DEVELOPING HIGH-THROUGHPUT CHEMICAL APPROACHES FOR PROTEOMIC PROFILING OF ASPARTIC PROTEASES AND PROTEIN KINASES

SHI HAIBIN

(M.Sc., SOOCHOW UNIVERSITY)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGEMENTS

I would like to give special thank to my helpful supervisor, **Prof. Yao Shao Qin** for his invaluable guidance, patience, encouragement and endless support throughout my four-year study. He has brought me into the fascination world of chemical biology and ignited in me an inexorable passion for science. The supervision and support that he gave truly help me make great progression and improvement in many aspects. His unparalleled passion, indefatigable spirits and professionalism have been deeply engraved upon my mind. My deepest appreciation is for Prof Yao -my teacher and mentor.

I sincerely thank all the labmates both in Chemistry and DBS Labs-Liu Kai, Chongjing, Xiamin, Grace, Jigang, Laypheng, Candy, Wu Hao, Pengyu, Mingyu, Li Lin, Jingyan, Liqian, Zhenkun, Su Ying, Zhengqiu, Xiaoyuan, Wendy, Ashley as well as some senior members in Yao lab: Mahesh, Kalesh, Hongyan, Raja, Souvik. I would like to take this opportunity to thank each of you for invaluable assistances, fruitful discussions and happy memories over these years. The cooperation is much appreciated. I would like to give special thanks to my collaborators, Liu Kai helped me to carry out many biological works, and I have learnt more from him. Besides, I am very thankful to Wu Hao's help. Thank Kalesh, Chongjing and Xiamin help me to synthesize the kinase probes, Jigang taught me how to carry out the protein identification work. Thanks again for your helps.

I also appreciate the support from Mdm Han Yanhui and Ms Peggy Ler from

NMR lab as well as other chemistry support labs.

Last but not least, I would like to express my heartful gratitude to my parents, my wife and my son for their love, understanding, encouragement, patience and moral supports over these years. This thesis will be dedicated to them.

I also acknowledge kind support from NUS for providing me research scholarship.

Table of Contents

Summary

With the advances from genome-sequencing projects, the functional annotation and characterization of newly discovered proteins have become increasingly important. As key regulators of virtually every biological process, even minor imbalances in the activity of certain enzymes might lead to severe pathological conditions. Deeply understandings of how an enzyme works at its molecular and cellular level will have direct relevance to drug discovery. To reach the goal of large scale studies of a significant number of enzymes present in any particular organism, development of high-throughput amenable chemistry and enzyme-screening tools are becoming very urgent. This thesis examines and addresses these challenges by combination of various synthetic approaches and introducing high throughput screening platforms. Chapter 2 describes a solution-cum-solid phase strategy for rapid assembly of diol-containing small molecules as potential HIV-1 inhibitors. Chapter 3 presents a high throughput synthetic strategy for rapid assembly of 475-member aspartic protease inhibitors library using “click” reaction. With this platform, new inhibitors against malaria were discovered. Chapter 4 and 5 present a small molecule microarray-facilated platform for proteome profiling of aspartic proteases in cellular lysates. Chapter 6 and 7 describes the cell permeable, photo-affinity drug-like probes for proteome profiling of cellular on/off-targets of drug dasatinib and staurosphorine.

List of Publication

(2006-2012)

1. Sun, H.; Lu, C. H. S.; **Shi, H.**; Gao, L.; Yao, S. Q.* Peptide Microarrays for High-throughput Studies of Ser/Thr Phosphatases. *Nat. Protocols*. (2008), *3*, 1485-1493.
2. **Shi, H.**; Liu, K.; Leong, W. Y. W.; Yao, S. Q.* Expedient Solid-phase Synthesis of Both Symmetric and Asymmetric Diol Libraries Targeting Aspartic Proteases. *Bioorg. Med. Chem. Lett.* (2009), *19*, 3945-3948.
3. **Shi, H.**; Liu, K.; Xu, A.; Yao, S. Q.* Small Molecule Microarray (SMM)-facilitated Screening of Affinity-based Probes (A*f*BPs) for γ-secretase**.** *Chem. Commun***.** (2009), 5030-5032.
4. Liu, K.; **Shi, H.**; Xiao, H.; Chong, A. G. L.; Bi, X.; Chang, Y. T., Tan, K.; Yada, R. Y.; Yao, S. Q.* Functional Profiling, Identification and Inhibition of Plasmepsins in Intraerythrocytic Malaria Parasites. *Angew. Chem. Intl. Ed*. (2009), *48*, 8293-8297.
5. Kalesh, K.A.; **Shi, H.**; Ge, J.; Yao, S. Q.* The Use of Click Chemistry in the Emerging Field of Catalomics. *Org. Biol. Chem.* (2010**)**, *8*, 1749-1762.
6. Loh, Y.; **Shi, H.**; Hu, M.; Yao, S. Q.* Click Synthesis of Small Molecule-Peptide Conjugates for Organelle-Specific Delivery and Inhibition of Lysomosal Cysteine Proteases. *Chem. Commun.* (2010), *46*, 8407-8409.

7. **Shi, H.**; Uttamchandani, M.; Yao, S. Q.* Applying Small Molecule Microarrays and Resulting Affinity Probe Cocktails for Proteome Profiling of Mammalian Cell Lysates. *Chem. -Asian. J. (*2011**)**, *6*, 2803-2815.

8. **Shi, H.;** Zhang, C. -J.; Wu, H.; Yao, S. Q.* Cell-Based Proteome Profiling of Potential Dasatinib™ Targets Using Affinity-Based Probes (A*f*BPs). *J. Am. Chem. Soc*. (2012), In press.

9. **Shi, H.;** Cheng, X.; Sze, S. K.; Yao, S. Q.* Proteome Profiling Reveals Potential Cellular Targets of Staurosporine Using a Clickable Cell-Permeable Probe. *Chem Commun*. (2011), *47*, 11306-11308.

BOOK CHAPTERS

Shi, H.; Uttamchandani, M.; Yao, S. Q. A Method for Small Molecule Microarray-based Screening for The Rapid Discovery of Affinity-based Probes. *Methods Mol. Biol.* (2010), *669*, 57-68.

List of Figures

List of Tables

List of Schemes

List of Abbreviations

δ	Chemical shift in ppm
AcOH	Acetic acid
AA	Amino acid
ABPP	Activity-based protein profiling
Thr	Threonine
Tyr	Tyrosine
Ala	Alanine
Asp	Aspartic acid
Ser	Serine
Val	Valine
Leu	Leucine
Phe	Phenylalanine
Lys	Lysine
Cys	Cysteine
Gly	Glycine
Boc	*tert*-Butoxycarbonyl
Br	Broad
^{t}Bu	*tert*-Butyl
BSA	Bovine serum albumin
Cy3	Cyanine dye3

CuAAC	Copper (I) catalyzed azide-alkyne cycloaddition
C-terminus	Carboxy terminus
Da	Dalton
DBU	1,8-Diazobicyclo[5.4.0]undec-7-ene
DCC	N, N'-Dicyclohexylcarbodiimide
DCM	Dichloromethane
dd	Doublet of doublet
DIC	N, N'-diisopropylcarbodiimide
DIEA	N, N'-diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
DNA	Deoxyribonucleic acid
dNTP	Deoxy Nucleotide Tri phosphate
DOS	Diversity oriented synthesis
EA	Ethyl acetate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide HCl
EDTA	Ethylenediaminetetracetic acid
ESI	Electron Spray Ionization
E. coli	Escherichia coli
Fmoc	9-Fluorenylmethoxycarbonyl
HCl	Hydrochloric acid

HATU	O-(7-azabenzotrizol-1-yl)-1,1,3,3,tetramethyluronium
HOBT	N-Hydroxybenzotriazole
HOAT	1-Hydroxy-7-azabenzotriazole
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High Performance Liquid Chromatography
HT	High-throughput
Hz	Hertz
IC_{50}	Half the maximal inhibitory concentration
K_d	Dissociation constant
LHMDS	Lithium bis(trimethylsilyl)amide
LB	Luria-bertani
M	Multiplet
Min	Minute
mmol	Millimole
MS	Mass spectrometry
NHS	N-Hydroxysuccinimide
nM	Nanomolar
NMR	N-Methylpyrrolidinone
NMP	N-methylpyrrolidone
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium

PyBrOP	Bromo-tris-pyrrolidino-phosphonium hexafluorophosphate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
*p*H	Negative logarithm of the hydroxonium ion concentration
q	Quartet
RBF	Round bottom flask
r.t.	Room temperature
s	Singlet
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SMM	Small molecule microarray
T	Triplet
TBTU	O-(Benzotriazol-1-yl)-N,N,N’,N’-tetramethyluronium
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Tris	Trishydroxymethylamino methane
Uv	Ultraviolet

List of Symbols

Å	angstrom
^{o}C	degree celsius
g	gram
h	hour
k	kilo
l	litre
w	wavelength
m	milli or meter
µ	micro
M	molar
min	minutes
mol	mole
n	nano
p	pico
s	seconds

Chapter 1

Introduction

1.1 Summary

In the last decade, much attention has been turned towards proteomic research after the successful completion of the human genome project.[1] Proteins is very crucial for cellular and metabolic functions. They play crucial roles in cell division, cell migration, cell-cell communication, signal transduction, cell death etc.[2] Dysregulation of certain protein activities has been implicated in many diseases including cancer, AIDS, malaria, Alzheimer's, diabetes, and so on. To date, it has been estimated that enzymes account for more than 20% of all drug targets. However, the physiological role, substrate specificity and downstream targets of many enzymes are poorly understood thus far. Therefore, it is very significant to develop chemical tools capables of studing enzyme activities in a high-throughput manner.[3]

High-throughput amenable reactions are mainly characterized by near-perfect, modular and robust and biocompatible nature and offer an efficient and rapid analysis of proteomes on a global scale. To date, many high-throughput methods have been applied to drug synthesis and proteome analysis, especially in the field of high-throughput enzymology.[4] This chapter will provide a brief overview of applications of high-throughput technologies in chemical proteomics. Much attention will be focused on studies of high-throughput strategies for proteomic profiling of aspartic proteases as well as protein kinases.

1.2 Aspartic Proteases as Therapeutic Targets

Proteases belong to one of the largest enzyme family encoded by the human genome with more than 500 known members. They catalyze the hydrolysis of peptide bonds in the proteins, which is related to many physiological and pathological processes such as cell proliferation, tissue remodelling, embryonic development, blood coagulation, blood pressure control, protein activation and maturation, protein catabolism, protein transport, inflammation, infection, and cancer. The importance of peptide bond cleavage in biological systems is also reflected by the finding that nature has separately invented the necessary catalytic machinery multiple times with different solutions. This provides the basis for the categorization of proteases into five classes: serine, threonine, cysteine, metallo and aspartic proteases (Figure. 1.1).[5]

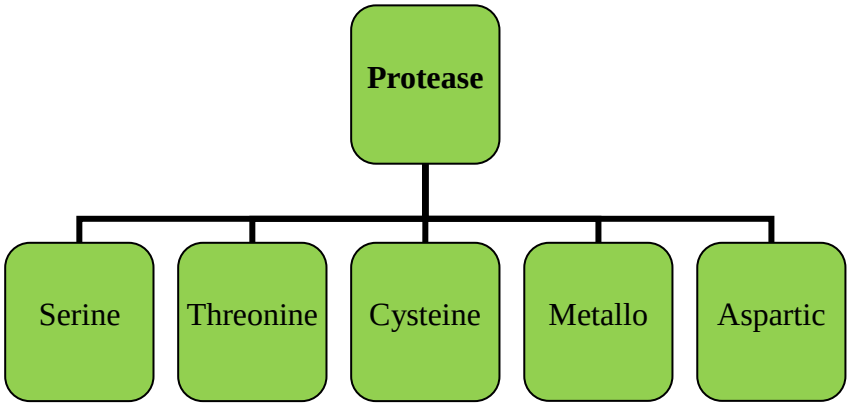


Figure 1.1 Classification of protease family.

In fact, aspartic protease is the smallest class with only 15 members in the human genome. However, they have received considerable attention as potential targets for pharmaceutical intervention since many have been shown to play important roles in physiological and pathological processes, which includes renin in hypertension,[6] cathepsin D in breast cancer,[7] γ, ß-secretase in Alzheimer disease,[8] plasmepsins in

malaria[9] and Human Immunodeficiency Virus-1 protease (HIV-1 PR) in AIDS[10] (Table 1.1).

Table 1.1 Some representative aspartic proteases and their functions in diseases.

Name	Source	Function	Disease
Pepsin	Animal stomach	Digestion of dietary proteins	Digestion disease
Chymosin	Animal stomach	Digestion of dietary proteins	Digestion disease
Cathepsin D,E	Spleen, liver	Lysosomal digestion of proteins	Breast cancer
Renin	Kidney	Regulation of blood pressure	Hypertension
HIV-1	AIDS virus	Processing of AIDS virus proteins	AIDS
Plasmepsins	Food vacuole	Degrade hemoglobin	Malaria
β, γ-Secretase	Brain	Leading to production of amyloid-β	Alzheimer's disease

1.2.1 Catalytic Mechanism of Aspartic Proteases

Proteases are known to play essential roles in many biological processes. They catalyze the hydrolysis of peptide bonds with high sequence selectivity and catalytic efficiency. These enzymes carry out their catalysis by different mechanisms. Aspartic proteases are characterized by two catalytic aspartic acid residues located in their active site. Although the catalytic mechanism of aspartic proteases is poorly understood, it has been generally accepted that peptide bond cleavage occurs by a general acid-base catalytic mechanism (Figure 1.2). One of the two catalytic aspartic residues is protonated in the enzyme substrate complex. The other aspartic residue acts as a general base activating a water molecule which then attacks the carbonyl carbon of the scissile amide bond resulting in the formation of a tetrahedral diol

intermediate. Subsequent deprotonation of the hydroxyl group by one of the catalytic aspartates and simultaneous activation of the leaving amine by the other protonated aspartic residue ultimately leads to peptide bond cleavage.[11]

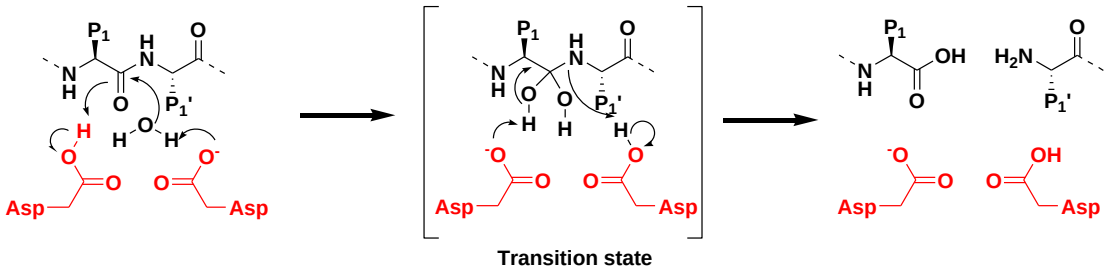


Figure 1.2 Catalytic mechanism of aspartic proteases.

1.2.2 Inhibitor Development for Aspartic Proteases

Aspartic proteases generally bind 6-10 amino acid regions of their peptide substrates which are typically cleaved with the aid of two catalytic aspartic acid residues in the active site.[12] Thus far, most of aspartic protease inhibitors are transition state analogues which mimic the transition state without becoming hydrolyzed. It has been proved that transition state analogue inhibitors are typically more efficient than substrate analogue inhibitors since the transition state is bound with a much higher affinity. In the last decade, many inhibitor scaffolds have been developed based on the tetrahedral intermediate of hydrolysis of aspartic protease substrates (Figure. 1.3). In this study, inhibitors with hydroxyethylamine and dihydroxylethelene have been mainly emphasized.

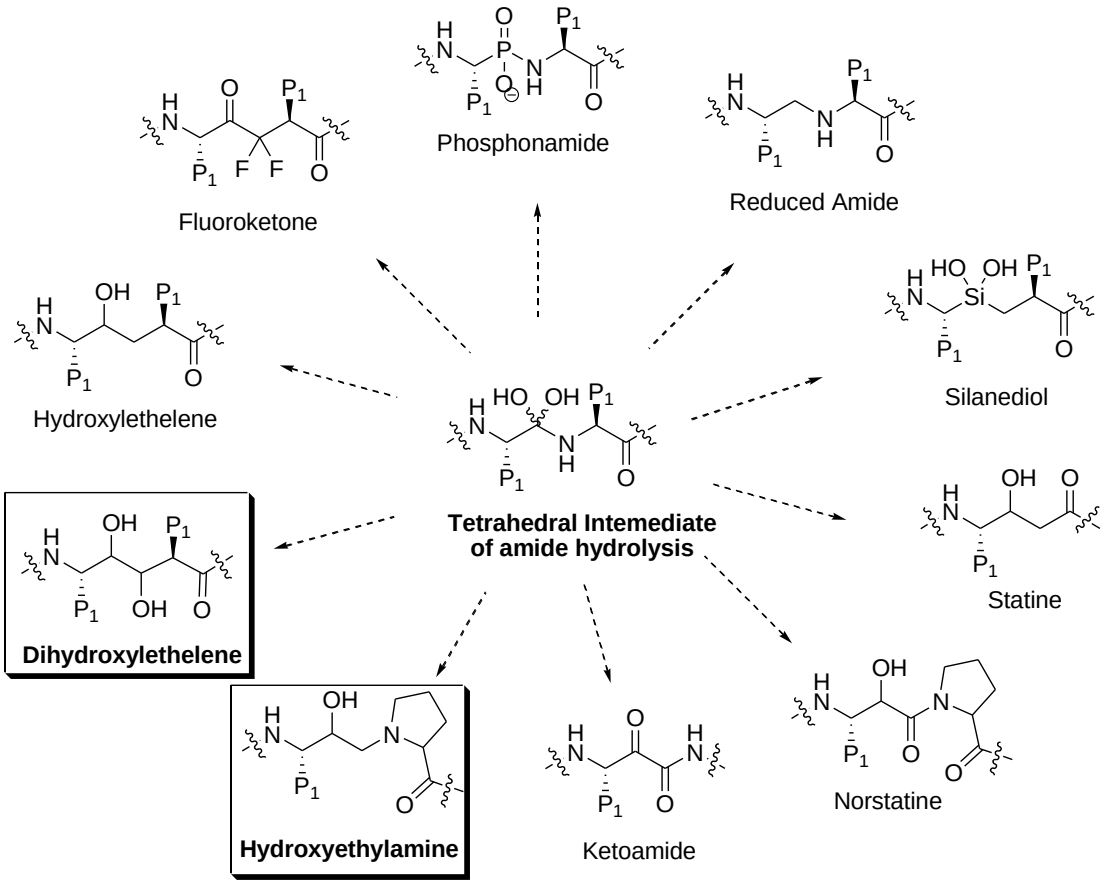


Figure 1.3 Some well-known transition state scaffolds for aspartic protease inhibitors.

Pepstatin A, a natural product first isolated in 1970, was first found to be a potent inhibitor of pepsin (K_i = 56 pM).[13] It was subsequently shown that pepstatin A is a generic aspartic protease inhibitor.[14] While these hydroxyethylene-based inhibitors in general lead to potent enzymatic inhibition, their use as therapeutic agents is, however, often hampered by their unfavorable biopharmaceutical properties due to their peptidic character. The protease of the human immunodeficiency virus (HIV-1 protease) has proved to be an attractive drug target due to its essential role in the replicative cycle of HIV. So far, nine HIV protease inhibitor drugs have been approved by FDA and are clinically available: Saquinavir, Nelfinavir, Ritonavir, Lopinavir, Indinavir, Amprenavir, Fosamprenavir, Atazanavir, and Tipranavir.

Among them, Amprenavir, a hyrdroxyethylamine transition state analogure, is one of the most potent HIV-1 inhibitors with an IC_{50} value of 0.6 nM.[15] In addition, some potent inhibitors for other aspartic proteases (Renin, cathepsin D, β, ɤ-secretase .etc) have also been developed. Some of them have become drugs for treatment of diseases.[5a]

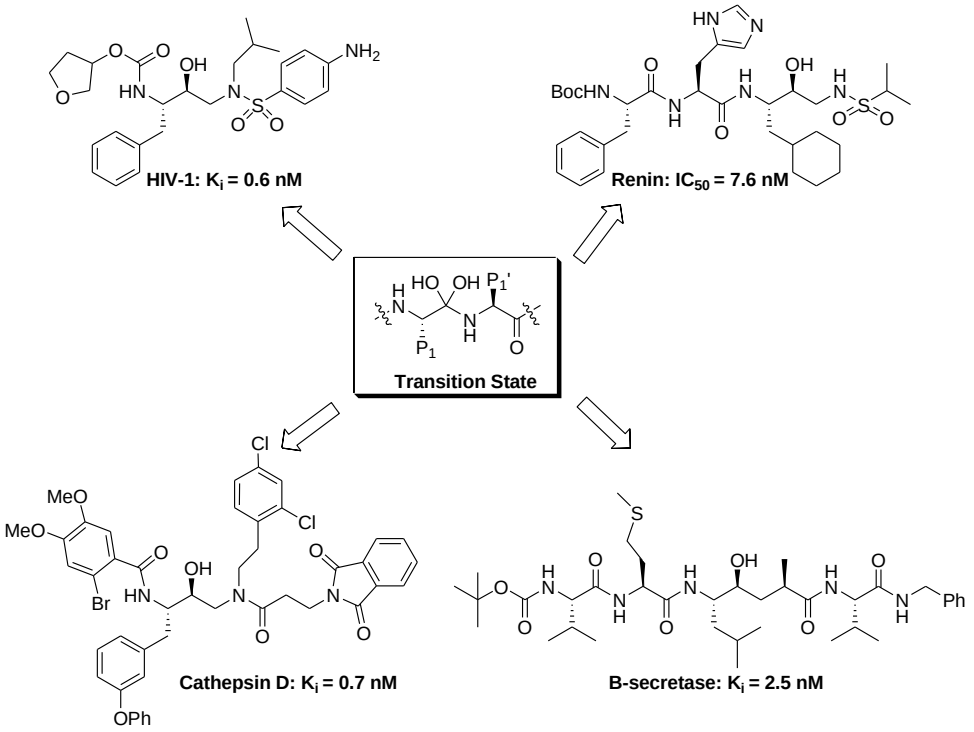


Figure 1.4 Some representative aspartic protease inhibitors.

1.3 Protein Kinases as Therapeutic Targets

Protein kinases (PKs) play a key role in signal transduction pathways, which regulate a varity of cellular processes including growth, division, differentiation and metabolism. Dysregulation of cellular kinase activities has been implicated in many diseases (including cancer, HIV, malaria, and so on). Thus, protein kinases have emerged as a major class of drug targets in recent years.[16,17] To date, approximately 30 distinct kinase targets have been developed to the level of a phase I clinical trial,

and around 518 protein kinases are encoded in human genome. Most of them share a catalytic domain conserved in sequence and structure but are notably different in how their catalysis is regulated.[18] The majority of kinase targets are being investigated for the treatment of cancer. However, dysregulation of kinase function has also been implicated in other disorders, including immunological, neurological, metabolic and infectious diseases.[18]

1.3.1 Catalytic Mechanism and Classification of Protein Kinases

Protein kinases are ATP-dependent phospho-transferases that modified other proteins by chemically adding phosphate groups to them. They require an essential divalent metal ion, usually Mg^{2+}, to facilitate the phosphoryl transfer reaction and assist in ATP binding. Classical protein kinases have a catalytic domain (~250 amino acids), which is constituted by a small N-terminal lobe of β-sheets and a larger C-terminal lobe of α-helices. ATP first binds in a cleft between the two lobes so that the adenosine moiety is buried in a hydrophobic pocket with the phosphate backbone orientated outwards towards the solution. Subsequently, the protein substrate binds along the cleft and a set of conserved residues within the kinase catalytic domain catalyse the transfer of the terminal γ-phosphate of ATP to the hydroxyl oxygen of the Ser, Thr or Tyr residue of the substrate (Figure 1.5). Although all protein kinases share a common mechanism, some of them are different because of the charge and hydrophobicity of surface residues. Certainly, the order of the steps also differs for different kinases. For example, some kinases bind to their protein substrates before binding ATP and others release ADP before releasing the protein substrate.[19]

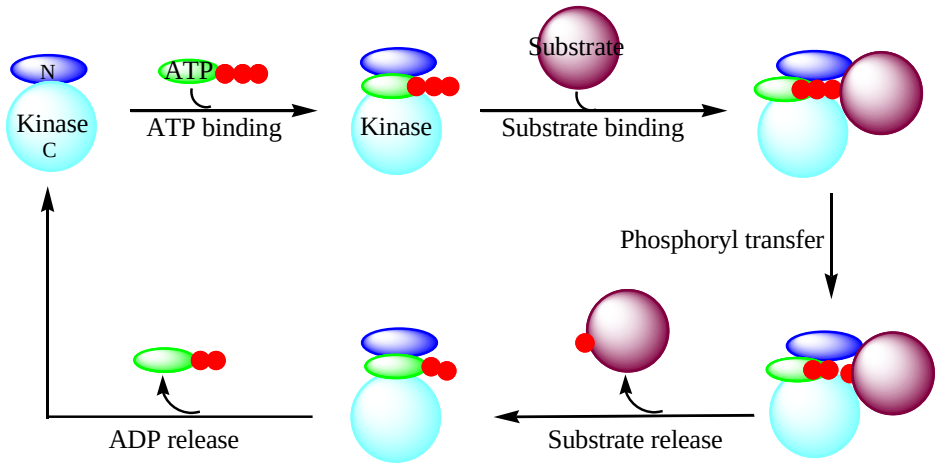


Figure 1.5 A common mechanism for protein kinases.

In human genome about 518 human protein kinases have been identified (constituting about 1.7% of human genes).[17, 20] Within this large protein kinase family, at least 90 tyrosine kinase genes have been identified [58 receptor tyrosine kinases (RTKs) and 32 nonreceptor tyrosine kinases (NRTKs)]. Thus far, a large number of tyrosine kinases (both receptor and non-receptor types) are associated with cancer. Clinical studies suggest that overexpression/deregulation of tyrosine kinases may cause great effects on patients.

1.3.2 Development of Protein Kinase Inhibitors

Small molecule inhibitors of protein kinases are widely used in signal transduction research and constitute a novel class of drugs for therapeutic intervention in a variety of human diseases.[21] To date, approximately 80 inhibitors have been advanced to some stages of clinical evaluation. Most currently known kinase inhibitors discovered are ATP competitive and present one to three hydrogen bonds to the amino acids located in the hinge region of the target kinases, thereby mimicking the hydrogen bonds that are normally formed by the adenine ring of ATP. Generally,

the kinase inhibitors are classified into several types based on different interactions between inhibitors and protein kinases.[22] The most popular type is ATP-competitive inhibitors which only recognize the active conformation (DFG-In) of kinases. Dasatinib (BMS-354825) was identified as a highly potent, ATP-competitive inhibitor of Src and Abl kinases.[23] Dasatinib was also found to inhibit a wide variety of other kinases. Studies have shown the activity of dasatinib against Bcr-Abl-positive leukemic cell lines as well as epithelial tumor cell lines including human prostate and breast cancer cells.[23b,24] As a result, this compound has become a promising therapeutic agent for preventing cancer cell growth. By contrast, the type II inhibitors recognize the inactive conformation (DFG-Out). For example, the phenylaminopyrimidine compound imatinib mesylate (Gleevec, STI571) was the first tyrosine kinase inhibitor approved by FDA, and this drug has been successfully used in the treatment of chronic myeloid leukemia (CML), whereas resistance formation and subsequent therapy failure was observed in patients with advanced CML.[25] The third class of inhibitors binds outside the ATP-binding site at an allosteric site. The most well characterized allosteric kinase inhibitor is CI-1040, which inhibits MeK1 and MeK2 by occupying a pocket adjacent to the ATP binding site.[26] The fourth class of kinase inhibitors is capable of forming an irreversible, covalent bond to the kinase active site, most frequently by reacting with a nucleophilic cysteine residue.[27] Clinically, the classical irreversible kinase inhibitors of epidermal growth factor receptor (EGFR) is HKI-272 and CL-387785.[28] Thus far, most small molecule inhibitors are primarily developed from a combination of methods, including

high-throughput screening using biochemical or cellular assays, analogue synthesis, structure-guided design and fragment-based assembly strategies.

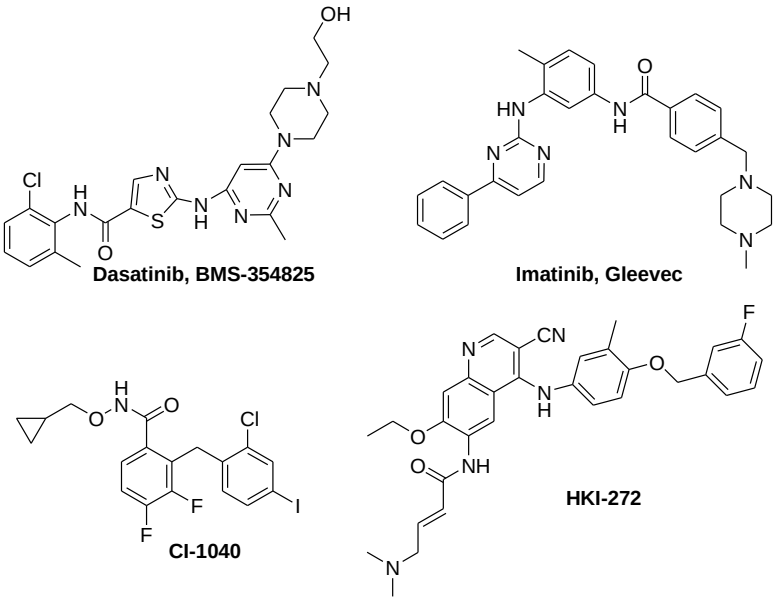


Figure 1.6 Several representative kinase inhibitors.

1.4 High-throughput Amenable Chemistry

The term "Catalomics" is used to define an emerging "-omics" field in which high-throughput studies of enzymes and other biocatalysts are carried out by using advanced chemical biology/proteomics approaches (Figure 1.7).[29] The progress in catalomics has made a very important effect on drug discovery, and it heavily relies on both powerful assay technologies such as microarrays and recent developments in some of the most reliable and robust chemical reactions. Thus far, one of the main challenges in the field of Catalomics is the development of high-throughput (HT) amenable chemical reactions that allow rapid synthesis of diverse chemical libraries for the interrogation of different classes of enzymes. Herein, we have rapidly and efficiently synthesized the enzyme inhibitor libraries by using high-throughput amenable reactions (Click chemistry and solid-phase) due to their near-perfect,

modular, robust and biocompatible nature.

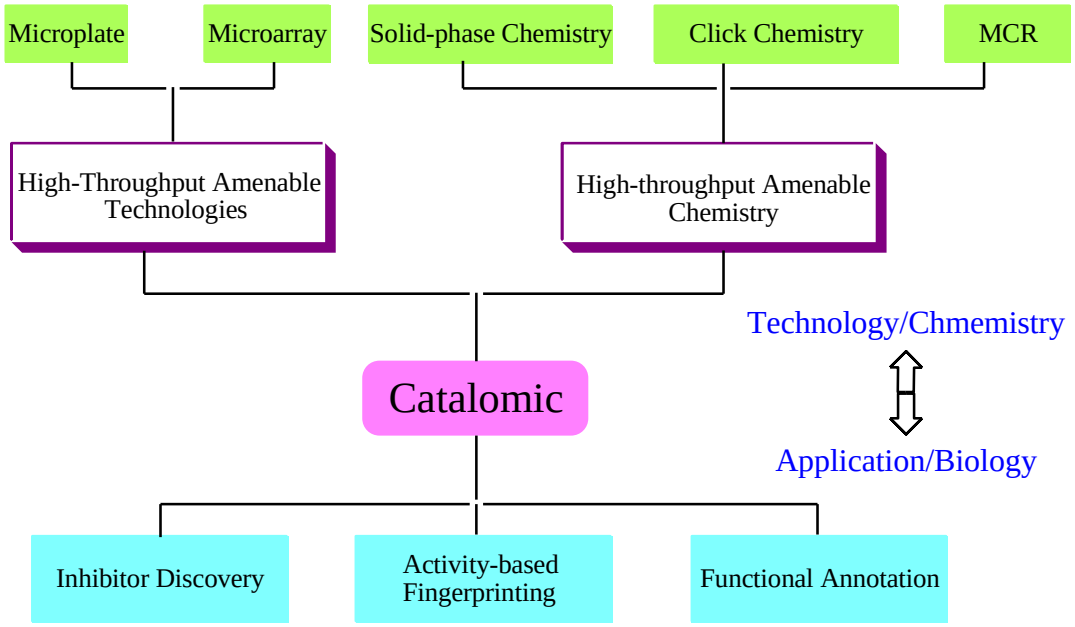


Figure 1.7 Overview of "Catalomics".

1.4.1 Solid-phase Chemistry

Combinatorial chemistry has emerged as a powerful new technology for chemists to synthesize a large number of compounds for biological evaluation. One of the techniques used in combinatorial chemistry is solid-phase organic synthesis. The first landmark paper on solid-phase synthesis of tetrapeptide was published by Merrifield in 1963,[30] and Ellman was the first to report the solid-phase combinatorial synthesis of non-peptide based small molecules, benzodiazepines.[31] Solid-phase synthesis has several advantages over solution-phase synthesis, which are listed below:

(1) Can be automated easily and amenable to high-throughput synthesis

(2) Isolation of the compound is very simple, usually by simple filtration

(3) High product purity

(4) Easier to generate a large library of compounds in a shorter time

(5) In some cases, the reaction can be forced to completion by using excess of reagent

(6) In some cases (especially peptide synthesis), solubility problems are minimized However, the main limitations of solid-phase synthesis are difficulty in monitoring the reaction, usage of excess reagents and solvents, difficulty in multistep stereo-selective synthesis.

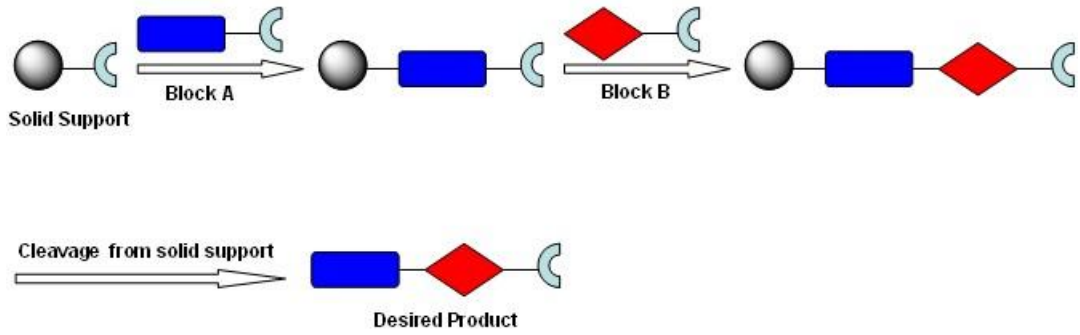


Figure 1.8 General principle of solid-phase synthesis.

We have adopted a traceless solid-phase methodology to synthesize an alkyne sub-library for assembly of aspartic protease inhibitors in our study. Traceless synthesis can be defined as a synthetic route which yields compounds composed of inherent atoms only from the particular target molecules.[32] Synthesis using solid-phase traceless resins leaves no residues on the final compound after the cleavage from the solid support. We have explored the solid-phase amide bond formation reaction using an aldehyde traceless resin (PL-FMP),[33] and a 19-member alkyne building block library was efficiently generated.

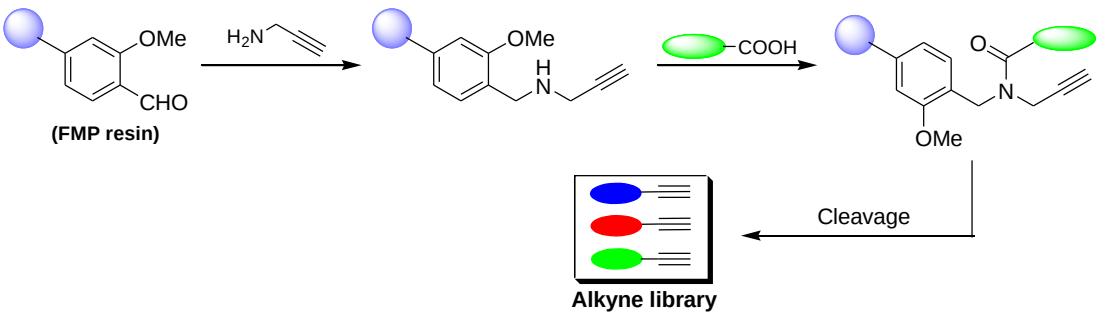


Figure 1.9 Traceless synthesis of alkyne building blocks.

1.4.2 Click Chemistry

A set of chemical reactions, known as bioorthogonal reactions, that are orthogonal to most functional groups in biological systems have so far shown promising use in biological research. Similarly, click chemistry, defined by Sharpless,[34] has been identified to have minimum synthetic demands, a wide range of solvent compatibilities and high yields. Of the different click reactions, the Cu (I) catalyzed version of Huisgen 1, 3-dipolar cycloaddition reaction between azides and terminal alkynes, discovered independently by Meldal[35] and Sharpless,[36] has emerged as a powerful tool in chemical biology and proteomic applications. This reaction is characterized by its high chemo-selectivity, modularity, near-perfect yield and biocompatibility in the aqueous conditions. As a result, click chemistry has become an attractive tool in many research fields ranging from materials sciences, biology to medicinal chemistry/chemical biology. It also has emerged as an integral part of the drug discovery pipeline by providing a high-throughput amenable chemical reaction platform for compound synthesis.

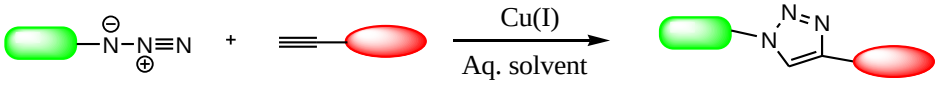


Figure 1.10 Cu (I) catalyzed azide-alkyne ligation.

In addition, “Click” reaction has some other features: (1) highly exothermic reaction (ΔH = -45 to -55 Kcal/mol); (2) high kinetic barrier–due to the high energy content of the azide and alkyne (25 kcal/mol); (3) in the presence of Cu (I) the reaction is not significantly affected by the steric and electronic properties of the groups attached to the azide and alkyne reactive centers; (4) the reaction is unaffected by water and other functional group, the rate of Cu (I) catalyzed is 10^7 times faster than the uncatalyzed one; (5) triazole ring is very stable, inert to oxidation, hydrolysis, and reduction conditions even at high temperature; (6) strong dipole moment, aromatic, good H-bonding acceptor. [37]

Development of enzyme inhibitors is one of most important areas where click chemistry plays an active role. It has been identified as a convenient strategy towards fragment-based inhibitor assembly, where large libraries of potential bidentate inhibitors are generated with minimum synthetic efforts. This approach is powerful especially against protein targets which possess multiple binding pockets in their active sites (e.g., Proteases, Phosphatases, and Kinases, etc). Currently, three methods have mainly been used to assist the assembly of compounds: (1) the tethering method developed by Wells *et al*[38] (2) the NMR-based SAR strategy[39] and (3) the *in situ* screening approach developed by Wong and co-workers[40] which was largely based on the “click” chemistry pioneered by Sharpless *et al*. Among them, the click chemistry approach is highly versatile due to its several advantages: (1) water compatible or non-toxic solvent should be used for the reaction; (2) reactions are usually carried out at microscale level; (3) products no need to be isolated or purified; (4) reactions have

very high yields; (5) no protection group required.

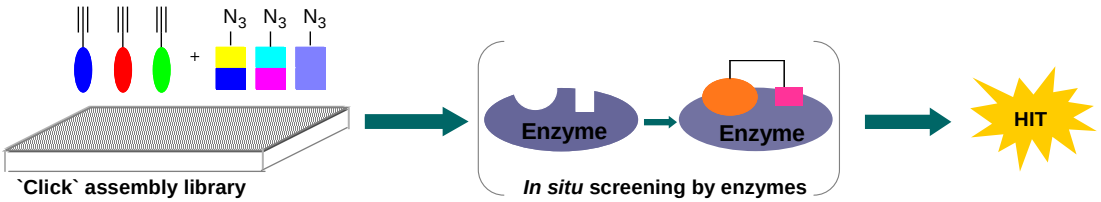


Figure 1.11 **"**Click**"** assembly and *in situ* screening.

Thus far, many groups including ours have utilized this powerful tool for the successful discovery of inhibitors against plasmepsins,[41] PTP1B,[42] matrix metalloproteases (MMPs),[43] caspases,[44] c-Abl,[45] HIV protease,[46] SARS 3CL protease,[47] α-1,3-fucosyltransferase,[48] acetylcholinesterase[49] and sulfotransferase[50] (Figure 1.12)

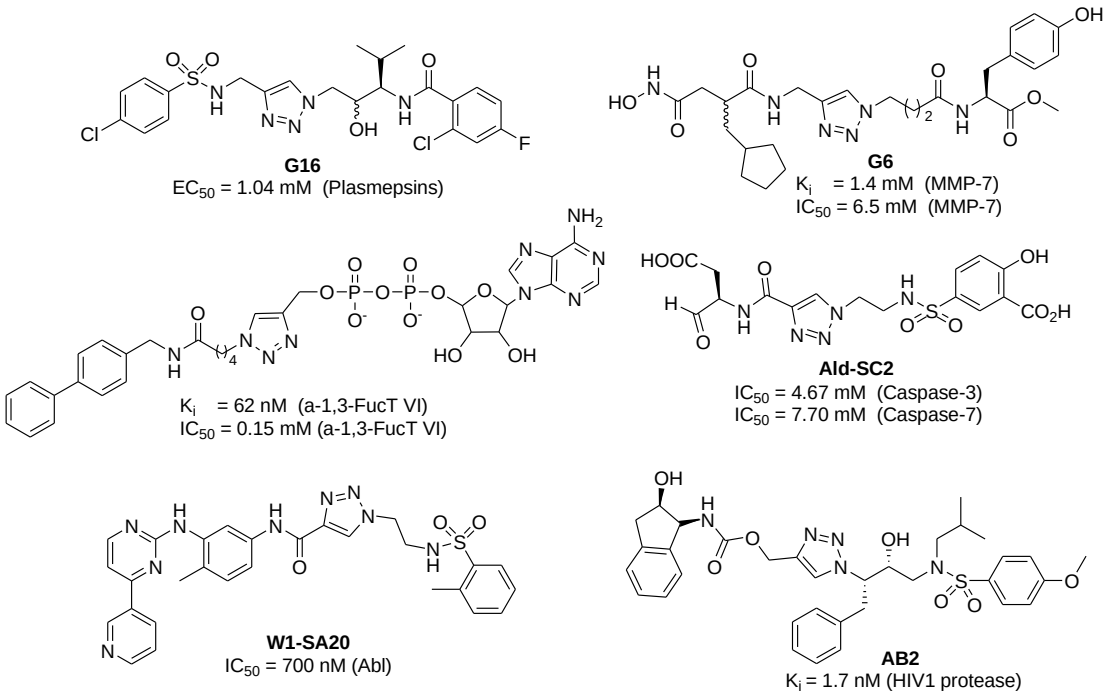


Figure 1.12 "Click" inhibitors targeting various enzyme classes.

Amide-bond formation reaction between an acid and an amine is also a very useful reaction and can be amenable to high-throughput synthesis and *in situ* screening except for "Click" chemistry[40] (Figure 1.13a). The most common coupling reagents used in this reaction is DCC/ EDC, HATU, HBTU, PyBop, etc. This efficient

reaction has recently been applied successfully in the rapid discovery of inhibitors against a number of enzymes, such as HIV protease,[41,51] cysteine proteases,[52] α-fucosidases,[53] SARS-3CL protease[54] and β-aryl sulfotransferase[55] (Figure 1.13b). This high-throughput amenable reaction has some key features, in that quantitative product formation can be achieved using powerful acylating/coupling reagents. Therefore, *in situ* biological screening may be carried out directly without product purification. However, the method is severely limited by the presence of excessive coupling reagents, byproduct and excessive starting materials in the reaction, which may lead to false positive results.[56]

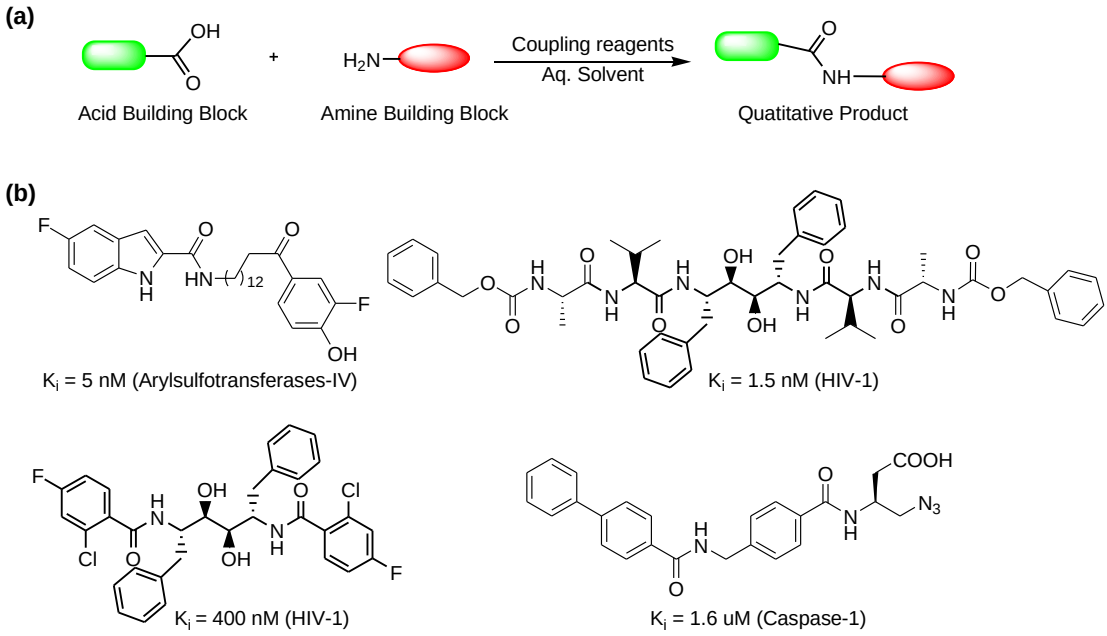


Figure 1.13 Traditional amide-bond formation reaction and representative inhibitors. (a) Amide-bond formation reaction. (b) Representative inhibitors assembled by amide-bond formation reaction.

1.4.3 Small Molecule Microarray

Over the last decade, microarray technology has become one of most popular high-throughput screening platforms for wide application from expression profiling,

mapping interaction networks to molecular fingerprinting and ligand discovery. Microarray technology was originally reported as peptide microarray by Fodor *et al*. in 1991[57] and later adopted by Affymetrix and the Brown group to create DNA microarray.[58] Subsequently, MacBeath *et al* developed small-molecule and protein microarrays for the first time.[59] Thus far, four key microarray-based platforms have been successfully developed for high-throughput study of enzymes. Those are protein-, peptide-, small molecule- and cell-based microarrays, respectively. Among them, small molecules and protein microarrays have witnessed great growth in recent years with significant technical and conceptual improvements.

Small molecule microarrays (SMMs) provide a valuable platform for the large-scale, quantitative determination of protein-ligand interactions in high-throughput. It offers a cheap and convenient method for the screening of thousands of compounds rapidly and has been successfully used in ligand identification, and in protein profiling. Furthermore, microarray-based technologies have also been used in the discovery of biosensors and biomarkers for disease diagnostics. The first small-molecule microarray (SMM) was developed in 1999 by MacBeath and Schreiber who showed that thousands of binding assays could be performed in parallel from molecules synthesized on individual beads.[59] Besides, it was shown that the SMM can also be used to study the weak binding pairs, such as FKBP12 and its ligand (K_d values in the μM range). Our group identified four novel ligands for human IgG from a library of 2,688 triazine compounds using SMM. The best hits showed the K_d values of 2.02 μM. In fact, immobilization strategies have

become critical for SMM.[60] In the last decade, great effort has been put on development of immobilization chemistry. For example, MacBeath and coworkers reported the Michael reaction in 1999.[59a] Schreiber group developed the reactions between a primary alcohol and silyl chloride, a phenolic derivative and diazobenzylidene in 2000 and 2003, respectively.[59b] Recently, Waldmann *et al*. and Raines *et al*. have independently developed methods based on the staudinger reaction that allow efficient reaction under mild conditions in the presence of water and oxygen.[61]

1.5 Activity-based Protein Profiling (ABPP)

Activity-based protein profiling (ABPP), developed by Cravatt and coworkers,[62] is a novel technique that is designed to address the proteome at the level of discrete enzyme families, providing a way to distinguish, for example, active enzymes from their inactive zymogen[63] and/or inhibitor-bound forms.[64] probes typically used in ABPP possess three general elements: (i) a binding group that promotes interactions with the active sites of specific classes of enzymes, (ii) a reactive group that covalently labels these active sites, and (iii) a reporter group (e.g., fluorophore or biotin) for the visualization and affinity purification of probe-labeled enzymes. ABPP has recently been shown to be a powerful tool in proteomic studies, which takes advantage of mechanism-based probes to record variations in the activity of enzymes in whole proteomes. To date, many ABPP probes have been developed for relevant enzyme classes, including serine hydrolases,[65,66] oxidoreductases[67] and cysteine proteases.[68,69] Notably, these probes have been shown to selectively label active

enzymes but not their inactive precursor or inhibitor-bound forms. Therefore, the current activity-based profiling strategies are limited in that they are applicable only to a few classes of enzymes which possess covalent intermediates in their catalytic mechanism.

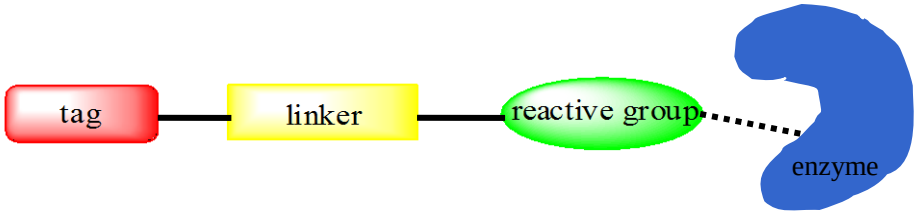


Figure 1.14 Structures of activity-based probes.

1.6 Affinity-based Protein Profiling (A*f*BP)

Affinity-based protein profiling (A*f*BP) of enzymes is a complementary strategy for activity-based proteomic profiling without the need of mechanism-based suicide inhibitors. The affinity unit comprises a known reversible inhibitor that binds tightly in a non-covalent manner to the active site of the target enzyme (or a specific class of target enzymes). Because the enzyme-probe interaction is mainly based on affinity, in order to promote selective binding and modification of proteins's active sites, an additional moiety, the photo-labile group, is used to covalently link the probe to the enzymes. The incorporation of a fluorescent tag eventually results in a trifunctional affinity-based probe for potential large-scale protein profiling experiments. Two photolabile groups, namely diazirine and benzophenone, have usually been used to covalently modify molecules in a variety of biological experiments. These photoactive labels operate by generating reactive intermediates such as carbenes, nitrenes, and ketyl biradicals and inserting them into the target molecule to form a

permanent linkage. The selected wavelength for UV irradiation is usually greater than 300 nm, thus preventing potential photochemically induced damages to the enzyme.[70] Our group recently reported some affinity-based probes which targeted other class of proteins, for example, aspartic proteases[41] and protein kinases.[45]

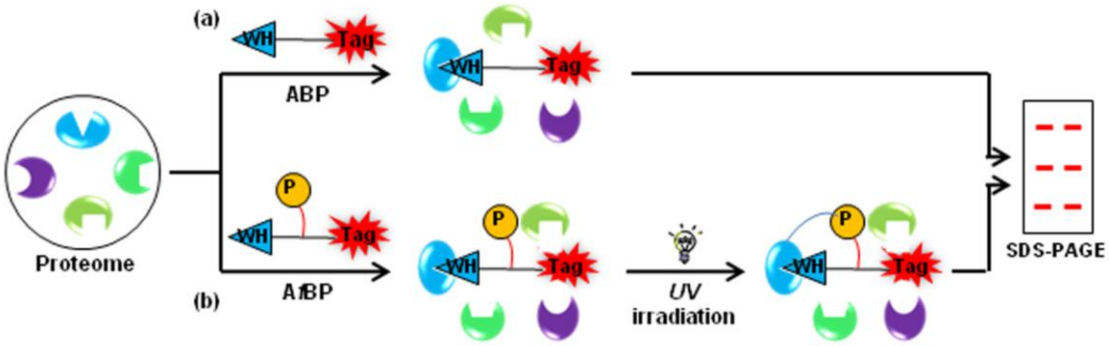


Figure 1.15 Proteome Profiling with activity-based probes (**ABPs**) (a) and affinity-based probes (**A*f*BPs**) (b).

1.7 Research Objectives

The aim of this thesis is to develop high-throughput amenable strategies for chemically profiling of aspartic proteases as well as protein kinases. The high-throughput synthetic and *in situ* screening strategy we developed may offer a useful method for drug discovery of diseases. It may also provide guidelines for rapid synthesis and validation of drugs in future. Additionally, our small molecule microarray (SMM)-facilitated strategy for the discovery of new A*f*BPs should provides a useful tool for high-throughput developments of novel small molecule probes and identification of new aspartic proteases as well as related biomarkers in future.

Chapter 2

Expedient Solid-phase Synthesis of Both Symmetric and Asymmetric Diol Libraries Targeting Aspartic Proteases

2.1 Summary

C_2-symmetric diols have been shown to be highly potent against HIV-1 protease (PR). However, gaining access to these compounds has been hampered by the need of multistep solution-phase reactions which are often tedious and inefficient. In this chapter we describe a solid-phase amide-forming approach for rapid assembly of both symmetric and asymmetric diol-containing small molecules as potential aspartic protease inhibitors. One potent and selective inhibitor of HIV-1 PR, **SYM-5** (K_i = 400 nM), was discovered which should possess better bioavailability properties (fewer amide bonds, lower MW) than its parental compounds. The clear advantage in our strategy lies in its feasibility for solid phase synthesis of asymmetric diols. When compared with other potent HIV-1 protease inhibitors, our compound pales in terms of its inhibitory activity.

2.2 Introduction

A key challenge in current drug discovery is the development of high-throughput (HT) amenable chemical reactions that allow rapid synthesis of diverse chemical libraries of enzyme inhibitors. In recent years, one reaction that has received much attention is the amide-bond formation between an amine and a carboxylic acid using

suitable activating/coupling reagents.[56] The reaction is highly efficient in that it often generates desired products in nearly quantitative yields, thus allowing direct *in situ* biological screening to be carried out. This strategy has been successfully adopted, in both solution- and solid-phase, by several groups for the rapid discovery of small molecule inhibitors.[51-55,71] The solution phase amide-forming strategy is comparatively easier to implement but has the following problems: (1) undesired chemicals (starting materials, reagents, by-products, etc) are present during the biological screening and therefore often cause false results; (2) multi-step reactions can not be carried out without proper purification of the intermediates; and (3) excessive starting materials/reagents normally cannot be used to drive an otherwise-difficult coupling to completion. Solid-phase synthesis could potentially solve many of these problems, and has been widely adopted in the synthesis of biopolymers (e.g., oligonucleotides and peptides) as well as small molecules (e.g., in combinatorial chemistry). In most cases, however, low-quality products are generated and need to be laboriously purified before screening. We have recently become interested in developing highly efficient solid-phase strategies that enable large-scale synthesis of high-quality compound collections which are suitable for direct *in situ* biological screening.[72] With these new chemical tools, we have successfully developed small molecule inhibitors targeting different classes of proteases, including cysteine and metallo-proteases. Herein, by making use of the amide-forming reaction, we report the solid-phase synthesis of both symmetric and asymmetric diols as putative inhibitors of aspartic proteases. The key advantages of this method include (i)

chemical modification of the library building blocks (i.e., diol core & acids) is unnecessary; (ii) it is solid-phase, enabling a large library to be constructed efficiently; (iii) it is robust, giving high-quality products.

Many aspartic proteases are known therapeutic targets.[72,73] For example, HIV-1 protease is one of the main targets of AIDS.[72] Renin is currently being used to treat hypertension.[73a] Cathepsin D is involved in breast cancer metastasis.[73b] β-Secretase is related to Alzheimer disease.[73c] Plasmepsins, key aspartic proteases from malaria involved in the hemoglobin degrading pathway in parasite-infected macrophages, are being pursued to treat malaria.[73d]

HIV-1 protease (PR) is one of the best known aspartic proteases, and an attractive target for the treatment of AIDS. The enzyme consists of two identical, non-covalently associated subunits of 99 amino acid residues formed in an C_2-symmetric fashion. The highly symmetrical active site is formed at the dimer interface. Over the years, many HIV-1 protease inhibitors (PI) have been developed.[72a] Two of the most effective PI are shown in Figure 2.1 (top two): they contain a C_2-symmetric, diol-containing moiety which mimics the transition state of amide hydrolysis. Wong's peptidic inhibitor developed in 1998,[72b] for example, has a K_i of 1.5 nM against HIV-1 PR. Its poor bioavailability (e.g., peptide-based, large MW) however makes it ineffective in cell-based assay. Abbott's Ritonavir,[72c] an FDA-approved drug and improved version of Wong's inhibitor, is nearly C_2-symmetric, and much more effective against AIDS due to its smaller size (MW = 820) and fewer amide bonds. The rapid emergence of drug-resistant HIV-1 PR,

however, has rendered many of these PI ineffective.

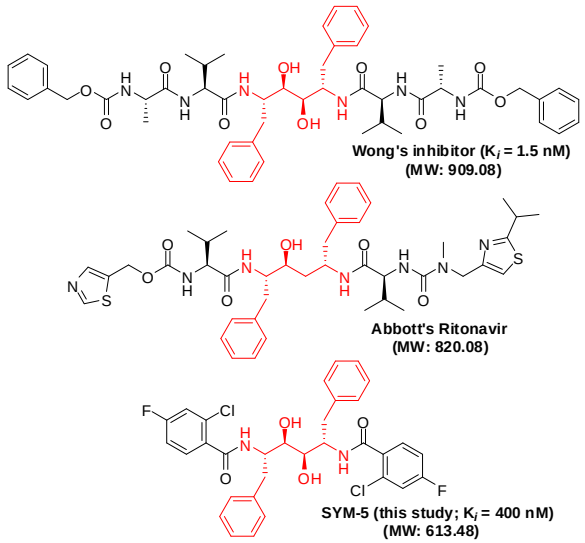
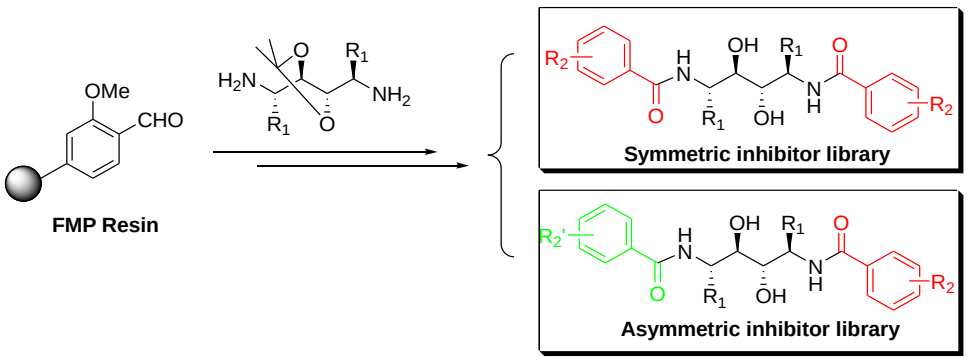


Figure 2.1 Structures of two known diol-containing HIV-1 protease inhibitors (top two) and the inhibitor identified from this study (bottom).

Thus, there is an urgent need to develop chemistry that permits rapid and efficient synthesis of new PI. Another paradigm shift in HIV PI research is the introduction of asymmetric PI which could be more effective against some of the drug-resistant HIV strains.[72d] In the current study, we have developed a solid-phase strategy for rapid synthesis of potential symmetric and asymmetric PI (Schemes 2.1 and 2.2). As a proof of concept, the strategy was successfully used to synthesize 75 diol-containing compounds, which, upon direct *in situ* screening, revealed a small (MW = 613) and potent HIV-1 PI (K_i = 400 nM).

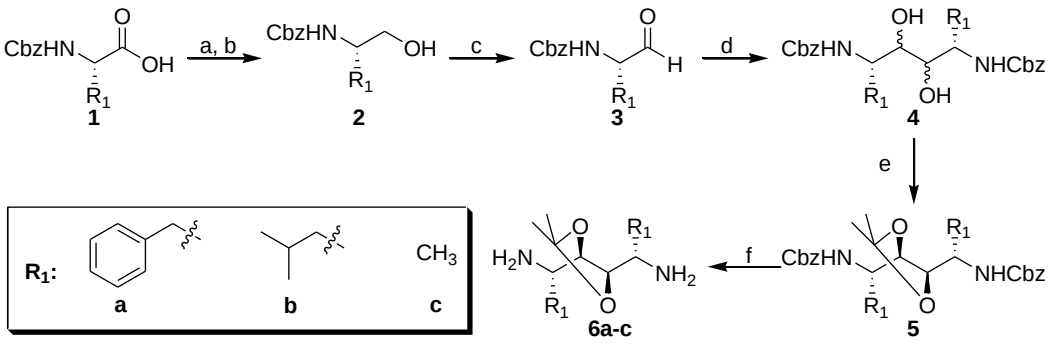


Scheme 2.1 Traceless synthesis of symmetric and asymmetric inhibitors targeting aspartic proteases.

2.3 Results and Discussion

2.3.1 Solution-phase Synthesis of Diamino Diols

In our strategy, we adopted a solution-cum-solid phase strategy, in which the most important component of the inhibitors-the diamino diol core group, was synthesized in solution (Scheme 2.2) and purified to homogeneity before being installed onto the commercially available 4-formyl-3-methoxyphenoxy (FMP) resin,[74] followed by diversification with a variety of acids (Scheme 2.3). As shown in Scheme 2.2, following previously published procedures,[72c] three different C_2-symmetric diaminodiols, **6a-c**, were made from commercially available Cbz-protected amino acids **1**. Upon reduction, the corresponding amino alcohols **2** was oxidized to the aldehydes **3** using Dess-Martin reagent. Subsequently, pinacol homocoupling reaction between two equivalents of **3** in the presence of VCl_3 afforded diastereomeric diols **4** in a single-step transformation (~20% yield). Subsequent protection of the dihydroxyl group with 2, 2-dimethoxypropane, flash chromatography to remove minor diastereomeric impurities, and deprotection of the Cbz groups using H_2 (in Pd/C) gave the resulting enantiomerically pure diaminodiols, **6a-c**, in moderate yield (50% in 2 steps).



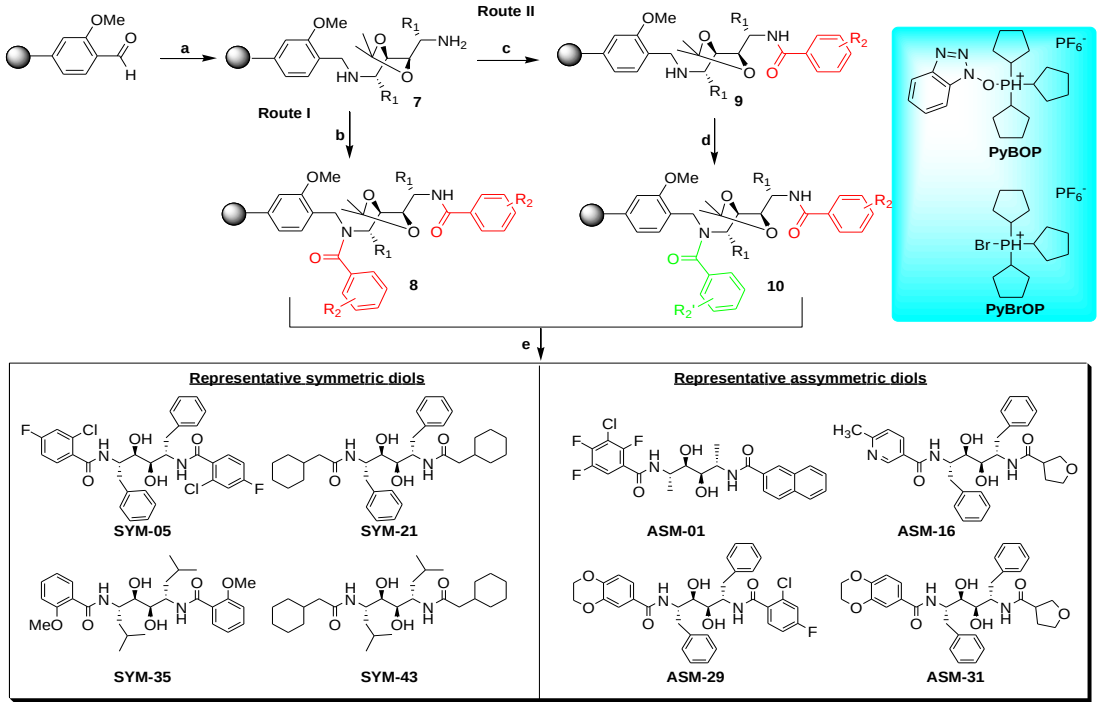
Scheme 2.2 Solution-phase synthesis of the diol warheads. (a) Isochloroformate,

NMM, DCE, -15 ^{o}C. (b) $NaBH_4$, THF, 0 ^{o}C, 70-77%. (c) Dess-Martin, DCM, r.t., 60-70%. (d) VCl_3, 1,3-dimethylimidazole, Zn, dry THF, reflux, 10-20%. (e) 2, 2-dimethoxypropane, *p*-TsOH, acetone, r.t., 60-70%. (f) H_2, Pd/C, MeOH/EA, r.t., 60-70%.

2.3.2 Traceless Synthesis of Symmetric and Asymmetric Inhibitor Library

To start the library assembly on solid phase, the aldehyde-containing FMP resin was used for capturing of **6** by reductive amination, followed by amide bond-forming reaction with acids. First, the three diaminodiols, **6a-c**, were treated with FMP resin in the presence of $Na(OAc)_3BH$/2% glacial acetic acid in DCE to give the resin-bound amines **7**, which contain an 1° and 2° amine each. Initially, simultaneous acylation of both the 1° and 2° amines with the same acid building blocks was met with some difficulties under most coupling conditions (e.g., DIC/HBTU/TBTU/HATU); in most cases, only the monoacylated product (at 1° amine position) was generated. This was clearly due to the chemical and steric difference of the two amines in **7**. We therefore took advantage of this to make both symmetric and asymmetric diols (Scheme 2.3). An exhaustive testing of (7 + acid) coupling under a variety of different coupling conditions finally gave rise to the following two sets of optimized conditions: (I) to make symmetric diols, Route I was used in which PyBrOP/ DIEA coupling with the acid (10 equiv) was carried out to gave **8**; (II) to make asymmetric diols, Route II was used in which PyBOP/HOAt coupling with the first acid (5 equiv) was performed, giving **9**, followed by PyBrOP/DIEA coupling with the second acid (5 equiv) to give 10. Finally, cleavage of the products from the resin using an optimized TFA cocktail (5:4:1 TFA/DCM/water) gave a total of 44 C_2-symmetric and 31 asymmetric diol inhibitors (representative compounds are shown in Scheme 2.3). To ensure the crude

products generated from our strategy were sufficiently pure for direct *in situ* screening, the compounds were further characterized by LCMS and NMR; in most cases the desired products were obtained with good purity (>90% and 60 – 95% for symmetric and asymmetric diols, respectively).



Scheme 2.3 Solid-phase synthesis of both symmetric and asymmetric diols. (a) $Na(OAc)_3BH$, 2% AcOH in DCE, r.t. (b) acid (10 eq), PyBrOP/DIEA in DMF, r.t. (c) acid 1 (5 eq), PyBOP/HOAt/DIEA in DMF, r.t. (d) acid 2 (5 eq), PyBrOP/DIEA in DMF, r.t. (e) TFA/DCM/H_2O (5/4/1), r.t.

2.3.3 Inhibitor Fingerprinting

Next, the inhibitory activity of these 75 diol-based inhibitors was determined against HIV-1 PR, plasmepsin I (PM I) and plasmepsin II (PM II) using a standard fluorescence microplate assay method. The activity of the aspartic proteases was determined by measuring the rate of hydrolysis from the commercial available fluorogenic substrates (AnaSpec, USA). The inhibition of all proteins were screened using black polypropylene flat-bottom 384-well microtiter plates (Greiner, Germany)

in a total reaction volume of 20 μl/well, monitored with a Tecan Infinite F200 microplate reader at excitation of 390 nm and emission of 492 nm prior to the addition of substrate. The results were shown in the heat maps below Figure 2.2. Eventually, an inhibitor fingerprint of the library against the aspartic proteases was obtained, from which six potential hits (SYM-5, -21, -35, -43 and ASM-16, -29) were identified.

2.3.4 IC_{50} Measurements of Selected Inhibitors

Besides, we also determined some representative compounds' IC_{50} values for HIV-1, plasmepsin I/II and pepsin based on Figure 2.2. As shown in Figure 2.5, compounds **SYM-05**, SYM-35 were identified to be relatively potent against HIV-1 PR and not against plasmepsin I, plasmepsin II and pepsin. Compound SYM-21 were identified to be relatively potent against both plasmepsin I and II but not against HIV-1 PR and pepsin. However, SYM-24, SYM-43 were found to specifically inhibit plasmepsin I but not plasmepsin II. Interestingly, ASM-16 and ASM-29 were against plasmepsin II but not against plasmepsin I. Detailed inhibition studies were then carried out to obtain the corresponding IC_{50} of these compounds, and the results are summarized in Table 2.1. The best inhibitor against HIV-1 PR was found to be **SYM-5**, with IC_{50} value of 395.4 nM. Significantly, it showed a very poor inhibition against PM I and PM II (>25 μM), indicating that this inhibitor is highly specific only towards HIV-1 PR. None of the asymmetric compounds was able to inhibit 50% of the HIV-1 PR activity at a concentration of 25 μM, which is expected since the assay was done with the C_2-symmetric wild type HIV-1 PR. However, two of these

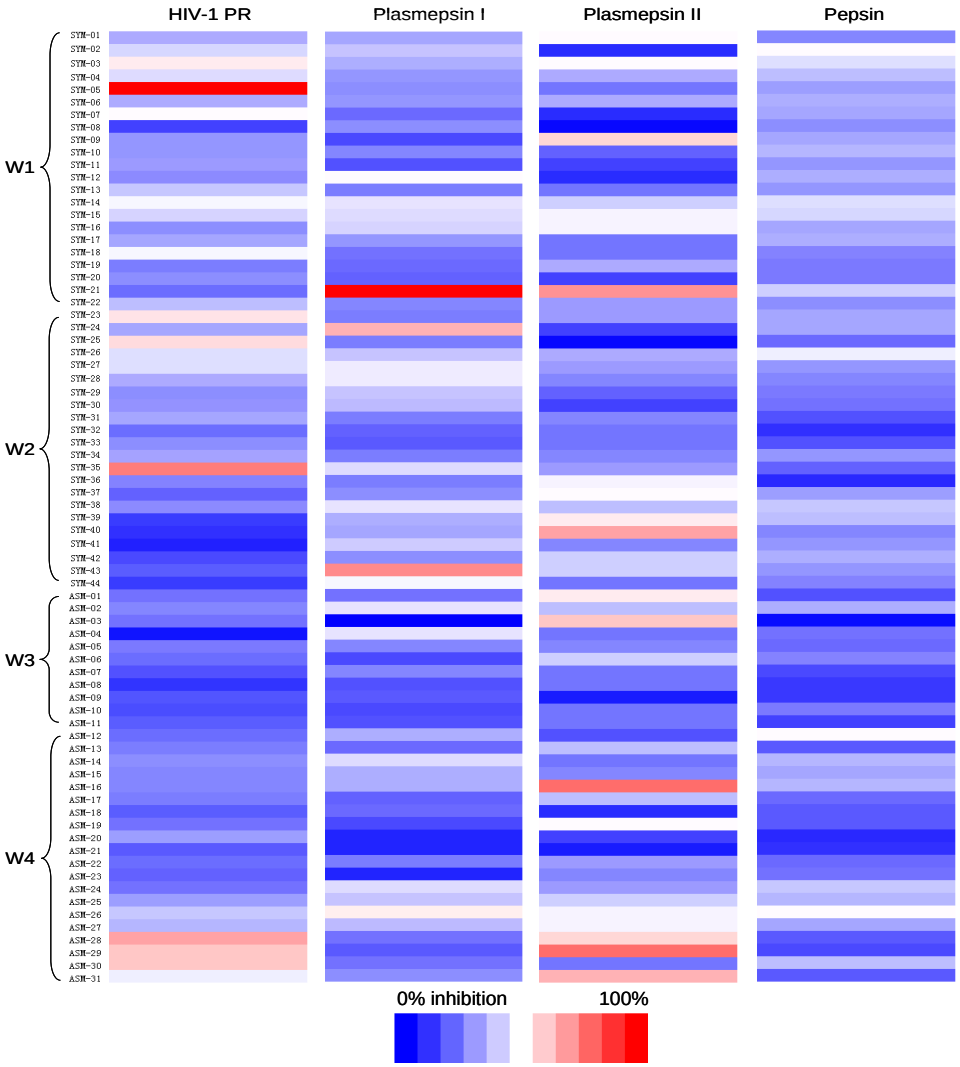


Figure 2.2 Inhibition profiles of the 75-member library against four aspartic proteases. Inhibitor concentration is kept constant (25 μM) with different enzymes.

compounds exhibited moderate activity against PM II (IC_{50} = 12.6 and 11.7 μM for ASM-16 and -29, respectively), which may be developed into potent inhibitors in future. SYM-21 and SYM-43 were identified to be moderate and selective inhibitors of PM II and PM I, respectively. These results thus validate our strategy as a feasible method for future discovery of other aspartic protease inhibitors.

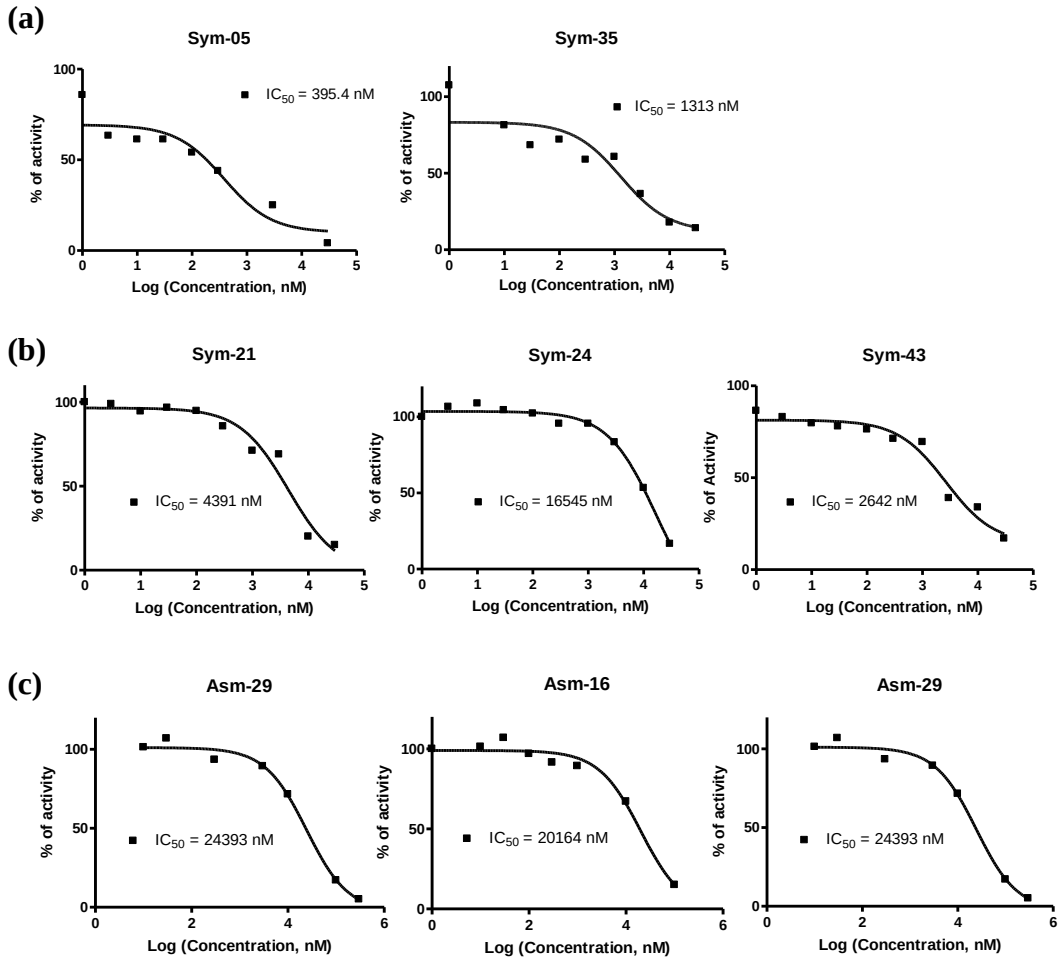


Figure 2.3 IC_{50} graphs for various inhibitors against (a) HIV-1 PR, (b) Plasmepsin I, and (c) Plasmepsin II.

2.3.5 K_i Measurements

To further investigate the potency of the strongest inhibitor **SYM-05** identified against HIV-1, a detailed kinetic evaluation was performed. Dose-dependent reactions were carried out by varying both the concentration of substrate and inhibitor while keeping the enzyme concentration constant. Briefly, a dilution series from

Table 2.1. Inhibition of the six selected inhibitors.

	IC_{50} (K_i) in nM					
Enzyme	**SYM-5**	**SYM-21**	**SYM-35**	**SYM-43**	**ASM-16**	**ASM-29**
HIV-1 PR	395.4 (400)	>25000	1313	>25000	>25000	>25000
PM I	>25000	4391	>25000	2642	>25000	>25000
PM II	>25000	3908	>25000	>25000	12657	11773

approximately 10000 nM to 100 nM (final concentration) was prepared for **SYM-05**. The enzymatic mix was incubated for 15 min before initiating the reaction with varied amounts of substrate (2.5 μM to 10 μM). The enzymatic reactions were immediately monitored at λ_{ex} = 340 nm and λ_{em} = 490 nm for a period of 10 min. The kinetic data obtained were analyzed using the Dixon Plot (1/initial velocity plotted against inhibitor concentration); affording four different linear graphs (corresponding to the different substrate concentration) that converged on the left of Y-axis to give a value that corresponds to K_i (Figure 2.4). Finally, it was found that K_i of **SYM-05** was 400 nM, which was quite consistent with its IC_{50} value.

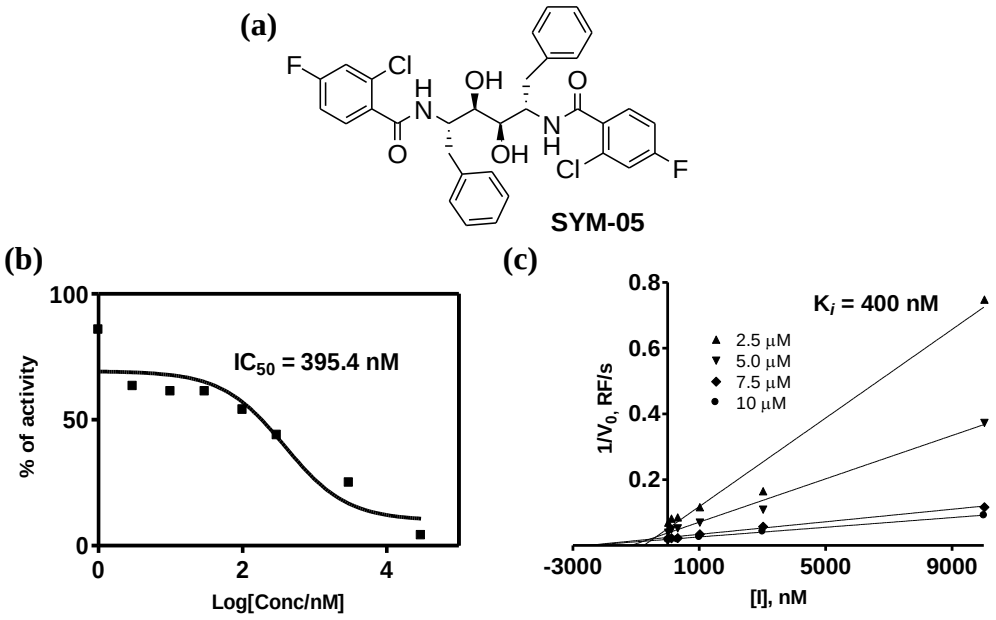


Figure 2.4 Characterization of inhibitor hit. Structure of **SYM-5** (a) IC_{50} (b) and K_i of **SYM-5** (c) against HIV-1 protease.

2.4 Conclusion

In summary, we have developed a solid-phase amide-forming approach for rapid assembly of both symmetric and asymmetric diol-containing small molecules as

potential aspartic protease inhibitors. One potent and selective inhibitor of HIV-1 PR was discovered which should possess better bioavailability properties (fewer amide bonds, lower MW) than its parental compounds. The clear advantage in our strategy lies in its feasibility for solid-phase synthesis of asymmetric diols. When compared with other potent HIV-1 protease inhibitors, our compound pales in terms of its inhibitory activity. Nevertheless, we believe the chemical strategy discolsed herein may be readily expanded in future for the construction of much bigger compound libraries which may lead to candidate compounds with improved biological activities.

Chapter 3

High-throughput Synthesis of Aspartic Protease Inhibitors and Probes for Proteome Profiling of Plasmapsins in Malaria Parasites

3.1 Summary

In this chapter we describe a high-throughput strategy to synthesize the 475-member inhibitor library against aspartic proteases using “click” chemistry. The strategy is robust, giving high-quality products which are spectroscopically pure enough to be used for biological screening without any further purification. Additionally, we have also developed the first affinity-based probes for functional profiling of all four plasmapsins (PMs) in intraerythrocytic malaria parasites. Subsequent *in situ* screening of parasites with these probes has led to the identification of a compound, **G16**, which showed good inhibition against all four PMs and parasite growth in infected RBCs. Molecular modelling indicates this inhibitor binds to the active site of the plasmepsins tested, as originally designed. Our finding indicates that feasibility of using A*f*BP approaches for identification of inhibitors against less-characterized enzymes (i.e. HAP) and those against multiple targets. We anticipate that these new chemical tools should facilitate discovery of new parasite biology and new anti-malaria drugs. In the current study, we were unable to detect previously predicted but unidentified aspartic proteases from the malaria

proteome. Thus, **G16** might have further targets, other than the four plasmepsins, in the malaria proteome.

3.2 Introduction

3.2.1 High-throughput Amenable Chemical Reactions in Inhibitor Discovery

High-throughput enzymology occupies a crucial role in modern drug discovery programs. The main challenge in this field is the development of high-throughput (HT) amenable chemical reactions that enable rapid generation of enzyme inhibitor libraries as well as probes. One such reaction is the Cu (I)-catalyzed, 1, 3-dipolar cycloaddition between an azide and an alkyne fragment, also known as "click" chemistry, which was pioneered by Sharpless and Meldal[35, 36] (Figure 3.1a). The major features of this reaction include its high efficiency, chemo-selectivity, modularity and biocompatibility, which allow direct *in situ* screening of the products formed without the need of further purification. To date, many research groups have successfully adopted this platform for the high-throughput screening and discovery of inhibitors against numerous enzymes, including HIV-1 protease,[46] sulfotransferases,[50] fucosyltransferases,[48] matrix metalloproteases,[43] acetylcholinesterase,[49] plasmepsins[41] and others. Another class of reactions possessing similar qualities is the amide bond forming reaction between an amine and a carboxylic acid using suitable activating/coupling reagents. Some research groups have recently used this reaction for solution-phase, rapid assembly of small molecule inhibitors against a variety of enzymes including HIV protease,[51] SARS-3CL protease,[54] and so on. This reaction are highly efficient, often generating the desired products in nearly quantitative yields,

thus allowing *in situ* biological screening to be carried out directly in some cases, even in the presence of excessive starting materials, reagents, or byproduct (Figure 3.1b). The effectiveness of this method, however, had recently been challenged, as it was discovered that unexpected byproduct may give rise to false results. Wong *et al*. recently reported that an intermediate benzotriazole ester of the amide bond-forming reaction was found to be responsible for the inhibitory activity toward SARS-3CL protease.[56] To avoid such limitation, we aim to develop a combinational synthetic approach by using solid-phase, amide bond-forming reactions and "click" chemistry.

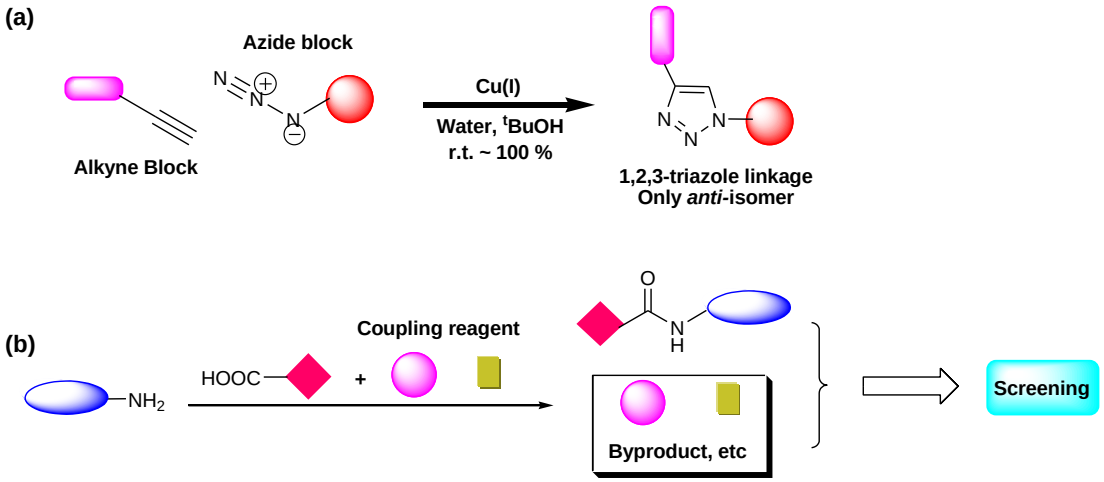


Figure 3.1 High-throughput amenable chemical reactions. (a) "Click" chemistry pioneered by Sharpless and Meldal. (b) Wong's solution-phase and *in situ* screening approach.

3.2.2 Functional Profiling Identification and Inhibition of Plasmepsins in Intraerythrocytic Malaria Parasites

Malaria is a global disease, affecting 300-500 million people annually and killing 1-2 million. The most deadly form of the disease is caused by the pathogen *Plasmodium falciparum*. Currently, quinolines and antifolates are the most common antimalaria drugs.[75] The cost of the drugs, as well as the emergence of multidrug

resistance, has however become a major problem, highlighting the need for newer and ideally cheaper drugs against this devastating disease.[76] Proteases, including cysteine (i.e. falcipains) and aspartic proteases (i.e. plasmepsins, or PMs), are required for parasite growth through the digestion of human hemoglobin and delivery of necessary nutrients. They have long been considered promising anti-malaria targets.[77] Genomic datas obtained for *P. falciparum* predict at least ten genes encoding aspartic proteases, four of which (PM-I, -II, -IV and the histo-aspartic protease or HAP) have been found so far, mostly in the food vacuole (FV) of the parasite. The other hypothetical aspartic proteases, however, have not been experimentally confirmed. In the last few years, drug discovery efforts on potential plasmepsin inhibitors have somewhat waned after gene knock-out experiments showed that parasites could still survive, albeit with a reduced growth rate, without most of the four functionally redundant FV plasmepsins.[78] it is now believed that the only effective way to kill the parasite with PM drugs is to look for potential inhibitors that can simultaneously target as many plasmepsins as possible.[79] At present, most inhibitors developed are only effective against selected PMs.[73c] This is due to difficulties associated with recombinant expression and insufficient biochemical characterizations of certain PMs (i.e. PM-I and HAP) *in vitro*,[80a,b] and the lack of methods that allow simultaneous screening of PMs' activity *in situ*.[80c] Previously, activity-based probes (ABPs) was successfully used for *in situ* screening of malaria cysteine proteases.[81] We report herein the first chemical proteomic approach for functional profiling of all four PMs in intraerythrocytic malaria parasites. This is made possible by the development of

affinity-based probes (A*f*BPs) against PMs (Figure 3.2).[82] *In situ* screening of PMs with these probes against a focused library of 152 hydroxyethylamine-containing small molecules has led to the identification of **G16** as an effective inhibitor against the parasite in infected RBC cultures.

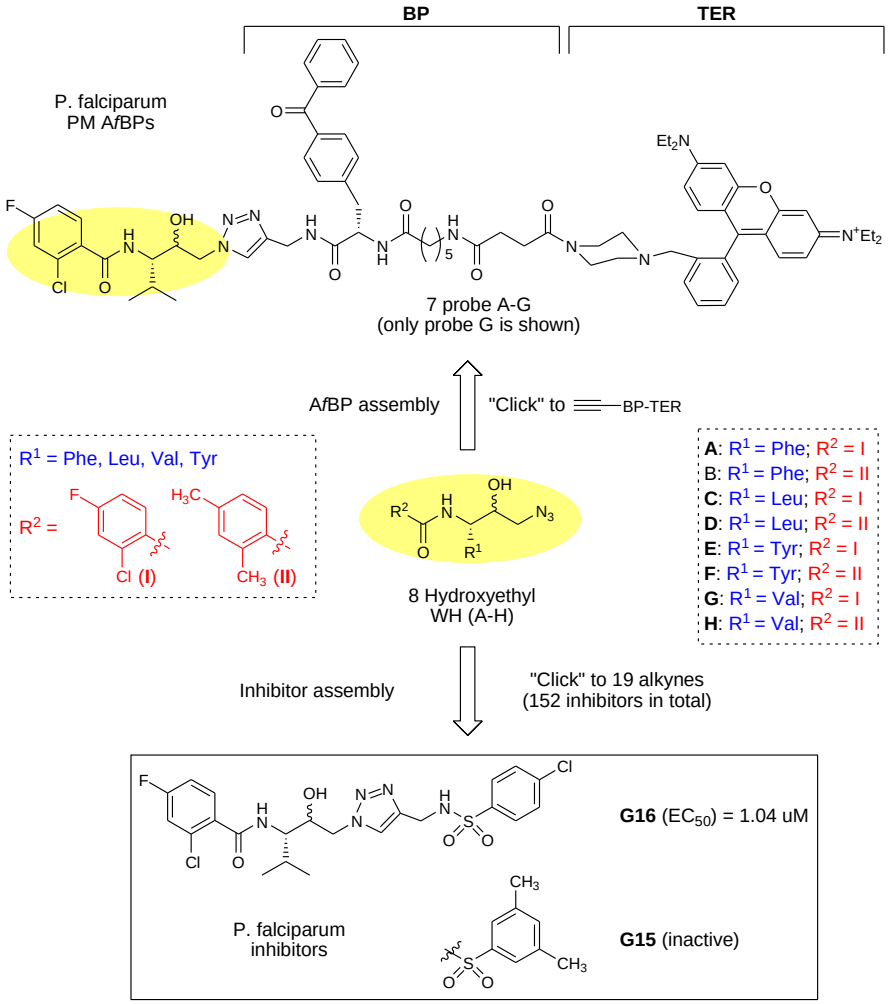


Figure 3.2 Assembly of affinity-based probes (A*f*BPs) and the 152-membered library of potential inhibitors against all four FV plasmepsins in *P. falciparum*. Click chemistry and *in situ* screening of inhibitor discovery.

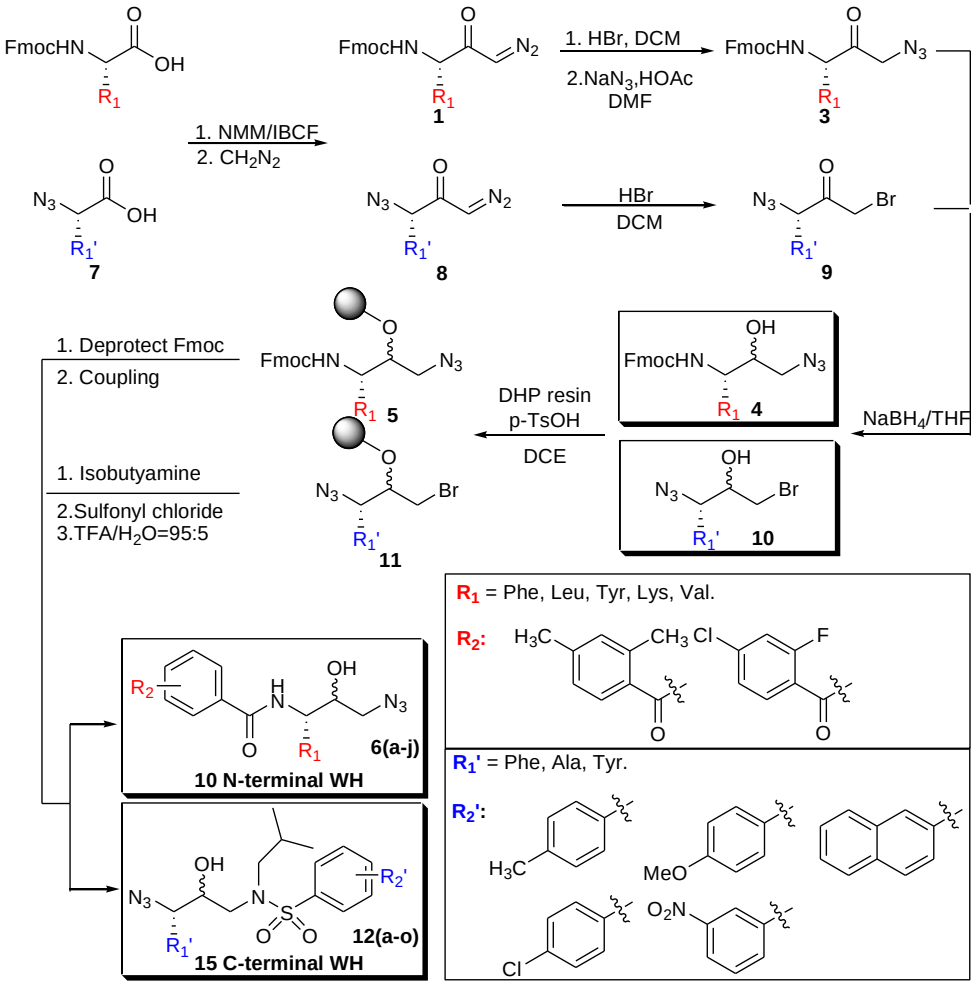
3.3 Results and Discussion

3.3.1 Solution-cum-solid Phase Synthesis of N, C-terminal Azides Warheads

Based on previous reported procedure with minor modification, we designed and synthesized a library of compounds which contained a hydroxyethylamine isosteres as

transition-state mimics, while diversifying the $P_2/P_{2'}$ residues to generate new inhibitors. Starting from commercial available amino acids, two different azide cores were prepared as summarized in Scheme 3.1. In the 10-member N-terminal warheads sublibrary, R_1 and R_2 positions, corresponding to the P_1 and P_2 residues, respectively, upon binding to the aspartic proteases, were varied with aromatic/aliphatic moieties (R_1 = Phe, Leu, Tyr, Lys, Val; R_2 = two aromatic acids as shown in Scheme 3.1). In the 15-member C-terminal warheads sublibrary, R_1' and R_2' positions, corresponding to the P_1 and $P_{1/2'}$ positions, respectively, were similarly varied. Briefly, *L*-amino acids first reacted with diazomethane prior to treatment with isobutyl choloroformate in order to activate acid group. Subsequently, bromination of acetone was carried out to afford compound **2** with good yield. Compound **2** was treated with NaN_3 overnight under acidic condition to give compound **3** in moderate yield. Subsequent reduction of ketone using $NaBH_4$ followed by coupling of the hydroxyl group with 4-dihydro-2H-pyran (DHP) resin under acid-catalyzed conditions gave resin **5**. Deprotection of the Fmoc group followed by standard acylation with two aromatic acids and finally cleavage by TFA to afford the 10-member N-terminal warhead sublibrary **6(a-j).** Similarly, starting from azido amino acids prepared following previous reported procedure, compound **10** was prepared following above the same four step reactions. Subsequently, it was attached to DHP resin followed by alkylation reaction with *i*-butylamine, acylation with five commercially available sulfonyl chlorides then TFA cleavage to afford C-terminal warhead sublibrary **12(a-o).** Eventually, the resulting 25 warheads were analyzed by LC-MS, which confirmed

most of them were sufficient purity (>90%) and were subsequently used for direct "click" with alkyne building blocks to generate the corresponding inhibitor library.

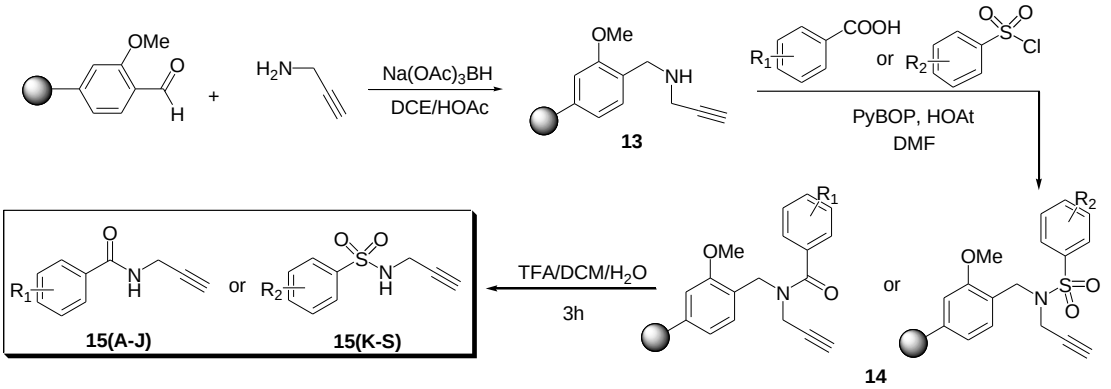


Scheme 3.1 Solution-cum-solid phase synthetic approach for synthesis of the N, C-terminal warheads (WH).

3.3.2 Traceless Synthesis of Alkyne Building Blocks.

To start the synthesis of alkyne building blocks on solid support, propargylamine was treated with FMP resin in the presence of $Na(OAc)_3BH$/2% glacial acetic acid in DCE to give the secondary amines **13** (Figure 3.2). Subsequent amide bond-forming reaction between the resin-bound secondary amines and the acid building blocks (**A-J**) was found to be highly challenging and required extensive optimizations. A variety of coupling reagents including HATU, PyBoP, HBTU, EDC and DIC were attempted,

only PyBop gave the desired products in satisfactory yield and purity. For sulfonyl amide bond-forming between sulfonyl chloride and the secondary amine was quite easy with excellent purity (> 90% in most cases). Finally, cleavage of the products from the resin using an optimized TFA cleavage cocktail (TFA:DCM:Water = 10:9:1) gave a total of 19 alkynes **15(A-S)**. Most crude products were characterized by LCMS/NMR and used for direct "click" with warheads to generate the corresponding inhibitor library.

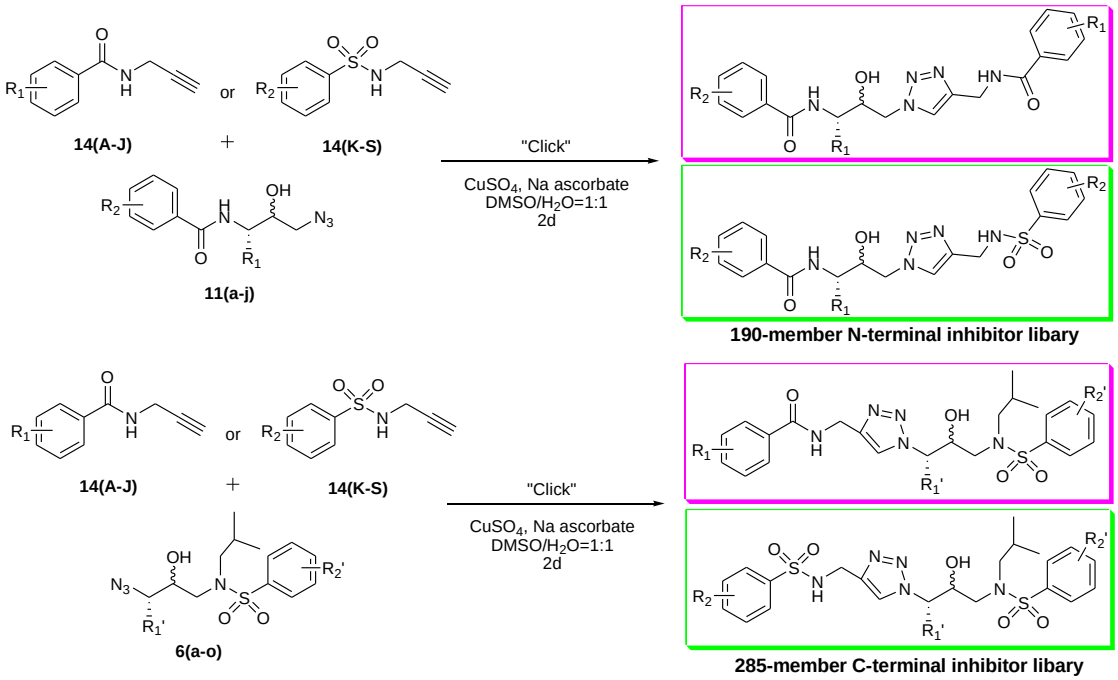


Scheme 3.2 Traceless synthesis of alkyne building blocks.

3.3.3 "Click" Assembly of Aspartic Protease Inhibitor Library

The Cu (I)-catalyzed "Click" reaction between azide warheads and alkyne building blocks was next carried out. After optimization for "click" condition, it was found that a mixed solvent system containing water and DMSO (1:1) enabled both the alkynes and the azides to be completely dissolved, and at the same time the 1, 3-dipolar coupling could be completed with extremely high efficiency. More importantly, the products could be taken directly without further purifications. 25 hydroxyethylamine warheads were clicked with 19 different alkynes to afford a 475 member library containing N-terminal and C-terminal sublibrary (Scheme 3.3).

LC-MS analysis of all coupling reactions indicated the complete consumption of the warheads and quantitative formation of the triazole products in most cases.



Scheme 3.3 "Click" assembly of the N, C-terminal inhibitor library.

3.3.4 Preliminary Screening Experiments with Aspartic Proteases

When the inhibitor library was in hand, their biological characterizations and utilities were further determined by my collaborators. The 475-member inhibitor library was firstly screened directly against several different aspartic proteases. Detailed procedure for the microplate based screening for HIV-1, pepsin, PM I and PM II was described below. The aspartic protease activities were determined by measuring the rate of hydrolysis of the commercial available fluorogenic substrates. The inhibition of proteins were measured using black polypropylene flat-bottom 384-well microtiter plate (Greiner, Germany) in a total reaction volume of 25 μL/well, monitored by Tecan Infinite F200 microplate reader at excitation of 390 nm and emission of 492 nm prior to the addition of substrate. One concentration of inhibitors

(2.5 μM) was first used in the assay, where necessary. Negative controls were performed in the absence of enzyme, and positive controls were also carried out in the presence of enzyme but without inhibitor. Consequently, some hits were picked out from the library. The inhibition results were shown in Table 3.1.

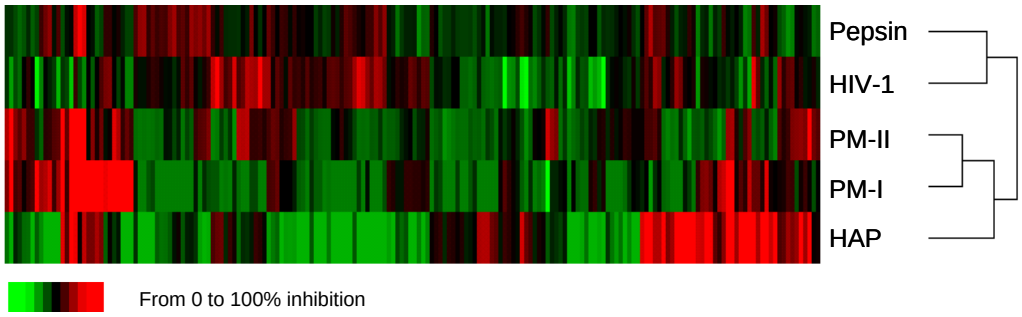


Figure 3.3 Inhibition profiles of the 475-member library against five aspartic proteases.

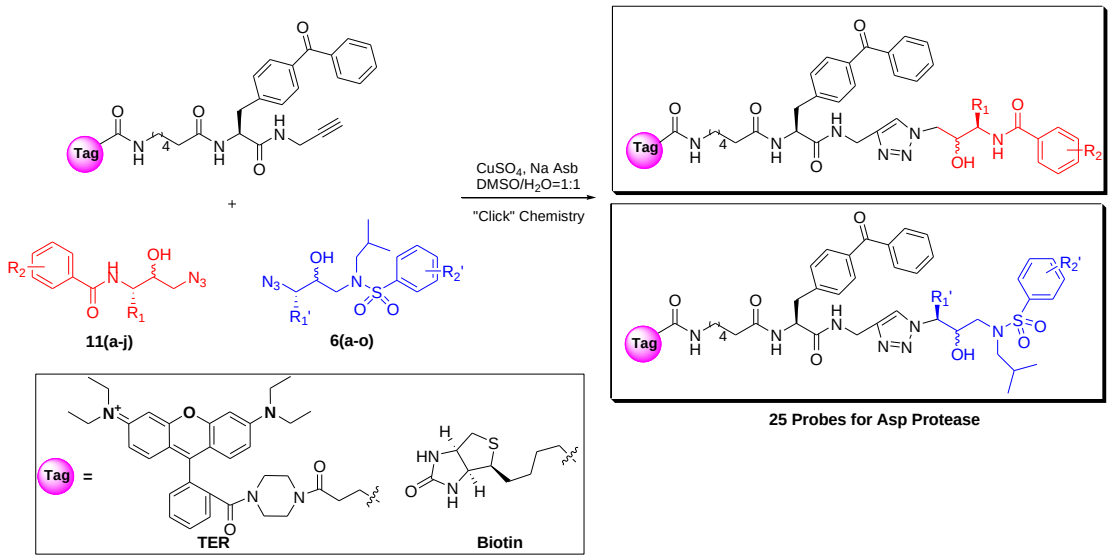
Table 3.1 Preliminary results of inhibition percentage for some hits.

Cmd	HIV-1			PM-II		PM I			
	Inhibition%(2.5 μ M)			Inhibition%(2.5 μ M)		Inhibition%(2.5 μ M)			
C-2	81.6	C-98	88.4			N-30	60	N-150	61
C-12	80.5	C-100	86.1	N-103	49.9	N-104	64	N-160	64
C-26	83.9	C-107	85.0	N-104	44.8	N-110	83	N-164	63
C-36	87.3	C-108	88.4	N-165	41.5	N-120	68	N-170	78
C-50	86.1	C-110	87.3	N-170	48.7	N-130	65	N-180	76
C-60	93.9	C-146	84.9	N-190	64.6	N-140	64	N-190	82
C-74	83.9	C-180	81.6						

3.3.5 "Click" Synthesis of Affinity-based Probes

Unlike other previously known aspartic protease probes, which were based on specific inhibitors against their respective targets (e.g. presenilin and β-secretase),[83] we aimed to establish a general approach that could be applicable for a variety of aspartic proteases. As shown in Figure 3.4, 25 probes (**A-G**), each containing a hydroxyethylamine warhead (WH) with a varied P_1 and P_2 group (assuming the orientation of probe-enzyme was as drawn), were assembled from the corresponding

azide-containing warheads and the alkynes, which contains a benzophenone (BP) photo-crossing unit and a tertraethylrhodamine (TER) or biotin reporter,[35,84] using above measured click chemistry. Hydroxyethylamine scaffolds are general transition state analogs of aspartic proteases. In all probes, aliphatic/aromatic groups were strategically chosen since they are preferred in some different aspartic proteases (for example, PM-I, -II & -IV, HAP, etc).



Scheme 3.4 "Click" assembly of affinity-based probes (A*f*BPs).

3.3.6 Labeling of Recombinant Aspartic Proteases

To demonstrate the suitability of the probes for the UV-initiated proteomic profiling of aspartic proteases, 15 probes which have good purities were characterized with various recombinant aspartic proteases, such as pepsin, renin, PM II and HAP, respectively. As shown in Figure 3.4, distinct labeling profiles of 15 probes against various aspartic proteases were observed. To further investigate the reactivities of our probes, 7 probes derivated from N-terminal warheads were characterized with previous several proteins. Because of the unknown catalytic mechanism of the

non-classic aspartic protease HAP, labeling reaction for HAP was carried out in both acidic and neutral *p*H. Each of the 7 probes was tested with each aspartic protease individually. Highly distinct labeling profiles against different aspartic proteases were observed. This result indicated that the variable R_1 and R_2 groups exerted a strong influence over specific enzyme/probe interactions. The labeling was abolished by treatment with pepstatin and mutations of key catalytic residues in the enzyme active site; thus, labeling was indeed dependent on enzyme activity.

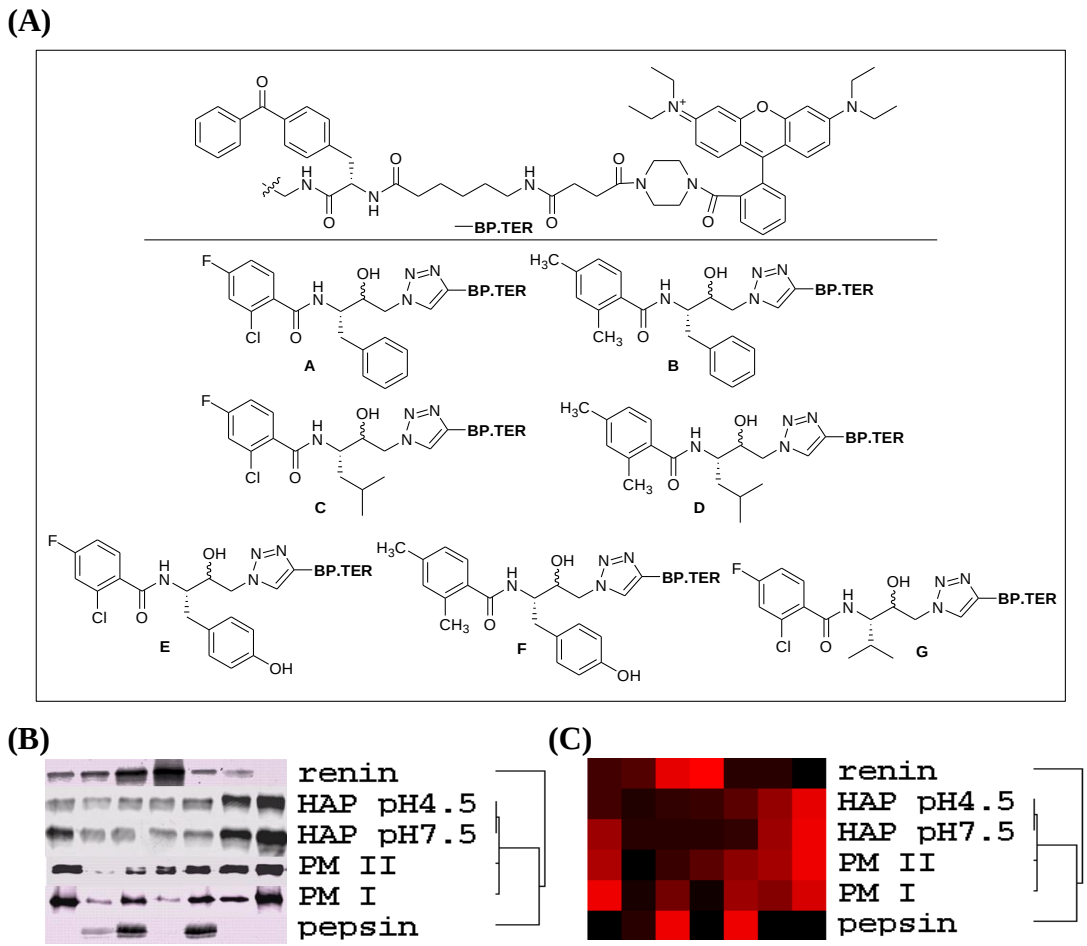


Figure 3.4 Profiling of *P. falciparum* aspartic proteases. (A) Structures of 7 affinity-based probes. (B) Characterization of recombinant PMs with seven A*f*BPs **A–G**. labeling profile of recombinant PMs with the A*f*BPs. (C) Tree-view representation of labeled bands (on the basis of their intensity).

3.3.7 Profiling of Aspartic Protease Activity Throughout Different of Blood

Stages *P. Falciparum*

To profile the aspartic protease activities in blood stage *plasmodium falciparum*, the probes were subsequently used to label proteomes of highly synchronized parasites obtained at different stages of parasite development (ring, trophozoite, and schizont; Figure 3.6a). A 37 kDa protein band, which corresponds to the molecular weight of the four known PMs, was highly visible across probes **A-G**, although the labeling intensities varied. Probe **G** consistently gave the strongest labeling profile and thus was chosen for further studies. Analysis by two-dimensional gel electrophoresis/MS and western blotting confirmed unambiguously that the 37 kDa band corresponds to all four labeled PMs The probes could therefore be used to profile all four known PMs and their enzymatic activities directly from the parasite. The detection of neither zymogens of the four PMs nor other proteins indicated the specificity of the probes in targeting only active PMs. Since it would be highly desirable if the probes could also detect the activity of other previously unidentified aspartic proteases in the malaria proteome (i.e. those predicted by genomic data), a deliberate effort was made in the 2D-GE/MS experiments to look for new protein spots. However, these attempts were unsuccessful; either these hypothetical aspartic proteases were present in very low abundance or our probes were not suitably designed to detect/label them. Subsequently, we determined the enzymatic activity of the four plasmepsins by labeling both detergent-soluble and detergent-insoluble lysates from each development stage (Figure 3.6B). The results indicated that PM activity was highly regulated and peaked at the trophozoite stage for the insoluble

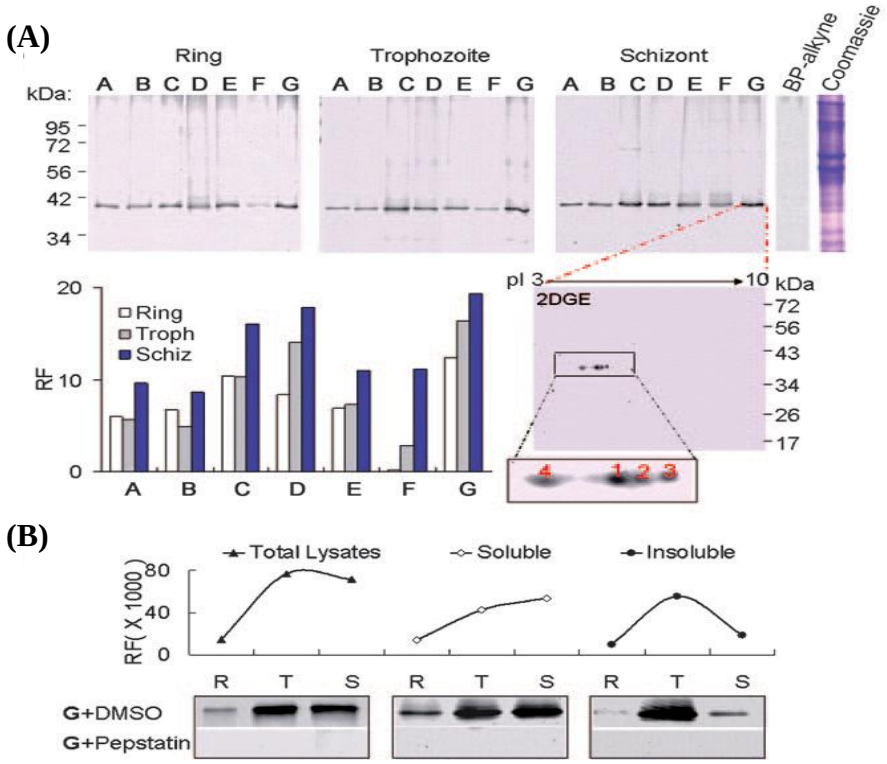


Figure 3.5 Profiling, identification and proteome characterization of *P. falciparum* aspartic proteases. (A) Parasite extracts at different stages were labeled with the seven A*f*BPs (the TER–BP–alkyne was used as a negative control). Top: in-gel fluorescence scanning showing the specific labeling of the 37 kDa band across different stages; bottom left: spectral counts of the labeled bands; bottom right: schizont-stage parasite extracts were labeled with probe **G** and then subjected to 2DGE/ MS analysis to identify the four PMs (spots 1-4 were identified as PM-II, PM-I, HAP, and PM-IV, respectively). pI = isoelectric point. (B) Characterization of plasmepsin activities from total lysates (left), NP-40-soluble fractions (middle), and insoluble fractions (right) across different intraerythrocytic stages of *P. falciparum*. R = ring stage, T = trophozoite stage, S = schizont stage. Top: spectral counts of the respective bands observed by in-gel fluorescence scanning.

fraction. This observation is consistent with the role of these PMs in hemoglobin degradation, which occurs in the food vacuole. The distinct profile of soluble-fraction activities peaked at the schizont stage; this peak probably indicates a change in the sub-cellular localization of the PMs (PM-II was previously shown to be released from the FV in the late schizont stage). Western blotting with antibodies against the four PMs also indicated that their absolute protein expression levels corroborated well with

labeling profiles observed with probe **G**. Thus, the utility of our probes for the accurate reporting of the activities of PMs from crude malaria proteomes was further confirmed. One of the key advantages of these A*f*BPs is the ability to use them for the simultaneous detection of the activities of multiple PMs in their native environment. This so-called *in situ* screening method, originally described by Cravatt and coworkers for other enzymatic systems,[63] enabled us to identify potential inhibitors against all four PMs directly from the malaria proteome without the recombinant production of every active PM.

3.3.8 Inhibitor Library Screening for PMs

To identify the most potent inhibitors against all four PMs, we preincubated each compound of the 152-membered hydroxyethylamine-based library from N-terminal sub-library with the parasite lysate and then added probe **G** and subjected the samples to UV irradiation. We determined the relative potency of each inhibitor by measuring the decrease in fluorescence intensity in the 37 kDa labeled band. A total of eight candidate "hits" were identified. These eight compounds were re-synthesized and purified, and their **G16**, showed an IC_{50} value of 937.5 nm (Figure 3.7A). In contrast, **G15**, a"false positive" identified from standard enzymatic assays by using selected recombinant PMs, showed significantly weaker inhibition (IC_{50} = 21.6 μM). This result underscores the importance of our *in situ* screening assay for the future discovery of general PM inhibitors. The inhibitory effect of these candidate compounds was tested with parasite-infected RBC cultures. The distinct activity/solubility profiles of PMs prompted us to test the inhibitory effect of the

inhibitors against schizontstage PMs, which showed the highest activity in the detergent nonyl phenoxylpolyethoxylethanol (NP-40). RBCs were treated with the inhibitors 40 h after parasite invasion (i.e. in the late schizont stage). Upon cell rupture, we measured the percentage of parasites in the ring and schizont stages (Figure 3.7B, C). Compound **G16**, but not **G15**, caused a marked decrease in the number of newly formed ring-stage parasites, and at the same time an increase in free extracellular merozoites. Thus, as well as blocking parasite development at the trophozoite/schizont stage, as one might expect, the inhibition of PM activity also caused the blockage of either the escape of the parasites from RBCs, or their reinvasion of RBCs. This speculation is further supported by previous findings that PM-II was able to digest an RBC membrane-skeleton protein in the late schizont stage at neutral *p*H, and that the invasion of *P. falciparum* merozoites was affected by treatment with pepstatin (a general aspartic protease inhibitor).[85] The inhibition of parasites by **G16** was dose-dependent, whereby the estimated EC_{50} value of 1.04 μM was similar to that obtained from the *in situ* proteomic screening (IC_{50} = 937.5 nM). In contrast, **G15** showed little or no inhibition towards infected RBCs, even at the highest concentration tested. Finally, **G16** showed no apparent cytotoxicity against common mammalian cell lines.

3.3.9 Identification of the 37-kDa Band (as PM-I, PM-II, HAP and PM-IV) by 2DGE & MS, and Western Blotting

To positively confirm the 37-kDa fluorescent band in the schizont stage parasite extract corresponds to the four known plasmepsins (PM-I, PM-II, HAP and PM-IV),

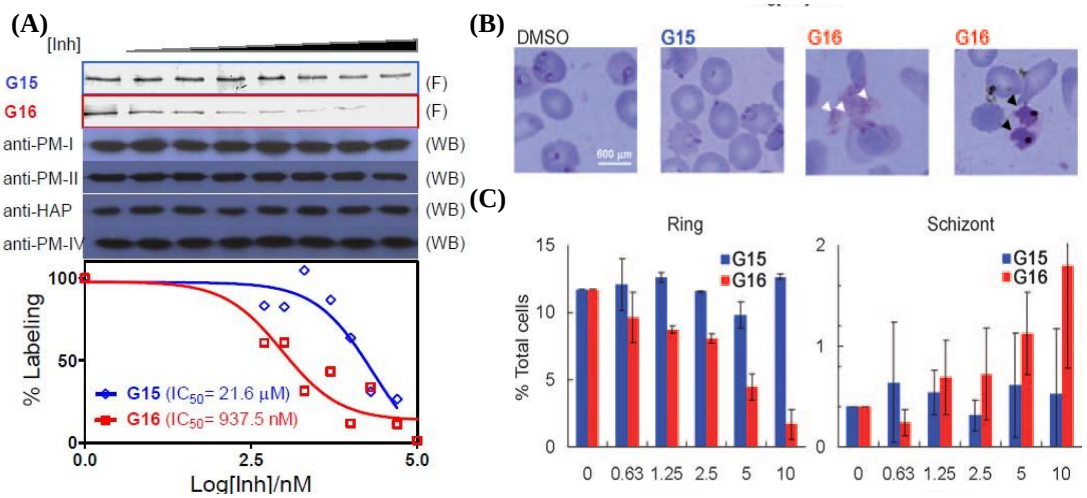


Figure 3.6 Inhibition of *P. falciparum* aspartic proteases. a) *In situ* screening assay and determination of the IC_{50} values of **G15** and **G16** against all four PMs in the parasite proteome. consistent PMs level detected by specific antibodies across all lanes(bottom four gels) show that the difference in labeling intensities were due to competitive inhibition of the PM labeling by **G16** rather than loading error.inh=inhibitor. (B) Representative images of parasite-infected RBCs treated with **G16** (10 mm; with dimethyl sulfoxide (DMSO) and **G15** as controls); arrows show the abnormal development of parasites. Scale bar: 600 mm. (C) Dosedependent-inhibition results from (B).

2-dimensional gel electrophoresis followed by MS analysis of the fluorescent spots was carried out. 1.0 mg of total proteins from the schizont stage parasite extract was pre-incubated in the reaction buffer (100 mM sodium acetate, 10 % glycerol, *p*H 5.0) and 5 μM of probe **G**. After labeling as earlier described, the sample was desalted and concentrated by acetone precipitation overnight at -20 ^{o}C. Precipitates were collected by centrifugation at 13,000 rpm. Samples were resuspended in 1× re-swelling buffer (8 M urea, 2% CHAPS and 2% IPG buffer and 0.002% bromphenol blue), and briefly sonicated. Just prior to use, DTT was added (final concentration: 3 mg/mL). Rehydration buffer and stock DTT (62.5 mg/mL) were stored at -20 ℃. Rehydrated IEF strips were isoelectrically focused at r.t. under low viscosity oil with a gradient voltage of 0–200 V for 1 min, 200–3,500 V for 1.5 h, and a constant voltage of 3,500

V for 1.5 h. After IEF separation, the gel strip was reduced and alkylated. The reduction step was performed for 15 min in 10 mL of equilibration buffer 1 (0.5% w/v DTT in 50 mM Tris-HCl *p*H 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). The alkylation step was performed for 15 min in 10 ml of equilibration buffer 2 (4.5% w/v iodoacetamide in 50 mM Tris-HCl *p*H 8.8, 6 M urea, 30%v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). The IEF strip was equilibrated in equilibration buffer stock solution (50 mM Tris-HCl *p*H 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). Next, the equilibrated IEF strips were each placed on a preparative well and sealed using 1% agarose plus 0.002% (w/v) bromphenol blue. After electrophoresis the two-dimensional gels were scanned with Typhoon fluorescence scanner. Fluorescence spots corresponding to each target (spots 1, 2, 3 and 4 for PM-II, PM-I, HAP and PM-IV, respectively, in Figure 3.6A) were excised directly from the gel trypsin digestion was performed with In-Gel Trypsin Digestion Kit from Pierce. After digestion, digested peptides were then extracted from the gel with 50% ACN and 1% formic acid. Tryptic peptide extracts were evaporated by speedvac and reconstituted with 10 μL of 0.1% TFA, a volume of 2 μL of the peptide extracts were manually spotted onto a Prespotted AnchorChip MALDI target plate for MALDI-TOF Mass Spectrometry (Bruker Daltonics) and incubated for 3 min, followed with washing step with 10 μL of 10 mM ammonium phosphate in 0.1% TFA, then dried. MALDI TOF mass spectra were recorded using Ultraflex III TOF/TOF mass spectrometer (Bruker Daltonics) with the compass 1.2 software package including flexControl 3.0 and

flexAnalysis 3.0, calibrated with PAC peptide calibration standards. MS/MS analysis for the major peaks in PMF spectra were carried out by auto LIFT on the MALDI-TOF/TOF instrument. MS and MS/MS Peak lists with intensity value were submitted to Matrix Science Mascot server (http://www.matrixscience.com/) through BioTools 3.0 (Bruker Daltonics) using database NCBInr with species of Plasmodium falciparum (malaria parasite), variable modifications of carbamidometyl on cysteine (C) and oxidation on methione (M), allowing maxium of 1 trypsin missed cleavage, peptide mass tolerance at 200 ppm; MS/MS mass tolerance of 0.7 Da.

3.3.10 Molecule Modeling G16 Binding to Plasmepsins

Lastly, to gain insight into the binding mode of **G16** with FV PMs, we carried out molecular-docking studies with three of the four PMs (Figure 3.8, the X-ray crystal structure of PM-I has not been reported). The results showed that the molecule bound well in the active site confined by the D34-D214 aspartic acid pair in the structure of PM-II and PM-IV, and by H32 and D214 in the structure of HAP (Figure 3.8a-c). Detailed analysis of the top 100 docked models identified the preferred conformations of **G16** in PM-II and PM-IV, as well as in nonclassical HAP (Figure 3.8g-i). The results also showed that the hydroxy group of G16 interacted closely with D34 and D214 in PM-II and PM-IV, as originally designed (Figure 3.8d, f). In binding to HAP, **G16** assumed a position in which its hydroxy group fitted between H32 and D214 in the enzyme active site (Figure 3.8e). A similar interaction was observed previously between the hydroxy group in pepstatin (a general aspartic protease inhibitor) and HAP.[86, 87]

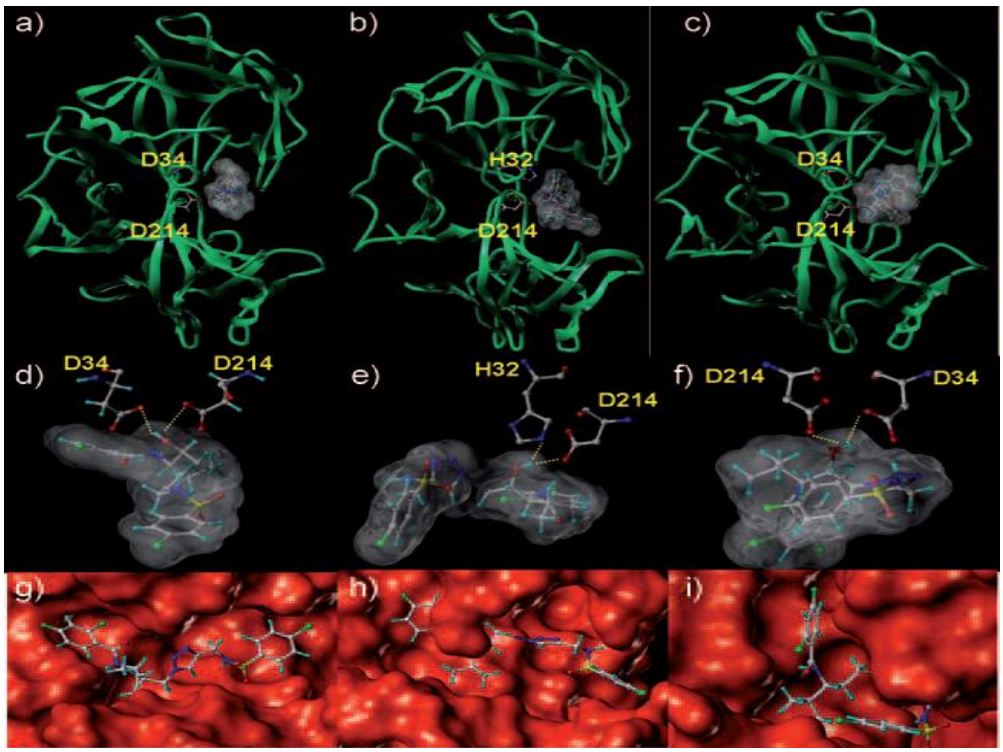


Figure 3.7 Molecular docking of **G16** in the active site of PM-II, HAP, and PM-IV. **G16** binds to the active-site binding pocket of a) PM-II, b) HAP, and c) PM-IV. The hydroxyl group of **G16** interacts closely with d) D34 and D214 in PM-II, e) H32 and D214 in HAP, and f) D34 and D214 in PM-IV. g–i) Preferred conformation of **G16** in the binding pocket of PM-II (g), HAP (h), and PM-IV (i).

3.4 Conclusion

In summary, we have developed a high throughput strategy for synthesis of aspartic protease inhibitors and affinity-based probes for the functional profiling of all four PMs in intraerythrocytic malaria parasites. Subsequent *in situ* screening of parasites with these probes led to the identification of a compound, **G16**, which showed good inhibition against all four PMs and parasite growth in infected RBCs. Molecular modeling indicated that this inhibitor bound to the active site of the plasmepsins tested, as originally designed. Our results indicate the feasibility of using A*f*BP approaches for the identification of inhibitors of less-characterized enzymes (such as HAP) and inhibitors of multiple targets. We anticipate that these new

chemical tools should facilitate the discovery of unknown parasite biology and new anti-malarial drugs. In the current study, we were unable to detect previously predicted but unidentified aspartic proteases from the malaria proteome. Thus, **G16** might have further targets, other than the four plasmepsins, in the malaria proteome. Future research will focus on the development of new chemical probes to address these issues.

Chapter 4

Small Molecule Microarray-facilitated Screening of Affinity-based Probe for γ-secretase

4.1 Summary

The ability to rapidly screen a large library of compounds represents a major bottleneck in the discovery of biologically relevant molecules and drug leads. This chapter extends the previously discussed high-throughput approach (Chapter 3) towards the rapid screening and identification of potent inhibitors and probes of enzymes. A 198-member library of putative aspartic protease inhibitors was screened with aspartic proteases. By incubating fluorescently labeled cellular lysates which over-expressed γ-secretase on small molecule microarray (SMM), we obtained inhibitor fingerprint of γ-secretase. A number of the most potent inhibitors were discovered and converted to affinity-based probes (A*f*BPs) using convenient, 1-step "click" chemistry. Overall this strategy offers not only a rapid method for inhibitor discovery, but also a viable method for high-throughput discovery of A*f*BPs against γ-secretase. We envisage that this platform will become a useful tool for high-throughput studies of proteomic profiling.

4.2 Introduction

The field of proteomics aims to develop methods for the comparative analysis of large numbers of proteins in complex biological samples. A chemical strategy referred

to as activity-based protein profiling (ABPP) utilizes active-site-directed probes to profile the functional state of enzymes in proteomes.[62] To date, ABPP probes have been developed for a number of enzyme families, and have been used to discover enzyme activities upregulated in aggressive cancer cells.

Microarray technologies based on surface-bound small molecules and proteins provide a convenient way to probe biomolecular interactions with a minimum amount of sample and therefore have been widely used in screening of primarily immobilized nucleic acids and proteins/peptides. Small molecule microarrays (SMMs) are becoming increasingly important for the large-scale determination of protein-ligand interactions, high-throughput screening of enzyme substrates/inhibitors and the development of biosensors and biomarkers for disease analysis.[88] Although purified proteins are usually used in binding assays on microarrays, this technology has been extended to the use of proteins from crude cell lysates. For example, Bradner *et al*. incubated small-molecule microarrays with cellular lysates containing over-expressed epitope-labeled proteins from mammalian cells.[99] Detection of small-molecule-protein interactions was achieved by treatment of the microarrays with fluorescently labeled antibodies that interact with the epitope tag.

Amyloid β-protein (Aβ), a central pathogenic feature of Alzheimer's disease (AD), is generated through proteolysis of amyloid precursor protein (APP) by membrane-associated aspartic proteases including β-and γ-secretases.[90] β-secretase has been well characterized; the study on γ-secretase, however, has been much more challenging, because the enzyme is a multimeric membrane protein complex,

comprising presenilin (PS), nicastrin, anterior pharynx defective 1 (Aph-1) and presenilin enhancer 2 (Pen-2). The proteolytically active form of PS is made up of two subunits, termed the N-terminal fragment (NTF) and C-terminal fragment (CTF). Despite recent advances, much remains to be understood about what the γ-secretase complex interacts with and how it carries out the proteolytic process. Due to well-known difficulties associated with biochemical studies of membrane proteins, much work on γ-secretase has so far relied on the development of the so-called affinity-based probes (A*f*BPs), which are typically generated from known inhibitors of γ-secretase by attaching a photo-cross linking group (i.e. benzophenone (BP) or diazirine) and a fluorescent dye (i.e. tetraethylrhodamine (TER)) to the inhibitors.[91] For example, A*f*BPs based on L685,458, a potent hydroxylethylene-based γ-secretase inhibitor, were first developed by Li *et al*.[91a] Recently, probes based on DAPT and arylsulfonamide derivatives have also been reported.[91(b-d)] The key advantage of all these probes is their ability to selectively bind the active subunits of γ-secretase complex (i.e. PS subunits) from the whole cell extract and, upon UV-photolysis and in gel-analysis, accurately report the active state of γ-secretase without the need of isolating the enzyme complex. As a result, these probes have provided invaluable tools for the studies of Alzheimer's disease, both in its pathology and potential therapeutics. Similar strategies are widely used for the development of A*f*BPs against other classes of enzymes in the field of activity-based protein profiling (ABPP) pioneered by Evans and Cravatt.

To convert an inhibitor into an effective A*f*BP, the following two steps are taken:

(1) identification of a suitable inhibitor (or an inhibitor scaffold) against the target enzyme (or the class of enzymes). This typically limits the A*f*BP strategy only to proteins for which inhibitors are known a priori. (2) Introduction of a BP-fluorophore linker into the molecule, which in many cases may abolish its target-binding ability and, as a result, leads to inactive probes.[101b] Consequently, a much larger number of potential probes (than the final probes obtained) are usually synthesized chemically and tested biologically. Clearly, there is an urgent need to develop new ways for rapid screening and identification of A*f*BPs against different protein targets, including both well-known and less-characterized ones. In our studies, we realized one of the key steps in most current SMM technologies requires the site-specific immobilization of small molecules onto the glass slide through the introduction of a suitable chemical "tag" (and the connecting linker) into the small molecules (e.g. biotin in Figure. 4.1).[4b] Upon SMM screening, the "hits" identified were in fact the "tagged" small molecules. We reasoned that the direct conversion of biotin into a linker containing both BP and a dye should retain the full protein binding ability of the "hit", thus making them suitable A*f*BPs against the target protein. So by combining concepts in microarray technologies and activity-based protein profiling, new A*f*xBPs may be rapidly identified. Herein, we report the first small molecule microarray (SMM)-facilitated approach for high throughput identification of A*f*BPs, in which γ-secretase was first used as our model target.

4.3 Results and Discussion

4.3.1 Synthesis of N, C-terminal Hydroxyethylamine Warheads

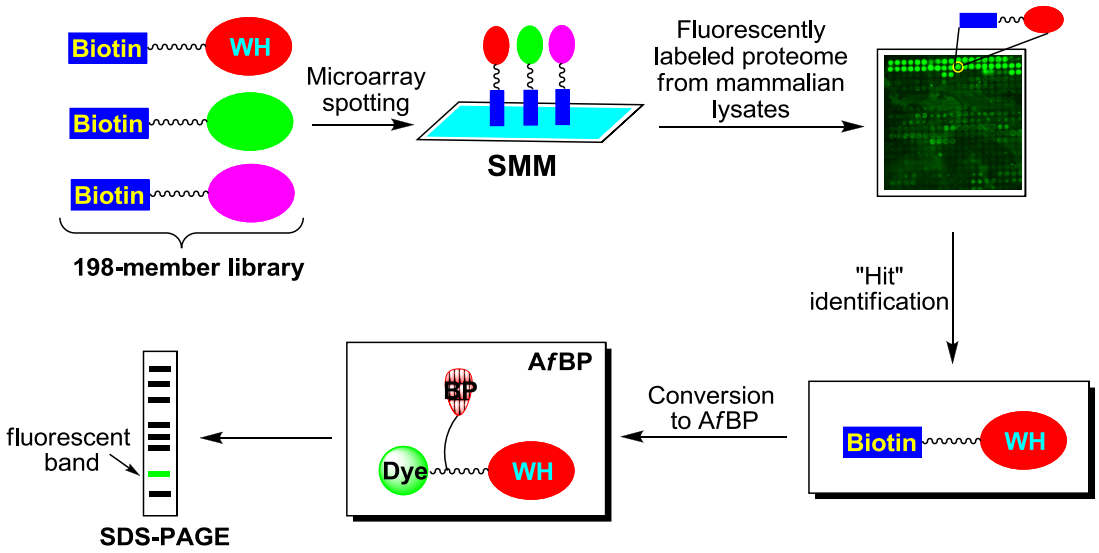
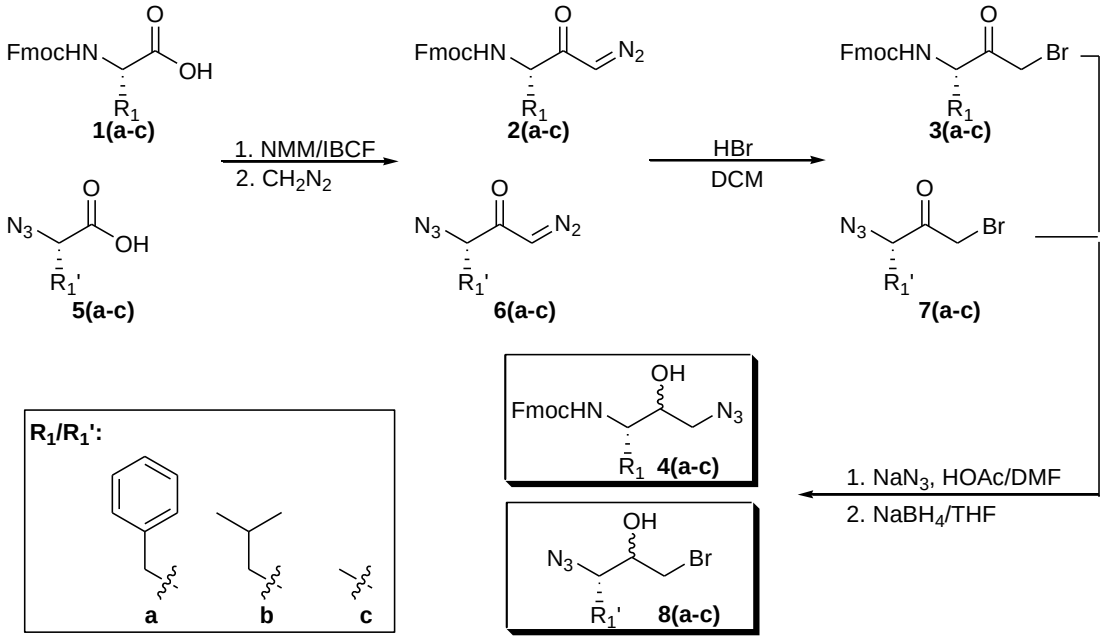


Figure 4.1 Overall strategy of the small molecule microarray (SMM)-facilitated platform for high-throughput identification of A*f*BPs.

To apply our strategy to the screening of A*f*BPs against γ-secretase as well as other cancer cells, we developed a solid-phase method for combinatorial synthesis of hydroxyethylamine-based small molecule inhibitors. Hydroxyethylamine are well-known scaffolds as transition state analogs of aspartic proteases, and have been included in the inhibitor design of numerous aspartic proteases including γ-secretase. In the 123-member N-terminal sublibrary, R_1 and R_2 positions, corresponding to the P_1 and P_2 residues, respectively, upon binding to the aspartic protease, were varied with aromatic/aliphatic moieties (R_1 = Phe, Leu, Ala; R_2 = aromatic/aliphatic acids). In the 75-member C-terminal sublibrary, $R_{1'}$ and $R_{2'}$ positions, corresponding to the P_1 and $P_{1/2'}$ positions, respectively, were similarly varied ($R_{1'}$ = Phe, Leu, Ala; $R_{2'}$ = Sulfonyl chloride). Starting from commercial available amino acids, two different azide cores were prepared as summarized in Scheme 4.1. In the both N, C-terminal warheads sublibrary, $R_1/R_{1'}$ positions, corresponding to the P_1 residues, respectively, upon binding to the aspartic proteases, were varied with aromatic/aliphatic moieties.

Briefly, Fmoc protective *L*-amino acids first reacted with diazomethane prior to treatment with isobutyl choloroformate in order to activate acid group. Subsequently, bromination of acetone was carried out to afford compound **3** with good yield. Compound **3** was treated with NaN_3 overnight under acidic condition and subsequently reduced by $NaBH_4$ to afford compound **4(a-c)**. Starting from the azido amino acids, compound **8(a-c)** were prepared following above the same step reactions.



Scheme 4.1 Synthesis of the N, C-terminal hydroxyethylamine transition state warheads **4(a-c)** and **8(a-c).**

4.3.2 Solid-phase Synthesis of Inhibitor Library

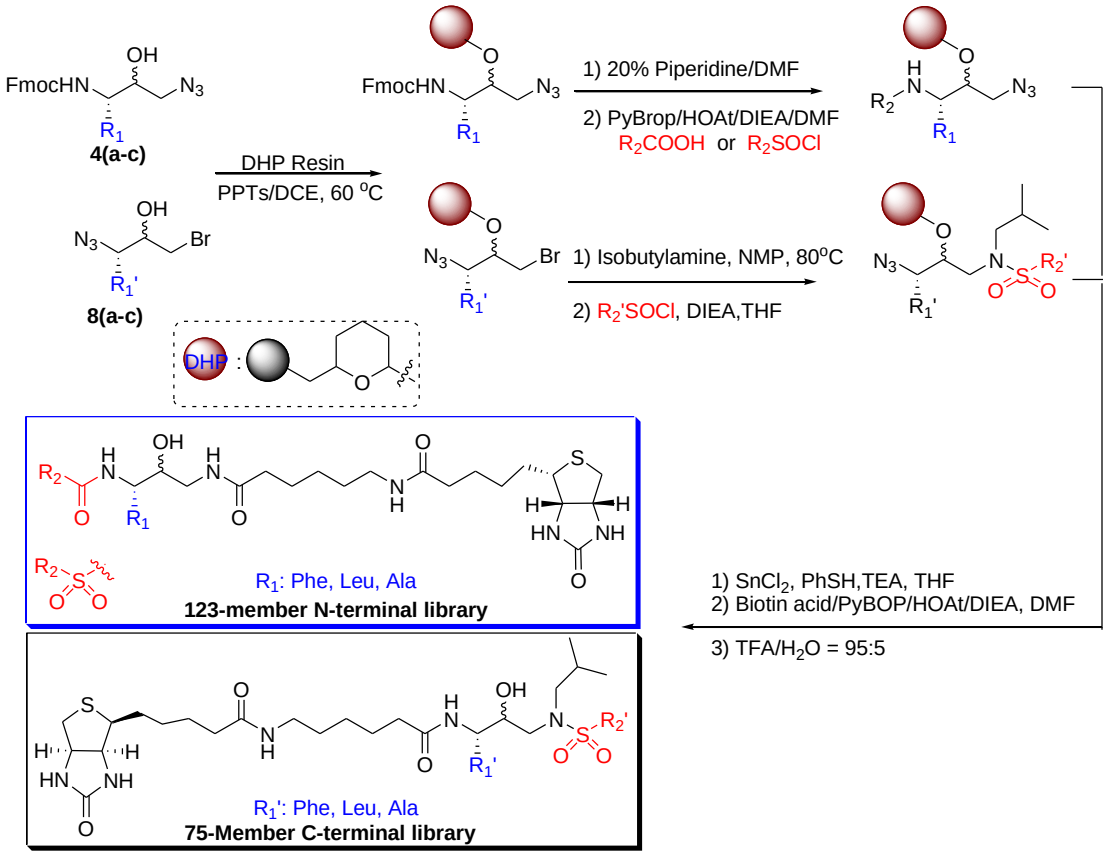
In both sublibraries, a biotin tag (and a connecting linker) was introduced in each compound for subsequent immobilization onto SMM. Since most known aspartic proteases recognize the four key residues (P_2, P_1, $P_{1'}$and $P_{2/3'}$) flanking the scissile bond of the substrates, our 198-member combined library could potentially be used for screening of other aspartic proteases besides γ-secretase. As shown in Scheme 4.2,

the synthetic strategy was so chosen because (1) it's solid-phase and amenable to large-scale synthesis of hydroxyethylamine-containing compounds from readily available Fmoc-protected *L*-amino acids; (2) the diversity of the library can be easily introduced throughout the P_2-$P_{2'}$ positions; (3) the key azido intermediates in the synthesis could be easily recovered and converted to the corresponding A*f*BPs using "click" chemistry. Briefly, **4(a-c)** were first immobilized at the -OH group onto 4-dihydro-2H-pyran (DHP) resin under acid-catalyzed conditions. Subsequently, deprotection of the Fmoc group followed by standard acylation with different acid building blocks and azide reduction, gave the corresponding amines. Finally, attachment of a biotin linker then TFA cleavage gave the 184-member N-terminal compounds. Similarly, **8(a-c)**, again from the corresponding three *L*-amino acids, were attached to DHP resin, followed by alkylation reaction with *i*-butylamine, acylation with commercially available sulfonyl chloride building blocks, reduction with $SnCl_2$/PhSH, attachment of the biotin linker then TFA cleavage, gave the 75-member C-terminal compounds. Finally, the resulting 198 compounds were analyzed by LC-MS, which confirmed most were of sufficient purity (>80%) and were subsequently used for direct immobilization onto avidin-functionalized slides to generate the corresponding small molecule microarrays.

4.3.3 Inhibitor Fingerprinting of HAP and Its Mutants

To assess whether the SMM could be used to detect aspartic protease-small molecule interactions, we carried out the preliminary screening with the recombinant purified protein, histo-aspartic protease (HAP), which is one of the aspartic proteases

from the parasite *Plasmodium falciparum*.[80b] HAP was fluorescently labeled with Cy3 before being applied onto the SMM (selected 198 compounds). As negative controls, three inactive mutant proteins of HAP (E278K, H34A and S37A mutants), previously shown to have minimum enzymatic activity, were used. Subsequently, the slides were scanned using an ArrayWoRx microarray scanner installed with the relevant filters (Cy_3: lex/em = 548/595 nm). Only HAP showed distinct binding profile against the SMM, while none of the three active-site mutants showed any significant binding (Figure 4.2B). This indicates that our SMM platform is able to sensitively report activity-based binding events of aspartic proteases towards their small molecule inhibitors.



Scheme 4.2 Solid-phase synthesis of the N- and C-terminal inhibitor libraries.

4.3.4 Inhibitors Fingerprinting of Cellular Lysates

Previous SMM strategies have mainly relied on incubation with a purified protein of interest. Recently, Koehler and coworkers have shown cell lysates expressing a target protein could be used directly for ligand identification in SMM as well.[99] In our case, since it is notoriously difficult to isolate γ-secretase, we decided to carry out our SMM screening using the fluorescently labeled membrane fraction of mammalian cell lysates. The γ-30 cell line, which is a γ-secretase overexpressing CHOK1 mammalian cell line, was used to ensure a sufficient amount of active γ-secretase was present in the lysate. The cells were grown as previously described.[92]

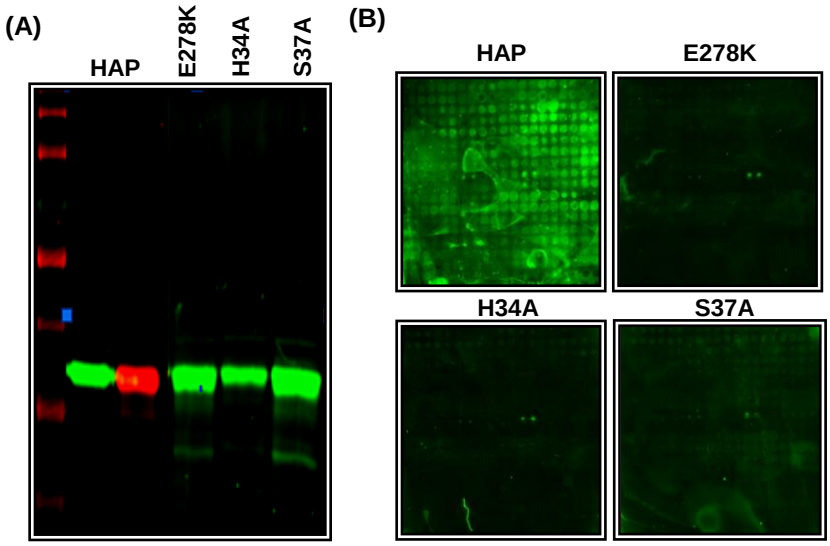


Figure 4.2 Microarray screening images. (A) Fluorescent gel images of the dye-labeled HAP and its mutants used in this study. Cy3-and Cy-5 labeled proteins were colored in green and red, respectively. (B) Small molecules microarray immobilized with the 198-member library was screened against fluorescently labeled HAP and its active-site mutants. Results indicated the ability of HAP binding to SMM was lost after mutation.

Subsequently, the membrane fraction was isolated, solubilized and fluorescently labeled with Cy3 before being applied onto the SMM. As shown in Figure. 4.3, highly reproducible and distinct binding profile of the 198-member library against the

membrane fraction of γ-30 lysates could be obtained. Interestingly, none of the N-terminal sublibrary showed any significant fluorescent binding. Of the strong binders, most were identified as members of the C-terminal sublibrary containing an alanine residue at the P_1 position. C-c-5 and C-c-24 showed the strongest relative fluorescence binding, and therefore were chosen for further studies. N-a-31, a non-binder, was also chosen as a negative control.

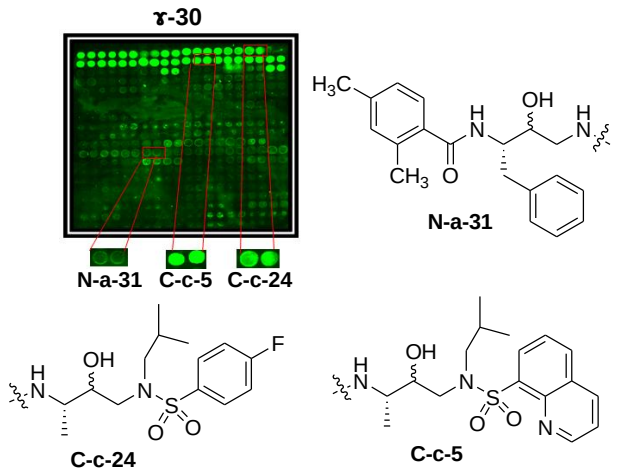
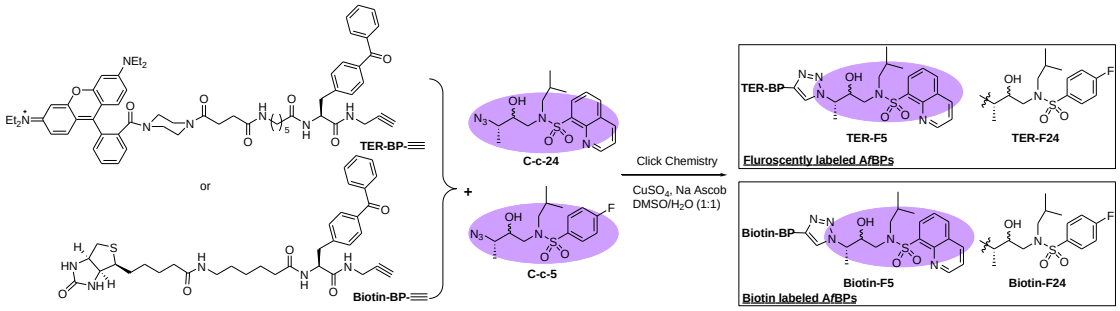


Figure 4.3 Fingerprinting of the 198-member library screened against the fluorescently labeled membrane fraction of γ-30 cell lysate, with (left) microarray image showing selected binders and non-binder and (right) their chemical structures. All compounds were printed in duplicate (horizontally).

4.3.5 "Click" Synthesis of TER/Biotin Probes

To make the corresponding A*f*BPs of above three compounds, their azido intermediates (shaded) were taken, and ‘‘click’’ assembled with TER-BP alkyne (for in gel activity-based profiling) and Biotin-BP alkyne (for pull-down experiments), as shown in Scheme 4.3. Previously, the modularity of ‘‘click" chemistry had been successfully explored to facilitate the synthesis of other classes of activity-based probes.[4d] Consequently, the final probes were purified by prep-HPLC and characterized/confirmed by NMR and LC-MS.



Scheme 4.3 "Click" assembly of affinity-based probes (A*f*BPs) against γ-secretase.

4.3.6 In-gel Fluorescence Scanning of Cellular Lysate with Probe and Validation

To demonstrate the ability of our SMM-derived A*f*BPs for UV-initiated, activity-based profiling of γ-secretase in cellular lysates, we carried out in-gel fluorescence labeling and pull-down experiments. As shown in Figure 4.4a, the two positive probes identified from SMM, F5 and F24 (corresponding to C-c-5 and C-c-24 in SMM, respectively) strongly labeled a 26-kDa protein in the total lysate of γ-30 mammalian cell line with high specificity. This band was subsequently confirmed as the N-terminal fragment of PS (PS-NTF) by both western blotting with anti-PS antibody and pull-down experiments (Figure. 4.4b). On the contrary, the total lysate of γ-30 treated with the negative control A31 (corresponding to N-a-31 in SMM) failed to give any noticeable labeled band. From our results, it is interesting to note only PS-NTF was labeled by our probes. Previous work on the γ-secretase complex had identified that both NTF and CTF are needed to form the catalytic core of PS.[90,91] A quick survey of other known A*f*BPs of γ-secretase revealed that most also labeled only PS-NTF. This thus unambiguously confirms our SMM-facilitated screening approach as a valuable tool for rapid discovery of A*f*BPs.

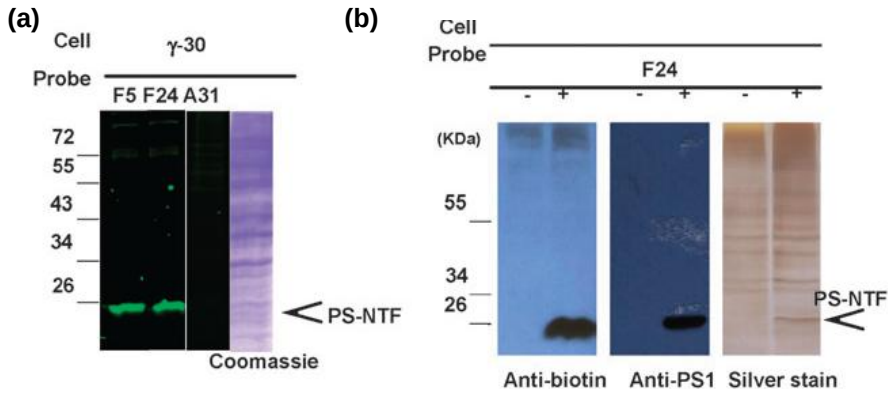


Figure 4.4 Validation of γ-secretase. (a) In-gel fluorescence profiling of γ -30 lysate using different A*f*BPs. The arrowed band was identified as PS-NTF. (b) After pull-down, western blotting showed the biotinylated protein in γ-30 cell lysate labeled by biotin-**F24** as PS-NTF.

4.4 Conclusion

In summary, a small molecule microarray-facilitated screening strategy has been developed for high-throughput identification of affinity-based probes. This approach does not require known inhibitors of a protein target and minimizes risks involved in the loss of protein-binding property of these inhibitors due to linker introduction. We further demonstrated the utility of this method by successfully identifying highly specific γ-secretase probes from a generic hydroxyethylamine small molecule library.

Chapter 5

Applying Small Molecule Microarrays and Resulting Affinity Probe Cocktails for Proteome Profiling of Mammalian Cell Lysates

5.1 Summary

Small molecule microarrays (SMMs) are proving to be increasingly important tools for assessing protein-ligand interactions, as well as in screening for enzyme substrates and inhibitors in a high-throughput manner. We previously described a SMM-facilitated screening strategy for the rapid identification of probes against γ-secretase, an aspartic protease. In this study, we extended upon this work with an expanded library of hydroxylethylamine-derivated inhibitors which non-exclusively target aspartic proteases. Our library was diversified across P_2, P_1, $P_{1'}$ and $P_{2'}$ positions. Accordingly, 86 new inhibitors were synthesized using a combinatorial, solid-phase synthetic approach, bringing the total library size to 284-biotinylated compounds, which were arrayed onto avidin slides. In order to elucidate enzymatic activity and profiles within complex biological samples, screens were performed using fluorescently-labeled mammalian cell lysates. This yielded reproducible profiles or binding fingerprints that corresponded with interactions from aspartic proteases present in the sample, or accessory proteins, as well as other interacting targets. The brightest microarray hits were converted to affinity-based probes (A*f*BPs) using convenient, 1-step "click" chemistry with benzophenone from the relevant building

blocks. Pull-down/mass spectrometric analysis with these probe cocktails yielded target lists that included well-known aspartic proteases such as cathepsin D, which was a clear marker for breast cancer cell lines, T47D. Many other hits were also identified, which may be secondary or tertiary interactors of aspartic proteases, or yet unreported off-targets of the hydroxyethylamine pharmacophore. This provides a candidate list of biomarkers for further investigation. Taken together, this SMM-Cocktail A*f*BP strategy should potentially be useful for high-throughput development of novel probes and the identification of new aspartic proteases as well as related biomarkers in future.

5.2 Introduction

Approaches in functional proteomics have classically included two-dimensional gel electrophoresis (2DE),[93] liquid chromatography-mass spectrometry (LC-MS)[94] and DNA microarray technology.[95] These methods have been applied to the large-scale characterization of cellular, subcellular or circulating proteins and biomarkers.[106] However, most such technologies are geared towards detecting changes in protein abundance, rather than actual functional activity.[4a,97] To address this gap, activity-based protein profiling (ABPP) has emerged as a powerful chemical proteomic alternative for assessing the differing functional states of enzymes and other protein sub-classes within a complex proteome.[62a,62b,98] To date, small molecule probes (the so-called activity-based probes, or ABPs) needed for ABPP studies have been successfully developed for a wide range of mechanistically distinct enzymes, such as serine hydrolases, cysteine proteases, aspartic proteases, metalloproteases,

phosphatases, kinases and others.[27a,41,45,65a,68,82,83b,99-104] Gel-based separation of complex samples labeled with ABPs however limits its utility in that only a few probes, or a narrow pool of samples, may be screened at any given time. Small molecule microarrays (SMMs), on the other hand, offer high-scalability in the assessment of ligand-protein interactions.[4c,88a,105] We and others previously described SMM-facilitated screenings, using combinatorial small molecule libraries, for the screening of purified enzymes and proteins.[106] This included a recent example for the high-throughput identification of affinity-based probes (A*f*BPs) against ɤ-secretase,[100] an aspartic protease implicated in the progression of Alzheimer's disease.[8a] More recently, we have successfully demonstrated the first example of SMMs for large-scale functional profiling of cysteine proteases present in native biological samples, including lysates from mammalian cells and parasite-infected red blood cells (RBCs).[107] Herein, we extend this SMM approach to the screening of a variety of mammalian cellular lysates from numerous tumor cell lines using small molecule probes derived from hydroxyethylamine-based pharmacophores. With the type of throughput provided by SMMs in the initial screening, we were able to quickly identify selected A*f*BPs suitable for the large-scale profiling of aspartic proteases and other potential biomarkers in mammalian proteomes. Additionally, unknown cellular targets identified from screening such probes may provide opportunities for future applications of hydroxyethylamine-based inhibitors in novel therapeutics (Figure 5.1).

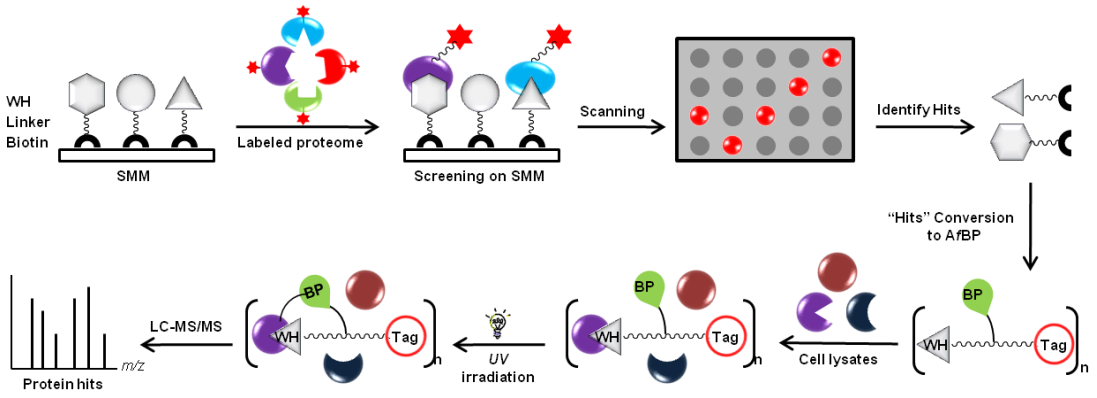


Figure 5.1 Overall strategy for small molecule microarray (SMM)-facilitated development of A*f*BPs for proteome profiling of aspartic proteases in mammalian cell lysates.

Aspartic proteases are the smallest of all protease classes in the human genome, with only 15 members documented thus far.[108] These enzymes, however, have received considerable attention as potential therapeutic targets since many have been shown to play important roles in the physiological and pathological processes.[6,7,9,10] For example, HIV-1 protease is an important target in the fight against AIDS.[10] Renin is a target in the treatment of hypertension.[6] Cathepsin D is implicated in breast cancer metastasis.[7] Plasmepsins are key mediator in malaria pathogenesis.[9] Aspartic proteases are characterized by two catalytic aspartic acid residues located in their active sites and in most cases, by the conserved Asp-Thr-Gly (DTG) sequence in the primary structure. Although the catalytic mechanism of aspartic proteases is poorly understood, it has been generally accepted that the aspartic residues in the enzyme active site bind a molecule of water through extensive H-bonding. Of the numerous known aspartic protease inhibitors, hydroxyethylamine-containing compounds are amongst the most potent and successful, some of which are already FDA-approved drugs.[15,109] For example, Amprenavir™ and Nelfinavir™ are just two examples of

hydroxyethylamine-based drugs currently used for treatments of AIDs (Figure 5.2A). By mimicing the transition state formed during the proteolytic reaction of aspartic proteases, the hydroxyl group on the hydroxyethylamine isostere binds tightly to the enzyme active site through H-bonding with the catalytic aspartic residues, thereby replacing the active site-bound water molecule. We previously showed that this pharmacophore could be strategically and conveniently diversified to generate inhibitor libraries that vary structurally at the corresponding P_2, P_1, P_1' and P_2' positions (Figure 5.2B).[100]

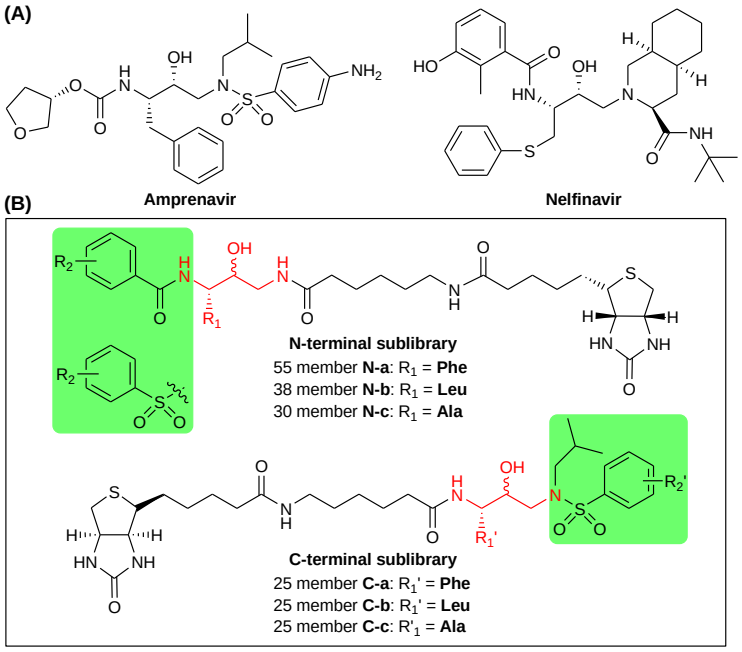


Figure 5.2 Hydroxyethylamine-derivated inhibitors targeting aspartic proteases. (A) Representative structures of two commercially available drugs against HIV protease, an aspartic protease. (B) The 184-member library against aspartic proteases based on the hydroxyethylamine pharmacophore.[100] In both N- and C-terminal sublibraries, the $P_{1/1'}$ positions were diversified by building blocks derived from three amino acid residues (Phe, Leu and Ala; in red). The $P_{2/2'}$ positions (highlighted in green boxes), were diversified by various carboxylic acids and sulfonyl chloride building blocks.

In the current study, with the aim to gain further insights into potential

hydroxyethylamine-interacting proteins in various mammalian proteomes, and to discovery new cancer biomarkers, we conceived a multi-step approach, as depicted in Figure 5.1; in the first stage, an SMM platform was established, on which structurally diverse library of hydroxyethylamine inhibitors, with each containing a biotin tag (and a connecting linker) for SMM immobilization, would be used to probe the binding profiles of endogenous aspartic proteases and other proteins from complex proteome samples. We had recently shown this is feasible with fluorescently labeled crude proteome lysates.[107] Subsequently, positive hits (e.g. bright spots) obtained from the SMM screening were identified, and the corresponding compounds were conveniently converted to affinity-based probes (A*f*BPs) via the Cu(I)-catalyzed 1,3-dipolar cycloaddition (Click chemistry or CuAAC).[4b, 110] This step was made possible due to the modular design of our compound libraries and the introduction of a Benzophenone (Bp) and a reporter tag to the vicinity of these hydroxyethylamine pharmacophores would enable the covalent capturing of any potential interacting proteins. In the next step, the resulting A*f*BPs, either as individual probes or a mixture (or "cocktail") of probes, would be used for gel-based proteome profiling against mammalian cell lysates. Finally, large-scale pull-down experiments followed by mass spectrometric analysis with selected probes would allow positive identification of hydroxyethylamine-binding protein targets which might be further developed into potential cancer biomarkers.

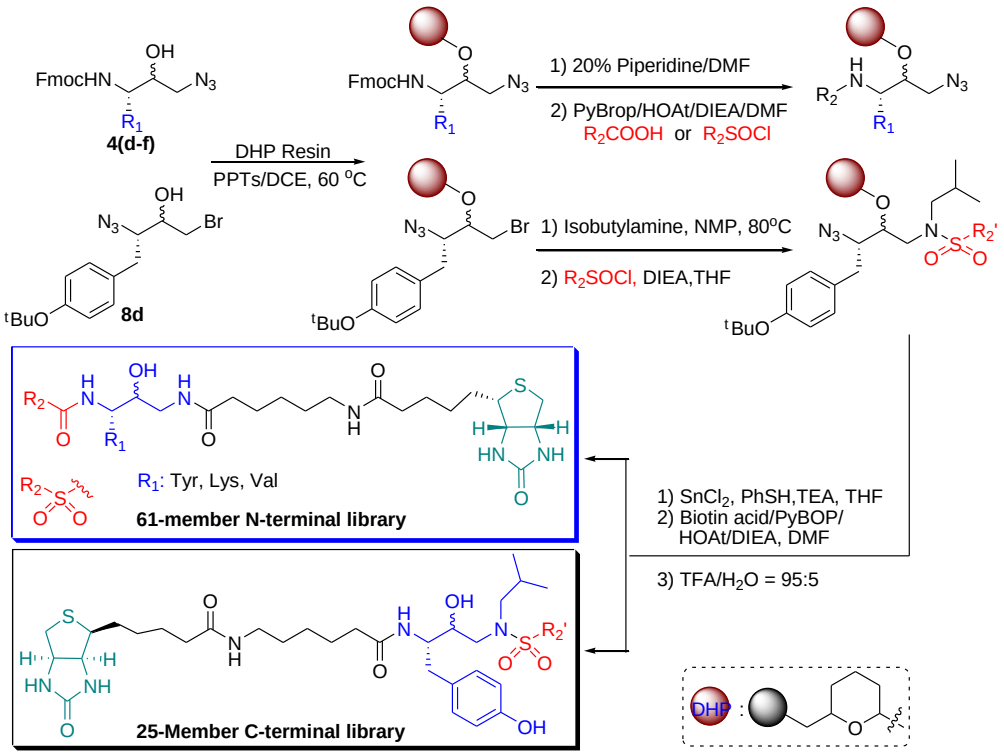
5.3 Results and Discussion

5.3.1 Design and Synthesis of Hydroxyethylamine-containing Inhibitor

Library

Hydroxyethylamine isosteres are transition-state mimics that were introduced as the backbone replacement of the amide bond located between the $P_1/P_{1'}$ position. We previously developed a solid-phase synthetic strategy capable of diversifying all four positions around the hydroxyethylamine core, i.e. $P_1/P_{1'}$ and $P_2/P_{2'}$ residues.[9] From that study, two sublibraries (N-terminal and C-terminal as shown in Figure 5.2B), with a total of 184 different compounds were already available. To furtehr expand the coverage of our library compouds to be spotted on the SMM, an additional eighty-six new compounds, were prepared in the current sutdy (Scheme 5.1). In the N-terminal sublibrary, the R_1 position, corresponding to the P_1 residue upon binding to aspartic proteases, was varied with three amino acids (Tyr, Lys and Val). The R_2 position, corresponding to the P_2 residue, was diversified with 24 aromatic/aliphatic moieties. In the C-terminal sublibrary, the $P_{1'}$ residue (denoted as R_1' in the inhibitor library) was fixed with Tyr for its general importance at this position amongst many known aspartic protease substrates/inhibitors. The $P_{2'}$ residue (denoted as R_2') were simiarly varied with a variety of commercially available sulfonyl chlorides. As shown in Scheme 5.1, the synthesis of these compounds followed our previously reported solid-phase approaches on different pre-loaded 3,4-dihydro-2H-pyran (DHP) resins.[110] Starting from scaffolds 4d-f and 8d, which were synthesized from the corresponding commercial available *L*-amino acids following previous reported procedure with some modifications,[51a,111] compounds 4d-f and 8d were immobilized at their –OH groups onto the DHP resin under acidic conditions, giving preloaded resins for N- and

C-terminal sublibraries, respectively. Next, for the N-terminal sublibrary, Fmoc deprotection of the resin was carried out, followed by standard acylation with 24 different acid building blocks. Subsequent azide reduction, which yielded the corresponding amines, attachment of a biotin linker and TFA cleavage yielded the 61-member N-terminal compounds. For the synthesis of C-terminal compounds, the corresponding preloaded DHP resin was first alkylated with *i*-butylamine, then acylated with commercially available sulfonyl chloride building blocks. Subsequent azide reduction $SnCl_2$/PhSH provided the corresponding amines, which upon coupling with the biotin linker and TFA cleavage, afforded the 25-member C-terminal compounds. The resulting 86 compounds were analyzed by LCMS and shown to be of sufficient purity (>80%) for subsequent microarray immobilization and screeening.



Scheme 5.1 Solid-phase synthesis of the N- and C-terminal libraries using preloaded DHP resins. A total of 86 new hydroxyethylamine-derived compounds were syntheisized.

5.3.2 Profiling of Mammalian Cell Lysates on SMMs

Above 86 compounds were combined with the 198 previously reported compounds, giving a total of 284 hydroxyethylamines. They were robotically spotted onto avidin-functionalized slides in duplicate to generate the corresponding small molecule microarray. Eight different human cell lines, seven of which are derived from tumor cells (e.g. human normal kidney cell line HEK293T, Ovary cancer cells

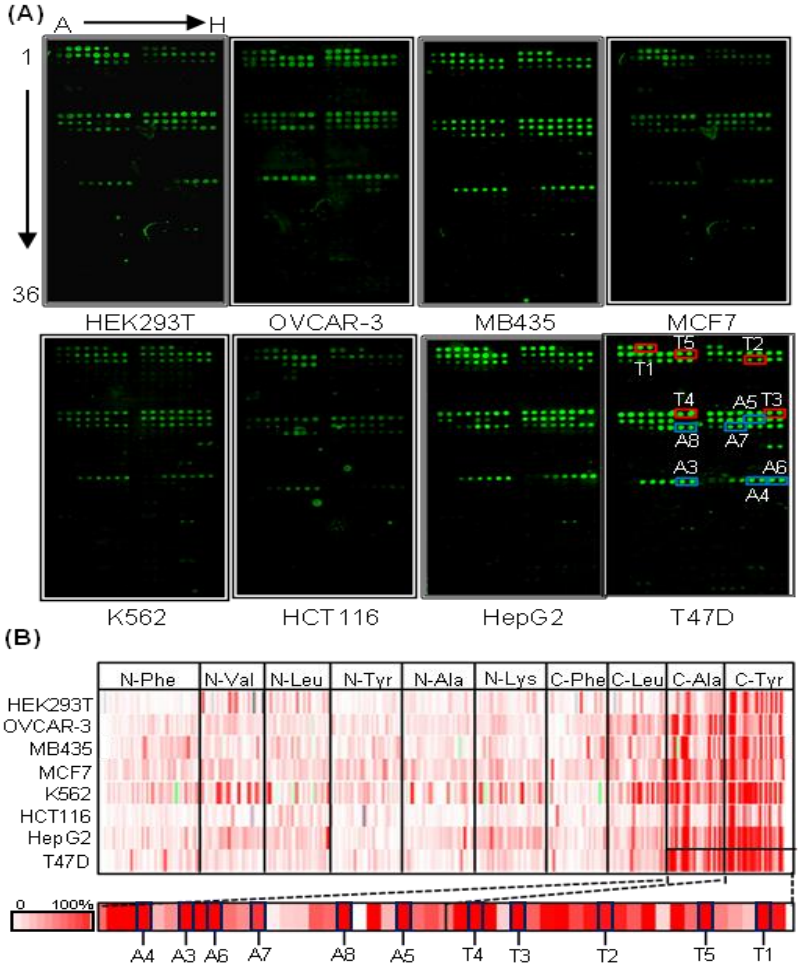


Figure 5.3 SMM profiles obtained by screening with eight different mammalian cell lysates. (A) The binding profiles of Cy3-labeled mammalian cell lysates obtained from screening against the 284-member hydroethylamine SMM. All 284 compounds were spotted in duplicate. Positions of 11 brightest inhibitor spots were highlighted on the SMM. (B) Color heat maps displaying binding of the 284-member compound library against eight Cy3-laleled mammalian cell lysates. The zoomed heat map shows inhibitor potencies of the C-terminal sublibraries presenting Ala and Tyr at the $P_{1'}$ position. The scale bar represents the relative potency ratio.

OVCAR-3, erythromyeloblastoid leukemia cells K562, colon cancer cells HCT116, hepatocellular liver carcinoma cells HepG2, and three breast cancer cell lines MDA-MB-435, MCF7 & T47D), were selected for SMM screening. Briefly, cellular proteome lysates from these cell lines were minimally labeled by an amino-reactive Cy3 dye and directly applied onto the microarray, based on previously developed procedures.[107] To ensure that all fluorescence signals were indeed attributable to activity-dependent interactions between the hydroxyethylamine inhibitors and their targeted proteins, the heat-denatured form of representative cellular lysates (i.e. T47D) was routinely applied to the same array as a negative conrol experiment (Figure 5.4). As shown in Figure 5.3A, all eight mammalian cell lysates produced highly distinctive and similar SMM binding profiles, with most of the brightest spots originated from selected members of the C-terminal sublibrary. We speculated that this is due to the presence of some commonly targeted proteins in both normal and cancer cells. Colored heat maps were generated from the corresponding SMM images in order to better visualize and compare different binding fingerprints obtained from the eight different cell lines (Figure 5.3B); each mammalian cell lines displayed characteristic patterns, and closer examination further revealed that spots which contained hydroxyethylamine compounds having Tyr and Ala at the $P_{1'}$ position provided the strongest binders. However, the effect of different substituent groups at the $P_{2'}$ position had a much less dramatic effect on the binding signals observed. Taken together, the SMM screening results provided a rapid screening tool to quickly delineate protein-ligand interaction and identify potential “hit” compounds for

subsequent follow-up studies. Additionally, the unique binding fingerprints obtained from these SMM experiments may also provide uesful clues for future design and diversification of hydroxyethylamine pharmacophores.[122]

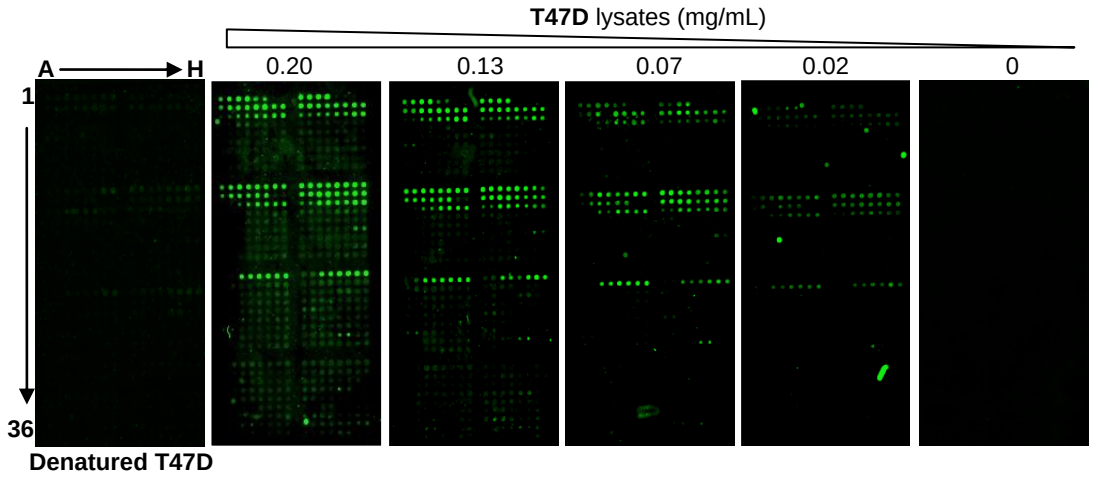
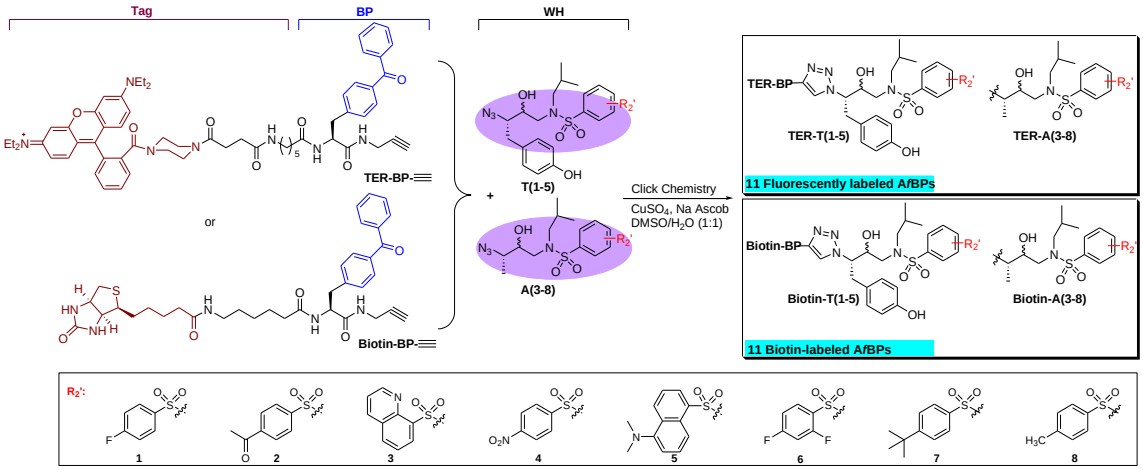


Figure 5.4 Dose-dependent screening of T47D lysate on SMM.

5.3.3 Click Assembly of Select A*f*BPs

To further examine the positive cellular protein targets that exhibited strong binding to members of our hydroxyethylamine library, the eleven brightest spots (boxed in Figure 5.3A) were chosen, and converted into the corresponding affinity-based probes (A*f*BPs). As earlier mentioned, this step was made possible due to the modular design of our compound libraries, and the implementation of Click chemistry into our design principle. The highly efficient and modular Cu(I)-catalyzed cycloaddition reaction between alkynes and azides (e.g. click chemistry) had previously been successfully exploited to facilitate the synthesis of other classes of activity-based probes.[4b,110] As shown in Scheme 5.2, selected azide-containing WH, **T(1-5)** and **A(3-8)**, which were actually key intermediates in the solid-phase synthesis of our hydroxyethylamine libraries (Scheme 5.1), were

"click" assembled with one of the two alkyne-containg photo-crosslinkers, TER-Bp-≡ and Biotin-Bp-≡, giving 11 fluorescently labeled A*f*BPs (for subsequent in-gel fluorescence-based activity-based profiling) and 11 biotin-labeled A*f*BPs (for subsequent pull-down/target identification). The introduction of the photoreactive moiety benzophenone (Bp) enabled us to covalently capture potential protein binders from within the proteome and perform MS analysis for *de novo* target identification. A variety of ''click'' conditions were investigated; eventually, conditions (0.1 eq $CuSO_4$/0.4 eq NaAsc with DMSO/H_2O as co-solvents) were found that afforded the desired 22 A*f*BPs with excellent yields and purities, within 12 h.



Scheme 5.2 Hit-to-probe conversion by click chemstry and structures of the twenty-two A*f*BPs.

5.3.4 Activity-Based Profiling of the A*f*BPs Library with Recombinant Cathepsin D

To demonstrate the ability of our SMM-derived A*f*BPs for UV-initiated, activity-based profiling of aspartic proteases, we first carried out in-gel fluorescence scanning of recombinant cathepsin D labeled with each of the eleven

tetraethylrhodamine (TER) A*f*BPs. Cathepsin D is a well-known lysosomal aspartic protease in mammalian cells and has been reported to play an important role in protein degradation and tumour progression. Following previously published protocols with some modifications,[100] individual probes were incubated with cathepsin D for 30 min before being exposed to UV irradiation (at ~350 nm) for 25 min. This was followed by SDS-PAGE separation and in-gel fluorescence scanning. As shown in Figure 5.5A, one distinct labeled band, which corresponded to cathepsin D, was highly visible across all eleven probes, with varying labeling intensities. It thus indicated that all selected probes could positively label recombinant cathepsin D. The labeling intensity of each probe also indicated the relative binding preference of the enzyme to the corresponding hydroxyethylamine pharmacophore. In order to investigate the detection limit of the labeling reaction, varied amounts of recombinant cathepsin D were labeled by a cocktail of the eleven probes (that is, a mixture of equal amount of each probe; Figure 5.6A); results showed proportional increases in flureoscence intensity of the labeled cathepsin D with increased amounts of the enzyme. We estimated that using this method, as little as 1.5 pmol of cathepsin D could be readily detected.

5.3.5 *In vitro* Proteome Profiling

To further investigate potential cellular binding targets of the probes within complex mammalian proteomes, we next carried out *in vitro* proteome profiling experiments. Four TER probes (**T1**, **T3**, **A3** and **A6**) were used to label the eight aforementioned mammalian cell lysates. Briefly, each probe (1 μM final probe

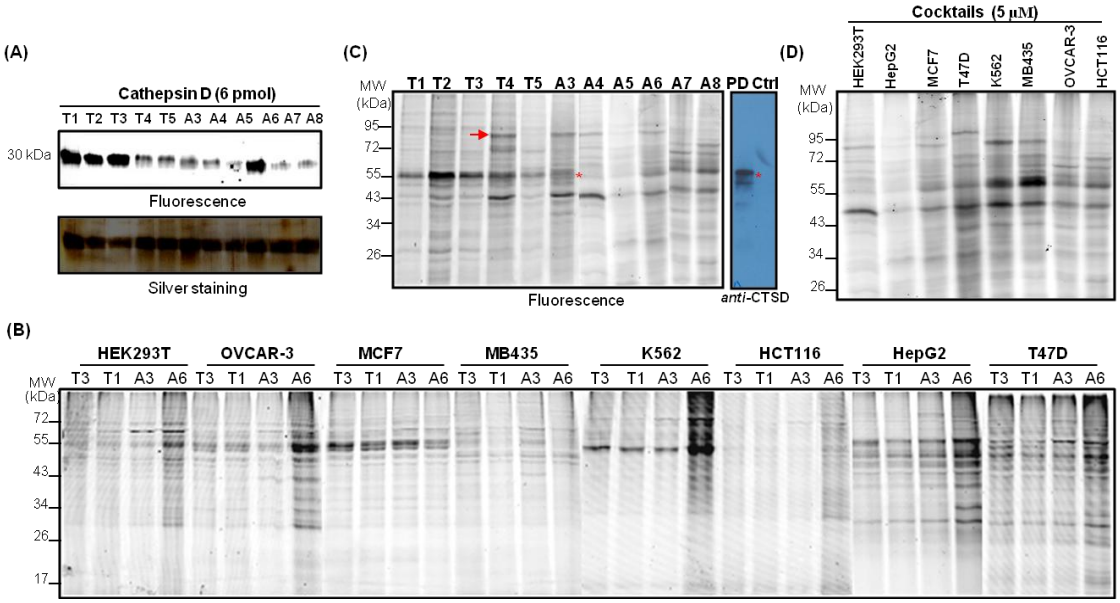


Figure 5.5 In-gel fluorescence images of recombinantly purified Cathepsin D and mammalian cell lysates labeled with different A*f*BPs (individual probe or/and mixture cocktails). (A) Fluorescence labelling profiles of recombinant cathepsin D with 11 individual fluorescently labeled TER probes. (B) Fluorescence labelling profiles of eight mammalian cell lysates with four representative TER probes (**T3**, **T1**, **A3** and **A6**; structures are shown in Scheme 5.2). The final concentration of each probe used in the labelling reaction was 1 μM. (C) *In vitro* proteome profiling of T47D cell lysates with 11 individual TER probes. The red asterisk indicates the protein band corresponding to cathepsin D, which was confirmed by pull-dwon/immunoblotting with *anti*-cathepsin D antibody (right gel). The red arrow indicates the distinct 90-kDa protein band that appeared only in selected cell lines. (D) *In vitro* proteome profiles of 8 different mammalian cell lysates with a mixture cocktail of 11 TER probes. The final concentration of each probe within the cocktail was 0.5 μM. Abbreviations: PD = positive pull-down assay; Ctrl = negative pull-down assay (avidin beads without probe).

concentration; equal amount of lysates was used in each labeling reaction) was treated with each of the eight mammalian cellular lysates for 30 min followed by 25-min UV irradiation. Subsequent SDS-PAGE separation and in-gel fluorescence scanning revealed the sub-proteome binding profiles of the mammalian lysates as a result of positive interactions between endogenous proteins and the hydroxyethylamine probe. As shown in Figure 5.5B, different human cell lines were each distinctly labeled desipte obvious similarities. For example, in T47D

and HepG2 cancer cell lines, many more positively labeled bands were observed, indicating significantly more positively interacting proteins that were present in the total cellular proteome of these cell lines. On the contrary, far fewer fluorescently labeled bands were observed for MDA-MB-435 and HCT116 cell lines (which was consistent with their earlier comparatively weaker fluorecense binding signals on the SMM; Figure 5.3). It is also interesting to note that, while the four TER probes did generate mostly similar proteome profiles within each mammalian cell line as expected, relative fluorescence intensities amongst the labeled bands varied from one probe to another.

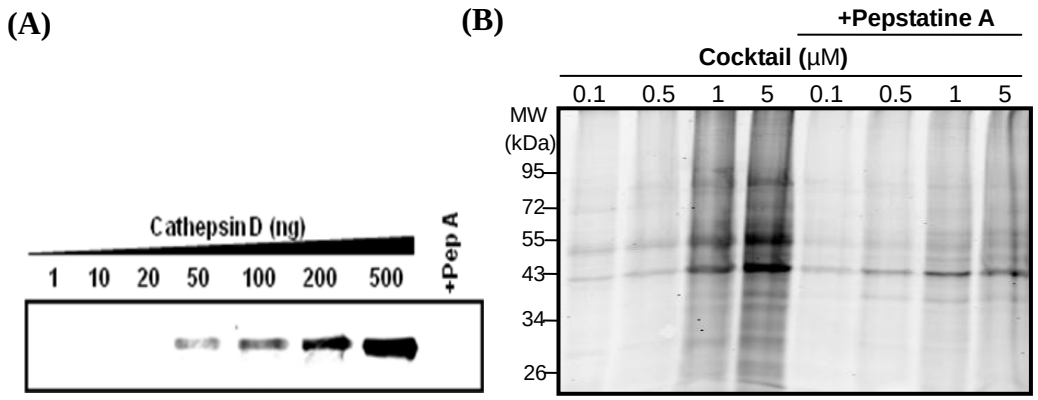


Figure 5.6 Proteome profiling of recombinant cathepsin D and T47D cell lysates by the probe cocktail. (A) Fluorescence labeling of recombinant cathepsin D was performed with the probes as a cocktail, and competitive experiments with pepstatine A. (B) T47D cell lysates was labeled by different amounts of the probe cocktails (0.1-5 μM) with (right) and without (left) the competition of pepstatine A (100 μM).

Next, all eleven probes were taken to label T47D cell lysates in order to obtain comprehensive proteome "fingerprinting" profiles as a result of hydroxyethylamine/protein binding interaction (Figure 5.5C); each probe again showed highly similar yet distinctive proteome labeling profiles, indicating that

the both $R_1/R_{2'}$ residues exerted a considerable influence over probe interactions with cellular protein targets. The experiments also revealed that each probe had different reactivity profiles. For example, a 90-kDa protein showed extremely strong labeling with probe **T4**, **A3**, **A4** and **A6**, but no labeling was observed for other probes (labeled with an arrow in Figure 5.5C). However, a 55 kDa-protein was positively labeled by most of probes except probe **A4** and **A5**. To validate whether this 55 kDa labeled band corresponded to endogenous cathepsin D, which is widely over-expressed in human breast cancer cells, we performed pull-down experiments using the biotinylated probe of **T2** which had the best labeling performance for T47D cell lysates based on the labeling profile. Upon labeling with probe, T47D cell lysates was subjected to affinity enrichment and subsequently analyzed by immunoblotting with cathepsin D antibody (right gel in Figure 5.5C); results clearly showed that cathepsin D was successfully labeled and isolated by the biotinylated probe of **T2**. This experiment thus demonstrates the feasibility of our affinity-based probe for profiling of endogenous cathepsin D in cellular lysates. At this stage, attempts were not made to determine whether other fluorescence labeled bands in our proteome profiling experiments corresponded to other endogenous aspartic proteases.

While the detailed proteome profiling analysis of T47D cell lysates with the 11 individual probes provided useful information about the interaction between cellular proteins and the hydroxyethylamine pharmacophore, the experiment itself was laborious, severely limiting its application in large-scale proteome profiling

against multiple mammalian cell proteomes. Previously, a cocktail-based approach, in which a mixture of several activity-based probes instead of individual probes, had been successfully applied in the profiling of metalloprotease activities in complex proteomes.[113] We speculated that a simiar approach may be adopted herein to increase the throughput of our profiling experiments, therefore enabling much more comprehensive proteome profiles of multiple mammalian cell lysates to be obtained. We therefore applied an equal amount of all eleven A*f*BPs as a single mixture cocktail in our profiling experiments. By applying the same labeling protocol, the cocktail (comprising 0.5 μM of each probe, giving an final probe concentration of ~5 μM) was applied to each of the eight mammalian cellular lysates (Figure 5.5D); in general, cancer cell line K562 and three human breast cancer cell lines, namely MCF-7, MDA-MB-435 and T47D, showed similar proteome labeling profiles, with the exception of an 100-kDa band which appeared only in the T47D cell line. Similar proteome labeling profiles were also observed for ovary cancer cell line OVCAR-3 and colon cancer cell line HCT116. Comparatively weak proteome labeling profiles were seen for HEK293T and HepG2 cell lines. Collectively, the cocktail approach shown herein indicates it may be more widely adopted in future for large-scale proteome profiling experiments, and in some cases to differentiate different types of mammalian cell lysates.

5.3.6 Pull-Down Experiments and Target Identification

Next, in order to identify the cellular targets of the known hydroxyethylamine

pharmacophore, we performed affinity pull-down experiments with the biotinylated version of the probe cocktail. The mixture of 11 biotinylated probes (0.5 μM of each probe) was incubated with T47D cell lysates for 30 min, followed by 25-min UV irradation to initiate covalent cross-linking of interacting protein-probe pairs. Upon precipitation with acetone and washing with methanol, the resulting biotinylated proteins were reconstituted with SDS buffer, then affinity-enriched with Neutravidin™ beads. Following SDS-PAGE gel separation and silver staining (Figure 5.7), the entire gel lane from the pull-down sample was then cut into slices and each slice was independently processed for in-gel trypsin digestion as described in the experimental section. As a negative pull-down control, the entire experiment was repeated with the same cell lysates treated with 5% DMSO instead of the actual probe cocktail. Following LC-MS/MS analysis of the trypsin-digested samples, all proteins were identified by setting a minimum

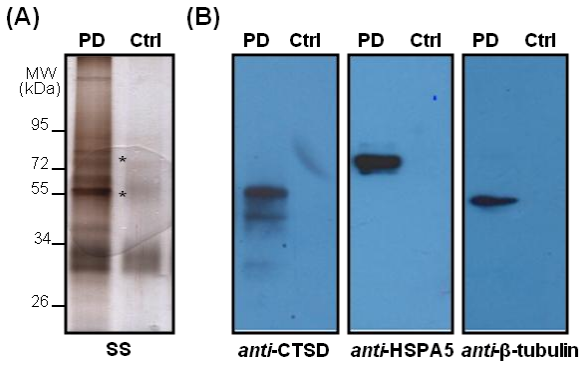


Figure 5.7. Pull-down results using the biotinylated probe cocktail and target validation. (A) Silver-stained gel of the samples from pull-down experiment. Asterisks show the expected locations of Cathepsin D, HSP5A and β-tubulin. (B) Western blotting analysis of pulled-down fractions of T47D cell lysates treated with the biotinylated probe cocktail (or negative controls without the cocktail; right lanes) with their respective antibodies. Abbreviations: SS, Silver stain; PD, positive pull-down assay; Ctrl, Neutravidin™ beads without probes (5% DMSO instead).

MOWSE score of 40 as the cut-off threshhold. In addition, proteins that appeared in the negative control pull-down experiment (beads only) were excluded as potential false positives. In order to further minimize false positives in our protein list and obtain more reliable putative hits, we also performed competitive pull-down experiments in the presence of pepstatin A which is a well-known general aspartic protease inhibitor. Briefly, the entire positive pull-down/LCMS experiment was repeated, with the introduction of a pre-incubation step of T47D lysates with pepstatin A (100 μM) for 30 min, prior to the addition of the probe cocktail. By further substracting protein list obtained from the pepstatin A-competitive pull-down experiment from that of the positive pull-down experiment, a total of 39 proteins were finally identified as highly confident putative proteins hits (Table 5.1). It should be noted that all MS-based results obtained herein (including all the putative protein hits listed in Table 5.1) should only be used as preliminary data. In our current study, deliberate efforts were made to improve our pull-down/LCMS results. Nevertheless, due to the highly complex cellular environment, the intrinsic limitation of affinity pull-down assay and mass spectrometry,[114,115] false positives/non-specific proteins binders could be minimized but not eliminated entirely from our results. Consequently, proper follow-up studies and validation experiments will be needed before any biological conclusion can be made to some of these protein hits. Of the 39 identified proteins, cathepsin D was one of them. This result is well expected as our previous immunoblotting results had independently verified this (Figure 5.5C). It is

interesting to note that missing from the list are other known mammalian aspartic proteases, which might have been casued by their relatively lower expression level in the T47D cell line. Among other proteins identified, both HSP5A and β-tubulin are two of the well-known interactors of cathepsin D.[116,117] Their positive interaction with our A*f*BPs were further confirmed by pull-down/western blotting with their respective antibodies (Figure 5.7B). Another two proteins known to be regulated by cathepsin D, namely superoxide dismutase (SOD) and eukaryotic translation elongation factor 1 delta isoform 4 (eEF1),[116] were also identified. Notably, a known pepstatin A-binding protein, disulfide-isomerase (PDI),[118] was also identified with a high MOWSE score. Additionally, a number of other targeted proteins were pulled down including one well-known serine protease (lon protease homolog or LONP1), several ribosomal proteins (60S ribosomal protein L5 or RPL5 and isoform 1 of plectin-1 or PLEC1), RNA recognition proteins (isoform 1 of heterogeneous nuclear ribonucleoprotein M or HNRNPM and Isoform 1 of heterogeneous nuclear ribonucleoprotein Q SYNCRIP), Leucyl-tRNA synthetase, cytoplasmic (or LARS) and DNA-directed RNA polymerase II subunit RPB2 (or POLR2B). At this stage, we do not have evidence to establish at whether these hits directly interacted with our probes and hydroxyethylamine pharmacophores in general, or were captured indirectly through interactions with other proteins. It might indicate that these hydroxyethylamines could potentially interact/target other protein classes besides aspartic proteases, and some of the hits identified herein may be considered potential off-targets of

hydroxyethylamine-derived drugs in cellular environments.[114c] Collectively, it demonstrates that our SMM-facilitated, A*f*BP strategy should potentially be useful for applying inhibitor libraries for comparative proteomic profiling, with the potential for novel and disease-related biomarker identification.

Table 5.1 Proteins identified by pull-down and mass spectrometry.

#	ID	Protein Description	MS	Score	Peptides matched
1	IPI00011229	Cathepsin D	45.0	163	4
2	IPI00023598	Tubulin beta-4 chain	50	1497	74
3	IPI00786995	protein kinase, DNA-activated, catalytic polypeptide	47	801	38
4	IPI00299402	Pyruvate carboxylase, mitochondrial	130.3	531	21
5	IPI00646779	TUBB6 protein	50.5	531	23
6	IPI00017454	Putative tubulin-like protein alpha-4B	27.8	510	18
7	IPI00024993	Enoyl-CoA hydratase, mitochondrial	31.8	412	14
8	IPI00171903	Isoform 1 of Heterogeneous nuclear ribonucleoprotein M	77.7	376	22
9	IPI00020984	highly similar to Calnexin	71.9	289	13
10	IPI00641706	TUBB6 46 kDa protein	46.3	259	22
11	IPI00000494	60S ribosomal protein L5	34.6	244	10
12	IPI00014898	Isoform 1 of Plectin-1	53.3	188	19
13	IPI00037070	HSPA8 54 kDa protein	53.7	187	10
14	IPI00010796	Protein disulfide-isomerase	57.5	181	6
15	IPI00420014	Isoform 1 of U5 small nuclear ribonucleoprotein 200 kDa helicase	24.6	180	6
16	IPI00292496	Tubulin beta-8 chain	50.3	177	17
17	IPI00003362	HSPA5 protein	72.5	175	10
18	IPI00003925	Isoform 1 of Pyruvate dehydrogenase E1 component subunit beta,	39.6	172	5
19	IPI00022891	ADP/ATP translocase 1	33.3	167	9
20	IPI00029737	Isoform Long of Long-chain-fatty-acid--CoA ligase 4	80.2	156	10
21	IPI00005040	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	47	128	4
22	IPI00219682	Erythrocyte band 7 integral membrane protein	31.9	125	5
23	IPI00943207	Ribonucleoside-diphosphate reductase	81.7	125	9
24	IPI00103994	Leucyl-tRNA synthetase, cytoplasmic	135.6	123	5
25	IPI00003348	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	38.05	107	5
26	IPI00027808	DNA-directed RNA polymerase II subunit RPB2	135.2	107	4
27	IPI00022314	Superoxide dismutase [Mn], mitochondrial	24.9	99	6
28	IPI00009634	Sulfide:quinone oxidoreductase, mitochondrial	50.2	94	3
29	IPI00031517	DNA replication licensing factor MCM6	93.8	92	3

30	IPI00413641	Aldose reductase	36.2	89	10
31	IPI00465233	Eukaryotic translation initiation factor 3, subunit E interacting protein	71.1	86	2
32	IPI00005158	Lon protease homolog, mitochondrial	106.9	81	4
33	IPI00021048	Isoform 1 of Myoferlin	23.6	81	2
34	IPI00009253	Alpha-soluble NSF attachment protein	33.7	80	4
35	IPI00018140	Isoform 1 of Heterogeneous nuclear ribonucleoprotein Q	69.8	80	3
36	IPI00290462	Carbonyl reductase [NADPH] 3	31.2	80	4
37	IPI00064086	Eukaryotic translation elongation factor 1 delta isoform 4	28.7	78	2
38	IPI00169383	Phosphoglycerate kinase 1	44.9	73	4
39	IPI00008986	Large neutral amino acids transporter small subunit 1	55.7	72	1

5.4 Conclusion

We have synthesized a new 86-member library of hydroxyethylamine pharmacophore derivated inhibitors in this study. Combined with previous published compound library, we have, for the first time, developed a small molecule microarray(SMM)-facilitated strategy for large-scale profiling of multiple mammalian cell lysates against many potential inhibitors. Based on the valuable SMM fingerprints, a panel of affinity-based probes were rapidly synthesized by “click” chemistry, from the top inhibitor hits. Subsequent *in vitro* proteome activity profiling and large-scale affinity pull-down/LCMS identifications of putative protein hits enabled us to identify the direct as well as potentially indirect targets of the hydroxyethylamine pharmacophore in the samples. Our sub-proteomic SMM–Cocktail A*f*BP strategy thus provides a viable solution towards for functionally profiling and differentiating complex proteomic samples.

Chapter 6

Developing Photo-affinity Probes for Proteomic Profiling of Cellular Targets of Dasatinib™

6.1 Summary

In this chapter we have successfully developed a novel chemical proteomic approach for identification of the celluar targets of dasatinib™ in native environment. Herein, we described the synthesis of two "clickable" dasatinib-like, affinity-based probes and further investigated their biological activities using c-Src as model protein. Our preliminary results demonstrated that these two probes exhibited comparative activities as dasatinib and positively labeled a wide of protein kinases in cellular lysates. More notably, we found that probe **DA-2** had good cell permeability and enabled to label endogenous c-Src, c-Abl as well as some other kinases within live cells. To further comprehensively study dasatinib's cellular binding targets, we performed *in vitro*/*situ* labelings and pull-down experiments with this probe. Consequently, a total of 54 kinases were identified as potential binding targets of dasatinib in K562 and HepG2 cancer cells. Among them, in addition to some known dasatinib binding targets (SRC family kinases, c-Abl, BTK, EGFR, p38α, and so on), several serine/threonine kinases, for the first time, were identified as putative dasatinib binding proteins, and some of them were further validated by pull-downs and immunoblottings. Thus, we envisage that this cell permeable probe, combined

with the bio-orthogonal click chemistry, should be broadly useful for the large-scale profiling of kinases in functional proteomics analyses.

6.2 Introduction

Protein kinases plays an important role in the development and progression of cancer by regulating cell growth, survival, invasion/metastasis and angiogenesis.[18] As one important class of protein kinases, tyrosine kinases are involved in a variety of physiological and pathological processes in cancer, immunology and solid tumor.[16] Thus, tyrosine kinases have become interest targets for biomarker discovery. As an example, Src family kinases (SFKs) have been reported to control a myriad of fundamental cellular processes, including cell proliferation, cell migration, cell invasion and survival.[119] Clinical studies have shown that Src is correlated with malignant progression of human tumors, and it has become a promising therapeutic target in treatment of tumors.[120]

Kinase inhibitors are a new class of therapeutics to inhibit multiple targets.[121] Their biological functions as anticancer therapies are not thoroughly defined thus far. Identification of cellular targets of kinase inhibitors has become an important step toward understanding how inhibitors interact with the human kinome. In a clinical setting, the identification of potential off-targets of kinase inhibitors may help explain any adverse effects that are observed. Dasatinib, an ATP-competitive inhibitor of Src and Abl kinases, has been used for treatment of imatinib-resistant chronic myelogenous leukemia (CML),[23,122] and recently it was also found to have activity to epithelial tumor cells including human prostate and breast cancer cells.[24] Studies on

off-targets of dasatinib, therefore, have become significant for therapy. The traditional method used for determining the binding targets of kinase inhibitors is the activity screening assays with purified kinases *in vitro*.[123] Although this approach has been extensively used in the past and produced important insight in kinase inhibitor specificity, it is impractical for profiling the large size of human kinome. In order to overcome this limitation, some alternative and unbiased approaches have been developed, for example, phage-expressed kinome,[13a, 124] affinity matrix,[13d] and so on. While affinity matrix, namely a resin with immobilized kinase inhibitor, has become currently the most prominent approach for identification of drug binding proteins from cell or tissue extracts. Oliver and his co-workers recently found that Btk and Tec kinases were another two major cellular binding targets of dasatinib besides Abl and Src kinases by using the dasatinib immobilized beads.[126] Additionally, a number of other dasatinib binding proteins have also been constantly identified by other groups using the similar approach.[127,128] However, theses approaches are still greatly challenged in several aspects (Figure 6.1A): first, due to the non-covalent interactions between drugs and proteins, weak washing condition must be applied, which usually introduces many nonspecific proteins; second, the solid support may potentially limit the accessibility of the immobilized drugs for interactions with the target proteins. Some proteins that cannot interact with the attached small molecules because of steric hindrance can be missed easily; third, the *in vitro* activities of proteins don't always reflect their real behaviors in native environment. With the aim to overcome some of these shortcomings, more recently, Bantscheff group have developed an improved

approach in which the immobilized broad-selectivity inhibitors as kinase capturing tools have been used to analyze the interactions of competing "free" dasatinib with their protein targets in solution. A number of previous unknown binding targets of dasatinib were identified besides some known targets.[127b] However, it still has become a great deal of interest in the development of new alternative approaches for proteome profiling of dasatinib binding targets. To address this gap, activity-based protein profiling (ABPP), pioneered by Cravatt, has emerged as a powerful chemical proteomic alternative for determining the bindings between small-molecule drugs and binding proteins under more physiological conditions.[62a] To date, ABPP probes have been successfully developed for many relevant enzyme classes, including serine hydrolases,[129] cysteine proteases,[68] aspartic proteases,[41,83b,100] metalloproteases,[82,101] phosphatases[102a] as well as protein kinases.[27a,103] Inspired by ABPP, we envision that the development of dasatinib-like probes which enable to extensively explore the on/off-targets of dasatinib in live cells may open up a new prospect for cancer treatment in future. To this end, we developed an alternative approach for proteome profiling of the binding targets of dasatinib in native environment. Herein, two dasatinib-like, affinity-based probes, for the first time, were synthesized and the abilities of profiling dasatinib cellular targets were further evaluated.

6.3 Results and Discussion

6.3.1 Design and Synthesis of Dasatinib-like Probes

Based on the well-known activity-based protein profiling (ABPP) approach, we have designed and synthesized two dasatinib-like probes. Our probes essentially have

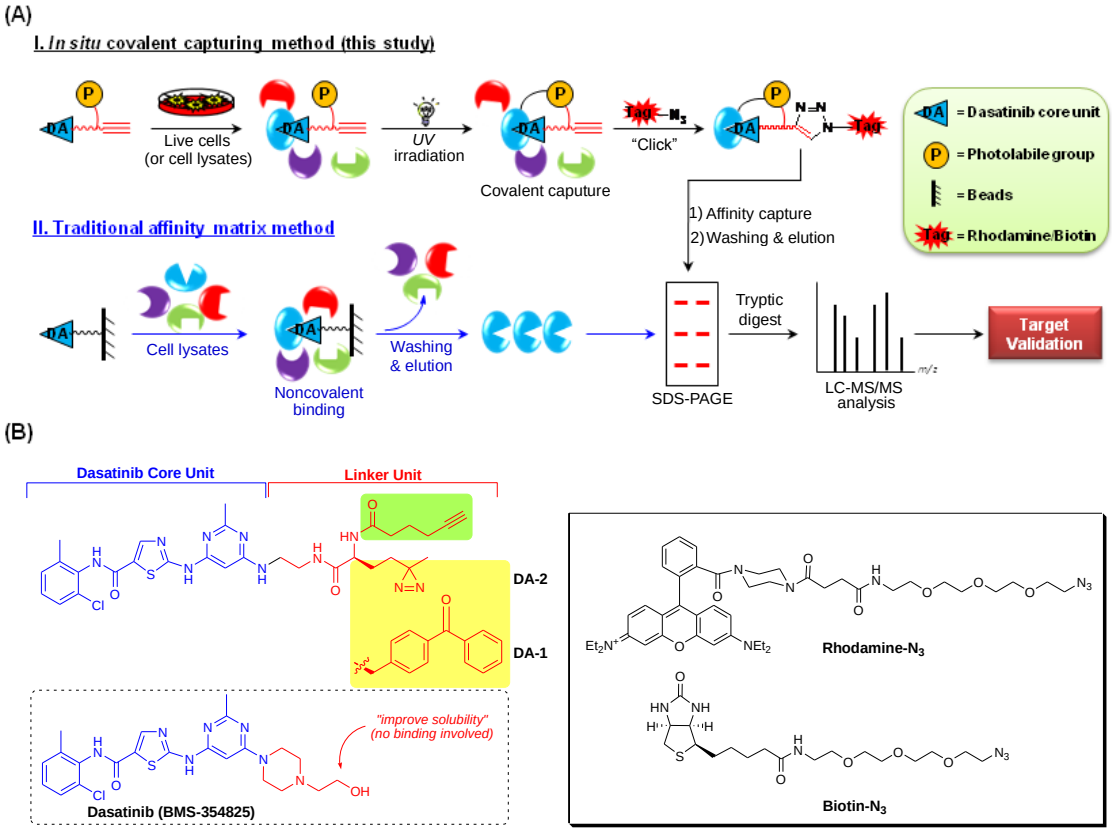


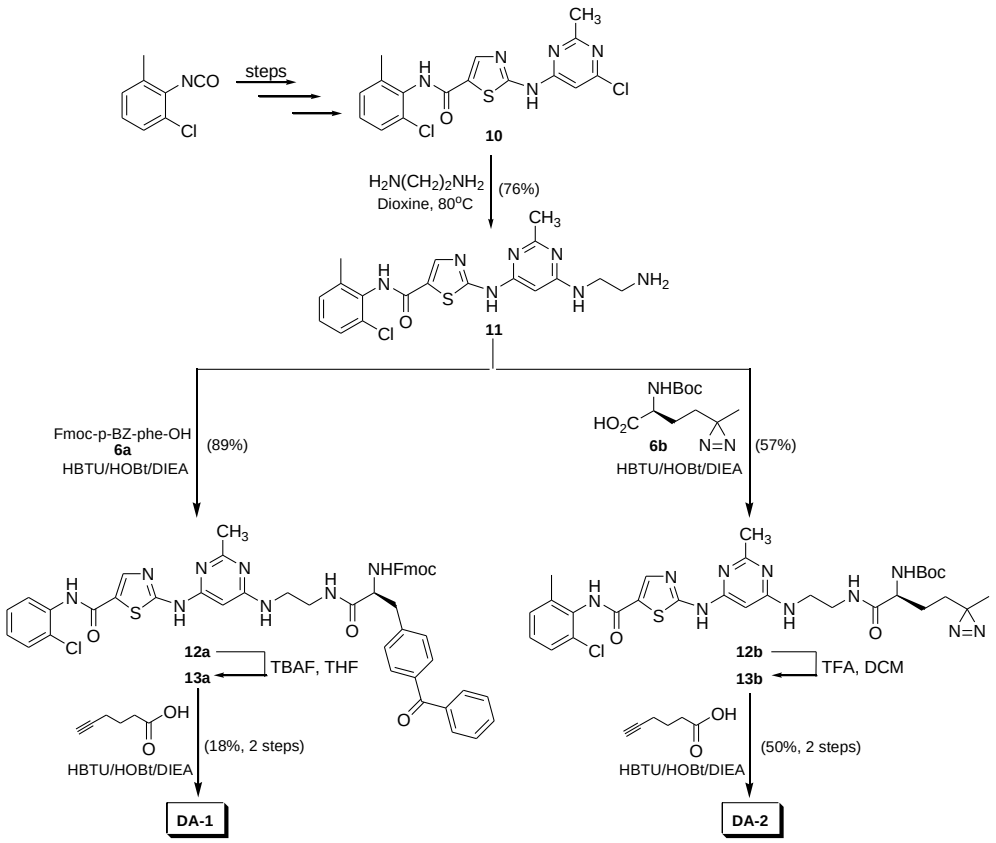
Figure 6.1 Activity-based proteome profiling**.** (A*) In situ* covalent capturing and the traditional affinity matrix method for identification of cellular targets of dasatinib. (B) Design of photoaffinity, dasatinib-like probes. The "Click"-based reporters used in the manuscript are rhodamine azide and biotin azide which were prepared as previously described.[13c]

three structural components (Figure 6.1B): (1) a protein recognition group derived from dasatinib; (2) a photo-reactive group which enables us to covalently capture the potential binders within the proteome and perform MS analysis for *de novo* target identification. In order to study the steric effect of photo-reactive group to the binding between probe and the ATP binding site of kinases, two photo-reactive groups (benzophenone and diazirine) were adopted in this study, and (3) an alkyne handle incorporating to a fluorescent dye or biotin group via the copper catalyzed Huisgen 1,3-dipolar cycloaddition.[34] For the general design of our probes, we took advantage of several key points: (1) dasatinib is a potent Src/Abl inhibitor with very low

nanomolar inhibition. Recent studies have showed that it is able to inhibit a wide of kinases;[127a,13b,130] (2) previously reported crystal structure of the Abl kinase domain complexed with dasatinib indicated that the hydroxyl side chain on the piperazine ring of dasatinib was not involved in the critical H-bonding and it localized at the protein surface. Additionally, prior experiences showed that dasatinib pharmacophore shared much of the same binding characters in the kinase catalytic domain.[130] We speculate replacement of piperazine part of dasatinib with photoreactive group should not disrupt the binding between inhibitor and the ATP binding site of kinase. Therefore, in order to maximally retain the activity of dasatinib, a minimal structure modification has been done in the synthesis of probes. The synthesis of the "clickable" dasatinib derivatives is shown in Scheme 6.1. Compound **10** was first synthesized following previously published method.[122] Next, compound **10** was reacted with 1, 2-diaminoethane in dioxane at reflux to afford **11**, and then subjected to coupling reagents with commercial available Fmoc-protected benzophenone/Boc-protected diazirine (prepared in six steps from Glutamic acid) to give **12a**/**12b**, respectively. Subsequent deprotections of amine protective groups with corresponding conditions, followed by couplings with hex-5-ynoic acid, gave the two probes, **DA-1** and **2** in 18% and 50% yield, respectively.

6.3.2 Molecule Modelling and Determination of IC_{50} Values

To obtain insight into the binding mode of two probes with kinases, we performed molecular docking analysis with c-Src (PDB ID 2SRC27)[131] and Abl (PDB ID 2GQG22)[132] crystal structures using autodock vina method.[133] The docking models



Scheme 6.1 Synthesis of probe **DA-1** and **DA-2**.

revealed that both probes bound well to the ATP binding sites of c-Src as well as c-Abl (Figure 6.2A), which is quite overlaid with dasatinib. In the binding modes of two probes with c-Src and c-Abl, a crucial pair of hydrogen bond was formed between the hydroxyl oxygen of Thr338 (or Thr315) and the carbonyl oxygen of thiazole amide**,** which is consistent with the reported crystal structure of Abl/dasatinib complex.[132] As expected, the photo-cross linkers, localized at the surface of proteins, don't have apparent effect for the interactions between probes and kinases.

Dasatinib is known to inhibit Src and Abl kinases with very low nanomolar IC_{50}. We compared the inhibitory potential of our two probes to dasatinib by measuring the kinase activities of the c-Src and Abl kinase domain in a drug concentration-dependent manner under optimized enzyme concentration of ~50 nM

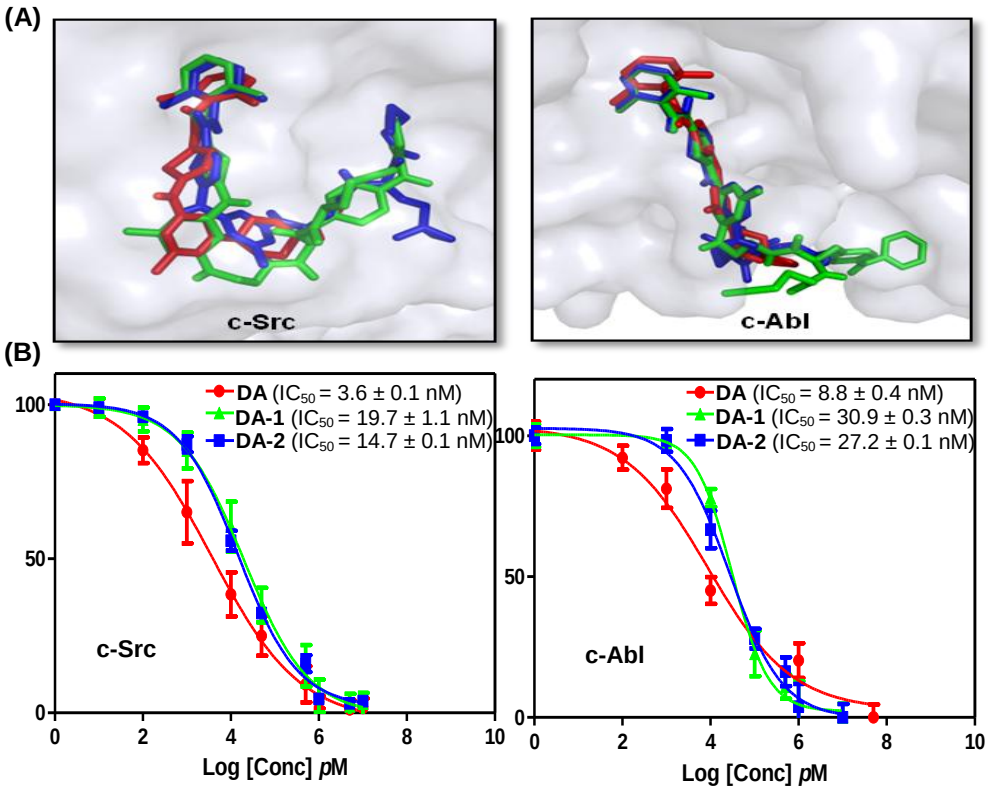


Figure 6.2 Determination of biological activities of **DA-1** and **2**. (A) Molecular docking of **DA-1** and **2** into the ATP binding pocket of c-Src and c-Abl. Dasatinib, **DA-1** and **2** were showed in red, green and blue, respectively. (B) IC_{50} values of two probe (**DA-1** and **2**) and Dasatinib for c-Src and c-Abl determined by Kinase-Glo® Plus Luminescent Kinase. Data represent the average (s.d. for two trials).

(Figure. 6.2B). Both probes maintained excellent potency versus c-Src and c-Abl, and were slightly less effective than dasatinib, which indicates the replacement of piperazine ring by photo-reactive group does not significantly alter the activities of probes. The observed IC_{50} values for c-Src inhibition by **DA-1** (19.7 nM) and **DA-2** (14.7 nM) closely resembled dasatinib (3.6 nM).[134] Similarily, the IC_{50} values for c-Abl inhibition by **DA-1** (30.9 nM) and **DA-2** (27.2 nM) were also close to dasatinib (8.8 nM).

6.3.3 Effects on Cell Proliferation and Phosphorylation of c-Src

Next, two probes were evaluated against dasatinib (as a positive control) for

potential biological activity. We carried out the cell proliferation assay using HepG2 and K562 as model cell lines. The activities of three compounds (**DA-1**, **DA-2** and dasatinib) against HepG2 and K562 cells were evaluated using XTT assay. Both probes showed inhibition of tumor cell proliferation with comparable potency, while they were slightly less potent than dasatinib (Figure 6.3B). The exact reason for the less cellular potency of the probes is not clear at the moment although we speculate that it could most likely be due either to the structure modification or poor cell permeability. In order to determine the cell permeability of two probes, we performed cell permeability assay together with dasatinib as positive control (Figure 6.3A). Results indicated that probe **DA-2** had comparable cell permeability with dasatinib, whereas **DA-1** was not cell permeable. Next, we performed comparative analysis of the compounds in their abilities to reduce the autophosphorylation of c-Src in Src-transduced CHOK1 cells. It is well known that dasatinib can inhibit the autophosphorylation of Y416 residue which is relative to the catalytic activity of c-Src. As shown in Figure 6.3A, CHOK1 cell treated with different amounts of dasatinib and **DA-2** showed almost similarly reduced c-Src autophosphorylation.

6.3.4 Cellular Imaging using DA-2

To further demonstrate our probes' cell permeability and utility for potential cellular imaging, we performed fluorescence microscopy to visualize the cells treated by probe **DA-2** (Figure 6.3C). Live HepG2 cells were first treated with respective probes followed by irradiation under UV, and then fixed with paraformaldehyde, permeabilized with Triton X-100, conjugated to rhodamine azide by click chemistry, and imaged (colored in red). Immunofluorescence was also carried out

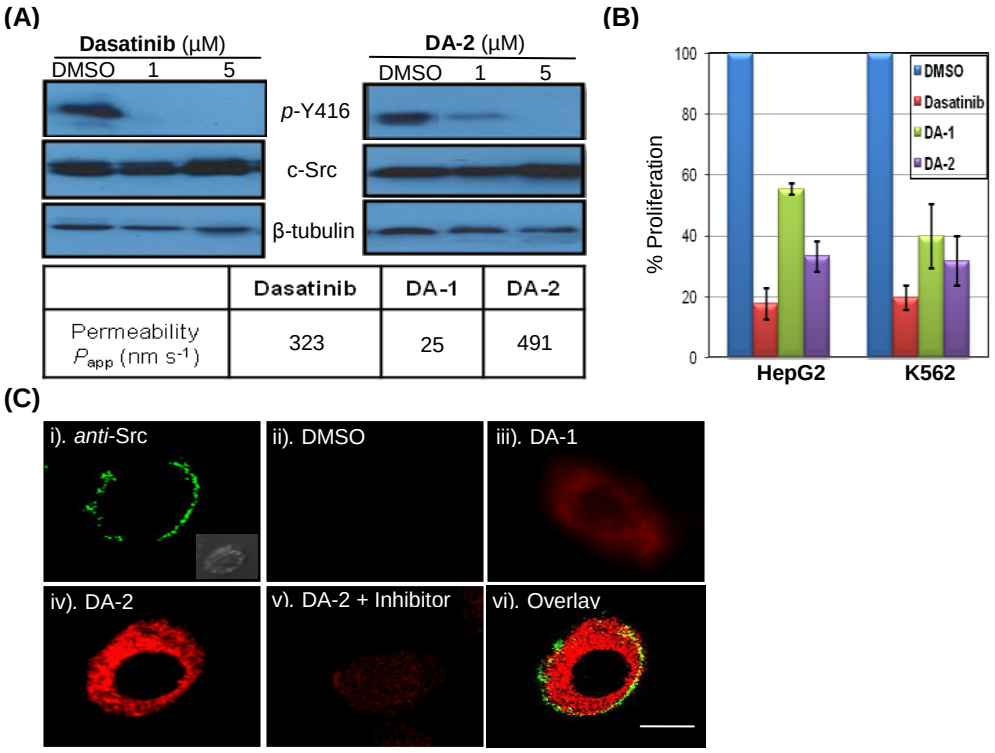


Figure 6.3 Biological activities of **DA-1** and **2** in live cells. (A) Comparison between dasatinib and **DA-2** on the inhibition of tyrosine phosphorylation in cells and the cell permeability assay. (B) Inhibition of HepG2 and K562 cells proliferation by 20 μM dasatinib, **DA-1** and **2** using XTT assay. (C) Cellular imaging of HepG2 cells with probe **DA-1** and **2**. All images were acquired in the same way. Scale bar = 10 μm.

on the same cells to visualize the localization of endogenous c-Src (colored in green). Results showed that minimal fluorescence was observed in cells treated with only DMSO as well as **DA-1**, again demonstrating that **DA-1** has poor cell permeability and labeling performance in intact cells. On the contrary, higher fluorescence was observed for cells treated with **DA-2** in which not only endogenous c-Src but also some other proteins were also labeled by **DA-2**, which is consistent with the multiple binding targets property of dasatinib. Taken together, probe **DA-2** has good cell permeability than **DA-1**, and it can be potentially used for bioimaging in future. Thus, **DA-2** was chosen to use for all the following labeling experiments.

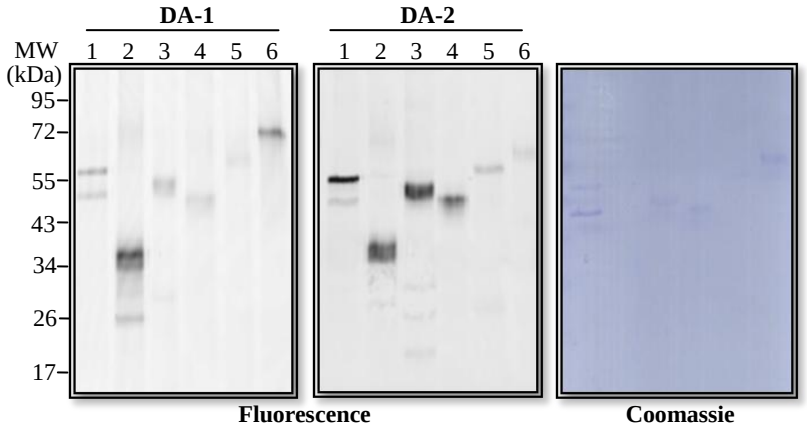


Figure 6.4 Fluorescent Labelling of different recombinant proteins (15 pmol, lanes 1 to 6 are Abl, c-Src, CSK, PKA, PAK6 and BSA, respectively) with **DA-1** and **2** in parallel.

6.3.5 *In vitro* Labeling of Purified Kinases

Next, to evaluate the labeling ability of the probe, dose-dependent experiments were carried out by varying the concentration of kinases in the reactions (Figure 6.5A). **DA-2** was incubated with different amount of c-Src and c-Abl in parallel for 30 min followed by irradiation under UV lamp at ~350 nm. Subsequently, the reactions were subjected to click reaction with rhodamine azide (TER-N_3) (structures shown in Figure 6.1B), after which the fluorescent labeled proteins were separated by SDS-PAGE and visualized in-gel with a fluorescence gel scanner. As expected, we observed increasing fluorescent intensity upon the addition of increasing concentration of kinases, indicating that the intensity of fluorescent bands observed is proportional to the amount of kinases used in the reactions. On the other hand, we could detect as little as 1 pmol c-Src or Abl by fluorescence scanning.

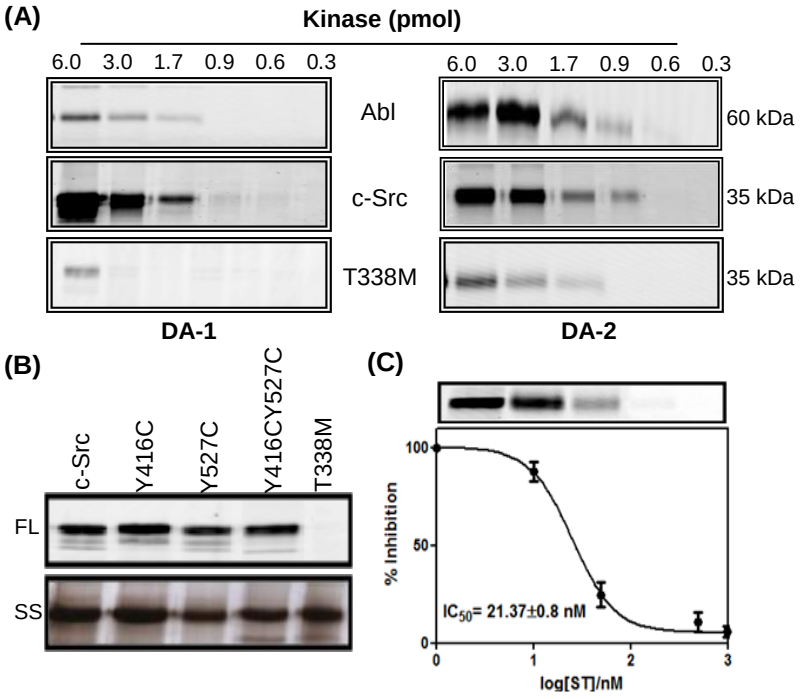


Figure 6.5 Fluorescent labeling of recombinantly purified protein kinases. (A) Labeling profiles of varied amount of c-Src, c-Abl and mutant SrcT338M by **DA-2** as well as **DA-1**. (B) Fluorescent labeling of c-Src and its four mutants (Y416C, Y527C, Y416CY527C and T338M) by **DA-2**. (C) Dose-dependent reduction of labeling of c-Src (100 nM) in the presence of staurosporine (STS). IC_{50} of STS was determined to be 21.37 ± 0.8 nM.

Tyrosine residue Y416 and Y527 are two important determinants for the activity of c-Src. Phosphorylation of Tyr527 in the C-terminal tail is found to down-regulate the kinase activity. On the contrary, phosphorylation of Tyr416 in the activation loop is necessary for full kinase activity. In order to investigate whether the tyrosine residue mutations affect the binding between probe and ATP binding site, four tyrosine mutants (SrcY416C, SrcY527C, SrcY416CY527C and T338M) were created by QuikChange mutagenesis method, and labeled by **DA-2**. As shown in Figure 6.5B, T338M gatekeeper mutation of c-Src resulted in apparent resistance to the probe, which is consistent with previously reported.[135] Other three tyrosine mutants were, however, almost equally labeled by **DA-2**, indicating that mutations of residue Y416

and Y527 don't affect the interaction between probes and ATP-binding site. Indeedly, the previous modeling results also indicated that these two Tyr residues were not involved in the binding between probe and c-Src.[131] Next, we performed labeling experiments in the presence of active-site competitive inhibitor-Staurosphorine. A dose-dependent reduction of fluorescence in c-Src labeling profile was observed (Figure. 6.5C), indicating the labeling is activity-based and points to the possible use of the probe for *in situ* screening of c-Src inhibitors.

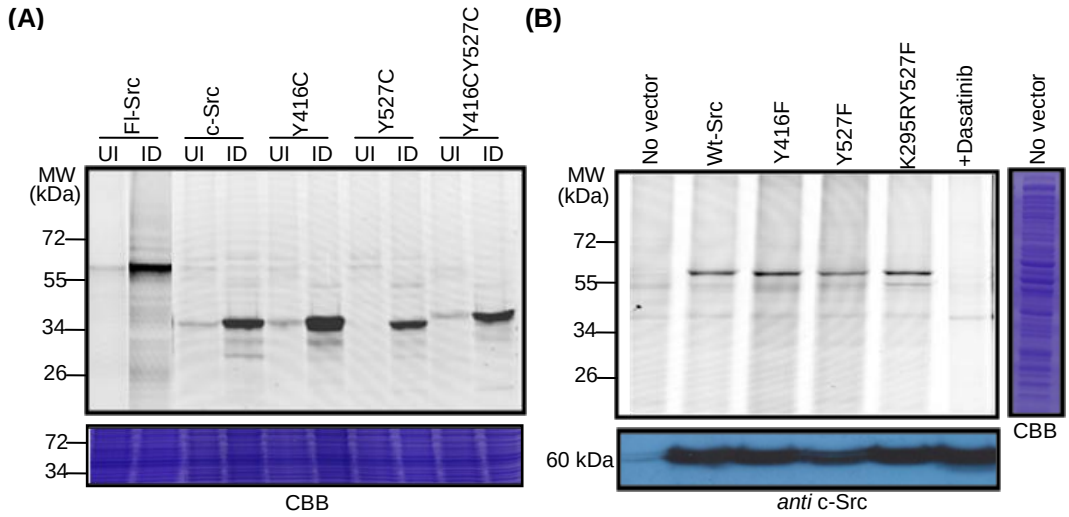


Figure 6.6 Labeling profiles of bacterial and mammalian proteomes. (A) Fluorescent labeling profiles of induced (ID) and un-induced (UI) bacteria expressing full length (FL)-Src, catalytic domain c-Src and its three mutants by **DA-2**. (B) Labeling profiles of CHOK1 lysates transfected with wide type c-Src and its three mutants together with competition experiment. The final concentration of **DA-2** is 1 μM in all experiments.

6.3.6 Labeling of c-Src in Bacterial and Mammalian Proteome

Next, to assess whether probe **DA-2** can selectively label c-Src in the presence of competing cellular proteins, we expressed the wide type c-Src (including full length and catalytic domain c-Src) and its three mutants in *E. coli* following reported method, respectively.[136] The corresponding bacterial lysates were then labeled by

DA-2 under previously optimized condition. As shown in Figure 6.6A, stronger labeling bands were seen for bacterial lysates treated with inducer, while almost no labeling was observed for un-induced lysates. These results indicated that **DA-2** could be potentially used to label and detect the expression levels of c-Src in different biological samples.

Next, in order to further test the ability of the probe to selectively label endogenous c-Src, we examined the reactivity of **DA-2** in CHOK1 cellular lysates which was transfected with the constructs encoding wild type full length c-Src and its three mutants (SrcY416F, Y527F and K295RY527F). The fluorescent profiles showed minimal background labeling of other proteins. We found that **DA-2** impartially labeled wide type c-Src and its three mutants, which were consistent with previously purified protein labeling results, and again indicating that dasatinib has activity against many mutants of c-Src. In addition, immunoblotting analysis with a monoclonal Src antibody identified comparable expression levels of wt c-Src and its mutants (Figure 6.6B). These datas collectively indicate again that Y416 and Y527 mutations in Src family kinase do not fundamentally affect the intrinsic interaction between the ATP binding site of Src family kinases and our probe. Additionally, the labeling could be completely blocked by competing inhibitor if the cells were prior to treat with dasatinib, again indicating the labeling is activity-based.

6.3.7 Labeling of Endogenous c-Src in Cancer Cell Lysates

To further test the ability of our probe to label endogenous c-Src, we carried out the labeling experiments for cell lysates (K562 and HepG2 cells) with **DA-2** in

dose-dependent manner. Briefly, **DA-2** was incubated with K562 and HepG2 lysates at a concentration range of 1-50 μM with or without excessive dasatinib for 0.5 h at r.t. The reactions were irradiated by UV followed by "click" with rhodamine azide, separated by SDS-PAGE gel, and analyzed by in gel fluorescence scanning. As shown in Figure 6.7A, in addition to the expected c-Src (60 kDa, marked with red asterisk) and c-Abl, confirmed by pull-down and immonoblotting with corresponding antibodies (Figure 6.7C), we also observed many other off-targets. Besides, those labeled bands could be competed away by treatment with excessive dasatinib, indicating that most labeled proteins are likely specific cellular targets of dasatinib.

6.3.8 *In situ* Labelling in Cultured Cells

DA-2 has been demonstrated to label Src family kinases in cellular proteome at low micro-molar concentration. We wondered whether our probe could also label c-Src in intact cells. To this end, we studied *in situ* proteome reactivity profiles of **DA-2** in both K562 and HepG2 live cells. **DA-2** was directly added to K562 and HepG2 cultured cells at a concentration range of 1-50 μM with or without excessive dasatinib (notably, DMSO should never exceed 1% in the final solution). After 5 h of incubation at 37 ℃/5% CO_2, the cells were washed gently to remove the excessive probe followed by UV irradiation (~350 nm) for 20 min on ice. The cells were then homogenized, incubated with rhodamine-azide under "click" conditions, separated by SDS PAGE gel, and analyzed by in-gel fluorescence scanning. As shown in Figure 6.7B, distinct labeling bands were observed. Comparing with *in vitro* labeling profile, apparent low background labeling profile was observed for *in situ* labeling.

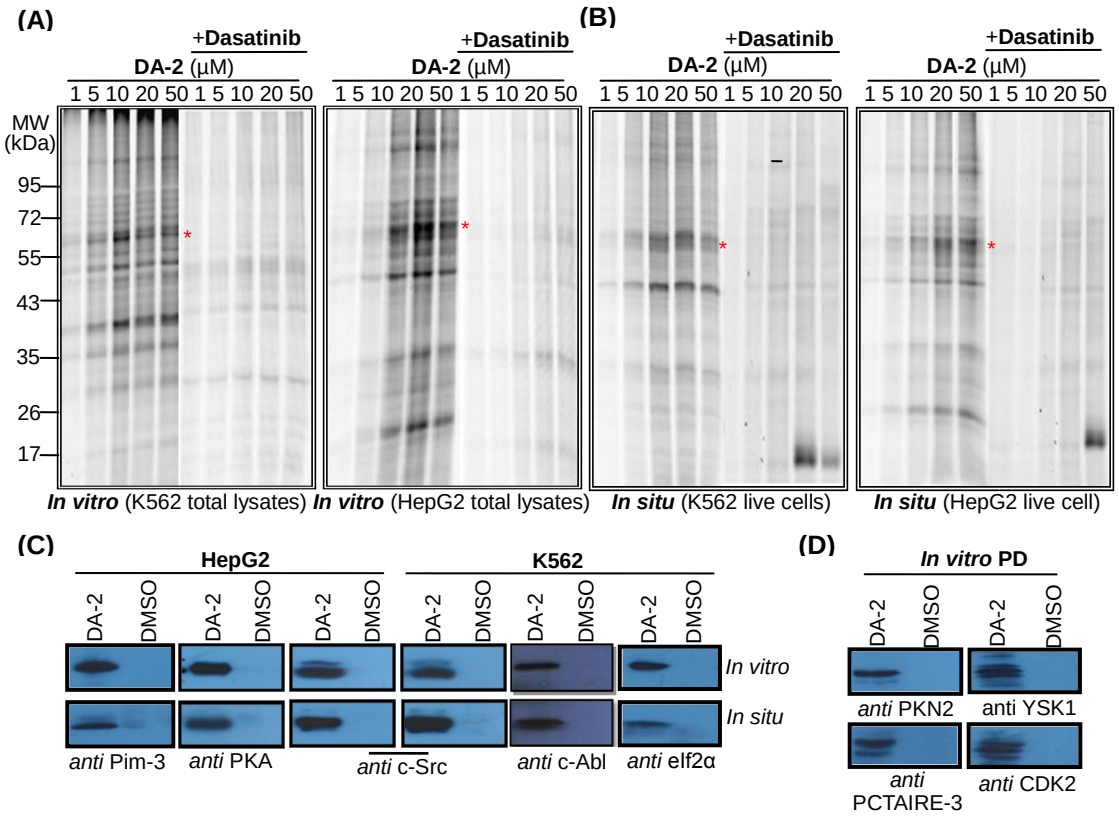


Figure 6.7 *In vitro* (A) and *in situ* (B) labeling profiling of K562 and HepG2 cells using **DA-2** (over a concentration range of 1-50 μM together with competitive experiments using dasatinib (100 μM). The red asterisks mark the position of endogenous c-Src. (C) Western blotting analysis of *in vitro* and *situ* pulled-down fractions of HepG2 and K562 cells treated with **DA-2** (or DMSO as negative controls; right lanes) with their respective antibodies. Biotin-azide was used in the click chemistry with avidin-agarose beads for pull-down experiments. (D) Western blotting validation of pulled-down fractions of HepG2 cellular lysates treated with **DA-2** (or DMSO as negative controls; right lanes) with several respective antibodies.

experiments even at the highest concentration of probe. Additionally, in order to test the labeling ability of our probe for endogenous c-Src and c-Abl in intact cells, *in situ* pull-down experiments were performed followed by immunoblotting with corresponding antibodies. Results showed that **DA-2** enabled to efficiently label endogenous c-Src and c-Abl in intact cells. Taken together, these data indicates that **DA-2** is a cell permeable probe which can be potentially used to profile the cellular targets of dasatinib in cultured cells.

6.3.9 Target Identification and Validation

In order to determine whether our probe can be used as an affinity reagent to isolate the binding proteins of dasatinib, the pull-down experiments were performed *in vitro* and *situ* with HepG2 and K562 cells using **DA-2.** In the meantime, imatinib probe (reported as a specific Abl probe)[45] and dasatinib immobilized beads were used as negative and positive controls, respectively. In brief, the labeled protein extracts obtained from *in vitro* and *in situ* labelings were subjected to click reaction conjugation with biotin azide (structure shown in Figure 6.1B) followed by enrichment with avidin-agarose beads, separated by SDS-PAGE gel and visualized by silver staining. Visual analysis of the samples from the respective pull-down experiments on a silver-stained SDS-PAGE gel revealed similar patterns of protein compositions. The entire gel lanes for each pull-down samples were then cut into two slices and each one was individually processed for in-gel trypsin digestion as standard procedure. As negative control, a large-scale proteomic experiment was also carried out with cells treated by DMSO instead of **DA-2**. Peptides obtained from each gel slice were eluted and subjected to nano-LC-MS/MS analysis. The LC-MS/MS datas were subsequently searched against IPI human protein database using an in-house MASCOT server (Matrix Science, London, UK) for protein identification.

Consequently, a total of 54 kinases (49 kinases from *in vitro* pull-downs, 25 kinases from *in situ* pull-downs) were identified as potential cellular targets of our probe. However, only 3 and 6 kinases were identified from imatinib probe and dasatinib matrix beads. Our experiments with **DA-2** versus matrix beads showed that

only few kinases were found in the matrix beads experiments while many more were found when our photo-affinity based approach was applied. Out of those 54 kinases, 21 proteins belong to protein kinases: 8 proteins are annotated as tyrosine kinases (SRC, LYN, YES, HCK, CSK, BTK, EGFR and PTK9), 13 proteins are serine/threonine kinases (p38α, DNA PKC, STK6, CDK2, and so on). The rests belong to other kinases, including STAT family, lipid kinases, pyruvate kinases, etc. Notably, in addition to several Src family kinases, CSK and BTK were consistently identified as the most prominent binding targets of dasatinib both in **DA-2** pull-down experiments and positive control experiments (on bead pull-down experiments) with high unique peptide count (Table 6.1). More notably, 13 serine/thronine family kinases were identified in our system. Among them, 7 protein kinases have been reported as dasatinib binding targets before,[127b] such as MAPKAPK2, p90-RSK-3,

Table 6.1 Some putative protein kinase targets of dasatinib identified by LC-MS.

Gene name	Description	ID	Unique peptides				Literature
			DA-2		Imatinib probe	Dasatinib matrix	
			In vitro	*In vivo*			
Src*	Isoform 1 Proto-oncogene tyrosine-protein kinase Src	IPI00641230	1	–	–	–	*PNAS*,(**2007**),104,13283
Src*	Isoform 2 Proto-oncogene tyrosine-protein kinase Src	IPI00328867	2	–	–	–	*PNAS*,(**2007**),104,13283
Yes*	Proto-oncogene tyrosine-protein kinase Yes	IPI00013981	3	–	–	–	*PNAS*,(**2007**),104,13283
Lyn*	Isoform Lyn A of tyrosine-protein kinase Lyn	IPI00298625	1	–	–	–	*PNAS*,(**2007**),104,13283
Fyn*	Isoform 1 Proto-oncogene tyrosine-protein kinase Fyn	IPI00219012	1	–	–	–	*PNAS*,(**2007**),104,13283
HCK*	Isoform p59-HCK of Tyrosine-protein kinase HCK	IPI00029769	1	–	–	–	*PNAS*,(**2007**),104,13283
CSK	Tyrosine-protein kinase CSK	IPI00013212	38	–	–	36	*PNAS*,(**2007**),104,13283
BTK	Tyrosine-protein kinase BTK	IPI00029132	41	–	–	41	*PNAS*,(**2007**),104,13283
EGFR	Isoform 1 of Epidermal growth factor receptor	IPI00018274	5	–	–	–	*Nat.Chem.Biol*.(**2010**),6,291
PTK9	Twinfilin-1 (protein tyrosine kinase 9)	IPI00183508	1	–	–	–	–
PCTAIRE-3	Isoform 2 of Serine/threonine-protein kinase PCTAIRE-3	IPI00394661	5	–	–	–	–
DNA-PKc	Isoform 1 of DNA-dependent protein kinase catalytic subunit	IPI00296337	60	6	1	–	*Nat.Biotech*.(**2007**),25,1035
MAPKAPK2	Isoform 1 of MAP kinase-activated protein kinase 2	IPI00026054	1	–	–	–	*Nat.Chem.Biol*.(**2010**),6,291
p38a	Isoform CSBP1 of Mitogen-activated protein kinase 14	IPI00221141	41	–	–	5	*Blood*,(**2007**),110,2007
YSK1	Serine/threonine-protein kinase 25	IPI00012093	3	2	–	–	–
p90-RSK-3	Ribosomal protein S6 kinase alpha-3	IPI00020898	2	–	–	–	*JPR*,(**2010**).9,806
eIF-2α	Interferon-induced, double-stranded RNA-activated protein kinase	IPI00019463	4	1	–	–	–
Pim-3	Serine/threonine-protein kinase Pim-3	IPI00848067	–	3	–	–	–
PKA C-α	Isoform 2 of cAMP-dependent protein kinase, alpha-catalytic subunit	IPI00217960	–	2	–	–	–
PKC-β	Isoform Beta-II of Protein kinase C beta type	IPI00219628	6	–	–	–	*Nat.Biotech*.(**2007**),25,1035
STK6	Serine/threonine-protein kinase 6	IPI00298940	36	–	–	–	*Nat.Biotech*.(**2007**),25,1035
CDK2	Cell division protein kinase 2	IPI00031681	5	–	–	–	*Nat.Biotech*.(**2007**),25,1035
PKN2	Serine/threonine-protein kinase N2	IPI00002804	55	–	–	–	–

- : no datas; ***** : sharing identical unique peptide

p38α, DNA-Pac, STK6, CDK2 and PKC-β. While the rest 6 kinases including PKN2, PCTAIRE-3, eIF-2α, Pim-3 and PKA, for the first time, were identified as previously unknown target proteins for dasatinib. In general, the majority of those 21 protein kinases identified by us are known to play important roles in signal transduction and cell cycle regulation. In addition to kinases, some non-kinase targets were also identified, including STAT1, STAT3,[137] GRB2,[127a] tyrosine phosphatases and serine/threonine-protein phosphatases, and so on. Most of them are likely to be interactors of dasatinib kinase targets.

In an effort to further validate our MS results, we performed *in vitro* inhibition assay for **DA-2** with purified CSK and PKA (Table 6.2). Results showed that IC_{50} of **DA-2** for CSK and PKA were 20 nM and 6 μM, respectively, which indicates that our probe is able to capture not only strong-affinity interactors but also weak-affinity ones. Additionally, we also performed pull-down and western blotting experiments with the respective antibodies (part D and E of Figure 5); results confirmed that, in addition to c-Src and Abl, PKA, Pim-3 and eI*f*2α could be indeedly labeled by **DA-2** both in cellular lysates and live cells. While other several kinases including CDK2, PKN2, PCTAIRE-3 and YSK1 were positively labeled by **DA-2** only in cellular lysates, which is quite consistent with our LCMS results. Taken together, our results further demonstrate that Abl, SRC family kinase, CSK and BTK are major binding targets of dasatinib. Likewise, Serine/Threonine kinases may be another important group of kinases directly interacting with dasatinib.

Table 6.2 IC_{50} values (nM) determined in biochemical enzyme assays.

Enzymes	IC_{50} (nM)				
	Dasatinib	**DA-1**	**DA-2**	**Staurosphorine**	**H-89**
c-Abl	8.8 ± 0.4	30.9 ± 0.3	27.2 ± 0.1	n.d.	n.d.
c-Src	3.6 ± 0.1	19.7 ± 1.1	14.7 ± 0.1	n.d.	n.d.
SrcT338M	319 ± 50	>5000	>5000	n.d.	n.d.
PKA	3620 ± 100	7041 ± 300	4759 ± 400	38.71 ± 0.3	81.73 ± 1.0
CSK	7.06 ± 0.3	26.28 ± 0.2	20.66 ± 0.09	26.88 ± 0.5	n.d.

n.d.: not determined.

6.4 Conclusion

We have developed a novel chemical proteomic approach which can be used to identify the cellular targets of dasatinib in native environment. In addition to Src and Abl kinase, a broad range of putative targeted kinases have been identified using our probe. Most notably, some serine/threonine protein kinases, for the first time, were identified and validated as potential dasatinib binding targets in this study. We predict that our chemical proteomic approach should be broadly useful for off-target identification against a number of existing drugs, and it may provide a powerful tool for biomarker discovery in future.

Chapter 7

A "Clickable", Cell-permeable Probe for Proteome Profiling of Potential Cellular Targets of Staurosphorine

7.1 Summary

This chapter describes a photo affinity-based approach for investigating the cellular on/off targets of staurosphorine. As in the previous chapter, where a cell permeable probe derived from dasatinib^{TM} was developed for proteome profiling of cellular kinases, herein a staurosphorine-like, cell permeable probe was designed and synthesized. We further evaluated its biochemical characterization and utility for proteome profiling of cellular targets of staurosphorine. The preliminary results demonstrated that this probe exhibited comparative activity as staurosphorine and positively labeled a wide of kinases both in cellular lysates and live cells. In order to investigate the cellular targets of staurosphorine, we performed large-scale proteomic/mass spectrometric experiments in HepG2 cells. A total of 43 kinases were identified (35 kinases from *in vitro* pull-downs, and 21 from *in situ* pull-downs) including a number of previously well-known staurosphorine binding targets. More notably, by comparing the MS results between *in vitro* and *situ* pull-downs, eight kinases, for the first time, were found to be uniquely labeled in live cells by our probe. It may provide important implication for studing the therapeutic function of staurosphorine in future.

7.2 Introduction

Protein kinases (PKs) play a key role in signal transduction pathways, which regulate a varity of cellular processes including growth, division, differentiation and metabolism. Dysregulation of cellular kinase activities has been implicated in many diseases (including cancer, HIV, malaria, and so on).[18] Kinase inhibitors are usually used to understand kinase function and as therapeutic compounds to treat cancer and other diseases.[138] As an example, staurosporine is the most potent general inhibitor for protein kinases, and it has been reported to target hundreds of kinase.[130,139] To date, staurosporine and many of its derivates have received considerable attention as potential anticancer drugs due to their significant biological effects.[140] With the aim to elucidate the potential uitility of staurosphorine as a therapeutic agent, it is very significant to identify its potential cellular off-targets. Although affinity matrix method has become the most popular approach for identification of kinase inhibitor binding targets from cellular or tissue extracts,[128b,141] this approach still limits in several aspects: first, the solid support may affect the accessibility of the immobilized drugs for interactions with the target proteins; second, due to the non-covalent interactions between drugs and proteins, it is hard to find a compromise condition between specific and nonspecific interactions; third, the false positive proteins could be easily identified due to the loss of cellular compartmentation *in vitro*. As an alternative approach, recently, Mathias and colleagues described a Capture Compound Mass Spectrometry (CCMS) method for comprehensive identification of

staurosphorine-binding kinases.[142] They used a trifunctional staurosphorine-derived probe to profile the kinome of the human liver-derived cell line in solution. However, due to the steric hindrance of reporters for the binding between probe and the active site of kinases, the less potency and poor cell permeability of this probe limit its utility only in the profing of cell lysates *in vitro*. In an attempt to develop a staurosphorine-derivated probes which can be used to proteome profile the cellular targets of staurosphorine in live cells, we designed and synthesized a "clickable", affinity-based probe (A*f*BP) based on the well known activity-based protein profiling (ABPP) approach which has been sucessfully applied for various classes of enzyme.[41,45,65a,68a,82,83b,101,102a] Our probe essentially contains three structural components: (1) a protein recognition group derived from staurosphorine; (2) a photo-cross linker which was diazirine in this study, and (3) an alkyne handle incorporating to a fluorescent dye or biotin group via the copper catalyzed Huisgen 1,3-dipolar cycloaddition.[34] Herein, we describe the design, synthesis and biochemical characterization of this probe, and further evaluate its utility for proteome profiling of cellular targets of staurosphorine.

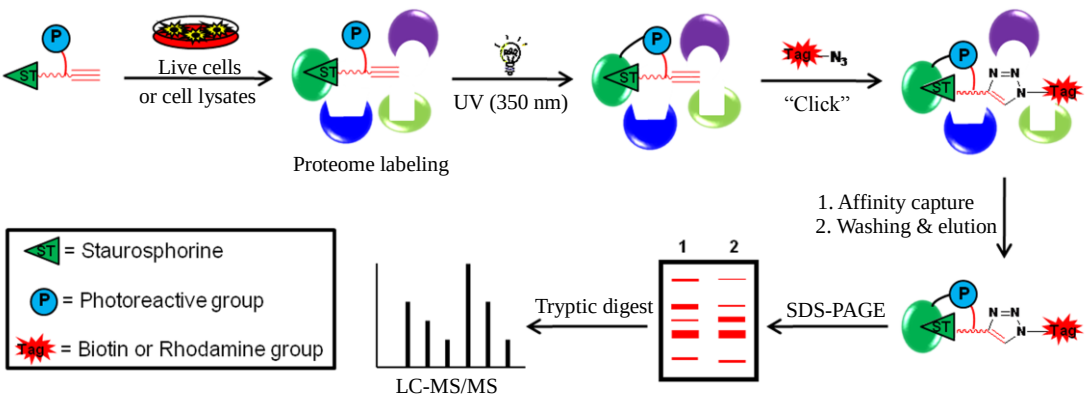
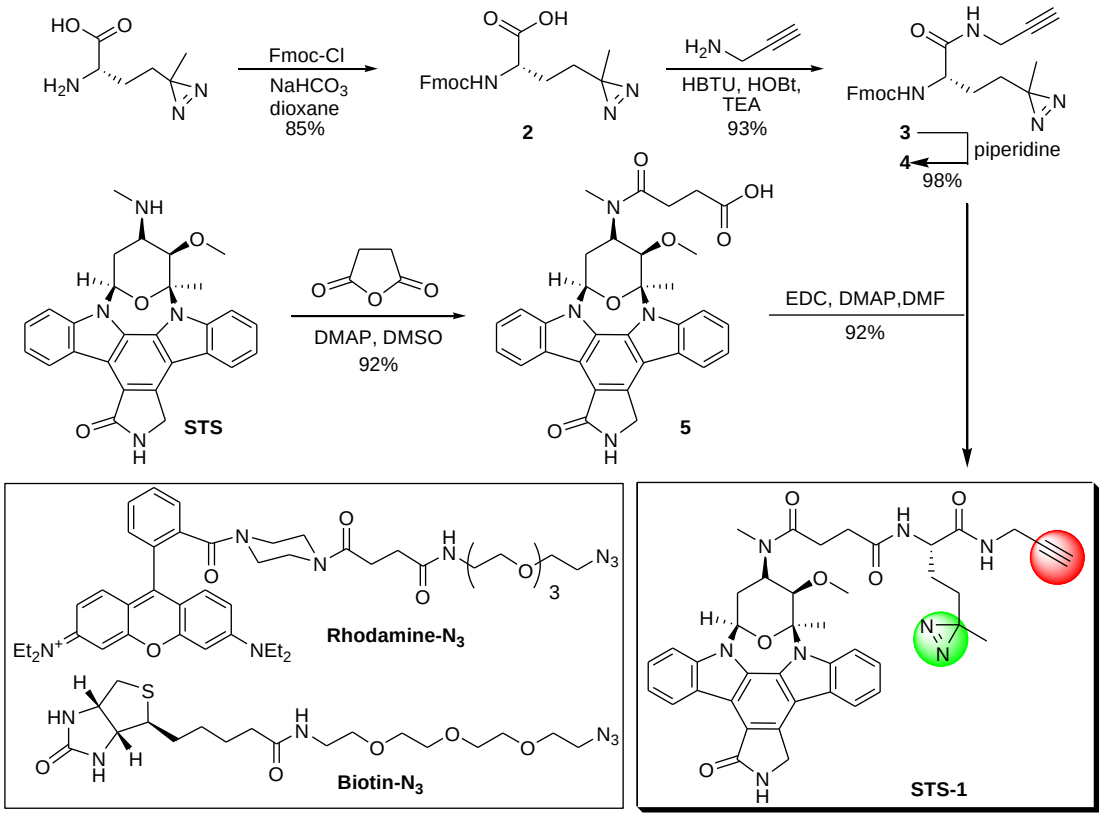


Figure.7 1 Overall strategy of affinity-based proteome profiling.

7.3 Results and Discussion

7.3.1 Design and Synthesis of Staurosphorine-like Probe

Towards developing affinity-based probes (A*f*BPs) against kinases, we chose staurosphorine core structure as the kinase recognition unit and diazirine as the photo-cross unit. An alkyne handle was used to incorporate to a fluorescent dye or biotin group for visualization/pull-downs. The general labeling strategy of the "clickable" staurosphorine-derived probe is shown in Figure. 7.1. To determine whether the introduction of the photo-cross linker affects the binding between probe and kinases, we carried out molecular docking study using PKA as model protein. The results showed that the staurosphorine warhead bound well to the ATP binding pocket (Figure. 7.2A), which is consistent with previously reported crystal structure of the PKA complexed with staurosphorine.[143] Additionally, the photo-cross linker points to the surface of protein as expected, which has negligible effect on the binding between staurosphorine and the ATP binding site. The synthesis of the probe is shown in Scheme 7.1. In brief, the diazirine unnatural amino acid was synthesized as previously reported method.[144] After protection of amine with Fmoc group, compound **2** was then coupled with propargylamine to give **3**. The Foc group in **3** was subsequently removed by treatment with 20% piperidine in DMF to afford the compound **4** in 91% yields over two steps. Commercial available staurosphorine (STS) was coupled to succinic anhydride, yielding compound **5** in 92% yields. Compound **5** was finally coupled to compound **4** to give the desired probe **STS-1** in 92% yield.



Scheme 7.1 Synthesis of the "clickable" staurosphorine-like probe **STS-1**. The "Click"-based reporters used are rhodamine azide and biotin azide which were prepared as previously described.[104c]

7.3.2 Biological Characterization of Probe STS-1

To characterize the potency of our probe for kinase, we compared the inhibitory potential of our probe to staurosphorine by measuring the kinase activities of PKA in a dose-dependent manner under optimized enzyme concentration of ~50 nM. As seen in Figure. 7.2B, IC_{50} of staurosphorine for PKA is 35.3 ± 0.5 nM, whereas IC_{50} of STS-1 is 62.2 ± 0.2 nM, which indicates the introduction of the photo-cross linker slightly affects staurosphorine activity, and probe **STS-1** exhibits comparative activity as staurosphorine.

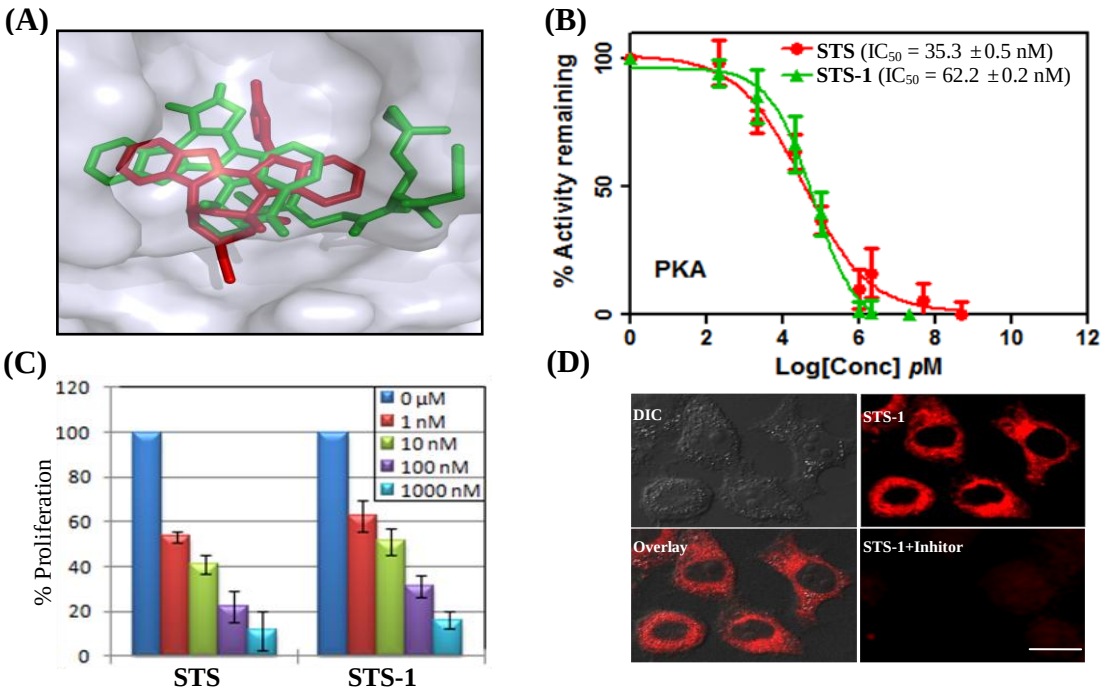


Figure 7.2 Biological activity of **STS-1** *in vitro* and *situ*. (A) Molecular docking of **STS-1** and staurosphorine in the ATP binding site of PKA. Protein, staurosphorine and **STS-1** are shown in grey, red and green, respectively. (B) IC_{50} values of two compounds (**STS-1** and staurosphorine) for PKA determined with ProFluor® PKA Assay. Data represent the average (s.d. for two trials). (C) Dose-dependent inhibition of HepG2 cells proliferation by STS and **STS-1** using XTT assay. (D) Cellular imaging of HepG2 cells with probe **STS-1**. All images were acquired in the same way. Scale bar = 10 µm.

Next, in order to evaluate the potential biological activity of **STS-1** against staurosphorine, we performed the cell proliferation assay using HepG2 cells. In the cell proliferation assay (Figure.7.2D), our probe showed a dose-dependent inhibition of tumor cell proliferation over a 72 h time period with comparable potency. On the other hand, this experiment also indicated that **STS-1** was cell permeable. To determine the cell permeability of our probe, we performed cell permeability assay together with staurosphorine as positive control. Results indicated that probe **STS-1** had comparable cell permeability with staurosphorine.

To further demonstrate our probe's cell permeability and utility for potential cellular imaging, we performed fluorescence microscopy to visualize the cells treated by **STS-1** (Figure 2D). Live HepG2 cells were first treated with **STS-1** followed by

irradiation under UV, and then fixed with paraformaldehyde, permeabilized with Triton X-100, conjugated to rhodamine azide by click chemistry, and imaged (colored in red). Competitive experiment was carried out in parallel. Results showed that minimal fluorescence was observed in cells treated with staurosphorine as competitor as well as DMSO. However, higher fluorescence was observed for cells treated with **STS-1**. Taken together, probe **STS-1** has good cell permeability, and it can be potentially used for bioimaging in future.

7.3.3 *In vitro* Labeling with Purified Kinase

Next, to evaluate **STS-1** as an activity-based probe for kinases, the same amount (~30 pmol) of three protein kinases, namely c-Src, CSK and PKA, were labeled by **STS-1**, respectively. **STS-1** was incubated with different kinases in parallel for 30 min followed by irradiation under UV lamp at ~350 nm. Subsequently, the reactions were subjected to click reaction with rhodamine azide (TER-N_3) (structures shown in Figure 7.1), after which the fluorescent labeled proteins were separated by SDS-PAGE and visualized in-gel with a fluorescence gel scanner. As shown in Figure. 3A, all kinases can be positively labeled, which is consistent with the multiple targets character of staurosphorine. To further assess the selectivity of the probe for PKA, we carried out the labeling experiments for recombinant PKA in the presense of bacterial proteome. Results showed that a stronger fluorescent band was observed for the bacterial lysates of bacteria treated with inducer, while almost no labeling was observed for un-induced lysates (Figure 3D). These results unambiguously indicate that **STS-1**

can be used to label and detect the expression levels of PKA from different biological samples.

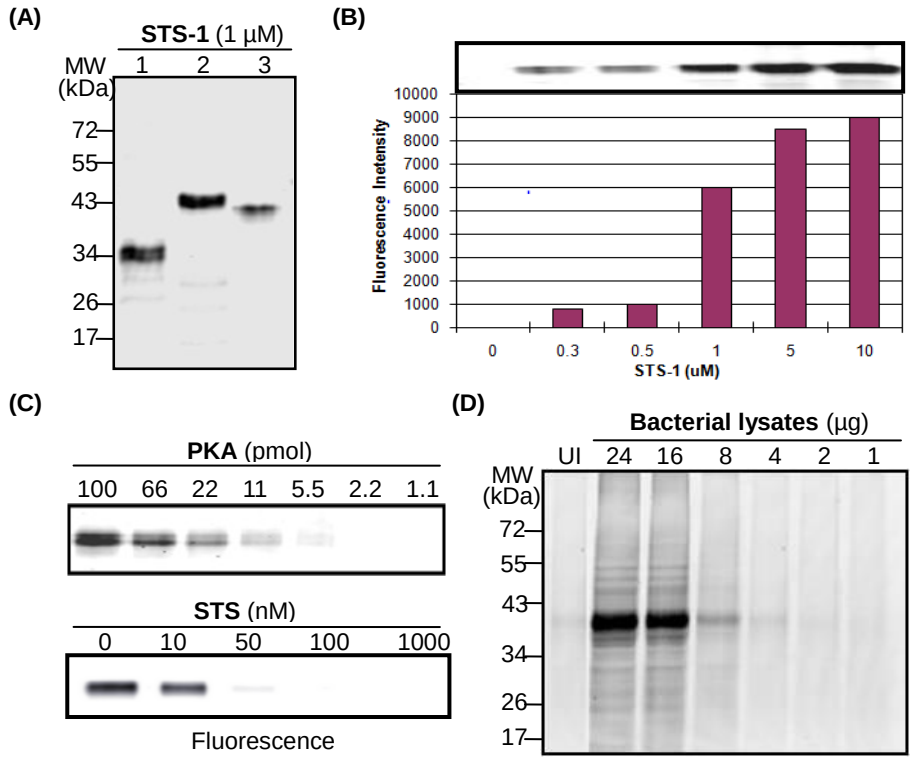


Figure 7.3 Fluorescent labeling of recombinant protein kinases. (A) Three purified kinases (30 pmol, lanes 1 to 3 are c-Src, CSK and PKA, respectively) were labeled by **STS-1**. (B) Recombinant PKA (100 nM) was labeled with increasing concentration of **STS-1** and quantification of fluorescent bands. (C) Dose-dependent labelings of PKA with **STS-1** (1 μM final concentration) (top) and competitive labeling experiments (bottom). (D) Bacterial lysates of induced and un-induced (UI) bacteria expressing PKA were labeled by **STS-1**(1 μM).

7.3.4 Proteome Profiling of Mammalian Cellular Lysates

Next, to study the labeling ability of the probe for endogenous PKA, we performed *in vitro* proteome profiling experiments with HepG2 cell lysates in dose-dependent manner together with staurosphorine as competitor. As shown in the left panel of Figure.7 4A, a distinct fluorescent band was observed at 40 kDa in the fluorescent gel. This band has been confirmed to be endogenous PKA by *in vitro* pull-down and immunoblotting with anti PKA antibody. Additionally, we

observed that, in addition to PKA, the probe also labeled a number of other proteins, and those labeled bands could be competed away by treatment with excessive staurosphorine, indicating that most labeled proteins are likely specific cellular targets. To further investigate the *in situ* proteome reactivity profiles of **STS-1** in HepG2 live cells, we then set out to perform the labeling experiments in HepG2 live cells. As shown in the right panel of Figure 7.4, similarly, in addition to the expected PKA band confirmed by immunoblotting, we also observed a number of staurosphorine-sensitive targets, that is, those labeled bands that were competed away by treatments with excessive staurosphorine. Taken together, these experiments thus demonstrate the feasibility of our probe for profiling of endogenous PKA.

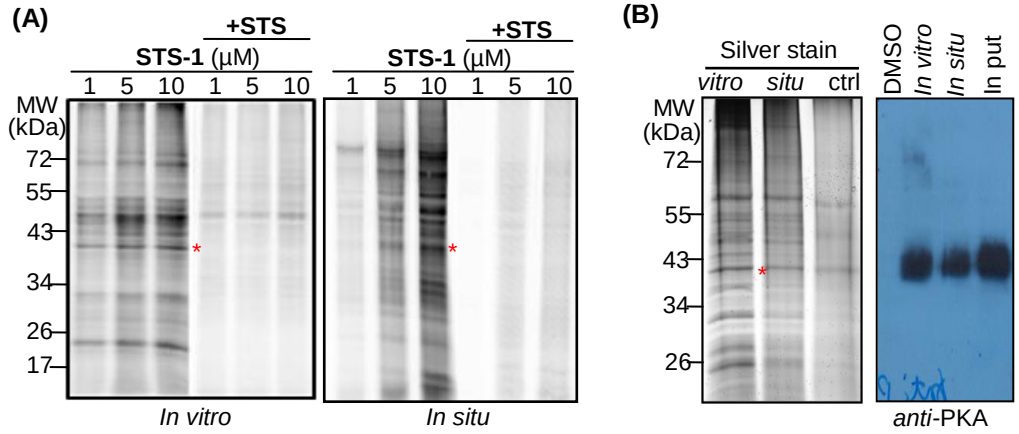


Figure 7.4 Proteome profilings of HepG2 cells and PKA validation. (A) *in vitro* and *situ* labeling of HepG2 cells by **STS-1** (over a concentration range of 1-10 μM together with competitive experiments with 100 μM staurosphorine (STS). The red asterisks mark the position of endogenous PKA; (B) Silver stain gel for *in vitro* and *situ* pull down fractions and western blotting analysis (DMSO as negative controls; left lanes) with anti PKA antibody.

7.3.5 Target Identification and Validation

In order to identify the cellular targets of staurosphorine, we performed

large-scale proteomic/mass spectrometric experiments in HepG2 cells. A total of 43 kinases were identified (35 kinases from *in vitro* pull-downs, and 21 from *in situ* pull-downs). Out of them, 14 kinases belong to protein kinases: 4 proteins are annotated as tyrosine kinases (SRC, YES, CSK and BTK), another 10 proteins are serine/threonine kinases (PCTK3, DNA-PKC, STK25, RSK-3, eIf2α, p38, PKA C-α, MST4, CDK2 and PKN2). The rest of kinases identified are non-protein kinases including lipid kinases, phosphofructokinases, pyruvate kinases, and so on. Some of them have been reported as known staurosphorine binding targets before.[142] By comparing the MS results between *in vitro* and *situ* pull-downs, 13 kinases were overlaid as common protein hits. Interestingly, eight unique kinases listed in Figure 7.5A were only identified from *in situ* pull-down experiments. In an effort to further validate these seven protein hits, we carried out *in vitro* and *situ* pull-down experiments followed by immunoblotting with respective antibodies. Results confirmed that five of them could be consistantly labeled by **STS-1** in live cells rather than cell lysates. Additionally, PKA, c-Src and PKN2 could be positively labeled both in cell lysates and live cells (Figure 7.5C). Taken together, our preliminary results herein thus may provide important information for studying the cellular off-targets of staurosphorine in future.

(A)

ID	Description	MS	Score	Peptides Matched	Classification
IPI00007123	RPS6KA6	84390	55	1	Ser/Thr Kinase
IPI00217960	PKA C-α	39910	47	2	Ser/Thr Kinase
IPI00219585	PFKM	82600	48	3	Phosphofructokinase
IPI00639981	PFKP	26400	70	1	Phosphofructokinase
IPI00219617	PRPS2	35330	254	3	Pyrophosphokinase
IPI00234368	ADK	39080	45	1	AK2 kinase
IPI00013862	CDC8	23980	62	1	Thymidylate kinase
IPI00022977	CKB	42900	319	17	Creatine kinase

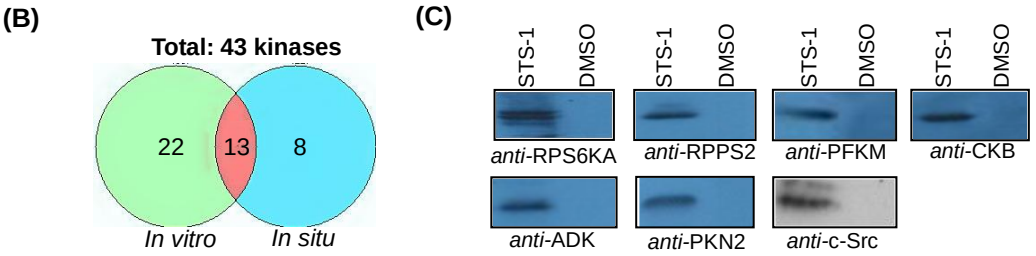


Figure 7.5 LCMS results and hit validation. (A) Eight kinases uniquely identified from *in situ* pull-down experiments by LC-MS but not *in vitro* pull-down. (B) Venn diagram illustrating the comparision of the number of kinases identified from HepG2 cell lysates and live cells. (C) Western blotting analysis of *in situ* pulled-down fractions of HepG2 live cells treated with **STS-1** (or DMSO as negative controls; right lanes) with the respective antibodies.

7.4 Conclusion

In summary, we have successfully synthesized a novel "clickable", cell permeable staurosphorine-derived probe which enble to positively label endogenous PKA not only in cellular lysates but also live cells. Most notably, the preliminary results demonstrated that our permeable probe, not like other probes, was very useful for proteomic profiling of the cellular targets of staurosphorine in live cells. Therefore, it should provide important implication for studing the therapeutic function of staurosphorine in future.

Chapter 8

Experiment Section

8.1 General Information

8.1.1 Materials

All chemicals were purchased from commercial vendors and used without further purification, unless otherwise noted. Either HPLC grade or distilled solvents were used in all the reactions. All moisture-sensitive reactions were performed under a positive pressure of nitrogen or argon. Concentration *in vacuum* was performed on a Büchi rotary evaporator. Analytical thin layer chromatography was performed using Merck silica gel plates (0.25 mm thickness) with fluorescent indicator UV254. Spots were visualized by ultraviolet illumination, iodine staining, $KMnO_4$ staining (for alkenes, aldehydes), Nindhydrin staining (for primary amines), or ceric molybdate staining. Flash chromatography was performed using Merck silica gel (40 μM particle size) and freshly distilled or AR grade solvents. Chemical shifts are reported as δ in parts per million (ppm) referenced with respect to residual solvent ($CDCl_3$ = 7.26 ppm, methanol-d4 = 3.3 ppm and DMSO-d6 = 2.50 ppm) or from internal standard tetramethylsilane (TMS = 0.00 ppm). ^{13}C NMR spectra are reported as δ in units of parts per million (ppm) relative from solvent signal: $CDCl_3$ (δ = 77.0, triplet) and methanol-d4 (δ = 49.0, singlet). The following abbreviations were used in reporting spectra: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of

doublets. Rink Amide resin, all Fmoc amino acids, O-benzotriazole-N,N,N,N-tetramethyl-uronium-hexafluoro-phosphate (HBTU), N-hydroxy-benzotriazole (HOBt), triisopropylsilane (TIS) and bromo-tris-pyrrolidino phosphonium hexafluorophosphate (PyBrOP) used in peptide synthesis were purchased from GL Biochem (China). 4-Formyl-3-methoxyphenoxy Resin (PL-FMP Resin, 0.9 mmol/g, 75~150 μm, Part no: 1465-799) was purchased from Polymer Laboratories (USA). All reactions requiring anhydrous conditions were carried out under an argon or nitrogen atmosphere using oven-dried glassware. Solid-phase reactions were performed on an IRORI AccuTagTM Combinatorial Chemistry System. Some enzymes were purchased from commercial sources. Pepsin and Renin (from Mucor miehei) were from Fluka. Enzyme stocks were made fresh using 10 mM HCl (for pepsin) or water (for renin). Plasmepsin I (PM-I), Plasmepsin II (PM-II), histoaspartic protease (HAP) and HAP mutant H34A, S37A, K78A, E278A were expressed and purified as described previously. Anti-PM-I mouse serum mAb1 C6-24, anti-PM-II rabbit serum 737, anti-PM-IV mouse mAb 13.9.2 were requested from MR4 (USA). The anti-HAP mouse mAb was kindly provided by Dr. Daniel E. Goldberg. Some enzymes were expressed in *E.coli* strain BL21-DE3 and purified as described previously. It includes Abl (human c-Abl, residues 229-512), c-Src and its mutants (chicken c-Src residues 251-533 or 83-533). Dasatinib and staurosphorine (98%) were purchased from invitrogen. Tris (2-carboxyethyl) phosphine (TCEP), and click chemistry ligand, tris[(1-benzyl- 1H-1,2,3-triazol-4-yl)methyl] amine (TBTA) were purchased from Sigma-Aldrich. Antibodies against c-Src (SC-8056), cathepsin D, HSPA5, PKA C-α

(SC-903), Rsk-4 (JS-31), PRPS2 (L-22), PFK-1 (G-11), Creatine kinase-B (N-20), ADK (F-5), PRK2 (C-6), DTYMK (C-17), Pim-3 (C-18), Cdk2 (AN4.3), YSK1 (E-5) and PCTAIRE-3 (P-23) were purchased from Santa Cruz Biotechnology, c-Abl, eIf2α, phospho-Y416 and phospho-Y527 antibodies were purchased from Cell Signaling Technologies (Beverly, MA). Antibody against β-tubulin was purchased from Abcam.

8.1.2 Instruments

Mass spectra were recorded on a Finnigan LCQ mass spectrometer, a Shimadzu LC-IT-TOF spectrometer or a Shimadzu LC-ESI spectrometer. Analytical and preparative RP-HPLC separations were performed on Phenomex C18 column (250×4.60 mm) and Phenomex C18 (250×21.20 mm) coloumns, respectively, using a waters 600E HPLC system equipped with a waters 600 controller and a waters 2487 UV detector. Shimadzu LCMS-IT-TOF system equipped with auto-sampler (cat. No. LCMS-2010EV) was used for the analysis of the inhibitor library. Eluents A (0.1% TFA/acetonitrile) and B (0.1% TFA/water) were used as the mobile phase. ^{1}H NMR and ^{13}C NMR spectra were recorded on a 300 MHz 89 Bruker ACF300 or Bruker *Avance* 500 MHz or DPX-300 NMR. For enzyme inhibition and IC_{50} measurements, Tecan microplate reader (Multimode Reader, Infinite®200) in luminescence mode with i-control™ software was used. Fluorescence scanning of the SDS-PAGE gels was carried out with Typhoon 9410 Variable Mode Imager scanner (Amersham), and where applicable, the bands were quantified with the ImageQuant 3.3 (Molecular Dynamics) software installed on the scanner. Microarrays were printed using a contact arrayer (Virtek Chipwriter, Ontario, Canada) and scanned on an ArrayWoRxTM

microarray scanner (Applied Precision, USA) equipped with the relevant filters for coumarin (lex/em: 360/457 nm), fluorescein (lex/em: 490/528 nm), Cy3 (lex/em: 548/595 nm) and Cy5 (lex/em: 633/685 nm).

8.2 Solution-phase Synthesis

8.2.1 General Procedure for Synthesis of Diol-amine Warheads

*General procedure for synthesis of N-benzyloxycarbonyl protection of amino acids 1 **(a-c)**:*

A solution of the free amino acid (1.0 eq.; 50 mmol) in THF (100 mL) was cooled to 0 ℃ followed by addition of a solution of Na_2CO_3 (3.0 eq.; 150 mmol; 15.90 g) in water (100 mL). Benzyloxycarbonyl chloride (1.1 eq.; 55 mmol; 7.82 mL) in THF (30 mL) was then added dropwise to the reaction mixture and the resulting mixture was stirred at 0 ℃ for 2 h, then warmed up to r.t. and continued overnight. Subsequently, the reaction was quenched with water followed by acidification with 10 % HCl to *p*H 1-2, and the aqueous layer was extracted with EA (3×100 mL). The combined organic layers were washed with brine (2×100 mL) and then dried over Na_2SO_4. Finally the solvent was evaporated under reduced pressure to afford the pure Cbz-protected amino acid without further purification.

(S)-2-[N-(benzyloxycarbonyl)amino]-3-phenyl-propionic acid (**1a**). White solid. Yield = 97 %. ^{1}H-NMR (300 MHz, $CDCl_3$), δ 2.80 (d, *J* = 7.05 Hz, 2H), 4.50 (m, 1H), 4.95 (s, 2H) , 5.05 (s, 1H) 7.28-7.17 (m, 10H); ESI-MS: *m/z* $[M-H]^-$ = 298.0.

(S)-2-[N-(benzyloxycarbonyl)amino]-4-methyl-pentanoic acid **(1b)**. Colourless oil.

Yield = 96 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.97-0.94 (m, 6H), 1.79-1.50 (m, 3H), 4.59 (m, 1H), 5.16 (s, 2H), 5.30 (s, 1H), 7.31-7.26 (m, 5H), 8.18 (br s, 1H); ESI-MS: *m/z* $[2M + Na]^+$ = 552.9.

(S)-2-(benzyloxycarbonylamino)propanoic acid **(1c).** Colourless oil. Yield = 94 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.42 (d, *J* = 7.32 Hz, 1H), 4.71-4.20 (m, 1H), 5.09 (s, 2H), 7.40-7.30 (m, 5H), 8.36 (s, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 175.6, 157.3, 136.5, 128.4-127.9, 60.4, 50.3, 18.4; ESI-MS: *m/z* $[M-1]^-$ = 222.0.

General procedure for preparation of Cbz-α-amino alcohols ***2(a-c):***

A solution of the Cbz-α-amino acid (1 eq.; 48 mmol) in dry dimethoxyethane (DME) (100 mL) was cooled to −15 °C under a nitrogen atmosphere. N-methylmorpholine (NMM) (1 eq.; 48 mmol; 6.26 mL) and isobutyl chloroformate (ICF) (1 eq.; 48 mmol; 5.28 mL) were added dropwise. After 10 min, the reaction mixture was allowed to warm up to r.t. (in ~30 min). The mixture was filtered under nitrogen atmosphere to remove the precipitated NMM hydrochloride. The filtrate was cooled to −15 °C and a solution of $NaBH_4$ (1.5 eq.; 72 mmol; 2.72 g) in water (10 mL) was added. After stirring for 1 h, water (100 mL) was added to quench the reaction. DME was then removed under reduced pressure. The aqueous layer was extracted with EA (3×50 mL) and the combined organic layers were washed with water and saturated NaCl solution and dried with Na_2SO_4. Finally the solvent was removed under reduced pressure. Purification was carried out by column chromatography on silica gel (EtOAc/Hexane = 1:1).

(S)-(1-benzyl-2-hydroxy-ethyl)-carbamic acid benzyl ester **(2a)**. White solid. Yield = 75 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 2.79 (d, *J* = 7.08 Hz, 1H), 3.53 (m, 2H), 3.89-3.84 (m, 1H), 4.93 (s, 1H), 5.00 (s, 2H), 7.31-7.11 (m, 10H); ^{13}C-NMR (75 MHz, $CDCl_3$): 156.6, 137.5, 136.2, 128.8-126.3, 66.5, 63.3,54.0, 37.4; ESI-MS: *m/z* $[M-H]^-$ = 284.1.

(S)-(1-hydroxymethyl-3-methyl-butyl)-carbamic acid benzyl ester **(2b)**. White solid. Yield: 77 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.90 (m, 6H), 1.30 (m, 2H), 1.64 (m, 1H), 3.56 (d, *J* = 8.01 Hz, 3H), 3.74 (s, 1H), 5.02 (s, 2H), 5.30 (d, 1H), 7.28 (s, 5H); ^{13}C-NMR (75 MHz, $CDCl_3$): 156.6, 136.3, 128.1-126.7, 65.2, 64.7,51.3, 40.2, 25.1, 22.0; ESI-MS: *m/z* $[2M+ Na]^+$ = 524.9.

(S)-(1-hydroxypropan)-carbamic acid benzyl ester **(2c)**. White solid. Yield: 70 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.17 (d, *J* = 6.9 Hz, 3H), 1.66 (s, 1H), 3.56-3.39 (m, 2H), 3.94-3.78 (m, 1H), 4.87 (s, 2H), 7.36-7.34 (m, 1H); ESI-MS: *m/z* $[M+Na]^+$ = 232.0.

General procedure for preparation of the Cbz-α-amino aldehydes ***3(a-c):***

A solution of **2** (1.0 eq.; 39.2 mmol) in DCM (100 mL) was added dropwise to a solution of DMP (1.3 eq.; 51 mmol; 21.6 g) in DCM (150 mL). The reaction mixture was stirred at r.t. for 6 h until TLC indicated complete conversion. Aqueous $Na_2S_2O_3$ (150 mL) was then added and stirring was continued for another 30 min. The organic layer was separated and the aqueous layer was extracted with DCM (3×50 mL). The combined organic layers were then washed with saturated $NaHCO_3$ (2×50 mL), water

(2×100 mL) and brine (2×100 mL) and then dried over Na_2SO_4. The solvent was evaporated under reduced pressure. The crude product was purified by flash chromatography (EA/hexane, 20-25 % v/v, with 0.1 % pyridine) to afford **3**.

(S)-(1-Benzyl-2-oxo-ethyl)-carbamic acid benzyl ester **(3a)**.Yellow solid. Yield: 70 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.08 (m, 2H), 4.45 (m, 1H), 5.08 (s, 2H), 5.40 (s, 1H), 7.33-7.21 (m, 10H), 9.57 (s, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 199.5, 156.5, 136.1, 129.4-127.7, 67.7, 61.7, 35.9; ESI-MS: *m/z* $[M-H]^-$ = 282.0.

(S)-(1-Formyl-3-methyl-butyl)-carbamic acid benzyl ester **(3b)**. Colorless oil. Yield: 65 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.95 (d, 6H), 1.41 (m, 1H), 1.77 (m, 2H), 4.27 (m, 1H), 5.10 (s, 2H), 5.44 (s, 1H), 7.33 (s, 5H), 9.54 (s, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 199.8, 156.1, 136.1, 128.4-128.0, 66.9, 58.7, 37.8, 24.5, 22.9, 21.7; ESI-MS: *m/z* $[M-H]^-$ = 248.1.

(S)-(1-formylpropan)-carbamic acid benzyl ester **(3c)**. White solid. Yield: 60 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.10 (d, *J* = 9.6 Hz, 3H), 3.54-3.48 (m, 1H), 5.09 (s, 2H), 7.35-7.30 (m, 5H).

*General procedure for synthesis of Cbz-protective diols **(4a-b)** by pinacol homocoupling reaction:*

VCl_3 (2 eq.; 60 mmol; 10 g) was refluxed in distilled THF (25 mL) under N_2 for 5 h. The mixture was cooled to r.t., followed by addition of Zn powder (2 eq.; 60 mmol; 4 g). The resulting mixture was stirred for 30 min, then heated to reflux before 1,3-dimethylimidazole (36 mL) was added, followed by the dropwise addition of **3** (1

eq.; 30 mmol) in distilled THF (20 mL). The reaction mixture was then refluxed for further 1 h. The solution was cooled to r.t. and diluted with 10 % aqueous citric acid. Extraction with EA was then carried out (3×50 mL) and the combined organic layers were washed with water (3×50 mL) and brine (2×50 mL), dried over Na_2SO_4, filtered and concentrated. The crude product was purified by recrystallization with ethanol to afford **4**.

(1-Benzyl-4-benzyloxycarbonylamino-2,3-dihydroxy-5-phenyl-pentyl)-carbamic-acid benzyl ester **(4a)**. White solid. Yield: 20 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 2.78-2.58 (m, 4H), 4.19 (d, *J* = 4.77 Hz, 2H), 454 (s, 2H), 4.93 (d, *J* = 8.70 Hz, 4H), 6.82 (d, *J* = 9.39 Hz, 2H), 7.36-7.32 (m, 20H); ^{13}C-NMR (75 MHz, DMSO-d_6): 155.85, 139.3, 137.5,129.2, 128.2, 127.9, 127.2, 125.7, 72.7, 64.9, 53.0, 39.5; LC-MS: *m/z* $[M+Na]^+$ = 591.2.

(4-Benzyloxycarbonylamino-2,3-dihydroxy-1-isobutyl-6-methyl-heptyl)-carbamic acid benzyl ester **(4b)**. White solid. Yield: 15 %. ^{1}H-NMR (300 MHz, DMSO-d_6) δ 0.83 (m, 12H), 1.12 (m, 2H), 1.44 (m, 4H), 3.17 (s, 2H), 3.91 (m, 2H), 4.34 (m, 2H), 5.02 (m, 4H), 6.59 (d, *J* = 8.70 Hz, 4H), 7.33 (s, 10H); ^{13}C-NMR (75 MHz, DMSO-d_6): 156.0, 137.6, 128.3, 127.6, 127.4, 72.9, 64.9, 49.6, 24.2, 23.3, 21.9; LC-MS: *m/z* $[2M+Na]^+$ = 523.3.

(4-Benzyloxycarbonylamino-2,3-dihydroxy-1-methyl)-carbamic acid benzyl ester **(4c)**. White solid. Yield: 10 %. ^{1}H-NMR (300 MHz, DMSO-d_6) δ 1.00 (d, *J* = 6.57 Hz, 6H), 3.20 (s, 2H), 3.77 (s, 2H), 4.36 (d, *J* = 5.43 Hz, 2H), 5.00 (s, 2H), 6.82 (d, *J* = 8.55

Hz , 2H), 7.35-7.32 (m, 10H); ^{13}C-NMR (75 MHz, DMSO-d_6): 155.8, 137.4, 128.3, 127.7, 127.6, 72.9, 65.0, 48.5, 18.0.

*General procedure for protection of Cbz-protected amino diols **4(a-c)** to **5(a-c):***

To a solution of **4** (1 eq.; 2.5 mmol) in acetone (15 mL), 2, 2-dimethoxypropane (10 mL) was added at r.t. A catalytic amount of *p*-TsOH (0.25 mmol) was added and the reaction mixture was stirred for 2 h. Upon solvent evaporation, the crude product was purified by flash column chromatography (EA/hexane, 20 % v/v) to isolate the desired product **5**.

{1-[5-(1-Benzyloxycarbonylamino-2-phenyl-ethyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-2-phenyl-ethyl}-carbamic acid benzyl ester **(5a)**. White solid. Yield: 50 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.39 (s, 6H), 2.80 (m, 4H), 3.72 (s, 2H), 3.13-3.95 (m, 2H), 5.02-4.75 (m, 4H), 7.26-7.23 (m, 20H), 7.33 (s, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$): 155.8, 137.4, 136.5, 129.9-127.5, 126.3, 108.5, 80.4, 66.6, 50.2, 39.7, 26.9; LC-MS: *m/z* $[M+23]^+$ = 631.3.

{1-[5-(1-Benzyloxycarbonylamino-3-methyl-butyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-3-methyl-butyl}-carbamic acid benzyl ester **(5b)**. White solid. Yield: 50 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.99-0.79 (t, *J* = 6.57 Hz, 2H), 1.35 (s, 6H), 1.61-1.51 (m, 4H), 3.92-3.65 (m, 2H), 5.29-4.97 (m, 4H), 7.34-7.27 (m, 10H); ^{13}C-NMR (75 MHz, $CDCl_3$): 156.1, 136.8, 128.5-127.9, 108.6, 81.8, 66.7, 47.6, 42.9, 27.4, 24.7, 24.4, 22.9; LC-MS: *m/z* $[M+23]^+$ = 563.3.

{1-[5-(1-Benzyloxycarbonylamino-3-methyl-butyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-1

-methyl)-carbamic acid benzyl ester **(5c)**. White solid. Yield: 60 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.23-1.19 (m, 6H), 1.34 (s, 6H), 3.61 (s, 2H), 3.92 (s, 2H), 5.09 (s, 4H), 7.34-7.29 (m, 10H); ^{13}C-NMR (75 MHz, $CDCl_3$): 155.8, 136.4, 128.3, 127.9, 127.8, 108.6, 80.1, 66.5, 60.1, 45.5, 26.8, 19.3, 13.9.

General procedure for deprotection of the Cbz-protected diols by hydrogenoloysis to ***6(a-c):***

To a solution of **5** in 1:1 MeOH:EA (50 mL) was added 10% Pd/C under H_2 conditions. The reaction mixture was stirred overnight at r.t. The solid was then filtered using celite and the filtrate was concentrated to yield the desired compound **6**.

1-[5-(1-Amino-2-phenyl-ethyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-2-phenyl-ethylamine **(6a)**. Yellow liquid. Yield: 85 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.50 (s, 6H), 2.14 m(s, 4H), 2.61-2.54 (m, 2H), 2.82-2.76 (dd, *J* = 4.77 Hz, 2H), 2.96 (d, *J* = 4.29 Hz, 2H), 4.02 (s, 2H), 7.24-7.14 (m, 10H); ^{13}C-NMR (75 MHz, $CDCl_3$): 134.0, 130.7, 131.5-128.0, 126.3, 108.8, 79.9, 53.6, 41.8, 27.4; LC-MS: *m/z* $[M+H]^+$ = 341.2.

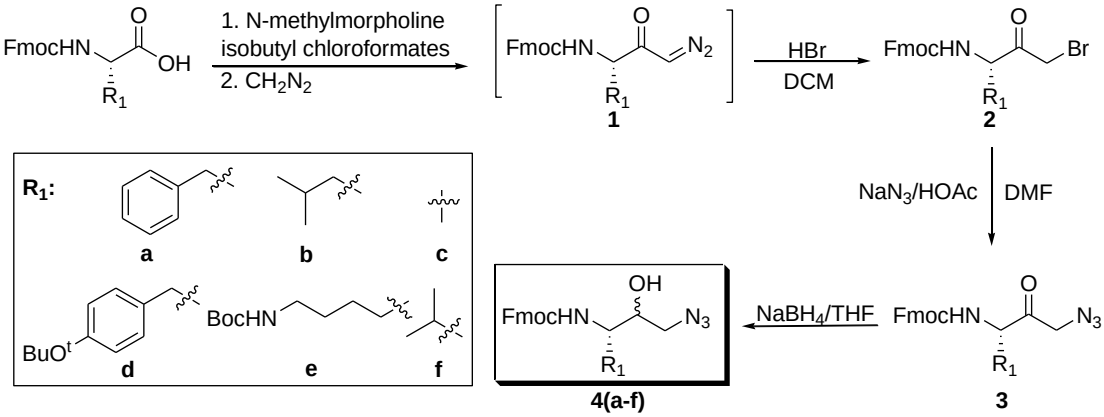
1-[5-(1-Amino-3-methyl-butyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-3-methyl-butylamine **(6b)**. Yellow liquid. Yield: 87 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.98-0.55 (m, 12H), 1.23-1.12 (m, 2H), 1.45-1.28 (m, 2H), 1.77 (s, 4H), 1.94-1.70 (m, 2H), 2.76 (d, *J* = 5.58 Hz, 2H), 3.74 (s, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$): 108.6, 83.3, 81.8, 50.5, 44.5, 27.3, 24.3, 23.4, 21.4; LC-MS: *m/z* $[M+H]^+$ = 273.2.

*1-[5-(1-Amino-1-methyl)-2,2-dimethyl-[1,3]dioxolan-4-yl]-1-methylamine***(6c)**.

Yellow liquid. Yield: 87 %. ^{1}H-NMR (300 MHz, CD_3OD) δ 3.84 (d, *J* = 4.62 Hz, 2H),

3.23-3.19 (m, 2H), 1.35 (s, 6H), 1.18 (d, J = 6.57 Hz, 6H); ^{13}C-NMR (75 MHz, CD_3OD): 111.9, 81.2, 50.3, 27.6, 17.4.

8.2.2 Procedure for Synthesis of N-terminal Azides



Scheme 8.1 Synthesis of the N-terminal warhead building blocks **4(a-f).**

*General procedure for synthesis of bromomethyl ketones **2(a-f):*** Compound **2** were synthesized according to published procedure.[145] To a solution of Fmoc-protection amino acids and N-methylmorpholine (6 mmol) in THF (40 mL) was added isobutyl chloro-formate (6 mmol) at -4 ^{o}C and with stirring. After 15 min, the reaction mixture was filtered. Diazomethane in ether (from 1.33 g, 6.2 mmol of 1-methyl-3-nitrol-nitrosoguanidine and 1.8 mL of 40 wt% aqueous KOH in 10 mL of EtOH) was added to the filtrate slowly and warmed up to room temperature slowly. After completion of the reaction, the solution was concentrated in vacuo to obtain diazo ketone **1** which was dissolved in 30 mL of CH_2Cl_2. 48% HBr (6 mmol) was then added slowly at 0 ^{o}C and stirred for further 15 min. The reaction mixture was finally diluted with EtOAc (40 mL) and washed with saturated sodium bicarbonate, and then washed with brine. The organic layer was dried over anhydrous Na_2SO_4,

filtered, and concentrated under reduced pressure to afford bromomethyl ketone **2(a-f)** in 80-88% yield.

N-α-Fmoc-phenylalanyldiazomethane (**1a**). Yield = 85%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 3.05 (d, *J* = 6.95 Hz, 2H), 4.21 (t, *J* = 6.00 Hz, 1H), 4.41-4.43 (m, 3H), 5.13 (s, 1H), 5.29 (s, 1H), 7.17 (d, *J* = 4.17 Hz, 1H), 7.23-7.32 (m, 6H), 7.40 (t, *J* = 4.53 Hz, 2H), 7.54 (t, *J* = 6.42 Hz, 2H), 7.78 (d, *J* = 7.49 Hz, 2H); ESI-MS: m/z $[M+Na]^+$ = 434.134.

N-α-Fmoc-phenylalanylbromomethylketone (**2a**). Yield = 86%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 3.09-3.14 (m, 2H), 3.77-3.93 (m, 2 H), 4.17 (d, *J* = 7.32 Hz, 1H), 4.42 (d, *J* = 6.30 Hz, 2H), 4.78 (d, *J* = 6.95 Hz, 1H), 5.31 (d, *J* = 6.95 Hz, 1H), 7.13 (d, *J* = 6.95 Hz , 4H), 7.25-7.32 (m, 5H), 7.40 (t, *J* = 7.55 Hz, 2H), 7.53 (t, *J* = 7.88 Hz, 2H), 7.76 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 200.3, 158.7, 143.5, 141.3, 135.5, 129.1, 127.7-124.7(m), 119.9, 74.1, 58.8, 54.0, 18.8; ESI-MS: m/z $[M+Na]^+$ = 486.06.

N-α-Fmoc-isobutylalanylbromomethylketone (**2b**). Yield = 89%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.93 (d, *J* = 10.02 Hz, 6H), 1.43 (t, *J* = 6.16 Hz, 2 H), 1.61 (t, *J* = 6.16 Hz, 2 H), 3.98 (d, *J* = 4.44 Hz, 2H), 4.18 (t, *J* = 6.33 Hz, 1H), 4.43 (t, *J* = 9.05 Hz, 2H), 4.52 (t, *J* = 12.66 Hz, 1H), 7.30-7.55 (m, 4H), 7. 65 (t, *J* = 3.36 Hz, 2H), 7.75 (d, *J* = 7.56 Hz, 2H).

N-α-Fmoc-methylalanylbromomethylketone (**2c**). Yield = 83%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.40 (d, *J* = 7.42 Hz, 3H), 3.93-4.00 (m, 1H), 4.18-4.25 (m, 2H), 4.38-4.66

(m, 3H), 5.39 (s, 2H), 7.29-7.43 (m, 4H), 7.59 (d, *J* = 5.07 Hz, 2H), 7.77 (d, *J* = 7.41 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 201.5, 143.6, 141.3, 127.7, 127.6, 127.0, 124.9, 119.9, 66.8, 53.5, 47.1, 45.9, 31.3, 18.7, 17.8, 17.5.

N-α-Fmoc-4-tert-butoxy-phenylalanylbromomethylketone (**2d**). Yield = 81%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.30 (s, 9H), 3.01 (d, *J* = 6.75 Hz, 2H), 3.68-3.76 (m, 2H), 4.18 (t, *J* = 6.33 Hz, 1H), 4.42 (d, *J* = 6.57 Hz, 2H), 4.64-4.73 (m, 1H), 5.32 (d, *J* = 7.38 Hz, 1H), 6.92 (d, *J* = 8.22 Hz, 2H), 7.02 (d, *J* = 8.22 Hz, 2H), 7.25-7.33 (m, 2H), 7.40 (t, *J* = 7.32 Hz, 2H), 7.53-7.56 (m, 2H), 7.76 (d, *J* = 7.41 Hz, 2H); ESI-MS: m/z $[M+Na]^+$ = 558.12.

N-α-Fmoc-4-Boc-butylalanylbromomethylketone (**2e**). Yield = 85%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.37 (s, 9H), 1.40 (s, 9H), 1.70 (s, 2H), 1.90 (s, 1H), 3.11 (s, 2H), 3.95-4.06 (m, 2H), 4.20 (t, *J* = 6.30 Hz, 1H), 4.39-4.51 (m, 2H), 4.57 (s, 1H), 5.54 (s, 1H), 7.32 (t, *J* = 7.55 Hz, 2H), 7.40 (t, *J* = 7.55 Hz, 2H), 7.59 (d, *J* = 6.95 Hz, 2H), 7.76 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 200.9, 156.1, 143.7, 141.4, 127.8, 127.1, 125.1, 120.0, 79.3, 66.93, 57.9, 47.3, 39.7, 32.0, 30.9, 29.7, 28.0, 22.2, 18.7.

N-α-Fmoc-isopropylalanylbromomethylketone (**2f**). Yield = 84%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.85 (d, *J* = 6.92 Hz, 3H), 1.01 (d, *J* = 6.92 Hz, 3H), 2.21-2.25 (m, 1 H), 3.96-3.4.04 (m, 2H), 4.22 (t, *J* = 6.95 Hz, 1H), 4.45 (t, *J* = 7.09 Hz, 2H), 4.56-4.58 (m, 1H), 5.27 (d, *J* = 3. 08 Hz, 1H), 7.33 (t, *J* = 7.55 Hz, 2H), 7.41 (t, *J* = 7.29 Hz, 2H), 7.59 (d, *J* = 3.75 Hz, 2H), 7.77 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz,

$CDCl_3$) δ 200.7, 155.5, 143.7, 141.4, 127.8, 127.1, 125.0, 124.9, 120.0, 67.0, 62.8, 47.3, 32.7, 30.2, 19.7, 17.0.

General procedure for synthesis of α-azido ketones ***3(a-f):*** α-azido ketones were synthesized according to published procedure with some modifications.[146] To a solution of α-bromo ketone **2** (5 mmol) and acetic acid (10 mmol) in DMF was added NaN_3 (7.5 mmol) in water at 0 ℃. After stirring for overnight at r.t., the reaction was quenched with water and the aqueous layer was extracted with EtOAc. The combined organic extracts were washed with brine, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by silica-gel chromatography with hexane/ethyl acetate (80:20) afforded α-azido ketones **3** in 85-90% yield.

N-α-Fmoc-phenylalanylazidoketone (**3a**). Yield = 79%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.03 (d, *J* = 3.39 Hz, 2H), 3.64 (d, *J* = 11.34 Hz, 1H), 3.90 (d, *J* = 11.34 Hz, 1H), 4.18 (t, *J* = 6.32 Hz, 1H), 4.44 (d, *J* = 6.42 Hz, 2H), 4.53-4.56 (m, 1H), 5.23 (d, *J* = 7.08 Hz, 1H), 7.12 (d, *J* = 6.90 Hz, 2H), 7.28-7.33 (m, 5H), 7.40 (t, *J* = 4.17 Hz, 3H), 7.54 (t, *J* = 5.10 Hz, 2H), 7.77 (d, *J* = 7.56 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 203.4, 155.8, 143.6, 141.1, 135.2, 129.0, 127.8-127.0(m), 124.9, 120.0, 66.9, 58.7, 56.6, 47.2, 37.6; ESI-MS: m/z $[M+Na]^+$ = 449.166.

N-α-Fmoc-isobutylalanylalanylazidoketone (**3b**). Yield = 87%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.94 (d, *J* = 1.90 Hz, 6H), 1.38-1.43 (m, 1H), 1.52-1.56 (m, 2H), 3.96 (d, *J* = 8.30 Hz, 2H), 4.37 (t, *J* = 8.20 Hz, 1H), 4.38-4.51 (m, 3H), 5.11 (d, *J* = 7.55 Hz,

1H), 7.31-7.34 (m, 2H), 7.41 (t, *J* = 7.57 Hz, 2H), 7.58 (t, *J* = 6.95 Hz, 2H), 7.77 (d, *J* = 7.55 Hz, 2H).

N-α-Fmoc-methylalanylazidoketone (**3c**). Yield = 75%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.35 (d, *J* = 6.90 Hz, 3H), 4.03 (d, *J* = 2.50 Hz, 2H), 4.20 (t, *J* = 6.30 Hz, 1H), 4.39-4.44 (m, 2H), 4.48-4.52 (m, 1H), 5.34 (d, *J* = 6.95 Hz, 1H), 7.32 (t, *J* = 6.95 Hz, 2H), 7.41 (t, *J* = 7.55 Hz, 2H), 7.59 (t, *J* = 6.30 Hz, 2H), 7.77 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 201.6, 144.5, 141.3, 126.9, 126.6, 126.0, 124.9, 119.6, 66.7, 53.1, 47.0, 45.9, 31.2, 18.7, 17.9, 17.7.

N-α-Fmoc-4-tert-butoxy-phenylalanylazidoketone (**3d**). Yield = 90%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.38 (s, 9H), 2.97 (d, *J* = 6.90 Hz, 2H), 3.58 (d, *J* = 8.60 Hz, 1H), 3.86 (d, *J* = 9.60 Hz, 1H), 4.08-4.19 (m, 1H), 4.42-4.54 (m, 3H), 5.28 (d, *J* = 7.23 Hz, 1H), 6.92 (d, *J* = 8.22 Hz, 2H), 7.01 (d, *J* = 7.22 Hz, 2H), 7.31 (t, *J* = 7.31 Hz, 2H), 7.40 (t, *J* = 7.40 Hz, 2H), 7.53-7.56 (m, 2H), 7.76 (d, *J* = 7.41 Hz, 2H)，^{13}C-NMR (75 MHz, $CDCl_3$) δ 203.6, 155.7, 154.8, 143.5, 141.3, 129.7, 127.7, 124.8, 119.9, 78.7, 66.9, 58.7, 56.7, 47.2, 37.1, 28.8; ESI-MS: m/z $[M+Na]^+$ = 521.217.

N-α-Fmoc-4-Boc-butylalanylazidoketone (**3e**). Yield = 85%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.23-1.28 (m, 2H), 1.34-1.36 (m, 2H), 1.49 (s, 9H), 1.63-1.71 (m, 2H), 3.09 (s, 2H), 4.02 (s, 2H), 4.1 (t, *J* = 6.96 Hz, 1H), 4.19 (t, *J* = 6.27 Hz, 1H), 4.31-4.43 (m, 1H), 4.49-4.60 (m, 2H), 7.31 (t, *J* = 5.87 Hz, 2H), 7.41 (t, *J* = 7.31 Hz, 2H), 7.59 (d, *J* = 6.96 Hz, 2H), 7.76 (d, *J* = 7.29 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 203.7, 156.2, 143.6, 141.3, 127.8, 127.1, 125.0, 119.9, 79.3, 66.8, 58.0, 55.4, 47.2, 39.5, 30.3, 29.7,

28.4, 22.1.

N-α-Fmoc-isopropylalanylazidoketone (**3f**). Yield = 72%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.86 (d, *J* = 6.30 Hz, 3H), 1.00 (d, *J* = 6.30 Hz, 3H), 2.13-2.17 (m, 1H), 4.01 (d, *J* = 7.55 Hz, 2H), 4.21 (t, *J* = 6.30 Hz, 1H), 4.32 (t, *J* = 4.10 Hz, 1H), 4.39-4.47 (m, 2H), 5.27 (d, *J* = 8.20 Hz, 1H), 7.33 (t, *J* = 7.55 Hz, 2H), 7.41 (t, *J* = 7.25 Hz, 2H), 7.58 (d, *J* = 5.65 Hz, 2H), 7.77 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 203.1, 156.4, 143.6, 141.4, 127.8, 127.1, 125.0, 120.0, 67.0, 62.6, 56.4, 47.3, 30.1, 19.6, 17.1.

General procedure for synthesis of α-azido alcohols ***4(a-f):***

α-azido ketones **3** (5 mmol) was dissolved in 20 mL of THF followed by addition of sodium borohydride (6 mmol) in 2 mL H_2O at 0 ^{o}C. The reaction mixture was then stirred for 1 h and neutralized with aqueous 1 N HCl. After extraction washed with brine and dried over anhydrous Na_2SO_4, and concentrated under reduced pressure. Purification by silica-gel chromatography with hexane/EtOAc (85:15) afforded bromomethyl alcohol **4** in 50-70%.

N-α-Fmoc-phenylalanylazidoalcohol (**4a).** Yield = 55%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 2.84-2.93 (m, 2H), 3.26 (t, *J* = 6.30 Hz, 1H), 3.35 (s, 1H), 3.69-3.89 (m, 2H), 4.12-4.17 (m, 1H), 4.40-4.43 (m, 2H), 4.76 (d, *J* = 6.95 Hz, 1H), 5.07 (d, *J* = 8.80 Hz, 1H), 7.14-7.31 (m, 7H), 7.40 (t, *J* = 7.25 Hz, 2H), 7.46-7.55 (m, 2H), 7.75 (d, *J* = 6.95 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 156.5, 143.7, 141.4, 137.0, 129.3-126.7(m), 124.9, 119.9, 72.2, 66.5, 54.9, 54.1, 47.3, 38.3; LC-MS: m/z $[M+Na]^{+}$ = 451.175.

N-α-Fmoc-isobutylalanylalanylazidoalcohol (**4b**). Yield =96%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.91 (d, *J* = 3.80 Hz, 6H), 1.24-1.36 (m, 2H), 1.48-1.57 (m, 1H), 3.19-3.31 (m, 2H), 3.71-3.76 (m, 2H), 4.19 (t, *J* = 5.97 Hz, 1H), 4.45-4.50 (m, 1H), 4.52-4.56 (m, 1H), 4.74 (t, *J* = 8.15 Hz, 1H), 4.86 (t, *J* = 9.50 Hz, 1H), 7.31 (t, *J* = 7.25 Hz, 2H), 7.39 (t, *J* = 6.95 Hz, 2H), 7.57 (d, *J* = 6.30 Hz, 2H), 7.75 (d, *J* = 7.55 Hz, 2H), ^{13}C-NMR (125 MHz, $CDCl_3$) δ 156.7, 143.7, 141.4, 127.7, 127.0, 124.8, 119.9, 73.5, 66.3, 53.7, 52.4, 47.4, 24.6, 23.2, 21.5; LC-MS: m/z $[M+Na]^+$ = 417.189.

N-α-Fmoc-methylalanylazidoalcohol (**4c)**. Yield = 75%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.18 (d, *J* = 5.91 Hz, 3H), 3.25-3.33 (m, 2H), 3.65-3.79 (m, 2H), 4.24 (t, *J* = 6.24 Hz, 1H), 4.55 (d, *J* = 6.72 Hz, 1H), 5.00 (s, 1H), 7.33-7.47 (m, 4H), 7.62 (d, *J* = 7.41 Hz, 2H), 7.81 (d, *J* = 7.56 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 157.5, 142.6, 139.4, 137.4, 128.9, 127.3, 121.4, 120.0, 109.6, 72.0, 53.3, 48.9, 16.7**.**

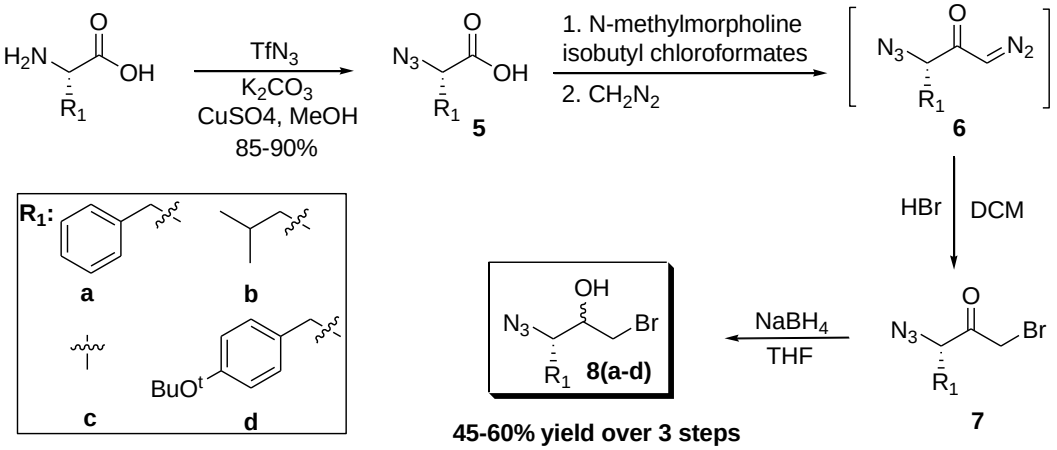
N-α-Fmoc-4-tert-butoxy-phenylalanylazidoalcohol (**4d**): Yield = 90%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.30 (s, 9H), 2.78-2.84 (m, 2H), 3.25-3.32 (m, 2H), 3.72-3.78 (m, 2H), 4.11-4.17 (m, 1H), 4.39-4.45 (m, 2H), 4.79 (d, *J* = 6.95 Hz, 1H), 5.08 (d, *J* = 8.85 Hz, 1H), 6.88 (d, *J* = 7.55 Hz, 2H), 7.02-7.07 (m, 2H), 7.30 (t, *J* = 7.23 Hz, 2H), 7.39 (t, *J* = 7.25 Hz, 2H), 7.51 (t, *J* = 7.55 Hz, 2H), ^{13}C-NMR (125 MHz, $CDCl_3$) δ 156.4, 154.2, 143.6, 141.4, 132.2-124.8(m), 124.2, 119.9, 78.4, 72.1, 66.5, 54.9, 54.2, 47.3, 35.2, 28.8; LC-MS: m/z $[M+Na]^+$ = 523.207.

N-α-Fmoc-4-Boc-butylalanylazidoalcohol (**4e**). Yield = 88%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.23-1.28 (m, 4H), 1.45 (s, 9H), 3.00-3.09 (m, 2H), 3.24-3.28 (m, 2H),

3.48-3.75 (m, 2H), 4.19 (t, *J* = 6.3 Hz, 1H), 4.42-4.61 (m, 2H), 5.01 (s, 1H), 5.13 (d, *J* = 8.85 Hz, 1H), 7.31 (t, *J* = 7.55 Hz, 2H), 7.40 (t, *J* = 7.55 Hz, 2H), 7.58 (d, *J* = 6.95 Hz, 2H), 7.76 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 156.5, 143.7, 141.3, 127.7, 127.0, 124.9, 119.9, 79.3, 73.1, 66.4, 60.4, 53.9, 47.4, 39.5, 31.3, 29.5, 22.4, 20.9; LC-MS: m/z $[M+H]^+$ = 410.205.

N-α-Fmoc-isopropylalanylazidoalcohol (**4f**). Yield = 79%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.72 (d, *J* = 9.55 Hz, 6H), 2.58 (d, *J* = 5.65 Hz, 1H), 3.19-3.41 (m, 2H), 3.53-3.58 (m, 1H), 3.67 (s, 1H), 4.49 (t, *J* = 5.98 Hz, 1H), 4.59-4.64 (m, 2H), 7.33 (t, *J* = 7.57 Hz, 2H), 7.40 (t, *J* = 3.78 Hz, 2H), 7.59 (t, *J* = 5.98 Hz, 2H), 7.77 (d, *J* = 7.55 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 157.1, 143.7, 141.5, 127.7, 127.1, 124.9, 120.0, 71.2, 66.3, 58.6, 54.4, 47.5, 28.0, 20.0; LC-MS: m/z $[M+Na]^+$ = 403.154.

8.2.3 Procedure for Synthesis of C-terminal Azides



Scheme 8.2 Synthesis of the C-terminal warhead building blocks **8(a-d).**

*General procedure for synthesis of α-azido acid **5(a-d):***

The diazo transfer reaction was performed according to the reported procedure.[111] *L*-Amino acids (10 mmol, 1.0 eqv) was combined with K_2CO_3 (15 mmol, 1.5 eqv) and

$CuSO_4$ pentahydrate (0.1 mmol, 0.01 eqv), distilled H_2O (20 mL), and CH_3OH (40 mL). Triflyl azide in CH_2Cl_2 (20 mmol, 2 equiv.) prepared according to the reported procedure was added with stirring, and the resulting mixture was stirred at ambient temperature overnight. Subsequently, the organic solvents were removed under reduced pressure and the aqueous slurry was diluted with H_2O (50 mL). This slurry was then acidified to *p*H 6 with 2N HCl aqueous solution and diluted with 0.25 M, *p*H 6.2 phosphate buffer (50 mL) and extracted with EtOAc to remove the sulfonamide byproduct. The aqueous phase was then acidified to *p*H 2 with 2N HCl aqueous solution. The product was obtained from another round of EtOAc extractions. The EtOAc extracts were combined, dried with anhydrous Na_2SO_4 and evaporated to dryness giving the α-azido acid **5(a-d)** in 70-85% yield.

2-azido-3-phenylpropanoic acid (**5a**). Yield = 84%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.04 (dd, *J* = 13.98, 8.88 Hz, 1H), 3.24 (dd, *J* = 13.98, 4.92 Hz, 1H), 4.13-4.18 (m, 1H), 7.25-7.33 (m, 5H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 175.6, 135.6, 129.2, 128.7, 127.4, 63.1, 37.5; LC-MS: m/z $[M-H]^-$= 190.1.

2-azido-4-methylpentanoic acid (**5b**). Yield = 73%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.99 (d, *J* = 6.95 Hz, 6H), 1.68-1.78 (m, 2H), 1.80-1.85 (m, 1H), 3.86-3.89 (m, 1H), 10.77 (br s, 1H); LC-MS: m/z 2$[M+H]^+$ = 313.0.

2-azidopropanoic acid (**5c**). Yield = 75%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.54 (d, *J* = 6.95 Hz, 3H), 4.02-4.06 (m, 1H), 10.6 (brs, 1H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 176.9, 57.0, 16.5.

2-azido-3-(4-tert-butoxyphenyl)propanoic acid (**5d**). Yield = 80%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.34 (s, 9H), 2.99 (dd, *J* = 13.92, 8.70 Hz, 1H), 3.19 (dd, *J* = 14.28, 4.89 Hz, 1H), 4.09-4.14 (m, 1H), 6.74 (d, *J* = 8.70 Hz, 2H), 7.15 (d, *J* = 8.37 Hz, 2H), 8.24 (brs, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 174.8, 154.4, 130.6, 129.7, 124.4, 78.9, 63.2, 36.9, 28.8.

General procedure for synthesis of α-azido bromomethyl ketones ***7(a-d):***

Isobutyl chloroformate (9.6 mmol, 1.2 eqv) was added to a solution of α-azido acid **5(a-d)** (8.0 mmol, 1.0 eqv) and N-methylmorpholine (9.6 mmol, 1.2 eqv) in THF (20 mL) at -40 ^{o}C, and the reaction mixture was stirred for 15 min. The reaction mixture was then filtered and diazomethane (4.0 eqv, made from 1-methyl-3-nitro1-nitrosoguanidine and of 40 wt% aqueous KOH in EtOH) was added slowly into the filtrate. The reaction mixture was stirred for 2-3 h at 0 ^{o}C. After completion of reaction, the solvent was evaporated under vacuum and the residue was dissolved in DCM followed by the treatment with 48% HBr aq (1.2 eqv) at 0 ^{o}C. The reaction mixture was then stirred for 3 h. After completion of the reaction, it was diluted with 15 wt% aqueous citric acid and extracted with DCM. The organic layer was washed with saturated sodium bicarbonate and brine. Finally, it was dried over anhydrous sodium sulfate, filtered, concentrated and purified by silica-gel chromatography with 80:20 hexane/EtOAc afforded α-azido bromomethyl ketones **7(a-d)** in 60-93%.

3-azido-1-bromo-4-phenylbutan-2-one（**7a**） Yield = 74%. ^{1}H-NMR (300 MHz,

$CDCl_3$) δ 2.99 (d, *J* = 8.55 Hz, 1H), 3.25 (d, *J* = 5.58 Hz, 1H), 3.86 (d, *J* = 4.59 Hz, 1H), 3.95 (d, *J* = 6.06 Hz, 1H), 4.35-4.40 (m, 1H), 7.25-7.36 (m, 5H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 198.3, 135.3, 129.2-128.6(m), 127.5, 66.8, 37.5, 32.3.

3-azido-1-bromo-5-methylhexan-2-one (**7b**). Yield = 65%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.73-1.17 (m, 6H), 1.62 (t, *J* = 7.48 Hz, 2H), 1.80-1.88 (m, 1H), 3.38 (t, *J* = 5.91 Hz, 1H), 4.05 (d, *J* = 6.27 Hz, 1H), 4.19 (d, *J* = 7.68 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 198.9, 64.7, 40.6, 32.7, 25.4, 23.5, 22.7.

3-azido-1-bromo-5-methy-2-one (**7c**). Yield = 60%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.52 (d, *J* = 6.30 Hz, 3H), 4.08-4.11 (m, 1H), 4.29-4.31 (m, 1H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 198.8, 61.0, 31.2, 15.8.

3-azido-1-bromo-4-(4-tert-butoxyphenyl)butan-2-one (**7d**). Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.40 (s, 9H), 2.93-3.01 (m, 1H), 3.09-3.23 (m, 1H), 3.79 (t, *J* = 11.26 Hz, 1H), 3.90-4.09 (m, 1H), 4.17-4.34 (m, 1H), 6.96(d, *J* = 8.22 Hz, 2H), 7.12 (d, *J* = 8.40 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 198.4, 154.8, 129.8, 124.4, 66.9, 36.9, 28.5, 18.8.

General procedure for synthesis of α-azido bromomethyl alcohol ***8(a-d):***

To the solution of α-azido bromomethyl ketones **7(a-d)** (8 mmol, 1 eqv) in 20 mL of 95:5 THF/H_2O, sodium borohydride (12 mmol. 1.5 eqv) was added gradually at 0 ^{o}C. The reaction mixture was stirred for 1 h and then was neutralized with aqueous 1N HCl. After extraction with EtOAc, the organic extracts were washed with brine, dried with Na_2SO_4, and concentrated under reduced pressure. Purification by

silica-gel chromatography with hexanes/EtOAc (80:20) afforded compound **8(a-d)** as white solid.

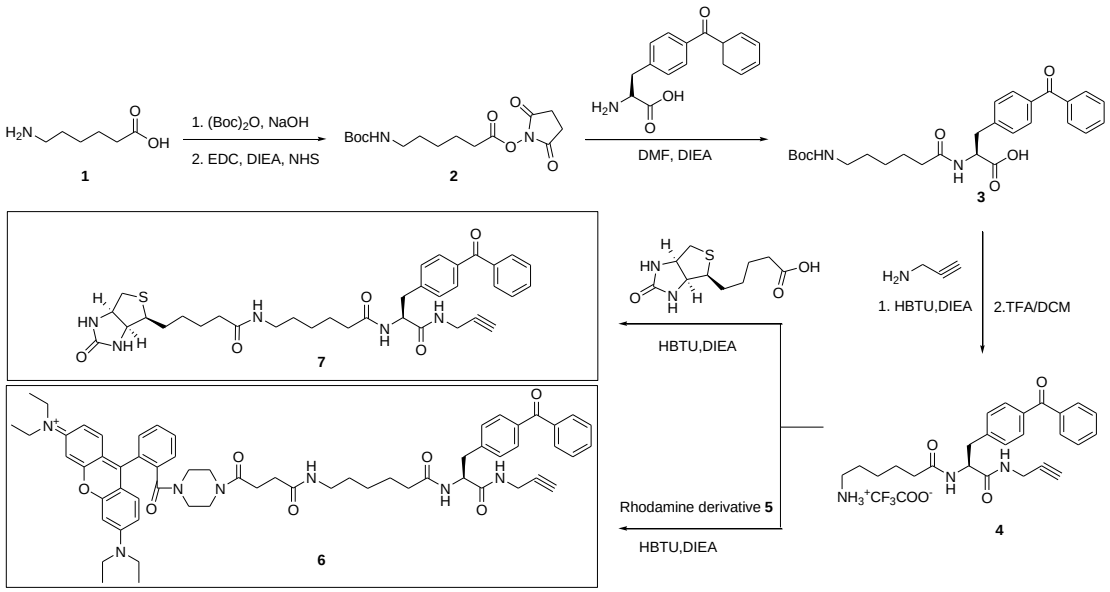
α-azido bromomethyl-Phe-alcohol (**8a**). Yield = 75%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 2.40 (d, *J* = 4.44 Hz, 1H), 2.80-3.18 (m, 2H), 3.48-3.52 (m, 1H), 3.65 (d, *J* = 6.09 Hz, 1H), 7.29-7.35 (m, 5H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 136.9, 136.6-128.6 (m), 127.0, 126.9, 72.1, 65.9, 64.8, 37.0, 35.2. LC-MS: m/z $[M+Na]^+$ = 293.990.

α-azido bromomethyl-Leu-alcohol (**8b**). Yield = 90%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 0.97 (d, *J* = 2.50 Hz, 6H), 1.38-1.49 (m, 1H), 1.60-1.69 (m, 1H), 1.73-1.84 (m, 1H), 2.53 (d, *J* = 5.65 Hz, 1H), 3.31-3.54 (m, 2H), 3.56-3.72 (m, 1H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 73.4, 62.9, 39.8, 36.6, 25.3, 23.3, 21.8. LC-MS: m/z $2[M]^+$ = 388.089.

α-azido bromomethyl-Ala-alcohol (**8c**). Yield = 80%. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 1.36 (d, *J* = 6.57 Hz, 3H), 2.37 (t, *J* = 5.43 Hz, 1H), 3.41-3.78 (m, 4H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 73.7, 59.3, 36.5, 34.9, 15.6, 14.9. LC-MS: m/z 2[M] = 388.089.

α-azido bromomethyl-Tyr(^{t}Bu)-alcohol (**8d**). Yield = 77%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.33 (s, 9H), 2.51 (dd, *J* = 5.07 Hz, 1H), 2.79 (d, *J* = 5.66 Hz, 1H), 2.90-3.14 (m, 2H), 3.52-3.59 (m, 1H), 3.63-3.75 (m, 2H), 6.96 (d, *J* = 8.40 Hz, 2H), 7.17 (d, *J* = 6.60 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 154.3, 146.9, 131.8-131.5(m), 129.8, 124.5, 72.2, 71.9, 69.2, 65.0, 37.3, 36.4, 28.8.

8.2.4 Synthesis of Rhodamine/biotin Bp Alkyne Compounds



Scheme 8.3 Synthetic scheme of rhodamine and biotin alkyne compounds.

Alkyne **6** and **7** were synthesized according to published procedures with some modifications. Couplings of **4** with rhodamine derivative **5** or biotin acid afford alkyne **6** and **7**, respectively. Column purification (5%-10% MeOH/DCM) was carried out to isolate the pure product.

NHS ester (**2**): To a solution of 6-aminohexanoic acid 1 (6.56 g, 50 mmol) and NaOH (2 M, 40 mL) in dioxane/water (2/1; 150 mL) at 0 ℃ was added $(Boc)_2O$ (13.1 g, 60 mmol). The reaction was allowed to proceed at r.t. overnight. Subsequently, dioxane was removed under reduced pressure and the resulting mixture was acidified to *p*H 2 with 1 M HCl, followed by extraction with EA (3×50 ml). The combined organic layers were dried over Na_2SO_4, filtered and concentrated to afford colorless oil, which was subsequently reacted with NHS (6.72 g, 58 mmol) and EDC (11.2 g, 58 mmol) in DMF at r.t. overnight. Upon column purification (20 - 40% EA/hexane), the desired product **2** was obtained as colorless oil (11.5 g, 35.0 mmol, 70%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 4.60 (br s, 1H), 3.16-3.05 (m, 2H), 2.82 (s, 4H), 2.60 (t, *J* = 7.49 Hz,

2H), 1.81-1.70 (m, 2H), 1.51-1.40 (m, 13H); ESI-MS: m/z $[M+Na]^+$ = 351.3.

3-(4-benzoylphenyl)-2-(6-(tert-butoxycarbonyl) hexanamido) propanoic acid (**3**): Compound **2** (1.64 g, 5 mmol) dissolved in 15 mL DMF was added H-p-Bz-Phe-OH (1.35 g, 5 mmol) and DIEA (1.05 ml, 6 mmol). The reaction mixture was stirred under N_2 overnight. After that, DMF was removed in vacuo. The residue was taken into EA and washed with 1 M HCl. The EA layer was dried over Na_2SO_4, filtered and concentrated. Upon column chromatography (5-10 % MeOH/DCM), the final product **3** was obtained as colorless oil (1.98 g, 4.10 mmol, 82%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.08 (s, 1H), 7.74-7.66 (m, 4H), 7.63-7.56 (m, 1H), 7.49-7.44 (m, 2H), 7.31-7.26 (m, 2H), 4.83 (br s, 1H), 3.41-2.96 (m, 2H), 2.21-2.05 (m, 2H), 1.61-1.47 (m, 2H), 1.47-1.32 (m, 11H), 1.30-1.14 (m, 2H); ESI-MS: m/z $[M-1]^-$ = 481.5.

6-amino-N-(1-(2-alkynethylamino)-3-(4-benzoylphenyl)-1-oxopropan-2-yl)hexanamide **(4)**: To compound **3** (1.93 g, 4 mmol) dissolved in DMF was added HBTU (1.90 g, 5 mmol) and DIEA (0.87 ml, 5 mmol) at 0 ^{o}C. The mixture was stirred for 10 min before addition of propylamine (0.28 g, 5 mmol). The mixture was further agitated for another 12 h at r.t., after which DMF was removed *in vacuo* and the residue was taken into ethyl acetate (50 mL). The organic layer was washed with saturated $NaHCO_3$ (2×30 mL), 1 M HCl (2×30 mL), brine (2×30 mL), dried over anhydrous Na_2SO_4, and concentrated under reduced pressure to afford a yellow oily product. Purification of this compound by flash chromatography (silica gel, ethyl acetate/hexane = 3:1) afforded the intermediate as a white solid (1.65 g, 3 mmol, 75%). ^{1}H-NMR (300 MHz,

$CDCl_3$) δ 7.77-7.69 (m, 4H), 7.58 (t, *J* = 7.49 Hz, 1H), 7.46 (t, *J* = 7.49 Hz, 2H), 7.29 (t, *J* = 8.19 Hz, 2H), 6.98 (t, *J* = 5.1 Hz, 1H), 6.63 (d, *J* = 7.65 Hz, 1H), 4.86-4.78 (m, 1H), 3.89-4.05 (m, 2H), 4.05- 3.89 (m, 2H), 3.21-3.02 (m, 4H), 2.20-2.16 (m, 2H), 1.62-1.52 (m, 2H), 1.30 (s, 12H), 1.27-1.23 (m, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 196.3, 173.3, 170.6, 141.3, 156.1, 141.5, 137.5, 136.3, 132.4, 130.4, 129.9, 129.3, 128.3, 79.0, 71.7, 53.8, 40.3, 38.3, 36.1, 29.6, 29.1, 28.4, 26.1, 25.0.

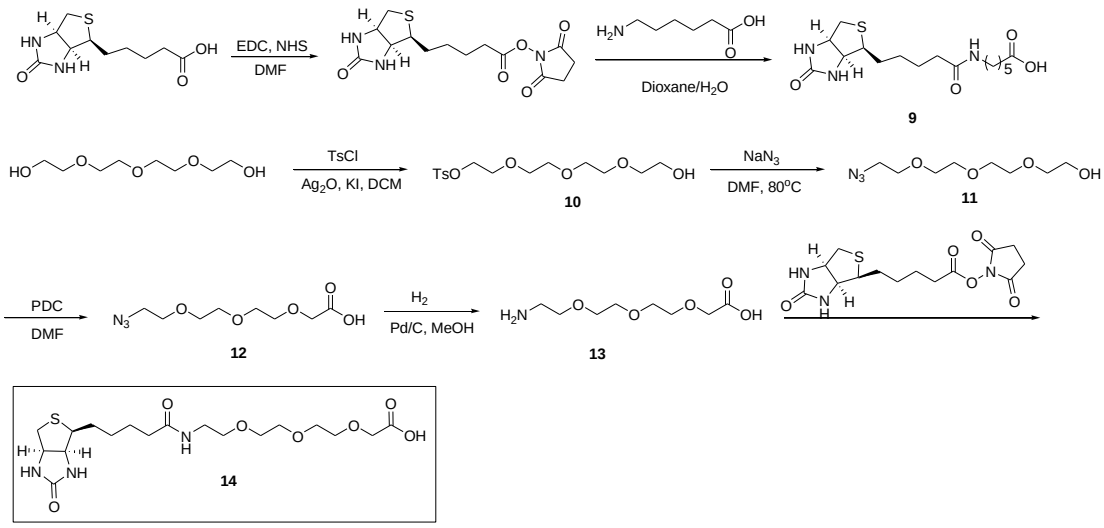
Deprotection of above intermediate (1.56 g, 3 mmol) with TFA (10 eq) in DCM gave **4** as yellow oil (1.26 g, 3 mmol, ~100%). ^{1}H-NMR (300 MHz, MeOD) δ 7.76-7.69 (m, 4H), 7.63 (d, *J* = 7.32 Hz, 1H), 7.52 (t, *J* = 7.49 Hz, 2H), 7.42 (d, *J* = 8.01 Hz, 2H), 4.85- 4.69 (m, 1H), 3.95 (t, *J* = 2.63 Hz, 2H), 3.31-3.19 (m, 1H), 3.30-2.94 (m, 1H), 2.91-2.85 (m, 2H), 2.63-2.58 (m, 1H), 2.21-2.16 (m, 2H), 1.55-1.38 (m, 4H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 198.3, 175.9, 172.9, 143.9, 138.9, 137.3, 133.8, 131.2, 130.9, 130.5, 129.5, 80.3, 72.3, 55.4, 42.0, 39.1, 36.6, 32.6, 31.0, 29.5, 27.9, 27.3, 26.5, 14.5.

Alkyne **6**: Yield: 50 %. ^{1}H-NMR (300 MHz, MeOD) δ 7.62-7.47 (m, 7H), 7.47-7.36 (m, 1H), 7.34 (d, *J* = 8.01 Hz, 1H), 7.21 (d, *J* = 9.39 Hz, 1H), 7.00 (d, *J* = 8.34 Hz, 1H), 4.15 (d, *J* = 5.58 Hz, 4H), 3.88 (s, 1H), 3.63-3.58 (m, 4H), 3.41-3.19 (m, 4H), 3.31-2.90 (m, 2H), 2.52 (s, 1H), 2.38-2.09 (m, 2H), 1.39-1.21 (m, 8H), 0.91-0.84 (m, 14H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 172.9, 169.3, 159.3, 157.1, 143.9, 138.9, 136.8, 133.6, 133.2, 131.5, 130.5, 129.7, 128.9, 115.4, 114.9, 97.4, 80.4, 72.3, 69.1, 55.4, 46.9, 40.2, 39.0, 36.6, 31.6, 30.7, 30.1, 29.7, 27.3, 26.4, 24.9, 23.9, 14.4, 12.8, 11.4.

ESI-MS: m/z $[M+1]^+$ = 1012.4459.

Alkyne **7**: Yield: 80 %. ^{1}H-NMR (300 MHz, DMSO-d_6) δ 9.05 (s, 1H), 8.67-7.34 (m, 3H), 8.16-8.09 (m, 6H), 8.04 (t, *J* = 7.44 Hz, 2H), 7.88 (d, *J* = 8.19 Hz, 2H), 5.82(s, 2H), 5.15-5.02 (m, 1H), 4.59-4.82 (m, 2H), 4.25-4.58 (m, 2H), 3.88-3.51 (m, 4H), 3.41-3.24 (m, 7H), 2.08-1.94 (m, 4H), 1.94-1.45 (m, 12H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 195.5, 171.9, 170.9, 162.5, 143.3, 137.3, 135.1, 133.3, 132.5, 129.4, 128.5, 124.5, 119.1, 115.4, 109.7, 80.9, 73.1, 61.1, 59.3, 55.4, 53.3, 40.3, 40.1, 38.9, 37.7, 36.5, 35.8, 34.1, 33.5, 30.8, 28.9, 28.1, 27.8, 26.6, 25.6, 24.7, 23.4, 17.9, 16.7, 12.2. ESI-MS: m/z $[M+1]^+$ = 646.2415.

8.2.5 Chemical Synthesis of Biotin Acid Linkers



Scheme 8.4 Synthesis of two biotin linkers**.**

*Biotin acid (**9**).* Compound **9** was synthesized according to previous report with some modification.[148] To a solution of D-biotin (12 g, 50 mmol) in DMF was added N-hydroxysuccinimide (7 g, 60 mmol) and EDC (11.5 g, 60 mmol). The reaction was allowed to proceed overnight. The resulting mixture was dried in vacuo to remove

DMF. The gel-like residue was recrystallized from EtOH/Acetic acid/H_2O (95:1:4) to afford biotin-NHS as a white solid (14 g; 80% yield). To a solution of 6-aminohexanoic acid (1.3 g, 10 mmol) in 30 ml of 1:1 Dioxane/water co-solvent, the pH of the solution was adjusted to 8~8.5 by using a 4 M NaOH solution at 0 ^{o}C. Biotin-NHS (3.4 g, 10 mmol) was subsequently added, and the reaction mixture was stirred at r.t. overnight. The resulting gelatinous solid formed was added ether, stirred for 5 minutes. The supernatant ether was decanted followed by addition of acetone and adjusting the *p*H of the resulting solution to 3 using 2 M HCl at 0 ^{o}C. Finally, the solid was filtered, washed several times with MeOH to afford the pure compound **9** as a white solid (2.85 g, 80%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.23-1.62 (m, 12H), 2.04 (t, *J* = 7.41 Hz, 2H), 2.19 (t, *J* = 7.23 Hz, 2H), 2.57 (d, *J* = 12.15 Hz, 1H), 2.82 (dd, *J* = 12.48, 4.92 Hz, 1H), 2.97-3.16 (m, 3H), 4.13 (t, *J* = 4.29 Hz, 1H), 4.30 (t, *J* = 7.38 Hz, 1H), 6.43 (s, 2H), 7.73 (t, *J* = 5.28 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 174.8, 172.2, 163.1, 61.4, 59.6, 55.8, 38.6, 35.6, 33.9, 29.2, 28.5, 26.3, 25.7, 24.6; LC-MS: m/z $[M+H]^+$ = 357.282.

*Biotin PEG acid(**14**).* Compound **14** was synthesized according to published procedures with some modifications,[147] as follows.

*Compound **10***. To a stirred solution of diol (1 mmol) in CH_2Cl_2 (10 mL) was added fresh Ag_2O (0.35 g, 1.5 mmol), TsCl (0.21 g, 1.1 mmol), and KI (0.033 g, 0.2 mmol). The reaction mixture was stirred at r.t. for 8 h, then filtered through a small pad of silica gel, and washed with EtOAc. Evaporation of the solvent, followed by column chromatography, gave the desired monotosylate product **10** (Yield = 70%). ^{1}H-NMR

(300 MHz, $CDCl_3$) δ 2.43 (s, 3H), 3.27(s, 1H), 3.57-3.68 (m, 14H), 4.15 (t, *J* = 4.44 Hz, 2H), 7.34 (d, *J* = 8.22 Hz, 2H), 7.78 (d, *J* = 8.04 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 144.4, 132.4, 129.4, 127.4, 72.0, 70.0-68.9(m), 68.1, 60.9, 21.1.

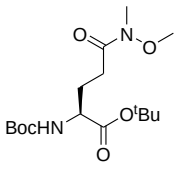
*Compound **11.*** To a solution of **10** (10 mmol) in DMF, NaN_3 (30 mmol) was added. The reaction was then stirred for 8 h at 80 ^{o}C. After completation of reaction, DMF was firstly removed under high vacuum, and the residue was purified by column chromatography to afford the desired product in 80% yield. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.39 (t, *J* = 5.10 Hz, 2H), 3.59 (t, *J* = 4.92 Hz, 2H), 3.67-3.71 (m, 10H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 72.4, 69.8-70.4(m), 61.3, 50.4.

*Compound **12**.* To a solution of **11** (10 mmol) in DMF, Pyridinium dichromate (PDC**)** (50 mmol) was added at 0 ^{o}C. Then the reaction was stirred overnight at r.t. After completation of reaction, a lot of brine was added to quench the reaction followed by EA extraction. The extrate was dried by Na_2SO_4, filtrated and evaporated under high vacuum. Finally the residue was purified by column chromatography to afford the desired product **12** in 55% yield. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.40 (t, *J* =4.65 Hz, 2H), 3.68-3.72 (m, 10H), 4.14 (s, 2H), 8.02 (s, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 172.1, 69.5-72.1(m), 68.0, 63.5, 50.2.

*Compound **13 & 14***. To a solution of **12** (10 mmol) in MeOH, Pd/C (20%w) was added at r.t. The reaction was then stirred overnight. After completation of reaction, Pd/C was filtrated by celite. The filtrate was evaporated and the residue was purified by column chromatography to afford compound **13** in 70% yields. To a solution of

above PEG amino acid in 30 mL of 1:1 dioxane/water co-solvent, the *p*H of the solution was adjusted to 8~8.5 by using a 4 M NaOH solution at 0 ^{o}C. Biotin-NHS (3.4 g, 10 mmol) was subsequently added, and the reaction mixture was stirred at r.t. overnight. The resulting gelatinous solid formed was added ether, stirred for 5 mins. The supernatant ether was decanted followed by addition of acetone and adjusting the *p*H of the resulting solution to 3 using 2 M HCl at 0 ^{o}C. Finally, the solid was filtered, washed several times with MeOH to afford the pure compound **14** as a white solid in 85% yield. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.34-1.69 (m, 6H), 2.15 (t, *J* = 6.99 Hz, 2H), 2.56-2.61 (m, 1H), 2.77-2.92 (m, 2H), 3.11-3.15 (m, 4H), 3.22-3.66 (m, 8H), 4.05 (s, 2H), 4.21-4.25 (m, 2H), 4.40-4.79 (m, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 176.2, 174.23, 167.7, 70.2-71.7(m), 69.2, 63.3, 61.6, 56.9, 41.0, 40.4, 36.7, 36.2, 33.8, 31.4, 29.5, 26.7, 25.6; LC-MS: m/z $[M+H]^+$ = 434.134.

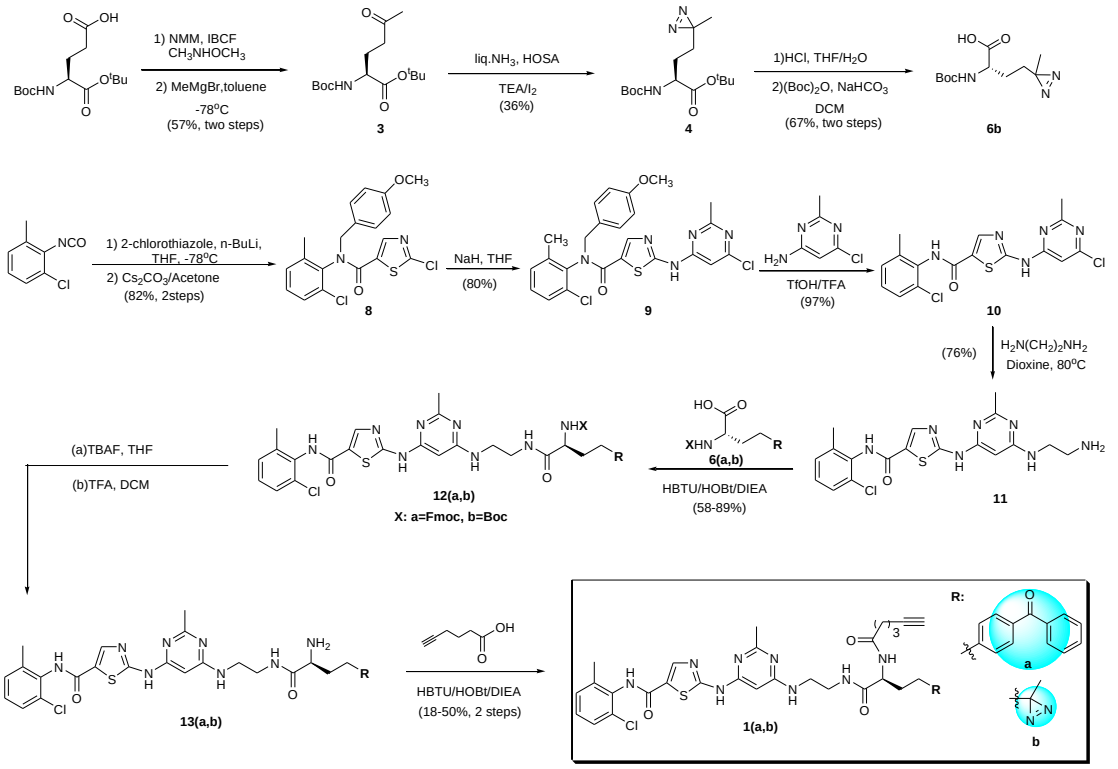
8.2.6 Chemical Synthesis of Dasatinib-like Probes



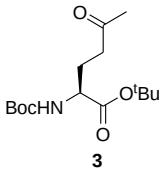
2

To a solution of 5-(tert-butoxy)-4-((tert-butoxycarbonyl)amino)-5-oxopentanoic acid (10.3 g, 33.9 mmol) in DCM, *N*-methylmorpholine (9.3 mL, 84.8 mmol) and isobutyl chloroformate (6.0 mL, 46.2 mmol) were added at 0 ℃. After stirring for 0.5 h at 0 ℃, *N*, *O*-dimethylhydroxylamine (3.9 g, 40.0 mmol) was added slowly, and the reaction was then warmed up to r. t. and stirred for further 4 h. After that, *p*H of the reaction mixture was adjusted to 7 with HCl (1 M) and extracted with DCM (300

mL×3). The organic layer was then dried with Na_2SO_4, concentrated and purified with flash chromatography (hexane: EtOAc=10:1/5:1) to give the product **2** (9.17 g, 78.4%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.24 (d, *J* = 8.1 Hz, 1H), 4.12 (d, *J* = 4.5 Hz, 1H), 3.61 (s, 3H), 3.10 (s, 3H), 2.44 (m, 2H), 2.05 (m, 1H), 1.90 (m, 1H), 1.40 (s, 9H), 1.37 (s, 9H); ^{13}C NMR (300 MHz, $CDCl_3$) δ 173.4, 171.3, 155.3, 81.6, 79.2, 61.0, 53.6, 32.0, 28.1, 27.7, 27.2.

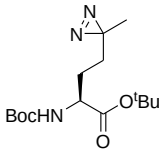


Scheme 8.5 Synthesis of dasatinib analogues (**1a, b)**



To a solution of compound **2** (9.17 g, 26.5 mmol) in toluene (200 mL), methylmagnesium bromide (40 mL, 3M solution in hexanes) was added at -78 ℃. The resulting mixture was then stirred for 3 h with the temperature up to -5 ℃. Subsequently, the reaction was quenched by addition of water (20 mL) at -5 ℃

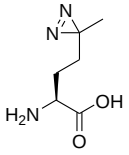
followed by extraction with EtOAc (200 mL×3). The separated organic layer was washed with brine (200 mL×3) and dried with Na_2SO_4, concentrated and purified with flash chromatography (hexane:EtOAc=10:1/5:1) to give the desired product **3** (5.83 g, 73.1%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.10 (s, 1H), 4.09 (s, 1H), 2.42-2.55 (m, 2H), 2.12 (s, 3H), 2.05-2.07 (m, 1H), 1.78-1.85 (m, 1H), 1.42 (s, 18 H); ^{13}C NMR (300 MHz, $CDCl_3$) δ 207.4, 171.3, 155.3, 81.8, 79.5, 53.3, 39.2, 29.8, 28.1, 27.8, 26.6.



4

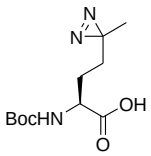
Compound **3** (5.83 g, 19.3 mmol) was added to liquid ammonia (20 mL) at -78 ℃. The reaction mixture was stirred overnight in the sealed tube. Subsequently, the reaction was cooled down to -78 ℃, and hydroxylamine-O-sulfonic acid (2.60 g, 23.0 mmol) in methanol (12 mL) was added over 30 min. The reaction mixture was then stirred overnight at r.t. while the ammonia was evaporated concurrently. The reaction mixture was filtrated and washed with methanol (30 mL×2). The combined methanol layer was then concentrated to half of the original volume. 50 mL of methanol was added to the residues and cooled down to 0 ℃ followed by addition of TEA (3.1 mL, 23.0 mmol). Subsequently, Iodine was added slowly until the brown color of solution formed. The reaction mixture was stirred for further 1 h at r.t, and concentrated under vacuum. The aqueous solution was then extracted with EtOAc (200 mL) and washed with brine (200 mL). After dried with Na_2SO_4, the organic layer was concentrated and purified with flash chromatography (hexane:EtOAc = 30:1/10:1) to give the product **4** (2.2 g, 36.4 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.08 (d, J = 7.8

Hz, 1H), 4.07 (s, 1H), 1.60-1.65 (m, 1H), 1.28-1.50 (m, 22H), 0.95 (s, 3H); ^{13}C NMR (300 MHz, $CDCl_3$) δ 171.1, 155.1, 81.8, 79.5, 53.1, 30.2, 28.1, 27.7, 27.1, 25.1, 19.4.



5

To a solution of compound **4** (2.2 g, 7.04 mmol) in THF (15 mL), HCl (4 M, 15 mL) was added slowly at 0 ℃. The resulting solution was stirred for 48 h at r. t. and the solvent was evaporated to give product **5** (1.0 g, 74.1%). ^{1}H NMR (300 MHz, D_2O) δ 3.94 (t, *J* = 6.6 Hz, 1H), 1.77-1.84 (m, 2H), 0.98-1.53 (m, 2H), 0.98 (s, 3H).



6b

To a solution of compound **5** (150 mg, 0.78 mmol) in water (5 mL), sodium hydroxide solution (80 mg, 2 mmol) was added at 0 ℃ followed by addition of $(Boc)_2O$ (170 mg, 0.78 mmol) in THF (10 mL). The resulting mixture was then stirred for 20 h at r.t. in the dark and concentrated under vacuum. The residue was re-dissolved in water (10 mL), and *p*H was adjusted to 6. Subsequently, the aqueous solution was extracted with EtOAc (20 mL×3) and the combined organic layer was dried with Na_2SO_4, concentrated, purified with flash chromatography to give the product **6b** (180 mg, 90.0 %). ^{1}H NMR (300 MHz, MeOD) δ 4.05 (s, 1H), 1.24-1.71 (m, 13H), 1.01 (s, 3H).

2-chloro-N-(2-chloro-6-methylphenyl) thiazole-5-carboxamide **(7).** To a solution of 2-chlorothiazole (7.96 g, 66.5 mmol) in THF (100 mL), n-butyllithium (1.6 M) in

hexanes (51 mL, 81.6 mmol) was added dropwise at -78 ℃. Subsequently, 1-chloro-2-isocyanato-3-methylbenzene (14.0 g, 83.8 mmol) in THF (20 mL) was added to the above mixture. After stirring for 2 h at -78 ℃, the reaction was quenched with saturated aqueous NH_4Cl (10 mL) followed by addition of water (300 mL) and EtOAc (500 mL). The organic layer was then separated and washed with brine (200 mL×2), concentrated and purified with flash chromatography (hexane: EtOAc=10:1/5:1) to give the desired product **7** (17.0 g, 89.0 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.08 (s, 1H), 7.65 (br, 1H), 7.26-7.29 (m, 1H), 7.18 (d, *J* = 3.93 Hz, 2H), 2.29 (s, 3H).; IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 285.98, Found: 286.98.

2-chloro-N-(2-chloro-6-methylphenyl)-N-(4-methoxybenzyl)thiazole-5-carboxamide **(8).** To a solution of **7** (153 mg, 0.53 mmol) in acetone, 4-methoxybenzyl chloride (124 mg, 0.8 mmol) and cesium carbonate (335 mg, 1.08 mmol) were added. The resulting mixture was refluxed. After 10 h, the reaction was concentrated and purified with flash chromatography (hexane: EtOAc = 5:1) to give desired product **8** (200mg, 92.5%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.32 (s, 2H), 7.28 (d, *J* = 7.8 Hz, 1H), 7.15 (d, *J* = 8.7 Hz, 2H), 7.10 (d, *J* = 7.8 Hz, 1H), 6.76 (d, *J* = 8.4 Hz, 2H), 5.42 (d, *J* = 13.8 Hz, 1H), 4.38 (d, *J* = 13.8 Hz, 1H), 3.77 (s, 3H), 1.57 (s, 3H). IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 407.04, Found: 407.03.

Buchwald-Hartwig coupling reaction for synthesis of compound **(9).** Sodium hydride was added to THF dropwise at 0 ℃ followed by addition of 4-amino-6-chloro-2-methylpyrimidine (11.2 g, 77.7 mmol). The mixture was then

stirred for 30 min at 0 ℃ followed by addition of compound **8** in THF (50 mL). The resulting mixture was refluxed for 4 h, and was cooled down to r.t. The reaction was then quenched with water (50 mL) at 0 ℃, and acidified with 1M HCl. The mixture was extracted with $CHCl_3$/MeOH (10:1) (200 mL×2). Subsequently, the combined organic layer was washed further with HCl (100 mL×2), dried over Na_2SO_4, concentrated to give the product **9** (11.4 g, 80.4%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.01 (s, 1H), 7.47 (d, *J* = 8.4 Hz, 1H), 7.42 (m, 2H), 7.30 (d, *J* = 8.4 Hz, 1H), 7.14 (d, *J* = 8.4 Hz, 2H), 6.81 (m, 3H), 5.21 (d, *J* = 14.1 Hz, 1H), 4.42 (d, *J* = 14.1 Hz, 1H), 3.46 (s, 3H), 2.44 (s, 3H), 1.72 (s, 3H). IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 514.09; Found: 514.06.

Deprotection of the 4-methoxybenzyl group to compound **(10).** To a solution of compound **9** in dichloromethane (80 mL), TFA (70 mL) and triflic acid (10 g, 66.6 mmol) were added at r.t. After stirring for 9 h, the reaction was poured into ice cautiously, and washed with $CHCl_3$ (200 mL×3). The combined organic layer was then concentrated and washed with the mixture of hexane and EtOAc (5:1, 200 mL). The solid was obtained by filtration, dried, to give the product **10** (8.2 g, 97.6 %). ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.22 (br, 1H), 10.0 (s, 1H), 8.30 (s, 1H), 7.39-7.42 (m, 1H), 7.25-7.28 (m, 2H), 6.94 (s, 1H), 2.58 (s, 3H), 2.23 (s, 3H). IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 394.03; Found: 394.02.

2-((6-((2-aminoethyl)amino)-2-methylpyrimidin-4-yl)amino)-N-(2-chloro-6-methylphenyl) thiazole-5-carboxamide **(11).** Compound **10** (500 mg, 1.26 mmol) was

suspended in the dioxane (2 mL) followed by addition of 1, 2-diaminoethane (8 mL). The resulting mixture was heated at 90 ℃ overnight. Subsequently, the solvent was evaporated under vacuum and potassium hydroxide aqueous solution (10 M) was added to *p*H 9. The mixture was finally filtrated and the resulting solid was washed with ether, dried, to give the yellow solid product **11** (400 mg, 76.2%.) ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.20 (s, 1H), 7.38-7.41 (m, 1H), 7.23-7.30 (m, 2H), 5.86 (s, 1H), 3.37 (s, 2H), 2.85 (s, 2H), 2.36 (s, 3H), 2.24 (s, 3H); ^{13}C NMR (300 MHz, DMSO-d_6) δ 18.69, 25.77, 41.12, 55.26, 83.71, 125.74, 127.34, 128.47, 129.36, 132.84, 133.94, 139.20, 141.28, 156.27, 160.38, 163.19, 163.60, 165.64. IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 418.12; Found: 418.11.

(9H-fluoren-9-yl)methyl(3-(4-benzoylphenyl)-1-((2-((6-((5-((2-chloro-6-methylpheny l)carbamoyl)thiazol-2-yl)amino)-2-methyl-pyrimi-din-4-yl)amino)ethyl)amino)-1-oxo propan-2-yl)carbamate **(12a).** To a solution of 2-((((9H-fluoren-9-yl)methoxy) carbonyl)amino)-3-(4-benzoylphenyl) propanoic acid **6a** (270 mg, 0.55 mmol), coupling reagent HOBt (74 mg, 0.55 mmol), HBTU (208 mg, 0.55 mmol) and DIEA (141 mg, 1.1 mmol) were added. Compound **11** (417 mg, 0.5 mmol) was then added to the above solution and the reaction was stirred at r.t. overnight. Subsequently, the reaction was quenched with water at 0 ℃ and extracted with the mixture of $CHCl_3$ and methanol (10:1, 100 mL). The organic layer was washed with water (20 mL), sodium bicarbonate solution (20 mL), ammonium chloride solution (20 mL), brine (20 mL). The organic layer was concentrated and precipitated with hexane and EtOAc (1:1), filtrated, to give the product **12a** (400 mg, 88.9%). ^{1}H NMR (300 MHz,

DMSO-d_6) δ 11.41 (s, 1H), 9.88 (s, 1H), 8.22 (s, 2H), 7.84 (d, *J* = 7.2 Hz, 2H), 7.58-7.67 (m, 9H), 7.21-7.44 (m, 13H), 5.89 (s, 1H), 4.23-4.28 (m, 2H), 4.10-4.12 (m, 2H), 3.11-3.32 (m, 1H), 2.72-3.08 (m, 2H), 2.35 (s, 3H), 2.22 (s, 3H). IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 891.28; Found: 891.26.

2-((6-((2-(2-amino-3-(4-benzoylphenyl)propanamido)ethyl)amino)-2-methylpyrimidin-4-yl)amino)-N-(2-chloro-6-methylphenyl)thiazole-5-carboxamide **(13a).** To a solution of compound **12a** (387 mg, 0.43 mmol) in DMF (10 mL), TBAF (1 M in THF, 1 mL) was added and the reaction was stirred for 0.5 h at r.t. The reaction was monitored by LC-MS. The product was used for next step without further purification. IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 669.22; Found: 669.19.

DA-1(1a). To a solution of hex-5-ynoic acid (50 mg, 0.44 mmol) in DMF (10 mL), coupling reagent HBTU (179 mg, 0.47 mmol), HOBt (67 mg, 0.5 mmol) and DIEA (110 mg, 0.92 mmol) were added. The resulting mixture was stirred for 0.5 h at r.t. and added to the previous step crude product. The resulting reaction was stirred overnight at r.t. followed by quenching with water (20 mL). The reaction was then extracted with $CHCl_3$: Methanol (10:1, 20 mL×3) and the combined organic layer was concentrated, purified with flash chromatography (DCM:MeOH = 15:1) to give the product **DA-1** (60 mg, 18.3 % for the two steps). ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.37 (s, 1H), 9.82 (s, 1H), 8.12-8.16 (m, 3H), 7.61-7.68 (m, 5H), 7.40-7.52 (m, 2H), 7.38 (d, *J* = 8.1 Hz, 2H), 7.26-7.36 (m, 2H), 7.15 (brs, 1H), 5.85 (s, 1H), 3.15 (m, 1H), 2.70-2.84 (m, 1H), 2.71 (s, 1H), 2.34 (s, 3H), 2.22 (s, 3H), 2.13 (s, 2H), 1.96-1.98 (m,

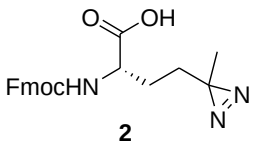
2H), 1.52-1.57 (m, 2H), 1.23 (s, 1H), 0.85 (m, 1H). ^{13}C NMR (300 MHz, DMSO-d_6) δ 195.4, 171.4, 171.1, 165.3, 163.0, 162.7, 159.9, 143.4, 140.8, 138.8, 137.2, 135.0, 133.5, 132.4, 129.5, 129.3, 129.0, 128.5, 127.0, 84.0, 71.3, 53.6, 37.8, 34.0, 25.4, 24.2, 22.0, 18.3, 17.2, 13.9. IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 763.26; Found: 763.24.

Tert-butyl(1-((2-((6-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)-2-methyl-pyrimidin-4-yl)amino)ethyl)amino)-4-(3-methyl-3H-diazirin-3-yl)-1-oxobutan-2-yl)carbamate **(12b).** To a mixture of 2-((tert-butoxycarbonyl)amino)-4-(3-methyl-3H-diazirin-3-yl) butanoic acid **6b** (180 mg, 0.7 mmol) and compound **11** (200 mg, 0.48 mmol) in DMF (15 mL), coupling reagent HOBt (105 mg, 0.78 mmol), HBTU (300 mg, 0.79 mmol) and DIEA (175 mg, 1.35 mmol) were added. The reaction was stirred for 24 h at r.t. in the dark and was then quenched by water (50 mL). Subsequently, the mixture was filtrated and the resulting solid was washed with water and ether, precipitated in the $CHCl_3$ and methanol (10:1) again to give the product **12b** (190 mg, 57.6%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.39 (s, 1H), 9.83 (s, 1H), 8.20 (s, 1H), 7.90 (s, 1H), 7.40 (d, J = 2.1 Hz, 1H), 7.22-7.38 (m, 2H), 7.12 (s, 1H), 6.84 (d, J = 7.5 Hz, 1H), 5.86 (s, 1H), 3.80 (s, 1H), 3.25 (m, 4H), 2.38 (s, 3H), 2.24 (s, 3H), 1.22-1.48 (m, 13 H), 0.95 (s, 3H). IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 657.25; Found: 657.30.

DA-2 (1b). To a solution of compound **12b** (188 mg, 0.28 mmol) in DCM (10 mL), TFA (10 mL) was added slowly at 0 ℃. After stirring overnight at r.t., the solvent was removed under vacuum and the residue was suspended to DMF (8 mL) followed by

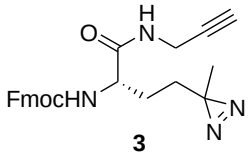
the addition of hex-5-ynoic acid (56 mg, 0.5 mmol), DIEA (240 mg, 1.86 mmol), HOBt (60 mg, 0.44 mmol) and HBTU (162 mg, 0.42 mmol). After stirring overnight at r.t. in the dark, the reaction was then quenched by water (30 mL) and filtrated. The solid was washed with water and ether, precipitated with $CHCl_3$ and methanol (10:1) to give the desired product **1b** (91 mg, 50 % for two steps). ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.40 (s, 1H), 9.84 (s, 1H), 8.20 (s, 1H), 7.92-7.997 (m, 2H), 7.39 (d, *J* = 6.9 Hz, 1H), 7.23-7.28 (m, 2H), 7.13 (br, 1H), 5.87 (s, 1H), 4.18-4.13 (m, 1H), 3.24-3.31 (m, 4H), 2.76 (s, 1H), 2.38 (s, 3H), 2.12-2.23 (m, 7H), 2.04 (s, 1H), 1.24-1.67 (m, 7H), 0.95 (s, 3H); ^{13}C NMR (300 MHz, DMSO-d6) δ 172.0, 171.8, 165.63, 163.1, 160.3, 141.1, 139.1, 133.9, 132.8, 129.3, 128.5, 127.3, 125.8, 84.4, 71.7, 52.3, 39.0, 30.6, 26.8, 26.0, 25.7, 24.5, 19.6, 18.6, 17.7. IT-TOF-MS: *m/z* $[M+1]^+$ calcd: 651.24; Found: 651.25.

8.2.7 Chemical Synthesis of Staurosphorine-derived Probe



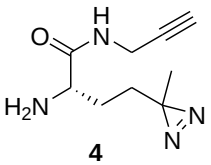
2

To a solution of diazirine free amino acid (100 mg, 0.7 mmol) and 10% Na_2CO_3 (5.2 mL) in dioxane (3.9 mL), Fmoc-Cl (258 mg, 1mmol) was added at r.t. After 20 h stirring, *p*H value of the reaction was adjusted to 2~3 using 1N HCl, and the mixture were extracted with ethyl acetate (5 mL × 3). The combined organic layer was dried with Na_2SO_4, concentrated, purified with flash chromatography to give the product **2** in 85% yield.



3

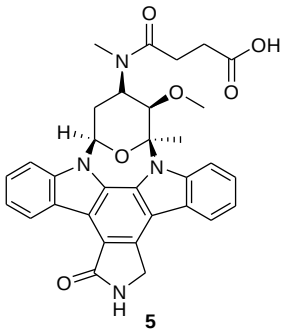
To a solution of of propargylamine (49 μL) in DMF (5 mL), coupling reagent HOBt (104 mg, 0.78 mmol), HBTU (288 mg, 0.78 mmol) and DIEA (200 mg, 1.56 mmol) were added. Compound **2** (190 mg, 0.5 mmol) was then added to the above solution and the reaction was stirred at r.t. overnight. Subsequently, the reaction was quenched with water and extracted with EA. The organic layer was washed with water (20 mL), sodium bicarbonate solution (20 mL), brine (20 mL). The organic layer was concentrated and purified by column chromatography to afford compound **3** in 93% yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.4 Hz, 2H), 7.56 (d, *J* = 7.4 Hz, 2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.34-7.29 (m, 2H), 6.41 (s, 1H), 5.37-5.34 (m, 1H), 4.44-4.42 (m, 2H), 4.22-4.17 (m, 1H), 4.15-4.11 (m, 1H), 4.00 (s, 2H), 2.17 (s, 1H), 1.49-1.42 (m, 2H), 1.32-1.21 (m, 2H), 1.01 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.62, 156.30, 143.54, 141.25, 127.74, 127.05, 124.89, 119.97, 78.80, 71.88, 67.00, 53.88, 47.05, 30.26, 29.17, 27.11, 19.54.



4

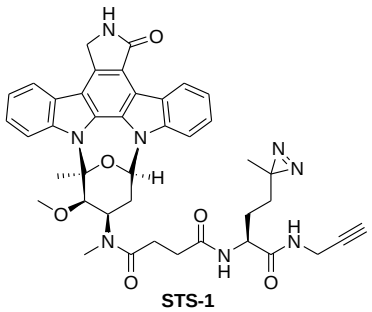
Compound **3** was dissolved into 20% piperidine in DMF, and stirred for 10 min at r.t. The reaction was concentrated and purified by column chromatography to give compound **4** in 98% yield. ^{1}H NMR (300 MHz, DMSO) δ 8.22 (s, 1H), 3.84 (s, 2H), 3.09-3.07 (m, 2H), 1.52-1.34 (m, 2H), 1.32-1.06 (m, 2H), 0.97 (s, 3H); ^{13}C NMR (75

MHz, DMSO) δ 174.9, 81.6, 73.2, 54.2, 40.7, 40.4, 40.1, 39.9, 39.6, 39.3, 39.0, 30.5, 29.8, 28.1, 26.2, 19.8.



5

To a solution of staurosphorine (**STS**) (5.0 mg, 10 μmol) in DMSO, succinic anhydride (1.5 mg, 15 μmol) and DMAP (61 μg, 20 μmol) were added under dark. After 30 h stirring, the mixture was precipitated with 0.1 % TFA in water, and the precipitate was triturated twice with 0.1 % TFA in water to afford compound **5** in 92% yield. ^{1}H NMR (300 MHz, DMSO-d6) δ 12.06 (s, 1H), 9.29 (d, *J* = 7.9 Hz, 1H), 8.59 (s, 1H), 8.05 (d, *J* = 7.7 Hz, 1H), 7.99 (d, *J* = 8.5 Hz, 1H), 7.67 (d, *J* = 8.2 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 2H), 7.35 (t, *J* = 7.5 Hz, 1H), 7.29 (t, *J* = 7.5 Hz, 1H), 7.03 (t, *J* = 7.5 Hz, 1H), 5.00 (s, 3H), 4.22 (s, 1H), 2.81 (s, 3H), 2.77 (s, 3H), 2.68 (d, *J* = 5.6 Hz, 1H), 2.60-2.56 (m, 2H), 2.33 (s, 3H), 2.28-2.17 (m, 1H). LCMS m/z calcd for $C_{32}H_{30}N_4O_6$ $[M+H]^+$ = 567.220, found 567.199 (IT-TOF).



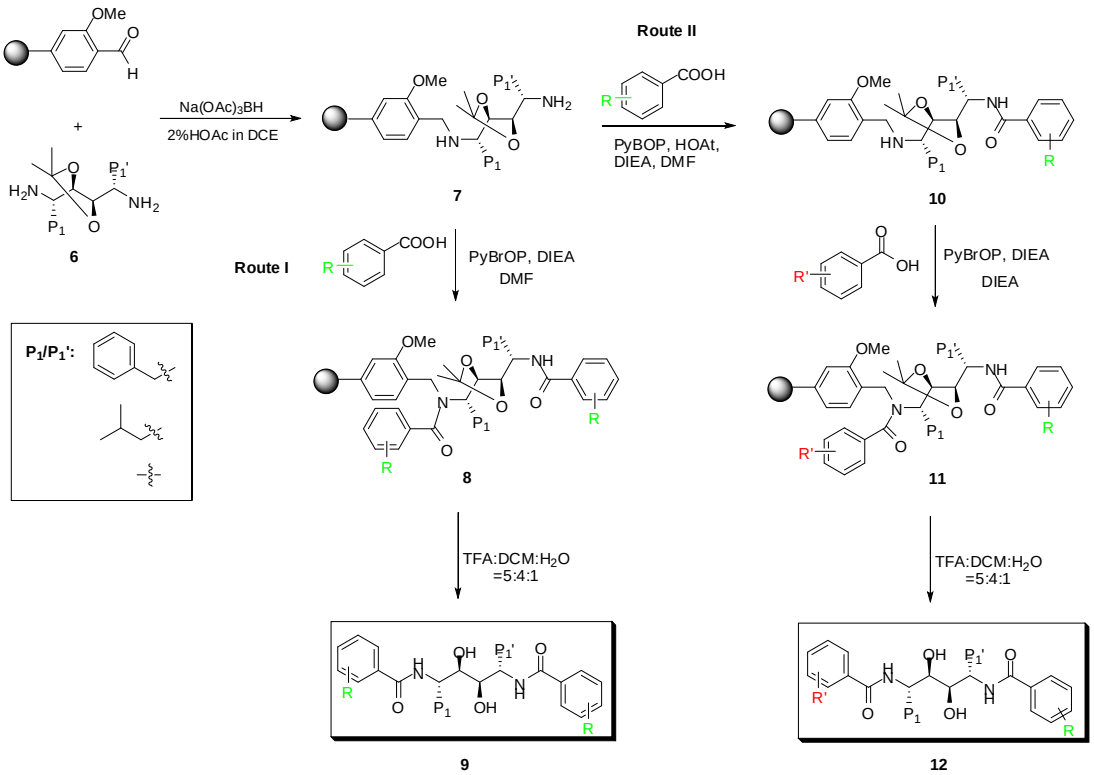
STS-1

To a solution of compound **5** in DMSO (1 mL), coupling reagent DMAP (2.4 mg) and EDC (15.4 mg) were added, and stirred for 10 min at r.t. Compound **4** was

subsequently added into the reaction at 0 °C. The mixture was then allowed to warm to r.t. After overnight stirring, the mixture was precipitated with 0.1 % TFA in water to afford compound **STS-1** in 92% yield. ^{1}H NMR (300 MHz, DMSO) δ 9.29 (d, *J* = 7.7 Hz, 1H), 8.61 (s, 1H), 8.33-8.31 (m, 1H), 8.16-8.11 (m, 1H), 8.08-8.03 (m, 1H), 7.99-7.94 (d, *J* = 8.2 Hz, 1H), 7.66 (t, *J* = 7.1 Hz, 1H), 7.49 (t, *J* = 7.1 Hz, 2H), 7.33 (dt, *J* = 18.4 Hz, 2H), 7.04-6.97 (m, 1H), 5.00 (s, 3H), 4.21-4.19 (m, 2H), 3.92-3.86 (m, 2H), 3.19-3.15 (m, 1H), 2.82-2.74 (m, 6H), 2.69-2.57 (m, 3H), 2.45-2.37 (m, 2H), 2.36-2.33 (m, 3H), 2.26-2.17 (m, 1H), 1.45-1.38 (m, 3H), 1.22-1.20 (s, 1H), 1.02 (s, 3H); LCMS m/z calcd for $C_{41}H_{43}N_8O_6$ $[M+H]^+$ = 743.330, found 743.322; $C_{41}H_{42}N_8O_6Na$ $[M+Na]^+$ = 765.3120, found 765.303 (IT-TOF).

8.3 Solid-phase Synthesis

8.3.1 Procedure for Synthesis of the Diol Library



Scheme 8.6 Traceless synthesis of symmetric and asymmetric diol inhibitors.

The construction of the inhibitor library was achieved by IRORI® split-and-pool directed sorting technology. Final products were released from the solid support by our optimized TFA cleavage protocol (below).

*General procedure for reductive amination with (**6**):*

PL-FMP resin (75× ~40 mg, 0.9 mmol/g) was taken into 75 MacroKan^{TM} reactors, with each containing a unique RF tag. The resin was swelled in 1, 2-dichloroethane (150 mL) for 2 h. Upon decanting of the solvent, the microreactors were then distributed into three different 50-mL tubes each containing 2 % acetic acid in 1,2-dichloroethane (20 mL). Each of the three diols, **6a-c** (5 eq, below) was added to each of the three tubes, respectively, followed by incubation for 3 h at r.t. with constant shaking. Subsequently sodium triacetoxyborohydride (6 eq) was added, and the tubes were shaken for another 8 h before the solutions were decanted. The reactors were combined, washed with DCM (3×50 mL), MeOH (3×50 mL) and THF (3×50 mL) and dried *in vacuo* to afford resin **7**. The success of the reaction was monitored by Nynhydrin and Chloranil test. Blue beads indicate the presence of primary or secondary amine and the attachment of the diol to the resin.

*General procedure for coupling/synthesis of symmetrical diols (**8**) (Route I):*

Selected microreactors containing resin **7** were swelled in DMF for 2 h. The solvent was decanted. A solution containing the acid building block (10.0 eq), PyBrOP (10.0 eq) and DIPEA (20.0 eq) in DMF was prepared under N_2 and mixed for 2 min before being added to the resin. The reaction mixture was shaken for 12 h at r.t., after which the solution was decanted and the resin was washed with DMF (3×20

mL), MeOH (3×20 mL) and (3×20 mL), following by drying *in vacuo* to afford resin **8**.

*General procedure for coupling/synthesis of asymmetrical diols (**11**) (Route **II**):*

Selected microreactors containing resin **7** were swelled in DMF for 2 h. The solvent was decanted. A solution containing the first acid building block (5.0 eq), PyBOP (5.0 eq), HOAt (5.0 eq) and DIPEA (10.0 eq) in DMF was prepared under N_2 and mixed for 2 min before being added to the resin. The reaction mixture was shaken for 12 h at r.t., after which the solution was decanted and the resin was washed with DMF (3×20 mL), MeOH (3×20 mL) and (3×20 mL) followed by drying *in vacuo* to afford the mono-acylated resin **10**. Subsequently, the resin was swelled in DMF for 2 h. The solvent was decanted. A solution containing the desired second acid building block (5.0 eq), PyBrOP (5.0 eq) and DIPEA (10.0 eq) in DMF was prepared under N_2 and mixed for 2 min before being added to the resin. The reaction mixture was shaken for 12 h at r.t., after which the solution was decanted and the resin was washed with DMF (3×20 mL), MeOH (3×20 mL) and (3×20 mL) followed by drying *in vacuo* to afford the asymmetrical diols **11**.

*Cleavage of resins (8) and (11) to give the final products (**9**) and (**12**):*

After sorting/decoding, each microreactor was treated with 2.0 mL of a freshly prepared cleavage cocktail of TFA/DCM/H_2O (5:4:1) for 3.5 h at r.t. The solutions were then transferred to an 96-deep well, polypropylene microplate, and concentrated with a GeneVac™ HT-4X concentrator to afford the final desired products in sufficiently pure form (w/o any purification). LC-MS, NMR characterizations and

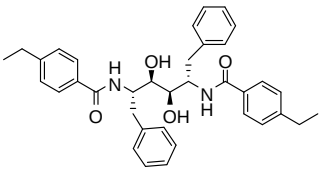
biological screening were carried out directly with these products. These compounds were typically re-dissolved in 0.5 mL of DMSO and kept in -20 ^{o}C as master stocks.

Library Characterizations:

For LCMS analysis, 40 μL of the samples from the master stocks were transferred to a 96-well polypropylene microplate, and LCMS was performed by injecting 3 μL using an autosampler. Most products analyzed were shown to be of sufficient purity (90-95% for symmetrical diols and 60-95% for asymmetrical diols; below). Randomly selected compounds were further characterized by ^{1}H and/or ^{13}C NMR.

^{1}H and/or ^{13}C-NMR data of representative crude products

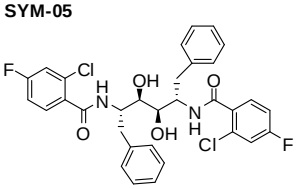
SYM-04



Exact Mass: 564.30

^{1}H NMR (300 MHz, $CDCl_3$) δ 1.15-1.23 (m, 6H), 2.07 (d, *J* = 4.71 Hz, 2H), 2.61-2.64 (m, 4H), 2.72-2.74 (m, 2H), 2.82-2.88 (m, 2H), 3.52-3.72 (m, 4H), 7.12-7.25 (m, 14H), 7.64 (d, *J* = 7.65 Hz, 2H), 7.77-7.94 (m, 2H). LCMS found for $C_{36}H_{40}N_2O_4$ $(M + H)^+$: m/z = 565.25.

SYM-05

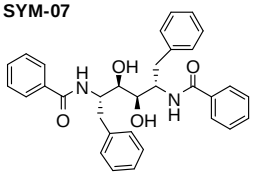


Exact Mass: 612.14

^{1}H NMR (500 MHz, DMSO-d_6) δ 2.77-2.81 (m, 2H), 2.88-2.93 (m, 2H), 4.64-4.69 (m, 2H), 4.76 (s, 2H), 7.03-7.06 (t, *J* = 7.25 Hz, 1H), 7.11-7.15 (t, *J* = 7.57 Hz, 3H), 7.26

(d, *J* = 6.95 Hz, 3H), 7.31-7.34 (m, 2H), 7.43-7.49 (m, 4H), 8.02 (d, *J* = 8.85 Hz, 1H). ^{13}C-NMR (75 MHz, DMSO-d_6): 164.9, 163.4, 160.2, 139.3, 137.0, 130.2, 129.2, 127.9, 125.7, 123.5, 117.9, 114.6, 72.2, 52.3, 37.4. LCMS found for $C_{32}H_{28}C_{l2}F_2N_2O_4$ $(M + Na)^+$: *m/z* = 613.12.

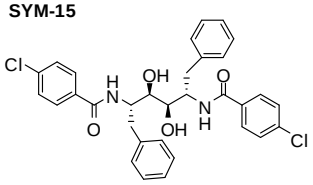
SYM-07



Exact Mass: 508.24

^{1}H NMR (300 MHz, $CDCl_3$) δ 2.72-2.81 (m, 2H), 2.88-2.90 (m, 2H), 4.05-4.26 (m, 4H), 4.69 (s, 2H), 7.05-7.28 (m, 11H), 7.40-7.72 (m, 6H), 7.80 (d, *J* = 6.96 Hz, 14H), 7.87-7.95 (m, 1H). LCMS found for $C_{32}H_{32}N_2O_4$ $(M + Na)^+$: m/z = 531.20.

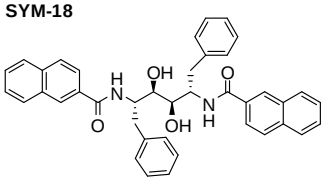
SYM-15



Exact Mass: 576.16

^{1}H NMR (300 MHz, $CDCl_3$) δ 2.72-2.80 (m, 2H), 2.80-2.96 (m, 2H), 4.08-4.29 (m, 4H), 4.67 (s, 2H), 7.03-7.27 (m, 10H), 7.46 (d, *J* = 8.70 Hz, 4H), 7.73 (d, *J* = 8.37 Hz, 3H), 8.01 (d, *J* = 9.06 Hz, 1H). LCMS found for $C_{32}H_{30}Cl_2N_2O_4$ $(M + Na)^+$: m/z = 599.12.

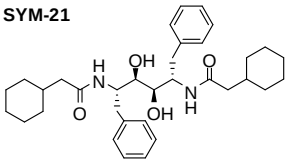
SYM-18



Exact Mass: 608.27

^{1}H NMR (300 MHz, $CDCl_3$) δ 2.73-2.60 (m, 2H), 2.82-2.87 (m, 2H), 4.09-4.39 (m, 4H), 4.80 (s, 2H), 6.98-7.43 (m, 11H), 7.58 (t, *J* = 3.81 Hz, 4H), 7.81-7.96 (m, 8H), 8.31 (s, 1H). LCMS found for $C_{40}H_{36}N_2O_4$ $(M + Na)^+$: m/z = 631.24.

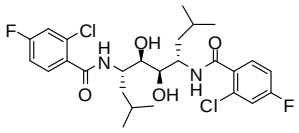
SYM-21



Exact Mass: 548.36

^{1}H NMR (500 MHz, $CDCl_3$) δ 0.67-0.86 (m, 3H), 1.04-1.28 (m, 5H), 1.20-1.28 (m, 3H), 1.46-1.65 (m, 8H), 1.81-1.86 (m, 3H), 1.90-2.07 (m, 1H), 2.74 (d, *J* = 8.81 Hz, 2H), 2.85 (d, *J* = 6.90 Hz, 2H), 3.97 (d, *J* = 8.85 Hz, 2H), 4.62 (s, 2H), 4.82-4.91 (m, 2H), 7.10-7.38 (m, 10H), 7.52 (d, *J* = 6.81 Hz, 1H), 7.77 (d, *J* = 6.35, Hz, 18H). LCMS found for $C_{34}H_{48}N_2O_4$ $(M + Na)^+$: m/z = 571.305.

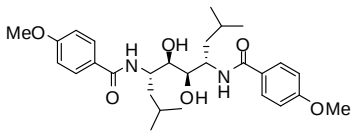
SYM-27



Exact Mass: 544.17

^{1}H NMR (500 MHz, DMSO-d_6) δ 0.84-0.86 (m, 12H), 1.23-1.25 (m, 4H), 1.55-1.64 (m, 2H), 3.25-3.66 (m, 4H), 4.40 (s, 2H), 7.32-7.39 (m, 1H), 7.46-7.51 (m, 2H), 7.63 (d, *J* = 9.45 Hz, 1H), 7.68 (d, *J* = 7.55 Hz, 1H), 7.90 (d, *J* = 7.55 Hz, 1H). ^{13}C-NMR (75 MHz, DMSO-d_6): 165.7, 163.6, 130.3, 123.6, 117.8, 114.1, 72.7, 49.3, 24.5, 23.4, 21.9. LCMS found for $C_{32}H_{28}Cl_2F_2N_2O_4$ $(M + H)^+$: *m/z* = 545.29.

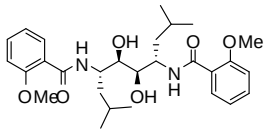
SYM-31



Exact Mass: 500.29

^{1}H NMR (500 MHz, DMSO-d_6) δ 0.83-0.85 (m, 12H), 1.16-1.19 (m, 4H), 1.64 (d, *J* = 7.65Hz, 4H), 1.93-1.99 (m, 2H), 2.04-2.05 (m, 2H), 2.85 (d, *J* = 6.95 Hz, 1H), 3.12 (s, 6H), 4.11-4.16 (m, 2H), 7.22 (d, *J* = 9.55 Hz, 1H), 7.39-7.55 (m, 1H), 7.50-7.55 (m, 2H), 7.68-7.79 (m, 2H), 8.05 (d, *J* = 7.59 Hz, 1H), 8.15 (d, *J* = 8.55 Hz, 1H). LCMS found for $C_{28}H_{40}N_2O_6$ $(M +H)^+$: *m/z* = 501.00.

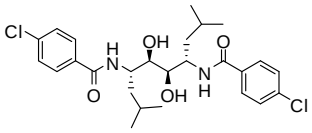
SYM-35



Exact Mass: 500.29

^{1}H NMR (500 MHz, DMSO-d_6) δ 0.81-0.83 (m, 12H), 1.17-1.21 (m, 4H), 1.62 (d, *J* = 7.60Hz, 4H), 1.94-1.99 (m, 2H), 2.06-2.08 (m, 2H), 2.85 (d, *J* = 6.95 Hz, 1H), 3.12 (s, 6H), 4.11-4.15 (m, 2H), 7.21 (d, *J* = 9.45 Hz, 1H), 7.39-7.50 (m, 1H), 7.50-7.56 (m, 2H), 7.66-7.77 (m, 2H), 7.95 (d, *J* = 7.55 Hz, 1H), 8.14 (d, *J* = 8.50 Hz, 1H). LCMS found for $C_{28}H_{40}N_2O_6$ $(M+H)^+$: *m/z* = 501.00.

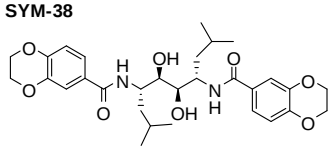
SYM-37



Exact Mass: 508.19

^{1}H NMR (300 MHz, $CDCl_3$) δ 0.80-0.90 (m, 12H), 1.40-1.87 (m, 6H), 4.00-4.28 (m, 4H), 7.42-7.62 (m, 3H), 7.77-7.94 (m, 3H), 8.05-8.30 (m, 2H). LCMS found for $C_{26}H_{34}Cl_2N_2O_4$ $(M + Na)^+$: m/z = 531.34.

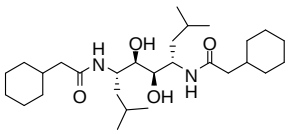
SYM-38



Exact Mass: 556.28

^{1}H NMR (300 MHz, $CDCl_3$) δ 0.87-0.92 (m, 12H), 1.52-1.78 (m, 6H), 2.72-3.16 (m, 4H), 4.45-4.52 (m, 8H), 7.57-7.60 (m, 2H), 7.94-7.98 (m, 4H). LCMS found for $C_{26}H_{34}Cl_2N_2O_4$ $(M + H)^+$: m/z = 557.26.

SYM-43

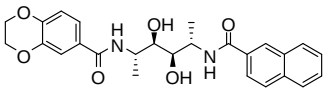


Exact Mass: 480.39

^{1}H NMR (500 MHz, $CDCl_3$) δ 0.78-0.87 (m, 12H), 1.08-1.24 (m, 10H), 1.40-1.52 (m,

4H), 1.50-1.52 (m, 14H), 1.89-2.08 (m, 4H), 3.04-3.14 (m, 2H), 4.12-4.27 (m, 2H), 1.40-1.52 (m, 4H), 4.93 (d, *J* = 12.6 Hz, 2H), 5.07 (d, *J* = 12.6 Hz, 2H), 7.21 (d, *J* = 9.45 Hz, 24H). LCMS found for Chemical Formula: $C_{28}H_{52}N_2O_4$ $(M + H)^+$: m/z = 481.35.

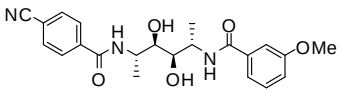
ASM-03



Exact Mass: 464.19

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.21-1.23 (m, 6H), 1.31-1.34 (m, 4H), 3.55 (d, *J* = 4.40 Hz, 1H), 4.17-4.52 (m, 3H), 7.57-7.61 (m, 3H), 7.95-8.01 (m, 6H), 8.46 (s, 1H) LCMS found for $C_{26}H_{28}N_2O_6$ $(M + Na)^+$: m/z = 487.15.

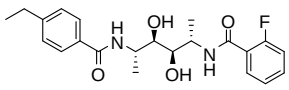
ASM-04



Exact Mass: 411.18

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.14-1.19 (m, 6H), 3.75 (s, 3H), 4.26-4.80 (m, 4H), 5.15-5.30 (m, 2H), 7.77 (d, *J* = 8.20 Hz, 1H), 7.87-8.10 (m, 7H). LCMS found for $C_{22}H_{25}N_3O_5$ $(M + Na)^+$: m/z = 434.15.

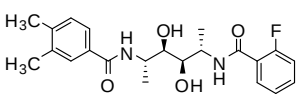
ASM-05



Exact Mass: 402.20

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.13-1.26 (m, 9H), 2.62-2.67 (m, 2H), 4.00-4.40 (m, 4H), 7.28 (d, *J* = 7.60 Hz, 4H), 7.75-7.81 (m,4H). LCMS found for $C_{22}H_{27}FN_2O_4$ $(M + Na)^+$: m/z = 425.15.

ASM-06

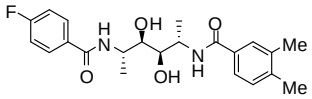


Exact Mass: 402.20

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.13 (d, *J* = 6.30 Hz, 6H), 2.25 (s, 6H), 4.09-4.38

(m, 4H), 7.19 (d, *J* = 8.20 Hz, 2H), 7.51-7.67 (m, 2H), 7.76 (d, *J* = 8.80 Hz, 2H). LCMS found for $C_{22}H_{27}FN_2O_4$ $(M + Na)^+$: m/z = 425.15.

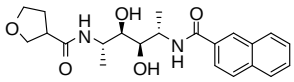
ASM-07



Exact Mass: 402.20

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.13 (d, *J* = 6.95 Hz, 6H), 2.35 (s, 6H), 3.77-4.25 (m, 4H), 7.17-7.29 (m, 3H), 7.89-7.98 (m,3H). LCMS found for $C_{22}H_{27}FN_2O_4$ $(M + H)^+$: m/z = 403.15.

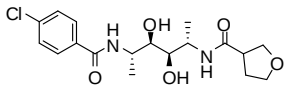
ASM-09



Exact Mass: 400.20

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.13 (d, *J* = 6.95 Hz, 6H), 1.39-1.48 (m, 2H), 2.16-2.19 (m, 1H), 3.77-4.26 (m, 4H), 7.17-7.28 (m, 3H), 7.89-7.95 (m, 4H). LCMS found for $C_{22}H_{28}N_2O_5$ $(M + H)^+$: m/z = 423.10.

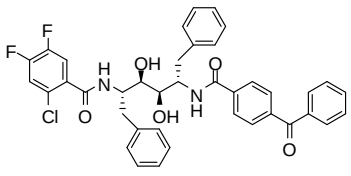
ASM-10



Exact Mass: 384.15

^{1}H NMR (500 MHz, DMSO-d_6) δ 1.14 (d, *J* = 6.95 Hz, 6H), 1.24-1.25 (m, 2H), 2.16-2.09 (m, 1H), 3.99-4.50 (m, 4H), 7. 85 (d, *J* = 8.80 Hz, 4H). LCMS found for $C_{18}H_{25}C_lN_2O_5$ $(M + Na)^+$: m/z = 407.05.

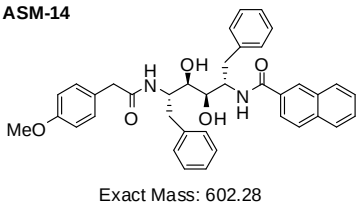
ASM-12



Exact Mass: 682.20

^{1}H NMR (300 MHz, DMSO-d_6) δ 2.72-2.89 (m, 4H), 3.76-3.88 (m, 4H), 6.59-6.63 (m, 4H), 6.94 (d, *J* = 5.94 Hz, 2H), 7.12-7.43 (m, 10H), 7.61-8.04 (m, 5H). LCMS found

for $C_{39}H_{33}C_lF_2N_2O_5$ $(M + H)^+$: m/z = 683.25.



^{1}H NMR (300 MHz, $CDCl_3$) δ 2.13-2.41 (m, 4H), 2.72-2.88 (m, 4H), 3.31 (s, 1H), 3.73-3.88 (m, 6H), 6.92-7.36 (m, 15H), 7.43-7.59 (m, 3H), 7.67-7.83 (m, 2H), 7.90-7.29 (m, 2H). LCMS found for $C_{39}H_{33}C_lF_2N_2O_5$ $(M + Na)^+$: m/z = 625.29.

8.3.2 Synthesis of Eight Hydroxyethylamine Azide Warheads

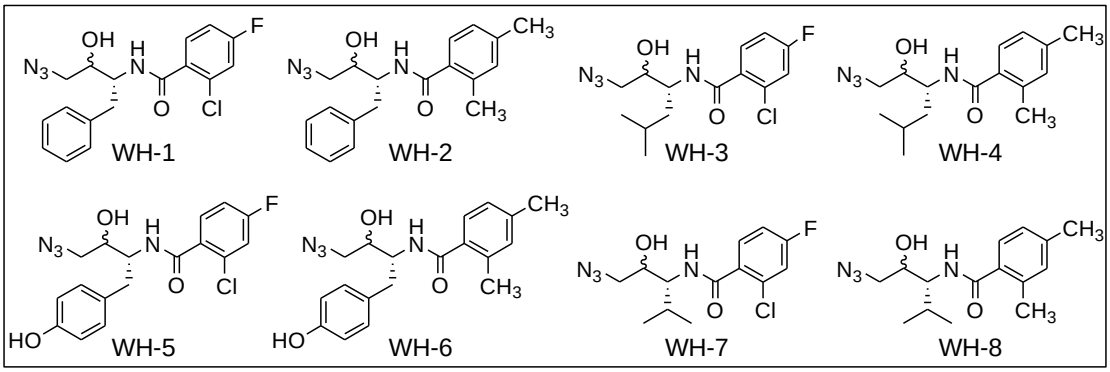


Figure 8.1 Structrues of the N-terminal azide warheads **(1-8).**

Warheads **1-8** were synthesized according to published procedures with some modifications.[100]

WH-1. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 2.83-3.09 (m, 2H), 3.29-3.35 (m, 2H), 3.76-3.84 (m, 1H), 4.31-4.35 (m, 2H), 6.62 (d, *J* = 8.22 Hz, 1H), 6.86-7.15 (m, 6H), 7.21-7.50 (m, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 166.4, 164.7, 161.4, 137.2, 131.9, 131.7, 131.3, 131.1, 130.9, 129.2, 129.1, 128.6, 126.7, 117.5, 114.2, 72.4, 70.0, 54.9, 54.5, 53.9, 38.6, 38.1, 34.6, 29.7; LC-MS: *m/z* $[M+H]^+$ = 363.093.

WH-2. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 2.14 (s, 3H), 2.27 (s, 3H), 2.74-2.98 (m, 2H),

3.23-3.34 (m, 2H), 3.77 (s, 1H), 4.23-4.38 (m, 2H), 6.28 (d, *J* = 10.0 Hz, 1H), 6.82-6.93 (m, 6H), 7.09 (d, *J* = 3.15 Hz, 2H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 171.0, 153.9, 139.9, 135.7, 127.1, 132.7, 132.4, 132.2, 131.5, 129.5, 129.4, 126.5, 126.1, 126.0, 124.2, 78.3, 72.8, 54.8, 53.8, 34.6, 28.6, 21.0, 19.4; LC-MS: *m/z* $[M+H]^+$ = 339.173.

WH-3. ^{1}H-NMR (300 MHz, DMSO-d_6) δ 0.94 (t, *J* = 4.20 Hz , 6H), 1.24-1.52 (m, 2H), 6.52 (d, *J* = 7.65 Hz, 1H), 6.98-7.10 (m, 1H), 7.12 (d, *J* = 2.07 Hz, 1H), 7.54-7.63 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 166.5, 166.0, 164.8, 161.5, 131.6, 117.6, 114.5, 73.3, 72.3, 54.9, 53.7, 51.8, 50.1, 41.1, 38.6, 24.8, 23.5, 23.0, 22.0, 21.4; LC-MS: *m/z* $[M+H]^+$ = 329.109.

WH-4. ^{1}H-NMR (300 MHz, DMSO-d_6) δ 0.88 (t, *J* = 3.90 Hz, 6H), 1.22-1.52 (m, 4H), 1.72-1.90 (m, 1H), 2.69 (s, 3H), 2.95 (s, 3H), 3.21-3.33 (m, 1H), 3.47-3.53 (m, 1H), 4.40 (s, 1H), 7.23-7.29 (m, 1H), 7.30-7.50 (m, 2H), 7.94 (s, 1H); ^{13}C-NMR (75 MHz, DMSO-d_6) δ 165.4, 163.6, 162.3, 160.3, 158.6, 157.6, 133.9, 130.7, 116.8, 114.3, 113.3, 72.8, 53.7, 50.0, 36.9, 30.8, 24.3, 23.6, 21.6; LC-MS: *m/z* $[M+H]^+$ = 305.190.

WH-5. ^{1}H-NMR (300 MHz, MeOD) δ 3.17-3.38 (m, 1H), 3.76-3.97 (m, 1H), 3.76-4.86 (m, 1H), 3.97-4.48 (m, 2H), 4.42-4.48 (m, 1H), 4.86-4.94 (m, 1H), 7.40 (d, *J* = 8.34Hz, 1H), 7.62-7.99 (m, 6H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 168.7, 165.6, 156.6, 133.8, 133.0, 131.2, 131.1, 130.3, 118.1, 117.8, 116.1, 115.8, 115.3, 114.8, 73.8, 71.7, 55.3, 36.2; LC-MS: *m/z* $[M+H]^+$ = 379.088.

WH-6. ^{1}H-NMR (300 MHz, MeOD) δ 2.76 (s, 3H), 2.95 (s, 3H), 3.22-3.64 (m, 2H),

4.00-4.13 (m, 2H), 4.45-4.51 (m, 1H), 4.93-4.99 (m, 1H), 7.39-7.41 (m, 2H), 7.62-7.66 (m, 3H), 7.73-7.80 (m, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 172.9, 156.7, 140.9, 140.7, 136.5, 134.8, 132.2, 132.1, 131.3, 130.4, 130.2, 127.8, 127.0, 126.9, 116.1, 74.1, 55.8, 54.8, 36.7, 21.2, 19.5; LC-MS: *m/z* $[M+H]^+$ = 355.167.

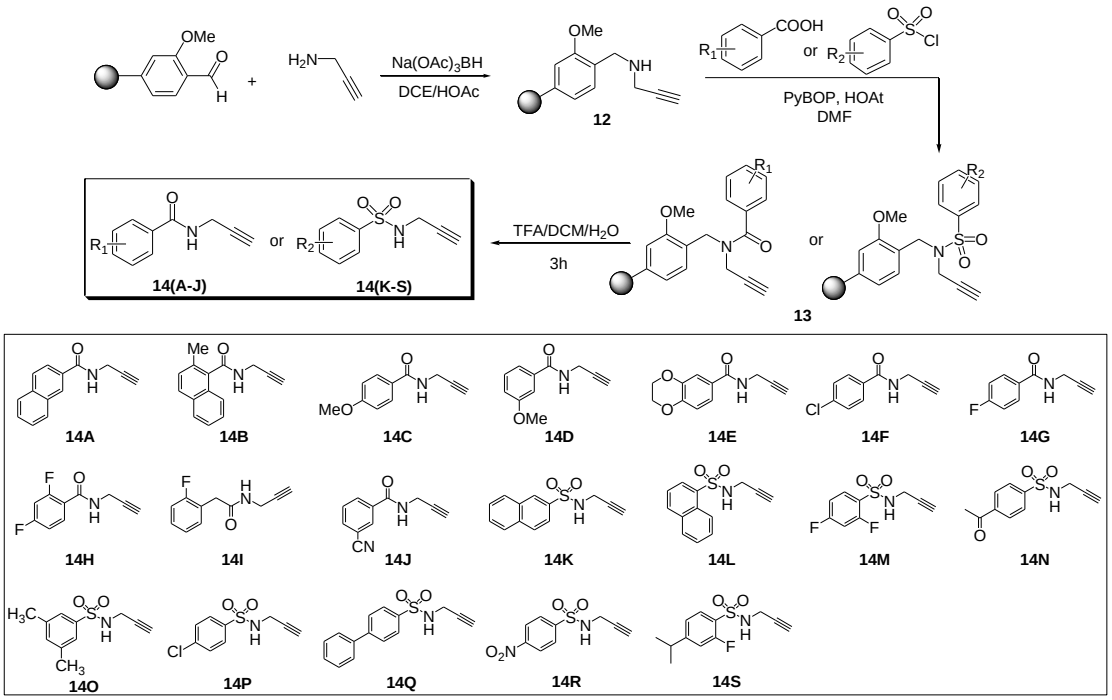
WH-7. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.96-0.99 (m, 6H), 2.07-2.18 (m, 1H), 3.37-3.46 (m, 2H), 3.83-3.89 (m, 1H), 4.08-4.14 (m, 1H), 6.62 (d, *J* = 9.42 Hz, 1H), 6.98-7.04 (m, 1H), 7.09-7.15 (m, 1H), 7.59 (d, *J* = 6.24 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 166.6, 164.8, 161.4, 131.1-131.7, 129.0, 116.8-117.6, 114.3-114.8, 71.2, 57.6, 54.3, 28.0, 20.2, 17.3; LC-MS: *m/z* $[M+H]^+$ = 315.092.

WH-8. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.97 (d, *J* = 4.95 Hz, 6H), 3.76-3.97 (m, 1H), 1.89-2.19 (m, 1H), 2.32 (d, *J* = 2.31 Hz, 3H), 2.38 (d, *J* = 1.98 Hz, 3H), 3.30 (d, *J* = 6.42 Hz, 1H), 3.38 (d, *J* = 3.30 Hz, 1H), 3.69 (t, *J* = 9.39 Hz, 1H), 3.84 (t, *J* = 6.06 Hz, 1H), 3.86-4.07 (m, 1H), 6.08 (d, *J* = 9.21 Hz, 1H), 6.96-7.02 (m, 2H), 7.22 (d, *J* = 7.56 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 171.4, 170.9, 140.1, 136.0, 133.1, 131.8, 126.4, 71.4,69.6, 57.5, 56.9, 55.3, 54.1, 29.78, 28.20, 21.1, 19.3-20.2, 17.7; LC-MS: *m/z* $[M+H]^+$ = 290.36.

8.3.3 Traceless Synthesis of Alkyne Building Blocks

Reductive amination: PL-FMP resin (150 mg) was swelled in DCE (15 ml) for 2 h. The solvent was removed, followed by the addition of 2 % glacial acetic acid solution in DCE (15 mL) and the propargylamine (10.0 eq). The reaction mixture was shaken for 3 h. Then $Na(OAc)_3BH$ (10.0 eq) was added and the reaction mixture was shaken for further 18 h. The resulting resin was washed with THF (3×20 mL), MeOH (3×20

mL) and DCM (3×20 mL) and dried under vacuum to afford **12**. The completeness of the reaction can be monitored by Chloranil test. Blue beads indicate the presence of secondary amine and the completeness of the reaction.



Scheme 8.7 Traceless synthesis of alkynes **14(A-S).**

Traceless N-acylation with Acids: The resulting resin firstly was swelled in DMF for 2 h. The solvent was removed. Then the resin was added a preactived solution of acids (5.0 eq), PyBOP (5.0 eq), HOAt (5.0 eq) and DIPEA (10.0 eq) in DMF (or solution of sulfonyl chloride (4.0 eq) and DIPEA (6.0 eq) in DCM. The reaction mixture was shaken overnight at room temperature and then the resin was washed with THF (3×), MeOH (3×) and CH_2Cl_2 (3×) and dried under vacuum to afford **13**. A negative chloranil test indicates the completeness of the reaction.

Cleavage of the product **13(A-S)** from solid support: Dried resin was treated with cleavage (3 mL) solution consisting of TFA:DCM:H_2O (5:4:1) and the mixture was shaken for 3 h at r.t. followed by filtration. The solution was concentrated to dryness

to afford warhead **14(A-S)**.

N-(prop-2-ynyl)-2-naphthamide (**14A).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.24 (s, 1H), 4.12-4.17 (m, 1H), 7.56-7.73 (m, 4H), 8.48 (s, 1H), 9.11 (t, *J* = 5.40 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 166.0, 134.2, 132.1, 131.2, 128.9-126.7(m), 124.5, 81.3, 72.9, 28.8; LC-MS: *m/z* $[M+H]^+$ = 210.083.

3-methoxy-N-(prop-2-ynyl)benzamide (**14D).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.17 (s, 1H), 3.82 (s, 3H), 4.04-4.29 (m, 2H), 7.08-7.11 (m, 1H), 8.92 (t, *J* = 5.23 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 165.7, 158.6, 135.2, 129.5, 119.9, 117.4, 112.3, 81.3, 72.8, 55.3, 27.0; LC-MS: *m/z* $[M+H]^+$ = 190.085.

N-(prop-2-ynyl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (**14E).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.17 (s, 1H), 3.82 (s, 3H), 4.03-4.06 (m, 2H), 4.26-4.29 (m, 2H), 4.34-4.37 (m, 2H), 6.85-7.00 (m, 2H), 7.27-7.36 (m, 1H), 8.45 (t, *J* = 5.24 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 164.2, 143.6, 141.8, 123.1, 122.1, 120.6, 119.7, 81.3, 72.7, 64.4, 63.6, 28.5; LC-MS: *m/z* $[M+H]^+$ = 218.077.

4-Fluoro-N-(prop-2-ynyl) benzamide (**14G).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.07 (s, 1H), 4.04-4.06 (q, 2H), 7.27 (t, *J* = 5.30 Hz, 1H), 7.93 (q, 2H), 8.95 (s, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 165.2, 165.1, 163.2, 159.0-158.1(m), 130.2, 115.3, 81.4, 72.9, 35.9, 30.8, 28.7; LC-MS: *m/z* $[M+H]^+$ = 178.058.

2-(2-fluorophenyl)-N-(prop-2-ynyl) acetamide (**14I).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.27 (s, 1H), 3.56 (s, 2H), 3.85-3.88 (m, 2H), 7.11-7.16 (m, 2H), 7.27-7.34 (m, 2H),

8.52 (s, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 168.9, 162.2, 158.9, 131.8, 128.6, 124.1, 114.9, 81.1, 73.0, 34.9, 28.5; LC-MS: *m/z* $[M+H]^+$ = 192.076.

4-Acetyl-N-(prop-2-ynyl) benzenesulfonamide (**14N).** ^{1}H-NMR (300 MHz, DMSO-d^6) δ 2.50 (s, 3H), 3.02 (s, 1H), 7.75 (q, *J* = 3.48 Hz, 2H), 7.94 (d, *J* = 8.34 Hz, 2H), 8.12 (d, *J* = 8.34 Hz, 1H), 8.36 (t, *J* = 5.91 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 197.4, 144.3, 139.5, 128.9, 127.1, 79.1, 74.9, 31.9, 27.1.

N-(prop-2-ynyl)naphthalene-2-sulfonamide (**14K).** ^{1}H-NMR (300 MHz, $CDCl_3$) δ 1.26 (s, 1H), 3.88 (s, 2H), 7.58-7.67 (m, 2H), 7.84-7.98 (m, 4H), 8.48 (s, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 136.2, 134.9, 132.0, 129.5-128.8(m), 122.4, 77.4, 73.0, 32.9.

N-(prop-2-ynyl)biphenyl-4-sulfonamide (**14Q).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.06 (t, *J* = 2.25 Hz, 1H), 3.74 (t, *J* = 2.79 Hz, 2H), 7.41-7.53 (m, 3H), 7.74 (d, *J* = 7.32 Hz, 2H), 7.89 (d, 4H), 8.20 (t, *J* = 5.75 Hz, 2H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 144.0, 139.2, 138.5, 132.1, 129.1, 127.4-127.1(m), 79.4, 74.8, 32.0.

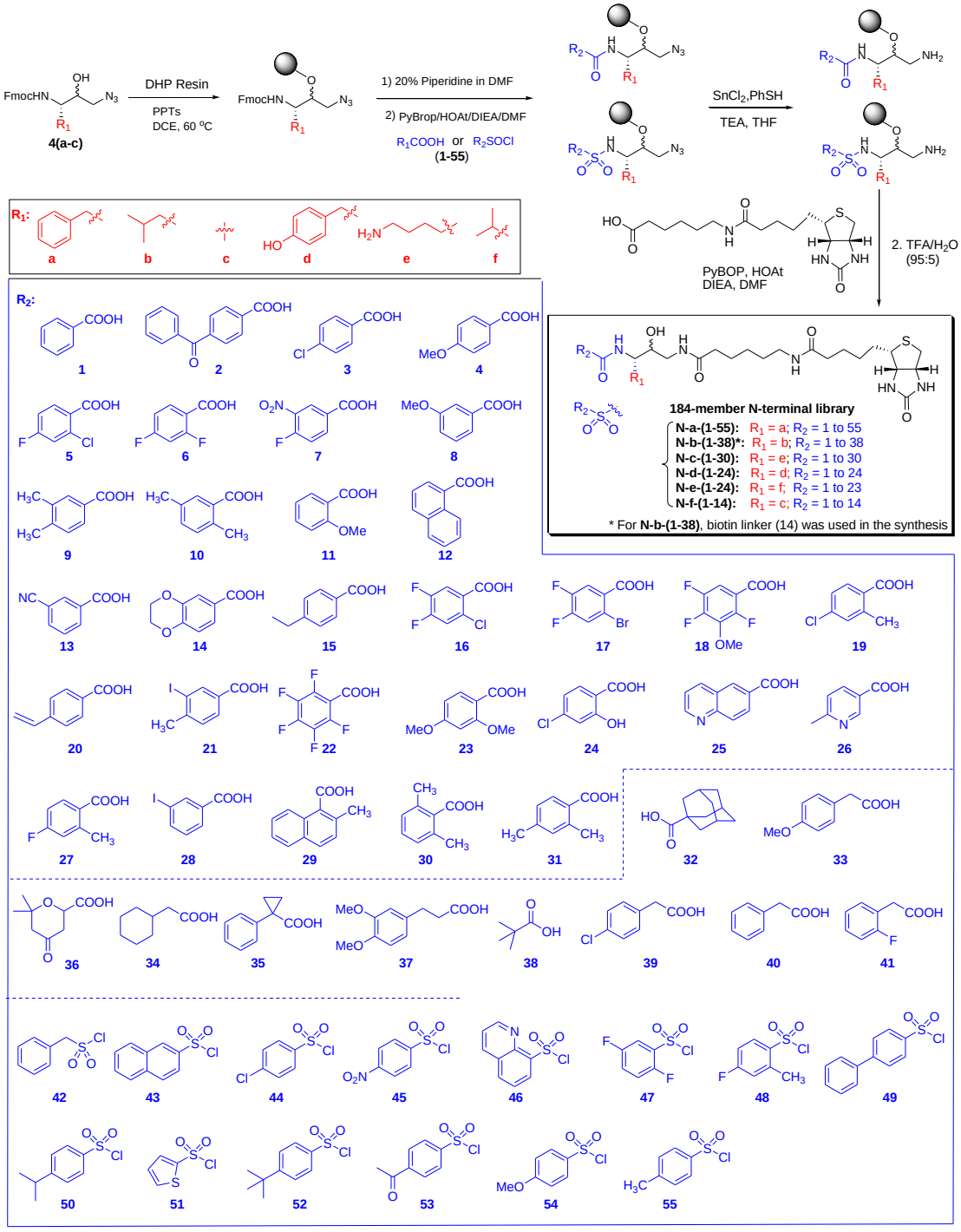
4-nitro-N-(prop-2-ynyl) benzenesulfonamide (**14R).** ^{1}H-NMR (300 MHz, DMSO-d6) δ 3.01 (s, 1H), 3.56 (d, *J* = 6.0 Hz, 2H), 8.06 (d, *J* = 8.70 Hz, 2H), 8.41 (d, *J* = 8.70 Hz, 2H), 8.54 (t, *J* = 5.75 Hz, 2H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 149.6, 146.3, 128.5, 124.4, 78.8, 75.2, 31.9.

8.3.4 Procedure for Synthesis of the 198-member N-terminal Library

*Immobilization of **4(a-f)** onto DHP resin.* According to previously published

procedures.[146-148] HM DHP resin (40 mg) was first swelled in DCE (5 mL) for 2 h. The solvent was then removed and **4(a-f)**(4.0 eq) and catalytic amount of PPTs (Pyridinium p-toluenesulfonate) (1.5 eq) in DCE (10 mL) were added at r.t. The reaction was then strirred for 12 h at 60 ^{o}C. The resulting resin was washed with NMP (3×), THF (3×), CH_2Cl_2 (3×) and Et_2O (3×) and dried in vacuo, then stored at -20 ^{o}C.

Deprotection of Fmoc group. The resulting resin was suspended in a solution of 20% piperidine in NMP, shaked for 2 h at r.t. The solvent was then removed by filtration and washed with NMP (3×), THF (3×), CH_2Cl_2 (3×) and Et_2O (3×) and dried under high vacuum. The completeness of the reaction was monitored by ninhydrin test. Blue beads indicate the presence of primary amine and the completeness of the reaction.

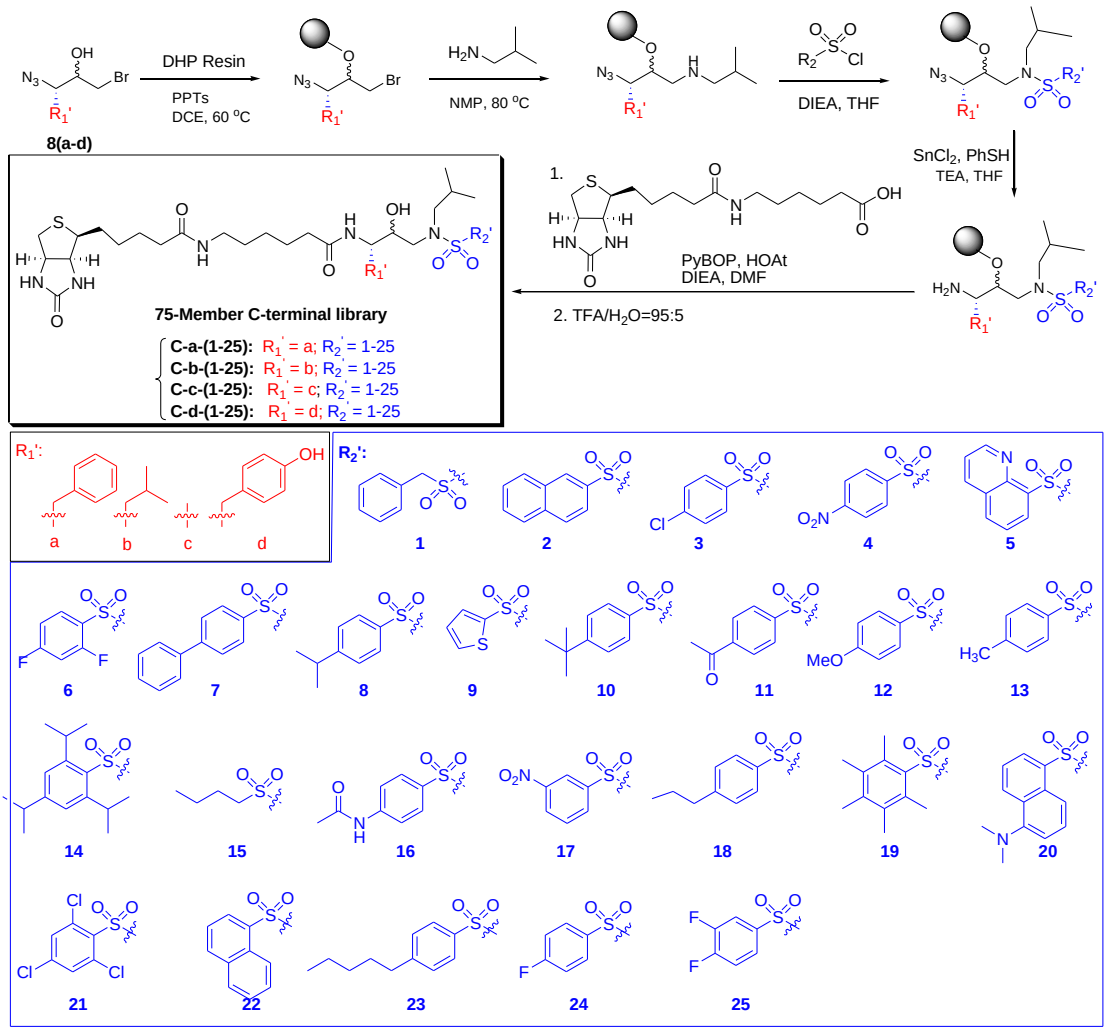


Scheme 8.8 Solid-phase synthesis of the N-terminal library.

Coupling reactions for acids and sulfonyl chlorides. The resulting resin firstly was swelled in dry THF for 2 h, and the solvent was removed. To the resin was added a preactived solution containing the corresponding acid (4.0 eq), PyBOP (4.0 eq), HOAt (4.0 eq) and DIPEA (8.0 eq) in THF or a solution of the corresponding

sulfonyl chloride (4.0 eq) and DIEA (8.0 eq) in dry THF. The reaction mixtures were shaken overnight at r.t. and the resin was finally washed with NMP (3×), THF (3×), CH_2Cl_2 (3×) and Et_2O (3×) and dried under vacuum. The completeness of the reaction was monitored by ninhydrin test. Negative test indicate the absence of primary amine and the completeness of the reaction.

8.3.5 Procedure for Synthesis of the 100-member C-terminal Library



Scheme 8.9 Solid-phase synthesis of the C-terminal library.

*Immobilization of **8(a-d)** onto DHP resin.* HM DHP resin (40 mg) was swelled in

DCE (5 mL) for 2 h. The solvent was then removed and **8(a-d)** (4.0 eq) and catalytic amount of PPTs (Pyridinium *p*-toluenesulfonate) (1.5 eq) in DCE (10 mL) were added at r.t. The reaction was then strirred for 12 h at 60 ^{o}C. The resulting resin was washed with NMP (3×), THF (3×), CH_2Cl_2 (3×) and Et_2O (3×) and dried in vacuo, then stored at -20 ^{o}C.

*Amine alkylation and coupling reactions for C-terminal library***.** Support-bound bromomethyl alcohols were added to a vial followed by addition of a solution of *i*-butylamine (10 eqv) in NMP (10 mL). The vial was then sealed and the reaction mixture was heated at 80 ^{o}C for 36 h. The resin was subsequently washed with NMP (3×), THF (3×), CH_2Cl_2 (3×), and ether (3×), and then dried in vacuo. Acylation with the sulfonyl chloride (4.0 eqv) and DIEA (8.0 eqv) in dry THF was carried out overnight. The resin was then washed with NMP (3×), THF (3×), CH_2Cl_2 (3×), and ether (3×), and then dried under vacuum.

*Reduction of azido group to primary amine***.** Reduction of the azide was accomplished using 0.2 M $SnCl_2$, 0.8 M PhSH, and 1.0 M Et_3N in THF (1 mL) for 4 h. The resin was then washed with 50 vol % aqueous THF solution (3×), THF (3×), CH_2Cl_2 (3×) and ether (3×), and then dried under vacuo. The reaction can be monitored by using ninhydrin test. If primary amine presents, the beads will become blue color after treatment with ninhydrin reagents.

*Acylation of Primary Amine with Long Chain Biotin Acids***.** The above resulting support-bound amine was then acylated in a preactived solution of biotin acid (4.0

eq), PyBOP (4.0 eq), HOAt (4.0 eq) and DIPEA (8.0 eq) in THF. The reaction mixture was shaken overnight at r.t., and the resins were washed with NMP (3×), THF (3×), CH_2Cl_2 (3×) and Et_2O (3×) and dried under vacuum. The completeness of the reaction was monitored by ninhydrin test. If the coupling reaction is completed, the beads won't have color changes after treatment with ninhydrin reagents.

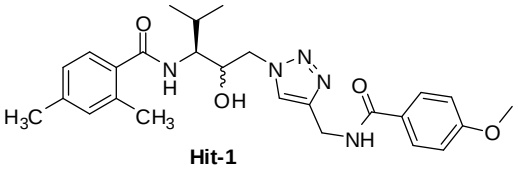
Cleavage of small molecules from solid support: The compound cleavage, microreactors containing of both N- & C-terminal sublibraries were decoded first, then cleaved with a 3 mL solution containing TFA (95%) and H_2O (5%) for 20 min at r.t. Following concentration in vacuo until >90% of cleavage cocktail was removed. Cold ether (chilled to -20 ℃) was added to the liquid residue to precipitate the compounds. The compounds were allowed to precipitate at -20 ℃ for overnight. The ether layer was discarded and the precipitates were dried thoroughly in vacuo. The solids were dissolved in 0.5 mL of DMSO and stored at -20 ℃ for future use. LCMS was performed to ensure the compounds were of correct mass and sufficient purity. Mass spectra were recorded on a Finnigan LCQ mass spectrometer, or a Shimadzu LC-IT-TOF spectrometer. Analytical HPLC was carried out on Shimadzu LC-IT-TOF system equipped with an autosampler, using reverse-phase Phenomenex Luna 5 μm C18(2) 100 Å 150×3.0 mm (for peptides) or 50×3.0 mm (for all other samples) columns. 0.1% TFA/H_2O and 0.1% TFA/acetonitrile were used as eluents. The flow rate was 0.6 ml/min. Each samples were run for 20 min under 10-100% ACN condition.

8.4 Click Chemistry Synthesis

8.4.1 Construction of the 152-member Inhibitor Library against PMs

The 152-membered library of aspartic protease inhibitors were assembled *in situ* in a 384-deep well microplate. Both the alkynes and azides were dissolved in a minimal amount of DMSO before loading into each well in proportions. The alkynes (1.2 eq) were added to the reaction in slight excess so as to ensure that all azides (1.0 eq) were consumed completely. Subsequently, 1 mL of a mixture of DMSO/H_2O (1:1 volume ratio) was applied to each well. The microplate was shaken for a few minutes, followed by sequential addition of catalytic amount of sodium ascorbate (~40%) and $CuSO_4$ (~10%) into each well to initiate the "Click" chemistry assembly, which was continued at r.t. for another 2 d with shaking. The assembled products were used directly for subsequent LC-MS analysis and enzymatic assays without any further purification. LC-MS analysis of all 152 *in situ*-assembled products indicated the complete consumption of the azides and quantitative formation of the correct triazole products in almost every case.

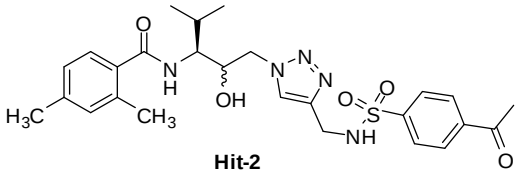
Charaterization of 8 hits for HAP:



Hit-1

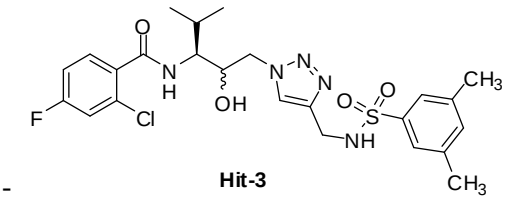
White solid. ^{1}H-NMR (500 MHz, DMSO-d6) δ 0.83-0.96 (m, 6H), 1.91-1.99 (m, 1H), 2.33 (s, 3H), 2.31 (s, 3H), 3.80 (s, 2H), 3.91-3.95 (m, 2H), 4.00-4.26 (m, 2H), 4.50 (s, 3H), 5.27(d, *J* = 7.55 Hz, 2H), 6.98 (d, *J* = 7.55 Hz, 2H), 7.05 (d, *J* = 8.80 Hz, 2H), 7.29 (dd, J_1 = 7.55 Hz, J_2 = 7.55 Hz, 1H), 7.85-7.87 (t, *J* = 4.40 Hz, 3H), 8.05 (s, 1H),

8.90 (s, 1H); ^{13}C-NMR (125 MHz, d6-DMSO): 169.6, 165.5, 161.6, 145.0, 138.6, 135.1, 134.9, 134.8, 130.9, 129.1, 127.1, 126.4, 126.3, 125.9, 123.6, 113.5, 69.6, 59.7, 56.8, 55.6, 55.3, 53.9, 53.7, 34.8, 28.6, 27.4, 20.7, 20.4, 19.7, 19.6, 19.5, 19.48, 15.9, 14.1; LC-MS: *m/z* $[M+H]^+$ = 480.00.



Hit-2

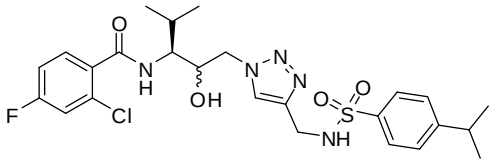
White solid. ^{1}H-NMR (500 MHz, DMSO-d6) δ 0.84-0.95 (m, 6H), 2.20-2.29 (m, 1H), 2.30-2.34 (m, 6H), 2.69 (s, 3H), 3.68-3.79 (m,1H), 3.89-3.94 (m, 1H), 4.09 (s, 2H), 4.11-4.20 (m, 1H), 4.43(dd, J_1 = 1.85 Hz, J_2 = 1.90 Hz, 1H), 5.75 (s, 1H), 7.06 (d, J = 6.30 Hz, 2H), 7.29 (dd, J_1 = 7.55 Hz, J_2 = 8.20 Hz, 1H), 7.80 (t, J = 9.45 Hz, 1H), 7.90-7.92 (m, 2H), 8.00 (d, J = 10.10 Hz, 2H); ^{13}C-NMR (125 MHz, DMSO-d6): 197.3, 169.6, 144.3, 143.0, 142.9, 139.3, 138.7, 135.1, 134.8, 134.5, 131.0, 128.9, 127.1, 126.8, 125.9, 124.0, 69.5, 61.2, 56.8, 55.6, 53.8, 50.5, 38.1, 28.5, 27.4, 27.0, 22.1, 20.8, 20.3, 19.7, 19.6, 19.5, 15.9; LC-MS: *m/z* $[M+H]^+$ = 528.218.



Hit-3

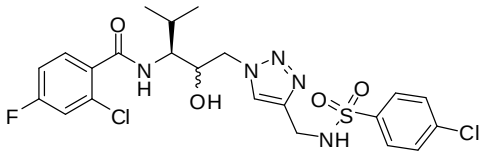
White solid. ^{1}H-NMR (300 MHz, DMSO-d6) δ 0.85 (d, J = 6.39 Hz, 3H), 0.95 (d, J = 6.57 Hz, 3H), 2.15-2.45 (m, 7H), 3.72-3.78 (m, 2H), 4.03 (d, J = 6.09 Hz, 1H), 4.11-4.19 (m, 1H), 4.52 (d, J = 12.99 Hz, 1H), 7.29-7.35 (m, 2H), 7.35 (s, 2H), 7.51-7.55 (m, 2H), 7.80 (s, 1H), 8.03 (t, J = 5.84 Hz, 1H), 8.34 (d, J = 9.39 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6): 166.1, 143.3, 140.2, 138.6, 134.2, 130.8, 124.0,

117.0, 114.3, 69.6, 56.0, 53.7, 39.0, 27.4, 20.7, 20.3, 15.7. LC-MS: *m/z* $[M+H]^+$ = 538.1544.



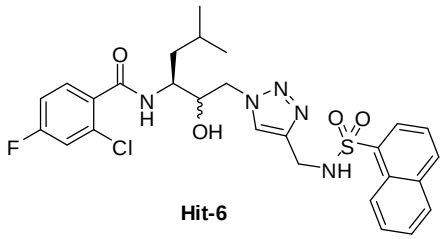
Hit-4

White solid. ^{1}H-NMR (300 MHz, DMSO-d6) δ 0.84 (d, *J* = 6.57 Hz, 3H), 0.95 (d, *J* = 6.72 Hz, 3H), 1.22 (d, *J* = 6.75 Hz, 6H), 2.12-2.30 (m, 1H), 2.93-2.99 (m, 1H), 3.77 (t, *J* = 9.56 Hz, 1H), 4.52 (d, *J* = 13.80 Hz, 1H), 5.33 (d, *J* =7.23 Hz, 1H), 7.31 (t, *J* = 7.40 Hz, 1H), 7.45 (d, *J* = 7.38 Hz, 2H), 7.52 (d, *J* = 8.22 Hz, 2H), 7.72 (d, *J* = 7.23 Hz, 2H), 7.83 (s, 1H), 8.06 (s, 1H), 8.34 (d, *J* = 9.54 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6): 166.1, 163.6, 160.3, 153.2,143.3,137.9, 134.1, 130.8, 126.8, 124.0, 116.8, 114.4, 69.6, 56.0, 53.7, 39.5, 33.4,27.4, 23.5, 20.3, 19.7, 15.7, 14.1. LC-MS: *m/z* $[M+H]^+$ = 554.1524.



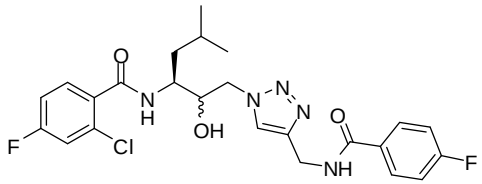
Hit-5

White solid. ^{1}H-NMR (300 MHz, DMSO-d6) δ 0.84 (d, *J* = 6.96 Hz, 3H), 0.95 (d, *J* = 6.63 Hz, 3H), 2.19 (m, 1H), 3.75 (t, *J* = 8.01 Hz, 2H), 3.86-3.93 (m, 2H), 4.01-4.06 (m, 3H), 4.12-4.19 (m, 1H), 4.51 (d, *J* = 12.54 Hz, 1H), 7.28-7.34 (m, 1H), 7.51-7.56 (m, 2H), 7.60 (d, *J* = 8.37 Hz, 2H), 7.76 (d, *J* = 8.70 Hz, 2H), 8.34 (d, *J* = 9.39 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6): 166.1, 163.6, 160.3, 143.8,140.3,136.7, 134.1, 131.0, 130.6, 124.0, 116.7, 114.4, 69.6, 59.8, 55.9, 54.9, 53.7,27.4, 20.5, 15.7, 14.1. LC-MS: *m/z* $[M+H]^+$ = 544.0840.



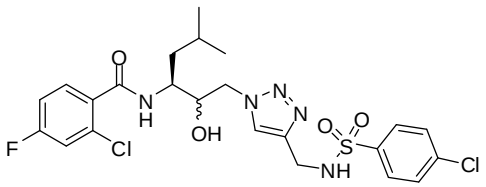
Hit-6

White solid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.91 (d, *J* = 3.12 Hz, 6H), 1.43-1.65 (m, 2H), 1.76 (s, 1H), 3.73 (d, *J* = 6.96 Hz, 2H), 3.99-4.18 (m, 2H), 4.25-4.57 (m, 1H), 4.62 (d, *J* = 5.58 Hz, 2H), 5.39 (d, *J* = 6.24 Hz, 1H), 7.28-7.40 (m, 2H), 7.44-7.57 (m, 4H), 7.75 (d, *J* = 7.65 Hz, 1H), 7.76-7.79 (m, 2H), 8.05 (d, *J* = 9.39 Hz, 1H), 8.45 (m, 1H), 9.03 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 168.6, 165.8, 165.7, 163.6, 160.3, 144.5, 134.5, 133.9, 131.3, 131.0, 130.9, 130.6, 130.0, 129.2, 127.2, 125.0, 123.7, 117.1, 116.7, 114.4, 72.7, 59.8, 53.0, 50.3, 34.6, 24.3, 23.6, 21.4, 20.8, 19.2, 14.1. LC-MS: *m/z* $[M+H]^+$ = 574.158.



Hit-7

White solid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.89 (d, *J* = 1.74 Hz, 6H), 1.43-1.55 (m, 2H), 1.71-1.73 (m, 1H), 3.70 (t, *J* = 7.32 Hz, 2H), 3.91-4.29 (m, 3H), 4.45-4.60 (m, 3H), 7.29 (t, *J* = 8.87 Hz, 3H), 7.48-7.52 (m, 2H), 7.92-7.97 (m, 3H), 8.34 (dd, J_1 = 19.17 Hz , J_2 = 9.03 Hz, 1H), 9.06 (t, *J* = 5.57 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 165.7, 165.3, 163.6, 162.3, 160.3, 144.7, 133.9, 131.0, 130.5, 129.9, 123.7, 117.0, 116.7, 115.2, 114.3, 72.7, 53.1, 50.3, 38.3, 24.3, 23.6, 21.4. LC-MS: *m/z* $[M+H]^+$ = 506.163.



Hit-8

White solid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 0.91 (d, *J* = 3.84 Hz, 6H), 1.31-1.49 (m, 2H), 1.52-1.56 (m, 1H), 3.67 (s, 1H), 4.07-4.23 (m, 4H), 4.49 (d, *J* =11.84 Hz, 1H), 5.33 (s, 1H), 7.30 (t, *J* = 9.00 Hz, 1H), 7.48-7.55 (m, 2H), 7.81 (dd, J_1 = 6.96 Hz, J_2 = 8.37 Hz, 3H), 8.32 (d, *J* = 9.06 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): 165.7, 163.6, 160.3, 142.9, 139.4, 137.2, 133.9, 131.1, 130.9, 130.6, 130.4, 128.8, 124.0, 117.0, 116.7, 114.4, 114.1, 72.6, 53.0, 50.3, 38.1, 24.0, 21.4. LC-MS: *m/z* $[M+H]^+$ = 558.109.

8.4.2 Synthesis of Affinity-based Probes Targeting Aspartic Proteases

Synthesis of all the probes targeting aspartic proteases was followed procedures similar to ones previously. In brief, the alkyne (1.2 eq) and the azide (1.0 eq) were dissolved in a minimal amount of DMSO. A mixture of DMSO/H_2O solution (1:1; 100 μL) was subsequently added and the reaction was shaken for a few minutes to obtain a clear solution. The "click" chemistry was initiated by sequential addition of catalytic amounts of sodium ascorbate (0.08 eq) and $CuSO_4$ (0.02 eq). The reaction was continued with shaking at r.t. for another 12 hrs. The reaction product was then directly injected on LC-MS; results indicated the complete consumption of the azide and quantitative formation of the triazole final product in all cases. The final probes (**A-G**) were subsequently characterized by LC-MS.

***Biotin-A3*.** ^{1}H-NMR (300 MHz, DMSO-d6) δ 0.70 (d, *J* = 5.85 Hz, 6H), 0.78 (d, *J* = 5.55 Hz, 2H), 0.90-1.04 (m, 3H), 1.23-1.55 (m, 10H), 1.97-2.03 (m, 64H), 2.27 (s,

2H), 2.66-3.00 (m, 10H), 3. 95-4.10 (m, 1H), 4.25-4.27 (m, 2H), 4.55-4.71 (m, 3H), 5.55 (s, 1H), 6.50 (s, 1H), 6.79 (s, 1H), 7.37 (d, J = 8.19 Hz, 3H), 7.52 (d, J = 6.99 Hz, 2H), 7.50-7.66 (m, 6H), 7.75-7.85 (m, 4H), 8.06 (s, 2H), 8.54 (s, 2H); LC-MS: m/z $[M+H]^+$ = 1023.150.

***Biotin-A1*.** ^{1}H-NMR (300 MHz, DMSO-d6) δ 0.72 (d, J = 5.85 Hz, 6H), 0.78 (d, J = 6.15 Hz, 2H), 1.13-1.41(m, 16H), 2.00 (t, J = 6.99 Hz, 4H), 2.66-2.78 (m, 6H), 2.85-3.15 (m, 24H), 2.29 (s, 2H), 5. 71 (s, 2H), 6.91 (s, 2H), 7.08 (s, 2H), 7.37 (s, 2H), 7.50-7.65 (m, 3H), 7.66-7.75 (m, 4H), 7.81 (s, 1H); LC-MS: m/z $[M+H]^+$ = 990.150.

Table 8.1 Summary of LC-MS Characterizations of TER/Biotin Probes.

Compd	MW_{calc}	MW_{obs}	Compd	MW_{calc}	MW_{obs}
TER-Bp-alkyne	1012.53	1012.4459	TER-A8	1352.70	1352.7050
Biotin-Bp-alkyne	645.80	646.2415	Biotin-T1	1081.46	1082.0000
TER-T1	1448.69	1449.4000	Biotin-T2	1105.48	1106.4935
TER-T2	1472.71	1472.1000	Biotin-T3	1114.48	1115.1500
TER-T3	1481.71	1482.2100	Biotin-T4	1108.45	1109.4665
TER-T4	1475.69	1476.3250	Biotin-A3	1022.45	1045.3755
TER-T5	1524.70	762.8499	Biotin-A4	1016.42	1039.2927
TER-A3	1389.69	1390.3000	Biotin-A5	1064.50	1065.4386
TER-A4	1383.70	1383.6530	Biotin-A6	1007.42	1008.4510
TER-A5	1431.73	1432.6370	Biotin-A7	1028.30	1028.2812
TER-A6	1374.70	1374.6310	Biotin-A8	985.46	1008.1500
TER-A7	1394.74	1394.6470	Biotin-A1	989.43	990.1500

HPLC conditions: 0-100% B for 10 min, then 100% B for 2 min. (Solvent A: 100% H_2O with 0.1 % TFA; Solvent B: 100% CH_3CN with 0.1 % TFA).

8.5 Microplate Assay

8.5.1 Inhibition Activity of Diols Library against Aspartic Proteases

The activities of the aspartic proteases were determined by measuring the rate of hydrolysis from the commercial available fluorogenic substrates (AnaSpec, USA).

The inhibition of all proteins were screened using black polypropylene flat-bottom 384-well microtiter plates (Greiner, Germany) in a total reaction volume of 20 μl/well, monitored with a Tecan Infinite F200 microplate reader at excitation of 390 nm and emission of 492 nm prior to the addition of substrate. Different concentrations of inhibitors were used in the assay (25 μM~2.5 μM), where necessary. Negative controls were performed in the absence of enzyme and positive controls were carried out in the presence of enzyme but without inhibitor. Pepsin and HIV-1 protease were commercially available. Plasmepsin I and Plasmepsin II were kindly provided by Dr. R. Yada (Guelph University, Canada). The assay conditions were given below:

<u>HIV-1</u>: HIV-1 (5 nM in 2 × Assay buffer) = 10 μL.

HIV Protease Substrate 1

(Arg-Glu(EDANS)-Ser-Gin-Asn-Tyr-Ile-Val-Gin-Lys(dabcyl-Arg) (5 μM in dd H_2O) = 5 μL.

Inhibitor (varied in 10% DMSO/H_2O) = 5 μL

[Assay buffer (1 ×) = 100 mM Sodium Acetate, 1.0 M NaCl, 1.0 mM EDTA, 1.0 mM DTT, 1 mg/mL BSA, *p*H = 4.7].

<u>Plasmepsin I:</u> Plasmepsin I (3 nM in 2 × Assay buffer) = 10 μL

Substrate EDANS-CO-CH_2-CH_2-CO-ALERMFLSFP-Dap-(DABCYL)-OH (1.5 μM in dd H_2O) = 5 μL

Inhibitor (varied in 10% DMSO/H_2O) = 5 μL

[Assay buffer (1 ×) = 100 mM sodium acetate, 10 % glycerol, *p*H 5.0].

<u>Plasmepsin II:</u> Plasmepsin II (3 nM in 2 × Assay buffer) = 10 μL

Substrate EDANS-CO-CH_2-CH_2-CO-ALERMFLSFP-Dap-(DABCYL)-OH (1.5 μM in dd H_2O) = 5 μL

Inhibitor (varied in 10% DMSO/H_2O) = 5 μL

[Assay buffer (1×) = 100 mM sodium acetate, 10 % glycerol, *p*H 5.0].

Pepsin: Pepsin (2 nM in 2×Assay buffer) = 10 μL

Substrate MOCAc-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(Dnp)-NH_2 (5 μM in dd H_2O) = 5 μL

Inhibitor (varied in 10% DMSO/H_2O) = 5 μL

[Assay buffer (1×) =10 mM HCl, *p*H 2.0].

The enzymatic reactions were allowed to incubate at r.t. for 15 min before being initiated by the addition of substrate, and the reaction was monitored at λ_{ex} = 340 nm and λ_{em} = 490 nm, period of 1 or 2 h. Eventually, an inhibitor fingerprint of the library against the aspartic proteases was obtained, from which six potential hits (SYM-5, -21, -35, -43 and ASM-16, -29) were identified.

8.5.2 Inhibition Activity of Dasatinib/staurosphorine Probes against Protein Kinases

Concentration-dependent experiments were performed to determine the inhibition potency and the binding affinity of the probes towards the catalytic domain of c-Src, Abl, and Csk. The inhibition assay was performed with Kinase-Glo® Plus Luminescent Kinase assay kit from Promega following the manufactures instructions. Briefly, recombinant proteins, corresponding peptide substrates (Src substrate

subsequence: KVEKIGEGTYGVVYK; Abl substrate sequence: EAIYAAPFAKKK; CSK substrate subsequence: KKKKEEIYFFF), ATP and the probe were mixed in the HEPES buffer (25 mM HEPES, *p*H 7.5, 150 mM NaCl, 2 mM $MgCl_2$) at a volume of 27.5 μL in a flat-bottom solid white 384-well plate. The incubation was allowed to continue for 20 min at 37 ℃ and the reaction was subsequently quenched by the addition of an equal volume of the Kinase-Glo reagent. After 5 min of incubation, the luminescence readouts from the wells were measured using Tecan microplate reader with *i*-control software. The ATP and substrate peptide concentrations used in the assay were 10 μM and 50 μM, respectively. The following control reactions were also performed simultaneously.

1. Enzyme control (No kinase, Buffer+Substrate+ATP),

2. Inhibitor control (No inhibitor, Buffer+Kinase+Substrate+ATP),

3. Substrate control (No substrate, Buffer+Kinase+ATP)

4. Buffer alone.

For the inhibition potency of staurosphorine probe, the inhibition assay was performed with ProFluor® PKA Assay Kit from Promega following the manufactures instructions. Briefly, recombinant PKA, Bisamide Rhodamine 110 peptide substrate (PKA R110 Substrate), ATP and probe **STS-1** were mixed in the 1× reaction buffer provided by kits at a volume of 10 μL in a flat-bottom black 384-well plate. The incubation was allowed to continue for 20 min at r.t. and 5 μL protease reagents in 1× termination buffer was added into the reaction, mixed and incubated for 30 min at r.t. Finally, the reaction was quenched by the addition of 5 μL stabilizer reagent in 1×

termination buffer, mixed and the fluorescence readouts from the wells were measured at 485/530 nm using Tecan microplate reader with *i*-control software.

The luminescence intensity from each well was measured and the inhibition potency was calculated using the following relation,

Potency = 1 - Remaining activity

Dose-dependent inhibition assays were performed by varying the concentration of the probe under optimized enzyme concentration of ~50 nM. The IC_{50} value of the probe was calculated from the percentage activity vs. log [concentration of probe] curves generated using GraphPad Prism software.

8.6 Microarray-based Screening

8.6.1 Preparation of Avidin Slides

25×75 mm glass slides (Sigma-Aldrich) were cleaned in piranha solution (sulfuric acid/hydrogen peroxide, 7:3). Amine functionality was incorporated onto the slides by silanization using a solution of 3% (aminopropyl) triethoxysilane in 2% water and 95% ethanol. After 1-2 h incubation, the slides were washed with ethanol and cured at 150 ℃ for at least 2 h. The resulting amine slides were incubated in a solution of 180 mM succinic anhydride in DMF for 30 min before being transferred to a boiling water bath for 2 min. The slides were rinsed with ethanol and dried under a stream of nitrogen. The carboxylic acid derivatized slide surface was activated with a solution of 100 mM of HBTU (O-Benzotriazole-N, N, N', N'-tetramethyl–uronium–hexafluoro-phosphate), 200 mM DIEA, and 100 mM N-hydroxysuccinimide in DMF, thus generating the NHS-derivatized slides. These surfaces were reacted with a

solution of 1 mg/mL avidin in 10 mM $NaHCO_3$ (*p*H = 9) for 40 min, washed with water, air dried. The unreacted NHS groups were quenched with a solution of 2 mM aspartic acid in a 0.5 M $NaHCO_3$ buffer, *p*H 9. These slides were washed with water, dried and stored at 4 ℃, ready for printing.

8.6.2 Microarray Preparation

All compounds stocks were prepared in a 1:1 DMSO/PBS spotting solution (to 5 mM final concentration), and were distributed in 384-well plates. All compounds were shown to be completely soluble in this spotting solution. One slide was spotted on an ESI SMA arrayer (Ontario, Canada) with the print head installed with 8 Stealth SMP8B Microspotting pins (Telechem, USA.). Spots generated were of approximately 350 μm diameter and were printed with a spot-spot spacing of 450 μm. The pins were rinsed in between samples using two cycles of wash (for 10 s) and sonication (for 10 s) in reservoirs containing 70 % ethanol followed by drying under reduced pressure (for 10 s). The slides were allowed to stand for 1 h on the printer platform and stored at 4 ℃ until use. Before incubation with the labeled proteins or lysates, the slides were rinsed with PBS (*p*H 7.4) for 10 min and blocked with PBS-containing 1 % BSA for 1 h. For studies with the 284-member library, all compounds were spotted on the same slide in duplicate.

8.6.3 Protein /proteome Labeling and Screening on SMM

Cellular lysates were minimally labeled with Cy3 dyes for 1 h on ice following the manufacturer's protocols. The excessive dye was quenched with a

10-fold molar excess of hydroxylamine for a further 1 h. The excess dye was removed by buffer exchange with a Microcon centrifugal filter (Millipore, USA) or extensive dialysis at 4℃ overnight (Amersham, GE Healthcare, USA). The labeled lysates was reconstituted in a final buffer volume of 30 µL of HEPES (*p*H 5.4) containing 1% bovine serum albumin. In a standard microarray experiment, the labeled lysates (4 µg in 30 µL HEPES) was applied under cover slip to the array. The samples were then incubated with the array in a humidified chamber for 1 h at r.t. before repeated rinses with PBST (PBS containing 0.05% of Tween 20), typically washes three times (one time/10 min) with gentle shaking. Slides were scanned using an ArrayWoRx microarray scanner installed with the relevant filters (Cy3, $\lambda_{ex/em}$ = 548/595 nm; Cy5, $\lambda_{ex/em}$ = 633/685 nm; fluorescein, $\lambda_{ex/em}$ = 490/528 nm).

8.6.4 Data Extraction and Analysis

Microarray data was extracted using the ArrayWoRx software. Values from duplicated points were background subtracted and averaged (Duplicated spots with a standard deviation > 0.2 were rejected). The data analysis was carried out with Microsoft Excel. Venn diagrams were generated to identify the potent binders against Aspartic proteases, using the Venn Diagram Generator (http://www.pangloss.com/seidel/Protocols/venn.cgi).

(a)

C-c-14	C-c-16	C-C-18	C-c-23	C-c-25	C-c-15	C-c-17	C-c-19	C-c-24	
C-c-4	C-c-6	C-c-8	C-c-10	C-c-12	C-c-5	C-c-7	C-c-9	C-c-11	C-c-13
C-a-20	C-b-21	C-b-23	C-b-25	C-c-2	C-a-21	C-b-22	C-b-24	C-c-1	C-c-3
C-a-10	C-a-12	C-a-14	C-a-16	C-a-18	C-a-11	C-a-13	C-a-15	C-a-17	C-a-19
N-c-44	C-a-2	C-a-4	C-a-6	C-a-8	C-a-1	C-a-3	C-a-5	C-a-7	C-a-9
N-b-50	N-b-52	N-b-31	N-c-2	N-c-42	N-b-51	N-b-53	N-c-1	N-c-3	N-c-43
N-b-37	N-b-40	N-b-42	N-b-44	N-b-47	N-b-38	N-b-55	N-b-43	N-b-46	N-b-49
N-a-45	N-a-47	N-b-20	N-b-22	N-b-33	N-a-46	N-a-48	N-b-21	N-b-32	N-b-34
N-a-35	N-a-37	N-a-39	N-a-41	N-a-43	N-a-36	N-a-38	N-a-40	N-a-42	N-a-44
N-a-25	N-a-27	N-a-29	N-a-31	N-a-33	N-a-26	N-a-28	N-a-30	N-a-32	N-a-34
C-b-15	C-b-17	C-b-19	C-c-20	C-c-22	C-b-16	C-b-18	C-b-20	C-c-21	
C-b-5	C-b-7	C-b-9	C-b-11	C-b-13	C-b-6	C-b-8	C-b-10	C-b-12	C-b-14
N-c-40	C-a-22	C-a-24	C-b-1	C-b-3	N-c-55	C-a-23	C-a-25	C-b-2	C-b-4
N-c-18	N-c-18	N-c-22	N-c-33	N-c-37	N-c-19	N-c-21	N-c-32	N-c-34	N-c-38
N-c-6	N-c-10	N-c-12	N-c-14	N-c-16	N-c-8	N-c-11	N-c-13	N-c-15	N-c-17
N-b-12	N-b-14	N-b-16	N-b-16	N-c-4	N-b-13	N-b-15	N-b17	N-b-19	N-c-5
N-a-55	N-b-2	N-b-4	N-b-6	N-b-10	N-b-1	N-b-3	N-b-5	N-b-8	N-b-11
N-a-21	N-a-23	N-a-49	N-a-51	N-a-53	N-a-22	N-a-24	N-a-50	N-a-52	N-a-54
N-a-11	N-a-13	N-a-15	N-a-17	N-a-19	N-a-12	N-a-14	N-a-16	N-a-18	N-a-20
N-a-1	N-a-3	N-a-5	N-a-7	N-a-9	N-a-2	N-a-4	N-a-6	N-a-8	N-a-10

(b)

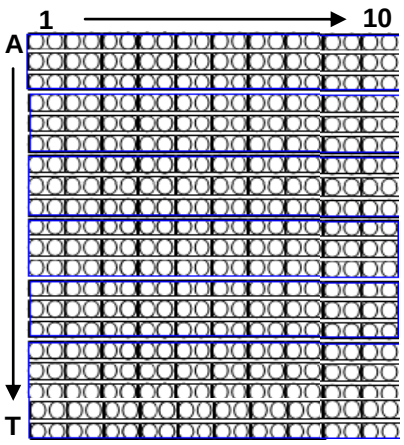
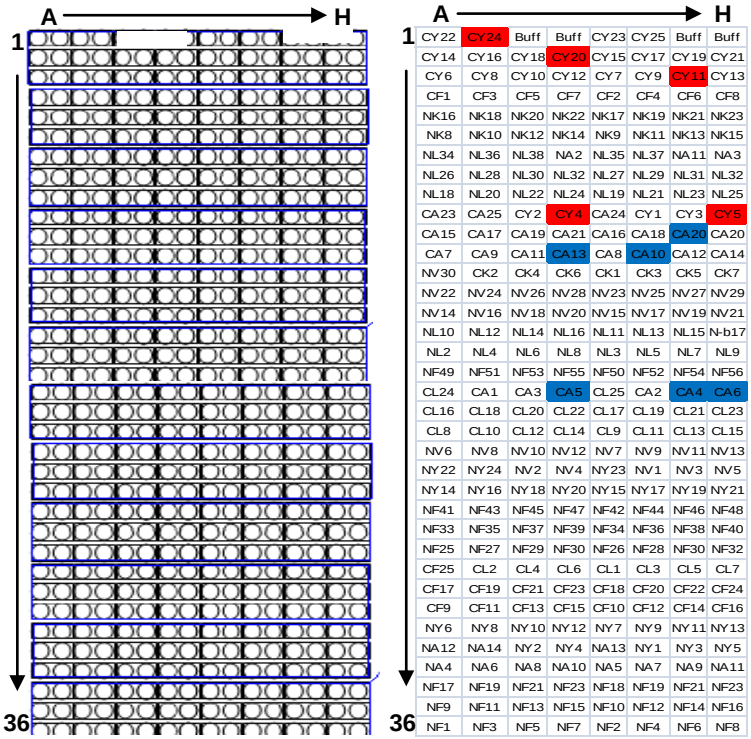


Figure 8.2 The spotting format of microarrays used in chapter 4 study. (a) The arrangement of the 198-member library on microarrary; (b) The spotting format (duplicate) of microarrays, unless otherwise specified.



CY22	CY24	Buff	Buff	CY23	CY25	Buff	Buff
CY14	CY16	CY18	CY20	CY15	CY17	CY19	CY21
CY6	CY8	CY10	CY12	CY7	CY9	CY11	CY13
CF1	CF3	CF5	CF7	CF2	CF4	CF6	CF8
NK16	NK18	NK20	NK22	NK17	NK19	NK21	NK23
NK8	NK10	NK12	NK14	NK9	NK11	NK13	NK15
NL34	NL36	NL38	NA2	NL35	NL37	NA11	NA3
NL26	NL28	NL30	NL32	NL27	NL29	NL31	NL32
NL18	NL20	NL22	NL24	NL19	NL21	NL23	NL25
CA23	CA25	CY2	CY4	CA24	CY1	CY3	CY5
CA15	CA17	CA19	CA21	CA16	CA18	CA20	CA20
CA7	CA9	CA11	CA13	CA8	CA10	CA12	CA14
NV30	CK2	CK4	CK6	CK1	CK3	CK5	CK7
NV22	NV24	NV26	NV28	NV23	NV25	NV27	NV29
NV14	NV16	NV18	NV20	NV15	NV17	NV19	NV21
NL10	NL12	NL14	NL16	NL11	NL13	NL15	N-b17
NL2	NL4	NL6	NL8	NL3	NL5	NL7	NL9
NF49	NF51	NF53	NF55	NF50	NF52	NF54	NF56
CL24	CA1	CA3	CA5	CL25	CA2	CA4	CA6
CL16	CL18	CL20	CL22	CL17	CL19	CL21	CL23
CL8	CL10	CL12	CL14	CL9	CL11	CL13	CL15
NV6	NV8	NV10	NV12	NV7	NV9	NV11	NV13
NY22	NY24	NV2	NV4	NY23	NV1	NV3	NV5
NY14	NY16	NY18	NY20	NY15	NY17	NY19	NY21
NF41	NF43	NF45	NF47	NF42	NF44	NF46	NF48
NF33	NF35	NF37	NF39	NF34	NF36	NF38	NF40
NF25	NF27	NF29	NF30	NF26	NF28	NF30	NF32
CF25	CL2	CL4	CL6	CL1	CL3	CL5	CL7
CF17	CF19	CF21	CF23	CF18	CF20	CF22	CF24
CF9	CF11	CF13	CF15	CF10	CF12	CF14	CF16
NY6	NY8	NY10	NY12	NY7	NY9	NY11	NY13
NA12	NA14	NY2	NY4	NA13	NY1	NY3	NY5
NA4	NA6	NA8	NA10	NA5	NA7	NA9	NA11
NF17	NF19	NF21	NF23	NF18	NF19	NF21	NF23
NF9	NF11	NF13	NF15	NF10	NF12	NF14	NF16
NF1	NF3	NF5	NF7	NF2	NF4	NF6	NF8

Figure 8.3 The spotting format of microarrays used in chapter 5 study, unless otherwise specified. 11 selected inhibitor hits were shaded in color.

8.7 Fluorescent Profiling of Aspartic Proteases and Protein Kinases

8.7.1 Labeling of Recombinant Aspartic Proteases

Detail of the labeling experiment was described below. Pepsin stocks were made

fresh using 10 mM HCl. Renin stock was made fresh by using water. Stocks of PM I (350 ng/mL in 100 mM NaOAc, 10% glycerol, *p*H 4.5), PM II (250 ng/mL in 100 mM NaOAc, 10% glycerol, *p*H 4.5) and HAP (1.5 mg/mL in 100 mM MES, 10% glycerol, *p*H 6.5) were diluted in appropriate buffer from the in -20 ^{o}C before use, respectively 100 mM NaOAc, 10% glycerol, *p*H 4.5 for PM I and PM II, 100 mM NaOAc, 10% glycerol, *p*H 6.5 for HAP. Labeling experiments were carried out with the following optimized conditions: desired enzyme amount, respectively 100 ng for pepsin, 100 ng for renin, 50 ng for PM I, PM II and HAP, were individually incubated with each probes (5 μM final concentration) in appropriate buffer for 20 min at r.t. For pepsin and renin, reaction buffer condition was as follow: 50 mM Tris Cl *p*H 2.0. For PM I, PM II and HAP, reaction buffer condition was 100 mM NaOAc, 10% glycerol, *p*H 4.5 or 100 mM $NaPO_4$, 10% glycerol *p*H 7,5. After 20 min incubation at r.t., samples were irradiated on ice for 25 min using a B100A lamp (UVP) at a distance of 5 cm. After irradiation, samples were boiled for 10 min with 4 μL of 6× loading buffer. The sample mixture (24 μL) were resolved by 12% SDS-PAGE followed by in-gel fluorescence scanning with the Typhoon gel scanner. Labeling bands arising from the experiment were spectral quantified by Image J software. Reactivity values were imported to cluster analysis software to generate heatmap and treeview for analysis.

8.7.2 Labeling and Identification of PM I, PM II and HAP

Infected RBC lysates were prepared as described previously. The 1.0 mg of total protein from infected RBC lysates was pre-incubated in labeling reaction buffer containing 100 mM sodium acetate, 10 % glycerol, *p*H 5.0 and 5 μM photo-affinity

probe (**P11**). After labeling, the reaction samples were desalted and concentrated by acetone precipitation overnight at -20 ^{o}C. Precipitated protein samples were collected by centrifugation at 13,000 rpm in a microcentrifuge. Samples were resuspended in 1× re-swelling buffer (8 M urea, 2% CHAPS and 2% IPG buffer and 0.002% bromphenol blue), and briefly sonicated. Just prior to use, DTT was added at a final concentration of 3 mg/mL. Rehydration buffer and stock DTT (62.5 mg/mL) were stored at -20 ℃. Rehydrated IEF strips were isoelectrically focused at r.t. under low viscosity oil with a gradient voltage of 0-200 V for 1 min, 200–3,500 V for 1.5 h, and a constant voltage of 3,500 V for 1.5 h. After IEF separation, the gel strip was reduced and alkylated. The reduction step was performed for 15 min in 10 mL of equilibration buffer **1** (0.5% w/v DTT in 50 mM Tris-HCl *p*H 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). The alkylation step was performed for 15 min in 10 mL of equilibration buffer **2** (4.5% w/v iodoacetamide in 50 mM Tris-HCl *p*H 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). The IEF strip was equilibrated in equilibration buffer stock solution (50 mM Tris-HCl *p*H 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, and 0.002% w/v bromphenol blue). Next, the equilibrated IEF strips were each placed on a preparative well and sealed using 1% agarose plus 0.002% (w/v) bromphenol. After electrophoresis the two-dimensional gels were scanned with Typhoon fluorescence scanner. Fluorescence spots corresponding to each target were excised directly from the gel and saved in -80 ^{o}C for future MS/MS analysis. Trypsin digestion was performed with In-Gel Trypsin Digestion Kit from Pierce. After digestion, digested

peptides were then extracted from the gel with 50% ACN and 1% formic acid. Extracts were concentrated by speedvec to 10 μL and analyzed by mass spectrometry. Spectra were analyzed by Mascot (www.mascot.com) using NCBInr database. Parameters used for Mascot searches were as follows: allow up to one missed proteolytic cleavage; peptide mass tolerance, 2 Da; production mass tolerance, 0.6 Da.

8.7.3 Characterization of Known Inhibitors for FV Plasmepsins

Further, we evaluated and validated the fidelity of this gel format approach as a surrogate assay for inhibitor discovery by testing several well-known plasmespins inhibitors against PM I and PM II and HAP**.** Inhibitor potency against well-characterized FV plasmepsins was assessed by incubation of those plasmepsins with each inhibitor followed by reaction with the aspartic protease activity–based probes. The potency of specific inhibitor scaffolds was measured as a ratio of the percent residual labeled proteases after inhibitor treatment relative to an untreated control.

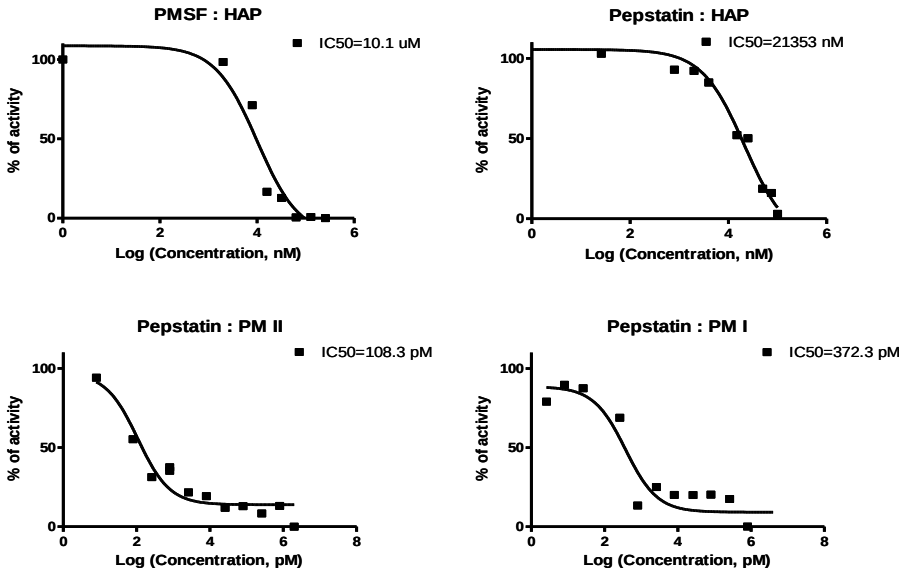


Figure 8.4 Characterization of two well-known inhibitors against recombinant FV plasmepsins in competitive labeling assay.

8.7.4 Characterization of A*f*BPs with γ-30 Cell Lysates

Details of the cell lysate labeling experiment are described below: a desired lysate amount (15 μg of total proteins as determined by Bradford assay) was incubated with the probes (5 μM final concentration; 5 % DMSO) in HEPES buffer for 30 min at r.t. After 30 min incubation, samples were irradiated on ice for 25 min using a B100A lamp (UVP) at a distance of 5 cm. After irradiation, samples were boiled for 10 min with 4 μL of 6× SDS loading buffer, resolved on a 12% SDS-PAGE followed by in-gel fluorescence scanning with a Typhoon gel scanner. Western blotting was carried out with anti-PS1 (Santa Cruz Biotechnology)

To further validate the labeling of PS-NTF by probe, biotin pull-down from labeled lysates experiments were carried out as described below: 2 mg of lysates was labeled by 5 μM probe **A1** in 2 mL reaction using previously described conditions. After labeling, the lysates in the reaction were acetone precipitated and resolubilized in 0.1% SDS in PBS with brief sonication. This resuspended sample was then incubated with avidin-agarose beads (100 μL/1 mg protein) at r.t. for 30 min. After centrifugation, supernatant were removed and the beads were washed with 1% SDS in PBS for 4 times. After washing, the beads were boiled in elution buffer (200 mM Tris *p*H 6.8, 400 mM DTT, 8% SDS). This pull-down sample from labeled lysates was then separated on 12% SDS PAGE gel together with pull-down sample from un-labeled lysates. After SDS-PAGE gels separation, proteins were then transferred to a PVDF membrane and subsequently blocked with 2.5% (w/v) BSA/PBST. Membranes were incubated for 1 h at r.t. with neutravidin conjugated HRP (1:10000)

or anti-PS-NTF (1/5000). After wash with PBST for three times, blots were further incubated with appropriate secondary antibody for anti-PS-NTF (1:5000) for 1 h at r.t. After incubation, blot was washed again with PBST for 3 times and the SuperSignal West Pico kit (Pierce) was used to develop the blot.

8.7.5 Fluorescent Labeling of Mammalian Cell Lysates

For individule probe labeling, cellular lysates from mammalian cells interested (~50 μg of total proteins as determined by Bradford assay) was incubated with individule TER probe (1 μM final concentration; 5 % DMSO) in HEPES buffer (50 mM HEPES, *p*H 5.4, 1 mM EDTA, 100 mM NaCl) for 30 min at r.t. After 30 min incubation, samples were irradiated on ice for 25 min using a B100A lamp (UVP) at a distance of 5 cm. The reactions were then resuspended in 50 μL 1× SDS-loading buffer and heated for 10 min at 95 ℃; around 20 μg of protein was loaded per gel lane for separation by SDS-PAGE (12% SDS gel), then visualized by in-gel fluorescence scanning using a Typhoon 9410 Variable Mode Imager scanner.

For probe cocktails labeling, cellular lysates from mammalian cells of interest (~50 μg of total proteins) were incubated with cocktails of 11 TER probes (0.5 μM of each probe, 5 % DMSO) in HEPES buffer (50 mM HEPES, *p*H 5.4, 1 mM EDTA, 100 mM NaCl) for 30 min at r.t. After 30 min incubation, samples were irradiated on ice for 25 min using a B100A lamp (UVP) at a distance of 5 cm. After irradiation, pre-chilled acetone (0.5 mL) was added to quench the reactions, and placed at -20 ℃ for 30 min followed by centrifuged at 13,000 rpm for 10 min

at 4 ℃ to precipitate proteins. The supernatant was discarded and the pellet was washed two times with 200 μL of pre-chilled methanol. The protein pellets were then resuspended in 50 μL 1× SDS-loading buffer and heated for 10 min at 95 ℃; around 20 μg of protein was loaded per gel lane for separation by SDS-PAGE (12% SDS gel), then visualized by in-gel fluorescence.

8.7.6 Proteome Profiling of Protein Kinases

8.7.6.1 Labeling of Kinases Present in the Bacterial Proteome

To evaluate the performance of probe in a complex bacterial proteome, the labeling experiments were carried out with kinase probe and bacterial lysates. Bacterial lysates were prepared using previous reported procedures with minor modification. The culture of LB (100 mL) with a single colony containing PKA gene, were grown at 37 ℃ with shaking to reach OD_{600} = 0.8. Subsequently, IPTG (0.2 mM final concentration) was added to induce protein expression and incubated for further 6 h at 37 ℃ with shaking (230 rpm). 5 mL LB was taken out before and after inducing, respectively, and harvested by centrifugation at 4000 rpm for 10 min at 4 ℃. The resulting pellets were resuspended in lysis buffer (25 mM HEPES, *p*H 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 50 μM PMSF) and sonicated (to complete lysis, 10 rounds of 1s on and 1s off, at 25 % amplitude), before centrifugation for 10 min (1000 rpm at 4 ℃). The resulting supernatant was further centrifuged for 15 min (13,000 rpm at 4 ℃). The total protein concentration of this lysate was then quantified by Bradford's Assay (Bio-Rad USA), and stored in -20 ℃, and used for all subsequent labeling experiments.

To different amount of above bacterial lysates, probe **(**final 1 μM**)** was added and incubated for 30 min at r.t. Subsequently, the reactions were irradiated with UV light (~350 nm) for 20 min followed by click-chemistry with rhodamine azide (TER-N_3), and separated by SDS PAGE gel, and analyzed by in-gel fluorescence scanning. Results showed that stronger labeling signal was observed only for the induced lysates, and dose-dependent increase of fluorescence band intensity was observed with increased amounts of bacterial lysates.

8.7.6.2 Labeling of Kinases Present in Mammalian Proteome

For *in vitro* proteome labeling, kinase probe was added to cell lysates (50 μg) in 50 μL of HEPEs buffer at a final concentration of 1-50 μM in the presence or absence of excess competitor (a final concentration of 100 μM). Unless indicated otherwise, samples were incubated for 30 min with varying concentrations of probe at r.t. After incubation, the samples were irradiated for 20 min under UV light (~350 nm). 4 μL of the freshly premixed click chemistry reaction cocktail in PBS [rhodamine-azide (100 μM, 1.0 mM stock solution in DMSO), Tris[(1-benzyl-1H-1,2,3-triazol-4-yl) methyl]amine (TBTA)(1.0 mM, 24 mM freshly prepared stock solution in deionized water), Sodium ascorbate (1 mM, 24 mM freshly prepared stock solution in deionized water) and $CuSO_4$ (1 mM, 24 mM freshly prepared stock solution in deionized water)] was added and vortexed, then incubated for further 2 h at r.t. with gentle mixing. The reactions were terminated by the addition of pre-chilled acetone (0.5 mL), placed at -20 ^oC for 30 min and centrifuged at 13,000 rpm for 10 min at 4 ^oC to precipitate proteins. The supernatant was discarded and the pellet was washed two times with

200 μL of pre-chilled methanol. The protein pellets were then resuspended in 20 μL 1× SDS-loading buffer and heated for 10 min at 95 ^{o}C; around 20 μg of protein was loaded per gel lane for separation by SDS-PAGE (10% SDS gel), then visualized by in-gel fluorescence scanning using a Typhoon 9410 Variable Mode Imager scanner.

For *in situ* labeling, cells were grown to 80-90% confluence in 24-well plates under the conditions described above. The medium was removed, and cells were washed twice with cold PBS, treated with 0.5 mL of DMEM-containing probe (1-50 μM), with or without competitive inhibitor (100 μM). Probes were applied from DMSO stocks whereby DMSO never exceeded 1% in the final solution. The same volume of DMSO was used as a negative control. After 5 h of incubation at 37 ^{o}C/5% CO_2, the medium was aspirated, and cells were washed twice gently with PBS to remove the excessive probe followed by UV irradiation (~350 nm) for 20 min on ice. The cells were then trypsined and pelleted by centrifugation. Eventually, the cell pellets were resuspended in PBS (50 μL), homogenized by sonication, and diluted to 1 mg/mL with PBS. Probe targets were detected by click chemistry with rhodamine-azide, SDS PAGE analysis, and in-gel fluorescence scanning.

8.8 Pull-down and Mass Spectrometry Identification

8.8.1 Pull-down Assay with Aspartic Protease Probes

Two sets of pull-down assays were carried out. For individule probe pull-downs, in brief, biotinylated aspartic protease probes (10 μM final concentration) was incubated with fresh T47D cell lysates (~5 mg) for 0.5 h at 4 ^{o}C in an acidic buffer (50 mM HEPES, *p*H 5.4, 1 mM EDTA, 100 mM NaCl). In

the meanwhile, the same volumn of DMSO was incubated with T47D cell lysates as a negative control experiment. Subsequently the samples were irradiated on ice for 25 min under UV lamp. The resulting mixtures were incubated with NeutrAvidin agarose beads for further 2 h at 4 ^{o}C. After that, the supernatant was removed by centrifugation, and the beads were washed with 0.1% SDS in PBS once, and washed with PBS for 4 times. After washing, the beads were boiled in the elution buffer (200 mM Tris at *p*H 6.8 containing 400 mM DTT and 4% SDS) for 15 mins. Eluted proteins were separated on 12% SDS-PAGE gel (together with negative pull-down control). Following SDS-PAGE separation, protein bands were visualized by coomassie blue/silver staining.

In order to identify the cell targets of hydroxyethylamine pharmacophore, pull-down experiments with cocktails of 11 biotinylated aspartic protease probes were also carried out. In brief, fresh T47D cell lysates (~5 mg) were incubated with the cocktails of 11 biotinylated probes (5 μM final concentration) for 0.5 h at 4 ^{o}C in an acidic buffer (50 mM HEPES, *p*H 5.4, 1 mM EDTA, 100 mM NaCl). Similarly, the same volumn of DMSO was incubated T47D lysates as a negative control. Subsequently the reactions were irradiated on ice for 25 min under UV lamp and quenched by pre-chilled acetone (10 mL) and placed at -20 ℃ for 30 min. The mixtures were then centrifuged at 13,000 rpm for 10 min at 4 ℃ to precipitate the proteins. The supernatant was discarded and the pellet was washed two times with 10 mL of pre-chilled methanol. The protein pellets were resolubilized into 0.1% SDS-PBS buffer. Next, The solutions were incubated with

NeutrAvidin agarose beads for 2 h at 4 ^{o}C. The supernatant was removed by centrifugation, and the beads were washed with 0.1% SDS in PBS once, and washed with PBS for 4 times. After washing, the beads were boiled in the elution buffer (200 mM Tris at *p*H 6.8 containing 400 mM DTT and 4% SDS) for 15 mins. Eluted proteins were separated on 12% SDS-PAGE gel (together with negative pull-down control). The protein bands were visualized by coomassie blue/silver staining.

8.8.2 Pull-down Assay with Kinase Probes

To identify *in vitro* and *in situ* targets of dasatinib/staurosphorine in cancer cell lines, pull-down experiments were carried out followed by immunoblotting and MS/MS identification experiments. Fresh cell lysates were prepared as described previously. Protein concentrations were determined according to Bradford. *In vitro* pull-down experiment: cellular lysates (5 mg) were supplemented with 200 μL 5× HEPEs buffer (125 mM HEPES at *p*H 7.5, 750 mM NaCl, 10 mM $MgCl_2$), the reaction volume will be adjusted to 1 mL with milli-Q water. Subsequently, a solution of probes (10 μM, 10 μL of a 1 mM stock soloution) was added, and equilibration was carried out for 2 h at 4 ^{o}C. The reaction mixture was then irradiated for 20 min using a 4W UV lamp. Subsequently, the reaction was reacted with biotin-azide (or biotin cleavable azide) by click chemistry under the conditions described before, acetone precipitated, and resolubilized in 0.1% SDS in PBS with brief sonication. This resuspended sample was then incubated with avidin-agarose beads (100 μL/mg protein) overnight at 4 ^{o}C. After centrifugation, supernatant were removed and the

beads were washed with 0.1% SDS once and PBS for 4 times. After washing, the beads were boiled in 1 × SDS loading buffer (200 mM Tris *p*H 6.8, 400 mM DTT, 8% SDS) for 15 min. If cleavable biotin azide was used, the beads were incubated for 1 h at r.t. with elution buffer (25 mM $Na_2S_2O_4$, 100 mM NH_4HCO_3, 0.5% SDS).

For *in situ* pull-down, briefly, probes (20 μM) were directly added in live cells followed by incubation for 5 h. DMSO should never exceed 1% in the final solution. After 5 h of incubation at 37 ^oC/5% CO_2, the medium was aspirated, and cells were washed twice gently with PBS to remove the excessive probe followed by UV irradiation (~350 nm) for 20 min on ice. The cells were then trypsined and pelleted by centrifugation. Eventually, the cell pellets were resuspended in PBS (50 μL), homogenized by sonication, and diluted to 1 mg/mL with PBS. The labeled lysates were then subjected to click reaction with biotin azide, and all the subsequent experiments just follow above *in vitro* pull-down procedure.

8.8.3 Mass Spectrometric Analysis

For MS/MS identification, proteins from pulled-down fractions were separated on 10% SDS-PAGE gel followed by silver staining or coomassie staining. Each gel lanes were cut into multiple slices. Subsequently, trypsin digestion (using In-Gel Trypsin Digestion Kit, Pierce) and peptide extraction (with 50% acetonitrile and 1% formic acid) were carried out. The samples were then dried in vacuo and stored at -20 ^oC. LCMS/MS analysis was performed using an LTQ-FT ultra mass spectrometer (Thermo Electron, Germany) coupled with an online Shimadzu UFLC system utilizing nanospray ionization. Peptides were first enriched with a Zorbax 300SB C18

column (5 mm × 0.3 mm, Agilent Technologies), followed by elution into an integrated nanopore column (75 μm × 100 mm) packed with C18 material (5 μm particle size, 300 Å pore size). Mobile phase A (0.1% formic acid in H_2O) and mobile phase B (0.1% formic acid in acetonitrile) were used to establish the 90-min gradient, comprising 3 min of 0-5% B, then 52 min of 5-30% B, followed by 12 min of 30-60% B; maintained at 80% B for 8 min before re-equilibrating at 5% B for 15 min. Sample was injected into the MS with an electrospray potential of 1.8 kV without sheath and auxiliary gas flow, ion transfer tube temperature of 180 ^{o}C, and collision gas pressure of 0.85 m Torr. A full-survey MS scan (350-2000 m/z range) was acquired in the 7-T FT-ICR cell at a resolution of 100,000 and a maximum ion accumulation time of 1000 ms. Precursor ion charge-state screening was activated. The linear ion trap was used to collect peptides where the 10 most intense ions were selected for collision-induced dissociation (CID) in MS2, performed concurrently with a maximum ion accumulation time of 200 ms. Dynamic exclusion was activated for this process, with a repeat count of one and exclusion duration of 30 s. For CID, the activation Q was set at 0.25, isolation width (m/z) 2.0, activation time 30 ms, and normalized collision energy 35%. The Extract-Msn (version 4.0) program found in Bioworks Browser 3.3 (Thermo Electron, Germany) was used to extract tandem MS spectra in the data format from the raw data of the LTQ-FT ultra. These data files were then converted into MASCOT generic file format using an in-house program. Intensity values and fragment ion m/z ratios were not manipulated. These data were used to obtain protein identities by searching against the corresponding database by means of an in-house

MASCOT server (version 2.2.03, Matrix Science, Boston, MA). The search was limited to a maximum of two missed trypsin cleavages, #13C of 2, mass tolerances of 10 ppm for peptide precursors and 0.8 Da for fragment ions. Only proteins with a MOWSE score higher than 40, corresponding to $p < 0.05$, were considered significant. The peptide/protein lists obtained were exported to an html file.

8.9 Cell Culture and Preparation of Lysates

8.9.1 Parasite Cultures

Plasmodium falciparum 3D7 strain was used in this study. Parasites were cultured in RPMI medium 1640 (Invitrogen, USA) supplemented with 0.29225 g *L*-glutamine, and 0.05 g hypoxanthine dissolved in 1 mL 1M NaOH. Parasites were synchronized twice 16 hrs apart at ring stage using 2.5% sorbitol. Cultures were stored at 37 ^{o}C after gassing with a 5% CO_2, 3% O_2 and 92% N_2 gas mixture and their hematocrit maintained at 2.5%. Parasitized red blood cells were treated with 0.1% Saponin in PBS for 15 min at r.t. with shake. The parasites were centrifuged and the pellet was resuspended in 50 mM Acetate *p*H 5.5, 1 mM DTT, 0.1% NP-40. Lysates were then centrifuged at 15,000 rpm for 5 min and split into a pellet fraction and a soluble fraction.

8.9.2 Mammalian Cell Cultures

Cell lines were obtained from the National Cancer Institute Developmental Therapeutics Program (NCI60 cell). γ-30, HEK293T, HepG2, OVCAR-3, T47D, MDA-MB-435, HCT116 and MCF7 cells were grown in DMEM (Invitrogen,

Carlsbad, CA) containing 10% heat-inactivated fetal bovine serum (FBS; Gibco Invitrogen), 100 U/mL penicillin and 100 μg/mL streptomycin (Thermo Scientific, Rockford, IL) and maintained in a humidified 37 ^{o}C incubator with 5% $CO_{2.}$ K562 cells was grown in RPMI 1640 medium (Invitrogen Corporation) supplemented with 10% heat-inactivated fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin and maintained in a humidified 37 ^{o}C incubator with 5% $CO_{2.}$

Mammalian lysates were prepared as following: fresh cell pellets were thawed on ice by suspension in 100 μL of lysis buffer (50 mM HEPES, *p*H 5.4, 1 mM EDTA, 100 mM NaCl, 10% Glycerol, 0.1% NP-40), incubated for 15 min on ice, and then spun for 20 min at 13,000 rpm at 4 ^{o}C. The cleared supernatant was decanted to an Eppendorf tube and kept at 4 ^{o}C. Protein concentration was quantified by Bradford assay (Bio-Rad, USA), aliquoted and stored in -80 ℃, and used for subsequent screening on array.

8.9.2.1 Preparation of Membrane Fractions of γ-30 cells

The cellular lysates were centrifuged at 800 rpm for 10 min to remove nuclei and cell debris. The postnuclear supernatant was further fractionated by sequential centrifugation at 100,000 rpm for 1 h .The collected pellet were resuspended in HEPEs buffer with 1% CHAPSO followed by centrifugation at 100,000 rpm for further 1 h. The supernant was collected with addition of 20% glycerol and stored at -80 ^{o}C until use. All centrifugation steps were carried out at 4 ^{o}C.

8.10 Western Blotting

For immunoblotting experiments, samples were separated on 10% SDS-PAGE gel and further transferred to PVDF membranes. Membranes were then blocked with 3% BSA in TBST (0.1% Tween in TBS) for 1 h at r.t. After blocking, membranes were incubated with the respective primary antibody at r.t., or overnight at 4 ^{o}C. The membranes were then washed with TBST (0.1% Tween in TBS) for four times (10 min/time) and incubated with an appropriate secondary antibody [anti-mouse conjugated HRP (1/3000), or anti-rabbit conjugated HRP (1/5000)]. After secondary incubation, blots were washed again with TBST for 4 times before the development with Dura (Pierce).

8.11 Cell Proliferation Assay

Cell viability was determined using the XTT colorimetric cell proliferation kit (Roche) following manufacturer's guidelines. Briefly, cells were grown to 20-30% confluence (since they will reach ~90% confluence within 48 to 72 h in the absence of drugs) in 96-well plates under the conditions described above. The medium was aspirated, and then washed with PBS, and treated in duplicate with 0.1 mL of the medium containing different concentrations of probes (1-50 μM) or drugs (1-50 μM, as a positive control). Probes were applied from DMSO stocks whereby DMSO never exceeded 1% in the final solution. The same volume of DMSO was used as a negative control. Fresh medium, along with two probes and dasatinib, were added every 24 h. However, for K562 cell only one time probes and drug were added. After a total treatment time of 72 h, proliferations were assayed using the XTT colorimetric cell

proliferation kit (Roche) following manufacturer's guidelines (read at 450 nm). Data represent the average (s.d. for two trials).

8.12 Protein Expression and Purification

All the genes including the chicken c-Src (residues 251-533), c-SrcY416C, c-SrcY527C, c-SrcY416C527C and c-SrcT338M genes were cloned into pET-28a vector (Novagen) and modified to yield an N-terminal, tobacco etch virus protease cleavable $(His)_6$ tag, a plasmid containing tyrosine phosphatase YopH cloned into pCDFDuet-1 (Novagen). All the c-Src mutants were generated by site-directed mutagenesis and confirmed by sequence analysis.

All chicken c-Src (residues 251-533) variants were co-expressed in *E. Coli* and purified as previously described.[136] Firstly, the two plasmids containing the c-Src (pET-28a) and the YOPH (pCDFDuet) were co-transformed into *Escherichia coli* BL21DE3 cells and plated on LB agar with kanamycin (50 mg/mL)/streptomycin (50 mg/mL) and grown overnight at 37 ^{o}C. The colonies from the plates were then resuspended in the expression media (LB, 50 mg/mL/streptomycin, 50 mg/mL kanamycin). Cultures were grown to an OD_{600} of 1.2 at 37 ^{o}C, cooled for 1 h with shaking at 18 ^{o}C prior to induction for 16 h at 18 ^{o}C with 0.2 mM IPTG. Cells were harvested by 10 min centrifugation at 4000 rpm at 4 ^{o}C and resuspended in 50 mM sodium phosphate (*p*H 8.0), 300 mM NaCl, 10 mM imidazole for immediate purification. Cells were lysed by sonication at 4 ^{o}C. Insoluble protein and cell debris was spinned down through a 30-min centrifugation at 13,000 rpm at 4 ^{o}C. The supernatant was loaded onto a Ni NTA affinity column which was prior washed three

times with lysis buffer. The loaded column was incubated 3 h at 4 ^{o}C followed by removing of flow through. The beads were washed ten times with washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazole). Finally, protein was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM Imidazole). The eluted fractions were analyzed by SDS-PAGE gel, and buffer was exchanged to PBS by dialysis.

The expression of the full length catalytically inactive Src (srcK295M) was based on previous procedure with a little modification.[149] The coding region for amino acid residues 83-533 of chicken c-src gene was subcloned into pRSET vector (Invitrogen) with BamHI and HindIII sites on the 5`-and 3`-ends, respectively. The constructed plasmid pRSET-dsrc was co-transformed into BL21(DE3) strain of *E. coli* with YOPH(pCDFDuet) in the presence of ampicillin (50 mg/mL) and streptomycin (50 mg/mL). The cells were grown in LB media at 37 ^{o}C to reach OD_{600} = 0.5 on a 1 L scale and then induced with 0.2 mM IPTG at 25 ^{o}C for 14 h. After harvesting and lysis, the proteins were purified followed above described procedure.

8.13 Site-directed Mutagenesis of c-Src to SrcT338M

Threonine 338 residues of c-Src were mutated to Methionine by using QuickChangeTM XL Site-directed Mutagenesis Kit (Strategene) with upstream primer 5'-AGCCCATCTACATCGTCATGGAGTACATGAGCAAGGGG-3' and downstream primer 5'-CCCCTTGCTCATGTACTCCATGACGATGTAGATGGGC T-3'. Primers were designed to introduce the desired mutation site. Appropriate primer sequences were chosen to ensure that the T_m melting temperature was above

78 ^{o}C. Codon usage of amino acid was calculated to ensure the desired protein would be well expressed in *E. coli* cells. PCR amplification was first carried out to amplify the whole sequence of gene with the primer containing desired mutation site. The PCR product was confirmed with the desired PCR band with gel analysis. Dpn I enzyme was used to digest the template which contains the methylation site. PCR product was then transformed into top 10 competent cells and cultured in LB+Kan agar plates at 37 ^{o}C. One of the clonies from the agar plate was picked up and cultured in 5 ml LB+ Amp for overnight. Miniprep was used to extract the genes and DNA sequencing was carried out to confirm the presence of mutation site and correct frame of gene.

8.14 Transient Transfection

Briefly, vectors were transfected with Endofectin^{TM} transfection reagents (Genecopoeia) at 80% confluence in CHOK1 cell. After 8 h, the medium was changed to fresh grown medium. And the cells were incubated further 24 h in 37 ^{o}C /5% CO_2 incubator. To generate protein lysates, cells were washed twice with cold phosphate buffered saline (PBS), and lysated by lysis buffer (25 mM HEPES at *p*H 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 0.1% NP-40) for 30 min on ice. The cells were then resuspended and collected by centrifugation. Protein concentration was determined by the Bradford assay.

8.15 Cellular Imaging

HepG2 cells were seeded onto 24-well plates containing sterile glass coverslips

and grown until 70-80% confluence. Further, cells were treated with 0.5 mL of DMEM with 20 μM probes or DMSO. After 5 h, the medium was removed, and cells were washed twice with PBS gently followed by UV irradiation (~350 nm) for 10 min. The cells were then fixed for 15 min at r.t. with 3.7% formaldehyde in PBS , and washed twice with cold PBS again, and permeabilized with 0.1% Triton X-100 in PBS for 10 min at r.t. The cells were then blocked with 2% BSA in PBS for 30 min at r.t., and then washed twice with PBS. Cells were subsequently treated with a freshly premixed click chemistry reaction solution in a 100 μL volume at final concentrations of the following reagents: 1 mM of $CuSO_4$, 1 mM of TCEP, 100 μM of TBTA, and 10 μM of rhodamine-azide in PBS for 2 h at r.t. with vigorous shaking. Cells were washed with PBS three times. If the background was still very high, the longer washing time was necessary. For colocalizing *in situ* targets of dasatinib analogues with c-Src, cells were further incubated with anti-c-Src primary antibodies (1:50) for 1 h at r.t. (or overnight at 4 ^{o}C), and washed twice with PBS. The cells were then incubated with FITC conjugated anti-mouse IgG (1:500) for 1 h, washed again. Cells were stained with 1 μg/mL Hoechst for 10 min at r.t., washed again before mounting.

8.16 Molecular Modelling Experiments

In order to explore the binding mode of probes with kinases, molecular docking studies were carried out using Autodock Vina software.[133] The Protein structures were obtained from the PDB databank (Abl PDB ID 2GQG[132], c-Src PDB ID 2SRC[131] and PDB ID 1STC[150]). Explicit hydrogen atoms were added, all water molecules were then deleted. The peptide ligand was removed and the protein structure was processed

using AutoDock Tools. All compounds were prepared for docking using AutoDock Tools to assign AD4 atom types, calculate Gasteiger charges, and set all rotatable bonds as active torsions. The ligand was docked into the protein using AutoDock Vina (version 1.11, The Scripps Research Institute). The exhaustiveness parameter was set to 100 (default = 8, linear scale); all other default settings were used. The macromolecule molecular surface and secondary structure were displayed by PyMol (version 0.99, DeLano Scientific LLC.

8.17 Cell Permeability Assays

The assay was carried out as previously described.[151] MDCK (Madin-Darby canine kidney) cells were used for testing the cell permeability. MDCK was seeded with 600,000~700,000 cells per cm^2 (0.11 cm^2 per well insert, Millipore® #PSHT004R1) and cultured for 3 days before test. When cultures are confluent, remove medium and rinse inserts with Hanks' Balanced Salt Solution (HBSS, Gibco® #14025). Transport assay donor solutions consisted of 100 μM wide type drugs, probes and caffeine in transport medium containing 60 μM Lucifer Yellow (LY) and 1% DMSO. Transport assays were conducted using 75 μL of apical (AP) donor solution and 235 μL of basolateral (BL) acceptor solution (transport medium, *p*H 7.4, following manufacture descriptions of Millipore® #PSHT004R1). Monolayer was incubated with donor and acceptor solutions for 60 min at 37 ℃, 95% humidity. After 60 min, the donor and acceptor solutions of each compound was collected and quantified by HPLC. Lucifer Yellow (LY) was quantified using a fluorescence 96-well plate reader (BioTek® Synergy 4 fluorescence plate reader at E_x = 485 nm

and E_m = 539 nm). Papp (apparent permeability) values were calculated according to the following equation:

$$Papp = \left(\frac{dQ}{dt}\right) \times \frac{1}{Co} \times \frac{1}{A}$$

where dQ/dt is the permeability rate, C_0 is the initial concentration in the donor compartment, and *A* is the surface area of the filter. Permeability rates were calculated by plotting the percent of initial AP drug mass (peak area) found in the BL compartment versus time and determining the slope of the line. Lucifer yellow (LY) results were used as an internal control for each monolayer to verify tight junction integrity during the entire assay period. Accordingly, LY P*app* values were quantified from 60 min basolateral samples after background subtraction. As a quality control, results from MDCK monolayer with LY P*app* > 30 nm/s were not used.

Table 8.2 The results of cell permeability assay.

	Positive control	**Negative control**	**Staurosphorine**	**STS-1**
Caffeine / 50 μM	+	-	-	-
Lucifer Yellow / 60 μM	+	+	+	+
Compound / 100 μM	-	-	+	+
DMSO	1%	1%	1%	1%
P*app*/nm.s^{-1}	762	n.d.	320	142

n.d.: not determined.

Chapter 9

Conclusion Remarks

In summary, we have described a series of technologies that span combinational chemistry and microarray-facilitated high-throughput enzyme profiling platform. Two classes of proteins which are aspartic proteases and protein kinases have been separately studied in this thesis with high throughput technologies.

The first objective of this study was to develop a solution-cum-solid phase strategy for synthesis of C_2-symmetric diols inhibitors targeting HIV-1 protease. It was found that both symmetric and asymmetric diol-containing small molecules as potential aspartic protease inhibitors could be synthesized and evaluated in a very short time. Compared with Wong's strategy, ours is more rapid and efficient. This is because all compounds including symmetric and asymmetric structures were generated by solid-phase synthesis rather than solution-phase synthesis which is quite time-consuming. The results have shown that the clear advantage in our strategy lies in its feasibility for solid-phase synthesis of asymmetric diols which may be quite useful for combating drug resistance of HIV-1 protease in future. Based on the results, it appears that the solution-cum-solid phase synthesis is highlighted by its robustness, product purity and diversification of a common core in a simple way. Thus, our chemical strategy may be readily expanded in future for the construction of much bigger compound libraries which may lead to candidate compounds with improved biological activities. It should find useful application in high-throughput drug discovery for different diseases.

The second objective in this thesis was to discover the potent inhibitors and

affinity-based probes for functional profiling of plasmepsins (PMs) in intraerythrocytic malaria parasites. After subsequent *in situ* screening of parasites with these probes, it was found that compound **G16** showed good inhibition against all four PMs and parasite growth in infected RBCs. Molecular modelling indicated that this inhibitor bound to the active site of the plasmepsins tested, as originally designed. Our results indicate the feasibility of using affinity-based probes approaches for the identification of inhibitors of less-characterized enzymes (such as HAP) and inhibitors of multiple targets. We anticipate that these new chemical tools should facilitate the discovery of unknown parasite biology and new anti-malarial drugs.

The third and fourth objectives were to identify affinity-based probes for aspartic proteases, for example, γ-secretase and cathepsin D, using small molecule microarray (SMM) platform. By incubating dye-labled, interest proteins with inhibitor library onto a protease-sensitive glass surface, a series of distinct binding profiles between inhibitor library and interest proteins or mammalian cell lysates were obtained. Some potent and specific affinity-based probes against aspartic proteases were subsequently generated from potent inhibitor hits by "click" chemistry. Our results demonstrated that SMM platform was able to sensitively report activity-based binding events of aspartic proteases towards their small molecule inhibitors. Previous small molecule microarray strategies have mainly relied on incubation with a purified protein of interest. In our case, since it is notoriously difficult to isolate some aspartic proteases, we carried out SMM screening which allows the direct detection of interactions between small molecules and target proteins expressed in mammalian cells without prior purification. This is because the cellular lysate has the optimum condition for proteins living compared with recombinant proteins. It was found that potential probes for aspartic proteases could be rapidly identified by using SMM platform.

Comparing with the development of traditional probes based on the known inhibitor, however, our approach does not require known inhibitors of a protein target and minimizes risks involved in the loss of protein-binding property of these inhibitors due to linker introduction. The results indicate that our strategy would provide comprehensive insight into the behaviors of virtually any cluster of enzymes. The functional profile obtained from inhibitor screening contributes towards the understanding of the enzymes and proteins and the development of predictive models for their functional perturbation.

In the fifth and sixth objectives, we have designed and synthesized two cell permeable, photo-affinity based probes targeting protein kinases based on well-known activity-based protein profiling (ABPP) approach. In order to maximally retain drug's native activity (Dasatinib^{TM} and Staurosphorine), minimal structure modifications have been done in the synthesis of each probe. With the aim to comprehensively study the cellular targets of dasatinib/staurosphorine, we performed pull-down/mass spectrometric analysis in large scale manners. Consequently, some putative cellular off-targets of dasatinib and staurosphorine have been identified using our probes in live cells, respectively. The preliminary results demonstated that our probes enabled to proteomic profile the cellular on/off-targets of dasatinib/staurosphorine in live cells. We predict that our chemical proteomic approach should be broadly useful for off-target identification against a number of existing drugs, and it may provide a powerful tool for biomarker discovery in future.

Considering that research in proteomic area is still at an infant stage, the following aspects are recommended for future research. Firstly, due to lack of experimental condition, we were unable to detect previously predicted but unidentified aspartic proteases from the malaria proteome. Thus, probe **G16** might

have further targets, other than the four plasmepsins, in the malaria. Identification of some other further targets for probe **G16** will be an interesting area for future work. In addition, it is necessary to develop other chemical probes for profiling of malaria.

Secondly, for the microarray work, it should be pointed out that proteins or cell lysates may lose activity after they are labeled by dyes. It will cause a lot of non-specific bindings between proteins and inhibitors. Thus, it is significant to develop other technologies which can be used to proteome profile the interactions between inhibitors and the binding proteins in native environment.

Thirdly, several aspartic proteases related to human diseases have been investigated in this thesis. Studies and explorations of some other aspartic proteases which are related to cancers will be our interesting objectives in future.

Fourthly, our cell permeable, photo-affinity probes can sucessfully label a wide of kinases in live cells. Some of the protein hits, for the first time, were identified as putative cellular targets of dasatinib and staurosphorine in our studies. They may be considered potential off-targets of dasatinib and staurosphorine in cellular environment, which should provide important implication for studing the therapeutic functions of dasatinib and staurosphorine. Thus, proper follow-up studies and validation experiments will be performed in future.

In summary, these high-throughput chemical approaches developed in this thesis should provide useful platforms for enzyme activity profiling and drug discovery.

Chapter 10

Reference

1 Venter, J. C.; Adams, M. D.; Myers, E. W. *Science* **2001**, *291*, 1304.

2 Creighton, T. E. Proteins: structure and molecular properties, 2nd edn. Freeman, New York **1993**.

3 Sun, H.; Chattopadhaya, S.; Wang, J.; Yao, S. Q. *Anal. Bioanal. Chem.* **2006**, *386*, 416.

4 a) Uttamchandani, M.; Lu, C. H. S.; Yao, S. Q. *Acc. Chem. Res.* **2009**, *42*, 1183; b) Kalesh, K. A.; Shi, H.; Ge, J.; Yao, S. Q. *Org. Biol. Chem.* **2010**, *8*, 1749; c) Uttamchandani, M.; Yao, S. Q. *Curr. Pharm. Des.* **2008**, *14*, 2428; d) Kalesh, K. A.; Yang, P. -Y.; Srinivasan, R.; Yao, S. Q. *QSAR Comb. Sci.* **2007**, *26*, 1135.

5 a) Jorg, E.; Ulrich, H.; Frederic, C.; Bruno, M.; Bernd, G. *Current Pharmaceutical Design.* **2007**, *13*, 271; b) Barrett, A. J.; Rawlings, N. D.; Woessner, J. F. *Handbook of Proteolytic Enzymes.* **2004**; 2nd edition.

6 Tigerstedt, R.; Bergman, P. G.; Niereund, K. S. *Arch Physiol.* **1898**, *8*, 223.

7 Rochefort, H.; Capony, F.; Garcia, M.; Cavailles, V.; Freiss, G.; Chambon, M. *J. Cell. Biochem.* **1987**, *35*, 17.

8 a) Hardy, J.; Selkoe, D. J. *Science*. **2002**, *297*, 353; b) De, S. B. *Neuron*. **2003**, *38*, 9.

9 Boss, C.; Richard-Bildstein, S.; Weller, T.; Fischli, W.; Meyer, S.; Binkert, C. *Curr. Med. Chem.* **2003**, *10*, 883.

10 a) Gallo, R. C.; Sarin, P. S.; Gelmann, E. P.; Robert-Guroff, M.; Richardson, E.; Kalyanaraman, V. S. *Science.* **1983**, *220*, 865; b) Barre-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J. *Science.* **1983**, *220*, 868.

11 a) Dunn, B. M. *Chem. Rev*. **2002**, *102*, 4431; b) Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790; c) Veerapandian, B.; Cooper, J. B.; Sali, A.; Blundell, T. L.; Rosati, R. L.; Dominy, B. W.; Damon, D. B.; Hoover, D. *J. Protein. Sci.* **1992**, *1*, 322.

12 James, M. N. G.; Sielecki, A. R.; Jurnak, F. A.; McPherson, A. E. *John Wiley & Sons*, New York, **1987**, *3*, 413.

13 Umezawa, H.; Aoyagi, T.; Morishima, H.; Matsuzaki, M.; Hamada, M. *J. Antibiot.* **1970**, *23*, 259.

14 Marciniszyn, J.; Hartsuck,, J. A.; Tang, J. *J. Biol. Chem.* **1976**, *251*, 7088.

15 Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 1181.

16 Blume-Jensen, P.; Hunter, T. *Nature*. **2001**, *411*, 355.

17 Cohen, P. *Nat. Rev., Drug Discov*. **2002**, 1, 309.

18 Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science.* **2002**, *298*, 1912.

19 Joseph, A. A. *Chem. Rev.* **2001,** *101*, 2271.

20 Krupa, A.; Srinivasan, N. *Genome. Biol.* **2002**, *3*, 1.

21 Sawyers, C. L. *Genes. Dev.* **2003**, *17*, 2998.

22 Zhang, J.-M.; Yang, P. L.; Gray, N. S. *Nature Reviews. Cancer*. **2009**, *9*, 28.

23 a) Dorsey, J. F.; Jove, R.; Kraker, A. J. *J. Cancer Res*. **2000**, *60*, 3127; b) Shah, N. P.; Tran, C.; Lee, F. Y.; Chen, P.; Norris, D.; Sawyers, C. L. *Science*. **2004**, *305*, 399.

24 a) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Borzilleri, R. M. *J Med Chem*. **2004**, *47*, 6658; b) Kantarjian, H.; Jabbour, E.; Grimley, J.; Kirkpatrick, P. *Nat Rev Drug Discov*. **2006**, *5*, 717.

25 a) Druker, B. J.; Tamura, S.; Buchdunger, E.; Ohno, S.; Segal, G. M.; Fanning, S.; Zimmermann, J.; Lydon, N. B. *Nature Med.* **1996**, *2*, 561; b) Druker, B. J.; Lydon, N. B. *J. Clin. Invest.* **2000**, *3*. 105.

26 Ohren, J. F.; Chen, H. F.; Pavlovsky, A.; Whitehead, C.; Zhang, E.; Kuffa, P.; Yan, C. H.; McConnell, P. Spessard, C.; Banotai, C.; Mueller, W. T.; Delaney, A.; Omer, C.; Leopold, J. S.; Dudley, D. T.; Leung, I. K.; Flamme, C.; Warmus, J.; Kaufman, M.; Barrett, S.; Tecle, H.; Hasemann, C. A. *Nature Struct. Mol. Biol*. **2004**, *11*, 1192.

27 a) Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. *Science.* **2005**, *308*, 1318; b). Kwak, E. L.; Sordella, R.; Bell, D. W.; Heymann, N. G.; Okimot, R. A.; Brannigan, B. W.; Harris, P. L.; Driscoll, D. R.; Fidias, Panos.; Lynch, T. J.; Rabindran, S. K.; McGinnis, J. P.; Wissner, A.; Sharma, S. V.; Isselbacher, K. J.; Settleman, J.; Haber, D. A. *Proc. Natl Acad. Sci. USA*. **2005**, *102*, 7665.

28 a) Rabindran, S. K. *Cancer Res.* **2004**, *64*, 3958; b) Kobayashi, S. *Cancer Res.***2005**, *65*, 7096.

29 Hu, Y.; Uttamchandani, M.; Yao, S. Q. *Comb. Chem. High Throughput Screen*. **2006**, *9*, 203.

30 Merrifield, R. B. *J. Am. Chem. Soc*. **1963**, *85*, 2149.

31 Bunin, B. A.; Ellman, J. A. *J. Am. Chem. Soc*. **1992**, *114*, 10997.

32 Plunkett, M. J.; Ellman, J. A. *J. Org. Chem*. **1995**, *60*, 6006.

33 Fivush, A. M.; Willson, T. M. *Tetrahedron Lett.* **1997**, *38*, 7151.

34 a) Sharpless, K. B.; Manetsch, R. *Exp. Opin. Drug Discovery.* **2006**, *1*, 525; b) Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today***. 2003**, *8*, 1128.

35 Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.

36 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.

37 Wu, P.; Fokin, V. V. *Aldrich. Acta.* **2007**, *40*, 7.

38 Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. *Proc. Natl. Acad. Sci. USA.* **2000**, *97*, 9367.

39 Szczepankiewicz, B. G. *J. Am. Chem. Soc.* **2003**, *125*, 4087.

40 Birk, A.; Lin, Y.-C.; Elder, J.; Wong, C.-H. *Chem. Biol.* **2002**, *9*, 891.

41 Liu, K.; Shi, H.; Xiao, H.; Chong, A. G. L.; Bi, X.; Chang, Y. T., Tan, K.; Yada, R. Y.; Yao, S. Q. *Angew. Chem., Intl. Ed.* **2009**, *48*, 8293.

42 a) Srinivasan, R.; Uttamchandani, M.; Yao, S. Q. *Org. Lett.* **2006**, *8*, 713; (b) Xie, J.; Seto, C. T. *Bioorg. Med. Chem.* **2007**, *15*, 458.

43 a) Hu, M.; Li, J.; Yao, S. Q, *Org. Lett.* **2008**, *10*, 5529; b) Wang, J.; Uttamchandani, M.; Li, J.; Hu, M.; Yao, S. Q. *Org. Lett*. **2006**, *8*, 3821.

44 Ng, S. L.; Yang, P. -Y.; Chen, K. Y. -T.; Srinivasan, R.; Yao, S. Q, *Org. Biomol. Chem.* **2008**, *6*, 844.

45 Kalesh, K. A.; Sim, S. B. D.; Wang, J.; Liu, K.; Lin, Q.; Yao, S. Q. *Chem. Commun.* **2010**, *46*, 1118.

46 Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. –H.; Kolb, H. C. *Angew. Chem., Intl. Ed*. **2005**, *44*, 116.

47 Sears, P.; Wong, C, -H. *Cell. Mol. Life Sci.* **1998**, *54*, 223.

48 Wu, C.-Y.; Jan, J. T.; Ma, S. -H.; Kuo, C. J.; Juan, H. F.; Cheng, Y. S.; Hsu, H. H.; Huang, H. -H.; Wu, D.; Brik, A.; Liang, F. -S.; Liu, R. S.; Fang, J. -M.; Chen, S. -T.; Liang, P. -H.; Wong, C. -H. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 10012.

49 Krasinski, A.; Radic, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C, *J. Am. Chem. Soc*. **2005**, *127*, 6686.

50 Lee, L. V.; Mitchel, M. L.; Huang, S. J.; Fokin, V. V.; Sharpless, K. B.; Wong, C. -H. *J. Am. Chem. Soc*. **2003**, *125*, 9588.

51 a) Fokin, V. V.; Sharpless, K. B.; Wong, C.-H. *ChemBioChem*. **2003**, *4*, 1246; b) Brik, A.; Alexandratos, J.; Lin, Y.-C.; Elder, J. H.; Olson, A. J.; Wlodawer, A.; Goodsell, D. S.; Wong, C.-H. *ChemBioChem*. **2005**, *6*, 1167.

52 Yang, P. -Y.; Wu, H.; Lee, M. Y.; Xu, A.; Srinivasan, R.; Yao, S. Q, *Org. Lett.* **2008**, *10*, 1881.

53 Wu, C. -Y.; Chang, C. -F.; Chen, J. S.-Y.; Lee, S. -T.; Wong, C. -H.; Lin, C.-H. *Angew. Chem., Int. Ed.* **2003**, *42*, 4661.

54 Wu, C. -Y.; King, K. -Y.; Kuo, C. -J.; Fang, J. -M.; Wu, Y. -T.; Ho, M. -Y.; Liao, C. -L.; Shie, J. -J.; Liang, P. -H.; Wong, C.-H. *Chem. Biol.* **2006**, *13*, 261.

55 Best, M. D.; Brik, A.; Chapman, E.; Lee, L. V.; Wong, C.-H. *ChemBioChem* **2004**, *5*, 811.

56 Brik, A.; Wu, C.-Y.; Wong, C.-H. *Org. Biomol. Chem.* **2006**, *4*, 1446.

57 Fodor, S. P.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science*. **1991**, *251*, 767.

58 http://www.affymetrix.com; Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science*. **1995**, *270*, 467.

59 a) MacBeath, G.; Koehler, A. N.; Schreiber, S. L. *J. Am. Chem. Soc.* **1999**, *121*, 7967; b) MacBeath, G.; Schreiber, S. L. *Science*. **2000**, *289*, 1760.

60 a) Kuruvilla, F. G.; Shamji, A. F.; Sternson, S. M.; Hergenrother, P. J.; Schreiber, S. L. *Nature*. **2002**, *416*, 653; b) Uttamchandani, M.; Walsh, D. P.; Khersonsky, S. M.; Huang, X.; Yao, S. Q.; Chang, Y. T. *J. Comb. Chem.* **2004**, *6*, 862; c) Hergenrother, P. J.; Depew, K. M.; Schreiber, S. L. *J. Am. Chem. Soc.* **2000**, *122*, 7849; d) Seeman, D. B.; Park, S. B.; Koehler, A. N.; Schreiber, S. L. *Angew. Chem., Intl. Ed.* **2003**, *42*, 2376.

61 a) Köhn, M.; Wacker, R.; Peters, C.; Schröder, H.; Soulère, L.; Breinbauer, R.; Niemeyer, C. M.; Waldmann, H. *Angew. Chem., Intl. Ed.* **2003**, *42*, 5830; b) Soellner, M. B.; Dickson, K. A.; Nilsson, B. L.; Raines, R. T. *J. Am. Chem. Soc.* **2003**, *125*, 11790.

62 a) Evans, M. J.; Cravatt, B. F. *Chem. Re*v. **2006**, *106*, 3279; b) Uttamchandani, M.; Li, J.; Sun, H.; Yao, S. Q. *ChemBioChem.* **2008**, *9*, 667.

63 Leung, D.; Hardouin, C.; Boger, D. L.; Cravatt, B. F. *Nut. Biotechnol.* **2003**, *21*, 687.

64 Greenbaum, D.; Baruch, A.; Hayrapetian, L.; Darula, Z.; Burlingame, A.; Medzihradszky, K. F.; Bogyo, M. *Mol. Cell. Proteomics.* **2002**, *1*, 60.

65 a) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. *Proc. Natl. Acad. Sci*. *USA* **1999**, *96*, 14694; b) Kidd, D.; Liu, Y.; Cravatt, B. F. *Biochemistry* **2001**, *40*, 6107; c) Patricelli, M. P.; Giang, D. K.; Stamp. L. M.; Burbaum, J. *J. Proteomics.* **2001**, *1*, 1067; d) Jessani, N.; Liu, Y.; Humphrey, M.; Cravatt, B. F. *Proc. Natl. Acad. Sci. USA.* **2002**, *99*, 10335.

66 Jessani, N.; Cravatt, B. F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 54.

67 a) Adam, G. C.; Cravatt, B. F.; Sorensen, E. *J. Chem. Biol.* **2001**, *8*, 81; b) Adam, G. C.; Cravatt, B. F.; Sorensen, E. J. *Nat. Biotechnol***. 2002**, *20*, 805.

68 a) Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogyo, M. *Chem. Biol*. **2000**, *7*, 569; (b) Borodovsky, A.; Ovaa, H.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Ploegh, H. L.; Kessler, B. M. *Chem. Biol*. **2002**, *9*, 1149; c) Faleiro, L.; Kobayashi, R.; Fearnhead, H.; Lazebnik, Y. *EMBO J*. **1997**, *16*, 2271.

69 a) Nazif, T.; Bogyo, M. *Proc. Natl. Acad. Sci*. *USA*. **2001**, *98*, 2967; c) Wang, G.; Uttamchandani, M.; Chen, G. Y. J.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 737; d)

Lo, L. C.; Pang, T. L.; Kuo, C. H.; Chiang, Y. L.; Wang, H. Y.; Lin, J. J. *J. Proteome Res*. **2002**, *1*, 35.

70 Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol. BioSyst*. **2008**, *4*, 473.

71 Chang, C. -F.; Ho, C. -W.; Wu, C. -Y.; Chao, T. -A.; Wong, C. -H.; Lin, C. -H. *Chem. Biol.* **2004**, *11*, 1301.

72 a) Mastrolorenzo, A.; Rusconi, S.; Scozzafava, A.; Barbaro, G.; Supuran, C. T. *Curr. Med. Chem.* **2007**, *14*, 2734; b) Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. -H. *Proc.Natl. Acad. Sci. USA* **1998**, *95*, 939; d) Wilkerson, W. W.; Dax, S.; Cheatham, W. W. *J. Med. Chem.* **1997**, *40*, 4079.

73 a) Jensen, C.; Herold, P.; Brunner, H. R. *Nat. Rev. Drug Disc Today*. **2008**, *7*, 399; b) Dash, C.; Kulkarni, A.; Dunn, B. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *38*, 89; c) Hills, I. D.; Vacca, J. P. *Curr. Opin. Drug. Disc. Dev.* **2007**, *10*, 383; d) Ersmark, K.; Samuelsson, B.; Hallberg, A. *Med. Res. Rev.* **2006**, *26*, 626.

74 Jensen, K. J.; Alsina, J.; Songster, M. F.; Barany, G. *J. Am. Chem. Soc.* **1998**, *120*, 5441.

75 Snow, R. W.; Guerra, C. A.; Noor, A. M.; Hay, S. I. *Nature.* **2005**, *434*, 214.

76 Boss, C.; Corminboeuf, O.; Grisostomi, C.; Meyer, S.; Jones, A. F.; Prade, L.; Binkert, C.; Fischli, W.; Weller, T.; Bur, D. *ChemMedChem*. **2006**, *1*, 1341.

77 a) Fidock, D. A.; Eastman, R. T.; Ward, S. A.; Meshnick, S. R. *Trends Parasitol*. **2008**, *24*, 537; b) Rosenthal, P. J. *Int. J. Parasitol*. **2004**, *34*, 1489; c) Goldberg, D. E. *Curr. Top. Microbiol. Immunol*. **2005**, *295*, 275.

78 a) Liu, J.; Istvan, E. S.; Gluzman, I. Y.; Gross, J.; Goldberg, D. E. *Proc. Natl. Acad. Sci. USA*. **2006**, *103*, 8840; b) Drew, M. E.; Banerjee, R.; Uffman, E. W.; Gilbertson, S.; Rosenthal, P. J.; Goldberg, D. E. *J. Biol. Chem*. **2008**, *283*, 12870.

79 a) Hof, F.; Schutz, A.; Fah, C.; Meyer, S.; Bur, D.; Liu, J.; Goldberg, D. E.; Diederich, F. *Angew. Chem.* **2006**, *118*, 2193; *Angew. Chem., Int. Ed.* **2006**, *45*, 2138; b) Nezami, A.; Kimura, T.; Hidaka, K.; Kiso, A.; Liu, J.; Kiso, Y.; Goldberg, D. E.; Freire, E. *Biochemistry*. **2003**, *42*, 8459.

80 a) Xiao, H.; Sinkovits, A. F.; Bryksa, B. C.; Ogawa, M.; Yada, R. Y. *Protein Expression Purif.* **2006**, *49*, 88; b) Xiao, H.; Tanaka, T.; Ogawa, M.; Yada, R. Y. *Protein Eng. Des. Sel*. **2007**, *20*, 625; c) Greenbaum, D. C. *Trends Pharmacol*. *Sci*. **2008**, *29*, 51.

81 a) Greenbaum, D. C.; Baruch, A.; Grainger, M.; Bozdech, Z.; Medzihradszky, K. F.; Engel, J.; DeRisi, J.; Holder, A. A.; Bogyo, M. *Science*. **2002**, *298*, 2002; b) Arastu-Kapur, S.; Ponder, E. L.; Fonovic, U. P.; Yeoh, S.; Yuan, F.; Fonovic, M.; Grainger, M.; Phillips, C. I.; Powers, J. C.; Bogyo, M. *Nat. Chem. Biol.* **2008**, *4*, 203.

82 Chan, E. W. S.; Chattopadhaya, S.; Panicker, R. C.; Huang, X.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*,14435.

83 a) Weihofen, A.; Binns, K.; Lemberg, M. K.; Ashman, K.; Martoglio, B. *Science.* **2002**, *296*, 2215; b) Takahashi, Y.; Konno, Y.; Watanabe, N.;

Miyashita, H.; Sasaki, M.; Natsugari, H.; Kan, T.; Fukuyama, T.; Tomita, T.; Iwatsubo, T. *ACS Chem. Biol.* **2007**, *2*, 408.

84 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.* **2002**, *114*, 2708.

85 a) Bonniec, S. L.; Deregnaucourt, C.; Redeker, V.; Banerjeei, R.; Grellier, P.; Goldbergi, D. E.; Schrevel, J. *J. Biol. Chem*. **1999**, *274*, 14218; b) Dejkriengkraikhul, P.;. Wilairat, P. Z. *Parasitenkd.***1983**, *69*, 313.

86 Phaumik, P.; Xiao, H.; Parr, C. L.; Kiso, Y.; Gustchina, A.; Yada, R. Y.; Wlodawer, A. *J. Mol. Biol*. **2009**, *388*, 520.

87 According to the docking results, an S configuration was favored for the carbon center attached to the hydroxy group in **G16**.

88 a) Duffner, J. L.; Clemons, P. A.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2007**, *11*, 74; (b) Uttamchandani, M.; Wang, J.; Yao, S. Q. *Mol. BioSyst*. **2006**, *2*, 58.

89 James, E.; Bradner, O. M.; McPherson, R. M.; David, B.-S.; John, P. S.; Jasmeet, D.; Kristen, E.; Stevenson, J. L.; Duffner, S. B.; Park, D. S.; Neuberg, P. N.; Stuart, L S.; Koehler, A. N. *Chemistry & Biology* **2006**, *13*, 493.

90 Jakob-Roetne, R.; Jacobsen, H. *Angew. Chem., Int. Ed.* **2009**, *48*, 3030.

91 a) Li, Y.-M.; Xu, M.; Lai, M. T.; Huang, Q.; Castro, J. L. *Nature.* **2000**, *405*, 689; b) Chun, J.; Yin, Y. I.; Yang, G.; Tarassishin, L.; Li, Y.-M. *J. Org. Chem.* **2004**, *69*, 7344; c) Fuwa, H.; Hiromotoa, K.; Takahashia, Y.; Yokoshimaa, S.; Kan, T.; Fukuyamaa, T.; Iwatsuboa, T.; Tomitaa and, T.; Natsugari, H. *Bioorg. Med. Chem. Lett*. **2006**, *16*, 4184.

92 Kimberly, W. T.; LaVoie, M. J.; Ostaszewski, B. L.; Ye, W.; Wolfe, M. S.; Selkoe, D. J. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 6382.

93 Patton, W. F.; Schulenberg, B.; Steinberg, T. H. *Curr. Opin. Biotechnol.* **2002**, *13*, 321.

94 a) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269; b) Gstaiger, M.; Aebersold, R. *Nat. Rev. Genet*. **2009**, *10*, 617; c) Cravatt, B. F.; Simon, G. M.; Yates, J. R. *Nature*. **2007**, *450*, 991.

95 a) Miura, S.; Miura, K.; Masuzaki, H.; Miyake, N.; Yoshiura, K.; Sosonkina, N.; Harada, N.; Shimokawa, O.; Nakayama, D.; Yoshimura, S.; Matsumoto, N.; Niikawa, N.; Ishimaru, T. *J. Hum. Genet.* **2006**, *51*, 412; b) Brown, P. P.; Botstein. D. *Nat. Genet*. **1999**, *21*, 33.

96 a) Kramer, R.; Cohen, D. *Nat. Rev. Drug. Discov*. **2004**, *3*, 965; b) Patterson, S. D.; Aebersold, R. *Nat. Genet*. **2003**, *33*, 311.

97 a) Kobe, B.; Kemp, B. E. *Nature.* **1999**, *402*, 373; b) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem*. **2008**, *77*, 383.

98 Sadaghiani, A. M.; Verhelst, S. H. L.; Bogyo, M. *Curr. Opin. Chem. Biol.* **2007**, *11,* 20.

99 Uttamchandani, M.; Liu, K.; Panicker, R. C.; Yao, S. Q. *Chem. Commun.* **2007**, 1518.

100 Shi, H.; Liu, K.; Xu, A.; Yao. S. Q. *Chem. Commun.* **2009**, 5030.

101 Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt. B. F. *Proc. Natl. Acad. Sci. USA*. **2004**, *101*, 10000.

102 a) Kumar, S.; Zhou, B.; Liang, F.; Wang, W. Q.; Huang, Z.; Zhang, Z. Y. *Proc. Natl. Acad. Sci. USA*. **2004**, *101*, 7943; b) Zhu, Q.; Huang, X.; Chen, G. Y. J.; Yao, S. Q. *Tetrahedron Lett.* **2003**, *44*, 2669.

103 Cohen, M. S.; Hadjivassiliou, H.; Taunton, J. *Nat. Chem. Biol.* **2007**, *3*, 156.

104 a) Srinivasan, R.; Huang, X.; Ng, S. L.; Yao, S. Q. *ChemBioChem.* **2006**, *7*, 32; b) Blais, D. R.; Brûlotte, M.; Qian, Y.; Bélanger, S.; Yao, S. Q.; Pezacki, J. P. *J. Proteome. Res.* **2010**, *9*, 912; c) Yang, P.-Y.; Liu, K.; Ngai, M. H.; Lear, M. J.; Wenk, M.; Yao, S. Q. *J. Am. Chem. Soc.* **2010**, *132*, 656.

105 a) Tomizaki, K. Y.; Usui, K.; Mihara, H. *ChemBioChem*. **2005**, *6*, 783; b) Kodadek, T. *Trends. Biochem. Sci.* **2002**, *27*, 295.

106 a) Sun, H.; Tan, L. P.; Gao, L.; Yao, S. Q. *Chem. Commun.* **2009**, 677; b) Lu, C. H. S.; Sun, H.; Bakar, F. B. A.; Uttamchandani, M.; Zhou, W.; Liou, Y.-C.; Yao, S. Q. *Angew. Chem.; Int. Ed.* **2008**, *47*, 7438; c) Sun, H.; Lu, C. H. S.; Uttamchandani, M.; Xia, Y.; Liou, Y.-C.; Yao, S. Q. *Angew. Chem.; Int. Ed.* **2008**, *47*, 1698; d) Uttamchandani, M.; Lee, W. L.; Wang, J.; Yao, S. Q. *J. Am. Chem. Soc.* **2007**, *129*, 13110; e) Salisbury, C. M.; Maly, D. J.; Ellman, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 14868; f) Kohn, M.; Gutierrez-Rodriguez, M.; Jonkheijm, P.; Wetzel, S.; Wacker, R.; Schroeder, H.; Prinz, H.; Niemeyer, C. M.; Breinbauer, R.; Szedlacsek, S. E.; Waldmann, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 7700; g) Winssinger, N.; Damoiseaux, R.; Tully, D. C.; Geierstanger, B. H.; Burdick, K.; Harris, J. L. *Chem. Biol.* **2004**, *11*, 1351; h) Reddy, M. M.; Kodadek, T. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12672; i) Roska, R. L. W.;

Lama, T. G. S.; Henne, J. P.; Carlson, R. E. *J. Am. Chem. Soc.* **2009**, *131*, 16660; j) Usui, K.; Tomizaki, K.; Mihara, H. *Mol. BioSyst.* **2006**, *2*, 417.

107 Wu, H.; Ge, J.; Yang, P.-Y.; Wang, J.; Uttamchandani, M.; Yao, S. Q. *J. Am. Chem. Soc.* **2011**, *133*, 1946.

108 http://merops.sanger.ac.uk/

109 Kaldor, S. W.; Hammond, M.; Dressman, B. A.; Fritz, J. E.; Crowell, T. A.; Hermann, R. A. *Bioorg. Med. Chem. Lett*. **1994**, *4*, 1385.

110 a) Mamidyala, S. K.; Finn, M. G. *Chem. Soc. Rev.* **2010**, *39*, 1252; b) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem.; Int. Ed.* **2009**, *48*, 6974; c) Meldal, M.; Tornøe, C. W. *Chem. Rev.* **2008**, *108*, 2952.

111 Chino, M.; Wakao, M.; Ellman, J. A. *Tetrahedron*. **2002**, *58*, 6305.

112 Uttamchandani, M.; Wang, J.; Li, J.; Hu, M.; Sun, H.; Chen, K. Y.-T.; Liu, K.; Yao, S. Q. *J. Am. Chem. Soc*. **2007**, *129*, 7848.

113 Sieber, S. A.; Niessen, S.; Hoover, H. S.; Cravatt, B. F. *Nat. Chem. Biol*. **2006**, *2*, 274.

114 a) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252; b) Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. *Anal. Bioanal. Chem.* **2007**, *389*, 1017.

115 a) Zhu, W.; Smith, J. W.; Huang, C. M. *J. Biomed. Biotechnol.* **2010**, *2010*, 840518; b) Liu, H.; Sadygov, R. G.; Yates, J. R. *3rd, Anal. Chem.* **2004**, *76*, 4193; c) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.;

Mann, M. *Mol. Cell. Proteomics* **2005**, *4*, 1265; d) Shinoda, K.; Tomita1, M.; Ishihama, Y. *Bioinformatics.* **2010**, *26*, 576.

116 Bewley, M. A.; Pham, T. K.; Marriott, H. M.; Noirel, J.; Chu, H.-P.; Ow, S. Y.; Ryazanov, A. G.; Read, R. C.; Whyte, M. K. B.; Chain, B.; Wright, P. C.; Dockrell, D. H. *Mol. Cell. Proteomics.* **2011**, *10*, DOI: 10.1074/ mcp. M111.008193.

117 a) Bracco, F.; Banay-Schwartz, M.; Deguzman, T.; Lajtha, A. *Neurochem Int.* **1982**, *4*, 541; b) Koike, M.; Nakanishi, H.; Saftig, P.; Ezaki, J.; Isahara, K.; Ohsawa, Y.; Schulz-Schaeffer, W.; Watanabe, T.; Waguri, S.; Kametaka, S.; Shibata, M.; Yamamoto, K.; Kominami, E.; Peters, C.; Figura, K. V.; Uchiyama, Y. *J. Neurosci*. **2000**, *20*, 6898.

118 a) Naguleswarana, A.; Alaeddinea, F.; Guionaudb, C.; Vonlaufena, N.; Sondac, S.; Jenoed, P.; Mevissenb, M.; Hemphilla, A. *Int. J. Parasitol.* **2005**, *35*, 1459; b) Delom, F.; Mallet, B.; Carayon, P.; Lejeune, P. J. *J. Biol. Chem.* **2001**. *276*, 21337; c) Ryser, H. J.; Levy, E. M.; Mandel, R.; DiSciullo, G. J. *Proc. Natl. Acad. Sci. USA*. **1994**, *91*, 4559.

119 Kim, L. G.; Song, L.-X.; Haura, E. B. *Nature Reviews*. **2009**, *6*, 587.

120 Irby, R. B.; Yeatman, T. J. *Oncogene*. **2000**, *19*, 5636.

121 Baselga, J. *Science*. **2006**, *312*, 1175; (b) Sebolt-Leopold, J. S.; English, J. M. *Nature*. **2006**, *441*, 457.

122 Burgess, M. R.; Skaggs, B. J.; Shah, N. P.; Lee, F. Y.; Sawyers, C. L. *Proc. Natl. Acad. Sci. USA*. **2005**, *102*, 3395.

123 a) Davies, S. P.; Reddy, H.; Caivano, M.; Cohen, P. *Biochem. J.* **2000**, *351*, 95; b) Bain, J.; McLauchlan, H.; Elliott, M.; Cohen, P. *Biochem. J.* **2003**, *371*, 199.

124 Carter, T. A.; Wodicka, L. M.; Shah, N. P.; Velasco, A. M.; Fabian, M. A.; Treiber, D. K. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 11011.

125 Knockaert, M.; Wieking, K.; Schmitt, S.; Leost, M.; Grant, K. M.; Mottram, J. C.; Kunick, C.; Meijer, L.; *J. Biol. Chem.* **2002**, *277*, 25493.

126 Hantschel, O.; Rix, U.; Schmidt, U.; Burckstummer, T.; Kneidinger, M.; Schutze, G.; Colinge, J.; Bennett, K. L.; Ellmeier, W.; Valent, Peter.; Giulio, S. F. *Proc. Natl. Acad. Sci. USA.* **2007**, *14*, 13283.

127 a) Uwe, R.; Oliver, H.; Gerhard, D.; Lily, L.; Remsing, R.; Melanie, P.; Nora, V. F.; Ines, K.; Bennett, K. L.; Peter, V.; Jacques, C.; Thomas, K.; Giulio, S. F. *Blood*, **2007**, *110*, 4055; b) Bantscheff, M.; Eberhard, D.; Abraham, Y.; Bastuck, S.; Boesche, M.; Hobson, S.; Mathieson, T.; Perrin, J.; Raida, M.; Christina, R.; Reader, V.; Sweetman, G.; Bauer, A.; Bouwmeester, T.; Hopf, C.; Kruse, U.; Neubauer, G.; Ramsden, N.; Rick, J.; Kuster, B.; Drewes, G.; *Nature Biotechnology.* **2007**, *25*, 1035; c) Li, J. H.; Rix, U.; Fang, B.; Bai1, Y.; Edwards, A.; Colinge, J.; Bennett, K. L.; Gao, J. C.; Song, L. X.; Eschrich, S.; Giulio, S. F.; Koomen, J.; Haura, E. B. *Nature Chemical Biology*. **2010**, *6*, 291.

128 a) Fabian, M. A.; William, H. B.; Daniel, K. T.; Corey, E. A.; Mihai, D. A.; Michael, G. B.; Todd, A. C.; Pietro, C.; Philip, T. E.; Mark, F. *Nat. Biotechnol.* **2005**, *23*, 329; b) Szardenings, K.; Li, B.; Ma, L.; Wu, M. *Drug. Discov. Today: Technologies.* **2004**, *3*, 9; c) Godl, K.; Wissing, J.; Kurtenbach, A.; Habenberger,

P.; Blencke, S.; Gutbrod, H.; Salassidis, K.; Stein-Gerlach, M.; Missio, A.; Cotten, M.; Daub, H. *Proc. Natl. Acad. Sci. USA*. **2003**, *100*, 15434; d) Knockaert, M.; Gray, N.; Damiens, E.; Chang, Y.-T.; Grellier, P.; Grant, K.; Fergusson, D.; Mottram, J.; Soete, M.; Dubremetz, J.-F.; Roch, K. L.; Doerig, C.; Schultz, P. G.; Meijeret, L. *Chem. Biol.* **2000**, *7*, 411.

129 a) Jessani, N.; Humphrey, M.; McDonald, W. H., Niessen, S.; Masuda, K.; Gangadharan, B.; Yates, J. R.; Mueller, B. M.; Cravatt, B. F. *Proc. Natl. Acad. Sci. USA*. **2004**, *101*, 13756; b) Kidd, D.; Liu, Y.; Cravatt, B. F. *Biochemistry*. **2001**, *40*, 4005.

130 Karaman, M. W.; Herrgard, Sanna.; Treiber, D. K.; Gallant, Paul.; Atteridge, C. E.; Zarrinkar, P. P. *Nat. Biotechnol*. **2008**, *26*, 127.

131 Xu, W. Q.; Doshi, A.; Lei, M.; Eck, M. J.; Harrison, S. C. *Mol. Cell.* **1999**, *3*, 629.

132 Tokarski, J. S.; Newitt, J. A.; Chang, C. Y.; Cheng, J. D.; Wittekind, M.; Kiefer, S. E.; Kish, K.; Lee, F. Y.; Borzillerri, R.; Lombardo, L. *Cancer Res.* **2006**, *66*, 5790.

133 a) Sanner, M. F. *J. Mol. Graphics Mod.* **1999**, *17*, 57; b) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455.

134 Nam, S.; Kim, D. H.; Cheng, J. Q.; Zhang, S. M.; Lee, J. H.; Janni Mirosevich, R. B.; Lee, F. Y.; Jove, R. *Cancer Res.* **2005**, *65*, 9185.

135 a) Apsel1, B.; Blair, J. A.; Gonzalez, B.; Nazif, T. M.; Feldman, M. E.; Aizenstein, B.; Hoffman, R.; Williams, R. L.; Shokat, K. M.; Knight, Z. A.

Nature Chemical Biology. **2008**, *4*, 691; b) Matthaus, G.; Christian, G.; Jeffrey, R.; Simard, S. K.; Matthias, R.; Haridas, B.; Rode, A. R.; Daniel, R. *J. Med. Chem.* **2009,** *52*, 3915.

136 Seeliger, M. A.; Young, M. M.; Henderson, N.; Pellicena, P.; King, D. S.; Falick, A. M.; Kuriyan, J. *Protein Science.* **2005**, *14*, 3135.

137 Amandine, C.; Sophie, L.; Edwige, V.; Laurent, G.; Patrice, D.; Paulo, D. S. *J. Biol. Chem*. **2011**, *286*, 5956.

138 a) Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. *Nat*. *Rev*. *Drug Discov*, **2002**, *1*, 493; b) Muhsin, M.; Graham, J.; Kirkpatrick, P. *Nat*. *Rev*. *Drug Discov*. **2002**, *2*, 515.

139 Tamaoki, T. *Methods Enzymol*. **1991**, *201*, 340.

140 Cragg, G. M.; Newman, D. J.; Snader, K. M.; *J. Nat. Prod.* **1997**, *60*, 52.

141 Oppermann, F. S.; Gnad, F.; Olsen, J. V.; Hornberger, R.; Greff, Z.; Keri, G.; Mann, M.; Daub, H. *Mol. Cell. Proteomics,* **2009**, *8*, 1751.

142 Fischer, J. J.; Graebner, O. Y.; Dalhoff, C.; Michaelis, S.; Schrey, A. K.; Ungewiss, J.; Andrich, K.; Jeske, D.; Kroll, F.; Glinski, M.; Sefkow, M.; Dreger and, M.; Koester, H. *J. Proteome. Res.* **2010**, *9*, 806.

143 Prade, L.; Engh, R. A.; Girod, A.; Kinzel, V.; Huber, R.; Bossemerer, D. *Structure*, **1997**, *5*, 1627.

144 MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J. *J. Am. Chem. Soc.* **2007**, *129*, 14560.

145 Heinsoo, A.; Raidaru, G.; Linask, K.; Jarv, J.; Zetterstrom, M.; Langel, U. *Tetrahedron*: *Asymmetry*. **1995**, *6*, 2245.

146 Jeroen, S.; Dickschat, H. R.; Irene, W. D.; Stefan, S. *Eur. J. Org. Chem.* **2005**, 4141.

147 Abderrahim, B.; Gilles, S. *Organic Lett.* **2002**, *4*, 2329.

148 Wang, J.; Uttamchandani, M.; Sun, L. P.; Yao, S. Q. *Chem. Commun.* **2006**, 717.

149 Wang, D. -X.; Huang, X. -Y.; Cole, P. -A. *Biochemistry*. **2001**, *40*, 2004.

150 Lars, P.; Richard, A. E.; Andreas, G.; Volker, K.; Robert, H.; Dirk, B. *Structure*. **1997**, *5*, 1628.

151 Wu, H.; Ge, J.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2010**, *49*, 6528.

Chapter 11

Appendix

11.1 Supplemental Tables

Table 11.1. Summary of Diol-based Inhibitors Characterizations.

#	ID of Product	LCMS			^{1}H NMR(Crude Product)
		Est % Purity	Cal MW	Obs MW	
1	SYM-01	>95%	716.29	717.262	-
2	SYM-02	>95%	568.26	569.233	-
3	SYM-03	>95%	558.23	559.205	-
4	SYM-04	~70%	564.30	587.255	Y
5	SYM-05	~60%	612.14	635.197	Y
6	SYM-06	>95%	564.30	565.274	-
7	SYM-07	>95%	508.24	531.197	Y
8	SYM-08	>95%	684.10	707.059	-
9	SYM-09	>95%	568.24	569.233	-
10	SYM-10	>95%	634.19	635.18	-
11	SYM-11	~90%	568.26	569.236	-
12	SYM-12	~50%	576.21	577.192	-
13	SYM-13	>95%	568.26	569.231	-
14	SYM-14	~50%	596.29	597.265	-
15	SYM-15	~80%	576.16	599.118	Y
16	SYM-16	>95%	624.25	625.255	-
17	SYM-17	~50%	666.13	689.082	-
18	SYM-18	-	608.27	-	Y
19	SYM-19	~70%	496.26	497.231	-
20	SYM-20	-	538.26	-	-
21	SYM-21	-	548.36	-	-
22	SYM-22	>95%	628.28	629.271	-
23	SYM-23	~90%	648.32	649.291	-
24	SYM-24	~80%	500.29	523.249	-
25	SYM-25	>95%	490.26	513.22	-
26	SYM-26	>95%	496.33	519.287	-
27	SYM-27	~90%	544.17	567.192	Y
28	SYM-28	-	496.33	-	-
29	SYM-29	~90%	440.27	463.233	-

30	SYM-30	>95%	616.13	617.107	-
31	SYM-31	~90%	500.29	523.246	-
32	SYM-32	70%	551.20	551.280	-
33	SYM-33	~80%	500.29	523.249	-
34	SYM-34	~90%	508.24	509.174	-
35	SYM-35	~90%	487.28	509.174	-
36	SYM-36	-	513.30	-	-
37	SYM-37	-	508.19	-	Y
38	SYM-38	~60%	541.25	541.273	Y
39	SYM-39	-	598.16	-	-
40	SYM-40	-	540.30	-	-
41	SYM-41	-	428.29	-	-
42	SYM-42	~90%	470.29	471.270	-
43	SYM-43	-	480.39	-	-
44	SYM-44	-	560.31	-	-
45	ASM-01	~60%	494.12	495.15	-
46	ASM-02	>90%	365.20	366.15	-
47	ASM-03	~70%	464.19	487.15	Y
48	ASM-04	~80%	411.18	434.15	Y
49	ASM-05	~70%	402.20	403.15	Y
50	ASM-06	~90%	402.20	403.15	Y
51	ASM-07	~80%	402.20	403.15	Y
52	ASM-08	~80%	466.13	467.20	-
53	ASM-09	~80%	400.20	401.15	Y
54	ASM-10	~90%	384.15	-	Y
55	ASM-11	~80%	456.13	457.10	-
56	ASM-12	-	682.20	683.25	Y
57	ASM-13	~60%	612.14	613.205	-
58	ASM-14	~80%	602.28	625.29	Y
59	ASM-15	>95%	646.18	647.1477	-
60	ASM-16	>95%	517.26	519.2506	-
61	ASM-17	~90%	554.20	577.149	-
62	ASM-18	~80%	616.26	639.2038	-
63	ASM-19	~60%	625.24	624.213	-
64	ASM-20	~80%	563.24	564.2102	-
65	ASM-21	>95%	558.22	559.2800	-
66	ASM-22	~60%	554.26	555.2268	-
67	ASM-23	~80%	554.26	555.2307	-
68	ASM-24	~70%	596.25	619.322	-
69	ASM-25	~50%	588.22	587.341	-
70	ASM-26	-	596.25	-	-
71	ASM-27	-	554.26	-	-
72	ASM-28	-	630.13	-	-

73	ASM-29	~80%	618.19	619.145	-
74	ASM-30	~60%	552.26	553.207	-
75	ASM-31	~70%	560.25	561.207	-

Table 11.2 Identity of each compound as well as its quality confirmed by LC/MS.

#	Product ID	Azide			Alkyne	LCMS			NMR & Scale up
		Warhead	R_1	R_2	R_3	Est % Purity	Cal. MW	Obs. MW	
1	N-A1	A	Phe	I	1	>90	571.17	572.168	-
2	N-B1	B		II		>90	547.25	548.253	-
3	N-C1	C	Leu	I		>95	537.19	538.183	-
4	N-D1	D		II		>90	513.27	514.266	-
5	N-E1	E	Tyr	I		>90	587.17	588.187	-
6	N-F1	F		II		>90	563.25	564.268	-
7	N-G1	G	Val	I		>90	523.17	524.191	-
8	N-H1	H		II		>95	499.25	500.296	-
9	N-A2	A	Phe	I	2	>90	585.19	586.222	-
10	N-B2	B		II		>90	561.27	562.275	-
11	N-C2	C	Leu	I		>90	551.21	552.222	-
12	N-D2	D		II		-	527.29	-	-
13	N-E2	E	Tyr	I		-	601.19	-	-
14	N-F2	F		II		-	577.27	-	-
15	N-G2	G	Val	I		-	537.19	-	-
16	N-H2	H		II		-	513.27	-	-
17	N-A3	A	Phe	I	3	>95	551.17	552.187	-
18	N-B3	B		II		>90	527.25	528.261	-
19	N-C3	C	Leu	I		>95	517.19	518.202	-
20	N-D3	D		II		-	493.27	-	-
21	N-E3	E	Tyr	I		-	567.17	-	-
22	N-F3	F		II		>90	543.25	544.258	-
23	N-G3	G	Val	I		-	503.17	-	-
24	N-H3	H		II		-	479.25	-	ν
25	N-A4	A	Phe	I	4	>95	551.17	552.179	-
26	N-B4	B		II		>95	527.25	528.261	-
27	N-C4	C	Leu	I		>90	517.19	518.198	-
28	N-D4	D		II		>90	493.27	494.277	-
29	N-E4	E	Tyr	I		-	567.17	-	-
30	N-F4	F		II		>85	543.25	544.258	-
31	N-G4	G	Val	I		-	503.17	-	-
32	N-H4	H		II		-	479.25	-	-
33	N-A5	A	Phe	I	5	>90	579.16	580.173	-
34	N-B5	B		II		>85	555.24	556.258	-

35	N-C5	C	Leu	I		>95	545.18	546.172	-
36	N-D5	D		II		-	521.26	-	-
37	N-E5	E	Tyr	I		>90	595.16	596.15	-
38	N-F5	F		II		>90	571.24	572.236	-
39	N-G5	G	Val	I		>85	531.16	532.192	-
40	N-H5	H		II		>90	507.24	508.236	-
41	N-A6	A	Phe	I	6	>95	555.12	556.111	-
42	N-B6	B		II		>95	531.2	532.169	-
43	N-C6	C	Leu	I		>95	521.14	522.126	-
44	N-D6	D		II		-	497.22	-	-
45	N-E6	E	Tyr	I		-	571.12	-	-
46	N-F6	F		II		>85	547.2	548.191	-
47	N-G6	G	Val	I		>90	507.12	508.116	-
48	N-H6	H		II		-	483.2	-	-
49	N-A7	A	Phe	I	7	>90	539.15	540.144	-
50	N-B7	B		II		>90	515.23	516.226	-
51	N-C7	C	Leu	I		>90	505.17	506.163	ν
52	N-D7	D		II		>90	481.25	504.221	-
53	N-E7	E	Tyr	I		>85	555.15	556.136	-
54	N-F7	F		II		>90	531.23	532.215	-
55	N-G7	G	Val	I		-	491.15	-	-
56	N-H7	H		II		>90	467.23	468.226	-
57	N-A8	A	Phe	I	8	-	557.14	-	-
58	N-B8	B		II		-	533.22	-	-
59	N-C8	C	Leu	I		-	523.16	-	-
60	N-D8	D		II		-	499.24	-	-
61	N-E8	E	Tyr	I		-	573.14	-	-
62	N-F8	F		II		-	549.22	-	-
63	N-G8	G	Val	I		>90	509.14	510.134	-
64	N-H8	H		II		-	485.22	-	-
65	N-A9	A	Phe	I	9	>95	553.16	554.16	-
66	N-B9	B		II		-	529.24	-	-
67	N-C9	C	Leu	I		>90	519.18	542.177	-
68	N-D9	D		II		-	495.26	-	-
69	N-E9	E	Tyr	I		>85	569.16	570.153	-
70	N-F9	F		II		>95	545.24	-	-
71	N-G9	G	Val	I		>90	505.16	506.16	-
72	N-H9	H		II		-	481.24	-	-
73	N-A10	A	Phe	I	10	>90	546.15	547.147	-
74	N-B10	B		II		-	522.23	-	-
75	N-C10	C	Leu	I		>95	512.17	513.169	-
76	N-D10	D		II		>95	488.25	489.242	-

77	N-E10	E	Tyr	I		>90	562.15	563.158	-
78	N-F10	F		II		>90	538.23	539.221	-
79	N-G10	G	Val	I		>85	498.15	499.167	-
80	N-H10	H		II		-	474.23	-	-
81	N-A11	A	Phe	I	11	>85	607.14	608.1371	-
82	N-B11	B		II		>90	583.22	584.2211	-
83	N-C11	C	Leu	I		>90	573.16	574.1565	-
84	N-D11	D		II		>95	549.24	550.2378	-
85	N-E11	E	Tyr	I		>95	623.14	624.132	-
86	N-F11	F		II		>95	599.22	600.213	-
87	N-G11	G	Val	I		>95	559.14	560.136	-
88	N-H11	H		II		>85	535.22	536.222	-
89	N-A12	A	Phe	I	12	>90	607.14	608.137	-
90	N-B12	B		II		>90	583.22	584.219	-
91	N-C12	C	Leu	I		>90	573.16	574.158	ν
92	N-D12	D		II		>90	549.24	550.24	-
93	N-E12	E	Tyr	I		>95	623.14	624.147	-
94	N-F12	F		II		>95	599.22	600.21	-
95	N-G12	G	Val	I		>90	559.14	560.149	-
96	N-H12	H		II		>90	535.22	536.222	-
97	N-A13	A	Phe	I	13	>95	593.11	594.11	-
98	N-B13	B		II		>95	569.19	570.191	-
99	N-C13	C	Leu	I		>90	559.13	560.129	-
100	N-D13	D		II		>95	535.21	536.26	-
101	N-E13	E	Tyr	I		>85	609.11	610.105	-
102	N-F13	F		II		>85	585.19	586.189	-
103	N-G13	G	Val	I		>90	545.11	546.112	-
104	N-H13	H		II		>90	521.19	522.191	-
105	N-A14	A	Phe	I	14	>90	599.14	600.136	-
106	N-B14	B		II		>90	575.22	576.223	-
107	N-C14	C	Leu	I		>90	565.16	566.157	-
108	N-D14	D		II		>90	541.24	542.232	-
109	N-E14	E	Tyr	I		-	615.14	-	-
110	N-F14	F		II		>85	591.22	592.217	-
111	N-G14	G	Val	I		>85	551.14	552.138	-
112	N-H14	H		II		>90	527.22	528.218	ν
113	N-A15	A	Phe	I	15	>90	585.16	586.159	-
114	N-B15	B		II		>85	561.24	562.236	-
115	N-C15	C	Leu	I		>90	551.18	552.177	-
116	N-D15	D		II		>90	527.26	528.254	-
117	N-E15	E	Tyr	I		>95	601.16	602.161	-
118	N-F15	F		II		>90	577.24	578.235	-

119	N-G15	G	Val	I		>90	537.16	538.1544	ν
120	N-H15	H		II		>85	513.24	514.248	-
121	N-A16	A	Phe	I	16	>90	591.09	592.078	-
122	N-B16	B		II		>90	567.17	568.161	-
123	N-C16	C	Leu	I		>90	557.11	558.109	ν
124	N-D16	D		II		>85	533.19	534.184	-
125	N-E16	E	Tyr	I		>95	607.09	608.081	-
126	N-F16	F		II		>90	583.17	584.168	-
127	N-G16	G	Val	I		>95	543.09	544.084	ν
128	N-H16	H		II		>85	519.17	520.161	-
129	N-A17	A	Phe	I	17	>90	633.16	634.157	-
130	N-B17	B		II		>95	609.24	610.239	-
131	N-C17	C	Leu	I		>95	599.18	600.148	-
132	N-D17	D		II		>90	575.26	576.239	-
133	N-E17	E	Tyr	I		>95	649.16	650.134	-
134	N-F17	F		II		>95	625.24	626.209	-
135	N-G17	G	Val	I		>95	585.16	586.139	-
136	N-H17	H		II		>95	561.24	562.221	-
137	N-A18	A	Phe	I	18	>90	602.11	603.094	-
138	N-B18	B		II		>90	578.19	579.173	-
139	N-C18	C	Leu	I		>90	568.13	569.11	-
140	N-D18	D		II		>90	544.21	545.191	-
141	N-E18	E	Tyr	I		>90	618.11	619.088	-
142	N-F18	F		II		-	594.19	-	-
143	N-G18	G	Val	I		>90	554.11	555.097	-
144	N-H18	H		II		-	530.19	-	-
145	N-A19	A	Phe	I	19	>90	599.17	600.149	-
146	N-B19	B		II		>90	575.25	576.224	-
147	N-C19	C	Leu	I		>90	565.19	566.174	-
148	N-D19	D		II		>90	541.27	542.245	-
149	N-E19	E	Tyr	I		>90	615.17	616.136	-
150	N-F19	F		II		>95	591.25	592.226	-
151	N-G19	G	Val	I		-	551.17	-	ν
152	N-H19	H		II		-	527.25	-	-

Table 11.3 Summary of LCMS Characterizations of the 284-member Library.

#	ID	Compound	Warhead	LCMS		
				Est % Purity	MW_{Calc}	MW_{Obs}
1	A01	N-A1		>90	623.31	624.238
2	A02	N-A2		>90	727.34	728.246
3	A03	N-A3		~85	657.28	658.191
4	A04	N-A4		>90	653.32	654.238

5	A05	N-A5		~80	675.27	676.179
6	A06	N-A6		~85	659.30	660.209
7	A07	N-A7		~50	686.29	688.193
8	A08	N-A8		>90	653.32	654.238
9	A09	N-A9		~70	651.36	652.258
10	A10	N-A10		>90	651.35	652.261
11	A11	N-A11		~85	653.32	654.247
12	A12	N-A12		>90	673.33	674.241
13	A13	N-A13		>90	648.31	649.226
14	A14	N-A14		>90	681.32	682.225
15	A15	N-A15		>90	651.35	652.261
16	A16	N-A16		~80	693.26	694.167
17	A17	N-A17		>90	737.21	740.110
18	A18	N-A18		~80	707.30	708.203
19	A19	N-A19		>90	671.29	672.200
20	A20	N-A20		~85	649.33	650.243
21	A21	N-A21		>90	763.23	764.119
22	A22	N-A22		~90	713.27	-
23	A23	N-A23		~80	683.34	684.249
24	A24	N-A24		>90	673.27	674.244
25	B01	N-A25		~90	674.33	674.243
26	B02	N-A26		~50	638.33	639.244
27	B03	N-A27		~80	655.32	656.238
28	B04	N-A28		~80	749.21	750.108
29	B05	N-A29		~85	687.35	688.259
30	B06	N-A30		~70	651.35	652.261
31	B07	N-A31	**N-Phe**	~80	651.35	652.257
32	B08	N-A32		~50	681.39	682.299
33	B09	N-A33		>90	667.34	668.255
34	B10	N-A34		~60	643.38	644.291
35	B11	N-A35		~50	663.35	664.260
36	B12	N-A36		~50	673.35	672.244
37	B13	N-A37		>90	711.37	712.272
38	B14	N-A38		~70	603.35	604.270
39	B15	N-A39		~70	671.29	673.223
40	B16	N-A40		~80	637.33	-
41	B17	N-A41		~85	655.32	656.238
42	B18	N-A42		~85	673.30	675.204
43	B19	N-A43		~80	709.30	710.203
44	B20	N-A44		~80	693.24	694.15
45	B21	N-A45		~80	704.27	-
46	B22	N-A46		~85	710.29	711.198

47	B23	N-A47		~80	695.26	696.165
48	B24	N-A48		~60	691.29	692.196
49	C01	N-A49		>90	735.31	736.216
50	C02	N-A50		>90	701.33	702.242
51	C03	N-A51		>90	665.24	666.150
52	C04	N-A52		~70	715.34	716.245
53	C05	N-A53		>90	701.29	702.199
54	C06	N-A54		>90	689.29	690.204
55	C07	N-A55		>90	673.30	674.211
56	C08	N-B1		~80	665.35	666.258
57	C09	N-B2		~70	769.37	770.262
58	C10	N-B3		~70	699.31	700.208
59	C11	N-B4		~80	695.36	696.259
60	C12	N-B5		~70	717.30	718.206
61	C13	N-B6		~80	701.33	702.231
62	C14	N-B8		~70	695.36	696.256
63	C15	N-B10		~80	693.38	694.283
64	C16	N-B11		~80	695.36	696.259
65	C17	N-B12		~70	751.33	-
66	C18	N-B13		~80	690.34	691.246
67	C19	N-B14		~70	723.35	724.247
68	C20	N-B15		~70	693.38	694.282
69	C21	N-B16		~70	569.16	570.153
70	C22	N-B17		~70	815.20	-
71	C23	N-B18		~70	749.33	750.222
72	C24	N-B19		~80	713.32	714.227
73	D01	N-B20		~80	691.36	692.264
74	D02	N-B21		~60	805.26	806.143
75	D03	N-B22		-	791.27	-
76	D04	N-B32		~70	723.42	746.299
77	D05	N-B33		~80	709.37	710.278
78	D06	N-B34		~70	685.41	686.312
79	D07	N-B37		~85	753.40	754.294
80	D08	N-B38	**N-Leu**	~60	645.38	646.292-
81	D09	N-B40		~70	679.36	680．265
82	D10	N-B55		~70	715.33	716.229
83	D11	N-B42		~70	715.33	-
84	D12	N-B43		~70	751.33	752.231
85	D13	N-B44		~60	735.27	736.172
86	D14	N-B46		~60	752.32	753.215
87	D15	N-B47		-	737.29	-
88	D16	N-B49		~60	777.34	778.241

89	D17	N-B50		~60	743.36	744.255
90	D18	N-B51		~80	707.27	708.174
91	D19	N-B52		~70	757.38	758.269
92	D20	N-B53		~70	743.32	744.217
93	D21	N-B31		~80	693.38	694.332
94	D22	N-C1		~60	575.31	576.349
95	D23	N-C2		~90	679.34	680.353
96	D24	N-C3		~80	609.28	610.284
97	E01	N-C4		~80	605.32	606.334
98	E02	N-C5		-	703.28	-
99	E03	N-C6		~60	611.30	612.307
100	E04	N-C8		~80	605.32	606.336
101	E05	N-C10	**N-Val**	-	679.36	-
102	E06	N-C11		-	681.34	-
103	E07	N-C12		~80	625.33	626.345
104	E08	N-C13		-	676.33	-
105	E09	N-C14		~80	633.32	634.334
106	E10	N-C15		~70	603.35	604.355
107	E11	N-C16		~40	645.26	646.265
108	E12	N-D1		~40	639.31	640.315
109	E13	N-D2		>90	743.34	744.349
110	E14	N-D3		>90	673.27	674.279
111	E15	N-D4		~50	669.32	670.328
112	E16	N-D5		-	691.26	-
113	E17	N-D6		~50	675.29	676.304
114	E18	N-D8		~60	669.32	670.338
115	E19	N-D10		~80	667.34	690.366
116	E20	N-D11		~60	669.32	670.336
117	E21	N-D12		~70	689.32	690.339
118	E22	N-D13		-	664.30	-
119	E23	N-D14		>90	697.31	698.334
120	E24	N-D15		>90	667.34	668.359
121	F01	N-D16	**N-Tyr**	~70	709.25	710.258
122	F02	N-D17		~50	753.20	754.177
123	F03	N-D18		~80	723.29	724.293
124	F04	N-D19		~90	687.29	688.297
125	F05	N-D20		~90	665.32	666.345
126	F06	N-D21		~50	779.22	780.231
127	F07	N-D22		>95	729.26	-
128	F08	N-D55		-	689.29	-
129	F09	N-D42		-	689.29	-
130	F10	N-D43		-	725.29	-

131	F11	N-D44		~50	709.24	710.263
132	F12	N-E1		~70	547.28	548.299
133	F13	N-E2		~80	651.31	652.310
134	F14	N-E3		~60	581.24	582.284
135	F15	N-E4		~80	577.29	578.301
136	F16	N-E5		~70	599.23	600.245
137	F17	N-E6		~70	583.26	584.279
138	F18	N-E8		~80	577.29	578.340
139	F19	N-E10		~70	575.31	576.330
140	F20	N-E11		~80	577.29	578.301
141	F21	N-E12		-	597.30	-
142	F22	N-E13		~80	572.28	573.288
143	F23	N-E14		~70	605.29	606.305
144	F24	N-E15		~70	575.31	576.327
145	G01	N-E16		~50	617.23	618.247
146	G02	N-E17	**N-Ala**	-	661.17	-
147	G03	N-E18		~40	631.27	632.251
148	G04	N-E19		~50	595.26	596.270
149	G05	N-E20		~40	573.30	574.333
150	G06	N-E21		~90	687.20	688.212
151	G07	N-E22		-	637.24	-
152	G08	N-E32		-	603.35	-
153	G09	N-E33		~80	591.31	592.331
154	G10	N-E34		~40	567.35	568.328
155	G11	N-E37		~60	635.34	636.382
156	G12	N-E38		~60	527.31	528.328
157	G13	N-E40		-	597.27	-
158	G14	N-E55		-	597.27	-
159	G15	N-E42		-	633.27	-
160	G16	N-E43		~50	617.21	618.272
161	G17	N-E44		-	634.26	-
162	G18	N-F1		~80	604.34	605.355
163	G19	N-F2		~50	708.37	709.379
164	G20	N-F3		~70	638.30	639.342
165	G21	N-F4		~70	634.35	635.368
166	G22	N-F5		~70	656.29	657.296
167	G23	N-F6		~70	640.32	641.331
168	G24	N-F8		~70	634.35	635.364
169	H01	N-F10		-	632.37	-
170	H02	N-F11	**N-Lys**	~70	634.35	635.363
171	H03	N-F12		~70	654.36	655.368
172	H04	N-F13		~80	629.34	630.355

173	H05	N-F14		~70	662.35	663.364
174	H06	N-F15		~70	632.37	633.364
175	H07	N-F16		~70	674.28	675.293
176	H08	N-F17		~70	718.23	721.243
177	H09	N-F18		~70	688.32	689.340
178	H10	N-F19		~70	652.32	653.333
179	H11	N-F20		~60	630.36	631.374
180	H12	N-F21		~60	744.25	745.266
181	H13	N-F22		-	694.29	-
182	H14	N-F55		~50	654.32	655.332
183	H15	N-F42		-	654.32	-
184	H16	N-F43		-	690.32	-
185	H17	C-G1		-	674.27	-
186	H18	C-G2		~70	765.36	766.370
187	H19	C-G3		>90	749.30	750.318
188	H20	C-G4		-	760.33	-
189	H21	C-G5		~80	766.35	767.375
190	H22	C-G6		~60	751.32	752.341
191	H23	C-G7		~70	791.38	792.392
192	H24	C-G8		~70	757.39	758.410
193	I01	C-G9		~50	721.30	722.318
194	I02	C-G10		~70	771.41	772.422
195	I03	C-G11		~60	757.35	758.362
196	I04	C-G12		-	745.35	-
197	I05	C-G13		~80	729.36	730.369
198	I06	C-G14	**C-Phe**	-	841.48	-
199	I07	C-G15		-	695.38	-
200	I08	C-G16		>90	772.37	773.381
201	I09	C-G17		-	760.33	-
202	I10	C-G18		~50	757.39	758.409
203	I11	C-G19		-	785.42	-
204	I12	C-G20		~50	808.40	809.415
205	I13	C-G21		-	817.23	-
206	I14	C-G22		~60	765.36	766.372
207	I15	C-G23		~80	785.42	786.430
208	I16	C-G24		~50	733.33	734.345
209	I17	C-G25		~50	751.32	752.333
210	I18	C-H1		~70	695.38	696.352-
211	I19	C-H2		~80	731.38	732.391
212	I20	C-H3	**C-Leu**	~85	715.32	716.335
213	I21	C-H4		-	726.34	-
214	I22	C-H5		~80	732.37	733.380

215	I23	C-H6		~30	717.34	718.354
216	I24	C-H7		~80	757.39	758.406
217	J01	C-H8		-	723.41	-
218	J02	C-H9		-	687.32	-
219	J03	C-H10		~60	737.42	738.438
220	J04	C-H11		~70	723.37	724.390
221	J05	C-H12		~70	711.37	712.408
222	J06	C-H13		~80	695.38	696.383
223	J07	C-H14		-	807.50	-
224	J08	C-H15		-	661.39	-
225	J09	C-H16		~80	738.38	739.405
226	J10	C-H17		-	726.34	-
227	J11	C-H18		~80	723.41	724.417
228	J12	C-H19		-	751.44	-
229	J13	C-H20		-	774.42	-
230	J14	C-H21		-	783.24	-
231	J15	C-H22		~60	731.38	732.389
232	J16	C-H23		~80	751.44	752.441
233	J17	C-H24		~70	699.35	700.364
234	J18	C-H25		~85	717.34	718.357
235	J19	C-I1		-	653.33	-
236	J20	C-I2		>90	689.33	690.338
237	J21	C-I3		>90	673.27	674.281
238	J22	C-I4		-	684.30	-
239	J23	C-I5		~70	690.32	691.332
240	J24	C-I6		~85	675.29	676.307
241	K01	C-I7		>90	715.34	716.352
242	K02	C-I8		~85	681.36	-
243	K03	C-I9		~90	645.27	646.280
244	K04	C-I10		>95	695.38	696.386
245	K05	C-I11		>90	681.32	682.339
246	K06	C-I12		>90	669.32	670.341
247	K07	C-I13		>95	653.33	654.33
248	K08	C-I14		-	765.45	-
249	K09	C-I15		-	619.34	-
250	K10	C-I16		~80	696.33	697.343
251	K11	C-I17	**C-Ala**	-	684.30	-
252	K12	C-I18		>90	681.38	682.369
253	K13	C-I19		~90	709.39	710.409
254	K14	C-I20		~90	732.37	733.389
255	K15	C-I21		-	741.20	-
256	K16	C-I22		~90	689.33	690.335

257	K17	C-I23		~90	709.39	710.397
258	K18	C-I24		>90	657.30	658.312
259	K19	C-I25		~90	675.29	676.309
260	K20	C-J1		~80	745.35	746.369
261	K21	C-J2		~90	781.35	782.365
262	K22	C-J3		~90	765.30	766.317
263	K23	C-J4		-	776.32	-
264	K24	C-J5		~90	782.35	783.361
265	L01	C-J6		~80	767.32	768.335
266	L02	C-J7		~90	807.37	808.383
267	L03	C-J8		~90	773.39	-
268	L04	C-J9		>90	737.30	738.313
269	L05	C-J10		>90	787.40	788.409
270	L06	C-J11		~85	773.35	774.366
271	L07	C-J12		>90	761.35	762.367
272	L08	C-J13	**C-Tyr**	~85	745.35	746.367
273	L09	C-J14		-	857.48	-
274	L10	C-J15		-	711.37	-
275	L11	C-J16		~90	788.36	789.366
276	L12	C-J17		-	776.32	-
277	L13	C-J18		>90	773.39	774.398
278	L14	C-J19		-	801.42	-
279	L15	C-J20		~90	824.40	825.420
280	L16	C-J21		-	833.22	-
281	L17	C-J22		~90	781.35	782.365
282	L18	C-J23		>90	801.42	802.425
283	L19	C-J24		~90	749.33	750.342
284	L20	C-J25		~80	767.32	768.339

Table 11.4 Summary of the 75-member, C-terminal Library.

#	ID	Structure (Figure S5)		LCMS		
		R1' (AA)	R2' (Sulfonyl chloride)	% purity	Cal MW	OBS MW
1	C-a-1	Phe	1	-	674.27	-
2	C-a-2	Phe	2	~70	765.36	766.370
3	C-a-3	Phe	3	>90	749.30	750.318
4	C-a-4	Phe	4	-	760.33	-
5	C-a-5	Phe	5	~80	766.35	767.375
6	C-a-6	Phe	6	~60	751.32	752.341
7	C-a-7	Phe	7	~70	791.38	792.392
8	C-a-8	Phe	8	~70	757.39	758.410
9	C-a-9	Phe	9	~50	721.30	722.318

10	C-a-10	Phe	10	~70	771.41	772.422
11	C-a-11	Phe	11	~60	757.35	758.362
12	C-a-12	Phe	12	-	745.35	-
13	C-a-13	Phe	13	~80	729.36	730.369
14	C-a-14	Phe	14	-	841.48	-
15	C-a-15	Phe	15	-	695.38	-
16	C-a-16	Phe	16	>90	772.37	773.381
17	C-a-17	Phe	17	-	760.33	-
18	C-a-18	Phe	18	~50	757.39	758.409
19	C-a-19	Phe	19	-	785.42	-
20	C-a-20	Phe	20	~50	808.40	809.415
21	C-a-21	Phe	21	-	817.23	-
22	C-a-22	Phe	22	~60	765.36	766.372
23	C-a-23	Phe	23	~80	785.42	786.430
24	C-a-24	Phe	24	~50	733.33	734.345
25	C-a-25	Phe	25	~50	751.32	752.333
26	C-b-1	Leu	1	~70	695.38	696.352-
27	C-b-2	Leu	2	~80	731.38	732.391
28	C-b-3	Leu	3	~85	715.32	716.335
29	C-b-4	Leu	4	-	726.34	-
30	C-b-5	Leu	5	~80	732.37	733.380
31	C-b-6	Leu	6	~30	717.34	718.354
32	C-b-7	Leu	7	~80	757.39	758.406
33	C-b-8	Leu	8	-	723.41	-
34	C-b-9	Leu	9	-	687.32	-
35	C-b-10	Leu	10	~60	737.42	738.438
36	C-b-11	Leu	11	~70	723.37	724.390
37	C-b-12	Leu	12	~70	711.37	712.408
38	C-b-13	Leu	13	~80	695.38	696.383
39	C-b-14	Leu	14	-	807.50	-
40	C-b-15	Leu	15	-	661.39	-
41	C-b-16	Leu	16	~80	738.38	739.405
42	C-b-17	Leu	17	-	726.34	-
43	C-b-18	Leu	18	~80	723.41	724.417
44	C-b-19	Leu	19	-	751.44	-
45	C-b-20	Leu	20	-	774.42	-
46	C-b-21	Leu	21	-	783.24	-
47	C-b-22	Leu	22	~60	731.38	732.389
48	C-b-23	Leu	23	~80	751.44	752.441
49	C-b-24	Leu	24	~70	699.35	700.364
50	C-b-25	Leu	25	~85	717.34	718.357
51	C-c-1	Ala	1	-	653.33	-
52	C-c-2	Ala	2	>90	689.33	690.338

53	C-c-3	Ala	3	>90	673.27	674.281
54	C-c-4	Ala	4	-	684.30	-
55	C-c-5	Ala	5	~70	690.32	691.332
56	C-c-6	Ala	6	~85	675.29	676.307
57	C-c-7	Ala	7	>90	715.34	716.352
58	C-c-8	Ala	8	~85	681.36	-
59	C-c-9	Ala	9	~90	645.27	646.280
60	C-c-10	Ala	10	>95	695.38	696.386
61	C-c-11	Ala	11	>90	681.32	682.339
62	C-c-12	Ala	12	>90	669.32	670.341
63	C-c-13	Ala	13	>95	653.33	654.33
64	C-c-14	Ala	14	-	765.45	-
65	C-c-15	Ala	15	-	619.34	-
66	C-c-16	Ala	16	~80	696.33	697.343
67	C-c-17	Ala	17	-	684.30	-
68	C-c-18	Ala	18	>90	681.38	682.369
69	C-c-19	Ala	19	~90	709.39	710.409
70	C-c-20	Ala	20	~90	732.37	733.389
71	C-c-21	Ala	21	-	741.20	-
72	C-c-22	Ala	22	~90	689.33	690.335
73	C-c-23	Ala	23	~90	709.39	710.397
74	C-c-24	Ala	24	>90	657.30	658.312
75	C-c-25	Ala	25	~90	675.29	676.309

Table 11.5 The list of some representative proteins identified from the LCMS experiments. Proteins are sorted by classification. Unique peptides are the number of all unique peptides observed for a particular protein. Protein tyrosine kinases, Ser/Thr kinase, other kinases and nonkinases are labeled in red, blue, pink and black, respectively. -, no data.

No	Kinase Name	ID	MW (kDa)	Unique peptides count						
				HepG2		K562		Imatinib	Matrix	Family
				In vitro	*In situ*	*In vitro*	*In situ*			
1	Isoform 1 Proto-oncogene tyrosine-protein kinase Src*	IPI00641230	61.06	1	-	-	-	-	-	Tyrosine kinase
2	Isoform 2 Proto-oncogene tyrosine-protein kinase Src	IPI00328867	61.06	-	-	2	-	-	-	Tyrosine kinase
3	Proto-oncogene tyrosine-protein kinase Yes*	IPI00013981	61.28	1	-	2	-	-	-	Tyrosine kinase
4	Isoform Lyn A of tyrosine-protein kinase Lyn*	IPI00298625	58.99	1	-	0	-	-	-	Tyrosine kinase
5	Isoform 1 Proto-oncogene tyrosine-protein kinase Fyn*	IPI00219012	61.23	1	-	0	-	-	-	Tyrosine kinase
6	Isoform p59-HCK of Tyrosine-protein kinase HCK*	IPI00029769	57.68	1	-	0	-	-	-	Tyrosine kinase
7	Tyrosine-protein kinase CSK	IPI00013212	51.24	-	-	38	-	-	36	Tyrosine kinase
8	Tyrosine-protein kinase BTK	IPI00029132	76.92	-	-	41	-	-	41	Tyrosine kinase
9	Isoform 1 of Epidermal growth factor receptor (EGFR)	IPI00018274	137.6	-	-	5	-	-	-	Tyrosine kinase
10	Twinfilin-1 (protein tyrosine kinase 9)	IPI00183508	44.12	1	-	0	-	-	-	Tyrosine

										kinase
11	Isoform 2 of Serine/threonine-protein kinase PCTAIRE-3 (PCTK3)	IPI00394661	54.73	2	-	3	-	-	-	Serine/threonine kinase
12	Isoform 1 of DNA-dependent protein kinase catalytic subunit (DNA-PKc)	IPI00296337	47.37	1	6	60	5	1	-	Serine/threonine kinase
13	Isoform 1 of MAP kinase-activated protein kinase 2 (MAPKAPK2)	IPI00026054	45.94	1	-	-	-	-	-	Serine/threonine kinase
14	Isoform CSBP1 of Mitogen-activated protein kinase 14 (p38α)	IPI00221141	47.41	-	-	41	-	-	5	Serine/threonine kinase
15	Serine/threonine-protein kinase 25 (STK25)	IPI00012093	48.31	2	-	1	2	-	-	Serine/threonine kinase
16	Ribosomal protein S6 kinase alpha-3 (p90-RSK-3)	IPI00020898	84.03	2	-	-	-	-	-	Serine/threonine kinase
17	Interferon-induced, double-stranded RNA-activated protein kinase (eIF-2A)	IPI00019463	62.51	2	-	2	1	-	-	Serine/threonine kinase
18	Serine/threonine-protein kinase Pim-3 (PIM3)	IPI00848067	33.23	-	3	-	-	-	-	Serine/threonine kinase
19	Isoform 2 of cAMP-dependent protein kinase, alpha-catalytic subunit (PKA C-α)	IPI00217960	39.91	-	2	-	3	-	-	Serine/threonine kinase
20	Isoform Beta-II of Protein kinase C beta type (PKC- β)	IPI00219628	78.16	-	-	6	-	-	-	Serine/threonine kinase
21	Serine/threonine-protein kinase 6 (STK6)	IPI00298940	46.18	-	-	36	-	-	3	Serine/threonine kinase
22	Cell division protein kinase 2 (CDK2)	IPI00031681	34.08	-	-	5	-	-	-	Serine/threonine kinase
23	Serine/threonine-protein kinase N2 (PKN2)	IPI00002804	130.1	-	-	55	-	-	-	Serine/threoni

			9							ne kinase
24	Isoform Del-701 of Signal transducer and activator of transcription 3(STAT3)	IPI00306436	88.72	1	3	3	3	-	-	STAT Family
25	Isoform Alpha of Signal transducer and activator of transcription 1-alpha/beta (STAT1)	IPI00030781	87.9	4	-	-	-	-	-	STAT Family
26	Isoform A of Bifunctional 3'-phosphoadenosine 5'-phosphosulfate	IPI00030009	70.03	2	-	-	-	-	-	APS kinase family
27	Isoform 1 of Acylglycerol kinase, mitochondrial precursor	IPI00019353	47.56	-	1	3	-	-	-	Lipid kinase
28	Isoform M2 of Pyruvate kinase isozymes M1/M2	IPI00479186	58.54	69	22	37	154	28	14	Pyruvate kinase
29	Phosphoglycerate kinase 1	IPI00169383	44.98	67	11	17	20	11	4	Pyruvate kinase
30	Isoform 1 of Pyridoxal kinase	IPI00013004	35.31	17	3	2	2	-	-	Pyridoxal kinase
31	Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	IPI00294380	71.45	7	2	7	4	-	-	Pyruvate kinase
32	6-phosphofructokinase type C	IPI00009790	86.45	5	2	10	4	-	-	phosphofructokinase
33	Isoform 1 of 6-phosphofructokinase, liver type	IPI00332371	85.76	8	4	1	5	-	-	phosphofructokinase
34	Isoform 1 of Adenylate kinase 2, mitochondrial	IPI00215901	26.69	5	2	1	-	-	-	AK2 subfamily
35	Ketosamine-3-kinase	IPI00099986	34.62	1	-	1	-	-	-	Fructosamine kinase
36	Isoform 1 of Nucleoside diphosphate kinase A	IPI00012048	17.31	2	-	3	1	-	-	NDK family
37	Isoform 3 of Adenylate kinase isoenzyme 2, mitochondrial	IPI00172460	22.42	-	2	5	-	-	-	AK2

										subfamily
38	CKMT1B Isoform 1 of Creatine kinase U-type, mitochondrial	IPI00658109	47.41	-	-	5	-	-	-	Creatine kinase
39	Ribose-phosphate pyrophosphokinase 1	IPI00219616	35.33	4	-	-	11	-	-	pyrophospho kinase
40	Isoform 2 of ADP-dependent glucokinase	IPI00329593	54.27	1	-	1	-	-	-	ADP-dependent kinase
41	Isoform 1 of Ribose-phosphate pyrophosphokinase 2	IPI00219617	35.33	3	-	4	11	-	-	pyrophospho kinase
42	Isoform Short of Adenosine kinase	IPI00234368	39.08	-	1	3	2	-	-	AK2 Kinase
43	Isoform 1 of Inositol-trisphosphate 3-kinase B (IP3K B)	IPI00021449	103.1	-	1	-	-	-	-	IPK Kinase
44	Ribose-phosphate pyrophosphokinase 3	IPI00218371	35.39	-	2	-	-	-	-	pyrophospho kinase
45	Creatine kinase B-type	IPI00022977	42.90	-	2	16	5	-	-	Creatine kinase
46	Isoform 1 of Hexokinase-1	IPI00018246	103.6	-	-	1	-	-	-	Hexokinase
47	NME1 Nucleoside diphosphate kinase	IPI00604590	32.91	-	-	3	-	-	-	NDK family
48	LOC727761 Thymidylate kinase	IPI00013862	23.98	-	-	1	-	-	-	Thymidylate kinase
49	Isoform 2 of SH3-containing GRB2-like protein 3	IPI00216201	39.28	-	-	3	-	-	-	SH3 domain
50	GK Isoform 1 of Glycerol kinase	IPI00027424	58.25	-	-	3	-	-	-	Glycerol kinase
51	Phosphoglycerate kinase 2	IPI00219568	45.17	-	-	6	17	-	-	Phosphoglycerate kinase
52	cDNA FLJ56047, highly similar to A kinase anchor protein	IPI00022585	102.4	-	-	1	-	-	-	-
53	Serine-threonine kinase receptor-associated protein	IPI00294536	40.01	1	-	6	1	-	-	-

54	Isoform 1 of Uridine-cytidine kinase 2	IPI00065671	29.45	1	-	1	-	-	-	Uridine-cytidine kinase
55	Receptor-type tyrosine-protein phosphatase beta	IPI00295577	224.3	1	-	-	-	-	-	PTP family
56	Protein tyrosine phosphatase-like protein PTPLAD1	IPI00008998	43.36	2	-	4	-	-	-	HACD family
57	Tyrosine-protein phosphatase non-receptor type 12	IPI00289082	88.92	-	-	1	-	-	-	PTP family
58	Tyrosine-protein phosphatase non-receptor type 21	IPI00004388	134.1		-	1	59	-	-	PTP family
59	Isoform Gamma-1 of Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	IPI00005705	37.70	4	2	19	2	-	-	PP-1 family
60	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B beta isoform	IPI00020850	52.08	-	-	44	-	-	-	Phosphatase 2A family
61	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	IPI00554737	66.06	2	-	6	-	--	-	Phosphatase 2A family
62	Alkaline phosphatase, placental-like	IPI00290380	57.63	5	16	8	-	-	-	Alkaline phosphatase
63	Alkaline phosphatase, placental type precursor	IPI00007289	59.93	-	16	5	8	-	-	Alkaline phosphatase
64	Protein phosphatase 1F	IPI00291412	49.97	-	-	2	-	-	-	PP2C family
65	Isoform 1 of LIM and SH3 domain protein 1	IPI00000861	30.10	-	-	5	-	-	-	SH3 domain
66	Src substrate cortactin	IPI00029601	61.72	-	2	2	-	-	-	SH3 domain
67	DNA damage-binding protein 1	IPI00293464	128.14	2	7	14	-	-	-	DDB1 family
68	Proliferation-associated protein 2G4	IPI00299000	44.10	4	-	28	-	-	-	Peptidase M24 family
69	Isoform 2 of CDK5 and ABL1 enzyme substrate 1	IPI00178886	42.24	-	-	33	-	-	-	Cyclin family
70	Isoform 1 of Brain acid soluble protein 1	IPI00299024	22.68.	3	-	3	-	-	-	BASP1

										family
71	Isoform 2 of Tumor protein D54	IPI00221178	19.95	3	-	2	-	-	-	TPD52 family
72	Isoform B of Phosphate carrier protein, mitochondrial	IPI00215777	40.39	22	1	74	-	-	-	Mitochondrial carrier
73	Calcium-binding mitochondrial carrier protein Aralar2	IPI00007084	74.53	6	-	30	-	-	-	Mitochondrial carrier
74	Isoform 1 of ATP-binding cassette sub-family F member 3	IPI00465160	80.09	1	-	1	-	-	-	ABCF family
75	Lung cancer oncogene 7	IPI00641950	38.38	-	-	6	-	-	-	-
76	Proteasome subunit beta type-4 precursor	IPI00555956	29.24	-	-	6	-	-	-	Peptidase T1B family
77	Isoform 4 of Cancer-associated gene 1 protein	IPI00514516	99.42	-	-	-	-	-	-	-
78	Novel protein similar to Pre-B cell enhancing factor	IPI00472879	53.65	-	-	-	7	-	-	-
79	Isoform 1 of Large proline-rich protein BAT3	IPI00465128	119.79	3	-	1		-	-	Ubiquitin-like domain
80	Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B, alpha isoform	IPI00442821	53.55	-	-	4	-	-	-	-
81	Fructose-bisphosphate aldolase C	IPI00418262	49.06	1	20	188	24	-	-	Fructose-bisphosphate
82	Isoform B of Mannose-6-phosphate receptor-binding protein 1	IPI00303882	47.19	2		26		-	-	Perilipin family
83	Human immunodeficiency virus type I enhancer-binding protein 2	IPI00144243	27.11	6	-	-	-	-	-	-
84	Isoform 1 of SAPK substrate protein 1	IPI00027378	33.42	-	-	5	-	-	-	-

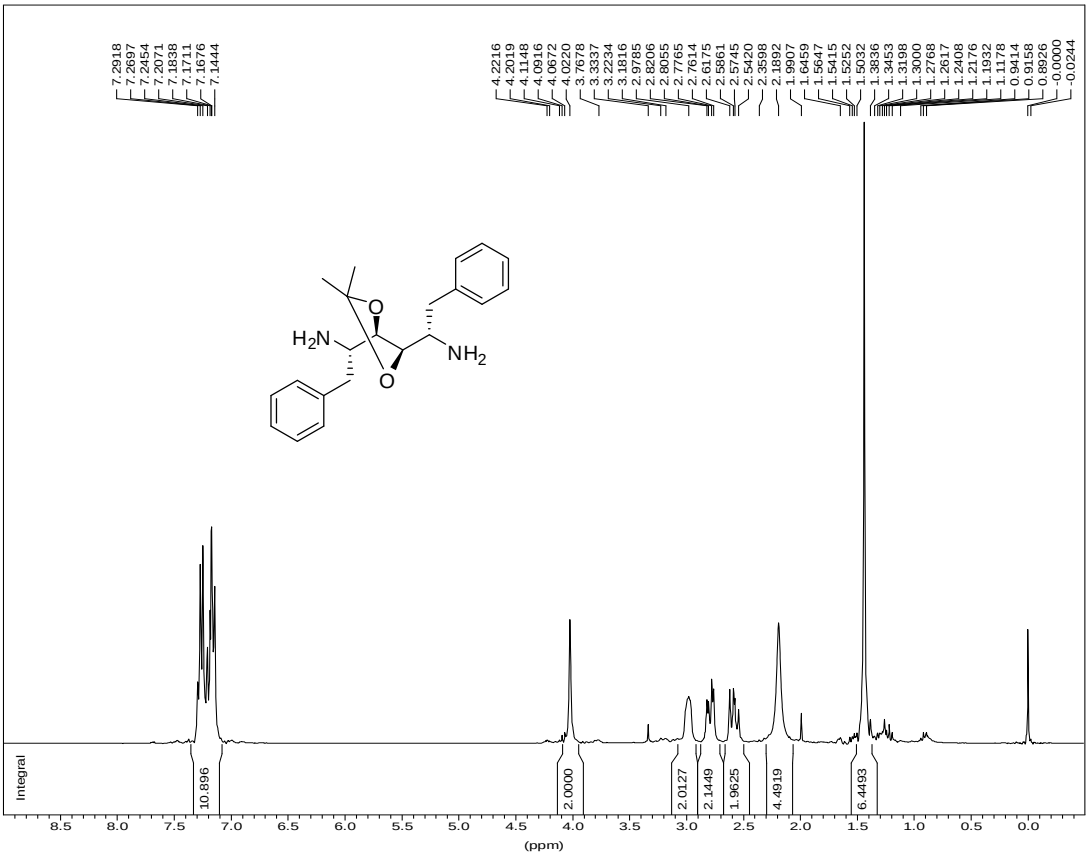
Table 11.6 The list of kinases identified from the LCMS experiments. Unique peptides are the number of all unique peptides observed for a particular protein. Protein kinases are shaded in blue , other kinases are labeled in blue. The unique kinases identified from *in situ* pulldown were shaded yellow. -, no data.

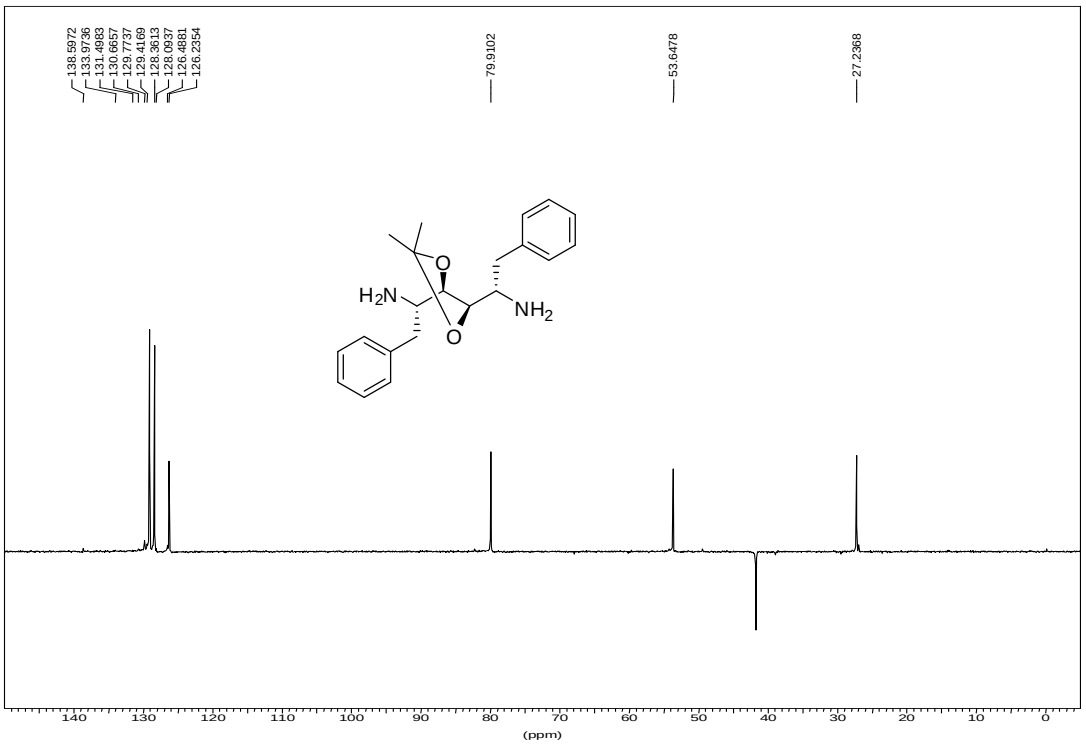
No	Kinase Name	ID	MW (KDa)	Unique peptides count		Family
				HepG2		
				In vitro	In situ	
1	Isoform 2 Proto-oncogene tyrosine-protein kinase Src	IPI00328867	61.06	2	-	Tyrosine kinase
2	Proto-oncogene tyrosine-protein kinase Yes	IPI00013981	61.28	2	-	Tyrosine kinase
3	Tyrosine-protein kinase CSK	IPI00013212	51.24	38	-	Tyrosine kinase
4	Tyrosine-protein kinase BTK	IPI00029132	76.92	41	-	Tyrosine kinase
5	Isoform 2 of Serine/threonine-protein kinase PCTAIRE-3 (PCTK3)	IPI00394661	54.73	3	-	Serine/threonine kinase
6	Isoform 1 of DNA-dependent protein kinase catalytic subunit (DNA-PKc)	IPI00296337	47.37	112	129	Serine/threonine kinase
7	Serine/threonine-protein kinase 25 (STK25)	IPI00012093	48.31	1	-	Serine/threonine kinase
8	Ribosomal protein S6 kinase alpha-3 (p90-RSK-3)	IPI00020898	84.03	3	1	Serine/threonine kinase
9	Ribosomal protein S6 kinase alpha-6	IPI00007123	84.39	-	1	Serine/threonine kinase
10	Interferon-induced, double-stranded RNA-activated protein kinase (eIF-2 α)	IPI00019463	62.51	2	2	Serine/threonine kinase
11	Isoform 2 of cAMP-dependent protein kinase, alpha-catalytic subunit (PKA C-α)	IPI00217960	39.91	-	2	Serine/threonine kinase
12	serine/threonine protein kinase MST4 isoform	IPI00182383	39.86	1	-	Serine/threonine kinase
13	Cell division protein kinase 2 (CDK2)	IPI00031681	34.08	5	-	Serine/threonine kinase
14	Serine/threonine-protein kinase N2 (PKN2)	IPI00002804	131.9	2	-	Serine/threonine kinase
15	Isoform 1 of Dephospho-CoA kinase domain-containing protein	IPI00291417	26.65	2	-	CoaE family

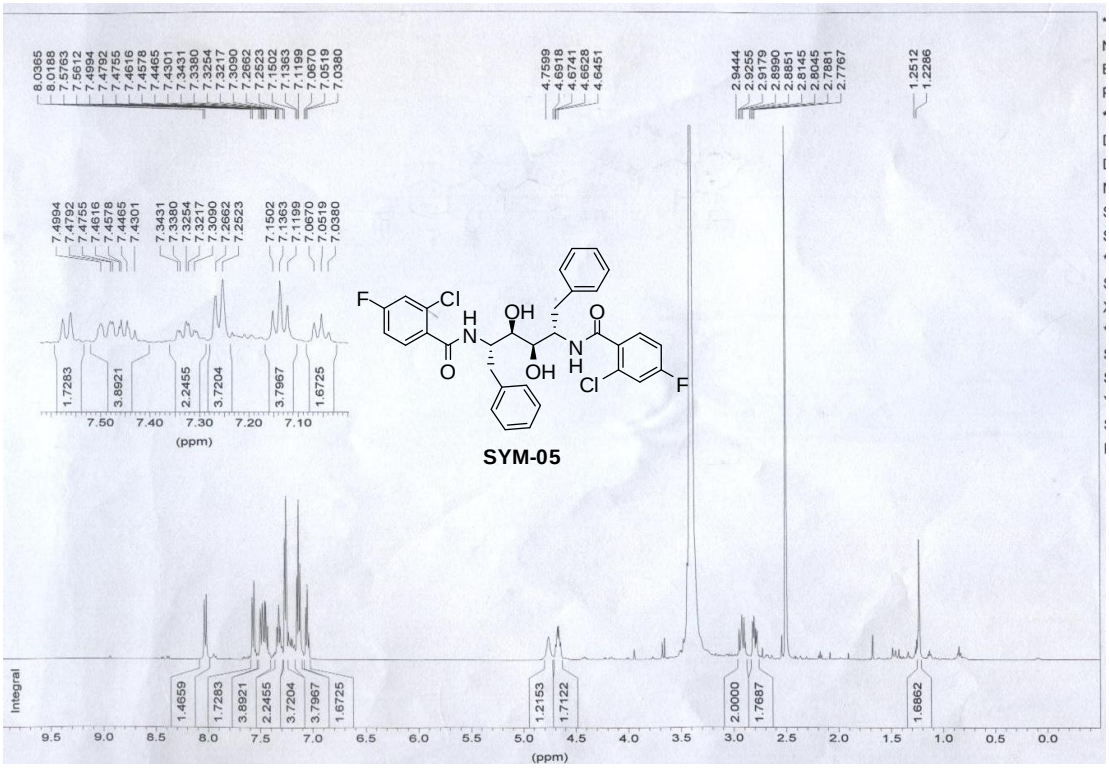
16	Myristoylated alanine-rich C-kinase substrate	IPI00219301	31.71	2	-	MARCKS family
17	Isoform 1 of Nucleoside diphosphate kinase B	IPI00026260	17.4	3	-	NDK family
18	Isoform 1 of Acylglycerol kinase, mitochondrial precursor	IPI00019353	47.56	2	1	Lipid kinase
19	Isoform M2 of Pyruvate kinase isozymes M1/M2	IPI00479186	58.54	60	1069	Pyruvate kinase
20	Phosphoglycerate kinase 1	IPI00169383	44.98	15	301	Pyruvate kinase
21	Isoform 1 of Pyridoxal kinase	IPI00013004	35.31	35	41	Pyridoxal kinase
22	Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	IPI00294380	71.45	7	-	Pyruvate kinase
23	6-phosphofructokinase type C	IPI00009790	86.45	2	3	phosphofructokinase
24	Isoform 1 of 6-phosphofructokinase, liver type	IPI00332371	85.76	4	9	phosphofructokinase
25	Isoform 2 of 6-phosphofructokinase, muscle type	IPI00219585	82.6	-	3	phosphofructokinase
26	Phosphofructokinase, platelet	IPI00639981	26.4	-	1	phosphofructokinase
27	Isoform 1 of Adenylate kinase 2, mitochondrial	IPI00215901	26.69	12	20	AK2 subfamily
28	Ketosamine-3-kinase	IPI00099986	34.62	1	-	Fructosamine kinase
29	Isoform 1 of Nucleoside diphosphate kinase A	IPI00012048	17.31	1	1	NDK family
30	Creatine kinase B-type	IPI00022977	42.90	0	17	Guanido phosphotransferase
31	Ribose-phosphate pyrophosphokinase 1	IPI00219616	35.33	7	4	pyrophosphokinase
32	Isoform 2 of ADP-dependent glucokinase	IPI00329593	54.27	1	-	ADP-dependent kinase
33	Ribose-phosphate pyrophosphokinase 2	IPI00219617	35.33	-	3	pyrophosphokinase
34	Isoform Short of Adenosine kinase	IPI00234368	39.08	-	1	AK2 Kinase
35	Isoform 1 of Hexokinase-1	IPI00018246	103.6	20	-	Hexokinase
36	LOC727761 Thymidylate kinase	IPI00013862	23.98	-	1	Thymidylate kinase
37	Isoform 2 of SH3-containing GRB2-like protein 3	IPI00216201	39.28	48	-	SH3 domain
38	Isoform 1 of Uridine-cytidine kinase 2	IPI00065671	29.45	1	-	Uridine-cytidine kinase
39	Isoform 1 of Inhibitor of Bruton tyrosine kinase	IPI00792333	15.06	-	2	Tyrosine kinase inhibitor
40	Isoform R-type of Pyruvate kinase isozymes R/L	IPI00027165	62.19	11	-	Pyruvate kinase

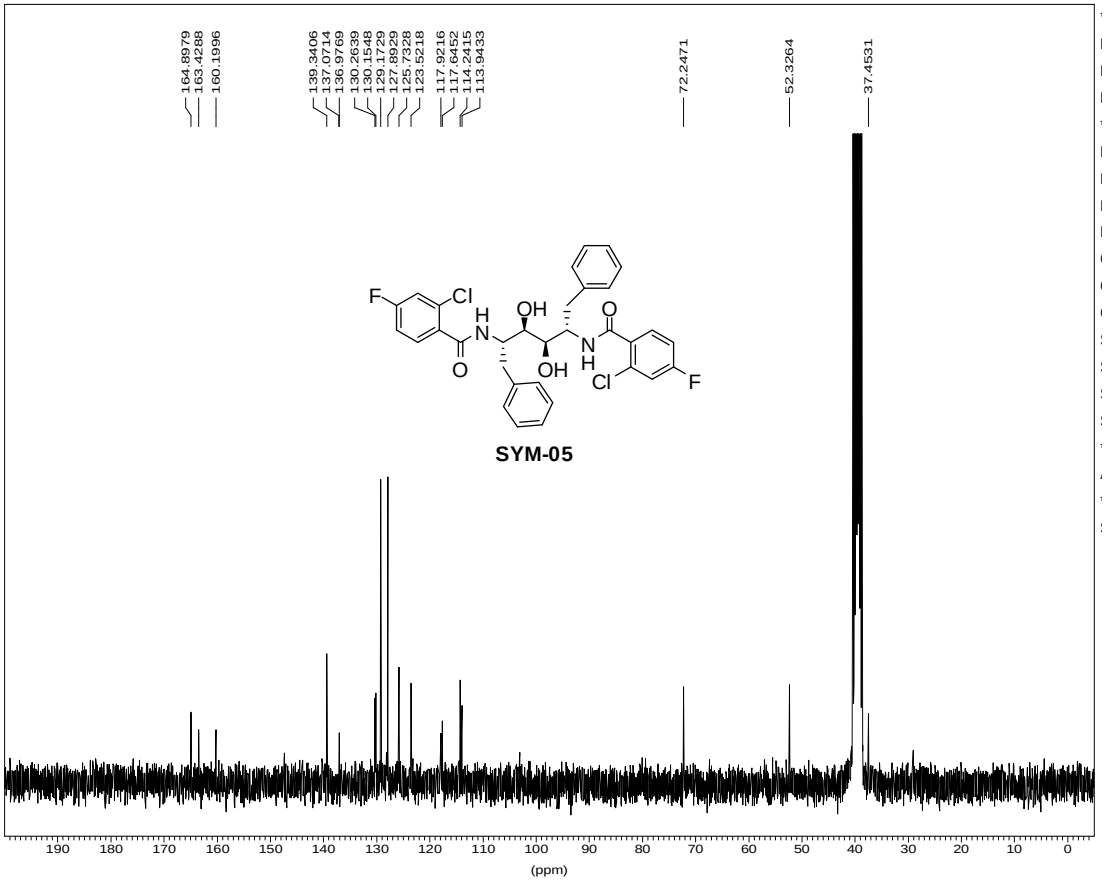
41	FN3KRP Ketosamine-3-kinase	IPI00099986	34.6	1	-	Fructosamine kinase
42	Thymidine kinase 1, soluble	IPI00299214	30.56	2	-	Thymidine kinase
43	Isoform 1 of S-phase kinase-associated protein 1A	IPI00301364	18.82	1	-	-

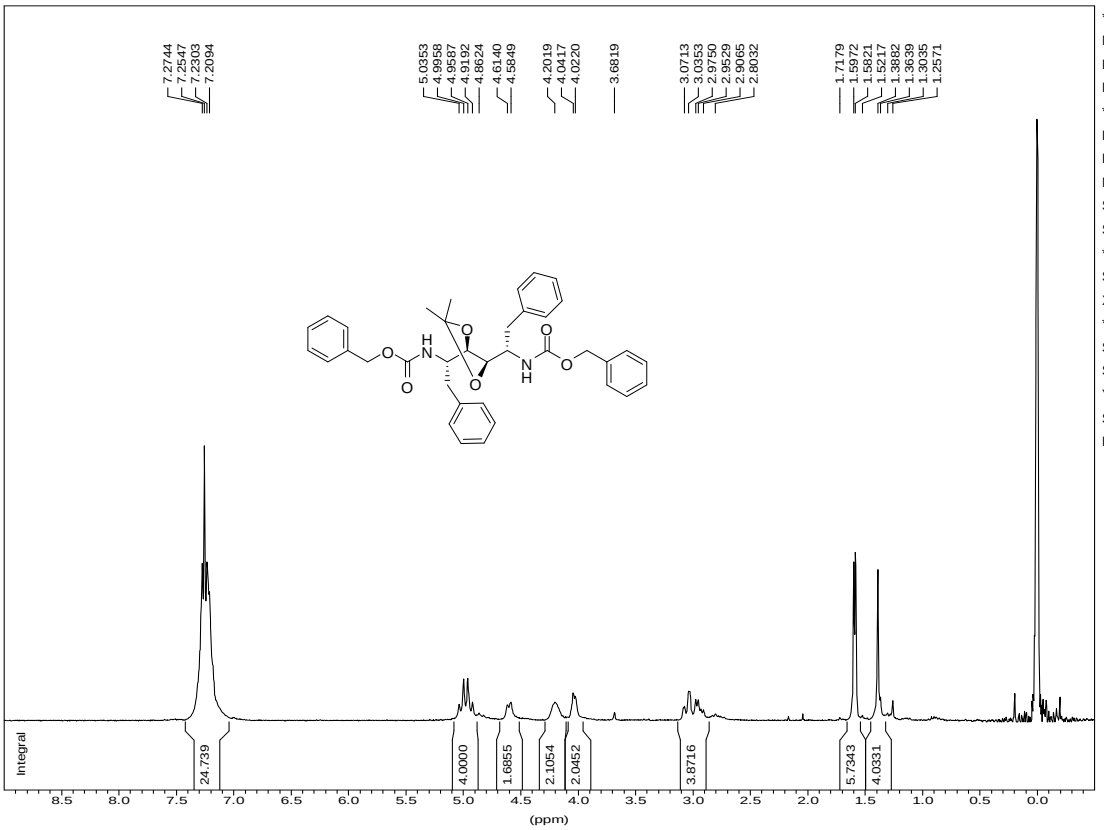
11.2 Supplemental Spectra

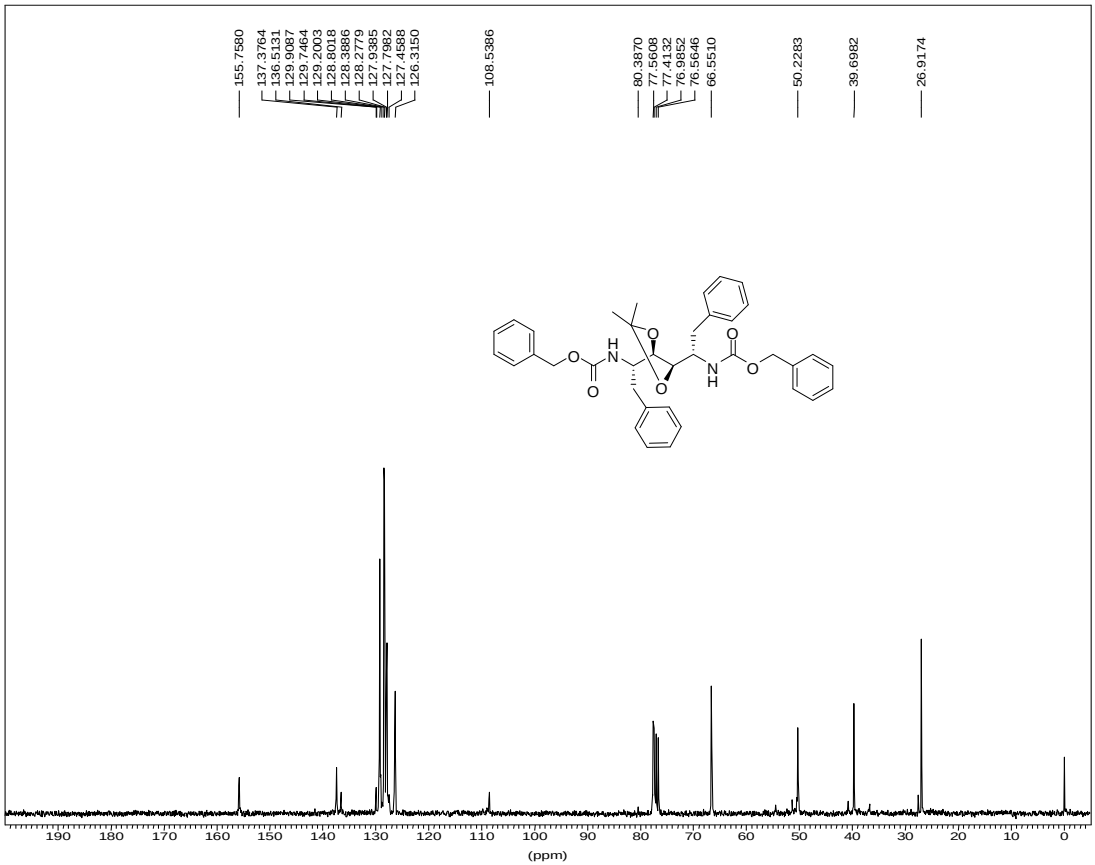


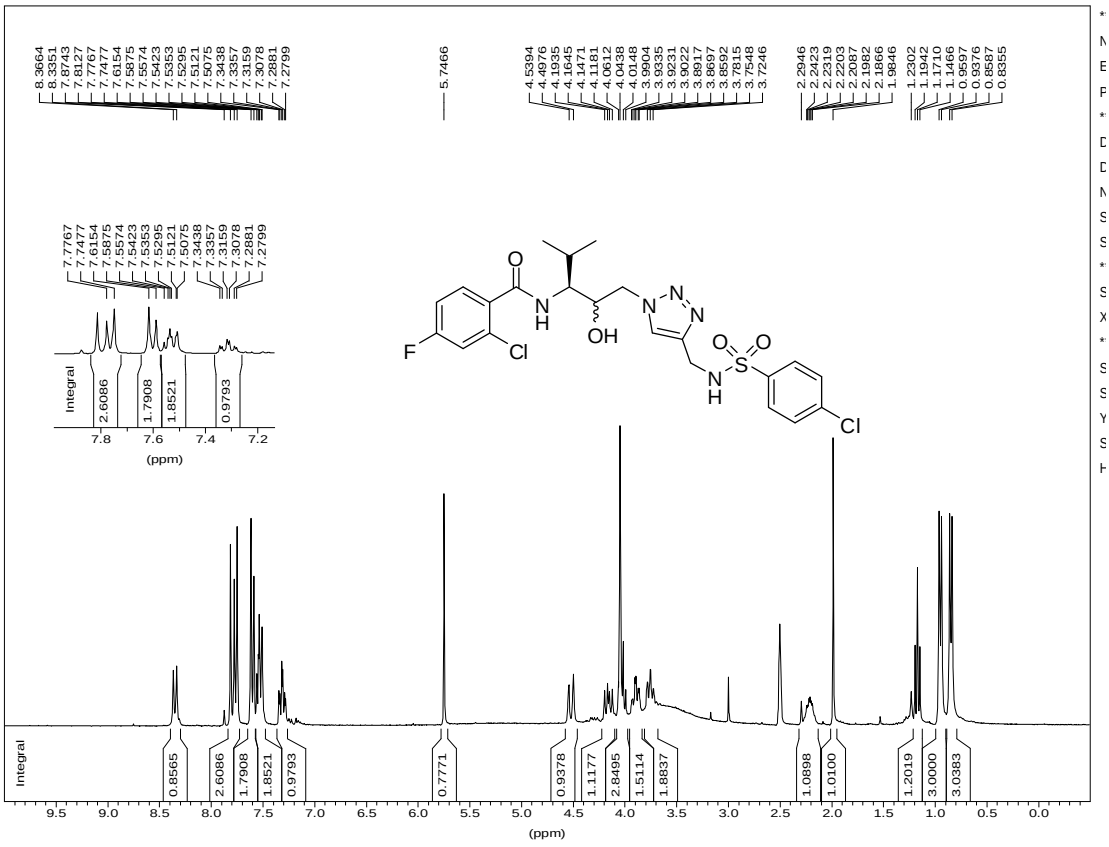


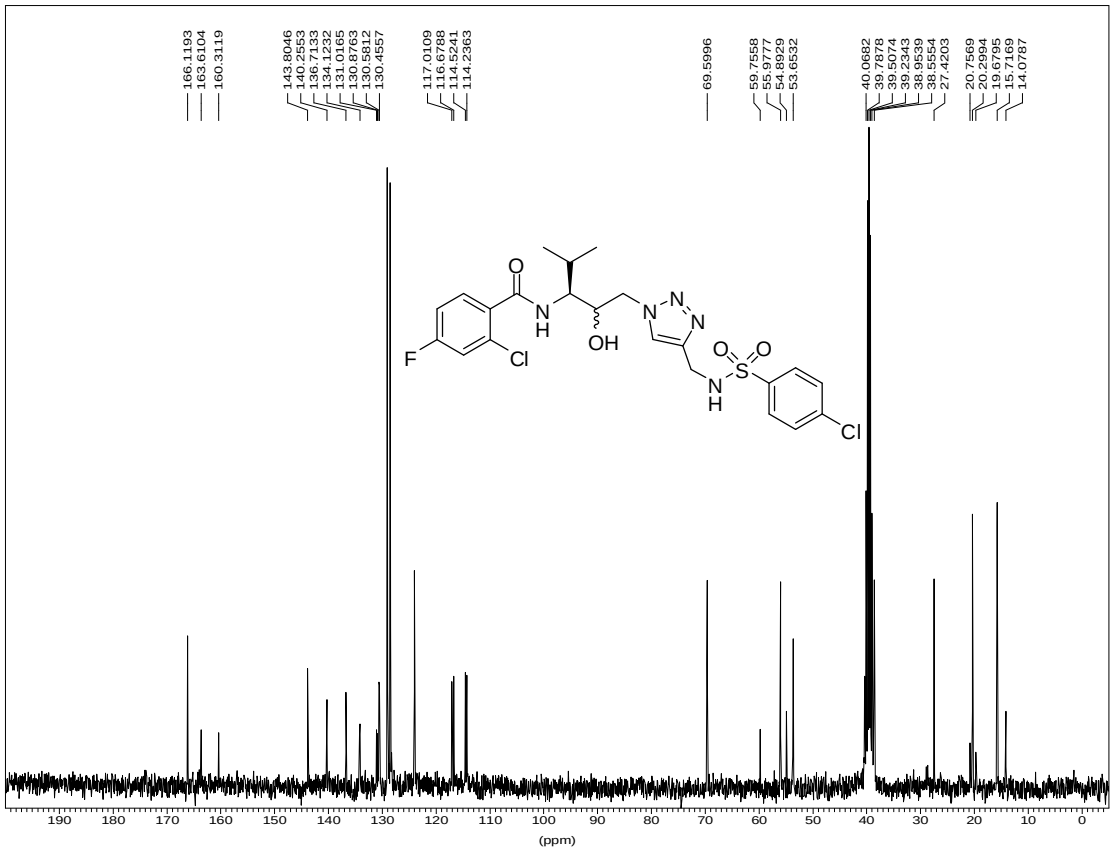

SYM-05
(ppm)
Integral

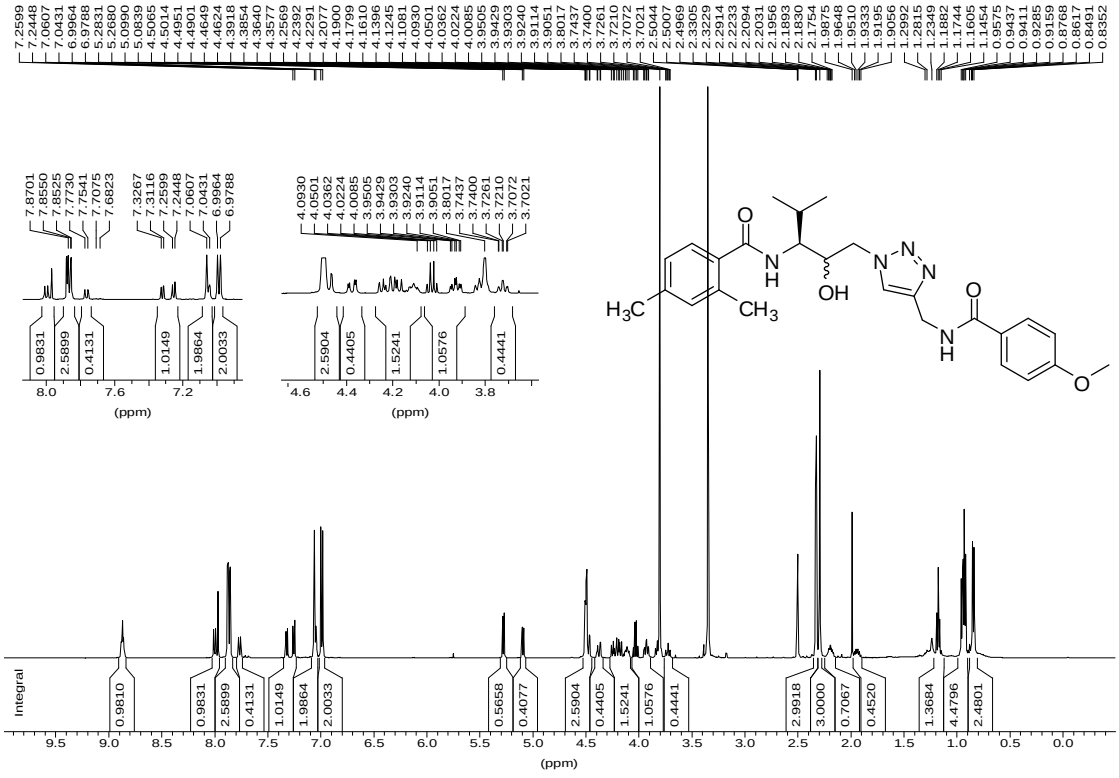

SYM-05
(ppm)

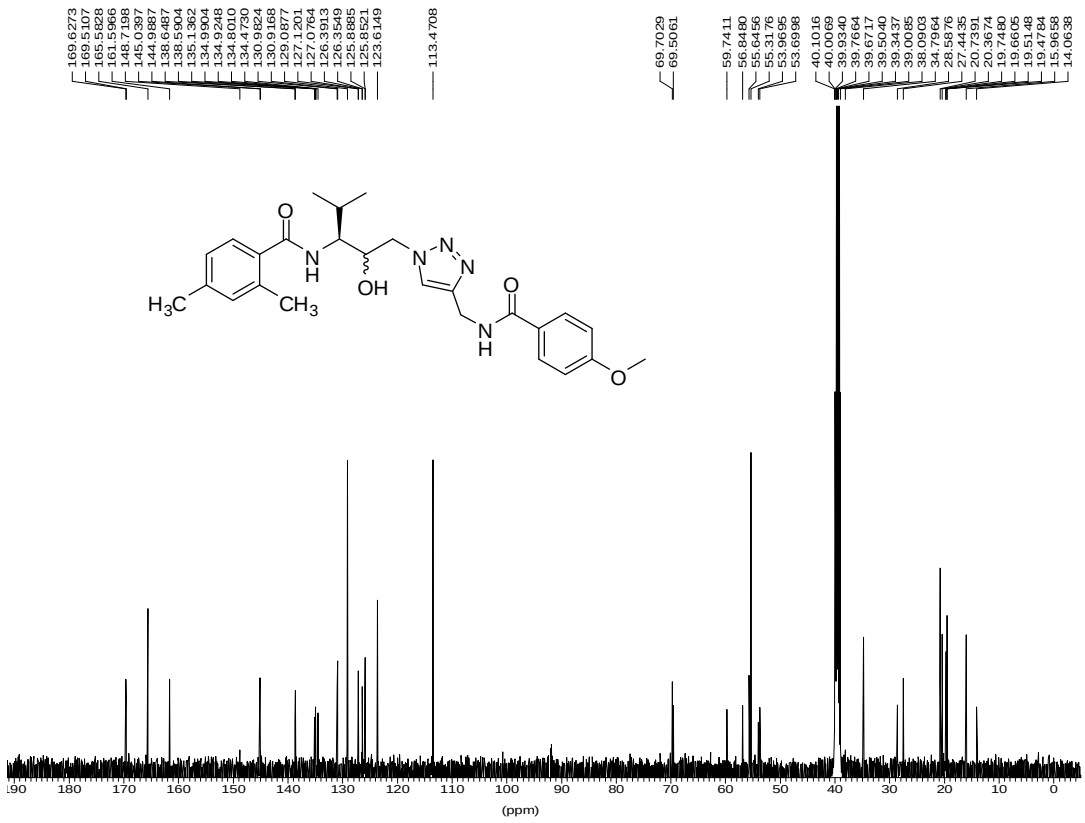

7.2744
7.2547
7.2303
7.2094
5.0353
4.9958
4.9587
4.9192
4.8624
4.6140
4.5849
4.2019
4.0417
4.0220
3.6819
3.0713
3.0353
2.9750
2.9529
2.9065
2.8032
1.7179
1.5972
1.5821
1.5217
1.3882
1.3639
1.3035
1.2571
Integral
24.739
4.0000
1.6855
2.1054
2.0452
3.8716
5.7343
4.0331
(ppm)

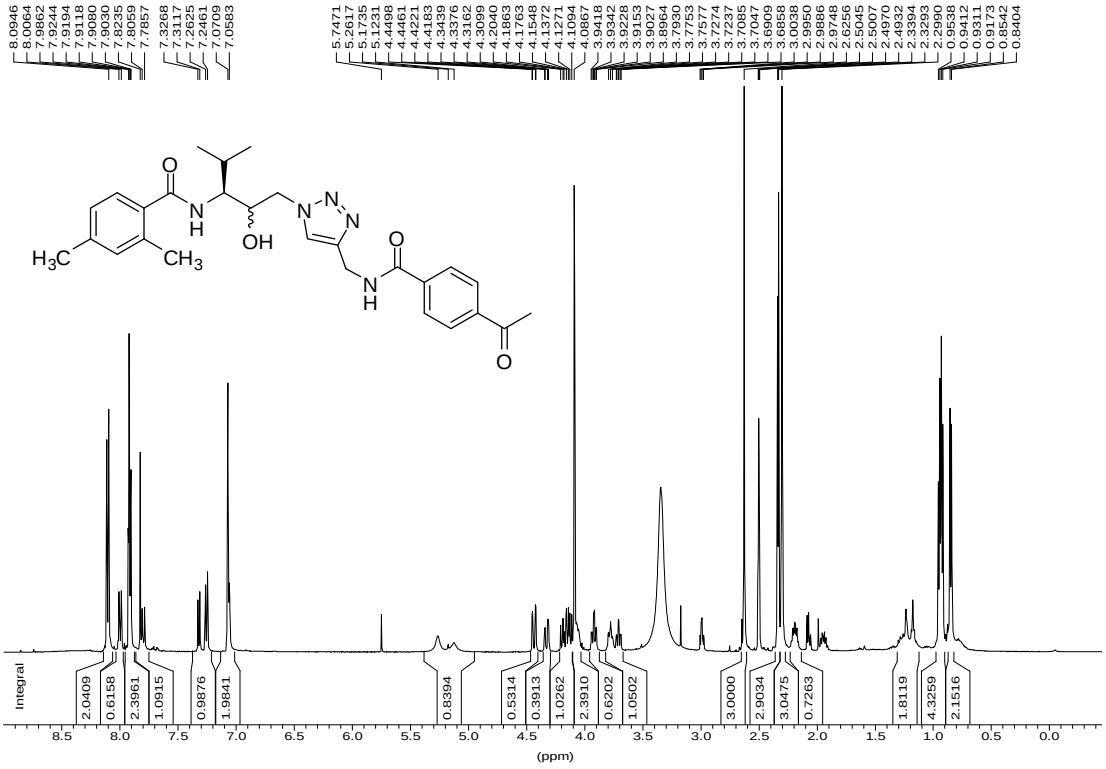

155.7580
137.3764
136.5131
129.9087
129.7464
129.2003
128.8018
128.3886
128.2779
127.9385
127.7982
127.4588
126.3150
108.5386
80.3870
77.5608
77.4132
76.9852
76.5646
66.5510
50.2283
39.6982
26.9174
(ppm)

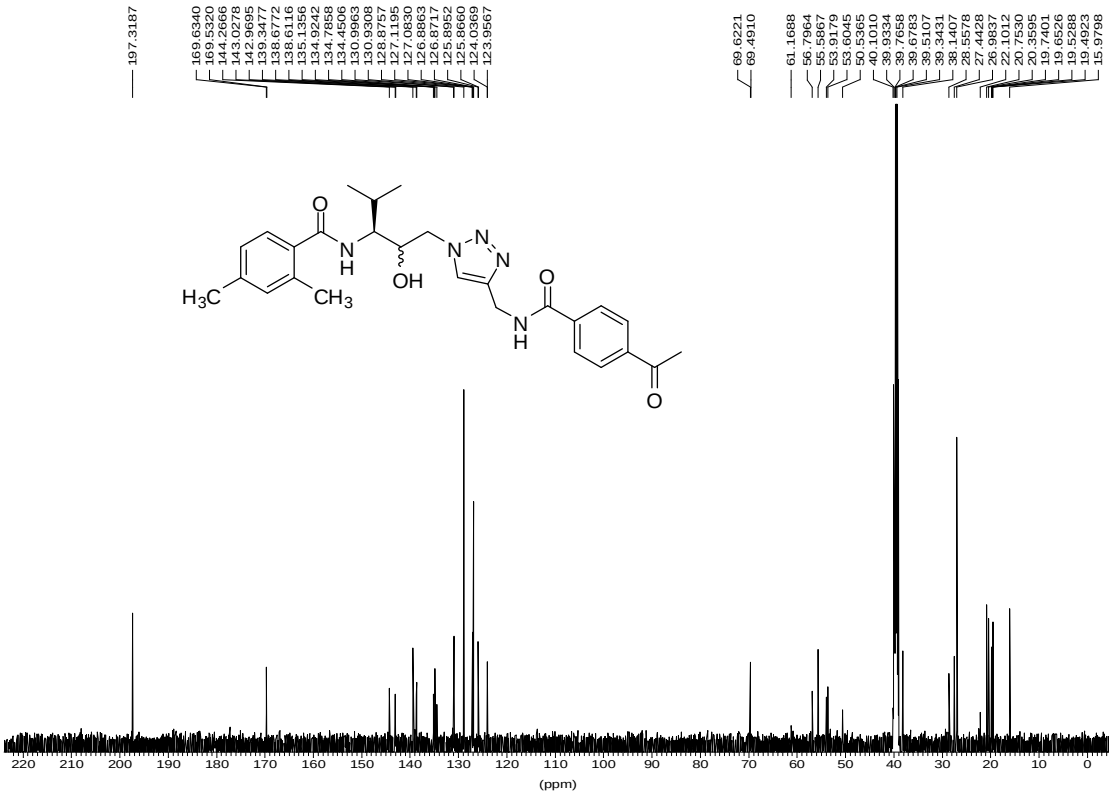

8.3664
8.3351
7.8743
7.8127
7.7767
7.7477
7.6154
7.5875
7.5574
7.5423
7.5353
7.5295
7.5121
7.5075
7.3438
7.3357
7.3159
7.3078
7.2881
7.2799
5.7466
4.5394
4.4976
4.1935
4.1645
4.1471
4.1181
4.0612
4.0438
4.0148
3.9904
3.9335
3.9231
3.9022
3.8917
3.8697
3.8592
3.7845
3.7548
3.7246
2.2946
2.2423
2.2319
2.2203
2.2087
2.1982
2.1866
1.9846
1.2302
1.1942
1.1710
1.1466
0.9597
0.9376
0.8587
0.8355
Integral
2.6086
1.7908
1.8521
0.9793
(ppm)
Integral
0.8565
2.6086
1.7908
1.8521
0.9793
0.7771
0.9378
1.1177
2.8495
1.5114
1.8837
1.0898
1.0100
1.2019
3.0000
3.0383
(ppm)

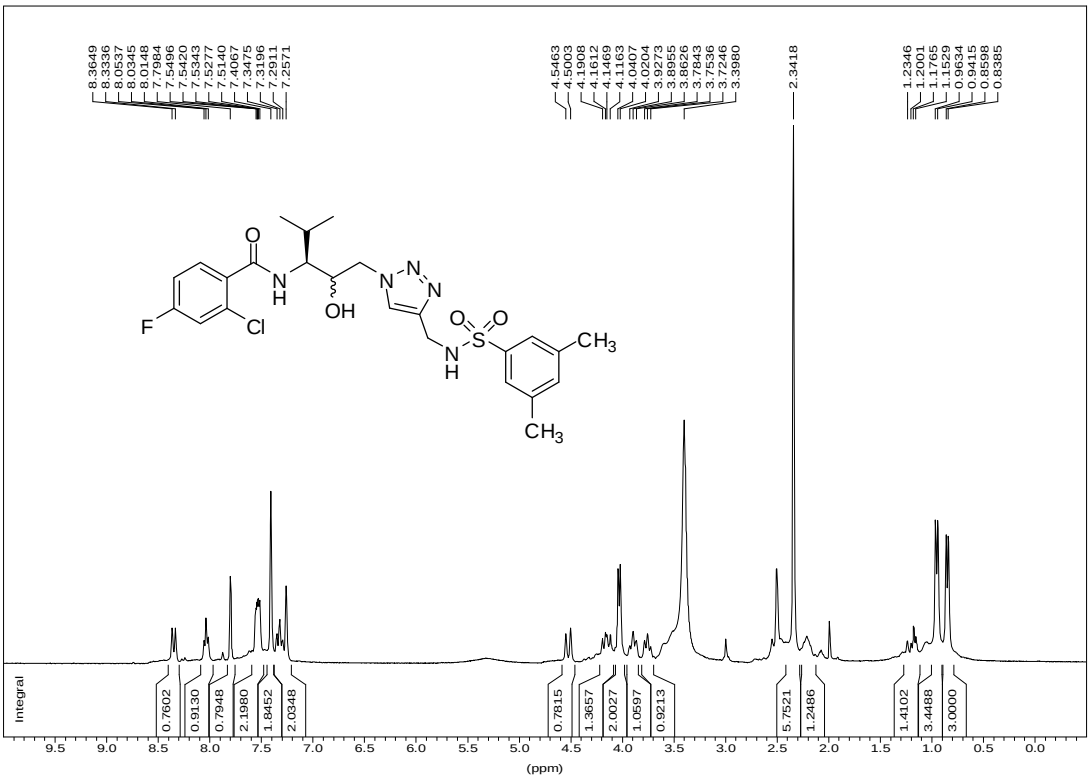

166.1193
163.6104
160.3119
143.8046
140.2553
136.7133
134.1232
131.0165
130.9763
130.5812
130.4557
117.0109
116.6788
114.5241
114.2363
69.5996
59.7558
55.9777
54.8929
53.6532
40.0682
39.7878
39.5074
39.2343
38.9539
38.5554
27.4203
20.7569
20.2994
19.6795
15.7169
14.0787
(ppm)

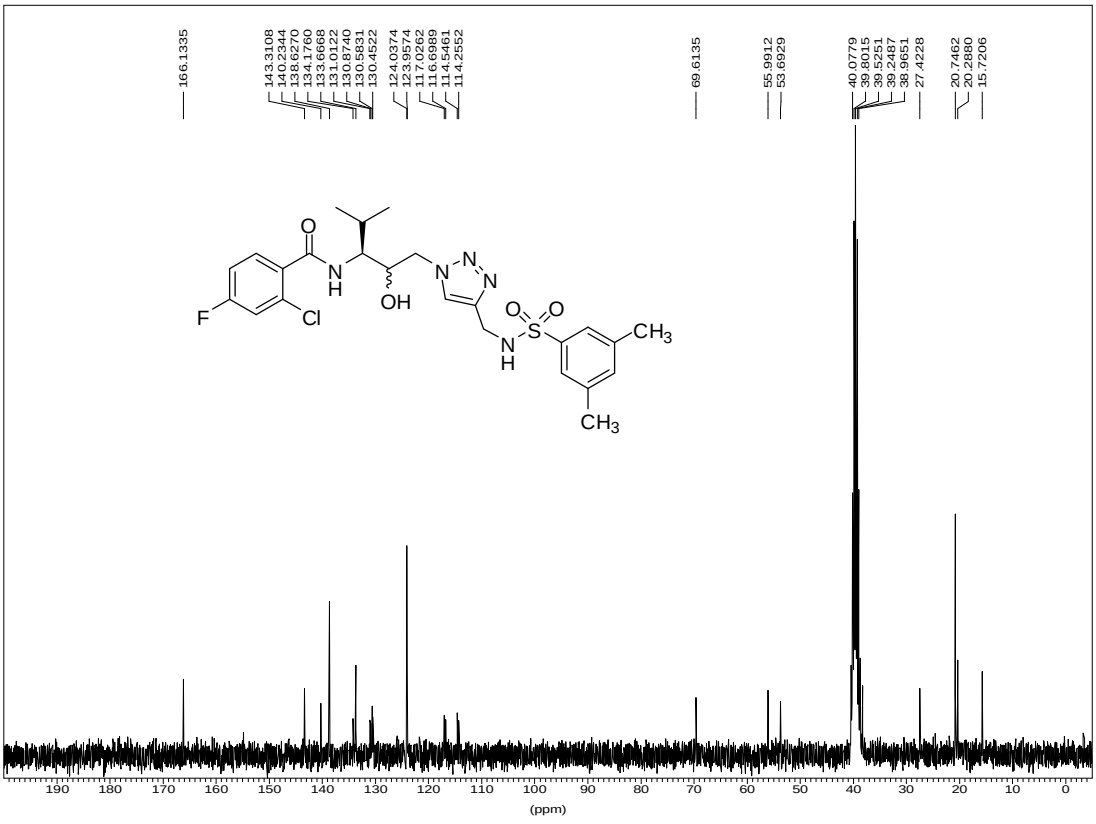

Integral
(ppm)

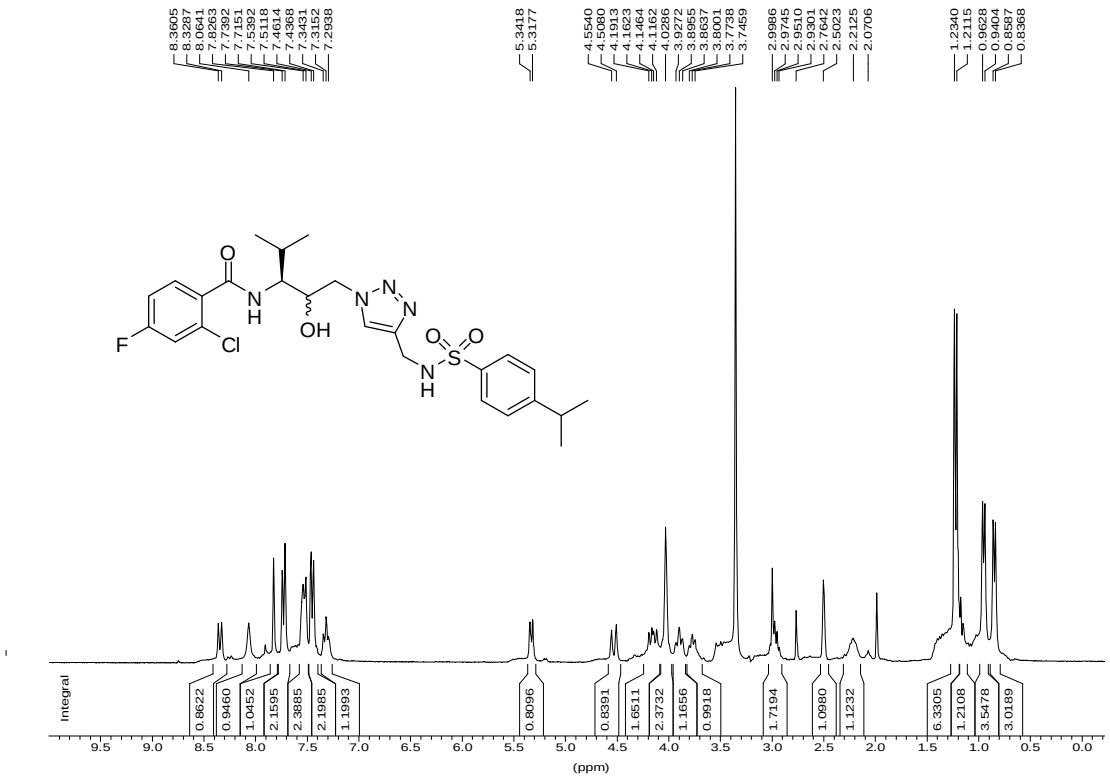

(ppm)

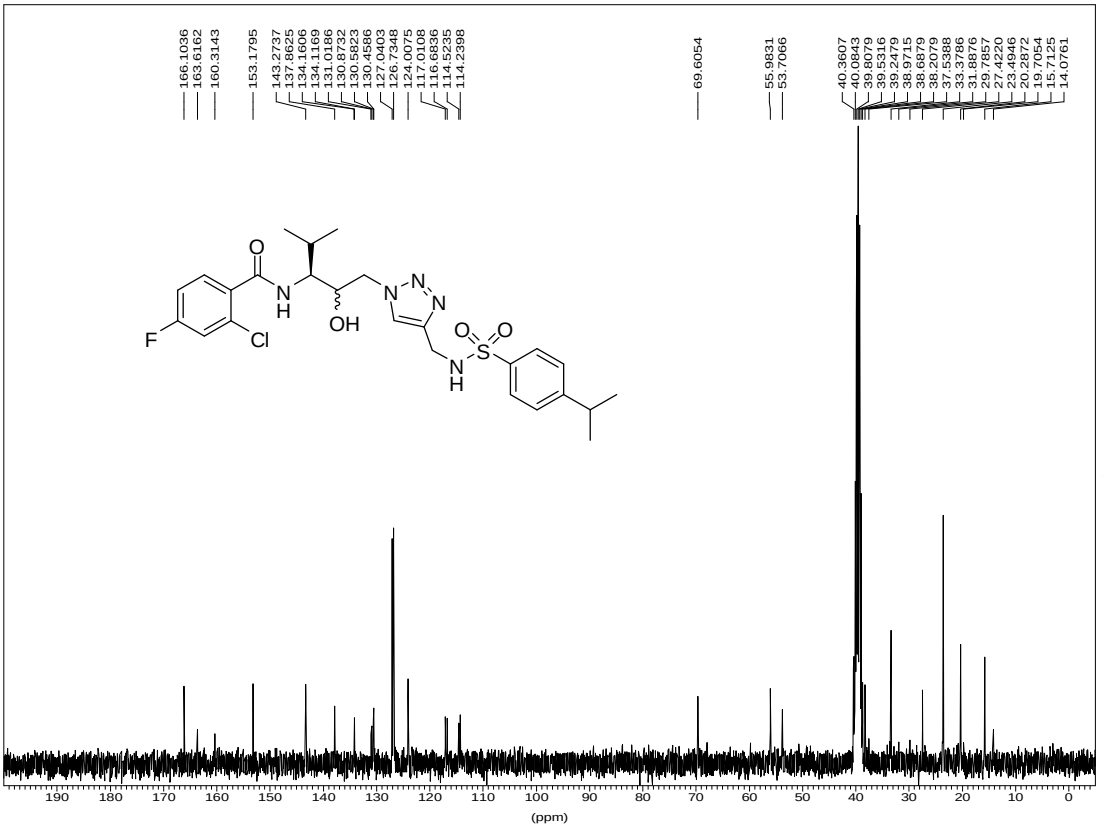

8.0946
8.0064
7.9862
7.9244
7.9194
7.9118
7.9080
7.9030
7.8235
7.8059
7.7857
7.3268
7.3117
7.2625
7.2461
7.0709
7.0583
5.7471
5.2617
5.1735
5.1231
4.4498
4.4461
4.4221
4.4183
4.3439
4.3376
4.3162
4.3099
4.2040
4.1863
4.1763
4.1548
4.1372
4.1271
4.1094
4.0867
3.9418
3.9342
3.9228
3.9153
3.9027
3.8964
3.7930
3.7753
3.7577
3.7274
3.7237
3.7085
3.7047
3.6909
3.6858
3.0038
2.9950
2.9986
2.9748
2.9756
2.5045
2.5007
2.4970
2.4932
2.3394
2.3293
2.2990
0.9538
0.9412
0.9311
0.9173
0.8542
0.8404
integral
2.0409
0.6158
2.3961
1.0915
0.9876
1.9841
0.8394
0.5314
0.3913
1.0262
2.3910
0.6202
1.0502
3.0000
2.9034
3.0475
0.7263
1.8119
4.3259
2.1516
(ppm)

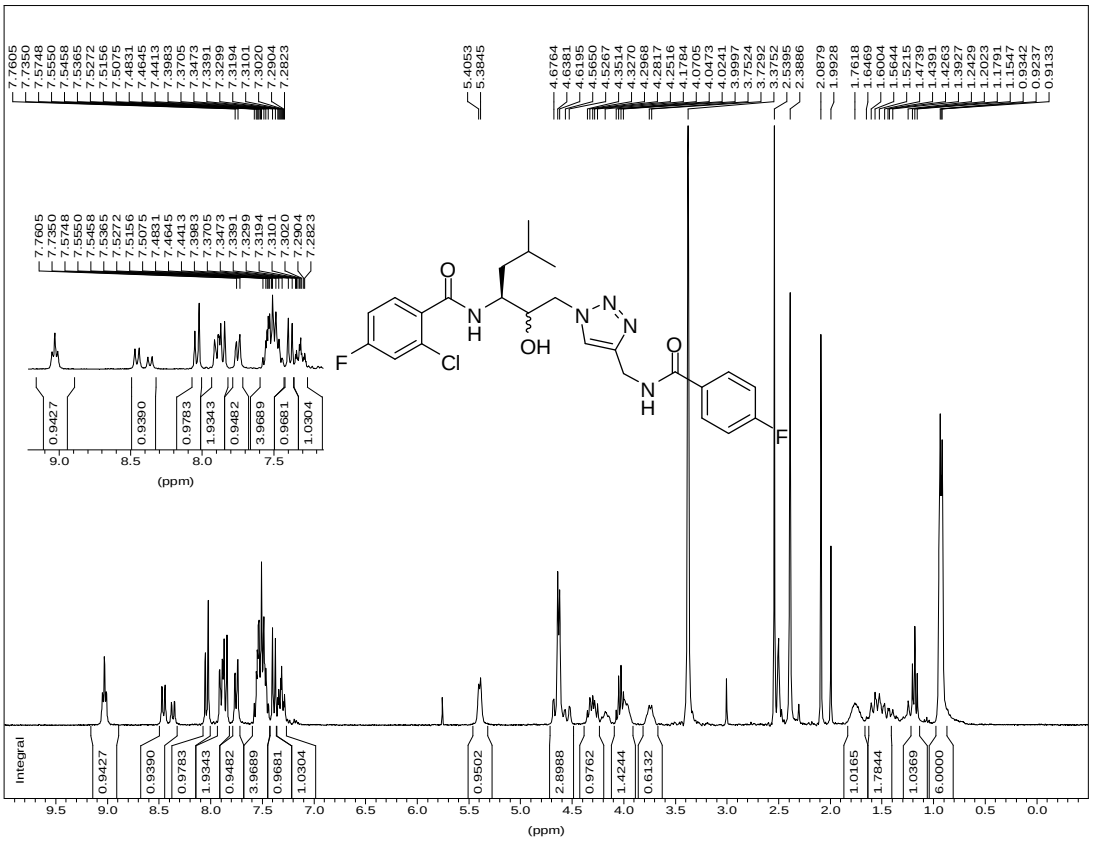

197.3187
169.6340
169.5320
144.2666
143.0278
142.9695
139.3477
138.6772
138.6116
135.1356
134.7642
134.7258
134.4506
130.9963
130.9308
128.8757
127.1195
127.0830
126.8863
126.8717
125.8952
125.8660
124.0369
123.9567
69.6221
69.4910
61.1688
56.7964
55.5867
53.9179
53.6045
53.5365
40.1010
39.9334
39.7658
39.6783
39.5107
39.3431
38.1407
28.5578
27.4428
26.9837
22.1012
20.7530
20.3595
19.7401
19.6526
19.5288
19.4923
15.9798
(ppm)

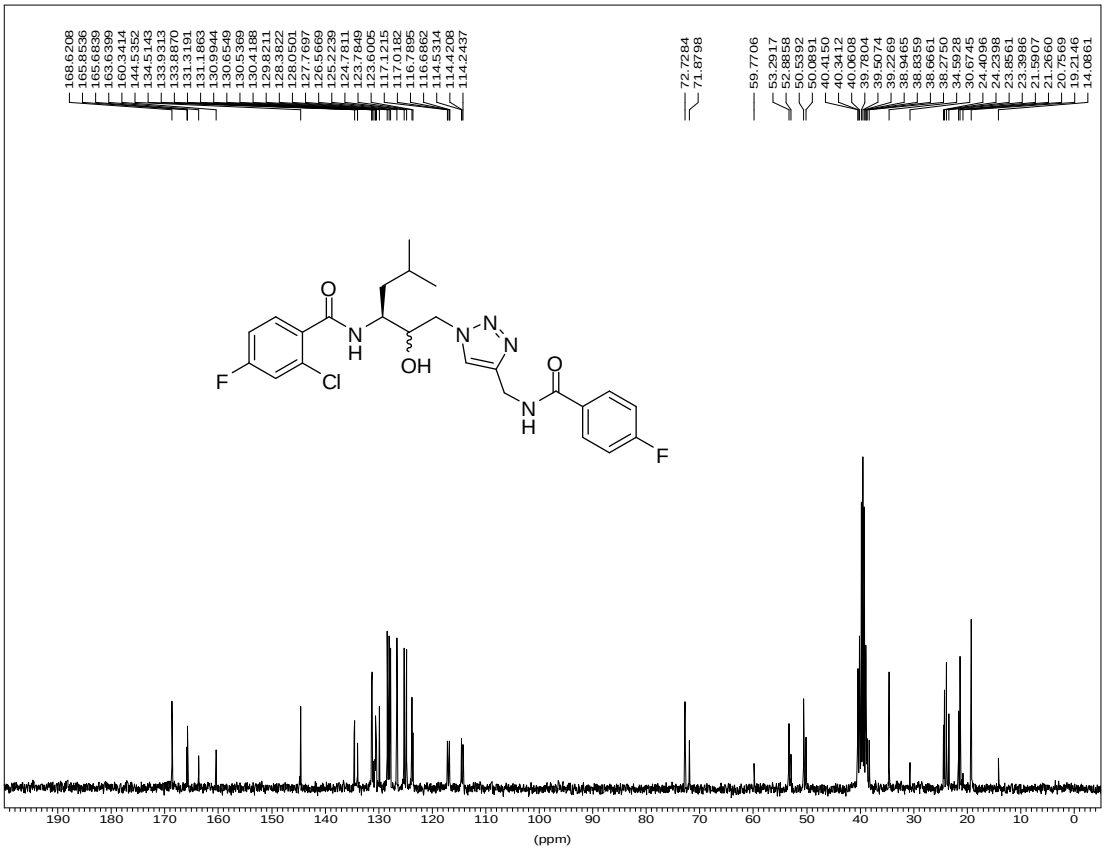

8.3649
8.3436
8.0537
8.0345
8.0148
7.7984
7.5496
7.5420
7.5343
7.5277
7.5140
7.4067
7.3475
7.3196
7.2911
7.2571
4.5463
4.5003
4.1908
4.1612
4.1469
4.1163
4.0407
4.0204
3.9273
3.8955
3.8626
3.7843
3.7536
3.7246
3.3980
2.3418
1.2346
1.2001
1.1765
1.1529
0.9634
0.9415
0.8598
0.8385
Integral
0.7602
0.9130
0.7948
2.1980
1.8452
2.0348
0.7815
1.3657
2.0027
1.0597
0.9213
5.7521
1.2486
1.4102
3.4488
3.0000
(ppm)


166.1335
143.3108
140.2344
138.6270
134.1760
133.6668
131.0122
130.8740
130.6531
130.4522
124.0374
123.9574
117.0262
116.6989
114.5461
114.2552
69.6135
55.9912
53.6929
40.0779
39.8015
39.5251
39.2487
38.9651
27.4228
20.7462
20.2880
15.7206
(ppm)

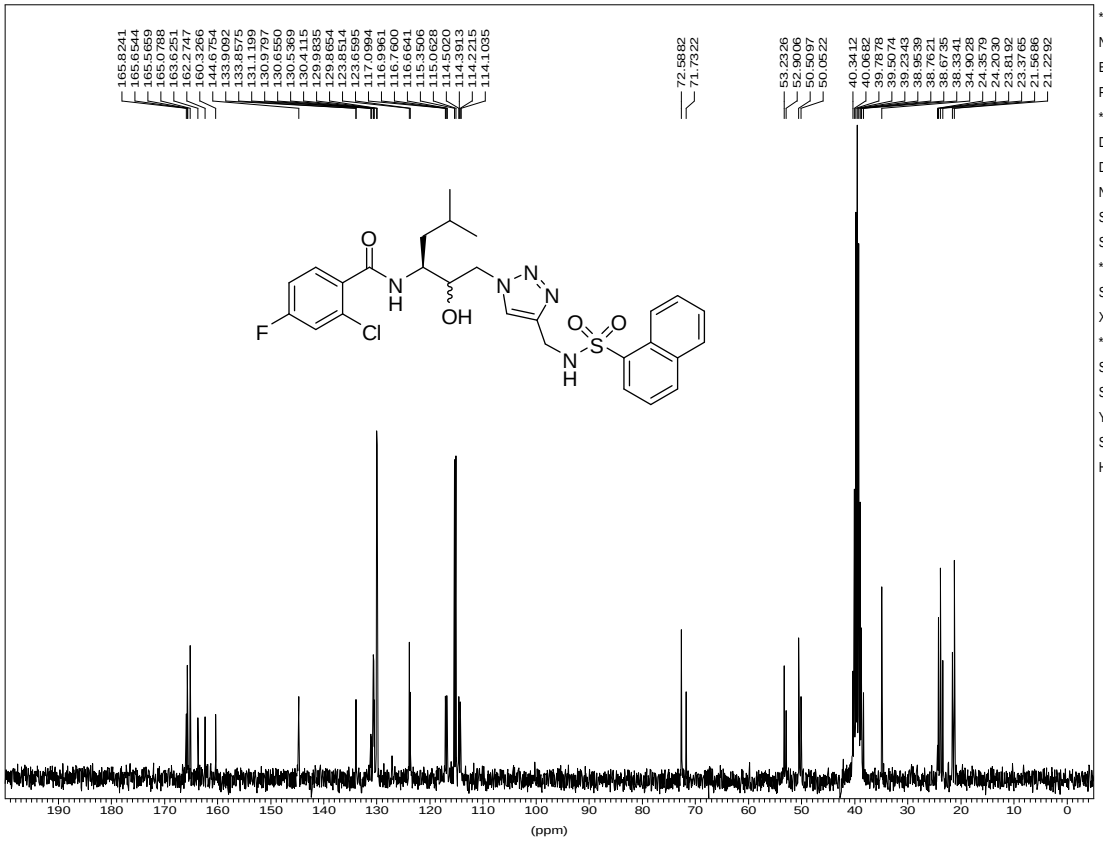

8.3605
8.3287
8.0641
7.8263
7.7392
7.7151
7.5392
7.5118
7.4614
7.4368
7.3431
7.3152
7.2938
5.3418
5.3177
4.5540
4.5080
4.1913
4.1623
4.1464
4.1162
4.0296
3.9272
3.8955
3.8637
3.8001
3.7738
3.7459
2.9986
2.9745
2.9510
2.9301
2.7642
2.5023
2.2125
2.0706
1.2340
1.2115
0.9628
0.9404
0.8587
0.8368
Integral
0.8622
0.9460
1.0452
2.1595
2.3885
2.1985
1.1993
0.8096
0.8391
1.6511
2.3732
1.1656
0.9918
1.7194
1.0980
1.1232
6.3305
1.2108
3.5478
3.0189
(ppm)

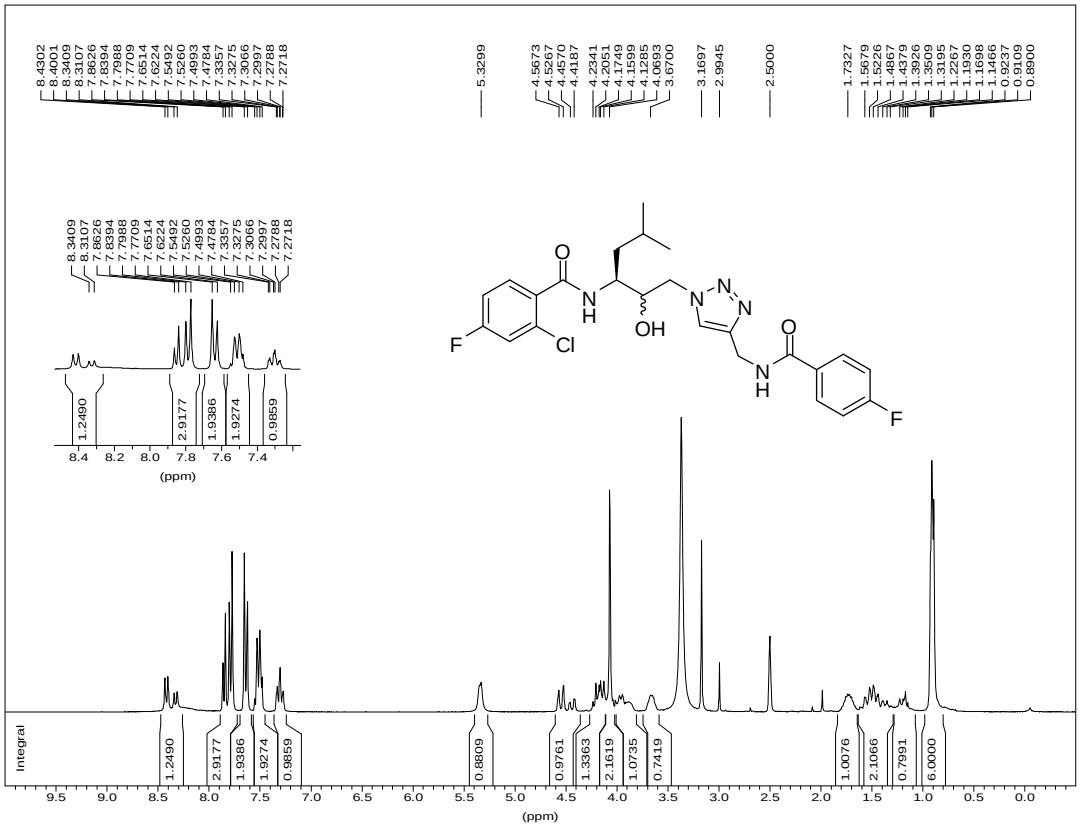

166.1036
163.6162
160.3143
153.1795
143.2737
137.8625
134.1606
134.1169
131.0186
130.8732
130.5823
130.4586
127.0403
126.7348
124.0075
117.0108
116.6836
114.5235
114.2398
69.6054
55.9831
53.7066
40.3607
40.0843
39.8079
39.5316
39.2479
38.9715
38.6879
38.2079
37.5388
33.3786
31.8876
29.7857
27.4220
23.4946
20.2872
19.7054
15.7125
14.0761
(ppm)

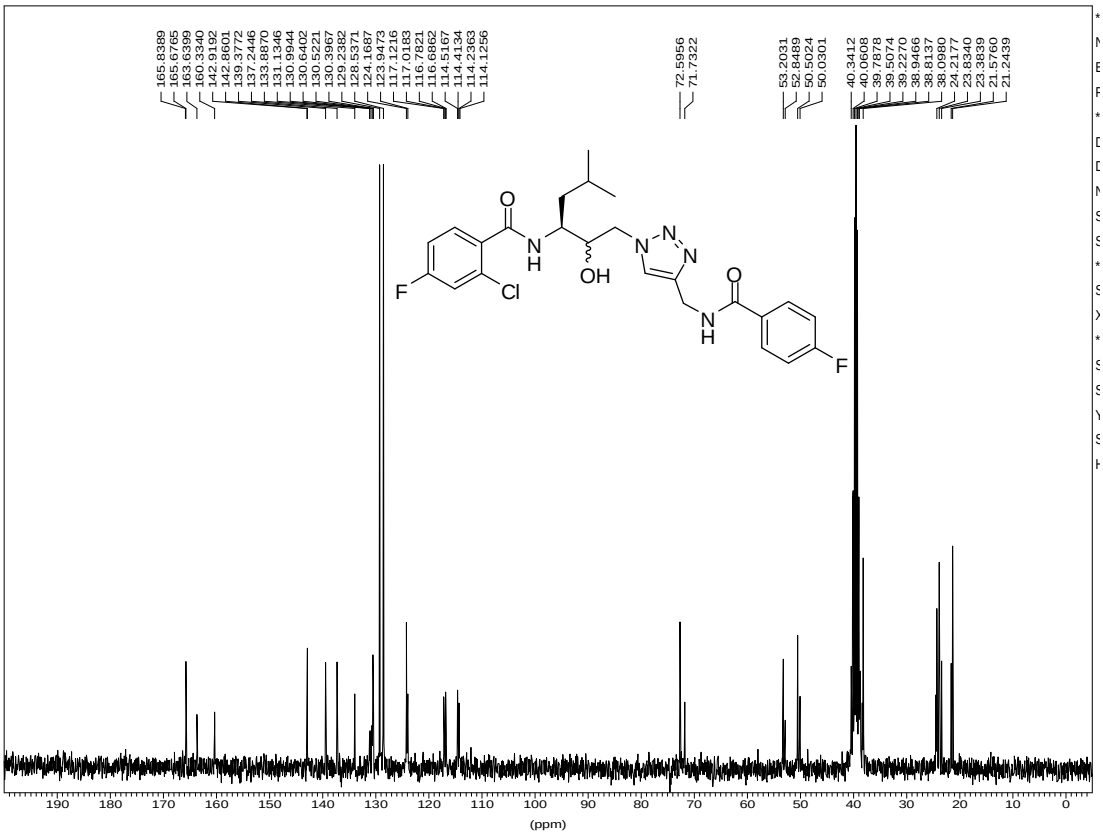

7.7605
7.7350
7.5748
7.5550
7.5458
7.5365
7.5272
7.5156
7.5075
7.4831
7.4645
7.4413
7.3983
7.3705
7.3473
7.3391
7.3299
7.3194
7.3101
7.3020
7.2904
7.2823
5.4053
5.3845
4.6764
4.6381
4.6195
4.5650
4.5267
4.3514
4.3270
4.2968
4.2817
4.2516
4.1784
4.0705
4.0473
4.0241
3.9997
3.7524
3.7292
3.3752
2.5395
2.3886
2.0879
1.9928
1.7618
1.6469
1.6004
1.5644
1.5215
1.4739
1.4391
1.4263
1.3927
1.2429
1.2023
1.1791
1.1547
0.9342
0.9237
0.9133
0.9427
0.9390
0.9783
1.9343
0.9482
3.9689
0.9681
1.0304
(ppm)
Integral
0.9427
0.9390
0.9783
1.9343
0.9482
3.9689
0.9681
1.0304
0.9502
2.8988
0.9762
1.4244
0.6132
1.0165
1.7844
1.0369
6.0000
(ppm)

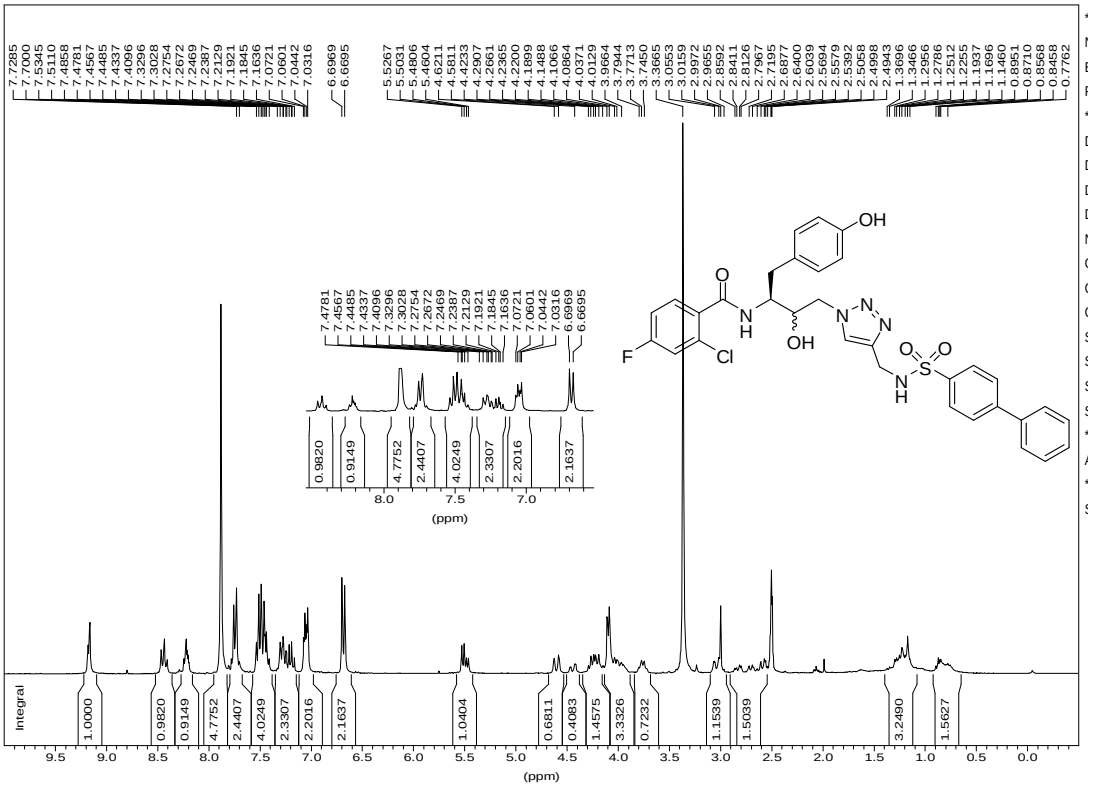

168.6208
165.8536
165.6339
163.6399
160.3414
144.5352
134.5143
133.9313
133.8870
131.3191
131.1863
130.9944
130.6549
130.5369
130.4188
129.8211
128.3822
128.0501
127.7697
126.5669
125.2239
124.7811
123.7849
123.6005
117.1215
117.0182
116.7895
116.6862
114.5314
114.4208
114.2437
72.7284
71.8798
59.7706
53.2917
52.8858
50.5392
50.0891
40.4150
40.3412
40.2673
39.7804
39.5074
39.2269
38.9465
38.8359
38.6661
38.2750
34.5928
30.6745
24.4096
24.2398
23.8561
23.3986
21.5907
21.2660
20.7569
19.2146
14.0861
(ppm)

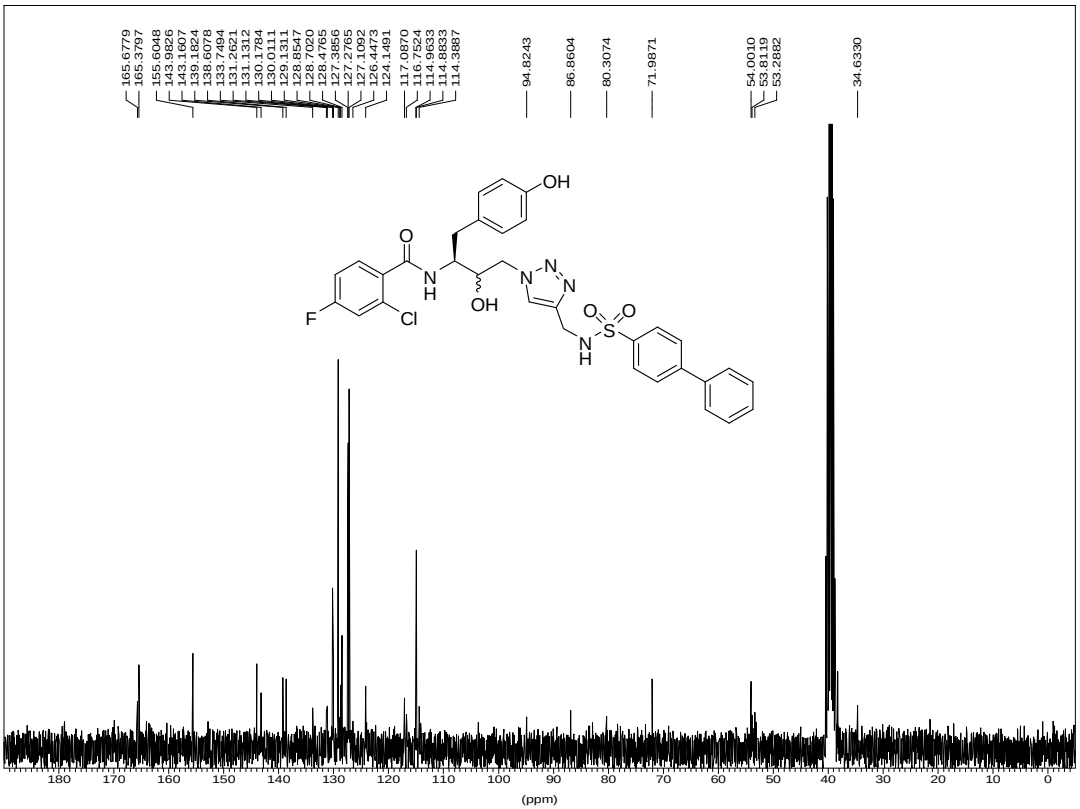

Integral
0.9173
0.9156
2.8460
1.8739
2.8723
2.9029
1.5246
1.3159
0.8483
0.9718
3.0178
6.0000
(ppm)

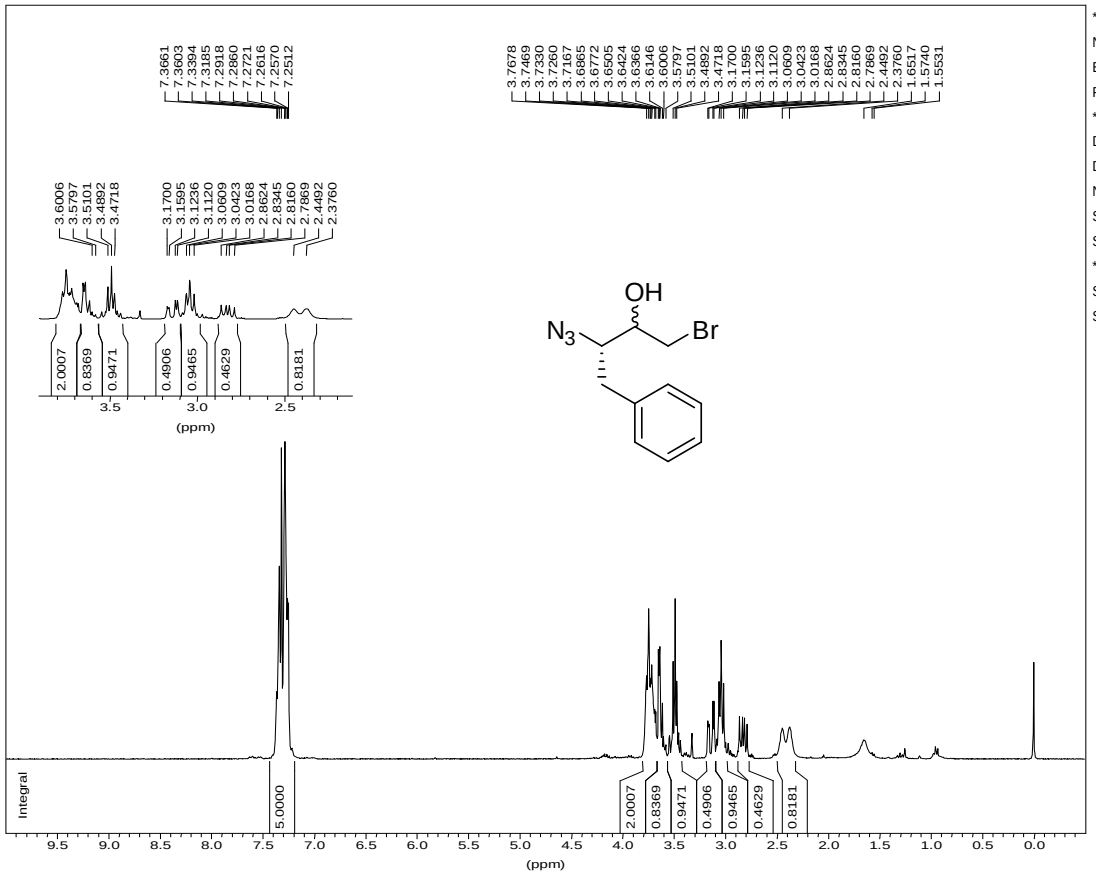

(ppm)

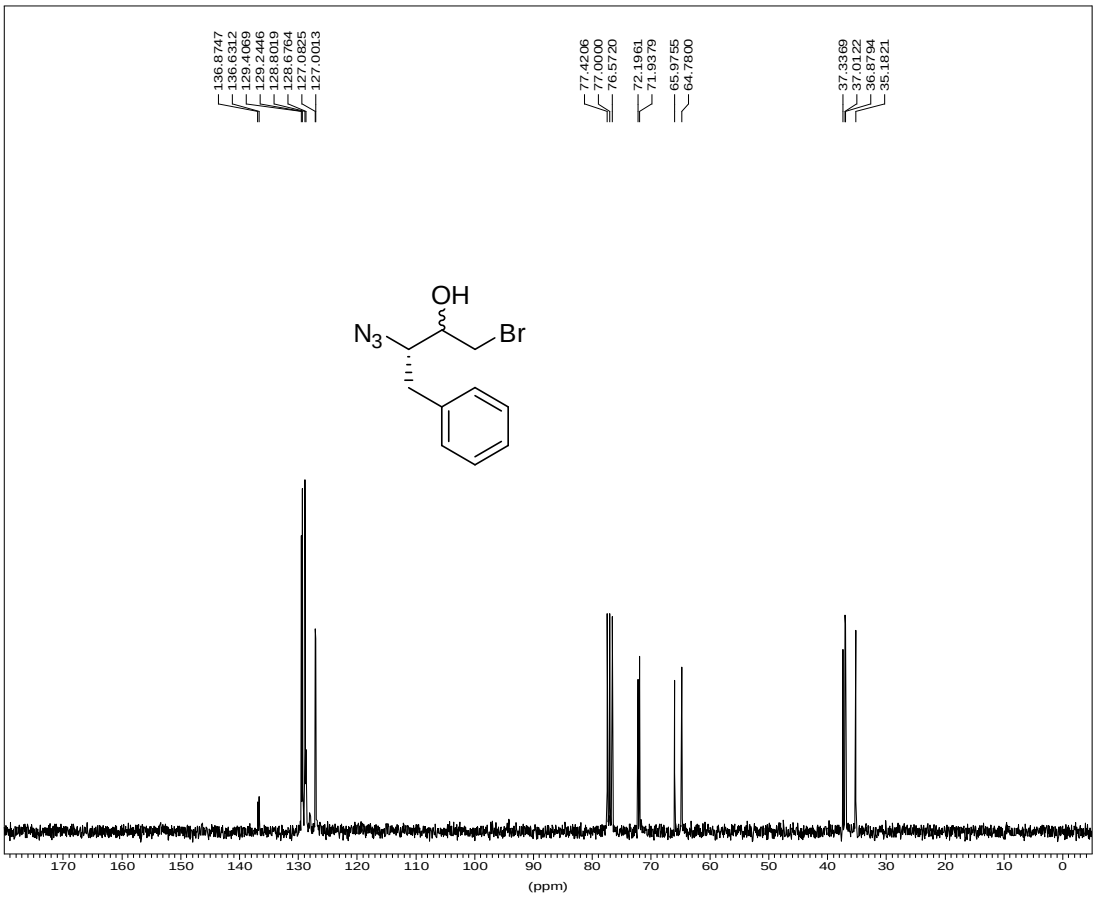

8.4302
8.4001
8.3409
8.3107
7.8626
7.8394
7.7988
7.7709
7.6514
7.6224
7.5492
7.5260
7.4993
7.4784
7.3357
7.3275
7.3066
7.2997
7.2788
7.2718
5.3299
4.5673
4.5267
4.4570
4.4187
4.2341
4.2051
4.1749
4.1599
4.1285
4.0693
3.6700
3.1697
2.9945
2.5000
1.7327
1.5679
1.5276
1.4867
1.4379
1.3926
1.3509
1.3195
1.2267
1.1930
1.1698
1.1466
0.9237
0.9109
0.8900
O
N
H
OH
N
N
N
F
Cl
O
N
H
F
1.2490
2.9177
1.9386
1.9274
0.9859
8.4
8.2
8.0
7.8
7.6
7.4
(ppm)
Integral
1.2490
2.9177
1.9386
1.9274
0.9859
0.8809
0.9761
1.3363
2.1619
1.0735
0.7419
1.0076
2.1066
0.7991
6.0000
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
(ppm)

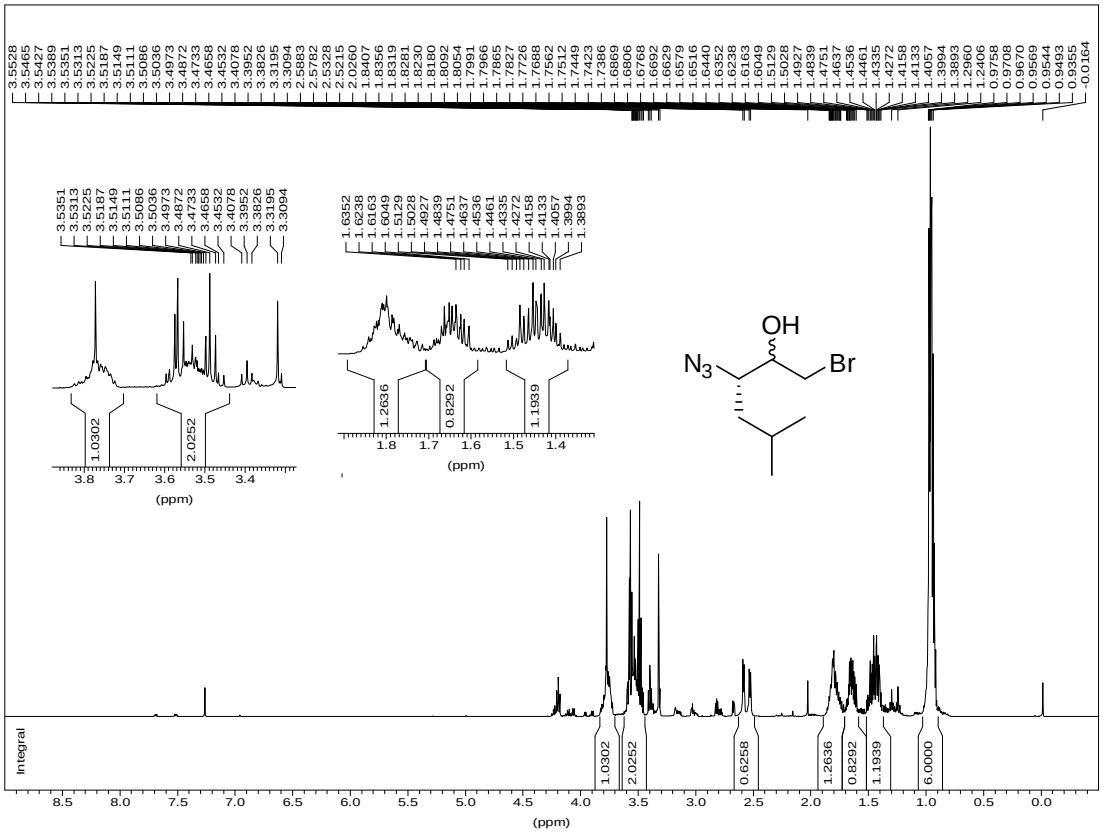

165.8389
165.6765
163.6399
160.3340
142.9192
142.8601
139.3772
137.2446
133.8870
131.2246
130.9944
130.6402
130.5221
130.3967
129.2382
128.5371
124.1687
123.9473
117.1216
117.0183
116.7821
116.6662
114.5167
114.4134
114.2363
114.1256
72.5956
71.7322
53.2031
52.8499
52.5024
50.0301
40.3412
40.0608
39.7878
39.5074
39.2270
38.9466
38.8137
38.0980
24.2177
23.8340
23.3839
21.5760
21.2439
O
N
H
OH
N
N
N
F
Cl
O
N
H
F
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

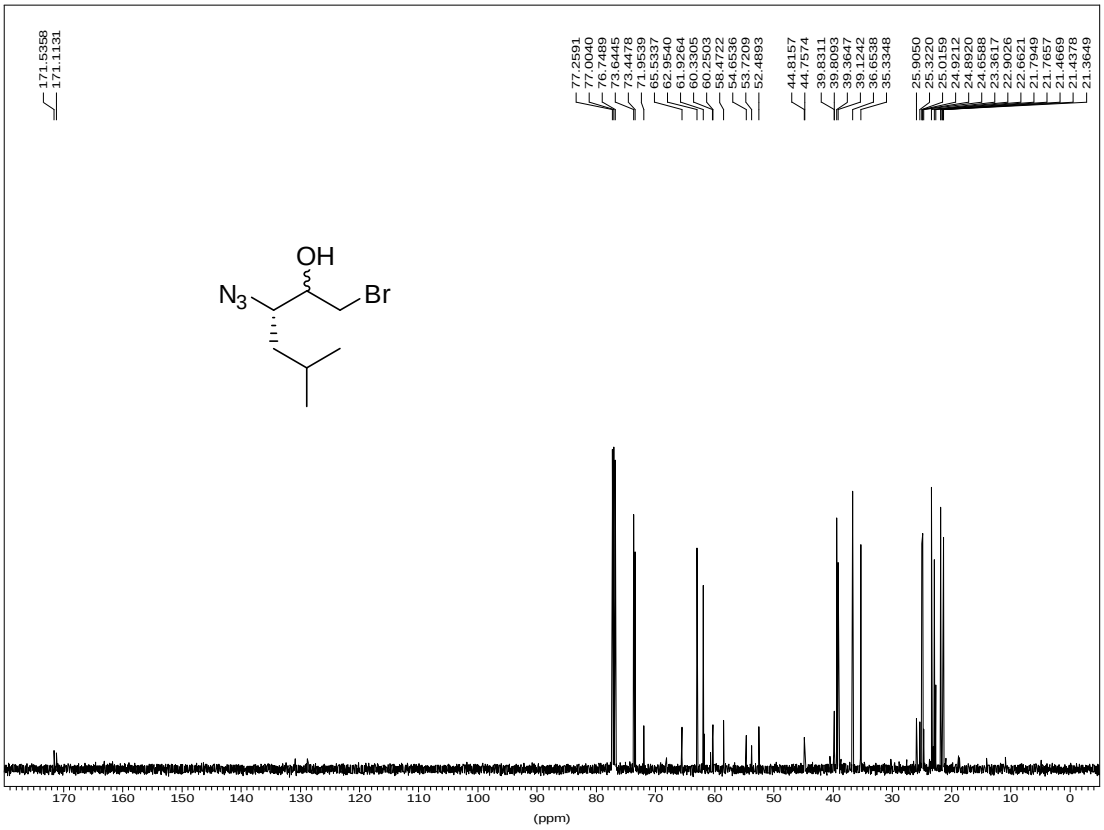

integral
(ppm)
OH
O
N
H
OH
F
Cl
N
N
N
O
O
S
N
H

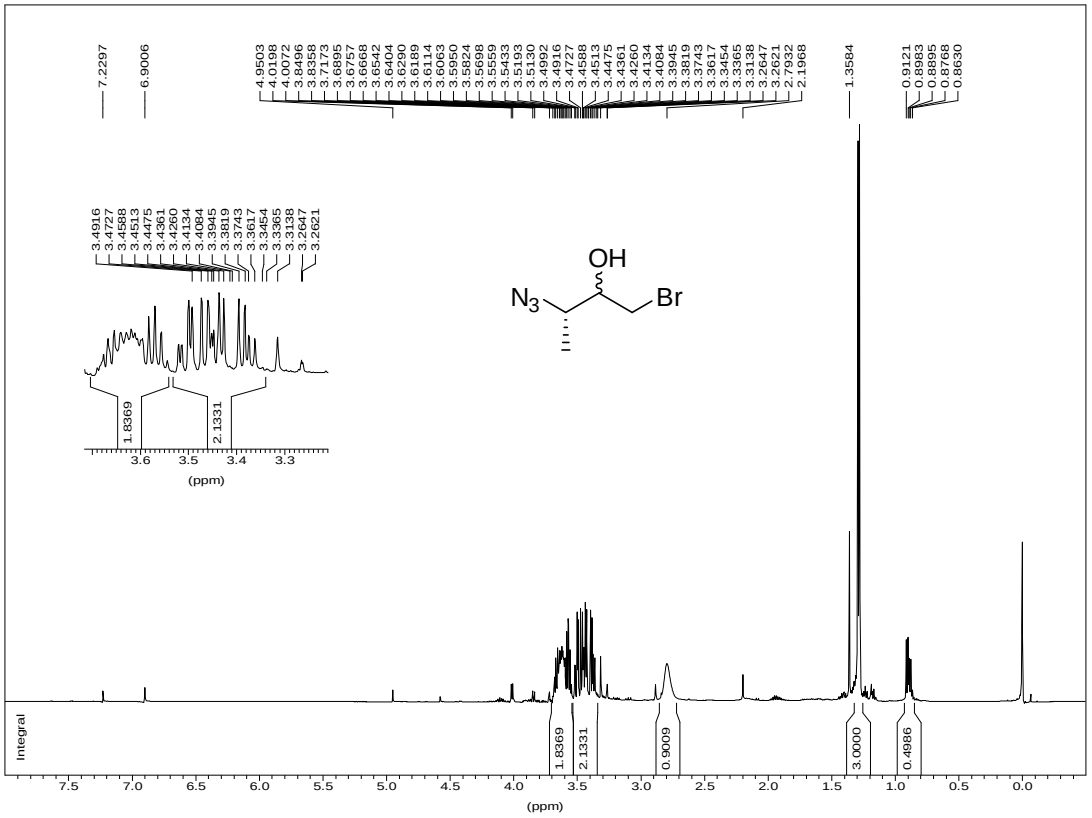

(ppm)
OH
O
N
H
OH
F
Cl
N
N
N
O
O
S
N
H

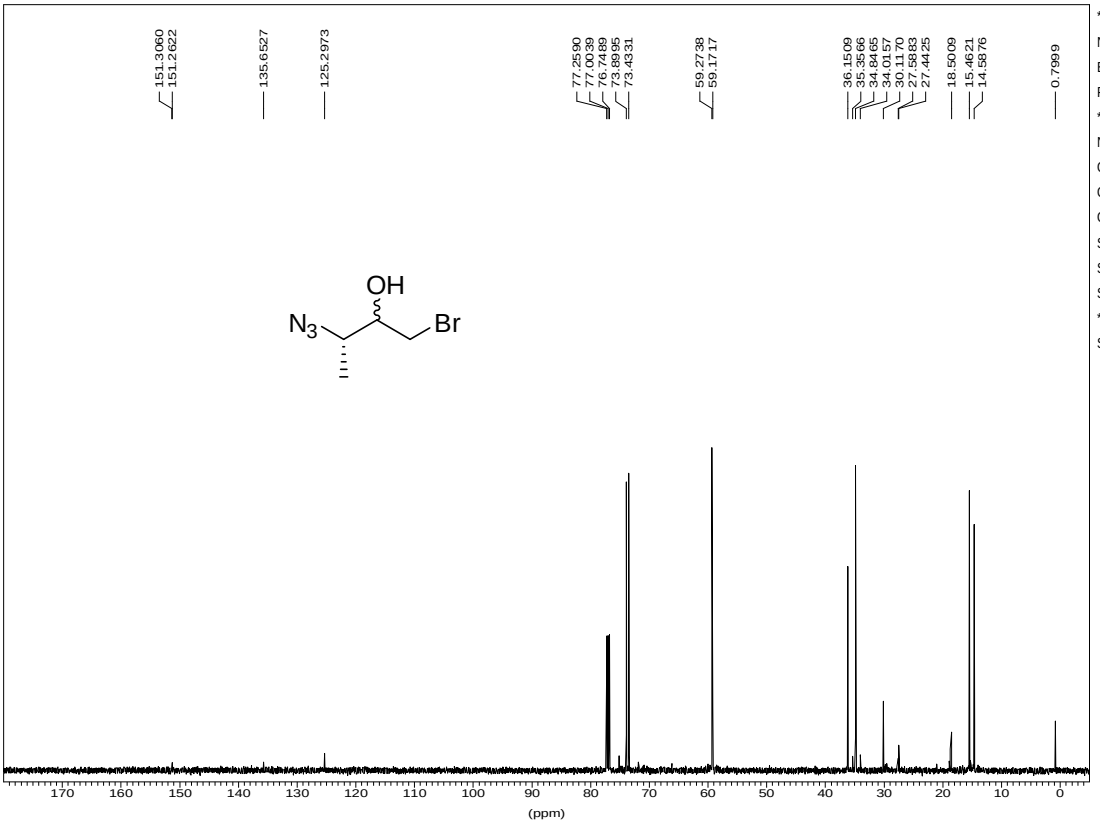

7.3661
7.3603
7.3394
7.3185
7.2918
7.2860
7.2721
7.2616
7.2570
7.2512
3.7678
3.7469
3.7330
3.7260
3.7167
3.6865
3.6772
3.6505
3.6424
3.6366
3.6146
3.6006
3.5797
3.5101
3.4892
3.4718
3.1700
3.1595
3.1236
3.1120
3.0609
3.0423
3.0168
2.8624
2.8345
2.8160
2.7869
2.4492
2.3760
1.6517
1.5740
1.5531
OH
N3
Br
Integral
5.0000
2.0007
0.8369
0.9471
0.4906
0.9465
0.4629
0.8181
(ppm)

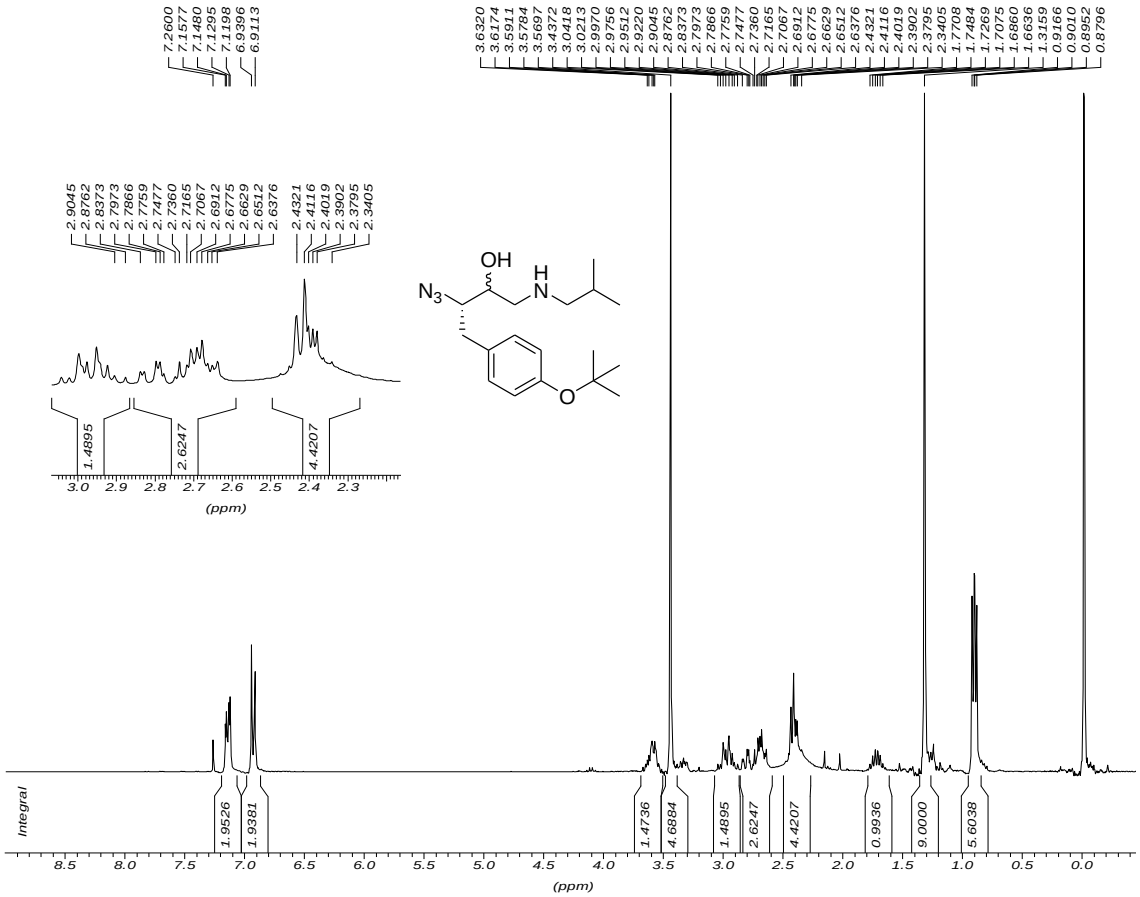

136.8747
136.6312
129.4069
129.2446
128.8019
128.6764
127.0825
127.0013
77.4206
77.0000
76.5720
72.1961
71.9379
65.9755
64.7800
37.3369
37.0122
36.8794
35.1821
OH
N3
Br
(ppm)

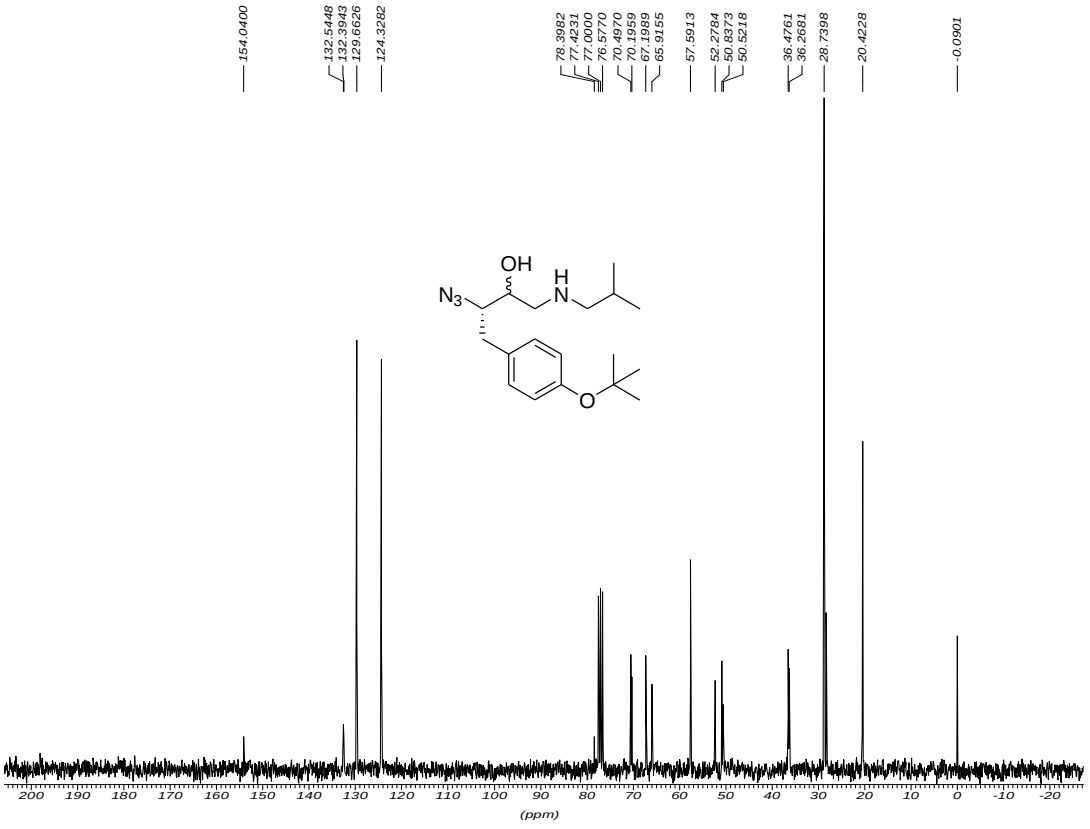

OH
N3
Br
Integral
1.0302
2.0252
0.6258
1.2636
0.8292
1.1939
6.0000
(ppm)

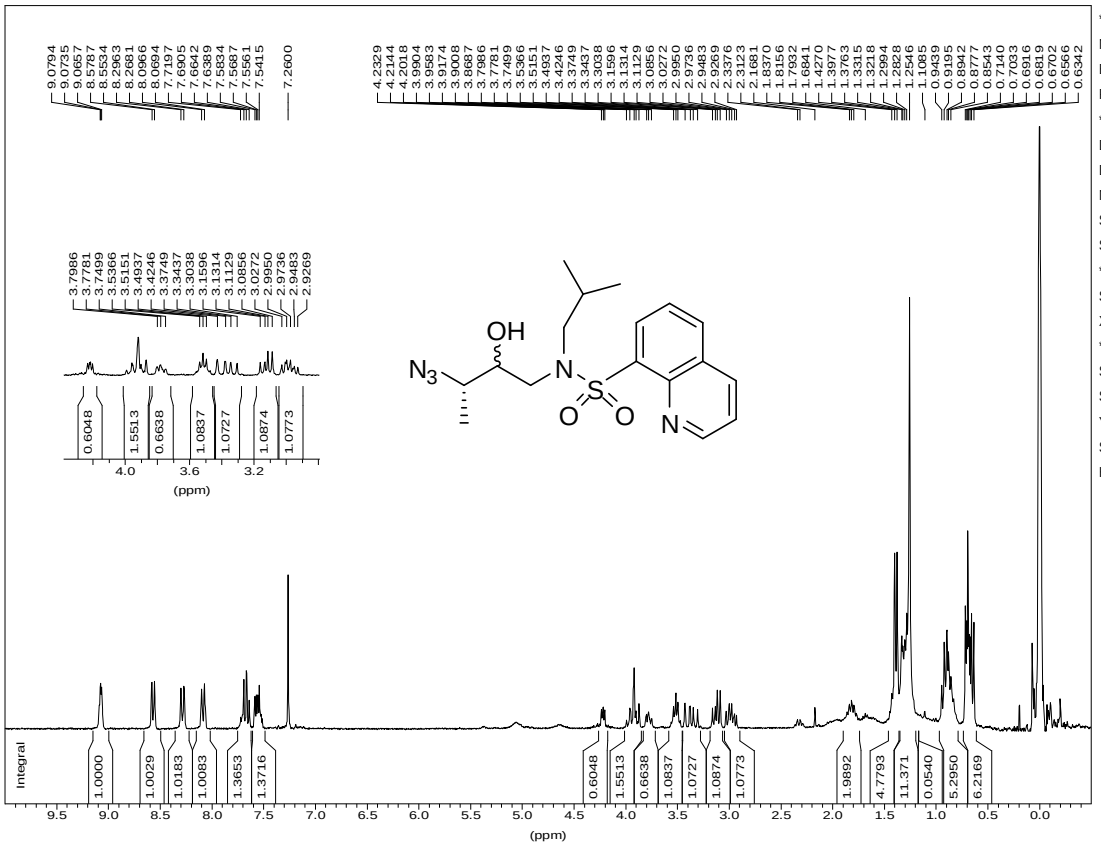

OH
N3
Br
(ppm)

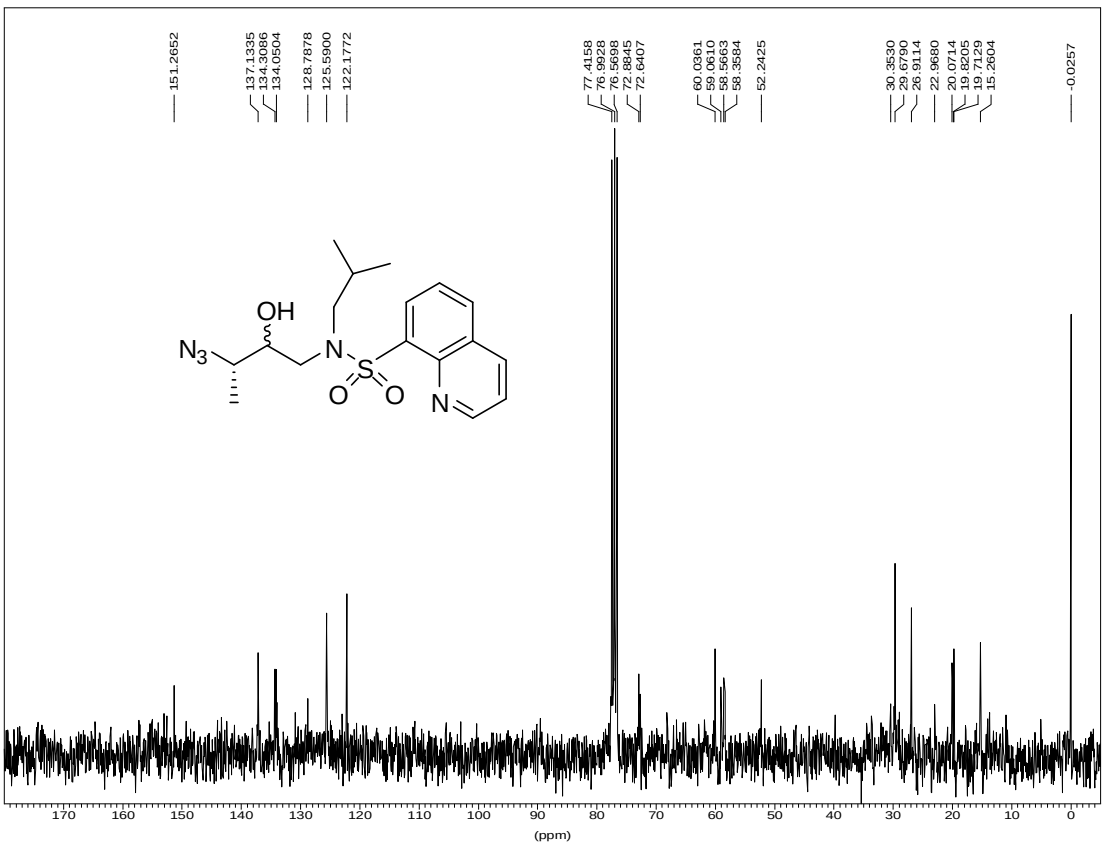

7.2297
6.9006
4.9503
4.0198
4.0072
3.8496
3.8358
3.7173
3.6895
3.6757
3.6668
3.6542
3.6404
3.6290
3.6189
3.6114
3.6063
3.5950
3.5824
3.5698
3.5559
3.5433
3.5193
3.5130
3.4992
3.4916
3.4727
3.4588
3.4513
3.4475
3.4361
3.4260
3.4134
3.4084
3.3945
3.3819
3.3743
3.3617
3.3454
3.3365
3.3138
3.2647
3.2621
2.7932
2.1968
1.3584
0.9121
0.8983
0.8895
0.8768
0.8630
OH
N3
Br
1.8369
2.1331
(ppm)
Integral
1.8369
2.1331
0.9009
3.0000
0.4986
(ppm)

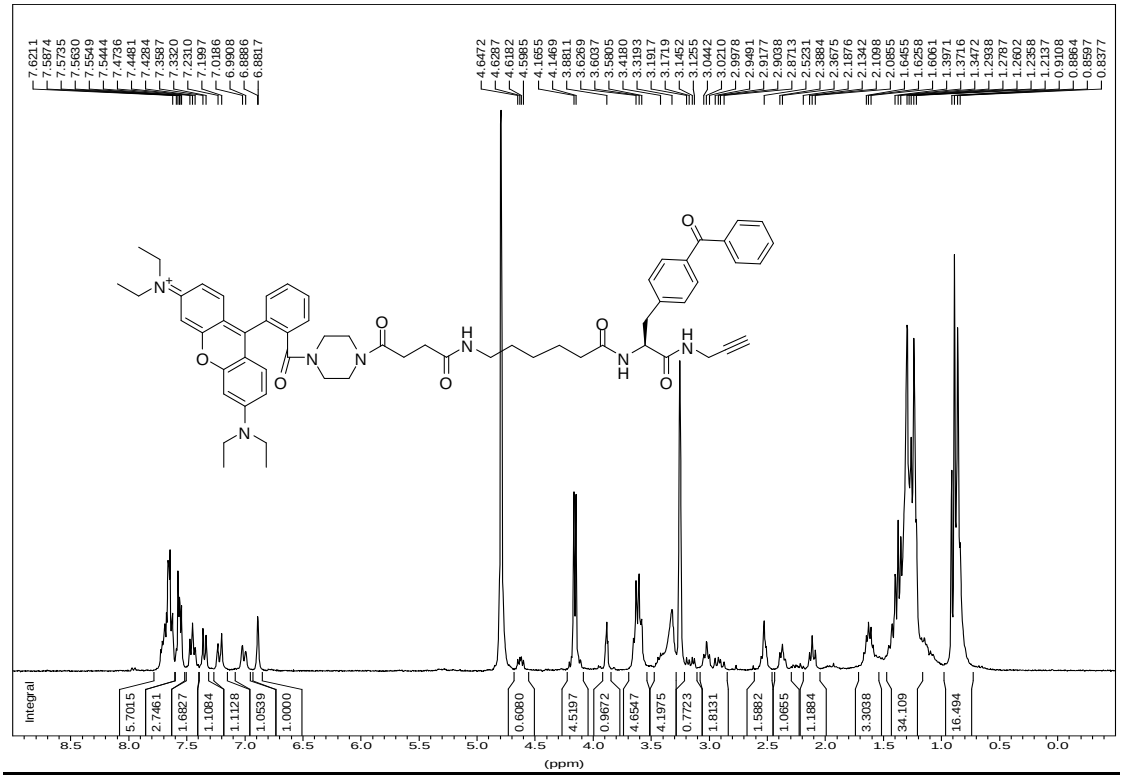

151.3060
151.2622
135.6527
125.2973
77.2590
77.0039
76.7488
73.8995
73.4331
59.2738
59.1717
36.1509
35.3566
34.8465
34.0157
30.1170
27.5983
27.4425
18.5009
15.4621
14.5876
0.7999
OH
N3
Br
(ppm)

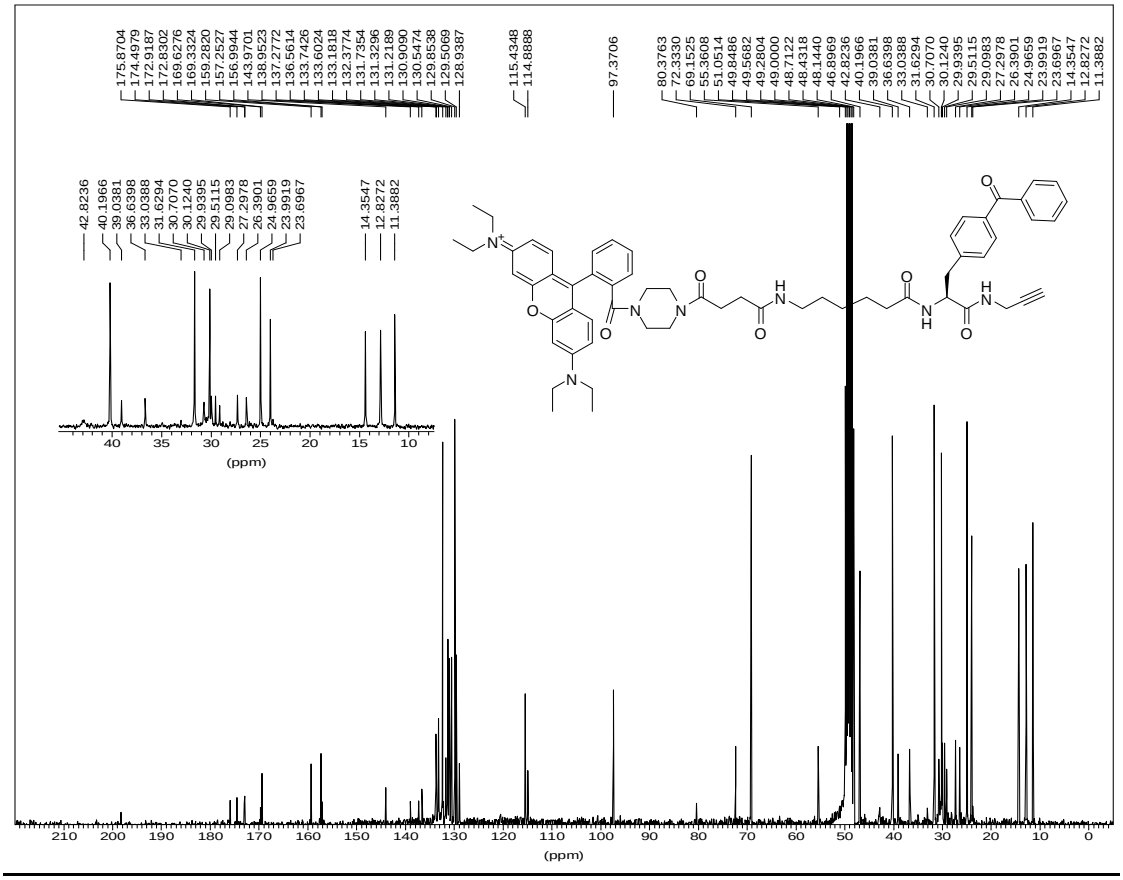

OH
H
N
N3
O
(ppm)
Integral

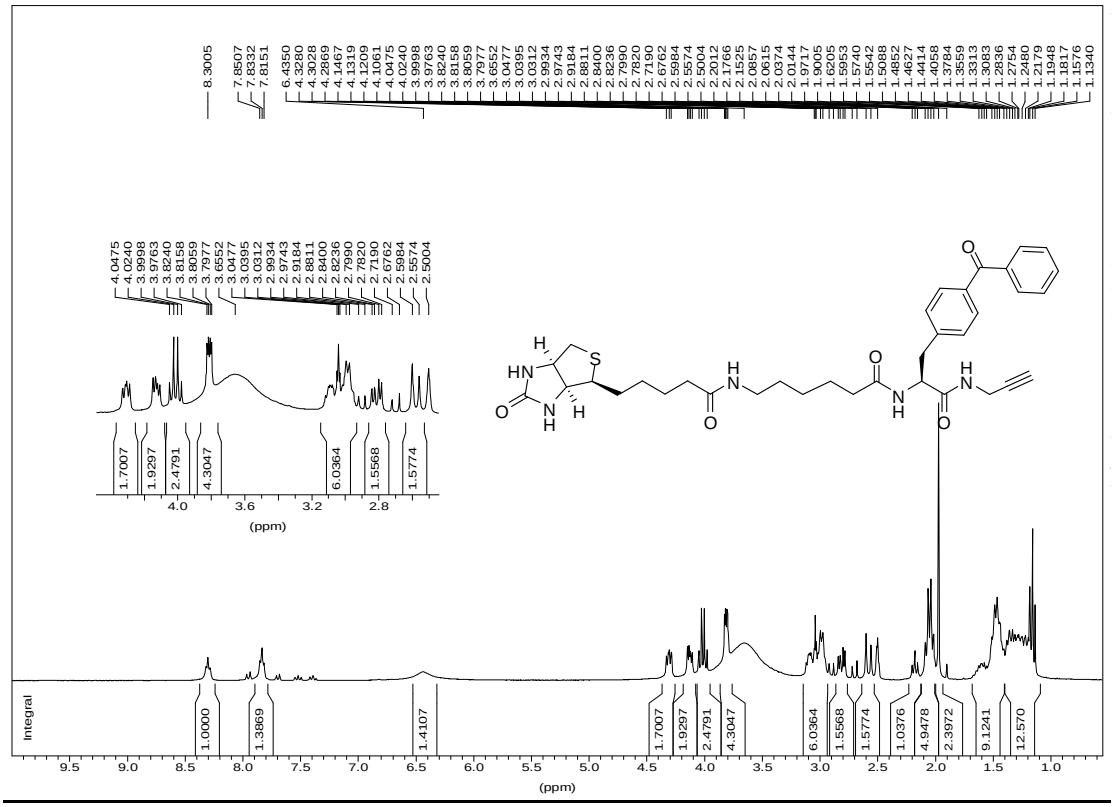

OH
H
N
N3
O
(ppm)

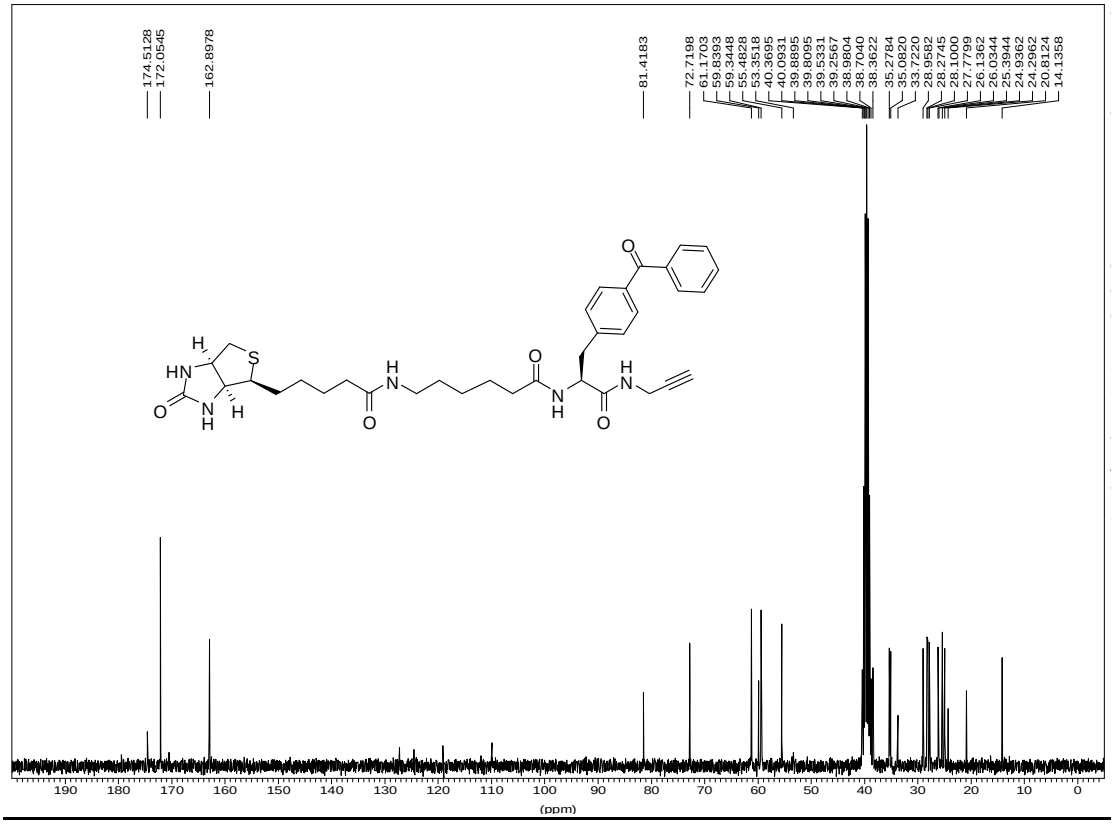

OH
N3
N
S
O
O
N
Integral
(ppm)

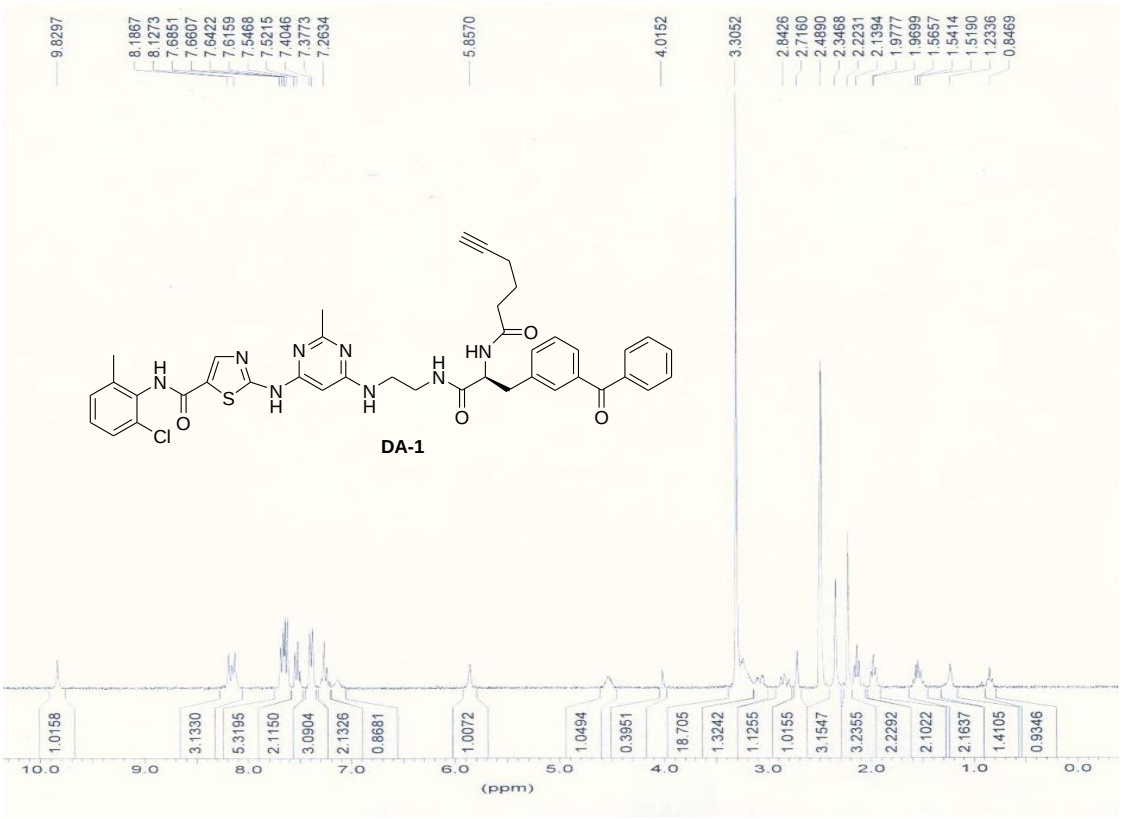

OH
N3
N
S
O
O
N
(ppm)



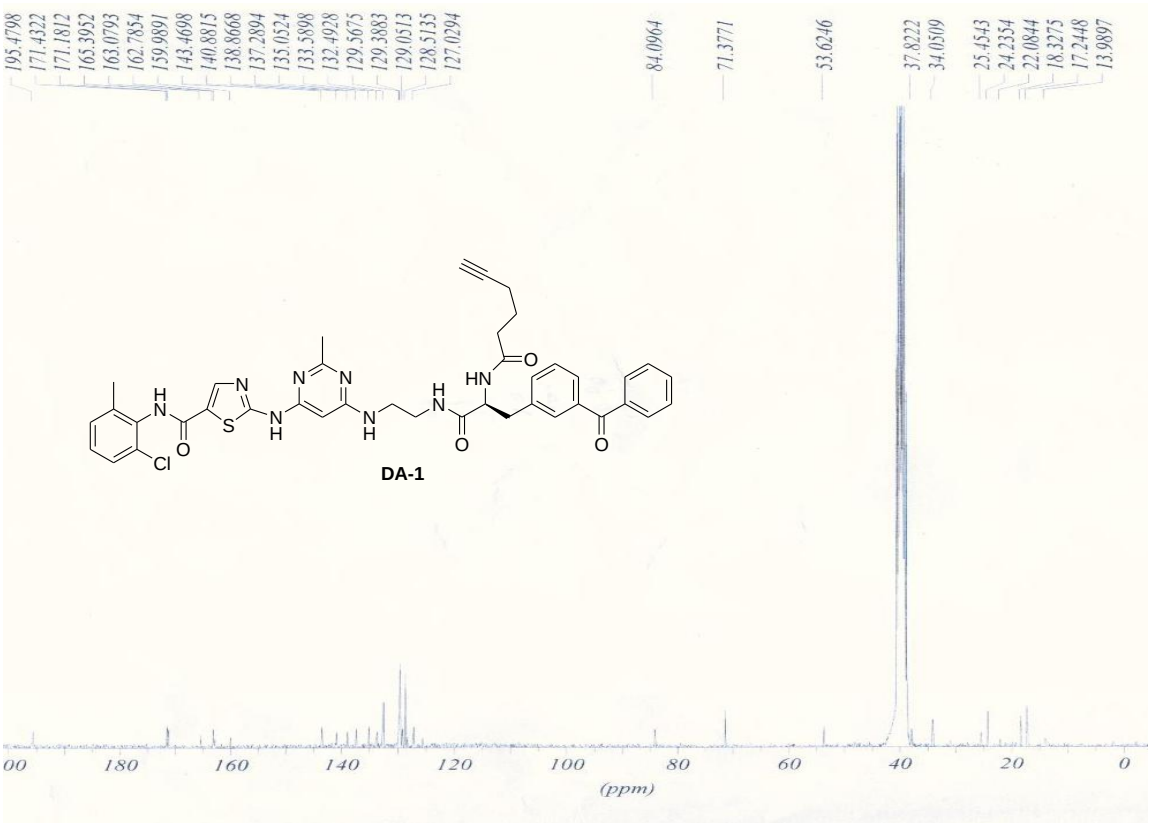




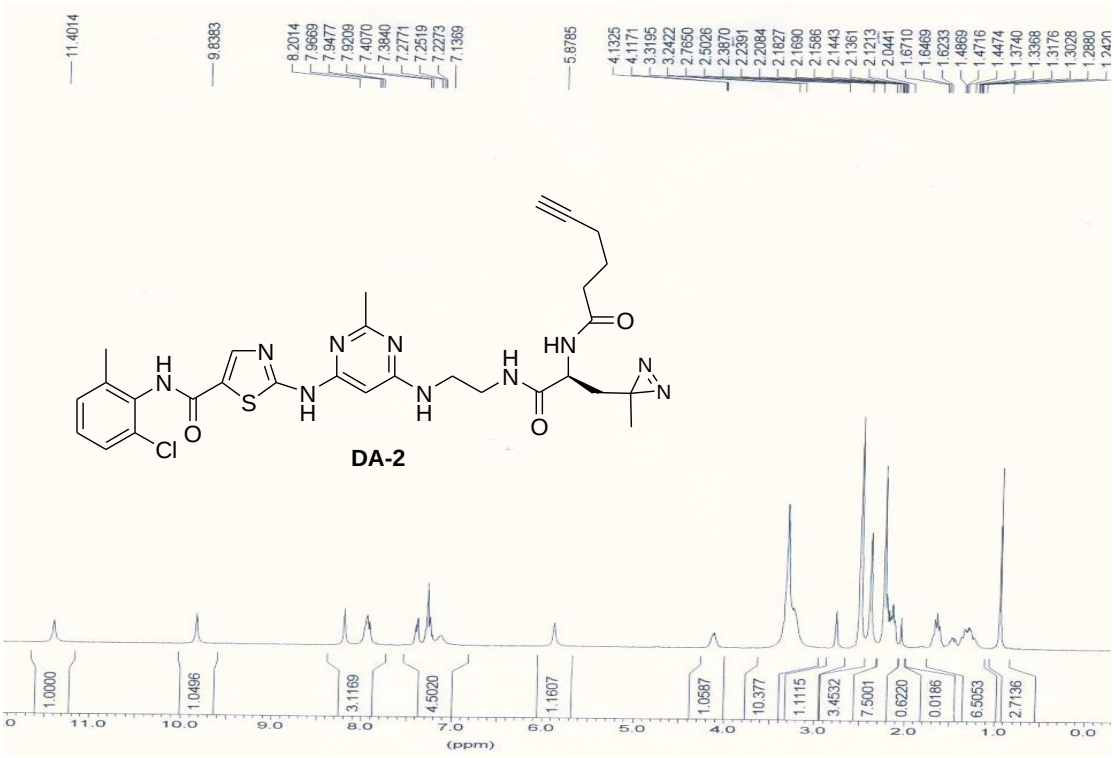



9.8297
8.1867
8.1273
7.6851
7.6607
7.6422
7.6159
7.5468
7.5215
7.4046
7.3773
7.2634
5.8570
4.0152
3.3052
2.8426
2.7160
2.4890
2.3468
2.2231
2.1394
1.9777
1.9699
1.5657
1.5414
1.5190
1.2336
0.8469
1.0158
3.1330
5.3195
2.1150
3.0904
2.1326
0.8681
1.0072
1.0494
0.3951
18.705
1.3242
1.1255
1.0155
3.1547
3.2355
2.2292
2.1022
2.1637
1.4105
0.9346
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
(ppm)

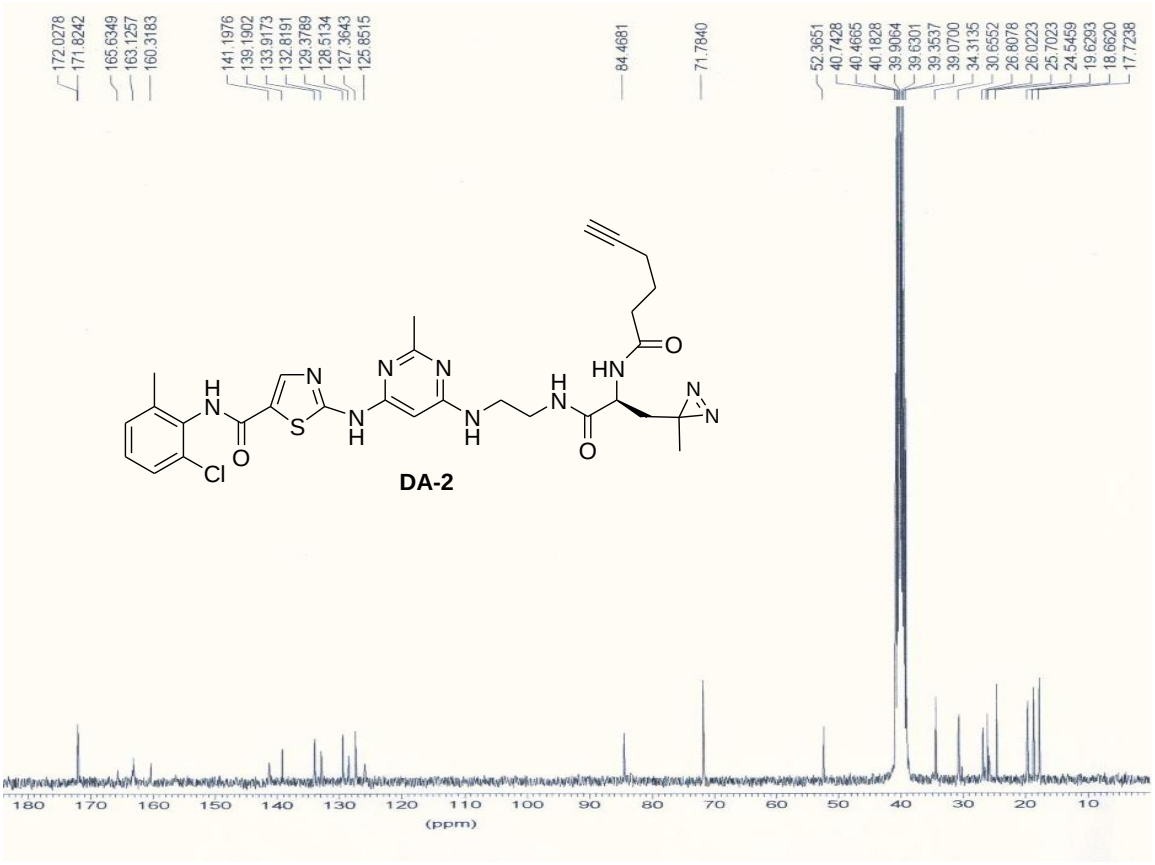
DA-1


193.4798
171.4322
171.1812
165.3952
163.0793
162.7854
159.9891
143.4698
140.8815
138.8668
137.2894
135.0524
133.5898
132.4928
129.5675
129.3883
129.0513
128.5135
127.0294
84.0964
71.3771
53.6246
37.8222
34.0509
25.4543
24.2354
22.0844
18.3275
17.2448
13.9897
180
160
140
120
100
80
60
40
20
0
(ppm)

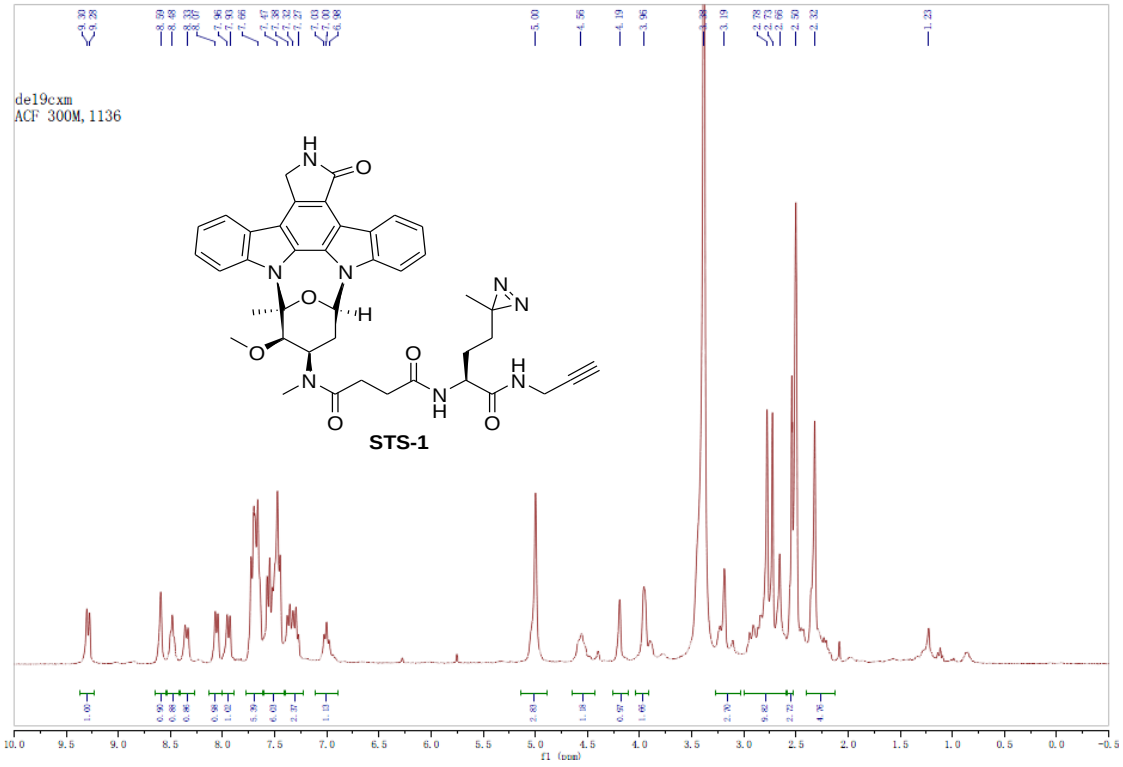
DA-1

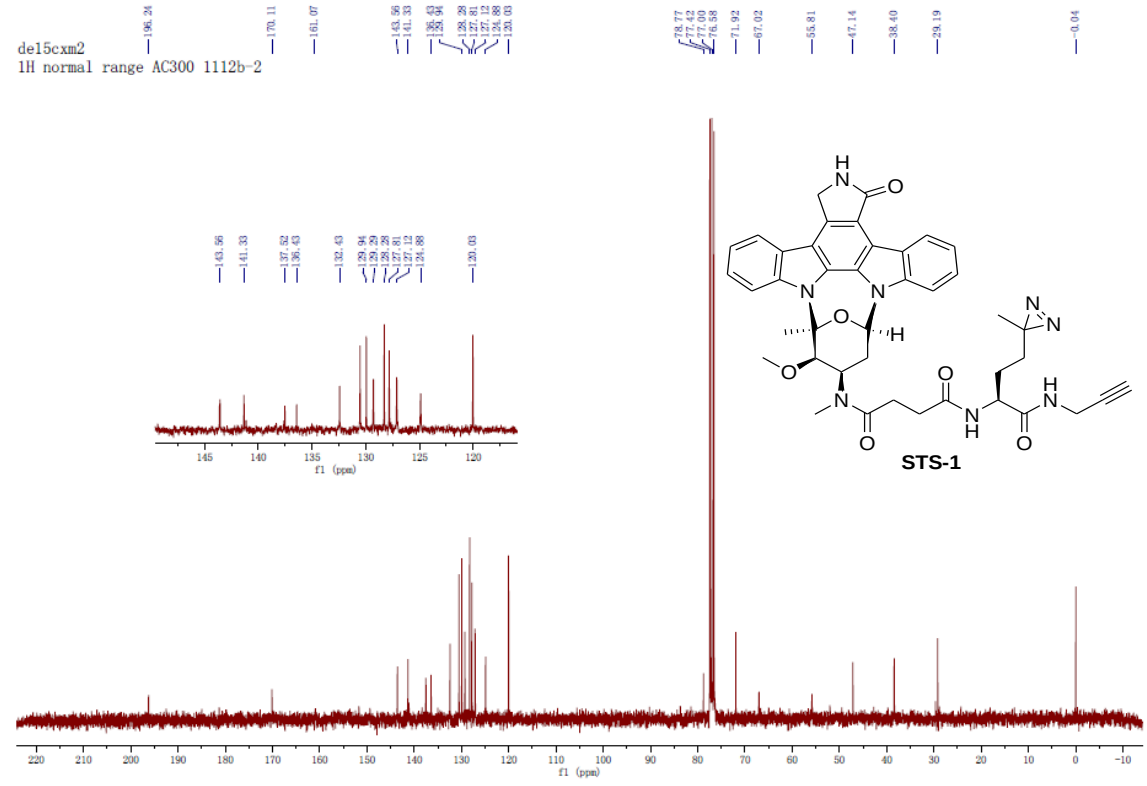

11.4014
9.8383
8.2014
7.9669
7.9477
7.9209
7.4070
7.3840
7.2771
7.2519
7.2273
7.1369
5.8785
4.1325
4.1171
3.3195
3.2422
2.7650
2.5026
2.3870
2.2391
2.2084
2.1827
2.1690
2.1586
2.1443
2.1361
2.1213
2.0441
1.6710
1.6469
1.6233
1.4869
1.4716
1.4474
1.3740
1.3368
1.3176
1.3028
1.2880
1.0000
1.0496
3.1169
4.5020
1.1607
1.0587
10.377
1.1115
3.4532
7.5001
0.6220
0.0186
6.5053
2.7136
(ppm)

DA-2


172.0278
171.8242
165.6349
163.1257
160.3183
141.1976
139.1902
133.9173
132.8191
129.3789
128.5134
127.3643
125.8515
84.4681
71.7840
52.3651
40.7428
40.4665
40.1828
39.9064
39.6301
39.3537
39.0700
34.3135
30.6552
26.8078
26.0223
25.7023
24.5459
19.6293
18.6620
17.7238
(ppm)

DA-2


de19cxm
ACF 300M, 1136
STS-1
f1 (ppm)


de15cxm2
1H normal range AC300 1112b-2
STS-1
f1 (ppm)