Nature Communications

https://doi.org/10.1038/s41467-025-67403-2

Article in Press

Enantioselective protein affinity selection mass spectrometry (E-ASMS)

Received: 6 February 2025

Accepted: 26 November 2025

Cite this article as: Wang, X., Sun, J., Ahmad, S. *et al.* Enantioselective protein affinity selection mass spectrometry (E-ASMS). *Nat Commun* (2025). https://doi.org/10.1038/s41467-025-67403-2

Xiaoyun Wang, Jianxian Sun, Shabbir Ahmad, Diwen Yang, Fengling Li, U. Hang Chan, Hong Zeng, Conrad V. Simoben, Stuart R. Green, Madhushika Silva, Scott Houliston, Aiping Dong, Albina Bolotokova, Elisa Gibson, Maria Kutera, Pegah Ghiabi, Ivan Kondratov, Tetiana Matviyuk, Alexander Chuprina, Danai Mavridi, Christopher Lenz, Andreas C. Joerger, Benjamin D. Brown, Richard B. Heath, Wyatt W. Yue, Lucy K. Robbie, Tyler S. Beyett, Susanne Müller, Stefan Knapp, Dafydd R. Owen, Rachel Harding, Matthieu Schapira, Peter J. Brown, Vijayaratnam Santhakumar, Suzanne Ackloo, Cheryl H. Arrowsmith, Aled M. Edwards, Hui Peng & Levon Halabelian

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

Enantioselective Protein Affinity Selection Mass Spectrometry (E-ASMS)

Xiaoyun Wang[1,#], Jianxian Sun[1,2,#], Shabbir Ahmad[2,#], Diwen Yang[3], Fengling Li[2], U Hang Chan[2,4], Hong Zeng[2], Conrad V. Simoben[2], Stuart R. Green[2], Madhushika Silva[2], Scott Houliston[5], Aiping Dong[2], Albina Bolotokova[2], Elisa Gibson[2], Maria Kutera[2,5,6], Pegah Ghiabi[2], Ivan Kondratov[7,8,9], Tetiana Matviyuk[7,10], Alexander Chuprina[7], Danai Mavridi[11,12], Christopher Lenz[11,12], Andreas C. Joerger[11,12], Benjamin D. Brown[13], Richard B. Heath[13], Wyatt W. Yue[13], Lucy K. Robbie[14], Tyler S. Beyett[14], Susanne Müller[11,12], Stefan Knapp[11,12], Dafydd R. Owen[15], Rachel Harding[2,4,5,16], Matthieu Schapira[2,4,5], Peter J. Brown[17], Vijayaratnam Santhakumar[2], Suzanne Ackloo[2], Cheryl H. Arrowsmith[2,5,6], Aled M. Edwards[2,5], Hui Peng[1,2,3,5*], Levon Halabelian[2,4,5*]

[1]Department of Chemistry, University of Toronto, Toronto, ON, Canada

[2]Structural Genomics Consortium, University of Toronto, Toronto, ON, Canada

[3]Department of Physical & Environmental Sciences, University of Toronto Scarborough, Toronto ON, Canada

[4]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada

[5]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

[6]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

[7]Enamine Ltd., Winston Churchill Street 78, 02094 Kyiv, Ukraine

[8]Enamine Germany GmbH, Industriepark Hoechst, G837, 65926 Frankfurt am Main, Germany

[9]V. P. Kukhar Institute of Bioorganic Chemistry and Petrochemistry, National Academy of Sciences of Ukraine, Akademik Kukhar Street 1, 02094 Kyiv, Ukraine

ARTICLE IN PRESS

ARTICLE IN PRESS

[10]Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA, USA.

[11]Institute of Pharmaceutical Chemistry, Goethe University, Frankfurt am Main, Germany

[12]Structural Genomics Consortium, Goethe University, Frankfurt am Main, Germany.

[13]Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

[14]Emory University School of Medicine, Atlanta, GA, USA

[15]Pfizer Research and Development, Cambridge, MA 02139, USA

[16]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

[17]Structural Genomics Consortium, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[#]These authors contributed equally to this work: Xiaoyun Wang, Jianxian Sun, Shabbir Ahmad

*Corresponding authors: hui.peng@utoronto.ca; l.halabelian@utoronto.ca

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract

We report an enantioselective protein affinity selection mass spectrometry screening approach (E-ASMS) that enables the detection of weak binders, informs on selectivity, and generates orthogonal confirmation of binding. After method development with control proteins, we screen 31 human proteins against a designed library of 8,217 chiral compounds. We identify 16 binders to 12 targets, including many proteins predicted to be "challenging to ligand", and confirm their interactions through orthogonal biophysical assays. Seven binders to six targets display enantioselective binding, with K_D values ranging from 3 to 20 μM. Binders for four targets (DDB1, WDR91, WDR55, and HAT1) are selected for in-depth characterization using X-ray crystallography. In all four cases, the mechanisms underlying enantioselectivity are readily explained. These results demonstrate that E-ASMS enables the identification and characterization of selective and weakly binding ligands for novel protein targets with unprecedented throughput and sensitivity.

ARTICLE IN PRESS

ARTICLE IN PRESS

Introduction

Bioactive small molecules are invaluable reagents in basic research and applied fields such as biomedicine, agriculture, and microbiology. The discovery of bioactive small molecules typically begins with screening of synthetic chemical or natural product libraries. Often, these libraries are screened in high-throughput or phenotypic screening[1-3] to identify compounds that alter biochemical or cellular functions. Alternatively, target-centric approaches can be employed, such as methods that measure binding to the target protein directly *in vitro*, including fragment-based lead discovery (FBLD).[4-6] However, FBLD requires significant investment of medicinal chemistry resource to optimize the fragments into high-affinity ligands. DNA-encoded library (DEL) selection[7] and affinity selection mass spectrometry (AS-MS)[8] are two other well-established assays for hit discovery. While these high-throughput screening approaches have been widely used for ligand discovery, a significant challenge across all screening strategies is to develop orthogonal assays to distinguish true hits from false positives. Indeed, particularly when the initial hits bind weakly (> 10 μM K_D), hit verification often demands more time and resources than the initial screen. As a result, development of assays with improved sensitivity and selectivity remains essential to enable ligand discovery for large numbers of proteins.

Small molecules with a stereocenter can bind to their targets in an enantioselective way[9], which can be leveraged to identify potential stereoselective protein binders when screening libraries containing chiral compounds, as well as to develop negative controls for chemoproteomics studies and for experiments with chemical probes[10-12]. To exploit this phenomenon, in this work, we develop a scalable orthogonal screening strategy, termed "Enantioselective Protein Affinity Selection Mass Spectrometry (E-ASMS)", to both identify and characterize chemical ligands for previously unliganded proteins.

ARTICLE IN PRESS

Results

Design and benchmarking of an E-ASMS platform. The E-ASMS screening concept workflow (Fig. 1a) is a variation of previous methods used for bioactive natural product discovery[13-15], by including the orthogonal enantioselectivity evidence. First, pools of ~600 compounds from an 8,217-member screening library, comprising racemic mixtures of drug-like compounds at a concentration of 0.1 µM, are incubated with purified and quality-controlled polyhistidine-tagged proteins immobilized on magnetic nickel beads. At this concentration, all compounds are far below their solubility limit in aqueous solution. To embed a measure of selectivity in the screen, 8-12 different proteins are screened in parallel at a concentration of 1 µM in solution. After washing the beads, compounds that remain bound to the proteins and the beads, which comprise a combination of real and nonspecific binders, are eluted in methanol. A small aliquot of each methanol eluate is subject to liquid chromatography (LC) followed by high-resolution mass spectrometry (MS) to identify any compounds in the eluates, and to estimate their abundance relative to the eluates from other targets screened in parallel. Our concept is that some measure of binding specificity can be inferred by observing strong binding of a compound to one protein but not to the other proteins screened in parallel. For the compounds that appear to bind one target preferentially over the others, another aliquot from the same methanol eluate is analyzed by chiral chromatography, and the ratio of the two enantiomers in the protein eluate is assessed. If one of the two enantiomers is enriched in the protein eluate compared to their ratio in the screening library, then this provides orthogonal evidence for compound binding.

ARTICLE IN PRESS

Proof of concept for E-ASMS with known enantioselective protein-ligand pairs. To pilot the E-ASMS concept, we selected four chiral compounds known to bind to their respective protein partners in an enantioselective manner. The first test case was lenalidomide, a chiral analogue of thalidomide that binds enantioselectively to the cereblon (CRBN) protein[16-18]. As shown in Fig. 1b, when screened against four proteins (CRBN, DCAF1, PRMT5, and USP21), lenalidomide was selectively enriched in the methanol eluate from CRBN beads; no significant enrichment was observed on beads containing any of the three other human proteins, which served as negative controls. The CRBN methanol eluate was then subjected to chiral chromatography. We recovered (*S*)-lenalidomide in the eluate from CRBN-containing beads at far greater quantities compared to its (*R*-) enantiomer, despite both being present in equal amounts in the input. Enantioselective enrichment was also observed for (*S*)-OICR-6766, (*S*)-LLY283, and (*R*)-BAY805 to their cognate protein targets: DCAF1[19], PRMT5[20], and USP21[21] (Fig. 1c, 1d, and 1e).

Racemate library design and characterization. With the proof of concept in hand, we set out to apply the method to novel proteins in screening mode. To this end, we assembled a chemically diverse screening library with drug-like properties and high MS sensitivity, comprising 8,217 commercially available racemates (Supplementary Fig. 1 and Supplementary Table 1). The E-ASMS library was designed to maximize chemical diversity, encompassing 7,307 distinct scaffolds (Supplementary Fig. 2), with the top 10 scaffolds highlighted in Fig. 1f. The designed E-ASMS library adheres to Lipinski's rule of 5 (*e.g.*, MW, 119 – 520 Da; logP, -1.82 – 5.81; nHBD, 0 – 6; nHBA, 1 – 10; Fig. 1g). Chromatographic methods were developed to resolve the enantiomers for as many of the racemates as was practical. To accomplish this, each of the 8,217 racemic mixtures was subject to chiral liquid chromatography under eight different conditions

(four columns × two LC methods). In total, ~7,000 enantiomer pairs could be resolved under one of these conditions. The most effective chromatographic steps involved using IG and IA columns (Fig. 1h). Thus, for any of these ~7,000 compounds, it is possible to compare the ratio of enantiomers in the screening library and the protein eluant, as described in the workflow outlined above (Fig. 1a).

E-ASMS identifies low affinity binders. Rapid size-exclusion chromatography (SEC) is the most common approach used to resolve protein binders and non-binders in AS-MS library screening[8]. However, this method is known to be limited by the inability to detect weakly bound compounds, perhaps due to the challenges in retaining compounds with fast off-rates during chromatography. We explored whether the E-ASMS platform in which the protein is concentrated on beads and employs rapid wash steps, had the potential to enhance the sensitivity of AS-MS screens. Using 109 previously characterized ligands across eight proteins (Supplementary Table 2), we tested the sensitivity of E-ASMS. The E-ASMS platform successfully captured nearly all high-affinity (K_D < 1 μM) ligands, half of the moderate-affinity (1 μM ≤ K_D < 10 μM) ligands, and 20% of the low-affinity (K_D ≥ 10 μM) ligands (Fig. 1i) without substantially increasing the false positive rate. In contrast, all moderate- and low-affinity ligands were lost by the SEC-coupled AS-MS approach (Fig. 1i). The apparent sensitivity of E-ASMS is a significant advantage for chemical ligand discovery, or to assess the ligandability of a protein, as binders can be rapidly identified using a relatively small chemical library and minimal amounts of protein in a protein-agnostic manner.

Application of the E-ASMS platform to novel targets. With evidence of increased sensitivity and an ability to potentially provide orthogonal evidence for specificity using enantioselectivity,

we set out to explore the application of the E-ASMS platform to new proteins. We selected 31 human proteins with diverse biological functions, 3D structures, and with varying levels of known or predicted 'ligandability' (Fig. 2a) as assessed by their drug-like density (DLID) scores[22]. We purposely included several targets (*e.g.*, WDR91 and DCAF1) known to be 'ligandable'[19,23,24]. Among the 31 proteins, three were predicted to have high ligandability (DLID > 1), 11 of medium ligandability (0.5 < DLID < 1), and 16 of low ligandability (DLID < 0.5) (Supplementary Fig. 3). The 31 proteins were grouped in four batches of eight proteins and screened against 14 pools of ~600 compounds. In total, 118 compounds (candidate binders) exhibited > 5-fold enrichment for one protein compared to the others screened in parallel, corresponding to an average hit rate of 0.05% (Fig. 2b and Supplementary Table 3). Candidate binders were detected for all three high ligandable targets, seven of the eleven targets with medium ligandability scores, and nine of the sixteen targets predicted to be the most challenging. Interestingly, these binders were distributed across the entire chemical space of the E-ASMS library (Supplementary Fig. 4).

To confirm binding, we tested the binding of 101 candidate binders to each of the 19 targets using orthogonal biophysical methods. We used surface plasmon resonance (SPR), a technique that requires relatively small amounts of protein and can provide quantitative orthogonal confirmation of binding in a pocket-agnostic manner. The binding of 16 hits to 12 targets was confirmed by SPR, ranging from 1-4 hits per target and with K_D values ranging from 2 to 87 μM (Fig. 2c, 2d, and Supplementary Table 3). Notably, 12 of the 16 SPR-confirmed hits exhibited low-affinity (K_D ≥ 10 μM), which might have been overlooked by the conventional SEC-coupled AS-MS method (Fig. 1i). According to orthogonal evidence, we categorized 118 candidate binders into four confidence levels (Fig. 2c). Interestingly, the degree of enrichment in the primary E-ASMS screen

ARTICLE IN PRESS

showed a strong correlation with the K_D values determined for all SPR-validated hits (Fig. 2e), demonstrating the potential application of the E-ASMS method for semi-quantitative read-outs. Ideally, SPR assay methods are developed using a positive control binder, which we lacked for nearly all the proteins. Accordingly, positive results by SPR can be interpreted, but negative results are inconclusive. The 85 candidate binders that were unable to be confirmed using SPR are not necessarily false positives for three reasons. First, they might not be able to be detected because the assay for the protein target could not be optimized using a positive control compound. Second, the compounds might be insoluble at the concentrations required and conditions used for SPR (up to 0.2 mM of the compound) and might confound the read-outs[25]. This is a common issue when attempting to characterize weakly binding compounds by any biophysical method. Third, they might also be *bona fide* binders with affinities beyond the sensitivity of SPR, or not able to be detected under the non-optimized solution conditions used. Indeed, we tested nine candidate hits with spectral shift (SpS) method, and two of them were confirmed by SpS but not by SPR (Supplementary Table 4). These results clearly demonstrate that our E-ASMS success rate might be underestimated by using SPR only.

ARTICLE IN PRESS

Enantioselective enrichment provides orthogonal evidence of compound binding. Our E-ASMS method, which creates an effective protein concentration of ~100 µM on the beads, enables the screening of compounds at a concentration of 100 nM, far below their solubility limits. In support of solubility being a confounding issue for SPR orthogonal validation, we observed that SPR confirmation rates improved significantly for E-ASMS hits with higher fold enrichment values, which likely correlates with affinity. For E-ASMS hits without enantioselectivity that were enriched > 5-fold, the SPR-confirmation rate was 11.1%; this increased to 20.0% for more potent

ARTICLE IN PRESS

hits with fold enrichment values > 20 (Fig. 2f). For the more potent binders that exhibited enantioselective enrichment, the SPR confirmation rate further increased to 50%. Even subject to the caveats of SPR, this trend clearly demonstrates the value of enantioselective enrichment in hit confirmation.

Among the 118 candidate binders identified in the screen of 31 proteins, clear enantioselective binding was detected for 32 binders to 14 targets (Fig. 2c). Of these targets, we were successful in developing an SPR assay for 12 and could calculate K_D values for seven binders (Fig. 2d), with affinities ranging from 3-20 μM (Supplementary Figs. 5-9). There were 25 candidate binders for which we were unable to generate convincing binding data using either SPR under standard conditions or with other biophysical methods, such as differential scanning fluorimetry or ^{19}F-NMR. At this point, we cannot be certain if these 25 enantioselective candidate binders are false positives or are true positives that are challenging to assay. Indeed, two enantioselective hits binding to DDB1 and WDR91 were confirmed by SpS but not by SPR (Supplementary Table 4).

In-depth characterization of enantioselective binding to four targets. Four targets – DDB1, WDR91, WDR55, and HAT1 – predicted to have 'medium' or 'low' ligandability scores, as measured by their DLID scores (Fig. 2a and Supplementary Fig. 3), were selected for further analysis.

Damage-specific DNA-binding protein 1 (DDB1) is a multidomain protein involved in protein homeostasis[26]. We screened histidine-tagged DDB1 and other proteins in parallel against the racemate library and identified a compound (XS381952) specifically enriched in the methanol eluate from DDB1 beads (Fig. 3a). When XS381952 in the DDB1-bead eluate was analyzed by chiral chromatography, one of the two enantiomers was clearly enriched compared with the ratio

of the enantiomers in the starting library, providing compelling evidence of specific binding (Fig. 3b). To rigorously characterize the binding, we resolved and characterized the two enantiomers of XS381952 using preparative chiral chromatography and electronic circular dichroism (ECD) spectroscopy (Supplementary Fig. 10), and measured their binding to DDB1 using SPR. We found that (*S*)-XS381952 bound DDB1 with a K_D of 2 μM (Supplementary Fig. 5), significantly more potent than its (*R*)-counterpart (estimated K_D > 73 μM). To elucidate the mechanism of binding, we determined the crystal structure of (*S*)-XS381952 with DDB1 (Supplementary Fig. 11a and Supplementary Table 5) and found that (*S*)-XS381952 is sandwiched between two aromatic residues, W1047 and Y1114, via π-stacking interactions (Fig. 3c). The enantioselectivity was explained by the positioning of the ethoxyphenyl moiety linked to the chiral carbon, which adopted a nearly 90° angle from the compound backbone, orienting toward V1132 and stabilizing against the 1129-1140 α-helix of DDB1. The (*R*)-enantiomer would not be able to fit into the binding site due to steric hindrance.

WD40 repeat containing protein 91 (WDR91) is a 747-residue protein with a C-terminal WDR domain that plays a critical role in endosomal maturation[27]. After screening the WDR domain of WDR91 with the racemate library, four compounds were enriched in WDR91-bead eluates compared with the other proteins screened in parallel (Supplementary Fig. 6). Two of these hits, XS838489 (Fig. 3d and 3e) and XS837729, bound in an enantioselective manner. The remaining two hits, XS381295 and XS381186, did not show enantioselectivity in its binding. The binding of all four hits was characterized by SPR, and three hits displayed single digit micromolar K_D values (Supplementary Fig. 6), while the fourth hit, XS381186 had a K_D value of 29 μM (Supplementary Fig. 6). To elucidate the binding mode of racemic XS838489, we co-crystallized it with the WDR domain of WDR91 (residues 392-747), revealing the (*R*)-enantiomer bound to WDR91 (Fig. 3f,

ARTICLE IN PRESS

Supplementary Fig. 11b, and Supplementary Table 5). This binding pocket lies between two β-propellers and is surrounded by hydrophobic residues, including L465, L467, L477, and A459. The amide nitrogen linked to the chiral carbon forms a hydrogen bond with the backbone oxygen of T547, favoring the (*R*)-enantiomer for binding. The co-crystal structure of WDR91 with XS381295 showed it binding to the same side pocket also in an enantioselective manner (Supplementary Fig. 11c and 12), even though this was not obviously apparent by E-ASMS. In the case of XS381295, its chlorophenyl ring fits into a hydrophobic pocket formed by the aliphatic side chains of L477, L465, L467, A459, as well as T547 and M550 (Supplementary Fig. 12). We suspect that the inability to detect enantioselective binding for XS381295 by chiral chromatography may be due to its rapid racemization. To evaluate the chiral stability of the 8,217 screened compounds under E-ASMS assay conditions, we performed hydrogen-deuterium (H/D) exchange analysis (Supplementary text). Deuterium incorporation was detected in 533 compounds, indicating their potential for spontaneous racemization (Supplementary Table 6). Among the 118 hits identified by E-ASMS, only XS381295 showed deuterium incorporation, which explains its lack of enantioselective enrichment in E-ASMS due to spontaneous racemization.

WD40 repeat containing protein 55 (WDR55) is a 383-residue WDR protein that adopts a seven-bladed β-propeller fold and functions as a nucleolar protein involved in ribosomal biogenesis[28]. Although WDR55 is predicted to be a challenging target according to its DLID score, racemic compound XS381774 was enriched (Fig. 3g and 3h), with a subsequently confirmed apparent K_D of 11 μM, as measured by SPR, and displayed enantioselectivity (Supplementary Fig. 7). The enantioselective binding of XS381774 was of interest, because the enantiomers differed by the orientation of a single methyl group. The mechanism of enantioselective binding was revealed by co-crystallization of racemic XS381774 with WDR55 (Supplementary Fig. 11d and Table 5). The

(*S*)-enantiomer of XS381774 preferentially co-crystallized in a hydrophobic side pocket of WDR55, and binding was mediated through hydrophobic interactions with the aliphatic side chains of residues I100, I108, V110, V122, L141, L135, and W153 (Fig. 3i). A methyl group which forms the chiral center points toward W153 and I108, enhancing hydrophobic interactions and that explained the enantioselective enrichment. To quantify the degree of selectivity, we separated and purified the individual enantiomers and evaluated their binding affinities to WDR55 using SPR. The (*S*)-enantiomer of XS381774 exhibited at least 5-fold higher binding affinity compared to the (*R*)-enantiomer, with K_D values of 5 μM and > 25 μM, respectively (Supplementary Fig. 7).

Histone Acetyltransferase 1 (HAT1) is an enzyme that acetylates histone lysine residues, a modification associated with a transcriptionally active chromatin state[29]. Despite two decades of efforts in both academia and industry, there are few, if any, small molecule ligands for HAT1. In our E-ASMS screen, XS380871 was selectively enriched by HAT1 (Fig. 3j). Interestingly, XS380871 contains two chiral centers, and two of its four stereoisomers, were selectively enriched by HAT1 (Fig. 3k). The binding of XS380871 to HAT1, was confirmed by SPR (K_D = 12 μM) as well as ^{19}F NMR (Supplementary Fig. 8). To determine the molecular basis of binding, we incubated and co-crystallized the racemic mixture of XS380871 with HAT1 (Supplementary Fig. 11e and Supplementary Table 5). The co-structure revealed that the compound binds at the acetyl-Coenzyme A (acetyl-CoA) binding site (Fig. 3l). The chiral carbon (*S*) linked to the nitrophenyl moiety exhibited an enantioselective binding mode, with the nitrophenol substituent forming hydrogen bonds with the backbone nitrogen of G253 (via its nitro group) and the backbone nitrogen of G249 (via its hydroxyl group). There would however be enough space elsewhere in the structure to accommodate the (*R*)-enantiomer, albeit with potentially reduced affinity. In contrast, the fluorophenyl moiety at the second chiral carbon bound to a hydrophobic subpocket

ARTICLE IN PRESS

lined by residues A275, V238, M241, and Y282; here the other enantiomer would be incompatible with binding due to severe steric clashes with the region around M241. This observation is consistent with the enantioselectivity data obtained from the E-ASMS analysis.

ARTICLE IN PRESS

ARTICLE IN PRESS

Discussion

Despite the extensive efforts over the past decades[30], approximately 80% of all human proteins still lack characterized chemical ligands[31-33]. The slow pace of chemical ligand discovery is largely due to the intrinsic limitations of current hit identification strategies, including resource-intensive hit verification processes. To address these challenges, we introduce the E-ASMS approach as a scalable platform for chemical ligand discovery.

E-ASMS offers several key advantages for screening at scale. First, the ability of the method to capture weak-affinity binders ensures a high success rate, even with a relatively small chemical library (~8,000 compounds). Second, the use of enantioselective binding as an orthogonal readout of binding specificity, potentially obviates the need for labour-intensive biophysical confirmation at higher compound concentrations. Thus, E-ASMS has the potential to decouple the confirmation step from traditional orthogonal assay development, enabling rapid hit triage at scale. Third, E-ASMS may also provide semi-quantitative insights. As shown in Supplementary Fig. 13, strong correlations were observed between the recovered MS signals of confirmed hits and their K_D values, suggesting the potential to estimate binding affinity directly from MS readouts. This raises the possibility for generating high-quality datasets to support artificial intelligence (AI)-driven drug discovery. As outlined in our recent roadmap paper, the E-ASMS platform will be employed on ~2,000 proteins by 2030,[34] with all data shared in our established AIRCHECK database (https://aircheck.ai/).

Not all E-ASMS candidate binders were confirmed by biophysical methods, such as SPR. Our limited data indicates this discrepancy may be due to fundamental differences in the underlying biophysical principles of each method. SPR requires the use of compound concentrations above the K_D, which in turn necessitates sufficient aqueous solubility, whereas E-ASMS is insensitive to

ARTICLE IN PRESS

chemical concentrations. This discrepancy is most apparent for weak-affinity E-ASMS hits, which may fail validation in SPR due to poor solubility rather than lack of binding. Nonetheless, E-ASMS also carries certain limitations. The chemical library design is more challenging than conventional high-throughput assays, and only compounds with sufficient MS response and chiral separation can be included. In addition, as E-ASMS measures only binding events, substantial further investments are needed to develop hits into chemical probes with high-affinity, selectivity, and cell permeability for functional studies.

ARTICLE IN PRESS

ARTICLE IN PRESS

Methods

Reagents. Methanol (HPLC grade), ultrapure water (HPLC grade), acetonitrile (HPLC grade), DMSO, glycerol, and formic acid were purchased from Fisher Scientific (Ottawa, ON, CA). Ni-NTA magnetic agarose beads and Triton X-100 were obtained from Sigma-Aldrich (St. Louis, MO, USA). Tris base, sodium chloride (NaCl), and Tris(2-Carboxyethyl)phosphine (TCEP) were sourced from BioShop Canada Inc. (Burlington, ON, CA). Imidazole was provided by Bio Basic Canada Inc. (Markham, ON, CA). The chemical library was purchased from Enamine US Inc. (Monmouth Jct., NJ, USA) and ChemDiv (San Diego, CA, USA). Stock solutions of the chemicals were prepared in DMSO and stored at -20 ^{o}C in the dark until use.

E-ASMS protein affinity selection experiments. The E-ASMS library of 8,217 chemicals was divided into 14 chemical pools by using a robotic system, with an average of ~600 chemicals in each pool. Each of the 14 chemical pools was incubated with purified recombinant His-tagged human proteins on 96-well plates, in the binding buffer (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM TCEP, 0.5% (v/v) glycerol, 0.01% (v/v) Triton X-100, and pH 7.5). The addition of DMSO-dissolved compounds resulted in a final DMSO concentration of 0.5% (v/v) in each well. 5 µL of Ni-NTA magnetic beads was added to the binding buffer, adjusted to a final volume of 250 µL per sample. To determine the optimal protein-to-ligand concentration ratio, we evaluated several conditions during the assay development phase using eight proteins and 109 ligands (Supplementary Table 2), including 10:1 (1 µM protein, 0.1 µM ligand), 4:1 (0.4 µM protein, 0.1 µM ligand), 6.7:1 (2 µM protein, 0.3 µM ligand), and 20:1 (2 µM protein, 0.1 µM ligand). The 10:1 ratio provided the most consistent signal and hit recovery. Therefore, we adopted 1 µM protein and 0.1 µM ligand as the standard condition, balancing assay robustness with compatibility

for screening at scale. All E-ASMS screenings were performed in triplicate. Incubation was carried out on a rotary mixer at 4 ^{o}C for 30 min. After incubation, the 96-well plate was placed on a magnetic plate to separate the beads, and the incubation buffer was removed. The beads were then washed twice with 100 µL of washing buffer 1 (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM TCEP, 0.5% glycerol, 0.01% Triton X-100, 5 mM imidazole, and pH 7.5), followed by one wash with 100 µL of washing buffer 2 (150 mM NaCl, 50 mM Tris-HCl, 5 mM imidazole, and pH 7.5). The beads and washing buffer 2 were transferred to a new plate. After removing the washing buffer 2, 120 µL of methanol was added to each well to denature proteins and release chemical ligands. The methanol extracts separated from magnetic beads were directly subjected to LC-MS analysis. In parallel, the proteins were also subjected to SDS-PAGE to confirm the presence of target proteins on the beads.

Identification of putative hits using LC-MS. The E-ASMS extracts were analyzed using an Orbitrap Exploris 240 mass spectrometer equipped with a Vanquish UHPLC system (Thermo Fisher Scientific, CA, USA). Chromatographic separation was conducted on an Accucore Vanquish C_{18} reverse-phase column (50 mm × 2.1 mm × 3 µm) at a flow rate of 0.3 mL/min. The injection volume was 1 µL. Ultrapure water with 0.1% formic acid (A) and methanol with 0.1% formic acid (B) were used as the two mobile phases. Gradient elution started with 5% B, increased to 80% at 1.5 min, then to 100% at 3.0 min. The composition was held at 100% for 1.4 min, rapidly returned to 5% B over 0.1 min, and was maintained at 5% for column re-equilibration until 6.0 min. The column temperature was maintained at 40 ^{o}C, and the sample compartment at 7 ^{o}C. Data acquisition was performed in full MS^1 scan mode (150-520 *m/z*) with a resolution of R = 60,000, in both positive and negative ionization modes. The detailed instrumental parameters of LC-MS

system are provided in Supplementary Table 7. Each tentative hit was further confirmed by injecting the compounds to LC-MS individually to avoid potential artifacts of chemical degradation or rearrangement.

Chiral analysis of putative hits. (1) Building the chiral separation database. The accurate prediction of suitable chiral separation conditions for resolving the enantiomers of a given compound remains a challenge[35,36]. To address this, we decided to establish the chiral separation database for the E-ASMS chemical library, under eight conditions (four columns × two LC gradient methods, see the details below). Each of the 14 chemical pools were individually injected under eight conditions to build the database. (2) Eight chiral separation conditions. Four columns from DAICEL Chemical Industries, LTD. were employed: CHIRALPAK IA (250 × 4.6 mm), CHIRALPAK IBN (250 × 2.1 mm), CHIRALPAK IC (250 × 2.1 mm), and CHIRALPAK IG (250 × 2.1 mm). These columns were selected because they have been demonstrated to achieve sufficient separations for the biggest number of chiral compounds[35].

Two LC gradient methods were employed for each column. In **method 1**, acetonitrile with 0.1% formic acid (A) and methanol with 0.1% formic acid (B) were used as mobile phases. Gradient elution started at 40% B, held for 1 min, then increased to 80% at 3 min, reached 100% at 8 min, held for 1 min, followed by a return to 40% over 2 min, and held for another 1 min. The flow rate was set at 0.3 mL/min for columns IBN, IC, IG, and at 1.0 mL/min for column IA. In **method 2**, H_2O with 0.1% formic acid (A) and methanol with 0.1% formic acid (B) were used as mobile phases. Gradient elution began with 40% B, increased to 80% at 6 min, reached 100% at 18 min, held for 6 min, and then return to 40% over 6 min. The flow rate was 0.2 mL/min for columns

ARTICLE IN PRESS

IBN, IC, IG, and 1.0 mL/min for column IA. The column temperature was maintained at 40 °C, and the sample compartment was kept at 7 °C. The injection volume was 1 µL.

(3) Chiral detection of putative hits under optimal conditions. Once putative hits were detected from initial E-ASMS screening, the optimal separation condition for the hit compounds was selected by searching against the chiral separation database established in step (2) above. The same E-ASMS extracts were then analyzed under the optimal conditions. To exclude potential interferences, enantiomers were further confirmed by matching MS^2 spectra.

Automatic data processing. A total of 1,302 E-ASMS screening samples (31 proteins × 3 replicates × 14 pools = 1,302) were completed in the current study, leading to 145 GB of raw mass spectrometry data. We have established an automatic data processing workflow to detect E-ASMS hits.

(1) Building the MS database for the E-ASMS library. To build the database, each of the 14 chemical pools was subjected to LC-MS analysis. Raw mass spectrometry files were converted to the mzXML format. Peak features from each chemical pool were identified using the 'XCMS' R package with a mass tolerance of 2.5 ppm[29]. Only the peak features with peak an intensity > 10^5 and abundances at least 10 times higher than methanol were considered as library compounds and retained for subsequent analysis. Isotopic peaks and adducts were excluded by matching chromatographic peaks and theoretical mass difference. Detected peak features were then matched to the E-ASMS library with a mass tolerance of 3 ppm. Potential inferences and misassigned compounds were further excluded via manual inspections. Then, a MS database of 8,217 E-ASMS library compounds was established, with *m/z*, retention time and SMILES information recorded.

ARTICLE IN PRESS

(2) Detecting compounds from E-ASMS features by matching to the MS database. To detect library compounds from the E-ASMS features, we matched each of the 8,217 E-ASMS compounds against each E-ASMS sample. A mass tolerance of 3 ppm, and a retention time matching window of 0.5 min were used. A final data matrix with 8,217 rows (corresponding to library compounds) and 24 columns (corresponding to E-ASMS samples from each batch) was created.

(3) Detecting putative E-ASMS hits. To detect putative E-ASMS hits, we calculated the enrichment fold and *p* values of each library compound against each protein by using seven other proteins from the same batch as the control (equation 1).

$$\text{Enrichment Fold} = \frac{A_{POI} \times N}{\sum A_i} \quad (1)$$

A_{POI} represents the peak intensity of the putative hit compound enriched by the protein of interest; N represents the number of background proteins from the same batch (N = 7); A_i represents the peak intensity of the same hit compound enriched by the i^{th} protein from the same batch. This normalization process avoids the impacts from varied instrumental signals of compounds.

Eventually, only the compounds with enrichment fold > 5, and *p* value < 0.05 were considered as putative hit compounds. Each putative hit was further manually inspected by matching to the chemical library, to exclude potential interferences misassigned by the algorithms.

Surface plasmon resonance. SPR hit confirmation and validations were performed using a Biacore 8K instrument at 20 °C. The biotinylated respective protein constructs were immobilized on the active flow cells of a streptavidin-coated SA sensor chip after initial conditioning of both reference and active flow cells with 50 mM NaOH for 3 × 60 s with a flow rate of 10 μL/min. Each protein solution (30-150 μg/mL) was injected through the respective active flow cells for 60-660 s with a flow rate of 5 μL/min to obtain a protein immobilization level estimated to yield ~30 RU

of signal at 100% compound binding. Equilibration was performed after protein immobilization by flowing running buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.001% Tween 20, 0.2% PEG3350, 0.5 mM TCEP, and 3% DMSO) over the flow cells with a flow rate of 50 µL/min until a stable baseline was observed. Initial start-up cycles, blank cycles and wash (50% DMSO wash to flush the needles) steps were included in the SPR compound analysis methods. The compound injections were performed over the reference and active flow cells using multi-cycle kinetics at a flow rate of 40 µL/min with a 55 s association time and a 120 s dissociation time for dose-response titration. The stock compound concentration series was performed in 100% DMSO with 2-3-fold serial dilution and prepared the samples using the running buffer by maintaining final DMSO concentration of 3% (v/v) across the tested concentration range. Initially, the compounds were titrated between 0-45 µM compound concentrations and later followed up between 0-15 µM, 0-30 µM, 0-45 µM, and 0-90 µM compound concentrations based on the affinity and solubility of the compounds. Solvent correction cycles were also included across each run to adjust high bulk responses from the solvent. Double referencing of the data was introduced by subtraction of the reference flow cell and the respective zero compound concentration cycles. Affinity fitting was performed by applying a 1:1 equilibrium binding model to the data using Biacore Intelligent Analysis tool provided with the Biacore Insight Evaluation software.

ARTICLE IN PRESS

Spectral shift (SpS) measurement. The spectral shift method requires target labeling with a fluorophore using a variety of coupling chemistries, *e.g.*, lysine reactive NHS labeling or cysteine reactive maleimide labeling, or site-specific affinity labeling technology, *e.g.*, His-tagged or biotinylated target labeling. Here, the protein labeling was performed through site-specific affinity labeling using the biotinylated target labeling kit (Cat # NT-L120; NanoTemper Technologies

GmbH). Initially, the biotinylated target proteins were diluted in the assay buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Tween-20, 0.5 mM TECEP) and set up a 16-point serial dilution with a top concentration of 500 nM protein in 10 µL volume for each point in a 384-well plate. An equal volume of the 4 nM labeling dye dissolved in the assay buffer was added to the protein wells in the serial dilution. The protein and the dye mix were incubated at room temperature in the dark for 15 minutes. The affinity between the biotinylated target proteins and the labeling dye was then measured by spectral shift method. The final protein concentration in the compound screening assays was determined based on the observed K_D between the target protein and the labeling dye. For compound screening and affinity measurement, the compounds were dissolved in the assay buffer at a 60 µM (30 µM final) compound concentration (for single-dose) or at a varying compound concentration (for dose-response titration) by keeping the 6% DMSO (final 3% after mixing with protein mix) concentration in the compound mix. The protein mix was prepared by adding 40 nM protein (final 20 nM) and 4 nM (final 2 nM) dye in the assay buffer. The protein mix was then incubated for 15 min at room temperature for target labeling. 10 µL of each of the compound mix and the protein mix was mixed into the assay wells in a 384-well assay plate. The assay mixture was then incubated for 30 minutes at room temperature before measuring the emission intensities using the Dianthus instrument equipped with dual-emission (650 nm and 670 nm) detection optics (NanoTemper Technologies GmbH). The affinity (K_D) of the compounds was calculated by plotting the ratios of the fluorescence intensities at two fixed wavelengths as a function of the compound concentration and fitting the data to a 1:1 binding model using DI.Screening Analysis v2.2 software.

ARTICLE IN PRESS

ARTICLE IN PRESS

Chiral Stability Assessment via Hydrogen-Deuterium (H/D) Exchange. To evaluate the chemical stability of stereogenic centers under assay conditions, we performed a global H/D exchange screen using high-resolution MS^1 data. All 8,217 chiral chemicals were incubated in the binding buffer (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM TCEP, 0.5% (v/v) glycerol, 0.01% (v/v) Triton X-100, and pH 7.5) with D_2O at 4 ^oC for 8 h to identify compounds exhibiting deuterium incorporation. Deuterium incorporation was determined by detecting a +1.0063 Da shift in the monoisotopic mass with mass accuracy within 3 ppm and retention time alignment within 1 second across replicates. To ensure a conservative evaluation, we considered all compounds showing +1.0063 Da shifts as potential H/D exchange events, although we acknowledge that exchange may also occur at non-stereogenic positions.

^{19}F NMR studies. The binding of XS380871 to HAT1 was confirmed using ^{19}F NMR by looking for the broadening and/or shifting of its ^{19}F resonance upon the addition of HAT1. Spectra of the compound were acquired on a Bruker Avance III spectrometer operating at 600 MHz, equipped with a QCI probe at 293K, and collected at 20 µM, with and without the presence of HAT1 (~34 µM, buffered in 50 mM Tris 7.0, 200 mM NaCl, 4 mM DTT). TFA (100 µM) was added as an internal reference. 1k transients were acquired over a sweep width of 150 ppm; an exponential window function (LB = 5 Hz) was applied prior to Fourier transformation.

Proteins used for E-ASMS and SPR. Protein expression and purification summaries are detailed in Supplementary Table 8.

ARTICLE IN PRESS

Crystallization and structure determination. Crystallization of E-ASMS hits with their respective targets are summarized in Supplementary Table 9.

Quality Assurance and Quality Control (QA/QC) of E-ASMS. To assure the high-quality data for scalable E-ASMS screening, we conducted stringent QA/QC from both protein and chemical perspectives.

Protein QA/QC. To assure the selected proteins were properly folded and compatible with our E-ASMS workflow, we used SDS-PAGE to semi-quantitatively assess the recovery of each protein through our E-ASMS procedure. In brief, SDS-PAGE analysis was conducted for each protein before and after E-ASMS. The band intensity of each protein was compared, and those proteins with low recoveries were excluded for E-ASMS analysis. The SDS-PAGE results for all proteins are provided in Supplementary Fig. 14.

Chemical QA/QC. During LC-MS analysis, the E-ASMS chiral library compounds were injected every batch to ensure the retention time, *m/z*, and intensity are not shifting for database matching. All the chiral compounds were detected by the LC-MS system. A QC standard of small subset (~600 chemicals) was injected after every 12 samples to monitor instrument stability and LC (chiral) separations. The instrument response deviation for analytes in the QC standard remained below 10% throughout the analysis. A solvent blank was injected after every 12 samples to monitor and prevent carryover.

ARTICLE IN PRESS

Data Availability

Atomic coordinates and structure factors for all crystal structures have been deposited in the Protein Data bank under the accession codes: 9EJO [https://doi.org/10.2210/pdb9EJO/pdb], 9EJP [https://doi.org/10.2210/pdb9EJP/pdb], 9EJQ [https://doi.org/10.2210/pdb9EJQ/pdb], 9EKP [https://doi.org/10.2210/pdb9EKP/pdb], and 9MJG [https://doi.org/10.2210/pdb9MJG/pdb]. All raw data generated in this work have been deposited in the NIH Common Fund's National Metabolomics Data Repository (NMDR) under Project ID NMDR: PR002550. The dataset is publicly available and can be accessed via the project DOI [http://dx.doi.org/10.21228/M81K02]. Source Data are provided with this paper.

Code Availability

The source code and demo datasets can be found via GitHub at [https://github.com/huiUofT/EAS-MS].

ARTICLE IN PRESS

ARTICLE IN PRESS

References

1 Moffat, J. G., Rudolph, J. & Bailey, D. Phenotypic screening in cancer drug discovery - past, present and future. *Nature reviews. Drug discovery* **13**, 588-602, doi:10.1038/nrd4366 (2014).

2 Eder, J., Sedrani, R. & Wiesmann, C. The discovery of first-in-class drugs: origins and evolution. *Nature reviews. Drug discovery* **13**, 577-587, doi:10.1038/nrd4336 (2014).

3 Macarron, R. *et al.* Impact of high-throughput screening in biomedical research. *Nature reviews. Drug discovery* **10**, 188-195, doi:10.1038/nrd3368 (2011).

4 Parker, C. G. *et al.* Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **168**, 527-541 e529, doi:10.1016/j.cell.2016.12.029 (2017).

5 Abbasov, M. E. *et al.* A proteome-wide atlas of lysine-reactive chemistry. *Nature chemistry* **13**, 1081-1092, doi:10.1038/s41557-021-00765-4 (2021).

6 Offensperger, F. *et al.* Large-scale chemoproteomics expedites ligand discovery and predicts ligand behavior in cells. *Science* **384**, eadk5864, doi:10.1126/science.adk5864 (2024).

7 Peterson, A. A. & Liu, D. R. Small-molecule discovery through DNA-encoded libraries. *Nature reviews. Drug discovery* **22**, 699-722, doi:10.1038/s41573-023-00713-6 (2023).

8 Prudent, R., Annis, D. A., Dandliker, P. J., Ortholand, J. Y. & Roche, D. Exploring new targets and chemical space with affinity selection-mass spectrometry. *Nature reviews. Chemistry* **5**, 62-71, doi:10.1038/s41570-020-00229-2 (2021).

9 Wang, Y. *et al.* Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nature chemistry* **11**, 1113-1123, doi:10.1038/s41557-019-0351-5 (2019).

10 Favalli, N. *et al.* Stereo- and regiodefined DNA-encoded chemical libraries enable efficient tumour-targeting applications. *Nature chemistry* **13**, 540-548, doi:10.1038/s41557-021-00660-y (2021).

11 Tao, Y. *et al.* Targeted Protein Degradation by Electrophilic PROTACs that Stereoselectively and Site-Specifically Engage DCAF1. *Journal of the American Chemical Society* **144**, 18688-18699, doi:10.1021/jacs.2c08964 (2022).

12 Ogasawara, D. *et al.* Chemical tools to expand the ligandable proteome: diversity-oriented synthesis-based photoreactive stereoprobes. *bioRxiv : the preprint server for biology*, doi:10.1101/2024.02.27.582206 (2024).

13 Yang, D. *et al.* Widespread formation of toxic nitrated bisphenols indoors by heterogeneous reactions with HONO. *Science advances* **8**, eabq7023, doi:10.1126/sciadv.abq7023 (2022).

14 Liu, J. *et al.* Diindoles produced from commensal microbiota metabolites function as endogenous CAR/Nr1i3 ligands. *Nat Commun* **15**, 2563, doi:10.1038/s41467-024-46559-3 (2024).

15 Liu, J. *et al.* The omega-3 hydroxy fatty acid 7(S)-HDHA is a high-affinity PPARα ligand that regulates brain neuronal morphology. *Science signaling* **15**, eabo1857, doi:10.1126/scisignal.abo1857 (2022).

16 Ito, T. *et al.* Identification of a primary target of thalidomide teratogenicity. *Science* **327**, 1345-1350, doi:10.1126/science.1177319 (2010).

17 Mori, T. *et al.* Structural basis of thalidomide enantiomer binding to cereblon. *Sci Rep* **8**, 1294, doi:10.1038/s41598-018-19202-7 (2018).

18 Chamberlain, P. P. *et al.* Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nature structural & molecular biology* **21**, 803-809, doi:10.1038/nsmb.2874 (2014).

19 Li, A. S. M. *et al.* Discovery of Nanomolar DCAF1 Small Molecule Ligands. *J Med Chem* **66**, 5041-5060, doi:10.1021/acs.jmedchem.2c02132 (2023).

20 Bonday, Z. Q. *et al.* LLY-283, a Potent and Selective Inhibitor of Arginine Methyltransferase 5, PRMT5, with Antitumor Activity. *ACS medicinal chemistry letters* **9**, 612-617, doi:10.1021/acsmedchemlett.8b00014 (2018).

21 Göricke, F. *et al.* Discovery and Characterization of BAY-805, a Potent and Selective Inhibitor of Ubiquitin-Specific Protease USP21. *J Med Chem* **66**, 3431-3447, doi:10.1021/acs.jmedchem.2c01933 (2023).

22 Sheridan, R. P., Maiorov, V. N., Holloway, M. K., Cornell, W. D. & Gao, Y. D. Drug-like density: a method of quantifying the "bindability" of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank. *J Chem Inf Model* **50**, 2029-2040, doi:10.1021/ci100312t (2010).

23 Ahmad, S. *et al.* Discovery of a First-in-Class Small-Molecule Ligand for WDR91 Using DNA-Encoded Chemical Library Selection Followed by Machine Learning. *J Med Chem* **66**, 16051-16061, doi:10.1021/acs.jmedchem.3c01471 (2023).

24 Ackloo, S. *et al.* A Target Class Ligandability Evaluation of WD40 Repeat-Containing Proteins. *J Med Chem*, doi:10.1021/acs.jmedchem.4c02010 (2024).

25 Helmerhorst, E., Chandler, D. J., Nussio, M. & Mamotte, C. D. Real-time and Label-free Bio-sensing of Molecular Interactions by Surface Plasmon Resonance: A Laboratory Medicine Perspective. *The Clinical biochemist. Reviews* **33**, 161-173 (2012).

26 Jackson, S. & Xiong, Y. CRL4s: the CUL4-RING E3 ubiquitin ligases. *Trends in biochemical sciences* **34**, 562-570, doi:10.1016/j.tibs.2009.07.002 (2009).

27 Casanova, J. E. & Winckler, B. A new Rab7 effector controls phosphoinositide conversion in endosome maturation. *The Journal of cell biology* **216**, 2995-2997, doi:10.1083/jcb.201709034 (2017).

28 Iwanami, N. *et al.* WDR55 is a nucleolar modulator of ribosomal RNA synthesis, cell cycle progression, and teleost organ development. *PLoS genetics* **4**, e1000171, doi:10.1371/journal.pgen.1000171 (2008).

29 Nagarajan, P. *et al.* Histone acetyl transferase 1 is essential for mammalian development, genome stability, and the processing of newly synthesized histones H3 and H4. *PLoS genetics* **9**, e1003518, doi:10.1371/journal.pgen.1003518 (2013).

30 Bunnage, M. E., Chekler, E. L. & Jones, L. H. Target validation using chemical probes. *Nat Chem Biol* **9**, 195-199, doi:10.1038/nchembio.1197 (2013).

31 Edwards, A. M. *et al.* Too many roads not taken. **470**, 165 (2011).

32 Müller, S. *et al.* Target 2035 - update on the quest for a probe for every protein. *RSC medicinal chemistry* **13**, 13-21, doi:10.1039/d1md00228g (2022).

33 Oprea, T. I. *et al.* Unexplored therapeutic opportunities in the human genome. *Nature reviews. Drug discovery* **17**, 317-332, doi:10.1038/nrd.2018.14 (2018).

34 Edwards, A. M. *et al.* Protein–ligand data at scale to support machine learning. *Nature Reviews Chemistry*, In press, doi:https://doi.org/10.1038/s41570-025-00737-z (2025).

35 Hong, Y., Welch, C. J., Piras, P. & Tang, H. Enhanced Structure-Based Prediction of Chiral Stationary Phases for Chromatographic Enantioseparation from 3D Molecular Conformations. *Anal Chem* **96**, 2351-2359, doi:10.1021/acs.analchem.3c04028 (2024).

ARTICLE IN PRESS

ARTICLE IN PRESS

36 Xu, H., Lin, J., Zhang, D. & Mo, F. Retention time prediction for chromatographic enantioseparation by quantile geometry-enhanced graph neural network. *Nat Commun* **14**, 3095, doi:10.1038/s41467-023-38853-3 (2023).

ARTICLE IN PRESS

ARTICLE IN PRESS

Acknowledgements

We would like to thank the advice and discussions on method development from Dafydd Owen (Pfizer), Judith Guenther (Bayer), Scott Johnson (Bristol Myers Squib), Oliver Kraemer (Boehringer Ingelheim), Julian Schmid (Merck), Juncai Meng (Janssen), Lawrence Szewczuk (Janssen), Xidong Feng (Pfizer), Wenyi Hua (Pfizer), Anja Giese (Bayer), Matteo Aldeghi (Bayer), Nidhi Arora (Takeda), James Kiefer (Genentech), Ingo Hartung (Merck), and Matthew Troutman (Pfizer). The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUbOPEN grant 875510], Janssen, Pfizer, and Takeda. Research was supported by a Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant, Ontario Early Researcher Award, and the National Institute of Health (NIH) (U54AG065187). The authors acknowledge the support of instrumentation grants from the Canada Foundation for Innovation, the Ontario Research Fund, and an NSERC Research Tools and Instrument Grant. ACJ, SM, and SK are also grateful for support by the German Cancer Aid grant TACTIC. DM was supported by the Onassis Foundation [Scholarship ID: F ZU 049-1/ 2024-2025]. TSB acknowledges startup funding provided by the Emory University School of Medicine and the Winship Cancer Institute. AME holds the Temerty Nexus Chair of Health Technology and Innovation at the University of Toronto.

Author Contributions

XW, JS, DY, AB, SA, PJB, TM, VS, LH, DRO, HP, AME, CHA: E-ASMS method development

JS, XW: E-ASMS screening

ARTICLE IN PRESS

HZ, PG, EG, MK, SRG, LKR, DM, MSi, CL, RBH, BDB: Protein production and QC

ShA, MK, FL, UHC, SRG, DM, CL: SPR, DSF confirmation

SH: F-NMR characterization

TM, IK, AC: Library design

HZ, UHC, AD: X-ray crystallography

MSc, CVS: Druggability assessment

SA, VS, PJB: Project management

WWY, RH, TSB, SM, ACJ, SK, LH, HP, AME, CHA: Supervision, review and editing

XW, JS, LH, HP: Writing, reviewing and editing with input from all

These authors contributed equally: XW, JS, ShA

These authors jointly supervised this work: HP, LH

Competing Interests

The authors declare no competing interests.

ARTICLE IN PRESS

ARTICLE IN PRESS

Figure Legends/Captions

Fig. 1 | Development of the E-ASMS platform for scalable protein ligandability discovery. (**a**) Schematic representation of the E-ASMS screening workflow. (**b-e**) Benchmarking the E-ASMS platform using four positive control ligands binding to CRBN, DCAF1, PRMT5, and USP21, respectively. Data are presented as mean values ± SD (n = 3). (**f**) Chemical space of the E-ASMS library. A two-dimensional uniform manifold approximation and projection (UMAP) visualization of SMILES descriptor. The displayed structures represent the top 10 scaffolds within the E-ASMS chemical library. (**g**) Molecular properties of the E-ASMS chemical library: molecular weight (MW), logP, hydrogen bond acceptors (HBA) and donors (HBD). **(h)** Chiral separation of the E-ASMS library using four chiral columns (IA, IB, IC, and IG), and two mobile phases (acetonitrile/methanol, yellow; water/methanol, green). **(i)** Detection rate of 109 known ligands against eight proteins by SEC (green) and E-ASMS (yellow) methods, respectively. Purple dots represent the K_D distribution of chemical ligands discovered through E-ASMS screening against 31 proteins.

Fig. 2 | Scalable chemical ligand profiling across 31 diverse proteins. (**a**) Overview of the 31 screened proteins. 'High': drug-like density (DLID) > 1, 'Medium': 0.5 < DLID < 1, 'Low': DLID < 0.5. (**b**) Scatter plot of all E-ASMS library compounds interacting with 31 proteins. Dots above the horizontal threshold represent hits with enrichment fold > 5. Red dots indicate hits with enrichment fold > 5 and $P < 0.05$. Statistical significance was assessed using a two-tailed unpaired t-test (n = 3), and no multiple-comparison adjustment was applied. (**c**) Confidence levels of all identified hits according to orthogonal evidence. (**d**) UMAP visualization of the chemical space for hits with enrichment fold > 5. (**e**) Correlation between K_D and enrichment fold. Data are

ARTICLE IN PRESS

presented as mean values ± SEM (n = 3). (**f**) SPR validation rate of hits with or without enantioselectivity, across different enrichment fold ranges.

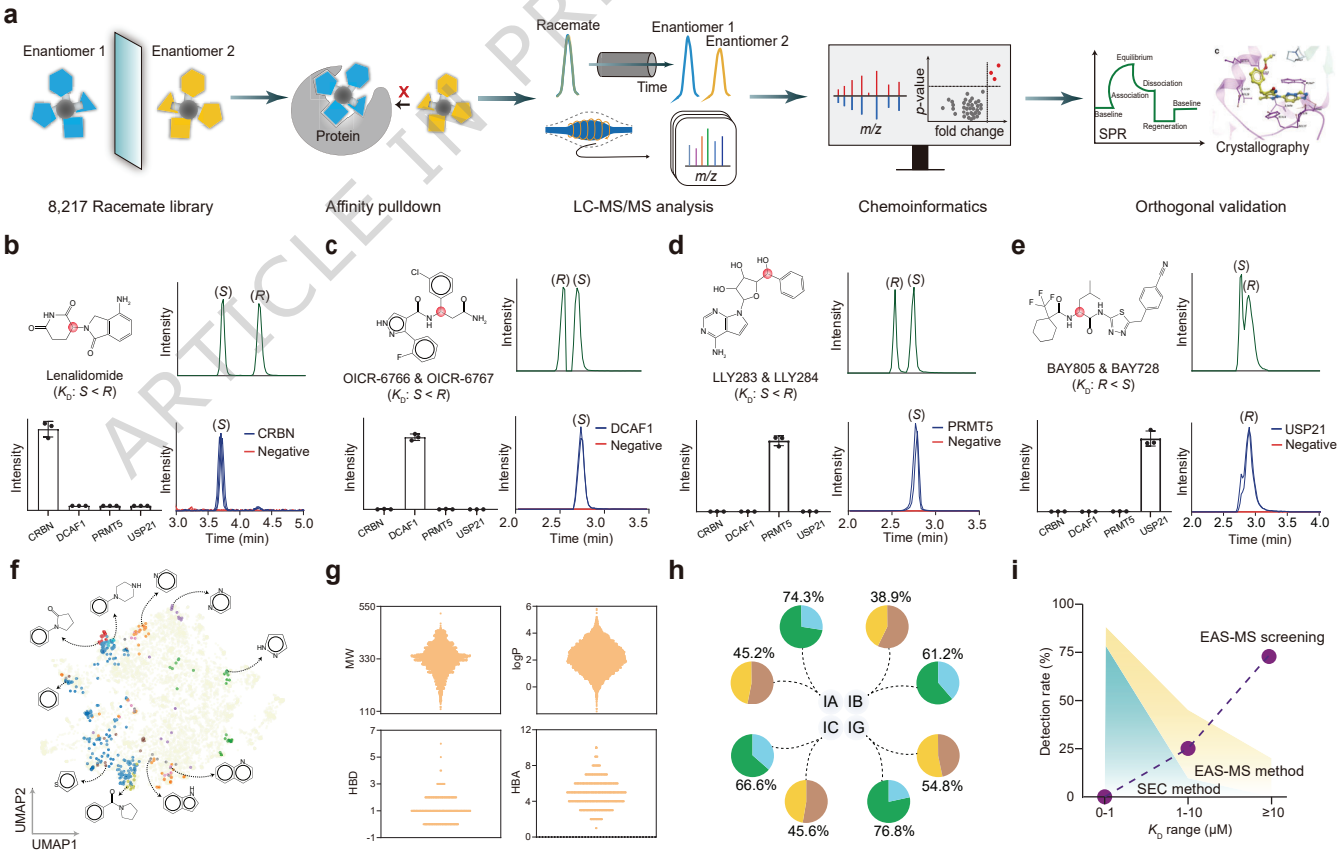
Fig. 3 | Identification of enantioselective hits for DDB1, WDR91, WDR55, and HAT1. (**a**) Scatter plot showing XS381952 pulled down by DDB1. (**b**) Chiral chromatograms of the XS381952 enantiomers before (upper) and after E-ASMS (bottom). (**c**) Co-crystal structure of DDB1 with XS381952. The hydrogen bond between the protein and the ligand is shown as a green dashed line. (**d**) Scatter plot showing XS838489 pulled down by WDR91. (**e**) Chiral chromatograms of the XS838489 enantiomers. (**f**) Co-crystal structure of WDR91 with XS838489. Hydrogen bonds between the protein and the ligand are shown as green dashed lines. (**g**) Scatter plot showing XS381774 pulled down by WDR55. (**h**) Chiral chromatograms of the XS381774 enantiomers. (**i**) Co-crystal structure of WDR55 with XS381774. (**j**) Scatter plot showing XS380871 pulled down by HAT1. (**k**) Chiral chromatograms of the XS380871 enantiomers. (**l**) Co-crystal structure of HAT1 with XS380871. Hydrogen bonds between the protein and the ligand are shown as green dashed lines. Different colors in chiral chromatograms represent three replicates.

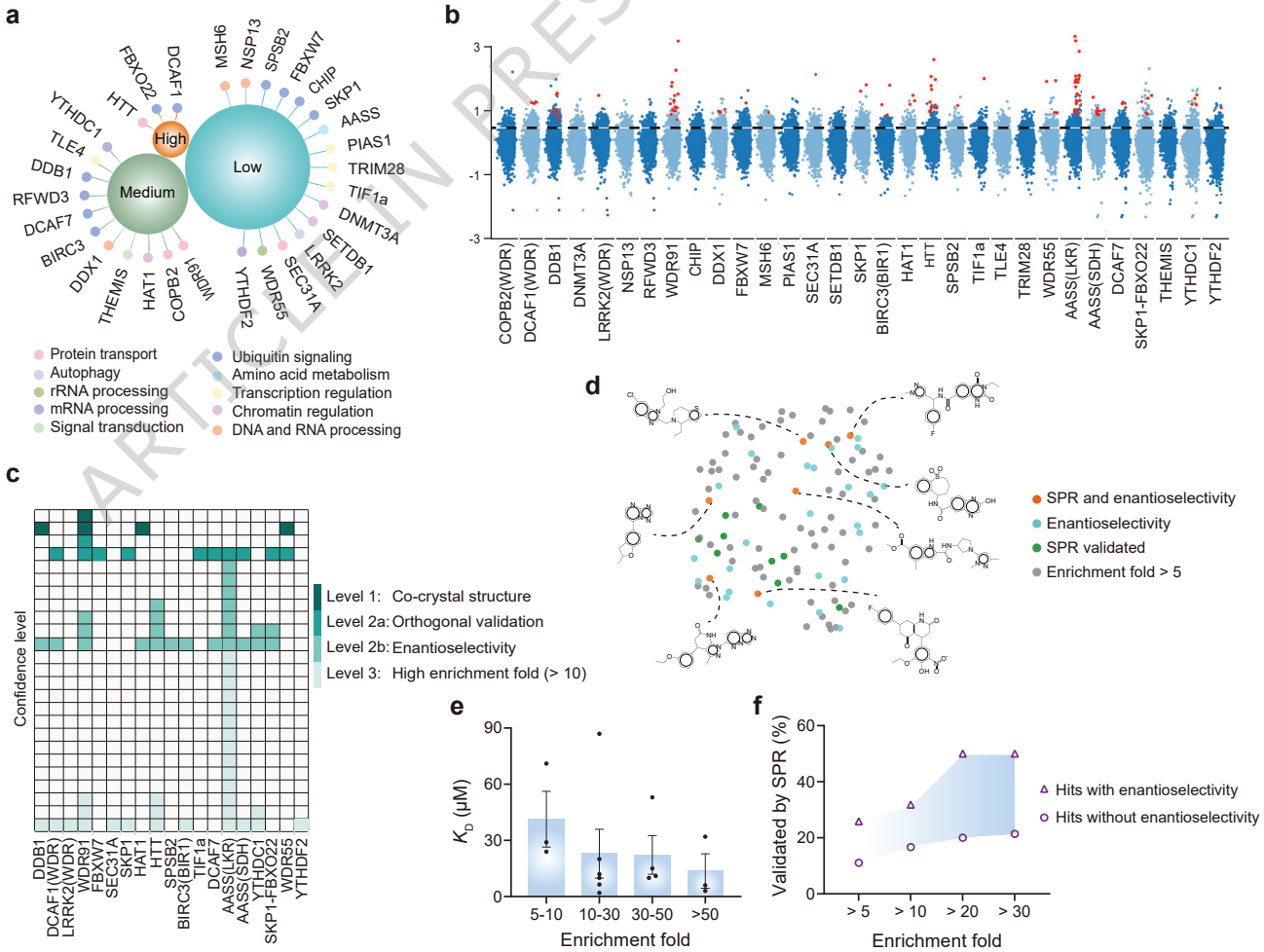
ARTICLE IN PRESS

Editor's Summary
High-throughput chemical ligand discovery is challenged by false positives. Here, authors introduce a scalable enantioselective affinity-selection mass spectrometry approach for proteome-wide ligand discovery with high sensitivity and selectivity

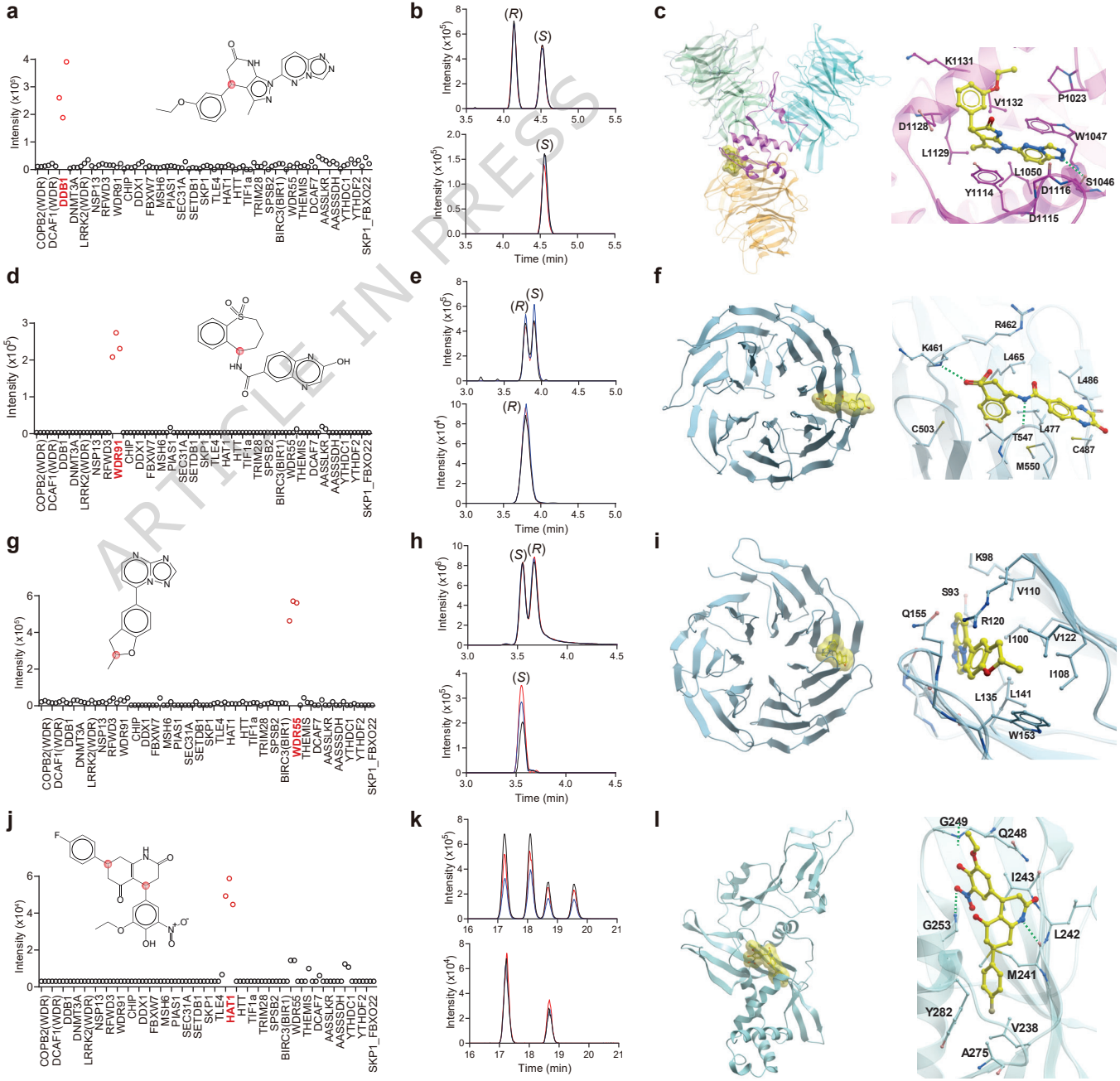
Peer review information: *Nature Communications* thanks Khaled Barakat, Susumu Uchiyama, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

ARTICLE IN PRESS





ARTICLE IN PRESS



ARTICLE IN PRESS

ARTICLE IN PRESS

High-throughput chemical ligand discovery is challenged by false positives. Here, authors introduce a scalable enantioselective affinity-selection mass spectrometry approach for proteome-wide ligand discovery with high sensitivity and selectivity

ARTICLE IN PRESS