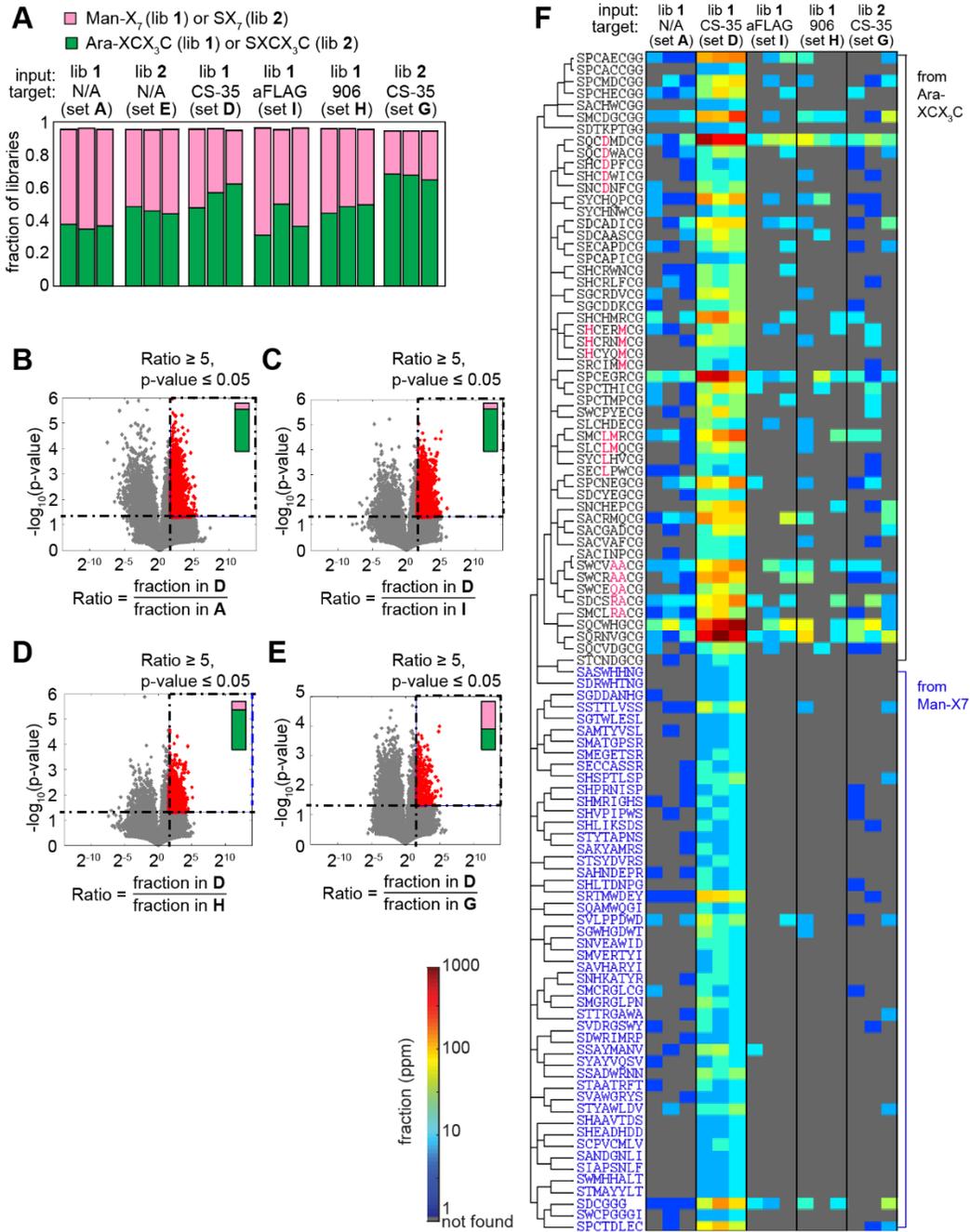


Figure 4-9. Post-selection analysis on Ara-XCX₃C selected against CS-35. (A) Composition of SX₇ and SXCX₃C, or Man-X₇ and Ara-XCX₃C library in each selection output. (B)-(E) Volcano plots showing sequences that are significantly enriched ($R \geq 3$, $p < 0.05$) using set I, H, and G as controls. We abbreviate the set of these sequences as D/A, D/I, D/H, and D/G. The composition of the enriched sequences was shown on the top right of each volcano plot. (D) Heat map showing putative hits from the intersection of D/A, D/I, D/H, and D/G in (D).

library may contribute only minimally to the specific binding interactions with CS-35. I believe this result leads to open possibilities to further refine the structure of the SXCX₃C library, for instance, incorporating a second carbohydrate moiety such as mannose or beta-arabinose, to obtain more specific ligands to CS-35.

4.5 Conclusion

In summary, I described a methodology to chemically modify phage libraries with two arabinose residues via two orthogonal reactions. These macrocyclic libraries provide potential for more efficient ligand search towards the discovery of specific binding epitopes of anti-LAM antibodies. I developed a deep-panning protocol that involves a single-round selection and Illumina sequencing. This system eliminates the need for multiple rounds of panning and amplification of phages; instead, phage binders were amplified solely through PCR and subjected to deep sequencing. Placing internal control inputs and model targets in panning allowed the demonstration that the system was functional, although the enrichment of the spiked binding sequences from the input library was significantly weaker than anticipated.

Once the deep-panning system was validated on model targets, I used a mixture of mannose-modified and arabinose-modified phage libraries to select against anti-LAM antibody CS-35. Differential enrichment analysis identified putative hits that originated from both Man-X₇ and Ara-XCX₃C libraries. The composition of the enriched sequences indicated that mannose may be useful for discovering specific binding ligands. Analysis of the composition of the recovered library after selection suggested that SXCX₃C and

Ara-XCX3C library could be recognized by CS-35 at a similar extent. Moreover, there was only minimal enrichment of the overall Ara-XCX3C library when comparing selection on CS-35 and closely-related anti-LAM 906.4321. I believe that the Ara-XCX₃C library may not contain *specific* binders, and a more refined library tailored towards CS-35 is necessary. Example of such library, according to our aforementioned hypothesis and data from sequencing, include mannosylated library or bis-modified Ara-M-Ara library.

In conclusion, this work describes a strategy to produce macrocyclic glycopeptide libraries with two distinct carbohydrate moieties which, in combination with the developed deep-panning technology, may offer a rapid route to discover specific binding ligands for serological detection of TB.

4.6 Experimental procedures

4.6.1 Materials and general information

Tris buffer was prepared as a solution of 50 mM Tris at pH of 8.5. HEPES buffer contains 10 mM HEPES, 150 mM NaCl, 0.01 mM CaCl₂, and 0.01 mM MnCl₂ with pH adjusted to 7.4 after preparation. All solutions used for phage work are sterilized filtration through 0.22 µm filters.

The synthesis of Man-ONH₂ was reported previously.⁵² αAra-ONH₂ was synthesized by R. Brunton from the Lowary Group. Aminoxy-biotin (AOB, #100009350) was purchased from Cayman Chemical. 1,3-Dichloroacetone (DCA, #168548) was purchased from Sigma-Aldrich. Biotin-PEG₂-iodoacetamide (BIA, #21334)

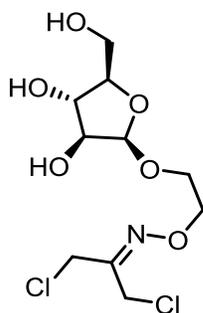
was purchased from Thermo Fisher Scientific. Product purification was accomplished with RP-HPLC on Waters HPLC system equipped with a Waters 1525 EF binary pump, a Waters FlexInject manual injector (dual mode) and a Waters 2489 tunable UV detector. SymmetryPrep™ C18 semi-preparative column (19 × 50 mm, particle size 5 μm, pore size 100 Å) was used for the purifications of DCO-derivatives at a typical flow rate of 12 mL/min. ¹H and ¹³C NMR spectra were acquired on Agilent/Varian VNMRS 600 MHz spectrometers. Chemical shifts (δ) are reported in ppm and coupling constants (J) are given in Hz. The following abbreviations classify the multiplet peaks in the ¹H NMR: s = singlet, d = doublet, t = triplet, m = multiplet or unresolved. HRMS (ESI) spectra were recorded on Agilent 6220 oaTOF mass spectrometer using either positive or negative ionization mode. Characterization of reaction crude was performed with UPLC-MS using a C18 column (Phenomenex Kinetex 1.7 μm EVO C18, 2.1×50 mm) running with a gradient of water/acetonitrile with 0.1% formic acid from 98/2 at 0 min to 40/60 at 5 min under a flow rate of 0.5 mL/min.

Production of mAb IgG of CS-35 and 906.4321 was performed as previously reported.²²⁴ antiFLAG antibody (A00187-200) was purchased from GenScript. The phage libraries SXCX₃C was prepared according to literature protocol²⁸⁹ using synthetic DNA oligos purchased from TriNucleotide Technologies. The use of SXCX₃C libraries in the synthesis of DCO-modified libraries was previously reported¹⁷⁸ The phage displaying sequence DYKDDDDKENLYFQS of the N-terminal portion of the pIII protein, where DYKDDDDK is the FLAG epitope and ENLYFQS is TEV-protease-cleavable sequence, was cloned by K. Tjhung in our published report.²⁸⁷ Streptavidin agarose resins (#20359) were purchased from Thermo Scientific. Protein G beads (G747A) were purchased from

Promega. ConA (C7275) from *Canavalis ensiformis* sp. was purchased from Sigma-Aldrich.

4.6.2 Synthesis of dichloro-oxime- α -arabinose (DCO- α Ara)

α Ara-ONH₂ (24 μ mol, 1.5 eq.) was added to 1,3-dichloroacetone (16 μ mol, 1 eq.) dissolved in MeOH. After 24 h reaction at RT, the reaction mixture was injected to a semi-preparative RP-HPLC system, using gradient indicated in Appendix to afford the product as oil (3.4 mg, 27% yield).



DCO- α Ara: ¹H NMR (600 MHz, D₂O): δ = 5.06 (d, J = 1.2 Hz, 1 H), 4.46 (s, 2 H), 4.40 (s, 2 H), 4.37 (t, J = 4.7 Hz, 2 H), 4.11 (dd, J=3.4 Hz, 1.8 Hz, 1 H), 4.07 (td, J = 5.8 Hz, 3.3 Hz, 1 H), 4.02 – 3.98 (m, 1H), 3.96 (dd, J = 6.2 Hz, 3.3 Hz, 1 H), 3.86 – 3.82 (m, 2 H), 3.71 (dd, J = 12.3 Hz, 5.6 Hz, 1 H); ¹³C NMR (176 MHz, D₂O) δ = 204.5, 108.3, 84.6, 51.8, 77.3, 74.4, 66.9, 62.1, 42.7, 34.4; HRMS (ESI) calcd. for C₁₀H₁₇Cl₂NNaO₆ [M+Na]⁺ m/z = 340.0325, found 340.033.

4.6.3 Reaction of SHCVWWDC with DCO- α -Ara

4.6.3.1 Synthesis of oxidized (disulfide-containing) SHCVWWDC

Peptide SHCWWDC was synthesized on Rink Amide resin using an according to standard protocol.⁵² Reduced SHCWWDC (4.8 μ mol) was dissolved in 160 μ L TFA and 4.8 mL of H₂O was added. 0.1 M I₂ in MeOH was added dropwise until yellow color sustained. Solid ascorbic acid was then added to the reaction mixture until yellow color disappeared. NH₄OH (15 N) was added until solution reached pH 4, as determined by universal pH paper. The solution was injected to semi-preparative RP-HPLC with gradient indicated in Appendix. The product was obtained as fluffy powder after lyophilization (3.7 mg, 74%).

4.6.3.2 Synthesis of α Ara-DCVWWDC

Disulfide-containing SHCWWDC (3.6 mg, 3.5 μ mol, 1 eq.) was dissolved in PBS (130 μ L, pH 7.4) and MeCN (150 μ L) in a 1.5-mL microcentrifuge tube. Sodium periodate (0.9 mg, 4.2 μ mol, 1.2 eq.) was added to the microcentrifuge tube and the reaction mixture was incubated for 1 h at RT. α Ara-ONH₂ (1 mg, 5.3 μ mol, 1.5 eq.) dissolved in 200 mM anilinium acetate (240 μ L, pH 4.7) was added to the reaction mixture. The oxime ligation was carried out for 1 h at RT. The reaction mixture was then injected to a RP-HPLC system using gradients described in Appendix. Lyophilization yielded α Ara-DCVWWDC as white solid (1.2 mg, 30%). The purity and identity of the product was confirmed with by UPLC-MS (see Appendix for spectra).

4.6.3.3 Synthesis of α Ara-M- α Ara

Glycopeptide α Ara-HCVWWDC (1.2 mg, 1 μ mol, 1 eq.) was dissolved in Tris (200 μ L, 50 mM, pH 8.5) in a 1.5-mL microcentrifuge tube. TCEP (2 μ mol, 2 eq.) was added, followed by the addition of Tris to adjust the pH to 8.5. The reduced cysteines were characterized by UPLC-MS. DCO- α Ara (0.48 mg, 1.5 μ mol, 1.5 eq.) was added to the reaction mixture and incubated at RT for 3 h. The purity and identity of the product was determined with UPLC-MS. Product was not isolated.

4.6.4 Generation of α Ara-M- α Ara libraries on phage

4.6.4.1 Bulk amplification of SXCX₃C and *N*-SerX₇ phage library

Libraries were amplified by adding the initial library (SXCX₃C: 100 μ L, 2.8×10^{10} PFU or *N*-SerX₇: 1 μ L, 2×10^{10} PFU) into 5 mL of ER2738 culture ($OD_{600} = 0.2$). The phage and bacterial mixtures were incubated for 40 min at 37 °C with vigorous shaking. The mixture was then transferred to 45 mL pre-warmed LB at 37 °C and proceeded to shaking for 3 h. The cultures were centrifuged (10 min, 12000 g) at 4 °C to pellet the bacterial cells. The supernatant was transferred to a fresh centrifuge tube and precipitated with PEC/NaCl at 4 °C overnight. After spinning for 15 min at 12000 g, the phage pellet was resuspended in 1 mL PBS and transferred to a 1.5-mL microcentrifuge tube. The solution of phage was then centrifuged (5 min, 14K rpm, 4 °C) to remove any insoluble byproducts; the supernatant was added to 200 μ L PEG/NaCl to precipitate for 2 h. The mixture was centrifuged (10 min, 14K rpm) and the phage pellet was resuspended

in 200 mL PBS. The final titer of the amplified libraries: SXCX₃C: 2.5×10^{13} PFU; *N*-SerX₇: 3×10^{12} PFU.

4.6.4.2 Generation of α Ara-XCX₃C library

SXCX₃C library was generated according to the published protocol.²⁸⁹ Prior to the chemical modification, the phage library was amplified in bulk following the procedure described in section 4.6.2 “Bulk amplification of SXCX₃C and *N*-SerX₇ phage library.” For the attachment of the first α Ara-ONH₂ residue at the N-terminus of the displayed peptide, the SXCX₃C phage library (500 μ L, 1.5×10^{12} PFU) was oxidized with 0.06 mM sodium periodate on ice for 5 min. The oxidation was quenched by adding 5 μ L of 100 mM solution of glutathione in MQ water and incubation at RT for 10 min. The oxidized library was split in two portions, each treated with equal volume of aminoxy- α -arabinose (2 mM α Ara-ONH₂ in 200 mM anilinium acetate buffer, pH 4.5) or aminoxy-biotin (2 mM AOB in 20 mM anilinium acetate buffer, pH 4.5) to yield modified libraries designated α Ara-XCX₃C and Bio-XCX₃C, respectively. The reaction mixtures were incubated for 1 h at RT, after which, the modified library α Ara-XCX₃C was split in two portions, one of which was buffer exchanged twice in Amicon® Ultra-4 Centrifugal filter units 10K MWCO using PBS buffer, and the other was buffer exchanged using Tris buffer (pH 8.5). Each buffer exchange was performed at 4700 rpm on Heraeus™ Multifuge™ X3R Centrifuge for 15 min. The recovered α Ara-XCX₃C phage in PBS buffer was used for panning experiments. The recovered α Ara-XCX₃C phage in Tris buffer was used for subsequent reaction to form Ara-M-Ara libraries.

To monitor the oxidation, a small portion of the Bio-X₇ library was diluted and captured in a biotin-capture assay as previously described.¹⁸⁰ Typically, 30% of the fractions of phage library were successfully oxidized. To quantify the modification with α Ara-ONH₂, a small portion of the α Ara-XCX₃C library was treated with AOB and characterized via biotin-capture assay as previously described.¹⁸⁰ Typically, 80% of the fractions of phage library has been modified with the α Ara-ONH₂.

4.6.4.3 Modification of α Ara-XCX₃C library with DCO- α Ara

α Ara-XCX₃C in Tris buffer pH 8.5 was prepared from 4.6.5.1 to yield a final titer of 6.2×10^{10} pfu/mL. To the mixture of phage (650 μ L), TCEP (6 μ L, 50 mM solution in water) was added and incubated at RT for 10 min. DCO- α Ara (6 μ L, 100 mM in water) was added to the reaction mixture and incubated at RT for 3 h. Then, an aliquot of the mixture (1 μ L) was diluted immediately by 100-fold with Tris buffer containing 0.5 mM TCEP, followed by the addition of biotin-PEG2-iodoacetamide (BIA, 1 μ L, 100 mM in DMF). The reaction mixture was incubated at RT for 30 min and the reaction was quenched by diluting 10^6 times. The modification efficiency was quantified by biotin capture assay as described previously.¹⁷⁸

4.6.5 Preparation of input mixture for selection

Five different phage inputs were used for selection. 1) α Ara-XCX₃C library (section 4.6.5.1); 2) SXCX₃C library (section 4.6.4); 3) Man-X₇;⁵² 4) Man-WYDLYH

modified with aminoxy-mannose using previously published procedure;⁵² 5) *N*-SerX₇ library (section 4.6.4); 6) DYKDDDDK_{ENLYFQS} phage.²⁸⁷

A mixture of α Ara-XCX₃C (3×10^{10} PFU/mL), Man-X₇ (3×10^{10} PFU/mL), Man-WYDLYH (4×10^5 PFU/mL), DYKDDDDK (1.8×10^5 PFU/mL), and blocking phage (6×10^{10} PFU/mL) in HEPES buffer was prepared. This mixture, designated as the test selection input reached a final titer of 5.2×10^{10} PFU/mL (excluding blocking phage, which appears white). Another mixture of SXCX₃C (2×10^{10} PFU/mL), SX₇ (2×10^{10} PFU/mL) DYKDDDDK_{ENLYFQS} (1.8×10^5 PFU/mL), and blocking phage (6×10^{10} PFU/mL) in HEPES buffer was prepared and designated as the control selection input. This mixture had a final titer of 4.8×10^{10} PFU/mL (excluding blocking phage, which appears in white).

4.6.6 Phage selection and extraction of ssDNA

First, I blocked streptavidin resin (20 μ L) and protein G beads (20 μ L) individually with 1 mL blocking solution (1% BSA, 10^{11} PFU/mL blocking phage, 0.1% tween in HEPES buffer) overnight on a rotator at 4 °C in 1.5-mL microcentrifuge tubes. We performed three technical replicates for each selection-target pair (test input-ConA-CS-35, control input-ConA-CS-35, test input-ConA-906.4321, control input-ConA-906.4321 test input-ConA-aFLAG, control input-ConA-aFLAG).

To each replicate, I added 10 μ g of biotinylated ConA, 10 μ g antibody (CS-35, 906.4321, or aFLAG), and 100 μ L of test selection input or control selection input prepared in section 4.6.5. Incubation for 1.5 h at 37 °C allowed the antibodies and the

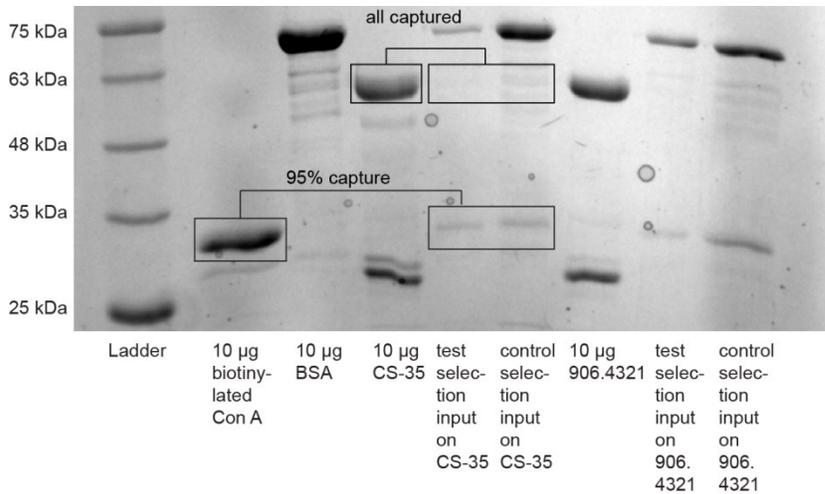


Figure 4-10. SDS-PAGE gel of selection supernatant after capturing antibody-phage mixture on protein G-coated beads.

biotinylated ConA to bind the phage. After incubation for 1.5 h, I transferred the phage-target mixture to preblocked streptavidin resin and allowed capturing of biotinylated ConA for 10 min on a shaker at 1200 rpm. The mixture was then centrifuged at 21,000 g for 1 min, and the supernatant was transferred to 20 uL preblocked protein G beads. After shaking for 10 min on a shaker at 1200 rpm, we used a magnetic stand to capture the protein G magnetic beads and discarded the supernatant. The efficiency of capture of ConA and aFLAG was confirmed by SDS-PAGE of the supernatant remaining after depletion of ConA. A minor amount of ConA was still present after the capture (Figure 4-10). I then washed the streptavidin resins and protein G beads 5 times using 1 mL of 0.1 % tween in HEPES buffer in each wash. After the final wash, I resuspended the beads in 30 µL 10 mM Tris, pH 8.5, added 30 µL of hexane, and shook the mixture at 3000 rpm for 10 min to extract phage DNA. Hexane was evaporated for 10 min on a heat block set

at 70 °C. The extracted phage ssDNA was subjected to PCR immediately according to the procedure in section 4.9.

4.6.7 Preparation for Illumina sequencing

The ssDNA extracted from section 4.8 was converted to Illumina-compatible dsDNA amplicon by PCR. Briefly, the ssDNA was combined with 1x Phusion[®] buffer, 200 µM dNTPs (each), 0.5 µM forward and reverse primers, and one unit Phusion[®] High-Fidelity DNA Polymerase in a total volume of 50 µL. Forward (F) and reverse (R) primer sequences are described in Chapter 2.

The temperature cycling protocol for PCR was as follows: 1) hold at 65 °C; 2) 98 °C for 3 min; 3) 98 °C for 10 s; 4) 50 °C for 20 s; 5) 72 °C for 30 s; 6) repeat 3-5 for 10 times; 7) 98 °C for 10 s; 8) 72 °C for 30 s; 9) repeat 7-8 for 20 times; 10) 72 °C for 5 min; 11) hold at 4 °C. The resulting amplicons were characterized by gel electrophoresis using 2% agarose gel, pooled together (20 ng per sample) and purified on E-Gel[®] SizeSelect[™] 2% agarose gel (Invitrogen, #G6610-02). Sequencing was performed using the Illumina NextSeq platform (Molecular Biology Service Unit, University of Alberta) and analyzed as described in the subsequent section 4.10 “Analysis of the deep-sequencing data.”

4.6.8 Analysis of the deep-sequencing data

Matlab scrips for analysis of the data were analogous to those used in section 2. Filtered data subjected to DE-analysis yielded 500-2000 putative hits, which were analyzed manually to characterize the fraction of C3C and SX₇ motifs.

Chapter 5: Conclusion and outlook

5.1 Conclusion

This thesis describes an enabling technology to facilitate the discovery of carbohydrate-based antigens for antibodies associated with mycobacterial infections. This technology serves as a promising avenue for the development of rapid, efficient TB serodiagnostics. Existing commercially available TB serological tests that rely on antigen-antibody interactions are shown variable in specificity and sensitivity,^{203, 215, 290} hindering correct TB diagnoses. Global TB pandemic remains an unsolved issue.

The diagnostic potential of a serological test is highly dependent on the antigens used in the test. Often, these antigens come from native²⁹¹⁻²⁹³ or recombinant proteins²⁹⁴⁻²⁹⁶ relevant to mycobacteria pathogenicity, or glycolipids present in mycobacterial cell wall²⁹⁷⁻²⁹⁹. The latter has attracted substantial amount of interest as antigens for diagnostic platforms, because it was shown that certain fragments of LAM can be recognized by specific antibodies in infected individuals.^{212, 300} Several groups conducted studies towards identifying structures in LAM that can detect antibodies associated with mycobacterial infections. Some approaches focus on testing synthetic complex carbohydrate libraries,²²⁴ while others seek to discover carbohydrate mimotopes via screening random peptide libraries.^{109, 301}

The former approach requires libraries of structurally complex glycans, which inevitably involves multistep syntheses and complicated chemical manipulations. As a result, the overall discovery can be challenged when compounds need to be scaled-up for validation purposes. On the other hand, the latter approach replaces complex glycan

libraries with random peptide libraries. However, these discovered peptides, in several cases, showed suboptimal diagnostic potential when applied on real-world serum samples. One fundamental reason is limited chemical diversity in peptide libraries (section 4.1). Chemical modifications can introduce topological constraints and expand functionalities. Several groups have shown successful incorporation of post translational modifications to genetically-encoded peptide libraries (section 4.1), which make them attractive sources for the discovery of disease-specific antigens.

Previously, we showed a GE-FBD that employs glycopeptide libraries displayed on phage can accelerate ligand discovery process for carbohydrate-binding proteins.⁵² The key of the design is to identify a potential second site located in close proximity to the glycan-binding site and link the fragment at the second site to the glycan for improved binding affinity. We believe that in a similar strategy which combines the carbohydrate moiety with phage-displayed random peptide libraries (diversity $> 10^8$), we can target specifically to the rare carbohydrate-binding antibodies in serum samples that are associated with mycobacteria infections.

In Chapter 2, I have demonstrated that a hybrid carbohydrate-peptide library can be a useful source for rapid discovery of antigens for antibodies associated with mycobacterial infections. Using a model anti-LAM antibody as target, hexasaccharide attached to the peptide libraries guides the selection and yields glycopeptide ligands that bind to target in a synergistic fashion. Without the hexasaccharide anchor, these hits could not be identified from the selection of the unmodified peptide library, and without the peptide portion, the hexasaccharide exhibited compromised specificity. The most promising glycopeptide showed binding affinity 10 fold greater than the hexasaccharide

alone. These selected glycopeptide ligands also showed enhancement of both affinity and specificity when incorporated into multivalent assays, a format which is employed in most serological tests.

In Chapter 3, I explored the structure-activity relationship between the truncated hexasaccharides from the selected glycopeptides (from Chapter 2) and the model target anti-LAM antibody. Unfortunately, the glycopeptides with monosaccharides truncated from the original hexasaccharides failed to demonstrate synergistic binding interactions to the target. As a result, no binding signals was observed when incorporated into multivalent assay. The loss of affinity suggests that for these peptide sequences tested, additional fragments from the hexasaccharide are required to retain binding interactions with anti-LAM antibodies. In order to discover monosaccharide-containing glycopeptides as antigens, truncating existing hexasaccharide-containing glycopeptides may not work. Instead, monosaccharide-containing glycopeptide antigens have to be discovered *de novo* from screening monosaccharide-containing glycopeptide libraries.

In Chapter 4, I describe an optimized methodology for dual modifications on phage libraries. We anticipate the resulting macrocyclic glycopeptide libraries that carry two carbohydrate fragments to be useful in fragment-based antigen discovery for TB. To facilitate the selection process, I developed a deep-panning system that compacts the traditional multiple round panning to a single round. It holds promise for unbiased selection. With internal controls spiked in the procedure that do not interfere with ligand-target interactions, consistency in selection outcome, which is especially important when new targets and inputs are used, can be monitored repeatedly.

5.2 Future directions

With the advancement of display technology, GE-FED is emerging rapidly. Several projects based on GE-FED are currently on-going in our lab. These projects include selection of galactose-containing peptides for the galectin superfamily, selection of decafluoro-diphenylsulfone-containing macrocyclic peptides³⁰² for Nodal protein, and selection of macrocyclic peptide binders for DC-SIGN.

While epitope mapping of anti-LAM antibody CS-35 was successful in Chapter 2, it is unlikely to provide predictive epitopes for real world TB(+) serum samples. Mapping of epitopes in serum is inherently difficult because a library of ligands is screened against a library of antibodies. The presence of an Ara anchor on every peptide epitope will focus the library-vs-library (LvL) screen and maximize the identification of the relevant epitope. Initiated with the development of site-specific dual modifications that allows the formation of macrocyclic glycopeptides that carry two distinct carbohydrate moieties, my next step will be screening the established Ara-M-Ara libraries on serum samples. Using the deep-panning system that I developed (in Chapter 4), we bypass amplification bias¹⁸⁷ (the primary reason for the failure of LvL selection) and thus improve the quality of LvL selection.

The WHO stresses the importance of targeted research for serological and/or point-of-care tests with improved accuracy for TB-diagnoses. The technology and strategy developed in this thesis can address this issue. While low-cost devices and implementation of diagnostics is out of scope of this thesis, I build my projects around simple-to-make glycopeptide-based ligands. In the future, when integrated into a rapid

diagnostic platform, these components can yield a test that allows for broad population-base screening and alleviate low accuracy problems from existing serological tests.

References

1. Money, D., The laboratory diagnosis of bacterial vaginosis. *Can. J. Infect. Dis.* **2005**, *16* (2), 77-79.
2. Xu, Y.; Moser, C.; Al-Soud, W. A.; Sørensen, S.; Høiby, N.; Nielsen, P. H.; Thomsen, T. R., Culture-Dependent and -Independent Investigations of Microbial Diversity on Urinary Catheters. *J. Clin. Microbiol.* **2012**, *50* (12), 3901-3908.
3. Hodinka, R. L.; Kaiser, L., Is the Era of Viral Culture Over in the Clinical Microbiology Laboratory? *J. Clin. Microbiol.* **2013**, *51* (1), 2-8.
4. Leland, D. S.; Ginocchio, C. C., Role of Cell Culture for Virus Detection in the Age of Technology. *Clin. Microbiol. Rev.* **2007**, *20* (1), 49-78.
5. Lindström, M.; Korkeala, H., Laboratory Diagnostics of Botulism. *Clin. Microbiol. Rev.* **2006**, *19* (2), 298-314.
6. Suchett-Kaye, G.; Morrier, J. J.; Barsotti, O., Clinical usefulness of microbiological diagnostic tools in the management of periodontal disease. *Res. Microbiol.* **2001**, *152* (7), 631-9.
7. Tanner, A. C. R.; Listgarten, M. A.; Ebersole, J. L.; Strzempko, M. N., *Bacteroides forsythus* sp. nov., a Slow-Growing, Fusiform *Bacteroides* sp. from the Human Oral Cavity. *Int. J. Syst. Evol. Microbiol.* **1986**, *36* (2), 213-221.
8. Kook, J. K.; Sakamoto, T.; Nishi, K.; Kim, M. K.; Seong, J. H.; Son, Y. N.; Kim, D. K., Detection of *Tannerella forsythia* and/or *Prevotella intermedia* might be useful for microbial predictive markers for the outcome of initial periodontal treatment in Koreans. *Microbiol. Immunol.* **2005**, *49* (1), 9-16.
9. Ishii, Y.; Imamura, K.; Kikuchi, Y.; Miyagawa, S.; Hamada, R.; Sekino, J.; Sugito, H.; Ishihara, K.; Saito, A., Point-of-care detection of *Tannerella forsythia* using an antigen-antibody assisted dielectrophoretic impedance measurement method. *Microb. Pathog.* **2015**, *82*, 37-42.
10. Arnon, S. S.; Schechter, R.; Inglesby, T. V.; et al., Botulinum toxin as a biological weapon: Medical and public health management. *JAMA* **2001**, *285* (8), 1059-1070.
11. Ma, H.; Zhou, B.; Kim, Y.; Janda, K. D., A cyclic peptide-polymer probe for the detection of *Clostridium botulinum* neurotoxin serotype A. *Toxicon : official journal of the International Society on Toxinology* **2006**, *47* (8), 901-908.
12. Brown, K. C., Peptidic tumor targeting agents: the road from phage display peptide selections to clinical applications. *Curr. Pharm. Des.* **2010**, *16* (9), 1040-54.
13. Smith, G. P.; Petrenko, V. A., Phage Display. *Chem. Rev.* **1997**, *97* (2), 391-410.
14. Rakonjac, J.; Bennett, N. J.; Spagnuolo, J.; Gagic, D.; Russel, M., Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Curr. Issues Mol. Biol.* **2011**, *13* (2), 51-76.
15. Hamzeh-Mivehroud, M.; Alizadeh, A. A.; Morris, M. B.; Church, W. B.; Dastmalchi, S., Phage display as a technology delivering on the promise of peptide drug discovery. *Drug Discov. Today* **2013**, *18* (23-24), 1144-57.
16. Kehoe, J. W.; Kay, B. K., Filamentous phage display in the new millennium. *Chem. Rev.* **2005**, *105* (11), 4056-72.
17. Danner, S.; Belasco, J. G., T7 phage display: A novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl. Acad. Sci. USA* **2001**, *98* (23), 12954-12959.

18. Krumpe, L. R.; Atkinson, A. J.; Smythers, G. W.; Kandel, A.; Schumacher, K. M.; McMahon, J. B.; Makowski, L.; Mori, T., T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. *Proteomics* **2006**, *6* (15), 4210-22.
19. Ishi, K.; Sugawara, F., A facile method to screen inhibitors of protein-protein interactions including MDM2-p53 displayed on T7 phage. *Biochem. Pharmacol.* **2008**, *75* (9), 1743-50.
20. Krumpe, L. R.; Mori, T., T7 lytic phage-displayed peptide libraries: construction and diversity characterization. *Methods Mol. Biol.* **2014**, *1088*, 51-66.
21. Talwar, H.; Rosati, R.; Li, J.; Kissner, D.; Ghosh, S.; Fernández-Madrid, F.; Samavati, L., Development of a T7 Phage Display Library to Detect Sarcoidosis and Tuberculosis by a Panel of Novel Antigens. *EBioMedicine* **2015**, *2* (4), 341-350.
22. Gamkrelidze, M.; Dąbrowska, K., T4 bacteriophage as a phage display platform. *Arch. Microbiol.* **2014**, *196* (7), 473-479.
23. Ren, Z. J.; Lewis, G. K.; Wingfield, P. T.; Locke, E. G.; Steven, A. C.; Black, L. W., Phage display of intact domains at high copy number: a system based on SOC, the small outer capsid protein of bacteriophage T4. *Protein Sci.* **1996**, *5* (9), 1833-43.
24. Ren, Z.-j.; Black, L. W., Phage T4 SOC and HOC display of biologically active, full-length proteins on the viral capsid. *Gene* **1998**, *215* (2), 439-444.
25. Sternberg, N.; Hoess, R. H., Display of peptides and proteins on the surface of bacteriophage lambda. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (5), 1609-1613.
26. Mikawa, Y. G.; Maruyama, I. N.; Brenner, S., Surface display of proteins on bacteriophage lambda heads. *J. Mol. Biol.* **1996**, *262* (1), 21-30.
27. Gupta, A.; Onda, M.; Pastan, I.; Adhya, S.; Chaudhary, V. K., High-density functional display of proteins on bacteriophage lambda. *J. Mol. Biol.* **2003**, *334* (2), 241-54.
28. Beghetto, E.; Gargano, N., Lambda-Display: A Powerful Tool for Antigen Discovery. *Molecules* **2011**, *16* (4), 3089.
29. Pavoni, E.; Vaccaro, P.; D'Alessio, V.; De Santis, R.; Minenkova, O., Simultaneous display of two large proteins on the head and tail of bacteriophage lambda. *BMC Biotechnol.* **2013**, *13* (1), 79.
30. Bradbury, A. R. M.; Marks, J. D., Antibodies from phage antibody libraries. *J. Immunol. Methods* **2004**, *290* (1-2), 29-49.
31. Scott, J. K.; Smith, G. P., Searching for peptide ligands with an epitope library. *Science* **1990**, *249* (4967), 386-90.
32. Kay, B. K.; Adey, N. B.; Yun-Sheng, H.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M., An M13 phage library displaying random 38-amino-acid peptides as a source of novel sequences with affinity to selected targets. *Gene* **1993**, *128* (1), 59-65.
33. Parmley, S. F.; Smith, G. P., Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **1988**, *73* (2), 305-18.
34. Petrenko, V. A.; Smith, G. P.; Gong, X.; Quinn, T., A library of organic landscapes on filamentous phage. *Protein Eng.* **1996**, *9* (9), 797-801.
35. Enshell-Seijffers, D.; Smelyanski, L.; Gershoni, J. M., The rational design of a 'type 88' genetically stable peptide display vector in the filamentous bacteriophage fd. *Nucleic Acids Res.* **2001**, *29* (10), e50-e50.

36. Zacher Iii, A. N.; Stock, C. A.; Golden Ii, J. W.; Smith, G. P., A new filamentous phage cloning vector: fd-tet. *Gene* **1980**, *9* (1–2), 127-140.
37. Smith, G. P.; Fernandez, A. M., Effect of DNA copy number on genetic stability of phage-displayed peptides. *Biotechniques* **2004**, *36* (4), 610-4, 616, 618.
38. Loset, G. A.; Sandlie, I., Next generation phage display by use of pVII and pIX as display scaffolds. *Methods* **2012**, *58* (1), 40-6.
39. Sioud, M.; Hansen, M.; Dybwad, A., Profiling the immune responses in patient sera with peptide and cDNA display libraries. *Int. J. Mol. Med.* **2000**, *6* (2), 123-8.
40. Tian, F.; Tsao, M. L.; Schultz, P. G., A phage display system with unnatural amino acids. *J. Am. Chem. Soc.* **2004**, *126* (49), 15962-3.
41. Day, J. W.; Kim, C. H.; Smider, V. V.; Schultz, P. G., Identification of metal ion binding peptides containing unnatural amino acids by phage display. *Bioorg. Med. Chem. Lett.* **2013**, *23* (9), 2598-600.
42. Sandman, K. E.; Benner, J. S.; Noren, C. J., Phage Display of Selenopeptides. *J. Am. Chem. Soc.* **2000**, *122* (5), 960-961.
43. Heinis, C.; Rutherford, T.; Freund, S.; Winter, G., Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502-507.
44. Meyer, S. C.; Shomin, C. D.; Gaj, T.; Ghosh, I., Tethering small molecules to a phage display library: discovery of a selective bivalent inhibitor of protein kinase A. *J. Am. Chem. Soc.* **2007**, *129* (45), 13812-3.
45. Woiwode, T. F.; Haggerty, J. E.; Katz, R.; Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Cwirla, S. E., Synthetic compound libraries displayed on the surface of encoded bacteriophage. *Chem. Biol.* **2003**, *10* (9), 847-58.
46. Rentero Rebollo, I.; Heinis, C., Phage selection of bicyclic peptides. *Methods* **2013**, *60* (1), 46-54.
47. Ng, S.; Jafari, M. R.; Derda, R., Bacteriophages and viruses as a support for organic synthesis and combinatorial chemistry. *ACS Chem. Biol.* **2012**, *7* (1), 123-38.
48. Fukunaga, K.; Hatanaka, T.; Ito, Y.; Taki, M., Gp10 based-thioetherification (10BASEd-T) on a displaying library peptide of bacteriophage T7. *Mol. Biosyst.* **2013**, *9* (12), 2988-2991.
49. Arai, K.; Tsutsumi, H.; Mihara, H., A monosaccharide-modified peptide phage library for screening of ligands to carbohydrate-binding proteins. *Bioorg. Med. Chem. Lett.* **2013**, *23* (17), 4940-3.
50. Jafari, M. R.; Deng, L.; Kitov, P. I.; Ng, S.; Matochko, W. L.; Tjhung, K. F.; Zeberoff, A.; Elias, A.; Klassen, J. S.; Derda, R., Discovery of light-responsive ligands through screening of a light-responsive genetically encoded library. *ACS Chem. Biol.* **2014**, *9* (2), 443-50.
51. Dwyer, M. A.; Lu, W.; Dwyer, J. J.; Kossiakoff, A. A., Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol.* *7* (4), 263-274.
52. Ng, S.; Lin, E.; Kitov, P. I.; Tjhung, K. F.; Gerlits, O. O.; Deng, L.; Kasper, B.; Sood, A.; Paschal, B. M.; Zhang, P.; Ling, C.-C.; Klassen, J. S.; Noren, C. J.; Mahal, L. K.; Woods, R. J.; Coates, L.; Derda, R., Genetically Encoded Fragment-Based Discovery of Glycopeptide Ligands for Carbohydrate-Binding Proteins. *J. Am. Chem. Soc.* **2015**, *137* (16), 5248-5251.

53. Carlsson, F.; Trilling, M.; Perez, F.; Ohlin, M., A dimerized single-chain variable fragment system for the assessment of neutralizing activity of phage display-selected antibody fragments specific for cytomegalovirus. *J. Immunol. Methods* **2012**, *376* (1-2), 69-78.
54. Hamers-Casterman, C.; Atarhouch, T.; Muyldermans, S.; Robinson, G.; Hamers, C.; Songa, E. B.; Bendahman, N.; Hamers, R., Naturally occurring antibodies devoid of light chains. *Nature* **1993**, *363* (6428), 446-8.
55. Arbabi Ghahroudi, M.; Desmyter, A.; Wyns, L.; Hamers, R.; Muyldermans, S., Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* **1997**, *414* (3), 521-6.
56. Smith, K. A.; Nelson, P. N.; Warren, P.; Astley, S. J.; Murray, P. G.; Greenman, J., Demystified... recombinant antibodies. *J. Clin. Pathol.* **2004**, *57* (9), 912-917.
57. McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J., Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **1990**, *348* (6301), 552-554.
58. Baird, C. L.; Fischer, C. J.; Pefaur, N. B.; Miller, K. D.; Kagan, J.; Srivastava, S.; Rodland, K. D., Developing recombinant antibodies for biomarker detection. *Cancer Biomark.* **2010**, *6* (5-6), 271-9.
59. Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R., Making Antibodies by Phage Display Technology. *Annu. Rev. Immunol.* **1994**, *12* (1), 433-455.
60. Hammers, C. M.; Stanley, J. R., Antibody Phage Display: Technique and Applications. *J. Invest. Dermatol.* **2014**, *134* (2), 1-5.
61. Huang, J.; Ru, B.; Zhu, P.; Nie, F.; Yang, J.; Wang, X.; Dai, P.; Lin, H.; Guo, F. B.; Rao, N., MimoDB 2.0: a mimotope database and beyond. *Nucleic Acids Res.* **2012**, *40* (Database issue), D271-7.
62. He, B.; Chai, G.; Duan, Y.; Yan, Z.; Qiu, L.; Zhang, H.; Liu, Z.; He, Q.; Han, K.; Ru, B.; Guo, F.-B.; Ding, H.; Lin, H.; Wang, X.; Rao, N.; Zhou, P.; Huang, J., BDB: biopanning data bank. *Nucleic Acids Res.* **2015**.
63. Wu, J.; Zeng, X. Q.; Zhang, H. B.; Ni, H. Z.; Pei, L.; Zou, L. R.; Liang, L. J.; Zhang, X.; Lin, J. Y.; Ke, C. W., Novel phage display-derived H5N1-specific scFvs with potential use in rapid avian flu diagnosis. *J. Microbiol. Biotechnol.* **2014**, *24* (5), 704-13.
64. Griep, R. A.; Prins, M.; van Twisk, C.; Keller, H. J. H. G.; Kerschbaumer, R. J.; Kormelink, R.; Goldbach, R. W.; Schots, A., Application of Phage Display in Selecting Tomato spotted wilt virus-Specific Single-Chain Antibodies (scFvs) for Sensitive Diagnosis in ELISA. *Phytopathology* **2000**, *90* (2), 183-190.
65. McElhiney, J.; Drever, M.; Lawton, L. A.; Porter, A. J., Rapid isolation of a single-chain antibody against the cyanobacterial toxin microcystin-LR by phage display and its use in the immunoaffinity concentration of microcystins from water. *Appl. Environ. Microbiol.* **2002**, *68* (11), 5288-95.
66. McElhiney, J.; Lawton, L. A.; Porter, A. J. R., Detection and quantification of microcystins (cyanobacterial hepatotoxins) with recombinant antibody fragments isolated from a naïve human phage display library. *FEMS Microbiol. Lett.* **2000**, *193* (1), 83-88.
67. Strachan, G.; McElhiney, J.; Drever, M. R.; McIntosh, F.; Lawton, L. A.; Porter, A. J. R., Rapid selection of anti-hapten antibodies isolated from synthetic and semi-synthetic antibody phage display libraries expressed in *Escherichia coli*. *FEMS Microbiol. Lett.* **2002**, *210* (2), 257-261.

68. Garet, E.; Cabado, A. G.; Vieites, J. M.; González-Fernández, Á., Rapid isolation of single-chain antibodies by phage display technology directed against one of the most potent marine toxins: Palytoxin. *Toxicon* **2010**, *55* (8), 1519-1526.
69. Bublil, E. M.; Yeger-Azuz, S.; Gershoni, J. M., Computational prediction of the cross-reactive neutralizing epitope corresponding to the [corrected] monoclonal [corrected] antibody b12 specific for HIV-1 gp120. *FASEB J.* **2006**, *20* (11), 1762-74.
70. Enshell-Seijffers, D.; Denisov, D.; Groisman, B.; Smelyanski, L.; Meyuhas, R.; Gross, G.; Denisova, G.; Gershoni, J. M., The mapping and reconstitution of a conformational discontinuous B-cell epitope of HIV-1. *J. Mol. Biol.* **2003**, *334* (1), 87-101.
71. Huang, J. H.; Liu, Z. Q.; Liu, S.; Jiang, S.; Chen, Y. H., Identification of the HIV-1 gp41 core-binding motif--HXXNPF. *FEBS Lett.* **2006**, *580* (20), 4807-14.
72. Jellis, C. L.; Cradick, T. J.; Rennert, P.; Salinas, P.; Boyd, J.; Amirault, T.; Gray, G. S., Defining critical residues in the epitope for a HIV-neutralizing monoclonal antibody using phage display and peptide array technologies. *Gene* **1993**, *137* (1), 63-8.
73. Scala, G.; Chen, X.; Liu, W.; Telles, J. N.; Cohen, O. J.; Vaccarezza, M.; Igarashi, T.; Fauci, A. S., Selection of HIV-Specific Immunogenic Epitopes by Screening Random Peptide Libraries with HIV-1-Positive Sera. *J. Immunol.* **1999**, *162* (10), 6155-6161.
74. Palacios-Rodriguez, Y.; Gazarian, T.; Rowley, M.; Majluf-Cruz, A.; Gazarian, K., Collection of phage-peptide probes for HIV-1 immunodominant loop-epitope. *J. Microbiol. Methods* **2007**, *68* (2), 225-35.
75. de Haard, J. J.; Kazemier, B.; Oudshoorn, P.; Boender, P.; van Gemen, B.; Koolen, M. J.; van der Groen, G.; Hoogenboom, H. R.; Arends, J. W., Selection of human anti-human immunodeficiency virus type 1 envelope single-chain antibodies from a peripheral blood cell-based phage repertoire. *J. Gen. Virol.* **1998**, *79* (12), 2883-2894.
76. Sticht, J.; Humbert, M.; Findlow, S.; Bodem, J.; Muller, B.; Dietrich, U.; Werner, J.; Krausslich, H. G., A peptide inhibitor of HIV-1 assembly in vitro. *Nat. Struct. Mol. Biol.* **2005**, *12* (8), 671-7.
77. Ferrer, M.; Harrison, S. C., Peptide ligands to human immunodeficiency virus type 1 gp120 identified from phage display libraries. *J. Virol.* **1999**, *73* (7), 5795-802.
78. Boots, L. J.; McKenna, P. M.; Arnold, B. A.; Keller, P. M.; Gorny, M. K.; Zolla-Pazner, S.; Robinson, J. E.; Conley, A. J., Anti-human immunodeficiency virus type 1 human monoclonal antibodies that bind discontinuous epitopes in the viral glycoproteins can identify mimotopes from recombinant phage peptide display libraries. *AIDS Res. Hum. Retroviruses* **1997**, *13* (18), 1549-59.
79. Dervillez, X.; Klaukien, V.; Durr, R.; Koch, J.; Kreutz, A.; Haarmann, T.; Stoll, M.; Lee, D.; Carlomagno, T.; Schnierle, B.; Mobius, K.; Konigs, C.; Griesinger, C.; Dietrich, U., Peptide ligands selected with CD4-induced epitopes on native dualtropic HIV-1 envelope proteins mimic extracellular coreceptor domains and bind to HIV-1 gp120 independently of coreceptor usage. *J. Virol.* **2010**, *84* (19), 10131-8.
80. Zhai, W.; Davies, J.; Shang, D. Z.; Chan, S. W.; Allain, J. P., Human recombinant single-chain antibody fragments, specific for the hypervariable region 1 of hepatitis C virus, from immune phage-display libraries. *J. Viral Hepat.* **1999**, *6* (2), 115-24.
81. Chan, C. E.; Zhao, B. Z.; Cazenave-Gassiot, A.; Pang, S. W.; Bendt, A. K.; Wenk, M. R.; MacAry, P. A.; Hanson, B. J., Novel phage display-derived mycolic acid-

- specific antibodies with potential for tuberculosis diagnosis. *J. Lipid Res.* **2013**, *54* (10), 2924-32.
82. Hu, Z.-Q.; Li, H.-P.; Zhang, J.-B.; Huang, T.; Liu, J.-L.; Xue, S.; Wu, A.-B.; Liao, Y.-C., A phage-displayed chicken single-chain antibody fused to alkaline phosphatase detects *Fusarium* pathogens and their presence in cereal grains. *Anal. Chim. Acta* **2013**, *764*, 84-92.
83. He, Y.; Wang, Y.; Struble, E. B.; Zhang, P.; Chowdhury, S.; Reed, J. L.; Kennedy, M.; Scott, D. E.; Fisher, R. W., Epitope mapping by random peptide phage display reveals essential residues for vaccinia extracellular enveloped virion spread. *Viol. J.* **2012**, *9*, 217.
84. Walper, S. A.; Liu, J. L.; Zabetakis, D.; Anderson, G. P.; Goldman, E. R., Development and Evaluation of Single Domain Antibodies for Vaccinia and the L1 Antigen. *PLoS One* **2014**, *9* (9), e106263.
85. Miller, L.; Michel, J.; Vogt, G.; Dollinger, J.; Stern, D.; Piesker, J.; Nitsche, A., Identification and characterization of a phage display-derived peptide for orthopoxvirus detection. *Anal. Bioanal. Chem.* **2014**, *406* (29), 7611-21.
86. Wu, D.; Li, G.; Qin, C.; Ren, X., Phage displayed peptides to avian H5N1 virus distinguished the virus from other viruses. *PLoS One* **2011**, *6* (8), e23058.
87. Rajik, M.; Jahanshiri, F.; Omar, A. R.; Ideris, A.; Hassan, S. S.; Yusoff, K., Identification and characterisation of a novel anti-viral peptide against avian influenza virus H9N2. *Viol. J.* **2009**, *6*, 74.
88. Luo, W.; Chen, Y.; Wang, M.; Chen, Y.; Zheng, Z.; Song, H.; Chen, H.; Guan, Y.; Ng, Mun H.; Zhang, J.; Xia, N., Peptide mimics of a conserved H5N1 avian influenza virus neutralization site. *Biochem. J* **2009**, *419* (1), 133-139.
89. Yuan, Q.; Jordan, R.; Brlansky, R. H.; Istomina, O.; Hartung, J., Development of single chain variable fragment (scFv) antibodies against *Xylella fastidiosa* subsp. *pauca* by phage display. *J. Microbiol. Methods* **2015**, *117*, 148-54.
90. Feliciano, N. D.; Ribeiro Vda, S.; Santos Fde, A.; Fujimura, P. T.; Gonzaga, H. T.; Goulart, L. R.; Costa-Cruz, J. M., Bacteriophage-fused peptides for serodiagnosis of human strongyloidiasis. *PLoS Negl. Trop. Dis.* **2014**, *8* (5), e2792.
91. Kouzmitcheva, G. A.; Petrenko, V. A.; Smith, G. P., Identifying diagnostic peptides for lyme disease through epitope discovery. *Clin. Diagn. Lab. Immunol.* **2001**, *8* (1), 150-60.
92. Rogers, J. D.; Ajami, N. J.; Fryszczyn, B. G.; Estes, M. K.; Atmar, R. L.; Palzkill, T., Identification and characterization of a peptide affinity reagent for detection of noroviruses in clinical samples. *J. Clin. Microbiol.* **2013**, *51* (6), 1803-8.
93. Chan, S. W.; Bye, J. M.; Jackson, P.; Allain, J. P., Human recombinant antibodies specific for hepatitis C virus core and envelope E2 peptides from an immune phage display library. *J. Gen. Virol.* **1996**, *77* (Pt 10), 2531-9.
94. Lu, X.; Yao, M.; Zhang, J. M.; Yang, J.; Lei, Y. F.; Huang, X. J.; Jia, Z. S.; Ma, L.; Lan, H. Y.; Xu, Z. K.; Yin, W., Identification of peptides that bind hepatitis C virus envelope protein E2 and inhibit viral cellular entry from a phage-display peptide library. *Int. J. Mol. Med.* **2014**, *33* (5), 1312-8.
95. Zhang, X. X.; Deng, Q.; Zhang, S. Y.; Liu, J.; Cai, Q.; Lu, Z. M.; Wang, Y., Broadly cross-reactive mimotope of hypervariable region 1 of hepatitis C virus derived

- from DNA shuffling and screened by phage display library. *J. Med. Virol.* **2003**, *71* (4), 511-7.
96. Tarr, A. W.; Owsianka, A. M.; Timms, J. M.; McClure, C. P.; Brown, R. J.; Hickling, T. P.; Pietschmann, T.; Bartenschlager, R.; Patel, A. H.; Ball, J. K., Characterization of the hepatitis C virus E2 epitope defined by the broadly neutralizing monoclonal antibody AP33. *Hepatology* **2006**, *43* (3), 592-601.
97. Prezzi, C.; Nuzzo, M.; Meola, A.; Delmastro, P.; Galfre, G.; Cortese, R.; Nicosia, A.; Monaci, P., Selection of antigenic and immunogenic mimics of hepatitis C virus using sera from patients. *J. Immunol.* **1996**, *156* (11), 4504-13.
98. Ferrieu-Weisbuch, C.; Bettsworth, F.; Becquart, L.; Paranhos-Baccala, G.; Michel, S.; Arnaud, M.; Jolivet-Reynaud, C., Usefulness of the phage display technology for the identification of a hepatitis C virus NS4A epitope recognized early in the course of the disease. *J. Virol. Methods* **2006**, *131* (2), 175-83.
99. Hong, H. W.; Lee, S. W.; Myung, H., Selection of peptides binding to HCV e2 and inhibiting viral infectivity. *J. Microbiol. Biotechnol.* **2010**, *20* (12), 1769-71.
100. Grihalde, N. D.; Chen, Y. C.; Golden, A.; Gubbins, E.; Mandecki, W., Epitope mapping of anti-HIV and anti-HCV monoclonal antibodies and characterization of epitope mimics using a filamentous phage peptide library. *Gene* **1995**, *166* (2), 187-95.
101. Brigati, J.; Williams, D. D.; Sorokulova, I. B.; Nanduri, V.; Chen, I. H.; Turnbough, C. L., Jr.; Petrenko, V. A., Diagnostic probes for Bacillus anthracis spores selected from a landscape phage library. *Clin. Chem.* **2004**, *50* (10), 1899-906.
102. Williams, D. D.; Benedek, O.; Turnbough, C. L., Species-Specific Peptide Ligands for the Detection of Bacillus anthracis Spores. *Appl. Environ. Microbiol.* **2003**, *69* (10), 6288-6293.
103. Knurr, J.; Benedek, O.; Heslop, J.; Vinson, R. B.; Boydston, J. A.; McAndrew, J.; Kearney, J. F.; Turnbough, C. L., Jr., Peptide ligands that bind selectively to spores of Bacillus subtilis and closely related species. *Appl. Environ. Microbiol.* **2003**, *69* (11), 6841-7.
104. Sainath Rao, S.; Mohan, K. V.; Nguyen, N.; Abraham, B.; Abdouleva, G.; Zhang, P.; Atreya, C. D., Peptides panned from a phage-displayed random peptide library are useful for the detection of Bacillus anthracis surrogates B. cereus 4342 and B. anthracis Sterne. *Biochem. Biophys. Res. Commun.* **2010**, *395* (1), 93-8.
105. Mourez, M.; Kane, R. S.; Mogridge, J.; Metallo, S.; Deschatelets, P.; Sellman, B. R.; Whitesides, G. M.; Collier, R. J., Designing a polyvalent inhibitor of anthrax toxin. *Nat. Biotechnol.* **2001**, *19* (10), 958-61.
106. Walper, S. A.; Anderson, G. P.; Brozozog Lee, P. A.; Glaven, R. H.; Liu, J. L.; Bernstein, R. D.; Zabetakis, D.; Johnson, L.; Czarnecki, J. M.; Goldman, E. R., Rugged Single Domain Antibody Detection Elements for Bacillus anthracis Spores and Vegetative Cells. *PLoS One* **2012**, *7* (3), e32801.
107. Gough, K. C.; Cockburn, W.; Whitlam, G. C., Selection of phage-display peptides that bind to cucumber mosaic virus coat protein. *J. Virol. Methods* **1999**, *79* (2), 169-80.
108. He, X.; Liu, S.; Perry, K. L., Identification of epitopes in cucumber mosaic virus using a phage-displayed random peptide library. *J. Gen. Virol.* **1998**, *79* (Pt 12), 3145-53.

109. Sharma, A.; Saha, A.; Bhattacharjee, S.; Majumdar, S.; Das Gupta, S. K., Specific and Randomly Derived Immunoactive Peptide Mimotopes of Mycobacterial Antigens. *Clin. Vaccine Immunol.* **2006**, *13* (10), 1143-1154.
110. Casey, J. L.; Sanalla, A. M.; Tamvakis, D.; Thalmann, C.; Carroll, E. L.; Parisi, K.; Coley, A. M.; Stewart, D. J.; Vaughan, J. A.; Michalski, W. P.; Luke, R.; Foley, M., Peptides specific for Mycobacterium avium subspecies paratuberculosis infection: diagnostic potential. *Protein Eng. Des. Sel.* **2011**, *24* (8), 589-96.
111. Hell, R. C.; Amim, P.; de Andrade, H. M.; de Avila, R. A.; Felicori, L.; Oliveira, A. G.; Oliveira, C. A.; Nascimento, E.; Tavares, C. A.; Granier, C.; Chavez-Olortegui, C., Immunodiagnosis of human neurocysticercosis using a synthetic peptide selected by phage-display. *Clin. Immunol.* **2009**, *131* (1), 129-38.
112. Gazarian, K.; Rowlay, M.; Gazarian, T.; Vazquez Buchelli, J. E.; Hernandez Gonzales, M., Mimotope peptides selected from phage display combinatorial library by serum antibodies of pigs experimentally infected with Taenia solium as leads to developing diagnostic antigens for human neurocysticercosis. *Peptides* **2012**, *38* (2), 381-8.
113. Manoutcharian, K.; Sotelo, J.; Garcia, E.; Cano, A.; Gevorkian, G., Characterization of cerebrospinal fluid antibody specificities in neurocysticercosis using phage display peptide library. *Clin. Immunol.* **1999**, *91* (1), 117-21.
114. Iniguez, P.; Zientara, S.; Marault, M.; Machin, I. B.; Hannant, D.; Cruciere, C., Screening of horse polyclonal antibodies with a random peptide library displayed on phage: identification of ligands used as antigens in an ELISA test to detect the presence of antibodies to equine arteritis virus. *J. Virol. Methods* **1998**, *73* (2), 175-83.
115. Park, H. Y.; Park, H. C.; Yoon, M. Y., Screening for peptides binding on Phytophthora capsici extracts by phage display. *J. Microbiol. Methods* **2009**, *78* (1), 54-8.
116. Rao, S. S.; Mohan, K. V. K.; Gao, Y.; Atreya, C. D., Identification and evaluation of a novel peptide binding to the cell surface of Staphylococcus aureus. *Microbiol. Res.* **2013**, *168* (2), 106-112.
117. Yang, G.; Gao, Y.; Dong, J.; Liu, C.; Xue, Y.; Fan, M.; Shen, B.; Shao, N., A Novel Peptide Screened by Phage Display Can Mimic TRAP Antigen Epitope against Staphylococcus aureus Infections. *J. Biol. Chem.* **2005**, *280* (29), 27431-27435.
118. Santos, E. M.; Cardoso, R.; Souza, G. R.; Goulart, L. R.; Heinemann, M. B.; Leite, R. C.; Reis, J. K., Selection of peptides for serological detection of equine infectious anemia. *Genet. Mol. Res.* **2012**, *11* (3), 2182-99.
119. Soykut, E. A.; Dudak, F. C.; Boyacı, İ. H., Selection of staphylococcal enterotoxin B (SEB)-binding peptide using phage display technology. *Biochem. Biophys. Res. Commun.* **2008**, *370* (1), 104-108.
120. Urushibata, Y.; Itoh, K.; Ohshima, M.; Seto, Y., Generation of Fab Fragment-Like Molecular Recognition Proteins against Staphylococcal Enterotoxin B by Phage Display Technology. *Clinical and Vaccine Immunology : CVI* **2010**, *17* (11), 1708-1717.
121. Sainath Rao, S.; Mohan, K. V. K.; Atreya, C. D., A Peptide Derived from Phage Display Library Exhibits Antibacterial Activity against E. coli and Pseudomonas aeruginosa. *PLoS One* **2013**, *8* (2), e56081.
122. Lu, X.; Weiss, P.; Block, T., A phage with high affinity for hepatitis B surface antigen for the detection of HBsAg. *J. Virol. Methods* **2004**, *119* (1), 51-54.

123. Deng, Q.; Zhai, J. W.; Michel, M. L.; Zhang, J.; Qin, J.; Kong, Y. Y.; Zhang, X. X.; Budkowska, A.; Tiollais, P.; Wang, Y.; Xie, Y. H., Identification and characterization of peptides that interact with hepatitis B virus via the putative receptor binding site. *J. Virol.* **2007**, *81* (8), 4244-54.
124. Deng, Q.; Zhuang, M.; Kong, Y.-Y.; Xie, Y.-H.; Wang, Y., Screening for PreS specific binding ligands with a phage displayed peptides library. *World Journal of Gastroenterology : WJG* **2005**, *11* (26), 4018-4023.
125. Wang, W.; Liu, Y.; Zu, X.; Jin, R.; Xiao, G., Blocking peptides against HBV: preS1 protein selected from a phage display library. *Biochem. Biophys. Res. Commun.* **2011**, *412* (4), 633-7.
126. D'Mello, F.; Partidos, C. D.; Steward, M. W.; Howard, C. R., Definition of the primary structure of hepatitis B virus (HBV) pre-S hepatocyte binding domain using random peptide libraries. *Virology* **1997**, *237* (2), 319-26.
127. Germaschewski, V.; Murray, K., Identification of polyclonal serum specificities with phage-display libraries. *J. Virol. Methods* **1996**, *58* (1-2), 21-32.
128. Folgori, A.; Tafi, R.; Meola, A.; Felici, F.; Galfre, G.; Cortese, R.; Monaci, P.; Nicosia, A., A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *EMBO J.* **1994**, *13* (9), 2236-43.
129. Chen, Y. C.; Delbrook, K.; Dealwis, C.; Mimms, L.; Mushahwar, I. K.; Mandecki, W., Discontinuous epitopes of hepatitis B surface antigen derived from a filamentous phage peptide library. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (5), 1997-2001.
130. Ho, K. L.; Yusoff, K.; Seow, H. F.; Tan, W. S., Selection of high affinity ligands to hepatitis B core antigen from a phage-displayed cyclic peptide library. *J. Med. Virol.* **2003**, *69* (1), 27-32.
131. Dyson, M. R.; Murray, K., Selection of peptide inhibitors of interactions involved in complex protein assemblies: association of the core and surface antigens of hepatitis B virus. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (6), 2194-8.
132. Tan, W. S.; Tan, G. H.; Yusoff, K.; Seow, H. F., A phage-displayed cyclic peptide that interacts tightly with the immunodominant region of hepatitis B surface antigen. *J. Clin. Virol.* **2005**, *34* (1), 35-41.
133. Pini, A.; Spreafico, A.; Botti, R.; Neri, D.; Neri, P., Hierarchical affinity maturation of a phage library derived antibody for the selective removal of cytomegalovirus from plasma. *J. Immunol. Methods* **1997**, *206* (1-2), 171-182.
134. Williamson, R. A.; Lazzarotto, T.; Sanna, P. P.; Bastidas, R. B.; Dalla Casa, B.; Campisi, G.; Burioni, R.; Landini, M. P.; Burton, D. R., Use of recombinant human antibody fragments for detection of cytomegalovirus antigenemia. *J. Clin. Microbiol.* **1997**, *35* (8), 2047-2050.
135. Srivastava, N.; Zeiler, J. L.; Smithson, S. L.; Carlone, G. M.; Ades, E. W.; Sampson, J. S.; Johnson, S. E.; Kieber-Emmons, T.; Westerink, M. A. J., Selection of an Immunogenic and Protective Epitope of the PsaA Protein of *Streptococcus pneumoniae* Using a Phage Display Library. *Hybridoma* **2000**, *19* (1), 23-31.
136. Susi, P.; Ziegler, A.; Torrance, L., Selection of Single-Chain Variable Fragment Antibodies to Black Currant Reversion Associated Virus from a Synthetic Phage Display Library. *Phytopathology* **1998**, *88* (3), 230-233.

137. Van Nieuwenhove, L. C.; Rogé, S.; Balharbi, F.; Dieltjens, T.; Laurent, T.; Guisez, Y.; Büscher, P.; Lejon, V., Identification of Peptide Mimotopes of *Trypanosoma brucei* gambiense Variant Surface Glycoproteins. *PLoS Negl. Trop. Dis.* **2011**, *5* (6), e1189.
138. Tsui, P.; Tornetta, M. A.; Ames, R. S.; Silverman, C.; Porter, T.; Weston, C.; Griego, S.; Sweet, R. W., Progressive epitope-blocked panning of a phage library for isolation of human RSV antibodies. *J. Immunol. Methods* **2002**, *263* (1–2), 123-132.
139. Casey, J. L.; Coley, A. M.; Street, G.; Parisi, K.; Devine, P. L.; Foley, M., Peptide mimotopes selected from a random peptide library for diagnosis of Epstein-Barr virus infection. *J. Clin. Microbiol.* **2006**, *44* (3), 764-71.
140. Morton, J.; Karoonuthaisiri, N.; Charlermroj, R.; Stewart, L. D.; Elliott, C. T.; Grant, I. R., Phage Display-Derived Binders Able to Distinguish *Listeria monocytogenes* from Other *Listeria* Species. *PLoS One* **2013**, *8* (9), e74312.
141. Kirsch, M. I.; Hülseweh, B.; Nacke, C.; Rülker, T.; Schirrmann, T.; Marschall, H.-J.; Hust, M.; Dübel, S., Development of human antibody fragments using antibody phage display for the detection and diagnosis of Venezuelan equine encephalitis virus (VEEV). *BMC Biotechnol.* **2008**, *8*, 66-66.
142. Lavilla, M.; De Luis, R.; Pérez, M. D.; Calvo, M.; Sánchez, L., Selection of high affine peptide ligands for detection of *Clostridium Tyrobutyricum* spores. *J. Microbiol. Methods* **2009**, *79* (2), 214-219.
143. van Wyngaardt, W.; Mashau, C.; Wright, I.; Fehrsen, J., Serotype- and serogroup-specific detection of African horsesickness virus using phage displayed chicken scFvs for indirect double antibody sandwich ELISAs. *J. Vet. Sci.* **2013**, *14* (1), 95-98.
144. Meissner, F.; Maruyama, T.; Frentsch, M.; Hessel, A. J.; Rodriguez, L. L.; Geisbert, T. W.; Jahrling, P. B.; Burton, D. R.; Parren, P. W., Detection of antibodies against the four subtypes of ebola virus in sera from any species using a novel antibody-phage indicator assay. *Virology* **2002**, *300* (2), 236-43.
145. Goodchild, S. A.; Dooley, H.; Schoepp, R. J.; Flajnik, M.; Lonsdale, S. G., Isolation and characterisation of Ebolavirus-specific recombinant antibody fragments from murine and shark immune libraries. *Mol. Immunol.* **2011**, *48* (15–16), 2027-2037.
146. Wang, X.; Li, G.; Ren, Y.; Ren, X., Phages Bearing Affinity Peptides to Bovine Rotavirus Differentiate the Virus from Other Viruses. *PLoS One* **2011**, *6* (12), e28667.
147. Tungtrakanpoung, R.; Pitaksajakul, P.; Na-Ngarm, N.; Chaicumpa, W.; Ekpo, P.; Saengjaruk, P.; Froman, G.; Ramasoota, P., Mimotope of *Leptospira* from phage-displayed random peptide library is reactive with both monoclonal antibodies and patients' sera. *Vet. Microbiol.* **2006**, *115* (1-3), 54-63.
148. Goldman, E. R.; Pazirandeh, M. P.; Charles, P. T.; Balighian, E. D.; Anderson, G. P., Selection of phage displayed peptides for the detection of 2,4,6-trinitrotoluene in seawater. *Anal. Chim. Acta* **2002**, *457* (1), 13-19.
149. Barelle, C.; Gill, D. S.; Charlton, K., Shark novel antigen receptors--the next generation of biologic therapeutics? *Adv. Exp. Med. Biol.* **2009**, *655*, 49-62.
150. Goldman, E. R.; Anderson, G. P.; Conway, J.; Sherwood, L. J.; Fech, M.; Vo, B.; Liu, J. L.; Hayhurst, A., Thermostable llama single domain antibodies for detection of botulinum A neurotoxin complex. *Anal. Chem.* **2008**, *80* (22), 8583-91.

151. Ngubane, N. A. C.; Gresh, L.; Ioerger, T. R.; Sacchettini, J. C.; Zhang, Y. J.; Rubin, E. J.; Pym, A.; Khati, M., High-Throughput Sequencing Enhanced Phage Display Identifies Peptides That Bind Mycobacteria. *PLoS One* **2013**, *8* (11), e77844.
152. Sorokulova, I. B.; Olsen, E. V.; Chen, I. H.; Fiebor, B.; Barbaree, J. M.; Vodyanoy, V. J.; Chin, B. A.; Petrenko, V. A., Landscape phage probes for *Salmonella typhimurium*. *J. Microbiol. Methods* **2005**, *63* (1), 55-72.
153. Cudic, M.; Condie, B. A.; Weiner, D. J.; Lysenko, E. S.; Xiang, Z. Q.; Insug, O.; Bulet, P.; Otvos, L., Jr., Development of novel antibacterial peptides that kill resistant isolates. *Peptides* **2002**, *23* (12), 2071-83.
154. Lakshmanan, R. S.; Guntupalli, R.; Hu, J.; Petrenko, V. A.; Barbaree, J. M.; Chin, B. A., Detection of *Salmonella typhimurium* in fat free milk using a phage immobilized magnetoelastic sensor. *Sensors Actuators B: Chem.* **2007**, *126* (2), 544-550.
155. Strachan, G.; Whyte, J. A.; Molloy, P. M.; Paton, G. I.; Porter, Development of Robust, Environmental, Immunoassay Formats for the Quantification of Pesticides in Soil. *Environ. Sci. Technol.* **2000**, *34* (8), 1603-1608.
156. Block, T.; Miller, R.; Korngold, R.; Jungkind, D. A., Phage-linked immunoabsorbent system for the detection of pathologically relevant pathogens. *Biotechniques* **1989**, (7), 756-761.
157. Mammen, M.; Choi, S.-K.; Whitesides, G. M., Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* **1998**, *37* (20), 2754-2794.
158. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R., Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. *Nature* **2000**, *403* (6770), 669-72.
159. Branson, T. R.; Turnbull, W. B., Bacterial toxin inhibitors based on multivalent scaffolds. *Chem. Soc. Rev.* **2013**, *42* (11), 4613-22.
160. Bernardi, A.; Jimenez-Barbero, J.; Casnati, A.; De Castro, C.; Darbre, T.; Fieschi, F.; Finne, J.; Funken, H.; Jaeger, K.-E.; Lahmann, M.; Lindhorst, T. K.; Marradi, M.; Messner, P.; Molinaro, A.; Murphy, P. V.; Nativi, C.; Oscarson, S.; Penades, S.; Peri, F.; Pieters, R. J.; Renaudet, O.; Reymond, J.-L.; Richichi, B.; Rojo, J.; Sansone, F.; Schaffer, C.; Turnbull, W. B.; Velasco-Torrijos, T.; Vidal, S.; Vincent, S.; Wennekes, T.; Zuilhof, H.; Imberty, A., Multivalent glycoconjugates as anti-pathogenic agents. *Chem. Soc. Rev.* **2013**, *42* (11), 4709-4727.
161. Jain, D.; Kaur, K.; Sundaravadivel, B.; Salunke, D. M., Structural and Functional Consequences of Peptide-Carbohydrate Mimicry: Crystal Structure Of A Carbohydrate-Mimicking Peptide Bound To Concanavalin A. *J. Biol. Chem.* **2000**, *275* (21), 16098-16102.
162. Scott, J. K., Discovering peptide ligands using epitope libraries. *Trends Biochem. Sci* **1992**, *17* (7), 241-5.
163. Folgori, A.; Tafi, R.; Meola, A.; Felici, F.; Galfré, G.; Cortese, R.; Monaci, P.; Nicosia, A., A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *The EMBO Journal* **1994**, *13* (9), 2236-2243.
164. Motti, C.; Nuzzo, M.; Meola, A.; Galfré, G.; Felici, F.; Cortese, R.; Nicosia, A.; Monaci, P., Recognition by human sera and immunogenicity of HBsAg mimotopes selected from an M13 phage display library. *Gene* **1994**, *146* (2), 191-198.

165. Xu, G. J.; Kula, T.; Xu, Q.; Li, M. Z.; Vernon, S. D.; Ndung'u, T.; Ruxrungtham, K.; Sanchez, J.; Brander, C.; Chung, R. T.; O'Connor, K. C.; Walker, B.; Larman, H. B.; Elledge, S. J., Comprehensive serological profiling of human populations using a synthetic human virome. *Science* **2015**, *348* (6239).
166. Doran, T. M.; Gao, Y.; Mendes, K.; Dean, S.; Simanski, S.; Kodadek, T., Utility of Redundant Combinatorial Libraries in Distinguishing High and Low Quality Screening Hits. *ACS. Comb. Sci.* **2014**, *16* (6), 259-270.
167. Derda, R.; Tang, S. K.; Li, S. C.; Ng, S.; Matochko, W.; Jafari, M. R., Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules* **2011**, *16* (2), 1776-803.
168. Reddy, M. M.; Wilson, R.; Wilson, J.; Connell, S.; Gocke, A.; Hynan, L.; German, D.; Kodadek, T., Identification of Candidate IgG Biomarkers for Alzheimer's Disease via Combinatorial Library Screening. *Cell* **2011**, *144* (1), 132-142.
169. Wu, C.-H.; Liu, I. J.; Lu, R.-M.; Wu, H.-C., Advancement and applications of peptide phage display technology in biomedical science. *J. Biomed. Sci.* **2016**, *23*, 8.
170. Matsubara, T., Potential of Peptides as Inhibitors and Mimotopes: Selection of Carbohydrate-Mimetic Peptides from Phage Display Libraries. *J. Nucleic Acids* **2012**, *2012*, 15.
171. Fukuda, M. N., Peptide-displaying phage technology in glycobiology. *Glycobiology* **2012**, *22* (3), 318-25.
172. Ivarsson, Y.; Arnold, R.; McLaughlin, M.; Nim, S.; Joshi, R.; Ray, D.; Liu, B.; Teyra, J.; Pawson, T.; Moffat, J.; Li, S. S.; Sidhu, S. S.; Kim, P. M., Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (7), 2542-7.
173. Freckleton, G.; Lippman, S. I.; Broach, J. R.; Tavazoie, S., Microarray Profiling of Phage-Display Selections for Rapid Mapping of Transcription Factor–DNA Interactions. *PLoS Genet.* **2009**, *5* (4), e1000449.
174. Larman, H. B.; Zhao, Z.; Laserson, U.; Li, M. Z.; Ciccina, A.; Gakidis, M. A.; Church, G. M.; Kesari, S.; Leproust, E. M.; Solimini, N. L.; Elledge, S. J., Autoantigen discovery with a synthetic human peptidome. *Nat. Biotechnol.* **2011**, *29* (6), 535-41.
175. Larman, H. B.; Laserson, U.; Querol, L.; Verhaeghen, K.; Solimini, N. L.; Xu, G. J.; Klarenbeek, P. L.; Church, G. M.; Hafler, D. A.; Plenge, R. M.; Nigrovic, P. A.; De Jager, P. L.; Weets, I.; Martens, G. A.; O'Connor, K. C.; Elledge, S. J., PhIP-Seq characterization of autoantibodies from patients with multiple sclerosis, type 1 diabetes and rheumatoid arthritis. *J. Autoimmun.* **2013**, *43*, 1-9.
176. Ng, S.; Tjhung, K.; Paschal, B.; Noren, C.; Derda, R., Chemical Posttranslational Modification of Phage-Displayed Peptides. In *Peptide Libraries*, Derda, R., Ed. Springer New York: 2015; Vol. 1248, pp 155-172.
177. Dwyer, M. A.; Lu, W.; Dwyer, J. J.; Kossiakoff, A. A., Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol.* **2000**, *7* (4), 263-74.
178. Ng, S.; Derda, R., Phage-displayed macrocyclic glycopeptide libraries. *Org. Biomol. Chem.* **2016**, *14* (24), 5539-5545.
179. Kalhor-Monfared, S.; Jafari, M. R.; Patterson, J. T.; Kitov, P. I.; Dwyer, J. J.; Nuss, J. M.; Derda, R., Rapid biocompatible macrocyclization of peptides with decafluoro-diphenylsulfone. *Chemical Science* **2016**.

180. Ng, S.; Jafari, M. R.; Matochko, W. L.; Derda, R., Quantitative Synthesis of Genetically Encoded Glycopeptide Libraries Displayed on M13 Phage. *ACS Chem. Biol.* **2012**, *7* (9), 1482-1487.
181. Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R., Silent Encoding of Chemical Post-Translational Modifications in Phage-Displayed Libraries. *J. Am. Chem. Soc.* **2016**, *138* (1), 32-35.
182. Liu, C. C.; Mack, A. V.; Brustad, E. M.; Mills, J. H.; Groff, D.; Smider, V. V.; Schultz, P. G., Evolution of Proteins with Genetically Encoded "Chemical Warheads". *J. Am. Chem. Soc.* **2009**, *131* (28), 9616-9617.
183. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108-60.
184. Robinson, W. H.; Steinman, L., Human peptidome display. *Nat. Biotechnol.* **2011**, *29* (6), 500-502.
185. Matochko, W. L.; Cory Li, S.; Tang, S. K. Y.; Derda, R., Prospective identification of parasitic sequences in phage display screens. *Nucleic Acids Research* **2014**, *42* (3), 1784-1798.
186. Christiansen, A.; Kringelum, J. V.; Hansen, C. S.; Bogh, K. L.; Sullivan, E.; Patel, J.; Rigby, N. M.; Eiwegger, T.; Szepfalusi, Z.; de Masi, F.; Nielsen, M.; Lund, O.; Dufva, M., High-throughput sequencing enhanced phage display enables the identification of patient-specific epitope motifs in serum. *Sci. Rep.* **2015**, *5*, 12913.
187. Matochko, W. L.; Cory Li, S.; Tang, S. K. Y.; Derda, R., Prospective identification of parasitic sequences in phage display screens. *Nucleic Acids Res.* **2013**.
188. Rentero Rebollo, I.; Sabisz, M.; Baeriswyl, V.; Heinis, C., Identification of target-binding peptide motifs by high-throughput sequencing of phage-selected peptides. *Nucleic Acids Res.* **2014**.
189. 't Hoen, P. A. C.; Jirka, S. M. G.; ten Broeke, B. R.; Schultes, E. A.; Aguilera, B.; Pang, K. H.; Heemskerk, H.; Aartsma-Rus, A.; van Ommen, G. J.; den Dunnen, J. T., Phage display screening without repetitious selection rounds. *Anal. Biochem.* **2012**, *421* (2), 622-631.
190. Getz, J. A.; Schoep, T. D.; Daugherty, P. S., Chapter four - Peptide Discovery Using Bacterial Display and Flow Cytometry. In *Methods Enzymol.*, Wittrup, K. D.; Gregory, L. V., Eds. Academic Press: 2012; Vol. Volume 503, pp 75-97.
191. Dutta, S.; Gulla, S.; Chen, T. S.; Fire, E.; Grant, R. A.; Keating, A. E., Determinants of BH3 binding specificity for Mcl-1 versus Bcl-xL. *J. Mol. Biol.* **2010**, *398* (5), 747-62.

192. Chen, T. S.; Palacios, H.; Keating, A. E., Structure-based redesign of the binding specificity of anti-apoptotic Bcl-x(L). *J. Mol. Biol.* **2013**, *425* (1), 171-85.
193. Daugherty, P. S., Protein engineering with bacterial display. *Curr. Opin. Struct. Biol.* **2007**, *17* (4), 474-80.
194. Pantazes, R. J.; Reifert, J.; Bozekowski, J.; Ibsen, K. N.; Murray, J. A.; Daugherty, P. S., Identification of disease-specific motifs in the antibody specificity repertoire via next-generation sequencing. *Sci. Rep.* **2016**, *6*, 30312.
195. Bidlingmaier, S.; Liu, B., Construction and application of a yeast surface-displayed human cDNA library to identify post-translational modification-dependent protein-protein interactions. *Mol. Cell. Proteomics* **2006**, *5* (3), 533-40.
196. Bidlingmaier, S.; Liu, B., Identification of Posttranslational Modification-Dependent Protein Interactions Using Yeast Surface Displayed Human Proteome Libraries. In *Yeast Surface Display: Methods, Protocols, and Applications*, Liu, B., Ed. Springer New York: New York, NY, 2015; pp 193-202.
197. Derda, R.; Gitaka, J.; Klapperich, C. M.; Mace, C. R.; Kumar, A. A.; Lieberman, M.; Linnes, J. C.; Jores, J.; Nasimolo, J.; Ndungu, J.; Taracha, E.; Weaver, A.; Weibel, D. B.; Kariuki, T. M.; Yager, P., Enabling the Development and Deployment of Next Generation Point-of-Care Diagnostics. *PLoS Negl. Trop. Dis.* **2015**, *9* (5), e0003676.
198. Kumar, A. A.; Hennek, J. W.; Smith, B. S.; Kumar, S.; Beattie, P.; Jain, S.; Rolland, J. P.; Stossel, T. P.; Chunda-Liyoka, C.; Whitesides, G. M., From the Bench to the Field in Low-Cost Diagnostics: Two Case Studies. *Angew. Chem. Int. Ed.* **2015**, *54* (20), 5836-5853.
199. Martinez, A. W.; Phillips, S. T.; Whitesides, G. M.; Carrilho, E., Diagnostics for the Developing World: Microfluidic Paper-Based Analytical Devices. *Anal. Chem.* **2010**, *82* (1), 3-10.
200. Fu, E.; Yager, P.; Floriano, P. N.; Christodoulides, N.; McDevitt, J., Perspective on diagnostics for global health. *IEEE pulse* **2011**, *2* (6), 40-50.
201. Yager, P.; Domingo, G. J.; Gerdes, J., Point-of-Care Diagnostics for Global Health. *Annu. Rev. Biomed. Eng.* **2008**, *10* (1), 107-144.
202. Veenstra, T. D., *Proteomic Applications in Cancer Detection and Discovery*. Wiley: 2013; p 320.
203. Steingart, K. R.; Flores, L. L.; Dendukuri, N.; Schiller, I.; Laal, S.; Ramsay, A.; Hopewell, P. C.; Pai, M., Commercial Serological Tests for the Diagnosis of Active Pulmonary and Extrapulmonary Tuberculosis: An Updated Systematic Review and Meta-Analysis. *PLoS Med* **2011**, *8* (8), e1001062.
204. Ratnam, S., The laboratory diagnosis of syphilis. *Can. J. Infect. Dis.* **2005**, *16* (1), 45-51.
205. Johnson, B. J. B., *Lyme Disease: an Evidence-based Approach*. 2011.
206. Flynn, J. L.; Chan, J., Immunology of Tuberculosis. *Annu. Rev. Immunol.* **2001**, *19* (1), 93-129.
207. Spencer, J. S.; Kim, H. J.; Wheat, W. H.; Chatterjee, D.; Balagon, M. V.; Cellona, R. V.; Tan, E. V.; Gelber, R.; Saunderson, P.; Duthie, M. S.; Reece, S. T.; Burman, W.; Belknap, R.; Mac Kenzie, W. R.; Geluk, A.; Oskam, L.; Dockrell, H. M.; Brennan, P. J.; Consortium, o. b. o. t. I., Analysis of Antibody Responses to Mycobacterium leprae Phenolic Glycolipid I, Lipoarabinomannan, and Recombinant Proteins To Define Disease

- Subtype-Specific Antigenic Profiles in Leprosy. *Clin. Vaccine Immunol.* **2011**, *18* (2), 260-267.
208. Nishimura, T.; Hasegawa, N.; Fujita, Y.; Yano, I.; Ishizaka, A., Serodiagnostic Contributions of Antibody Titers against Mycobacterial Lipid Antigens in Mycobacterium avium Complex Pulmonary Disease. *Clin. Infect. Dis.* **2009**, *49* (4), 529-535.
209. Scollard, D. M.; Adams, L. B.; Gillis, T. P.; Krahenbuhl, J. L.; Truman, R. W.; Williams, D. L., The Continuing Challenges of Leprosy. *Clin. Microbiol. Rev.* **2006**, *19* (2), 338-381.
210. Ignatov, D.; Kondratieva, E.; Azhikina, T.; Apt, A., Mycobacterium avium-triggered diseases: pathogenomics. *Cell. Microbiol.* **2012**, *14* (6), 808-818.
211. Flores, L. L.; Steingart, K. R.; Dendukuri, N.; Schiller, I.; Minion, J.; Pai, M.; Ramsay, A.; Henry, M.; Laal, S., Systematic Review and Meta-Analysis of Antigen Detection Tests for the Diagnosis of Tuberculosis. *Clin. Vaccine Immunol.* **2011**, *18* (10), 1616-1627.
212. Tong, M.; Jacobi, C. E.; van de Rijke, F. M.; Kuijper, S.; van de Werken, S.; Lowary, T. L.; Hokke, C. H.; Appelmelk, B. J.; Nagelkerke, N. J. D.; Tanke, H. J.; van Gijlswijk, R. P. M.; Veuskens, J.; Kolk, A. H. J.; Raap, A. K., A multiplexed and miniaturized serological tuberculosis assay identifies antigens that discriminate maximally between TB and non-TB sera. *J. Immunol. Methods* **2005**, *301* (1-2), 154-163.
213. Singh, S.; Singh, J.; Kumar, S.; Gopinath, K.; Balooni, V.; Singh, N.; Mani, K., Poor Performance of Serological Tests in the Diagnosis of Pulmonary Tuberculosis: Evidence from a Contact Tracing Field Study. *PLoS One* **2012**, *7* (7), e40213.
214. Imaz, M. S.; Comini, M. A.; Zerbini, E.; Sequeira, M. D.; Latini, O.; Claus, J. D.; Singh, M., Evaluation of Commercial Enzyme-Linked Immunosorbent Assay Kits for Detection of Tuberculosis in Argentinean Population. *J. Clin. Microbiol.* **2004**, *42* (2), 884-887.
215. Steingart, K. R.; Ramsay, A.; Dowdy, D. W.; Pai, M., Serological tests for the diagnosis of active tuberculosis: relevance for India. *Indian J. Med. Res.* **2012**, *135* (5), 695-702.
216. Gevorkian, G.; Segura, E.; Acero, G.; Palma, José P.; Espitia, C.; Manoutcharian, K.; López-Marín, Luz M., Peptide mimotopes of Mycobacterium tuberculosis carbohydrate immunodeterminants. *Biochem. J* **2005**, *387* (Pt 2), 411-417.
217. Barenholz, A.; Hovav, A.-H.; Fishman, Y.; Rahav, G.; Gershoni, J. M.; Bercovier, H., A peptide mimetic of the mycobacterial mannosylated lipoarabinomannan: characterization and potential applications. *J. Med. Microbiol.* **2007**, *56* (5), 579-586.
218. Oldenburg, K. R.; Loganathan, D.; Goldstein, I. J.; Schultz, P. G.; Gallop, M. A., Peptide ligands for a sugar-binding protein isolated from a random peptide library. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5393-5397.
219. Jain, D.; Kaur, K.; Sundaravadivel, B.; Salunke, D. M., Structural and functional consequences of peptide-carbohydrate mimicry. Crystal structure of a carbohydrate-mimicking peptide bound to concanavalin A. *J. Biol. Chem.* **2000**, *275* (21), 16098-102.
220. Goel, M.; Krishnan, L.; Kaur, S.; Kaur, K. J.; Salunke, D. M., Plasticity within the antigen-combining site may manifest as molecular mimicry in the humoral immune response. *J. Immunol.* **2004**, *173* (12), 7358-67.

221. Oyelaran, O.; McShane, L. M.; Dodd, L.; Gildersleeve, J. C., Profiling human serum antibodies with a carbohydrate antigen microarray. *J. Proteome Res.* **2009**, *8* (9), 4301-10.
222. Liang, P.-H.; Wu, C.-Y.; Greenberg, W. A.; Wong, C.-H., Glycan arrays: biological and medical applications. *Curr. Opin. Chem. Biol.* **2008**, *12* (1), 86-92.
223. Oyelaran, O.; Gildersleeve, J. C., Application of carbohydrate array technology to antigen discovery and vaccine development. *Expert Rev Vaccines* **2007**, *6* (6), 957-69.
224. 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.; Lowary, T. L., Ligand Specificity of CS-35, a Monoclonal Antibody That Recognizes Mycobacterial Lipoarabinomannan: A Model System for Oligofuranoside-Protein Recognition. *J. Am. Chem. Soc.* **2007**, *129* (34), 10489-10502.
225. Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K., Structural insights into antibody recognition of mycobacterial polysaccharides. *J Mol Biol* **2009**, *392* (2), 381-92.
226. Nishi, J. S.; Shury, T.; Elkin, B. T., Wildlife reservoirs for bovine tuberculosis (*Mycobacterium bovis*) in Canada: Strategies for management and research. *Vet. Microbiol.* **2006**, *112* (2-4), 325-338.
227. Kitov, P. I.; Vinals, D. F.; Ng, S.; Tjhung, K. F.; Derda, R., Rapid, Hydrolytically Stable Modification of Aldehyde-Terminated Proteins and Phage Libraries. *J. Am. Chem. Soc.* **2014**, *136* (23), 8149-8152.
228. Matochko, W. L.; Chu, K.; Jin, B.; Lee, S. W.; Whitesides, G. M.; Derda, R., Deep sequencing analysis of phage libraries using Illumina platform. *Methods* **2012**, *58* (1), 47-55.
229. Deiss, F.; Matochko, W. L.; Govindasamy, N.; Lin, E. Y.; Derda, R., Flow-Through Synthesis on Teflon-Patterned Paper To Produce Peptide Arrays for Cell-Based Assays. *Angew. Chem. Int. Ed.* **2014**, *53* (25), 6374-6377.
230. Clamp, J. R.; Hough, L., The periodate oxidation of amino acids with reference to studies on glycoproteins. *Biochem. J* **1965**, *94* (1), 17-24.
231. D'Souza, F. W.; Lowary, T. L., The first total synthesis of a highly branched arabinofuranosyl hexasaccharide found at the nonreducing termini of mycobacterial arabinogalactan and lipoarabinomannan. *Org. Lett.* **2000**, *2* (10), 1493-5.
232. Noren, K. A.; Noren, C. J., Construction of High-Complexity Combinatorial Phage Display Peptide Libraries. *Methods* **2001**, *23* (2), 169-178.
233. Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R., Silent Encoding of Chemical Post-Translational Modifications in Phage-Displayed Libraries. *J. Am. Chem. Soc.* **2016**, *138* (1), 32-5.
234. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S., Reliable Determinations of Protein-Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass. Spectrom.* **2012**, *23* (3), 431-441.
235. Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S., Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **2006**, *78* (9), 3010-8.
236. Raman, R.; Tharakaraman, K.; Sasisekharan, V.; Sasisekharan, R., Glycan-protein interactions in viral pathogenesis. *Curr. Opin. Struct. Biol.* **2016**, *40*, 153-162.

237. Weis, W. I.; Drickamer, K.; Hendrickson, W. A., Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **1992**, *360* (6400), 127-34.
238. Kanagawa, M.; Liu, Y.; Hanashima, S.; Ikeda, A.; Chai, W.; Nakano, Y.; Kojima-Aikawa, K.; Feizi, T.; Yamaguchi, Y., Structural basis for multiple sugar recognition of Jacalin-related human ZG16p lectin. *J. Biol. Chem.* **2014**, *289* (24), 16954-65.
239. Naismith, J. H.; Emmerich, C.; Habash, J.; Harrop, S. J.; Helliwell, J. R.; Hunter, W. N.; Raftery, J.; Kalb, A. J.; Yariv, J., Refined structure of concanavalin A complexed with methyl alpha-D-mannopyranoside at 2.0 Å resolution and comparison with the saccharide-free structure. *Acta Crystallogr. D Biol. Crystallogr.* **1994**, *50* (Pt 6), 847-58.
240. Kiessling, L. L.; Pohl, N. L., Strength in numbers: non-natural polyvalent carbohydrate derivatives. *Chem. Biol.* **1996**, *3* (2), 71-77.
241. Deng, S. J.; MacKenzie, C. R.; Hirama, T.; Brousseau, R.; Lowary, T. L.; Young, N. M.; Bundle, D. R.; Narang, S. A., Basis for selection of improved carbohydrate-binding single-chain antibodies from synthetic gene libraries. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (11), 4992-6.
242. Indik, Z. K.; Park, J. G.; Hunter, S.; Schreiber, A. D., The molecular dissection of Fc gamma receptor mediated phagocytosis. *Blood* **1995**, *86* (12), 4389-99.
243. Mammen, M.; Helmerson, K.; Kishore, R.; Choi, S.-K.; Phillips, W. D.; Whitesides, G. M., Optically controlled collisions of biological objects to evaluate potent polyvalent inhibitors of virus-cell adhesion. *Chem. Biol.* **1996**, *3* (9), 757-763.
244. Mowery, P.; Yang, Z. Q.; Gordon, E. J.; Dwir, O.; Spencer, A. G.; Alon, R.; Kiessling, L. L., Synthetic glycoprotein mimics inhibit L-selectin-mediated rolling and promote L-selectin shedding. *Chem. Biol.* **2004**, *11* (5), 725-32.
245. Magnani, J. L.; Patton, J. T.; Sarkar, A. K., Glycomimetic inhibitors of the PA-II lectin, PA-III lectin or both the lectins from pseudomonas. Google Patents: 2012.
246. Ernst, B.; Magnani, J. L., From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.* **2009**, *8* (8), 661-677.
247. Bhatia, S.; Dimde, M.; Haag, R., Multivalent glycoconjugates as vaccines and potential drug candidates. *MedChemComm* **2014**, *5* (7), 862-878.
248. Charles A Janeway, J.; Travers, P.; Walport, M.; Shlomchik, M. J., *Immunobiology, 5th edition*. Garland Science: New York, 2001.
249. Mohan, J. F.; Unanue, E. R., Unconventional recognition of peptides by T cells and the implications for autoimmunity. *Nat. Rev. Immunol.* **2012**, *12* (10), 721-728.
250. Geysen, H. M.; Meloen, R. H.; Barteling, S. J., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (13), 3998-4002.
251. Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G., General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **1991**, *37* (6), 487-93.
252. Houghten, R. A., General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (15), 5131-5.
253. Smith, G. P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228* (4705), 1315-7.

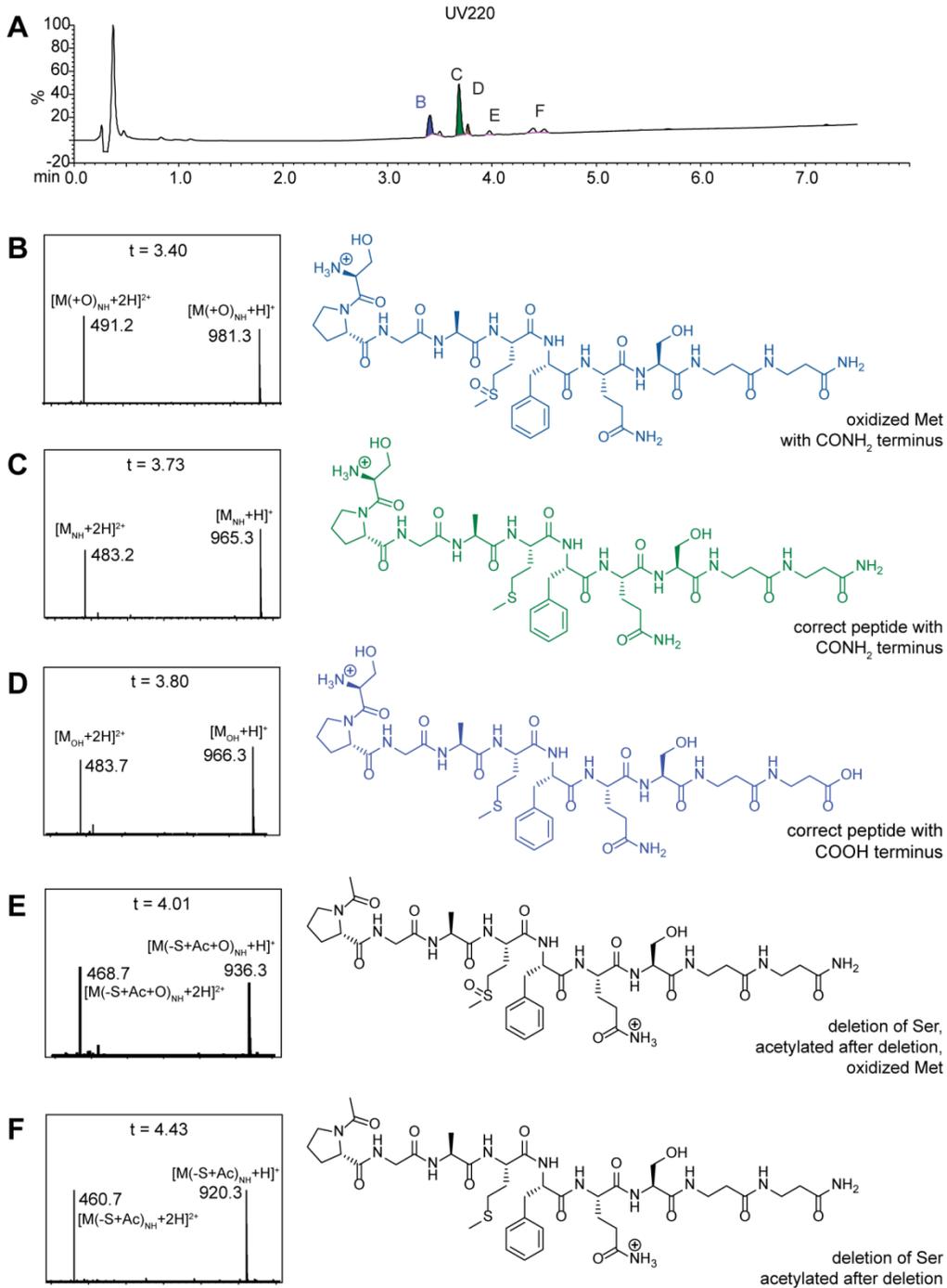
254. Carlos F. Barbas III; Dennis R. Burton; Jamie K. Scott; Silverman, G. J., *Phage Display: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2001.
255. Harris, S. L.; Craig, L.; Mehroke, J. S.; Rashed, M.; Zwick, M. B.; Kenar, K.; Toone, E. J.; Greenspan, N.; Auzanneau, F.-I.; Marino-Albernas, J.-R.; Pinto, B. M.; Scott, J. K., Exploring the basis of peptide–carbohydrate crossreactivity: Evidence for discrimination by peptides between closely related anti-carbohydrate antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (6), 2454-2459.
256. Irving, M. B.; Craig, L.; Menendez, A.; Gangadhar, B. P.; Montero, M.; van Houten, N. E.; Scott, J. K., Exploring Peptide Mimics for the Production of Antibodies Against Discontinuous Protein Epitopes. *Mol. Immunol.* **2010**, *47* (5), 1137.
257. Phalipon, A.; Folgori, A.; Arondel, J.; Sgaramella, G.; Fortugno, P.; Cortese, R.; Sansonetti, P. J.; Felici, F., Induction of anti-carbohydrate antibodies by phage library-selected peptide mimics. *Eur. J. Immunol.* **1997**, *27* (10), 2620-5.
258. Kieber-Emmons, T.; Luo, P.; Qiu, J.; Agadjanyan, M.; Carey, L.; Hutchins, W.; Westerink, M. A.; Stepkowski, Z., Peptide mimicry of adenocarcinoma-associated carbohydrate antigens. *Hybridoma* **1997**, *16* (1), 3-10.
259. Kügler, J.; Zantow, J.; Meyer, T.; Hust, M., Oligopeptide M13 Phage Display in Pathogen Research. *Viruses* **2013**, *5* (10), 2531-2545.
260. Rifai, N.; Gillette, M. A.; Carr, S. A., Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotech* **2006**, *24* (8), 971-983.
261. t Hoen, P. A.; Jirka, S. M.; Ten Broeke, B. R.; Schultes, E. A.; Aguilera, B.; Pang, K. H.; Heemskerk, H.; Aartsma-Rus, A.; van Ommen, G. J.; den Dunnen, J. T., Phage display screening without repetitious selection rounds. *Anal. Biochem.* **2012**, *421* (2), 622-31.
262. Glanville, J.; D'Angelo, S.; Khan, T. A.; Reddy, S. T.; Naranjo, L.; Ferrara, F.; Bradbury, A. R. M., Deep sequencing in library selection projects: what insight does it bring? *Curr. Opin. Struct. Biol.* **2015**, *33*, 146-160.
263. Rentero Rebollo, I.; Sabisz, M.; Baeriswyl, V.; Heinis, C., Identification of target-binding peptide motifs by high-throughput sequencing of phage-selected peptides. *Nucleic Acids Res.* **2014**, *42* (22), e169-e169.
264. Rondas, D.; Crevecoeur, I.; D'Hertog, W.; Ferreira, G. B.; Staes, A.; Garg, A. D.; Eizirik, D. L.; Agostinis, P.; Gevaert, K.; Overbergh, L.; Mathieu, C., Citrullinated glucose-regulated protein 78 is an autoantigen in type 1 diabetes. *Diabetes* **2015**, *64* (2), 573-86.
265. van Venrooij, W. J.; Pruijn, G. J. M., Citrullination: a small change for a protein with great consequences for rheumatoid arthritis. *Arthritis Res.* **2000**, *2* (4), 249-251.
266. Cohen, P., The role of protein phosphorylation in human health and disease. *Eur. J. Biochem.* **2001**, *268* (19), 5001-5010.
267. Pinho, S. S.; Reis, C. A., Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **2015**, *15* (9), 540-555.
268. Somers, V.; Somers, K.; Stinissen, P., Citrullination specific phage display. Google Patents: 2011.
269. Schmitz, R.; Baumann, G.; Gram, H., Catalytic specificity of phosphotyrosine kinases Blk, Lyn, c-Src and Syk as assessed by phage display. *J. Mol. Biol.* **1996**, *260* (5), 664-77.

270. Arai, K.; Tsutsumi, H.; Mihara, H., A monosaccharide-modified peptide phage library for screening of ligands to carbohydrate-binding proteins. *Bioorg. Med. Chem. Lett.* **2013**, *23* (17), 4940-4943.
271. Millward, S. W.; Takahashi, T. T.; Roberts, R. W., A general route for post-translational cyclization of mRNA display libraries. *J. Am. Chem. Soc.* **2005**, *127* (41), 14142-3.
272. Horiya, S.; Bailey, J. K.; Temme, J. S.; Guillen Schlippe, Y. V.; Krauss, I. J., Directed Evolution of Multivalent Glycopeptides Tightly Recognized by HIV Antibody 2G12. *J. Am. Chem. Soc.* **2014**, *136* (14), 5407-5415.
273. Wang, W.; Hirano, Y.; Uzawa, T.; Liu, M.; Taiji, M.; Ito, Y., In vitro selection of a peptide aptamer that potentiates inhibition of cyclin-dependent kinase 2 by purvalanol. *MedChemComm* **2014**, *5* (9), 1400-1403.
274. Wals, K.; Ovaa, H., Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins. *Front Chem* **2014**, *2*, 15.
275. Chin, J. W., Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem* **2014**, *83*, 379-408.
276. Zhang, W. H.; Otting, G.; Jackson, C. J., Protein engineering with unnatural amino acids. *Curr. Opin. Struct. Biol.* **2013**, *23* (4), 581-7.
277. Piscotta, F. J.; Tharp, J. M.; Liu, W. R.; Link, A. J., Expanding the chemical diversity of lasso peptide MccJ25 with genetically encoded noncanonical amino acids. *Chem. Commun.* **2015**, *51* (2), 409-412.
278. Wang, L.; Xie, J.; Schultz, P. G., Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225-49.
279. Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu. Rev. Biochem* **2010**, *79*, 413-44.
280. Sun, S. B.; Schultz, P. G.; Kim, C. H., Therapeutic applications of an expanded genetic code. *ChemBioChem* **2014**, *15* (12), 1721-1729.
281. Liu, C. C.; Mack, A. V.; Tsao, M.-L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V., Protein evolution with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (46), 17688-17693.
282. Terasaka, N.; Suga, H., Flexizymes-facilitated Genetic Code Reprogramming Leading to the Discovery of Drug-like Peptides. *Chem. Lett.* **2014**, *43* (1), 11-19.
283. Heinis, C.; Winter, G., Encoded libraries of chemically modified peptides. *Curr. Opin. Chem. Biol.* **2015**, *26*, 89-98.
284. Egger, J.; Weckerle, C.; Cutting, B.; Schwardt, O.; Rabbani, S.; Lemme, K.; Ernst, B., Nanomolar E-Selectin Antagonists with Prolonged Half-Lives by a Fragment-Based Approach. *J. Am. Chem. Soc.* **2013**, *135* (26), 9820-9828.
285. Calosso, M.; Tambutet, G.; Charpentier, D.; St-Pierre, G.; Vaillancourt, M.; Bencheqroun, M.; Gratton, J.-P.; Prévost, M.; Guindon, Y., Acyclic Tethers Mimicking Subunits of Polysaccharide Ligands: Selectin Antagonists. *ACS Med. Chem. Lett.* **2014**, *5* (9), 1054-1059.
286. Assem, N.; Ferreira, D. J.; Wolan, D. W.; Dawson, P. E., Acetone-Linked Peptides: A Convergent Approach for Peptide Macrocyclization and Labeling. *Angew. Chem. Int. Ed.* **2015**, *54* (30), 8665-8668.

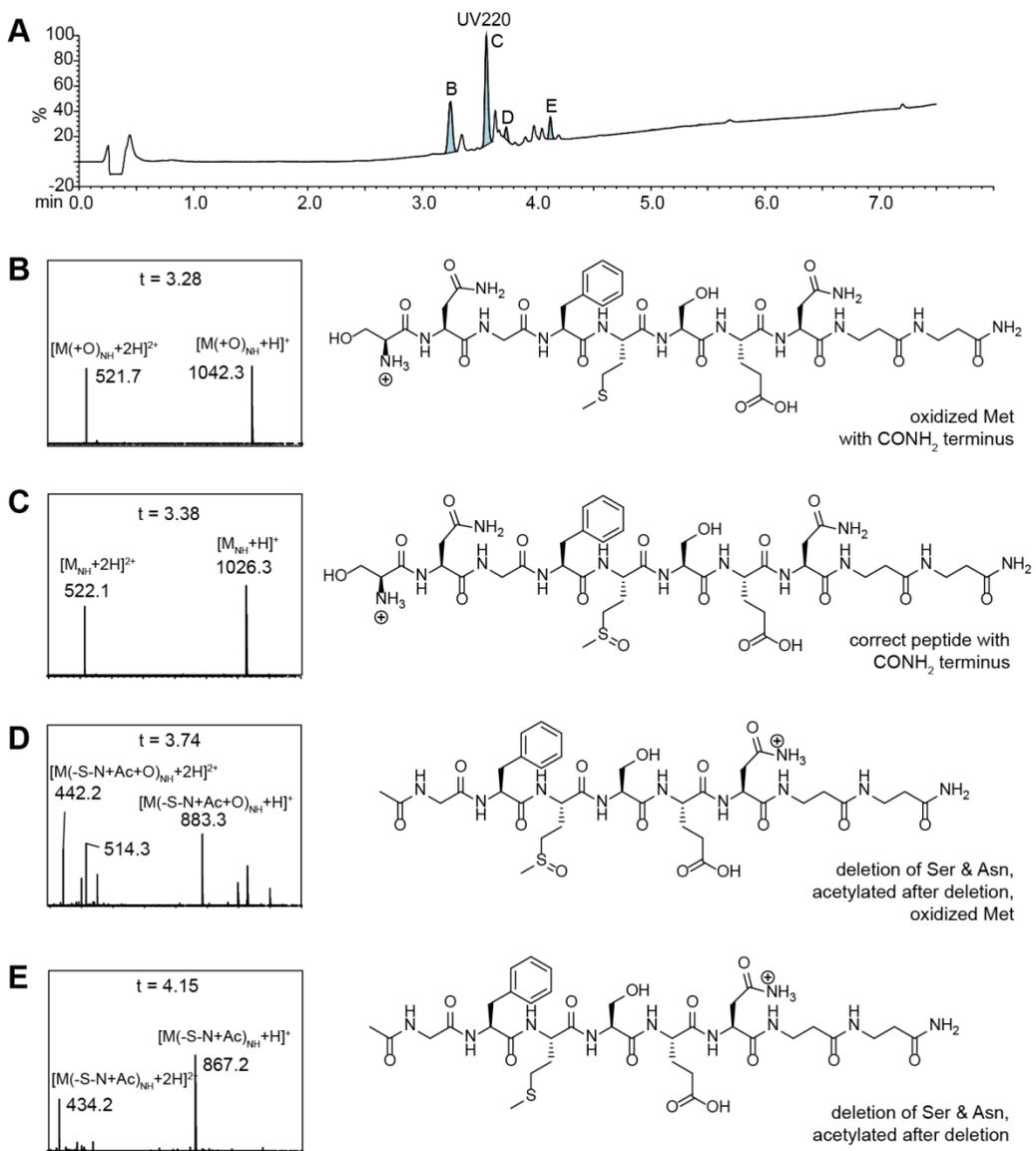
287. Tjhung, K. F.; Deiss, F.; Tran, J.; Chou, Y.; Derda, R., Intra-domain phage display (ID-PhD) of peptides and protein mini-domains censored from canonical pIII phage display. *Front. Microbiol.* **2015**, *6* (340).
288. Kaur, D.; Lowary, T. L.; Vissa, V. D.; Crick, D. C.; Brennan, P. J., Characterization of the epitope of anti-lipoarabinomannan antibodies as the terminal hexaarabinofuranosyl motif of mycobacterial arabinans. *Microbiology* **2002**, *148* (Pt 10), 3049-57.
289. S. Ng; K. Tjhung; B. Paschal; C. Noren; Derda, R., *Peptide libraries*. 2015/03/12 ed.; Springer New York: 2015; Vol. 1248.
290. Achkar, J. M.; Lawn, S. D.; Moosa, M.-Y. S.; Wright, C. A.; Kasprovicz, V. O., Adjunctive Tests for Diagnosis of Tuberculosis: Serology, ELISPOT for Site-Specific Lymphocytes, Urinary Lipoarabinomannan, String Test, and Fine Needle Aspiration. *J. Infect. Dis.* **2011**, *204* (suppl 4), S1130-S1141.
291. Tsibulkin, A. P.; Khaertanova, I. M.; Urazov, N. G.; Khaertinov, K. S., The Screening Diagnostic potential of Native Protein Fractions of Mycobacterium Tuberculosis using Technique of Immune blotting. *Klin. Lab. Diagn.* **2016**, *61* (2), 90-2, 102.
292. Ireton, G. C.; Greenwald, R.; Liang, H.; Esfandiari, J.; Lyashchenko, K. P.; Reed, S. G., Identification of Mycobacterium tuberculosis Antigens of High Serodiagnostic Value. *Clin. Vaccine Immunol.* **2010**, *17* (10), 1539-1547.
293. Samanich, K. M.; Keen, M. A.; Vissa, V. D.; Harder, J. D.; Spencer, J. S.; Belisle, J. T.; Zolla-Pazner, S.; Laal, S., Serodiagnostic Potential of Culture Filtrate Antigens of Mycobacterium tuberculosis. *Clin. Diagn. Lab. Immunol.* **2000**, *7* (4), 662-668.
294. Van-Lume, D. S. M.; De Souza, J. R.; Cabral, M. M. L.; Rego, J. C.; Balbino, V.; Saad, M. H.; Schindler, H. C.; Abath, F. G. C.; Montenegro, S. M. L., Immunological Diagnosis of Tuberculosis Based on Recombinant Antigens ESAT-6 and CFP-10 in Children from an Endemic Area in Northeast Brazil. *Scand. J. Immunol.* **2010**, *72* (5), 460-468.
295. Zhou, F.; Xu, X.; Wu, S.; Cui, X.; Fan, L.; Pan, W., Protein array identification of protein markers for serodiagnosis of Mycobacterium tuberculosis infection. *Sci. Rep.* **2015**, *5*, 15349.
296. Kulshrestha, A.; Gupta, A.; Verma, N.; Sharma, S. K.; Tyagi, A. K.; Chaudhary, V. K., Expression and purification of recombinant antigens of Mycobacterium tuberculosis for application in serodiagnosis. *Protein Expr. Purif.* **2005**, *44* (1), 75-85.
297. Tiwari, R. P.; Tiwari, D.; Garg, S. K.; Chandra, R.; Bisen, P. S., Glycolipids of Mycobacterium tuberculosis Strain H37Rv Are Potential Serological Markers for Diagnosis of Active Tuberculosis. *Clin. Diagn. Lab. Immunol.* **2005**, *12* (3), 465-473.
298. Bezerra, J. M.; Beck, S. T.; Kanunfre, K. A.; Leite, O. M.; Ferreira, A. W., A study of iga antibody response to different mycobacterium tuberculosis antigens in the diagnosis and monitoring of pulmonary tuberculosis. *Braz. J. Infect. Dis.* **2009**, *13*, 53-58.
299. Morioka, T.; Loik, N. D.; Hipolito, C. J.; Goto, Y.; Suga, H., Selection-based discovery of macrocyclic peptides for the next generation therapeutics. *Curr. Opin. Chem. Biol.* **2015**, *26*, 34-41.
300. Torgal-Garcia, J.; David, H. L.; Papa, F., Preliminary evaluation of a Mycobacterium tuberculosis phenolglycolipid antigen in the serologic diagnosis of tuberculosis. *Annales de l'Institut Pasteur / Microbiologie* **1988**, *139* (3), 289-294.

301. Gevorkian, G.; Segura, E.; Acero, G.; Palma, J. P.; Espitia, C.; Manoutcharian, K.; Lopez-Marin, L. M., Peptide mimotopes of Mycobacterium tuberculosis carbohydrate immunodeterminants. *Biochem. J* **2005**, *387* (Pt 2), 411-7.
302. Kalhor-Monfared, S.; Jafari, M. R.; Patterson, J. T.; Kitov, P. I.; Dwyer, J. J.; Nuss, J. M.; Derda, R., Rapid biocompatible macrocyclization of peptides with decafluoro-diphenylsulfone. *Chem. Sci.* **2016**, *7* (6), 3785-3790.

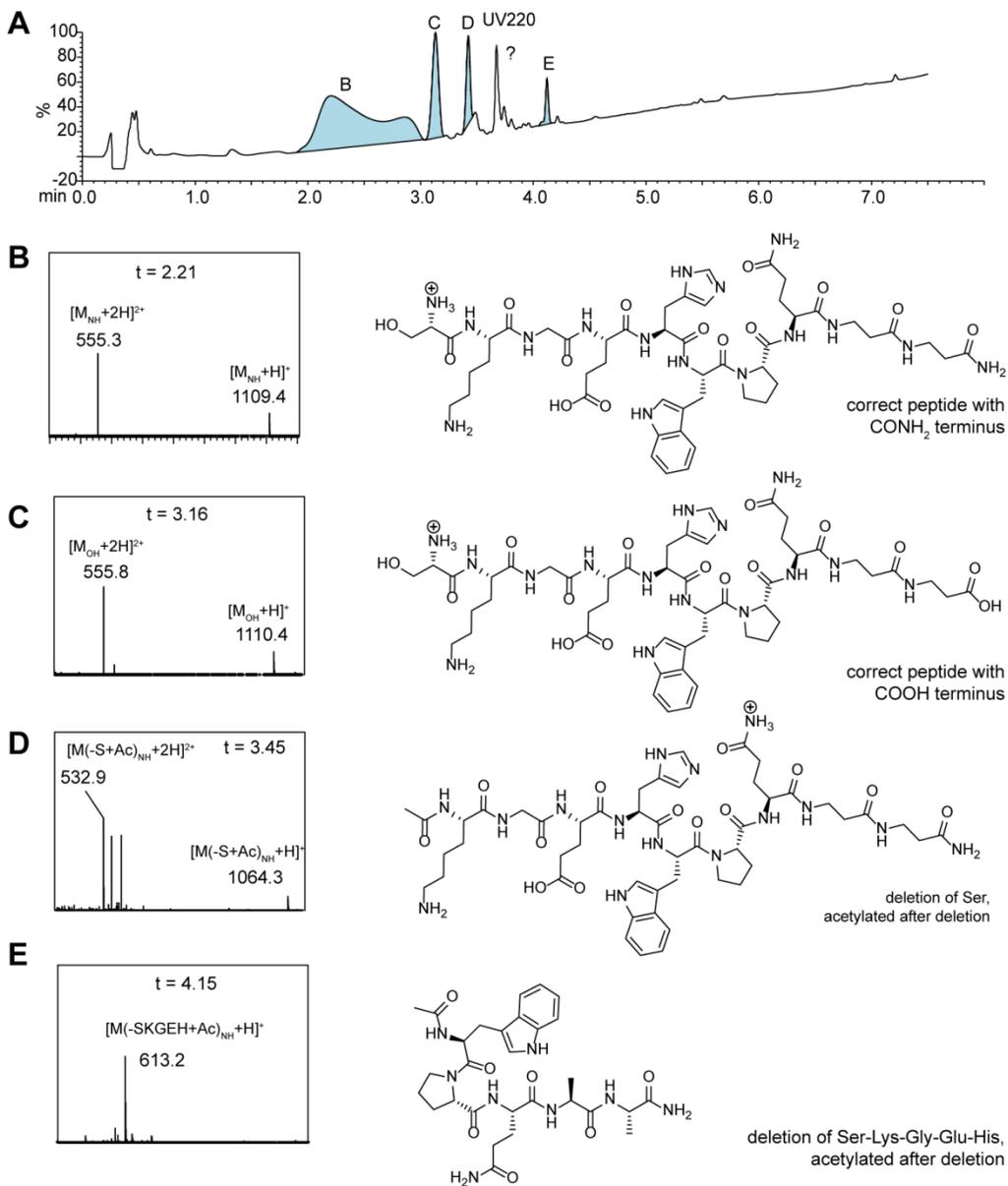
Appendix A: Supporting information for chapter 2



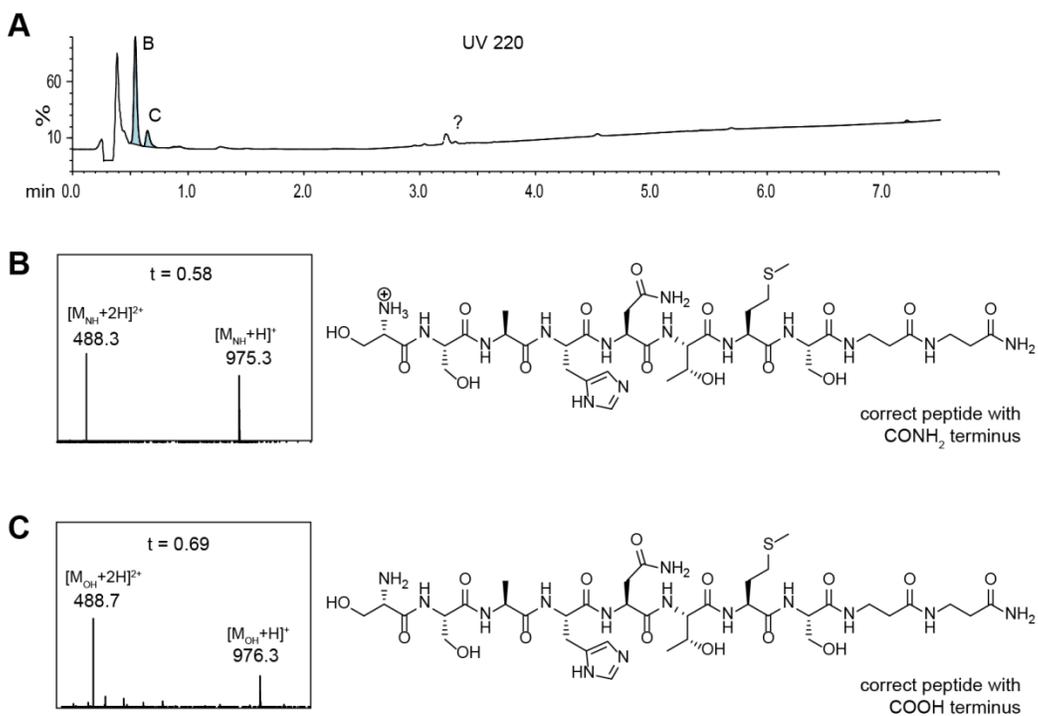
Appendix A-1. LCMS analysis of peptide Ser-Pro-Ala-Gly-Ala-Met-Phe-Gln-Ser-βAla-βAla synthesized on paper.



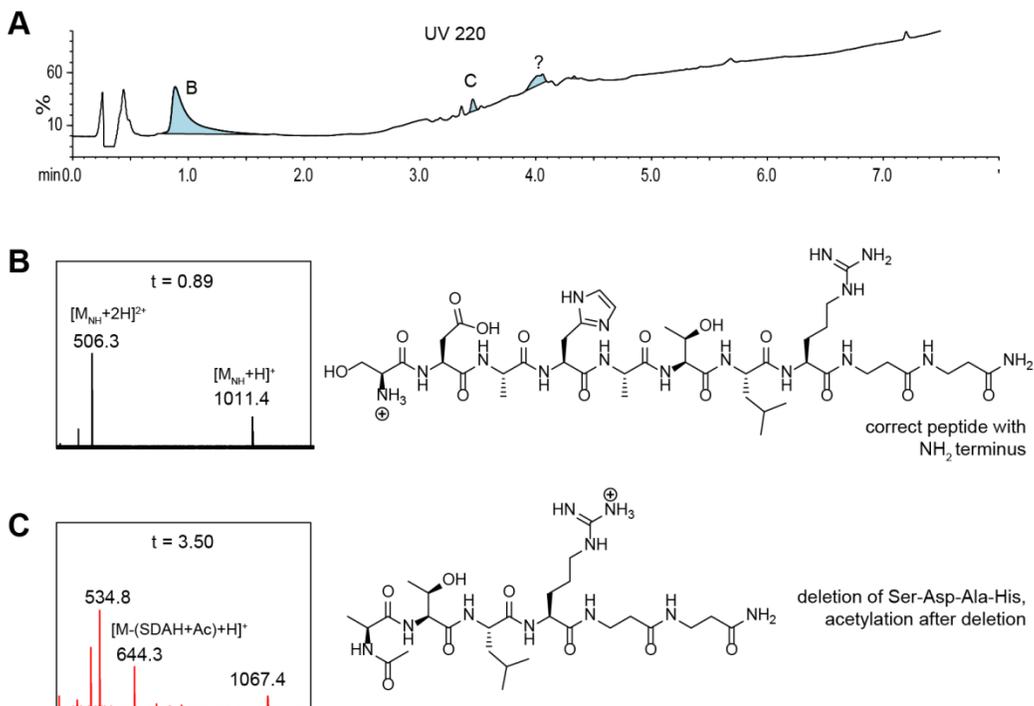
Appendix A-2. LCMS analysis of peptide Ser-Asn-Gly-Phe-Met-Ser-Glu-Asn- β Ala- β Ala synthesized on paper.



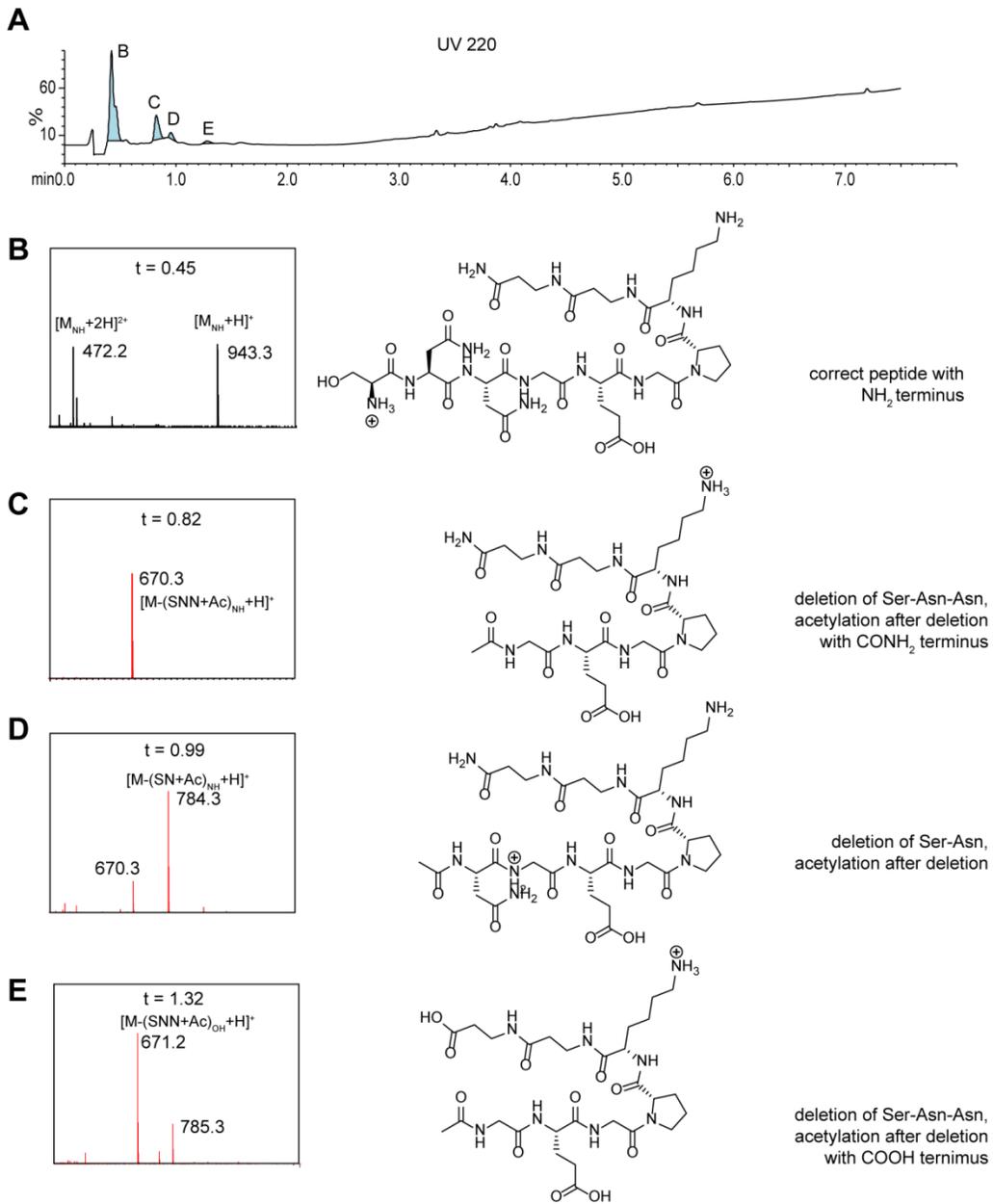
Appendix A-3. LCMS analysis of peptide Ser-Lys-Gly-Glu-His-Trp-Pro-Gln-βAla-βAla synthesized on paper.



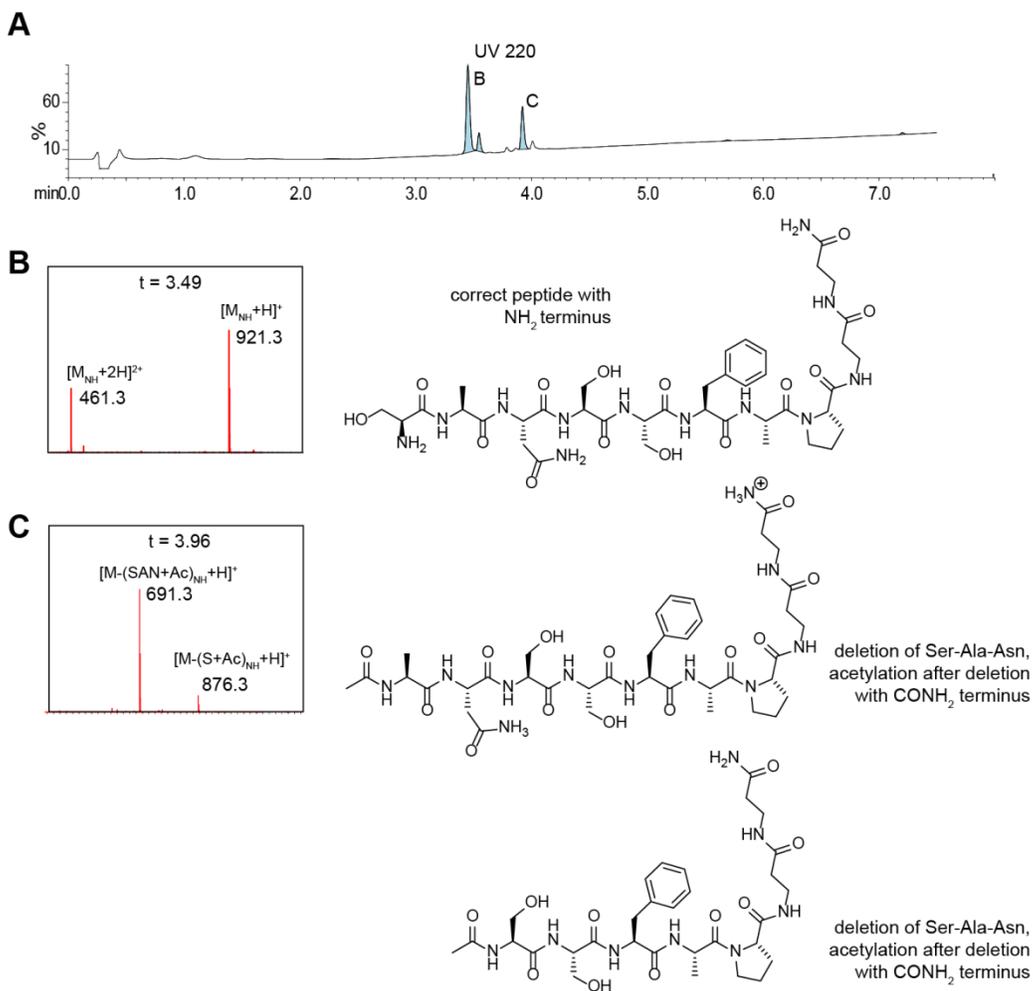
Appendix A-4. LCMS analysis of peptide Ser-Ser-Ala-His-Asn-Thr-Met-Ser- β Ala- β Ala synthesized on paper.



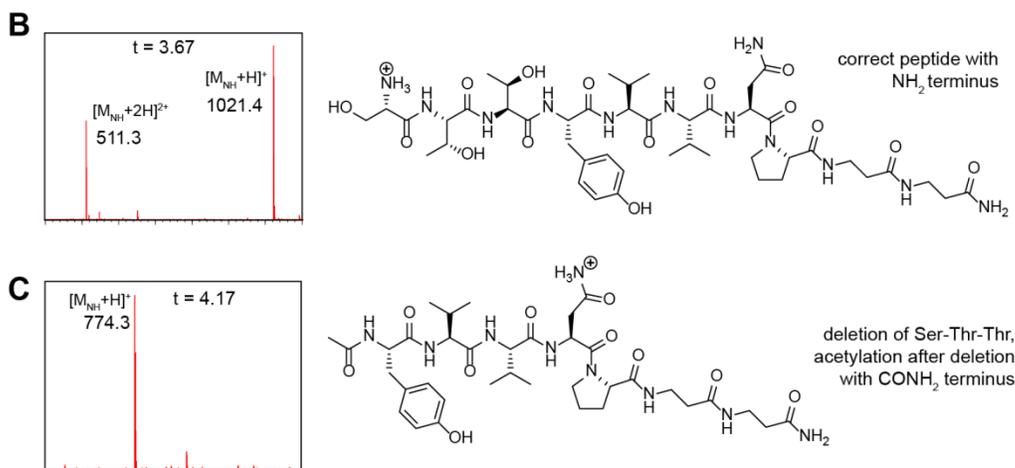
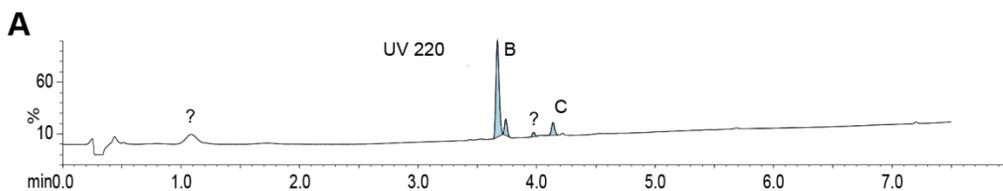
Appendix A-5. LCMS analysis of peptide Ser-Asp-Ala-His-Ala-Thr-Leu-Arg- β Ala- β Ala synthesized on paper.



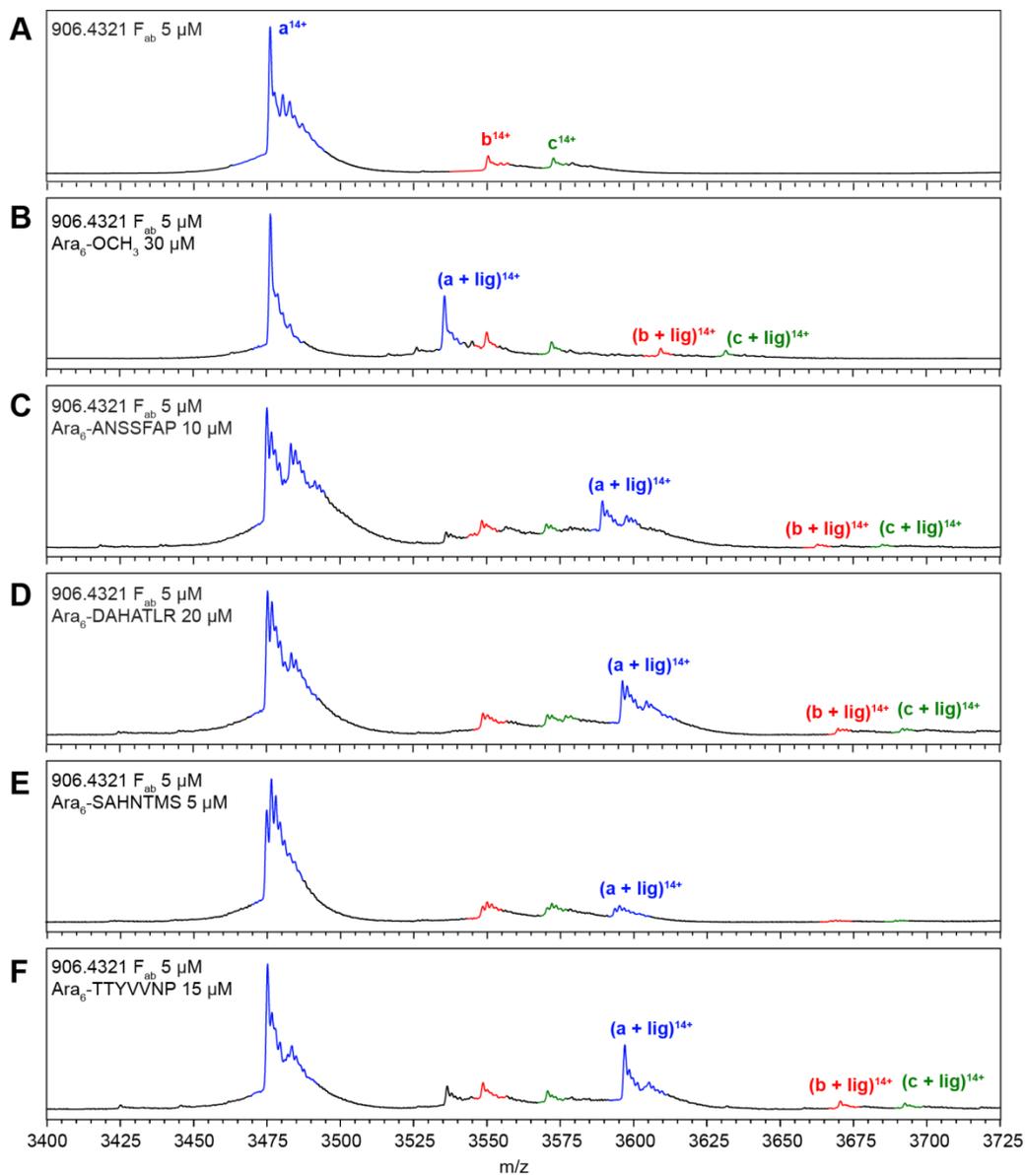
Appendix A-6. LCMS analysis of peptide Ser-Asn-Asn-Gly-Glu-Gly-Pro-Lys- β Ala- β Ala synthesized on paper.



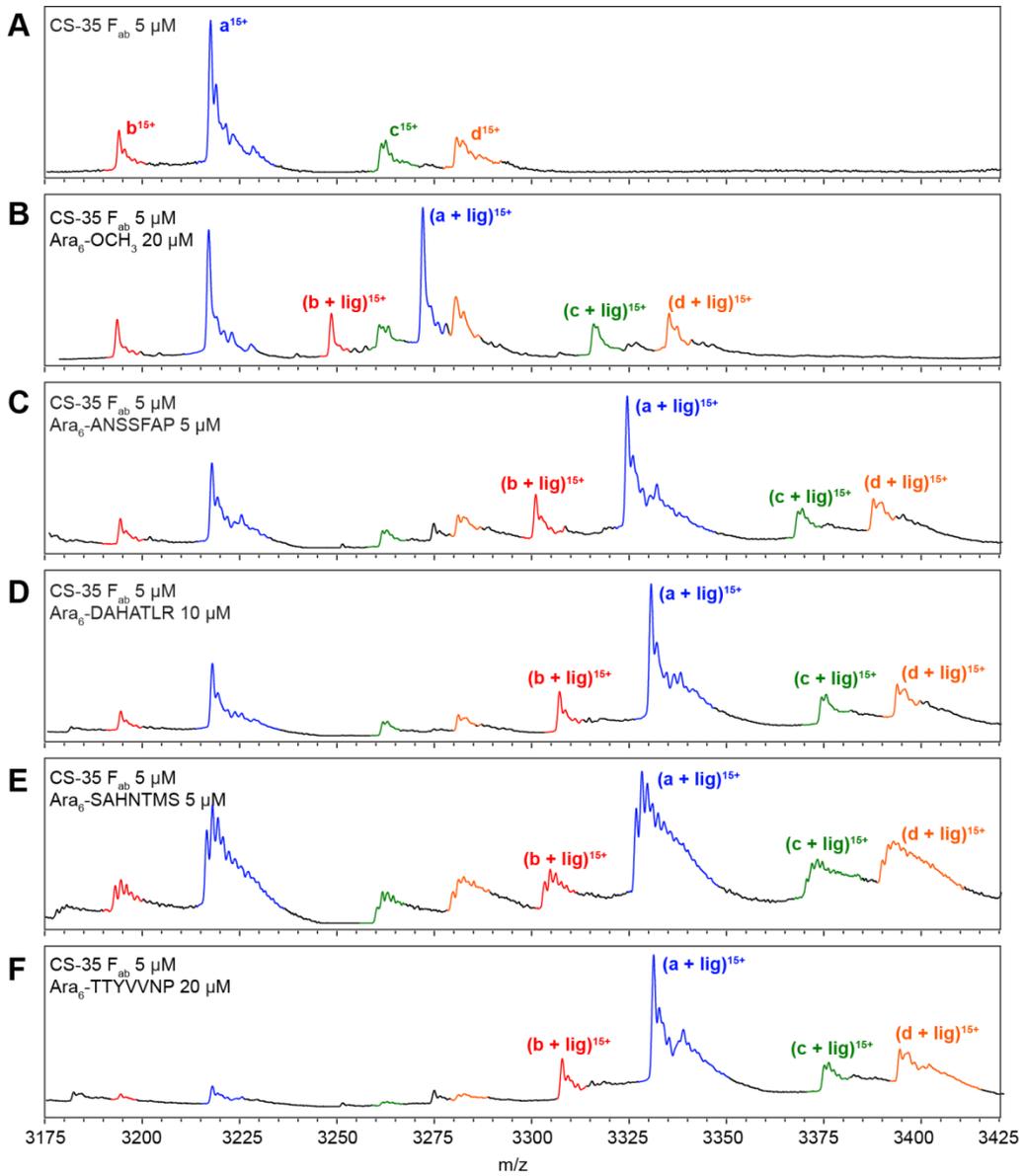
Appendix A-7. LCMS analysis of peptide Ser-Ala-Asn-Ser-Ser-Phe-Ala-Pro- β Ala- β Ala synthesized on paper.



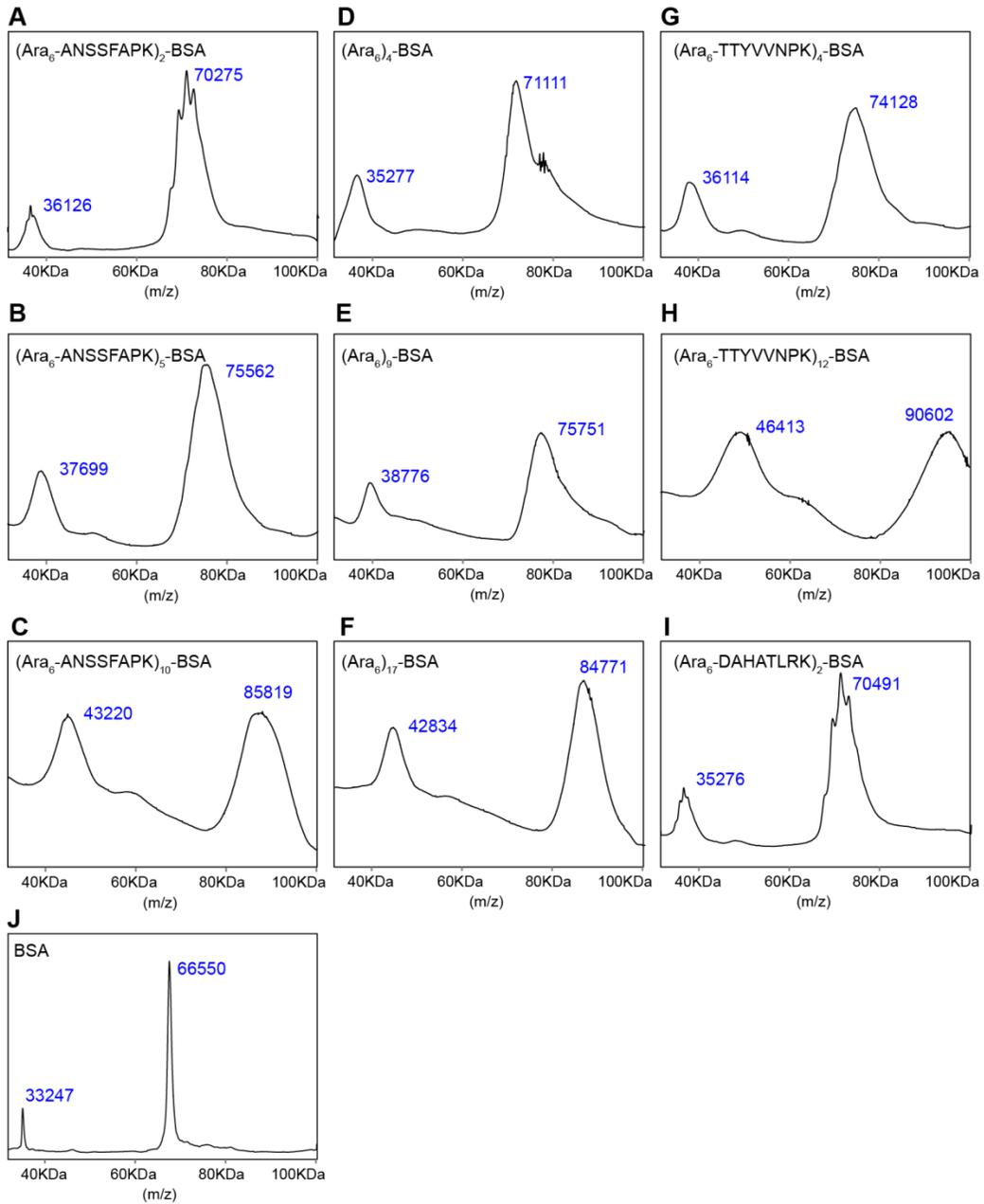
Appendix A-8. LCMS analysis of peptide Ser-Thr-Thr-Tyr-Val-Val-Asn-Pro- β Ala- β Ala synthesized on paper.



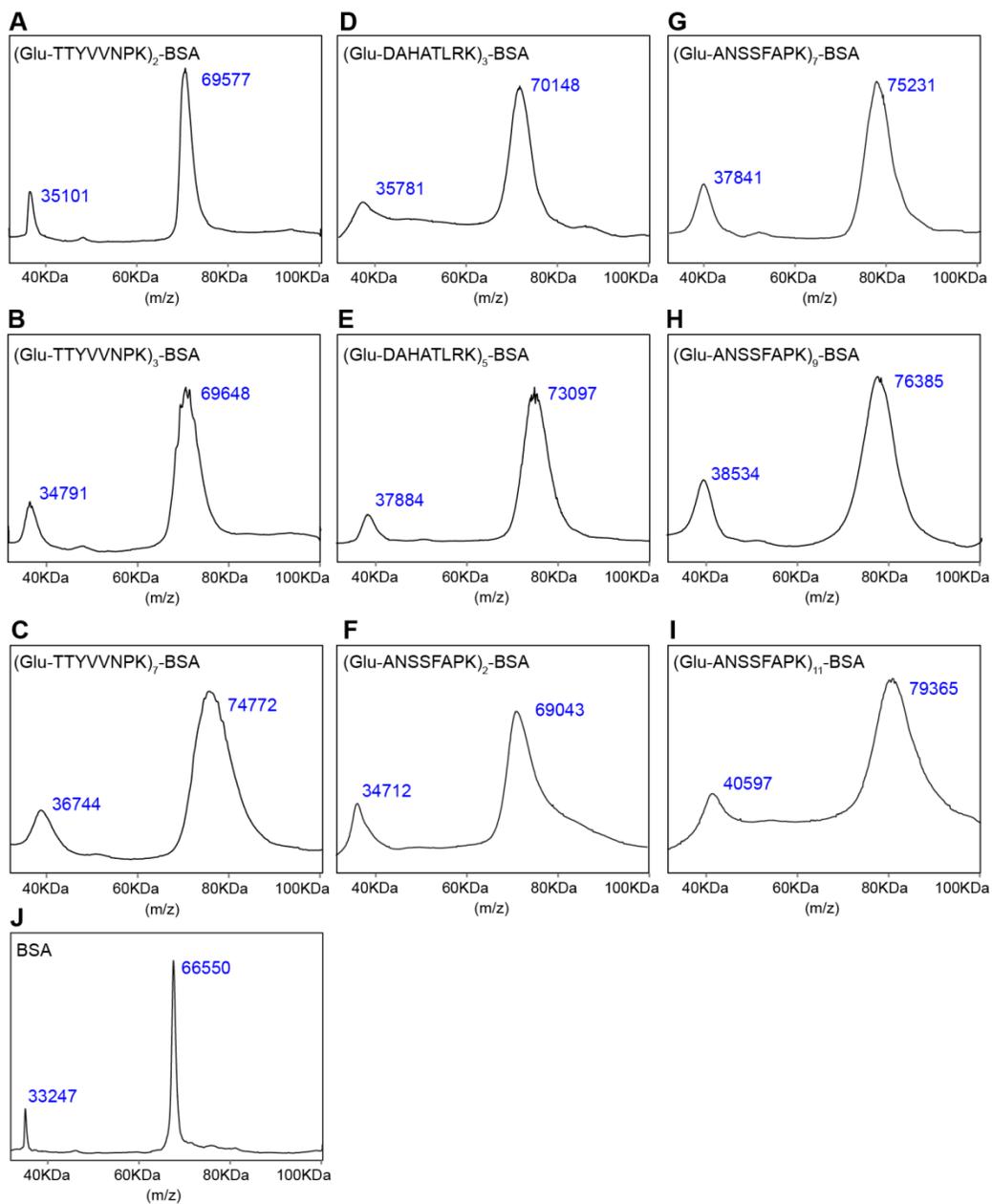
Appendix A-9. ESI-MS binding data of Ara₆-glycopeptides to 906.4321 F_{ab}.



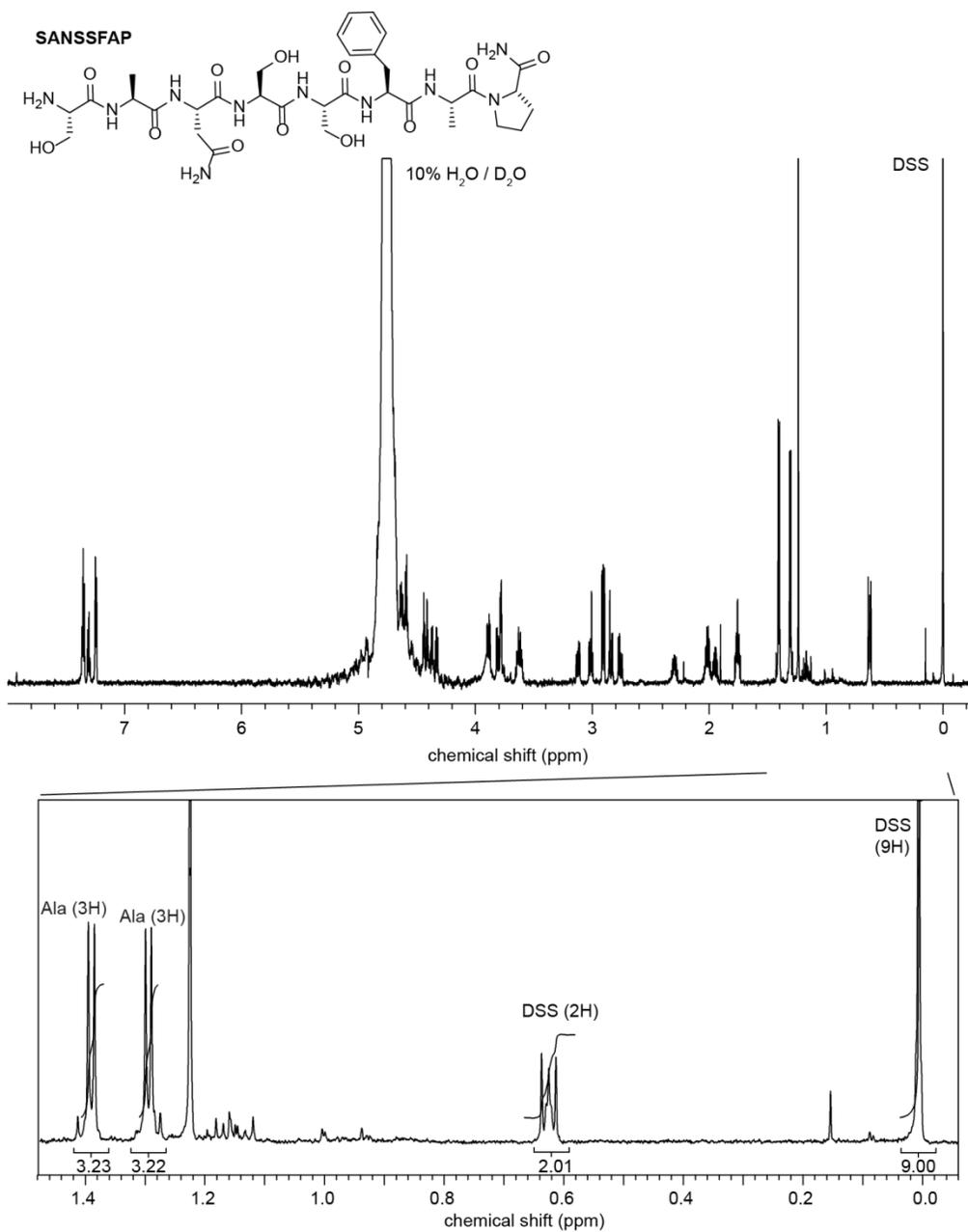
Appendix A-10. ESI-MS binding of data Ara₆-glycopeptides to CS-35 F_{ab}.



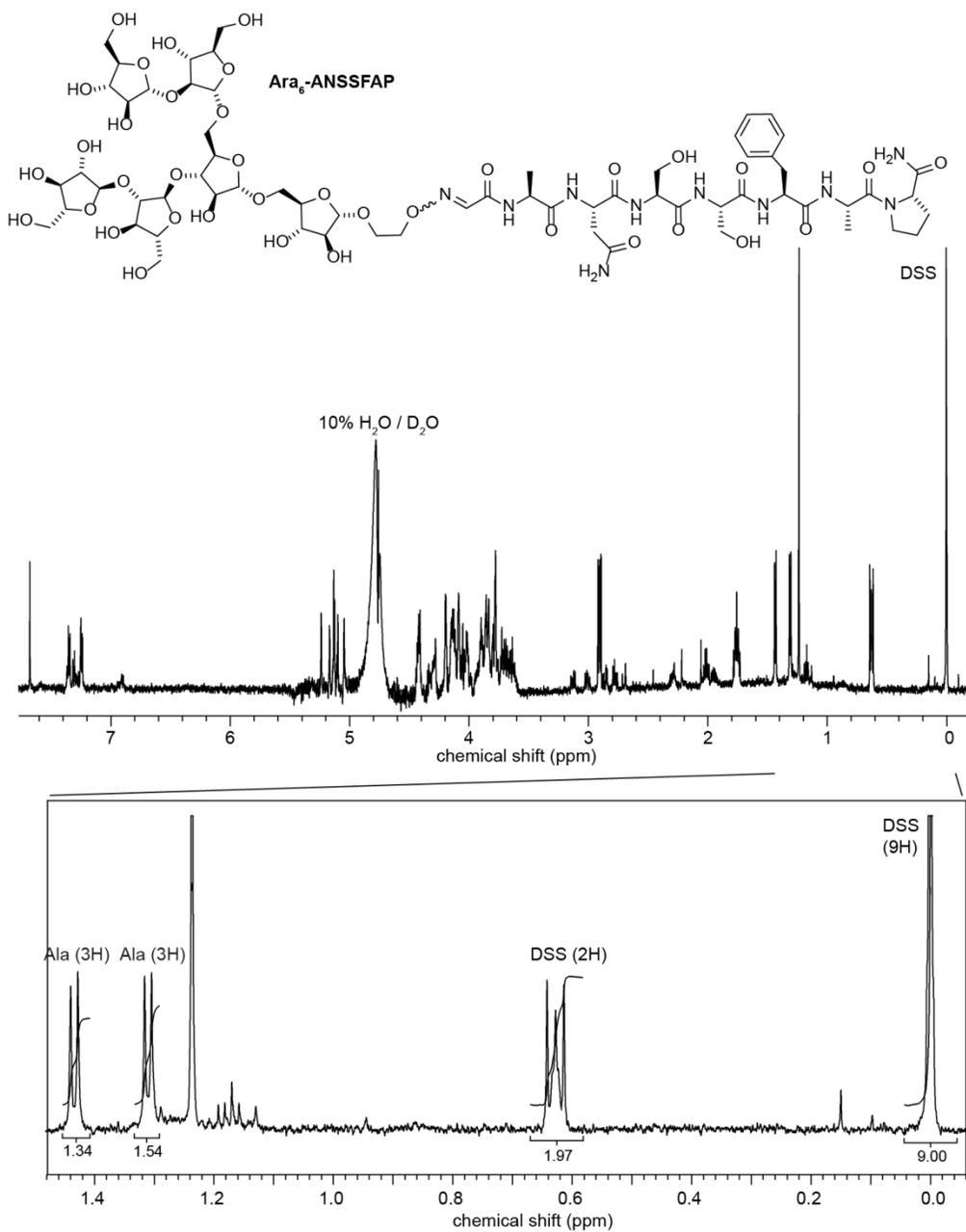
Appendix A-11. MALDI of synthesized Ara₆glycopeptide BSA conjugates.



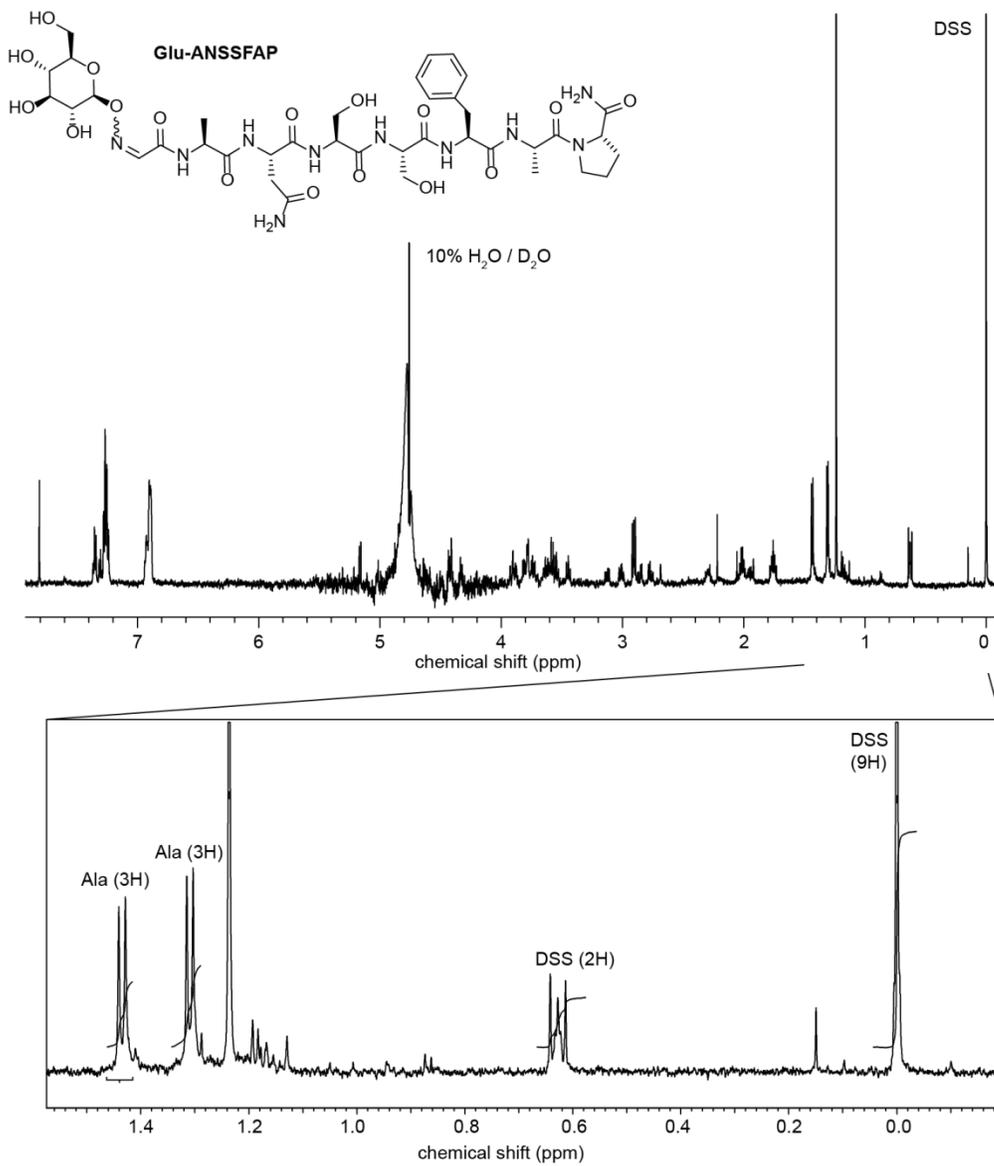
Appendix A-12. MALDI of synthesized Glu-glycopeptide BSA conjugates.



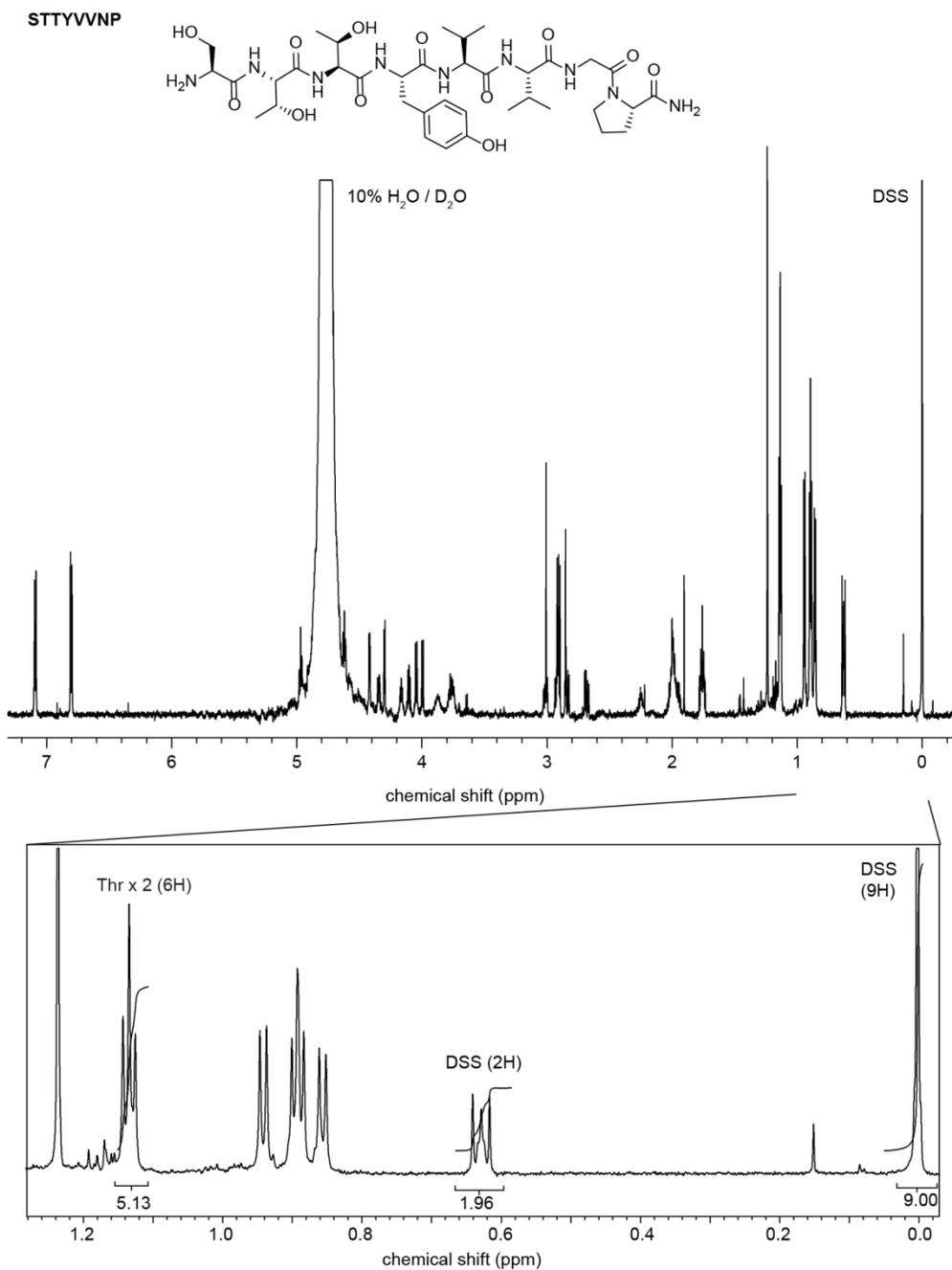
Appendix A-13. ¹H NMR (700 MHz, D₂O) spectra of SANSSFAP.



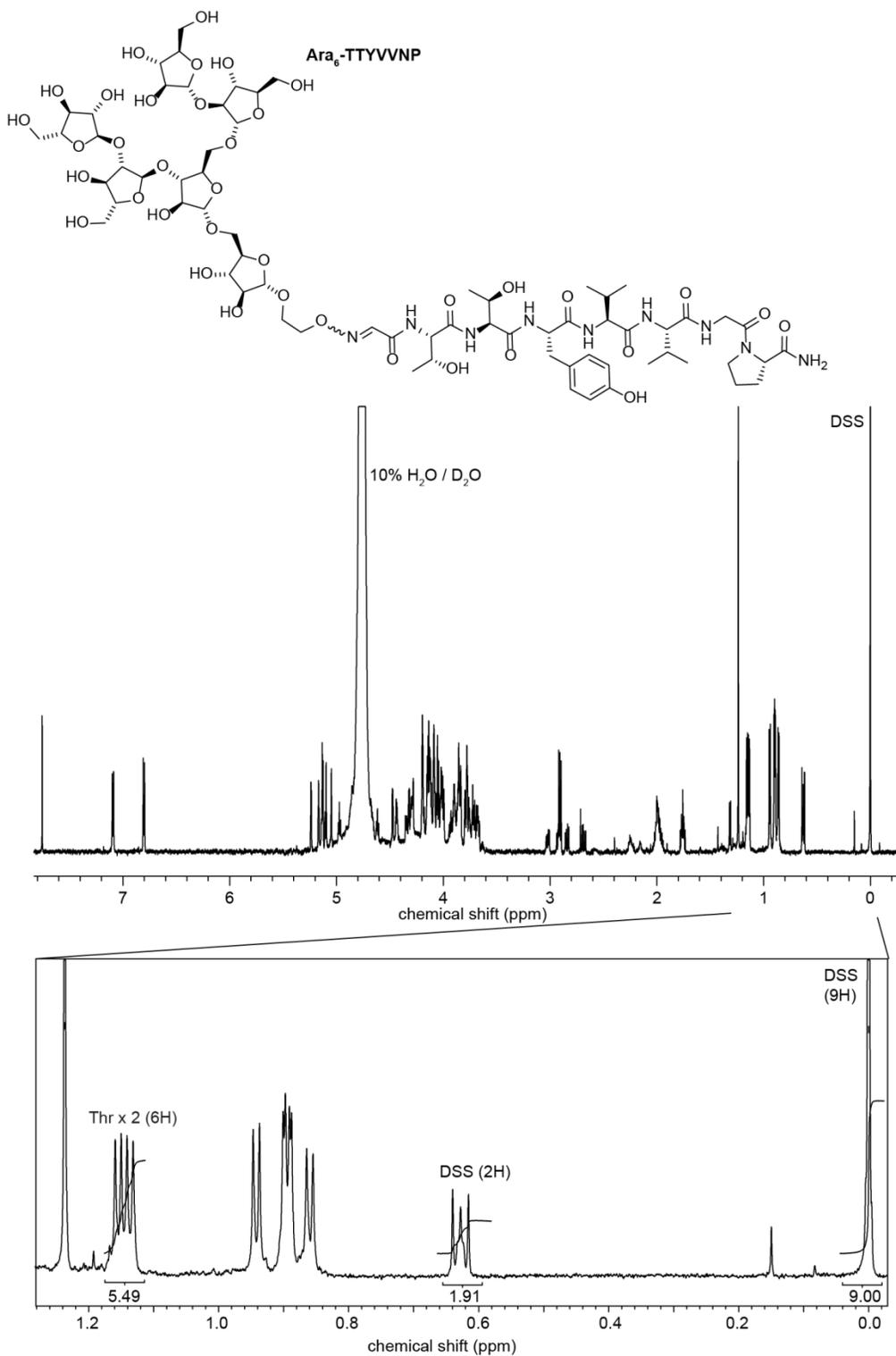
Appendix A-14. ¹H NMR (700 MHz, D₂O) spectra of Ara₆-ANSSFAP.



Appendix A-15. ^1H NMR (700 MHz, D_2O) spectra of Glu-ANSSFAP.

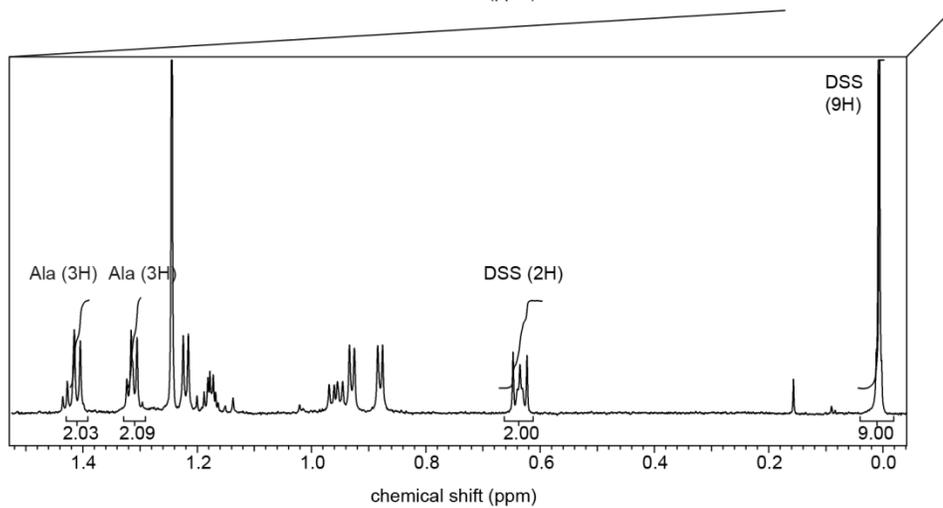
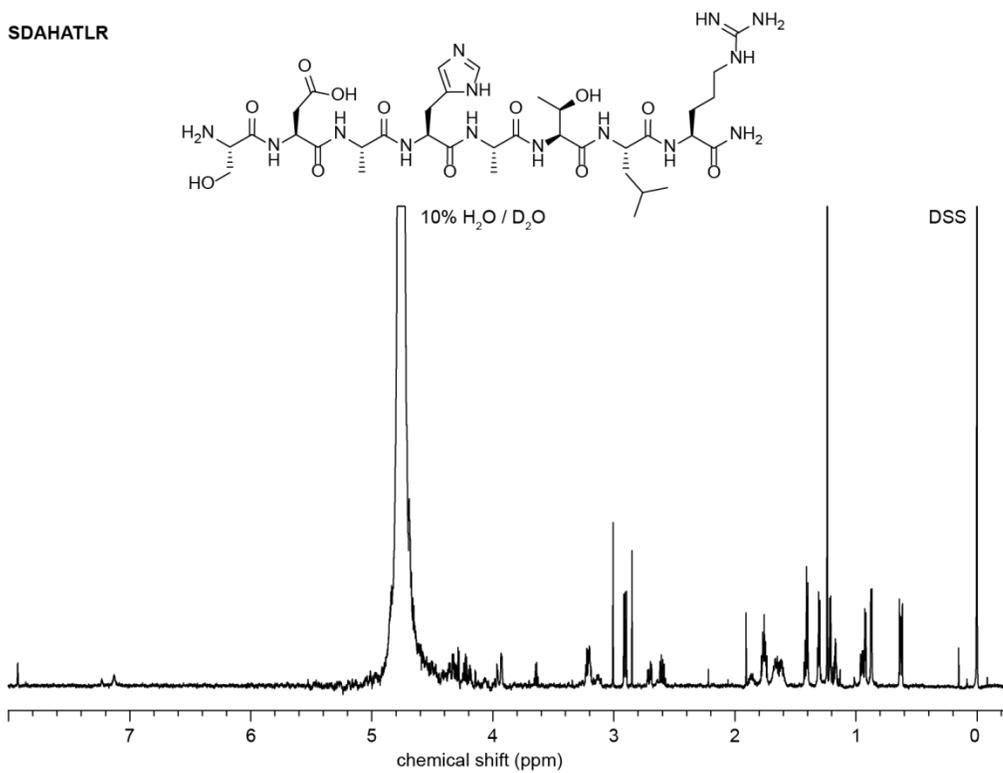


Appendix A-16. ¹H NMR (700 MHz, D₂O) spectra of STTYVVNP.

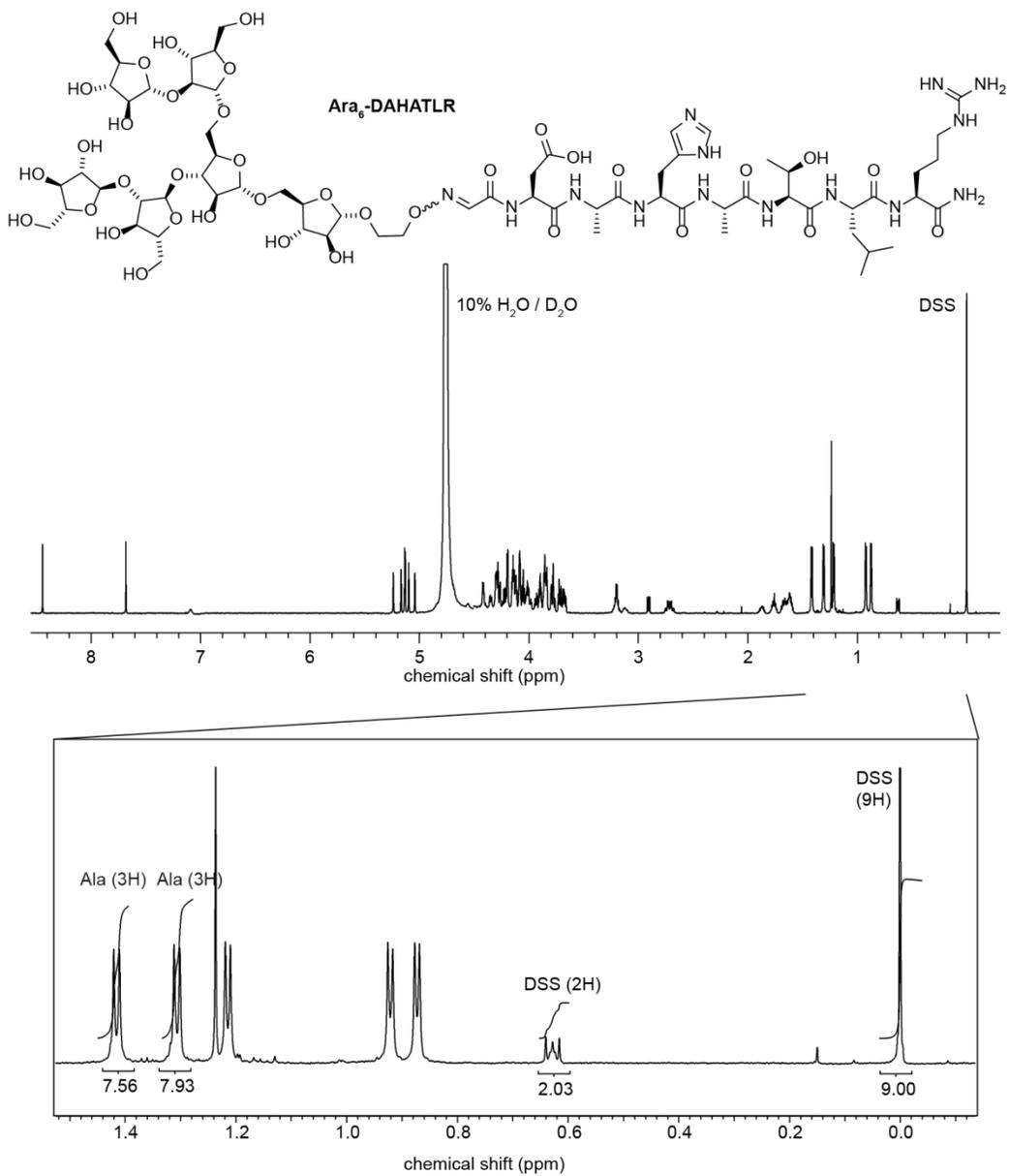


Appendix A-17. ¹H NMR (700 MHz, D₂O) spectra of Ara₆-TTYVVNP.

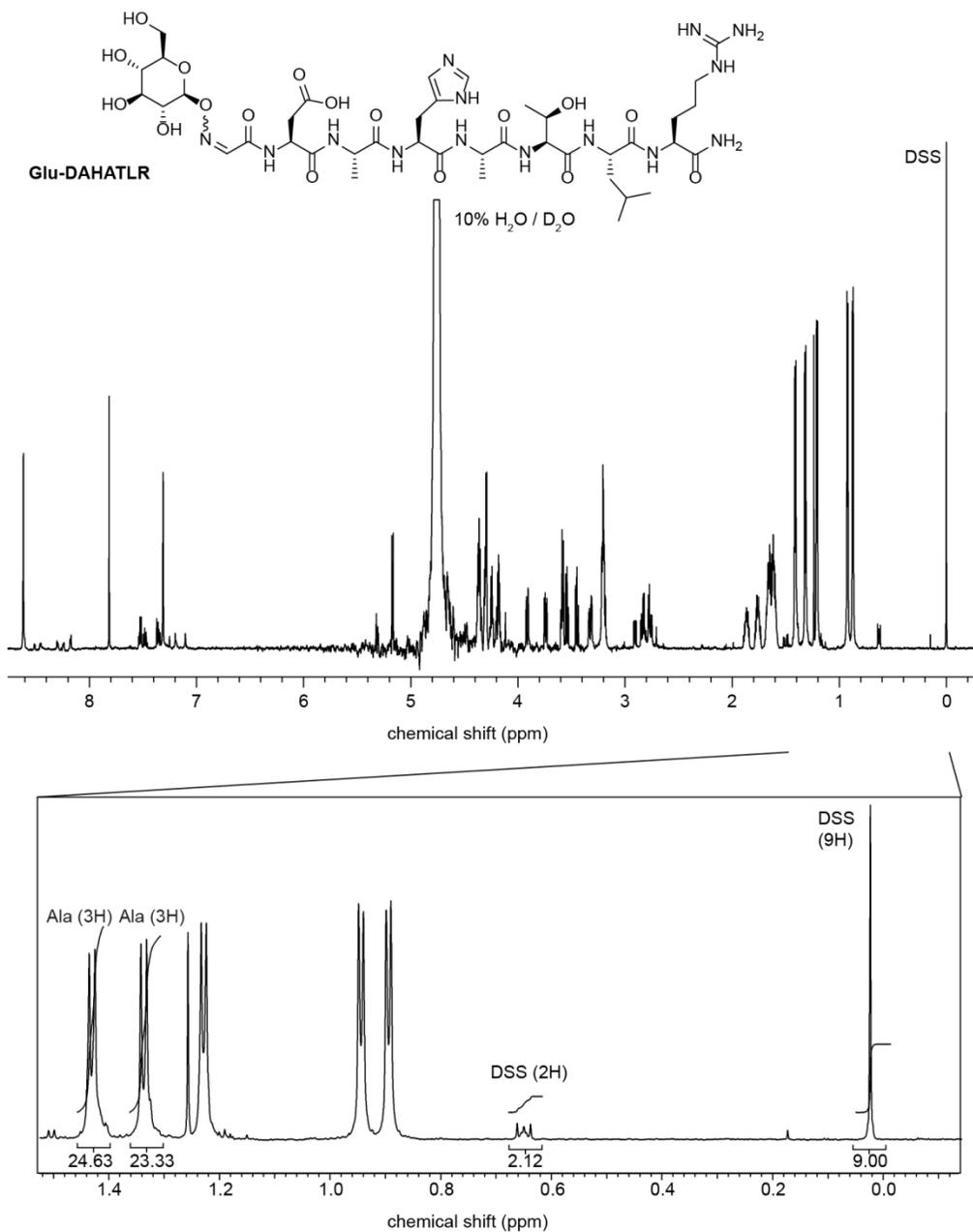
SDAHATLR



Appendix A-19. ¹H NMR (700 MHz, D₂O) spectra of SDAHATLR.

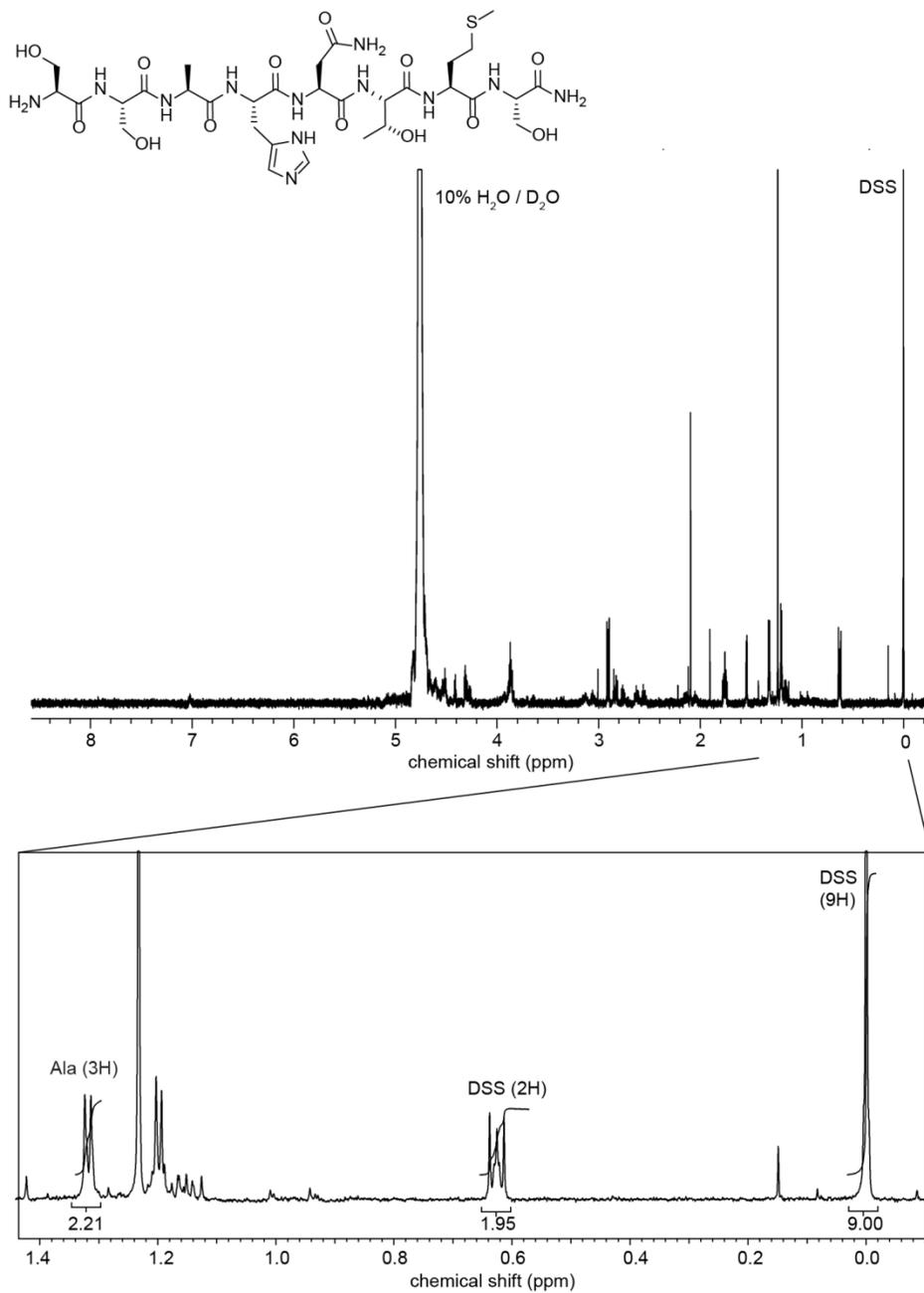


Appendix A-20. ¹H NMR (700 MHz, D₂O) spectra of Ara₆-DAHATLR.

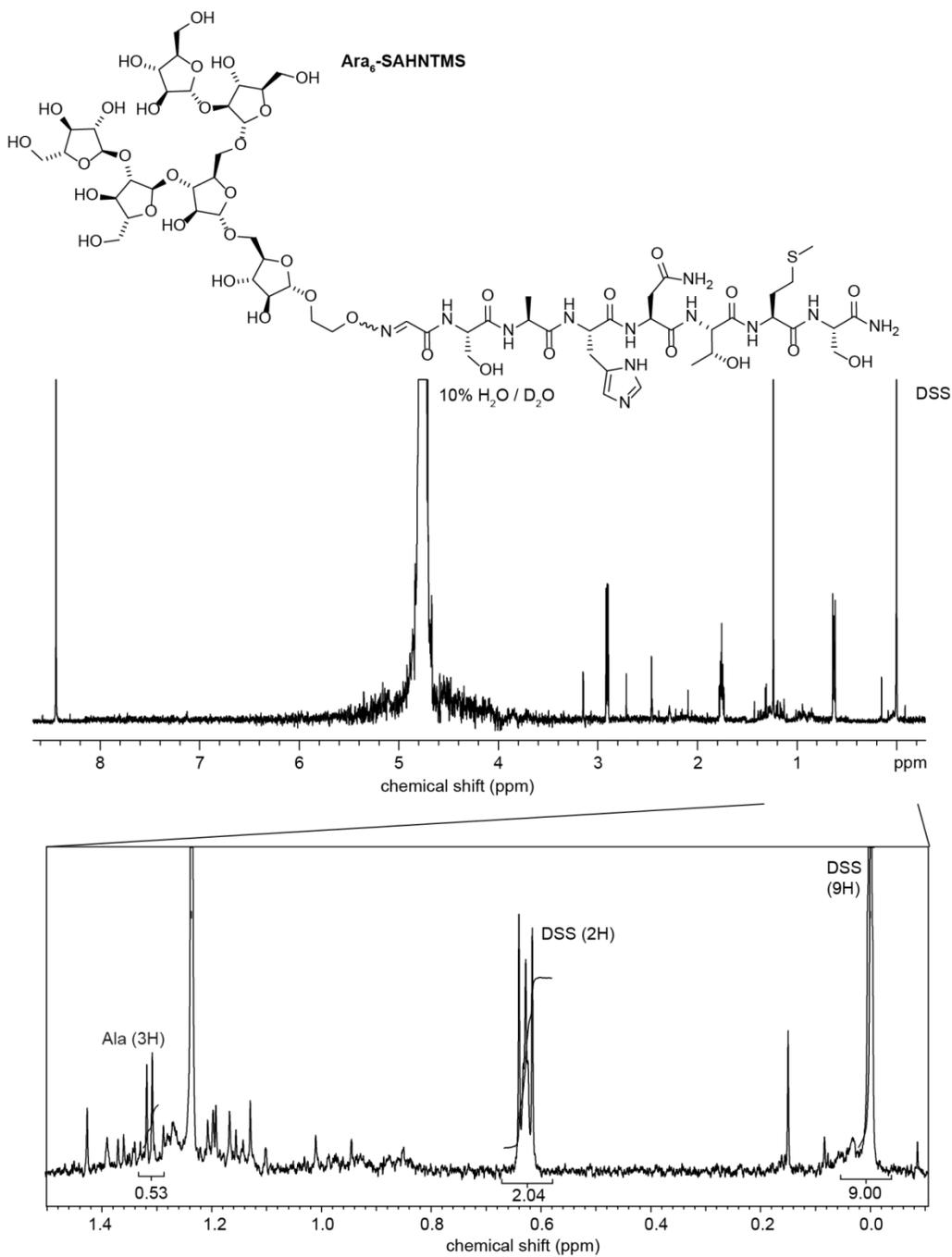


Appendix A-21. ¹H NMR (700 MHz, D₂O) spectra of Glu-DAHATLR.

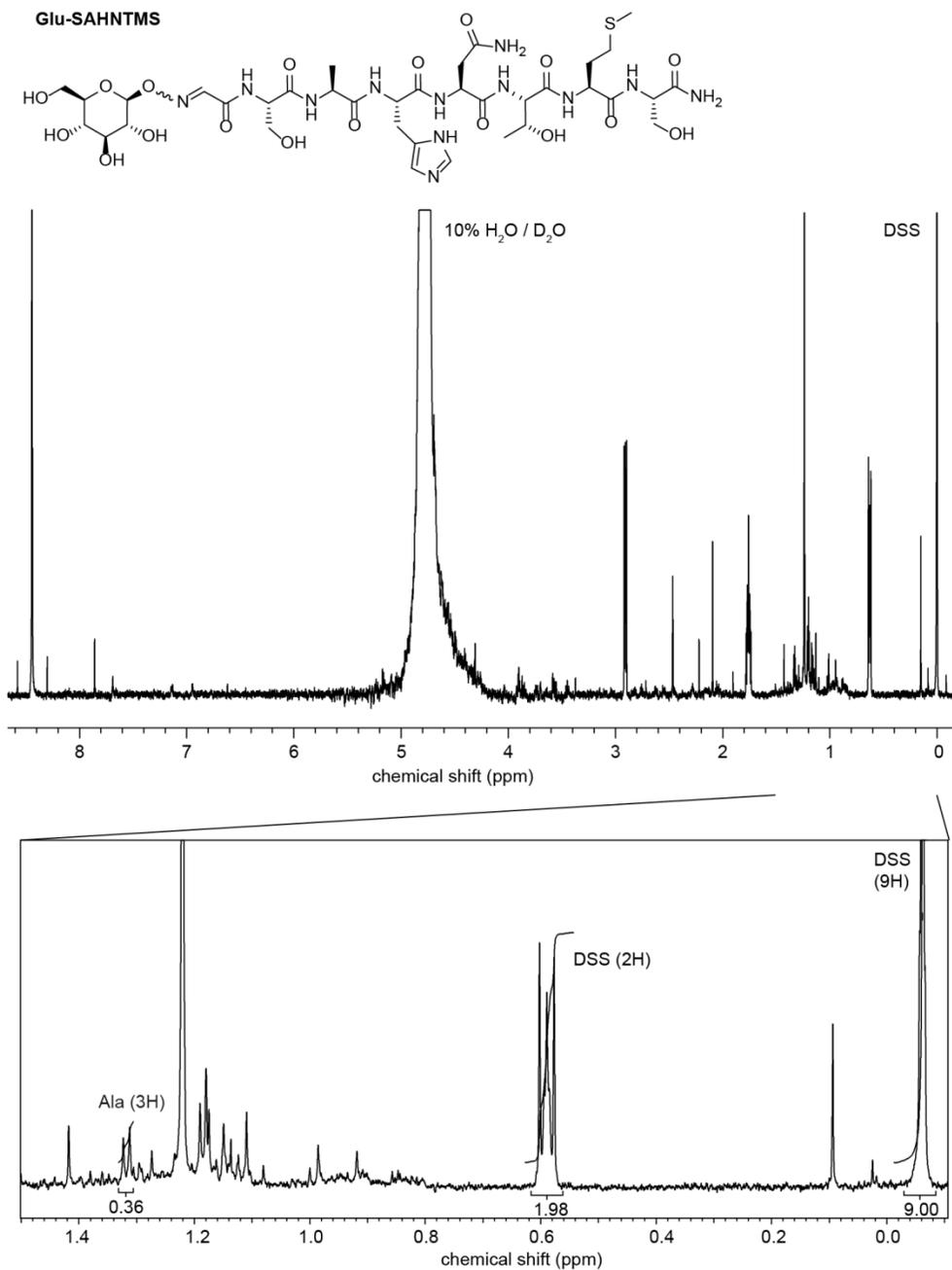
SSAHNTMS



Appendix A-22. ¹H NMR (700 MHz, D₂O) spectra of SSAHNTMS.



Appendix A-23. ¹H NMR (700 MHz, D₂O) spectra of Ara₆-SAHNTMS.



Appendix A-24. ¹H NMR (700 MHz, D₂O) spectra of Glu-SAHNTMS.

Appendix A-25. MatLab script Makefigure2_4.m

```
clear all; close all;

Dir='';
File = 'SequencingDataForFigure2-4.txt';
SAVEto = [File(1:end-4) ' Analysis.txt']; % keep blank if don't want to save

SET{1} = [1 2 3 4]; % RI Ara6-modified on Cs-35
SET{2} = [5 6 7 8]; % RI Gal-modified on CS-35
SET{3} = [9 10 11 12]; % RI Ara6-modified on blank well
SET{4} = [13 14 15]; % RII ara6-CS35
SET{5} = [16 17 18]; % RII gal-CS35
SET{6} = [19 20 21]; % RII ara6-blank
SET{7} = [22 23 24]; % RII unmod-CS35

TEST_SET = 4;
CONTROL_SETS = [1 2 3 5 6 7];

HITS2DISPLAY = 100; % maximum number of hits to display
SHOWaminoACIDS = [1 2 3 4 5 6 7 8];

CLUSTERbyH = 1; % 1 if you want your hits to be clustered by Hamming dist.

PLOT_VOLCANO = 1; % set to 1 if you want to see the actual volcano plot

%%%%%% volcano plot parameters here %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
p_cutoff = 0.05; % p-value cutoff
R_cutoff = 4; % ratio cutoff
MaxX=14; % maximum on the X-scale (if plotting volcano)
vert_cutoff = 0.00001; % maximum on the Y-scale (if plotting volcano)

[Nuc, AA, Fr] = readMulticolumn('Dir', Dir, 'File', File, ...
                               'column', 1:max(cell2mat(SET)),...
                               'skip', 2, 'output', 'normalized+1');

% select only the aminoacids you want to see
cAA = char(AA);
AA=cellstr(cAA(:,SHOWaminoACIDS));

SQUARE=zeros(size(Fr,1),1);
IX=zeros(size(Fr,1),numel(CONTROL_SETS));

i=0;

for j=CONTROL_SETS
    i=i+1;
    ratio(:,i) = mean(Fr(:,SET{TEST_SET}), 2) ./ mean(Fr(:,SET{j}), 2);

    [~,confi(:,i)] = ttest2(Fr(:,SET{TEST_SET})',Fr(:,SET{j})',...
                           p_cutoff,'right','unequal');

    IX(:,i) = (confi(:,i) <= p_cutoff) & (ratio(:,i) >= R_cutoff);

    SQUARE = SQUARE + ratio(:,i).^2;

    if PLOT_VOLCANO
        subplot(1,numel(CONTROL_SETS),i);

        plot(log2(ratio(:,i)),...
             -log10(confi(:,i)),'d',...

```

```

        'MarkerSize',4,...
        'MarkerFaceColor',0.5*[1 1 1],...
        'MarkerEdgeColor',0.5*[1 1 1]); hold on;

    plot( log2 ( ratio(find(IX(:,i)),i)),...
        -log10( confi( find(IX(:,i)),i)), 'd',...
        'MarkerSize',4,...
        'MarkerFaceColor','r',...
        'MarkerEdgeColor','r'); hold on;

    line([log2(R_cutoff) MaxX],[-log10(p_cutoff) -log10(p_cutoff)]);
    line([ log2(R_cutoff)    log2(R_cutoff)],...
        [-log10(p_cutoff) -log10(vert_cutoff)]);

    xlim([-MaxX MaxX]);
end
end

R2 = sqrt(SQUARE);

IXall = find( (sum(IX,2)>=5) ); %hits that satisfy 5 criteria

hits    = char(AA(IXall,:));
Rhits   = ratio(IXall,:);
R2hits  = R2(IXall);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
figure(2)
if size(hits,1)>3
    Y = pdist(hits,'hamming');
    Z = linkage(Y,'complete');
    [H,T,perm] = dendrogram(Z,0,'colorthreshold',20);
    set(H,'LineWidth',2)

    for i =1:size(hits,1)
        label{i} = i;
    end
    set(gca,'XTick', 1:1:size(hits,1), 'XTickLabel',label);

    hits    = hits(perm,:);
    Rhits   = Rhits(perm,:);
    R2hits  = R2(perm);
    IXall   = IXall(perm);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% display all results as heat map
figure(3)

if size(IXall,1)>=HITS2DISPLAY
    N=HITS2DISPLAY; % display only the first or defined number of hits
else
    N=size(IXall,1); %display all
end

FrPPM = round(10^6*Fr); % convert normalized fraction frequency to PPM

imagesc( log10([FrPPM(IXall(1:N),:) ratio(IXall(1:N),:) ]+1) );

set(gca,'YTick', 1:1:N, 'YTickLabel',cellstr(hits(1:N,:)),'TickDir','out',...
    'FontName','Courier New','FontSize',14);
set(gca,'XTick', 1:1:size(Fr,2)+4, 'TickDir','out');
jet1=jet;
jet1(1,:)=[0.4 0.4 0.4];

```

```

colormap(jet1);
colorbar;

% generate a plain text table for saving or copy from command line

S = char(32*ones(size(hits,1),2));
L = [ S(:,1) char(124*ones(size(hits,1),1)) S(:,1)];
F = FrPPM(IXall,:); % display frequency in ppm

toSave = [hits      S  ];

for i=1:numel(SET)
    for j=1:numel(SET{i})
        toSave = [toSave num2str(F(:,SET{i}(j))) S];
    end
    toSave = [toSave L];
end
toSave = [toSave S  num2str(round(Rhits)) L];

disp(toSave);

if ~isempty(SAVEto)
    fs = fopen(fullfile(Dir,SAVEto),'w');
    RET = char(10*ones(size(toSave,1),1));
    fprintf( fs, '%s\r\n', [toSave RET]');
    fclose all;
end

% or you can just copy paste the results from the command line

```

Appendix A-26. MatLab script MakeFigure2_9.m

```
%using signal from empty spots as background
%using median signal

clear all;
close all;

% define colors for plotting
teal = [56 142 142]/256;
chartreuse = [113 198 113]/256;
turquoise = [ 0 134 139]/256;
crimson = [220 20 60 ]/256;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% defined the plotting appearance here
Xmin = 0.01; % smallest non-zero value of X
Xmax = 100;
Ymin = -5000; % smallest value of Y
Ymax = 60000;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

file = 'YC-XII-39_NoHeader';
DIR = '';
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

A = xlsread(fullfile(DIR,file)); %assign file contents to matrix A
NofIgGPoints = size(A,1); % total number of data points

row = A(:,3); %row number
column = A(:,2); %column number
block = A(:,1); %block number

IgG = A(:,9); % F532 Median

block_CS35 = [ 1, 4, 7, 10, 13, 16, 19, 22];
conc_CS35 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01];

block_CS40 = [ 2, 5, 8, 11, 14, 17, 20, 23];
conc_CS40 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01];

block_906 = [ 3, 6, 9, 12, 15, 18, 21, 24];
conc_906 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01];

% define structures that contains row, column, and name information

sample(1).R = 2;
sample(1).C = 1:3;
sample(1).name = 'Ara6-low';

sample(2).R = 2;
sample(2).C = 4:6;
sample(2).name = 'Ara6-med';

sample(3).R = 3;
sample(3).C = 1:3;
sample(3).name = 'Ara6-high';

sample(4).R = 3;
sample(4).C = 4:6;
sample(4).name = 'Ara6-ANSSFAPK-low';

sample(5).R = 4;
sample(5).C = 1:3;
sample(5).name = 'Ara6-ANSSFAPK-med';
```

```
sample(6).R = 4;
sample(6).C = 4:6;
sample(6).name = 'Ara6-ANSSFAPK-high';
```

```
sample(7).R = 6;
sample(7).C = 1:3;
sample(7).name = 'Ara6-TTYVVNPK-low';
```

```
sample(8).R = 6;
sample(8).C = 4:6;
sample(8).name = 'Ara6-TTYVVNPK-high';
```

```
sample(9).R = 7;
sample(9).C = 1:3;
sample(9).name = 'Ara6-DAHATLRK-low';
```

```
%%%%%%%%
```

```
sample(10).R = 7;
sample(10).C = 4:6;
sample(10).name = 'Glu-ANSSFAPK-low';
```

```
sample(11).R = 8;
sample(11).C = 1:3;
sample(11).name = 'Glu-ANSSFAPK-med';
```

```
sample(12).R = 8;
sample(12).C = 4:6;
sample(12).name = 'Glu-ANSSFAPK-high';
```

```
sample(13).R = 9;
sample(13).C = 1:3;
sample(13).name = 'Glu-TTYVVNPK-low';
```

```
sample(14).R = 9;
sample(14).C = 4:6;
sample(14).name = 'Glu-DAHATLRK-low';
```

```
sample(15).R = 10;
sample(15).C = 1:3;
sample(15).name = 'Glu-DAHATLRK-med';
```

```
%%%%%%%%
```

```
sample(16).R = 10;
sample(16).C = 4:6;
sample(16).name = 'A-antigen';
```

```
sample(17).R = 11;
sample(17).C = 1:3;
sample(17).name = 'B-antigen';
```

```
sample(18).R = 11;
sample(18).C = 4:6;
sample(18).name = 'O-antigen';
```

```
sample(19).R = 5;
sample(19).C = 1:3;
sample(19).name = 'alpha-Gal';
```

```
sample(20).R = 12;
sample(20).C = 1:6;
sample(20).name = 'BSA';
```

```
sample(21).R = 5;
sample(21).C = 4:6;
sample(21).name = 'buffer';
```

```

%%%% get empty spots intensity for each block

empty_block=zeros([1 24]);

for i=1:24 %loop through each block
    empty = zeros([1 7]);

    for j=1:NofIgGPoints %get F532 at specific column and row (that contains
empty spots)

        if block(j)==i && row(j)==1 && column(j)==3
            empty(1)=IgG(j); %then record value to "empty" matrix
        end

        if block(j)==i && row(j)==1 && column(j)==4
            empty(2)=IgG(j);
        end

        if block(j)==i && row(j)==1 && column(j)==5
            empty(3)=IgG(j);
        end

        if block(j)==i && row(j)==13 && column(j)==2
            empty(4)=IgG(j);
        end

        if block(j)==i && row(j)==13 && column(j)==3
            empty(5)=IgG(j);
        end

        if block(j)==i && row(j)==13 && column(j)==4
            empty(6)=IgG(j);
        end

        if block(j)==i && row(j)==13 && column(j)==5
            empty(7)=IgG(j);
        end
    end
    average = mean(empty); %average all empty values

    %assign averaged empty values according to block number
    empty_block(i)=average;

end

% Get IgG intensity from every sample in CS-35 blocks, then plot
figure(1);

for jj = 1:numel(sample) % loop through all samples

    for B = block_CS35 %block_CS35 %go through all data that has desired block
number

        rep = 0; % blank the number of replicates

        for C = sample(jj).C % loop replicates in current block

            rep = rep + 1; % increase the replicate by 1

            for i = 1:NofIgGPoints %look at all points in matrix A

                %if block, row, column is desired
                if row(i)==sample(jj).R && block(i)==B && column(i)==C

                    % find location of the block value in the block_CS35

```



```

        %array at the location = L

        sample(jj).signal_906(rep,L) = IgG(i) - empty_block(B);
        %disp(sample(jj).signal_906(rep,L))
    end
end
end
end
end
end

%Fitting

for jj = 1:numel(sample) % loop through all samples

    if jj<=9
        figure(1)
        subplot(3,3,jj);
    else
        figure(2)
        subplot(3,4,jj-9);
    end

    hold on;

    SIZE = 4;

    plot( conc_CS35, sample(jj).signal_CS35, ...
          'ok', ...
          'MarkerFaceColor', 0.8*[1 0 0],...
          'MarkerSize', SIZE );

    plot( conc_CS40, sample(jj).signal_CS40, ...
          'ok', ...
          'MarkerFaceColor', turquoise,...
          'MarkerSize', SIZE );

    plot( conc_906, sample(jj).signal_906, ...
          'ok', ...
          'MarkerFaceColor', chartreuse,...
          'MarkerSize', SIZE );

    title([ 'Row: ' num2str(sample(jj).R) ...
            ' Columns: ' num2str(sample(jj).C) ' ' ...
            sample(jj).name], 'FontSize', 6 );

    set( gca, 'YScale', 'linear', 'XScale', 'log',...
          'TickDir', 'out','XMinorTick','on', 'YMinorTick','on',...
          'XTick', [0.01 0.1 1 10 100], 'FontSize', 6);

    ylim([Ymin Ymax]);
    xlim([Xmin Xmax]);

end

%%
%%% this is where we re-loop through all the samples after we have plotted
%%% them and fit each subplot to an IC50 equation.
figure(1);

% first define the grid that will be used for plotting of the solutions

N = 100; % # of points to plot (pick a larger number for a smooth curve
grid = 10.^(log10(Xmin): (log10(Xmax)-log10(Xmin))/N :log10(Xmax));

```

```

% for jj = 1:9      % instead of numel(sample), lets just loop through first 9
for jj = 1:numel(sample) % loop through all samples

    % first you have to define the X points (concentrations) and the
    % y-points (intensities) as one continous vector

    if jj<=9
        figure(1)
        subplot(3,3,jj);
    else
        figure(2)
        subplot(3,4,jj-9);
    end

    y1 = []; y2 = []; y3 = [];
    x1 = []; x2 = []; x3 = [];

    for i = 1:3
        y1 = [y1 sample(jj).signal_CS35(i,1:end) ];
        y2 = [y2 sample(jj).signal_CS40(i,1:end) ];
        y3 = [y3 sample(jj).signal_906(i,1:end) ];

        x1 = [x1 conc_CS35(1:end) ];
        x2 = [x2 conc_CS40(1:end) ];
        x3 = [x3 conc_906(1:end) ];
    end

    %%%% define the fit options      %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    min=0;
    MAX=10000;
    EC50=1;
    N=1;

    s = fitoptions('Method','NonlinearLeastSquares',...
        'Lower',    [0,    0,    0,    0],...
        'Upper',    [1000, 60000, 2, 1],...
        'Startpoint',[EC50, MAX, N, min],...
        'TolFun', 1e-10 );

    %%%% define the fit equation

    equation = ('(min + (MAX-min)/(1+(EC50/x)^N)')');

    ft = fittype( equation,'options',s );

    %%%% fit the data      %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    [sample(jj).fit_CS35, sample(jj).gof2_CS35] = fit(x1',y1',ft);
    sample(jj).CON_CS35 = confint(sample(jj).fit_CS35);

    [sample(jj).fit_CS40, sample(jj).gof2_CS40] = fit(x2',y2',ft);
    sample(jj).CON_CS40 = confint(sample(jj).fit_CS40);

    [sample(jj).fit_906, sample(jj).gof2_906 ] = fit(x3',y3',ft);
    sample(jj).CON_906 = confint(sample(jj).fit_906);

    plot( grid, sample(jj).fit_CS35(grid),'r');
    plot( grid, sample(jj).fit_CS40(grid),'color',turquoise);
    plot( grid, sample(jj).fit_906(grid) , 'color',chartreuse);

    TL{1} = [' IC50 (CS35) = ' ...
        num2str( sample(jj).fit_CS35.EC50 ,'%0.1f') ...
        ' [' num2str( sample(jj).CON_CS35(:,1) ,'% 0.1f') ' ]'];

    TL{2} = [' IC50 (CS40) = ' ...

```

```

        num2str( sample(jj).fit_CS40.EC50 ,'%0.1f') ...
    ' [' num2str( sample(jj).CON_CS40(:,1) ,'% 0.1f')  ']''];

TL{3} = [' IC50(906) = ' ...
        num2str( sample(jj).fit_906.EC50 ,'%0.1f') ...
        ' [' num2str( sample(jj).CON_906(:,1) ,'% 0.1f')  ']''];

legend off;

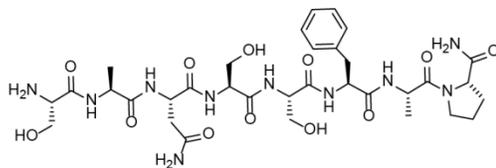
text(Xmin,0.85*Ymax, char(TL),'FontSize',6);

drawnow;

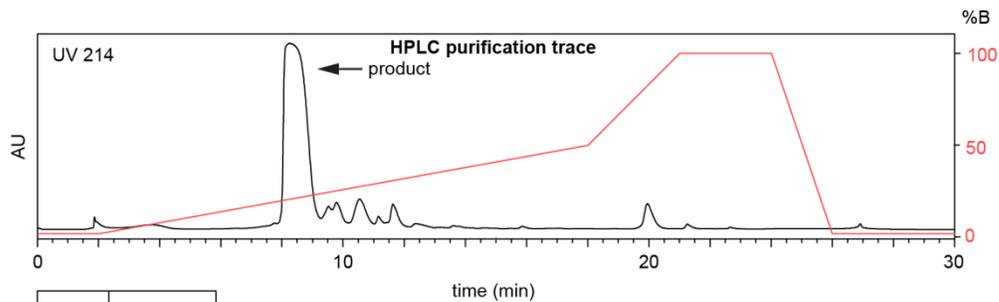
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end

```

SANSSFAP

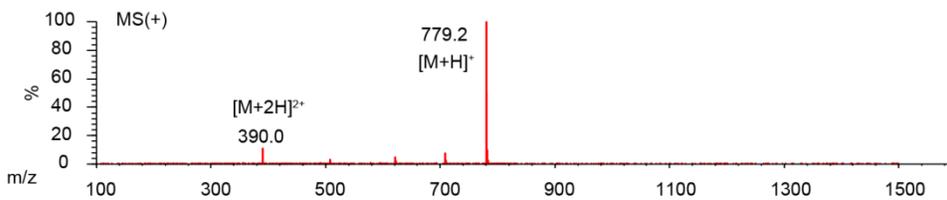
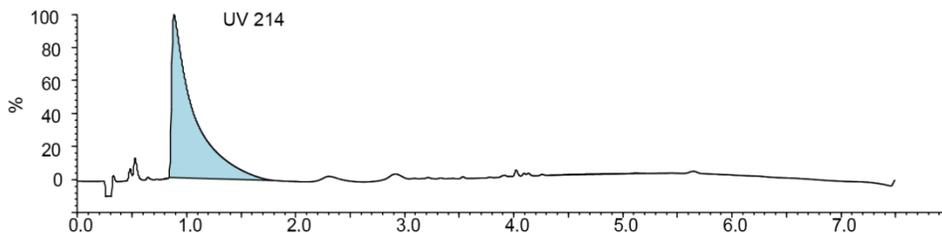


Chemical Formula: $C_{33}H_{50}N_{10}O_{12}$
Exact Mass: 778.36

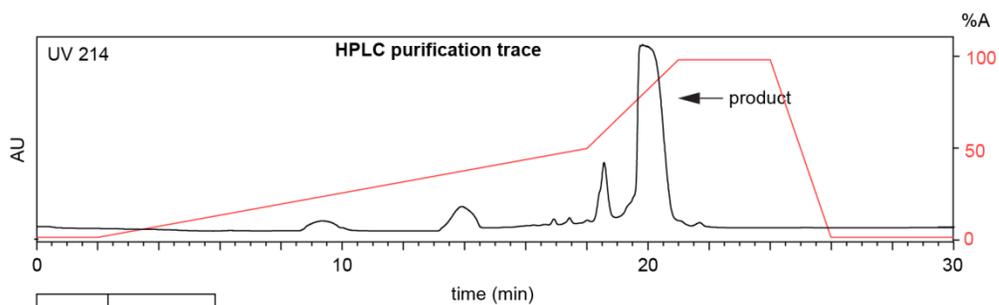
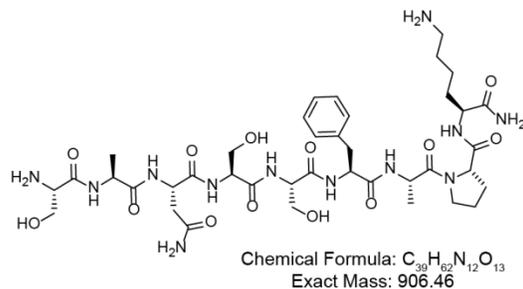


Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

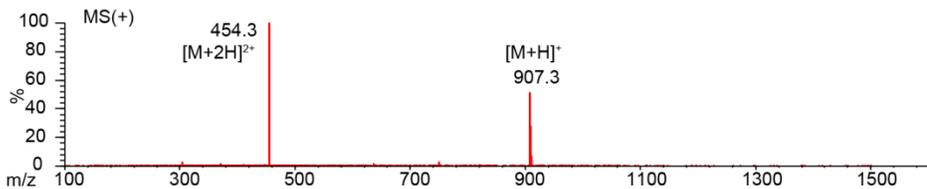
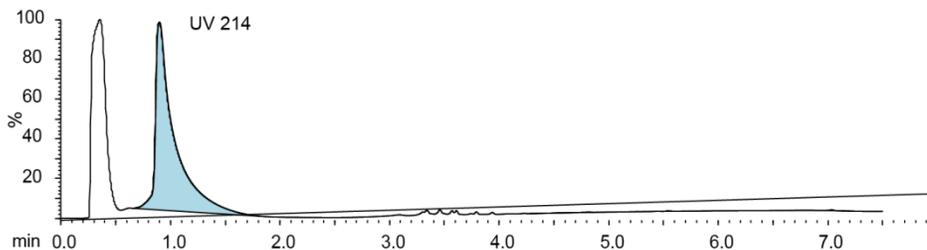


SANSSFAPK

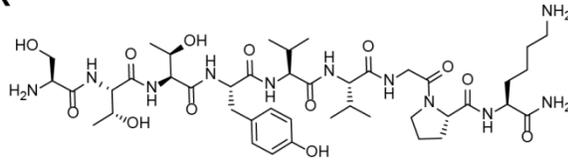


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

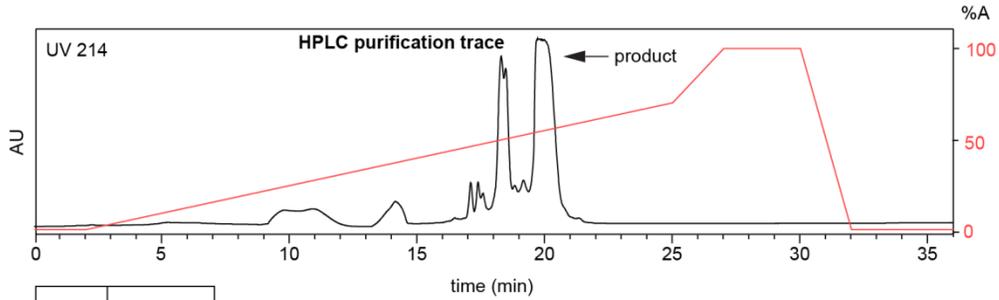
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μ m, 4.6 mm X 250 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



STTYVNPVK

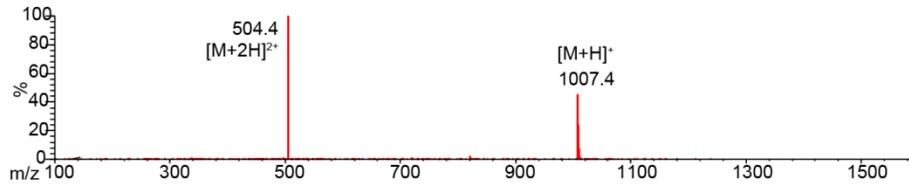
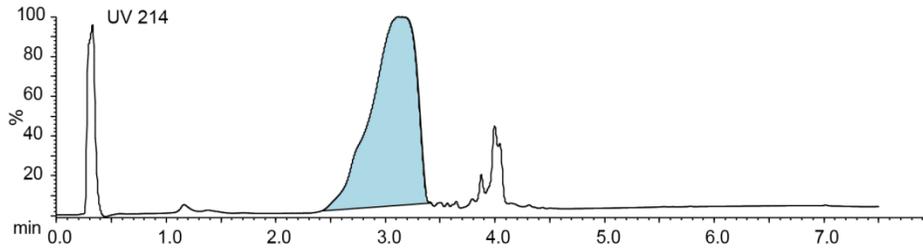


Chemical Formula: $C_{45}H_{74}N_{12}O_{14}$
 Exact Mass: 1006.54

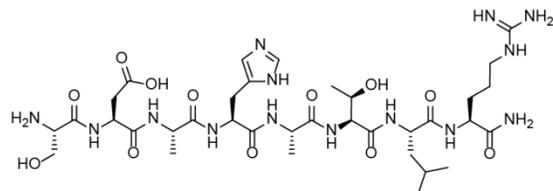


Time (min)	Solvent A (%)
0	0
2	0
25	70
27	100
30	100
32	0
36	0

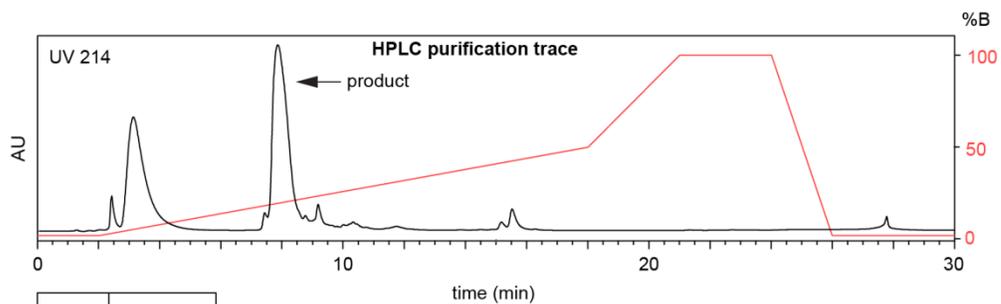
Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μ m, 4.6 mm X 250 mm)
 Solvent A: $H_2O + 0.1\%$ (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



SDAHATLR

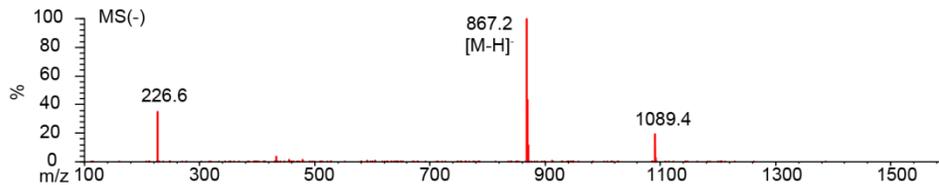
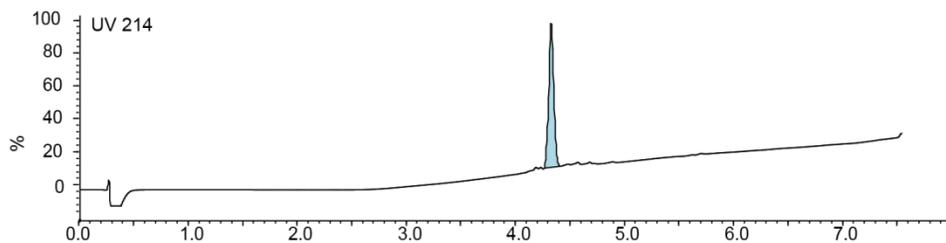


Chemical Formula: $C_{35}H_{60}N_{14}O_{12}$
Exact Mass: 868.45

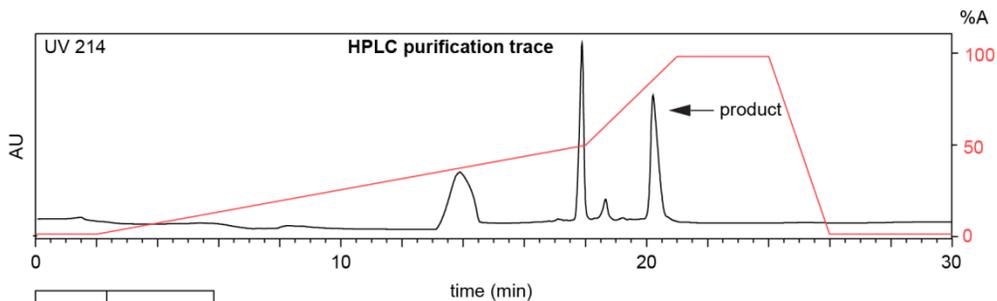
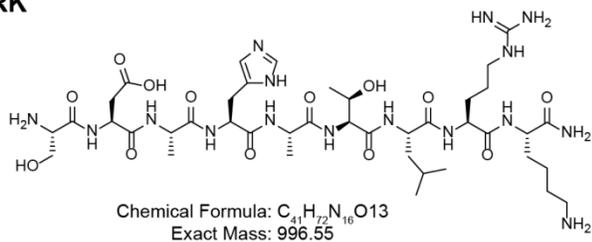


Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

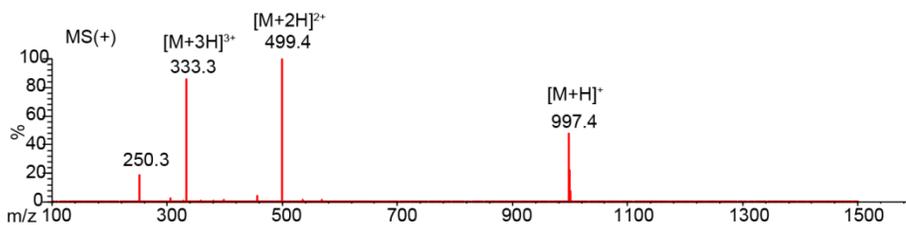
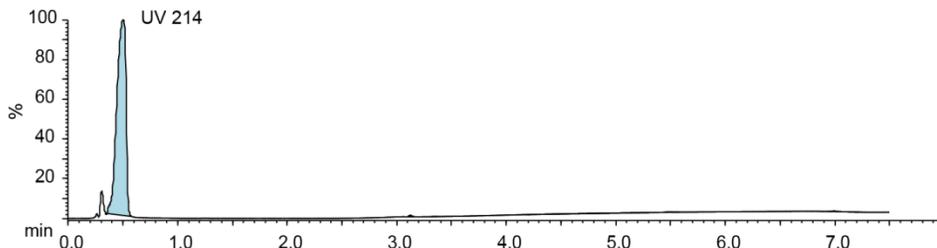


SDAHATLRK

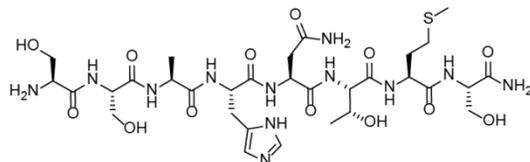


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

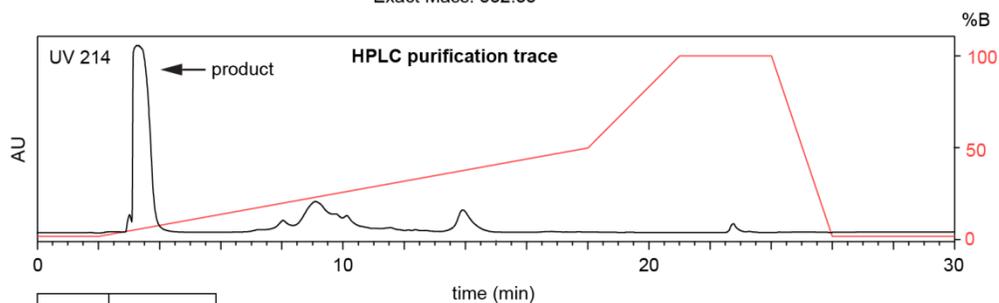
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μm, 4.6 mm X 250 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



SSAHNTMS

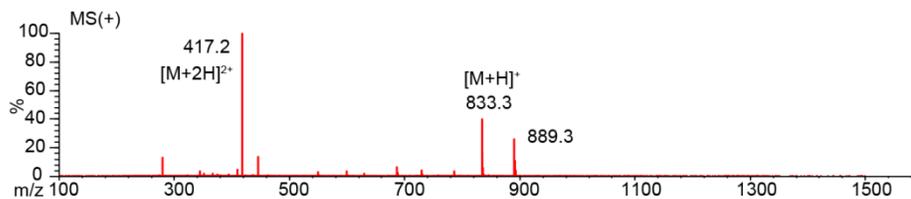
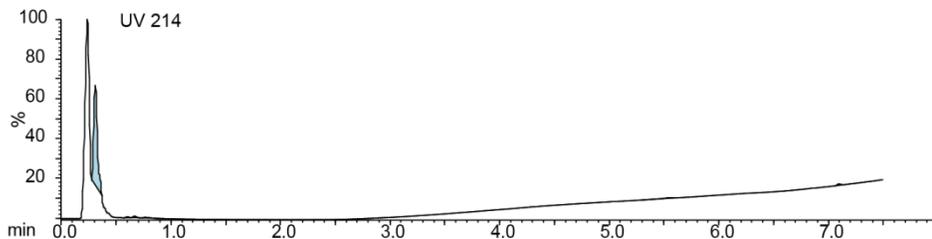


Chemical Formula: C₃₁H₅₂N₁₂O₁₃S
 Exact Mass: 832.35

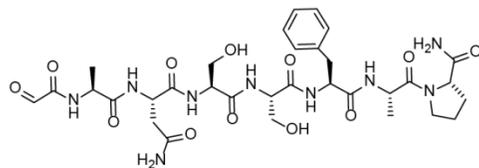


Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

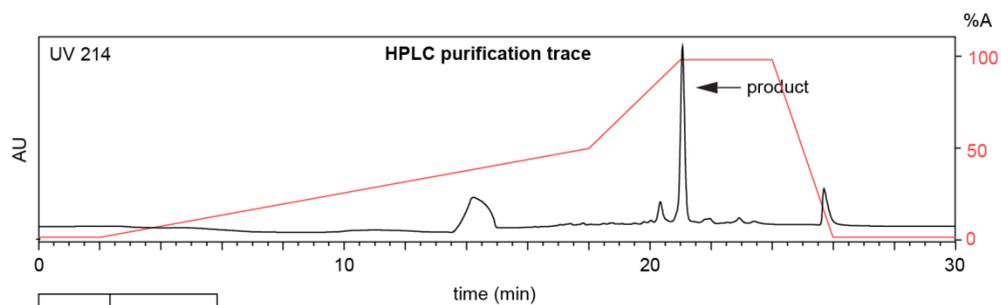
Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



ox-ANSSFAP



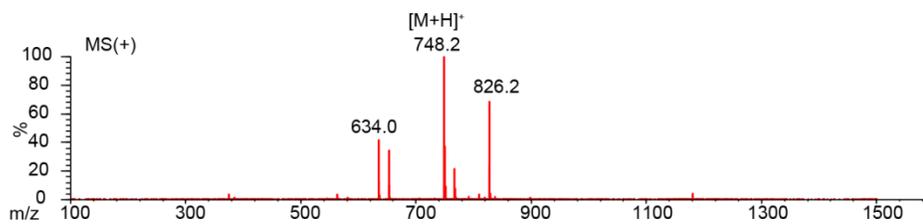
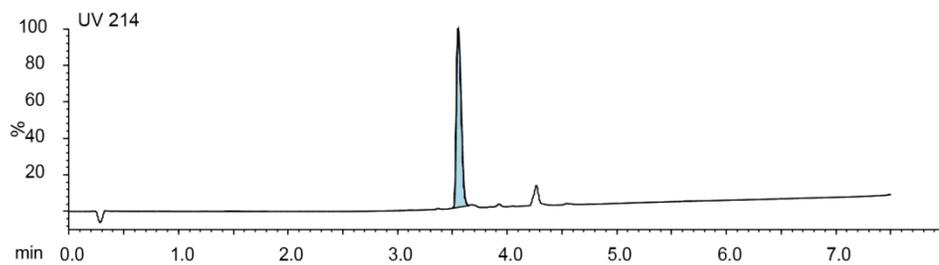
Chemical Formula: $C_{32}H_{45}N_9O_{12}$
Exact Mass: 747.32



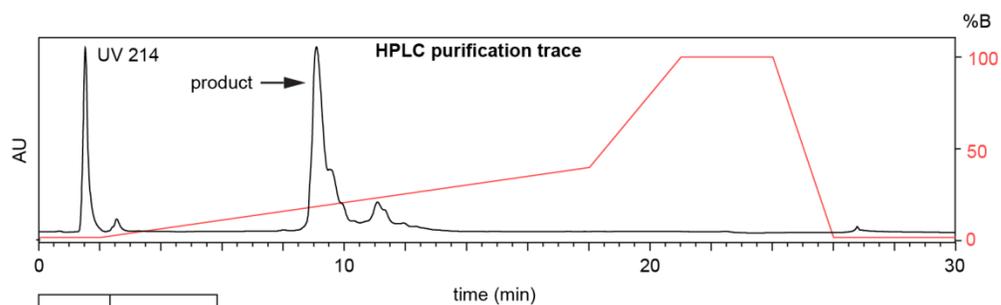
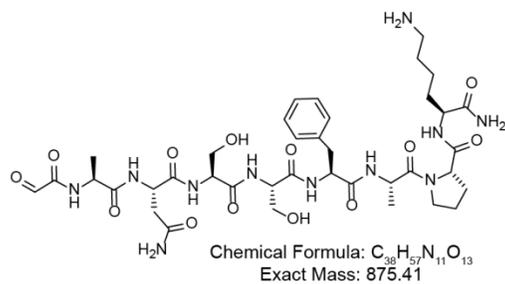
Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μm, 4.6 mm X 250 mm)

Solvent A: $H_2O + 0.1\%$ (v/v) TFA
Solvent B: $MeCN + 0.1\%$ (v/v) TFA

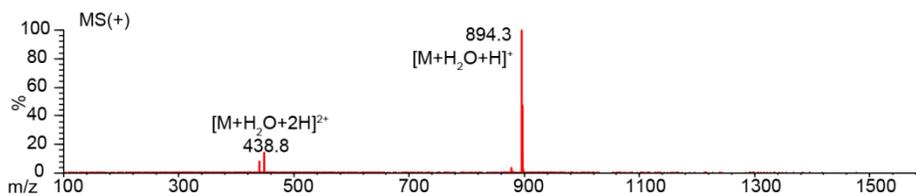
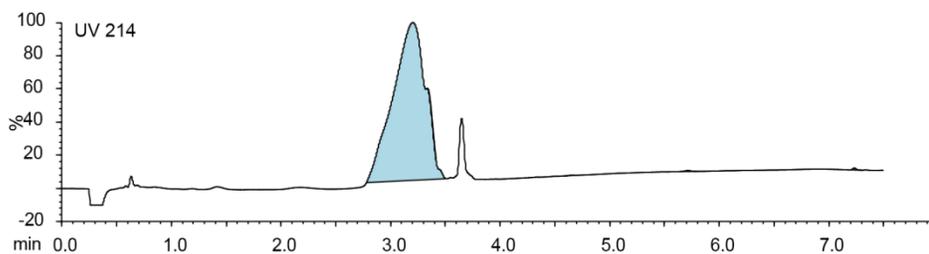


ox-ANSSFAPK

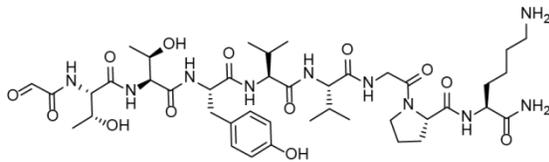


Time (min)	Solvent B (%)
0	2
2	2
18	40
21	100
24	100
26	2
30	2

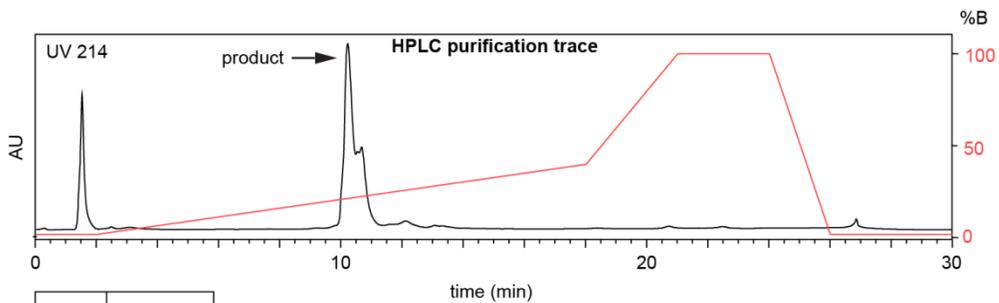
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



ox-TTYVVNPK

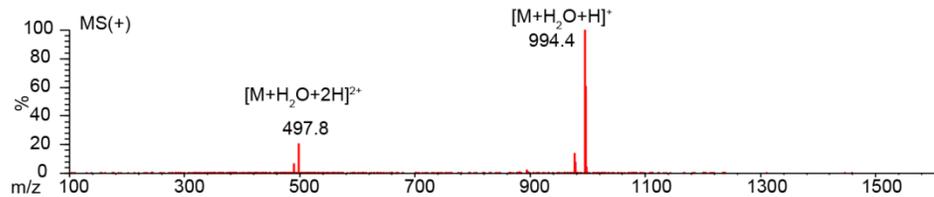
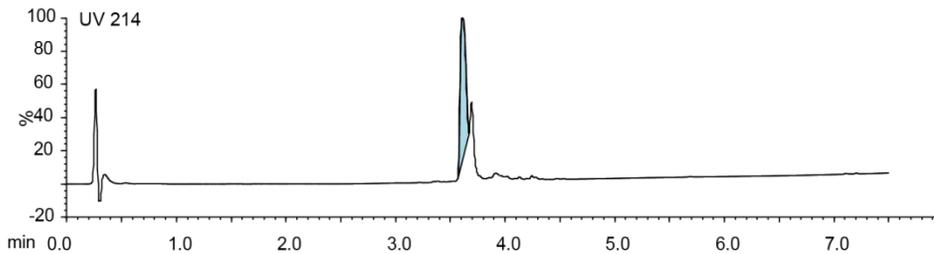


Chemical Formula: $C_{44}H_{69}N_{11}O_{14}$
 Exact Mass: 975.50

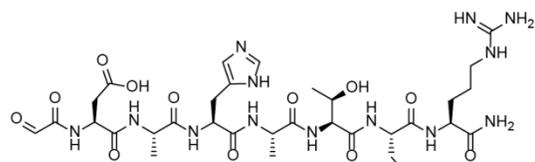


Time (min)	Solvent B (%)
0	2
2	2
18	40
21	100
24	100
26	2
30	2

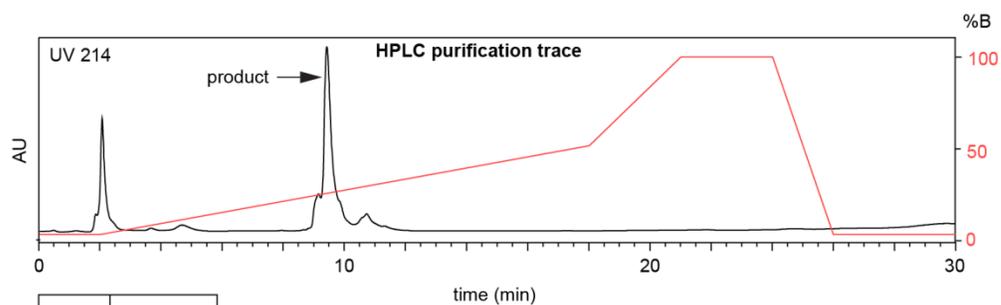
Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



ox-DAHATLR

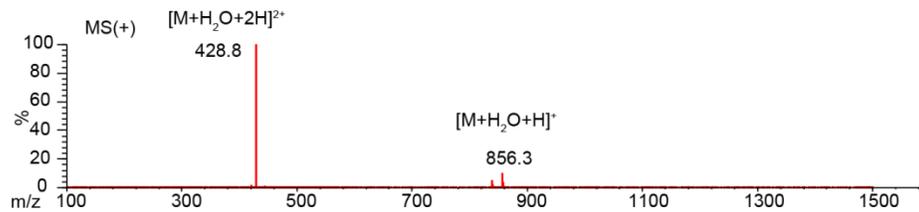
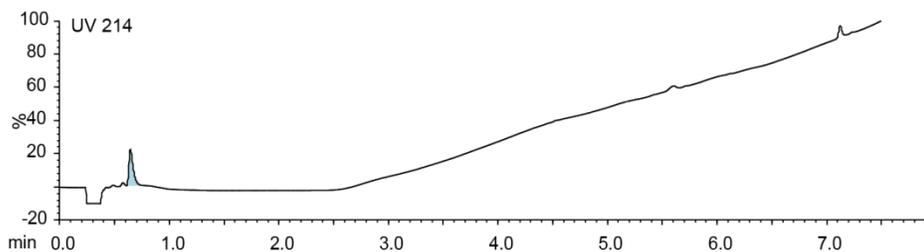


Chemical Formula: $C_{34}H_{55}N_{13}O_{12}$
 Exact Mass: 837.41

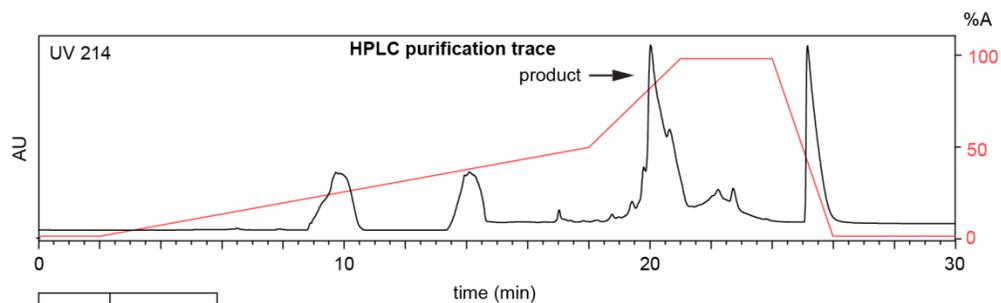
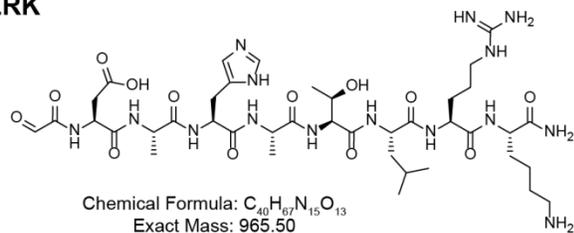


Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: $H_2O + 0.1\%$ (v/v) TFA
 Solvent B: $MeCN + 0.1\%$ (v/v) TFA

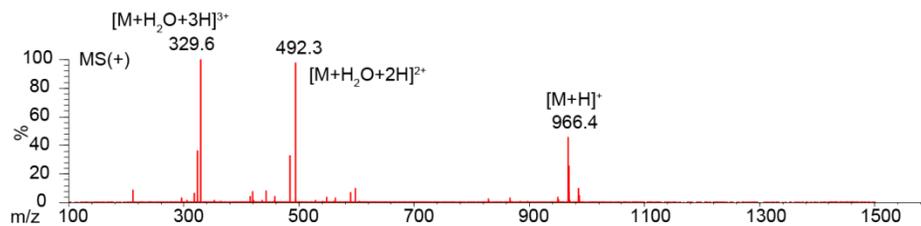
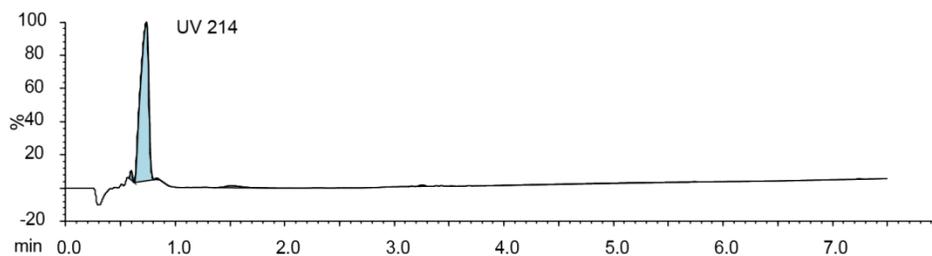


ox-DAHATLRK

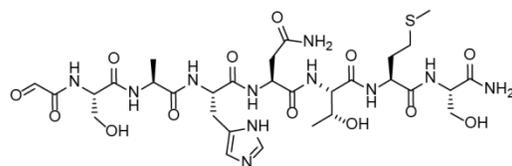


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

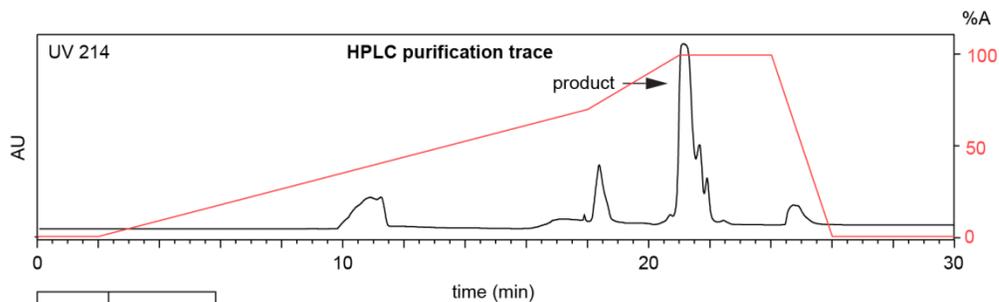
Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μm, 4.6 mm X 250 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



ox-SAHNTMS

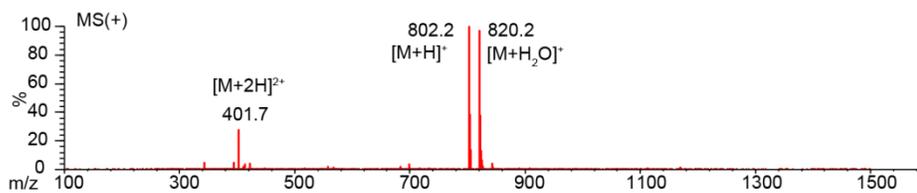
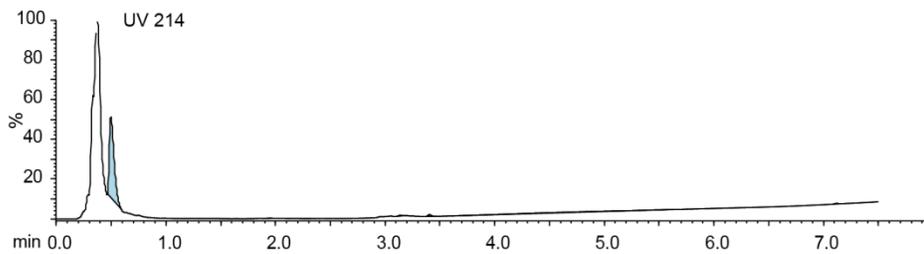


Chemical Formula: $C_{30}H_{47}N_{11}O_{13}S$
Exact Mass: 801.31

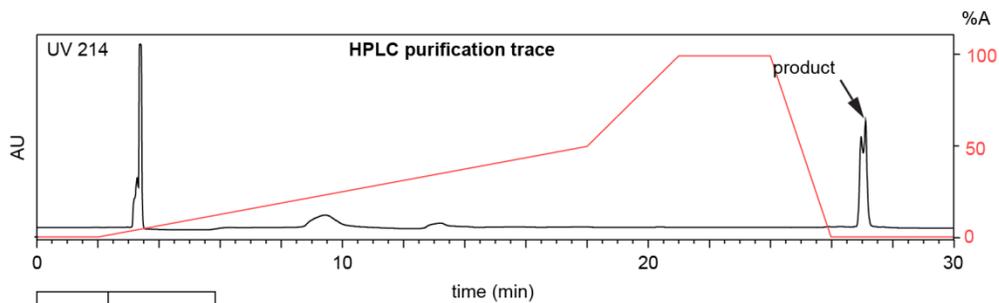
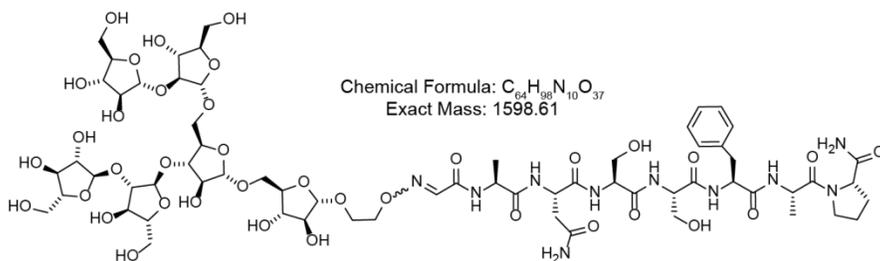


Time (min)	Solvent A (%)
0	0
2	0
18	70
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μ m, 4.6 mm X 250 mm)
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



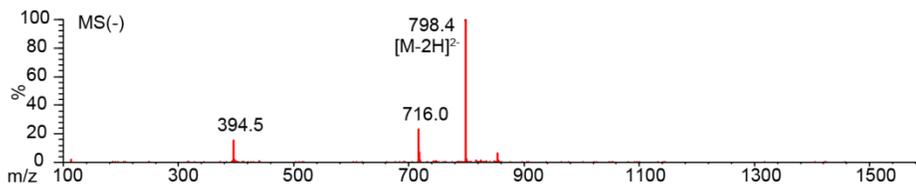
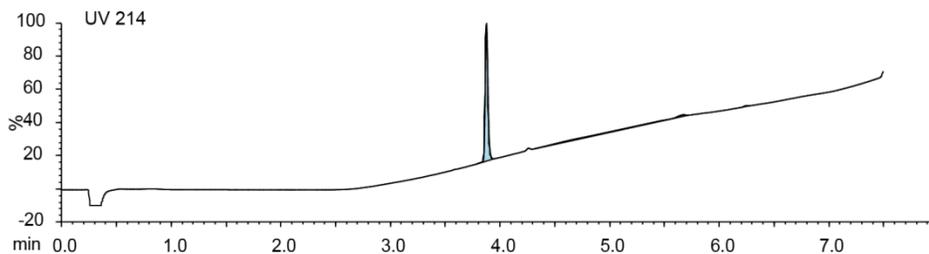
Ara₆-ANSSFAP



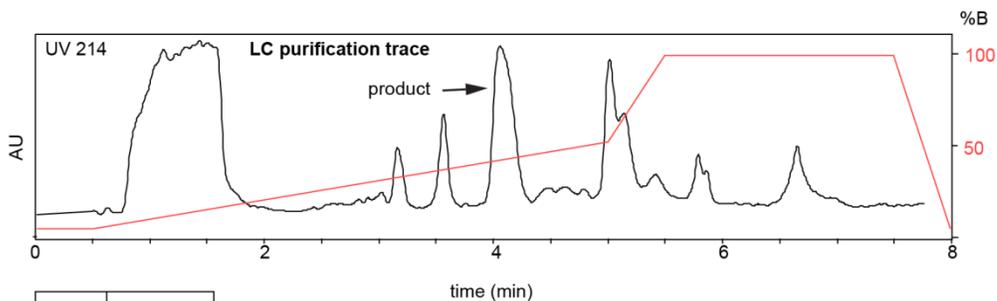
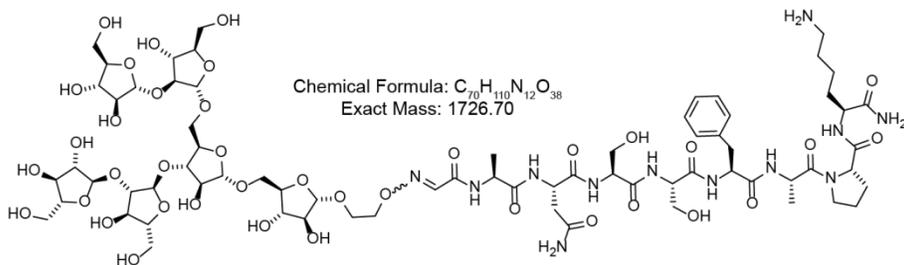
Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	2
30	2

Flow rate: 1 mL/min
 Amide Column
 (100Å, 3.5 μm, 4.6 mm X 75 mm)

Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

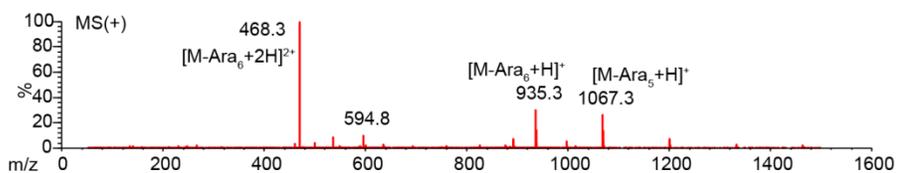
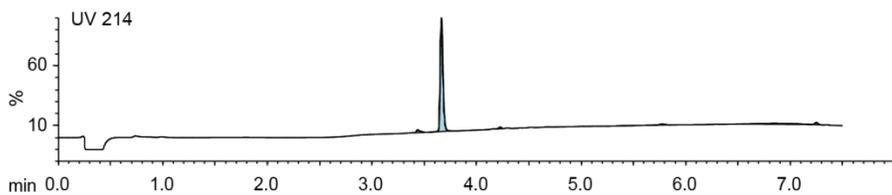


Ara₆-ANSSFAPK

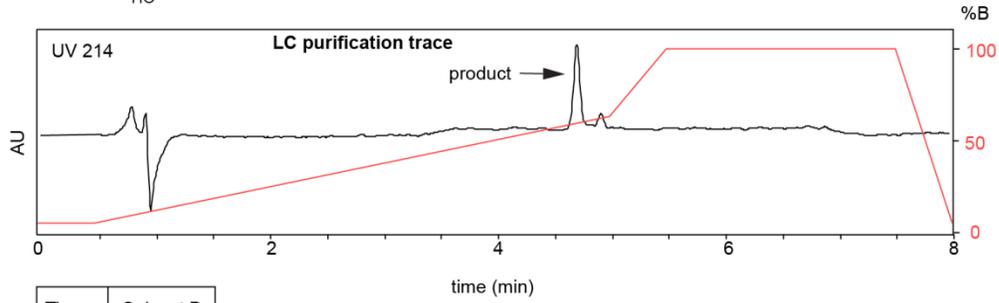
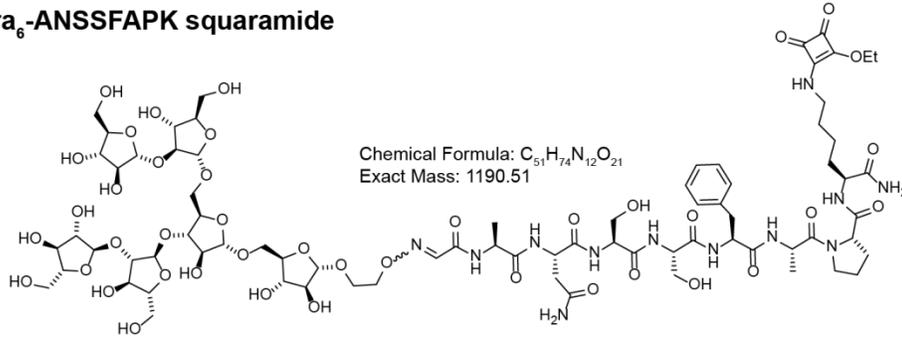


Time (min)	Solvent B (%)
0	5
0.5	5
5	50
5.5	95
7.5	95
8	5

Flow rate: 1 mL/min
 Symmetry C18 Column
 (100Å, 3.5 μm, 4.6 mm X 75 mm)
 Solvent A: H₂O + 0.1% (v/v) FA
 Solvent B: MeCN + 0.1% (v/v) FA

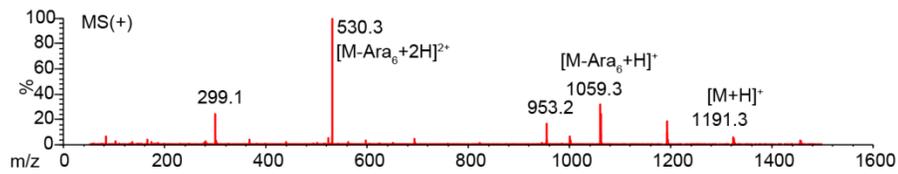
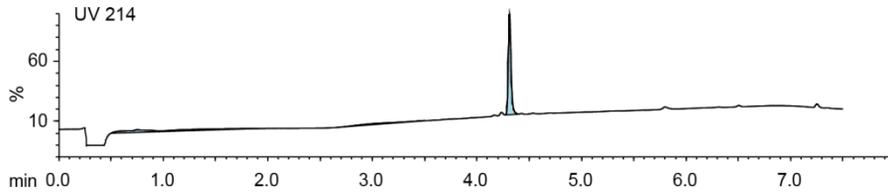


Ara₆-ANSSFAPK squaramide

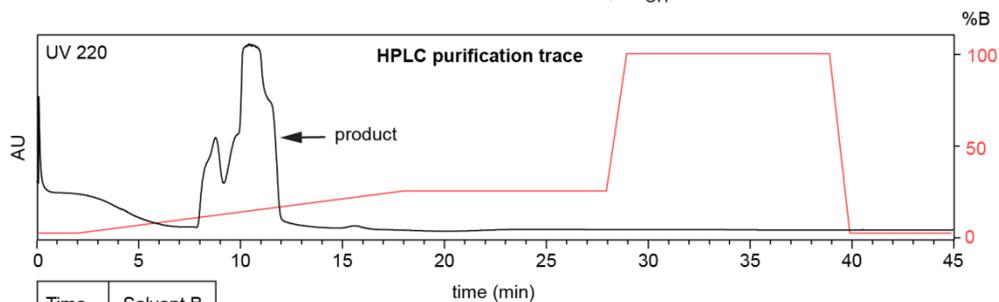
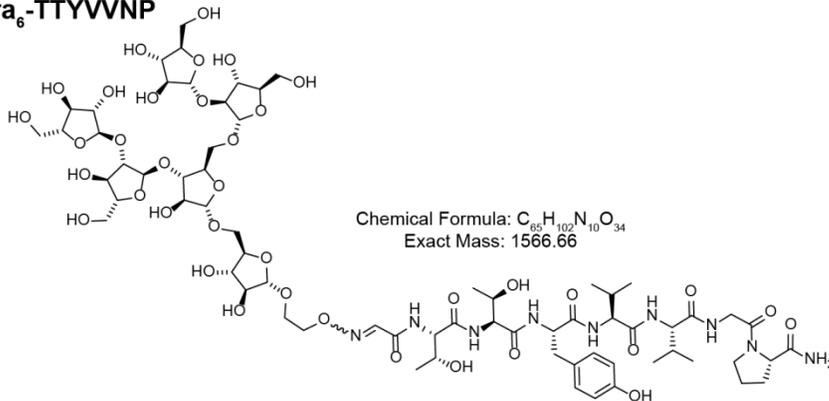


Time (min)	Solvent B (%)
0	5
0.5	5
5	60
5.5	95
7.5	95
8	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

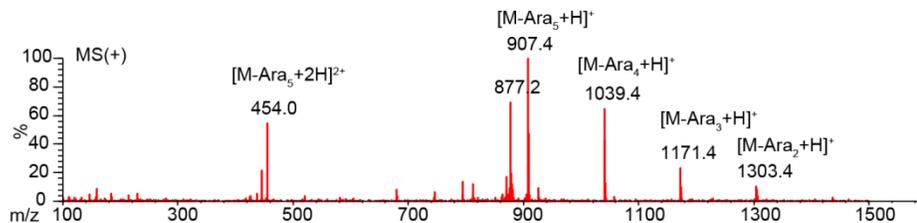
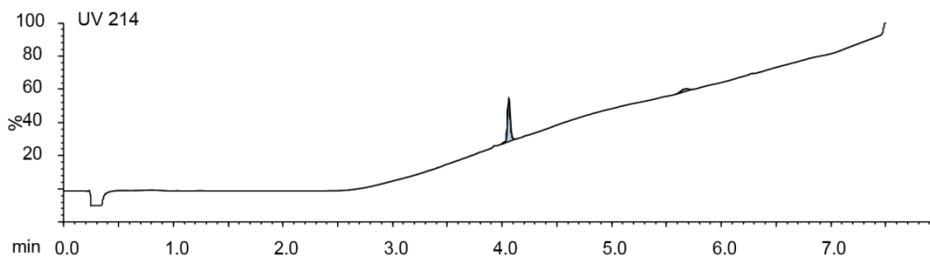


Ara₆-TTYWNP

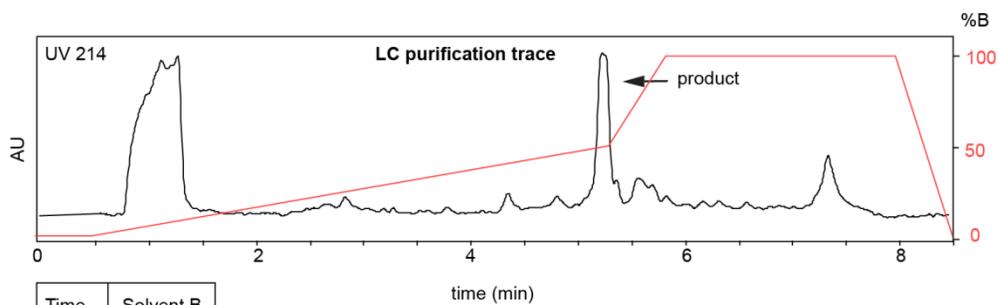
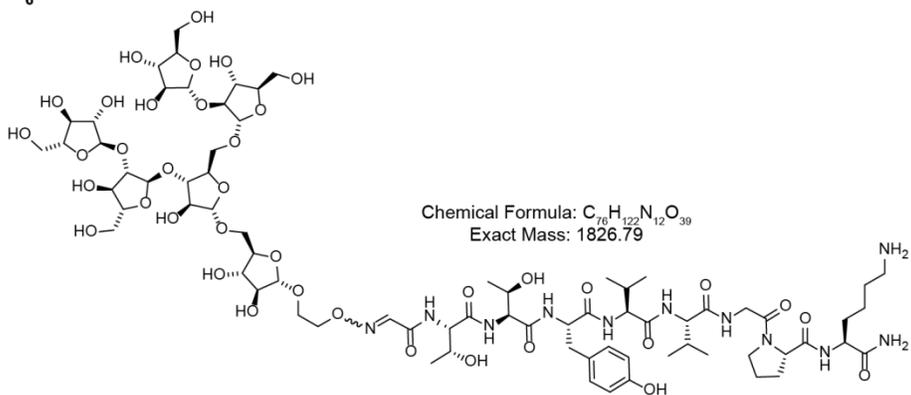


Time (min)	Solvent B (%)
0	2
2	2
18	25
28	25
29	100
39	100
40	2
45	2

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



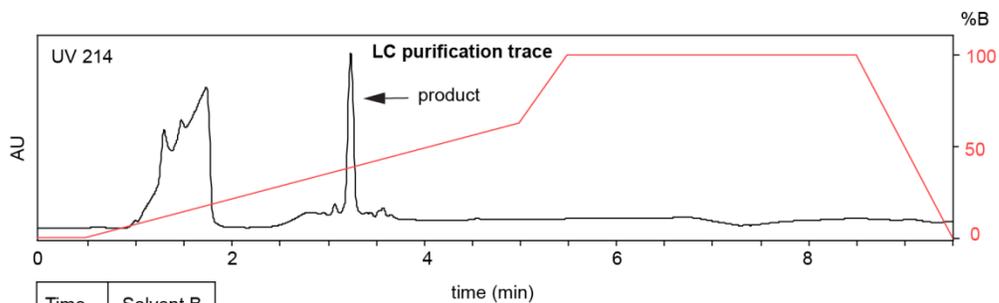
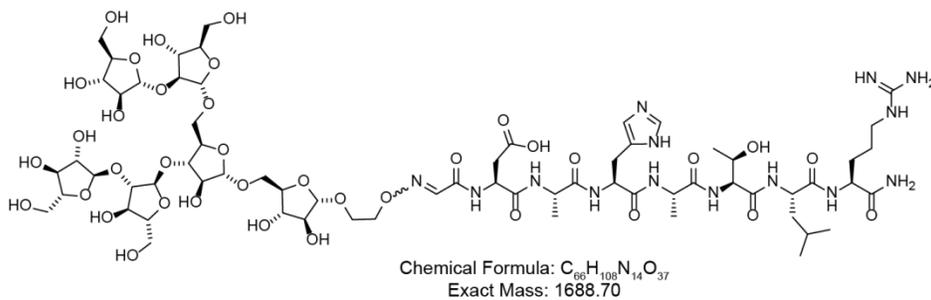
Ara₆-TTYWNPK



Time (min)	Solvent B (%)
0	5
0.5	5
5	50
5.5	95
7.5	95
8	5

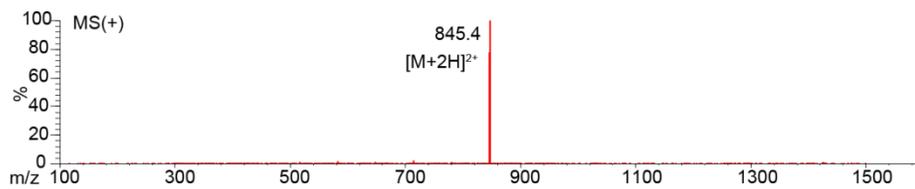
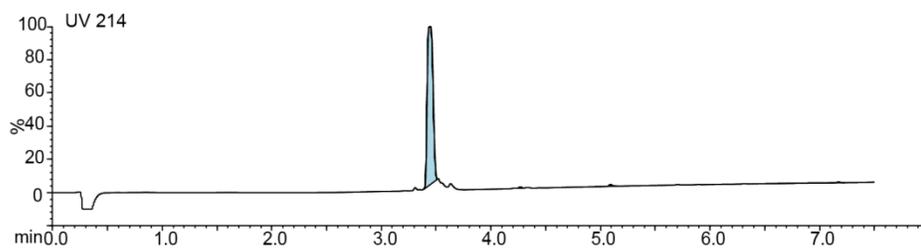
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

Ara₆-DAHATLR

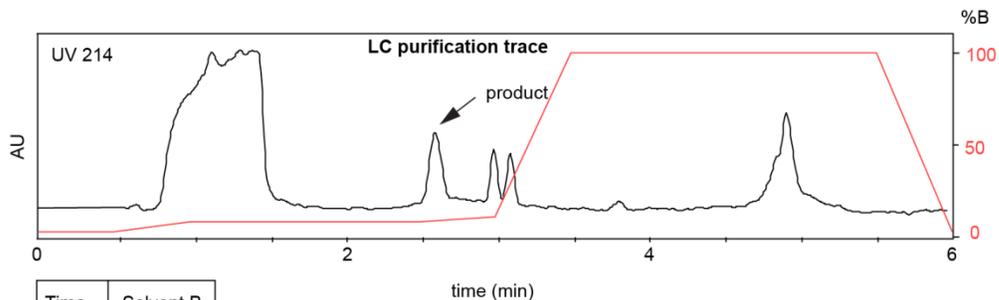
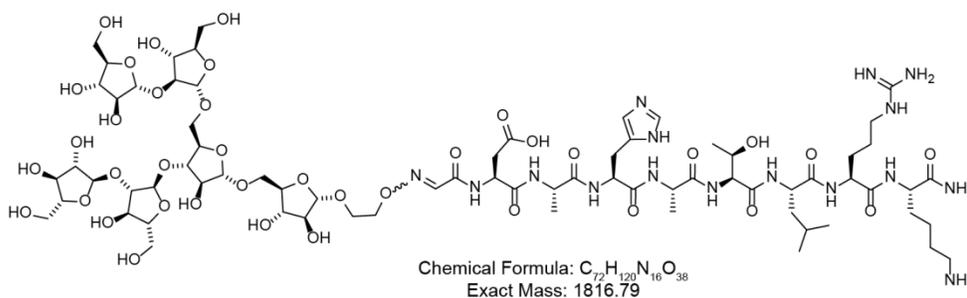


Time (min)	Solvent B (%)
0	1
0.5	1
5	60
5.5	95
8.5	95
9.5	1

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA



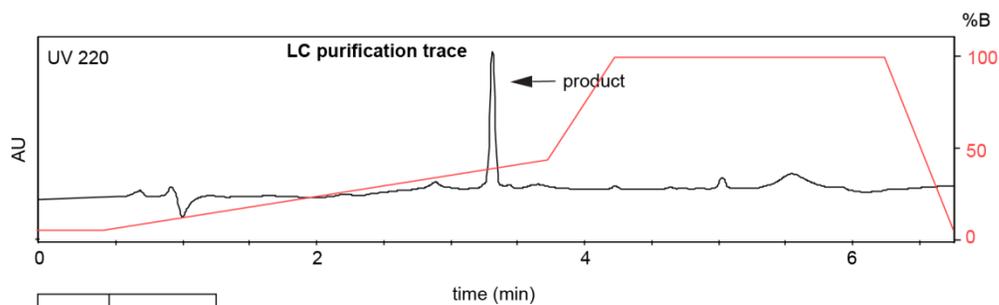
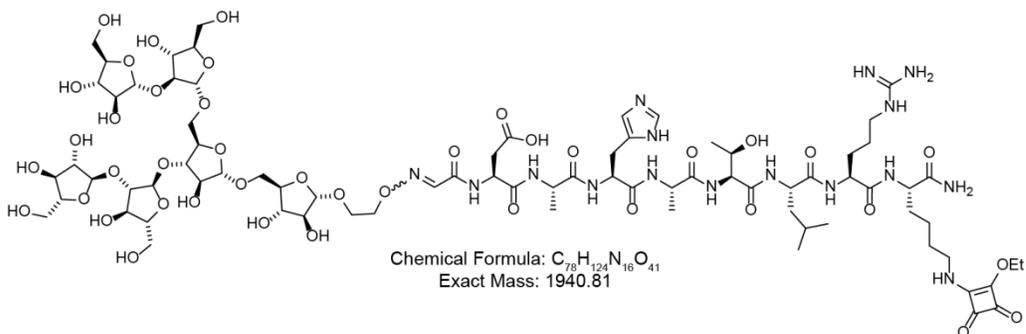
Ara₆-DAHATLRK



Time (min)	Solvent B (%)
0	5
0.5	5
1	10
2.5	10
3	12.5
3.5	95
5.5	95
6	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

Ara₆-DAHATLRK squaramide

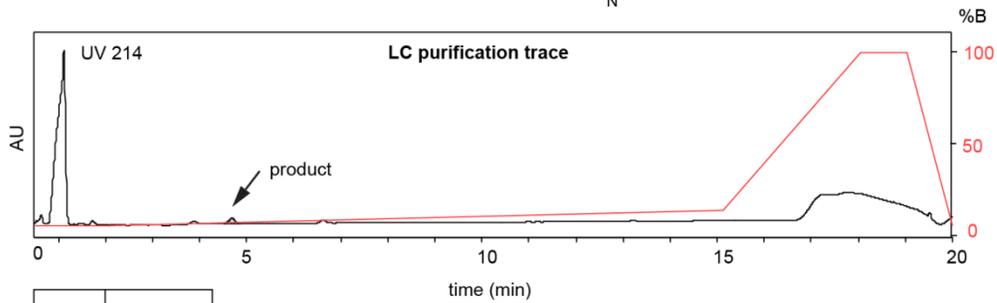
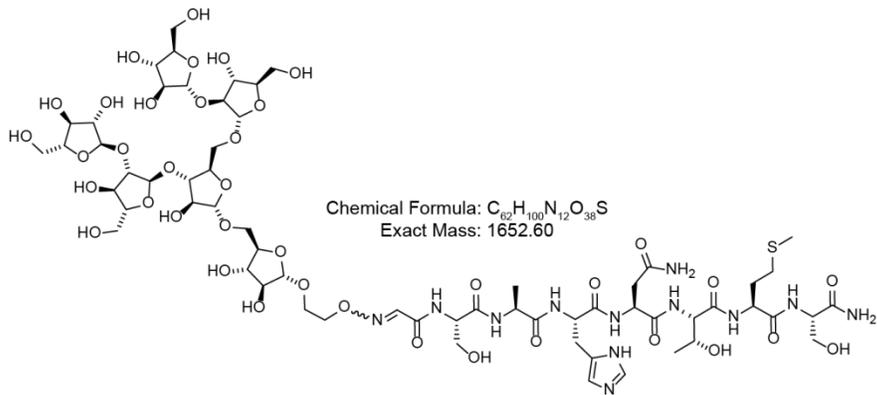


Time (min)	Solvent B (%)
0	5
0.5	5
3.8	41.6
4.3	95
6.3	95
6.8	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)

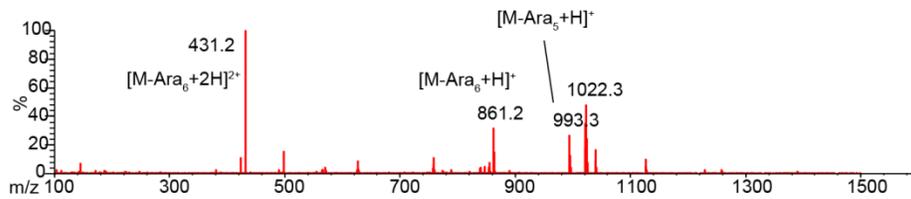
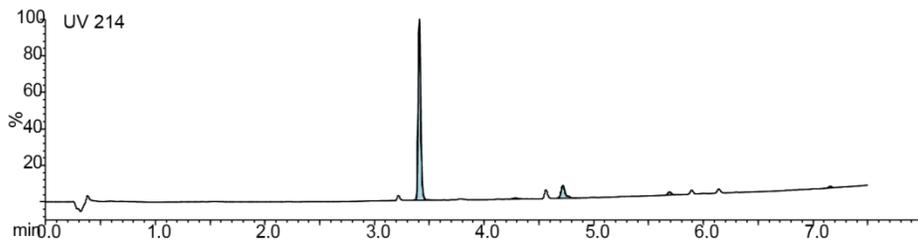
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

Ara₆-SAHTMS

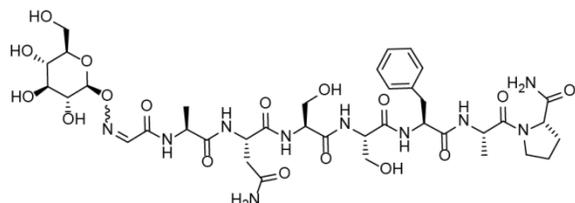


Time (min)	Solvent B (%)
0	5
1.5	5
15	13
18	95
19	95
20	5

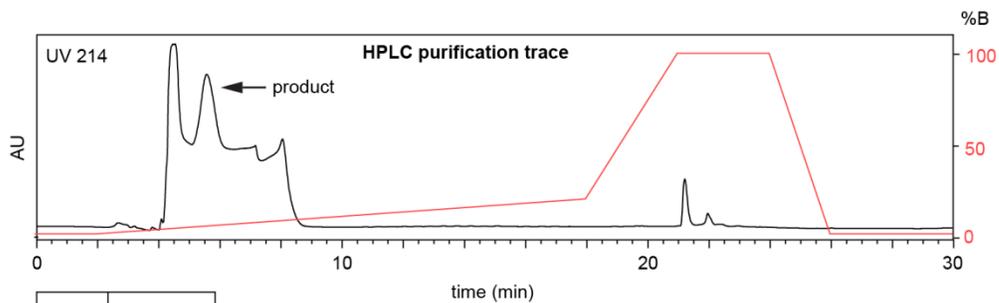
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μm, 4.6 mm X 75 mm)
Solvent A: H₂O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA



Glu-ANSSFAP

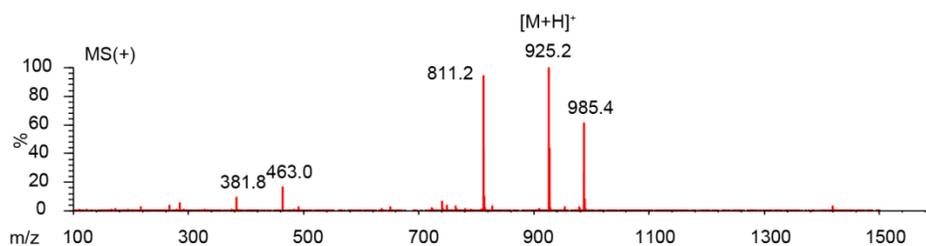
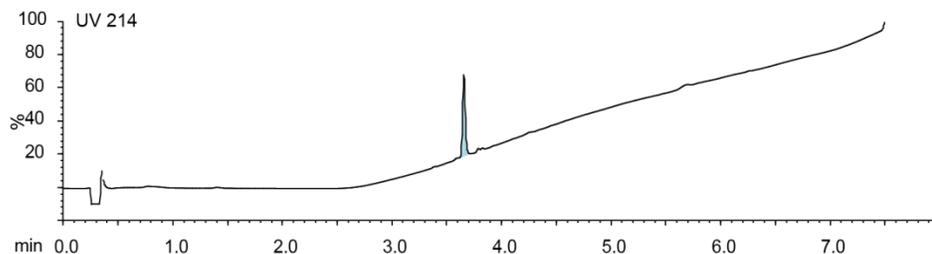


Chemical Formula: $C_{38}H_{50}N_{10}O_{17}$
 Exact Mass: 924.38

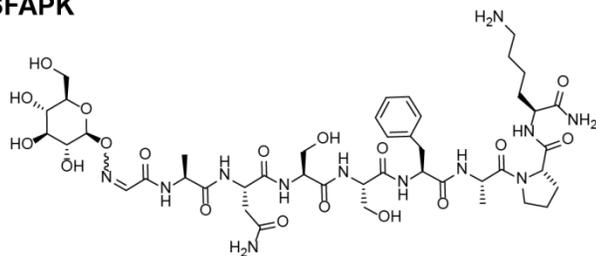


Time (min)	Solvent B (%)
0	2
2	2
18	20
21	100
24	100
26	2
30	2

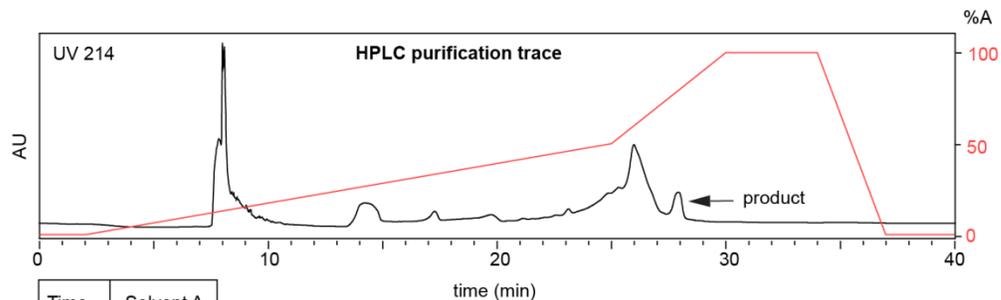
Flow rate: 1 mL/min
 Symmetry C18 Column
 (100Å, 5 µm, 4.6 mm X 150 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



Glu-ANSSFAPK

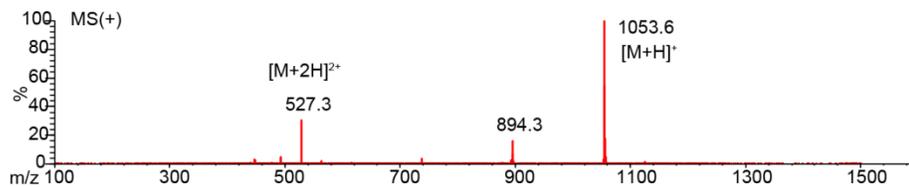
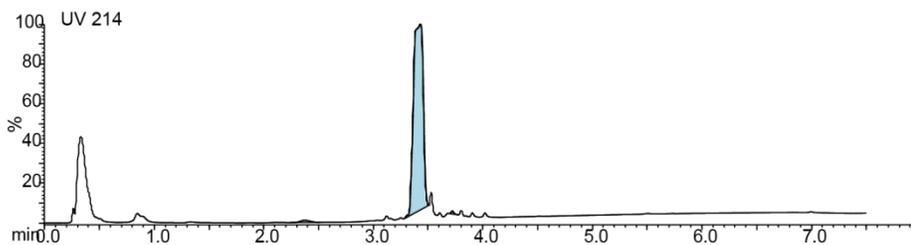


Chemical Formula: $C_{44}H_{68}N_{12}O_{18}$
 Exact Mass: 1052.48

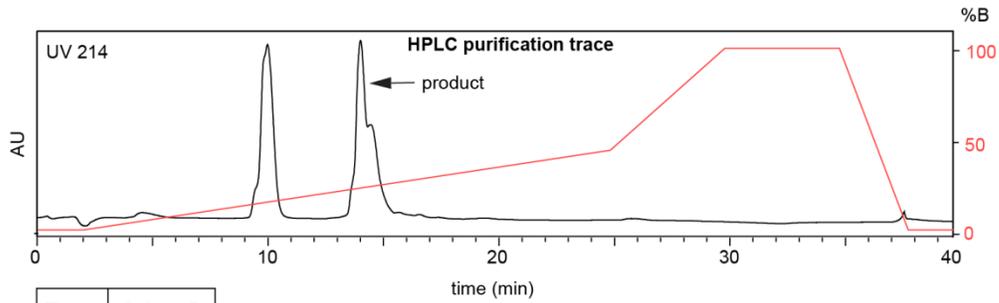
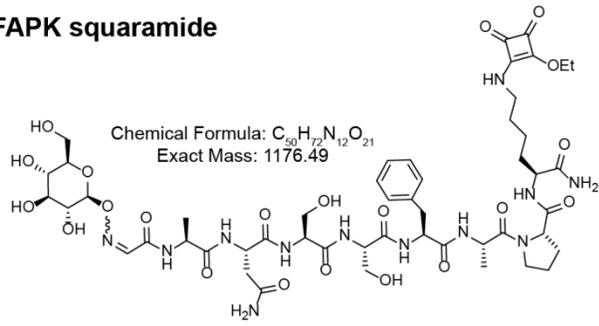


Time (min)	Solvent A (%)
0	0
2	0
25	50
30	100
34	100
37	0
40	0

Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μm, 4.6 mm X 250 mm)
 Solvent A: $H_2O + 0.1\%$ (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

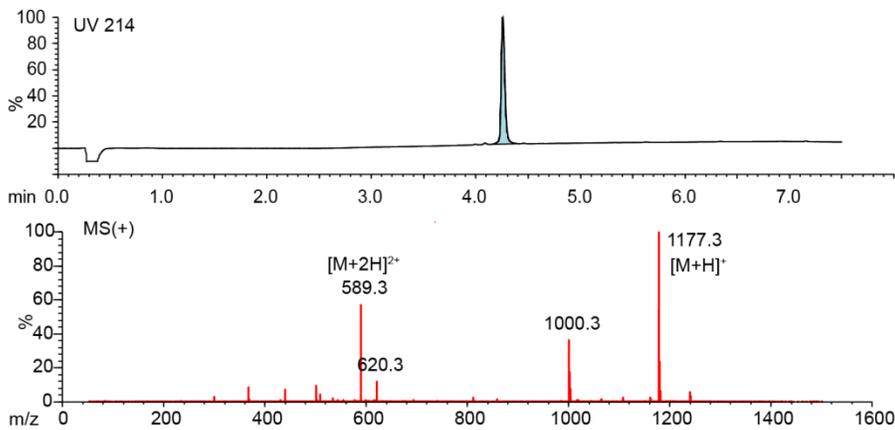


Glu-ANSSFAPK squaramide

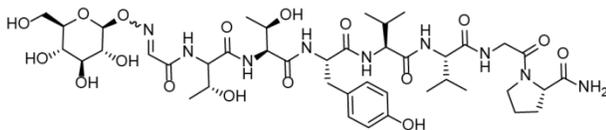


Time (min)	Solvent B (%)
0	2
2	2
22	35
27	100
32	100
36	2
40	2

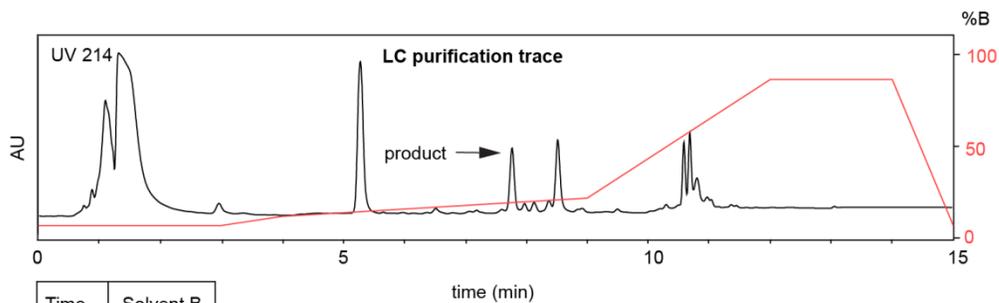
Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



Glu-TTYVVNP

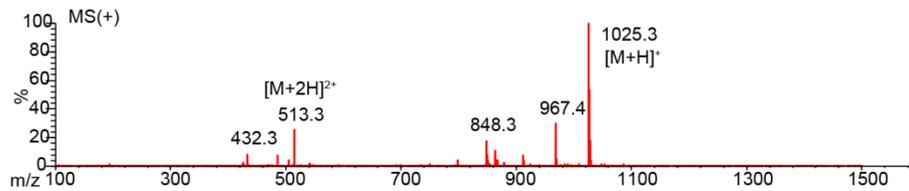
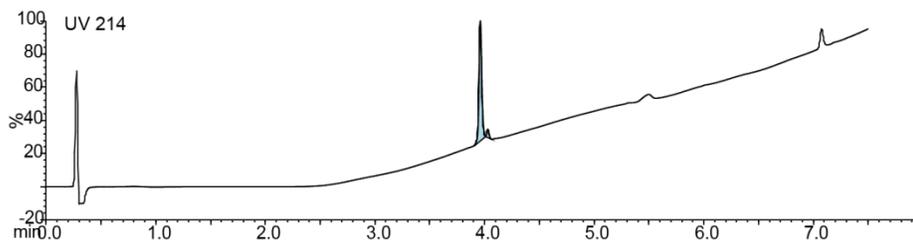


Chemical Formula: $C_{441}H_{687}N_{10}O_{18}$
 Exact Mass: 1024.47

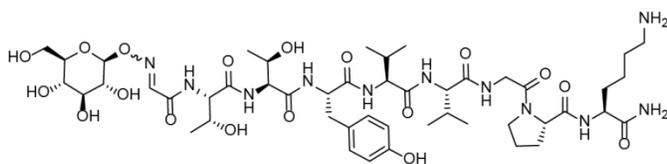


Time (min)	Solvent B (%)
0	5
3	5
4	10
9	20
12	85
24	85
15	5

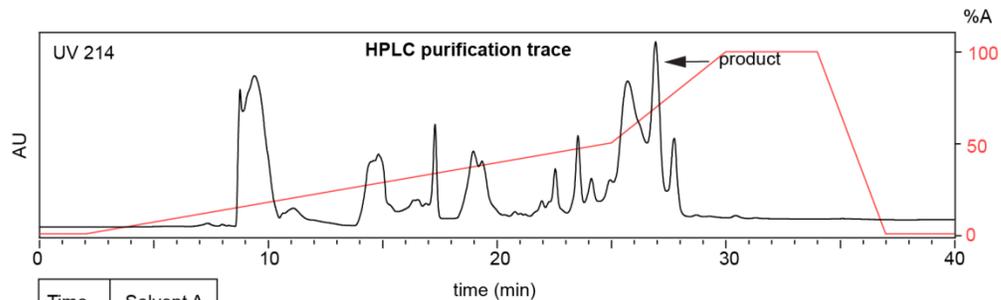
Flow rate: 1 mL/min
 Symmetry C18 Column
 (100Å, 3.5 μm, 4.6 mm X 75 mm)
 Solvent A: H₂O + 0.1% (v/v) FA
 Solvent B: MeCN + 0.1% (v/v) FA



Glu-TTYVVNPK

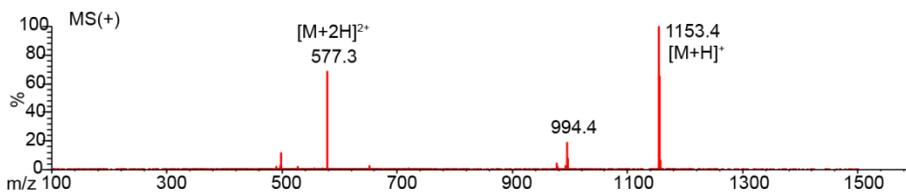
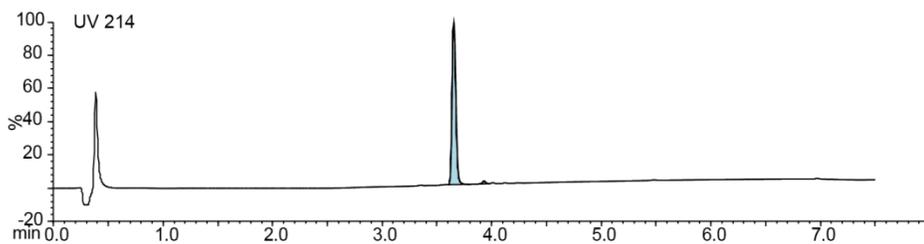


Chemical Formula: $C_{50}H_{80}N_{12}O_{19}$
Exact Mass: 1152.57

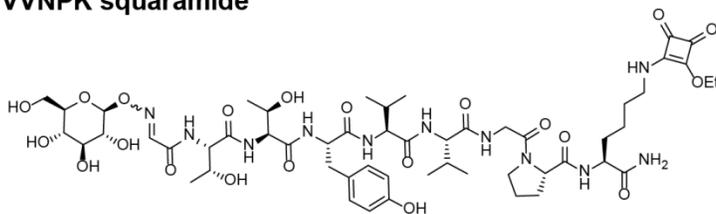


Time (min)	Solvent A (%)
0	0
2	0
25	50
30	100
34	100
37	0
40	0

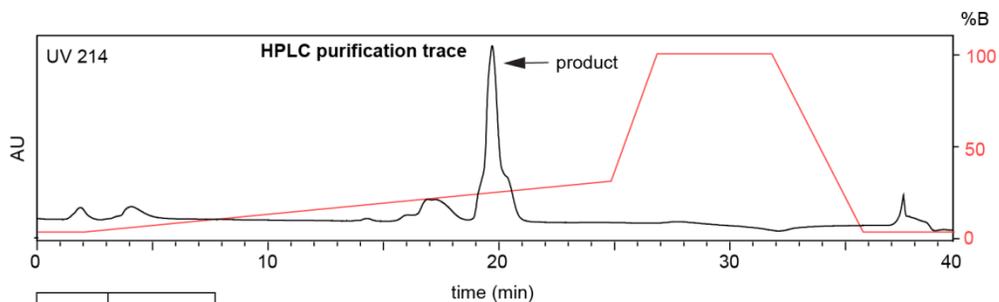
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μ m, 4.6 mm X 250 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



Glu-TTYVVNPK squaramide

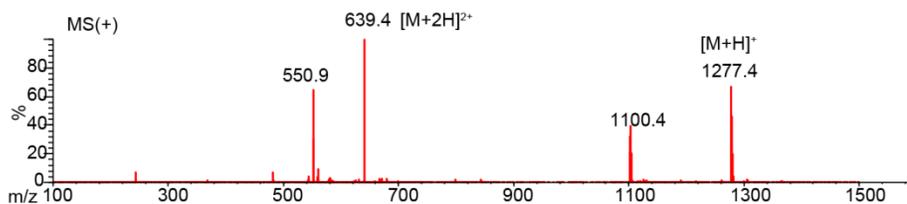
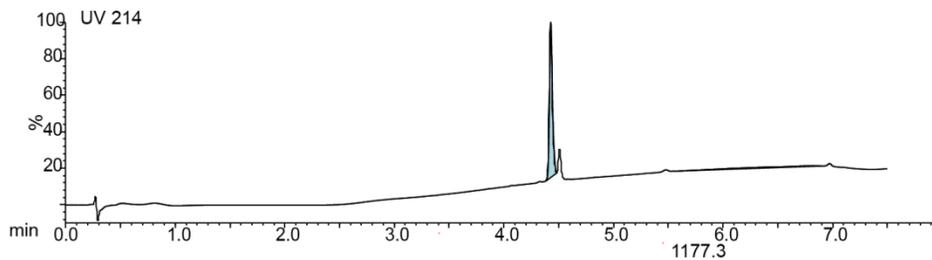


Chemical Formula: $C_{96}H_{84}N_{12}O_{22}$
 Exact Mass: 1276.58

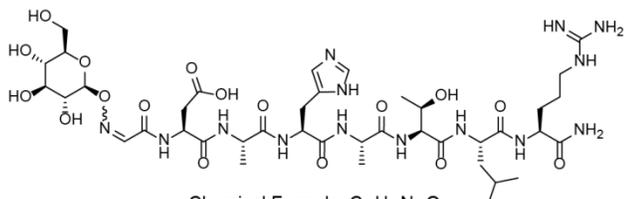


Time (min)	Solvent B (%)
0	2
2	2
25	30
27	100
32	100
36	2
40	2

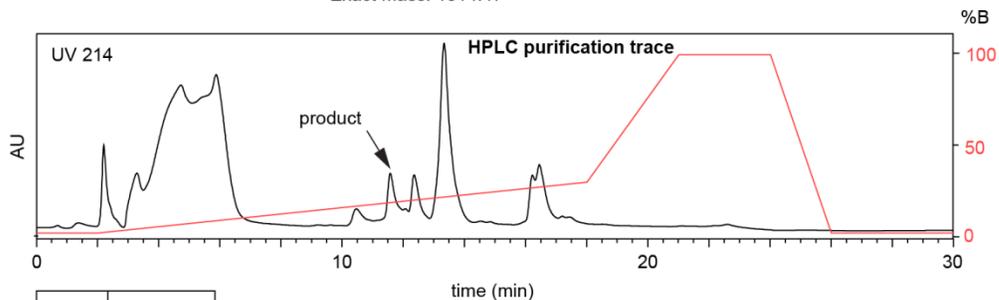
Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μm, 4.6 mm X 250 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



Glu-DAHATLR

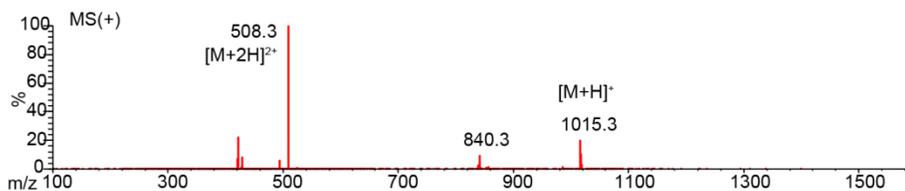
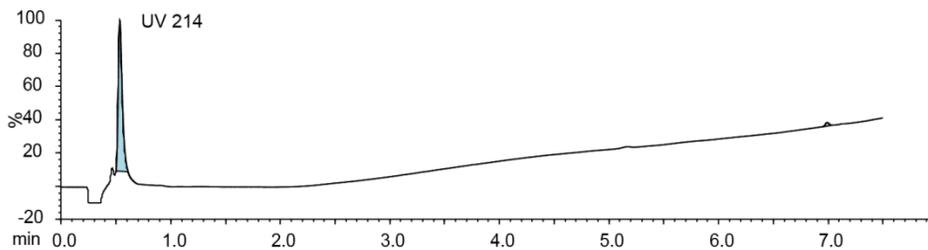


Chemical Formula: $C_{90}H_{66}N_{14}O_{17}$
 Exact Mass: 1014.47

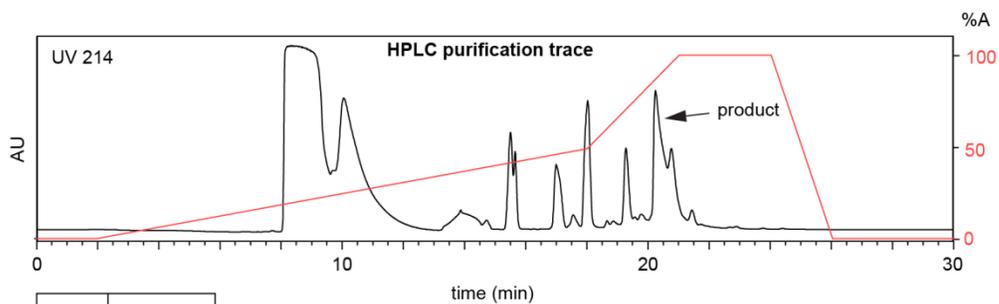
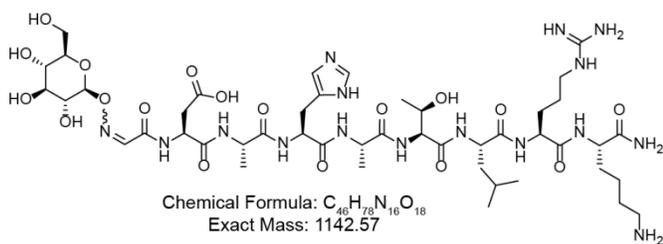


Time (min)	Solvent B (%)
0	2
2	2
18	30
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: $H_2O + 0.1\%$ (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

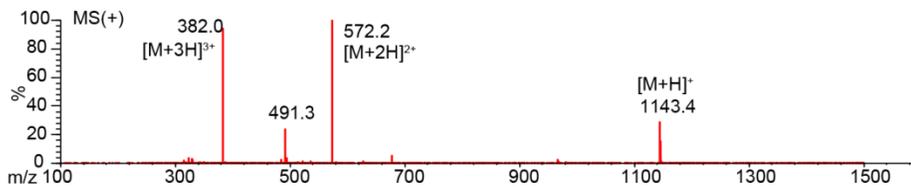
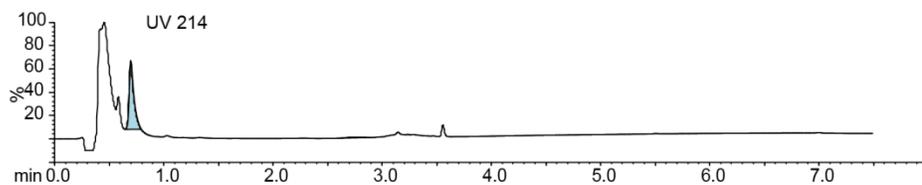


Glu-DAHATLRK

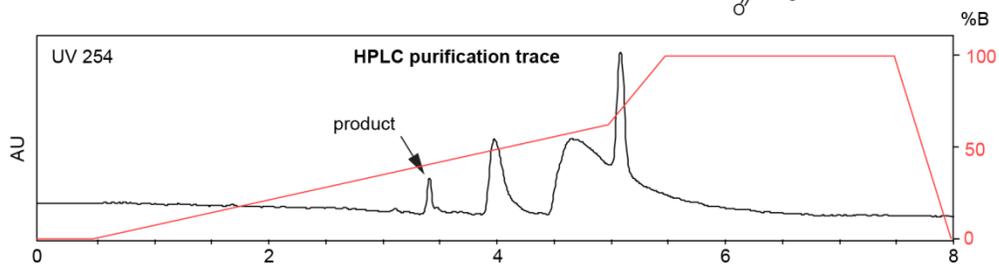
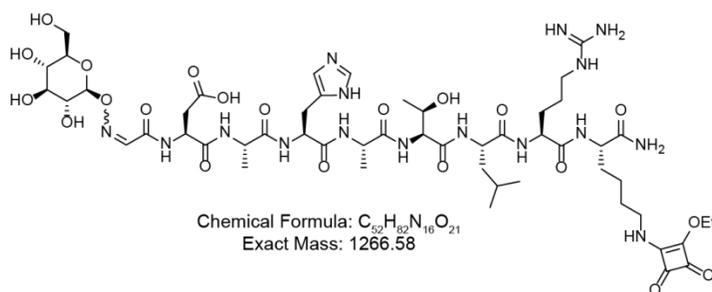


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μ m, 4.6 mm X 250 mm)
 Solvent A: H_2O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

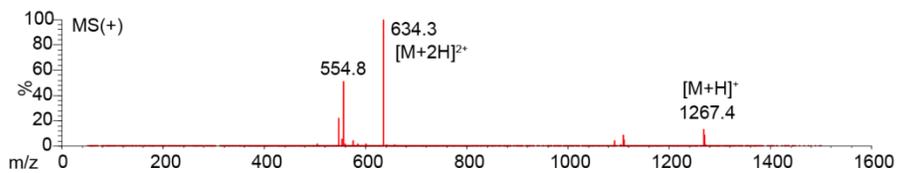
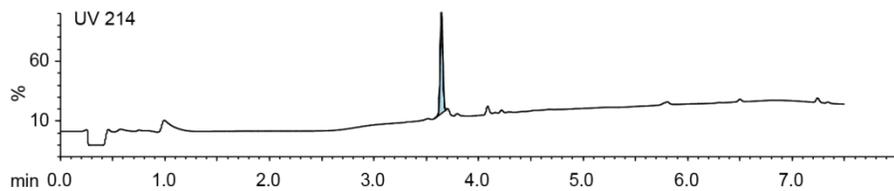


Glu-DAHATLRK squaramide

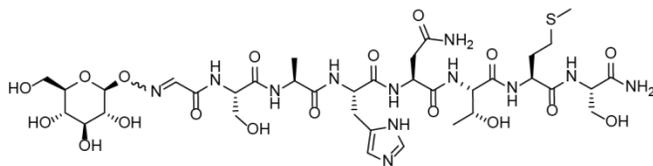


Time (min)	Solvent B (%)
0	2
0.5	2
5	60
5.5	95
7.5	95
8	2

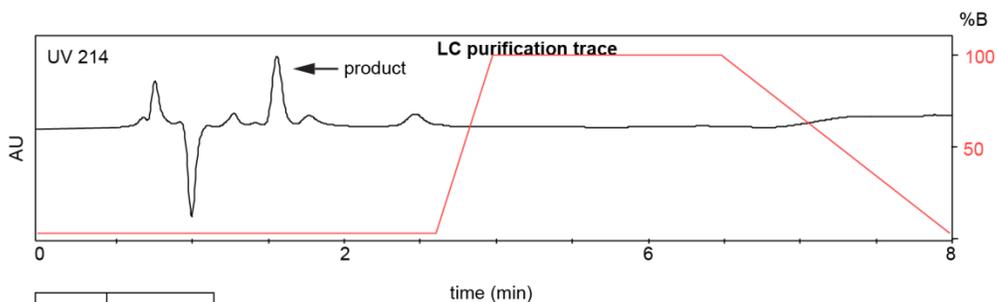
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 μ m, 4.6 mm X 75 mm)
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



Glu-SAHNTMS

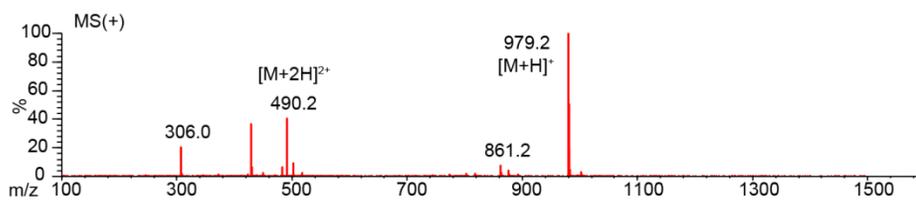
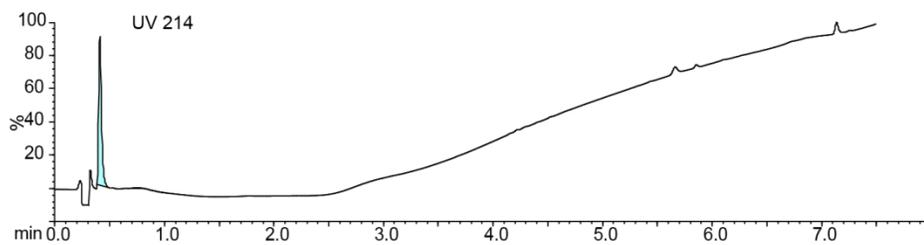


Chemical Formula: $C_{36}H_{56}N_{12}O_{10}S$
 Exact Mass: 978.37

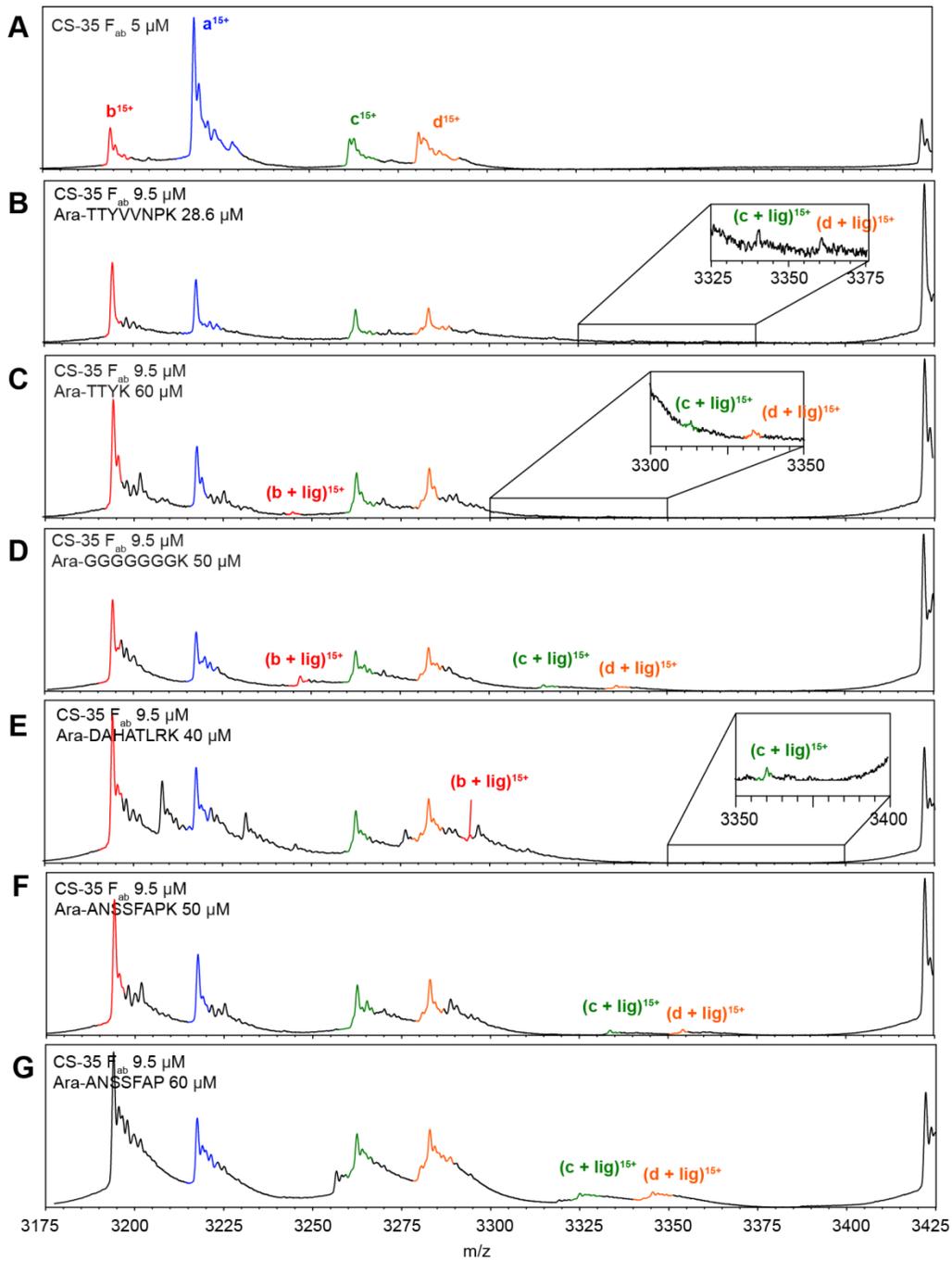


Time (min)	Solvent B (%)
0	2
3.5	2
4	90
6	90
8	2

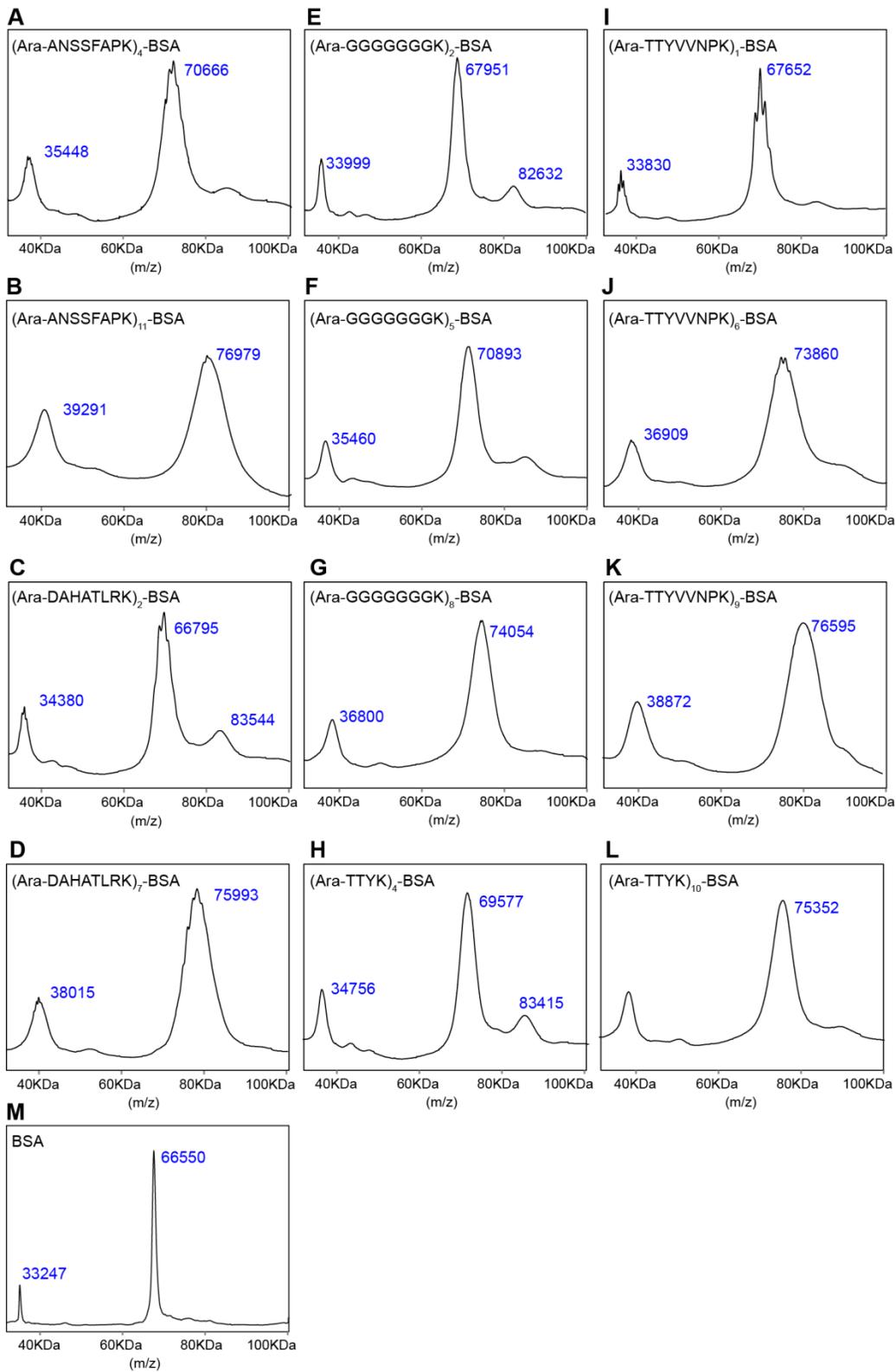
Flow rate: 1 mL/min
 Symmetry C18 Column
 (100Å, 3.5 μ m, 4.6 mm X 75 mm)
 Solvent A: H₂O + 0.1% (v/v) FA
 Solvent B: MeCN + 0.1% (v/v) FA



Appendix B: Supporting information for chapter 3



Appendix B-1. ESI-MS data of Ara-glycopeptides binding to CS-35 F_{ab}.



Appendix B-2. MALDI of synthesized Ara-glycopeptide BSA conjugates.

Appendix B-3. MatLab script MakeFigure3_5.m

```
%using signal from empty spots as background
%using average signal

clear all;
close all;

% define some swanky colors for plotting
teal      = [56      142      142]/256;
chartreuse = [113     198     113]/256;
turquoise  = [ 0      134     139]/256;
crimson    = [220     20      60 ]/256;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% defined the plotting appearance here
Xmin = 0.01;
Xmax = 30;
Ymin = -5000; % smallest value of Y
Ymax = 60000;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

file = 'Nov28_NoHeader';
DIR = '';
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

A = xlsread(fullfile(DIR,file)); %assign file contents to matrix A
NofIgGPoints = size(A,1); % total number of data points

row = A(:,3); %row number
column = A(:,2); %column number
block = A(:,1); %block number

IgG = A(:,10); % F532 average

block_CS35 = [ 1, 4, 7, 10, 13, 16, 19, 22];
conc_CS35 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 ];

block_CS40 = [ 2, 5, 8, 11, 14, 17, 20, 23];
conc_CS40 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 ];

block_906 = [ 3, 6, 9, 12, 15, 18, 21, 24];
conc_906 = [ 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 ];

% define structures that contains row, column, and name information

sample(1).R = 2;
sample(1).C = 1:2;
sample(1).name = 'Ara6-4';

sample(2).R = 2;
sample(2).C = 3:4;
sample(2).name = 'Ara6-9';

sample(3).R = 2;
sample(3).C = 5:6;
sample(3).name = 'Ara6-17';

sample(4).R = 3;
sample(4).C = 1:2;
sample(4).name = 'Ara6-ANSSFAPK-2';

sample(5).R = 3;
sample(5).C = 3:4;
```

```
sample(5).name = 'Ara6-ANSSFAPK-5';

sample(6).R = 3;
sample(6).C = 5:6;
sample(6).name = 'Ara6-ANSSFAPK-10';

sample(7).R = 4;
sample(7).C = 1:2;
sample(7).name = 'Ara6-TTYVVNPK-4';

sample(8).R = 4;
sample(8).C = 3:4;
sample(8).name = 'Ara6-TTYVVNPK-12';

sample(9).R = 4;
sample(9).C = 5:6;
sample(9).name = 'Ara6-DAHATLRK-2';

sample(10).R = 11;
sample(10).C = 1:2;
sample(10).name = 'Ara1-ANSSFAPK-4'; %c25

sample(11).R = 9;
sample(11).C = 5:6;
sample(11).name = 'Ara1-ANSSFAPK-11'; %c24

sample(12).R = 11;
sample(12).C = 5:6;
sample(12).name = 'Ara1-DAHATLRK-2'; %c27

sample(13).R = 11;
sample(13).C = 3:4;
sample(13).name = 'Ara1-DAHATLRK-7'; %c26

sample(14).R = 12;
sample(14).C = 5:6;
sample(14).name = 'Ara1-TTYVVNPK-1'; %c30

sample(15).R = 12;
sample(15).C = 3:4;
sample(15).name = 'Ara1-TTYVVNPK-6'; %c29

sample(16).R = 12;
sample(16).C = 1:2;
sample(16).name = 'Ara1-TTYVVNPK-9'; %c28

sample(17).R = 13;
sample(17).C = 3:4;
sample(17).name = 'Ara1-TTYK-4'; %c32

sample(18).R = 13;
sample(18).C = 1:2;
sample(18).name = 'Ara1-TTYK-10'; %c31

sample(19).R = 14;
sample(19).C = 3:4;
sample(19).name = 'Ara1-GGGGGGK-2'; %c35

sample(20).R = 14;
sample(20).C = 1:2;
sample(20).name = 'Ara1-GGGGGGK-5'; %c34

sample(21).R = 13;
sample(21).C = 5:6;
sample(21).name = 'Ara1-GGGGGGK-8'; %c33

%%%%%%%%%
```

```
sample(22).R = 5;
sample(22).C = 1:2;
sample(22).name = 'Glu-ANSSFAPK-2'; %c10

sample(23).R = 5;
sample(23).C = 3:4;
sample(23).name = 'Glu-ANSSFAPK-7'; %c11

sample(24).R = 14;
sample(24).C = 5:6;
sample(24).name = 'Glu-ANSSFAPK-9'; %c36

sample(25).R = 5;
sample(25).C = 5:6;
sample(25).name = 'Glu-ANSSFAPK-11'; %c12

sample(25).R = 6;
sample(25).C = 1:2;
sample(25).name = 'Glu-TTYVVNPK-2'; %c13

sample(26).R = 15;
sample(26).C = 3:4;
sample(26).name = 'Glu-TTYVVNPK-3'; %c38

sample(27).R = 15;
sample(27).C = 1:2;
sample(27).name = 'Glu-TTYVVNPK-7'; %c37

sample(28).R = 6;
sample(28).C = 3:4;
sample(28).name = 'Glu-DAHATLRK-3'; %c14

sample(29).R = 6;
sample(29).C = 5:6;
sample(29).name = 'Glu-DAHATLRK-5'; %c15

sample(30).R = 10;
sample(30).C = 1:2;
sample(30).name = 'Buffer';

sample(31).R = 10;
sample(31).C = 3:4;
sample(31).name = 'BSA';

sample(32).R = 10;
sample(32).C = 5:6;
sample(32).name = 'aGal';

%%%%%%%%%%

sample(33).R = 7;
sample(33).C = 1:2;
sample(33).name = 'Ara1-ANSSFAPK'; %c16

sample(34).R = 7;
sample(34).C = 3:4;
sample(34).name = 'Ara1-DAHATLRK'; %c17

sample(35).R = 7;
sample(35).C = 5:6;
sample(35).name = 'Ara1-GGGGGGK'; %c18

sample(36).R = 8;
sample(36).C = 1:2;
sample(36).name = 'Ara1-TTYVVNPK'; %c19
```

```

sample(37).R = 8;
sample(37).C = 3:4;
sample(37).name = 'Aral-TTYK'; %c20

sample(38).R = 8;
sample(38).C = 5:6;
sample(38).name = 'Glu-TTYVVNPK'; %c21

sample(39).R = 9;
sample(39).C = 1:2;
sample(39).name = 'Glu-ANSSFAPK'; %c22

sample(40).R = 9;
sample(40).C = 3:4;
sample(40).name = 'Glu-DAHATLRK'; %c23

%%% get empty spots intensity for each block

empty_block=zeros([1 24]);

for i=1:24 %loop through each block
    empty = zeros([1 9]);

    for j=1:NofIgGPoints %get F532 at specific column and row (that contains empty spots)

        if block(j)==i && row(j)==1 && column(j)==3
            empty(1)=IgG(j); %then record value to "empty" matrix
        end

        if block(j)==i && row(j)==1 && column(j)==4
            empty(2)=IgG(j);
        end

        if block(j)==i && row(j)==1 && column(j)==5
            empty(3)=IgG(j);
        end

        if block(j)==i && row(j)==15 && column(j)==5
            empty(4)=IgG(j);
        end

        if block(j)==i && row(j)==15 && column(j)==6
            empty(5)=IgG(j);
        end

        if block(j)==i && row(j)==16 && column(j)==2
            empty(6)=IgG(j);
        end

        if block(j)==i && row(j)==16 && column(j)==3
            empty(7)=IgG(j);
        end

        if block(j)==i && row(j)==16 && column(j)==4
            empty(8)=IgG(j);
        end

        if block(j)==i && row(j)==16 && column(j)==5
            empty(9)=IgG(j);
        end
    end
    %disp(empty)
    average = mean(empty); %average all empty values

    %assign averaged empty values according to block number
    empty_block(i)=average;
end

```

```

        %display(average)

end
%display(empty_block)

% Get IgG intensity from every sample in CS-35 blocks, then plot
figure(1);

for jj = 1:numel(sample) % loop through all samples

    for B = block_CS35 %block_CS35 %go through all data that has desired block number

        rep = 0; % blank the number of replicates

        for C = sample(jj).C % loop replicates in current block

            rep = rep + 1; % increase the replicate by 1

            for i = 1:NofIgGPoints %look at all points in matrix A

                %if block, row, column is desired
                if row(i)==sample(jj).R && block(i)==B && column(i)==C

                    % find location of the block value in the block_CS35
                    % variable
                    L = find(block_CS35 == B);

                    %extract IgG value and place it into the plotting
                    %array at the location = L

                    sample(jj).signal_CS35(rep,L) = IgG(i) - empty_block(B);
                    %disp(sample(jj).signal_CS35(rep,L))

                end
            end
        end
    end
end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for jj = 1:numel(sample) % loop through all samples

    for B = block_CS40 %go through all data that has desired block number

        rep = 0; % blank the number of replicates

        for C = sample(jj).C % loop 3 replicates in current block

            rep = rep + 1; % increase the replicate by 1

            for i = 1:NofIgGPoints %look at all points in matrix A

                %if block, row, column is desired
                if row(i)==sample(jj).R && block(i)==B && column(i)==C

                    % find location of the block value in the block_CS35
                    % variable
                    L = find(block_CS40 == B);

                    %extract IgG value and place it into the plotting
                    %array at the location = L

                    sample(jj).signal_CS40(rep,L) = IgG(i) - empty_block(B);

                end
            end
        end
    end
end

```

```

        end
    end
end
end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for jj = 1:numel(sample) % loop through all samples

    for B = block_906 %go through all data that has desired block number

        rep = 0; % blank the number of replicates

        for C = sample(jj).C % loop 3 replicates in current block

            rep = rep + 1; % increase the replicate by 1

            for i = 1:NofIgPoints %look at all points in matrix A

                %if block, row, column is desired
                if row(i)==sample(jj).R && block(i)==B && column(i)==C

                    % find location of the block value in the block_CS35
                    % variable
                    L = find(block_906 == B);

                    %extract IgG value and place it into the plotting
                    %array at the location = L

                    sample(jj).signal_906(rep,L) = IgG(i) - empty_block(B);
                    %disp(sample(jj).signal_906(rep,L))

                end
            end
        end
    end
end
end
end

% this is where we loop through all samples to fit

for jj = 1:numel(sample) % loop through all samples

    if jj<=21
        figure(1)
        subplot(4,6,jj);
    else
        figure(2)
        subplot(4,5,jj-21);
    end

    hold on;

    SIZE = 2;

    plot( conc_CS35, sample(jj).signal_CS35, ...
          'ok', ...
          'MarkerFaceColor', 0.8*[1 0 0],...
          'MarkerSize', SIZE );

    plot( conc_CS40, sample(jj).signal_CS40, ...
          'ok', ...
          'MarkerFaceColor', turquoise,...
          'MarkerSize', SIZE );

    plot( conc_906, sample(jj).signal_906, ...
          'ok', ...

```

```

        'MarkerFaceColor', chartreuse,...
        'MarkerSize', SIZE );

title([ 'Row: ' num2str(sample(jj).R) ...
        ' Columns: ' num2str(sample(jj).C) ' ' ...
        sample(jj).name], 'FontSize', 4 );

set( gca, 'YScale', 'linear', 'XScale', 'log',...
'TickDir', 'out','XMinorTick','on', 'YMinorTick','on',...
'XTick', [0.01 0.1 1 10 100], 'FontSize', 4);

ylim([Ymin Ymax]);
xlim([Xmin Xmax]);

end

%%
%%% this is where we re-loop through all the samples after we have plotted
%%% them and fit each subplot to an IC50 equation.
figure(1);

% first define the grid that will be used for plotting of the solutions

N = 100;      % # of points to plot (pick a larger number for a smooth curve
grid = 10.^(log10(Xmin): (log10(Xmax)-log10(Xmin))/N :log10(Xmax));

% for jj = 1:9      % instead of numel(sample), lets just loop through first 9
for jj = 1:numel(sample) % loop through all samples

    % first you have to define the X points (concentrations) and the
    % y-points (intensities) as one continuous vector

    if jj<=21
        figure(1)
        subplot(4,6,jj);
    else
        figure(2)
        subplot(4,5,jj-21);
    end

    y1 = []; y2 = []; y3 = [];
    x1 = []; x2 = []; x3 = [];

    for i = 1:2
        y1 = [y1 sample(jj).signal_CS35(i,1:end) ];
        y2 = [y2 sample(jj).signal_CS40(i,1:end) ];
        y3 = [y3 sample(jj).signal_906(i,1:end) ];

        x1 = [x1 conc_CS35(1:end) ];
        x2 = [x2 conc_CS40(1:end) ];
        x3 = [x3 conc_906(1:end) ];
    end

    %%%% define the fit options  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    min=0;
    MAX=20000;
    EC50=1;
    N=1;

    s = fitoptions('Method','NonlinearLeastSquares',...
        'Lower', [0, 0, 0, -1000],...
        'Upper', [1000, 100000, 2, 1],...

```

```

                'Startpoint',[EC50, MAX, N, min],...
                'TolFun', 1e-10 );

%%%%% define the fit equation

equation = ('(min + (MAX-min)/(1+(EC50/x)^N))');

ft = fitype( equation,'options',s );

%%%%% fit the data %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[sample(jj).fit_CS35, sample(jj).gof2_CS35] = fit(x1',y1',ft);
sample(jj).CON_CS35 = confint(sample(jj).fit_CS35);

[sample(jj).fit_CS40, sample(jj).gof2_CS40] = fit(x2',y2',ft);
sample(jj).CON_CS40 = confint(sample(jj).fit_CS40);

[sample(jj).fit_906, sample(jj).gof2_906 ] = fit(x3',y3',ft);
sample(jj).CON_906 = confint(sample(jj).fit_906);

disp(['trying to plot sample number ' num2str(jj) ]);
plot( grid, sample(jj).fit_CS35(grid),'r');
plot( grid, sample(jj).fit_CS40(grid),'color',turquoise);
plot( grid, sample(jj).fit_906(grid) , 'color',chartreuse);

TL{1} = [' IC50(CS35) = ' ...
        num2str( sample(jj).fit_CS35.EC50 ,'%0.1f') ...
        ' [' num2str( sample(jj).CON_CS35(:,1)' ,'% 0.1f') ' ]'];

TL{2} = [' IC50(CS40) = ' ...
        num2str( sample(jj).fit_CS40.EC50 ,'%0.1f') ...
        ' [' num2str( sample(jj).CON_CS40(:,1)' ,'% 0.1f') ' ]'];

TL{3} = [' IC50(906) = ' ...
        num2str( sample(jj).fit_906.EC50 ,'%0.1f') ...
        ' [' num2str( sample(jj).CON_906(:,1)' ,'% 0.1f') ' ]'];

legend off;

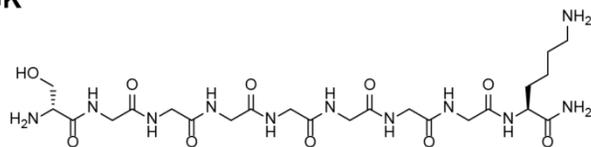
text(Xmin,0.85*Ymax, char(TL),'FontSize',4);

drawnow;

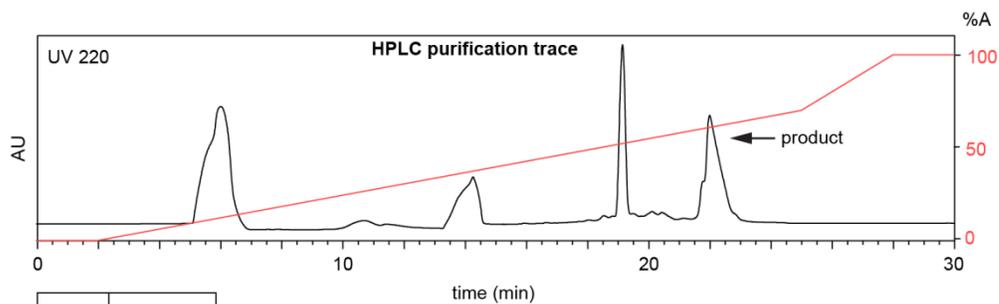
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end

```

SGGGGGGK

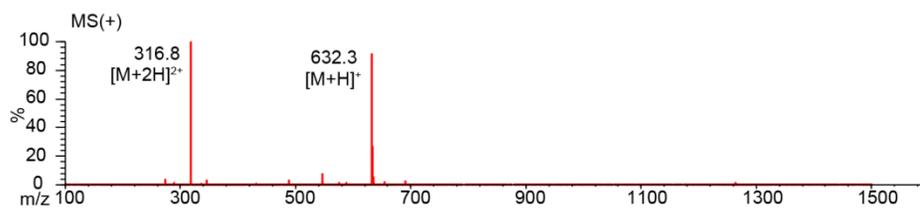
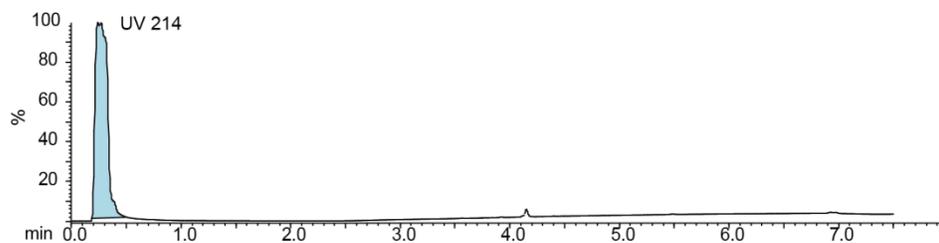


Chemical Formula: $C_{23}H_{41}N_{11}O_{10}$
Exact Mass: 631.30

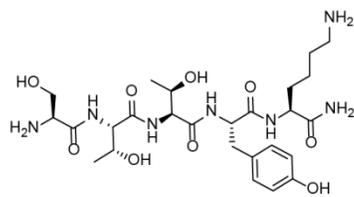


Time (min)	Solvent A (%)
0	0
2	0
25	70
28	100
30	100
34	0
38	0

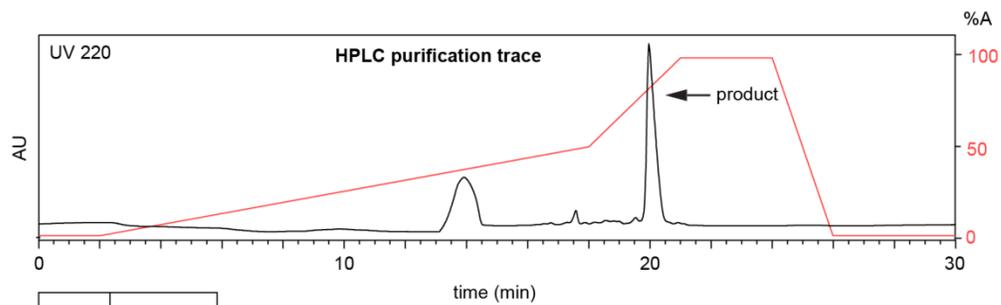
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μ m, 4.6 mm X 250 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



STTYK

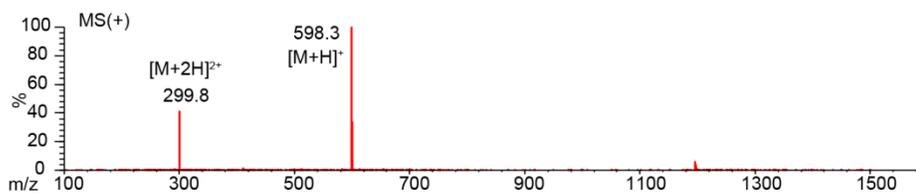
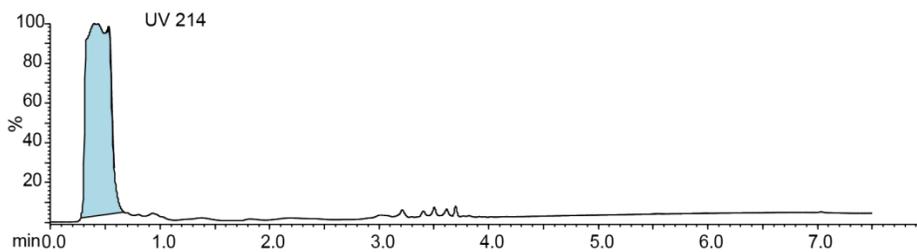


Chemical Formula: $C_{26}H_{43}N_7O_9$
Exact Mass: 597.31

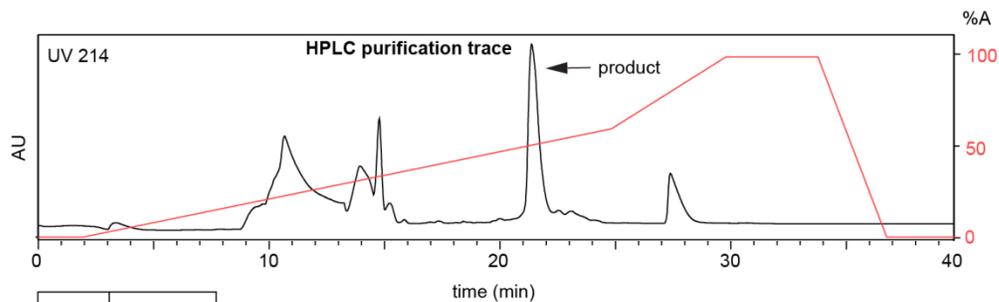
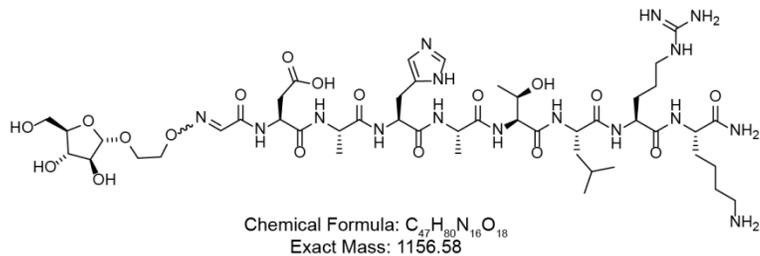


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μ m, 4.6 mm X 250 mm)
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

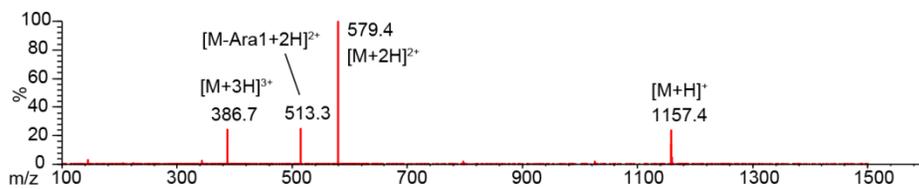
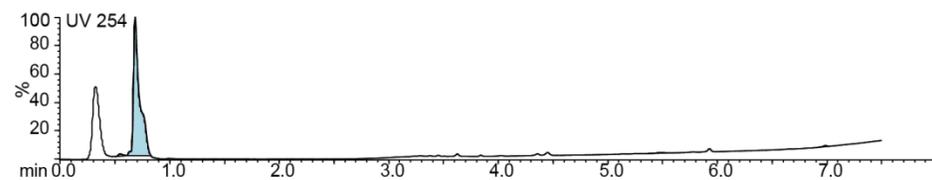


Ara-DAHATLRK

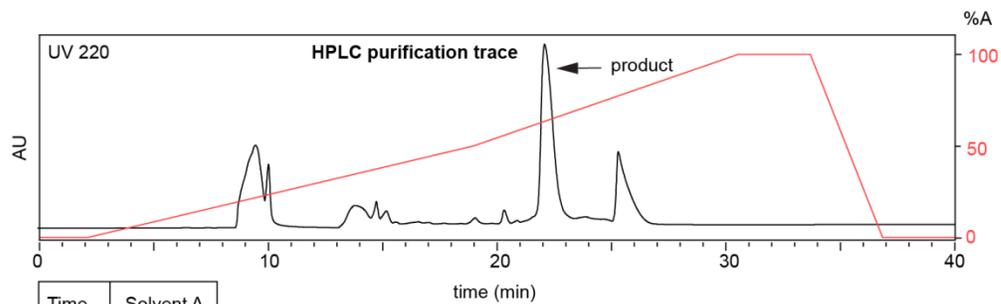
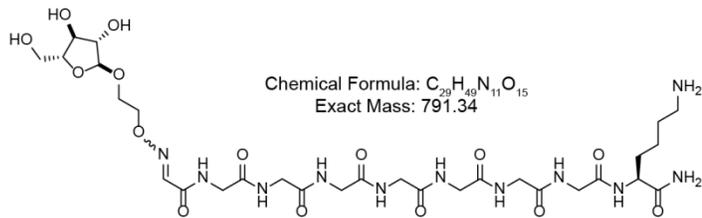


Time (min)	Solvent A (%)
0	0
2	0
25	60
30	100
34	100
37	0
40	0

Flow rate: 12 mL/min
 Amide Column
 (130A, 3.5 μ m, 4.6 mm X 250 mm)
 Solvent A: H_2O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

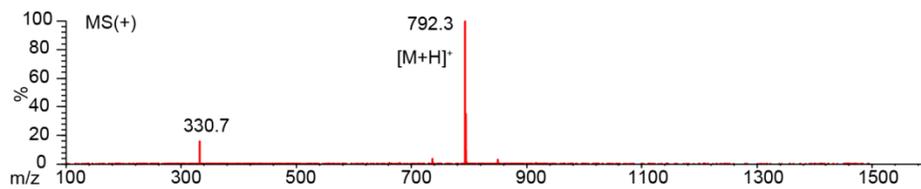
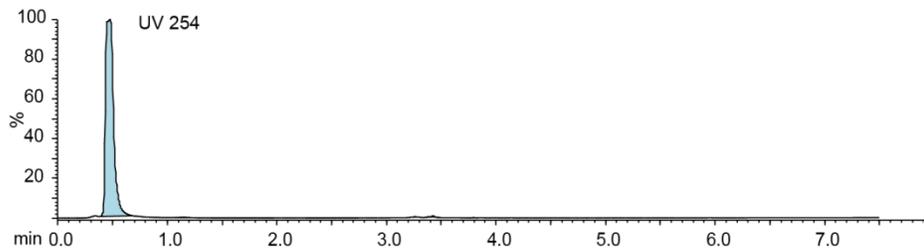


Ara-GGGGGGK

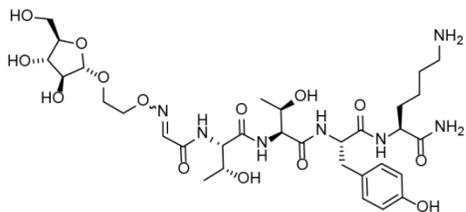


Time (min)	Solvent A (%)
0	0
2	0
18	50
29	100
32	100
35	2
38	2

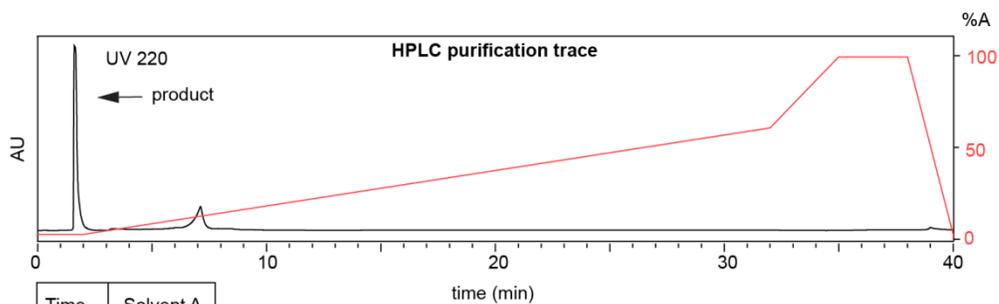
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 μm, 4.6 mm X 250 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



Ara-TTYK

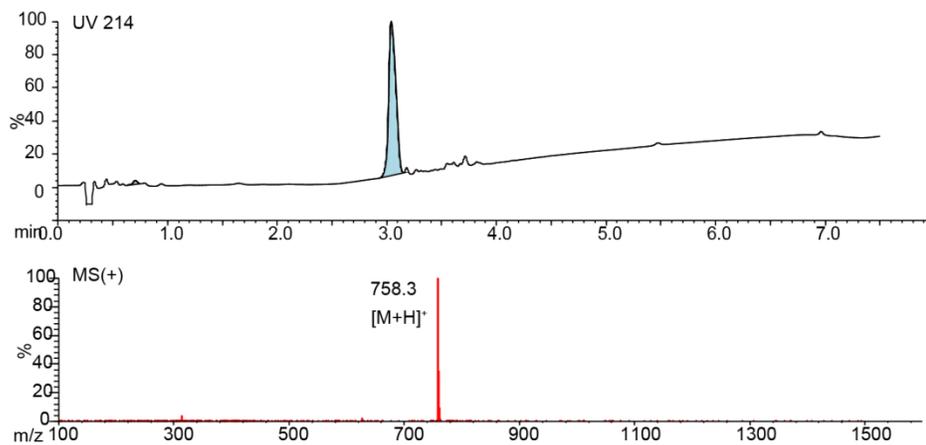


Chemical Formula: $C_{32}H_{51}N_7O_{14}$
 Exact Mass: 757.35

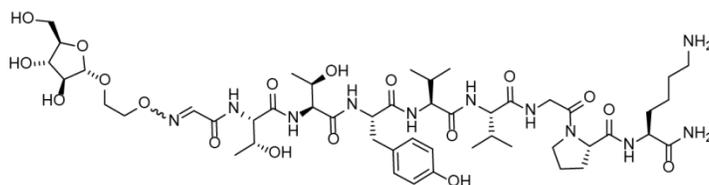


Time (min)	Solvent A (%)
0	0
2	0
32	60
35	100
38	100
41	0
46	0

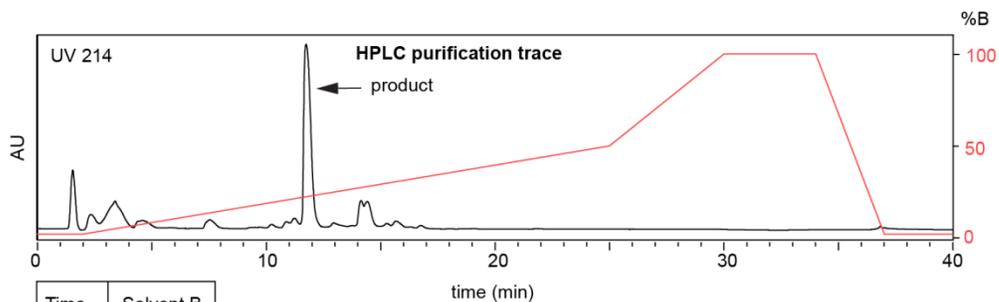
Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μm, 4.6 mm X 250 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA



Ara-TTYVVNPK



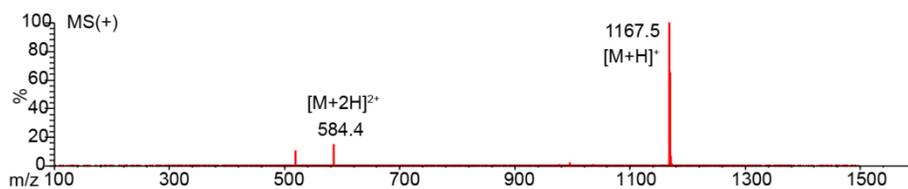
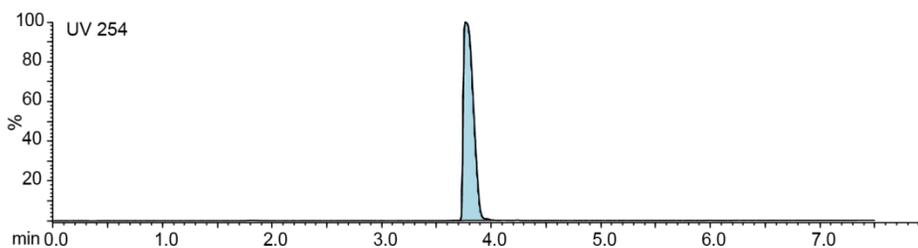
Chemical Formula: $C_{311}H_{482}N_{12}O_{19}$
 Exact Mass: 1166.58



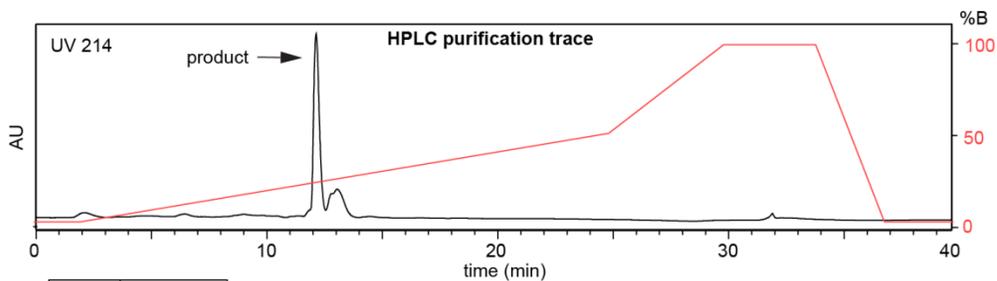
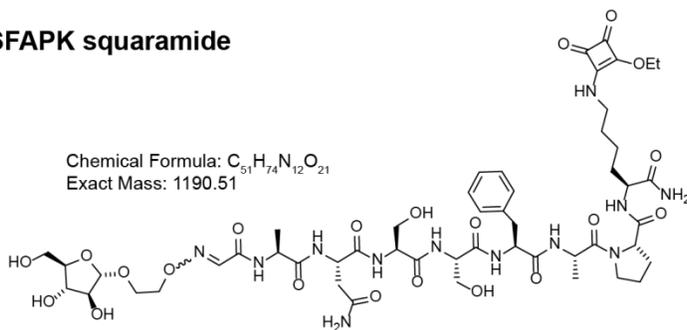
Time (min)	Solvent B (%)
0	2
2	2
25	50
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)

Solvent A: H₂O + 0.1% (v/v) FA
 Solvent B: MeCN + 0.1% (v/v) FA

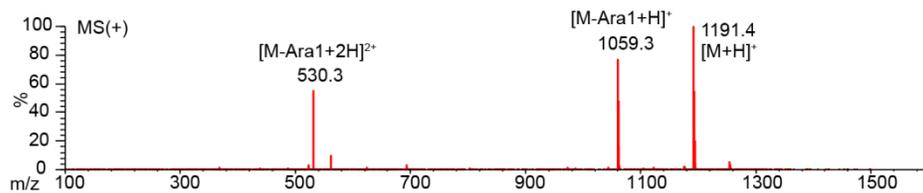
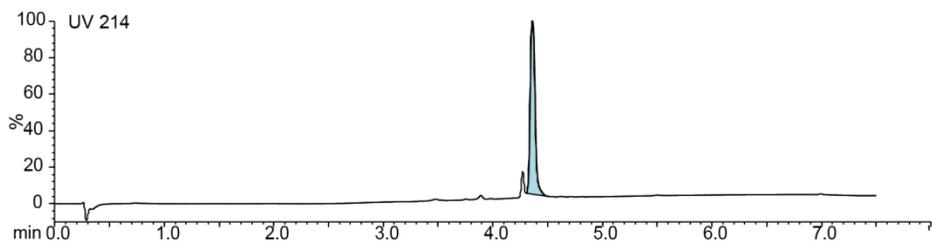


Ara-ANSSFAPK squaramide

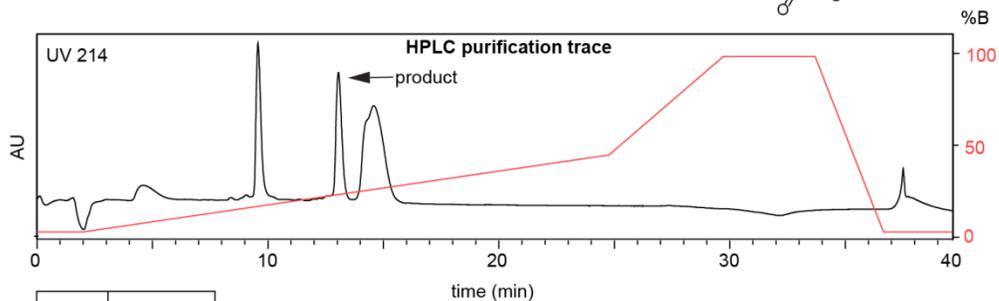
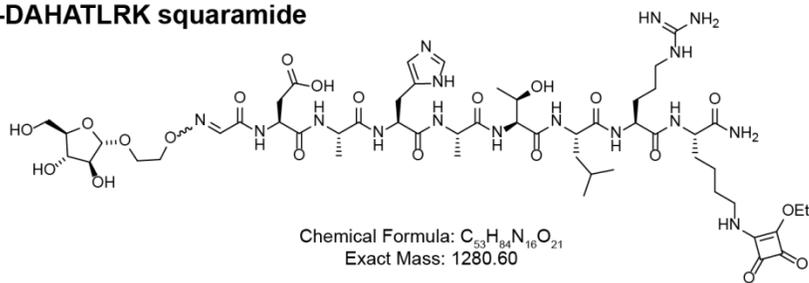


Time (min)	Solvent B (%)
0	2
2	2
25	50
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 µm, 19 mm X 50 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

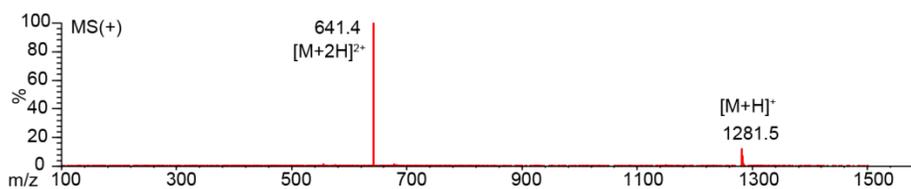
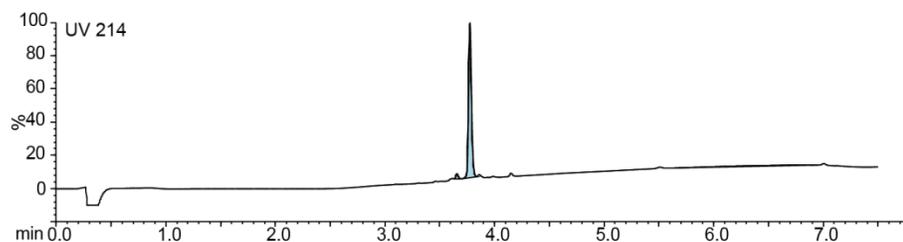


Ara-DAHATLRK squaramide

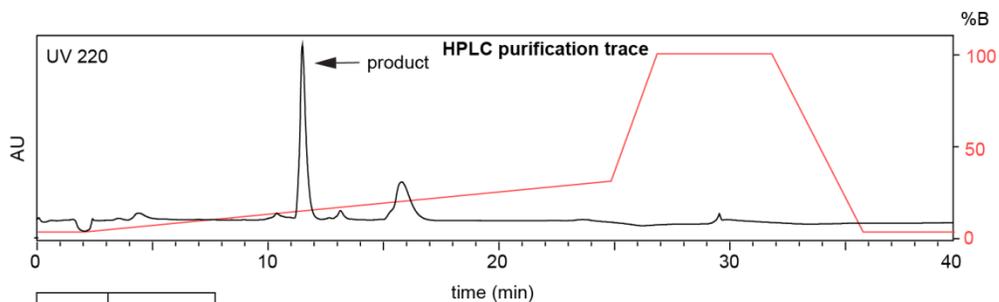
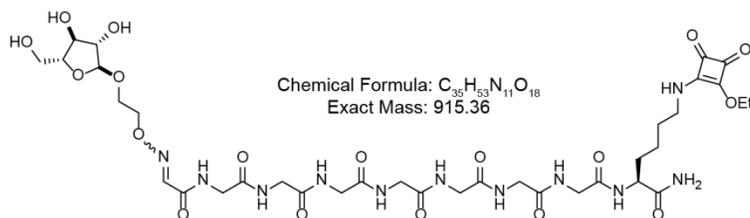


Time (min)	Solvent B (%)
0	2
2	2
25	45
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

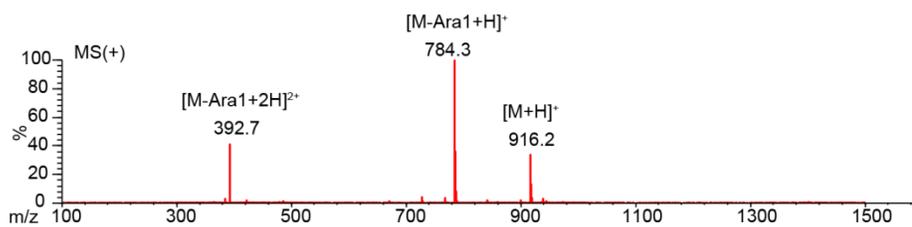
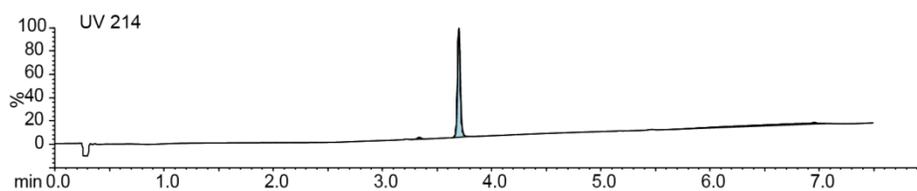


Ara-GGGGGGK squaramide

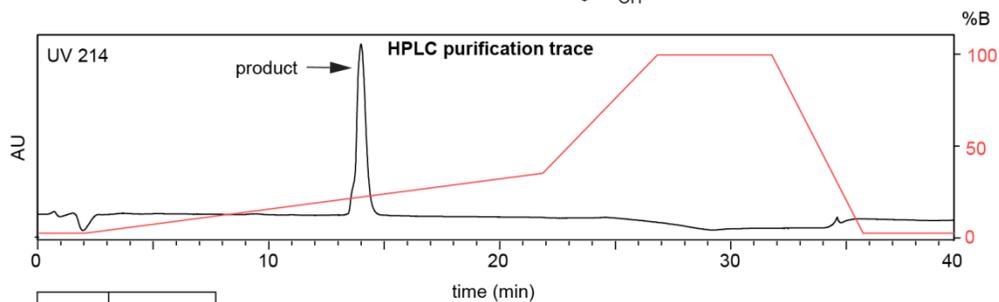
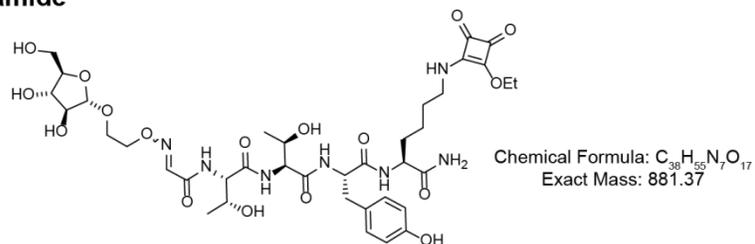


Time (min)	Solvent B (%)
0	2
2	2
25	30
27	100
32	100
36	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 μm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

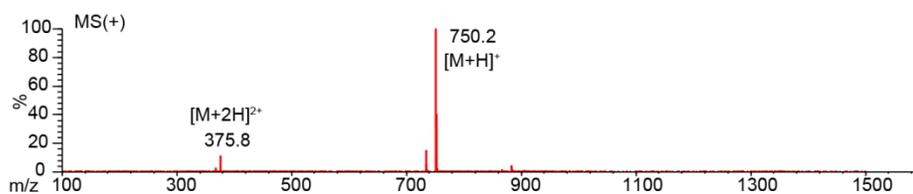
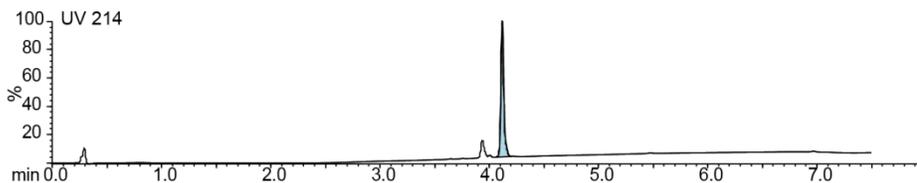


Ara-TTYK squaramide

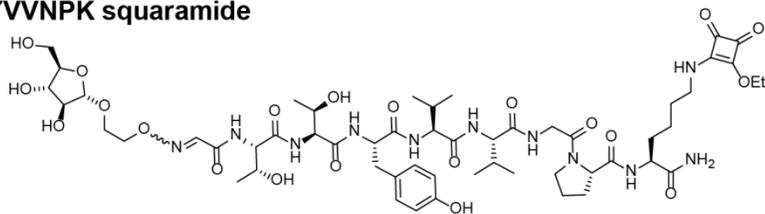


Time (min)	Solvent B (%)
0	2
2	2
22	35
27	100
32	100
36	2
40	2

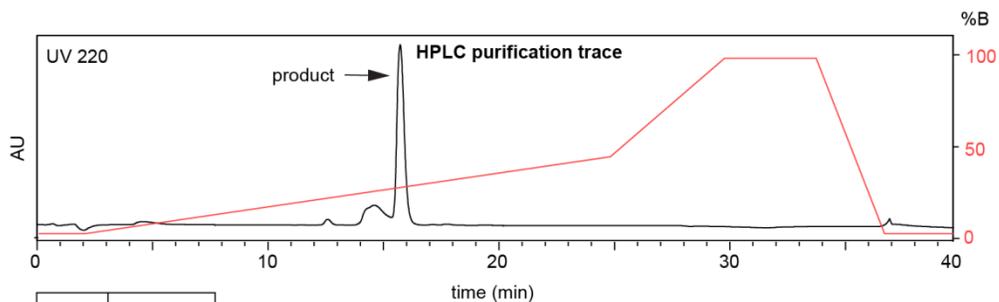
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)
Solvent A: H₂O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA



Ara-TTYVVNPK squaramide

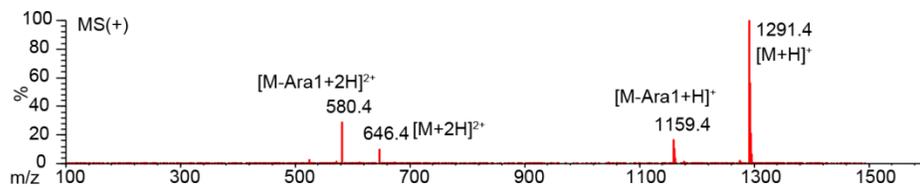
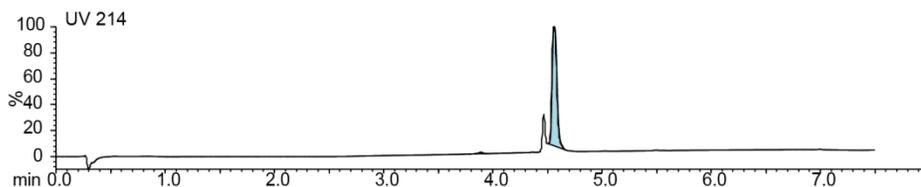


Chemical Formula: $C_{57}H_{86}N_{12}O_{22}$
 Exact Mass: 1290.60

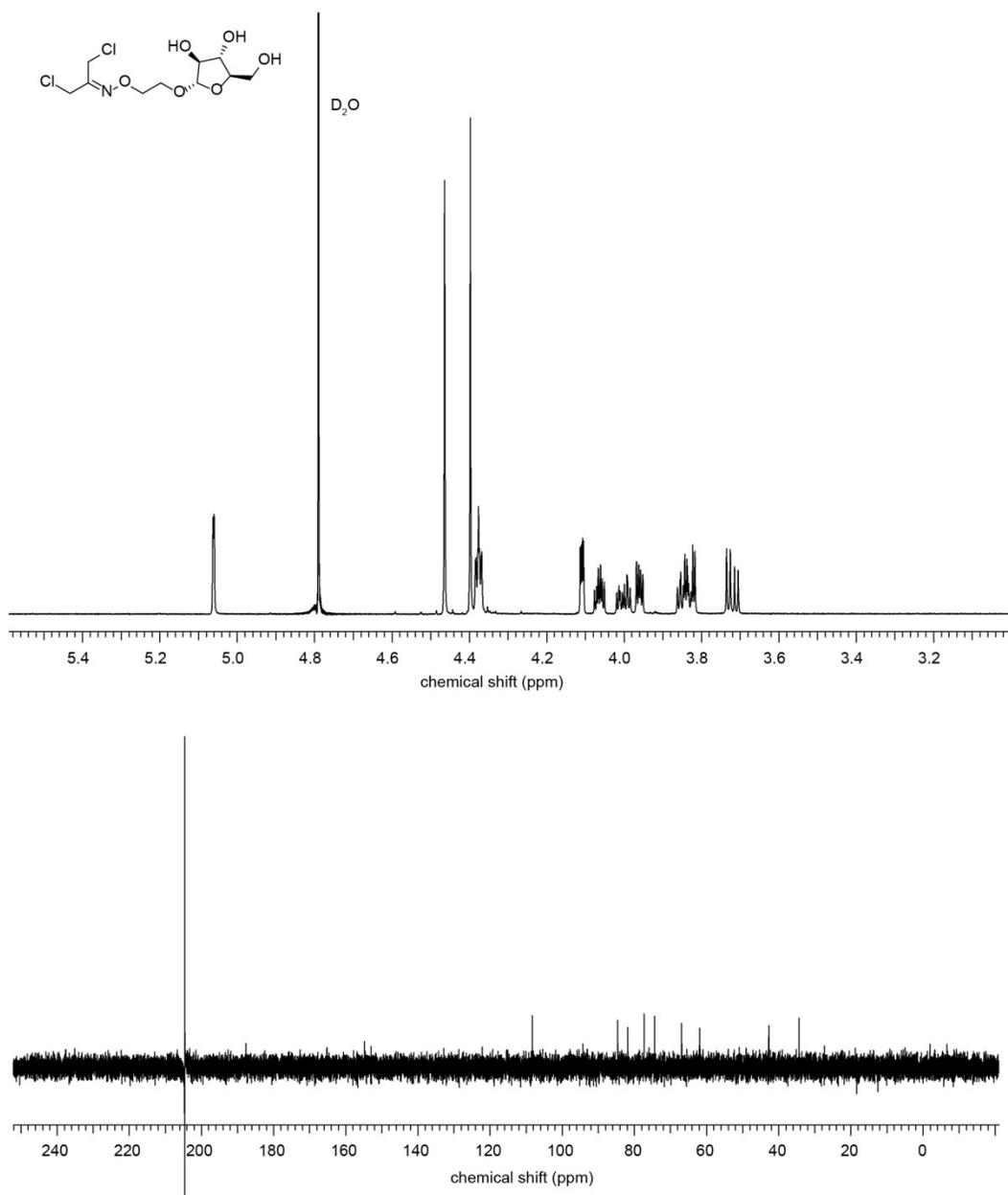


Time (min)	Solvent B (%)
0	2
2	2
25	45
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
 Amide Column
 (130Å, 3.5 μm, 4.6 mm X 250 mm)
 Solvent A: H₂O + 0.1% (v/v) TFA
 Solvent B: MeCN + 0.1% (v/v) TFA

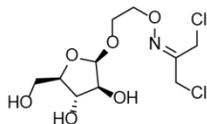


Appendix C: Supporting information for chapter 4

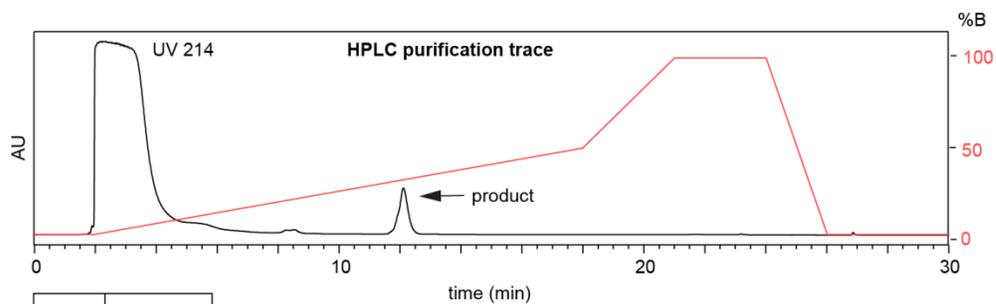


Appendix C-1. ^1H NMR (600 MHz, D_2O) and ^{13}C NMR (196 MHz, D_2O) spectra of DCO- α -Ara.

DCO- α -Ara



Chemical Formula: $C_{10}H_{17}Cl_2NO_6$
 Exact Mass: 317.04



Time (min)	Solvent B (%)
0	2
2	2
18	60
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
 Symmetry C18 Prep Column
 (100Å, 5 μ m, 19 mm X 50 mm)

Solvent A: H_2O
 Solvent B: MeCN

