FROM HIGH THROUGHPUT INHIBITOR DISCOVERY TO HIGHLY EFFICIENT CELLULAR DELIVERY OF THERAPEUTIC PROTEINS AND DRUGS

FU JIAQI

NATIONAL UNIVERSITY OF SINGAPORE

2015

FROM HIGH THROUGHPUT INHIBITOR DISCOVERY

TO HIGHLY EFFICIENT CELLULAR DELIVERY OF

THERAPEUTIC PROTEINS AND DRUGS

FU JIAQI

(B. Sc. Zhejiang University)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety, under the supervision of Professor Yao Shao-Qin, (in the laboratory S5-03-15), Chemistry Department, National University of Singapore, between 5/08/2011 and 30/09/2015.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____ _____ _____

Name   Signature   Date

Acknowledgements

I feel very fortunate that I have studied as a graduate student in NUS for the past four years. My sincere gratitude goes to plenty of people who played very important roles on the completion of this thesis as well as in my life. This section is dedicated to them.

This first person I would like to express my deepest gratitude to is my supervisor, Prof. Yao Shaoqin. I feel honoured to be able to work in his group. Without his guidance and support, it is impossible for me to complete my PhD projects. His wide knowledge and personal encouragements are of great value to me. He gave lots of constructive comments to help me solve the tough problems in my experiments. He also offered me lots of good opportunities to practice and develop my research skills. I extremely appreciate his help for providing me a wonderful research experience for the past four years.

I would like to thank all members in Prof Yao's group - both past and present. Mahesh, Jingyan, Li Lin, Zhengqiu, Zhenkun, Lay Pheng, Xiamin, Su Ying, Chongjing, Changmin, Xiaofei, Yanli, Hailong, Biwei, Linghui, Sijun, Peng Bo, Danyang, Shubo, Xiaoyuan, Chelsea, Maoxin, Shikun, Dongsheng, Wang Ping, Yu Sheng who made this time so special and enjoyable. I thanks specially the following person: Mahesh and Zhenkun, who taught me microarray-related technology; Jingyan and Li Lin who not only mentored me on research but helped me on my life in Singapore; Linghui and Sijun who discussed with me and helped me to have a better understanding about research; Dr Hao Wu, Ms Madelin Teo and Ms Kaijia Peh, who assisted with the library synthesis and initial screens against SH2 domains; Pengbo, who helped me on

the synthesis of tri-functional linker; Changmin and Shikun, who helped me on Chemistry and MSNs experiment.

Finally, I would like to thank my parents and Mr Liu Jiang for their understanding, supports and encouragement throughout these years.

Big Thanks to all of you.

Table of Contents

Summary

Drug discovery is the process through which potential new medicines are identified and brought into market. In 20th century, advances in technology have accelerated the processes of drug discovery. For instance, human genome project has identified more than 500+ proteins as druggable target. Combinatorial chemistry and solid phase chemistry have speed up the rate of compounds synthesis in 800 folds. It is an opportunity, but meanwhile a challenge. When look at above numbers, it is not easy to study the libraries and target proteins at such large scales especially considering the selectivity issues. To solve the problem, high-throughput screening (HTS) platform is required for lead compounds discovery. In this thesis, we have made use of microarray as a HTS platform for the specific ligand discovery of SH2 domains. Besides, our microarray has branched away from the exclusive pursuits of small molecule "hits" to proteomics, the large-scale studies of the performances of bioactive compounds in live cells.

Once identified using HTS, hits need further optimizations to generate leads which have acceptable drug properties. However, in spite of the efforts, approximately 40% of currently marketed drugs and up to 75% of compounds currently under development have been suggested to be poorly water-soluble. The difficulty is more pronounced for macromolecules such as proteins and DNAs/siRNAs, which are regarded as emerging therapeutic agents. Therefore, we have established a highly efficient delivery system of drugs including both small and large molecules using cell penetrate poly(disulfide)s.

List of Publications

1. Foong, Y. M., <u>Fu, J.,</u> Yao S. Q., Uttamchandani, M. "Current Advances in Peptide and Small Molecule Microarray Technologies." *Curr. Opin. Chem. Biol.,* **2012**, *16*, 234-242.

2. <u>Fu, J.;</u> Na, Z.; Uttamchandani, M.; and Yao, S.Q. "Profiling human Src homology 2 (SH2) domain proteins and ligand discovery using a peptide-hybrid small molecule microarray." *Chem. Commun*., **2013**, 49, 9660-9662.

3. Wang, P.; Na, Z.; <u>Fu, J.;</u> Tan, C.Y.J.; Zhang, H.; Yao, S.Q.; Sun, H. "Microarray immobilization of biomolecules using a fast trans-cyclooctene (TCO)-tetrazine reaction." *Chem Commun.*, **2014**,*50*, 11818-1182.

4. <u>Fu, J.;</u> Yu, C.; Li, L.; Yao, S.Q. "Intracellular Delivery of Functional Proteins and Native Drugs by Cell-Penetrating Poly(disulfide)s" *J. Am.Chem. Soc.*, **2015***, 137*, 12153-12160

5. <u>Fu, J.;</u> Na, Z.; Peng, B.; Uttamchandani, M.; and Yao, S.Q. "Accelerated cellular on- and off-target screening of bioactive compounds using microarrays" *Chem Commun, submitted.*

6. Yu, C.; Qian, L.; Uttamchandani, M.; Ge, J.; <u>Fu, J.;</u> Yuan, P.; Yao, S.Q. "Single Vehicular Co-delivery of Antagomir and Native Drugs for Enhanced Inhibit miR-21 Function" *Angew. Chem. Int. Ed.* Manuscript in preparation.

List of Tables

List of Figures

List of Schemes

List of Abbreviations and Symbols

μM	Micromolar
Å	Angstrom
CPZ	Chlorpromazine
ABP	Activity based probe
ABPP	Activity-based probe profiling
Abs	Antibodies
AD	Alzheimer's disease
ADME	Absorption, distribution, metabolism and excretion
A*f*BP	Affinity-based probe
BJH	Barrett-Joyner-Halenda
Boc	tert-Butoxycarbonyl
br	Broad
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue-stained gel
CC	Click chemistry
CM	Carboxymethyl
CPD	Cell-Penetrating Poly(disulfide)s
CPPs	Cell-penetrating peptides
C-terminus	Carboxyl terminus
CuAAC	Cu(I)-catalyzed 1,3-dipolar cycloaddition
DCM	Dichloromethane
dd	Doublet of doublets
DIC	N,N'-diisopropylcarbodiimide
DIEA	N, N'-diisopropylethylamine
DLS	dynamic light scattering
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide

DTNB	5,5'-dithioobis-2-nitrobenzoic acid
DTT	Dithiothreitol
E. coli	Escherichia coli
EA/EtOAc	Ethyl acetate
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide HCl
EDTA	Ethylenediamine tetraacetic acid
eq.	Equivalent
ESI	Electron Spray Ionization
Et_2O	Diethyl ether
EtOH	Ethanol
FA	Folic acid
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FL	Fluorescent gel
Fmoc	9-Fluorenylmethoxycarbonyl
FP	fluorescence polarization
g	Gram
h	Hours
HBA	Hydroxybenzoic acid
HBTU	O-Benzotriazole-N,N,N',N'-tetramethyl-uroniumhexafluorophosphate
HCl	Hydrogen chloride
HCS	High-content screening
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
$HHG2C_{18}$	1, 5-dioctadecyl-L-glutamyl 2-histidyl-hexahydrobenzoic acid
HOBT	N-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HTP	High-throughput
HTS	High-throughput screening
Hz	Hertz

K_D	Dissociation constant
$KHCO_3$	potassium hydrogen carbonate
K_m	Michaelis constant
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
M	Molar
m	Multiplet
MAP	Mitogen-activated protein
MeOH	Methanol
METH	Methamphetamine
mg	Milligram
min	Minute
mM	Millimolar
mmol	Millimole
MS	Mass
MSNs	Mesoporous silica nanoparticles
MW	Molecular weight
MβCD	methyl-β-cyclodextrin
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
NaOH	Sodium hydroxide
Ni-NTA	Nickel-nitrilotriacetic acid
nM	Nanomolar
NMR	Nuclear magnetic resonance
NPs	nanoparticles
NTA	nitrilotriacetic acid
N-terminus	Amino terminus
$^{\circ}C$	degree celsius
PAE	poly(amino-co-ester)s

PAMAM	polymer polyamidoamine
PBS	Phosphate buffered saline
PD	pull-down
PDI	protein disulfide isomerases
PDI	polydispersity index
PEI	polyethylenimine
pH	Negative logarithm of the hydroxonium ion concentration
PLA	Polylactic acid
PLL	poly(L-lysine)
PPIs	Protein-protein interactions
ppm	Parts per million
PTMs	Post-translational modifications
pTyr.	Phosphotyrosine
PyBrOP	Bromotripyrrolidinophosphonium hexafluorophosphate
q	Quartet
RPPA	reverse-phase protein array
s	Singlet
s	seconds
S.D.	Standard deviation
SAR	structure-activity relationship
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SH2	Src homology 2
siCPDs	substrate-initiated, cell-penetrating poly-(disulfide)s
SMDs	small molecule drugs
SMM	small molecule microarray
SPC	soy phosphatidylcholine
SPPS	Solid phase peptide synthesis
SPR	surface plasmon resonance
STS	Staurosporine

t	Triplet
TAT	transcription trans-activating protein
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCEP	Tri(2-carboxyethyl)phosphine
TCO	*trans*-cyclooctyne
TEM	Transmission electron micrographs
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TOF	Time of flight
TPh A	triphenyl compound A
Tris	Tris(hydroxymethyl) amino methane
TSA	tyramide signal amplification
UV	Ultraviolet
V_{max}	maximum velocity
w	wortmannin
WB	Western blot
XTT	2,3-Bis-(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide
λabs	Absorption wavelengt
λem	Emission wavelength
λ_{ex}	Exitation wavelength

List of Amino Acids

One Letter	Three Letter	Amino Acid
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	Trp	Tryptophan
Y	Tyr	Tyrosine

Chapter 1.

Introduction

Summary

Drug discovery is the process through which potential new medicines are identified and brought into market. The advent of genomic science, DNA sequencing, combinatorial chemistry, cell-based assays, and automated high-throughput screening (HTS) has led to a new opportunity of drug discovery.[1]
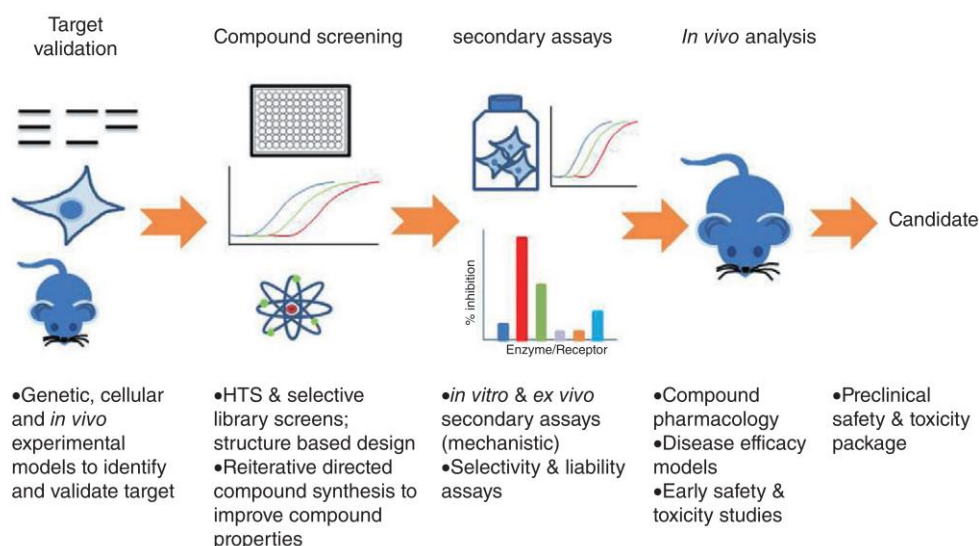
Figure 1.1. Overview of typical drug discovery process[2]

For instance, human genome project has identified more than 500+ proteins as druggable targets.[3] Combinatorial chemistry and solid phase chemistry have speeded up the rate of compounds synthesis in 800 folds. All these advances in technologies made it possible for rapid drug identification.

It is an opportunity, meanwhile a challenge. When we look at above numbers, it is not easy to study the libraries and target proteins at such large

scales especially considering the selectivity issues. To solve this problem, HTS technologies are required, which are capable of screening many protein interactions concurrently and sensitively at minimum cost. Over the last decade, microarray has been established as a robust HTS platform for rapid screening, lead discovery and molecular characterization.[4-8]

Once the compounds (termed as "hits") are identified via microarray screening, they would undergo evaluation in cell or animal models to validate their efficiencies in complex environment. Meanwhile, relevant drug properties such as toxicity, solubility, permeability, absorption, distribution, metabolism and excretion (ADME) are also tested. It is unlikely that the perfect compounds could be found within one-round screening. Always, further optimization related to SAR investigations is required.[2]

In spite of the efforts on drug optimization, around 40% of currently marketed drugs and up to 75% of compounds currently under development have been suggested to be poorly water-soluble.[9] It is not easy to deliver drugs to the target positions. The difficulty is more pronounced for macromolecules such as proteins and DNAs/siRNAs, which are considered as emerging therapeutic agents.[10] Therefore, companies start to make use of macromolecular delivery systems such as cell-penetrating peptides (CPPs), liposomes and poly-cationic formulations to address these issues.

After several rounds of optimizations, the surviving hits (now termed as "leads") undergo preclinical and clinical trials. As shown in Figure 1.2, it is not until now that safety issues such as side-effects and "off-targets" are extensively investigated. Often, it is costly when drugs fail at these stages. To avoid the loss,

the development of platform for broad-based, proteome-wide screening at the outset of drug discovery is needed urgently. During this process, the target binding spectrums and key biological pathways affected by the leads are revealed at the early stage and thus miniaturized the loss.

Identification of active leads
Identify molecules that are active on the target protein.

Optimisation phase
Optimise the molecules' properties so that they are safe and effective.

Preclinical development
Preclinical trials in order to document that the compound is safe for entry into human trials.

Confirmed phase III trials
Pivotal trials to confirm efficacy and safety in larger patient populations.

Identification and validation of target proteins
Identify and validate the biological mechanism behind a disease.

Identification of hit compounds
Identify classes of molecules that have the potential to become pharmaceuticals.

Candidate drugs
Selection of substances with optimised properties for further development.

Phase I and phase II
The first clinical trial is carried out on healthy volunteers in order to document pharmaceutical safety, followed by patient trials in order to define the safe and effective dose.

Registration and market launch

Figure 1.2. overview of pre-clinical and clinical phase

In recent years, the development of activity-based protein profiling (ABPP),[11] to some extent, alleviates the problem. By minimally modifying those drug-like compounds, the resulting probes enable the large-scale interrogation of protein small molecule interactions and the rapid identification of potential cellular targets in live cells.

In this chapter, I would first introduce the protein and small molecule microarrays and their recent applications related to drug discovery. Secondly, I would give the detailed descriptions in the current strategies of protein delivery systems including CPPs, polymers and liposomes as well as their pros and cons.

1.1 Microarray

Over the last decade, microarray screening has transformed the life science research landscape. Novel applications such as molecular

fingerprinting, ligand discovery, protein expression profiling and interaction networks mapping, have significantly impacted both basic and applied spheres of research. Creative ideas by biologists, chemists and engineers are propelling this interdisciplinary technology to interesting new levels. The numerous successful examples have inspired a growing following of scientists to take on high-throughput, discovery-driven research, drawing impetus towards accelerated information assimilation and knowledge growth.

Microarrays are miniaturised assemblies of molecule library organized across a planar surface.[8] The size of library ranges from hundreds to tens of thousands. Each compounds of the library is identified based on the physical position on planner surface. Breaking away from tradition microtiter-based (or well-based) assays, the core feature of microarray is miniaturization and parallelization, allowing many different biological moieties available within a compact surface for massively parallel screening against multiple targets of choice to generate informative and comparable results while consuming low quantities of precious sample and analytes.

Microarray technology is first introduced as miniaturized DNA assemblies on chips.[12] After that, it does not take long before further pioneering efforts make it possible to immobilize small molecules, peptides and subsequently proteins in addressable grids for rapid analysis.[13,14] Since then, we have witnessed the development in microarray technology and the emergency of innovative application based on various microarray-based platforms.

Traditional microarrays have been applied to the discovery of small molecule ligands or substrates for protein targets of interest. For example, peptide and small molecule serve as short substrates, inhibitors or ligands, which can be readily targeted by enzymes or other receptors. Large-scale screening on microarrays with those libraries against target proteins facilitates the discovery of specific binding motif, as well as the binding spectrums and preference of proteins and consequently accelerates the drug discovery process.

In recent years, microarray has branched away from the exclusive pursuits of small molecule "hits", to proteomics, the large-scale studies of the structures and functions of proteins.[4,5,15] As reported, proteins regulate virtually every cellular process. The majority of diseases, including cancers, are characterized by the abnormal functions of certain proteins. The systems-wide analysis of these abnormal functions leads to a comprehensive understanding of the intricate singling transduction pathways in diseased states, before and after treatment with drugs. This valued information facilitates the identification of target protein, validation of drug's efficiency, understanding of drug mechanisms, and discovery of biomarkers that predict or guide therapeutic response.

1.1.1 Discovery of Ligand, Inhibitor and Substrate

Microarray-assisted "hit" discovery strategies can be classified into two categories: those are screened against recombinant proteins, and those are against cellular lysate.[16]

For the first one, microarrays have been incubated with purified target proteins and the bindings are detected using readouts such as surface

plasmon resonance (SPR), fluorescently labelled antibodies against the proteins or the fluorescent tags on the proteins. This method has been used to identify probes of various types of proteins, including enzymes (such as kinase,[17-19] proteases[20-22] and decetylases[23-25]) as well as classically challenging proteins (such as growth factors and BRCT domains of BRCA1 proteins).[26]

In one recent example that highlighted the discovery of potent enzyme inhibitor, Lee and Park developed Jeffamine-coated slide which facilitated the reduction of nonspecific binding and thus enhanced the sensitivity. With such slides, 2, 4, 4'-trihydroxychalcone (Figure 1.3) was identified as a potent inhibitor of tyrosinase, a major regulator of melanin production (K_D=0.4 μM).[27] The activity was confirmed by an inhibition assay that measured the catalytic rate of the tyrosinase-assisted conversion of $_L$DOPA to DOPAchrome (IC_{50}=0.76 μM). Further, it demonstrated that 2, 4, 4'-trihydroxychalcone was able to treat melasma and hyperpigmentation, which were skin disorders resulted from abnormally high level of melanin.

Apart from inhibitors which target the active sites of enzymes, microarray has also been applied to the study of protein-protein interactions (PPIs). PPIs control important regulatory pathways in virtually all cellular process. Inappropriate interactions between proteins are known to cause many human diseases. Compared to mainstream drugs, which usually bind to the active-sites, and have historically been the focus of pharmaceutical research, such inhibitors especially those made of small and cell-permeable molecules are attractive due to their generally better selectivity profiles.

However, the development of protein-protein interactions (PPIs) is perceived to be highly challenging. When referring to PPIs as targets in drug discovery, it appears to be that large protein-protein (PP)-contact surfaces (i.e. ~1500-3000 $Å^2$) are involved, most of which are flat and often lacking in suitable pockets for small molecules to bind. [28,29]

In spite of the challenges, a number of successful examples have started to prove that it is possible to overcome these hurdles and develop PPI modulators as drugs.[26,30-32] By integrating small molecule and peptide hybrid libraries and microarray technologies, our group has successfully discovered the cell-permeable inhibitor (**15**, Figure 1.3) against BRCT domain of BRCA1 protein with K_D value from 250 nM to 1 μM range. In this setup, we first determined the preferred BRCA1 (BRCT)$_2$-binding motif as RVF*p*SPVF using 1000-member combinatorial phosphor-peptide microarray. Then, the two flanking peptide fragments of *p*S residue were replaced by a library of commercial available acid and amine building blocks, representatively. A 51-membrane N-terminal library and a 50-member C-terminal hybrid library were synthesized and anchored onto array to screen for potent binders of BRCT. After potent binders from each sub-library were identified, the "hits" amine and acid building blocks were reconstituted to yield the final non-peptide small molecule inhibitors. Together with its non-hydrolysable analogues **15a** and **15b**, these small molecule like dipeptide compounds were shown to possess both in vitro and in cell antitumor activities via the disruption of BRCA1 (BRCT)$_2$/protein interaction.

Compared to purified protein screening, recently, applying cell lysate to microarray has become more common. This format allows hits to bind targets with relevant post transcriptional modifications (PTMs) and interact with protein complex. This format also enables the discovery for proteins that cannot be purified.

An impressive example of hits development for proteins using such a strategy related to nuclear protein pirin.[33] Pirin is known to bind to Bcl3, whose overexpression level enhances the proliferation, survival and migration of many tumors. Osada and coworkers sought to identify small molecule probes of pirin in order to further characterize the cellular role of pirin. Using SMMs, the team screened >20,000 compounds from the RIKEN Natural Products Repository against pirin residing in cell lysate. Slides were incubated with HEK293T cell lysate with the overexpression of DsRed-fused pirin and the binder were detected via fluorescent readout. The positive hits, triphenyl compound A (TPh A, Figure 1.3) was used for further evaluation. Isothermal titration calorimetry (ITC) and crystal structure further illustrated the binding potency of TPh A to pirin ($K_D = 0.6$ μM) and corresponding binding site (near Fe^{2+}-binding site). Subsequent pull-down (PD) assay and cell-based experiments showed that TPh A was able to disrupt binding of pirin to Bcl3 and consequently inhibited cell migration.

Recent, the applications of identified hit are not limited to compound lead. In certain instances, the hits were also transformed into probes, which enabled the broader identification and pull-down of the target proteins in cellular lysate.[34-36] Also, they are converted to fluorescence polarization

(FP)-compatible binding probes for high-throughput assay, facilitating the discovery of small molecular inhibitors of PPIs.[30]
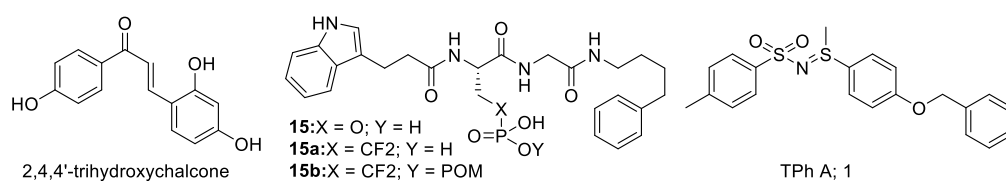
Figure 1.3 Structure of ligand, inhibitor or substrate[26,33,37]

1.1.2. Proteomic

In the last decade alone, microarrays have evolved from being used primarily as basic analytical research tools into now viable options for more sophisticated applications in protein profiling, molecule interaction mapping and biomarker discovery. In this part, we elaborate on the studies of microarray applications which focus on two areas relevant to proteomics research: functional annotation and biomarker discovery.

1.1.2.1. Functional Annotation

With traditional protein screening assay, the functional annotation of target proteins is usually performed by incubating them with corresponding substrates in individual well. The activities of proteins are reported in forms of absorbance, fluorescence or luminescence signals. Similarly, microarray requires a platform where the enzyme activities can be converted to readable signals with minimal interferences from neighbouring proteins. In principle, protein microarray can achieve the goal, where thousands of proteins are immobilized and simultaneously characterized, and the activities can be identified in a high throughput way. However, since enzymes are catalysts, the corresponding substrates will not bond to the immobilized enzymes but

are easily washed away. That means the strategy used in microplate is not compatible with the fluorescence detection method used in slide-based microarray technologies.

Activity based probes (ABPs), originally developed by Cravatt *et al.* in 1999 to study serine hydrolases, a mainly used to identify and characterized the activities of enzymes both *in vitro* and *in vivo*.[38] As shown in Figure 1.4, such probes normally contain three key elements 1) warhead, which is able to direct the probe to the enzyme (class) and form a covalent bond with the target enzyme(s) based on the catalytic activity of enzyme(s); 2) a reporter group, typically fluorophore tag for the visualization of labelled proteins; 3) a linker, which provides flexibility between reporter and warhead. Often, the activities of proteins are analysed by fluorescent gel or Western blot.
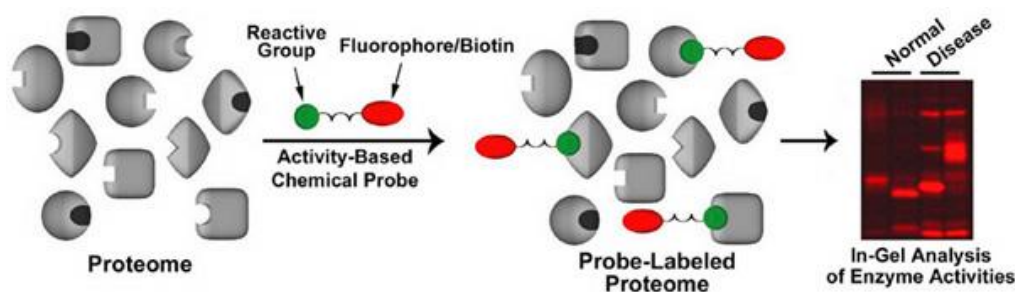
Figure 1.4. Structure of ABPs and ABPP process[39]

Inspired by ABPs, we developed the first microarray-based strategy for function annotation of enzyme activities in a rapid and reliable way. In this strategy, protein microarrays where active enzymes immobilized was treated with different ABPs, leading to selective covalent bindings between enzymes and the probes.[40] The activities were identified based on the fluorescent signals from APBs (Figure 1.5). Once developed, our strategy

has been applied to different classes of enzymes including phosphatases, cysteine proteases and serine hydrolases.[41,42] This enzyme assay was performed to reach two goals. First, this strategy allowed the high-throughput identification and assignment of protein function by virtue of their enzymatic activities. Second, inhibition studies could be conducted using potential enzyme inhibitors, either general or specific ones, allowing simultaneously characterize and compare the efficiency of inhibitors towards different enzymes. Based on this strategy, Eppinger et al. and Funeriu et al. have developed a strategy to quantitatively determine kinetic constants on the chip according to the fluorescent readouts.[43,44]
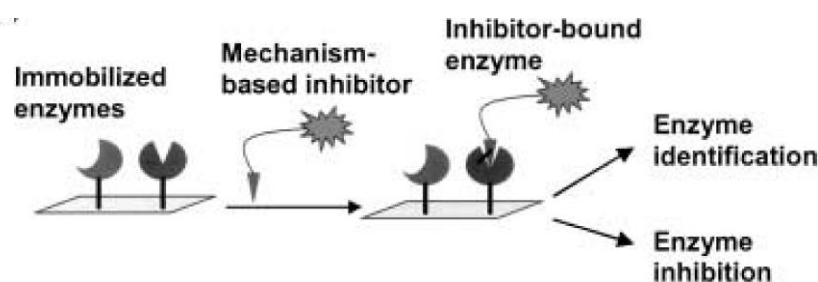
Figure 1.5. Principle of the activity-based detection of enzymes in a protein microarray with fluorescently-labelled, mechanism-based inhibitors.[40]

Recently, a new surface concentration-based assay was development to determine the kinetic parameters of proteases on microarray.[45] In this methods, a panel of rhodamine conjugated peptide substrates for MMP-3, calpain-1, caspase-3 and caspase-9 were immobilized onto maleimide-functionalized slide. The activities of enzymes were quantified by calculating the surface concentration of substrate peptides cleaved on the array (dye-off measurement). According to the solid-phase standard curve, this new assay was able to generate kinetic data such as Michaelis constant (K_m) and maximum velocity (V_{max}).

By integrating ABPs with antibody microarray, Cravatt and co-workers provided an alternative approach which enabled the proteomic profiling of enzyme activities on microarray platform.[46] Proteomic was first incubated with fluorescence-tagged probes. Then, the labelled enzymes were then probed and identified by the antibodies on array.

With those comprehensive understanding of the catalytic activity of enzymes, microarrays facilitated the definitions of new targets for the treatment of human diseases. By studying the phosphorylation patterns of paediatric brain tumors, Sikkema et al. have identified the elevated levels of Src family kinase activity in the tumors which could be targeted by the addition of kinase inhibitors, PP1 and dasatinib, leading to selectively induced death of these brain tumor cells.[47] This strategy allowed a focused analysis into disease states, providing the opportunity to guide further therapeutic interventions.

Microarray technologies also allowed the comprehensive understanding of intricate singling transduction pathways under disease states or with stimulations.

For instant, antibody arrays could be used to detect the changes in protein expression levels and PMTs, which are induced by the presence or absence of bioactive compounds. Sokolov and Cadet have made used of protein microarrays to study the correlation between the expression levels of different proteins and the behavioural phenotypes of mice which were treated with methamphetamine (METH).[48] As reported, METH abuse has been shown to stimulate aggressive behaviours in humans and in other animals. The author found that the repeat treatment of METH on mice induced the

deregulation of mitogen-activated protein (MAP) kinase pathway and finally increased aggressiveness and hyper-locomotion. In this work, 378 antibodies specific for proteins related to signal transduction, cell cycle regulation, cell structure, apoptosis, oncogene products and neurobiology were used to prepared microarray. The antibody microarray was incubated with fluorescent labelled protein extractions from METH treated mice. The results indicated that the level of Erk2 and 14-3-3e decreased in the striata of the mice treated with METH. Since Erk2 was regarded as a principle component of the MAP kinase pathway when protein 14-3-3e was an inhibitor of kinase C, the reduction in these two proteins suggested that repeated exposure to METH might alter MAP kinase-related pathway involved in behavioural changes.

In another example, Lackner and co-workers have developed a reverse-phase protein array (RPPA) to test the phosphorylation status of 100 proteins in different breast cancer cell lines.[49] Cell lysates from different cell lines were immobilized onto slide in serial dilutions and detected by various phosphorylated-recognized antibodies. This study allowed the group to carry out signalling pathway analysis in HTP way and classify breast cancer cell lines into different subtypes according to the slight differences in fingerprint. Furthermore, such analysis provided valuable information on the deregulated signalling pathways in individual cancers.

1.1.2.2. Biomarker Discovery

Furthermore, the availability of systematic and general HTP platforms to measure the changes across cellular states is helpful to understand the characteristics of unique diseases at the proteomic level.

Readdy et al. compared serum samples from both healthy persons and patients with Alzheimer's disease (AD) using peptide microarrays.[50] The samples were incubated on microarrays comprised 15,000 structurally novel peptides, and IgG binding patterns before visualized using secondary antibody. Three AD peptides were identified from the screening, and further characterization revealed at least two candidate auto-antibody biomarkers for AD.

Microarrays, as mentioned above, are HTS platforms which meet the requirement of large scale studies. In this part, we mainly talk about small molecule and protein microarrays. At early stage, the applications of microarrays mainly focus on the discovery of inhibitors, substrates and ligands. In recent years, microarray technologies are not limited to recombinant proteins but be used in more complex systems, helping researcher to have comprehensive understanding about the native cellular environment.

1.2 Cellular Delivery

Often, the drug candidates from microarray screening suffer from hydrophobicity and low water solubility issues. First of all, combinatorial synthesis and HTS are carried out in non-aqueous media (DMSO etc.) which decreases the possibilities of aqueous hits discovery. Second, the demands of increase potency, coupled with the theory that the binding potency is mediated by hydrophobic interactions, further magnifies the likehood of hydrophobic drug candidates. Third, current drug targets, such as lipids, process proteins or highly lipophilic ligands, which prefer to interacting with highly lipophilic,

poorly water-soluble drug candidates, amplifies the requirement. Therefore, in spite of efforts made during lead optimization, approximately 40% of currently marketed drugs and up to 75% of compounds currently under development have been suggested to be poorly water-soluble.[51,52] Due to the hydrophobicity, such drug candidates tend to be low or variable bioavailable, causing variability in clinical responses. Hence, it is urgent to develop methods for efficient delivery of such drugs into cells.

Beside small molecule drugs (SMDs), the other class is "large molecule" biologics including sugars, nuclei acids, recombinant therapeutic proteins or complex combination of these substances. Among these, therapeutic proteins are attractive and have been used to treat a variety of medical conditions where SMDs are not available. Compared to SMDs, proteins often generate highly binding affinity and specificity that cannot be mimicked by simple chemical compounds. Also, due to their highly specificities, there is less potential for protein therapeutics to interfere with off-target proteins or cause side-effect. Therefore, protein therapeutics have a significant role in the field of medicine. However, the applications of such strategies are limited by the low cellular delivery efficiency due to the high molecular weight, the polarity of therapeutic proteins and their poor stability against proteolytic and hydrolytic degradation.

Since the first discovery of TAT peptide, a variety of methods have been developed in the last 20 years, including the use of cell-penetrating peptides (CPPs),[53] super-charged proteins,[54] liposomes,[55] nanoparticles (NPs)[56] and polymers.[57] Based on their principle, these methods can be classified into three

catalogues, mechanical delivery methods, covalent protein modification and supramolecular delivery systems.[58]

1.2.3. Mechanical Delivery Methods

Mechanical delivery methods such as microinjection and electroporation are the most traditional strategy. These strategies enable the delivered protein to be direct access to the cytosol which are very useful for *in vitro* investigation. However, such strategies are low-throughput, invasive and require specialized equipment. All these dramatically limited the applications of mechanical delivery. Further, the transit cell permeabilization might introduce undesired biomolecules into cells and thus cause side effect.

1.2.4 Covalent Protein Modification

1.2.4.1 Cell-Penetrating Peptides (CPPs)

CPPs are generally peptides of less than 30 amino acids derived from natural or unnatural proteins or chimeric sequences that have the capability to cross cellular membranes via energy-dependent and/or independent machanisms.[53,59]

It is first found by Frankel and Pabo, who show that the transcription trans-activating protein (Tat) of HIV-1 could enter cells and translocated into nuclear.[60] In 1991, the group of Prochiantz demonstrated the similar property of Drosophila Antennapedia homeodomain.[61] Three year later, the first CPPs, a 16-mer-peptide driven from the third helix of the homeodomain termed penetratin, was discovered. After that, a batch of CPPs has been developed. However, it was until 1999 that Dowdy proved that CPPs enabled the delivery of small peptide and large proteins.[62] Until now, CPPs have been applied for the

delivery of synthetic and biological components into cells including enzymes, cytokines, apoptotic molecules, protein hormones, molecular chaperones and cell-signalling proteins.[58]

Generally speaking, most CPPs are positively charged peptides containing a few of basic residues and hydrophobic residues. Among these amino acids, basic residues facilitate CPPs to bind to the negatively charged proteoglycans and phospholipids on the cell surface via electrostatic interaction. Hydrophobic residues play a major role in the interaction with the plasma membrane bilayer via hydrophobic interaction. It has been demonstrated that both basic and hydrophobic interactions strongly affect the translocation efficiency and mechanism.[63,64]

In spite of these knowledges, the mechanism of cell uptake of CPP-tagged materials remains a matter of debate.[65] Original hypothesis suggested that CPPs worked by penetrating the cell membrane through translocation which enabled the direct access to cytosol. But further experiment indicated that the uptake occurred through endocytic mechanism. Currently, the more acceptable theory is that both endocytosis and direct translocation across the plasma membrane are involved in the uptake process. Besides, recent study suggests that a receptor-mediated entry is not ruled out for some CPP-conjugates. Generally speaking, several parameters including molecular weight, the nature of linkage, the cell line and the concentration affect uptake mechanism.[66-68] Nevertheless, endosomal entrapment and subsequently lysosomal degradation are the most challenging issues for CPPs assisted strategies. It is reported that CPPs linked proteins tend to rapidly concentrate

inside the endocytic organelles. As a result, the delivery efficiencies are dramatically decreased since only little protein is able to escape from endosome, reach its extra-endosomal target and exert its biological activity. Furthermore, macromolecules are subjected to degradation by acidic pH or hydrolases when they are trapped in late endosomes or lysosomes.[69]

To address this issue, several strategies have been developed to improve endosome escaping efficiency, including pH- and temperature-induced modulators, synthetic endosomal lysis agents (cyclic CPPs and multivalent CPPs), and photo-induced physical disruption.[58,69,70] However, the efficiencies of cytosolic access of these strategies are in doubt.

1.2.5. Supramolecular Delivery System

1.2.5.1. Liposome

Liposomes is one the nano-carrier (typically 50-500 nm in diameter), consisting of a lipid bialayer.[71] The formation of liposomes is a spontaneous process enabled by the interactions between water molecules and amphiphilic lipid molecules. Once completing, an aqueous volume becomes trapped within the core of liposome, giving liposome the ability to selectively sequester solvent for encapsulation, and thus form the basis for drug delivery.

Compared to the rest of delivery systems, the most interesting feature of liposome is the mechanism. As the bilayer of liposome closely mimic those of cells, they can be directly fused to the membrane. Even if the liposome is uptaken through the endocytosis, the bilayers enable the cargos to escape from endosome either by disrupting endosome membrane or by fusing with the membrane.

It is able to fabricate different types of liposome drug delivery vehicles based on the composition of the bilayers. By incorporating antibodies, ligands, polymers and peptides into the bilayers, liposome delivery can achieve pH-, temperature and photo-induced controllable release as well as target delivery.[72-76]

In recent study, Ping and co-workers have devised a novel type of multistage pH-responsive liposomes (HHG2C$_{18}$-L) for tumor target drug delivery (Figure 1.6).[77] The key component was 1, 5-dioctadecyl-L-glutamyl 2-histidyl-hexahydrobenzoic acid (HHG2C$_{18}$), a zwitterionic oligopeptide lipid. This lipid could mixed with soy phosphatidylcholine (SPC) and cholesterol to generate the smart liposomes, whose surface charge changed based on the environment pH value. Under physiological pH value (7.2-8.0), such liposome had strong negative charge surface charge (-22.8 mV) but changed sharply to +6.3 mV due to hydroxybenzoic acid (HBA) near tumor microenvironment. Consequently, the charge conversion facilitated the tumor cellular uptake. After that, a second-stage pH response occurred when the liposome entered endosomes and/or lysosomes where pH value decreased to 5.5. As a results 1) The imidazole group of HHG2C$_{18}$ facilitated proton influx to endolysosomes, resulting in endolysosomal escaping and 2) HBA was cleaved in the acidic environment to generate HHG2C$_{18}$-L which facilitated the mitochondria-target delivery of anticancer drugs.
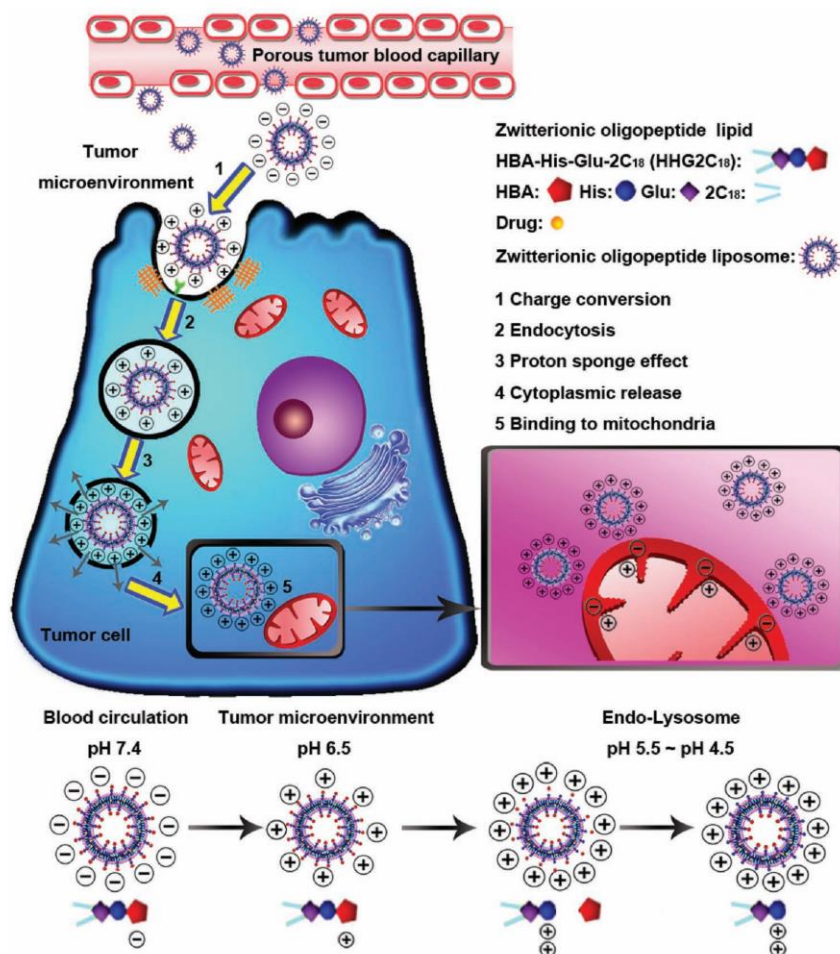
Figure 1.6 Schematic design of the smart liposomes (HHG2C$_{18}$-L) with multistage pH response to the tumor extracellular matrix and intracellular compartments for mitochondrial-targeted anticancer drug delivery.[77]

Since the first example demonstrated the liposome-mediated carrier as a useful tool for functional protein delivery,[78] it has proved that liposome-based methods enable the delivery of various target proteins into both adherent and suspension cells while retain the activities of those cargos. However, their delivery efficiencies vary significantly depending on protein size/charge since liposome-based methods make use of electrostatic/hydrophobic interactions for complex formation with target proteins.[79]

Further, such strategies are limited by its low loading amount (generally < 30% for active loading) due to the limited entrapment volume, the protein solubility and the poor stability.[80] A report indicates that a freeze-thaw cycling process would enhance the loading amount, but such a process is risky with the regard to loss of protein activities.

1.2.5.2. Polymer

Cationic polymers including polyethylenimine (PEI), poly(L-lysine) (PLL) have been widely used for drug delivery via endocytosis.[81] Once delivered, the carriers enable cargos to escape from endolysosome through "proton sponge effect", during which multivalent amine-containing polymer absorb protons, build up osmotic pressure in acidic organelles and finally cause the broken of endosome or lysosome. However, in spite of its wide applications, cationic polymers are somehow cytotoxic due to its large molecular weight and high surface cationic charge density.

To alleviate the problem, block copolymers containing a cationic block and a PEG block are introduced. In the copolymers, cationic blocks are responsive for cellular delivery while the PEG blocks form an exterior hydration layer that shields the cationic charge as well as stabilizes the complex. Further, moieties that facilitate the uptake or targeting capability can also be modified on PEG, making such a design an effective way for protein delivery. For instant, PLL-co-PEG block polymer was synthesized with folic acid (FA) linked to the end of PEG chaine.[82] In the system, PLL was responsible for the cellular uptake while folic acid was a targeting agent that recognized the folic acid receptor overexpressed by a variety of cancer cells. As a result, the cellular uptake was

largely enhanced in folate receptor overexpressed mouth epidermal carcinoma cells.

Beside, a variety of other biodegradable polymer backbones (polymer polyamidoamine (PAMAM), poly(amino-co-ester)s (PAE)) and structures (linear, branched, and dendritic architectures) have been also tested.[83,84]

In recent years, the development of stimuli-sensitive systems for protein delivery has been a subject of significant pharmaceutical importance, as the systems can recognize environmental changes, induce structural changes, and release therapeutic proteins in a responsive manner. With different building-up monomers, various stimuli-responsive polymer-assisted delivery strategies including temperature, pH, photo, redox potential bio-responsive systems have been built up.[85-87]

A particular interesting example of dendrimer-based delivery is shown by Yao and colleagues, who used carboxymethyl (CM) chitosan-poly(amidoamine) (PAMAM) dendrimer core-shell nanoparticles to load and release lysozyme.[88] In this system, PAMAM formed a hydrophobic negatively charged core which facilitated the loading of positively charged lysosome at physiological pH value. Outside, the negatively charged CM chitosan served as shell at pH 7.4. When pH decreased to 5.1, the protonation of amino groups in CM-chitosans and tertiary amines in PAMAM led to a strong electrostatic repulsion in CM-chitosan-PAMAM, triggering the release of encapsulated drugs. Compared to previous generated CM-chitosan polyion micelles, these new dendrimer nanoparticles have better loading capacity, higher pH sensitivity and less cytotoxicity.
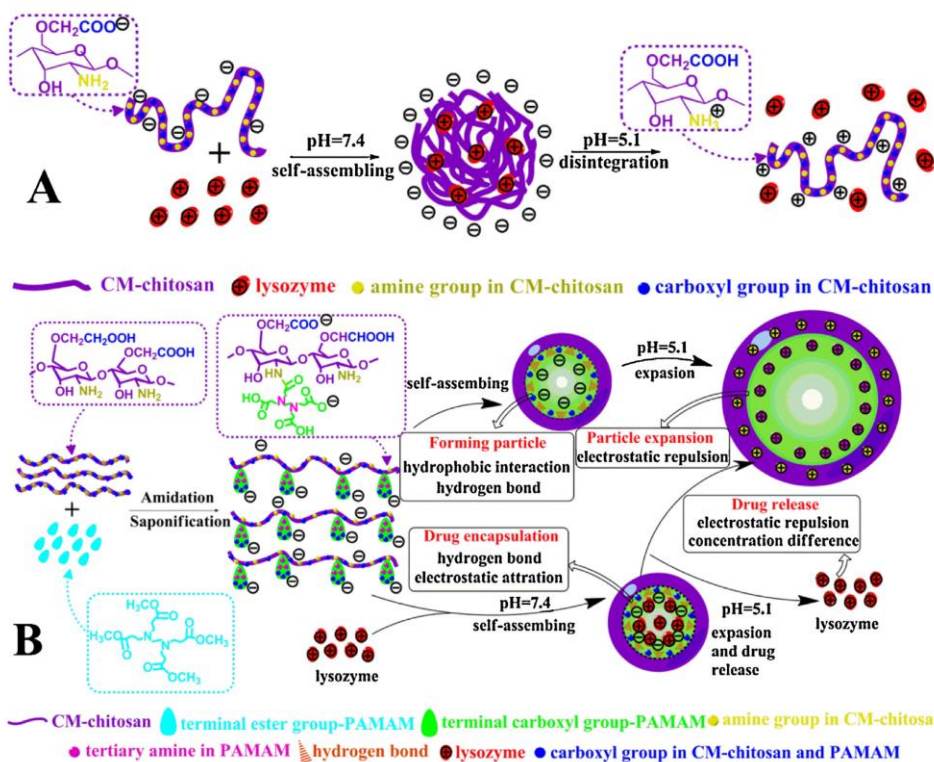
Figure 1.7. Overall strategy of CM chitosan-PAMAM delivery system. (A) Schematic representation of the encapsulation and release mechanism of lysozyme from CM-chitosan/lysozyme PIC nanoparticles. (B) Schematic representation of synthesis and self-assembling for CM-chitosan-PAMAM dendrimer nanoparticles, and the encapsulation and release mechanism of lysozyme from CM-chitosan PAMAM/lysozyme PIC nanoparticles.[88]

Further, with sophisticate design, blocks with different functions can be combined together to generate multi-controllable co-polymer system.[89] Redox sensitivity of polymers may be combined with *p*H-responsivity by copolymerization of polyethylenimine (PEI) with polylactic acid (PLA) via redox sensitive disulfide bonds to generate dual responsive polymer system. Under physiological *p*H value (7.4), such polymers had negative charge surface charge at around -10 mV but changed sharply to +15 mV when *p*H came to 6.8. Consequently, the charge conversion together with the folate ligand (FA) facilitated the tumor cell uptake. When polymers entered endosomes and/or lysosomes, their surface charge further increased to +20 mV, facilitating

endosome escaping of polymers through "proton sponge effect". Once into cytosol, disulfide bonds would be cleaved by GSH of high concentration, resulting in the high efficient release of anticancer drugs.
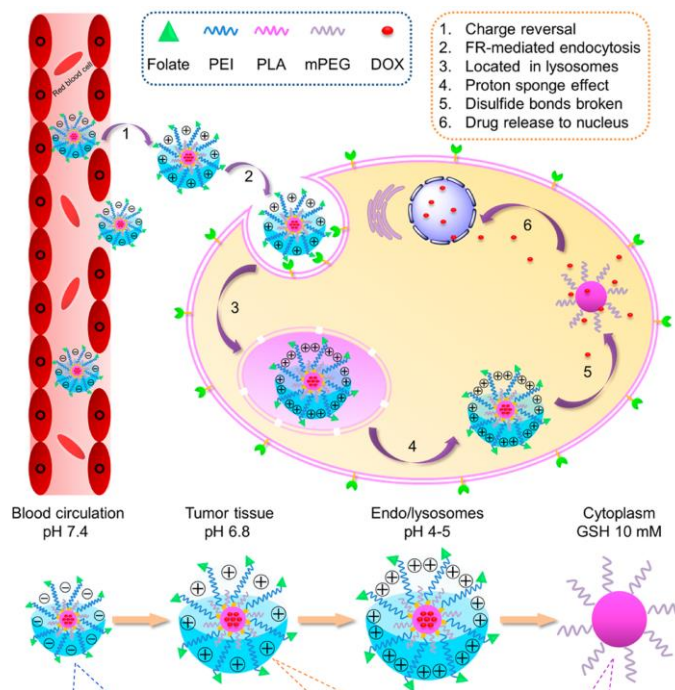
Figure 1.8. Schematic design of the targeted micelles (PELE/DA-FA) with charge reversal to the tumor extracellular matrix and intracellular GSH-triggered drug release.[89]

1.3 Objective

The objective of this thesis is to use SMMs as HTS strategy for the identifications of the specific ligands which are able to disrupt PPIs (Chapter 2). Further, microarrays are applied to proteomic profiling. By uniting protein microarrays with *in situ* profiling, we have established a platform for high throughput identifications of on- and off- targets of those drug candidates in live cells (Chapter 3). In addition, the highly efficient delivery system of drugs

including both small and large molecule is established with cell penetrate poly(disulphide)s (Chapter 4).

Chapter 2.

Profiling human Src homology 2 (SH2) domain proteins and ligand discovery using a peptide-hybrid small molecule microarray

Summary
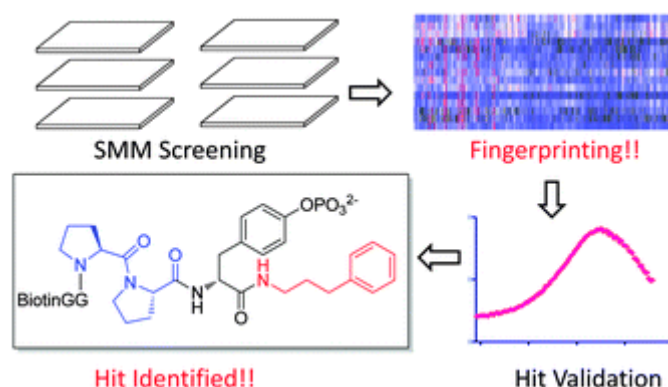
In this chapter, a 396-member peptide-hybrid small molecule microarray was fabricated to profile small molecule ligands across 15 SH2 proteins, yielding hits against Lck and Grb2.

2.1 Introduction

Src homology 2 (SH2) domains are readers of tyrosine kinase phosphorylation. By binding phosphorylated target sequences, these domains relay signals downstream to effect cellular regulation and gene expression control.[90-92]As a result, they form critical components of cell signalling networks. Over 110 human proteins contain SH2 domains, including kinases, phosphatases, transcription factors and adaptor proteins.[90,91]These proteins are over-expressed in various diseases, including cancer, and hence represent key targets for modulation using therapeutics.[90-92] Through *in vitro* and *in silico*

screens, various peptide and small molecule inhibitors have been identified against SH2 domains.[93,94] Herein, we present a small molecule microarray (SMM) strategy for rapid fingerprinting and profiling of SH2 proteins, which enables rapid molecular characterization and comparative inhibitor profiling.

Various methods have been applied to map the interaction preference of SH2 domains,[95-97] particularly more recently through the use of microarrays.[98-101] For example, Huang *et al.* have mapped the binding motifs for 76 SH2 domains against phosphopeptides using SPOT synthesis/macroarrays.[99] Using this map, SH2 domains have been classified into three groups (I, II, III). While the binding specificity of SH2 domains is attributed largely to the C-terminal segment of the phosphorylated targets, residues at the N-terminus also contribute to overall binding potency.[95-97,99]

In order to further examine the binding preferences of various SH2 domains and facilitate inhibitor discovery, we have built a novel phosphopeptide-hybrid small molecule microarray to investigate this class proteins. We designed and fabricated a 396-member library of small molecules containing a centrally located phosphotyrosine flanked by 9 N-terminal dipeptides and 44 C-terminal small molecule building blocks (Figure 2.1). The objectives of this strategy were to 1) significantly expand the specific binding preference of SH2 C-terminal pockets from traditional peptide-centric approaches to small molecule-centric strategies, 2) profile the binding spectra across multiple SH2 domains and identify selective small molecule building blocks which can bind the C-terminal pocket, and 3) identify complementary N-terminal sequences that could enhance binding efficacy and selectivity.
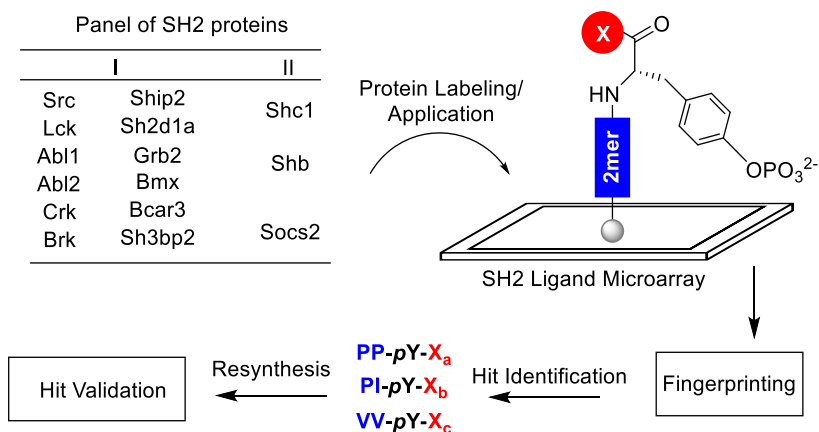
Figure 2.1. Overall strategy of the screening SH2 domain containing proteins against a ligand microarray for high-throughput profiling.[102]

2.2 Library Synthesis

The general strategy is shown in Figure. 2.1. A panel of SH2 domain proteins were individually expressed and purified using a one-step labelling and purification approach. The library was synthesized using 9 different N-terminal sequences from known SH2 binding targets and 44 commercially available amine building blocks whose structures cover aromatic, aliphatic, cyclic and bicyclic compounds (X_{1-44}; Figure. 2.2). A Biotin-Gly-Gly linker was added to facilitate compound immobilization onto avidin- functionalized glass surfaces. Synthesis was carried out on solid phase using PL-FMP resin and Irori Microkan[TM] reactors (Scheme 2.1). Following acidic cleavage and ether precipitation, representative compounds were dissolved in DMF and run on LC-MS, to ensure purity was > 80%, and the MS of the final product was correct.
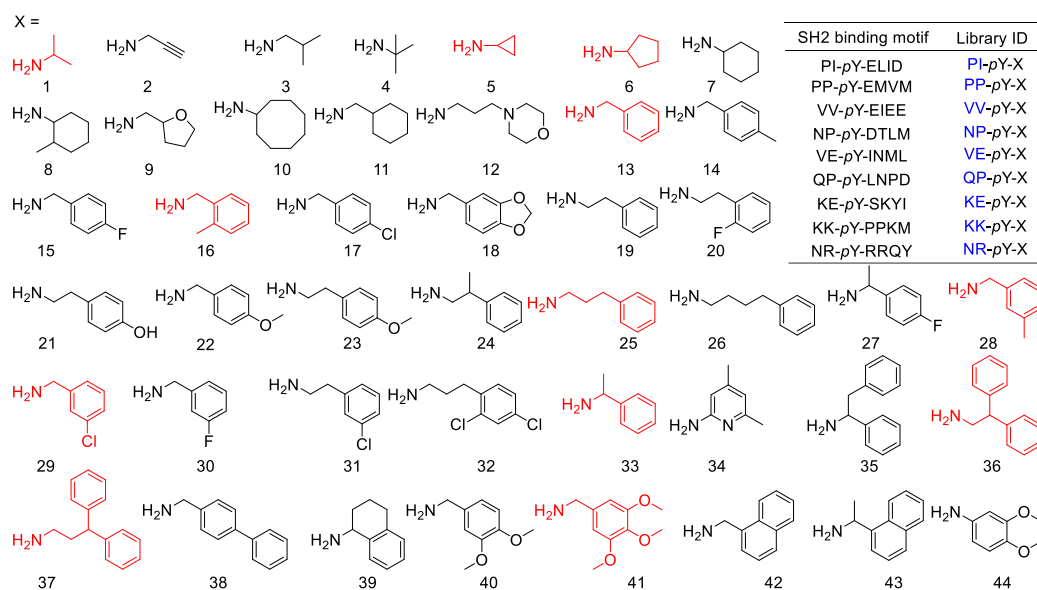
Figure 2.2. Design and synthesis of the 396-member SH2 ligand library.[102] (A) C-terminal substituents comprising of 44 different amine building blocks. Highlighted in red are the C-terminal structures from the 14 hits identified using the SMM screening results. (B) Original peptide sequence motifs for the selected 9 SH2 N-terminal substituents (in blue).
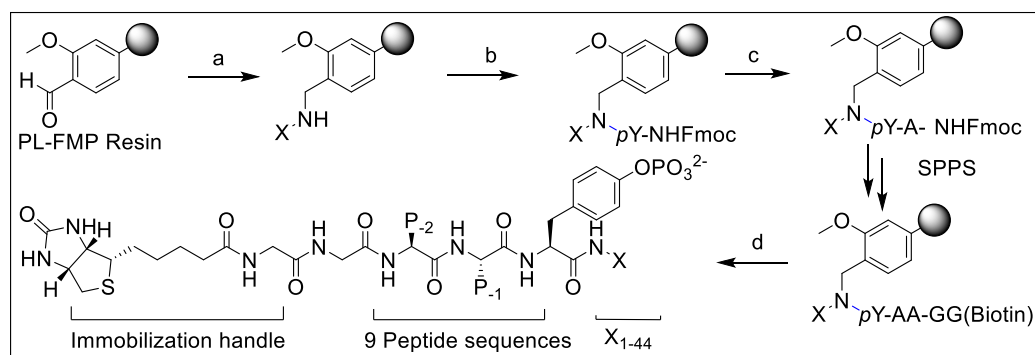
Scheme 2.1. SH2 library synthesis.[102] (a)(i) amine building block/EDC, 2 h; (ii) 1% HOAc/Na(OAc)$_3$BH, 16 h; (b) (i) Fmoc-Tyr[PO(OBzl)OH]-OH/PyBrOP/DIEA, 12 h; (c) (i) 20% piperidine/DMF 1 h; (ii) Fmoc-AA-OH/HOBt/HBTU/DIEA/DMF, 12 h; (d) TFA/TIS/H$_2$O (95 : 2.5 : 2.5).

2.3 Microarray Fabrication and Screening

Microarrays were spotted with the 396-member library in duplicate. To confirm immobilization, slides were stained with a phosphor specific binding dye, ProQ Diamond®. All spots displayed a positive signal, confirming

immobilization of the library across the slides (Figure. 2.3). The intensities of the spots were not uniform. This phenomenon was observed in previous experiment. It might be because of the residues next to phosphor-tyrosine that affected the binding affinity of ProQ. Proteins were expressed in *E.coli* with hexa-histidine tags, and were labelled on beads with Cy5-NHS (GE Healthcare), before being eluted with imidazole and size separated (using a G-25 Nap column, GE Healthcare), facilitating labelling and purification in a combined step. A total of 15 fluorescent labelled SH2 proteins were prepared in this manner (Figure 2.4), and screened against this small molecule library on microarrays. A 10 μM concentration of protein was applied to each slide. Following a wash cycle with PBS containing 0.1% Tween, the fluorescent scanned images were obtained (Figure 2.5). The microarray data was extracted, background subtracted, normalized, and displayed in a coloured heatmap (Figure 2.6). The labelling efficiency varied among different SH2 domains. We further confirmed that when the arrays were dephosphorylated through alkaline phosphatase treatment, no binding was observed. This highlighted that binding to the SMM was dependent on SH2 domain interactions, as well as on the presence of a phosphorylated Tyr residue (data not shown).
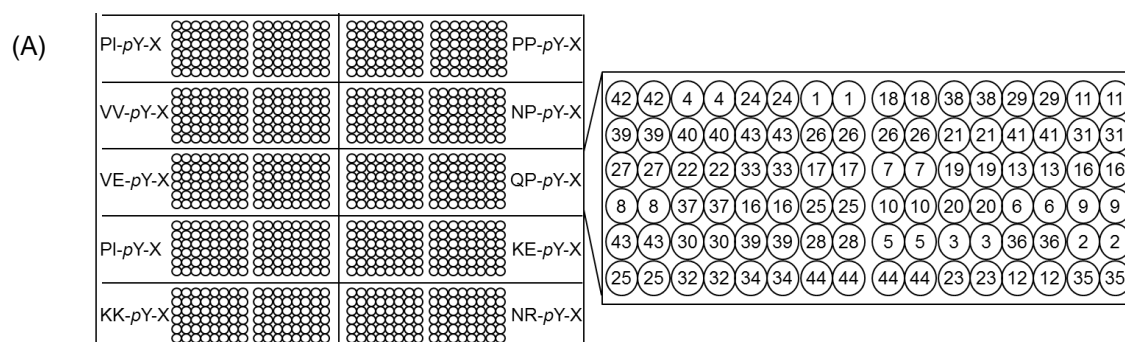
(B)

Figure 2.3. Microarray fabrication.[102] (A) Print code of the 396 library (B) ProQ image of 396 member library.
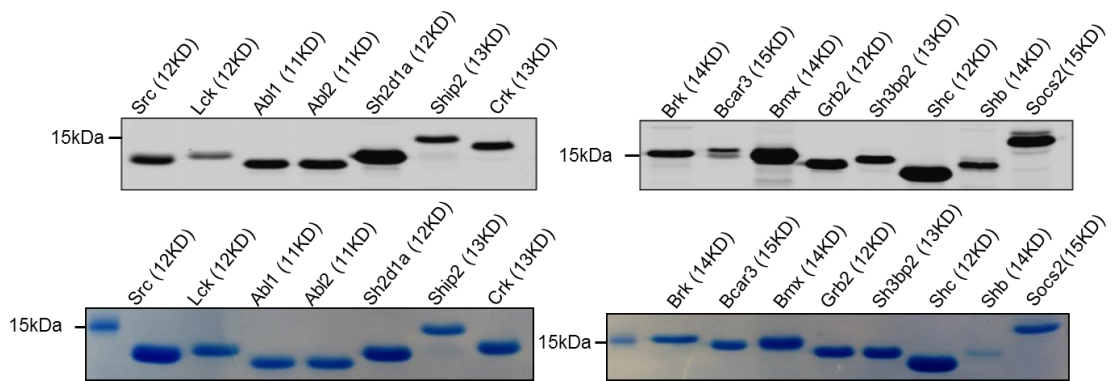
Figure 2.4. Florescent gel of SH2 domains purified and labelled by "one-step" labelling.[102]
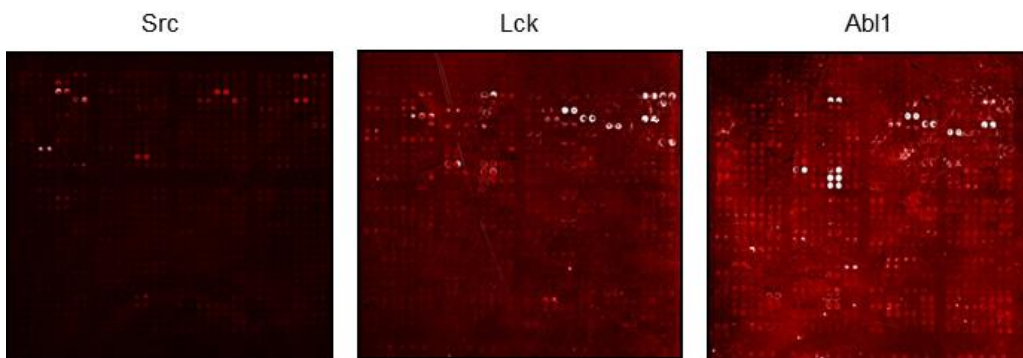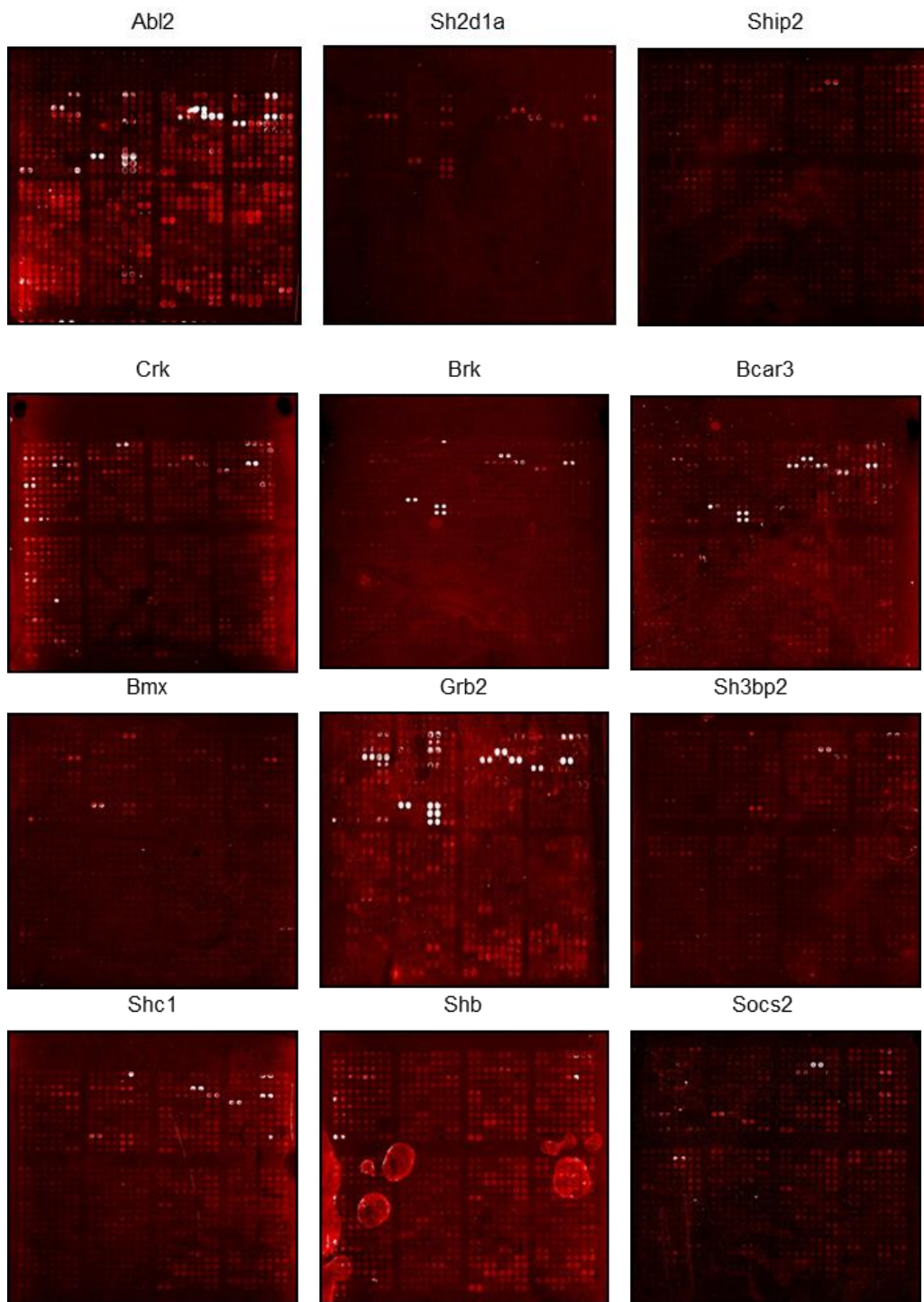
Figure 2.5. Microarray image of 15 SH2 proteins[102]

2.4 Ligand Fingerprint

The complete dataset is shown in Figure. 2.6 as a coloured heat map. Each of the 15 SH2 domains exhibited characteristic fingerprints. Based on

these results, several broad conclusions could be drawn. Generally it appeared that positions with N-terminal PP, VV or PI sequences generated unique hits, more so than any other N-terminal sequence. Also, the heat map showed that the class II SH2 domain containing proteins exhibited a binding pattern that was rather distinctive.
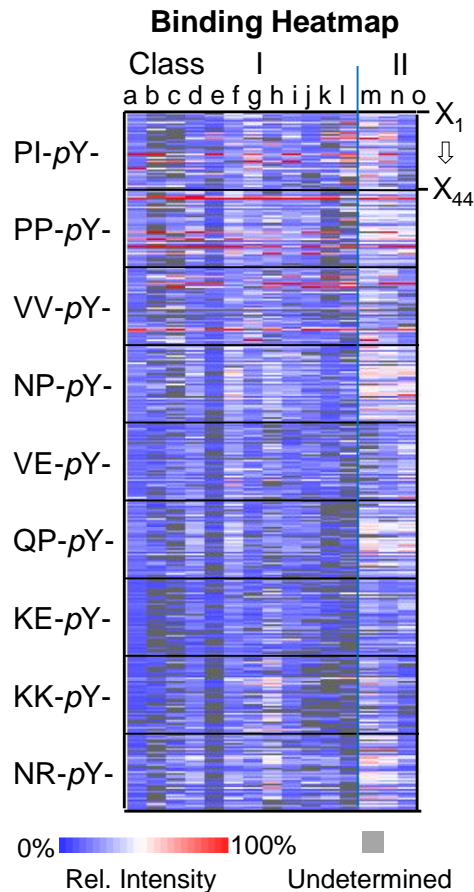
Figure 2.6. Coloured heat maps displaying intensities with each of the ligand (scale inset).[102] Src (a), Lck (b), Abl1 (c), Abl2 (d), Sh2d1a (e), Ship2 (f), Crk (g), Brk (h), Bcar3 (j), Bmx (i), Grb2 (k), Sh3bp2 (l), Shc1 (m), Shb (n), Socs2 (o).

As the dataset was extensive, we further simplified the results for analysis by first averaging the complete microarray dataset formed by contributions of signals from the respective N-terminal and C-terminal substituents (Figure. 2.7). When this analysis was performed by grouping the

SH2 domains into their classes (12 proteins in class I and 3 proteins in class II, Figure 2.1), it was revealed that a common set of C-terminal substituents for class I proteins, namely **12**, **25**, **36** and **37**, which were long chain ring containing compounds, while **1**, **5**, **6** and **10** which contained cyclic rings linked proximately to amine displayed good affinity across the proteins tested. Compared to those of class I, class II proteins displayed a distinct preference for binding to aromatic groups with short chains such as **16**, **17**, **22**, and **31**. This is consistent with a previous study claiming that class II SH2 domains prefer a hydrophobic residue at the P1 position.[99]
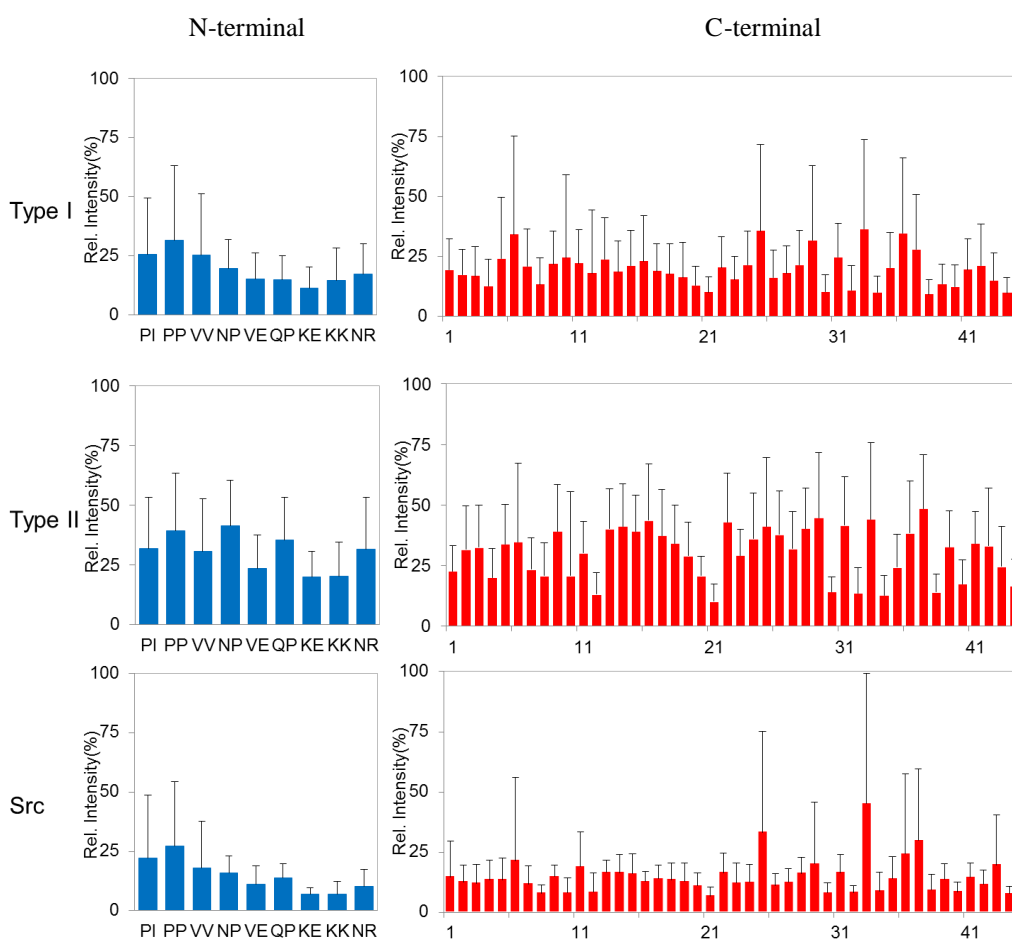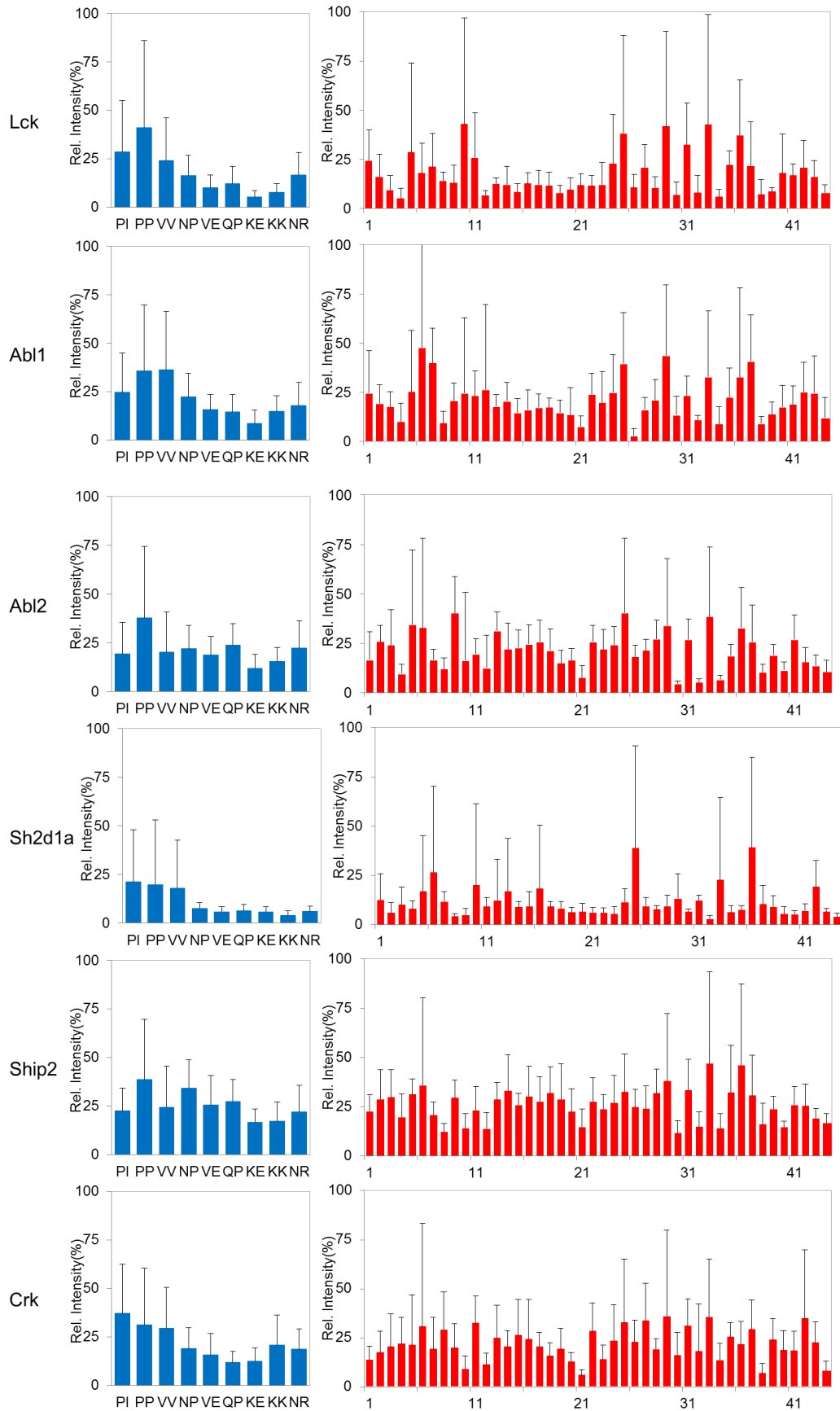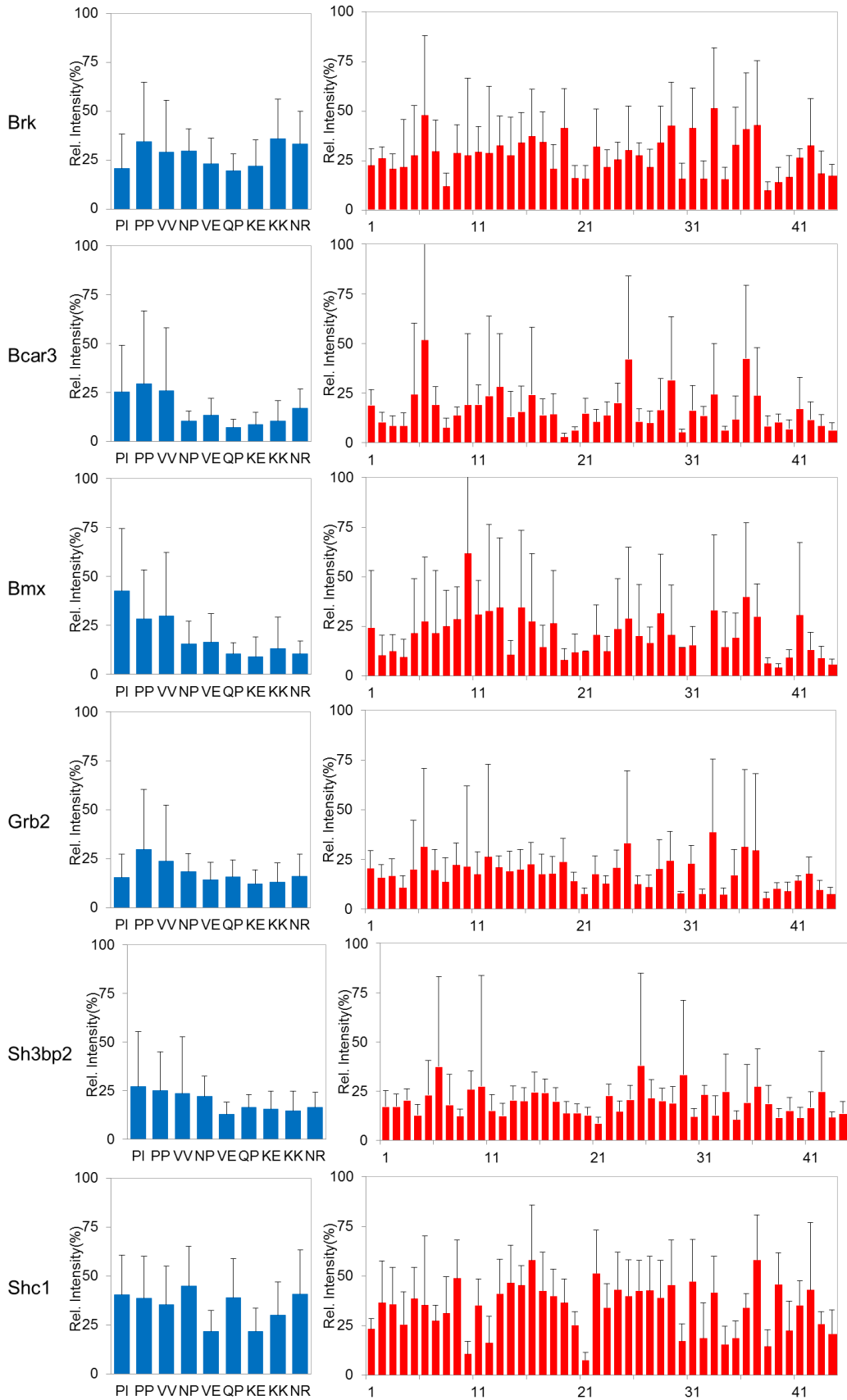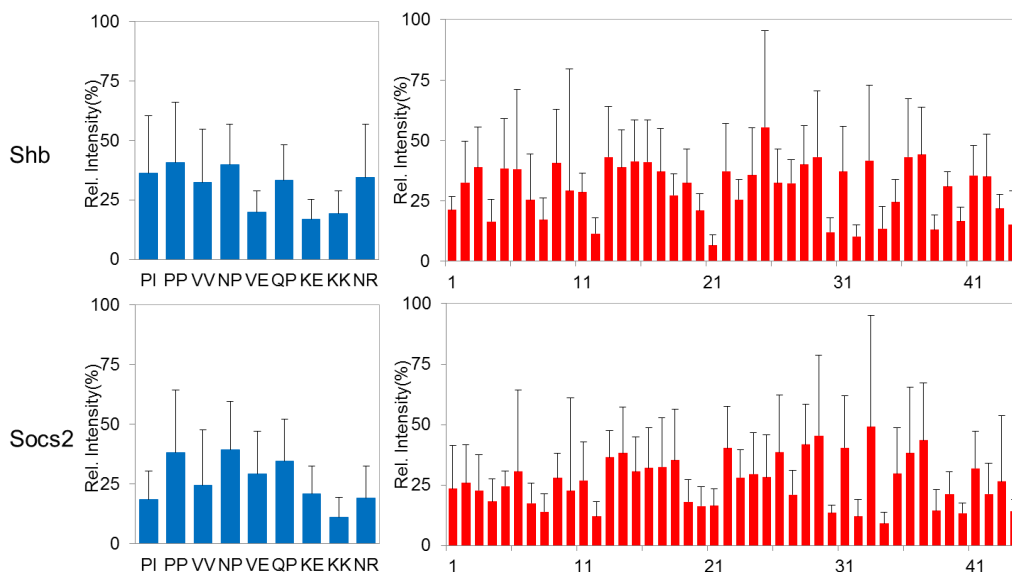
Figure 2.7. Average binding contributions across N-terminal peptide sequences (Left) and 44 C-terminal small molecule building blocks (Right). Each bar represents averaged binding across in the library presenting the relevant peptide sequences. The error bar denotes the standard deviation across each group.[102]

In addition, these averaged results also showed that the N-terminal positions also contributed significantly to the binding, as evidenced by the sizeable error bars (Figure 2.7). Across class I SH2 proteins, PP, PI or VV sequences had a clearly higher contribution to overall binding, compared to the other 6 N-terminal sequences tested. For class II SH2 domains, there was a broader preference with proline containing P1 members contributing towards the strongest relative binding. This result was consistent with references describing favourable contribution of hydrophobic residues on the N-terminus to overall inhibitor potencies.[99,103] To further resolve the groups according to binding preference, hierarchical clustering was carried out. This, unlike sequence homology clustering usually performed,[90] categorizes SH2 domains according to their binding preferences against ligand libraries.[104]
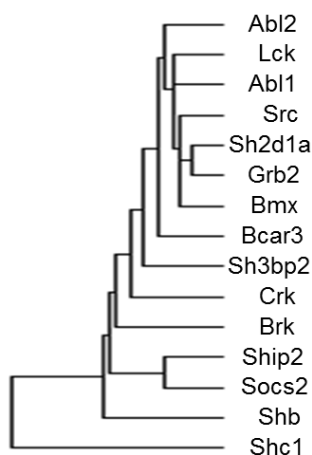
Figure 2.8. Cladograms of SH2 domains based on binding fingerprints[102]

Our clusters generally grouped class I and class II proteins separately (Figure 2.8). Class I SH2 domains all contained a Tyr or a Phe aromatic residue at βD5, which has been shown to influence binding at P1 and P3 positions.[96] These included namely Src, Abl1 and Lck. Lck, Src, Abl1 and Abl2 from group I share a common pY-E(L/E)I binding preference,4b which can account for the similarity observed. Shb, Socs2 and Shc1were grouped most distantly from the majority of other group 1 SH2s, as might have been expected. In particular, Socs2 and Ship2 formed a sister pair indicating a highly similar binding profile, as did Grb2 and Sh2d1a, highlighting the clustering based on the microarray fingerprints correlated closely with traditional substrate grouping across these SH2 domains.[99]

2.5 Ligand Identification and Validation

In addition to fingerprinting, we also identified several strongly binding molecules that appeared to be potential hits. They showed significant levels of fluorescence on the microarray screens, across various SH2 proteins. A total of

14 such hits were identified for follow-up validations and tests (Table 2.1). These hits were resynthesized, purified and LC-MS analysis was conducted, to confirm purity and identity. These molecules were re-spotted on smaller sub-grids of avidin slides, and subjected to concentration-dependent applications of various SH2 proteins. As a result, apparent K_D values were computed for these hits against the various SH2 domains, as shown against Lck (Figure 2.9). Based on these results, PP-pY-X$_5$ and PP-pY-X$_{25}$ were identified as hits that bound to both Lck and Grb2, and PI-pY-X$_1$ appeared to be a potentially selective hit for Grb2.

Table 2.1. Structures of 14 potential binders identified[102]

Figure 2.9. Concentration dependent application and apparent K_D results for selected hits.[102] (A) Grid arrangement. (B) Pro-Q image of the potential binding hits indicating successful immobilization (top grid in green). Concentration-dependent screening of the microarray against Cy5-labeled Lck(bottom grids in red). (C) K_D analysis of selected compounds. (D) Summary of K_D values obtained from the microarray experiment. TMR - tetramethylrhodamine (spotted dye reference).

Further validation was carried out by a protein thermal shift assay using thermal shift measurements for both Lck and Grb2, using methods previously described.[105] Lck without a ligand acted as a negative control, with the melt temperature of about 55.4 1 °C. After adding the ligand, the melting point shift increased due to the stability of the protein–ligand complex. Compared to the melting point of the negative control, all ligand–protein complexes presented a right shift ranging from 55.7 to 58.6 1 °C. The shift range was proportional to the binding potency of the hits. For instance, PI-pY-X1 produced a small thermal shift (0.3 °C) while the stronger Lck binding hits PP-pY-X$_5$, PP-pY-X$_6$ and PP-pY-X$_{25}$ contributed to larger shifts (~3 °C). The results indicated that

the latter three hits were true binders. Similarly, the thermal shift results of Grb2 matched that expected from the microarray results. Among the hits tested, PP-pY-X$_5$ and PP-pY-X$_{25}$ and PI-pY-X$_1$ presented as hits for Grb2. Taken together, PP-pY-X$_5$ and PP-pY-X$_{25}$ broadly bound Grb2 and Lck, while PI-pY-X$_1$ bound Grb2, but only weakly bound Lck (Figure 2.10). Future docking or co-crystallization experiments could confirm the binding pockets and orientation of these hits.
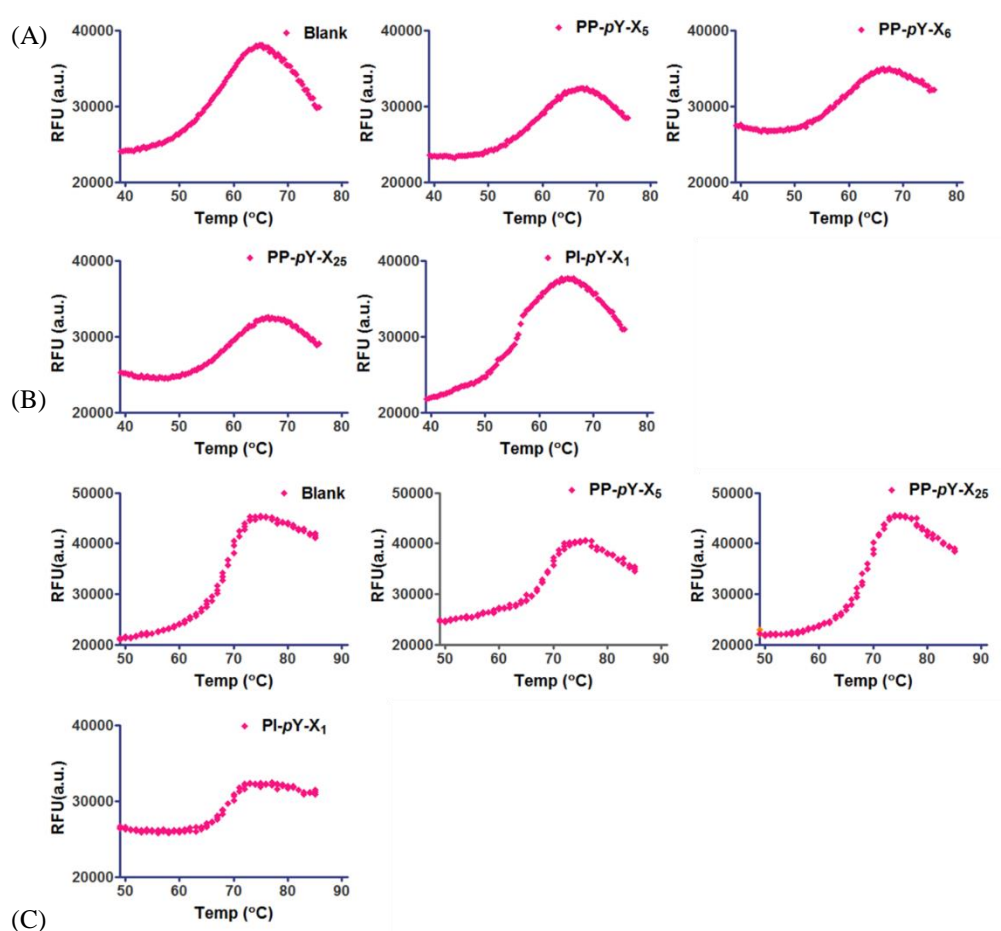
ID	Lck(40μM)		Grb2(40μM)	
	T_m(°C)	ΔT_m(°C)	T_m(°C)	ΔT_m(°C)
Control	55.4		67.6	
PP-pY-X$_5$	58.4	3.0±0.14	69.2	1.5±0.40
PP-pY-X$_6$	58.2	2.8±0.11	N.D.	N.D.
PP-pY-X$_{25}$	58.6	3.2±0.12	68.5	0.9±0.10
PI-pY-X$_1$	55.7	0.3±0.10	69.4	1.8±0.13

Figure 2.10. Results of thermal shift assay.[102] (A) Representative unfolding curves of Lck-SH2 domain obtained from protein thermal shift assay. (B) Representative unfolding curves of Grb2-SH2 domain obtained from protein thermal shift assay. (C) Result of protein thermal shift assay

2.6 Conclusion

In conclusion, we have successfully fabricated a small molecule microarray platform which enables, for the first time, convenient profiling of SH2 domains with small molecule building blocks. In the present format, the binding profiles were sufficiently distinct to produce informative fingerprints that clustered the various SH2 domains according to their expected functional classes. In addition, we have identified putative small molecule hits against Grb2 and Lck. These peptide-hybrid compounds, though not explicitly demonstrated herein, might in future be converted into cell permeable small molecule protein-protein interaction (PPI) inhibitors.[106] It will be possible to monitor slight differences in binding using a multi-coloured labelling and application strategy.[107,108] Overall this SMM strategy offers an attractive approach for performing global selectivity screens of SH2 domains against a much larger set of diversity elements.

Chapter 3.

Accelerated cellular on- and off-target screening of bioactive compounds using microarrays

Summary
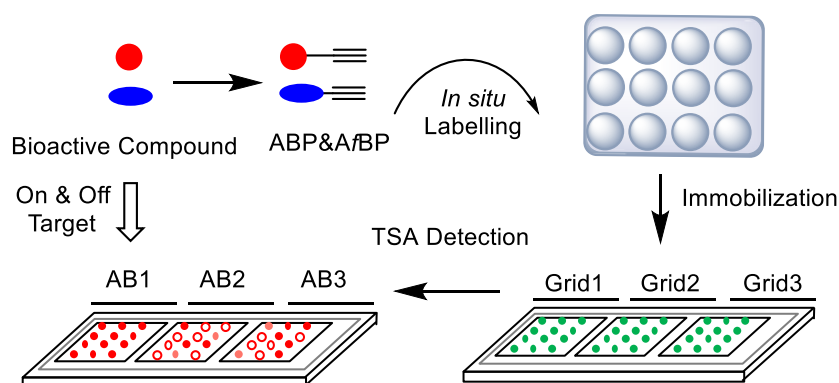
In this chapter, *in situ* proteome labeling was carried out with 9 drug-like probes in live mammalian cells, with the corresponding cellular targets captured on microarrays and simultaneously screened using a diverse set of antibodies, revealing potential on- and off-targets.

3.1 Introduction

Drug discovery is a long-drawn and expensive process, partly due to the unexpected behaviour and off-target binding spectrum of drug candidates.[109] Often, side-effects become evident only at the late stages beyond lead development, particularly during pre-clinical and clinical trials, by when failures are costly. The ability to perform broad-based, proteome-wide screening at the outset of drug discovery introduces a new paradigm, where the

target binding spectrum and key biological pathways affected are revealed and, if necessary, tuned upfront. Such a drug development approach will enhance candidate success, and also provide an informative toolbox for pre-empting multi-drug reactions and even facilitating combination drug therapy.[110,111] In the last several years, by borrowing from concepts in the field of activity-based protein profiling (ABPP),[11] *in situ* drug profiling has become possible using drug-like chemicals minimally modified from their parental compounds. Such drug-like probes enable the large-scale interrogation of protein−small molecule interactions and can rapidly identify potential cellular targets, both on- and off-targets, in live cells.[112-114] One of the main challenges, however, is with the target validation step (Figure 3.1A; workflow on the right); upon protein labelling, pull-down (PD) experiments are usually required, followed by Western blot (WB) analysis, to confirm the genuine *in situ* target-probe interactions. The process is cumbersome, expensive and low-throughput.[115-117]

Microarrays are ideal for selectivity screening, because of their low cost, vast scalability and throughput.[5,46] However, there exist no approaches for microarrays to contribute to functional drug selectivity screening *in situ*. Herein we report a variant of reverse phase protein microarrays,[118] to unite the *in situ* drug profiling and microarray approaches for the first time, in order to rapidly identify and/or validate on- and off-targets of bioactive compounds (Figure 3.1A; left).

3.2 Microarray Fabrication

The proposed microarray target screening approach is depicted in Figure 3.1A. Nine bioactive compounds were first minimally tagged with an alkyne handle, with or without an alkyl diazirine photo-crosslinker to generate the corresponding affinity-based probes (A*f*BPs) or activity-based probes (ABPs), respectively (Scheme 3.1).[117,119,120] The A*f*BPs/ABPs were incubated with recombinant proteins, spiked cell lysates or living cells, to initiate protein labelling. Thereafter, the labelled reaction was clicked with an azide-containing trifunctional linker (Figure 3.1B), using copper-catalyzed alkyne-azide cycloaddition (CuAAC).[121,122] As the trifunctional linker also contained rhodamine and biotin units (for imaging and protein capture, respectively), the resulting samples could be directly spotted onto avidin-coated slides, where labelled proteins were captured across multiple sub-grids. The arrays were then probed with antibodies (Abs), representing putative targets. To ensure high-sensitivity detection, tyramide signal amplification (TSA) was employed.[11] Potential advantages of this strategy include 1) scalability and throughput in probing drug-target interactions *in vivo*, 2) simultaneous and quantitative comparisons across various dimensions (protein targets, compounds and cell types), and 3) ease of use with minimal quantities of reagents and antibodies required. Similar experiments, if conducted by classical PD experiments, followed by in-gel fluorescence scanning and WB analysis (Figure 3.1A; right), remain cumbersome, and require large quantities of expensive antibodies.
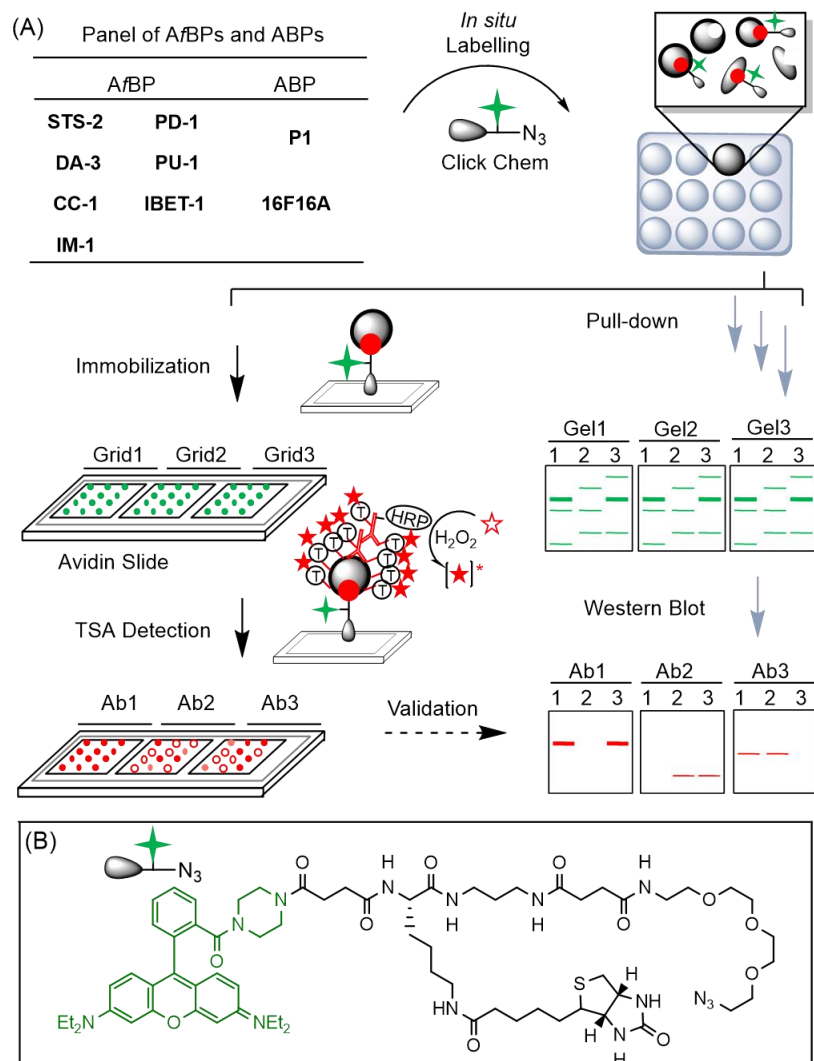
Figure 3.1. (A) Overall strategy of the microarray-guided, target profiling/validation of drug-like compounds. (Left) Following click reactions and immobilizations on to microarrays, the samples were probed with multiple antibodies (Abs) and detected with tyramide signal amplification (TSA). (Right) Traditional PD/WB workflow for target validation. (B) Structure of the tri-functional linker containing rhodamine, azide and biotin units.

The 9 probes used in this study were derivatives of bioactive compounds targeting protein kinases (**STS-2, DA-3, PU-1, PD-1, IM-1** and **CC-1**),[117] protein disulfide isomerases (PDI) (**P1** and **16F16A**) and bromodomains (**IBET-1**) (Scheme 3.1).[119,120] These classes of proteins are vital targets for a host of human diseases, ranging from cancer to neurodegenerative disorders.

The compounds were modified minimally from the original inhibitor using an aliphatic alkyne linker, which in the case of reversible inhibitors also included a diazirine moiety to facilitate covalent, UV dependent, target capture. The linker was inserted at the solvent accessible site, which preserved target binding properties of the compounds.[117,119,120]
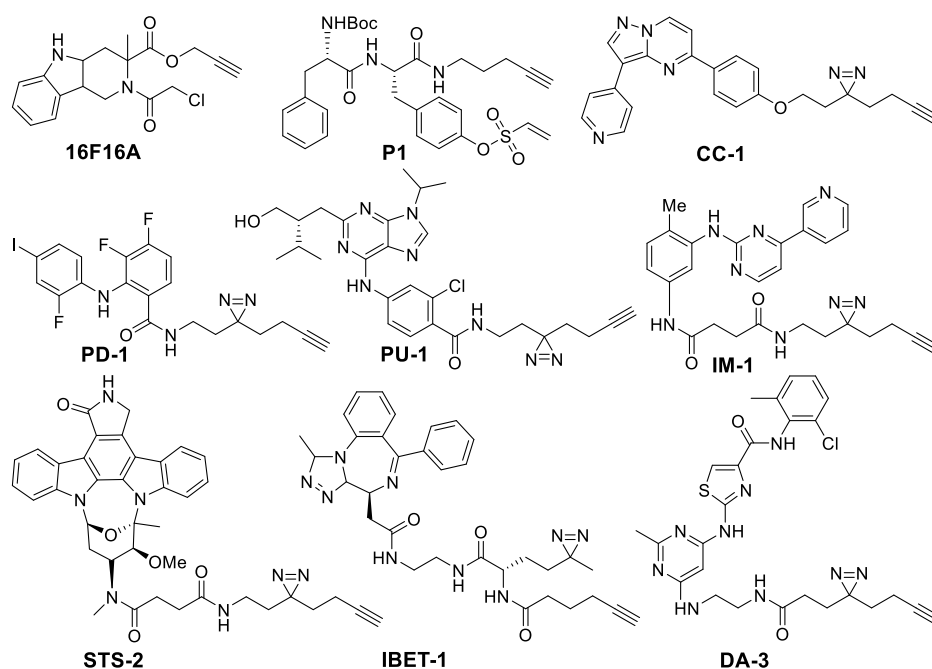
Scheme 3.1. Structures of 9 minimally modified drug-like compounds for target profiling. All probes were previously reported.[117,119,120]

3.3 Pure Protein and Spiked Lysate Profiling

To prove the concept, we first selected a protein tyrosine kinase, c-Src and its corresponding antibody, together with 3 A*f*BPs (**STS-2**, **DA-3** and **IBET-1**). c-Src was expected to bind both **STS-2** and **DA-3** while **IBET-1** was expected to exhibit low or neglibible affinity with this biological target. c-Src, at a concentration of 1 mg/ml, was incubated with the probes. Following UV initiation and the click reaction,[115-117] excess unreacted probes were removed by

acetone precipitation. The labelled samples were serially diluted and spotted onto avidin-coated glass slides in a concentration dependent manner (1, 10, 100, 1,000, 10,000 ng/ml).[104] Following a 2 h incubation, the slides were extensively washed with TBST (0.5% Tween20), probed with *anti*-c-Src antibody, and visualized with TSA-Cy5 detection system.[35] The set of experiments were conveniently replicated across multiple subgrids on the same slide, with good reproducibility (Figure 3.2A). The array results were also plotted graphically (right); c-Src was shown to be captured most effectively by **STS-2**, followed by **DA-3**, with readable signals even at the lowest 1 ng/ml concentration. At the highest concentration of c-Src used (i.e. 10,000 ng/ml), only slight signals were detectable with the **IBET-1** probe. Nevertheless, there was a clear concentration-dependent trend. (Figure. 3.2) These binding profiles with each of the 3 probes were consistent with in-gel fluorescence scanning results (Figure 3.2), which was obtained by detecting the rhodamine from the trifunctional linker. While this initial experiment was able to produce good sensitivity, at concentration ranges of proteins we might expect *in vivo*, we anticipated that performance could vary in response to the sensitivity and specificity of antibodies used, and were hence selective in the choice of the antibodies we used
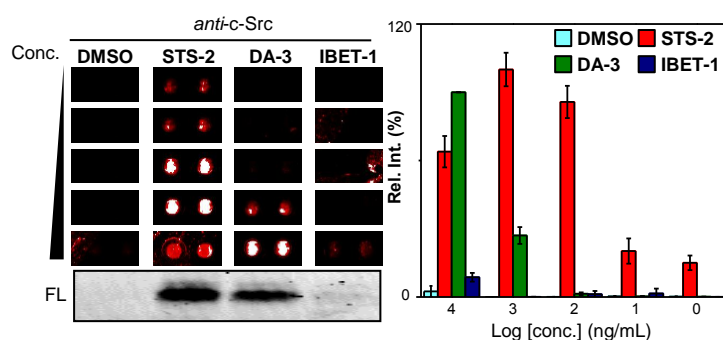
Figure 3.2. Recombinant c-Src was labelled with three probes (**STS-2**, **DA-3** and **IBET-1**) and a DMSO control, captured on the microarray and visualized with *anti*-c-Src Abs and a TSA Cy5 kit. Bottom to top (ng/ml): 10,000, 1000, 100, 10, 1. Array was scanned under PMT 140. In-gel fluorescence scanning of the samples was analysed in a gel format (bottom left).

We next performed spike-in experiments, to assess the potential to perform this capture and profiling within the context of a complex cellular proteome. HepG2 lysates were spiked with recombinant c-Src at varying concentrations (1, 10, 100, 1,000, 10,000 ng/ml). Each of these mixtures were treated with 3 different probes respectively (Figure 3.3A). Encouragingly, the results were generally consistent with those obtained with pure protein and in-gel fluorescent scanning. Background signals were evident with DMSO at 10,000 ng/ml c-Src spiked concentrations, due to non-specific capture of proteins on the array even after extentively washing. However, the background was negligible at lower protein target concentrations (Figure 3.3A). Both **DA-3** and **STS-2** were able to capture c-Src from complex cell lysates, with signals diminished with decreasing quantities of spiked protein. Under a higher PMT to improve signal to noise, it was observed that as low as 10 ng/ml of c-Src in the lysate could successfully be detected with **DA-3** by using the array strategy (Figure 3.3B). Based on the database, the concentration of endogenous c-Src ranged from 1 to 100 ng/ml. The sensitivity of our strategy appeared to be sufficient to detect the endogeous c-Src. We next compared the sensitivity of microarray system with that of the fluorescent gel. The sensitivity attained on microarrays was around 10 ng/mL, appearing to be one order of magnitude higher than that obtained through imaging on fluorescent gels (Figure 3.3B), and this is likely attributable to the TSA strategy adopted, which offered signal

enhancement. In addition, the spectrum of c-Src profiles in lysate screened with the other 6 probes remained were largely within expectations. Another kinase probe (**IM-1**) was able to capture c-Src and showed detectable signals on the arrays (Figure 3.4A, B). In addition, **P1** a probe that was designed for PDIs, was capable of binding c-Src within cell lysates, while **PU-1** and **16F16A** presented weaker binding with c-Src. **PD-1** and **CC-1** generally did not bind c-Src, even at high protein concentrations.
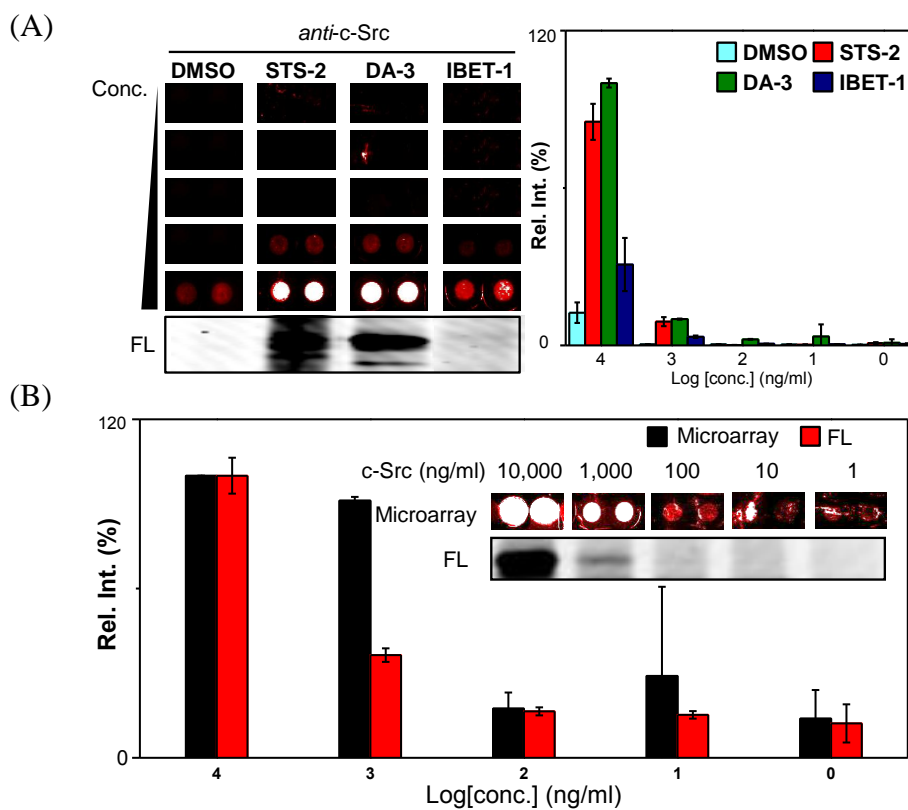
Figure 3.3. Spike-in lysate labelling and profiling. (A) HepG2 lysates (1 mg/ml) spiked with the same concentration gradient of recombinant c-Src (10,000, 1,000, 100, 10, 1 ng/ml) was labelled with three probes (**STS-2**, **DA-3** and **IBET-1**) and a DMSO control, captured on the microarray and visualized with *anti*-c-Src Abs and a TSA Cy5 kit. Bottom to top (ng/ml): 10,000, 1000, 100, 10, 1. Array was scanned under PMT 140. In-gel fluorescence scanning of the samples was analysed in a gel format (bottom left). (B) Concentration-dependent gel and microarray results of recombinant c-Src labelled with **DA-3**. The corresponding fluorescent gel (inset) was imaged in the Cy3 channel. The error bars represent standard deviations across two independent experimental replicates. Array was scanned under PMT 180.
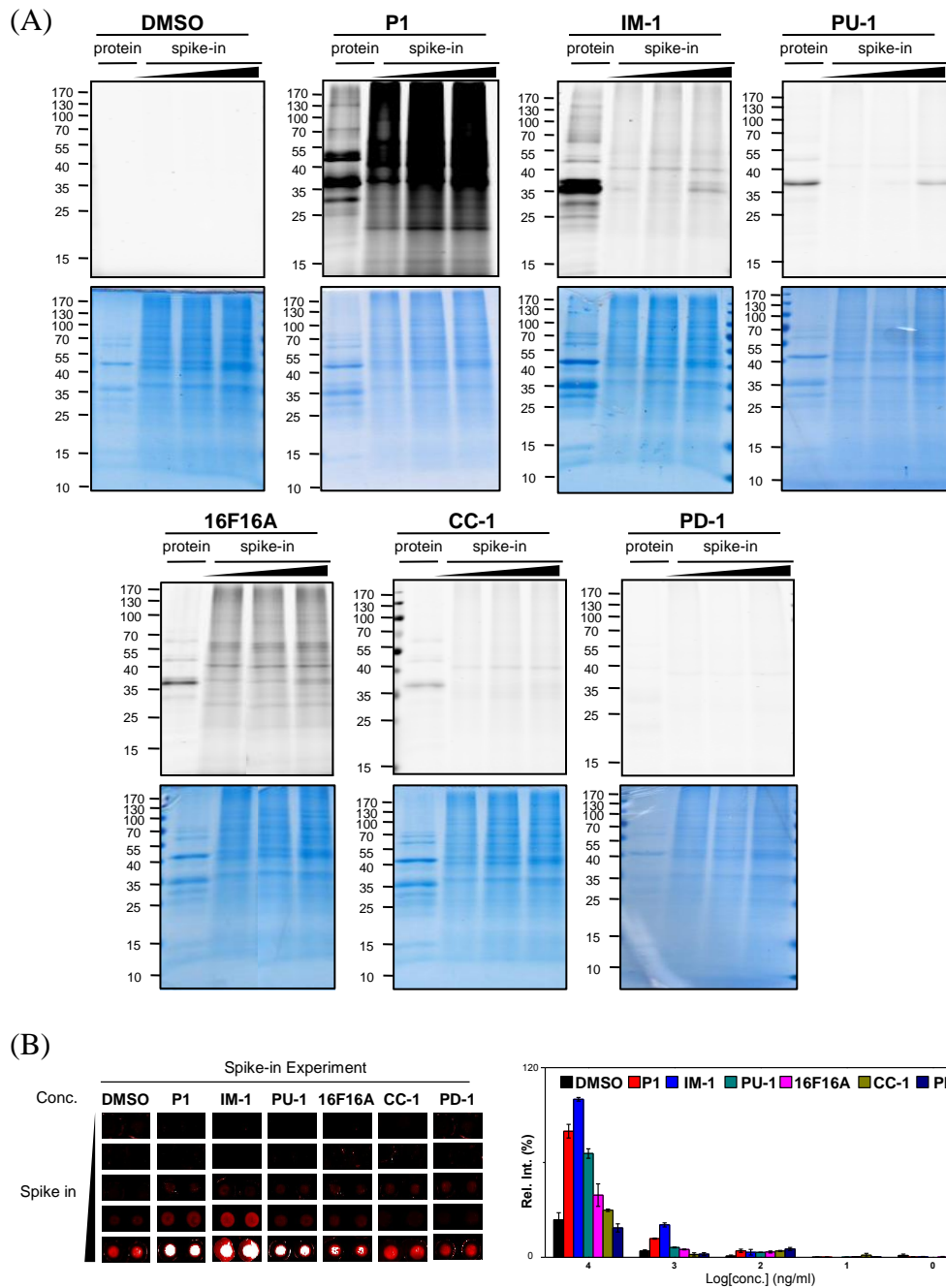
Figure 3.4. (A) Recombinant c-Src and spiked cell lysate (1, 100, 10,000 ng/ml) are labelled with **P1**, **IM-1**, **PU-1**, **16F16A**, **CC-1** and **PD-1**. Fluorescent gel (Top) and Coomassie stain (Bottom). (B) HepG2 lysate samples were capture on microarry and visualized with anti-c-Src Abs and a TSA Cy5 kit. Bottom to top (ng/ml): 10,000, 1000, 100, 10, 1.

3.4 *In Situ* Labelling Profiling

Having demonstrated that we could successfully adopt this strategy *in vitro* with c-Src, we went ahead to also test this approach in live cells. To increase the captured targets as much as possible, we used our compounds at high concentration (10 μM), with HepG2 cells under standard growth conditions. The cells were then isolated and lysed. The resulting labelled proteome was immobilized onto slides at optimal concentration (2 mg/ml) and detected against a repertoire of six antibodies. The complete array result was summarized as a coloured heat map (Figure 3.5). Based on these results, some initial conclusions could be drawn. Firstly, each antibody exhibited distinctive binding fingerprints, and each probe displayed unique binding affinity to corresponding targets. Secondly, most probes exhibited potent binding ability with their potential targets. For instance, **PU-1** and **DA-3** successfully captured their respective targets CDK1 and c-Src (Figure 3.5). **STS-2** appeared to capture a range of kinases (PKA, CDK1, c-Src and MEK1), as should be expected.[123] **16F16A** successfully bound to its target PDI, as did **P1**. **CC-1** consistently exhibited weak binding against the range of protein targets.
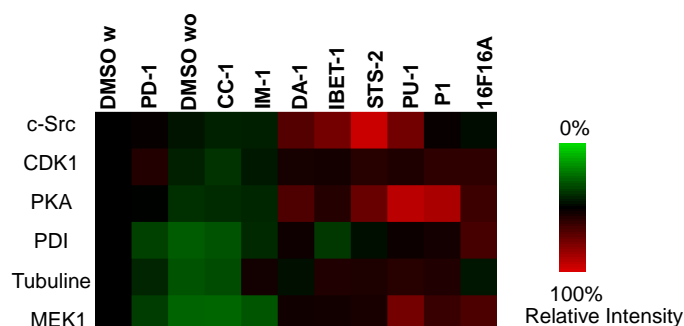
Figure 3.5. Heatmap of nine probes against six antibodies.

3.5 Target Validation

Figure 3.6. *In situ* labelling with 9 probes, captured on the microarray and target validation (A) Microarray image of 6 probes profiling against 4 selective antibodies (B) Pull-down result of corresponding probes for target validation. (C) Preliminary off-targets validation for **P1** (left) and **PU-1** (right). Concentration dependent labelling of **P1** and **PU-1** with recombinant PKA.

To further evaluate the accuracy of heatmap obtained from microarrays, *in situ* labelling and PD/WB target identification with 4 of the antibodies and six of the probes were carried out under same conditions (Figure 3.6A, B). The results were largely consistent, as visualized in the similar heatmap patterns. For instance, both **STS-2** and **DA-3** exhibited interactions with c-Src. This was consistent with spike-in experiments. Apart from c-Src, **STS-2**, also interacted with PKA and CDK1, as is characteristic of staurosporine (STS), being a broad-spectrum kinase inhibitor. **P1** and **16F16A**, as the strongest PDI probes based on microarray results, exhibited strong interaction with PDI in PD experiment

while the rest of the probes consistently showed weak binding to PDI, confirming the microarray result. In addition, **P1** also exhibited interactions with various kinases including PKA, c-Src and CDK1. **P1** displayed a much wider spectrum of interactions than initially expected, potentially due to its electrophilic phenyl vinyl sulfonate. Similarly, the results of **PD-1** and **PU-1** from WB matched those from microarray. In particular, **PD-1** bound CDK1. **PU-1** as a probe for CDK1, exhibited binding interactions with PKA and c-Src as well. With these preliminary results, we picked up both ABP (**P1**) and A*f*BP (**PU-1**) for further validation. As for **P1**, *in vitro* concentration dependent labelling of recombinant PKA was carried out. As shown in Figure 3.6C, **P1** was able to label PKA, with signals detectable even at low concentrations of 0.1 μM, suggesting strong interactions between **P1** and PKA. Similarly, **PU-1** bound with recombinant PKA at a low concentration of 0.2 μM (Figure 3.6C).

3.6 Conclusion

In conclusion, we have developed the first microarray-based platform to our knowledge, capable of rapidly identifying on- and off-targets of drug-like compounds from live cells. Compared to gel-based scans and WB, microarrays consume less samples but can sensitively detect endogenous proteins. In this work, we have successfully profiled a panel of drug-like ABP and A*f*BP bioactive compounds against various protein targets, and have likely identified the off-targets of ABP (**P1**) and A*f*BP (**PU-1**). It should be highlighted that the quality of the antibodies used are critical to ensure robust interpretations of only the target proteins in the array. Overall, this microarray-based strategy would enhance and complement high-content screening approaches, for early stage

lead-optimization, and off-target screening, in drug discovery applications. A single experiment thus performed on a microarray slide could replace dozens of WBs/PDs, and consequently, accelerate target profiling.

Chapter 4.

Intracellular Delivery of Functional Proteins and Native Drugs by Cell-Penetrating Poly(disulfide)s

Summary
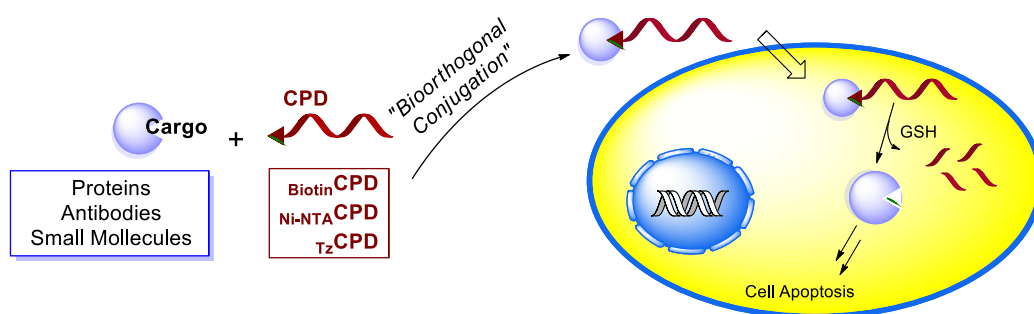
The efficient delivery of bioactive compounds into cells is a major challenge in drug discovery. This chapter summarized the development of novel methods for intracellular delivery of functional proteins (including antibodies) and native small molecule drugs by making use of cell-penetrating poly(disulfide)s, or CPDs. CPDs were recently shown to be rapidly taken up by mammalian cells in endocytosis-independent pathways, but their applications for delivery of proteins and native small molecule drugs have not been demonstrated. With our newly developed, CPD-assisted approaches, rapid and "bioorthogonal" loading of cargos was carried out with pre-synthesized CPDs, in two-steps and in a matter of minutes under aqueous conditions. The resulting CPD-cargo conjugates were used immediately for subsequent cell delivery studies. With the versatility and flexibility of these methods, we further showed they could be used for immediate delivery of a variety of functional cargos with minimum chemical and genetic manipulations. The minimal cell cytotoxicity of

these CPDs and their cargo-loaded conjugates further highlights the unique advantage of this new cell-transduction method over other existing strategies, and ensures our entire delivery protocol is compatible with subsequent live cell experiments and biological studies.

4.1 Introduction

The efficient delivery of bioactive compounds, including nucleic acids, peptides/proteins and small molecules, into cells is a major challenge in drug discovery.[124] The difficulty is more pronounced for large molecules such as proteins and DNAs/RNAs,[125-130] but due to the hydrophobicity and poor water solubility of many small molecule drug candidates, such compounds don't enter cells readily either, without proper formulation and/or delivery ve-hicles.[131-133] To achieve intracellular delivery of proteins, a variety of methods have been developed in the last 20 years, including the use of cell-penetrating peptides (CPPs), super-charged proteins, liposomes, nanoparticles (NPs) and polymers.[134-141] Despite significant progress, these methods have their share of shortcomings, including low/variable delivery efficiency, the need for protein modification, high cytotoxicity, and perhaps most importantly, ineffective endosomal/lysosomal escape.[130] Take CPPs for example, it is well-documented that, while CPP-conjugated small- and medium-size cargos may be efficiently transduced into cells via non-endocytic pathways, large cargos such as proteins are mostly taken up by endocytosis, leading to subsequent endosomal trapping and lysosomal degradation.[142,143] To deliver hydrophobic small molecules intracellularly, chemical modifications may be used to make analogs that possess improved physicochemical and pharmacokinetic profiles,[144] but this is

costly, time-consuming and worse, often results in alteration of the compound's biological properties.3b Recent ad-vances in material chemistry have provided alternatives, where native drugs are directly "loaded" into a suitable "container" without the need of chemical modifica-tions.[133,145-147] For example, the use of small molecule-encapsulated mesoporous silica nanoparticles (MSNs) for drug delivery is worth-noting, in part due to the numerous desirable properties that MSNs possess, including high loading capacity, biocompatibility and "zero premature release".[145,146] In order to minimize unwanted leakage of the encapsulated drug and improve cellular uptake of MSNs, the surfaces of these nanoparticles are "capped" with CPPs and other forms of chemicals,[148-150] which, in most cases, also leads to severe endosome trapping and ineffective drug release.[148,149]

Herein, we focus on the development of novel methods for intracellular delivery of functional proteins (including antibodies) and native small molecule drugs by making use of cell-penetrating poly(disulfide)s, or CPDs (Figure 4.1A).[151-154] CPDs could be considered synthetic mimics of poly-arginine CPPs, in which the polypeptide backbone was replaced with poly(disulfide)s. Upon cellular uptake, CPDs are rapidly degraded in the cytosol by glutathione (GSH)-assisted depolymerization and show minimal cytotoxicity.[151-156] Importantly, Matile *et al* showed in a recent study that, CPDs made of thiol-modified small fluorophores (as initiators/cargos), a guanidinium propagating monomer (e.g. **M** in Figure 4.1A) and a terminator (e.g. **T**), rapidly enter mammalian cells via thiol-mediated pathways.[152] The major issue of endosomal trapping commonly associated with CPPs and other means of delivery was thus minimized.[154] Subsequently, the authors proposed this substrate-initiated, cell-penetrating

poly-(disulfide)s (siCPDs) may be used for intracellular delivery of other thiol-containing (or modified) cargos. This hypothesis was, however, never demonstrated experimentally. Furthermore, during siCPD synthesis, given the obligatory role of the thiol-containing initiator (**I**), the need of millimolar concentrations of initiator/monomer/terminator as well as organic co-solvents, it is not trivial how a protein could be directly used as an initiator , nor is it
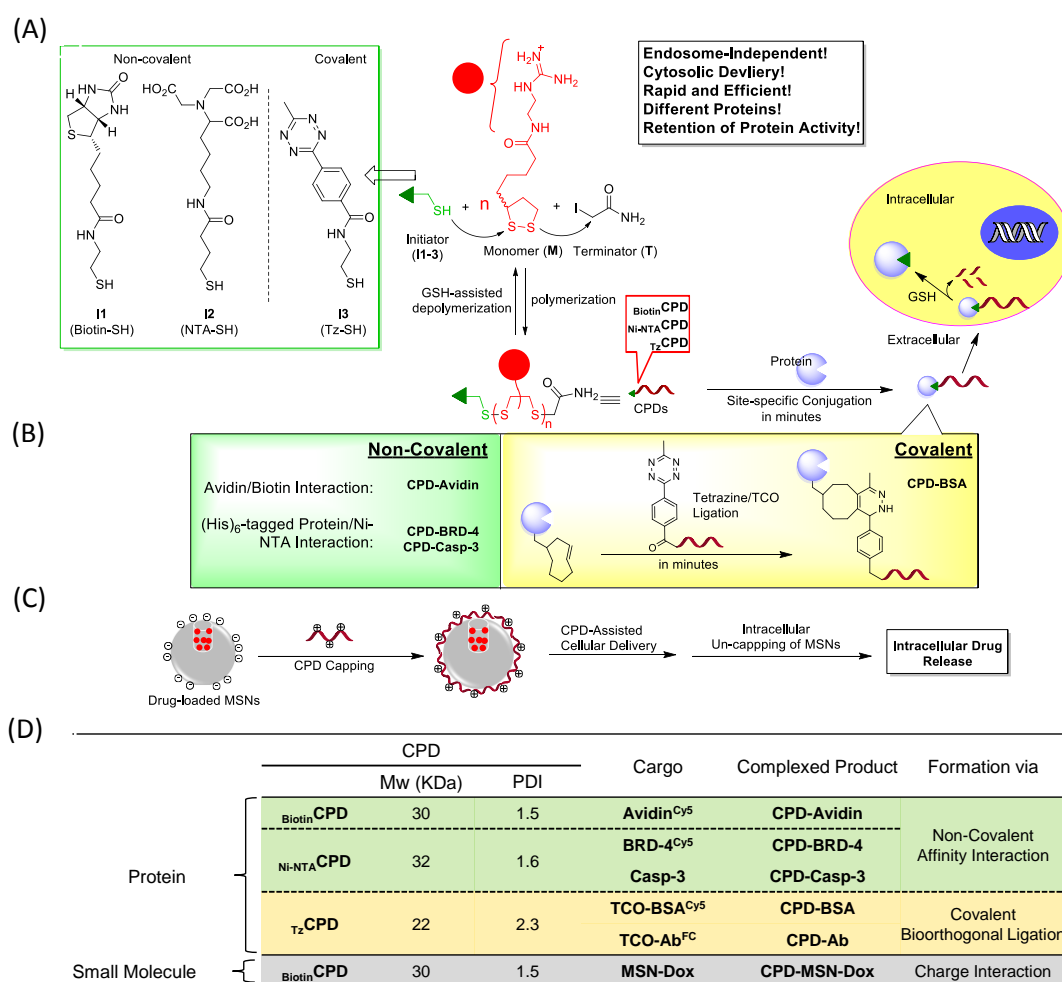
Figure 4.1. Overview of CPD-facilitated intracellular delivery of proteins (including antibodies) and native small molecule drugs.[157] (A) Newly developed initiators (**I1/I2/I3**), Monomer (**M**), terminator (**T**), the polymerization/depolymerization process of CPDs, and the two-step approach for "conjugation" of protein cargos with CPDs. (B) Summary of non-covalent and covalent approaches for bioorthogonal attachment of CPDs to proteins. The highly efficient site-specific tetrazine-TCO ligation reaction is highlighted. (C) Schematic summary of intracellular delivery of native small molecules by drug-loaded MSNs capped with **Biotin**CPD. (D) Table summary of all CPDs and cargos used in the current study.

practical to use thiol-modified small molecule drugs, as few small molecule drugs contain native thiols in their structures.

With the current work, we have confirmed that, for the first time, proteins conjugated with CPDs, either covalently (via bioorthogonal chemistry) or non-covalently (via affinity interaction), could be rapidly and efficiently delivered into the cytosol of different mammalian cells without being trapped by endocytic vesicles. Similarly, by making use of CPD-capped MSNs, we have successfully achieved intracellular delivery of native small molecule drugs (e.g. doxorubicin, or Dox). Our results indicate these novel protein and small molecule delivery methods possess the following favorable properties when compared to most existing strategies: (1) fast delivery, with cell entry in less than 15 min, (2) flexible, enabling convenient delivery of different types of cargos, (3) less cytotoxic and applicable to different types of mammalian cells, (4) efficient, facilitating cargo delivery at nanomolar concentrations, and (5) immediately available upon cell entry due to rapid CPD degradation, thus retaining the biological activity of the delivered cargo.

4.2 Design and Synthesis of New CPDs.

We were intrigued by the excellent properties of siCPDs (e.g. minimal endosome trapping and cytotoxicity), as reported by Matile *et al*,[151-153,156] and wondered whether robust, CPD-mediated methods could be developed to facilitate intracellular delivery of functional proteins, therapeutic antibodies and native small molecule drugs (that is, without any form of thiol modification). Instead of the substrate-initiated CPD synthesis as originally proposed,[152] we envisaged such cargos could be more conveniently appended, in two steps, to

pre-synthesized, functionally decorated CPDs by using suitable "conjugation" chemistries that are already available for recombinant proteins/antibodies and small molecules (Figure 4.1). Three types of thiol-containing initiators, **I1/I2/I3**, were thus designed, containing biotin, nitrilotriacetic acid (NTA) and tetrazine (Tz), respectively. Upon polymerization, the corresponding CPDs (**BiotinCPD**, **Ni-NTACPD** & **TzCPD**; Figure 4.1D), could be obtained. **Ni-NTACPD** could be attached bioorthogonally to readily available (His)$_6$-tagged proteins via non-covalent affinity interaction (K_D of Ni-NTA/(His)$_6$ < 10^{-7} M). **BiotinCPD** was designed to test whether it could be used to deliver avidin via similar non-covalent, but substantially stronger, interaction (K_D of biotin/avidin < 10^{-15} M). For a therapeutic antibody (Ab), which might not possess a his-tag,[158] we used the well-known bis-sulfone chemistry that enables site-specific introduction of a *trans*-cyclooctyne (TCO) into the native disulfide present in the antibody (*vide infra*).[159] Subsequent bioorthogonal covalent attachment of **TzCPD** to the TCO-modified antibody would result in quantitative formation of **CPD-Ab** within minutes by the highly efficient tetrazine-TCO ligation (Figure 4.1B).[160] To "append" a native small molecule drug to the CPD, we would use a positively charged CPD (**BiotinCPD** in this case) to cap the negatively charged, drug-loaded MSNs (i.e. **MSN-Dox**) by electrostatic interaction, giving **CPD-MSN-Dox** (Figure 4.1C-D); the capping of nanoparticles with CPDs is unprecedented and we were hopeful that, if CPDs could facilitate endocytosis-independent cellular uptake of large, nanometer-size cargos such as MSNs, then a variety of hydrophobic drugs could potentially be delivered intracellularly in their native form, in a controllable manner.[145-147]

All three initiators were conveniently synthesized (two of which in their respective disulfide forms; Scheme 4.1), and treated/reduced with TECP immediately prior to polymerization. The monomer (**M),** as well as subsequent polymerized products with the corresponding initiators (and capped with **T**; Figure 4.1A), were synthesized according to published protocols.[152]
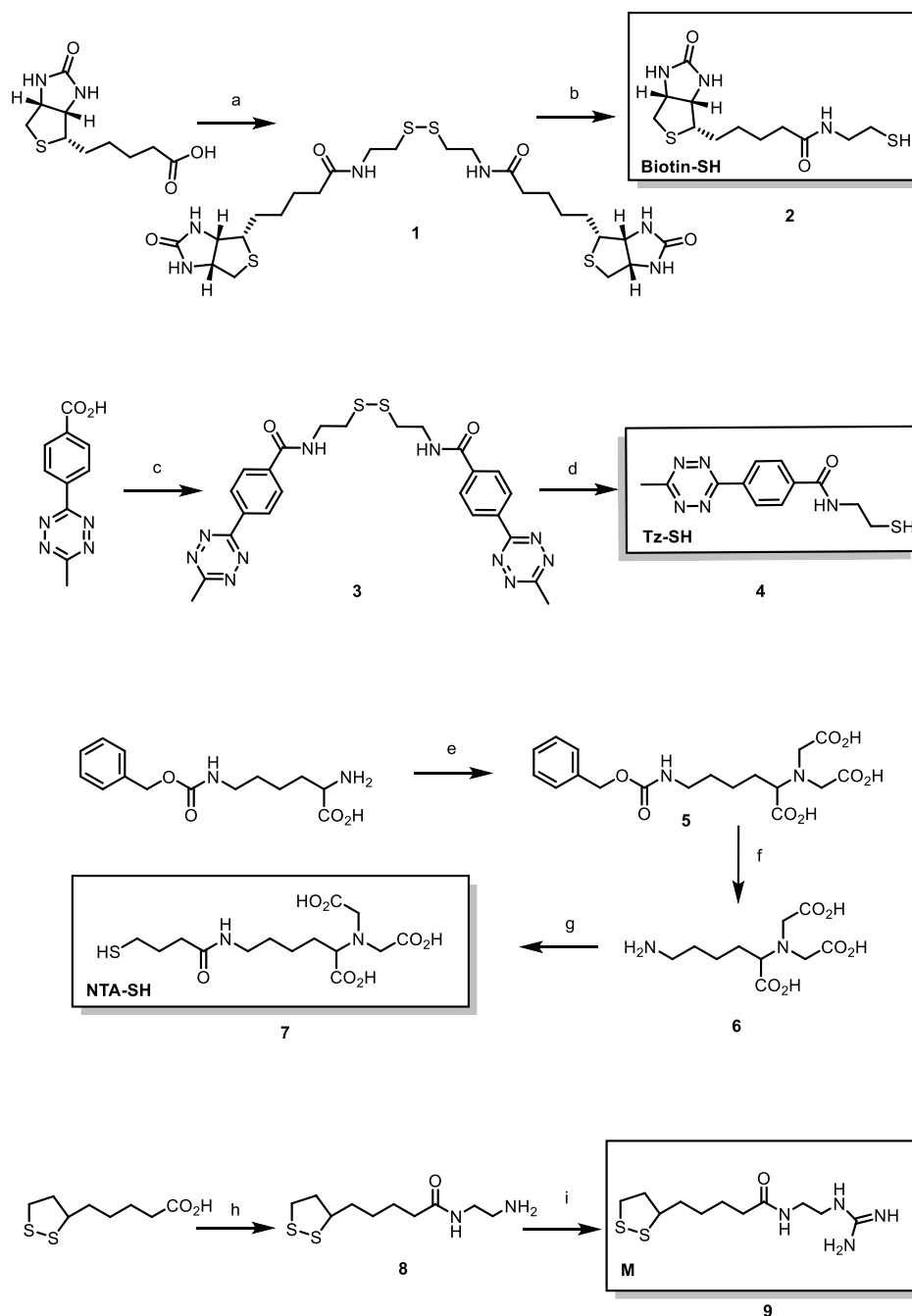
Scheme 4.1. Initiators synthesis.[157] (a) EDC·HCl, cystamine dihydrochloride, trimethylamine, DMF, 18 h, r.t., 91%; (b) TCEP, 50% DMF/H$_2$O, 30 min, r.t.; (c) EDC·HCl, cystamine dihydrochloride, trimethylamine, DMF, overnight, r.t., 91%; (d) TCEP, 50% DMF/H$_2$O, 30 min, r.t. ; (e) Bromoacetic acid, 2 M aqueous NaOH, 2 h, r.t., then 19 h, 50 ℃, 80%; (f) H$_2$, Pd/C, MeOH, 12 h, r.t., 75%; (g) 4-butyrothiolactone, NaHCO$_3$, 72 ℃, 15 h, 70%; (h) CDI, ethylene diamine, DCM, 2 h, r.t., 78%; (i) 1H-pyrazole-1-carboxamidine hydrochloride, DCM, 4 h, r.t., 56%. r.t. = room temperature

Upon purification, the resulting CPDs were further characterized by gel permeation chromatography (GPC), giving an average molecular weight (Mw) of 22-32 KDa with a polydispersity index (PDI) of 1.5-2.3 (Figure 4.1D & 4.2A). The concentration of **TzCPD** stock solution was determined by measurement with UV-Vis spectroscopy at the tetrazine absorbance ($\lambda_{520\,nm}$),[161] and for the other two CPDs, they were similarly estimated.
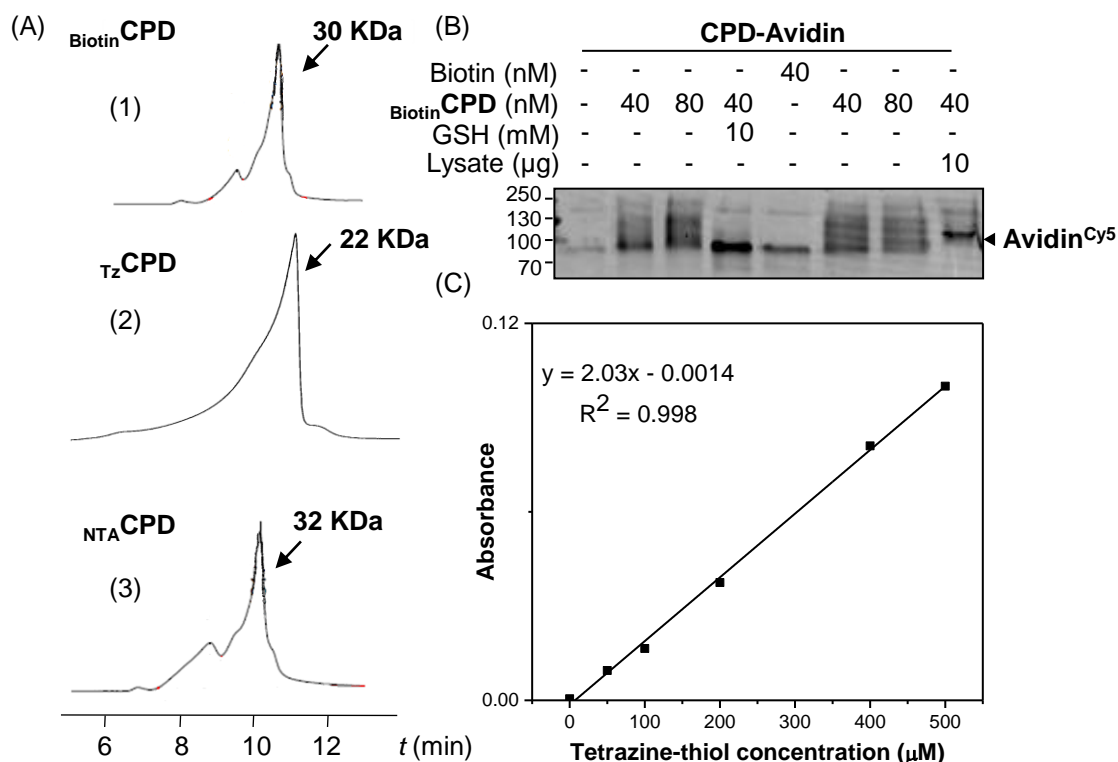

Figure 4.2. Polymer synthesis and characterization.[157] (A) Representative GPC profiling of polymers (1) **BiotinCPD**, (2) **TzCPD**, (3) **NTACPD**. (B) Polymer depolymerization tested by fluorescent in-gel scanning. 10 nM **CPD-Avidin** (treated without and with GSH (10 mM) or HeLa lysate (1

mg/ml) for 1 h at 37 °C before SDS-PAGE (DTT-free) and in-gel fluorescence scanning. **Avidin^{Cy5}** (CPD-free) was used as negative control. (C) UV-Vis measurement of $_{Tz}$**CPD** concentration using tetrazine disulfide (**3**) of different concentrations (50, 100, 200, 400, 500 μM) to generate standard curve (y = 2.03 x -0.0014, R^2 = 0.998). The $_{Tz}$**CPD** concentration of obtained stock was 200 μM.

4.3 Protein Attachment to CPDs.

We next chose three fluorescently labeled recombinant proteins having different molecular weights, **Avidin^{Cy5}** (~80 KDa in its tetrameric form), **TCO-BSA^{Cy5}** (~66 KDa) and **BRD-4^{Cy5}** (~15 KDa; an epigenetic reader protein[19]) as model cargos for delivery by $_{Biotin}$**CPD**, $_{Tz}$**CPD** and $_{Ni-NTA}$**CPD**, respectively. Fluorescent labeling of these proteins was done by standard protein conjugation chemistry with commercially available Cy5 dyes (Figure 4.3), thus allowing the entire cargo delivery process to be monitored by fluorescence microscopy and quantified by flow cytometry. Attachment of CPDs, either non-covalently or covalently, to the proteins was next done by simple "mix-and-go" protocols, giving **CPD-Avidin**, **CPD-BSA** and **CPD-BRD-4**, respectively (Figures 4.4). It should be highlighted that our choices of highly bioorthogonal "conjugation" chemistries between CPDs and proteins were critical, not only for convenient sake and generality, but more importantly to allow protein complexes be directly used in subsequent cell-based experiments. Furthermore, since most CPD/protein-complexing reactions couldn't be reliably monitored by SDS-PAGE without addition of DTT, which in turn caused CPD degradation (data not shown), we needed such conjugation chemistries to be free of failure! Nevertheless, we found that, for **Avidin^{Cy5}** and **CPD-Avidin**, they were well separated under modified DTT-free SDS-PAGE conditions (presumably due to extremely strong biotin/avidin interaction and avidin stability[162]), and were thus

chosen as the model system to monitor the processes of CPD/protein complex formation and GSH-assisted intracellular depolymerization/cargo release (Figure 4.2B); upon incubation with a freshly prepared HeLa lysate (1 mg/mL) for 1 h at 37 °C, the higher-order **CPD-Avidin** complex was found to have significantly depolymerized, indicating our CPD-loaded protein cargos would also be released readily in cytosolic environments.
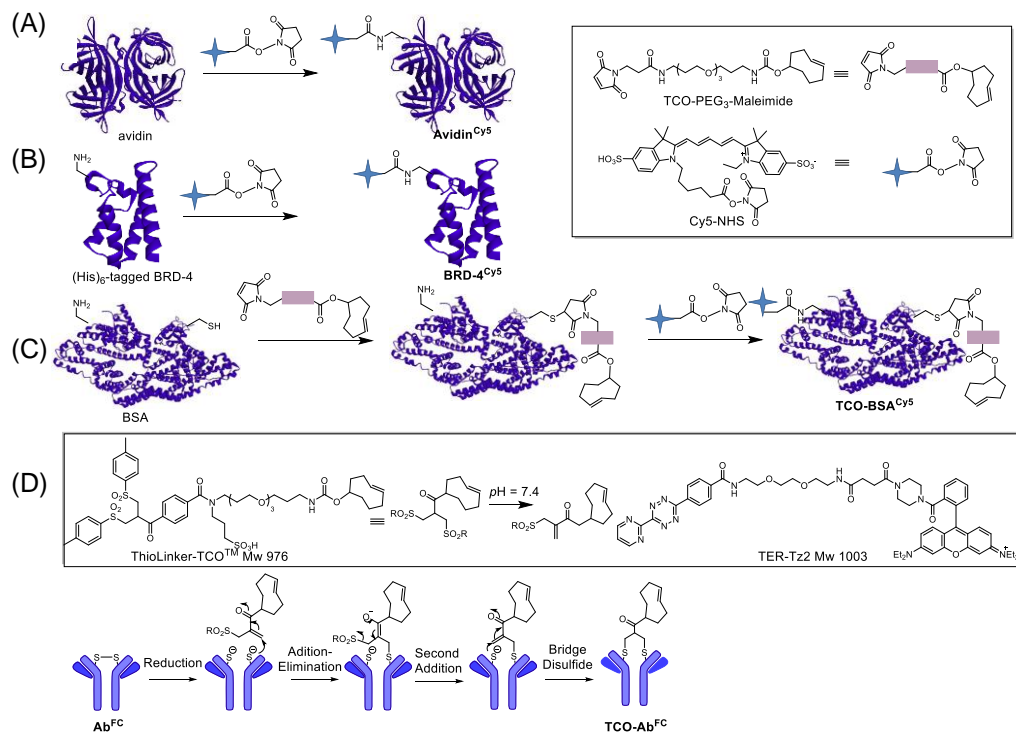
Figure 4.3. Labeling process of avidin, BRD-4, BSA and **Ab^FC**.[157] (A) Recombinant avidin was labeled with Cy5-NHS to generate **Avidin^Cy5**. (B) Recombinant (His)$_6$-tagged BRD-4 was labeled with Cy5-NHS to generate **BRD-4^Cy5**. (C) Recombinant BSA was first labeled with TCO via site-specific conjugation between TCO-PEG$_3$-Maleimide and the single free cysteine present in the protein. Next, TCO-BSA was fluorescently labeled with Cy5-NHS to obtain **TCO-BSA^Cy5**. (D) For antibody conjugation, a commercially available FC-labeled antibody (i.e. **Ab^FC**) was used, and the TCO moiety was site-specifically introduced to the antibody by first reducing the disulfide bond linkage in the antibody, followed by treatment with the ThioLinker-TCO^TM reagent to obtain **TCO-Ab^FC**.
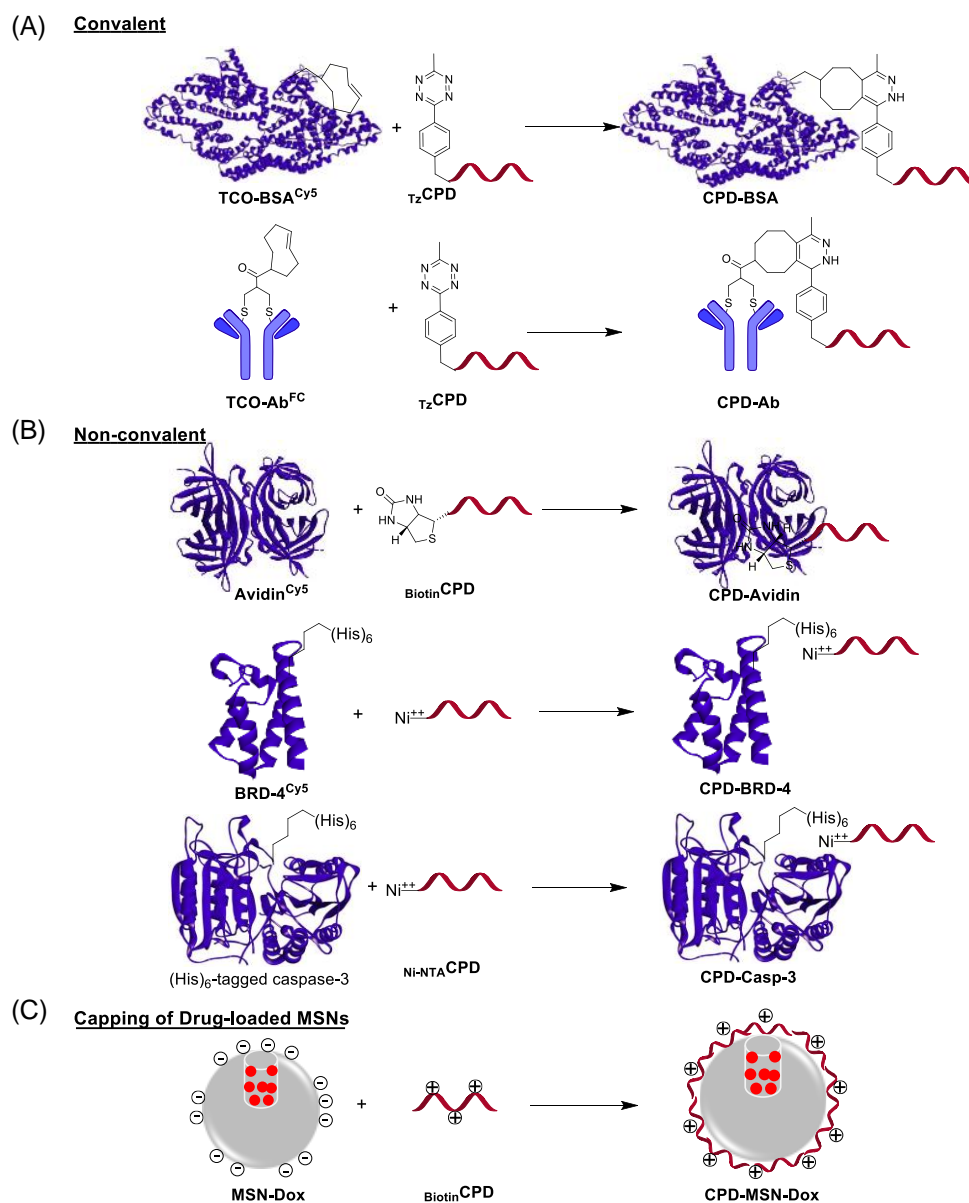
Figure 4.4. Formation of **CPD-Protein** and **CPD-MSN-Dox**.[157] (A) The formation of **CPD-BSA** between TzCPD and **TCO-BSA**Cy5, and **CPD-Ab** between TzCPD and **TCO-Ab**FC, via TCO-tetrazine bioorthogonal ligation. (B) Non-covalent formation of **CPD-Avidin** between BiotinCPD and **Avidin**Cy5, **CPD-BRD-4** between Ni-NTACPD and **BRD-4**Cy5, and **CPD-Casp-3** between Ni-NTACPD and (His)$_6$-tagged caspase-3, via affinity interaction. (C) "Capping" of small molecule drug-loaded **MSNs** with CPD via electrostatic charge-charge interaction between the negatively charged **MSN-Dox** and positively charged CPD (in this case, BiotinCPD was used).

4.4 Cellular Uptake.
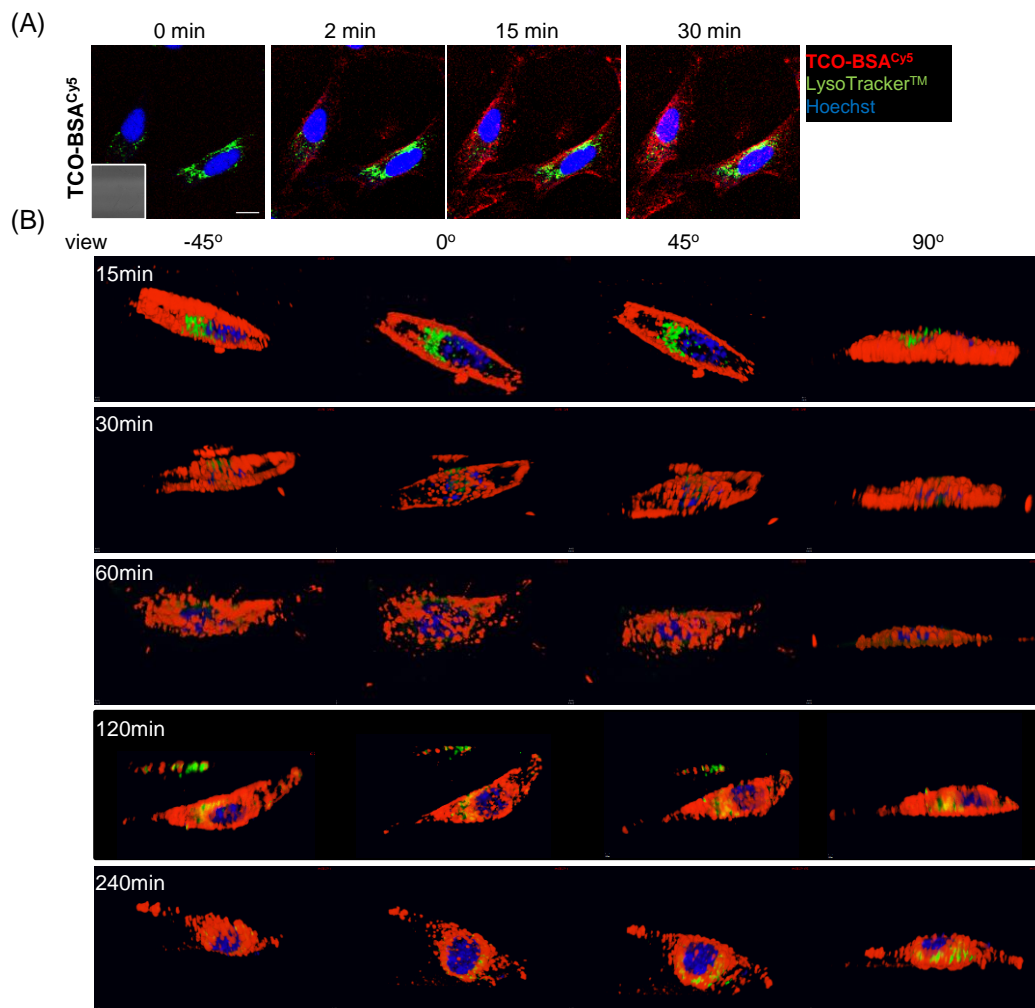
Figure 4.5. Cellular uptake of CPD-conjugated proteins.[157] (A) Real-time CLSM of live HeLa cells treated with 50 nM of **CPD-BSA** (red)**,** LysoTracker™ (green) and Hoechst (blue) at indicated time intervals. (Inset) bright-field images of the corresponding fluorescence channels. Scale bar = 20 μm. (B) 3D projections of *z*-stack images (left to right: different perspectives) of live HeLa cells treated with 50 nM **CPD-BSA** for 15, 30, 60, 120 and 240 min (step size, 0.186 μm). Cells were co-stained with Hoechst and LysoTracker™ for 15 min at 37 °C prior to the termination of protein delivery.

In order to unequivocally demonstrate the cellular uptake of the CPD-conjugated proteins, and their subcellular localization, we used confocal laser scanning microscopy (CLSM) with live HeLa cells. It was previously found that the use of fixed cells are not suitable for such studies;[142] upon fixation, cargos

delivered by CPPs and other means often artificially "escape" from endocytic vesicles, thus resulting in misleading conclusions.[130,163,164] Real-time imaging experiments together with different fluorescence organelle trackers (CellMask™ membrane tracker/LysoTracker™: pseudocolored in green; Hoechst nuclear stain: pseudocolored in blue) were carried out. As shown in Figure 4.5A, 2 min upon addition of 50 nM of **CPD-BSA** (pseudocolored in red) to the culture medium at 37 °C, green fluorescence started to accumulate around the cell membrane. After 15 min, a substantial amount of **CPD-BSA** was observed to have successfully been transduced and evenly distributed throughout the cytosolic space, with no evidence of endosome/lysosome trapping. This trend persisted for the next 2 h (Figures 4.5B). With prolonged incubation (> 4 h), however, we started to observe some merged green/red fluorescence signals, indicating some delivered protein had been destined for lysosomal degradation, presumably due to high protein concentration or unfolding.[130]

We next directly compared the cellular uptake efficiency of this **CPD**-assisted strategy with that of the Pro-Ject™ reagent, a commercially available liposome-based protein delivery system.[79] In addition to CLSM (Figure 4.6A), we analyzed the successfully delivered and depolymerized proteins by SDS-PAGE/in-gel fluorescence scanning of lysates from treated cells (Figure 4.6C). We further quantified protein uptake by flow cytometry analysis (Figure 4.6C). In order to minimize false readings of cells derived from membrane-bound, but not internalized, fluorescent proteins, cells were washed with heparin-containing PBS before analysis.[152] In all cases, with all three protein cargos (**TCO-BSA**[Cy5], **Avidin**[Cy5] & **BRD-4**[Cy5]) having

different types of "conjugation" chemistries (covalent and non-covalent), their resulting CPD-complexes (**CPD-BSA**, **CPD-Avidin** & **CPD-BRD-4**) were delivered into HeLa cells more efficiently than the Pro-Ject™ delivery method. Successful intracellular uptake of a cargo was found to be completely dependent on its complex formation with the corresponding CPD, as none of the cargos alone (Figure 4.6A-B) could enter cells. We found that, unlike the liposome-based Pro-Ject™ method which makes use
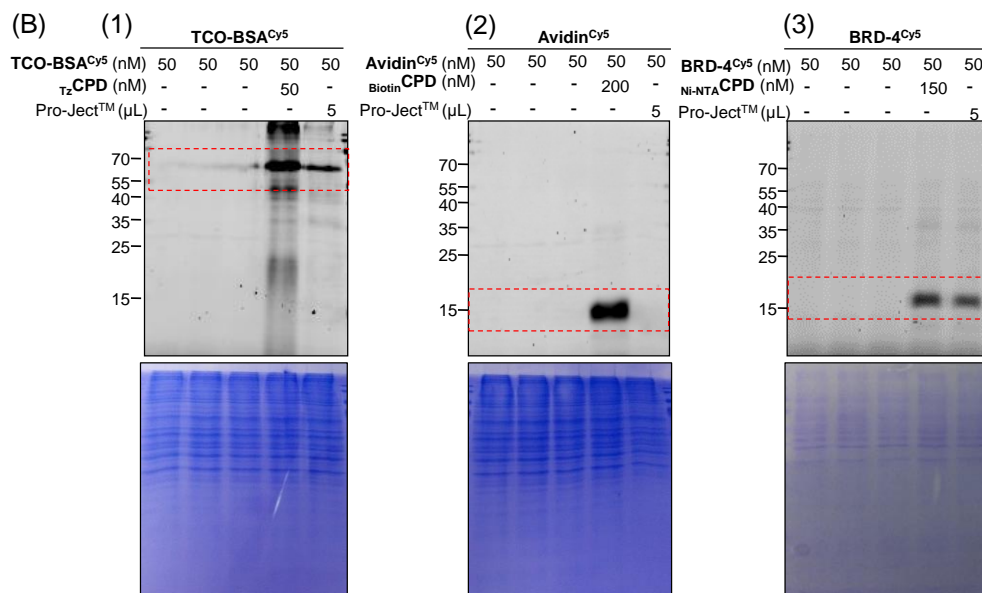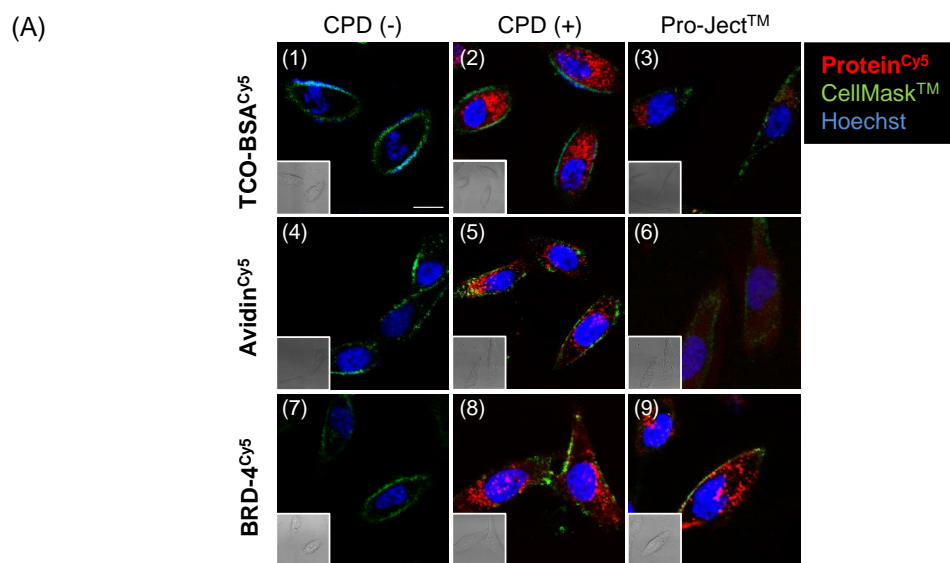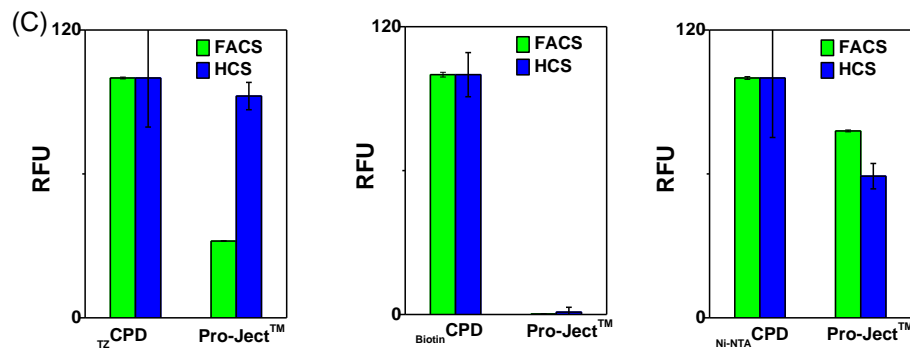
Figure 4.6. Comparison of the delivery efficiency of CPDs versus Pro-Ject[TM].[157] (A) Confocal image of HeLa cells treated with 50 nM **CPD-Protein** or 50 nM **protein**[Cy5]/Pro-Ject[TM] complex for 1 h. Cells were co-stained with Hoechst and CellMask[TM] after delivery (red, **CPD-Protein**, λ_{ex} = 643 nm, λ_{em} = 665-740 nm; green, CellMask[TM], λ_{ex} = 543 nm, λ_{em} = 555-650 nm; blue, Hoechst, λ_{ex} = 405 nm, λ_{em} = 440-470 nm). (1) 50 nM **TCO-BSA**[Cy5], (2) 50 nM **CPD-BSA**, (3) 50 nM **TCO-BSA**[Cy5] + Pro-Ject[TM], (4) 50 nM **Avidin**[Cy5], (5) 50 nM **CPD-BSA**, (6) 50 nM **Avidin**[Cy5] + Pro-Ject[TM], (7) 50 nM **BRD-4**[Cy5], (8) 50 nM **CPD-BRD-4**, (9) 50 nM **BRD-4**[Cy5] + Pro-Ject[TM]. (Insert) Bright-field images of the corresponding fluorescence channels. Scale bar = 20 μm. (B) Comparison of the delivery efficiency of CPDs versus Pro-Ject[TM] measured with FACS (green) and HCS (blue). HeLa cells were incubated with 50 nM **CPD-Protein** complex or 50 nM **protein**[Cy5]/Pro-Ject[TM] for 1 h before Hoechst stained, imaged and sorted (**CPD-Protein**, λ_{ex} = 632 ± 22 nm, λ_{em} = 684 ± 25 nm; Hoechst, λ_{ex} = 390 ± 18 nm, λ_{em} = 432.5 ± 48 nm). (1) **CPD-BSA**, (2) **CPD-Avidin**, (3) **CPD-BRD-4**. (C) In-gel fluorescence scanning and coomassie gel of corresponding HeLa lysates. (1) **CPD-BSA**, (2) **CPD-Avidin**, (3) **CPD-BRD-4**. (Lanes): 1. Blank; 2. **protein**[Cy5] in DMEM; 3. **protein**[Cy5] in serum-free DMEM; 4. **CPD-Protein**; 5. **protein**[Cy5] + Pro-Ject[TM]. Each set of data was independently normalized to HeLa treated with **CPD-Protein** at 37 °C (set to 100) and HeLa treated with **protein**[Cy5] at 37 °C (set to 0).

of electrostatic/hydrophobic interaction for complex formation with a target protein (thus delivery efficiency varies significantly with protein size/charge; see Figures 4.6B-C),[79] proteins of different sizes and charges were efficiently delivered by the CPD-assisted method. Gratifyingly, even (His)$_6$-tagged proteins formed by comparatively moderate non-covalent interaction with $_{Ni-NTA}$**CPD** (e.g. **CPD-BRD-4**) could be delivered, thus setting the stage for more widespread applications of this new protein transduction technology in future.
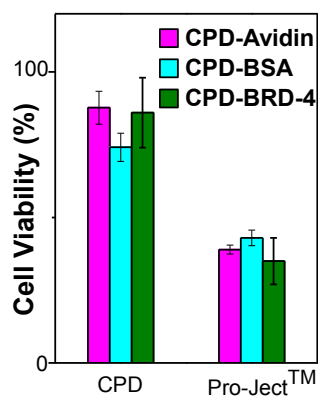
Figure 4.7. Comparison of the cytotoxicity of CPDs versus Pro-Ject[TM].[157] Cell viability measured with XTT assay for HeLa cells treated with a protein (50 nM; 1 h incubation) delivered by either CPD or Pro-Ject[TM] method. buffer-treated cells were used for normalization (as 100% viability).

In addition, we found the CPD method was significantly less cytotoxic than the Pro-Ject™ approach for protein delivery (Figure 4.7). This is in good agreement with previous cell-based experiments with fluorophore-loaded siCPDs.[152] In fact, the 10-20% cell death observed in our experiments was likely caused by trace amount of residual iodoacetamide from the polymerization reaction, which could not be completely removed with current purification method.
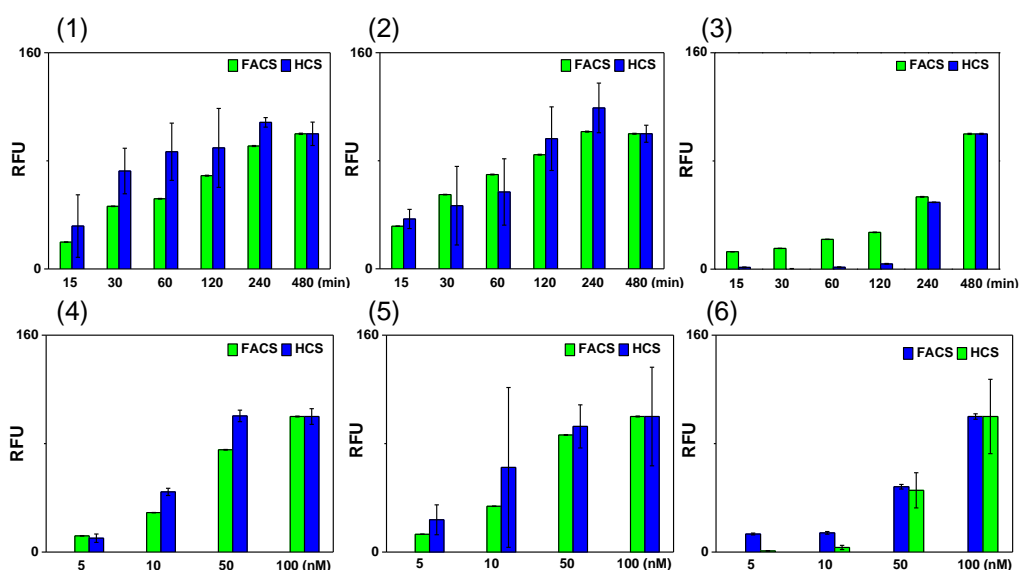
Figure 4.8. Time- and concentration-dependent delivery of **CPD-protein**.[157] Time-dependent delivery of (1) **CPD-BSA**, (2) **CPD-Avidin** and (3) **CPD-BRD-4** with HeLa Cells as measured by FACS (green) and HCS (blue). Concentration-dependent delivery of (4) **CPD-BSA**, (5) **CPD-Avidin** and (6) **CPD-BRD-4** with HeLa cells as measured by FACS (green), HCS (blue). For concentration-dependent experiment, HeLa cells were incubated with **CPD-Protein** (5, 10, 50 and 100 nM) for 8 h, followed by Hoechst stained, imaged and sorted. For time-dependent experiment, HeLa cells were incubated with 50 nM **CPD-Protein** for 15, 30, 60, 120, 240, and 480 min before imaged and sorted. Co-staining was done with Hoechst 15 min prior to the termination of protein delivery. (**CPD**-**Protein**, $\lambda_{ex} = 632 \pm 22$ nm, $\lambda_{em} = 684 \pm 25$ nm; Hoechst, $\lambda_{ex} = 390 \pm 18$ nm, $\lambda_{em} = 432.5 \pm 48$ nm).

Further optimizations of **CPD-Protein** delivery were done by concentration- and time-dependent experiments (Figures 4.8). In addition to flow cytometry, we used imaging-based, high-content screening (HCS) for comparative studies. HCS could be used to simultaneously analyze many live cells in the same experiment, without cell detachment/fixation which might cause artifacts, thus was a quantitative complement to our CLSM results. As expected, both longer incubation time and higher cargo loading led to increases in the amount of delivered proteins. With the issues of potential lysosomal degradation and cytotoxicity in mind, we recommend the optimal conditions for this CPD delivery method to be 25-50 nM protein loading and 1-2 h incubation, which are still more efficient than existing protein delivery strategies.[134-141]
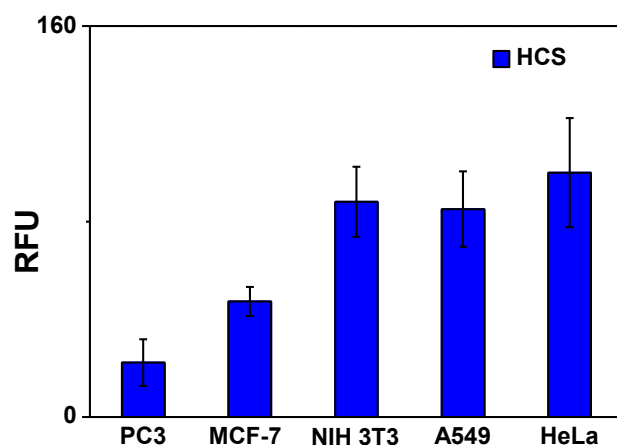
Figure 4.9. CPD delivery in different mammalian cells.[157] Quantification of cellular uptake efficiency of different cell lines (PC3, MCF-7, NIH 3T3, A549, HeLa) as calculated by HCS. Cells were incubated with 50 nM of **CPD-BSA** (red; **CPD-BSA**, λ_{ex} = 632 ± 22 nm, λ_{em} = 684 ± 25 nm) for 1 h before being imaged/quantified by In-Cell Analyzer.

The CPD method was further tested on other mammalian cell lines (NIH 3T3, MCF-7, A549 and PC3; Figure 4.9); in all cases, **CPD-BSA** was successfully taken up intracellularly, albeit at varying degrees of efficiency.
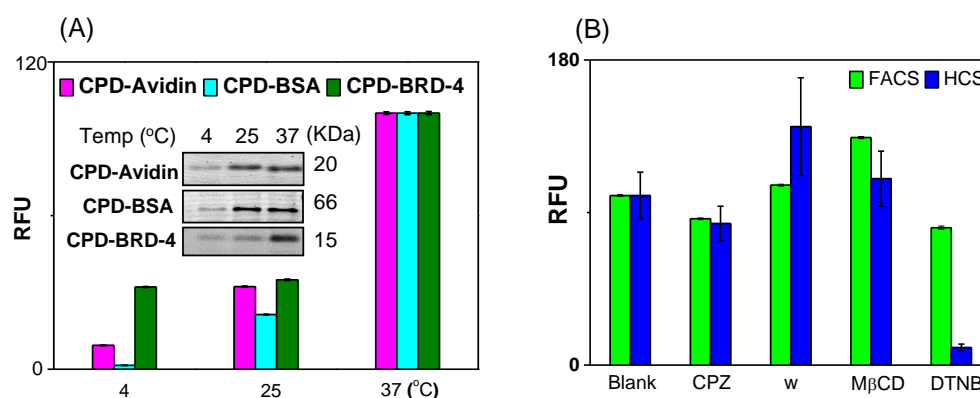
Figure 4.10. Temperature and endosome inhibition assay.[157] (A) Temperature-dependent protein uptake by HeLa cells (50 nM protein; 1 h treatment), as determined by flow cytometry. Data were normalized to those obtained at 37 °C. (Inset) SDS-PAGE/in-gel fluorescence scanning of lysates from treated cells. (B) Flow cytometry/HCS quantification of protein uptake (50 nM of **CPD-BSA**; 1 h incubation) by HeLa cells treated with different inhibitors, including chlorpromazine (CPZ), wortmannin (w), methyl-β-cyclodextrin (MβCD) and 5,5'-dithioobis-2-nitrobenzoic acid (DTNB). Data were normalized to those of HeLa cells treated with **CPD-BSA** only (Blank).

Our earlier CLSM results showed even cytosolic distribution of intracellular fluorescence upon protein delivery (Figure 4.5-4.6), indicative of endocytosis-independent pathways facilitated by these CPDs as previously proposed.[152] In order to further confirm this, we carried our detailed uptake studies of **CPD-Protein** by HeLa cells at different temperatures and in the presence of endocytosis inhibitors. In general, cell uptake profiles observed with these CPD-conjugated proteins were similar to what was previously reported

with small fluorophore-modified siCPDs.[152] Reduced temperature decreased protein delivery efficiency but did not block the process completely (Figure 4.10 A). Low temperature might decrease the fluidity of the bilayer membrane and thus affected the transduction efficiency. The insensitivity of protein delivery with endocytosis-related inhibitors used (chlorpromazine, wortmannin & methyl-β-cyclodextrin), ruled out the endocytosis pathway. On the contrary, blocking exofacial thiols on the cell surface with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB) significantly suppressed protein uptake, further supporting thiol-mediated cargo delivery mechanisms.[152] During this process, CPDs would first accumulated on membranes via counter-ion interaction and then covalently bond to exo-facial thiol facilitating the formation of micellar pores. Finally, proteins crossed the membrane through micellar pores. In our hands, flow cytometry gave less consistent results, presumably due to non-specific, surface-bound **CPD-Protein** and the need of cell detachment (Figure 4.10B). HCS on the other hand, was more reliable, enabling direct quantification of live cells having only intracellular fluorescence signals.

4.5 CPD-Assisted Transduction of Functionally Active Caspase-3

Having confirmed the effect of this newly developed CPD method for intracellular protein delivery with minimal endosome trapping, we next investigated whether it could deliver functional, therapeutic proteins. Caspase-3 (a cysteine protease) is a promising therapeutic protein owing to its key role in cell apoptosis.[165] Intracellular delivery of active caspase-3 to tumor cells

render them hypersensitive toward treatment by anticancer drugs such as Dox.[166] Functionally active, recombinant (His)$_6$-tagged caspase-3 was prepared as previously described.[167] Upon mixing with $_{Ni-NTA}$CPD, the resulting **CPD-Casp-3** was formed. Mindful of the trace amount of iodoacetamide from CPD preparation, and that absence of DTT might further reduce the enzymatic activity of caspase-3,[167] we carried out normalization of **CPD-Casp-3** (Figure 4.11); results indicate that even without DTT, **CPD-Casp-3** was able to retain > 20% of the original caspase-3 activity. The activity of the complex was partially restored upon DTT treatment, presumably due to either CPD depolymerization or/and DTT reduction of active-site cysteine in the enzyme. Therefore for **CPD-Casp-3**, upon cell entry, its enzymatic activity would also be restored under the highly reduced cytosolic environment.
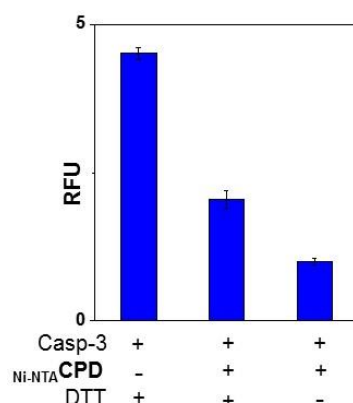
Figure 4.11. Enzymatic assay of **CPD-Casp-3** in HEPES buffer or DTT-free HEPES buffer.[157] 5 μM active caspase-3 was incubated with 15 μM $_{Ni-NTA}$CPD in HEPES buffer (10 mM HEPES, pH = 7.9) for 15 min on ice to form the **CPD-Casp-3** complex. Subsequently, 50 pM of **CPD-Casp-3** was incubated with 1 μM Ac-DEVD-AMC in HEPES buffer at 25 °C. Activity of 50 pM native (His)$_6$-tagged caspase-3 was measured concurrently as positive control. HEPES buffer with substrate only was used as negative control.

HeLa cells were incubated with 50 nM of **CPD-Casp-3**, and the resulting cells were imaged for intracellular caspase-3 activity upon treatment

with Ac-DEVD-AMC (Figure 4.12A) for 2 h;[168,169] significant fluorescence signals (from the liberated AMC dye) throughout the cytosol of **CPD-Casp-3**-treated cells, but not in control cells, were detected, indicating successful cytosolic delivery of the functionally active protein. These results were further confirmed by *in vitro* enzymatic determination of caspase-3 activity, as well as Western blotting (WB) analysis, of the corresponding cell lysates (Figure 4.12B);[150,170] in all cases, the presence of intracellular caspase-3 and its activity were unequivocally established. Intracellular activation of caspase-3 is known to promote subsequent translocation of the active enzyme into nucleus, cleave PARP1, and finally cause cell death by apoptosis.[171] We observed similar phenomena in our "artificially" induced apoptotic cells as well (Figure 4.12B inset & Figure 4.12C). These results further suggest that cytosolic proteins delivered by this CPD system are not necessarily confined to their original destination upon cell entry. Instead, they may be further sorted/translocated in manners similar to endogenous proteins.
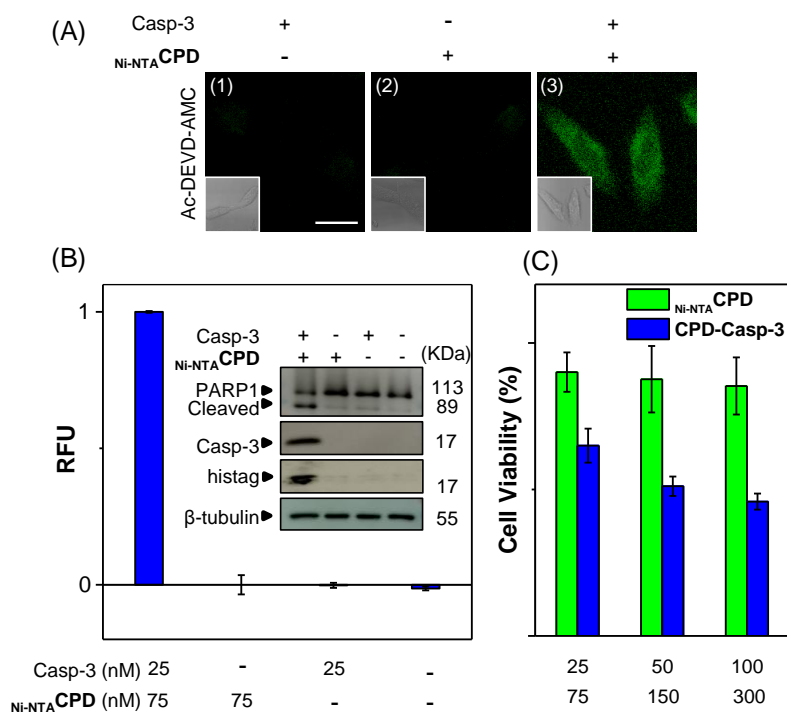
Figure 4.12. CPD-assisted delivery of functionally active caspase-3 to HeLa cells.[157] (A) CLSM images of HeLa cells treated with **CPD-Casp-3** (50 nM) for 2 h, followed by incubation with Ac-DEVD-AMC (40 μM; 2 h). Scale bar = 20 μm. (Inset) bright-field images. (B) RFU of *in vitro* enzymatic caspase-3 assay of treated HeLa lysates.[150,170] (Inset) WB results of delivered active caspase-3 and cleaved endogenous PARP1 in HeLa lysates from the same cells. (C) Cell viability/apoptosis caused by delivered active caspase-3 as measured by XTT assay. Since it took some time for intracellular caspase-3 to be translocated into the nucleus and cleave PARP1,[79] for (B & C), HeLa cells were treated with **CPD-Casp-3** for 8 h prior to analysis.

4.6 CPD-Assisted Antibody Delivery.

Monoclonal antibodies (mAbs) constitute one of the largest classes of therapeutic proteins.[172] There are currently more than 30 antibody-based, FDA-approved drugs, most of which target cancer. The number could have been much higher, had an effective means for intracellular delivery of antibodies been available.[158] Common protein delivery methods are even more problematic for intracellular delivery of antibodies, as most of them are large (> 150 KDa).

Realizing the highly efficient, endosome-independent features endowed by our newly developed CPD protein delivery method, we anticipated it might be ideally suited for delivery of therapeutic antibodies, for which many robust conjugation chemistries are already available in the forms of antibody-drug conjugates (ADCs) and PEGylation without compromising antibodies' activity and stability.[158,159,173,174] We used Alexa Fluor® 488 goat anti-rabbit IgG (**Ab^FC**) as a model antibody and modified it with a commercially available bis-sulfone reagent, ThioLinker-TCO™ (Figure 4.13A); upon TCEP reduction of inter-chain disulfides in the antibody, the TCO reagent underwent double elimination-addition reactions with two cysteines from the same reduced disulfide to form a three-carbon bridge, thus successfully introducing TCO into the antibody while leaving it structurally intake. The same approach has been used for site-specific PEGylation and ADCs in various therapeutic proteins/antibodies.[173,175] As shown in Figure 4.13B, while **Ab^FC** showed up on a DTT-free SDS-PAGE gel as a 250-KDa fluorescent band under the FITC channel (Lanes 4), successful introduction of the TCO moiety (giving the resulting **TCO-Ab^FC**) followed by ligation with a tetrazine-containing tetraethylrhodamine reporter (TER-Tz2[161]) shifted the band to higher molecular-weight regions where they became detectable under both FITC and TER channels (Lanes 2).

With **TCO-Ab^FC** being successfully prepared, we next conjugated it to **Tz CPD**. as earlier described, to generate **CPD-Ab** which was subsequently used for confirmation of successful intracellular antibody delivery by CLSM. As shown in Figure 4.13C, live HeLa cells treated with 50 nM of **CPD-Ab** for 1 h

were shown to have taken up the fluorescently labeled antibody and distributed it throughout the cytosol (right panel), whereas no fluorescence was detected in cells treated with the antibody alone (without conjugation to **Tz**CPD; left panel). We thus conclude that this CPD-assisted method could be used for intracellular delivery of antibodies as well.
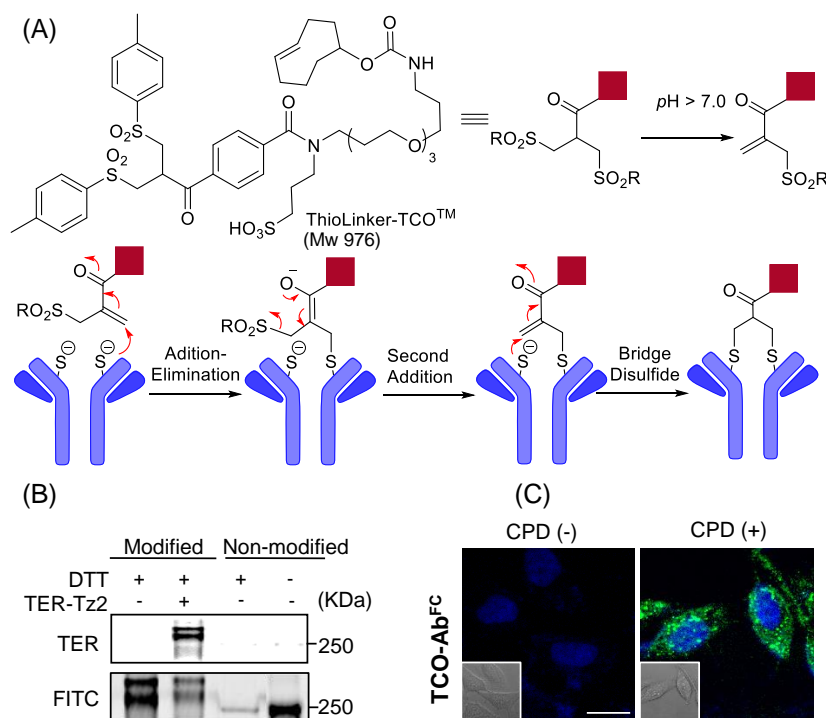
Figure 4.13. CPD-assisted antibody delivery.[157] (A) Labeling mechanism of antibodies by ThioLinker-TCO™. (B) **TCO-Ab^{FC}** reacting with or without TER-Tz2, before being analyzed by SDS-PAGE (with or without 10 mM DTT in the loading dye). The gel was visualized under both FITC and TER channels. (C) HeLa cells treated with 50 nM of **CPD-Ab** (1 h at 37 °C) before being imaged. **CPD-Ab** (green); Hoechst (blue). Scale bar = 20 μm.

4.7 Small Molecule Drug Delivery by **CPD-MSNs**

Delivery of small molecule drugs to tumor sites often suffers from low efficiency due to hydrophobicity and poor water solubility. Approximately 40% of all current commercial drugs and up to 75% of drug candidates are classified

as being poorly water-soluble.[132] We reasoned the siCPD method previously proposed by Matile *et al* would have been impractical, as it would require chemical modification of small molecules with a thiol handle which is not available in most native drugs.[152] We therefore turned to mesoporous silica nanoparticles (MSNs), which are widely used for intracellular delivery of native small molecule drugs. Due to the large size of MSNs (> 100 nm in diameter), they are often not efficiently taken up by cells unless surface modification with CPPs or other chemicals is introduced.[145-150] In most cases, however, this leads to endocytosis and poor cytosolic release of MSN-encapsulated drugs. We wondered whether the unique endocytosis-independent mechanisms of the CPD-assisted delivery could be successfully emulated in drug-loaded MSNs, e.g. **CPD-MSN-Dox**.

To make **CPD-MSN-Dox**, PO_4^--modified MSNs were first prepared according to published protocols.[176] Such MSNs, due to their negatively charged surface, were known to be minimally taken up by mammalian cells. To follow the cellular uptake of MSNs, they were first doped with a small amount of fluorescein.[150] We next loaded Dox followed by capping the surface of the resulting drug-loaded MSNs with positively charged ᴮⁱᵒᵗⁱⁿ**CPD** via charge-charge interaction. The resulting MSNs were shown to be highly monodisperse (mean diameter: ca. 155 nm) and possess well-defined pore sizes with high specific areas (Figure 4.14-4.17). Reversal of Zeta potentials from negatives in the PO_4^--modified MSNs to positives in the CPD-capped MSNs was evident (Figure 4.18). *In vitro* GSH-induced depolymerization/uncapping followed by release of Dox was successfully observed (Figure 4.19).
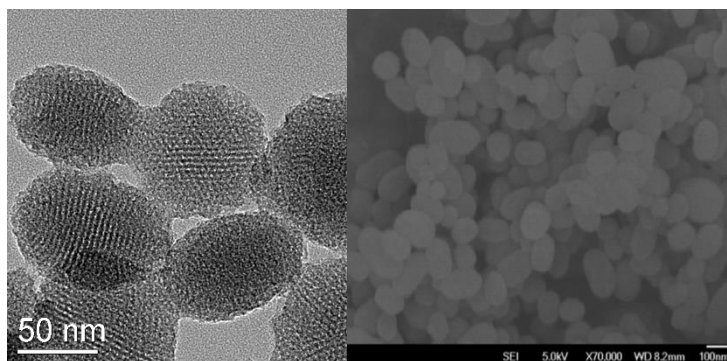
Figure 4.14. TEM (left) and SEM (right) images of **CPD-MSN**.[157] The resulting porous silica nanoparticle is about 120 nm in diameter and arranged pore structure can be observed.
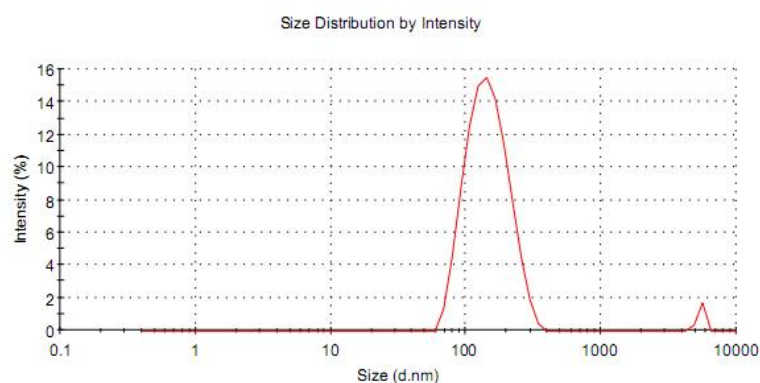
Figure 4.15. Hydrodynamic size distributions of **CPD-MSN** in PBS buffer (pH = 7.4).[157] The diameter of the **CPD-MSN** in the buffer is about 155 nm.
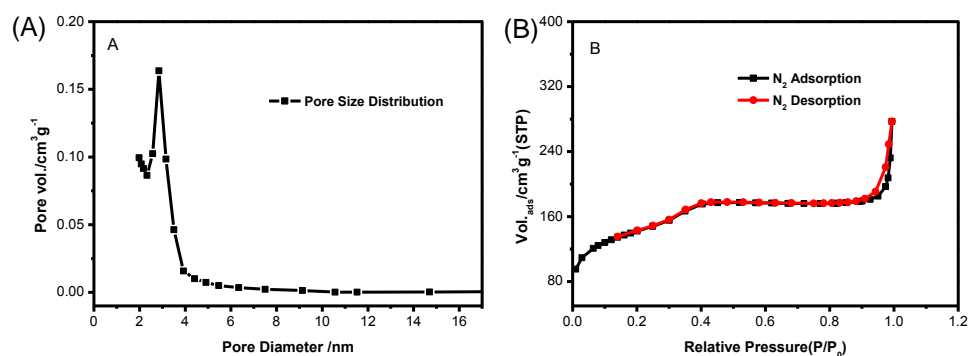
Figure 4.16. BET nitrogen adsorption/desorption isotherms (A) and BJH pore size distribution (B) of the nanoparticles. BET specific surface values, pore volumes, and pore sizes calculated from the nitrogen adsorption-desorption isotherms were 674.8 m^2/g, 0.320 cm^3/g, 2.45 nm for MSN-PO$_4^-$ [157]
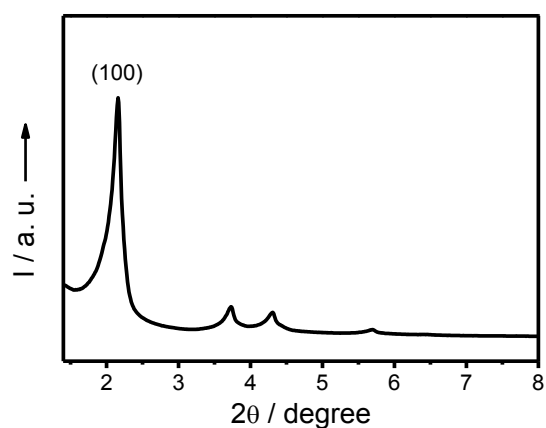
Figure 4.17. Powder X-ray diffraction pattern of the particles.[157]
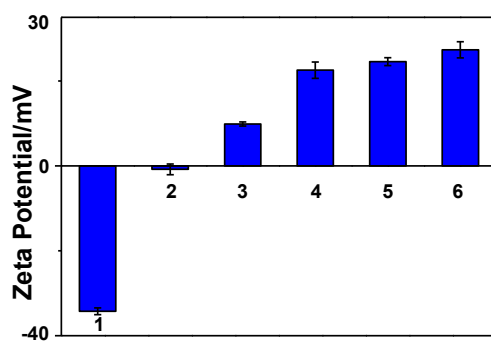
Figure 4.18. Zeta potentials of MSN-PO$_4^-$ capped with $_{Biotin}$**CPD** of different concentrations (1 to 6 represented 0, 10, 20, 50, 100 and 150 μM $_{Biotin}$**CPD** respectively).[157]
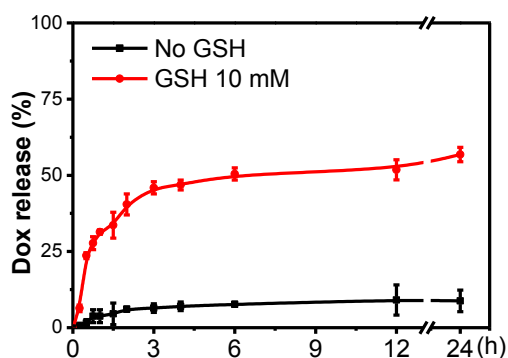
Figure 4.19. The percentage of Dox released from **CPD-MSN-Dox** (0.1 mg/mL) versus the incubation time in PBS with or without 10 mM GSH.[157]

To follow the entire process, from cellular uptake of **CPD-MSN-Dox** in live HeLa cells, the subsequent uncapping/depolymerization of CPD to

cytosolic release of Dox, we directly added to cell medium 20 μg/mL of the drug-loaded MSNs and the resulting cells were imaged over 24 h by CLSM, followed by WB and apoptosis analysis (Figure 4.20). **MSN-Dox**, the drug-loaded nanoparticles without capping with **BiotinCPD**, was used as a negative control. Since Dox is an intrinsically fluorescent compound, its intracellular distribution could be conveniently monitored by CLSM. It is also one of the most effective anti-cancer drugs. By intercalating to nuclear DNA of target tumors, it is known to cause caspase-3 activation, PARP1 cleavage and cell apoptosis.[177] As shown in Figure 5A, we observed accumulation of most **CPD-MSN-Dox** inside the cytosol of treated cells after only 3 h of incubation, at which point substantial release of the MSN-encapsulated Dox (in red) was also observed (panels 1-3). The slower cellular uptake of **CPD-MSN-Dox**, when compared to that of **CPD-Protein**, was likely due to the much larger size of MSNs, but was nevertheless still faster than that of CPP-capped MSNs.[149,150] Over the course of the next 21 h, more Dox was released from **CPD-MSN-Dox** and entered cell nucleus (panels 6, 10 & 14 in Figure 4.20A), resulting in apoptosis in > 70% of cells (Figure 4.20B-C; 24-h treatment). Successful activation of endogenous caspase-3 activity and cleavage of PARP1 were observed as well. For cells treated with **MSN-Dox** (i.e no **BiotinCPD** capping; 12 and 24 h incubation), intracellular fluorescence signals were detected in the red but not green channel, indicating cellular uptake of **MSN-Dox** was unsuccessful due to a lack of the surface-bound **BiotinCPD** (Figure 4.21), but leaked Dox from the nanoparticles was able to subsequently enter cells freely and cause cell death. Again, a small percentage of cell death detected in cells treated with

CPD-MSN (i.e. no Dox; Figure 4.20C) was attributed to the trace amount of

iodoacetamide present in BiotinCPD.
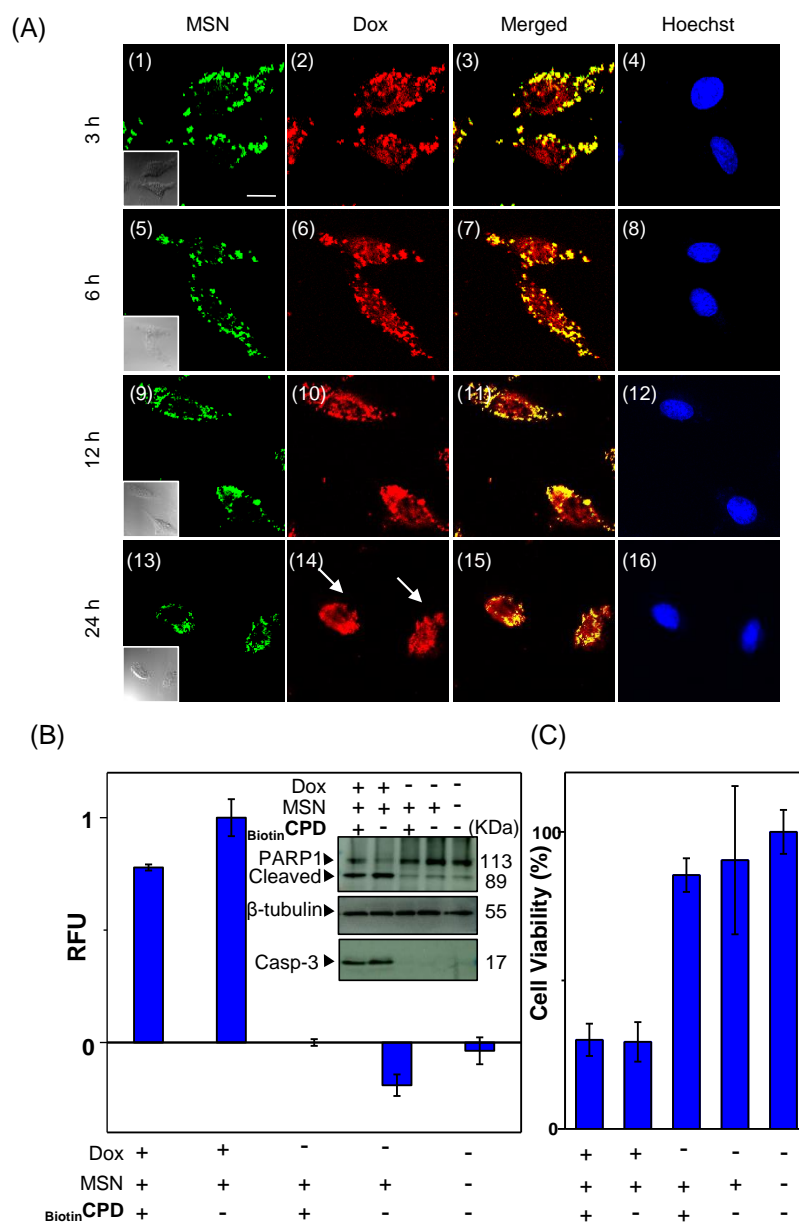
Figure 4.20. CPD-assisted MSN delivery.[157] (A) Confocal images of HeLa cells treated with **CPD-MSN-Dox** (20 μg/mL) over 24 h at 37 °C. Dox (red), MSN (green) and Hoechst (blue) were colored accordingly. Arrowed: apoptotic cells. Scale bar = 20 μm. (B) *In vitro* enzymatic assay of the cell lysates (after 24-h cell treatment). RFU was recorded after 5 h incubation with Ac-DEVD-AMC (1 μM) with HeLa lysate (25 μg/reaction). (Inset) WB results showing the successful activation of endogenous caspase-3 and cleavage of PARP1 of treated cells. (C) Cell viability/apoptosis caused by release of intracellular Dox in (A & B) as measured by XTT assay. Control

experiments were concurrently done with cells treated with MSNs alone, **CPD-MSN** or **MSN-Dox**.
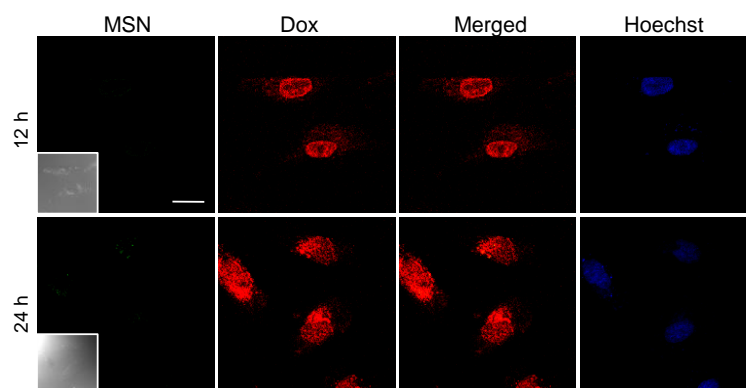
Figure 4.21. Confocal images of HeLa cells treated with **MSN-Dox** (20 μg/mL) over 24 h at 37°C.[157] (red, Dox, λ_{ex} = 488 nm, λ_{em} = 580-650 nm; green, MSN, λ_{ex} = 488 nm, λ_{em} = 500-560 nm; blue, Hoechst, λ_{ex} = 405 nm, λ_{em} = 440-470 nm). (Inset) DIC images. Scale bar = 20 μm.

4.8 Conclusion

In this study, we have successfully designed and synthesized several novel cell-penetrating poly(disulfide)s. These CPDs (BiotinCPD, Ni-NTACPD & TzCPD), upon highly efficient bioorthogonal "conjugation", either non-covalently or covalently, to readily available cargos including recombinant proteins and suitably modified antibodies, were able to rapidly and efficiently deliver these cargos into different mammalian cells via endocytosis-independent pathways. Rapid intracellular CPD depolymerization of the delivered cargos under highly reduced cytosolic environments subsequently released the proteins in their functionally active form, which may be further translocated to their intended subcellular organelles for additional biological processes. The successful intracellular delivery of antibodies by TzCPD indicates this method may be more broadly applicable in future for effective cellular delivery of many other therapeutic antibodies, which at present could

not be adequately achieved with other protein transduction methods.[158] Unlike the siCPDs approach recently developed by Matile *et al*,[152] who suggested thiol-containing small molecule drugs and probes may be directly "grown" onto CPDs during polymerization, we have successfully developed CPD-capped MSNs for encapsulation of native small molecule drugs without the need of introducing thiol handles. With doxorubicin as an example, we found **CPD-MSN-Dox** entered mammalian cells rapidly and was able to subsequently release free Dox into the cytosol. While more studies are needed to investigate the utilities of CPDs as novel "capping" agents for MSNs and other types of nanoparticles, our preliminary finding herein indicates that they may be more widely used in future for intracellular delivery of otherwise difficult-to-deliver drugs in a highly controllable manner.[131-133,145-147]

One of the key features of our two-step, CPD-assisted approaches is their versatility and flexibility, enabling immediate delivery of a variety of cargos (recombinant proteins, antibodies and native small molecule drugs) with minimum chemical and genetic manipulations. The other feature is the rapid and "bioorthogonal" cargo-loading process – with different types of pre-synthesized CPDs in hand, the resulting CPD-cargo conjugates could be prepared in a matter of minutes under aqueous conditions, and used immediately for subsequent cell delivery studies. The minimal cell cytotoxicity of these new CPDs and their cargo-loaded conjugates further highlights the unique advantage of this new cell-transduction method over other existing strategies, and ensures our entire delivery protocol is compatible with live cell experiments. Future work will focus on the expansion of the types of CPDs by using other conjugation chemistries, development of better CPD purification protocols, and

application of these CPDs for cell type-specific delivery of other therapeutically important drugs (including proteins, antibodies and small molecules).

Chapter 5.

Experiment Section

5.1 Profiling human Src homology 2 (SH2) domain proteins and ligand discovery using a peptide-hybrid small molecule microarray

5.1.1 General Information

All chemicals were purchased from company vendors and used without further purification. Fmoc-acid, HBTU, HOBt, TIS, PyBrOP and Rink Amide resin were purchased from GL Biochem (China). PL-FMP 4-Formyl-3-methoxyphenoxy Resin (PL-FMP Resin, 0.9 mmol/g, 75~150 μm, Part no:1465-799) was purchased from Polymer Laboratories (USA). HPLC grade solvents were used for peptide synthesis. The reaction was carried out at room temperature and monitored by ninhydrin or chloranil test. HPLC and Mass spectra profiling were recorded on Shimadzu LC-ITTOP system using reverse-phase Phenomenex Luna 2.6 μm C18 100 Å 50 × 3.0 mm columns. Preparative HPLC was carried out on Gilson preparative HPLC system using Trilution software and reverse-phase Phenomenex Luna 5 μm C18(2) 100 Å 50 × 30.00 mm column, to purify selected hits. 0.1% TFA/H2O and 0.1% TFA/CH3CN were used as eluents for both systems. The flow rates were 0.6 ml/min for HPLC and 8ml/min for preparative HPLC. SH2 domain were expressed, purified and labelled on beads. Gels were scanned on a Typhoon fluorescence gel scanner (GE Healthcare, USA). Plain glass slides were purchased from Sigma Aldrich

(USA) and the generation of avidin slides were performed as described previously.1 Microarray slides were scanned using a Tecan Launch LS Reloaded Microarray Scanner (TecanTrading AG, Switzerland) installed with suitable lasers: Cy3: λEx/Em= 532/575 nm; Cy5: λEx/Em= 633/692 nm.

5.1.2 Expression and Labelling of Protein

The mammalian SH2 domain collection was cloned into modified pET28 bacterial expression vectors with His tagged was obtained from Open Biosystems. All expression constructs were transformed into *E.coli* BL21 (DE3) competent cells. The transformed cells were then plated onto the LB-kanamycin agar plates and placed in the incubator at 37 ℃ for 16 hours. Overnight cultures were diluted 1:100 in LB media supplemented with 50μg/ml of kanamycin and grown at 37℃ with shaking. When OD600 was about 0.6-0.8, expression was induced with 0.1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and cultures were grown further at 18 ℃ for 18 h. After cell harvest (4 ℃, 10000 g, 30 min), lysis buffer (*p*H 7.4, 50 mM Tris-HCl, 300 mM NaCl) and 1mg/ml lysozyme were added. After incubation on ice for 20 min, lysis was performed by sonication (12×pulses of 10 s each at half maximal power, on ice), the cell debris was pelleted by centrifugation at 10000 g, 20 min, 4 ℃, which the supernatant was loaded into a column containing 50-80 μl of Ni-NTA resin preequilibrated with wash buffer (*p*H 7.4, 50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole). Following incubation for 2 h at 4 ℃, the resin was washed 3 times with the wash buffer and the protein was finally eluted with elution buffer (*p*H 7.4, 50mM Tris-HCl, 300 mM NaCl, 200 mM imidazole). Fractions containing the desired proteins were pooled and dialysed

onto Microcon® centrifugal filter device (3kDa cutoff) and stored at -20 °C in PBS containing 20% glycerol. Protein concentration was determined from a standard curve using the Bradford protein assay (Bio-Rad) with BSA protein. Protein purity was determined by separation on a 15% SDS-PAGE gel by Coomassie® brilliant blue staining.

5.1.2.1. Solution-phase Protein Labelling

Protein samples were labelled with Cy5 dye (Amersham, G.E. Healthcare, USA) for 1 hour on ice, following manufacturer's protocol.[107] A 50 µl protein sample at 1 mg/ml was quenched with 5 µl quench solution (3M hydroxylamine, 2.5M NaOH) for further 30min. The excess dye was removed with a Microcon® centrifugal filter device (3KDa cutoff). The labelled protein was determined by separation on a 15% SDS-PAGE gel by Coomassie® brilliant blue staining.

5.1.2.2. Protein Labelling while on Solid-support

Following washing the Ni-NTA resin with wash buffer, the resin was washed twice with PBS. To the resulting protein-bound resin, a 200 µl Cy5-NHS ester solution(Amersham, G.E. Healthcare, USA) was added to label the protein. Following incubation on ice for 1 hour, the unreacted dye was quenched with 20 µl quench solution for a further 30 min. 15ml PBS was used to wash the labelled-resin thrice to remove the excess dye. The protein was finally eluted with elution buffer (pH 7.4, 50 mM Tris-HCl, 300 mM NaCl, 200 mM imidazole). Fractions containing the desired protein were pooled and dialysed onto Microcon® centrifugal filter device (for 3 KDa) and stored at -20 °C with

PBS containing 20% glycerol. The labelled protein was determined by separation on a 15% SDS-PAGE gel by fluorescence scan on Typhoon.

5.1.3. Synthesis of Peptide-small molecule Hybrid Library
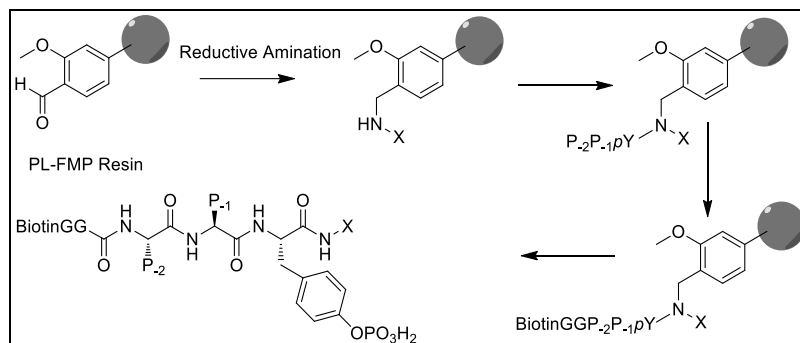
Scheme 5.1. Synthesis of the peptide-small molecule library[102]

The library was synthesized using the standard Fmoc strategy combined with IRORI^{TM} technology (Scheme 5.1). Briefly, the 396-member library was synthesized on PL-FMP resin 50 mg resin was used for each compound. The resin was first swelled in DCE for 2 h and then coupled the amine building blocks (X_{1-44}) onto the PL-FMP resin by reductive animation. Four molar equivalents of the amine building block was added in DCE and the incubated with shaking for 2 h. Next, HOAc(1%) and Na(OAc)$_3$ (6.0 eq.) in DCE was added and the reaction was further shaken for 16 h. Subsequently, the resin was washed thoroughly with THF, MeOH and DCM and dried in high vacuo. To confirm completion, the reaction was monitored using chlorinal test. Next, the resin was coupling with Fmoc-Tyr[PO(OBzl)OH]-OH(4.0 eq.), PyBrOP(4.0 eq.) and DIEA(8.0 eq.) in DMF. After reacting for overnight at room temperature, the result resin was extensive washed with DMF and DCM and dried. To confirm completion, the reaction was also monitored by chlorinal test. Subsequently, the Fmoc group was removed by 20% piperidine and ready for next step. Repeat the cycle using HOBt/HBTU/DIEA coupling method until the

Biotin-GG linker was coupled. Cleavage was performed with TFA/H$_2$O/TIS (95:2.5:2.5, total 2 ml) for each micro-reactor. After continuous shaking for 6 h at room temperature, the resin was separated and the filtrate was collected and concentrated under reduced pressure. Cold ether (chilled to -20 ℃, ~5 ml) was added to the liquid concentrates to precipitate the peptide-small molecule compounds. The mixture was stored at -20 ℃ overnight. Upon centrifugation, the ether layer was decanted off and washed with ether twice. The precipitates were then dried thoroughly *in vacuo*, dissolved in DMSO (0.5 ml) and stored at -20 ℃ for future use. The 14 individual hits were also re-synthesised using the procedure above. Each micro-reactor reactor contained 200 mg of PL-FMP resin. Unique RF tags were used to facilitate sorting. The peptide-small molecule compounds were further purified using preparative HPLC (Gilson) and lyophilised. LC-MS was performed to ensure the peptide-small molecules were of correct mass and pure for subsequent microarray experiments, the spectrums and MS results are shown in Figure 5.1

5.1.4. Pro-Q Staining and Detection

The spotted slide was washed with distilled water and stained with Pro-Q™ Diamond dye for 1 h at room temperature in a humidified chamber. The slide was then de-stained with a solution of 20% acetonitrile in sodium acetate (*p*H 4) for 25 min, rinsed with water, dried and scanned under the Cy3 channel ($\lambda_{ex/em}$: 532/575 nm).

5.1.5. Microarray Preparation

The peptide-small molecular hybrid library stock solutions were prepared to approximately 1.0 mM in 50% DMSO and 50% PBS and were

distributed in Genetix384 well plates. Stock spotting plates were not recycled more than 5 times, to minimize variability. Avidin coated glass slides (75x25 mm) were spotted on an OmniGrid® Accent (DigiLab, USA) microarray spotter with the print head installed with four Stealth SMP8B Micro-spotting pins (TeleChem USA). Spots generated were approximately 295 μm diameter and were printed with a spot-spot spacing of 450 μm. The pins were rinsed in between samples using two cycles of wash (for 5 s) and sonication (for 5 s) in reservoirs containing 70% ethanol followed by drying under reduced pressure (for 5 s). The slides were allowed to stand for overnight on the printer platform and stored at 4 °C until use (spotted slides were stable for 6 months under these storage conditions). Before incubation with the labelled protein, the slides were rinsed with PBS (pH 7.4) for 20 min and blocked with PBS containing 1% BSA for 1 h. For studies with the 396-member library, all were spotted on the slide in duplicate. For the K_D experiments, a new microarray was fabrication where up to 8 identical subarrays were created on the same slide using the same set of 14 selected hits. ProQ assays were performed to ensure batch-batch variability.

5.1.6 Data extraction and analysis

Microarray data was extracted using the Array-Pro® software. Values from duplicated points and was background subtracted and averaged (Duplicated spots with a standard deviation >0.8 were rejected) and were depicted in grey in heatmaps, to indicate data not obtained.

5.1.7. K_D Analysis of Selected High Binders

14 hits including the previous synthesized hits and the newly synthesized were spotted onto the same slide were up 8 identical subarray were

generated on the same platform, allowing consistent/uniform screening and binding of Lck SH2 domain. By using dose-dependent experiments, as previous described,[101] we extracted the binding data of Lck SH2 domain. The corresponding K_D was generated by fitting the data to the following equation, under the assumption that equilibrium was achieved during the incubation period:

$$\text{Observed fluorescence of x}=\frac{(Maxium\ Fluorescence, X) \times [Protein\ Concentration]}{K_D + [Protein\ Concentration]}$$

Saturation dynamics observed when plotting Observed Fluorescence against the applied Protein Concentration were fitted the above equation using the Graph pad Prism software ver 4.03 (GraphPad, San Diego, USA) revealing the binding dissociation constant, K_D.[101]

5.1.8. Thermfluor Based Melting Point Shift Assay

Thermal shift assay was determined by the Thermal Shift Assay Dye kit (Applied Biosystem, #4461146) followed by the manufacturer's guidelines. The Lck and Grb2 SH2 domain protein was added to a final concentration of 40 μM in buffer. (5 μl Protein Thermal ShiftTM buffer, 2.5 μl Protein Thermal ShiftTM dye). The final volume was 20 μl/well in RT-PCR 96-well white microplates. The temperature gradient was performed in the range of 30-80 ºC, using a 7500 real-time PCR instrument from Applied Biosystems (Carlsbad, CA) with a standard ramp of 1% over the course of 60 min. The compound screening was performed at the 400 μM (in 4% DMSO) concentration. The detection of protein unfolding was performed with an excitation wavelength of 567 nm and an emission of 591 nm. The Prism software ver 4.03 (GraphPad, San Diego, USA) was used to analysed the result.[178]

5.1.9. HPLC Result

All compounds were purified and characterized by LCMS and were shown to be of correct molecular weight and sufficient purity to be used. Below are the LCMS profiles of the 14 potential hits.

PP-pY-X$_5$
Exact Mass: 834.31

PP-pY-X$_6$
Exact Mass: 862.34

PP-pY-X$_{17}$
Exact Mass: 918.29

PP-*p*Y-X$_1$

Exact Mass: 852.36

PP-*p*Y-X$_{33}$

Exact Mass: 898.34

PP-*p*Y-X$_{37}$

Exact Mass: 988.39

PI-*p*Y-X$_{29}$

Exact Mass: 934.32

PP-*p*Y-X₂₅

Exact Mass: 912.36

PI-*p*Y-X₁₆

Exact Mass: 914.38

PI-*p*Y-X₂₅

Exact Mass: 928.39

PI-*p*Y-X₂₈

Exact Mass: 914.38
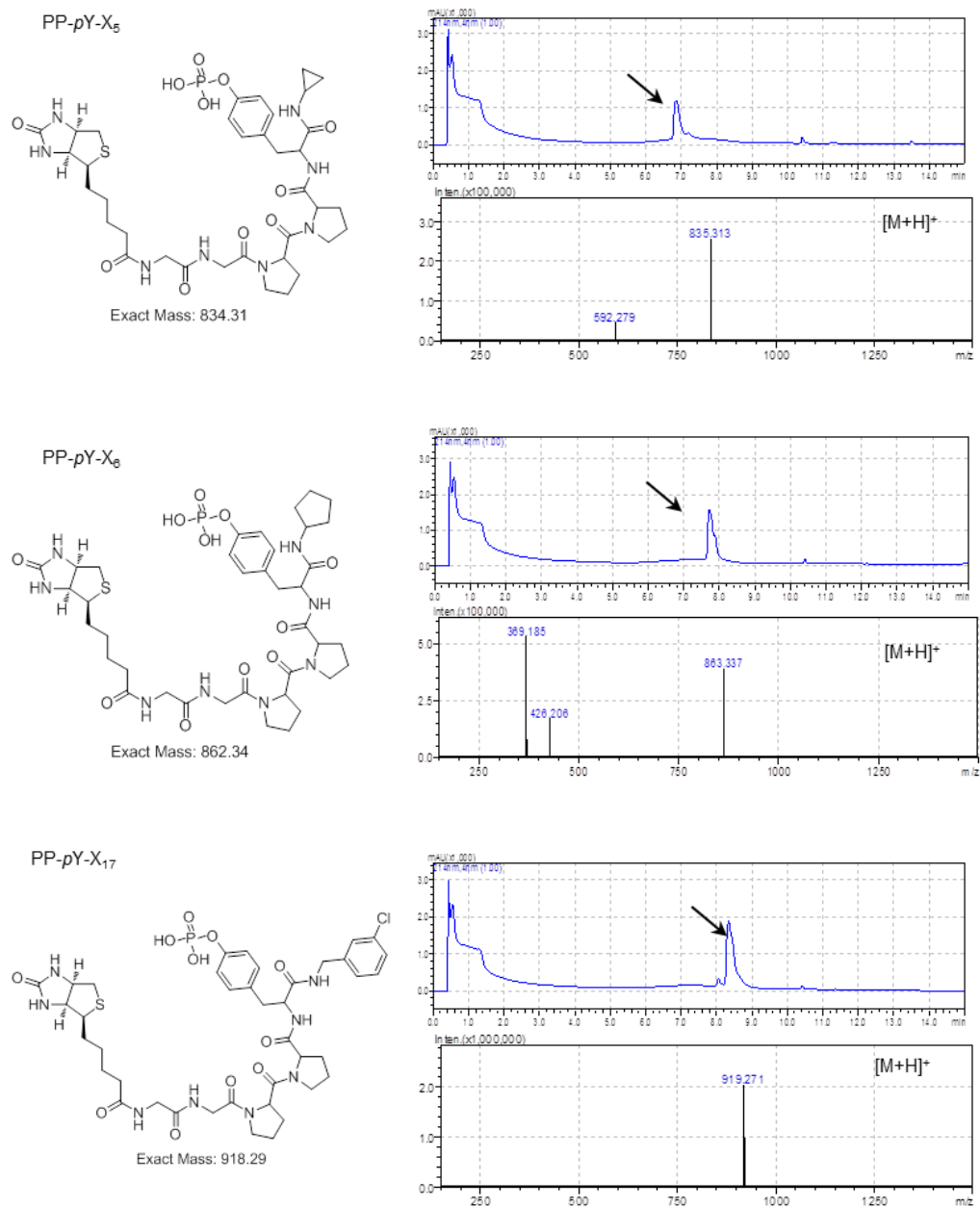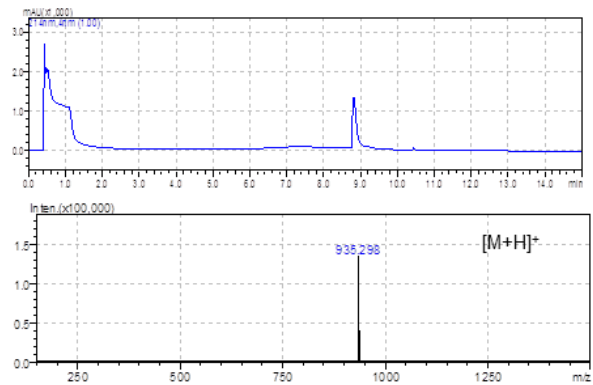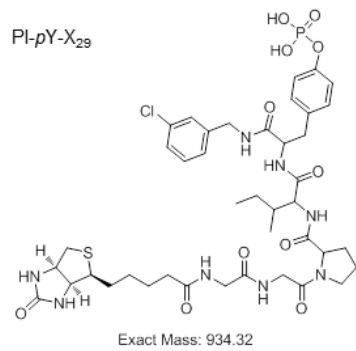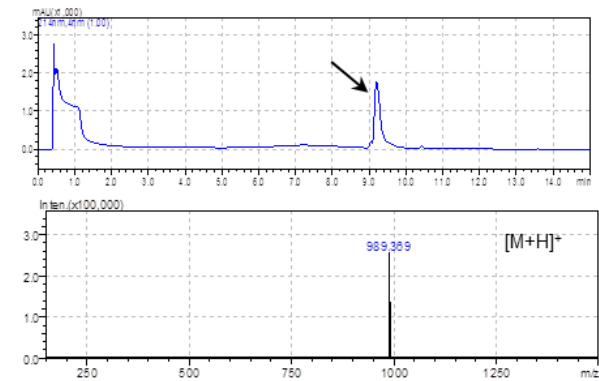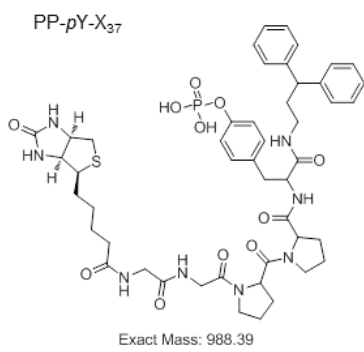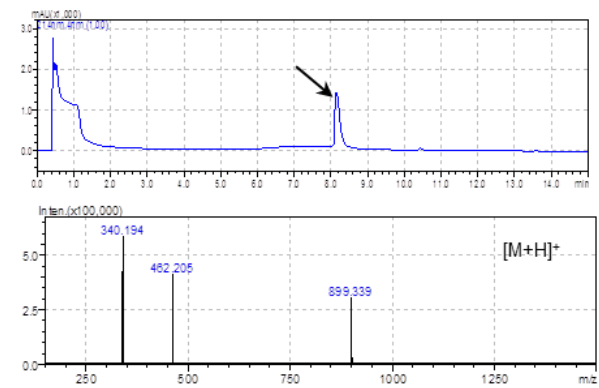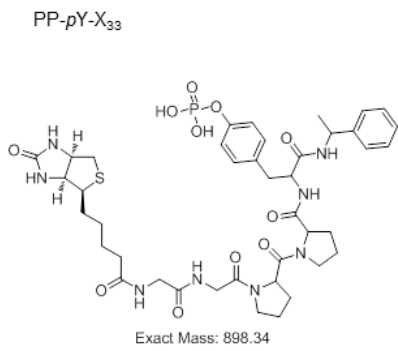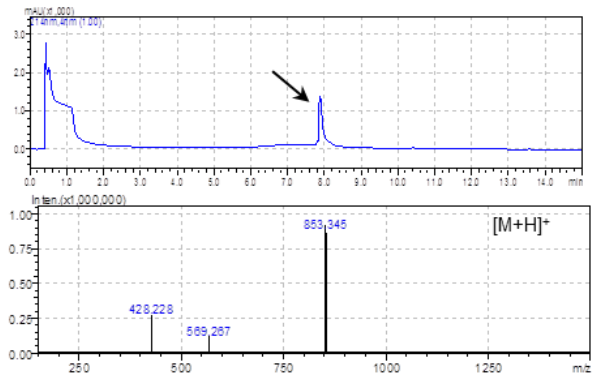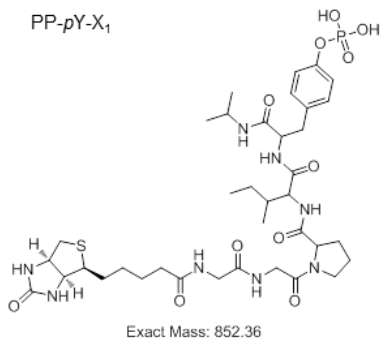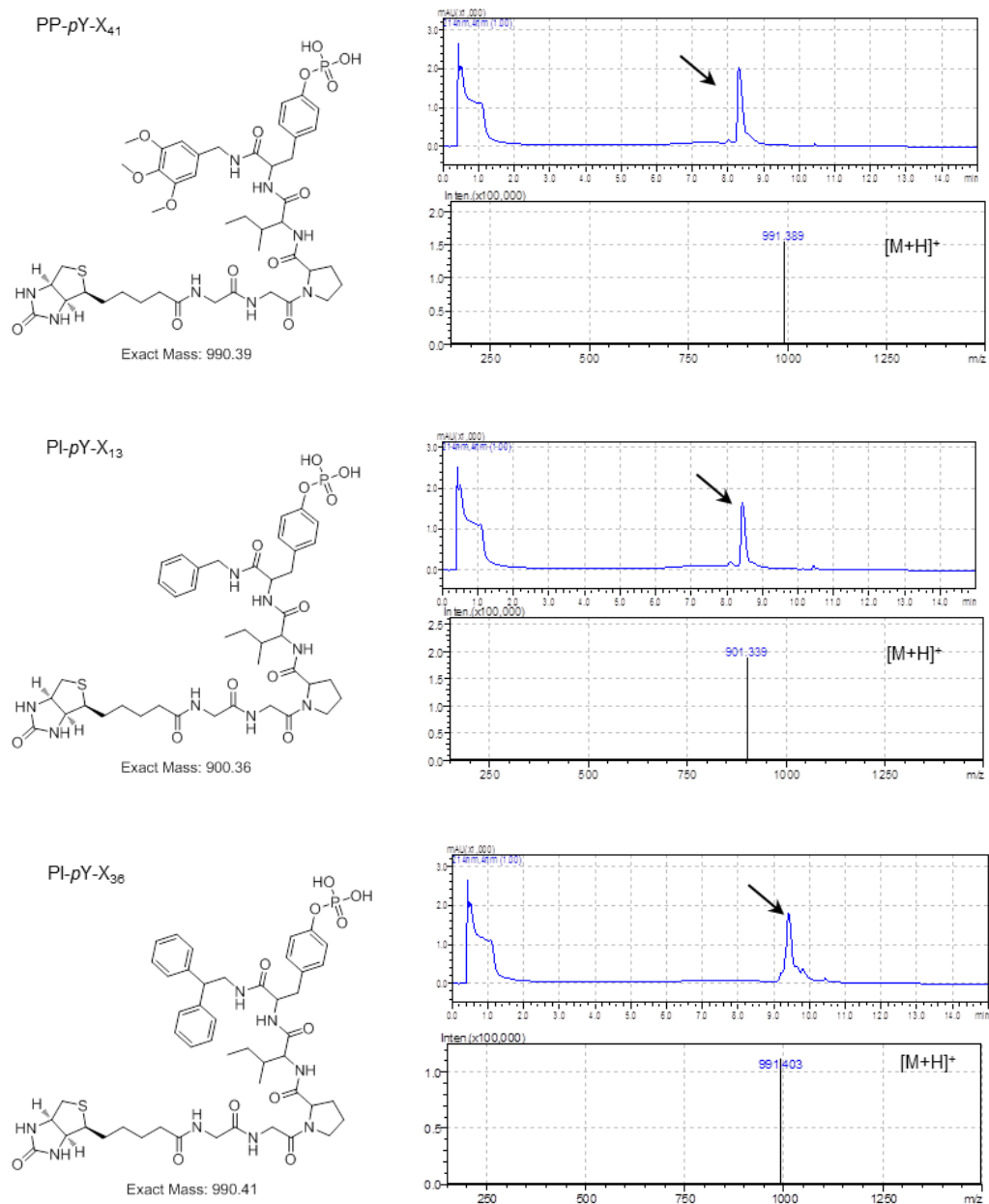
Figure 5.1. LC-MS profiling of resynthesized 14 hits[102]

5.2. Accelerated cellular on- and off-target screening of bioactive compounds using microarrays

5.2.1. General Information

All probes were synthesized as reported.[117,119] Other chemicals were purchased from vendors and used without further purification, unless indicated

otherwise. HBTU, HOBT, EDC were purchased from GL Biochem (China). Tyramide signal amplification kits were purchased from Invitrogen (USA). Gels were scanned on a Typhoon fluorescence gel scanner (GE Healthcare, USA). Microarray slides were scanned using a Tecan Launch LS Reloaded Microarray Scanner (Tecan Trading AG, Switzerland) installed with suitable lasers: Cy3: $\lambda_{Ex/Em}$= 532/575 nm; Cy5: $\lambda_{Ex/Em}$= 633/692 nm.

5.2.2. Expression and Labelling of Protein

Proteins were purified from bacterial cell lysates, as described.[179] The pre-cultures of LB with single colony containing c-Src (32 kDa) were grown at 37 °C to reach OD_{600} 0.6~0.8. The expression was then induced by 0.1mM IPTG and cultures were further grown at 18°C for 18h. The harvested pellets were resuspended with lysis buffer and incubated for 20min. After sonication, the insoluble cell debris was removed by centrifugation. The resulting supernatant was incubated with Ni-NTA beads for 2 h at 4 °C. After washing with wash buffer, target proteins were eluted in elution buffer. The desired proteins were dialysed with Microcon® centrifuge filter device and stored at -20°C in Hepes containing 20% glycerol.

5.2.3. Structure of the Tri-functional Linker and Probes
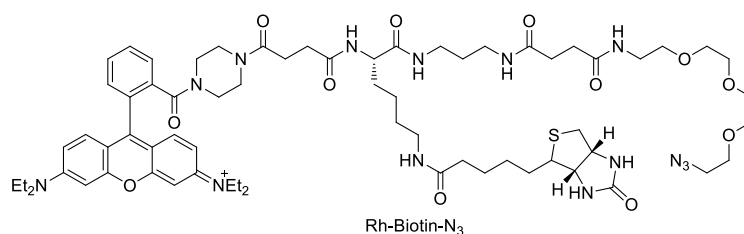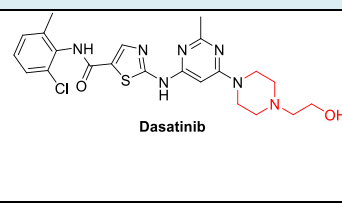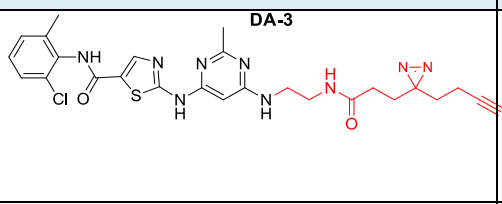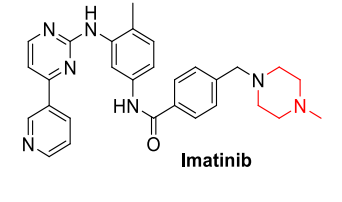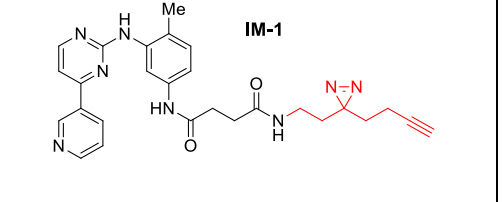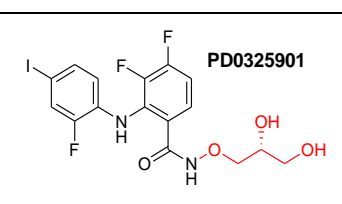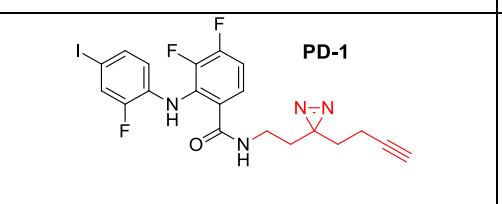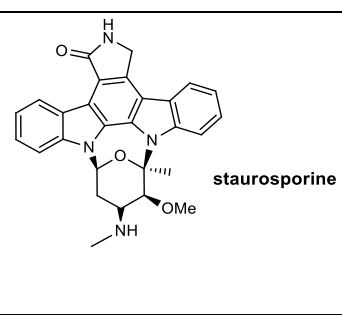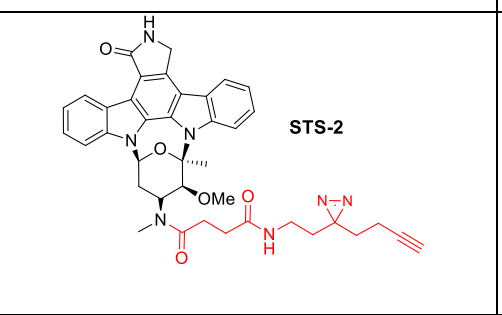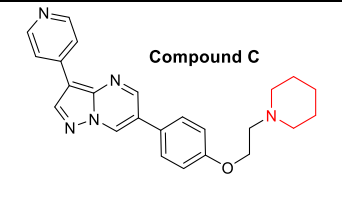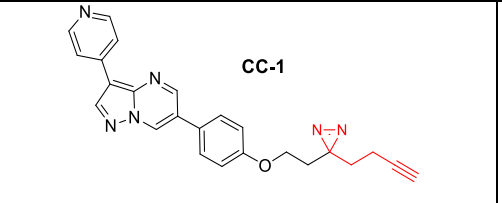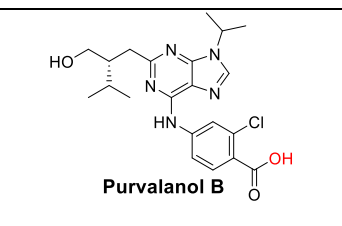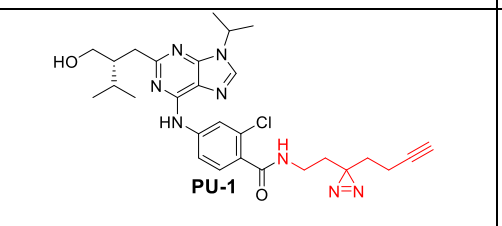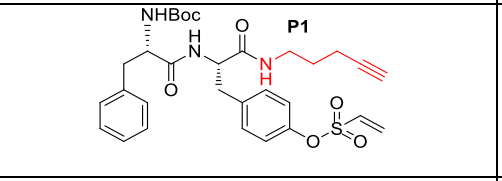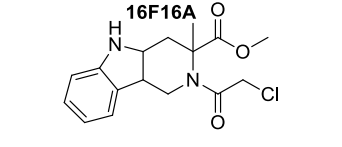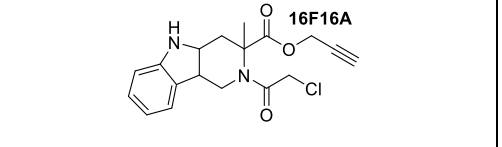
Figure 5.2. Structure of Biotin-TER-N_3[180]

Structure of Inhibitor	Structure of Probe	Known protein targets	Ref.
Dasatinib	**DA-3**	ABL1, SRC	[181]
Imatinib	**IM-1**	ABL, Kit, PDGFR	[182]
PD0325901	**PD-1**	MEK1	[183]
staurosporine	**STS-2**	PRKCH pan-kinase	[184]
Compound C	**CC-1**	AMPK, SRC	[185]
Purvalanol B	**PU-1**	MEK1	[186]
NA	**P1**	PDI	[119]
16F16A	**16F16A**	PDI	[119]
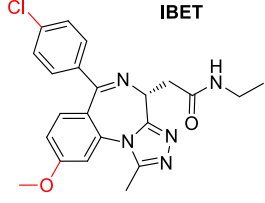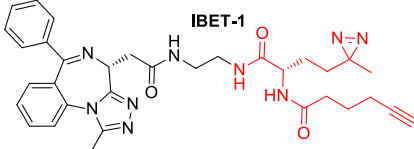
Table 5.1. Design and structures of ABP and A*f*BP

5.2.4. Pure Protein Labelling Experiment

1 mg/ml pure protein (c-Src) was incubated with 10 μM probes for 2 h at room temperature following UV irradiation for 20 min for A*f*BPs. The resulting proteins were tagged with TER-Biotin-N$_3$ via click chemistry. After 2 h of click reaction, labelled proteins were washed with acetone (1×) and methanol (2×) to remove excess dye. After washing, 2% SDS in TBS was added to dissolve proteins; final concentration was around 1mg/ml. In gel fluorescent scanning was used to visualize the labelled protein bands. Both in-gel florescence scanning (FL) and Coomassie staining (CBB) were always carried out on the gels, following separation of labelled samples by SDS-PAGE.

5.2.5. Spike-In and *In-situ* Proteome Labelling

The procedure is similar to our previous reported methods but with further modification as described.[119] For spike in labelling, probes were added to a final concentration of 10 μM into 1mg/ml of fresh cell lysate with or without spiked proteins and incubated for 2 h at room temperature. For activity-based probes, 20 min UV irradiation was required. The resulting lysate was tagged with TER-Biotin-N3 via click chemistry. The reactions were further incubated for 2 h at room temperature with gentle mixing before the reaction was terminated by addition of acetone. Precipitated proteins were subsequently washed with methanol twice to remove the excess dye following dissolved in

TBS with 2% SDS. Around 20 μg proteins were separated by SDS-PAGE gel and then visualized by in-gel fluorescence scanning. For *in situ* labelling, cells were grown between 80% to 90% confluence. The growth medium was removed and the cells were washed with PBS twice. Then cells were treated with DMEM containing 10 μM probes. After 5 h incubation at 37°C/5% CO2, the medium was aspirated and cells with activity-based probes were irradiated by UV for 20 min on ice. The cells were trypsinized and spun down. The resulting cell pellets were resuspended in lysis buffer and homogenized by sonication, and brought with final concentration of 4-8mg/ml in Hepes buffer. All subsequent procedures were performed in the same was as described for the *in vitro* labelling experiments.

5.2.6. Array-based Target Identification and Validation

Avidin slides were prepared as previously reported.[4] Slides were blocked with 1% BSA for 1 h at room temperature before sample spotted to reduce non-specific binding as much as possible. The labelled proteins were dissolved in Spotting Buffer (20% DMSO in TBS) in 1:1 ratio and spotted onto avidin slides. Spike-in and *in situ* labelling samples were spotted onto slides with a final concentration of 2 or 4 mg/ml. After incubation for 2 h, avidin slides were washed with 0.5% TBST for 6 times, 5 min per time. Slides with three sub-grids were first incubated with different antibodies separately for 1.5 h at room temperature. After washing with TBST, avidin slides were incubated with secondary antibody for 1 h at room temperature following washed with TBST. Targets were identified by Cy5 tyramide signal amplification kits according to manufacturer's protocol with slight modifications. Generally, tyramide was

diluted into 1:400 with amplification buffer containing 0.015% H_2O_2. A 200 μl reagent was applied onto each slide under cover slip with 15 min exposure time at room temperature. After development, the resulting slides were washed with TBST for 5 times and scanned by Tecan Launch LS Reloaded Microarray Scanner in the Cy5 channel. Slides were generally scanned under a PMT setting of 140.

Antibody	Company	WB	Microarray	Product code
anti-c-Src	Santa Cruz	1：500	1：100	sc-8056
anti-PDI	Santa Cruz	1：500	1：50	sc-166474
anti-PKA	Santa Cruz	1：2000	1：500	sc-903
anti-MEK1	Abcam	1：2000	1：100	ab32091
anti-CDK1	Abcam	1：10000	1：1000	ab133327
anti-Tubulin	Abcam	1：5000	1：500	ab6064
anti-rabbit	Cell signalling	1：5000	1：500	7074S
anti-mouse	Santa Cruz	1：5000	1：500	sc-358917

Table 5.2. Antibodies used for screening and their correspoding dilution

5.2.7. Data Extraction and Analysis

Microarray data was extracted using the Array-Pro®software. Values from duplicated points were background subtracted and averaged. For each antibody, signals were first normalized by its mean and subtracted by DMSO treated controls. The coloured array heatmaps were produced using Cluster 3.0 and Java Treeview 3.0 (http://sourceforge.net/projects/jtreeview/files).

5.2.8. Pull-down Experiment

To identify potential cellular targets and *off*-targets of the probes and their parental compounds, pull-down (PD) experiments were carried out, followed by Western blotting (WB). The general pull-down procedure was

based on previously reported procedures.[161,187] *In situ* labelling was carried out as described above. After labelling, the reaction was reacted by click chemistry with TER-biotin-N_3 before acetone precipitated, methanol washed and resolubilized in 2% SDS in TBS with brief sonication. This resuspended sample was then incubated with avidin-agarose beads (100 μL/mg protein) for 3 h at room temperature. After centrifugation, supernatant were removed. The beads were washed with 0.5% SDS once and 0.1% SDS PBS for four times. After washing, the beads were boiled in $1 \times$ SDS loading buffer (200 mM Tris pH 6.8, 400 mM DTT, 8% SDS) for 15 min. Control PD using DMSO was carried out concurrently. WB experiments were carried out as previously described using the corresponding antibodies.

5.2.9. Recombinant PKA Labelling

1 μg/ml pure protein (PKA) was incubated with probes (**P1** and **PU-1**) of varying concentrations (0.1, 0.2, 0.5, 1, 2, 5 μM) for 2 h at room temperature before UV irradiation for 20 min for A*f*BPs. The resulting proteins were tagged with TER-Biotin-N_3 via click chemistry for 2 h. SDS-PAGE gel profiling were carried out for the sample without further purification before silver stain.

5.3. Intracellular Delivery of Functional Proteins and Native Drugs by Cell-Penetrating Poly(disulfide)s

5.3.1. General Information.

All chemicals were purchased from commercial vendors and used without further purification. All ^{1}H NMR and ^{13}C NMR spectra were carried out on a Bruker ACF-300 or 500 MHz NMR spectrometer. Chemical shifts were

reported in parts per million relative to internal standard tetramethylsilane ($Si(CH_3)_4 = 0.00$ ppm) or residual solvent peaks (DMSO-d_6 = 2.50 ppm, DMF-d_7 = 8.03, 2.92, 2.75 ppm, $CDCl_3$ = 7.26 ppm, D_2O = 4.79 ppm). ^{1}H NMR data was reported as follows: chemical shift in ppm from the respective internal standard, multiplicity (s = singlet, br = broad, d = doublet, t = triplet, m = multiplet, dd = doublet of doublets, ddd = doublet of doublets of doublets, dt = doublet of triplets or overlap of nonequivalent resonances, dq = doublet of quartets), integration, coupling constant. Mass spectra were recorded on a Finnigan LCQ mass spectrometer, a Shimadzu LC-IT-TOF spectrometer or a Shimadzu LC-ESI spectrometer. Analytical HPLC system was equipped with an autosampler, using reverse-phase Phenomenex Luna 5 μm C_{18} 100 Å 50 × 3.0 mm columns. 0.1% TFA/H_2O and 0.1% TFA/acetonitrile were used as eluents for all HPLC experiments. Analytical GPC was carried out on a Waters Alliance e2965 model with an Waters™ Ultra-hydrogel Linear Column (10 μm, 7.8 mm × 300 mm, 500-10 M) equipped with a guard column (200 Å, 6 μm, 6 mm × 40 mm) and a Waters™ RI 2414 detector. Pure deionized water was used as the mobile phase at the flow rate of 0.8 mL/min. In-gel fluorescent scanning was carried out with a Typhoon 9410 fluorescence gel scanner (GE Amersham). High-content screening (HCS) were acquired with an In-Cell Analyzer 2200 (GE Healthcare Life Science) equipped with a solid-state, multi-wavelength illuminator (λ_{ex} = 390 ± 18 nm; λ_{ex} = 438 ± 24 nm; λ_{ex} = 475 ± 28 nm; λ_{ex} = 513 ± 17 nm; λ_{ex} = 542 ± 27 nm; λ_{ex} = 575 ± 25 nm; λ_{ex} = 632 ± 22 nm). Images were processed and quantified with In-Cell Developer Toolbox software 1.9.2. Confocal images were acquired on a Leica TCS SP5X Confocal Microscope System equipped with Leica HCX PL APO 40×/0.85 Dry CORR CS, 405 nm

diode laser, argon ion laser, white laser (470 nm to 670 nm, with 1 nm increments, with 8 channels AOTF for simultaneous control of 8 laser lines, each excitation wavelength provides 1.5 mV), a PMT detector ranging from 420 nm to 700 nm for steady state fluorescence. Images were processed with Leica Application Suite Advanced Fluorescence (LAS AF). Antibodies were purchased from the following vendors: anti-PARP1 (#9532, Cell Signaling Technology) Rabbit, anti-cleaved caspase-3 (#9661, Cell signaling Technology) Rabbit, anti-His (34460, QIAGEN) HRP, anti-β-tubulin, (#ab6064, abcam) Rabblit, HRP-conjugated anti-Rabbit IgG (#7074, Cell signaling Technology), Alexa Fluor® 488 Goat Anti-Rabbit IgG (#A11008, Invitrogen). Western blotting was done by using the ECL Plus Western Kit (GE Healthcare). TCO-PEG$_3$-Maleimide and ThioLinker-TCO^{TM} were purchase from http://www.clickchemistrytools.com/. Endosome inhibitors 5,5'-dithiobis-2-nitrobenzoic acid (DTNB), chlorpromazine (CPZ), wortmannin (w), and methyl-β-cyclodextrin (mβCD) were purchased from Sigma-Aldrich. Fluorescence-activated cell sorting (FACS) experiments were carried out on a BD Accuri™ C6 cell analyzer (Beckman Coulter). Recombinant (His)$_6$-tagged caspase-3 and BRD-4 were expressed, purified as previous described.[1] The Tetrazine-containing dye reporter, TER-Tz2 (in Figure S1D) was prepared as previously described.[2] Tetraethylorthosilicate (TEOS), 3-(trihydroxysilyl)propyl methylphosphonate monosodium (TPMP), N-cetyltrimethylammonium bromide (CTAB) were purchased from Sigma. Doxorubicin hydrochloride (Dox) was obtained from Selector Ltd. All other reagents were analytical grade and used without purification. All aqueous solutions were prepared using diethylpyrocarbonate (DEPC)-treated ultrapure

water from a Milli-Q system. Transmission electron micrographs (TEM) were carried out on a JEOL JEM 200CX transmission electron microscope with an accelerating voltage of 200 kV. The ordered structures of the MSN materials were confirmed by small angle powder XRD on a Thermo ARL^{TM} SCINTAG X'TRA diffractometer using Cu-Kα radiation. The specific surface areas were calculated from the adsorption data in the low pressure range using the BET model and pore size was determined using the Barrett-Joyner-Halenda (BJH) method. UV-Vis absorption and fluorescence spectra were measured by using Shimadzu UV-Vis spectrometer. The particle size and distribution was determined by dynamic light scattering (DLS) on a Malvern Nano-ZS90 particle size analyzer.

5.3.2. Chemistry.

All CPD building blocks were synthesized by using procedures modified from published protocols,[151,188-191] as described in detail below.
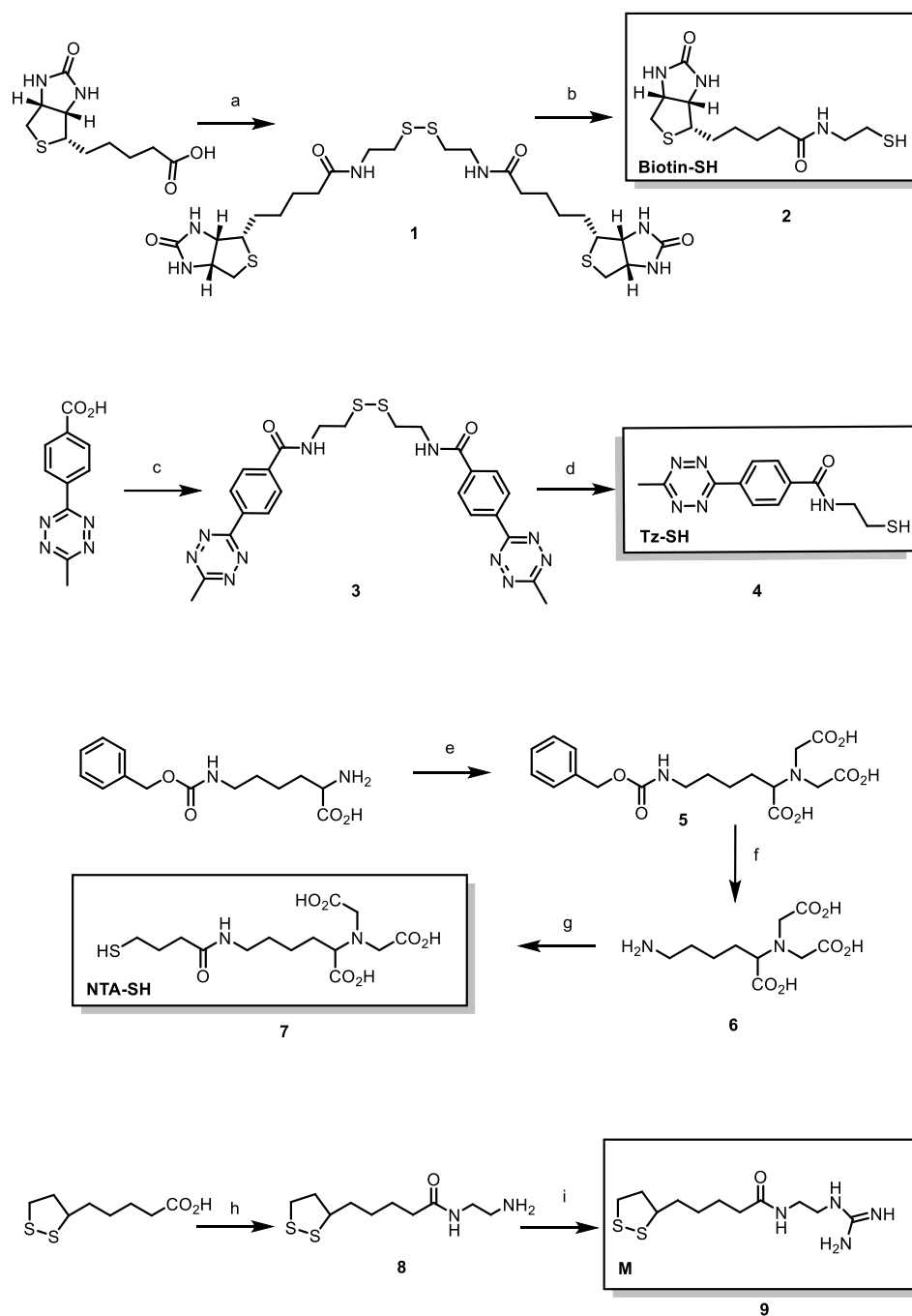
Scheme 5.2. Initiators synthesis.[157] (a) EDC.HCl, cystamine dihydrochloride, trimethylamine, DMF, 18 h, r.t., 91%; (b) TCEP, 50% DMF/H2O, 30 min, r.t.; (c) EDC.HCl, cystamine dihydrochloride, trimethylamine, DMF, overnight, r.t., 91%; (d) TCEP, 50% DMF/H2O, 30 min, r.t. ; (e) Bromoacetic acid, 2 M aqueous NaOH, 2 h, r.t., then 19 h, 50 ℃, 80%; (f) H2, Pd/C, MeOH, 12 h, r.t., 75%; (g) 4-butyrothiolactone, NaHCO3, 72 ℃, 15 h, 70%; (h) CDI, ethylene diamine, DCM, 2 h, r.t., 78%; (i) 1H-pyrazole-1-carboxamidine hydrochloride, DCM, 4 h, r.t., 56%. r.t. = room temperature

Biotin disulfide (1). Biotin disulfide was synthesized and characterized by following a reported procedure.[188] ^{1}H NMR (300 MHz, DMF-d7) δ 8.10-7.96 (m, 2H), 6.34 (dd, *J* = 23.5, 6.8 Hz, 4H), 4.47 (dd, *J* = 7.7, 5.1 Hz, 2H), 4.32-4.28 (m, 2H), 3.47 (dd, *J* = 12.8, 6.5 Hz, 4H), 3.31-3.11 (m, 2H), 2.93-2.91 (m, 4H), 2.74 (s, 4H), 2.32 (dt, *J* = 29.1, 7.3 Hz, 4H), 1.80-1.42 (m, 12H).

Initiator Biotin-SH (I1; 2). 50 mM compound **2** in 50% DMF/water was acquired by 30 min vigorous agitation at room temperature of an equal volume mixture of 1 M TCEP (in H_2O) and 50 mM biotin disulfide (in DMF), and immediately used, without further purification, for subsequent polymer synthesis.

Tetrazine disulfide (3). A solution of 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid (100 mg, 0.46 mmol), EDC·HCl (106.4 mg, 0.56 mmol), cystamine dihydrochloride (57 mg, 0.25 mmol) and triethylamine (211 μL, 1.52 mmol) in DMF (5 mL) was kept at room temperature for 18 h, then concentrated *in vacuo*. The resulting residue was triturated with H_2O, providing the pure tetrazine disulfide as a pink solid (115 mg, 91%). ^{1}H NMR (500 MHz, DMSO-d6) δ 8.90 (t, *J* = 5.5 Hz, 2H), 8.53 (d, *J* = 8.5 Hz, 4H), 8.09 (d, *J* = 8.6 Hz, 4H), 3.62 (dd, *J* = 12.8, 6.4 Hz, 4H), 3.12-2.90 (m, 10H). ^{13}C NMR (75 MHz, DMSO-d6) δ 167.28 (s), 165.57 (s), 162.83 (s), 137.49 (s), 134.24 (s), 128.17 (s), 127.33 (s), 37.09 (s), 20.86 (s). One carbon peak was merged with solvent pick. MS (ESI): m/z [M+Na]$^+$ calcd 571.1417, found 571.1413.

Initiator Tz-SH (I3; 4). 50 mM compound **4** in 50% DMF/H_2O was acquired by 30 min vigorous agitation at room temperature of an equal volume mixture of 1 M TCEP (in H_2O) and 50 mM tetrazine disulfide (in DMF), and

immediately used, without further purification, for subsequent polymer synthesis.

N-(5-Carbobenzyloxyamino-1-carboxypentyl) iminodiacetic acid (5). Compound **5** was synthesized and characterized according to a published protocol.[4] ^{1}H NMR (300 MHz, DMSO-d_6): δ 7.42-7.13 (m, 5H), 5.00 (s, 2H), 3.61-3.39 (m, 4H), 3.35 (t, 1H, J = 7.2 Hz), 2.96 (d, 2H, J = 5.9 Hz), 1.79-1.01 (m, 6H).

N-(5-Amino-1-carboxypentyl) iminodiacetic acid (6). Compound **6** was synthesized and characterized according to a published protocol.[4] ^{1}H NMR (500 MHz, D_2O): δ 3.93 (s, 5H), 3.01 (t, 2H, J = 7.2 Hz), 1.96-1.56 (m, 6H).

Initiator NTA-SH (I2; 7). Compound **7** was synthesized and characterized according to a published protocol.[5] ^{1}H NMR (300 MHz, D_2O): δ 3.86 (s, 5H), 3.22 (s, 2H), 2.75 (t, 2H), 2.38 (t, 2H, J = 7.1 Hz), 2.03-1.59 (m, 8H).

N-(2-aminoethyl)-5-(1,2-dithiolan-3-yl) pentanamide (8). Compound **8** was synthesized and characterized according to a published procedure.[6] ^{1}H NMR (500 MHz, CDCl$_3$): δ 5.94 (sbr, 1H), 3.59-3.56 (m, 1H), 3.30 (dt, J = 11.4 Hz, 2H), 3.21-3.08 (m, 2H), 2.88-2.79 (m, 2H), 2.55-2.31 (m, 1H), 2.22 (t, 2H, J = 7.4 Hz), 1.91-1.84 (m, 1H), 1.76-1.40 (m, 8H).

Monomer M (9). Compound **9** was synthesized and characterized based on a published procedure.[151,191] ^{1}HNMR (500 MHz, MeOD): δ 8.44 (d, 1H, J = 3.0 Hz), 7.94 (d, 1H, J = 1.5 Hz), 7.57 (d, 1H, J = 1.7 Hz), 6.71 (dd, 1H, J = 3.0, 1.6 Hz), 6.29 (t, 1H, J = 2.1 Hz,), 3.99 (s, 1H), 3.67-3.40 (m, 1H), 3.30-3.22 (m, 4H), 3.20-2.98 (m, 2H), 2.54-2.33 (m, 1H), 2.21 (t, J = 7.4 Hz, 2H), 1.94-1.76 (m, 1H), 1.77-1.32 (m, 6H).

General procedure for polymer synthesis and characterizations. The CPD polymers were synthesized and characterized according to published procedures,[151,191] and modified where applicable. Briefly, stock solutions of the monomer (**M**; 2 M in DMF), initiator precursor-disulfide (50 mM in DMF), terminator (**T**; iodoacetamide, 0.5 M in H_2O, fresh), TCEP (1 M in H_2O) and TEOA buffer (1 M, pH = 7.0) were prepared. The initiator (**I1/I2/I3**) was first generated *in situ*, by mixing 5 μL of the initiator precursor stock solution with 5 μL of the TCEP solution (**NTA-SH** was treated with TCEP as well, to ensure any possible disulfide formation) followed by vigorous shaking for 30 min at room temperature. Subsequently, 80 μL of TEOA buffer (for **Biotin-SH** and **NTA-SH** initiated polymerization) or DMF/TEOA 50% mixture (for **Tz-SH** initiated polymerization) and 10 μL of the monomer stock solution were added to the reaction mixture. After 30 min of agitation at room temperature, the polymerization reaction was quenched by addition of 1.9 mL of the terminator stock solution. The resulting polymer was purified with a NAP™-5 desalting column (GE Healthcare) against H_2O in the same day, by following protocols provided by the vendor. The elution was lyophilized and kept in -20 ℃. The molecular weight of each polymer was determined by analytical GPC under conditions as described in the General Information, with standard molecular weight markers used for calibration. Quantification of **ᴛᴢCPD** was achieved by UV-Vis measurements of the tetrazine absorbance at the wavelength of 520 nm, and the tetrazine disulfide (**3**) of different concentrations (50, 100, 200, 400 and 500 μM) was used to generate the corresponding calibration curve (Figure S3C). The polymer concentration of the obtained stock solution was estimated to be ~200 μM. The concentrations of other CPDs were assumed to be the same. We

tried to make the free thiol of various initiators, but found they were oxidized rapidly, and therefore were of limited practical use. Therefore we developed the TCEP-reducing protocols to make the thiols in situ and use them immediately. We took pain in optimizing the amount of TCEP (ensuring min TCEP needed was used). We initially tried 1-fold and 3-fold TCEP to reduce tetrazine disulphide. But it takes more than half an hour to complete the reduction process. Therefore, 10-fold of TCEP was used to make sure the complete and fast reduction of initiator. We found that this amount of excess TCEP did not affect the subsequent polymerization reaction, as during polymerization, the concentration of TCEP was significantly diluted. Furthermore, during polymerization, the amount of monomer used was much higher than that of TCEP (4 folds). Finally, as soon as the polymerization reaction was completed, the newly formed polymer was immediately purified by dialysis which would remove all remaining TCEP.

5.3.3. Protein Bioconjugation by CPD.

Figure 5.3. Labeling process of avidin, BRD-4, BSA and **Ab^{FC}**.[157] (A) Recombinant avidin was labeled with Cy5-NHS to generate **Avidin^{Cy5}**. (B) Recombinant (His)₆-tagged BRD-4 was labeled with Cy5-NHS to generate **BRD-4^{Cy5}**. (C) Recombinant BSA was first labeled with TCO via site-specific conjugation between TCO-PEG₃-Maleimide and the single free cysteine present in the protein. Next, TCO-BSA was fluorescently labeled with Cy5-NHS to obtain **TCO-BSA^{Cy5}**. (D) For antibody conjugation, a commercially available FC-labeled antibody (i.e. **Ab^{FC}**) was used, and the TCO moiety was site-specifically introduced to the antibody by first reducing the disulfide bond linkage in the antibody, followed by treatment with the ThioLinker-TCO™ reagent to obtain **TCO-Ab^{FC}**.

Figure 5.4. Formation of **CPD-Protein** and **CPD-MSN-Dox**.[157] (A) The formation of **CPD-BSA** between **TzCPD** and **TCO-BSA^{Cy5}**, and **CPD-Ab** between **TzCPD** and **TCO-Ab^{FC}**, via TCO-tetrazine bioorthogonal ligation. (B) Non-covalent formation of **CPD-Avidin** between **BiotinCPD** and **Avidin^{Cy5}**, **CPD-BRD-4** between **Ni-NTACPD** and **BRD-4^{Cy5}**, and **CPD-Casp-3** between **Ni-NTACPD** and (His)$_6$-tagged caspase-3, via affinity interaction. (C) "Capping" of small molecule drug-loaded **MSNs** with CPD via electrostatic charge-charge interaction between the negatively charged **MSN-Dox** and positively charged CPD (in this case, **BiotinCPD** was used).

20 μL of NaHCO$_3$ (500 mM in H$_2$O) and 10 μL of Cy5-NHS (GE Healthcare Life Science in DMSO) were added to an avidin solution (10 mg/mL, 100 μL in H$_2$O). After 30 min incubation on ice, the excess dye was removed

with a MicroSpin G25 Column (GE Healthcare Life Sciences) by following the vendor's protocol (Figure 5.3A). The dye-to-protein ratio in the labeling reaction of **Avidin^{Cy5}** was determined to be 0.68 by UV/fluorescence measurement (at 650 nm) by using a Nanodrop™ fluorescence spectrometer (Thermo Scientific). Similarly, 20 μL of NaHCO$_3$ (500 mM in H$_2$O) and 10 μL of Cy5-NHS (GE Healthcare Life Science in DMSO) were added to a (His)$_6$-tagged BRD-4 solution (10 mg/mL, 100 μL in PBS).[32,192] After 30 min incubation on ice, the excess dye was removed by MicroSpin G25 Column (Figure 5.3B). The dye to protein ratio of **BRD-4^{Cy5}** was determined to be 0.52 as determined by UV/fluorescence measurement (at 650 nm). For labeling of BSA, 1 mg of the protein dissolved in 100 μL of PBS (*p*H = 7.4) was reacted with 2.5 μL of TCO-PGE$_3$-Maleimide (stock: 2 mM in DMSO) for 4 h at room temperature. Subsequently, excessive TCO-PGE$_3$-Maleimide was removed by G25 Column. To the resulting protein fraction, 20 μL of NaHCO$_3$ (500 mM) and 10 μL of Cy5-NHS were next added. After incubation on ice for 30 min, the excessive dye was removed by G25 column (Figure 5.3C). The dye to protein ratio of **TCO-BSA^{Cy5}** was determined to be 0.86 determined as earlier described.

The labeling of the commercially available, fluorescently labeled antibody was carried out by following published protocols.[159] Briefly, 1 mg/mL of the **Ab^{FC}** (8 nM) in PBS buffer (100 mM sodium phosphate, 150 mM sodium chloride, 10 mM EDTA, *p*H = 7.5) was treated with 10 mM TCEP (stock: 500 mM in H$_2$O) for 30 min at room temperature, to reduce the disulfide bond. The excessive TCEP was subsequently removed by G25 column as earlier described. The resulting desalted **Ab^{FC}** was subsequently incubated with 40 nM of the

ThioLinker-TCO™ (from stock solution; 4 mM in DMSO) for 6 h. At the end, the excessive ThioLinker-TCO™ reagent was removed by G25 spin column. The labeling efficiency of **TCO-Ab^FC** was determined by in-gel fluorescence scanning of the labeled protein upon conjugation with a Tetrazine-containing dye reporter (TER-Tz2 in Figure 5.3D), as previously described.[187] As protocols reported, the ratio of ThioLinker-TCO™ to on **Ab^FC** is 4 to 1. Once conjugated with TER-Tz2, there was a shift around 10 KDa of **TCO-Ab^FC** in SDS-PAGE gel.

As shown in Figure 5.4, for **CPD-BSA**, 5 μM of **Tz CPD** was mixed with 5 μM of **TCO-BSA^Cy5** in PBS buffer, resulting in instantaneous and quantitative formation of the desired product, which was then used directly without further purifications. For **CPD-Avidin**, 20 μM of **Biotin CPD** was incubated with 5 μM of **Avidin^Cy5** for 10 min at room temperature in PBS to obtain the desired product. For the generation of **Ni-NTA CPD**, 15-fold excess of NiCl$_2$ was added to **NTA CPD** in 10 mM HEPES (pH = 7.0), and the resulting mixture was incubated overnight at 4 °C.[193] Free nickel ion was removed by dialysis with a slide-A-lyzer™ (3.5K Mw cutoff; Thermo Scientific) against 10 mM HEPES (pH = 7.9) for three times. Subsequently, 15 μM of the resulting **Ni-NTA CPD** was incubated with 5 μM **BRD-4^Cy5** to give **CPD-BRD-4**. The preparation of successful CPD conjugates and CPD depolymerization on selected protein samples were analyzed by SDS-PAGE gel and in-gel fluorescence scanning (Figure S3B). Briefly, 10 nM of **CPD-Avidin** was treated with GSH (10 mM) or HeLa lysate (1 mg/mL, freshly prepared in PBS) for 1 h at 37 °C.[153] Subsequently, DTT-free loading dye was added to the mixture, and without boiling (to prevent dissociation of avidin/biotin complex), the samples were directly analyzed by

15% SDS-PAGE gel (DTT-free) followed by in-gel fluorescence scanning. **Avidin**$^{\textbf{Cy5}}$ (CPD-free) was carried out concurrently as negative control. For preparation of **CPD-Casp-3**, **CPD-Ab** and **CPD-MSN-Dox**.

5.3.4. Biochemical and Cellular Experiments

5.3.4.1 Confocal Laser Scanning Microscope (CLSM)

HeLa cells were seeded at 4-well glass-bottom dish (Greiner Bio-One) and grown until 50~60% confluency. Upon medium removal, cells were treated with 250 μL of 50 nM **CPD-Protein** (from 5 μM stock, freshly prepared in PBS) in DMEM medium. HeLa cells incubated with **protein**$^{\textbf{Cy5}}$ (CPD-free) were done concurrently as negative controls. The cells were incubated for 1 h at 37 ^{o}C before washing with PBS (containing 0.1 mg/mL heparin) for three times. For protein transfection experiment, an equal amount of **protein**$^{\textbf{Cy5}}$ (12.5 pmol) was transfected by using the commercially available Pro-Ject$^{\text{TM}}$ Protein Transfection reagent (Thermo Scientific) at the recommended volume by following the vendor's protocols. Briefly, 12.5 pmol of **protein**$^{\textbf{Cy5}}$ in 25 μL PBS was mixed with 2.5 μL of the Pro-Ject$^{\text{TM}}$ reagent for 5 min at room temperature. The complex was then added to an imaging-compatible petri dish, followed by incubation further for 1 h in serum-free DMEM medium, and washing with PBS (containing 0.1 mg/mL heparin). Where applicable, HeLa cells were further co-stained with CellMask$^{\text{TM}}$ Orange (180 nM) and Hoechst (for an additional 15 min) prior to the completion of the above protein delivery protocol. Subsequently, DMEM medium was re-introduced into the cells, which were then imaged in a Leica TCS SP5X Confocal Microscope System at different detection channels (**CPD-BSA** channel: λ_{ex} = 643 nm, λ_{em} = 665-740 nm;

CellMask™ membrane tracker channel: λ_{ex} = 543 nm, λ_{em} = 555-650 nm; nuclear-staining Hoechst channel: λ_{ex} = 405 nm, λ_{em} = 440-470 nm).

For temperature-dependent experiments, all buffers and media were kept at the respective temperature (i.e. 4 ℃, 25 ℃ or 37 ℃) for 1 h prior to use. HeLa cells were pre-incubated for 1 h at 4 ℃, 25 ℃ or 37 ℃, respectively, followed by treatment with 50 nM of **CPD-BSA** for an additional hour. Upon washing with PBS (containing 0.1 mg/mL heparin) for three times, then re-introduction of DMEM, CLSM was immediately performed. For real-time imaging of **CPD-BSA** cellular uptake and sub-cellular localization, HeLa cells were first stained with Hoechst and LysoTracker™ DND-99 for 15 min at 37 ℃. Upon washing, the cells were treated with 50 nM of **CPD-BSA** in DMEM medium, and immediately imaged over the course of 30 min (LysoTracker™ channel: λ_{ex} = 585 nm, λ_{em} = 600-650 nm).

For 3D imaging, HeLa cells were treated with 50 nM of **CPD-BSA** for different periods of incubation time (15, 30, 60, 120, 240 min). Co-staining was done with LysoTracker™ and Hoechst 15 min prior to the termination of protein delivery. Upon washing (with PBS containing 0.1 mg/mL heparin) and re-introduction of DMEM medium, cells were imaged with the FV1000-X81 confocal microscope (Olympus) equipped with a 40× water immersion objective and the corresponding filters. Images were processed with the VOLOCITY software. Step size: 0.186 μm. Detection channels used were mostly similar (**CPD-BSA** channel: λ_{ex} = 640 nm, λ_{em} = 660-750 nm) (LysoTracker™ channel: λ_{ex} = 561 nm, λ_{em} = 580-650 nm) (Hoechst channel: λ_{ex} = 405 nm, λ_{em} = 415-475 nm).

5.3.4.2. Flow Cytometry, HCS and In-gel Fluorescence Scanning

For Flow Cytometry (FACS) experiments, HeLa cells were seeded in a 12-well plate (Greniner CELLSTAR®) and cultured overnight. Upon medium removal, the cells were washed three times with PBS and treated with 50 nM of **CPD-Protein** in DMEM medium. The cells were incubated for 1 h at 37 °C and washed three times with PBS (containing 0.1 mg/mL heparin). For protein transfection experiments, an equal amount of **protein^{Cy5}** was transfected with the Pro-Ject^{TM} Protein Transfection reagent as earlier described. Co-staining with Hoechst was done, where necessary, and the resulting cells were first imaged with In-Cell Analyzer 2200 (for HCS image acquisition), where applicable, before being detached from the plate by treatment with 200 μL of 0.1% trypsin-EDTA at 37 °C for 2 min. The detached cells were collected by centrifugation (at 700 × g for 2 min) at 4 °C. Upon further washing with cold PBS (200 μL) three times, the cells were suspended in 500 μL of PBS. Experiments with HeLa cells treated with only **protein^{Cy5}** (CPD- and transfection-free) were concurrently carried out as negative controls. Cells were analyzed and counted/quantified on laser excitation at 640 nm on a BD Accuri^{TM} C6 cell analyzer (min. 10000 cells were counted; λ_{ex} = 640 nm). For analysis of protein content, where applicable, the corresponding SDS-PAGE separation followed by in-gel fluorescence scanning was concurrently done on lysates collected from the same batches of cells.

For temperature-dependent experiments, all buffers and media were kept at the respective temperature (i.e. 4 °C, 25 °C or 37 °C) for 1 h prior to use. HeLa cells were pre-incubated for 1 h at 4 °C, 25 °C or 37 °C, respectively, followed by treatment with 50 nM of **CPD-Protein** for an extra hour. Upon

washing with PBS (containing 0.1 mg/mL heparin) for three times, cells were treated in the same way as described above. For time-dependent experiment, cells were treated with **CPD-Protein** (50 nM) for 15, 30, 60, 120, 240, and 480 min, respectively, before washing. HeLa cells treated with 50 nM of **protein^{Cy5}** (CPD- and transfection-free) alone for 480 min were done concurrently as negative controls. Then, the cells were analyzed in the same way as described above. For concentration-dependent experiments, cells were treated with **CPD-Protein** (5, 10, 50 and 100 nM) for 480 min before washing. Cells treated with 100 nM of **protein^{Cy5}** (CPD- and transfection-free) were done concurrently as negative controls. Then, the cells were analyzed in the same way as describe above.

For CPD experiments in different cell lines (PC3, MCF-7, NIH 3T3, A549, HeLa), the corresponding cells were incubated with 50 nM **CPD-BSA** (red; **CPD-BSA**, $\lambda_{ex} = 632 \pm 22$ nm, $\lambda_{em} = 684 \pm 25$ nm) for 1 h before washed and imaged by In-Cell Analyzer 2200.

For imaging-based, high-content screening (HCS) experiments with live cells, HeLa cells were seeded at 12-well plate and grown until 60~70 % confluency and then were treated at desired concentrations and time. After incubation, HeLa cells were washed with PBS (containing 0.1 mg/ml Heparin) three times. DMEM were then introduced before images were acquired. The In-Cell Analyzer automated fluorescence imaging system (GE Healthcare) was used for automated image acquisition. Images were acquired with a $20\times$ objective. The instrument acquired three images of each well randomly with a laser-based autofocus system Hoechst ($\lambda_{ex} = 390 \pm 18$ nm, $\lambda_{em} = 432.5 \pm 48$ nm), **CPD-Protein** ($\lambda_{ex} = 632 \pm 22$ nm, $\lambda_{em} = 684 \pm 25$ nm) and Bright-field ($\lambda_{ex} =$

542 ± 27 nm). All images were taken at the same exposure time for the same batch of experiments ranging from 1500 to 3000 ms, depending on the intensity of fluorescence. The signals would be neither too weak to detect nor overexposed. On average, around 300 cells were analyzed in each well. To quantify the fluorescence signals, images were processed with In-Cell Developer Toolbox software 1.9.2 to identify individual cells and **CPD-Protein** within these cells by following vendor's protocols. Both the nucleus and **CPD-Protein** were segmented via intensity method to remove background first. Then they were filtered via sieve operator to refine a segmented image (nucleus: greater than 20 μm^2; **CPD-Protein**: greater than 50 μm^2) to remove objects or artifacts belonging to a particular size range. Finally, the nucleus were measured by the counts and **CPD-Protein** was measured by the sum of fluorescent intensity.[10]

$$\frac{\text{Sum of } \textbf{CPD} - \textbf{protein} \text{ Intensity}}{\text{Count of Hoechst}}$$

The above formula was used to calculate the average intensity of delivered **CPD-Protein** in each cell. The intensity of three images was used to calculate standard errors and the average intensity of those images was regarded as the fluorescence readout for a particular sample.

5.3.4.3. Cytotoxicity Assay

Cell viability was determined using XTT colorimetric cell proliferation kit (Roche) following manufacturer's guidelines.[7,169,194] HeLa were seeded in a 96-well plate and grown to 50~60% confluency. Cells were treated in duplicate with **CPD-Protein** complex (50 nM) in 0.1 mL DMEM medium. Equal amounts of **protein**[Cy5] were transfected with Pro-Ject[TM] reagent in serum-free medium. Control experiment were done concurrently with PBS buffer alone.

The cells were incubated for 1 h at 37 °C before washing with PBS (containing 0.1 mg/mL heparin) for three time. Subsequently, the DMEM medium was re-introduced into the cells followed by further incubation for 12 h. Then proliferation was assayed by using the XTT colorimetric cell proliferation kit (Roche) following the vendor's protocol (read at 450 nm, ref. at 650 nm). Data represented the average (s.d. for two trials).

5.3.4.4. Mechanistic Studies of CPD-Assisted Protein Uptakes.

HeLa cells were seeded in a 12-well dish and grown until 50~60% confluency. Upon removal of the media, the cells were incubated for 30 min at 37 °C with one of the specific inhibitors at the specified concentration (4.8 mM for 5, 5'-dithioobis-2-nitrobenzoic acid, 10 μg/mL for chlorpromazine, 50 nM for wortmannin and 50 μM for methyl-β-cyclodextrin) in DMEM medium. Then, cells were treated with 50 nM of **CPD-Protein** complex, during which period the inhibitor concentrations were kept constant.[195-197] Upon 1 h incubation at 37 °C, the cell medium was removed by aspiration. Upon washing with PBS (containing 0.1 mg/mL heparin) for three times, the cells were imaged by HCS with In-Cell Analyzer 2200, analyzed by FACS as earlier described.

5.3.5. CPD-Assisted Delivery of Functional Caspase-3

For **CPD-Casp-3** preparations, 5 μM of recombinantly purified, functionally active $(His)_6$-tagged caspase-3 (stock: 20 μM in buffer containing 50 mM HEPES, 10% glycerol, 100 mM NaCl, 5 mM β-ME, pH = 7.0)[1] were incubated with 15 μM $_{Ni-NTA}$**CPD** for 15 min on ice in a HEPES buffer (10 mM HEPES, pH = 7.9) to form **CPD-Casp-3** complex. As a cysteine protease, the availability of the active-site cysteine residue is essential to the caspase activity. Due to the presence of trace amount of the iodoacetamide terminator (**T**) in $_{Ni-}$

NTACPD which could not be removed completely after the polymer was synthesized and purified, the enzymatic activity of **CPD-Casp-3** was partially inactivated prior to being delivered to cells, as a result of alkylation of the active-site cysteine by iodoacetamide. Furthermore, the presence of DTT in the enzymatic assay might help to restore some of the caspase enzymatic activity, by successful reduction of any oxidized active-site cysteine. Therefore, caspase-3 enzymatic assays were performed by following published protocols, with and without DTT.[198] Briefly, 50 pM of **CPD-Casp-3** was incubated with 1 μM of Ac-DEVD-AMC (#BML-P411, Enzo Life Sciences) in HEPES buffer (50 mM HEPES, 50 mM NaCl, 0.1% CHAPS, 10 mM EDTA, 5% glycerol, 10 mM DTT, pH = 7.2), as well as in DTT-free HEPES buffer, at 25 °C, respectively. Liberation of fluorescence (λ_{ex} = 360 ± 40 nm and λ_{em} = 460 ± 40 nm) was recorded by using a BioTek Synergy 4 plate reader until the fluorescence signals were saturated. All measurements were performed in duplicate. Control experiments were done with recombinant (His)$_6$-tagged caspase-3 (the one prior to complexation with **Ni-NTACPD**) of equal amounts and incubated with the same substrate under the same conditions in HEPES buffer). HEPES buffer (with DTT) incubated with substrate under the same condition were carried out as negative controls. For protein transduction experiments, different concentrations of **CPD-Casp-3** (final Conc.: 25, 50, 100 nM) were incubated with HeLa cells in a 96-well plate. Control experiments were done concurrently with HeLa cells treated with **Casp-3** (CPD-free) or **Ni-NTACPD** alone, respectively. After 8 h incubation, XTT assay was carried out to test the extent of cell death/apoptosis caused by the successfully intracellulary delivered active caspase-3. For imaging-based determination of intracellular caspase-3

activity, HeLa cells grown in a 4-well imaging petri dish were similarly treated, by incubation with **CPD-Casp-3** of different concentrations (50 nM) for 2 h. Control experiments were done concurrently with cells treated with buffer, caspase-3 (50 nM) or Ni-NTA**CPD** alone, respectively. At the end of the experiments, cells were washed with PBS (containing 0.1 mg/mL heparin) for three times, then incubated with 40 μM Ac-DEVD-AMC for 2 h at 37 °C before being imaged on a fluorescence microscope (λ_{ex} = 405 nm, λ_{em} = 420-500 nm).[13a] In experiments where mammalian lysates were used to determine caspase-3 activity and the amount of cleaved PARP1 formation, the lysates from above treated cells were prepared, and used for 1) caspase-3 enzymatic assay (25 μg of lysate, 1 μM Ac-DEVD-AMC, same conditions as above), and 2) SDS-PAGE followed by Western blotting (WB) analysis of cleaved PARP1 and (His)$_6$-tagged caspase-3 as previously described.[30,150]

5.3.6. CPD-Assisted Antibody Delivery

For **CPD-Ab** preparation, 5 μM of **TCO-Ab^{FC}** was mixed with 20 μM of Tz**CPD.** The product was formed within minutes and used directly (Figure S2). HeLa cells were treated with 50 nM of **CPD-Ab** for 1 h. Upon washing with PBS (containing 0.1 mg/mL heparin) for three times, cells were stained with Hoechst (15 min) and imaged (**CPD-Ab** channel: λ_{ex} = 488 nm, λ_{em} = 500-560 nm; Hoechst channel: λ_{ex} = 405 nm, λ_{em} = 440-470 nm).

5.3.7. CPD-Assisted Delivery of Small Molecule-Loaded MSNs

MSNs were synthesized according to previously reported methods with some modification.[150,176] Most other experiments were performed based on published protocols with some necessary modifications.[150] Briefly, CTAB (0.1 g, 0.28 mmol), dissolved in a mixture of H$_2$O (48 mL) and NaOH (2 M, 0.35

mL), was stirred and heated to 80 °C under a nitrogen atmosphere. TEOS (0.5 mL, 2.24 mmol) was then added to the solution and stirred continually. After 15 min reaction, TPMP (0.13 mL, 0.68 mmol) was slowly added to the mixture. After the solution was stirred at 80 °C for 2 h, the formed nanoparticles were filtered, washed with MeOH and H_2O before dried in the freeze dryer overnight. In order to remove the surfactant from the nanopores of the nanoparticles, the as-synthesized nanoparticles (0.1 g) were suspended in MeOH (10 mL) and HCl (0.6 mL, 37.4%), and the solution was heated under reflux for 24 h. the nanoparticles were obtained via centrifugation and washed with ethanol for three times and water for three times. The product was dried in freeze dryer overnight to generate MSN-PO_4^-. For the FITC-labeled mesoporous silica nanoparticles, first, (3-Aminopropyl) triethoxysilane (APTES, 25 μL) was reacted with fluorescein isothiocyanate (FITC, 2 mg) in 0.5 mL ethanol overnight. CTAB (0.1 g) was first dissolved in distilled H_2O (48 mL) and NaOH (2 M, 0.35 mL), followed by heating the solution to 80 °C. TEOS (0.5 mL) was added to the solution. 15 min later, the above APTES-FITC solution (30 μL) and TPMP (0.13 mL, 0.68 mmol) were added while stirring was continued. After the stirring for 2 h, the nanoparticles were obtained by the filter. Finally, MSNs with CTAB template were refluxed for 6 h in a nitrogen atmosphere followed by extensive washes with methanol and deionized water and CTAB was removed by refluxing in acidic methanol solution to give FITC-labeled MSN-PO_4^-.

The above MSN-PO_4^- (1.0 mg/mL) was incubated in 1 mM of Dox stock solution in PBS buffer (10 mM, pH = 7.4) for 24 h, respectively. After that, the solutions were centrifuged (10000 rpm, 10 min) and washed with water twice

to remove the Dox molecules absorbed physically on the outer surface of the silica. The loading efficiency were determined by measuring the UV-vis spectrum of the supernatant and the stock solutions. Then the Dox-loaded MSNs (**MSN-Dox**) were redisposed in the PBS buffer before capped with **Biotin**CPD of different concentrations (10, 20, 50, 100, 150 μM). The mixtures were stirred for 2 h at 37 °C. After that, the resulting **CPD-MSN-Dox** were obtained by centrifugation and washed twice with PBS buffer. The nanoparticles were dried in the freeze dryer overnight. The nanoparticles were characterized via SEM, TEM, FTIR, DLS and Zeta potential. For following experiment, we use **CPD-MSN-Dox** that was capped with 50 μM **Biotin**CPD.

In a typical experiment, **CPD-MSN-Dox** was dispersed in PBS buffer at 37 °C (final Conc.: 0.1 mg/mL). Aliquots were taken from the suspension in a certain time and the delivery of Dox from the pore to the buffer solution was monitored via the UV absorbance at 480 nm.

HeLa cells were treated with 20 μg/mL of **CPD-MSN-Dox** (1.0 mg/mL in PBS) in DMEM for 3 h, 6h, 12 h or 24 h. Cells treated with **MSN-Dox** for 12 h and 24 h were carried out concurrently as controls. After incubation, HeLa cells were washed with PBS (containing 0.1 mg/mL heparin) three time before being imaged (Dox channel: λ_{ex} = 488 nm, λ_{em} = 580-650 nm; MSN channel: λ_{ex} = 488 nm, λ_{em} = 500-560 nm; Hoechst Channel: λ_{ex} = 405 nm, λ_{em} = 440-470 nm). The same experiment was carried out in 6-well plate. After 24 h incubation time, the pallets were collected for subsequent enzymatic assays (25 μg lysate, 1 μM Ac-DEVD-AMC) as described above. WB analysis of caspase-3 and cleaved PARP1 was also carried out for the same batch of HeLa lysates. XTT assay was carried out on HeLa cells treated with **CPD-MSN-Dox** (24 h

treatment) to evaluate the percentage of cell apoptosis caused by the intracellular release of Dox, as previously described.[150]

5.3.8 NMR
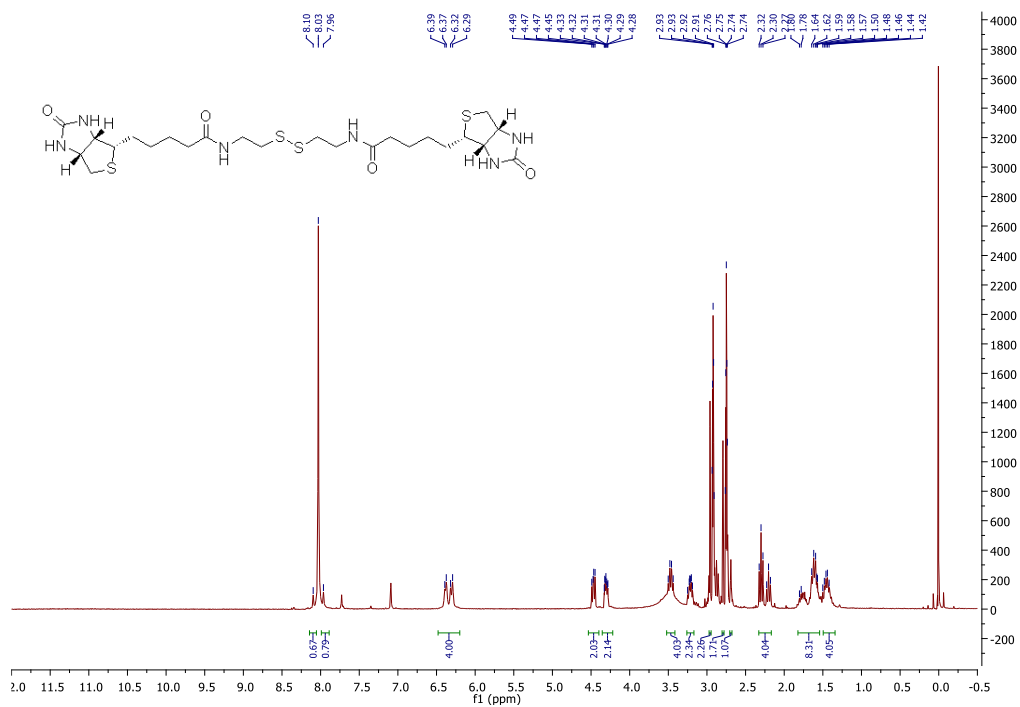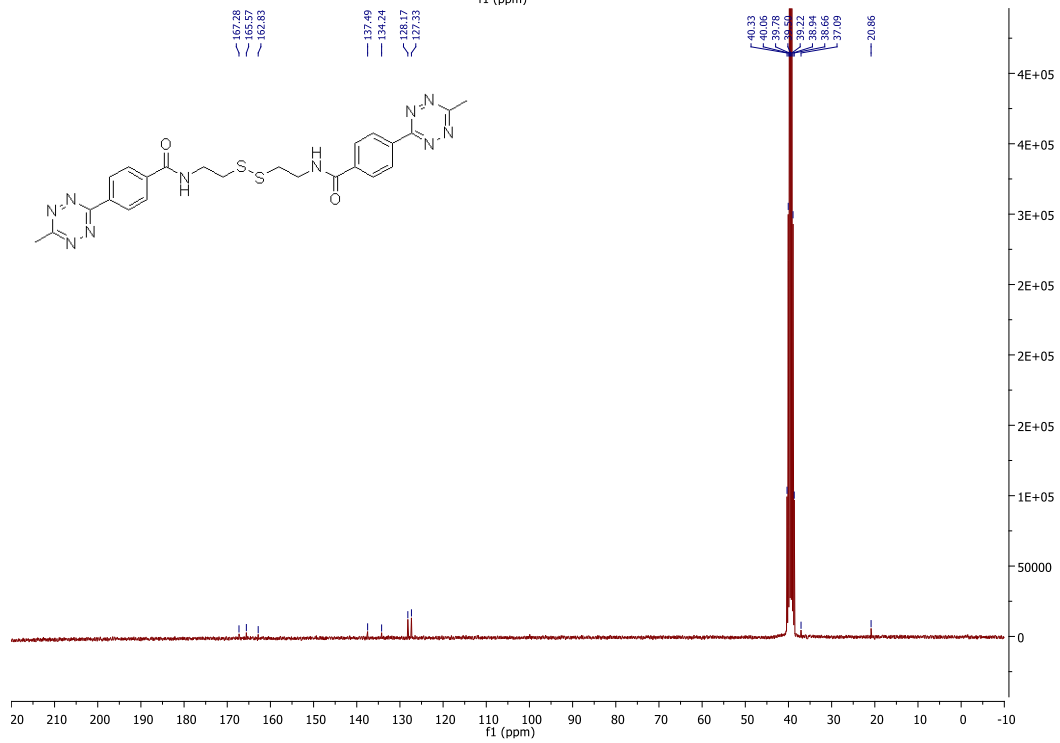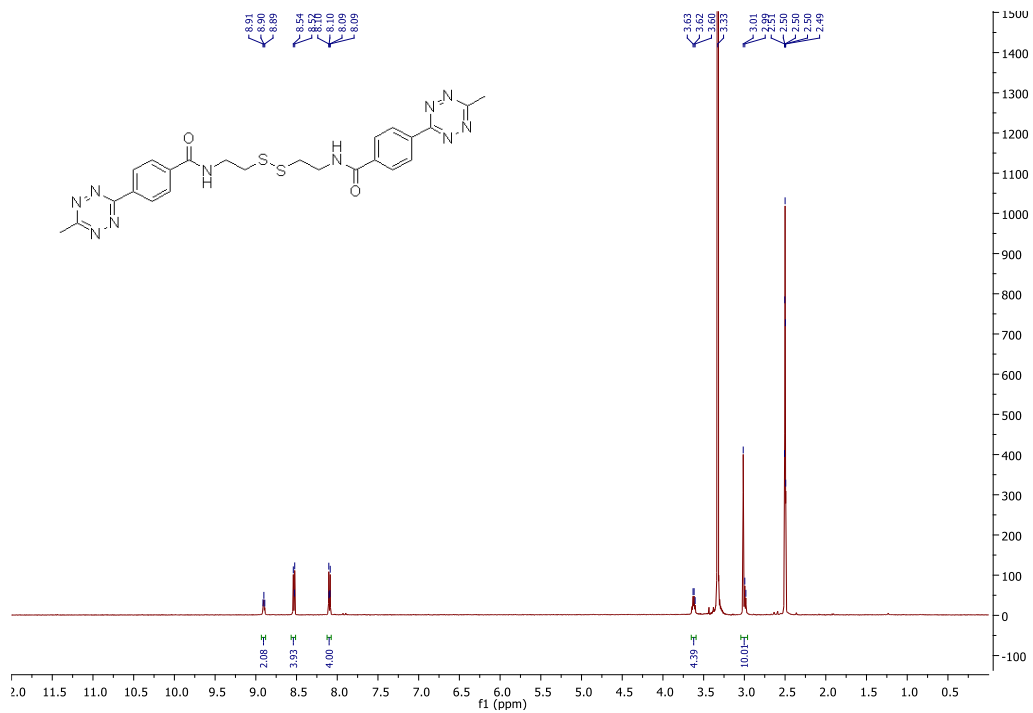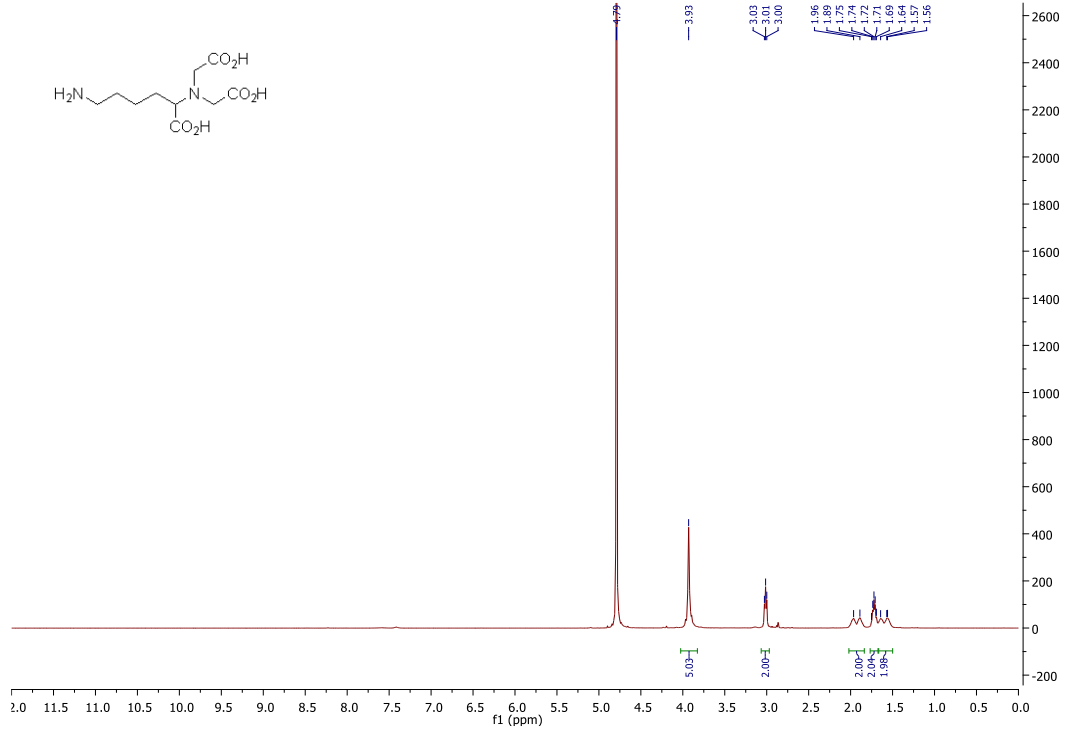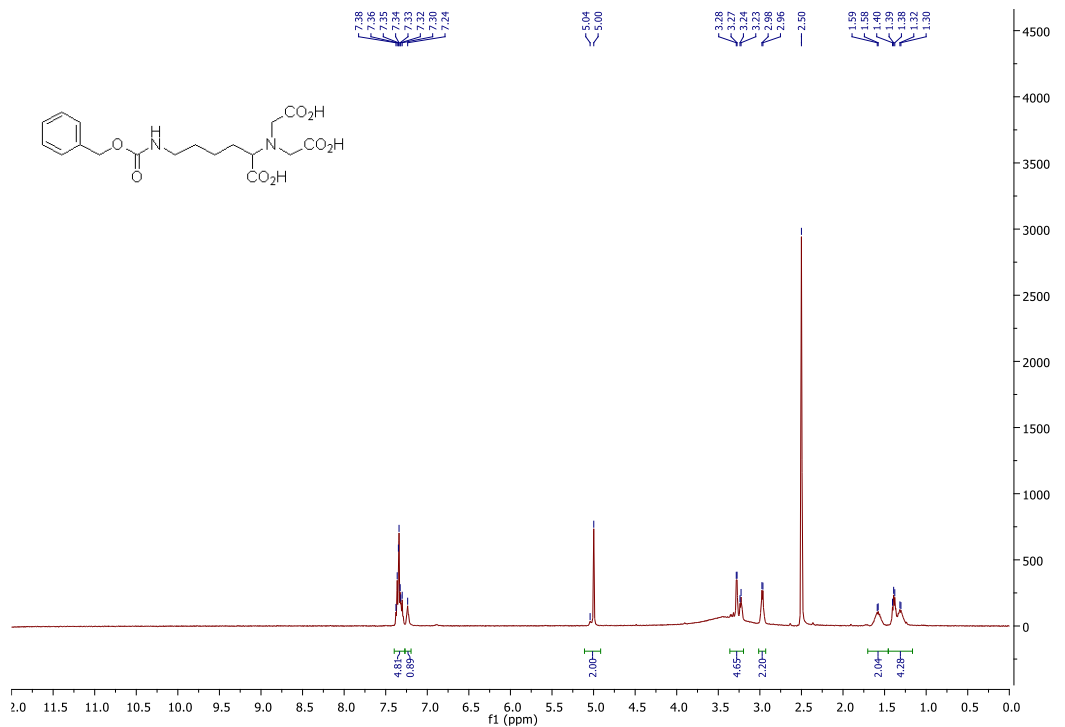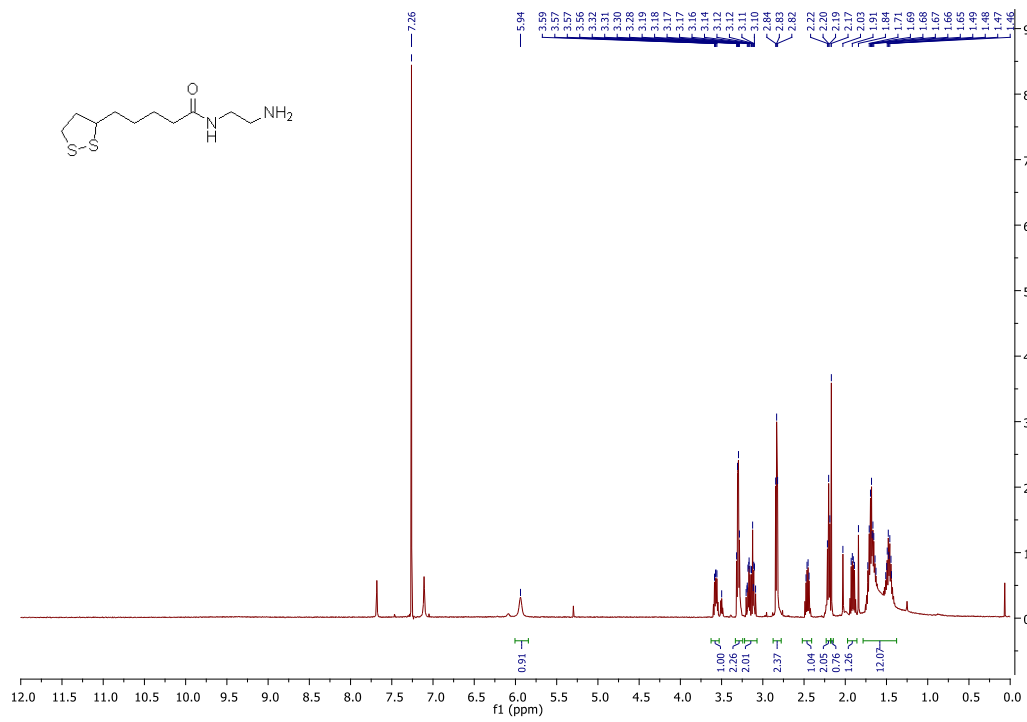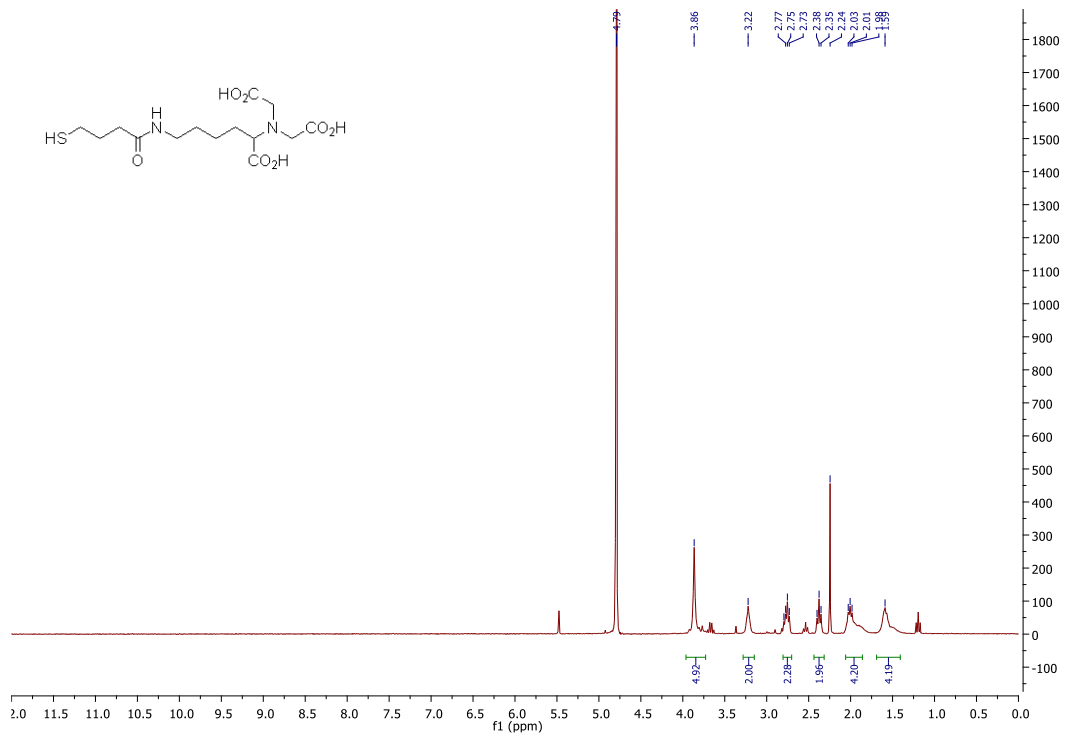
5.3.9 High-resolution Mass Spectrometry

Mass Spectrum SmartFormula Report

Analysis Info

Analysis Name	D:\Data\Chemistry\2015 Sample\Apr 2015\#1.d	Acquisition Date	4/6/2015 11:36:44 AM
Method	YCH-150-1800.m		
Sample Name	#1	Operator	default user
Comment	A/P Yao Shao Qin	Instrument / Ser#	micrOTOF-Q II 10269

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	2.0 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	50 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1800 m/z	Set Collision Cell RF	200.0 Vpp	Set Divert Valve	Waste

Meas. m/z	#	Formula	m/z	err [ppm]	rdb	e⁻ Conf	N-Rule
571.1413	1	C 24 H 24 N 10 Na O 2 S 2	571.1417	0.7	17.5	even	ok

Bruker Compass DataAnalysis 4.0   printed: 4/6/2015 11:51:34 AM   Page 1 of 1
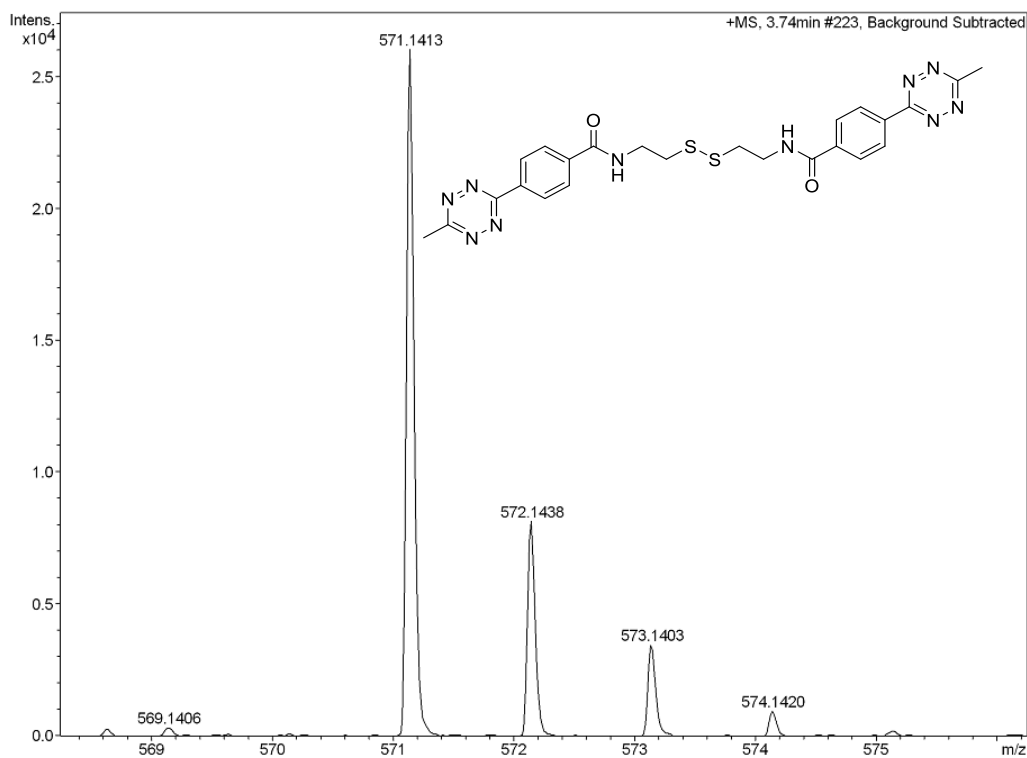
Chapter 6.

Conclusion remarks

The aim of the thesis include the development of novel HTP platforms for drug discovery and new strategies for highly efficient intracellular drug delivery.

Advances in genomic science and combinatorial chemistry provide the opportunities and possibilities for rapid drug development. Therefore, it is urgent to develop the high throughput screening strategies for the lead discovery and following up validation.

In chapter 2, we have successfully fabricated a 396-member peptide-hybrid small molecule microarray as a high-throughput platform for convenient profiling of SH2 domains. By screening the library against 15 SH2 domains, an informative fingerprint was generated which clustered the various SH2 domains according to their functional class. Further, ligands PP-pY-X$_5$, PP-pY-X$_{25}$, PP-pY-X$_1$ were identified as putative small molecule hits against Lck and Grb2. In future, these peptide-hybrid compounds might be converted into small molecule-based cell-permeable protein-protein inhibitors. Further, a dual colour strategy might be applied to monitor the slight differences in SH2 domains.

In chapter 3, the first microarray-based platform, enabling rapidly identifying on- and off-target of drug like compound from live cells was established. Compared to gel-based scans and PDs/WBs, our microarray provided the possibility of high-throughput validation with less sample consumption but comparable sensitivity. In this experiment, we have

successfully profiled a panel of drug-like ABP and A*f*BP bioactive compounds against six antibodies, and have likely identified PKA as the off-target for **P1** and **PU-1**. It should be noticed that the quality of antibodies used plays critical roles in the reliability of the readout. Further large-scale validation such as WBs/PDs were required for the potential off-target identified from microarray.

Overall, microarray-assisted strategies provide possibility for high-throughput screening approaches, for hit-identification, lead-optimization and off-target screening in drug discovery application.

Once lead compounds are identified, it requires further optimization in terms of solubility and cell-permeability. In spite of the effort on optimization, approximately 40% of currently available drugs and up to 75% of compounds under development have been estimated as poor water-soluble. The issue become more pronounced for macromolecules such as proteins and DNAs/siRNAs which emerge as another branch of therapeutic drugs. Therefore, the development of efficiently drug delivery strategy is urgently needed.

In chapter 4, we have successfully designed and synthesized several novel cell-penetrating poly(disulfide)s. These CPDs, upon conjugate to cargos including recombinant proteins and suitably modified antibodies via either covalent or non-covalent biorthogonal strategies, are able to rapidly and efficiently deliver those cargos into different mammalian cells via endocytosis-independent pathway. Rapid intracellular CPD depolymerization of the delivered cargos under highly reduced cytosolic environments subsequently released the protein in their functionally active form, which may be further translocated to their intended subcellular organelles for additional biological

process. The successful delivery of antibodies indicates this strategy may be more broadly applicable in further for effective intracellular delivery of other therapeutic antibodies. Besides, we have successfully developed CPD-capped MSNs for encapsulation of native small molecule drugs without the need of chemical modification. With doxorubicin as an example, we found **CPD-MSN-Dxo** entered mammalian cells rapidly and was able to subsequently release free Dox into cytosol.

Compared to existing strategies, the key advantages are their versatility and flexibility, enabling immediate delivery of a variety of cargos with minimum chemical modification. Further, the rapid and "biorthogonal" cargo-loading process enables the preparation of CPD-cargo conjugation in a matter of minutes under aqueous conditions, accelerating the cell delivery study. The minimal cell cytotoxicity further highlights the unique advantage of this new cell-transduction method over others.

Future work will focus on the expansion of the type of CPDs by using other conjugation chemistries, application of these CPDs for cell type specific delivery of other therapeutically important drugs and further animal-based therapeutic drug delivery.

Chapter 7.

Reference

(1)Triggle, D. J. *Med Princ Pract* **2007**, *16*, 1-14.

(2)Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. *Br J Pharmacol* **2011**, *162*, 1239-1249.

(3)*Science*, *287*.

(4)Foong, Y. M.; Fu, J.; Yao, S. Q.; Uttamchandani, M. *Curr Opin Chem Biol* **2012**, *16*, 234-242.

(5)Sun, H.; Chen, G. Y.; Yao, S. Q. *Chem Biol* **2013**, *20*, 685-699.

(6)Uttamchandani, M.; Lu, C. H.; Yao, S. Q. *Acc Chem Res* **2009**, *42*, 1183-1192.

(7)Sun, H.; Lu, C. H.; Uttamchandani, M.; Xia, Y.; Liou, Y. C.; Yao, S. Q. *Angew Chem Int Ed Engl* **2008**, *47*, 1698-1702.

(8)Wu, H.; Ge, J. Y.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2011**, *47*, 5664-5670.

(9)Williams, H. D.; Trevaskis, N. L.; Charman, S. A.; Shanker, R. M.; Charman, W. N.; Pouton, C. W.; Porter, C. J. H. *Pharmacological Reviews* **2013**, *65*, 315-499.

(10)Juliano, R. L.; Astriab-Fisher, A.; Falke, D. *Mol Interv* **2001**, *1*, 40-53.

(11)Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu Rev Biochem* **2008**, *77*, 383-414.

(12)Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science* **1995**, *270*, 467-470.

(13)MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289*, 1760-1763.

(14) MacBeath, G.; Koehler, A. N.; Schreiber, S. L. *Journal of the American Chemical Society* **1999**, *121*, 7967-7968.

(15) Moll, J. Target Identification and Validation in Drug Discovery: Methods and Protocols, Springer: New York, **2013**, pp187-213.

(16) Hong, J. A.; Neel, D. V.; Wassaf, D.; Caballero, F.; Koehler, A. N. *Curr Opin Chem Biol* **2014**, *18*, 21-28.

(17) Kwon, S. J.; Lee, M. Y.; Ku, B.; Sherman, D. H.; Dordick, J. S. *Acs Chemical Biology* **2007**, *2*, 419-425.

(18) Mok, J.; Im, H.; Snyder, M. *Nature Protocols* **2009**, *4*, 1820-1827.

(19) Han, X. M.; Sonoda, T.; Mori, T.; Yamanouchi, G.; Yamaji, T.; Shigaki, S.; Niidome, T.; Katayama, Y. *Combinatorial Chemistry & High Throughput Screening* **2010**, *13*, 777-789.

(20) Urbina, H. D.; Debaene, F.; Jost, B.; Bole-Feysot, C.; Mason, D. E.; Kuzmic, P.; Harris, J. L.; Winssinger, N. *Chembiochem* **2006**, *7*, 1790-1797.

(21) Harris, J.; Mason, D. E.; Li, J.; Burdick, K. W.; Backes, B. J.; Chen, T.; Shipway, A.; Van Heeke, G.; Gough, L.; Ghaemmaghami, A.; Shakib, F.; Debaene, F.; Winssinger, N. *Chemistry & Biology* **2004**, *11*, 1361-1372.

(22) Sun, H. Y.; Tan, L. P.; Gao, L. Q.; Yao, S. Q. *Chemical Communications* **2009**, 677-679.

(23) Kemp, M. M.; Weiwer, M.; Koehler, A. N. *Bioorganic & Medicinal Chemistry* **2012**, *20*, 1979-1989.

(24) Vegas, A. J.; Bradner, J. E.; Tang, W. P.; McPherson, O. M.; Greenberg, E. F.; Koehler, A. N.; Schreiber, S. L. *Angewandte Chemie-International Edition* **2007**, *46*, 7960-7964.

(25) Gurard-Levin, Z. A.; Kilian, K. A.; Kim, J.; Bahr, K.; Mrksich, M. *Acs Chemical Biology* **2010**, *5*, 863-873.

(26) Na, Z.; Pan, S.; Uttamchandani, M.; Yao, S. Q. *Angew Chem Int Ed Engl* **2014**, *53*, 8421-8426.

(27) Lee, H. Y.; Park, S. B. *Molecular Biosystems* **2011**, *7*, 304-310.

(28) Zinzalla, G.; Thurston, D. E. *Future Medicinal Chemistry* **2009**, *1*, 65-93.

(29) Mullard, A. *Nat Rev Drug Discov* **2012**, *11*, 173-175.

(30) Na, Z.; Peng, B.; Ng, S.; Pan, S.; Lee, J. S.; Shen, H. M.; Yao, S. Q. *Angew Chem Int Ed Engl* **2015**, *54*, 2515-2519.

(31) Wu, H.; Ge, J.; Yao, S. Q. *Angew Chem Int Ed Engl* **2010**, *49*, 6528-6532.

(32) Zhang, C. J.; Tan, C. Y.; Ge, J.; Na, Z.; Chen, G. Y.; Uttamchandani, M.; Sun, H.; Yao, S. Q. *Angew Chem Int Ed Engl* **2013**, *52*, 14060-14064.

(33) Miyazaki, I.; Simizu, S.; Okumura, H.; Takagi, S.; Osada, H. *Nature Chemical Biology* **2010**, *6*, 667-673.

(34) Na, Z. K.; Li, L.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2012**, *48*, 7304-7306.

(35) Wu, H.; Ge, J. Y.; Yang, P. Y.; Wang, J. G.; Uttamchandani, M.; Yao, S. Q. *Journal of the American Chemical Society* **2011**, *133*, 1946-1954.

(36) Shi, H. B.; Uttamchandani, M.; Yao, S. Q. *Chemistry-an Asian Journal* **2011**, *6*, 2803-2815.

(37) Lee, H. Y.; Park, S. B. *Mol Biosyst* **2011**, *7*, 304-310.

(38) Liu, Y. S.; Patricelli, M. P.; Cravatt, B. F. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 14694-14699.

(39) Kidd, D.; Liu, Y.; Cravatt, B. F. *Biochemistry* **2001**, *40*, 4005-4015.

(40) Chen, G. Y.; Uttamchandani, M.; Zhu, Q.; Wang, G.; Yao, S. Q. *ChemBioChem*, **2003**, *4*, 336-339.

(41) Uttamchandani, M.; Liu, K.; Panicker, R. C.; Yao, S. Q. *Chem Commun (Camb)* **2007**, 1518-1520.

(42) Srinivasan, R.; Huang, X.; Ng, S. L.; Yao, S. Q. *Chembiochem* **2006**, *7*, 32-36.

(43) Funeriu, D. P.; Eppinger, J.; Denizot, L.; Miyake, M.; Miyake, J. *Nature Biotechnology* **2005**, *23*, 622-627.

(44) Eppinger, J.; Funeriu, D. P.; Miyake, M.; Denizot, L.; Miyake, J. *Angewandte Chemie-International Edition* **2004**, *43*, 4389-4389.

(45) Jung, S. H.; Kong, D. H.; Park, S. W.; Kim, Y. M.; Ha, K. S. *Analyst* **2012**, *137*, 3814-3820.

(46) Sieber, S. A.; Mondala, T. S.; Head, S. R.; Cravatt, B. F. *Journal of the American Chemical Society* **2004**, *126*, 15640-15641.

(47) Sikkema, A. H.; Diks, S. H.; den Dunnen, W. F. A.; ter Elst, A.; Scherpen, F. J. G.; Hoving, E. W.; Ruijtenbeek, R.; Boender, P. J.; de Wijn, R.; Kamps, W. A.; Peppelenbosch, M. P.; de Bont, E. S. J. M. *Cancer Research* **2009**, *69*, 5987-5995.

(48) Sokolov, B. P.; Cadet, J. L. *Neuropsychopharmacology* **2006**, *31*, 956-966.

(49) Boyd, Z. S.; Wu, Q. J.; O'Brien, C.; Spoerke, J.; Savage, H.; Fielder, P. J.; Arnler, L.; Yan, Y. B.; Lackner, M. R. *Molecular Cancer Therapeutics* **2008**, *7*, 3695-3706.

(50) Reddy, M. M.; Wilson, R.; Wilson, J.; Connell, S.; Gocke, A.; Hynan, L.; German, D.; Kodadek, T. *Cell* **2011**, *144*, 132-142.

(51) Di, L.; Fish, P. V.; Mano, T. *Drug Discovery Today* **2012**, *17*, 486-495.

(52) Di, L.; Kerns, E. H.; Carter, G. T. *Current Pharmaceutical Design* **2009**, *15*, 2184-2194.

(53) Heitz, F., Morris, M. C., Divita, G. *Br J Pharmacol* **2009**, *157,* 195-206.

(54) Cronican, J. J., Thompson, D. B., Beier, K.T., McNaughton, B. R., Cepko, C. L., Liu, D. R. *ACS Chem Biol* **2010**, *5,* 747-752.

(55) Allen, T. M.; Cullis, P. R. *Adv Drug Deliv Rev* **2013**, *65*, 36-48.

(56) Chou, L. Y.; Ming, K.; Chan, W. C. *Chem Soc Rev* **2011**, *40*, 233-245.

(57) Zhao, M.; Liu, Y.; Hsieh, R. S.; Wang, N.; Tai, W.; Joo, K. I.; Wang, P.; Gu, Z.; Tang, Y. *J Am Chem Soc* **2014**, *136*, 15319-15325.

(58) Fu, A.; Tang, R.; Hardie, J.; Farkas, M. E.; Rotello, V. M. *Bioconjug Chem* **2014**, *25*, 1602-1608.

(59) Bechara, C.; Sagan, S. *FEBS Lett* **2013**, *587*, 1693-1702.

(60) Frankel, A. D.; Pabo, C. O. *Cell* **1988**, *55*, 1189-1193.

(61) Joliot, A.; Pernelle, C.; Deagostinibazin, H.; Prochiantz, A. *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 1864-1868.

(62) Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F. *Science* **1999**, *285*, 1569-1572.

(63) Fischer, P. M.; Zhelev, N. Z.; Wang, S.; Melville, J. E.; Fahraeus, R.; Lane, D. P. *Journal of Peptide Research* **2000**, *55*, 163-172.

(64) Mishra, A.; Lai, G. H.; Schmidt, N. W.; Sun, V. Z.; Rodriguez, A. R.; Tong, R.; Tang, L.; Cheng, J. J.; Deming, T. J.; Kamei, D. T.; Wong, G. C. L. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 16883-16888.

(65) Raagel, H.; Saalik, P.; Pooga, M. *Biochimica Et Biophysica Acta-Biomembranes* **2010**, *1798*, 2240-2248.

(66) Jiao, C. Y.; Delaroche, D.; Burlina, F.; Alves, I. D.; Chassaing, G.; Sagan, S. *Journal of Biological Chemistry* **2009**, *284*, 33957-33965.

(67) Belting, M. *Trends in Biochemical Sciences* **2003**, *28*, 145-151.

(68) Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R. *Traffic* **2007**, *8*, 848-866.

(69) Erazo-Oliveras, A.; Muthukrishnan, N.; Baker, R.; Wang, T. Y.; Pellois, J. P. *Pharmaceuticals (Basel)* **2012**, *5*, 1177-1209.

(70) Qian, Z. Q.; LaRochelle, J. R.; Jiang, B. S.; Lian, W. L.; Hard, R. L.; Selner, N. G.; Luechapanichkul, R.; Barrios, A. M.; Pei, D. H. *Biochemistry* **2014**, *53*, 4034-4046.

(71) Sun, T.; Zhang, Y. S.; Pang, B.; Hyun, D. C.; Yang, M.; Xia, Y. *Angew Chem Int Ed Engl* **2014**, *53*, 12320-12364.

(72) Al-Ahmady, Z. S.; Al-Jamal, W. T.; Bossche, J. V.; Bui, T. T.; Drake, A. F.; Mason, A. J.; Kostarelos, K. *Acs Nano* **2012**, *6*, 9335-9346.

(73) Chen, K. J.; Liang, H. F.; Chen, H. L.; Wang, Y. C.; Cheng, P. Y.; Liu, H. L.; Xia, Y. N.; Sung, H. W. *Acs Nano* **2013**, *7*, 438-446.

(74) Koren, E.; Apte, A.; Jani, A.; Torchilin, V. P. *Journal of Controlled Release* **2012**, *160*, 264-273.

(75) Thamphiwatana, S.; Fu, V.; Zhu, J. Y.; Lu, D. N.; Gao, W. W.; Zhang, L. F. *Langmuir* **2013**, *29*, 12228-12233.

(76) Kelly, C.; Jefferies, C.; Cryan, S. A. *J Drug Deliv* **2011**, *2011*, 727241.

(77) Mo, R.; Sun, Q.; Xue, J.; Li, N.; Li, W.; Zhang, C.; Ping, Q. *Adv Mater* **2012**, *24*, 3659-3665.

(78) Debs, R. J.; Freedman, L. P.; Edmunds, S.; Gaensler, K. L.; Duzgunes, N.; Yamamoto, K. R. *J Biol Chem* **1990**, *265*, 10189-10192.

(79) Zelphati, O.; Wang, Y.; Kitada, S.; Reed, J. C.; Felgner, P. L.; Corbeil, J. *J Biol Chem* **2001**, *276*, 35103-35110.

(80) Akbarzadeh, A.; Rezaei-Sadabady, R.; Davaran, S.; Joo, S. W.; Zarghami, N.; Hanifehpour, Y.; Samiei, M.; Kouhi, M.; Nejati-Koshki, K. *Nanoscale Research Letters* **2013**, *8*.

(81) Tabujew, I.; Peneva, K. **2014**, 1-29.

(82) Kim, S. H.; Jeong, J. H.; Chun, K. W.; Park, T. G. *Langmuir* **2005**, *21*, 8852-8857.

(83) Kim, M. S.; Gao, G. H.; Kang, S. W.; Lee, D. S. *Macromolecular Bioscience* **2011**, *11*, 946-951.

(84) Cohen, S.; Coue, G.; Beno, D.; Korenstein, R.; Engbersen, J. F. J. *Biomaterials* **2012**, *33*, 614-623.

(85) Mura, S.; Nicolas, J.; Couvreur, P. *Nat Mater* **2013**, *12*, 991-1003.

(86) Islam, M. R.; Gao, Y.; Li, X.; Serpe, M. J. *Journal of Materials Chemistry B* **2014**, *2*, 2444.

(87) Tian, L.; Kang, H. C.; Bae, Y. H. *Biomacromolecules* **2013**, *14*, 2570-2581.

(88) Zhang, X.; Zhao, J.; Wen, Y.; Zhu, C.; Yang, J.; Yao, F. *Carbohydr Polym* **2013**, *98*, 1326-1334.

(89) Guo, X.; Shi, C.; Yang, G.; Wang, J.; Cai, Z.; Zhou, S. *Chemistry of Materials* **2014**, *26*, 4405-4418.

(90) Liu, B. A.; Jablonowski, K.; Raina, M.; Arce, M.; Pawson, T.; Nash, P. D. *Molecular Cell* **2006**, *22*, 851-868.

(91) Machida, K.; Mayer, B. J. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2005**, *1747*, 1-25.

(92) Yaffe, M. B. *Nature Reviews Molecular Cell Biology* **2002**, *3*, 177-186.

(93) Kraskouskaya, D.; Duodu, E.; Arpin, C. C.; Gunning, P. T. *Chemical Society Reviews* **2013**, *42*, 3337-3370.

(94) Liu, B. A.; Engelmann, B. W.; Nash, P. D. *Proteomics* **2012**, *12*, 1527-1546.

(95) Miller, M. L.; Hanke, S.; Hinsby, A. M.; Friis, C.; Brunak, S.; Mann, M.; Blom, N. *Molecular & Cellular Proteomics* **2008**, *7*, 181-192.

(96) Songyang, Z.; Shoelson, S. E.; McGlade, J.; Olivier, P.; Pawson, T.; Bustelo, X. R.; Barbacid, M.; Sabe, H.; Hanafusa, H.; Yi, T.; Ren, R.; Baltimore, D.; Ratnofsky, S.; Feldman, R. A.; Cantley, L. C. *Molecular and Cellular Biology* **1994**, *14*, 2777-2785.

(97) Eck, M. J.; Shoelson, S. E.; Harrison, S. C. *Nature* **1993**, *362*, 87-91.

(98) Wu, H.; Ge, J.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2011**, *47*, 5664-5670.

(99) Huang, H.; Li, L.; Wu, C.; Schibli, D.; Colwill, K.; Ma, S.; Li, C.; Roy, P.; Ho, K.; Zhou, S.; Pawson, T.; Gao, Y.; Li, S. S. C. *Molecular & Cellular Proteomics* **2008**, *7*, 768-784.

(100) Jones, R. B.; Gordus, A.; Krall, J. A.; MacBeath, G. *Nature* **2006**, *439*, 168-174.

(101) Gordus, A.; MacBeath, G. *Journal of the American Chemical Society* **2006**, *128*, 13668-13669.

(102) Fu, J. Q.; Na, Z. K.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2013**, *49*, 9660-9662.

(103) Zhou, S. Y.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; Neel, B. G.; Birge, R. B.; Fajardo, J. E.; Chou, M. M.; Hanafusa, H.; Schaffhausen, B.; Cantley, L. C. *Cell* **1993**, *72*, 767-778.

(104) Uttamchandani, M.; Wang, J.; Li, J.; Hu, M.; Sun, H.; Chen, K. Y. T.; Liu, K.; Yao, S. Q. *Journal of the American Chemical Society* **2007**, *129*, 7848-7858.

(105) Ambaye, N. D.; Gunzburg, M. J.; Lim, R. C. C.; Price, J. T.; Wilce, M. C. J.; Wilce, J. A. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 693-701.

(106) Wu, H.; Ge, J.; Yao, S. Q. *Angewandte Chemie-International Edition* **2010**, *49*, 6528-6532.

(107) Lu, C. H. S.; Sun, H.; Abu Bakar, F. B.; Uttamchandani, M.; Zhou, W.; Liou, Y.-C.; Yao, S. Q. *Angewandte Chemie-International Edition* **2008**, *47*, 7438-7441.

(108) Gao, L.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2012**, *48*, 2240-2242.

(109) Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. *Angewandte Chemie-International Edition* **2013**, *52*, 2744-2792.

(110) Dar, A. C.; Das, T. K.; Shokat, K. M.; Cagan, R. L. *Nature* **2012**, *486*, 80-U101.

(111) Simon, G. M.; Niphakis, M. J.; Cravatt, B. F. *Nature Chemical Biology* **2013**, *9*, 200-205.

(112) Su, Y.; Ge, J.; Zhu, B.; Zheng, Y. G.; Zhu, Q.; Yao, S. Q. *Curr Opin Chem Biol* **2013**, *17*, 768-775.

(113) Eder, J.; Sedrani, R.; Wiesmann, C. *Nat Rev Drug Discov* **2014**, *13*, 577-587.

(114) Harrison, C. *Nature Reviews Drug Discovery* **2014**, *13*.

(115) Yang, P. Y.; Liu, K.; Ngai, M. H.; Lear, M. J.; Wenk, M. R.; Yao, S. Q. *Journal of the American Chemical Society* **2010**, *132*, 656-666.

(116) Shi, H. B.; Zhang, C. J.; Chen, G. Y. J.; Yao, S. Q. *Journal of the American Chemical Society* **2012**, *134*, 3001-3014.

(117) Li, Z. Q.; Hao, P. L.; Li, L.; Tan, C. Y. J.; Cheng, X. M.; Chen, G. Y. J.; Sze, S. K.; Shen, H. M.; Yao, S. Q. *Angewandte Chemie-International Edition* **2013**, *52*, 8551-8556.

(118) Sevecka, M.; MacBeath, G. *Nature Methods* **2006**, *3*, 825-831.

(119) Ge, J. Y.; Zhang, C. J.; Li, L.; Chong, L. M.; Wu, X. Y.; Hao, P. L.; Sze, S. K.; Yao, S. Q. *Acs Chemical Biology* **2013**, *8*, 2577-2585.

(120) Li, Z. Q.; Wang, D. Y.; Li, L.; Pan, S. J.; Na, Z. K.; Tan, C. Y. J.; Yao, S. Q. *Journal of the American Chemical Society* **2014**, *136*, 9990-9998.

(121) Sletten, E. M.; Bertozzi, C. R. *Angewandte Chemie-International Edition* **2009**, *48*, 6974-6998.

(122) Kalesh, K. A.; Shi, H. B.; Ge, J. Y.; Yao, S. Q. *Organic & Biomolecular Chemistry* **2010**, *8*, 1749-1762.

(123) Fabian, M. A.; Biggs, W. H.; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. *Nature Biotechnology* **2005**, *23*, 329-336.

(124) Torres, A. G.; Gait, M. J. *Trends Biotechnol.* **2012**, *30*, 185-190.

(125) Lindsay, M. A. *Curr. Opin. Pharmacol.* **2002**, *2*, 587-594.

(126) Kostarelos, K.; Miller, A. D. *Chem. Soc. Rev.* **2005**, *34*, 970-994.

(127) Mintzer, M. A.; Simanek, E. E. *Chem. Rev. (Washington, DC, U. S.)* **2009**, *109*, 259-302.

(128) van den Berg, A.; Dowdy, S. F. *Curr. Opin. Biotechnol.* **2011**, *22*, 888-893.

(129) Nguyen, J.; Szoka, F. C. *Acc. Chem. Res.* **2012**, *45*, 1153-1162.

(130) Fu, A.; Tang, R.; Hardie, J.; Farkas, M. E.; Rotello, V. M. *Bioconjugate Chem.* **2014**, *25*, 1602-1608.

(131) Jorgensen, W. L.; Duffy, E. M. *Adv. Drug Delivery Rev.* **2002**, *54*, 355-366.

(132) Williams, H. D.; Trevaskis, N. L.; Charman, S. A.; Shanker, R. M.; Charman, W. N.; Pouton, C. W.; Porter, C. J. H. *Pharmacol. Rev.* **2013**, *65*, 315-499, 185 pp.

(133) Sun, T.; Zhang, Y. S.; Pang, B.; Hyun, D. C.; Yang, M.; Xia, Y. *Angew. Chem., Int. Ed.* **2014**, *53*, 12320-12364.

(134) Heitz, F.; Morris, M. C.; Divita, G. *Br. J. Pharmacol.* **2009**, *157*, 195-206.

(135) Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R. *ACS Chem. Biol.* **2010**, *5*, 747-752.

(136) Allen, T. M.; Cullis, P. R. *Adv. Drug Delivery Rev.* **2013**, *65*, 36-48.

(137) Montrose, K.; Yang, Y.; Sun, X.; Wiles, S.; Krissansen, G. W. *Sci. Rep.* **2013**, *3*, 1661, 1667 pp.

(138) Han, D. H.; Na, H.-K.; Choi, W. H.; Lee, J. H.; Kim, Y. K.; Won, C.; Lee, S.-H.; Kim, K. P.; Kuret, J.; Min, D.-H.; Lee, M. J. *Nat. Commun.* **2014**, *5*, 5633.

(139) Sarker, S. R.; Hokama, R.; Takeoka, S. *Mol. Pharmaceutics* **2014**, *11*, 164-174.

(140) Nischan, N.; Herce, H. D.; Natale, F.; Bohlke, N.; Budisa, N.; Cardoso, M. C.; Hackenberger, C. P. R. *Angew. Chem., Int. Ed.* **2015**, *54*, 1950-1953.

(141) Zuris, J. A.; Thompson, D. B.; Shu, Y.; Guilinger, J. P.; Bessen, J. L.; Hu, J. H.; Maeder, M. L.; Joung, J. K.; Chen, Z.-Y.; Liu, D. R. *Nat. Biotechnol.* **2015**, *33*, 73-80.

(142) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. *J. Biol. Chem.* **2003**, *278*, 585-590.

(143) Tuinnemann, G.; Martin, R. M.; Haupt, S.; Patsch, C.; Edenhofer, F.; Cardoso, M. C. *Faseb J.* **2006**, *20*, 1775-1784.

(144) Michaudel, Q.; Journot, G.; Regueiro-Ren, A.; Goswami, A.; Guo, Z.; Tully, T. P.; Zou, L.; Ramabhadran, R. O.; Houk, K. N.; Baran, P. S. *Angew. Chem., Int. Ed.* **2014**, *53*, 12091-12096.

(145) Li, Z.; Barnes, J. C.; Bosoy, A.; Stoddart, J. F.; Zink, J. I. *Chem. Soc. Rev.* **2012**, *41*, 2590-2605.

(146) Coll, C.; Bernardos, A.; Martinez-Manez, R.; Sancenon, F. *Acc. Chem. Res.* **2013**, *46*, 339-349.

(147) Nicolas, J.; Mura, S.; Brambilla, D.; Mackiewicz, N.; Couvreur, P. *Chem. Soc. Rev.* **2013**, *42*, 1147-1235.

(148) Slowing, I.; Trewyn, B. G.; Lin, V. S. Y. *J. Am. Chem. Soc.* **2006**, *128*, 14792-14793.

(149) Zhang, P.; Cheng, F.; Zhou, R.; Cao, J.; Li, J.; Burda, C.; Min, Q.; Zhu, J.-J. *Angew. Chem., Int. Ed.* **2014**, *53*, 2371-2375.

(150) Yu, C.; Qian, L.; Uttamchandani, M.; Li, L.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2015**, *54*, 10574-10578.

(151) Bang, E.-K.; Gasparini, G.; Molinard, G.; Roux, A.; Sakai, N.; Matile, S. *J. Am. Chem. Soc.* **2013**, *135*, 2088-2091.

(152) Gasparini, G.; Bang, E.-K.; Molinard, G.; Tulumello, D. V.; Ward, S.; Kelley, S. O.; Roux, A.; Sakai, N.; Matile, S. *J. Am. Chem. Soc.* **2014**, *136*, 6069-6074.

(153) Chuard, N.; Gasparini, G.; Roux, A.; Sakai, N.; Matile, S. *Org. Biomol. Chem.* **2015**, *13*, 64-67.

(154) Gasparini, G.; Bang, E.-K.; Montenegro, J.; Matile, S. *Chem. Commun. (Cambridge, U. K.)* **2015**, *51*, 10389-10402.

(155) Son, S.; Namgung, R.; Kim, J.; Singha, K.; Kim, W. J. *Acc. Chem. Res.* **2012**, *45*, 1100-1112.

(156) Gasparini, G.; Sargsyan, G.; Bang, E.-K.; Sakai, N.; Matile, S. *Angew. Chem., Int. Ed.* **2015**, *54*, 7328-7331.

(157) Fu, J.; Yu, C.; Li, L.; Yao, S. Q. *J Am Chem Soc* **2015**, *137*, 12153-12160.

(158) Marschall, A. L. J.; Frenzel, A.; Schirrmann, T.; Schüngel, M.; Dubel, S. *mAbs* **2014**, *3*, 3-16.

(159) Khalili, H.; Godwin, A.; Choi, J.-w.; Lever, R.; Brocchini, S. *Bioconjugate Chem.* **2012**, *23*, 2262-2277.

(160) Devaraj, N. K.; Weissleder, R. *Acc. Chem. Res.* **2011**, *44*, 816-827.

(161) Li, Z.; Wang, D.; Li, L.; Pan, S.; Na, Z.; Tan, C. Y. J.; Yao, S. Q. *J. Am. Chem. Soc.* **2014**, *136*, 9990-9998.

(162) Katz, B. A. *J. Mol. Biol.* **1997**, *274*, 776-800.

(163) Fuchs, S. M.; Raines, R. T. *ACS Chem. Biol.* **2007**, *2*, 167-170.

(164) McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 6111-6116.

(165) Denault, J.-B.; Salvesen, G. S. *Chem. Rev. (Washington, DC, U. S.)* **2002**, *102*, 4489-4499.

(166) Yang, X.-H.; Sladek, T. L.; Liu, X.; Butler, B. R.; Froelich, C. J.; Thor, A. D. *Cancer Res.* **2001**, *61*, 348-354.

(167) Ng, S. L.; Yang, P.-Y.; Chen, K. Y. T.; Srinivasan, R.; Yao, S. Q. *Org. Biomol. Chem.* **2008**, *6*, 844-847.

(168) Li, J.; Yao, S. Q. *Org. Lett.* **2009**, *11*, 405-408.

(169) Hu, M.; Li, L.; Wu, H.; Su, Y.; Yang, P.-Y.; Uttamchandani, M.; Xu, Q.-H.; Yao, S. Q. *J. Am. Chem. Soc.* **2011**, *133*, 12009-12020.

(170) Na, Z.; Peng, B.; Ng, S.; Pan, S.; Lee, J.-S.; Shen, H.-M.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2015**, *54*, 2515-2519.

(171) Chaitanya, G. V.; Steve, A. J.; Babu, P. P. *Cell Commun. Signaling* **2010**, *8*, 31.

(172) Weiner, L. M.; Surana, R.; Wang, S. *Nat. Rev. Immunol.* **2010**, *10*, 317-327.

(173) Shaunak, S.; Godwin, A.; Choi, J.-W.; Balan, S.; Pedone, E.; Vijayarangam, D.; Heidelberger, S.; Teo, I.; Zloh, M.; Brocchini, S. *Nat. Chem. Biol.* **2006**, *2*, 312-313.

(174) Lyon, R. P.; Meyer, D. L.; Setter, J. R.; Senter, P. D. *Methods Enzymol.* **2012**, *502*, 123-138.

(175) Badescu, G.; Bryant, P.; Bird, M.; Henseleit, K.; Swierkosz, J.; Parekh, V.; Tommasi, R.; Pawlisz, E.; Jurlewicz, K.; Farys, M.; Camper, N.; Sheng, X.; Fisher, M.; Grygorash, R.; Kyle, A.; Abhilash, A.; Frigerio, M.; Edwards, J.; Godwin, A. *Bioconjugate Chem.* **2014**, *25*, 1124-1136.

(176) Zhao, Y.-L.; Li, Z.; Kabehie, S.; Botros, Y. Y.; Stoddart, J. F.; Zink, J. I. *J. Am. Chem. Soc.* **2010**, *132*, 13016-13025.

(177) Tacar, O.; Sriamornsak, P.; Dass, C. R. *J. Pharm. Pharmacol.* **2013**, *65*, 157-170.

(178) Lea, W. A.; Simeonov, A. *Plos One* **2012**, *7*.

(179) Liu, K.; Kalesh, K. A.; Ong, L. B.; Yao, S. Q. *Chembiochem* **2008**, *9*, 1883-1888.

(180) Cheng, X. M.; Li, L.; Uttamchandani, M.; Yao, S. Q. *Chemical Communications* **2014**, *50*, 2851-2853.

(181) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A. M.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S. H.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. *Journal of Medicinal Chemistry* **2004**, *47*, 6658-6661.

(182) Kalesh, K. A.; Sim, D. S. B.; Wang, J. G.; Liu, K.; Lin, Q. S.; Yao, S. Q. *Chemical Communications* **2010**, *46*, 1118-1120.

(183) Barrett, S. D.; Bridges, A. J.; Dudley, D. T.; Saltiel, A. R.; Fergus, J. H.; Flamme, C. M.; Delaney, A. M.; Kaufman, M.; LePage, S.; Leopold, W. R.;

Przybranowski, S. A.; Sebolt-Leopold, J.; Van Becelaere, K.; Doherty, A. M.; Kennedy, R. M.; Marston, D.; Howard, W. A.; Smith, Y.; Warmus, J. S.; Tecle, H. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 6501-6504.

(184) Shi, H. B.; Cheng, X. M.; Sze, S. K.; Yao, S. Q. *Chemical Communications* **2011**, *47*, 11306-11308.

(185) Machrouhi, F.; Ouhamou, N.; Laderoute, K.; Calaoagan, J.; Bukhtiyarova, M.; Ehrlich, P. J.; Klon, A. E. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 6394-6399.

(186) Yenugonda, V. M.; Deb, T. B.; Grindrod, S. C.; Dakshanamurthy, S.; Yang, Y. H.; Paige, M.; Brown, M. L. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 2714-2725.

(187) Li, Z.; Wang, D.; Li, L.; Pan, S.; Na, Z.; Tan, C. Y.; Yao, S. Q. *J Am Chem Soc* **2014**, *136*, 9990-9998.

(188) Lo Conte, M.; Pacifico, S.; Chambery, A.; Marra, A.; Dondoni, A. *Journal of Organic Chemistry* **2010**, *75*, 4644-4647.

(189) Mackiewicz, N.; Surendran, G.; Remita, H.; Keita, B.; Zhang, G.; Nadjo, L.; Hagege, A.; Doris, E.; Mioskowski, C. *Journal of the American Chemical Society* **2008**, *130*, 8110-+.

(190) Du Roure, O.; Debiemme-Chouvy, C.; Malthete, J.; Silberzan, P. *Langmuir* **2003**, *19*, 4138-4143.

(191) Nam, J.; Won, N.; Jin, H.; Chung, H.; Kim, S. *Journal of the American Chemical Society* **2009**, *131*, 13639-13645.

(192) Ng, S. L.; Yang, P. Y.; Chen, K. Y. T.; Srinivasan, R.; Yao, S. Q. *Organic & Biomolecular Chemistry* **2008**, *6*, 844-847.

(193) Laboria, N.; Wieneke, R.; Tampe, R. *Angew Chem Int Ed Engl* **2013**, *52*, 848-853.

(194) Li, L.; Ge, J. Y.; Wu, H.; Xu, Q. H.; Yao, S. Q. *Journal of the American Chemical Society* **2012**, *134*, 12157-12167.

(195)Khalil, I. A.; Kogure, K.; Akita, H.; Harashima, H. *Pharmacol Rev* **2006**, *58*, 32-45.

(196) Aubry, S.; Burlina, F.; Dupont, E.; Delaroche, D.; Joliot, A.; Lavielle, S.; Chassaing, G.; Sagan, S. *FASEB J* **2009**, *23*, 2956-2967.

(197) Tang, H. Y.; Yin, L. C.; Kim, K. H.; Cheng, J. J. *Chemical Science* **2013**, *4*, 3839-3844.

(198) Kim, M. S.; Buisson, L. A.; Heathcote, D. A.; Hu, H. P.; Braddock, D. C.; Barrett, A. G. M.; Ashton-Rickardt, P. G.; Snyder, J. P. *Organic & Biomolecular Chemistry* **2014**, *12*, 8952-8965.