THE NITRILIMINE-ALKENE CYCLOADDITION MECHANISM AND PHAGE-DISPLAYED CYCLIC PEPTIDE LIBRARIES FOR DRUG DISCOVERY

A Dissertation

by

XIAOSHAN WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Wenshe Liu
Committee Members,	Tadhg Begley
	Paul Lindahl
	Frank Raushel
Head of Department,	Simon North

May 2018

Major Subject: Chemistry

ABSTRACT

This study is composed of two parts. In the first part, we discussed nitrilimine-alkene cycloaddition for protein labeling. The mechanism of this nitrilimine-alkene cycloaddition was proposed, and thereby the best experimental condition for this protein labeling approach was investigated. The transient formation of nitrilimine in aqueous conditions is greatly influenced by pH and chloride. In basic conditions (pH 10) with no chloride, a diarylnitrilimine precursor readily ionizes to form diarylnitrilimine that reacts almost instantly with an acrylamidecontaining protein and fluorescently labels it.

In the second part, a novel method for the synthesis of phage-displayed cyclic peptide libraries is presented. Cyclic peptide drugs are appealing in the drug discovery research area due to their unique advantages including high affinity, high specificity, low toxicity, and high cellular and serum stability. In order to identify cyclic peptides as therapeutic agents, during my graduate study I have developed a phage display-based methodology that integrates the genetic noncanonical amino acid (ncAA) mutagenesis technique for the synthesis of novel phage-displayed cyclic peptides through simultaneous 1,4-addition between a cysteine thiol group and acrylamide moiety in N^{ε}-acryloyl-lysine (AcrK), a ncAA. Both cysteine and AcrK are genetically coded. The success of using a cysteine and an AcrK to cyclic a peptide in a model protein and phages was validated by various approaches. In order to afford a library, a phage-displayed cyclic peptide library was constructed by inserting a consecutive but sequence-randomized 6-mer peptide

flanked by an amino side cysteine and a carboxyl side AcrK. Panning of the synthesized phage-displayed cyclic peptide library was performed against two target proteins that are tobacco etch virus (TEV) protease and histone deacetylase 8 (HDAC8). A lot of high-affinity phage clones were isolated and collected. DNA sequencing of these selected clones led to the identification of several peptides that potentially inhibit TEV protease and HDAC8. To confirm their potencies as inhibitors, abundant peptides and their fluorophore-conjugated derivatives were synthesized through solid-phase peptide synthesis (SPPS). Measurements of fluorescence polarization change and IC50 value of these peptides when they bound to TEV protease and HDAC8 were performed.

Overall, we have mechanistically characterized the nitrilimine-alkene cycloaddition reaction and developed a novel approach for the synthesis of phage-displayed cyclic peptide libraries. The selection of displayed peptides against TEV protease and HDAC8 has resulted multiple peptides that display high potencies against these two enzymes.

ACKNOWLEDGEMENTS

First of all, I am extremely thankful to my committee chair, Dr. Wenshe Liu. Without his supervision, guidance and support, I would not have made great progress in this project.

I would also like thank my committee members, Dr. Begley, Dr. Lindahl, and Dr. Raushel, for their conscientious guidance and support throughout the course of my graduate studies.

I would also like to thank all my lab colleagues for their help with my research and life.

Finally, I would extend my gratitude to my parents, my husband, and friends for their love, encouragement and support during my Ph.D. program.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This study was supervised by a dissertation committee consisting of Dr. Wenshe Liu (Chair) of the Department of Chemistry; Dr. Tadhg Begley of the Department of Chemistry (committee member)' Dr. Paul Lindahl (committee member) of the Department of Chemistry, and the Department of Biochemistry and Biophysics; Dr. Frank Raushel (committee member) of the Department of Chemistry, and the Department of Biochemistry and Biophysics.

In Chapter II, the synthesis of organic compounds and the construction of plasmids which are used for expression of AcrKS2GPF were performed by Dr. Yan-Jiun Lee of the Department of Chemistry.

In Chapter III, mutation of M13K07 helper phage and construction of phage-related synthetase plasmids were performed by Dr. Catrina Reed of the Department of Chemistry. Synthesis of selected peptides was with the help from Dr. Jean-Philippe Pellois of the Department of Biochemistry and Biophysics. I would also like to thank Dr. Jeffery Tharp of the Department of Chemistry for his help on phage selection.

Funding Sources

In Chapter II, this work was supported by the National Science Foundation (grant CHEM-1148684) and the Welch Foundation (grant A-1715)

NOMENCLATURE

AcrK	N^{ε}-acryloyl-L-lysine
AMC	7-Amino-4-methylcoumarin
CrtK	N^{ε}-crotonyl-L-lysine
FAM	5-Carboxyfluorescein
FP	Fluorescence polarization
HATs	Histone acetyltransferases
HDACs	Histone deacetylases
NCAA	Non-canonical amino acids
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
SPPS	Solid phase peptide synthesis
TEV	Tobacco etch virus
Ni-NTA	Nickel nitrilotriacetic acid
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
sfGFP	Superfolder green fluorescent protein
SAHA	Suberoylanilide hydroxamic acid
Tris	Tris(hydroxymethyl)aminomethane
TSA	Trichostatin A
WT	Wild type

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

Cyclic Peptide

Small molecules and biologics are two categories of currently widely used drugs. Small molecule drugs are defined as small size drugs with molecular weights less than 500 Da, while biologics have molecular weights larger than 5000 Da. Compared to large biologics such as insulin and antibodies, small molecules have better cell permeability and metabolic stability. Small molecules are more suitable for oral delivery while protein therapeutics require injection or intranasal delivery. However, because of less structure complexity, small molecules don't perform well in target selectivity as protein therapeutics and they may have more side-effects.[1] Thus, there is a growing interest in the discovery of a new kind of drug.

With a medium molecular weight between small molecules and large biologics (500 Da - 5000 Da), peptide-based molecules are recognized as great success in novel therapeutics with high selectivity and affinity. It combines favorable properties of small molecule drugs such as good bioavailability and stability, and advantages of biologics such as high specificity. Moreover, compared to small molecules drugs, peptide-based molecules have lower toxicity and low accumulation in tissues due to amino-acid based structure. Peptide-based molecules began to emerge in mid-20th century. After decades of research and development, more than 7000 peptides have been identified, more than 500

peptide-based drugs are in preclinical development, and around 140 peptide-based drugs have been applied in clinical trials.[2]

Among all peptides, cyclic peptide is a special category of polypeptide with ring structure. Compared to traditional linear counterparts, cyclic peptides have more advantages in stability because macrocyclic structures is helpful in forming intramolecular hydrogen bonds, shielding polar atoms from solvent, and thereby have limited conformational flexibility.[3] The rigidity of structure also decreases the entropy and allows high binding affinity. In addition, due to lack of cleavable terminal amino group and carboxyl group, amide-linked cyclization provides protection against proteolytic degradation from exo- and endo-peptideases.[4] Because of these dominant advantages, more and more cyclic peptide drugs emerge in therapeutic market. Currently, more than 40 cyclic peptide therapeutics are in clinical trials.[5]

There are several common kinds of macrocyclization method, head-to-tail, side-chain-to-tail and side-chain-to-side-chain reactions. Head-to-tail cyclization formed amide bond through lactonization. Side-chain-to-side-chain cyclization includes reactions between sulfhydryl groups of cysteines, reactions between an amino group of lysine and a carboxyl group of aspartic acid or glutamic acid. It may also occur between canonical amino acid and non-canonical amino acid or between non-canonical amino acids. Recently Dr. Liu and co-workers has demonstrated the success of 1,4-addition of N^{ε}-acryloyl-l-lysine, a non-canonical amino acid, with thiol nucleophiles in protein labeling.[6] This study expands the chemical diversity of peptide cyclization approach.

Phage Display

The phage display system, as first described by George Smith in 1985 [7], is one of the most powerful methodologies in drug discovery. The phage display technology is designed to express foreign peptides as fusion proteins on the surface of filamentous phages. In general, a phage library with randomized amino acids at one or more positions is capable of having large diversity, containing as many as 10^{13} different variants. Phage selection is also rapid, convenient and efficient in producing large variants and amounts of peptides or proteins. The selection and isolation of specific peptides with high affinity can be achieved through biopanning of the very diverse phage libraries against immobilized targets. Usually, the enrichment of selected phages can be generated by passage and expression through an *E. coli* cell host. To enable high-affinity binding with a specific disease-related target, iterative rounds of the selection are performed. Finally, the selected phages are subjected to DNA sequencing to determine the peptide sequences. Further study of protein-ligand interactions potentially leads to pharmaceutical applications.

In the phage display system, the best-studied filamentous bacteriophage is F pilus-specific phage, also known as f1, fd and M13 (**Figure 1**).[8] The most commonly used approach is to use a filamentous phage to express a peptide fused to a major coat protein (pVIII) or minor coat protein (pIII). The pVIII display system has a high valency (about 2700 copies per virion) and only 10% can be utilized as fusion peptides.[9] In comparison, the pIII display system has lower valency than the pVIII system (less than 5 copies per virion). A high-valency system typically results in a selection of low affinity ligands, while lower valency ensures a selection of high affinity ligands. Moreover, during the expression

step, pIII can be knocked out while pVIII cannot. Therefore, to guarantee the high-affinity selection, most peptides and proteins are displayed with pIII proteins. In genetic engineering of phage libraries, phagemid and phages are two mainly used vectors, containing replication origins, gene 3 and antibiotic resistance gene. Compared to phages, phagemids are more commonly used for several advantages, such as smaller genomes, more restriction enzyme recognition sites for gene engineering, higher efficiency and easier control in DNA level.[10] However, phagemids cannot produce progeny phage particles by themselves. For the conversion of phagemids to phage particles, helper phages such as M13K07 and VSM13 are required to provide packaging of phagemid DNAs.

Life circles of filamentous phages include bacterial infection and replication. Unlike other DNA phages which inject their DNA into the host cells, filamentous phages inject the entire phage particles. For a host cell like *Escherichia coli.*, F pilus enables the binding of pIII protein, and membrane protein TolA, TolQ and TolR are receptors. After entry into the host cell, the ssDNA of M13 enables the replication and amplification.

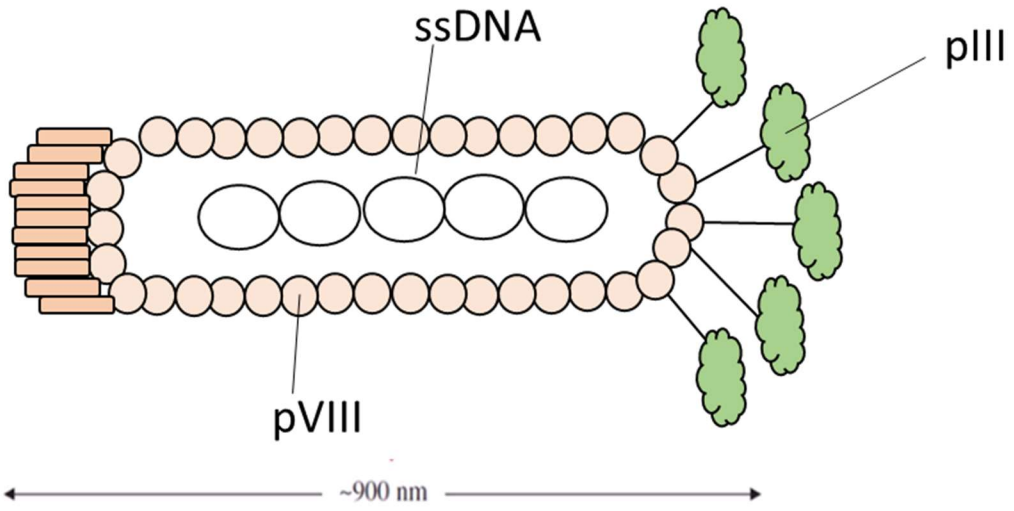


Figure 1 Structure of M13 phage.

Cyclization of Side Chains in Phage Display

In recent years, there is much current interest in expanding cyclization methods through side chains in phage display system. Although in natural, head-to-tail lactonization is a common way to generate macrocyclic peptides, it is not favorable by phage-displayed cyclic peptides due to the lack of free C-terminus on phages.[7] Two strategies have been reported to generate phage-displayed macrocyclic libraries, one through internal side-chain reactions and the other through chemical linkers. For the former one, the most widely used approach is through a disulfide bond between cysteines of a fusion peptide. Ghosh and co-workers demonstrated a method of constructing a 6-mer cyclic library Cys-$(Xxx)_6$-Cys (Xxx represents randomized amino acids) and screened this library against protein kinase A (**Figure 2A**).[11] Results show that cyclo(CTFRVFGC) is the most abundant sequence

with IC_{50} = 57 ± 3 μM. However, this disulfide bond is not stable in reducing environment, and thereby some organic linkers are designed to connect two cysteines.

To make cyclic phage-displayed libraries through ligation of chemical linkers, Szostak and co-workers recently described a method to generate a cyclic phage library containing 10 randomized amino acids through the reaction between two cysteines and an organic linker dibromoxylene(**Figure 2B**).[12] Winter and co-workers also described a phage strategy to generate a bicyclic peptide library through an organic linker tris-(bromomethyl)benzene which reacts with three cysteines (**Figure 2C**).[13] The peptide sequence of the phage library they designed is Cys-$(Xxx)_6$-Cys-$(Xxx)_6$-Cys. After the cyclization with tris-(bromomethyl)benzene, the selected bicyclic peptide has k_i = 1.5 nM towards human plasma kallikrein. To expand the chemical diversity of cyclization, more non-canonical amino acids are incorporated into the peptide. Schultz and co-workers have already incorporated several non-canonical amino acids into the phage display system[14], showing that the incorporation of some certain non-canonical amino acids could dramatically improve the ligand binding affinity towards specific targets.[15] Besides bromo-cysteine reactions, Derda and co-workers described an efficient way of cyclization using dichloro-oxime (DCO) derivatives as a cyclization linker.[16] The reaction is quite clean and highly chemo- and regio-selective.

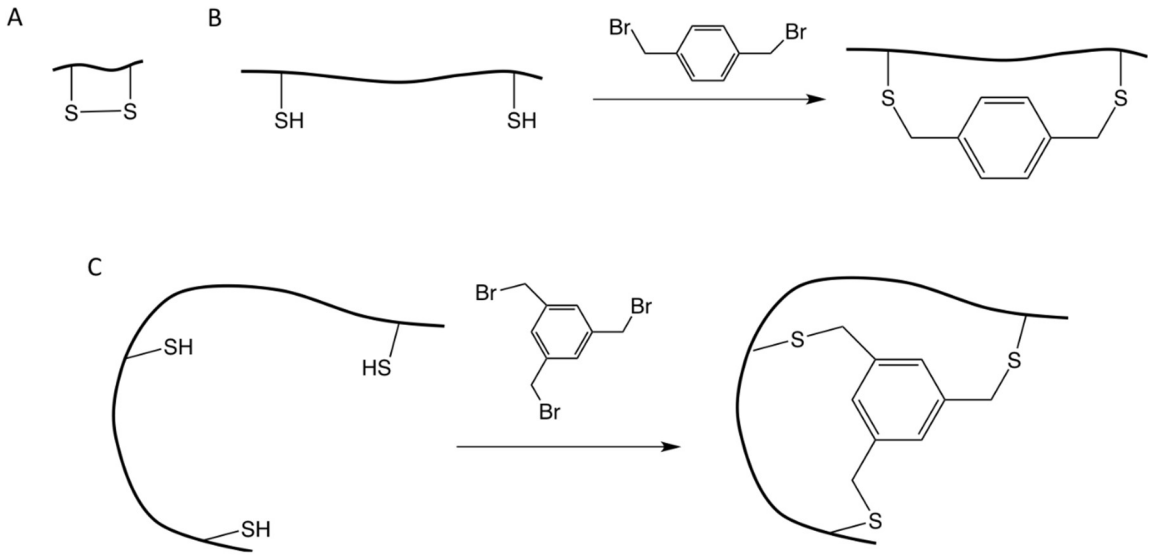


Figure 2 Cyclization methods to construct cyclic phage library. **(A)** Cyclization through a disulfide bond; **(B)** Cyclization through an organic linker dibromoxylene between two cysteines; **(C)** Cyclization through an organic linker tris-(bromomethyl)benzene.

Genetic Incorporation of Non-Canonical Amino Acids

There are twenty canonical amino acids found in natural. To expand the chemical diversity of amino acids, peptides and proteins, efforts have been made in site-specific incorporation of non-canonical amino acids (NCAAs) both *in vitro* and *in vivo*. Early efforts are focused on canonical machinery to suppress sense codon and incorporate NCAAs which are similar to canonical amino acids. This method has been limited to the strict similarity between incorporated NCAAs and canonical amino acids, as well as the competition from canonical amino acids. Nonsense codon suppression including stop codon (amber codon UAG, opal codon UGA and ochre codon UAA) suppression and quadruplet codon suppression means reading the nonsense codon as the sense codon to insert NCAAs via an orthogonal aminoacyl-tRNA synthetase (AARS)-tRNA pairs. Compared to sense codon suppression, this strategy has greater flexibility and has been applied in incorporation of more than 150 different NCAAs. Pioneering efforts made by Schultz and co-workers utilized a MjTyrRS-$tRNA_{CUA}^{Tyr}$ pair and allowed the site-specific insertion of NCAAs at amber codon and bulk expression of proteins in *vivo* via host cell's own translational machinery.[17]

The tRNA of the 22nd proteinogenic amino acid, pyrrolysine, is also chosen for its success in reassign UAG codon to encode a NCAA. In 2002, pyrrolysine was firstly discovered in Methanosarcina barkeri and a specific aminoacyl-tRNA synthetase, Pyrrolysyl-tRNA synthetase (PylRS), was reported to encode pyrrolysine with the amber codon[18,19] . Due to its high substrate side chain promiscuity and low selectivity toward α-amine and the tRNA anti-codon, PylRS and its orthogonal suppressor $tRNA^{Pyl}$ have been

harnessed for the genetic code expansion.[19] Hereby more NCAAs with hard-to-introduced functional group can be incorporated to the recombinant proteins by PylRS-$tRNA^{Pyl}$ pairs or their mutants. As achieved by Liu and co-workers, the mutants of PylRS and their cognitive tRNA pairs have been successfully used in the incorporation of diverse lysine and phenylalanine derivatives.[6] This translation process can be described as **Figure 3**. Under the catalysis of PylRS or its mutant, Pyl or other corresponding NCAA is transferred to the amber codon tRNA. The aminoacylated tRNA is delivered to the A site of the ribosome by the elongation factor, and further to the P site to couple its uncharged amino acid. Translation will not stop until the ribosome encounters a "real" stop codon.

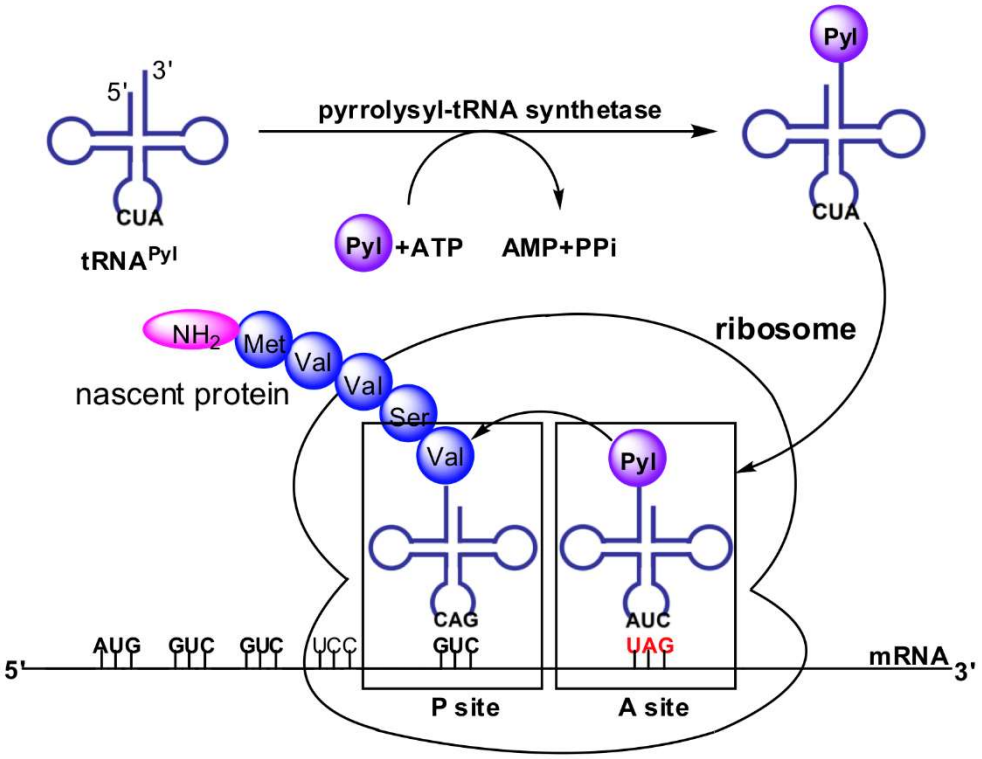


Figure 3 Machinery of NCAA incorporation using a pyrrolysine model

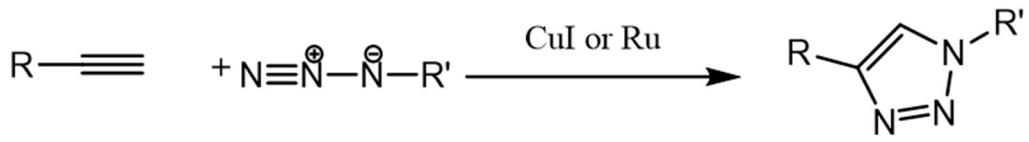
Biorthogonal Reactions in Protein Labeling

Selective labeling with chemical probes has emerged as an indispensable tool to investigate structure, function and internal or external interactions of proteins. Success of protein labeling requires biorthogonal reactions only with target functional group at a defined site. Reaction conditions should also be mild and have little or no toxicity to living cells. Additionally, reaction conditions such as aqueous solvent, neutral pH and proper temperature are also needed. Moreover, the incorporation of NCAAs bearing diverse biorthogonal functionalities allows site-specific labeling of recombinant proteins *in vitro* and *in vivo*.

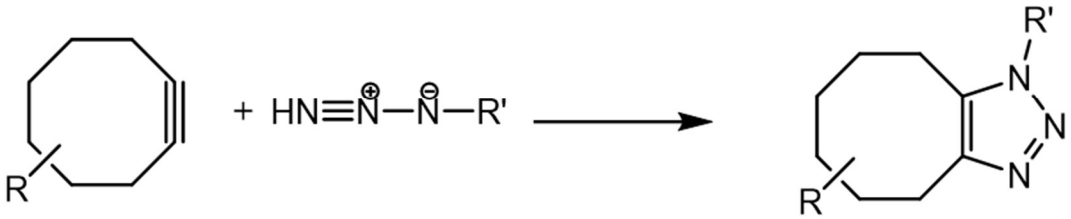
In recent decades, researchers have made tremendous progress on the discovery of biorthogonal reactions with chemo- and regio-selectivity. Amongst these reactions, click reactions have raised much focus due to their high selectivity, reliability and the ability to meet criterion of protein labeling as mentioned earlier. Click reactions include 1,3-dipolar cycloaddtion, Staudinger reaction, copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC), Diels–Alder cycloaddition, strain-promoted azide alkyne cycloaddition (SPAAC) and so on.

Amongst these click reactions, aizde derivatives play very important roles due to their unique advantages including small size, inert nature and high selectivity when reacting with phosphines, alkynes and alkenes. General ligations with azides include Staudinger ligation, copper catalyzed azide-alkyne cycloaddition (CuAAC) and copper-free azide-alkyne cycloaddition (**Figure 4**).[20]

Metal-catalyzed azide-alkyne 1,3-cycloaddition (CuAAC)



Copper-free azide-alkyne 1,3-cycloaddtion (SPAAC)



Staudinger ligation

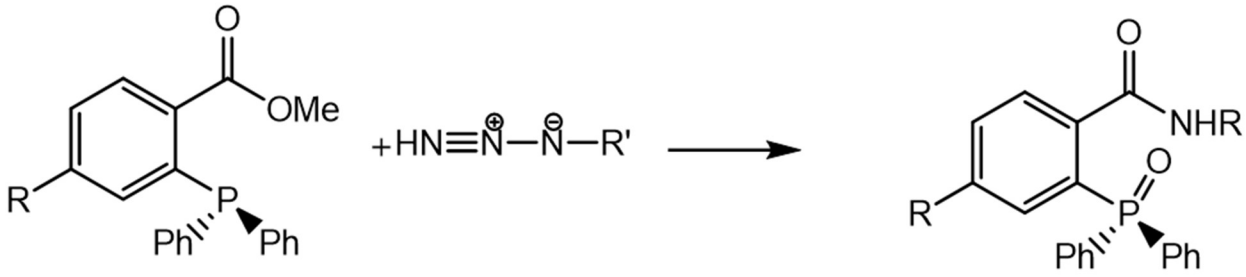


Figure 4 General ligation of azide derivatives

Staudinger ligation between azides and triarylphosphines with a ester group to yield amines, reported by Bertozzi and coworkers in 2004, is the first bioconjugation methodology applied in living system.[21] Later, the traceless Staudinger ligation was developed, where phosphine oxide is excised during the formation of amide bond. However, the application of Staudinger ligation as a biorthogonal reaction is limited for the low reaction rate (~ 10^{-3} M^{-1} s^{-1}) and phosphine reagents' susceptibility to oxidation. [22]

Huisgen's pioneering efforts in 1,3-dipolar cycloaddition reaction, where azides act as 1,3-dipoles and react with alkyes to yield triazoles under the catalysis of copper(I) salts, has had a profound impact in protein labeling. This CuAAC reaction is at least 25 times faster than the Staudinger ligation.[22] However, the toxicity of copper which may be harmful to living systems is still a big problem. Thereby reactions without catalysis from copper have raised more focus. These reactions are the so-called "Cu-free click reactions". One example is strain-promoted alkyne-azide cycloaddition (SPAAC) which can avoid copper catalysis by introducing ring strain into the alkyne.[23]

Nitrile imine is also an essence synthetic intermediate of 1,3-dipolar cycloadditions leading to pyrazolines, pyrazoles and other organic compounds.[24] Non-stabilized nitrile imines have six different structures: 1,3-dipolar, allenic, propargylic, reverse 1,3-dipolar, 1,3-diradical and carbonic, as shown in **Figure 5**.[25] Sibi and co-workers firstly reported the highly regio- and enantio-selective asymmetric 1,3-diplar cycloaddition of nitrile imines to olefins in 2005.[26]

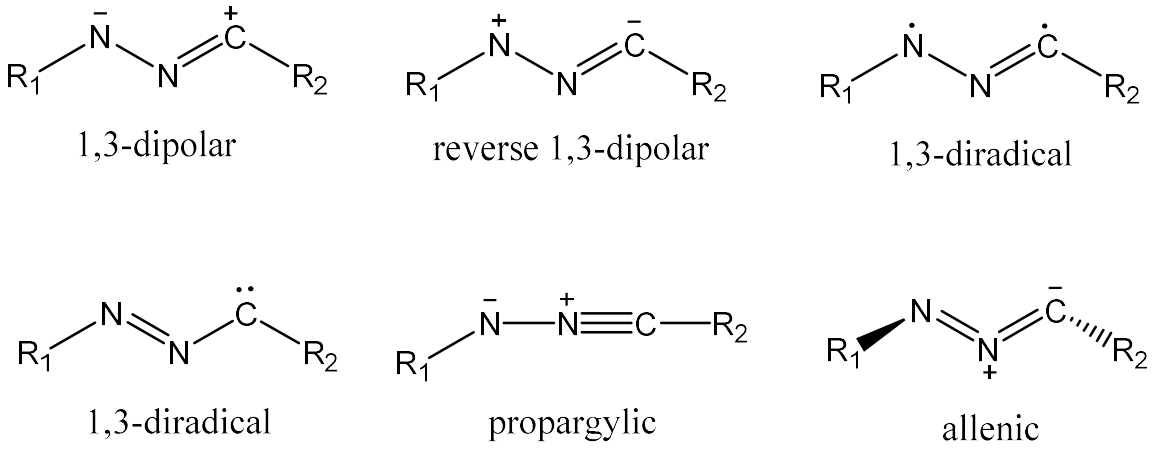


Figure 5 Fundamental structures of nitrile imine

In recent years researchers have reported a variety of methods to generate nitrile imine from different precursors, including 1) heating or photolysis of tetrazoles, 2) catalytic oxidation of aldehyde, 3) based-induced dehydrohalogenation of hydrazonoyl chlorides (**Figure 6**).[27] These generated nitrile imines are able to react with a number of dipolarophiles, leading to various cycloaddition products.

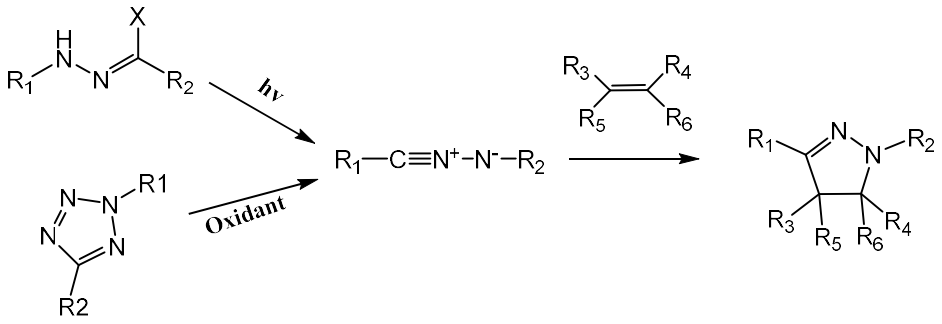


Figure 6 Generations of nitrile imine and related reactions.

Fluorescence Protein Labeling Strategy and Fluorogenic Probes

Fluorescence microscopy and fluorescent labeling are powerful tools for biological research. Fluorescent labeling includes direct labeling with fluorophore probes and labeling with fluorescence turn-on probes. For the former one, an ideal fluorophore should have small size, intense brightness, good spectral properties and long-term photostability.[28] For the latter one, fluorogenic reactions lead to fluorescent products from two non-fluorescent reagents. The turn-on fluorescence comes from the new fluorophore scaffold synthesized through cycloaddition of starting materials, or from activation of fluorescence resonant energy transfer (FRET) from an energy-transfer agent.[29] Reactions with the dienophile and removal of a functional group which quenches fluorescence of a fluorophore may also increase inherent fluorescence signal from original probes. These activatable fluorescent turn-on probes are highly desirable for the rapid imaging of biomolecules, low background signal and simple labeling step without removal of unreacted reagents.[29]

However, because of the difficulty of design, only a few probes have been reported in recent years. Lin and co-workers firstly discovered a photo-induced tetrazole-alkene cycloaddition and termed it as "photoclick reaction" in 2008 (**Figure 7**).[30] Photoclick reaction is biocompatible, easier to trigger, more convenient to monitor, as well as highly photoactivable. Additionally, it provides spatial and temporal control over the labeling process. [31] Fluorescence from the photo-induced cycloaddition product was observed after irradiation of UV lamp at 302 nm. The fast reaction requires reactive intermediate species triggered by light, and can be finished in 1 min. Moreover, the facile products from facile

alkenes and alkynes also show the similar efficiency. In addition, when reacting with tetrazole, cyclic alkenes tend to have higher yields than acyclic counterpart.[32] Reaction has been reported in the labeling of protein or peptide *in vitro* and *in vivo*.[33] Recently, Guo and co-workers demonstrated a similar fluorogenic probe based on styrene–tetrazine cycloaddition for in live cell labeling.[34]

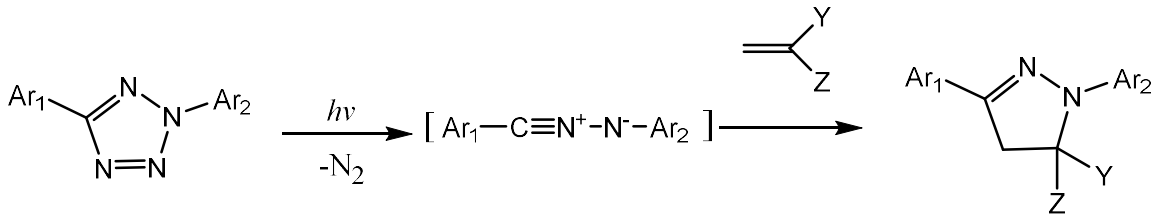


Figure 7 "Photoclick" 1,3-dipolar cycloaddition reaction

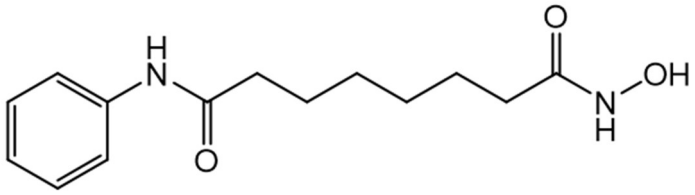
Histone Deacetylases

Introduction

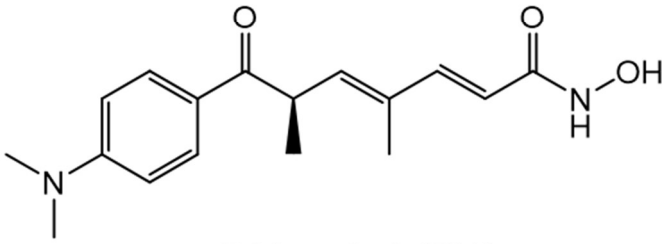
Acetylation of histones is a very important posttranslational modification (PTM) process in various organisms. The highly reversible acetylation process occurs at lysine residues on histones, especially amino-terminal tail of histones. Under catalysis of histone acetyltransferases (HATs), lysine residues on histones become acetylated. Histone acetylation affects chromatin structure and gene expression, allowing the acetylated chromatin more accessible to interacting proteins.[35] As opposed to acetylation, histone deacetylation catalyzed by histone deacetylases (HDACs) removes of the acetyl group from an ε-N-acetyl lysine on a histone. To date, 18 HDACs have been identified. Based on their phylogenetic similarity, human HDACs can be classified into four subclasses: Class I (HDAC1, HDAC2, HDAC3, and HDAC8), Class II (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10), Class III (SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, and SIRT7), and Class IV (HDAC11). Unlike Class I, II and IV which contains metal ion catalysis cofactor, Class III HDACs require NAD^+ as a cofactor for deacetylation. Histone deacetylation has impact on chromatin structure by making the DNA wrapped by histone more tightly, as well as transcription, protein-protein interaction and dynamics of histone crosstalk. Thus, abnormal HDACs or aberrant deacetylation process contributes to many human diseases, including cancer and diverse disorders.

Discovery of HDAC inhibitors is important to the cellular regulation and anti-cancer therapeutic development. To date, FDA has approved four HDAC inhibitor drugs for anti-cancer purpose. At the same time, five HDAC inhibitor drugs are in clinical trial process. The first HDAC inhibitor drug Vorinostat, also known as suberoylanilide

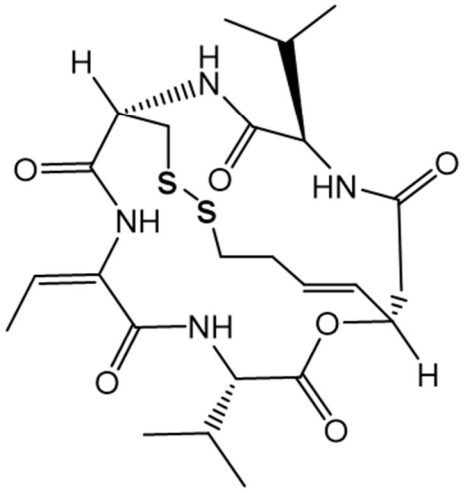
hydroxamic acid (SAHA, **Figure 8**), is approved by FDA for the treatment of cutaneous T-cell lymphoma (CTCL) in 2006.[36] The hydroxamic acid group can chelate the active-site zinc in HDAC and inhibit binding from other substrates. Later, Romidepsin (FK228, **Figure 8**) was approved in 2009 for its striking treatment in peripheral T-cell lymphoma (PTCL). It is notable that Romidepsin is a cyclicpeptide drug isolated from *Chrossmobacterium violaceum*. The intramolecular disulfide bond of Rommidepsi can be reduced in cells, yielding a thiol side chain that can chelate the active-site zinc in HDAC.[37] Another example is Trichostatin A (TSA, **Figure 8**) , a widely used HDAC inhibitor with 30-fold stronger inhibitory activity than SAHA.[38, 39] However, TSA has not been applied to clinical trial yet. For most of these FDA-approved drugs, a big problem is nonselectivity towards most HDACs, which may inhibit non-target HDAC or other non-target proteins, leading to side effects including cardiac arrhythmia and thrombocytopenia.[40,41]



Suberoylanilide Hydroxamic Acid (SAHA)



Trichostatin A (TSA)



Romidepsin (FK228)

Figure 8 Commonly used HDAC8 inhibitors

HDAC8 structure

Among all HDAC homologues, HDAC8 is the best studied in both mechanism and structure. HDAC8 is a zinc-dependent class I deacetylase, containing 377 amino acids. It has been demonstrated that the overexpression of HDAC8 affects cAMP responsive element binding protein (CREB) phosphorylation in cells. The overexpression is also reported in multiple cancer tissues, including colon, breast, lung and pancreas. Substrates of HDAC8 include not only acetylated histones (H2A/H2B, H3 and H4 histones), but also some non-histone substrates. Examples of histone substrates include acetylated lysines at position 14 (Kac14), 16 (Kac16) and 20 (Kac20) on histone H4 or its derivatives.[42] Besides histone substrates, some acetylated non-histone peptides can also be catalyzed by HDAC8, such as arginine-Kac129 in Estrogen-Related Receptor α (ERRα), RHKK in p53, RSKacFE in inv(16) fusion protein, and transcription factor CREB. It has been demonstrated that peptides with a N-terminal arginine at the -1 position to the acetyllysine or a C-terminal phenylalanine at the +1 position are potential substrates of HDAC8.[42] More non-histone substrates are being studied for specificity of deacetylation sites.

HDAC8 is the second smallest enzyme in HDAC family. Unlike other HDACs catalytically active as large complexes, HDAC8 functions as a small complex and is therefore chosen as a good target for the study of deacetylase. [43] In general, the structure of HDAC8 is a head-to-head dimer, with a zinc ion and two potassium ions in each monomer part.[44] The zinc ion is pentacoordinated by Asp178, His180, Asp267 and one or two water molecules.[44] The monomer HDAC8 consists of eight parallel β-sheets surrounded by eleven α-helixes. In the active site, there is a long hydrophobic tunnel leading to a narrow pocket. The tunnel and the surrounding nine loops have been

demonstrated to accommodate the binding of acetylated substrates. Crystal structure of HDAC8 and Trichostatin A (TSA), a well-studied HDAC inhibitor, illustrates how TSA interact with HDAC via the binding of the tunnel and three nearby loops (**Figure 9**). [36]

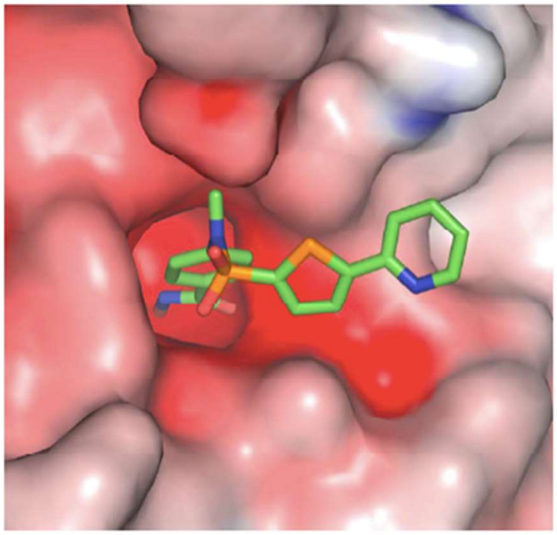


Figure 9 Surface of HDAC8 with two TSA molecules near active site. PDB ID:1t64. The first TSA molecule binds to the active site inside the tunnel. The second TSA molecule binds to the cavity nearby.[45]

Catalytic Mechanism of HDAC8

Figure 10 illustrates a proposed mechanism of HDAC8. As mentioned previously, the Zn^{2+} ion has a 5-coordinate square pyramidal geometry, including two asparagines, a histidine and two water molecules. During the catalysis process, one of these two water molecules is replaced by the acetyl lysine substrate. His143 functions as a base and accepts a proton from the zinc-chelated water molecule. The rate-determining step is where the deprotonated water molecule nucleophilically attacks the chelated carbonyl oxygen from acetylated lysine substrates. A hydrogen bond is generated between Tyr306 and the carbonyl group, forming a tetrahedron oxyanion intermediate. Later, the deacetylated lysine product leaves after getting a proton from the protonated base. Finally, the acetate product leaves after replaced by a water molecule.[36] To prove the key role of His143 and Tyr306, Christianson and co-workers mutated His143 to arginine. Additionally, Marco and coworkers also mutated Tyr306 to phenylalanine, respectively.[46,47] Results show that both of the mutated enzymes are inactive to their substrates.

Due to the two monovalent cation (MVC) sites in HDAC8, the HDAC8 catalytic activity can be modulated by the concentration of cations such as K^+ and Na^+. Fierke and co-workers investigated the two MVCs and demonstrated that the first MVC near the active site enhances HDAC8 catalysis, while the second MVC, which is 20 Å from the active site, decreases catalytic activity.[48] In general, addition of low concentration of K^+ increases the catalysis activity ($K_{1/2,\ act}$ = 14 mM), while the high concentration decreases the catalysis activity ($K_{1/2,\ int}$ = 130 mM). [36] Additionally, other divalent metal ions can also enhance the HDAC8 activity, following the order $Co^{2+} > Fe^{2+} > Zn^{2+} > Ni^{2+}$.[49]

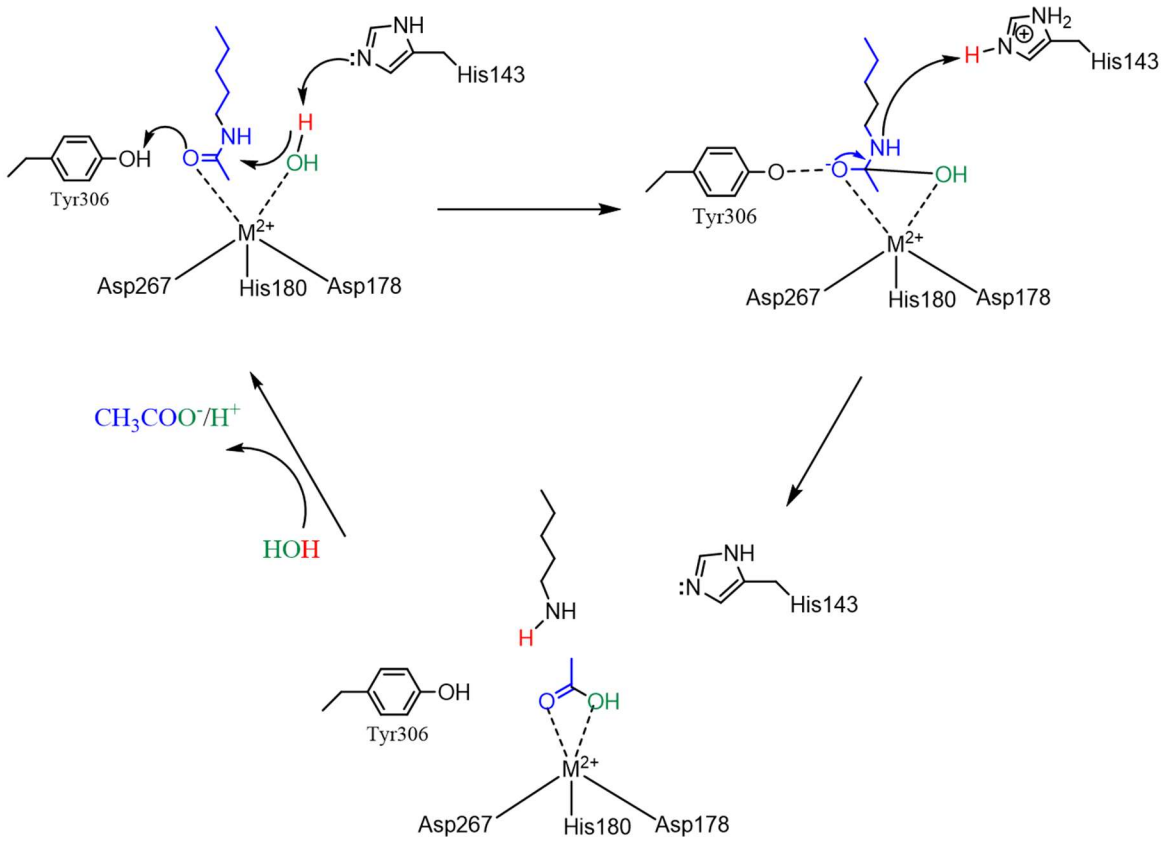


Figure 10 Proposed mechanism for HDAC catalysis.

CHAPTER II

THE NITRILIMINE-ALKENE CYCLOADDITION IS AN ULTRA RAPID CLICK REACTION*

Introduction

A recent mining of organic reactions for click labelling of proteins has revamped tetrazine-based inverse electron-demand Diels-Alder cycloaddition[50–54] and cyclooctyne-based 1,3-dipolar cycloaddition.[55–58] Unlike the Cu^{+}-catalysed click reaction,[59–61] both tetrazine- and cyclooctyne-involved cycloadditions undergo spontaneously in aqueous conditions, avoiding side reactions potentially induced by a transition metal catalyst.[62] A noteworthy advantage of tetrazine-based click reaction is its fast reaction kinetics. A hitherto fastest reported tetrazine-transcyclooctene reaction has a second-order rate constant as 2.8×10^{6} $M^{-1}s^{-1}$.[63] Cyclooctyne was originally explored for labeling proteins with azide and recently extended to react with nitrone and tetrazine functionalities for click labeling of proteins.[56,64,65] Cyclooctyne reacts rapidly with tetrazine.[66] Derivatives of cyclooctyne that react with azide and nitrone with a second-order rate constant higher than 50 $M^{-1}s^{-1}$ have also been developed.[67–69] Another copper-free click reaction that has been recently explored for protein labeling but not yet highly appreciated is the nitrilimine-alkene cycloaddition.[70] On contrary to tetrazine and cyclooctyne that stably exist in aqueous conditions, nitrilimine reacts

---

*Reprinted (adapted) with permission from "The nitrilimine–alkene cycloaddition is an ultra rapid click reaction" by Wang, X. S., Lee, Y. -J., and Liu, W. R. *Chem. Commun* **2014**, 50, 3176-3179. Copyright 2014 The Royal Society of Chemistry.

with water, therefore needs to be formed transiently.[6] There are two general methods to transiently form nitrilimine in aqueous conditions. One is the photolysis of tetrazole and the other is the ionization of hydrozonoyl halide.[70,71] Lin and coworkers have extended the first approach for photoclick labeling of proteins that contain the alkene functionality in living cells.[72–74] The second approach has been recently explored to undergo fluorescent turn-on labeling of proteins incorporated with norbornene and acrylamide moieties.[6,75,76] Reaction kinetics of the nitrilimine-alkene cycloaddition that involved both tetrazole and hydrozonoyl chloride as nitrilimine precursors were previously characterized.[6,73] All these characterizations were performed in a PBS-acetonitrile (1:1) buffer. The high concentration of chloride (140 mM) in PBS potentially offsets the transient formation of nitrilimine and consequently curbs its reaction with alkene. Here we report a comprehensive study of effects of pH and chloride concentrations on the nitrilimine-alkene cycloaddition reaction kinetics and demonstrate the nitrilimine-alkene cycloaddition is an ultra rapid click reaction for protein labeling at a slightly basic condition (pH 10) with no chloride.

Experimental Methods

1. Kinetic Analysis

Acrylamide was added to a solution of hydrazonoyl chloride (**1** in **Figure 11**, 5 μM, 2mL) in 1:1 acetonitrile-50 mM phosphate buffer at different pH and varying concentrations of chloride. The florescence increases at 480 nm due to the formation of pyrazoline product was monitored by PTI QM-40 fluorescent spectrophotometer with 320 nm excitation.

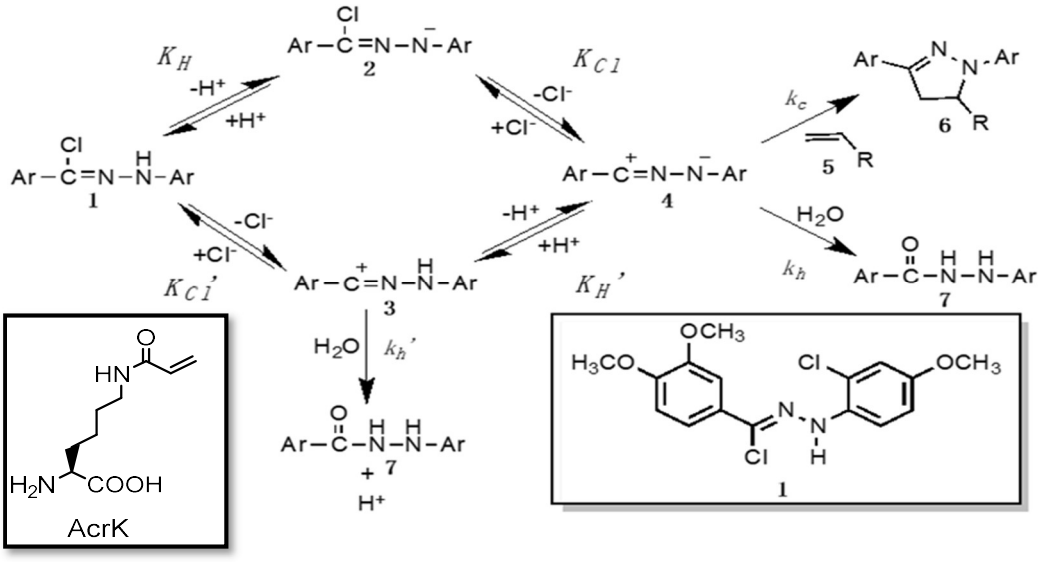


Figure 11 The nitrilimine-alkene reaction in the presence of chloride

2. Equation Derivation

As indicted by **Figure 11**, when $[H^+] >> K_H$ and $[Cl^+] >> K_{Cl}$, the formation rates of **6** and **7** and the consumption rate of **1** follow a relation defined by $\frac{d[1]}{dt} = \frac{d[6]}{dt} + \frac{d[7]}{dt}$, in which $\frac{d[1]}{dt} = k_c \cdot [5] \cdot [4]$, $\frac{d[7]}{dt} = k_h \cdot [4] + k_{h'} \cdot [3]$, $[3] = [1] \cdot \frac{K_{Cl'}}{[Cl^-]}$ and $[4] = [1] \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]}$. Therefore the consumption rate of **1** can be defined as $\frac{d[1]}{dt} = \left(k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} \cdot [5] + k_h \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} + k_{h'} \cdot \frac{K_{Cl'}}{[Cl^-]}\right) \cdot [1]$ that can be integrated to give $[1] = [1]_0 \cdot e^{-\left(k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} \cdot [5] + k_h \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} + k_{h'} \cdot \frac{K_{Cl'}}{[Cl^-]}\right)t}$. At anytime, $[6] + [7] = [1]_0 - [1]$ and $\frac{[7]}{[6]} = \frac{k_h \cdot [4] + k_{h'} \cdot [3]}{k_c \cdot [5] \cdot [4]} = \frac{k_h \cdot [4] + k_{h'} \cdot \frac{[H^+]}{K_H} \cdot [4]}{k_c \cdot [5] \cdot [4]} = \frac{k_h + k_{h'} \cdot \frac{[H^+]}{K_H}}{k_c \cdot [5]}$ that can be simplified to give $[6] = \frac{k_c \cdot [5]}{k_c \cdot [5] + k_h + k_{h'} \cdot \frac{[H^+]}{K_{H'}}} \cdot [1]_0 \cdot \left(1 - e^{-\left(k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} \cdot [5] + k_h \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} + k_{h'} \cdot \frac{K_{Cl'}}{[Cl^-]}\right)t}\right)$.

3. Protein Expression and Characterization

DNA sequence of sfGFP2TAG

Atgtagaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaattttctgtcc

gtggagagggtgaaggtgatgctacaaacggaaaactcacccttaaatttatttgcactactggaaaactacctgttccgtggcca

acacttgtcactactctgacctatggtgttcaatgcttttcccgttatccggatcacatgaaacggcatgacttttcaagagtgccat

gcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggacctacaagacgcgtgctgaagtcaagtttgaaggt

gataccttgttaatcgtatcgagttaaagggtattgattttaaagaagatggaaacattcttggacacaaactcgagtacaactttaa

ctcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaacttcaaaattcgccacaacgttgaagatggt

tccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccagacaaccattacctgtcgaca

caatctgtcctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagtttgtaactgctgctgggattacacatggc atggatgagctctacaaaggatcccatcaccatcaccatcactaa (The mutated site is highlighted in red)

sfGFPS2Acrk Expression and Purification

The pEVOL-PylT-PrKRS was constructed as previously reported.[6] Plasmid pEVOL-PylT-PrKRS plasmid and pET-sfGFPS2TAG were cotransformed with BL21 (DE3) and plate on agar plate containing ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). A single colony was picked and inoculated into 5 mL of LB medium with ampicillin and chloramphenicol. This overnight culture was used to inoculate 100 mL of LB medium. Cells were grown at 37℃ in an incubator (250 r.p.m.). When OD_{600}, 200 mM N^{ε}-acryloyl-lysine (AcrK), 1 mM IPTG and 0.2% arabinose were added. After 8 hour induction, cells were harvested by being centrifuged at 4000 g for 15 min, and resuspended in a lysis buffer (50 mM NaH_2PO_4, 250 mM NaCl, 10 mM imidazole, pH 8.0. The cell lysate was sonicated in an ice bath six times (2 min each pulse, 5 min interval for cooling), and centrifuged at 1000g for 60 min (4 ℃). The supernatant was isolated and incubated with 1 mL Ni-NTA resin (Qiagen) (1.5 h, 4 ℃), and then loaded to a column. The protein-resin mixture was washed with 50 mL of a wash buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 10 mM imidazole (pH 8.0), and eluted by a elution buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 250 mM imidazole (pH 8.0). The purified protein was concentrated and dialyzed into a buffer containing 10 mM ammonium bicarbonate. The protein was analyzed by 15% SDS-PAGE and stored at -80℃.

4. Protein Labeling with Hydrazonoyl Chloride **1**

pH dependence of the labeling reaction in the presence of chloride

Hydrazonoyl chloride **1**(5 mM, 15 μL) was added to the solution of sfGFP2Acrk (5 μM, 500 μL) in 1:1 acetonitrile-50 mM phosphate buffer with 50 mM chloride (pH 6-10) incubated at RT for 10 min and then quenched by 500 mM acrylamide. The labeled-protein was further purified using Ni-NTA resin (5 μL). The protein-bound resin was centrifuged (10 min, 13.4k) and washed with water for 4 times. After boiling the resin in 6X protein loading buffer (375 mM Tris-HCl, 10% SDS, 30% Glycerol, 0.03% Bromophenol blue, 600 mM DTT) and filtering, the labeled protein was released and subjected to 15% SDS-PAGE analysis. In-gel fluorescence detection was performed using BioRad ChemiDoc XRS+ Imaging system before the gel was stained by coomassie blue.

Chloride dependence of the labeling reaction at pH7

Hydrazonoyl chloride **1** (5 mM, 15 μL) was added to the solution of sfGFP2Acrk (5 μM, 500 μL) in 1:1 acetonitrile-50 mM phosphate buffer (pH7), with chloride concentration from 0 to 200 mM recpectively, incubated at RT for 30 min and then quenched by 500 mM acrylamide. The labeled-protein was further purified using Ni-NTA resin(5 μL). The protein-bound resin was centrifuged (10 min, 13.4k) and washed with water for 4 times. After boiling the resin in 6X protein loading buffer (375 mM Tris-HCl, 10% SDS, 30% Glycerol, 0.03% Bromophenol blue, 600 mM DTT) and filtering, the labeled protein was released subjected to 15% SDS-PAGE analysis. In-gel fluorescence detection was performed using BioRad ChemiDoc XRS+ Imaging system before the gel was stained by coomassie blue.

The labeling reaction at pH 10 without chloride

Hydrazonoyl chloride **1** (5 mM, 15 μL) was added to the solution of sfGFP2Acrk (5 μM, 500 μL) in a 1:1 acetonitrile-50mM phosphate buffer (pH10 without chloride), incubated at for 1, 2, 3, 4 min respectively and then quenched by 500 mM acrylamide. The labeled-protein was further purified using Ni-NTA resin(5 μL). The protein-bound resin was centrifuged (10 min, 13.4k) and washed with water for 4 times. After boiling the resin in 6X protein loading buffer (375 mM Tris-HCl, 10% SDS, 30% Glycerol, 0.03% Bromophenol blue, 600 mM DTT) and filtering, the labeled protein was released and subjected to 15% SDS-PAGE analysis. In-gel fluorescence detection was performed using BioRad ChemiDoc XRS+ Imaging system before the gel was stained by coomassie blue.

Labelling with different proteins

Hydrazonoyl chloride **1** (5 mM, 15 μL) was added to different solutions: 1: sfGFP2AcrK; 2: sfGFP-p53 peptide fusion; 3: BSA; 4: lysozyme; 5: sfGFP; 6: sfGFP (all proteins are 5 μM, 500 μL) in a 1:1 acetonitrile-50mM phosphate buffer (pH10 without chloride), incubated at for 1, 2, 3, 4 min respectively and then quenched by 500 mM acrylamide. The labeled-protein was further purified using Ni-NTA resin (5 μL). The protein-bound resin was centrifuged (10 min, 13.4k) and washed with water for 4 times. After boiling the resin in 6X protein loading buffer (375 mM Tris-HCl, 10% SDS, 30% Glycerol, 0.03% Bromophenol blue, 600 mM DTT) and filtering, the labeled protein was released and subjected to 15% SDS-PAGE analysis. In-gel fluorescence detection was performed using BioRad ChemiDoc XRS+ Imaging system before the gel was stained by coomassie blue.

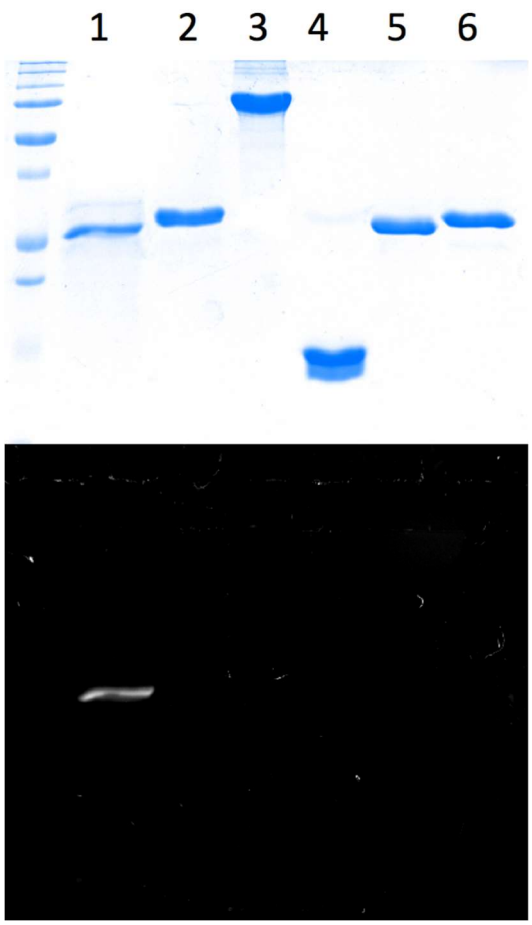


Figure 12 The selective labeling of sfGFP2AcrK with diarylnitrilimine. Proteins (1: sfGFP2AcrK; 2: sfGFP-p53 peptide fusion; 3: BSA; 4: lysozyme; 5: sfGFP; 6: sfGFP incorporated with a meta-triflоromethyl-phenylalanine at its S2 site) were incubated with 5 mM hydrazonyl chloride 1 for 20 min before they were analyzed by SDS-PAGE. The top panel shows the Coomassie-blue staining of the gel and the bottom panel shows the fluorescent imaging of the gel before it was Coomassie-blue stained. The data clearly demonstrated the specific reaction between acrylamide and nitrilimine.

Results and Discussion

We chose a hydrozonoyl chloride (**1** in **Figure 13**) as a nitrilimine precursor for our kinetic analysis due to the difficulty of quantitative photolysis of a tetrazole to form a nitrilimine that nonetheless reacts with water and chloride in aqueous conditions to form a hydrozonoyl chloride. In an aqueous buffer with a high concentration of chloride ion, **1** presumably undergoes two parallel two-step ionization processes to lose a proton cation and a chloride anion to generate a nitrilimine **4** as shown in **Figure 13**. **4** then reacts either with **5** to form a fluorescent cycloaddition product **6** or with water to produce a hydrazide **7**. **7** can also be made from the reaction of water with the immediate dechlorination product **3** of **1**. These two parallel ionization processes of hydrozonoyl halide in aqueous conditions were studied and demonstrated previously.[77] Assuming the ionization of **1** to form **4** is a fast equilibrium, the formation of **6** will follow

$$[\mathbf{6}] = \frac{k_c \cdot [\mathbf{5}]}{k_c \cdot [\mathbf{5}] + k_h + k_h' \cdot \frac{[H^+]}{K_{H'}}} \cdot [\mathbf{1}]_0 \cdot \left(1 - e^{-\left(k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} \cdot [\mathbf{5}] + k_h \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} + k_h' \cdot \frac{K_{Cl'}}{[Cl^-]}\right) t}\right) \quad (1)$$

when **5** is excessive, $[H^+] \gg K_H$, and $[Cl^-] \gg K_{Cl}$. In eqn (1), $[\mathbf{1}]_0$ represents the initial concentration of **1**. Based on eqn (1), the pseudo first-order reaction rate constant of the **6** formation can be described as

$$k_{app} = k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} \cdot [\mathbf{5}] + k_h \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]} + k_h' \cdot \frac{K_{Cl'}}{[Cl^-]}. \quad (2)$$

This equation can be further simplified as

$$k_{app} = k_{c(obs)} \cdot [\mathbf{5}] + k_{h(obs)} \quad (3)$$

Where

$$\boldsymbol{k_{c(obs)} = k_c \cdot \frac{K_H \cdot K_{Cl}}{[H^+] \cdot [Cl^-]}}. \quad (4)$$

As **6** is highly fluorescent, its formation can be facilely detected using a fluorescent spectrophotometer and analysed to obtain k_{app}. The determined k_{app} values at varied concentrations of **5,** at a given pH and a chloride concentration can be readily fitted to eqn (3) to obtain $k_{c(obs)}$. In principle, the determined $k_{c(obs)}$ values at varying pH and chloride concentrations will allow to assess k_c, the second-order rate constant of the nitrilimine-alkene cycloaddition in aqueous conditions.

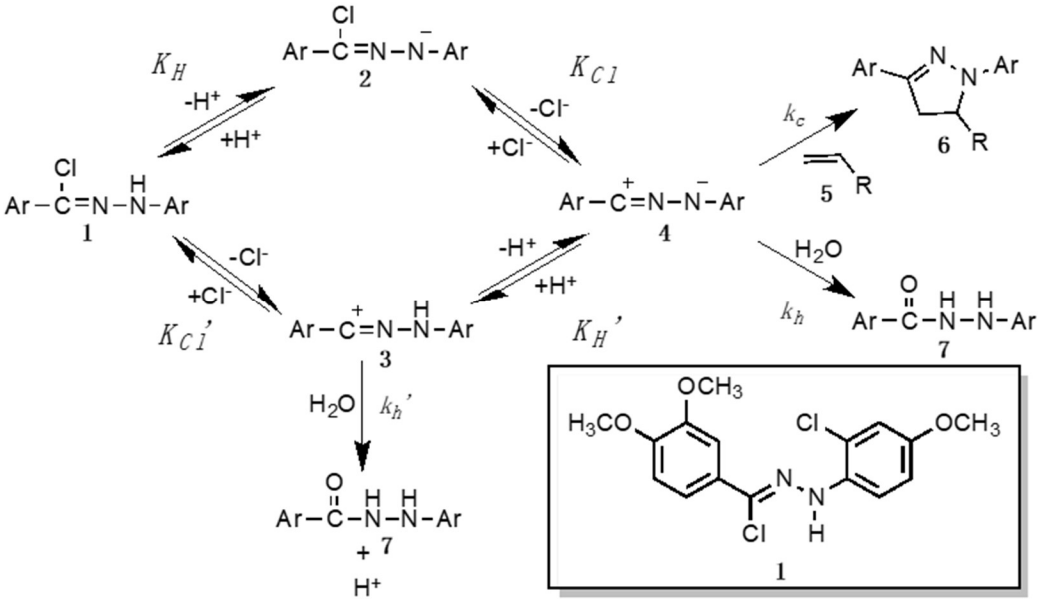


Figure 13 The ionization of diarylhydrozonoyl chloride to form diarylnitrilimine in aqueous conditions with chloride and the subsequent reactions with alkene and water.

We first studied the pH dependence of $k_{c(obs)}$ at 50 mM chloride. Reactions between 5 M **1** and varying concentrations of acrylamide at five given pH (6-10) were monitored using a PTI QM-40 fluorescent spectrophotometer with an excitation light at 320 nm and an emission detection at 480 nm. The fluorescence incremental data were fitted to a single exponential increase equation to obtain k_{app}. The resolved k_{app} values were then plotted against the acrylamide concentrations. As shown in **Figure 14**, k_{app} is linearly dependent on the acrylamide concentration at a given pH and the data were readily used to determine $k_{c(obs)}$, validating the mechanism proposed in **Figure 13**. Although log ($k_{c(obs)}$) shows a linear dependence on pH (**Figure 14A**) as eq. (4) predicts, (eq. (4) can be transformed as $\boldsymbol{log\left(k_{c(obs)}\right) = log\left(\frac{k_c \cdot K_H \cdot K_{Cl}}{[Cl^-]}\right) + pH}$), the data can not be simply fitted to eqn (4) and best fitted to

$$\boldsymbol{k_{c(obs)} = k_c \cdot \frac{(K_H)^x \cdot K_{Cl}}{[H^+]^x \cdot [Cl^-]}} \quad (5)$$

with an x value as 0.64±0.01 and $\boldsymbol{k_c \cdot \frac{(K_H)^{0.64} \cdot K_{Cl}}{[Cl^-]}}$ as (1.42±0.01) $\times 10^{-5}$. Any deviation from eq. (4) may be due to the presence of chloride that changes the proton activity during the ionization process. This is highly possible as a similar deviation was not observed during the kinetic analyses in reaction conditions without chloride, which will be presented later. **Figure 14A** clearly shows that the observed cycloaddition rate constant increased about 200 folds when pH was changed from 6 to 10. Therefore, when an acrylamide-containing protein is separately labeled with **1** at different pH, faster labeling rates are expected at higher pH values. To approve this, we performed the labeling of sfGFP2AcrK (a superfolder green fluorescent protein with N^{ε}-acryloyl-

lysine (AcrK) incorporated at its S2 position) by 150μM **1** for 15 min at 50 mM chloride and pH from 6 to 10. AcrK is a noncanonical amino acid with an acrylamide moiety that can be genetically incorporated into proteins at amber mutations sites in *E. coli* using an evolved pyrrolysyl-tRNA synthetase-$\mathbf{tRNA}_{\mathbf{CUA}}^{\mathbf{Pyl}}$ pair. The expression of sfGFP2AcrK followed a method described in a previous publication.[6] Presented in **Figure 14B**, the labeling efficiency is clearly pH dependent, with higher pH leading to more efficient labeling.

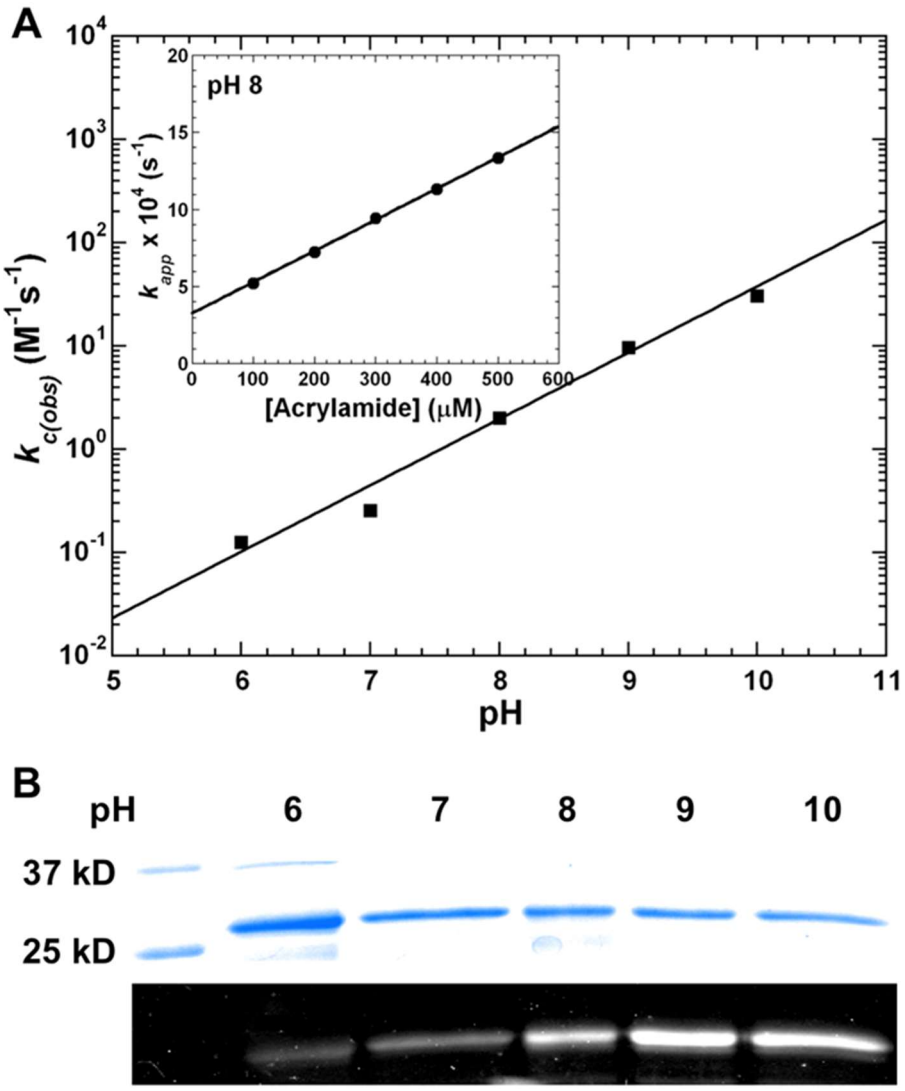


Figure 14 (**A**) the pH dependence of $k_{c(obs)}$. The inset shows the acrylamide concentration dependence of k_{app} at pH 8 and 50 mM chloride in acetonitrile-50 mM phosphate buffer (1:1). (**B**) The labeling efficiency of sfGFP2AcrK by **1** at different pH. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50mM phosphate buffer (1:1) for 15 min before 500 mM acrylamide was added to sequestrate **1** from reacting with sfGFP2AcrK and then the labeling solutions were analyzed by SDS-PAGE. The top panel shows the Coomassie blue stained gel and the bottom panel presents a fluorescent image of the same gel before it was stained by Coomassie blue. The fluorescent imaging was performed with a BioRad ChemDoc XRS system under UV irradiation.

Eqn (5) also indicates an inverse linear dependence of k_{app} on the chloride concentration, which has been approved by our kinetic studies performed in varying chloride concentrations and pH 9. At a particular chloride concentration (10-100 mM), the determined k_{app} values are linearly dependent on the acrylamide concentrations, which were used to obtain $k_{c(obs)}$. Plotting $k_{c(obs)}$ against 1/[Cl^-] shows that this dependence is linear (**Figure 15 A**). We also did similar kinetic analyses at 1 mM chloride. Although the determined k_{app} values are much higher than those determined at higher chloride concentrations, the k_{app} values at different acrylamide concentrations are almost constant, and therefore not valid for the calculation of $k_{c(obs)}$. It is possible that at a low chloride concentration, the two dechlorination processes (**2** to **4** and **1** to **3**) do not reach fast equilibria, invalidating **Figure 13** and eqn (1) in data analysis. This study clearly shows a strong inhibitory effect of chloride on the nitrilimine-alkene cycloaddition, indicating that applying the nitrilimine-alkene cycloaddition for protein labeling needs to avoid running the reaction at a high chloride concentration. This is exactly what we observed in the labeling reactions of sfGFP2AcrK by **1** at pH 7 and different chloride concentrations (**Figure 15B**). A 30 min labeling reaction in the absence of chloride led to an intensely fluorescently labeled protein. The labeling efficiency gradually diminished when the chloride concentration was increased from 0 to 200 mM.

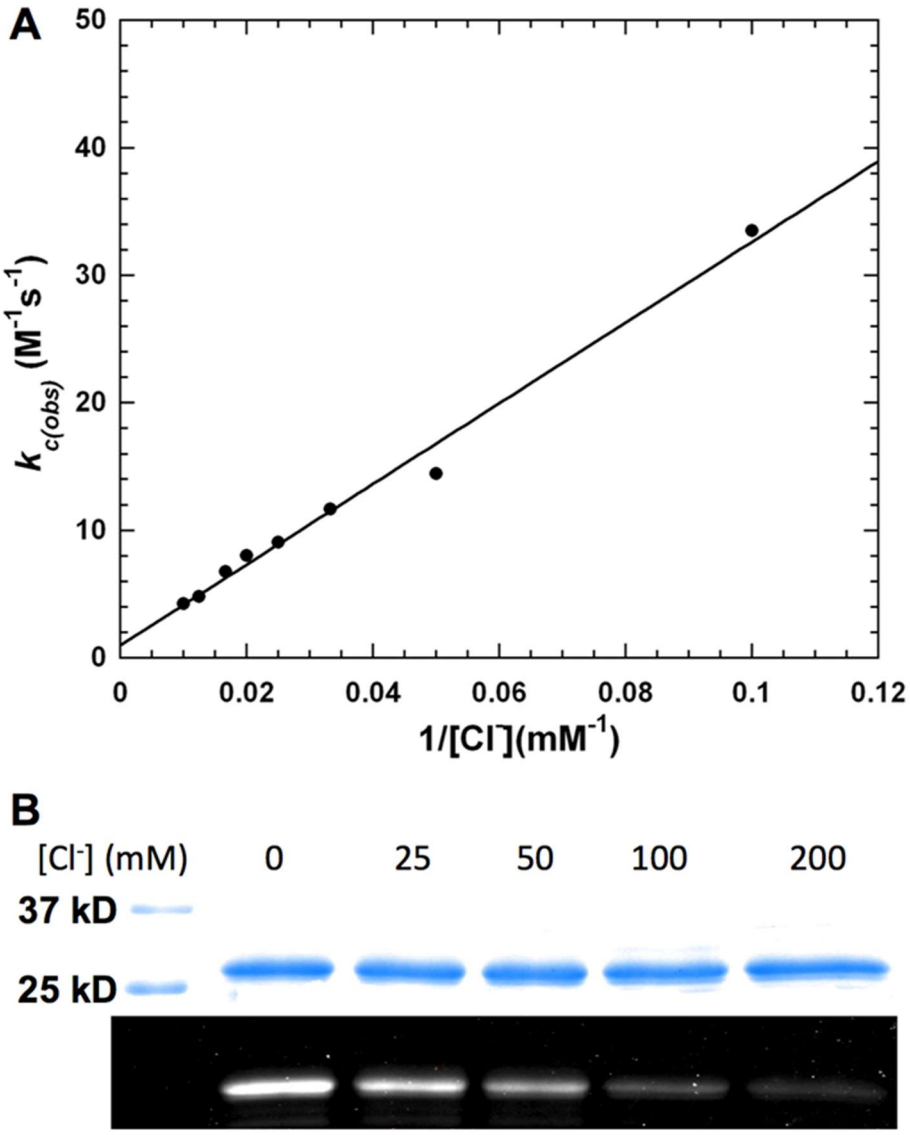


Figure 15 (**A**) the chloride dependence of $k_{c(obs)}$. (**B**) The labeling efficiency of sfGFP2AcrK by **1** at pH 7 and different chloride concentrations. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50mM phosphate buffer (1:1), pH 7, and varying chloride concentrations for 30 min before 500 mM acrylamide was added to sequestrate **1** from reacting with sfGFP2AcrK and then the labeling solutions were analyzed by SDS-PAGE. The top panel shows the Coomassie blue stained gel and the bottom panel presents a fluorescent image of the same gel before it was stained by Coomassie blue.

The aforementioned kinetic analyses at 1 mM chloride prompted us to look into the reaction kinetics of the nitrilimine-alkene cycloaddition in the absence of chloride. Without chloride, the ionization of **1** and subsequent reactions with water and alkene will presumably follow a mechanism presented in **Figure 16**. The two dechlorination steps become rate limiting. Since the ionization of **1** will generate an equal amount of chloride in reaction conditions, providing **1** at a concentration much lower than 1 mM will not eviscerate the mechanism shown in **Figure 16**. **Figure 16** depicts the formation of **6** following

$$[\mathbf{6}] = \frac{k_c \cdot [\mathbf{5}]}{k_c \cdot [\mathbf{5}] + k_h + k_{h'} \cdot \frac{[H^+]}{K_{H'}}} \cdot [\mathbf{1}]_0 \cdot (\mathbf{1} - e^{-\left(k_{Cl} \cdot \frac{K_H}{[H^+]} + k_{Cl'}\right)t}) \quad (6)$$

that results in an apparent rate constant defined as

$$k_{app} = k_{Cl} \cdot \frac{K_H}{[H^+]} + k_{Cl'}. \quad (7)$$

Eqn (7) shows that k_{app} is inversely proportional to the proton centration but not related to the provided acrylamide, which was observed during kinetic analyses of the nitrilimine-alkene cycloaddition in the absence of chloride. When reactions between 5 μM **1** and different acrylamide concentrations were performed at a specific pH without chloride, they all resulted in a same reaction rate constant (inset of **Figure 17A**). Raising the pH value led to higher reaction rate constants. The logarithms of determined k_{app} values as a function of pH are presented in **Figure 17A**, which can be well fitted to eqn (7). At pH 10 with chloride, the determined k_{app} is 0.111±0.002 s^{-1}. Since this rate constant is not related to the concentrations of both **1** and the acrylamide, so using **1** to label a protein with an acrylamide moiety at any concentrations of **1** and

the protein will have a labeling half life close to 6 s when the chloride anion is absent in labeling conditions, further resulting in almost instantaneous protein labeling. To demonstrate this rapid labeling process, we tested the labeling of sfGFP2AcrK by **1** at pH 10 for different lapses of time. As shown in **Figure 17B**, labeling sfGFP2AcrK with **1** for 1 min to 4 min all led to an intensely fluorescently labeled protein with an equally fluorescent intensity, implying that the labeling reaction was mostly completed within 1 min.

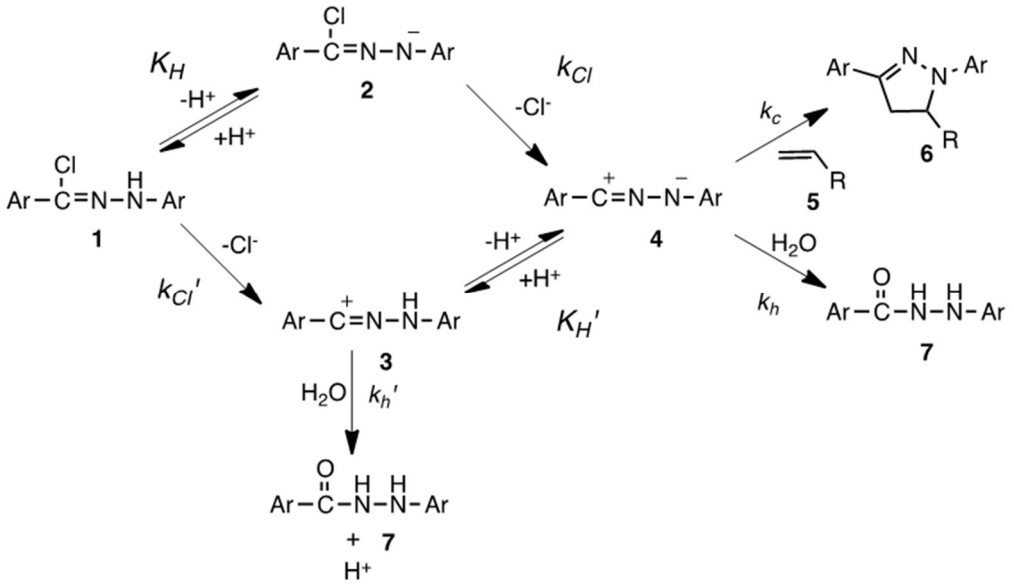


Figure 16 The ionization of diarylhydrozonoyl chloride to form diarylnitrilimine in aqueous conditions without chloride and the subsequent reactions with alkene and water.

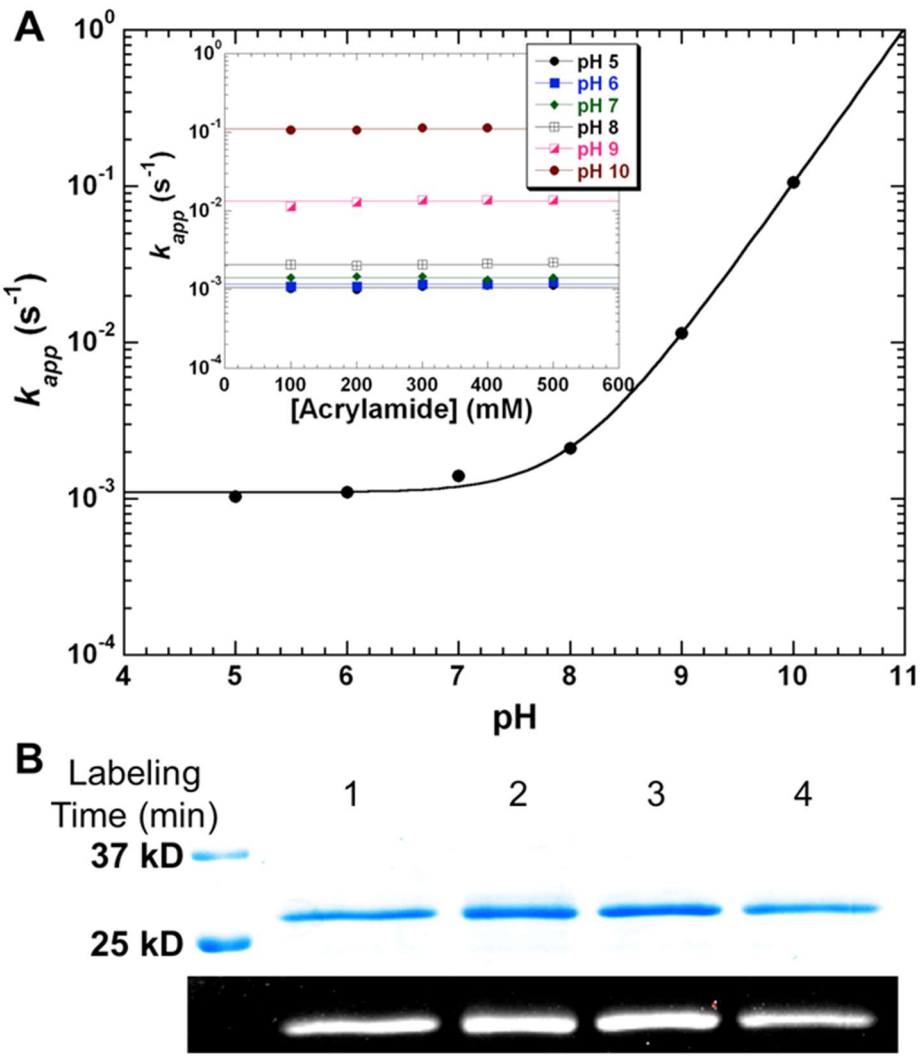


Figure 17 (**A**) the pH dependence of k_{app} in the absence of chloride. The data were fitted to eqn (7). The inset shows the acrylamide concentration dependence of k_{app} at pH 5-10 in acetonitrile-50 mM phosphate buffer (1:1) but absence of chloride. All determined k_{app} values at a given pH are similar. (**B**) The labeling efficiency of sfGFP2AcrK by **1** at pH 10 without chloride. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50mM phosphate buffer (1:1) with no chloride provided for different lapses of time (1-4 min) before 500 mM acrylamide was added to sequestrate **1** from reacting with sfGFP2AcrK and then the labeling solutions were analyzed by SDS-PAGE. The top panel shows the Coomassie blue stained gel and the bottom panel presents a fluorescent image of the same gel before it was stained by Coomassie blue.

Conclusions

Being a catalyst-free click reaction type, nitrilimine-alkene cycloaddition has been explored for click and photo-click labeling of proteins. All previous kinetic characterizations of the nitrilimine-alkene cycloaddition were completed in PBS buffers that contained high concentrations of chloride. The current study clearly shows that all previously measured second-order rate constants of the nitrilimine-alkene cycloaddition are apparent second-order rate constants that are significantly influenced by pH and chloride concentrations. Based on eqn (5), one would need to determine K_H and K_{Cl} to calculate k_c, the second-order rate constant of the nitrilimine-alkene cycloaddition. When we derived eqn (1), we applied the following preconditions; $[H^+] >> K_H$ and $[Cl^-] >> K_{Cl}$. When these conditions do not hold, the determined apparent rate constant in theory will in theory follow

$$\boldsymbol{k_{c(obs)} = k_c \cdot \frac{K_H \cdot K_{Cl}}{([H^+]+K_H)\cdot([Cl^-]+K_{Cl})}} \quad (8)$$

but should be best described as

$$\boldsymbol{k_{c(obs)} = k_c \cdot \frac{(K_H)^{0.64} \cdot K_{Cl}}{([H^+]+K_H)^{0.64}\cdot([Cl^-]+K_{Cl})}} \quad (9)$$

due to the proton activity deviation from what is indicated by the pH. As indicated by eqn (9), in a specific chloride concentration, $k_{c(obs)}$ will reach a plateau when $[H^+] << K_H$. As we did not observe the trend of $k_{c(obs)}$ becoming saturated to pH 10, a safe guess of a K_H value is small than 10^{-12}. Similarly, $k_{c(obs)}$ showed an inversely proportional dependence of the chloride concentration when lower than 10 mM. A safe estimate of a K_{Cl} value is small than 10^{-3}. We have determined that at 50 mM chloride, $\boldsymbol{k_c \cdot \frac{(K_H)^{0.64} \cdot K_{Cl}}{[Cl^-]}}$ is 1.42×10^{-5}. With

two estimated values of K_H and K_{Cl}, we can easily determine a k_c value higher than 3.4×10^4 $M^{-1}s^{-1}$. This reaction rate constant is comparable to that of the rapid transcyclooctene-tetrazine cycloaddition and makes the nitrilimine-alkene cycloaddition one of the fastest click reactions. Another implication of our study is the possibility of different labeling kinetics in extracellular and intracellular spaces when the nitrilimine–alkene reaction is applied for in *vivo* labeling. Mammalian cells maintain an intracellular chloride concentration which is much lower than their extracellular environments.[78] This large chloride concentration variation may allow for the application of the nitrilimine–alkene reaction to specifically achieve intracellular protein sensitization.

CHAPTER III

PHAGE-DISPLAYED CYCLIC PEPTIDE LIBRARIES FOR THE IDENTIFICATION OF INHIBITORS OF HISTONE DEACETYLASES

Introduction

Histone deacetylases (HDAC8) comprise an indispensable enzyme class for humans. Aberrant overexpression of HDACs has been demonstrated to be related with many human diseases such as cancers. Thus, discovery of HDAC inhibitors are very import to therapeutic studies. Current widely used HDAC inhibitors includes vorinostat (SAHA), trichostatin A (TSA), and Romidepsin (FK228). Although these inhibitors have shown excellent inhibition ability through chelation of metal ion at active site, selective inhibition is a drawback due to non-specificity to most HDACs.

To investigated more potential HDAC inhibitors with high affinity and specificity, focus has been raised on peptide-based molecules. Due to advantages such as high binding affinity, cell permeability, less toxicity, high target specificity, peptide-based molecules have emerged as indispensable leads for drug discovery. Moreover, development of solid-phase peptide synthesis (SPPS) makes peptide-based molecules synthesized efficiently and conveniently.[79] Amongst these peptide drugs, cyclic peptides are very attractive. Compared to linear counterparts, cyclic peptides have dominant advantages including conformational rigidity and stability against proteolysis.

Recent years have seen numerous reports on the construction of cyclic peptide libraries. Amongst these methodologies, phage display is a highly powerful tool in high-throughput screening of enzyme inhibitors. Phage display expresses foreign peptides on

the surface of filamentous phages. By linking phenotype to genotype, phage display can be used to build peptide library from the construction of phage library in DNA level. The ease of genetic manipulation makes phage display an outstanding biological platform. However, one pitfall of the phage-displayed peptide library is the limitation of 20 canonical amino acids. To overcome this shortage, one strategy is using chemoselective bioconjugation and surface chemistry techniques. Another strategy is to incorporate non-canonical amino acids which can efficiently expand chemical diversity as well as provide more strategies of cyclization. Compared to bioconjugation methodologies which require the installation of biorthogonal reagents, the incorporation of NCAA is more straightforward and easily manipulated.

The thiol group on cysteine is very reactive and undergoes a wide variety of reactions, such as Michael addition and nucleophilic substitution. To generate a cyclic peptide, many researchers have focused on bis-cysteine cyclic peptides [11]. The most well-studied approach of the cyclization was through an intramolecular disulfide bridge between two internal cysteine thiol groups. O'neil and co-workers firstly reported phage-displayed cyclic peptide library consists of two cysteines and six randomized amino acids in between.[80] High-affinity cyclic peptides with disulfide bonds are selected as inhibitors of platelet glycoprotein IIb/IIIa.[80,81] However, the disulfide bond is reducible, and therefore its application *in vivo* is limited. To overcome the limitation of disulfide bonds, we explored reactions using a cysteine with other functional groups instead of a reaction between two cysteines. To expand the chemical diversity of cyclic peptides, we used a non-canonical amino acid (NCAA), rather than any canonical amino acid, as a cyclization reagent. The incorporation of NCAA into a peptide is based on stop codon suppression. In

this approach, a nonsense stop codon (such as UAG) is used to direct the incorporation of NCAAs when a particular NCAA-corresponding aminoacyl-tRNA synthetase (aaRS)-tRNA pair is present in cells.

In this dissertation the well-studied 1,4-addition between cysteine and the acrylamide moiety is used as cyclization approach **(Figure 18**). The second order rate constant of the reaction between acrylamide and β-mercaptoethanol (β-ME) is 0.013 ± 0.001 M^{-1} s^{-1} at pH 7.4[11]. Our group recently synthesized two acrylamide-containing NCAAs, N^{ε}-acryloyl-L-lysine (AcrK) and N^{ε}-crotonyl-L-lysine (CrtK), and incorporated them into proteins at amber mutation sites. Then the reaction between AcrK and β-ME was investigated. Results show that in a slightly alkali environment (pH~8), both proteins bearing these NCAAs can be labeled with β-ME efficiently[6]. Hereby we used this efficient 1,4-cycloaddition for the construction of phage-displayed cyclic peptide library.

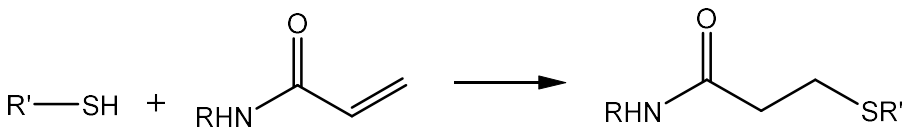


Figure 18 1,4–addition between cysteine and the acrylamide moiety

Experimental Methods

1. Synthesis of NCAAs

Synthesis of N^{ε}-acryloyl-L-lysine (AcrK)[6]

To a solution of N-hydroxysuccinimide (1.3 g, 11.3 mmol) in anhydrous dichloromethane (25 mL) was added N,N-diisopropylethylamine (1.5 mL, 8.9 mmol), followed by dropwise addition of acryloyl chloride (0.8 mL, 9.3 mmol) in ice bath over 10 min. The mixture was then stirred for 10 hours at room temperature. The mixture was extracted with ethyl acetate (50 mL), washed with saturated NH_4Cl solution and brine, and dried with anhydrous sodium sulfate. The solution was filtered and evaporated under vacuum to give **2** (1.5 g) as yellow oil.

To a solution of copper(II) sulfate pentahydrate (1.0 g, 4.0 mmol) in water (50 mL) was added lysine hydrochloride (1.5g, 8.0 mmol) and sodium bicarbonate (1.9g, 22.4 mmol). The mixture was stirred at room temperature for 20 min. A solution of compound **2** in acetone (20 mL) was added to the reaction. The reaction mixture continued stirring at room temperature for additional 10 h. The blue mixture was filtered, and the blue filter cake was washed with water (three times) and acetone, dissolved in a solution of water and chloroform (v/v = 1:1, 100 mL), stirred at room temperature for 20 min. To the suspension was added 8-hydroxyquinoline (1.6 g, 11.0 mmol). The solution was stirred at room temperature for 30 min. The green suspension was filtered, and the filtrate was washed with chloroform for three times, concentrated under reduced pressure, subjected to ion-exchange chromatography for further purification, and give N^{ε}-acryloyl-L-lysine as a white powder (1.2 g, 60%).

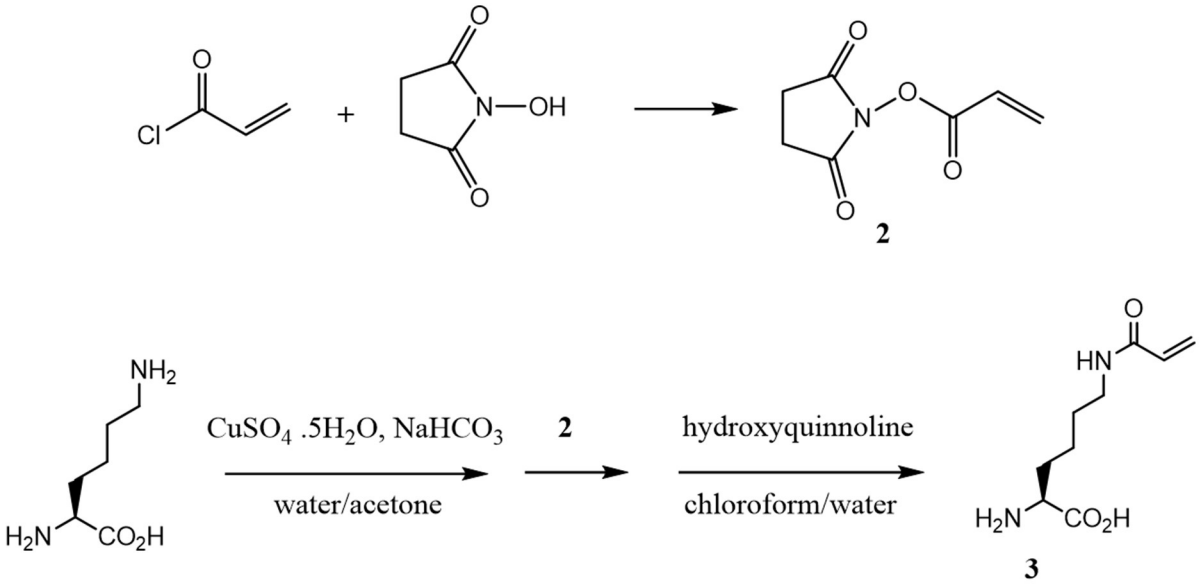


Figure 19 Synthesis of AcrK.

Synthesis of N^{ε}-crotonyl-L-lysine (CrtK)[6]

A suspension of Boc-Lys(Cbz)-OMe (4.00 g, 10.1 mmol) and palladium on carbon (10% Pd, 0.67g, 0.64 mmol) was stirred under a hydrogen balloon at roomt temperature for 3 h. The mixture was filtered over a celite pad, and the filtrate was concentrated to give **4** (2.51 g) as a clear oil. The product is pure enough for the next step.

To a solution of **4** (2.51 g, 9.6 mmol) and crotonic acid (1.70g, 19.2 mmol) in anhydrous dichloromethane (50 mL) was added dimmethylaminolpyridine (0.048 g, 3.8 mmol), followed by dropwise addition of 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (2.3 g, 12 mmol) in anhydrous dichloromethane (10 mL) over 10 min.The mixture was stirred at room temperature for 15 h, extracted by ethyl acetate (50 mL), washed with brine (50 mL x 3), dried with anhydrous $MgSO_4$, evaporated and subjected to chromatography to afford **5** (2.66 g) as a clear oil for next step without further purification.

To a solution of **5** (2.66 g, 8.0 mmol) in tetrahydrofuran (25 mL) was added lithium hydroxide solution (1 M, 40 mL, 40 mmol) in an ice-water bath. The reaction was stirred in the ice-water bath for 2 h. The reaction mixture was diluted with water (25 mL), extracted with ethyl acetate (25 mL x 3) to remove impurities. The aqueous layer was adjusted to pH 3 with 6 M hydrochloric acid solution and extracted with ethyl acetate (25 mL x 2). The combined ethyl acetate solution was washed with brine (25 mL x 3), dried with anhydrous $MgSO_4$, concentrated under reduced pressure to afford a colorless oil for next step without further purification.

To a solution of hydrochloric acid in dioxane (4 M, 10 mL, 40 mmol) was added the oil from previous step. The reaction was stirred at room temperature for 4 h, diluted

with dichloromethane and evaporated under reduced pressure to give **6** (1.35g, 70% for three steps) as a white powder.

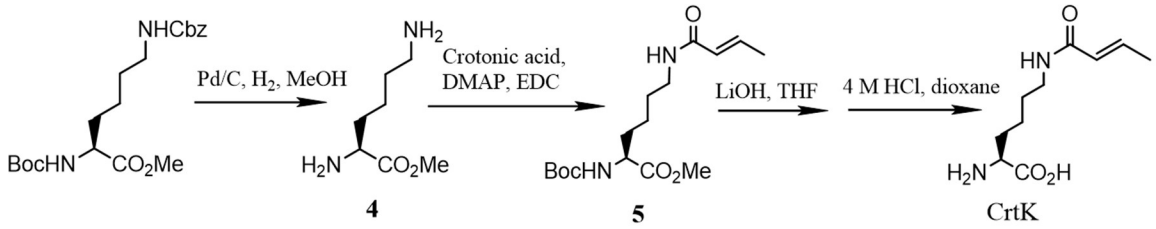


Figure 20 Synthesis of CrtK

2. sfGFP mutant Expression, Purification and Labeling

DNA sequence of Met-Ala-(Ala)$_5$-TAG-sfGFP

atggccgccgccgccgccgcctagaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaat
gggcacaaattttctgtccgtggagagggtgaaggtgatgctacaaacggaaaactcacccttaaatttatttgcactactggaaa
actacctgttccgtggccaacacttgtcactactctgacctatggtgttcaatgcttttcccgttatccggatcacatgaaacggcatg
acttttttcaagagtgccatgcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggacctacaagacgcgtgct
gaagtcaagtttgaaggtgataccttgttaatcgtatcgagttaaagggtattgattttaaagaagatggaaacattcttggacaca
aactcgagtacaactttaactcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaacttcaaaattcgc
cacaacgttgaagatggttccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccagac
aaccattacctgtcgacacaatctgtcctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagtttgtaactgctg
ctgggattacacatggcatggatgagctctacaaaggatcccatcaccatcaccatcactaa (mutated sites are highlighted in red)

DNA sequence of Met-Cys-(Ala)$_5$-TAG-sfGFP

atgtgcgccgccgccgccgcctagaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaat
gggcacaaattttctgtccgtggagagggtgaaggtgatgctacaaacggaaaactcacccttaaatttatttgcactactggaaa
actacctgttccgtggccaacacttgtcactactctgacctatggtgttcaatgcttttcccgttatccggatcacatgaaacggcatg
acttttttcaagagtgccatgcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggacctacaagacgcgtgct
gaagtcaagtttgaaggtgataccttgttaatcgtatcgagttaaagggtattgattttaaagaagatggaaacattcttggacaca
aactcgagtacaactttaactcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaacttcaaaattcgc
cacaacgttgaagatggttccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccagac
aaccattacctgtcgacacaatctgtcctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagtttgtaactgctg

ctgggattacacatggcatggatgagctctacaaaggatcccatcaccatcaccatcactaa (mutated sites are highlighted in red)

sfGFP Derivatives Expression and Purification

Plasmid pEVOL-PylT-PrKRS and pEVOL-PylT-BuKRS were constructed as previously reported.[6] pEVOL-PylT-PrKRS plasmid and pET-Met-Cys-$(Ala)_5$-TAG-sfGFP or pET-Met-Ala-$(Ala)_5$-TAG-sfGFP were cotransformed with BL21 (DE3) and plate on agar plate containing ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). A single colony was picked and inoculated into 5 mL of LB medium with ampicillin and chloramphenicol. This overnight culture was used to inoculate 100 mL of LB medium. Cells were grown at 37℃ in an incubator (250 r.p.m.). When OD_{600}, 200 mM AcrK or CrtK, 1 mM IPTG and 0.2% arabinose were added. After 8-hour induction, cells were harvested by being centrifuged at 4000 g for 15 min and resuspended in a lysis buffer (50 mM NaH_2PO_4, 250 mM NaCl, 10 mM imidazole, pH 8.0. The cell lysate was sonicated in an ice bath six times (2 min each pulse, 5 min interval for cooling), and centrifuged at 1000g for 60 min (4 ℃). The supernatant was isolated and incubated with 1 mL Ni-NTA resin (Qiagen) (1.5 h, 4 ℃), and then loaded to a column. The protein-resin mixture was washed with 50 mL of a wash buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 10 mM imidazole (pH 8.0), and eluted by a elution buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 250 mM imidazole (pH 8.0). The purified protein was concentrated and dialyzed into a buffer containing 10 mM ammonium bicarbonate. The protein was analyzed by 15% SDS-PAGE and stored at -80℃.

Protein Labeling with Hydrazonoyl Chloride 1

Hydrazonoyl chloride **1** (5 mM, 15 μL) was added to four different solutions of Met-Cys-$(Ala)_5$-AcrK-sfGFP, Met-Ala-$(Ala)_5$-AcrK-sfGFP, Met-Cys-$(Ala)_5$-CrtK-sfGFP and Met-Ala-$(Ala)_5$-CrtK-sfGFP after (5 μM, 500 μL) in a 1:1 acetonitrile-50mM phosphate buffer (pH10 without chloride), incubated at for 10 min respectively and then quenched by 500 mM acrylamide. The labeled-protein was further purified using Ni-NTA resin(5 μL). The protein-bound resin was centrifuged (10 min, 13.4k) and washed with water for 4 times. After boiling the resin in 6X protein loading buffer (375 mM Tris-HCl, 10% SDS, 30% Glycerol, 0.03% Bromophenol blue, 600 mM DTT) and filtering, the labeled protein was released and subjected to 15% SDS-PAGE analysis. In-gel fluorescence detection was performed using BioRad ChemiDoc XRS+ Imaging system before the gel was stained by Coomassie blue.

3. Target Protein Expression and Purification

Protein Sequence of HDAC8

meepeepadsgqslvpvyiyspeyvsmcdslakipkrasmvhslieayalhkqmrivkpkvasmeematfhtdaylqh
lqkvsqegdddhpdsieyglgydcpategifdyaaaiggatitaaqclidgmckvainwsggwhhakkdeasgfcylnda
vlgilrlrrkferilyvdldlhhgdgvedafsftskvmtvslhkfspgffpgtgdvsdvglgkgryysvnvpiqdgiqdekyy
qicesvlkevyqafnpkavvlqlgadtiagdpmcsfnmtpvgigkclkyilqwqlatlilggggynlantarcwtyltgvilg
ktlsseipdhefftaygpdyvleitpscrpdrnephriqqilnyikgnlkhvlenlyfqgdydipttlehhhhhh

HDAC8 Expression and Purification

Plasmid pHD4-HDAC8-TEV-His6 was transformed with BL21-CodonPlus(DE3). A single colony was picked and inoculated into 5 mL of 2xYL supplemented with

ampicillin (Amp) (100 μg/mL). This overnight culture was used to inoculate 500 mL of autioinduction TB medium (24 g/L yeast extract, 12 g/L tryptone, 8 g/L tris, 4 g/L lactose, 1 g/L glycerol, pH 7.5) supplemented with ampicilin (100 μg/mL) and 200 μM $ZnSO_4$. Cells were grown at 37℃ in an incubator (250 r.p.m.). After 20-hour auto-induction, cells were harvested by being centrifuged at 4000 g for 15 min, and resuspended in a lysis buffer (50 mM NaH_2PO_4, 250 mM NaCl, 10 mM imidazole, pH 8.0. The cell lysate was sonicated in an ice bath six times (3 min each pulse, 6 min interval for cooling), and centrifuged at 1000g for 60 min (4 ℃). The supernatant was isolated and incubated with 1 mL Ni-NTA resin (Qiagen) (1.5 h, 4 ℃), and then loaded to a column. The protein-resin mixture was washed with 50 mL of a wash buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 10 mM imidazole (pH 8.0), and eluted by a elution buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 250 mM imidazole (pH 8.0). Eluted fractions were collected, concentrated and subjected to a Q Sepharose column (GE Healthcare) for further purification. The purified protein was dialyzed into a dialysis buffer (25 mM Tris - HCl, 300 mM NaCl, 200 μM $ZnSO_4$, 5 μM KCl, pH7.5). The protein was stored as 5 μM aliquots at -80℃.

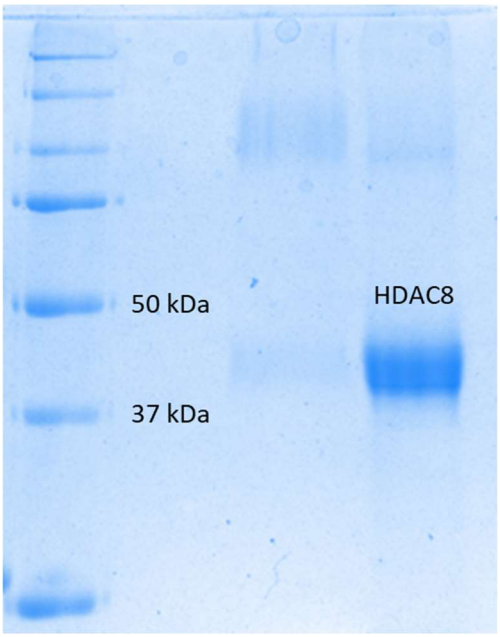


Figure 21 15% SDS Page gel imaging of purified HDAC8.

TEV Protease Expression and Purification

BL21(DE3) transformed with N-terminal His-tagged TEV protease was grown in 5 mL of LB medium at 37ºC. This overnight culture was used to inoculate 500 mL of 2xYT medium supplemented with ampicilin (100 μg/mL). Cells were grown at 37℃ in an incubator (250 r.p.m.) and induced with 800 μM IPTG after OD_{600} reached 0.4 ~ 0.6. After 4 hour induction, cells were harvested by being centrifuged at 4000 g for 15 min, and resuspended in a lysis buffer (50 mM NaH_2PO_4, 250 mM NaCl, 10 mM imidazole, pH 8.0. The cell lysate was sonicated in an ice bath six times (2 min each pulse, 5 min interval for cooling), and centrifuged at 1000g for 60 min (4 ℃). The supernatant was isolated and incubated with 1 mL Ni-NTA resin (Qiagen) (1.5 h, 4 ℃), and then loaded to a column. The protein-resin mixture was washed with 50 mL of a wash buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 10 mM imidazole (pH 8.0), and eluted by a elution buffer containing 50 mM NaH_2PO_4, 250 mM NaCl and 250 mM imidazole (pH 8.0). The purified protein was concentrated and dialyzed into a buffer containing 10 mM ammonium bicarbonate. The protein was analyzed by 15% SDS-PAGE and stored at -80℃.

4. Construction of Phage Library

DNA sequencing of pADLg3-TGC-(NNK)6-TAG phagemid

gcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccct
gataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgcct
tcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactgg
atctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgc

ggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccag

tcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca

acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgtt

gggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgca

aactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccactt

ctgcgctcggcgcttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcact

ggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacaga

tcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcattt

ttaatttaaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga

ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctacc

agcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactg

ttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagt

ggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct

gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaga

aagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagg

gagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgt

caggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgac

ccgacaccatcgaatggcgcaaaaccttcgcggtatggcatgatagcgcccggaagagagtcaattcagggtggtgaatgtg

aaaccagtaacgttatacgatgtcgcagagtatgccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacg

tttctgcgaaaacgcgggaaaaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggc

gggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccgtcgcaaattgtcgcggcgattaaatct

cgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcac

aatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcctgcac

taatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacggtacgcgactgggc

gtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctgg

ctggctggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtccggttttca

acaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctgggcgcaatgc

gcgccattaccgagtccgggctgcgcgttggtgcggacatctcggtagtgggatacgacgataccgaagacagctcatgttata

tcccgccgttaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggcc

aggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaaaaccaccctggcgcccaatacgcaaaccgcctct

ccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcggtacccgataa

aagcggcttcctgacaggaggccgttttgttttgcagcccacctcaacgcaattaatgtgagttagctcactcattaggcacccca

ggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgat

tacgaatttctagataacgagggcaaatcatgaaatacctattgcctacggcggccgctggattgttattactcgcggcccagccg

gccatggcctcgnnknnknnknnknnknnktaggcccgggaggccaaggcggtggttctgagggtggtggctccctcg

agggcgcgccagccgaaactgttgaaagttgtttagcaaaacctcatacagaaaattcatttactaacgtctggaaagacgacaa

aactttagatcgttacgctaactatgagggctgtctgtggaatgctacaggcgttgtggtttgtactggtgacgaaactcagtgttac

ggtacatgggttcctattgggcttgctatccctgaaaatgagggtggtggctctgagggtggcggttctgagggtggcggttctga

gggtggcggtactaaacctcctgagtacggtgatacacctattccgggctatacttatatcaaccctctcgacggcacttatccgc

ctggtactgagcaaaaccccgctaatcctaatccttctcttgaggagtctcagcctcttaatactttcatgtttcagaataataggttcc

gaaataggcagggtgcattaactgtttatacgggcactgttactcaaggcactgaccccgttaaaacttattaccagtacactcctg

tatcatcaaaagccatgtatgacgcttactggaacggtaaattcagagactgcgctttccattctggctttaatgaggatccattcgtt

tgtgaatatcaaggccaatcgtctgacctgcctcaacctcctgtcaatgctggcggcggctctggtggtggttctggtggcggctc

tgagggtggcggctctgagggtggcggttctgagggtggcggctctgagggtggcggttccggtggcggctccggttccggt

gattttgattatgaaaaaatggcaaacgctaataagggggctatgaccgaaaatgccgatgaaaacgcgctacagtctgacgcta

aaggcaaacttgattctgtcgctactgattacggtgctgctatcgatggtttcattggtgacgtttccggccttgctaatggtaatggt

gctactggtgattttgctggctctaattcccaaatggctcaagtcggtgacggtgataattcacctttaatgaataatttccgtcaatat

ttaccttctttgcctcagtcggttgaatgtcgcccttatgtctttggcgctggtaaaccatatgaattttctattgattgtgacaaaataaa

cttattccgtggtgtctttgcgtttcttttatatgttgccacctttatgtatgtattttcgacgtttgctaacatactgcgtaataaggagtct

taatcaagctttaatattttgttaaaattcgcgttaaatttttgttaaatcagctcatttttttaaccaataggccgaaatcggcaaaatccc

ttataaatcaaaagaatagaccgagatagggttgagtgttgttccagtttggaacaagagtccactattaaagaacgtggactcca

acgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaaccatcaccctaatcaagttttttggggtcgaggt

gccgtaaagcactaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcgaacgtggcgagaa

aggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaaccaccacacc

cgccgcgcttaatgcgccgctacagggcgcgtcaggtg (mutated sites are highlighted in red. n presents any of a, g, c, t, and k presents g or t)

Construction of Phagemid Library

The phagemid library was constructed and amplified by Polymerase chain reaction (PCR) using forward primer: 5'- GGT CCG ATG GCC NNK NNK NNK NNK NNK NNK GGC CCG GG -3', and reverse primer: 5'- CCA CGG CCA TGG CCG GCT GGG CCG CG -3'. The PCR product was digested by NCoI restriction enzyme, and ligated by T4 DNA ligase. The ligated plasmids were electroporated into competent *Ecoli.* Top10 cell, incubated with 1 mL of LB medium and then inoculated into 5 mL of LB medium containing 100 μg/mL ampicillin. After OD_{600} reached 1.0, 0.5 mL of cell culture was mixed with with 50% glycerol and stored in -80 °C. Several of aliquots are made to for the total coverage of more than 10^{11} cfu. To collect the phagemids, normalized amount of cell stocks was used to guarantee equal amount of phagemids from each aliquot.

Phage Expression

Q350 at M13KO7 was mutated to stop codon TAA to knock out expression of pIII protein from helper phage. Phagemid library constructed from previous step was electro-porated into *Ecoli.* Top10 electro competent cells with M13KO7g3TAA and pEVOL-CloDF-AcrkRS. Transformed cells were inoculated into 100 mL of 2YT medium. Fusion protein expression was induced with 0.2% arabinose, 0.5 mM IPTG, 2 mM AcrK when OD_{600} reached 0.5. After 12-hour induction, cell pellets were precipitated by centrifuge and discarded. The supernatants containing phages were precipitated by chilled polyethylene glycol, and then subjected to centrifuge (15 min, 10,000 g, 4 ℃). Phage pellets were collected and dissolved in PBS buffer. The total number of phage was calculated as following: 10 uL of phages was incubated at 65 ºC water bath for 15 min to kill all *Ecoli.* Top10 cell. 90 uL of Top10F' (OD_{600} = 1.0) are infected by these phages and subjected to the agar plate with serial dilution for calculation. The total yield is around 10^{10} cfu per 100 mL of LB medium, sufficient to cover library diversity (theoretical diversity of a 6-mer library is $20 = 6.4 \times 10^6$)

5. Phage Selection Against TEV Protease and HDAC8

For selection, streptavidin magnetic beads that only bind to biotin for selection were used. To generate biotinylated protein in aqueous solution, biotin sulfosuccinimidyl ester kit (thermos fisher scientific) was used since a succinimidyl ester moiety reacts with primary amine group (**Figure 22**). A 15μM purified target protein was incubated with 30μM biotin succinimidyl ester in 50 mM phosphate buffer for 2 hours at room temperature. The reaction was quenched with addition of 10 mM lysine and subjected to

the protein purification kit (Bio-RAD). 5 uM purified biotinylated protein are incubated with streptavidin magnetic beads (Pierce) for 1 hour in PBS buffer. Unreacted protein was washed away.

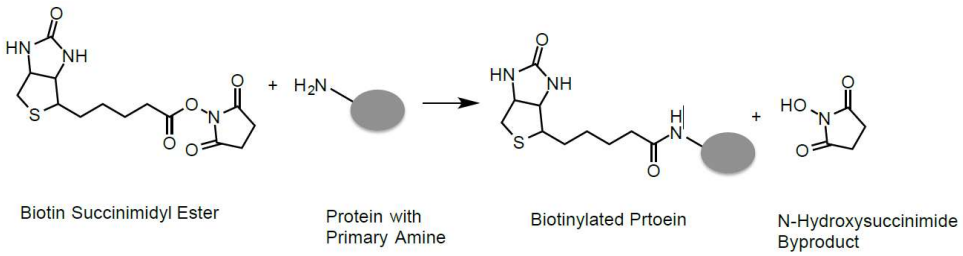


Figure 22 Reaction between biotin succinimidyl ester and protein with primary amine

In selection process, to remove individuals capable of non-specific binding, phage library was pre-incubated with only streptavidin magnetic beads for every round of selection. Later, phage library was incubated with protein-binding streptavidin magnetic beads for 10 min. The beads were washed 10 times by phosphate buffered saline with Tween-20 (PBST, 8mM Na_2HPO_4, 150mM NaCl, 3 mM KCl, 2mM KH_2PO_4, 0.05% Tween-20, pH 7.4), eluted with Glycine-HCl buffer (pH 2.2), and neutralized with Tris buffer (pH 9.1). Top10F' cells were infected by the eluted phage to calculated the number of phage particles. To amplify the selected phage library, the phagemid from infected Top10F' cells were extracted. Cell transformation, phage expression and phage selection were repeated for consecutive three rounds. For better comparison, controls of background binding phages with streptavidin magnetic beads were also studied in each round of selection.

Figure 23 illustrates the number of eluted phage for every round, of TEV protease and HDAC8, respectively. A dramatic increase of eluted phages was observed after three rounds, for both TEV protease and HDACs, while only a slightly enrichment of controls was observed.

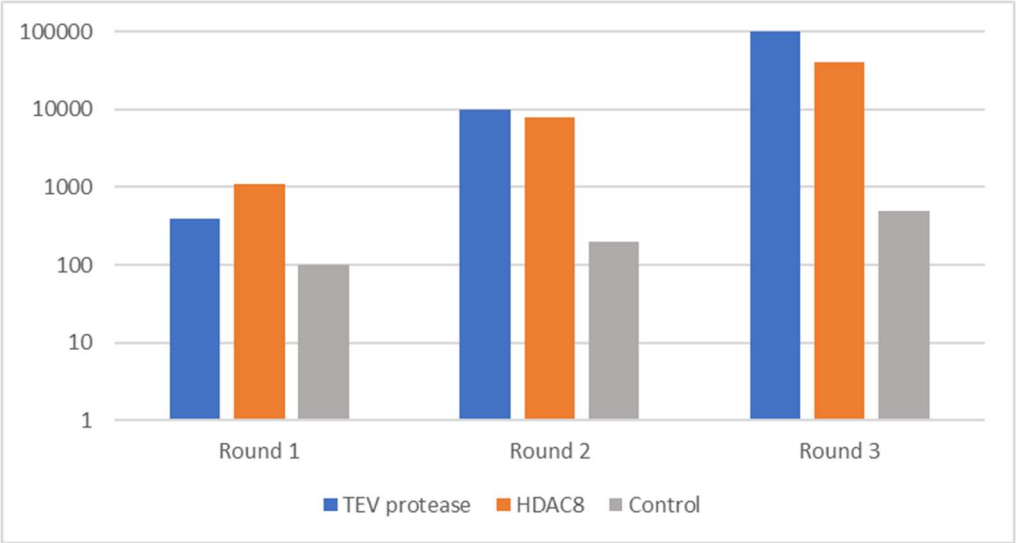


Figure 23 Number of eluted phages for TEV protease and HDAC8 in each selection round. Number of input phages is 10^9 in each round.

6. Synthesis of Selected Peptide CX_6-AcrK (X designates any canonical amino acid)

General information

With advantages such as high yield, ease of handling and simple isolation steps, solid phage peptide synthesis (SPPS) strategy has been demonstrated as the best alternative to synthesize peptides.[79,82] In this strategy, the peptide sequences are synthesized from C-terminal to N-terminal (**Figure 24**). Special resins are used as solid phase to couple amino acids one by one. The amino acids with protected side chains are coupled to the α-amino group. In this dissertation, Fmoc-protected amino acids were used in synthesis of selected. After the removal of Fmoc protecting group from N-terminal and activation of C-terminal, coupling of each amino acid is usually finished in a few minutes to a few hours. Unreacted reagents and byproducts were then washed away by DMF and dichloromethane. Finally, peptides were cleaved from resin by 95% TFA, precipitated by cold ether and subjected to further characterization. To monitor the coupling process, Kaiser test is the most commonly used approach. Similar to ninhydrin test, only primary amines turn blue when they reacts with Kaiser test kit (**Figure 25**).[83,84]

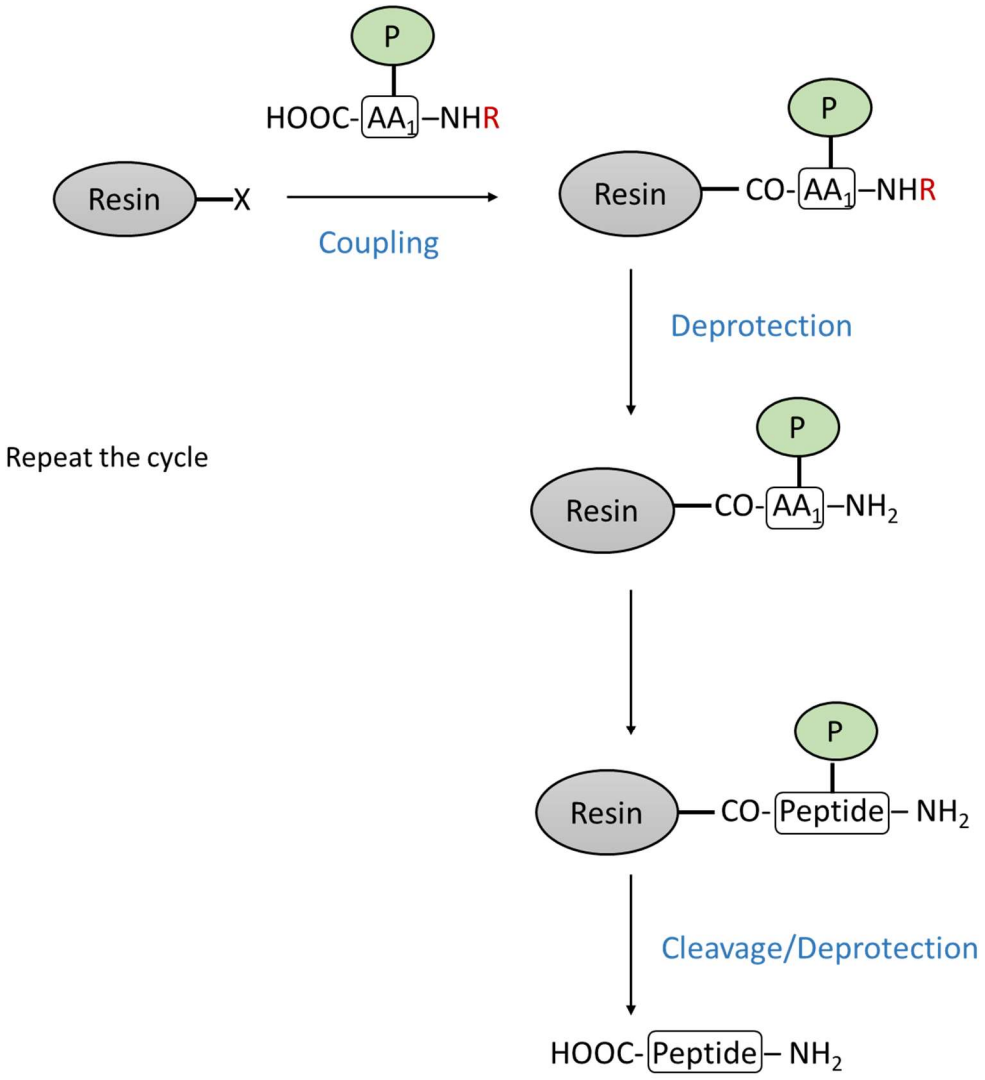


Figure 24 General procedure of solid phage peptide synthesis (SPPS)

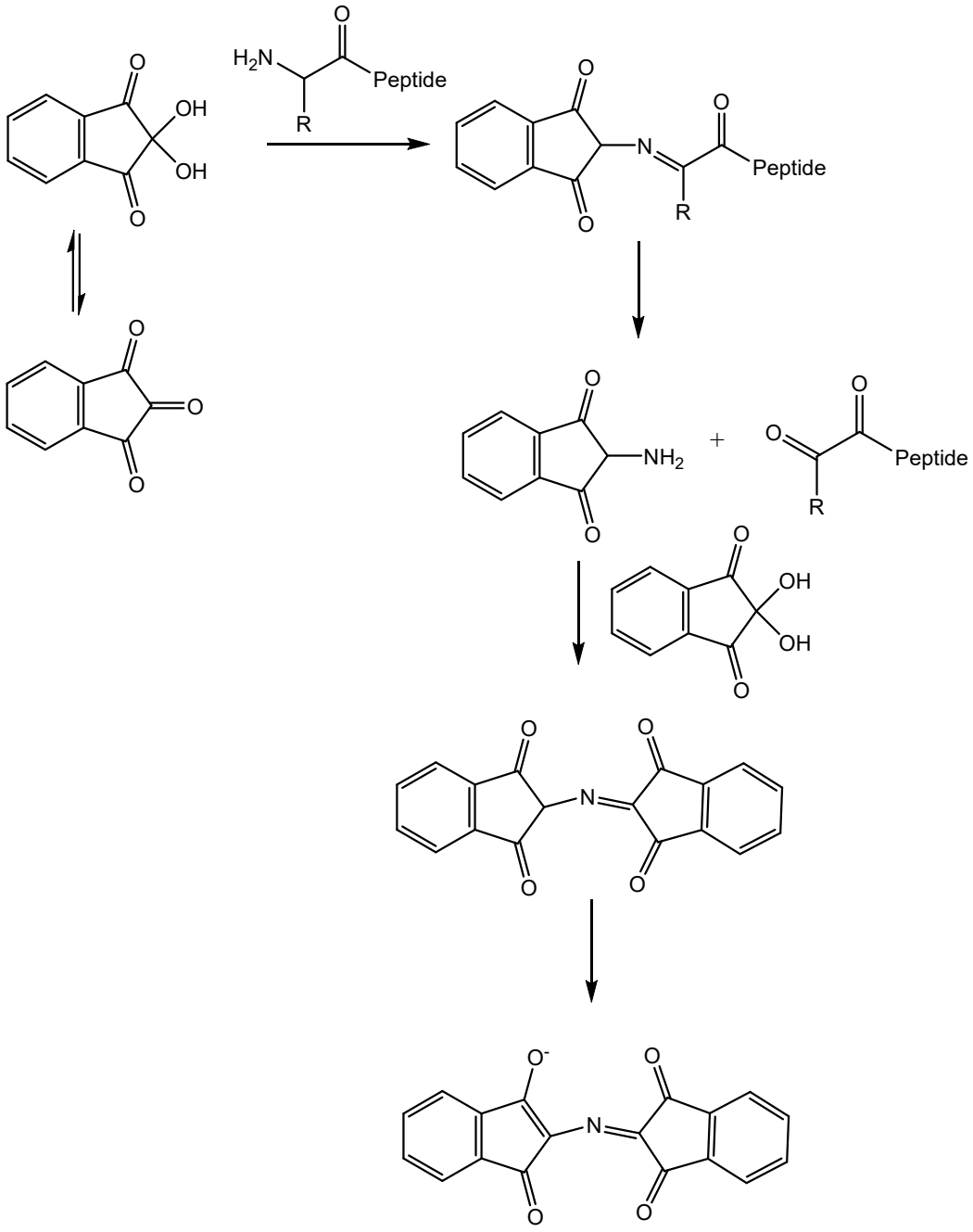


Figure 25 Reaction of Kaiser test.

Couple the first lysine to the resin

200 mg rink amide MBHA resin (Novabiochem) in DMF was added to a poly vessel, swelled for 1 h. Fmoc group of the resin was then deprotected by 20% (vol/vol) piperidine in DMF for 30 minutes, and then washed with DMF, dichloromethane (DCM) and methanol. Fmoc-Lys(mtt)-OH (4 equiv), tetramethyluronium hexafluorophosphate (HBTU, 4 equiv) and diisopropyl-ethylamine (DIEA, 10 equiv) were dissolved in DMF. The solution was added to the reaction vessel under nitrogen and mixed with the resin. Coupling was not finished until Kaiser-ninhydrin test became negative.

Couple the remaining amino acids to the resin

Fmoc-protected amino acids (4 equiv), tetramethyluronium hexafluorophosphate (HBTU, 4 equiv) and diisopropyl-ethylamine (DIEA, 10 equiv) were dissolved in DMF (10 mL). The solution was added to the reaction vessel under nitrogen and mixed with the resin. Reaction was not stopped until Kaiser-ninhydrin test became negative. The last amino acid we used is Boc-Cys(trt)-OH. There is no additional deprotection steps after the final coupling.

Synthesis of N-succinimidyl acrylate[6]

To a solution of N-hydroxysuccinimide (1.3 g, 11.3 mmol) in anhydrous dichloromethane (25 mL) was added N,N-diisopropylethylamine (1.5 mL, 8.9 mmol), followed by dropwise addition of acryloyl chloride (0.8 mL, 9.3 mmol) in ice bath in 10 min. The mixture was then stirred for 10 hours at room temperature. The mixture was extracted with 50 mL of ethyl acetate, washed with saturated NH_4Cl solution (x3) and

brine, and dried with anhydrous Na_2SO_4. The solution was filtered and evaporated under vacuum to give a yellow oil (1.5 g).

Couple N-succinimidyl acrylate to the first lysine

After the coupling of cysteine, use % TFA in dichloromethane (vol/vol) and 1% triisopropylsilane (TIS) in dichloromethane (vol/vol) to remove -mtt group. N-Succinimidyl acrylate (2 equiv) and DIEA (5 equiv) in DMF were added to the resin and coupled until Kaiser test became negative.

Cleavage of the peptide from the resin

4 mL of cleavage solution containing 92.5% trifluoroacetic acid (TFA), 2.5% TIS, 2.5% water, 2.5% 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) was incubated with 200 mg resin for 2~3 hours. The peptide products were precipitated with 10 volume cold diethyl ether. The cleaved peptides were collected by centrifuge, washed with cold diethyl ether and purified by HPLC. The purified products were lyophilized and subjected to MALDI-TOF analysis.

Preparation of cyclic peptide

Purified peptide was dissolved in PBS buffer and incubated at room temperature for 4 hours, and subjected to HPCL for purification. Eluted peptide solution was lyophilized to give a white powder. Compared to linear peptides, ^{1}H NMR shows the appearance of δ 5.89 (m, 1H), proving the success of cyclization between side chain of cysteine and the acryloyl moiety of AcrK.

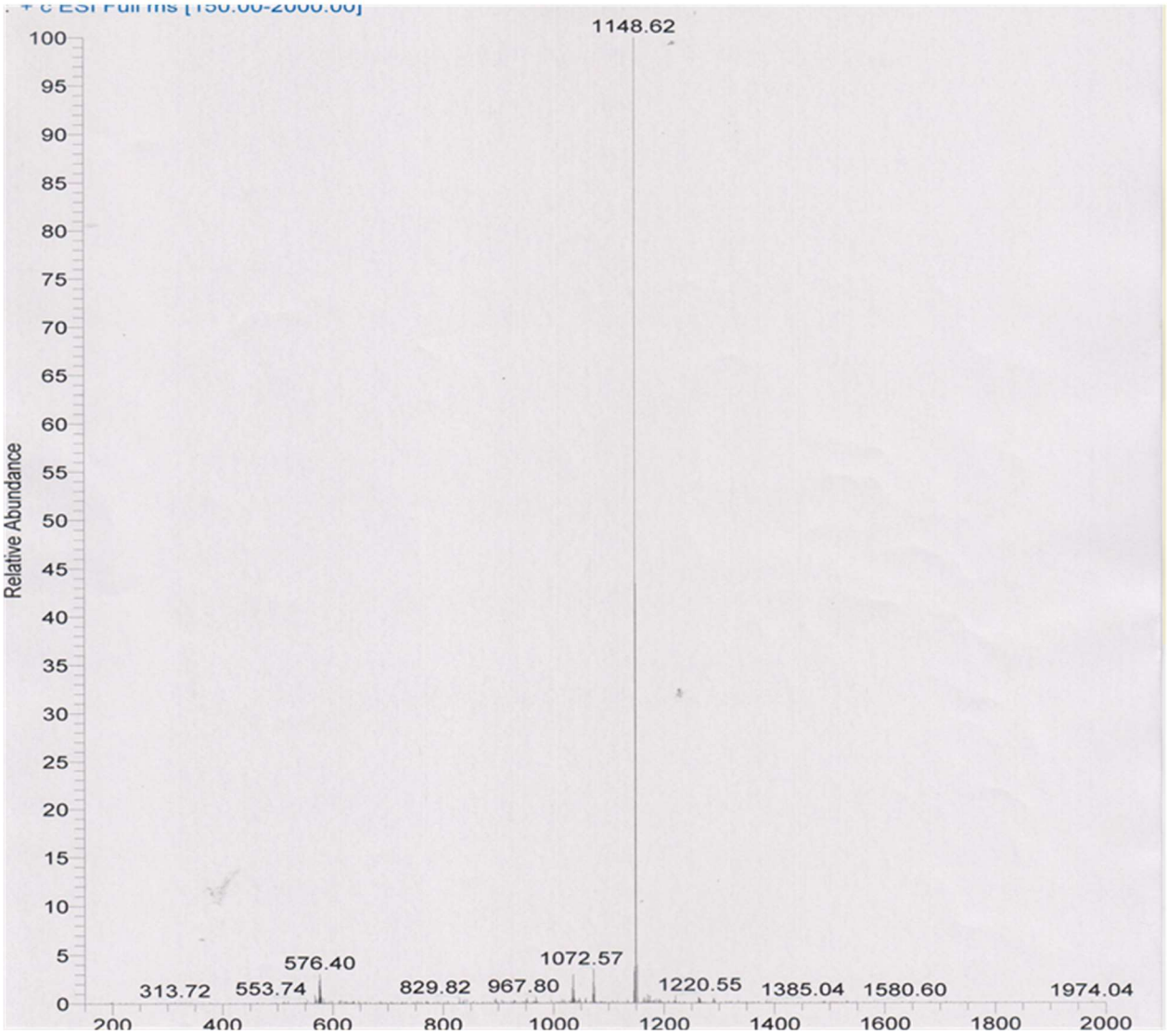


Figure 26 MALDI-TOF spectrum of CQWFSHR-AcrK, M.W.:1144.02 g/mol

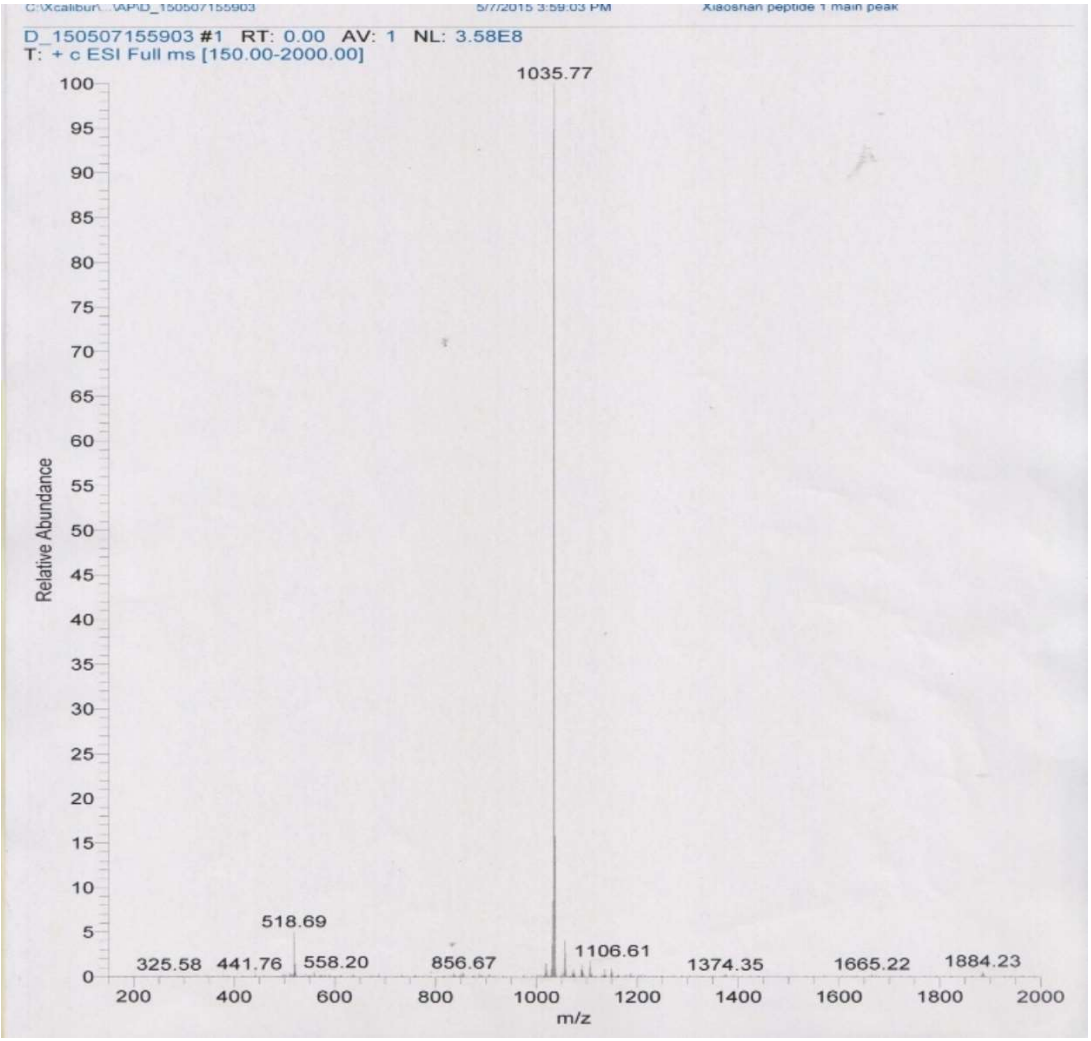


Figure 27 MALDI-TOF spectrum of CGTWLKF-AcrK, M.W.:1035.28 g/mol

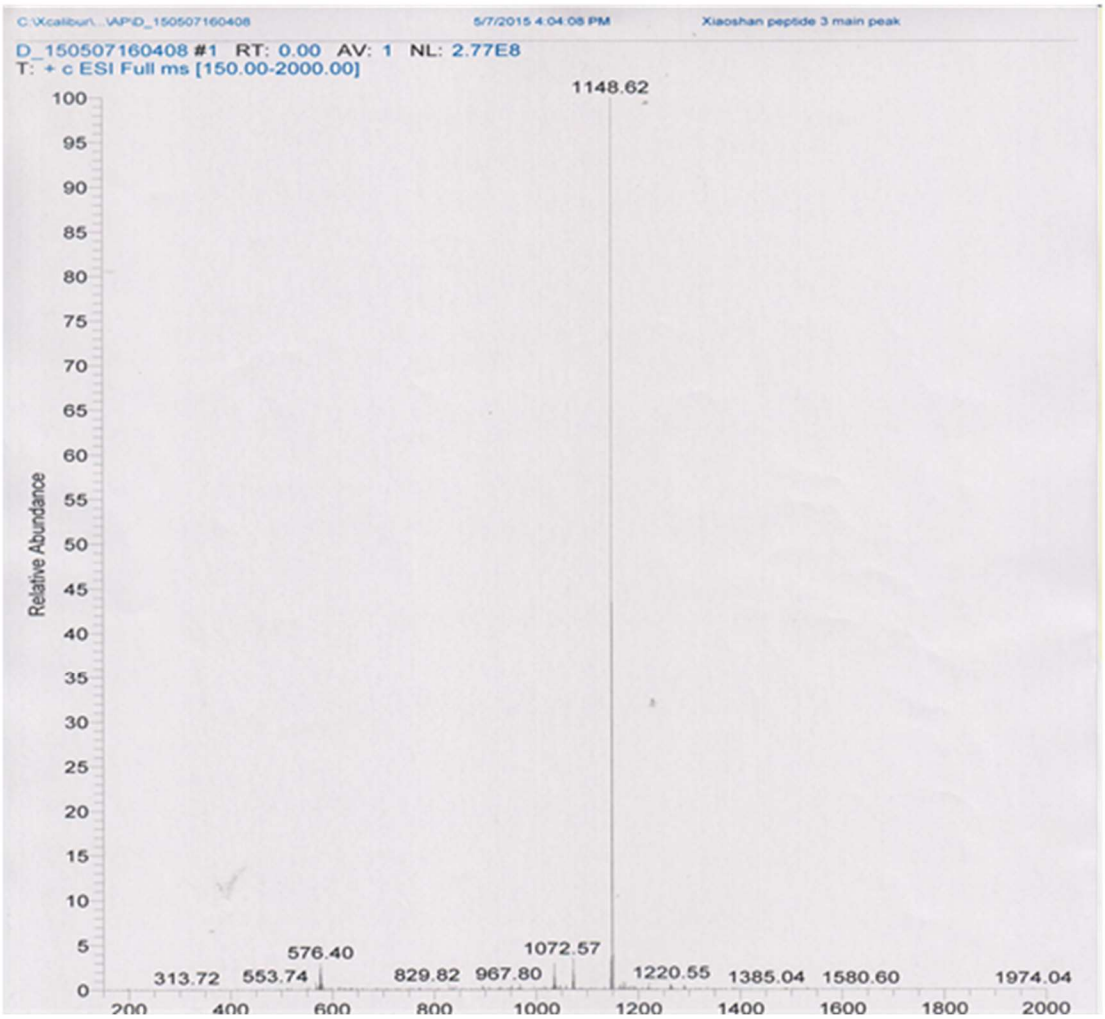


Figure 28 MALDI-TOF spectrum of CWRDYLI-AcrK, M.W.:1149.38 g/mol

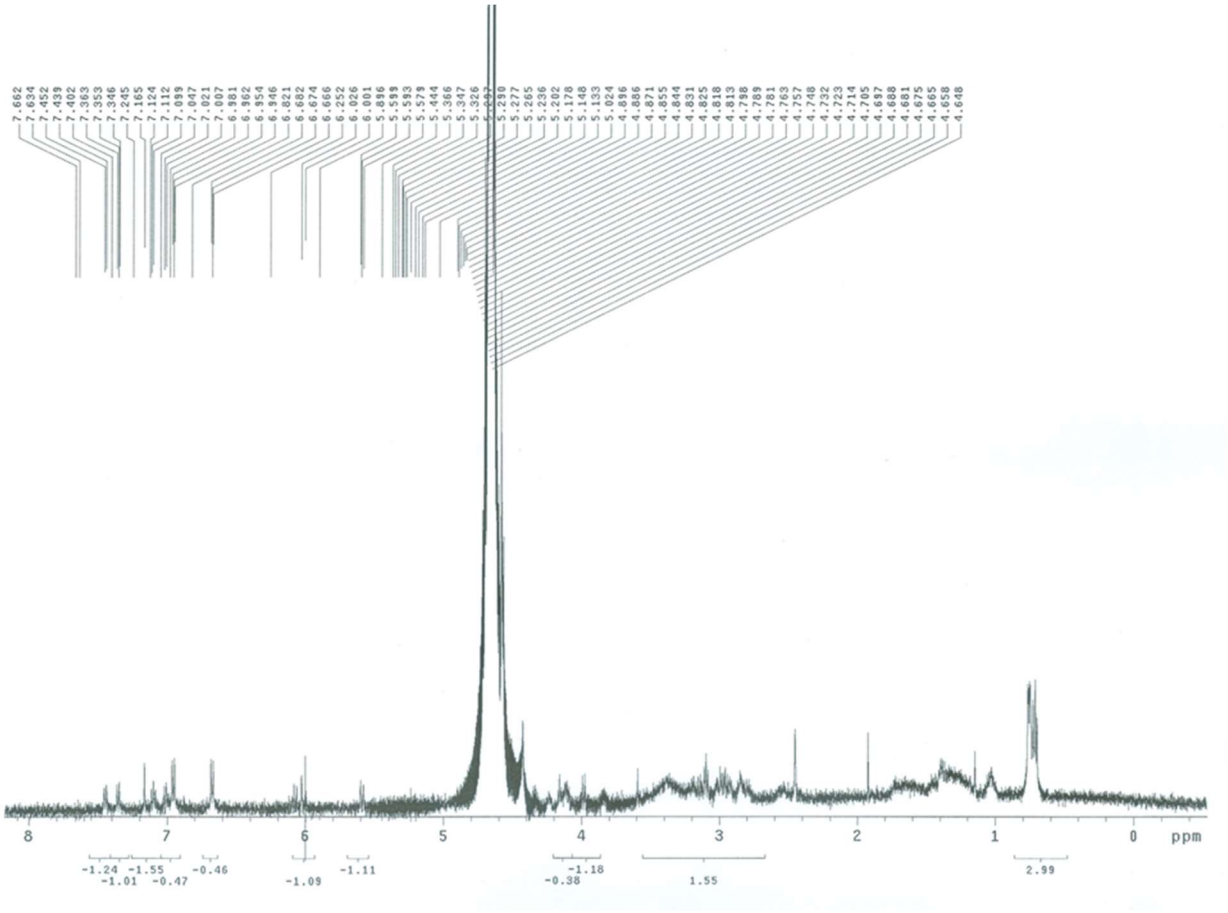


Figure 29 ^{1}H NMR spectrum for cyclo(CWRDYLI-AcrK)

7. Fluorescence Polarization Measurement

Synthesis of 5-carboxyfluorescein (FAM) -conjugated peptide (CX_6-AcrK-K(5-FAM), X designates any amino acid)

Couple the first lysine to the resin: 200 mg rink amide MBHA resin (Novabiochem) in DMF was added to a poly vessel, swelled for 1 h. Fmoc group of the resin was then deprotected by 20% (vol/vol) piperidine in Dimethylformamide (DMF) for 30 minutes, and then washed with DMF, dichloromethane (DCM) and methanol. Fmoc-Lys(mtt)-OH (4 equiv), tetramethyluronium hexafluorophosphate (HBTU, 4 equiv) and diisopropyl-ethylamine (DIEA, 10 equiv) were dissolved in DMF. The solution was added to the reaction vessel under nitrogen and mixed with the resin. Coupling was not finished until Kaiser-ninhydrin test became negative.

Couple 5-Carboxyfluorescein to lysine: -Mtt protecting group from the first lysine was removed by repeated washing of 1% TFA in dichloromethane (vol/vol) and 1% triisopropylsilane (TIS) in dichloromethane (vol/vol). Couple 5-Carboxyfluorescein (2 equiv) and DIEA (5 equiv) in DMF were added to the resin and coupled until Kaiser test became negative.

Couple the remaining amino acids to the resin: Fmoc-protected amino acids (4 equiv), tetramethyluronium hexafluorophosphate (HBTU, 4 equiv) and diisopropyl-ethylamine (DIEA, 10 equiv) were dissolved in DMF. The solution was added to the reaction vessel under nitrogen and mixed with the resin. Reaction was not finished until Kaiser-ninhydrin test became negative. The last amino acid we used is Boc-Cys(trt)-OH. There is no additional deprotection steps after the final coupling.

Couple N-Succinimidyl acrylate to the second lysine: after the coupling of cysteine, use % TFA in dichloromethane (vol/vol) and 1% triisopropylsilane (TIS) in dichloromethane (vol/vol) to remove -mtt group from the second lysine. N-Succinimidyl acrylate (2 equiv) and DIEA (5 equiv) in DMF were added to the resin and coupled until Kaiser test became negative.

Cleavage of the peptide from the resin: 4 mL of cleavage solution containing 92.5% trifluoroacetic acid (TFA), 2.5% TIS, 2.5% water, 2.5% 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) was incubated with 200 mg resin for 2 hours. The peptide products were precipitated with 10 volume cold diethyl ether. The cleaved peptides were collected by centrifuge, washed with cold diethyl ether and subjected to HPCL for purification. The purified products were lyophilized and subjected to MALDI-TOF analysis.

Preparetion of cyclic peptide: Purified peptide was dissolved in PBS buffer and incubated at room temperature for 4 hours, and subjected to HPCL for purification. Eluted peptide solution was lyophilized to give a white powder.

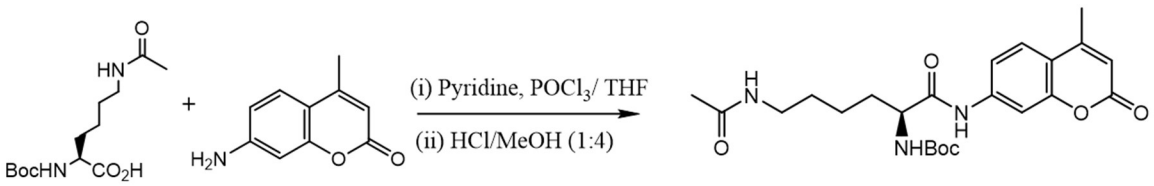
Fluorescence polarization measurement

25 nM 5-FAM-conjugated cyclic peptide and different concertation of target proteins (160 nM to 160 μM) were incubated at black 96-well plates. PBS buffer were added to make the final volume 200 μL in each well. Fluorescence polarization was measured by a microplate reader at Ex/Em = 490 nm / 520 nm.

8. I_{50} Value Measurement

Synthesis of Boc-lys(Ac)-AMC [85]

Boc-lys(Ac)-OH (2.0 mmol, 576.7 mg) and 7-Amino-4-methylcoumarin (2.0 mmol,350.4 mg) were dissolved in ice-chilled anhydrous THF (50 mL). Pyridine (20.0 mmol, 1.6 mL) was added to the solution dropwise, followed by the addition of phosphoryl chloride (8.4 mmol, 0.8 mL). The mixture was stirred in ice-water bath for 3 hours. Later, the reaction was quenched by the addition of saturated sodium bicarbonate solution (50 mL). The mixture was concentrated to 50 mL under reduced pressure. The mixture was extracted with 25 mL dichloromethane for three times, and washed with 25 mL saturated NaCl solution and 0.5 M HCl solution (4 x 50 mL). The combined dichloromethane extraction was dried over anhydrous $MgSO_4$, concentrated under reduced pressure, and dissolved in HCl/MeOH (1:4 v/v). The solution was stirred at room temperature for 24 hours and concentrated under reduced pressure, affording desired product (489.1 mg, 55% for two steps) as a yellow powder.



Boc-Lys(Ac)-AMC

Figure 30 Synthesis of Boc-Lys(Ac)-AMC

IC_{50} value measurement

Peptide CQSLWMN-AcrK was pre-incubated at PBS buffer for 4 hours for cyclization, generating cyclo(CQSLWMN-AcrK). Different concentrations (1 nM – 1000 μM) of cyclo(CQSLWMN-AcrK) and 5 μM HDAC8 were added to a black 96-well plate (Pierce). PBS buffer was added to make the final volume in each well 200 μL. The plate was incubated at 30℃ for 10 min. 50 μM Boc-Lys(Ac)-AMC was added to each well. After 1-hour incubation at 30℃, the HDAC-catalyzed deacetylation was terminated by addition of trichostatin A (TSA, 1 μM), followed by addition of trypsin (0.5 mg/mL) to the reaction solution. After additional 1-hour incubation at 30℃, the fluorescence of coumarin was measured by a microplate reader with Ex/Em = 360 nm/ 460 nm. Experiments of IC_{50} measurements were repeated three times.

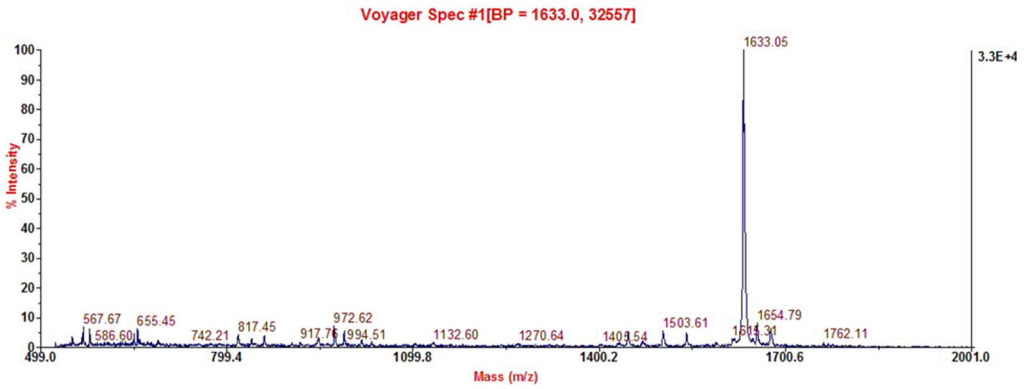


Figure 31 MALDI-TOF spectrum of CWRDYLI-AcrK-K(5-FAM), M.W.: 1636 g/mol

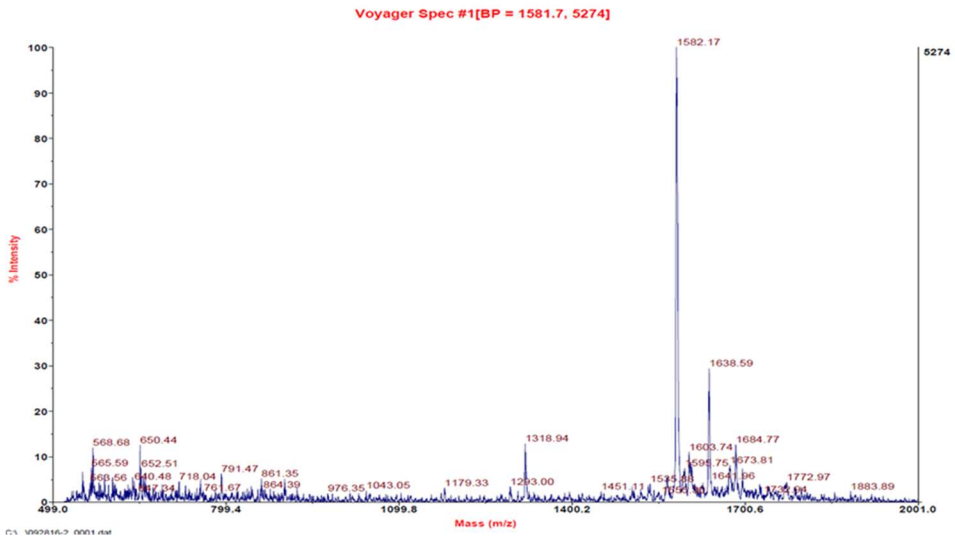


Figure 32 MALDI-TOF spectrum of CWRDYLIKK(5-FAM), M.W.: 1581 g/mol

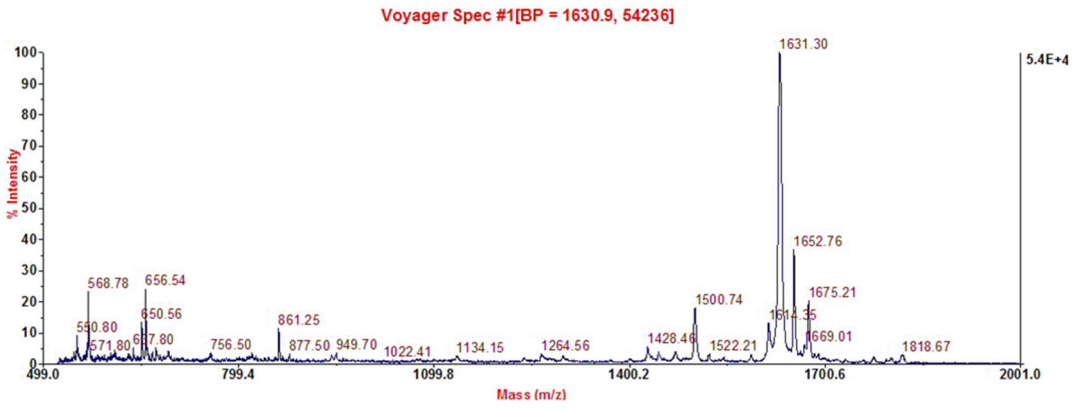


Figure 33 MALDI-TOF spectrum of CWRDYLI-AcrK-K(5-FAM), M.W.: 1631 g/mol

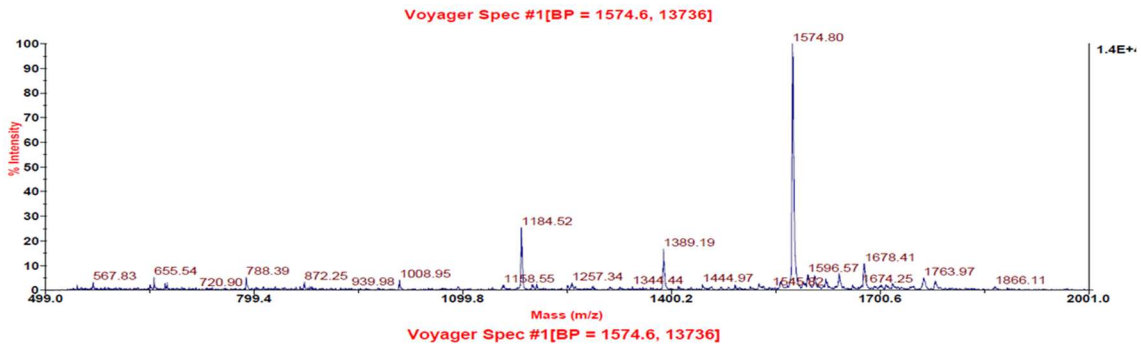


Figure 34 MALDI-TOF spectrum of CWRDYLIK-K(5-FAM), M.W.: 1575 g/mol

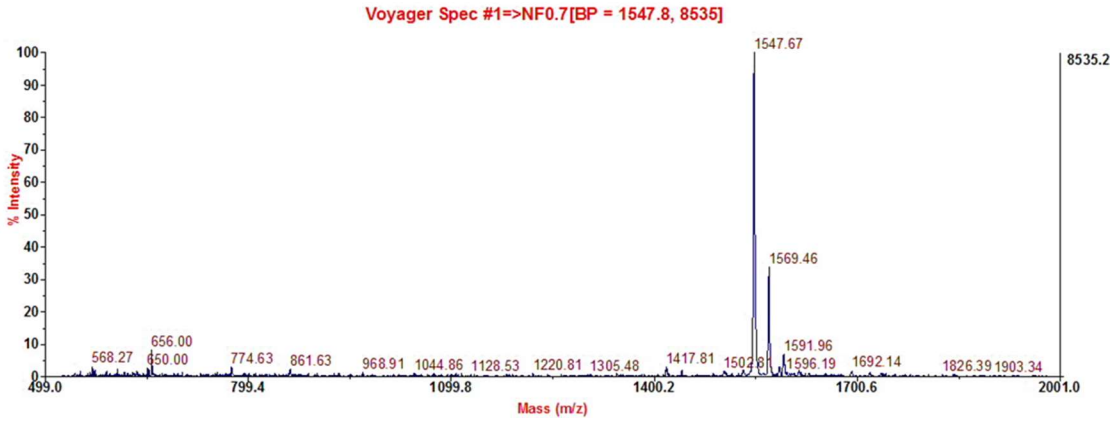


Figure 35 MALDI-TOF spectrum of CQSLWMN-AcrK-K(5-FAM), M.W.: 1549 g/mol

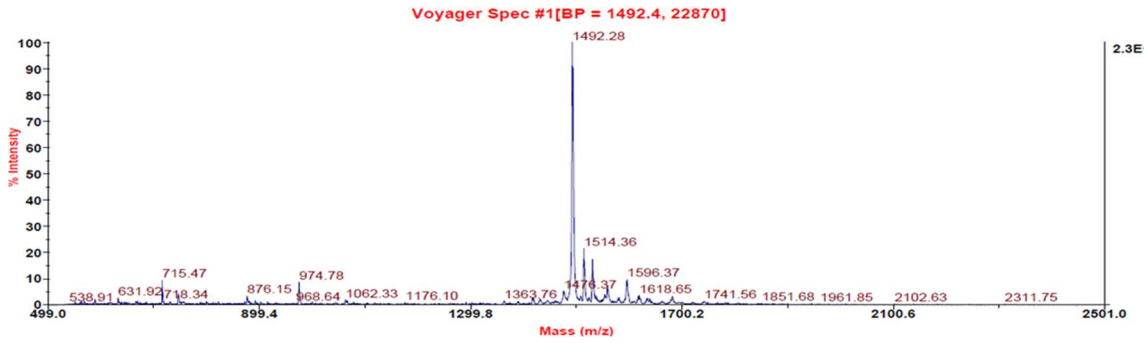


Figure 36 MALDI-TOF spectrum of CQSLWMNK-K(5-FAM), M.W.: 1493 g/mol

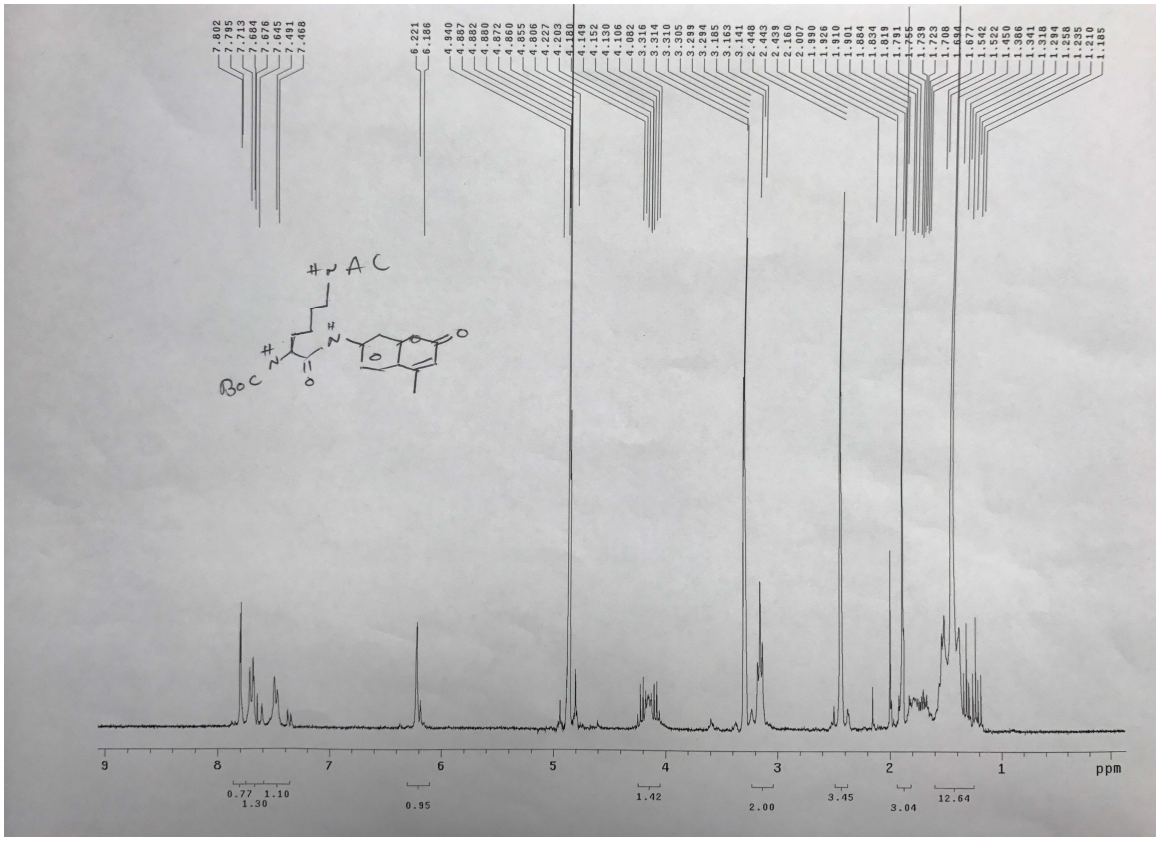


Figure 37 ^{1}H NMR spectrum for Boc-Lys(Ac)-AMC

Results and Discussion

1. Characterization of 1, 4-cycloaddition in sfGFP

Before the construction of peptide libraries, we firstly tested 1, 4-cycloaddition between cysteine and the acrylamide moiety in model protein Met-Cys-$(Ala)_5$-AcrK-sfGFP and Met-Cys-$(Ala)_5$-CrtK-sfGFP, respectively. N^{ε}-acryloyl-L-lysine (AcrK) and N^{ε}-crotonyl-L-lysine (CrtK) (**Figure 38**), two NCAAs both bearing an acrylamide moiety, were incorporated at amber TAG position by two mutants of PylRS, PrKRS (Y384W MmAcKRS1 mutant) and BuKRS (Y384W PylRS mutant), respectively. For better comparison, another two sfGFP mutants without cysteine, Met-$(Ala)_6$-AcrK-sfGFP and Met-$(Ala)_6$-CrtK-sfGFP, were generated in the same way and used as linear controls. Theoretically, these two linear controls have no free cysteine and cannot undergo cyclization.

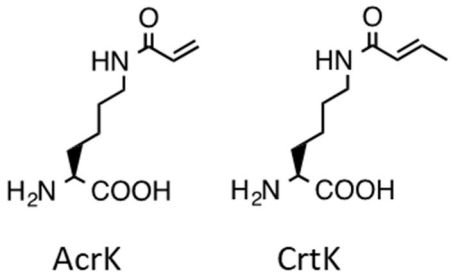


Figure 38 Structure of AcrK and CrtK.

To test the success of cyclization, hydrozonoyl chloride **1** acts as a fluorogenic probe. If there is any unreacted acrylamide moiety in Met-Cys-$(Ala)_5$-AcrK-sfGFP or Met-Cys -$(Ala)_5$-CrtK-sfGFP, the remaining acrylamide moiety undergoes a fluorescence turn-on nitrile imine cycloaddition a with hydrozonoyl chloride **1**, leading to a fluorescent cycloaddition product from those non-fluorescence starting materials (**Figure 39A**). Thereby monitoring fluorescence is an approach to detect the cyclization process. Details of this nitrile imine cycloaddition have been demonstrated in Chapter II. The best condition for the reaction is in acetonitrile-50 mM phosphate buffer (vol/vol = 1:1) at pH 10. The reaction is very efficient and can be finished in one minute. In this way, sfGFP mutants bearing AcrK and CrtK were both incubated with hydrozonoyl chloride in acetonitrile-50 mM phosphate buffer (1:1), pH 10 for 10 min, and then subjected to SDS page analysis (**Figure 39B**). Both Met-Cys-$(Ala)_5$-AcrK-sfGFP(a) and Met-Cys-$(Ala)_5$-CrtK-sfGFP (d) show no fluorescence, verifying that the cyclization through 1,4-addition is already completed during sfGFP expression and purification steps. In contrast, both of the two sfGFP mutants without cysteine show fluorescence. Additionally, the linear sfGFP mutants incorporated with CrtK shows weaker fluorescence than the mutants with AcrK, demonstrating the fact that CrtK was less reactive that AcrK. This is because the methyl group in CrtK residue hinders the attack from cysteine thiol due to steric effect, and thus the reaction rate of CrtK with hydrozonoyl chloride is slower than that of AcrK.

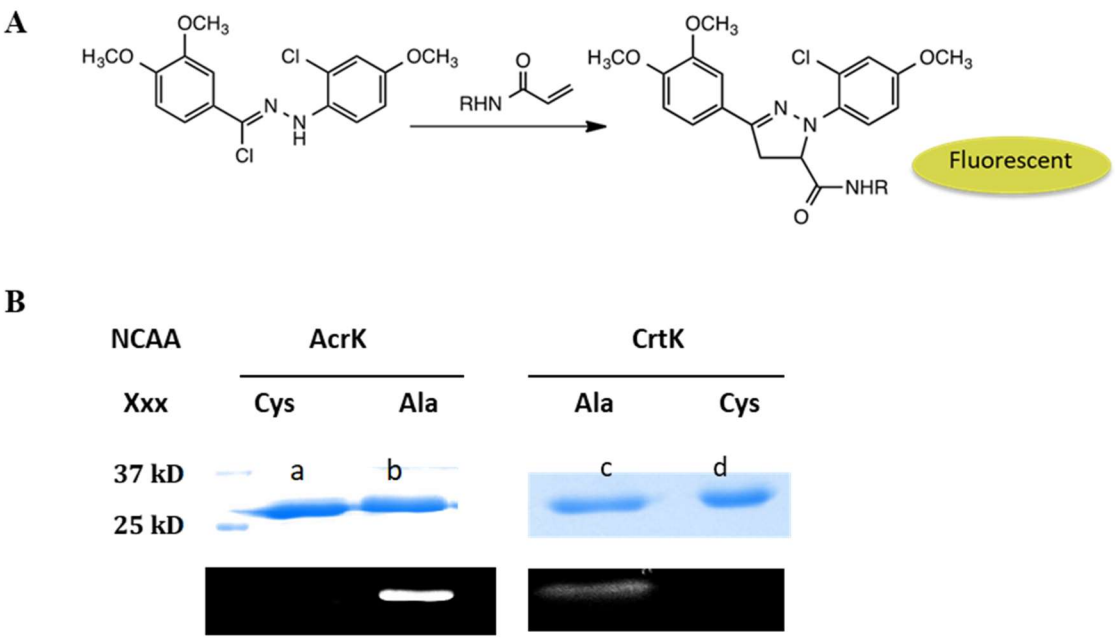


Figure 39 (A) Click reaction between hydrozonoyl chloride and alkene; **(B)** SDS-Page analysis of a. Met-Cys-$(Ala)_5$-AcrK-sfGFP and control b. Met-$(Ala)_6$-AcrK-sfGFP, control c.Met-$(Ala)_6$-CrtK-sfGFP and d. Met-Cys-$(Ala)_5$-CrtK-sfGFP after reaction with hydrozonoyl chloride in acetonitrile-50 mM phosphate buffer (1:1) with pH 10. The top gels were stained by Coomassie blue; the bottom are the same gels but visualized by fluorescence before stained by Coomassie blue.

2. Construction of Phage-displayed Cyclic Peptide Library

Encouraged by the success of cyclization reactions on model proteins, we extended this approach to the construction of phage-displayed cyclic peptide library bearing a conserved cysteine and AcrK. To incorporate NCAAs into phage library, our lab has developed a strategy based on the amber stop codon suppression. In our methodology, pADLg3 phagemid was used to produce full-length pIII coat protein, where the foreign peptide incorporated with an NCAA is fused to. Compared to the other phage vectors, pADL phagemids have notable advantages such as low contamination risk, PelB secretion and single pDNA package. A helper phage lacking pIII is required to pack phagemid, and guarantee that the only pIII derived from pADL phagemid. After transformation with cells (e.g. *Escherichia coli*) and expression, single-strand phagemid DNA is infected by helper phages and efficiently packed into viral particles. Then the packed fusion proteins will be released from host cells.

For the genetic construction, pADL library, a TAG-enriched M13 phagemid library of pIII fusion proteins, was constructed by flanking diversity of 6 residue positions between a cysteine codon and the amber stop codon, each residue position randomizing 20 canonical amino acids. Polymerase chain reaction (PCR) enables the insertion of six randomized codons (NNKs) between pelB and pIII in the pADL phagemid. The amino acid sequence is: (PelB)-Ala-Cys-$(Xxx)_6$-TAG-(pIII) (Xxx designates a randomized canonical amino acid). More than 10^9 clones were obtained to guarantee the coverage for the possibility of all 20 amino acids at each randomized position ($20^6 = 6.4x10^7$).

Next, we investigated the expression of phage library. Given that two plasmids with a same replication origin are incompatible during expression, CloDF origin was used

instead of p15A in synthetase plasmid to avoid the conflict, since the helper phage M13K07 has a p15A origin. To incorporate AcrK into the phage library, synthetase plasmid pEVOL-CloDF-PrKRS were constructed, as reported previously.[6] The M13K07 helper phage is used for phage assembly and infectivity during expression to generate a fusion peptide on the phage surface. In addition, genetic modification of helper phages is necessary for the pIII knockout to guarantee that pIII only comes from phagemid. Herein, we mutated Q350 to stop codon TAA to delete pIII from helper phage. The cloned pADLg3 phagemid, M13K07g3TAA and the synthetase plasmid were electroporated into *E. Coli* Top10 competent cells and expressed with AcrK. Thus, the pIII fusion Ala-Cys-$(Xxx)_6$-AcrK peptide was displayed on the surface of M13 filamentous phage and cyclized through 1,4-addition between cysteine and AcrK. The expression yield for purified phage is 10^{10} colony-forming units/L.

3. Screening a Phage-displayed Cyclic Library Against Target Proteins: TEV Protease and HDAC8

With the 6-mer cyclic phage library in hand, we performed the selection against two target proteins: a well-studied target protein tobacco etch virus (TEV) protease and a cancer related protein HDAC8, respectively. A biotin capture reagent, here biotin-sulfosuccinimidyl ester, which could react with a primary amine, was used for the biotinylation of target proteins. Next, the 6-mer cyclic phage library with more than 10^9 different pIII fusion peptides was subjected to three consecutive rounds of selection against two biotinylated target proteins immobilized on streptavidin magnetic beads, respectively. To remove impact from non-specific binding, phage was pre-incubated with streptavidin

magnetic beads before every round of selection. After each round of selection, Top10F' cells are infected by the selected phage, and were used to extract selected phagemid DNA. Next, these selected phagemids were amplified, and then subjected to the next round of selection (**Figure 40**).

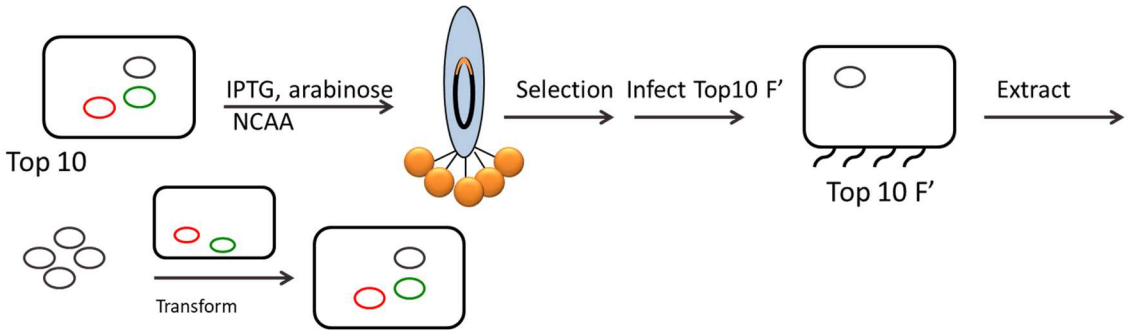


Figure 40 Phage generation and amplification

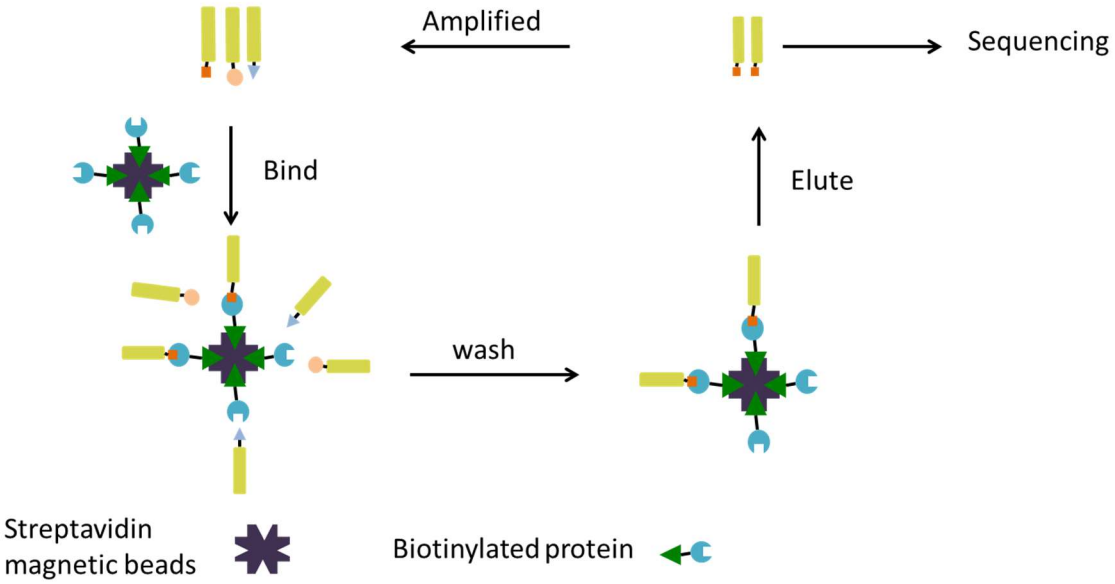


Figure 41 General process of phage selection

After three rounds of selection against streptavidin mobilized biotinylated-TEV protease, dramatically enriched phage-displayed peptide sequences were observed. Eluted number of phages increase 250 folds from the first round to the last round. A number of clones were chosen and sent for DNA sequencing to identify the sequences. Sequencing results show that the cyclic peptide sequence, CWRDYLI-AcrK, is the most abundant sequence, followed by CQWFSHR-AcrK and CGTWLKF-AcrK (**Table 1**). Later these three peptide sequences were synthesized through solid phase peptide synthesis (SPPS) method. Solid phase peptide synthesis is based on a solid resin and synthesized from C-terminal to N-terminal using N-terminal protected amino acids. HPLC allows the purification of the final peptide products. Isothermal thermal titration (ITC) was used to characterize the binding affinity between selected cyclic peptides and target protein TEV protease.

For selection against HDAC8, phage enrichment was also observed after three rounds of selection against streptavidin mobilized succinimidyl-HDAC8. Eluted number of phage particles increases from 40 folds from the first round to the third round. Out of all identified phagemids sequence, CQSLWMN-AcrK is the most abundant sequence, followed by CKHSLWV-AcrK and CLSDCRU-AcrK (**Table 1**).

Table 1 Selected peptide sequencing result*

Target protein	Peptide	Number of clones
TEV protease	CWRDYLI-AcrK	10
	CQWFSHR-AcrK	8
	CGTWLKF-AcrK	7
HDAC8	CQSLWMN-AcrK	8
	CKHSLWV-AcrK	2
	CLSDCRU-AcrK	1

*36 clones were sequenced after the 3rd round of selections.

4. Fluorescence Polarization Study of HDAC8 and TEV Protease

Having identified the selected peptide sequence, we characterized interaction and inhibition ability of them against target proteins. Fluorescence polarization (FP) is a sensitive nonradioactive approach for the measurement of interactions between macromolecules when one of the reactant is relatively small and fluorophore-labeled. When excited by polarized light, the fluorescent labeled molecule emits light with a degree of polarization that is inversely proportional to the molecular rotation of fluorophore-labeled molecule.[86] Moreover, the molecular rotation is related to the molecular weight. A lager molecule has a slower rotation.[86,87] Thus, when a small fluorophore binds to a macromolecule, association rate or dissociation rate could be determined by monitoring the change of polarization. In general, a small fluorophore is less than 1500 Da. Sometimes

a fluorophore can be up 5000 Da if the binding protein is much larger than itself. Quantitatively, FP is termed as the difference between the parallel and perpendicular emission intensity divided by the total of the parallel and perpendicular emission intensity:

$$FP = (I_{parallel} - I_{perpendicular}) / (I_{parallel} - I_{perpendicular})$$

where $I_{parallel}$ is the intensity of emission light which is parallel to the excitation light plane and $I_{perpendicular}$ is the intensity of emission light which is perpendicular to the excitation light plane.

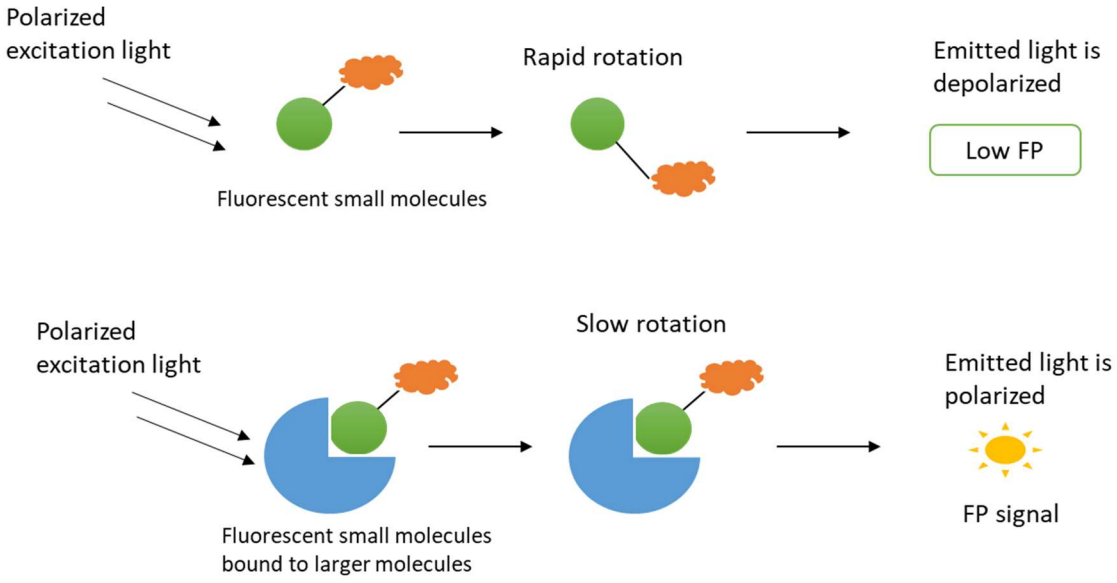


Figure 42 Mechanism of fluorescence polarization

In this dissertation, small peptides conjugated with 5-FAM fluorophore were used as fluorescence polarization assay. For TEV protease, we chose two most abundant peptide sequences and synthesized their fluorophore-labeled derivatives: cyclo(CWRDYLI-AcrK-K(5-FAM)) and cyclo(CQWFSHR-AcrK-K(5-FAM)). For HDAC8, we chose the most abundant peptide and synthesized its fluorophore-labeled derivative: cyclo(CQSLWMN-AcrK-K(5-FAM)). At the same time, linear counterparts without acryloyl group on the lysine were also synthesized as linear controls. To determine the interaction between fluorophore-labeled peptide probes and target proteins, fluorophore-labeled peptide probes in PBS buffer were added into a 96-well, black-bottom plate. Increasing concentrations of target proteins were added to different well, and mixed with previously added fluorophore-labeled peptide probe. The total reaction volume in each well is 200 uL. The final concentration of each fluorophore-labeled peptide probe is fixed to 25 nM, while the final concentration of each target proteins ranges from 100 nM to 120 μM. The plates were subjected to BioTek Synergy H1 microplate reader, with a 490 nm excitation filter and a 520 nm emission filter at 30 °C.

The shape of relation curve of polarization and protein concentration is sigmoidal. From the curve of polarization vs protein concentration, dissociation rate (k_d) was calculated. Compared to liner counterparts without acryloyl group in the lysine, all fluorophore-labeled cyclic peptide probes have better binding affinity with 5-7 folds smaller k_d (**Figure 43**). Thus, we have established a methodology for the discovery of cyclic peptide inhibitors with significant increased affinity and better specificity.

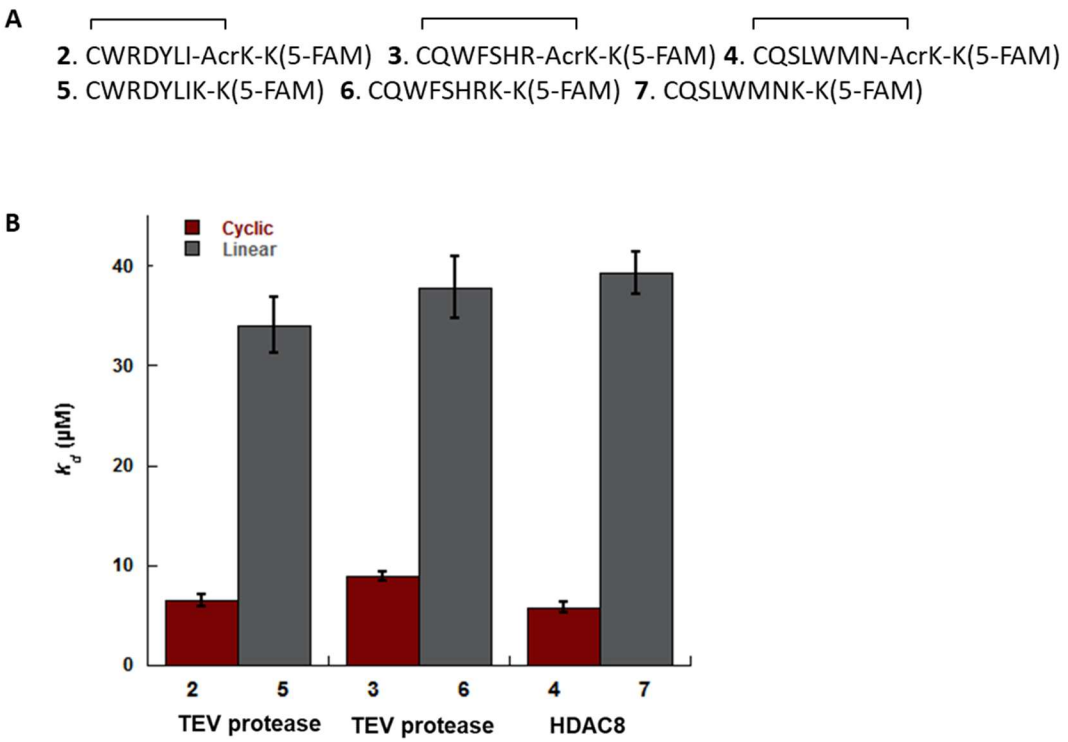


Figure 43 (**A**) Selected cyclic peptides conjugated to 5-FAM fluorophore as FP probes (**2** - **4**); linear peptides conjugated to 5-FAM fluorophore as FP probes (**5** - **7**). (B) Binding affinity of FP probes (**2**-**7**) to TEV protease and HDAC8.

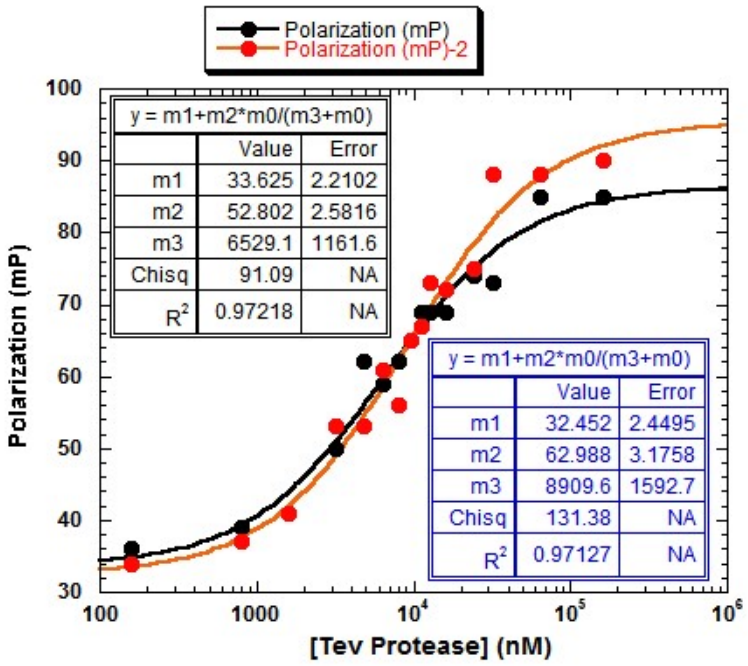


Figure 44 Fluorescence polarization measurement of cyclic peptides CWRDYLI-AcrK-K(5-FAM) and CQWFSHR-AcrK-K(5-FAM) with TEV Protease

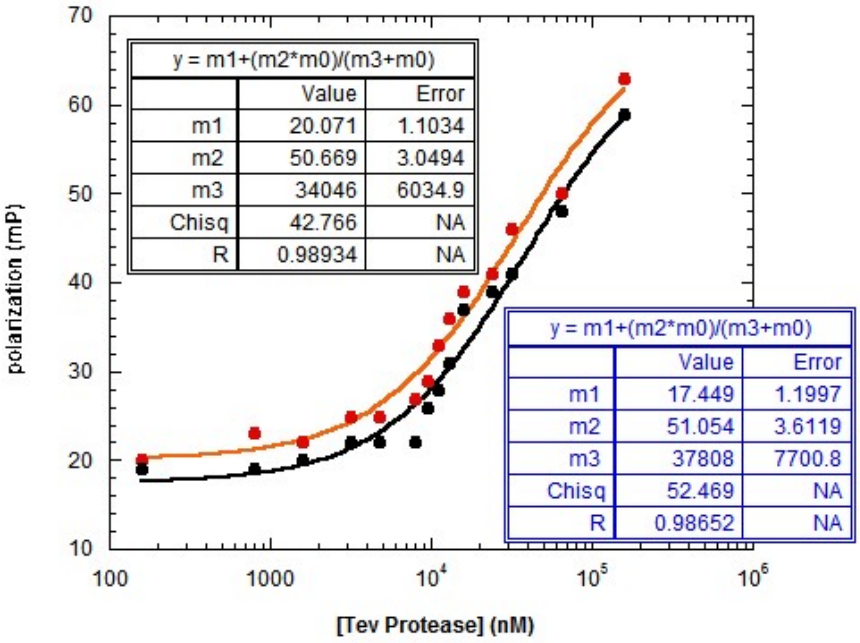


Figure 45 Fluorescence polarization measurement of linear peptides CWRDYLIK-K(5-FAM) and CQWFSHRK-K(5-FAM) with TEV protease

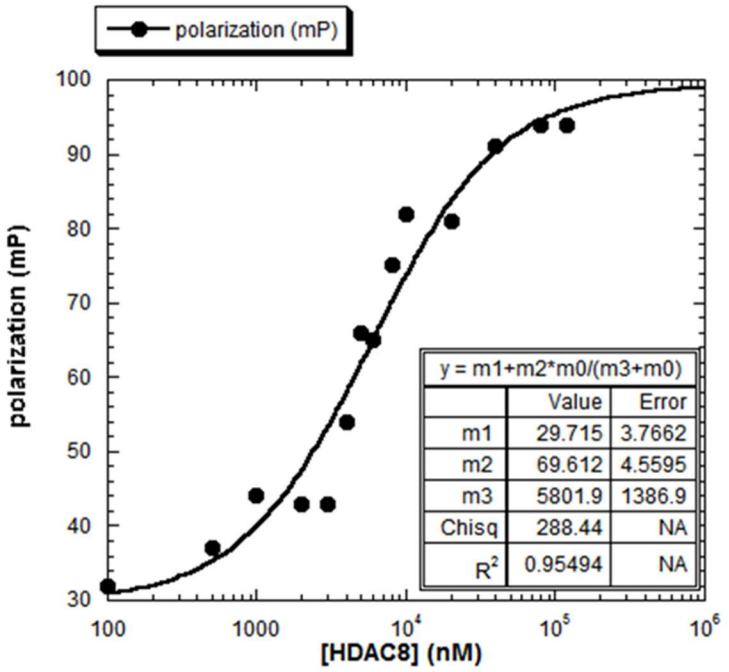


Figure 46 Fluorescence polarization measurement of cyclic peptide CQSLWMN-AcrK-K(5-FAM) with HDAC8

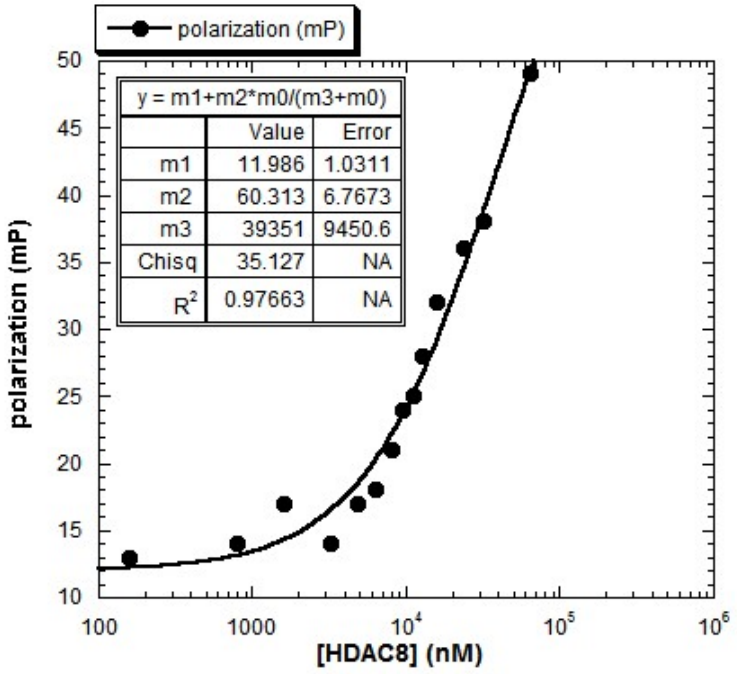


Figure 47 Fluorescence polarization measurement of linear peptide CQSLWMNK-K(5-FAM) with HDAC8

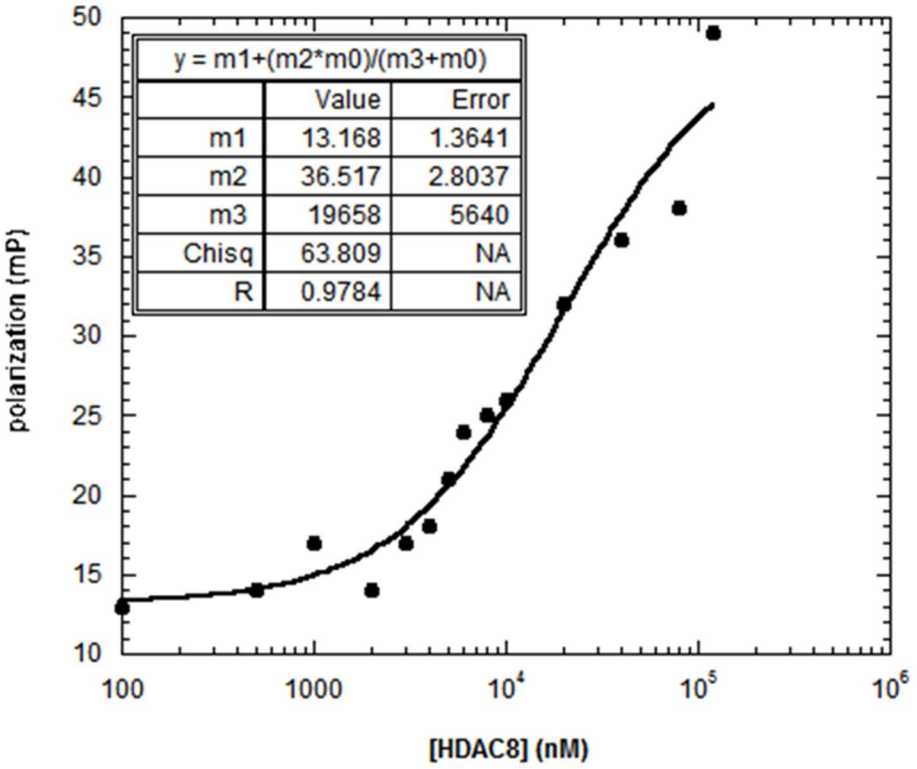


Figure 48 Fluorescence polarization measurement of cyclic peptide CWRDYLI-AcrK-K(5-FAM) with HDAC8

5. Inhibition Ability Measurement

Inhibition ability of selected peptide with HDAC8 was investigated by the measurement of half maximal inhibitory concentration (IC_{50} value). In this dissertation, a fluorophore 7-Amino-4-methylcoumarin (AMC) derivative, Boc-lys(Ac)-AMC, was used to obtain IC_{50} value of HDAC8.[85,88,89] Boc-lys(Ac)-AMC was synthesized from α-BOC-ε-acetyl lysine and 7-Amino-4-methylcoum using $POCl_3$ and pyridine.[90] Different concentrations ranging from 1 nM to 100 μM of selected cyclic peptide, CQSLWMN-AcrK, were mixed with 2 μM of purified HDAC8 in black 96-well plate. PBS buffer was added to make the total reaction volume in each well 200 μL. 50 μM Boc-lys(Ac)-AMC was added to the solution. After 1 hour incubation, 2 μM TSA was added to the solution to stop the HDAC8 inhibition activity, followed by the addition of trypsin (0.5 mg/mL). The plate was incubated at 30 °C for 2 hours, allowing the cleavage of trypsin of deacetylated assay and releasing free AMC fluorophore. Intensity was normalized for the inhibition ability. IC_{50} value for CQSLWMN-AcrK against HDAC8 is 9.7±0.4 μM. Compared to well-studied HDAC inhibitors such as SAHA (IC_{50} = 306 nM) and TSA (IC_{50} = 110 nM), the inhibition ability is not very strong.[91] However, as the library only contains 6 randomized amino acids, the inhibition ability for this small size library is good. We expect by screening against a cyclic phage library with more positions with randomized amino acids, potential inhibitors would be selected with lower IC_{50} value and increased inhibition ability.

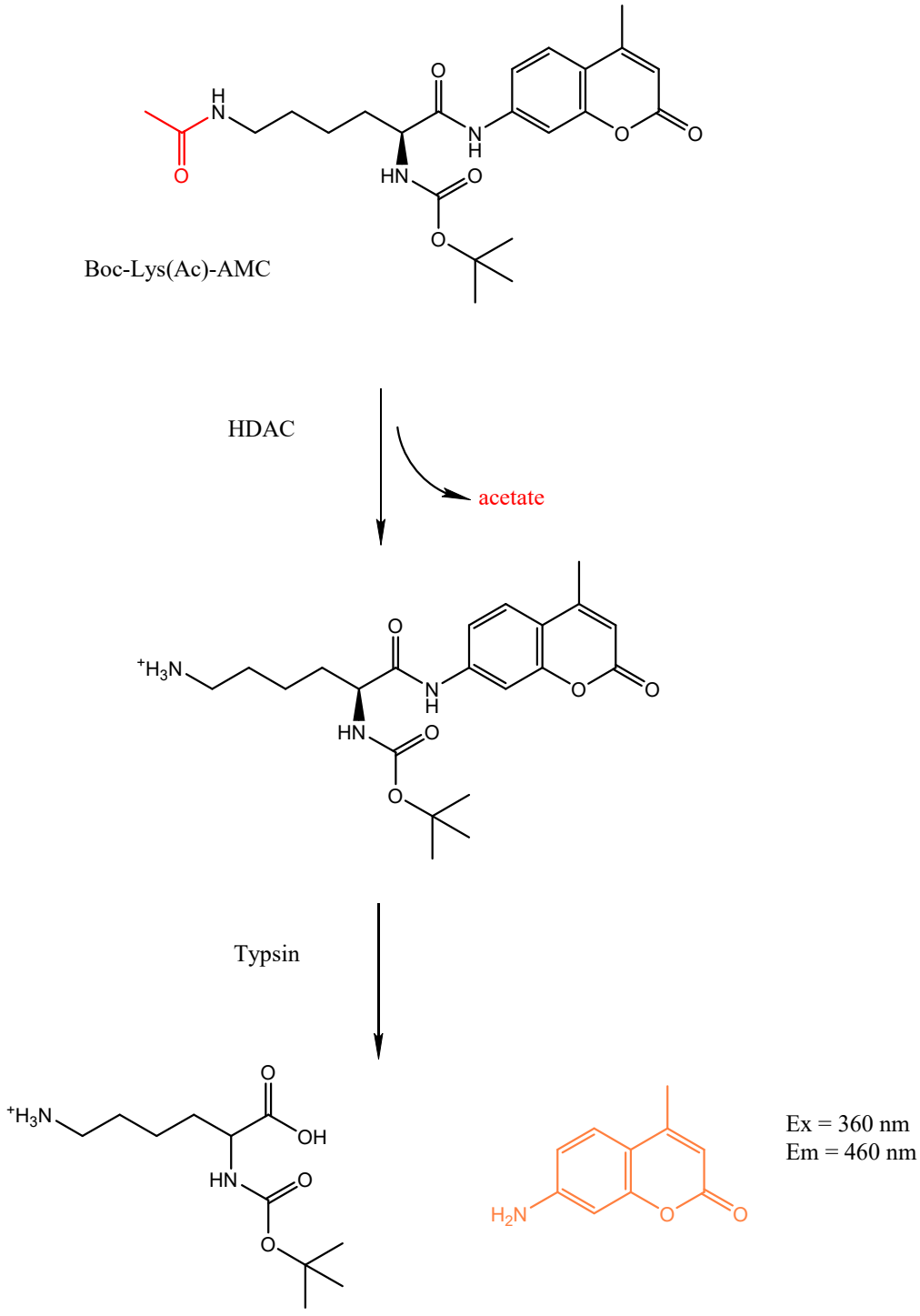


Figure 49 Scheme of Boc-Lys(Ac)-AMC assay. Inhibition activity is determined by monitoring fluorescence form methylcoumarin fluorophores.

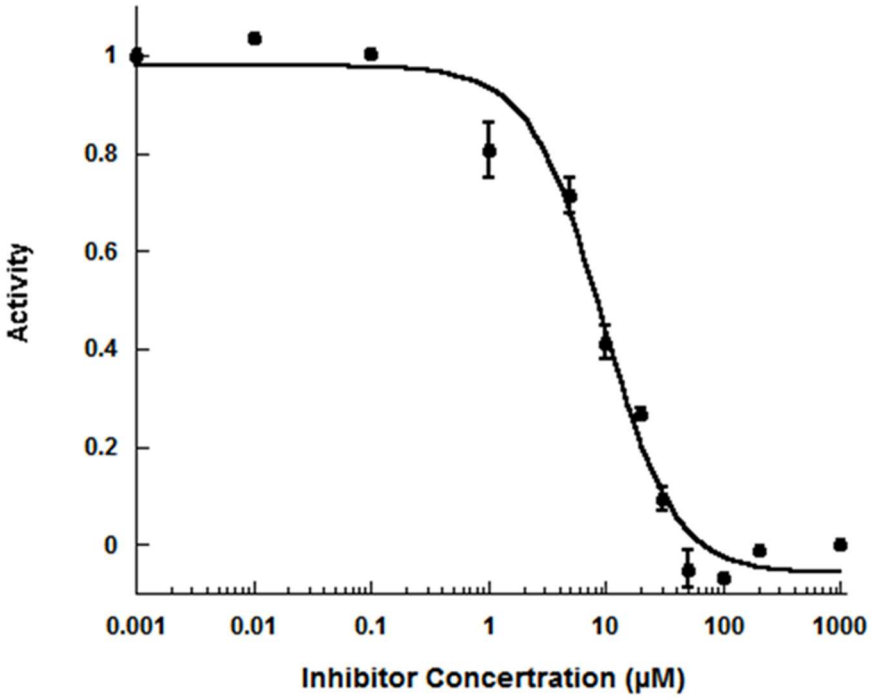


Figure 50 I_{50} measurement of cyclic peptide CQSLWMN-AcrK with HDAC8

CHAPTER IV

CONCLUSIONS

In the first part of this study, we proposed a mechanism for nitrile-imine cycloaddition in the absence of chloride. Then we performed a serial of kinetic study and protein labeling experiment, demonstrating that the nitrile imine cycloaddition is an ultra rapid reaction. Also, we found an optimized reaction condition, pH 10 without chloride, for efficient protein labeling which can be finished in 1 min.

In the second part, we have established a methodology to select bivalent inhibition ligands from cyclic phage library against specific target proteins with high affinity, better specificity and increased inhibition ability. In this study, we incorporated non-canonical amino acid AcrK and constructed 6-mer cyclic peptide phage library. The cyclization was through 1,4-addition between cysteine and acryloyl moiety. Phage display methodology was used for the selection of target proteins via three consecutive rounds of screening against streptavidin mobilized biotinylated TEV protease and biotinylated HDAC8, respectively. Fluorescence polarization measurements have indicated the high specificity and affinity of selected peptides with target proteins. IC_{50} value of selected peptide, CQSLWMN-AcrK, against HDAC8 is 9.7 ± 0.4 μM. We expect larger library such as 12-mer cyclic phage library would have increased inhibition ability. Also, further studies will focus on inhibition mode based on protein structure, cellular permeability and delivery, and *in vivo* study. Moreover, we will also apply this methodology to discover peptide-based drugs for more HDACs and other human disease related proteins.

REFERENCES

1. Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The Future of Peptide-based Drugs. *Chem. Biol. Drug Des.* **81,** 136–147 (2013).

2. Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* **20,** 122–128 (2015).

3. Nielsen, D. S. *et al.* Orally Absorbed Cyclic Peptides. *Chem.Rev.* **117,** 8128 (2017).

4. Joo, S. H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **20,** 19–26 (2012).

5. Zorzi, A., Deyle, K. & Heinis, C. Cyclic peptide therapeutics: past, present and future. *Curr. Opin. Chem. Biol.* **38,** 24–29 (2017).

6. Lee, Y.-J. *et al.* A Genetically Encoded Acrylamide Functionality. *ACS Chem. Biol.* **8,** 1664–1670 (2013).

7. Berth-Jones, J. The use of ciclosporin in psoriasis. *J. Dermatolog. Treat.* **16,** 258–277 (2005).

8. Rakonjac, J., Bennett, N. J., Spagnuolo, J., Gagic, D. & Russel, M. Filamentous Bacteriophage: Biology, Phage Display and Nanotechnology Applications. *Curr. Issues Mol. Biol.* **13,** 51–76 (2011).

9. Smith, G. P. & Petrenko, V. A. Phage Display. *Chem. Rev.* **97,** 391–410 (1997).

10. Qi, H., Lu, H., Qiu, H.-J., Petrenko, V. & Liu, A. Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. *J. Mol. Biol.* **417,** 129–143 (2012).

11. Meyer, S. C., Shomin, C. D., Gaj, T. & Ghosh, I. Tethering Small Molecules to a Phage Display Library: Discovery of a Selective Bivalent Inhibitor of Protein Kinase A. *J. AM. CHEM. SOC* **129,** 13812–13813 (2007).

12. Schlippe, Y. V. G., Hartman, M. C. T., Josephson, K. & Szostak, J. W. In Vitro Selection of Highly Modified Cyclic Peptides That Act as Tight Binding Inhibitors. *J. AM. CHEM. SOC* **134,** 10469–10477 (2012).

13. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **7,** 502–507 (2009).

14. Tian, F., Tsao, M.-L. & Schultz, P. G. A Phage Display System with Unnatural Amino Acids. *J. Am. Chem. Soc.* **126,** 15962–15963 (2004).

15. Lee, H. S., Spraggon, G., Schultz, P. G. & Wang, F. Genetic Incorporation of a Metal-Ion Chelating Amino Acid into Proteins as a Biophysical Probe. *J. Am. Chem. Soc.* **131,** 2481–2483 (2009).

16. Ng, S. & Derda, R. Organic & Biomolecular Chemistry Phage-displayed macrocyclic glycopeptide libraries. *Org. Biomol. Chem.* **14,** (2016).

17. Wang, L., Magliery, T. J., Liu, D. R. & Schultz, P. G. A New Functional Suppressor tRNA/Aminoacyl−tRNA Synthetase Pair for the in Vivo Incorporation of Unnatural Amino Acids into Proteins. *J. Am. Chem. Soc.* **122,** 5010–5011

(2000).

18. Ng, S. & Derda, R. Organic & Biomolecular Chemistry Phage-displayed macrocyclic glycopeptide libraries. *Org. Biomol. Chem* **14,** (2016).

19. Wan, W., Tharp, J. M. & Liu, W. R. Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta - Proteins Proteomics* **1844,** 1059–1070 (2014).

20. Schilling, C. I., Jung, N., Biskup, M., Schepers, U. & Brae, S. Bioconjugation via azide–Staudinger ligation: an overview. *Chem. Soc. Rev. Chem. Soc. Rev* **40,** 4840–4871 (2011).

21. Prescher, J. A., Dube, D. H. & Bertozzi, C. R. Chemical remodelling of cell surfaces in living animals. *Nature* **430,** 873 (2004).

22. Lang, K. & Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem.Rev.* **114,** 4764–4806 (2014).

23. McKay, C. S. & Finn, M. G. Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chem. Biol.* **21,** 1075–1101 (2014).

24. Beue, D., Qiao, G. G. & Wentrup, C. Nitrile Imines: Matrix Isolation, IR Spectra, Structures, and Rearrangement to Carbodiimides. *J. Am. Chem. Soc* **134,** 5339−5350 (2012).

25. Beue, D. & Wentrup, C. Carbenic Nitrile Imines: Properties and Reactivity. *J. Org. Chem.* **79,** 1418–1426 (2014).

26. Sibi, M. P., Stanley, L. M. & Jasperse, C. P. An Entry to a Chiral Dihydropyrazole Scaffold: Enantioselective [3 + 2] Cycloaddition of Nitrile Imines. *J. AM. CHEM. SOC* **127,** 8276–8277 (2005).

27. Zhang, Y., Liu, W. & Zhao, Z. K. Nucleophilic Trapping Nitrilimine Generated by Photolysis of Diaryltetrazole in Aqueous Phase. *Molecules* **19,** 306–315 (2014).

28. Toseland, C. P. Fluorescent labeling and modification of proteins. *J. Chem. Biol.* **6,** 85–95 (2013).

29. Vázquez, A. *et al.* Mechanism-Based Fluorogenic *trans* -Cyclooctene-Tetrazine Cycloaddition. *Angew. Chemie Int. Ed.* **56,** 1334–1337 (2017).

30. Song, W., Wang, Y., Qu, J., Madden, M. M. & Lin, Q. A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angew. Chem., Int. Ed.* **47,** 2832–2835 (2008).

31. Ramil, C. P. & Lin, Q. Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. *Curr. Opin. Chem. Biol.* **21,** 89–95 (2014).

32. Ramil, C. P. & Lin, Q. Bioorthogonal chemistry: strategies and recent developments. *Chem. Commun.* **49,** 11007–11022 (2013).

33. Song, W., Wang, Y., Qu, J. & Lin, Q. Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via Photoclick Chemistry in Bacterial Cells. *J. Am. Chem. Soc.,* **130,** 9654–9655 (2008).

34. Shang, X. *et al.* Fluorogenic protein labeling using a genetically encoded

unstrained alkene. *Chem.Sci* **8,** (2017).

35. Eberharter, A. & Becker, P. B. Histone acetylation: a switch between repressive and permissive chromatin Second in review series on chromatin dynamics. *EMBO Rep.* **3,** 224–229 (2002).

36. Wolfson, N. A., Ann Pitcairn, C. & Fierke, C. A. HDAC8 substrates: Histones and beyond. *Biopolymers* **99,** 112–126 (2013).

37. Seto, E. & Yoshida, M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harb Perspect Biol* **6,** 1–26 (2014).

38. Marks, P. A. *et al.* Histone deacetylases and cancer: causes and therapies. *Nat. Rev. Cancer* **1,** 194 (2001).

39. Finnin, M. S. *et al.* Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* **401,** 188 (1999).

40. Negmeldin, A. T., Knoff, J. R. & Pflum, M. K. H. The structural requirements of histone deacetylase inhibitors: C4-modified SAHA analogs display dual HDAC6/HDAC8 selectivity. *Eur. J. Med. Chem.* **143,** 1790–1806 (2018).

41. Kelly, W. K. *et al.* Phase I Study of an Oral Histone Deacetylase Inhibitor, Suberoylanilide Hydroxamic Acid, in Patients With Advanced Cancer. *J. Clin. Oncol.* **23,** 3923–3931 (2005).

42. Chakrabarti, A. *et al.* HDAC8: a multifaceted target for therapeutic interventions. *Trends Pharmacol. Sci.* **36,** 481–492 (2015).

43. Alam, N. *et al.* Structure-Based Identification of HDAC8 Non- histone Substrates Structure-Based Identification of HDAC8 Non-histone Substrates. *Structure* 458–468 (2016). doi:10.1016/j.str.2016.02.002

44. Vannini, A. *et al.* Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 15064–15069 (2004).

45. Somoza, J. R. *et al.* Structural Snapshots of Human HDAC8 Provide Insights into the Class I Histone Deacetylases Treatment with HDIs causes tumor cells to cease growth and to either differentiate or become apoptotic. These findings have led to interest in the use of HDIs as an. *Structure* **12,** 1325–1334 (2004).

46. Vannini, A. *et al.* Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex. *EMBO Rep.* **8,** 879–884 (2007).

47. Dowling, D. P., Gantt, S. L., Gattis, S. G., Fierke, C. A. & Christianson, D. W. Structural Studies of Human Histone Deacetylase 8 and Its Site-Specific Variants Complexed with Substrate and Inhibitors,. *Biochemistry* **47,** 13554–13563 (2008).

48. Gantt, S. L., Joseph, C. G. & Fierke, C. A. Activation and Inhibition of Histone Deacetylase 8 by Monovalent Cations. *J. Biol. Chem.* **285,** 6036–6043 (2010).

49. Gantt, S. L., Gattis, S. G. & Fierke, C. A. Catalytic Activity and Inhibition of Human Histone Deacetylase 8 Is Dependent on the Identity of the Active Site Metal Ion. *Biochemistry* **45,** 6170–6178 (2006).

50. Liu, D. S. *et al.* Diels–Alder Cycloaddition for Fluorophore Targeting to Specific

Proteins inside Living Cells. *J. Am. Chem. Soc.* **134,** 792–795 (2012).

51. Devaraj, N. K., Upadhyay, R., Haun, J. B., Hilderbrand, S. A. & Weissleder, R. Fast and Sensitive Pretargeted Labeling of Cancer Cells through a Tetrazine/trans-Cyclooctene Cycloaddition. *Angew. Chemie Int. Ed.* **48,** 7013–7016 (2009).

52. Taylor, M. T., Blackman, M. L., Dmitrenko, O. & Fox, J. M. Design and Synthesis of Highly Reactive Dienophiles for the Tetrazine–trans-Cyclooctene Ligation. *J. Am. Chem. Soc.* **133,** 9646–9649 (2011).

53. Blackman, M. L., Royzen, M. & Fox, J. M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity. *J. Am. Chem. Soc.* **130,** 13518–13519 (2008).

54. Devaraj, N. K., Hilderbrand, S., Upadhyay, R., Mazitschek, R. & Weissleder, R. Bioorthogonal Turn-On Probes for Imaging Small Molecules inside Living Cells. *Angew. Chemie Int. Ed.* **49,** 2869–2872 (2010).

55. Chang, P. V *et al.* Copper-free click chemistry in living animals. *Proc. Natl. Acad. Sci. U. S. A.,* **107,** 1821–1826 (2010).

56. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U. S. A.,* **104,** 16793–16797 (2007).

57. Sletten, E. M., Nakamura, H., Jewett, J. C. & Bertozzi, C. R. Difluorobenzocyclooctyne: Synthesis, Reactivity, and Stabilization by β-Cyclodextrin. *J. Am. Chem. Soc.* **132,** 11799–11805 (2010).

58. Codelli, J. A., Baskin, J. M., Agard, N. J. & Bertozzi, C. R. Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. *J. Am. Chem. Soc.* **130,** 11486–11493 (2008).

59. Rostovtsev, V. V, Green, L. G., Fokin, V. V & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective 'Ligation' of Azides and Terminal Alkynes. *Angew. Chemie Int. Ed.* **41,** 2596–2599 (2002).

60. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie Int. Ed.* **40,** 2004–2021 (2001).

61. Chan, T. R., Hilgraf, R., Sharpless, K. B. & Fokin, V. V. Polytriazoles as Copper(I)-Stabilizing Ligands in Catalysis. *Org. Lett.* **6,** 2853–2855 (2004).

62. Wu, B., Wang, Z., Huang, Y. & Liu, W. R. Catalyst-Free and Site-Specific One-Pot Dual-Labeling of a Protein Directed by Two Genetically Incorporated Noncanonical Amino Acids. *ChemBioChem* **13,** 1405–1408 (2012).

63. Rossin, R. *et al.* Highly Reactive trans-Cyclooctene Tags with Improved Stability for Diels–Alder Chemistry in Living Systems. *Bioconjug. Chem.* **24,** 1210–1217 (2013).

64. Ning, X. *et al.* Protein Modification by Strain-Promoted Alkyne–Nitrone Cycloaddition. *Angew. Chemie Int. Ed.* **49,** 3065–3068 (2010).

65. Plass, T. *et al.* Amino Acids for Diels–Alder Reactions in Living Cells. *Angew. Chemie Int. Ed.* **51,** 4166–4170 (2012).

66. Lang, K. *et al.* Genetic Encoding of Bicyclononynes and trans-Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic Diels–Alder Reactions. *J. Am. Chem. Soc.* **134,** 10317–10320 (2012).

67. McKay, C. S., Moran, J. & Pezacki, J. P. Nitrones as dipoles for rapid strain-promoted 1,3-dipolar cycloadditions with cyclooctynes. *Chem. Commun.* **46,** 931–933 (2010).

68. Gordon, C. G. *et al.* Reactivity of Biarylazacyclooctynones in Copper-Free Click Chemistry. *J. Am. Chem. Soc.* **134,** 9199–9208 (2012).

69. Mckay, C. S., Chigrinova, M., Blake, J. A. & Pezacki, J. P. Organic & Biomolecular Chemistry Kinetics studies of rapid strain-promoted [3 + 2]-cycloadditions of nitrones with biaryl-aza-cyclooctynone. *Org. Biomol. Chem* **10,** (2012).

70. Huisgen, R. 1,3-Dipolar Cycloadditions. Past and Future. *Angew. Chemie Int. Ed. English* **2,** 565–598 (1963).

71. Huisgen, R. *et al.* 1.3-Dipolare Additionen, II. Synthese von 1.2.4-Triazolen aus Nitriliminen und Nitrilen. *Justus Liebigs Ann. Chem.* **653,** 105–113 (1962).

72. Song, W., Wang, Y., Qu, J. & Lin, Q. Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via 'Photoclick Chemistry' in Bacterial Cells. *J. Am. Chem. Soc.* **130,** 9654–9655 (2008).

73. Wang, Y., Song, W., Hu, W. J. & Lin, Q. Fast Alkene Functionalization In Vivo by Photoclick Chemistry: HOMO Lifting of Nitrile Imine Dipoles. *Angew. Chemie*

Int. Ed. **48,** 5330–5333 (2009).

74. Song, W., Wang, Y., Qu, J., Madden, M. M. & Lin, Q. A Photoinducible 1,3-Dipolar Cycloaddition Reaction for Rapid, Selective Modification of Tetrazole-Containing Proteins. *Angew. Chemie Int. Ed.* **47,** 2832–2835 (2008).

75. Yu, Z., Pan, Y., Wang, Z., Wang, J. & Lin, Q. Genetically Encoded Cyclopropene Directs Rapid, Photoclick-Chemistry-Mediated Protein Labeling in Mammalian Cells. *Angew. Chemie Int. Ed.* **51,** 10600–10604 (2012).

76. Kaya, E. *et al.* A Genetically Encoded Norbornene Amino Acid for the Mild and Selective Modification of Proteins in a Copper-Free Click Reaction. *Angew. Chemie Int. Ed.* **51,** 4466–4469 (2012).

77. Hegarty, A. F., Cashman, M. P. & Scott, F. L. The kinetics of nitrilimine formation in base-catalysed hydrolysis of hydrazonyl halides. *J. Chem. Soc, Perkin Trans. 2* 44–52 (1972).

78. Billups, D. & Attwell, D. Control of intracellular chloride concentration and GABA response polarity in rat retinal ON bipolar cells. *J. Physiol.* **545,** 183–198 (2002).

79. Behrendt, R., White, P. & Offer, J. Advances in Fmoc solid-phase peptide synthesis. *J. Pept. Sci.* **22,** 4–27 (2016).

80. O'Neil, K. T. *et al.* Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library. *Proteins Struct. Funct. Bioinforma.* **14,** 509–515 (1992).

81. Deyle, K., Kong, X.-D. & Heinis, C. Phage Selection of Cyclic Peptides for Application in Research and Drug Development. *Acc. Chem. Res.* **50,** 1866–1874 (2017).

82. Palomo, J. M. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC Adv.* **4,** 32658–32672 (2014).

83. Sarin, V. K., Kent, S. B. H., Tam, J. P. & Merrifield, R. B. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. *Anal. Biochem.* **117,** 147–157 (1981).

84. Kaiser, E., Colescott, R. L., Bossinger, C. D. & Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34,** 595–598 (1970).

85. Galleano, I., Nielsen, J. & Madsen, A. S. Letter Syn lett Scalable and Purification-Free Synthesis of a Myristoylated Fluoro- genic Sirtuin Substrate. **28,** 2169–2173 (2017).

86. Arkin, M. R., Glicksman, M. A., Fu, H., Havel, J. J. & Du, Y. Inhibition of Protein-Protein Interactions: Non- Cellular Assay Formats. (2012).

87. Ng, Y. Z., Baldera-Aguayo, P. A. & Cornish, V. W. Fluorescence Polarization Assay for Small Molecule Screening of FK506 Biosynthesized in 96-Well Microtiter Plates. *Biochemistry* **56,** 5260–5268 (2017).

88. Wegener, D., Wirsching, F., Riester, D. & Schwienhorst, A. A Fluorogenic Histone Deacetylase Assay Well Suited for High-Throughput Activity Screening.

Chem. Biol. **10,** 61–68 (2003).

89. Grozinger, C. M. & Schreiber, S. L. Deacetylase Enzymes: Biological Functions Review and the Use of Small-Molecule Inhibitors. *Chem. Biol.* **9,** 3–16 (2018).

90. Hoffmann, K., Brosch, G., Loidl, P. & Jung, M. A non-isotopic assay for histone deacetylase activity. *Nucleic Acids Res.* **27,** (1999).

91. Huber, K. *et al.* Inhibitors of Histone Deacetylases CORRELATION BETWEEN ISOFORM SPECIFICITY AND REACTIVATION OF HIV TYPE 1 (HIV-1) FROM LATENTLY INFECTED CELLS. *jounal Biol. Chem.* **286,** 22211–22218 (2011).