TARGET PROFILING OF ELECTROPHILIC COMPOUNDS BY USING CHEMICAL PROTEOMICS STRATEGIES

YANG YANG

PhD

The Hong Kong Polytechnic University

2023

The Hong Kong Polytechnic University

Department of Applied Biology and Chemical Technology

TARGET PROFILING OF ELECTROPHILIC COMPOUNDS BY USING CHEMICAL PROTEOMICS STRATEGIES

YANG YANG

A thesis submitted in partial fulfilment of the

requirements for the degree of Doctor of Philosophy

Aug 2023

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

YANG Yang

Abstract

Target identification elucidates the mode of action of bioactive molecules and therefore is critical in modern drug discovery. Various methods for target identification have been developed in recent years including chemical proteomic, genomic methods, and bioinformatic prediction. However, most methods investigate compounds individually. It remains challenging to efficiently examine multiple compounds in parallel, particularly for structurally distinct compounds. Celastrol (Cel) has intriguing bioactivities and has been recognised as one of five promising natural products for drug discovery. In this study we have achieved two objectives (i) to employ a classic compound-centric chemical proteomics strategy to identify the molecular target of Cel; (ii) to extend target identification to other compounds that are similar to Cel, by developing a novel chemical proteomics-genomics strategy for high-throughput target profiling of electrophilic compounds in parallel.

First, we identified catechol-O-methyltransferase (COMT) as a binding target of Cel by using compound-centric chemical proteomics. After comprehensive characterization of Cel-COMT interactions, Cel was

demonstrated to inhibit the enzymatic activity of COMT and thus increased dopamine level in neuroendocrine chromaffin cells significantly, which explained the known dopaminergic and neuroprotective effects of Cel. Our study not only identified a novel binding target of Cel, but also provided a new scaffold and a cysteine hot spot for developing a new generation of COMT inhibitors to treat neurological disorders.

Next, we established a chemical proteomics-genomics strategy for multiplexed target profiling of bioactive small molecules. We selected Cel as a representative electrophilic bioactive compound, and other three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their perturbations of global transcription. An reactive cysteine profiling method was developed by using two complementary chemical probes to scrutinize reactive cysteines as potential binding sites of tested electrophiles. Then, target profiling of the four electrophilic compounds was performed by using the chemical proteomics approach, leading to identification of known and novel binding targets. Our results demonstrated that the cysteine reactivity of electrophilic compounds is

associated with their influence on gene transcription. Therefore, the integrated approach combining genomics and chemical proteomics enables multiplexing of target identification of structurally distinct compounds.

List of publications

Journal papers

1. **Yang, Y.**; Zhang, Q.; Yang, C.; Tse, Y. S.; Wong, K. Y.; Lam, T. C.; Zhao, Q.*, A chemical proteomics-genomics approach for multiplexed target profiling of electrophiles. (In submission)
2. Guo, H. #; **Yang, Y.** #; Zhang, Q. #; Deng, J. R.; Yang, Y.; Li, S.; So, P. K.; Lam, T. C.; Wong, M. K.; Zhao, Q.*, Integrated mass spectrometry reveals celastrol as a novel catechol-O-methyltransferase inhibitor. ACS Chem Biol 2022, 17 (8), 2003-2009. (#co-first authors)
3. Shen, C. #; Zhang, D. #; Xu, F. #; **Yang, Y.** #; Tan, Y.; Zhao, Q.*; Li, L.; Ding, K.; Li, Z.*, Two-photon fluorescent turn-on probes for highly efficient detection and profiling of thiols in live cells and tissues. Biol Chem 2022, 403 (4), 445-451. (#co-first authors)
4. Leung, C. O. N.; **Yang, Y.**; Leung, R.W.H.; SO, K. K. H.; Guo, H.; Lei M.M.L.; Yun J.; Ma, S.; Zhao, Q.*; Lee, T. K. W.*, Broad-spectrum kinome profiling identifies CDK6 upregulation as a driver of Lenvatinib. Nat Comm 2023. (In revision).

5. Li, J.; Dong, J.; Wang, W.; Yu, D.; Fan, X.; Hui, Y. C.; Lee, C. S. K.; Lam, W. H.; Alary, N.; **Yang, Y.**; Zhang, Y.; Zhao, Q.; Chen, C. L.; Tye, B. K.*; Dang, S.*; Zhai, Y.*, The human pre-replication complex is an open complex. Cell 2023, 186 (1), 98-111 e21.
6. Leeman-Neill, R. J.; Song, D.; Bizarro, J.; Wacheul, L.; Rothschild, G.; Singh, S.; **Yang, Y.**; Sarode, A. Y.; Gollapalli, M. C.; Whisenant, D. E.; Bhavsar, S.; Lim, J.; Swerdlow, S. H.; Bhagat, G.; Zhao, Q.; Berchowitz, L. E.; Lafontaine, D.; Wang, J.*; Basu, U.* Acquired non-coding mutations retarget super-enhancer activity during B cell lymphoma progression. Nature Genetics 2023. (Minor revision)
7. Yang, Y.; Wang, H.; Zhang, Y.; Chen, L.; Chen, G.; Bao, Z.; **Yang, Y.**; Xie, Z.; Zhao, Q.*, An Optimized Proteomics Approach Reveals Novel Alternative Proteins in Mouse Liver Development. Mol Cell Proteomics 2023, 22 (1), 100480.
8. Chen, L.; Zhang, Y.; Yang, Y.; **Yang, Y.**; Li, H.; Dong, X.; Wang, H.; Xie, Z.; Zhao, Q.*, An Integrated Approach for Discovering Noncanonical MHC-I Peptides Encoded by Small Open Reading Frames. J Am Soc Mass Spectrom 2021, 32 (9), 2346-2357.

9. Wang, H.; Wang, Y.; Yang, J.; Zhao, Q.; Tang, N.; Chen, C.; Li, H.; Cheng, C.; Xie, M.; **Yang, Y.**; Xie, Z.*, Tissue- and stage-specific landscape of the mouse translatome. Nucleic Acids Res 2021, 49 (11), 6165-6180.

Conference papers

1. **Yang, Y.**; Zhao, Q*. Integrative multi-omics strategy for target profiling of electrophilic compounds. 1st Chinese American Society for Mass Spectrometry (CASMS), 2021. (Poster)
2. **Yang, Y.** #; Guo, H. #; Zhang, Q. #; Deng, J. R.; Yang, Y.; Li, S.; So, P. K.; Lam, T. C.; Wong, M. K.; Zhao, Q.*, Integrated Mass Spectrometry Reveals Celastrol As a Novel Catechol-O-methyltransferase Inhibitor. 2nd Chinese American Society for Mass Spectrometry (CASMS), 2022. (Poster & Invited lightning presenters)
3. **Yang, Y.** #; Guo, H. #; Zhang, Q. #; Deng, J. R.; Yang, Y.; Li, S.; So, P. K.; Lam, T. C.; Wong, M. K.; Zhao, Q.*, Integrated Mass Spectrometry Reveals Celastrol As a Novel Catechol-O-methyltransferase Inhibitor. The 29th Symposium on Chemistry Postgraduate Research in Hong Kong, 2022. (Poster)

4. **Yang, Y.**; Zhao, Q.*. An integrative strategy for target profiling of electrophilic compounds. The 28th Symposium on Chemistry Postgraduate Research in Hong Kong, 2021. (Poster)
5. **Yang, Y.**; Zhao, Q.*. Chemoproteomics and quantitative proteomics method development to facilitate biological studies. The 27th Symposium on Chemistry Postgraduate Research in Hong Kong, 2020. (Oral Presentation & Poster)

Acknowledgements

Firstly, I would like to express my utmost gratitude to my supervisor Dr. Zhao Qian for her patient guidance, constant understanding, and continuous encouragement. Dr. Zhao is an extraordinary, conscientious, and responsible scientist who always prioritizes students' requests, in not only research but also personal matters. I have gained valuable experiences and knowledge of mass spectrometry and proteomics under her guidance. Grateful acknowledgement is also made to my co-supervisor Prof. Lo Chun-lap, Samuel for his kind help.

Also, my sincere thanks are given to our collaborators Prof. Wong Man-kin, Prof. Lee Kin-wah Terrence, Prof. Zhao Yanxiang, Prof. Li Zhengqiu, Dr. Ko Chi-bun Ben, Dr. Lam Chuen Thomas, Dr. Wong Kin-Yau Alex, Dr. Leung Oi-ning Carmen, and Miss Zhou Jessica for their trusts and supports.

I also owe a special debt of gratitude to the tremendous supports from Department of Applied Biology and Chemical Technology (ABCT), The University Research Facility in Chemical and Environmental Analysis (UCEA) and The University Research Facility in Life

Sciences (ULS). I would like to specially thank Dr. So Pui-kin and Dr. Wong Lai-king Iris for sharing knowledge and experience in techniques, and Ms. Kwok Fung-yee Peggy for her nice assistance in answering questions processing documents during my graduation and thesis submission.

In addition, I deeply appreciate to my team members and friends Dr. Zhang Qi, Dr. Chen Lei Alyssa, Dr. Faleti Oluwasijibomi Damola, Dr. Zhu Yanting, Mr. Guo Haijun, Miss Li Shuqi, Miss Yang Ying, Miss Shang Jin, Mr. Zhang Yuanliang, Mr. Li Kecheng, Miss Tse Yin-suen Chloe, Miss Yang Chenxi, Miss Feng Yu, Miss Gao Lei for their kind support for research and life during the study period. I am truly fortunate to have my friends.

Finally, I would like to dedicate this thesis to my beloved parents for their unconditional love and patience. At last but not least, special thanks go to my dearest Miss Xu Sha for her unwavering love and constant encouragement throughout my studies to strive through the days and nights.

Grant Support

This study was generously financed by,

1. RGC General Research Fund (15304819)
2. RGC Collaborative Research Fund Equipment (C5033-19E)

I sincerely thank for the generous support from the above-mentioned grants.

List of abbreviations

Full form	Abbreviation
Acetonitrile	ACN
Auranofin	Aur
Activity-based protein profiling	ABPP
Bicinchoninic acid	BCA
Celastrol	Cel
Cysteine	Cys
Connectivity Map	CMap
Catechol-O-methyltransferase	COMT
Cellular thermal shift assay	CETSA
Copper(II) sulfate	$CuSO_4$
Data-dependent acquisition	DDA
Dithiothreitol	DTT
Dopamine	DA
Ethynylbenziodoxolone	EBX
Electrospray ionization	ESI
Formic acid	FA
Higher-energy C-trap dissociation	HCD

Heat shock protein 90 beta	Hsp90β
Iodoacetamide	IAM
Iodoacetamide alkyne	IAA
Inhibitor of nuclear factor kappa-B kinase subunit beta	IKKB
Liquid chromatogram	LC
Label free quantitation	LFQ
Methanol	MeOH
Mass spectrometry	MS
3-Methoxytyramine	3-MT
Peroxiredoxin-1	PRDX1
Phosphate-buffered Saline	PBS
Phorbol 12-myristate 13-Acetate	PMA
Parallel reaction monitoring	PRM
Room temperature	RT
Sodium dodecyl sulfate	SDS
Stable isotope labelling by amino acids in cell culture	SILAC

Tetramethylrhodamine	TAMRA
Tris(2-carboxyethyl)phosphine	TCEP
Tris(benzytriazolylmethyl)amine	TBTA
Triptolide	TL
Ultra performance liquid chromatography	UPLC
Ultraviolet	UV
Vinpocetine	Vin
Withaferin-a	WA

Table of content

List of figures

List of tables

Chapter 1: Overview

1.1 Drug target identification strategies with chemical proteomics

1.1.1 Introduction to chemical proteomics

Bioactive small molecules such as natural products serve as a treasure mine of drug development. Identification of the cellular targets of bioactive small molecules has been the mainstay but has also presented as one of the biggest challenges in drug discovery[1, 2]. To meet the challenge, several systematic and nonbiased target profiling strategies have been developed, including the chemical proteomic approach[3, 4], the genomic method[5, 6], and the bioinformatic prediction[7, 8].

Chemical proteomics is a useful approach to characterize small molecule mode of action and protein function[9]. As a rapidly evolving area of chemical biology, chemical proteomics utilises synthetic chemistry to design and develop chemical probes with high selectivity to a specific protein or a group of proteins. In conjunction with modern mass spectrometry (MS) techniques, the identification of protein binding targets or partners of small molecules in drug target discovery

is enabled[10, 11]. Chemical proteomics comes with two different flavours: (i) activity-based probe profiling (ABPP) and (ii) compound-centric chemical proteomics (CCCP)[12].

Typical ABPP begins with the design of small-molecule probes to capture specific enzyme families in proteomes[13, 14]. The synthetic probes contain a reporter group, a linker and a reactive group. For example, alkyne or azide as widely used functional groups are available for click reaction. Reactive groups are required to interact with the target proteins, forming a stable covalent bond[15, 16]. After protein targets are labelled by the probes, subsequent bioorthogonal ligation is performed between an alkyne and azide. Copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) is the most widely used bioorthogonal click reaction, enabling azides and alkynes to generate 1,4-disubstituted 1,2,3-triazoles under mild conditions[17, 18]. It is demonstrated that probes designed with alkynes groups have a smaller size in synthesis and show lower labelling background than that with azides[19]. When CuAAC ligation is used to conjugate a fluorescent label, the labelled proteins can be visualized in sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel[20, 21]. When

CuAAC ligation is used to conjugate with an affinity label like biotin, it allows for the enrichment and pull-down of labelled proteins in streptavidin beads or agarose. The strong binding affinity ($K_d = 10^{-14}$ M^{-1}) to streptavidin of biotin makes it feasible to isolate the complexes of activity-based probe and protein targets from proteome.

Alternatively, a competitive ABPP strategy has emerged as a popular tool in early-stage target identification of small molecules[22-24]. In competitive ABPP, the lysate is incubated with the tested compounds prior to labelling with an activity-based probe. If the probe is able to label particular proteomes and the tested compound binds one or more of these with the same site as the probe, a decreased signal of target proteins can be observed by using MS. The schematic diagram of the typical ABPP, as well as the competitive ABPP, is presented in **Figure 1-1**.

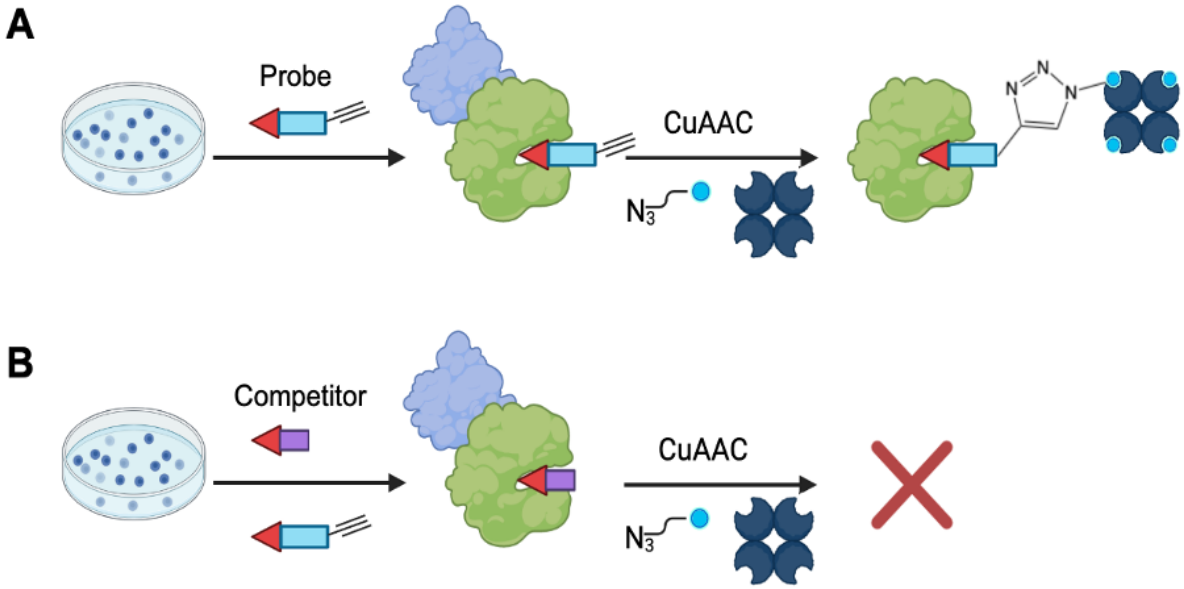


Figure 1-1 A schematic diagram of the typical ABPP and the competitive ABPP. (A) Target proteins in live cells or lysate are labelled with the activity-based probes, conjugated to biotin via CuAAC ligation, and captured on an affinity matrix such as the streptavidin beads. (B) Target proteins in live cells or lysate are pre-reacted with the competitor and subsequently with the activity-based probe. By competing for the common targets, the probe labeling of the target proteins is inhibited.

Typical CCCP begins with the structural modification of the tested compound. The first step is to modify the compound with an affinity tag (e.g., biotin), an functional group (e.g., alkyne), or immobilize it to

a solid matrix (e.g., agarose beads) for affinity chromatography[3, 25]. A key principle is that the structural modification should not affect the pharmacological activity of the parent compound of interest. It is critical to follow up the modified compound with biochemical assays to evaluate if the pharmacological activity is still retained[26]. In a classic workflow, target proteins are captured by compound-immobilized agarose beads and subsequently identified by MS[27, 28] (**Figure 1-2A**), or labelled by a compound-modified probe, reacted by CuAAC ligation, and enriched on an affinity matrix like streptavidin beads (**Figure 1-2B**).

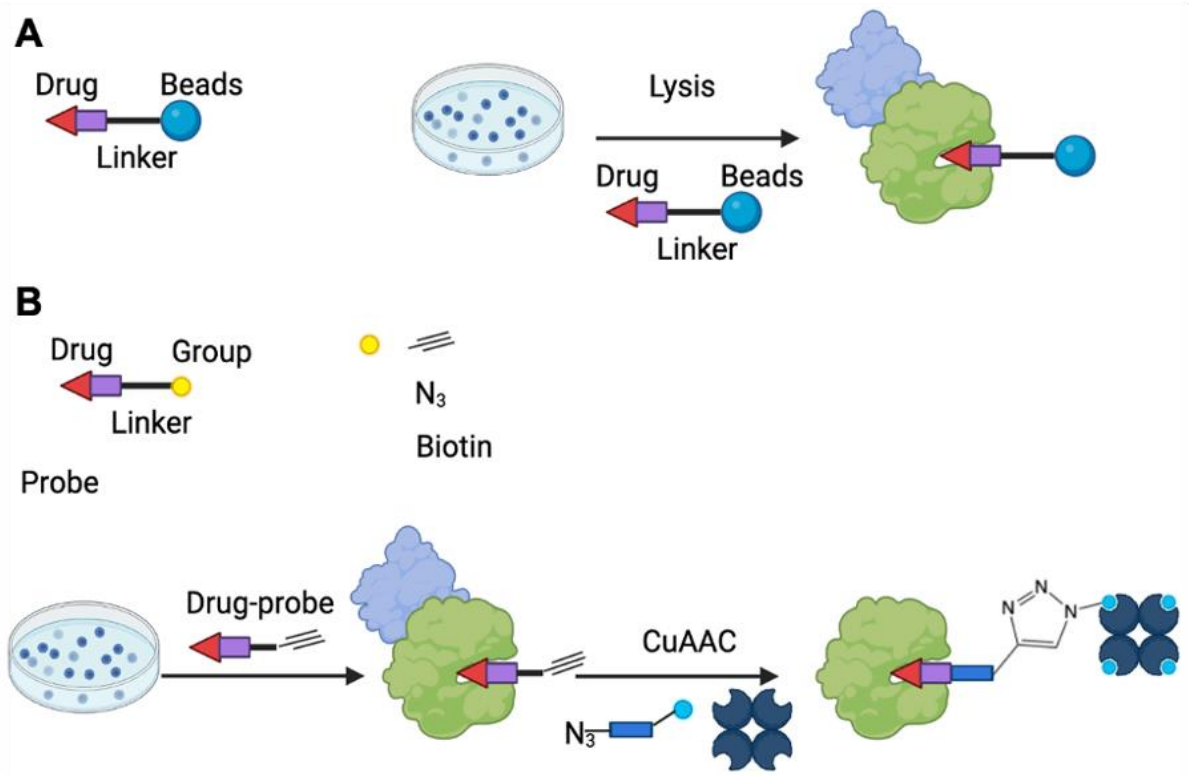


Figure 1-2 A schematic diagram of the typical CCCP. (A) The compound is immobilized with a biocompatible matrix. The compound-immobilized matrix is then incubated with proteomes to capture target proteins for further analysis. (B) The compound-modified probe is incubated with live cells or lysate to label target proteins, followed by CuAAC ligation and affinity matrix enrichment.

Comparison of ABPP and CCCP in drug target identification

In general, CCCP is a more unbiased method than ABPP in target identification, because it allows for the comprehensive identification of

theoretically all protein targets, without prior knowledge on function or classfication of the target proteins. However, CCCP highly relies on the synthesis of structural modification in tested compounds. The major limitation of CCCP is the strict requirement of the compound-centric synthesis to retain the compound's pharmacological activity. ABPP can be used to detect specific targets based on the activity of particular proteins and can even to identify the binding site directly. However, non-specific binding of the probes is the greatest limitation of most chemical proteomics strategies, which can be sometimes addressed in the competitive ABPP strategy[29, 30].

1.1.2 Cleavable linkers in chemical proteomics

In chemical proteomics, the identity of target protein is typically inferred from its constituent peptides by using a bottom-up proteomics strategy. When the target protein is revealed, the probe-modified residue/peptide could get lost and become undetectable. However, the direct binding site information of the tested compound is important for mechanism interpretation[31]. So far, there are several designed azide-biotin form that incorporate a cleavable linker, enabling the release of

the probe-labelled peptide and identification of the direct binding site of the target protein[32-34]. The schematic diagram of the site mapping strategy via the cleavable linker is shown in **Figure 1-3**.

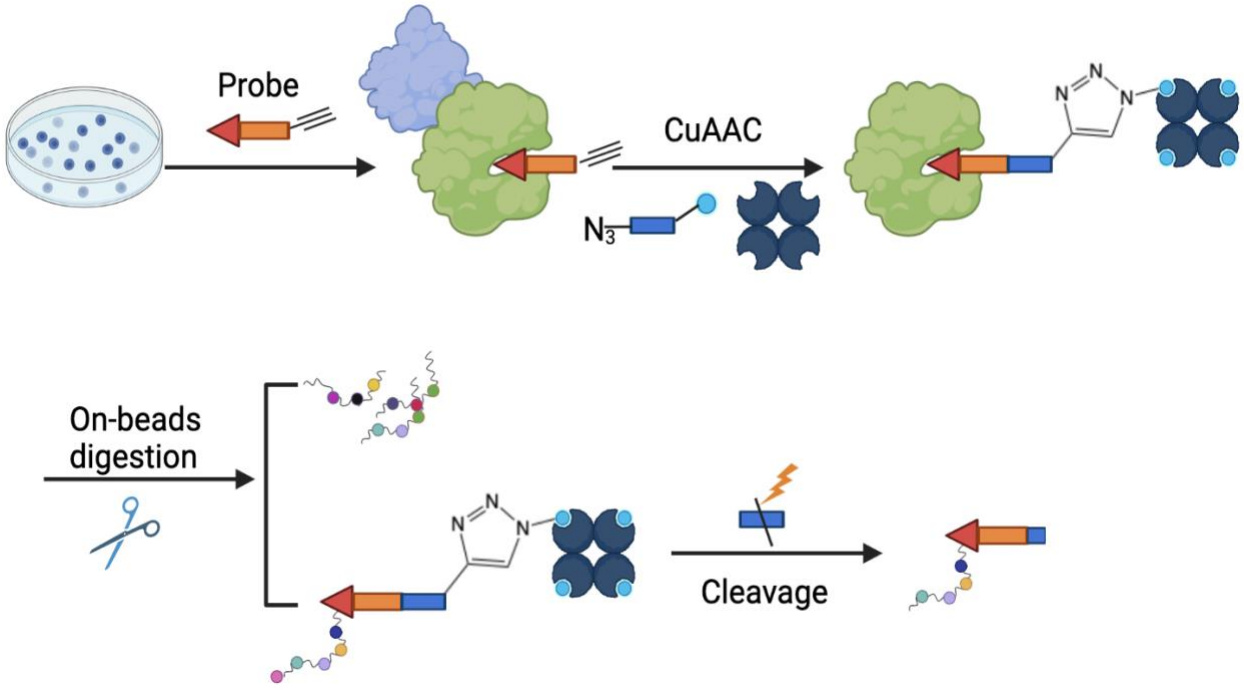


Figure 1-3 The schematic diagram of site mapping strategy via the cleavable linker in chemical proteomics. Target proteins in live cells or lysate are labelled by the activity-based probe, reacted by CuAAC ligation with a cleavable azide-biotin, and captured on an affinity resin such as streptavidin beads for further on-beads digestion. The modified peptide can be selectively cleaved from beads for further analysis.

Several cleavable azide-biotins have been reported previously, including acid-cleavable dialkoxydiphenylsilane (DADPS) biotin azide[35], hydrazine-cleavable DDE biotin azide[36], enzyme-cleavable tobacco etch virus protease (TEV) biotin azide[37], and ultraviolet (UV) cleavable biotin azide[38].

Acid-cleavable DADPS biotin azide

An ideal acid-cleavable linker should be sensitive and responsive to minor changes in Pondus Hydrogenii (pH). A typical acid-cleavable biotin probe, like DADPS biotin azide designed by Szychowski *et al.*[39], links a biotin with an azide moiety through a 'DADPS' linker. Captured proteins can be efficiently released from beads under mild conditions (5~10 % formic acid for one hour) and a small fragment (143 Da) is left from the labelled protein after the acid cleavage. In Szychowski's study, cleavage of the DADPS biotin azide was proved to have 98% cleavage efficiency when using 10% formic acid, while the previous acid-cleavable probe required above 95% trifluoroacetic acid to get a comparative cleavage. Rabalski *et al.*[35] conducted a head-to-head comparison of two cleavable biotin azide DADPS and DIAZO in

reactive cysteine identification. When DADPS biotin azide was used in CuAAC for cysteine mapping, it achieved a higher identification number of cysteine-contained peptides (7866) than DIAZO biotin azide (4326). Using DADPS biotin azide in cysteine profiling, over 10,000 unique cysteine residues were detected and quantified successfully in cancer cell lines.

Hydrazine-cleavable DDE biotin azide

DDE biotin azide is made comprised of a biotin moiety connected to an azide group by a spacer arm that contains a hydrazine-cleavable DDE moiety. With a mild condition of 2% aqueous hydrazine, the DDE liner can be cleaved, releasing the target fragments for MS identification. Griffin *et al.*[40] employed an alkyne-DDE-biotin linker in the mapping of N-acetylglucosamine monosaccharide (O-GlcNAcylation) in cell lysates. They also compared the DDE biotin azide with a photocleavable biotin azide in terms of the labelling efficiency and the recovery rate of O-GlcNAcylated proteins. The result suggested that the DDE biotin linker can be cleaved more efficiently than the photocleavable linker.

Enzyme-cleavable biotin azide

It has been reported that virus protease-cleavable startegy offers higher cleavage specificities because of the strict requirement of the amino acid sequence as the recognition site[41]. Cravatt *et al.*[42] developed a TEV azide-biotin that links the cleavable sequence ENLYFQG for selective release of the probe-modified peptides. They incubated cell lysates with an iodoacetamide alkyne probe to capture cystine-containing proteins. Followed by click reaction and digestion, the cysteine-containing peptides could be released after a 12-hour TEV protease cleavage. The developed method, isotopic tandem orthogonal proteolysis-ABPP (isoTOP-ABPP), offers a high-throughput way of predicting functional cysteines in complex biological samples. This outstanding work subsequently inspired more applications in target identifications of electrophilic natural products[43, 44].

UV-cleavable biotin azide

The UV-cleavable linker is selective and stable even under both acidic and basic conditions. The wavelength absorptions for the UV-cleavable biotin azide are in the near-UV (300–365 nm) range of the spectrum.

The cleavage condition is under irradiation with 365 nm UV light for 1 or 2 hours. Yang *et al.*[38] developed a multiplexed thiol reactivity profiling method using a UV-cleavable linker for the target identification of seven electrophilic and structurally diversified natural lactones, revealing potential cysteine targets in a multiplexed manner.

1.1.3 Label-free and label-based quantitative MS approaches

Quantitative proteomics is a powerful approach to drug discovery and target identification. With the growing sensitivity and resolution of modern mass spectrometers, quantitative proteomics has become the forerunner of chemical biology[45]. The strategies in quantitative proteomics can be broadly classified to label-free and label-based.

Label-free quantification

Label-free quantification (LFQ) allows for the determination and quantification of the relative changes in at least two biological samples. The number of tested samples is not limited in principle[46]. A general label-free quantitative proteomics workflow contains three basic and fundamental steps, including sample preparation, sample analysis by

liquid chromatography with tandem MS (LC-MS/MS) and data processing[23, 47, 48]. Sample preparation in label-free chemical proteomics studies follows a classical bottom-up strategy: extracted proteins from biological samples turn into peptides via proteolytic digestion. A peptide fractionation is next performed to reduce the complexity of the peptide mixtures. Then the desalted peptides are analysed by MS. The quantification can be inferred by the comparison of the spectra counts of a given protein or the peak intensities from the precursor ion of the peptide.

The following recent reports of chemical probes highlight the application of LFQ chemical proteomics in target identification of natural products. Wang *et al*.[49] designed an alkyne-tagged probe for antimalarial drug artemisinin to study the targets and mechanisms of artemisinin activation. Artemisinin is extracted from the plant *Artemisia annua* and used in the treatment of malaria. However, the target protein and the mode of action of artemisinin are not fully understood. Wang *et al*. identified 124 covalently binding protein targets and visualised artemisinin activation levels both *in vitro* and *in situ*. They also found that haem is the predominant factor in artemisinin

activation rather than free ferrous iron. Zheng *et al.*[50] synthesised several alkyl chain analogues of natural antibiotic cerulenin as chemical probes to investigate its targets in melanoma cells. They identified over 200 proteins as cellular targets of the cerulenin probe. Zheng *et al.* found that around 140 probe-labelled proteins can be abolished by both palmitoyl acyltransferase inhibitors and cerulenin, indicating that these proteins belonged to potential targets of cerulenin related with protein palmitoylation function.

Quantification with chemical labeling of the peptides

Although LFQ shows the utility to quantify the relative abundance of a given protein in a straightforward way and allows non-limiting of sample numbers, it is challenging to overcome the technical variance during sample preparation and the deficiency in detecting low abundance proteins or peptides. In contrast, label-based methods utilise stable isotope labels, which can conduct quantitation among samples in the same run rather than in different runs, to achieve high accuracy and reproducibility of quantitative results[51]. Because a high-resolution mass spectrometer is able to characterise the mass-to-charge difference

between the labelled peptide and its corresponding unlabelled peptide, quantification is performed by comparing the MS signal intensities[52]. Label-based methods can be classified into *in vivo* labelling and *in vitro* labelling. A typical technique of *in vivo* labelling is stable isotope labelling using amino acids in cell culture (SILAC)[53, 54]. Tandem mass tag (TMT) and isobaric tags for relative and absolute quantification (iTRAQ) are two widely used methods of *in vitro* labelling[55, 56]. The major concerns of label-based approaches are the limited multiplexing capacity and the cost of isobaric tags.

The reports described below of chemical probes highlight the application of label-based chemical proteomics in target identification of natural products. Kreuzer *et al*.[57] designed functionalized probes of natural product acivicin for SILAC-ABPP in a liver cancer cell line, indicating a specific interaction of ALDH4A1 with acivicin. Further target validation suggested that acivicin may inhibit ALDH4A1 activity through a reaction with its catalytic site cysteine348. Gleissner *et al*.[58] synthesised the alkynylated probe of the natural product neocarzilin A and applied SILAC-ABPP for target discovery. The results revealed that neocarzilin A targets the largely uncharacterized

synaptic vesicle membrane protein VAT-1. In addition, VAT-1 was demonstrated to play a role in cancer cell migration and neocarzilin A was demonstrated to be a potent inhibitor of cell motility. Sharma *et al*.[59] used ATP-analogue probes for kinase profiling by a TMT-ABPP quantitative approach in four BRAF inhibitor-naïve or resistant cell lines. The result revealed that BRAF inhibitor resistance is correlated with the enhancement of proteins related to cytoskeletal organisation and adhesion as well as the decrease of proteins implicated in cell metabolic processes. In addition, different cell lines showed high heterogeneity in kinase utilisation but similar phenotypic adaptations, indicating a challenge to solving therapy resistance in the clinic.

1.1.4 Other approaches

Cellular thermal shift assay (CETSA)

The thermal shift assay (TSA) has become widely used in drug target research over the last 20 years. However, the application of TSA is limited to recombinant proteins. A novel but conceptually similar technology called CETSA was developed by Molina *et al*.[60] to broaden the applicability of TSA. Molina *et al*. established CETSA by

determining compound-induced stabilisation in complex biological systems. CETSA conducts the direct measurement of protein-small molecule interactions. In addition, CETSA is a modification-free method that does not require additional efforts in chemical synthesis.

In a typical CETSA protocol, live cells are first treated with either vehicle or tested compounds. Next, the cells are heated at various temperatures. With the temperature rises, proteins start to denature, and finally aggregate and precipitate out. Then, the soluble protein is able to be separated after rapid chilling, cell lysis and centrifugation. The semiquantitative result of the target protein can be visualized by western blotting. In brief, for cells treated with various concentrations of tested compounds while keeping the temperature constant, it is expected to observe a significant difference of targeted protein signal in western blotting results between the control and the compound-treated groups[61, 62].

Later Savitski *et al.*[63] extended the readout of traditional CETSA from western blotting to MS. First, the authors compared the differences in drug responses to thermal shift by using cell lysates or live cells,

suggesting that cell integrity has a very important impact on protein stability. In their study, staurosporine treatment produced thermal shifts for above 50 targets of 7000 detected proteins. Moreover, the results showed that the treatment of dasatinib induced thermal shifts in known downstream effectors. The result also demonstrated that ferrochelatase was a novel protein target for the ALK inhibitor alectinib and the BRAF inhibitor vemurafenib.

Thermal proteome profiling (TPP)

Traditional CETSA requires prior knowledge of interested targets and measures the stability shifts of target proteins in cellular samples. Thus, it is not available for unbiased target screening and cannot be performed at proteome level. To address these issues, CETSA has been developed into an unbiased proteome-wide screening method, called TPP[64], combining CETSA with multiplexed quantitative MS for target identification and validation. Using TMT labelling technology based on MS, TPP can perform simultaneous measurements of ten temperature points to show the shift for each detected protein. Similar to CETSA, TPP can be conducted both in live cells and lysates.

However, it should be noted that proteins showing thermal stability shifts in live cells but not in cell lysates are likely to be indirect targets[65].

Drug affinity responsive target stability (DARTS)

In 2009, Lomenick *et al.*[66] developed a method called DARTS for target identification. Firstly, cell lysates are treated with or without tested compounds. Then, a protease (e.g., thermolysin) is added to the lysate samples to trigger proteolysis. Enriched target proteins can then be further analysed through immunoblotting or MS. Pronase is preferred in DARTS experiments as some proteins are not sensitive towards thermolysin. In this study, the authors confirm known drug-target interactions by DARTS to demonstrate its feasibility. Recently, it was reported that a similar approach[67] was built to study the activity and mechanism of Grape Seed Extracts (GSE) in *in vivo* models. The result showed that endoplasmic reticulum stress response proteins might be potential targets of GSE. In summary, DARTS is straightforward and efficient. However, the major drawback is the difficulty of selecting the optimal proteolysis condition for an unknown target. Moreover, proteins are usually undergoing a limited digestion

under a native condition in DARTS. However, some proteins are folded in native so that proteases is unable to reach their cleavage sites, it has been demonstrated that those proteins can be kept intact even after digestion for several days[68]. Thus, DARTS is normally unable to screen these natively folded proteins in target discovery.

Stability of proteins from rates of oxidation (SPROX)

SPROX was created by West *et al*.[69]. In SPROX, the stability of proteins is measured by the oxidation rates of methionine residues. Cell lysates from each analysis are diluted into a series of buffers containing guanidinium hydrochloride (GdmCl), a chemical denaturant, in varying concentrations. The identical amount of hydrogen peroxide is then added to the aliquoted protein samples in GdmCl buffer. After quenching the protein oxidation reaction, quantitative MS analysis is able to quantify the non-oxidised and oxidised methionine-containing peptides. The target proteins, including direct or indirect binding with a compound, show transition midpoint shifts between the control group and the compound-treatment groups. Three important values, protein folding free energy, binding free energy and disassociation constant,

can be calculated by the transition midpoint shifts. In complex biological systems, SPROX can evaluate protein folding and compound-binding characteristics simultaneously[70]. In this experiment, SPROX was used to assess the folding and chemical binding characteristics of 327 proteins in yeast cell lysate.

SPROX is compatible with multiple types of quantitative approaches to quantify the non-oxidised and oxidised methionine-containing peptides. Nevertheless, taking into account the SPROX protocol's numerous denaturation steps, SPROX prefers TMT, iTRAQ, and SILAC. For example, recent research using SPROX with a TMT-based proteomics approach was reported to quantify the folding properties of about 10,000 unique peptides of about 3000 proteins in human cell lysate[71]. However, SPROX is only available for the proteins containing methionine residues and is not available for methionine-absence proteins. Therefore, methods involving chemical alteration of other amino acid residues were subsequently suggested to solve the issue. For instance, an amine reactive reagent, called S-methylthio acetimidate, can be applied to measure the protein folding and ligand binding properties via SPROX in proteins containing lysine residues[72].

1.1.5 Current and future challenges

Chemical proteomics enables an unbiased, global and quantitative profile of protein-compound interactions[23]. However, there are some unavoidable challenges. On the one hand, when chemical probes are used at high concentrations, there are non-specifically labelled proteins. Some proteins can interact with the affinity matrix such as streptavidin resin, increasing the background of profiling. On the other hand, because the true targets may not always belong to the highly abundant protein portions identified by MS and are likely to be masked by background, it is difficult to choose a cut-off to obtain target candidates for further studies[73, 74]. Designing a competitive experiment is one approach to assess whether the identification is probe-specific. The optimisation of the concentration of the tested compound as well as the probe also provides more reliable results in chemical proteomics. Furthermore, synthesizing and testing multiple probes simultaneously can reduce the risk of probe-dependent off-targets and cross-validate target identifications from different probes.

Although probes in theory can be derived to mimic natural products and provide insights into potential targets, it is still challenging to synthesise the ideal probes for some natural products with complex structures[75, 76]. In addition, it is also intriguing that most clinically approved drugs are found to bind membrane receptors in live cells[77]. Particularly, only a few probes have been successfully synthesised and applied to address membrane proteins[78, 79].

The advances of cleavable linkers coupled with modern MS technology make chemical proteomics more applicable in target profiling of natural products[80, 81]. Reducing the risk of off-targets, expanding the multiplexing quantitative capacity, and developing precise mappings of direct binding sites continue to be critical and at the cutting edge of MS-based chemical proteomics[82, 83]. The combination of multiple approaches for target identification or validation brings benefits to overcome the limitations of a single approach (**Table 1-1**). Furthermore, other modification-free approaches including CETSA, TPP, SPROX, and DARTS can serve as complementary strategies in target identification and validation of natural products[84, 85].

Table 1-1 Summary of target identification approaches.

Method	Sample Type	Detection	Quantitative Strategies	Biophysical information	Limitations
ABPP	Intact cells Cell lysates	Western blotting In-gel fluorescence scan MS	Protein or peptide	Specific enzyme families captured by the activity-based probe	Only allows identification in pre-defined protein classes
CCCP	Cell lysates	Western blotting In-gel fluorescence scan MS	Protein or peptide	Comprehensive target capture by the direct compound-immobilized matrix	Highly rely on the synthesis Non-specific binding
CETSA	Intact cells Cell lysates	Western blotting	Protein	Thermal melting curves	Only available for thermal sensitive proteins Only for known interested targets Cannot be applied at proteome level Cannot reveal binding sites

Method	Sample Type	Detection	Quantitative Strategies	Biophysical information	Limitations
TPP	Intact cells Cell lysates	MS	Protein	Thermal melting curves	Only available for thermal sensitive proteins Cannot reveal binding sites Incompatible with membrane proteins
DARTS	Cell lysates	Western blotting MS	Protein	Protease susceptibility	Only for known interested targets (for Western blotting) Cannot reveal binding sites
SPROX	Cell lysates	MS	Peptide	Protein folding energy Binding free energy	Relies on the exposure of reactive amino acid residues Limited to some amino acid residues

1.2 Electrophilic compound Celastrol (Cel) and its bioactivities

Cel, a pentacyclic triterpene with electrophilic quinone methide extracted from the Chinese medicine *Tripterygium wilfordii*, was recognised as one of the five most promising natural products for drug discovery[86]. Many studies has been conducted over the last few decades to explore Cel's wide variety of activities including anticancer[87, 88], anti-obesity[89, 90], anti-inflammation[91, 92] and neuroprotection[93, 94].

The anti-cancer effects and the related molecular targets modulated by Cel have been widely studied in various tumor cell lines and animal models (**Figure 1-4**). Recent studies have revealed that Cel can attenuate tumour growth by inhibiting cell proliferation[95, 96], inducing apoptosis[97, 98] autophagy[99], suppressing angiogenesis[96], and cell metastasis[100]. Although considerable efforts have been made to reveal potential binding targets of Cel[101-103], many anti-cancer effects of Cel cannot be explained with the known protein targets and clear binding stoichiometry. The major known targets as well as the corresponding key pathways are shown in **Figure 1-5**.

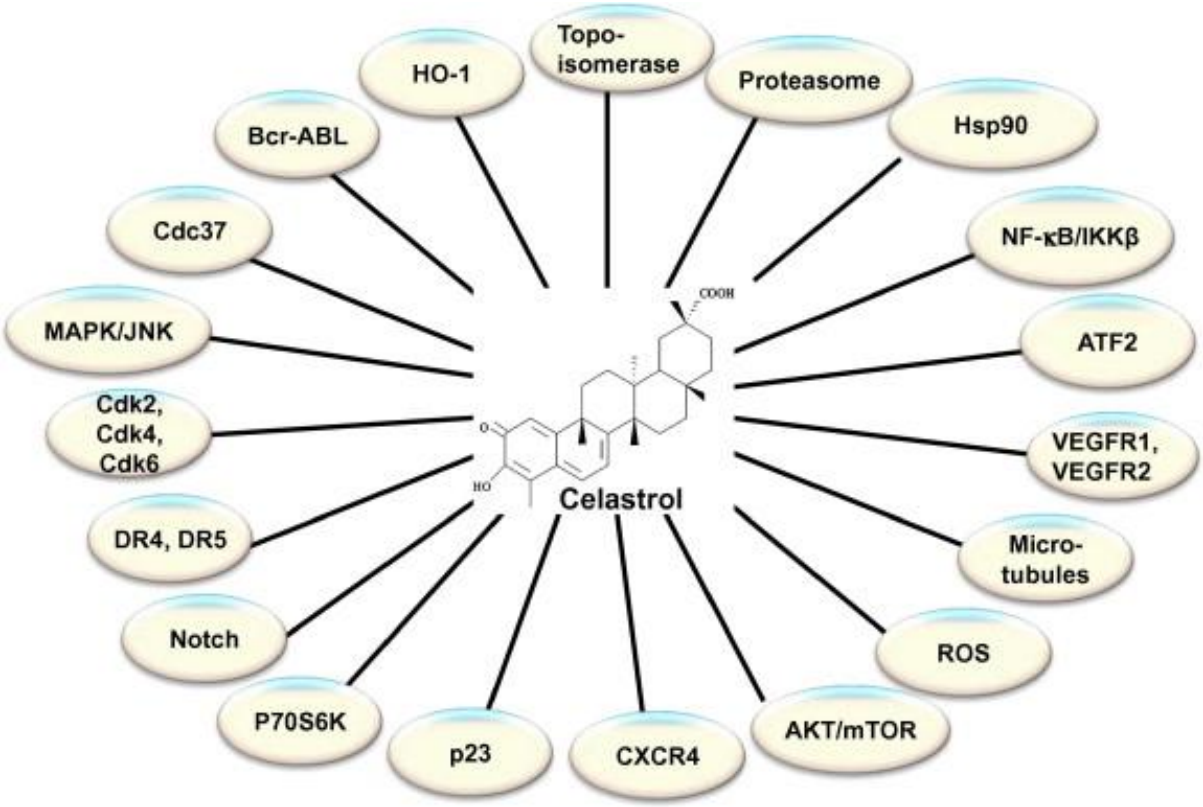


Figure 1-4 Molecular targets modulated by Cel in various cancer cell lines cells and animal models[104]. Permission from Elsevier and Copyright Clearance Center, please see Appendix for the copyright agreement.

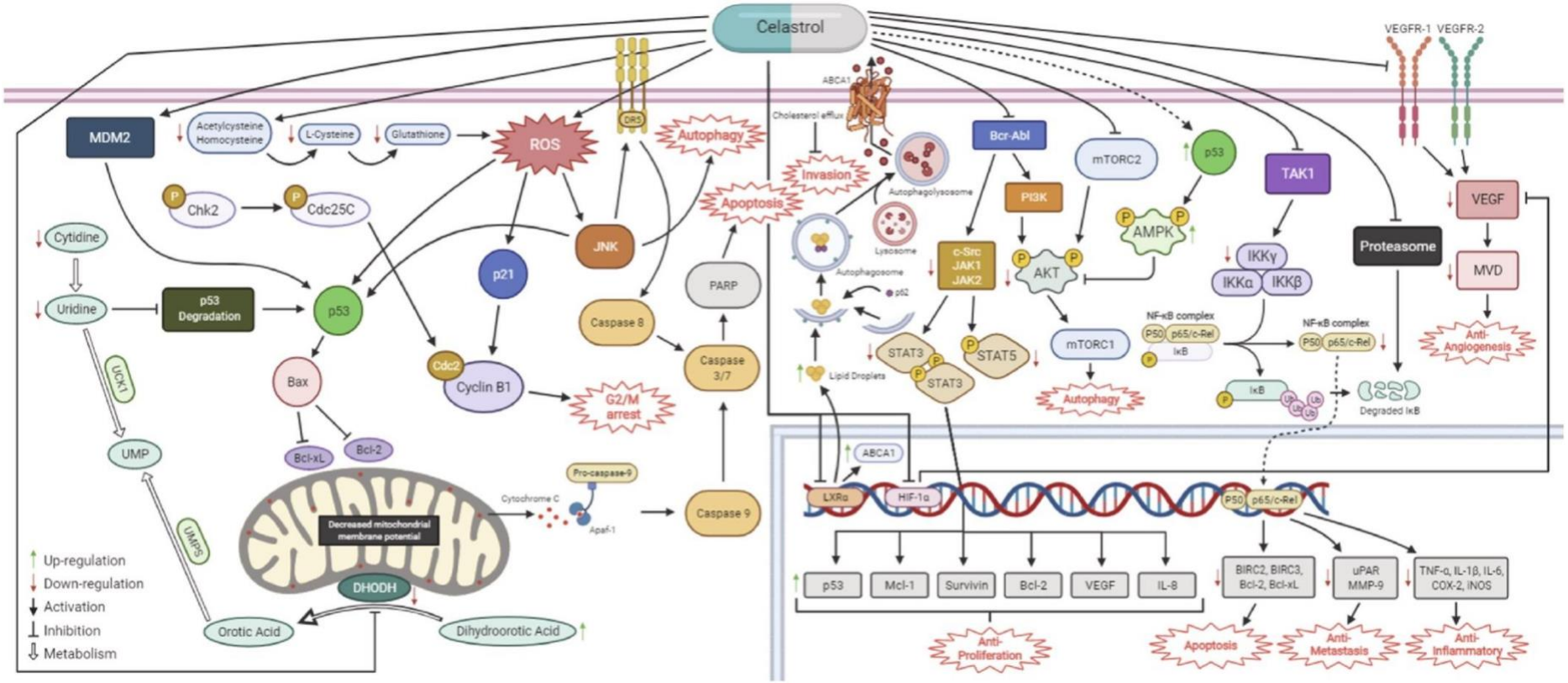


Figure 1-5 Molecular targets and pathways of Cel's anti-cancer mechanisms[105]. Permission from Elsevier and Copyright Clearance Center, please see Appendix for the copyright agreement.

Intriguingly, recent works since 2015 demonstrate that Cel plays a role in the treatment of obesity. Dietary Cel in mice reduces intestinal lipid absorption and the expression level of lipid transporters, as a leptin sensitizer, resulting in decreased body weight[89, 106]. The neuroprotection and anti-inflammatory effects of Cel in cerebral ischemia-reperfusion injury have been revealed by quantitative chemical proteomics studies, suggesting HMGB1[93] as a potential target of Cel. In addition, Cel shows a dopaminergic neuroprotective effect in Drosophila Parkinson's disease[107], reducing the loss of dopaminergic neurons in the brain. However, the mechanism of Cel's neuroprotective function remains unknown, thus, the search for the corresponding targets is imperative.

1.3 Outline of this thesis

The work is divided into two parts:

1. Identification and validation of the molecular target of Cel
2. Development of a high-throughput strategy for multiplexed target profiling of electrophiles

In the first part, we identified catechol-O-methyltransferase (COMT) as a binding target of Cel with compound-centric chemical proteomics. We developed a clickable probe, that retained the bioactivity of Cel and bound to its cellular targets in live cells. Target profiling indicated a strong interaction between COMT and Cel. We also analysed the binding stoichiometry between Cel and COMT to characterise the major binding residue. Furthermore, Cel was demonstrated to inhibit the enzymatic activity of COMT and thus increased dopamine level in neuroendocrine chromaffin cells significantly, which explained the known dopaminergic and neuroprotective effects of Cel. Our study not only identified a novel binding target of Cel, but also provided a new scaffold and a cysteine hot spot for developing a new generation of COMT inhibitors to treat neurological disorders. (Chapter 2~4).

In the second part, we developed a chemical proteomics-genomics strategy for multiplexed target profiling of electrophiles. Inspired by the applications of the L1000 transcriptomics platform in target identification, we asked whether it could serve as a sensor for target profiling via clustering or distinguishing small molecules based on their global transcription similarities. We selected Cel as a representative electrophilic compound, and other three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their cellular perturbations of global transcription. Next, an reactive cysteine profiling method was performed by using two complementary chemical probes to scrutinize reactive cysteines as potential binding sites of tested electrophiles. By using the chemical proteomics approach, known and novel binding targets of tested electrophiles were identified and validated. We demonstrated the cysteine reactivity of electrophilic compounds is associated with their cellular perturbations of global transcription. The integrated approach combining genomics and chemical proteomics enables multiplexing of target identification of structurally distinct compounds. (Chapter 5~8).

Chapter 2: Target identification of Cel with compound-centric chemical proteomics

2.1 Introduction

Cel is recognised as one of five promising natural products for drug discovery[108, 109]. However, various biological activities of Cel are not fully clarified with the known protein targets or clear binding stoichiometry. For instance, the mechanism through which Cel induces neuroprotection in the brain is still unknown, thus a comprehensive target profiling of Cel is called for.

2.2 Materials and methods

2.2.1 Chemicals and reagents

Cel (HY-13067) was supplied by MCE. $^{13}C6^{15}N2$ L-lysine-HCl (608041), phorbol 12-myristate 13-acetate (PMA), L-arginine monohydrochloride (A6969), L-lysine monohydrate (L9037) as well as compounds used in probe-synthesis including alkynyl amide and propargyl bromide were ordered from Sigma-Aldrich. $^{13}C6^{15}N4$ L-arginine-HCl (TJ271440), $CuSO_4$ (209198), Triton X-100, NP-40

(IGEPAL® CA-630), sodium dodecyl sulfate (SDS), iodoacetamide (IAM), protease inhibitor (11873580001), benzonase (70746-4), dithiothreitol (DTT), Tris(2-chloroethyl phosphate (TCEP), fetal bovine serum (FBS, including dialyzed FBS), penicillin, streptomycin, tris (hydroxymethyl) aminomethane hydrochloride (Tris-HCl), hydroxyethyl piperazine ethane sulfonic acid (HEPES), sodium chloride (NaCl), phosphate buffered saline (PBS), dimethyl sulfoxide (DMSO), acetonitrile (ACN), and chemiluminescent substrate (34580) were purchased from Thermo Fisher Scientific. All antibodies in the experiments includes COMT antibody (ab126618: rabbit source from Abcam and sc-137253: mouse source from Santa Cruz), β-actin antibody (AF7018, Affinity), IκBα antibody (sc-1643, Santa Cruz), goat anti-mouse HRP (7076S, CST) and goat anti-rabbit HRP (AS014, ABclonal).

2.2.2 Synthesis and characterization of Cel probes

To synthesise probes C1a, C1b and C1c, Cel, hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) and amine were mixed at the 1:1.5:1.5 molar ratios. The mixture was stirred for 16 hours at

room temperature (RT). For probe C2a synthesis, Cel, $NaHCO_3$ and propargyl bromide toluene solution were mixed at 1:3:2 molar ratios in dry DMF and stirred at RT for 3 hours. The mixture was concentrated and processed via flash column chromatography on silica gel. Ethyl acetate/n-hexane (1:4) as eluent was employed to collect the desired products C1a (orange solid, 68% yield), C1b (orange solid, 72% yield), C1c (orange solid, 63% yield) and C2a (orange solid, 92% yield).

C1a ^{1}H NMR (400 MHz, $CDCl_3$, 298 K) δ 7.00 (d, J = 6.3 Hz, 2H), 6.50 (s, 1H), 6.33 (d, J = 7.2 Hz, 1H), 5.99 (t, J= 4.8 Hz, 1H), 4.00 (ddd, J = 17.6, 5.4, 2.5 Hz, 1H), 3.83 (ddd, J = 17.6, 4.4, 2.5 Hz, 1H), 2.41 (d, J = 15.6 Hz, 1H), 2.19 (s, 3H), 2.17 (t, J = 2.5 Hz, 1H), 2.15-1.96 (m, 3H), 1.93 -1.78 (m, 3H), 1.72-1.62 (m, 3H), 1.60-1.45 (m, 4H), 1.42 (s, 3H), 1.24 (s, 3H), 1.15 (s, 3H), 1.11 (s, 3H), 1.06-0.96 (m, 1H), 0.63 (s, 3H).

^{13}C NMR (100 MHz, $CDCl_3$, 298 K) δ 178.46, 177.56, 170.32, 164.91, 146.13, 134.23, 127.51, 119.56, 118.17, 117.24, 79.71, 71.93, 45.20, 44.49, 43.11, 40.32, 39.49, 38.28, 36.44, 35.02, 33.64, 33.56, 31.69, 31.39, 30.89, 30.13, 29.63, 29.43, 28.77, 21.87, 18.55, 10.37.

HRMS (ESI) [M+H]+ calculated for $C_{32}H_{42}NO_3$ 488.3159, found 488.3161.

C1b ^{1}H NMR (400 MHz, $CDCl_3$, 298 K) δ 7.00 (d, J = 7.2 Hz, 2H), 6.50 (s, 1H), 6.32 (d, J = 7.1 Hz, 1H), 6.17 (t, J = 5.8 Hz, 1H), 3.35-3.17 (m, 2H), 2.78 (s, 2H), 2.44 (d, J = 15.4 Hz, 1H), 2.34-2.25 (m, 2H), 2.19 (s, 3H), 2.15-1.95 (m, 4H), 1.94-1.78 (m, 3H), 1.71-1.45 (m, 5H), 1.41 (s, 3H), 1.24 (s, 3H), 1.14 (s, 3H), 1.10 (s, 3H), 1.00 (d, J = 13.9 Hz, 1H), 0.62 (s, 3H).

^{13}C NMR (100 MHz, $CDCl_3$, 298 K) δ 178.41, 178.04, 170.40, 164.89, 146.10, 134.25, 127.48, 119.58, 118.14, 117.23, 81.75, 70.22, 45.14, 44.46, 43.11, 40.48, 39.43, 38.29, 37.92, 36.46, 35.01, 33.89, 33.60, 31.72, 31.10, 30.88, 30.22, 29.52, 28.74, 21.82, 19.15, 18.47, 10.35.

HRMS (ESI) [M+H]+ calculated for $C_{33}H_{44}NO_3$ 502.3316, found 502.3308.

C1c ^{1}H NMR (400 MHz, $CDCl_3$, 298 K) δ 7.00 (d, J = 6.9 Hz, 2H), 6.52 (s, 1H), 6.33 (d, J = 7.1 Hz, 1H), 6.09 (t, J = 5.1 Hz, 1H), 3.33-3.15 (m, 2H), 2.45 (d, J = 15.0 Hz, 1H), 2.25-2.09 (m, 6H), 2.06-1.76 (m, 6H), 1.75-1.45 (m, 9H), 1.43 (s, 3H), 1.24 (s, 3H), 1.14 (s, 3H), 1.11 (s, 3H), 0.99 (d, J = 13.4 Hz, 1H), 0.63 (s, 3H).

^{13}C NMR (100 MHz, $CDCl_3$, 298 K) δ 178.45, 177.92, 170.36, 164.86, 146.13, 134.13, 127.52, 119.68, 118.14, 117.15, 83.94, 69.48, 45.16, 44.48, 43.12, 40.43, 39.48, 39.33, 38.30, 36.48, 35.17, 33.95, 33.59, 31.71, 31.17, 30.96, 30.24, 29.56, 28.80, 27.76, 21.88, 18.46, 16.60, 10.37.

HRMS (ESI) $[M+H]^+$ calculated for $C_{34}H_{45}NO_3$ 516.3472, found 516.3473.

C2a ^{1}H NMR (400 MHz, $CDCl_3$, 298 K) δ 7.04- 6.96 (m, 2H), 6.54-6.49 (m, 1H), 6.34 (d, J = 7.2 Hz, 1H), 4.63 (dd, J = 15.7, 2.4 Hz, 1H), 4.49 (dd, J = 15.7, 2.4 Hz, 1H), 2.50-2.38 (m, 2H), 2.26-2.18 (m, 4H), 2.17-2.11 (m, 1H), 2.10-1.99 (m, 1H), 1.94-1.78 (m, 3H), 1.76-1.63 (m, 3H), 1.61-1.45 (m, 3H), 1.44 (s, 3H), 1.43-1.33 (m, 1H), 1.26 (s, 3H), 1.20 (s, 3H), 1.09 (s, 3H), 1.03 – 0.93 (m, 1H), 0.56 (s, 3H).

^{13}C NMR (100 MHz, $CDCl_3$, 298 K) δ 178.50, 177.44, 170.08, 164.92, 146.15, 134.18, 127.56, 119.62, 118.27, 117.24, 77.62, 75.11, 52.01, 45.19, 44.36, 43.08, 40.53, 39.59, 38.37, 36.46, 34.80, 33.61, 32.70, 31.69, 30.79, 30.65, 29.96, 29.74, 28.77, 21.78, 18.87, 10.38.

HRMS (ESI) $[M+H]^+$ calculated for $C_{32}H_{41}O_4$ 489.2999, found 489.2997.

2.2.3 Cell culture

HeLa S3, MCF-7 and HEK-293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS. Jurkat cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium with 10% FBS. PC-12 cells were cultured in DMEM supplemented with 5% FBS and 10% horse serum. All mediums were supplemented with 100 U/mL penicillin and 100 μg/mL streptomycin. All cells were cultured in a humidified 37 °C incubator with 5% CO_2.

2.2.4 Cell viability assay

Jurkat cells were seeded into a 96-well plate and incubated for 24 hours in a 37°C incubator for complete adherence. Cel or probe that was dissolved in DMSO was added into RPMI 1640 medium at the final concentration of 10 nM to 50 μM and incubated for another 24 hours. To analyse the cell viability, 20 μL of MTS reagents were pipetted into 100 μL cultured cell samples and incubated for another 2 hours. After that, the absorbance at 490 nm was measured. Wells treated with DMSO and those without cells served as positive and negative groups.

IC50 values were fitted in a dose-response model by using GraphPad Prism 8.0.

2.2.5 Assessment of degradation of IκBα in Jurkat cells

Jurkat cells were pre-treated with 2.5 μM Cel or its probes (C1a, C1b, C1c, and C2a) for 30 minutes (mins), and subsequently stimulated with 50 ng/mL PMA for 60 mins. Then, cells were lysed with lysis buffer (0.2% SDS, 0.1% benzonase and 1 × protease inhibitor). 50 μg extracted proteins were used to perform western blotting analysis with IκBα antibody (SC-4094, Santa Cruz).

2.2.6 Proteome labelling and in-gel fluorescence visualisation

Live cells with 80−90% confluency in a 6-well plate pre-treated with DMEM-containing Cel was then treated with probes for 30 mins at 37 °C in a 5% CO_2 incubator. All cell pellets were lysed and subjected to a click chemistry reaction with 50 μM TAMRA-azide, 0.1 mM TBTA, 1 mM TCEP, and 1 mM $CuSO_4$. After incubation for 1.5 hours at RT, the reaction was terminated by adding pre-chilled acetone. Precipitated proteins were subsequently collected by centrifugation and

loaded in SDS-PAGE to be visualized by in-gel fluorescence scanning. Gels were scanned by Typhoon 9400 under Cy3 and Cy5 channels for fluorescence bands and protein markers, respectively. All figures were applied with a 100-pixel image scanning model.

2.2.7 SILAC pull-down assay

HeLa S3 cells were lysed and subjected to a click reaction with biotin-azide at a final concentration of 100 μM for 2 hours. For SILAC pull-down assays, HeLa S3 cells were exposed to isotopically labelled amino acids. $^{13}C6$ $^{15}N2$ L-lysine-HCl and $^{13}C6$ $^{15}N4$ L-arginine-HCl labelled cells as heavy cells (H), L-lysine monohydrate and L-arginine monohydrochloride labelled cells as light cells (L). These heavy or light cells with 80−90% confluency in a 100 mm cell culture dish were treated with or without Cel followed by C1a treatment. Then, H and L cells were pooled together with the ratio of H/L=1. After click reaction, samples were mixed with 5-fold of cold acetone and stored overnight (-20 °C). Precipitated proteins were washed twice with 5 mL cold methanol, re-suspended in 1 mL PBS (pH 7.5) containing 1% SDS, and desalted by passing through a NAP-10 column (GE). Labelled proteins

were then immobilised with magnetic streptavidin beads (Thermo Fisher Scientific). After washing with freshly prepared 6 M urea, beads were rinsed with digestion buffer containing 20 mM Tris-HCl, pH 8.0, and boiled in SDS loading buffer for immune-blotting. For trypsin digestion, after DTT (5 mM) reduction and IAM (10 mM) alkylation, samples were digested with trypsin overnight. Digestion was quenched by adding 1% formic acid. The peptides were desalted with home-made C18 stage-tip and eluted with 50% acetonitrile and 0.1% formic acid. The elution was concentrated and then re-suspended in 0.1% formic acid for LC-MS/MS analysis.

2.2.8 Immunostaining assay of C1a and COMT in live cells

Live HeLa S3 cells were seeded into glass bottom dishes. After reaching 30% confluency, cells were treated with 0.5 mL DMEM with probes at the final concentrations of 5 μM. Those treated with DMSO served as negative control. After incubation for 1 hour at 37 °C, the medium was removed and cells were washed with 0.5 mL 1x PBS twice, followed by fixation at room temperature with 3.7% formaldehyde in PBS for 1 hour, and permeabilized with 0.2% Triton X-100 in PBS for

10 minutes. Then, cells were blocked with blocking buffer for 30 minutes followed by click reaction for 2 hours at room temperature with vigorous shaking. Subsequently, cells were washed with PBS once and 0.1% Tween-20 in PBS for three times. 1/100 dilution of COMT antibody (Santa Cruz) was incubated with cells at room temperature for 1 hour, followed by incubation with secondary antibody at 1:200 dilution. Finally, the cells were stained with DAPI (1:10000 dilution in 1 × PBS) for 10 minutes at room temperature in dark prior to imaging. Images were collected by using the Leica TCS SPE confocal microscope system equipped with Leica 40×/1.15 oil objective, 405/488/561 nm diode laser, white laser (470−670 nm, with 1 nm increments) and a photomultiplier tube (PMT) detector ranging from 410 to 700 nm for steady state fluorescence.

2.3 Results and discussion

2.3.1 Design and optimization of probes

First, we developed “clickable” probes that retained the bioactivity of Cel and could bind to its cellular targets in living cells. We synthesised

four probes by derivatizing the carboxyl terminus of Cel with terminal alkyne groups using various linkers (**Figure 2-1** and **Figure 2-2**).

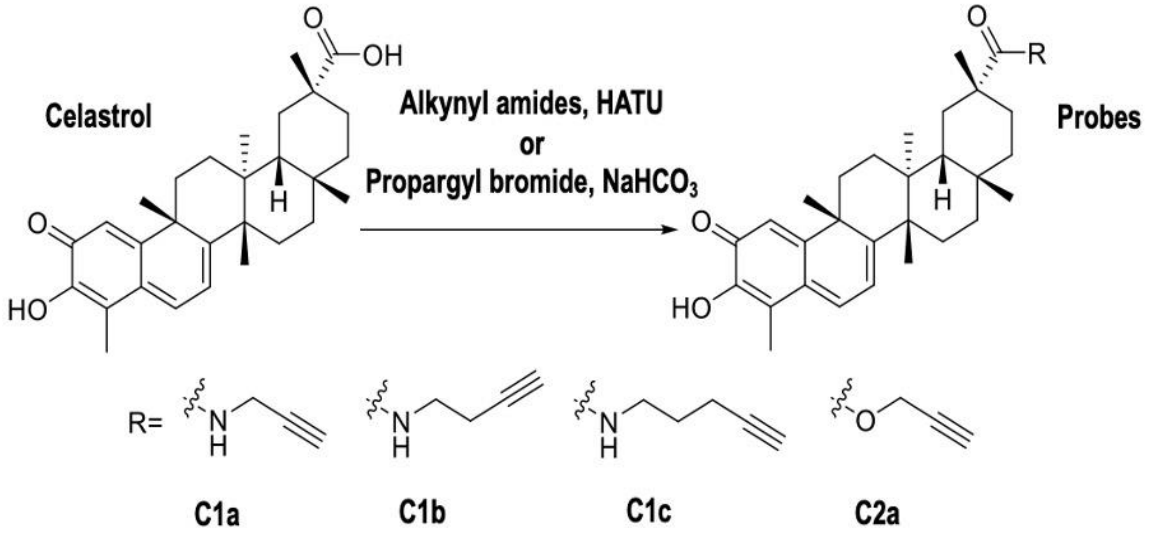
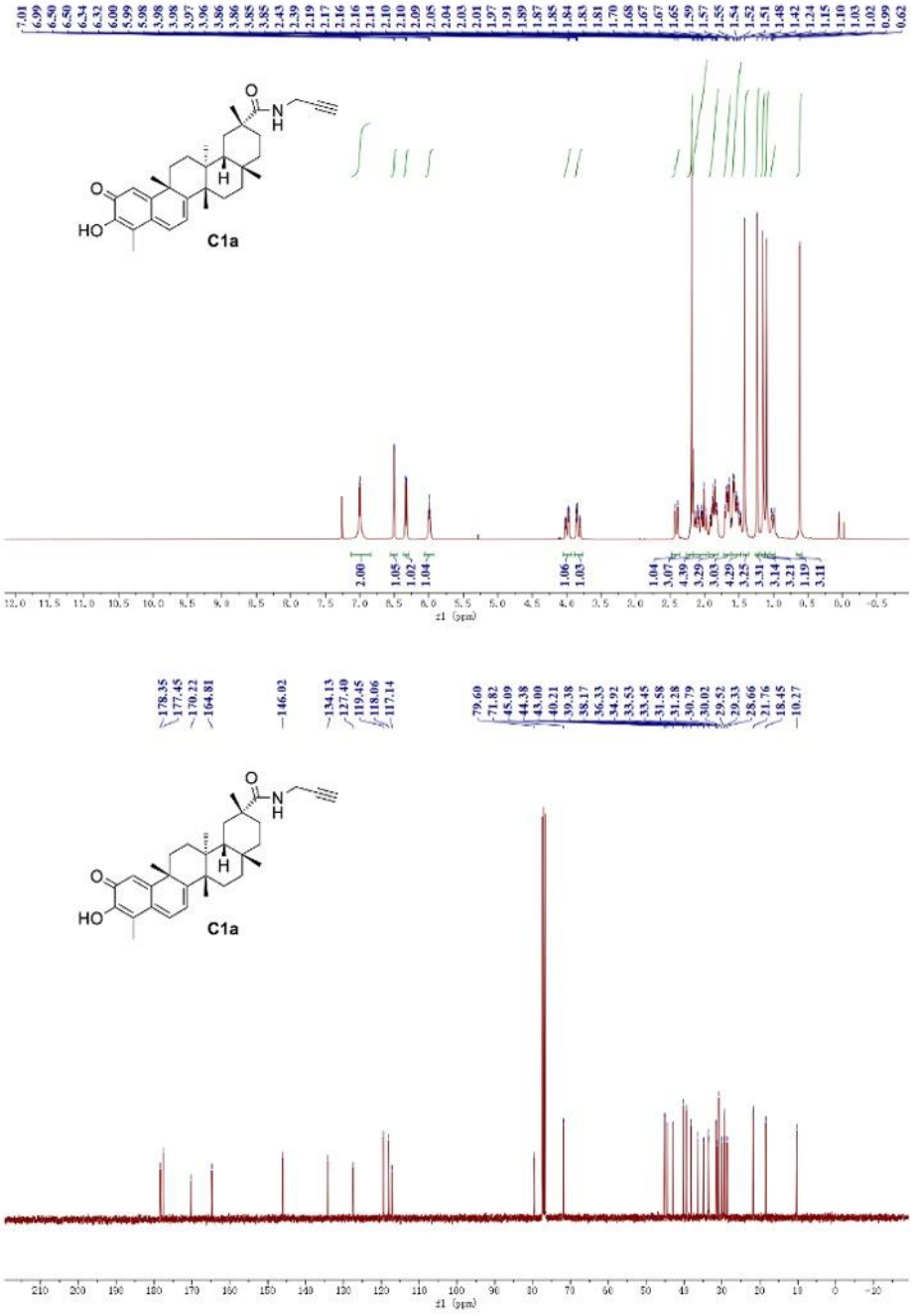
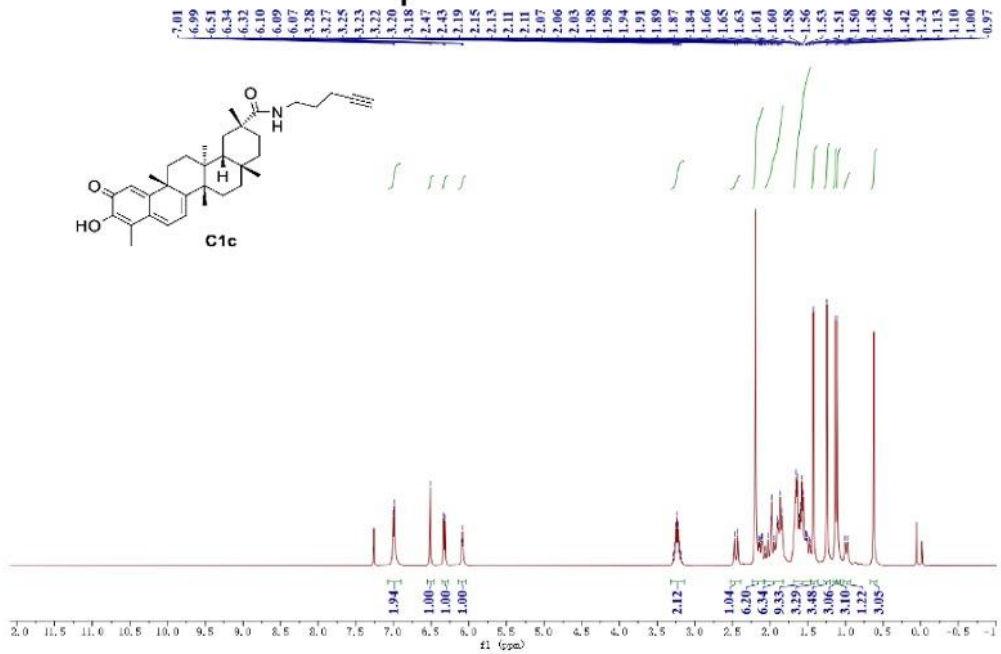


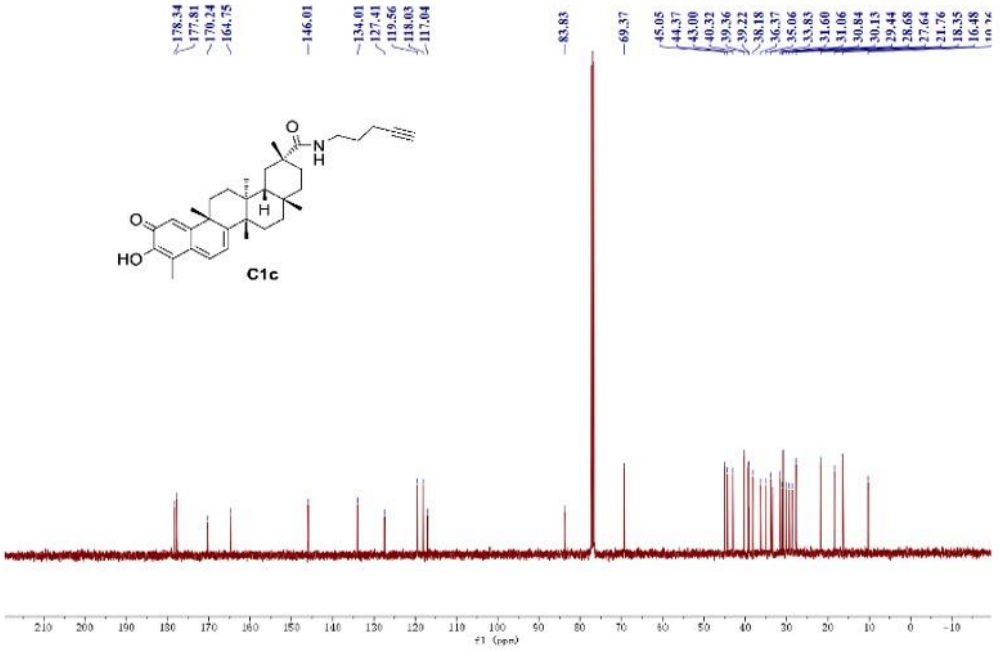
Figure 2-1 Structures and simplified synthesis routes of chemical probes.

A ^{1}H and ^{13}C NMR spectra of **C1a**

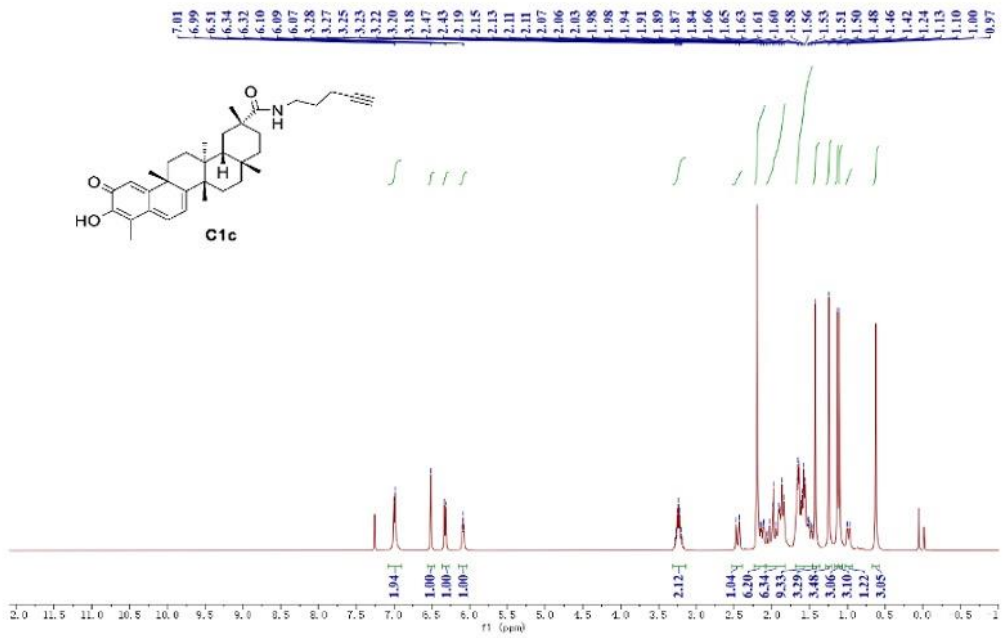


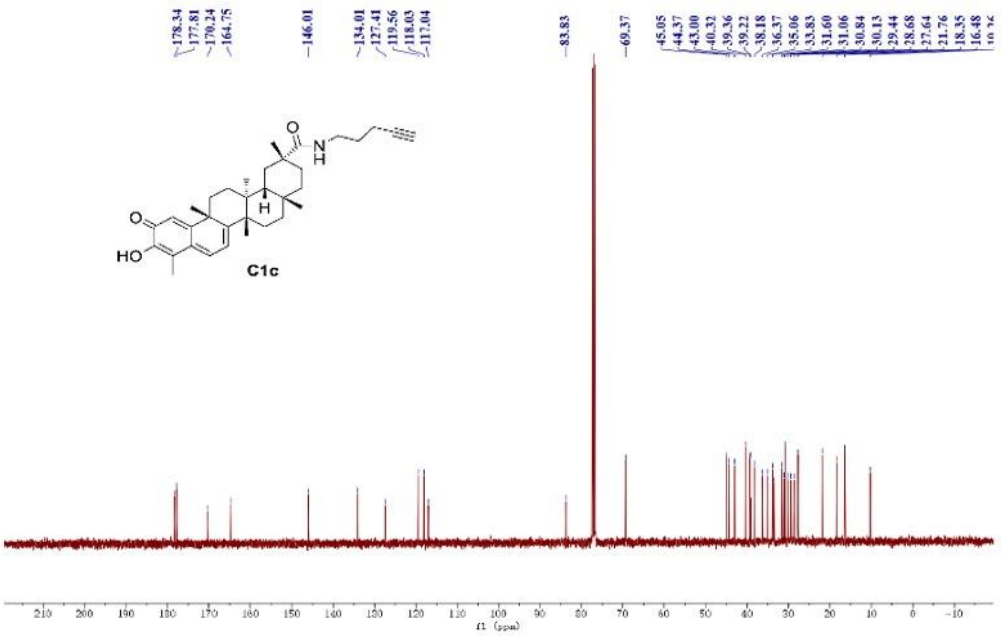
B ^{1}H and ^{13}C NMR spectra of **C1b**





C ^{1}H and ^{13}C NMR spectra of **C1c**





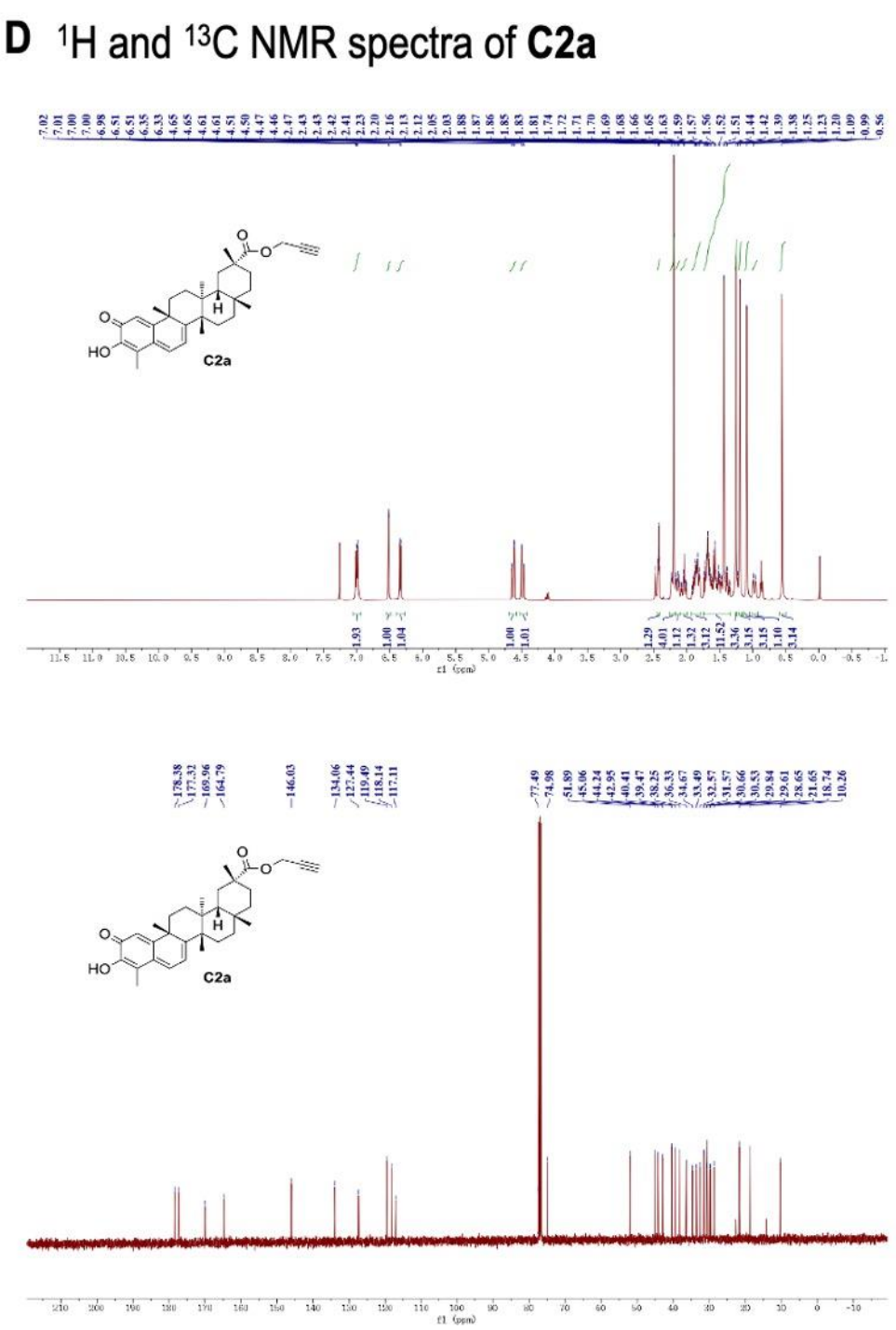


Figure 2-2 The ^{1}H and ^{13}C NMR spectra of C1a, C1b, C1c and C2a probes.

Cel is well-known for reducing inflammation by inhibiting the NF-κB pathway, which is characterised by IκBα degradation[110]. This

experiment revealed that probes C1a, C1b and C1c were presented to inhibit PMA-induced IκBα degradation, confirming their anti-inflammatory properties (**Figure 2-3**).

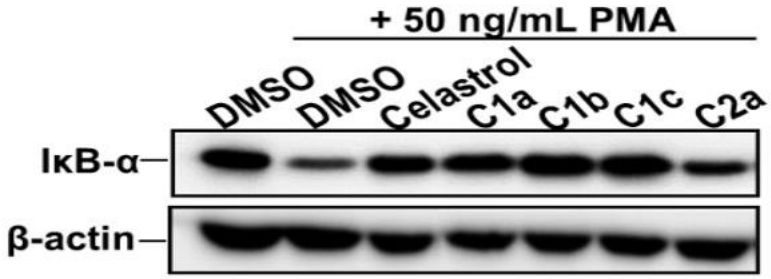


Figure 2-3 Western blotting of PMA-induced degradation of IκBα as a hallmark of inflammation to test the effects of chemical probes in Jurkat cells.

Next, the half maximal inhibitory concentrations (IC50) of the designed probes were tested. The IC50 values of probes in Jurkat cells varied from 1.17 μM to 3.11 μM when Cel and probes were employed to inhibit cell viability, which were equivalent to that of Cel (9.51 μM) (**Figure 2-4**).

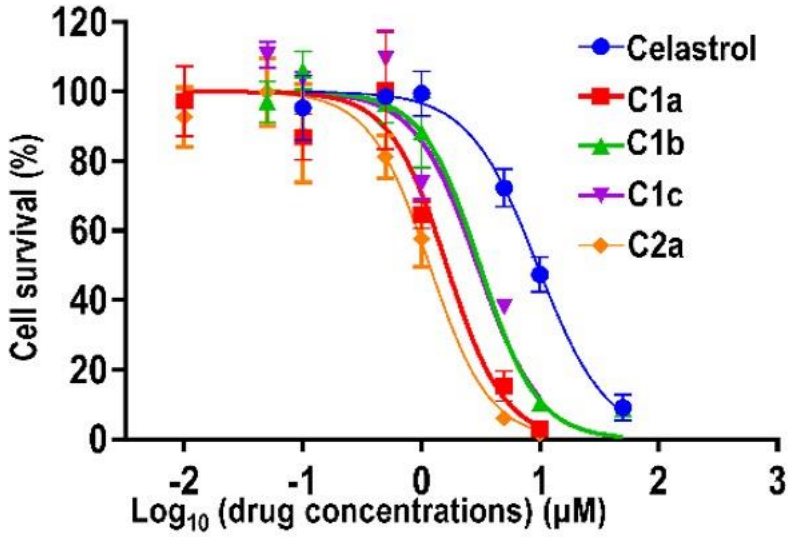


Figure 2-4 Cell survival of Jurkat cells after 24 hours of incubation with Cel or indicated probes. IC50 of Cel, C1a, C1b, C1c and C2a are 9.51±0.89 μM, 1.64±0.37 μM, 3.11±0.84 μM, 2.97±0.35 μM and 1.17±0.18 μM, respectively.

Then, we compared the four probes on the labelling efficiency and specificity for the target proteins. Live Jurkat cells were treated with vehicle or excess Cel before labelled with the appropriate probe. C1a had the highest intensity in the fluorescence scan, and Cel was able to compete with many proteins that were labelled by C1a, demonstrating the highest labelling efficiency of C1a (**Figure 2-5**).

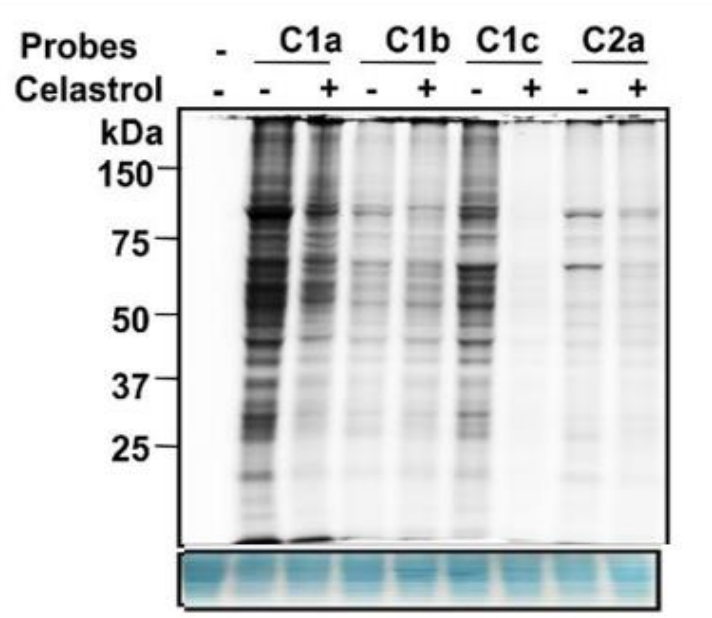


Figure 2-5 Representative in-gel fluorescence scan of proteins labelled by probes in Jurkat cells. Jurkat cells with or without pre-treatment of Cel (20 μM, 1 hour) were incubated with 2 μM C1a, C1b, C1c or C2a probes for 30 mins at 37 °C. Those without the probe treatment in the absence of Cel served as negative groups. In-gel fluorescence scanning (upper) and Coomassie blue gel staining (lower) were carried out following the click reaction.

We assessed three cell lines since a considerable background was observed in probe-labelled Jurkat cells. The lowest background was displayed in HeLa S3 cells, which also had a notable protein band labelled at around 30 kDa (**Figure 2-6**). Additional, Cel virtually

abolished the probe labelling of this protein band in all three cell lines, indicating great specificity and high affinity of the interaction between Cel and this protein (shown by the red arrow). Hence, we decided to carry out the subsequent target profiling using HeLa S3 cells and the C1a probe.

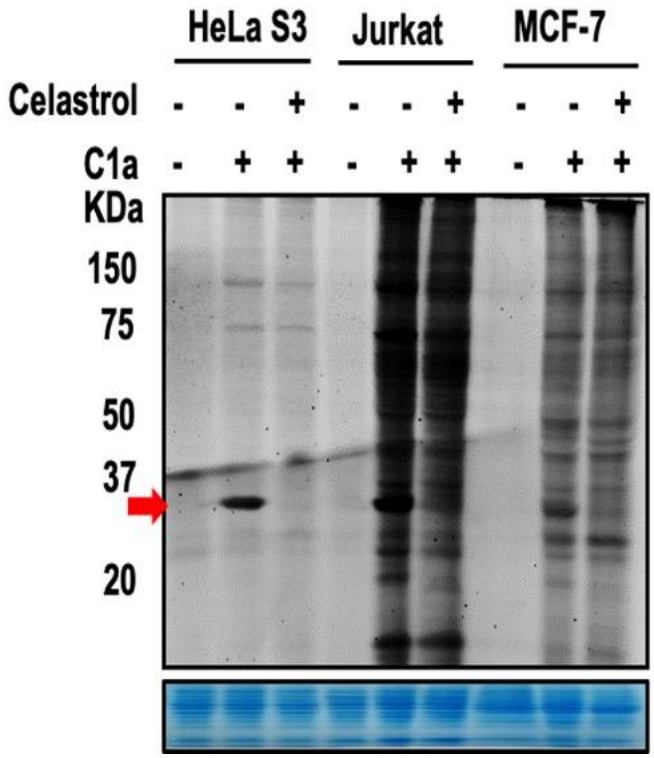


Figure 2-6 In-gel fluorescence scan of proteins labelled with 2 μM C1a probe after 20 μM Cel treatment in different cell lines. The red arrow indicates the potential C1a-labelled protein target.

2.3.2 SILAC MS reveals COMT as a target of Cel

Later, we performed target profiling of Cel in HeLa S3 cells using the SILAC approach. In the forward SILAC, heavy isotope-labelled cells (Heavy, H) were treated with C1a and regular cells (Light, L) were treated with Cel followed by C1a. In the reciprocal experiment, the treatment was regarded as the reverse SILAC. After treatment, cell lysates were collected and coupled to biotin-azide through CuAAC. The probe-labelled proteins were enriched with streptavidin magnetic beads, subsequently digested, and submitted to LC-MS/MS analysis. The pull-down MS results revealed that COMT was discovered to be one of the significantly highly enriched proteins with MS intensities over 1E8. Meanwhile, the intensity of COMT decreased more than 10-fold under Cel treatment, pointing to a significant interaction between COMT and the parent compound Cel. Such findings remained consistent for both forward and reserve SILAC experiments (**Figure 2-7** and **Table 2-1**).

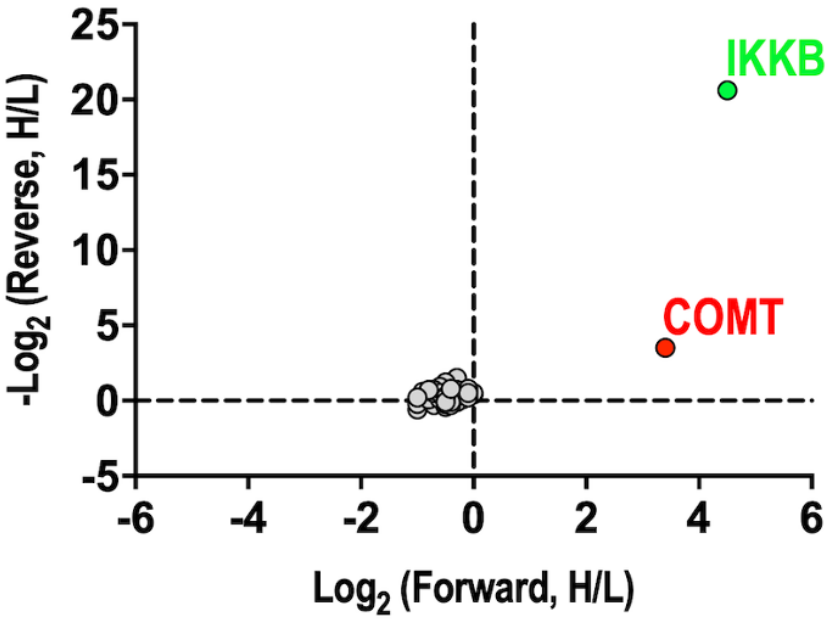


Figure 2-7 SILAC ratios (H/L) of each quantified protein. Forward: heavy cells treated with 2 μM C1a probe, light cells treated with 20 μM Cel followed by C1a probe labelling. Reverse: light cells treated with C1a probe, heavy cells treated with Cel followed by C1a probe labelling.

Table 2-1 Proteins identified and quantified by SILAC. Proteins with intensity over 1E8 in probe-enriched samples are included.

Protein IDs	*Gene names*	*Unique peptides Forward*	*Unique peptides Reverse*	*Mol. weight [kDa]*	*H/L Ratio (Probe/Probe+Cel) Forward*	*H/L Ratio (Probe+Cel/Probe) Reverse*	*Intensity Log_{10}Probe Average*
Q71DI3	HIST2H3A	3	3	15.388	0.63	0.80	9.88
Q5VTE0	EEF1A1P5	8	7	50.184	0.77	0.76	9.58
P40429	RPL13A	8	8	23.577	0.90	0.73	9.34
P31327	CPS1	46	44	164.94	0.64	0.87	9.32
P0DMV9	HSPA1B	9	10	70.051	0.83	0.82	9.26
Q99880	HIST1H2BL	2	2	13.952	0.77	0.84	9.20
P09651	HNRNPA1	8	9	38.746	0.65	0.69	9.10
P04083	ANXA1	18	16	38.714	0.72	0.70	8.94
P07355;A6NMY6	ANXA2;ANXA2P2	18	17	38.604	0.45	0.78	8.91
P62987	UBA52	3	3	14.728	0.70	1.00	8.84
Q9NZM1;O75923	MYOF	43	43	234.71	0.73	0.84	8.83
P63244	GNB2L1	13	13	35.076	0.64	0.70	8.82
P26447	S100A4	3	3	11.728	0.80	0.66	8.81
O15427	SLC16A3	9	10	49.469	0.81	1.06	8.78
Q16719	KYNU	17	15	52.351	1.03	0.73	8.73
P62854;Q5JNZ5	RPS26;RPS26P11	2	2	13.015	0.78	0.81	8.69
P41091;Q2VIR3	EIF2S3;EIF2S3L	14	12	51.109	0.81	0.83	8.65
P05023	ATP1A1	22	18	112.89	0.81	0.74	8.65
P21964	COMT	11	7	30.037	10.56	0.09	8.63

Protein IDs	*Gene names*	*Unique peptides Forward*	*Unique peptides Reverse*	*Mol. weight [kDa]*	*H/L Ratio (Probe/Probe+Cel) Forward*	*H/L Ratio (Probe+Cel/Probe) Reverse*	*Intensity Log_{10}Probe Average*
Q16658	FSCN1	13	10	54.529	0.73	0.88	8.62
P06744	GPI	12	10	63.146	0.94	0.75	8.60
P00966	ASS1	13	8	46.53	0.80	0.67	8.58
P31939	ATIC	11	6	64.615	0.56	0.01	8.58
P15121;C9JRZ8	AKR1B1	10	8	35.853	0.82	0.64	8.56
P06703	S100A6	3	2	10.18	0.72	0.68	8.54
Q96N66	MBOAT7	7	5	52.764	0.64	0.75	8.52
P52209	PGD	11	8	53.139	0.81	0.36	8.51
Q9UBM7	DHCR7	7	6	54.489	0.65	0.69	8.50
P36952	SERPINB5	14	11	42.1	0.87	0.85	8.48
Q9H0U4;Q92928	RAB1B;RAB1C	3	2	22.171	0.74	0.79	8.47
P49748	ACADVL	13	16	70.389	0.77	0.68	8.47
Q92616	GCN1L1	38	34	292.75	0.63	1.27	8.46
P49588	AARS	18	17	106.81	0.74	0.84	8.44
P63241;Q6IS14	EIF5A;EIF5AL1	5	6	16.832	0.71	0.89	8.43
P00505	GOT2	12	12	47.517	0.62	0.86	8.42
P46977	STT3A	9	10	80.529	0.76	0.71	8.41
P27695	APEX1	9	6	35.554	0.72	0.58	8.41
P09429;B2RPK0	HMGB1;HMGB1P1	4	5	24.893	0.82	0.89	8.41
P05388	RPLP0	10	6	34.273	0.80	0.83	8.40
O14975	SLC27A2	18	12	70.311	0.72	0.42	8.40
Q99613;B5ME19	EIF3C;EIF3CL	12	13	105.34	0.89	0.79	8.38
P61619;Q9H9S3	SEC61A1	4	5	52.264	0.72	1.29	8.38
Q8TCJ2	STT3B	11	10	93.673	1.01	0.72	8.37
P30086	PEBP1	7	8	21.057	0.91	0.65	8.35

Protein IDs	*Gene names*	*Unique peptides Forward*	*Unique peptides Reverse*	*Mol. weight [kDa]*	*H/L Ratio (Probe/Probe+Cel) Forward*	*H/L Ratio (Probe+Cel/Probe) Reverse*	*Intensity Log_{10}Probe Average*
Q9Y6C9	MTCH2	7	5	33.331	0.49	1.47	8.34
Q8NBJ5	COLGALT1	12	14	71.635	0.86	0.90	8.33
O43390	HNRNPR	10	7	70.942	0.75	0.88	8.33
Q9Y6K5	OAS3	14	14	121.17	0.48	1.15	8.32
Q8TCT9	HM13	5	5	41.488	0.80	0.64	8.31
O00571;O15523	DDX3X;DDX3Y	15	9	73.243	0.69	0.58	8.31
Q9UNX3;P61254	RPL26L1;RPL26	5	4	17.256	0.60	0.66	8.30
P42330;P17516	AKR1C3	5	3	36.853	0.86	0.87	8.28
Q16881	TXNRD1	11	9	70.905	0.80	1.07	8.27
P11413	G6PD	10	12	59.256	0.58	0.91	8.25
B0I1T2	MYO1G	16	13	116.44	0.64	0.93	8.25
P04792	HSPB1	6	6	22.782	0.68	0.62	8.25
O14920	IKBKB	12	10	86.563	22.67	0.00	8.25
P40925	MDH1	9	8	36.426	0.65	0.78	8.24
Q9Y617	PSAT1	11	8	40.422	0.79	0.81	8.24
Q9BW60	ELOVL1	4	3	32.662	0.75	0.74	8.24
O43776	NARS	10	10	62.942	0.66	0.52	8.22
P35610	SOAT1	9	12	64.734	0.53	0.65	8.22
Q14692	BMS1	14	11	145.81	0.93	0.59	8.20
P51149	RAB7A	9	8	23.489	0.75	0.84	8.19
Q09161	NCBP1	9	10	91.838	0.60	0.64	8.18
P30043	BLVRB	4	4	22.119	0.74	0.67	8.17
P51659	HSD17B4	9	6	79.685	0.64	0.71	8.16
Q01650;Q9UM01	SLC7A5	2	4	55.01	0.73	1.13	8.16
Q92973;O14787	TNPO1	12	12	102.35	0.74	1.23	8.16

Protein IDs	*Gene names*	*Unique peptides Forward*	*Unique peptides Reverse*	*Mol. weight [kDa]*	*H/L Ratio (Probe/Probe+Cel) Forward*	*H/L Ratio (Probe+Cel/Probe) Reverse*	*Intensity Log_{10}Probe Average*
P07305	H1F0	5	4	20.863	0.78	1.10	8.14
Q96QD8	SLC38A2	4	2	56.025	0.96	0.90	8.11
O00767	SCD	5	5	41.522	1.44	1.91	8.11
O75131	CPNE3	5	5	60.13	0.80	0.54	8.11
P30044	PRDX5	5	5	22.086	0.71	0.79	8.10
Q9P0L0	VAPA	4	7	27.893	0.56	0.81	8.10
Q16666;Q6K0P9	IFI16	8	11	88.255	0.69	0.88	8.09
P09960	LTA4H	8	12	69.284	0.79	0.83	8.09
P08559;P29803	PDHA1	5	6	43.295	0.72	0.69	8.08
P35237	SERPINB6	7	7	42.621	0.58	0.63	8.07
P31949	S100A11	3	2	11.74	0.88	0.94	8.07
P84077;P61204	ARF1;ARF3	2	2	20.697	0.85	0.74	8.07
P00367;P49448	GLUD1;GLUD2	9	8	61.397	0.68	0.80	8.07
Q96G23	CERS2	5	5	44.876	0.70	1.09	8.07
P23634;Q16720	ATP2B4	4	2	137.92	0.50	0.87	8.07
O75874	IDH1	9	7	46.659	0.91	0.72	8.06
P42765	ACAA2	6	6	41.924	0.53	0.68	8.06
P51572	BCAP31	4	5	27.991	0.57	0.54	8.06
P14324	FDPS	4	5	48.275	0.74	0.63	8.05
Q14573;Q14571	ITPR3	12	12	304.1	0.76	0.58	8.05
O14818;Q8TAA3	PSMA7;PSMA8	7	5	27.887	1.00	1.21	8.03
Q9BUN8	DERL1	4	5	28.8	0.44	0.78	8.03
P24539	ATP5F1	7	4	28.908	0.76	0.90	8.02
O15258	RER1	4	4	22.958	0.78	0.95	8.01
Q969Q0;P83881	RPL36AL;RPL36A	2	2	12.469	0.76	0.55	8.00

Protein IDs	*Gene names*	*Unique peptides Forward*	*Unique peptides Reverse*	*Mol. weight [kDa]*	*H/L Ratio (Probe/Probe+Cel) Forward*	*H/L Ratio (Probe+Cel/Probe) Reverse*	*Intensity Log_{10}Probe Average*
P21291	CSRP1	4	5	20.567	1.61	1.26	8.00
Q9Y6N5	SQRDL	5	6	49.96	0.60	0.86	8.00
Q5T4S7	UBR4	18	15	573.83	0.74	0.82	8.00
P14550	AKR1A1	2	3	36.573	0.78	0.71	7.99
Q15629	TRAM1	5	5	43.071	0.67	1.00	7.99
P49755	TMED10	5	5	24.976	0.52	1.07	7.99
P06737	PYGL	14	11	97.147	0.69	0.35	7.99
P61019;Q8WUD1	RAB2A	6	7	23.545	0.58	0.44	7.99
P02786	TFRC	6	4	84.87	1.29	0.41	7.99
Q15392	DHCR24	6	4	60.101	0.58	0.50	7.99
P08237	PFKM	10	7	85.182	0.51	0.60	7.98
Q00796	SORD	4	4	38.324	0.86	0.64	7.97
P61803	DAD1	3	3	12.497	0.96	1.12	7.97
P43490	NAMPT	10	8	55.52	0.74	0.82	7.97
Q03519;Q9NUT2	TAP2	5	8	75.663	1.01	1.08	7.97
O43592	XPOT	12	13	109.96	0.49	1.25	7.97
P00491	PNP	7	9	32.118	0.62	0.88	7.96
P37837	TALDO1	7	7	37.54	1.19	0.81	7.96

Notes: For forward SILAC experiment, the heavy cells were treated with 2 μM probe and the light cells were treated with 20 μM Cel followed by 2 μM probe. For reverse SILAC experiment, the heavy cells were treated with 20 μM Cel followed by 2 μM probe while the light cells were treated with 2 μM probe.

In addition, we carried out three target identification experiments utilising LFQ strategy separately. Despite the diverse enrichment methods used in the three LFQ experiments, COMT was consistently identified as the target of Cel among the top-ranked candidates and was the only one with a high degree of reproducibility (**Figure 2-8** and **Table 2-2**). We also found certain known targets of Cel such as inhibitor of nuclear factor kappa-B kinase subunit beta (IKKB) (**Figure 2-7**), as well as other potential novel targets **(Table 2-1 and Table 2-2)**.

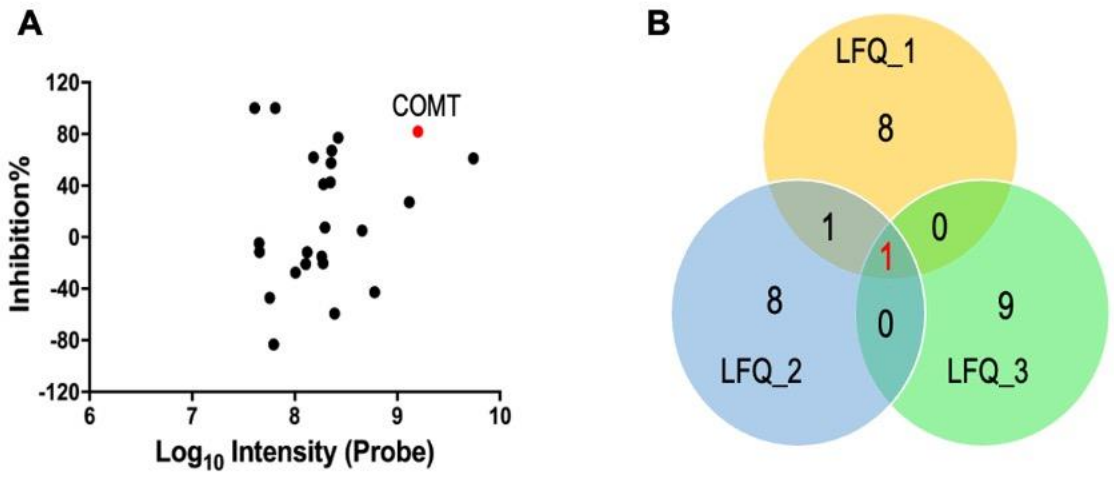


Figure 2-8 Target screening by LFQ proteomics in HeLa S3 cells. (A) Scatter plot of one representative LFQ experiment. (B) Venn diagram of protein targets from three LFQ experiments. For LFQ_1 and LFQ_2, biotin azide (1265-5, Click Chemistry Tools) and DIAZO biotin azide (1240-10, Click Chemistry Tools) were used for click reaction,

respectively, and in-gel tryptic digestion were applied; for LFQ_3, Tev biotin azide (GV-12, GL Biochem) was used for click reaction and on-beads tryptic digestion was applied.

Table 2-2 Protein list of three independent LFQ experiments. Top ten proteins with high intensities are listed.

Protein ID	Protein Name	Gene ID	Mol. weight [kDa]	Intensity DMSO (a)	Intensity Probe (b)	Intensity Cel+Probe (c)	Inhibition% (1.0-c/b)%
LFQ Dataset_1							
P63104	14-3-3 protein zeta/delta	YWHAZ	27.745	3.2E+07	3.4E+08	3.0E+08	10.5
P62258	14-3-3 protein epsilon	YWHAE	29.174	1.2E+07	3.2E+08	3.3E+08	2.0
P21964	Catechol O-methyltransferase	COMT	30.037	0.0E+00	1.3E+08	1.2E+07	90.9
P31946	14-3-3 protein beta/alpha	YWHAB	28.082	3.7E+06	9.4E+07	8.1E+07	13.6
P18669	Phosphoglycerate mutase 1	PGAM1	28.804	0.0E+00	4.6E+07	4.4E+07	2.9
P61981	14-3-3 protein gamma	YWHAG	28.302	0.0E+00	3.1E+07	1.9E+07	37.3
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2	HSD17B10	26.923	7.1E+06	2.9E+07	1.8E+07	39.4
P31947	14-3-3 protein sigma	SFN	27.774	0.0E+00	2.6E+07	1.6E+07	36.7
O00299	Chloride intracellular channel protein 1	CLIC1	26.922	0.0E+00	2.0E+07	1.6E+07	19.7
Q8NCW5	NAD(P)H-hydrate epimerase	APOA1BP	31.674	0.0E+00	1.4E+07	1.0E+07	26.5
LFQ Dataset_2							
P21964	Catechol O-methyltransferase	COMT	30.037	0.0E+00	6.2E+08	0.0E+00	100.0
P27348	14-3-3 protein theta	YWHAQ	27.764	2.6E+08	4.4E+08	9.7E+07	78.0

Q15056	Eukaryotic translation initiation factor 4H	EIF4H	27.385	4.1E+07	1.3E+08	3.4E+07	73.8
O95833	Chloride intracellular channel protein 3	CLIC3	26.648	3.2E+07	5.3E+07	1.2E+07	77.5
Q8NCW5	NAD(P)H-hydrate epimerase	APOA1BP	31.674	2.0E+07	4.3E+07	6.1E+06	85.9
Q04917	14-3-3 protein eta	YWHAH	28.218	2.1E+07	4.1E+07	6.8E+06	83.5
P10768	S-formylglutathione hydrolase	ESD	31.462	6.8E+06	3.4E+07	7.0E+06	79.5
Q9Y3A5	Ribosome maturation protein SBDS	SBDS	28.763	2.3E+07	2.7E+07	2.9E+06	89.4
Q96C19	EF-hand domain-containing protein D2	EFHD2	26.697	3.1E+06	2.7E+07	1.0E+07	61.4
P29218	Inositol monophosphatase 1	IMPA1	30.188	4.7E+06	2.6E+07	0.0E+00	100.0
LFQ Dataset_3							
P28070	Proteasome subunit beta type-4	PSMB4	29.204	1.0E+08	2.7E+08	6.2E+07	76.6
P30048	Thioredoxin-dependent peroxide reductase	PRDX3	27.692	0.0E+00	2.1E+08	8.4E+07	60.4
P21964	Catechol O-methyltransferase	COMT	30.037	0.0E+00	1.6E+08	0.0E+00	100.0
Q9UL46	Proteasome activator complex subunit 2	PSME2	27.401	0.0E+00	5.9E+07	0.0E+00	100.0
Q9P031	Thyroid transcription factor 1-associated protein 26	CCDC59	28.669	0.0E+00	5.1E+07	0.0E+00	100.0

P33316	Deoxyuridine 5-triphosphate nucleotidohydrolase	DUT	26.563	0.0E+00	4.9E+07	0.0E+00	100.0
P28066	Proteasome subunit alpha type-5	PSMA5	26.411	0.0E+00	4.1E+07	0.0E+00	100.0
Q15056	Eukaryotic translation initiation factor 4H	EIF4H	27.385	0.0E+00	3.8E+07	0.0E+00	100.0
P15559	NAD(P)H dehydrogenase [quinone] 1	NQO1	30.867	0.0E+00	3.3E+07	0.0E+00	100.0
Q9BTT0	Acidic leucine-rich nuclear phosphoprotein 32 family member E	ANP32E	30.692	0.0E+00	3.0E+07	1.7E+07	45.2

2.3.3 Target validation of Cel

We verified COMT as a potential target by using multiple independent assays. First, we re-analysed the pull-down products with a targeted MS approach called PRM. Two unique COMT peptides (^{152}MVDFAGVK^{159} and ^{252}EVVDGLEK^{259}) were selected and precisely measured. Both peptides were enriched by C1a and inhibited by 10-fold Cel (**Figure 2-9**).

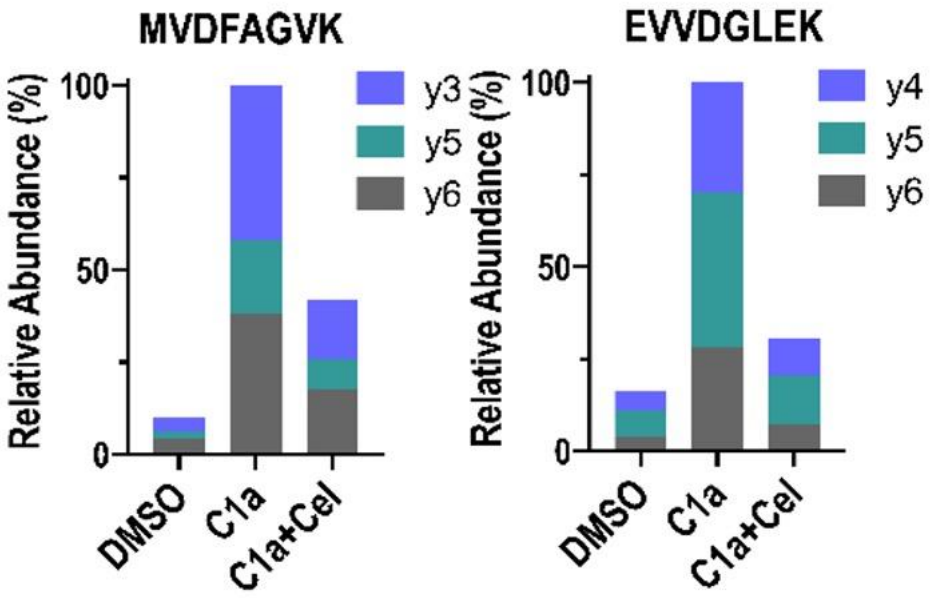


Figure 2-9 Targeted MS analysis of two unique peptides of COMT. Cells were treated with 20 μM Cel followed by 2 μM C1a probe.

The mammalian COMT gene encodes both the membrane bound MB-COMT and the soluble isoform S-COMT. S-COMT is mainly found in

the cytosolic fraction of periphery whereas MB-COMT is thought to have an important role in the central nervous system[111]. In order to distinguish the interaction between Cel and COMTs, we knocked down the expression of endogenous COMTs in HeLa S3 cells and observed the changes by the fluorescence signal of C1a labelling. We found that both expressions of MB-COMT and S-COMT reduced, resulting in decreased probe-labelled fluorescence signals (**Figure 2-10A**). Moreover, a chemical precipitation experiment was performed, which involved identifying COMT in pulldown products using an anti-COMT antibody (**Figure 2-11**). Despite endogenous S-COMT being detectable in western blotting, the probe labelling of C1a in S-COMT was much weaker than that of MB-COMT. This observation was consistent with the results in **Figure 2-10**, where S-COMT was barely detectable with C1a labelling. Next, in HEK293T cells with overexpressed MB-COMT, an additional protein band with a strong fluorescence signal was found, indicating effective labelling by C1a (**Figure 2-10B**). Therefore, we concluded that in living cells, Cel binds to MB-COMT rather than S-COMT.

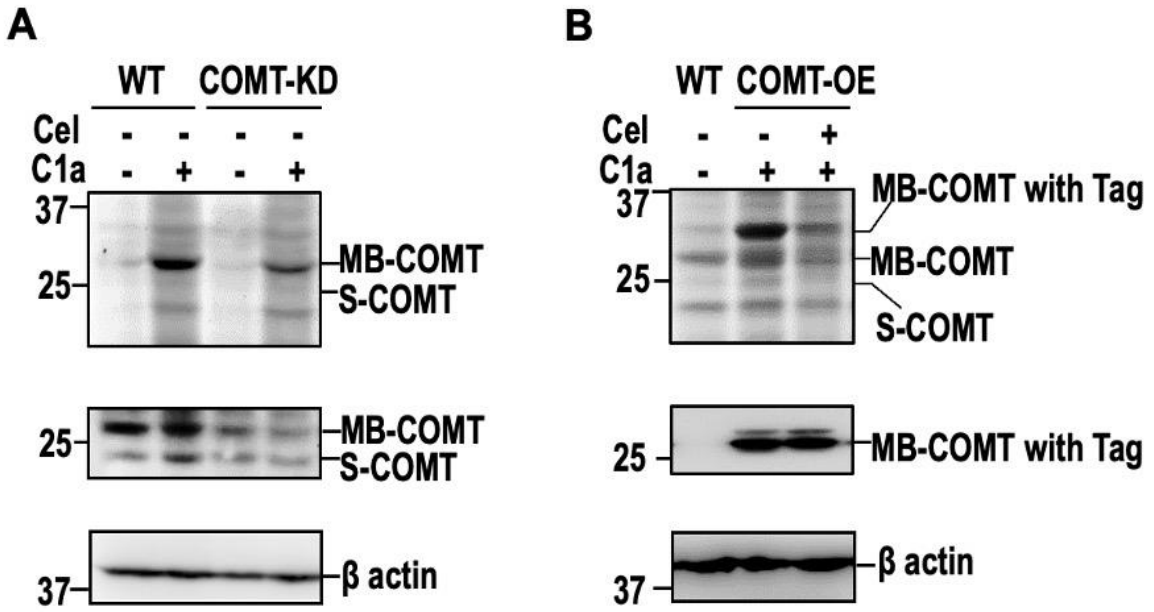


Figure 2-10 Fluorescence labelling (upper) of C1a and western blotting (lower) of COMT. (A) COMT knockdown (KD) in HeLa S3 cells. (B) COMT overexpression (OE) in HEK293T cells.

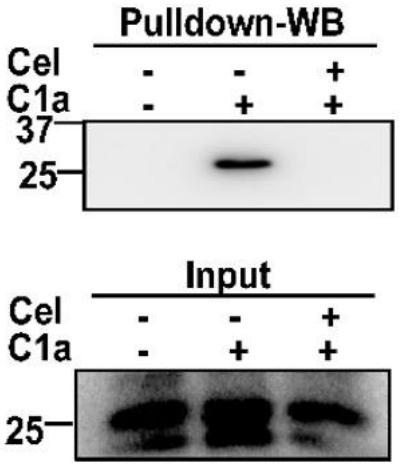


Figure 2-11 Western blotting analysis of pull down products detected by using an anti-COMT antibody. Live HeLa S3 cells were treated with

C1a probe in the presence or absence of Cel, followed by click reaction and chemical precipitation.

2.3.4 Cellular localisation of Cel

We used fluorescent labelling in HeLa cells to visualise the cellular localization of Cel. After immunostaining with the anti-COMT antibody, COMT was recognised with green fluorescence under a confocal microscope. C1a-labelled proteins were labelled as red after click reaction with tetramethylrhodamine (TAMRA)-azide. Co-localisation of Cel and COMT in HeLa cells was shown by the overlay of endogenous COMT (green) and C1a (red) (**Figure 2-12**). According to a prior report[112], MB-COMT is found both in the plasma membrane and endoplasmic reticulum (ER). COMT and the fluorescent signal from C1a were co-localized, and probably in the membrane structures.

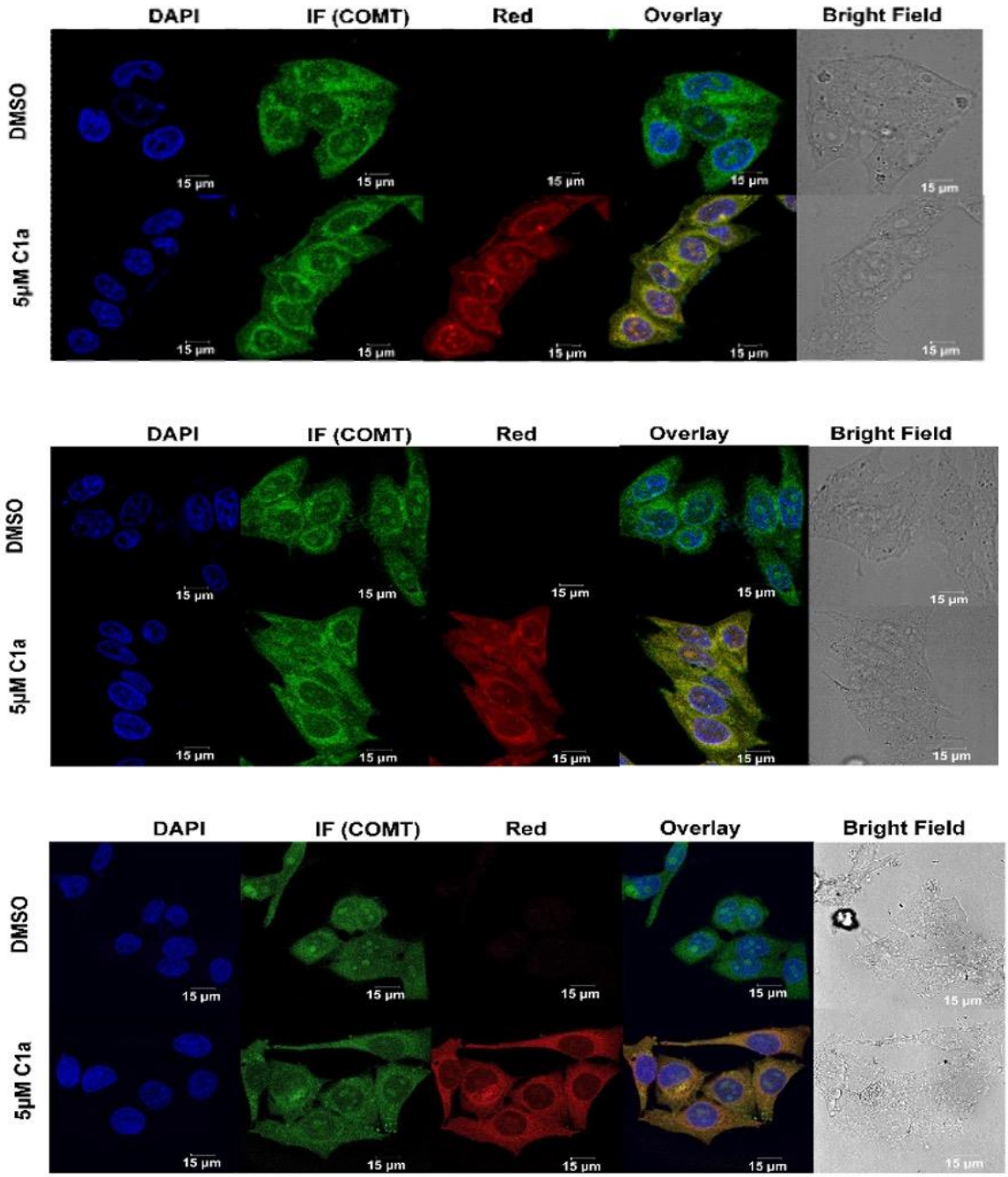


Figure 2-12 Immunofluorescence analysis of endogenous COMT and Cel in HeLa cells. Blue channel: DAPI. Green channel: COMT. Red channel: C1a. Yellow channel: Overlap of COMT and C1a.

2.4 Conclusion

We synthesised chemical probes to mimic Cel and used probe C1a to perform target identification of Cel with a chemical proteomics strategy. Based on the specifically labelled target band in the in-gel fluorescence scan and the MS identification by both SILAC and LFQ, we identified COMT as the major binding target of Cel. Then, we verified the interaction of Cel and COMT by employing a targeted MS approach. We also found that Cel preferentially bound to MB-COMT over S-COMT in live cells. In addition, to visualise the cellular localization of Cel, we found it co-localised with COMT through an immunofluorescence analysis.

Chapter 3: Characterisation of Cel's binding sites on COMT

3.1 Introduction

The electrophilic quinone methide moiety of Cel has been reported to react with nucleophilic amino acid residues to generate covalent bonds via Michael addition[113-115]. The donor in Michael reaction is usually an enolate or other nucleophile and the Michael acceptor is usually an α,β-unsaturated carbonyl. The conjugate addition produces a Michael adduct by forming a carbon-carbon bond at the β-carbon position of the Michael acceptor[116]. As cysteines serve as the most nucleophilic amino acid and the most useful hot spots for irreversible inhibitors, we reasoned that Cel can bind with the cysteine sidechain of COMT by Michael addition[117, 118]. Hence, in this chapter, we utilised a series of assays to further investigate the binding stoichiometry of Cel and COMT.

3.1 Materials and methods

3.1.1 Chemicals and reagents

Ammonium nitrate (NH_4NO_3), Luria-Bertani broth powder, dimethyl sulfoxide (DMSO), Magnesium chloride ($MgCl_2$), all antibodies in the experiments included COMT antibody (ab126618: rabbit source from Abcam and sc-137253: mouse source from Santa Cruz), β-actin antibody (AF7018, Affinity), IκBα antibody (sc-1643, Santa Cruz), 488-FITC antibody (sc-516140, mouse source, Santa Cruz), goat anti-mouse HRP (7076S, CST), and goat anti-rabbit HRP (AS014, ABclonal). All gel casting reagent and protein markers were purchased from Bio-Rad.

3.1.2 The *in vitro* and *in situ* competition assay to validate/characterise Cel-COMT interactions

For the *in vitro* assay, 0.24 μg purified S-COMT in 20 μL incubation buffer (50 mM HEPES, pH 7.5, 150 mM NaCl and 2 mM $MgCl_2$) was incubated with Cel at the concentrations of 50 nM to 50 μM. After overnight incubation at 4 °C, 2 μM C1a probe was added, click reaction was performed for the in-gel fluorescence scan. For the *in situ* assay, HeLa S3 cells pre-treated with Cel (20 nM to 20 μM) for 1 hour were

treated with 2 μM C1a probe for 30 mins. All figures with 100-pixel image scanning model, 400 PMT laser power was applied for scanning.

3.1.3 UV-visible absorption assay of Cel

Cel (50 μM, pH 7.4) was added into 0.25 molar equivalents of wild-type or mutated S-COMT for 30 mins at RT. Cel incubated with PBS and β-mercaptoethanol (β-ME) served as negative and positive control groups, respectively. The absorption of Cel was measured at 440 nm.

3.1.4 In-gel fluorescence assays on wild-type and mutated S-COMT

To determine cysteine-binding sites, 1.0 μg wild-type or mutated S-COMT was incubated with or without 20 μL of 12 μM C1a probe at 4 °C overnight. In-gel fluorescence scan was performed after click reaction. For all figures with a 100-pixel image scanning model, 400 PMT laser power was applied.

3.1.5 Dilution experiment with COMT-Cel adducts

Cel (50 μM, pH 7.4) was incubated with PBS, 15 μM purified S-COMT, or 100 μM β-ME in PBS for 30 mins at RT. Then, these solutions were

10-fold diluted in PBS containing 15 μM purified S-COMT, 100 μM of β-ME, or PBS alone, respectively. The mixtures were incubated at RT for 30 mins and the disassociation of adducts was detected by a UV-visible spectrometer. The absorption of Cel was measured at 440 nm. The assay was performed in triplicate.

3.1.6 ESI-Q-TOF assay on S-COMT

Cel (17.5 μM) was incubated with purified wild-type S-COMT or mutated S-COMT (2.5 μM). Binding assays were pre-carried out in 50 mM HEPES buffer (150 mM NaCl and 2 mM $MgCl_2$, pH 7.5). After 2 hours of incubation at 4 °C, ESI-Q-TOF analysis was subsequently performed on the complex to identify potential cysteine binding sites. Specifically, ESI-Q-TOF assays were performed using an Agilent 6540 quadrupole time-of-flight (Q-TOF) mass spectrometry system (Agilent Technologies, Santa Clara, CA). Samples were separated using an Agilent 1290 UPLC system. The samples could be separated on an analytical column (2.1 × 100 mm) packed with 1.7 μm C4 reversed-phase material and eluted with the following gradient: 5% for 2 mins; 5% - 95% B for 8 mins; 95% B for 2 mins (solvent A: 0.1% formic acid

in water, solvent B: 0.1% formic acid in ACN). The eluent was introduced to the mass spectrometer at a flow rate of 300 μL/min and 40 °C operating temperature. The instrument parameters were as follows: gas temperature 320 °C, drying gas 8 L/min, nebulizer 45 psi and capillary 4.5 kV (0.039 μA). MS analysis was performed in the positive ion mode using a mass range of *m/z* 500–3200. The instrument control software was Agilent mass hunter workstation data acquisition 6.0 software. Raw data were processed using mass hunter qualitative analysis B.07.00 software. MS spectra were deconvoluted using a maximum entropy algorithm.

3.2 Results and discussion

3.2.1 Illustration of the COMT-Cel interaction

Reactive cysteines serve as hot spots for irreversible inhibitors to bind[119, 120]. Through Michael addition[113-115], Cel was reported to covalently react with the thiol groups of cysteine residues. We hypothesised that the cysteine sidechain of COMT reacts with Cel at Cel's C6 position via Michael addition (**Figure 3-1A**). In order to

examine the interactions of COMT and Cel, wild type (WT) S-COMT, WT MB-COMT, and mutant S-COMT were purified (**Table 3-1**).

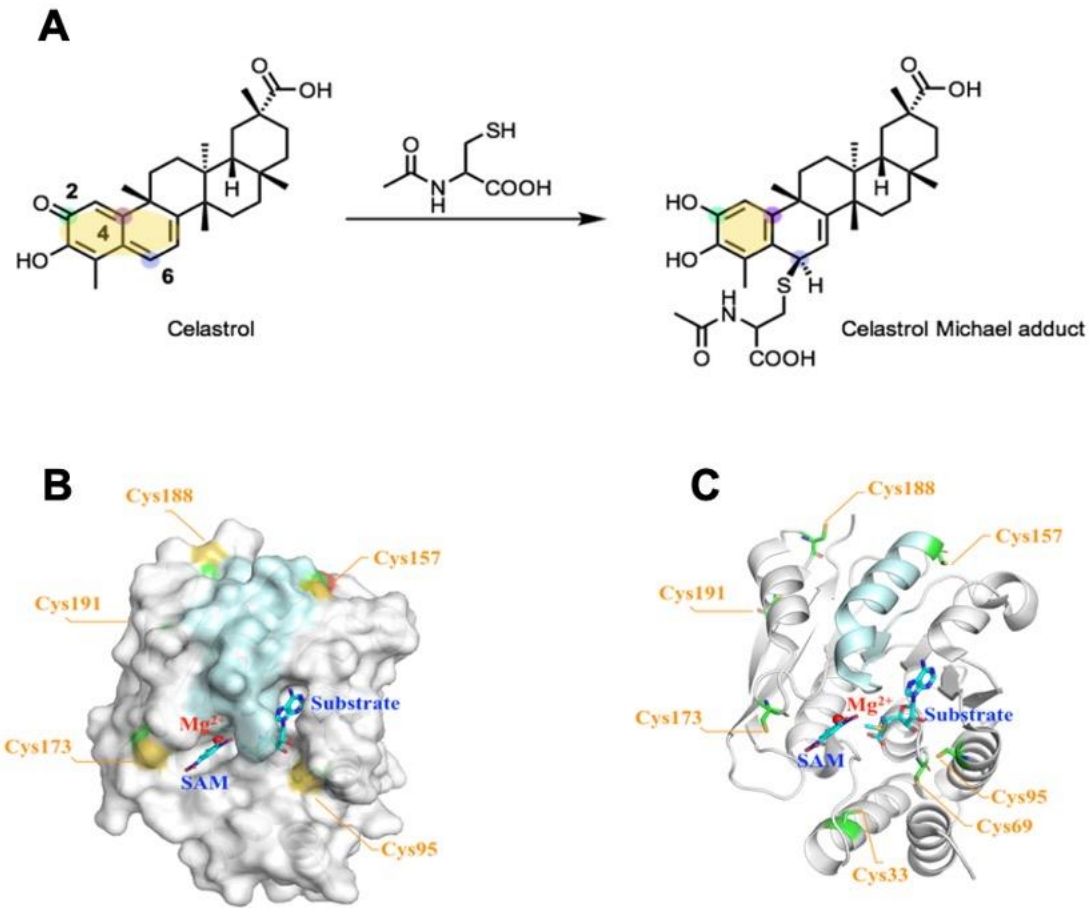


Figure 3-1 Illustration of COMT-Cel interaction. **(**A) Chemical reaction of COMT-Cel interaction. (B) The cartoon diagram of S-COMT shows details of seven cysteine residues. Five cysteines are clearly located on the surface. (C) The SAM and substrate pockets are

separated by a helix (D141-C157, cyan area) which has Cys157 as its terminal residue on the far end.

Table 3-1 Primers involved in the experiment. Mutated nucleotides were highlighted in red area.

Primer names	Primer sequences
COMT-C95A-F	5'- CCCGACGCTGCCGCCATCACCCAG -3'
COMT-C95A-R	5'- CGGCAGCGTCGGGGTTGATCTCGATGG -3'
COMT-C157A-F	5'- GAGGAAGCTGGCCTGCTGCGGAAGGG G -3'
COMT-C157A-R	5'- CAGGCCAGCTTCCTCCAAGAGAAGCGTG -3'
COMT-C173A-F	5'- CGTGATCGCCCCAGGTGCGCCAGAC -3'
COMT-C173A-R	5'- CTGGGGCGATCACGTTGTCAGCCAGTAG -3'
COMT-C188A-F	5'- GCAGCGCCTTTGAGTGCACACACTACCAATC -3'
COMT-C188A-R	5'- CTCAAAGGCGCTGCTCCCGCGCACG -3'
COMT-C191A-F	5'- CTTTGAGGCCACACACTACCAATCGTTCCTGG -3'
COMT-C191A-R	5'- GTGTGTGGCCTCAAAGCAACTACTCCCGCG -3'

Next, we employed a UV-visible spectroscopy approach to monitor the interactions of COMT and Cel. The UV spectrum of Cel exhibits a peak at 440 nm in physiological condition[113]. When the quinone methide of Cel reacts with thiol groups, the resulting formation of a covalent bond causes a decrease in Cel's UV-visible absorption. To further investigate

the exact binding residues, five cysteines (Cys, C) that are exposed on the surface of S-COMT were mutated to alanine (A) to create C95A, C157A, C173A, C188A and C191A (**Figure 3-1B** and **3-1C**). After WT S-COMT or mutants were incubated with Cel, the absorption of free Cel was significantly decreased, suggesting new bond creation during interaction (**Figure 3-2**). Compared to WT S-COMT, mutants of C157A and C95A have lower ability to lower Cel's UV absorption, showing that Cys157 and Cys95 of S-COMT are probable Cel binding sites.

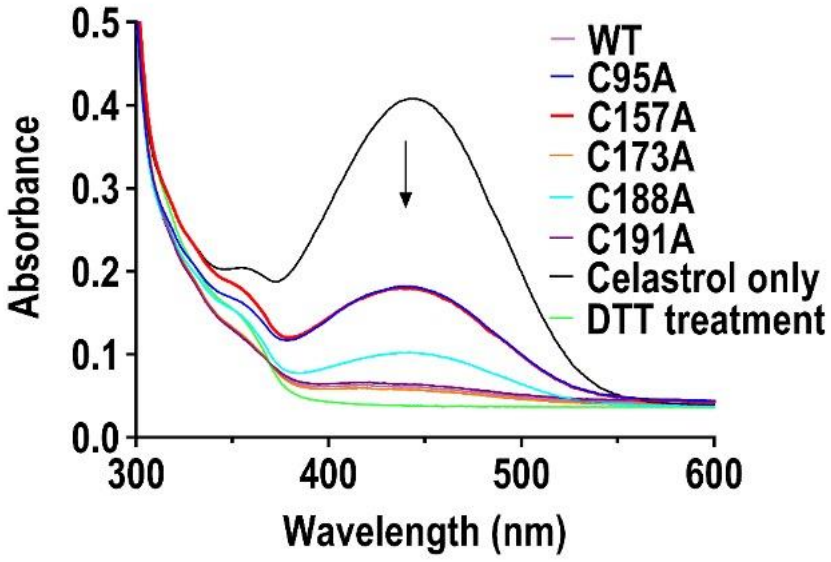


Figure 3-2 UV absorption of 50 μM Cel at 440 nm after incubation with indicated recombinant proteins.

To further characterise the irreversible or reversible covalent bond formation, we performed dilution assays using the Cel-S-COMT adduct or Cel alone in PBS. As a comparison, the reversible adduct formed by Cel and β-Mercaptoethanol (β-ME) was diluted with PBS. Consistent with the earlier report[113], the reversible adduct formed by β-ME and Cel disintegrated when it was diluted by PBS, owing to the equilibrium of the reaction was reversed with the dilution of the reactant (β-ME). In contrast, the UV absorption, an indicator of Cel-COMT adduct formation, remained unchanged regardless of whether the Cel-COMT adduct was diluted in PBS or excess Cel, suggesting that Cel and COMT form an irreversible covalent bond under the tested conditions (**Figure 3-3**). Moreover, we performed an in-gel fluorescence scan in COMT or mutants with C1a labelling, showing a weaker interaction with C1a in C95A and C157A (**Figure 3-4**). The result is consistent with the result in **Figure 3-2**. Notably, C157A displayed the greatest reduction in fluorescence signal of all mutants, indicating Cys157 was Cel's main binding residue.

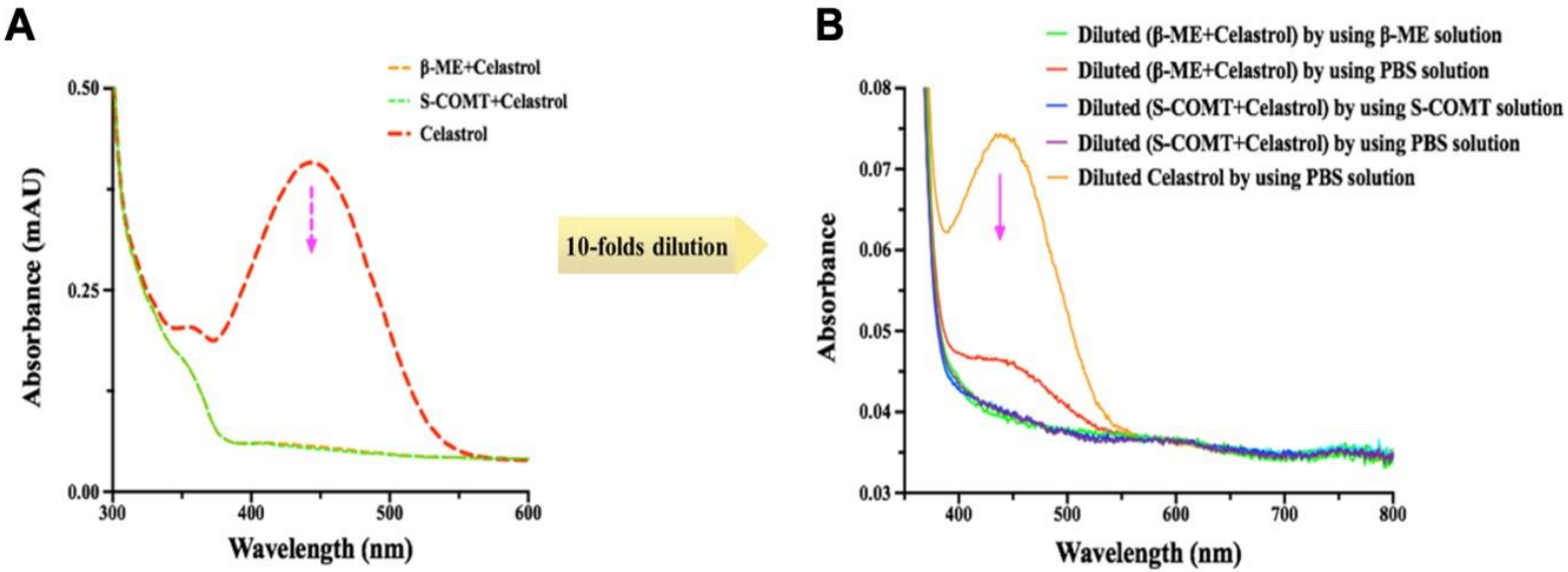


Figure 3-3 Cel irreversibly binds with S-COMT. Dilution experiment with S-COMT-Cel adduct was performed by observing the change in absorption of free Cel at 440 nm. After incubation with S-COMT, the absorbance of Cel was hardly observed and had not re-appeared even in 10-fold dilution with PBS solution, indicating the covalent bonding was irreversible under the tested conditions.

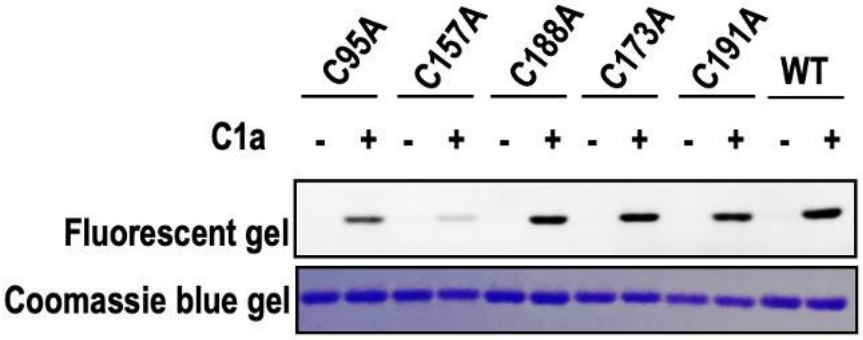


Figure 3-4 In-gel fluorescence scan of recombinant COMT or mutants. The concentration of recombinant COMT or mutants was 1.67 μM while the concentration of C1a was 12 μM. Incubation of proteins and C1a was performed at 4 °C overnight.

3.2.2 Intact protein MS

We next used ESI-Q-TOF MS to examine the binding stoichiometry of Cel and COMT in intact recombinant S-COMT and mutants pre-treated with Cel. Intact protein MS is a direct and powerful tool for the identification of protein-compound interaction via assessing protein's total molecular weight without prior digestion by mass spectrometer[121, 122]. In **Figure 3-5A** and **3-5B**, new peaks appeared with mass shifts of 450.4, 901.1, and 1351.8 Da appeared after Cel treatment, respectively representing the conjugation of one, two, and three Cel molecules.

Then, using five S-COMT mutants (C95A, C157A, C188A, C173A, and C191A), we carried out the same intact protein MS analysis. As shown in **Figure 3-5C** and **3-5D**, both C173A and C191A COMT still bound to Cel molecules. For C95A and C188A, the complex bound to one or two Cel molecules considerably decreased while the complex bound to three Cel completely disappeared (**Figure 3-6**). Remarkably, a single mutation of the residue Cys157 entirely disrupted all interactions between Cel and COMT, again proving that the particular residue Cys157 should serve as the primary binding site (**Figure 3-6**).

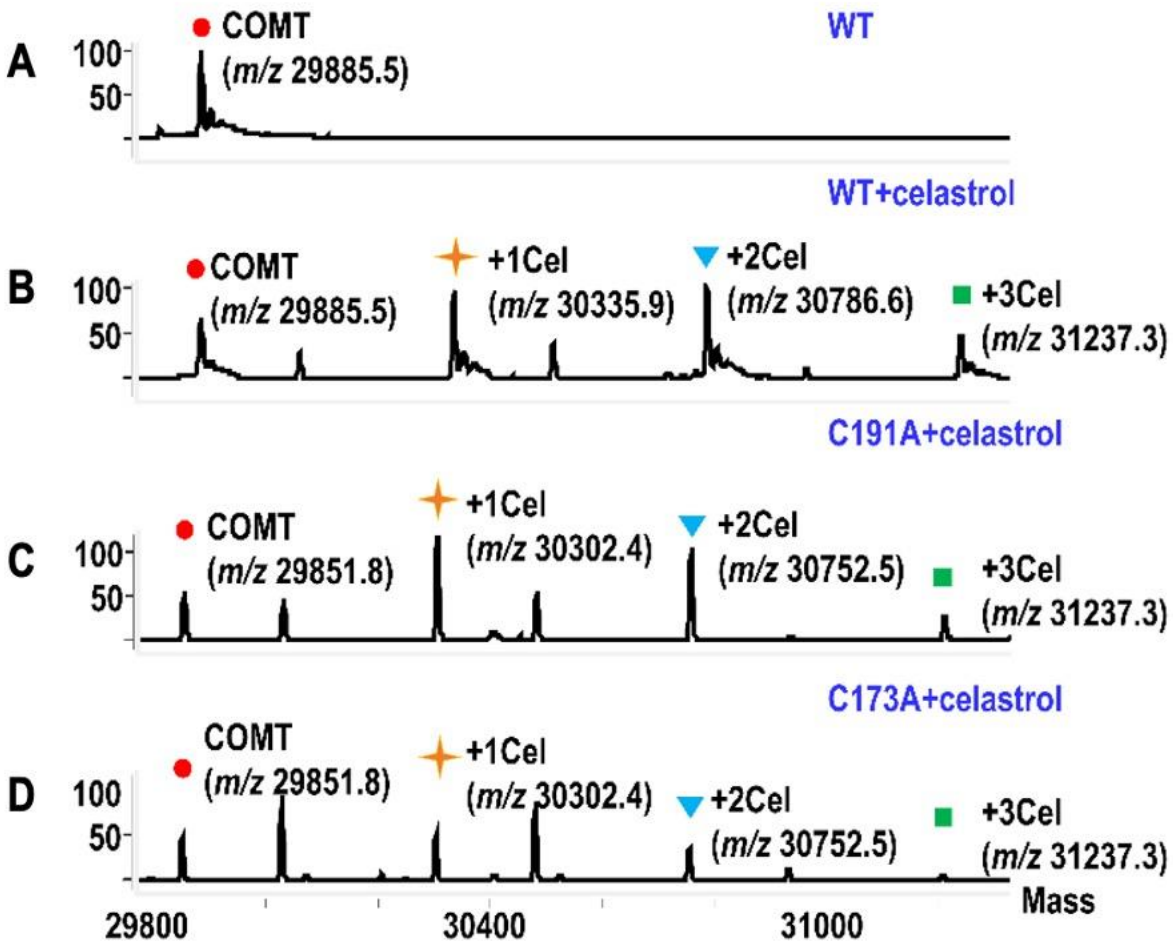


Figure 3-5 ESI-Q-TOF analysis on WT S-COMT or mutated S-COMT (C191A and C173A). S-COMT is marked with the circle. After incubation for 2 hours at 4 °C with Cel, the peaks indicating S-COMT with one Cel, two Cel and three Cel are labelled as asterisks, triangles, and squares, respectively.

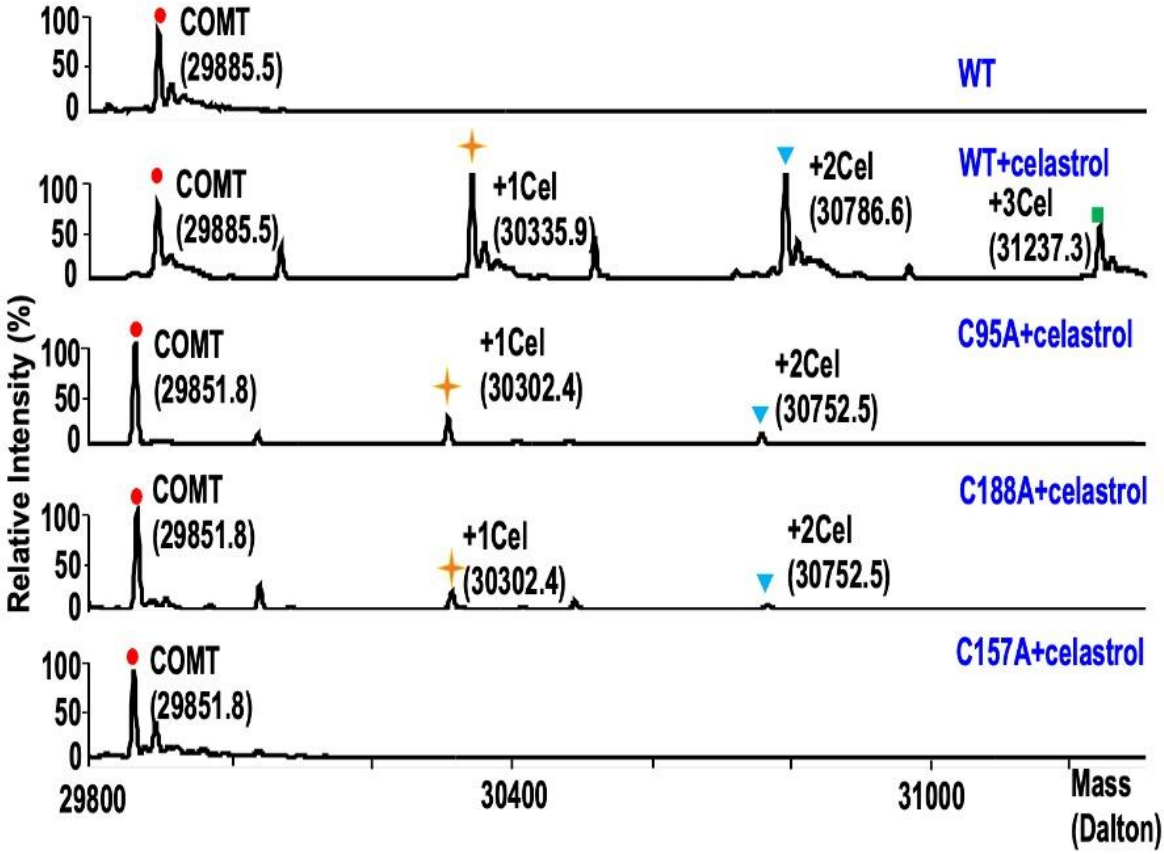


Figure 3-6 ESI-Q-TOF analysis on WT or mutant S-COMT (C95A, C188A, and C157A) with the Cel treatment. S-COMT is marked with the circle. After incubation for 2 hours at 4 °C with Cel, the peaks indicating S-COMT with one Cel, two Cel and three Cel are labelled as asterisks, triangles, and squares, respectively.

3.2.3 Direct binding site identification by using tandem MS

In order to demonstrate the direct binding sites of Cel, a tandem MS strategy was performed. S-COMT treated with Cel was used for trypsin digestion. Peptides were subjected into LC-MS/MS for analysis. By tandem MS, the fragment ion (MS2) spectra from the precursor ion offer structural information for peptide identification[123]. The unique peptide YLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVR was identified with a mass shift of Cel modification on Cys157. The typical MS spectra in **Figure 3-7** again demonstrated that Cel directly targeted Cys157 of S-COMT.

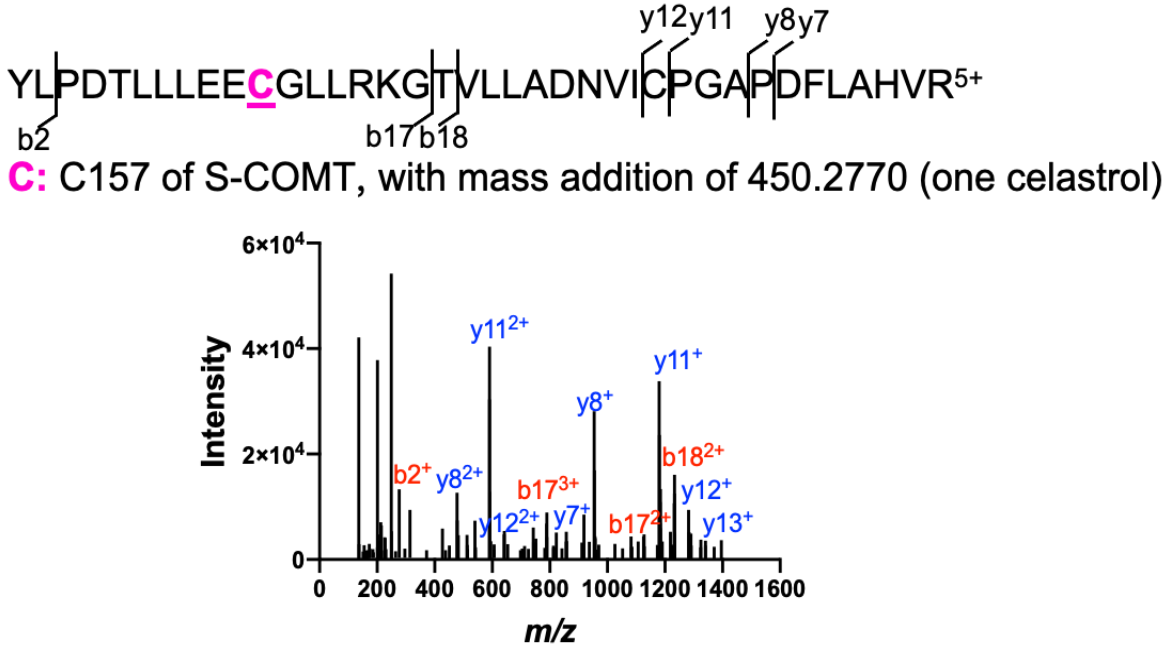


Figure 3-7 Tandem MS analysis on purified WT S-COMT. 2.5 μM COMT was incubated with 7-fold Cel for 2 hours at 4 ^{o}C.

3.3 Conclusion

Cysteines in COMT were reasoned as potential targets of Cel. To confirm the exact binding residue, recombinant and mutant S-COMT were purified for a series of binding stoichiometry assays. The result suggested there was an irreversible covalent interaction between Cel and S-COMT. Based on the results of the UV absorption assay, in-gel fluorescence scan, intact protein MS, and tryptic tandem MS, Cys157 of S-COMT was demonstrated as the primary binding site of Cel.

Chapter 4: The effect of Cel on enzymatic activity of COMT *in vitro* and *in situ*.

4.1 Introduction

COMT is a methyltransferase enzyme that catalyses catechol estrogens and catechol neurotransmitters such as dopamine (DA)[124, 125]. Genetic studies indicate that the function of COMT associates with various cognitive abilities in human[126]. COMT codes two forms including soluble S-COMT (26 kDa) and membrane-bound MB-COMT (30 kDa)[127]. S-COMT displays higher catalytic activity than MB-COMT, whereas MB-COMT has higher substrate affinity[128].

In the clinical, COMT inhibitors are widely used in combination therapy with levodopa to treat Parkinson's disease[129]. In addition, COMT inhibitors have also been applied in the therapy of Alzheimer's disease, schizophrenia, and deficit hyperactivity disorder[130-132]. The second-generation COMT inhibitors in clinical mainly involve entacapone and tolcapone[133]. However, present COMT inhibitors have been limited in clinical use due to their insufficient brain penetration and safety issues[134, 135].

In this chapter, we employed both *in vitro* and *in situ* assays to evaluate the effect of Cel on COMT enzymatic activity.

4.2 Materials and methods

4.2.1 Chemicals and reagents

3,4-dihydroxybenzylamine hydrobromide (DHBA, HY-N3023) was supplied by MedChemExpress. Dopamine (DA, H8502), 3-methoxytyramine (3-MT, M4251), 3,4-dihydroxyacetophenone (DHAP) were ordered from Sigma-Aldrich.

4.2.2 *In vitro* enzymatic activity assays

The enzymatic activity was evaluated by observing the change in absorbance of DHAP after the COMT-based catalysing reaction. Specifically, 1 μg purified S-COMT and its variants were first pre-treated with 0.1 to 100 μM Cel in treatment buffer (50 mM NaCl and 20 mM Tris-HCl, pH 7.4) at RT for 1 hour. Aliquots with and without 1 μg purified S-COMT and its variants served as negative and positive control groups. Subsequently, these pre-treated S-COMT and its variants (50 μL) were mixed with 50 μL 0.2 M N-Tris(hydroxymethyl)-

methyl-2-aminoethane sulphonic acid buffer containing 1 mM DHAP, 10mM S-adenosyl-L-methionine (SAM), 40 mM DTT and 12 mM $MgCl_2$. After incubation at 37 °C for 1 hour, the reaction was stopped by adding 0.4 M sodium borate (pH 10). To analyse the time-dependent inhibition of Cel, a similar protocol was performed but with different incubation times at RT. The relative enzyme activity was calculated as [(Abs sample, 340 – Abs negative, 340) / (Abs positive, 340 – Abs negative, 340)] × 100%.

4.2.3 Quantitative analysis of DA and 3-MT *in situ*

PC-12 cells (1×10^5 cells, 500 μL/well) in a 12-well plate were incubated with Cel or tolcapone at indicated concentrations. After 24 hours of incubation, 50 μL samples including cells and cultured medium were collected. Then, 200 μL cold methanol and 2 μL of 5 μM DHBA (internal standard) were added. The mixture was vortexed for 5 mins at 1,200 rpm, followed by centrifugation for 10 mins at 12,000 rpm. A volume of 250 μL supernatant was mixed with 125 μL 100 mM sodium tetraborate for 2 mins. Then, 125 μL 2% benzoyl chloride in acetonitrile (v/v) was added to each sample and mixed for 5 mins. Next,

125 μL 1% formic acid was added to quench the reaction. 5 μL of each sample was injected and analysed using a SCIEX 6500+ liquid chromatography/electrospray ionization (ESI) - QTrap mass spectrometer coupled with ExionLC AD series UHPLC instrument in positive ion mode. Solvents were 0.1% FA and 5 mM ammonium formate in water (A) and acetonitrile (B). LC separation was performed with a 12 min gradient using an ACQUITY UPLC HSS T3 Column, 1.8 μm, 2.1 mm × 100 mm starting with 5% B. The flow rate was 300 μL/min. The LC gradient was set as: 0−1 min, 5% B; 1−4 min, 49% B; 4−6 min, 85% B; 6−7.5 min, 95% B; 7.5-10 min, 95% B; 10.2 min, 5% B. Precursor ion of DA and 3-MT were m/z 466.1649 and m/z 376.1543, respectively. Product ion of DA and 3-MT were both m/z 105.0330. The MS parameters were as follows: 4.5 kV ion spray voltage; 600 °C ion source temperature; 50 psi nebulizer gas (GS 1) and heater gas (GS 2). Statistical calculations were performed using Prism 8.1.1. *$p < 0.05$ was considered as statistical significance.

4.2.4 COMT stability analysis

Live PC-12 cells were treated with 1 μM Cel. After incubation for 24 hours, western blotting was performed to analyse the level of COMT in PC-12 cells.

4.2.5 Statistical analysis

All calculations of significance were based on $p^* < 0.05$, $p^{**} < 0.01$ and $p^{***} < 0.001$ unless otherwise stated.

4.3 Results and discussion

4.3.1 *In vitro* assay

As COMT is a well-known methyltransferase, we questioned whether Cel can affect its enzymatic activity. Given that COMT catalyses the methyl transfer from SAM to DHAP, the absorption of DHAP enables the monitoring of COMT's enzymatic activity[136, 137]. According to our findings, Cel inhibited COMT's enzymatic activity in a dose-dependent manner. Compared to S-COMT with an IC50 of 6.22±0.34 μM, MB-COMT showed slightly higher sensitivity to Cel treatment with an IC50 of 3.81±0.34 μM (**Figure 4-1A**). In a time-dependent experiment, the observed half-effective time of Cel to inhibit S-COMT and MB-COMT

were 15.6 and 12.6 mins respectively (**Figure 4-1B**). Additionally, whereas mutation of Cys157 dramatically suppressed Cel's inhibitory ability to COMT enzymatic activity, mutations of Cys173, Cys188 and Cys191 had no effect on the Cel inhibition of COMT enzymatic activity, demonstrating that Cel primarily targeted Cys157 to inhibit COMT enzymatic activities (**Figure 4-2**).

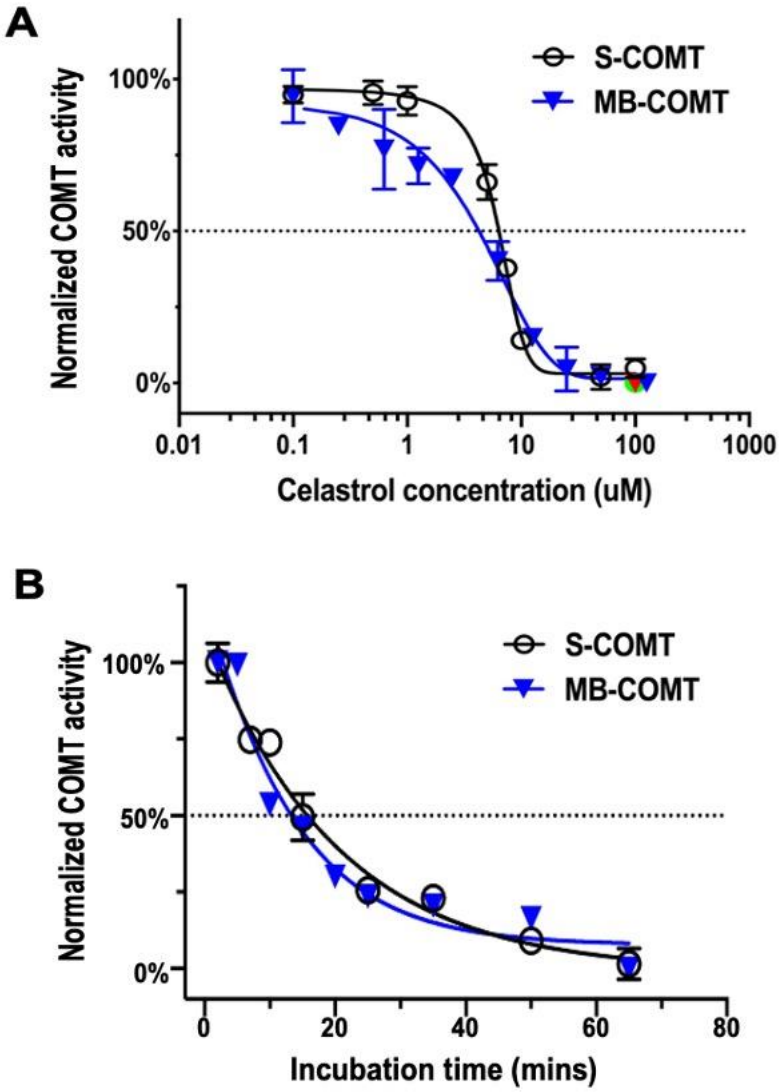


Figure 4-1 The enzymatic activity of COMT *in vitro*. (A) The enzymatic activities of COMTs after Cel treatment at indicated concentrations. The red triangles and green circle indicated the inhibition effect of 100 μM entacapone on MB-COMT and S-COMT, respectively. (B) The kinetics of 50 μM Cel-COMT incubation.

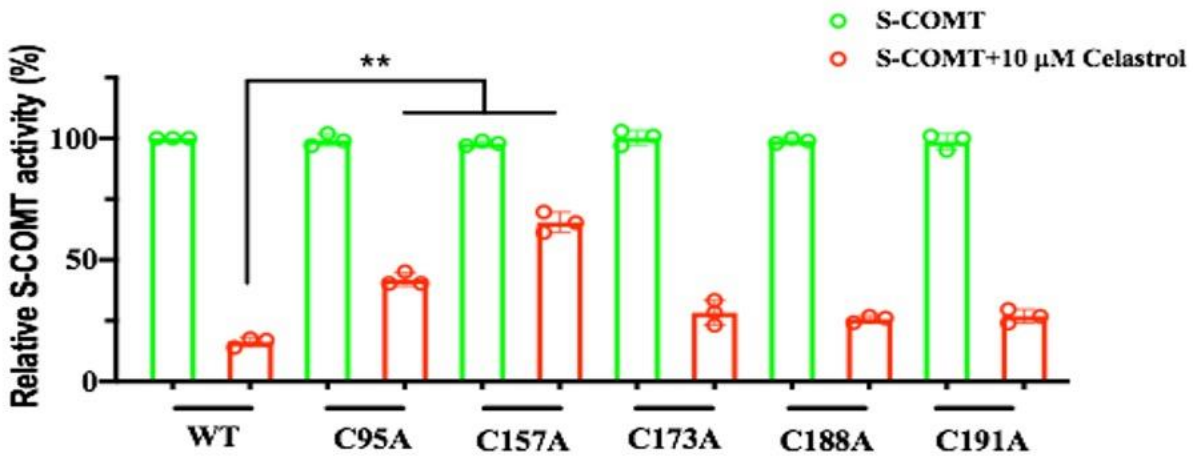


Figure 4-2 Enzymatic activities of S-COMT and mutants in response to Cel treatment. Activity was performed by assessing the transfer of methyl from SAM into DHAP catalysed by COMT.

4.3.2 *In situ* assay

DA is known to act as a catecholamine neurotransmitter involved in multiple physiological processes. 3-MT is the major metabolite of DA mediated by COMT[138, 139]. PC-12 is a rat adrenal pheochromaocytoma-derived cell line that has been widely applied as a model system for the synthesis and release of DA[140]. We then performed targeted MS analysis of DA and its metabolite 3-MT in PC-12 cells. Tolcapone is the FDA-approved COMT inhibitor[141], which is regarded as the positive control in this assay. As shown in **Figure 4-3**, upon Cel treatment, a significant increase of DA as well as a significant decrease

of 3-MT were observed in a dose-dependent manner. Notably, the amount of DA increased by 6-fold after treatment with 1 μM Cel. The effect of Cel in DA and 3-MT was close to the positive drug tolcapone, suggesting Cel showed an equivalent inhibitory effect in COMT enzymatic activity with tolcapone. In addition, in PC-12 cells, Cel did not influence the expression or stability of COMT (**Figure 4-4**). It was conceivable that the alteration in DA and its COMT-catalysed metabolite 3-MT resulted from the inhibition of COMT enzymatic activity by Cel treatment.

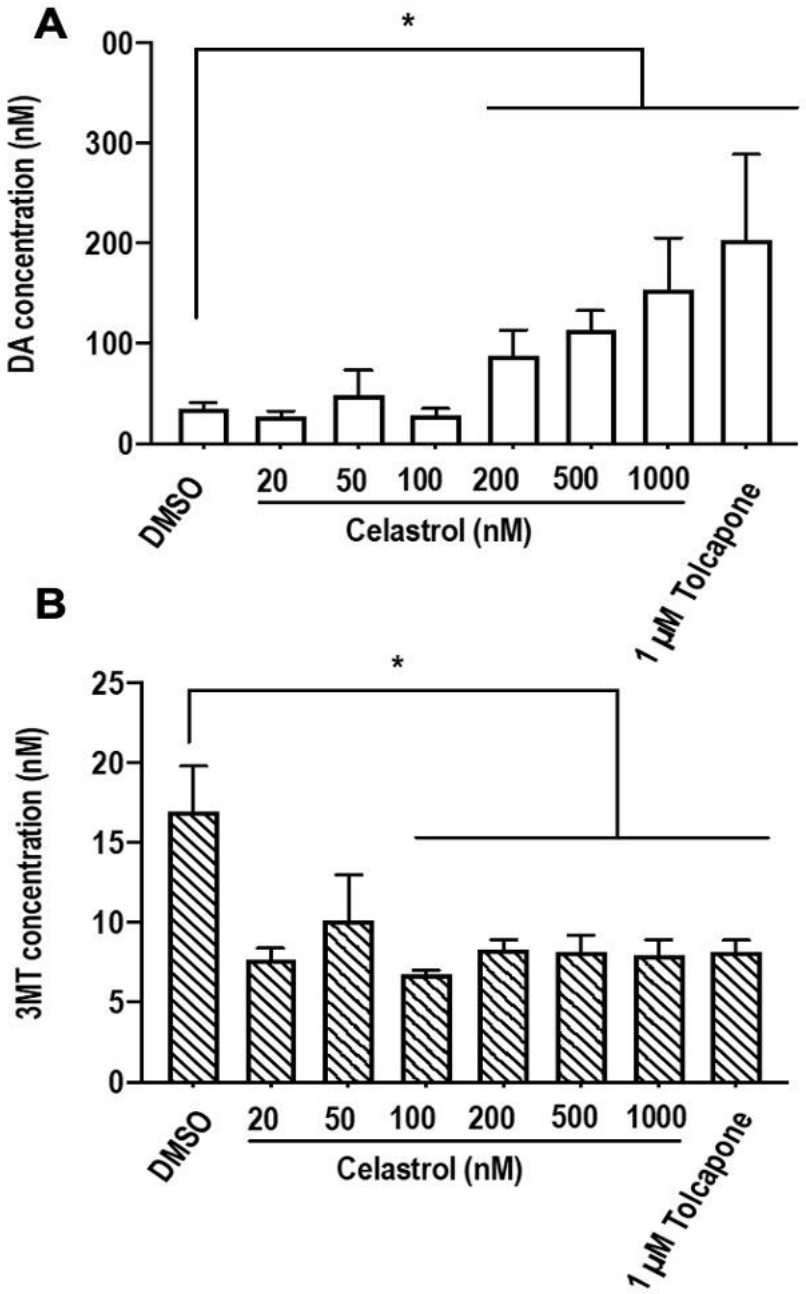


Figure 4-3 Quantitative analysis of DA and 3-MT with targeted MS in PC-12 cells. *p<0.05.

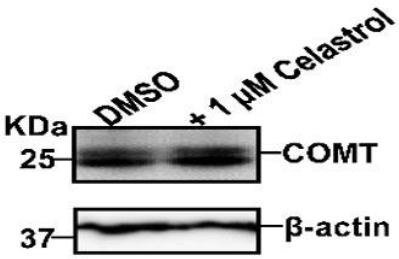


Figure 4-4 The effect of Cel on COMT stability in PC-12 cells. Stability of COMT was evaluated by comparing the level of COMT in PC-12 cells exposed to DMSO or 1 μM Cel for 24 hours.

4.4 Conclusion

We utilised functional enzymatic assays to evaluate the effects of Cel on COMT. On the one hand, by monitoring the enzymatic activities of COMT and its mutations under Cel treatment *in vitro*, the result revealed that Cel inhibited the enzymatic activities of COMT by mainly targeting Cys157. On the other hand, through determining the cellular levels of DA and its metabolite 3-MT via targeted MS *in situ*, we concluded that the change of DA and 3-MT were triggered by Cel's inhibition of COMT enzymatic activity. Admittedly, the main limitation is the lack of COMT knockdown experiment on PC-12 cells for *in-situ* assay. With the knockdown experiment, it can further verify

the enzymatic activity of COMT and the inhibitory effect of Cel on the enzymatic activity of COMT.

Chapter 5: Multiplexed target profiling of electrophiles with a chemical proteomics-genomics approach

5.1 Introduction

Identification of the cellular targets of bioactive small molecules has been the mainstay and the biggest challenge in drug discovery. Chemical proteomics is a growing area of chemical biology that employs chemical probes to understand protein-small molecule interactions. As cysteines serve as the most nucleophilic amino acid and electrophiles can interact with reactive cysteines via covalent attachment in proteins, chemical proteomic approaches using cysteine-reactive probes and click reactions by cleavable biotins has been successfully developed for target profiling[42, 142]. Traditional target identification of bioactive molecules was mostly on an individual basis. There is still a lack of systematic strategy to efficiently examine multiple compounds in parallel, particularly for structurally distinct compounds.

Using gene expression signatures to explore the mode of action of bioactive molecules presents the power of genomic and bioinformatic

methods in drug discovery[143, 144]. L1000 Connectivity Map (CMap) platform, that is profiling gene expression changes of 978 landmark genes with pharmacologic or genetic (knockdown or over-expression) perturbations in human cell lines[145]. The L1000 CMap has the advantage of high throughput, can thus be utilised to screen and rank small molecules according to their similarity of cellular perturbations on gene expression levels. We asked whether the L1000 CMap could serve as a sensor for target profiling via clustering or distinguishing small molecules based on their genomic similarities. In this chapter, we selected three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their induced cellular genomic perturbations. Next, a cysteine profiling approach using two chemical probes was developed to map reactive cysteines. We are asking if there is a correlation between the cysteine reactivity of electrophilic compounds and their perturbations on gene expression levels.

5.2 Materials and methods

5.2.1 Chemicals and reagents

Cel (HY-13067), Auranofin (SKF-39162), Withaferin A (HY-N2065), Triptolide (HY-32735) and Vinpocetine (HY-13295) were supplied by MedChemExpress. Alkynyl iodoacetamide (IAA, EVU111) and UV cleavable biotin-azide (EVU102) were purchased from Kerafast. TMS-ethynylbenziodoxolone (EBX) was a generous gift from Prof. Jerome Waser (EPFL, Ecole Polytechnique Federale de Lausanne). All gel-casting reagents and protein markers were purchased from Bio-Rad. TAMRA-azide, biotin-azide, Tris((1-benzyl-4-triazolyl) methyl) amine (TBTA), DDE biotin azide (1136-5), and DADPS biotin azide (1330-5) were ordered from Click Chemistry Tools.

5.2.2 Cell culture

HeLa and HeLa S3 were grown in DMEM. Jurkat, NCI-H460, and PC-12 cells were grown in RPMI 1640 medium. PC3 cell was grown in Ham's F12K medium.

5.2.3 Fluorescence gel imaging using IAA or EBX

For HeLa S3 cell lysate (2 μg/uL, 50 uL), probe labelling was performed with IAA (10, 50, 100 μM) or EBX (10, 50, 100 μM) for 1

hour at RT. Click chemistry was initiated by sequential addition of the following to each lysate: 100 μM TAMRA-azide, 1 mM TCEP, 100 μM TBTA ligand in 1:4 DMSO:t-butyl alcohol, and 1 mM CuSO4. After incubation for 1.5 hours at RT, reactions were quenched by acetone protein precipitation. Protein pellets were resuspended in the sample buffer, resolved and separated by SDS-PAGE. Fluorescence gels were scanned by using Typhoon Imaging System (GE), then stained with Coomassie blue.

5.2.4 Western blotting

Protein lysates were loaded onto SDS-PAGE and transferred to a polyvinylidene difluoride membrane. The antibody being used was: HRP-conjugated streptavidin (1:5000, N100, Thermo Fisher Scientific), After overnight incubation at 4 °C, signals were visualised by using the chemiluminescence method.

5.2.5 Compound treatment and cell lysis in PC3 cells

PC3 cells were grown until 90% confluency, treated with 10 μM compounds or vehicle for 1 hour at 37 °C and then rinsed two times

with 1 × PBS. Cells were lifted by 0.25% trypsin and pellets were harvested by centrifugation at 1,400 g for 3 min. Cell pellets were lysed by sonication in cold 1 × PBS with 1 × protease inhibitor (Roche), pH 7.4. The sonifier (SFX 550, Branson) worked with a pulsed regime for 2 mins (10-second sonication, 15-second pause) at 30% amplitude under ice cooling. After centrifugation at 16,000 g for 20 mins at 4 °C, the supernatant was collected. For HeLa, HeLa S3, Jurkat, and NCI-H460 cell lines without compound treatment, the procedure is the same with PC3 cells.

5.2.6 Preparation of mouse tissue sample

Mouse brain tissues were harvested and immediately frozen in liquid nitrogen. The tissues were then homogenised in 200 μL 1 × PBS supplemented with 1 × protease inhibitor (Roche), pH 7.4. The tissues were homogenized by an automated homogeniser (Bertin Technologies) by running 4 cycles of 10 seconds at 5,800 rpm, with a 30-second break between cycles. After centrifugation at 16,000 g for 20 mins at 4 °C, the supernatant was collected.

5.2.7 Probe labelling and cysteine profiling sample preparation (click reaction in proteins)

Protein lysates were labelled with 100 μM IAA for 1 hour at RT. Excess IAA was removed by passing the sample through a NAP-5 column (GE, Healthcare). Click chemistry was performed in the presence of 300 μM UV biotin azide or DDE biotin azide, 2 mM TCEP, 300 μM TBTA ligand in 1:4 DMSO:t-butyl alcohol and 2 mM CuSO4 at RT for 1 hour in the dark. Protein precipitation was performed with a methanol-chloroform system (water /methanol/chloroform, 4:4:1 (v/v/v)). The protein pellets were resuspended by sonication in 1 × PBS containing 0.2 M urea to incubate with streptavidin agarose for 4 hours at RT. Streptavidin agarose was washed with 2 M urea in 1 × PBS and ultra-pure water using a micro bio-spin column to remove non-specific binding proteins and salts. The washed agarose beads were resuspended in 1 × PBS, reduced with 10 mM DTT, incubated at 45 °C for 30 mins, and further alkylated with 30 mM IAM at RT for 30 mins in the dark. Following reduction and alkylation, the on-beads digestion with MS-grade Lys-C (Wako) at an enzyme/protein ratio of 1:200 (w/w) was allowed to proceed for 2 hours at 37 °C. A secondary digestion was

performed with sequencing-grade trypsin (Promega) at an enzyme/protein ratio of 1:50 (w/w) for additional 12 hours at 37 °C. The digest was isolated from the beads using a micro bio-spin column. The beads were resuspended and washed with 2 M urea in 1 × PBS, 1 × PBS, and ultra-pure water. For UV-biotinylated peptides, the washed agarose beads were resuspended in ultra-pure water and irradiated with 365 nm UV light for 15 mins under gentle shaking on ice. For DDE-biotinylated peptides, the agarose beads were resuspended in 2% hydrazine and incubated for 1 hour at RT under gentle shaking. The supernatant was collected by centrifugation (1,200 g, 2 mins), cleaned by self-packed C18 stage-tip, dried under vacuum for further MS analysis.

5.2.8 Probe labelling and cysteine profiling sample preparation (click reaction in peptides)

Protein lysates were labelled with 100 μM IAA or 50 μM EBX for 1 hour at RT. Proteins were then reduced with 10 mM DTT, incubated at 45 ºC for 30 mins, and further alkylated with 30 mM IAM at RT for 30 mins in the dark. Protein precipitation was performed with a methanol-

chloroform system (water/methanol/chloroform, 4:4:1 (v/v/v)). The protein pellets were resuspended by sonication in 25 mM ammonium bicarbonate containing 0.2 M urea. Lysates were digested by MS-grade Lys-C (Wako) at an enzyme/protein ratio of 1:200 (w/w) for 2 hours at 37 °C. A secondary digestion was performed with sequencing-grade trypsin (Promega) at an enzyme/protein ratio of 1:50 (w/w) for additional 12 hours at 37 °C. Then the tryptic digests were desalted with HLB extraction cartridges (Waters) and dried. The dried peptide mixtures were then dissolved in 50 μL 30% ACN. Click chemistry was performed in the presence of 300 μM UV biotin azide or DDE biotin azide, 2 mM TCEP, 300 μM TBTA ligand in 1:4 DMSO:t-butyl alcohol, and 2 mM $CuSO_4$ for 1 hour at RT in the dark. Strong cation exchange (SCX) was used to remove the excess biotin. The eluent in 5 mM NaH_2PO_4, 0.4 M NaCl, 25% ACN (PH=3.0) was dried and then reconstituted in 1 × PBS to interact with pre-washed streptavidin agarose for 4 hours at RT. Streptavidin agarose was washed with 2 M urea in 1 × PBS, 1 × PBS, and ultra-pure water (twice) using a micro bio-spin column to remove non-specific binding peptides and salts. The washed agarose beads were resuspended in ultra-pure water and

irradiated with 365 nm UV light for 15 mins under gentle shaking for UV-biotinylated peptides. For DDE-biotinylated peptides, the agarose beads were resuspended in 2% hydrazine and incubated for 1 hour at RT under gentle shaking. The supernatant was collected by centrifugation (1,200 g, 2 mins), cleaned by self-packed C18 stage-tip, dried under vacuum, and stored at −80 °C until mass spectrometry analysis.

5.2.9 LC-MS/MS of cysteine profiling

LC-MS/MS analyses were performed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) coupled with an UltiMate 3000 UPLC system (Thermo Fisher Scientific). An RSLC C18 analytical column (75 μm × 250 mm, 2.0 μm, 100 Å) (Thermo Fisher Scientific) was employed for LC separation. Mobile phase A is 0.1% FA in water and mobile phase B is 0.1% FA in ACN. A 120-minute gradient with a 300 nL/min flow rate and an initial 2% mobile phase B was used. Mobile phase B was increased to 4% at 12 mins, 30% at 88 mins, 85% at 104 mins and held for 5 mins. Then, mobile phase B was back to 2% at 110 mins and maintained this composition until

120 mins. Data was collected in data-dependent acquisition (DDA) mode. The top ten precursor ions with a charge state of 2+ or higher were fragmented by HCD. The MS1 Orbitrap resolution was set at 60,000 and the MS1 AGC target was set at 4×10^5. The MS2 Orbitrap resolution was set at 30,000 while the MS2 AGC target and the maximum injection time were set at 1×10^5 and 50 ms.

5.2.10 MS data process of cysteine profiling

The DDA raw data were analysed with Proteome Discoverer 2.4. The searching was against the homo sapiens UniProt database (Version June 2020, 20368 entries). The precursor mass tolerance was set to 10 ppm with a fragment tolerance of 0.02 Da. The maximum number of modifications was four. Methionine oxidation, carboxyamidomethylation on cysteine, custom modification from cleavage of cleavable biotin (IAA-UV adduct *m/z* 252.1222; EBX-UV adduct *m/z* 181.0851; IAA-DDE adduct *m/z* 315.1443; EBX-DDE adduct *m/z* 243.0994) were specified as variable modifications, no fixed modifications were specified. The identified proteins were

filtered with a false discovery rate of 1% and normalization was performed against the total peptide amount.

5.2.11 Mathematical measurement

We compared the reactivity profiles of a reference compound with a target compound by the difference in cumulative mean of the logit-transformed reactivity ratios. Let p be the total number of cysteine sites and $(r_1, \dots, r_p)$ be the reactivity ratio (r, r = $I_{electrophile}/I_{DMSO}$. $I_{electrophile}$: Intensity of electrophile treatment. I_{DMSO}: Intensity of DMSO treatment). We transformed the reactivity ratio using the logit-transformation by the function:

$$logit(r) = ln\frac{r}{1-r}$$

and call the transformed value the reactivity score where smaller score values correspond to higher reactivity.

To avoid extreme score values, the reactivity ratios are truncated at 0.01 and 0.99. Let $(x_1, \dots, x_p)$ be the reactivity scores of the cysteine sites under the reference compound where the cysteine sites are ranked in

ascending order of x_j's and $(y_1, \cdots, y_p)$ be the reactivity scores of the corresponding cysteine sites under the target compound. We plot the function:

$$f(t) = \frac{\sum_{j=1}^{p} I(x_j \leq t)(x_j - y_j)}{\sum_{j=1}^{p} I(x_j \leq t)}$$

At any t the function value is computed by taking the cysteine sites with reactivity score smaller than t under the reference compound and computing the difference in mean reactivity scores among these cysteine sites under the reference compound and the target compound.

5.3 Results and discussion

5.3.1 Method establishment for cysteine profiling

First, we established a cysteine profiling method to capture reactive cysteines. We optimised our method by assessing three cleavable azide-biotins and two thiol-reactive probes (**Figure 5-1**). IAA and EBX were utilised, the former is the most popular cysteine-reactive chemical probe in the last decades, and the latter is a novel cysteine-reactive probe developed by Prof. Jerome Waser. Three cleavable biotin azide DDE, UV and DADPS enable the tagging of alkyne modified

substrates via copper-catalysed click chemistry, while the captured target can be released under 2% aqueous hydrazine, photolysis conditions at 365 nm and 10% formic acid, respectively.

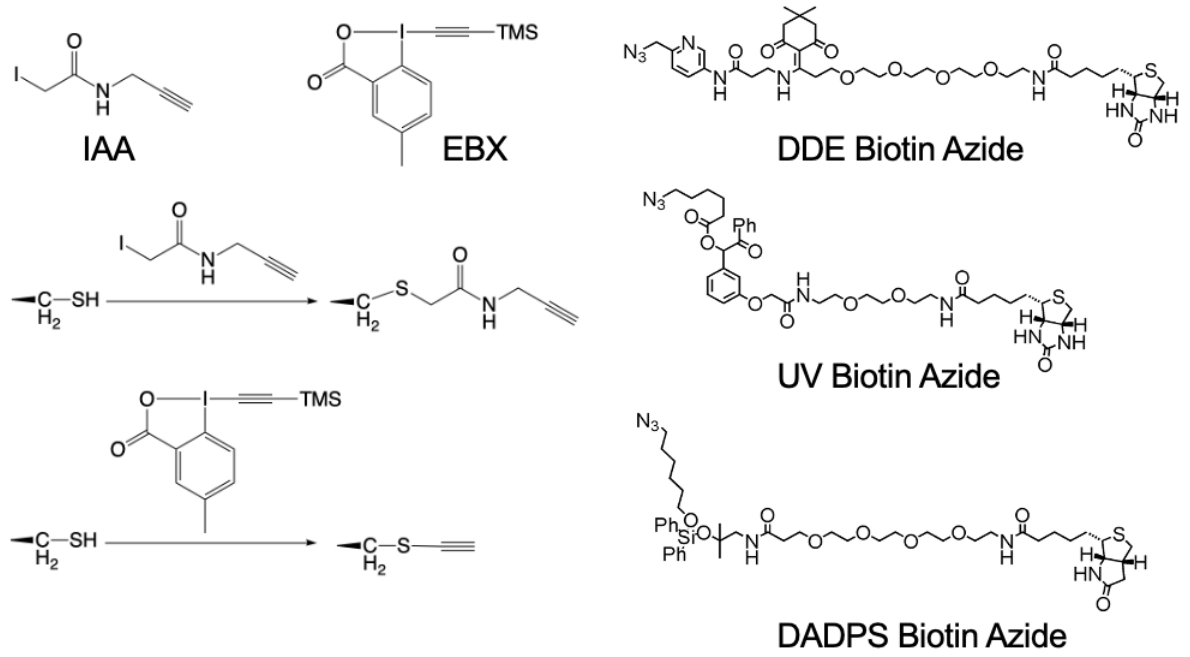


Figure 5-1 Chemical structures of probes and cleavable biotin azide.

Given the result shown in **Figure 5-2A**, we determined the optimal working concentrations of IAA (100 μM) and EBX (50 μM). Additionally, cleavable biotin-azide was optimised to be UV cleavable biotin azide with a higher labelling efficiency (**Figure 5-2B**).

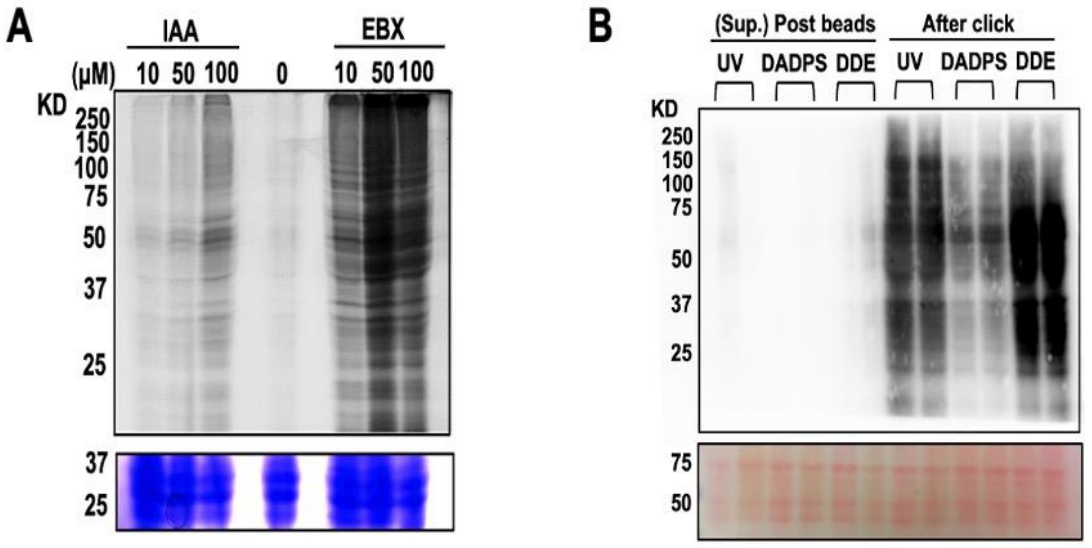


Figure 5-2 Establishment of method for cysteine labelling. (A) Determination of optimal concentration of probe labelling. HeLa S3 lysate was labelled by IAA or EBX with indicated concentrations for 0.5 hour at RT. Lysate was then subjected to click reaction with 100 μM TAMRA-azide and in-gel fluorescence analysis. Upper: fluorescence scan, below: Coomassie blue staining. (B) Comparison of labelling efficiency of three cleavable biotins. HeLa S3 lysate was labelled by 100 μM IAA for 0.5 hour at RT. Lysate was then subjected to click reaction with 100 μM indicated biotin azide, followed by enrichment of magnetic streptavidin beads, boiling beads and releasing labelled-proteins, resolving in SDS-PAGE and western blotting by HRP-conjugated streptavidin antibody. The supernatant after magnetic

streptavidin beads enrichment was also collected to show the enrichment efficiency as one quality control (left part of gel). Upper: western blotting scan, below: ponceau's stain.

We also assessed alterations in cysteine profiling, showing that a higher number of reactive cysteine identification was obtained when the click reaction occurred in peptides (**Figure 5-3**).

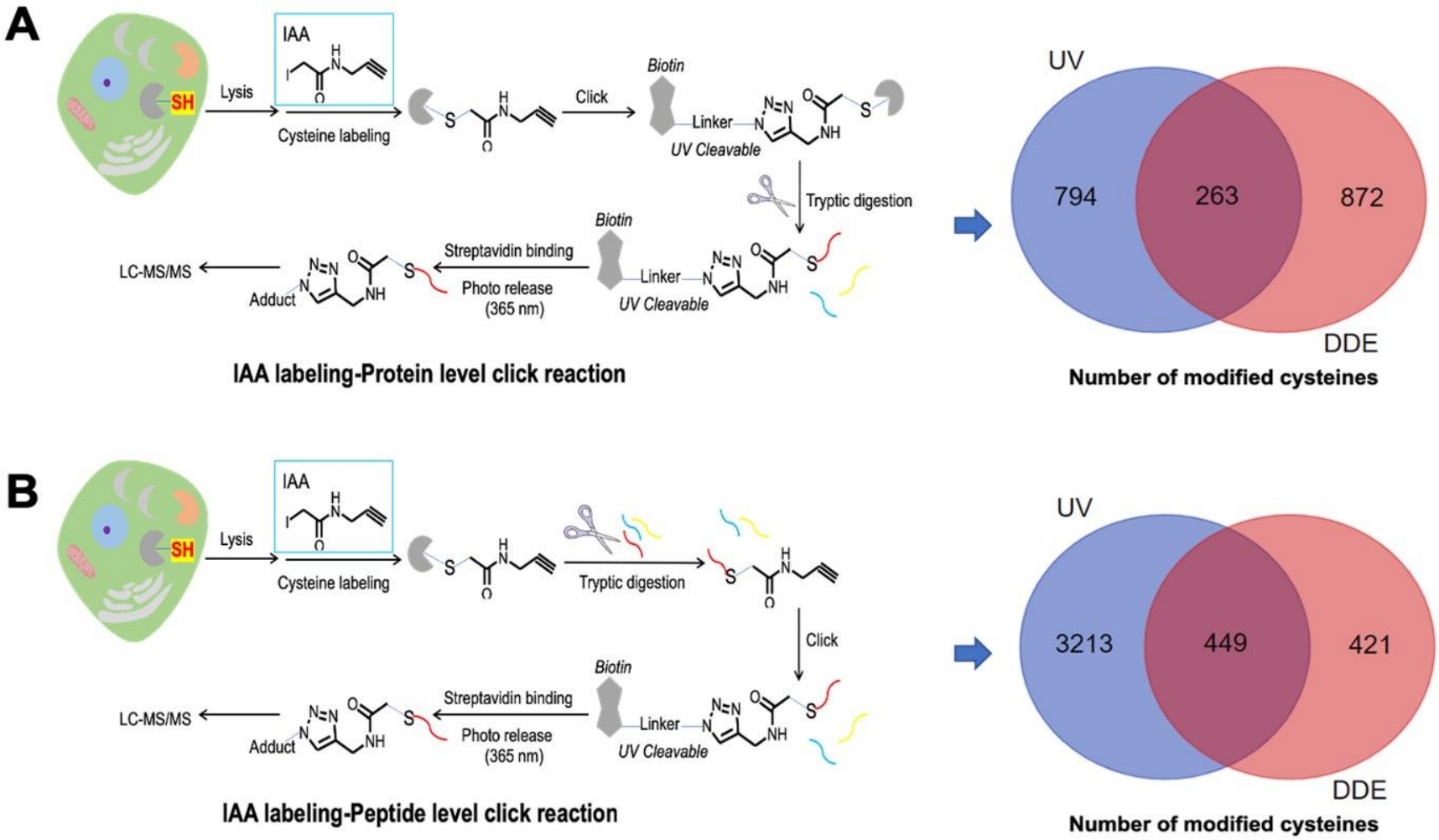

A
SH
Lysis
IAA
Cysteine labeling
Click
Biotin
Linker
UV Cleavable
Tryptic digestion
Biotin
Streptavidin binding
Photo release
(365 nm)
LC-MS/MS
Adduct
Linker
UV Cleavable
IAA labeling-Protein level click reaction
UV
794
263
872
DDE
Number of modified cysteines
B
SH
Lysis
IAA
Cysteine labeling
Tryptic digestion
Click
Biotin
Streptavidin binding
Photo release
(365 nm)
LC-MS/MS
Adduct
Linker
UV Cleavable
IAA labeling-Peptide level click reaction
UV
3213
449
421
DDE
Number of modified cysteines

Figure 5-3 Method optimisation of cysteine profiling by MS. General workflow and identification number of cysteine profiling (A) with protein-level click reaction using IAA probe and UV biotin azide, (B) with peptide-level click reaction using IAA probe and UV biotin azide. 500 μg HeLa S3 lysate was used for the proteomic analysis using the indicated workflow. Experiments were performed in biological duplicates (n=2).

5.3.2 Method evaluation with cell lines and tissue samples

The cysteine profiling coverage was also improved by simultaneously applying the two cysteine-reactive probes. Compared to a single chemical probe whose cysteine mapping capacity is limited by its own structure, the simultaneous use of IAA and EBX can achieve the assessment of sterically buried cysteines. In addition, we found that the number and distribution of probe-specific cysteines in different types of samples varied (**Figure 5-4**), demonstrating that our method can capture more probe-specific cysteines and achieve a better profiling coverage.

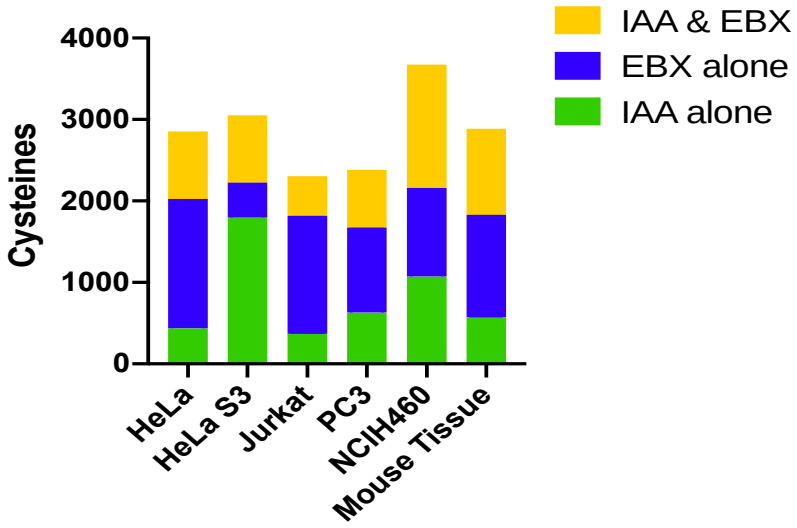


Figure 5-4 Identification number of cysteines from different cell lines and mouse tissue samples. Cysteines were divided into three types: labelled by IAA alone, labelled by EBX alone and labelled by two probes. Experiments were performed in biological duplicates (n=2).

The availability and reproducibility of this strategy were also validated in five human cell lines, with correlation coefficients over 0.90 across different cell lines (**Figure 5-5A** and **5-5B**). Furthermore, the cysteines that can be labelled by both IAA and EBX may represent cysteines with higher reactivity. The typical spectra of IAA-labelled and EBX-labelled cysteine-containing peptides were shown in **Figure 5-5C** and **5-5D**, respectively. Taken together, our results indicated that using the two complementary chemical probes coupled with UV biotin azide

followed by click reaction occurring in peptides enabled the highest number of reactive cysteine identification.

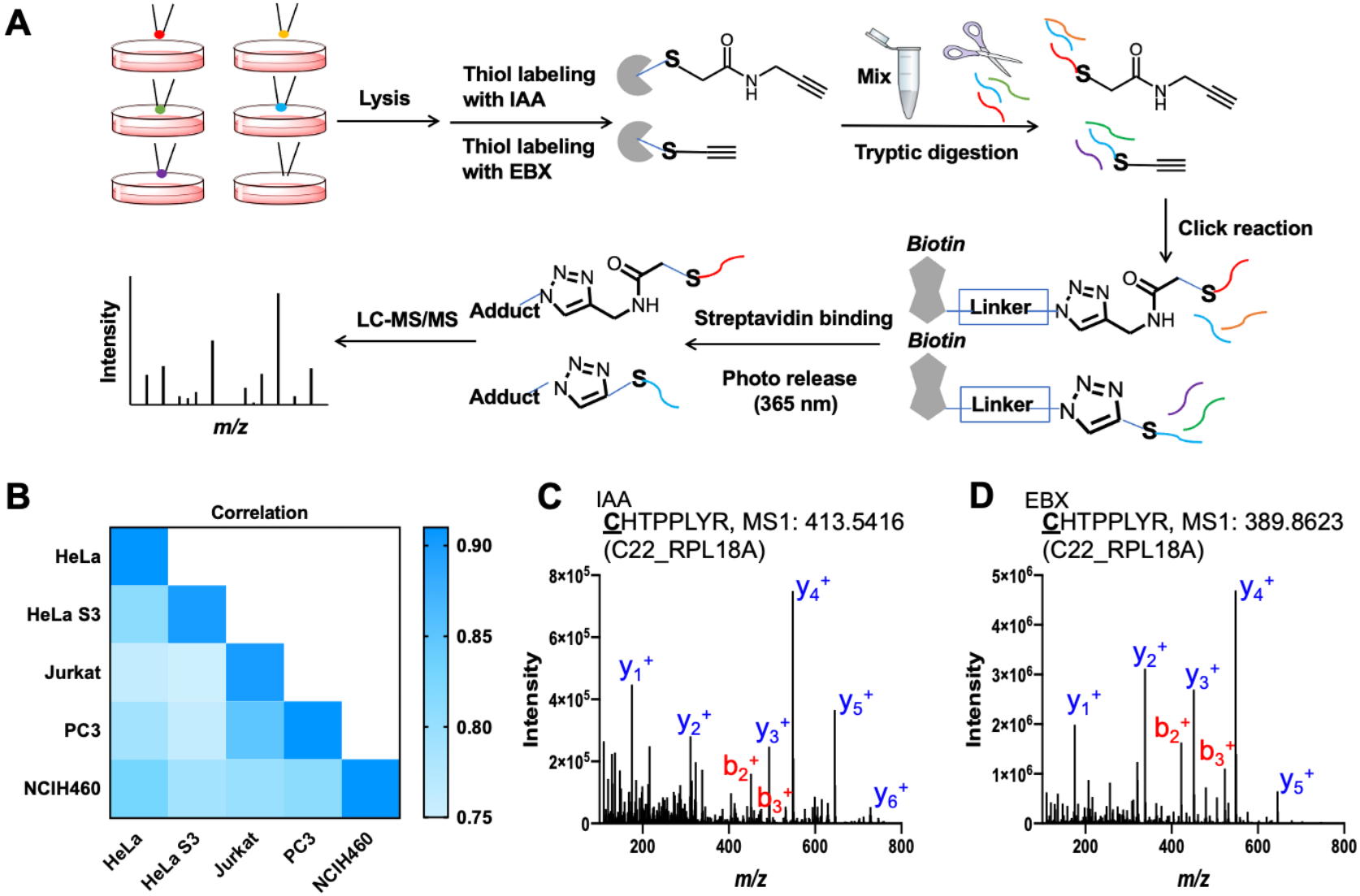

A
Lysis
Thiol labeling with IAA
Thiol labeling with EBX
Mix
Tryptic digestion
Click reaction
Biotin
Linker
Biotin
Linker
Streptavidin binding
Photo release (365 nm)
Adduct
Adduct
LC-MS/MS
Intensity
m/z
B
Correlation
HeLa
HeLa S3
Jurkat
PC3
NCIH460
0.90
0.85
0.80
0.75
C
IAA
CHTPPLYR, MS1: 413.5416
(C22_RPL18A)
D
EBX
CHTPPLYR, MS1: 389.8623
(C22_RPL18A)

Figure 5-5 Method establishment and evaluation for cysteine profiling. (A) General workflow of optimised cysteine profiling method. (B) Correlation analysis of our method in five cell lines. (C and D) Typical spectra of probe labelled peptides.

5.3.3 Selection of electrophiles from chemical genomics database

We next verified associations between the cysteine reactivity of electrophilic compounds and their perturbations on gene expression levels. We selected Cel as a representative electrophilic bioactive compound, and other three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their perturbation on global gene expression. According to the L1000 CMap database, when regarding Cel as a reference in PC3 cell line, auranofin (Aur), withaferin-a (WA), triptolide (TL) were selected with a connectivity score of 98, 96, 56, respectively. Vinpocetine (Vin) with a connectivity score of 0 was regarded as the negative compound. (**Figure 5-6A and 5-6B**). Indicatively, Aur and WA are the most similar in gene expression profile to Cel, TL is of medium similarity and Vin is the least similar.

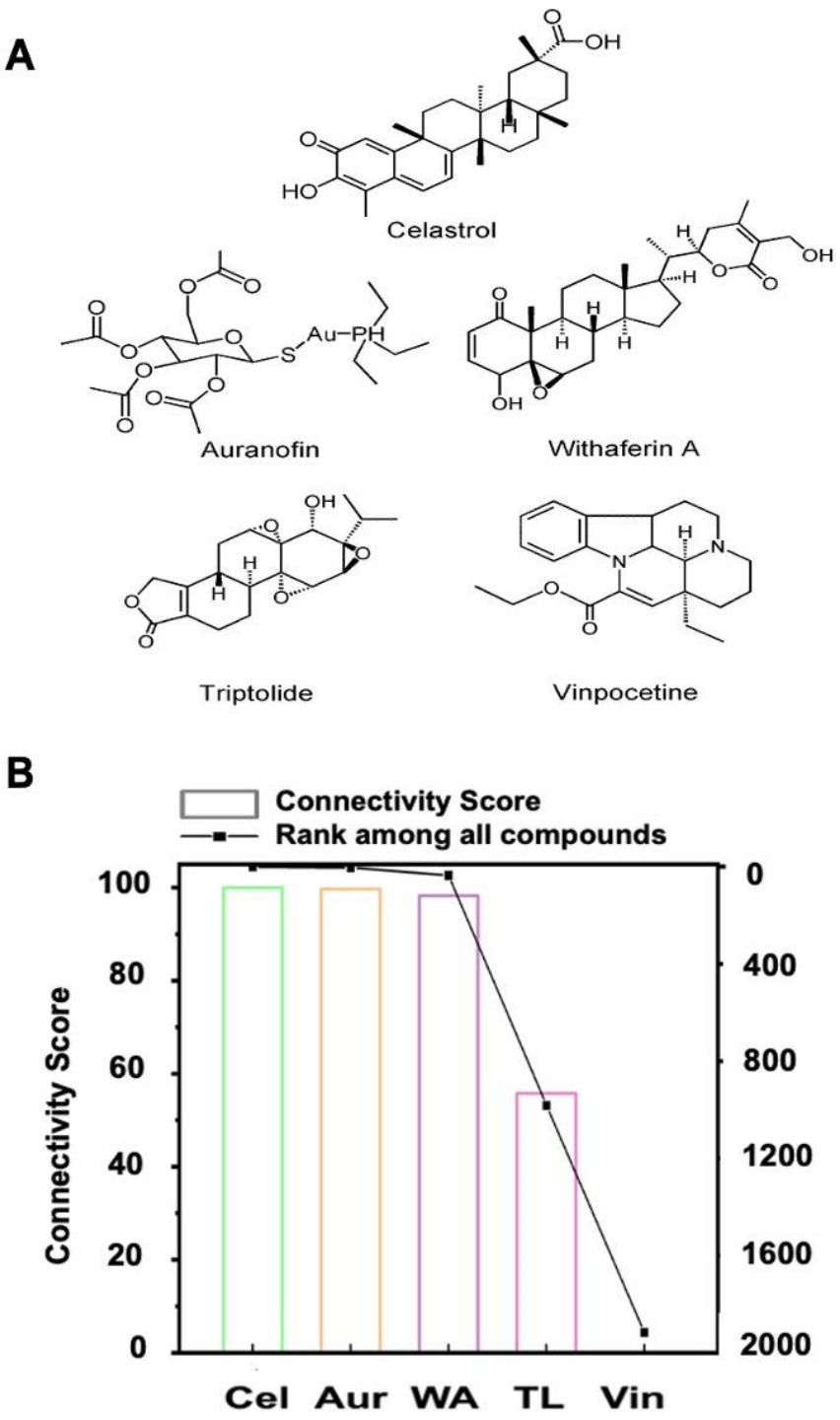


Figure 5-6 Structures and the connectivity score and rank of indicated electrophiles. (A) Structures of five electrophiles. (B) The connectivity score and rank of the indicated electrophile with Cel as the reference,

based on chemical genomics analysis. Score of 100 or rank as top 1 indicates extreme similarity to Cel.

We then applied the optimised cysteine profiling strategy for proteomic investigations. As shown in **Figure 5-7A**, Cel, Aur and WA presented similar global patterns with higher %inhibition, while TL and Vin showed lower %inhibition. According to the distribution plot of %inhibition (**Figure 5-7B**), the order of median values is Cel (74) > Aur (72) > WA (71) > TL (12) > Vin (7). As the cysteine with higher %inhibition represents greater reactivity to the electrophiles, our results indicated that Cel, Aur and WA targeted reactive cysteines with stronger binding abilities while TL and Vin showed relatively weaker cysteine-binding abilities.

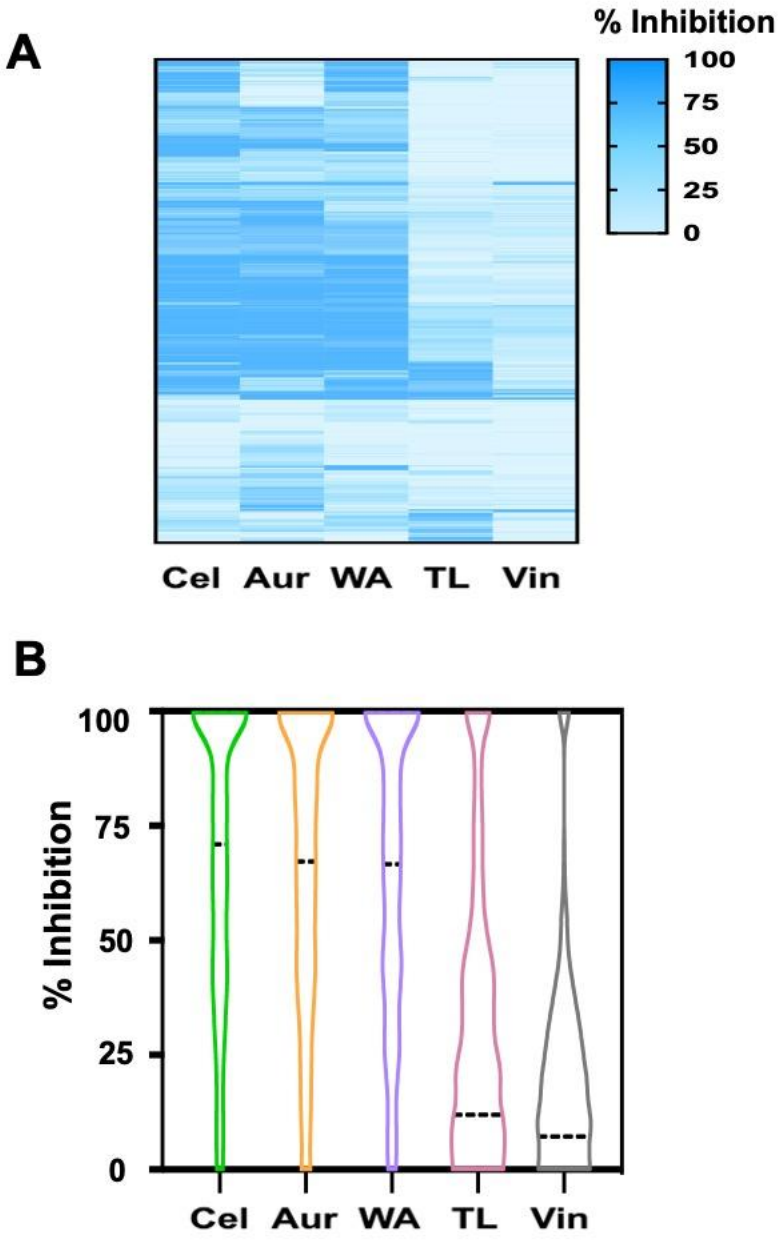


Figure 5-7 A heatmap (A) and a violin plot (B) show the results of a competition experiment and proteomics quantification. Live cells were incubated with the indicated compound followed by probe labelling. Cysteines that react with electrophiles will no longer get labelled by probe and thus cause decreased MS intensity. The percentage reduction of probe labelling by electrophiles was calculated for each cysteine as

follows: % inhibition = $100(1-I_{electrophile}/I_{DMSO})$. $I_{electrophile}$: Intensity of electrophile treatment. I_{DMSO}: Intensity of DMSO treatment. Dotted line indicates the median value of % inhibition.

To further evaluate the cysteine-binding abilities of electrophiles, we constructed a mathematical measurement (**Figure 5-8**). When taking Cel as the reference, its cysteine-binding ability was set as the baseline y=0, and any curve of compound below y=0 indicated lower reactivity between the compound and a specific cysteine compared to the reference. The curves of Cel, Aur and WA were close to each other and kept a similar pattern in cysteines with high reactivity (equal to %inhibition > 50) and low reactivity (equal to %inhibition < 50).

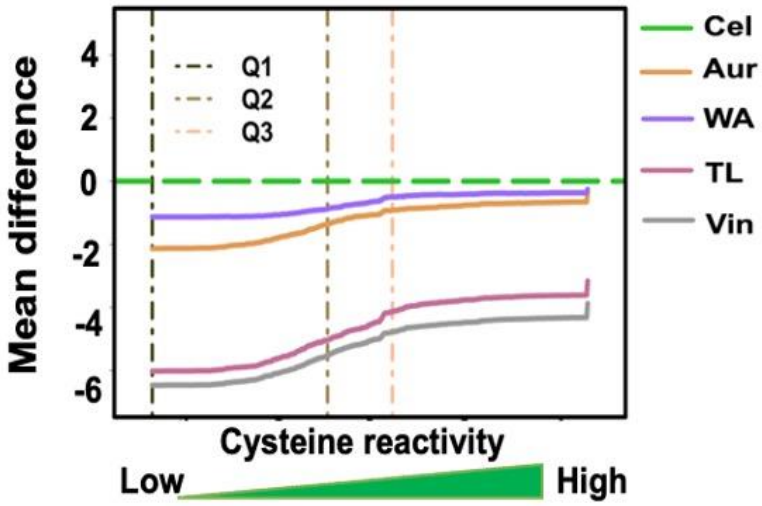


Figure 5-8 Evaluation of the cysteine-binding abilities of electrophiles with a mathematical measurement. Cel was set as the reference with y=0, and mean difference with a minus value showed lower cysteine-binding ability than the reference.

5.4 Conclusion

In this chapter, we first established a cysteine profiling method for target identification of electrophiles, by using two cysteine-reactive probes simultaneously. We selected Cel as a representative electrophilic bioactive compound, and other three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their cellular perturbations of global transcription.

Our results demonstrated that the cysteine reactivity of electrophilic compounds associated with their influence on gene transcription. We reasoned that such correlation can be potentially adopted to develop a novel method for multiplexing the target identification of compounds.

Chapter 6: Target validation of electrophiles with targeted MS

6.1 Introduction

Shotgun MS is the most widespread technique of proteome analysis, in which proteins are digested into peptides, optionally separated by chromatogram and finally determined by mass spectrometers[146, 147]. Although this approach can certainly provide high throughput with dynamic range in MS detection, it has a bottleneck in sensitivity and robustness. Targeted MS is a promising fervor that can address the drawbacks of the shotgun MS method. For instance, selected reaction monitoring (SRM) is a typical targeted MS technique where a particular ion and a product ion fragmented from the precursor ion are both selected for MS detection[148]. With the development of a quadrupole-Orbitrap mass spectrometer, a targeted proteomics strategy called parallel reaction monitoring (PRM) is generated where all fragments from the targeted peptide are monitored in parallel based on high-resolution MS[149-151]. PRM can serve as a validation tool for the identified candidates or biomarkers from discovery-based proteomics, with high specificity to confirm the identification[152, 153].

6.2 Materials and methods

6.2.1 Cell culture

PC3 cells were grown in Ham's F12K medium supplemented with 10% HIFBS, 100 unit/mL penicillin and 100 μg/mL streptomycin. Cells were cultured at 37°C with 5% CO_2.

6.2.2 PRM analysis

PRM analyses were performed on the same LC and MS system as described in section 5.2.9. A RSLC C18 analytical column (75 μm × 250 mm, 1.6 μm, 120 Å) (Aurora, Ion opticks) was employed for LC separation. A 120-min gradient at a flow rate of 300 nL/min and an initial 8% mobile phase B was used. Mobile phase B was increased to 10% at 2 mins, 32% at 82 mins, 90% at 100 mins and held for 8 mins. Then, mobile phase B was back to 8% at 110 min and maintained this composition until 120 min. PRM acquisition methods were directly developed based on our DDA data. The DDA data was imported into Skyline to select targeted peptides. The relative information including precursor *m/z*, charge, and retention time window of the selected peptides was exported from Skyline into Xcalibur software to edit the

PRM method. The targeted MS1 parameters were as follows: resolution, 120,000; AGC target, 3.0×10^5; and maximum injection time, 100 ms. PRM scanning was performed at 60,000 resolution, 1×10^5 AGC target, maximum injection time was set as auto mode and 1.0 *m/z* isolation window. Fragmentation was performed with a normalised collision energy of 30.

6.2.3 PRM data process with Skyline

PRM data was imported into Skyline (version: 21.2.0.369) for data processing. The spectral library was built from the DDA data and the spectra of precursors were manually inspected. Good results can be achieved with sharp, defined peaks without coelution. To get reliable identification and quantification, idotp (isotope dot product) and dotp (dot product) should be above 0.80. The top three product ions were summed up to represent the peptide abundance.

6.3 Results and discussion

6.3.1 Benchmark the method: reproducible identification of reported drug-target interactions

It is noteworthy that reported target proteins were identified with our approach. Here we selected six reactive cysteines from reported target proteins to further investigate our strategy. Their cysteine reactivities identified from our cysteine profiling results were summarised in **Table 6-1**. The percentage reduction of probe labelling by electrophiles is calculated for each cysteine as follows: % inhibition = 100 (1-$_{Ielectrophile}$/I_{DMSO}). $I_{electrophile}$: Intensity of electrophile treatment. I_{DMSO}: Intensity of DMSO treatment.

As a proof-of-concept of our approach, we next applied targeted MS on reactive cysteines by PRM strategies to validate our cysteine profiling results.

Firstly, we verified interactions between electrophiles and known binding sites in reported target proteins. For examples, Cys40 of prostaglandin E synthase 3 (PTGES3) was reported as one of the Cel-modified cysteines to disrupt PTGES3 function[154]. Cys83 and Cys173 of peroxiredoxin-1 (PRDX1) have been reported as binding sites of TL[155]. As shown in Figure 3A, compared to the DMSO group, reduced intensities were observed after treatment with interacting compounds.

Our chemical proteomics results revealed that Cel, Aur and WA showed similar cysteine reactivities (**Table 6-1**). Our results (**Figure 6-1**) again verified the bindings of Cel on Cys40 of PTGES3, TL on Cys83 and Cys173 of PRDX1.

Table 6-1 Cysteine reactivities of six identified cysteines from reported target proteins, represented by %inhibition. The cysteine with higher %inhibition represents greater reactivity to the indicated electrophiles.

	Cel	Aur	WA	TL	Vin
Cys40_PTGES3	71.0	75.6	72.2	31.5	15.2
Cys83_PRDX1	60.5	47.2	58.0	48.4	14.3
Cys173_PRDX1	69.3	40.7	55.5	50.5	12.0
Cys564_Hsp90β	40.3	64.2	31.0	6.6	8.0
Cys328_Vimentin	86.4	87.0	81.3	10.4	3.7
Cys100_PRDX5	56.0	83.1	80.0	28.2	14.5

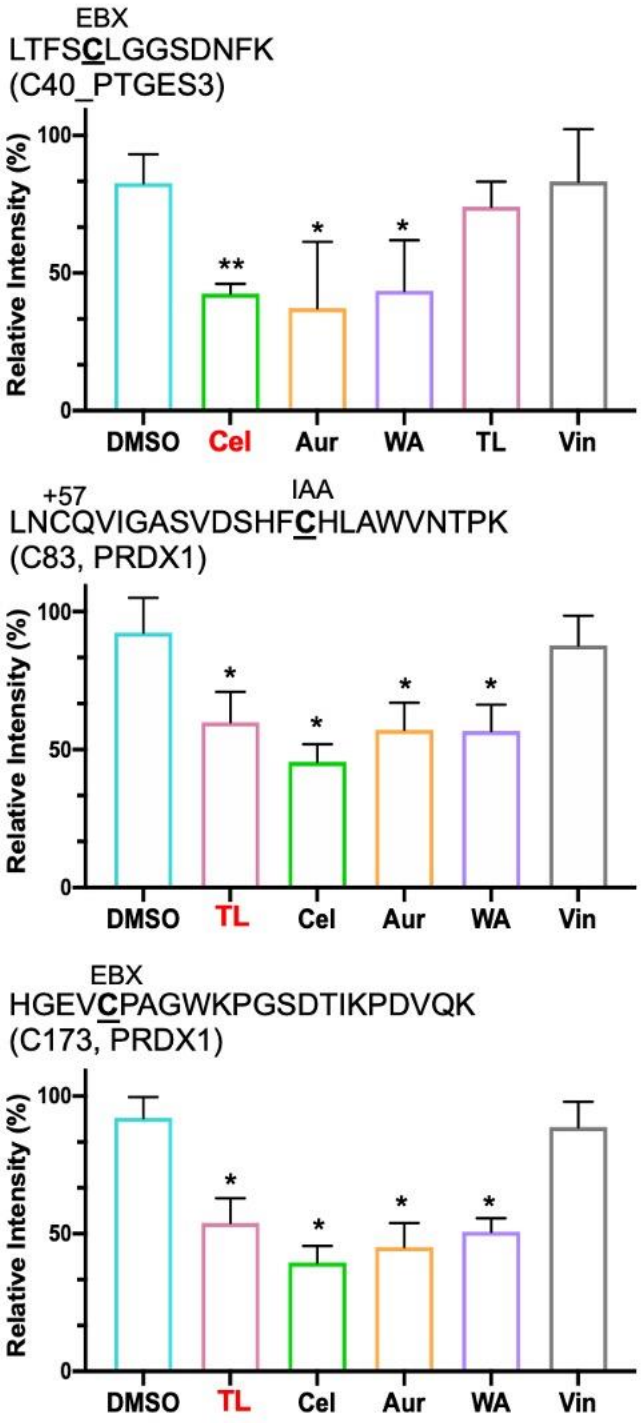


Figure 6-1 Detection of known interactions between electrophiles and proteins with targeted MS. Detected interactions between electrophiles and known binding sites in reported target proteins. Red colour highlights known ligands of the binding sites. Data were presented as

mean values ± SD, n = 3. Statistical significance was assessed via unpaired Student's t-test (two-tailed) comparing to DMSO group, *p<0.05. **p<0.01.

6.3.2 Novel binding sites of electrophiles

As for the known targets with unknown cysteine binding sites, for examples, as shown in **Figure 6-2**, vimentin was reported as the protein target of WA[156], peroxiredoxin-5 (PRDX5) was annotated as the protein target of Aur[157], and C-terminal domain of heat shock protein 90β (Hsp90β) was predicted as the target of several Hsp90β inhibitors such as Cel[158] and WA[159]. Based on our cysteine profiling results, we identified Cys564 of Hsp90β, Cys328 of vimentin and Cys100 of PRDX5 as reactive cysteines bound with the tested electrophiles. Targeted MS results further demonstrated that Cel and WA interacted with Cys564 of Hsp90β, WA, Cel and Aur interacted with Cys328 of vimentin, Aur, Cel and WA interacted with Cys100 of PRDX5, which are consistent with the cysteine profiling results (**Table 6-1**). Overall, targeted MS results (**Figure 6-1 and 6-2**) showed consistent with the cysteine profiling results (**Table 6-1**).

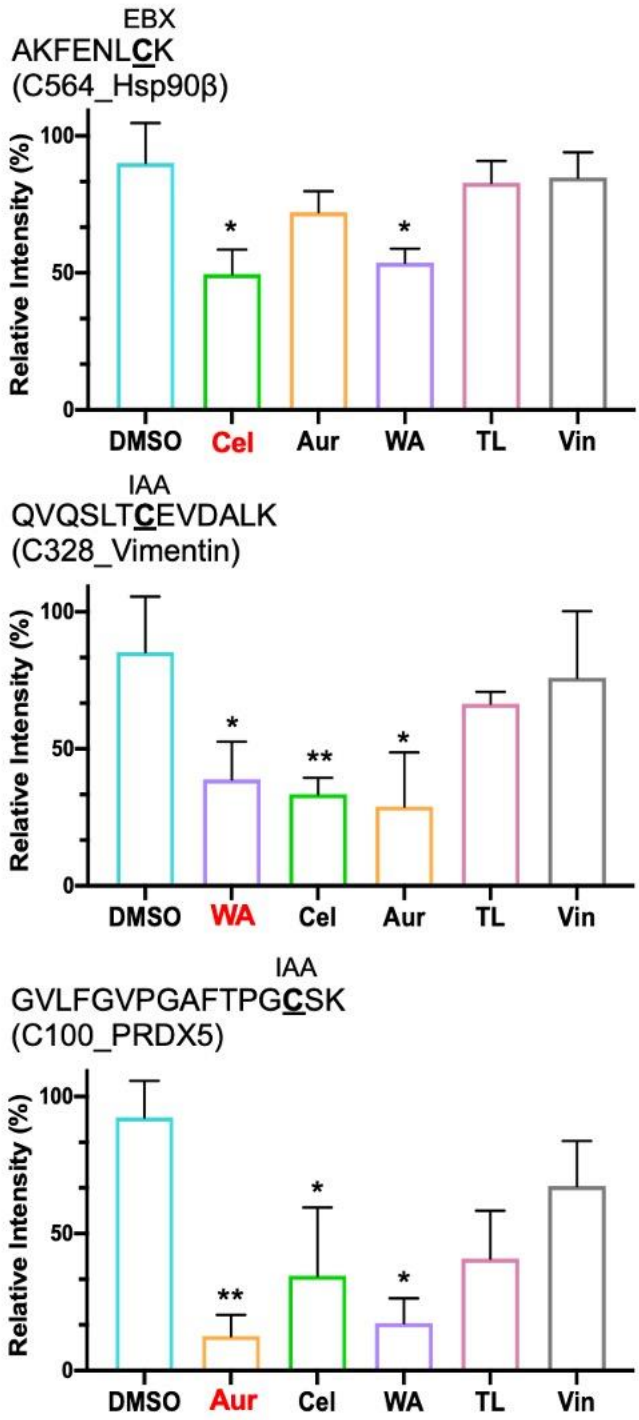


Figure 6-2 Detection of novel interactions between electrophiles and proteins with targeted MS. Detected interactions between electrophiles and novel binding sites of target proteins. Red colour highlights reported ligands of the indicated proteins. Data were presented as mean

values ± SD, n = 3. Statistical significance was assessed via unpaired Student's t-test (two-tailed) comparing to DMSO group, *p<0.05. **p<0.01.

6.4 Conclusion

Targeted MS results demonstrated that Cel, Aur and WA shared common targets, not only corresponded to known binding sites like Cys83 and Cys173 of PRDX1, but also provided clues for novel binding site information. Overall, targeted MS verified our strategy of chemical proteomics-genomics-based target profiling, demonstrating its applicability for multiplexed target profiling of electrophiles.

Chapter 7: Verification of electrophile-target protein (PRDX1) interactions identified with the novel method

7.1 Background

PRDX1 was selected as the first model protein for validation assays. PRDX1 encodes a member of the peroxiredoxin family of antioxidant enzymes[160]. Zhao *et al.*[155] identified Cys83 and Cys173 of PRDX1 as the binding sites of Cel, WA and TL in a covalent manner. TL was demonstrated to selectively inhibit the chaperone activity of PRDX1 but not its peroxidase activity. In addition, Cel and WA showed more potent than TL in the chaperone activity assay. Based on our MS results (Cysteine profiling MS in Chapter 5 and targeted MS in Chapter 6), we found interactions of Cel, Aur and WA with Cys83 or Cys173 of PRDX1. Here, we further employed biochemical assays to verify our findings in PRDX1.

7.2 Materials and methods

7.2.1 Chemicals and reagents

Isopropy-β-D-thiogalactoside (IPTG) was ordered from Sigma-Aldrich. Ampicillin was purchased from Thermo Fisher Scientific. Competent cells (E. coli BL21 (DE3)) and Mini-Prep kit (DP103-03) were ordered from TianGen. Antibody of PRDX1 (sc-137222) was from Santa Cruz Biotechnology.

7.2.2 PRDX1 plasmid construction, protein expression and purification

Plasmids with full length genes encoding PRDX1, PRDX1_C83S and PRDX1_C173S were from our own collection and transformed into E. coli BL21 (DE3). The single colony was taken up and injected into LB medium that was made with 100 g/mL ampicillin after overnight culture on an agar plate at 37 °C. Next, overnight cultures were amplified with fresh LB medium containing 100 μg/mL ampicillin. Cells were subsequently cultured 0.2 mM IPTG at 20 °C for 20 hours. Cell pellets were lysed by sonication in lysis buffer (100 mM Tris-HCl, pH=8.0, 1 mM EDTA and 150 mM NaCl) at 4 °C. The lysates were filtered and transfered to 5 mL Strep-Tactin® XT columns (Beyotime Biotechnology). The protein was eluted out by lysis buffer with 50 mM

biotin. The elutions were concentrated and injected to Superdex 75 (GE Healthcare) for further purification. The fractions involved target protein PRDX1 was combined for the subsequent studies.

7.2.3 UV-visible absorption assay of Cel

50 μM Cel in PBS was incubated with wild-type or mutated PRDX1 for 30 mins at RT. Cel incubated with PBS or DTT served as the positive control and the negative control, respectively. The absorption of Cel was measured at 440 nm.

7.2.4 Fluorescence gel imaging using probe Cy3-TL in PRDX1 and its mutants

Wild-type PRDX1 (5 μM) and the mutant-PRDX1-C83S (5 μM), mutant-PRDX1-C173S (5 μM) were incubated with 50 μM compounds for 4 hours at 4 ^{o}C. Probe Cy3-TL was then added to a final concentration of 10 μM for additional 0.5 hour at RT and subsequently resolved by SDS-PAGE. Gels were scanned for fluorescently labelled proteins (Typhoon Imaging System, GE), then stained with Coomassie blue.

7.2.5 Thermal shift assay

PC3 cells were treated with 10 μM compounds or vehicle for 1 hour at 37°C and then rinsed two times with 1 × PBS. Cell pellets were lysed by sonication in cold 1 × PBS with 1 × protease inhibitor (Roche), pH 7.4. The sonifier (SFX 550, Branson) worked with a pulsed regime for 2 min (10 second sonication, 15 second pause) at 30% amplitude under ice cooling. After centrifugation at 16,000 g, 4 ºC for 20 min, the supernatant fraction was collected and adjusted to 1.0 μg/μL by BCA assay. After heating at the indicated range of temperatures using a thermal cycler, the lysates were centrifuged at 15,000 g for 20 min at 4 °C and subsequently resolved by SDS-PAGE for western blotting.

7.3 Results and discussion

7.3.1 Competitive in-gel fluorescence assay

Firstly, we adopted biochemical assays to validate the aforementioned electrophile-target protein interactions, using PRDX1 as the first model protein.

As shown in **Figure 7-1**, in-gel fluorescence scan revealed that PRDX1 was marginally labelled by neither IAA nor EBX after incubation with a low concentration (5 μM) of Cel or WA, indicating strong compound-cysteine interactions. On the other hand, Aur and TL showed a concentration-dependent manner of interaction, presenting strong binding competitiveness only at a high concentration (50 μM). Consistent with our targeted-MS result, Vin marginally inhibited the probe-labelling cysteines of PRDX1 even at high concentration.

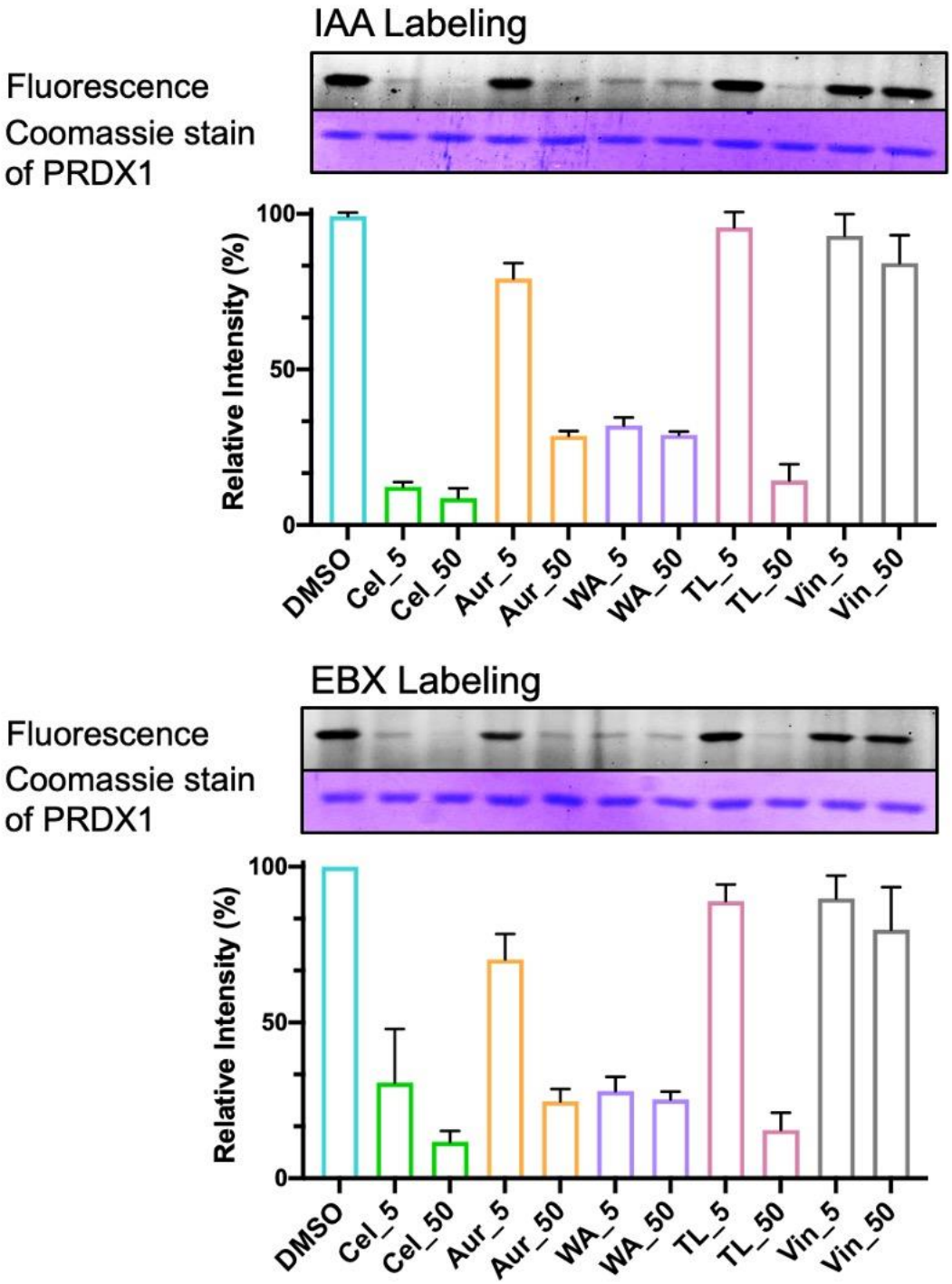


Figure 7-1 In-gel fluorescence scan of recombinant PRDX1 labelled with IAA or EBX probe, in presence or absence of the indicated

compounds as competitors. A low concentration (5 μM) and a high concentration (50 μM) of compounds were used.

To verify the binding cysteines of the tested compounds, we subsequently carried out a *in-vitro* competition experiment using wild-type and mutant recombinant PRDX1 and Cy3-TL (**Figure 7-2**). Cy3-TL was used as a specific probe to label PRDX1 and conjugate to fluorophore for visualization of the interactions. Cel, Aur, WA and TL were found to inhibit the probe-labelled band of wild-type PRDX1 while Vin did not show any inhibition. After Aur and TL treatment, the C83S mutant of PRDX1 showed increased labelling intensity compared to that of wild-type PRDX1, indicating Cys83 in PRDX1 as binding target of Aur and TL. Similarly, after Cel, Aur, WA and TL treatment, the C173S mutant of PRDX1 showed increased labelling compared to wild-type PRDX1, indicating Cys173 in PRDX1 as binding target of Cel, Aur, WA and TL. This result illustrated that both Cys83 and Cys173 play important roles in the formation of the PRDX1-compound complex.

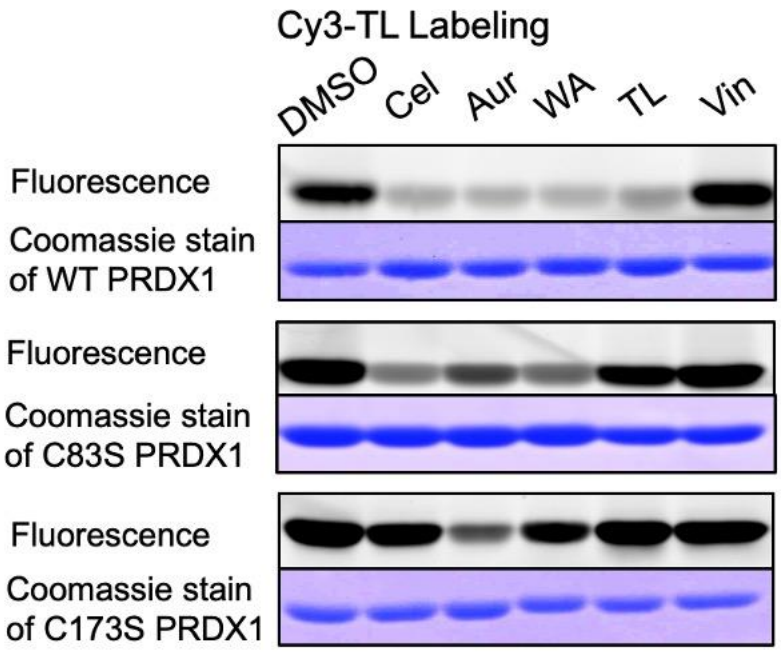


Figure 7-2 In-gel fluorescence scan of recombinant PRDX1 and its mutants labelled with Cy3-TL probe in presence or absence of the indicated compounds as competitors.

7.3.2 Validation of the compound-target interaction with CETSA

Additionally, in the thermal shift assay (**Figure 7-3**), Cel, Aur, WA and TL efficiently stabilised PRDX1 within the tested temperature range, whereas the vehicle (DMSO) and Vin failed to do so, denoting direct binding of Cel, Aur, WA and TL to PRDX1.

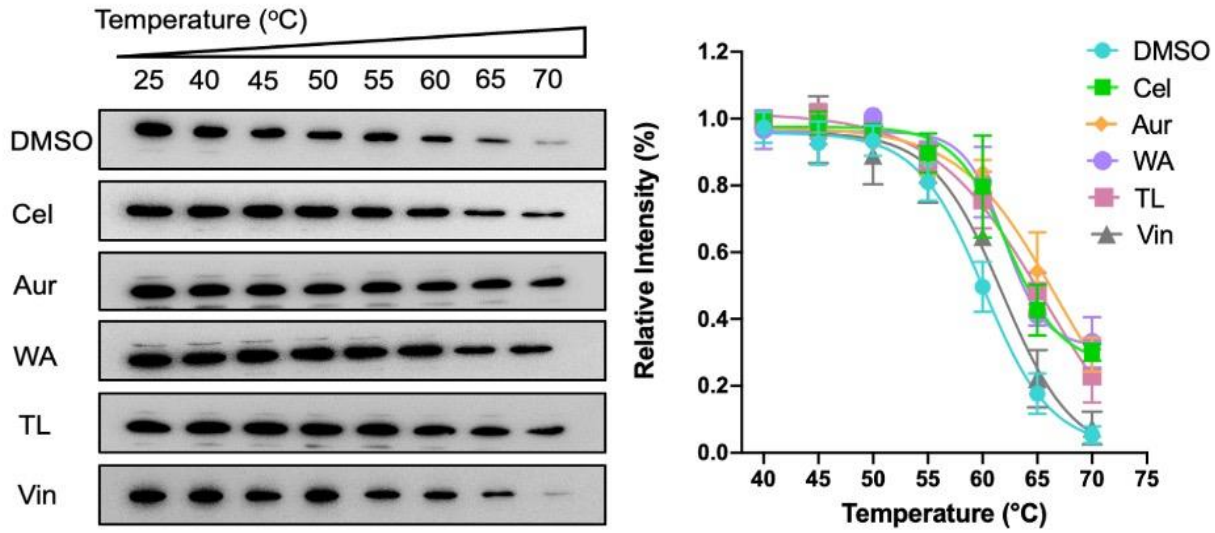


Figure 7-3 CETSA shows the thermostability change of PRDX1 after incubation with the indicated compounds or vehicle (right panel). One representative western blot was shown (left panel) with quantification based on three independent experiments.

7.3.3 Validation of the Cel-target interaction with the UV absorption assay

Besides, a UV-visible assay showed a decrease in Cel's absorption at 440 nm after addition of PRDX1, while an increase in absorbance after addition of two mutants C83S and C173S PRDX1, indicating covalent binding of Cel on Cys83 and Cys173 of PRDX1 (**Figure 7-4**).

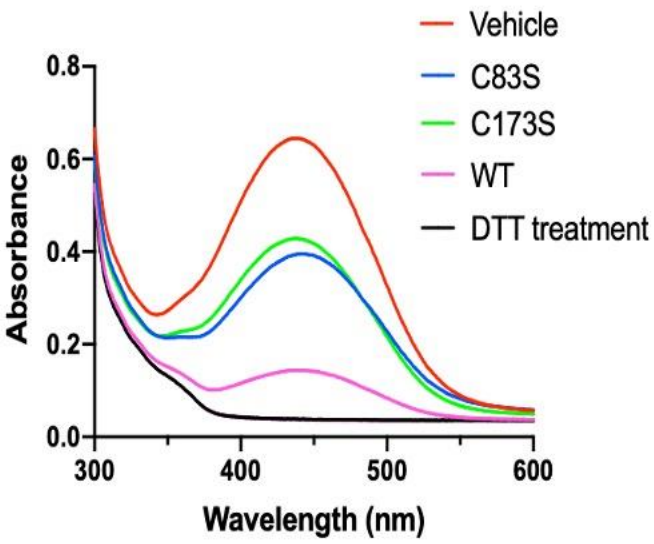


Figure 7-4 UV absorption of Cel after incubation with recombinant PRDX1, its mutants or DTT. The decrease in absorption indicates interaction and disruption of the chromophore of Cel.

7.4 Conclusion

By the competitive in-gel fluorescence scan, CETSA and the UV absorption assay in recombinant PRDX1 and its mutants, we verified several interactions between electrophiles and PRDX1, again demonstrating that Cel, Aur and WA shared common targets in PRDX1.

Chapter 8: Verification of electrophile-target protein (Hsp90β) interactions identified with the novel method

8.1 Background

Hsp90β was selected as the second model protein for validation assays. Hsp90β is a member of the molecular chaperone Hsp90 family and regulates the refolding of proteins, cellular homeostasis and maturation of proteins in signal transduction[161-163]. Zhang *et al.* [158] found that Cel may represent a novel Hsp90 inhibitor by modifying the C-terminal domain of Hsp90 to regulate its chaperone activity. Yu *et al.*[159] demonstrated that WA can bind to Hsp90β to inhibit its chaperone activity through an ATP-independent mechanism. However, the active site of Hsp90β that binds with Cel or WA is unknown. Based on our MS results (Cysteine profiling MS in Chapter 5 and targeted MS in Chapter 6), we found interactions of Cel and WA on Cys564 of Hsp90β. Here, we further employed biochemical assays to verify our findings in Hsp90β.

8.2 Materials and methods

8.2.1 Chemicals and reagents

IPTG was ordered from Sigma-Aldrich. Ampicillin was purchased from Thermo Fisher Scientific. Competent cells (E. coli BL21 (DE3)) and Mini-Prep kit (DP103-03) were ordered from TianGen.

8.2.2 Plasmid construction, protein over expression and purification of Hsp90β

Plasmids with full-length genes encoding Hsp90β and Hsp90β-C564S were from our own collection and transformed into E. coli BL21 (DE3). The procedures for protein expression and purification are as mentioned in section 7.2.2. The purified protein Hsp90β was combined for the subsequent studies.

8.2.3 Fluorescence gel imaging using probe IAA or EBX in Hsp90β and its two mutants

WT Hsp90β (5 μM) and the mutant-Hsp90β-C564S (5 μM) were incubated with 50 μM compounds for 4 hours at 4 ^{o}C. Probe IAA or EBX was then added with a final concentration of 2 μM for additional 0.5 hour at RT and subsequently resolved by SDS-PAGE. Gels were

scanned for fluorescently labelled proteins (Typhoon Imaging System, GE), then stained with Coomassie blue.

8.2.4 Molecular docking

Hsp90β molecular chaperone was picked up as the docking receptor from its Cro-EM structure with its co-chaperones (Cdc3, Cdk4) (PDB code: 5FWM). Ligand complexes (Cel and WA) were presented in ChemBioOffice (version 13) and preformed the MOE quick-preparation program before docking. While all hydrogen atoms were being added, all incorrect charges and broken bonds were corrected. The parameters and charges were consigned with the MMFF94x force field and an induced fit model based on covalent docking was selected in followed docking studies. The sulfur atom of Cys564 of Hsp90β was selected as the docking atom and a related reaction was drawn according to the standard MOE reaction method. All rotamer libraries of Cys564 and ligands were considered to see more possible combinations of the docking poses. The highest ten docking scores were selected. A similar docking program was preformed between ribosomal protein S6 kinase alpha-3 and its covalent bonding ligand

(PDB code: 4JG7) as the benchmark for comparison. All docking runs were also performed with MOE2020.9.

8.3 Results and discussion

8.3.1 Competitive in-gel fluorescence assay

Similarly, regarding Hsp90β, it was indicated that Cel was able to compete the signal of labelled cysteine at a low concentration (5 μM) regardless of what kind of probe we used (**Figure 8-1**). The competition of Aur and WA was concentration-dependent, showing competition at a high concentration (50 μM). TL and Vin did not inhibit the probe-labelling cysteines of Hsp90β even at a high concentration.

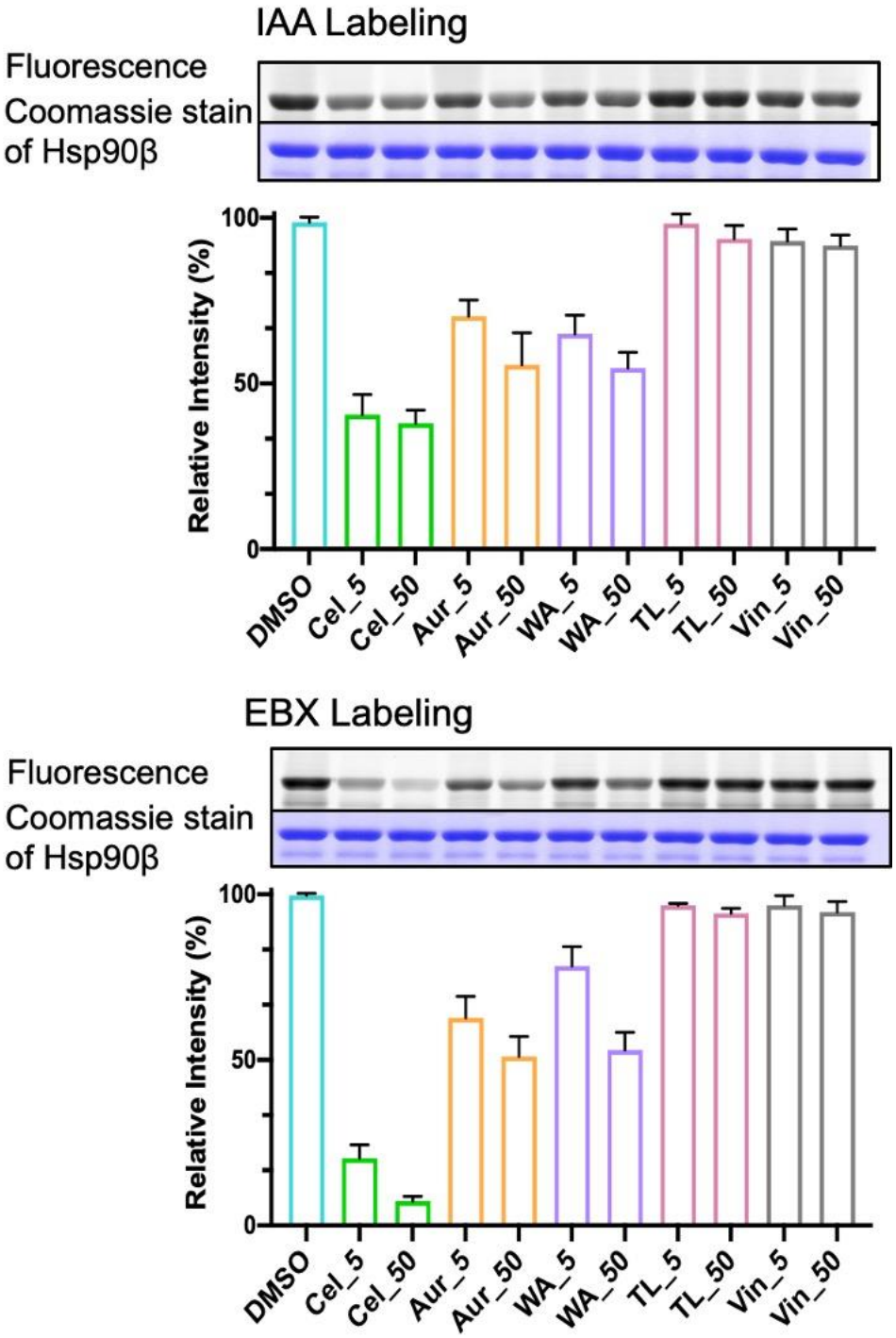


Figure 8-1 In-gel fluorescence scan of recombinant Hsp90β treated with indicated compounds at 5 μM or 50 μM followed by IAA or EBX labelling.

To verify the binding cysteine of the tested compound in Hsp90β, we also carried out a competition experiment using wild-type and mutant recombinant Hsp90β (**Figure 8-2**). After Cel or WA treatment, the C564A Hsp90β showed substantially increased labelling intensity compared to wild-type Hsp90β, while the signal after Aur treatment remained at a low level. These were consistent with our targeted MS results, which together suggested Cel and WA can target Cys564 of Hsp90β, while Aur may not mainly target Cys564 of Hsp90β. The results denoted that Cys564 played important roles in the formation of the Hsp90β–compound complex.

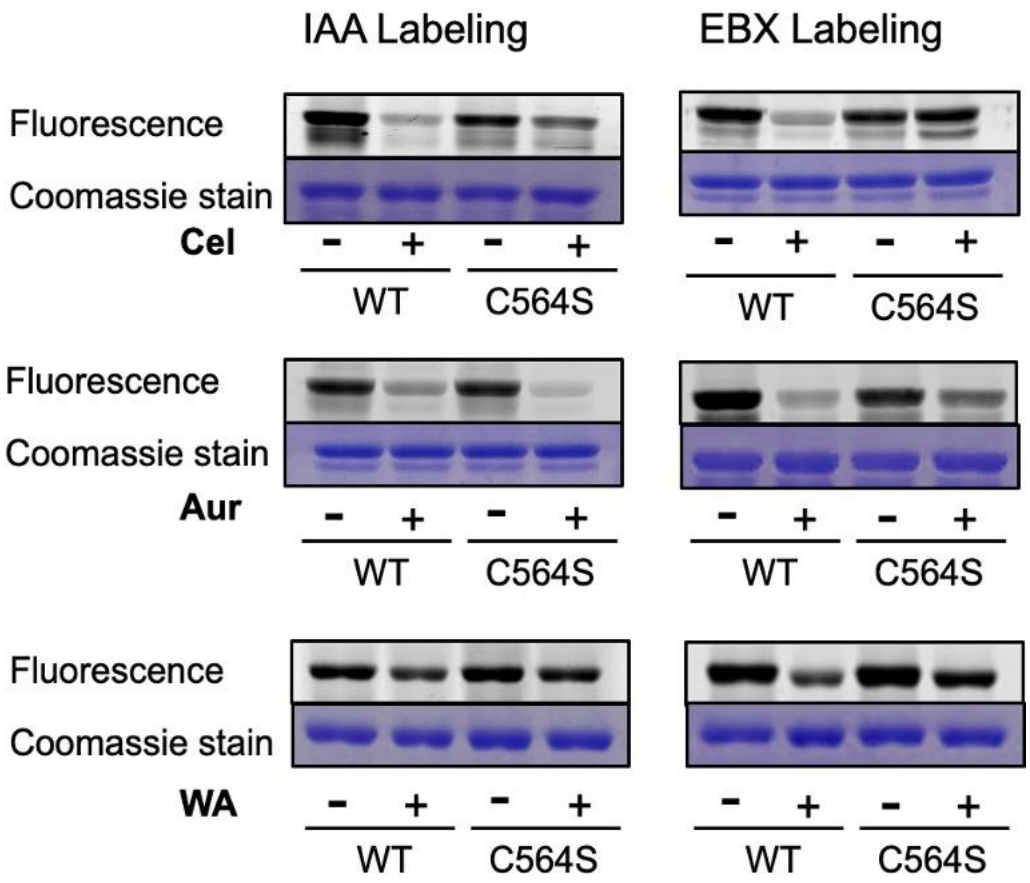


Figure 8-2 In-gel fluorescence scan of recombinant Hsp90β and its mutants treated with indicated electrophiles followed by probe labelling.

8.3.2 Validation of the compound-target interaction with the molecular docking

We further revealed the selectivity mechanism of Cel and WA adduct to Cys564 in Hsp90β by molecular docking (**Figure 8-3**). We found that both Cel and WA possessed low docking scores (around -5) with

Cys564 in Hsp90β, revealing the probability of ligand interactions, which may contribute to the disruption of Hsp90β function.

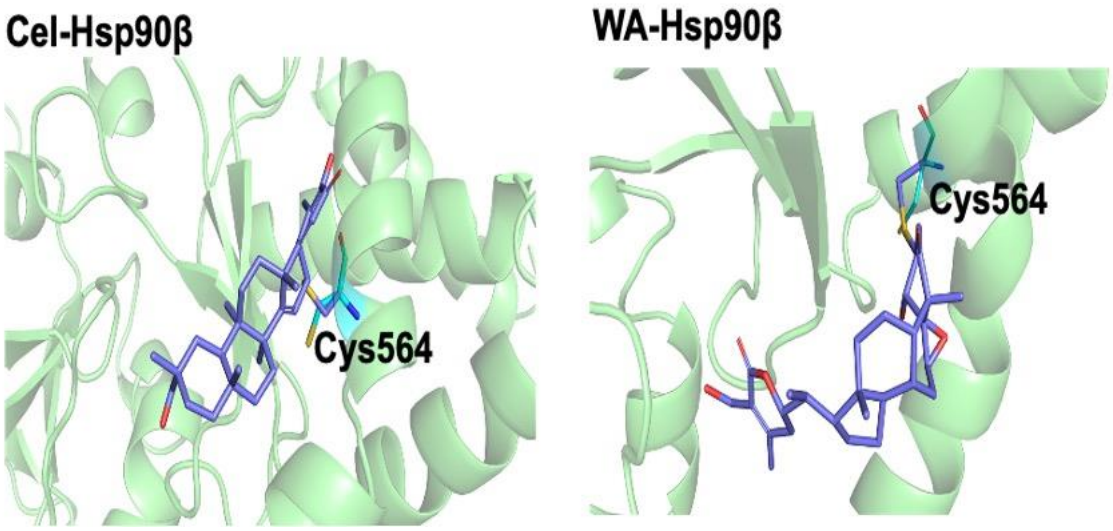


Figure 8-3 Molecular docking shows the interactions of Cel and WA with Cys564 of Hsp90β by forming a covalent bond.

8.3.3 Validation of the Cel-target interaction with the UV absorption assay

Moreover, the UV-visible assay result showed a decrease in Cel's absorption at 440 nm after addition of Hsp90β, while an increase in absorbance after addition of mutant Hsp90β (**Figure 8-4**), also indicated covalent binding of Cel with Cys564 of Hsp90β.

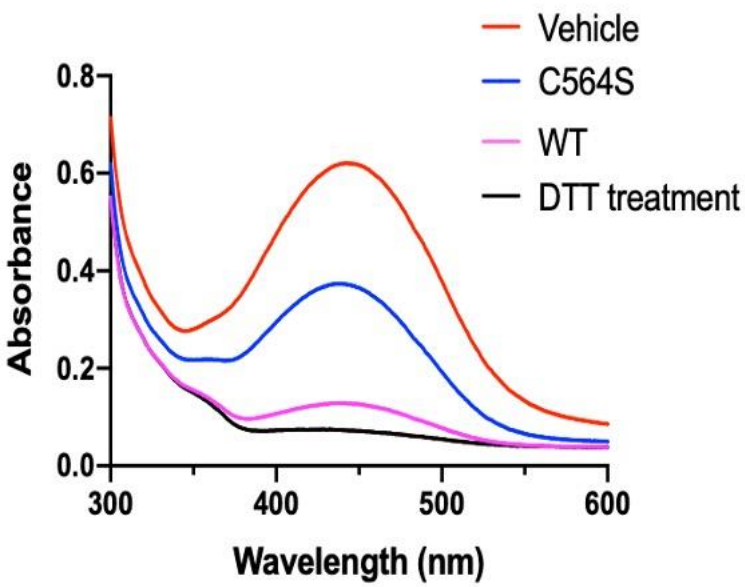


Figure 8-4 UV absorption of Cel after incubation with the indicated compound or proteins. Decrease of absorption indicates interaction with Cel and disruption of its chromophore.

8.4 Conclusion

By the competitive in-gel fluorescence scan, molecular docking and the UV absorption assay in recombinant Hsp90β and its mutants, we verified several interactions between electrophiles and Hsp90β, demonstrating Cys564 of Hsp90β may be the major binding site of Cel and WA.

Chapter 9: Conclusion and future perspective

9.1 Conclusion

First, we focused on Cel and identified COMT as one of its major binding targets. We synthesised chemical probes to mimic the natural product Cel and applied chemical proteomics for target identification. The interaction between Cel and COMT was validated with independent assays including targeted MS and immunostaining. Cel bound to both S-COMT and MB-COMT and inhibited the enzymatic activities of both isoforms. Unlike the first-generation and second-generation COMT inhibitors[164], Cel is able to inhibit COMT enzymatic activity by binding to Cys157 which is at one terminal of the α-helix (D141-C157) that bridges the SAM-binding pocket and substrate-binding pocket. The conversion of DA to its downstream metabolite 3-MT, a process catalysed by COMT, was inhibited by Cel in neuroendocrine chromaffin cells. The Cys157 residue of COMT can serve as a targetable hot spot for developing novel COMT inhibitors, highlighting the favourable properties of Cel for neuroprotection. (**Figure 9-1**).

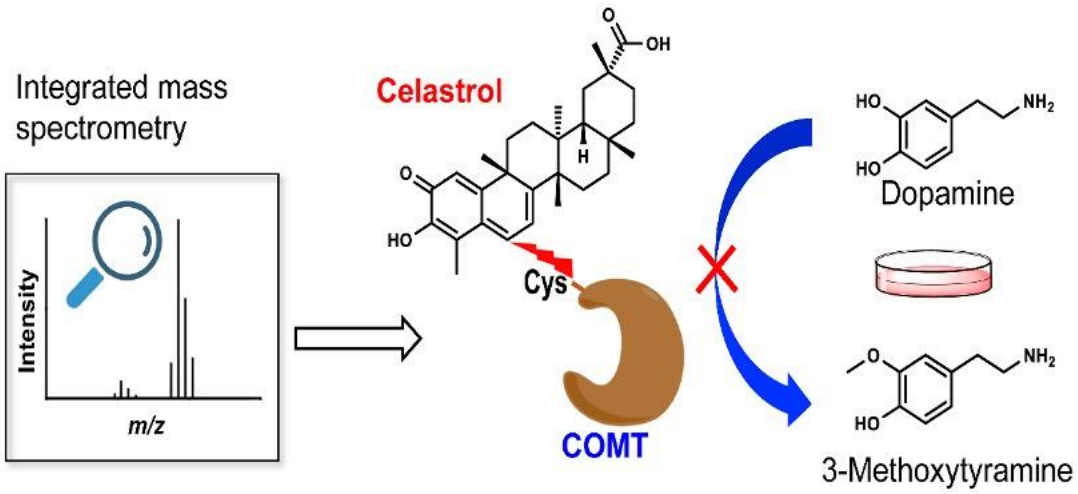


Figure 9-1 Integrated strategies reveal Cel as a novel COMT inhibitor.

Then, we carried out a proof-of-concept study for establishing a strategy for multiplexed target profiling of structurally distinct electrophiles. We selected Cel as a representative electrophilic bioactive compound, and other three structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their cellular perturbation of gene transcription. A cysteine profiling approach using two complementary chemical probes was developed to enable scrutiny of reactive cysteines as potential binding sites of tested electrophiles. In addition, target profiling of the four electrophilic compounds was performed by using the chemical

proteomics approach, leading to identification of known and novel binding targets. Furthermore, targeted MS as well as biochemical assays demonstrated that Cel, Aur and WA shared common targets, not only known interactions like TL on Cys83 of PRDX1, but also novel binding sites like WA on Cys564 of Hsp90β. Notably, we demonstrated that the cysteine reactivity of electrophilic compounds is associated with their influence on gene transcription. Therefore, the integrated approach combining genomics and chemical proteomics can serve as a high-throughput and easy-to-use sensor for multiplexed target identification of structurally distinct compounds. We strongly believe that the presented platform is of great prominence to facilitate future development of multiplexing target profiling of electrophilic compounds (**Figure 9-2**).

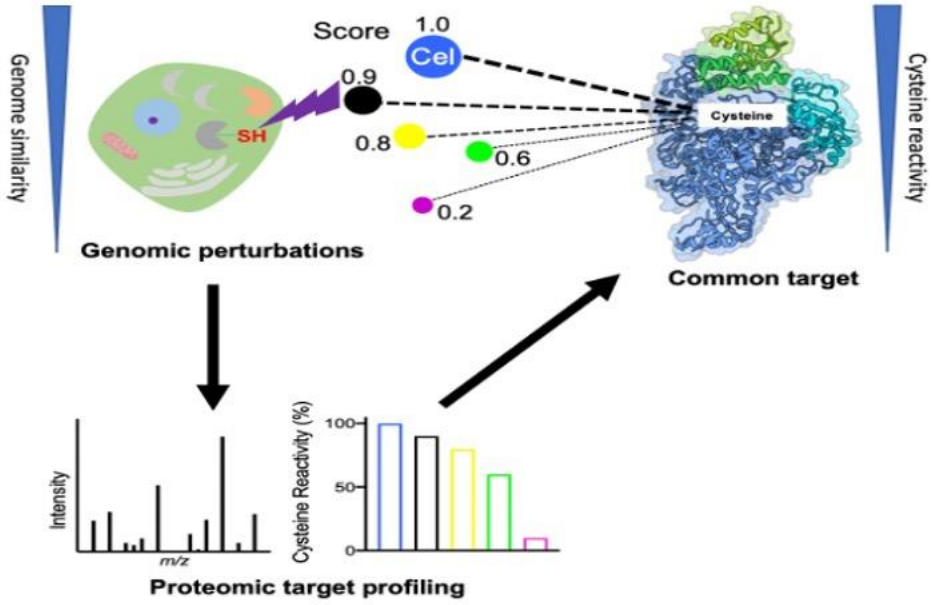


Figure 9-2 A chemical proteomics-genomics approach for multiplexed target profiling of electrophiles.

9.2 Future perspective

Projecting forward, we anticipate the presented platform to show great potential and facilitate future development of target profiling of electrophilic compounds.

Chemical probes and cleavable linkers are important tools in chemical biology, constantly evolving and developing to achieve the identification of modification sites of target proteins. Our compound-

centric chemical proteomics as well as chemical proteomics-genomics strategies are feasible to employ specific chemical probes and cleavable linkers to capture target proteomes.

Besides cysteine-reactive probes, our strategy has showed the availability in kinase mapping by kinase probes, nascent protein profiling by O-propargyl-puromycin probes, and has obtained preliminary progress with our collaborators. Novel-designed probes for protein target profiling and/or site mapping are drawing our interest. Furthermore, the binding stoichiometry of target protein is critical in understanding of mode of action of the tested compounds. Like the binding stoichiometry study of Cel and COMT in our study, several independent assays including intact MS, tandem MS, and in-gel fluorescence scan in recombinant proteins can provide evidence for binding site identification. The integrated strategies with specific chemical probes as well as cleavable linkers enable precise mapping of the direct binding sites of compounds.

References

1. Ha, J.; Park, H.; Park, J.; Park, S. B., Recent advances in identifying protein targets in drug discovery. *Cell Chem Biol* **2021,** *28* (3), 394-423.

2. Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A., Target identification and mechanism of action in chemical biology and drug discovery. *Nature Chemical Biology* **2013,** *9* (4), 232-240.

3. Rix, U.; Superti-Furga, G., Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* **2009,** *5* (9), 616-24.

4. McCloud, R. L.; Yuan, K.; Mahoney, K. E.; Bai, D. L.; Shabanowitz, J.; Ross, M. M.; Hunt, D. F.; Hsu, K.-L., Direct Target Site Identification of a Sulfonyl–Triazole Covalent Kinase Probe by LC-MS Chemical Proteomics. *Analytical Chemistry* **2021,** *93* (35), 11946-11955.

5. Spreafico, R.; Soriaga, L. B.; Grosse, J.; Virgin, H. W.; Telenti, A., Advances in Genomics for Drug Development. *Genes (Basel)* **2020,** *11* (8).

6. Jost, M.; Weissman, J. S., CRISPR Approaches to Small Molecule Target Identification. *ACS Chemical Biology* **2018,** *13* (2), 366-375.

7. Katsila, T.; Spyroulias, G. A.; Patrinos, G. P.; Matsoukas, M.-T., Computational approaches in target identification and drug discovery. *Computational and Structural Biotechnology Journal* **2016,** *14*, 177-184.

8. Sydow, D.; Burggraaff, L.; Szengel, A.; van Vlijmen, H. W. T.; Ijzerman, A. P.; van Westen, G. J. P.; Volkamer, A., Advances and Challenges in Computational Target Prediction. *Journal of Chemical Information and Modeling* **2019,** *59* (5), 1728-1742.

9. Fedorov, II; Lineva, V. I.; Tarasova, I. A.; Gorshkov, M. V., Mass Spectrometry-Based Chemical Proteomics for Drug Target Discoveries. *Biochemistry (Mosc)* **2022,** *87* (9), 983-994.

10. Wright, M. H.; Sieber, S. A., Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat Prod Rep* **2016,** *33* (5), 681-708.

11. Bar-Peled, L.; Kemper, E. K.; Suciu, R. M.; Vinogradova, E. V.; Backus, K. M.; Horning, B. D.; Paul, T. A.; Ichu, T.-A.; Svensson,

R. U.; Olucha, J.; Chang, M. W.; Kok, B. P.; Zhu, Z.; Ihle, N. T.; Dix, M. M.; Jiang, P.; Hayward, M. M.; Saez, E.; Shaw, R. J.; Cravatt, B. F., Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* **2017,** *171* (3), 696-709.e23.

12. Drewes, G.; Knapp, S., Chemoproteomics and Chemical Probes for Target Discovery. *Trends Biotechnol* **2018,** *36* (12), 1275-1286.

13. Pan, S.; Zhang, H.; Wang, C.; Yao, S. C.; Yao, S. Q., Target identification of natural products and bioactive compounds using affinity-based probes. *Natural product reports* **2016,** *33* (5), 612-620.

14. Liu, G.-Y.; Nie, S.; Zheng, X.; Li, N., Activity-Based Protein Profiling Probe for the Detection of Enzymes Catalyzing Polysorbate Degradation. *Analytical Chemistry* **2022,** *94* (24), 8625-8632.

15. Verhelst, S. H. L.; Bonger, K. M.; Willems, L. I., Bioorthogonal Reactions in Activity-Based Protein Profiling. *Molecules* **2020,** *25* (24).

16. Benns, H. J.; Wincott, C. J.; Tate, E. W.; Child, M. A., Activity- and reactivity-based proteomics: Recent technological advances and applications in drug discovery. *Current Opinion in Chemical Biology* **2021,** *60*, 20-29.

17. Parker, C. G.; Pratt, M. R., Click Chemistry in Proteomic Investigations. *Cell* **2020,** *180* (4), 605-632.

18. Zhu, L.; Brassard, C. J.; Zhang, X.; Guha, P. M.; Clark, R. J., On the Mechanism of Copper(I)-Catalyzed Azide-Alkyne Cycloaddition. *Chem Rec* **2016,** *16* (3), 1501-17.

19. Yao, T.; Xu, X.; Huang, R., Recent Advances about the Applications of Click Reaction in Chemical Proteomics. *Molecules* **2021,** *26* (17).

20. Lebraud, H.; Wright, D. J.; East, C. E.; Holding, F. P.; O'Reilly, M.; Heightman, T. D., In-gel activity-based protein profiling of a clickable covalent ERK1/2 inhibitor. *Molecular BioSystems* **2016,** *12* (9), 2867-2874.

21. Litwin, K.; Crowley, V. M.; Suciu, R. M.; Boger, D. L.; Cravatt, B. F., Chemical proteomic identification of functional cysteines with atypical electrophile reactivities. *Tetrahedron Letters* **2021,** *67*, 152861.

22. Wang, S.; Tian, Y.; Wang, M.; Wang, M.; Sun, G. B.; Sun, X. B., Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Front Pharmacol* **2018,** *9*, 353.

23. Tsuboi, K.; Bachovchin, D. A.; Speers, A. E.; Spicer, T. P.; Fernandez-Vega, V.; Hodder, P.; Rosen, H.; Cravatt, B. F., Potent and Selective Inhibitors of Glutathione S-Transferase Omega 1 That Impair Cancer Drug Resistance. *Journal of the American Chemical Society* **2011,** *133* (41), 16605-16616.

24. Marchenko, M.; Thomson, A.; Ellis, T. N.; Knuckley, B.; Causey, C. P., Development of a clickable activity-based protein profiling (ABPP) probe for agmatine deiminases. *Bioorganic & Medicinal Chemistry* **2015,** *23* (9), 2159-2167.

25. Liu, Y.; Guo, M., Chemical proteomic strategies for the discovery and development of anticancer drugs. *Proteomics* **2014,** *14* (4-5), 399-411.

26. Wang, K.; Yang, T.; Wu, Q.; Zhao, X.; Nice, E. C.; Huang, C., Chemistry-based functional proteomics for drug target deconvolution. *Expert Review of Proteomics* **2012,** *9* (3), 293-310.

27. Baker, M. A.; Schneider, E. K.; X. Huang, J.; Cooper, M. A.; Li, J.; Velkov, T., The Plasma Protein Binding Proteome of Ertapenem: A Novel Compound-Centric Proteomic Approach for Elucidating

Drug–Plasma Protein Binding Interactions. *ACS Chemical Biology* **2016,** *11* (12), 3353-3364.

28. Zhang, Y.; Wen, J.; Liu, D.; Qiu, Z.; Zhu, Q.; Li, R.; Zhang, Y., Demethylenetetrahydroberberine alleviates nonalcoholic fatty liver disease by inhibiting the NLRP3 inflammasome and oxidative stress in mice. *Life Sciences* **2021,** *281*, 119778.

29. Wang, W.; Tekcham, D. S.; Yan, M.; Wang, Z.; Qi, H.; Liu, X.; Piao, H.-L., Biochemical reactions in metabolite-protein interaction. *Chinese Chemical Letters* **2018,** *29* (5), 645-647.

30. Heal, W. P.; Dang, T. H. T.; Tate, E. W., Activity-based probes: discovering new biology and new drug targets. *Chem. Soc. Rev.* **2011,** *40* (1), 246-257.

31. Li, Z.; Liu, K.; Xu, P.; Yang, J., Benchmarking Cleavable Biotin Tags for Peptide-Centric Chemoproteomics. *J Proteome Res* **2022,** *21* (5), 1349-1358.

32. Yang, Y.; Fonović, M.; Verhelst, S. H., Cleavable Linkers in Chemical Proteomics Applications. *Methods Mol Biol* **2017,** *1491*, 185-203.

33. Beard, H. A.; Korovesis, D.; Chen, S.; Verhelst, S. H. L., Cleavable linkers and their application in MS-based target identification. *Mol Omics* **2021,** *17* (2), 197-209.

34. Yang, Y.-Y.; Grammel, M.; Raghavan, A. S.; Charron, G.; Hang, H. C., Comparative Analysis of Cleavable Azobenzene-Based Affinity Tags for Bioorthogonal Chemical Proteomics. *Chemistry & Biology* **2010,** *17* (11), 1212-1222.

35. Rabalski, A. J.; Bogdan, A. R.; Baranczak, A., Evaluation of Chemically-Cleavable Linkers for Quantitative Mapping of Small Molecule-Cysteinome Reactivity. *ACS Chem Biol* **2019,** *14* (9), 1940-1950.

36. Phillips, N. J.; Vinaithirthan, B. M.; Oses-Prieto, J. A.; Chalkley, R. J.; Burlingame, A. L., Capture, Release, and Identification of Newly Synthesized Proteins for Improved Profiling of Functional Translatomes. *Mol Cell Proteomics* **2023,** *22* (3), 100497.

37. Speers, A. E.; Cravatt, B. F., A tandem orthogonal proteolysis strategy for high-content chemical proteomics. *J Am Chem Soc* **2005,** *127* (28), 10018-9.

38. Tian, C.; Sun, R.; Liu, K.; Fu, L.; Liu, X.; Zhou, W.; Yang, Y.; Yang, J., Multiplexed Thiol Reactivity Profiling for Target Discovery of Electrophilic Natural Products. *Cell Chem Biol* **2017,** *24* (11), 1416-1427 e5.

39. Szychowski, J.; Mahdavi, A.; Hodas, J. J.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A., Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J Am Chem Soc* **2010,** *132* (51), 18351-60.

40. Griffin, M. E.; Jensen, E. H.; Mason, D. E.; Jenkins, C. L.; Stone, S. E.; Peters, E. C.; Hsieh-Wilson, L. C., Comprehensive mapping of O-GlcNAc modification sites using a chemically cleavable tag. *Mol Biosyst* **2016,** *12* (6), 1756-9.

41. Speers, A. E.; Cravatt, B. F., A Tandem Orthogonal Proteolysis Strategy for High-Content Chemical Proteomics. *Journal of the American Chemical Society* **2005,** *127* (28), 10018-10019.

42. Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010,** *468* (7325), 790-795.

43. Desai, H. S.; Yan, T.; Yu, F.; Sun, A. W.; Villanueva, M.; Nesvizhskii, A. I.; Backus, K. M., SP3-Enabled Rapid and High Coverage Chemoproteomic Identification of Cell-State-Dependent Redox-Sensitive Cysteines. *Mol Cell Proteomics* **2022,** *21* (4), 100218.

44. Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; Cravatt, B. F.; Gygi, S. P., Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat Biotechnol* **2021,** *39* (5), 630-641.

45. Taverna, D.; Gaspari, M., A critical comparison of three MS-based approaches for quantitative proteomics analysis. *Journal of Mass Spectrometry* **2021,** *56* (1), e4669.

46. Moulder, R.; Goo, Y. A.; Goodlett, D. R., Label-Free Quantitation for Clinical Proteomics. *Methods Mol Biol* **2016,** *1410*, 65-76.

47. Schubert, O. T.; Röst, H. L.; Collins, B. C.; Rosenberger, G.; Aebersold, R., Quantitative proteomics: challenges and opportunities in basic and applied research. *Nature Protocols* **2017,** *12* (7), 1289-1294.

48. Wang, R.; Zhou, P.; Pan, Y.; Zheng, L.; Dong, X.; Shen, R.; Lan, P., Label-Free Quantitative Proteomics in Plant. *Methods Mol Biol* **2023,** *2665*, 75-83.

49. Wang, J.; Zhang, C. J.; Chia, W. N.; Loh, C. C.; Li, Z.; Lee, Y. M.; He, Y.; Yuan, L. X.; Lim, T. K.; Liu, M.; Liew, C. X.; Lee, Y. Q.; Zhang, J.; Lu, N.; Lim, C. T.; Hua, Z. C.; Liu, B.; Shen, H. M.; Tan, K. S.; Lin, Q., Haem-activated promiscuous targeting of artemisinin in Plasmodium falciparum. *Nat Commun* **2015,** *6*, 10111.

50. Zheng, B.; Zhu, S.; Wu, X., Clickable Analogue of Cerulenin as Chemical Probe to Explore Protein Palmitoylation. *ACS Chemical Biology* **2015,** *10* (1), 115-121.

51. Anand, S.; Samuel, M.; Ang, C. S.; Keerthikumar, S.; Mathivanan, S., Label-Based and Label-Free Strategies for Protein Quantitation. *Methods Mol Biol* **2017,** *1549*, 31-43.

52. Tian, X.; Permentier, H. P.; Bischoff, R., Chemical isotope labeling for quantitative proteomics. *Mass Spectrom Rev* **2023,** *42* (2), 546-576.

53. Deng, J.; Erdjument-Bromage, H.; Neubert, T. A., Quantitative Comparison of Proteomes Using SILAC. *Current Protocols in Protein Science* **2019,** *95* (1), e74.

54. Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F., Quantitative proteomics using SILAC: Principles, applications, and developments. *PROTEOMICS* **2015,** *15* (18), 3175-3192.

55. Xu, D.; Zhu, X.; Ren, J.; Huang, S.; Xiao, Z.; Jiang, H.; Tan, Y., Quantitative proteomic analysis of cervical cancer based on TMT-labeled quantitative proteomics. *Journal of Proteomics* **2022,** *252*, 104453.

56. Chen, X.; Sun, Y.; Zhang, T.; Shu, L.; Roepstorff, P.; Yang, F., Quantitative Proteomics Using Isobaric Labeling: A Practical Guide. *Genomics, Proteomics & Bioinformatics* **2021,** *19* (5), 689-706.

57. Kreuzer, J.; Bach, N. C.; Forler, D.; Sieber, S. A., Target discovery of acivicin in cancer cells elucidates its mechanism of growth inhibition. *Chemical Science* **2015,** *6* (1), 237-245.

58. Gleissner, C. M. L.; Pyka, C. L.; Heydenreuter, W.; Gronauer, T. F.; Atzberger, C.; Korotkov, V. S.; Cheng, W.; Hacker, S. M.; Vollmar, A. M.; Braig, S.; Sieber, S. A., Neocarzilin A Is a Potent

Inhibitor of Cancer Cell Motility Targeting VAT-1 Controlled Pathways. *ACS Central Science* **2019,** *5* (7), 1170-1178.

59. Sharma, R.; Fedorenko, I.; Spence, P. T.; Sondak, V. K.; Smalley, K. S.; Koomen, J. M., Activity-Based Protein Profiling Shows Heterogeneous Signaling Adaptations to BRAF Inhibition. *J Proteome Res* **2016,** *15* (12), 4476-4489.

60. Martinez Molina, D.; Jafari, R.; Ignatushchenko, M.; Seki, T.; Larsson, E. A.; Dan, C.; Sreekumar, L.; Cao, Y.; Nordlund, P., Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **2013,** *341* (6141), 84-7.

61. Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; Lundbäck, T.; Nordlund, P.; Martinez Molina, D., The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat Protoc* **2014,** *9* (9), 2100-22.

62. Caballero, I. M.; Lundgren, S., A Shift in Thinking: Cellular Thermal Shift Assay-Enabled Drug Discovery. *ACS Med Chem Lett* **2023,** *14* (4), 369-375.

63. Savitski, M. M.; Reinhard, F. B.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Martinez Molina, D.; Jafari, R.;

Dovega, R. B.; Klaeger, S.; Kuster, B.; Nordlund, P.; Bantscheff, M.; Drewes, G., Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **2014,** *346* (6205), 1255784.

64. Mateus, A.; Määttä, T. A.; Savitski, M. M., Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome Sci* **2016,** *15*, 13.

65. Franken, H.; Mathieson, T.; Childs, D.; Sweetman, G. M.; Werner, T.; Tögel, I.; Doce, C.; Gade, S.; Bantscheff, M.; Drewes, G.; Reinhard, F. B.; Huber, W.; Savitski, M. M., Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* **2015,** *10* (10), 1567-93.

66. Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A.; Wohlschlegel, J. A.; Vondriska, T. M.; Pelletier, J.; Herschman, H. R.; Clardy, J.; Clarke, C. F.; Huang, J., Target identification using drug affinity responsive target stability (DARTS). *Proc Natl Acad Sci U S A* **2009,** *106* (51), 21984-9.

67. Derry, M. M.; Somasagara, R. R.; Raina, K.; Kumar, S.; Gomez, J.; Patel, M.; Agarwal, R.; Agarwal, C., Target identification of grape seed extract in colorectal cancer using drug affinity responsive target stability (DARTS) technique: role of endoplasmic reticulum stress response proteins. *Curr Cancer Drug Targets* **2014,** *14* (4), 323-36.

68. Park, C.; Zhou, S.; Gilmore, J.; Marqusee, S., Energetics-based protein profiling on a proteomic scale: identification of proteins resistant to proteolysis. *J Mol Biol* **2007,** *368* (5), 1426-37.

69. West, G. M.; Tang, L.; Fitzgerald, M. C., Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass spectrometry-based strategy. *Anal Chem* **2008,** *80* (11), 4175-85.

70. West, G. M.; Tucker, C. L.; Xu, T.; Park, S. K.; Han, X.; Yates, J. R., 3rd; Fitzgerald, M. C., Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc Natl Acad Sci U S A* **2010,** *107* (20), 9078-82.

71. Walker, E. J.; Bettinger, J. Q.; Welle, K. A.; Hryhorenko, J. R.; Ghaemmaghami, S., Global analysis of methionine oxidation provides

a census of folding stabilities for the human proteome. *Proc Natl Acad Sci U S A* **2019,** *116* (13), 6081-6090.

72. Xu, Y.; Falk, I. N.; Hallen, M. A.; Fitzgerald, M. C., Mass spectrometry- and lysine amidination-based protocol for thermodynamic analysis of protein folding and ligand binding interactions. *Anal Chem* **2011,** *83* (9), 3555-62.

73. Keskin, O.; Tuncbag, N.; Gursoy, A., Predicting Protein–Protein Interactions from the Molecular to the Proteome Level. *Chemical Reviews* **2016,** *116* (8), 4884-4909.

74. Castoreno, A. B.; Eggert, U. S., Small Molecule Probes of Cellular Pathways and Networks. *ACS Chemical Biology* **2011,** *6* (1), 86-94.

75. Resnick, E.; Bradley, A.; Gan, J.; Douangamath, A.; Krojer, T.; Sethi, R.; Geurink, P. P.; Aimon, A.; Amitai, G.; Bellini, D.; Bennett, J.; Fairhead, M.; Fedorov, O.; Gabizon, R.; Gan, J.; Guo, J.; Plotnikov, A.; Reznik, N.; Ruda, G. F.; Díaz-Sáez, L.; Straub, V. M.; Szommer, T.; Velupillai, S.; Zaidman, D.; Zhang, Y.; Coker, A. R.; Dowson, C. G.; Barr, H. M.; Wang, C.; Huber, K. V. M.; Brennan, P. E.; Ovaa, H.; von Delft, F.; London, N., Rapid Covalent-Probe

Discovery by Electrophile-Fragment Screening. *Journal of the American Chemical Society* **2019,** *141* (22), 8951-8968.

76. Mertens, R. T.; Gukathasan, S.; Arojojoye, A. S.; Olelewe, C.; Awuah, S. G., Next Generation Gold Drugs and Probes: Chemistry and Biomedical Applications. *Chemical Reviews* **2023,** *123* (10), 6612-6667.

77. Rask-Andersen, M.; Almén, M.; Schi?Th, H. B., Trends in the exploitation of novel drug targets. *Nature Reviews Drug Discovery* **2011,** *10* (8), 579.

78. Suciu, R. M.; Cognetta, A. B., III; Potter, Z. E.; Cravatt, B. F., Selective Irreversible Inhibitors of the Wnt-Deacylating Enzyme NOTUM Developed by Activity-Based Protein Profiling. *ACS Medicinal Chemistry Letters* **2018,** *9* (6), 563-568.

79. Weichert, D.; Gmeiner, P., Covalent molecular probes for class A G protein-coupled receptors: advances and applications. *Acs Chemical Biology* **2015,** *10* (6), 1376-1386.

80. Kozhinov, A. N.; Johnson, A.; Nagornov, K. O.; Stadlmeier, M.; Martin, W. L.; Dayon, L.; Corthésy, J.; Wühr, M.; Tsybin, Y. O., Super-Resolution Mass Spectrometry Enables Rapid, Accurate, and

Highly Multiplexed Proteomics at the MS2 Level. *Analytical Chemistry* **2023,** *95* (7), 3712-3719.

81. Jiao, F.; Salituro, L. J.; Yu, C.; Gutierrez, C. B.; Rychnovsky, S. D.; Huang, L., Exploring an Alternative Cysteine-Reactive Chemistry to Enable Proteome-Wide PPI Analysis by Cross-Linking Mass Spectrometry. *Analytical Chemistry* **2023,** *95* (4), 2532-2539.

82. Yang, F.; Gao, J.; Che, J.; Jia, G.; Wang, C., A Dimethyl-Labeling-Based Strategy for Site-Specifically Quantitative Chemical Proteomics. *Anal Chem* **2018,** *90* (15), 9576-9582.

83. van Bergen, W.; Hevler, J. F.; Wu, W.; Baggelaar, M. P.; Heck, A. J. R., Site-Specific Activity-Based Protein Profiling Using Phosphonate Handles. *Mol Cell Proteomics* **2023,** *22* (1), 100455.

84. Gao, P.; Liu, Y.-Q.; Xiao, W.; Xia, F.; Chen, J.-Y.; Gu, L.-W.; Yang, F.; Zheng, L.-H.; Zhang, J.-Z.; Zhang, Q.; Li, Z.-J.; Meng, Y.-Q.; Zhu, Y.-P.; Tang, H.; Shi, Q.-L.; Guo, Q.-Y.; Zhang, Y.; Xu, C.-C.; Dai, L.-Y.; Wang, J.-G., Identification of antimalarial targets of chloroquine by a combined deconvolution strategy of ABPP and MS-CETSA. *Military Medical Research* **2022,** *9* (1), 30.

85. Wang, M.-R.; Huang, L.-F.; Guo, C.; Yang, J.; Dong, S.; Tang, J.-J.; Gao, J.-M., Identification of NLRP3 as a covalent target of 1,6-O,O-diacetylbritannilactone against neuroinflammation by quantitative thiol reactivity profiling (QTRP). *Bioorganic Chemistry* **2022,** *119*, 105536.

86. Corson, T. W.; Crews, C. M., Molecular understanding and modern application of traditional medicines: triumphs and trials. *Cell* **2007,** *130* (5), 769-74.

87. Zhu, B.; Wei, Y., Antitumor activity of celastrol by inhibition of proliferation, invasion, and migration in cholangiocarcinoma via PTEN/PI3K/Akt pathway. *Cancer Med* **2020,** *9* (2), 783-796.

88. Zeng, X.; Zhu, X.; Tian, Q.; Tan, X.; Sun, N.; Yan, M.; Zhao, J.; Wu, X.; Li, R.; Zhang, Z.; Zeng, H., Celastrol-conjugated chitosan oligosaccharide for the treatment of pancreatic cancer. *Drug Deliv* **2022,** *29* (1), 89-98.

89. Liu, J.; Lee, J.; Salazar Hernandez, M. A.; Mazitschek, R.; Ozcan, U., Treatment of obesity with celastrol. *Cell* **2015,** *161* (5), 999-1011.

90. Xu, S.; Feng, Y.; He, W.; Xu, W.; Xu, W.; Yang, H.; Li, X., Celastrol in metabolic diseases: Progress and application prospects. *Pharmacol Res* **2021,** *167*, 105572.

91. Wang, Y.; Li, C.; Gu, J.; Chen, C.; Duanmu, J.; Miao, J.; Yao, W.; Tao, J.; Tu, M.; Xiong, B.; Zhao, L.; Liu, Z., Celastrol exerts anti-inflammatory effect in liver fibrosis via activation of AMPK-SIRT3 signalling. *J Cell Mol Med* **2020,** *24* (1), 941-953.

92. Bian, M.; Du, X.; Cui, J.; Wang, P.; Wang, W.; Zhu, W.; Zhang, T.; Chen, Y., Celastrol protects mouse retinas from bright light-induced degeneration through inhibition of oxidative stress and inflammation. *J Neuroinflammation* **2016,** *13*, 50.

93. Liu, D. D.; Luo, P.; Gu, L.; Zhang, Q.; Gao, P.; Zhu, Y.; Chen, X.; Guo, Q.; Zhang, J.; Ma, N.; Wang, J., Celastrol exerts a neuroprotective effect by directly binding to HMGB1 protein in cerebral ischemia-reperfusion. *J Neuroinflammation* **2021,** *18* (1), 174.

94. Chen, S.; Gu, C.; Xu, C.; Zhang, J.; Xu, Y.; Ren, Q.; Guo, M.; Huang, S.; Chen, L., Celastrol prevents cadmium-induced neuronal cell death via targeting JNK and PTEN-Akt/mTOR network. *J Neurochem* **2014,** *128* (2), 256-266.

95. Lee, J.-H.; Won, Y.-S.; Park, K.-H.; Lee, M.-K.; Tachibana, H.; Yamada, K.; Seo, K.-I., Celastrol inhibits growth and induces apoptotic cell death in melanoma cells via the activation ROS-dependent mitochondrial pathway and the suppression of PI3K/AKT signaling. *Apoptosis* **2012,** *17*, 1275-1286.

96. Pang, X.; Yi, Z.; Zhang, J.; Lu, B.; Sung, B.; Qu, W.; Aggarwal, B. B.; Liu, M., Celastrol Suppresses Angiogenesis-Mediated Tumor Growth through Inhibition of AKT/Mammalian Target of Rapamycin PathwayCelastrol Inhibits Tumor Angiogenesis. *Cancer research* **2010,** *70* (5), 1951-1959.

97. Lu, Z.; Jin, Y.; Qiu, L.; Lai, Y.; Pan, J., Celastrol, a novel HSP90 inhibitor, depletes Bcr–Abl and induces apoptosis in imatinib-resistant chronic myelogenous leukemia cells harboring T315I mutation. *Cancer letters* **2010,** *290* (2), 182-191.

98. Zhang, X.; Yang, J.; Chen, M.; Li, L.; Huan, F.; Li, A.; Liu, Y.; Xia, Y.; Duan, J.-a.; Ma, S., Metabolomics profiles delineate uridine deficiency contributes to mitochondria-mediated apoptosis induced by celastrol in human acute promyelocytic leukemia cells. *Oncotarget* **2016,** *7* (29), 46557.

99. Lee, H.-W.; Jang, K. S. B.; Choi, H. J.; Jo, A.; Cheong, J.-H.; Chun, K.-H., Celastrol inhibits gastric cancer growth by induction of apoptosis and autophagy. *BMB reports* **2014,** *47* (12), 697.

100. Yang, Y.; Cheng, S.; Liang, G.; Honggang, L.; Wu, H., Celastrol inhibits cancer metastasis by suppressing M2-like polarization of macrophages. *Biochemical and biophysical research communications* **2018,** *503* (2), 414-419.

101. Chen, X.; Zhao, Y.; Luo, W.; Chen, S.; Lin, F.; Zhang, X.; Fan, S.; Shen, X.; Wang, Y.; Liang, G., Celastrol induces ROS-mediated apoptosis via directly targeting peroxiredoxin-2 in gastric cancer cells. *Theranostics* **2020,** *10* (22), 10290-10308.

102. Ye, S.; Luo, W.; Khan, Z. A.; Wu, G.; Liang, G., Celastrol Attenuates Angiotensin II-Induced Cardiac Remodeling by Targeting STAT3. *Circulation Research* **2020,** *126* (8).

103. Celastrol Promotes Weight Loss in Diet-Induced Obesity by Inhibiting the Protein Tyrosine Phosphatases PTP1B and TCPTP in the Hypothalamus. *Journal of Medicinal Chemistry* **2018,** *61* (24), 11144-11157.

104. Kannaiyan, R.; Shanmugam, M. K.; Sethi, G., Molecular targets of celastrol derived from Thunder of God Vine: Potential role in the treatment of inflammatory disorders and cancer. *Cancer Letters* **2011,** *303* (1), 9-20.

105. Lim, H. Y.; Ong, P. S.; Wang, L.; Goel, A.; Ding, L.; Li-Ann Wong, A.; Ho, P. C.-l.; Sethi, G.; Xiang, X.; Goh, B. C., Celastrol in cancer therapy: Recent developments, challenges and prospects. *Cancer Letters* **2021,** *521*, 252-267.

106. Zhang, T.; Zhao, Q.; Xiao, X.; Yang, R.; Hu, D.; Zhu, X.; Gonzalez, F. J.; Li, F., Modulation of Lipid Metabolism by Celastrol. *J Proteome Res* **2019,** *18* (3), 1133-1144.

107. Faust, K.; Gehrke, S.; Yang, Y.; Yang, L.; Beal, M.; Lu, B., Neuroprotective effects of compounds with antioxidant and anti-inflammatory properties in a Drosophila model of Parkinson's disease. *Bmc Neuroscience* **2009,** *10* (1), 109-109.

108. Corson, T. W.; Crews, C. M., Molecular understanding and modern application of traditional medicines: triumphs and trials. *Cell* **2007,** *130* (5), 769-774.

109. Zhou, B.; Yuan, Y.; Shi, L.; Hu, S.; Wang, D.; Yang, Y.; Pan, Y.; Kong, D.; Shikov, A. N.; Duez, P., Creation of an anti-inflammatory, leptin-dependent anti-obesity celastrol mimic with better druggability. *Frontiers in Pharmacology* **2021,** *12*, 705252.

110. Lee, J.-H.; Koo, T. H.; Yoon, H.; Jung, H. S.; Jin, H. Z.; Lee, K.; Hong, Y.-S.; Lee, J. J., Inhibition of NF-κB activation through targeting IκB kinase by celastrol, a quinone methide triterpenoid. *Biochemical pharmacology* **2006,** *72* (10), 1311-1321.

111. Su, Y.; DePasquale, M.; Liao, G.; Buchler, I.; Zhang, G.; Byers, S.; Carr, G. V.; Barrow, J.; Wei, H., Membrane bound catechol-O-methytransferase is the dominant isoform for dopamine metabolism in PC12 cells and rat brain. *European journal of pharmacology* **2021,** *896*, 173909.

112. Myöhänen, T. T.; Männistö, P. T., Distribution and functions of catechol-O-methyltransferase proteins: do recent findings change the picture? *Int Rev Neurobiol* **2010,** *95*, 29-47.

113. Zhou, Y.; Li, W.; Wang, M.; Zhang, X.; Zhang, H.; Tong, X.; Xiao, Y., Competitive profiling of celastrol targets in human cervical

cancer HeLa cells via quantitative chemical proteomics. *Molecular BioSystems* **2017,** *13* (1), 83-91.

114. Zhang, D.; Chen, Z.; Hu, C.; Yan, S.; Li, Z.; Lian, B.; Xu, Y.; Ding, R.; Zeng, Z.; Zhang, X.-K.; Su, Y., Celastrol binds to its target protein via specific noncovalent interactions and reversible covalent bonds. *Chemical Communications* **2018,** *54* (91), 12871-12874.

115. Peng, B.; Xu, L.; Cao, F.; Wei, T.; Yang, C.; Uzan, G.; Zhang, D., HSP90 inhibitor, celastrol, arrests human monocytic leukemia cell U937 at G0/G1 in thiol-containing agents reversible way. *Molecular Cancer* **2010,** *9* (1), 79.

116. Little, R. D.; Masjedizadeh, M. R.; Wallquist, O.; Mcloughlin, J. I., The Intramolecular Michael Reaction. In *Organic Reactions*, pp 315-552.

117. Zhang, D.; Chen, Z.; Hu, C.; Yan, S.; Li, Z.; Lian, B.; Xu, Y.; Ding, R.; Zeng, Z.; Zhang, X.-K.; Su, Y., Celastrol binds to its target protein via specific noncovalent interactions and reversible covalent bonds. *Chemical Communications* **2018,** *54* (91), 12871-12874.

118. Mather, B. D.; Viswanathan, K.; Miller, K. M.; Long, T. E., Michael addition reactions in macromolecular design for emerging technologies. *Progress in Polymer Science* **2006,** *31* (5), 487-531.

119. Serpa, J., Cysteine as a Carbon Source, a Hot Spot in Cancer Cells Survival. *Front Oncol* **2020,** *10*, 947.

120. Al-Khawaldeh, I.; Yasiri, M. A.; Aldred, G.; Basmadjian, C.; Bordoni, C.; Harnor, S.; Heptinstall, A.; Hobson, S.; Jennings, C.; Khalifa, S., An Alkynylpyrimidine-Based Covalent Inhibitor That Targets a Unique Cysteine in NF-κB-Inducing Kinase. *Journal of medicinal chemistry* **2021,** *64* (14), 10001-10018.

121. Kang, L.; Weng, N.; Jian, W., LC–MS bioanalysis of intact proteins and peptides. *Biomedical Chromatography* **2020,** *34* (1), e4633.

122. Pacholarz, K. J.; Garlish, R. A.; Taylor, R. J.; Barran, P. E., Mass spectrometry based tools to investigate protein–ligand interactions for drug discovery. *Chemical Society Reviews* **2012,** *41* (11), 4335.

123. Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R., Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **1986,** *83* (17), 6233-6237.

124. Harrison, P. J.; Tunbridge, E. M., Catechol-O-Methyltransferase (COMT): A Gene Contributing to Sex Differences in Brain Function, and to Sexual Dimorphism in the Predisposition to Psychiatric Disorders. *Neuropsychopharmacology* **2008,** *33* (13), 3037-3045.

125. Cumming, P.; Brown, E.; Damsma, G.; Fibiger, H., Formation and Clearance of Interstitial Metabolites of Dopamine and Serotonin in the Rat Striatum: An In Vivo Microdialysis Study. *Journal of Neurochemistry* **1992,** *59* (5), 1905-1914.

126. Barnett, J. H.; Scoriels, L.; Munafò, M. R., Meta-analysis of the cognitive effects of the catechol-O-methyltransferase gene Val158/108Met polymorphism. *Biological psychiatry* **2008,** *64* (2), 137-144.

127. Parkin, G. M.; Udawela, M.; Gibbons, A.; Scarr, E.; Dean, B., Catechol-O-methyltransferase (COMT) genotypes are associated with varying soluble, but not membrane-bound COMT protein in the human prefrontal cortex. *Journal of Human Genetics* **2018,** *63* (12), 1251-1258.

128. Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J., Kinetics of human soluble and membrane-

bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* **1995,** *34* (13), 4202-4210.

129. Kiss, L. E.; Soares-da-Silva, P., Medicinal Chemistry of Catechol O-Methyltransferase (COMT) Inhibitors and Their Therapeutic Utility. *Journal of Medicinal Chemistry* **2014,** *57* (21), 8692-8717.

130. Sweet, R. A.; Devlin, B.; Pollock, B. G.; Sukonick, D. L.; Kastango, K. B.; Bacanu, S. A.; Chowdari, K. V.; Dekosky, S. T.; Ferrell, R. E., Catechol-O-methyltransferase haplotypes are associated with psychosis in Alzheimer disease. *Molecular Psychiatry* **2005,** *10* (11), 1026.

131. Egan, M. F.; Goldberg, T. E.; Kolachana, B. S.; Callicott, J. H.; Mazzanti, C. M.; Straub, R. E.; Goldman, D.; Weinberger, D. R., Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (12), 6917-6922.

132. Paloyelis, Y.; Asherson, P.; Mehta, M. A.; Faraone, S. V.; Kuntsi, J., DAT1 and COMT Effects on Delay Discounting and Trait Impulsivity in Male Adolescents with Attention

Deficit|[sol]|Hyperactivity Disorder and Healthy Controls. *Neuropsychopharmacology* **2010,** *35* (12), 2414-2426.

133. Haasio, K., Toxicology and Safety of Comt Inhibitors. In *International Review of Neurobiology*, Nissinen, E., Ed. Academic Press: 2010; Vol. 95, pp 163-189.

134. Xia, Y.; Pang, H.; Dou, T.; Wang, P.; Ge, G., Interspecies comparison in the COMT-mediated methylation of 3-BTD. *Rsc Advances* **2018,** *8* (29), 16278-16284.

135. Silva, T. B.; Borges, F.; Serro, M. P.; Soares-Da-Silva, P., Liver says no: the ongoing search for safe catechol O-methyltransferase inhibitors to replace tolcapone. *Drug Discovery Today* **2020,** *25* (10).

136. Bartl, J.; Palazzesi, F.; Parrinello, M.; Hommers, L.; Riederer, P.; Walitza, S.; Grünblatt, E., The impact of methylphenidate and its enantiomers on dopamine synthesis and metabolism in vitro. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **2017,** *79*, 281-288.

137. Wang, F.-Y.; Wang, P.; Zhao, D.-F.; Gonzalez, F. J.; Fan, Y.-F.; Xia, Y.-L.; Ge, G.-B.; Yang, L., Analytical methodologies for

sensing catechol-O-methyltransferase activity and their applications. *Journal of Pharmaceutical Analysis* **2021,** *11* (1), 15-27.

138. Czarnota, S.; Johannissen, L. O.; Baxter, N. J.; Rummel, F.; Wilson, A. L.; Cliff, M. J.; Levy, C. W.; Scrutton, N. S.; Waltho, J. P.; Hay, S., Equatorial Active Site Compaction and Electrostatic Reorganization in Catechol-O-methyltransferase. *ACS Catal* **2019,** *9* (5), 4394-4401.

139. Cockerham, R.; Liu, S.; Cachope, R.; Kiyokage, E.; Cheer, J. F.; Shipley, M. T.; Puche, A. C., Subsecond Regulation of Synaptically Released Dopamine by COMT in the Olfactory Bulb. *J Neurosci* **2016,** *36* (29), 7779-85.

140. Zhang, G.; Buchler, I. P.; DePasquale, M.; Wormald, M.; Liao, G.; Wei, H.; Barrow, J. C.; Carr, G. V., Development of a PC12 Cell Based Assay for Screening Catechol-O-methyltransferase Inhibitors. *ACS Chem Neurosci* **2019,** *10* (10), 4221-4226.

141. Ashare, R. L.; Wileyto, E. P.; Ruparel, K.; Goelz, P. M.; Hopson, R. D.; Valdez, J. N.; Gur, R. C.; Loughead, J.; Lerman, C., Effects of tolcapone on working memory and brain activity in abstinent

smokers: A proof-of-concept study. *Drug and Alcohol Dependence* **2013,** *133* (3), 852-856.

142. Pace, N. J.; Weerapana, E., Diverse Functional Roles of Reactive Cysteines. *ACS Chemical Biology* **2013,** *8* (2), 283-296.

143. Musa, A.; Ghoraie, L. S.; Zhang, S. D.; Glazko, G.; Yli-Harja, O.; Dehmer, M.; Haibe-Kains, B.; Emmert-Streib, F., A review of connectivity map and computational approaches in pharmacogenomics. *Brief Bioinform* **2017,** *18* (5), 903.

144. Jiang, H.; Hu, C.; Chen, M., The Advantages of Connectivity Map Applied in Traditional Chinese Medicine. *Front Pharmacol* **2021,** *12*, 474267.

145. Subramanian, A.; Narayan, R.; Corsello, S. M.; Peck, D. D.; Natoli, T. E.; Lu, X.; Gould, J.; Davis, J. F.; Tubelli, A. A.; Asiedu, J. K.; Lahr, D. L.; Hirschman, J. E.; Liu, Z.; Donahue, M.; Julian, B.; Khan, M.; Wadden, D.; Smith, I. C.; Lam, D.; Liberzon, A.; Toder, C.; Bagul, M.; Orzechowski, M.; Enache, O. M.; Piccioni, F.; Johnson, S. A.; Lyons, N. J.; Berger, A. H.; Shamji, A. F.; Brooks, A. N.; Vrcic, A.; Flynn, C.; Rosains, J.; Takeda, D. Y.; Hu, R.; Davison, D.; Lamb, J.; Ardlie, K.; Hogstrom, L.; Greenside, P.; Gray,

N. S.; Clemons, P. A.; Silver, S.; Wu, X.; Zhao, W. N.; Read-Button, W.; Wu, X.; Haggarty, S. J.; Ronco, L. V.; Boehm, J. S.; Schreiber, S. L.; Doench, J. G.; Bittker, J. A.; Root, D. E.; Wong, B.; Golub, T. R., A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **2017,** *171* (6), 1437-1452 e17.

146. Ye, M.; Jiang, X.; Feng, S.; Tian, R.; Zou, H., Advances in chromatographic techniques and methods in shotgun proteome analysis. *TrAC Trends in Analytical Chemistry* **2007,** *26* (1), 80-84.

147. Meyer, J. G.; Niemi, N. M.; Pagliarini, D. J.; Coon, J. J., Quantitative shotgun proteome analysis by direct infusion. *Nature Methods* **2020,** *17* (12), 1222-1228.

148. Yocum, A. K.; Chinnaiyan, A. M., Current affairs in quantitative targeted proteomics: multiple reaction monitoring–mass spectrometry. *Briefings in Functional Genomics* **2009,** *8* (2), 145-157.

149. Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J., Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Molecular & Cellular Proteomics* **2012,** *11* (11), 1475-1488.

150. Rauniyar, N., Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *International Journal of Molecular Sciences* **2015,** *16* (12), 28566-28581.

151. Ronsein, G. E.; Pamir, N.; von Haller, P. D.; Kim, D. S.; Oda, M. N.; Jarvik, G. P.; Vaisar, T.; Heinecke, J. W., Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *Journal of Proteomics* **2015,** *113*, 388-399.

152. Park, J.; Oh, H. J.; Han, D.; Wang, J. I.; Park, I. A.; Ryu, H. S.; Kim, Y., Parallel Reaction Monitoring-Mass Spectrometry (PRM-MS)-Based Targeted Proteomic Surrogates for Intrinsic Subtypes in Breast Cancer: Comparative Analysis with Immunohistochemical Phenotypes. *Journal of Proteome Research* **2020,** *19* (7), 2643-2653.

153. Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J., Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics*. *Molecular & Cellular Proteomics* **2012,** *11* (11), 1475-1488.

154. Chadli, A.; Felts, S. J.; Wang, Q.; Sullivan, W. P.; Botuyan, M. V.; Fauq, A.; Ramirez-Alvarado, M.; Mer, G., Celastrol inhibits Hsp90 chaperoning of steroid receptors by inducing fibrillization of the Co-chaperone p23. *J Biol Chem* **2010,** *285* (6), 4224-4231.

155. Zhao, Q.; Ding, Y.; Deng, Z.; Lee, O. Y.; Gao, P.; Chen, P.; Rose, R. J.; Zhao, H.; Zhang, Z.; Tao, X. P.; Heck, A. J. R.; Kao, R.; Yang, D., Natural products triptolide, celastrol, and withaferin A inhibit the chaperone activity of peroxiredoxin I. *Chem Sci* **2015,** *6* (7), 4124-4130.

156. Bargagna-Mohan, P.; Hamza, A.; Kim, Y. E.; Khuan Abby Ho, Y.; Mor-Vaknin, N.; Wendschlag, N.; Liu, J.; Evans, R. M.; Markovitz, D. M.; Zhan, C. G.; Kim, K. B.; Mohan, R., The tumor inhibitor and antiangiogenic agent withaferin A targets the intermediate filament protein vimentin. *Chem Biol* **2007,** *14* (6), 623-34.

157. Saei, A. A.; Gullberg, H.; Sabatier, P.; Beusch, C. M.; Johansson, K.; Lundgren, B.; Arvidsson, P. I.; Arnér, E. S. J.; Zubarev, R. A., Comprehensive chemical proteomics for target deconvolution of the redox active drug auranofin. *Redox Biol* **2020,** *32*, 101491.

158. Zhang, T.; Li, Y.; Yu, Y.; Zou, P.; Jiang, Y.; Sun, D., Characterization of celastrol to inhibit hsp90 and cdc37 interaction. *Journal of Biological Chemistry* **2009,** *284* (51), 35381-35389.

159. Yu, Y.; Hamza, A.; Zhang, T.; Gu, M.; Zou, P.; Newman, B.; Li, Y.; Gunatilaka, A. L.; Zhan, C.-G.; Sun, D., Withaferin A targets heat shock protein 90 in pancreatic cancer cells. *Biochemical pharmacology* **2010,** *79* (4), 542-551.

160. Neumann, C. A.; Krause, D. S.; Carman, C. V.; Das, S.; Dubey, D. P.; Abraham, J. L.; Bronson, R. T.; Fujiwara, Y.; Orkin, S. H.; Van Etten, R. A., Essential role for the peroxiredoxin Prdx1 in erythrocyte antioxidant defence and tumour suppression. *Nature* **2003,** *424* (6948), 561-565.

161. Mishra, S. J.; Liu, W.; Beebe, K.; Banerjee, M.; Kent, C. N.; Munthali, V.; Koren, J., 3rd; Taylor, J. A., 3rd; Neckers, L. M.; Holzbeierlein, J.; Blagg, B. S. J., The Development of Hsp90β-Selective Inhibitors to Overcome Detriments Associated with pan-Hsp90 Inhibition. *J Med Chem* **2021,** *64* (3), 1545-1557.

162. Mishra, S. J.; Liu, W.; Beebe, K.; Banerjee, M.; Kent, C. N.; Munthali, V.; Koren, J., III; Taylor, J. A., III; Neckers, L. M.;

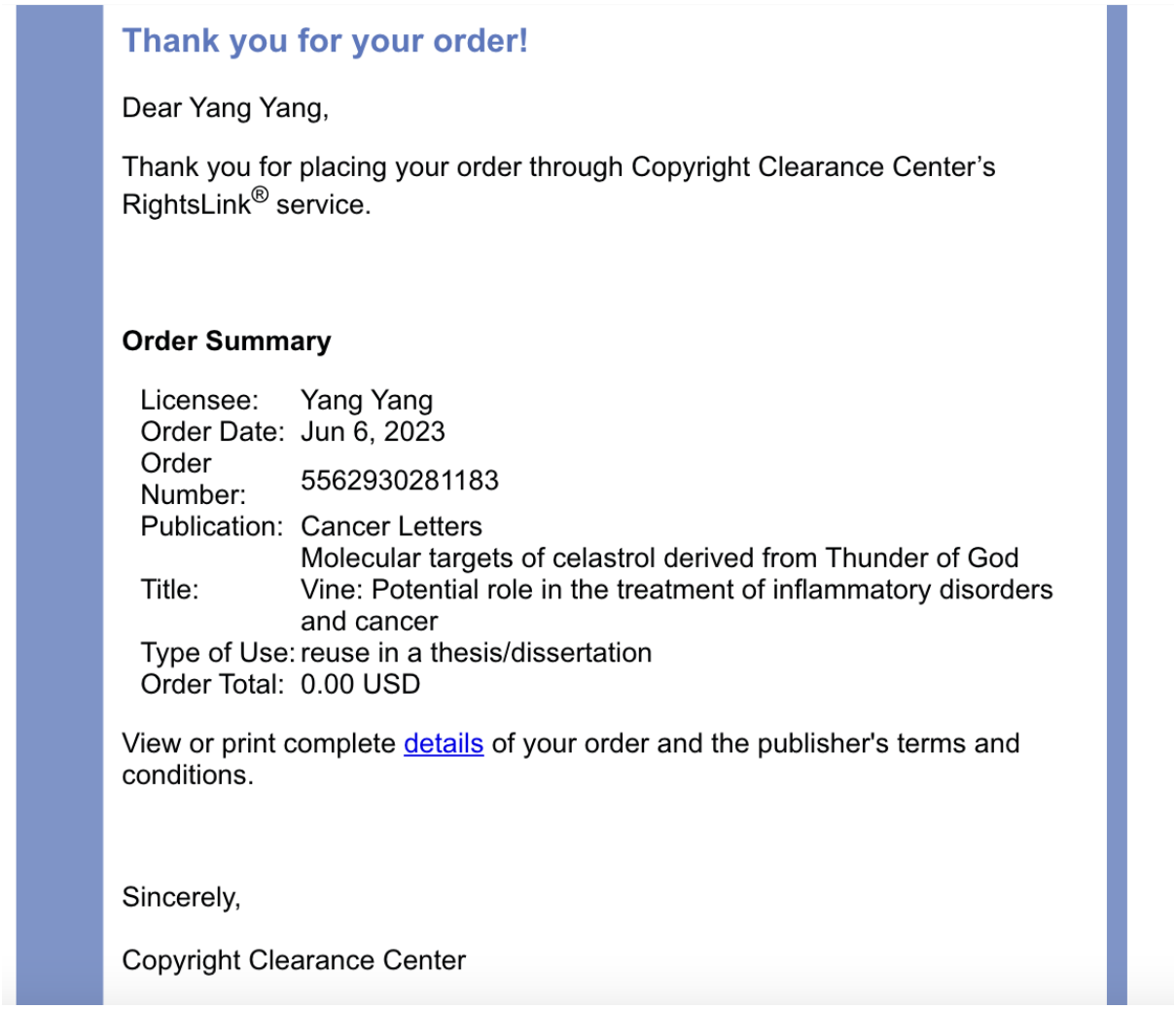
Holzbeierlein, J.; Blagg, B. S. J., The Development of Hsp90β-Selective Inhibitors to Overcome Detriments Associated with pan-Hsp90 Inhibition. *Journal of Medicinal Chemistry* **2021,** *64* (3), 1545-1557.

163. He, M. Y.; Xu, S. B.; Qu, Z. H.; Guo, Y. M.; Liu, X. C.; Cong, X. X.; Wang, J. F.; Low, B. C.; Li, L.; Wu, Q.; Lin, P.; Yan, S. G.; Bao, Z.; Zhou, Y. T.; Zheng, L. L., Hsp90β interacts with MDM2 to suppress p53-dependent senescence during skeletal muscle regeneration. *Aging Cell* **2019,** *18* (5), e13003.

164. Roy, R. K.; Patra, N., Prediction of COMT Inhibitors Using Machine Learning and Molecular Dynamics Methods. *The Journal of Physical Chemistry B* **2022,** *126* (19), 3477-3492.

Appendices

Copyright approval letter for use of **Figure 1-4**



Thank you for your order!

Dear Yang Yang,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Yang Yang
Order Date:	Jun 6, 2023
Order Number:	5562930281183
Publication:	Cancer Letters
Title:	Molecular targets of celastrol derived from Thunder of God Vine: Potential role in the treatment of inflammatory disorders and cancer
Type of Use:	reuse in a thesis/dissertation
Order Total:	0.00 USD

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Copyright approval letter for use of **Figure 1-5**

Thank you for your order!

Dear Yang Yang,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Yang Yang
Order Date:	Jun 6, 2023
Order Number:	5562940920514
Publication:	Cancer Letters
Title:	Celastrol in cancer therapy: Recent developments, challenges and prospects
Type of Use:	reuse in a thesis/dissertation
Order Total:	0.00 USD

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center