Multiplexed target profiling with integrated chemical genomics and chemical proteomics

Yang Yang[a,b], Yin-suen Tse[a,c], Qi Zhang[b], Kin-yau Wong[d], Chenxi Yang[a], Ying Yang[a], Shuqi Li[a], Kin-wa Lau[d], Trevor C. Charles[b,e], Thomas C. Lam[b], Qian Zhao[a]*

a. State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China

b. Centre for Eye and Vision Research, 17W Hong Kong Science Park, Hong Kong, SAR 999077, China

c. Laboratory for Synthetic Chemistry and Chemical Biology Limited, Hong Kong, SAR 999077, China

d. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China

e. Department of Biology, University of Waterloo, Waterloo, ON N2L 3G1, Canada

ABSTRACT Target identification is crucial for elucidating the mechanisms of bioactive molecules in drug discovery. However, traditional methods assess compounds individually, making it challenging to efficiently examine multiple compounds in parallel, especially for structurally diverse compounds. This study reports a novel strategy called chemical genomics-facilitated chemical proteomics (CGCP) for multiplexing target identification of bioactive small molecules. CGCP correlates compounds' perturbation of global transcription, or chemical genomic profiles, with their reactivity towards target proteins, enabling simultaneous identification of targets. We demonstrated the utility of CGCP by studying the targets of celastrol (Cel) and four other electrophilic compounds with varying levels of similarity to Cel based on their chemical genomic profiles. We identified multiple novel targets and binding sites shared by the compounds in a single experiment. CGCP enabled multiplexity and improved the efficiency of target identification for structurally distinct compounds, indicating its potential to accelerate drug discovery.

INTRODUCTION

Identification of the cellular targets of bioactive small molecules is an essential but challenging step in drug discovery[1]. Various chemical proteomics strategies have been developed for target identification in a large scale. Compound-centric chemical proteomics methods employ synthetic probes to mimic the bioactive molecules and enrich their binding proteins for subsequent mass spectrometry (MS) identification[2, 3]. These methods offer sensitivity by enriching specific targets of the given compounds, however, usually investigate compounds on an individual basis. High throughput methods such as thermal proteome profiling (TPP)[4], drug affinity responsive target stability (DARTS)[5], multiplexed thiol reactivity profiling (MTRP)[6], and target responsive accessibility profiling (TRAP)[7] that are based on the changes of target stability or cysteine/lysine druggability allow the target screening of multiple compounds in parallel. However, either the compound needs to be examined on individual basis, or only compounds with high structural similarity can be analyzed together. We still lack an effective approach that utilizes chemical probes for specific target enrichment, and meanwhile allow multiplexed target identification of compounds with diversified structures.

To achieve this goal, clustering compounds not based on structural similarity but rather their functionality is important. In chemical genomics, differential gene expression profiles under

compound perturbation are recorded as signatures. Signatures with high similarity represent connections between compounds with similar functions regardless of structural similarity[8,9]. The L1000 Connectivity Map (CMap) developed by Broad Institute is a comprehensive collection of genome-wide transcriptional expression data derived from cultured human cells treated with bioactive small molecules, along with advanced pattern-matching algorithms. The L1000 CMap database documents chemical genomic signatures of more than 5000 small molecules and provides their connectivity scores to one another[10]. By integrating these components, CMap facilitates the discovery of functional connections among drugs, genes, and diseases, highlighting shared gene-expression changes that occur in response to various treatments. Researchers can easily access the public transcriptome data of various compounds in different cell lines, bypassing the process of collecting gene expression data by themselves. Given that CMap data was used to successfully predict compounds with the same functions[9,11], we reason that signatures could have correlations with compounds' binding targets and potential to facilitate target identification process.

Bioactive electrophile celastrol (Cel), a quinone methide that exhibits anti-inflammatory[12], anti-oxi-dant[13], anti-cancer activities[14], etc[11,15], was recognized as one of the top-five promising drug candidates from natural products[16]. As Cel has attractive functions and multiple targets, we selected it as a representative compound to evaluate the new strategy CGCP, together with other electrophiles (Auranofin (Aur), Withaferin A (WA), Triptolide (TL), and Vinpocetine (Vin)) that are of high, medium, and low similarity to Cel according to chemical genomics. Aur is a gold compound, approved by Food and Drug Administration (FDA) of the USA for treatment of arthritis by decreasing inflammation[17]. WA is a natural steroidal lactone, known for its diverse pharmacological activities, including its role as a tumor growth inhibitor and its anti-inflammatory properties[18]. TL is a natural diterpenoid, exhibiting antirheumatic, anti-inflammatory, immunomodulatory, and antitumor pharmacological effects[19]. Vin is a synthetic derivative of the vinca alkaloid vincamine, has been used for treatment of cerebrovascular disorders such as stroke and dementia for over three decades[20]. In recent years, several protein targets of these electrophiles have been identified, for example, Cel was found to inhibit the enzymatic activity of catechol-O-methyltransferase (COMT)[21]. However, there are limited reports detailing the specific binding sites. Mapping the binding sites of bioactive compounds enables researchers to predict how effectively a compound will bind to its target, which is essential for assessing its efficacy and safety. By identifying these sites, scientists can design more effective drugs, optimize existing compounds, and minimize side effects. Site mapping also deepens insights into drug mechanisms of action and can facilitate the discovery of new therapeutic targets[22-23].

In this study, CGCP offered an innovative approach to compound clustering for multiplexed target profiling. We grouped compounds according to their similarities on the transcriptome level. The transcriptome data used in CGCP were from CMap database. After compound clustering, target identification was then conducted by cysteine profiling. CGCP provided novel targets and novel binding sites, which were validated with complementary approaches. We anticipated broader utility of CGCP for efficient target profiling of a variety of compounds in future.

RESULTS AND DISCUSSION

Optimization of a chemical proteomics method for comprehensive profiling of ligandable targets. First, we optimized the chemical proteomics method to comprehensively capture reactive cysteines, which served as potential binding sites of electrophilic compounds. Two complementary cysteine-reactive probes IAA[6,24,25]and EBX[26], along with three cleavable biotin-azide were tested to achieve comprehensive screening of ligandable cysteines (Figure S1). An optimal working concentration of IAA at 100 μM and EBX at 50 μM were determined to enable saturate labeling (Figure S2A). Compared to IAA, EBX showed higher proteome reactivity according to the gel-based fluorescence results (Figure S2A), which was consistent with the previous report[25]. Additionally, the UV cleavable

biotin-azide showed the highest labeling efficiency (Figure S2B) when the click reaction was performed with tryptic peptides (Figure S2C and S2D). The coverage of reactive cysteines across the proteome was improved by the simultaneous use of the two cysteine-reactive probes because the two probes can label complementary cysteinome (Figure S3A). Across five cell liens and mouse tissue sample, while 2998 (49%) cysteines were commonly found in both probes, 1825 cysteines (30%) were only enriched by EBX, and 1302 cysteines (21%) were only enriched by IAA (Figure S3B). The reproducibility and robustness of this method was tested with five human cell lines, HeLa, HeLa S3, Jurkat, PC3, and NCIH460. The correlation coefficients between biological replicates in any one of the five cell lines were all above 0.90 (Figure 1A-B). Representative spectra generated from IAA-labeled and EBX-labeled cysteine-containing peptides exhibited high quality, as shown in Figure 1C, 1D and S4. Taken together, the combination of two complementary chemical probes with UV biotin-azide facilitated the most comprehensive coverage of ligandable cysteines.

Multiplexed target profiling of electrophiles with the approach chemical genomics-facilitated chemical proteomics (CGCP). We next verified the associations between the reactivity of electrophilic compounds towards cysteines are of high, medium, and low similarity to Cel, in terms of their perturbations on global gene expression. Using Cel as a reference compound in L1000 CMap database, Aur, WA, and TL get a connectivity score of 98, 96, 56, respectively. Vin, with a connectivity score of 0, was regarded as the negative control (Figure 2A). Notably, Aur and WA demonstrated the most similarity to Cel in terms of global gene expression regulation, while TL showed medium similarity, and Vin exhibited the least similarity. When we applied a cysteine profiling strategy to these compounds simultaneously, Aur, and WA presented similar profiles to that of Cel in the heatmap (Figure 2B), with broad-spectrum cysteines as potential binding sites (Table S1). The result was consistent when we analyzed the %inhibition upon compound treatment, with the order of median values being Cel (74) > Aur (72) > WA (71) > TL (12) > Vin (7) (Figure 2C). There were 642 cysteines with high reactivities shared between Cel and Aur groups, 698 cysteines with high reactivities shared between Cel and WA groups, 560 cysteines with high reactivities shared by Cel, Aur and WA groups (Figure 2D). TL bound fewer reactive cysteines than Cel, with 134 cysteines in common. Vin showed the minimum inhibition on cysteine reactivity to IAA, with few common targets with Cel. To further evaluate the cysteine-binding reactivity quantitively, we adopted a mathematic formula to evaluate the cysteine-binding abilities based on the chemical proteomics results (Figure 2E and S5). When Cel was used as the reference, its cysteine-binding ability was set as the baseline y=0, and any curve of the compound below y=0 indicated lower reactivity between the compound and a specific cysteine compared to the reference. No matter which compound was set as the reference, the curves of Cel, Aur, and WA were close to each other and kept a similar pattern in cysteines with both high %inhibition and low %inhibition. Taken together, we found that the cysteine reactivity with electrophilic compounds is associated with their induced cellular perturbations of global gene expression.

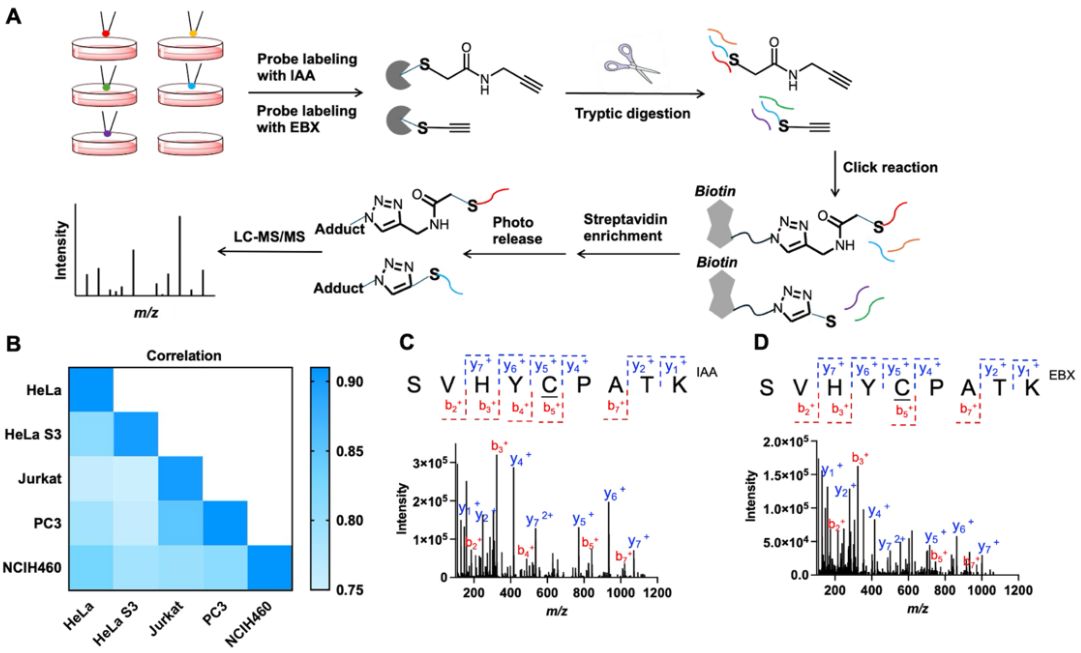

Figure 1. Method establishment of cysteine profiling by the simultaneous use of two chemical probes. (A) General workflow of optimized cysteine profiling method. (B) Correlation analysis of our method in five cell lines. (C and D) Typical spectra of probe-labeled peptides of protein MCM3.

Targeted MS strategy to validate the small molecule-protein interactions. With the CGCP approach, over 1800 cysteines were identified and over 1200 reactive cysteines with %inhibition > 15 upon treatment were regarded to be potential binding sites of the selected electrophiles. We evaluated the effectiveness of the CGCP approach by examining reported interactions between electrophiles and target proteins as benchmarks (Figure 3A, 3B, and S6). The known binding site of Cel, Cys40 of prostaglandin E synthase 3 (PTGES3)[27], and the known binding sites of TL, Cys83 and Cys173 of peroxiredoxin-1 (PRX1)[28] were monitored with the targeted MS method parallel reaction monitoring (PRM). Compared to the DMSO group, reduced intensities upon compound treatment indicate binding occupancy of the compound in individual cysteine. As expected, the binding of Cel to Cys40 of PTGES3, and TL to Cys83/Cys173 of PRX1 were observed by PRM, which agree with the results obtained using CGCP. We also examined the reported target proteins in which the binding sites were unknown. CGCP effectively revealed potential bindings sites of the electrophiles in their reported target proteins, including Cys328 of vimentin as the target of WA[29], and Cys100 of peroxiredoxin-5 (PRX5) to bind Aur[30]. Previously the C-terminal domain of heat shock protein 90β (Hsp90β) was predicted as the target region of Hsp90β inhibitors such as Cel[31] and WA[32]. Based on the CGCP results, Cys564 of Hsp90β was identified to be the potential binding site (Figure 3B).

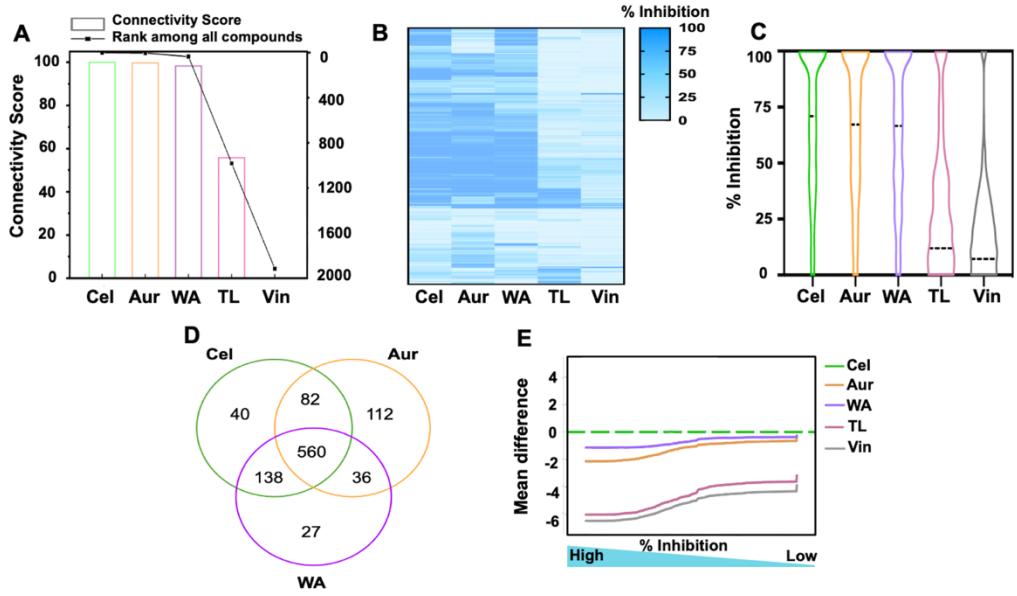

Figure 2. Cysteine profiling to evaluate the correlation between genome perturbations of electrophiles and their reactivity towards cysteines. (A) The connectivity score and rank of indicated electrophile with Cel as a reference, based on chemical genomics analysis. Score= 100 or rank as top 1 indicates extreme similarity to Cel. (B) A heatmap and (C) a violin plot shows the results of a competition experiment and proteomics quantification. Live cells were incubated with the indicated compounds followed by probe labeling. Cysteines that react with electrophiles will no longer get labeled by probe and thus cause decreased MS intensity. The percentage reduction of probe labeling 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. We defined cysteines with %inhibition > 50 as cysteines with high reactivities to the compound. (D) A Venn diagram shows shared cysteines with high reactivities (%inhibition > 50) upon Cel, Aur and WA treatment. (E) Evaluation of cysteine-binding abilities of electrophiles with a mathematical measurement. Cel was set as the reference with y=0, and mean difference with minus value showed lower cysteine-binding ability than the reference.

Biochemical assays in two representative proteins PRX1 and Hsp90β to verify binding interactions. The novel electrophile-target protein interactions identified with CGCP were further validated with several independent assays. As shown in Figure 4A, in-gel fluorescence scanning revealed that probe labeling of PRX1 by IAA or EBX were abolished after incubation with Cel or WA at a relatively low concentration (5 μM), indicating strong interactions. On the other hand, Aur and TL exhibited a concentration-dependent manner of interaction, presenting strong binding competitiveness at high concentration (50 μM). Vin showed minimal competition with the probe for PRX1 labeling even at the high concentration. The in-gel fluorescence results were in agreement with the MS results (Figure S7). In a thermal shift assay, four binders of PRX1 stabilized the endogenous protein and changed the thermal denaturation temperature, whereas the vehicle (DMSO) and Vin failed to do so (Figure 4B).

To further confirm the binding sites identified with CGCP, we compared the binding of electrophiles to wild-type and mutant target proteins (Figure 4C and 4D). Cys83 and Cys173 of PRX1 were identified to be binding sites shared by TL, Aur, WA and Cel. A synthetic probe Cy3-TL[28] was used to label recombinant wild-type and mutant PRX1 for in-gel fluorescence visualization (Figure 4C). Consistent with the results shown in Figure 4A, Cel, Aur, WA, and TL effectively inhibited probe labeling

of wild-type PRX1 while Vin did not show any inhibition. After Aur, WA, and TL treatment, the C83S mutant of PRX1 showed increased labeling compared to wild-type PRX1, indicating Cys83 in PRX1 as a binding target of Aur, WA, and TL. Similarly, after Cel, Aur, WA and TL treatment, the C173S mutant of PRX1 showed increased labeling than wild-type PRX1, indicating Cys173 in PRX1 as a binding target of Cel, Aur, WA, and TL. These findings illustrated the important roles of both Cys83 and Cys173 in the formation of the PRX1-compound complex. The increased fluorescence intensities in the C83S and C173S mutants compared to the wild-type PRX1 in DMSO group were likely due to structural changes of PRX1, which dissociated from its multimeric form[33] upon Cys83 and Cys173 mutations, exposing additional cysteine residues that can bind the probe. Cel, which has a strong absorption at 440 nm and is orange in color, has its chromophore disrupted by covalent binding to reactive cysteines in target protein, leading to decreased absorption in a UV-visible assay[34]. The wild-type PRX1 dramatically decreased Cel's absorption while C83S and C173S mutants had only partial inhibition, indicating the importance of both cysteines for binding (Figure 4D).

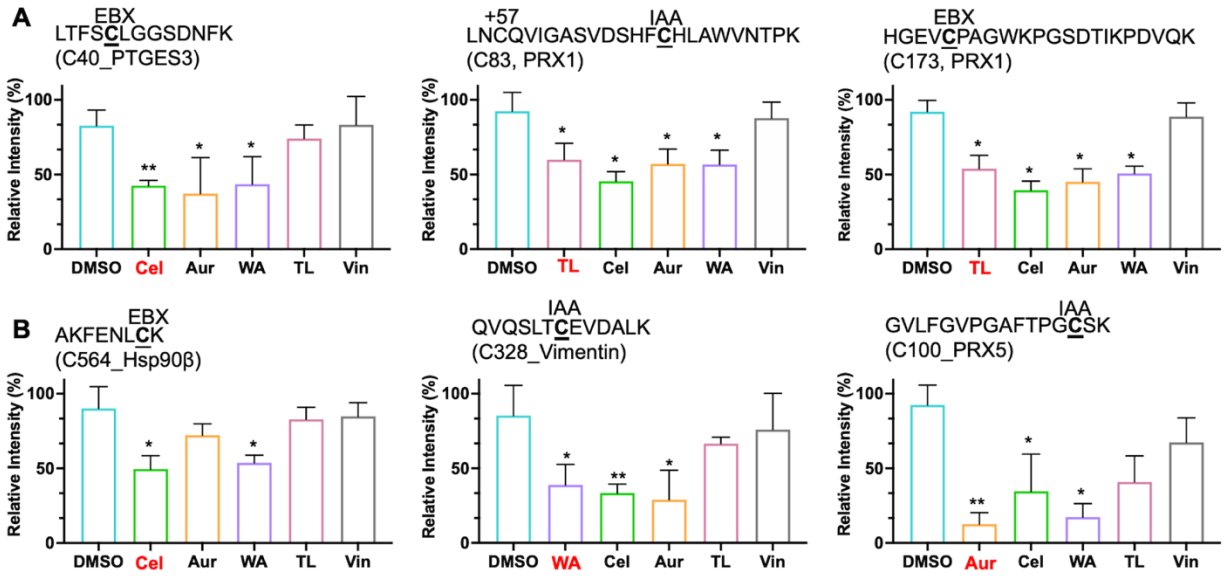

Figure 3. Verification of known and novel interactions between tested electrophiles and proteins with targeted MS. (A) Detected interactions between electrophiles and known binding sites in reported target proteins. Red color highlights known ligands of the binding sites. (B) Detected interactions between electrophiles and novel binding sites of target proteins. Red color highlights reported ligands of the indicated proteins. Data are 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.

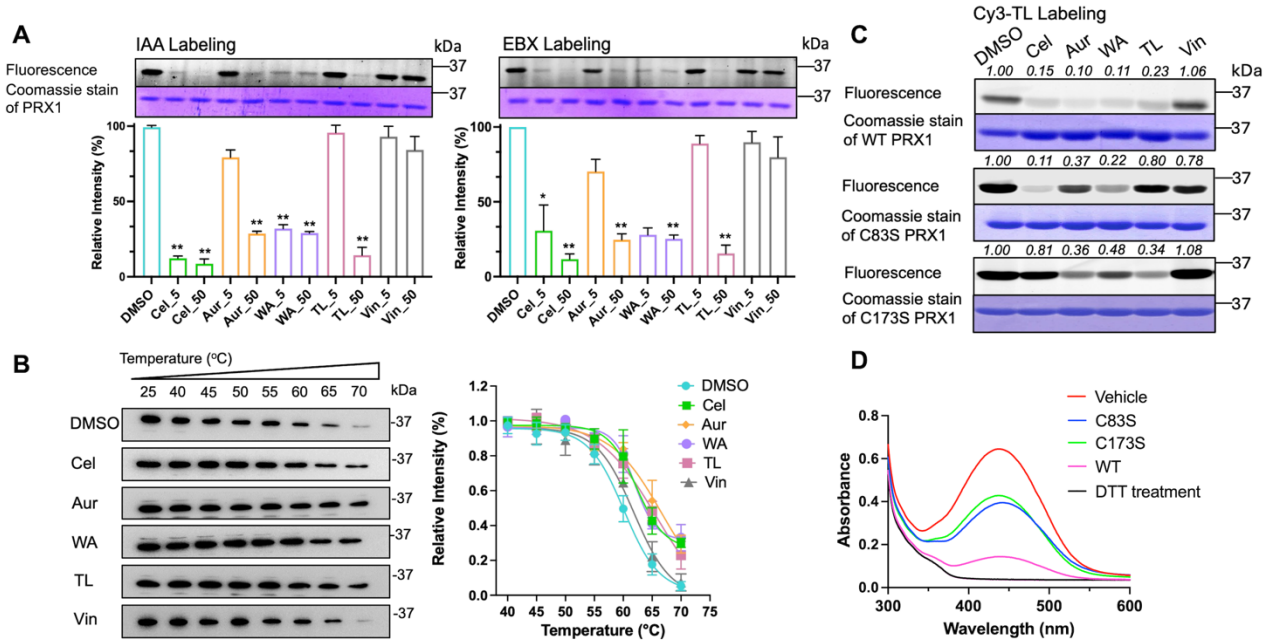

Figure 4. Verification of electrophile-target protein interactions in PRX1. (A) In-gel fluorescence scan of recombinant PRX1 labeled with IAA or EBX probe, in presence or absence of the indicated compounds as competitors. (B) Cellular thermal shift assay (CETSA) shows the thermostability change of PRX1 after incubation with indicated compounds (50 μM) or vehicle (right panel). One representative western blot was shown (left panel) quantification based on three independent experiments. (C) In-gel fluorescence scan of recombinant PRX1 and its mutants labeled with Cy3-TL probe in presence or absence of the indicated compounds (50 μM) as competitors. For each gel image, the relative intensities were calculated, when the DMSO-treated group was regarded as 1.00. (D) UV absorption of Cel after incubation with recombinant PRX1, its mutants or DTT. Decrease of absorption indicates interaction and disruption of the chromophore of Cel.

As another representative, wild-type and mutant Hsp90β were used to validate the novel electrophile-target protein interactions. As shown in Figure 5A, it was indicated that the probe labeling of Hsp90β by IAA or EBX were inhibited after incubation with Cel at a relatively low concentration (5 μM), suggesting strong interactions. On the other hand, Aur and WA showed a concentration-dependent manner of interaction, presenting strong binding competitiveness at a high concentration (50 μM). In agreement with the MS results shown in Figure 3B, TL and Vin barely competed with the probe to label Hsp90β even at the high concentration. The selectivity mechanism of Cel and WA adduct to Cys564 in Hsp90β was further revealed by molecular docking (Figure 5B). We found that both Cel and WA possessed low docking scores (around -5) with Cys564 in Hsp90β, revealing the probability of ligand interactions.

Next, to confirm the binding site identified with CGCP and predicted by docking, we compared the binding of electrophiles to wild-type and mutant target proteins (Figure 5C and 5D). Cel, Aur, and WA effectively inhibited the probe labeling of wild-type Hsp90β. After Cel and WA treatment, the C564S mutant of Hsp90β showed increased labeling compared to wild-type Hsp90β, indicating Cys564 in Hsp90β as a binding target of Cel and WA (Figure 5C). These findings were consistent with our targeted MS results (Figure 3B), which together suggested that Cel and WA can target Cys564 of Hsp90β, while Aur may not mainly target Cys564 of Hsp90β. In addition, with an UV-visible assay, the wild-type Hsp90β dramatically decreased Cel's absorption while the C564S mutant had only partial inhibition, indicating the importance of Cys564 for binding (Figure 5D).

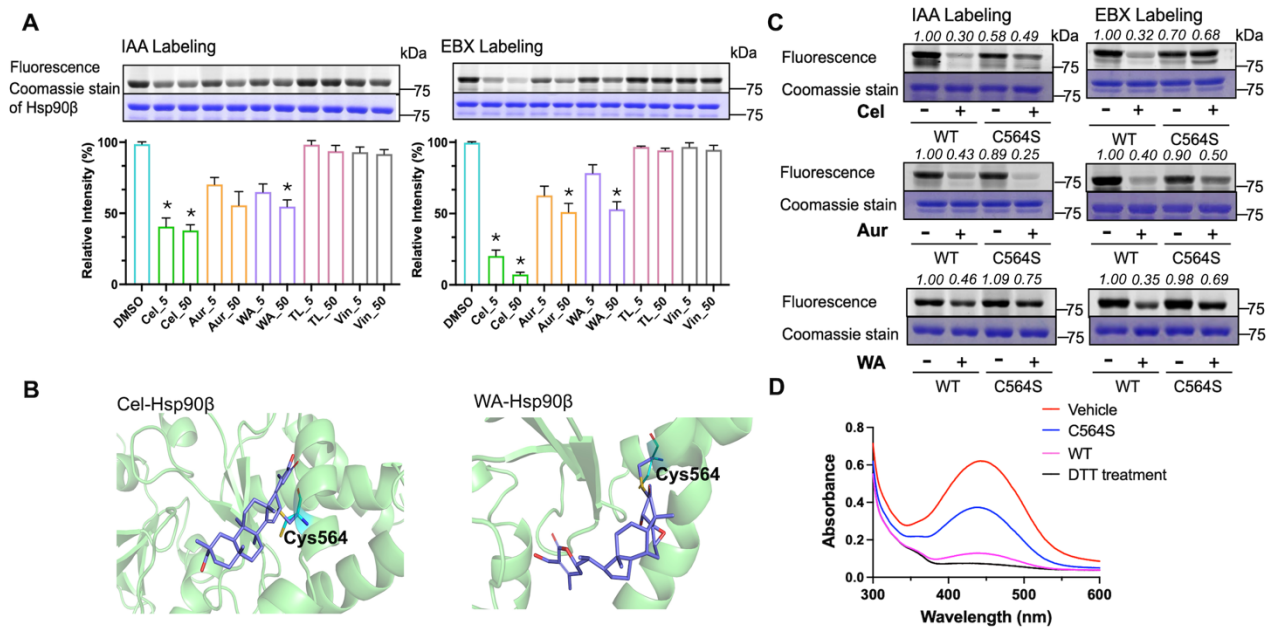

Figure 5. Verification of novel compound-protein interactions in Hsp90β. (A) In-gel fluorescence scan of recombinant Hsp90β treated with indicated compounds at 5 μM or 50 μM followed by IAA or EBX probe labeling. (B) Molecular docking shows the interactions of Cel or WA with the Cys564 by forming a covalent bond (PDB entry 5FWM). (C) In-gel fluorescence scan of recombinant Hsp90β and its mutants treated with indicated electrophiles (50 μM) followed by probe labeling. For each gel image, the relative intensities were calculated, when the intensity of DMSO-treated group in WT was regarded as 1.00. (D) UV absorption of Cel after incubation with the indicated compound or proteins. Decrease of absorption indicates interaction with Cel and disruption of its chromophore.

CONCLUSIONS

In this work, we developed an approach named CGCP that integrates chemical genomics and chemical proteomics for multiplexed target profiling of structurally distinct electrophiles. As a proof-of-concept study, we demonstrated the utility of CGCP by target profiling of Cel and other structurally distinct electrophilic compounds that are of high, medium, and low similarity to Cel, according to their cellular perturbation of gene expression. Cysteine profiling with two complementary chemical probes led to the identification of both known and novel binding targets of the tested electrophiles. Notably, we found that the cysteine reactivity with electrophilic compounds is associated with their cellular perturbations of global gene expression. Using CGCP, we identified common targets of Cel, Aur, WA and TL, and validated the newly identified binding proteins and binding sites with targeted MS and biochemical assays.

CGCP enables multiplexed chemical proteomics and thus enhance the efficiency of target identification, particularly for structurally diverse compounds. Comparing to the traditional chemical proteomics strategies for target identification, which require examining compounds individually or only analyzing those with high structural similarity, CGCP offers an innovative approach to compound clustering for multiplexed target profiling. It is compatible to traditional chemical proteomics strategies for target identification such as activity-based protein profiling (ABPP) and many reported chemical probes. The current version of CGCP can be readily used when the compounds of interest as well as the cell lines are documented in L1000 database, so their chemical genomic profiles are most relevant. If other compounds and cell lines are the study focus, researchers need to

acquire RNA sequencing data and perform the data analysis as described previously[10]. In addition, the site identification number in CGCP can be improved by using labeled quantification techniques like tandem mass tag (TMT)[35].

It is foreseeable that the CGCP strategy should be applicable to a broad spectrum of compounds be-yond the proof-of-concept cases shown in the present study. For example, lysine-targeting molecules could be examined with lysine-reactive chemical probes[36] and proteome profiling methods for CGCP target identification. Theoretically, CGCP should also be applicable to compounds without obvious electrophilic properties. There are many good methods for target profiling of noncovalent drugs, such as by using photo-crosslinking probes, or TPP[4], TRAP[7], etc. CGCP is compatible with these methods because the major difference lays in the classification of compounds. In CGCP, compounds are grouped based on their chemical genomic profiles, which is independent of their selectivity of binding residues or mode of actions. Therefore, simply adaption or integration with other methods/probes will enable broader applications of CGCP to study more compounds. Following target profiling by CGCP, pharmacists can validate the efficacy of small molecules for specific protein targets through pharmacological phenotyping studies, thereby enhancing the drug development process.

EXPERIMENTAL SECTION

Cell culture. HeLa, HeLa S3 were grown in DMEM medium. Jurkat and NCI-H460 cells were grown in RPMI-1640 medium. PC3 cell was grown in Ham's F12K medium. All mediums were supplemented with 10% fetal bovine serum (FBS), 100 unit/mL penicillin and 100 μg/mL streptomycin. All cells were cultured at 37 °C with 5% CO_2.

Query in L1000 CMap database. There is a web-based searching engine and here is the searching link using Celastrol as the query: https://clue.io/command?q=/conn%20%22celastrol%22. The CMap con-nectivity score is a standardized measure that ranges from -100 to 100. For each perturbagen in the query results, the score reflects similarity to the current query. The data is available for download, al-lowing for easy and user-friendly access.

In-gel fluorescence scan. For HeLa S3 cell lysate (2 μg/uL, 50 uL), probe labeling was performed with alkynyl iodoacetamide (IAA) (10, 50, 100 μM) or TMS-ethynylbenziodoxolone (EBX) (10, 50, 100 μM) for 1 hour at room temperature (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 $CuSO_4$. After incubation for 1.5 hours at RT, reactions were quenched by acetone protein precipitation. Protein pellets were resuspended in sample buffer and resolved by SDS-PAGE. Gels were scanned for fluorescently labeled proteins (Typhoon Imaging System, GE), then stained with Coomassie blue.

Compound treatment and cell lysis. PC3 cells were grown until 90% confluency, treated with 10 μM compounds or vehicle for 60 min at 37°C and then rinsed two times with 1 × Phosphate Buffered Saline (PBS). The purity of compounds was determined to be >95% by HPLC. 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 min (10 second sonication, 15 second pause) at 30% amplitude under ice cooling. After centrifugation at 16,000 g, 4 °C for 20 min, the supernatant fraction was collected and adjusted to 2.0 μg/μL by BCA assay. For HeLa, HeLa S3, Jurkat and NCI-H460 cell lines without compound treatment, the procedure is same with PC3 cells.

Probe labeling and proteomic sample preparation (click reaction in peptides). Protein lysates were labeled with 100 μM IAA or 50 μM EBX for 1 hour at RT. Proteins were then reduced with 10 mM dithiothreitol (DTT) and incubated at 45 ºC for 30 min, and further alkylated with 30 mM iodoacetamide (IAM) at RT for 30 min in the dark. Protein precipitation was performed with a

methanol-chloroform system (aqueous phase/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 by adding sequencing-grade trypsin (Promega) at an enzyme/protein ratio of 1:50 (w/w) for additional 12 hours at 37 °C. Then the 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 ultraviolet (UV)-cleavable biotin-azide or DDE biotin-azide, 2 mM TCEP, 300 μM TBTA ligand in 1:4 DMSO:t-butyl alcohol and 2 mM $CuSO_4$. Samples were allowed to react at RT for 1 h in the dark with rotation. The excess biotin was removed by strong cation exchange (SCX). 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 min under gently 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 gently shaking. The supernatant was collected by centrifugation (1,200 g, 2 min), cleaned by self-packed C18 stage-tip, dried under vacuum, and stored at −80 °C until MS analysis.

Liquid chromatography (LC)-MS/MS analysis and data processing. LC-MS/MS analyses were performed on Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) coupled with an UltiMate 3000 UPLC system (Thermo Fisher Scientific). A RSLC C18 analytical column (75 μm × 250 mm, 2.0 μm, 100 Å) (Thermo Fisher Scientific) was employed for LC separation. Mobile phases A and B consist of 0.1% FA in water and 0.1% FA in ACN, respectively. A 120 min length gradient at a flow rate of 300 nL/min and an initial 2% mobile phase B was used. Mobile phases B was increased to 4% at 12 min, 30% at 88 min, 85% at 104 min and held for 5 min. Then, mobile phase B was back to 2% at 110 min and maintained this composition until 120 min. 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, and the MS2 AGC target and the maximum injection time were set at 1×10^5 and 50 ms. The acquired label-free DDA data were searched against the homo sapiens UniProt database (Version June 2020, 20368 entries) using the SEQUEST algorithm (Proteome Discoverer 2.4, Thermo Fisher Scientific). Precursor ion mass tolerance and fragmentation tolerance were set as 10 ppm and 0.02 Da for the database search. 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%.

ASSOCIATED CONTENT

Supporting Information

Experimental section (including chemicals and materials, sample preparation of cell lysate and mouse tissue, mathematical measurement building and evaluation, validation of protein targets using PRM-based assay, plasmid construction, protein over expression and purification, UV-visible assay, thermal shift assay, molecular docking studies, and purity determination), and figures (including chemical structures of probes, cleavable biotins, and compounds; method optimization of cysteine profiling; Identification number of cysteines from different cell lines and mouse tissue sample; characterization of reactive cysteines by a mathematical model.)

Predicted binding model of Celastrol with Hsp90β (PDB)
Predicted binding model of Withaferin-a with Hsp90β (PDB)
Molecular formula strings (CSV)
Table of identified reactive cysteines by MS in PC3 cells.

AUTHOR INFORMATION

*Corresponding Author
Qian Zhao – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China; orcid.org/ 0000-0003-2244-6516; Email: q.zhao@polyu.edu.hk
Authors
Yang Yang – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China; Centre for Eye and Vision Research, 17W Hong Kong Science Park, Hong Kong, SAR 999077, China; orcid.org/ 0000-0002-1269-2804.
Yin-suen Tse – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China; Laboratory for Synthetic Chemistry and Chemical Biology Limited, Hong Kong, SAR 999077, China.
Qi Zhang – Centre for Eye and Vision Research, 17W Hong Kong Science Park, Hong Kong, SAR 999077, China
Kin-yau Wong – Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China; orcid.org/0000-0001-9066-1619.
Chenxi Yang – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China.
Ying Yang – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China.
Shuqi Li – State Key Laboratory of Chemical Biology and Drug Discovery, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China.
Kin-wa Lau – Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, SAR 999077, China.
Trevor C. Charles– Department of Biology, University of Waterloo, Waterloo, Canada; orcid.org/0000-0002-0344-5932.
Thomas C. Lam – Centre for Eye and Vision Research, 17W Hong Kong Science Park, Hong Kong, SAR 999077, China; orcid.org/0000-0002-3511- 5620.
Notes: The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank Prof. Jerome Waser for providing EBX reagent. We acknowledge the funding support from NSFC 21705136, Research Grants Council-GRF 15304819, GRF 15307122, CRF C5033-19E, R5008-22, and ITC MRP/043/21. We thank the support from Centre for Eye and Vision Research (CEVR) and Laboratory for Synthetic Chemistry and Chemical Biology Limited (LSCCB) under the Health@InnoHK Programme launched by ITC, HKSAR. We thank the PolyU Research Facilities UCEA and ULS, RCSV, RiFood, RCMI for technical support.

ABBREVIATIONS USED

ABPP, activity-based protein profiling; Aur, auranofin; Cel, celastrol; CMap, Connectivity Map; CETSA, cellular thermal shift assay; CGCP, chemical genomics-facilitated chemical proteomics; DARTS, drug affinity responsive target stability; EBX, ethynylbenziodoxolone; Hsp90β, heat shock protein 90β; IAA, alkynyl iodoacetamide; MS, mass spectrometry; MTRP, multiplexed thiol reactivity profiling; PRM, parallel reaction monitoring; PRX1, peroxiredoxin-1;TPP, thermal proteome profiling; TRAP, target responsive accessibility profiling; TL, triptolide; Vin, vinpocetine; WA, withaferin-a.

REFERENCES

(1) Schenone, M.; Dancik, V.; Wagner, B. K.; Clemons, P. A., Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* **2013,** *9* (4), 232-40.

(2) 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.

(3) Chen, X.; Wang, Y.; Ma, N.; Tian, J.; Shao, Y.; Zhu, B.; Wong, Y. K.; Liang, Z.; Zou, C.; Wang, J., Target identification of natural medicine with chemical proteomics approach: probe synthesis, target fishing and protein identification. *Signal Transduct Target Ther* **2020,** *5* (1), 72.

(4) 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.

(5) Pai, M. Y.; Lomenick, B.; Hwang, H.; Schiestl, R.; McBride, W.; Loo, J. A.; Huang, J., Drug affinity responsive target stability (DARTS) for small-molecule target identification. *Methods Mol Biol* **2015,** *1263*, 287-98.

(6) 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.

(7) Yan, W.; Wang, D.; Wan, N.; Wang, S.; Shao, C.; Zhang, H.; Zhao, Z.; Lu, W.; Tian, Y.; Ye, H.; Hao, H., Living Cell-Target Responsive Accessibility Profiling Reveals Silibinin Targeting ACSL4 for Combating Ferroptosis. *Anal Chem* **2022**, *94* (43), 14820-14826.

(8) 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* **2018,** *19* (3), 506-523.

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

(10) 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.

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

(12) Venkatesha, S. H.; Dudics, S.; Astry, B.; Moudgil, K. D., Control of autoimmune inflammation by celastrol, a natural triterpenoid. *Pathog Dis* **2016,** *74* (6).

(13) Gao, Q.; Qin, H.; Zhu, L.; Li, D.; Hao, X., Celastrol attenuates collagen-induced arthritis via inhibiting oxidative stress in rats. *Int Immunopharmacol* **2020,** *84*, 106527.

(14) 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 Lett* **2011,** *303* (1), 9-20.
(15) Zhang, B.; Zhong, Q.; Chen, X.; Wu, X.; Sha, R.; Song, G.; Zhang, C.; Chen, X., Neuroprotective Effects of Celastrol on Transient Global Cerebral Ischemia Rats via Regulating HMGB1/NF-kappaB Signaling Pathway. *Front Neurosci* **2020,** *14*, 847.
(16) Shi, J.; Li, J.; Xu, Z.; Chen, L.; Luo, R.; Zhang, C.; Gao, F.; Zhang, J.; Fu, C., Celastrol: A Review of Useful Strategies Overcoming its Limitation in Anticancer Application. *Front Pharmacol* **2020,** *11*, 558741.
(17) Cox, A. R.; Masschelin, P. M.; Saha, P. K.; Felix, J. B.; Sharp, R.; Lian, Z.; Xia, Y.; Chernis, N.; Bader, D. A.; Kim, K. H.; Li, X.; Yoshino, J.; Li, X.; Li, G.; Sun, Z.; Wu, H.; Coarfa, C.; Moore, D. D.; Klein, S.; Sun, K.; Hartig, S. M., The rheumatoid arthritis drug auranofin lowers leptin levels and exerts antidiabetic effects in obese mice. *Cell Metab* **2022**, *34* (12), 1932-1946.e7.
(18) Hassannia, B.; Logie, E.; Vandenabeele, P.; Vanden Berghe, T.; Vanden Berghe, W., Withaferin A: From ayurvedic folk medicine to preclinical anti-cancer drug. *Biochem Pharmacol* **2020**, *173*, 113602.
(19) Song, J.; He, G.-N.; Dai, L., A comprehensive review on celastrol, triptolide and triptonide: Insights on their pharmacological activity, toxicity, combination therapy, new dosage form and novel drug delivery routes. *Biochem Pharmacol* **2023**, *162*, 114705.
(20) Zhang, Y.; Li, J.; Yan, C., An update on vinpocetine: New discoveries and clinical implications. *Eur J Pharmacol* **2018**, *819*, 30-34.
(21) 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.
(22) Tian, Y.; Wan, N.; Zhang, H.; Shao, C.; Ding, M.; Bao, Q.; Hu, H.; Sun, H.; Liu, C.; Zhou, K.; Chen, S.; Wang, G.; Ye, H.; Hao, H., Chemoproteomic mapping of the glycolytic targetome in cancer cells. *Nat Chem Biol* **2023**, *19* (12), 1480-1491.
(23) Boike, L.; Henning, N. J.; Nomura, D. K., Advances in covalent drug discovery. *Nat Rev Drug Discov* **2022**, *21* (12), 881-898.
(24) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010,** *468* (7325), 790-5.
(25) Abegg, D.; Frei, R.; Cerato, L.; Prasad Hari, D.; Wang, C.; Waser, J.; Adibekian, A., Proteome-Wide Profiling of Targets of Cysteine reactive Small Molecules by Using Ethynyl Benziodoxolone Reagents. *Angew Chem Int Ed Engl* **2015,** *54* (37), 10852-7.
(26) Tessier, R.; Nandi, R. K.; Dwyer, B. G.; Abegg, D.; Sornay, C.; Ceballos, J.; Erb, S.; Cianferani, S.; Wagner, A.; Chaubet, G.; Adibekian, A.; Waser, J., Ethynylation of Cysteine Residues: From Peptides to Proteins in Vitro and in Living Cells. *Angew Chem Int Ed Engl* **2020,** *59* (27), 10961-10970.
(27) 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.
(28) 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.
(29) 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.
(30) 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 Biology* **2020,** *32*, 101491.

(31) Zhang, T.; Li, Y.; Yu, Y.; Zou, P.; Jiang, Y.; Sun, D., Characterization of celastrol to inhibit hsp90 and cdc37 interaction. *J Biol Chem* **2009,** *284* (51), 35381-9.
(32) Yu, Y.; Hamza, A.; Zhang, T.; Gu, M.; Zou, P.; Newman, B.; Li, Y.; Gunatilaka, A. A.; Zhan, C. G.; Sun, D., Withaferin A targets heat shock protein 90 in pancreatic cancer cells. *Biochem Pharmacol* **2010,** *79* (4), 542-51.
(33) Barranco-Medina, S.; Lázaro, J.-J.; Dietz, K.-J., The oligomeric conformation of peroxiredoxins links redox state to function. *FEBS Letters* **2009**, *583* (12), 1809-1816.
(34) 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.
(35) Vinogradova, E. V.; Zhang, X.; Remillard, D.; Lazar, D. C.; Suciu, R. M.; Wang, Y.; Bianco, G.; Yamashita, Y.; Crowley, V. M.; Schafroth, M. A.; Yokoyama, M.; Konrad, D. B.; Lum, K. M.; Simon, G. M.; Kemper, E. K.; Lazear, M. R.; Yin, S.; Blewett, M. M.; Dix, M. M.; Nguyen, N.; Shokhirev, M. N.; Chin, E. N.; Lairson, L. L.; Melillo, B.; Schreiber, S. L.; Forli, S.; Teijaro, J. R.; Cravatt, B. F., An Activity-Guided Map of Electrophile-Cysteine Interactions in Primary Human T Cells. *Cell* **2020**, *182* (4), 1009-1026.e29.
(36) Cuesta, A.; Taunton, J., Lysine-Targeted Inhibitors and Chemoproteomic Probes. *Annu Rev Biochem* **2019**, *88*, 365-381.

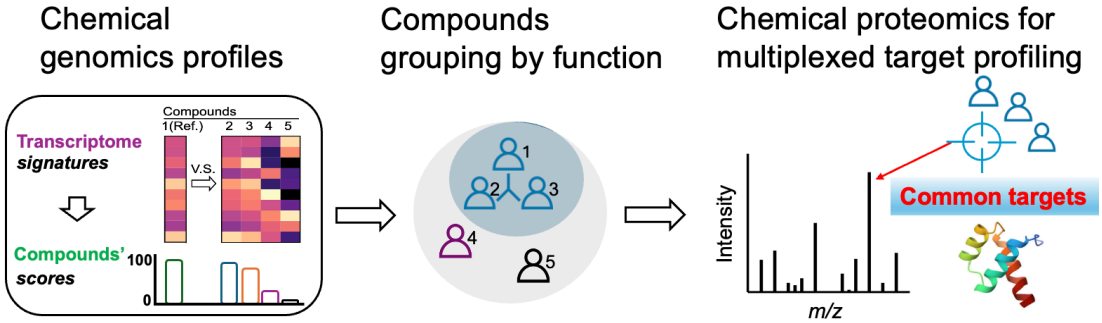

Table of Content