



ProxyBind: A compendium of binding sites for proximity-induced pharmacology

Evianne Rovers^a, Lihua Liu^a, Matthieu Schapira^{a,b,*}

^aStructural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada

^bDepartment of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada



ARTICLE INFO

Article history:

Received 23 September 2022

Received in revised form 4 November 2022

Accepted 4 November 2022

Available online 8 November 2022

Keywords:

PROTAC

Proteome

Proximity pharmacology

Kinases

Phosphatases

Post-translational modification

ABSTRACT

Proximity-induced pharmacology (ProxPharm) is a novel paradigm in drug discovery where a small molecule brings two proteins in close proximity to elicit a signal, generally from one protein onto another. The potential of ProxPharm compounds as a new therapeutic modality is firmly established by proteolysis targeting chimeras (PROTACs) that bring an E3 ubiquitin ligase in proximity to a target protein to induce ubiquitination and subsequent degradation of the target. The concept can be expanded to induce other post-translational modifications via the recruitment of different types of protein-modifying enzymes. To survey the human proteome for opportunities in proximity pharmacology, we systematically mapped non-catalytic drug binding pockets on the structure of protein-modifying enzymes available from the Protein Databank. In addition to binding sites exploited by previously reported ProxPharm compounds, we identified putative ligandable non-catalytic pockets in 236 kinases, 45 phosphatases, 37 deubiquitinases, 14 methyltransferases, 11 acetyltransferases, 13 glycosyltransferases, 4 deacetylases, 7 demethylases and 2 glycosidases, including cavities occupied by chemical matter that may serve as starting points for future ProxPharm compounds. This systematic survey confirms that proximity pharmacology is a versatile modality with largely unexplored and promising potential and reveals novel opportunities to pharmacologically rewire molecular circuitries. All data is available from the ProxyBind database at <https://polymorph.sgc.utoronto.ca/proxybind/index.php>.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Proteolysis targeting chimeras (PROTACs) are bifunctional small molecules that simultaneously bind an E3 ubiquitin ligase and a target protein, thereby inducing the ubiquitination and subsequent proteasomal degradation of the protein target [1]. This type of molecules has evolved over the past 20 years from a chemical biology curiosity to a promising therapeutic modality, with clear dose-dependent degradation of therapeutic targets such as AR, IRAK4 or BTK observed in man (clinicaltrials.gov identifiers NCT03888612, NCT04772885, NCT04830137), and the question is no longer whether but when the first PROTAC will be approved for therapeutic use by regulatory agencies. Proximity-induced Pharmacology

(ProxPharm) is an extension of targeted protein degradation, where chemically induced proximity with proteins beyond E3 ligases can be used to rewire the molecular circuitry of cells for chemical biology applications or therapeutic benefit [2,3]. Indeed, ProxPharm compounds were recently reported that recruit a phosphatase, two kinases, an acetyltransferase, and a deubiquitinase to post-translationally modify neosubstrates [4–7].

Structural studies have shown that PROTACs are not simply acting as chemical linkers but rather stabilize non-natural protein-protein interactions between E3 ligases and target proteins [8]. Because compatible protein interfaces do not always exist between two proteins, a prevailing notion is that a collection of chemical handles binding a diverse array of E3 ligases will be necessary to productively induce the degradation of any given protein. Additionally, the tissue expression profile and subcellular localization of the E3 ligase must match that of the target protein for a PROTAC to be active. Finally, PROTACs recruiting E3 ligases with disease-specific tissue expression profiles can avoid adverse effects associated with the indiscriminate inhibition of the protein target. For

Abbreviations: ProxPharm, Proximity Pharmacology; PROTAC, Proteolysis Targeting Chimera.

* Corresponding author at: Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada.

E-mail address: matthieu.schapira@utoronto.ca (M. Schapira).

<https://doi.org/10.1016/j.csbj.2022.11.010>

2001-0370/© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

example, a senolytic PROTAC exploits the restricted expression profile of the E3 ligase CRBN to avoid toxicity associated with the adverse inhibition of the target protein, Bcl-xl, in platelets [9]. Similar rules are expected to apply to ProxPharm compounds beyond PROTACs, emphasizing the need to identify chemical handles for a diverse array of protein-modifying enzymes.

To uncover novel opportunities for the development of future ProxPharm compounds, we searched for non-catalytic ligandable pockets (structural cavities that can be occupied by small-molecule ligands) in all experimental structures of human protein-modifying enzymes, including kinases, phosphatases, acetyltransferases, deacetylases, methyltransferases, demethylases, glycosyltransferases, glycosidases and deubiquitinases. These pockets need to be non-catalytic to preserve the catalytic activity of the protein-modifying enzyme which is necessary for the ProxPharm-induced response. The ligandability of E3 ligases was previously reviewed and not considered in this analysis which is focused on opportunities for proximity pharmacology beyond PROTACs [1,10–14]. We identified non-catalytic pockets in 369 human enzymes, including those recruited by previously reported ProxPharm compounds. This analysis further confirms the rich potential of proximity pharmacology for chemical biology applications.

2. Methods

2.1. Mapping binding pockets

A list of enzymes was compiled from the ExPasy ENZYME database and the UniprotKB database [15] and mapped to corresponding PDB codes. The 3D structures were extracted from the PDB and the biologically relevant oligomeric state was generated with ICM (Molsoft, San Diego). The *icmPocketfinder* module was run against each converted ICM object using default settings. The pockets were categorized as non-catalytic based on the following two approaches.

2.2. Interpro domain analysis

The domain architecture of each enzyme was extracted from the InterPro database [16]. The domains were marked either as catalytic or non-catalytic based on GO ontology or literature. Residues within 2.8 Å of the pocket mesh generated by ICM were considered as lining the pocket, and the N- and C-terminal boundaries of this selection were used to define a ‘pseudo’ sequence for the pocket. These sequences were aligned and compared with the domain architecture of the enzyme to determine the domain location of the pocket. If the pocket was in a manually curated non-catalytic domain, the pocket was marked non-catalytic.

2.3. Catalytic residues proximity analysis

For each enzyme, the corresponding catalytic residue information was extracted from either the Mechanism and Catalytic Site Atlas database [17] or UniprotKB database [15]. If the catalytic residues were present in the structure, the distance between the pockets and the catalytic residues were measured. If the pocket was more than 7 Å away from the catalytic residues, it was categorized as non-catalytic.

2.4. Additional filters

Nucleotide binding residues and co-factor binding residues information was extracted from the UniprotKB database to determine which pockets corresponded to nucleotide or co-factor bind-

ing sites. For example, the ATP binding site in protein kinases or the acetyl-CoA binding site in acetyltransferases. If the distance between the pocket and nucleotide/co-factor binding residues was <7 Å, the pocket was filtered out. If the pocket was in proximity (<5 Å) of unresolved residues in the structure due to poor electron density, the pocket was not included for further analysis. If the catalytic residues were among the missing residues, pockets were excluded as well. Pockets were also excluded when located at the interface of inhibitor proteins and enzyme complexes. Next, pockets were filtered for duplicates: when two structures representing the same enzyme had a similar pocket, the largest pocket was retained. Pockets predicted unligandable were also removed: ligandability was determined using the pocket properties generated by ICM (volume: 155.7–661.1 Å³, area: 155–655 Å², hydrophobicity: >0.44, buriedness: 0.6–0.95, DLID [18]: >-1). Cutoff values were based on properties of experimentally proven druggable pockets. Lastly, the pockets were grouped based on their domains. A list of manually curated non-catalytic domains was formed, from which non-catalytic domains necessary for the catalytic activity were excluded.

2.5. Cysteine reactivity

The predicted reactivity of cysteine sidechains lining pockets was predicted using the *ReactiveCys* module of ICM. The method is based on reactivity data for 34 reactive and 184 non-reactive cysteines from isoTOP-ABPP (isotopic tandem orthogonal proteolysis activity-based protein profiling) [19] and a nonredundant set of PDB protein structures (resolution < 2.5 Å) with covalently-modified cysteines (272 reactive).

3. Results

To assemble a database of druggable binding pockets that may be exploited by ProxPharm compounds, all high-resolution structures of human protein-modifying enzymes beyond E3 ligases in the PDB were analyzed with the cavity mapping tool *IcmPocketFinder* (Molsoft, San Diego). Only structural cavities with properties (volume, area, hydrophobicity, buriedness and drug-like density (DLID)) within a pre-defined range (detailed in the Methods section) were deemed ligandable and were considered further. A permissive definition of ligandability was used to reflect the fact that chemical handles for ProxPharm compounds do not need to bind with high potency to their target. Indeed, ligands with up to 10 μM affinity have been successfully used to make PROTACs [20]. When a ligandable cavity was found in a non-catalytic domain, the domain was also deemed ligandable in the context of enzymes not in the PDB, but with a low confidence score. When enzymes were bound to other proteins in the PDB, cavities were also searched at the protein interface. Pockets that may be exploited by ProxPharm compounds could be divided into three categories: 1) those located in non-catalytic domains, 2) those found at non-catalytic sites of the catalytic domain, 3) those mapping at the interface of protein complexes (Fig. 1).

Potentially ligandable non-catalytic pockets were found in 236 kinases, 45 phosphatases, 37 deubiquitinases, as well as several writers and erasers of methyl, acetyl and glycosyl groups (Fig. 1, Table S1–3). In the following section, we review in detail each protein family.

3.1. Protein kinases

Ligandable non-catalytic pockets were found in the catalytic domain of 170 kinases (Fig. 1, Table S2). For instance, in 86 kinases, a pocket is found in the α-lobe of the kinase domain (Fig. 2A,

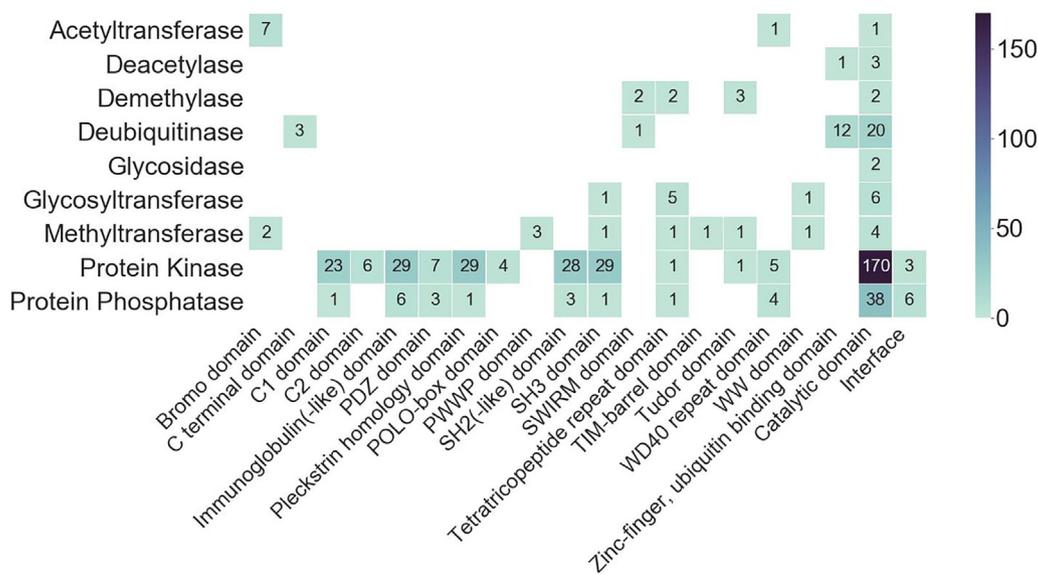


Fig. 1. Distribution of non-catalytic pockets in human protein-modifying enzymes. The number of proteins with a putative ligandable non-catalytic pocket is shown for each structural domain and each protein family. For each enzyme family, the total number of proteins in human can be found in Table S1.

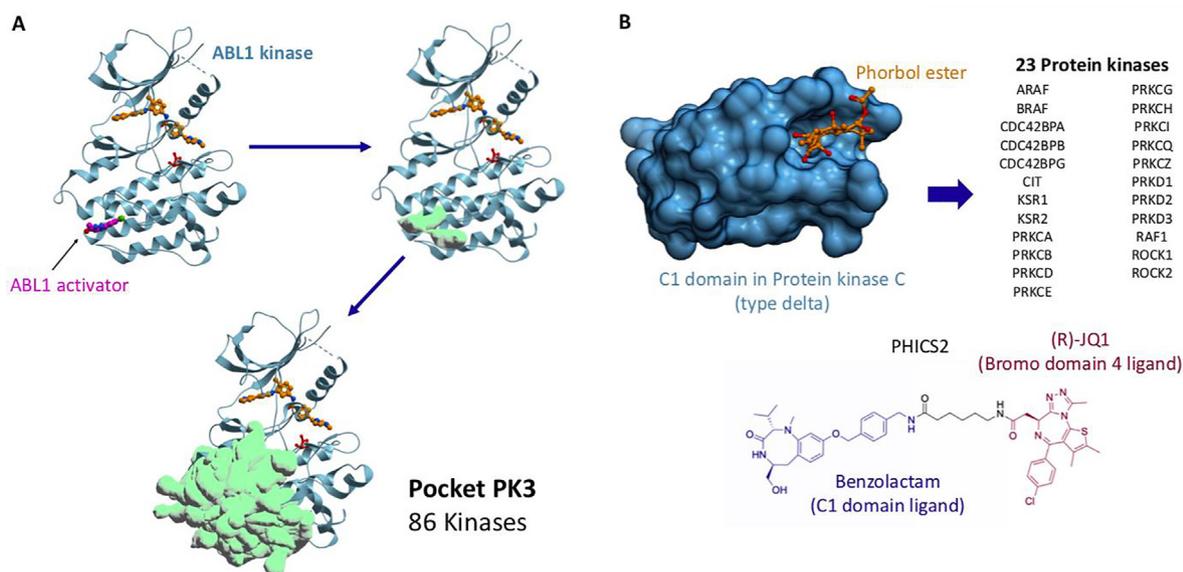


Fig. 2. Recurrent non-catalytic pockets found in kinases. A) Pockets found in the kinase domain. ABL1 (blue) bound to catalytic inhibitor (orange) is used as a canonical reference structure (PDB: 6NPU [21]). An allosteric activator is shown in purple. A similar pocket is found in 86 protein kinases. B) Phorbol ester bound to the C1 domain of KPCD (PDB: 1PTR [28]). A C1 domain can be found in 23 protein kinases. Benzolactam binds to C1 domains and is used as chemical handle to recruit protein kinase C (type delta) to Bromodomain-containing protein 4 (PHICS2) [4]. The full list of kinases and pockets summarized here is provided in Table S2, S3 and in the database ProxyBind.

Pocket PK3) and, in the context of Abelson kinase, is exploited by an activating compound located over 15 Å away from the imatinib-occupied active site (Figure S1, PDB 6NPU, Pocket PK3) [21]. Other pockets are recurrently found at five other locations and could potentially be exploited to pharmacologically hijack kinases (Figure S1). In particular, 47 kinases share a cavity below the sub-activation loop (Figure S1, Pocket PK4) which is occupied by a small molecule in the MAP kinase p38 α [22] (PDB 3HVC). A β -lobe cavity is found in another 25 kinases (Figure S1, Pocket PK5), where, in PDK1, a cysteine is covalently engaged by fragment inhibitors or activators (PDB 3ORZ) [23] and a different β -lobe pocket is identified in 9 kinases (Figure S1, Pocket PK2) and occupied by a fragment molecule in the context of CDK2 (PDB 6Q4D) [24].

Ligandable pockets were also found at the interface of the catalytic domain of 3 kinases (PRKAA1, PRKAA2, CDK5) and cofactor proteins (Figure S1). For example, a pharmacological activator is sandwiched at the interface of the β -lobe of PRKAA1 and its cofactor PRKAB1 (Fig. 2B, Pocket PK1) [25]. Interestingly, this chemical scaffold was recently linked to an inhibitor of Bruton's tyrosine kinase (BTK), leading to the phosphorylation of BTK by PRKAA1 in cells, in what was the first example of a phosphorylation-inducing chimeric small molecule (PHIC) [4]. Another pocket is found at the interface of CDK5 and CDK5R1, a neural-specific CDK5 activator protein, raising the possibility to develop brain-specific phosphorylating agents that would exploit this site (Figure S1).

Multiple potentially ligandable cavities were also identified in non-catalytic domains of kinases (Table S2, Figure S2). For example, a cavity was found in the non-catalytic C1 domain of 23 kinases such as BRAF, CDC42 binding kinases, or PKC kinases (Fig. 2). Binding of diacylglycerol to this pocket leads to translocation from the cytosol to the membrane of PKC kinases, and catalytic activation [26]. The cavity was successfully targeted by drug-like molecules such as V8-benzolactams [27], which can be used as PKC-recruiting handles in heterobifunctional PHICS. Using this strategy, Siriwardena et al. could induce the phosphorylation of BRD4 by PKC [4]. A similar strategy may be applied to the other kinases where we also identified a C1-domain cavity.

A membrane-targeting C2 domain is also present in 6 protein kinases, including PKC kinases, but the ligandability of its phosphatidylserine binding pocket is unclear. A tyrosine-lined pocket conserved in the POLO domain of PLK kinases participates in substrate recognition and was targeted by weak compounds that would need to be optimized to serve as ProxPharm handles (Figure S2) [29]. Five kinases contain a WD-40 repeat (WDR), which is a β -propeller domain with a druggable central cavity [30]. For instance, the WDR domain of LRRK2 could be exploited by future PHICS to phosphorylate targets in the brain, where it is expressed.

Other protein domains of potential interest were identified in human kinases, but even though cavities meeting our selection criteria were found, the general ligandability of these domains remains to be supported experimentally. For instance, 29 kinases contain an immunoglobulin-like domain (Figure 1 and S2). Small molecule ligands were shown to bind to the immunoglobulin-like domain of the unrelated protein RAGE, but ligands were prohibitively weak [31]. Another 28 kinases contain both SH2 and SH3 domains (Figure 1 and S2), known to participate in the formation of an auto-inhibitory state and contribute to substrate recruitment of Src family kinases. Despite sustained efforts, potent, drug-like, cell-penetrant ligands remain to be found for these domains. Nevertheless, they may be sufficiently ligandable for the discovery of weak compounds that may serve as valid chemical handles for kinase-recruiting ProxPharm molecules. In another example, the poorly characterized kinase STK31 includes a Tudor domain (Fig. 1, Table S2), generally found in proteins involved in chromatin-mediated signaling. This domain was targeted by a potent chemical probe in the context of the methyltransferase SETDB1 [32] and may be ligandable in STK31.

3.2. Protein phosphatases

Non-catalytic pockets were found in 51 protein phosphatases (Table S3). Among these, 38 protein phosphatases have a non-catalytic pocket in the catalytic domain and 20 enzymes in juxtaposed domains (Fig. 1, Table S2). Some of the non-catalytic cavities were recurrently found in the phosphatase domain: 13 tyrosine-protein phosphatases share a cavity 15 Å from the catalytic site (Fig. 3, Pocket PP3), which, in the context of PTPN5, is occupied by an allosteric activator (PDB 6H8R) [33]. The recurrent pocket locations are labelled as pocket Protein Phosphatase # (Pocket PP#). Other recurrent cavities are found at four other locations of the catalytic domain and could potentially be exploited to recruit tyrosine-protein phosphatases to target proteins (Fig. S3A). Furthermore, 5 serine/threonine-protein phosphatases have 4 recurrent non-catalytic cavities in their catalytic domain (Fig. S3B).

Non-catalytic pockets were also found at multiple protein–protein interfaces (locations denoted as Pocket Protein Phosphatases Interface # (Pocket PPI#)), including a cavity located at the interface of the three subunits of the protein phosphatase 2A (PP2A) heterotrimer, and occupied by a small molecule activator [34] (Fig. 3, Pocket PPI1, Figure S4). Heterobifunctional compounds derived from this activator could potentially be used for targeted

dephosphorylation. This hypothesis is further supported by the fact that another phosphatase, PP2A, was successfully recruited to dephosphorylate the kinases AKT or EGFR by linking kinase inhibitors to peptidic ligands that exploit the tetratricopeptide repeat domain in PP2A [5].

Cavities are also found in the PDZ domain of protein phosphatases PTPN3, PTPN4 and PTPN14 (Figure S5). The ligandability of these pockets is not experimentally validated, but they are occupied by the C-terminal leucine or valine of pentameric peptide ligands [35,36], and a similar pocket in the PDZ domain of the unrelated protein PICK1 was crystallized in complex with a small molecule binding with sub-micromolar potency [37]. Finally, pockets with unclear ligandability were found in the SH2 domain of phosphatases PTPN6, PTPN11 and TNS2, and the tetratricopeptide repeat of PPP5C (Figure S5).

3.3. Protein methyltransferases

Protein methyltransferases (PMTs) are typically large multi-modular proteins where chromatin-binding modules are often found juxtaposed to the catalytic domain. For instance, a PWWP domain is found in the NSD subfamily of PMTs (NSD1, NSD2 and NSD3) and chemical probes were reported for the PWWP domain of NSD2 and NSD3 (Fig. 4) [38,39]. The NSD3 ligand was recently used as the chemical handle of an NSD3-degrading PROTAC [40]. These ligands - which do not inhibit the enzymatic activity - could also potentially serve as chemical moieties to recruit NSD2 or NSD3 for the methylation of new protein substrates.

SETDB1, another multi-modular PMT, includes a non-catalytic Tudor domain selectively targeted by a potent chemical probe that may be linked to other ligands to methylate non-natural protein substrates (Fig. 4) [32]. Interestingly, recurrent genetic aberrations drive the overexpression of NSD2 in multiple myeloma and pediatric leukemia, and of NSD2, NSD3 and SETDB1 in lung cancer [41–44], which could possibly offer an opportunity for targeted protein methylation in cells presenting a specific disease-associated genetic profile. Putative ligandable cavities were found in a few other non-catalytic domains of PMTs, including the bromodomain of KMT2A and ASH1L (bromodomains are typically druggable (Fig. 1) [45], but no ligand was reported for these domains.

A recurrent pocket was also found in the catalytic domain of two protein arginine methyltransferases, PRMT3 and PRMT8, which is located more than 17 Å away from the catalytic site (Figure S6, Pocket M1). Other unique non-catalytic pockets were found in the methyltransferase catalytic domain of 3 PMTs (PRMT3, SETD7, CARM1) (Table S3). These cavities met our ligandability criteria but so far, their chemical tractability was not validated experimentally.

3.4. Lysine demethylases

A number of non-catalytic domains of lysine demethylases include potentially ligandable pockets. KDM4A, KDM4B and KDM4C all have a Tudor domain, which was shown to be chemically tractable in the context of SETDB1. The Tudor domain of KDM4A was crystallized in complex with a low-affinity chemical fragment ($K_D \sim 80 \mu\text{M}$) that may be optimized into a stronger-binding chemical handle towards the development of a demethylase-recruiting bifunctional molecules (Fig. 4) [46]. Putative ligandable pockets were also found in the tetratricopeptide repeat of KDM6A and UTY and the SWIRM domain of KDM1A and KDM1B (Fig. 1, Table S2), but no ligand was so far reported for these domains.

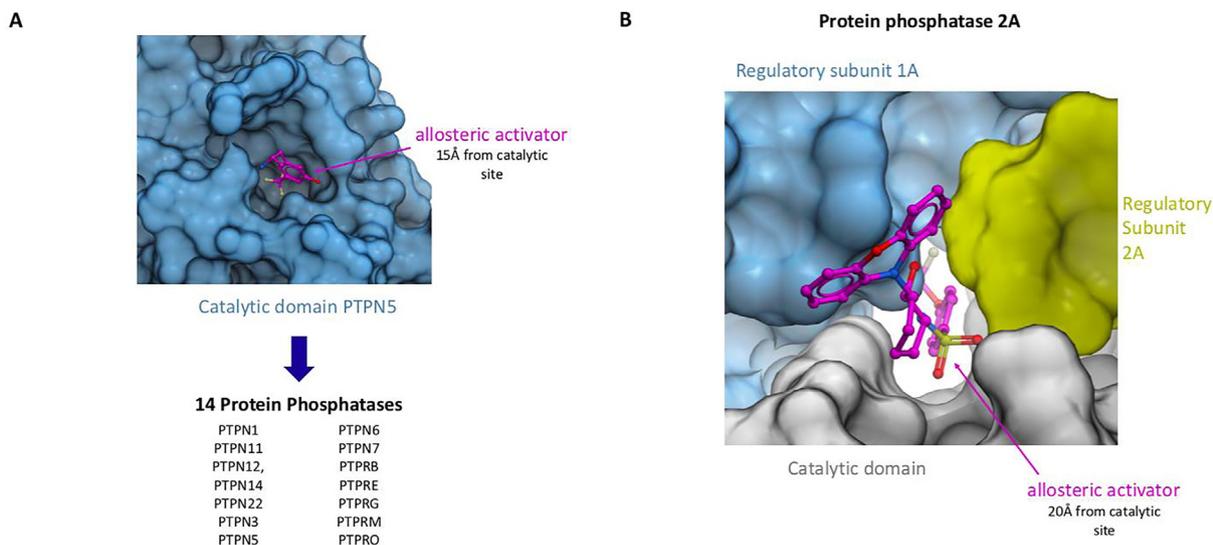


Fig. 3. Recurrent non-catalytic pockets mapping at protein phosphatase domains. A) Pocket in the catalytic domain of tyrosine-protein phosphatases 5 bound to an allosteric activator (purple) 15 Å from the catalytic site (PDB 6H8R). A similar pocket (labelled PP3 in Table S3 and the Proxybind database) is recurrent in 14 tyrosine-protein phosphatases. B) Pocket found at the interface of the protein phosphatase 2A catalytic domain (gray) and interacting subunits (blue and yellow) with allosteric activator (PDB: 6NTS) [34].

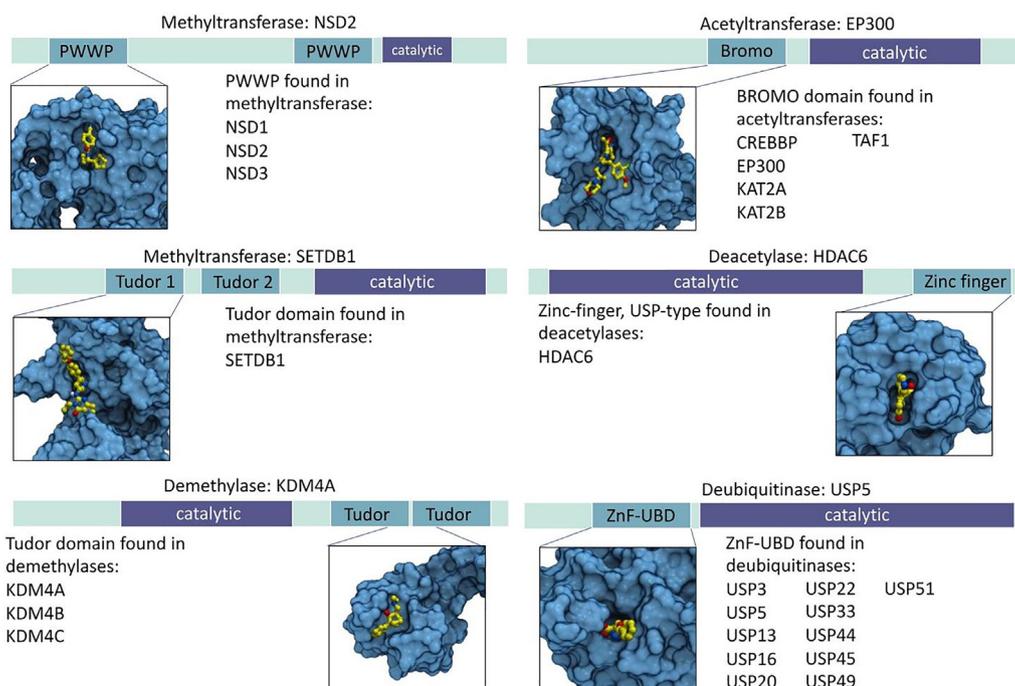


Fig. 4. Examples of pockets in non-catalytic domains of methyltransferases, acetyltransferases, demethylases, deacetylases and deubiquitinases. PWWP domain in NSD2 with small-molecule ligand (PDB: 6UE6 [55]), bromodomain in EP300 with small-molecule ligand (PDB: 5BT3), Tudor domain in SETDB1 with small-molecule ligand (PDB: 7CJT [32]), zinc-finger, USP-type in HDAC6 with small-molecule ligand (PDB: 5KH7 [56]), Tudor domain in KDM4A with chemical fragment (PDB: 5VAR [46]), ZnF-UBD in USP5 with small-molecule ligand (PDB: 6DXT [51]).

3.5. Lysine acetyltransferases

With over 3000 acetylated lysine sidechains across 1700 human proteins, acetylation is a ubiquitous post-translational modification involved in a diverse array of cellular machineries such as the regulation of gene expression, splicing or cell cycle [47,48]. Out of 35 lysine acetyltransferases in the human genome, we found non-catalytic ligandable pockets in 9 (Fig. 1, Table S3). Several acetyltransferases include an acetyl-lysine binding bromodomain, five of which were crystallized in complex with multiple small-

molecule ligands (EP300, CREBBP, KAT2A, KAT2B and TAF1) (Fig. 4) [45]. A compound targeting the bromodomain of one of these, EP300, was chemically linked to an FKBP12-binding molecule to successfully induce the acetylation of FKBP12-fusion proteins by EP300, thereby confirming that acetyltransferases are amenable to proximity pharmacology, and strongly suggesting that bromodomain ligands could be used as chemical handles to recruit other acetyltransferases to neosubstrates [7].

A WDR domain is also found in GTF3C4, a poorly characterized acetyltransferase (Fig. 1, Table S2). The structure of this domain

was not experimentally solved, but WDR domains are ligandable in the context of other proteins [30,49] and this enzyme could potentially be harnessed for targeted acetylation.

3.6. Lysine deacetylases

Deacetylases have a limited number of non-catalytic domains and a ligandable site was found in only one of them: the zinc-finger ubiquitin-binding domain (Znf-UBD) of HDAC6 (Fig. 4). This binding pocket recognizes the C-terminal extremity of ubiquitin and was successfully targeted by small molecule ligands [50] representing excellent chemical handles for proximity pharmacology applications. Non-catalytic pockets were also found in the catalytic domain of three other deacetylases: HDAC4, HDAC8 and HDAC1, but the ligandability of these sites remains to be experimentally validated (Figure S7).

3.7. Deubiquitinases

Deubiquitinases (DUBs) typically remove ubiquitin tags deposited by E3 ligases. When these tags are signalling for proteasomal degradation, DUBs deubiquitinate and rescue their protein substrates from the ubiquitin–proteasome system and have a stabilizing effect on their target. Chemical handles binding non-catalytic pockets of DUBs may therefore enable the recruitment of DUBs for targeted protein stabilization. As a proof-of-concept, a bifunctional molecule linking a ligand that covalently engages the DUB OTUB1 to a chemical moiety that binds Δ F508-CFTR in cystic fibrosis could stabilize Δ F508-CFTR in an OTUB1-dependent manner [6]. There is no structural information on the N-terminal domain of OTUB1 that is covalently recruited by this chimeric compound, but structures of other non-catalytic domains in DUBs reveal other opportunities for targeted protein stabilization.

The most recurrent ligandable non-catalytic domain of DUBs is the Znf-UBD, found in 12 ubiquitin-specific proteases (USPs, a class of DUBs) (Figure S7, Table S2). Low micromolar ligands were reported for the Znf-UBD of USP5, but these compounds were shown to inhibit the catalytic activity of USP5 and therefore cannot be used as chemical handles to productively recruit USP5 to neo-substrates [51]. However, the function of the Znf-UBD of DUBs is poorly understood in other USPs, and ligands targeting this domain may still be valid handles for targeted protein stabilization in the context of other DUBs.

Ligandable pockets were also found in a tandem ubiquitin-like domain located at the C-terminus of four DUBs: USP7, 11, 15 and 25 (Fig. 1, Table S2). In the context of USP7, this domain binds and activates the catalytic domain [52]. In the absence of structure of full-length USP7 in its activated form, it is unclear whether ligands occupying this C-terminal binding pocket would preserve the activation mechanism of USP7 and could be used to productively recruit USP7 for targeted protein stabilization. Another non-catalytic domain present in deubiquitinases is a SWIRM domain in MYM1. Chemical ligands have not yet been reported for this domain. Non-catalytic pockets were recurring at six locations of eight USPs within the peptidase C19-type catalytic domain (Figure S8A, Table S3). Another non-catalytic cavity is observed in the peptidase C12-type catalytic domain of UCHL1 and UCHL5 (Figure S8B, Table S3). As above, the ligandability of these pockets needs to be confirmed experimentally.

3.8. Glycosyltransferases

Glycosylation is a post-translational modification that is most common in excreted and extracellular membrane-associated proteins and is frequently dysregulated in diseases, such as cancer or bacterial infection [53]. Proof of principle for proximity-

induced glycosylation of target proteins was established by fusing substrate-targeting nanobodies to the glycosyltransferase O-GlcNAc transferase (OGT), which effectively induced the glycosylation of the desired protein targets [54]. Putative ligandable pockets in the tetratricopeptide repeat of OGT and TMTC1-4 may be exploited to chemically recruit these glycosyltransferases to neo-substrates. Similarly, the SH3 domain of FUT8 and WW domain of GALNT9 may be considered for the chemical recruitment of these enzymes. Non-catalytic cavities in the glycosyltransferase domain of ST8SIA3, B3GAT1-3, and POFUT2 were also found but, as above, their ligandability should be confirmed experimentally.

3.9. Glycosidases

Similar to glycosyltransferases, protein constructs have been developed using O-GlcNAcase or sialidase connected to nanobody to artificially induce deglycosylation [57–59]. There are limited structures and domain information available for glycosidases, but ligandable pockets are found in the catalytic domain of OGA and MAN1B1 that could be explored for deglycosylation-inducing chimeras.

3.10. Reactive cysteines

PROTACs covalently engaging an E3-ligase have demonstrated that covalent binding is a valid strategy for proximity-induced post-translational modification of target proteins [60–64]. For instance, covalent recruitment of only a small fraction of the cellular pool of the E3-ligase DCAF16 is sufficient to support targeted degradation [61]. A deubiquitinase-targeting chimera also forms a covalent bond with a cysteine of the DUB OTUB1 [6]. Electrophilic chemical handles enable the covalent recruitment of domains otherwise not considered ligandable, such as the RING domain of the E3-ligase RNF4 [60], and can be advantageous to enhance potency or selectivity. We used ICM to evaluate the reactivity of cysteine sidechains found in non-catalytic pockets of human protein-modifying enzymes (see Methods section for details).

Reactive cysteines were predicted in multiple proteins (Figure S9, Table S3). For instance, C576 is lining a pocket in the UBL domain of USP7 C-terminal to the catalytic domain, C210 is found at an ectopic site of the STK16 kinase domain, C266 at a non-catalytic site of the PP2BA phosphatase domain, and C1030 at a cavity remote from the active site of the deacetylase HDAC4 (Figure S9). It would be interesting to screen such proteins with electrophilic fragments to find covalent adducts that may serve as a starting point for novel proximity-pharmacology applications.

4. Discussion

Our systematic structural survey of the human proteome reveals numerous opportunities for the pharmacological recruitment of protein-modifying enzymes beyond E3 ligases to non-natural substrates. The predicted ligandability of a binding pocket can vary from one method to another and is not a conclusive metric. Here, we use a permissive definition based on volume, area, hydrophobicity, buriedness and DLID values. We first note that, in addition to the hundreds of new pockets identified, this approach does retrieve binding sites for known ProxPharm compounds, including a protein–protein interface pocket used to recruit the kinase PRKAA (Fig. S1B, Pocket PKI1) [4] and a bromodomain pocket used to recruit the acetyltransferase EP300 (Fig. 4) [7].

Among the collection of binding sites that we compiled, we assigned the highest confidence (confidence level 1, Table S3–4) to the ones for which a high-affinity ligand was already reported.

Table with ligandability confidence scale can be found in the [Supplementary information \(Table S4\)](#). For instance, V8-benzolactams bind the C1 domain of protein kinase C (Fig. 2B) [4,27], UNC6934 binds the PWWP domain of NSD2 (Fig. 4) [39] and compound R734 binds a protein interface of the kinase AMPK (Fig. S1B) [4,25]. A number of non-catalytic pockets were also found that are targeted by weak ligands that may be valid starting points for the development of ProxPharm compounds (confidence level 2, Table S3–4). These include compounds and peptides found in the POLO-box domain of PLK1 (Figure S2) [29] and the PDZ domain of PTPN3 (Figure S5) [35]. Because ProxPharm compounds induce the formation of ternary complexes where direct protein–protein interactions contribute to the overall energy of the system, proximity-inducing compounds can be derived from chemical handles binding with relatively weak affinity (up to 10 μ M or more) to their target (Han et al. 2019). Binding sites with this lower confidence level therefore remain of potential interest. Less reliable, but still promising are domains for which no ligand was reported in the context of the protein of interest, but that were shown to be chemically tractable in other proteins (confidence level 3, Table S3–4). For example, low nanomolar ligands targeting the WDR domains of EED and WDR5 are in pre-clinical [65–67] or clinical development (EED clinicaltrials.gov identifier NCT02900651) and WDR domains are found in the kinases LRRK1, LRRK2, MET, MST1R, PIK3R4 and the acetyltransferase GTF3C4 (Table S2). Similarly, Tudor domains are found in demethylases (KDM4A, KDM4B, KDM4C) and protein kinase STK31 (Table S2), and share a canonical aromatic cage with the Tudor domain of SETDB1 targeted by a high-affinity ligand (K_D 90 nM) [32]. Finally, sites that meet our ligandability criteria but for which no ligands were found in the protein of interest or close homologues are less reliable (confidence level 4, Table S3–4).

A limitation of our analysis is that we focused exclusively on the structures of enzymes that add or remove chemical or peptidic tags to proteins. In the future, we believe it would be interesting to expand to other enzymes, such as proteases. Pockets in non-catalytic protein subunits of enzymatic complexes may also serve as starting points for ProxPharm development. Lastly, proteins without enzymatic activity, for example transcription factors, could be analyzed for ligandable pockets that could be recruited with ProxPharm compounds. While beyond the scope of this work, such studies could be undertaken using methodologies similar to the ones presented here. We also limited our approach to proteins (and homologs) with structural information in the protein databank, but recent breakthroughs in protein structure predictions [68–70] may enable a future expansion of the analysis to the entire human proteome. Finally, the design of hetero-bifunctional molecules with favourable ADME profiles able to productively induce protein–protein interactions in a structural arrangement that allows enzymatic activity of one protein onto the other remains a challenging trial-and-error process. The necessary combinatorial synthesis of candidate molecules with varying chemical handles and linkers, and subsequent screening are non-trivial experimentally. Recent signs of progress in the rational design of PROTACs should nevertheless be noted [71–75].

In spite of these limitations, the compendium of binding sites – including some with chemical starting points – for proximity-induced pharmacology assembled here reveals a multitude of avenues to harness protein modifying enzymes involved in epigenetic mechanisms, splicing, protein homeostasis and other cellular machineries. For example, recruiting histone methyltransferases or acetyltransferases to specific genomic loci may up-regulate repressed genes, while targeted glycosylation may be a mechanism to control inflammatory and viral immune responses. Now that molecular proof-of-concept was established for a number of enzyme classes, it is the time to fully explore and test the bound-

aries of this promising modality for chemical biology and drug discovery applications.

CRediT authorship contribution statement

Evianne Rovers: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Lihua Liu:** Data curation, Visualization. **Matthieu Schapira:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was enabled in part by support provided by Compute Ontario (<https://www.computeontario.ca/>) and the Digital Research Alliance of Canada (alliancecan.ca).

Funding

MS gratefully acknowledges support from NSERC [Grant RGPIN-2019-04416], CQDM (Quantum Leap-176) and MITACS accelerate (IT13051). The Structural Genomics Consortium is a registered charity (no: 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196], EU/EPPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUOPEN grant 875510], Janssen, Merck KGaA (aka EMD in Canada and US), Pfizer and Takeda.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2022.11.010>.

References

- [1] Pettersson M, Crews CM. PROteolysis TArgeting Chimeras (PROTACs) – Past, present and future. *Drug Discov Today Technol* 2019;31:15–27. <https://doi.org/10.1016/j.ddtec.2019.01.002>.
- [2] Gerry CJ, Schreiber SL. Unifying principles of bifunctional, proximity-inducing small molecules. *Nat Chem Biol* 2020;16:369–78. <https://doi.org/10.1038/s41589-020-0469-1>.
- [3] Conway SJ. Bifunctional Molecules beyond PROTACs. *J Med Chem* 2020;63:2802–6. <https://doi.org/10.1021/acs.jmedchem.0c00293>.
- [4] Siriwardena SU, Munkanatta Godage DNP, Shoba VM, Lai S, Shi M, Wu P, et al. Phosphorylation-Inducing Chimeric Small Molecules. *J Am Chem Soc* 2020;142:14052–7. <https://doi.org/10.1021/jacs.0c05537>.
- [5] Yamazoe S, Tom J, Fu Y, Wu W, Zeng L, Sun C, et al. Heterobifunctional Molecules Induce Dephosphorylation of Kinases—A Proof of Concept Study. *J Med Chem* 2020;63:2807–13. <https://doi.org/10.1021/acs.jmedchem.9b01167>.
- [6] Henning NJ, Boike L, Spradlin JN, Ward CC, Liu G, Zhang E, et al. Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat Chem Biol* 2022;18. <https://doi.org/10.1038/s41589-022-00971-2>.
- [7] Wang WW, Chen LY, Wozniak JM, Jadhav AM, Anderson H, Malone TE, et al. Targeted Protein Acetylation in Cells Using Heterobifunctional Molecules. *J Am Chem Soc* 2021;143:16700–8. <https://doi.org/10.1021/jacs.1c07850>.
- [8] Gadd MS, Testa A, Lucas X, Chan KH, Chen W, Lamont DJ, et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat Chem Biol* 2017;13(13):514–21. <https://doi.org/10.1038/nchembio.2329>.
- [9] He Y, Zhang X, Chang J, Kim H-N, Zhang P, Wang Y, et al. Using proteolysis-targeting chimera technology to reduce navitoclax platelet toxicity and improve its senolytic activity. *Nat Commun* 2020;11:1996. <https://doi.org/10.1038/s41467-020-15838-0>.

- [10] Schneider M, Radoux CJ, Hercules A, Ochoa D, Dunham I, Zalmas LP, et al. The PROTACtable genome. *Nat Rev Drug Discov* 2021;20:789–97. <https://doi.org/10.1038/S41573-021-00245-X>.
- [11] Schapira M, Calabrese MF, Bullock AN, Crews CM. Targeted protein degradation: expanding the toolbox. *Nat Rev Drug Discov* 2019;18:949–63. <https://doi.org/10.1038/S41573-019-0047-Y>.
- [12] Nalawansha DA, Crews CM. PROTACs: An Emerging Therapeutic Modality in Precision Medicine. *Cell Chem Biol* 2020;27:998–1014. <https://doi.org/10.1016/j.CHEMBIOL.2020.07.020>.
- [13] Sun X, Gao H, Yang Y, He M, Wu Y, Song Y, et al. PROTACs: great opportunities for academia and industry. *Signal Transduct Target Ther* 2019;4. <https://doi.org/10.1038/S41392-019-0101-6>.
- [14] Cruz Walma DA, Chen Z, Bullock AN, Yamada KM. Ubiquitin ligases: guardians of mammalian development. *Nat Rev Mol Cell Biol* 2022;0123456789. <https://doi.org/10.1038/s41580-021-00448-5>.
- [15] Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, et al. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 2021;49:D480–9. <https://doi.org/10.1093/NAR/GKAA1100>.
- [16] Blum M, Chang HY, Chuguransky S, Grego T, Kandasamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* 2021;49:D344–54. <https://doi.org/10.1093/NAR/GKAA977>.
- [17] Ribeiro AJM, Holliday GL, Furnham N, Tyzack JD, Ferris K, Thornton JM. Mechanism and Catalytic Site Atlas (M-CSA): a database of enzyme reaction mechanisms and active sites. *Nucleic Acids Res* 2018;46:D618–23. <https://doi.org/10.1093/nar/gkx1012>.
- [18] Sheridan RP, Maiorov VN, Holloway MK, Cornell WD, Gao YD. 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* 2010;50:2029–40. <https://doi.org/10.1021/C1100312T>.
- [19] Backus KM, Correia BE, Lum KM, Forli S, Horning BD, González-Páez GE, et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* 2016;534:570–4. <https://doi.org/10.1038/NATURE18002>.
- [20] Chi J (Jack), Li H, Zhou Z, Izquierdo-Ferrer J, Xue Y, Wavelet CM, et al. A novel strategy to block mitotic progression for targeted therapy. *EBioMedicine* 2019;49:40–54. <https://doi.org/10.1016/j.EBIOM.2019.10.013>.
- [21] Simpson GL, Bertrand SM, Borthwick JA, Campobasso N, Chabanet J, Chen S, et al. Identification and Optimization of Novel Small c-Abl Kinase Activators Using Fragment and HTS Methodologies. *J Med Chem* 2019;62:2154–71. <https://doi.org/10.1021/ACS.JMEDCHEM.8B01872>.
- [22] Perry JJP, Harris RM, Moiani D, Olson AJ, Tainer JA. p38 α MAP Kinase C-Terminal Domain Binding Pocket Characterized by Crystallographic and Computational Analyses. *J Mol Biol* 2009;391:1–11. <https://doi.org/10.1016/j.jmb.2009.06.005>.
- [23] Sadowsky JD, Burlingame MA, Wolan DW, McClendon CL, Jacobson MP, Wells JA. Turning a protein kinase on or off from a single allosteric site via disulfide trapping. *Proc Natl Acad Sci U S A* 2011;108:6056–61. <https://doi.org/10.1073/pnas.1102376108>.
- [24] Wood DJ, Lopez-Fernandez JD, Knight LE, Al-Khawaldeh I, Gai C, Lin S, et al. FragLites - Minimal, Halogenated Fragments Displaying Pharmacophore Doubles. An Efficient Approach to Druggability Assessment and Hit Generation. *J Med Chem* 2019;62:3741–52. https://doi.org/10.1021/ACS.JMEDCHEM.9B00304/SUPPL_FILE/JM9B00304_SI_003.PDF.
- [25] Yan Y, Edward Zhou X, Novick SJ, Shaw SJ, Li Y, Brunzelle JS, et al. Structures of AMP-activated protein kinase bound to novel pharmacological activators in phosphorylated, non-phosphorylated, and nucleotide-free states. *J Biol Chem* 2019;294:953–67. <https://doi.org/10.1074/jbc.RA118.004883>.
- [26] Hennings H, Blumberg PM, Pettit GR, Herald CL, Shores R, Yuspa SH. Bryostatins 1, an activator of protein kinase C, inhibits tumor promotion by phorbol esters in SENCAR mouse skin. *Carcinogenesis* 1987;8:1343–6. <https://doi.org/10.1093/CARCIN/8.9.1343>.
- [27] Mach UR, Lewin NE, Blumberg PM, Kozikowski AP. Synthesis and pharmacological evaluation of 8- and 9-substituted benzolactam-v8 derivatives as potent ligands for protein kinase C, a therapeutic target for Alzheimer's disease. *ChemMedChem* 2006;1:307–14. <https://doi.org/10.1002/CMDC.200500068>.
- [28] Zhang G, Kazanietz MG, Blumberg PM, Hurley JH. Crystal Structure of the Cys2 Activator-Binding Domain of Protein Kinase C δ in Complex with Phorbol Ester. *Cell* 1995;81:917–24.
- [29] Sharma P, Mahen R, Rossmann M, Stokes JE, Hardwick B, Huggins DJ, et al. A cryptic hydrophobic pocket in the polo-box domain of the polo-like kinase PLK1 regulates substrate recognition and mitotic chromosome segregation. *Sci Rep* 2019;9. <https://doi.org/10.1038/S41598-019-50702-2>.
- [30] Schapira M, Tyers M, Torrent M, Arrowsmith CH. WD40 repeat domain proteins: a novel target class? *Nat Rev Drug Discov* 2017;16:773–86. <https://doi.org/10.1038/NRD.2017.179>.
- [31] Kozlyuk N, Gilston BA, Salay LE, Gogliotti RD, Christov PP, Kim K, et al. A fragment-based approach to discovery of Receptor for Advanced Glycation End products inhibitors. *Proteins* 2021;89:1399–412. <https://doi.org/10.1002/PROT.26162>.
- [32] Guo Y, Mao X, Xiong L, Xia A, You J, Lin G, et al. Structure-Guided Discovery of a Potent and Selective Cell-Active Inhibitor of SETDB1 Tudor Domain. *Angew Chem Int Ed Engl* 2021;60:8760–5. <https://doi.org/10.1002/ANIE.202017200>.
- [33] Tautermann CS, Binder F, Büttner FH, Eickmeier C, Fiegen D, Gross U, et al. Allosteric Activation of Striatin-Enriched Protein Tyrosine Phosphatase (STEP, PTPN5) by a Fragment-like Molecule. *J Med Chem* 2019;62:306–16. <https://doi.org/10.1021/ACS.JMEDCHEM.8B00857>.
- [34] Leonard D, Huang W, Izadmeh S, O'Connor CM, Wiredja DD, Wang Z, et al. Selective PP2A Enhancement through Biased Heterotrimer Stabilization. *Cell* 2020;181:688–701.e16. <https://doi.org/10.1016/j.cell.2020.03.038>.
- [35] Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* 1996;85:1067–76. [https://doi.org/10.1016/S0092-8674\(00\)81307-0](https://doi.org/10.1016/S0092-8674(00)81307-0).
- [36] Genera M, Samson D, Raynal B, Haouz A, Baron B, Simenel C, et al. Structural and functional characterization of the PDZ domain of the human phosphatase PTPN3 and its interaction with the human papillomavirus E6 oncoprotein. *Sci Rep* 2019;9. <https://doi.org/10.1038/S41598-019-43932-X>.
- [37] Lin EYS, Silvan LF, Marcotte DJ, Banos CC, Jow F, Chan TR, et al. Potent PDZ-Domain PICK1 Inhibitors that Modulate Amyloid Beta-Mediated Synaptic Dysfunction. *Sci Rep* 2018;8. <https://doi.org/10.1038/S41598-018-31680-3>.
- [38] Böttcher J, Dilworth D, Reiser U, Neumüller RA, Schleicher M, Petronczki M, et al. Fragment-based discovery of a chemical probe for the PWWP1 domain of NSD3. *Nat Chem Biol* 2019;15:822–9. <https://doi.org/10.1038/S41589-019-0310-X>.
- [39] Dilworth D, Hanley RP, Ferreira de Freitas R, Allali-Hassani A, Zhou M, Mehta N, et al. A chemical probe targeting the PWWP domain alters NSD2 nucleolar localization. *Nat Chem Biol* 2021. <https://doi.org/10.1038/S41589-021-00898-0>.
- [40] Xu C, Meng F, Park K-S, Storey AJ, Gong W, Tsai Y-H, et al. A NSD3-targeted PROTAC suppresses NSD3 and cMyc oncogenic nodes in cancer cells. *Cell Chem Biol* 2021. <https://doi.org/10.1016/j.CHEMBIOL.2021.08.004>.
- [41] Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* 2003;101:1520–9. <https://doi.org/10.1182/BLOOD-2002-06-1675>.
- [42] Yuan G, Flores NM, Hausmann S, Lofgren SM, Kharchenko V, Angulo-Ibanez M, et al. Elevated NSD3 histone methylation activity drives squamous cell lung cancer. *Nature* 2021;590:504–8. <https://doi.org/10.1038/S41586-020-03170-Y>.
- [43] Rodriguez-Paredes M, Martinez De Paz A, Simó-Riudalbas L, Sayols S, Moutinho C, Moran S, et al. Gene amplification of the histone methyltransferase SETDB1 contributes to human lung tumorigenesis. *Oncogene* 2014;33:2807–13. <https://doi.org/10.1038/ONC.2013.239>.
- [44] Sengupta D, Zeng L, Li Y, Hausmann S, Ghosh D, Yuan G, et al. NSD2 dimethylation at H3K36 promotes lung adenocarcinoma pathogenesis. *Mol Cell* 2021;81. <https://doi.org/10.1016/j.MOLCEL.2021.08.034>.
- [45] Filippakopoulos P, Knapp S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat Rev Drug Discov* 2014;13:337–56. <https://doi.org/10.1038/NRD4286>.
- [46] Upadhyay AK, Judge RA, Li L, Pithawalla R, Simanis J, Bodelle PM, et al. Targeting lysine specific demethylase 4A (KDM4A) tandem TUDOR domain - A fragment based approach. *Bioorg Med Chem Lett* 2018;28:1708–13. <https://doi.org/10.1016/j.BMCL.2018.04.050>.
- [47] Verdin E, Ott M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* 2015;16:258–64. <https://doi.org/10.1038/NRM3931>.
- [48] Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 2009;325:834–40. <https://doi.org/10.1126/SCIENCE.1175371>.
- [49] Song R, Wang ZD, Schapira M. Disease Association and Druggability of WD40 Repeat Proteins. *J Proteome Res* 2017;16:3766–73. <https://doi.org/10.1021/ACS.JPROTEOME.7B00451>.
- [50] Ferreira De Freitas R, Harding RJ, Franzoni I, Ravichandran M, Mann MK, Ouyang H, et al. Identification and Structure-Activity Relationship of HDAC6 Zinc-Finger Ubiquitin Binding Domain Inhibitors. *J Med Chem* 2018;61:4517–27. <https://doi.org/10.1021/ACS.JMEDCHEM.8B00258>.
- [51] Mann MK, Franzoni I, De Freitas RF, Tempel W, Houliston S, Smith L, et al. Discovery of Small Molecule Antagonists of the USP5 Zinc Finger Ubiquitin-Binding Domain. *J Med Chem* 2019;62:10144–55. <https://doi.org/10.1021/ACS.JMEDCHEM.9B00988>.
- [52] Faesen AC, Dirac AMG, Shanmugham A, Ova H, Perrakis A, Sixma TK. Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol Cell* 2011;44:147–59. <https://doi.org/10.1016/j.MOLCEL.2011.06.034>.
- [53] Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nat Rev Nephrol* 2019;15:346–66. <https://doi.org/10.1038/S41581-019-0129-4>.
- [54] Ramirez DH, Aonbangkhen C, Wu HY, Naftaly JA, Tang S, O'Meara TR, et al. Engineering a Proximity-Directed O-GlcNAc Transferase for Selective Protein O-GlcNAcylation in Cells. *ACS Chem Biol* 2020;15:1059–66. <https://doi.org/10.1021/ACS.CHEMBIO.0C00074>.
- [55] Ferreira De Freitas R, Liu Y, Szewczyk MM, Mehta N, Li F, McLeod D, et al. Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2. *J Med Chem* 2021;64:1584–92. <https://doi.org/10.1021/ACS.JMEDCHEM.0C01768>.
- [56] Harding RJ, Ferreira De Freitas R, Collins P, Franzoni I, Ravichandran M, Ouyang H, et al. Small Molecule Antagonists of the Interaction between the Histone Deacetylase 6 Zinc-Finger Domain and Ubiquitin. *J Med Chem* 2017;60:9090–6. https://doi.org/10.1021/ACS.JMEDCHEM.7B00933/SUPPL_FILE/JM7B00933_SI_012.CSV.
- [57] Ge Y, Ramirez DH, Yang B, D'Souza AK, Aonbangkhen C, Wong S, et al. Target protein deglycosylation in living cells by a nanobody-fused split O-GlcNAcase. *Nat Chem Biol* 2021;17:593–600. <https://doi.org/10.1038/S41589-021-00757-Y>.

- [58] Xiao H, Woods EC, Vukojicic P, Bertozzi CR. Precision glycolyx editing as a strategy for cancer immunotherapy. *Proc Natl Acad Sci U S A* 2016;113:10304–9. <https://doi.org/10.1073/PNAS.1608069113>.
- [59] Gray MA, Stanczak MA, Mantuano NR, Xiao H, Pijnenborg JFA, Malaker SA, et al. Targeted glycan degradation potentiates the anticancer immune response in vivo. *Nat Chem Biol* 2020;16:1376–84. <https://doi.org/10.1038/s41589-020-0622-x>.
- [60] Ward CC, Kleinman JI, Brittain SM, Lee PS, Chung CYS, Kim K, et al. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem Biol* 2019;14:2430–40. <https://doi.org/10.1021/ACSCHEMBO.8B01083>.
- [61] Zhang X, Crowley VM, Wucherpennig TG, Dix MM, Cravatt BF. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat Chem Biol* 2019;15:737–46. <https://doi.org/10.1038/s41589-019-0279-5>.
- [62] Spradlin JN, Hu X, Ward CC, Brittain SM, Jones MD, Ou L, et al. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat Chem Biol* 2019;15:747–55. <https://doi.org/10.1038/s41589-019-0304-8>.
- [63] Luo M, Spradlin JN, Boike L, Tong B, Brittain SM, McKenna JM, et al. Chemoproteomics-enabled discovery of covalent RNF114-based degraders that mimic natural product function. *Cell Chem Biol* 2021;28:559–566.e15. <https://doi.org/10.1016/j.CHEMBOI.2021.01.005>.
- [64] Tong B, Luo M, Xie Y, Spradlin JN, Tallarico JA, McKenna JM, et al. Bardoxolone conjugation enables targeted protein degradation of BRD4. *Sci Rep* 2020;10. <https://doi.org/10.1038/s41598-020-72491-9>.
- [65] He Y, Selvaraju S, Curtin ML, Jakob CG, Zhu H, Comess KM, et al. The EED protein–protein interaction inhibitor A-395 inactivates the PRC2 complex. *Nat Chem Biol* 2017;13(13):389–95. <https://doi.org/10.1038/nchembio.2306>.
- [66] Qi W, Zhao K, Gu J, Huang Y, Wang Y, Zhang H, et al. An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nat Chem Biol* 2017;13(13):381–8. <https://doi.org/10.1038/nchembio.2304>.
- [67] Grebien F, Vedadi M, Getlik M, Giambrodo R, Grover A, Avellino R, et al. Pharmacological targeting of the Wdr5–MLL interaction in C/EBP α N-terminal leukemia. *Nat Chem Biol* 2015;11(11):571–8. <https://doi.org/10.1038/nchembio.1859>.
- [68] Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein–sequence space with high-accuracy models. *Nucleic Acids Res* 2021. <https://doi.org/10.1093/NAR/GKAB1061>.
- [69] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nat* 2021;596:583–9. <https://doi.org/10.1038/s41586-021-03819-2>.
- [70] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* (80-) 2021;373:871–6. https://doi.org/10.1126/SCIENCE.AB18754/SUPPL_FILE/AB18754_MDAR_REPRODUCIBILITY_CHECKLIST.PDF.
- [71] Drummond ML, Williams CI. In Silico Modeling of PROTAC-Mediated Ternary Complexes: Validation and Application. *J Chem Inf Model* 2019;59:1634–44. https://doi.org/10.1021/ACS.JCIM.8B00872/SUPPL_FILE/C18B00872_SI_001.PDF.
- [72] Drummond ML, Henry A, Li H, Williams CI. Improved accuracy for modeling ProTAC-mediated ternary complex formation and targeted protein degradation via new in silico methodologies. *J Chem Inf Model* 2020;60:5234–54. https://doi.org/10.1021/ACS.JCIM.0C00897/SUPPL_FILE/C10C00897_SI_001.PDF.
- [73] Zaidman D, Prilusky J, London N. PROsettaC: Rosetta based modeling of PROTAC mediated ternary complexes. *J Chem Inf Model* 2020;60:4894–903. https://doi.org/10.1021/ACS.JCIM.0C00589/SUPPL_FILE/C10C00589_SI_001.PDF.
- [74] Dixon T, MacPherson D, Mostofian B, Dauszhenka T, Lotz S, McGee D, et al. Atomic-Resolution Prediction of Degradation-mediated Ternary Complex Structures by Combining Molecular Simulations with Hydrogen Deuterium Exchange. *BioRxiv* 2021:2021.09.26.461830.
- [75] Liao J, Nie X, Unarta IC, Erickson SS, Tang W. In Silico Modeling and Scoring of PROTAC-Mediated Ternary Complex Poses. *J Med Chem* 2022. <https://doi.org/10.1021/acs.jmedchem.1c02155>.