Developing Functional Peptides as Synthetic Receptors, Binders of Protein and Probes for Bacteria Detection

Wenjian Wang

A Dissertation

Submitted to the Faculty of Chemistry

in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Boston College

Morissey College of Arts and Sciences

Graduate School

February 2021

Developing functional peptides as synthetic receptors, binders of protein, and probes for bacteria detection

Wenjian Wang

Advisor: Jianmin Gao

Nature has developed a generous number of peptides carrying out various essential functions in all living organisms. Human body produces peptides as signaling molecules, such as hormones, to transmit messages from cell to cell and regulate metabolic homeostasis. Microbes synthesize peptides as antibiotics to inhibit the growth of other microorganisms. These peptides display an exceeding diversity of amino acid composition, peptide sequence, secondary structure and post-translational modification. Inspired by nature, researchers have developed peptides as a unique modality of therapeutics, combining the best attributes of small-molecule drugs and protein-based biopharmaceuticals. This work has sought to explore the potential of peptides as synthetic receptors, binders of protein and probes for bacteria detection.

The research started from a foldable cyclic peptide scaffold, prolinomycin, a proline-rich analogue of valinomycin. The peptide can chelate a potassium ion folding into a drum like structure, which provides a platform to display and preoganize functional side chains for target binding. We first investigated its folding behavior under physiological conditions. We demonstrate that the metal-assisted folding of the prolinomycin scaffold tolerates various side chain mutations. The stability of the structure can be improved by introducing crosslinking moieties. Based on this scaffold, we rationally designed synthetic receptors of various amines by utilizing iminoboronate chemistry with acetylphenyl boronic acid (APBA).

Furthermore, I pursued phage display, a powerful technique to develop high affinity peptide binders of protein targets. Proteins are the most appealing targets for drug development and disease biomarkers discovery. We chose sortase A (SrtA) as a model target protein to screen for potent peptide binders. A peptide inhibitor of sortase A with single-digit micromolar affinity was identified from a cyclic peptide library displayed by phage. In addition, from the chemically modified phage display peptide library presenting APBA motifs, peptide binders with specificity and micromolar affinity towards SrtA were discovered. Instead of binding to the active site, the peptide could recognize the surface of the protein.

Additionally, to further expand the chemical space of phage display, I constructed a phage display peptide library presenting N-terminal cysteine (NCys) which can undergo site-specific chemical modifications. Two pieces of chemistry were applied, including thiazolidino boronate (Tzb) mediated acylation reaction of NCys and 2-cyanobenzothiazole (CBT)-NCys condensation. The site-specific dual modifications on NCys and internal Cys of phage-encoded peptides were achieved. Furthermore, a strategy to N, S-doubly label NCys via an alternative pathway of CBT condensation was reported, which presents a significant addition to the toolbox for site-specific protein modifications.

Finally, by functionalizing graphene field effect transistors (G-FET) with peptide probes, we developed the first selective, electrical detection of the pathogenic bacterial species *Staphylococcus aureus* and antibiotic resistant *Acinetobacter baumannii* on a single platform.

Overall, peptides provide enormous opportunities for therapeutics development. Research herein demonstrated principles of peptide design for specific molecular

recognition. Novel chemistry strategies have been developed to expand the molecular diversity of peptide libraries. We believed that the advances in peptide design and screening would promote peptide-based drug discovery.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Schemes

List of abbreviations

A. baumannii	*Acinetobacter baumannii*
AFM	atomic force microscopy
Ala, A	Alanine
Alloc	Allyloxycarbonyl
AMPs	Antimicrobial peptides
APBA	Acetylphenyl boronic acid
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic Acid
B2pin2	Bis(pinacolato)diboron
BCBT	Biotin-CBT
BLyS	B-lymphocyte stimulator
BMS	Bristol-Myers Squibb
Boc	Tert-butyloxycarbonyl
Bpin	Pinacol ester
BSA	Bovine serum albumin
BuK	N-butyryl-lysine
CBT	2-Cyanobenzothiazole
cfu	Colony Forming Units
CuAAC	Copper (I) Azide-Alkyne Click Chemistry
CVD	Chemical vapor deposition
Cys, C	Cysteine
Dab	Diaminobutyric acid
Dap	Diaminopropionic acid
Dbz	Diaminobenzoic acid
DCM	Dichloromethane
DIPEA	N, N'-Diisopropylethylamine
DMF	Dimethylformamide

DMMA	(Z)-3-(2,5-dimethoxyphenyl)-2-(4-methoxyphenyl) acrylonitrile
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
E. coli	Eschericia coli
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
Et3N	Triethylamine
FAM	Carboxyfluorescein
FDA	Food and Drug Administration
FITC	Fluorescein isothiocyanate
Fmoc	Fluorenylmetholoxycarbonyl
FP	fluorescence polarization
FPBA	Formylphenyl boronic acid
G-FET	graphene field effect transistors
GlcNAc	N-Acetylglucosamine
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
HA	Human influenza hemagglutinin
HBTU	Hexafluorophosphate Benzotriazole Tetramethyl Uronium
HCl	Hydrochloride
His, H	Histidine
HRP	Horseradish peroxidase
Hyp	hydroxyproline
IA	Iodoacetamide
IDT	Integrated DNA Technologies
IL	Interleukin
Ile, I	Isoleucine
IPTG	Isopropyl β-D-1-thioglactopyranoside

iTCEP	Immobilized TCEP
K2CO3	Potassium carbonate
KCl	Potassium chloride
Kd	Equilibrium dissociation constant
KOAc	Potassium acetate
LC-MS	Liquid chromatography-mass spectrometry
Leu, L	Leucine
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
Lys, K	Lysine
Lys-PG	Lysyl phosphatidylglycerol
MDM	Mouse double minute 2
MeOH	Methanol
Met, M	Methionine
MMP	Metalloproteinase
mRNA	Messenger RNA
MS	Mass spectrometry
NaCl	Sodium chloride
NCL	Native chemical ligation
NCys	N-terminal cysteine
NEB	New England Biolabs
NF-ATc	Nuclear factor of activated T cells
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B
NHS	N-hydroxysuccinimide
NMM	N-Methylmorpholine
NMR	Nuclear Magnetic Resonance spectroscopy
NTA	Nitrilotriacetate
OD600	Optical density measured at 600 nm
Orn	Ornithine
p53	Tumor protein p53

PA	Protective antigen
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Tween
$Pd(dppf)Cl_2$	1,1′-Bis(diphenylphosphino)ferrocene] dichloropalladium(II)
PDL	Programmed death-ligand
PDMS	Polydimethylsiloxane
PEG	Polyethylene glycol
pfu	Plaque-forming units
Phe, F	Phenylalanine
PK	Plasma kallikrein
PPIs	Protein-protein interactions
Pro, P	Proline
RaPID	Random non-standard Peptide Integration Discovery
RNA	Ribonucleic acid
RP-HPLC	Reversed-phage high-performance liquid chromatography
S. aureus	Staphylococcus aureus
SA	Salicylic acid
Ser, S	Serine
SIRT2	Human sirtuin2
SPPS	Solid-phase peptide synthesis
SrtA	Sortase A
TAMRA	Carboxytetramethylrhodamine
TATA	1,3,5-triacryloyl-1,3,5-triazinane
TBMB	Tris(bromomethyl)benzene
TBS	Tris buffered saline
TBST	Tris-buffered saline with Tween
TCEP	Tris(2-carboxyethyl) phosphine
TEV	Tobacco Etch Virus
TFA	Trifluoroacetic acid

THF	Tetrahydrofuran
Thr, T	Threonine
TIS	Triisopropylsilane
TLC	Thin layer chromatography
TLR	Toll-like receptor
TNF-α	Tumor necrosis factor alpha
TPOR	Thrombopoietin receptor
Trp, W	Tryptophan
TSAT	(Tris-(succinimidyl)aminotriacetate)
Tyr, Y	Tyrosine
Tzb	Thiazolidino boronate
uPA	urokinase-type plasminogenactivator
UV-Vis	Ultraviolet-visible
Val, V	Valine
V_D	Dirac voltage
VEGF-A	Vascular endothelial growth factor A
WT	Wild type
Xgal	5-bromo-4-chloro-2-indolyl-β-D-galactopyranoside

Acknowledgements

Firstly, I would like to thank Dr. Jianmin Gao for his support and encouragement over past five years. I am always inspired by his great passion for science. It is his high standard for novel science that really raised my level of critical thinking. The difficult questions he asked made me think much deeper and wider, which has been extremely rewarding to my intellectual growth. I also really appreciate the opportunities he gave me to present my work in many conferences, where I learnt how to communicate with people in science.

I'd like to thank Dr. Eranthie Weerapana, Dr. Abhishek Chatterjee and Dr. Jia Niu for all of their help and suggestions during my time in graduate school. Their interest in my research always encouraged me. I appreciate their effort to invite great scientists to BC and give amazing seminars. I also learnt a lot from the Cume Exams they gave. Additionally, I want thank their lab members who are always willing to help me and answer my questions.

Additionally, I'd like to thank my co-workers and lab mates. Dr. Azade Hosseini is a great mentor who taught me how to push a project forward strongly. Dr. Kelly McCarthy, Dr. Samantha Cambray, Dr. Kaicheng Li and Dr. Mike Kelly always provide me with help and advices. I really appreciate their effort to keep every equipment running in the lab. I want to thank Dr. Narendra Kumar for being a great collaborator, without whom I couldn't have finished the amazing G-FET project. Furthermore, I'd like to thank Dr. Mengmeng Zheng and Dr. Mi Zhou, who brought a lot of fun to my life. They are not only supportive mentors but also great friends.

Moreover, I’d like to thank my friends. Dr. Miao Qi is my best friend at BC, who always accompanies me in difficult times. My life would be much bleak without her. Junyan Liu stands by me all the time, who won't be impatient to listen to my complaints. Xiaonan Ren, although not around, but always shares her life with me and support me remotely. Jialin Cao, a very positive person, always encourages me, with whom I can always talk about dreams. I also want to thank Xiangxuan Deng who helped me a lot with thesis writing. I’m grateful to have Shu Wang around, who is very intelligent and humorous. I’d like to thank my cousins, Zimin Wang and Yuanyu Chang, who brighten my life all the time. Last but not least, I’d like to thank my boyfriend, Sixu Liu, who is a real delight to me. He is so funny and intelligent. Sixu is the one who makes me laugh when I’m depressed, the one who calms me down when I’m anxious, the one who points the way when I don’t know what to do.

More importantly, I’d like to thank my parents for their spiritual and financial support. They are the kindest and strongest people I have ever seen. It is their love and encouragement that push me forward to pursue my dreams. They never ask anything from me. Instead, they always tell me not to work too hard. From them, I know what courage is, what love is, and who I am.

Chapter 1 Introduction

1.1 Peptides as a new modality of therapeutics

Nature has evolved peptides with numerous functions. They can be hormones regulating essential life processes, or antibiotics fighting bacterial infections. Inspired by nature, researchers have developed peptide drugs as powerful therapeutics, such as hormone derivatives octreotide and leuprorelin used to treat cancer, and liraglutide used to treat type 2 diabetes. More than 60 peptide-based drug have been approved worldwide and about 20 new peptides enter into clinical trials every year[1, 2]. In addition to developing natural peptide derivatives, a significant effort has been put into the *de novo* generation of peptide ligands. Owing to advances in the various screening techniques and rational design techniques, various peptides have been prepared with diverse structures, suitable sizes, conformational restraint, preorganization of multiple functional groups and decent pharmacological properties[3-5]. This allows peptides to emerge as a new modality of therapeutics and reimagine medicine.

In this chapter, I will first introduce three categories of peptide drugs inculing antibiotics, enzyme inhibitors and protein-protein inhibitors. Furhtermore, the methods for developing functional peptides will be presented, focusing on rational design, phage display and mRNA display.

1.1.1. Peptides as antibiotics

Hundreds of peptide antibiotics have been discovered since mid-twentieth century. In general, they can be classified into two groups, nonribosomally synthesized and ribosomally synthesized peptides[6]. Nonribosomally synthesized peptides contain at least two amino acid moieties and are made on multienyzme complexes of bacteria and fungi, instead of on ribosomes. These peptides are diverse in residue, sequence, and structure. In

addition to L- or D- amino acids, they also contain unusual amino acids, which can be modified by acylation, N methylation, glycosylation and cyclic ring formation. For example, vancomycin and teicoplanin belong to this class, which have sugar-substituted peptide backbones (Figure 1-1). In addition, gramicidin S, polymyxin E, and bacitracin fall into this category, which are widely used in clinic. Some of these peptides, like polymycins, are cationic and can bind to and permeabilize the outer membrane, and further destabilize the cytoplasmic membrane of gram negative bacteria[7]. Other peptides such as bacitracin are gram positive specific and able to inhibit the synthesis of cell wall, peptidoglycan and proteins[8]. Even though these nonribosomally synthesized peptides have been used to treat a variety of infectious diseases, they show systemic toxicity are restricted to topical

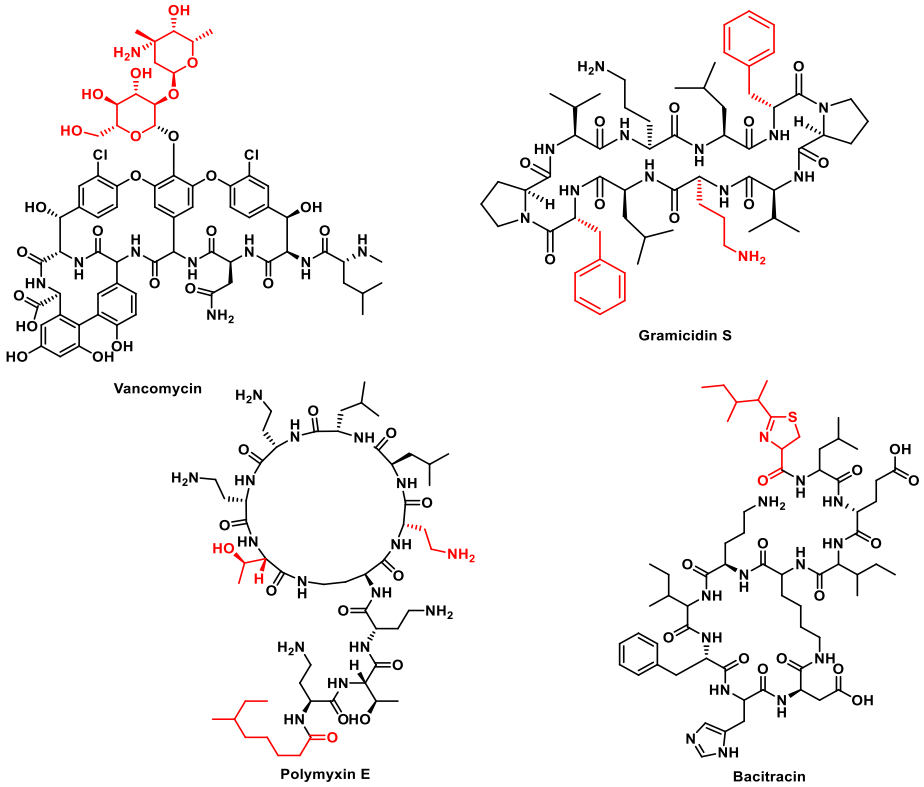


Figure 1-1 . Chemical structures of nonribosomally synthesized peptide antibiotics. Unique modifications are highlighted by red color.

applications. Therefore, there is a need to improve the selectivity of these peptides for bacteria than for mammalian cells and decrease their toxicity.

Antimicrobial, ribosomally synthesized, peptides have been identified from mammals, amphibians, insects, plants, bacteria and viruses. These peptides tend to be amphipathic and form β-structures, α-helices and loops. These antimicrobial peptides usually contain a large number of positively charged amino acids, lysine and arginine, together with hydrophobic amino acid residues, but few negatively charged amino acids. For example, human defensins show antimicrobial functions. They have a high arginine and lysine content and form β-sheet structures that is stabilized by disulfide bonds (Figure 1-2a)[9, 10]. In addition, some peptides tend to form amphipathic α-helices, such as melittin (Figure 1-2b)[11] isolated from bee venom and amphibian peptide magainins. It is believed that cationic nature of the peptides promotes interaction with bacterial outer and cytoplasmic membranes and the hydrophobic groups can provide hydrophobic interactions with lipid acyl chains[12]. Some peptides can also bind to lipopolysaccharide (LPS) and perturb the

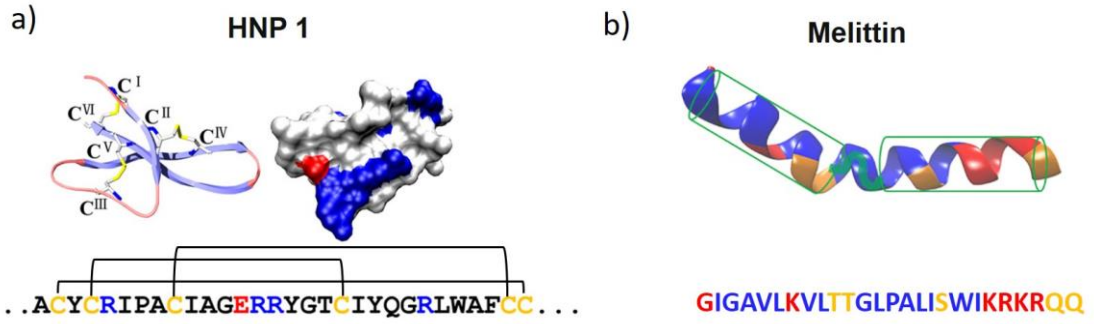


Figure 1-2. Structures of antimicrobial peptides. a) Sequence and structure of human neutrophil peptide (HNP-1). Adapted with permission from Ref. [10]. b) Sequence and structure of melittin. Different types of residues are colored differently: blue for hydrophobic residues, orange for hydrophilic ones, and red for charged. Adapted with permission from Ref. [11].

outer membrane[13]. However, they also destruct mammalian cells including red blood cells, T cells and neuronal cells[14].

It is still challenging to predict the activity and toxicity of peptide antibiotics, however, these natural peptides provide elegant templates and great starting points for researchers to design novel synthetic antibiotics.

1.1.2. Peptides as enzyme inhibitors

Most enzyme inhibitors are "small molecules" with molecular weight less than 500 Da, following the "rule of 5"[15]. The only peptide-based enzyme inhibitor approved by FDA so far is cyclosporine A, an inhibitor of calcineurin. Cyclosporine A is an 11-amino acid cyclic peptide (Figure 1-3), originally isolated from the fungus *Tolypocladim inflatum* and works as a potent immunomodulatory agent for the treatment of autoimmune diseases. It can form a complex with cyclophilin, which then inhibits the enzymatic activity of calcineurin phosphatase. As a result, the cytoplasmic component of nuclear factor of

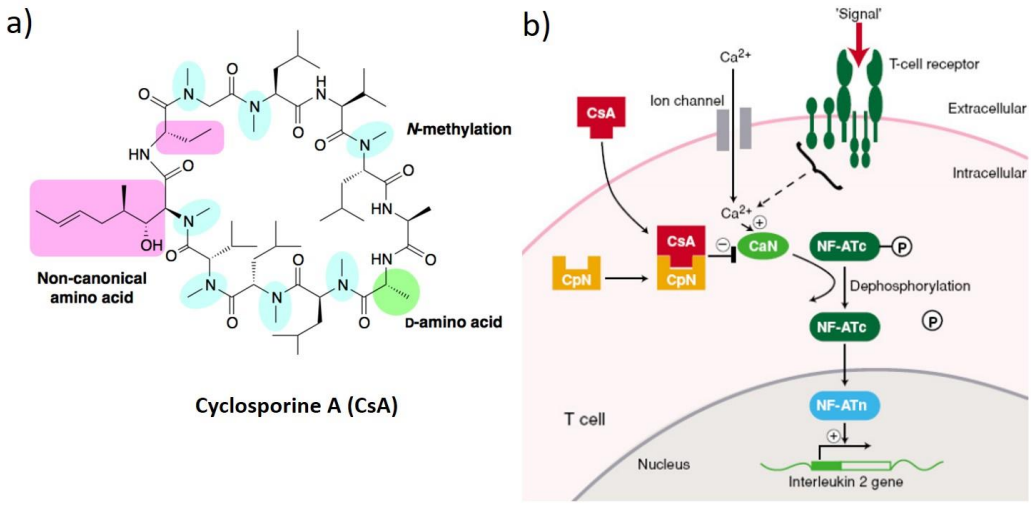


Figure 1-3. Enzyme inhibitor cyclosporine. a) Structure of Cyclosporine A. Adapted from Ref. [17]. b) Mechanism of cyclosporine action. Adapted from Ref. [16].

activated T cells (NF-ATc) can't be dephosphorylated, which blocks the transport of NF-ATc into the nucleus and inhibits the production of IL-2, a cytokine involved in the regulation of T-cell activation (Figure 1-3b)[16]. Cyclosporin A presents a highly modified macrocyclic peptide scaffold such as N methylations, non-canonical sidechains, and D-alanine residue (Figure 1-3a). It is these unique features that determine the cell membrane permeability, proteolytic stability and oral bioavailability of cyclosporine A. Cyclosporin A highlights the potential of macrocyclic peptides as drugs in the chemical space beyond the "rule of 5".

Inspired by natural products, non-traditional macrocyclic peptides containing non-canonical side chains and N-methylated backbones have been developed as protein inhibitors[17]. For example, Suga group has developed a Random non-standard Peptide Integration Discovery (RaPID) system to construct mRNA-encoding libraries of macrocyclic peptides containing diverse non-canonical amino acids. From these libraries, potent inhibitors of AKT2[18], human sirtuin2 (SIRT2)[19], and ubiquitin ligase E6AP[20] have been identified (Figure 1-4). Targeting these proteins have been challenging due to their highly conserved binding sites and large binding pockets. The peptide inhibitors have a thioether-linked macrocyclic backbone which allows them to interact with the target via a larger surface area. For targeting SIRT2, the lysine analogue presenting a trifluoroacetyl group on the ε-amino group (K^{Tfa}) was incorporated into the peptides, which promotes the discovery of low nanomolar potency inhibitors (Figure 1-4b). Remarkably, these peptides also showed isoform selectivity of greater than 10-fold relative to the other isoforms. In addition, N-methyl amino acids have been introduced into the peptide library, from which macrocyclic N-methyl-peptide inhibitors of E3 ubiquitin ligase with was single-digit

nanomolar affinity were obtained (Figure 1-4c)[20]. The macrocyclic peptide was able to inhibit E6-dependent ubiquitination of P53 and the E6-independent ubiquitination of Peroxiredoxin 1. Moreover, peptide-based inhibitors of proteases have been well investigated and will be discussed further in Chapter 2.

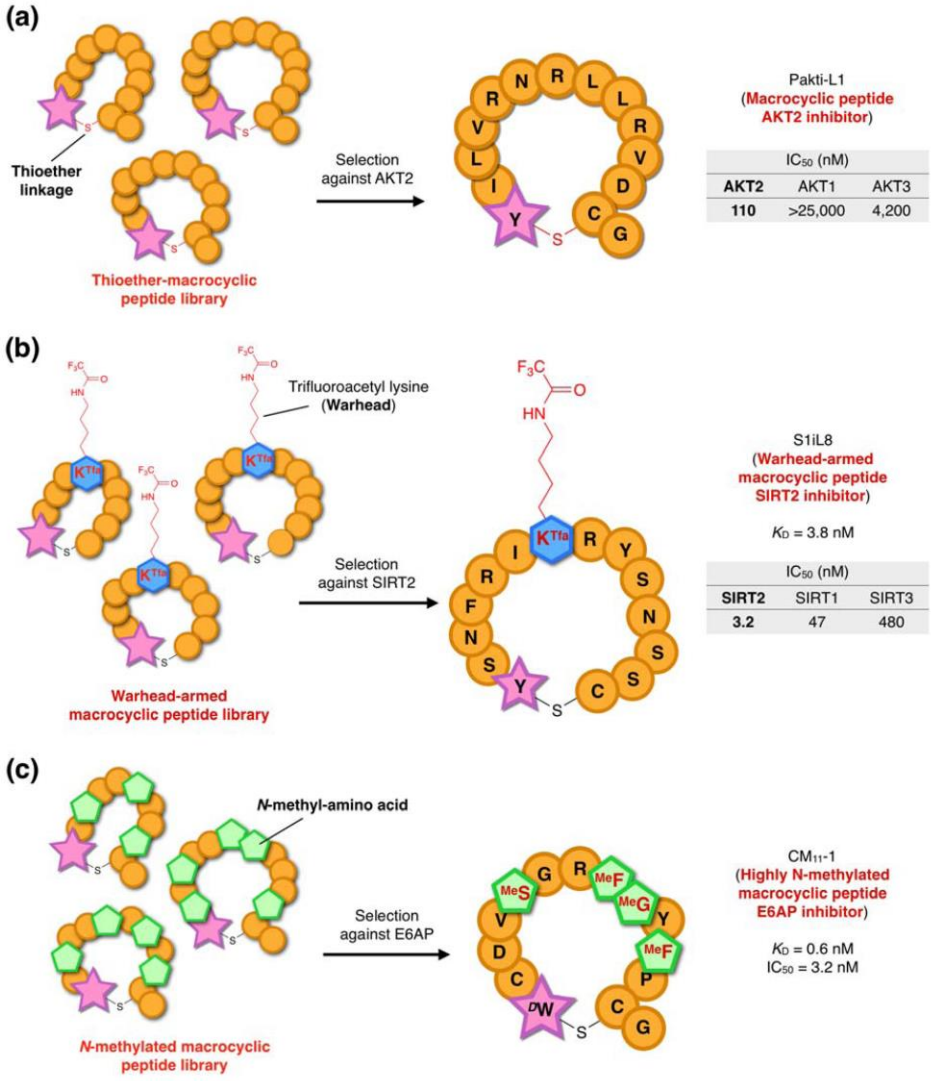


Figure 1-4. Macrocyclic peptide-based enzyme inhibitors discovered using the RaPID system. a) AKT2 isoform selective peptide inhibitor. b) SIRT2 isoform selective peptide inhibitor presenting a trifluoroacetyl Lys (KTfa) warhead. c) N-methylated thioether macrocyclic peptide targeting E6AP. Adapted from Ref. [17].

1.1.3. Peptides as protein-protein interaction inhibitors

Protein-protein interactions (PPIs) play various essential roles in life processes and are associated with diverse diseases, such as cancer, autoimmune diseases, and infectious diseases. Despite significant efforts, drugs that can modulate PPIs are still rare. This could attribute to the large interfaces of PPIs which usually reach 1000-3000 $Å^2$ and can even extend to 6000 $Å^2$ [21], however, small molecules can only afford a contact area of 300-1000 $Å^2$ flat and flexible[22]. In addition, the flat and shallow shape, or even intrinsically disordered, binding sites make PPIs "undruggable" for traditional drugs[22]. To overcome the undruggability, an increasing number of antibodies have been developed as a complement to the traditional small-molecule drugs[23]. Protein therapeutics are advantageous owing to their complex functions, safety, superior specificity and high affinity. However, the large size and solubility of antibodies may limit their route of administration, distribution and diffusion across the cell membrane[24]. Moreover, it is hard to produce the proteins with complex post-translational modifications in a large scale, which results in a high manufactory cost. As an alternative, peptides have been designed to inhibit PPIs. The molecular weight of peptide is between those of small molecules and antibodies. Therefore, peptide can interact with the protein target through larger and more extended surface area resulting in high affinity and selectivity, and meanwhile, peptides have better cell and tissue permeability than antibodies do[25]. These unique features make peptides as appealing drug candidates for modulating PPIs.

One example of peptide-based PPI modulator in clinical trials is ALRN-6924, a stapled α-helical peptide MDM2/p53 inhibitor, developed by Aileron. It shows dual MDMX/MDM2 inhibition and significantly activates p53-dependent transcription[26].

Stapled α-helical peptides have been developed as a promising modality of therapeutics. A group of MDM2 peptide inhibitors were reported by Chang et al. based on stapled peptides that mimic the key α-helical structure of p53 to bind MDM2/ MDMX (Figure 1-5)[27]. ATSP peptides afford low nanomolar affinity. The most potent peptide ATSP-7041 shows K_i of 0.9 nM and demonstrates promising p53-dependent tumor growth suppression in MDM2/MDMX-overexpressing xenograft cancer models.

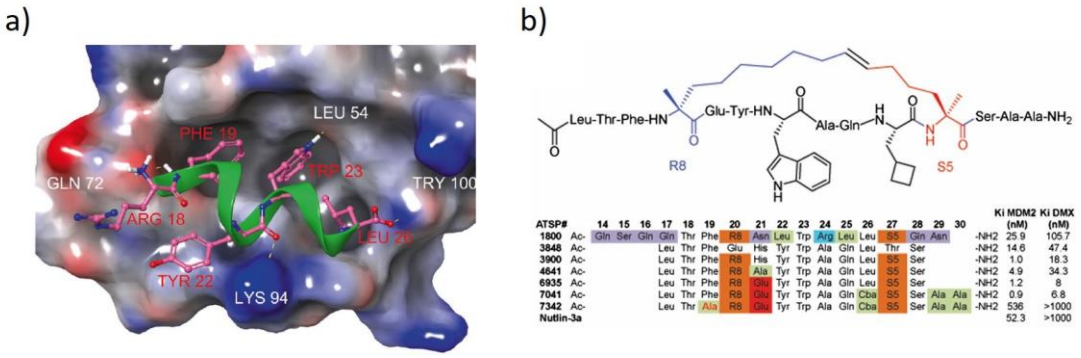
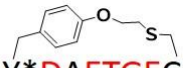


Figure 1-5. Peptide based MDM2/p53 inhibitor. a) MDM2 (surface)-p53 peptide (green) complex (PDB:1T4F). Adapted from Ref. [25]. b) Sequences and binding properties to MDM2 and MDMX of ATSP stapled peptides. Adapted from Ref. [27]. Copyright (2021) National Academy of Sciences.

Furthermore, a variety of peptide inhibitors for Keap1/ Nrf2 interaction (Kelch-like ECH-associated protein 1/nuclear factor erythroid-2) interaction, the PPI regulating cytoprotective responses to oxidative and xenobiotic stresses, have been discovered[28-30]. These peptides usually share a consensus motif DXETGE, which is part of Nrf2 and has strong binding affinity to the Kelch domain of Keap1 (Table 1-1)[31, 32]. For example, a head-to-tail cyclic peptide, Peptide 3 with the sequence of GQLDPETGEFL, was discovered by Lu et al. affording K_d of 18 nM (Figure 1-6)[28]. To improve the cell permeability of this peptide, a cyclic cell-penetrating peptide (CPP) was conjugated to it through a flexible linker[33]. The resulting peptide demonstrated good cell permeability and

biological activity at low micromolar concentrations in cell culture. In addition, Suga and his coworkers identified a linear peptide, Astp1, that has low nanomolar K_d to Keap1 by the peptidomic mRNA display. The peptide has the consensus peptide sequence (CDPETGECLCY) derived from astrotactin-1(ASTN1) and shows stronger affinity to Keap1 than the peptide derived from Nrf2 does[30]. More recently, Fasan group developed a strategy to construct libraries of macrocyclic peptides constrained by a thioether bridge on phage. By screening these libraries against Keap1, a high-affinity binder with the sequence of YDAETGEC was discovered (K_d = 40 nM).

Table 1-1. High-affinity peptide inhibitors of Keap1/Nrf2 interaction

Peptide	Sequence	K_d (nM)
Peptide 3	cyclo(GQLDPETGEFL)	18
Astp1	CDPETGECLCY	75.9
Astp1-C8S/C10S	CDPETGESLSY	13.4
KKD(6X)-m1	 Y*DAETGEC*	40

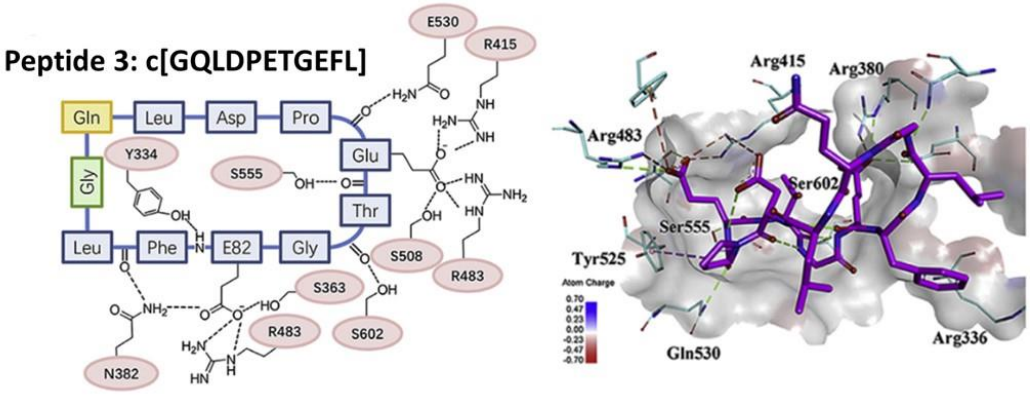


Figure 1-6. A head-to-tail cyclic peptide as potent Keap1-Nrf2 PPI inhibitor. Adapted with permission from Ref. [28].

1.2. Methods for developing functional peptides

Several strategies have been developed to identify peptide leads for drug development. When targeting a receptor with known peptide ligands, the natural peptides can provide a starting point for rational design. Otherwise, *de novo* peptide ligands need to be identified from large libraries of peptides. Biological display techniques, such as phage display and mRNA display, have been developed to discover lead peptides. These techniques encode the peptide library with DNA or RNA enabling rapid deconvolution of hits. Other techniques such as one-bead-one-compound peptide libraries[34] and intracellular peptide libraries based on SICLOPPS (split-intein circular ligation of peptides and proteins)[35] have also been successful in peptide hits identification, which are reviewed in the references and will not be discussed in this chapter.

1.2.1. Rational design

An increasing number of high-resolution structures of protein-protein and protein-peptide complexes shed light on the principles of biomolecule recognition. Nowadays,

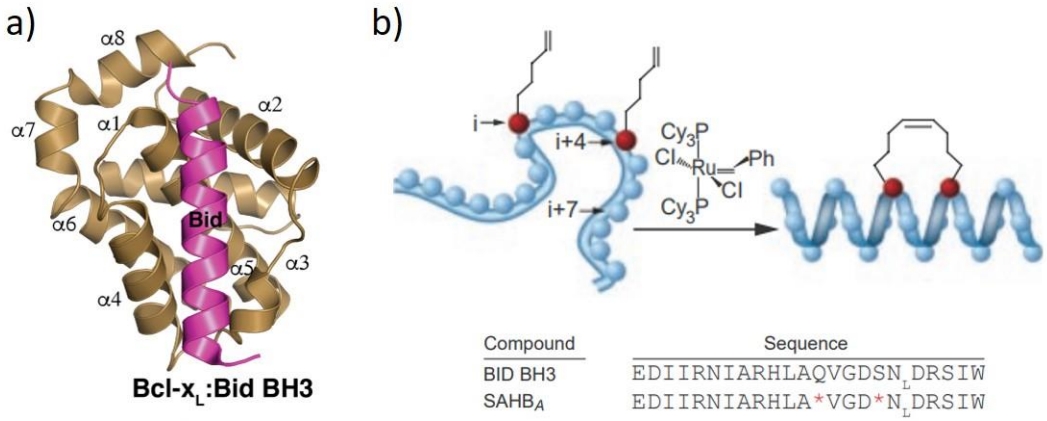


Figure 1-7. Rational design of stapled α-helical peptide. a) The 3D structure of Bcl-xL: Bid BH3 complex. b) Illustration of the strategy of stapling a peptide mimicking BID BH3 sequence.

almost 200,000 of 3D structures are available in the Protein Data Bank (PDB: www.pdb.org), which promotes the rational drug design. For example, peptide fragments derived from the crystallographic interface of a protein-protein interaction provide great starting points for rational design. For example, the amphipathic α-helical BH3 segment is essential to protein interactions between BCL-2 members (Figure 1-7a)[36]. Based on this α-helical segment, the stapled peptide $SAHB_A$ has been designed and synthesized, which demonstrated high affinity (K_d of 38.8 nM), protease-resistance and cell-permeability (Figure 1-7b)[37].

Furthermore, protein docking is a widely used tool for peptide design. Peptide-protein docking strategies fall into two categories, local or global docking. A potential binding conformation of peptide at a defined binding site of its target receptor can be identified by local docking. Many methods have been developed to refine peptide-binding sites including DynaRock, Rosetta FlexPepDock and PepCrawler. In contrast, global docking strategies are used to search for the peptide-binding site at the target protein. Docking methods have been summarized and compared in the reference [1].

There are still challenges despite recent advances in the ration design of peptide drugs. First, only 50% of all drug targets have structural information that is hard to obtain for membrane proteins[38]. Therefore, peptide ligands of protein targets with unknown structure have rarely been designed. In addition, the design of long peptide is a challenge for current computational methods. Methods for predicting structural conformation, cell permeability and pharmacokinetics of a peptide drug are still under development[39, 40].

1.2.2. Phage display

In 2018, half of the Nobel Prize in Chemistry was awarded to George P. Smith and Gregory P. Winter for the invention and development of the phage display. Phage display (Figure 1-8), first introduced by George Smith in 1985[41], is a powerful method to evolve new peptides and proteins using bacteriophage. M13 is the most widely used filamentous phage for phage display, which has a circular single-stranded DNA genome and only infects *E. coli* strains expressing the F pilus[42]. The DNA fragment of foreign peptide encoding sequence is inserted into the genome of the phage that expresses the peptide as a fusion to the coat protein, such as pIII. A phage display library contains a heterogeneous mixture of phage clones carrying different foreign DNA inserts. The peptide sequence can be randomized by inserting degenerate DNA oligonucleotides. Up to 10^9 variants of

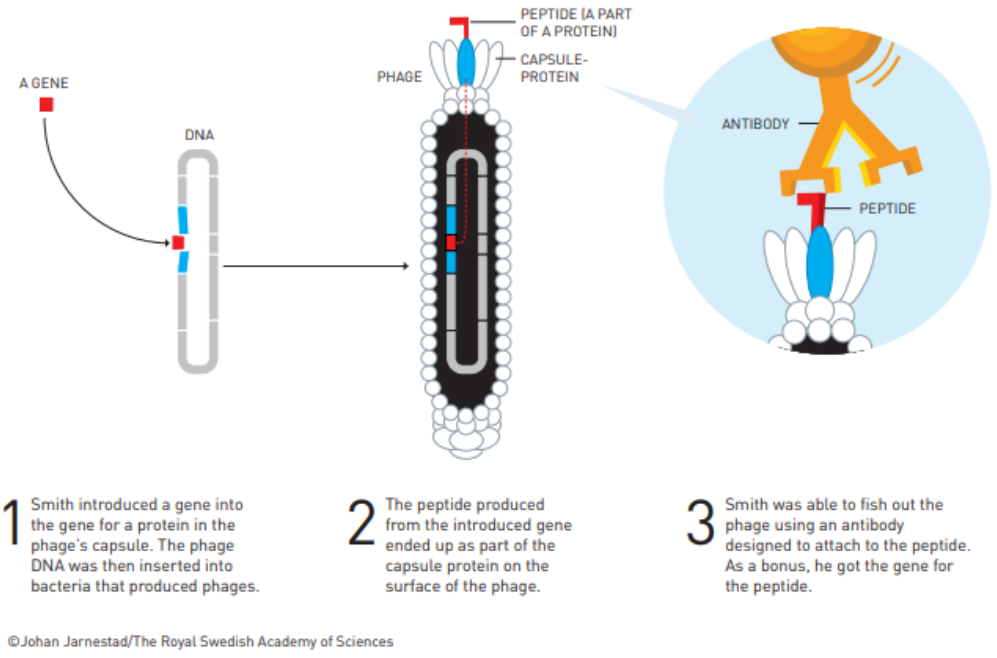


Figure 1-8. Phage display. The DNA fragment of foreign peptide encoding sequence is inserted into the genome of the phage expresses the peptide as a fusion to the coat protein.

peptides can be displayed in one library providing a great diversity. Biopanning is done by screening phage libraries against the target proteins that are usually immobilized on a solid support (Figure 1-9). The phage mixture is incubated with the target and then subjected to vigorous washing to remove unbound phage and weak binders. Then the bound phage will be eluted from the target and undergo amplification followed by another round of biopanning. The phage clones are analyzed by DNA sequencing after three to six rounds of biopanning, which will show the convergence of peptide sequence[43].

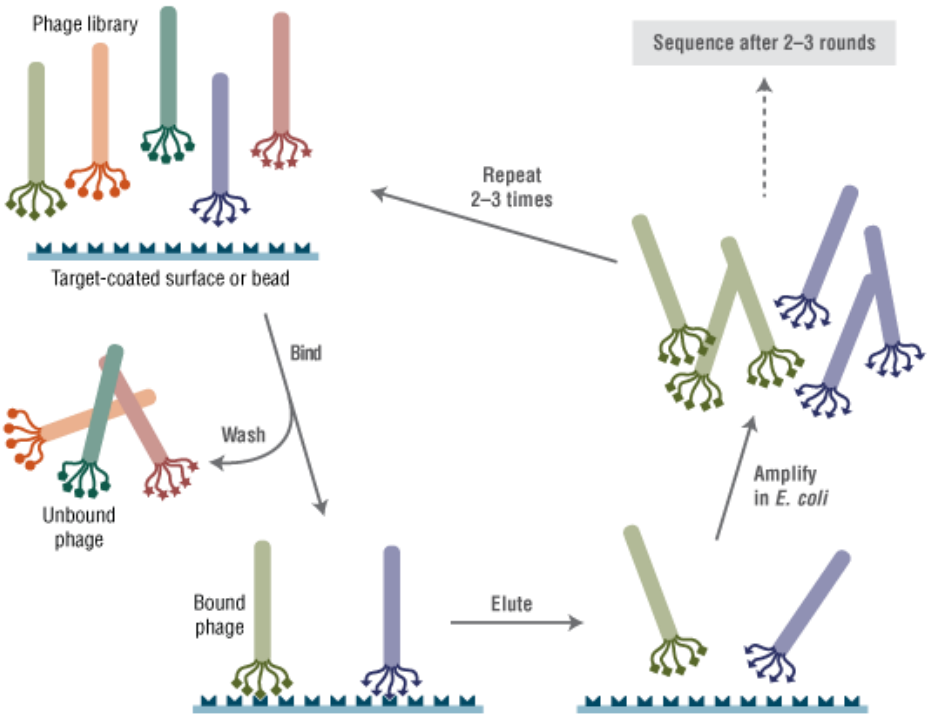


Figure 1-9. Screening of phage display library against immobilized target. Adapted from NEB-phage display instruction manual.

Table 1-2. Drugs derived from phage display

Name	Sponsor company	Target	Indication(s)	Approved time
Adalimumab (Humira)	Abbott Laboratories	Tumor necrosis factor-α (TNF)	Rheumatoid arthritis, juvenile idiopathic arthritis, psoriatic arthritis, ankylosing spondylitis, crohn's disease, ulcerative colitis, plaque psoriasis	2002
Ranibizumab (Lucentis)	Genentech	Vascular endothelial growth factor A (VEGF-A)	Neovascular (wet) age-related macular degeneration	2006
Romiplostim (Nplate)	Amgen	Thrombopoietin receptor (TPOR)	Immune thrombocytopenic purpura	2008
Ecallantide (Kalbitor)	Dyax Corp.	Plasma kallikrein	Hereditary angioedema	2009
Belimumab(Benlysta)	GlaxoSmithKline	B-lymphocyte stimulator (BLyS)	Autoantibody-positive, systemic lupus	2011
Raxibacumab (ABthrax)	Discovered by CAT technology and Human Genome SciencesGlaxoSmithKline (2012)	Protective antigen (PA) component of anthrax (Bacillus anthracis)	Prophylaxis and treatment of anthrax	2012
Necitumumab (Portrazza)	Eli Lilly and Company	EGFR	Metastatic squamous non-small-cell lung carcinoma (NSCLC)	2015
Ramucirumab (CYRAMZA)	Eli Lilly and Company	VEGFR2 (KDR)	Stomach cancer,non-small cell lung cancer (NSCLC), colorectal cancer (CRC), hepatocellular carcinoma (HCC)	2020

Greg Winter constructed a phage display library of antibodies and identified high affinity binders of target proteins such as human tumor necrosis factor α (TNFα)[44]. The first pharmaceutical derived from phage display, adalimumab, was approved in 2002, for the treatment of rheumatoid arthritis and is now also used for treating different types of psoriasis and inflammatory bowel diseases. Since then, over 40 antibody and peptide-based drugs or drug candidates have been developed from phage display for treating cancer and autoimmune diseases. The approved drugs are listed in Table 1-2 and others are summarized in the reference [45]. However, the only peptide-based drug is Ecallantide, a 60-amino acid polypeptide inhibitor of kallikrein (serine protease) for the treatment of hereditary angioedema. Developing peptide-based therapeutics with low molecular weight is still a challenge. Indeed, phage display has intrinsic limitations including that peptide encoded by phage is restrict to the canonical amino acids and short peptides are flexible, unstructured and susceptible to proteolytic degradation.

In recent years, various strategies have been developed to enhance phage display. For example, the amber stop codon suppression approach has been used to incorporate various non-canonical amino acids into phage-encoded peptides (Figure 1-10a). Schultz and his coworkers introduced sulfotyrosine[46] and bidentate metal binding amino acid bipyridylalanine[47] into the phage display system and identified high affinity anti-gp120 antibodies and Ni binding peptides with low micromolar K_d respectively. Similarly, Liu group build up a phage display library presenting N-butyryl-lysine (BuK) for targeting SIRT2[48], which yielded low nanomolar inhibitors (K_d of 10 nM) of SIRT2. These non-canonical amino acids can provide a strong driving force for ligand binding and expand the chemical space of traditional phage display. On the other hand, chemical modification of

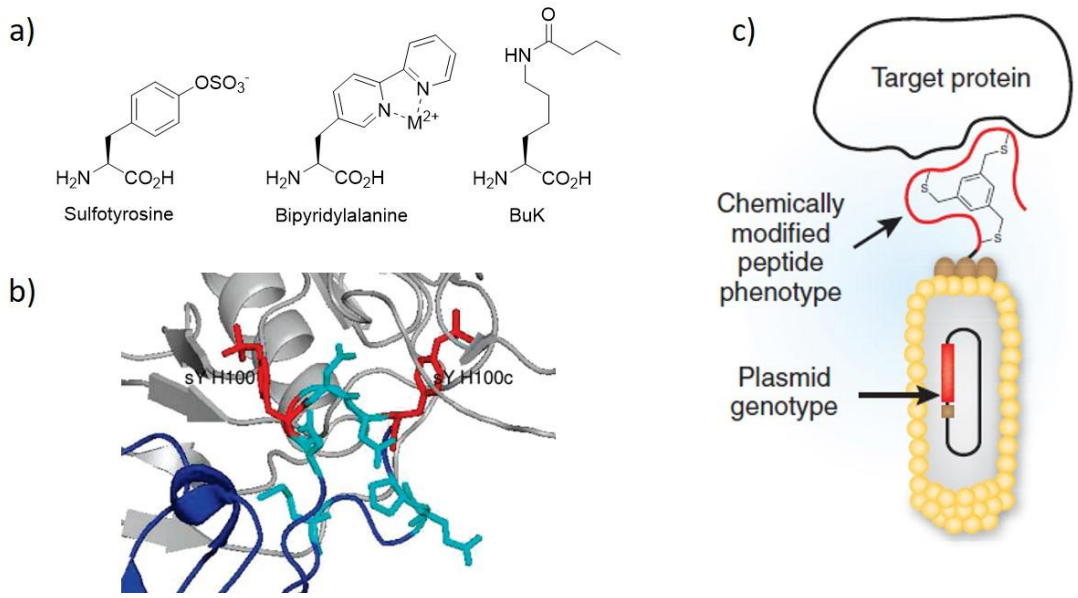


Figure 1-10. Chemical enhancement of phage display. a) Noncanonical amino acids incorporated into phage display by amber codon suppression approaches. b) Demonstration of sulfotyrosines (red) at the binding interface between anti-gp120 antibody (blue) and gp120 (gray). Residues chosen for library randomization by sitesaturation mutagenesis are colored cyan. Adapted from Ref. 46. Copyright (2009), American Chemical Society. c) Illustration of phage-encoded bicyclic peptide library.

phage is an alternative way to enhance the chemical and structural diversity of the peptides displayed by phage. The library of bicyclic peptides was developed by Heinis by chemically cyclizing random peptides displayed on phage (Figure 1-10c)[49]. By using this strategy, Bicycle Therapeutics has developed a drug candidate BT1718 for the treatment of advanced solid tumors and currently is under Phase I/IIa trial evaluating. BT1718 is a peptide-drug conjugate containing a bicyclic peptide binder of type 1 matrix metalloproteinase and a toxin[5]. The peptide ligand can deliver the toxin specifically to tumors minimizing systemic toxicity of the drug. Modification of phage-encoded peptide library through the site-specific chemistry doesn't need the sophisticated technologies for unnatural amino acid incorporation and more examples will be further demonstrated in Chapter 3.

1.2.3. mRNA display

mRNA display is another in vitro display method where peptides or proteins are covalently attached to their coding mRNA progenitor via a puromycin linkage. Two separate groups, Szostak and Roberts as well as Yanagawa and co-worker, reported this method in 1997[50, 51]. A DNAs library encoding peptide is first prepared and translated into mRNA using recombinant in vitro transcription and translation mixtures that contain ribosomes, translation factors, transfer RNA (tRNA) synthetases and amino acids. After transcription, a strand of DNA with a puromycin molecule at the 3' end is attached to each mRNA. In the translation step, the ribosome reads mRNA and synthesized the peptide from

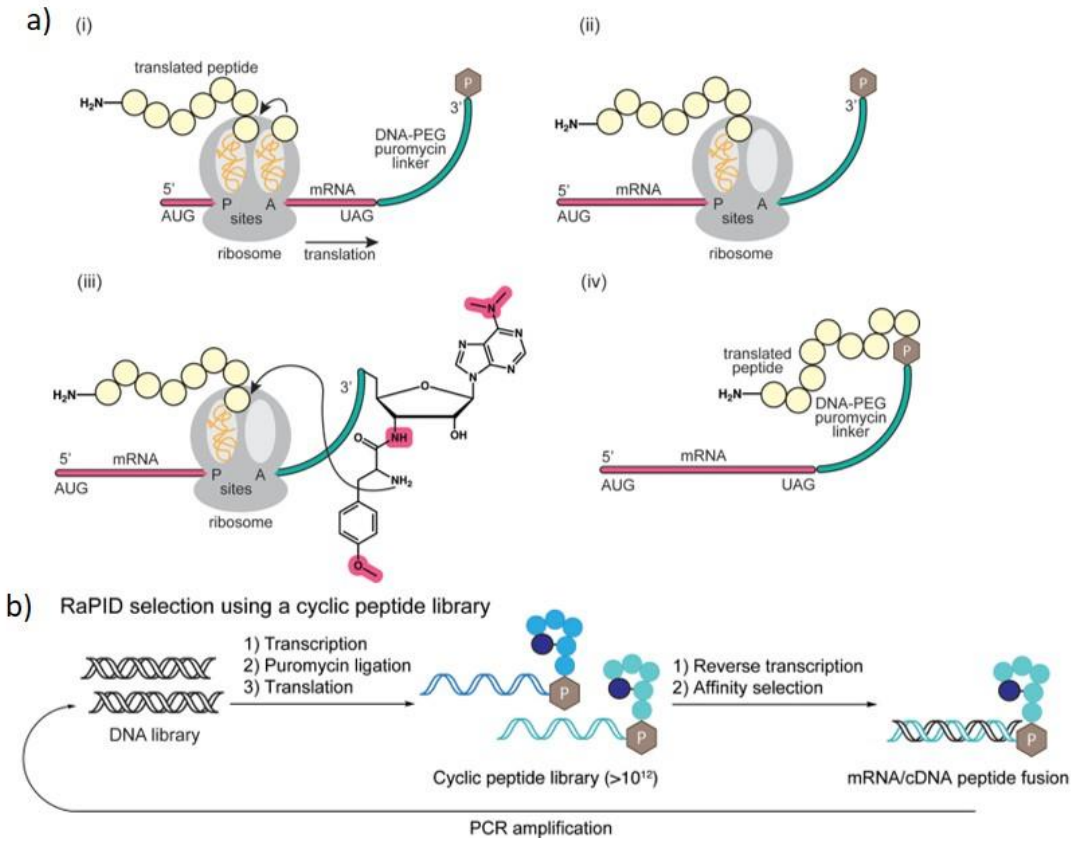


Figure 1-11. mRNA display. a) The generation of mRNA-peptide fusions via a puromycin linker. b) Illusion of RaPID system that allows the synthesis of cyclic peptide libraries. Reprinted with permission from Ref. 49. Copyright (2019) American Chemical Society.

5’ to 3’. At the end of translation, the ribosome reaches the stop codon and gets close to the puromycin molecule that is then attached covalently onto the peptide chain in the P site of the ribosome. Therefore, each peptide is linked to its encoding mRNA strand (Figure 1-11a). Then, revers transcription of the mRNA-peptide fusions gives a library of mRNA-cDNA-peptide fusions that undergoes the screening against the immobilized protein target. The cDNA of bound peptides is amplified by PCR and serves as a template in the next round of panning to enrich the peptide binders. Finally, strong peptide binders can be identified by analyzing the DNA sequences.

So far, various mRNA libraries have been built to display linear peptides, disulfide-rich peptides, macrocyclic peptides with thioether cyclization and peptides containing non-canonical amino acids, summarized in a recent review[52]. Especially, the RaPID (random nonstandard peptide integrated discovery) system established by Suga and co-workers have led to great success in developing high-affinity macrocyclic peptide binders of diverse protein targets (Figure 1-11b). In this system, flexizymes are used, which can recognize a variety of activated non-canonical amino acids and charge them onto the desired tRNA. Therefore, unnatural amino acids can be incorporated into the peptide library by reprogramming the genetic codons. The combination of flexizyme and PURE (The Protein synthesis Using Recombinant Elements) components leads to the generation of libraries presenting D-amino acids, N-methylated amino acids, β-amino acids, γ-amino acids and warheads.

Based on advanced mRNA display technologies, several companies have been established and developed drugs candidates in the clinical trials. For example, Peptidream, a biotech company found in 2006, established a Peptide Discovery Platform System based

on mRNA display to discover macrocyclic peptide-based drugs[53]. Currently, it has about 100 ongoing programmes, collaborating with 18 companies, such as Bristol-Myers Squibb (BMS) and Novartis. Partnered with BMS, the first drug candidate entered the clinic trail in 2016, which is a constrained peptide against the checkpoint protein PDL1[54]. Additionally, RA Pharmaceuticals developed a macrocyclic peptide inhibitor of complement C5 inhibitor, Zilucoplan (RA101495), which is under Phase III trial for the treatment of generalized myasthenia gravis. Zilucoplan can bind C5 with subnanomolar affinity and allosterically inhibit the cleavage of C5 into C5a and C5b[55, 56]. In addition, it can also block the binding of C5b to complement component C6. This is the first-in-class non-antibody C5 inhibitor.

mRNA display systems have some advantages in contrast to phage display. First, mRNA display doesn't need a transformation step that limits the diversity of library, and therefore libraries with larger diversities ($> 10^{12}$) can be constructed. Furthermore, PCR-based amplification is easy and robust which avoids the growth bias accumulated during phage amplification. In addition, by reprogramming the genetic codons, numerous non-canonical amino acids can be incorporated efficiently. However, the negatively charged mRNA tag may interact with highly positively charged targets resulting in the nonspecific binding[57]. Also, mRNA/cDNA fusion can interact with proteins that can bind nucleic acids like transcription factors.

1.3. Conclusions

Hundreds of peptide-based therapeutics have been developed or are being developed to treat infectious diseases, metabolic disorders and cancer. This field is now attracting unprecedented attention and interest from both academia and pharmaceutical industry.

With the new technologies, a wide variety of biological targets, including those used to be considered as undruggable, have been successfully modulated by peptides. The incorporation of non-canonical amino acids, macrocyclic structures and chemical warheads into the peptide libraries has significantly promoted the development of peptide-based drugs.

1.4. References

1. Lee, A.C.-L., et al., A comprehensive review on current advances in peptide drug development and design. International journal of molecular sciences, 2019. **20**(10): p. 2383.
2. Henninot, A., J.C. Collins, and J.M. Nuss, The Current State of Peptide Drug Discovery: Back to the Future? Journal of Medicinal Chemistry, 2018. **61**(4): p. 1382-1414.
3. Tsomaia, N., Peptide therapeutics: Targeting the undruggable space. European Journal of Medicinal Chemistry, 2015. **94**: p. 459-470.
4. Vinogradov, A.A., Y. Yin, and H. Suga, Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. Journal of the American Chemical Society, 2019. **141**(10): p. 4167-4181.
5. Deyle, K. and C. Heinis, Drugs Based on de novo-developed Peptides are Coming of Age mune is a. CHIMIA International Journal for Chemistry, 2018. **72**(6): p. 426-427.
6. Hancock, R.E. and D.S. Chapple, Peptide antibiotics. Antimicrobial agents and chemotherapy, 1999. **43**(6): p. 1317-1323.

7. Wiese, A., T. Gutsmann, and U. Seydel, Towards antibacterial strategies: studies on the mechanisms of interaction between antibacterial peptides and model membranes. J Endotoxin Res, 2003. **9**(2): p. 67-84.

8. Stone, K.J. and J.L. Strominger, Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate. Proceedings of the National Academy of Sciences of the United States of America, 1971. **68**(12): p. 3223-3227.

9. Ganz, T. and R.I. Lehrer, Antimicrobial peptides of leukocytes. Curr Opin Hematol, 1997. **4**(1): p. 53-8.

10. Xie, Z., et al., Human α-defensins are immune-related Kv1.3 channel inhibitors: new support for their roles in adaptive immunity. Faseb j, 2015. **29**(10): p. 4324-33.

11. Hong, J., et al., How Melittin Inserts into Cell Membrane: Conformational Changes, Inter-Peptide Cooperation, and Disturbance on the Membrane. Molecules, 2019. **24**(9): p. 1775.

12. Wieprecht, T., et al., Modulation of membrane activity of amphipathic, antibacterial peptides by slight modifications of the hydrophobic moment. FEBS letters, 1997. **417**(1): p. 135-140.

13. Hancock, R.E., Peptide antibiotics. The Lancet, 1997. **349**(9049): p. 418-422.

14. Radermacher, S., V. Schoop, and H. Schluesener, Bactenecin, a leukocytic antimicrobial peptide, is cytotoxic to neuronal and glial cells. Journal of neuroscience research, 1993. **36**(6): p. 657-662.

15. Lipinski, C.A., et al., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of

original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25.1. Advanced Drug Delivery Reviews, 2001. **46**(1): p. 3-26.

16. Colombo, D. and A. Di Pietro, Systemic Cyclosporin in the Treatment of Psoriasis. 2012: INTECH Open Access Publisher.

17. Ito, K., T. Passioura, and H. Suga, Technologies for the Synthesis of mRNA-Encoding Libraries and Discovery of Bioactive Natural Product-Inspired Non-Traditional Macrocyclic Peptides. Molecules, 2013. **18**(3): p. 3502-3528.

18. Hayashi, Y., J. Morimoto, and H. Suga, In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. ACS Chemical Biology, 2012. **7**(3): p. 607-613.

19. Morimoto, J., Y. Hayashi, and H. Suga, Discovery of Macrocyclic Peptides Armed with a Mechanism-Based Warhead: Isoform-Selective Inhibition of Human Deacetylase SIRT2. Angewandte Chemie International Edition, 2012. **51**(14): p. 3423-3427.

20. Yamagishi, Y., et al., Natural Product-Like Macrocyclic N-Methyl-Peptide Inhibitors against a Ubiquitin Ligase Uncovered from a Ribosome-Expressed De Novo Library. Chemistry & Biology, 2011. **18**(12): p. 1562-1570.

21. Smith, M.C. and J.E. Gestwicki, Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity. Expert Reviews in Molecular Medicine, 2012. **14**: p. e16.

22. Scott, D.E., et al., Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. Nature Reviews Drug Discovery, 2016. **15**(8): p. 533-550.

23. Scott, A.M., J.D. Wolchok, and L.J. Old, Antibody therapy of cancer. Nature Reviews Cancer, 2012. **12**(4): p. 278-287.

24. Leader, B., Q.J. Baca, and D.E. Golan, Protein therapeutics: a summary and pharmacological classification. Nature Reviews Drug Discovery, 2008. **7**(1): p. 21-39.

25. Lu, H., et al., Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 213.

26. Carvajal, L.A., et al., Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. Science Translational Medicine, 2018. **10**(436): p. eaao3003.

27. Chang, Y.S., et al., Stapled α−helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. Proceedings of the National Academy of Sciences, 2013. **110**(36): p. E3445-E3454.

28. Lu, M.-C., et al., Discovery of a head-to-tail cyclic peptide as the Keap1-Nrf2 protein-protein interaction inhibitor with high cell potency. European Journal of Medicinal Chemistry, 2018. **143**: p. 1578-1589.

29. Owens, A.E., et al., MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles. ACS Central Science, 2020. **6**(3): p. 368-381.

30. Hirose, H., et al., A Case Study on the Keap1 Interaction with Peptide Sequence Epitopes Selected by the Peptidomic mRNA Display. ChemBioChem, 2019. **20**(16): p. 2089-2100.

31. Lo, S.-C., et al., Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. The EMBO Journal, 2006. **25**(15): p. 3605-3617.

32. Tong, K.I., et al., Keap1 Recruits Neh2 through Binding to ETGE and DLG Motifs: Characterization of the Two-Site Molecular Recognition Model. Molecular and Cellular Biology, 2006. **26**(8): p. 2887-2900.

33. Salim, H., et al., Development of a Cell-Permeable Cyclic Peptidyl Inhibitor against the Keap1–Nrf2 Interaction. The Journal of Organic Chemistry, 2020. **85**(3): p. 1416-1424.

34. Bononi, F.C. and L.G. Luyt, Synthesis and cell-based screening of one-bead-one-compound peptide libraries. Peptide Libraries, 2015: p. 223-237.

35. Sohrabi, C., A. Foster, and A. Tavassoli, Methods for generating and screening libraries of genetically encoded cyclic peptides in drug discovery. Nature Reviews Chemistry, 2020. **4**(2): p. 90-101.

36. Banjara, S., et al., Structural Insight into African Swine Fever Virus A179L-Mediated Inhibition of Apoptosis. Journal of Virology, 2017. **91**(6): p. e02228-16.

37. Walensky, L.D., et al., Activation of Apoptosis in Vivo by a Hydrocarbon-Stapled BH3 Helix. Science, 2004. **305**(5689): p. 1466-1470.

38. Tanrikulu, Y. and G. Schneider, Pseudoreceptor models in drug design: bridging ligand- and receptor-based virtual screening. Nature Reviews Drug Discovery, 2008. **7**(8): p. 667-677.

39. Hosseinzadeh, P., et al., Comprehensive computational design of ordered peptide macrocycles. Science, 2017. **358**(6369): p. 1461-1466.

40. Wolfe, J.M., et al., Machine Learning To Predict Cell-Penetrating Peptides for Antisense Delivery. ACS central science, 2018. **4**(4): p. 512-520.

41. Smith, G.P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, 1985. **228**(4705): p. 1315-1317.

42. Pande, J., M.M. Szewczyk, and A.K. Grover, Phage display: Concept, innovations, applications and future. Biotechnology Advances, 2010. **28**(6): p. 849-858.

43. Smith, G.P. and V.A. Petrenko, Phage Display. Chemical Reviews, 1997. **97**(2): p. 391-410.

44. Jespers, L.S., et al., Guiding the Selection of Human Antibodies from Phage Display Repertoires to a Single Epitope of an Antigen. Bio/Technology, 1994. **12**(9): p. 899-903.

45. Nixon, A.E., D.J. Sexton, and R.C. Ladner, Drugs derived from phage display: from candidate identification to clinical practice. mAbs, 2014. **6**(1): p. 73-85.

46. Liu, C.C., et al., Mutagenesis and Evolution of Sulfated Antibodies Using an Expanded Genetic Code. Biochemistry, 2009. **48**(37): p. 8891-8898.

47. Day, J.W., et al., Identification of metal ion binding peptides containing unnatural amino acids by phage display. Bioorganic & Medicinal Chemistry Letters, 2013. **23**(9): p. 2598-2600.

48. Tharp, J.M., et al., An amber obligate active site-directed ligand evolution technique for phage display. Nature Communications, 2020. **11**(1): p. 1392.

49. Heinis, C., et al., Phage-encoded combinatorial chemical libraries based on bicyclic peptides. Nature chemical biology, 2009. **5**(7): p. 502.

50. Roberts, R.W. and J.W. Szostak, RNA-peptide fusions for the in vitro selection of peptides and proteins. Proceedings of the National Academy of Sciences, 1997. **94**(23): p. 12297-12302.

51. Nemoto, N., et al., In vitro virus: bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. FEBS letters, 1997. **414**(2): p. 405-408.

52. Huang, Y., M.M. Wiedmann, and H. Suga, RNA Display Methods for the Discovery of Bioactive Macrocycles. Chemical Reviews, 2019. **119**(17): p. 10360-10391.

53. Ishizawa, T., et al., TRAP Display: A High-Speed Selection Method for the Generation of Functional Polypeptides. Journal of the American Chemical Society, 2013. **135**(14): p. 5433-5440.

54. Morrison, C., Constrained peptides' time to shine? Nature Reviews Drug Discovery, 2018. **17**(8): p. 531-533.

55. Ricardo, A., et al., Preclinical Evaluation of RA101495, a Potent Cyclic Peptide Inhibitor of C5 for the Treatment of Paroxysmal Nocturnal Hemoglobinuria. Blood, 2015. **126**(23): p. 939-939.

56. Howard, J.F., Jr, et al., Clinical Effects of the Self-administered Subcutaneous Complement Inhibitor Zilucoplan in Patients With Moderate to Severe Generalized Myasthenia Gravis: Results of a Phase 2 Randomized, Double-Blind, Placebo-Controlled, Multicenter Clinical Trial. JAMA Neurology, 2020. **77**(5): p. 582-592.

57. Lamboy, J.A., et al., Chemical and genetic wrappers for improved phage and RNA display. ChemBioChem, 2008. **9**(17): p. 2846-2852.

Chapter 2 Developing synthetic receptors by rational design of prolinomycin-based scaffold

2.1. Introduction

Nature has evolved proteins and RNA to carry out complicated chemical processes, such as catalysis, specific binding and directed flow of electrons. The characteristic compact conformations of these large biological polymers are critical for their functions. Inspired by nature, chemists have a great passion to develop low-molecular weight biomimetic polymers for specific molecular recognition[1-3]. Peptides and peptidomimetics are promising in this aspect. On the one hand, peptides provide a modular scaffold to display diverse functional groups to engage the target through multiple interactions. On the other hand, the conformation and flexibility of peptides can be modulated which allow them to form stable secondary structures.

In this chapter, I will first introduce peptide-based scaffold including β-peptide foldamers and peptide macrocycles. Then, a novel scaffold based on prolinomycin will be introduced. In addition, the investigation and improvement we have done on the prolinomycin scaffold will be described. The research is focusing on two aspects: structure and function. In respect of structure, the rigidity and potassium binding propensity of prolinomycin mutants are improved by introducing proline modifications and cross-linking moieties. The function of the scaffold is extended to interact with small molecules by incorporating the warheads targeting amine groups via iminoboronate formation. Furthermore, the dimerization of a pair of prolinomycin mutants was also explored.

2.1.1. *β-Peptide foldamers*

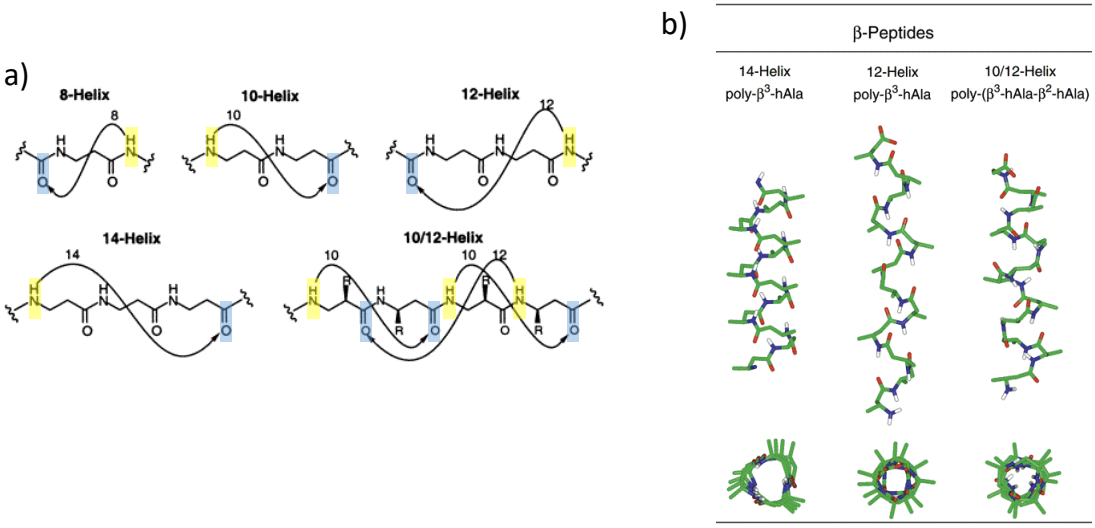


Figure 2-1. β-peptide helical conformations. a) Demonstration of H-bonds formed in β-peptide helical conformations. b) Structure of the 14-helix, 12-helix, and 10/12-helix. Carbon atoms are shown in green, nitrogen in blue, and oxygen in red. Adapted with permission from Ref. [2]. Copyright © 2001, American Chemical Society

Foldamers are synthetic oligomers that are able to form compact conformations to mimic folding patterns of biomolecules like proteins and nucleic acid. Since the mid-1990s, many types of oligomers have been developed to adopt secondary structures that are specified by backbone H-bonding or π-π stacking[4, 5]. β-peptide foldamer, containing exclusively β-amino acid residues, is one of the most thoroughly characterized types of foldamer. Through intramolecular H-bonds, β-peptides tend to adopt helices and hairpin-type structures. Based on hydrogen-bonding patterns, various helical conformations have been formed by β-Peptide, including 8-, 10-, 12-, 14- and10/12-helix (Figure 2-1) [2]. These stable conformations allow researchers to design functional mimics of natural peptides and proteins.

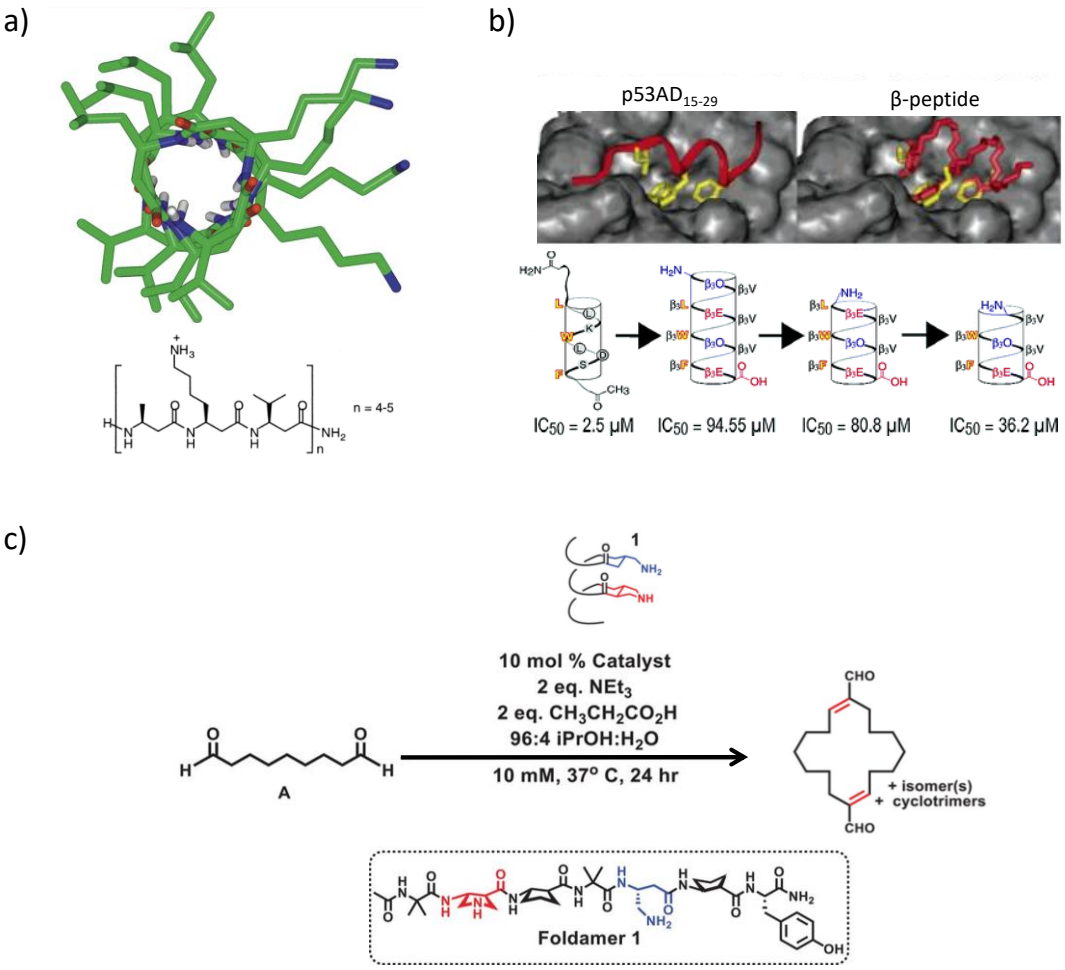


Figure 2-2. Fuction of β-peptid. a) Molecular model of an amphiphilic β-peptide. The hydrophobic and positively charged residues are displayed on opposite sides of the helix. Carbon atoms are shown in green, nitrogen in blue, and oxygen in red. Adapted with permission from Ref. [8]. Copyright © 1999, American Chemical Society. b) The β-decapeptide inhibitor of the p53-hDM2 interaction presents its leucine, tryptophan, and phenylalanine side chains in a similar arrangement to the corresponding α-peptide. Adapted with permission from Ref. [9]. Copyright © 2004, American Chemical Society c) Foldamer-catalyzed cyclodimerization of dialdehyde A. Adapted with permission from Ref. [11].

Antibacterial and antifungal β-peptides have been developed by mimicking natural membrane-active peptide toxins and antibiotics[6, 7]. The 14-helix structure of β-peptides allows the segregation of hydrophobic and hydrophilic residues on the opposite side of the helix, which is critical for their biological activities (Figure 2-2a)[8]. In addition, β-peptides

are able to mimic helices involved in protein-protein interactions. For example, β-peptide inhibitors of the p53-hDM2 interaction were designed by presenting the leucine, tryptophan, and phenylalanine side chains on the same face of 14-helix (Figure 2-2b). Therefore, these three residues are arranged in the similar way to the corresponding α-peptide and successfully mimic the functional epitope[9]. Furthermore, β-peptides have also been developed as foldamer catalysts for formation or cleavage of carbon−carbon bonds[10]. Recently, a foldamer catalyst for macrocycle formation was designed by displaying a catalytic primary amine–secondary amine diad in a specific arrangement on an α/β peptide foldamers, the oligomers containing both α- and β-amino acid (Figure 2-2c) [11]. This contribute demonstrated the robust utility of this foldamer scaffold to achieve the conformational preorganization of reactive groups.

Overall, β-peptide foldamers provide a powerful platform to arrange functional groups in a predictable way. The well-defined secondary structures allow the scaffold to mimic protein fragments and harness principles of enzyme function for catalyst design.

2.1.2. Peptide macrocycles

In addition to foldamers, another group of oligomers that has a strong tendency to form rigid structures is peptide and peptidomimetic macrocycles[12]. Macrocyclization is able to preorganize diverse functionality on the restricted ring structure which can lead to high affinity to the target molecules by reducing the entropy cost for binding[13, 14]. For example, a structurally well-defined macrocyclic scaffold with antiparallel β-sheets was used to

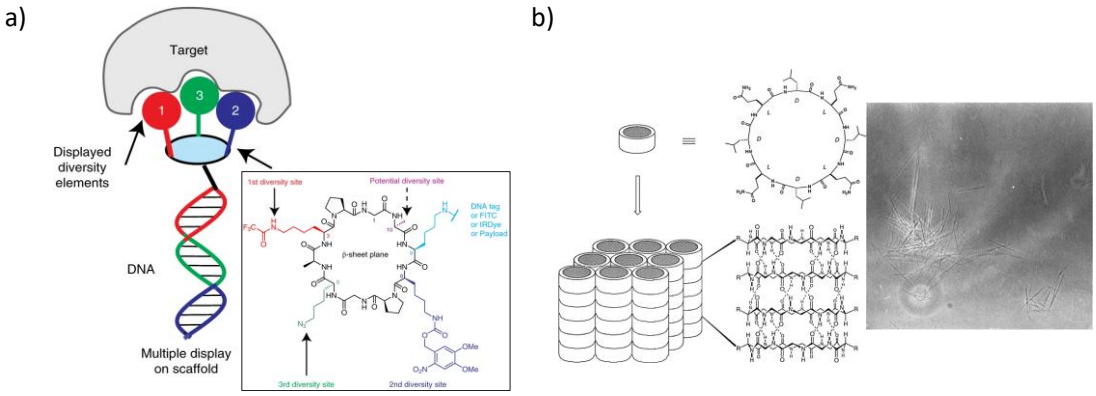


Figure 2-3. Peptide macrocycles scaffold. a) DNA-encoded chemical libraries based on a well-defined macrocyclic scaffold. Adapted with permission from Ref. [15]. b) Nanotubes constructed by cyclic peptide with flat ring-shaped conformations through highly oriented intersubunit hydrogen-bonding. Adapted with permission from Ref. [16]. Copyright © 1996 American Chemical Society.

identify chemical probes of various protein targets (Figure 2-3a)[15]. The scaffold provided multiple diversity sites for chemical modifications, which allows the synthesis of DNA-encoded chemical libraries. The diversity elements were displayed on the macrocyclic scaffold in a predefined spatial orientation which was crucial for binding.

The cyclic peptides consisting of alternating D- and L-amino acid residues are able to adopt a flat ring-shaped conformation, which promotes the self-assembling of peptides through contiguous intermolecular H-bonding (Figure 2-3b)[16]. Nanotubes were constructed by these peptide self-assembly subunits, which showed good stabilities in

water and common organic solvent. The highly oriented backbone-backbone hydrogen-bonding network is necessary for nanotube's stability. This scaffold provides a strategy to rationally design peptide nanotubes with predictable size and structure.

2.1.3. Prolinomycin

In our lab, a new strategy has been developed by taking the advantage of the metal-nucleated folding of the prolinomycin-based scaffold. Prolinomycin is a 12-residue cyclized peptide, an invention of Bruce Merrifield which was first reported as an analogue of valinomycin, a potassium ion transporter (Figure 2-4)[17]. Replacing the six hydroxy acids with prolines gives prolinomycin, however proline residues prevent the peptide from dissociating the metal ion efficiently due to the restricted freedom of rotation, making it a poor K^+ transporter[18]. Nevertheless, this proline rich cyclic peptide gives us a rigid, stable and predictable structure to explore alternative applications.

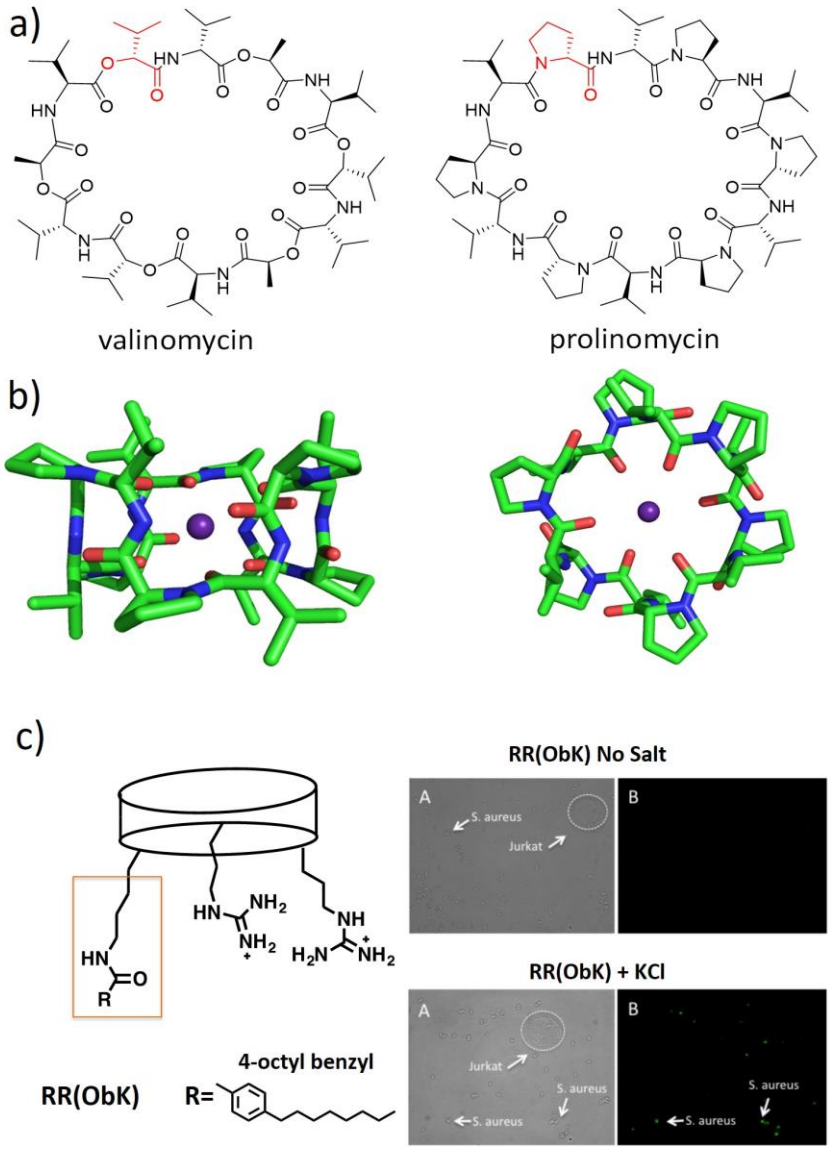


Figure 2-4. Prolinomycin as a peptide scaffold. a) Structure of valinomycin and prolinomycin. b) Side view (left) and top view (right) of the prolinomycin-K^+ complex. C: green; N: blue; O: red; K: purple. c) K^+-dependent staining of S. aureus by the prolinomycin mutant RR(ObK). Adapted with permission from Ref. [19].

In the previous study, the scaffold has been found to tolerate various side chain mutations at multiple positions without interfering the metal-induced folding. The prolinomycin mutants adopt a drum like structure preorganizing side chains for target binding. Encouragingly, cationic mutants were successfully developed to selectively bind anionic vesicles and bacterial cells under physiological conditions[19]. Attracted by its

beautiful structure, we believed that prolinomycin scaffold provides a good starting point to design peptides with various functions.

2.2. Improve prolinomycin potassium binding ability

Various prolinomycin mutants have been made presenting Arg, Lys, and Trp at multiple non-proline positions. Their metal-induced folding was determined by ^{1}H-NMR titration. When potassium was added into the peptide solution, sharp peaks showed up in the amide

Table 2-1. Potassium binding affinity of prolinomycin mutants

Mutant	Sequence	K_d
VVV	Cyclo-[tPro-Val-Pro-Val-Pro-Val-Pro-Val-Pro-Val-Pro-Ala-]	5mM
KKK	Cyclo-[tPro-Lys-Pro-Val-Pro-Lys-Pro-Val-Pro-Lys-Pro-Ala-]	8mM
RRR	Cyclo-[tPro-Arg-Pro-Val-Pro-Arg-Pro-Val-Pro-Arg-Pro-Ala-]	2mM
VRR-4R	Cyclo-[tPro-Val-Pro-Arg-Pro-Arg-Pro-Val-Pro-Arg-Pro-Ala-]	1mM
RRR-4W	Cyclo-[tPro-Arg-Pro-Trp-Pro-Arg-Pro-Val-Pro-Arg-Pro-Ala-]	0.5mM
RRRWT	Cyclo-[tPro-Arg-Pro-Trp-Pro-Arg-Pro-Thr-Pro-Arg-Pro-Ala-]	0.1mM

K_d values were determined in 10% D_2O, 50% CD_3OH NaPi solution

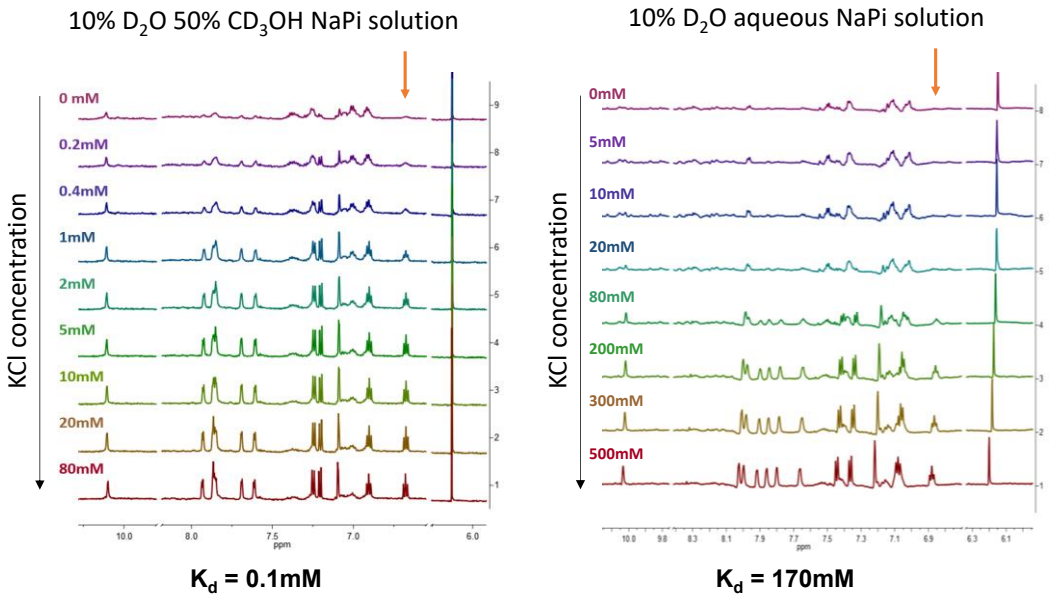


Figure 2-5. ^{1}H NMR of RRRWT titrated with potassium in water: methanol 1:1 mixture (left) versus pure aqueous buffer (right).

region which were consistent with the amide hydrogens forming intramolecular hydrogen bond, indicating the folding of peptide. The potassium binding curve was obtained by plotting the KCl concentration versus the integrated peak area. The K_d values were determined from the binding curves. As shown in Table 2-1, in the water: methanol 1:1 mixture the prolinomycin mutants can fold with low millimolar potassium ions (Table 2-1). Especially, RRRWT showed the best potassium binding affinity with K_d of 0.1 mM. However, in the pure aqueous solutions hundreds of millimolar potassium ions are necessary to induce the folding of the scaffold, which is not ideal for biological applications (Figure 2-5).

2.2.1. Modifications on the backbone proline residues

To improve the potassium binding ability, we chose RRRWT as the model peptide and first investigated the effect of modifications on the proline residues. Proline residues are essential for the structure of prolinomycin, which constitute the backbone of the scaffold. Compared to the wild type prolinomycin, our synthetic analogues have a thiol group on one proline residue (tPro) which may affect peptide's ability to fold.

To test this hypothesis, a dethiolated mutant, RRRWTdeSH, was obtained through the raney nickel reduction of RRRWT (Figure 2-6a). The potassium binding ability was evaluated by KCl titration performed by NMR spectroscopy (Figure 2-6b). The results showed the dethiolation of RRRWT increased K_d by 5 folds, which indicated that the thiol on the proline could enhance the peptide's propensity to bind potassium (Figure 2-6c).

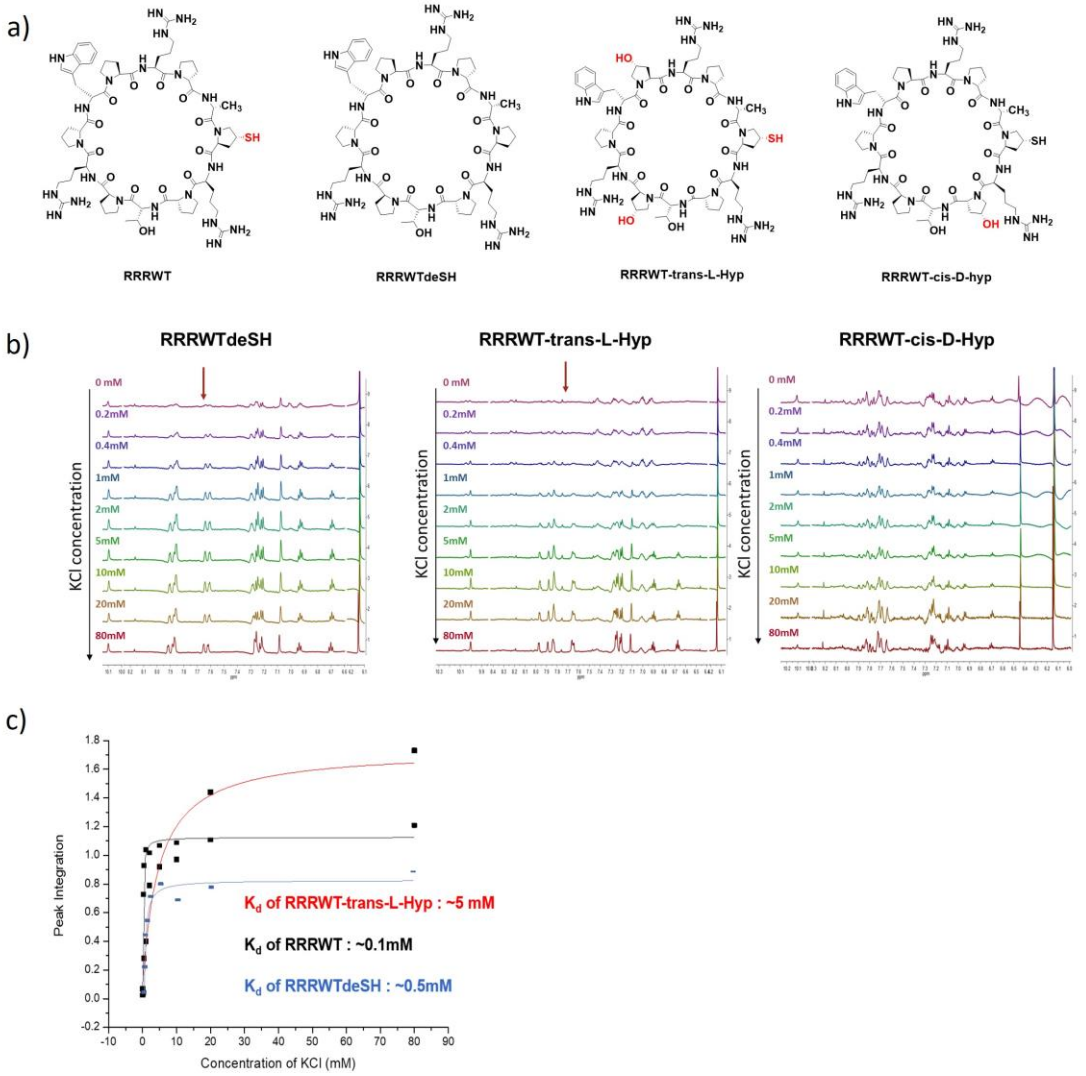


Figure 2-6. Prolinomycin mutants with proline residue modifications. a) Structures of prolinomycin mutants. b) K^+-nucleated folding of RRRWTdeSH, RRRWT-trans-L-Hyp, and RRRWT-cis-D-Hyp monitored by ^{1}H-NMR in the 10% D_2O 50% CD_3OH NaPi solution. c) Potassium binding curves and K_d values.

It is well known that the collagen triple helix is stabilized by (2S, 4R) 4-hydroxyproline (Hyp) since the hydroxy group can lock the conformation of the proline ring and favor triple helix formation[20]. The thiol group on the proline might behave in the similar mechanism to promote the folding of prolinomycin mutants. Based on this hypothesis, two (2S, 4R) 4-hydroxyprolines were incorporated into RRRWT to further improve its

potassium binding ability. Surprisingly, the resulting mutant, called RRRWT-trans-L-Hyp, showed a much higher K_d value than RRRWTdeSH and RRRWT which suggested that the conformation locked by this trans-L-Hyp impeded the folding of peptide (Figure 2-6b). Therefore, we turned to another hydroxyproline, (2R, 4R) 4-hydroxyproline which should lock the proline ring in a different conformation. RRRWT-cis-D-Hyp was made, however, KCl titration didn't make a significant difference on its NMR spectrum indicating that the potassium couldn't induce the folding anymore (Figure 2-6b). The cis-D-Hyp may force the peptide to form other conformations that can't respond to the potassium.

In conclusion, the prolinomycin scaffold is sensitive to the modifications on the proline. The tPro can increase the peptide's propensity to bind potassium and fold, however, the Hyp has an opposite effect. RRRWT shows the smallest K_d value among these mutants, about 170mM in the pure aqueous solution which still need to be improved.

2.2.2. Cross-linked prolinomycin mutants

To further improve the potassium binding ability of the scaffold, another strategy was developed by introducing cross-linking moieties. Taking advantage of scaffold's high tolerance towards mutations on the non-proline residues, we proposed that by cross-linking three residues on the same deck of the drum structure, its potassium binding ability could be enhanced. The simple reaction between N-hydroxysuccinimide esters (NHS esters) and primary amines was utilized. Two cross-linkers were designed, one of which is smaller and more flexible called TSAT (tris-(succinimidyl)aminotriacetate) and the other one is more rigid generated from the trimesic acid (Figure 2-7a). Three lysine analogue Dab residues were first introduced into the peptide on the same deck. DabdeSH, CL1DabdeSH and CL2DabdeSH were synthesized and their potassium binding abilities were determined in

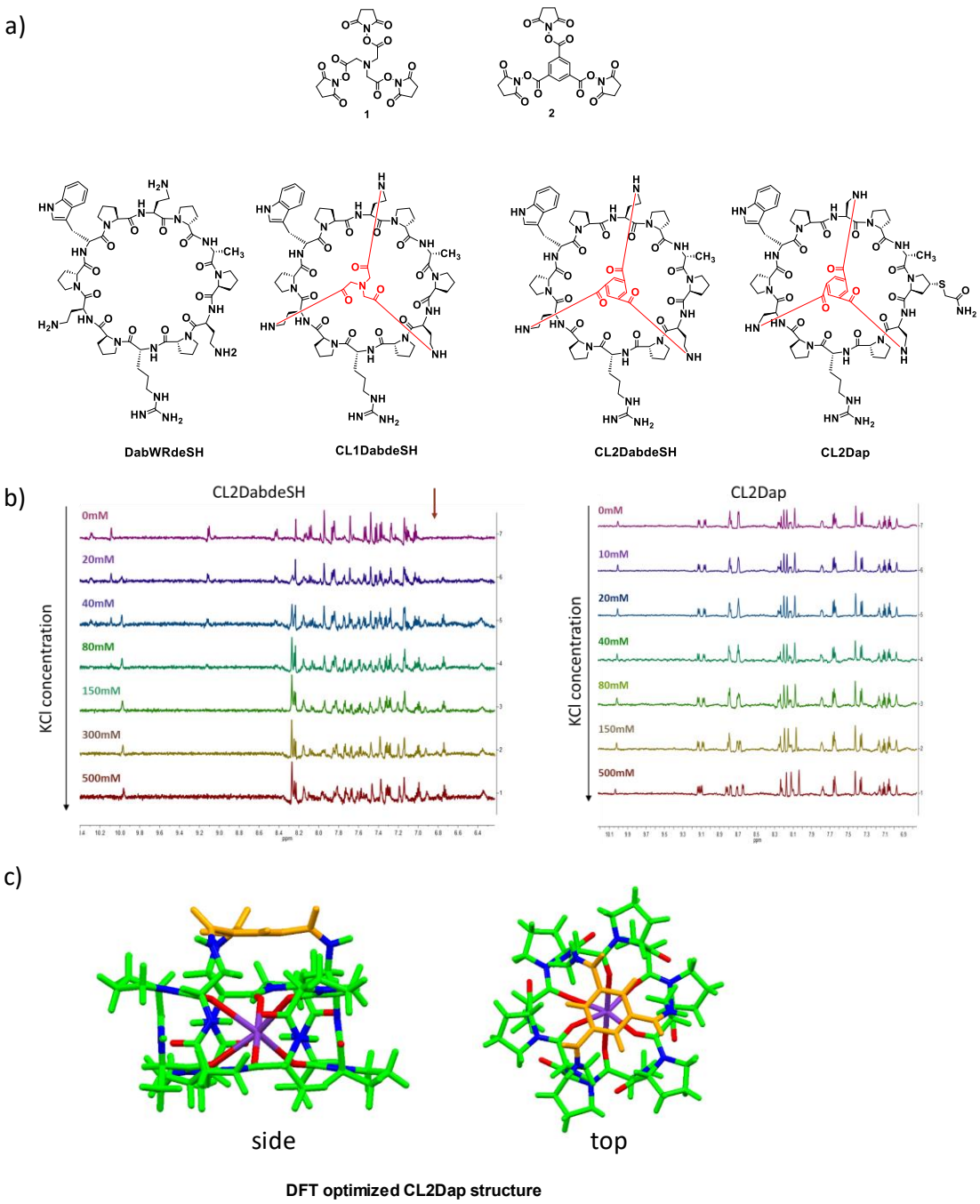


Figure 2-7. Cross-linking prolinomycin mutants. a) The structure of cross-linkers and cross linked prolinomycin mutants. b) K^+-nucleated folding of CL2DabdeSH (left) and CL2Dap (right) monitored by ^{1}H-NMR in 10% D_2O NaPi solution. c) DFT optimized 3D structure of CL2Dap.

the aqueous solution as well as the methanol: water 1:1 solution. The peptide without cross-linker has a K_d > 500mM in the pure aqueous solution. CL1DabdeSH has a 4-fold smaller

K_d value indicating the cross-linker 1 improved the potassium binding but the K_d value in the pure water solution is still high, which is around 150 mM. Encouragingly, in the pure aqueous solution CL2DabdeSH showed an impressive K_d value, 46 mM which is a 10-fold enhancement (Figure 2-7b). This rigid cross-linker seems to be able to stabilize the structure significantly.

To further improve this property, Dab was replaced by Dap, which has a shorter chain and could give a more rigid structure. Therefore, CL2Dap was synthesized and the potassium titration was done. It was surprising that ^{1}H-NMR spectrum of the peptide showed very well-defined peaks even in the absence of potassium ion (Figure 2-7b). The high concentration of K^+ did not change the ^{1}H-NMR spectrum of CL2Dap much, which indicates that CL2Dap has a strong tendency to adopt a well-defined conformation in pure aqueous solutions (Figure 2-7c).

From these data, we can draw the conclusion that the cross-linker indeed can improve the potassium binding ability of the prolinomycin scaffold by preorganizing the peptide for the drum structure, especially the rigid cross-linker. The prolinomycin mutant with Dap and trimesic acid cross-linker forms a rigid structure in the pure aqueous solution even in the absence of potassium ion.

2.3. Design and evaluate prolinomycin mutants as synthetic receptors

Encouraged by the unique structure features of prolinomycin mutants, we are motivated to endow the scaffold with interesting chemical functions. Given the size of the scaffold, we envision that it has the potential to be a synthetic receptor for small molecules. Towards this goal, we first investigate the iminoboronate chemistry as a powerful mechanism for ligand binding.

2.3.1. Covalent recognition of biological amines by iminoboronate chemistry

The specificity of molecular interactions is realized by the cooperative action of multiple noncovalent mechanisms, including hydrophobic packing, hydrogen bonding and electrostatic interactions. Recently, the use of reversible covalent chemistry has been recognized as an appealing and powerful strategy to target biomolecules (Figure 2-8a) as the reversibility minimizes the "off-target" effects caused by irreversible covalent bond formation. Specifically, we and others have introduced the iminoboronate chemistry as a powerful way to target various biological amines, including those of proteins, lipids and aminosugars[21, 22]. In comparison to typical imines, an iminoboronate enjoys greater

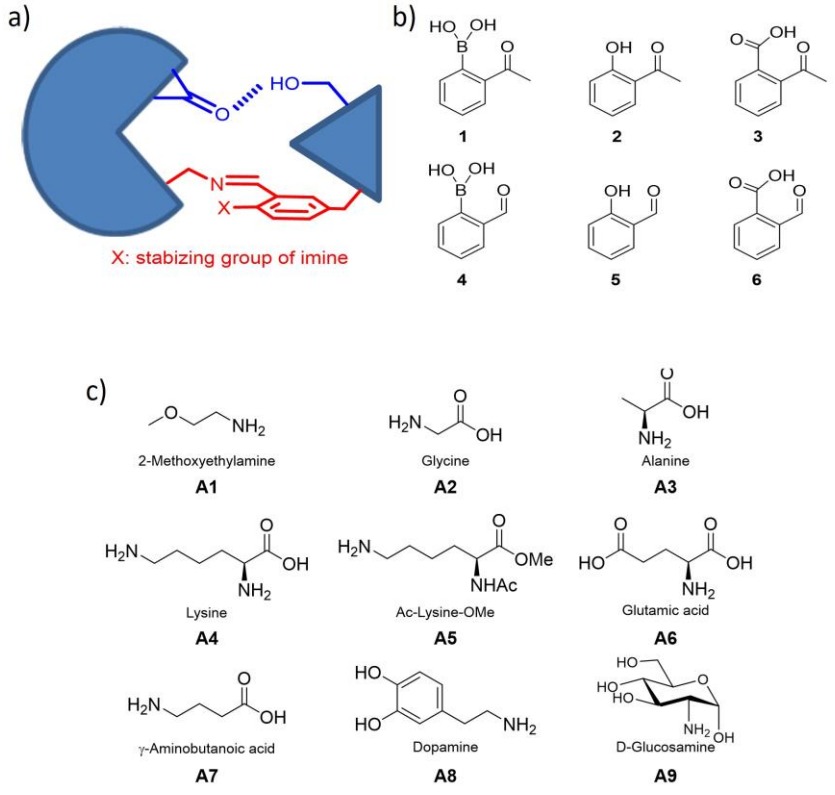


Figure 2-8. Imine formation as a mechanism for molecular recognition. a) Illustration of using reversible imine formation in conjugation with hydrogen bonding to target biomolecules. b) Derivatives of benzaldehyde and acetophenone designed to elicit imine formation. c) A selection of abundant amines in biology used for this study.

thermodynamic stability and potentially faster kinetics of formation[23]. We envision that the prolinomycin scaffold can display these reversible covalent warheads to target amines.

To better understand the iminoboronate chemistry, we first performed a comparative study of imine formation, in which (2-formylphenyl) boronic acid (2-FPBA) and (2-acetylphenyl) boronic acid (2-APBA) are compared to their analogues carrying a hydroxyl or a carboxylic acid group respectively (Figure 2-8b). A group of nine biological amines including representative amino acids, aminosugars and small molecule neurotransmitters (Figure 2-8c) were analyzed for conjugation to these derivatives of benzaldehyde and acetophenone.

Thermodynamic Analysis of the imine formation by UV-Vis and ^{1}H-NMR

The imine formation elicited significant changes in the UV-vis absorption, which provided a convenient way for monitoring the reactions. 2-APBA (**1**) and 2-FPBA (**4**) give an absorption maximum at 254 nm, while their conjugation product with glycine absorb at 280 nm. Salicylic acid (SA) (**5**) has two absorption peaks at 254 nm and 330 nm, both of which shift to 285 nm and 393 nm respectively when forming imine (Figure 2-9). By monitoring the absorption changes, we have assessed the thermodynamic equilibrium of various imine-forming pairs. We have further assessed these conjugation reactions with ^{1}H-NMR spectroscopy. The acetyl hydrogen of 2-APBA shifts from 2.6 ppm to 2.4 ppm when forming the iminoboronate conjugate. Similarly, the aldehyde hydrogen of 2-FPBA and SA also shift to the upfield (Figure 2-10). ^{1}H-NMR spectroscopy gives consistent results as the UV-vis based measurements.

Summarized in Table 2-2 are the dissociation constants of the reactants in various combinations. Consistent with the previous literature reports, both 2-APBA (**1**) and 2-FPBA (**4**) were found to conjugate with various amines at millimolar concentrations. In sharp contrast, no conjugation was observed for the 2-APBA analogues **2** and **3**, indicating the *ortho*-hydroxyl or carboxylic acid substituents cannot stabilize the imine product to the

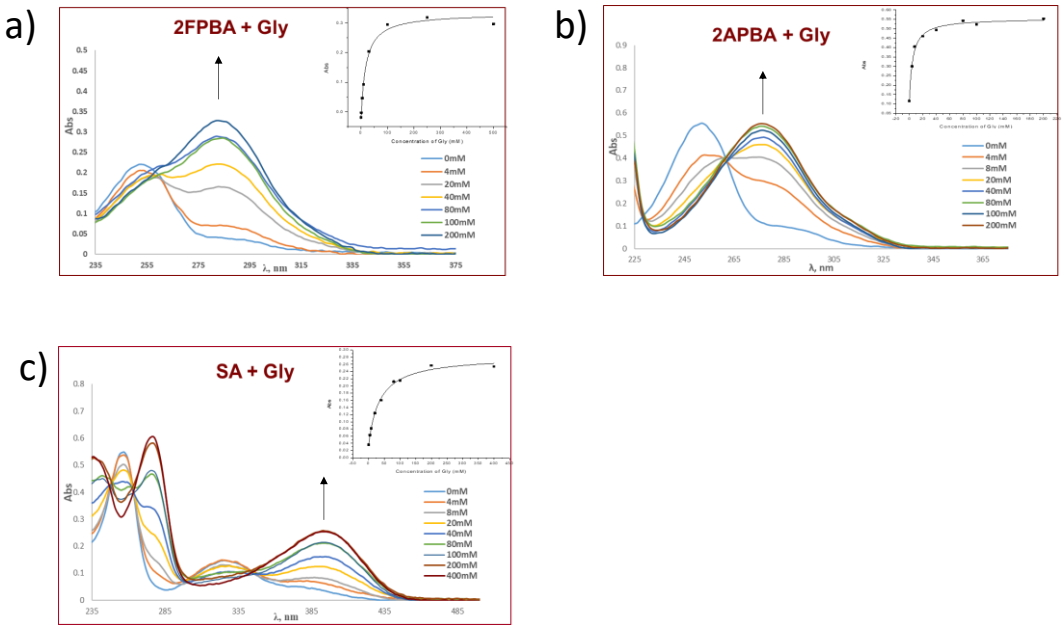


Figure 2-9. Typical UV-Vis spectrophotometric titrations of 50μM (a) 2FPBA, (b) 2APBA and (c) SA by Gly (0-200mM) at pH=7.

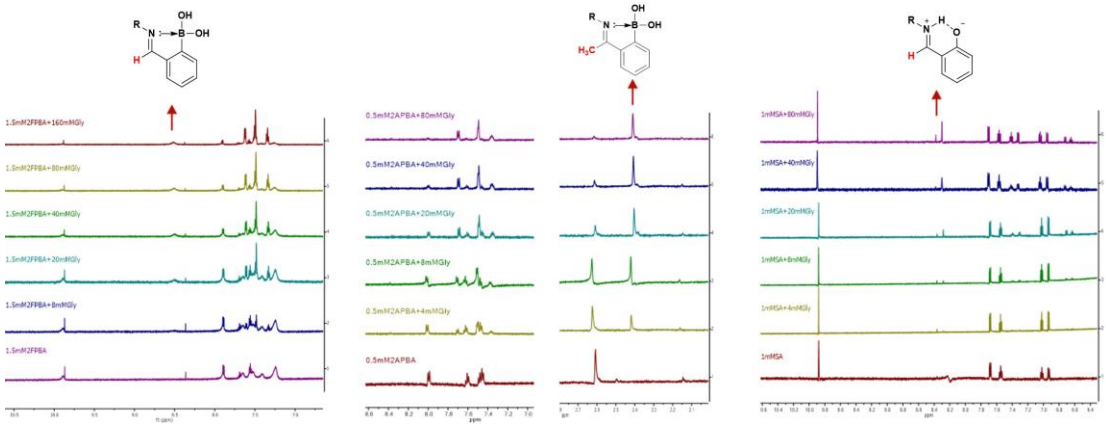
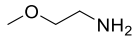
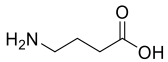
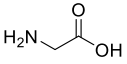
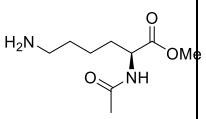
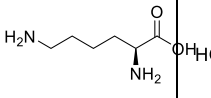
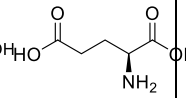
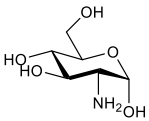
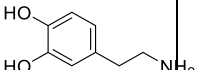
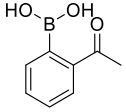
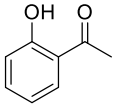
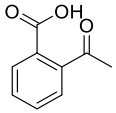
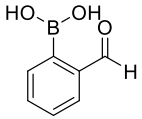
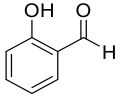
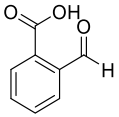


Figure 2-10. Typical ^{1}H-NMR spectra of (a)1.5 mM 2-FPBA, (b)0.5 mM 2-APBA, (c) 1 mM SA in 10% D2O at pH=7 with increasing amount of Gly (0-160 mM)

same extent as the boronic acid. Interestingly, 2-APBA showed diverse affinities toward various amine partners. For example, 2-APBA gave less than 10 mM K_d values with unhindered amines including 2-methoxyamine (A1), glycine (A2), lysine (A4), and γ-aminobutyric acid (A7). In contrast, the more sterically hindered amines showed diminished affinity to bind 2-APBA, with alanine (A3) yielding a K_d value of nearly 400 mM and free glutamic acid showing no measurable imine formation at all. The conjugation of lysine to 2-APBA presumably occurs on the side chain (ε-amine) as a main chain-protected derivative (A5) yielded a comparable K_d to that of the free lysine.

Similar to the ketones, 2-formylphenylboronic acid (2-FPBA, **4**) conjugates with the amines with millimolar K_d (1.7 to 125 mM), while the carboxylate derivative yielded no measurable conjugation. In contrast to compound **2**, the *ortho*-hydroxyl substituted benzaldehyde (SA, **5**) effectively elicited imine formation, yielding K_d values comparable to those of 2-FPBA. For example, the conjugation of SA with glycine gives a K_d of 36 mM, which is essentially the same as that of 2-FPBA (33 mM). SA appeared to conjugate with the remaining unhindered amines (A1, A5, and A7) with lower propensity than 2-FPBA by 2-4 folds. In contrast, SA appears to be more favorable for conjugation to the sterically hindered amines (A3, A6).

Table 2-2. Dissociation constants of the reactants in various combinations

	2-Methoxyethylamine 	γ-Aminobutyric acid 	Glycine 	Alanine 	Methyl acetyl-*L*-lysinate 	Lysine 	Glutamate 	D-(+)-Glucosamine 	Dopamine 
 1	6.6mM	26.4mM	5mM	396.6mM	6.0mM	8.4mM	N	39.6mM	11.4mM
 2	N	N	N	N	N	N	N	N	N
 3	N	N	N	N	N	N	N	N	N
 4	9.0mM	26.5mM	33.3mM	140mM	10.5mM	11.8mM	125mM	17.5mM	1.7mM
 5	28.9mM	117.1mM	36.2mM	91.1mM	24.2mM	15.9mM	69.0mM	20.9mM	15.7mM
 6	N	N	N	N	N	N	N	N	N

This is perhaps not surprising given the larger size of the boronic acid moiety. Dopamine (A9) gives the largest K_d difference between 2-FPBA and SA; this is actually due to the catechol moiety directly binding to the boronic acid to form boronate esters (Figure 2-11).

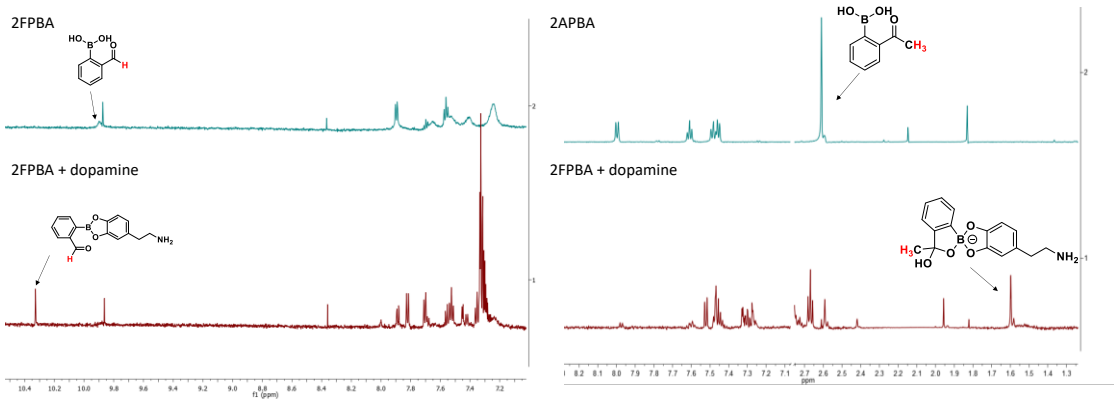


Figure 2-11. ^{1}H-NMR spectra of 2-FPBA and 2-APBA-dopamine conjugate.

Kinetic Analysis of the imine formation

The conjugation reactions of 2-APBA and 2-FPBA reached equilibrium instantaneously (< 10s). In contrast, SA binds to amines much slower. For instance, in the titration with 20mM Glu, within twenty minutes peaks at 282 nm and 393 nm are continuously increasing

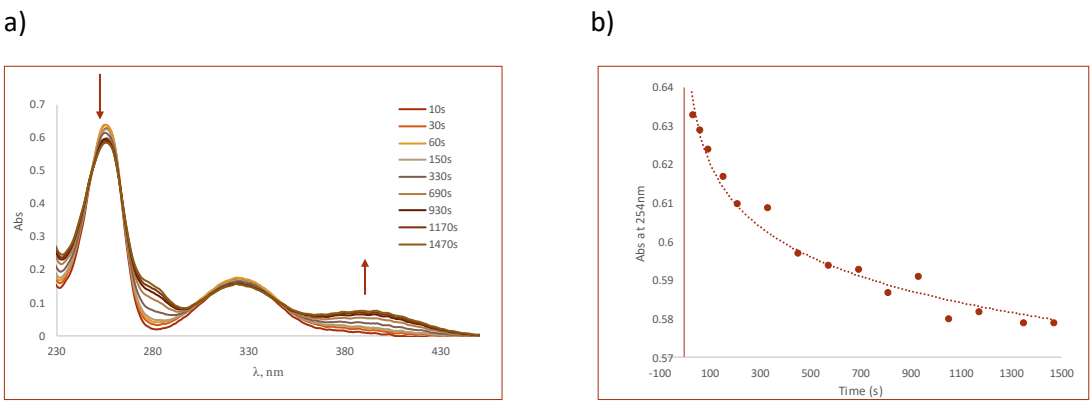


Figure 2-12. Kinetic study of interaction between SA and Glu. a) UV-vis spectrum of 50 µM SA and 20 mM Glu in 20 min; b) Decrease of Abs versus time.

with the decreasing of peaks at 254 nm and 330 nm. It shows a rate constant about 8×10^{-3} $M^{-1}s^{-1}$. Mechanistically, the boronic acid moiety facilitates the nucleophilic attack of the carbonyl by stabilizing the tetrahedral intermediate via boroxole formation (Figure 2-12).

Conclusions

This part documents our detailed comparative study of the strategies to elicit imine formation as a way to target amine-presenting biomolecules. The results unequivocally demonstrate the advantages of having an *ortho*-boronic acid moiety in promoting formation of imines. Specifically, the boronic acid group improves on the thermodynamic stability of imines by forging a B-N dative bond giving rise to iminoboronates. Installing an *ortho*-boronic acid to benzaldehyde and acetophenone enables facile imine (iminoboronate) formation with a variety of biological amines at millimolar concentrations. In contrast, having an *ortho*-carboxylic acid group afforded no conjugation under the same conditions. Interestingly, incorporating an *ortho*-hydroxyl group allowed imine formation of benzaldehyde, but not the ketone analogue. In addition to the improved thermodynamic stability, formation of an iminoboronate occurs instantaneously at neutral conditions. The kinetic acceleration presents additional advantage of the iminoboronate chemistry over the hydrogen bond stabilized imines. Finally, our results reveal the dynamic interplay of iminoboronate formation and boronate ester formation by 2-APBA and 2-FPBA. Strategic utilization of these reversible covalent reactions may greatly empower chemists to target biomolecules. In the following study, we will incorporate FPBA and APBA warheads into the peptides for target binding.

2.3.2. Prolinomycin mutant targeting glutamic acid

Glutamate is a major excitatory neurotransmitter in the central nervous system, whose misregulation can cause neuronal dysfunction and degeneration[24]. Synthetic receptors that can recognize the glutamate selectively under the physiological condition could have a therapeutic potential. WW-Pro1 was first designed and synthesized to target glutamate molecule (Figure 2-13a). FPBA moiety was installed onto the scaffold as a tether, which had intrinsic affinity and fast kinetics to bind the amine group of glutamate due to the iminoboronate formation (Figure 2-13b). Two arginine analogues were also incorporated on the same deck to enhance the affinity by interacting with the carboxyl groups of the glutamate.

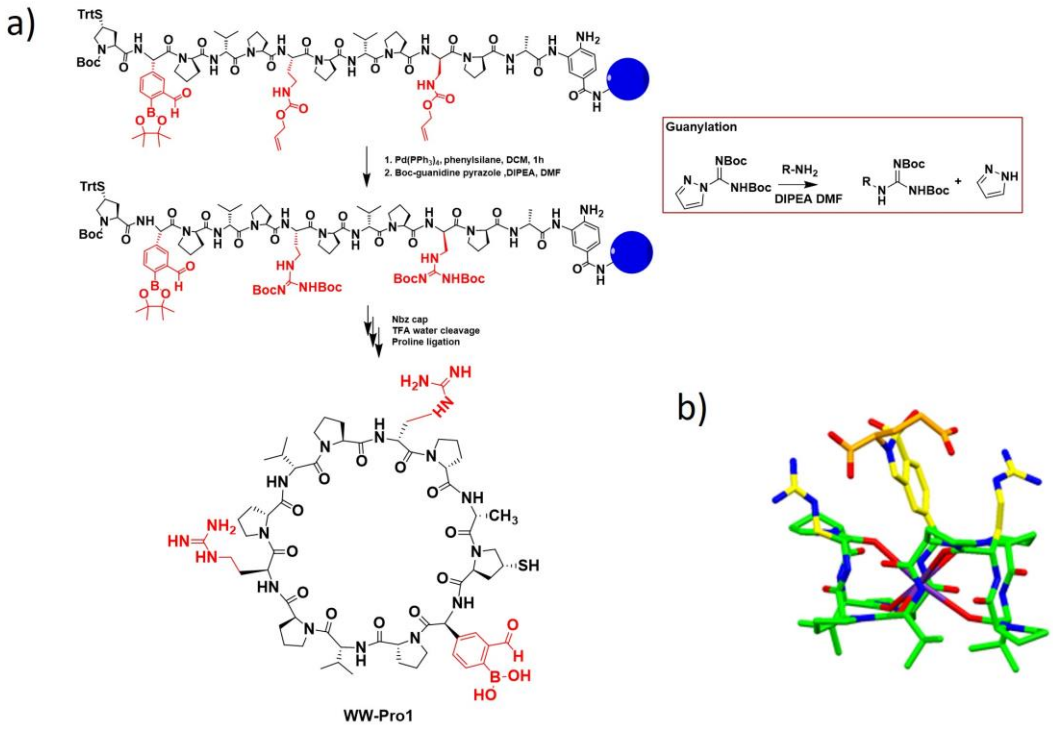


Figure 2-13. Synthetic glutamic acid receptor WW-Pro1. a) The synthesis and structure of WW-Pro1. b) Glutamic acid-receptor conjugate structure optimized by DFT.

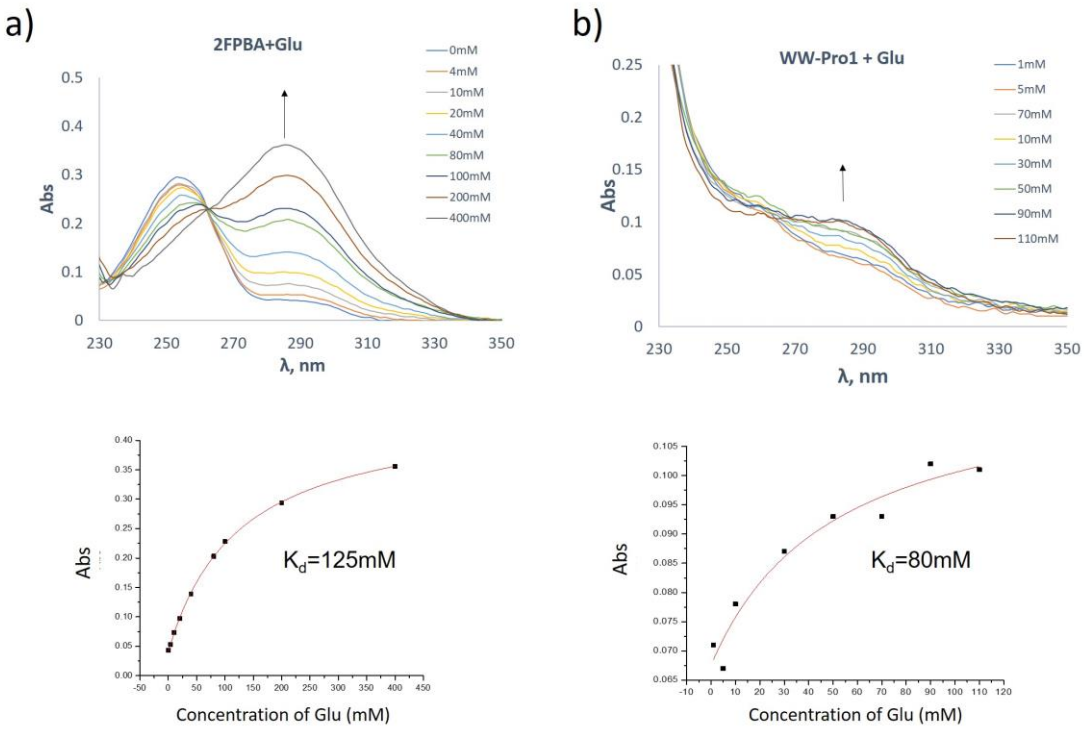


Figure 2-14. UV-Vis titration of Glu against 2FPBA (a) and WW-Pro1 (b).

The affinity was estimated by UV-vis spectrophotometric titrations in the neutral aqueous solution. Surprisingly, no FPBA characteristic absorbance peak showed up around 250 nm in the UV-Vis spectrum of WW-Pro1, suggesting that the formation of the boroxole. When adding glutamate and 2-methoxyethylamine (MEA), a positive control, the absorbance at 280 nm increased indicating the formation of the iminoboronate. K_d values were determined from the binding curve. Compared to the FPBA alone (K_d = 125mM), WW-Pro1 showed a little improvement on the affinity to the glutamate (K_d = 80mM) (Figure 2-14).

The undesired affinity can be attributed to the boroxole formation which may lower the affinity of FPBA moiety towards glutamate. On the other hand, the benefit of two arginine analogues was not clear in these experiments. One concern is that the chirality of one

arginine analogue was changed from L to D, which was found to make the peptide lose the ability to chelate potassium ion therefore it couldn't fold into the desired receptor structure.

2.3.3. Prolinomycin mutant targeting amino acids

Taking lessons from WW-Pro1, we further designed WW-Pro2 which contains two L-Arg residues and one APBA warhead on the same deck (Figure 2-15a). The potassium binding ability was confirmed by ^{1}H-NMR KCl titration. In the 50% CD_3OH NaPi buffer,

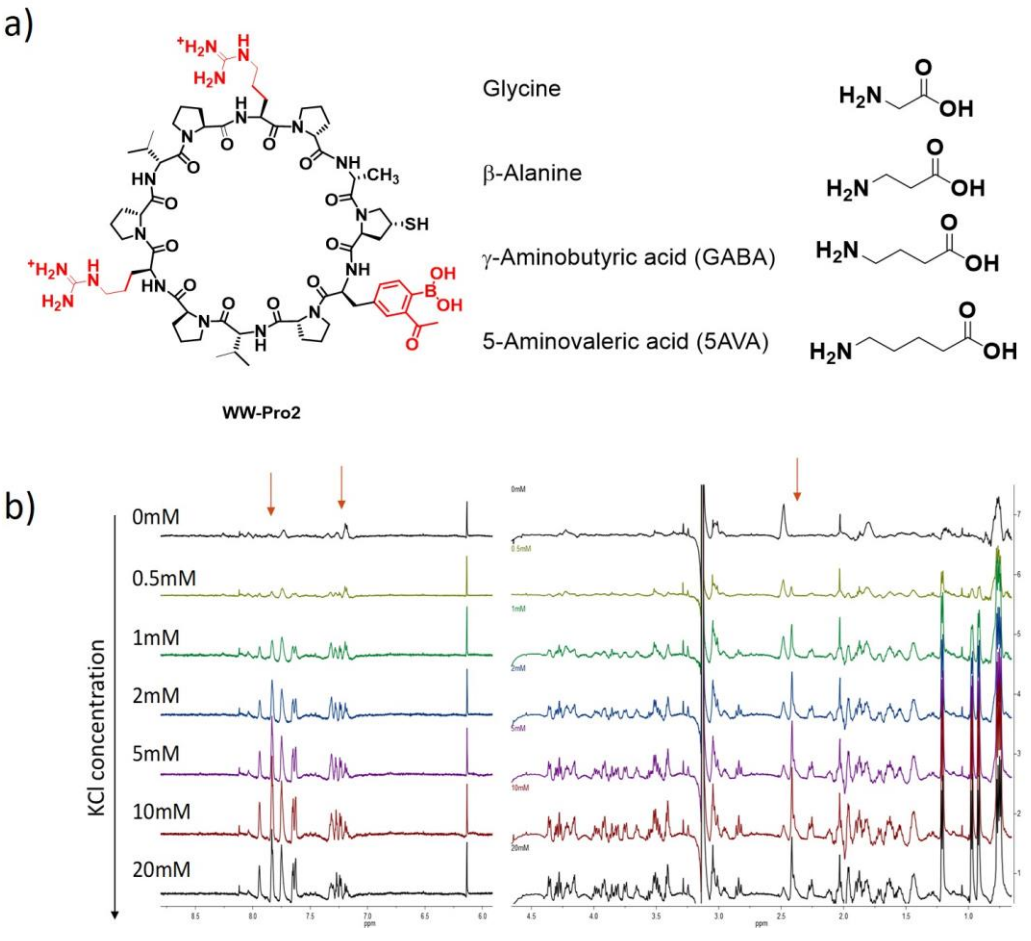


Figure 2-15. WW-Pro2 targeting amino acids. a) The structure of WW-Pro2 and amino acids ligands. b) ^{1}H NMR spectra of WW-Pro2 with increasing concentration of potassium ion.

the addition of KCl led to the appearance of sharp peaks in the amide region and a sharp peak at 2.4 ppm which correspond to the acetyl hydrogens of APBA motif (Figure 2-15b).

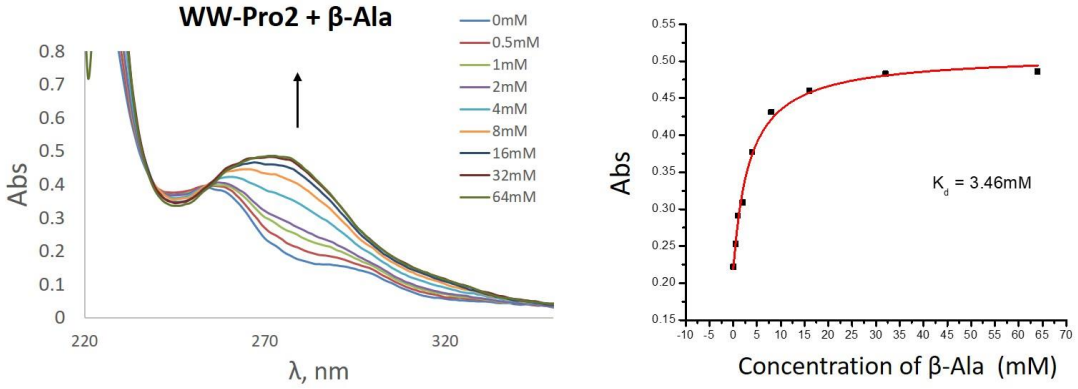
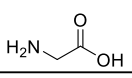
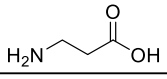
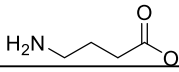
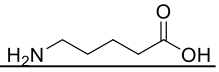


Figure 2-16. UV-Vis titration of β-Ala against WW-Pro2 and K_d value.

Table 2-3. Dissociation constants of WW-Pro2 against various amino acids (mM)

	Gly	β-Ala	GABA	5AVA
				
2APBA	1.38	3.35	2.62	3.38
WW-Pro2	1.29	3.46	2.71	3.54

To answer the question whether charge-charge interactions between guanidine groups of the Arg residues and carboxylic acid groups of the ligands can enhance binding, we tested WW-Pro2's binding affinity against a series of amino acids including glycine, β-alanine, γ-aminobutyric acid (GABA) and 5-aminovaleric acid (5AVA). The titrations were done in the solution containing 20 mM KCl, 50% methanol and 50% PBS buffer (pH 7). As amines were added, the absorbance around 280 nm increased corresponding to the formation of iminoboronates (Figure 2-16). By plotting the absorbance at 280 nm versus the concentrations of amines, K_d values can be evaluated. As shown in Table 2-3, WW-

Pro3 has the similar K_d to APBA, which indicates that charge-charge interactions are not significant on the prolinomycin scaffold. This is perhaps due to the water exposed environment of the prolinomycin surface which diminishes the effect of weak noncovalent interactions such as electrostatic interactions. Therefore, we next explored the effect of hydrophobic interactions which can provide a stronger driving force for ligand binding.

2.3.4. Prolinomycin mutant targeting amines with hydrophobic groups

To take advantage of hydrophobic interactions, a homophenylalamine residue and an alanine residue were incorporated into WW-Pro3 together with the APBA warhead (Figure 2-17). It is proposed that through hydrophobic interaction homophenylalanine residue can enhance the affinity of APBA against the amines with hydrophobic groups.

The conformation of WW-Pro3 was first confirmed with KCl titration monitored by ^{1}H NMR (Figure 2-18a). We selected cyclohexanemethylamine, benzylethylamine, and 3-phenyl-1-propylamine as hydrophobic ligands of WW-Pro3. 2-APBA served as a control. The titrations were done in the solution containing 200 mM KCl, 50% methanol and 50%

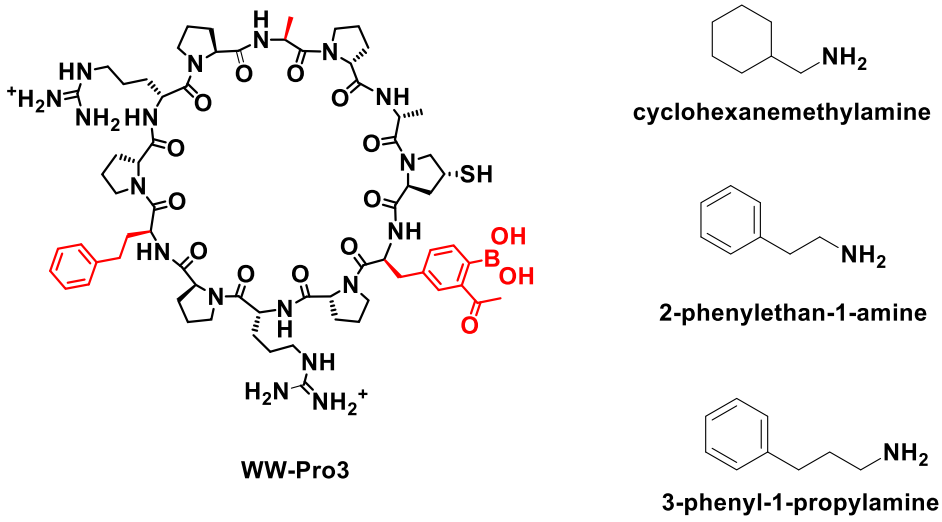


Figure 2-17. The structure of WW-Pro3 and hydrophobic amine ligands.

PBS buffer (pH 7). As amines were added, the absorbance around 275 nm increased corresponding to the formation of iminoboronates (Figure 2-18b). By plotting the absorbance at 275 nm versus the concentrations of the amines, K_d values can be evaluated

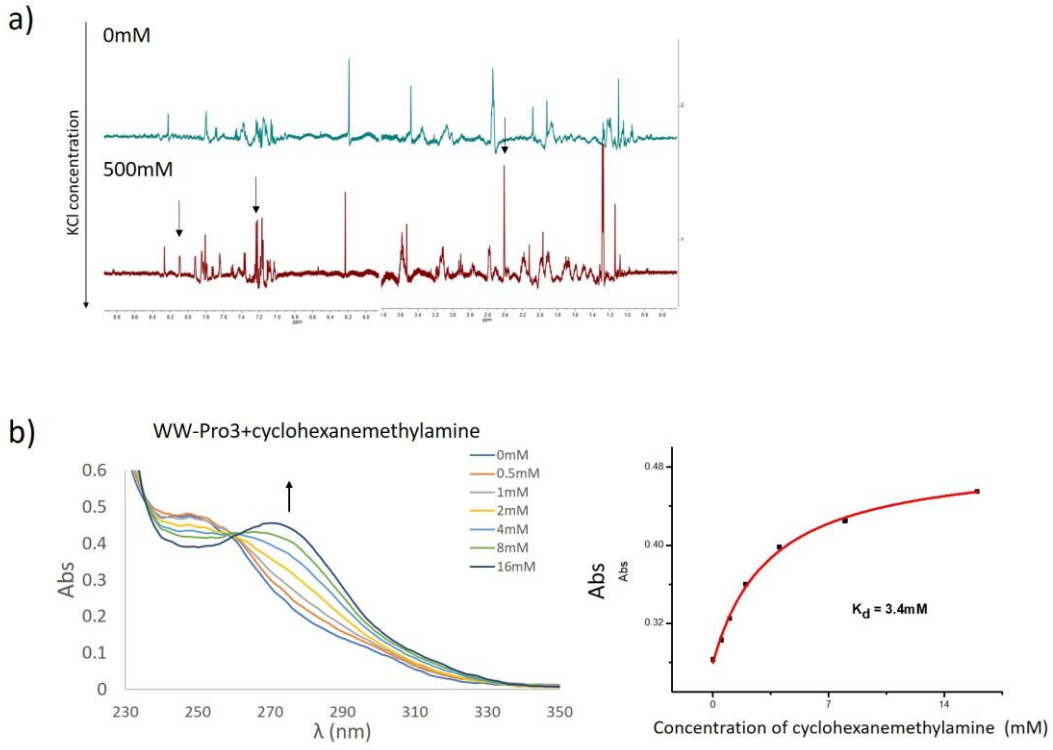
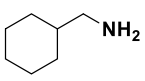
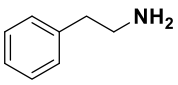
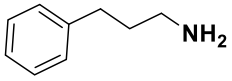


Figure 2-18. Characterization of WW-Pro3. a) ^{1}H NMR spectra of WW-Pro3 with increasing concentration of potassium ion. b) UV-Vis titration of cyclohexanemethylamine against WW-Pro3 and K_d value.

Table 2-4. Dissociation constants of WW-Pro3 against various amines (mM)

	Cyclohexanemethylamine 	**Benzylethylamine** 	**3-phenyl-1-propylamine** 
2APBA	2.9	2	2
WW-Pro3	3.4	2.4	1.6

(Table 2-4). The K_d values were around 2-3 mM, however, there was no significant difference between WW-Pro3 and 2APBA.

The results indicate that the effect of homophenylalanine residue is minimal. The hydrophobic amines tested probably are too small to interact with APBA and homophenylalanine residue simultaneously. Also, one hydrophobic residue on the prolinomycin may not provide enough binding affinity.

2.3.5. Prolinomycin mutant targeting lysine residues of short peptides

Finally, we redesigned a prolinomycin mutant, WW-Pro4, to site-specifically target lysine residues (Figure 2-19a). Besides APBA warhead, WW-Pro4 consists two pentafluorobenzene moieties placed on the same face of the scaffold to target the aromatic amino acid residues around the lysine via π-π stacking (Figure 2-19b). We believed that two hydrophobic residues would enhance the binding better than one residue did.

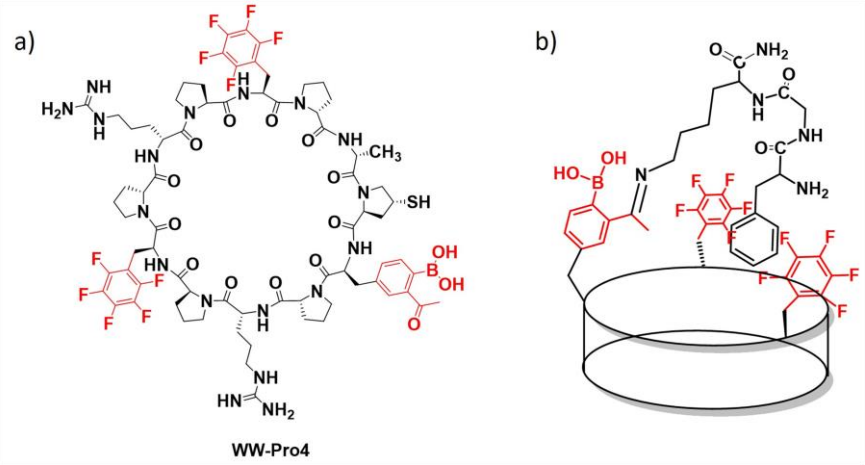
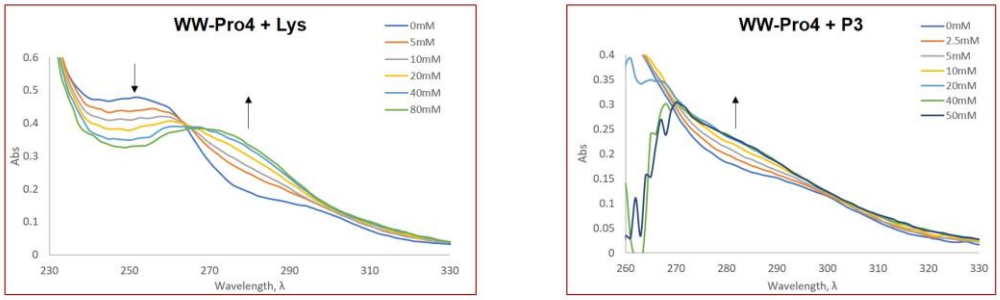


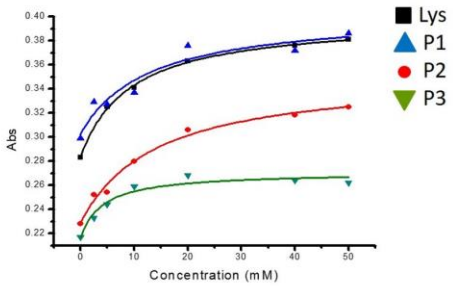
Figure 2-19. a) Structure of WW-Pro4. b) Proposed interactions between WW-Pro4 and lysine containing peptide.

We also synthesized a series of short peptides containing lysine and phenylalanine including Phe-Lys (P1), Phe-Pro-Lys (P2) and Phe-Gly-Lys (P3). By doing so, we can vary the length and conformation of the ligand peptides. The affinity, K_d values, were evaluated by UV-spectrum titrations. The results suggested that WW-Pro4 had the similar affinities to Lys, P1 and P2. Interestingly, WW-Pro4 showed a preference for P3 whose K_d, 4.1mM, is 2-fold smaller than that of Lys and 3-fold smaller than those of P1 and P2 (Figure 2-20). The Gly residue probably provides the suitable flexibility promoting the interaction between P3 and WW-Pro4. We can conclude that the cooperative action of all side chains

a)



b)



Targets	K_d (mM)
Lys	8.2
Phe-Lys (P1)	14.0
Phe-Pro-Lys (P2)	12.6
Phe-Gly-Lys (P3)	4.1

Figure 2-20. Binding between WW-Pro4 and short peptides. a) UV-Vis titration of lysine containing peptides against WW-Pro4. b) Binding curves and K_d values between WW-Pro4 and ligands.

allows WW-Pro4 to show specific high affinity for the lysine residue in a certain short peptide.

2.4. Dimerization of prolinomycin mutants

The unique structure of prolinomycin makes it possible to display three copies of APBA moieties and lysine side chains on the same deck. We hypothesized that this pair of prolinomycin mutants would have high affinity to each other and form a stable dimer. Based on this specific and reversible dimerization, a fluorescence sensing system can be developed to detect important biomolecules containing amino group.

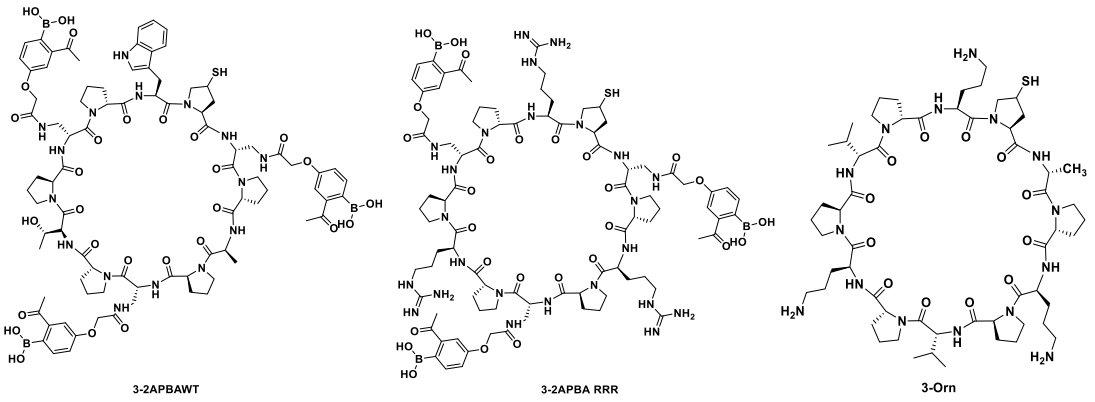


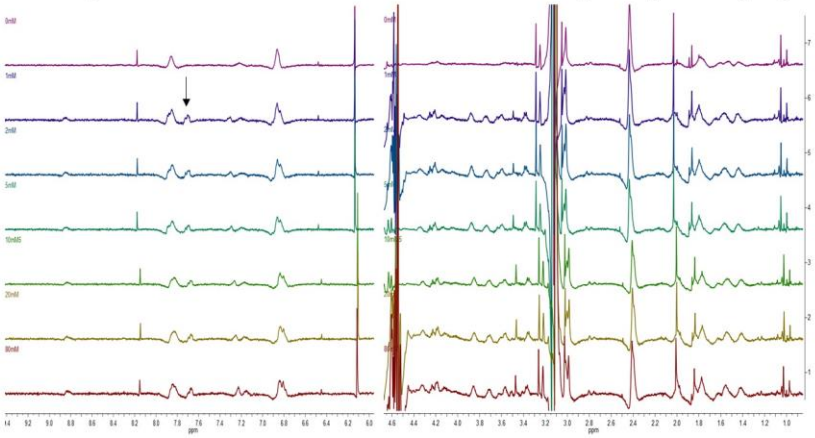
Figure 2-21. The structure of 3-2APBAWT, 3-2APBA RRR and 3-Orn.

First, two prolinomycin mutants with three 2APBA moieties was made (Figure 2-21), one having tryptophan, threonine and alanine on the other side (3-2APBAWT) and one having three arginines (3-2APBA-RRR). The prolinomycin mutant with triple Orn residues was also made.

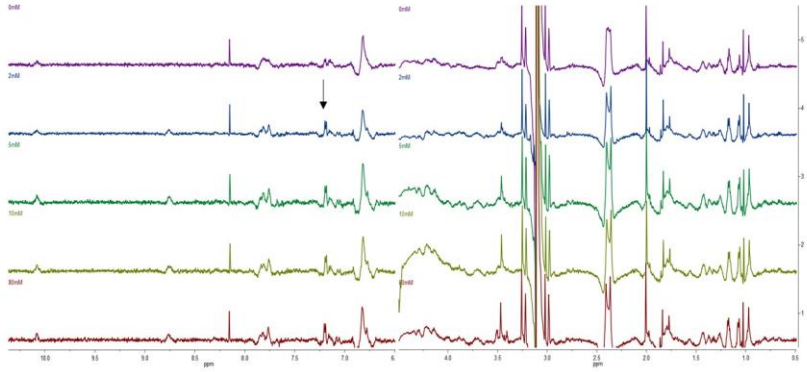
^{1}H-NMR studies were done to characterize the structures of 3-2APBA mutants. First KCl titration was done in in the 50% CD_3OH Napi buffer pH4 (Figure 2-22a and b). Without potassium, there were two broad peaks in the aromatic region which might be the aromatic hydrogens of 2APBA moieties and a relatively sharp peak at 2.4ppm which

should be the acetyl hydrogen of 2APBA moieties. With KCl added, several broad peaks appeared both in the aromatic region and aliphatic region. Even though no sharp peak appeared, the appearance of new peaks indicates that the peptide can respond to the KCl.

(a) ^{1}H-NMR spectra of KCl titration of 3-2APBA RRR in the 50% CD_3OH Napi buffer (pH4)



(b) ^{1}H-NMR spectra of KCl titration of 3-2APBA WT in the 50% CD_3OH Napi buffer (pH4)



(c) ^{1}H-NMR spectra of KCl titration of 3Orn in the 50% CD_3OH Napi buffer (pH4)

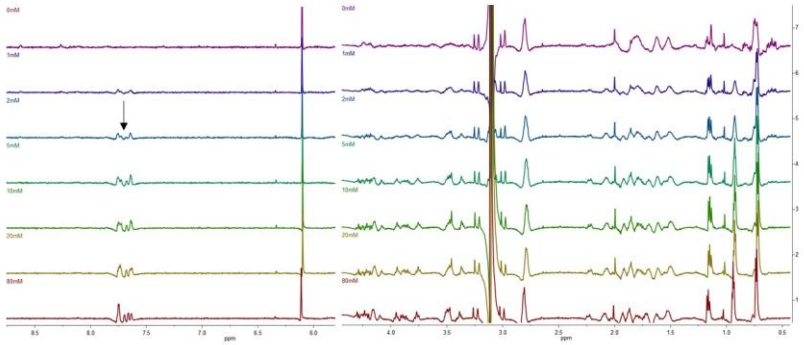


Figure 2-22. ^{1}H-NMR analysis of KCl titratin against. (a) 3-2APBA RRR, (b) 3-2APBA WT, and (c) 3Orn.

The same KCl titration was done in the 3Orn mutant (Figure 2-22c). Several sharp amide and aliphatic peaks appeared which suggests that with potassium 3Orn mutant folds into a well-defined structure.

3Orn was titrated into 3-2APBA RRR solution at pH 7 with 80 mM KCl to determine the affinity between two peptides. As 3Orn was added, the broad aromatic peaks and the acetyl peak around 2.4 ppm decreased, which was due to the precipitation of 3-2APBA RRR in the presence of high concentration of 3Orn. Several small peaks appeared in the aromatic region and a small sharp appeared at 2.5 ppm, which might indicate the formation of iminoboronates (Figure 2-23). However, the peaks were too small to be interpreted.

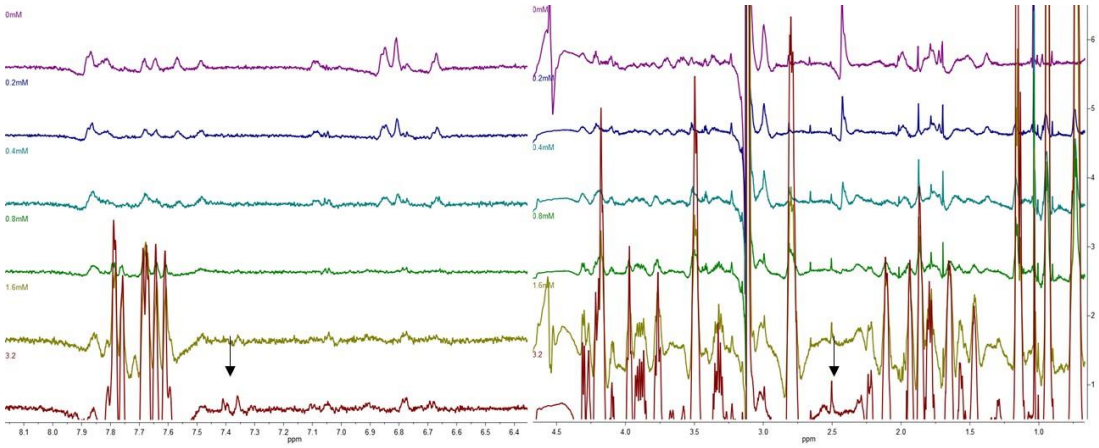


Figure 2-23. ^{1}H-NMR spectra of 3Orn titration in 3-2APBA RRR in the 50% CD_3OH Napi buffer (pH7 with 80mM KCl)

To further determine the binding between 3-2APBA mutants and 3Orn, the UV-Vis titration was done (Figure 2-24). Lysine was titrated as a control. The peptide showed a peak around 280 nm and as Lys was added the absorbance decreased. When 3Orn was

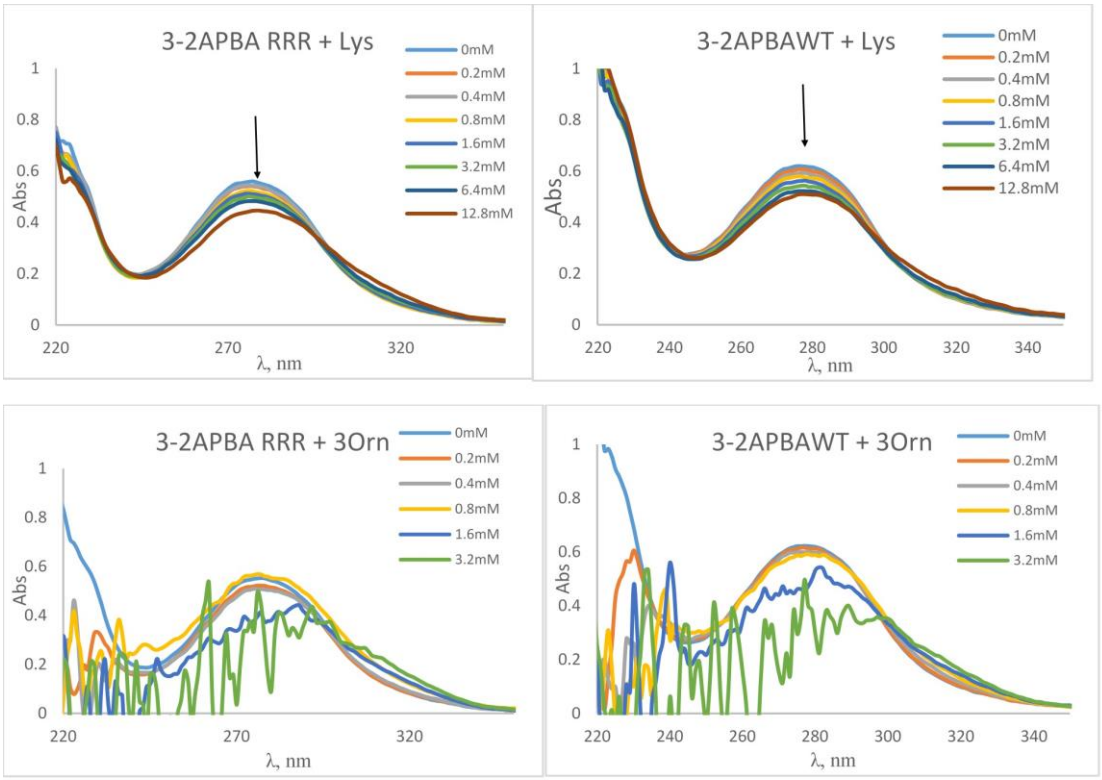


Figure 2-24. UV-Vis spectrophotometric titrations of 3-APBA mutants with lysine and 3Orn.

added the absorbance at 280 nm also decreased a little bit, however, as 3Orn itself had a large absorbance around 280 nm, the absorbance reading at high concentration of 3Orn couldn't be obtained accurately. It could be hard to monitor the affinity by UV-Vis spectroscopy.

To make sure the peptide can bind potassium and fold into the organized structure, we made 3-2APBA-Orn which has an Orn linker between 2APBA moiety and the backbone (Figure 2-25a). We could see sharp peaks on the ^{1}H-NMR spectra when potassium was titrated into the peptide solution indicating the peptide could fold into the well-defined structure (Figure 2-25b).

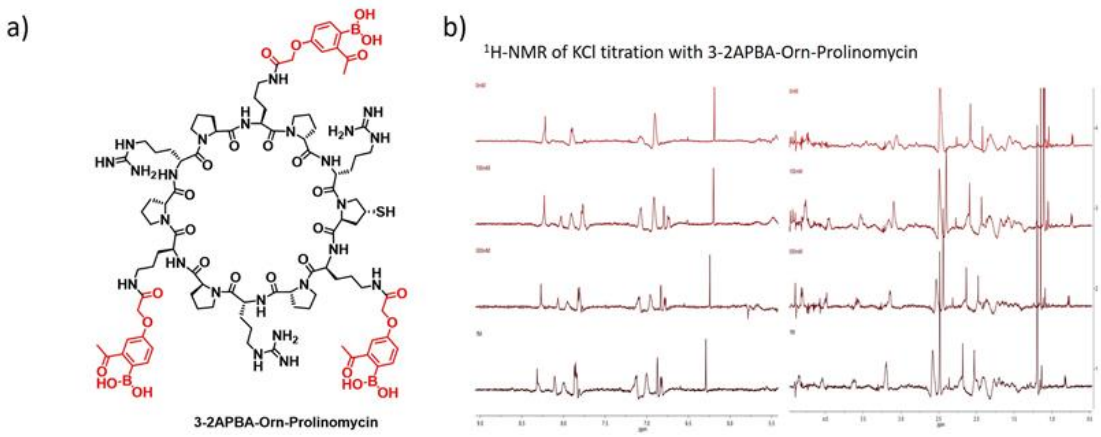


Figure 2-25. 3-2APBA-Orn-Prolinomycin. a) structure, b) KCl titration monitored by ^{1}H NMR spectra.

Due to the difficulties of UV-Vis and NMR titrations, we decided to measure the binding by monitoring the fluorescence by labeling 3-2APBA-Orn by Alexa 488 and 3Orn by DABCYL which is a quencher of Alexa 488 (Figure 2-26a). Therefore, the fluorescence of Alexa 488 would decrease when the peptides bind to each other. From the results we could see that when dabcyl labeled 3Orn was titrated into Alexa 488 labeled 3-2APBA-Orn solution, the fluorescence at 530 nm decreased. However, adding the Dabcyl acid alone also resulted in the decrease of the fluorescence in the similar manner (Figure 2-26b). Therefore, no significant binding was observed until 60 µM 3Orn was added which suggested the K_d must be larger than 60 µM.

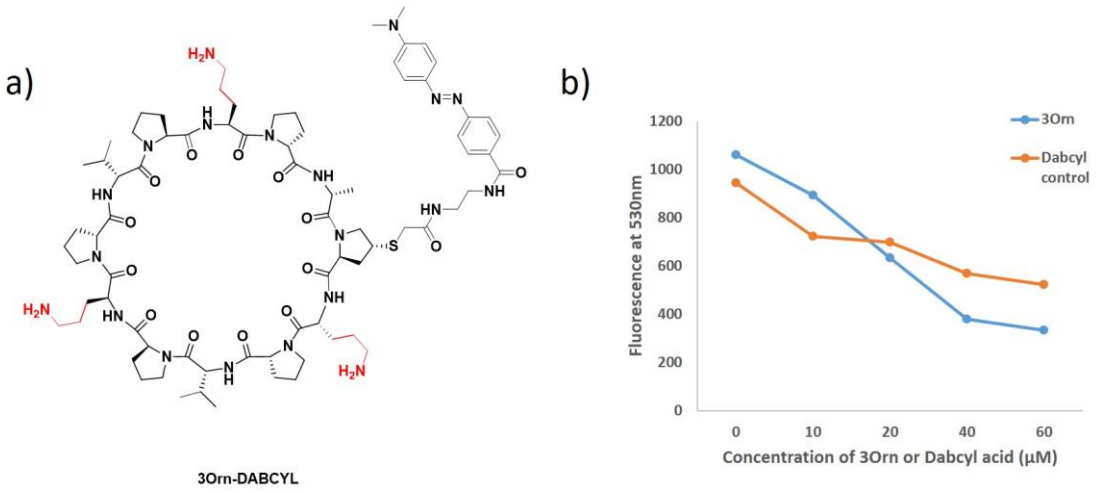


Figure 2-26. Measurement of the binding between 3Orn and 3APBA-Orn-prolinomycin monitored by fluorescence. a) structure of 3Orn-DABCYL, b) Fluorescence measurement of 3-2APBA-Orn-Alexa titrated with 3Orn-DABCYL and Dabcyl control.

The binding between 3-2APBA mutants and 3Orn was also investigated by a competition assay. First, the affinity of 3-2APBA prolinomycin mutants against BSA were determined by fluorescence anisotropy (Figure 2-27a and b). The K_d of 3-2APBA-RRR was 0.42 mg/mL and that of 3-2APBA-Orn was a little bit higher, 1.5 mg/mL. 3Orn was added into the mixture of 200bnM 3-2APBA-RRR and 1mg/mL BSA. As the 3Orn concentration increased, the anisotropy decreased, which indicated 2APBA trimer begun to bind 3Orn instead of BSA (Figure 2-27c). The IC_{50} was determined as 380 μM. In contrast, when 3Orn was added into the mixture of 200 nM 3-2APBA-Orn and 5mg/mL BSA, the anisotropy didn't change indicating the 3Orn didn't bind to 3-2APBA-Orn. The

long Orn linker between 2APBA moiety and the back bone might be too flexible and reduced the affinity comparing with the 3-2APBA-RRR with Dap linker.

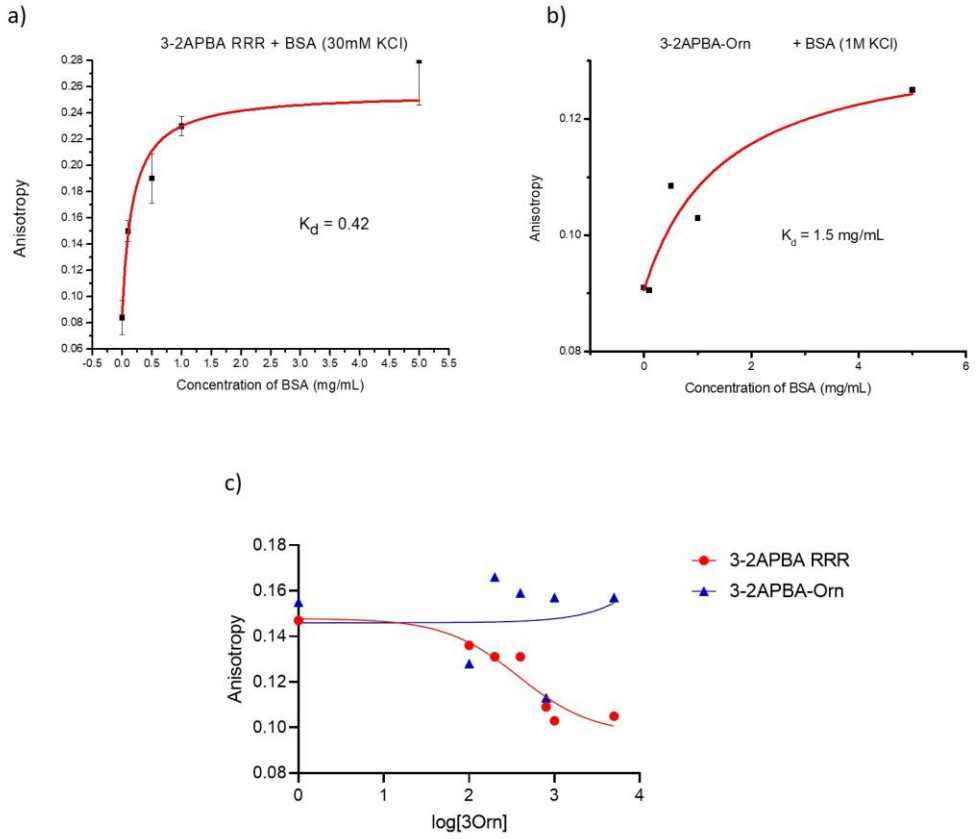


Figure 2-27. Fluorescent competition assay to determine the binding between 3-2APBA mutants and 3Orn. Binding curve of 3-2APBA RRR (a) and 3-2APBA-Orn (b) against BSA monitored by fluorescence anisotropy. (c) Inhibition of the binding between 3-2APBA mutants and BSA with 3Orn.

2.5. Conclusions

With the development of solid phase synthesis and organic chemistry, several ways to modulate peptide conformation and flexibility have been developed, including (1) combination of L and D residues; (2) introducing constrained residues like proline; and (3) cyclization of the peptide. The peptide possessing a well-defined conformation provides a versatile platform for designing functional molecules, such as antimicrobials, protein-protein inhibitors and catalysts. Prolinomycin, a cyclic peptide that folds into a drum like

structure upon potassium binding, was studied in this chapter. We envision that the prolinomycin scaffold can display a variety of key functional side chains to bind target molecules.

First, the potassium-nucleated folding behavior of prolinomycin was thoroughly investigated. It was found that the scaffold tolerates various mutations at non-proline positions. Both natural and unnatural amino acids have been incorporated onto the scaffold without affecting the potassium induced folding. However, in the case of WW-Pro4 and 3-2APBA mutants, the well-defined conformations are not obvious because some ^{1}H NMR peaks were still broad with high concentration of potassium. Perhaps the aromatic residues on the same face of the scaffold can interact with each other which prohibits the folding. Therefore, increasing the distance between side chains and backbone can avoid this interaction and resume potassium binding of the scaffold, as shown in the case of 3-2APBA-Orn mutant.

The proline residues are crucial for the conformation. We found that (2S, 4R) 4-thioproline can stabilize the structure, however, the hydroxyproline can't. Especially, the incorporation of cis-D-Hyp destroyed the conformation completely. We hypothesize that the hydroxyl group might form H-bonds with backbone carbonyl groups and inhibit the chelation of potassium ion.

A strategy to enhance the folding propensity of the peptide has been developed by introducing cross-linkers. The side chains of three residues on the same face of the scaffold can be cross-linked efficiently. The K_d value of potassium binding was decreased to 46 mM from hundreds millimolar, more than 10-fold improvement. Furthermore, a prolinomycin mutant modified with short side chains and the rigid cross-linker, CL2Dap,

showed very rigid structure in pure aqueous solution even without potassium ion. This is very encouraging and indicates that we can use this crosslinking strategy to design other rigid peptide scaffolds, such as the prolinomycin scaffold with larger ring size.

Based on the understanding of the prolinomycin structure, the potential of the peptide to be a synthetic receptor was explored. We believed that covalent bonds and non-covalent interactions can happen simultaneously between the prolinomycin mutant and its target. We introduced the APBA warhead to target amine groups via reversible covalent iminoboronate formation. In addition, charge-charge interactions and hydrophobic interactions were also studied, however, the benefit of them is minimal. This might be attribute to the water exposed environment around the prolinomycin side chains and mismatched conformation of the target molecules and receptors. By incorporating a pair of pentafluorobenzene moieties, we proposed that through π-π stacking the prolinomycin mutant can recognize a lysine residue around an aromatic residue of short peptides. We are delight to see that this cooperative action of side chains on the same deck allowed WW-Pro4 to bind a special lysine residue (Phe-Gly-Lys) with higher affinity than others. It is also intriguing to see if a prolinomycin mutant can site-specifically recognize the lysine residue in a larger environment where more interactions could happen between the ligand and prolinomycin scaffold.

Lastly, we investigated the dimerization of a pair of prolinomycin mutants. We developed a fluorescence-based method to determine the binding affinity between 3-2APBA and 3Orn mutants. The multivalent display of 2-APBA on the same face of the scaffold decreased the K_d from about 10 mM to submillimolar. Even though the binding is still relatively week, it can be optimized by introducing other stronger interactions.

Overall, we learnt the principles of tfolding behavior of prolinomycin and explored the interactions that can happen simultaneously on the same face of the scaffold. These studies pave the way to the rational design of functional molecules based on peptide scaffolds.

2.6. Experimental Procedures

2.6.1 General Methods

Dawson Dbz resin, Fmoc-Osu and HBTU were purchased from Novabiochem. Fmoc-protected amino acids were purchased from Advanced Chemtech or Chem-Impex Int'l Inc. Dry dioxane, dry dichloromethane, diethanolamine and methoxyethylamine were purchased from Acros Organics. Bispinacolato diboron (B2pin2) was purchased from Matrix Scientific. Trifluoroacetic acid was purchased from Protein Technologies. Other chemicals were obtained from Fisher Scientific or Sigma Aldrich. Peptide synthesis was carried out on a Tribute peptide synthesizer (Protein Technologies). ^{1}H NMR measurements were performed on a Varian INOVA 500 and 600 MHz spectrometer. A Nanodrop UV-vis spectrometer was used to measure the concentration of the fluorescein labelled compound. Mass-spec data were generated by an Agilent 6230 LC TOF mass spectrometer. Peptide synthesis was carried out on a Tribute peptide synthesizer from Protein Technologies. HPLC purification was done on Waters 2695 Alliance HPLC system.

2.6.2 UV-Vis Titration Analysis

50 μM of each derivatives of benzaldehyde and acetophenone in 1X PBS pH 7.4 buffer was mixed with increasing concentrations of each amine overnight before determining the readings. The absorbance readings from 220 nm to 500 nm were taken. The K_d values were calculated by fitting the binding curves into Hyperbola equation.

For the prolinomycin mutants containing 2APBA, WW-Pro1, 2 and 4 were dissolved in 1X PBS pH 7.4 buffer with 200 mM KCl and WW-Pro3 was in the same buffer containing 50% methanol. 1 M stock of each amine was prepared and titrated into the cuvette

containing 0.5 ml peptide solution. The background absorbance of amines was determined and subtracted.

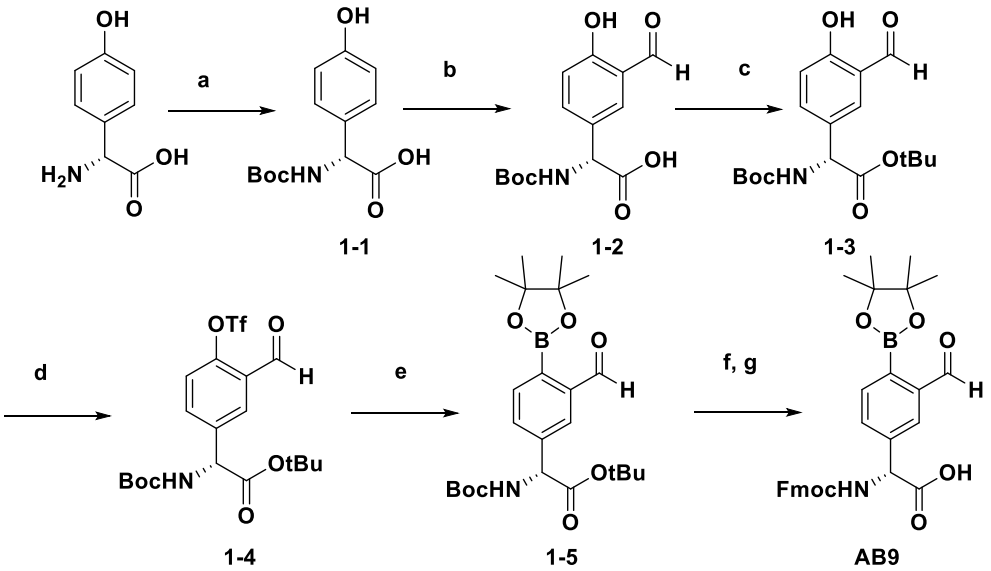
2.6.3 ^{1}H NMR Iminoboronate Conjugate Analysis

0. 2 mM of each derivatives of benzaldehyde and acetophenone in 1X PBS pH 7.4 buffer with 10% D_2O was mixed with increasing concentrations of each amine. ^{1}H NMR was obtained with water suppression and 128 scans.

2.6.4 Kinetic Analysis of the imine formation by UV-Vis

Each amine was added into the cuvette containing 0.5 mL 50 μM of each derivatives of benzaldehyde and acetophenone in 1X PBS pH 7.4 buffer. The absorbance reading was taken every 1 min for 25 min after the amine was added.

2.6.5 Synthesis of AB9



Scheme 2-1. Synthesis of AB9

a. Synthesis of 1-1: Phenylglycine (18 mmol) was dissolved in 5% Na_2CO_3 (75 mL) and mixed with Di-tert-butyl dicarbonate (18 mmol) dissolved in THF (75 mL). The

reaction was allowed to react overnight with stirring at room temperature. The reaction was acidified with 1N HCl and extracted with ethyl acetate for 3 times. The combined organic layers were washed with brine and dried over sodium sulfate. The solvent was then evaporated under vacuum. Hexane and chloroform were added and evaporated to get compound 1-1 in the form of white powder (95%).

1H NMR (500 MHz, $CDCl_3$): δ 7.13 (m, 2H), 6.69 (d, 2H), 5.0-5.2 (m, 1H), 1.40 (s, 5H), 1.29 (s, 4H)

b. Synthesis of 1-2: Compound 1-1 (6.76 mmol) was dissolved in $CHCl_3$ (25 mL) at 70 ºC for 5 min. Meanwhile, NaOH (27.7 mmol) was dissolved in water (4 mL) at 70 ºC for 5 min. Then, NaOH solution was added into 1-1 solution dropwise and stirred at 70 ºC for 2 h. The reaction was cooled down for 1 h and the solvent was evaporated under vacuum. The mixture was diluted with water (150 mL) and ethyl acetate (150 mL). 1N HCl was added to acidify the mixture to pH 1. The resulting mixture was extracted with ethyl acetate. The combined organic layer was washed with brine and dried over sodium sulfate. The crude product was purified on silica gel using dichloromethane: methanol (95:5) as eluent. The desired product was obtained as yellow oil (40%).

   1H NMR (500 MHz, Acetone-d_6): δ 10.06 (s, 1H), 7.86 (d, 1H), 7.69-7.72 (q, 1H), 7.00 (d, 1H), 6.64 (s, 1H), 5.30-5.31 (d, 1H), 1.40-1.41 (m, 9H).

c. Synthesis of 1-3: Compound 1-2 (1.35 mmol) was dissolved in dry ethyl acetate (12 mL). tert-Butyl 2,2,2-trichloroacetimidate (2.71 mmol) was added. The reaction was stirred at 65 ºC for an overnight. Ethyl acetate was evaporated. The crude product was

purified on silica gel using gradient elution of hexane: ethyl acetate mixture from 5% to 20% ethyl acetate. The desired product was obtained as light yellow oil (47%).

1H NMR (500 MHz, $CDCl_3$): δ 11.01 (1, 1H), 9.89 (s, 1H), 7.56 (s, 1H), 7.53-7.51 (q, 1H), 6.97 (d, 1H), 5.67 (d, 1H), 5.17 (d, 1H), 1.43 (s, 9H), 1.39 (s, 9H)

d. Synthesis of 1-4: Compound 1-3 (3.13 mmol) was dissolved in anhydrous dichloromethane under argon and cooled in an acetone-dry ice bath. N-Methylmorpholine (12.52 mmol) was added slowly with stirring. After 5 min, trifluoromethanesulfonic anhydride (6.26 mmol) was added dropwise. After 10 min, the reaction was warmed up to room temperature and stirred for 30 min. Saturated sodium bicarbonate (30 mL) was added to stop the reaction. The resulting mixture was extracted with dichloromethane for 3 times. The combined organic layer was washed with brine, dried over sodium sulfate and concentrated. The crude product was purified on silica gel using hexane: ethyl acetate (9:1) as eluent. The desired product was obtained as white solid. (90%)

   1H NMR (500 MHz, $CDCl_3$): δ 10.26 (1, 1H), 8.01 (d, 1H), 7.75-7.73 (d, 1H), 7.40 (d, 1H), 5.76 (d, 1H), 5.30 (d, 1H), 1.43 (s, 9H), 1.39 (s, 9H)

e. Synthesis of 1-5: Compound 1-4 (0.21 mmol), B_2Pin_2 (0.525 mmol), [1,1'-Bis(diphenylphosphino)ferrocene]palladium(II) dichloride (0.017 mmol) and potassium acetate (0.63 mmol) were mixed together in anhydrous dioxane (1.24 mL) under argon. The mixture was stirred and heated at 87 ºC for 15 min and cooled down for 20 min. Water was added to stop the reaction. The mixture was extracted with diethyl ether for 3 times. The combined organic layer was washed with brine, dried over sodium sulfate and concentrated. The crude product was purified on silica gel

using gradient elution of 5% to 15% diethyl ether in hexane. The product was obtained as white powder. (72%)

1H NMR (500 MHz, Acetone-d6): δ 10.54 (s, 1H), 8.00 (d, 1H), 7.88-7.86 (d, 1H), 7.50-7.73 (d, 1H), 6.65 (d, 1H), 5.30 (d, 1H), 1.40-1.20 (m, 30H).

f. Synthesis of AB9: Compound 1-5 (0.26 mmol) was dissolved in dichloromethane (1 mL). Trifluoroacetic acid (TFA, 2 mL) was added and stirred for 1.5 h. Then, the solvent was evaporated under vacuum. TFA (1.5 mL) was added again and stirred for another hour. TFA was removed by evaporation and yellow oil intermediate was obtained. Dichloromethane (3 mL), N-methylmorpholine (3 mmol) and Fmoc-Osu (0.24 mmol) were added and the reaction was stirred for 3 h at room temperature. The mixture was acidified with 2N HCl and extracted with ethyl acetate for 3 times. The combined organic layer was washed with brine and dried over sodium sulfate. The solvent was evaporated under vacuum. The crude product was purified on silica gel using hexane: ethyl acetate: acetic acid (50:50:1) as eluent. The product was white powder (65%).

   1H NMR (500 MHz, Acetone-d_6): δ 7.76 (s, 2H), 7.74 (s, 1H), 7.71 (d, 1H), 7.58 (d, 1H), 7.39 (t, 2H), 7.34 (d, 1H), 7.30 (t, 3H), 6.48 (s, 1H), 5.80 (s, 1H), 5.43 (d, 1H), 4.36 (t, 2H), 4.21 (t, 1H), 1.34 (s, 12H).

   MS-ESI$^+$ (m/z): Calculated for C30H30BNO7 $[M+H]^+$ 528.21, found 528.22.

2.6.6 Peptide synthesis

Peptides were synthesized according to the reference[19]. The prolinomycin linear precursor peptides were synthesized on Dawson Dbz resin through Fmoc/tBu chemistry by

the peptide synthesizer. Unnatural amino acids including N-Boc-(2S,4R)-tPro(Trt)-OH, AB3 and AB9 were synthesized and coupled onto the resin manually at 2 equivalents for 2 h. Short-chain analogues of Arg, 2-amino-4-guanidino-butyric acid (Agb) and 2-amino-3-guanidino-propionic acid (Agp), were synthesized on resin by incorporating Fmoc-Dab(Alloc)-OH and Fmoc-Dap(Alloc)-OH as precursors. For 0.025 mmol scale of peptide synthesis, the alloc group was deprotected using a mixture of 30 mg tetrakis(triphenylphosphine)-palladium ($Pd(PPh_3)_4$) and 0.2 mL phenylsilane in 2 mL dichloromethane for 1 h. Then, 4 equivalents of N, N′-Di-Boc-1H-pyrazole-1-carboxamidine was added in DMF (1 mL) with DIPEA (20 μL) for 2 h to accomplish the guanylation reaction.

After synthesizing the linear peptide, Dbz resin was converted to the Nbz moiety by mixing the resin with 5 equivalents of 4-nitrophenylchloroformate in DCM (3 mL) for 1 hour, followed by 20 min incubation with 3 mL of 0.5 M NMM in DMF. The peptides were cleaved off the resin and globally deprotected with 2 mL reagent K (80% TFA, 5% H_2O, 2.5% EDT, 5% thioanisole and 7.5% phenol). For the peptide containing 2-APBA and 2-FPBA, the reagent with 95% TFA and 5% H_2O was used instead. The crude peptides were subjected to the proline ligation reaction directly.

The crude peptide (2-5 mM) was dissolved in 200 μL DMF and mixed with 300 μl ligation buffer (0.2 M sodium phosphate, 3 M guanidinium chloride, 20 mM 4-mercaptophenylacetic acid, 10 mM TCEP, pH 8) for an overnight. The reaction was monitored by LC-MS analysis. After completion, the cyclized peptide was purified by RP-HPLC (Waters Prep LC, Phenomenex C18 Column). The fractions containing the peptides were combined and lyophilized.

2.6.7 Desulfurization of prolinomycin mutants

First, Raney nickel reagent was prepared freshly. Nickel acetate (200 mg) was dissolved in water (5 mL). sodium borohydride (60 mg) was added slowly with stirring for 10 min. The desired product was obtained as black precipitate by vacuum filtration and washed by water. 2 mg peptide was dissolved in 20% acetic acid aqueous solution with 1 mg TCEP and 20 mg Raney nickel. A balloon containing hydrogen gas was attached and the reaction was stirred for 5 h. The reaction was monitored by LC-MS analysis. The desired product (50%) was purified by RP-HPLC and lyophilized.

2.6.8 Synthesis of crosslinker 2

The crosslinker 2 was synthesized according to the reference[25]. Briefly, 1,3,5-benzenetricarboxylic acid (2.41 mmol) was dissolved in THF (100 mL). N-Hydroxysuccinimide (7.90 mmol) and N, N'-dicyclohexylcarbodiimide (8 mmol) were added and the reaction was stirred overnight at room temperature. White suspension was filtered. The supernatant was concentrated to yield white solid which was recrystallized in isopropanol (300 mL). The desired product was obtained as white solid (80%).

1H NMR (500 MHz, $CDCl_3$) 9.14 (s, 3H), 2.94 (s, 12H).

2.6.9 Synthesis of cross-linked prolinomycin mutants

The thiol group was first removed by Raney nickel reagent or capped by iodoacetamide. The reaction was monitored by LC-MS analysis. The peptide was purified by RP-HPLC and lyophilized. 1 mg peptide and 1 equivalent of the cross-linker (stock in DMF) was mixed in 1XPBS buffer pH 7.5 containing 500 mM KCl. The reaction was left to sit for 2 h and analyzed by LC-MS. The desired product was purified by RP-HPLC.

2.6.10 LC-MS analysis of prolinomycin mutants

The mass spec data of prolinomycin mutants synthesized are summarized in Table 2-5.

Table 2-5. Mass spec data of prolinomycin mutants

Mutants	Calculated $[M+H]^+$	Observed $[M+H]^+$
RRRWT	1441.76	1441.76
RRRWTdeSH	1409.78	1409.78
RRRWT-trans-L-Hyp	1473.75	1473.76
RRRWT-cis-D-hyp	1457.75	1457.76
DabWRdeSH	1296.73	1296.64
CL1DabdeSH	1433.74	1434.74
CL2DabdeSH	1452.71	1453.71
CL2Dap	1499.66	1499.66
WW-Pro1	1359.60	1359.67
WW-Pro2	1429.77	1411.72
WW-Pro3	1464.75	1463.83
WW-Pro4	1705.66	1705.66

2.6.11 ^{1}H NMR characterization of prolinomycin mutants

Peptide samples for NMR analysis were prepared in 50 mM sodium phosphate buffer, pH 4.0, with 10% D_2O or 1:1 mixture of CD_3OH and 50 mM sodium phosphate buffer, pH 4.0, with 10% D_2O. The concentration of peptide was 0.2 mM. 0.2 mM maleic acid was added as the internal standard. ^{1}H NMR spectra of the peptides were collected on a Varian INOVA 500 or 600 MHz NMR spectrometer. Water suppression was done to purge and subtract the water peaks by using a PRESAT experiment. 128 scans were done in total for each sample. For KCl titration experiment, NMR spectra were collected after sequentially adding 4M KCl stock into the sample. NMR data were processed and analyzed by MestReNova.

2.6.12 Fluorophore labeling of prolinomycin mutants

3 mM peptide was dissolved in DMF with 6% NMM by volume and 1 equivalent of TCEP. 2 equivalents of Alexa Fluor 488 C5 maleimide were added into the solution and the reaction was allow to sit for 2 h in the dark. The reaction was monitored by LC-MS analysis. The fluorophore labeled peptides were purified by RP-HPLC (Waters Prep LC, Phenomenex C18 Column) and lyophilized.

2.6.13 Fluorescence analysis of prolinomycin dimerization

200 nM of Alexa Fluor 488 labeled peptide in 1XPBS buffer 7.4 with 1 M KCl was added into Corning® black/clear bottom 96 well plate with increasing concentrations of Dabcyl labeled peptide (0-60 μM). The fluorescence readings were obtained on the plate reader (SpectraMax M5) with excitation wavelength 495 nm and emission wavelength 530 nm.

2.6.14 BSA binding assay and competition assay using fluorescence anisotropy

200 nM of Alexa Fluor 488 labeled peptide 1X PBS buffer 7.4 with 1 M KCl was added into Corning® opaque 96 well plate with increasing concentrations of BSA (0-10 mg/mL). The fluorescence anisotropy was obtained on the plate reader (SpectraMax M5) with excitation wavelength 495 nm and emission wavelength 530 nm.

For the competition assay, 200 nM of Alexa Fluor 488 labeled peptide 1X PBS buffer 7.4 with 1 M KCl was mixed with 1 mg/mL BSA first in the 96 well plate. Increasing concentrations of 3Orn peptide was added. The fluorescence anisotropy was obtained on the plate reader (SpectraMax M5) with excitation wavelength 495 nm and emission wavelength 530 nm. The inhibition curves and IC_{50} were calculated by software Graphpad Prism8.

2.7. References

1. Gellman, S.H., *Foldamers: A Manifesto.* Accounts of Chemical Research, 1998. **31**(4): p. 173-180.
2. Cheng, R.P., S.H. Gellman, and W.F. DeGrado, *β-Peptides: from structure to function.* Chemical reviews, 2001. **101**(10): p. 3219-3232.
3. Kirshenbaum, K., R.N. Zuckermann, and K.A. Dill, *Designing polymers that mimic biomolecules.* Current opinion in structural biology, 1999. **9**(4): p. 530-535.
4. Gellman, S.H., *Foldamers: a manifesto.* Accounts of Chemical Research, 1998. **31**(4): p. 173-180.
5. Horne, W.S. and S.H. Gellman, *Foldamers with heterogeneous backbones.* Accounts of chemical research, 2008. **41**(10): p. 1399-1408.
6. Hamuro, Y., J.P. Schneider, and W.F. DeGrado, *De novo design of antibacterial β-peptides.* Journal of the American Chemical Society, 1999. **121**(51): p. 12200-12201.
7. Karlsson, A.J., et al., *Antifungal Activity from 14-Helical β-Peptides.* Journal of the American Chemical Society, 2006. **128**(39): p. 12630-12631.
8. Liu, D. and W.F. DeGrado, *De Novo Design, Synthesis, and Characterization of Antimicrobial β-Peptides.* Journal of the American Chemical Society, 2001. **123**(31): p. 7553-7559.
9. Kritzer, J.A., et al., *Helical β-Peptide Inhibitors of the p53-hDM2 Interaction.* Journal of the American Chemical Society, 2004. **126**(31): p. 9468-9469.
10. Girvin, Z.C. and S.H. Gellman, *Foldamer Catalysis.* Journal of the American Chemical Society, 2020. **142**(41): p. 17211-17223.

11. Girvin, Z.C., et al., *Foldamer-templated catalysis of macrocycle formation.* Science, 2019. **366**(6472): p. 1528-1531.

12. Yoo, B., et al., *Peptoid Macrocycles: Making the Rounds with Peptidomimetic Oligomers.* Chemistry – A European Journal, 2010. **16**(19): p. 5528-5537.

13. Driggers, E.M., et al., *The exploration of macrocycles for drug discovery--an underexploited structural class.* Nat Rev Drug Discov, 2008. **7**(7): p. 608-24.

14. Mallinson, J. and I. Collins, *Macrocycles in new drug discovery.* Future Med Chem, 2012. **4**(11): p. 1409-38.

15. Li, Y., et al., *Versatile protein recognition by the encoded display of multiple chemical elements on a constant macrocyclic scaffold.* Nature Chemistry, 2018. **10**(4): p. 441-448.

16. Hartgerink, J.D., et al., *Self-Assembling Peptide Nanotubes.* Journal of the American Chemical Society, 1996. **118**(1): p. 43-50.

17. Gisin, B. and R. Merrifield, *Synthesis of a hydrophobic potassium binding peptide.* Journal of the American Chemical Society, 1972. **94**(17): p. 6165-6170.

18. Gisin, B., et al., *Selective ion binding and membrane activity of synthetic cyclopeptides.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 1978. **509**(2): p. 201-217.

19. Hosseini, A.S., et al., *Metal-Assisted Folding of Prolinomycin Allows Facile Design of Functional Peptides.* ChemBioChem, 2017. **18**(5): p. 479-482.

20. Holmgren, S.K., et al., *Code for collagen's stability deciphered.* Nature, 1998. **392**(6677): p. 666-667.

21. Cambray, S. and J. Gao, *Versatile Bioconjugation Chemistries of ortho-Boronyl Aryl Ketones and Aldehydes.* Accounts of chemical research, 2018. **51**(9): p. 2198-2206.

22. Cal, P.M.S.D., et al., *Iminoboronates: A New Strategy for Reversible Protein Modification.* Journal of the American Chemical Society, 2012. **134**(24): p. 10299-10305.

23. Gutiérrez-Moreno, N.J., F. Medrano, and A.K. Yatsimirsky, *Schiff base formation and recognition of amino sugars, aminoglycosides and biological polyamines by 2-formyl phenylboronic acid in aqueous solution.* Organic & biomolecular chemistry, 2012. **10**(34): p. 6960-6972.

24. Greenamyre, J.T., et al., *Alterations in L-glutamate binding in Alzheimer's and Huntington's diseases.* Science, 1985. **227**(4693): p. 1496-1499.

25. Megens, R.P., et al., *Multinuclear non-heme iron complexes for double-strand DNA cleavage.* Chemistry, 2009. **15**(7): p. 1723-33.

Chapter 3 Developing peptide binders of proteins by phage display

3.1. Introduction

Over the past two decades, peptides have emerged as a versatile modality of therapeutics to modulate various protein targets that are challenging, or even undruggable, for other current pharmaceuticals[1]. In comparison to traditional small-molecule drugs, peptides possess larger sizes and more extended structures, and hence tend to interact with targets via a wider surface area leading to improved selectivity and affinity[3]. On the other hand, peptides can also be superior over large biologics in terms of immunogenicity, safety, cell permeability and manufactory cost[5]. Therefore, they have the potential to integrate the best attributes of small molecules and large biologics to provide an alternative strategy to reimagine medicine.

In this chapter, I will first introduce protein targets uniquely amenable to peptide inhibition. Furthermore, the latest strategies of chemically modified phage display to develop bioactive peptides will be described. Our work on applying phage display to target a model protein, Sortase A (SrtA), will be demonstrated. The structure-activity relation of peptide inhibitors of SrtA has been investigated. Last but not least, we will discuss the remaining challenges and future directions of chemically modified phage display for targeting proteins.

3.1.1. Protein targets amenable to peptides

Proteins are appealing targets for drug development and disease biomarkers discovery. About 30% of human protein-coding gene were predicted to encode secreted proteins or membrane-bound proteins with extracellular domains which are crucial to many biological processes[6]. Cytokines, coagulation factors, hormones and growth factors are examples of important extracellular proteins secreted by certain tissue cells. Ion channels, transporters, and receptors comprise the majority of membrane proteins. These extracellular proteins are particular appealing to peptide-based drugs in contrast to their intracellular counterparts,

which would require peptides of high membrane permeability. Herein we first summarize several families of extracellular proteins that have been successfully targeted by peptides.

Extracellular proteases are critical players in extracellular matrix remodeling, blood coagulation, complement fixation and fibrinolysis[7, 8]. They have been associated with many diseases including coagulopathies, inflammation, infectious diseases, cancer and degenerative diseases[9, 10]. Due to their highly conserved catalytic domains and shallow substrate binding sites, it is difficult to develop small molecule inhibitors of sufficient potency and specificity[11, 12]. Antibodies have been generated with increased specificity and affinity to inhibit these proteases, however, their utilization has been limited by their poor tissue penetration capabilities[13]. Peptides with mono-[14] and bicyclic[15, 16] structures have been identified as uPA inhibitors from phage display libraries. The bicyclic peptide hit has the 10-fold stronger affinity than the monocyclic comparison does, presumably due to the fact that the bicyclic peptides forge a wider contact surface with the target protein uPA (Figure 3-1a). For example, the bicyclic peptide UK18 cross-linked by the trifunctional tris(bromomethyl)benzene (TBMB) presents 53 nM potency (in contrast to 500 nM for the monocyclic peptide)[15]. Another class of challenging extracellular proteases is matrix metalloproteinases (MMPs). Similar to trypsin-like serine proteases, MMPs share conserved, large and open active sites. Researchers have spent much effort to develop selective inhibitors of specific MMPs but with little success[17]. Screening of the bicyclic phage library yielded an initial hit for the matrix metalloproteinase MMP-2. Further modification/optimization of the initial hit with a zinc ion chelating moiety resulted in the first synthetic MMP-2 inhibitor with single-digit nanomolar potency and target selectivity (Figure 3-1b)[18]. Furthermore, the phage-encoded library based on double-

bridged peptides enable them to identify peptide inhibitors of coagulation factor XIa[19]. These structurally constrained cyclic peptides are able to cover surface areas of 600-700 Å^2, over which multiple hydrogen bonds and charge-charge interactions are made to afford

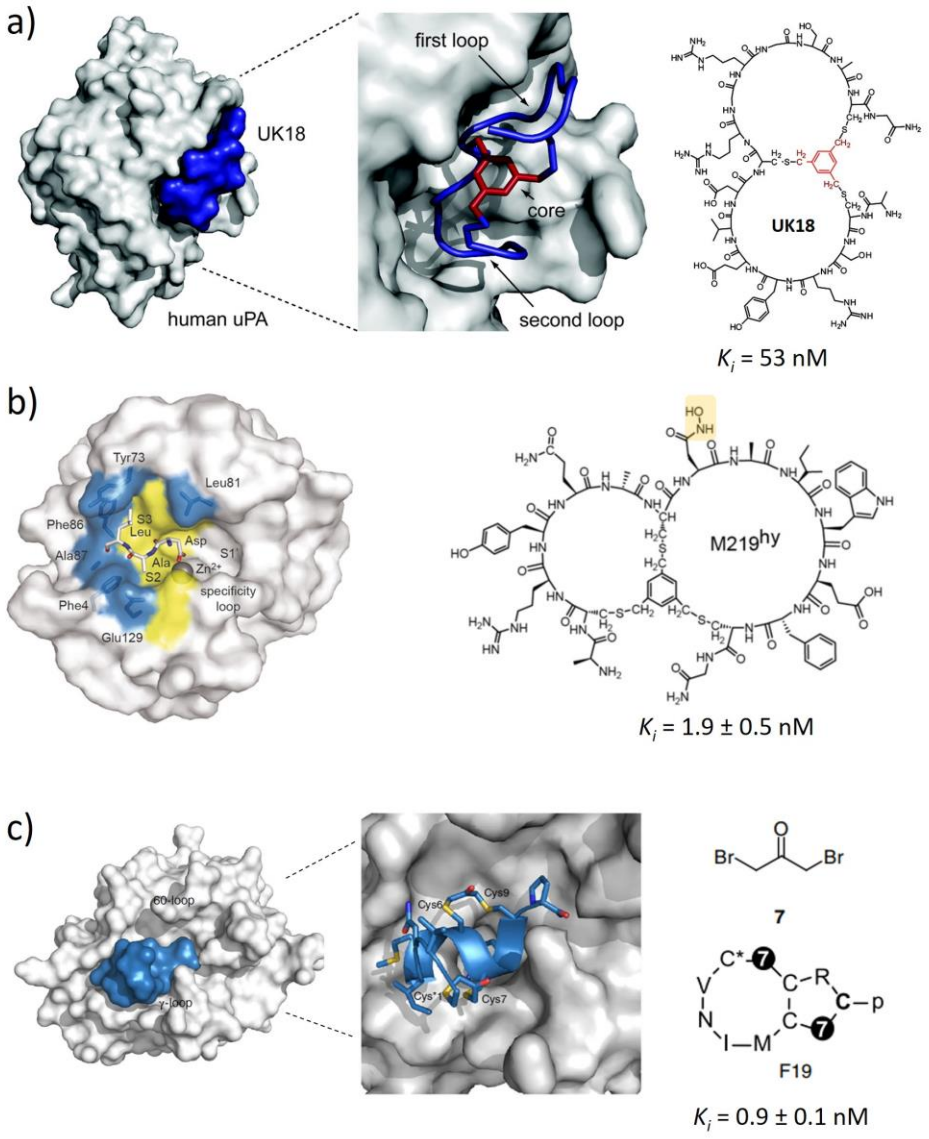


Figure 3-1. Extracellular proteases amenable to peptide inhibitors.a) Crystal structure of human uPA in complex with the bicyclic peptide UK18. Adapted with permission from Ref. [15]. Copyright (2019) American Chemical Society. b) Chemical structure of M219hy that can bind MMP-2 by mimicking Asp-Ala-Leu motif of the amyloid precursor protein. Adapted with permission from Ref. [18]. c) Crystal structure of double-bridged peptide F19 bound to the active site of FXIa. Adapted with permission from Ref. [19].

low nanomolar affinity and specificity (Figure 3-1 c)[15]. Importantly, cyclization with non-natural crosslinking motifs can significantly enhance the metabolic stability of peptides.

Transmembrane receptors play many essential roles in various cellular signaling pathways which are associated with cell proliferation, migration, differentiation, immune

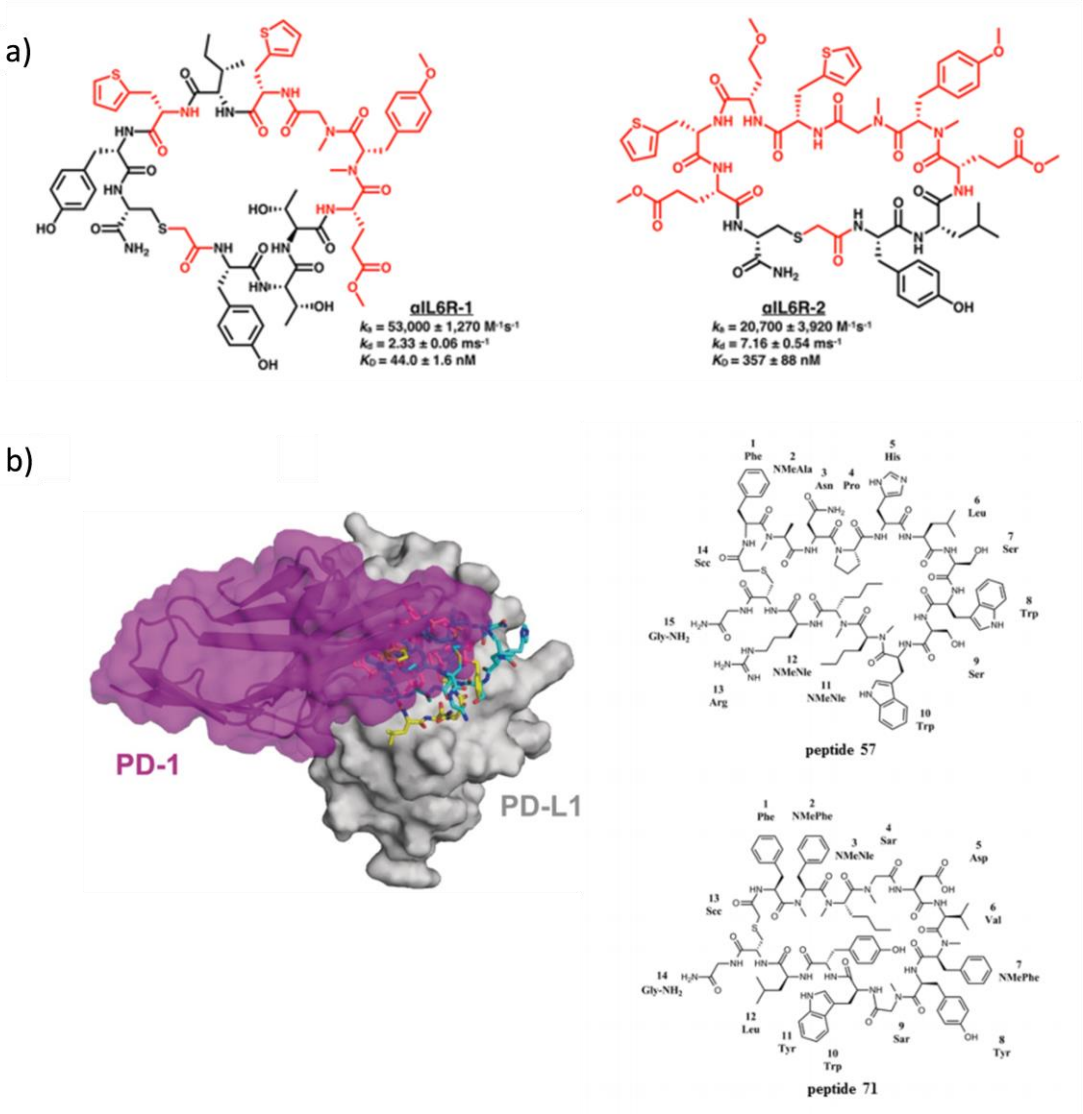


Figure 3-2. Peptide inhibitors of transmembrane receptors. a) Macrocyclic peptides generated by RaPID system that can inhibit IL6R. Adapted with permission from Ref. [22]. Copyright (2019) American Chemical Society. b) Macrocyclic peptides can inhibit PD-1/PD-L1 interaction by binding to PD-L1 at the same site as PD-1. Peptides-57 (blue) and -71 (yellow). Adapted with permission from Ref. [25].

response and angiogenesis. Some natural endogenous peptides including insulin, oxytocin and calcitonin, have been discovered as ligands for receptors and used as therapeutics for a long time[20]. Recently, more and more challenging extracellular receptors have been targeted by macrocyclic peptides discovered by various advanced display screening techniques. Take interleukin-6 receptor (IL6R) as an example which is a cancer-related target and has only been drugged by antibody-based therapeutics. Disulfide-rich peptides isolated from c-DNA display library[21] and monocyclic peptides presenting unnatural amino acid residues generated by RaPID system[22] were able to bind IL6R with dissociation constants in the nanomolar range (Figure 3-2a). In addition, an increasing amount of interests and effort have been focused on the inhibition of immune checkpoint PD-1/PD-L1 pathway to treat cancer. The immunotherapies based on antibodies have been developed as a powerful strategy, however, high production cost and severe immune-related adverse events limit their application[23]. As an alternative to antibodies, macrocyclic peptides were investigated to inhibit the interaction between PD-1 and PD-L1[24]. Bristol-Myers-Squibb reported macrocyclic peptides which can inhibit PD-1/PD-L1 interaction with nanomolar activity by binding to PD-L1[25], which are peptide-57 and -71 with EC_{50} values of 566 nM and 293 nM respectively (Figure 3-2b). The crystal structure of PD-L1/peptide complex showed that the peptide could mimic the hydrophobic interactions of the antibody residues leading to the high affinity.

70% of human proteome are intracellular including most of the therapeutically relevant PPIs. However, intracellular targets only comprise 30% of targets for approved drug[6]. The inconsistency between the number of the total intracellular targets and the number of ones having been drugged indicates the challenge to develop drugs against intracellular proteins.

One reason is that the plasma membrane can prevent the entry of biologics and many charged molecules, which therefore limits the scope of drug candidates. Moreover, small-molecule drugs often have low affinity towards PPIs interface, off-target activities and toxicity, making them insufficient to target a large group of intracellular targets.

Unlike large biologics, a large number of peptides can cross the plasma membrane by various mechanisms including passive diffusion, heparin sulfate-facilitated diffusion, receptor-mediated endocytosis, pinocytosis, and transporter-mediated cell entry[26, 27].

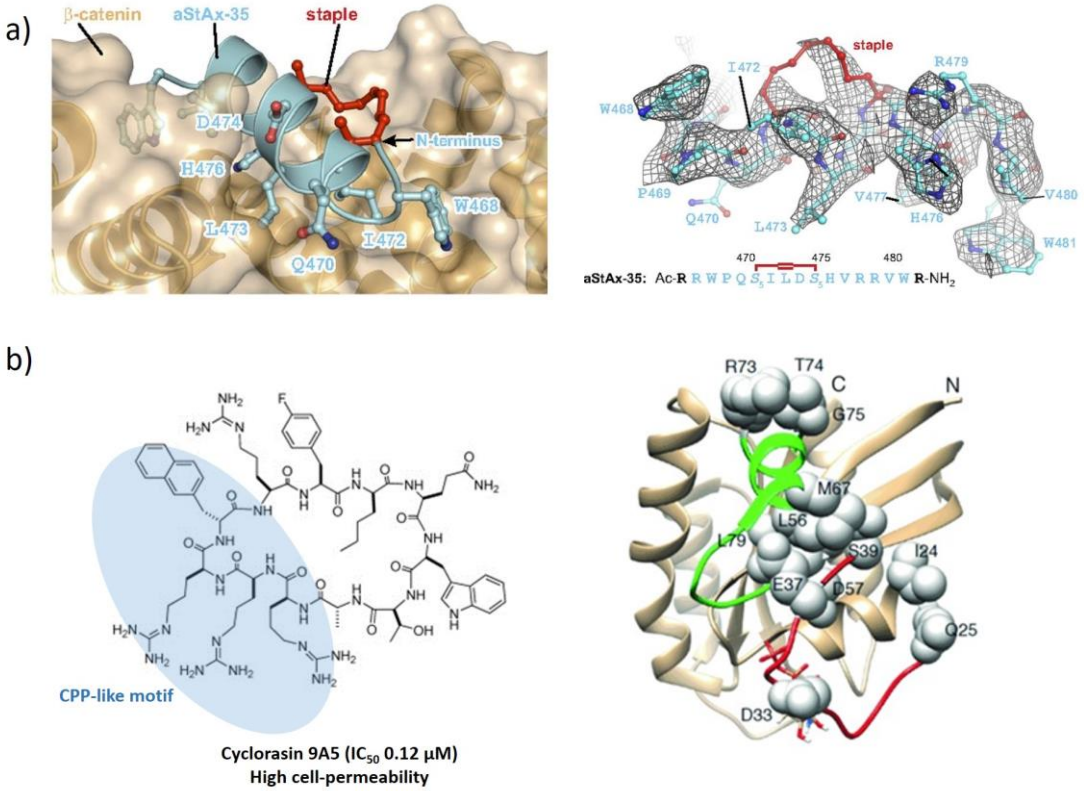


Figure 3-3. Intracellular PPIs based on peptides. a) Crystal structure of the complex of β-catenin and aStAx-35, a stapled peptide. Adapted with permission from Ref. [33]. Copyright (2012) National Academy of Sciences. b) A cell permeable peptide inhibitor of K-Ras. K-Ras residues perturbed by 9A5 are determined by NMR spectra and shown as spheres. Adapted with permission from Ref. [38].

Particularly, macrocyclization, hydrophobicity, positively charged amino acid residues and N-methylation can improve cellular uptake of peptides[28]. Therefore, an increasing number of intracellular challenging proteins and PPIs have been targeted by macrocyclic peptides[29].

For instance, β-catenin works as an important mediator in the Wnt signaling pathway regulating cell proliferation and differentiation. β-catenin has been considered as a potential target for cancer therapy but never been drugged[30]. Various macrocyclic peptides identified by in silico design[31], phage display[32] and stapled peptides were reported to bind β-catenin (Figure 3-3a)[33, 34]. Particularly, the cell permeability of a stapled peptide was optimized by polar N-terminal modifications together with the hydrophobic core sequence[35]. Hence, the peptide could inhibit the interaction between β-catenin and T cell factor/lymphoid enhancer-binding factor (TCF/LEF) transcription factors leading to the inhibition of Wnt-dependent cell lines proliferation and migration. Similarly, other PPIs such as p53-MDM2 and BH3-MCL-1 have all been targeted by various stapled peptides[36]. Furthermore, the cell penetrating peptides (CPPs) can work as membrane transporter to deliver macrocyclic peptides into the cytoplasm to selectively inhibit PPIs and enzymes such as protein-tyrosine phosphatase 1B and a peptidyl-prolyl cis−trans isomerase Pin[37]. Remarkably, K-Ras, a long-term undruggable cancer related target, was successfully targeted by a family of macrocyclic peptides containing the CPP-like motif[38]. These cell-permeable peptides can inhibit the interaction between K-Ras and effector protein by binding to Ras-GTP with high potency (IC_{50} = 0.12 μM) and selectivity.

Last but not least, the ubiquitin (Ub) system, a promising target for cancer therapy, can also be recognized and inhibited by macrocyclic peptides. A previously undruggable E3

ubiquitin ligase was targeted by a macrocyclic N-methyl-peptide with single-digit nanomolar affinity[39]. The peptide could prevent proteins from the polyubiquitination catalyzed by E6AP. Furthermore, an even more challenging target, K48-linked Ub chain was modulated by a family of cell-permeable macrocyclic peptides, which could protect Ub-tagged proteins from degradation and consequently induce apoptosis in cancer cells[40, 41]. The macrocyclic peptide possessed the antibody-like affinity (nanomolar K_d) and linkage specificity together with the excellent membrane permeability.

3.1.2. Chemically modified phage display

Phage display is a powerful technique to develop peptide inhibitors of protein targets. However, phage-encoded peptide libraries are limited to display the 20 canonical amino acids, which restricts the chemical space that can be explored. In addition, short peptides are usually flexible and unstructured resulting in high entropic energy required to bind the targets. Disulfide bonds are frequently introduced to constrain the conformation of peptides, however, these bonds are not stable and can be reduced. Moreover, such peptides suffer from degradation by proteases in gastrointestinal tract and blood plasma, which is the major obstacle for oral peptide drug development.

Chemical modification is able to provide facile solutions to the above problems. One promising example is phage-encoded bicyclic peptides library developed by Winter and Heinis (Figure 3-4 a)[42]. The libraries were generated by cyclizing peptides containing randomized amino acids in between three cysteine residues with the chemical linkers like 1,3,5-tris(bromomethyl)-benzene (TBMB) and 1,3,5-triacryloyl-1,3,5-triazinane (TATA). From these bicyclic peptide libraries, inhibitors of various proteases including plasma kallikrein (PK)[42], carthepsin G, urokinase-type plasminogen activator (uPA)[15] and coagulation factor XIIa[43], as well as Her2[44] and Notch 1 receptors, have been identified affording nanomolar affinity and superb selectivity. More recently, Heinis and his coworkers developed a new strategy to generate libraries of genetically encoded double-bridged peptides (Figure 3-4b)[19, 45]. In this strategy, peptides contain four cysteine

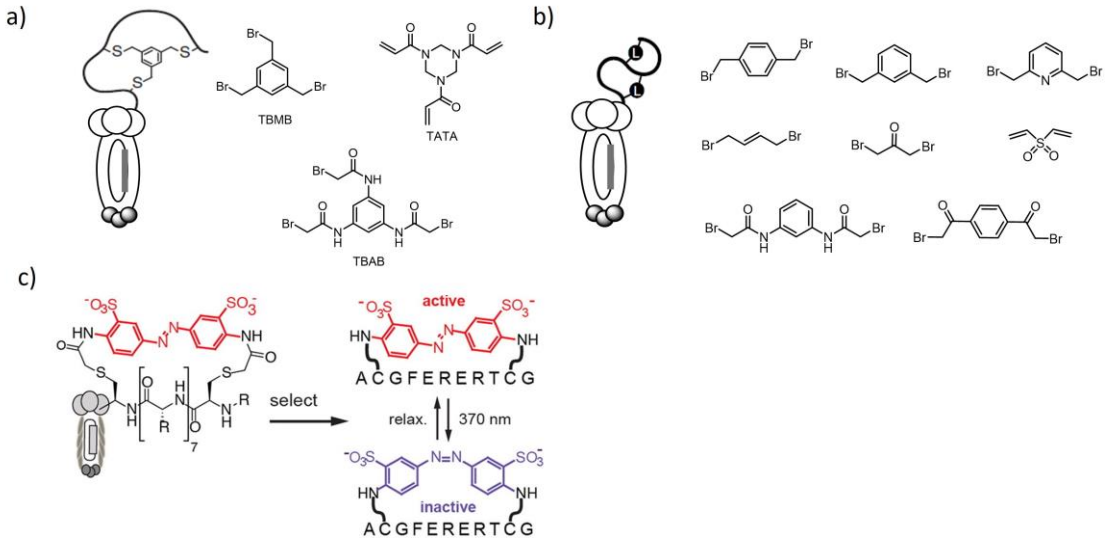


Figure 3-4. Phage displayed cyclic peptide library. a) The bicyclic peptides library generated by crosslinking three cysteine residues by various chemical linkers. b) The double-bridged peptides library generated by cyclizing two pairs of cysteine residues by a panel of chemical linkers. c) The photoswitchable peptide library. Adapted with permission from Ref. [37]. Copyright (2014) American Chemical Society.

residues were cyclized with chemical linkers like bis(bromomethyl)benzene, forming two bridges that segment the peptides into three parts containing random amino acids. These libraries yielded short-peptide based inhibitors of PK and Factor XIa with subnanomolar affinity and selectivity larger than 1,000-fold. Remarkably, the double-bridged peptides showed resistance to proteolytic degradation, which proves a powerful strategy to develop oral peptide drugs. In addition, photoswitchable peptide ligands were identified by Derda group from phage display libraries cyclized with an azobenzene linker (Figure 3-4c)[46]. The affinity of the peptides against the target, streptavidin, can be modulated by UV exposure, showing K_d of 452 μM in the relaxed trans conformation and >2000 μM after UV exposure.

Not only have the chemical linkers advanced phage display, covalent warheads have also been incorporated into the phage-encoded peptide libraries by modifying the reactive

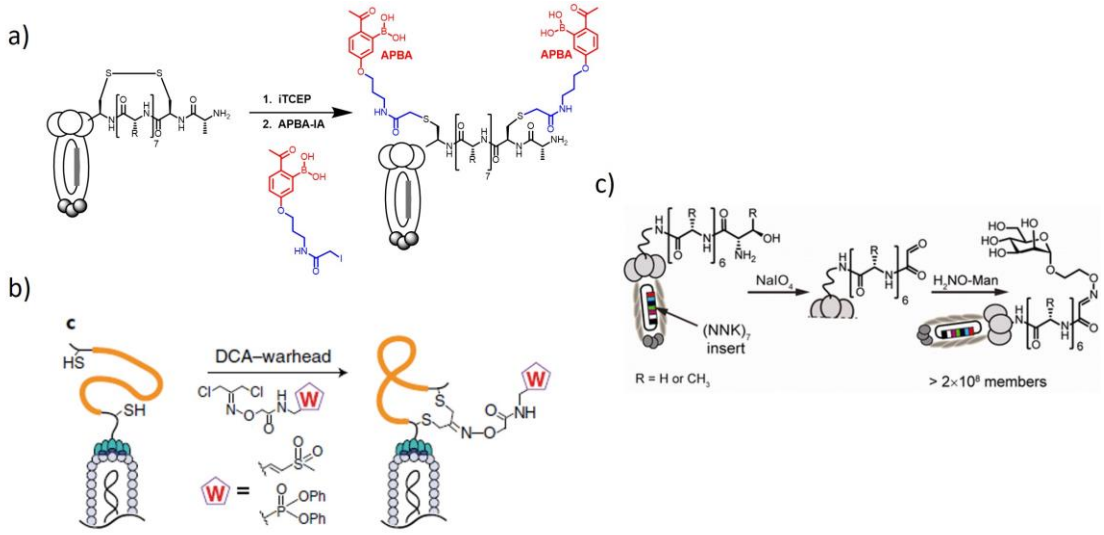


Figure 3-5. Phage-encoded peptide library presenting chemical warheads. a) The phage display of dynamic covalent binding motifs, APBA. Adapted with permission from Ref. [46]. Copyright (2018) American Chemical Society. b) The phage display of vinyl sulfone and diphenylphosphonate warheads. Adapted with permission from Ref. [48]. c) The phage-encoded glycopeptides library.

cysteine residues. Covalent bond formation provides a strong driving force for ligand binding. For example, our group reported the phage display of dynamic covalent binding motifs, APBA, for targeting the bacterial cell surface (Figure 3-5a)[47, 48]. The iminoboronate chemistry provides a promise binding mechanism leading to submicromolar potency peptide binders of bacteria including *S. aureu* and colistin-resistant bacterial strains. The peptides possess remarkable selectivity that they can differentiate drug-resistant bacterial strains from their wild-type counterparts. In addition, Bogyo and his coworkers recently installed a cysteine-reactive α-amino vinyl sulfone warhead and a serine-reactive diphenylphosphonate warhead to target a cysteine protease and a serine hydrolase respectively (Figure 3-5b)[49]. The peptide inhibitors of TEV protease and FphF hydrolase with submicromolar potency and exceptional specificity were identified. This method provides an effective way to develop selective covalent inhibitors for protein targets.

The N-terminal Ser/Thr of peptides displayed by phage also provide a unique site for chemical modification. Derda group has introduced glycan moieties into the peptide library via an oxime ligation following oxidation of N-terminal Ser/Thr (Figure 3-5c)[50, 51]. For example, by incorporating a mannose moiety into the phage-displayed library, glycopeptide ligands of a mannose-binding protein, concanavalin A, were identified. The ligands afforded low micromolar affinity that is 40 to 50-fold enhancement over the mannose moiety itself.

In summary, chemical modification of peptides encoded by phage display significantly expanded the chemical space of traditional phage display technique. This strategy can avoid the use of sophisticated technologies for unnatural amino acid incorporation.

Meanwhile, it can enhance both scaffold and functionality diversities, which allow phage to display peptides with unique structures and various warheads. Not only have the affinity and selectivity of peptides been improved, but also the proteolytic stability is remarkably boosted.

3.2. Sortase A as a target protein

Sortase A is a protease-transpeptidase that anchors LPXTG-containing proteins to lipid II on the cell wall of Gram-positive bacteria (Figure 3-6a, b). Its substrates play many important roles in bacterial infection such as adhesion, immune evasion, blood clotting and

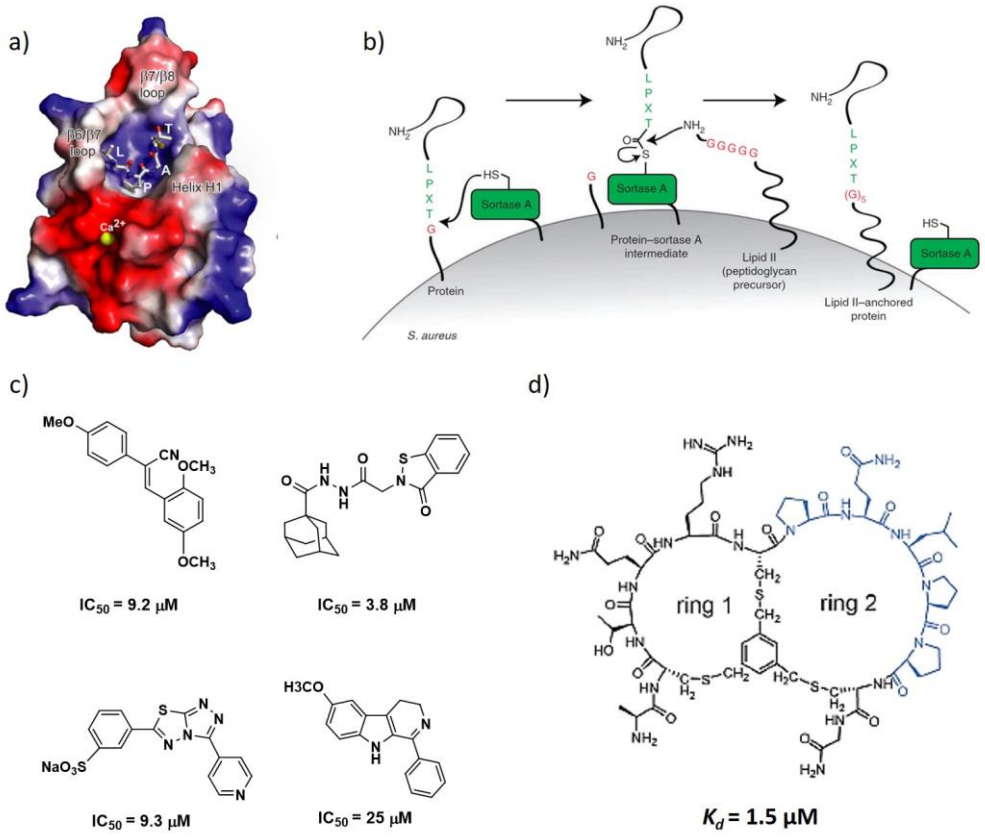


Figure 3-6. Structures of *S. aureus* Sortase A and its inhibitors. a) Structure of the covalent $SrtA_{\Delta N59}$-LPAT* complex. Adapted from Ref. [2]. b) Schematic representation of how proteins are anchored to the cell wall via sortase in Gram-positive bacteria. Adapted with permission from Ref. [4]. c) Structures and IC_{50} of some small molecule SrtA inhibitors. d) Structure and K_d of a bicyclic peptide inhibitor of SrtA.

mammalian cell invasion[52]. SrtA inhibitors even though don't affect bacterial growth, they can prevent infections cause by Gram-positive bacteria like *Staphylococcus aureus*.

A variety of small molecule inhibitors of SrtA have been developed and summarized in the reference [53] (Figure 3-6c). However, most of the current compounds are not potent enough, giving high micromolar IC_{50}. Due to the large and open catalytic site, it is challenging to identify small molecule drugs with high potency. Potent and selective inhibitors of SrtA were identified from the bicyclic phage-encoded peptide library. The peptides contain the consensus motif Leu-Pro-Pro that is similar to the natural substrates of SrtA and can inhibit SrtA activity at low micromolar (Figure 3-6d)[54]. The macrocyclic peptides are believed to bind to the active site of the target through larger surface area and therefore have higher affinity than small molecules do. This work indicates that SrtA is amenable to peptide-based inhibitors and can serve as a model protein to study the interactions between peptide and proteins, from which we could learn the principles of peptide-base inhibitors design.

3.2.1. Preparation of SrtA

The expression plasmid (pET28a-SrtAdelta59) for N-terminal his tagged *S. aureus* SrtA delta 59 was obtained from Addgene and transformed into BL21(DE3) competent *E. coli* for overexpression. The protein was purified on a Ni-NTA affinity column and desalted with a PD-10 column. The purity of the protein was confirmed by SDS-PAGE and LC-MS analysis (Figure 3-7). About 20 mg protein per liter culture was obtained.

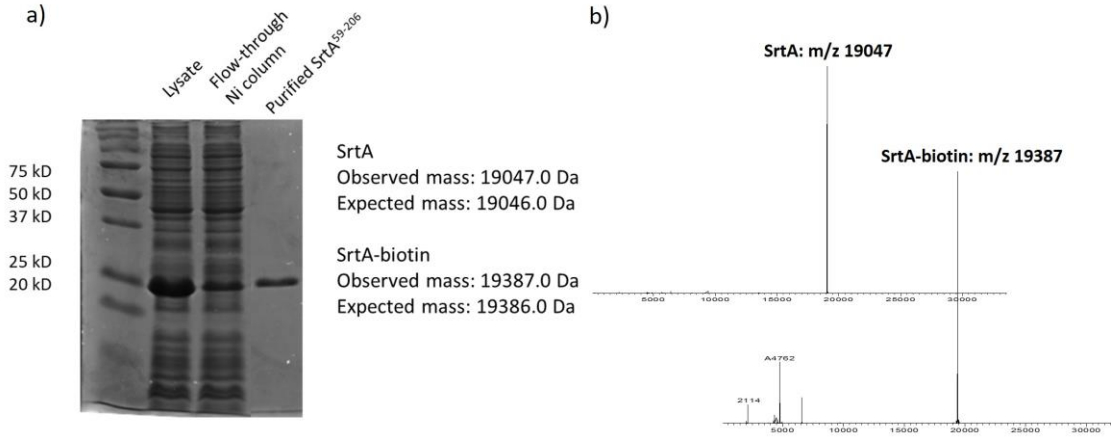


Figure 3-7. SDS-PAGE and ESI-MS analysis of pure SrtA. a) SDS-PAGE (15%) of crude cell lysate, Ni column flow-through and purified SrtA. b) Deconvoluted spectra of SrtA and SrtA-biotin.

To immobilize the target onto the streptavidin beads for screening, SrtA was biotinylated with 1.5 equivalents EZ-link™ Sulfo-NHS-LC-biotin in PBS pH 6.5. By controlling the equivalent and lowering the pH, SrtA was only labeled by one biotin motif which left the most lysine residues unmodified and could be targeted by APBA warheads of peptides. The biotinylation was monitored by LC-MS showing SrtA with one biotin was the major product (Figure 3-7b).

3.2.2. *Enzymatic activity assay*

The activity of SrtA was determined by an enzymatic assay using the fluorescence-quenched substrate Dabcyl-LPETG-Edans. Cleavage of this FRET substrate by SrtA generates the fluorescent signal (Abs/Em = 340/490 nm). Hence, we can determine the initial rate (V_0) of SrtA-catalyzed reaction, which increases with the concentration of substrate (Figure 3-8a). The inhibition activity of a small molecule inhibitor of SrtA with (Z)-Diarylacrylonitrile structure, compound DMMA, was determined to confirm our assay (Figure 3-8b). Increasing the concentration of DMMA resulted in the decrease of the initial

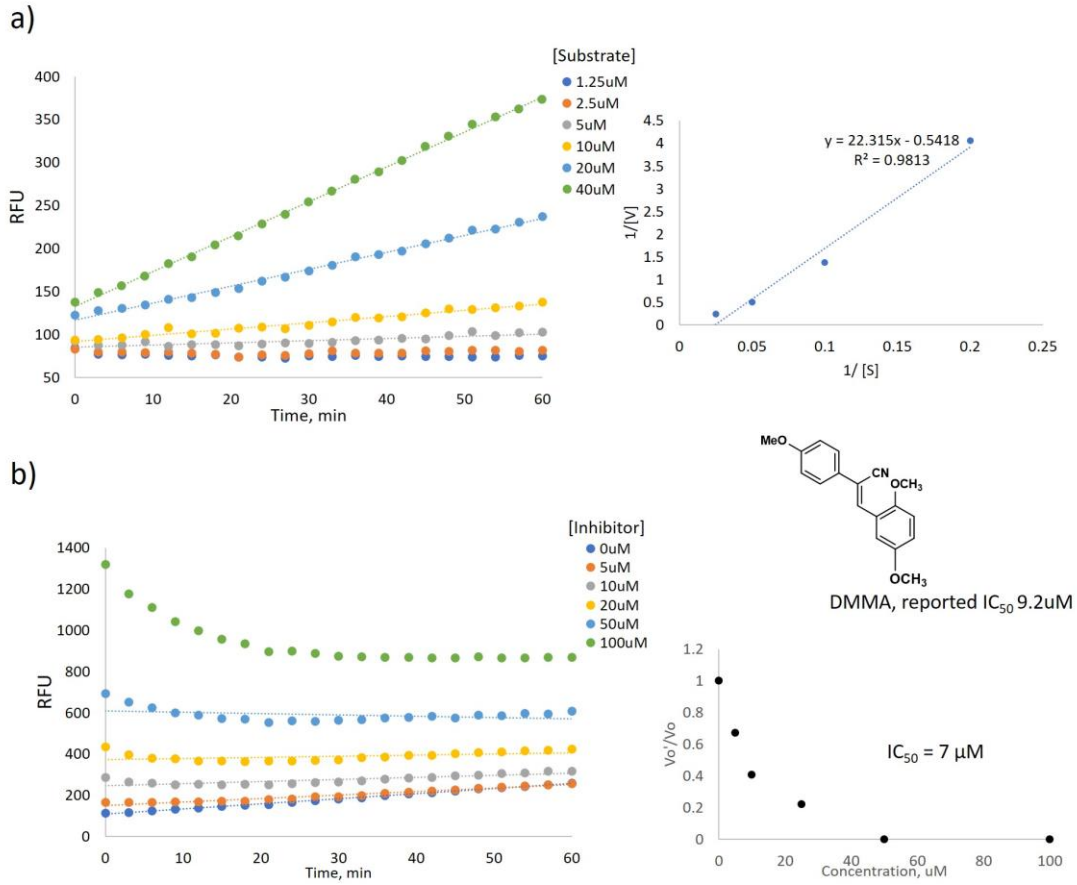


Figure 3-8. Enzymatic activity assay of SrtA. a) Initial rates of SrtA with different concentrations of the substrate. b) Inhibition assay of DMMA.

rate (V_0'). The IC_{50} was calculated as 7 µM that is consistent with the reported IC_{50} of 9.2 µM[55]. Overall, we have developed the enzymatic assay to determine the activity of SrtA.

3.3. Screening of Ph.D.™-C7C phage display peptide library against sortase A

3.3.1. Screening platform

We used Ph.D.™-C7C Phage Library from NEB, which is a combinatorial library of peptides with 7 randomized amino acid residues in a disulfide constrained loop, to screen against SrtA. The screening platform is shown in Figure 3-9a. The biotinylated protein target was immobilized on the streptavidin magnetic beads and incubated with the phage library (10^{11} phage particles). Weak and unspecific binders were washed away by vigorous washing. The strong binders were eluted by acidic buffer and amplified for another rounds

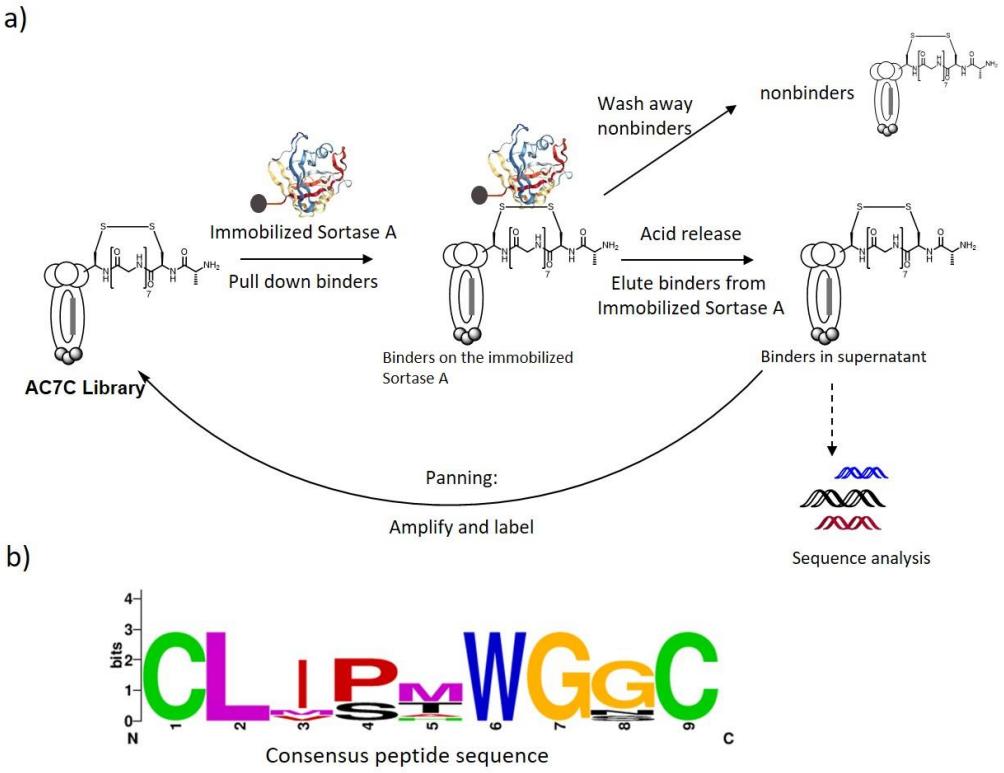


Figure 3-9. AC7C phage library screening against SrtA. a) Schematic representation of panning protocol. b) Consensus peptide sequence showed up after 3 rounds of panning.

of panning. To avoid the binders of streptavidin, a negative screening was done starting from the second round, in which the phage was incubated with the streptavidin beads and the phage in the supernatant was collected to screen against SrtA. In total, three rounds of panning were done. Phage titer after each round was determined and there was an increase in the output phage which indicated the enrichment of strong binders (Table 3-1). We randomly sequenced some phage clones after the second and third round. Interestingly, the sequencing results showed consensus peptide sequence, $CL^{I}/_{M}{}^{P}/_{S}XWGXC$ (Figure 3-9b, Table 3-2). Especially, the sequence ACLIPTWGGC showed up multiple times indicating it could be a real hit.

Table 3-1. Phage titer after each round of AC7C library panning against SrtA (pfu).

1st	2nd	3rd
$4*10^5$	$2*10^6$	$5*10^6$

Table 3-2. Sequencing analysis of output phage after the 2nd and 3rd round panning.

2nd round
CLIPTWGGC
CLIPTWGGC
CLVSMWGGC
3 blanks

3rd round
CLISMWGGC
CLMPMWGGC
CLIPHWGGC
CLIPMWGNC
CLIPTWGGC
CLISMWGGC
CLIPAWGSC
1 blank

3.3.2. Hit validation

Peptide with the sequence ACLIPTWGGCGGGDap* (W7) was synthesized to validate its binding against SrtA. The Dap residue was labeled with fluorescein and cysteine residues were oxidized to form the disulfide bond. The binding was determined by fluorescence polarization (FP) assay. From the results, we could see that increasing the concentration of SrtA led to the increase in FP and K_d was calculated as 1.5 µM (Figure 3-10a). In contrast, the addition of BSA didn't increase FP indicating that W7 couldn't bind BSA. Encouraged by its high affinity, we assessed its inhibition activity. SrtA activity was measured with increasing concentrations of nonfluorescent W7 (ACLIPTWGGC), which inhibited SrtA with IC_{50} of 3.4 µM (Figure 3-10b). The IC_{50} is close to its K_d, which further confirms that the peptide binds to the active site of SrtA.

In summary, a cyclic peptide inhibitor of SrtA was identified from AC7C phage library. The peptide showed high affinity and selectivity. The K_d of 1.5 µM makes W7 one of the most potent binders of SrtA so far, which could provide a good starting point to design more potent SrtA inhibitors.

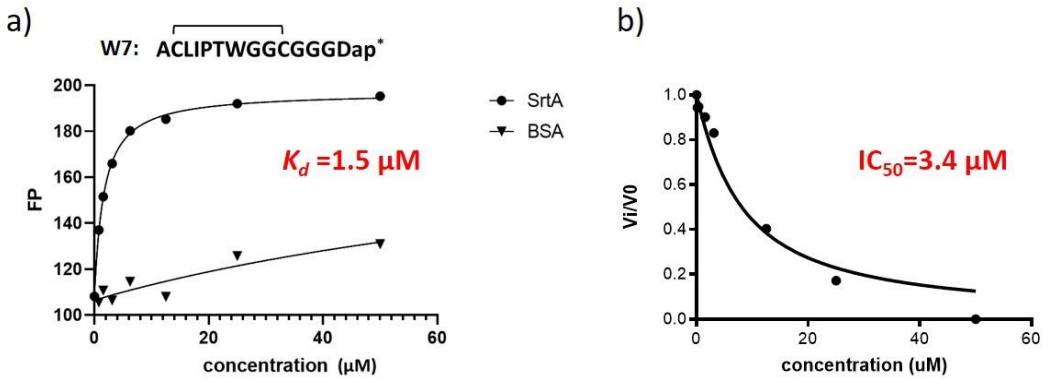
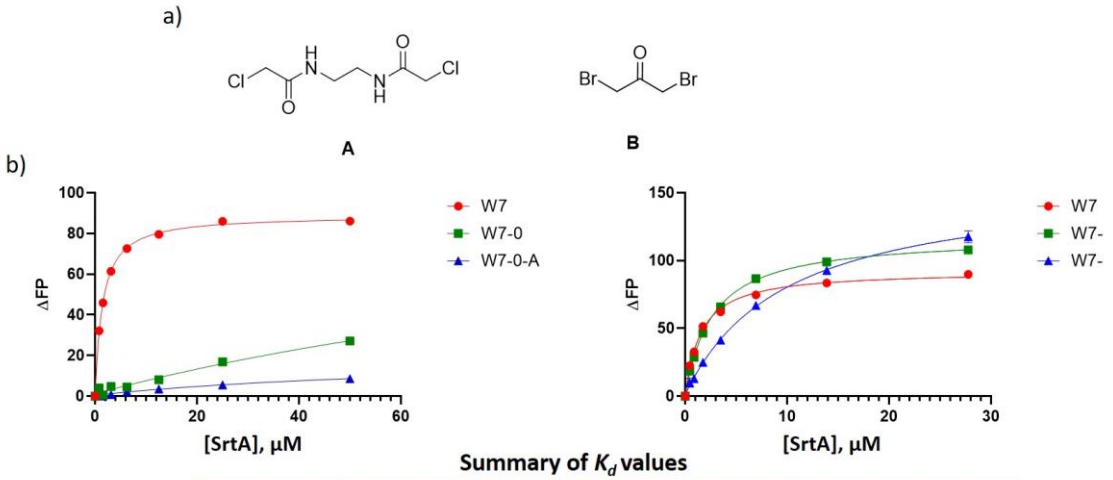


Figure 3-10. Binding affinity of W7. a) FP analysis of W7 and SrtA revealed the K_d of 1.5 µM. b) Enzymatic activity assay showed that W7 can inhibit SrtA with IC_{50} of 3.4 µM.

3.3.3. *Structure-activity relation of SrtA peptide binder*

To better understand the binding mode of W7, we synthesized a series of peptides to study its structure activity relationship. First, we removed one and two Gly residues close to the C terminus of W7, yielding W7-1 and W7-0 respectively. We presumed that Gly residues might not provide much binding affinity. From the FP results (Figure 3-11) we can see that W7-1 still had high binding affinity, however, W7-0 couldn't bind SrtA anymore. Gly might be important to retain the ring size and conformation of W7. Next, we chose two cross-linkers A and B (Figure 3-11a) to modify W7-0 and W7-1 respectively to test whether the disulfide bond can be replaced. Even though W7-0-A and W7-1-B have similar ring size to W7, they didn't bind SrtA as well as W7. W7-0-A didn't show any binding and W7-1-B had K_d of 9.6 μM, 6 times larger than that of W7.

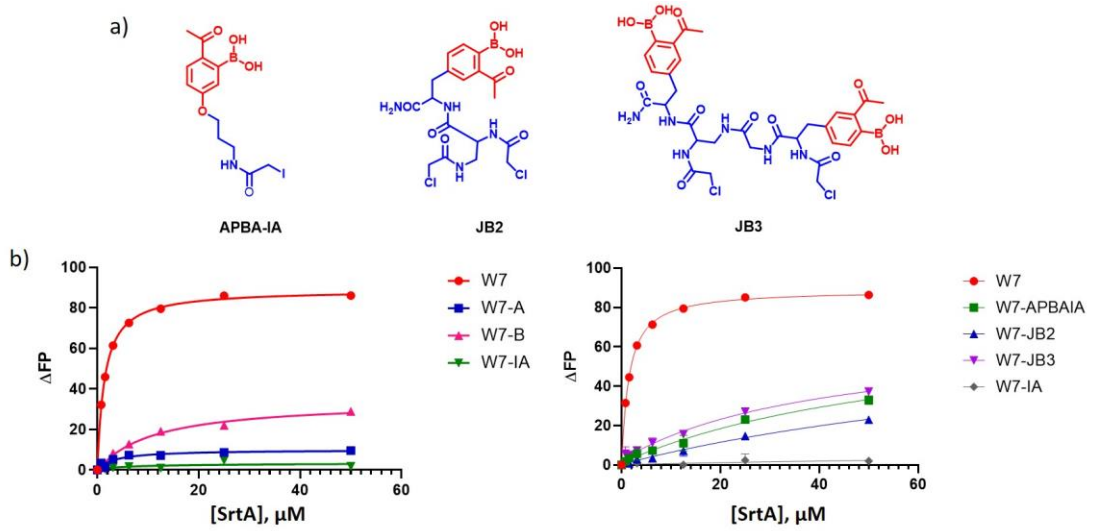


Summary of K_d values

Name	Sequence	Modifier	K_d (μM)
W7-0	ACLIPTWC	Disulfide bond	161.6
W7-0-A	ACLIPTWC	Crosslinker A	N/A
W7-1	ACLIPTWGC	Disulfide bond	2.6
W7-1-B	ACLIPTWGC	Crosslinker B	9.8

Figure 3-11. Binding affinity of W7-0 and W7-1 derivatives. a) Structures of the crosslinkers. b) FP analysis and K_d values of W7-0 and W7-1 derivatives.

We believe that the disulfide bond can constrain the peptide conformation that is crucial for binding. This is further confirmed by the linear peptide W7-IA whose cysteine residues were alkylated by iodoacetamide (Figure 3-12). Without the cyclic structure, W7-IA didn't show any binding to SrtA. Similarly, when W7 was modified by cross-linker A, the increase of FP was not obvious, indicating that the peptide couldn't bind SrtA as well as before (Figure 3-12). The shorter cross-linker B made the peptide W7-B to bind SrtA to some extent but the affinity was significantly lower.



Summary of K_d values

Name	Sequence	Modifier	K_d (μM)
W7	ACLIPTWGGC	Disulfide bond	1.4
W7-A	ACLIPTWGGC	Crosslinker A	N/A
W7-B	ACLIPTWGGC	Crosslinker B	10.75 (ΔFP is small)
W7-JB2	ACLIPTWGGC	JB2	112.9
W7-JB3	ACLIPTWGGC	JB3	41.65
W7-APBAIA	ACLIPTWGGC	Linear, APBA dimer	64.52
W7-IA	ACLIPTWGGC	Linear, IA	N/A

Figure 3-12. Binding affinity of W7 derivatives. a) Structures of the APBA presenting modifiers. b) FP analysis and K_d values of W7 derivatives.

We were wondering whether the incorporation of the APBA warhead into the peptide can further enhance the binding. To test this hypothesis, we used APBA-IA, JB2 and JB3 to modify the peptide (Figure 3-12a). APBA-IA and JB3 labeled the peptide with two APBA warheads but made it more flexible due to the linear structure and large ring size. JB2 only installed one APBA warhead and extended the peptide structure in the way similar to cross-linker A. As shown in Figure b, W7-JB3 showed the best binding affinity among these peptides with K_d of 42 μM that is still much lower than W7. These results led us to draw the conclusion that even though APBA could help the peptide bind to SrtA through forming iminoboronate with lysine residues, peptide itself lost affinity to the binding pocket.

In summary, Ph.D.^{TM}-C7C phage library allows us to identify a high affinity peptide inhibitor of SrtA. The motif LIPTWG is important for binding together with the conformation formed by the disulfide bond. When the disulfide bond was opened up and modified by cross-linkers or APBA presenting modifiers, the binding affinity dropped significantly. We could further rationally design peptide inhibitors of SrtA based on these results.

3.4. Screening of APBA presenting phage display peptide libraries

As mentioned before, we reported the phage display of dynamic covalent binding motifs, APBA moieties, for targeting the bacterial cell surface[47, 48]. The APBA-IA molecule was synthesized to modify the reduced cysteine residues of phage-encoded peptides to yield an APBA dimer library (Figure 3-13a). In this work we explore the utility of APBA presenting phage display peptide libraries for targeting the surface of proteins.

3.4.1. Targeting the surface of Sortase A by APBA dimer library

SrtA presents 19 surface lysine residues out of a total of ~200 (Figure 3-13b). The high abundance of lysine bodes well for our iminoboronate-based targeting strategy with the

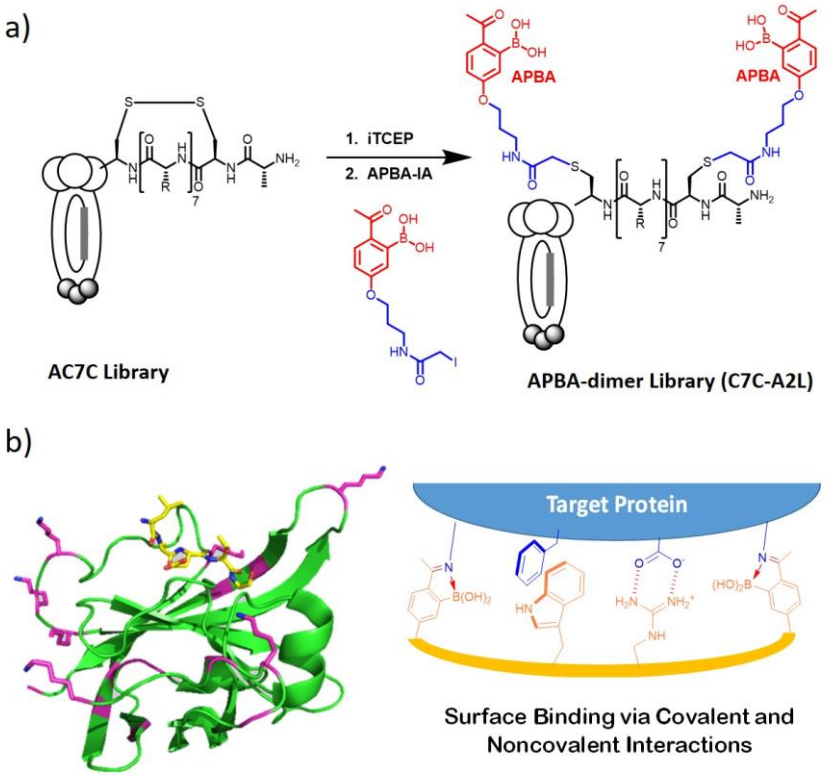


Figure 3-13. Targeting the surface of Sortase A by APBA dimer library. a) Illustraion of APBA-dimer library (C7C-A2L). b) Demonstration of lysine residues on the surface of SrtA and iminoboronate formation between lysine side chains and APBA presenting peptides.

APBA-dimer library, namely C7C-A2L. AC7C library was reduced by TCEP and modified by APBA-IA. As a negative control, the same library was modified by iodoacetamide (IA). Phage panning was carried out following the same protocol as described above for AC7C library screening against SrtA, except that streptavidin agarose beads (Pierce^{TM}) were used in this case. The phage titer was monitored after each step and shown in Table 3-3. The input population was around 10^{10} to 10^{11} pfu and the output population was typically on the scale of 10^4 to 10^5 pfu. Starting from the second round, negative screening was done to remove the nonspecific binder of streptavidin beads by incubating the phage with beads

Table 3-3. Phage titer after each step of C7C-A2L library panning against SrtA (pfu).

Modifier	1st round input	1st round output	2nd round input	2nd round after negative screening	2nd round output	3rd round input	3rd round after negative screening	3rd round output
APBA-IA	10^{10}	$1.5*10^5$	$5*10^8$	$2.5*10^8$	$1.2*10^4$	$6*10^{10}$	$6*10^9$	$1.2*10^5$
IA	10^{10}	$1.2*10^5$	$2.5*10^8$	$1.2*10^8$	$1.2*10^4$	$6*10^{10}$	$9*10^9$	$1.2*10^5$

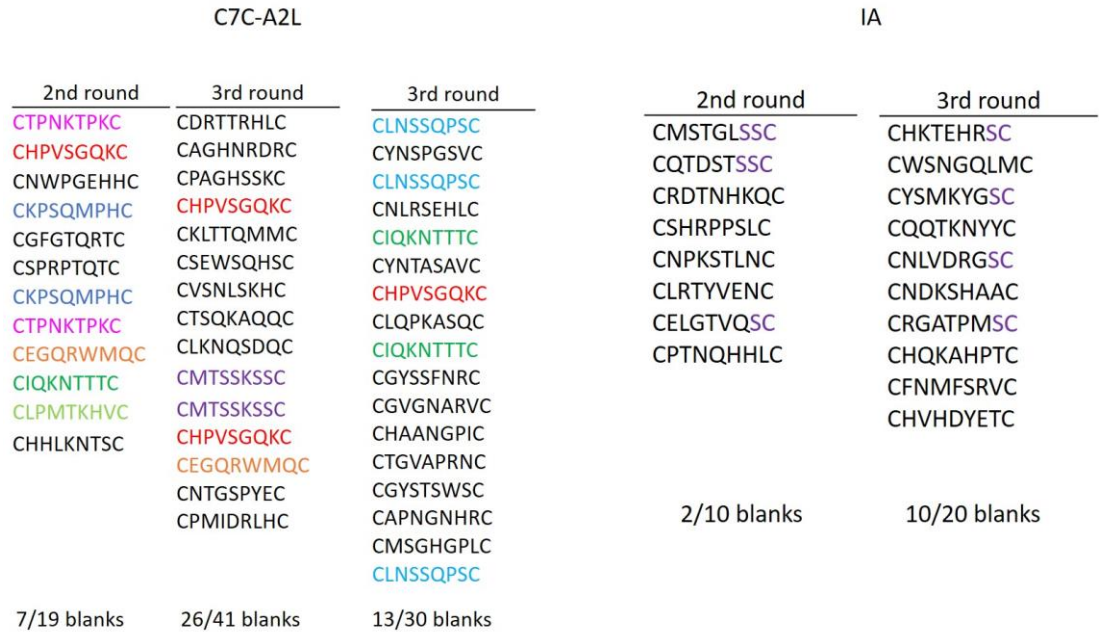


Figure 3-14. Sequencing analysis of C7C-A2L and control library (IA).

not coated with SrtA. Random phage clones were sequenced which showed convergence after round 2. After round 3, more repeating sequences showed up (Figure 3-14). In contrast, three rounds of panning of the control library yielded no recurring sequences at all, indicating no potent binders of SrtA exist in this control library.

Table 3-4. Peptides chosen for validation

Peptides	Sequence	Repeating times
W1	$AC_mHPVSGQKC_mGGGDap$-FAM	4
W2	$AC_mLNSSQPSC_mGGGDap$-FAM	3
W3	$AC_mIQKNTTTC_mGGGDap$-FAM	3
W4	$AC_mTPNKTPKC_mGGGDap$-FAM	2

3.4.2. Hit validation

The four peptide hits (W1-4, Table 3-4) that appeared the most times were chosen for validation. The peptides were synthesized and labeled with fluorophore. The cysteine residues were modified with APBA-IA or IA as negative control. To determine the binding between SrtA and the peptides, the fluorescently labeled peptides were incubated with SrtA-coated streptavidin beads (SrtA +). Under a fluorescence microscope, W2 at 10 μM showed strong fluorescence staining of the SrtA-coated beads, while other peptides gave marginal fluorescence of the beads (Figure 3-15a). Therefore, we focused on W2 in the following study. Importantly, W2 afforded little binding to the beads without SrtA coating, indicating the peptide selectively bound SrtA instead of streptavidin. Remarkably, the addition of 10 mg/mL BSA didn't affect the binding (Figure 3-15b). Furthermore, the negative control peptide W2-IA elicited no binding to SrtA coated beads under the same conditions, which highlighted the necessity of the APBA moieties (Figure 3-15c). This is

consistent with the fact that no peptide hits were identified from screening of the IA modified library.

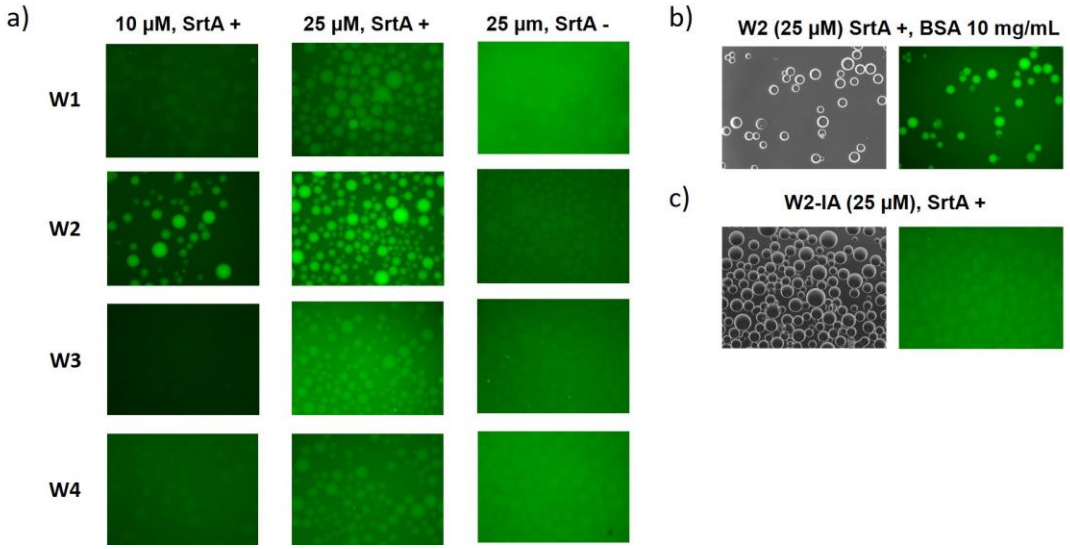


Figure 3-15. Characterization of the peptides' binding to SrtA via fluorescence imaging of SrtA-coated beads. a) Fluorescent imaging of W1-4 staining SrtA coated beads and streptavidin beads. b) Addition of 10 mg/mL BSA didn't inhibit the W2's binding to SrtA. c) W2-IA couldn't stain SrtA coated beads.

To further quantify the binding potency, we examined W2 of varied concentrations for staining SrtA-coated beads (Figure 3-16a). Plotting the mean fluorescence intensity of the beads against peptide concentration gave the binding curves, fitting which yielded the K_d value of 14 μM (Figure 3-16b).

The peptide-SrtA binding was further analyzed using fluorescence polarization, for which SrtA of increasing concentrations was added to W2 solution and the fluorescence anisotropy was recorded on a fluorimeter. Fitting the binding curves generated through fluorescence polarization experiments gave K_d values of 12 μM, consistent with the result from the microscopy experiments (Figure 3-17a). However, the increase in FP is small so the real K_d might be larger than that. The false positive W3 as well as the control peptide

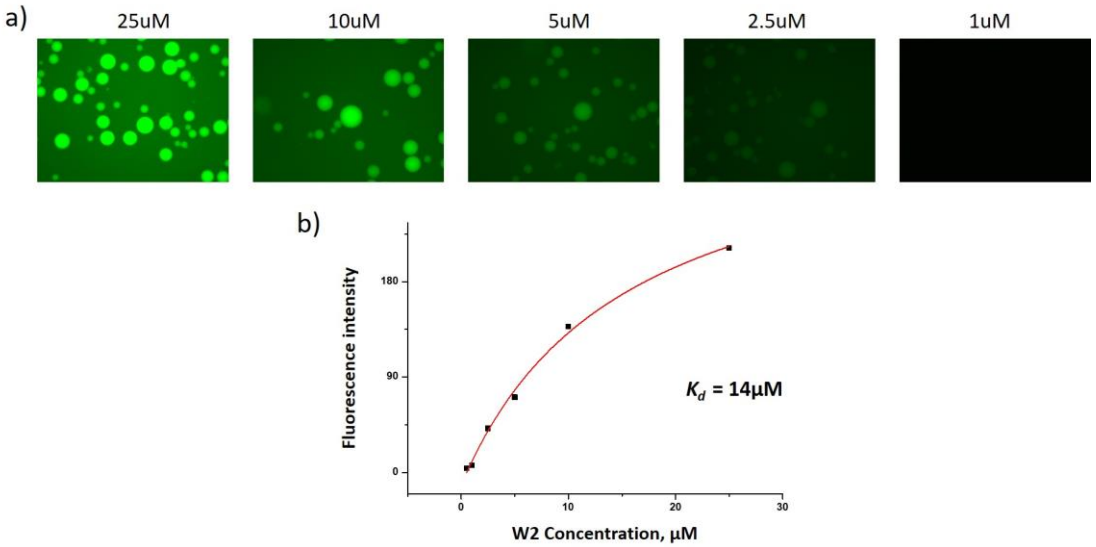


Figure 3-16. Quantification of the binding potency by plotting the mean fluorescence intensity of the beads against peptide concentration. a) Fluorescent imaging of W2 staining SrtA coated beads at different concentration. b) The binding curve generated by plotting the mean fluorescence intensity of the beads against peptide concentration.

W2-IA showed little binding (Figure 3-17a, b). In addition, the specificity of W2 was investigated by titrating in other proteins including BSA and lysozyme. W2 could bind BSA with K_d of 20 µM but didn't show much binding to lysozyme. This is not surprising considering that BSA displays about 60 lysine residues on its surface that could interact with the APBA moieties of W2. However, from the fluorescent imaging experiments we could see that 10 mg/mL had little inhibition of W2's binding to SrtA.

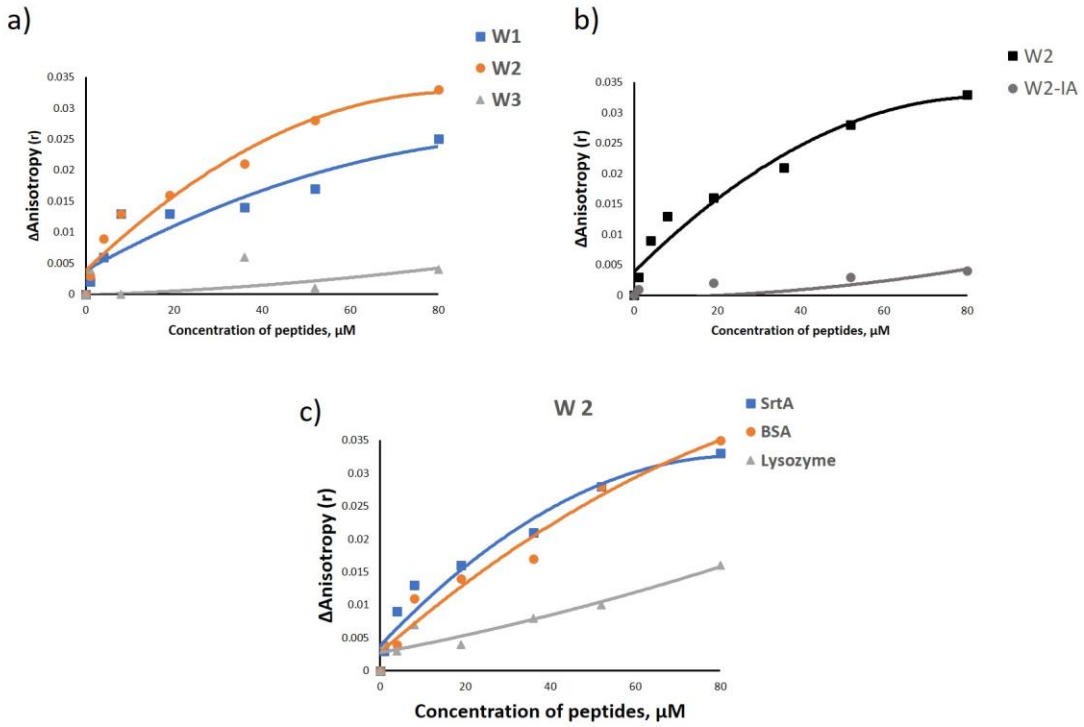


Figure 3-17. Binding affinity determined by fluorescence polarization. a) Binding curve of W1-3 against SrtA. b) Comparison of the binding curves of W2 and W2-IA. c) Binding curves of W2 against SrtA, BSA and lysozyme.

3.4.3. Peptide's binding site

To probe the peptides' binding site on SrtA, we performed a competition experiment with the inhibitor of SrtA, DMMA (Figure 3-18a). Interestingly, DMMA addition elicited no inhibition of the SrtA binding by W2. These results suggest that instead of binding to the active site, W2 probably binds to the surface of SrtA. This is further confirmed by the enzymatic assay, in which W2 showed little inhibition of SrtA activity even at hundred micromolar concentration (Figure 3-18b).

To better understand the binding mode of W2, we expressed ^{15}N labeled SrtA and carried out ^{1}H-^{15}N HSQC protein NMR experiments. We presumed that the binding of W2 would

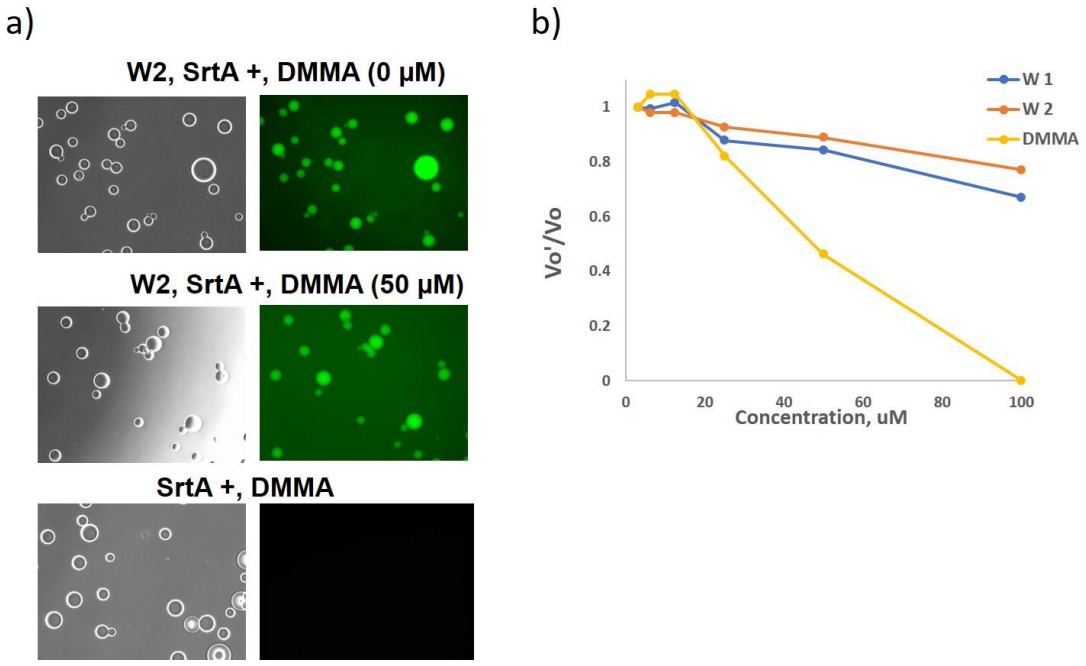


Figure 3-18. Probing the binding site of W2. a) Fluorescent imaging of SrtA beads strained by W2 with DMMA or not. b) Enzymatic activity of SrtA in the presence of W1-2 and DMMA. W2 couldn't inhibit SrtA's activity.

make the peaks of amino acids around the binding site shift on the spectrum. To test our hypothesis, SrtA was first incubated with 2APBA (2 mM) and we could see some peaks shift, indicating the iminoboronate formation can be detected by HSQC experiment (Figure 3-19a). Then, nonfluorescent W2 (2 mM) was incubated with SrtA for NMR study. By overlapping the spectra, we could see that some peaks shifted when W2 and APBA were added and some peaks only shifted in the presence of W2 (Figure 3-19b). However, the negative control peptide W2-IA (2 mM) also led to the peak shift under the same conditions and the spectra of W2 and W2-IA almost overlapped (Figure 3-19c). These results indicated that the peak shifts were due to the nonspecific interactions with the peptides that were at millimolar concentration. We tried to decrease the concentration of peptides to micromolar but it resulted in no shift on the spectra. This could be attribute to the low binding affinity of W2 that has the K_d in the medium micromolar range.

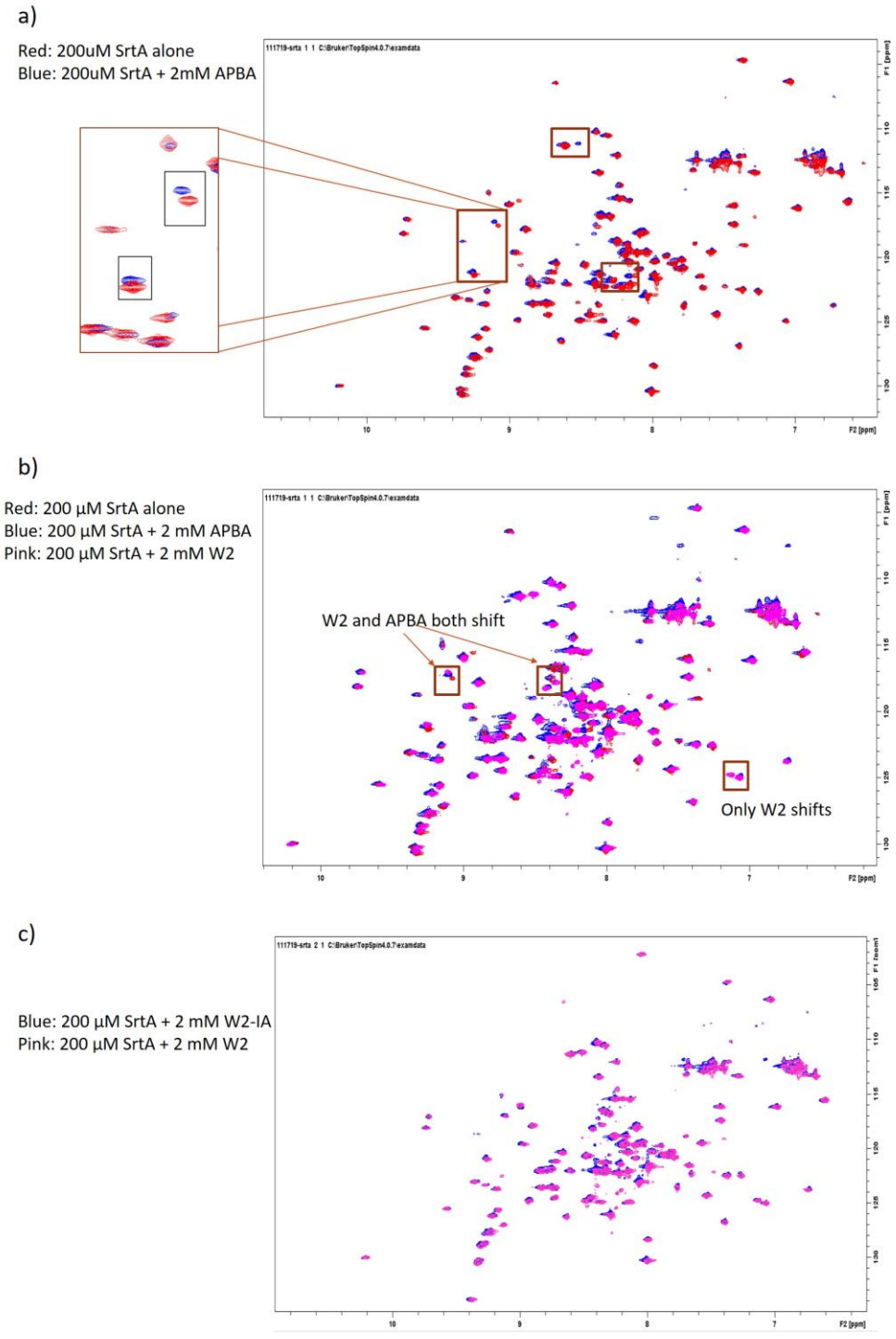


Figure 3-19. ^{1}H-^{15}N HSQC protein NMR experiments. The experiments were done in the NMR buffer containing 20 mM sodium acetate, 50 mM NaCl, and 10 mM $CaCl_2$, pH 6.0, 10% D_2O, 2% DMSO-d_6. a) The overlay of spectra of ^{15}N-SrtA alone and ^{15}N-SrtA and APBA mixture. b) The overlay of spectra of ^{15}N-SrtA, ^{15}N-SrtA and W2 mixture, and ^{15}N-SrtA and APBA mixture. c) The overlay of spectra of mixtures of ^{15}N-SrtA and W2 or W2-IA.

The lysine residues that are prone to react with APBA moieties were also investigated by LC/MS and LC/LC-MS/MS. SrtA was incubated with APBA (2 mM) and $NaBH_4$ (100 mM) was added to reduce the imine bond and yield a stable conjugate. The reaction was first analyzed by LC/MS, which showed about 30% SrtA conjugated with one APBA molecule (Figure 3-20a). Then the mixture was subjected to dialysis, denaturation, trypsin digestion and LC/LC-/MS/MS analysis that revealed 6 lysine residues out of 19 modified by APBA (figure 3-20b). Taken K34 as an example, the fragmentation (MS2) spectrum demonstrated clear masses corresponding to the APBA labeled peptide fragments (Figure 3-20c). These results gave us a hint of which lysine residues are more likely to be modified by APBA moiety (Figure 3-20d).

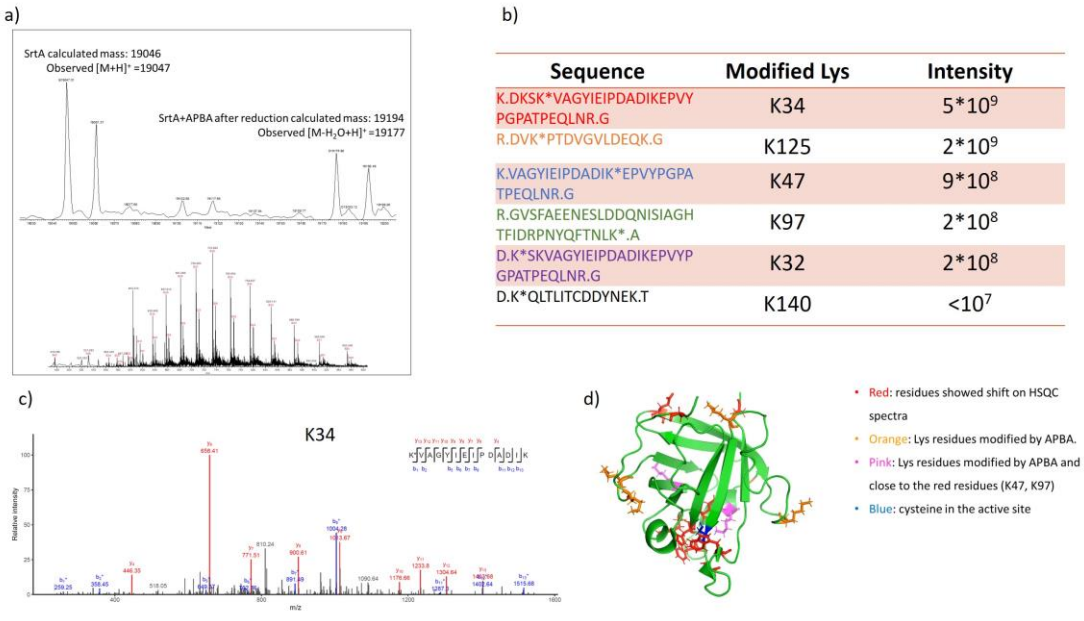


Figure 3-20. LC-MS/MS analysis of SrtA-APBA conjugates. a) Deconvoluted (top) and raw (bottom) ESI-MS spectra of SrtA-APBA conjugates. b) Lysine residues modified with APBA identified by MS/MS analysis. c) Fragmentation (MS2) spectrum revealed the modified K34. d) Demonstration of lysine residues of SrtA that are likely to react with APBA and W2.

3.4.4. Comparison of magnetic and agarose beads as solid support

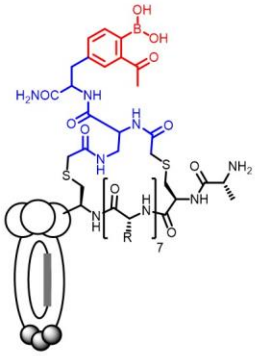
The choice of solid support is important to the success of a screening. We used the streptavidin agarose beads (Pierce™) when screening A7C-A2L against SrtA and found that due to the high biotin capacity a large amount of biotinylated SrtA was needed to saturate the beads. For example, about 50 μg of SrtA was added to coat 50 μL agarose beads, which resulted in 5 μM as the final concentration of SrtA during panning (Table 3-5). In contrast, the capacity of streptavidin magnetic beads (Dynabeads™) is about 10-fold smaller, and therefore the concentration of SrtA during panning is around 0.5 μM that is more closed to the value reported in other's protocol[56]. We suspect that the high concentration of the target could allow weak binders to be enriched. Especially, the weak binders with growth advantages may take over the strong binders. Lowering the concentration of target means increasing the stringency of screening, which enables the high affinity binders to be identified. Additionally, another advantage of the magnetic beads is that they are easy to be separated from the solution by a strong magnet, which makes the washing step more efficient with little loss of the beads. Therefore, we prefer the magnetic beads to the agarose beads in the following study.

Table 3-5. Comparison of magnetic and agarose beads

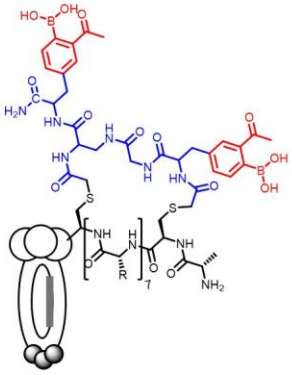
	Pierce™ Agarose	**Dynabeads™ magnetic**
Separation method	Centrifuge	Magnet
Size	45 to 165 μm	2.8 μm
Biotin capacity	1 to 3 mg biotinylated BSA/mL resin	0.1 mg biotinylated antibody/ mL beads
Target concentration	50 μL beads (50 μg SrtA) in 500 μL = 5 μM	50 μL beads (5 μg SrtA) in 500 μL =0.5 μM

3.5. Targeting the surface of Sortase A by cyclic peptide libraries presenting APBA

The peptide cyclization leads to more compact peptide structures and can minimize unfavorable entropy penalty upon binding to the targets. We presume that by cyclizing the phage-encode peptides we could identity more potent binders of SrtA. Towards this goal, JB2 and JB3 were used to cyclize the phage-encoded peptides and meanwhile incorporate APBA moieties (Figure 3-21). Ph.D.^{TM}-C7C phage library was reduced by TCEP and modified by JB2 and JB3 to generate library C7C-A1C and C7C-A2C respectively in the



Mono-APBA cyclic peptides
(C7C-A1C)



Bi-APBA cyclic peptides
(C7C-A2C)

C7C-A1C

2nd Round		3rd Round	
CLLSELMGC	CTTWAGLSC	CHGSNSHTC	CSHNLSHLC
CTPNKFARC	CVDMWTDNC	CTGGTQSSC	CWPPSSQWC
CSRSMDSTC	CSAHMSTNC	CGTNPIKKC	CTKTSSNLC
CDGRPDRAC	CQLTVAHQC	CTDKASSSC	CEKRVGTQC
CTDKASSSC	CPFDGPKVC	CEQRFLAHC	CREPTYNQC
CQWPLVKNC		CFGTNHRPC	CKHSSLKLC

C7C-A2C

2nd Round	3rd Round
CHMYHNATC	CHMYHNATC
CTTRPTIPC	CDGRPDRAC
CTPHNTSSC	CDGRPDRAC
CGHGDFVLC	CDGRPDRAC
CLMT*RSTC	
CDGRPDRAC	

Figure 3-21. Illusion of C7C-A1C and C7C-A2C libraries and their sequencing results.

Table 3-6. Phage titer C7C-A1C and C7C-A2C library panning against SrtA (pfu).

	1st input	1st output	2nd input	2nd output	3rd input	3rd output
C7C-A1C	10^{10}pfu	$7*10^{4}$pfu	10^{10}pfu	$7*10^{6}$pfu	10^{10}pfu	$5*10^{5}$fu
C7C-A2C	10^{10}pfu	$3*10^{6}$pfu	10^{10}pfu	$2*10^{7}$pfu	10^{10}pfu	$2*10^{7}$pfu

same way to generate C7C-A2L library. Screening against SrtA was done following the previous protocol using streptavidin magnetic beads as the solid support. The phage titer of each round's input and output was listed in Table 3-6. Interestingly, C7C-A2C library generally showed a larger number of phages in the output of each round than C7C-A1C and A2L libraries did. However, the sequencing analysis of the round 2 and 3 output phage revealed that no obvious consensus motif was obtained. The phage with the sequence of CDGRPDRAC or CTDKASSSC also showed up multiple times in the screening against other targets and believed as "parasitic phages" that were enriched due to the amplification bias. Unfortunately, we couldn't identify any strong binder of SrtA from these two libraries.

To improve the library, one direction is the optimization of the length of peptides. Seven amino acids in between a pair of cysteine only give a nonapeptide that might be too short to bind the target. Therefore, we'll construct libraries of longer peptides, like C8C, C9C and C10C. On the other hand, we'd like to design more rigid cross-linkers which can improve the rigidity of peptides. This is consistent with the concept of double bridged peptides library, in which the conformation of the peptide can be stabilized by the cross-linkers. Also, the effort to reduce the amplification bias is valuable through optimizing the amplification time and conditions.

3.6. Conclusions

In summary, we demonstrated that peptide binders of SrtA were identified from the peptide library displayed by phage. From the commercially available Ph.D.^{TM} C7C phage library, W7 was identified and possessed single-digit micromolar affinity to SrtA. The peptide can inhibit the enzymatic activity of SrtA, which confirms it's binding to the active site. Furthermore, the structure-activity relationship study of W7 shows that the compact cyclic structure is critical to its binding. The peptide probably adopts a certain conformation in the binding site, which enables the side chains of consensus amino acids (Leu, Pro and Trp) to interact with the target.

We also applied the phage libraries presenting covalent warheads, APBA, to screening against SrtA. From the linear peptide library C7C-A2L, W2 was identified showing medium micromolar affinity to SrtA. Interestingly, instead of binding to the active site, W2 is more likely to anchor on the surface of SrtA. This is consistent with the binding mechanism of APBA that reacts with the lysine residue on the surface of protein. The micromolar affinity exhibited by the W2 although needs to be improved, showcases the potential of specific protein recognition by targeting protein surfaces, which provides a great opportunity to modulate the undruggable targets.

3.7. Experimental Procedures

3.7.1 General Methods

Rink Amide resin, Fmoc-Osu and HBTU were purchased from Novabiochem. Fmoc-protected amino acids were purchased from Advanced Chemtech or Chem-Impex Int'l Inc. Trifluoroacetic acid was purchased from Protein Technologies. PD 10 Desalting columns and Amersham NAP-5 columns were obtained from GE Healthcare and USA Scientific respectively for desalting and buffer exchange. Bacterial Sortase Substrate I, FRET was purchased from AnaSpec, Inc. Other chemicals were obtained from Fisher Scientific or Sigma Aldrich. Expression plasmid for *S. aureus* Sortase A (SrtA) was purchased from Addgene. GeneJET Plasmid Miniprep Kit from Thermo Fisher was used to prepare plasmids for sequencing. HisPur™ Ni-NTA Resin was obtained from Thermo Fisher Scientific. Sanger Sequencing was done by Genewiz or Eton Bioscience Inc. Peptide synthesis was carried out on a Tribute peptide synthesizer (Protein Technologies). ^{1}H-^{15}N HSQC protein NMR experiments were performed on the Bruker AVANCE NEO 800 spectrometera at Brandeis NMR Facility. A Nanodrop UV-vis spectrometer was used to measure the concentration of the fluorescein labelled compound. Mass-spec data were generated by an Agilent 6230 LC TOF mass spectrometer. HPLC purification was done on Waters 2695 Alliance HPLC system. Fluorescent polarization experiments were done on a Cary Eclipse Fluorescence Spectrophotometer from Agilent Technologies or a SpectraMax M5 plate reader from Molecular Devices.

3.7.2 Expression and purification of SrtA

Expression and purification were followed the methods described in the reference[4]. The protein plasmid was transformed into *E. coli* BL21(DE3) cells. 10 mL overnight cell

culture was inoculated into 1 L LB media with kanamycin (50 mg/mL) in the 4L flask. The cells were cultured at 37°C with 250 rpm shaking and induced at OD600 = 0.5-0.8 by the addition of isopropyl β-D-1-thiogalactopyr-anoside (IPTG, 1 mM). The cells were cultured room temperature, 250 rpm for 16h before harvested by centrifugation (5000g for 20 min) at 4 °C. The cell pellet was suspended in 20 mL of buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 10 mM imidazole) and lysed by sonication. The insoluble cell debris was removed by centrifugation (6500 rpm for 40 min followed by 12000 g for 40 min) at 4 °C. The supernatant was cleared with a 0.22 μm filter and the protein was purified by a Ni-NTA affinity column. 1.5 mL Ni-NTA resin was loaded into a column and washed with 5 mL buffer A. The protein solution was loaded onto the column. The resin was washed with 5 mL buffer B (Wash buffer contains 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 16 mM imidazole). The protein was eluted with 2.5 mL elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 250 mM imidazole). The protein was desalted by PD-10 column and finally stored in the DPBS buffer with calcium and magnesium pH 7.0. The concentration of protein was determined by *DC* Protein Assay Reagents from Bio-Rad.

To analyze the purity of the protein, 15% SDS-polyacrylamide gel was prepared. Precision Plus

Protein Unstained Standards were obtained from Bio-Rad as reference. The protein samples were diluted in 4x Laemmli Sample Buffer (Bio-Rad) and loaded onto the gel. The gel was run in the Electrophoresis Chamber (Bio-Rad) with 180 V for 30 min, stained with Coomassie Blue buffer for 30 min and destained overnight.

LC-MS analysis was used to determine the mass of the protein. 100 ng protein was subjected into the LC-MS system with Agilent Extend C18 column.

3.7.3 Biotinylation of SrtA

EZ-linke Sulfo-NHS-LC-biotin 10 mM stock was prepeared in DMF. To 0.5 mL SrtA (32 μM) solution in PBS pH 6.5, 2.5 μL of 10 mM EZ-linke Sulfo-NHS-LC-biotin stock (48 μM) was added. The mixture was incubated at 4 ºC overnight. The reaction was monitored by LC-MS. The product was purified by an Amersham NAP-5 column with DPBS buffer.

3.7.4 Enzymatic activity assay of SrtA

2.5 μM SrtA was added into each well of the 96 well plate containing 0.5 mM triglycine and varous concentrations of fluorogenic substrate (1, 2, 4, 8 and 16 μM) in the TBS buffer pH 7.4 with 2 mM $CaCl_2$. The reactions were incubated at 37 ºC. The fluorescence readings at 495 nm were obtained every 1 min for 1 h, on the plate reader with Ex 350 nm and cut off at 475 nm.

To determine the IC_{50} of SrtA inhibitors, increasing concentrations of inhibitors (0-0.1 mM) were added into the reactions containing 0.2 mM fluorogenic substrate, 0.5 mM triglycine and 2.5 μM SrtA. The fluorescence readings were obtained following the same protocol described above.

3.7.5 Chemical modification of phage

5 μL phage stock (2×10^{13} pfu/mL) was diluted in 0.2 mL NH_4HCO_3 buffer pH 8.5 and incubated with 1 mM TCEP for 1 h at room temperature. The reduction can also be done with immobilized TCEP gel (Thermo Fisher Scientic) as previously reported[48]. 1/5 volume PEG (20%)/ 2.5M NaCl was added to precipitate phage on ice for 30 min. The phage pellet was collected by centrifuge (14000g, 4 ºC, 20 min) and resuspended in 0.2 mL NH_4HCO_3 buffer pH 8.5. 1 μL chemical modifier, APBA-IA, JB2 or JB3 (200 mM

stock in DMF), was added into the phage solution. The reaction was mixed on a rotating wheel (10 rpm) at room temperature for 2 h. The excess reagents were removed by PEG/NaCl precipitation. Finally, phage was suspended in DPBS pH 7.0 buffer.

3.7.6 Panning protocols

5 μg of biotinylated SrtA was incubated with 30 μl magnetic streptavidin beads (Dynabeads^{TM}, M-280 from Invitrogen) with rotating at room temperature for 30 min. The beads were washed 3 times with DPBST buffer (0.1 % Tween in DPBS buffer) and blocked with 1 mL 2 mg/mL BSA for 1 h at room temperature. Meanwhile, phage was incubated with 2 mg/mL BSA. The blocking buffer was removed from beads and phage solution was mixed with the blocked beads mixed on a rotating wheel (10 rpm) for 30 min. Then the phage in the supernatant was removed and the beads were washed 8 times with 1 mL DPBST buffer and 2 times with 1 mL DPBS buffer. The beads were resuspended thoroughly by vortexing in each washing. The phage was eluted by incubating the beads with 1 mL glycine buffer (50 mM, pH 2.2) for 10 min and transferred to 200 μL of 1 M Tris-Cl buffer, pH 9.1. Phage titer of each round output was determined.

3.7.7 Phage amplification

1 mL phage was added into 20 mL LB medium with 2 mL overnight culture of ER 2738 (NEB). The culture was shaking at 250 rpm at 37 ºC for 4-4.5 h. The bacteria were pelleted by centrifuge (5000g, 20 min, 4 ºC). The phage was obtained from the supernatant by PEG/NaCl precipitation at 4 ºC overnight. The amplified phage was finally resuspended in 0.2 mL DPBS buffer and usually gave a concentration of 5×10^{12} pfu/mL.

3.7.8 Peptide synthesis

Peptides were synthesized on Rink Amide resin through Fmoc/tBu chemistry by the peptide synthesizer. Fmoc-Dap(Alloc)-OH was installed at the C-terminus for fluorescein labeling. For 0.025 mmol scale of peptide synthesis, the alloc group was deprotected using a mixture of 30 mg tetrakis(triphenylphosphine)-palladium (Pd(PPh3)4) and 0.2 mL phenylsilane in 2 mL dichloromethane for 1 h. 5(6)-Carboxyfluorescein (30 mg) and HBTU (28 mg) were dissolved in 1.5 mL 0.4M NMM in DMF and then mixed with the resin. Then the Fmoc group at the N-terminus was deprotected by 20% Piperidine in DMF. The peptide was cleaved from the resin with 2 mL of 90% TFA, 5%water, 2.5% triiopropylsilane and 2.5% 1,2-ethanedithiol and then precipitated by could diethyl ether. The crude peptide was dissolved in HPLC buffer (water and acetonitrile mixture) and purified by RP-HPLC. The fractions were analyzed by LC-MS and the one containing the pure peptide was lyophilized. 1-2 mg peptide was then subjected into the oxidation to form disulfide bond in $NaHCO_3$ pH 8.6 buffer with 10% DMSO overnight, or modified with 3 equivalents of APBA-IA with 5% DIPEA in $NaHCO_3$ pH 8.6 buffer. For JB2 and JB3 modification, only 1.5 equivalent reagents were used. The reactions were monitored by LC-MS and purified by RP-HPLC.

3.7.9 Fluorescence polarization experiment

75 µL of 0.4 mM fluorescein labeled peptide in DPBST (0.1% Tween) was incubated with 75 µL increasing concentrations of protein (0-100 µM) in the 96 well plate (Black opaque) for 20 min at room temperature. The fluorescence polarization readings were obtained on the plate reader or spectrophotometer with Ex 485 nm/ Em 535. The binding curve and K_d values were calculated by plotting the FP versus protein concentration and fitting the curve with the following equation in the software Graphpad Prism8:

Y=FPmin+(FPmax-FPmin)*(K_d+0.2+x^((K_d +0.2+x)^2-4*0.2*x))/(2*0.2)

in which Y is FP, x is the concentration of protein.

3.7.10 Fluorescence imaging experiment

10 μL SrtA coated agarose beads and 10 μL fluorescein labeled peptide in DPBS pH 7.0 buffer were mixed and incubated for 20 min with rotating. Then the beads were spun down by centrifuge and resuspended in 20 μL DPBS buffer of which 2 μL was quickly transferred onto the glass slide and imaged by fluorescence microscopy (fluorescein channel) at 10X magnification. The same procedure was carried out for streptavidin agarose beads. BSA and DMMA was added when beads were incubated with the peptide to study their effect on peptide's binding to SrtA.

3.7.11 ^{1}H-^{15}N HSQC protein NMR experiment

^{15}N labeled SrtA was prepared following the same protocol for normal SrtA except that *E. coli* BL21 (DE3) with SrtA expression plasmid was grown in M9 medium supplemented with $^{15}NH_4Cl$.

4.8mg ^{15}N-labeled SrtA was obtained from 100mL culture. The molecular weight was confirmed by LC-MS analysis that showed 19280 (MW calculated: 19287).

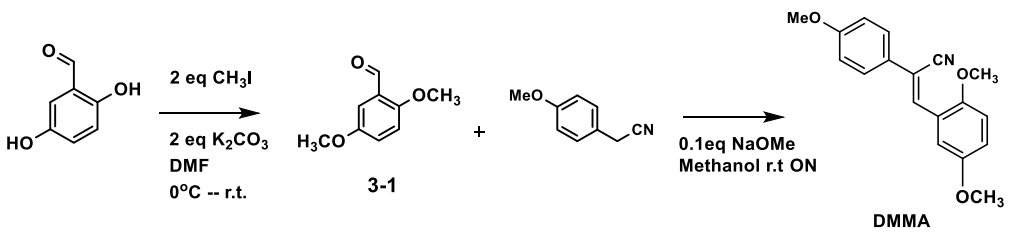
The ^{15}N-labeled SrtA was dissolved in 0.5 mL NMR buffer containing 20 mM sodium acetate, 50 mM NaCl, and 10 mM $CaCl_2$, pH 6.0, 10% D_2O. Peptides or APBA were dissolved in DMSO-d_6 and then mixed with ^{15}N-labeled SrtA resulting in 2% DMSO-d_6. ^{1}H-^{15}N HSQC experiments were done on the Bruker AVANCE NEO 800 spectrometer with a 5 mm TCI cryoprobe (1H, {13C, 15N}).

3.7.12 LC/LC-MS/MS analysis of SrtA-APBA conjugates

SrtA (20 μM) and 2-APBA (2 mM) was mixed in PBS pH 7. Sodium borohydride was added to a concentration of 100 mM. The reaction was incubated at room temperature for 2 h. The sample was diluted in water and analyzed by LC/MS that showed mass 19176.9 corresponding to one 2-APBA adduct.

13 μL of 100% trichloroacetic acid in PBS pH 7.0 was added to 0.25 mL SrtA-APBA sample containing about 95 μg protein. The mixture was vortexed vigorously and incubated at -80 ºC for 1 h. The sample was thawed ad centrifuged at 14000 RPM for 10 min. The supernatant was removed and the pellet was resuspended in 100 μL of could acetone by vortexing. The mixture was centrifuged again at 5000 RPM for 10 min and acetone was removed. The pellet was air dried and resuspended in 15 μL of 8 M urea in PBS pH 7.5. 35 μL of 100 mM ammonium bicarbonate was added together with 0.75 μL of 1 M DTT. The mixture was incubated in a 65 ºC water bath for 15 min. 1.25 μL of 500 mM iodoacetamide was added. The mixture was left at room temperature for 30 min. 25 μL of the mixture was diluted with 60 μL of PBS and 4 μL of trypsin (0.5 μg/ μL) was added with 1 μL of 100 mM $CaCl_2$. The trypsin digest was agitated at 37 ºC overnight and quenched by 5 μL of formic acid. The undigested and precipitated protein was pelleted by centrifuge (14000 RPM, 20 min). The supernatant was collected of which 10 μL was loaded onto C18 column and analyzed by Thermo LTQ Orbitrap Discovery mass spectrometer coupled to an Agilent 1200 series HPLC. The mass addition of 130.0590 was used to identify the APBA modifications. The data analysis was done by Eranthie Weerapana lab as shown in the reference [57].

3.7.13 Synthesis of DMMA



Scheme 3-1. Synthesis of DMMA

a. Synthesis of 3-1: 2,5-dihydroxybenzaldehyde (0.5 g, 3.62 mmol) and potassium carbonate (1 g, 7.24 mmol) were mixed in DMF (5 mL) and cooled on ice. Nitrogen gas balloon was attached to avoid water and oxygen. Iodomethane (452.3 µL, 7.24 mmol) was added into the mixture slowly. The reaction was stirred at room temperature for an overnight. Water (100 mL) was added into the reaction which was extracted with ethyl acetate for 3 times. The combined organic layers were washed with brine and dried over sodium sulfate. The solvent was then evaporated under vacuum. The crude product was purified on silica gel using hexane: ethyl acetate (8:2) as eluent. The desired product was obtained as white solid (85%).

   ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.44 (s, 1H), 7.33 (d, J = 3.3 Hz, 1H), 7.14 (dd, J = 9.1, 3.3 Hz, 1H), 6.94 (d, J = 9.1 Hz, 1H), 3.90 (s, 3H), 3.80 (s, 3H).

b. Synthesis of DMMA: Compound 3-1 (100 mg, 0.6 mmol) and 2-(4-methoxyphenyl)acetonitrile (81 µL, 0.6 mmol) were dissolved in anhydrous ethyl alcohol. Sodium methoxide (12 µL of 5M solution, 0.06 mmol) was added dropwise. The reaction was stirred at room temperature overnight. The solvent was dried under vacuum. The crude product was purified on silica gel with 10% ethyl acetate in hexane as elution. The product was obtained as white solid (81%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.83 (s, 1H), 7.75 (d, J = 3.0 Hz, 1H), 7.65 – 7.61 (m, 2H), 6.97 – 6.93 (m, 3H), 6.87 (d, J = 9.0 Hz, 1H), 3.85 (s, 3H), 3.85 (d, J = 1.2 Hz, 3H), 3.84 (s, 3H).

3.8. References

1. Zorzi, A., K. Deyle, and C. Heinis, *Cyclic peptide therapeutics: past, present and future.* Current Opinion in Chemical Biology, 2017. **38**: p. 24-29.
2. Suree, N., et al., *The Structure of the Staphylococcus aureus Sortase-Substrate Complex Reveals How the Universally Conserved LPXTG Sorting Signal Is Recognized **.* Journal of Biological Chemistry, 2009. **284**(36): p. 24465-24477.
3. Villar, E.A., et al., *How proteins bind macrocycles.* Nature Chemical Biology, 2014. **10**(9): p. 723-731.
4. Guimaraes, C.P., et al., *Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions.* Nature Protocols, 2013. **8**(9): p. 1787-1799.
5. Tsomaia, N., *Peptide therapeutics: Targeting the undruggable space.* European Journal of Medicinal Chemistry, 2015. **94**: p. 459-470.
6. Uhlén, M., et al., *Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
7. Bond, J.S., *Proteases: History, discovery, and roles in health and disease.* Journal of Biological Chemistry, 2019. **294**(5): p. 1643-1651.
8. López-Otín, C. and J.S. Bond, *Proteases: multifunctional enzymes in life and disease.* Journal of Biological Chemistry, 2008. **283**(45): p. 30433-30437.
9. López-Otín, C. and L.M. Matrisian, *Emerging roles of proteases in tumour suppression.* Nature reviews cancer, 2007. **7**(10): p. 800-808.

10. Turk, B., *Targeting proteases: successes, failures and future prospects.* Nature reviews Drug discovery, 2006. **5**(9): p. 785-799.

11. Drag, M. and G.S. Salvesen, *Emerging principles in protease-based drug discovery.* Nature reviews Drug discovery, 2010. **9**(9): p. 690-701.

12. Dormán, G., et al., *Matrix metalloproteinase inhibitors: a critical appraisal of design principles and proposed therapeutic utility.* Drugs, 2010. **70**(8): p. 949-64.

13. Sgier, D., et al., *Isolation and characterization of an inhibitory human monoclonal antibody specific to the urokinase-type plasminogen activator, uPA.* Protein Engineering, Design & Selection, 2010. **23**(4): p. 261-269.

14. Hansen, M., et al., *A urokinase-type plasminogen activator-inhibiting cyclic peptide with an unusual P2 residue and an extended protease binding surface demonstrates new modalities for enzyme inhibition.* Journal of Biological Chemistry, 2005. **280**(46): p. 38424-38437.

15. Angelini, A., et al., *Bicyclic peptide inhibitor reveals large contact interface with a protease target.* ACS chemical biology, 2012. **7**(5): p. 817-821.

16. Chen, S., et al., *Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries.* Journal of the American Chemical Society, 2013. **135**(17): p. 6562-6569.

17. Winer, A., S. Adams, and P. Mignatti, *Matrix Metalloproteinase Inhibitors in Cancer Therapy: Turning Past Failures Into Future Successes.* Molecular Cancer Therapeutics, 2018. **17**(6): p. 1147-1155.

18. Maola, K., et al., *Engineered Peptide Macrocycles Can Inhibit Matrix Metalloproteinases with High Selectivity.* Angewandte Chemie International Edition, 2019. **58**(34): p. 11801-11805.

19. Kong, X.-D., et al., *De novo development of proteolytically resistant therapeutic peptides for oral administration.* Nature Biomedical Engineering, 2020. **4**(5): p. 560-571.

20. Hruby, V.J., *Designing peptide receptor agonists and antagonists.* Nature Reviews Drug Discovery, 2002. **1**(11): p. 847-858.

21. Yamaguchi, J., et al., *cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA-protein fusions.* Nucleic Acids Res, 2009. **37**(16): p. e108.

22. Passioura, T., et al., *Display Selection of Exotic Macrocyclic Peptides Expressed under a Radically Reprogrammed 23 Amino Acid Genetic Code.* Journal of the American Chemical Society, 2018. **140**(37): p. 11551-11555.

23. Tan, S., C.W.H. Zhang, and G.F. Gao, *Seeing is believing: anti-PD-1/PD-L1 monoclonal antibodies in action for checkpoint blockade tumor immunotherapy.* Signal Transduction and Targeted Therapy, 2016. **1**(1): p. 16029.

24. Guzik, K., et al., *Development of the Inhibitors that Target the PD-1/PD-L1 Interaction-A Brief Look at Progress on Small Molecules, Peptides and Macrocycles.* Molecules (Basel, Switzerland), 2019. **24**(11): p. 2071.

25. Magiera-Mularz, K., et al., *Bioactive Macrocyclic Inhibitors of the PD-1/PD-L1 Immune Checkpoint.* Angewandte Chemie International Edition, 2017. **56**(44): p. 13732-13735.

26. Vinogradov, A.A., Y. Yin, and H. Suga, *Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges.* Journal of the American Chemical Society, 2019. **141**(10): p. 4167-4181.

27. Dougherty, P.G., A. Sahni, and D. Pei, *Understanding Cell Penetration of Cyclic Peptides.* Chemical Reviews, 2019. **119**(17): p. 10241-10287.

28. Lättig-Tünnemann, G., et al., *Backbone rigidity and static presentation of guanidinium groups increases cellular uptake of arginine-rich cell-penetrating peptides.* Nat Commun, 2011. **2**: p. 453.

29. Qian, Z., P.G. Dougherty, and D. Pei, *Targeting intracellular protein–protein interactions with cell-permeable cyclic peptides.* Current Opinion in Chemical Biology, 2017. **38**: p. 80-86.

30. Cui, C., et al., *Is β-Catenin a Druggable Target for Cancer Therapy?* Trends in Biochemical Sciences, 2018. **43**(8): p. 623-634.

31. Schneider, J.A., et al., *Design of Peptoid-peptide Macrocycles to Inhibit the β-catenin TCF Interaction in Prostate Cancer.* Nat Commun, 2018. **9**(1): p. 4396.

32. Bertoldo, D., et al., *Phage Selection of Peptide Macrocycles against β-Catenin To Interfere with Wnt Signaling.* ChemMedChem, 2016. **11**(8): p. 834-839.

33. Grossmann, T.N., et al., *Inhibition of oncogenic Wnt signaling through direct targeting of β-catenin.* Proceedings of the National Academy of Sciences, 2012. **109**(44): p. 17942-17947.

34. Takada, K., et al., *Targeted Disruption of the BCL9/β-Catenin Complex Inhibits Oncogenic Wnt Signaling.* Science Translational Medicine, 2012. **4**(148): p. 148ra117-148ra117.

35. Dietrich, L., et al., *Cell Permeable Stapled Peptide Inhibitor of Wnt Signaling that Targets β-Catenin Protein-Protein Interactions.* Cell Chemical Biology, 2017. **24**(8): p. 958-968.e5.

36. Ali, A.M., et al., *Stapled Peptides Inhibitors: A New Window for Target Drug Discovery.* Computational and structural biotechnology journal, 2019. **17**: p. 263-281.

37. Lian, W., et al., *Cell-Permeable Bicyclic Peptide Inhibitors against Intracellular Proteins.* Journal of the American Chemical Society, 2014. **136**(28): p. 9830-9833.

38. Upadhyaya, P., et al., *Inhibition of Ras signaling by blocking Ras-effector interactions with cyclic peptides.* Angew Chem Int Ed Engl, 2015. **54**(26): p. 7602-6.

39. Yamagishi, Y., et al., *Natural Product-Like Macrocyclic N-Methyl-Peptide Inhibitors against a Ubiquitin Ligase Uncovered from a Ribosome-Expressed De Novo Library.* Chemistry & Biology, 2011. **18**(12): p. 1562-1570.

40. Nawatha, M., et al., *De novo macrocyclic peptides that specifically modulate Lys48-linked ubiquitin chains.* Nature Chemistry, 2019. **11**(7): p. 644-652.

41. Huang, Y., et al., *Affinity Maturation of Macrocyclic Peptide Modulators of Lys48-Linked Diubiquitin by a Twofold Strategy.* Chemistry – A European Journal. **n/a**(n/a).

42. Heinis, C., et al., *Phage-encoded combinatorial chemical libraries based on bicyclic peptides.* Nature chemical biology, 2009. **5**(7): p. 502.

43. Baeriswyl, V., et al., *Development of a Selective Peptide Macrocycle Inhibitor of Coagulation Factor XII toward the Generation of a Safe Antithrombotic Therapy.* Journal of Medicinal Chemistry, 2013. **56**(9): p. 3742-3746.

44. Diderich, P. and C. Heinis, *Phage selection of bicyclic peptides binding Her2.* Tetrahedron, 2014. **70**(42): p. 7733-7739.

45. Kale, S.S., et al., *Cyclization of peptides with two chemical bridges affords large scaffold diversities.* Nature Chemistry, 2018. **10**(7): p. 715-723.

46. Jafari, M.R., et al., *Discovery of Light-Responsive Ligands through Screening of a Light-Responsive Genetically Encoded Library.* ACS Chemical Biology, 2014. **9**(2): p. 443-450.

47. Kelly, M., et al., *Peptide Probes of Colistin Resistance Discovered via Chemically Enhanced Phage Display.* ACS Infectious Diseases, 2020.

48. McCarthy, K.A., et al., *Phage display of dynamic covalent binding motifs enables facile development of targeted antibiotics.* Journal of the American Chemical Society, 2018. **140**(19): p. 6137-6145.

49. Chen, S., et al., *Identification of highly selective covalent inhibitors by phage display.* Nature Biotechnology, 2020.

50. Ng, S., et al., *Quantitative Synthesis of Genetically Encoded Glycopeptide Libraries Displayed on M13 Phage.* ACS Chemical Biology, 2012. **7**(9): p. 1482-1487.

51. Ng, S., et al., *Genetically Encoded Fragment-Based Discovery of Glycopeptide Ligands for Carbohydrate-Binding Proteins.* Journal of the American Chemical Society, 2015. **137**(16): p. 5248-5251.

52. Maresso, A.W. and O. Schneewind, *Sortase as a target of anti-infective therapy.* Pharmacological reviews, 2008. **60**(1): p. 128-141.

53. Cascioferro, S., et al., *Sortase A inhibitors: recent advances and future perspectives.* Journal of medicinal chemistry, 2015. **58**(23): p. 9108-9123.

54. Rentero Rebollo, I., et al., *Development of potent and selective S. aureus sortase A inhibitors based on peptide macrocycles.* ACS medicinal chemistry letters, 2016. **7**(6): p. 606-611.

55. Oh, K.-B., et al., *Discovery of Diarylacrylonitriles as a Novel Series of Small Molecule Sortase A Inhibitors.* Journal of Medicinal Chemistry, 2004. **47**(10): p. 2418-2421.

56. Rentero Rebollo, I. and C. Heinis, *Phage selection of bicyclic peptides.* Methods, 2013. **60**(1): p. 46-54.

57. Peeler, J.C., et al., *Generation of Recombinant Mammalian Selenoproteins through Genetic Code Expansion with Photocaged Selenocysteine.* ACS Chemical Biology, 2020. **15**(6): p. 1535-1540.

Chapter 4 Developing a novel phage display peptide library with site-specific modification on N-terminal cysteine

4.1 Introduction

Site-specific modification of proteins is a powerful strategy for understanding their function in complex biological milieu and developing protein-based therapeutics[1, 2]. Synthetic and enzyme-based techniques have been developed to conjugate proteins with fluorophores, reporters and drugs[3-5]. In the previous chapter, it was illustrated that site-specific modification of phage-encoded peptide libraries can significantly expand the chemical space of phage display. Not only have a variety of unnatural functionalities been incorporated providing strong driving force for ligand binding, but also diverse peptide conformations were stabilized by chemical modifiers.

In this chapter, I will first briefly introduce site-selective chemistry with the focus on N-terminal cysteine modification. Later, I will demonstrate how we applied these site-selective chemistries for novel phage display peptide libraries construction. Finally, the development of a strategy to introduce two modifiers on N-terminal cysteine selectively will be described.

4.1.1 Site-specific modification of proteins

As one of the least abundant amino acids, cysteine has been frequently used to modify proteins in a site-specific manner. For example, Tsien and colleagues developed a tetra-cysteine motif (Cys-Cys-Pro-Gly- Cys-Cys) to enable targeted and site-specific modification of a protein of interest in cells[6, 7]. Ligand 4',5'-bis(1,3,2-dithioarsolan-2-yl)fluorescein, called FLAsH-EDT_2, was synthesized to specifically label the tetracysteine peptide and generate bright fluorescence (Figure 4-1a). This system allows the imaging of fluorescently labeled proteins in intact cells[8]. More recently, Pentelute and coworkers reported that a "π-clamp" (Phe-Cys-Pro-Phe) tag that can be selectively labeled by a

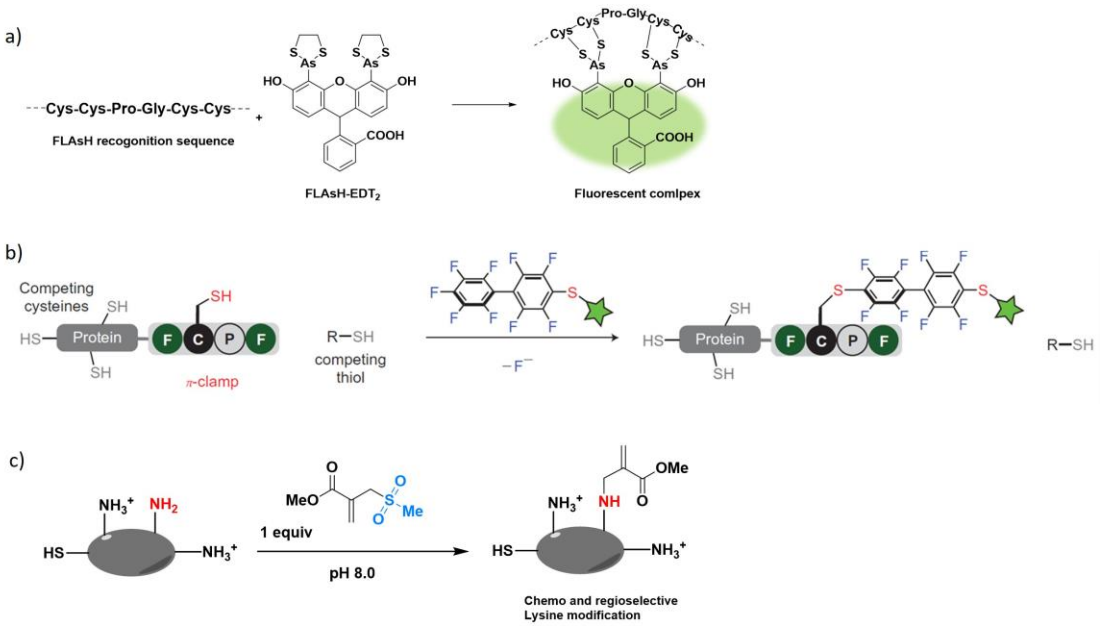


Figure 4-1. Site-specific modification of proteins. a) A tetracysteine peptide recognized by FLAsH-EDT2. b) A "π-clamp" tag selectively labeled by a perfluoroaromatic probe. Adapted with permission from Ref. [9]. c) Chemo- and regioselective lysine modification by sulfonyl acrylate reagents. Adapted from Ref. [10].

perfluoroaromatic probe (Figure 4-1b) [9]. Inserting the "π-clamp" in to antibodies enabled the synthesis of site-specific antibody-drug conjugates. Both works used the microenvironment to fine-tune cysteine reactivity for site-specific ligations. In addition to cysteine, lysine provides reactive amine group for labeling. Bernards group designed sulfonyl acrylate reagents for the labeling of a single lysine residue on native protein sequences (Figure 4-1c) [10]. The lysine with the lowest pK_a was kinetically favored and got modified. This method allows precise protein labeling for bioimaging and regioselective antibody conjugation.

4.1.2 N-terminal cysteine reactivity

Besides the side chains of cysteine and lysine residues, the N termini of proteins are popular targets for site-specific modification since they are usually solvent exposed and

can be extended without interfering the overall protein structure. Also, the pK_a of N-terminal α-amines is lower than that of other aliphatic amines which could favor the selective modification of N-terminal amines at low-to-neutral pH. Rosen and Francis have summarized the site-selective modification on N termini of protein in the reference[10]. In this chapter we will focus on the reactivity of N-terminal cysteine residues.

The unique reactivity of N terminal cysteines (NCys) has given rise to the native and expressed protein ligation technologies (Figure 4-2a) [11, 12], which allow chemical synthesis of proteins for various applications. However, the native chemical ligation gives a relative small rate constant around 0.1 M^{-1} s^{-1}[13]. Furthermore, 2-cyanobenzothiazole (CBT) is found to rapidly conjugate with an NCys under physiologic conditions to give a

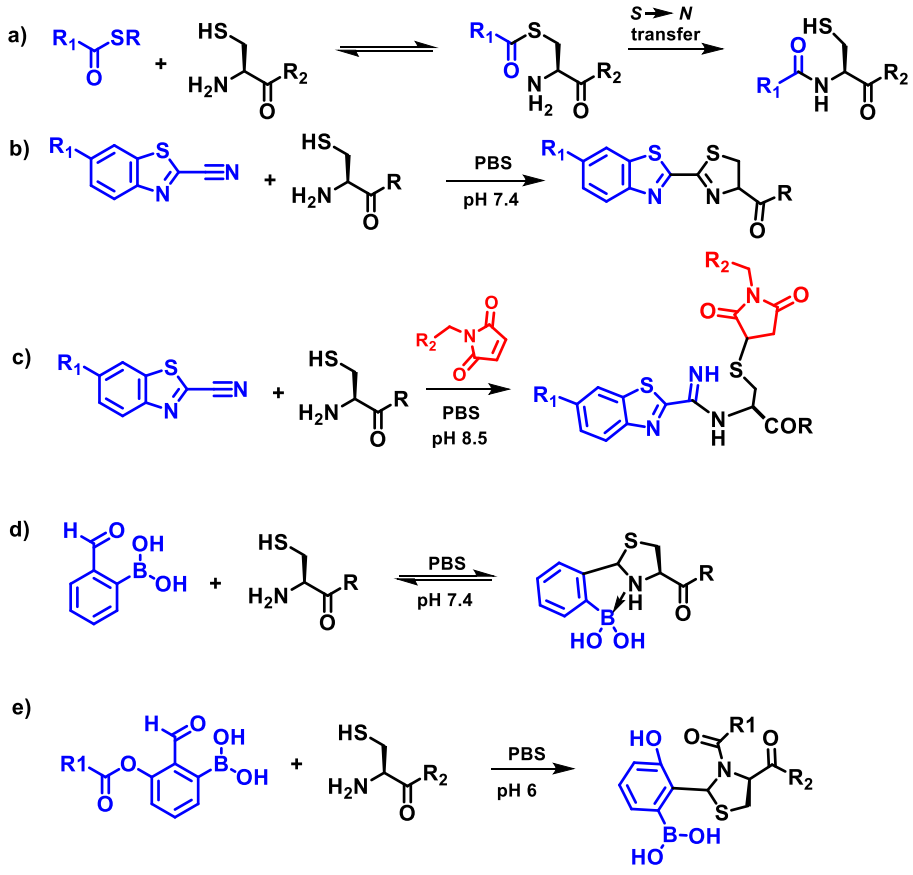


Figure 4-2. Reactions for selective NCys modification

luciferin product, which provides a powerful tool for protein labeling (Figure 4-2b) [14]. This condensation reaction shows much faster kinetic, k_2 = 26.8 M^{-1} s^{-1}, but suboptimal NCys selectivity. Lin and his coworkers reported a peptide tag with an internal cysteine for CBT ligation, which can stabilize the thioimidate adduct by the sequence environment[15]. This work indicates that some internal cysteine residues can also react with CBT resulting in nonspecific modifications. Recently, we also discovered an alternative reaction pathway of CBT condensation, which leads to an N-terminal amidine rendering the side chain thiol available for further modification (Figure 4-2c) [16]. Based on this mechanism, we developed a strategy for facile N, S-double labeling of a protein which will be demonstrated in this chapter. Another chemotype found to selectively conjugate with NCys in complex biological media is 2-formylphenylboronic acid (2-FPBA), which gives a thiazolidino boronate (TzB) complex (Figure 4-2d) [17, 18]. The reaction shows very fast kinetics with a rate constant of 5000 $M^{-1}s^{-1}$, which is several orders of magnitude faster than NCL or CBT conjugation. In addition, this conjugation proceeds with exquisite selectivity which enables facile protein modification. However, the reaction is reversible and the conjugate dissociates over an hour. To overcome this problem, recently our group reported a TzB-mediated acylation reaction of NCys that forms stable conjugates (Figure 4-2e). We'll show the application of this conjugation to modify phage library in this chapter.

4.2 Construction of a phage display peptide library presenting N-terminal cysteine

4.2.1 Plasmid construction

We fused an HA tag at N-terminus of the pIII protein before the factor Xa cleavage site (IEGR) to monitor the cleavage of factor Xa (Figure 4-3a). The protease cleavage site is followed by a $C(X)_7C$ peptide library and a short spacer (GGGS). The peptide sequence displayed on phage was YPYDVPDYAAIEGRC-X7-CGGGS. The Ph.D. Peptide Display Cloning System from NEB was used to construct the peptide library (Figure 4-3b). The insert oligonucleotide was purchased from IDT. The insert was prepared by annealing the M13 extension primer to the oligonucleotide, extending with Klenow fragment, and digesting with EagI-HF and KpnI-HF (Figure 4-4a). The M13KE vector was also digested

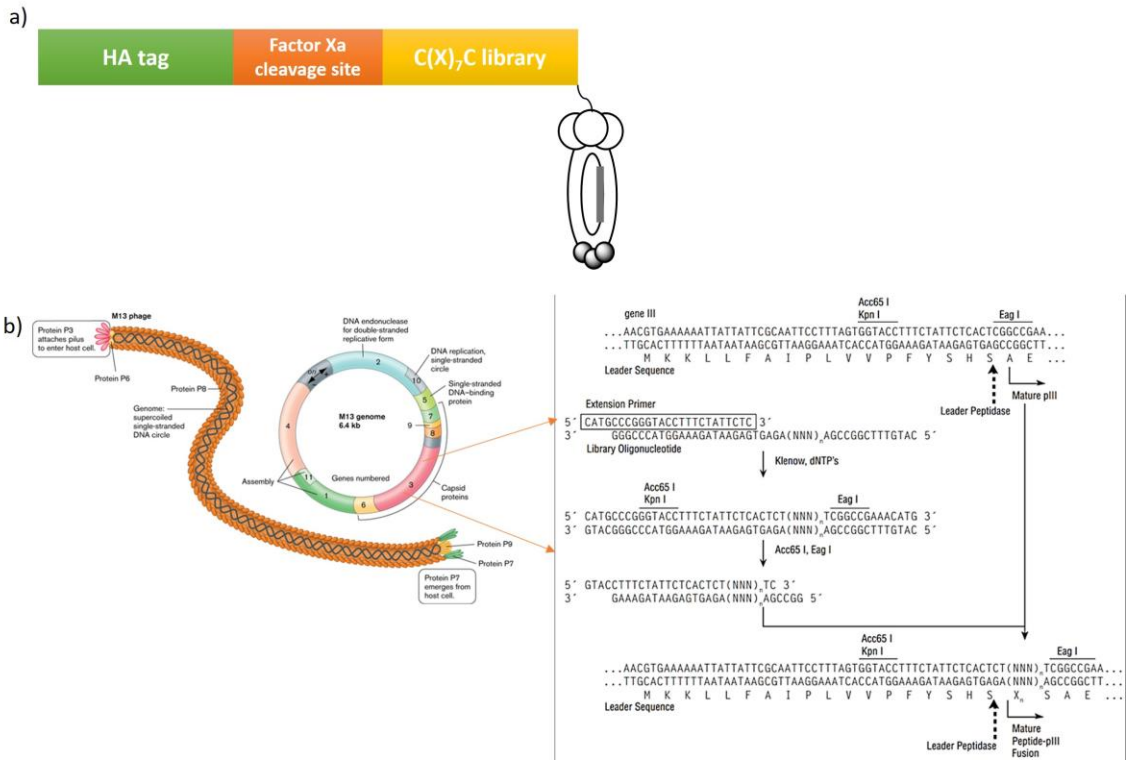


Figure 4-3. Illustration of the library design(a) and plasmid construction (b).

by EagI-HF and KpnI-HF (Figure 4-4b). Both insert and vector were purified from the gel and ligated by T4 DNA ligase (Figure 4-4c). We transformed TOP10 competent cells to construct the library that covered the complexity of 10^8 clones.

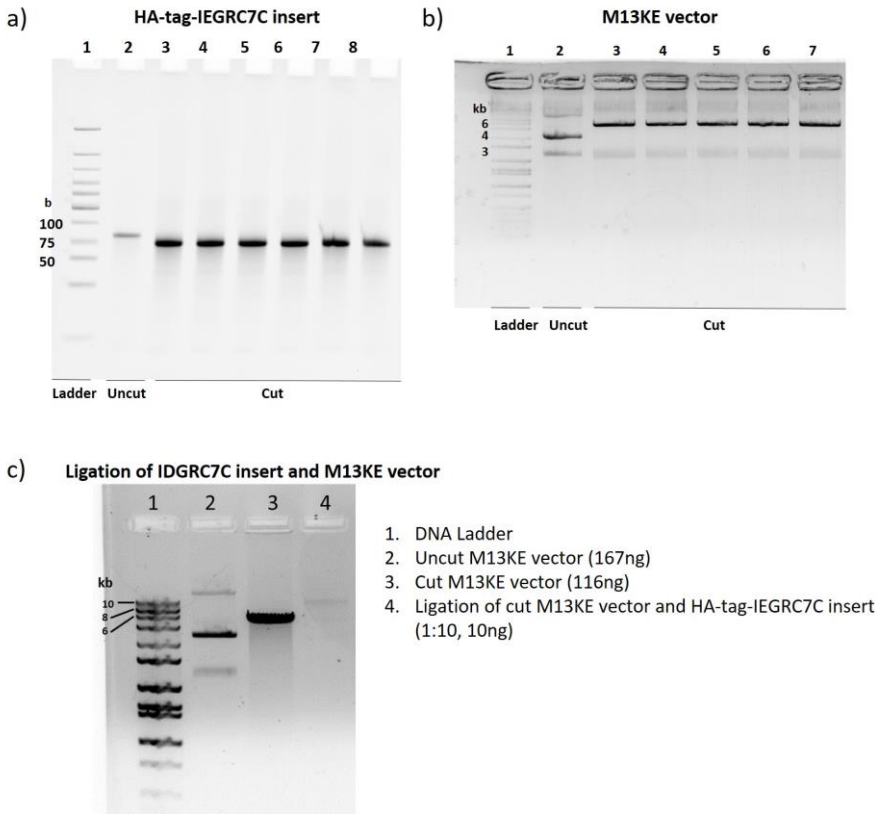


Figure 4-4. Gel Electrophoresis analysis of DNA. a) Library insert before (lane 2) and after (lane 3-8) restriction enzyme digestion was separated and analyzed on an 8% PAGE gel. b) M13KE vector before (lane 2) and after (lane 3-7) restriction enzyme digestion was separated and analyzed on 1% agarose gel. c) 1% agarose gel analysis of the ligation of M13KE vector and insert.

4.2.2 Validation of phage display peptide library quality

To validate the diversity of the library, 26 phage clones were amplified for DNA sequencing (Table 4-1). Each of them contained a different insert sequence which indicates an excellent diversity. The quality of the library was further tested by screening against streptavidin. After 2 rounds of panning, the consensus sequence for streptavidin-binding peptides, His-Pro-Gln (HPQ), already showed up 4 times out of 5 clones (Table 4-2). Also,

the phage output (10^8 pfu) significantly increased after the second round of panning, which indicates the enrichment of the real binders. Therefore, a phage displayed peptide library with 10^8 complexity and good quality was successfully constructed.

In the same way, a library with 5 randomized amino acid residues in between a pair of cysteine residues was also constructed, labeled as $C(X)_5C$ library.

Table 4-1. The sequencing results of HA tag-IEGR-C7C phage library

YPYDVPDYAAIEGRCTAISMMTC	YPYDVPDYAAIEGRCTSIHNAIC	YPYDVPDYAAIEGRCQVINKNSC
YPYDVPDYAAIEGRCHNLSVIEC	YPYDVPDYAAIEGRCVIDNRVQC	YPYDVPDYAAIEGRCRVTNPLMC
YPYDVPDYAAIEGRCHVSQQKGC	YPYDVPDYAAIEGRCEVYTPTWC	YPYDVPDYAAIEGRCTNTLDHQC
YPYDVPDYAAIEGRCSAIPAPTC	YPYDVPDYAAIEGRCTNTLLYSC	YDVPDYAAIEGRCMKYNLFVC
YPYDVPDYAAIEGRCSPFPDSRC	YPYDVPDYAAIEGRCPNIRYNEC	YDVPDYAAIEGRCTGGTELTC
YPYDVPDYAAIEGRCGSGVYNHC	YPYDVPDYAAIEGRCKVKSEGMC	YPYDVPDYAAIEGRCSGKTVRFC
YPYDVPDYAAIEGRCYLFINENC	YPYDVPDYAAIEGRCHVLSLNHC	
YPYDVPDYAAIEGRCLVSMKHNC	YPYDVPDYAAIEGRCPSTPHMLC	
YPYDVPDYAAIEGRCTASWTMIC	YPYDVPDYAAIEGRCTHSRTLLC	
YPYDVPDYAAIEGRCMYAPDAFC	YPYDVPDYAAIEGRCAQTISTSC	

Table 4-2. Sequencing analysis of HA tag-IEGR-C7C phage library screening against streptavidin.

CHPQFPNSC
CHPQNNSIC
CNEGWHPVC
CHPQFPNSC
CHPQAGXWC

4.2.3 Factor Xa cleavage

Factor Xa was used to cleave the peptide and generate the N-terminal cysteine on phage. To confirm FXa cleavage, we carried out an ELISA assay to confirm the removal of HA tag from phage (Figure 4-5a). M13 major coat protein antibody was immobilized into each well of 96 well plate to capture phage. HA tag monoclonal antibody-HRP conjugate was used to detect the HA tag on the N-terminus of the peptide. If FXa successfully removes the HA tag, then there won't be any ELISA signal. From the result, we can see that no HA tag was detectable after FXa cleavage (Figure 4-5b). In contrast, without FXa treatment, high ELISA signal showed up. These results indicate that FXa cleavage was efficient on phage. In addition, phage titer before and after FXa treatment didn't change much.

Overall, we have developed a strategy to construct phage-encoded peptide libraries presenting N-terminal cysteine. The libraries possess high diversity and good quality, which allow us to do selection against protein targets and identify peptide binders.

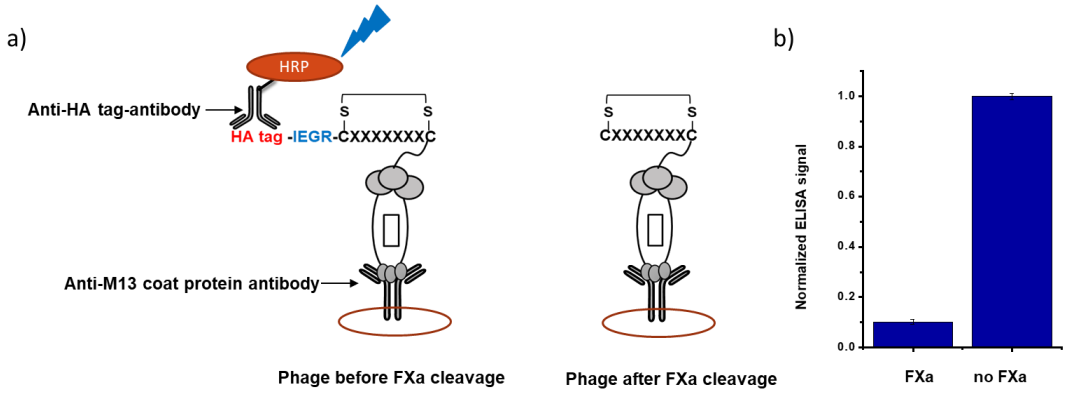


Figure 4-5. ELISA for HA tag detection. a) Illustration the ELISA experimental design b) ELISA results showing that no HA tag was detectable after factor Xa cleavage. These results indicate that FXa cleavage was efficient on phage.

4.3 Site-specific modifications on N-terminal cysteine of phage

With the library presenting N-terminal cysteine in hand, we move forward to investigate site specific modification of the peptide on phage. We applied two methods, thiazolidino

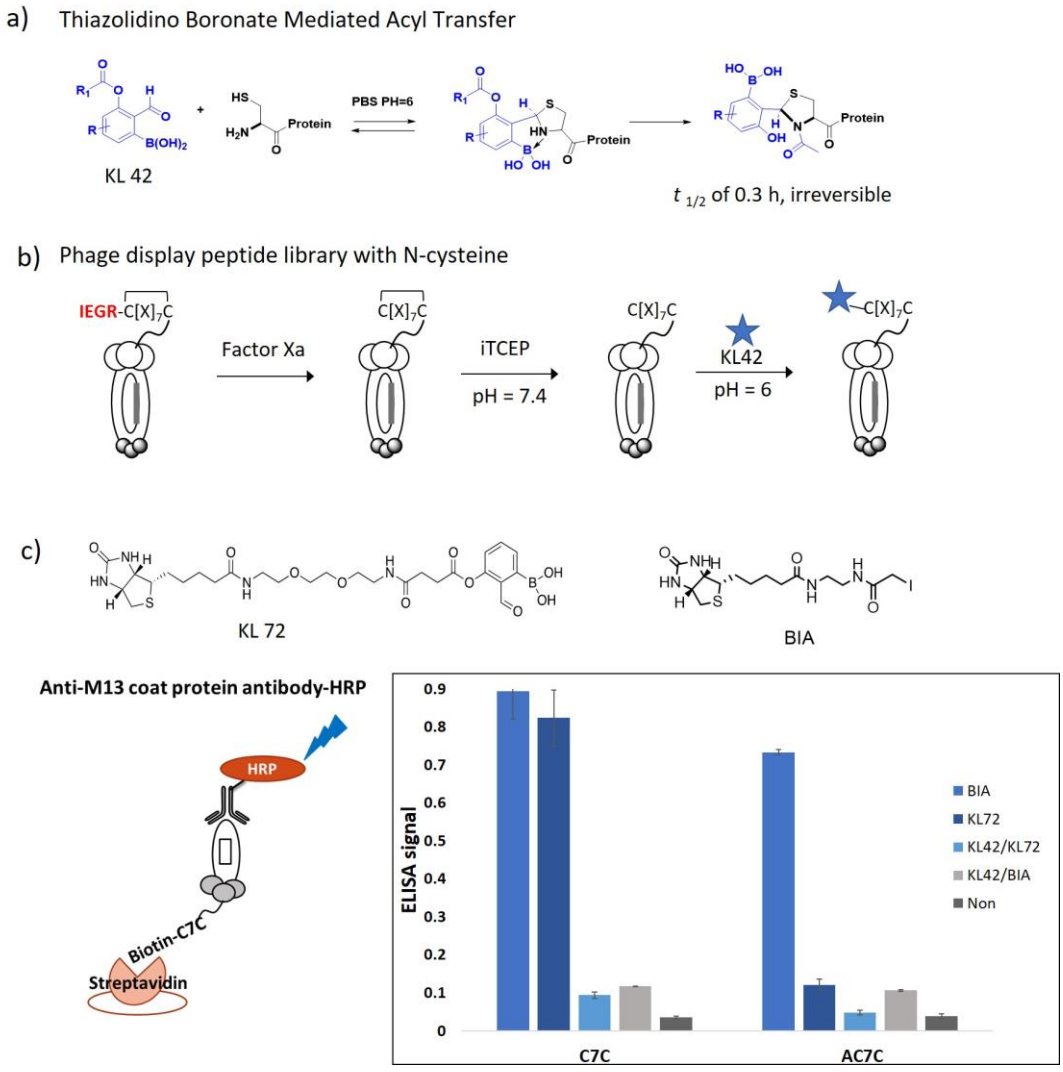


Figure 4-6. Chemical modification of a M13 phage library through TzB mediated NCys conjugation. a) Stable NCys modification with KL42. b) Illustration of site-selective modification of phage-encode peptide library. c) Illustration of ELISA experimental design and ELISA results showing that KL72 modifies NCys-bearing phage with high efficiency and specificity.

boronate (Tzb) mediated acylation reaction of NCys and 2-cyanobenzothiazole (CBT) condensation.

4.3.1 Acetyl-FPBA ester modification

We recently reported a new method for NCys modification based on a TzB mediated acylation reaction of NCys. An acetyl ester of 2-FPBA (KL42) was synthesized which can rapidly form a TzB intermediate with NCys (Figure 4-6a). The intramolecular acyl transfer happened afterwards to give a stable thiazolidine product. The fast kinetics of the conjugation allows clean NCys modification with an equimolar amount of the labeling reagent. In addition, the reaction also shows superb NCys selectivity. Therefore, we decided to apply this method to modify NCys of the peptides displayed on phage (Figure 4-6b).

To monitor the reaction happened on phage, we developed an ELISA method. Streptavidin was immobilized onto the well plate and M13 coat protein antibody-HRP conjugate was used to detect the phage bound to streptavidin (Figure 4-6c). If the phage is labeled with biotin containing reagents, then the ELISA signal should be high (Figure). KL 72, a biotin derivative of KL42, was synthesized which allows us to monitor the modification of phage-encoded peptides by ELISA.

FXa treated phage (C7C) was incubated with (1) BIA as the positive control; (2) KL72; and (3) KL42 followed by KL72. The phage without NCys was used as control (AC7C phage). Excitingly, the ELISA results show that KL72 elicited efficient biotinylation of the FXa treated phage (Figure 4-6c), giving comparable readout to the positive controls. Furthermore, the KL72-elicited phage pulldown can be effectively blocked by pre-treatment with KL42, which presumably modifies NCys without biotinylation. Importantly,

treating the control phage (AC7C) with KL72 afforded little phage pulldown as expected for the absence of NCys. Collectively, these results showcase the efficiency and specificity of KL72 for modifying bacteriophage with N-terminal cysteines.

The pH plays an important role in this reaction. The optimal pH was found to be 6 so we controlled the pH at 5-6 when modified phage with KL42 or KL72. However, we found that phage lost viability during the iTCEP reduction at acidic condition. Therefore, we need to carry out the iTCEP reduction at neutral pH and carefully adjust the pH before adding labeling reagents. It is also found that KL72 could nonspecifically label phage at neutral pH, which can be explained by the acylation of reactive lysine residues on the phage surface. Overall, controlling the pH at 5-6 is the key to site-specific modification of NCys on phage by acetyl esters of 2-FPBA.

Note that NCL has been previously applied to modify phage libraries, but much compromised phage viability was observed, presumably due to non-specific reactions[19]. In contrast, KL72 treatment did not noticeably compromise the phage's infectivity since comparable titering results were obtained for phage samples before and after KL72 treatment.

4.3.2 CBT modification

2-cyanobenzothiazole (CBT) has been found to selectively rapidly conjugate with an NCys under physiologic conditions to give a luciferin product (Figure 4-7a). It has proven to be a powerful tool for protein labeling, but has never been applied to phage. We envision that by using CBT to modify the NCys, we can further label the internal Cys with a different functionality, and therefore to generate a phage-encoded peptide library with dual modifications (Figure 4-7b).

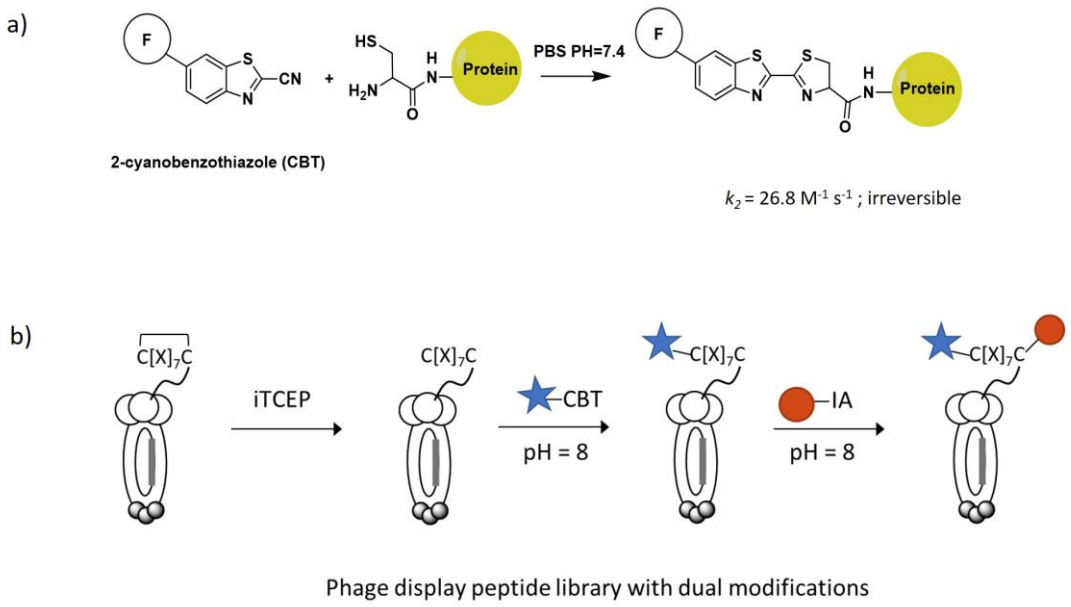


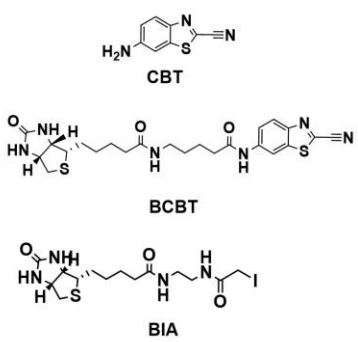
Figure 4-7. Chemical modification of a M13 phage library through CBT mediated NCys conjugation. a) CBT-NCys conjugation forming a luciferin product. b) Illustration of phage display peptide library with dual modifications.

We first studied CBT condensation reaction on a model peptide with the sequence YPYDVPDYAAIEGRCQVINKNSC. The peptide was treated with FXa followed by a series of modifications, including (1) BIA as the positive control; (2) BCBT; and (3) CBT first followed by BIA to achieve the dual-modification on NCys and internal Cys (Figure 4-8a). The peptide not treated with FXa underwent the same modifications as controls. The reactions were monitored by LC-MS.

As shown in the Figure 4-8b, BIA modified the peptide efficiently as expected. However, two CBT motifs adduct showed up when the peptide reacted with CBT or BCBT. It is not surprising because CBT can label the internal cysteine forming thioimidate, but this reaction is reversible. Therefore, BIA could replace CBT motif to modify the internal

cysteine irreversibly as shown in trial 4 and 8. Free cysteine can be used to quench the reaction between CBT and internal cysteine. Overall, CBT can modify NCys efficiently and the internal Cys can also be labeled by BIA. Dual labeling was achieved on the model peptide.

a)



b)

	Modifications	Products
~100uM peptide + FXa	1. BIA, 1mM, 2h	CQVINKNSC+2BIA
	2. B-CBT, 0.5mM, 2h	CQVINKNSC+2BCBT
	3. CBT, 0.5mM, 2h + BIA, 1mM, 2h	CQVINKNSC+2 CBT; CQVINKNSC+CBT+BIA
~100uM peptide	4. BIA, 1mM, 2h	YPYDVPDYAAIEGRCQVINKNSC+ 2BIA
	5. B-CBT, 0.5mM, 2h	YPYDVPDYAAIEGRCQVINKNSC+ 2BCBT
	6. CBT, 0.5mM, 2h + BIA, 1mM, 2h	YPYDVPDYAAIEGRCQVINKNSC+ 2BIA

Figure 4-8.CBT condensation reaction of a model peptide. a) Structures and nomenclatures of reagents. b) Products after each modification.

Encouraged by these results, we applied the modifications on phage. iTCEP was used to reduce the disulfide bond of $C(X)_7C$. After reduction, the phage was incubated with following reagents, (1) BIA as the positive control; (2) BCBT; (3) CBT first followed by BCBT to monitor the completion of modification; and (4) CBT first followed by BIA (Figure 4-9a). The modifications were done on phage treated with FXa, labeled as "NCys", as well as phage not treated with FXa, labeled as "no NCys". From the results (Figure 4-

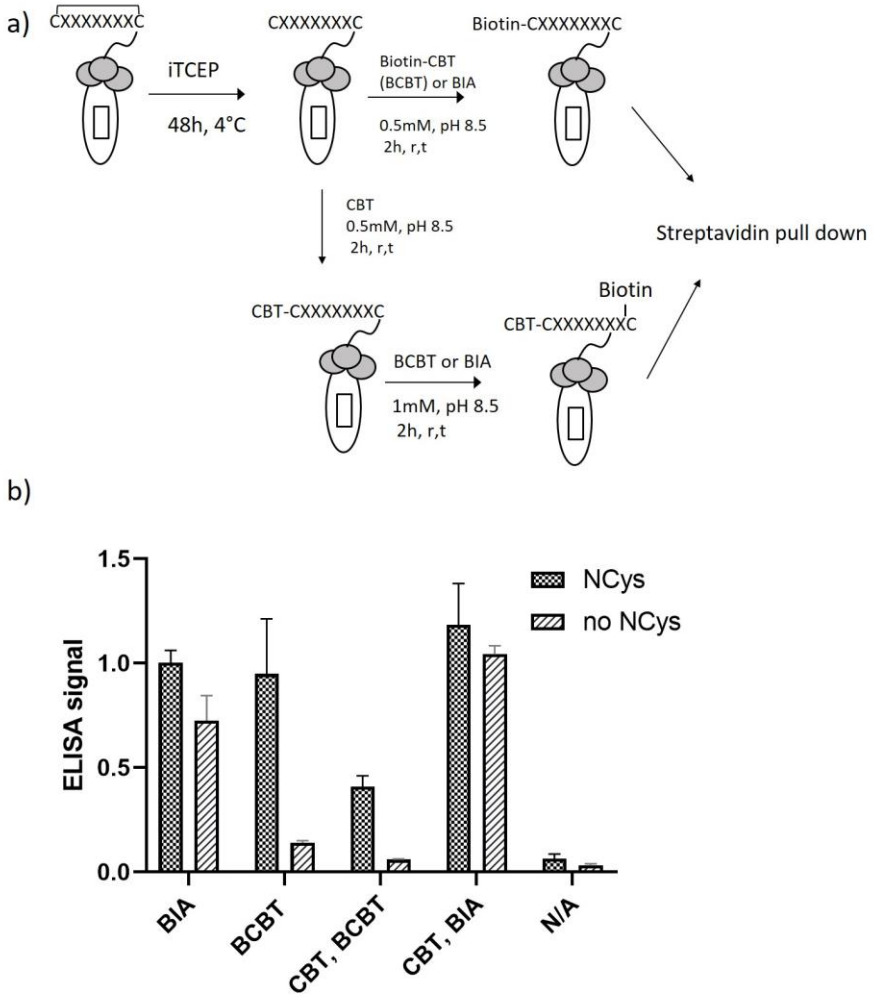


Figure 4-9. Phage display peptide library with dual modifications. a) Illustration of dual modifications of phage and control experiments. b) ELISA results showing that CBT NCys-bearing phage with high efficiency and specificity. Iodoacetamide reagent can further modify the internal cysteine to achieve the dual modification of phage-encoded peptide.

9b) we can see that, phage was successfully labeled by BIA indicating the reduction of the disulfide bond was efficient. It is exciting to see that the NCys phage modified by BCBT showed as high signal as the positive control. In contrast, the phage without NCys showed little signal after the incubation with BCBT, which highlights the site specificity of CBT labeling on NCys. In the case that phage was first modified with CBT for 2 h and then BCBT, the signal decreased about 60%, which indicates that 2 h incubation with CBT is sufficient to modify at least 60% phage. The incubation time can be longer to achieve the completion of modification. Furthermore, the addition of BIA after CBT can also generate high ELISA signal indicating the internal Cys can be further modified by BIA.

To sum up, we investigated the CBT condensation reaction to modified peptide library presenting NCys on phage. The ELISA results demonstrated that CBT reagent can site specifically label NCys leaving the internal Cys to be alkylated by iodoacetamide reagent. We believe that this dual-modification strategy can expand the chemical space of phage-encoded peptide library.

4.3.3 Phage library with dual modifications screening against Sortase A

To explore the utility of the phage library with dual modifications, we decided to attach two distinct warheads onto the peptides and apply the library to screen against Sortase A (SrtA). A small molecule inhibitor of SrtA with (Z)-Diarylacrylonitrile structure, compound DMMA, drew our attention (Figure 4-10a), which has an IC_{50} of 9.2 μM[20]. According to the docking model of A within the SrtA, the methoxy group at the 5-position of the phenyl ring points outside the active site, which allows us to conjugate a CBT handle at this position without impairing the binding. Therefore, we synthesized the Diarylacrylonitrile-CBT (DCBT) with IC_{50} of 30 μM to modify the NCys. As for the internal cysteine, we choose to use APBA-IA for modification. We presumed that the APBA warhead and peptide side chains could work collectively to enhance the affinity of the small molecule inhibitor against SrtA, so that we could identify a strong binder of SrtA.

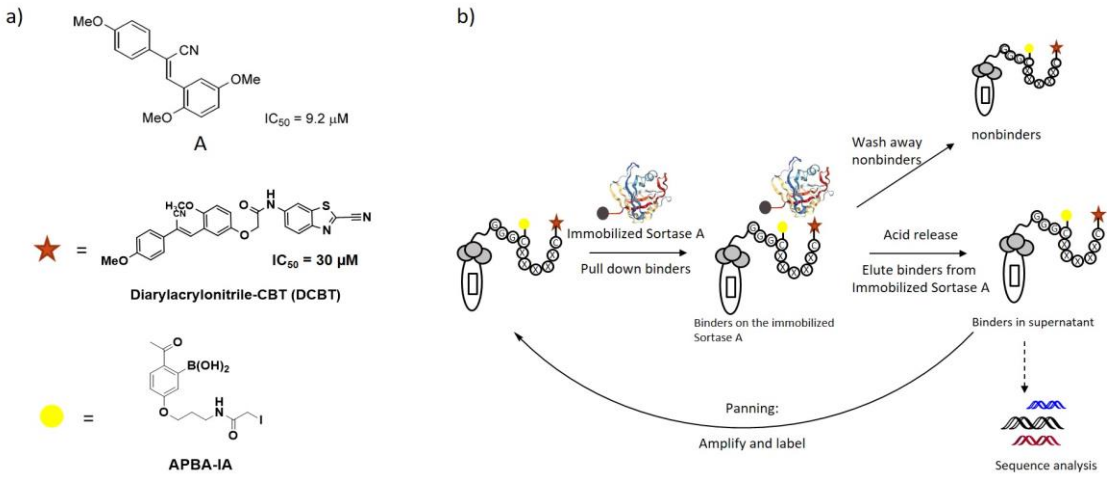


Figure 4-10. Phage library with dual modifications screening against Sortase A. a) Structures and nomenclatures of reagents. b) Schematic representation of panning against SrtA.

We used DCBT and APBA-IA to modify the $C(X)_5C$ phage library sequentially and then incubated the phage with SrtA which was immobilized onto the magnetic beads. The beads were washed 10 times to remove the weak binders. And then phage binders were eluted from the beads and amplified for the next rounds of panning (Figure 4-10b). In total, three rounds of panning were done. The phage titer of each round output was in the range of 10^5 to 10^6 pfu, however, there was no significant increase (Table 4-3). We sequenced some phage clones after the second and third round, however, no repeating sequence showed up (Table 4-4). Unfortunately, these results indicate that no specific and high affinity binder has been identified.

Table 4-3. Phage titer after each round of panning against SrtA (pfu)

1st	2nd	3rd
$3*10^6$	$4*10^5$	$3*10^6$

Table 4-4. Sequencing analysis of screening against SrtA.

2nd round
CHDDRMC
CATRGHC
CIDKGAC
CVAHKNC
CTQRGYC

3rd round
CGHNAWC
CPKSMEC
CNAAKTC
CKSKTGC
CTFNRGC
CGSSGRC
CHWGILNKC
CSANSAC
CGWTRDC

We can improve the library from the following aspects. First, the solubility of DBCT in aqueous solution is very poor which results in the incomplete modification on phage. This might decrease the chance to identify the strong binder. To solve this problem, we need to increase the hydrophilicity of DCBT, probably by introducing PEG linker in between diarylacrylonitrile and CBT moieties. Second, the length of the peptide can be optimized. Five amino acid residues may be too short to provide enough interactions to bind the target.

We also constructed the library with seven and nine random amino acids in between a pair of cysteine residues, which provide higher diversity and more driving force for binding. Lastly, the peptides with dual modifications are linear, which unlike cyclic peptides are less prone to form a well-defined structure. The entropy penalty needs to be minimized for ligand binding. Therefore, the strategy to introduce cyclic structures will be investigated.

4.4 Discovery of a strategy to N, S-doubly label N-terminal cysteine of peptides and proteins via an alternative pathway of 2-cyanobenzothiazole (CBT)

Conjugation of 2-cyanobenzothiazole (CBT) with N-terminal cysteines (NCys) has proven to be a powerful tool for site-selective labeling of proteins. While CBT-NCys conjugation typically gives a luciferin product, we herein report an alternative reaction pathway leading to an N-terminal amidine product with the side chain thiol available for further modification. This non-luciferin pathway is favored under basic conditions and dependent on the peptide sequence downstream of the NCys. In particular, we report a tripeptide tag CIS that allows facile N, S-double labeling of a protein of interest with excellent yield. This alternative reaction pathway of CBT-NCys condensation presents a significant addition to the toolbox for site-specific protein modifications.

4.4.1 Mechanism of amidine formation in CBT-NCys conjugation

We serendipitously discovered that mixing CIY (0.1 mM) with five equivalents of CBT at pH 8.5 gave 50% of the luciferin product D according to LC-MS analysis (Figure 4-11a). An additional product was observed at 50% that corresponds to a bis-CBT modified peptide F (Figure 4-11b). ^{1}H-NMR characterization of the reaction showed two new sets of aromatic peaks in addition to the CBT resonances (Figure 4-11c). Mechanistically speaking, this bis CBT-labelled product can result from CBT conjugation with the amidine product **E**, which arises from the elimination of the side chain thiol of the 2-aminothiazolidine

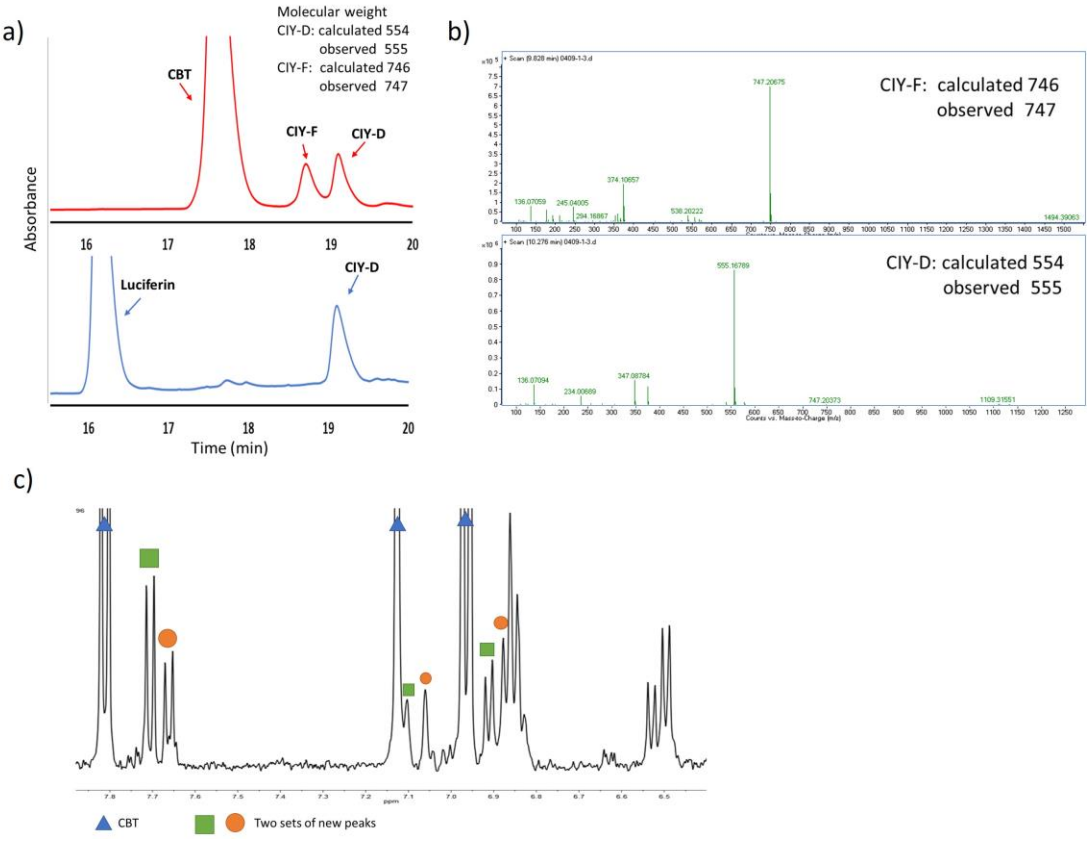
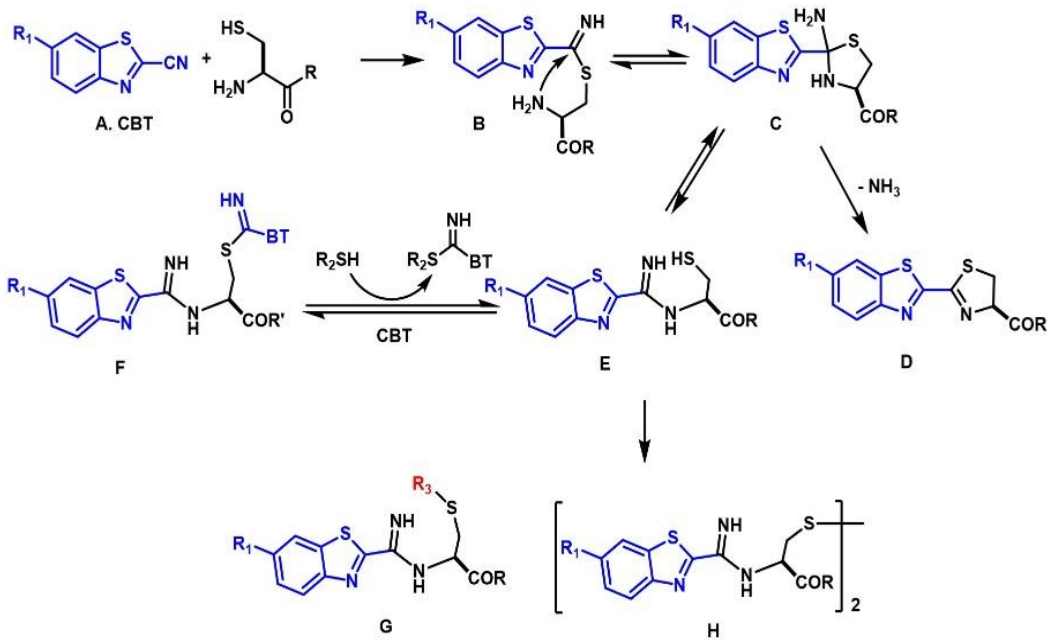


Figure 4-11. Characterization of the reaction between CIY and CBT. a) LC-MS analysis of 100 µM CIY treated with 5 equiv of CBT at pH 8.5 for 2h (top) and then with 5 equiv of cysteine (bottom). b) Mass spec analysis of CIY-F and CIY-D. c) ^{1}H-NMR spectrum of CIY and 5 eq CBT (pH 8.5) reaction mixture. The sample was prepared in 50% DMSO-d_6 for NMR study.

intermediate **C** (Scheme 4-1). Through this alternative elimination pathway of **C**, the CBT moiety is passed along from the side chain thiol to main chain amine under our experimental conditions. The formation of the bis adduct **F** is further supported by treating the conjugate with 1 mM free cysteine, which triggered the disappearance of **F** and increased formation of **D** according to LC-MS analysis (Figure 4-11a, bottom). This result is consistent with the labile nature of the thioimidate moiety in **F**. It also indicates that the amidine product **E** can revert to **C** and further transform into the luciferin product **D**. The reversibility of these reaction pathways presents a previously underappreciated aspect of the CBT-NCys conjugation chemistry. A similar thioimidate-to-amidine transformation was recently reported between internal cysteine and lysine side chains by Bertozzi and coworkers, which also supports this mechanism of amidine formation in CBT-NCys conjugation[21].



Scheme 4-1. Postulated reaction pathways of CBT-NCys conjugation.

The CBT-NCys conjugation is initiated with the thiol addition into the cyano functionality to give a thioimidate B, which then reacts intramolecularly with the NCys amine to give the 2-aminothiazolidine C. Elimination of C can proceeds via three possible routes: elimination of exocyclic amine gives the luciferin product D, elimination of the endocyclic amine gives back C, and elimination of the thiol gives the amidine product E. With a free thiol, E could capture another CBT to give a bis-CBT adduct F. Alternatively, the free thiol of E could be captured by an alkylating reagent

4.4.2 The effect of pH and equivalent of CBT

To better understand the luciferin versus nonluciferin reaction pathways, we tested the effect of pH on F formation in the CBT-CIY conjugation (Figure 4-12a). At pH 6.5, the conjugation gives predominantly the luciferin product, yet the nonluciferin product increases with pH and reaches 55% at pH 9.5. This is consistent with the fact that basic conditions disfavor the elimination of ammonia which is the rate determining step (C to D) to form luciferin. Meanwhile, the nucleophilicity of thiol group increases with pH which

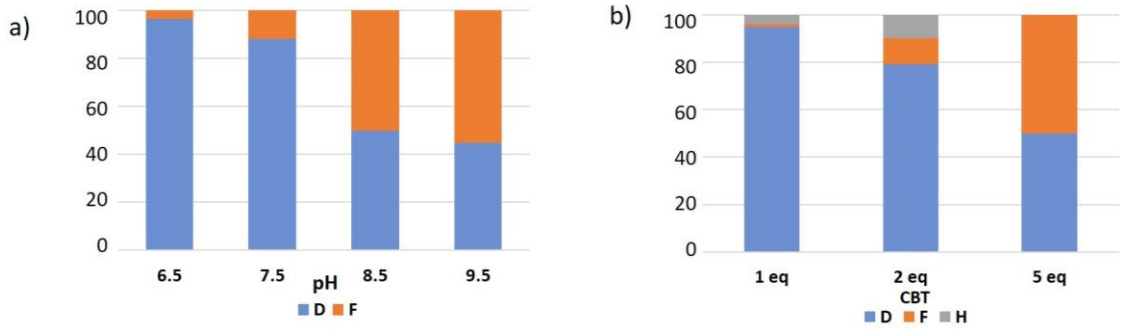


Figure 4-12. The effect of pH and equivalent of CBT. a) pH dependence of CIY-F formation. 100 μM CIY was incubated with 5 equiv of CBT at pH 6.5, 7.5, 8.5, and 9.5 respectively for 2 h before LC-MS analysis. (b) Effect of CBT equivalency on CIY-F formation. 100 μM CIY was incubated at pH 8.5 for 2 h with 1, 2, and 5 equiv of CBT, respectively, and then subjected to LC-MS analysis.

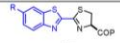
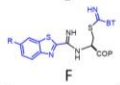
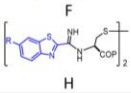
favors the thioimidate formation. The pH dependence presumably explains why the non-luciferin product has eluded characterization in earlier studies, which were primarily done at neutral pH.

In addition to pH, the amount of the CBT used has a significant effect on the outcomes of the CBT-CIY conjugation. In contrast to the bis-CBT adduct F seen at 50% with 5 equivalents of CBT, CBT at 1 equivalent yielded only 3% F and 87% of the luciferin product D. Consistently, 2 equivalents of CBT gave 15% F and 73% D (Figure 4-12b). In addition, a peptide dimer H (~10%) was observed because of disulfide formation of E. The observed dimerization can be rationalized by the slower E-to-F conversion caused by the low concentrations of CBT used.

4.4.3 Reactions between CBT and CXX short peptides

With the deepened understanding of the reaction mechanisms, we further postulated that the CBT-NCys reaction outcome might be tunable with peptide sequences downstream of the NCys. To test this hypothesis, we have synthesized a panel of short peptides and analyzed their conjugation with CBT (Table 4-5). The first group contains CIX mutants, in which X consists of a hydrophobic amino acid. They elicited 22% to 50% F formation. In contrast, CIA under the same experimental conditions gave exclusively the luciferin product D. It appears that hydrophobic residues can favor the bis-CBT adduct formation. We next mutated Ile to Ala mutation (CIY vs CAY), which had minimal effect on F formation. However, Ile-to-Pro mutation abolished formation of F, giving luciferin as the sole product. Due to its unique structure, Pro residue may constrain the conformation of the short peptide and prohibit the hydrophobic interaction between the Tyr side chain and the CBT moiety. Collectively, these results suggest that these hydrophobic interactions favor the nonluciferin pathway.

Table 4-5. Outcomes of CBT conjugation with various peptides.

Peptides	CIA	CIY	CIL	CIW	CIF	CICha	CIT	CAY	CPY	CY	CIS	CIDap	CIK	CVS	CAS	CPS	CS
 D	100	50	78	56	76	51	96	54	100	87	17	100	100	80	100	100	60
 F	-	50	22	43	24	49	4	46	-	13	-	-	-	-	-	-	-
 H	-	-	-	-	-	-	-	-	-	-	83	-	-	20	-	-	40

The reaction was performed in PBS buffer pH 8.5 with 100 uM peptides and 500 uM CBT. D, F and H (%) formation was determined by LC-MS.

We further introduced hydrophilic amino acids into X position. As expected, CIK and CI(Dap) both gave the luciferin product D exclusively. However, we found that CIS (Cys-Ile-Ser) conjugated with 5 equivalents CBT to give major product as a disulfide linked dimer H (Figure 4-13a, Table 4-5). CIS-H is very stable and can be purified by RP-HPLC, which was characterized by ^{1}H-NMR and UV-Vis spectrophotometer (Figure 4-13). As shown in Figure 4-8 b, the aromatic peaks of CIS-H and CIS-D had different chemical shifts. According to the absorbance at 280 nm, we calculated the molar extinction coefficient of CIS-D and CIS-H respectively to determine the percentage of CIS-H

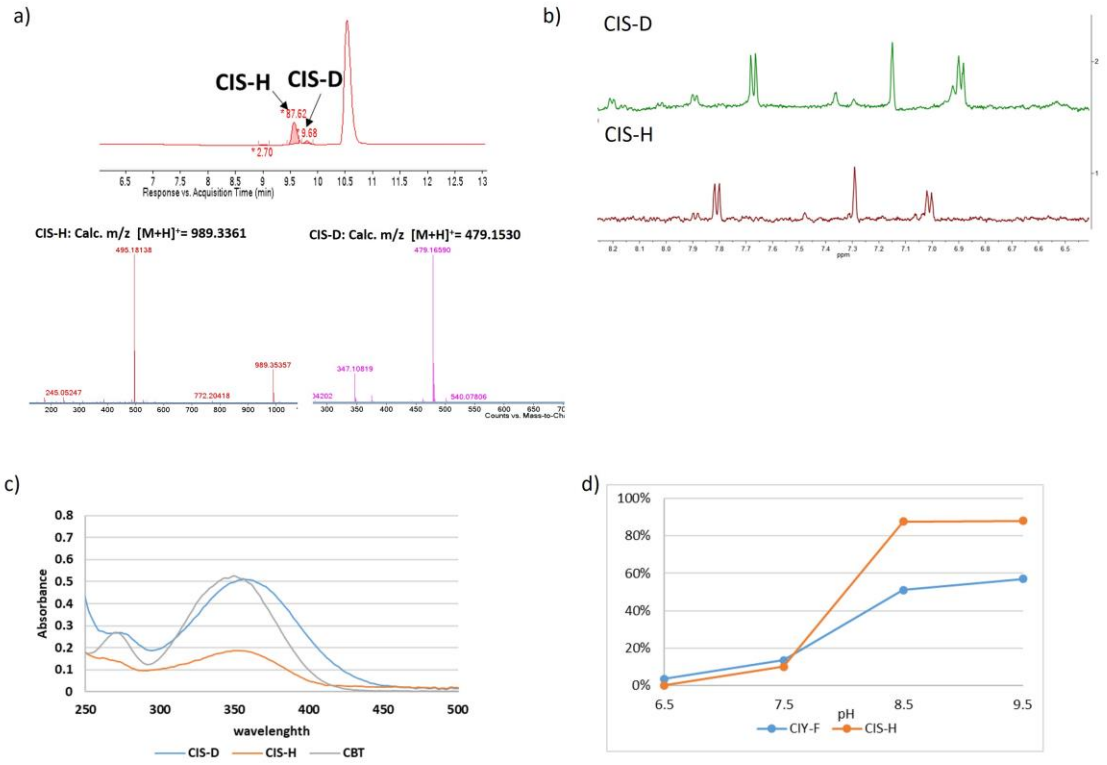


Figure 4-13. Characterization of the reaction between CIS and CBT. a) LC-MS analysis of the reaction between CIS and 5 eq CBT at pH 8.5. b) ^{1}H-NMR spectra of CIS-D and CIS-H. (c) UV-vis spectra of CIS-D, CIS-H and CBT in PBS buffer (40 μM, pH 7.4). c) Comparison of CIY and CIS in terms of the pH dependence of their conjugation with CBT.

formation (Figure 4-13c). In contrast to the hydrophobic CIX variants, only 17% CIS formed the luciferin product D, with the remaining product (83%) appearing as a CIS-H (Figure 4-13, Table 4-5). This surprising result indicates the CIS peptide strongly disfavors luciferin formation by promoting thiol elimination of C to allow disulfide bond formation.

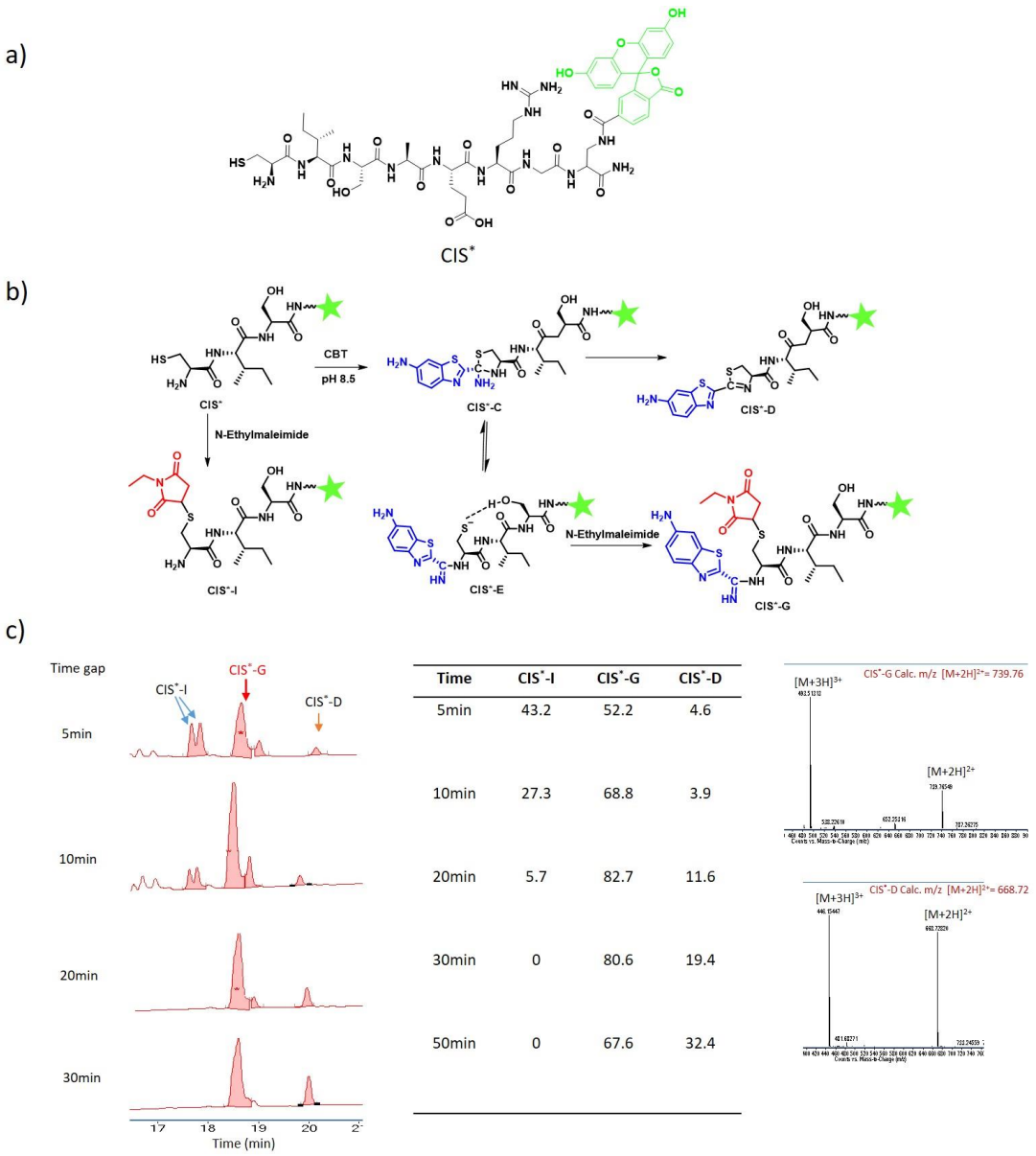
Remarkably, CIS is quite unique in this behavior as Mutating CIS to CVS dramatically reduced the percentage of H formation to 20%, and further CAS and CPS did not give H at all. We postulated that the CIS tripeptide adopts a unique structure that favors hydrogen bonding interaction between the Ser side chain and the thiolate and this stabilization favors thiol elimination over luciferin formation. The proposed mechanism is supported by the sharp pH dependence of the CIS-CBT conjugation: the disulfide dimer H was observed at only 7% (93% luciferin product D) at pH 7.5, in contrast to 83% at pH 8.5 (Figure 4-13d). In comparison, a much shallower pH dependence was observed for the CBT conjugation with CIY, which gave the bis adduct CIY-F as the main product.

Overall, these results revealed a significant sequence dependence in terms of reaction outcomes. The hydrophobic interactions and hydrogen bond can favor nonluciferin pathway with specific conformations.

4.4.4 Dual modification of peptides and proteins with a CIS tag

We envisioned that the transiently formed E, instead of forming disulfide or conjugating with another CBT, could be captured by a strategically designed electrophile to afford an N, S-doubly labelled NCys (Figure 4-14). We first test this hypothesis on a fluorophore labelled model peptide with CIS on its N-terminus (Figure 4-14a). TCEP was added in the reaction mixture to avoid oxidation and only 2 equivalents of CBT were used. N-ethylmaleimide was used as a test reagent to capture the free thiol. Due to the necessity of

the NCys thiol for the initiation of the CBT conjugation, N-ethylmaleimide was added a later time to allow CBT conjugation with NCys. Not surprisingly, the time gap between reaction initiation and N-ethylmaleimide addition is critical (Figure 4-14c): adding N-ethylmaleimide too early (5 min) resulted in significant formation of CIS-I (43%), which is no longer able to conjugate with CBT. Adding N-ethylmaleimide too late (50 min) results in increased presence of the luciferin product CIS-D. 20 to 30 min of time gap appeared to be optimal, yielding over 80% of the N, S- double labelled product CIS-G, a small percentage of the luciferin product CIS-D, and little to no CIS-I (Figure 4-14c). The results in Figure 4-9c, also indicates a $t_{1/2}$ of ~5min for the amidine formation. Based on this, we estimate a rate constant of ~20 M^{-1} s^{-1} for the amidine formation, which is on par with the CBT conjugation to give luciferins.

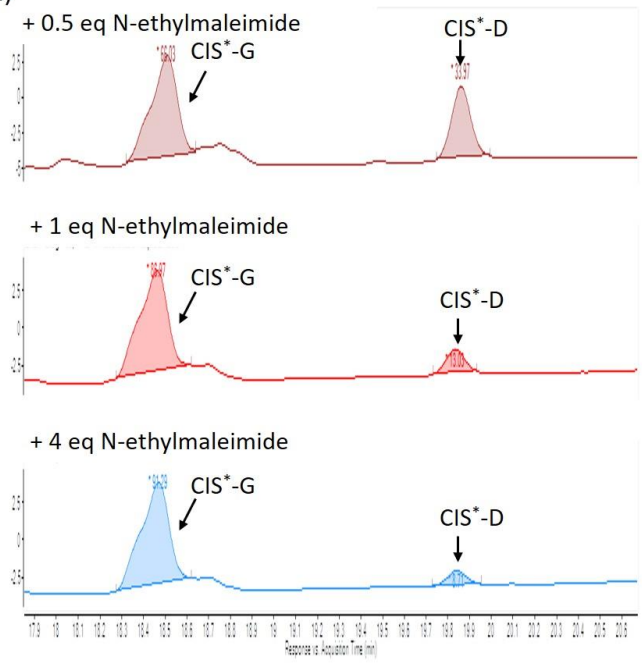


Time	CIS*-I	CIS*-G	CIS*-D
5min	43.2	52.2	4.6
10min	27.3	68.8	3.9
20min	5.7	82.7	11.6
30min	0	80.6	19.4
50min	0	67.6	32.4

Figure 4-14. N, S-double labeling of a CIS peptide. a) Structure of the fluorescein labeled CIS peptide. b) Postulated reaction pathways of CBT conjugation with a CIS* peptide leading to varied reaction outcomes. c) LC-MS characterization of CIS* (100 μM) sequentially treated with CBT (200 μM) and N-ethylmaleimide (100 μM) with varied time gap. CIS*-I appeared as a pair of peaks presumably due to the generation of diastereomers at the newly generated chiral center on the succinimide moiety of CIS*-I. Percentages of product composition of the reactions shown in the table.

We also optimized the equivalent of N-ethylmaleimide in the reaction. As shown in Figure 4-15, when there was not enough alkylating reagent, 0.5 equivalent, a fraction of CIS* would form CIS*-D instead. Use of N-ethylmaleimide at 4 equiv increased the yield of CIS*-G to 91%.

a)



b)

Equivalence of N-ethylmaleimide	CIS*-G	CIS*-D
0.5	66.0	34.0
1	87.0	13.0
4	91.3	8.7

Figure 4-15. The concentration effect during the thiol-Michael addition to the cysteine. (a) LC-MS (Method III) analysis of double labeling of CIS* peptide with 0.5 eq, 1 eq and 4 eq N-ethylmaleimide addition. The data were recorded by monitoring absorbance at 450 nm (b) Percentages of product composition of the reactions.

We further studied the effect of a competitor CIA. Interestingly, the efficiency of CIS* double labeling was only slightly reduced with the addition of 2 equivalents of CIA as a competitor (Figure 4-16), indicating CBT preferentially conjugates with CIS* over other NCys residues. Finally, and importantly, we found that the doubly labeled product CIS*-G displays robust stability under physiologic conditions with no decomposition over 24 h, even in the presence of 10 mM glutathione (Figure 4-17).

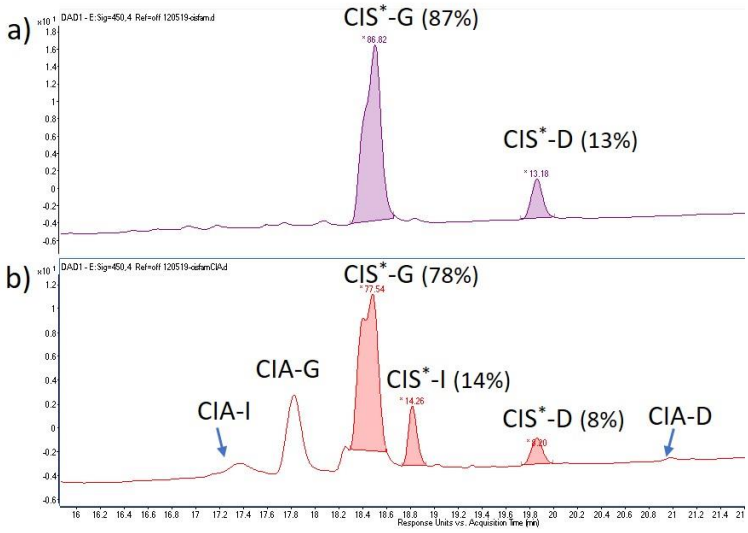


Figure 4-16. The effect of CIA peptide under reaction conditions. a) CIS* peptide (50 μM) was incubated with 2eq (100 μM) CBT at pH 8.7 for 30 min followed by the addition of 4 eq (200 μM) N-ethylmaleimide. Under same conditions, 2eq (100 μM) CIA was added in (b) before the addition of CBT. The product composition of the reactions was analyzed by LC-MS. The data were recorded by monitoring absorbance at 450 nm. In the presence of 2 eq CIA, 15.6% CIS* didn't react with CBT and therefore formed CIS*-I.

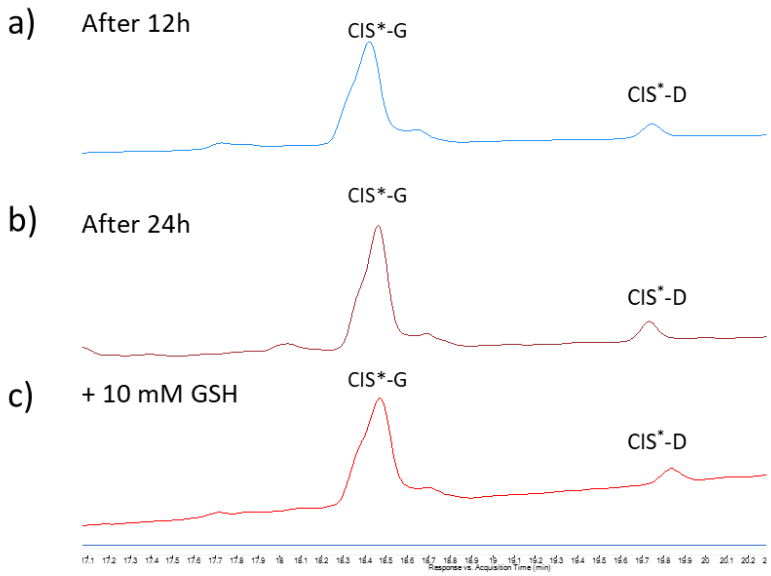


Figure 4-17. Stability of CIS*-D. a) LC trace of a doubly labeled CIS* (treated with CBT and then N-ethylmaleimide) after 12 h incubation. The data were recorded by monitoring absorbance at 480 nm and (b) the same reaction of (a) after 24 h incubation. c) The same reaction after 24 h incubation with 10 mM GSH. The data presented in this figure show the stability of the doubly labeled product.

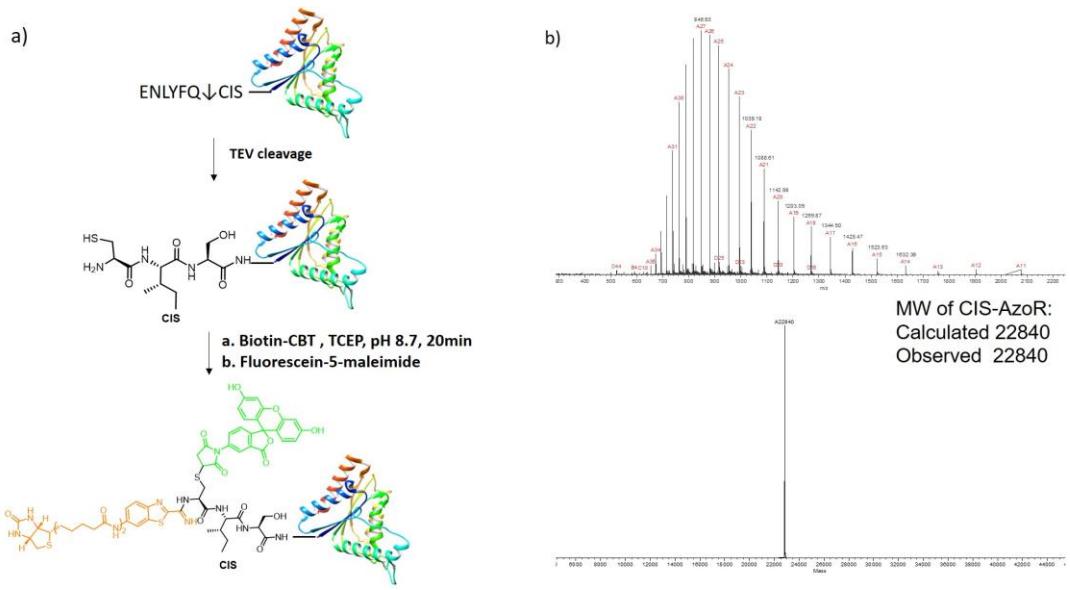


Figure 4-18. N, S-double labeling of proteins with an N-terminal CIS tag. a) Schematic illustration of the preparation and labeling of CIS-AzoR. b) Raw and deconvoluted mass spectrum of CIS-AzoR.

To further demonstrate the utility of this CIS dual labeling strategy for protein modification, a CIS tag was introduced at the N-terminus of a model protein, *E. coli* azoreductase (AzoR, Figure 4-18a). The tobacco etch virus (TEV) recognition epitope (ENLYFQ↓CIS) was fused to the N-terminus of AzoR followed by the peptide tag CIS. The recombinant protein was expressed and purified by Ni-NTA column. TEV protease cleavage was done under the dialysis conditions to generate the AzoR with an N-terminal CIS tag (CIS-AzoRs), which was finally purified by gel filtration and confirmed with ESI-MS (Figure 4-18b). To demonstrate N, S-double labeling of CIS-AzoR, the protein was first incubated with biotin-CBT at pH 8.7 with TCEP for 20 min followed by the addition of fluorescein-5-maleimide. LC-MS analysis showed that the major product (>90%) corresponds to the doubly labelled CIS-AzoR with a minimal luciferin product (Figure 4-19). The high efficiency of N, S-double labeling was further corroborated using a gel shift assay, in which the labeled CIS-AzoR was incubated with streptavidin and then subjected

to SDS-PAGE analysis (Figure 4-19c). The fluorescence image of the gel revealed little protein remaining at the expected molecular weight of AzoR. Instead, the protein

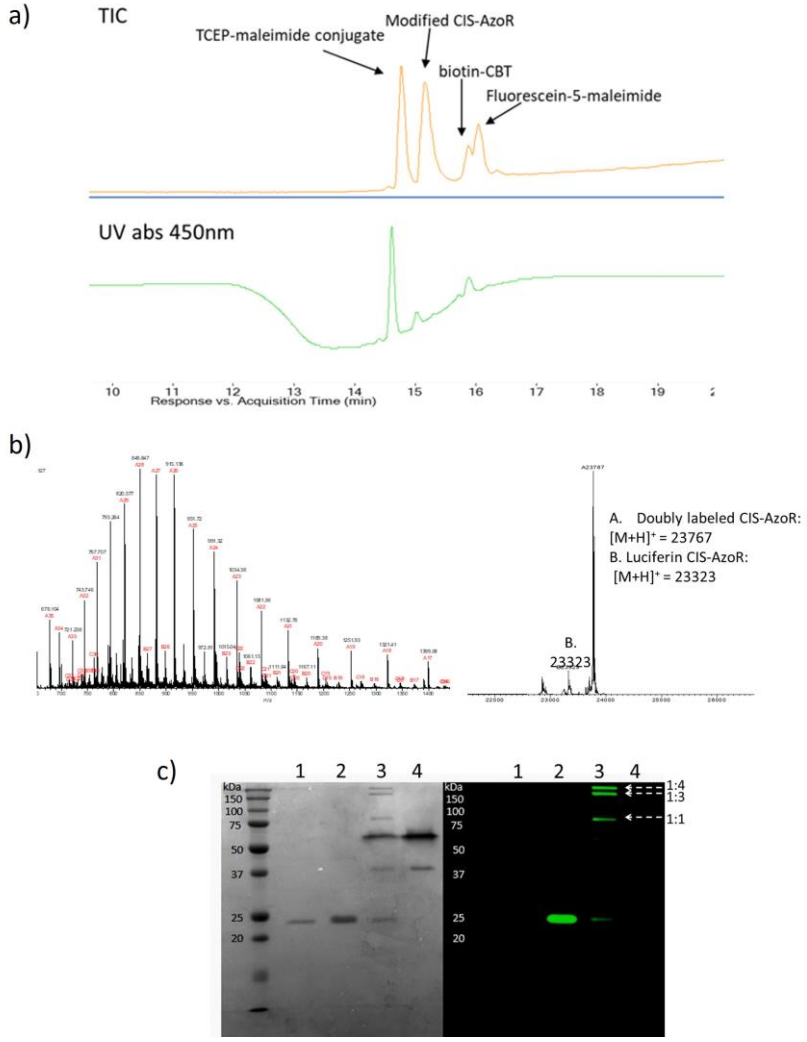


Figure 4-19. N, S-double labeling of CIS-AzoR. a) LC-MS analysis of double labeling of CIS-AzoR. The data were recorded by monitoring the TIC signal (top) and the absorbance at 450 nm (bottom) (b) Raw mass-spec data and (c) the deconvoluted mass. The modified CIS-AzoR. Peak A corresponds to the doubly labeled CIS-AzoR with the calculated mass 23766 and observed mass 23767. Peak B corresponds to luciferin CIS-AzoR with the calculated mass 23322 and observed mass 23323. (c) Fluorescence (right) and Coomassie-staining (left) images of a gel loaded with CIS-AzoR after sequential labeling by Biotin-CBT and fluorescein-5-maleimide. Lane 1, CIS-AzoR; lane 2, CIS-AzoR subjected to double labeling conditions; lane 3, CIS-AzoR subjected to double labeling and then mixed with streptavidin; lane 4, streptavidin alone. Lane 3 shows the formation of 1:1, 1:3, and 1:4 complexes between streptavidin and the doubly labeled CIS-AzoR.

predominantly (>90%) appeared as complexes with streptavidin (Figure 4-19c). This result confirms efficient double labeling of CIS-AzoR as only the biotinylated protein can bind streptavidin to form complexes. In comparison, only the luciferin product was observed for the negative control Cys-AzoR, which presents a single cysteine instead of the CIS tag on the N-terminus (Figure 4-20).

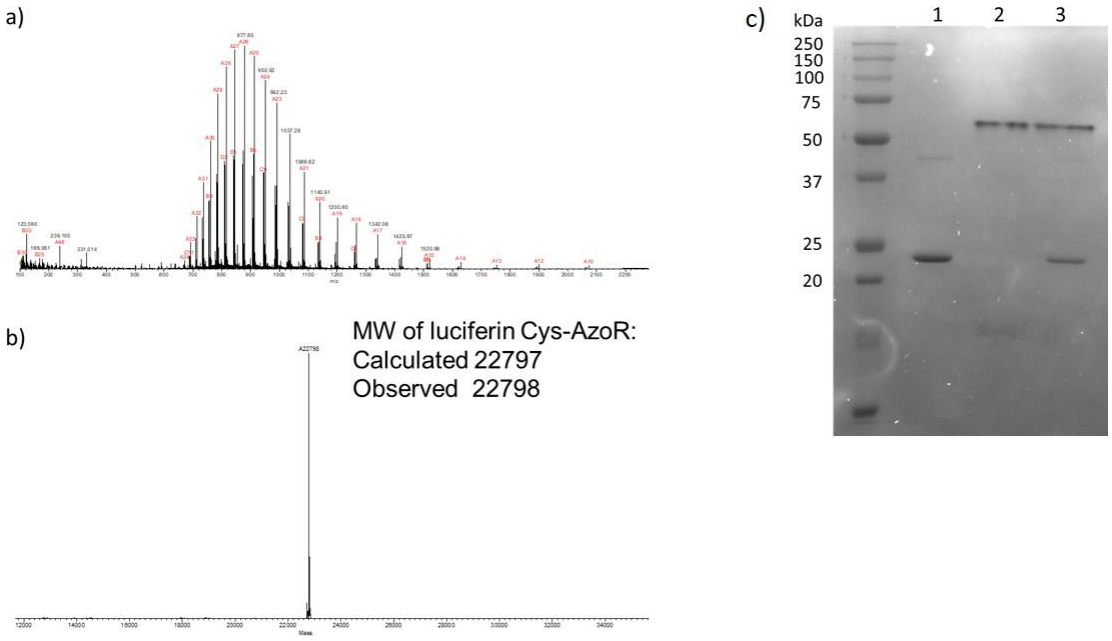


Figure 4-20. Modification of Cys-AzoR. Raw mass-spec data. a) and the deconvoluted mass (b) of luciferin Cys-AzoR. b) Coomassie-staining images of a gel loaded with unmodified CIS-AzoR. Lane 1: CIS-AzoR; lane 2: streptavidin alone; lane 3: CIS-AzoR mixed with streptavidin. No formation of complexes between streptavidin and unmodified CIS-AzoR was shown in lane 3.

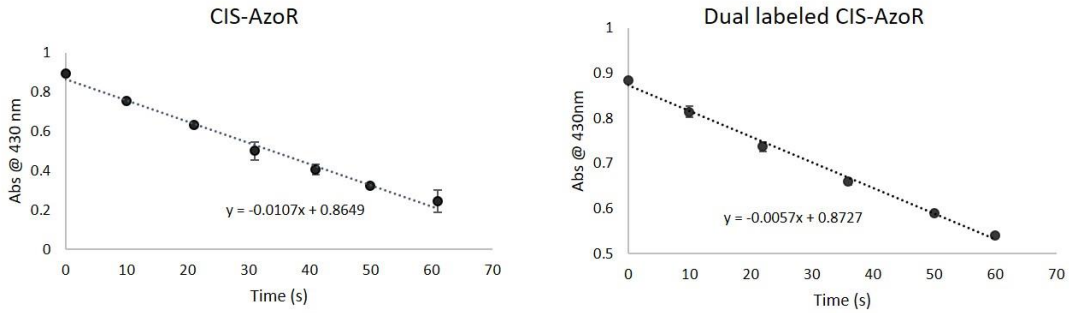


Figure 4-21. Activity of CIS-AzoR and Dual labeled CIS-AzoR. The decrease in absorbance at 430 nm versus time was monitored.

To better define the scope of applications of this NCys double labeling strategy, we tested whether the double labeling protocol would have any negative impact on the enzymatic activity of AzoR. The E. coli AzoR is an FMN-dependent NADH-azoreductase, which can degrade methyl red. We determined the enzyme activity by monitoring the decrease in absorbance at 430 nm. To our satisfaction, the doubly labeled CIS-AzoR showed noncompromised enzymatic activity (73.2 units/mg) in comparison to the unmodified CIS-AzoR (68.7 units/mg) (Figure 4-21). These values are consistent with a prior literature report, where the wild type enzyme activity was estimated to be 63.9 units/mg. These results show that the N, S-double labeling protocol is well tolerated by protein structures, which suggests potentially broad applications of this chemistry.

4.5 Conclusions

In summary, this chapter described the strategy to construct a phage display peptide library presenting NCys for site-specific modification. The library showed good quality and large diversity. The protease can recognize and cleave the phage-encoded peptides efficiently. To chemically modify this library, a novel conjugation reaction for fast and

highly specific NCys modification was applied, which proceeds via rapid formation of a TzB intermediate followed by intramolecular acyl transfer to give stable N-acylthiazolidines. This reaction shows the fast kinetics and exquisite NCys selectivity, and enables efficient and site-specific phage modification with low micromolar concentration of reagents. In addition, CBT-NCys condensation reaction was also performed to install non-natural functionalities to phage in a site-specific manner. Furthermore, the internal Cys can be subsequently alkylated to install anther functionality. Therefore, a strategy to introduce two distinct functionalities to phage display library in a site-specific way was developed.

Even though CBT-NCys conjugation usually gives a luciferin product at neutral pH, we discovered an alternative reaction pathway that proceeds under basic conditions and leads to rapid formation of N-terminal amidine, leaving the side chain thiol free for further modification. Exploration of the peptide sequence dependence reveals that a tripeptide tag CIS strongly favors the amidine conjugation pathway, which enables N, S double labeling of peptides and proteins in excellent yields. The utility of the CBT-CIS conjugation is demonstrated by facile double labeling of azoreductase (AzoR) as a model protein. Remarkably, the mild labeling conditions elicited little reduction of the enzymatic activity of AzoR. We believe that this unique reaction pathway of CBT-NCys condensation presents a significant addition to the toolbox for site-specific protein modifications and is beneficial for creating chemically modified phage libraries.

4.6. Experimental Procedures

4.6.1 General Methods

Dawson Dbz resin, Fmoc-Osu and HBTU were purchased from Novabiochem. Fmoc-protected amino acids were purchased from Advanced Chemtech or Chem-Impex Int'l Inc. Trifluoroacetic acid was purchased from Protein Technologies. ProTEV Plus was purchased from Promega Corporation. Illustra NAP-5 columns were obtained from GE Healthcare for desalting and buffer exchange. The Q5 Site-Directed Mutagenesis Kit was obtained from New England biolabs (NEB). Bacterial Sortase Substrate I, FRET was purchased from AnaSpec, Inc. Other chemicals were obtained from Fisher Scientific or Sigma Aldrich. Expression plasmid for *S. aureus* Sortase A (SrtA) was purchased from Addgene. DNA oligomers were ordered from Integrated DNA Technologies (IDT). Restriction enzymes were obtained from New England Biolabs (NEB). GeneJET Plasmid Miniprep Kit from Thermo Fisher was used to prepare plasmids for sequencing. Sanger Sequencing was done by Genewiz or Eton Bioscience Inc. HA Tag Monoclonal Antibody-HRP was obtained from Invitrogen. Peptide synthesis was carried out on a Tribute peptide synthesizer (Protein Technologies). ^{1}H NMR measurements were performed on a Varian INOVA 500 and 600 MHz spectrometer. A Nanodrop UV-vis spectrometer was used to measure the concentration of the fluorescein labelled compound. Mass-spec data were generated by an Agilent 6230 LC TOF mass spectrometer. HPLC purification was done on Waters 2695 Alliance HPLC system.

4.6.2 Construction of HA tag-IEGR-C(X)$_7$C phage library

The Ph.D. Peptide Display Cloning System (NEB #E8101S) was used to construct the peptide library. The insert oligonucleotide was purchased from IDT with the following

sequence:

CATGTTTCGGCCGAACCTCCACCACAMNNMNNMNNMNNMNNMNNMNNAC

ACCGGCCCTCAATTGCAGCGTAGTCTGGAACGTCGTAGGGGTAAGAGTGAGA

ATAGAAAGGTACCCGGG, in which the underlined positions were randomized. The insert was prepared by annealing the M13 extension primer to the oligonucleotide, extending with Klenow fragment (NEB #M0210), and digesting with EagI-HF (NEB #R3505) and KpnI-HF (NEB #R3142). The gel-purified insert was ligated to the EagI-HF/KpnI-HF digested M13KE vector by T4 DNA Ligase (NEB #M0202). 3 μg ligated DNA was electroporated into TOP10 Competent Cells (Thermo Fisher) to cover the complexity of 10^8 clones (electroporation efficiency: 10^8 pfu/μg DNA). Phage amplification was accomplished by incubating the electroporated cells with 1 L of early-log (OD600 0.01-0.05*) E. coli* ER2738 (NEB #E4104) cells in LB media with vigorous aeration (250 rpm) at 37°C for 4.5-5 h. Phage was recovered from the supernatant by adding 1/6 volume of 20% PEG/2.5 M NaCl, incubating overnight at 4°C and pelleting by centrifugation at 5000g for 45 min at 4°C. The phage was resuspended in TBS buffer pH 7.4 and stored at 4°C.

4.6.3 FXa cleavage on the phage display peptide library

5 μl of 1mg/mL FXa (NEB #P8010) was added into 1 mL phage solution (5×10^{12} pfu/mL) in the reaction buffer (20 mM tris-HCl, 100 mM NaCl, 2 mM $CaCl_2$, pH 8.0) and incubated for 6 h at room temperature. The phage after cleavage was collected by adding 1/6 volume of 20% PEG/2.5 M NaCl, incubating for 2h at 4°C and pelleting by centrifugation. The phage was resuspended in PBS buffer pH 7.4 and stored at 4°C. The phage titer was calculated according to the NEB M13 Titer Protocol.

To confirm FXa cleavage, we carried out an ELISA assay to confirm the removal of HA tag from phage. 100 µl 10 µg/mL M13 Major Coat Protein Antibody (RL-ph1, SCBT sc-53004) was added into each well of 96 Well EIA/RIA Assay Microplate (Corning #3361) and incubated overnight at 4 °C. The solution was removed by slapping plate face-down onto a paper towel. 100 µl blocking buffer (5 mg/mL BSA in TBS buffer pH 8.5) was added in the plate and incubated for 1 hr at room temperature followed by 6 times of washing with TBST. 100 µL 10^{11} pfu/mL phage treated with FXa or not was added into the blocked plate respectively and incubated for 1 hr at room temperature followed by 6 times of washing with TBST. 100 µL 1 µg/mL HA Tag Monoclonal Antibody-HRP (Invitrogen, # 26183-HRP) in blocking buffer was added in the plate and incubated for 1 h at room temperature followed by 6 times of washing with TBST. Finally, 100 µL TMB (Thermo scientific #34028) was added in the plate and incubated for 30 min at room temperature. 100 µL 2M sulfuric acid was added to stop the reaction. The absorbance of each well at 450 nm was measured as ELISA signal.

4.6.4 Chemical modification of phage display peptide library

The HA tag-IEGR-C7C phage after FXa cleavage (50 µL of 5×10^{12} pfu/mL) was subjected to reduction by mixing with 25 µL iTCEP (Thermo scientific #77712) in a total volume of 200 µL in PBS buffer (pH 7.4). The mixture was incubated at 4°C for 48 hr. After reduction, the phage solution was adjusted to pH 5 and then KL 72 (1 mM stock solution in DMF) was added to a final concentration of 20 µM. The labeling reaction was allowed to proceed at room temperature for 2 hr. The labeled phage was separated form iTCEP through centrifugation and the supernatant containing phage was precipitated by adding 1/6 volume of 20% PEG/2.5 M NaCl (2 hr at 4°C). The precipitated phage was

resuspended in 200 µl PBS buffer (pH 7.4) and the phage titer was found to be ~5×10^{11} pfu/mL.

As a positive control, the reduced phage was incubated with 1mM Biotin-IA (BIA, from a 100 mM DMF stock) for 2 hr in PBS buffer (pH 8.5) followed by precipitation. The phage was resuspended in 200 µl PBS buffer (pH 7.4) and the phage titer was found to be ~5×10^{11} pfu/mL.

As a negative control, the reduced phage was first treated with KL42 (20 µM, from a 1 mM DMF stock) at pH 5, room temperature for 2 h followed by subsequent labeling with KL72 (20 µM, from a 1 mM DMF stock). The phage was precipitated and resuspended in 200 µL PBS buffer (pH 7.4) and the phage titer was found to be ~3×10^{11} pfu/mL.

The Ph.D.-C7C Phage Library from NEB, which has an NAla instead of NCys, was used as an additional negative control. The AC7C phage was subjected to the same reduction and chemical modification as described above.

The phage was modified with CBT reagents in the same way except that 0.5 mM CBT or BCBT was used. 1mM free cysteine was added afterwards to quench the reaction. 2 mM BIA was added to modified the internal cysteine. Due to the low solubility of DCBT, only 50 µM DCBT was added with 10% DMF in the PBS buffer.

4.6.5 ELISA for assessing phage modification

To assess the extent of phage modification, a phage ELISA was designed and performed. 100 µL 100 µg/mL Streptavidin (NEB #N7021) was added into each well of 96 Well EIA/RIA Assay Microplate (Corning #3361) and incubating overnight at 4°C. The solution was removed by slapping plate face-down onto a paper towel. 100 µL blocking buffer (5

mg/ml BSA in TBS buffer pH 8.5) was added in the plate and incubated for 1 h at room temperature followed by 6 times of washing with TBST. 100 μL 10^9 pfu/mL chemically modified phage was added into the blocked plate and incubated for 1 h at room temperature followed by 6 times of washing with TBST. 100 μL 1000×dilution of M13 Major Coat Protein Antibody-HRP (GE Healthcare, 27-9421-01) in blocking buffer was added in the plate and incubated for 1 h at room temperature followed by 6 times of washing with TBST. Finally, 100 μL TMB (Thermo scientific #34028) was added in the plate and incubated for 30 min at room temperature. 100 μL 2M sulfuric acid was added to stop the reaction. The absorbance of each well at 450nm was measured as ELISA signal.

4.6.6 Conjugation of CBT and CXX peptides

Short peptides were synthesized on Rink Amide MBHA resin and purified by HPLC. 1 μL of 50 mM CBT stock solution in DMF was added into 100 μL of 100 μM CXX peptides solution in PBS buffer (final CBT concentration was 500 μM). The reaction was analyzed by LC-MS (Method II) after 2 h. CIS-D and CIS-H were purified by HPLC. HRMS (ESI) of CIS-D, calcd for $C_{20}H_{27}N_6O_4S_2^+$ 479.1530, found 479.1659. HRMS (ESI) of CIS-H, calcd for $C_{40}H_{57}N_{14}O_8S_4^+$ 989.3361, found 989.3536.

The product abundance was calculated by the UV absorbance at 280 nm by using the molar extinction coefficients ($M^{-1}cm^{-1}$) as following: Tyrosine, 1280; Tryptophan, 5690; Luciferin conjugate, 2964; CBT amidine conjugate, 2221.

4.6.7 Protocol for double labeling of a model CIS peptide CISAERGDap-FAM

Solid phase peptide synthesis was performed on a rink amide resin using Fmoc chemistry. An alloc-protected diaminopropionic acid (Dap) residue was installed at the C-terminus

for on-resin coupling of the fluorophore, 5(6)-FAM. The peptide was purified by HPLC and confirmed by LC-MS. The calculated mass was 1177.45 and the observed mass was 1178.38. 100 μM peptide solution with 1 mM TCEP in PBS buffer pH 8.5 was prepared. To 500 μL peptide solution, 5 μL of 20 mM CBT stock solution in DMF was added (final concentration was 200μM). The mixture was divided into 5 microcentrifuge tubes and 1 μL of 50 mM N-ethylmaleimide stock solution in DMF was added into each tube after 5, 10, 20, 30 and 50 min respectively. The reactions were analyzed by LC-MS after 1 h. Selected samples were further analyzed after 12 and 24 h incubation to probe the stability of the conjugates. The stability in the presence of 10mM GSH was also analyzed.

4.6.8 Preparation of azoreductase with N-terminal cysteine (C-AzoR) and CIS tag (CIS-AzoR)

pET28a-TEV-Cys-AzoR was obtained as a gift from Prof. Hua Lu at Peking University. The protein has an N-terminal GMENLYFQ↓C peptide sequence and a C-terminal His_6 tag. The plasmid was transformed into BL21(DE3) competent *E. coli*. The cells were cultured in 125 mL LB media with kanamycin (100 μg/mL) to OD_{600} = 0.5-0.8 followed by the addition of isopropyl β-D-1-thiogalactopyr-anoside (IPTG, 1 mM). The cells were cultured at 30 ºC, 250 rpm for 5 h and harvested by centrifugation (5000 g for 20 min) at 4 ºC. The cell pellet was resuspended in 3 mL resuspension buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM imidazole, 10% (vol/vol) glycerol) and lysed by sonication. The cell debris were removed by centrifugation (5000 g for 20 min) at 4 ºC. The filtered supernatant was loaded onto Ni-NTA column, washed by 1mL washing buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM imidazole, 10% (vol/vol) glycerol)) and eluted

by 600 μL elution buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 250 mM imidazole, 10% (vol/vol) glycerol). The expression yield of TEV-Cys-AzoR was 30 mg protein per liter of LB media. The molecular weight of the protein was confirmed by LC-MS: calculated 23624, found 23624.

The plasmid of TEV-CIS-AzoR was constructed by Q5® Site-Directed Mutagenesis Kit. The forward primer sequence was: ATTTCTGGTATGAGCAAGGTATTAG. The reverse primer sequence was: ACATTGGAAGTACAGGTTC. The protein was produced in the same way with the yield of 72 mg per liter of LB media. The molecular weight of the protein was confirmed by LC-MS: calculated 23824, found 23824.

The TEV protease cleavage of TEV-Cys-AzoR and TEV-CIS-AzoR was done under the dialysis conditions. In a 1000MWCO dialysis bag, 0.5mg protein and 50U of Pro TEV Plus protease were mixed in 100μL 1XPBS buffer pH 7.4 containing EDTA (0.5 mM), and DTT (1 mM). The dialysis bag was gently stirred in 300mL dialysate (1XPBS buffer pH 7.4, 0.5mM EDTA, and 1mM DTT) for 3h at room temperature. The reaction was monitored by LC-MS and showed quantitative conversion. The proteins were further purified by gel filtration (illustra NAP-5 columns). The molecular weight of the Cys-AzoR was confirmed by mass-spec: calculated 22640, found 22641. The molecular weight of the CIS-AzoR was confirmed by mass-spec: calculated 22840, found 22840.

Sequences of recombinant proteins:

TEV-Cys-AzoR

GMENLYFQCGMSKVLVLKSSILAGYSQSNQLSDYFVEQWREKHSADEITVRDLA
ANPIPVLDGELVGALRPSDAPLTPRQQEALALSDELIAELKAHDVIVIAAPMYNFN

ISTQLKNYFDLVARAGVTFRYTENGPEGLVTGKKAIVITSRGGIHKDGPTDLVTP

YLSTFLGFIGITDVKFVFAEGIAYGPEMAAKAQSDAKAAIDSIVSAHHHHHH

TEV-CIS-AzoR

GMENLYFQCISGMSKVLVLKSSILAGYSQSNQLSDYFVEQWREKHSADEITVRD

LAANPIPVLDGELVGALRPSDAPLTPRQQEALALSDELIAELKAHDVIVIAAPMY

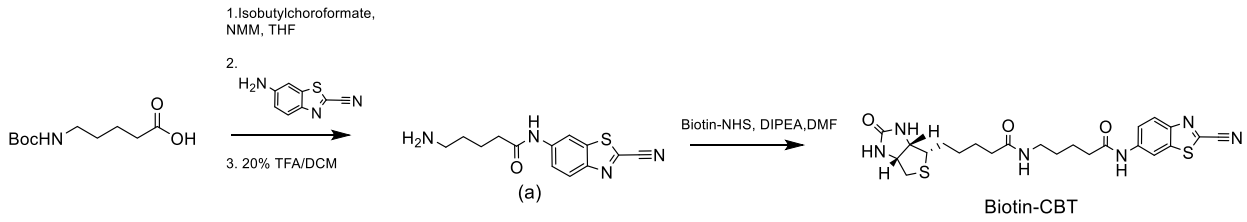
NFNISTQLKNYFDLVARAGVTFRYTENGPEGLVTGKKAIVITSRGGIHKDGPTDL

VTPYLSTFLGFIGITDVKFVFAEGIAYGPEMAAKAQSDAKAAIDSIVSAHHHHHH

4.6.9 Protocol for double labeling of CIS-AzoR

1 μL of 5 mM CBT or Biotin-CBT stock solution in DMF was added into 20 μL of 18 μM CIS-AzoR solution in PBS buffer pH8.7 with 1 mM TCEP. After 20 min, 1 μL of 8 mM N-ethylmaleimide or Fluorescein-5-maleimide was added. The reaction was monitored by LC-MS (Method I). As a negative control, the reaction was carried out with Cys-AzoR in the same way.

4.6.10 Synthesis of Biotin-CBT



Scheme 4-2. Synthesis of Bio-CBT.

a. Synthesis of 5-amino-N-(2-cyanobenzo[d]thiazol-6-yl)pentanamide (**a**): 52mg 5-((tert-butoxycarbonyl)amino)pentanoic acid (0.24mmol) was dissolved in 3mL THF with 38μL NMM (0.35mmol). 25μL isobutyl chloroformate (0.19mmol) was

added at 0°C under N_2 and the reaction was stirred for 30min on ice. 20mg 6-amino-CBT was added into the reaction mixture and stirred for 2h on ice and then room temperature overnight. Saturated sodium bicarbonate solution was added to the reaction. The reaction mixture was extracted with ethyl acetate, which was washed by brine, dried by Na_2SO_4 and evaporated. The crude product was dissolved in 2mL 20% TFA in DCM and stirred for 2h at room temperature. TFA and DCM was evaporated. 5-amino-N-(2-cyanobenzo[d]thiazol-6-yl)pentanamide was purified by HPLC.

^{1}H NMR (600 MHz, CD_3OH): 8.55 (s, 1H), 8.01 (d, J = 9.0 Hz, 1H), 7.60 (d, J = 9.0, 1H), 3.22 (s,2H), 2.91 (t, J = 7.2 Hz, 2H), 2.45 (t, J = 6.9 Hz, 2H), 1.77 – 1.63 (m, 4H).

^{13}C NMR (150 MHz, CD_3OH): 175.3, 150.9, 142.1, 139.3, 137.8, 127.2, 123.3, 115.3, 113.7, 41.7, 38.2, 29.3, 24.4.

HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{13}H_{15}N_4OS^+$ 275.0961; Found 275.0940.

b. Synthesis of Biotin-CBT: 19mg tert-butyl (5-((2-cyanobenzo[d]thiazol-6-yl)amino)-5-oxopentyl)carbamate (0.07mmol) and 25.8uL DIPEA (0.14mmol) were dissolved in 1mL DMF. 23.6mg Biotin-NHS (0.07mmol) was added and stirred for 2h at room temperature. Biotin-CBT was purified by HPLC.

^{1}H NMR (600 MHz, DMSO-d_6): 10.40 (s, 1H), 8.75 (s, 1H), 8.18 (d, J = 9.1 Hz, 1H), 7.79 (t, J = 5.6 Hz, 1H), 7.74 – 7.70 (m, 1H), 6.41 (s, 1H), 6.35 (s, 1H), 4.28 (dd, J = 7.1, 5.7 Hz, 1H), 4.11 (dd, J = 8.4, 3.6 Hz, 1H), 3.07 (dt, J = 13.6, 6.5 Hz,

3H), 2.79 (dd, J = 12.4, 5.1 Hz, 1H), 2.55 (d, J = 12.4 Hz, 1H), 2.39 (t, J = 7.4 Hz, 2H), 2.05 (t, J = 7.4 Hz, 2H), 1.66 – 1.56 (m, 3H), 1.53 – 1.40 (m, 5H), 1.36 – 1.21 (m, 2H).

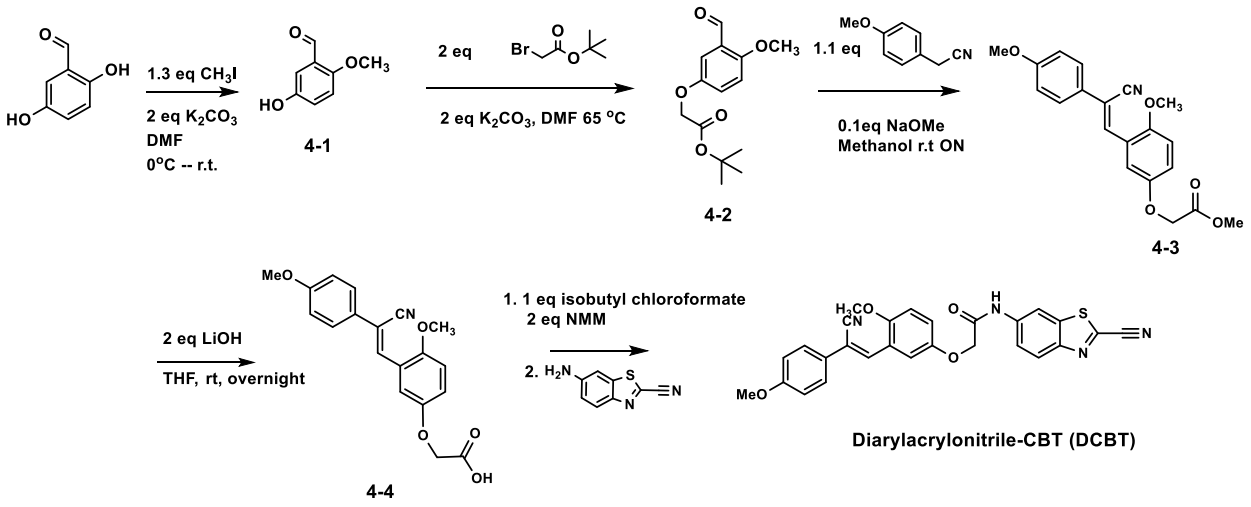
^{13}C NMR (150MHz, DMSO-d_6) δ 171.6, 171.6, 162.4, 147.2, 139.6, 136.5, 134.5, 124.5, 120.4, 113.3, 110.7, 60.8, 58.9, 55.2, 37.8, 35.8, 35.0, 28.5, 27.9, 27.8, 25.1, 22.2.

HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{23}H_{29}N_6O_3S_2^+$ 501.1737; Found 501.16914.

4.6.11 Azoreductase activity assay

The assay was done according to the reference 23 with following modifications. The activity of AzoR was determined in 500μL of 25 mM TrisHCl (pH 7.4) buffer containing 25 μM methyl red, 0.1 mM NADH, 20 μM FMN, and enzyme (0.2μg of CIS-AzoR, 0.1μg of dual labeled CIS-AzoR). The enzyme was added to initiate the reaction. The decrease in absorbance at 430 nm in the first 1 min at room temperature was used to calculate the initial reaction rate (the molar absorption coefficient: 23360 $M^{-1}cm^{-1}$). One unit of methyl red reductase activity was defined as the amount catalyzing the degradation of 1μmol of methyl red/min at room temperature.

4.6.12 Synthesis of Diarylacrylonitrile-CBT



Scheme 4-3. Synthesis of Diarylacrylonitrile-CBT

g. Synthesis of 4-1: 2,5-dihydroxybenzaldehyde (0.5 g, 3.62 mmol) and potassium carbonate (0.75 g, 5.43 mmol) were mixed in DMF (5 mL) and cooled on ice. Nitrogen gas balloon was attached to avoid water and oxygen. Iodomethane (294 µL, 4.71 mmol) was added into the mixture slowly. The reaction was stirred at room temperature for an overnight. Water (100 mL) was added into the reaction which was extracted with ethyl acetate for 3 times. The combined organic layers were washed with brine and dried over sodium sulfate. The solvent was then evaporated under vacuum. The crude product was purified on silica gel using hexane: ethyl acetate (8:2) as eluent. The desired product was obtained as white solid (55%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.41 (s, 1H), 7.30 (d, J = 3.2 Hz, 1H), 7.09 (dd, J = 8.9, 3.2 Hz, 1H), 6.91 (d, J = 8.9 Hz, 1H), 4.92 (s, 1H), 3.89 (s, 3H).

h. Synthesis of 4-2: compound 4-1 (100 mg, 0.65 mmol) and potassium carbonate (99 mg, 0.72 mmol) were mixed in DMF (1 mL). Nitrogen gas balloon was attached to avoid water and oxygen. Tert-Butyl bromoacetate (139 mg, 0.72 mmol) was added

slowly. The reaction was stirred at 65 ºC overnight. After the reaction was cooled to the room temperature, water (10 mL) was added. The mixture was extracted with ethyl acetate for 3 times. The combined organic layers were washed with brine and dried over sodium sulfate. The solvent was then evaporated under vacuum. The crude product was purified on silica gel using hexane: ethyl acetate (9:1) as eluent. The desired product was obtained as white solid (98%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.42 (s, 1H), 7.28 (d, J = 3.3 Hz, 1H), 7.21 (dd, J = 9.0, 3.3 Hz, 1H), 6.95 (d, J = 9.1 Hz, 1H), 4.51 (s, 2H), 3.90 (s, 3H), 1.49 (s, 9H).

i. Synthesis of 4-3: compound 4-2 (415 mg, 1.56 mmol) and 2-(4-methoxyphenyl)acetonitrile (210 µL, 1.56 mmol) were mixed in methanol (3 mL) with sodium methoxide (31 µL, 0.156 mmol). The reaction was stirred at room temperature overnight. The solvent was removed under vacuum. The mixture was purified on silica gel using gradient elution of 10% to 50% ethyl acetate in hexane. The desired product was obtained as Chartreuse green solid (23.4%).

   ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (s, 1H), 7.69 (d, J = 3.0 Hz, 1H), 7.63 – 7.58 (m, 2H), 7.02 (dd, J = 9.0, 3.0 Hz, 1H), 6.96 – 6.92 (m, 2H), 6.87 (d, J = 9.0 Hz, 1H), 4.67 (s, 2H), 3.84 (d, J = 4.0 Hz, 6H), 3.81 (d, J = 5.8 Hz, 3H).

j. Synthesis of 4-4: compound 4-3 (38 mg, 0.11 mmol) and lithium hydroxide (10.86 mg, 0.24 mmol) were dissolved in tetrahydrofuran (0.6 mL) with a drop of water. The reaction was stirred at room temperature overnight. LC-MS analysis showed no starting material left. 1N HCl was added to acidify the reaction which was then extracted with ethyl acetate. The combined organic layers were washed with brine and dried over sodium sulfate. The solvent was then evaporated under vacuum. The crude

product was purified on RP-HPLC. The fraction containing the desired product was collected and lyophilized. The product was obtained as green-yellow powder (90%). MS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{19}H_{18}NO_5^+$ 340; Found 340.

k. Synthesis of DCBT: compound 4-4 (22 mg, 0.06 mmol) was dissolved in tetrahydrofuran (1mL) with NMM (16 µL, 0.15mmol). Isobutyl chloroformate (7.9 µL, 0.06 mmol) was added at 0°C under N_2 and the reaction was stirred for 30min on ice. 6-amino-CBT (10 mg, 0.05 mmol) was added into the reaction mixture and stirred for 2h on ice and then room temperature overnight. Green-yellow precipitate was filtered, washed with dichloromethane and dried (80%). MS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{27}H_{21}N_4O_4S^+$ 497; Found 497.

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.71 (s, 1H), 8.74 (s, 1H), 8.16 (t, J = 8.1 Hz, 1H), 7.83 (s, 1H), 7.80 (d, J = 12.9 Hz, 1H), 7.61 (d, J = 8.6 Hz, 2H), 7.55 (d, J = 2.3 Hz, 1H), 7.16 – 7.09 (m, 1H), 7.07 (d, J = 9.1 Hz, 1H), 7.04 – 6.98 (m, 2H), 4.76 (s, 2H), 3.76 (dd, J = 17.7, 7.7 Hz, 6H).

4.7. References

1. Krall, N., et al., *Site-selective protein-modification chemistry for basic biology and drug development.* Nature Chemistry, 2016. **8**(2): p. 103-113.
2. Lambert, J.M. and A. Berkenblit, *Antibody–drug conjugates for cancer treatment.* Annual review of medicine, 2018. **69**: p. 191-207.
3. Xue, L., et al., *Imaging and manipulating proteins in live cells through covalent labeling.* Nature chemical biology, 2015. **11**(12): p. 917-923.
4. Volgraf, M., et al., *Allosteric control of an ionotropic glutamate receptor with an optical switch.* Nature chemical biology, 2006. **2**(1): p. 47-52.
5. Agarwal, P. and C.R. Bertozzi, *Site-specific antibody–drug conjugates: the nexus of bioorthogonal chemistry, protein engineering, and drug development.* Bioconjugate chemistry, 2015. **26**(2): p. 176-192.
6. Griffin, B.A., S.R. Adams, and R.Y. Tsien, *Specific covalent labeling of recombinant protein molecules inside live cells.* Science, 1998. **281**(5374): p. 269-272.
7. Adams, S.R., et al., *New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications.* Journal of the American Chemical Society, 2002. **124**(21): p. 6063-6076.
8. Hoffmann, C., et al., *Fluorescent labeling of tetracysteine-tagged proteins in intact cells.* Nature protocols, 2010. **5**(10): p. 1666-1677.
9. Zhang, C., et al., *π-Clamp-mediated cysteine conjugation.* Nature chemistry, 2016. **8**(2): p. 120-128.

10. Matos, M.J., et al., *Chemo- and Regioselective Lysine Modification on Native Proteins.* Journal of the American Chemical Society, 2018. **140**(11): p. 4004-4017.

11. Dawson, P.E., et al., *Synthesis of proteins by native chemical ligation.* Science, 1994. **266**(5186): p. 776-779.

12. Muralidharan, V. and T.W. Muir, *Protein ligation: an enabling technology for the biophysical analysis of proteins.* Nature methods, 2006. **3**(6): p. 429-438.

13. Saito, F., H. Noda, and J.W. Bode, *Critical Evaluation and Rate Constants of Chemoselective Ligation Reactions for Stoichiometric Conjugations in Water.* ACS Chemical Biology, 2015. **10**(4): p. 1026-1033.

14. Ren, H., et al., *A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins.* Angewandte Chemie International Edition, 2009. **48**(51): p. 9658-9662.

15. Ramil, C.P., et al., *Sequence-specific 2-cyanobenzothiazole ligation.* Journal of the American Chemical Society, 2016. **138**(17): p. 5499-5502.

16. Wang, W. and J. Gao, *N, S-Double Labeling of N-Terminal Cysteines via an Alternative Conjugation Pathway with 2-Cyanobenzothiazole.* The Journal of Organic Chemistry, 2020. **85**(3): p. 1756-1763.

17. Bandyopadhyay, A., S. Cambray, and J. Gao, *Fast and selective labeling of N-terminal cysteines at neutral pH via thiazolidino boronate formation.* Chemical science, 2016. **7**(7): p. 4589-4593.

18. Faustino, H., et al., *Iminoboronates are efficient intermediates for selective, rapid and reversible N-terminal cysteine functionalisation.* Chemical science, 2016. **7**(8): p. 5052-5058.

19. Dwyer, M.A., et al., *Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity.* Chemistry & biology, 2000. **7**(4): p. 263-274.

20. Oh, K.-B., et al., *Discovery of Diarylacrylonitriles as a Novel Series of Small Molecule Sortase A Inhibitors.* Journal of Medicinal Chemistry, 2004. **47**(10): p. 2418-2421.

21. Keyser, S.G.L., A. Utz, and C.R. Bertozzi, *Computation-Guided Rational Design of a Peptide Motif That Reacts with Cyanobenzothiazoles via Internal Cysteine–Lysine Relay.* The Journal of Organic Chemistry, 2018. **83**(14): p. 7467-7479.

Chapter 5 Developing a peptide-based G-FET sensor for detection of antibiotic resistant bacteria

5.1 Introduction

The over prescription and misuse of antibiotics is causing a surge in the number of antibiotic resistant bacterial infections around the world, which not only requires the development of new drugs, but also rapid, cheap, scalable, and accurate diagnostics. Label free biosensors have attempted to meet these requirements, however, the trade-off between selectivity and sensitivity of such sensors remains a key challenge. Towards achieving this goal, we applied peptide probes to the graphene field effect transistors (G-FETs) to electrically detect bacteria in a highly specific manner.

In this chapter, I will first briefly introduce G-FETs focusing on their applications for bacteria detection. Then the role of peptides as elegant probes to detect bacteria with specificity will be demonstrated. In collaboration with Dr. Narendra Kumar and Dr. Juan C. Ortiz-Marquez, we developed a novel G-FET-based biosensor equipped with peptide probes[1]. Utilizing our new biosensor and procedures, we demonstrated the first selective, electrical detection of the pathogenic bacterial species *Staphylococcus aureus* (*S. aureus*) and antibiotic resistant *Acinetobacter baumannii* (*A. baumanni*) on a single platform.

5.1.1. Graphene field effect transistors (G-FETs) for the detection of bacteria

Graphene field-effect transistor (GFET) is a kind of graphene-based bioelectronics that can transduce biomolecule charges or cellular voltage signals into a change in their current–voltage (*I–V*) characteristics[2]. Its high sensitivity, scalability, biocompatibility and ease

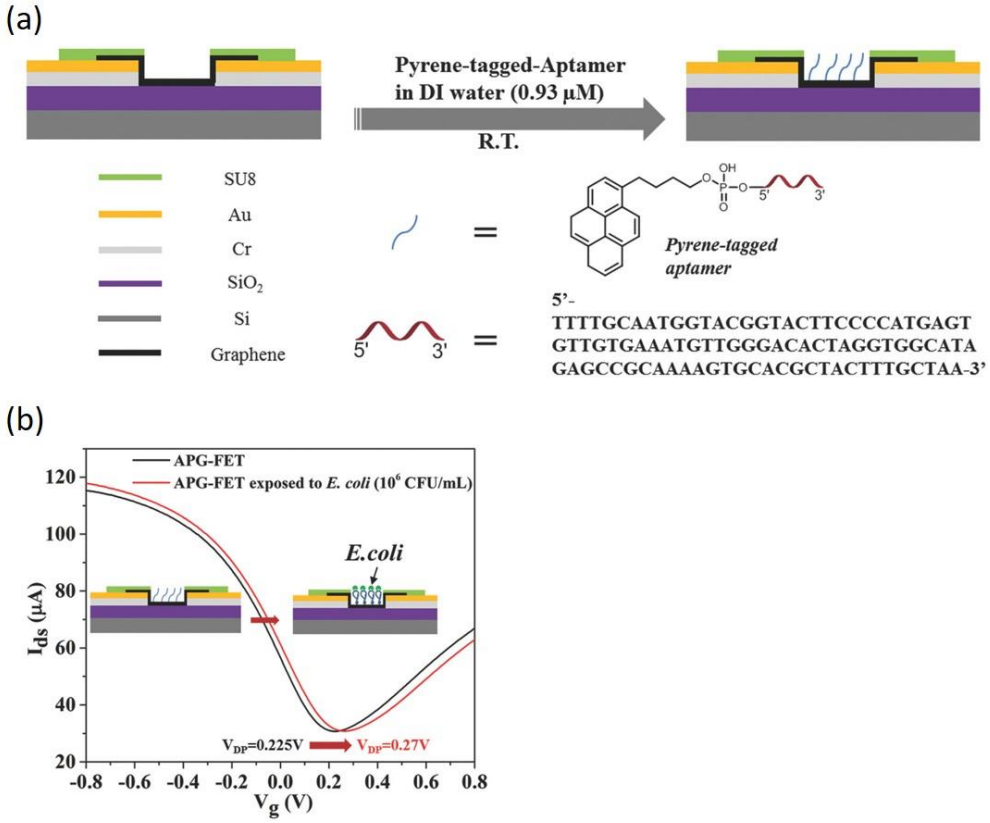


Figure 5-1. Aptamer-modified graphene field-effect transistors for the detection of *Escherichia coli*. a) Illustration of the construction of the G-FET. b) The attachment of *E. coli* induced the signal shift. Adapted from with permission Ref. [8].

of incorporation allow for the detection of charged biomolecules such as proteins and DNAs[3-6]. Given the negative surface charge of bacteria at physiological pH, G-FETs provide a promising platform for bacteria detection. Huang et al. demonstrated the detection of *E. coli* using the antibody-modified G-FET that[7]. A low detection limit of 10 cfu/mL was achieved in this work. In addition, an aptamer-modified G-FET was developed by Wu et al. to detect *E. coli* with excellent selectivity and a low detection limit of 100 cfu/mL (Figure 5-1)[8]. However, there is still an impediment to the broad use of G-FETs for detecting other bacterial species except *E. coli*. This is due to the lack of suitable probes which should be readily available, easily handled (simple preparation and/or long shelf

life), species, strain, and resistance-specificity[9, 10]. An equally crucial challenge is enhancing the sensitivity to achieve detection at a clinically relevant cell density. Even though antibodies have high affinity and selectivity to bacteria, it suffers from batch-to-batch variations and are not easy to store and transport. Also, the large size of antibody could make the target beyond the Debye screening length of the solution, resulting in reduced sensitivity[2]. Aptamer as an alternative has good stability and high affinity. However, the large charge induced by the probes themselves, may reduce the sensitivity to the additional charge of the target. Therefore, smaller size, neutral nature and high selectivity would be the desired features of optimal probes for utilization in G-FET based biosensors.

5.1.2. Peptide probes for bacteria

Peptide probes are superior to antibodies or aptamers, due to their small size, neutral nature (no net charge), long stability, and easy synthesis[11-14]. For example, antimicrobial

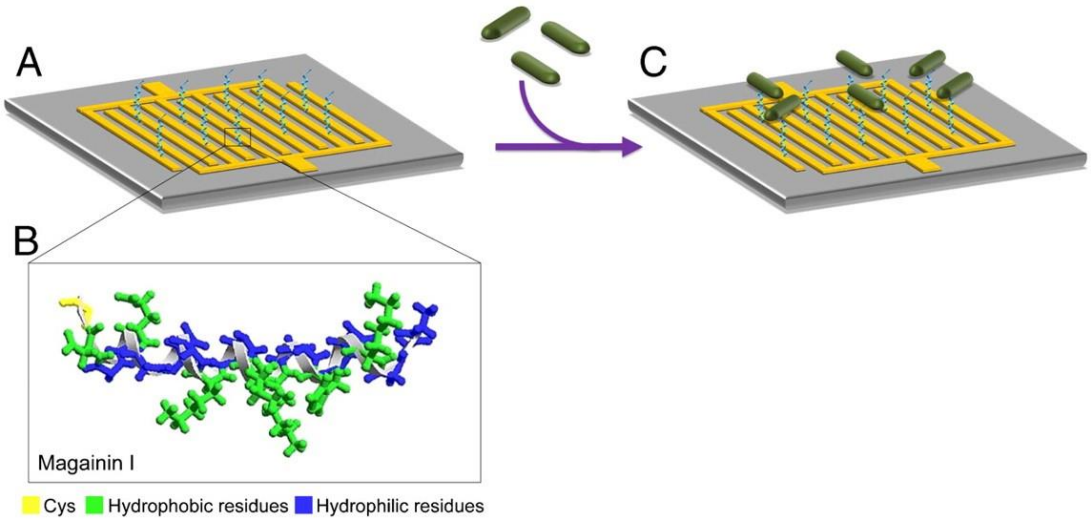


Figure 5-2. AMP-based electrical detection of bacteria. A. Schematic of AMPs magainin I immobilized on gold microelectrodes array. (*B*) Structure of the magainin I in helical form, modified with a terminal cysteine residue. (*C*) Demonstration of binding of target cells to the immobilized AMPs. Adapted with permission from Ref. [12].

peptides (AMPs) have been used as probes for bacteria detection. McAlpine and his coworkers were attracted by the positively charged magainin I that has broad-spectrum activity against Gram-negative bacteria. By immobilizing magainin I on gold microelectrodes via a C-terminal cysteine residue, detection of *E. coli* (1 cell/ μL) was achieved (Figure 5-2)[12]. Later, they functionalized graphene with odorranin-HP, an AMP with activity towards both the Gram-negative bacteria *E. coli* and *Helicobacter pylori*, and the Gram-positive bacteria *S. aureus*, and achieved the detection of bacteria at single-cell levels[15]. Furthermore, Etayash et al. successfully detected Gram-positive bacteria, such as Listeria monocytogenes, with the AMP leucocin A immobilized on the gold microelectrodes[16]. Even though these sensors showed high sensitivity, the strain selectivity still needs improvement. It is hard for AMPs to differentiate bacterial species or strains.

To solve this problem, our group developed a phage display platform that can rapidly select for small peptides that recognize and bind specific bacterial species or strains[17]. A

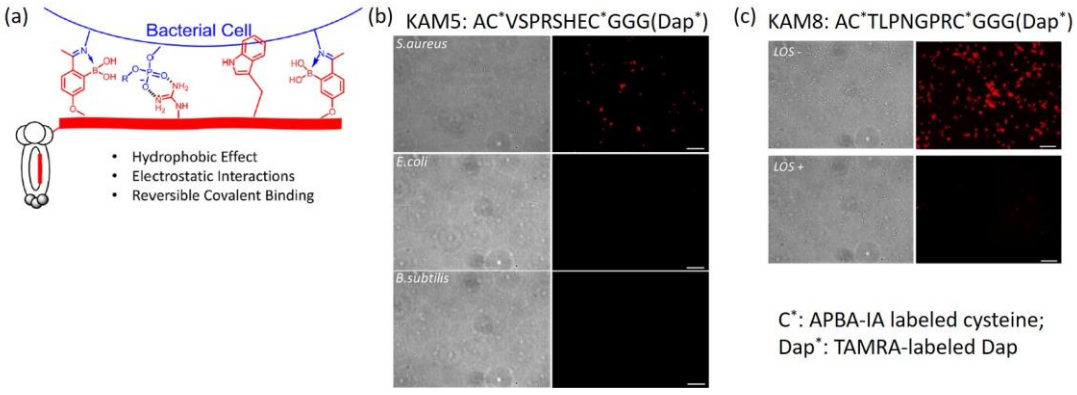


Figure 5-3. Peptide probes specific for S. aureus and a colistin-resistant strain of A. baumannii. a) Illustration of a modified phage binding to bacterial cells via iminoboronate formation. b) Microscopic images of bacterial cells stained with 10 μM KAM5 showed specific staining of S. aureus cells. c) KAM8 specifically stained A. baumannii (LOS-) strain. Adapted with permission from Ref. [17].

pair of 2-acetylphenylboronic acid (APBA) moieties was installed onto the peptides to bind biological amines via dynamic formation of iminoboronates (Figure 5-3a). The peptide KAM5 showed submicromolar affinity (EC_{50} of 1.5 μM) and high specificity for clinical strains of *S. aureus* (Figure 5-3b). Additionally, a peptide probe of a colistin-resistant strain of *A. baumannii* (LOS-) was identified, named KAM8. It gave a high affinity, EC50 of ~0.3 μM. Remarkably, KAM8 targets bacteria in a strain-specific manner as it showed no binding to the wild-type *A. baumannii* (LOS+) (Figure 5-3b). Therefore, we believed that G-FETs modified with these peptide probes would enable the detection of clinically relevant pathogenic bacteria and antibiotic resistant strains with high specificity.

5.2 G-FET device construction and baseline measurements.

G-FET devices are constructed by Narendra Kumar in Kenneth S. Burch lab, which consist of a low-pressure chemical vapor deposition (CVD) graphene on a standard SiO_2/Si substrate, and etched into an active area of 20 x 50 μm, with Cr/Au source and drain. The

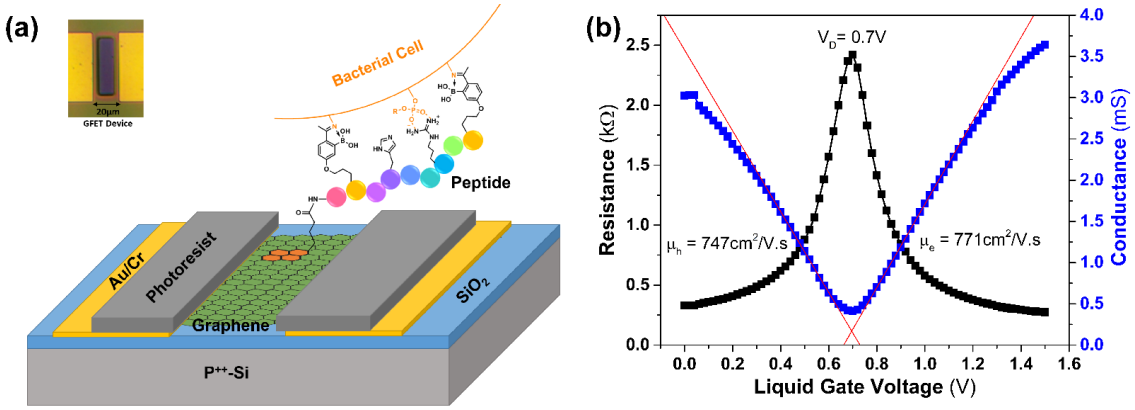
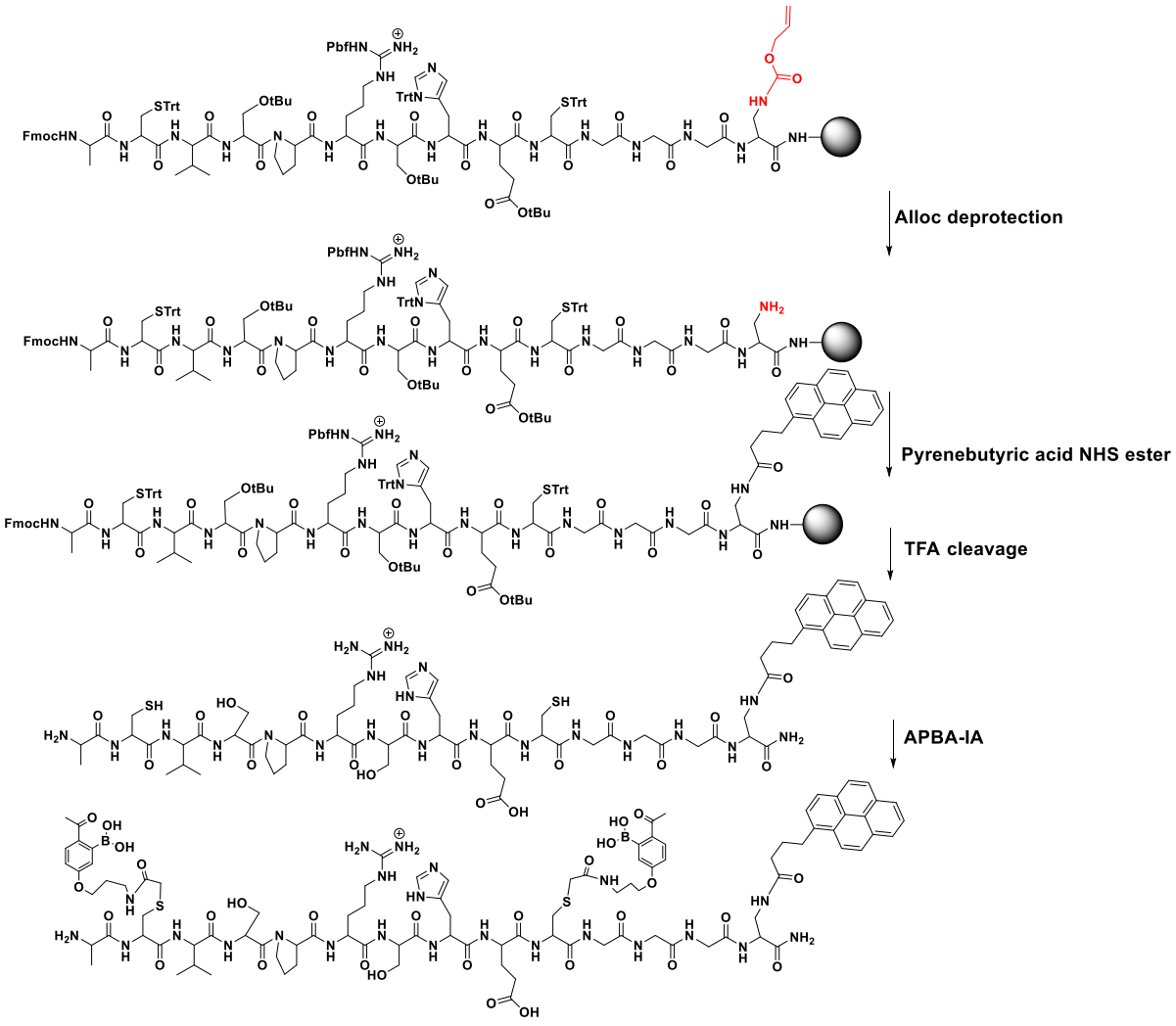


Figure 5-4. Scheme of functionalization of G-FETs. (a) A schematic of a G-FET functionalized with a pyrene-conjugated peptide probe binding to the surface of a bacterium. The inset shows a light microscopy image of G-FET, with an active area of 10 x 40 μm, located in between two gold contacts. (b) Resistance/conductance vs voltage plots of G-FET representing the Dirac voltage (0.7 V), hole and electron mobilities of 747 and 771 cm^2/V.s.

contacts were passivated and the sensing area (10 x 40 μm) was defined easily with a conventional hard baked photoresist (S1805) (Figure 5-4a). The baseline conductance/resistance of a device was measured in liquid gate mode. A Pt wire works as the reference electrode and 0.01x PBS as electrolyte (Figure 5-4b). The measured average Dirac voltage (V_D) of the fabricated G-FETs is around 0.7±0.16 V, consistent with the reported value[18]. Then, the mobility was calculated and consistent with reported values for CVD graphene on SiO_2 substrates[8, 19].

5.3 Design and synthesis of peptide probes for graphene field effect transistors (G-FET) modification

To immobilize the peptide probe onto the graphene without perturbing its electronic properties, we installed a pyrene moiety onto the C-terminus of the peptide, which allows the noncovalent functionalization of graphene through π-π stacking[20]. We synthesized a



Scheme 5-1. Synthesis of P-KAM5_Probe

peptide-pyrene conjugate (P-KAM5_Probe), by on-resin coupling of 1-pyrenebutyric acid N-hydroxysuccinimide ester onto the diaminopropionic acid residue installed at the C-

terminus of the peptide (Scheme 5-1). APBA-IA was used to modified the peptide. The pyrene-conjugated peptide dissolved in PBS buffer pH7.4 was incubated on the device for 2 h followed by a wash step. The uniform functionalization was characterized with atomic force microscopy (AFM; Figure 5-5). The height of graphene functionalized with P-KAM5_Probe increased by ~2.5 nm as compared to the bare graphene surface, which is consistent with peptides attached to carbon nanotubes or graphene oxide[21, 22]. The peptide-pyrene conjugates thus facilitate simplified graphene functionalization by a single step process, which enables rapid and easy preparation of the device as well as reduced fabrication cost.

P-KAM8_Probe was synthesized in the same way. Negative control peptides, named P-KAM5/8_Control were also made as comparison, which were modified by iodoacetamide instead of APBA-IA and incapable of binding bacteria.

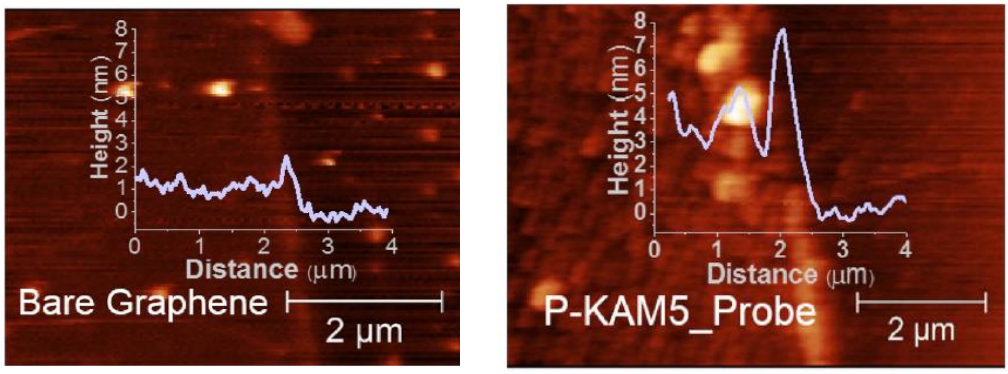


Figure 5-5. AFM image of the patterned graphene before and after peptide functionalization. After functionalization, the coverage of the graphene channel by the P-KAM5_Probe peptide is shown as an increase in height of ~2.5 nm (right panel) compared with the bare graphene surface (left panel).

5.4 Bacteria detection

5.4.1 Detection of S. aureus

First, we tested the detection of *S. aureus* with P-KAM5_Probe functionalized G-FETs. As shown in Figure 5-6a, the attachment of peptide itself didn't generate any shift in V_D, consistent with their charge neutral structure at pH 7. The incubation of *S. aureus* at 10^7 cells/ml on the G-FET resulted in a voltage shift of 300 mV (Figure 5-6a). The observed positive shift in the V_D is attributed to the negatively charged surface of bacterial cells which increase the hole carrier density in graphene[8]. In contrast, no notable voltage shift was observed when *S. aureus* was incubated on devices functionalized with P-KAM5_Control (Figure 5-6b). The capture of bacteria can be visualized and analyzed with optical microscopy. As shown in the Figure, the black dots represent individual bacterial cells that were only observed on devices functionalized with P-KAM5_Probe and not P-

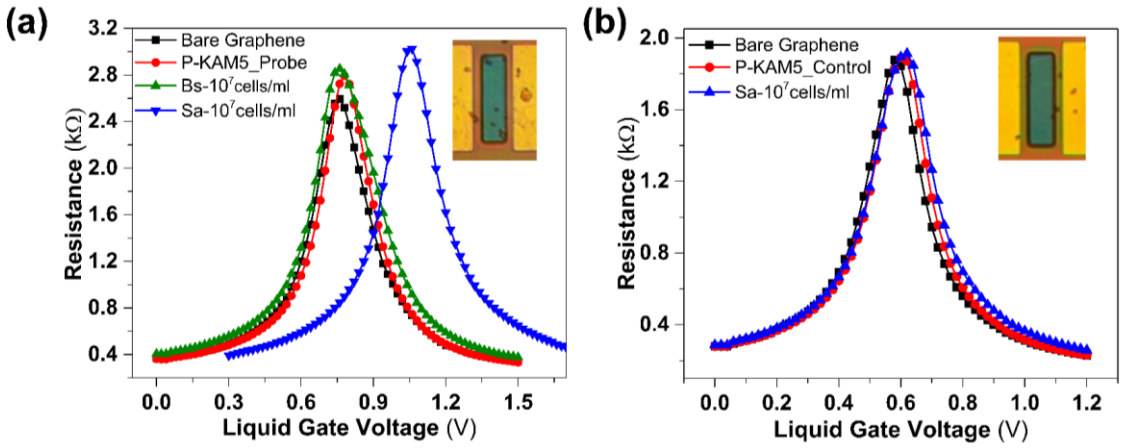


Figure 5-6. Resistance vs voltage plots of G-FET for detection of *S. aureus*. (a) G-FET functionalized with Probe peptides (P-KAM5_Probe) and incubated with *B. subtilis* and *S. aureus* at a concentration of 10^7cells/ml. No shift was observed with peptides and *B. subtilis* while a shift of ~300 mV is seen with *S. aureus* as well as the attachment of bacterial cells to the graphene (see image in inset). (b) G-FET functionalized with control peptides (P-KAM5_Control) and incubated with *S. aureus* at a concentration of 10^7 cells/ml. No voltage shift or attachment of bacterial cells was observed after the incubations as well as with bacteria.

KAM5_Control (inset of Figure 5-6a-b). In addition, we found a strong correlation between the number of bacteria that are bound by the probe to the graphene and the registered voltage shift after measuring 20 devices. As shown in Figure 5-7, a linear shift of V_D was seen with increasing number of attached bacteria with a sensitivity of 56.3 ± 7.3 mV/bacteria. Surprisingly, the devices are able to detect a single bacterium with an average

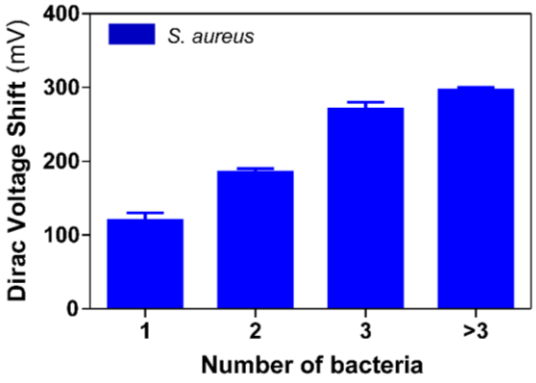


Figure 5-7. Measured Dirac voltage shift of G-FETs having different number of bacteria (*S. aureus*) attached. Devices having single bacterium attached show an average shift of ~130 mV and linear increase in voltage shift is observed with increased number of bacteria attached. (Data represents average and standard error of at least 3 independent replicates).

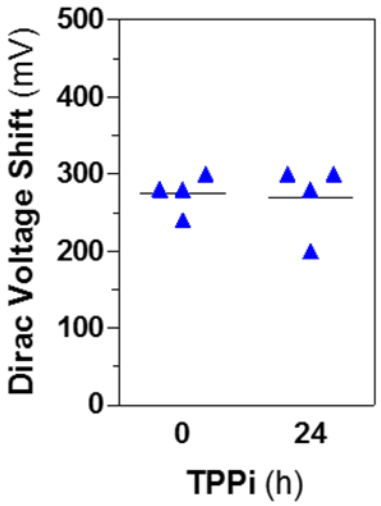


Figure 5-8. Stability test of G-FETs functionalized with P-KAM5_Probe. G-FETs were stored for 24h and then detection of *S. aureus* at 10^7 cells/ml was performed. Dirac voltage shift of different G-FETs after 0h and 24h time post probe incubation (TPPi) showed almost same values.

voltage shift of 128±18 mV, a nearly 20% increase in the measured V_D over the as-prepared G-FET. Importantly, the voltage shifts of ~130 → 300 mV (Figure 5-7) are much higher than those reported for *S. aureus* (~25 mV) and *E. coli* (~60 mV) using silicon based FET sensors[23, 24]. Furthermore, the peptides functionalized over G-FET remains stable tested after storing for 24 h in PBS which showed detection capability similar to those used immediately after functionalization (Figure 5-8).

To probe the postulated *S. aureus* specificity of our G-FET, other bacteria species were also tested, including *Bacillus subtilis* (*B. subtilis*) a different Gram-positive species and *E. coli* a representative Gram-negative bacterium. As expected, no significant shift in V_D was observed when the devices were incubated with either species under the same conditions

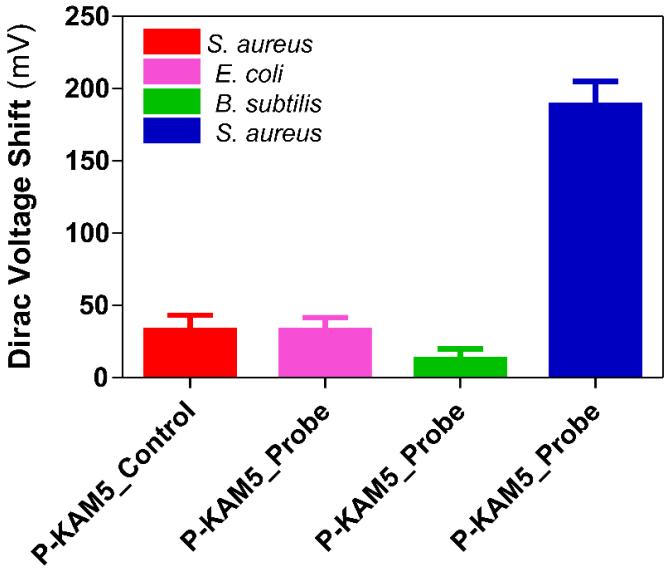


Figure 5-9. Specificity of P-KAM5_Probe functionalized G-FETs. No notable shift was observed when G-FET were functionalized with P-KAM5_Probe and incubated with unspecific bacteria (*E. coli* and *B. subtilis*) at a concentration of 10^7 cells/ml. An average shift of ~190 mV was observed when G-FET functionalized with P-KAM5_Probe and incubated with *S. aureus* at a concentration of 10^7 cells/ml. (Data represents average and standard deviation of at least 6 independent replicates)

used for *S. aureus* (Figure 5-9). Furthermore, after rinsing with water, the same devices subsequently incubated with *S. aureus* can still generate an average shift in V_D of ~190 mV, indicating that the devices functionalized with P-KAM5_Probe are specific to *S. aureus* and insensitive to other Gram-positive and negative species.

In summary, these results confirm that G-FETs functionalized with P-KAM5_Probe are capable of detecting *S. aureus* with high specificity and sensitivity, at the single cell level. The small size and neutral charge of peptide-based probe can reduce the Debye screening effect and background noise, respectively. Therefore, the sensitivity of the graphene to the charge of the bacteria is significantly enhanced.

5.4.2 Detection of antibiotic resistant strain of A. baumannii

Encouraged by the above results, we moved forward to test the specific detection of antibiotic-resistant pathogenic strains with P-KAM8_Probe modified G-FETs. Similar to the P-KAM5_Probe, functionalizing G-FET with P-KAM8_Probe caused no shift in V_D (Figure 5-10a). After incubation with 10^7 cells/ml of the colistin-resistant strain of *A. baumannii* (AB5075 LOS−; AbR), a V_D shift ranging between 280-460mV was observed. In contrast, no obvious shift was measured with P-KAM8_Control (Figure 5-10b). This confirms that P-KAM8_Probe effectively captures AbR cells onto the graphene surface triggering a change in V_D. Additionally, similar to what was observed for the interaction between *S. aureus* and P-KAM5_Probe, the measured voltage shifts correlate with the number of bacterial cells attached to the graphene surface (Figure 5-10c). We found a single *A. baumannii* produced a V_D shift of ~200mV, which is comparatively higher than that obtained with *S. aureus* which likely results from a higher density of surface charge displayed by a Gram-negative bacterium in comparison to Gram-positives[25]. In total, 14

different devices were tested using suspensions of 10^7 cells/ml and 10^6 cells/ml of AbR, obtaining average Dirac voltage shifts of about 350 and 280 mV, respectively (Figure 5-10d).

To test the strain specificity of P-KAM8_Probe, non-colistin resistant wild-type strain of *A. baumannii* (AB5075; AbW) was loaded onto the device at the same condition. No shift was observed in Dirac voltage indicating that P-KAM8_Probe does not capture the wild-

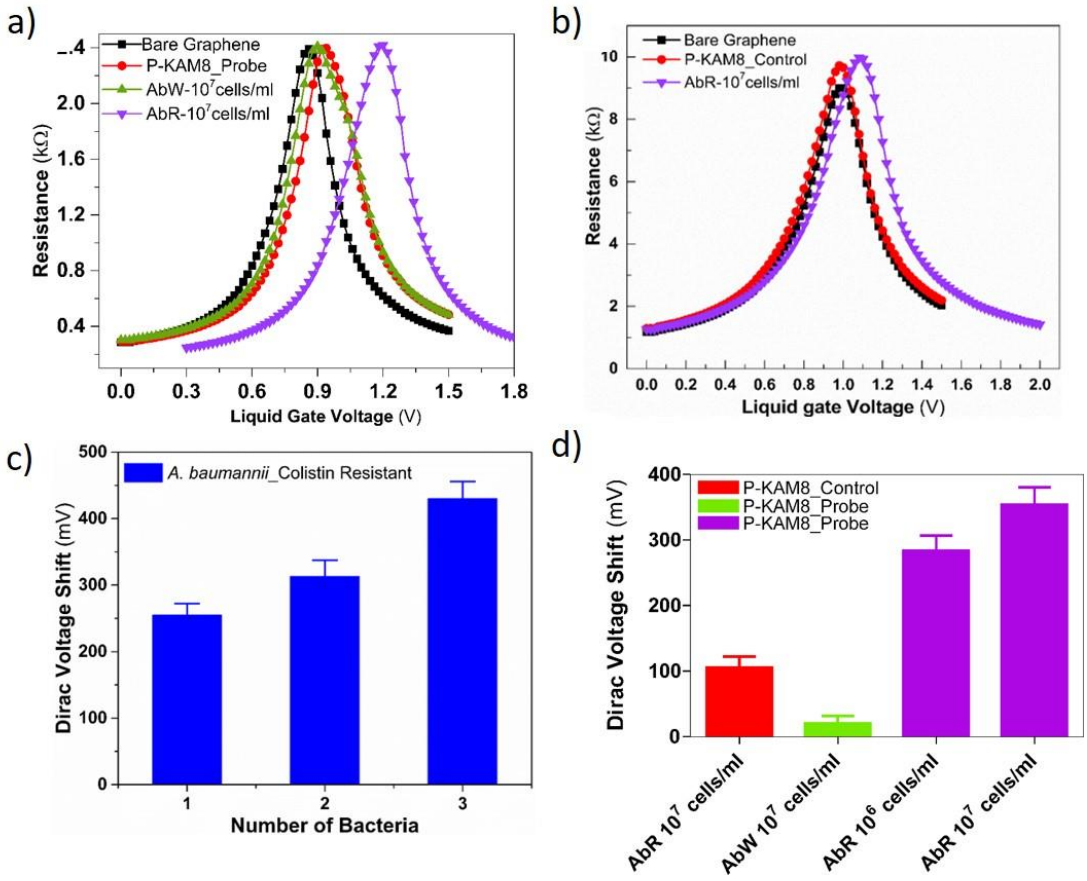


Figure 5-10. Specific detection results of *A. baumannii*. (a) Resistance vs voltage plots of G-FET for detection of *A. baumannii* with probe peptides KAM8_Probe. No shift was observed when the colistin sensitive wild type *A. baumannii* strain (AbW) was exposed to the device, while a ~300 mV shift occurs in the presence of the colistin resistant strain AbR. (b) G-FET functionalized with control peptides (P-KAM8_Control) didn't show notable voltage shift when incubated with *A. baumannii* of concentration 10^7cells/ml. (c) Device having single bacterium attached showed an average shift of ~250mV and linear increase in voltage shift is observed with increased number of bacteria attached. (d) Only shifts in Dirac voltage are registered when P-KAM-Probe is combined with AbR. This confirms that devices functionalized with P-KAM8_Probe are specific for AbR with average voltage shifts of 280 mV and 350 mV at concentrations 10^6 and 10^7 cells/ml respectively. (Data represents average and standard deviation of at least 4 independent replicates)

type *A. baumannii (Figure 5-10a)*. To our satisfaction, the subsequent incubation of the colistin-resistant strain of *A. baumannii* on the same device led to a shift of 280mV confirming the strain specificity of the probe.

The effect of pH on bacterial binding was also investigated. G-FETs functionalized with P-KAM8_Probe were tested with *A. baumannii* suspended in PBS altered to pH 6 and pH 8. Both cases showed the detection of bacteria (Figure 5-11), however, a decrease in voltage shift was noticed in comparison with that obtained with incubating in standard PBS pH 7.4. Overall, these results demonstrate the potential of the peptide-functionalized G-FETs for specific detection of antibiotic resistant strains of bacterial pathogens.

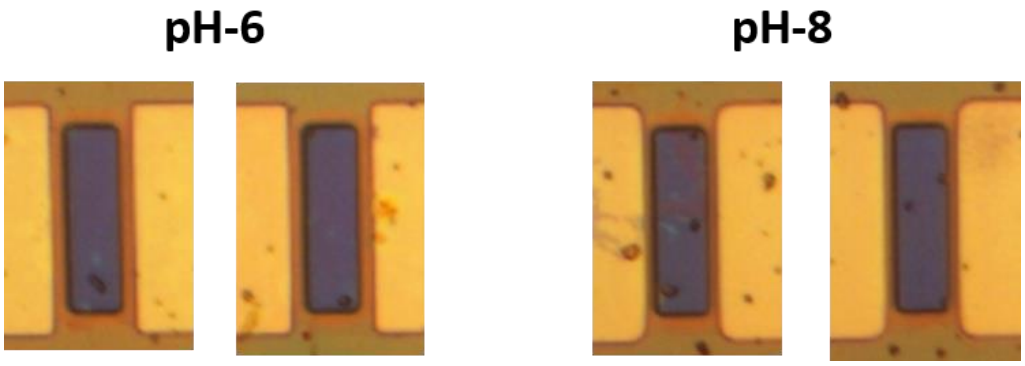


Figure 5-11. Effect of pH on binding of bacteria *A. baumannii*. G-FETs functionalized with P-KAM8_Probe were tested with *A. baumannii* suspended in PBS pH6 and pH8. Attachments were observed in both cases with variation in amount. A decrease in voltage shift was also noticed in comparison with that obtained with incubating in standard PBS (pH 7.4).

5.4.3 Enhancement on the detection limit

Although a single bacterial cell can generate a significant voltage shift on our device, a high bacterial cell density (10^7 cells/ml of *S. aureus*) is necessary to facilitate the capture of a single bacterium at the graphene surface. This is due to the relatively small size of the device where bacterial cells are distributed in an area of 2.5 x 2.5 mm and only 20 μL of

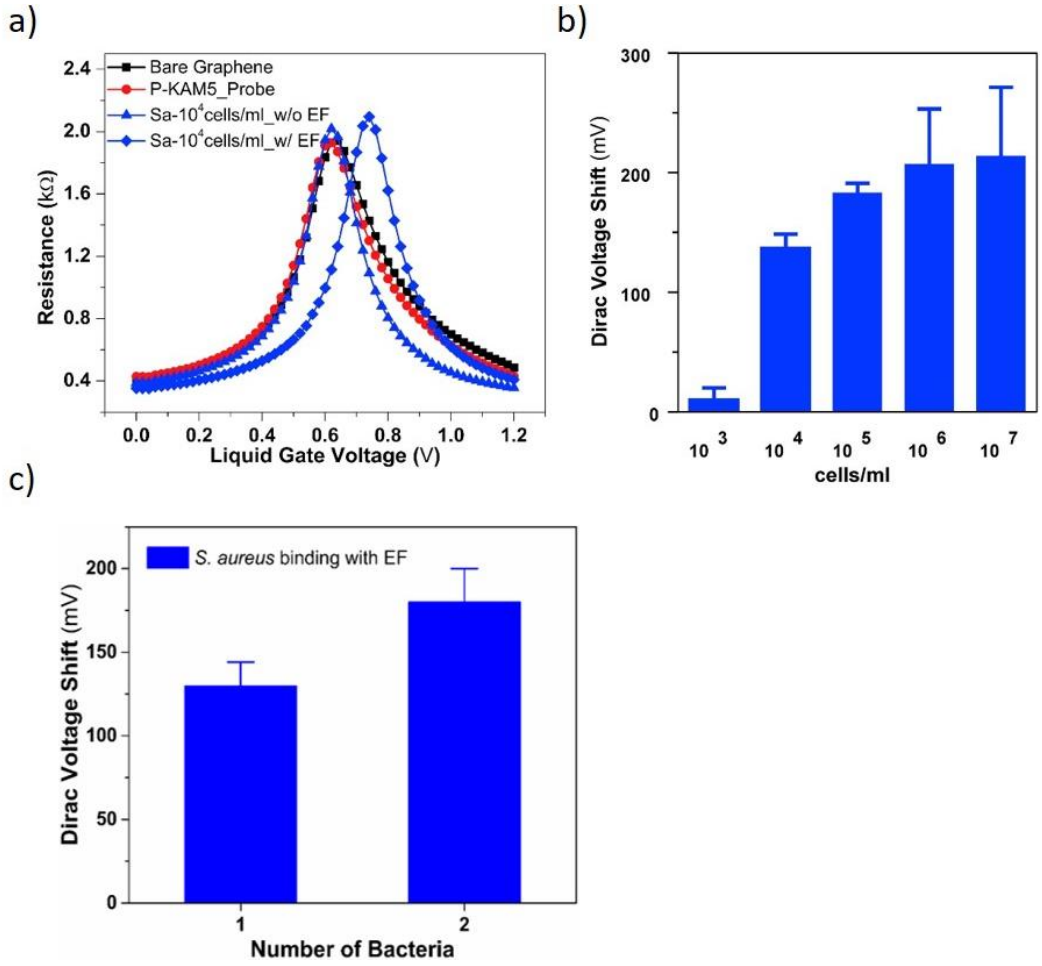


Figure 5-12. Improve the sensitivity of G-EFTs with electric field assisted binding. Resistance versus voltage plots of G-FET functionalized with Probe peptides (P-KAM5_Probe) (a) before (blue triangle) and after (blue diamond) electric field (EF) assisted binding of *S. aureus* at 10^4 cells/ml. (b) The chart shows the average Dirac voltage shift obtained with different concentrations of *S. aureus* after electric field assisted binding. (c) Measured Dirac voltage shift of G-FETs having different number of *S. aureus* cells attached obtained with electric field assisted binding at a concentration of 10^4 and 10^5 cells/ml.

cell sample is loaded each time. To improve the sensitivity of G-FET, we hypothesized that by applying voltage pulses from the top of the well that holds the sample, the charge of the bacteria could be exploited to drive them to the graphene surface[26]. Specifically, a negative voltage of -0.5 V was applied to the Pt electrode with five pulses, 10 seconds in duration to minimize potential damage to the bacteria. We applied the voltage to *S. aureus* sample with concentration of 10^4 cells/mL and then determine the shift in V_D. As shown in Figure 5-12a, a clear shift in V_D was observed, indicating attachment of bacteria to the graphene. Increased average V_D was observed with increasing concentration of cells, indicating enhanced probability of bacterial binding (see Figure 5-12b). A saturation in shift of V_D was observed when the bacteria count reaches to more than 4 per device. However, this only occurs when the concentrations are above 10^6 cells/ml, well beyond the clinically relevant limit[27]. Furthermore, the incubation time was decreased from 45 min to 5min with this electric-field assisted binding strategy. The Dirac voltage shift found in the electric-field assisted attachment is still dependent on the number of bacteria on the device (Figure 5-12c). No shift or bacterial attachment was observed for *B. subtilis* and *E. coli* using P-KAM5_probe devices, and *S. aureus* using P-KAM5_Control after applying the voltage, indicating the selectivity of the devices is not affected by applying the voltage (Figure 5-13).

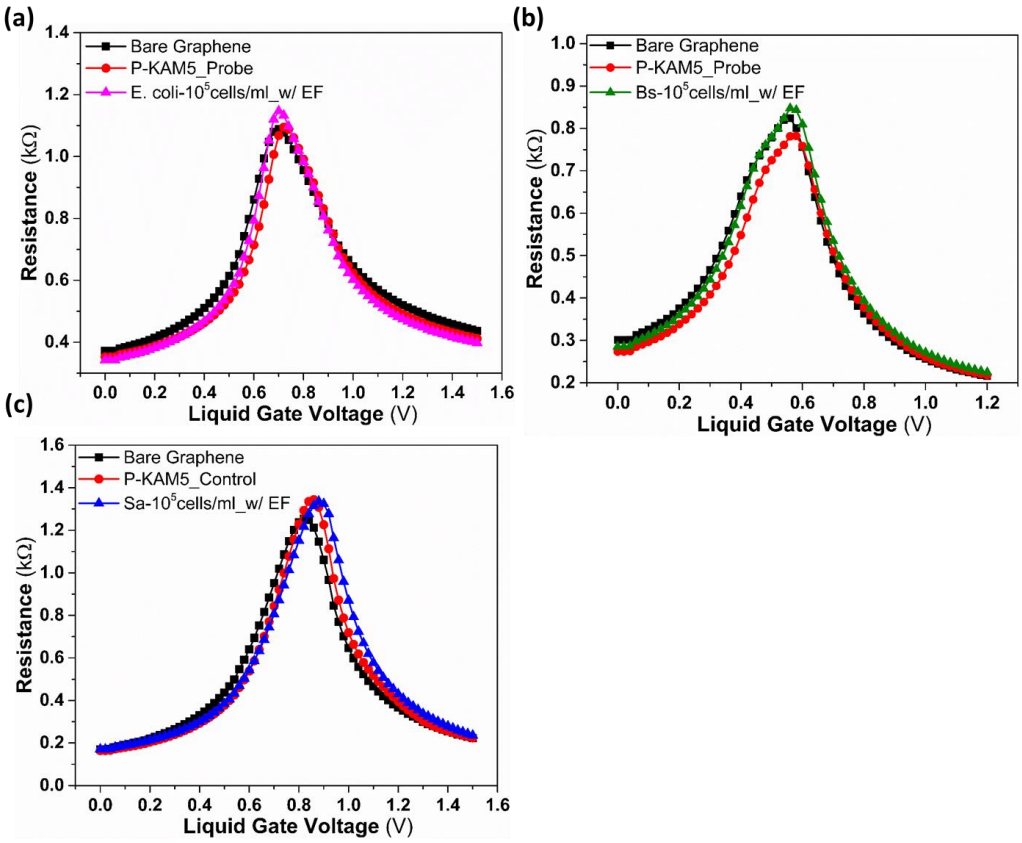


Figure 5-13. The specificity of G-FETs was not affected by electric field assisted binding. Resistance versus voltage plots of G-FET functionalized with Probe peptides (P-KAM5_Probe) and after electric field assisted binding of E. coli at 10^5 cells/ml (a), B. subtilis (b). Resistance versus voltage plots of G-FET functionalized with control peptides (P-KAM5_control) and after electric field assisted binding of S. aureus at 10^5 cells/ml (c).

The electric-field assisted binding method was applied to AbR sample with concentration of 10^4 cells/ml on the P-KAM8_Probe functionalized devices. However, -0.5V was not enough to push the AbR bacteria onto the graphene, which worked for *S. aureus*. We tried to apply -1V for 100s, but this seemed to damage the electrodes in the devices. Then we slowly increased the voltage from 0 to -1 V with a step voltage of 10mV, resulting in detection of AbR at cell densities as low as 10^4 cells/mL (Figure 5-14), which is lower than the reported threshold value of 10^5 cfu/mL as per the CDC reports[28]. Furthermore, by

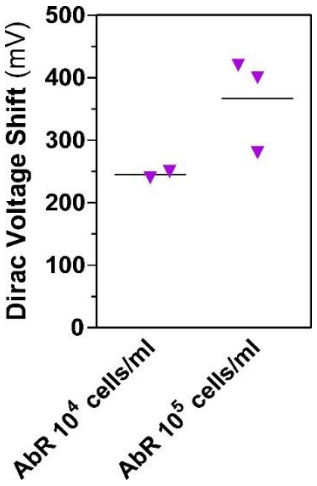


Figure 5-14. Bar chart shows average Dirac voltage shift versus concentrations obtained with electric field assisted binding of bacteria *A. baumannii*.

applying more cycles of incubation and voltage on the same device we achieved bacterial detection even at 10^3 cells/ml. However, these results showed more variability between replicates, it highlights that further sensitivity improvements are possible. The reason for the variability at 10^3 cells/mL could be that only ~20 bacterial cells are present in the 20 μL of sample that is loaded onto the device. Therefore, the bacterial density is low and the travel distance between a bacterium and the graphene surface might be long. To further improve the detection sensitivity, we will optimize the geometry of the PDMS well, resist surface and voltage application process, and/or integrate the system into a PDMS microfluidics chip.

5.5 Conclusions

In this chapter, we demonstrated an electronic, label free biosensor of clinically relevant bacteria by implementing two new capabilities in G-FETs; dielectrophoresis capture of the target and highly specific synthetic peptides. This enabled smaller size G-FETs, which are able to selectively detect a single bacterium on the device in just 5 min. Furthermore, this is achieved with a small (20 μL) sample volume at a detection limit of 10^4 cells/mL. Utilizing our new biosensor and procedures, we demonstrate the first selective, electrical detection of the pathogenic bacterial species *Staphylococcus aureus* and antibiotic resistant *Acinetobacter baumannii* on a single platform.

5.6 Experimental procedures

5.6.1 General Methods

Fmoc-protected amino acids were purchased from Advanced Chemtech or Chem-Impex Int'l Inc. HBTU were purchased from Novabiochem. Trifluoroacetic acid was purchased from Protein Technologies. Other chemicals were obtained from Fisher Scientific or Sigma Aldrich. Mass-spec data were generated by an Agilent 6230 LC TOF mass spectrometer. HPLC purification was done on Waters 2695 Alliance HPLC system.

5.6.2 G-FET Fabrication and Characterization

This part was done by Dr. Narendra Kumar. G-FETs were fabricated on CVD monolayer graphene transferred over SiO_2/Si substrates. Monolayer graphene was grown on copper via low pressure chemical vapor deposition. The copper foil (Alfa Aesar) was pre-treated in Ni etchant (Transene) to remove any coatings or oxide layers from its surface. The tube furnace was evacuated to a read pressure of 200 mTorr with a constant flow of H_2 (10 sccm). Prior to growth, the foil was annealed at 1010 ºC (ramp rate 25 ºC/min) for 35 minutes. Growth was done at 1010 ºC with 68 sccm of H_2 and 3.5 sccm of CH_4 for 15 minutes. After growth, a polymethyl methacrylate (PMMA) layer was spin coated on one side of the copper foil and baked for 60 seconds at 60 ºC. To facilitate smooth and fast etching of the copper foil, the backside graphene was etched out using oxygen plasma with 60-watt power for 60 seconds. The exposed copper was etched away in Ni etchant for 2h at 60 ºC. The remaining PMMA/graphene structure was washed in 2 water baths, the first water bath for 60 seconds and the second for 30 minutes, to rinse away left over etchant. The PMMA/graphene was transferred onto SiO_2/Si chips of size 1 x 1 cm. Any leftover water was slowly dried out with nitrogen gas, and finally the PMMA was dissolved in acetone

vapors; isopropanol alcohol (Fischer) was used for a final wash. The chips were baked at 300 ºC for 8h in vacuum followed by deposition of 3 nm AlOx at room temperature by evaporating aluminum at oxygen pressure of 7.5 x 10^5 mbar. Substrates were baked at 175 ºC for 10 minutes before lithography process. The electrodes patterning was done using bilayer photoresist (LOR1A/S1805) and laser mask writer (Heidelberg Instruments) followed by Au/Cr (45 nm/5 nm) deposition and lift off using remover PG (MicroChem). After that the graphene patterning was done with lithography using same bilayer resist and oxygen plasma etching. Devices were cleaned with remover PG and rinsed with IPA, DI water and dried with Argon. In order to protect the electrodes and edges of the graphene for liquid gating, photolithography was done using S1805 to open the sensing area (10 x 40 μm) and contact pads while leaving remaining chip covered. The developing time was increased to 90 seconds to etch away the AlOx layer deposited in the beginning to protect the graphene from photoresist. Finally, the chips were baked at 150 ºC for 5 minutes and then temperature increased to 200 ºC and baked for 5 more minutes to harden the photoresist. To hold the solution for the measurement, two PDMS wells of size 2.5 x 2.5 mm were fabricated and placed over the chip having two sets of the devices with three devices in each well.

5.6.3 Peptide synthesis

Solid phase peptide synthesis was performed on a rink amide resin using Fmoc chemistry. An alloc-protected diaminopropionic acid residue was installed at the C-terminus for on-resin coupling of pyrene. The alloc protecting group was selectively removed by tetrakis (triphenylphosphine) palladium (0) and phenylsilane in DCM for 1 h. 2 equivalents of 1-Pyrenebutyric acid N-hydroxysuccinimide ester in 20% v/v DIPEA/DMF was added. The

coupling was done in 2 h at room temperature. The peptides were cleaved off resin and globally deprotected with 90% TFA, 5% H_2O, 2.5% triisopropylsilane, 2.5% 1,2-ethanedithiol for 2 h. Crude peptides were obtained via cold ether precipitation and purified by RP-HPLC. For cysteine alkylation, the peptides were treated with 3 equivalents of APBA-IA or IA in 5% v/v DIPEA/DMF for 3 h and purified via RP-HPLC. All peptides were characterized with LC-MS to confirm their identities and excellent purities (>95%). The pyrene-conjugated peptide was dissolved in DMF as stock and then diluted in 1×PBS pH 7.4 buffer.

5.6.4 Bacterial strains and culture conditions

Detections were made using the following strains: *S. aureus* (ATCC 6538), wild-type *A. baumannii* (AB5075)[29], colistin resistant and LOS deficient *A. baumannii* (5075 LOS−)[30], *B. subtilis*, and *E. coli (BL 21)*. All bacteria were cultured overnight in LB broth at 37 °C with 220 rpm constant shaking. The overnight culture was diluted 10^2 times in fresh media and grown to an OD_{600} of 0.5-1.0. These fresh cultures were then washed and diluted with 1x PBS (pH 7.4) buffer to obtain the desired concentrations.

5.6.5 G-FET functionalization and measurement conditions

This part was done by Dr. Narendra Kumar. G-FETs were functionalized with peptides by incubating with 10 μM concentration of P-KAM5_Probe and P-KAM8_Probe for optimized durations of 2 h and 16 h, respectively. To minimize the noise in the electrical measurement, G-FETs were characterized by measuring the resistance using a digital multimeter by sweeping Liquid gate voltage between 0 V to a maximum of 1.7 V. The test current was limited to 10 μA reduce the effects of heating the device and prevent failure.

The maximum (Dirac voltage) in resistance versus voltage plot was chosen as reference point and shift in the maximum was measured upon bacterial binding. Functionalized G-FETs were incubated with 20 µl desired bacterial solutions in 1x PBS (pH-7.4), while the measurements were performed in 0.01x PBS diluted in DI water to maximize the signal by reducing the Debye screening effect[8, 24]. A platinum wire of 0.5 mm diameter was used for liquid gating.

5.7 References

1. Kumar, N., et al., *Dielectrophoresis assisted rapid, selective and single cell detection of antibiotic resistant bacteria with G-FETs.* Biosensors and Bioelectronics, 2020. **156**: p. 112123.
2. Donnelly, M., et al., *Graphene field-effect transistors: the road to bioelectronics.* Journal of Physics D: Applied Physics, 2018. **51**(49): p. 493001.
3. Fu, W., et al., *Sensing at the surface of graphene field-effect transistors.* Advanced Materials, 2017. **29**(6): p. 1603610.
4. Ping, J., et al., *Scalable production of high-sensitivity, label-free DNA biosensors based on back-gated graphene field effect transistors.* ACS nano, 2016. **10**(9): p. 8700-8704.
5. Xu, S., et al., *Real-time reliable determination of binding kinetics of DNA hybridization using a multi-channel graphene biosensor.* Nature communications, 2017. **8**: p. 14902.
6. Fu, W., et al., *Biosensing near the neutrality point of graphene.* Science advances, 2017. **3**(10): p. e1701247.
7. Huang, Y., et al., *Graphene-based biosensors for detection of bacteria and their metabolic activities.* Journal of Materials Chemistry, 2011. **21**(33): p. 12358-12362.
8. Wu, G., et al., *Graphene Field-Effect Transistors for the Sensitive and Selective Detection of Escherichia coli Using Pyrene-Tagged DNA Aptamer.* Advanced healthcare materials, 2017. **6**(19): p. 1700736.
9. Templier, V., et al., *Ligands for label-free detection of whole bacteria on biosensors: A review.* TrAC Trends in Analytical Chemistry, 2016. **79**: p. 71-79.

10. Rubab, M., et al., *Biosensors for rapid and sensitive detection of Staphylococcus aureus in food.* Biosensors and Bioelectronics, 2018. **105**: p. 49-57.

11. McGregor, D.P., *Discovering and improving novel peptide therapeutics.* Current opinion in pharmacology, 2008. **8**(5): p. 616-619.

12. Mannoor, M.S., et al., *Electrical detection of pathogenic bacteria via immobilized antimicrobial peptides.* Proceedings of the National Academy of Sciences, 2010. **107**(45): p. 19207-19212.

13. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* nature, 2002. **415**(6870): p. 389.

14. Mannoor, M.S., et al., *Graphene-based wireless bacteria detection on tooth enamel.* Nature communications, 2012. **3**: p. 763.

15. Mannoor, M.S., et al., *Graphene-based wireless bacteria detection on tooth enamel.* Nature Communications, 2012. **3**(1): p. 763.

16. Etayash, H., et al., *Impedimetric Detection of Pathogenic Gram-Positive Bacteria Using an Antimicrobial Peptide from Class IIa Bacteriocins.* Analytical Chemistry, 2014. **86**(3): p. 1693-1700.

17. McCarthy, K.A., et al., *Phage display of dynamic covalent binding motifs enables facile development of targeted antibiotics.* Journal of the American Chemical Society, 2018. **140**(19): p. 6137-6145.

18. Chen, T.-Y., et al., *Label-free detection of DNA hybridization using transistors based on CVD grown graphene.* Biosensors and Bioelectronics, 2013. **41**: p. 103-109.

19. Kireev, D., et al., *Graphene transistors for interfacing with cells: towards a deeper understanding of liquid gating and sensitivity.* Scientific reports, 2017. **7**(1): p. 6658.

20. Georgakilas, V., et al., *Functionalization of graphene: covalent and non-covalent approaches, derivatives and applications.* Chemical reviews, 2012. **112**(11): p. 6156-6214.

21. Kuang, Z., et al., *Biomimetic Chemosensor: Designing Peptide Recognition Elements for Surface Functionalization of Carbon Nanotube Field Effect Transistors.* ACS Nano, 2010. **4**(1): p. 452-458.

22. Liu, C., et al., *A graphene oxide nanosensor enables the co-delivery of aptamer and peptide probes for fluorescence imaging of a cascade reaction in apoptotic signaling.* Analyst, 2018. **143**(1): p. 208-214.

23. Nikkhoo, N., et al., *Rapid Bacterial Detection via an All-Electronic CMOS Biosensor.* PloS one, 2016. **11**(9): p. e0162438.

24. Formisano, N., et al., *Inexpensive and fast pathogenic bacteria screening using field-effect transistors.* Biosensors and Bioelectronics, 2016. **85**: p. 103-109.

25. Kłodzińska, E., et al., *Effect of zeta potential value on bacterial behavior during electrophoretic separation.* Electrophoresis, 2010. **31**(9): p. 1590-1596.

26. Belgrader, P., et al., *PCR detection of bacteria in seven minutes.* Science, 1999. **284**(5413): p. 449-450.

27. Habimana, J.d.D., J. Ji, and X. Sun, *Minireview: trends in optical-based biosensors for point-of-care bacterial pathogen detection for food safety and clinical diagnostics.* Analytical Letters, 2018. **51**(18): p. 2933-2966.

28. Bulens, S.N., et al., *Carbapenem-nonsusceptible Acinetobacter baumannii, 8 US metropolitan areas, 2012–2015.* Emerging infectious diseases, 2018. **24**(4): p. 727.

29. Jacobs, A.C., et al., *AB5075, a highly virulent isolate of Acinetobacter baumannii, as a model strain for the evaluation of pathogenesis and antimicrobial treatments.* MBio, 2014. **5**(3): p. e01076-14.

30. Boll, J.M., et al., *A penicillin-binding protein inhibits selection of colistin-resistant, lipooligosaccharide-deficient Acinetobacter baumannii.* Proceedings of the National Academy of Sciences, 2016. **113**(41): p. E6228-E6237.

Chapter 6 Conclusions

My research herein demonstrated the development of peptides with various functions. We applied different techniques, including rational design based on the structure of peptide scaffold, and screening technique phage display. Meanwhile, we also developed new chemistry methods for site-specific protein and peptide modification, which allowed us to expand the chemical space of phage display.

Prolinomycin provided us an excellent platform to study the folding properties of cyclic peptides. The proline residues could constrain the structure but high concentration of potassium ion is needed to induce the folding of the peptide in the aqueous buffer. The rigid cross-linkers were introduced to further stabilize the structure, which is similar to the strategy researchers has been used to develop stapled peptides. We introduced several non-canonical amino acids into the peptide including APBA moiety aiming to develop synthetic receptors for small molecules containing the amine group. However, we found that on the water-exposed surface of prolinomycin, the charge-charge interaction and hydrophobic interaction were not able to enhance the binding between the scaffold and ligands. Additionally, it is hard to predict the correct conformations of peptide-ligand complexes, which may require the advanced computational tool. The π-π stacking between pentafluorobenzene moieties and phenylalanine residue of a short peptide (Phe-Gly-Lys) did enhance the binding by 2-3 folds, indicating the cooperative action of side chains on the same deck of the scaffold could be beneficial.

In addition to structure-based rational design, we explored the screening technology to discover functional peptides. The advantage of this method is that we are not limited by one peptide scaffold, rather we can construct a large library of peptides, containing $10^{9\text{-}10}$ different peptide sequences, by phages and *E. coli* cells, which provides us a higher chance

of success. In addition, it requires less peptide synthesis, which could save the time and labor cost. The power of phage display was demonstrated herein by screening against sortase A (SrtA). The consensus motif of peptide sequence showed up after three rounds indicating the success of the screening. A cyclic peptide W7 was identified from NEB Ph.D.^{TM}-C7C phage display peptide library, showing single-digit micromolar affinity to SrtA, which makes it one of the most potent inhibitors of SrtA. We are surprised by how important the ring size and conformation of the peptide are to its binding affinity.

Furthermore, we chemically modified this peptide library encoded by phage and introduced the covalent warheads, APBA moieties. Interestingly, totally different screening results came out. Instead of a consensus motif, some repeating sequences showed up. We found a peptide W2 showing medium micromolar affinity to the surface of SrtA. This was consistent to the mechanism of iminoboronate formation between APBA warheads and the lysine side chains on the surface of SrtA. The linear conformation might make the peptide too flexible to bind to the active site of the target. Therefore, we further explored the library of cyclic peptides presenting APBA moieties. However, to our disappointment, no hit was identified from these cyclic peptide libraries. The potential reason could be that cyclic peptides modified with the cross linkers were too large or flexible and the conformations did match the binding pocket of SrtA. In the future study, we will optimize the size and rigidity of the cross linkers. Moreover, the length of the peptide will also be optimized. Overall, the sequence, length and conformation of the peptide play essential roles for its binding to a target. We also believed that targeting different proteins, different peptide libraries are required, depending on the size and

geography of the binding pocket, as well as the reactive amino acids around the binding site that could be covalently targeted by warheads.

In order to further expand the chemical space of phage display and enhance the molecular diversity of peptide library, we developed a novel strategy to site-specifically modify N-terminal cysteine and internal cysteine by two distinct functionalities. Consequently, a peptide library displayed by phage presenting dual modifications was constructed. This was achieved by taking the advantage of two chemistry methods, including thiazolidino boronate (TzB) mediated acylation reaction of NCys and 2-cyanobenzothiazole (CBT)-NCys condensation.

Based on CBT chemistry, we discovered an alternative reaction pathway that proceeds under basic conditions and leads to rapid formation of N-terminal amidine, which allowed for the N, S double labeling of peptides and proteins in excellent yields. We believe that this unique reaction pathway of CBT-NCys condensation presents a significant addition to the toolbox for site-specific protein modifications and can be applied to create novel chemically modified phage libraries.

Finally, we presented an application of peptide probes of bacteria detection. In the collaboration with the experts in physics and biology, peptide-modified G-FETs as electronic, label free biosensors of clinically relevant bacteria were developed. This device only required a small amount of sample (20 μL) and a short operation time (5 min). Remarkably, we demonstrate the first selective, electrical detection of the pathogenic bacterial species *Staphylococcus aureus* and antibiotic resistant *Acinetobacter baumannii* on a single platform.

Overall, this work presented the potential of peptides as synthetic receptor, protein inhibitor and probe for detection. It also deepened our understanding of the principles of molecular recognitions, including the interactions between peptides and small molecules, peptides and proteins, as well as peptides and bacteria. Last but not least, the importance of new chemical reactions for the improvement of phage display was highlighted here, which encouraged us to further develop novel biocompatible site-specific reactions.