

DEVELOPING CHEMICAL BIOLOGY APPROACHES FOR
THE ACTIVITY-BASED INVESTIGATIONS OF
REVERSIBLE PROTEIN PHOSPHORYLATION-
MEDIATING ENZYMES

KARUNAKARAN NAIR A. KALESH

(M.Sc, Indian Institute of Technology, Madras, India)

A THESIS SUBMITTED FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF CHEMISTRY
NATIONAL UNIVERSITY OF SINGAPORE

2010

Acknowledgement

First and foremost, I express my deepest gratitude to my supervisor A/P Yao Shao Qin for being nothing less than a wonderful research advisor. Words are too few to express how much he has influenced me and how much he has inspired me in this journey. He has given me an incredibly encouraging and motivating environment to do science, he taught me how to find answers to my questions, and every scientific discussion with him has fueled my passion for science. His unparalleled commitment and dedication to science, professionalism and quick and intelligent approaches to problem solving have deeply influenced me and I hope they will guide me in my scientific journey in the years ahead.

Thanks are due to my colleagues in the Yao lab (Chemistry and Biology) for all your help, collaborations, discussions and most importantly your friendship which turned all the inevitable difficulties in research into wonderful learning experience, which I will cherish for ever. Souvik, Mingyu, Raja, Junqi, Liu Kai, Haibin, Lay Pheng, Candy, Jingyan, Hongyan, Bahulayan, Kitty, Liquian, Pengyu, Jigang, Wu Hao, Mahesh, Joo Leng, Li Bing, Derek, Wee Liang, Grace, Farhana, Wang Jun, Li Lin, Chongjing, Xiamin, Zhenkun, Su Ying, Cathy, Catherine, Ching Tian, Su Ling, Shen Yuan – working with all of them have been great experience. From day one, I thank Raja for introducing me to the Chemistry lab, for showing me for the first time how to run a column, for all your support at every stage of my life in the lab. I thank Souvik for all his help, both professional and personal, throughout my life at Yao lab. He was there always to discuss science, to help me troubleshoot bio-experiments, with utmost sincerity, nothing less than what one could expect from the closest friend or relative.

Special thanks are due to my collaborators- Liu Kai for providing me all the kinases, for the PTP-pull-down experiments and for all the biological experiments with the NDA-AD cross-linker. Lay Pheng for providing me all the PTPs and for helping me in the PTP-labeling experiments, Joo Leng for helping me in the synthesis of the caged PTP-probes, Li Bing and Wee Liang for helping me in the synthesis of the NDA-AD cross-linker, Derek for helping me in the synthesis of the dialdehyde 7 and Liquian and Hongyan for their help in the peptide synthesis - I have been extremely fortunate to have worked all of you.

There are a number of people outside Yao lab who made this journey more enjoyable- Santhosh, Rajesh, Abhilash....I thank you all for your true friendship.

I thank all staff from the Chemistry office, in particular Suria. I appreciate the support of the laboratory staff from the NMR and MS labs for the training and technical assistance.

I would never have accomplished this without the support, prayers and sacrifices of my parents. Kala-my sister, I would never have overcome difficulties in life without her support. Words are too few to express how much she has helped me to stabilize, emotionally, at all difficult times - both in personal life and in professional life. I dedicate this thesis to you Kala- my dear sister.

Last but not the least I thank the NUS for financial support in the form of the research scholarship.

Table of Contents

Contents	Page
Chapter 1: Introduction	1
1. 1 Proteomic approaches for the global analysis of protein expression and functions.	3
1. 1. 1 Methods based on liquid chromatography-tandem mass spectrometry (LC- MS/MS)	3
1. 1. 2 Methods based on isotope coded affinity tagging-tandem mass- spectrometry (ICAT- MS/MS)	5
1. 1. 3 Yeast-two-hybrid assays	7
1. 1. 4 Activity-based protein profiling (ABPP)	8
1. 1. 4. 1 General design considerations of ABPs	9
1. 1. 4. 2 “Label-free” versions of ABPs	13
1. 1. 4. 3 “Label-free” clickable versions of A/BPs	15
1. 1. 4. 4 “Non-directed approaches” in ABP designs	16
1. 2 Protein phosphorylation - An important post-translational modification (PTM)	16
1. 2. 1 Protein kinases	18
1. 2. 2 Protein phosphatases	20

1. 2. 3 Catalytic mechanism of Protein tyrosine phosphatases (PTPs)	21
1. 3 Enzyme inhibitor developments - Fragment-based approaches and high-throughput chemistry	22
1. 3. 1 “Click chemistry” in enzyme inhibitor developments	23
1. 3. 2 “ <i>In-situ</i> click chemistry” facilitated enzyme inhibitor developments	24
Chapter 2: Development of Peptide-Based Activity-Based Probes (ABPs) for Protein Tyrosine Phosphatases (PTPs)	27
Summary	27
2. 1 Introduction	27
2. 2 Synthesis of the unnatural amino acid, 2-FMPT	30
2. 3 Solid-phase synthesis of substrate peptides and peptide-based activity-based probes	31
2. 4 Expression and purification of PTPs	34
2. 5 Results and discussions	36
2. 5. 1 Labeling experiments with purified proteins	36
2. 5. 2 Detection limits of the probes	39
2. 5. 3 Labeling experiments with mutant PTPs	39
2. 5. 4 Effect of H ₂ O ₂ on PTP activity assessed with the probes	40
2. 5. 5 Kinetic characterizations and substrate specificities of the probes	

and the corresponding phosphopeptides	42
2. 5. 6 Labeling experiments in the presence of complex proteomes	47
2. 5. 6. 1 Labeling in the presence of bacterial cell lysates	47
2. 5. 6. 2 Labeling in the presence of mammalian proteome	48
2. 6 Conclusions	50
2. 7 General procedures for sample preparations and labeling experiments	
using proteomes	51
2. 7. 1 Preparation of bacterial cell lysates and labeling experiments	
using the lysates	51
2. 7. 2 Procedure for ‘Western Blot’ analysis	51
2. 7. 3 Procedure for ‘Pull-down’ of biotinylated probe labeled PTP	52
2. 8 Synthetic details and characterizations of compounds	53
Chapter 3: Caged Activity-Based Probes for Protein Tyrosine Phosphatases	
Summary	59
3. 1 Introduction	59
3. 2 Synthesis of caged 2-FMPT	65
3. 3 Synthesis of caged peptide-based activity-based probes	67
3. 4 Results and discussion	69
3. 5 Conclusions	71

3. 6 Synthetic details and chemical characterizations	72
Chapter 4: High-Throughput Synthesis of Abelson Tyrosine Kinase (Abl) Inhibitors using Click Chemistry	77
Summary	77
4. 1 Introduction	77
4. 2 Results and discussions	79
4. 2. 1 The first-generation kinase click inhibitors	79
4. 2. 2 The second-generation kinase click inhibitors	82
4. 2. 3 Kinase inhibition assays	84
4. 2. 3. 1 Screening of the inhibitor library and generation of heat-map	84
4. 2. 3. 2 IC ₅₀ evaluation of the click-inhibitors against Abl and Src kinases	87
4. 2. 4 Cell culturing and anti-proliferative assay	89
4. 3 Conclusions	99
4. 4 General experimental procedures	100
4. 4. 1 The click-assembly of inhibitors	100
4. 4. 1. 1 General procedures for the click-assembly of 344-member library formed from ADP-alkyne and azides	100

4. 4. 1. 2 General procedures for the click-assembly of 90-member Imatinib analogue library formed from the two warheads (W1 & W2) and azides	100
4. 4. 2 General procedures for Kinase inhibition assays	101
4. 4. 3 General procedures for cell-culturing and anti-proliferation assays	102
4. 5 Synthetic details and characterizations of compounds	103
Chapter 5: A Mechanism-Based Cross-Linker for Protein Kinase-Substrate Complexes	114
Summary	114
5. 1 Introduction	114
5. 2 Synthesis of the cross-linkers	117
5. 2. 1 Synthesis of OPA-AD	117
5. 2. 2 Synthesis of NDA-AD	118
5. 3 Synthesis of peptide pseudosubstrates	119
5. 4 Results and discussions	120
5. 5 Conclusions	123
5. 6 Synthetic details and characterizations of compounds	124
Chapter 6 Small-Molecule Probes that Target Abl Kinase	130
Summary	130

6. 1 Introduction	130
6. 2 Synthesis of the probes	134
6. 3 Results and discussions	136
6. 3. 1 Labeling experiments with the dialdehyde-7	136
6. 3. 1. 1 Labeling experiments with pure kinases and kinase spiked in cellular lysates	136
6. 3. 1. 2 pH-dependence of labeling reaction	138
6. 3. 1. 3 Effect of exogenous thiols on the efficiency of labeling	139
6. 3. 1. 4 Effect of exogenous amines on the efficiency of labeling	140
6. 3. 1. 5 IC ₅₀ evaluation of the probe	140
6. 3. 2 Labeling experiments with the photo cross-linkers	142
6. 3. 2. 1 Comparative labeling experiments	142
6. 3. 2. 2 Detection limit of pure Abl with the photo-cross-linker 6-13	146
6. 3. 2. 3 Labeling experiments with the clickable probe (6-13) in the presence of K562 cell lysate	146
6. 4 Conclusions	148
6. 5 Synthetic details and characterizations of compounds	149

Chapter 7: Future directions	153
Summary	153
7. 1 Protein-based PTP probes to identify/validate the PTPs responsible for dephosphorylating a given substrate protein	153
7. 2 Synthesis of a scaffold for the development of affinity-based probes (AfBPs) and bidentate inhibitors of protein kinases with a compact gatekeeper residue	161
Chapter 8: Concluding remarks	169
Chapter 9: References	172
Appendix	190

Summary

The reversible phosphorylation of proteins catalyzed by the opposing actions of protein kinases (PKs) and protein phosphatases (PPs) has been identified as one of the major post-translational modes (PTMs) of cellular signal transduction. These two classes of enzymes and their extremely intricate protein interaction networks and associated signal cascades play the most crucial roles in maintaining the normal cellular physiology. Being the key mediators of several cellular communications, the activities of members in these two classes of enzymes are tightly controlled by a variety of mechanisms and in many cases imbalances in such a control and the resultant aberrant activities of some of these proteins have been identified as the root causes of several pathological conditions in humans. Hence detailed investigations of individual members in these two classes of enzymes, both in their purified and isolated form (i e. *in vitro*) and in their native cellular environment (i e. *in vivo*), is of paramount importance both in terms of our better understanding of their roles in the cellular functioning and in developing more selective and effective therapeutic agents. Although conventional proteomic methods provide valuable information regarding the expression levels of the proteins, relatively newer approaches such as Activity-Based Protein Profiling (ABPP) provide more insights into the functional states of these proteins, which are of more relevance in the cellular physiology and pathology. This dissertation reports certain chemical approaches developed towards better understanding and manipulations of some important members in these two classes of proteins.

In Chapter 2, the design and development of a panel of peptide-based Activity-Based Probes (ABPs) for protein tyrosine phosphatases (PTPs) with a key PTP-

reactive unnatural amino acid has been described. Labeling reactions with the panel of probes using purified and isolated proteins showed activity-based labeling specificity consistent with the known substrate preferences of different PTPs. The strategy has also been found to be useful for efficient labeling reactions of PTPs from highly complex biological samples. A caged version of the unnatural amino acid with a photolabile *o*-nitrobenzyl group on the phosphate moiety (caged-2-FMPT) was subsequently synthesized and incorporated into peptides to generate peptide-based, caged, ABPs (Chapter 3). Using these probes, with PTP1B as a model system, the concept of photo-uncaging followed by activity-based labeling was validated. Chapter 4 describe the development of a synthetic strategy using the modular and efficient nature of the Cu (I) catalyzed click-reaction to rapidly assemble inhibitor libraries of Abelson (Abl) tyrosine kinase. Biochemical assays using the click-inhibitor library revealed a set of moderately potent and selective inhibitors of the Abl kinase. In Chapter 5, the synthesis and biochemical evaluation of an improved mechanism-based cross-linker, naphthalene 2,3-dicarboxaldehyde-adenosine (NDA-AD), for the identification of kinase-substrate interactions from crude proteomes is described. The cross-linker NDA-AD, in addition to its improved labeling performances from crude proteomes was found to be suitable for the detection of kinase-pseudosubstrate interactions of both tyrosine-specific and serine/threonine-specific protein kinases. In Chapter 6, the development of selective small molecule-based ABPs for the Abl kinase using two different strategies namely a dialdehyde-based cross-linking and photo-affinity labeling is described. Chapter 7 provides a brief outlook to some of the future developments possible in line with the ABPP- and inhibitor-developments of PKs and PTPs discussed in the previous chapters.

It is hoped that the kinase- and phosphatase-directed ABPP and inhibitor-development approaches presented as part of this thesis, would provide a guideline for the future developments of more powerful tools for the investigation of these extremely important signalling enzymes.

List of Publications

1. Kalesh, K. A.; Sim, S. B. D.; Wang, J.; Liu, K.; Lin, Q.; Yao, S. Q.; “Small Molecule Probes that Target Abl Kinase”, *Chem. Commun.*, **46**, 1118-1120 (2010).
2. Kalesh, K. A.; Tan, L. P.; Liu, K.; Gao, L.; Wang, J.; Yao, S. Q.; “Peptide-Based Activity-Based Probes (ABPs) for Target-Specific Profiling of Protein Tyrosine Phosphatases (PTPs)”, *Chem. Commun.*, **46**, 589-591 (2010).
3. Kalesh, K. A.; Liu, K.; Yao, S. Q.; “Rapid Synthesis of Abelson Tyrosine Kinase Inhibitors Using Click Chemistry”, *Org. Biomol. Chem.*, **7**, 5129-5136 (2009).
4. Liu, K.; Kalesh, K. A.; Ong, L. B.; Yao, S. Q.; “An Improved Mechanism-Based Cross-Linker for Multiplexed Kinase Detection and Inhibition in a Complex Proteome”, *ChemBioChem*, **9**, 1883-1888 (2008).
5. Tan, L. P.; Wu, H.; Yang, P. -Y.; Kalesh, K. A.; Zhang, X.; Hu, M.; Srinivasan, R.; Yao, S.Q.; “High-Throughput Discovery of Mycobacterium Tuberculosis Protein Tyrosine Phosphatase (MptpB) Inhibitors Using Click Chemistry”, *Org. Lett.*, **11**, 5102-5105 (2009).
6. Srinivasan, R.; Tan, L. P.; Wu, H.; Yang, P. -Y.; Kalesh, K. A.; Yao, S. Q.; “High-Throughput Synthesis of Azide Libraries Suitable for Direct Click Chemistry and in situ Screening”, *Org. Biomol. Chem.*, **7**, 1821-1828 (2009).
7. Srinivasan, R.; Li, J.; Ng, S.L.; Kalesh, K.A.; Yao, S.Q. “Methods of Using Click Chemistry in the Discovery of Enzyme Inhibitors – Potential Application in Drug Discovery and Catalomics”, *Nat. Protoc.*, **2**, 2655-2664 (2007).

8. Kalesh, K. A.; Yang, P. -Y.; Srinivasan, R.; Yao, S. Q.; “Click Chemistry as a High-Throughput Amenable Platform in Catalomics”, *QSAR Comb. Sci.*, **26**, 1135-1144 (2007).
9. Kalesh, K. A.; Shi, H.; Ge, J.; Yao, S. Q.; “The Use of Click Chemistry in the Emerging Field of Catalomics”, *Org. Biomol. Chem.*, **8**, 1749-1762 (2010).

List of Abbreviations

AcOH	Acetic acid
AA	Amino acid
ABP	Activity-based probe
ABPP	Activity-based protein profiling
Boc	<i>tert</i> -Butoxycarbonyl
br	Broad
BSA	Bovine serum albumin
tBu	<i>tert</i> -Butyl
CA	Chloroacetamide
CBD	Chitin binding domain
Cbz	Benzyloxycarbonyl
Cy3	Cyanine dye ³
C-terminal	Carboxy terminal
Da	Dalton
DAST	Diethylamino sulfurtrifluoride
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
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EA	Ethyl acetate
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetracetic acid
EPL	Expressed protein ligation
ESI	Electrospray ionization
Fmoc	9-Fluorenylmethoxycarbonyl
FMP	4-Formyl-3-methoxyphenoxy resin
HATU	O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	O-benzotriazole-N,N,N',N'-tetramethyluronium hexafluoro phosphate

HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOBT	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
Hz	Hertz
ICAT	Isotope coded affinity tagging
IC ₅₀	Half maximal inhibitory concentration
J	NMR coupling constant
K _D	Dissociation constant
K _M	Michaelis-Menten constant
LC-MS	Liquid chromatography-Mass spectrometry
m	Multiplet
m-CPBA	m-Chloroperbenzoic acid
min	Minute
mmol	Millimole
MMP	Matrix metalloprotease
MP	Metalloprotease
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry

MSNT	1-(Mesitylene-2-sulfonyl)-3-nitro-1,2,4-triazole
MudPIT	Multidimensional protein identification technology
MW	Molecular weight
NaCl	Sodium chloride
NaHCO ₃	Sodium bicarbonate
Na ₂ SO ₄	Sodium sulphate
NCL	Native chemical ligation
NHS	N-Hydroxysuccinimide
nM	Nanomolar
NMP	N-methylpyrrolidone
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
pI	Isoelectric point
PKA	Protein Kinase A
PTP	Protein tyrosine phosphatases
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

q	Quartet
RBF	Round bottom flask
RF	Relative fluorescence
RP	Reverse phase
RT	Room temperature
s	Singlet
SDS-PAGE electrophoresis	Sodium dodecyl sulfate-polyacrylamide gel
SE	Sulfonate ester
SrtA	Sortase A
t	Triplet
TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetraborofluorate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMSI	Trimethylsilyliodide
Tof	Time of flight
Tris	Trishydroxymethylamino methane

UV	Ultraviolet
VS	Vinyl sulfone
Y2H	Yeast two hybrid

List of 20 Natural Amino Acids

Single Letter Code	Three Letter Code	Full Name
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Try	Tryptophan
Y	Tyr	Tyrosine

List of Schemes

Scheme	Page
2.1 Synthesis of the unnatural amino acid, 2-FMPT	31
2.2 Solid-phase synthesis of 10 phosphopeptides and 11 peptide-based ABPs	33
2.3 Schematic representation of different PTP constructs used	36
3.1 Proposed mechanism for light-mediated uncaging of <i>o</i> -nitrobenzyl caged molecules	60
3.2 Synthesis of caged 2-FMPT	65
3.3 Synthesis of caged peptide-based ABPs	67
4.1 Synthesis of ADP-alkyne warhead	81
4.2 Synthesis of the two warheads (W1 & W2) for Imatinib-based click library	84
5.1 Scheme showing the three-component cross-linking reaction of kinase with its pseudosubstrate and NDA-AD	116
5.2 Synthesis of the cross-linker, OPA-AD	118
5.3 Synthesis of the cross-linker, NDA-AD	118
6.1 Synthesis of the Abl-directed probes	135
7.1 Scheme for constructing a protein-based PTP-probe using Expressed Protein Ligation (EPL)	157

7.2	Solid-phase synthesis of the peptide ligation-partners	159
7.3	Synthesis of a clickable inhibitor scaffold (compound 7-9) for the potential development of Af/BPs and bidentate inhibitors of protein kinases with a compact gatekeeper residue	163

List of Figures

Figure		Page
1.1	Overview of Catalomics	2
1.2	Schematic representation of 2D-LC coupled to MS/MS	4
1.3	Chemical structures of ICAT reagents	5
1.4	Schematic representation of ICAT-MS-based protein quantification and identification strategy	6
1.5	Overview of Yeast two-hybrid assay	8
1.6	Schematic of ABPP showing two different approaches using either (a) ABPs (activity-based probes) or (b) Af/BPs (affinity-based probes)	10
1.7	Schematic of ABPP using clickable activity-based probes	14
1.8	Schematic of Af/BPP using clickable photo-reactive probes	15
1.9	Schematic representation of reversible protein phosphorylation mediated by protein kinases and protein phosphatases	18
1.10	Catalytic mechanism of PTPs with PTP1B as a representative	22
1.11	Click assembly of enzyme inhibitors	24

1.12	Schematic of <i>in situ</i> click chemistry	26
2.1	(a) Structures of known ABPs (top) of PTPs and 2-FMPT , and its corresponding peptide-based ABPs (boxed)	30
	(b) Proposed mechanism of activity-based labeling using 2-FMPT incorporated peptide-based probes	30
2.2	Activity-based labeling of different proteins using a representative probe	37
2.3	(a) Fluorescent labeling profiles of five different PTPs (top to bottom) with the panel of probes (left to right)	38
	(b) Quantified relative fluorescence intensity of labeling of the 5 different PTPs against the panel of probes	38
2.4	Detection limit of PTP1B with probe P3	39
2.5	(a) Comparative labeling profiles of mutant, denatured and active PTP1B versions with probe P3	40
	(b) Microplate-based enzymatic assay of PTP1B and mutants using DiFMUP as the fluorogenic enzyme substrate	40
2.6	Effect of H ₂ O ₂ on PTP1B activity assessed with the probe	41
2.7	Determination of the kinetics of inactivation of PTP1B using the probes	44
2.8	(a) Standard curve of phosphate detection using Malachite green assay	45
	(b) Determination of k_{obs} of the phosphopeptides for reaction with PTP1B	45
2.9	Comparison of relative activity of the 10 probes against PTP1B as Determined from quantitative analysis of the fluorescent gels,	

1/K _i values from inactivation kinetic experiments and data from the dephosphorylation of ten phosphopeptides	46
2.10 Labeling fingerprint of PTP1B spiked in the bacterial proteome using the panel of 10 probes	47
2.11 (a) Labeling of spiked PTP1B in the presence of bacterial proteome	49
(b) Labeling of PTPs in mammalian cell lysates. Left panel, in-gel Fluorescence analysis of global PTP activity profiles obtained from total cell lysates of HEK293T cells and NIH3T3 cells with probe P3 . Right panel, Anti PTP1B blots of the two labeled lysates	49
(c) Pull-down results using the biotinylated probe P11 from HEK293T cell lysate showing detection of endogenous PTP1B	49
3.1 Chemical structures of reagents used in the synthesis of caged phosphates	63
3.2 Proposed mechanism of photo-uncaging of the peptide-based probe followed by labeling of PTP	64
3.3 UV-irradiation-dependant labeling of PTP1B with the probes c-P1 and c-P2	69
3.4 (a) Scheme representation photo-uncaging of the caged-peptide-based probe	70
(b) HPLC profiles of the caged probe, c-P1 with increasing time of UV-irradiation	70
4.1 Click-assembly of the 344-member inhibitor library generated using the ADP-alkyne and azides	81
4.2 The heat-map of the inhibition assay obtained with the 344-member	

inhibitor library against Src and Abl kinases	82
4.3 General scheme for luminescence-based kinase assay	85
4.4 Heat-map showing the relative inhibition of the 90-member Imatinib-based bisubstrate inhibitor library against Src and Abl kinases	86
4.5 IC ₅₀ evaluation of selected click-inhibitors and the warheads (W1 & W2) against Abl/Src kinases	88
4.6 Antiproliferation assay of K-562 cells in the presence of click-inhibitors	90
5.1 Chemical structure of ATP, OPA-AD and NDA-AD	116
5.2 Comparative labeling profiles of OPA-AD and NDA-AD for the kinase PKA (with PKA-pseudosubstrate) in the presence of bacterial cell lysate	121
5.3 Fluorescence-scanned gels showing cross-linking profiles of NDA-AD against Tyr-specific and Ser/Thr-specific protein kinases	122
6.1 Two different strategies to develop Abl-selective probes.	
(a) A three-component (kinase, pseudosubstrate and dialdehyde) reaction- mediated labeling of Abl by the dialdehyde 7 (compound 6-7)	132
(b) Clickable photo-affinity probe (6-13) mediated labeling of Abl	132
6.2 Proposed mechanism for the labeling of Abl by (a) dialdehyde 7 and (b) clickable photo-affinity probe (6-13)	136
6.3 Fluorescence scanned gels showing the labeling of Abl and Csk kinases with the dialdehyde 7	138
6.4 pH-dependence of the labeling of Abl by the dialdehyde 7	138

6.5	Effect of β -mercaptoethanol (BME) on the three-component reaction of Abl kinase with the dialdehyde 7 and Pseudo-Abltide	139
6.6	Effect of exogenous lysine on the labeling reaction	140
6.7	IC ₅₀ evaluation of the cross-linker 7 for Abl kinase inhibition	142
6.8	(a) Labeling of Abl with the photo-cross-linkers 6-12 and 6-13	145
	(b) Abl-labeling specificity of 6-13 evaluated with purified proteins	145
	(c) Labeling of different amounts of spiked Abl in the presence of CHO-K1 mammalian proteome	145
	(d) Dose-dependant reduction of labeling of Abl with 6-13 in the presence of the generic kinase inhibitor Staurosporine	145
6.9	Evaluation of detection limit of pure Abl with compound 6-13	146
6.10	Labeling of different amounts of spiked Abl in the presence of K-562 mammalian proteome	147
7.1	Domain structure of c-Src kinase	156
7.2	Chemical structures of the peptide ligation partners for the construction of the protein-based PTP-probes. (A) the peptide for pTRAP probe and (B) the peptide for the ppCAP probe	158
7.3	LC-MS profiles of the purified pTRAP and pCAP peptides	160

List of Tables

Table		Page
2.1	The 11 probes and their AA sequences and substrate preferences	34
2.2	Comparison of the substrate specificity of PTP1B obtained from the probe-mediated inactivation of the enzyme (K_i values) and that from the Malachite green phosphatase assay of the corresponding phosphopeptides	45
3.1	Amino acid sequences in the two peptide-based caged PTP probes	69
4.1	Inhibition data of selected click-inhibitors against Abl and Src kinases	91
4.2	Azide library used for the synthesis of the Imatinib analogue click-library	92

Chapter 1: Introduction

With the advancements in the genome sequencing projects, functional annotations of an innumerable number of proteins, already known and newly predicted, gained accelerated research interests. In particular, understanding the catalytic roles and interaction networks of enzymes, the most important bio-catalysts, has become a very active area of research. Being the key regulators of virtually every aspect of cellular physiology, even minor imbalances in certain enzyme activities are known to have profound implications in many pathological conditions.¹ Most enzyme families consist of, several sub-families and classes and with the structure and functions of majority of members remaining largely unexplored, in-depth understanding of molecular configurations that are recognized and accepted by individual enzymes or enzyme classes, which has direct relevance to drug developments, functional annotations and even identification of their signalling cascades, remains a daunting task. To meet these goals one of the major requirements is amenable synthetic chemistry in combination with high-throughput enzyme screening and characterization technologies. These combinations of different powerful chemical and technological advances directed towards elucidating and modulating the catalytic functions of enzymes are described in a unified platform termed Catalomics (Figure 1.1).²

Context-dependant post-translational protein modifications (PTMs) have long been known to play key roles in regulating the activities and functions of many proteins.³ Of the different PTMs, the reversible protein phosphorylation catalyzed by the opposing actions of protein kinases (PKs) and protein phosphatases (PPs) has been identified as one of the most important modes of cellular signal transduction.⁴ Hence,

detailed investigations on various aspects; such as the expression levels, sub-cellular localizations, catalytic activities and functional roles of these reversible phosphorylation-mediating enzymes are of paramount importance in our understanding of their roles in maintaining the normal cellular physiology. Together with several conventional proteomic techniques, relatively newer approaches such as activity-based protein profiling (ABPP) are providing unprecedented advancements in this field in recent years. Furthermore, as the aberrant activities and expression levels of many of these enzymes are characterized, and in many more cases implicated, as root causes for several pathological conditions, selective and powerful ways of modulating their activities, e.g. with inhibitors or activators, remains another important area in protein kinase/phosphatase research. This chapter gives a brief introduction into the fundamental conceptual and technical framework, upon which much of my research reported in this dissertation is based.

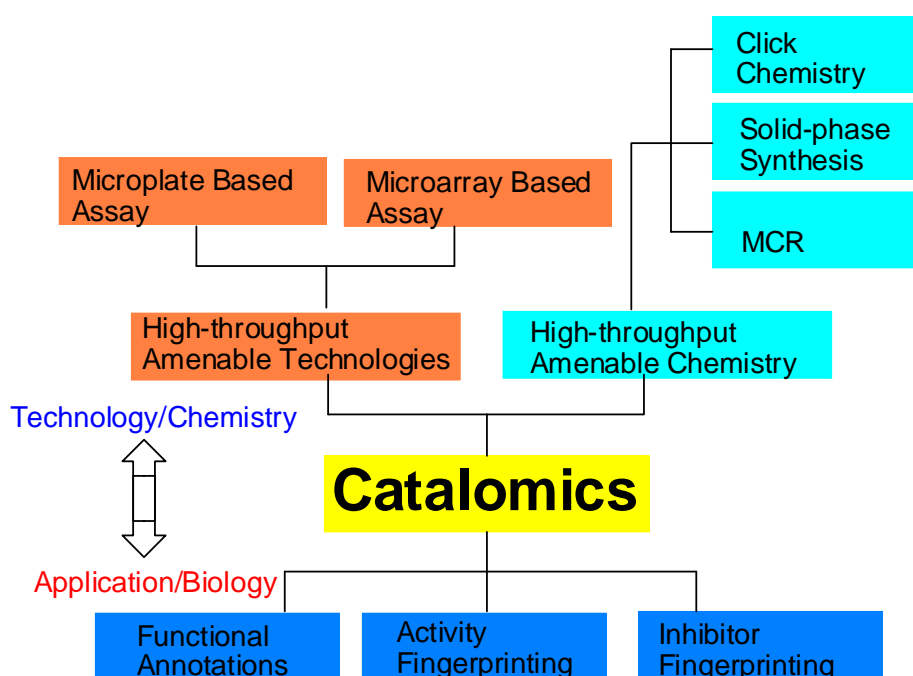


Figure 1. 1 Overview of Catalomics

1. 1 Proteomic approaches for the global analysis of protein expression and functions.

1. 1. 1 Methods based on liquid chromatography-tandem mass spectrometry (LC-MS/MS)

The field of proteomics is primarily involved in the large scale analysis of proteins in complex biological samples such as normal and cancer cells, tissues and fluids. Such a global analysis of proteins is of fundamental importance in identifying the functional roles of proteins and protein complexes, their interaction networks, expression levels and co-operative effects in the cellular physiology. The simplest method to resolving proteins from mixtures is the use of one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) where denatured and sodium dodecyl sulfate (SDS) capped proteins are separated according to their sizes in a cross-linked and porous polymeric gel matrix under an electrical potential difference across the gel bed. However given the high heterogeneity and finite quantity of biological samples, the resolution of 1D-PAGE is typically not sufficient for global protein analysis. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled to mass spectrometry (MS) is perhaps the most widely used method for the large scale protein separation and identification. In 2D-PAGE, proteins are first separated in one dimension based on their differences in the isoelectric points (pI) and then in the second dimension based on their differences in the molecular weight, thus giving rise to a higher resolution. Although widely used, the 2D-PAGE is low-throughput and in many cases the sample preparations, extraction of spots, digestion and analysis of each spot are all tedious and time consuming. Moreover proteins of extreme pI, very

high or very low molecular weight proteins, low abundance proteins and membrane-bound proteins are typically not tractable with this technique.

The shortcomings of the gel-based 2D separation are eliminated in an alternative 2D separation method developed by Link et al (Figure 1.2).⁵ The researchers developed a two-dimensional liquid chromatography (2D-LC) system equipped with a microcapillary column packed with two independent chromatographic stationary phases (a strong cation exchange resin phase and a reverse phase) for the 2D separation of peptides generated from the tryptic digestion of denatured protein complexes and an online coupling of the LC unit to tandem mass spectrometry (MS/MS) and comparison of the obtained MS data with existing protein MS database provides a quick identification of the corresponding proteins from which the peptide fragments are generated. The strategy has been successfully employed for the large scale analysis of proteome samples from various biological sources.⁶ Such 2D-LC coupled to MS/MS is also known as multidimensional protein identification technology (MudPIT) and the technique, by virtue of its enhanced resolution and sensitivity, has been identified as a powerful tool for detailed proteomic analysis of complex biological samples.

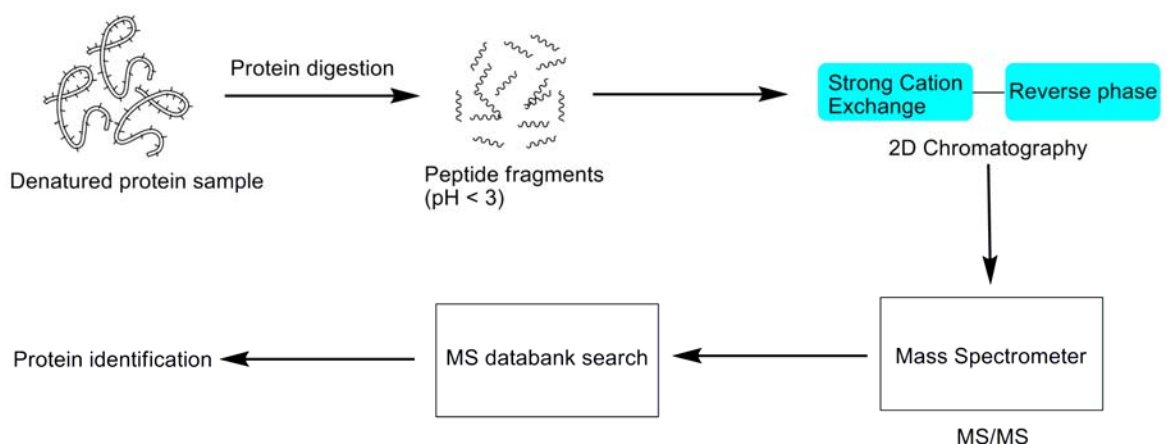


Figure 1. 2 Schematic representation of 2D-LC coupled to tandem MS for protein analysis of complex samples

1. 1. 2 Methods based on isotope coded affinity tagging – tandem mass spectrometry (ICAT- MS/MS)

Isotope coded affinity tagging (ICAT) is a mass spectrometry-based technique developed towards quantitative proteomics.⁷ The technique employs reagents equipped with a thiol reactive group, a biotin tag and an isotopically coded linker. The reagent exists in light and heavy isotopic forms (Figure 1.3). Protein samples in two different cell states are separately treated with the two isotopic forms of the reagent, upon which the thiol reactive group covalently labels the sulfhydryl residues in every protein. The protein mixtures are then combined, digested with trypsin and the tryptic peptides are isolated using affinity purification with avidin beads. The labeled peptides are separated by microcapillary liquid chromatography (μ -LC) followed by quantification and sequence identification using tandem mass spectrometry (MS/MS).

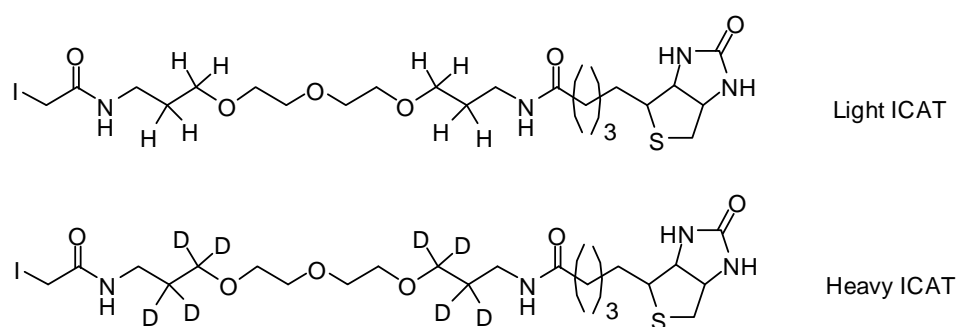


Figure 1. 3 Chemical structures of the light and heavy isotopic forms of the ICAT reagents

A pair of ICAT labeled identical peptides from the two different cell state samples remain chemically identical and coelute but are distinguished in the mass spectrometry due to the mass difference between the isotopically labeled probes they carry. Thus a simple measurement of peak relative abundance ratio of each peptide in

the mass spectrum provides a quantitative measure of the corresponding relative protein levels in the two samples. Several improvements in the basic ICAT-MS/MS technique such as solid phase labeling methods with isotopically labeled amino acids and photocleavable linkers have been reported.⁸ Similar mass spectrometry based techniques such as Stable isotope labeling by amino acids in cell culture (SILAC)⁹ which employs isotopically labeled amino acids in cell growth medium and Isobaric tagging for relative and absolute quantification (iTRAQ)¹⁰ which employs amine reactive N-hydroxysuccinimide esters are also increasingly used in the large scale proteomic research.

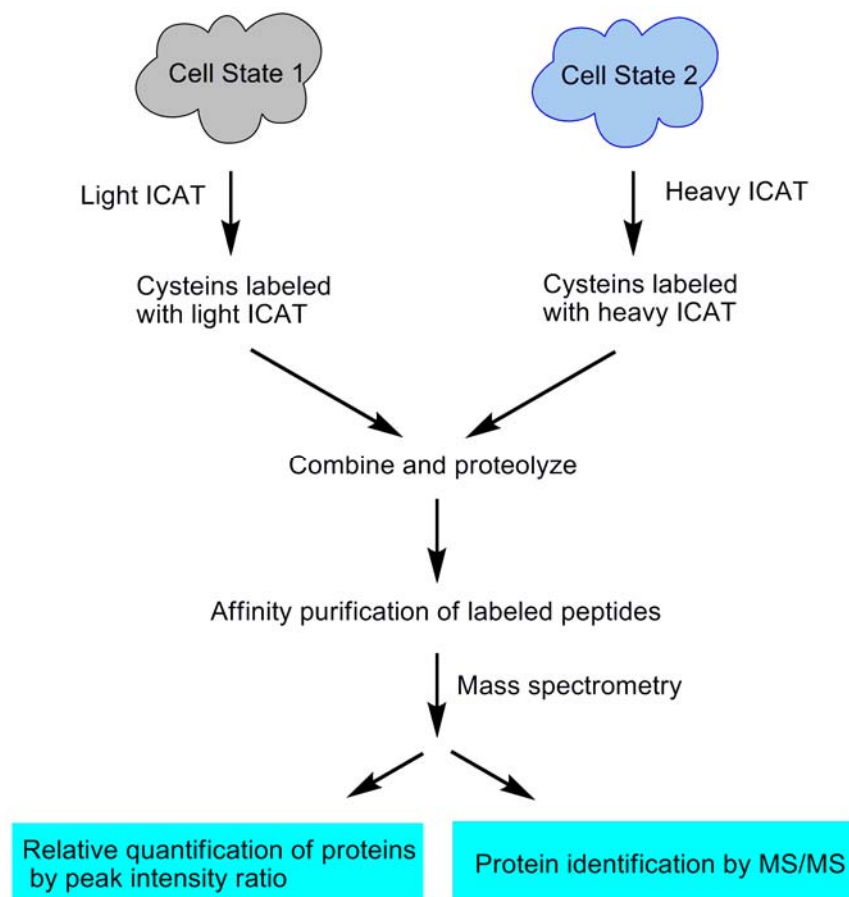


Figure 1. 4 Schematic representation of ICAT-MS-based protein quantification and identification strategy

1. 1. 3 Yeast two-hybrid assays

The metabolism and homeostasis of cells depends on protein-protein interactions. Thus a molecular level understanding of the life process requires detailed knowledge about such intricate protein interaction networks. Several standard techniques such as the use of glutathione-s-transferase fusion proteins, coimmunoprecipitation, use of chemical cross-linkers, phage-display methods and yeast two-hybrid assays have been developed to characterize protein-protein interactions. The yeast two-hybrid (Y2H) system (Figure 1.5) is one of the most widely used techniques for mapping out protein interaction networks.¹¹ This technique is based on the modular domain structure of the transcription factor, GAL4, of yeast which is made up of a DNA binding domain and a transcription activation domain and the two domains become functional in close proximity to each other even in the absence of a direct binding between the two. In the Y2H assay, a protein whose interaction partner has to be identified (termed bait) is expressed as a hybrid with the DNA binding domain of the yeast GAL4 while the potential interaction partner (termed prey) is expressed with the activation domain. If the two proteins (the bait and the prey) interact, it functionally reconstitutes the GAL4 which in turn induces the expression of reporter genes. On the other hand if the proteins do not interact, the transcription of the reporter gene does not take place. The Y2H system also allows the screening of the protein of interest expressed with the DNA binding domain with a library of proteins fused to the activation domains from which the potential interaction partners can be verified.

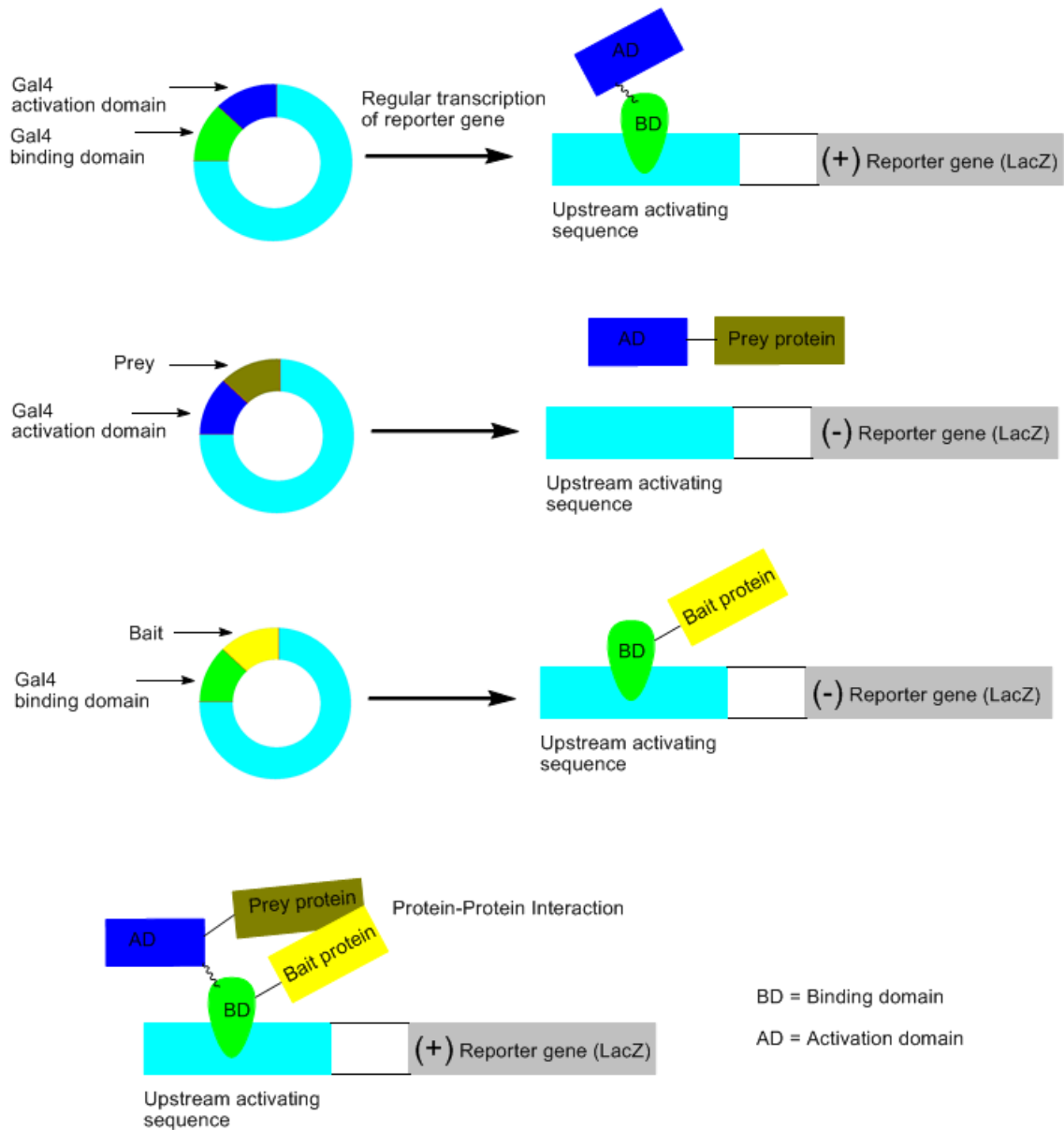


Figure 1. 5 Overview of Yeast two-hybrid assay

1. 1. 4 Activity-based protein profiling (ABPP)

The comparative proteomic methods discussed above primarily depend on the expression levels of proteins; therefore the study of low abundant proteins remains extremely cumbersome with these techniques. Moreover these methods, although provide valuable information about the global protein expression, protein interaction networks and protein abundance, are inherently limited in their ability to directly

report the functional states and activities of proteins in their native cellular environment. Since most of the proteins/enzymes activities are regulated by autoinhibitory domains or endogenous inhibitors/activators and a myriad of post-translational events with a direct correlation to the normal cellular physiology and pathology, complementary protein-profiling strategies capable of reporting the functional states of the proteins rather than their mere abundance are of extreme importance. Towards this, a novel chemical proteomic strategy called Activity-Based Protein Profiling (ABPP) has been emerged.¹² This technique employs active site directed chemical probes which report the functional integrity and activity of the target protein. These Activity-Based Probes (ABPs), not only are able to identify their target proteins from a crude proteome but more importantly distinguishes the functionally active form of the target protein from its zymogen or inhibitor-bound forms. As such, competitive profiling of enzymes with ABPs in the presence of enzyme inhibitors allows a convenient means to identify potent and selective inhibitors which could provide impetus in drug discovery efforts toward several diseases. Furthermore, large-scale comparative proteomic profiling of cellular lysates with ABPs may helps in the functional annotations of uncharacterized proteins as well as in the identification of functional roles of known proteins in normal cellular physiology versus pathological cellular conditions.

1. 1. 4. 1 General design considerations of activity-based probes

The general designs of the ABPs contain three parts (Figure 1.6a). (1) active-site targeting reactive group, sometimes called warhead (W), that directs the probe to the active site and causes covalent reaction with the protein. (2) a reporter unit, which is typically a fluorophore or biotin for the direct read out of the protein's

activity and (3) a linker, which minimize the possible disruption from the reporter tag in the protein recognition of the reactive unit. The early stage developments of ABPs solely relied on known electrophilic irreversible inhibitor scaffolds capable of forming stable covalent bonds with nucleophilic residues near the active-site of the target proteins. This approach, although highly desirable, encounter significant challenge from proteins that lack such active-site targeting covalent binding scaffolds. Photo-affinity reagents¹³ offered a powerful alternative strategy to probe such proteins and several research groups have developed photo-affinity based probes (AfBP) for different proteins via appending a photoreactive molecule such as benzophenone, alkyl or aryl diazirine or aryl azides, to tight-binding reversible inhibitor scaffolds targeting those proteins.¹⁴ The AfBP first recognizes the active site of the target protein using the reversible inhibitor scaffold and subsequently, upon irradiation with UV light, a reactive intermediate is generated from the photo-reactive group which causes covalent cross-linking of the probe with a suitable proximal residue near the active-site of the protein (Figure 1.6b). These photo-affinity-based probes typically label the target enzyme in an activity-dependant manner like the traditional ABPs.

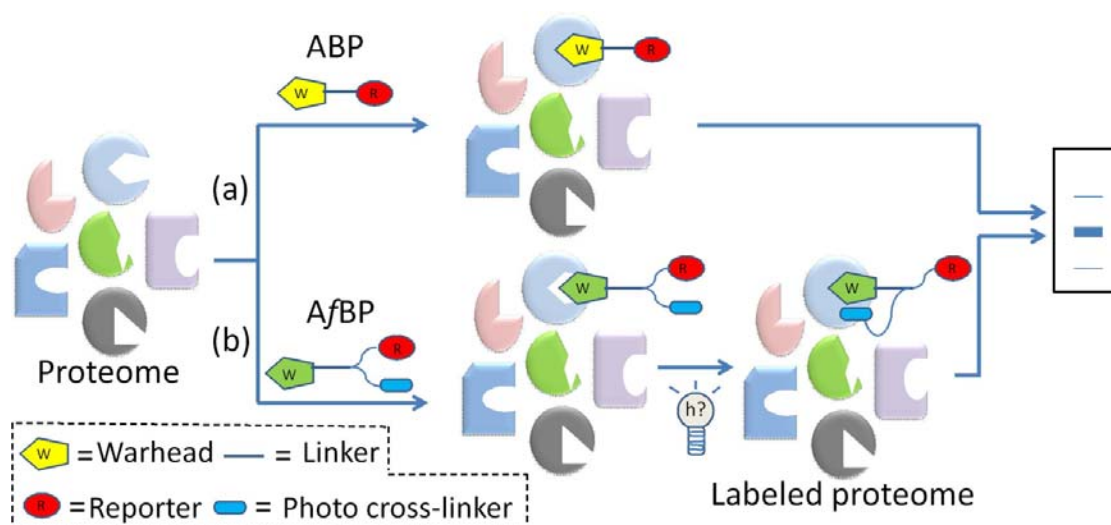


Figure 1. 6 Schematic of ABPP showing two different approaches using either ABPs (activity-based probes) or AfBPs (affinity-based probes). (a) In this approach, a crude proteome is treated with the activity-based probe which bears a reactive group or warhead (**W**, shown in yellow) and the reporter unit (**R**, shown in red). The probe covalently reacts with the active site of the target protein and the labeling is visualized by in-gel fluorescence scanning following separation of proteins on SDS-PAGE. (b) In this approach, the crude proteome is treated with an AfBP whose warhead (**W**, shown in green) first recognizes the target protein's active site via non-covalent interactions. Subsequently upon UV irradiation, a reactive intermediate generated from the photo-reactive unit (depicted as blue oval) covalently reacts with suitable residues near the active site. The labeled protein is separated by SDS-PAGE and visualized by in-gel fluorescent scanning.

Although the primary role of the linker portion is to minimize the possible steric hindrance from the reporter tag in the protein active-site recognition of the reactive group, variations in linker designs, in many cases, lead to altered labeling performances of the probes. For instance, long chain alkyl linkers causes hydrophobic interactions which are favourable in certain cases of protein recognition while it limits aqueous solubility of the probes. On the other hand, polyethylene glycol (PEG) linkers facilitates hydrophilic interactions and increases the aqueous solubility of the probes. Peptide-based linkers have been designed to perform isoform-selective labeling of certain proteins where the amino acid sequences in the linkers exploit the substrate specificity of different proteins within the same family.¹⁵ A recent improvement in the linker designs is due to the introduction of cleavable units which could be conveniently removed using a suitable trigger after the proteome labeling. The cleavage trigger could be an acid (e. g. TFA for acid-cleavable linkers),¹⁶ a protease (for protease-cleavable linkers),¹⁷ UV light (for photo-cleavable linkers)¹⁸ or a mild reducing agent (e. g. sodium hydrosulfite for diazobenzene-based reductively cleavable linkers).¹⁹ The use of cleavable linkers eliminate the need of harsh elution conditions required in

the case of biotinylated probe-labeled proteins as well as non-specific bindings from endogenously biotinylated proteins in the post-labeling protein enrichment using avidine beads.

Many ABPs use biotin as the reporter tag as it facilitates, in addition to visualization of the labeled protein by streptavidine blot, purification and enrichment of the labeled protein by means of proteomic pull-down with streptavidine/neutralavidine coated magnetic/agarose beads. But the incorporation of biotin into probes could cause adverse effects, in particular several non-specific bindings and membrane impermeability. Since the biotin-avidine interaction is one of the strongest interactions known in biology ($K_d \approx 10^{-15}$ M), harsh elution conditions are required to separate the labeled protein from the avidine beads, which sometimes lead to protein loss and labeling artifacts. Furthermore, the biotin probe labeled protein is visualized indirectly with anti-biotin antibodies or neutralavidine-HRP complexes and such assays are limited in sensitivity and have a very narrow dynamic range, making quantitative labeling assessments and comparisons of different protein samples extremely cumbersome. The use of fluorescent tags eliminates most of these difficulties. Fluorescently labeled proteins can be directly visualized by in-gel fluorescent scanning following protein separation by SDS-PAGE. Moreover, fluorescence is much more sensitive and the availability of fluorophores with a large range of spectral properties provides a large dynamic range of protein detection and facilitates proteomic labeling of complex biological samples using probes with spectrally distinct tags. The most commonly used fluorophores include rhodamine derivatives, fluorescein, BODIPY dyes and cyanine dyes such as Cy3 and Cy5.

1. 1. 4. 2 “Label-free” versions of ABPs

Many ABPs, due to the presence of their reporter tags, suffer obvious limitations both from synthetic point of view and in their labeling performance. Fluorophores are expensive, and typically the synthetic steps of the probes with reporter tags are difficult to purify due to the peculiar solubility and polarity properties of most of the fluorophores/biotin. Even more serious drawbacks appear in the biological labeling performance of these probes mainly due to their poor membrane permeability. Moreover, the bulky reporter units, in many cases, disrupt the proper binding interaction between the reactive unit and the protein’s active site. To override these problems, the labs of Cravatt and Overkleeft independently reported the first “label-free” versions of ABPs, where a two-step protein profiling strategy making use of a bioorthogonal labeling of the probe treated proteome was employed.²⁰ Overkleeft and co-workers used an azide functionalized vinyl sulfone probe to target proteasomes in cell lysates as well as in live cells. After the probe treatment, the lysates were subjected to Staudinger ligation with a biotinylated triarylphosphine reagent and the labeling was visualized by immunoblotting. The Cravatt’s lab, on the other hand, utilized click chemistry for the labeling. In their original work, Cravatt and co-workers treated cell lysates from COS7 cells overexpressing glutathione S-transferases (GSTO 1-1) with an azide derivatized phenyl sulfonate (PS) probe. The probe treated sample upon click-conjugation with a rhodamine alkyne followed by SDS-PAGE and in-gel fluorescence scanning identified intense labeling of the target GSTO 1-1.

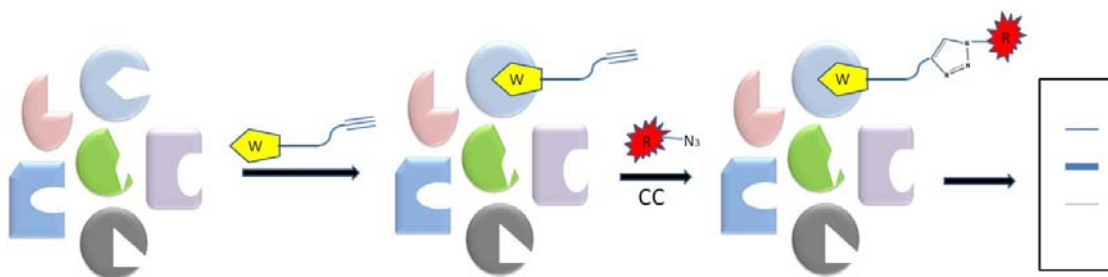


Figure 1. 7 Schematic of ABPP using clickable activity-based probes. The proteome is treated with the clickable probe whose warhead **W** (yellow) covalently reacts with the active site of the target protein. The reporter **R** (shown in red) is then incorporated into the probe-protein complex using CC, which enables subsequent visualization of the labeled protein with in-gel fluorescent scanning.

With the initial success, the label-free clickable ABPP strategy has been quickly adopted for other classes of proteins. Alexander and Cravatt utilized an N-alkynyl derivative of a known tight binding inhibitor against fatty acid amide hydrolases (FAAHs), named JP104, for the comprehensive evaluation of the *in vivo* proteome reactivity of these enzymes.²¹ Similarly in an effort to determine protein targets of bioactive natural products, Cravatt and co-workers performed cell-based screening with a small-molecule clickable probe library bearing different natural product scaffolds.²² The researchers identified one probe, MJE3, which possesses an 1-oxa-spiro[2,5]octane scaffold, showing significant inhibition of breast cancer cell proliferation. The probe was found to covalently label a glycolytic enzyme called phosphoglycerate mutase 1 (PGAM1) in cancer cells. Cravatt and co-workers recently reported a set of clickable probes for the enzymes Cytochrome P450s.²³ The researchers observed that their probes were able to detect both inhibition and activation of cytochrome P450s by different small molecules. Similarly, Taunton and co-workers developed a clickable probe, called fmk-pa, to study the autoactivation mechanisms of the p90 ribosomal protein S6

(RSK) kinase.²⁴

1. 1. 4. 3 “Label-free” clickable versions of AfBPs

Similar to ABPs, clickable versions of photo affinity-based probes (AfBPs) have also been developed. The first design of a clickable AfBPs was from the Pieters group. Based on the known binding interaction of β -galactoside containing motifs to galectins, the group designed a clickable photo-affinity probe which was shown to bind purified galectins.²⁵ The same group subsequently reported a divalent version of the original probe, which possesses improved detection limits and enabled specific labeling of endogenous galectin 3 in human colon carcinoma lysate.²⁶ Sieber *et al* adopted the same strategy for targeting metalloproteases using hydroxamate-containing peptide-based probes.²⁷ In another recent development, Cravatt’s laboratory reported a very sensitive clickable photoreactive probe, named, SAHA-BPyne, for detection of histone deacetylases (HDACs).^{14b} The SAHA-BPyne was found to detect endogenous class-1 and class-2 HDACs as well as some of the HDAC-interacting proteins in the close proximity of the probe in highly complex whole cell proteomes and live cells.

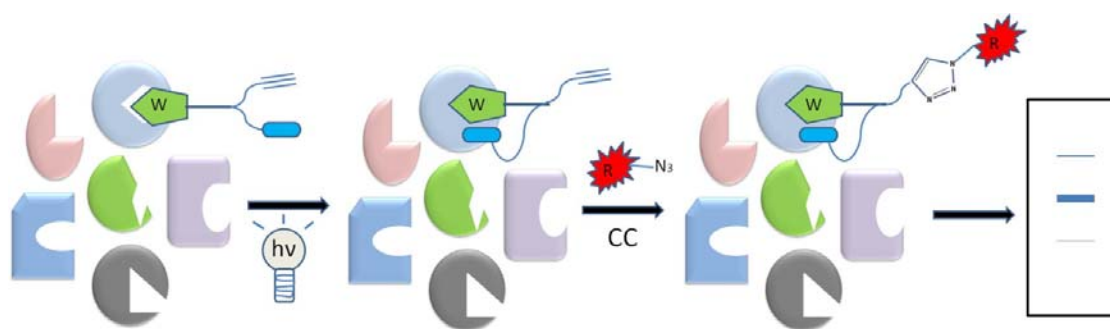


Figure 1. 8 Schematic of AfBPP using clickable photo-reactive probes. The warhead, **W** (shown in green), which first occupies the target protein’s active site. Subsequent UV irradiation leads to covalent cross-linking of the probe with the target protein. The reporter is then incorporated into the probe-protein complex using CC and the labeled proteins are visualized by in-gel fluorescent scanning.

1. 1. 4. 4 Non-directed approaches in ABP designs

Besides the rationally designed probes with known/predicted protein targets, the ABPP strategy has been applied in a conceptually different manner using the so called “non-directed probe libraries”. In this approach, a probe library equipped with several structurally different electrophilic warheads are applied to a complex proteome where each protein/protein class may exhibit its own characteristic labeling profiles with the different warheads and the comparative labeling profiles thus generated would be useful for the functional characterization of unidentified proteins and also in the identification of preferred reactive units of each protein. Cravatt and co-workers employed a non-directed library of sulfonate ester (SE) probes bearing different aryl/alkyl binding groups against various proteome samples.²⁸ The researchers identified six mechanistically different enzyme classes as the probes targets with each target exhibiting unique labeling profile with the probes. Recently dipeptidic α -chloroacetamide (α -CA)²⁹ and spiroepoxide²² reactive-groups based non-directed probe libraries have also been reported.

1. 2 Protein phosphorylation - An important post-translational modification (PTM)

The most fundamental knowledge-bearing molecular entities in the cells are the genes; however the genome carries nothing but the information to generate primary structures of proteins. It is the proteins and their highly intricate and tightly controlled interaction networks that regulate and mediate virtually every aspect of cellular functions. Post-translational modifications (PTMs) of proteins, which involve both reversible and irreversible incorporation of a wide range of chemical moieties such as

phosphates, lipids, sugars, alkyl, acyl and sulphates is one of the most important cellular mechanisms that control the structure, localization and functions of proteins. It is one of the three fundamental modes of cellular signal transduction; the other two being direct contact-based interactions such as the interaction between G-proteins and receptors and communications mediated through diffusible signalling molecules such as hormones, neurotransmitters, cytokines etc. Among the different PTMs, the reversible protein phosphorylation has been identified as the most dynamic and ubiquitous mechanism for intracellular communications. Also, several exogenous signals transmitted through receptor proteins trigger intracellular protein phosphorylation/dephosphorylation cascades by which the cell decipher and process external stimuli/signals. Thus not only the intracellular signal transductions but the cellular response to several exogenous signals and hence the associated intercellular signal transduction events also depend on the protein phosphorylation.

Phosphorylation of proteins is catalyzed by enzymes called protein kinases which facilitate the transfer of a single phosphate (γ -phosphate) from ATP molecules to the side chain hydroxyls of serine, threonine and tyrosine residues in proteins. The reverse reaction, i.e. removal of phosphate from the side chain hydroxyls of phosphorylated serine, threonine and tyrosine residues in proteins, is catalyzed by a group of enzymes called protein phosphatases. The co-operative actions of these two classes of enzymes mediate a large array of vital cellular processes but the specific outcome of a protein's phosphorylation/dephosphorylation event is typically context-dependant. Thus although the reversible phosphorylation acts as a binary code for the cell to decipher and process many signals, the exact meaning of the signal depends on the context, and in many cases the same phosphorylation/dephosphorylation event

conveys different meanings and triggers different cellular processes, making the study of phosphorylation highly challenging.

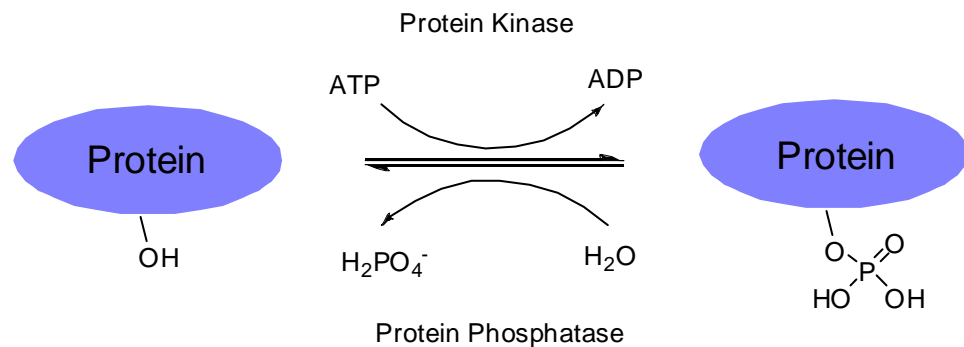


Figure 1. 9 Reversible protein phosphorylation mediated by protein kinases and protein phosphatases. Protein kinases catalyze the transfer of a phosphate from ATP molecules to hydroxyls of Serine, Threonine and Tyrosine residues in proteins while protein phosphatases catalyze the dephosphorylation of such phosphorylated residues.

1. 2. 1 Protein kinases

Protein kinases are classified into two main groups⁴ 1) the Ser/Thr – kinases, which constitute around 80% of the protein kinases and 2) the Tyr – kinases. The two classes of kinases, although phosphorylate different residues, have similar catalytic domain structure which consists of a bi-lobed structure with a small N-terminal lobe predominantly made up of β -sheets and a larger C-terminal lobe predominantly made up of α -helices with the ATP-binding site located in between the two lobes. The tyrosine-specific kinases have a deeper catalytic cleft to accommodate the tyrosine residue which is presumably the major reason for the Tyr vs. Ser/Thr substrate selectivity between the two classes of kinases. However this selectivity is not exclusive; several Ser/Thr kinases are known to phosphorylate tyrosine residues as well whereas Tyr kinases tend to show more rigorous preference for tyrosine residues.

Approximately 2% of the entire eukaryotic genome encodes protein kinases and it is estimated that at any given time approximately 30% of cellular proteins are phosphorylated on at least one residue. Around 17% of amino acid contribution for proteins being Ser, Thr and Tyr residues, for every given protein kinase there are hundreds of thousands of potential phosphorylation sites, yet a typical kinase is known to phosphorylate from one to a few hundred protein substrates only. Such a degree of specificity in substrate recognition from a large potential substrate background is the result of a variety of mechanisms including polar, hydrophobic and hydrogen bonding interactions with the substrate in and around the catalytic pocket and also due to distal docking interactions between motifs on the substrate and interaction domains on the kinase.

The kinase catalyzed phosphoryl transfer from ATP to substrates requires an essential divalent metal ion, typically Mg^{2+} or Mn^{2+} , and the phosphoryl transfer occurs via a transient interaction with the substrate without the formation of any stable intermediate. Protein kinase activities are regulated via several mechanisms including phosphorylation/dephosphorylation events of activation loop and/or tail ends, allosteric interactions, interactions with regulatory units and domains, fatty acid acylation, interactions with secondary messengers and subcellular localization. The aberrant regulation of kinase activities have been implicated in many human diseases including several cancers.^{1c,30} Thus the kinome remains an important target for many pharmaceutical drugs, besides the interests in knowing the important roles of these enzymes in cellular signalling cascades.

1. 2. 2 Protein phosphatases

The protein phosphatases are classified into two main classes based on their substrate specificity and protein structure; they are 1) the Ser/Thr protein phosphatases (PPs) and 2) the protein tyrosine phosphatases (PTPs). The PP class is further divided into two distinct gene families namely PPP and PPM. The two classes of phosphatases (i.e. PPs and PTPs) use fundamentally different catalytic mechanisms. Most PTPs utilize a nucleophilic cysteine residue to initiate the formation of an enzyme-phosphate intermediate whereas PPs utilize a dinuclear metal ion centre which is proposed to activate a water molecule for direct hydrolysis of serine/threonine phosphoester bonds in protein substrates.

Protein tyrosine phosphatases (PTPs) are important signalling enzymes that catalyze the dephosphorylation of phosphotyrosine residues in protein substrates. A set of 107 genes in the human genome have been identified to encode the PTPs. The PTP family is classified into four sub-families based on the amino acid sequence of their catalytic domains.³¹ They are 1) Class 1 cysteine-based phosphatases which include both the classical PTPs, which dephosphorylate only phosphotyrosine (pTyr) residues in substrates and dual specific protein phosphatases (DSPs), which can dephosphorylate both pTyr and phosphoserine/phosphothreonine (pSer/pThr) residues, 2) Class 2 cysteine-based PTPs, which include the pTyr specific low molecular weight phosphatases which appear to have evolved from a different ancestor compared to the class 1 cysteine-based phosphatases, 3) Class 3 cysteine-based phosphatases which are both pTyr and pThr specific phosphatases and 4) Class 4 aspartic acid-based phosphatases, which unlike the other three classes of PTPs use an aspartic acid residue and a cation to carry out the catalysis.

1. 2. 3 Catalytic mechanism of PTPs

All the three major PTP classes share the same catalytic mechanism.^{31,32} The PTP signature motif has been identified to be CX₅R(S/T) where the catalytic cysteine residue is of unusually low pKa value (~ 4.5) which enables it an efficient nucleophile to initiate a nucleophilic attack on the phosphorous atom of the phosphorylated substrate. When the substrate binds to the catalytic pocket the enzyme undergoes a conformational change which brings the pTyr residue of the substrate in close proximity to the catalytic cysteine and a nucleophilic attack on the phosphorous atom leads to the formation of a pentacoordinate intermediate whose geometry is stabilized by ionic interaction with the positively charged arginine residue at the active site. An aspartic acid residue near the catalytic site may further stabilize the negative charge in the intermediate via protonating the phosphate. The aspartic acid residue plays multiple roles in the catalysis. It first acts as a general acid to protonate the phenolic hydroxyl in the tyrosine, making it a good leaving group and then once the thiophosphoryl intermediate is formed, the aspartate acts as a general base to activate a water molecule to attack the intermediate to release the cysteine and regenerate the enzyme after the product separation. The role of the conserved serine/threonine residue is presumably the stabilization of the negative charge of the catalytic cysteine thiolate through hydrogen bonding interactions with the side chain hydroxyls and the same interactions may also facilitate the hydrolysis of the thiophosphoryl intermediate.

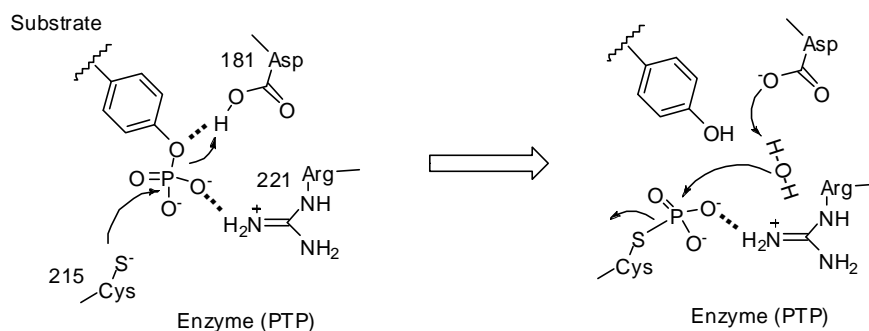


Figure 1.10 Catalytic mechanism of PTPs with PTP1B as a representative^{32b,c}

1. 3 Enzyme inhibitor developments - Fragment-based approaches and high-throughput chemistry

Enzymes, which encompass some of the most valuable drug targets for the potential treatment of several human diseases, receives substantial research interest in the context of potent and selective inhibitor developments. Since every major enzyme family consists of many sub-families and classes, with the structure and functions of the majority of members remaining largely unexplored, developing highly selective and potent inhibitor for a particular member, which is typically the key step in developing many drugs, is often a very challenging task. To meet this requirement, the most inevitable component is synthetic chemistry that can keep up with pace of protein target identification through rapid and efficient synthesis of a large number of diverse chemical compounds. Conventional drug development efforts typically rely on medicinal chemistry both at the very first stage of *lead finding* and at the next stage of *lead optimization*. High-throughput synthetic methods such as combinatorial approaches and synthesis using solid supports together with amenable chemical reactions such as the “click” reactions and multicomponent reactions could potentially expedite these processes.

Many enzymes are known to possess multiple binding pockets yet conventional inhibitor developments generally focus more on only the active site. However, in many cases the secondary/allosteric binding sites confer selectivity as well as potency. To tackle this issue, new strategies such as fragment-based assembly have been emerged and gained significant attention in recent years.³³ Fragment-based assembly is a novel approach which enables high-throughput identification of molecules through the exploration of N^2 possibilities from N+N combination of fragments. Thus in the case of inhibitor developments, N+M combination of inhibitor fragments lead to the generation of N×M potential bidentate inhibitors. It typically involves a two step process namely the fragment identification and the fragment linkage. NMR/X-ray-based strategies^{33,34} and MS-based tethering strategies³⁵ are well-known existing approaches which are now routinely used to facilitate the fragment-based assembly in drug discovery.

1. 3. 1 “Click chemistry” in enzyme inhibitor developments

Fragment-based assembly using click chemistry has drawn significant attention for enzyme inhibitor developments and accelerating the proteomic research in general.³⁶ Click chemistry is a term coined by Sharpless to represent a set of reactions with features such as excellent chemoselectivity, good solvent compatibilities, modularity, minimum synthetic demands and high yields.³⁷ The best known click reaction is perhaps the Cu (I) catalyzed version of the Huisgen 1,3-dipolar cycloaddition reaction³⁸ between azides and terminal alkynes to afford 1,2,3-triazoles, discovered independently by Meldal and co-workers³⁹ and Sharpless and co-workers.⁴⁰ This reaction is highly chemoselective, bioorthogonal (as typically there is no terminal alkynes or azides in living systems) and suitable for rapid and modular assembly of a

large number of compounds. It exploits the inherent kinetic stability of terminal alkynes and azides under a variety of reaction conditions and regioselective formation of 1,4-regio isomer of the triazole is formed in the presence of catalytic amount of Cu (I). Because of the high efficiency and water-compatible nature of this reaction, in most cases, the assembled products could be directly screened for inhibition without the need of any purification.

Various research groups including ours have successfully used click chemistry to assemble inhibitors against important proteins such as protein tyrosine phosphatases,⁴¹ protein kinases,⁴² fucosyltransferases,⁴³ metalloproteases,⁴⁴ aspartic proteases^{14g,45} and caspases.⁴⁶

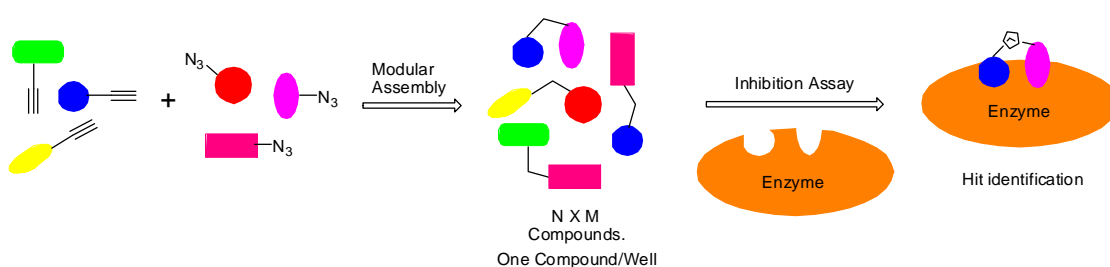


Figure 1. 11 CC facilitates a highly modular and rapid assembly of potential bidentate inhibitors ($N \times M$ compounds) from N number of alkynes and M number of azides. This high-throughput inhibitor assembly can be performed in a 96- or 384-well plate (one compound per well) and the inhibitors in each well can be directly screened with the enzyme (*in situ* screening) without the need to purify the compounds.

1. 3. 2 “*In-situ click chemistry*” facilitated enzyme inhibitor developments

The azide-alkyne cycloaddition reaction has been utilized in a conceptually and technically different manner for the assembly of potent bidentate inhibitors. In this

technique called “*in-situ* click chemistry” the target protein is used as a template for the proximity enhanced cycloaddition of the two fragments. Previously, ingenious approaches such as molecular imprinting techniques⁴⁷ and dynamic combinatorial chemistry (DCC),⁴⁸ both based on proteins acting as a template for the assembly of molecules, have been developed to identify tight-binding partners of proteins. In the former case irreversible linkages are formed between reactive partners that typically recognize protein surface sites whereas the later strategy is active-site directed, where reversible reactions between mixtures of interconverting species are employed. The key to the active-site directed approaches is the ability of the protein to choose and stabilize molecular combinations with highest binding affinity under thermodynamic conditions. Based on Mock’s finding that the very high kinetic barrier in the 1,3-dipolar cycloaddition reaction between azides and alkynes could be tremendously reduced when the two components are brought together into close proximity in a Cucurbituril (a nonadecacyclic cage compound) host structure,⁴⁹ Sharpless and co-workers designed *in-situ* click chemistry as a powerful strategy towards kinetically controlled (unlike the thermodynamic control in the conventional DCC) target-guided synthesis of enzyme inhibitors. In this approach the non-protein host structure in Mock’s original designs were replaced with the active sites of enzymes where the otherwise slow cycloaddition reaction gets heavily accelerated. Sharpless and co-workers have successfully developed this strategy to identify extremely potent inhibitors of enzymes acetylcholine esterase,⁵⁰ carbonic anhydrase II⁵¹ and HIV-1 protease.⁵²

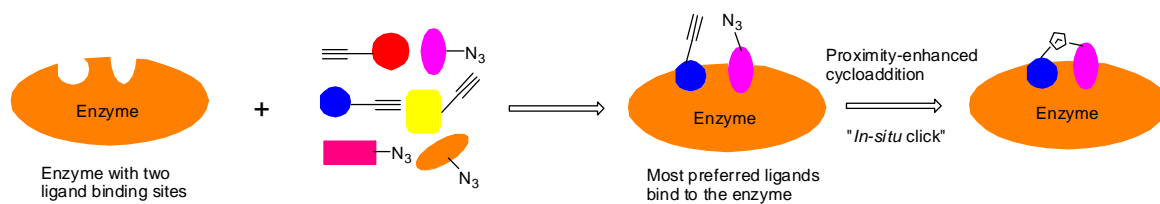


Figure 1. 12 Schematic of *in situ* CC. In the *in situ* CC, the enzyme is incubated with a mixture of alkyne- and azide-equipped ligand scaffolds from which the enzyme chooses the most preferred binders under the kinetic control. In the enzyme active site, the two components react to form the triazole linkage and hence a stable bidentate inhibitor due to proximity-enhanced CC.

Chapter 2: Development of Peptide-Based Activity-Based Probes (ABPs) for Protein Tyrosine Phosphatases (PTPs)

Summary:

This chapter summarizes the development of a panel of peptide-based activity-based probes that allow target-specific profiling of the important signalling enzymes protein tyrosine phosphatases. At the heart of the probe development, we designed and synthesized a new tyrosine phosphatase-reactive unnatural amino acid, termed 2-fluoromethylphosphotyrosine (2-FMPT), and incorporated it into the consensus substrate peptide sequences of different tyrosine phosphatases. The probes, by virtue of the presence of the unnatural amino acid unit covalently react with the enzymes while the presence of the peptide sequence provide an additional degree of specificity for each probe towards its most preferred enzyme. The probes were equipped with fluorescent tags which facilitate the visualization of the labeled protein band by in-gel fluorescence scanning. Labeling reactions with the panel of probes using purified and isolated proteins showed activity-based labeling specificity consistent with the known substrate preferences of the PTPs and the probes were found to be useful for efficient labeling reactions from highly complex biological samples.

2. 1 Introduction

The functional state of PTPs in a cell is highly dynamic and tightly regulated. In human, there are well over 100 known PTPs, which, together with several hundred protein kinases, control the phosphorylation of as many as 1-2% of all human proteins. In order to maintain an intricate balance of this highly complex phosphoproteome network, PTPs carry out dephosphorylation with a high degree of specificity inside

the cell; although *in vitro* they have shown only moderate specificity towards synthetic peptide substrates.⁵³ Consequently, chemical and biological methods that report not only the global PTP activities, but more importantly the precise enzymatic activity of individual PTPs, may offer unprecedented views on how these enzymes carry out their biological functions under native physiological settings.

Activity-based protein profiling (ABPP) is a chemical proteomics strategy that employs active-site directed chemical probes to detect active enzymes present in a crude proteome. Over the past few years, this technique has been successfully applied to different classes of enzymes including phosphatases. In order to develop activity-based probes (ABPs) that detect global PTP activities, Lo *et al* first explored small molecules containing a 4-fluoromethylphenyl phosphate (FMPP) moiety which serves as a phosphotyrosine mimic, and upon dephosphorylation by a PTP, generates a highly reactive quinone methide intermediate that subsequently alkylates nucleophiles present in the PTP active site.⁵⁴ Yao *et al* made similar PTP probes using 2-difluoromethyl phenyl phosphate (DFPP) which presumably works via a similar mechanism as FMPP.⁵⁵ Unfortunately, due to the diffusible nature of the quinone methide generated, both classes of probes possess poor PTP specificity and cross-react with other nearby proteins in the crude proteome. Zhang *et al* recently developed a PTP probe that uses α -bromobenzylphosphonate (BBP) as a nonhydrolyzable phosphotyrosine (*p*Tyr) mimic.⁵⁶ The probe was shown to be specific towards PTPs in proteomic experiments, but the highly reactive/unstable nature of the probe renders it impractical for wide-spread use. The same group developed a newer version of PTP probes by using phenyl vinyl sulfone/sulfonate as the enzyme-targeting warhead.⁵⁷ These probes operate via an electrophilic capture of the active site Cys residue via a 1,4-conjugate addition and these probes, compared to the BBP based probes, have

better hydrolytic stability at pH greater than 7 and were shown to have better membrane permeability as well. Vinyl sulfones/sulfonates are, however, well-established irreversible inhibitors of other cysteine-utilizing enzymes (i.e. cysteine proteases), and are likely to cross-label them as well. Notwithstanding, all four existing classes of PTP probes have so far failed to address one of the most important issues in PTP biology - how can one monitor/profile the active state of individual PTPs in a crude proteome, where these enzymes' native cellular environment is closely emulated?

To address these issues, a novel unnatural amino acid, 2-FMPT was designed, whose incorporation into peptide-based probes could provide a potential means for the target-specific activity-based profiling of PTPs (Figure 2.1a). Apart from an additional 2-fluoromethyl group located in the aromatic ring, 2-FMPT is structurally identical to *p*Tyr and should cause minimal disruption in PTP recognition. The most significant advantage of 2-FMPT over other existing PTP probes is that, with its N- and C-terminus (as in the case of naturally occurring amino acids), essential peptide recognition elements which occupy the proximal positions of *p*Tyr in a naturally occurring PTP substrate could be introduced. With this, the corresponding peptide-based probes could achieve target-specific binding to/profiling of different PTPs. As shown in Figure 2.1b, the peptide-based probe works by first binding to the active site of the target PTP in a sequence-specific manner, and is subsequently dephosphorylated (I; as in the case of a synthetic PTP substrate). The ensuing charge delocalization followed by spontaneous elimination of $-F$ generates the reactive quinone methide intermediate which, upon attack by a nearby nucleophilic residue in the enzyme, would give rise to the covalent enzyme-probe complex (II). We anticipated that, in our newly developed probes, introduction of the additional peptide

fragments specific towards individual PTPs should improve probe binding affinity as well as specificity, thus overcoming the diffusion problem observed in the previous ABPs (i.e. FMPP/DFPP).

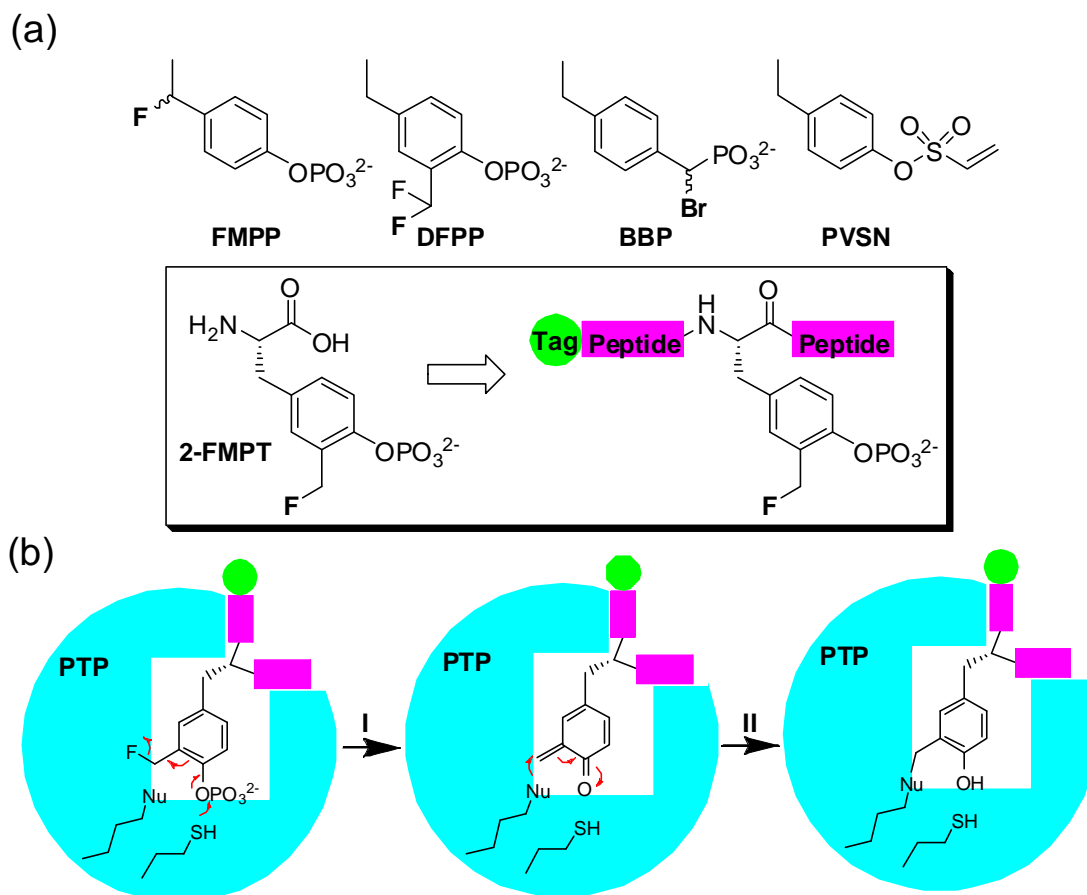
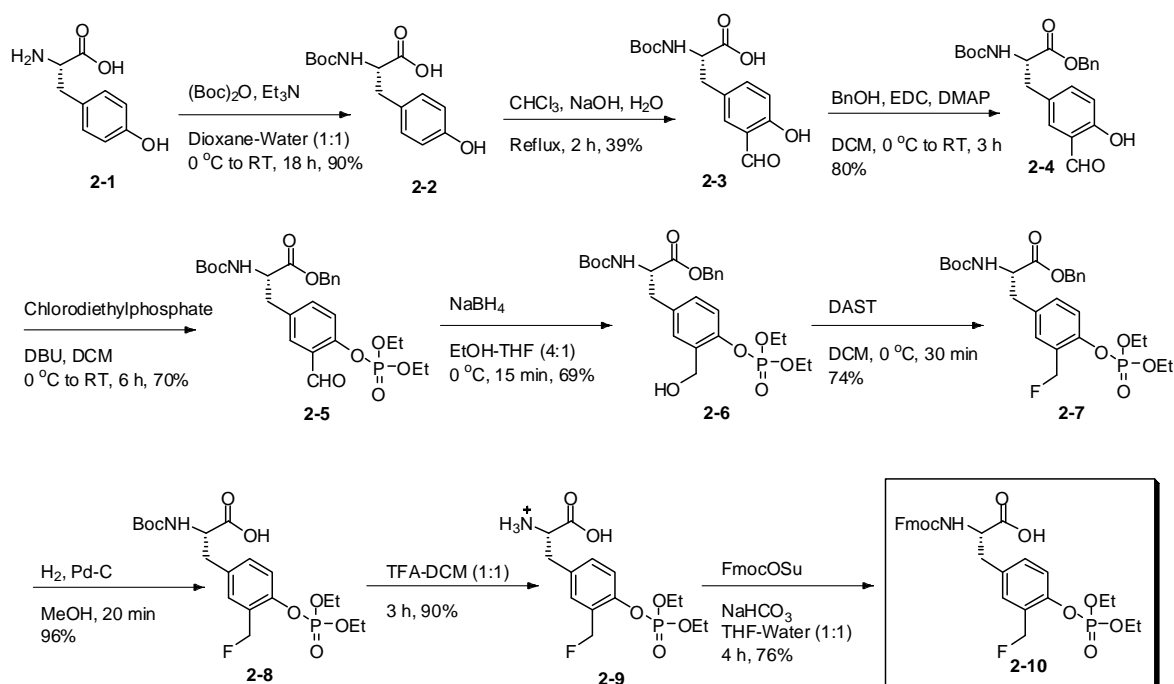


Figure 2. 1 (a) Structures of known ABPs (top) of PTPs and the unnatural amino acid, 2-FMPT, and its corresponding peptide-based ABPs (boxed). (b) Proposed mechanism of the sequence-specific, activity-based labeling of the peptide-based probes.

2. 2. Synthesis of the unnatural amino acid, 2-FMPT

The synthesis of the Fmoc-protected 2-FMPT (**2-10** in Scheme 2.1) was carried out from L-tyrosine. Briefly, the free amino group of L-tyrosine was protected with Boc to get **2-2** (90%). Formylation of **2-2** at the ortho position of the phenolic group by Reimer-Tiemann reaction gave **2-3** in moderate yield (39%).⁵⁸ Subsequently, **2-3** was

protected to give the corresponding benzyl ester **2-4** (80%), followed by phosphorylation with chlorodiethylphosphate to give **2-5** (70%). Subsequent reduction of the CHO group in **2-5** with NaBH₄ gave **2-6** (69%) followed by fluorination with DAST, giving **2-7** (74%). Finally, deprotection of the benzyl ester in **2-7**, followed by Boc-to-Fmoc conversion gave the target compound **2-10**.

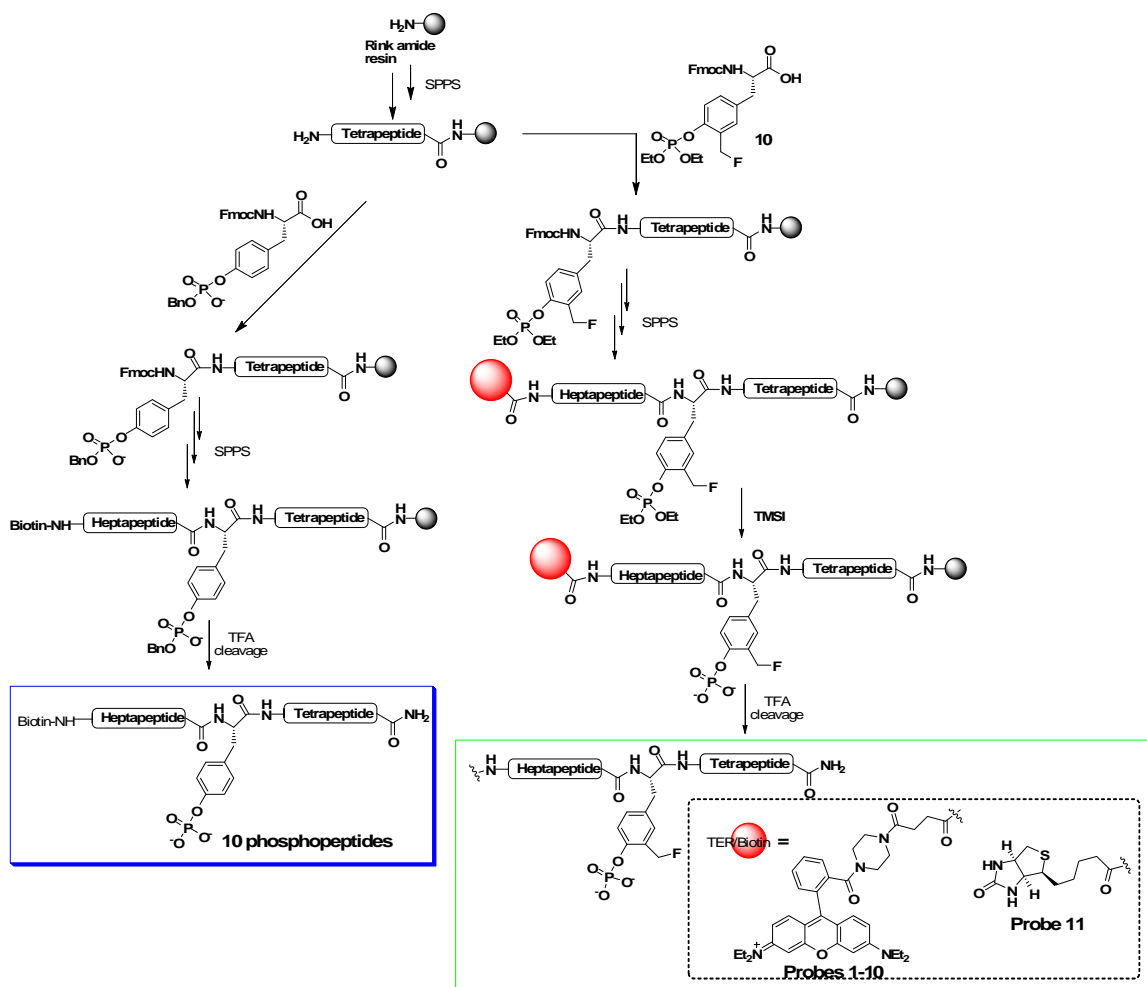


Scheme 2. 1 Synthesis of the unnatural amino acid (2-FMPT)

2. 3 Solid-phase synthesis of substrate peptides and peptide-based activity-based probes

The unnatural amino acid was incorporated into peptides using standard solid-phase peptide synthesis (SPPS) protocols. Eleven peptide probes, designed against five different PTPs (PTP1B, TCPTP, SHP1, SHP2 and LMWPTP), were synthesized using rink amide resin (Scheme 2.2 & Table 2.1). For comparison, ten corresponding

phosphopeptides were also synthesized, in which the naturally occurring *p*Tyr was used in place of 2-FMPT. The SPPS was performed on an automated peptide synthesizer (Chemspeed Technologies ASW2000) utilizing standard Fmoc chemistry on rink amide resin (ChemPep Inc.). All the natural Fmoc amino acids (4 eq.) (Advanced ChemTech) were activated by DIC/HOBt (4 eq. each) (Advanced ChemTech) and each coupling reaction was performed for 2 h in the presence of 8 eq. of DIEA. For coupling of the unnatural amino acid, pre-activation of the unnatural amino acid with HOBt/HBTU (4 eq. each) in the presence of 8 eq. of DIEA for 10 min was performed before being added to the growing peptide chain on the resin. After washing with DMF (6X), the Fmoc-protecting group was removed using 20% piperidine in DMF, followed by another washing with DMF (6X). The N-terminus of each peptide was coupled to a Rhodamine dye containing an acid linker (or biotin) employing the same pre-activation method as earlier described for the unnatural amino acid. Finally deprotection of the phosphate ethyl ester groups was performed by treating the resin-bound peptide with 10 eq. of TMSI in dry DCM, followed by washings with DMF (6X) and DCM (6X). Prior to cleavage of the peptide, the resin was washed with MeOH (6X), DCM (6X), and MeOH (6X), and dried *in vacuo*. Peptide cleavage was carried out with an 1-mL cleavage cocktail (95:2.5:2.5 of TFA:H₂O:TIS) at RT for 2 h. The synthesis and cleavage of the phosphopeptides were done as previously described.⁵⁹ After filtration and precipitation with cold ether, the precipitated peptides were washed twice with cold ether, followed by drying *in vacuo* (Genevac). Each peptide was purified by HPLC and characterized by LC-MS (IT-TOF) (Shimadzu). Appropriate standard solutions were made in DMSO and stored at -80 °C freezer prior to use.



Scheme 2. 2. Solid-phase synthesis of the 10 phosphopeptides and 11 peptide-based ABPs.

Probe	Putative Target	Original Protein Source	Sequence ³
P1	SHP1	Siglec	Rh-GGDEGIHXSELI-NH ₂
P2	PTP1B/TCPTP	Stat3	RhGGGSAAPX ³ LTK-NH ₂
P3	PTP1B	Stat5b	Rh-GGKAVD ³ GXVKPQ-NH ₂
P4	LMWPTP	Zap70	Rh-GGLNSD ³ GXTPEP-NH ₂
P5	SHP2	Ptk2b	Rh-GGLPPEGX ³ VVVV-NH ₂
P6	SHP1/SHP2	CD31	Rh-GGNSD ³ VQXTEVQ-NH ₂
P7	PTP1B	Tyk2	Rh-GGPEGHEX ³ pYRVR-NH ₂
P8	SHP2/PTP1B	Janus	Rh-GGPQDKEX ³ YVK-NH ₂
P9	SHP2/PTP1B	EGFR	Rh-GGVDAEX ³ LIPQ-NH ₂
P10	PTP1B	Zhang's ²	Rh-GGELEFX ³ MDE-NH ₂
P11	PTP1B	Stat5b	Rh-K(Biotin)-KAVD ³ GXVKPQI-NH ₂

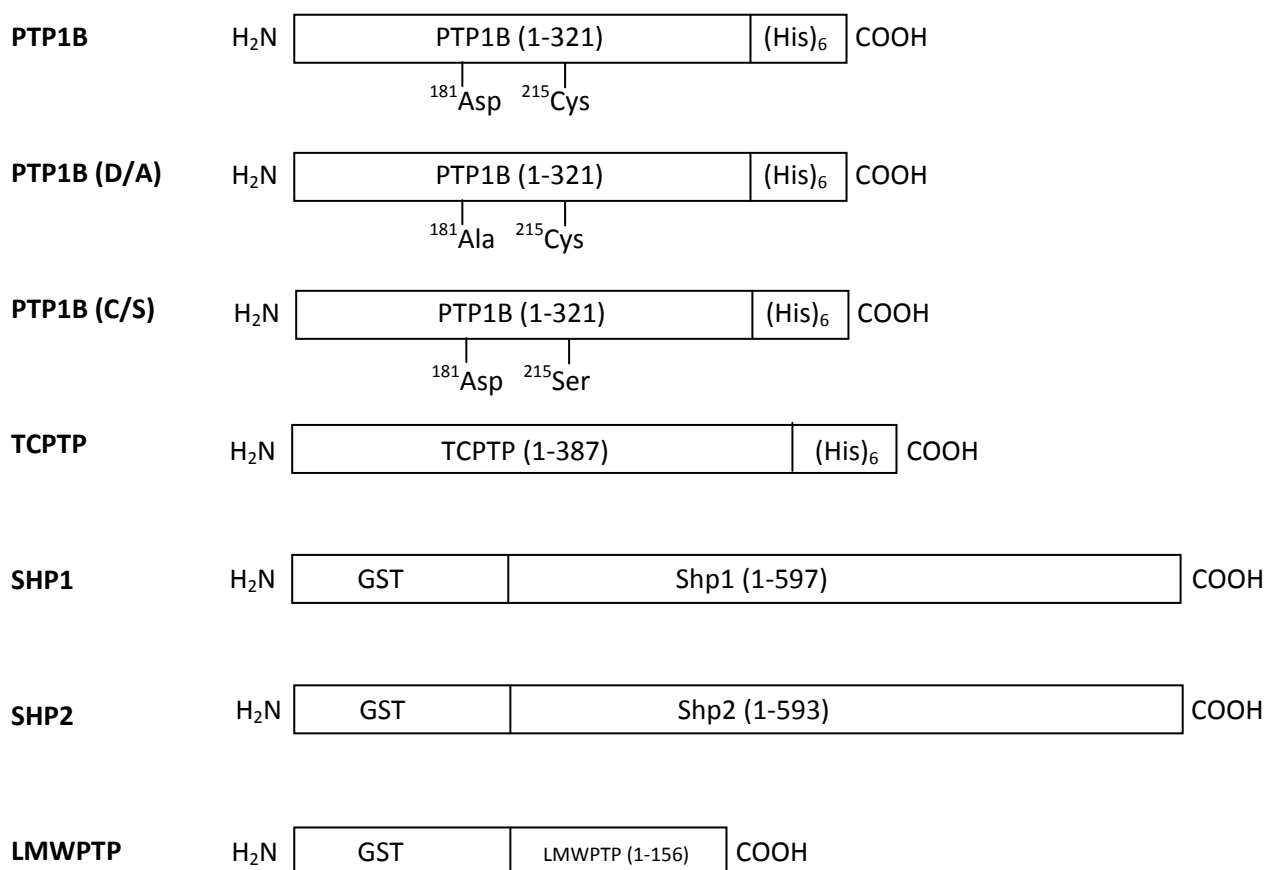
³X = 2-FMPT.

Table 2. 1 The 11 probes and their AA sequences and substrate preferences

2. 4 Expression and purification of PTPs

The catalytic domain (1-321) of PTP1B was amplified by polymerase chain reaction (PCR) and sub-cloned using Nde1 and BamH1 restriction sites into a pET-28b vector to give PTP1B with a C-terminal hexahistidine tag. The two PTP1B mutants (D/A and C/S mutants) were obtained by site-directed mutagenesis using the QuikChange Mutagenesis Kit (Stratagene) following manufacturer's instructions. The final constructs were verified by DNA sequencing and transformed into *Escherichia coli* BL21 DE3 cells for protein expression. A single colony was grown overnight at 37 °C in Luria Bertani media that contain the antibiotic Kanamycin at a final concentration of 50 µg/ml. Subsequently, the overnight culture was diluted 1:100 and grown at 37 °C to an optical density of 0.6-0.8 at 600 nm. Isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.4 mM and grown for a

further 4 hours at 30 °C. The culture was harvested and lysed by sonication on ice, in the presence of complete proteases inhibitor (Roche). The histidine-tagged PTP1B proteins were purified from the clarified lysates on Nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography matrices (Qiagen) according to manufacturer's instructions. The plasmid encoding TC-PTP was a kind gift of Prof Harry Charbonneau and was modified to insert a C-terminal hexahistidine tag in-frame with the TCPTP gene. The mutagenesis was carried out using the QuikChange Mutagenesis Kit (Stratagene) following manufacturer's instructions. The final construct was verified by DNA sequencing and were transformed into BL21 DE3 cells for protein expression. TC-PTP-(His)₆-protein was expressed and purified following the same protocol as for the PTP1B proteins, with the only difference being that IPTG was added to a final concentration of 50 μM and grown for a further 12-16 h at 18 °C. The plasmids encoding Shp1 and Shp2 were obtained from Addgene.org (Plasmid 8578, Plasmid 8322), while the plasmid encoding LMW-PTP was a kind gift of Prof. Tomas Mustelin. All purification steps were performed at 4 °C following the same protocol described above. Briefly, the pelleted cells were resuspended in lysis buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl, 0.2 mM DTT and 1X complete proteases inhibitors. Sonication was done on ice and lysate was clarified by centrifuging at 12000 g for 30 min. The supernatant was then incubated with Glutathione Sepharose 4B beads (Amersham Pharmacia Biotech) for 30 min, with rotation. Subsequently, beads were washed 4X 5 min with 1X PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 1.8 mM KH₂PO₄, pH 7.4) wash buffer, complemented with 0.2 mM DTT and 0.2% Tween 20. Elution was carried out using 20 mM Tris, pH 8.0, and 10 mM reduced glutathione.



Scheme 2. 3 Schematic representation of different PTP constructs used in the work.

2. 5 Results and Discussions

2. 5. 1 Labeling experiments with purified proteins

Labeling reactions with PTPs (e.g. PTP1B, TCPTP, YOPH, LMWPTP) and non-PTP proteins (e.g. 14-3-3 protein, Sortase, BSA & CnA), were initiated by adding 1 μ L of the probe **7** (a representative probe) to a pre-equilibrated enzyme/protein solution in an appropriate buffer at 25 °C. The final probe concentrations in all the reactions were 10 μ M and the protein amounts were 1 μ g. Tris buffer (50 mM, pH = 7.5) with ionic strength adjusted to 150 mM with NaCl was employed for PTP1B and TCPTP and YOPH. Sodium acetate buffer (50 mM, pH = 5.2) was used for LMWPTP.

Labeling reactions with non-PTP proteins were performed in Tris buffer (50 mM, pH = 7.5, 150 mM NaCl). All labeling reactions were run for 1 h and subsequently quenched by the addition of 6X SDS loading dye followed by heating at 95 °C for 10 min. The proteins were resolved on a 10% SDS-PAGE gel and the labeled bands were visualized by in-gel fluorescence scanning with Typhoon 9200 fluorescence gel scanner using green (533 nm) laser with optimum emission filters and the scanned images were quantified with the ImageQuant 3.3 (Molecular Dynamics) software. As shown in Figure 2.2, under standard labeling conditions, the probes were found to label only PTPs in activity-dependant manner.

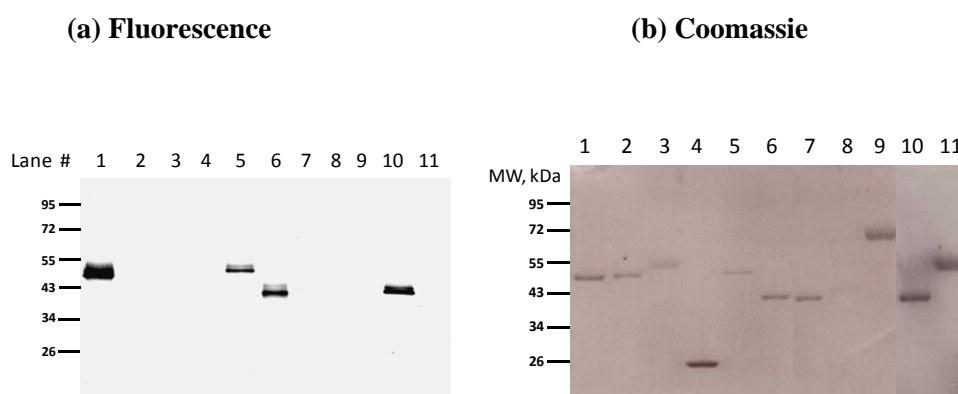


Figure 2. 2 Activity-based labeling of proteins using probe **7**. lane 1- TCPTP; lane 2-Boiled & cooled TCPTP; lane 3- CnA; lane 4- Sortase; lane 5- YOPH; lane 6- PTP1B; lane 7- Boiled & Cooled PTP1B; lane 8- empty; lane 9- BSA; lane 10- LMWPTP; lane 11- 14-3-3 protein.

Activity-based labeling of the ten rhodamine-containing probes, **P1** to **P10**, with five different PTPs was tested to generate the so-called activity-based fingerprints.⁶⁰ As shown in Figure 2.3, clearly distinctive labeling profiles of individual PTPs were observed, and to a large extent corroborated well with their known substrate

preferences (Table 2.1). For example, the top-5 probes (**P2**, **P3**, **P7**, **P9** & **P10**) which generated the strongest labeling against PTP1B, were derived expectedly from known PTP1B substrates. TCPTP, known to have similar substrate preferences as PTP1B, gave a correspondingly similar activity-based fingerprint. For SHP1, two of the most active probes are **P1** and **P6**. Of the two SHP2-specific probes (**P8** & **P10**), one was derived from its known substrate. Similarly, LMWPTP also gave a distinctive labeling profile, which is in good agreement with its substrate preference.⁶¹

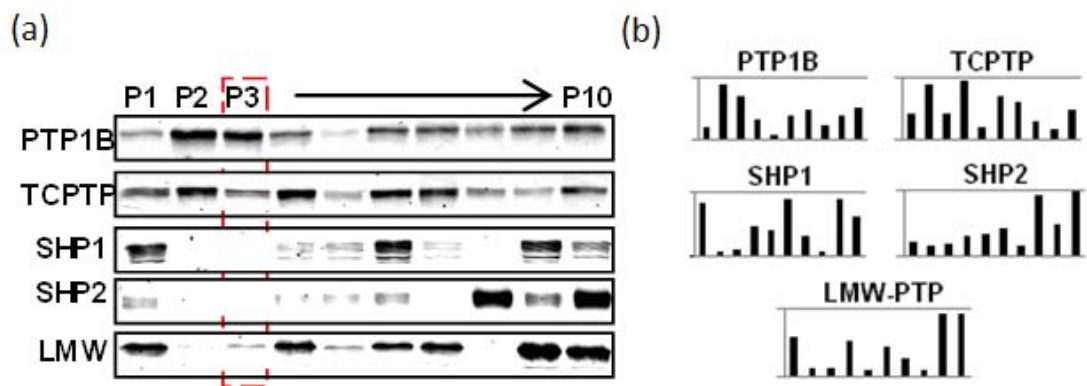


Figure 2. 3 (a) Fluorescent labeling profiles of five different PTPs (top to bottom) with the ten peptide probes (left to right). The fluorescent gels clearly showed a sequence-dependent labeling of different PTPs. The labeling profiles of **P3** were highlighted (boxed in Red). (b) The fluorescent intensity of each band is quantified using ImageQuant software and the relative fluorescence intensity (Y-axis) of labeling of the 5 different PTPs were plotted against the panel of 10 probes (X-axis, P1 to P10 from left to right in each graph).

2. 5. 2 Detection limits of the probes

In order to determine the fluorescence detection limit of the probes, 200 ng of PTP1B was incubated with probe **P3** (one of the highly selective probes for PTP1B) at a final probe concentration of 2 mM in the reaction buffer (50 mM Tris buffer, pH 7.5, 150 mM NaCl) for 1 h. After quenching the reaction, the labeled PTP1B was serially diluted with the same buffer and separated from excess of the probe on SDS-PAGE and the labeled bands were visualized by in-gel fluorescence scanning

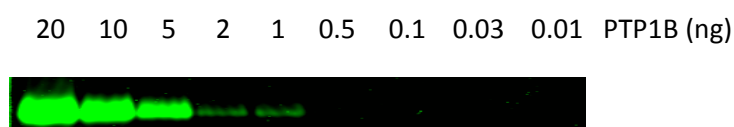


Figure 2. 4 Detection limit of PTP1B with Probe **P3**.

2. 5. 3 Labeling experiments with mutant PTPs

PTPs possess a highly conserved Cys residue in the active site that plays a key role in the catalytic mechanism via initiating a nucleophilic attack on the phosphotyrosine of the substrate, leading to the formation of a thiophosphoryl enzyme intermediate which eventually collapses to give the dephosphorylated product. PTPs with mutations at this position (e.g. ²¹⁵Cys→Ser mutant of PTP1B) are catalytically inactive. The substrate binding into the active site of PTPs is followed by a well characterized conformational change in the highly conserved “WPD” loop in which it flips over the phosphotyrosine residue such that the aspartic acid in this loop gets close enough to the phosphoryl-cysteine intermediate for effective general acid mechanism (protonating the leaving phenolic oxygen). For PTP1B, mutation of this Asp in the WPD motif to Ala (e.g. ¹⁸¹Asp→Ala mutant shown in Scheme 2.3) was shown to significantly reduce the k_{cat} of the enzyme. In order to get more insight into

the activity-based labeling of PTP1B using our probes, we performed labeling experiments with both C/S and D/A mutants of PTP1B (Figure 2.5a). As anticipated, the labeling result showed a complete lack of labeling with the mutants as well as the heat-denatured PTP1B, showing that the probes are indeed activity-based. Microplate-based enzymatic assay of the mutants using DiFMUP as substrate, revealed no catalytic activity for the C/S mutant while the D/A mutant retained a very low activity (Figure 2.5b).

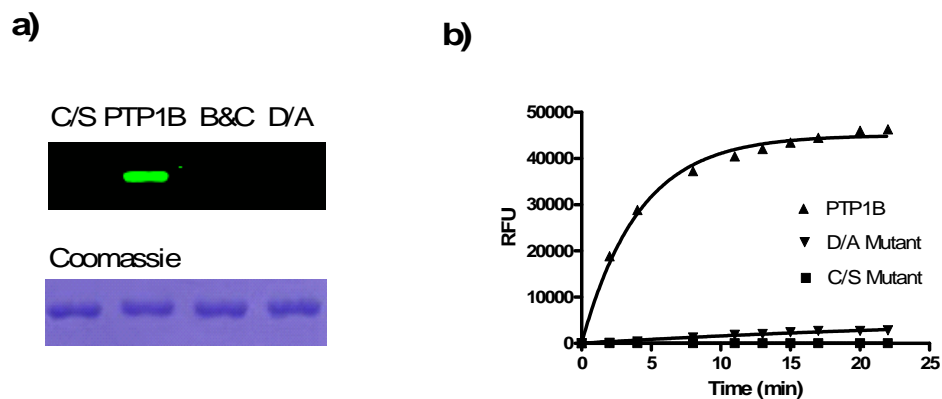


Figure 2. 5 (a) Labeling experiment with mutant PTPs. lane 1. PTP1B C/S mutant; lane 2. PTP1B; lane 3. Boiled and Cooled PTP1B; lane 4. PTP1B D/A mutant. (b) Microplate-based enzymatic assay of PTP1B and mutants used with DiFMUP as the fluorogenic enzyme substrate.

2. 5. 4 Effect of H₂O₂ on PTP activity assessed with the probes

The catalytic Cys residue in the active site of PTPs exhibit an unusually low pKa value which makes it an efficient nucleophile and at the same time makes the enzyme susceptible to oxidative inactivation. H₂O₂ generated during cellular insulin stimulation is shown to cause inactivation of PTPs via oxidizing the catalytic Cys to

cysteine sulfenic acid (Cys-SOH) or a cyclic sulfenamide species.⁶² Since cysteine sulfenic acid and the cyclic sulfenamide species can be reduced back to free cysteine by various cellular reductants, this mode of oxidative inactivation of PTPs may have a significant role in controlling the tyrosine phosphorylation-dependant signal transduction events. If the PTP is exposed to higher levels of oxidants the catalytic Cys residue gets irreversibly oxidized to either sulfinic acid (Cys-SO₂H) or sulfonic acid (Cys-SO₃H) which results in abrogation of enzymatic activity. Thus in order to assess the feasibility of employing our activity-based probes for monitoring the H₂O₂-mediated PTP inactivation, PTP1B (1 μg) was treated with 1 mM H₂O₂ in 20 μL reaction at 25 °C for 5 min followed by treatment with the Probe **3** for 1 h. The reaction was quenched with 6X SDS loading dye followed by heating at 95 °C for 10 min. A control experiment without H₂O₂ treatment was also performed simultaneously. As shown in Figure 2.6, H₂O₂-treated PTP1B didn't not show any trace of labeling showing the effectiveness of the probe in distinguishing the active versus oxidatively inactivated forms of the enzyme.

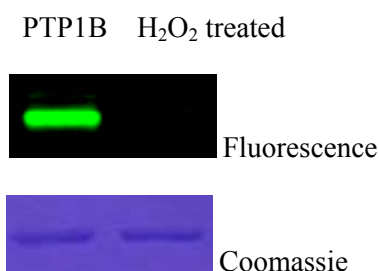


Figure 2. 6 Effect of H₂O₂ on PTP1B activity assessed with the probe

2. 5. 5 Kinetic characterizations and substrate specificities of the probes and the corresponding phosphopeptides

Kinetic characterizations of the labeling reactions were carried out, with PTP1B as the model, by time- and concentration-dependent measurements of the enzyme's inactivation in the presence of the ten probes, as well as measurements of PTP1B-mediated dephosphorylation with ten corresponding phosphopeptides. The kinetic experiments were performed at 25 °C in HEPES buffer (50 mM, pH = 7.5, 150 mM NaCl). Appropriate dilutions of the probes (0.2 mM to 2 mM) in the reaction buffer (45 μ L) were treated with 5 μ L aliquots of PTP1B stock solution. From this, 2 μ L was withdrawn at appropriate time intervals (1 to 60 min) and added into 48 μ L of a 20 μ M solution of DiFMUP in the same reaction buffer in a black 384 well flat-bottom plate. The fluorescence intensity in relative fluorescence units (RFU) were continuously measured for 15 min (e.g. until a consistent trend is observed) using a Tecan fluorescence plat reader at excitation wavelength of 360 nm and emission wavelength of 460 nm at a gain of 65. Background corrections for non-enzymatic hydrolysis of the DiFMUP were made. The experiments were performed in triplicate and the data for each probe were averaged and used to derive the kinetic parameters. The inactivation reactions follow Kitts and Wilson kinetics⁶³ and the pseudo-first order inactivation constants (k_{obs}) at various concentrations of the probes were obtained from the initial slope of the semilogarithmic curves of $\ln(v_t/v_0)$ versus preincubation time, where v_t and v_0 are the enzyme activities at time t and zero, respectively. A non-linear regression fit of the double-reciprocal plot of k_{obs} versus probe concentration gave the maximal inactivation rate constant (k_i) and dissociation constant K_i according to the following equation.

$$k_{\text{obs}} = k_i \times [I] / K_i + [I] \quad \text{Eq. 1}$$

Ten phosphopeptides with the same amino acid sequence as those in the probes were also synthesized (with the only difference being that in the 10 phosphopeptides, phosphotyrosine was used in place of the unnatural amino acid, and biotin in place of the rhodamine dye) and the dephosphorylation preference of these peptides by PTP1B was measured in a microplate-based enzymatic assay using the malachite green phosphatase assay.⁶⁴ First, a standard phosphate assay was performed in order to quantify the absorbance values (Figure 2.8a). For the phosphate standard curve, 20 μL aliquots of different concentrations of phosphate (0.02 to 120 μM) prepared in TETBS buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.05% Tween20) were treated with 5 μL of malachite green reagent in a 384-well plate (transparent flat-bottom) at 25 $^{\circ}\text{C}$ for 15 min and the absorbance at 620 nm was measured using the Tecan plate reader. The absorbance was found to be linear in the range of 30 to 800 picomoles of phosphate (inset in Figure 2.8a). Next, the phosphatase assay was performed with 3 $\mu\text{g}/\text{mL}$ of PTP1B and a 100 μM solution of each of the 10 phosphopeptides in TETBS buffer. The reaction was incubated for 10 min at 25 $^{\circ}\text{C}$ in a final reaction volume of 20 μL . The phosphatase reaction was quenched by adding 5 μL of malachite green reagent containing hydrochloric acid and the absorbance at 620 nm was measured after 15 min. The experiments were carried out in duplicate and averaged values were used to derive the number of nanomoles of inorganic phosphate released per min per mg of the enzyme (Table 2.2). Corrections were made for the non enzymatic hydrolysis of the peptides. In order to further verify the substrate specificity, the k_{obs} values for the dephosphorylation reaction were obtained via following the time course of dephosphorylation (Figure 2.8b & Table 2.2). Briefly,

the Absorbance values were measured at 620 nm over 60 min. and fitted into the following equation,

$$A = A_{\infty} [1 - \exp(-k_{\text{obs}} \cdot t)] \quad \text{Eq. 2}$$

where $k_{\text{obs}} = (k_{\text{cat}}/K_M)$. E , represents the observed k_{cat}/K_M at a given enzyme concentration (E). The assay was run under conditions where the enzyme and peptide substrate concentrations were well below the K_M , such that the relative rates of dephosphorylation (k_{obs}) truly reflect the k_{cat}/K_M .

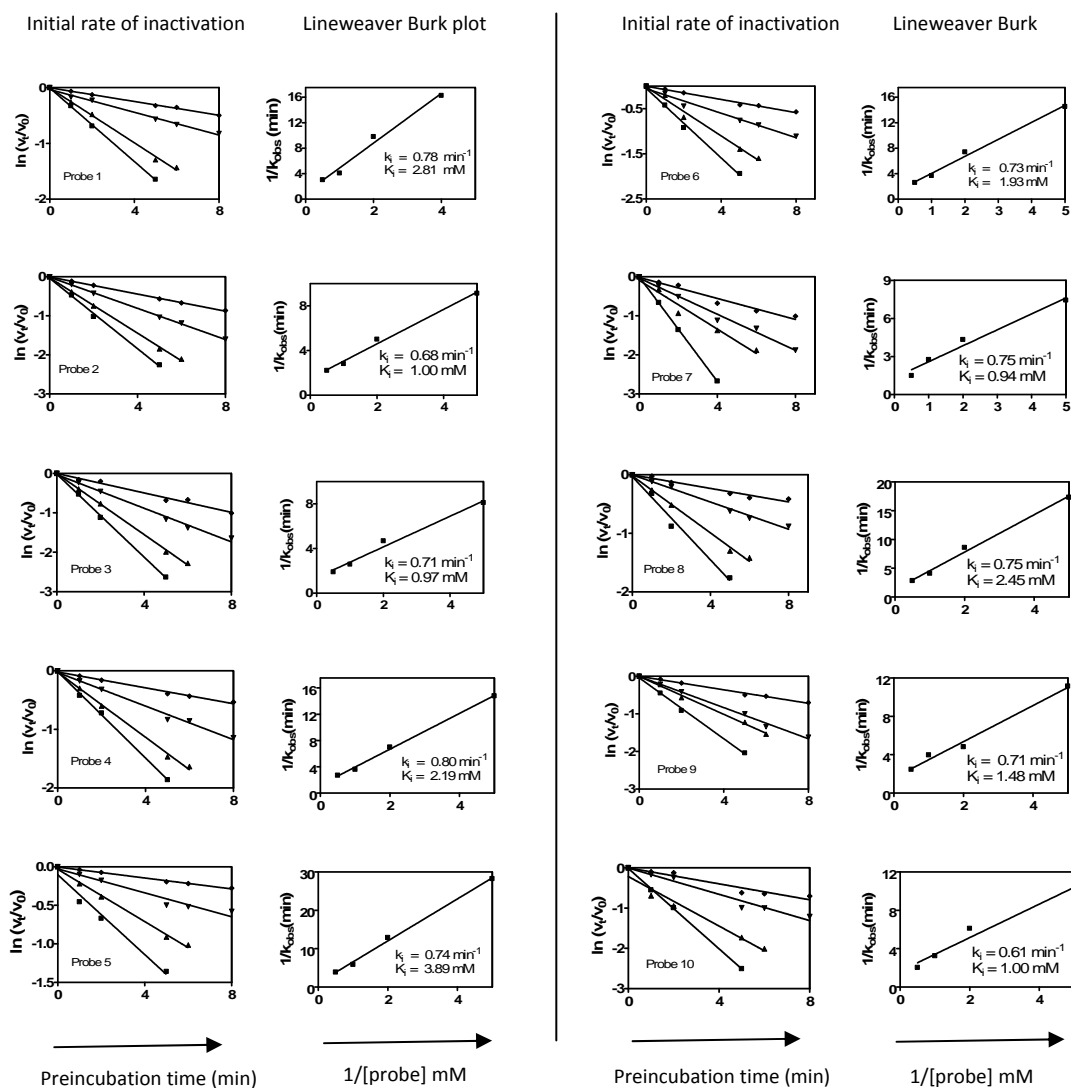


Figure 2. 7 Determination of the kinetics of inactivation for the 10 probes on PTP1B

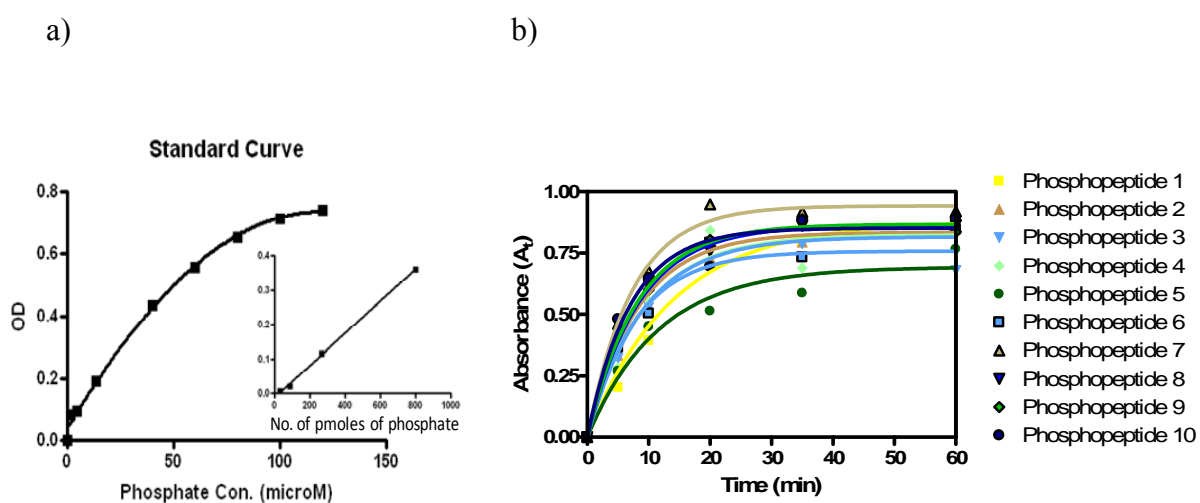


Figure 2. 8 (a) Standard curve of phosphate detection using the Malachite green assay, with inset showing the linear portion (30 to 800 pmoles of phosphate). (b) Time-dependant variation in the absorbance value for the 10 phosphopeptides with PTP1B. The k_{obs} values were derived from the curves using eq.2.

Probe	K_i (mM)	Phosphopeptide	Phosphate released (nmol/min/mg)	k_{obs}
1	2.81	1	6741	0.072
2	1.00	2	10524	0.121
3	0.97	3	13549	0.127
4	2.19	4	9083	0.109
5	3.89	5	7528	0.086
6	1.93	6	8707	0.111
7	0.94	7	14419	0.137
8	2.48	8	11180	0.107
9	1.48	9	12397	0.123
10	1.00	10	15749	0.143

Table 2. 2 Comparison of the substrate specificity of PTP1B obtained from the probe mediated inactivation of the enzyme (K_i values) and that from the standard Malachite green phosphatase assay of the corresponding phosphopeptides. Red indicates preferred substrate, Yellow for substrate and Blue indicates no substrate or very week substrate.

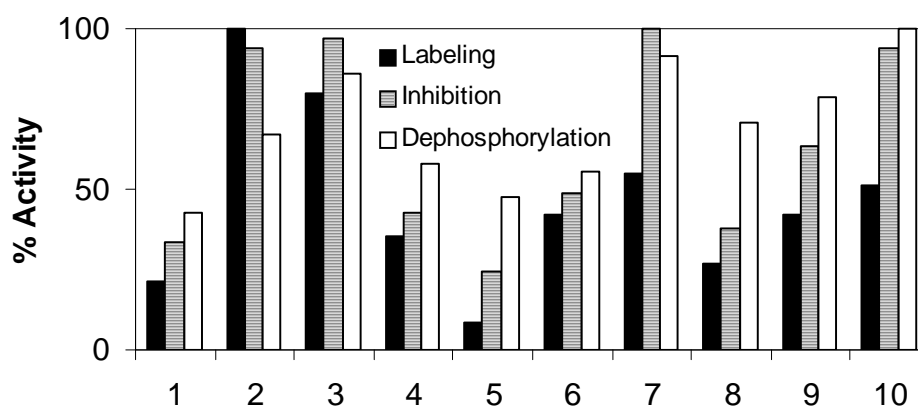


Figure 2. 9 Comparison of relative activity of the 10 probes against PTP1B as determined from (■) quantitative analysis of the fluorescent gel shown in Figure 2.3a, (▨) $1/K_i$ values obtained from inactivation kinetic experiments and (□) data obtained from the dephosphorylation of ten corresponding phosphopeptides catalyzed by PTP1B. All three sets of data were internally normalized. % Activity: relative extent of labeling/inactivation/dephosphorylation for each probe/peptide (100%: most active probe/peptide within each set).

Analysis of pseudo-first-order inactivation rate constants, k_{obs} , as a function of probe concentration showed time- and concentration-dependant inactivation with saturation kinetics typical of irreversible inhibition. The dissociation constants of the probes reflect their substrate preferences, with highest $1/K_i$ values correlated well with the most active probes (in their fluorescent labeling), as well as the most preferred phosphopeptide substrates (in the amount of PO_3^{2-} release). All these lines of evidence suggest our peptide-based activity-based probes can be used to report not only PTP enzymatic activities, but also the substrate specificities of individual PTPs.

2. 5. 6 Labeling experiments in the presence of complex proteomes

2. 5. 6. 1 Labeling experiments in the presence of bacterial cell lysate

To evaluate the performance of the probes in a complex proteome, we set up the labeling reaction with spiked PTP1B in a bacterial lysate. Cell lysates from *E. coli* strain BL21 (DE3) were used as there is no known PTP present in this bacterial genome. The procedures for bacterial culturing and cell-lysate preparation are explained in section 2.6.1. Selectively labeling of spiked PTP1B was observed even in the presence of large excess of other proteins in the bacterial proteome, demonstrating the highly selective nature of the probe's PTP trapping reaction. Also labeling reactions with the panel of the 10 probes (10 μ M in reaction) in the presence of the bacterial proteome (10 μ g total protein) with 0.1 μ g of spiked PTP1B generated labeling fingerprint (Figure 2.10) identical with the fingerprint obtained using pure PTP1B, demonstrating the usefulness of the probes in a complex proteome.

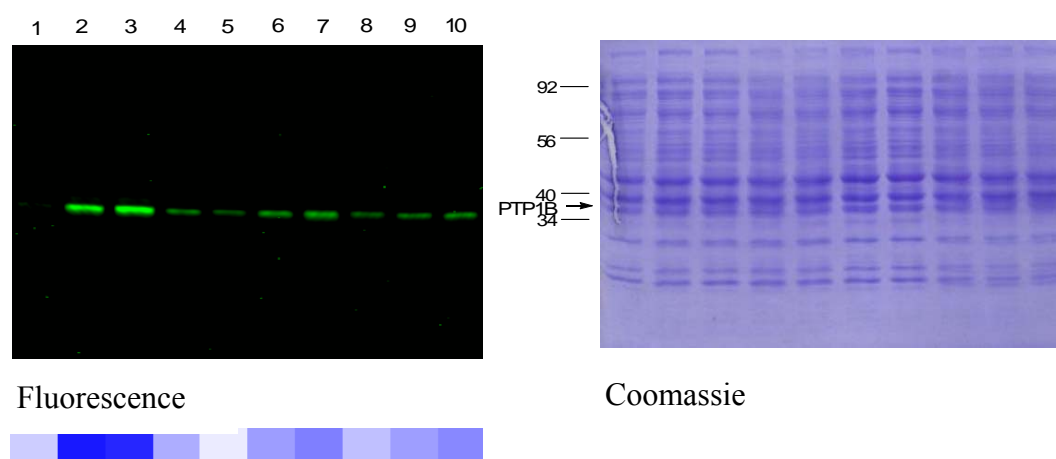


Figure 2. 10 Labeling fingerprint of PTP1B spiked in the bacterial proteome using the panel of the 10 probes.

2. 5. 6. 2 Labeling experiments with mammalian cell proteome

Total cell lysates from two different mammalian cells (HEK293T and NIH3T3) were prepared and treated with **P3** (25 μ M in reaction) in Tris buffer (50 mM Tris, pH=7.5) at 25 °C for 1h (Figure 2. 11b). The proteins were separated on SDS-PAGE and the labeled bands were visualized by in-gel fluorescence scanning (left panel) and Western blotting using anti-PTP1B antibody (right panel); Western blotting indicated the presence of endogenously expressed PTP1B in HEK293T (~52 kDa) but not in NIH3T3 cells. A corresponding fluorescent band was observed in the fluorescence gel from HEK293T cell lysate (lane 1, highlighted in red box). Neither the probe nor the antibody detected PTP1B in the NIH3T3 cells (lanes 2 in Figure 2. 11b). This is in agreement with the previously reported extremely low or undetectable expression levels of endogenous PTP1B in NIH3T3 cell lines.⁶⁵ In addition to the endogenous PTP1B, the probe also labeled a number of other proteins in the HEK293T cell lysates. This might have originated from other endogenous PTPs which were present in the cell lysate and accepted **P3** as a substrate. To further validate the labeling of endogenous PTP1B, we performed pull-down experiments using **P11** - the biotinylated version of **P3**. As shown in Figure 2.11c, upon labelling with **P11**, the HEK293T cell lysate was subjected to affinity enrichment and subsequently analyzed by Western blotting; endogenous PTP1B was successfully detected (lane 2). This experiment thus demonstrates the feasibility of our peptide-based activity-based probes for target-specific profiling of PTP activities in cellular lysates.

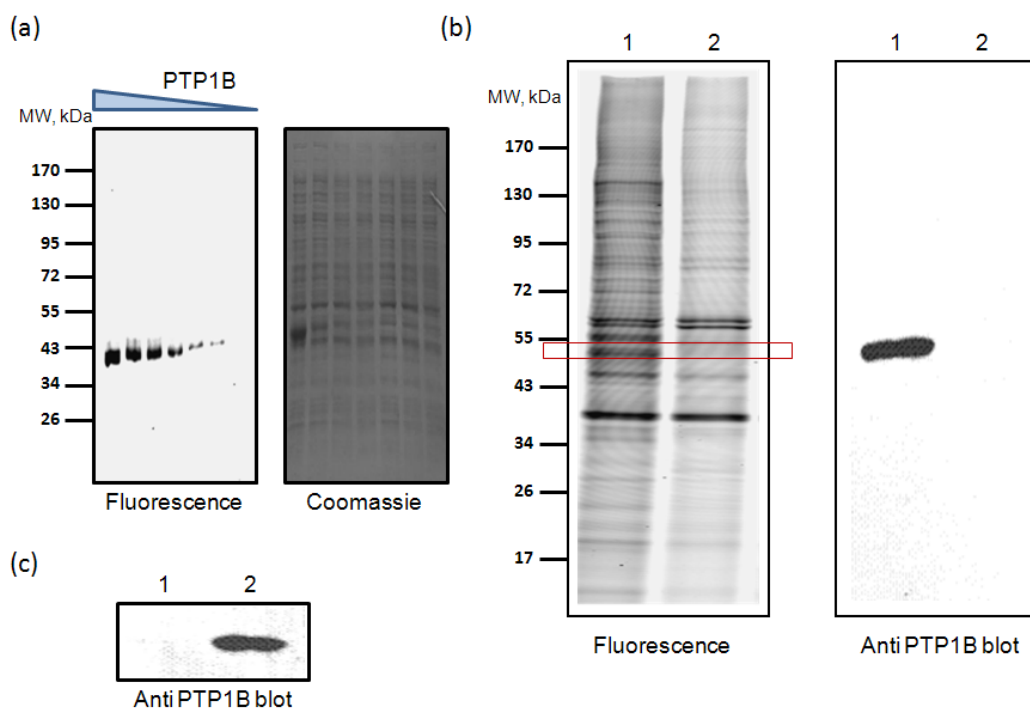


Figure 2. 11 (a) Labeling of spiked PTP1B in the presence of bacterial proteome. Different amounts (left to right: 1000 ng, 400 ng, 200 ng, 100 ng, 20 ng, 5 ng & 0.5 ng) of recombinant PTP1B was spiked to 10 µg BL21 bacterial proteome and the labeling was performed with 10 µM probe **P3** for 1 h. (b) Labeling of PTPs in mammalian cell lysates. Left panel, in-gel fluorescence analysis of global PTP activity profiles obtained from total cell lysates of HEK293T cells and NIH3T3 cells (20 µg total proteins in each lane) with probe **P3**. Right panel, Anti PTP1B blot of the two labeled lysates detecting endogenous PTP1B at ~52 kDa from the HEK293T cell lysate (lane 1) while no PTP1B was detected from the NIH3T3 cells (lane 2). (c) Pull-down results using the biotinylated probe **P11** with HEK293T cell lysate followed by enrichment with NeutrAvidine agarose beads. Anti PTP1B WB of the probe treated HEK293T cell lysate after pull-down experiments detected the endogenous PTP1B at ~52 kDa (lane 2) while no PTP1B was detected from the untreated pull-down fraction (lane 1).

2. 6 Conclusions

In conclusion, a novel unnatural amino acid (2-FMPT) which is a close mimic of phosphotyrosine was synthesized and incorporated into peptide sequences to generate peptide-based probes for PTPs. With the peptide-based probe design, target-specific profiling of PTPs present in mammalian cell lysates was demonstrated using PTP1B as an example. Further improvement of the probes may be achieved by incorporation of longer peptide recognition sequences or by genetically or semi-synthetically incorporating the unnatural amino acid into suitable protein-based substrates of PTPs.⁶⁶ Ultimately, it might be possible to extend the strategy to *in vivo* profiling of PTP/substrate interactions in living cells.

2. 7 General procedures for sample preparations and labeling experiments using proteomes

2. 7. 1 *Preparation of bacterial cell lysates and labeling experiments using the lysates*

Bacterial lysates were prepared as follows. Briefly, an overnight culture of LB (100 ml) inoculated with a single colony of the bacteria, and grown at 37 °C with shaking (300 rpm), was harvested by centrifugation at 4000 rpm for 10 min at 4 °C. The resulting Pellets were suspended in the lysis buffer (50 mM Tris at pH 7.5, 150 mM NaCl) and sonicated (to complete lysis, 20 rounds of 1s on and 5s off, at 25 % amplitude), before centrifugation for 10 min (1000 rpm at 4 °C). The resulting supernatant was further centrifuged for 15 min (13000 rpm at 4 °C). The total protein concentration of this lysate was then quantified by Bradford's Assay (Biorad), and adjusted to 6.0 mg/ml, aliquoted and stored in -20 °C, and used for all subsequent labeling experiments.

In a typical labeling reaction, different amounts of PTP1B (1 µg to 0.05 ng) was spiked into 10 µg of the bacterial proteins in Tris buffer (50 mM Tris, pH 7.5, 150 mM NaCl) at a reaction volume of 19 µL at 25 °C. To this was added 1 µL of the probe 3 (10 µM in the reaction) and incubated for 1h at 25 °C before protein separation on SDS-PAGE and in-gel fluorescence scanning.

2. 7. 2 *Procedure for Western Blot Analysis*

The proteins in the gel after in-gel fluorescence scanning were electrotransferred overnight at 4 °C to a PVDF membrane. The membrane was blocked with 5% non-fat dry milk in Tris-Buffered Saline with 0.1% Tween 20 (TBS-T) for 1 h at 25 °C. After

two quick washes with TBS-T, the membrane was treated with PTP1B antibody (abcam, product code: ab52650, EP1837Y) in 5% BSA in TBS-T (1:10⁵ dilution) for 2 h at 25 °C. The membrane was then washed three times (5 min for each wash) with TBS-T and subsequently incubated with goat anti-rabbit HRP antibody (1:2000 dilution in TBS-T) for 1 h. The secondary antibody treated blot was then washed three times (10 min for each wash) with TBS-T, rinsed with TBS and subsequently treated with HRP substrate. Excess chemiluminescence reagent was removed and the image was acquired using dark room development facilities.

2. 7. 3 Procedure for Pull-down of biotinylated probe labeled PTP

The pull-down followed by immunoblot experiments were carried out as described below. HEK293 cell lysates were prepared as previously described. Before PTP probe labeling, lysates was incubated with washed NeutrAvidin-Agarose beads (Pierce) to remove endogenously biotinylated proteins. This flow-through fraction was used for PTP probe labeling. After 1 h reaction at RT, proteins were precipitated from the lysate by acetone and re-dissolved in solubilizing buffer (PBS with 1% SDS). Before incubation with NeutrAvidin Agarose Beads, samples volumes were adjusted by adding PBS to dilute SDS to 0.1%. After incubation with washed beads for 1 h at RT, beads were washed with 10 bed volumes of PBS with 0.1%SDS for 3 times. Bounded protein was eluted by boiling in SDS sample buffer. For immunoblot analysis, this pull down sample was then separated on 12% SDS-PAGE gel, transferred to PVDF membrane and probed with anti-PTP1B antibody. 52 kDa PTP1B band was observed in the pull-down elute fraction while not in the negative control (lysate without probe treatment) pull-down fraction.

2. 8 Synthetic details and characterizations of compounds

2-(*tert*-Butoxycarbonylamino)-3-(4-hydroxyphenyl) propanoic acid (2-2):

To a solution of L-Tyrosine, **2-1** (Aldrich, 40 g, 220.8 mmol) in 1/1 dioxane/water (250 mL each) was added triethylamine (46 mL, 1.5 eq., and 331.2 mmol). The reaction flask was cooled to 0 °C with an ice bath and di-*tert*-butyl-dicarbonate (53 g, 1.1 eq., 331.2 mmol) was added. After 1 h, the cold bath was removed and the reaction mixture was stirred at ambient temperature for 20 h. The reaction mixture was then concentrated on a rotary evaporator and the residue diluted with water and ethyl acetate. The aqueous layer was acidified to pH 1 with 1 N HCl and back extracted with ethyl acetate. The organic layer was washed with brine, dried over Na₂SO₄ and evaporated to give the protected amino acid, Compound **2-2**, as a white solid (56 g, 90%) which was used without further purifications. ¹H- NMR (300 MHz, CDCl₃) δ 6.95 (2H, d, *J* = 8.22 Hz), 6.70 (2H, d, *J* = 7.41 Hz), 5.92 (1H, bs, OH), 5.16 (1H, bs, NH), 4.53 (1H, unresolved m, α H), 3.01 (2H, unresolved m, βH), 1.40 (9H, s, Boc); ESI-MS: *m/z* = 280.2 [M-H]⁻

2-(*tert*-Butoxycarbonylamino)-3-(3-formyl-4-hydroxyphenyl) propanoic acid (2-3):

Compound **2-2** (40 g, 142.3 mmol) in 600 mL of CHCl₃ was heated to 65 °C. To this was added a hot (65 °C) solution of NaOH (22.8 g, 4 eq., 569 mmol) in 80 mL of water and the mixture was refluxed for 1.5 h. After cooling, CHCl₃ was evaporated off and the reaction mixture was diluted with water and EtOAc, and the aqueous layer was acidified to pH 1 with 1 N HCl and back extracted with EtOAc. The organic layer was washed with brine, dried over Na₂SO₄ and concentrated. Flash column chromatography (silica gel, 12/1 CHCl₃/MeOH with 1% acetic acid eluent) afforded

2-3 as a pale yellow oil (17.1 g, 39%). Unreacted starting material **2-2** was also recovered. ¹H-NMR (300 MHz, CDCl₃) δ 10.9 (1H, s, OH), 9.8 (1H, s, CHO), 7.36 (1H, s, C2-H), 7.32 (1H, d, *J* = 8.37 Hz), 6.91 (1H, d, *J* = 8.37 Hz, C5-H), 5.27 (1H, bs, NH), 4.57 (1H, unresolved m, α H), 3.15 (1H, unresolved m, β H), 3.02 (1H, unresolved m, βH), 1.38 (9H, s, Boc); ¹³C-NMR (75.5 MHz, CDCl₃) δ 196.52, 177.03, 160.52, 155.29, 138.08, 134.23, 127.62, 120.37, 117.72, 80.43, 54.16, 36.83, 28.16; ESI-MS: *m/z* = 308.1 [M-H]⁻

Benzyl 2-(*tert*-butoxycarbonylamino)-3-(3-formyl-4-hydroxyphenyl) propanoate (2-4):

To a solution of **2-3** (10 g, 32.36 mmol) in CH₂Cl₂ was added benzyl alcohol (6.73 mL, 64.72 mmol, 2 eq.) followed by DMAP (0.6 g, 0.15 eq.) and the solution was cooled to 0 °C with an ice-water bath while stirring. EDC (7.45 g, 38.83 mmol, 1.2 eq.) was added and after 30 min the ice bath was removed, and continued stirring at room temperature for 3 h. The reaction mixture was then extracted with water, brine, dried over Na₂SO₄ and concentrated. Flash column chromatography (silica gel, CHCl₃/MeOH, 12/1) afforded **2-4** as a pale yellow oil (13.8 g, 80%) ¹H-NMR (300 MHz, CDCl₃) δ 10.84 (1H, s, OH), 9.59 (1H, s, CHO), 7.33-7.24 (5H, unresolved m), 7.12 (2H, m), 6.8 (1H, d, *J* = 8.4 Hz) 5.27 (1H, bs, NH), 5.10 (2H, m, CH₂-benzyl), 4.57 (1H, unresolved m, α H), 3.15 (1H, unresolved m, βH), 3.02 (1H, unresolved m, βH), 1.38 (9H, s, Boc) ; ¹³C-NMR (75.5 MHz, CDCl₃) δ 196.37, 171.37, 160.40, 154.99, 137.99, 135.04, 134.07, 128.57, 127.56, 120.26, 117.58, 79.95, 67.08, 54.34, 37.03, 28.21

Benzyl-2-(*tert*-butoxycarbonylamino)-3-(4-(diethoxyphosphoryloxy)-3-formylphenyl) propanoate (2-5):

Compound **2-4** (8 g, 20 mmol) was dissolved in CH₂Cl₂ (90 mL). To this was added DBU (6 mL, 40 mmol, 2 eq.) and the reaction mixture was cooled to 0 °C with an ice-water bath while stirring under N₂ gas. After 15 min, diethylchlorophosphate (6.4 mL, 44 mmol, 2.2 eq.) was added and the ice bath was continued for 45 min, and after that the ice bath was removed and stirring continued at room temperature for 8 h. The reaction was extracted with 5% citric acid solution, brine, dried over Na₂SO₄ and concentrated. Flash column chromatography (silica gel, Hexane/EtOAc, 10/2) afforded **2-5** as a pale yellow oil (7.5 g, 70%). ¹H-NMR (300 MHz, CDCl₃) δ 10.29 (1H, s, CHO), 7.35-7.28 (5H, unresolved m), 7.27-7.23 (2H, m C2-H & C6-H), 7.21 (1H, d, J = 8.37 Hz, C5-H), 5.10 (2H, m, CH₂-benzyl), 4.55 (1H, unresolved m, α H), 4.22 (4H, q, CH₂-phosphate) 3.14 (1H, unresolved m, βH), 2.98 (1H, unresolved m, βH), 1.33 (9H, s, Boc), 1.28 (6H, t, phosphate) ; ³¹P-NMR (121.5 MHz, CDCl₃) δ = -6.0902 (s); ESI-MS: m/z = 558.2 [M+Na]⁺

Benzyl-2-(tert-butoxycarbonylamino)-3-(4-(diethoxyphosphoryloxy)-3-(hydroxymethyl) phenyl) propanoate (2-6):

To a solution of the aldehyde **2-5** (6 g, 11.2 mmol) in absolute EtOH (90 mL), THF (20 mL) and water (900 μL) was added NaBH₄ (0.43 g, 11.2 mmol) at 0 °C and stirred for 15 min. The reaction mixture was then diluted with acidic water (pH = 4; 80 mL) and stirring continued for 10 min. The mixture was extracted with EtOAc (2 × 90 mL) and concentrated under reduced pressure. The oily compound obtained was then coevaporated with MeOH three times, again dissolved in EtOAc, extracted with brine, dried over Na₂SO₄ and concentrated at reduced pressure. It was then purified by silica gel flash chromatography (DCM/ MeOH 10/1) to provide **2-6** as a pale yellow oil (4.2 g, 69.7%) ¹H-NMR (300 MHz, CDCl₃) δ 7.35-7.27 (5H, unresolved m), 7.15 (1H, s) 7.05 (1H, d, J = 8.22 Hz), 6.92 (1H, d, J = 8.04 Hz), 5.1 (2H, m, CH₂-benzyl), 4.62

(2H, s, CH₂-alcohol), 4.55 (1H, m, α H), 4.21 (4H, m, CH₂-phosphate) 3.14 (1H, unresolved m, βH), 2.98 (1H, unresolved m, βH), 1.33 (9H, s, Boc), 1.28 (6H, t, CH₃-phosphate); ¹³C-NMR (75.5 MHz, CDCl₃) δ 171.72, 155.31, 146.97, 135.35, 133.66, 130.56, 129.26, 128.53, 120.11, 79.63, 66.98, 64.85, 59.33, 54.70, 37.27, 28.23, 16.07; ³¹P-NMR (121.5 MHz, CDCl₃) δ = -5.1757(s)

Benzyl-2-(*tert*-butoxycarbonylamino)-3-(4-(diethoxyphosphoryloxy)-3-(fluoromethyl) phenyl) propanoate (2-7):

A solution of the alcohol **2-6** (3.23 g, 6 mmol) in dry CH₂Cl₂ (40 mL) was cooled to 0 °C and diethylaminosulfurtrifluoride (2.36 mL, 18 mmol, 3 eq.) was added and stirred for 30 min to complete the reaction. Saturated NaHCO₃ solution (40 mL) was added to quench the DAST and stirring continued for another 10 min. The reaction was then extracted with water, brine, dried over Na₂SO₄ and concentrated on a rotary evaporator. The yellow oily compound obtained was purified by flash column chromatography on silica gel (Hexane/ EtOAc 70/30) to obtain **2-7** as pale yellow oil (2.4 g, 74.1%). ¹H-NMR (300 MHz, CDCl₃) δ 7.33-7.27 (5H,unresolved m), 7.25 (1H,s) 7.03 (1H, d, *J* = 8.22 Hz), 7.05 (1H, d, *J* = 8.22 Hz), 5.46-5.30 (2H, d, *J* = 47.50 Hz, CH₂F), 5.12 (2H, m, CH₂-benzyl), 4.58 (1H, m, α H), 4.18 (4H, m, CH₂-phosphate), 3.1-2.97 (2H, unresolved m, βH), 1.39 (9H, s, Boc), 1.30 (6H, t, CH₃-phosphate); ¹³C-NMR (75.5 MHz, CDCl₃) δ 171.35, 154.99, 147.21, 135.11, 133.24, 130.76, 130.19, 128.23, 127.54, 119.70, 80.54, 78.32, 66.83, 64.62, 54.40, 36.96, 28.02, 15.87; ³¹P-NMR (121.5 MHz, CDCl₃) δ = -5.8829 (s); ¹⁹F-NMR (288.2 MHz, CDCl₃) δ = -138.41 (t, *J* = 47.45 Hz); ESI-MS: *m/z* = 562.2 [M+Na]⁺

2-(*tert*-butoxycarbonylamino)-3-(4-(diethoxyphosphoryloxy)-3-(fluoromethyl) phenyl) propanoic acid (2-8):

To a degassed solution of **2-7** (2 g, 3.7 mmol) in dry MeOH (40 mL) was added 10% Pd/C (0.2 g) and the solution was stirred under hydrogen atmosphere for 20 min to complete the reaction. The Pd/C was removed by filtration through a celite and the filtrate was concentrated under reduced pressure to afford **2-8** as pale yellow oil (1.6 g, 96%). It was used in the next reaction without further purification ¹H-NMR (300 MHz, CDCl₃) δ 7.32 (1H, m), 7.27 (1H, s) 7.19 (1H, d, *J* = 8.22 Hz), 5.51-5.35 (2H, d, *J* = 47.50 Hz, CH₂F), 4.52 (1H, m, α H), 4.22 (4H, m, CH₂-phosphate), 3.22-3.01 (2H, unresolved m, βH), 1.41 (9H, s, Boc), 1.34 (6H, t, CH₃-phosphate); ¹³C-NMR (75.5 MHz, CDCl₃) δ 173.55, 155.41, 147.31, 133.84, 131.20, 130.75, 130.66, 119.79, 80.82, 78.61, 65.23, 54.19, 37.16, 28.28, 15.97; ³¹P-NMR (121.5 MHz, CDCl₃) δ = -6.1999 (s); ¹⁹F-NMR (288.2 MHz, CDCl₃) δ = -138.25 (t, *J* = 47.41 Hz); ESI-MS: *m/z* = 448 [M-H]⁻

2-amino-3-(4-(diethoxyphosphoryloxy)-3-(fluoromethyl) phenyl) propanoic acid (2-9):

The N-Boc unnatural amino acid **2-8** (1.6 g, 3.56 mmol) was dissolved in a mixture of TFA/DCM (1/1; 40 mL each) and the solution was stirred for 3 h in an open RBF. TFA and DCM are evaporated off under reduced pressure to obtain **2-9** as pale yellow oil which was used in the next reaction without further purification. ¹H-NMR (300 MHz, CDCl₃) δ 12.90 (1H, COOH), 7.57 (2H, bs, NH₂), 7.37 (2H, m), 7.15 (1H, m), 5.47-5.31 (2H, d, *J* = 48 Hz, CH₂F), 5.27 (1H, m, α H), 4.25 (4H, t, CH₂-phosphate), 3.32-3.04 (2H, unresolved m, βH), 1.37 (6H, t, CH₃-phosphate); ³¹P-NMR (121.5 MHz, CDCl₃) δ = -7.2485 (s)

2-(((9H-fluoren-9-yl) methoxy) carbonyl amino)-3-(4-(diethoxyphosphoryloxy)-3-(fluoromethyl) phenyl) propanoic acid (2-10):

Compound **2-9** (1.6 g, 3.45 mmol) was dissolved in a mixture of THF/Water (1/1; 30 mL each) and the solution was adjusted to pH = 9 by adding NaHCO₃ (0.87 g, 10.35 mmol, 3.3 eq). To this was added FmocOSu (1.28 g, 3.8 mmol, 1.1 eq) and stirring continued for 4 h. THF from the reaction mixture was evaporated off and the solution was diluted with EtOAc. The aqueous layer was then acidified to pH = 2 by 1 N HCl and back extracted with EtOAc. The organic layer was then extracted with brine, dried over Na₂SO₄ and concentrated. Flash column chromatography (silica gel, 90/10 hexane/EtOAc to 97/3 EtOAc/MeOH with 0.1% acetic acid) afforded the N-Fmoc unnatural amino acid **2-10** as an off white foam (1.51 g, 76.5%). ¹H-NMR (300 MHz, CDCl₃) δ 10.38 (1H, s, COOH), 7.72 (2H, d, *J* = 7.38 Hz), 7.56 (2H, t, *J* = 7.08 Hz), 7.36 (1H, s), 7.31 (1H, s), 7.28 (4H, m), 7.16 (1H, d, *J* = 7.89 Hz), 6.93 (1H, m), 5.48-5.33 (2H, d, *J* = 47.00 Hz, CH₂F), 4.69 (1H, unresolved m, α H), 4.20 (4H, m, CH₂-phosphate), 3.23-3.06 (2H, unresolved m, βH), 1.31 (6H, t, CH₃-phosphate); ¹³C-NMR (75.5 MHz, CDCl₃) δ 173.11, 163.32, 155.77, 147.33, 143.69, 141.20, 133.39, 131.18, 130.53, 127.31, 127.00, 125.04, 119.61, 80.73, 78.53, 66.97, 65.04, 47.03, 36.77, 15.97; ³¹P-NMR (121.5 MHz, CDCl₃) δ = -6.2853 (s); ¹⁹F-NMR (288.2 MHz, CDCl₃) δ = -137.57 (t, *J* = 47.43 Hz); ESI-MS: *m/z* = 570.3 [M-H]⁻

Chapter 3: Caged Activity-Based Probes for Protein Tyrosine Phosphatases (PTPs)

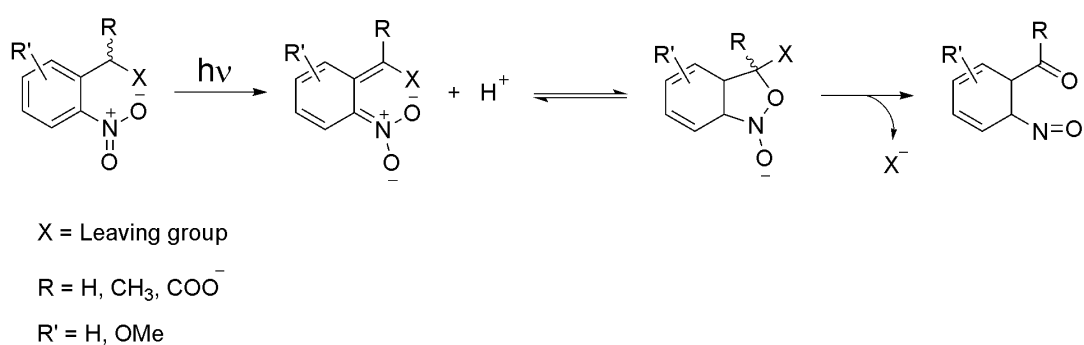
Summary:

This chapter summarizes the design and development of peptide-based light-controllable activity-based probes for PTPs. A novel caged unnatural amino acid (caged-2-FMPT) bearing photolabile *o*-nitrobenzyl group on the phosphate moiety of 2-fluoromethyl phosphotyrosine was synthesized and the unnatural amino acid was incorporated into peptides to generate peptide-based, caged, activity-based probes. With PTP1B as a model system, the concept of photo-uncaging followed by activity-based labeling was validated. The caged ABPs, with further improvements in design, could potentially be used for the spatially and temporally resolved investigation of PTP activities in cells.

3. 1 Introduction

The term, “caged compounds” in chemical biology contexts represents molecules in which a critical reactive functionality is masked with a chemical entity, which can be irreversibly unmasked using a suitable physical or chemical trigger such that it allows the spatially and temporally controlled release of a concentration burst of the reactive species, making possible a corresponding spatio-temporal resolution in the investigations of the downstream signalling events associated with the reactive functionality.⁶⁷ Photolabile groups are used extensively for this purpose mainly because of the precise focussing and manipulations of photons possible with the developments in laser and microscope technologies. Historically, the term caging was coined by J. F. Hoffman et al. in 1978 when the γ -phosphate of ATP was protected

with a photolabile *o*-nitrobenzyl group.⁶⁸ The caged ATP was found to be inert towards Na, K- ATPases while upon UV-irradiation the release of free ATP occurred and the photo-uncaging was characterized by the consumption of Na, K- ATPases and Na⁺ efflux from the cells through the Na:K pump. Indeed, a cyclic AMP derivative with a photolabile *o*-nitrobenzyl protecting group on the phosphate was reported by Engels et al. prior to Hoffman's report of the caged-ATP, although the use of photo-uncaging was not the focus of their paper.⁶⁹



Scheme 3. 1 Proposed mechanism for light-mediated uncaging of *o*-nitrobenzyl caged molecules.⁷⁴

The selection of a suitable caging group for a particular biological application is based on a number of criteria such as 1) synthetic considerations of the caged molecule 2) physical and chemical properties of the cage (the cage should not alter the fundamental properties of the system to be investigated) 3) photo-physical properties of the uncaging process such as the quantum yield of uncaging, the wavelength and molar extinction coefficient of excitation and the kinetics of uncaging 4) toxicity and chemical reactivity properties of photo-uncaging byproducts. The most extensively used caging groups in biological studies is the *o*-nitrobenzyl and its derivatives, as these protecting groups possess several desirable properties such as synthetic ease, good quantum yield of uncaging, inertness to cellular environment and water soluble

photo-uncaging byproducts. Several other molecules such as coumarin derivatives,⁷⁰ cinammic acid derivatives,⁷¹ hydroxyphenacyl derivatives,⁷² ketoprofen derivatives⁷³ etc. have also been used for caging applications in different biological experiments.

Caged versions of several biologically relevant molecules such as co-factors, neurotransmitters, secondary messengers, hormones, nucleic acids, lipids, peptides and even full-length proteins have been developed and they find great values in unravelling the roles of these important molecules in cellular functions which are otherwise difficult to achieve. Given the paramount importance of the reversible protein phosphorylation in cellular signal transduction events, significant research efforts have been dedicated to develop caged compounds useful in the study of protein kinases. In 1998, Lawrence and co-workers and Bayley and co-workers independently reported the first caged versions of c-AMP-dependent protein kinase (PKA).⁷⁵ While Bayley and co-workers employed a thiol reactive small molecule, 2-nitrobenzyl bromide (NBB), to covalently attach the photolabile group to a cysteine residue near the catalytic site of PKA, Lawrence and co-workers employed a peptide-based affinity label equipped with the thiol reactive 2-nitrobenzyl bromide at the C-terminal side to covalently label the cysteine residue (Cys-199) near the active site of the kinase. These caged versions of PKA remain catalytically inactive until exposure to UV light, while the photo-irradiation causes photo-uncaging and recovery of at least 50% of the kinase activity. In the same year Lawrence and co-workers reported a peptide-based caged inhibitor of PKA which was shown to inhibit the kinase activity only when uncaged with UV light.⁷⁶ In 2003, Lawrence and co-workers reported the first caged probe for intracellular PKC activity.⁷⁷ The probe design was based on a substrate peptide sequence of PKC which was equipped with an environmentally sensitive fluorophore at the N-terminal side and an adjacent serine residue (which acts

as the phosphorylation site for the kinase) was caged with 2-nitro-4, 5-dimethoxy benzyl group. Upon photo-uncaging, the serine residue gets receptive for phosphorylation and the phosphoryl transfer to which causes changes in the fluorescence of the environmentally sensitive fluorophore. Thus a simple measurement of fluorescence change following the photo-uncaging provides readout of the kinase activity. In 2003, Imperiali and co-workers designed a caged phosphopeptide derived from the autophosphorylation region of focal adhesion kinase (FAK).⁷⁸ Photo-uncaging of the phosphopeptide was used to cause temporally controlled lamellar arrest in a rat bladder tumour cell line. Nagamune and co-workers recently reported light-controlled activation of the lipid kinase, phosphatidylinositol-3-kinase (PIK), using a caged phosphopeptide.⁷⁹ Lawrence and co-workers recently reported a photodeactivatable bisubstrate-based inhibitor of Src tyrosine kinase.⁸⁰ The inhibitor was equipped with a photolabile site positioned at the β -Ala tether separating the catalytic (SH1) and SH2 domains targeting peptide sequences and upon photo-irradiation, the inhibitor gets split to its component peptides which no longer serves as an effective bisubstrate inhibitor of the kinases, thus providing a photodeactivatable version of a bisubstrate inhibitor of the kinase.

Although caging has been applied as a useful method for regulating the activity of various kinases, the strategy is relatively less explored in the study of protein phosphatases. A significant study involves the development of an imaging approach based on Förster Resonant Energy Transfer (FRET) between a GFP-tagged PTP1B and a fluorophore incorporated synthetic phosphopeptide to spatially resolve the enzyme-substrate interactions in live cells.⁸¹ By using a caged phosphopeptide substrate, Bastiaens and co-workers performed temporally controlled live-cell imaging of PTP1B-substrate interactions which revealed the establishment of a

steady-state enzyme-substrate gradient across the cell and spatially regulated PTP1B activity.

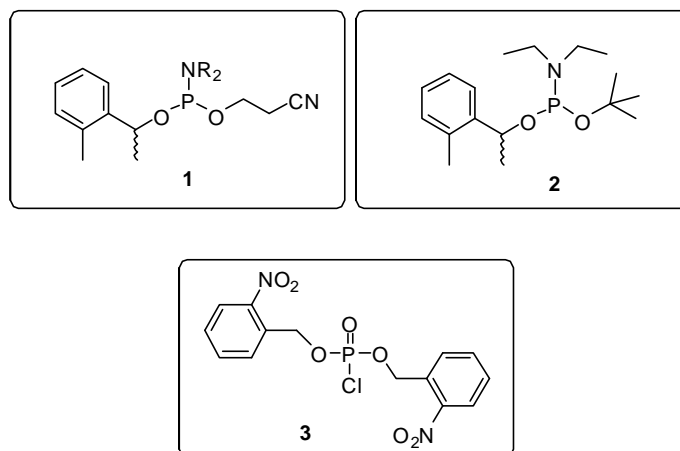


Figure 3. 1 Chemical structures of reagents used in the synthesis of caged phosphates

In line with the development of the peptide-based ABPs for PTPs, discussed in the previous chapter, we sought to develop caged versions of the probes which could be used for temporally controlled ABPP. Similar to the 2-FMPT based peptide probes, we decided to make a caged 2-FMPT, which could be incorporated into peptide synthesis to generate the caged ABPs. Imperiali and co-workers have previously reported the use of phosphoramidites such as **1** and **2** (Figure 3.1) to synthesize caged versions of phospho-serine, phospho-threonine and phospho-tyrosine amino acids and such phospho-caged amino acids have been shown to be compatible with standard solid-phase peptide synthesis protocols.⁸² However, attempts to make caged version of 2-FMPT using the phosphoramidite strategy were unsuccessful presumably because of the steric bulk near the reaction site of the unnatural tyrosine substrate. So we decided to make the chlorophosphate reagent **3** for generating the caged 2-FMPT, which was indeed proven to be a convenient reagent to accomplish the target molecule (Scheme 3.2). The caged 2-FMPT thus synthesized was subsequently incorporated into solid-phase peptide synthesis to generate peptide-based caged-ABPs.

The key idea of photo-uncaging followed by PTP-trapping is schematically shown in Figure 3.2. The presence of *o*-nitrobenzyl group on the phosphate protects the caged probe from reacting with PTPs while upon UV-irradiation, uncaging occurs generating the PTP-reactive 2-FMPT-based peptide probe. A nucleophilic attack of the catalytic cysteine residue from the active site of the PTP on the phosphate leads to dephosphorylation of the probe. Subsequent charge delocalization followed by elimination of F⁻ generates the reactive quinone methide intermediate which, upon attack by a nearby nucleophilic residue in the enzyme, would give rise to the covalent enzyme-probe complex.

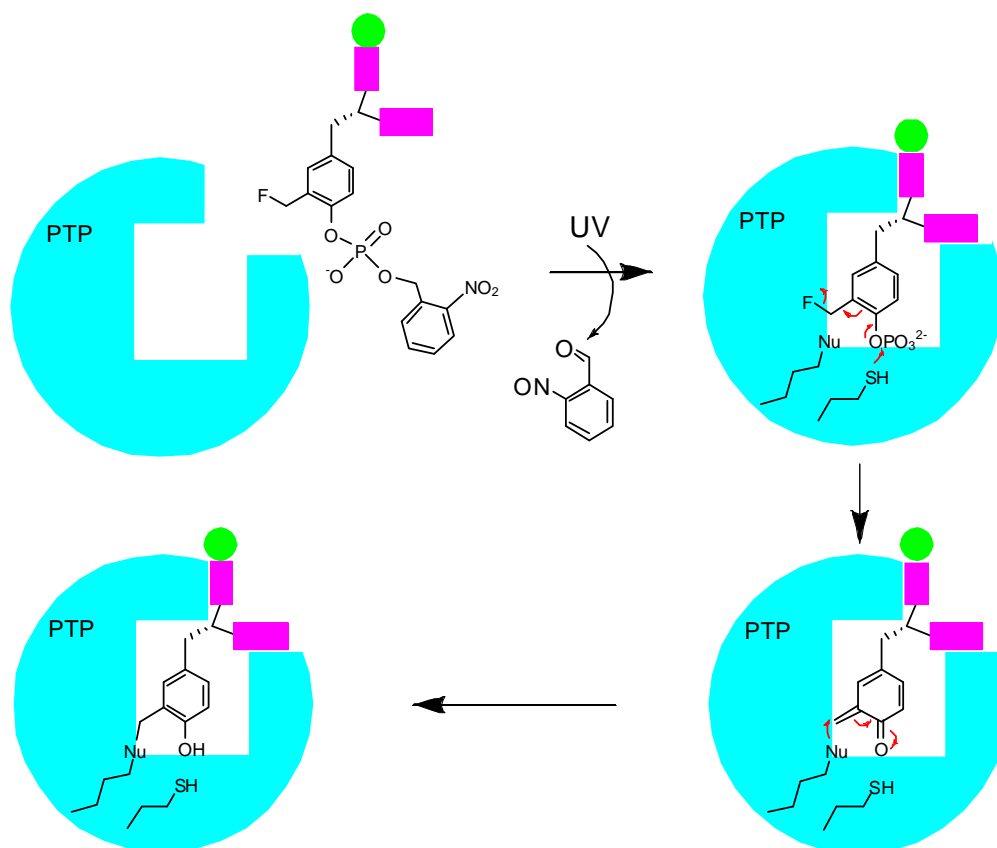
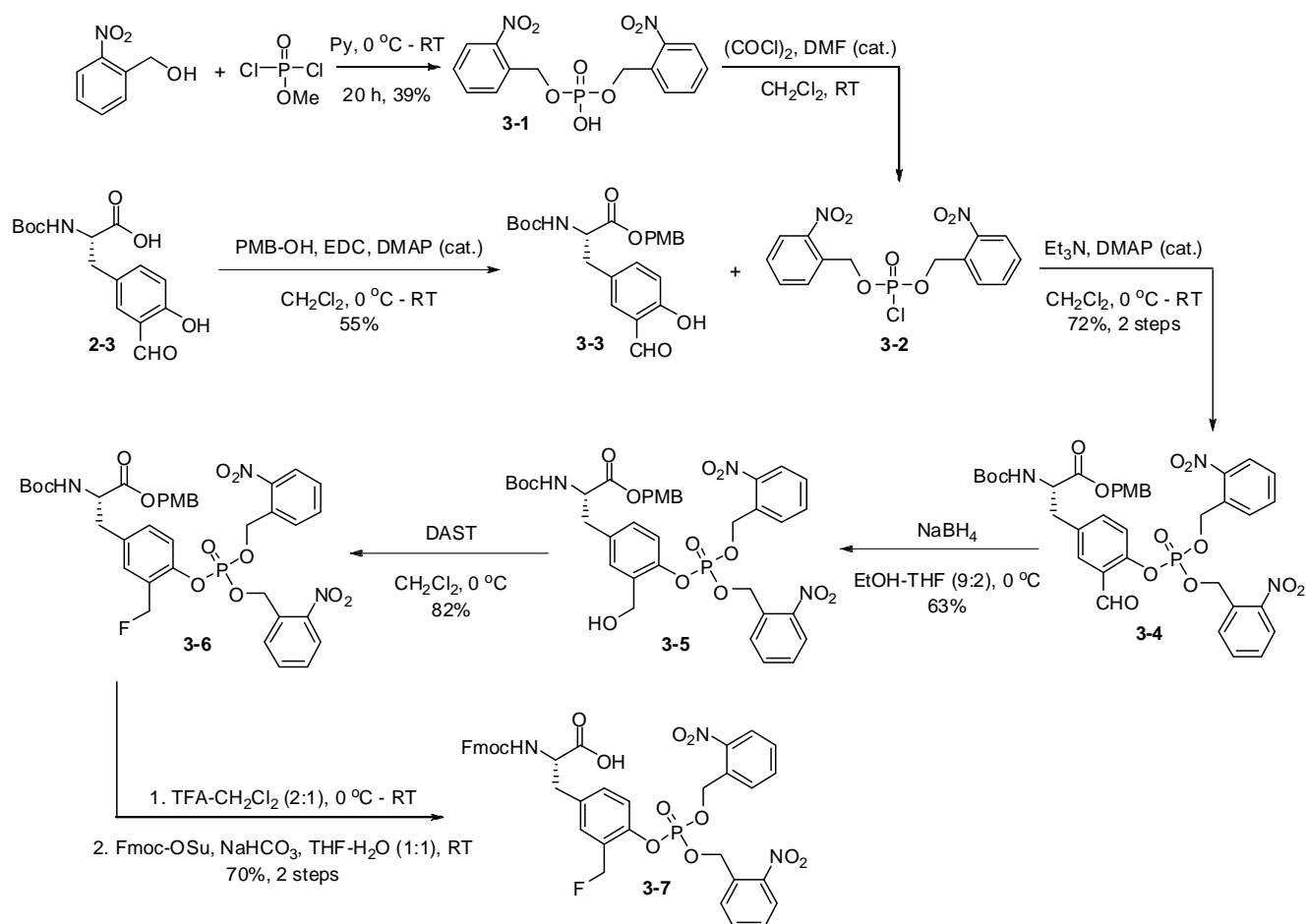


Figure 3. 2 Proposed mechanism of photo-uncaging of the peptide-based probe followed by labeling of PTP. The presence of *o*-nitrobenzyl group on the phosphate protects the caged probe from reacting with PTPs while upon UV-irradiation, uncaging of the phosphate occurs and the probe reacts with PTP forming covalent PTP-Probe complex.

3. 2 Synthesis of Caged 2-FMPT

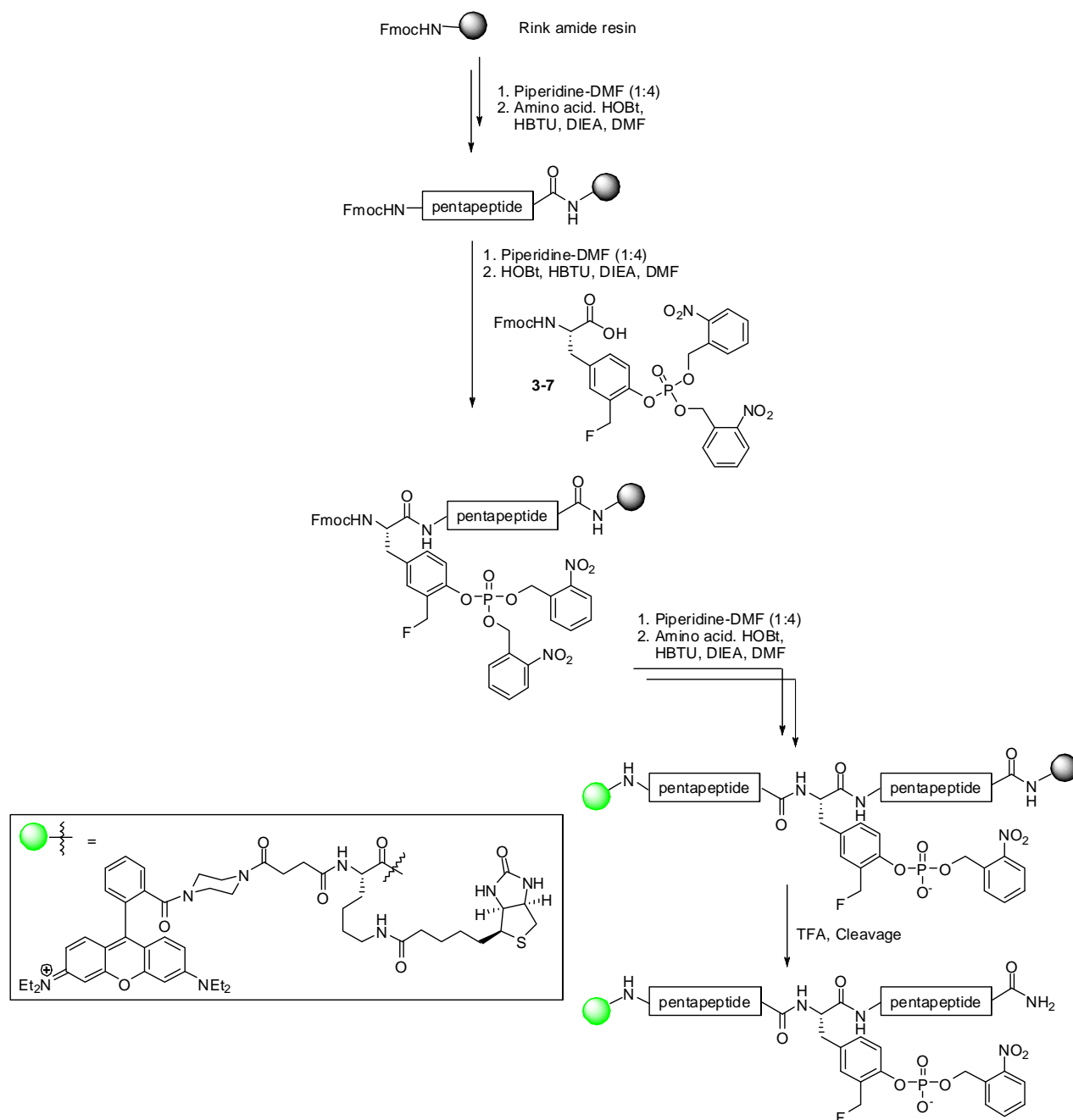


Scheme 3. 2 Synthesis of caged 2-FMPT

The synthesis of caged 2-FMPT was started with compound **2-3** which was an intermediate in the synthesis of 2-FMPT (the synthesis of 2-FMPT is discussed in chapter 2). The carboxylic acid group in **2-3** was protected with *p*-methoxybenzyl ester (PMB) to get **3-3** at 55% yield. In the next step the hydroxyl group of the unnatural tyrosine **3-3** was phosphorylated with the chlorophosphate reagent **3-2** to get the caged phosphorylated amino acid **3-4**. The chlorophosphate reagent **3-2** was

prepared from bis (2-nitrobenzyl) hydrogen phosphate **3-1** via treatment with oxalyl chloride and compound **3-1** was indeed prepared from commercially available 2-nitrobenzyl alcohol and methyl dichlorophosphate in pyridine at 39% yield according to published procedure.⁸³ The formyl group in **3-4** was subsequently reduced with NaBH₄ (63% yield) to get compound **3-5** which was then fluorinated with DAST (82% yield) afforded compound **3-6**. The PMB and Boc protective groups in compound **3-6** were simultaneously removed via treatment with TFA and the free amino group was then protected with Fmoc using FmocOSu afforded the final caged unnatural amino acid (caged-2-FMPT), compound **3-7**, with 14.3% yield over 7 steps.

3. 3 Synthesis of Caged Peptide-Based Probes



Scheme 3. 3 Synthesis of caged peptide-based ABPs

Solid-phase synthesis of peptide-based probes with caged 2-FMPT were carried out using rink amide resin following standard Fmoc-based strategy. Two caged

peptide-based probes were synthesized. The sequences of the probes (Table 3.1) were derived from putative PTP1B substrates reported in the Human Protein Reference Databank. The resin (in microreactors) was swelled in DMF for 30 min after which the Fmoc group was deprotected by treatment with 20% piperidine in DMF for 1 h followed by washing with DMF (3 × 10 min), DCM (3 × 10 min) and DMF again (3 × 10 min). The microreactors were then introduced into a reaction vessel containing Fmoc amino acid (4.0 eq., 2.0 eq. when **3-7** was used), pre-activated with HOBt (4 eq.), HBTU (4 eq.) and diisopropylethylamine (DIEA, 8 eq.) in DMF (2 mL). Coupling reactions were carried out overnight following which the microreactors were collected and washed with DMF (3 × 10 min), dichloromethane (3 × 10 min) and DMF again (3 × 10 min) followed by the next cycle of Fmoc deprotection and coupling. The N-terminus of each peptide was coupled to a Lys-(biotin) (4.0 eq.) unnatural amino acid followed by a rhodamine dye derivative (4.0 eq.) using the same pre-activation method as described before. Following the coupling of the rhodamine dye, the resin was washed extensively and dried thoroughly *in vacuo*. The peptides were cleaved from the resin for 3 h with a cocktail of TFA (95%), TIS (2.5%) and H₂O (2.5%). After cleavage, the solution was concentrated *in vacuo* until >80% of the cleavage cocktail was removed. To this was added cold ether (-20 °C) to precipitate the peptides and the mixture was left at -80 °C overnight to complete the precipitation. The peptides were then washed with cold ether (3 ×) and dried under reduced pressure. The probes were further purified by preparative HPLC and stock solutions were prepared in DMSO. Notably, it was found that one of the *o*-nitrobenzyl protecting groups on caged 2-FMPT was removed during the course of peptide synthesis, possibly during treatment of peptides with 20% piperidine in DMF.

Peptide Probe	Original Protein Source	Sequence ¹
c-P1	Stat5b	Rho-Lys(Biotin)KAVDG X VKPQI
c-P2	Tyk2	Rho-Lys(Biotin)PEGHE X YRVRE

¹ = Obtained from www.hprd.org X = caged 2-FMPT

Table 3. 1 Amino acid sequences in the two peptide-based caged PTP probes

3. 4 Results and Discussions

In order to evaluate the strategy of photo-uncaging followed by activity-based labeling, we performed time- and UV-irradiation dependant labeling reactions with pure PTP1B using the two caged probes c-P1 and c-P2. As shown in Figure 3.3, no labeling of PTP1B was detected during incubation of the probes with the enzyme in the absence of UV-irradiation whereas irradiation-time-dependant increase in labeling intensity was observed after exposure of the sample to UV-light.

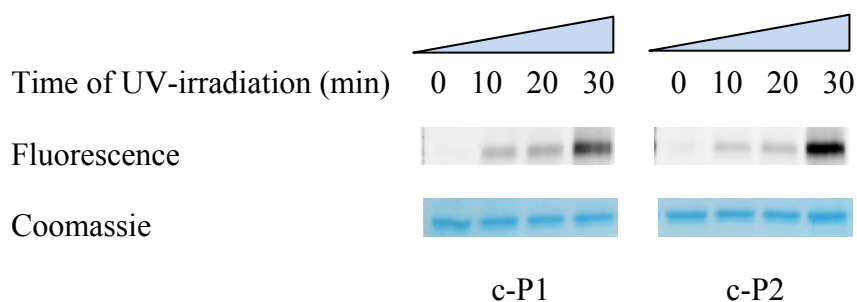


Figure 3. 3 UV-irradiation-dependant labeling of PTP1B (1 μ g, Tris buffer, pH = 7.5) with the probes c-P1 and c-P2 (4 μ M each in reaction). A handheld UV-light source (4W, \sim 360 nm light) was used for the irradiating the samples.

The photo-uncaging of the peptide probes (600 μ M solutions in DMSO) were monitored by HPLC, which showed complete uncaging of the probes in approximately 120 minutes of UV-irradiation (Figure 3.4).

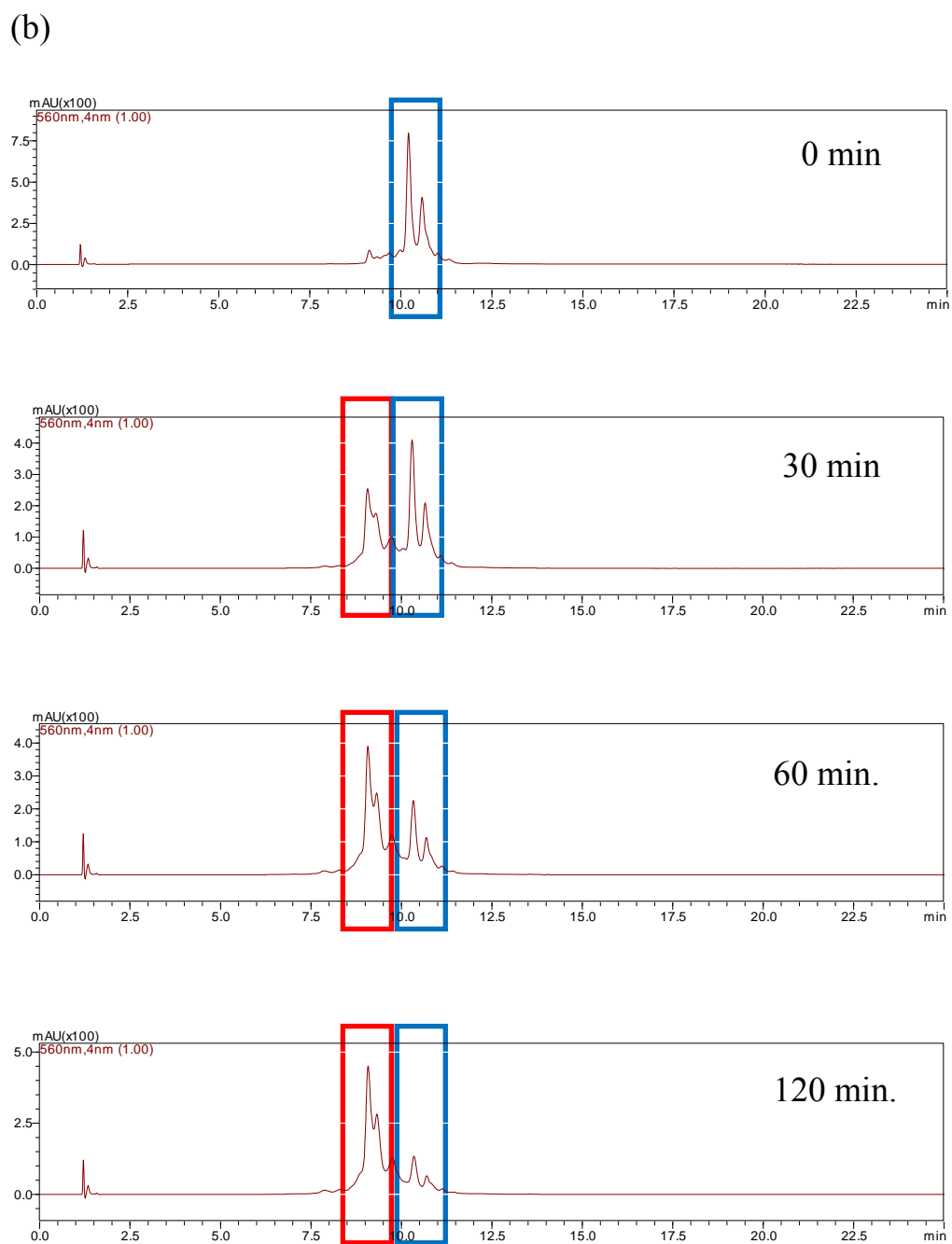
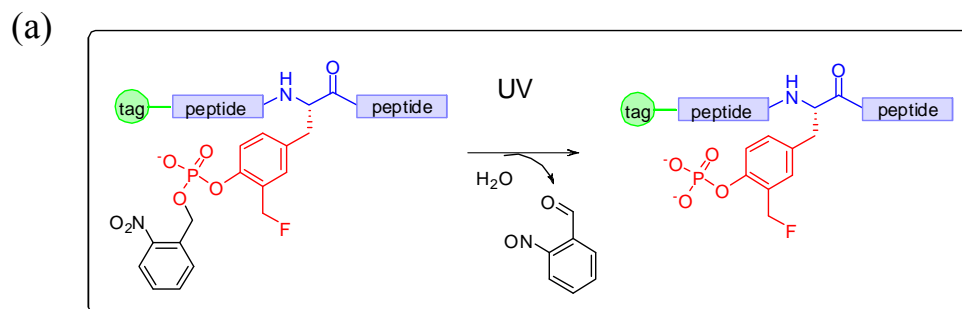


Figure 3. 4 (a) Scheme showing the photo-uncaging of the caged-peptide-based probe (b) HPLC profiles of the caged probe, c-P1, after increasing time (0 min, 30 min, 60 min and 120 min) of UV-irradiation. The peak highlighted in the blue box corresponds to the caged probe whereas the peak highlighted in the red box corresponds to the uncaged probe.

3. 5 Conclusions

In conclusion, a novel caged unnatural amino acid (caged 2-FMPT) was synthesized, which was successfully incorporated into solid-phase peptide synthesis to generate peptide-based caged activity-based probes. With PTP1B as a model system, the concept of photo-uncaging followed by activity-based labeling was validated *in vitro* and confirmed that light-mediated deprotection of the *o*-nitrobenzyl protecting group of the phosphate in the unnatural amino acid is essential for the activity-based labeling of the enzyme. Spatially and temporally resolved investigations of individual PTP activities in intact cells remains a formidable task within the current molecular and cellular biology research premises. Even a partial accomplishment of this extremely challenging task could be the key to unravel the critical roles that various PTPs play in the cellular signal transduction networks. It is hoped that the design of the caged-ABPs for PTPs discussed in this chapter could provide the guideline for the future development of useful chemical biology tools for addressing this highly challenging task.

3. 6 Synthetic details and chemical characterizations

Bis (2-nitrobenzyl) hydrogen phosphate (3-1):

Compound **3-1** was synthesized according to published procedures with some modifications.⁸³ To a solution of anhydrous pyridine (10 mL) at 0 °C was added dropwise methyl dichlorophosphate (1.0 g, 7.0 mmol). The mixture was kept at 0 °C for a further 15 min after which 2-nitrobenzyl alcohol (2.6 g, 17.5 mmol) was added. The ice bath was removed and the sealed reaction mixture was stirred overnight at room temperature. After 20 h, the mixture was poured into 10% aqueous NaHCO₃ (50 mL) and extracted with diethyl ether. The aqueous layer was acidified with 4N HCl to pH 1, filtered, washed with cold H₂O and dried *in vacuo* to give **3-1** as a white solid (1.01 g, 39%). ¹H-NMR (300 MHz, DMSO) δ 8.1 (d, 2H, *J* = 9.0 Hz), 7.8 (m, 4H), 7.6 (t, 2H, *J* = 7.4 Hz), 5.3 (d, 4H, *J* = 7.2 Hz). ¹³C-NMR (75 MHz, DMSO) δ 146.7, 134.2, 132.7, 132.6, 129.0, 128.5, 124.8, 64.5, 64.4. ³¹P-NMR (121 MHz, DMSO) δ -0.6. ESI-MS: *m/z* [(M-H⁺)/1] calculated for C₁₄H₁₂N₂O₈P: 367.0, found: 367.0.

4-Methoxybenzyl-2-(tert-butoxycarbonylamino)-3-(3-formyl-4-hydroxyphenyl)propanoate (3-3):

To a solution of 2-(tert-butoxycarbonylamino)-3-(3-formyl-4-hydroxyphenyl)propanoic acid, **2-3** (5.7 g, 18.4 mmol) in dichloromethane (50 mL) was added p-methoxybenzyl alcohol (3.4 mL, 27.7 mmol) and DMAP (0.3 g, 2.8 mmol). The solution was cooled to 0 °C followed by the addition of 1-ethyl-3-(3-dimethylaminopropyl)carbo-diimide hydrochloride (EDC-HCl, 4.2 g, 22.1 mmol). The reaction was stirred at 0 °C for 1 h after which the ice bath was removed and the

reaction was stirred for 3 h. After 3 h, the reaction was extracted with water, washed with brine and dried over anhyd. Na₂SO₄. Following the removal of the solvent, the crude product was purified by flash column chromatography (3:1 hexane:ethyl acetate) to give **3-3** as a white solid (4.33 g, 55%). ¹H-NMR (300 MHz, CDCl₃) δ. 10.9 (s, 1H), 9.6 (s, 1H), 7.2 (d, 2H, *J* = 8.7 Hz), 7.1 (m, 2H), 6.9 (d, 2H, *J* = 8.7 Hz), 6.8 (d, 1H, *J* = 8.4 Hz), 5.1 (m, 2H), 4.6 (m, 1H), 3.8 (s, 3H), 3.0 (m, 2H), 1.4 (s, 9H). ¹³C-NMR (75 MHz, CDCl₃) δ 196.4, 171.4, 160.6, 160.0, 155.0, 138.1, 134.2, 130.6, 127.5, 127.2, 120.4, 117.7, 114.0, 80.1, 67.1, 55.3, 54.4, 37.3, 28.3. ESI-MS: *m/z* [(M+Na⁺)/1] calculated for C₂₃H₂₇NNaO₇: 452.2, found: 452.1.

4-Methoxybenzyl-3-(4-(bis (benzyloxy) phosphoryloxy)-3-formylphenyl)-2-(tert-butoxy-carbonylamino) propanoate (3-4):

To a suspension of **3-1** (100 mg, 0.3 mmol) in distilled dichloromethane (3 mL) was added a catalytic amount of DMF (2 μL). Oxalyl chloride (172.5 mg, 1.4 mmol) was then added dropwise at room temperature to the reaction mixture. The reaction mixture was allowed to stir for 1 h after which the solvents were completely removed *in vacuo* to give **3-2** as a light yellow oil. The product was used immediately in the next step without further purification. Compound **3-2** was dissolved in distilled dichloromethane (2 mL) followed by the addition of DMAP (1.6 mg, 0.01 mmol) and **3-3** (58 mg, 0.14 mmol). The solution was cooled to 0 °C and triethylamine (30.2 mg, 0.3 mmol) was added dropwise. The ice bath was removed and the reaction mixture was stirred overnight. After 20 h, the reaction mixture was diluted and extracted with 0.25 N HCl. The organic layer was dried over anhyd. Na₂SO₄ and concentrated. The crude product was purified by flash column chromatography (1:1 hexane:ethyl acetate) to give **3-4** as an oil (76 mg, 72%). ¹H-NMR (300 MHz, CDCl₃) δ 10.2 (s, 1H), 8.1 (d,

2H, $J = 8.0$ Hz), 7.7 (m, 4H), 7.5 (m, 3H), 7.34 (d, 1H, $J = 8.4$ Hz), 7.26 (m, 3H), 6.9 (d, 2H, $J = 8.7$ Hz), 5.7 (d, 4H, $J = 7.9$ Hz), 5.1 (m, 2H), 4.6 (m, 1H), 3.8 (s, 3H), 3.1 (m, 2H), 1.4 (s, 9H). ^{13}C -NMR (75 MHz, CDCl_3) δ 187.8, 171.2, 159.9, 155.0, 150.8, 150.7, 146.8, 136.6, 134.3, 134.2, 131.5, 131.3, 130.6, 130.5, 129.4, 128.6, 127.1, 127.0, 126.9, 125.2, 121.0, 114.0, 80.1, 67.24, 67.18, 55.3, 54.3, 37.3, 28.2. ^{31}P -NMR (121 MHz, CDCl_3) δ -6.5. ESI-MS: m/z [(M+Na⁺)/1] calculated for $\text{C}_{37}\text{H}_{38}\text{N}_3\text{NaO}_{14}\text{P}$: 802.2, found: 802.2.

(S)-4-Methoxybenzyl-3-(4-(bis(2-nitrobenzyloxy)phosphoryloxy)-3-(hydroxymethyl) phenyl)-2-(tert-butoxycarbonylamino) propanoate (3-5):

Compound **3-4** (0.8 g, 1.0 mmol) was dissolved in a mixture of ethanol (18 mL), THF (4 mL) and water (180 μL) and cooled to 0 °C. NaBH_4 (38 mg, 1.0 mmol) was then added and the reaction mixture was stirred at 0 °C for 5 min. The reaction was quenched with 0.25 N HCl and stirred for another 15 min after which the organic solvents were removed *in vacuo*. The residue obtained was diluted with water and extracted ethyl acetate. The combined organic extracts were washed with brine and dried over anhyd. Na_2SO_4 and concentrated. The crude product was purified by flash column chromatography (1:1 hexane:ethyl acetate) to give **3-5** as a white solid (0.50 g, 63%). ^1H -NMR (300 MHz, CDCl_3) δ 8.1 (d, 2H, $J = 8.2$ Hz), 7.7 (m, 4H), 7.5 (m, 2H), 7.2 (m, 3H), 7.1 (d, 1H, $J = 8.4$ Hz), 6.9 (d, 1H, $J = 8.1$ Hz), 6.8 (d, 2H, $J = 8.6$ Hz), 5.6 (d, 4H, $J = 7.7$ Hz), 5.1 (m, 2H), 4.6 (m, 3H), 3.8 (s, 3H), 3.0 (m, 2H), 1.4 (s, 9H). ^{13}C -NMR (75 MHz, CDCl_3) δ 171.6, 159.8, 155.1, 146.7, 134.2, 134.1, 132.7, 132.6, 131.6, 131.5, 131.2, 130.4, 129.8, 129.2, 128.6, 127.3, 125.1, 120.2, 113.9, 79.9, 67.1, 67.0, 59.7, 55.2, 54.5, 37.4, 28.2. ^{31}P -NMR (121 MHz, CDCl_3) δ -5.3. ESI-MS: m/z [(M+Na⁺)/1] calculated for $\text{C}_{37}\text{H}_{40}\text{N}_3\text{NaO}_{14}\text{P}$: 804.2, found: 804.1.

(S)-4-Methoxybenzyl 3-(4-(bis(2-nitrobenzyloxy)phosphoryloxy)-3-(fluoromethyl)phenyl)-2-(tert-butoxycarbonylamino)propanoate (3-6):

To a solution of **3-5** (0.45 g, 0.6 mmol) in distilled dichloromethane (5 mL) at 0 °C was added diethylaminosulfurtrifluoride (DAST, 0.25 g, 1.5 mmol). The reaction mixture was stirred for 1 h at 0 °C and a further 20 min at room temperature. Following which the reaction was quenched with saturated NaHCO₃ and stirred at room temperature for an additional 15 min. The reaction mixture was partitioned between dichloromethane and water, the organic extracts were washed with brine and dried over anhyd. Na₂SO₄. Following concentration, the crude product was purified by flash column chromatography (3:2 hexane:ethyl acetate) to give **3-6** as an oil (0.37g, 82%). ¹H-NMR (300 MHz, CDCl₃) δ 8.1 (d, 2H, *J* = 8.6 Hz), 7.7 (m, 4H), 7.5 (m, 2H), 7.3 (m, 1H), 7.2 (d, 2H, *J* = 8.7 Hz), 7.1 (m, 1H), 7.0 (d, 1H, *J* = 8.2 Hz), 6.9 (d, 2H, *J* = 8.7 Hz), 5.6 (d, 4H, *J* = 7.7 Hz), 5.4 (d, 2H, *J* = 46.3 Hz), 5.0 (m, 2H), 4.6 (m, 1H), 3.8 (s, 3H), 3.0 (m, 2H), 1.4 (s, 9H). ¹³C-NMR (75 MHz, CDCl₃) δ 171.3, 159.9, 157.5, 155.0, 154.9, 147.3, 147.2, 146.7, 135.0, 134.2, 133.8, 132.2, 131.9, 131.8, 131.7, 131.6, 131.3, 130.9, 130.8, 130.5, 129.2, 129.1, 128.7, 128.5, 128.2, 127.2, 126.2, 125.1, 119.8, 114.0, 109.5, 80.7, 80.0, 78.5, 67.9, 67.1, 67.0, 66.9, 66.8, 55.5, 55.2, 54.8, 54.4, 37.8, 37.3, 28.4, 28.3. ¹⁹F-NMR (282 MHz, CDCl₃) δ -137.0 (t, *J* = 47.4 Hz). ³¹P-NMR (121 MHz, CDCl₃) δ -5.8, -6.4. ESI-MS: *m/z* [(M+Na⁺)/1] calculated for C₃₇H₃₉FN₃NaO₁₃P: 806.2, found: 806.1.

(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-(bis(2-nitrobenzyloxy)phosphoryloxy)-3-(fluoromethyl)phenyl)propanoic acid (3-7):

To a solution of **3-6** (30 mg, 0.04 mmol) in dichloromethane (1 mL) at 0 °C was added water (2 μL) and TFA (2 mL). The reaction was stirred for another 5 min at 0

°C following which the ice bath was removed and stirring was continued at room temperature. After 4 h, dichloromethane and TFA were removed completely *in vacuo*. The residue obtained was dissolved in THF (1 mL) and water (1 mL) following which solid NaHCO₃ was added to adjust the pH to 8. To this mixture was added N-(9-fluorenylmethoxycarbonyloxy) succinimide (FmocOSu, 14.2 mg, 0.04 mmol) and stirred overnight. After 21 h, THF was removed *in vacuo* and the residue was diluted with water and extracted with ethyl acetate. The aqueous layer was then acidified to pH 1 with 1 N HCl and extracted with ethyl acetate. The combined organic extracts were washed with brine, dried over anhyd. Na₂SO₄ and concentrated. The crude product was purified by flash column chromatography (2:3 hexane:ethyl acetate with 1 % acetic acid) to give **3-7** as an off-white solid (21 mg, 70%). ¹H-NMR (300 MHz, CDCl₃) δ 8.1 (d, 2H, *J* = 8.2 Hz), 7.7 (d, 2H, *J* = 7.4 Hz), 7.6 (m, 4H), 7.5 (m, 2H), 7.4 (m, 2H), 7.3 (t, 2H, *J* = 7.3 Hz), 7.2 (m, 4H), 7.1 (d, 1H, *J* = 7.9 Hz), 5.6 (d, 4H, *J* = 7.4 Hz), 5.3 (d, 2H, *J* = 47.4 Hz), 4.6 (m, 1H), 4.3 (m, 2H), 4.1 (t, 1H, *J* = 6.9 Hz), 3.1 (m, 2H). ¹³C-NMR (75 MHz, CDCl₃) δ 173.7, 156.0, 147.2, 147.2, 146.7, 143.7, 141.3, 134.2, 131.5, 131.4, 131.2, 131.1, 129.2, 128.6, 127.7, 127.1, 125.1, 120.0, 80.8, 78.6, 67.9, 67.2, 67.1, 55.4, 54.8, 54.7, 47.1, 37.1. ¹⁹F-NMR (282 MHz, CDCl₃) δ -136.2 (t, *J* = 46.3 Hz). ³¹P-NMR (121 MHz, CDCl₃) δ -6.6. HRMS (ESI): *m/z* [(M+Na⁺)/1] calculated for C₃₉H₃₃FN₃NaO₁₂P: 808.1678, found: 808.1679 (0.10 ppm).

Chapter 4: High-Throughput Synthesis of Abelson Tyrosine Kinase (Abl) Inhibitors using Click Chemistry

Summary:

This chapter summarizes the development of a synthetic strategy using the modular and efficient nature of the Cu (I) catalyzed Huisgen 1, 3-dipolar cycloaddition reaction between azides and terminal alkynes (or click chemistry) to rapidly assemble inhibitors of Abelson (Abl) tyrosine kinase. A generic kinase inhibitor library using the click assembly of an ADP-alkyne and a 344-member library of azides was first synthesized. Kinase inhibition assays with this generic inhibitor library revealed a preference for short chain linkers by the Abl kinase. Subsequently, a second generation of Abl inhibitor library was assembled using click chemistry with azides of short chain linkers and alkyne warheads derived from the Abl-selective drug Imatinib. Biochemical assays revealed a set of moderately potent and selective inhibitors, some of which have comparable potency as Imatinib against Abl. This work highlights the power of click chemistry to quickly generate inhibitor libraries of clinically relevant proteins.

4. 1 Introduction

Protein kinases catalyze the phosphorylation of serine, threonine and tyrosine residues in proteins. The reversible phosphorylation is one of the major modes of cellular signal transduction and it plays regulatory roles in most metabolic pathways. Hence, virtually every aspect of cellular physiology is inherently linked to the proper functioning of these enzymes. Aberrant regulation of kinase activity has been implicated in many diseases including several cancers. Thus new strategies in kinase

inhibitor design remain an active area of research with direct relevance to drug development. The majority of kinase inhibitors targets the ATP-binding site of the enzymes and, due to the highly conserved nature of the ATP-binding pocket amongst the various kinases, such inhibitors usually lack selectivity. An inhibitor that exploits binding pockets other than the ATP-binding site could potentially override this problem as such secondary binding pockets (e.g. substrate-binding or allosteric pockets) are less likely to be conserved. But there exist relatively few strategies for the rational design of such inhibitors. The most commonly used approach makes use of large libraries of synthetic compounds and natural products coupled with cell-based assays but so far, these methods, yielded only a limited number of hits. (Examples of this class of compounds include a substrate competitive inhibitor, ON012380,⁸⁴ of Bcr-Abl kinase and an allosteric inhibitor, GNF-2,⁸⁵ targeting the myristoyl binding pocket of the same kinase).

The recent development of certain kinase inhibitors with bidentate mode of inhibition⁸⁶ exploiting both the ATP-binding site and an allosteric site or a substrate-binding site opens up exciting opportunities for the rational design of bidentate kinase inhibitors. The strategy works by simply appending an inhibitor scaffold that targets the ATP-binding site of a kinase with another inhibitor scaffold that targets the allosteric/substrate-binding site, resulting in a bidentate inhibitor with improved selectivity and potency. Successful implementation of this method, however, largely depends on the available knowledge of secondary site-binding scaffolds on the target kinase, which is so far quite limited. Thus combinatorial approaches are necessary to identify such scaffolds and to generate new bidentate inhibitors.

We were particularly interested in Abl kinase as the aberrant activity of this kinase in a fusion protein form (Bcr-Abl) has been directly implicated in the

myeloproliferative disorder called Chronic Myelogenous Leukaemia (CML).⁸⁷ The cellular form of Abl is a close relative of the Src family of tyrosine kinases. It consists of approximately 1150 residues with two regulatory domains (SH3 and SH2) and a catalytic domain (SH1) with ~ 47% sequence homology with Src.⁸⁸ But unlike Src, the c-Abl kinase adopts a unique autoinhibitory mechanism involving the interaction of an N-terminal myristoyl group with a C-terminal myristoyl binding cleft.^{88,89} This autoinhibition is of paramount importance as in the absence of such a regulatory mechanism the kinase remains turned on always and the aberrant kinase activity leads to most of the pathological causes in the CML. The remarkable clinical success of the drug Imatinib (or ST1571, Gleevec™) and the recent cases of drug resistance due to mutations in the BCR-Abl gene have accelerated the search for new Abl inhibitors.⁹⁰ New development of Imatinib analogues and molecules possessing unique inhibitory mechanisms against Abl may offer promising approaches to tackle the drug resistance problem.^{90,91} We decided to develop a click chemistry-based strategy to generate potential bidentate inhibitors against Abl as the strategy had previously been shown to be highly modular and efficient for rapid assembly of bidentate inhibitors against many other enzymes.⁹²

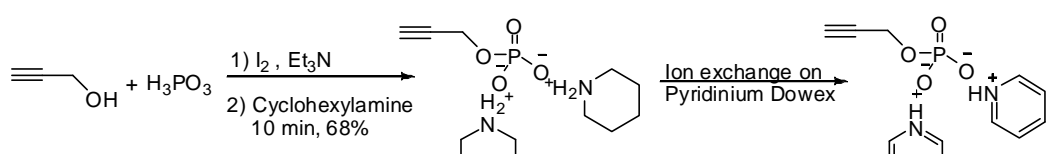
4. 2 Results and Discussions

4. 2. 1 The First-Generation of Kinase Click Inhibitors

At first, the possibility of assembling kinase inhibitors using click chemistry with ADP-alkyne (as a promiscuous ATP-surrogate) and a 344-member library of azides (Fig. 4.1) was investigated. The ADP-alkyne was synthesized (Scheme 4.1) using procedures similar to the synthesis of GDP-alkyne employed by Wong *et al.*⁹³ Briefly, propargyl alcohol was phosphorylated with crystalline phosphorous acid in presence

of iodine as oxidizing agent and the resultant propargyl phosphate **4-1** was isolated as a cyclohexylamine salt via precipitation from acetone containing cyclohexyl amine. The propargylphosphate-cyclohexylamine salt was exchanged for pyridinium ion form **4-2** using Dowex 50W-X2 resin in pyridinium form and subsequently coupled with AMP-morpholidate **4-3** using 1-H tetrazole as coupling reagent in anhydrous pyridine, affording the ADP-alkyne **4-4** in 21.5% yield over 4 steps. The 344-member library of azides was synthesized via a highly efficient solid-phase traceless protocol recently developed in the Yao's laboratory.^{41c} The click assembly of the two fragments was then carried out in a 384-well plate in a mixed solvent system containing *t*-BuOH, Water and DMSO (1:1:0.1) with CuI (0.1 eq.) as catalyst and tributylamine (0.2 eq.) as Cu (I) stabilizing base. The triazole formation was found to complete in 24 h. The crude reactions were evaporated to dryness and redissolved in DMSO to prepare standard solutions for kinase inhibition assays and LC-MS characterizations.

Upon “click” assembly, the 344-member library was screened for inhibition against Abl and Src kinases using a non-radioactive luminescence-based kinase assay; unique and distinctive inhibitor fingerprints were obtained for the library against the two kinases (Fig. 4.2). Detailed analysis of the inhibition profiles revealed a clear preference for short-chain linkers by Abl inhibitors whereas no such linker length preference was observed for Src inhibitors. The preference for short-chain azides by Abl kinase may point to the presence of secondary binding sites not too far from its ATP-binding site. This generic 344-member inhibitor library may be used as a platform for high-throughput identifications of preliminary bidentate inhibitors against other protein kinases.



Scheme 4. 1 Synthesis of the ADP-alkyne warhead

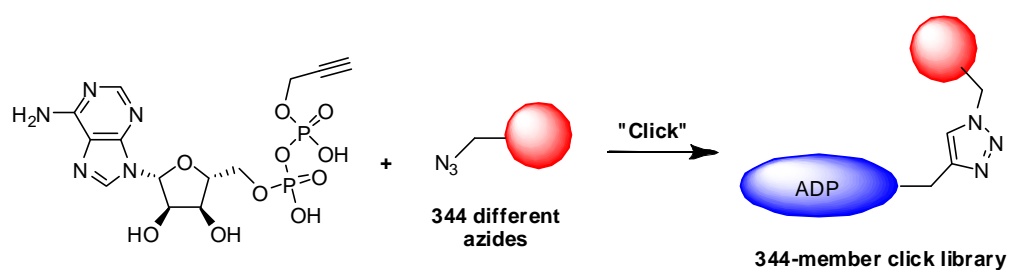


Figure 4. 1 Schematic representation of the click-assembly of the 344-member inhibitor library generated using the ADP-alkyne and azides

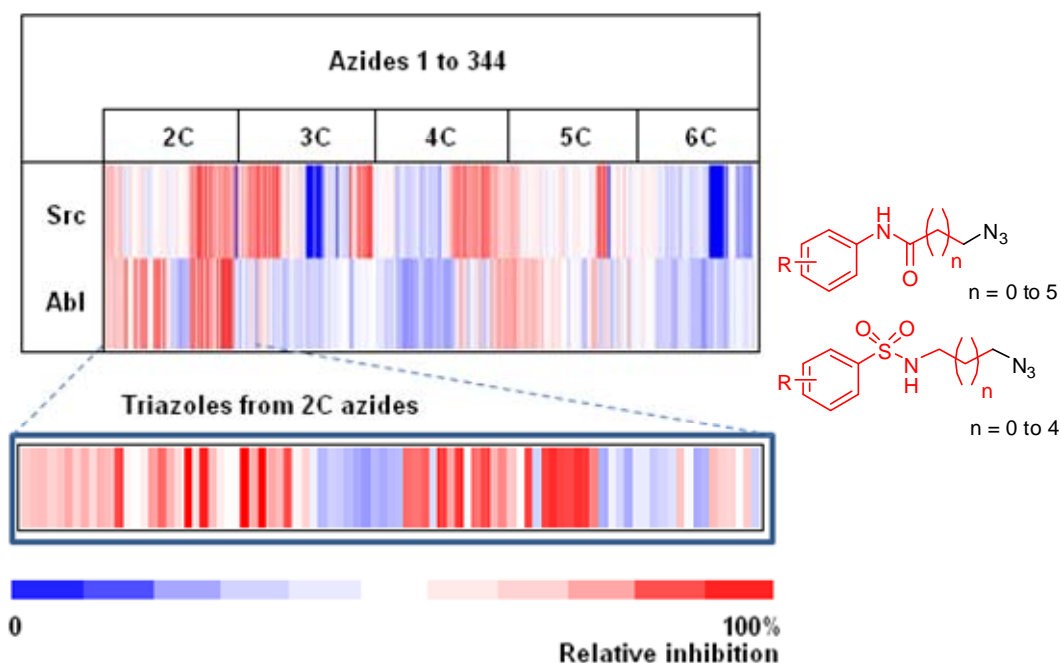


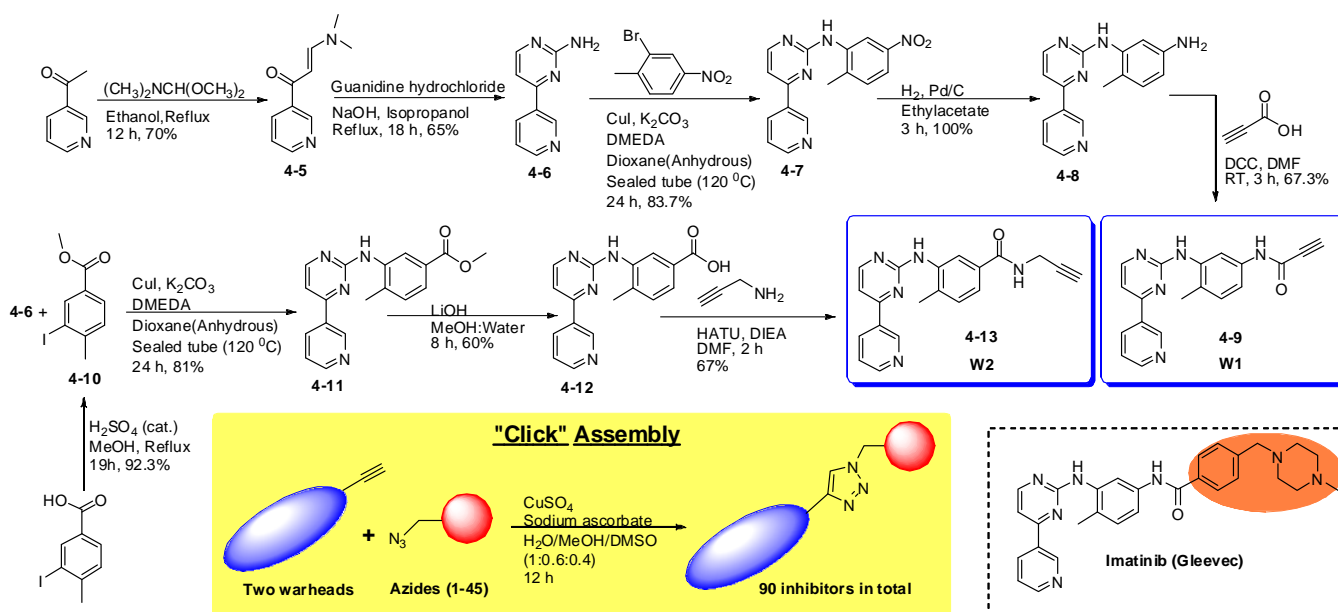
Figure 4. 2 The heat-map of the inhibition assay obtained with the 344-member inhibitor library against Src and Abl kinases. The preference for short-chain linkers by the Abl kinase is highlighted

4. 2. 2 The Second-Generation of Kinase Click Inhibitors

Having successfully established the platform for click-assembly of bidentate kinase inhibitors, the focus turned to Abl kinase. As shown in Scheme 4.2, the promiscuous, weak binding, ADP-alkyne warhead was replaced with the head portion of Imatinib, as this was reasoned to improve the selectivity as well as potency of the resulting inhibitors. Compared to the ADP-alkyne click library, the new click products should also be more ‘drug-like’. Short-chain (**2C**) azides were chosen based on the results obtained in Figure 4.2. Two clickable warheads (**W1** and **W2**) containing the Imatinib core structure were designed by replacing the piperazine-linked benzene ring (Scheme 4.2; shaded in orange) in the drug with an

alkyne group. Previously, results found that the piperazine moiety of Imatinib does not contribute significantly to the inhibition of Abl.⁹⁰ Thus it was conceived that, upon click chemistry, the warhead will guide the binding of the assembled inhibitor to the ATP-binding site and, at the same time, project the diversity element (from the azide) to the adjacent secondary binding site in the kinase (i.e. substrate-binding or allosteric site). The triazole ring formed should closely mimic the benzene ring in Imatinib, minimizing potential structural disruption of the inhibitor. **W1** was synthesized in 5 steps starting from acetylpyridine. The enamination reaction of acetylpyridine with N,N-dimethylformamide dimethylacetal in refluxing ethanol provided the enaminone **4-5**, which was converted to the pyrimidinyl amine **4-6** by reacting with guanidine hydrochloride in refluxing isopropanol in basic medium. Compound **4-6** was converted to **4-7** via a Cu(I)-catalyzed “Ullmann-Type” reaction as previously reported.⁹⁴ Subsequent reduction of **4-7** via catalytic hydrogenation gave **4-8**. Standard coupling of **4-8** with propiolic acid using DCC in DMF yielded **4-9 (W1)** in 25.6% overall yield (in 5 steps). **W2** was synthesized similarly from the key intermediate **4-11**. Methyl ester deprotection of **4-11** gave **4-12** which was then coupled with propargyl amine using HATU in DMF to afford **4-13 (W2)** in 13.6% yield over 6 steps.

Next, a 90-member inhibitor library based on Imatinib was assembled, by click chemistry, using the two warheads and 45 azides (**2C**) (Scheme 4.2). Briefly, in each reaction a warhead was mixed with an azide in a MeOH/H₂O/DMSO mixed solvent with catalytic amounts of CuSO₄ and sodium ascorbate. The click chemistry proceeded in quantitative yields in ~ 12 h at room temperature. The solvents were evaporated off and the products were redissolved in DMSO and characterized by LC-MS.



Scheme 4.2 Synthesis of the two warheads (W1 & W2) for Imatinib-based click library

4. 2. 3 Kinase Inhibition Assays

4. 2. 3. 1 Screening of the inhibitor library and generation of heat-map

A non-radioactive, luminescence based kinase inhibition assay (Kinase-Glo® Plus from Promega)⁹⁵ was employed to identify the inhibition potency of each compound in the library. The assay was based on the amount of ATP left after the kinase catalyzed phosphorylation of a substrate peptide (Figure 4.3). The ATP left after the kinase reaction is utilized for the monooxygenation of Luciferin catalyzed by the enzyme Luciferase and this reaction is accompanied with liberation of light. The Luciferase reaction produces one photon per turnover and the kinase activity is inversely proportional to the luminescence intensity.

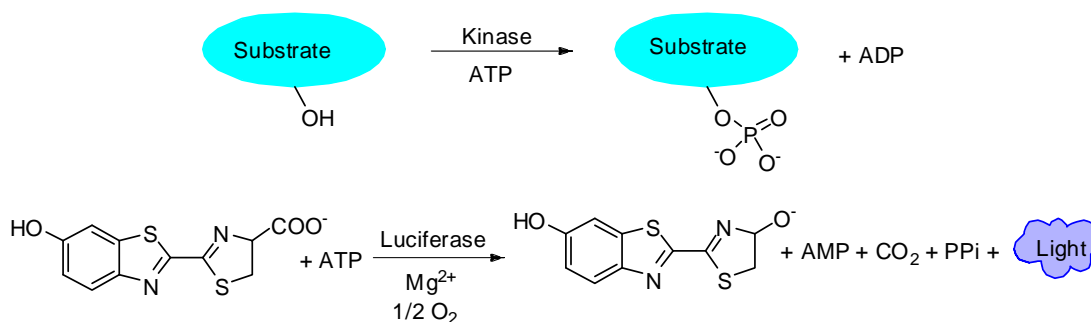


Figure 4.3 Luminescence-based kinase assay

The click-inhibitors were directly used for enzymatic screening against Src and Abl kinases. The results are shown in Figure 4.4. Detailed analysis of the inhibition profiles revealed that many inhibitors showed better inhibition against Abl over Src kinase. Eleven of these inhibitors were further identified (Figure 4.4; structures shown on right). They were scaled up, purified to homogeneity (by preparative HPLC), characterized (by LC-MS) before quantitative inhibition assays were carried out to obtain their IC₅₀ values against Abl kinase (Table 4.1).

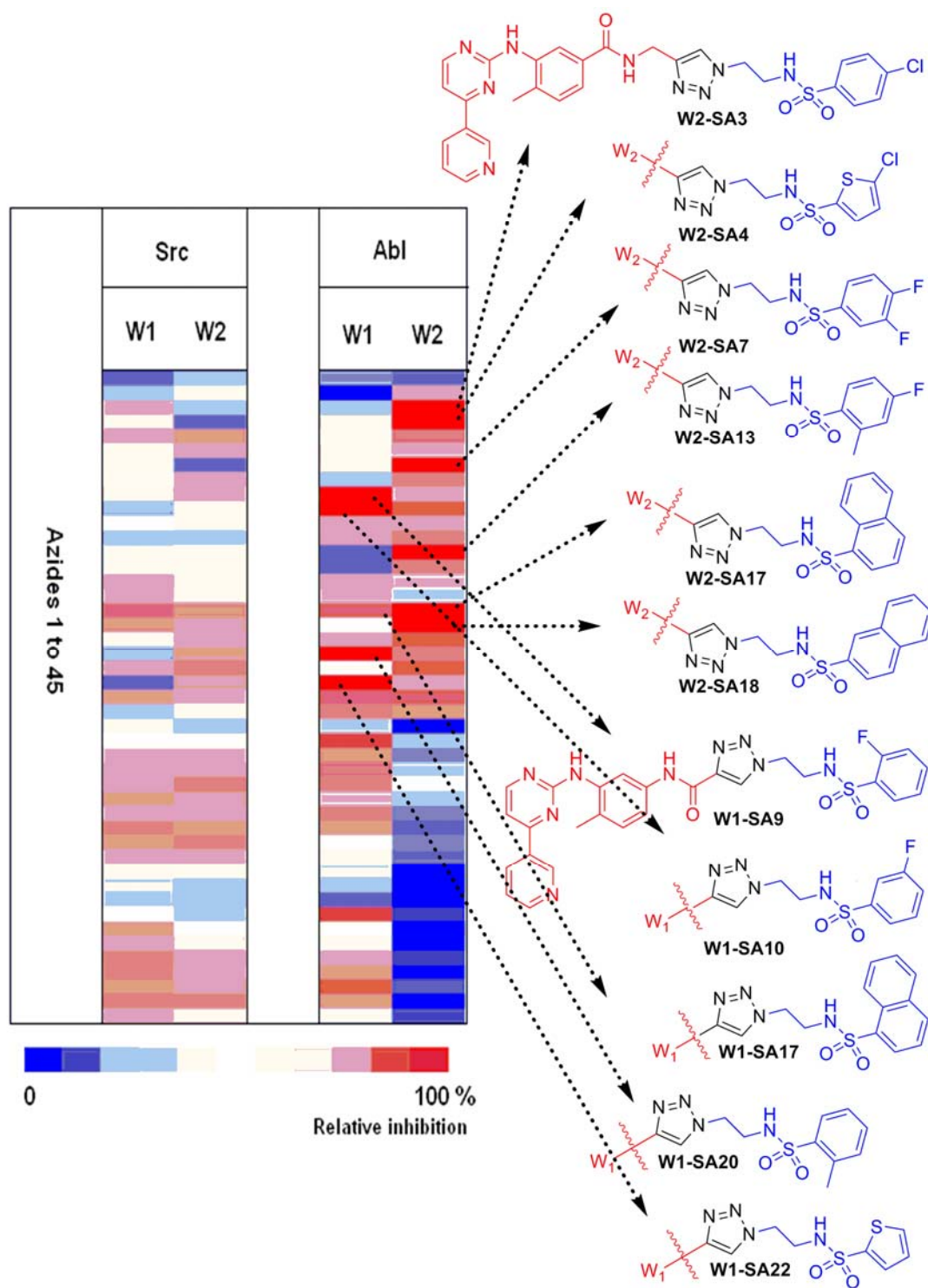
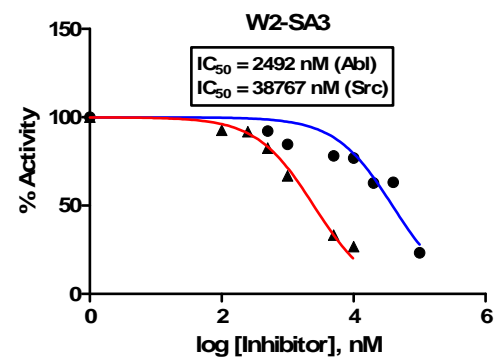
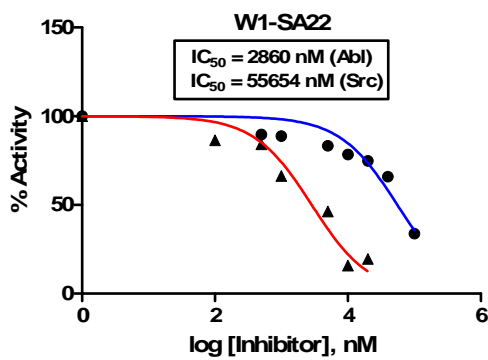
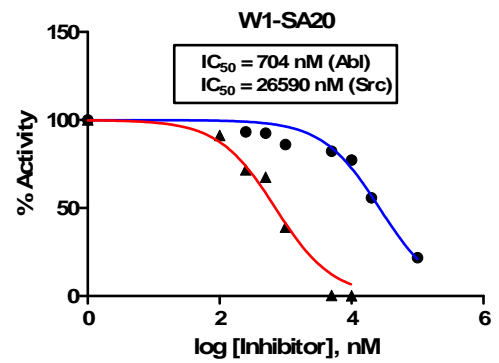
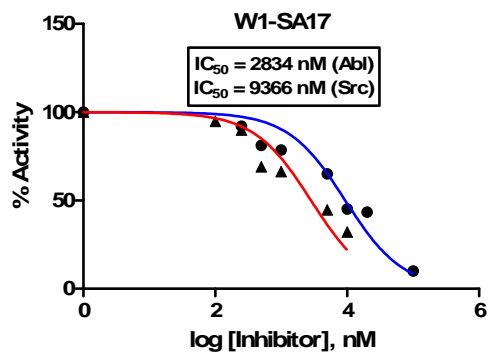
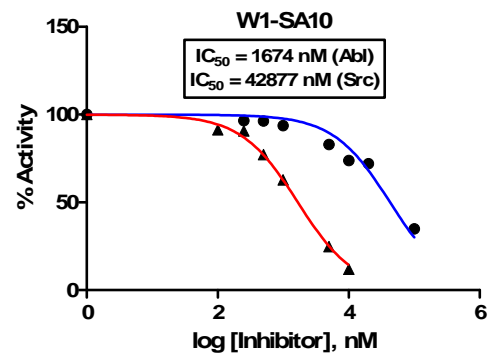
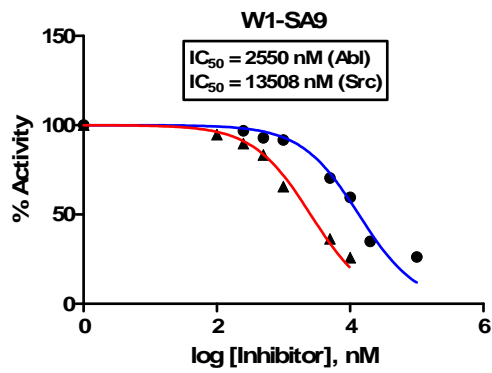


Figure. 4. 4 Heat-map showing the relative inhibition of the 90-member Imatinib-based bisubstrate inhibitor library against Src and Abl kinases. Structures of representative examples are shown on the right.

4. 2. 3. 2 IC_{50} evaluation of the click-inhibitors against Abl/Src kinases

Concentration-dependent experiments were performed to confirm the potency of the identified initial hits (structures shown above in Figure 4.4). All the 11 hits were purified by Semi-Prep reverse-phase HPLC using water with 0.1% TFA and acetonitrile with 0.1% TFA as eluents on a Luna 5 μ C-18 (2) 100A column (50 \times 30 mm, 5 micron) at a flow rate of 10 ml/min before subjecting to the detailed inhibition measurements. The dose-dependent inhibition assays were performed by varying the concentration of the inhibitors under fixed enzyme concentration of 50 nM. The IC_{50} values for each inhibitor were calculated from the percentage activity vs. log [concentration of inhibitor] curves generated using the GraphPad Prism software.

The IC_{50} results showed that most of the inhibitors exhibited 10-20 folds in selectivity for Abl over Src kinase (Figure 4.5). None of the azide components were found to exhibit any significant inhibition against either kinase. The warhead **W1** alone was more potent ($IC_{50} = 1.36 \mu\text{M}$) than **W2** ($IC_{50} = 27.3 \mu\text{M}$) against Abl kinase; this was presumably contributed by the proper positioning of the amide bond in **W1** which resulted in more favourable Hydrogen bonding interactions with residues located in the ATP-binding pocket of the kinase. The same phenomenon had previously been observed in the Imatinib-Abl complex as well.⁸⁹ **W1** was nearly as potent as Imatinib ($IC_{50} = 0.8 \mu\text{M}$) under identical assay conditions, indicating indeed the piperazine moiety in Imatinib was not essential for binding to Abl. Although the five inhibitors derived from **W1** showed only marginal improvement in potency over **W1** against Abl, with the best inhibitor, **W1-SA20**, having an IC_{50} of 0.70 μM (2-fold improvement over **W1**), the six **W2**-derived bisubstrate inhibitors were > 10-fold more potent than **W2** alone.



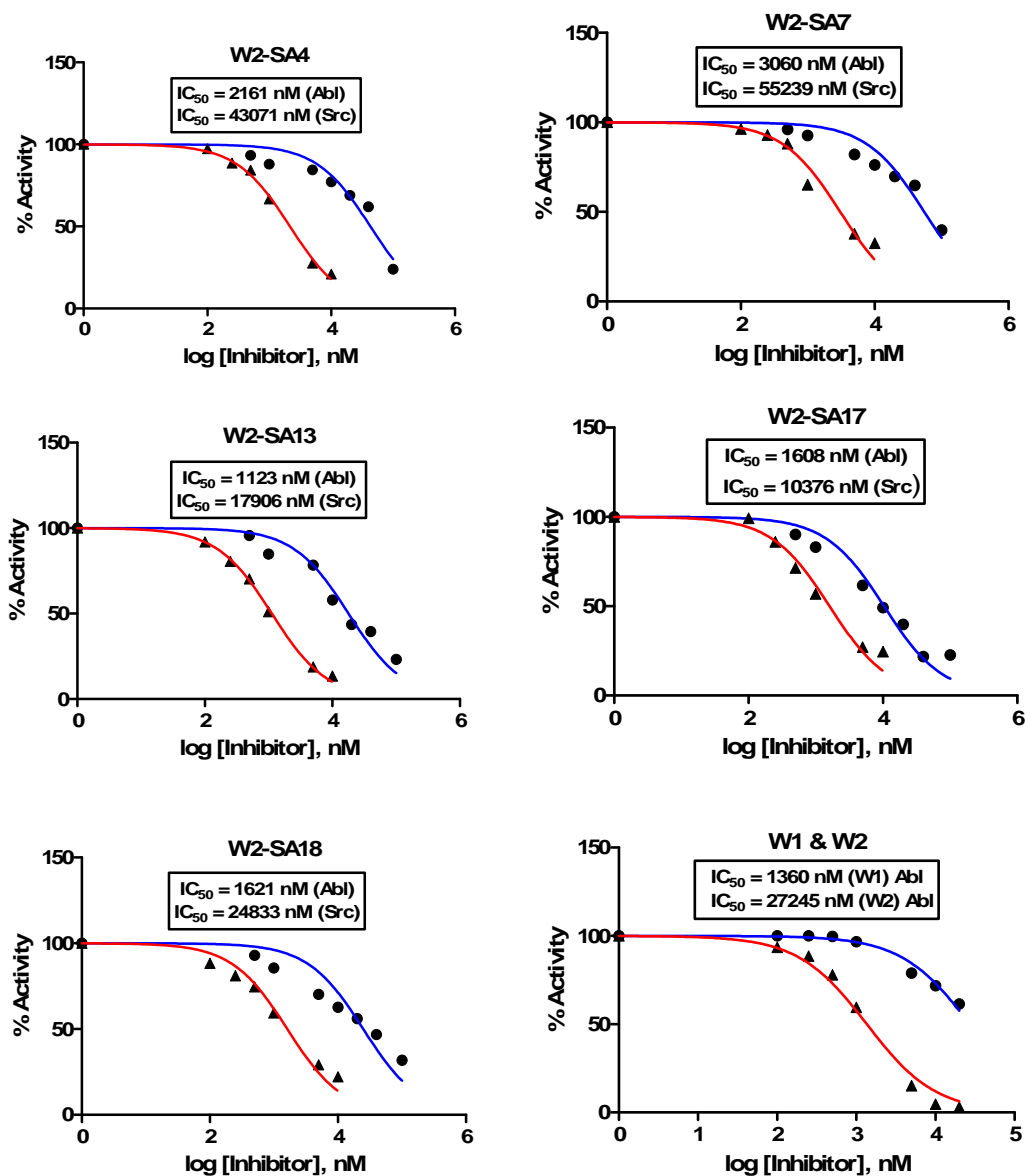


Figure 4. 5 IC_{50} evaluation of the selected click-inhibitors and the warheads against Abl/Src kinases

4. 2. 4 Cell culturing and anti-proliferation assay

Human leukemic cell lines K-562 (p-210 Bcr-Abl expressing chronic myelogenous leukemia) were treated with the 11 inhibitor hits identified against recombinant Abl from the in-vitro assay and the effect of the compounds on cell viability was tested with an anti-proliferation assay (XTT-assay).⁹⁶ The compounds, unfortunately, did not

show significant effect on the cell viability. Even the best inhibitor identified in the cell-based assay (**W2-SA17**, $IC_{50} = 1608$ nM in in-vitro assay) shows only 60% inhibition of the cell proliferation at 50 μ M concentration (All other compounds including the best inhibitor **W1-SA20** from the in-vitro assay were found to be less potent with less than 50% inhibition of cell proliferation at 50 μ M concentration. The exact reason for the lack of cellular potency of the compounds is not clear, although we speculate that it could most likely be due either to the poor solubility of the compounds in the assay medium or the cell-impermeability of the compounds.

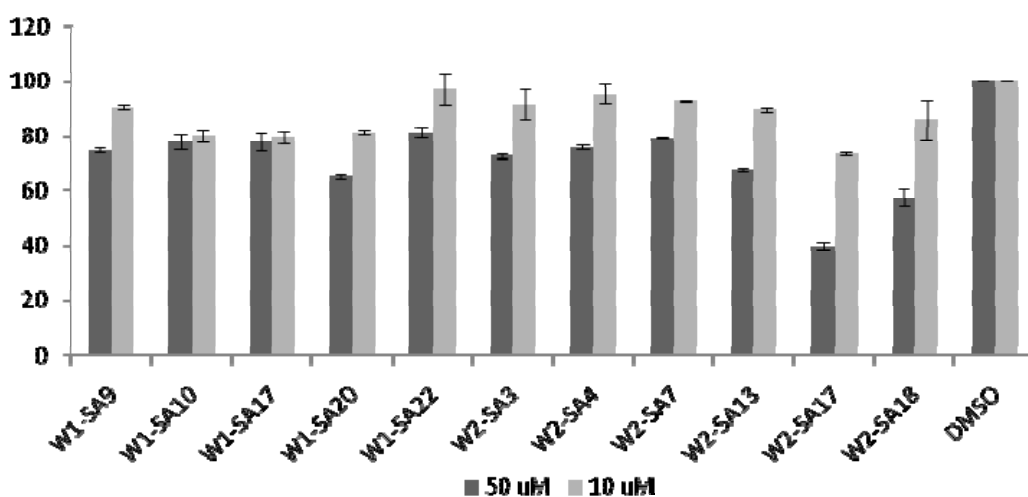


Figure 4. 6 Antiproliferation assay of K-562 cells in the presence of the 11 click-inhibitors at two different concentrations (50 μ M and 10 μ M). The Y axis represents the cell proliferation normalized with respect to the no inhibitor control (DMSO alone). Shorter column heights indicate stronger antiproliferative property.

Inhibitor	IC ₅₀ (in μ M)		Cell-based assay (% inhibition) ^a
	Abl	Src	
Imatinib	0.8 ^b		
W1	1.36	>100	-
W1-SA9	2.55	13.5	-
W1-SA10	1.67	42.9	-
W1-SA17	2.83	9.37	-
W1-SA20	0.70	26.6	35
W1-SA22	1.00	55.7	19
W2	27.3	>100	-
W2-SA3	2.49	38.8	-
W2-SA4	2.16	43.1	-
W2-SA7	3.06	55.2	-
W2-SA13	1.12	17.9	32
W2-SA17	1.61	10.4	60
W2-SA18	1.62	24.8	-

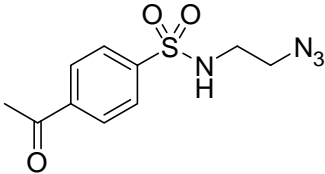
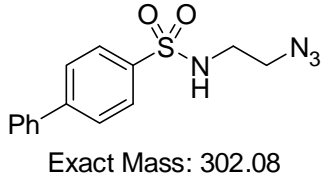
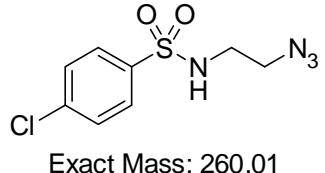
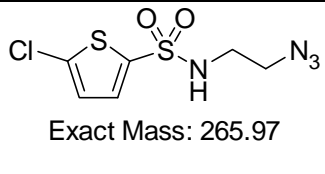
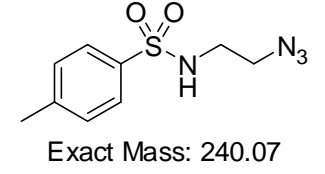
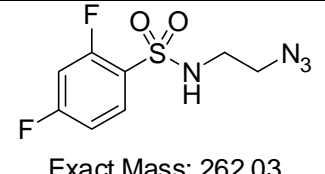
^aThe cell-based assay was done at 50 μ M concentration of inhibitors.

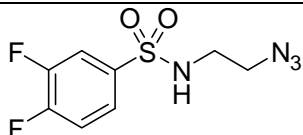
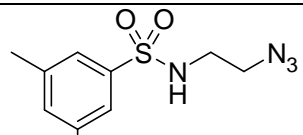
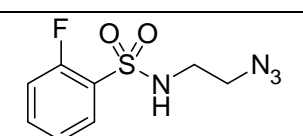
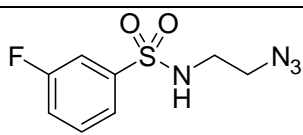
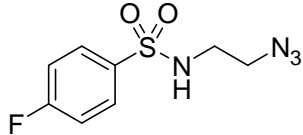
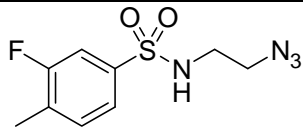
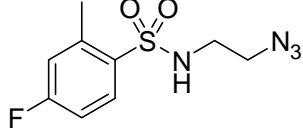
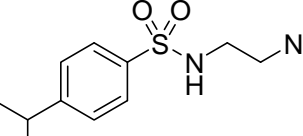
^bData obtained from reference 97. “-“ = Not determined.

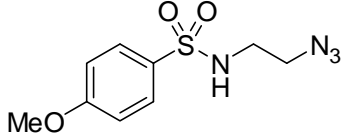
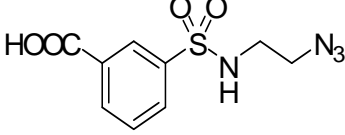
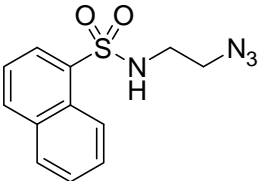
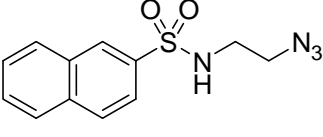
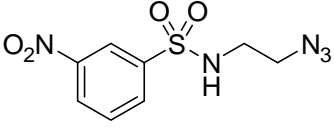
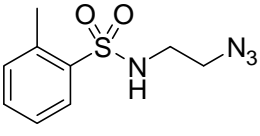
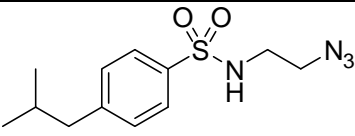
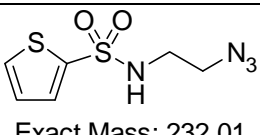
Table 4. 1 Inhibition data of selected compounds against Abl and Src kinases

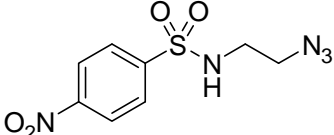
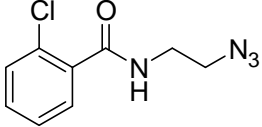
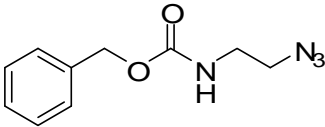
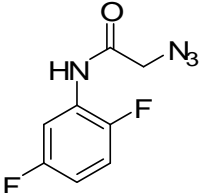
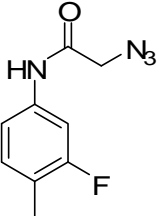
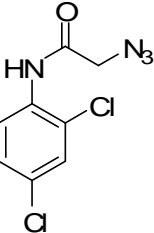
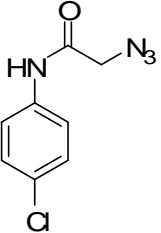
Table 4. 2 Azide library used for the synthesis of the Imatinib analogue click-library

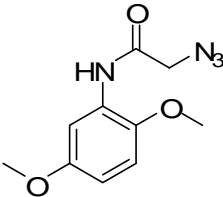
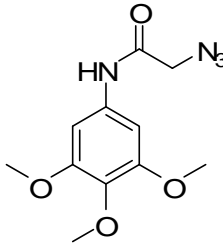
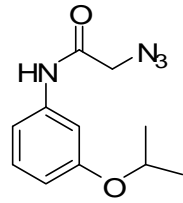
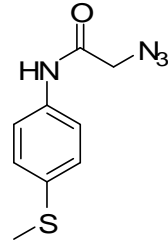
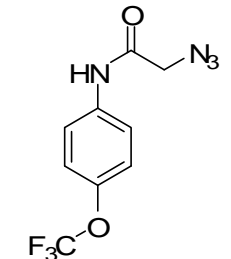
The azides were synthesized via a highly efficient traceless solid-phase method we recently developed.^{41c}

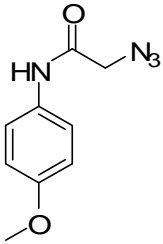
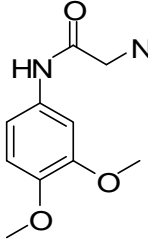
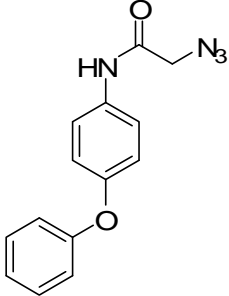
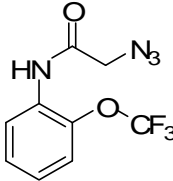
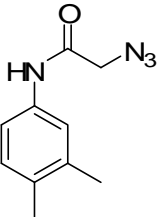
Entry	Compounds	Identity
1	 <p>Exact Mass: 268.06</p>	SA-1
2	 <p>Exact Mass: 302.08</p>	SA-2
3	 <p>Exact Mass: 260.01</p>	SA-3
4	 <p>Exact Mass: 265.97</p>	SA-4
5	 <p>Exact Mass: 240.07</p>	SA-5
6	 <p>Exact Mass: 262.03</p>	SA-6

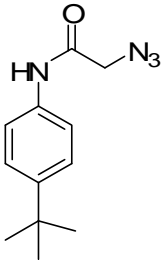
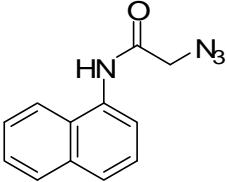
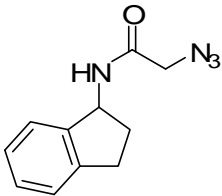
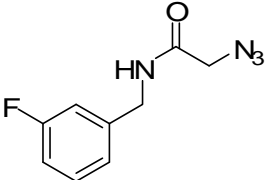
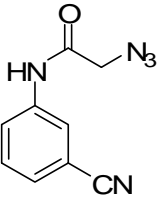
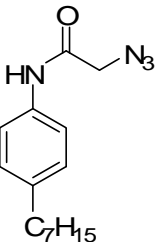
7	 <p>Exact Mass: 262.03</p>	SA-7
8	 <p>Exact Mass: 254.08</p>	SA-8
9	 <p>Exact Mass: 244.04</p>	SA-9
10	 <p>Exact Mass: 244.04</p>	SA-10
11	 <p>Exact Mass: 244.04</p>	SA-11
12	 <p>Exact Mass: 258.06</p>	SA-12
13	 <p>Exact Mass: 258.06</p>	SA-13
14	 <p>Exact Mass: 268.10</p>	SA-14

15	 <p>Exact Mass: 256.06</p>	SA-15
16	 <p>Exact mass: 270.04</p>	SA-16
17	 <p>Exact Mass: 276.07</p>	SA-17
18	 <p>Exact Mass: 276.07</p>	SA-18
19	 <p>Exact Mass: 271.04</p>	SA-19
20	 <p>Exact Mass: 240.07</p>	SA-20
21	 <p>Exact Mass: 282.12</p>	SA-21
22	 <p>Exact Mass: 232.01</p>	SA22

23	 <p>Exact Mass: 271.04</p>	SA-23
24	 <p>Exact Mass: 224.05</p>	SA-24
25	 <p>Exact mass: 220.09</p>	A-25
26	 <p>Exact mass: 212.05</p>	A-26
27	 <p>Exact mass: 208.07</p>	A-27
28	 <p>Exact mass: 243.99</p>	A-28
29	 <p>Exact mass: 210.03</p>	A-29

30	 <p>Exact mass: 236.09</p>	A-30
31	 <p>Exact mass: 266.10</p>	A-31
32	 <p>Exact mass: 234.11</p>	A-32
33	 <p>Exact mass: 222.05</p>	A-33
34	 <p>Exact mass: 260.05</p>	A-34

35	 <p>Exact mass: 206.08</p>	A-35
36	 <p>Exact mass: 236.09</p>	A-36
37	 <p>Exact mass: 268.09</p>	A-37
38	 <p>Exact mass: 260.05</p>	A-38
39	 <p>Exact mass: 204.10</p>	A-39

40	 <p>Exact mass: 232.13</p>	A-40
41	 <p>Exact mass: 226.08</p>	A-41
42	 <p>Exact mass: 216.10</p>	A-42
43	 <p>Exact mass: 208.07</p>	A-43
44	 <p>Exact mass: 201.06</p>	A-44
45	 <p>Exact mass: 274.17</p>	A-45

4.3 Conclusions

In conclusion, the first click-chemistry approach for rapid assembly of potential bidentate inhibitors of Abelson tyrosine kinase have been developed. Selected compounds identified from this “click-based” library, though showed only comparable potency as that of Imatinib in their inhibition of Abl kinase, they nevertheless offer the first glimpse of the generality and utility of the click strategy in future development of bidentate inhibitors against a variety of other protein kinases. The click-products are more ‘drug-like’ compared to the existing peptide based bisubstrate kinase inhibitors and have lot more scope to further optimize their pharmacological properties. Therefore, the strategy should provide a useful tool in the fields of kinase research and drug discovery, as well as in the emerging field of Catalomics.

4. 4 General experimental procedures

4. 4. 1 *The click-assembly of inhibitors*

4. 4. 1. 1 *General procedure for the click-assembly of 344-member library formed from ADP-alkyne and azides*

The click-assembly of the two fragments (ADP-alkyne and azides) was carried out in a 384-well plate in a mixed solvent system containing *t*-BuOH, Water and DMSO (1:1:0.1) with CuI (0.1 eq.) as catalyst and tributylamine (0.2 eq.) as Cu(I) stabilizing base. Other commonly employed click-catalytic systems such as CuSO₄-Sodium ascorbate and Cu-CuSO₄ were not successful. The triazole formation was found to complete in 24 h. The crude reactions were evaporated to dryness and redissolved in DMSO to prepare standard solutions for kinase inhibition assays and LC-MS characterizations.

4. 4. 1.2 *General procedure for the click-assembly of 90-member*

Imatinib analogue library formed from the two warheads (W1 & W2) and azides

The following optimized condition was used for the click-assembly of the inhibitors:

The alkyne warhead (**W1** or **W2**) = 5 mM in DMSO, 20 μ L,

The Azide = 5 to 7.5 mM in DMSO, 20 μ L,

CuSO₄ = 0.1 mM, Sodium ascorbate = 1 mM,

Reaction volume = 200 μ L (100 μ L water + 40 μ L DMSO + 60 μ L Methanol)

Briefly, the alkyne warhead and the azide were mixed in the above mentioned solvent system in the presence of indicated amounts of CuSO₄ and sodium ascorbate. The reactions were shaken for 12 h at room temperature upon which the entire warhead was converted to the triazole (monitored by LC-MS). The reactions were concentrated and redissolved in DMSO to prepare standard solutions for enzyme inhibition assays.

4. 4. 2 General procedure for Kinase inhibition assays

The kinase inhibition assay was performed with Kinase-Glo[®] Plus Luminescent Kinase assay kit from Promega following the manufactures instructions. Briefly, the kinase, the peptide substrate, ATP and the inhibitor were mixed in the kinase reaction buffer (100 mM Tris, pH=7.5, 10 mM MgCl₂) at a volume of 25 µL in a flat-bottom solid white 384-well plate (or 55 µL in a flat-bottom solid white 96-well plate). The incubation was allowed to continue for 30 min at 37 °C and the kinase reaction was then 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 5 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.

The inhibition assays with Abl and Src kinases were performed at nearly equal activities of the two kinases at a concentration of ~ 50 nM for Abl and ~ 90 nM for Src kinases respectively. The substrate peptides used are shown below.

Abl substrate peptide: **KKGEAIYAAPFA-NH₂**

Src substrate peptide: **KVEKIGEGTYGVVYK-NH₂**

The inhibitors were screened against the closely related kinases Abl and Src in a 384-well plate or 96-well plate as described above. The luminescence intensity from each well was measured and the inhibition potency was calculated using the following relation.

Potency = 1 - Remaining activity

The average value of potency across the entire library was calculated and the heat-maps were generated by assigning the standard deviation values to each compound.

4. 4. 3 General procedure for cell-culturing and anti-proliferation assay

Human leukemic cell lines K-562 (p-210 Bcr-Abl expressing chronic myelogenous leukemia) were maintained in RPMI-1640 medium supplemented with L-glutamine and 10% FBS. The cells ($0.2-0.4 \times 10^6$ per mL) were plated in duplicate in 96-well plates containing two different concentrations (50 and 10 μ M) of the 11 inhibitor hits identified against recombinant Abl from the in-vitro assay. After incubation at 37 °C in 5% CO₂ for 48 h, the effect of the compounds on the cell proliferation was determined by the XTT (sodium 3' - [1-(phenylaminocarbonyl)-3, 4-tetrazolium]-bis (4-methoxy-6-nitro) benzene sulfonic acid hydrate) colorimetric

dye reduction method. Briefly, the procedure for the assay was as follows. The XTT was dissolved in hot RPMI media (37 °C) at a concentration of 1 mg/mL. Immediately before use, the electron coupling reagent PMS (N-methyl dibenzopyrazine methyl sulfate) was added to the XTT solution giving a PMS concentration of 125 μM (The PMS was stored as 100 mM stock solution in saline at 4 °C). 25 μL of this XTT/PMS solution was added to each well in the 96-well plate containing ~ 100 μL per well culture giving a final concentration of 0.2 mg/mL XTT and 25 μM PMS. After incubation at 37 °C in 5% CO₂ for 2-6 h the absorbance of each well was measured at a wavelength of 450 nm using Tecan microplate reader and effect of the compounds on the cell proliferation was determined by comparing with the DMSO (no inhibitor) controls.

4. 5 Synthetic details and characterizations of compounds

Prop-2-ynyl phosphate cyclohexylamine salt (4-1):

Compound **4-1** was synthesized as previously reported.⁹³ Briefly, to a solution of propargyl alcohol (30 mL, 516 mmol) and triethylamine (7.5 mL, 54 mmol) was added phosphorous acid (1.23 g, 15 mmol). To this was added 5.7 g of iodine (22.47 mmol) in small portions over 5 min with stirring. The stirring was continued for 10 more minutes and then the reaction mixture was pored into 400 mL of acetone containing 15 mL cyclohexylamine. A white precipitate was formed immediately and the solution was kept for 2 h, then the precipitate was filtered, recrystallized from 95% ethanol containing few drops of cyclohexylamine. The precipitate obtained was dried to give 3.11 g (yield = 68%) of the propargyl cyclohexylamine salt **4-1** as a white solid. ¹H-NMR (300 MHz, D₂O) δ ppm = 1.1 (2H, m), 1.25 (8H, m), 1.61 (2H, m), 1.75 (4H, m), 1.93 (4H, m), 2.75 (1H, t, J =

2.46 Hz), 3.09 (2H, m), 4.34 (2H, dd, $J = 7.07$ Hz, 2.46 Hz). ^{31}P -NMR (121.5 MHz, D_2O) δ ppm = 4.17.

Prop-2-ynyl phosphate pyridinium salt (4-2):

The propargyl cyclohexylamine salt (0.83 g, 2.71 mmol) was dissolved in water (1.5 mL) and converted to the pyridinium ion form via ion exchange on a pyridinium form of Dowex 50W-X2 resin. The pyridinium form of Dowex resin was prepared from the commercially available H^+ form of the resin by repeated elution of a pre-washed (with water) column of the resin with 10% pyridine. The product **4-2** was eluted with 30 mL of water in nearly quantitative yield and concentrated on a rotary evaporator. ^1H -NMR (300 MHz, D_2O) δ ppm = 2.78 (1H, t, $J = 2.46$ Hz), 4.41 (2H, dd, $J = 7.07$ Hz, 2.46 Hz), 7.95 (t, $J = 7.23$ Hz), 8.50 (m), 8.65 (d, $J = 5.28$ Hz). ^{13}C -NMR (75 MHz, D_2O) δ ppm = 53.50, 75.72, 78.96, 127.29, 140.90, 147.07. ^{31}P -NMR (121.5 MHz, D_2O) δ ppm = 0.48.

AMP-morpholidate (4-3):

The AMP-morpholidate was synthesized as previously reported with minor modifications.⁹⁸ Briefly, adenosine monophosphate monohydrate (0.84 g, 2.3 mmol) was dissolved in a mixture of water (23 mL), *t*-butanol (23 mL) and distilled morpholine (0.78 mL, 11.4 mmol, 5 eq.). The mixture was heated under reflux and DCC (1.9 g, 9.2 mmol, 4 eq.) in *t*-butanol (35 mL) was added dropwise over 60 min and the reaction mixture was maintained under reflux for 2 h. The progress of the reaction was monitored by TLC using a 6:3:2 isopropanol/ammonia/water solvent system. Upon completion of reaction after 2 h, the reaction mixture was cooled to room temperature; dicyclohexyl urea was filtered off and washed with water. The filtrate was concentrated and extracted

with diethyl ether to remove excess of DCC. The aqueous solution was concentrated and purified by column chromatography on silica gel using isopropanol/ammonia/water (6:3:2) solvent system. The product was concentrated to give 0.86 g (yield = 90%) of **4-3** as an off-white solid. $^1\text{H-NMR}$ (300 MHz, D_2O) δ ppm = 2.84 (4H, m), 3.46 (4H, t, $J = 4.35$ Hz), 3.96 (2H, m), 4.24 (1H, m), 4.41 (1H, t, $J = 4.92$ Hz), 4.63 (1H, t, $J = 4.92$ Hz), 5.90 (1H, d, $J = 4.77$ Hz), 7.89 (1H, s), 8.21 (1H, s). $^{13}\text{C-NMR}$ (75 MHz, D_2O) δ ppm = 44.51, 63.65, 66.67, 70.01, 74.06, 83.44, 87.21, 117.93, 139.22, 148.21, 152.33, 154.81. ESI-MS (m/z) calculated = 417.128 [$\text{M}+\text{H}^+$], (m/z) observed = 417.120 [$\text{M}+\text{H}^+$].

The warhead, ADP-alkyne (4-4):

To the pyridinium salt form of the propargyl phosphate **4-2** (~ 2.5 mmol) was added trioctylamine (1.1 mL, 2.5 mmol) and anhydrous pyridine (30 mL) and the mixture was concentrated to an oil. It was subsequently subjected to repeated co-evaporations with anhydrous pyridine (3×4 mL) followed by the addition of AMP-morpholidate (1.04 g, 2.5 mmol). The mixture was again subjected to co-evaporations with anhydrous pyridine (3×4 mL) and finally concentrated to a minimum volume. The coupling reagent 1H-tetrazole (0.56 g, 7.96 mmol) and 4 mL of anhydrous pyridine were added and the reaction was stirred under Argon for 48 h. The solvent was then removed under reduced pressure and the reaction was quenched with saturated NaHCO_3 until no CO_2 was liberated. The mixture was washed with ethyl ether (3×20 mL), concentrated and purified by column chromatography on silica gel using isopropanol:ammonia:water (6:3:2) solvent system. The sample was further purified by Semi-Prep HPLC on C-8 column using acetonitrile and water as eluents with tributylamine (0.01%) as ion-pairing agent

afforded 0.41 g (yield = 35%) of the ADP-alkyne as off-white solid. ¹H-NMR (300 MHz, D₂O) δ ppm = 2.76 (1H, t, *J* = 2.46 Hz), 4.17 (3H, unresolved m), 4.29 (1H, m), 4.42 (1H, m), 4.47 (2H, unresolved m), 6.03 (1H, d, *J* = 5.43 Hz), 8.31 (1H, s), 8.49 (1H, s). ¹³C-NMR (75 MHz, D₂O) δ ppm = 54.03, 65.15, 70.07, 74.49, 75.69, 78.88, 83.94, 87.90, 118.25, 142.31, 144.60, 148.09, 149.63. ESI-MS (m/z) calculated = 466.052 [M+H⁺], (m/z) observed = 466.052 [M+H⁺]

3-(Dimethylamino)-1-(pyridin-3-yl) prop-2-en-1-one (4-5):⁹⁹

A 250 mL round-bottomed flask was charged with 6.4 mL (58.3 mmol) of 3-acetylpyridine, 9.4 mL (70 mmol, 1.2 eq.) of N,N-dimethylformamide dimethylacetal and 50 mL of ethanol. The reaction mixture was refluxed overnight, cooled to room temperature and the solvent was removed under reduced pressure. To the crude residue approximately 50 mL of diethyl ether was added and cooled to 0 °C. The product 3-(dimethylamino)-1-(pyridin-3-yl) prop-2-en-1-one was subsequently filtered off as yellow crystals (7.18 g, yield = 70%) and used in the subsequent steps without further purification. ¹H-NMR (300 MHz, CDCl₃) δ 2.92 (3H, s), 3.14 (3H, s), 5.65 (1H, d, *J* = 12.33 Hz), 7.32 (1H, m), 7.81 (1H, d, *J* = 12.33 Hz), 8.15 (1H, d, *J* = 9.18 Hz), 8.62 (1H, d, *J* = 7.38 Hz), 9.04 (1H, s); ¹³C-NMR (75 MHz, CDCl₃) δ 38.01, 45.85, 92.50, 123.93, 135.73, 136.30, 149.51, 152.05, 155.36, 187.02. ESI-MS (m/z) calculated = 177.102 [M+H⁺], (m/z) observed = 177.099 [M+H⁺]

4-(Pyridin-3-yl) pyrimidin-2-amine (4-6):¹⁰⁰

3-(dimethylamino)-1-(pyridin-3-yl) prop-2-en-1-one (2.64 g, 15 mmol) and guanidinium hydrochloride (1.5 g, 15.75 mmol, 1.05 eq.) were mixed in 25 mL of

2-propanol. To the suspension was added 0.7 g of NaOH (17.5 mmol, 1.17 eq.) and the mixture was refluxed for 18 h. The reaction was then cooled to 0 °C and the precipitate was filtered off, suspended in water, filtered off once more and washed with 2-propanol and diethyl ether. The residue was dried in oven at 60 to 70 °C and the desired product was obtained as a white powder (1.68 g, yield = 65%). ¹H-NMR (300 MHz, DMSO-d₆) δ 6.77 (2H, s), 7.19 (1H, d, *J* = 5.1 Hz), 7.52 (1H, m), 8.34 (1H, d, *J* = 5.1 Hz), 8.39 (1H, d, *J* = 7.89 Hz), 8.68 (1H, d, *J* = 6.57 Hz), 9.22 (1H, s); ¹³C-NMR (75 MHz, CDCl₃) δ 106.01, 123.71, 132.46, 134.11, 147.95, 151.10, 159.35, 161.56, 163.79. ESI-MS (*m/z*) calculated = 173.082 [M+H⁺], (*m/z*) observed = 173.079 [M+H⁺]

N-(2-Methyl-5-nitrophenyl)-4-(pyridin-3-yl) pyrimidin-2-amine (4-7):⁹⁴

3-(dimethylamino)-1-(pyridin-3-yl) prop-2-en-1-one **4-6** (0.456 g, 2.65 mmol, 1.1 eq.), CuI (0.12 g, 0.6 mmol, 0.25 eq.) and anhydrous K₂CO₃ (0.67 g, 4.82 mmol, 2 eq.) were added to a sealed tube (50 mL) equipped with a rubber septum and a magnetic stirring bar. The tube was filled with N₂ gas and subsequently was added a mixture of 2-bromo-1-methyl-4-nitrobenzene (0.518 g, 2.41 mmol, 1 eq.) and DMEDA (0.064 mL, 0.6 mmol, 0.25 eq.) in anhydrous dioxane (20 mL) by syringe at room temperature. The rubber septum was quickly replaced with a Teflon screw cap and the reaction mixture was stirred at 120 °C for 20 h. The reaction was cooled to room temperature and concentrated ammonia (10 mL) and saturated solution of NaCl (40 mL) were added then extracted with ethyl acetate (5 × 40 mL). The organic layers were dried over Na₂SO₄, concentrated under reduced pressure and the residue was purified by column chromatography on silica gel to give 0.62 g (yield = 83.7%) of **4-7** as a yellow powder. ESI-MS (*m/z*) calculated =

308.114 [M+H⁺], (m/z) observed = 308.120 [M+H⁺]

6-Methyl-N1-(4-(pyridin-3-yl) pyrimidin-2-yl) benzene-1,3-diamine (4-8):

To a degassed solution of 0.4 g (1.302 mmol) of N-(2-methyl-5-nitrophenyl)-4-(pyridin-3-yl) pyrimidin-2-amine **4-7** in 50 mL ethyl acetate was added 10% Pd/C (0.04 g) and stirred under a hydrogen atmosphere (balloon) for 3 h to complete disappearance of the starting material. The reaction was then filtered through celite, concentrated and the product was obtained as a yellow solid (0.36 g, quantitative yield) and used in the subsequent steps without further purification. ¹H-NMR (300 MHz, CDCl₃) δ 2.25 (3H, s), 6.40 (1H, m), 6.98 (1H, s), 7.01 (1H, m), 7.15 (1H, d, *J* = 5.1 Hz), 7.41 (1H, m), 7.60 (1H, s), 8.33 (1H, m), 8.49 (1H, d, *J* = 5.1 Hz), 8.71 (1H, m), 9.27 (1H, s). ESI-MS (m/z) calculated = 278.140 [M+H⁺], (m/z) observed = 278.101 [M+H⁺]

N-(4-Methyl-3-(4-(pyridin-3-yl) pyrimidin-2-ylamino) phenyl) propiolamide (4-9):

To a solution of propiolic acid (0.0665 mL, 1.08 mmol, 1.2 eq.) in 8 mL DMF was added 0.223 g (1.08 mmol, 1.2 eq.) of DCC followed by 0.25 g (0.902 mmol) of 6-methyl-N1-(4-(pyridin-3-yl) pyrimidin-2-yl) benzene-1,3-diamine and the reaction mixture was stirred at room temperature for 4h until the starting material was completely used up. The reaction was concentrated on a rotary evaporator and the crude product was purified by column chromatography on silica gel to give 0.2 g (yield = 67.3%) of **4-9** as a pale yellow solid. ¹³C-NMR (75 MHz, DMSO-d₆) δ 17.62, 76.90, 78.52, 107.75, 116.06, 116.45, 128.19, 130.28, 135.97, 136.18, 137.83, 146.87, 149.46, 149.94, 159.56, 160.96, 161.03. ESI-MS (m/z) calculated

= 330.135 [M+H⁺], (m/z) observed = 330.135 [M+H⁺]

Methyl 3-iodo-4-methylbenzoate (4-10):¹⁰¹

A solution of 3-iodo-4-methyl benzoic acid (0.75 g, 2.86 mmol) and catalytic amount of H₂SO₄ (~ 0.05 mL) in 20 mL methanol was refluxed for 20 h. The reaction mixture was cooled and the solvent was removed under reduced pressure. The residue obtained was dissolved in diethyl ether, washed with saturated NaHCO₃ (3×), water, brine, dried over Na₂SO₄ and the solvent was removed to give **4-10** as an orange viscous liquid (0.73 g, yield = 92.3%). ESI-MS (m/z) calculated = 276.972 [M+H⁺], (m/z) observed = 276.964 [M+H⁺]

Methyl 4-methyl-3-(4-(pyridin-3-yl) pyrimidin-2-ylamino) benzoate (4-11):

3-(dimethylamino)-1-(pyridin-3-yl) prop-2-en-1-one **4-6** (0.415 g, 2.41 mmol), CuI (0.12 g, 0.6 mmol, 0.25 eq.) and anhydrous K₂CO₃ (0.67 g, 4.82 mmol, 2 eq.) were added to a sealed tube (50 mL) equipped with a rubber septum and a magnetic stirring bar. The tube was filled with N₂ gas and subsequently was added a mixture of methyl 3-iodo-4-methylbenzoate **7** (0.67 g, 2.41 mmol, 1 eq.) and DMEDA (0.064 mL, 0.6 mmol, 0.25 eq.) in anhydrous dioxane (20 mL) by syringe at room temperature. The rubber septum was quickly replaced with a Teflon screw cap and the reaction mixture was stirred at 120 °C for 20 h. The reaction was cooled to room temperature and concentrated ammonia (10 mL) and saturated solution of NaCl (40 mL) were added and then extracted with ethyl acetate (5 × 40 mL). The organic layers were dried over Na₂SO₄, concentrated under reduced pressure and the residue was purified by column chromatography on silica gel to give 0.63 g (yield = 81%) of **4-11** as a pale yellow powder. ¹H-

NMR (300 MHz, CDCl₃) δ 2.39 (3H, s), 3.95 (3H, s), 7.19 (1H, d, $J = 5.25$ Hz), 7.29 (1H, m), 7.42 (1H, m), 7.72 (1H, m), 8.41 (1H, m), 8.49 (1H, d, $J = 5.1$ Hz), 8.71 (1H, m), 8.98 (1H, s), 9.29 (1H, s); ¹³C-NMR (75 MHz, CDCl₃) δ 18.86, 52.67, 108.90, 122.90, 124.20, 125.04, 129.20, 131.06, 133.01, 133.88, 135.10, 138.12, 149.15, 152.12, 159.75, 161.01, 163.00, 167.66. ESI-MS (m/z) calculated = 321.135 [M+H⁺], (m/z) observed = 321.130 [M+H⁺]

4-Methyl-3-(4-(pyridin-3-yl) pyrimidin-2-ylamino) benzoic acid (4-12):

LiOH (0.24 g, 10 mmol, 8 eq.) was dissolved in a 3:1 mixture of methanol and water (12 mL). To this was added the methyl ester **4-11** (0.4 g, 1.25 mmol) and the reaction mixture was stirred for 8 h at room temperature. The reaction was acidified to $pH = 1$ with 1 N HCl. The solvent was removed under reduced pressure and the residue obtained was purified by column chromatography on silica gel to obtain 0.23 g (yield = 60%) of **4-12** as an off-white solid. ¹H-NMR (300 MHz, DMSO-d₆) δ 2.30 (3H, s), 7.29 (1H, d, $J = 8.07$ Hz), 7.45 (1H, d, $J = 5.28$ Hz), 7.53 (1H, m), 7.64 (1H, m), 8.22 (1H, s), 8.45 (1H, m), 8.52 (1H, d, $J = 5.25$ Hz), 8.68 (1H, m), 9.01 (1H, s), 9.26 (1H, s). ESI-MS (m/z) calculated = 307.118 [M+H⁺], (m/z) observed = 307.075 [M+H⁺]

4-Methyl-N-(prop-2-ynyl)-3-(4-(pyridin-3-yl) pyrimidin-2-ylamino) benzamide (4-13):

The acid **4-12** (0.2 g, 0.653 mmol) was added to 10 mL of DMF followed by HATU (0.3 g, 0.784 mmol, 1.2 eq.) and DIEA (0.133 mL, 0.784 mmol, 1.2 eq.). After stirring for 5 min. was added propargylamine (0.054 mL, 0.784 mmol, 1.2 eq.) and the reaction was stirred for 2 h. The solvent was removed under reduced

pressure and the crude product obtained was purified by column chromatography on silica gel to give 0.15 g (yield = 67%) of the Warhead-2 as an off-white solid. ¹H-NMR (300 MHz, DMSO-d₆) δ 2.30 (3H, s), 3.10 (1H,s), 4.05 (2H, m), 7.35 (1H, d, *J* = 8.07 Hz), 7.54 (1H, d, *J* = 5.07 Hz), 7.61 (1H, m), 7.87 (1H, m), 8.13 (1H, s), 8.60 (1H, d, *J* = 5.25 Hz), 8.78 (1H, d, *J* = 8.07 Hz), 8.86 (2H, m), 9.21 (1H, s), 9.35 (1H, s); ¹³C-NMR (75 MHz, DMSO-d₆) δ 18.03, 28.45, 72.67, 81.41, 107.99, 123.26, 124.29, 125.50, 130.23, 131.76, 136.11, 137.68, 138.44, 144.64, 147.66, 158.57, 159.76, 160.10, 160.95, 165.60. ESI-MS (m/z) calculated = 344.151 [M+H⁺], (m/z) observed = 344.140 [M+H⁺]

N-(4-Methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenyl)-1-(2-(2-methylphenylsulfonamido)ethyl)-1H-1,2,3-triazole-4-carboxamide [W1-SA20]:

100 μl of a stock solution of **W1** (100 mM) in DMSO was mixed with a 50 mM stock solution of the azide **SA20** (200 μl in DMSO). To this was added 200 μl water, 100 μl methanol, 50 μL copper sulphate solution (10 mM stock solution in water) and 50 μl sodium ascorbate solution (50 mM stock solution in water). The reaction was left overnight on a shaking incubator at room temperature. Upon completion of the reaction (< 12 h), the crude reaction mixture was purified by reverse phase Semi-Prep HPLC on a C-18 column. ¹H-NMR (300 MHz, DMSO-d₆) δ 2.23 (3H, s), 2.51 (3H, s), 3.34 (2H, m), 4.48 (2H, t, *J* = 5.59 Hz), 7.20 (1H, d, *J* = 8.37 Hz), 7.35 (2H, m), 7.47 (2H, d, *J* = 5.25 Hz), 7.53 (1H, m), 7.67 (1H, m), 7.76 (1H, d, *J* = 7.71 Hz), 7.97 (1H, t, *J* = 5.91 Hz), 8.15 (1H, s), 8.55 (1H, d, *J* = 5.07 Hz), 8.56 (1H,s), 8.63 (1H, d, *J* = 7.86 Hz), 8.76 (1H, bs), 9.03 (1H, s), 9.32 (1H, bs), 10.32 (1H, s).

N-(4-Methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenyl)-1-(2-(2-methylphenylsulfonamido)ethyl)-1H-1,2,3-triazole-4-carboxamide [W2-SA13]:

100 μ l of a stock solution of **W2** (100 mM) in DMSO was mixed with a 50 mM stock solution of the azide **SA13** (200 μ l in DMSO). To this was added 200 μ l water, 100 μ l methanol, 50 μ L copper sulphate solution (10 mM stock solution in water) and 50 μ l sodium ascorbate solution (50 mM stock solution in water). The reaction was left overnight on a shaking incubator at room temperature. Upon completion of the reaction (< 12 h), the crude reaction mixture was purified by reverse phase Semi-Prep HPLC on a C-18 column. ¹H-NMR (300 MHz, DMSO-d₆) δ 2.29 (6H, s), 3.20 (2H, m), 4.37 (2H, t, J = 6.00 Hz), 4.49 (2H, d, J = 5.43 Hz), 7.33 (1H, d, J = 7.89 Hz), 7.49 (4H, m), 7.63 (1H, d, J = 7.89 Hz), 7.79 (1H, m), 7.95 (2H, m), 8.14 (1H, s), 8.57 (1H, d, J = 5.28 Hz), 8.72 (1H, d, J = 8.04 Hz), 8.83 (1H, bs), 8.94 (1H, t, J = 5.76 Hz), 9.17 (1H, s), 9.32 (1H, bs).

High Resolution Mass Spectrometry (HRMS) results of the selected click-inhibitors:

W1-SA9 [C₂₇H₂₄FN₉O₃S], (m/z)_{calculated} = 596.1599 [C₂₇H₂₄FN₉O₃SNa] (m/z)_{observed} = 596.1586 (ESI-HRMS)

W1-SA10 [C₂₇H₂₄FN₉O₃S], (m/z)_{calculated} = 596.1599 [C₂₇H₂₄FN₉O₃SNa] (m/z)_{observed} = 596.1590 (ESI-HRMS)

W1-SA17 [C₃₁H₂₇N₉O₃S], (m/z)_{calculated} = 606.2036 [C₃₁H₂₈N₉O₃S] (m/z)_{observed} = 606.2023 (ESI-HRMS)

W1-SA20 [C₂₈H₂₇N₉O₃S], (m/z)_{calculated} = 570.2036 [C₂₈H₂₈N₉O₃S] (m/z)_{observed} = 570.2031 (ESI-HRMS)

W1-SA22 [C₂₅H₂₃N₉O₃S₂], (m/z)_{calculated} = 584.1258 [C₂₅H₂₃N₉O₃S₂Na] (m/z)_{observed} = 584.1251 (ESI-HRMS)

W2-SA3 [C₂₈H₂₆ClN₉O₃S], (m/z)_{calculated} = 626.1460 [C₂₈H₂₆ClN₉O₃SNa] (m/z)_{observed} = 626.1441 (ESI-HRMS)

W2-SA4 [C₂₆H₂₄ClN₉O₃S₂], (m/z)_{calculated} = 632.1024 [C₂₆H₂₄ClN₉O₃S₂Na] (m/z)_{observed} = 632.0999 (ESI-HRMS)

W2-SA7 [C₂₈H₂₅F₂N₉O₃S], (m/z)_{calculated} = 628.1661 [C₂₈H₂₅F₂N₉O₃SNa] (m/z)_{observed} = 628.1692 (ESI-HRMS)

W2-SA13 [C₂₉H₂₈FN₉O₃S], (m/z)_{calculated} = 624.1912 [C₂₉H₂₈FN₉O₃SNa] (m/z)_{observed} = 624.1939 (ESI-HRMS)

W2-SA17 [C₃₂H₂₉N₉O₃S], (m/z)_{calculated} = 642.2006 [C₃₂H₂₉N₉O₃SNa] (m/z)_{observed} = 642.2013 (ESI-HRMS)

W2-SA18 [C₃₂H₂₉N₉O₃S], (m/z)_{calculated} = 642.2006 [C₃₂H₂₉N₉O₃SNa] (m/z)_{observed} = 642.2004 (ESI-HRMS)

Chapter 5: A mechanism-based cross-linker for protein kinase-substrate complexes

Summary:

This chapter summarizes the synthesis and biochemical evaluation of an improved mechanism-based chemical cross-linker, naphthalene 2,3-dicarboxaldehyde-adenosine (NDA-AD), for the identification of kinase-substrate interactions from crude proteome. The design of the cross-linker was based on a previously reported mechanism-based cross-linker, orthophthalaldehyde-adenosine (OPA-AD), developed by Shokat et al. that potentially allows researchers to use known phosphopeptides/proteins to identify their upstream protein kinases. Although OPA-AD could detect kinase-substrate interactions with purified kinases, the compound was found to produce several non-specific labelling in a complex proteome which severely limit its potential applications. This prompted us to develop the improved version of the cross-linker. In addition to improved labeling performances, the compound NDA-AD was found to be suitable for the detection of kinase-substrate interactions of both tyrosine-specific and serine/threonine-specific protein kinases.

5. 1 Introduction

The human genome encodes 518 protein kinases, and hundreds of thousands of potential phosphorylation sites are present in the human proteome.¹⁰² It is estimated that at any given time in the life cycle of a cell, approximately 30% of cellular proteins are phosphorylated on at least one residue.^{4b} Existing biochemical and proteomic methods have offered invaluable tools for the identification of phosphoproteins as well as phosphorylation sites in proteins.¹⁰³ Powerful methods for the identification of kinase substrates have also been developed.¹⁰⁴ However, the opposite problem of identification of the upstream kinase for a specific phosphorylation event/phosphoprotein remains a much more challenging one, mainly

due to the very transient and weak-binding nature of kinase-substrate interactions. Towards this, Shokat and co-workers developed a mechanism-based cross-linker, *o*-phthalaldehyde-adenosine (OPA-AD), which potentially allows researchers to use known phosphoproteins/phosphopeptides to identify their upstream kinases.¹⁰⁵ OPA and its derivatives have conventionally been used in amino acid/peptide analysis where a three-component reaction between primary amino group from amino acids, formyl functionality of OPA and a proximal thiol functional group or a nucleophilic cyanide ion leads to the formation of a stable isoindole product which facilitates sensitive detection of the primary amino functionality.¹⁰⁶ The compound OPA-AD enabled the researches to carryout this three-component reaction at the active site of kinases as the adenosine unit in the reagent drives the molecule to the ATP-binding pocket of the enzyme. Once the molecule occupies the ATP-binding pocket, the catalytic lysine residue (this lysine is conserved in most of the kinases) forms a reversible imine adduct with one of the formyl functionalities in OPA, which subsequently reacts with a thiol functionality (from a cysteine residue) from a pseudo-peptide substrate leading to the formation of an isoindole product, cross-linking the kinase with its cognate substrate partner.

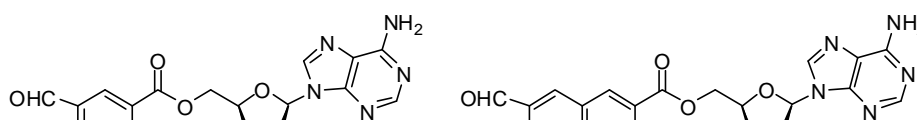
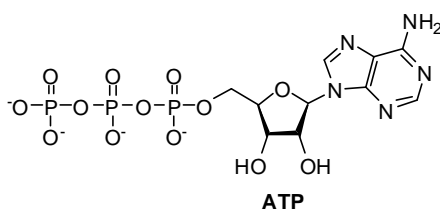
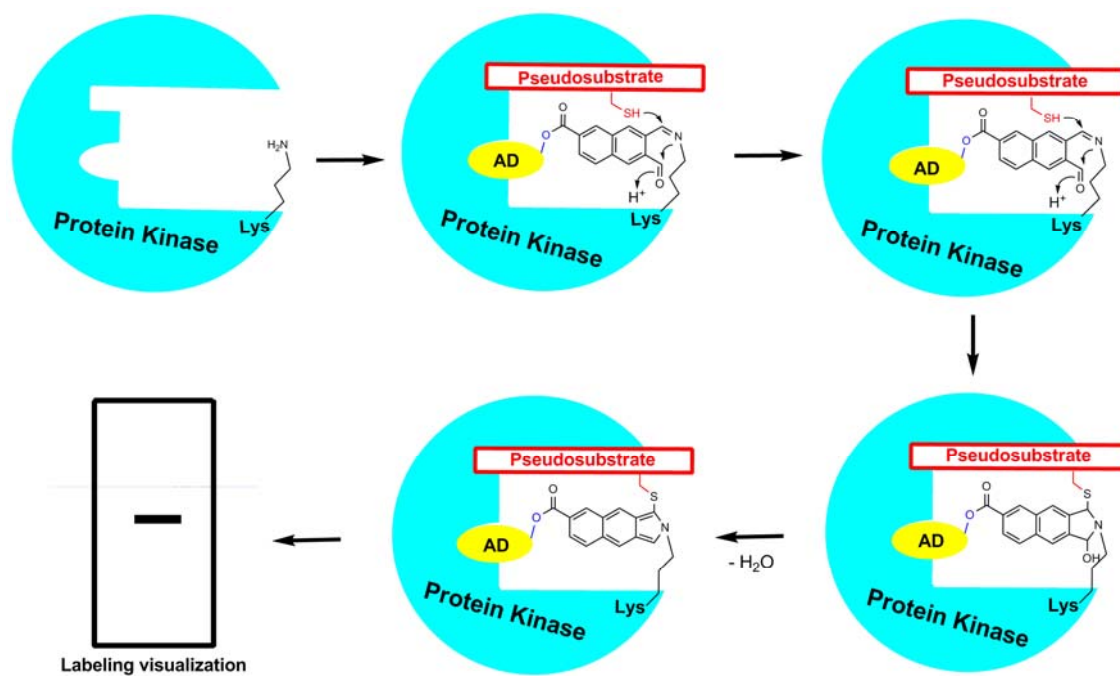


Figure 5. 1 Chemical structure of ATP, OPA-AD and NDA-AD.

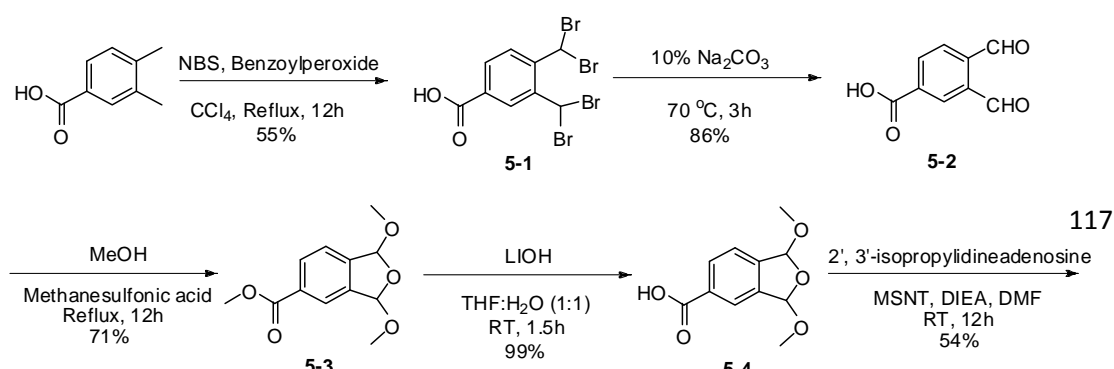


Scheme 5. 1 Scheme showing the three-component cross-linking reaction of the kinase with its pseudosubstrate and NDA-AD.

5. 2 Synthesis of the cross-linkers

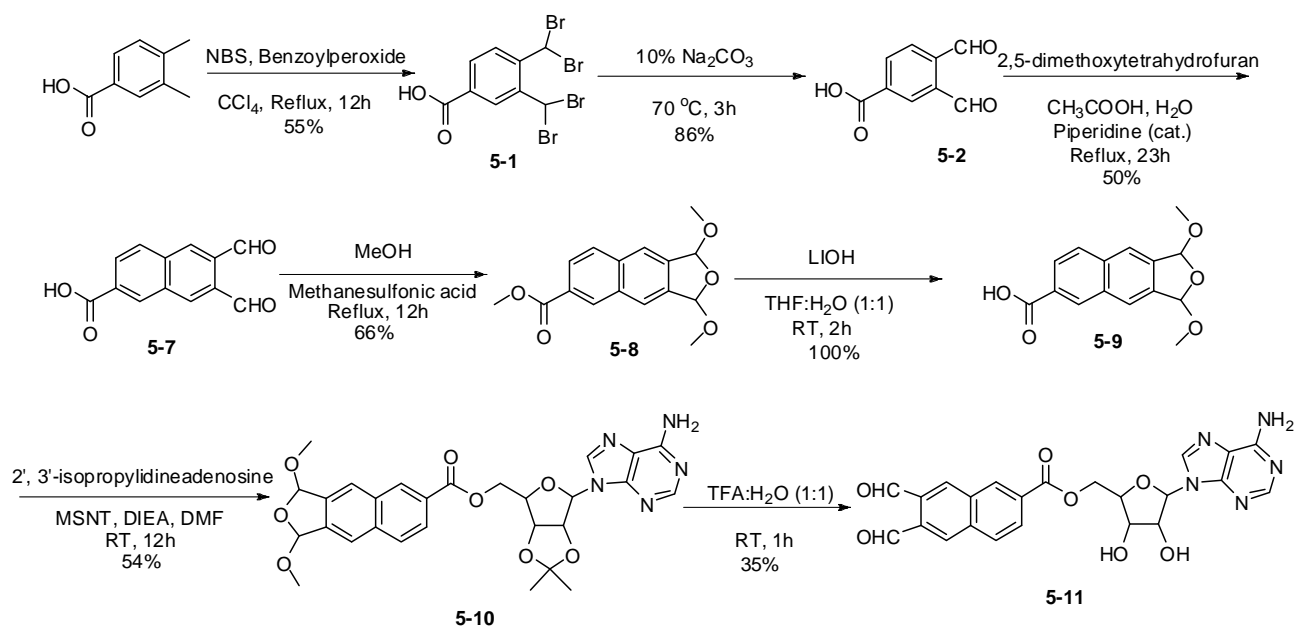
5. 2. 1 Synthesis of OPA-AD

The cross-linker OPA-AD was synthesized as previously reported.¹⁰⁵ It was resynthesized for a direct comparison of its labeling performance with our cross-linker NDA-AD. Briefly, the commercially available compound 3,4-dimethylbenzoic acid was converted to the tetrabromo product **5-1** via treatment with N-bromosuccinimide in the presence of benzoyl peroxide under reflux condition. Compound **5-1** was then converted to 3,4-diformylbenzoic acid (**5-2**) via treatment with a hot aqueous solution of sodium carbonate followed by acidic hydrolysis. The formyl groups in **5-2** were subsequently protected with a cyclic bis-acetal group via treatment with “super-dry” methyl alcohol in the presence of catalytic amount of methanesulfonic acid under reflux condition. The methyl ester in compound **5-3** was subsequently deprotected to free acid by treatment with LiOH yielded the compound **5-4**, which was used to couple to the 5'-hydroxyl group of the commercially available compound 2',3'-isopropylideneadenosine using the coupling reagent MSNT to give the compound **5-5**. Finally, the cyclic bis-acetal protecting group in **5-5** was deprotected by treatment with 1:1 TFA:H₂O afforded the target compound, OPA-AD (**5-6**), which was purified by Prep-HPLC using reverse-phase C-18 column with a suitable gradient of water with 0.1% TFA and acetonitrile with 0.1% TFA as the mobile phases.



Scheme 5. 2 Synthesis of the cross-linker, OPA-AD

5. 2. 2. Synthesis of NDA-AD



Scheme 5. 3 Synthesis of the cross-linker, NDA-AD

The synthesis of the cross-linker, NDA-AD, was started with the intermediate **5-2** from the OPA-AD synthetic scheme. A double aldol condensation reaction of **5-2** with 2,5-dimethoxytetrahydrofuran in a mixture of water and glacial acetic acid with

catalytic amount of piperidine afforded the key intermediate 6,7-diformyl-2-naphthoic acid (**5-7**). Treatment of **5-7** with “super-dry” methyl alcohol in the presence of catalytic amount of methanesulfonic acid under reflux condition yielded the cyclic bis-acetal compound **5-8**, which was subsequently treated with LiOH to get **5-9**. The free carboxylic acid group in compound **5-9** was subsequently coupled to the 5'-hydroxyl group of the commercially available compound 2',3'-isopropylideneadenosine using the coupling reagent MSNT yielded the compound **5-10**. Finally, the cyclic bis-acetal protecting group in **5-10** was deprotected by treatment with 1:1 TFA:H₂O afforded the target compound, NDA-AD (**5-11**), which was purified by Prep-HPLC using reverse-phase C-18 column with a suitable gradient of water with 0.1% TFA and acetonitrile with 0.1% TFA as the mobile phases.

5. 3 Synthesis of Peptide Pseudosubstrates

Solid phase peptide synthesis (SPPS) was performed on an automated peptide synthesizer (Chemspeed Technologies ASW2000) utilizing standard Fmoc chemistry on rink amide resin (ChemPep Inc.). Fmoc amino acids (Advanced ChemTech) were activated by DIC/HOBt (Advanced ChemTech) and the coupling reaction was performed for 2 hours with amino acids (4 eq) during each cycle. After washing with DMF (6X), the Fmoc-protecting group was removed using 20% piperidine in DMF, followed by another washing with DMF (6X). The N-terminus was capped by agitating the resin for 20 hours in the presence of a fluorescein tag (4eq), HATU (4eq) and DIEA (8eq), or a biotin tag (4eq), HATU (4eq) and DIEA (8eq). Prior to cleavage, the resin was washed with MeOH (6X), DCM (6X), and MeOH (6X). The synthesized peptides were cleaved from the resin and deprotected from its side chain protecting groups in a 1mL cleavage mixture containing 94% TFA, 2.5% H₂O, 2.5%

EDT and 1% TIS for 2.5 hours. After filtration and precipitation with ether, the precipitated peptides were washed once with TFA and twice with ether and dried using Genevac (Biomedica). Each peptide was then dissolved in 1mL DMF and characterized by LC-MS (Shimadzu).

Csk-pseudosubstrate (Csktide): Fluorescein-GG-KKKKKEEICFFF. Molecular weight calculated = 1960, LC-MS: $m/z [M+2H]/2 = 980.9$.

Erk 1 and Erk 2 -pseudosubstrates: Fluorescein-GG-ELVEPLCPSGEAPNQ. Molecular weight calculated = 2096, LC-MS: $m/z [M+2H]/2 = 1048.9$.

Src-pseudosubstrate: Fluorescein-GG-KVEKIGEGTCGVVYK. Molecular weight calculated = 2122, LC-MS: $m/z [M+2H]/2 = 1062.0$.

Abl-pseudosubstrate: Fluorescein-GG-EAICAAPFAKKK. Molecular weight calculated = 1790, LC-MS: $m/z [M+2H]/2 = 895.9$.

PKA-pseudosubstrate: Fluorescein-GG-LRRACLG. Molecular weight calculated = 1300, LC-MS: $m/z [M+2H]/2 = 651.3$.

5. 4 Results and Discussions

Although OPA-AD was found to be useful in cross-linking kinase-pseudosubstrate pairs with purified kinases, we observed that in the presence of complex proteome, the compound produces a large number of non-specific cross-linking bands in addition to the one corresponding to the desired kinase-pseudosubstrate pair. This non-selective cross-linking severely limits its potential applications. So we aimed to improve the selectivity of the cross-linker via modulating the reactivity of the dialdehyde unit. We hypothesized that replacing the highly reactive OPA moiety with a less reactive naphthalene-2,3-dicarboxaldehyde (NDA) group may improve the

labeling selectivity, which was indeed found to be true as evident from the much more cleaner cross-linking of Pka/Pkatide complex with the new cross-linker naphthalene-2,3-dicarboxaldehyde-adenosine (NDA-AD) in the presence of competing cellular proteins from *E. coli* DE3 cell lysate (Figure 5.2). The NDA moiety in the improved cross-linker is also a better structural fit in the kinase active site compared to OPA in the OPA-AD.

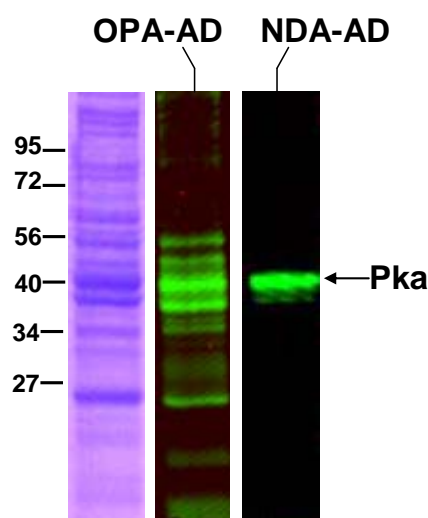


Figure 5. 2 Comparative labeling profiles of OPA-AD and NDA-AD for the kinase PKA (with PKA-pseudosubstrate) in the presence of bacterial cell lysate.

Next we assessed whether the NDA-AD could serve as a general mechanism-based cross-linker for both Tyrosine and Serine/Threonine kinases. The cross-linking reactions were tested with a set of 6 purified kinases, of which three are Tyr kinases (Csk, Src & Abl) and the other three Ser/Thr kinases (Erk1, Erk2 & Pka). All kinases were recombinantly expressed and tested to ensure their purity as well as enzymatic activities. Fluoresceine-labelled, cysteine-containing kinase pseudosubstrates were chemically synthesized based on their known peptide substrate sequences. As shown in Figure 5.3, incubation of each of the six kinases, regardless of whether they are Tyr

or Ser/Thr kinases, with their cognate pseudosubstrates in the presence of NDA-AD led to the successful cross-linking of kinase-substrate complex, as indicated by a fluorescence band on the SDS-PAGE. All three components (i.e. kinase, pseudosubstrate and NDA-AD) were necessary, as labelling was not observed in the absence of any of the components. No cross-linking was observed with heat-denatured kinases, indicating the cross-linking was indeed activity-dependent.

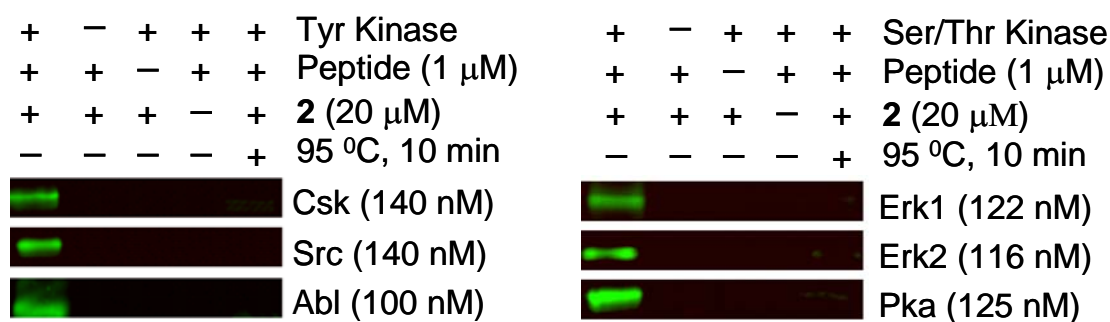


Figure 5. 3 Fluorescence-scanned gels showing cross-linking profiles of NDA-AD against three different Tyr kinases (left) and three different Ser/Thr kinases (right).

5. 5 Conclusions

An improved mechanism based kinase-pseudosubstrate cross-linker, NDA-AD, was synthesized with which we demonstrated a general and useful approach for identifying kinase activities from a crude proteome.¹⁰⁸ The cross-linker NDA-AD, not only showed improved labeling performances from crude cellular proteome but was found to be compatible with detecting the kinase-substrate pairs of both serine/threonine-specific and tyrosine-specific protein kinases. Thus given the fact that peptide pseudosubstrates of any sequence could be synthesized easily, this approach would be compatible for detecting activities of any diverse sets of kinases. The establishment of the NDA-adenosine guided kinase/peptide cross-linking reaction and the initial applications represents a significant step towards tackling the problem of identifying the upstream kinase for a given phosphoprotein of interest, although much more progress remains to be made to accomplish this goal. Further improvements in the strategy that allows cross-linking of protein substrates rather than peptide pseudosubstrates with full-length endogenous protein kinases in cellular lysates, and more desirably in intact live cells, would provide an invaluable tool in the detailed investigations of cellular signal transduction networks in which protein kinases play significant roles.

5. 6 Synthetic details and characterizations of compounds

3,4-Bis(dibromomethyl)benzoic acid (5-1):

3,4-dimethylbenzoic acid (15.33 g, 0.1 mol) was dissolved in hot carbon tetrachloride (250 mL). N-bromosuccinimide (73.70 g, 0.41 mol, 4.1 eq.) and benzoylperoxide (3 g, cat.amt.) were added and the reaction mixture was refluxed for 12 hours while stirring. The reaction was then cooled down to room temperature and filtered. The solid filtered out was washed with benzene (50 mL) three times and the residue obtained was dissolved in diethyl ether (250 mL) and stirred for 2 hours. The mixture was then filtered and the residue was again dissolved in diethyl ether (250 mL) and stirred for 2 hours. The mixture was filtered for the last time and the ether extracts were combined and concentrated under reduced pressure to dryness. The orange solid obtained was recrystallized from acetonitrile to yield 26.0 g (yield = 54.7%) of compound of **5-1** in white powdered form. ¹H-NMR (DMSO-d₆): δ 7.77 (s, 1H), 7.79 (s, 1H), 7.93-8.03 (m, 2H), 8.35 (s, 1H), 13.54 (bs, 1H, exchangeable with D₂O addition, COOH). ESI-MS: m/z [M-H]⁻ = 464.9.

3,4-Diformylbenzoic acid (5-2):

3,4-bis(dibromomethyl)benzoic acid **5-1** (20 g, 0.04 mol) was dissolved in 10% aqueous sodium carbonate solution (200 mL). The mixture was heated at 70°C for 3 hours and the precipitation evolved was filtered out. The filtrate obtained was acidified with concentrated HCl until pH 1 while cooling on ice. The acidic solution was then extracted with ethyl acetate three times. The organic layers were combined, washed with deionized water and dried with anhydrous sodium sulphate. Solvent was removed on a rotary evaporator to give 6.7 g (yield = 85.9%) of compound **5-2** as a pale yellow powder which was used for subsequent steps without further purification. ¹H-NMR (DMSO-d₆): δ 8.06 (d, 1H, *J* = 7.89 Hz), 8.34 (dd, 1H, *J*₁ = 7.98 Hz, *J*₂ =

1.74 Hz), 8.49 (d, 1H, $J = 1.65$ Hz), 10.50 (s, 1H), 10.54 (s, 1H), 13.53 (bs, 1H, exchangeable with D₂O addition, COOH). ESI-MS: m/z [M-H]⁻ = 177.0.

Methyl 1,3-dimethoxy-1,3-dihydroisobenzofuran-5-carboxylate (5-3):

3,4-diformylbenzoic acid **5-2** (0.45 g, 2.5 mmol) was dissolved in 'super-dry' methanol (62.5 mL) in a flame-dried rbf under nitrogen gas atmosphere. Catalytic amount of methanesulfonic acid was added and the reaction was refluxed overnight with stirring. The resulting reaction mixture was cooled and concentrated on a rotary evaporator to afford an amber coloured oil. The oil was dissolved in ethyl acetate and washed with water. The water layer was washed with ethyl acetate two times. The organic layers were combined and washed with water and brine, dried over anhydrous sodium sulphate and concentrated on a rotary evaporator to afford an amber coloured oily compound which was then purified by silica gel flash column chromatography to yield compound **5-3** as a yellowish oil 0.4 g (yield = 71.0%). ¹H-NMR (DMSO-d₆): δ 3.31 (s, 6H), 3.88 (s, 3H), 6.10 (s, 1H), 6.36 (s, 1H), 7.57 (d, 1H, $J = 7.89$ Hz), 7.95 (s, 1H), 8.05 (d, 1H, $J = 7.89$ Hz). ESI-MS: m/z [M+H]⁺ = 239.1.

1,3-Dimethoxy-1,3-dihydroisobenzofuran-5-carboxylic acid (5-4):

Methyl 1,3-dimethoxy-1,3-dihydroisobenzofuran-5-carboxylate **5-3** (0.5 g, 2.1 mmol) was dissolved in a 1:1 mixture of tetrahydrofuran and water (15 mL:15 mL). Lithium hydroxide (0.06 g, 2.5 mmol, 1.2 eq.) was added and the yellow coloured solution was stirred at room temperature for 1.5 hours. Solvent was then removed on a rotary evaporator and the remaining water was washed with ether. The pH of the aqueous layer was adjusted very cautiously and quickly to pH = 2 with 1N HCl, and then quickly extracted with ethyl acetate. The organic layer was washed with brine, dried over anhydrous sodium sulphate and concentrated to yield 0.46 g (yield 98.8%) of

compound **5-4** as a yellow powder. ¹H-NMR (DMSO-d₆): δ 3.31 (s, 3H), 3.35 (s, 3H), 6.10 (s, 1H), 6.35 (s, 1H), 7.53 (d, 1H, *J* = 7.89 Hz), 7.92 (s, 1H), 8.03 (d, 1H, *J* = 7.89 Hz). ESI-MS: *m/z* [M-H]⁻ = 223.2.

(6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl 1,3-dimethoxy-1,3-dihydroisobenzofuran-5-carboxylate (5-5):

1,3-dimethoxy-1,3-dihydroisobenzofuran-5-carboxylic acid **5-4** (140 mg, 0.6 mmol), 2',3'-isopropylideneadenosine (375 mg, 1.2 mmol, 2 eq.), DIEA (0.21 mL, 1.25 mmol, 2.08 eq.), and DMF (4 mL) were added into a flame-dried rbf under nitrogen gas atmosphere. The reaction mixture was cooled to 0°C and the coupling reagent MSNT (180 mg, 0.6 mmol, 1 eq.) was added portion-wise. The reaction mixture was stirred at room temperature for 12 hours and then concentrated. The resultant red oil was dissolved in ethyl acetate (100 mL), and extracted with 10% citric acid (50 mL), and 10% sodium bicarbonate (50 mL). The organic layer was then dried over anhydrous sodium sulphate, concentrated and purified over silica gel flash column chromatography (5% methanol in chloroform) to yield 0.17 g (yield 54.0%) of **5-5** as a crystalline solid. ¹H-NMR (CDCl₃): δ 1.42 (s, 3H), 1.63 (s, 3H), 3.44-3.47 (m, 6H), 4.49-4.66 (m, 3H), 5.17-5.20 (m, 1H), 5.55-5.62 (m, 1H), 6.02-6.10 (m, 2H), 6.27-6.31 (m, 1H), 7.37-7.43 (m, 1H), 7.86-8.01 (m, 3H), 8.32 (d, 1H, *J* = 5.76Hz). ESI-MS: *m/z* [M+H]⁺ = 514.7.

(5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl 3,4-diformylbenzoate (5-6):

Compound **5-5** (15 mg, 29 μmol) was dissolved in 1:1 TFA/H₂O (0.5 mL:0.5 mL) and stirred for 1 hour at room temperature. TFA was removed on a rotary evaporator and the resultant aqueous solution was purified using Prep-HPLC on a C-18 reverse-phase

column. (CH₃CN with 0.1% TFA and H₂O with 0.1% TFA as the mobile phases). The purified product was concentrated by lyophilisation to afford 5 mg (yield = 40%) of compound **5-6** as a white powder. ¹H-NMR (1:1 DMSO-d₆/D₂O-d₂): δ 4.30-4.45 (m, 4H), 4.65-4.78 (m, 3H), 5.90-5.92 (m, 1H), 6.15-6.18 (m, 1H), 6.43-6.44 (m, 1H), 7.41- 7.45 (m, 1H), 7.75-7.87 (m, 2H), 8.01-8.03 (m, 1H), 8.21-8.22 (m, 1H). ESI-MS: m/z [M+H]⁺ = 428.1.

6,7-Diformyl-2-naphthoic acid (5-7):¹⁰⁷

3,4-diformylbenzoic acid **5-2** (0.6 g, 3.37 mmol) was placed in a 25 mL rbf. Deionized water (3 mL), glacial acetic acid (2 mL), 2,5-dimethoxytetrahydrofuran (0.436 mL, 3.37 mmol, 1 eq.) and 3 drops of piperidine catalyst were added and the reaction mixture was refluxed for 23 hours with stirring. The orange precipitate formed was collected by filtration, washed twice with water, twice with 1N HCl, and once with ether. The solids were dried further under vacuum to yield 0.38 g (yield = 49.7%) of compound **5-7**. ¹H-NMR (DMSO-d₆): δ 8.23 (dd, 1H, *J*₁ = 8.55 Hz, *J*₂ = 1.65 Hz), 8.36 (d, 1H, *J* = 8.70 Hz), 8.67 (s, 1H), 8.81 (s, 1H), 8.88 (s, 1H), 10.52 (s, 1H), 10.56 (s, 1H).

Methyl 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylate (5-8):

6,7-diformyl-2-naphthoic acid **5-7** (0.38 g, 1.66 mmol) was dissolved in 'super-dry' methanol (60 mL) in a flame-dried rbf under nitrogen gas atmosphere. Catalytic amount of methanesulfonic acid was added and the reaction mixture was refluxed overnight with stirring. The reaction mixture was then cooled and concentrated to afford an amber coloured oil. The oily compound was dissolved in ethyl acetate and washed with water. The aqueous layer was washed with ethyl acetate two times. The organic layers were combined and washed with water and brine. The organic layer

was then dried over anhydrous sodium sulphate, concentrated and purified over silica gel flash column chromatography to yield 0.31 g (yield = 65.7%) of compound **5-8** as a yellowish oil. ¹H-NMR (CDCl₃): δ 3.50 (s, 3H), 3.52 (s, 3H), 3.99 (s, 3H), 6.21 (d, 1H, *J* = 3.78 Hz), 6.45 (s, 1H), 7.91 (s, 1H), 7.95 (s, 1H), 7.98 (s, 1H), 8.09 (dd, 1H, *J*₁ = 8.55 Hz, *J*₂ = 1.65 Hz), 8.65 (s, 1H).

1,3-Dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylic acid (5-9):

Methyl 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylate **5-8** (0.2 g, 0.7 mmol) was dissolved in a 1:1 mixture of tetrahydrofuran and water (10 mL:10 mL). Lithium hydroxide (0.04 g, 1.4 mmol, 2 eq.) was added and the yellow solution was stirred at room temperature for 2 hours. Solvent was then removed on a rotary evaporator and the remaining water was washed with ether. The pH of the aqueous layer was adjusted very cautiously and quickly to pH = 2 with 1N HCl, and then quickly extracted with ethyl acetate. The organic layer was washed with brine, dried over sodium sulphate and concentrated to yield 0.19 g (yield = 99.5%) of compound **5-9** as a yellow powder. ¹H-NMR (CDCl₃): δ 3.52 (s, 3H), 3.54 (s, 3H), 6.23 (s, 1H), 6.47 (s, 1H), 7.94-8.02 (m, 3H), 8.14 (d, 1H, *J* = 1.47 Hz), 8.76 (s, 1H). ESI-MS: *m/z* [M-H]⁻ = 273.2.

(6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylate (5-10):

1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylic acid **5-9** (100 mg, 0.36 mmol), 2',3'-isopropylideneadenosine (224 mg, 0.72 mmol, 2eq.), DIEA (0.13 mL, 0.76 mmol, 2.08 eq.), and DMF (4 mL) were added into a flame-dried rbf under nitrogen gas atmosphere. The reaction mixture was cooled to 0°C and MSNT (107 mg, 0.36 mmol, 1 eq.) was added portion-wise. The reaction mixture was stirred at

room temperature for 12 hours and then concentrated on a rotary evaporator. The resultant red oil was dissolved in ethyl acetate (100 mL), and extracted with 10% citric acid (50 mL) followed by 10% sodium bicarbonate (50 mL). The organic layer was then dried over anhydrous sodium sulphate, concentrated and purified over silica gel flash column chromatography to yield 0.11 g (yield = 54.3%) of **5-10** as a crystalline solid. ¹H-NMR (CDCl₃): δ 1.44 (s, 3H), 1.65 (s, 3H), 3.50-3.56 (m, 6H), 4.57-4.59 (m, 1H), 4.68-4.72 (m, 2H), 5.21-5.24 (m, 1H), 5.61-5.63 (m, 1H), 6.09-6.19 (m, 2H), 6.44 (m, 1H), 7.84-7.96 (m, 5H), 8.28 (s, 1H), 8.44-8.50 (m, 1H). ESI-MS: m/z [M+H]⁺ = 564.0.

(5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl 6,7-diformyl-2-naphthoate (5-11):

(6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylate **5-10** (100 mg, 0.18 mmol) was dissolved in 1:1 TFA/H₂O (7 mL:7 mL) and stirred for 2 hours at room temperature. TFA was removed under vacuum and the resultant aqueous solution was purified using Prep-HPLC on a reverse-phase C-18 column. The purified product was concentrated by lyophilisation to afford 30 mg (yield = 34.9%) of compound **5-11** as an off-white powder. ¹H-NMR (1:1 DMSO-d₆): δ 4.32 (m, 2H), 4.51-4.56 (m, 2H), 4.69-4.77 (m, 2H), 5.01 (s, 1H), 5.97 (s, 2H), 6.23 (s, 1H), 8.10-8.41 (m, 3H), 8.68-9.06 (m, 4H), 10.53 (s, 1H), 10.56 (s, 1H). ESI-MS: m/z [M+H]⁺ = 478.1.

Chapter 6: Small Molecule Probes that Target Abl Kinase

Summary:

This chapter summarizes the development of selective small molecule-based activity-based probes (ABPs) for the Abelson (Abl) tyrosine kinase using two different strategies namely a dialdehyde-based cross-linking and photo-affinity labeling. The probes were designed from the core structure of the Abl-selective inhibitor ImatinibTM. The dialdehyde-based probe, although found to be useful in detecting pure Abl kinase was found to have poor labeling performances in the presence of competing cellular proteins. A photo-affinity based probe (A/BP) with a fluorescent reporter tag was found to have more severe limitations in labeling performances even with pure proteins. This was attributed to the presence of the bulky reporter tag in the probe which might have disrupted the proper binding interactions of the probe with the active site of the kinase. On the other hand, a clickable “tag-free” version of the A/BP was found to have superior labeling performances both with pure protein and in the presence of competing cellular proteins. The probe also showed promising use for potential *in situ* screening of Abl inhibitors.

6. 1 Introduction

Active site-directed probes capable of detecting functionally active form of a target protein/protein class have been developed for many proteins. The ABPP approach allows one to directly interrogate active enzymes present in complex proteomes using the ABPs. Protein kinases present significant challenge to the development of kinase-selective ABPs partly because these classes of enzymes carry out catalysis with only transient and non-covalent interactions with the substrates and partly because of the absence of highly reactive residues near the active site of most of the kinases. In spite of these limitations, probes for the kinase RSK and EGFR, both based on irreversible

inhibitor scaffolds targeting a nucleophilic cysteine residue near the active sites of these kinases, have been reported.¹⁰⁹ This approach, however, is not general as many protein kinases, including clinically highly relevant kinases such as Abl, do not possess such a nucleophilic cysteine residue near the active site. Targeting the generally conserved catalytic lysine residue of kinases is another option. There are a handful of both selective and broad-spectrum kinase probes based on this approach. For example, the fungal metabolite Wortmannin has been identified as covalent inhibitor targeting the catalytic lysine residue of phosphoinositol-3-kinase (PI3K) family and based on its irreversible mode of inhibition, biotin or fluorophore conjugated Wortmannin analogues have been developed as probes for PI3Ks and mammalian polo-like kinase-1.¹¹⁰ Fluorosulfonylbenzoyl adenosine (FSBA) is a small molecule known to target the catalytic lysine of kinases and biotin conjugates of FSBA have been developed as general probes for protein kinases.¹¹¹ Acyl phosphate-based promiscuous protein kinase probes targeting the catalytic lysine residue have been recently reported by Kozarich et al.¹¹² Shokat and co-workers utilized the same lysine residue in a kinase to cross-link kinase/pseudopeptide substrate complex using an orthophthaldehyde (OPA)-derivatized ATP analogue.¹⁰⁵ We further improved this method by replacing the OPA portion with naphthalene-2,3-dicarboxaldehyde (NDA), resulting in a cross-linker with improved kinase specificity (discussed in chapter-5).¹⁰⁸ Although the original motivation with these dialdehyde-based cross-linkers was to identify unknown kinases from putative phosphopeptides/phosphoproteins, we reasoned the same strategy could be used to develop activity-based probes for a specific kinase, for instance Abl kinase, if the dialdehyde cross-linker could be made selective towards the kinase of interest.

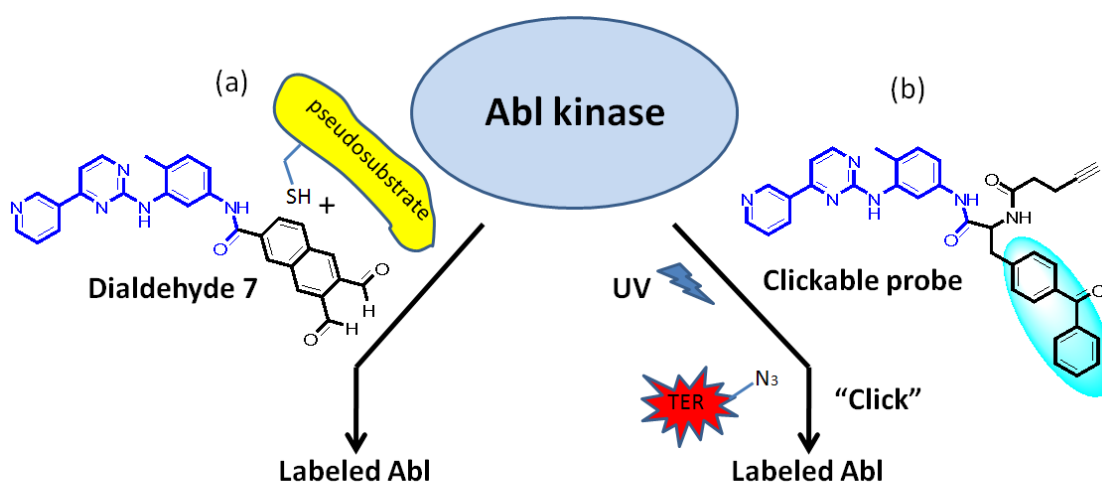


Figure 6. 1 Two strategies to develop Abl-selective probes. (a) A three-component (kinase, pseudosubstrate and dialdehyde) reaction mediated labeling of Abl by the dialdehyde **7**. (b) Clickable photo-affinity probe (Compound **6-13**) mediated labeling of Abl.

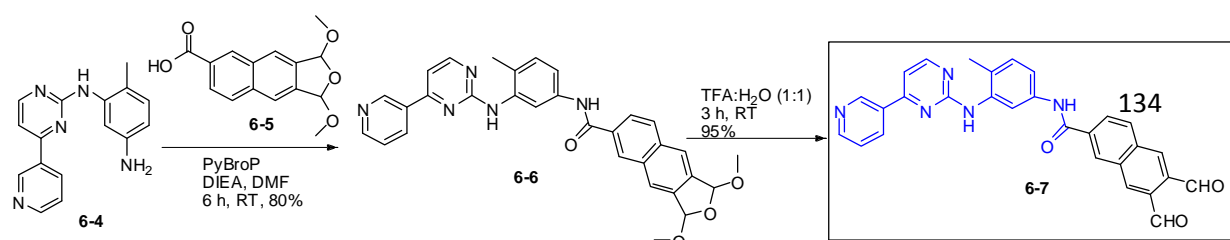
Affinity-based probes (A/BPs) offer another potential strategy for targeting kinases that lacks nucleophilic residues near the catalytic site. Such probes have been developed for various classes of enzymes including kinases.¹⁴ An affinity-based probe essentially has three structural components: 1) a protein recognition unit; usually derived from a potent reversible inhibitor scaffold 2) a photoreactive group; which could be benzophenone, alkyl or aryl diazirine or aryl azides and 3) a reporter unit; which is typically a fluorescent dye or biotin. We hypothesized that the remarkable selectivity of Imatinib for Abl could be taken as a guiding principle for developing Abl-selective small-molecule probes. This chapter reports the synthesis of such Abl-selective, active-site directed probes that utilize both design principles (i. e. dialdehyde-based cross-linking and photo-cross-linking) (see Figure 6.1). At the heart of the probe designs we used the Imatinib-core structure, essentially to ensure sufficient binding interaction of the probes with the active-site of the Abl kinase. At

first, the dialdehyde-based cross-linker (compound **6-7** in Scheme 6.1) was synthesized. It has two key structural elements; the Imatinib-core structure (for kinase recognition) and the NDA reactive unit (for mechanism-based cross-linking of the kinase with pseudo-substrates). The choice of NDA as the reactive unit was based on our previous results (chapter-5), where we observed improved kinase labeling with the NDA-reactive unit compared to OPA-reactive unit. The proposed mechanism of labeling of Abl kinase using this dialdehyde-based cross-linker is shown in Figure 6.2a. Briefly, the compound **6-7** first occupies the ATP-binding pocket (shaded) of the kinase by exploiting the high binding affinity and selectivity of the imatinib core structure (blue). The catalytic lysine residue from the kinase active site forms a reversible imine adduct by reacting with the formyl group in **6-7**. The cysteine residue from the pseudosubstrate attacks the imino carbon with the subsequent removal of a water molecule, leading to the formation of a stable isoindole product. The cross-linked Abl-pseudosubstrate pair is subsequently visualized by in-gel fluorescent scanning. For the development of photo-affinity-based probes we first synthesized compound **6-12**, with three distinctive structural units; once again the Imatinib core-structure for selectivity, a benzophenone-based photo-reactive group for light-mediated covalent cross-linking of the probe with the kinase and a TER-based fluorescent dye for visualization of the labeled protein with in-gel fluorescence scanning after protein-separation using SDS-PAGE. Compound **6-12**, however, exhibited very poor labeling performance which is presumably due to the presence of the bulky rhodamine dye unit in the probe which might have caused severe disruption in the probe's active-site recognition of the kinase. So we designed a "label-free", clickable version of the photo-affinity probe (Compound **6-13**) which was equipped with a flexible alkyne handle for click-conjugation with an azide-functionalized

rhodamine fluorophore following the actual covalent labeling of the kinase with the probe. This compound exhibited superior labeling performances, both with purified protein and in the presence of large excess of competing cellular proteins. The proposed mechanism of labeling of Abl kinase using the clickable cross-linker is shown in Figure 6.2b. Briefly, upon incubation with the kinase, compound **6-13** occupies the active site by virtue of the imatinib core structure. Subsequent UV irradiation generates a highly reactive diradical intermediate from the benzophenone unit of **6-13**, which quickly reacts with C-H bonds in the vicinity, leading to covalent binding of **6-13** with the kinase. The probe/kinase complex upon treatment with a rhodamine azide reporter (i.e. TER-N₃) under click chemistry conditions enables the fluorescence visualization of the labeled kinase with in-gel fluorescence scanning.

6. 2 Synthesis of the Probes

The structures of the probes are shown in Scheme 6.1 below. The dialdehyde **6-7** was synthesized first. Briefly, the key intermediate **6-4** was synthesized as previously reported with minor modifications.⁹⁴ Compound **6-4** was subsequently coupled with 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylic acid, **6-5**, to give **6-6**. The cyclic acetal protecting group in **6-6** was subsequently removed by treatment with TFA in water to afford the target compound **6-7** in 28.9% yield over six steps. For the synthesis of the affinity-based probes **6-12** and **6-13** (a clickable version of **6-12**; *vide infra*), Boc-protected benzophenone, **6-8**, was first coupled to **6-4**, yielding compound **6-9**. Subsequent deprotection with TFA, followed by coupling with either the rhodamine dye **6-11** or pent-4-ynoic acid, gave the two probes, **6-12** and **6-13** in 11.6% and 26.9% overall yield, respectively.



Scheme 6. 1 Synthesis of the small-molecule probes of Abl kinase

6. 3 Results and Discussions

The proposed mechanism of labeling of Abl kinase using the two different probes is shown below in Figure 6.2.

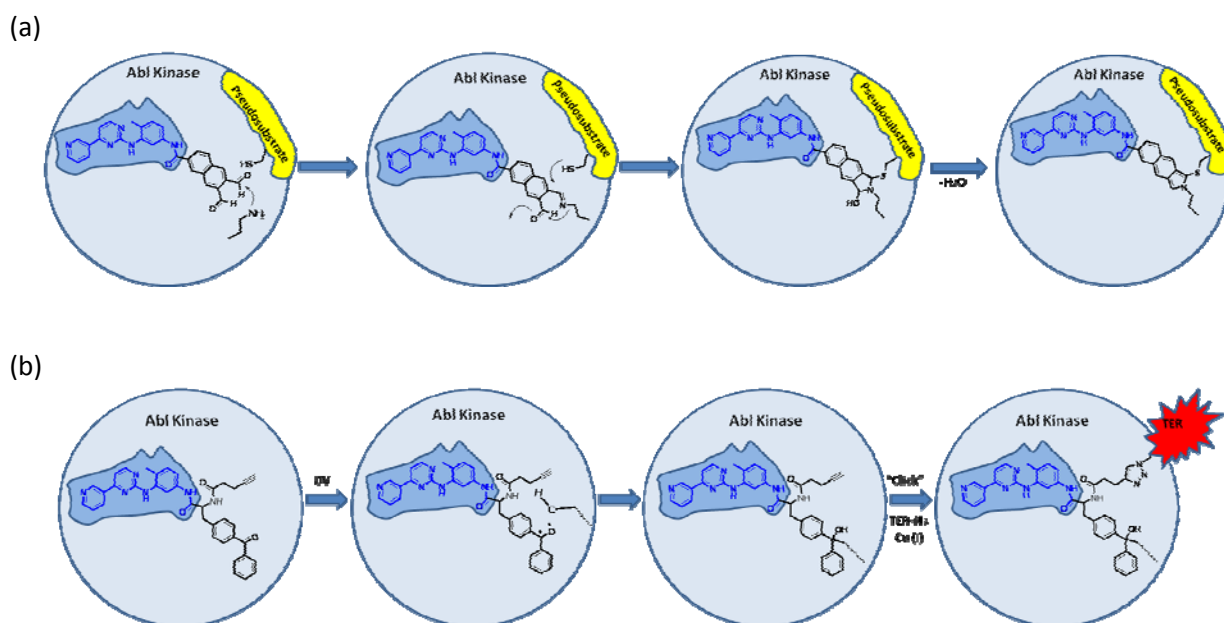


Figure 6. 2 Two different methods to label Abl kinase. (a) A three-component reaction (kinase + pseudosubstrate + dialdehyde **6-7**) mediated labeling of Abl kinase. (b) Clickable photo-affinity probe (compound **6-13**) mediated labeling of Abl.

6. 3. 1 Labeling Experiments with the Dialdehyde **7**

6. 3. 1. 1 Labeling Experiments using pure kinases and kinase spiked in cellular lysates

To evaluate the dialdehyde **7** (compound **6-7**) as an activity-based probe for Abl kinase, a three-component reaction was initiated by incubating the kinase and the dialdehyde **7** with an Abl-selective peptide pseudosubstrate Fluoresceine-K(biotin)-EAICAAPFAKKK. The peptide pseudosubstrate was equipped with a cysteine residue (underlined in the sequences), making the peptide suitable for the three- component reaction. Typical labeling reactions with the dialdehyde **7** were

carried out in the following optimized conditions: kinase (400 to 600 nM), pseudosubstrate (2 μ M) and dialdehyde **7** (10 μ M) in the reaction buffer (25 mM HEPES at pH = 7.5, 150 mM NaCl, 2 mM MgCl₂) were incubated for 20 min at RT before adding 6X loading dye (reducing). The sample was heated at 95°C for 10 min before SDS-PAGE and in-gel fluorescence scanning (excitation at 480 nm, emission collected at 520 nm filter). As shown in Figure 6.3a, intense fluorescent labeling of the Abl kinase was observed. Unlike Abl kinase, only a weak labeling was observed from the three-component reaction of Csk kinase with the dialdehyde **7** and the Csk pseudosubstrate, Fluoresceine-K(biotin)-KKKKEEICFFF (i.e. Csktide), thus demonstrating the Abl selective nature of the dialdehyde probe. It was also observed that, unlike the previous generations of kinase-directed dialdehyde cross-linkers, labeling of the Abl kinase with **6-7** showed little preference for the amino acid sequence in the peptide pseudosubstrates (see lanes 1 & 2). This lack of substrate selectivity could be due to the high binding affinity of the dialdehyde **7** towards Abl kinase (IC₅₀ = 194 nM) as a result of its imatinib core. Thus, unlike the previous dialdehyde-based kinase/substrate crosslinking strategy, which targets a kinase with the selectivity of the pseudopeptide substrate, the new dialdehyde probe (compound **6-7**) labels Abl kinase by taking advantage of the strong binding interaction of its kinase-binding scaffold. Next, we asked if the probe could selectively label Abl in the presence of competing cellular proteins. But unfortunately in a spike experiment with varying amounts of CHO-K1 cell lysate, the labeling was severely affected even in the presence of as little as 1 μ g of the mammalian proteome (Figure 6.3b).

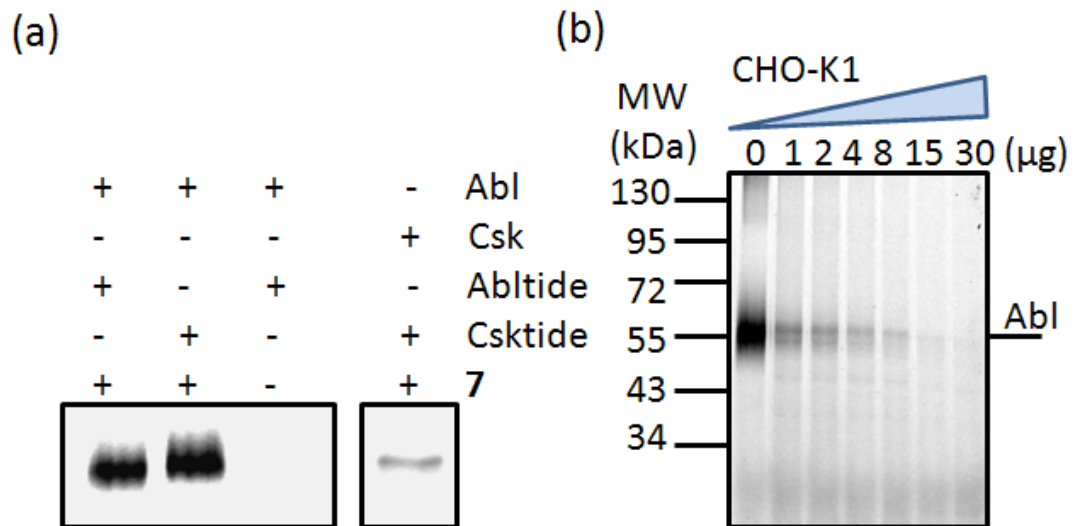


Figure 6. 3 Fluorescence scanned gels showing the labeling of Abl and Csk kinases (500 nM each) with the **6-7** (10 μM) in the presence of the pseudosubstrates (2 μM). (b) Labeling of Abl (500 nM) with **6-7** in the presence of increasing amounts of the CHO-K1 proteome.

6. 3. 1. 2 pH-dependence of the cross-linking reaction

HEPES buffer in the pH range of 6 to 9 were used to determine the pH dependence of the labeling reaction of Abl kinase with the dialdehyde **7**. As shown below, efficient labeling was observed at pH > 7 while a low pH was found to be unfavourable for the reaction.

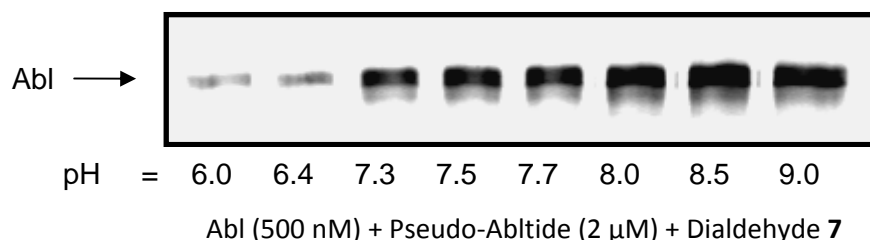


Figure 6. 4 pH-dependence of the labeling profiles of the **6-7** mediated cross-linking of Abl kinase with the pseudosubstrate

6. 3. 1. 3 Effect of exogenous thiols on the efficiency of the labeling

In order to determine the exogenous thiol tolerance of the dialdehyde-guided cross-linking reaction of the kinase with the thiol peptide pseudosubstrate, reactions were done at varying concentrations of an exogenous competing thiol, β -mercaptoethanol. Briefly, 500 nM of Abl kinase was incubated in the buffer (25 mM HEPES at pH = 7.5, 150 mM NaCl, 2 mM MgCl₂) with varying amounts of β -mercaptoethanol (BME). To this mixture was added the dialdehyde **7** (10 μ M) and the Abl-pseudosubstrate (2 μ M). The reactions were incubated for 20 min at RT and subsequently denatured and subjected to SDS-PAGE and in-gel fluorescence scanning. As shown in Figure 6.5, 5 to 10 fold excess of the exogenous thiol was found to have no significant effect in the cross-linking efficiency while higher concentrations (> 50 fold excess) of the thiol was found to significantly shield the kinase from cross-linking with the substrate peptide.

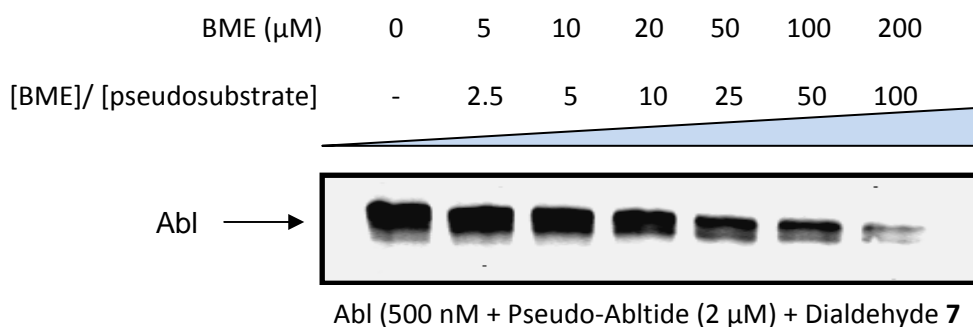


Figure 6. 5 Effect of β -mercaptoethanol (BME) on the three-component reaction of Abl kinase with compound **6-7** and Pseudo-Abltide.

6. 3. 1. 4 Effect of exogenous amines on the efficiency of the labeling

Since the labeling reaction was based on the selectivity of the dialdehyde **7** towards Abl kinase, we tested if the reaction is tolerated in the presence of exogenous amines (e.g. lysine). Thus cross-linking reactions were set up in the presence of varying amounts of lysine. As shown in Figure 6.6, even 800 fold excess of lysine was found to have no significant effect in the labeling efficiency. This indicates that the labeling reaction is selective towards the kinase and is highly shielded from exogenous amines.

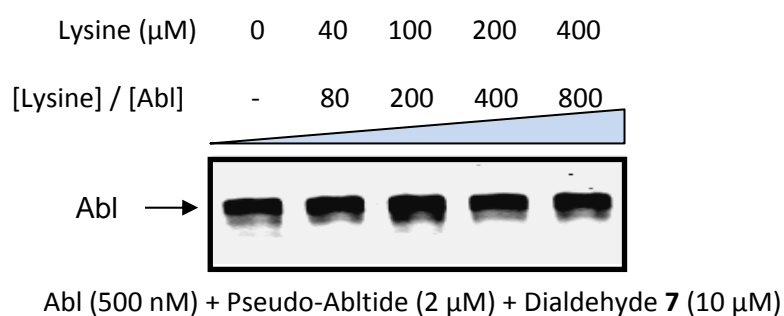


Figure 6. 6 Effect of exogenous lysine on the labeling reaction.

6. 3. 1. 5 IC_{50} evaluation of the probe

Concentration-dependent experiments were performed to determine the inhibition potency and hence the binding affinity of the probes towards the Abl kinase. The inhibition assay was performed with Kinase-Glo[®] Plus Luminescent Kinase assay kit from Promega following the manufactures instructions. Briefly, recombinant Abl kinase, a consensus Abl peptide substrate with sequence KKGEAIYAAPFA-NH₂, ATP and the probe were mixed in the kinase reaction buffer (100 mM Tris, pH = 7.5, 10 mM MgCl₂) at a volume of 55 μL in a flat-bottom solid white 96-well plate. The incubation was allowed to continue for 20 min at 37 °C and the kinase 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 μ M and 50 μ M, 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.

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

$$\text{Potency} = 1 - \text{Remaining activity}$$

Dose-dependent inhibition assays were performed by varying the concentration of the probes under fixed enzyme concentration of 50 nM. The IC₅₀ values of the probes were calculated from the percentage activity vs. log [concentration of probe] curves generated using GraphPad Prism software. The dialdehyde **7** was found to be a potent inhibitor of the Abl kinase as evident from its IC₅₀ value of 194 nM for Abl kinase inhibition (Figure 6.7).

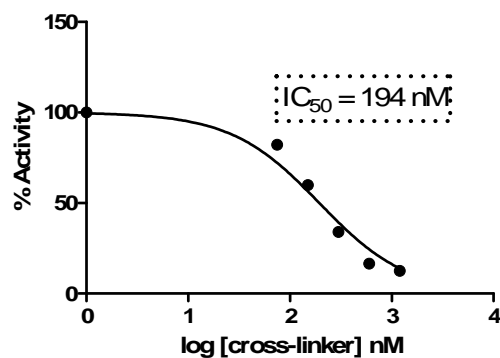


Figure 6. 7 IC₅₀ evaluation of compound **6-7** for Abl kinase inhibition

6. 3. 2 Labeling experiments with the photo-cross-linkers

6. 3. 2. 1 Comparative labeling experiments

The photo-cross-linkers **6-12** and **6-13** were tested for their ability to selectively label both recombinantly purified Abl kinase and Abl spiked in mammalian proteome (Figure 6.8). The general procedure for the labeling studies with the photo-affinity probe-1 (compound **6-12**) is as follows; 2 μL of the enzyme stock solutions (0.4 $\mu\text{g}/\mu\text{L}$) were diluted with 17.6 μL of Tris.HCl buffer (50 mM, pH = 7.5). Different concentrations of **6-12** (stock solutions in DMSO) were added (0.4 μL) and the reactions were incubated at room temperature in the dark for 30 min. Subsequently the reaction mixtures were irradiated with a handheld UV lamp under the long UV channel for 20 min. The reactions were quenched by the addition of 4 μL of 6X SDS loading buffer followed by boiling at 95 °C for 10 min. Samples were resolved on a 10% denaturing SDS-PAGE gel and fluorescence was detected with a fluorescence gel scanner (Typhoon 9200, Amersham). The general procedure for the photo cross-linking with subsequent click-chemistry (for probe **6-13**) is as follows; 2 μL of the enzyme stock solutions (0.1 to 0.5 $\mu\text{g}/\mu\text{L}$) were diluted with 17.6 μL of Tris.HCl buffer (50 mM, pH = 7.5). The clickable probe,

6-13, (stock solution in DMSO) was added (0.4 μL) and the reactions were incubated at room temperature in the dark for 30 min. Subsequently the reaction mixtures were irradiated with a handheld UV lamp under the long UV channel (~ 350 nm) for 20 min in the absence of other light sources. The reaction mixtures were then subjected to click-chemistry with a TER-azide dye in the following conditions. To each reaction was added 1 μL each of TER- N_3 fluorescent dye (concentration in reaction = 50 μM , stock solution in DMSO), CuSO_4 (concentration in reaction = 1 mM, stock solution in water), TBTA (concentration in reaction = 100 μM , stock solution in DMSO) and sodium ascorbate (concentration in reaction = 1 mM, stock solution in water) and incubated for 2 h at room temperature. The reactions were quenched by the addition of 5 μL of 6X SDS loading buffer followed by boiling at 95 $^\circ\text{C}$ for 10 min. Samples were analyzed on a 10% denaturing SDS-PAGE gel and fluorescence was detected with the fluorescence gel scanner.

Compound **6-12** had rather poor labeling performance and a very high concentration of the probe (up to 100 μM) was required in order for sufficient Abl labeling to be observed (see lanes 1 & 2, Figure 6.8a). Furthermore, the probe was found to nonspecifically label other proteins as well (see a labeled band of YOP, i.e. a phosphatase). We speculated that the poor performance of this probe was mostly due to the presence of the bulky rhodamine dye which might have disrupted the binding of **6-12** to the active site of Abl. On the other hand we were delighted to find that the “clickable” probe, **6-13**, performed much better at selective labeling of Abl kinase. **6-13** was equipped with an alkyne handle in place of the rhodamine. Upon labeling with Abl kinase by UV irradiation, the probe/kinase complex was subsequently reacted, via the well known click

chemistry, with a rhodamine azide reporter (i.e. TER-N₃ in Scheme 6.1), before visualization with in-gel fluorescence scanning. This two-step, sequential labeling approach had previously been shown to minimize the adverse effects caused by introduction of bulky reporter dyes in other ABPs. As shown in Figure 6.8a, a low concentration of **6-13** (0.5 μM; lane 3) was sufficient to selectively label Abl kinase. The contaminant YOP in the same reaction was not labeled. Other proteins including kinases c-Src and ERK-2 were also tested for labeling with probe **6-13** (Figure 6.8b); only Abl kinase was positively labeled, again indicating the selectivity of the probe towards Abl. We next tested the labeling of Abl spiked in a complex mammalian proteome with **6-13**. As shown in Figure 6.8c, although the Abl labeling was diminished in the presence of CHO-K1 proteome, less than 1% of Abl present in the total proteome lysate could be detected (see lanes 5 & 6). Finally, we performed labeling experiments in the presence of active-site competing inhibitor Staurosporine. Dose-dependent reduction in Abl labeling in the presence of the inhibitor (Figure 6.8d) was observed, indicating the labeling is activity-based and pointing to the possible use of the probe for *in situ* screening of Abl inhibitors.

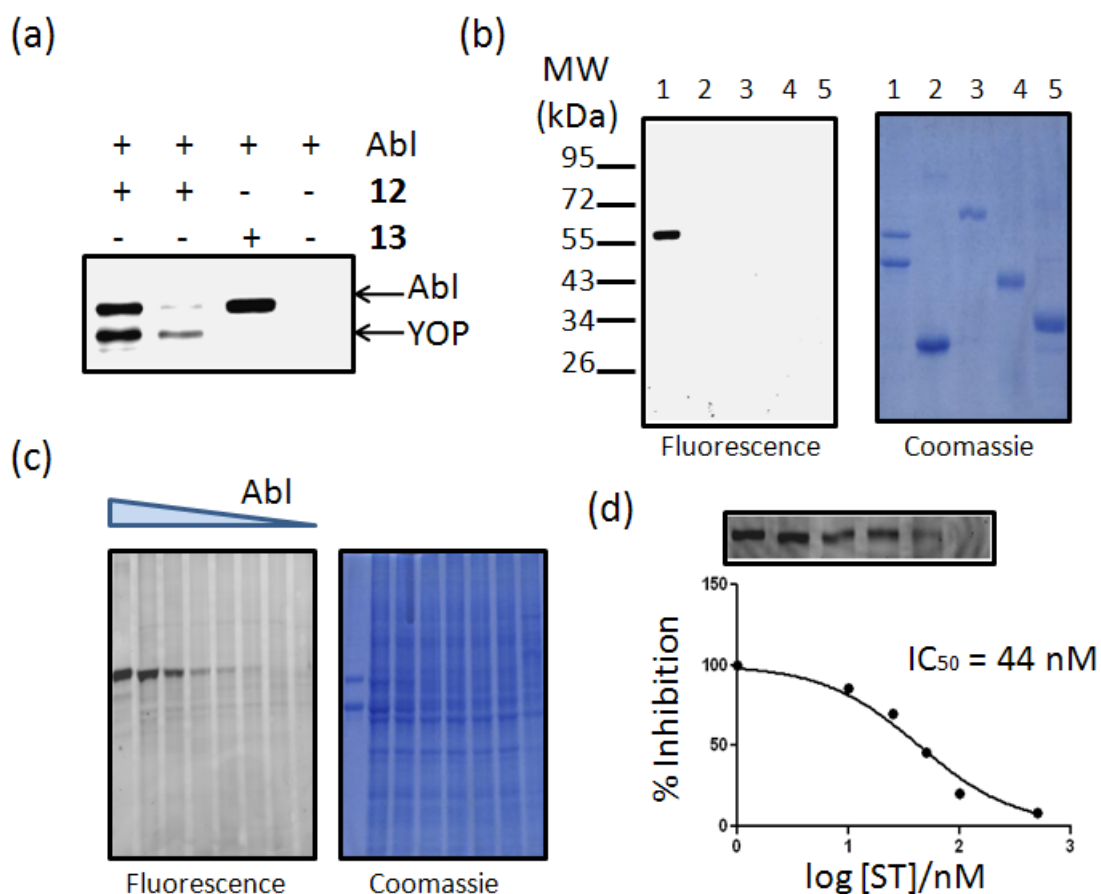


Figure 6. 8 (a) Fluorescence gels of recombinant Abl (600 nM) labeled with **6-12** (lanes 1 & 2) or **6-13** (lane 3). The concentrations of **6-12** in lanes 1 & 2 were 100 μ M & 20 μ M, respectively. The concentration of **6-13** in lane 3 was 0.5 μ M. The contaminant YOP was also labeled by **6-12** in lanes 1 & 2 as well. The two-step sequential click chemistry was initiated by adding (to the UV-irradiated protein sample) TER-N₃, Cu²⁺, TBTA and sodium ascorbate for 2 h before SDS-PAGE and fluorescence scanning. (b) Different proteins (lanes 1 to 5 are Abl, PTPB, BSA, ERK-2 and c-Src, respectively) were labeled with **6-13** (0.5 μ M). Only Abl was selectively labeled by the probe (see left gel). (c) 30 μ g of CHO-K1 proteome was spiked with different amounts of Abl kinase (giving the final % of Abl to be 100 (1000 nM), 3.3 (1000 nM), 2.6 (800 nM), 1.3 (400 nM), 0.65 (200 nM), 0.325 (100 nM), 0.162 (50 nM) and 0 (0 nM) from lane 1 to 8 respectively), then labeled with **6-13** (1 μ M). (d) Dose-dependant reduction of labeling of Abl (600 nM) with **6-13** (1 μ M) in the presence of Staurosporine (ST). IC₅₀ of ST was determined to be ~44 nM.

6. 3. 2. 2 Detection limit of the pure Abl with the photo-cross-linker **6-13**.

In order to determine the fluorescent detection limit of the Abl kinase by the clickable photo-affinity probe-2, photo-cross-linking reactions were set up as mentioned before with decreasing amounts of Abl. As shown in Figure 6.9 below, 0.3 to 0.2 μg of pure Abl kinase was readily detected with the probe at a concentration of 0.5 μM .

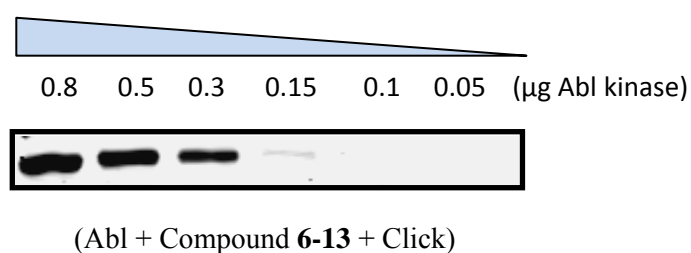


Figure 6. 9 Labeling of decreasing amounts of pure Abl kinase with 0.5 μM clickable probe (**6-13**)

6. 3. 2. 3 Labeling experiments with the clickable probe (**6-13**) in the presence of K562 cell lysate

Human leukemic cell line K562 is a p-210 Bcr-Abl expressing (Chronic Myelogenous Leukemia) cell line. Total cell lysates were prepared and used for labeling reactions with the clickable probe, **6-13**. Neither endogenous Bcr-Abl (MW = 210 kDa) nor endogenous c-Abl (MW = 130 kDa) were detected from the lysate with our probe. This is presumably due to the insufficient detection limit of the probe as well as due to the insufficient expression levels of these proteins in the cell lines. So we carried out labeling experiments from this mammalian proteome with varying amounts of spiked Abl kinase as follows. Decreasing amounts of recombinant Abl kinase was spiked to 30 μg of K562 total cell lysate (50 mM Tris buffer, $\text{pH} = 7.5$, 150 mM NaCl). The cell lysates were prepared from approximately 1.5×10^7 cells cultured in T75 cell culture flasks. To 30 μg of the lysate was added 0.5 μM clickable

probe (**6-13**) and the reaction was allowed to continue for 20 min. Subsequently the reactions were irradiated with UV light (~ 350 nm) for 20 min followed by click-chemistry with rhodamine azide (TER-N₃) as described above. After 2 h of click reaction, 6X SDS loading dye were added, heated to 95 °C for 10 min and the proteins were resolved by SDS-PAGE. In-gel fluorescent scanning was used to visualize the labeled protein bands. As shown below in Figure 6.10, up to ~ 1% of the Abl kinase in the total proteome could be readily detected using the probe.

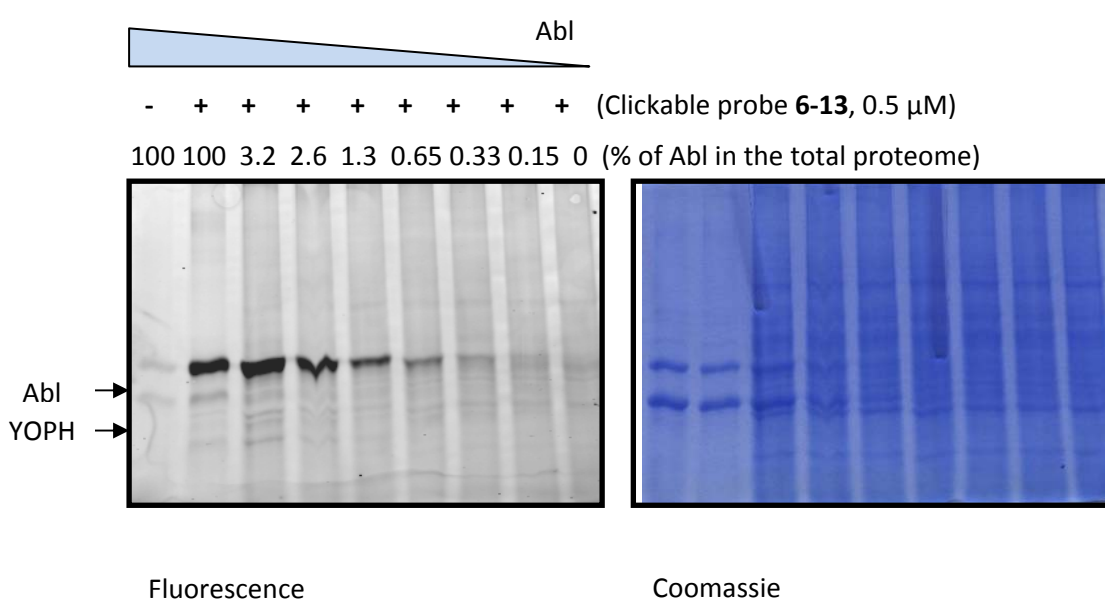


Figure 6. 10 Labeling of decreasing amounts of spiked Abl kinase in the presence of K562 cell lysate (30 μg total proteins per lane).

6. 4 Conclusions

In conclusion, two different strategies for active site-directed, selective labeling of Abl kinase have been developed. These probes are the first activity-based probes for Abl kinase. One of the probes, compound **6-13** (an AfBP), showed promising results for proteomic profiling of Abl and for potential *in situ* screening of Abl inhibitors, though further improvements need to be made in order for it to be more useful in a complex biological system. Ultimately, new chemical proteomic tools will help us unravel the complex signalling cascades of this extremely important kinase and will shed more light into the role of individual kinases in critical cellular events.

6. 5 Synthetic details and characterizations of compounds

1,3-Dimethoxy-N-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenyl)-1,3-dihydronaphtho[2,3-c]furan-6-carboxamide (6-6):

To a solution of 1,3-dimethoxy-1,3-dihydronaphtho[2,3-c]furan-6-carboxylic acid (**6-5**) (0.05 g, 0.180 mmol, 1 eq.) in 10 mL of DMF was added bromo-tris-pyrrolidino phosphonium hexafluorophosphate (PyBrOP) (0.126 g, 0.27 mmol, 1.5 eq.) and diisopropylethylamine (DIEA) (0.048 mL, 0.27 mmol, 1.5 eq.) and the mixture was stirred for 10 min. Compound **6-4** (0.05 g, 0.180 mmol, 1 equiv.) was subsequently added, and the mixture was stirred at room temperature overnight. The solvent was then removed *in vacuo* and the residue dissolved in ethyl acetate and washed with water and brine. The organic layer was dried with anhydrous Na₂SO₄ and concentrated. Purification by column chromatography afforded the pure product as a yellow powder (0.077 g, 80%). LC-MS: *m/z* found = 534.15 [M + H]⁺ calculated = 534.21

6,7-Diformyl-N-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenyl)-2-naphthamide (6-7):

Compound **6-6** (0.077 g, 0.144 mmol) was dissolved in a mixture of TFA and H₂O (1:1; total of 10 mL) and stirred at room temperature for 3 h, after which the solvent was removed under reduced pressure. Traces of TFA was removed by the addition of methylene chloride and then removing the solvent under reduced pressure, which was repeated for several times to yield the product as an orange powder (0.066 g, yield = 95%) which was further purified by reverse phase semi-prep HPLC using water and acetonitrile as eluents. ¹H-NMR (300 MHz, CDCl₃) δ 10.55 (s, 2H), 9.40 (s, 1H), 9.08

(s, 1H), 8.87 (s, 1H), 8.87 – 8.82 (m, 2H), 8.75 – 8.72 (m, 2H), 8.68 (s, 1H), 8.57 (d, $J = 5.1$ Hz, 1H), 8.40 – 8.26 (m, 2H), 8.20 – 8.15 (m, 1H), 7.77 – 7.68 (m, 1H), 7.56 – 7.50 (m, 2H), 7.26 (d, $J = 8.4$ Hz, 1H), 2.26 (s, 3H). ESI-MS: m/z found = 488.14 [$M + H$]⁺ calculated = 488.16.

3-(4-Benzoylphenyl)-2-(tert-butoxycarbonylamino) propanoic acid (6-8):

To 0.269 g (1 mmol) of commercially available 2-amino-3-(4-benzoylphenyl) propanoic acid (1 mmol) in 1:1 dioxane/water (5 mL each) was added 0.24 g of di-*tert*-butyl-dicarbonate (1.1 mmol) and 0.197 mL of triethylamine. The reaction was stirred for 4 h. The reaction mixture was then concentrated on a rotary evaporator and the residue was diluted with water and ethyl acetate. The aqueous layer was acidified to pH 1 with 1 N HCl and back extracted with ethyl acetate. The organic layer was washed with brine, dried over Na₂SO₄ and evaporated to give 0.332 g (yield = 90%) of the protected amino acid (compound **6-8**) as an off-white solid. The compound was used in the subsequent steps without further purification.

***Tert*-butyl 3-(4-benzoylphenyl)-1-(4-methyl-3-(4-(pyridin-3-yl) pyrimidin-2-ylamino) phenylamino)-1-oxopropan-2-ylcarbamate (6-9):**

A 50 mL round-bottomed flask was charged with 0.29 g of **6-8** (0.785 mmol), 0.328 g of HATU (0.863 mmol), 0.162 mL of DIEA and 15 mL of DMF. After stirring for 5 min, 0.217 g of **6-4** (0.785 mmol) was added and the reaction was allowed to run overnight. The solvent was removed under reduced pressure and the crude residue was purified by column chromatography on silica gel using hexane and ethyl acetate as eluents to give 0.4 g of the product **6-9** (yield = 81%) as a yellow solid, and use directly in next steps.

2-Amino-3-(4-benzoylphenyl)-N-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino) phenyl) propanamide (6-10):

Compound **6-9** (0.33 g, 0.525 mmol) was dissolved in a 1:1 mixture of TFA and DCM with 1% water (5 mL total volume) and stirred at room temperature in an open round-bottomed flask. The reaction was completed in 30 min and the solvents were removed under reduced pressure to give 0.33 g of the product **6-10** (yield = 92%) which was used in the subsequent steps without further purification. ESI-MS: m/z found = 529.19 $[M + H]^+$ calculated = 529.22

N-(9-(2-(4-(4-(3-(4-Benzoylphenyl)-1-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenylamino)-1-oxopropan-2-ylamino)-4-oxobutanoyl)piperazine-1-carbonyl)phenyl)-6-(diethylamino)-3H-xanthen-3-ylidene)-N-ethylethanaminium (6-12), Photo-Affinity Probe 1:

A 25 mL round-bottomed flask was charged with 0.115 g of rhodamine B 4-(3-carboxypropionyl) piperazine amide (**6-11**) (0.189 mmol), 0.079 g of HATU (0.206 mmol), 0.045 mL of DIEA (0.258 mmol) and 5 mL of DMF. After stirring for 5 min, 0.11 g of **6-10** (0.172 mmol) was added and the stirring was continued for 3 h upon which all of the starting material **6-10** was used up. The solvent was removed on a rotary evaporator and the crude residue was purified by Semi-Prep HPLC on a C₁₈ column, giving 0.08 g (yield = 41%) of the product **6-12** as a dark fine powder. ESI-MS: m/z found = 561.257 = $(M+2H^+)/2$, $[M]^+$ calculated = 1121.53

N-(3-(4-Benzoylphenyl)-1-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenylamino)-1-oxopropan-2-yl) pent-4-ynamide (Compound 6-13), Clickable Photo-Affinity Probe:

A 25 mL round-bottomed flask was charged with 0.018 g of pent-4-ynoic acid (0.189 mmol), 0.079 g of HATU (0.206 mmol), 0.045 mL of DIEA (0.258 mmol) and 5 mL of DMF. After stirring for 5 min, 0.11 g of **6-10** (0.172 mmol) was added and the stirring was continued for 1 h upon which all of the starting material **6-10** was used up. The solvent was removed on a rotary evaporator and the crude residue was purified by column chromatography on silica gel using hexane and ethyl acetate as eluents to give 0.1 g of the product **6-13** (yield = 95%) as off-white solid. ¹H NMR (300 MHz, DMSO-d₆) δ = 2.19 (s, 3H), 2.29 (m, 4H), 2.70 (s, 1H), 2.94-3.18 (m, 2H), 4.77 (m, 1H), 7.15 (d, 1H, *J* = 8.4 Hz), 7.26 (d, 1H, *J* = 8.22 Hz), 7.42 (d, 1H, *J* = 8.4), 7.46 (d, 2H, *J* = 8.22 Hz), 7.53 (m, 3H), 7.66 (m, 5H), 7.92 (s, 1H), 8.40 (d, 1H, *J* = 8.22 Hz), 8.45 (d, 1H, *J* = 8.04 Hz), 8.50 (d, 1H, *J* = 5.10 Hz), 8.67 (d, 1H, *J* = 5.10 Hz), 8.97 (s, 1H), 9.3 (s, 1H), 10.1 (s, 1H). ¹³C NMR (75 MHz, DMSO-d₆) δ = 14.10, 17.59, 33.98, 37.91, 54.46, 71.22, 83.68, 107.73, 115.87, 116.16, 124.72, 127.27, 128.53, 129.43, 129.49, 129.57, 130.22, 132.57, 133.04, 135.21, 136.67, 137.23, 137.73, 142.91, 149.42, 158.21, 159.58, 160.85, 160.98, 169.63, 170.45, 195.57. ESI-MS: *m/z* found = 609.24, [M+H]⁺ calculated = 609.25.

Chapter 7: Future directions

Summary

This chapter provides a brief outlook to some of the future developments possible in line with the ABPP- and inhibitor-developments of PKs and PTPs discussed in the previous chapters of this dissertation. The PTP-reactive unnatural amino acids (2-FMPT and caged-2-FMPT) could be incorporated into full-length substrate proteins of PTPs using the well-known Expressed Protein Ligation (EPL) method and the resultant protein-based probes may provide useful in the identification of the upstream/downstream PTP/PTPs responsible for a specific dephosphorylation event of the substrate protein. The construction of such a probe involves the generation of a truncated protein with a C-terminal thioester and the unnatural amino acid incorporated peptide-ligation partner with an N-terminal cysteine residue. The key concepts and the preliminary results for the synthesis of the peptides are discussed in this chapter. The second part of this chapter discuss the possibility of developing AfBPs and bidentate inhibitors for a subset of the protein kinase family with a compact threonine gatekeeper residue. The synthesis of a key inhibitory scaffold, equipped with a bulky *p*-tolyl substituent on a pyrrolopyrimidine unit and the potential possibilities of the strategy are discussed.

7. 1 Protein-based PTP probes to validate/identify the PTP/PTPs responsible for dephosphorylating a given substrate protein

As the metabolism and homeostasis of cells heavily relies on protein-protein interactions, several techniques have been developed to identify and characterize such interactions.¹¹³ Many of these techniques such as the yeast two-hybrid assays,

glutathione-s-transferase fusion proteins, co-immunoprecipitation, and phage-display methods have become routine in today's proteomic research. These techniques find significant use in phosphoproteomic research as well. However, protein kinases and protein phosphatases present significant challenges to such general techniques of mapping-out and characterizations of protein-interaction networks. In particular, many protein kinases and protein phosphatases, although display relatively tolerated substrate selectivities *in vitro* with synthetic peptide substrates, typically show high levels of substrate specificities in the living cell with their physiological protein substrate partner/partners and this property, in the light of the highly intricate nature of the signaling networks associated with these two classes of proteins, makes it more meaningful to characterize their physiological substrates rather than identifying/characterizing peptide substrates *in vitro*. In addition to contributions from sub-cellular localization and associated accessibility/proximity-driven interactions (also protein co-localization and associated variations in effective concentrations of interacting partners), long-range interactions or interactions that are distal from the site of phosphorylation/dephosphorylation are thought to play critical roles in determining the high degree of substrate specificities observed with protein partners *in vivo* for these classes of enzymes.¹¹³

As the catalysis via protein kinases lead to the incorporation of a phosphate unit on to the substrate protein, *in vivo* identification of a kinase substrate is, at least conceptually, relatively easy with standard techniques based on radioactivity or chemiluminescence. The same goal in the cases of phosphatases is much more challenging as there is no direct method for the detection of the removal of a phosphate group from an already phosphorylated protein substrate *in vivo*.

However, the recent development of an ingenious approach known as “substrate-trapping mutants” makes this highly challenging task possible at least in the case of several protein tyrosine phosphatases (PTPs).¹¹⁴ But the inverse problem, that of identifying the PTP responsible for dephosphorylating a particular substrate protein remains unsolved. We hypothesized that a “protein-based PTP-probe” could potentially address this problem as such a probe, by virtue of the presence of the full protein sequence could exploit most of the binding interactions between the original native substrate protein and the corresponding PTP/PTPs.

As a proof-of-concept, we decided to choose the membrane-associated non-receptor protein tyrosine kinase, c-Src, for making the protein-based PTP-probe. The structural domain of c-Src kinase consists of an N-terminal myristoyl group, one SH3 domain, one SH2 domain and a catalytic domain (Figure 7.1).¹¹⁵ In the cell the c-Src kinase is normally maintained in an inactive state via an intramolecular interaction between its SH2 domain and a phosphotyrosine residue 527 near the C-terminal side.¹¹⁵ The intracellular phosphorylation of this Tyr-527 is known to be catalyzed by another protein kinase called Csk.¹¹⁶ This negative regulation of c-Src kinase activity is of high importance as the aberrant activity of this kinase has been identified in a variety of cancer cells.¹¹⁷ Once activated the c-Src kinase could trigger the phosphorylation of an array of intracellular substrates with significant implications in several downstream signaling pathways affecting cell-division, cell-differentiation and cell-mobility. Although aberrant activity of this kinase is implicated in several cancers, its transient activation has been observed during important cellular events such as mitosis and growth factor receptor activation.¹¹⁸ The tyrosine phosphatase PTP1B has been shown to preferentially dephosphorylate the phospho-Tyr-527 in c-Src *in vitro*.¹¹⁹

Furthermore, elevated levels of PTP1B expression has been observed in several cancer cell lines with increased c-Src activity.¹¹⁹ These lines of evidences point to the potential roles of PTP1B in positively regulating the activity of endogenous c-Src along with several other activity-regulating mechanisms.

One could use the Expressed Protein Ligation (EPL) strategy¹²⁰ to construct the protein-based PTP-probe using a c-Src construct. The overall scheme for constructing such a protein-based PTP-probe is shown below (Scheme 7.1). Briefly, amino acid residues 83 to 524 of chicken c-Src (K295M) with an N-terminal His₆ tag and a C-terminal thioester moiety could be generated via intein-mediated protein splicing.¹²¹ The protein with the C-terminal thioester could undergo a Native Chemical Ligation (NCL)¹²² with the peptide bearing an N-terminal cysteine residue. In our probe design, this peptide would carry the PTP-reactive unnatural amino acid (indicated **X** in Scheme 7.1) so that the ligated protein could act as a PTP-probe.

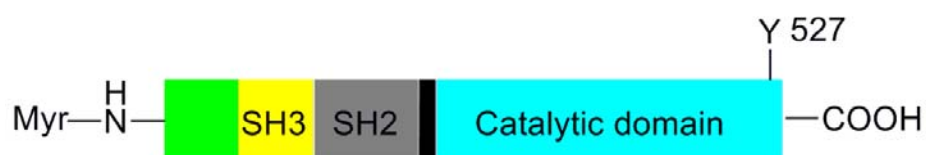
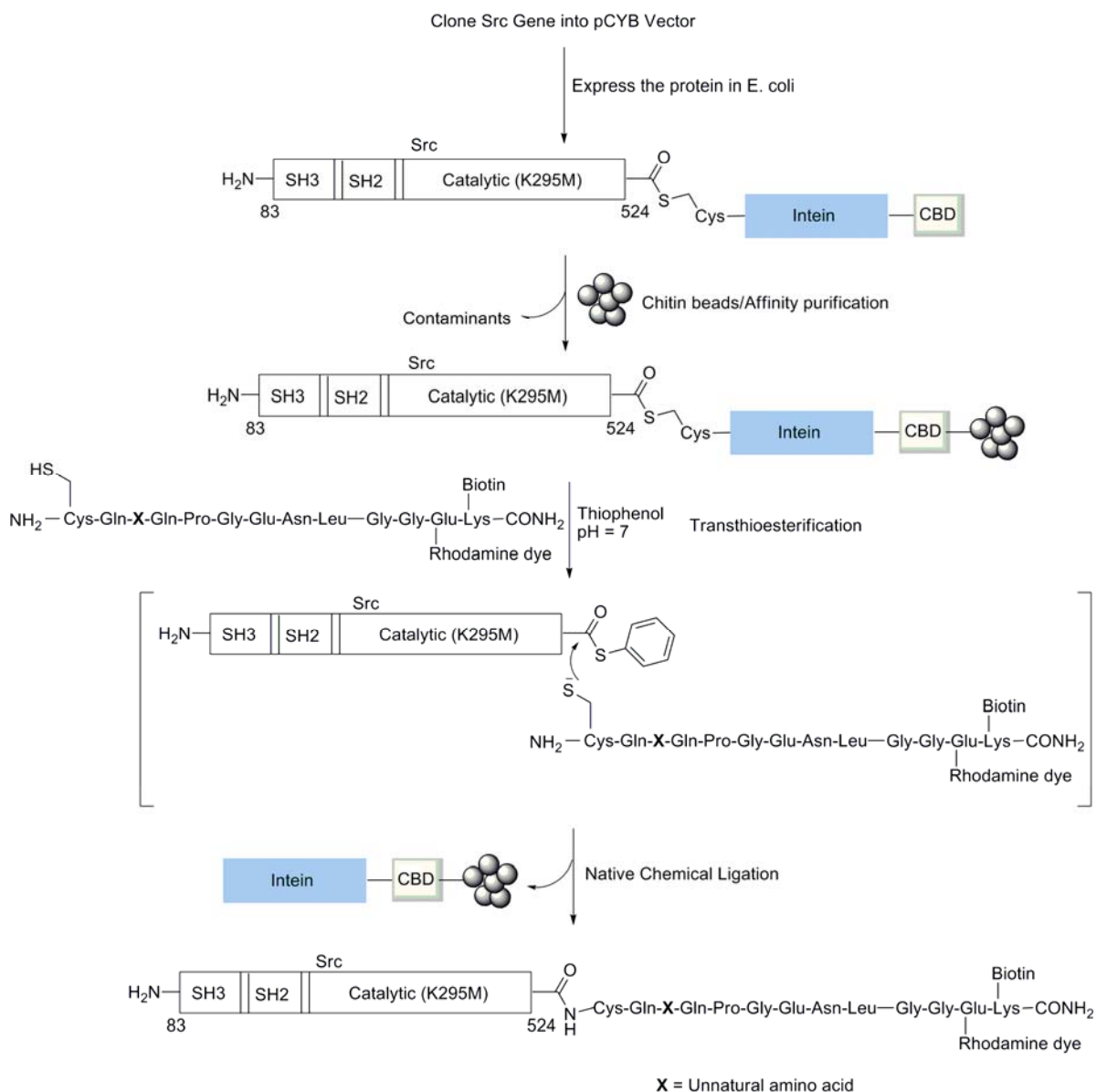


Figure 7. 1 Domain structure of c-Src kinase



Scheme 7. 1 Scheme for constructing the protein-based PTP-probe using Expressed Protein Ligation

We decided to make two different versions of the probes namely a PTP-trapping probe and a PTP-sensor. The PTP-trapping probe (designated as pTRAP) would be equipped with the unnatural amino acid 2-FMPT for covalently reacting with the PTP whereas the PTP-sensor probe¹²³ (designated as ppCAP) would be equipped with an unnatural amino acid pCAP which could report the PTP activity via fluorescence emission. The peptide ligation partner of the probes (Figure 7.2)

carry four distinctive structural units. They are 1. N-terminal cysteine residue (for ligating to the c-Src protein with C-terminal thioester), 2. The unnatural amino acid (for PTP-recognition), 3. A fluorescent dye (for visualization of the trapped PTP) and 4. A biotin (for affinity enrichments/purifications of labeled protein complex).

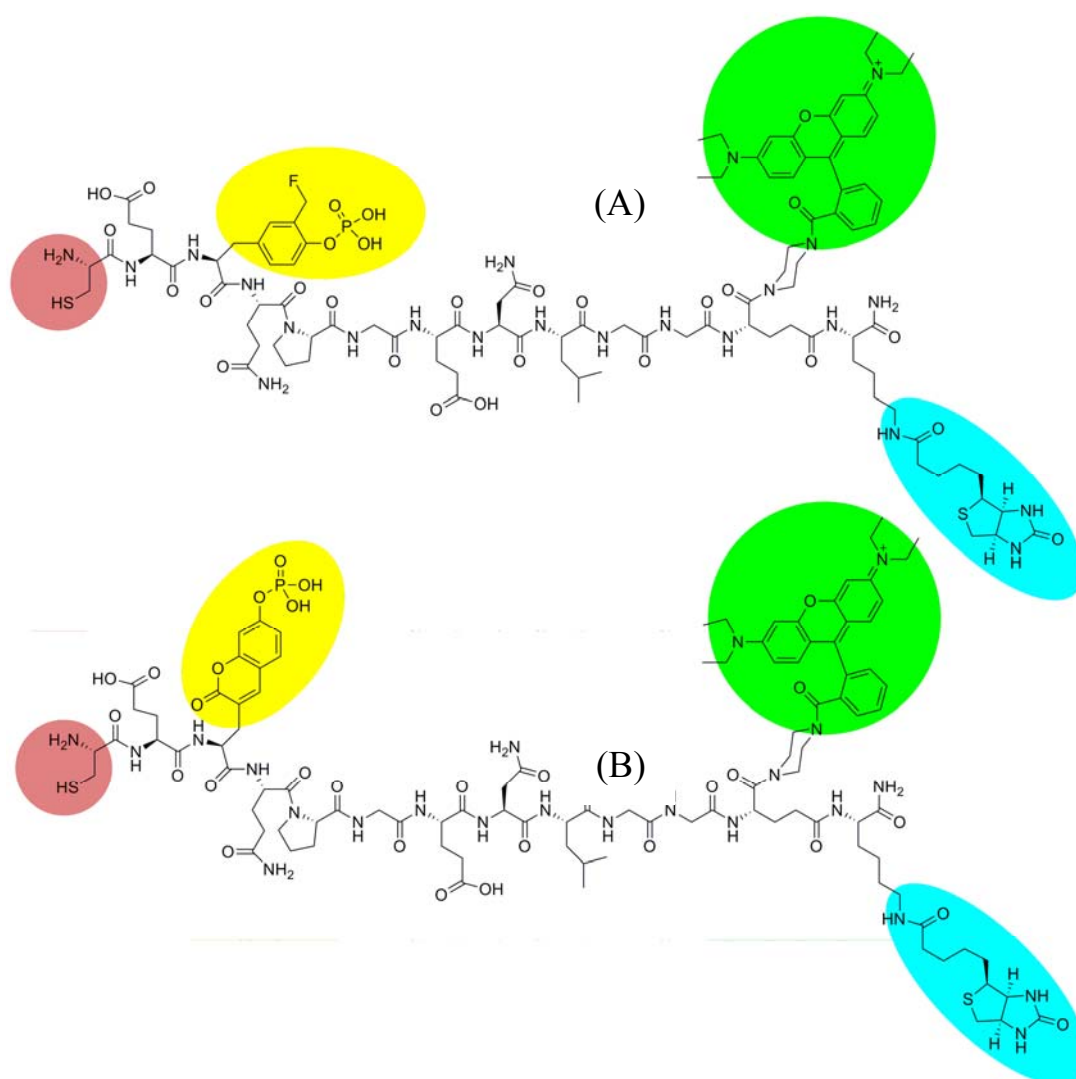
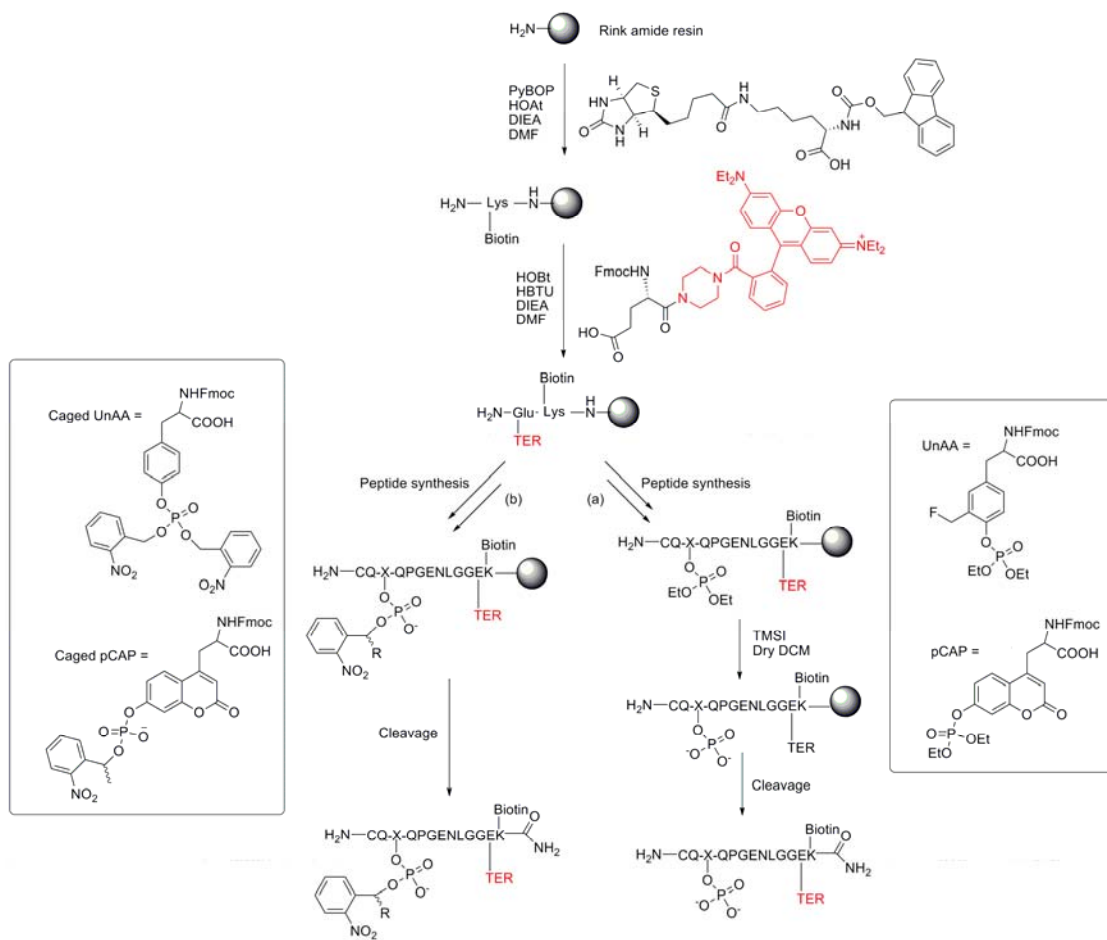


Figure 7. 2 The peptide ligation partners for the construction of the protein-based PTP-probes. (A) the peptide for pTRAP probe and (B) the peptide for the ppCAP probe. The four key structural units in the peptides are highlighted with colors



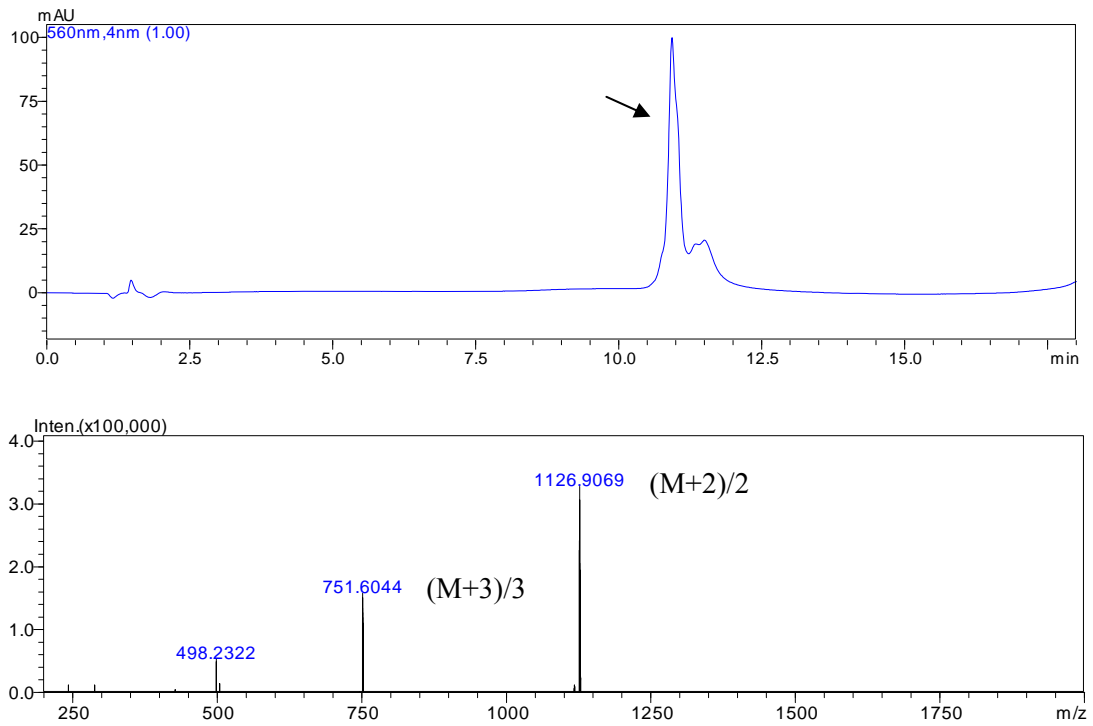
Scheme 7.2 Solid-phase synthesis of the peptide ligation-partners.

The peptides were synthesized following standard Solid-Phase Peptide Synthesis (SPPS) protocols using rink-amide resin and Fmoc-protected amino acids (Scheme 7.2a). The crude peptides were purified by Prep-HPLC and characterized by LC-MS (Figure 7.3). Further improvements in the synthesis may be possible using the caged versions of the unnatural amino acids (caged-2-FMPT discussed in chapter-3, and caged-pCAP) as the use of such caged amino acids eliminate the need of harsh deprotection condition of ethyl ester protection of phosphate moiety using TMSI (Scheme 7.2b).

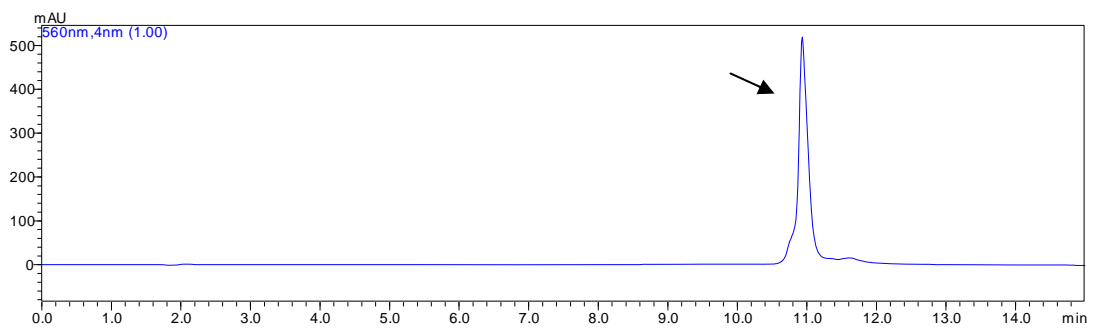
The future goals include validation of the protein-based probe concept with the

ultimate aim of using the probes to perform live-cell imaging and related experiments to identify the PTP/PTPs responsible for dephosphorylating a given protein substrate of interest.

(A) pTRAP peptide



(B) pCAP peptide



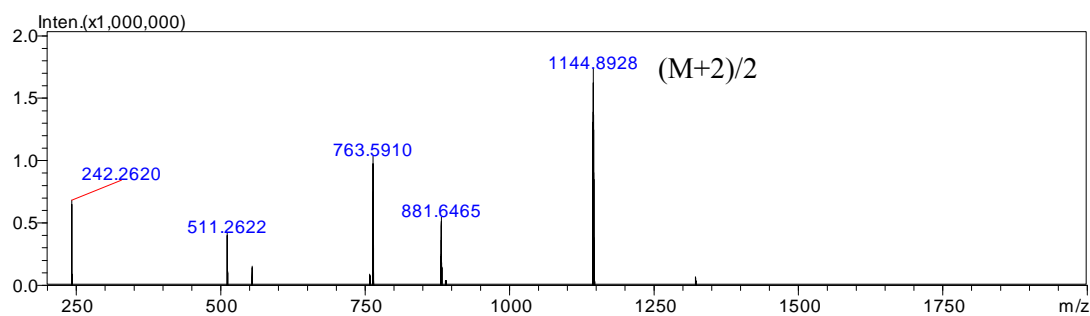


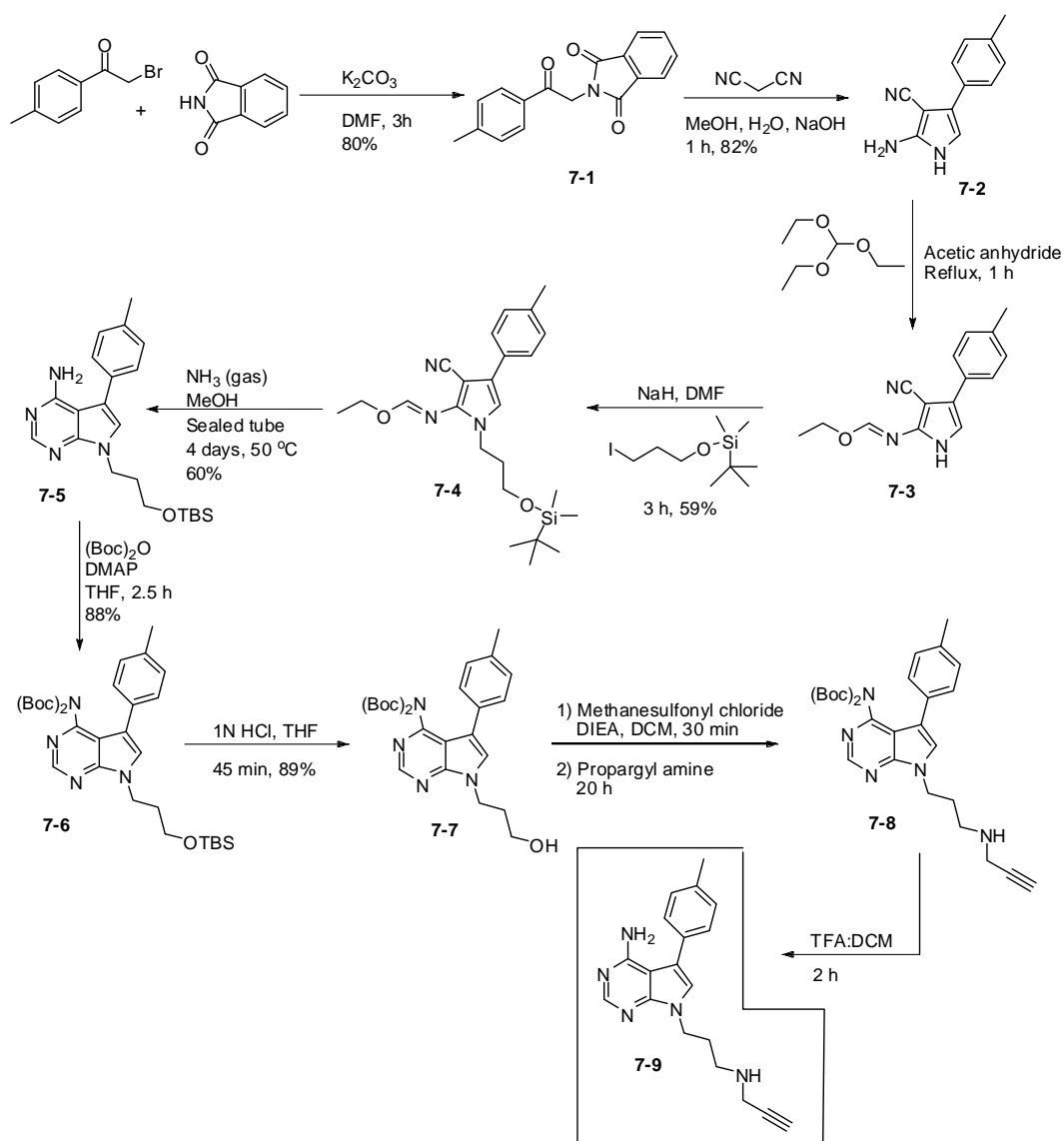
Figure 7. 3 LC-MS profiles of the purified (A) pTRAP and (B) pCAP peptides. LC-running condition: C-18 reverse phase column with 20 to 100% ACN (80 to 20% water) with 0.1% TFA for 15 min.

7. 2 Synthesis of a scaffold for the development of affinity-based probes (A/BPs) and bidentate inhibitors of protein kinases with a compact gatekeeper residue

Structural studies of kinases have revealed a “gatekeeper” residue, near a conserved $\beta 5$ strand, at the opening of the active-site and close to the hinge region connecting the N- and C-terminal domains of the protein.¹²⁴ The size of this gatekeeper residue has been shown to be the major determinant for selective accessibility of several small-molecule inhibitors to the active site of different kinases.¹²⁵ A compact gatekeeper residue such as threonine or glycine would allow inhibitors with large aromatic substituents whereas bulky gatekeepers such as phenyl alanine, leucine or isoleucine would restrict their access. It has been observed that approximately 20% of protein kinases possess threonine as the gatekeeper residue¹²⁶ and many kinase inhibitors exploit this compact residue as selectivity filter.

In line with the development of the kinase-selective inhibitors and ABPs

discussed in chapters 4 and 6 respectively, we sought to develop A/BPs and bidentate inhibitors for protein kinases with a compact gatekeeper residue. Towards this we first synthesized a key inhibitory scaffold (compound **7-9** in Scheme 7-3) which was equipped with a bulky *p*-tolyl substituent on the pyrrolopyrimidine unit for the potential exploitation of the compact threonine gatekeeper residues in certain kinases such as c-Src, Rsk, c-Abl, EGFR, HER, Fyn, Lck, c-Kit, p38 etc. The design of the scaffold was based on a highly selective and potent clickable irreversible inhibitor, fmk-pa, of the C-terminal domain of the p90 ribosomal protein S6 kinase (Rsk).^{109a} Compound **7-9** could be clicked with various azides to generate a library of potential bidentate kinase inhibitors targeting some of the members in the 20% of kinases possessing compact gatekeeper residue and potentially the structural variants from the azide components in the inhibitors may exploit differences in secondary-binding pockets adjacent to the ATP-binding pocket in those kinases, and this in turn may lead to more selective and potent inhibitors of such kinases. Structural variants of compound **7-9** with different linker length for the alkyne handle and different bulky aromatic substituents at the position of the *p*-tolyl substituent on the pyrrolopyrimidine unit could also be synthesized to explore more possibilities. Compound **7-9** could also be used for developing A/BPs for kinases via attaching photo-reactive units such as benzophenone or aryl diazirine and a fluorescent reporter tag.



Scheme 7. 3 Synthesis of a clickable scaffold (compound **7-9**) for the potential development of AfBPs and bidentate inhibitors of protein kinases with a compact gatekeeper residue

Synthetic procedures and characterizations of compounds in the Scheme 7.3

2-(2-Oxo-2-*p*-tolylethyl)isoindoline-1,3-dione (7-1):

To a solution of 3.79 g (25.81 mmol, 1.1 eq.) of phthalimide in 35 mL DMF was added 3.88 g (28.16 mmol, 1.2 eq.) of K₂CO₃ at RT and the solution was stirred for 45 min. To this was added 2-bromo-1-*p*-tolylethanone (5 g, 23.47 mmol, 1 eq.) and stirred at RT for 3h. The reaction was then poured into 750 mL water upon which a white precipitate was formed. The solution was filtered and the residue was purified by recrystallization from DCM:Ether to afford 5.3 g of compound **7-1** (yield = 80%) as a white solid. ESI-MS: $m/z = 280.1$ [M+H]⁺

2-Amino-4-*p*-tolyl-1H-pyrrole-3-carbonitrile (7-2):

Compound **7-2** was prepared as previously reported. Briefly, to a solution of malononitrile (0.461 g, 6.98 mmol, 1.3 eq.) in 6:1:1 MeOH:48% aq. NaOH:H₂O was added 1.5 g (5.37 mmol, 1 eq.) of compound **7-1** and the mixture was stirred at RT for 1h upon which the product was precipitated out of the solution. The reaction mixture was filtered and the residue was washed with H₂O, DCM and hexane to afford 0.87 g (yield = 82%) of **7-2** as an off-white flaky solid. ¹H-NMR (300 MHz, CD₃OD) δ 7.49 (d, 2H, $J = 7.6$ Hz), 7.08 (d, 2H, $J = 7.6$ Hz), 6.36 (s, 1H), 4.83 (br s, 2H), 2.27 (s, 3H). ESI-MS: $m/z = 198.1$ [M+H]⁺

Ethyl N-3-cyano-4-*p*-tolyl-1H-pyrrol-2-ylformimidate (7-3):

To a solution of compound **7-2** (0.5 g, 2.53 mmol) in triethyl orthoformate (5 mL) was added 70 μ L of acetic anhydride and the reaction mixture was refluxed for 1h. The reaction mixture was then cooled to RT and the solvent was removed under

reduced pressure. The crude pyrrole iminoether was azeotropically dried with toluene and used in the subsequent step without further purification.

Ethyl N-1-(3-(*tert*-butyldimethylsilyloxy)propyl)-3-cyano-4-*p*-tolyl-1H-pyrrol-2-ylformimidate (7-4):

To a suspension of NaH (0.056 g, 2.34 mmol) in 5 mL dry DMF was added ~ 2.34 mmol of the crude pyrrole iminoether (compound **7-3**) in 3 mL dry DMF and the solution was stirred at RT for 30 min. To this was then added 0.7 g (2.34 mmol) of 3-(*t*-butyldimethylsilyloxy)propyl iodide over 10 min and the reaction was stirred for 3h. The solvent was then removed under reduced pressure and the residue was purified by flash column chromatography on silica gel to afford 0.59 g (yield = 59%) of compound **7-4** as a brown oil. ¹H-NMR (300 MHz, CDCl₃) δ 8.50 (s, 1H), 7.55 (d, 2H, *J* = 8.04 Hz), 7.22 (d, 2H, *J* = 8.04 Hz), 6.69 (s, 1H), 4.40 (q, 2H, *J* = 7.23 Hz), 4.02 (t, 2H, *J* = 7.0 Hz), 3.65 (t, 2H, *J* = 5.76 Hz), 2.4 (s, 3H), 1.97-1.87 (m, 2H), 1.42 (t, 3H), 0.97 (s, 9H), 0.09 (s, 6H). ESI-MS: *m/z* = 426.7 [M+H]⁺

7-(3-(*tert*-Butyldimethylsilyloxy)propyl)-5-*p*-tolyl-7H-pyrrolo[2,3-*d*]pyrimidin-4-amine (7-5):

Compound **7-4** (0.32 g, 0.75 mmol) was dissolved in 10 mL MeOH and transferred to a 50 mL sealed-tube reaction vessel with a magnetic stirring bar. Argon gas was bubbled through the solution for 10 min. The tube was immersed in a cold bath of dry ice-MeOH mixture and NH₃ gas was bubbled through the solution for 15 min. The tube was quickly sealed with a Teflon screw cap and the reaction was allowed to warm to room temperature. The reaction was stirred at 50 °C for 4 days. The sealed tube was subsequently cooled on an ice bath and the NH₃ gas was slowly released. The solvent was removed under reduced pressure and the crude product was purified

by flash column chromatography to yield 0.18 g (yield = 60%) of compound **7-5**. $^1\text{H-NMR}$ (300 MHz, CDCl_3) δ 8.30 (s, 1H), 7.38 (d, 2H, $J = 7.9$ Hz), 7.25 (d, 2H, $J = 8.04$ Hz), 6.95 (s, 1H), 5.67 (br s, 2H), 4.33 (t, 2H, $J = 6.8$ Hz), 3.66 (t, 2H, $J = 5.8$ Hz), 2.40 (s, 3H), 2.14-2.04 (m, 2H), 0.92 (s, 9H), 0.06 (s, 6H). $^{13}\text{C-NMR}$ (75 MHz, CDCl_3) δ 157.9, 152.2, 150.9, 137.2, 132.5, 130.2, 129.2, 123.6, 116.4, 101.7, 60.3, 42.0, 33.5, 26.4, 21.6, 18.8, -4.8.

Compound 7-6:

To a solution of compound **7-5** (90 mg, 0.227 mmol) in 4 mL THF was added $(\text{Boc})_2\text{O}$ (0.149 g, 0.681 mmol, 3 eq.) and DMAP (5.6 mg, 0.045 mmol, 0.2 eq.) and the reaction was stirred at room temperature for 2.5 h. The solvent was evaporated off and the residue was dissolved in ethyl acetate, extracted with 0.5 N HCl, brine and dried over anhydrous Na_2SO_4 and concentrated. The residue was purified by flash column chromatography using 1:1 hexane:ethyl acetate pre-equilibrated column with fast elution to afford 0.12 g of (yield = 88%) compound **7-6**. $^1\text{H-NMR}$ (300 MHz, CDCl_3) δ 8.83 (s, 1H), 7.37 (d, 2H, $J = 7.89$ Hz), 7.30 (s, 1H), 7.20 (d, 2H, $J = 7.89$ Hz), 4.45 (t, 2H, $J = 6.9$ Hz), 3.67 (t, 2H, $J = 5.75$ Hz), 2.40 (s, 3H), 2.13-2.07 (m, 2H), 1.3 (s, 18H), 0.94 (s, 9H), 0.07 (s, 6H). $^{13}\text{C-NMR}$ (75 MHz, CDCl_3) δ 153.9, 151.8, 151.4, 151.2, 137.1, 131.1, 129.7, 129.1, 128.3, 116.5, 113.8, 83.5, 60.2, 42.3, 33.3, 28.2, 26.5, 21.7, 18.8, -4.8.

Compound 7-7:

Compound **7-6** (0.12 g, 0.201 mmol) was dissolved in a 3:1 (V/V) mixture of THF:1N HCl and stirred for 45 min at room temperature. The reaction mixture was then diluted with 40 mL ethyl acetate and washed with saturated NaHCO_3 followed by brine. The organic layer was dried with anhydrous Na_2SO_4 and concentrated. The

residue obtained was purified by flash column chromatography on silica gel with fast elution to afford 0.087 g (yield = 89%) of compound **7-7**. ¹H-NMR (300 MHz, CDCl₃) δ 8.80 (s, 1H), 7.36 (d, 2H, *J* = 7.89 Hz), 7.27 (s, 1H), 7.19 (d, 2H, *J* = 7.80 Hz), 4.45 (t, 2H, *J* = 6.30 Hz), 3.51 (t, 2H, *J* = 5.55 Hz), 2.39 (s, 3H), 2.06-2.02 (m, 2H), 1.29 (s, 18H). ¹³C-NMR (75 MHz, CDCl₃) δ 154.1, 152.2, 151.4, 151.1, 137.3, 130.7, 129.8, 129.1, 128.2, 117.1, 113.7, 83.7, 58.6, 41.7, 33.5, 28.2, 21.6.

Compound 7-8:

To a solution of 87 mg (0.18 mmol) of compound **7-7** in 3 mL DCM was added DIPEA (70 μL, 0.414 mmol, 2.3 eq.) and methanesulfonyl chloride (17 μL, 0.216 mmol, 1.2 eq.) at 0 °C. After 30 min the reaction mixture was diluted with ethyl acetate and extracted with saturated NH₄Cl solution followed by brine. The organic layer was dried over anhydrous Na₂SO₄, filtered and concentrated under reduced pressure. The crude mesylate was used for the reaction with propargyl amine without further purification.

To the crude mesylate was added 0.8 mL propargyl amine at room temperature and the reaction was stirred at 40 °C for 20h. The solvent was removed under reduced pressure and the crude product was attempted to purify by flash column chromatography on silica gel which could remove most of the impurities. The crude product **7-8** was used in the next step without further purification. ESI-MS *m/z* = 520.8 [M+H]⁺

Compound 7-9:

Compound **7-8** was dissolved in a 1:1 DCM:TFA mixture (1 mL each) and stirred for 2h at room temperature. Solvent was removed under reduced pressure and the crude

residue was purified by Prep-HPLC using reverse-phase C-18 column with a gradient of acetonitrile with 0.1% TFA and water with 0.1% TFA as mobile phases afforded approximately 5 mg of compound **7-9** as a pale brown oil. ¹H-NMR (300 MHz, CDCl₃) δ 10.66 (br s, 1H), 8.17 (s, 1H), 7.28 (m, 4H), 7.15 (s, 1H), 5.89 (br s, 2H), 4.41 (t, 2H, *J* = 6.40 Hz), 3.83 (s, 2H), 3.14 (t, 2H, *J* = 6.73 Hz), 2.5 (s, 1H), 2.40 (s, 3H), 2.39-2.33 (m, 2H). ¹³C-NMR (75 MHz, CDCl₃) δ 153.4, 148.3, 142.9, 139.5, 131.1, 129.4, 129.1, 125.3, 120.6, 100.5, 78.7, 73.2, 44.3, 42.9, 37.2, 27.4, 21.8.

Chapter 8: Concluding remarks

My doctoral dissertation has been focused on the development of certain chemical techniques that facilitate activity-based profiling and inhibitor developments of some important members in the protein tyrosine phosphatase (PTP) and protein kinase (PK) family. Reversible phosphorylation of proteins catalyzed by the opposing actions of these two classes of enzymes play key roles in the cellular signal transduction events and thus in turn in maintaining the normal cellular physiology. Conventional proteomic techniques, although provide informations regarding the expression levels and to some extent about the interaction networks of these proteins, are however, incapable of directly reporting the activities of these proteins and given the fact that the enzyme activity, and not the expression level, being the major determinant of the functional roles of these proteins, development of techniques that facilitate their activity-based profiling is of more significance in understanding their roles in the cellular functioning. Furthermore, being signaling enzymes, dysregulation of activities of many of these proteins have been identified as the root causes for several pathological conditions; hence development of selective and potent inhibitors of some of the members in these two classes of enzymes is another related area with direct relevance to the development of effective therapeutic agents.

As described in Chapter 2, we have designed a new “PTP-reactive unnatural amino acid”, 2-FMPT, which although exploits quinone methide chemistry similar to most of the existing PTP probes, has a very significant advantage that, being an amino acid, it could be easily incorporated into peptide sequences with essential substrate recognition elements at the proximal positions of *p*Tyr in a naturally occurring PTP substrate so that the corresponding peptide-based probes could

achieve target-specific profiling of different PTPs. A panel of peptide-based probes were synthesized and labeling reactions with the probes using purified and isolated proteins showed activity-based labeling specificity consistent with the known substrate preferences of the different PTPs. The strategy has also been found to be useful for efficient labeling reactions of PTPs from highly complex biological samples. Subsequently, a caged version of the unnatural amino acid with a photolabile *o*-nitrobenzyl group on the phosphate moiety (caged-2-FMPT) was synthesized and incorporated into peptides to generate peptide-based, caged, ABPs (Chapter 3). Using these caged probes, with PTP1B as a model system, the concept of photo-uncaging followed by activity-based labeling was validated.

In Chapter 4, we have demonstrated the potential of the “click chemistry” to rapidly assemble inhibitor libraries of the clinically highly relevant protein, Abelson (Abl) tyrosine kinase. Biochemical assays with the inhibitor library revealed a set of moderately potent and selective inhibitors of the Abl kinase. The *in vitro* hits identified from the library, however, did not show significant antiproliferative activities in cell-based assays. This was attributed mainly to the poor membrane permeability of the compounds and it is hoped that future designs that could account for this shortcomings may provide a highly useful and efficient strategy that could tackle most of the drug resistance problems observed in the treatment of Chronic Myelogenous Leukemia (CML) with Imatinib.

Similar to selective inhibition of a kinase of interest, selective detection of kinase-substrate interactions, in particular, detection of the upstream kinase responsible for a given phosphorylation event, remains a highly challenging task. Towards this, an improved mechanism-based kinase-peptide pseudosubstrate cross-linker, termed naphthalene 2,3-dicarboxaldehyde-adenosine (NDA-AD) has

been synthesized. The cross-linker, in addition to its improved labeling performances from crude proteomes compared to a previously reported cross-linker OPA-AD, was found to be suitable for the detection of kinase-pseudosubstrate interactions of both tyrosine-specific and serine/threonine-specific protein kinases. Further developments that allow cross-linking of protein substrates rather than peptide pseudosubstrates with full-length endogenous protein kinases in intact cells, would provide an invaluable tool in the study of signal transductions involving protein kinases.

The design and development of selective small molecule-based ABPs for the Abl kinase has been described in Chapter 6. Two different strategies namely a dialdehyde-based cross-linking and photo-affinity labeling were developed to selectively probe the Abl kinase. The dialdehyde-based probe, although useful for selective detection of purified Abl kinase, was found to have poor labeling performances in the presence of cellular lysates. On the other hand, a clickable “tag-free” version of the photo-affinity-based probe was found to have superior labeling performances both with purified protein and in the presence of competing cellular proteins.

It is hoped that the kinase- and phosphatase-directed ABP and inhibitor-development approaches described in this thesis, would provide useful chemical biology tools for the investigation of these extremely important signalling enzymes. They also provide a guideline for the future designs of more powerful methods such as the developments of protein-based probes, which could exploit not only the binding interactions near the active-site of these enzymes but also many distal binding interactions that facilitate more detailed investigations of these enzymes and their roles in critical cellular events.

Chapter 9: References

1. (a) C. M. Overall and O. Kleinfeld, *Nat. Rev. Cancer*, 2006, **6**, 227-239. (b) N. K. Tonks, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 833-846. (c) P. Cohen, *Nat. Rev. Drug Discovery*, 2002, **1**, 309-315. (d) B. Turk, *Nat. Rev. Drug Discovery*, 2006, **5**, 785-799.
2. (a) H. Sun, S. Chattopadhyaya, J. Wang and S. Q. Yao, *Anal. Bioanal. Chem.*, 2006, **386**, 416- 426. (b) M. Uttamchandani, C. H. S. Lu and S. Q. Yao, *Acc. Chem. Res.*, 2009, **42**, 1183-1192. (c) R. Srinivasan, J. Li, S. L. Ng, K. A. Kalesh and S. Q. Yao, *Nat. Protoc.*, 2007, **2**, 2655- 2664. (d) K. A. Kalesh, P.-Y. Yang, R. Srinivasan and S. Q. Yao, *QSAR Comb. Sci.*, 2007, **26**, 1135-1144.
3. K. E. Krueger and S. Srivastava, *Mol. Cell. Proteomics*, 2006, **5.10**, 1799- 1810.
4. (a) T. Hunter, *Cell*, 2000, **100**, 113-127. (b) J. A. Ubersax and J. E. Ferrell Jr, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 530-541.
5. A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik and J. R. Yates III, *Nat. Biotechnol.*, 1999, **17**, 676-682.
6. M. P. Washburn, D. Wolters and J. R. Yates III, *Nat. Biotechnol.*, 1999, **19**, 242-247.
7. S. P. Gygi, B. Rist, S. A. Gerbert, F. Turecek, M. H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, **17**, 994-999.
8. H. Zhou, J. A. Ranish, J. D. Watts and R. Aebersold, *Nat. Biotechnol.*, 2002, **20**, 512-515.

9. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann, *Mol. Cell. Proteomics*, 2002, **1**, 376-386.
10. P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson and D. J. Pappin, *Mol. Cell. Proteomics*, 2004, **3**, 1154-1169.
11. T. Ito, K. Ota, H. Kubota, Y. Yamaguchi, T. Chiba, K. Sakuraba and M. Yoshida, *Mol. Cell. Proteomics*, 2002, **1**, 561-566.
12. (a) M. J. Evans and B. F. Cravatt, *Chem. Rev.*, 2006, **106**, 3279-3301. (b) M. Uttamchandani, J. Li, H. Sun and S. Q. Yao, *ChemBioChem*, 2008, **9**, 667-675.
13. Y. Tanaka, M. R. Bond and J. J. Kohler, *Mol. BioSyst.*, 2008, **4**, 473-480.
14. (a) K. Liu, H. Shi, H. Xiao, A. G. L. Chong, X. Bi, Y. T. Chang, K. Tan, R. Y. Yada and S. Q. Yao, *Angew. Chem., Int. Ed.*, 2009, **48**, 8293-8297. (b) C. M. Salisbury and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1171-1176. (c) E. W. S. Chan, S. Chattopadhyaya, R. C. Panicker, X. Huang and S. Q. Yao, *J. Am. Chem. Soc.*, 2004, **126**, 14435-14446. (d) S. Chattopadhyaya, E. W. S. Chan and S. Q. Yao, *Tetrahedron Lett.*, 2005, **46**, 4053-4056. (e) K. R. Shreder, M. S. Wong, T. Nomanbhoy, P. S. Leventhal and S. R. Fuller, *Org. Lett.*, 2004, **6**, 3715-3718. (f) J. Wang, M. Uttamchandani, J. Li, M. Hu and S. Q. Yao, *Chem. Commun.*, 2006, 3783-3785. (g) H. Shi, K. Liu, A. Xu and S. Q. Yao, *Chem. Commun.*, 2009, 5030-5032. (h) M. Winnacker, S. Breeger, R. Strasser and T. Carell, *ChemBioChem*, 2009, **10**, 109-118. (i) Y.-M. Li, M. Xu, M.-T. Lai, Q. Huang, J. L. Castro, J. DiMuzio-Mower, T. Harrison, C. Lellis, A. Nadin, J. G. Neduvilil, R. B. Register,

- M. K. Sardana, M. S. Shearman, A. L. Smith, X.-P. Shi, K.-C. Yin, J. A. Shafer and S. J. Gardell, *Nature*, 2000, **405**, 689-694. (j) J. Chun, Y. I. Yin, G. Yang, L. Tarassishin and Y.-M. Li, *J. Org. Chem.*, 2004, **69**, 7344-7347.
15. (a) A. B. Berger, M. D. Witte, J. B. Denault, A. M. Sadaghiani, K. B. Sexton, G. S. Salvesen and M. Bogyo, *Mol. Cell.*, 2006, **23**, 509-521. (b) M. Bogyo, S. Shin, J. S. McMaster and H. L. Ploegh, *Chem. Biol.*, 1998, **5**, 307-320.
16. P. Van der Veken, E. H. Dirksen, E. Ruijter, R. C. Elgersma, A. J. Heck, D. T. Rijkers, M. Slijper and R. M. Liskamp, *ChemBioChem*, 2005, **6**, 2271-2280.
17. A. E. Speers and B. F. Cravatt, *J. Am. Chem. Soc.*, 2005, **127**, 10018-10019.
18. R. Orth and S. A. Sieber, *J. Org. Chem.*, 2009, **74**, 8476-8479.
19. (a) S. H. Verhelst, M. Fonovic and M. Bogyo, *Angew. Chem. Int. Ed.*, 2007, **46**, 1284-1286. (b) M. Fonoic, S. H. Verhelst, M. T. Sorum and M. Bogyo, *Mol. Cell. Proteomics*, 2007, **6**, 1761-1770.
20. (a) A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686-4687. (b) H. Ovaa, P. F. Van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebiger, A. M. C. H. van den Nieuwendijk, P. J. Galaray, G. A. van der Marcel, H. L. Ploegh and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2003, **42**, 3626-3629.
21. J. P. Alexander and B. F. Cravatt, *Chem. Biol.*, 2005, **12**, 1179-1187.
22. M. J. Evans, A. Saghatelian, E. J. Sorensen and B. F. Cravatt, *Nat. Biotechnol.*, 2005, **23**, 1303-1307.
23. A. T. Wright, J. D. Song and B. F. Cravatt, *J. Am. Chem. Soc.*, 2009, **131**, 10692-10700.

24. M. S. Cohen, H. Hadjivassiliou and J. Taunton, *Nat. Chem. Biol.*, 2007, **3**, 156-160.
25. L. Ballell, K. J. Alink, M. Slijper, C. Versluis, R. M. J. Liskamp and R. J. Pieters, *ChemBioChem*, 2005, **6**, 291-295.
26. L. Ballell, M. V. Scherpenzeel, K. Buchalova, R. M. J. Liskamp and R. J. Pieters, *Org. Biomol. Chem.*, 2006, **4**, 4387-4394.
27. S. A. Sieber, S. Niessen, H. S. Hoover and B. F. Cravatt, *Nat. Chem. Biol.*, 2006, **2**, 274-281.
28. (a) G. C. Adam, B. F. Cravatt and E. J. Sorensen, *Chem. Biol.*, 2001, **8**, 81-95. (b) G. C. Adam, E. J. Sorensen and B. F. Cravatt, *Nat. Biotechnol.*, 2002, **20**, 805-809.
29. K. T. Barglow and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 1523-1531.
30. (a) C. L. Sawyers, *Cancer Cell*, 2002, **1**, 13-15. (b) D. S. Krause, R. A. Van Etten, *N. Engl. J. Med.*, 2005, **353**, 172-187.
31. A. Alonso, J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon and T. Mustelin, *Cell*, 2004, **117**, 699-711.
32. (a) Z. -Y. Zhang, *Acc. Chem. Res.*, 2003, **36**, 385-392. (b) D. Barford, A. J. Flint and N. K. Tonks, *Science*, 1994, **263**, 1397-1404. (c) A. D. B. Pannifer, A. J. Flint, N. K. Tonks and D. Barford, *J. Biol. Chem.*, 1998, **273**, 10454-10462.
33. M. R. Arkin and J. A. Wells, *Nat. Rev. Drug Discovery*, 2004, **3**, 301-317.
34. B. G. Szczepankiewicz, G. Liu, P. J. Hajduk, C. Abad-Zapatero, Z. Pei, Z. Xin, T. H. Lubben, J. M. Trevillyan, M. A. Stashko, S. J. Ballaron, H. Liang, F. Huang, C.

- W. Hutchins, S. W. Fesik and M. R. Jirousek, *J. Am. Chem. Soc.*, 2003, **125**, 4087-4096.
35. D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 9367-9372.
36. H. C. Kolb and K. B. Sharpless, *Drug Discov. Today*, 2003, **8**, 1128-1137.
37. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.
38. R. Huisgen, in: A. Padwa (Ed.), *1,3-Dipolar Cycloaddition Chemistry*, Wiley, New York 1984, pp. 1-176.
39. (a) C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057-3064. (b) M. Meldal and C. W. Tornøe, *Chem. Rev.*, 2008, **108**, 2952-3015.
40. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596-2599.
41. (a) R. Srinivasan, M. Uttamchandani and S. Q. Yao, *Org. Lett.*, 2006, **8**, 713-716. (b) J. Xie and C. T. Seto, *Bioorg. Med. Chem.*, 2007, **15**, 458-473. (c) R. Srinivasan, L. P. Tan, H. Wu, P. -Y. Yang, K. A. Kalesh and S. Q. Yao, *Org. Biomol. Chem.*, 2009, **7**, 1821-1828. (d) L. P. Tan, H. Wu, P.-Y. Yang, K. A. Kalesh, X. Zhang, M. Hu, R. Srinivasan and S. Q. Yao, *Org. Lett.*, 2009, **11**, 5102-5105.
42. (a) K. A. Kalesh, K. Liu and S. Q. Yao, *Org. Biomol. Chem.*, 2009, **7**, 5129-5136. (b) M. Klein, P. Dinér, D. Dorin-Semlat, C. Doerig and M. Grötli, *Org. Biomol. Chem.*, 2009, **7**, 3421-3429. (c) A. J. Poot, J. van Ameijde, M. Slijper, A. van den

- Berg, R. Hilhorst, R. Ruijtenbeek, D. T. S. Rijkers and R. M. J. Liskamp, *ChemBioChem*, 2009, **10**, 2042-2051.
43. L. V. Lee, M. L. Mitchel, S. J. Huang, V. V. Fokin, K. B. Sharpless and C.-H. Wong, *J. Am. Chem. Soc.*, 2003, **125**, 9588-9589.
44. (a) J. Wang, M. Uttamchandani, J. Li, M. Hu and S. Q. Yao, *Org. Lett.*, 2006, **8**, 3821- 3824. (b) M. Hu, J. Li and S. Q. Yao, *Org. Lett.*, 2008, **10**, 5529-5531. (c) Po C. Chen, V. Patil, W. Guarrant, P. Green and A. K. Oyelere, *Bioorg. Med. Chem.*, 2008, **16**, 4839-4853. (d) J. Shen, R. Woodward, J. P. Kedenburg, X. Liu, M. Chen, L. Fang, D. Sun and P. G. Wang, *J. Med. Chem.*, 2008, **51**, 7417-7427. (e) W. S. Horne, C. A. Olsen, J. M. Beierle, A. Montero and M. R. Ghadiri, *Angew. Chem. Int. Ed.*, 2009, **48**, 4718-4724.
45. (a) H. Fuwa, Y. Takahashi, Y. Konno, N. Watanabe, H. Miyashita, M. Sasaki, H. Natsugari, T. Kan, T. Fukuyama, T. Tomita and T. Iwatsubo, *ACS Chem. Biol.*, 2007, **2**, 408-418. (b) F. Hof, A. Schutz, C. Fah, S. Meyer, D. Bur, J. Liu, D. E. Goldberg and F. Diederich, *Angew. Chem. Int. Ed.*, 2006, **4**, 2138-2141. (c) A. Brik, J. Muldoon, Y.-C. Lin, J. H. Elder, D. S. Goodsell, A. J. Olson, V. V. Fokin, K. B. Sharpless and C.-H. Wong, *ChemBioChem*, 2003, **4**, 1246-1248 (d) A. Brik, J. Alexandratos, Y.-C. Lin, J. H. Elder, A. J. Olson, A. Wlodawer, D. S. Goodsell and C.-H. Wong, *ChemBioChem*, 2005, **6**, 1167-1169.
46. S. L. Ng, P.-Y. Yang, K. Y. T. Chen, R. Srinivasan and S. Q. Yao, *Org. Biomol. Chem.*, 2008, **6**, 844-847.
47. G. Wulff, *Angew. Chem. Int. Ed.*, 1995, **34**, 1812-1832.

48. P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652-3711.
49. W. L. Mock and N.-Y. Shih, *J. Org. Chem.*, 1983, **48**, 3619-3620.
50. (a) W. G. Lewis, L. G. Greene, F. Grynszpan, Z. Radic, P. R. Carlier, P. Taylor, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 1053-1057. (b) R. Manetsch, A. Krasiniski, Z. Radic, J. Raushel, P. Taylor, K. B. Sharpless and H. C. Kolb, *J. Am. Chem. Soc.*, 2004, **126**, 10809-10818. (c) A. Krasiniski, Z. Radic, R. Manetsch, J. Raushel, P. Taylor, K. B. Sharpless and H. C. Kolb, *J. Am. Chem. Soc.*, 2005, **127**, 6686-6692.
51. V. P. Mocharla, B. Colasson, L. V. Lee, S. Roper, K. B. Sharpless, C. H. Wong and H. C. Kolb, *Angew. Chem. Intl. Ed.*, 2005, **44**, 116-120.
52. M. Whiting, J. Muldoon, Y. C. Lin, S. M. Silverman, W. Limstorm, A. J. Olson, H. C. Kolb, M. G. Finn, K. B. Sharpless, J. H. Elder and V. V. Fokin, *Angew. Chem. Int. Ed.*, 2006, **45**, 1435-1439.
53. (a) Z. -Y. Zhang, A. M. Thieme-Seffler, D. Maclean, D. J. McNamara, E. M. Dobrusin, T. K. Sawyer, J. E. Dixon, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 4446-4450. (b) S. W. Vetter, Y. -F. Keng, D. S. Lawrence, Z. -Y. Zhang, *J. Biol. Chem.*, 2000, **275**, 2265-2268. (c) C. Blanchetot, M. Chagnon, N. Dubé, M. Hallé, M. L. Tremblay, *Methods*, 2005, **35**, 44-53. (d) M. Köhn, M. Gutierrez-Rodriguez, P. Jonkheijm, S. Wetzel, R. Wacker, H. Schroeder, H. Prinz, C. M. Niemeyer, R. Breinbauer, S. E. Szedlacsek, H. Waldmann, *Angew. Chem. Int. Ed.*, 2007, **46**, 7700-7703.

54. L. C. Lo, T. L. Pang, C. H. Kuo, Y. L. Chiang, H. Y. Wang, J. J. Lin, *J. Proteome Res.*, 2002, **1**, 35-40.
55. Q. Zhu, X. Huang, G. Y. J. Chen, S. Q. Yao, *Tetrahedron Lett.*, 2003, **44**, 2669-2672.
56. (a) S. Kumar, B. Zhou, F. Liang, W. -Q. Wang, Z. Huang, Z. -Y. Zhang, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 7943-7948. (b) S. Kumar, B. Zhou, F. Liang, H. Yang, W. -Q. Wang, Z. -H. Zhang, *J. Proteome Res.*, 2006, **5**, 1898-1905.
57. S. Liu, B. Zhou, H. Yang, Y. He, Z. -X. Jiang, S. Kumar, L. Wu, Z. -Y. Zhang, *J. Am. Chem. Soc.*, 2008, **130**, 8251-8260.
58. M. E. Jung, T. I. Lazarova, *J. Org. Chem.*, 1997, **62**, 1553-1555.
59. (a) H. Sun, C. H. S. Lu, M. Uttamchandani, Y. Xia, Y. -C. Liou and S. Q. Yao, *Angew. Chem. Int. Ed.*, 2008, **47**, 1698. (b) H. Sun, L. P. Tan, L. Gao and S. Q. Yao, *ChemComm.* 2009, 677.
60. (a) H. Sun, S. Chattopadhyaya, J. Wang and S. Q. Yao, *Anal. Bioanal. Chem.*, 2006, **386**, 416-426. (b) R. Srinivasan, X. Huang, S. L. Ng and S. Q. Yao, *ChemBioChem*, 2006, **7**, 32-36.
61. (a) L. Bialy and H. Waldmann, *Angew. Chem. Int. Ed.*, 2005, **44**, 3814-3839. (b) S. Zhang and Z. -Y. Zhang, *Drug Disc. Today*, 2007, **12**, 373-381.
62. (a) P. Klatt and S. Lamas, *Eur. J. Biochem.*, 2000, **267**, 4928-4944. (b) A. Salmeen, J. N. Andersen, M. P. Myers, T. -C. Meng, J. A. Hinks, N. K. Tonks and D. Barford, *Nature*, 2003, **423**, 769-773. (c) R. L. M. van Montfort, M. Congreve, D. Tisi, R. Carr and H. Jhoti, *Nature*, 2003, **423**, 773-777.
63. R. Kitz and I. B. Wilson, *J. Biol. Chem.*, 1962, **237**, 3245-3249.

64. T. P. Geladopoulos, T. G. Sotiroudis and A. E. Evangelopoulos, *Anal. Biochem.*, 1991, **192**, 112-116.
65. L. V. Ravichandran, H. Chen, Y. Li and M. J. Quon, *Mol. Endocrinol.*, 2001, **15**, 1768-1780.
66. (a) J. M. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell. Biol.*, 2006, **7**, 775-782. (b) T. W. Muir, *Annu. Rev. Biochem.*, 2003, **72**, 249-289.
67. G. Mayer and A. Heckel, *Angew. Chem. Int. Ed.*, 2006, **45**, 4900-4921.
68. J. H. Kaplan, B. Forbush III and J. F. Hoffman, *Biochemistry*, 1978, **17**, 1929-1935.
69. J. Engels and E. –J. Schlaeger, *J. Med. Chem.*, 1977, **20**, 907-911.
70. (a) T. Furuta and M. Iwamura, *Methods in Enzymology*, 1998, **291**, 50-63. (b) T. Furuta, S. S. H. Wang, J. L. Dantzker, T. M. Dore, W. J. Bybee, E. M. Callaway, W. Denk and R. Y. Tsien, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 1193-1200. (c) A. Z. Suzuki, T. Watanabe, M. Kawamoto, K. Nishiyama, H. Yamashita, M. Ishii, M. Iwamura and T. Furuta, *Org. Lett.*, 2003, **5**, 4867-4870. (d) V. Hagen, S. Frings, B. Wiesner, S. Helm, U. B. Kaupp and J. Bendig, *ChemBioChem*, 2003, **4**, 434-442.
71. A. D. Turner, S. V. Pizzo, G. Rozakis and N. A. Porter, *J. Am. Chem. Soc.*, 1988, **110**, 244-250.
72. (a) R. S. Givens and C. –H. Park, *Tetrahedron Lett.*, 1996, **37**, 6259-6262. (b) C. –H. Park and R. S. Givens, *J. Am. Chem. Soc.*, 1997, **119**, 2453-2463. (c) R. S. Givens, A. Jung, C. –H. Park, J. Weber and W. Bartlett, *J. Am. Chem. Soc.*, 1997,

- 119**, 8369-8370. (d) G. Arabaci, X. -C. Guo, K. D. Beebe, K. M. Coggeshall and D. Pei, *J. Am. Chem. Soc.*, 1999, **121**, 5085-5086. (e) R. S. Givens, J. F. W. Weber, P. G. Conrad, G. Orosz, S. L. Donahue and S. A. Thayer, *J. Am. Chem. Soc.*, 2000, **122**, 2687-2697. (f) K. Zou, W. T. Miller, R. S. Givens and H. Bayley, *Angew. Chem. Int. Ed.*, 2001, **40**, 3049-3051.
73. M. Lukeman and J. C. Scaiano, *J. Am. Chem. Soc.*, 2005, **127**, 7698-7699.
74. (a) J. W. Walker, G. P. Reid, J. A. McCray and D. R. Trentham, *J. Am. Chem. Soc.*, 1988, **110**, 7170-7177. (b) J. E. T. Corrie, A. Barth, V. R. N. Munasinghe, D. R. Trentham and M. C. Hutter, *J. Am. Chem. Soc.*, 2003, **125**, 8546-8554.
75. (a) K. Curley and D. S. Lawrence, *J. Am. Chem. Soc.*, 1998, **120**, 8573-8574. (b) Chung-yu Chang, T. Fernandez, R. Panchal and H. Bayley, *J. Am. Chem. Soc.*, 1998, **120**, 7661-7662.
76. J. S. Wood, M. Koszelak, J. Liu and D. S. Lawrence, *J. Am. Chem. Soc.*, 1998, **120**, 7145-7146.
77. W. F. Veldhuyzen, Q. Nguyen, G. McMaster and D. S. Lawrence, *J. Am. Chem. Soc.*, 2003, **125**, 13358-13359.
78. D. Humphrey, Z. Rajfur, M. E. Vazquez, D. Scheswohl, M. D. Schaller, K. Jacobson and B. Imperiali, *J. Biol. Chem.*, 2005, **280**, 22091-22101.
79. T. Kawakami, H. Cheng, S. Hashiro, Y. Nomura, S. Tsukiji, T. Furuta and T. Nagamune, *ChemBioChem*, 2008, **9**, 1583-1586.
80. H. Li, Jung-Mi Hah and D. S. Lawrence, *J. Am. Chem. Soc.*, 2008, **130**, 10474-10475.

81. I. A. Yudushkin, A. Schleifenbaum, A. Kinkhabwala, B. G. Neel, C. Schultz and P. I. H. Bastiaens, *Science*, 2007, **315**, 115-119.
82. (a) D. M. Rothman, M. E. Vázquez, E. M. Vogel and B. Imperiali, *Org. Lett.*, 2002, **4**, 2865-2868. (b) D. M. Rothman, E. J. Petersson, M. E. Vázquez, G. S. Brandt, D. A. Dougherty and B. Imperiali, *J. Am. Chem. Soc.*, 2005, **127**, 846-847.
83. R. Reinhard and B. F. Schmidt, *J. Org. Chem.*, 1998, **63**, 2434-2441.
84. K. Gumireddy, S. J. Baker, S. C. Cosenza, P. John, A. D. Kang, K. A. Robell, M. V. R. Reddy and E. P. Reddy, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1992-1997.
85. F. J. Adrià, Q. Ding, T. Sim, A. Velentza, C. Sloan, Y. Liu, G. Zhang, W. Hur, S. Ding, P. Manley, J. Mestan, D. Fabbro and N.S. Gray, *Nat. Chem. Biol.*, 2006, **2**, 95-102.
86. (a) Y. Liu and N. S. Gray, *Nat. Chem. Biol.*, 2006, **2**, 358-364. (b) C. Pargellis, L. Tong, L. Churchill, P.F. Cirillo, T. Gilmore, A. G. Graham, P. M Grob, E. R. Hickey, N. Moss, S. Pav and J. Regan, *Nat. Struct. Biol.*, 2002, **9**, 268-272. (c) J. G. Cumming, C.L. McKenzie, S.G. Bowden, D. Campbell, D.J. Masters, J. Breed, and P. J. Jewsbury, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5389-5394. (d) P. W Manley, G. Bold, J. Brüggem, G. Fendrich, P. Furet, J. Mestan, C. Schnell, B. Stolz, T. Meyer, B. Meyhack, W. Stark, A. Strauss and J. Wood, *Biochim. Biophys. Acta.*, 2004, **1697**, 17-27. (e) P. T. C. Wan, M. J. Garnett, S. M. Roe, S. Lee, D. Niculescu-Duvaz, V. M. Good, C. M. Jones, C. J. Marshall, C. J. Springer, D. Barford and R. Marais, *Cell*, 2004, **116**, 855-867. (f) A. L. Gill, M. Frederickson, A. Cleasby, S. J. Woodhead, M. G. Carr, A. J. Woodhead, M. T.

- Walker, M. S. Congreve, L. A. Devine, D. Tisi, M. O'Reilly, L. C. A. Seavers, D. J. Davis, J. Curry, R. Anthony, A. Padova, C. W. Murray, R. A. E. Carr and H. Jhoti, *J. Med. Chem.*, 2005, **48**, 414-426. (g) A. Ricouart, J. C. Gesquiere, A. Tartar and C. Sergheraert, *J. Med. Chem.*, 1991, **34**, 73-78. (h) K. Parang, J. H. Till, A. J. Ablooglu, R. A. Kohanski, S. R. Hubbard and P.A. Cole, *Nat. Struct. Biol.*, 2001, **8**, 37-41. (i) J. H. Lee, S. Kumar and D.S. Lawrence, *ChemBioChem*, 2008, **9**, 507-517. (j) S. C. Meyer, C. D. Shomin, T. Gaj and I. Ghosh, *J. Am. Chem. Soc.*, 2007, **129**, 13812-13813. (k) E. Enkvist, D. Lavogina, G. Raidaru, A. Vaasa, I. Viil, M. Lust, K. Viht and A. Uri, *J. Med. Chem.*, 2006, **49**, 7150-7159.
87. M. W. N. Deininger, J. M. Goldman and J.V. Melo, *Blood*, 2000, **96**, 3343-3356.
88. B. Nagar, O. Hantschel, M. A. Young, K. Scheffzek, D. Veach, W. Bornmann, B. Clarkson, G. Superti-Furga and J. Kuriyan, *Cell*, 2003, **112**, 859-871.
89. T. Schindler, W. Bornmann, P. Pellicena, W. T. Miller, B. Clarkson and J. Kuriyan, *Science*, 2000, **289**, 1938-1942.
90. (a) P. W. Manley, S. W. Cowan-Jacob and J. Mestan, *Biochim. et Biophys. Acta*, 2005, **1754**, 3-13. (b) M. E. Gorre, M. Mohammed, K. Ellwood, N. Hsu, R. Paquette, P. N. Rao and C. L. Sawyers, *Science*, 2001, **293**, 876-880. (c) N. von Bubnoff, F. Schneller, C. Peschel and J. Duyster, *Lancet*, 2002, **359**, 487-491.
91. (a) T. O'Hare, D. K. Walters, M. W. N. Deininger and B. J. Druker, *Cancer Cell*, 2005, **7**, 117-119. (b) P. La Rosee, A. S. Corbin, E. P. Stoffregen, M. W. Deininger and B. J. Druker, *Cancer Res.*, 2002, **62**, 7149-7153. (c) J. S. Tokarski, J. A. Newitt, C. Y. J. Chang, J. D. Cheng, M. Wittekind, S. E. Kiefer, K. Kish, F. Y. F. Lee, R. Borzilleri, L. J. Lombardo, D. Xie, Y. Zhang and H. E. Klei,

- Cancer Res.*, 2006, **66**, 5790-5797. (d) P. W. Manley, W. Breitenstein, J. Brüggem, S.W. Cowan-Jacob, P. Furet, J. Mestan and T. Meyer, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5793-5797. (e) E. Weisberg, P. W. Manley, S. W. Cowan-Jacob, A. Hochhaus and J. D. Griffin, *Nat. Rev. Cancer*, 2007, **7**, 345-356.
92. For reviews, see: (a) K. A. Kalesh, P. -Y. Yang, R. Srinivasan and S. Q. Yao, *QSAR Comb. Sci.*, 2007, **26**, 1135-1144. (b) A. Birk, C. -Y. Wu and C. -H. Wong, *Org. Biomol. Chem.*, 2006, **4**, 1446-1457.
93. L.V. Lee, M.L. Mitchell, S-J. Huang, V.V. Fokin, K.B. Sharpless and C-H. Wong, *J. Am. Chem. Soc.*, 2003, **125**, 9588-9589.
94. Y. -F. Liu, C. -L. Wang, Y. -J. Bai, N. Han, J. -P. Jiao and X. -L. Qi, *Org. Process Res. Dev.*, 2008, **12**, 490-495.
95. (a) D. A. Walsh and D. B. Glass, *Methods in Enzymology*, 1991, **201**, 304-316. (b) H. Hidaka, M. Watanabe and K. Kobayashi, *Methods in Enzymology*, 1991, **201**, 328-339. (c) T. Tamaoki, *Methods in Enzymology*, 1991, **201**, 340-347.
96. N. W. Roehm, G. H. Rodgers, S. M. Hatfield and A. L. Glasebrook, *J. Immunol. Methods*, 1991, **142**, 257-265.
97. M. A. Seeliger, M. Young, M. N. Henderson, P. Pellicena, D. S King, A. M. Falick and J. Kuriyan, *Protein Sci.*, 2005, **14**, 3135-3139.
98. J. G. Moffatt and H. G. Khorana, *J. Am. Chem. Soc.*, 1961, **83**, 649-658.
99. Z. S. Szakács, S. Béni, Z. Varga, L. Örfi, G. Kéri and B. Noszál, *J. Med. Chem.*, 2005, **48**, 249-255.

100. G. B. Bennett, R. B. Mason, L.J. Alden and J. B. Roach, Jr, *J. Med. Chem.*, 1978, **21**, 623-628.
101. M. Lautens, J-F. Paquin, S. Piguel and M. Dahlmann, *J. Org. Chem.*, 2001, **66**, 8127-8134.
102. J. Villén, S. A. Beausoleil, S. A. Gerber and S. P. Gygi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1488-1493.
103. (a) D. T. McLachlin and B. T. Chait, *Curr. Opin. Chem. Biol.*, 2001, **5**, 591-602.
(b) Z. A. Knight, B. Schilling, R. H. Row, D. M. Kenski, B. W. Gibson and K. M. Shokat, *Nat. Biotechnol.*, 2003, **21**, 1047-1054.
104. (a) J. Ptacek and M. Snyder, *Trends Genet.*, 2006, **22**, 545-554. (b) Z. Songyang, S. Blechner, N. Hoagland, M. F. Hoekstra, H. Piwnica-Worms and L. C. Cantley, *Curr. Biol.*, 1994, **4**, 973-982. (c) J. Ptacek, *Nature*, 2005, **438**, 679-684. (d) J. J. Allen, M. Li, C. S. Brinkworth, J. L. Paulson, D. Wang, A. Hübner, W. H. Chou, R. J. Davis, A. L. Burlingame, R. O. Messing, C. D. Katayama, S. M. Hedrick and K. M. Shokat, *Nat. Methods*, 2007, **4**, 511-516. (e) J. D. Blethrow, J. S. Glavy, D. O. Morgan and K. M. Shokat, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1442-1447.
105. D. J. Maly, J. A. Allen, K. M. Shokat, *J. Am. Chem. Soc.*, 2004, **126**, 9160-9161.
106. P. Zuman, *Chem. Rev.*, 2004, **104**, 3217-3238.
107. S. P. Haris, Y. Zhang, B. L. Bourdonnec, C. R. McCurdy and P. S. Portoghese, *J. Med. Chem.*, 2007, **50**, 3392-3396.

108. K. Liu, K. A. Kalesh, L. B. Ong and S. Q. Yao, *ChemBioChem*, 2008, **9**, 1883-1888.
109. (a) M. S. Cohen, H. Hadjivassiliou and J. Taunton, *Nat. Chem. Biol.*, 2007, **3**, 156-160. (b) J. A. Blair, D. Rauh, C. Kung, C. -H. Yun, Q. -W. Fan, H. Rode, C. Zhang, M. J. Eck, W. A. Weiss and K. M. Shokat, *Nat. Chem. Biol.*, 2007, **3**, 229-238.
110. (a) M. -C. Yee, S. C. Fas, M. M. Stohlmeyer, T. J. Wandless and K. A. Cimprich, *J. Biol. Chem.*, 2005, **280**, 29053-29059. (b) Y. Liu, K. R. Shreder, W. Gai, S. Corral, D. K. Ferris and J. S. Rosenblum, *Chem. Biol.*, 2005, **12**, 99-107.
111. S. J. Ratcliffe, T. Yi and S. S. Khandekar, *J. Biomol. Screen.*, 2007, **12**, 126-132.
112. M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner and J. W. Kozarich, *Biochemistry*, 2007, **46**, 350-358.
113. (a) E. D. Lowe, I. Tews, K. Y. Cheng, N. R. Brown, S. Gul, M. E. Noble, S. J. Gamblin and L. N. Johnson, *Biochemistry*, 2002, **41**, 15625-15634. (b) W. F. Waas and K. N. Dalby, *J. Biol. Chem.*, 2002, **277**, 12532-12540. (c) B. E. Aubol, B. Nolen, D. Vu, G. Ghosh and J. A. Adams, *Biochemistry*, 2002, **41**, 10002-10009. (d) D. A. Critton, A. Tortajada, G. Stetson, W. Peti and R. Page, *Biochemistry*, 2008, **47**, 13336-13345.
114. C. Blanchetot, M. Chagnon, N. Dubé, M. Hallé and M. L. Tremblay, *Methods*, 2005, **35**, 44-53.
115. (a) M. T. Brown and J. A. Cooper, *Biochem. Biophys. Acta*, 1996, **1287**, 121-149. (b) S. M. Thomas and J. S. Brugge, *Annu. Rev. Cell Dev. Biol.*, 1997, **13**, 513-609.

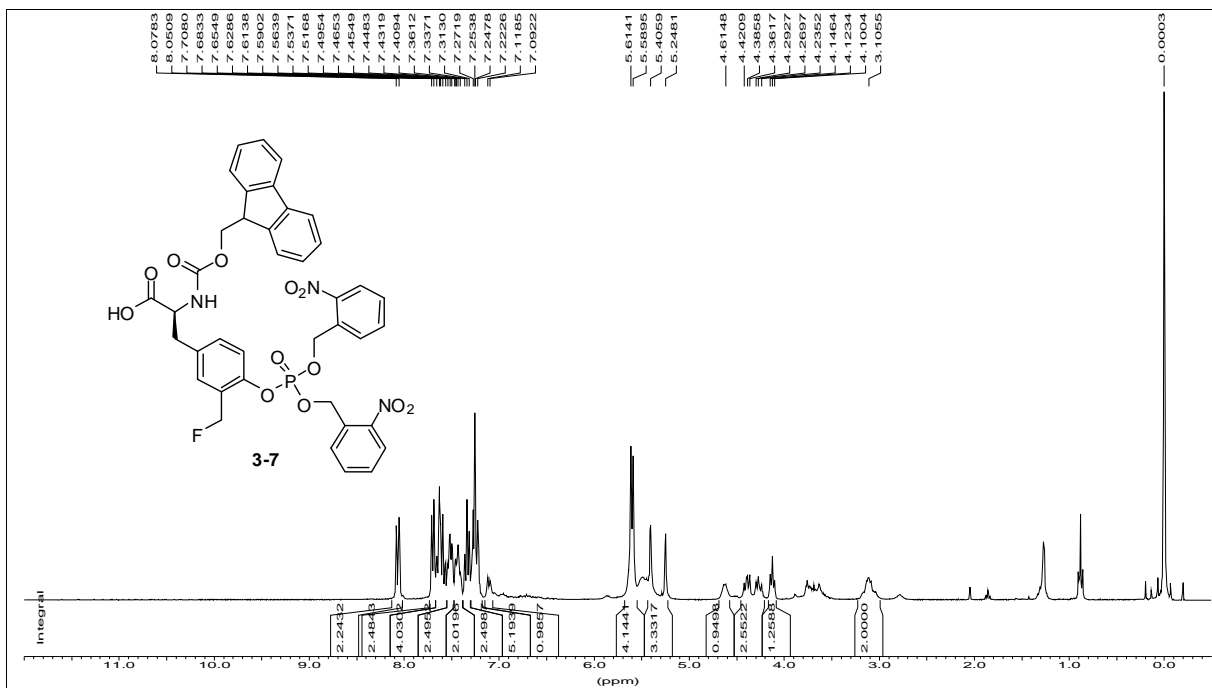
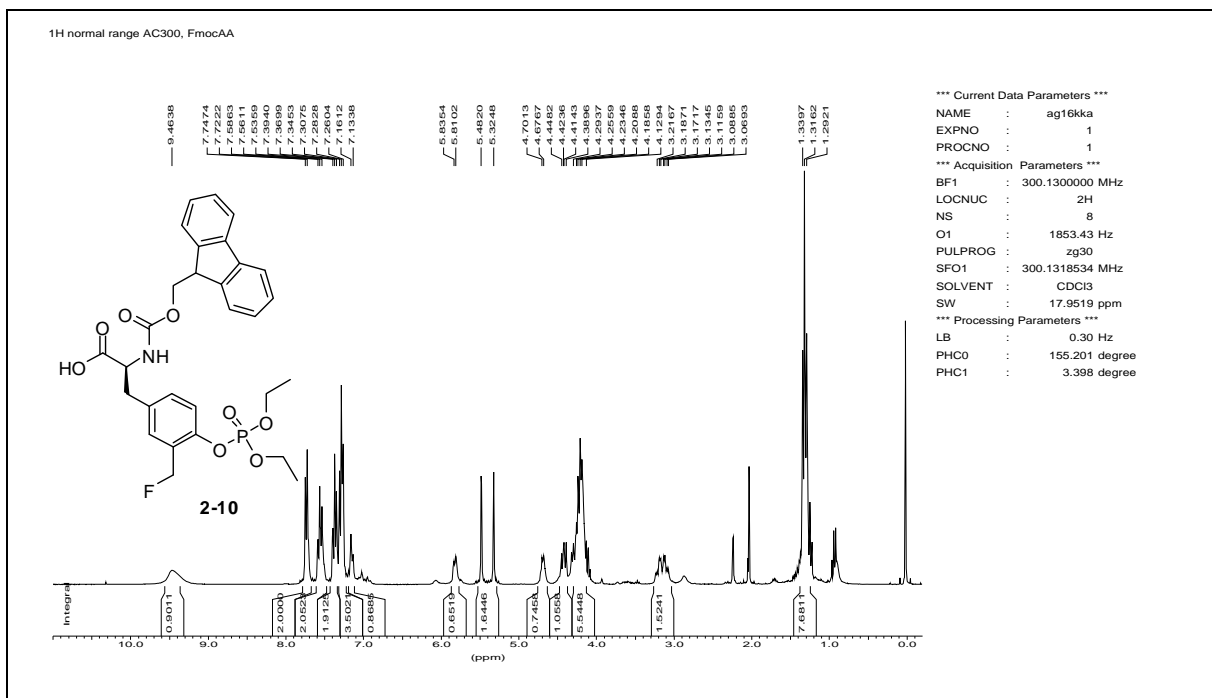
116. (a) S. Nada, M. Okada, A. MacAuley, J. A. Cooper and H. Nakagawa, *Nature*, 1991, **351**, 69-72. (b) R. R. Roussel, S. R. Brodeur, D. Shalloway and A. P. Laudano, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 10696-10700. (c) K. B. Bibbins, H. Boeuf and H. E. Varmus, *Mol. Cell. Biol.*, 1993, **13**, 7278-7287. (d) X. Liu, S. R. Brodeur, G. Gish, Z. Songyang, L. C. Cantley, A. P. Laudano and T. Pawson, *Oncogene*, 1993, **8**, 1119-1126. (e) M. Ruzzene, Z. Songyang, O. Marin, A. Donella-Deana, A. M. Brunati, B. Guerra, P. Agostinis, L. C. Cantley and L. A. Pinna, *Eur. J. Biochem.*, 1997, **246**, 433-439. (f) D. Sondhi, W. Xu, Z. Songyang, M. J. Eck and P. A. Cole, *Biochemistry*, 1998, **37**, 165-172. (g) P. A. Cole, K. Shen, Y. Qiao and D. Wang, *Curr. Opin. Chem. Biol.*, 2003, **7**, 580-585.
117. (a) C. Jacobs and H. Rubsamen, *Cancer Res.*, 1983, **43**, 1696-1702. (b) C. A. Cartwright, A. I. Meisler and W. Eckhart, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 558-562. (c) D. K. Luttrell, A. Lee, T. J. Lansing, R. M. Crosby, K. D. Jung, D. Willard, M. Luther, M. Rodriguez, J. Berman and T. M. Gilmer, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 83-87. (d) M. S. Talamonti, M. S. Roh, S. A. Curley and G. E. Gallick, *J. Clin. Invest.*, 1993, **91**, 53-60. (e) N. Resen, J. B. Bolen, A. M. Schwartz, P. Cohen, V. Deseau and M. A. Israel, *J. Biol. Chem.*, 1986, **261**, 13754-13759. (f) A. E. Ottenhoff-Kalff, G. Rijksen, E. A. van Beurden, A. Hennipman, A. A. Michels and G. E. Staal, *Cancer Res.*, 1992, **52**, 4773-4778. (g) C. Egan, A. Pang, D. Durda, H. C. Cheng, J. H. Wang and D. J. Fujita, *Oncogene*, 1999, **18**, 1227-1237.
118. (a) S. Bagrodia, I. Chackalaparampil, T. E. Kmiecik and D. Shalloway, *Nature*, 1991, **349**, 172-175. (b) I. Chackalaparampil and D. Shalloway, *Cell*, 1988, **52**, 801-810. (c) R. M. Kypta, Y. Goldberg, E. T. Ulug and S. A. Courtneidge, *Cell*, 1990, **62**, 481-492.

119. J. D. Bjorge, A. Pang and D. J. Fujita, *J. Biol. Chem.*, 2002, **275**, 41439-41446.
120. (a) T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705-6710. (b) D. Wang and P. A. Cole, *J. Am. Chem. Soc.*, 2001, **123**, 8883-8886. (c) K. Shen and P. A. Cole, *J. Am. Chem. Soc.*, 2003, **125**, 16172-16173.
121. (a) M. -Q. Xu and F. B. Perler, *EMBO J.*, 1996, **15**, 5146-5153. (b) S. Chong, F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus and Ming-Qun Xu, *Gene*, 1997, **192**, 271-281.
122. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776-779.
123. (a) S. Mitra and A. M. Barrios, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 5142-5145. (b) S. Mitra and A. M. Barrios, *Anal. Biochem.*, 2007, **370**, 249-251.
124. (a) K. Shah, Y. Liu, C. Deirmengian and K. M. Shokat, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 3565-3570. (b) Y. Liu, K. Shah, F. Yang, L. Witucki and K. M. Shokat, *Chem. Biol.*, 1998, **5**, 91-101. (c) T. Fox, J. Coll, X. Xie, P. Ford, U. Germann, M. Porter, S. Pazhanisamy, M. Fleming, V. Galullo, M. Su and K. Wilson, *Protein Sci.*, 1998, **7**, 2249-2255. (d) A. Bishop, C. Kung, K. Shah, L. Wituchi, K. M. Shokat and Y. Liu, *J. Am. Chem. Soc.*, 1999, **121**, 627-631. (e) K. Specht and K. M. Shokat, *Curr. Opin. Cell Biol.*, 2002, **14**, 155-159.
125. (a) K. Wilson, P. McCaffrey, K. Haiiao, S. Pazhanisamy, V. Galullo, G. Bemis, M. Fitzgibbon, P. Caron, M. Murcko and M. Su, *Chem. Biol.*, 1997, **4**, 423-431. (b) A. Bridges, *Chem. Rev.*, 2001, **101**, 2541-2572. (c) M. Gorre, M. Mohammed, K.

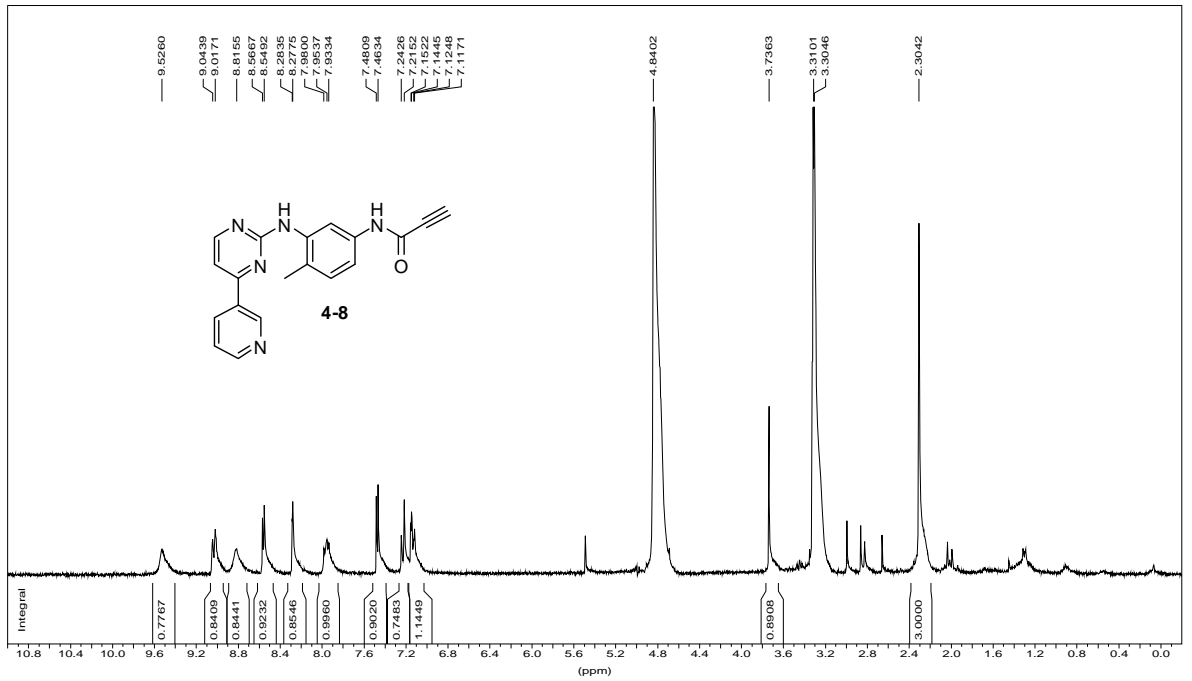
Ellwood, N. Hsu, R. Paquette, P. Rao and C. Sawyers, *Science*, 2001, **293**, 876-880.

126. (a) M. S. Cohen, C. Zhang, K. M. Shokat and J. Taunton, *Science*, 2005, **308**, 1318-1321. (b) J. A. Blair, D. Rauh, C. Kung, Cai-Hong Yun, Qi-Wen Fan, H. Rode, C. Zhang, M. J. Eck, W. A. Weiss and K. M. Shokat, *Nat. Chem. Biol.*, 2007, **3**, 229-238.

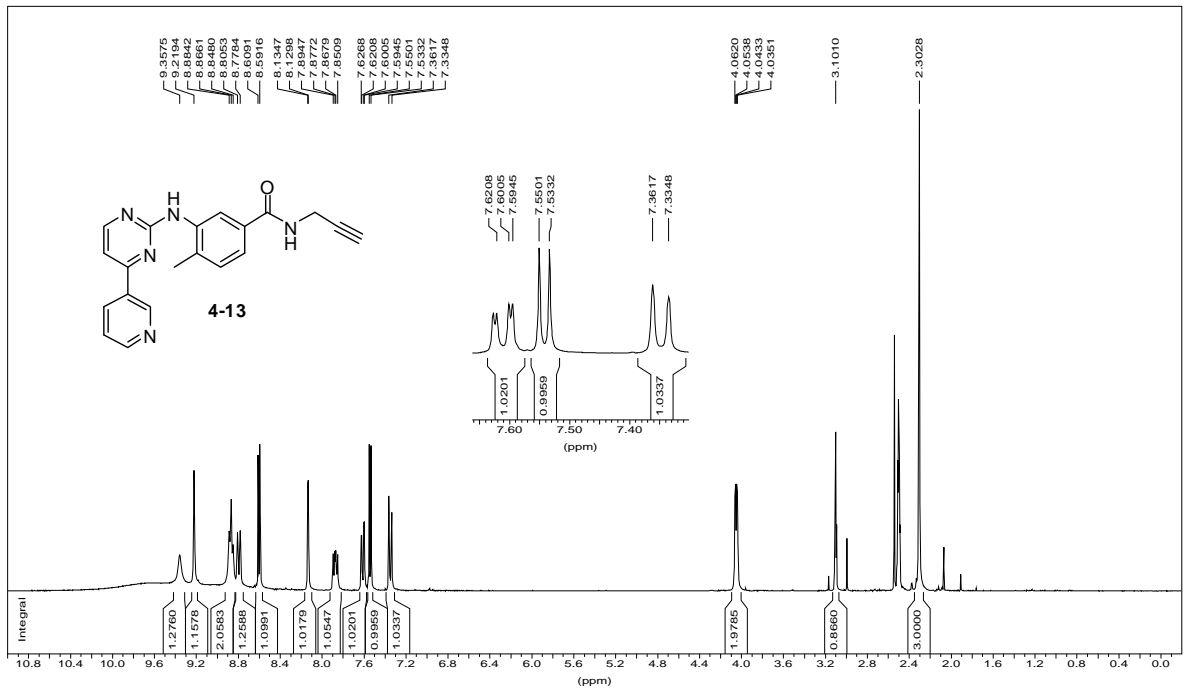
Appendix

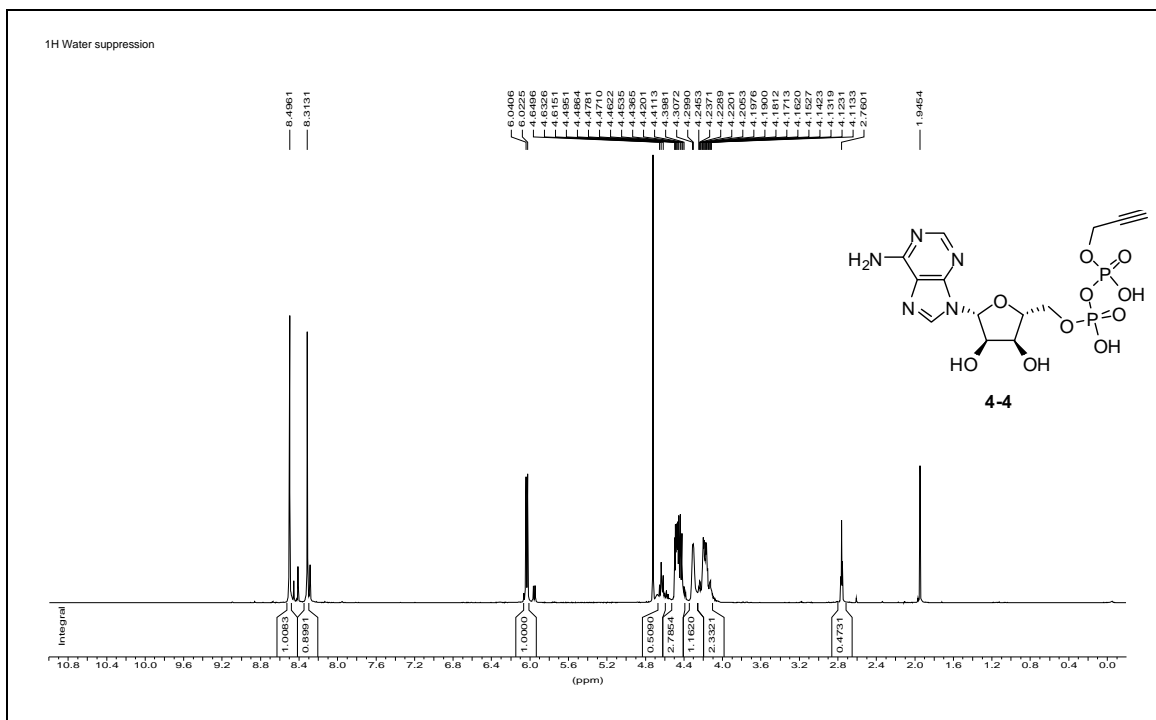


1H normal range AC300, B3-217



1H normal range AC300, W2





1H normal range AC300, W1-SA20

