JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Perspective

Advancing Targeted Protein Degradation via Multiomics Profiling and Artificial Intelligence

Miquel Duran-Frigola,* Marko Cigler, and Georg E. Winter*

Cite This: J. Am. Chem. Soc. 2023, 145, 2711−2732
Read Online
ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Only around 20% of the human proteome is considered to be druggable with small-molecule antagonists. This leaves some of the most compelling therapeutic targets outside the reach of ligand discovery. The concept of targeted protein degradation (TPD) promises to overcome some of these limitations. In brief, TPD is dependent on small molecules that induce the proximity between a protein of interest (POI) and an E3 ubiquitin ligase, causing ubiquitination and degradation of the POI. In this perspective, we want to reflect on current challenges in the field, and discuss how advances in multiomics profiling, artificial intelligence, and machine learning (AI/ML) will be vital in overcoming them. The presented roadmap is discussed in the context of small-molecule degraders but is equally applicable for other emerging proximity-inducing modalities.

Targeted protein degradation (TPD) is a pharmacologic modality that harbors the potential to address critical shortcomings of inhibitor-centric approaches. In order to deliver on these promises, it will be essential to continue to innovate and advance this field. In this perspective, we initially summarize how TPD is differentiated from small-molecule agonists and antagonists. Moreover, we are reflecting on some key challenges the field is currently facing, and how advances in multiomics profiling, artificial intelligence, and machine learning (AI/ML) can help researchers to address and overcome these challenges. In this context, we are reviewing strategies that have already been implemented, but also expand on approaches that should provide a tangible upside in the future.

1. INHIBITOR-CENTRIC VERSUS PROXIMITY-INDUCING PHARMACOLOGY

1.1. Concepts and Limitations of Inhibitor-Centric Pharmacology. Most ligand discovery efforts are focused on small molecules that bind to hydrophobic pockets and aim to block the biochemical activity encoded by the bound site. Ligand optimization is focused on improving binding potency and residence time to maximize depth and duration of target occupancy to achieve therapeutic efficacy. Consequently, inhibitor-centric concepts of ligand discovery have recently been summarized as "occupancy-driven" small-molecule modalities.[1] For many targets that are currently pursued in drug development, the occupancy-driven inhibitor discovery approach is in principle well suited. From a global perspective, these "low hanging fruit" aspects of ligand discovery however make up only around 20% of the human genome, thus leaving some of the most compelling targets outside the reach of small-molecule discovery approaches.[2,3] This includes, among others, transcription factors or other intrinsically disordered proteins, as well as scaffolding proteins.[4,5]

Upon closer inspection, even a protein that is classified as ligandable can pose significant challenges to the drug discovery process. The ligandable domain might for instance be shared by a plethora of other proteins, thus complicating the development of selective inhibitors. Another issue could emerge from insufficient target characterization. Often, we lack functional resolution of a particular protein target, meaning that we do not necessarily know if the pharmacologically addressable domain is also relevant for the disease-association. For both scenarios, protein kinases can serve as a well-established example. The highly conserved ATP-binding site of the more than 500 kinases encoded in the human genome makes the discovery of selective inhibitors a challenging and an often unachievable venture.[6,7] Moreover, it is also well-documented that many kinases feature scaffolding functions. This is best exemplified by the subfamily of pseudokinases, such as HER3. HER3 features very low catalytic activity but engages in oncogenic signaling by acting as a scaffold to enable receptor heterodimerization.[8−10] Consequently, occupying the, *per se* druggable, ATP binding pocket of HER3 has been shown to be insufficient to functionally disrupt HER3.[11] Importantly, even if a protein domain is not functionally compromised, benefits associated with small-molecule inhibitors might be limited. This is exemplified, among others, by bromodomain-containing proteins. The druggability of bromodomains was first established for the bromodomain and extra-terminal domain (BET) family of proteins.[12] These results established the

Published: January 27, 2023

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

motivation for ligand discovery efforts for other bromodomain-containing proteins such as the histone-acetyltransferases CBP/EP300,[13−17] the E3 ligase TRIM24,[18,19] and the chromatin remodelers SMARCA2/4.[20] While these efforts culminated in potent, specific, and well-characterized ligands for the individual bromodomains, ensuing functional analyses revealed that they are largely lacking cellular efficacy.[21] In all cases, the targeted proteins also harbored multiple other functional domains and are often participating in intricate protein−protein interactions that appeared to be bromodomain independent. Collectively, these examples outline that the pharmacologically addressable domain is not necessarily functionally involved in the observed phenotype. In other words: a protein that appears "druggable" on the first glance might still not be within reach of the conventional occupancy-driven and inhibitor-centric pharmacology.

Lastly, even if a protein target can be meaningfully blocked by a small-molecule inhibitor, a potential therapeutic window can be compromised by toxicity that emerges from target inhibition in other tissues or cell types. This is a particular challenge if functional inhibition of the target requires a high fraction of target occupancy and inhibitors need to be dosed at high concentrations leading to elevated plasma concentrations.

Over the past decade, we have experienced tremendous progress in the innovation or further development of small-molecule modalities that overcome some of the limitations of occupancy-driven, competitive inhibitors. These innovations range from a renaissance of covalent small molecule inhibitors to advancements in proximity-inducing pharmacology, most notably the emerging field of TPD.[22,23] In the following, we want to outline how TPD can address inherent challenges of occupancy-driven pharmacology. Next, we want to introduce three key challenges that TPD is currently facing and share our perspective on how these problems can be overcome. In particular, we argue that the advent of AI/ML technologies, coupled with multiomics profiling techniques for phenotypic screening, may pave the way toward efficient identification and rational design of small-molecule degraders.

1.2. Degraders are proximity inducers that can address inhibitor shortcomings. One avenue to overcome challenges of conventional inhibitors is to fundamentally change the perspective and perception of small-molecule action. An intriguing solution is to turn attention to small molecules that can change (rather than inhibit) protein function, allowing us to chemically rewire biological circuits. One way to change target protein function via small molecules is to induce naturally nonoccurring protein−protein interactions (PPIs) involving that target. This forms the conceptual foundation of chemically induced dimerization (CID), which is also referred to as proximity-inducing pharmacology.[24] Some of the best-studied small-molecules such as rapamycin, FK506, or cyclosporin function via stabilizing non-natural protein−protein interactions.[25] Collectively, these molecules have also been referred to as "molecular glues". The concept of recruiting sterically confined effector proteins to inhibit target protein function is still actively pursued both in basic research and in translational studies. Due to space constraints, we want to refer to excellent recent reviews covering this space.[26,27] The fundamental concept of leveraging proximity-inducing pharmacology to change biological circuits has been explored via a multitude of different avenues, also including biologicals.[28−31] Small molecules, mostly of a heterobifunctional design, have been leveraged to install PPIs with kinases,[32] phosphatases,[33] acetyltransferases,[34] and deubiquitinases,[35,36] thus enabling researchers to alter posttranslational modifications on a target protein to change protein function. Moreover, incorporation of RNA-binding small-molecules into a heterobifunctional design has afforded ribonuclease targeting chimeras (RIBOTACs): small molecules that induce proximity between an RNA of interest and a ribonuclease to prompt RNA degradation.[37−39]

A powerful embodiment of proximity-inducing drugs are small-molecule degraders. Significant attention has been given to degraders that hijack the autophagy machinery;[40−42] however, most efforts in the field focus on degraders that co-opt the ubiquitin-proteasome system (UPS).[43] For the purpose of this perspective, we will hence exclusively focus on these compounds. Typically, they function by inducing the proximity between a target protein of interest (POI) and an E3 ubiquitin ligase. If the ligase and the POI are positioned in a favorable manner, the POI gets poly-ubiquitinated and degraded by the proteasome. Degraders thus harbor catalytic potential, meaning that one equivalent of degrader can induce the degradation of multiple POI equivalents. Due to this characteristic feature, degraders can achieve cellular potency that is far below their K_D for the respective POI. Most small-molecule degraders fall into one of two categories. On the one hand, heterobifunctional proteolysis-targeting chimeras (PROTACs) leverage two different ligands to simultaneously engage the POI and the ligase. Both ligands are connected by a flexible linker element that ensures proper spacing and architecture of the resulting ternary complex. On the other hand, molecular glue degraders (MGDs) connect ligase and POI by modulating protein surface topologies and orchestrating multiple PPIs between the involved proteins. They typically only bind one component in isolation, and ternary complex formation is driven by overall binding cooperativity. This enables the degradation of proteins that are generally considered to be unliganded targets for small-molecule modulation. A more in-depth differentiation between PROTACs and MGDs is discussed elsewhere.[44]

1.3. Small-molecule degraders can address challenges of occupancy-driven pharmacology. The pharmacology of degraders enables them to address many challenges that are inherent to the occupancy-driven drug design. Most importantly, degraders can extend the reach of the druggable proteome. Focusing on PROTACs, this is exemplified by degraders of the aforementioned bromodomain proteins SMARCA2/4, TRIM24, and CBP. When ineffectual bromodomain ligands were furnished into PROTAC targeting warheads, potent and selective degradation could be achieved.[45−47] In sum, PROTACs hence allow ineffectual binders to be repurposed for conversion into functional degraders. PROTACs are therefore, in principle, able to address proteins that can be liganded, but fall short for proteins that are considered unliganded, such as the majority of transcription factors. Seminal work around thalidomide and its analogs (collectively referred to as "IMiDs") has revealed a blueprint to close this gap. In brief, IMiDs bind the E3 ligase cereblon (CRBN), thus complementing a surface patch that subsequently recruits a variety of C2H2 zinc finger transcription factors for ubiquitination and degradation.[48−51] Recruitment of these transcription factors (such as IKZF1 and IKZF3) happens in the absence of a measurable binding affinity between the IMiDs and IKZF1/3. Rather, IMiDs catalyze ternary complex formation by orchestrating multiple PPIs.[52,53] Additional MGDs have been reported, which induce the degradation of proteins devoid of obvious ligand-binding

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

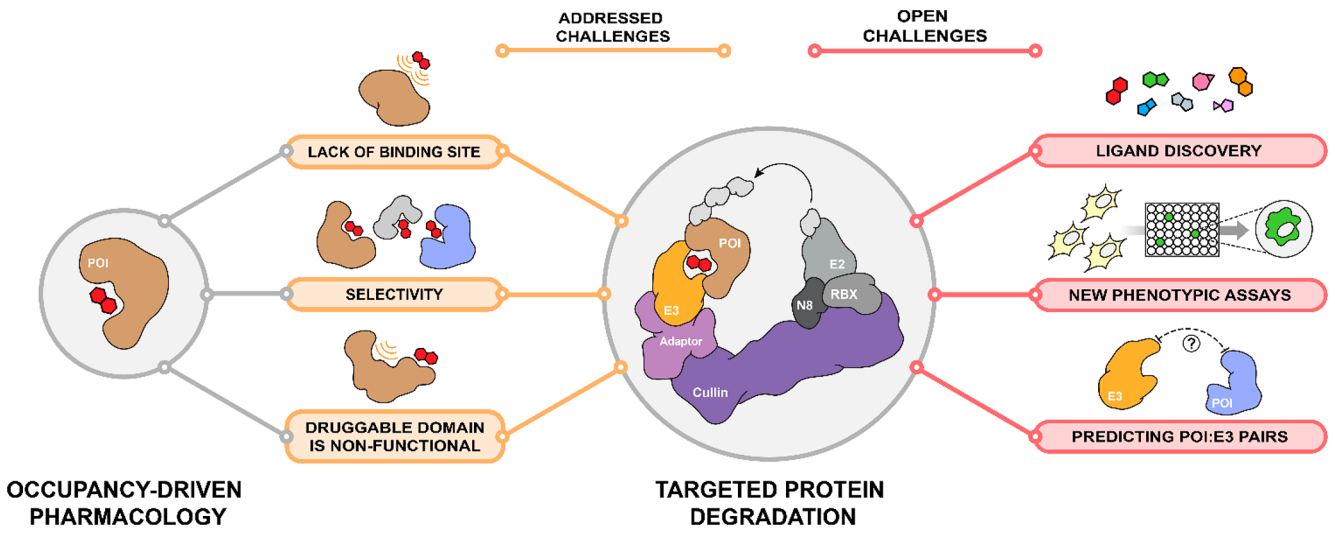

Figure 1. Addressed and open challenges of targeted protein degradation.

pockets, such as RBM39[54−57] or cyclin K (CycK).[58−60] In sum, MGDs have shown the potential to induce the degradation of proteins that are considered unligandable. Interestingly, recent evidence suggests that MGDs depend on a certain baseline affinity between E3 and POI in the absence of the respective small molecule.[61] Further research will be required to validate if this could be a predictive feature that can be exploited for *de novo* MGD design.

In addition to expanding the druggable space, degraders also address the challenge of inhibitor selectivity. There are multiple lines of evidence where promiscuous ligands could be furnished into highly selective PROTACs. Several mechanisms have been shown to contribute to this increase in selectivity. This includes steric incapability of some targets to form a productive ternary complex, or lack of a suitably positioned lysine residue for ubiquitination.[62−69] Another layer of degrader selectivity has been observed between mutated and *wild type* target proteins. Among others, degraders with comparable affinity for both forms were shown to preferentially degrade $BRAF^{V600E}$ over $BRAF^{WT}$.[70] In most reported cases, the observed selectivity of a particular degrader resulted from empirical optimization or was fortuitous, but first computational approaches empower rational design of selective degraders.[71] Lastly, degraders might also be able to address challenges in dose-limiting on-target toxicity. For instance, it was shown that a PROTAC that co-opts the E3 ligase VHL to induce the degradation of $BCL\text{-}X_L$ can spare the thrombocytopenia associated with systemic $BCL\text{-}X_L$ inhibition.[72] Mechanistically, this was rationalized by the low VHL levels in platelets. In addition, the catalytic mechanism of degraders can potentially omit the need for continuously high drug levels, thus providing another opportunity to limit toxicity. This advantage will be particularly evident for POIs with long half-life and slow protein resynthesis rates.

Collectively, small-molecule degraders offer solutions to address various limitations of occupancy-based inhibitor design. Arguably, many of these theoretical advantages and solutions need to be put into clinical practice. Pharmacokinetics (PK) and pharmacodynamics (PD) liabilities are likely to emerge for small-molecule degraders, especially in the case of the relatively large PROTAC compounds.[73] Moreover, the field needs to evolve from episodically identified points of differentiation to a rational framework to design potent, selective degraders that address clinically meaningful challenges. In the following, we want to highlight some of the associated key challenges.

1.4. Current Challenges in the TPD Field. Many of the challenges of degrader modalities will be associated with translating preclinical data into effective medicines that provide measurable benefit in clinical studies. Given the scope of this perspective, we will here however focus on challenges that are associated with broadening the scope of degraders in a rational and data-driven manner. To this end, we want to focus our attention on three topics: (i) Expanding the reach of PROTACs through advanced ligand discovery efforts; (ii) Identification of MGDs via phenotypic screens coupled to cutting-edge target identification; (iii) Predicting compatible POI-ligase pairs for degrader discovery (Figure 1).

A key challenge (i) for PROTACs has been the dependency on a limited number of E3 ligases that can effectively be co-opted via potent and well-characterized ligands, preventing target- and cell-type selective degradation. Moreover, most reported targeting ligands are leveraging well-described chemical matter, such as kinase inhibitors. The dependency on existing ligands critically stymies progress toward undrugged proteins that are *per se* ligandable, but where no chemical matter has been reported. The discovery of MGDs has, for the most part, been serendipitous (ii). While first rational approaches have been reported,[59,74] further scaling the discovery of MGDs will require innovations in phenotypic assays and in downstream target validation strategies. Finally, (iii) an important goal would be to develop a computational framework that can prioritize POI-ligase pairs that feature compatible surfaces, thus empowering the de novo design of degraders in a structure-informed manner.

2. EMPOWERING LIGAND DISCOVERY WITH PROTEOMICS AND AI/ML

2.1. Proteomics-Empowered Ligand Discovery Approaches. The paucity of ligands for E3 ligases and currently undrugged proteins critically hampers progress in TPD. Traditional ligand discovery efforts typically focus on recombinant protein assays, aiming to identify small molecules that either bind a structured site or block the encoded

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

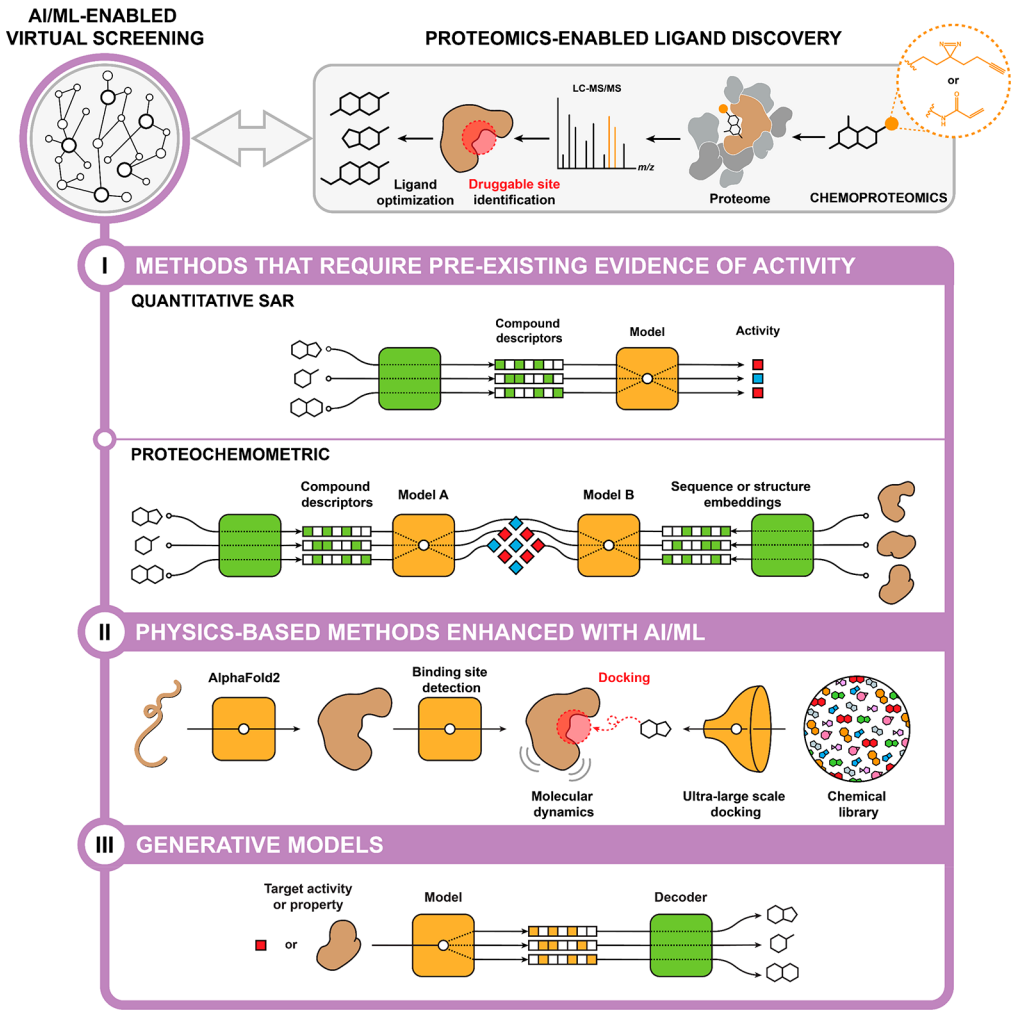

Figure 2. Advancing ligand discovery with proteomics and computational approaches. AI/ML-based methods can be divided into distinct categories (I–III) depending on the information they require as input.

biochemical activity. This includes, among many others, fluorescence polarization,[75,76] surface plasmon resonance,[77] thermal shift assays,[78,79] and binding-enrichment assays that are amenable to DNA-encoded chemical libraries (DELs).[80] These assays can inform ligand discovery when the assessed recombinant protein adopts a conformation that closely approximates the native fold found in a cellular environment. Unfortunately, many undrugged proteins do not fall in this category.

In contrast, chemoproteomics aims to map ligand–protein interaction on a proteome-wide scale, typically in intact cells. In principle, it couples a drug-affinity enrichment with mass spectrometry to identify proteins that bind to a given small molecule.[81−83] A key advantage of this approach over conventional ligand discovery strategies is the ability to probe proteins in their native environments. This enables the discovery of actionable binding sites that might be undetectable or nonobvious in isolated, recombinant systems. For instance, chemoproteomics profiling has led to the discovery that the natural product nimbolide covalently ligands a cysteine residue in a disordered region of the E3 ligase RNF114. This enabled the design of RNF114-based degraders, also expanding to other chemotypes.[84,85] Chemoproteomics is often employed to identify protein targets of phenotypic screening hits. Over the past decade, it has however also successfully been employed with fragment-like structures, aiming to prospectively identify specific fragment–protein interactions to enable ligand optimization. Initial work has predominantly focused on cysteine-targeting, covalent fragments.[86−88] These approaches have the advantage that they not only inform on the liganded protein but also precisely reveal the covalently bound cysteine residue. Of relevance for TPD, this has recently identified fragments that covalently engage E3 ligases.[84,89−93] Encouragingly, as mentioned above, these fragments could successfully be incorporated into functional PROTACs capable of inducing the degradation of BRD4 and other POIs. Future research will reveal if, in general, hits from covalent fragment screens can be advanced toward drug-like ligands for PROTAC or MGDs. Moreover, we will likely witness the continuation of a systematic expansion toward other amino acids including lysines or tyrosines.[94−97] Of note, proteomics-enabled ligand discovery is not limited to covalent compounds. Equipping noncovalent fragments with a photoreactive diazirine moiety similarly enables their proteome-wide target identification without the prerequisite of a particular amino acid in or nearby a ligandable site.[98−100] For both covalent and noncovalent compounds, the site of compound binding is typically revealed by introducing a cleavable connection between the fragment and the bead resin, or by using desthiobiotin.[88,101] This step is instrumental to

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

inform ensuing ligand optimization efforts. Below, we will outline how AI/ML-enabled docking studies can critically support this step, particularly for proteins that lack complete structural characterization.

In sum, proteomics-driven ligand discovery has the potential to significantly expand the reach of TPD and other proximity-inducing modalities. Ligands identified for undrugged POIs can be incorporated in a heterobifunctional PROTAC design. Identification of novel E3 ligands can empower further diversification of PROTACs. Elaboration of an E3-binding ligand into an E3-hijacking MGD is less obvious yet could be critically empowered by (protein−protein) docking[102,103] and molecular surface interaction fingerprinting,[104] as outlined in section 4 below. Next, we will discuss how purely computational methods can support ligand discovery, including identification of putative binding sites in protein structures, large-scale molecular docking, molecular dynamics simulations, and quantitative structure−activity relationship (QSAR) models.

2.2. Ligand Discovery Enabled and Empowered by AI/ML. Computational approaches to ligand discovery can be divided into two broad categories (Figure 2). On the one hand, there are methods that require pre-existing evidence of a certain activity (typically, binding affinity measurements for a target of interest). The classical representative of this category is QSAR modeling where structural/physicochemical properties of the small-molecules are correlated with experimentally measured values of activity. Based on the observed patterns of correlation, the activity of new compounds can be predicted. Translated to the field of AI/ML, QSAR can be seen as a "supervised" approach requiring training data. Indeed, a long repertoire of supervised AI/ML algorithms has been applied to the discovery of new ligand−protein pairs based on previous binding data,[105] borrowing techniques from the fields of natural language processing,[106] image recognition,[107] and network analysis.[108] In essence, AI/ML methods are able to dynamically engineer small-molecule features that are optimally correlated with the experimental outcome. Typically, AI/ML methods engineer molecular features based on a string (text) representation of the molecule (SMILES or, more recently, SELFIES[109]), or a small-molecule graph where nodes are atoms and edges are bonds annotated with their respective properties (atomic weight, valence, etc.). Modern supervised AI/ML algorithms have shown outstanding performance in a broad set of biophysical and physiological prediction tasks, and are now incorporated in QSAR benchmarks.[110] In TPD, QSAR models have for instance been employed to identify novel glutarimide analogs that promote selective degradation of IKZF3 and/or GSPT1 via CRBN.[111]

A fundamental limitation of supervised AI/ML is the requirement of previously available experimental data. This is aggravated by the fact that advanced AI/ML methods are generally data-greedy, meaning they require more data than the classical set of QSAR techniques. As a result, supervised AI/ML for ligand discovery is best suited for well-studied ligandable proteins. Indeed, recent efforts have broadened the applicability of supervised AI/ML models to low-data and no-data scenarios,[112,113] i.e. to under-studied and unliganded proteins. Conceptually, the improvement has been to simultaneously train over millions of experimental data points (spanning thousands of targets) contained within large chemogenomics databases such as ChEMBL,[114] BindingDB,[115] or PubChem BioAssay.[116] This kind of global training is laid out as a "bimodal" AI/ML model, where both the ligand structure and the protein sequence/structure are passed as inputs. Thus, the so-called proteochemometric methods focus their attention not only on the properties of the ligands but also on the proteins, and ultimately aim to discover the rules of protein−ligand recognition. In this context, algorithms to embed protein sequences into dense numerical representations that are amenable for AI/ML offer promising opportunities. These numerical representations have been proven to capture a wide spectrum of protein characteristics, including structural domains, family membership, and function.[117−119] One advantage of proteochemometric approaches, compared to the classical "one protein, one model" rationale, is that, at least in principle, ligand−protein interaction predictions can be done for any protein of interest, as long as their sequence is available. However, the applicability of existing proteochemometric methods is still limited to relatively well-studied regions of the protein space, enriched with ligandable proteins for which a substantial amount of experimental data is available. In practice, the performance of proteochemometric methods may vary depending on protein families, with a bias toward protein families that are already considered to be ligandable.[120] Extrapolation to genuinely novel targets, belonging to unliganded regions of the sequence space, is a major challenge for current supervised AI/ML methods. In this line, passing chemoproteomics data through these models may increase their domain of applicability,[121] since these high-throughput assays are likely to introduce new evidence for unseen targets and target families.

On the other hand, there are computational approaches that do not require previous knowledge of binding affinity (Figure 2). Traditionally, these methods have relied on physics-based calculations obtained through docking procedures and/or molecular dynamics simulations. An advantage of physics-based methods is that they are, in principle, applicable to any protein of interest as long as a high-resolution 3D structure is available. For years, the need for 3D structural data of the proteins has been perceived as a critical limitation, especially for membrane-bound proteins which are difficult to crystallize.[122] Ligand docking has been used with homology models of protein structures,[123] although the applicability is limited by the need for structural templates from close homologues,[124−126] since atomistic details in the binding site are crucial determinants of affinity. However, recent breakthroughs in proteome-wide AI/ML-based structure prediction solely based on sequence may help overcome this limitation. The unprecedented accuracy of AlphaFold2,[127] RosettaFold,[128] and others[129,130] offers, at least in principle, the opportunity to virtually screen any protein of interest with the structure-based approach, irrespective of its *a priori* ligandability (after AlphaFold2, the structural coverage of the human proteome increased from 48% to 76% of the total residues).[131] While there is ongoing debate as to whether e.g. AlphaFold2 structure predictions are of sufficient quality for protein−ligand docking,[132,133] it is clear that the possibility is within reach, even if some filters and extra preparation steps like molecular dynamics relaxation may be required.[134] In combination with knowledge gathered from chemoproteomics experiments, such as mapping of the ligand binding site, the accuracy and speed of structure-based simulations could be boosted, providing a tool to quickly screen compounds against fully structurally resolved proteomes.

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

Identification of putative binding sites in a protein structure is a first, advisable step in docking procedures. Narrowing the search space to specific regions of the protein allows for higher throughput compared to "blind" docking. Computationally, binding pockets can be identified through inspection of geometrical features in the protein surface, presence of putative hydrogen bond donors and acceptors, spots of hydrophobic contact, etc.[135−137] AI/ML models have been applied to the identification of these potential pockets, and especially to the up-ranking of the most ligandable ones.[138] Such predictions will be highly complementary to ligandability data obtained from chemoproteomics experiments. Following binding site identification (either fully computationally or supported by wet-lab evidence), each candidate ligand can be evaluated in the site region, where multiple binding poses are tried and scored. Evaluation can be done in flexible mode, allowing for induced fit between the ligand and the protein. Assessment of covalent binding is also possible,[139−141] which may be relevant to complement certain chemoproteomics screening platforms.[23] Recently, enhanced docking pipelines have been proposed to drastically improve the computational feasibility of docking procedures, enabling the screening of ultralarge "make-on-demand" chemical libraries such as the Enamine REAL collection.[122,142,143] For example, a 100-fold acceleration of the process was observed by docking a fraction of the chemical library in synchrony with an AI/ML ligand-based prediction of the remaining docking scores.[144]

2.3. Ligand Library Design Aided by AI/ML. Coupled with the augmented virtual screening capacity, we are witnessing an avalanche of "generative" AI/ML models capable of designing new small-molecule structures with desired target properties (Figure 2).[145−148] Generative models have been used to compile both focused and diversity-oriented chemical libraries,[149,150] to design inhibitors *de novo*,[151] and to perform hit-to-lead optimization,[152] typically in a multiobjective mode where e.g. selectivity, synthetic accessibility, and drug-like properties are improved simultaneously.[153,154] Of note, these algorithms can be interfaced with docking calculations in order to design small molecules with high docking scores.[155,156] This could simplify the design of compounds binding to unliganded proteins. Mature software tools now exist to run generative models on conventional computers,[154,157] and community-accepted benchmarks enable the systematic assessment of generated small-molecules, considering not only their predicted activity but also their novelty, validity, and diversity.[158,159]

Progressively, and despite initial limitations and criticisms,[160] AI/ML-based generative models are being incorporated into realistic experimental settings. For example, by constraining compounds with a subset of simple reactions and building blocks, a "synthesis-on-a-chip" procedure was devised where the AI/ML tool suggested compounds that could be immediately synthesized.[161] Other specialized methods focus on fragment elaboration,[162] which could be employed downstream of proteome-wide fragment screening experiments via aforementioned chemoproteomics. Moreover, existing platforms have been versioned to design optimal linkers connecting the two molecular subunits in a PROTAC.[163] Design of PPI inhibitors has also been attempted,[164] which may set the stage for the design of MGDs. As more degraders are discovered, it is likely that dedicated generative models will flourish, borrowing principles apprehended in generalistic models[165] and fine-tuning them to the degrader-specific tasks. Adaptability to a chemical space of interest by adjusting the inner parameters of a given model is a unique feature offered by AI/ML.

3. MOLECULAR GLUE DEGRADER DISCOVERY VIA ADVANCED PHENOTYPIC SCREENS

3.1. Phenotypic Discovery of MDGs. Historically, cell-based phenotypic assays have provided the richest source for MGD identification. The clinical evaluation of lenalidomide, one of the aforementioned CRBN-modulating IMiDs, was motivated by phenotypic assays for caspase cleavage of thalidomide analogs.[166] Similarly, a straightforward screen for IMiD-induced cytotoxicity in cells with low CRBN levels revealed the highly potent IMiD CC-92480,[167] which has recently entered clinical investigation. Here, the rationale for screening in CRBN low cells was to select for compounds with high potency and catalytic turnover that could overcome limited availability of CRBN. The discovery of CC-885 as a CRBN-dependent GSPT1 degrader[53] was likewise initially motivated by cancer cell line profiling studies, prompting the development of the clinical candidate CC-90009.[168,169] Cell-viability-based phenotypic assays have also led to the identification of several MGDs that induce the degradation of CycK. One study reported a strategy based on chemical profiling in cells deficient for UBE2M.[59] As UBE2M is required for the activity of a large fraction of all 250 cullin-RING ligases (CRL), this screen was poised to identify small molecules that functionally depend on a CRL. Another study analyzed correlative data to identify compounds whose activity correlates with expression of the CRL adaptor protein DDB1.[58] Also this study led to the identification of chemical matter that induces CycK degradation. Both studies converge on a shared mechanism of action where the identified compounds directly glue the CDK12-CycK complex to DDB1, thus prompting CycK degradation in the absence of a dedicated substrate receptor. Collectively, these examples demonstrate the relevance of cell-based assays as a starting point for MGD discovery. However, while increasingly more elaborate in their design, these studies fundamentally still relied on cell-viability-associated readouts. This comes with the limitation that the scope of protein targets is limited to those with an essential function in cell viability/cellular fitness. In this section we want to discuss how advanced phenotypic screens will enable the identification of next-generation MGDs, and how innovations in phenotypic readouts will require matched advancements in target identification strategies. In the following, we discuss how richer phenotypic readouts can contribute to a better understanding of TPD. We also present emerging AI/ML techniques that can contribute to efficient processing and integration of these complex data sets. In sum, we hypothesize that these advancements will be necessary in order to expand the reach of phenotypic assays to protein targets that are not essential for cellular fitness.

3.2. Introducing Diversity in Phenotypic Screens. Conceptually, we anticipate that advancements in phenotypic screens will emerge from introducing diversity that can be measured, quantified, interpreted, and predicted. Diversity can emerge from different genetic backgrounds, cell types, metabolic states, environmental conditions, or exposure to perturbagens, among others. In many examples that are mentioned in the following, this necessitates implementation of automated microscopy workflows that empower readouts on a single-cell level. Alternative methods are, for instance, based

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

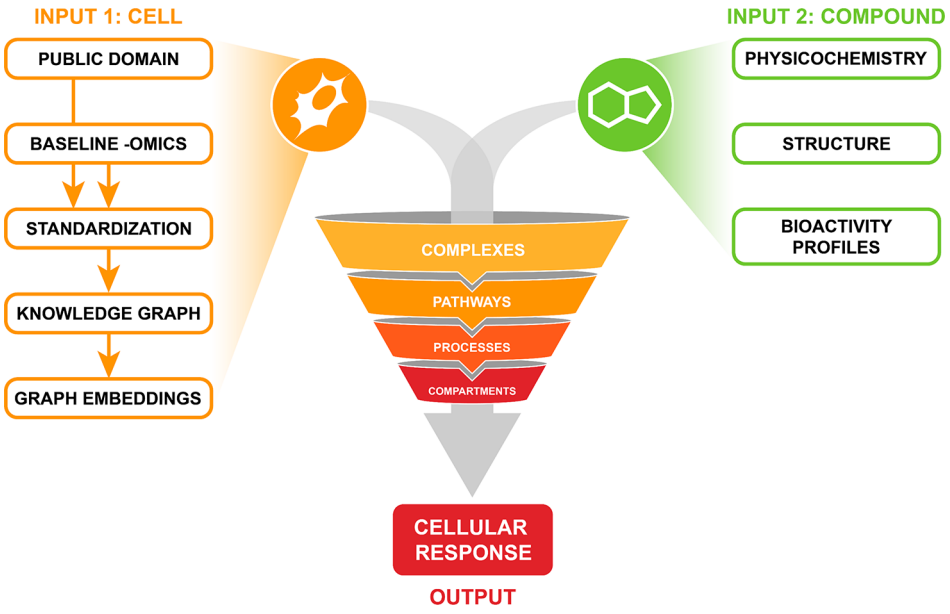

Figure 3. Preparing phenotypic screening data for AI/ML. AI/ML models enable the analysis of phenotypic screening data by translating cell- and compound-based features (input) into cellular response (output).

on screening approaches with a single-cell transcriptomic readout.[170] Segregation of different cell types/states in microscopy images can be conducted by staining,[171] by in situ sequencing,[172] or by segregation based on morphological features which is often enabled by AI/ML.[173,174]

One strategy to introduce diversity is to conduct phenotypic screens not in homogeneous populations of cancer cell lines, but instead in a cellular background with more physiological relevance. To that end, several examples have documented the power of chemical screens in organoids[175,176] or directly in patient material.[177−179] Another avenue to introduce heterogeneity is to genetically engineer cells in a pooled fashion. To identify novel degraders, a selection of several E3s could for instance be inactivated via pooled sgRNA transduction (one E3 ligase knockout per cell). The (degrader) drug-induced phenotype of interest, for instance cell differentiation, could thus be quantified as a function of E3 ligase activity. In other words, this approach would allow the identification of drugs that depend on a particular E3 to induce a phenotypic/morphologic change on a cellular level. Another strategy to introduce complexity is to endogenously tag libraries of POIs, for instance with a fluorescent protein or a self-labeling tag.[180,181] Again, this would enable quantification of drug effects on abundance or localization of multiple POIs in a pooled format. Identification of the individual clones can again be performed retrospectively, either by morphologic features specific to a particular clone (EGFP-POI localization and intensity) or via in situ sequencing. Recognition can further be advanced, for instance via weakly supervised deep learning.

3.3. Maximizing the Depth of Downstream Analysis. In addition to increasing the diversity of the screened cell populations, another strategy to exploit phenotypic screening is to maximize the depth of the downstream analysis. Linked to advanced data integration techniques, phenotypic readouts can provide functional insights about the compound, highlighting engaged cellular processes (e.g., degradation pathways) and dominant mechanisms of action. For example, the CDK inhibitor CR8 was identified as a CycK MGD by correlating gene expression levels of the E3 ligase adaptor DDB1 with cell sensitivity data.[58] AI/ML may unveil more complex connections between molecular and phenotypic events. Ideker and co-workers showed that AI/ML network architectures can be crafted following the logic of a hierarchy (ontology) that reflects the known machinery of cells.[182,183] In this network, inputs are genes, and calculations are forwarded from the more specific terms (e.g., protein complexes) to the more general ones (e.g., organelles or broad terms such as DNA repair), traversing pathways and biological processes. Applied to a genotype-phenotype association study of yeast genetic screening data, it was shown that the AI/ML network could map mutated genes to cellular growth rates.[184] More importantly, inner, specific subsystems were highlighted in the mapping, reflecting the cellular subsystems activated in response to the genetic perturbation. Similar strategies have been applied in the context of prostate cancer clinical data.[185] It is expected that more informative readouts, such as the aforementioned high-content microscopy data, will allow this kind of "visible networks" to interpret cellular responses beyond growth and proliferation. In turn, this will provide more detailed insights about the processes that connect a small-molecule perturbation to an observed phenotype. In this context, identification of degraders could be enabled by searching for compounds that trigger a cellular response that is functionally linked to the ubiquitin-proteasome system or any other protein degradation pathway. Indeed, PROTAC phenotypic signatures have already been identified through cell morphological profiling.[186]

3.4. Preparing Data for AI/ML-Enabled Analysis of Phenotypic Screening Data. At a high level, an AI/ML model based on phenotypic screening data will have two inputs (e.g., a cell and a small-molecule) and one output (i.e., a cellular response) (Figure 3). Pharmacogenomics cancer cell-line panels fit into this scheme. Cancer cell-line collections are deeply annotated, offering a wealth of data to the "cell" branch of the input, including mutational, gene expression, protein abundance, and epigenetic profiles, among others.[187,188] All of these -omics profiles express the baseline status of the cell, and are thus preperturbation data that can be passed to the AI/ML model either independently or in multiplexed form.[189−191]

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

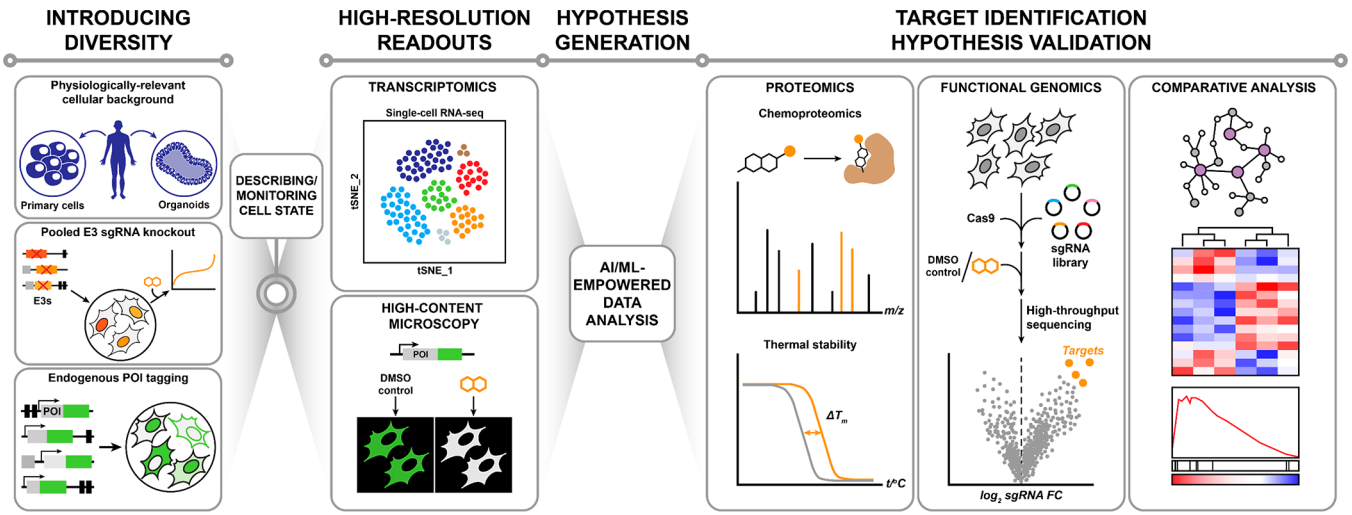

Figure 4. Schematic outline for advanced phenotypic screens. We envision that the innovation of the phenotypic discovery of MGDs will emerge from integrating changes in experimental screening approaches (through diversifying screened cell populations, implementing high-resolution single-cell readouts and further development of compatible target identification methods) with AI/ML-driven data interpretation.

Likewise, a broad set of traditional and modern small-molecule descriptors are able to capture physicochemical properties, 2D/3D structural information, and even *a priori* knowledge of compound bioactivities,[192,193] which means that, on the "small-molecule" branch of the input, there is also a large amount of information available. Converting raw -omics and molecular data into an AI/ML-friendly form is a challenging task due to the high dimensionality of the data sets, biases, and artifacts specific to the -omics technique, and inconsistencies in units and identifiers. The challenges are exacerbated when input data are collected from multiple sources in the public domain. To ease this task, large biological knowledge graphs have been suggested as an intuitive framework for data integration,[194−196] followed by graph embedding techniques capable of expressing the context of each entity (e.g., a compound or a gene) in a dense vectorial format. In brief, vector "embeddings" capture the interactions and statistical correlations found in the knowledge graph, such that highly connected nodes will have similar vector embeddings. With this procedure, heterogeneous large-scale data sets (e.g., compound-protein or gene-cell associations) can be formatted in a unified manner, and can be naturally compared or merged in multilayered analyses.[197,198] It has been shown that embedding techniques can reduce the dimensionality of the input data by several orders of magnitude without affecting the predictive capacity of downstream AI/ML models.[196] Moreover, the complexity of the downstream analysis is reduced substantially—in a recent analysis of a deubiquitinating enzymes (DUBs) genetic screen, DUB inhibitors could be identified based on simple similarity measures between embedded public data of drugs and the DUBs of interest.[193]

3.5. Processing Rich Phenotypic Readouts. Despite the rich annotation of small molecules and unperturbed cell lines, most pharmacogenomics panels have a very simple readout, typically a growth inhibition measure. This is reflected as a single-valued data point in the output branch of the AI/ML model. Even with such a simple readout, genes and pathways that determine drug response can be identified,[199,200] and AI/ML models can be pretrained and then fine-tuned with organoids or patient data to achieve clinically relevant predictions.[201] We anticipate that these notions will be adapted to the TPD field. By analyzing growth inhibition profiles across cell lines, it was for instance found that the drug efflux pump MDR1 confers resistance to PROTACs.[202]

Phenotypic screens are increasingly delivering richer outputs, including high-content microscopy readouts and postperturbation transcriptomics measures such as bulk- or single-cell RNA-seq (Figure 4). The goal of these data-rich readouts is to capture in more detail the cellular response triggered by a chemical compound. For example, observed changes in cellular morphology can help categorize compounds in broad mechanistic categories (e.g., alteration of intracellular metabolism, DNA transport, etc.)[203] and differential gene expression profiles analyzed with conventional bioinformatics techniques can reveal dysregulated cellular pathways.[204] Of note, the LINCS consortium, and in particular the L1000 Connectivity Map resource, provides a catalog of postperturbation transcriptomics profiles for thousands of compounds, offering a global map of the phenotypic response landscape of the medicinal chemistry space.[205] The common way of exploiting this resource is through "connectivity" analysis, i.e. by comparing gene expression signatures of interest to a reference set of signatures of well-annotated compounds.[206] It has been shown that compounds eliciting similar transcriptional signatures tend to share mechanisms of action.[170,207−209] A recognized limitation of the approach is the confounding effect of transcriptional patterns attributable to the cell line, which complicates the extraction of compound-specific effects. Advanced transcriptional similarity metrics derived through AI/ML approaches can help ameliorate this bias.[210]

In addition, the LINCS consortium is performing genetic shRNA, overexpression, and CRISPR screens. Mechanisms of action can be revealed by phenocopying compound transcriptional readouts and these genetic profiles. In this case, similarity measures can be confounded by intrinsic differences between small-molecule and genetic perturbations, which further justifies the need for AI/ML "metric learning" approaches. For the specific purpose of degrader discovery, these differences are less essential as pharmacologic induction of POI degradation often phenocopies the genetic perturbation via shRNA or sgRNA. Recent genome-wide CRISPR screens with a single cell RNA-seq readout are providing vast

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

collections of transcriptional phenotypes,[211] with a clearly identifiable protein degradation cluster that can be used as a reference for small-molecule degrader discovery. Likewise, comparison of high-content imaging outcomes (either in their raw image forms or as extracted geometric features) can be framed as a "metric learning" problem for which AI/ML is ideal. Recent efforts to map the phenotypic landscape of genetic perturbations[212,213] will provide a reference data set of imaging data to which compound phenotypic readouts can be matched and compared.

While similarity-based interpretation of transcriptomics and imaging profiles is likely to remain the by-default analytical framework for phenotypic screens, recent AI/ML-based approaches point toward standalone or direct mechanistic interpretation of phenotypic readouts. In particular, functional ontologies have been mapped to microscopy images,[214] and mechanistic models of transcriptional readouts are becoming established tools.[215] In summary, we believe that the necessary computational elements are already in-place for maximized, in-depth interpretation of phenotypic readouts. We anticipate that future degraders will frequently be discovered by phenotypic assays able to illuminate corners of the biological space that have not yet been explored with small-molecule perturbations.

3.6. Associated Requirements for Target Identification Approaches. Advanced phenotypic assays will also come with a need for experimental target identification (ID) approaches. In accordance with the increased resolution of phenotypic screens, target ID will similarly need to be compatible with single-cell formats. These might involve single-cell proteomics profiling,[216−218] spatial metabolomics,[219,220] or CRISPR/Cas9 screens with a scRNAseq readout.[211,221−223] Conceptually, we differentiate between target ID methods that aim to (i) identify the drug-bound protein target, strategies that (ii) functionally assign effector proteins to drug action, and methods that (iii) holistically describe drug impact in an unbiased manner and aim to develop a mechanistic hypothesis based on comparative analysis with existing, large scale databases (as outlined in section 3.5) (Figure 4).

Approaches to identify the target protein for a small molecule are, largely speaking, dominated by proteomics. Aforementioned chemoproteomics approaches are the gold-standard to identify protein target(s) bound by a given small-molecule hit of interest. A prototypic example is the identification of CRBN as the protein target of thalidomide, which was enabled by a chemoproteomics strategy that employed a tethered thalidomide analogue.[48] A potential downside of this approach is the necessity to chemically derivatize the compound of interest to allow tethering for ensuing affinity enrichments. In most cases, attachment points for an immobilization strategy might not be obvious. In other cases, the mode of molecular recognition of a compound might simply not allow any major derivatization. In such scenarios, orthogonal proteomics applications have emerged that are based on the notion that ligand binding (i) thermally stabilizes the target protein or (ii) protects a protein from proteolysis. Thermal stabilization (i) is leveraged in thermal proteomics profiling (TPP), an approach that has been employed to, for instance, identify the ferrochelatase FECH as a frequent off-target of kinase inhibitors,[224] or to uncover leucine aminopeptidase 3 as an off-target of several HDAC inhibitors.[225] The fact that ligand binding can protect from proteolysis forms the foundation of the limited proteolysis approach (LiP). Among others, this approach has identified the mechanism of action of a fungicide as a kinase inhibitor.[121] Particularly focusing on target identification for small-molecule degraders, proteomics has been key to identifying degraded targets.[226] In addition to global expression proteomics that assess changes in protein abundance after drug treatment, other approaches enable a more in-depth analysis of drug impact on nascent and mature proteins which allows approximation of drug effects on protein stability.[224,227,228] As of now, none of these target-ID focused proteomics approaches have been shown to be compatible with single-cell analysis. As a consequence, we lack the resolution to describe small-molecule interactomes or drug impact on the proteome at a level of granularity that matches transcriptomics approaches. However, recent traction in the field of single-cell proteomics promises to overcome these hurdles in the near future.[229−231] Target identification and elucidation of the mechanism of action of degraders via functional genomics was initially mostly driven by pooled CRISPR/Cas9 screens with a positive selection readout based on cellular fitness ("drug resistance screens").[168,232,233] Future readouts will increasingly focus on readouts that address POI abundance, or states of posttranslational POI modifications, such as (poly-) ubiquitination.[234,235] Of note, a recent breakthrough in high-speed fluorescence image-enabled cell sorting now opens up possibilities to functionally assess effects of gene knockouts on POI localization.[236] For the purpose of drug target identification, this strategy can be used to map all genes that are functionally required for a drug-induced change in POI localization such as nuclear export of a transcription factor. On another level, future target identification approaches will aim to increase the resolution of functional genomics from the protein level to the level of individual domains and amino acids. To this goal, base-editor screens, deep mutational scanning (DMS), and CRISPR tiling screens can play an important role. For instance, both DMS and CRISPR tiling have recently highlighted hotspots in ligase and POI that are functionally relevant for degrader activity.[237,238]

4. ENABLING PROSPECTIVE DEGRADER DISCOVERY

4.1. Leveraging Low-Affinity Protein−Protein Interactions. In parallel to phenotypic screening, rational design of small-molecule degraders is likely to be enabled by recent progress in AI/ML. As explained in section 2.3, *de novo* design and hit-to-lead optimization are possible through generative models, which are constrained or guided by a given task. In TPD, the task involves the assembly of a ternary complex consisting of the POI, the E3 ligase, and the PROTAC or MGD.

It has been suggested that baseline compatibility between the POI and the E3 is a necessary requirement for the formation of a ternary complex.[61,239] This implies that existing docking methods to discover PPIs and resolve interaction interfaces are likely to play a key role in the modeling of TPD mechanisms. However, there are several considerations to be made before attempting out-of-the-box application of docking in this scenario. First, the POI and E3 ligase are generally not cognate interaction partners. As a result, binding interfaces will likely be transient with weak affinity. Accurate modeling of low-affinity interactions is a recognized challenge for docking methods.[240] Moreover, widely used and computationally efficient template-based docking methods are probably not applicable in TPD, since aid from structural alignment to

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

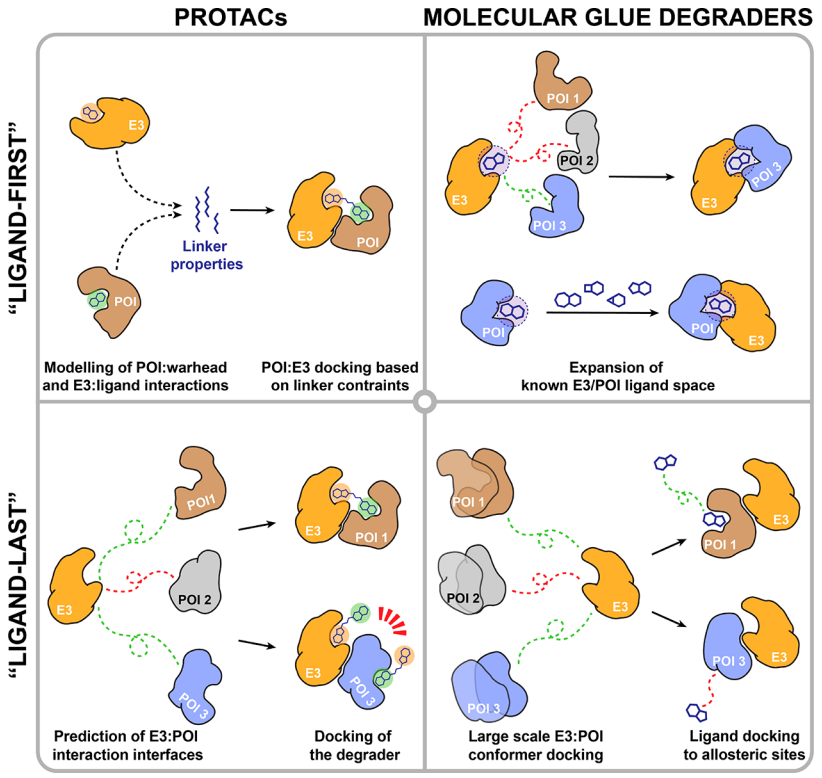

Figure 5. Computational approaches to prospective degrader discovery.

known homologues may not be available.[241] In general, ab initio docking methods will be more appropriate. One advantage of ab initio docking is that it allows for discriminating cognate and noninteracting protein pairs without the need for pre-existing evidence of a PPI. Another advantage is that it capitalizes primarily on physicochemical and shape complementary calculations, which does not impose constraints on known interologs. A major disadvantage is that the search space of docking poses can be unattainable computationally. For this reason, strong assumptions such as rigidity of the monomers,[242] or stepwise pipelines starting with the identification of putative binding sites in each monomer independently,[243] are often necessary to increase the feasibility of the approach.

Ultimately, in the context of TPD, the setup of a large-scale computational prediction would consist, on one side, of structures of the E3 ligases and, on the other side, of a set of POIs to be matched to the ligases. The first task would be to predict whether the POI-E3 PPI can potentially occur, and the second would be to provide a structural model for the interaction. Ab initio docking may naturally encompass both tasks, as demonstrated by recent approaches focused on PROTAC discovery.[244] Additionally, data-driven methods may help discard unfeasible POI-E3 pairs early in the pipeline. For example, methods to assess the PROTACtability of POIs have been suggested,[245] prioritizing over a thousand proteins with potential for becoming PROTAC targets. More generally, many PPI prediction methods exist, most of them based on protein sequences.[246] The current state-of-the-art in sequence-based modeling is the use of AI/ML embeddings pretrained on multimillion-scale protein collections.[118] These embeddings can capture evolutionary traits of the proteins, as well as functional and structural characteristics.[119] As such, contrary to more complicated data structures such as phylogenetic trees or sequence alignments, they can be used directly as numerical inputs for predictive tasks, and to uncover relationships between proteins through straightforward distance calculations. Moreover, at least two follow-up applications of AlphaFold2 suggest that AI/ML may play an important role in the structural modeling of POI-E3 pairs. First, PPIs have already been modeled with small adaptations to the original algorithm.[247,248] Second, protein design models have shown that new folds can be successfully predicted beyond the structural landscape explored by nature.[249] Applied to POI-E3 ligase pairs, this suggests that existence of coevolving residue pairs at the PPI interface may not be a requirement for accurate modeling of structural contacts.

4.2. Inclusion of the Small-Molecule Degrader Early in the Modeling Pipeline. One obvious limitation of considering the POI-E3 ligase pair in the absence of the degrader is the fact that structural conformations induced or stabilized by the ligand are not taken into account. While structural modeling with apo (unbound) forms of the POI and E3 monomers may be sufficient in some ternary PROTAC complexes, it is generally recognized that the ligand plays an important role in determining the PPI interface, either by reshaping it to increase affinity (as it happens in every MGD) or by imposing occupancy and distance constraints (as in the case of PROTACs). Thus, several TPD ternary complex modeling approaches favor the use of holo (bound) forms. Indeed, the holo form of CRBN was used in a large-scale docking run to predict C2H2 zinc-finger degrons.[51] Similarly, the computational pipeline PRosettaC starts by modeling POI-warhead and E3-recruitment ligand interactions, and then exploits the fact that PROTAC linkers have length constraints to drastically reduce the docking search space (Figure 5).[71] In this line, efficient optimization techniques have been proposed to sample the space of ternary complex candidates using

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

composite scores that take into account both PPI docking results and a PROTAC score that measures the feasibility of accommodating the ligand given the position of the binding pockets.[250] Additional biological constraints can be incorporated to further restrict the search space, such as the location of the lysine in the POI for ubiquitination, or of known residue mutations that modulate complex formation. Overall, though, given the abundance of rotatable bonds in the small-molecule degrader, and the importance of protein backbone flexibility in the coordinated formation of PROTAC ternary complexes, modeling these assemblies is likely to remain a computationally intensive task despite the constraints imposed by PROTAC-bridged complexes. The small number of experimentally resolved ternary complexes further complicates benchmarking of methods.

4.3. Toward Computational Pipelines for MGD Discovery. All of the challenges above are enhanced in the case of MGDs. Currently, there is no well-established computational framework to predict and evaluate compound-induced PPIs.[251] Therefore, the discussion presented below refers to potential, envisaged approaches, rather than a compendium of validated methods.

As a result of the scarcity of rational design guidelines, serendipity plays an important role in the discovery of MGDs.[252] Progressively, the chemical characteristics that make a good MGD are being identified, such as addition of solvent-facing functional groups in already known binders.[25,26,253] In addition, systematic screening efforts have revealed high glue hit rates in drug-like collections,[59] suggesting that glue-like traits may be more abundant than previously anticipated. Likewise, it has been shown that proteins have an intrinsic, evolved tendency to form complexes,[254] with random mutations easily prompting the emergence of new interfaces. Thus, a suggested way forward for the design of MGDs is to elaborate derivatives of known ligands located inside or in the vicinity of a PPI interface, with the goal of contributing additional hotspots to the interface and increasing affinity for new protein interactors (Figure 5). For example, thalidomide and other IMiD-inspired derivatives have been used to design focused libraries of putative MGDs. These molecules are expected to bind the E3 CRBN ligase but are diverse in their capacity to recruit POIs.[255]

Similar design principles could be applied to ligands bound in the POI side, instead of the E3 ligase side. In this case, drug design notions borrowed from the discovery of PPI modulators could be implemented.[256,257] In particular, detection of hotspots in flat hydrophobic (interface-like) regions of proteins, as well as fragment-based virtual screening methods, will provide initial hints to accommodate the seed ligand in the POI structure (Figure 5). Evolving these seed ligands into glues that recruit E3 ligases could be done by chemical library expansion strategies, or by rational modification of the solvent-exposed region of the seed molecule. Here, the goal would be to craft additional synthetic hotspots within the weak interaction interface, so that the surface can be recognized with sufficient affinity by the E3 ligase. Although no dedicated methods exist to achieve this goal, systematic analysis of residue−residue contacts observed in known PPIs, and the study of degrons,[258] may offer insights into the necessary pharmacophoric features to be met. Recently, an optimization cycle guided by docking was able to evolve a promiscuous ligand of 14−3−3 proteins to a selective PPI inhibitor.[259]

Finally, MGDs may also function from an allosteric site.[260] In this case, the small molecule does not participate directly in the interaction. Rather, it exposes or modifies a PPI interface by binding to a distant pocket that stabilizes a (non-native) conformation of the protein. Allostery has been suggested as a promising approach to modulating PPIs,[261] since allosteric sites are more ligandable than PPI interfaces, in general. Key to discovering allosteric sites is to consider proteins as ensembles of structures, capturing feasible, but transient, conformational states.[262,263] Translated to the problem of MGD discovery, and focusing on the POI, this involves identifying alternative conformations that reveal interesting PPI interfaces, followed by detection of allosteric cavities that would stabilize the given conformation in the presence of a ligand (Figure 5). While allostery may be a more far-fetched strategy to designing MGDs, there are several advantages in the approach that make it attractive from a computational viewpoint. First, the calculation can be thought of as a two-step procedure where the first step is independent of the small-molecule degrader. Given an E3 ligase, a large-scale docking run could be devised to screen each conformer of each POI independently. In a second step, satisfactory poses could be inspected to identify allosteric binding pockets and dock a ligand to those sites. While an approach like this one may be computationally intensive, it is easy to parallelize. Moreover, databases of structural ensembles of proteins exist[264] and, even though in its by-default form AlphaFold2 was not optimized to sample conformations beyond the native structure,[265] follow-up studies suggest that the tool has potential for generation of biologically relevant ensembles,[266] although extensive validation beyond canonical examples and benchmarking against molecular dynamics simulations are still lacking. Systematic annotation of putative allosteric binding cavities is also possible, and has already been attempted on a proteome-wide scale.[267]

Overall, computational methods to enable prospective and rational design of degraders are still scarce. The relatively small number of known PROTACs and MGDs enforces that design is done primarily with a structure-based approach, which is particularly challenging due the complexity of assembling a ternary complex *de novo*. Supervised AI/ML methods are likely to become more relevant as more instances of small-molecule degraders are discovered. In this line, the aforementioned screening assays, as well as scalable methods to discover drug-induced PPIs such as Y2H-based/related assays,[268] may generate the necessary data to include QSAR modeling in the loop of design. Recent modeling attempts have been done to predict half maximal (DC_{50}) and maximal degradation (D_{max}) concentrations of novel PROTACs.[269] In the future, protein−ligand and PPI prediction models will have to be evolved to quantify their transient nature and eventually be incorporated in kinetic models for ternary complex formation.[270] Despite its relevance to resolve complete and incomplete degradation behaviors, this remains a long-term goal.

5. OUTLOOK

In this perspective, we have attempted to illustrate the status quo of how -omics technologies interface with AI/ML in order to advance research in TPD. We admittedly cover a lot of ground, aiming to review the relevance for proteomics-empowered, as well as AI/ML-enabled ligand discovery, and how they will synergize and complement each other. We have

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

reflected on the successful history of phenotypic screens in the identification of MGDs and tried to illustrate how -omics and advanced imaging platforms will further diversify these approaches to enable researchers to unlock novel target classes. Importantly, we anticipate that AI/ML will play a key role in maximizing the actionable information content of these screens, possibly allowing the deduction of generalizable rules. Lastly, we aimed to reflect on the big challenge of prospective degrader design. While this is becoming a reality for the design of PROTACs, the *de novo* design of MGDs and the discussed approaches reflect, at this point, arguably rather a wish list that will need to be put into practice over the next couple of years. Success of these efforts will, among other factors, determine the overall influence of AI/ML for the TPD field.

What are other success factors? AI/ML is becoming a core component of drug discovery. Arguably, though, adoption of AI/ML methods in this field has been relatively slow and often received with skepticism, compared to other industrial applications.[271] One reason for this is the fact that, as standalone tools, most AI/ML methods do not offer sufficient evidence to support their predictions. Rather, they are presented as "black boxes" that ingest an input (e.g., a molecule structure) and produce an output (e.g., a binding affinity prediction), without providing physical or mechanistic insights. Making more transparent AI/ML models, for instance by highlighting the structural moieties that determine a binding event, has been identified as a key factor to gain trust in AI/ML.[272] Accordingly, "explainable" AI/ML benchmarks are flourishing, aimed at finding substructures in a molecule that determine bioactivity, in agreement with medicinal chemistry expertise.[273] Modern algorithms such as "transformer models", which have found tremendous success in the field of natural language processing and computer vision, have been applied to identify relevant atoms in a ligand or to pinpoint important residues in a protein binding site.[274] Similarly, statistical methods to quantify contribution of certain chemical features to the predicted outcome have been suggested, borrowing concepts from "game theory".[275] In addition, current efforts are focused on identifying "causal" links between the input and the output of an AI/ML model. This often reveals punctual modifications of the input that cause abrupt changes in the output (a.k.a. "counterfactuals"),[276] which should aid the rational design of small-molecules and the interpretation of -omics experiments, where many statistical correlations are discovered but only a few may have a mechanistic impact on the output.

Another limitation of most AI/ML models is that they do not assign a confidence interval to the predictions.[277,278] Naturally, AI/ML models will be more confident if the input small molecule belongs to the same chemical space that was used at training time, while predictions will be less reliable if the input has completely new characteristics previously unseen by the model. The same applies to e.g. transcriptomics profiles of cell lines.[279] Confidence estimation has been shown to be crucial to select true hits in virtual screening experiments, where a large proportion of molecules can fall outside the applicability domain of the AI/ML model. Indeed, in an AI/ML study based on kinase profiling data, considering confidence intervals remarkably increased the relevance of selected virtual hits.[280] Likewise, AlphaFold2 provides a per-residue confidence measure that allows researchers to judge the quality of the predicted structure.[281]

The rapid adoption of AlphaFold2 can only be explained by unprecedented postpublication efforts to make the technology accessible to the broad scientific community. In fact, precalculated structures for 200 million protein sequences are already available online,[282] which effectively removes any technical barrier to adoption. Unfortunately, though, the majority of existing AI/ML models remain in specialized journals, with poorly maintained code and variable rigor in validation.[283] Initiatives like the DREAM Challenges,[284] MoleculeNet,[110] and Therapeutic Data Commons[285] offer benchmark data sets across a wide spectrum of prediction tasks for pharmacology. With the emergence of user-friendly, readily available repositories of AI/ML models, it is likely that execution of AI/ML models will become routine practice in medicinal chemistry and chemical biology laboratories.

In conclusion, we anticipate a key role of AI/ML tools in answering central challenges in TPD research. In our opinion, the unique potential of AI/ML in this interdisciplinary context arises from three aspects. First, its capacity to learn numerical representations (embeddings) of small-molecules, proteins, and cellular phenotypes, which can then be unified in a single analytical framework. Second, its ability to decode these embeddings, yielding scalable generative models that can aid rational design. Third, its scalability in a virtual screening framework, holding promise for systematic discovery of hits across multiple proteins, binding sites, and conformations. Lowering the aforementioned barriers to adoption of AI/ML will be key to ensure an efficient wet-lab/dry-lab cycle and thus an important contribution of AI/ML in the progress of TPD and other proximity-inducing modalities.

AUTHOR INFORMATION

Corresponding Authors

Miquel Duran-Frigola − *CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria; Ersilia Open Source Initiative, CB1 3DE Cambridge, United Kingdom*; Email: miquel@ersilia.io

Georg E. Winter − *CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria*; orcid.org/0000-0001-6606-1437; Email: gwinter@cemm.oeaw.ac.at

Author

Marko Cigler − *CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.2c11098

Funding

The Winter lab is further supported by the Austrian Academy of Sciences, and by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement 851478), as well as by funding from the Austrian Science Fund (FWF, Projects P32125, P31690, and P7909).

Notes

The authors declare the following competing financial interest(s): G.E.W. is scientific founder and shareholder of Proxygen and Solgate. The Winter lab received research funding from Pfizer.

■ ACKNOWLEDGMENTS

We thank all members of the Winter lab for helpful scientific discussions.

■ REFERENCES

(1) Pettersson, M.; Crews, C. M. PROteolysis TArgeting Chimeras (PROTACs) - Past, present and future. *Drug Discov Today Technol.* **2019**, *31*, 15−27.
(2) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov* **2002**, *1* (9), 727−30.
(3) Oprea, T. I.; Bologa, C. G.; Brunak, S.; Campbell, A.; Gan, G. N.; Gaulton, A.; Gomez, S. M.; Guha, R.; Hersey, A.; Holmes, J.; Jadhav, A.; Jensen, L. J.; Johnson, G. L.; Karlson, A.; Leach, A. R.; Ma'ayan, A.; Malovannaya, A.; Mani, S.; Mathias, S. L.; McManus, M. T.; Meehan, T. F.; von Mering, C.; Muthas, D.; Nguyen, D. T.; Overington, J. P.; Papadatos, G.; Qin, J.; Reich, C.; Roth, B. L.; Schurer, S. C.; Simeonov, A.; Sklar, L. A.; Southall, N.; Tomita, S.; Tudose, I.; Ursu, O.; Vidovic, D.; Waller, A.; Westergaard, D.; Yang, J. J.; Zahoranszky-Kohalmi, G. Unexplored therapeutic opportunities in the human genome. *Nat. Rev. Drug Discov* **2018**, *17* (5), 317−332.
(4) Bushweller, J. H. Targeting transcription factors in cancer - from undruggable to reality. *Nat. Rev. Cancer* **2019**, *19* (11), 611−624.
(5) Biesaga, M.; Frigole-Vivas, M.; Salvatella, X. Intrinsically disordered proteins and biomolecular condensates as drug targets. *Curr. Opin Chem. Biol.* **2021**, *62*, 90−100.
(6) Klaeger, S.; Heinzlmeir, S.; Wilhelm, M.; Polzer, H.; Vick, B.; Koenig, P. A.; Reinecke, M.; Ruprecht, B.; Petzoldt, S.; Meng, C.; Zecha, J.; Reiter, K.; Qiao, H.; Helm, D.; Koch, H.; Schoof, M.; Canevari, G.; Casale, E.; Depaolini, S. R.; Feuchtinger, A.; Wu, Z.; Schmidt, T.; Rueckert, L.; Becker, W.; Huenges, J.; Garz, A. K.; Gohlke, B. O.; Zolg, D. P.; Kayser, G.; Vooder, T.; Preissner, R.; Hahne, H.; Tonisson, N.; Kramer, K.; Gotze, K.; Bassermann, F.; Schlegl, J.; Ehrlich, H. C.; Aiche, S.; Walch, A.; Greif, P. A.; Schneider, S.; Felder, E. R.; Ruland, J.; Medard, G.; Jeremias, I.; Spiekermann, K.; Kuster, B. The target landscape of clinical kinase drugs. *Science* **2017**, *358* (6367), eaan4368.
(7) Ferguson, F. M.; Gray, N. S. Kinase inhibitors: the road ahead. *Nat. Rev. Drug Discov* **2018**, *17* (5), 353−377.
(8) Wallasch, C.; Weiss, F. U.; Niederfellner, G.; Jallal, B.; Issing, W.; Ullrich, A. Heregulin-dependent regulation of HER2/neu oncogenic signaling by heterodimerization with HER3. *EMBO J.* **1995**, *14* (17), 4267−75.
(9) Yarden, Y.; Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2001**, *2* (2), 127−37.
(10) Littlefield, P.; Liu, L.; Mysore, V.; Shan, Y.; Shaw, D. E.; Jura, N. Structural analysis of the EGFR/HER3 heterodimer reveals the molecular basis for activating HER3 mutations. *Sci. Signal* **2014**, *7* (354), ra114.
(11) Xie, T.; Lim, S. M.; Westover, K. D.; Dodge, M. E.; Ercan, D.; Ficarro, S. B.; Udayakumar, D.; Gurbani, D.; Tae, H. S.; Riddle, S. M.; Sim, T.; Marto, J. A.; Janne, P. A.; Crews, C. M.; Gray, N. S. Pharmacological targeting of the pseudokinase Her3. *Nat. Chem. Biol.* **2014**, *10* (12), 1006−12.
(12) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. Selective inhibition of BET bromodomains. *Nature* **2010**, *468* (7327), 1067−73.
(13) Hammitzsch, A.; Tallant, C.; Fedorov, O.; O'Mahony, A.; Brennan, P. E.; Hay, D. A.; Martinez, F. O.; Al-Mossawi, M. H.; de Wit, J.; Vecellio, M.; Wells, C.; Wordsworth, P.; Muller, S.; Knapp, S.; Bowness, P. CBP30, a selective CBP/p300 bromodomain inhibitor, suppresses human Th17 responses. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (34), 10768−73.
(14) Picaud, S.; Fedorov, O.; Thanasopoulou, A.; Leonards, K.; Jones, K.; Meier, J.; Olzscha, H.; Monteiro, O.; Martin, S.; Philpott, M.; Tumber, A.; Filippakopoulos, P.; Yapp, C.; Wells, C.; Che, K. H.; Bannister, A.; Robson, S.; Kumar, U.; Parr, N.; Lee, K.; Lugo, D.; Jeffrey, P.; Taylor, S.; Vecellio, M. L.; Bountra, C.; Brennan, P. E.; O'Mahony, A.; Velichko, S.; Muller, S.; Hay, D.; Daniels, D. L.; Urh, M.; La Thangue, N. B.; Kouzarides, T.; Prinjha, R.; Schwaller, J.; Knapp, S. Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. *Cancer Res.* **2015**, *75* (23), 5106−5119.
(15) Taylor, A. M.; Cote, A.; Hewitt, M. C.; Pastor, R.; Leblanc, Y.; Nasveschuk, C. G.; Romero, F. A.; Crawford, T. D.; Cantone, N.; Jayaram, H.; Setser, J.; Murray, J.; Beresini, M. H.; de Leon Boenig, G.; Chen, Z.; Conery, A. R.; Cummings, R. T.; Dakin, L. A.; Flynn, E. M.; Huang, O. W.; Kaufman, S.; Keller, P. J.; Kiefer, J. R.; Lai, T.; Li, Y.; Liao, J.; Liu, W.; Lu, H.; Pardo, E.; Tsui, V.; Wang, J.; Wang, Y.; Xu, Z.; Yan, F.; Yu, D.; Zawadzke, L.; Zhu, X.; Zhu, X.; Sims, R. J., 3rd; Cochran, A. G.; Bellon, S.; Audia, J. E.; Magnuson, S.; Albrecht, B. K. Fragment-Based Discovery of a Selective and Cell-Active Benzodiazepinone CBP/EP300 Bromodomain Inhibitor (CPI-637). *ACS Med. Chem. Lett.* **2016**, *7* (5), 531−6.
(16) Crawford, T. D.; Romero, F. A.; Lai, K. W.; Tsui, V.; Taylor, A. M.; de Leon Boenig, G.; Noland, C. L.; Murray, J.; Ly, J.; Choo, E. F.; Hunsaker, T. L.; Chan, E. W.; Merchant, M.; Kharbanda, S.; Gascoigne, K. E.; Kaufman, S.; Beresini, M. H.; Liao, J.; Liu, W.; Chen, K. X.; Chen, Z.; Conery, A. R.; Cote, A.; Jayaram, H.; Jiang, Y.; Kiefer, J. R.; Kleinheinz, T.; Li, Y.; Maher, J.; Pardo, E.; Poy, F.; Spillane, K. L.; Wang, F.; Wang, J.; Wei, X.; Xu, Z.; Xu, Z.; Yen, I.; Zawadzke, L.; Zhu, X.; Bellon, S.; Cummings, R.; Cochran, A. G.; Albrecht, B. K.; Magnuson, S. Discovery of a Potent and Selective in Vivo Probe (GNE-272) for the Bromodomains of CBP/EP300. *J. Med. Chem.* **2016**, *59* (23), 10549−10563.
(17) Romero, F. A.; Murray, J.; Lai, K. W.; Tsui, V.; Albrecht, B. K.; An, L.; Beresini, M. H.; de Leon Boenig, G.; Bronner, S. M.; Chan, E. W.; Chen, K. X.; Chen, Z.; Choo, E. F.; Clagg, K.; Clark, K.; Crawford, T. D.; Cyr, P.; de Almeida Nagata, D.; Gascoigne, K. E.; Grogan, J. L.; Hatzivassiliou, G.; Huang, W.; Hunsaker, T. L.; Kaufman, S.; Koenig, S. G.; Li, R.; Li, Y.; Liang, X.; Liao, J.; Liu, W.; Ly, J.; Maher, J.; Masui, C.; Merchant, M.; Ran, Y.; Taylor, A. M.; Wai, J.; Wang, F.; Wei, X.; Yu, D.; Zhu, B. Y.; Zhu, X.; Magnuson, S. GNE-781, A Highly Advanced Potent and Selective Bromodomain Inhibitor of Cyclic Adenosine Monophosphate Response Element Binding Protein, Binding Protein (CBP). *J. Med. Chem.* **2017**, *60* (22), 9162−9183.
(18) Bennett, J.; Fedorov, O.; Tallant, C.; Monteiro, O.; Meier, J.; Gamble, V.; Savitsky, P.; Nunez-Alonso, G. A.; Haendler, B.; Rogers, C.; Brennan, P. E.; Muller, S.; Knapp, S. Discovery of a Chemical Tool Inhibitor Targeting the Bromodomains of TRIM24 and BRPF. *J. Med. Chem.* **2016**, *59* (4), 1642−7.
(19) Palmer, W. S.; Poncet-Montange, G.; Liu, G.; Petrocchi, A.; Reyna, N.; Subramanian, G.; Theroff, J.; Yau, A.; Kost-Alimova, M.; Bardenhagen, J. P.; Leo, E.; Shepard, H. E.; Tieu, T. N.; Shi, X.; Zhan, Y.; Zhao, S.; Barton, M. C.; Draetta, G.; Toniatti, C.; Jones, P.; Geck Do, M.; Andersen, J. N. Structure-Guided Design of IACS-9571, a Selective High-Affinity Dual TRIM24-BRPF1 Bromodomain Inhibitor. *J. Med. Chem.* **2016**, *59* (4), 1440−54.
(20) Fedorov, O.; Castex, J.; Tallant, C.; Owen, D. R.; Martin, S.; Aldeghi, M.; Monteiro, O.; Filippakopoulos, P.; Picaud, S.; Trzupek, J. D.; Gerstenberger, B. S.; Bountra, C.; Willmann, D.; Wells, C.; Philpott, M.; Rogers, C.; Biggin, P. C.; Brennan, P. E.; Bunnage, M. E.; Schule, R.; Gunther, T.; Knapp, S.; Muller, S. Selective targeting of the BRG/PB1 bromodomains impairs embryonic and trophoblast stem cell maintenance. *Sci. Adv.* **2015**, *1* (10), e1500723.
(21) Vangamudi, B.; Paul, T. A.; Shah, P. K.; Kost-Alimova, M.; Nottebaum, L.; Shi, X.; Zhan, Y.; Leo, E.; Mahadeshwar, H. S.; Protopopov, A.; Futreal, A.; Tieu, T. N.; Peoples, M.; Heffernan, T. P.; Marszalek, J. R.; Toniatti, C.; Petrocchi, A.; Verhelle, D.; Owen, D. R.; Draetta, G.; Jones, P.; Palmer, W. S.; Sharma, S.; Andersen, J. N. The SMARCA2/4 ATPase Domain Surpasses the Bromodomain as a Drug Target in SWI/SNF-Mutant Cancers: Insights from cDNA

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

Rescue and PFI-3 Inhibitor Studies. *Cancer Res.* **2015**, *75* (18), 3865−3878.
(22) Kiely-Collins, H.; Winter, G. E.; Bernardes, G. J. L. The role of reversible and irreversible covalent chemistry in targeted protein degradation. *Cell Chem. Biol.* **2021**, *28* (7), 952−968.
(23) Boike, L.; Henning, N. J.; Nomura, D. K. Advances in covalent drug discovery. *Nat. Rev. Drug Discov* **2022**, *21* (12), 881−898.
(24) Stanton, B. Z.; Chory, E. J.; Crabtree, G. R. Chemically induced proximity in biology and medicine. *Science* **2018**, *359* (6380), eaao5902.
(25) Gerry, C. J.; Schreiber, S. L. Unifying principles of bifunctional, proximity-inducing small molecules. *Nat. Chem. Biol.* **2020**, *16* (4), 369−378.
(26) Geiger, T. M.; Schäfer, S. C.; Dreizler, J. K.; Walz, M.; Hausch, F. Clues to molecular glues. *Current Research in Chemical Biology* **2022**, *2*, 100018.
(27) Schreiber, S. L. The Rise of Molecular Glues. *Cell* **2021**, *184* (1), 3−9.
(28) Deshaies, R. J. Multispecific drugs herald a new era of biopharmaceutical innovation. *Nature* **2020**, *580* (7803), 329−338.
(29) Banik, S. M.; Pedram, K.; Wisnovsky, S.; Ahn, G.; Riley, N. M.; Bertozzi, C. R. Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature* **2020**, *584* (7820), 291−297.
(30) Ahn, G.; Banik, S. M.; Miller, C. L.; Riley, N. M.; Cochran, J. R.; Bertozzi, C. R. LYTACs that engage the asialoglycoprotein receptor for targeted protein degradation. *Nat. Chem. Biol.* **2021**, *17* (9), 937−946.
(31) Cotton, A. D.; Nguyen, D. P.; Gramespacher, J. A.; Seiple, I. B.; Wells, J. A. Development of Antibody-Based PROTACs for the Degradation of the Cell-Surface Immune Checkpoint Protein PD-L1. *J. Am. Chem. Soc.* **2021**, *143* (2), 593−598.
(32) Siriwardena, S. U.; Munkanatta Godage, D. N. P.; Shoba, V. M.; Lai, S.; Shi, M.; Wu, P.; Chaudhary, S. K.; Schreiber, S. L.; Choudhary, A. Phosphorylation-Inducing Chimeric Small Molecules. *J. Am. Chem. Soc.* **2020**, *142* (33), 14052−14057.
(33) Chen, P. H.; Hu, Z.; An, E.; Okeke, I.; Zheng, S.; Luo, X.; Gong, A.; Jaime-Figueroa, S.; Crews, C. M. Modulation of Phosphoprotein Activity by Phosphorylation Targeting Chimeras (PhosTACs). *ACS Chem. Biol.* **2021**, *16* (12), 2808−2815.
(34) Wang, W. W.; Chen, L. Y.; Wozniak, J. M.; Jadhav, A. M.; Anderson, H.; Malone, T. E.; Parker, C. G. Targeted Protein Acetylation in Cells Using Heterobifunctional Molecules. *J. Am. Chem. Soc.* **2021**, *143* (40), 16700−16708.
(35) Henning, N. J.; Boike, L.; Spradlin, J. N.; Ward, C. C.; Liu, G.; Zhang, E.; Belcher, B. P.; Brittain, S. M.; Hesse, M. J.; Dovala, D.; McGregor, L. M.; Valdez Misiolek, R.; Plasschaert, L. W.; Rowlands, D. J.; Wang, F.; Frank, A. O.; Fuller, D.; Estes, A. R.; Randal, K. L.; Panidapu, A.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Nomura, D. K. Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat. Chem. Biol.* **2022**, *18* (4), 412−421.
(36) Liu, J.; Yu, X.; Chen, H.; Kaniskan, H. U.; Xie, L.; Chen, X.; Jin, J.; Wei, W. TF-DUBTACs Stabilize Tumor Suppressor Transcription Factors. *J. Am. Chem. Soc.* **2022**, *144* (28), 12934−12941.
(37) Costales, M. G.; Matsumoto, Y.; Velagapudi, S. P.; Disney, M. D. Small Molecule Targeted Recruitment of a Nuclease to RNA. *J. Am. Chem. Soc.* **2018**, *140* (22), 6741−6744.
(38) Haniff, H. S.; Tong, Y.; Liu, X.; Chen, J. L.; Suresh, B. M.; Andrews, R. J.; Peterson, J. M.; O'Leary, C. A.; Benhamou, R. I.; Moss, W. N.; Disney, M. D. Targeting the SARS-CoV-2 RNA Genome with Small Molecule Binders and Ribonuclease Targeting Chimera (RIBOTAC) Degraders. *ACS Cent Sci.* **2020**, *6* (10), 1713−1721.
(39) Bush, J. A.; Aikawa, H.; Fuerst, R.; Li, Y.; Ursu, A.; Meyer, S. M.; Benhamou, R. I.; Chen, J. L.; Khan, T.; Wagner-Griffin, S.; Van Meter, M. J.; Tong, Y.; Olafson, H.; McKee, K. K.; Childs-Disney, J. L.; Gendron, T. F.; Zhang, Y.; Coyne, A. N.; Wang, E. T.; Yildirim, I.; Wang, K. W.; Petrucelli, L.; Rothstein, J. D.; Disney, M. D. Ribonuclease recruitment using a small molecule reduced c9ALS/FTD r(G(4)C(2)) repeat expansion in vitro and in vivo ALS models. *Sci. Transl Med.* **2021**, *13* (617), eabd5991.
(40) Takahashi, D.; Moriyama, J.; Nakamura, T.; Miki, E.; Takahashi, E.; Sato, A.; Akaike, T.; Itto-Nakama, K.; Arimoto, H. AUTACs: Cargo-Specific Degraders Using Selective Autophagy. *Mol. Cell* **2019**, *76* (5), 797−810.
(41) Li, Z.; Wang, C.; Wang, Z.; Zhu, C.; Li, J.; Sha, T.; Ma, L.; Gao, C.; Yang, Y.; Sun, Y.; Wang, J.; Sun, X.; Lu, C.; Difiglia, M.; Mei, Y.; Ding, C.; Luo, S.; Dang, Y.; Ding, Y.; Fei, Y.; Lu, B. Allele-selective lowering of mutant HTT protein by HTT-LC3 linker compounds. *Nature* **2019**, *575* (7781), 203−209.
(42) Ji, C. H.; Kim, H. Y.; Lee, M. J.; Heo, A. J.; Park, D. Y.; Lim, S.; Shin, S.; Ganipisetti, S.; Yang, W. S.; Jung, C. A.; Kim, K. Y.; Jeong, E. H.; Park, S. H.; Bin Kim, S.; Lee, S. J.; Na, J. E.; Kang, J. I.; Chi, H. M.; Kim, H. T.; Kim, Y. K.; Kim, B. Y.; Kwon, Y. T. The AUTOTAC chemical biology platform for targeted protein degradation via the autophagy-lysosome system. *Nat. Commun.* **2022**, *13* (1), 904.
(43) Salami, J.; Crews, C. M. Waste disposal-An attractive strategy for cancer therapy. *Science* **2017**, *355* (6330), 1163−1167.
(44) Hanzl, A.; Winter, G. E. Targeted protein degradation: current and future challenges. *Curr. Opin Chem. Biol.* **2020**, *56*, 35−41.
(45) Gechijian, L. N.; Buckley, D. L.; Lawlor, M. A.; Reyes, J. M.; Paulk, J.; Ott, C. J.; Winter, G. E.; Erb, M. A.; Scott, T. G.; Xu, M.; Seo, H. S.; Dhe-Paganon, S.; Kwiatkowski, N. P.; Perry, J. A.; Qi, J.; Gray, N. S.; Bradner, J. E. Functional TRIM24 degrader via conjugation of ineffectual bromodomain and VHL ligands. *Nat. Chem. Biol.* **2018**, *14* (4), 405−412.
(46) Farnaby, W.; Koegl, M.; Roy, M. J.; Whitworth, C.; Diers, E.; Trainor, N.; Zollman, D.; Steurer, S.; Karolyi-Oezguer, J.; Riedmueller, C.; Gmaschitz, T.; Wachter, J.; Dank, C.; Galant, M.; Sharps, B.; Rumpel, K.; Traxler, E.; Gerstberger, T.; Schnitzer, R.; Petermann, O.; Greb, P.; Weinstabl, H.; Bader, G.; Zoephel, A.; Weiss-Puxbaum, A.; Ehrenhofer-Wolfer, K.; Wohrle, S.; Boehmelt, G.; Rinnenthal, J.; Arnhof, H.; Wiechens, N.; Wu, M. Y.; Owen-Hughes, T.; Ettmayer, P.; Pearson, M.; McConnell, D. B.; Ciulli, A. BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nat. Chem. Biol.* **2019**, *15* (7), 672−680.
(47) Vannam, R.; Sayilgan, J.; Ojeda, S.; Karakyriakou, B.; Hu, E.; Kreuzer, J.; Morris, R.; Herrera Lopez, X. I.; Rai, S.; Haas, W.; Lawrence, M.; Ott, C. J. Targeted degradation of the enhancer lysine acetyltransferases CBP and p300. *Cell Chem. Biol.* **2021**, *28* (4), 503−514.
(48) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. Identification of a primary target of thalidomide teratogenicity. *Science* **2010**, *327* (5971), 1345−50.
(49) Lu, G.; Middleton, R. E.; Sun, H.; Naniong, M.; Ott, C. J.; Mitsiades, C. S.; Wong, K. K.; Bradner, J. E.; Kaelin, W. G., Jr. The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **2014**, *343* (6168), 305−9.
(50) Kronke, J.; Udeshi, N. D.; Narla, A.; Grauman, P.; Hurst, S. N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X.; Ciarlo, C.; Hartman, E.; Munshi, N.; Schenone, M.; Schreiber, S. L.; Carr, S. A.; Ebert, B. L. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **2014**, *343* (6168), 301−5.
(51) Sievers, Q. L.; Petzold, G.; Bunker, R. D.; Renneville, A.; Slabicki, M.; Liddicoat, B. J.; Abdulrahman, W.; Mikkelsen, T.; Ebert, B. L.; Thoma, N. H. Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN. *Science* **2018**, *362* (6414), eaat0572.
(52) Petzold, G.; Fischer, E. S.; Thoma, N. H. Structural basis of lenalidomide-induced CK1alpha degradation by the CRL4(CRBN) ubiquitin ligase. *Nature* **2016**, *532* (7597), 127−30.
(53) Matyskiela, M. E.; Lu, G.; Ito, T.; Pagarigan, B.; Lu, C. C.; Miller, K.; Fang, W.; Wang, N. Y.; Nguyen, D.; Houston, J.; Carmel, G.; Tran, T.; Riley, M.; Nosaka, L.; Lander, G. C.; Gaidarova, S.; Xu, S.; Ruchelman, A. L.; Handa, H.; Carmichael, J.; Daniel, T. O.; Cathers, B. E.; Lopez-Girona, A.; Chamberlain, P. P. A novel cereblon modulator recruits GSPT1 to the CRL4(CRBN) ubiquitin ligase. *Nature* **2016**, *535* (7611), 252−7.

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

(54) Han, T.; Goralski, M.; Gaskill, N.; Capota, E.; Kim, J.; Ting, T. C.; Xie, Y.; Williams, N. S.; Nijhawan, D. Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. *Science* **2017**, *356* (6336), eaal3755.
(55) Ting, T. C.; Goralski, M.; Klein, K.; Wang, B.; Kim, J.; Xie, Y.; Nijhawan, D. Aryl Sulfonamides Degrade RBM39 and RBM23 by Recruitment to CRL4-DCAF15. *Cell Rep* **2019**, *29* (6), 1499−1510.e6.
(56) Faust, T. B.; Yoon, H.; Nowak, R. P.; Donovan, K. A.; Li, Z.; Cai, Q.; Eleuteri, N. A.; Zhang, T.; Gray, N. S.; Fischer, E. S. Structural complementarity facilitates E7820-mediated degradation of RBM39 by DCAF15. *Nat. Chem. Biol.* **2020**, *16* (1), 7−14.
(57) Bussiere, D. E.; Xie, L.; Srinivas, H.; Shu, W.; Burke, A.; Be, C.; Zhao, J.; Godbole, A.; King, D.; Karki, R. G.; Hornak, V.; Xu, F.; Cobb, J.; Carte, N.; Frank, A. O.; Frommlet, A.; Graff, P.; Knapp, M.; Fazal, A.; Okram, B.; Jiang, S.; Michellys, P. Y.; Beckwith, R.; Voshol, H.; Wiesmann, C.; Solomon, J. M.; Paulk, J. Structural basis of indisulam-mediated RBM39 recruitment to DCAF15 E3 ligase complex. *Nat. Chem. Biol.* **2020**, *16* (1), 15−23.
(58) Slabicki, M.; Kozicka, Z.; Petzold, G.; Li, Y. D.; Manojkumar, M.; Bunker, R. D.; Donovan, K. A.; Sievers, Q. L.; Koeppel, J.; Suchyta, D.; Sperling, A. S.; Fink, E. C.; Gasser, J. A.; Wang, L. R.; Corsello, S. M.; Sellar, R. S.; Jan, M.; Gillingham, D.; Scholl, C.; Frohling, S.; Golub, T. R.; Fischer, E. S.; Thoma, N. H.; Ebert, B. L. The CDK inhibitor CR8 acts as a molecular glue degrader that depletes cyclin K. *Nature* **2020**, *585* (7824), 293−297.
(59) Mayor-Ruiz, C.; Bauer, S.; Brand, M.; Kozicka, Z.; Siklos, M.; Imrichova, H.; Kaltheuner, I. H.; Hahn, E.; Seiler, K.; Koren, A.; Petzold, G.; Fellner, M.; Bock, C.; Muller, A. C.; Zuber, J.; Geyer, M.; Thoma, N. H.; Kubicek, S.; Winter, G. E. Rational discovery of molecular glue degraders via scalable chemical profiling. *Nat. Chem. Biol.* **2020**, *16* (11), 1199−1207.
(60) Lv, L.; Chen, P.; Cao, L.; Li, Y.; Zeng, Z.; Cui, Y.; Wu, Q.; Li, J.; Wang, J. H.; Dong, M. Q.; Qi, X.; Han, T. Discovery of a molecular glue promoting CDK12-DDB1 interaction to trigger cyclin K degradation. *Elife* **2020**, *9*, e59994.
(61) Cao, S.; Kang, S.; Mao, H.; Yao, J.; Gu, L.; Zheng, N. Defining molecular glues with a dual-nanobody cannabidiol sensor. *Nat. Commun.* **2022**, *13* (1), 815.
(62) Zengerle, M.; Chan, K. H.; Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **2015**, *10* (8), 1770−7.
(63) Gadd, M. S.; Testa, A.; Lucas, X.; Chan, K. H.; Chen, W.; Lamont, D. J.; Zengerle, M.; Ciulli, A. Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat. Chem. Biol.* **2017**, *13* (5), 514−521.
(64) Olson, C. M.; Jiang, B.; Erb, M. A.; Liang, Y.; Doctor, Z. M.; Zhang, Z.; Zhang, T.; Kwiatkowski, N.; Boukhali, M.; Green, J. L.; Haas, W.; Nomanbhoy, T.; Fischer, E. S.; Young, R. A.; Bradner, J. E.; Winter, G. E.; Gray, N. S. Pharmacological perturbation of CDK9 using selective CDK9 inhibition or degradation. *Nat. Chem. Biol.* **2018**, *14* (2), 163−170.
(65) Bondeson, D. P.; Smith, B. E.; Burslem, G. M.; Buhimschi, A. D.; Hines, J.; Jaime-Figueroa, S.; Wang, J.; Hamman, B. D.; Ishchenko, A.; Crews, C. M. Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chem. Biol.* **2018**, *25* (1), 78−87.e5.
(66) Huang, H. T.; Dobrovolsky, D.; Paulk, J.; Yang, G.; Weisberg, E. L.; Doctor, Z. M.; Buckley, D. L.; Cho, J. H.; Ko, E.; Jang, J.; Shi, K.; Choi, H. G.; Griffin, J. D.; Li, Y.; Treon, S. P.; Fischer, E. S.; Bradner, J. E.; Tan, L.; Gray, N. S. A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-kinase Degrader. *Cell Chem. Biol.* **2018**, *25* (1), 88−99.e6.
(67) Brand, M.; Jiang, B.; Bauer, S.; Donovan, K. A.; Liang, Y.; Wang, E. S.; Nowak, R. P.; Yuan, J. C.; Zhang, T.; Kwiatkowski, N.; Muller, A. C.; Fischer, E. S.; Gray, N. S.; Winter, G. E. Homolog-Selective Degradation as a Strategy to Probe the Function of CDK6 in AML. *Cell Chem. Biol.* **2019**, *26* (2), 300−306.e9.
(68) Smith, B. E.; Wang, S. L.; Jaime-Figueroa, S.; Harbin, A.; Wang, J.; Hamman, B. D.; Crews, C. M. Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase. *Nat. Commun.* **2019**, *10* (1), 131.
(69) Donovan, K. A.; Ferguson, F. M.; Bushman, J. W.; Eleuteri, N. A.; Bhunia, D.; Ryu, S.; Tan, L.; Shi, K.; Yue, H.; Liu, X.; Dobrovolsky, D.; Jiang, B.; Wang, J.; Hao, M.; You, I.; Teng, M.; Liang, Y.; Hatcher, J.; Li, Z.; Manz, T. D.; Groendyke, B.; Hu, W.; Nam, Y.; Sengupta, S.; Cho, H.; Shin, I.; Agius, M. P.; Ghobrial, I. M.; Ma, M. W.; Che, J.; Buhrlage, S. J.; Sim, T.; Gray, N. S.; Fischer, E. S. Mapping the Degradable Kinome Provides a Resource for Expedited Degrader Development. *Cell* **2020**, *183* (6), 1714−1731.e10.
(70) Alabi, S.; Jaime-Figueroa, S.; Yao, Z.; Gao, Y.; Hines, J.; Samarasinghe, K. T. G.; Vogt, L.; Rosen, N.; Crews, C. M. Mutant-selective degradation by BRAF-targeting PROTACs. *Nat. Commun.* **2021**, *12* (1), 920.
(71) Zaidman, D.; Prilusky, J.; London, N. PRosettaC: Rosetta Based Modeling of PROTAC Mediated Ternary Complexes. *J. Chem. Inf Model* **2020**, *60* (10), 4894−4903.
(72) Khan, S.; Zhang, X.; Lv, D.; Zhang, Q.; He, Y.; Zhang, P.; Liu, X.; Thummuri, D.; Yuan, Y.; Wiegand, J. S.; Pei, J.; Zhang, W.; Sharma, A.; McCurdy, C. R.; Kuruvilla, V. M.; Baran, N.; Ferrando, A. A.; Kim, Y. M.; Rogojina, A.; Houghton, P. J.; Huang, G.; Hromas, R.; Konopleva, M.; Zheng, G.; Zhou, D. A selective BCL-XL PROTAC degrader achieves safe and potent antitumor activity. *Nat. Med.* **2019**, *25* (12), 1938−1947.
(73) Bartlett, D. W.; Gilbert, A. M. Translational PK-PD for targeted protein degradation. *Chem. Soc. Rev.* **2022**, *51* (9), 3477−3486.
(74) Hanzl, A.; Barone, E.; Bauer, S.; Yue, H.; Nowak, R. P.; Hahn, E.; Pankevich, E. V.; Koren, A.; Kubicek, S.; Fischer, E. S.; Winter, G. E. E3-Specific Degrader Discovery by Dynamic Tracing of Substrate Receptor Abundance. *J. Am. Chem. Soc.* **2023**, *145* (2), 1176−1184.
(75) Owicki, J. C. Fluorescence polarization and anisotropy in high throughput screening: perspectives and primer. *J. Biomol Screen* **2000**, *5* (5), 297−306.
(76) Burke, T. J.; Loniello, K. R.; Beebe, J. A.; Ervin, K. M. Development and application of fluorescence polarization assays in drug discovery. *Comb Chem. High Throughput Screen* **2003**, *6* (3), 183−94.
(77) Frostell, A.; Vinterback, L.; Sjobom, H. Protein-Ligand Interactions Using SPR Systems. *Methods Mol. Biol.* **2013**, *1008*, 139−65.
(78) Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; Lundback, T.; Nordlund, P.; Martinez Molina, D. The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat. Protoc* **2014**, *9* (9), 2100−22.
(79) Henderson, M. J.; Holbert, M. A.; Simeonov, A.; Kallal, L. A. High-Throughput Cellular Thermal Shift Assays in Research and Drug Discovery. *SLAS Discov* **2020**, *25* (2), 137−147.
(80) Gironda-Martinez, A.; Donckele, E. J.; Samain, F.; Neri, D. DNA-Encoded Chemical Libraries: A Comprehensive Review with Succesful Stories and Future Challenges. *ACS Pharmacol Transl Sci.* **2021**, *4* (4), 1265−1279.
(81) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383−414.
(82) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010**, *468* (7325), 790−5.
(83) Bantscheff, M.; Drewes, G. Chemoproteomic approaches to drug target identification and drug profiling. *Bioorg. Med. Chem.* **2012**, *20* (6), 1973−8.
(84) Spradlin, J. N.; Hu, X.; Ward, C. C.; Brittain, S. M.; Jones, M. D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; Olzmann, J. A.; Bussiere, D. E.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; Schirle, M.; Maimone, T. J.; Nomura, D. K. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat. Chem. Biol.* **2019**, *15* (7), 747−755.

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

(85) Luo, M.; Spradlin, J. N.; Boike, L.; Tong, B.; Brittain, S. M.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Maimone, T. J.; Nomura, D. K. Chemoproteomics-enabled discovery of covalent RNF114-based degraders that mimic natural product function. *Cell Chem. Biol.* **2021**, *28* (4), 559−566.e15.
(86) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; Gonzalez-Paez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **2016**, *534* (7608), 570−4.
(87) Vinogradova, E. V.; Zhang, X.; Remillard, D.; Lazar, D. C.; Suciu, R. M.; Wang, Y.; Bianco, G.; Yamashita, Y.; Crowley, V. M.; Schafroth, M. A.; Yokoyama, M.; Konrad, D. B.; Lum, K. M.; Simon, G. M.; Kemper, E. K.; Lazear, M. R.; Yin, S.; Blewett, M. M.; Dix, M. M.; Nguyen, N.; Shokhirev, M. N.; Chin, E. N.; Lairson, L. L.; Melillo, B.; Schreiber, S. L.; Forli, S.; Teijaro, J. R.; Cravatt, B. F. An Activity-Guided Map of Electrophile-Cysteine Interactions in Primary Human T Cells. *Cell* **2020**, *182* (4), 1009−1026.e29.
(88) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; Cravatt, B. F.; Gygi, S. P. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **2021**, *39* (5), 630−641.
(89) Zhang, X.; Crowley, V. M.; Wucherpfennig, T. G.; Dix, M. M.; Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat. Chem. Biol.* **2019**, *15* (7), 737−746.
(90) Ward, C. C.; Kleinman, J. I.; Brittain, S. M.; Lee, P. S.; Chung, C. Y. S.; Kim, K.; Petri, Y.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; Schirle, M.; Nomura, D. K. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* **2019**, *14* (11), 2430−2440.
(91) Zhang, X.; Luukkonen, L. M.; Eissler, C. L.; Crowley, V. M.; Yamashita, Y.; Schafroth, M. A.; Kikuchi, S.; Weinstein, D. S.; Symons, K. T.; Nordin, B. E.; Rodriguez, J. L.; Wucherpfennig, T. G.; Bauer, L. G.; Dix, M. M.; Stamos, D.; Kinsella, T. M.; Simon, G. M.; Baltgalvis, K. A.; Cravatt, B. F. DCAF11 Supports Targeted Protein Degradation by Electrophilic Proteolysis-Targeting Chimeras. *J. Am. Chem. Soc.* **2021**, *143* (13), 5141−5149.
(92) Tao, Y.; Remillard, D.; Vinogradova, E. V.; Yokoyama, M.; Banchenko, S.; Schwefel, D.; Melillo, B.; Schreiber, S. L.; Zhang, X.; Cravatt, B. F. Targeted Protein Degradation by Electrophilic PROTACs that Stereoselectively and Site-Specifically Engage DCAF1. *J. Am. Chem. Soc.* **2022**, *144* (40), 18688−18699.
(93) Henning, N. J.; Manford, A. G.; Spradlin, J. N.; Brittain, S. M.; Zhang, E.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Rape, M.; Nomura, D. K. Discovery of a Covalent FEM1B Recruiter for Targeted Protein Degradation Applications. *J. Am. Chem. Soc.* **2022**, *144* (2), 701−708.
(94) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **2017**, *9* (12), 1181−1190.
(95) Hahm, H. S.; Toroitich, E. K.; Borne, A. L.; Brulet, J. W.; Libby, A. H.; Yuan, K.; Ware, T. B.; McCloud, R. L.; Ciancone, A. M.; Hsu, K. L. Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nat. Chem. Biol.* **2020**, *16* (2), 150−159.
(96) Brulet, J. W.; Borne, A. L.; Yuan, K.; Libby, A. H.; Hsu, K. L. Liganding Functional Tyrosine Sites on Proteins Using Sulfur-Triazole Exchange Chemistry. *J. Am. Chem. Soc.* **2020**, *142* (18), 8270−8280.
(97) Abbasov, M. E.; Kavanagh, M. E.; Ichu, T. A.; Lazear, M. R.; Tao, Y.; Crowley, V. M.; Am Ende, C. W.; Hacker, S. M.; Ho, J.; Dix, M. M.; Suciu, R.; Hayward, M. M.; Kiessling, L. L.; Cravatt, B. F. A proteome-wide atlas of lysine-reactive chemistry. *Nat. Chem.* **2021**, *13* (11), 1081−1092.
(98) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; Narvaiza, I.; Saez, E.; Cravatt, B. F. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **2017**, *168* (3), 527−541.e29.
(99) Gao, J.; Mfuh, A.; Amako, Y.; Woo, C. M. Small Molecule Interactome Mapping by Photoaffinity Labeling Reveals Binding Site Hotspots for the NSAIDs. *J. Am. Chem. Soc.* **2018**, *140* (12), 4259−4268.
(100) Wang, Y.; Dix, M. M.; Bianco, G.; Remsberg, J. R.; Lee, H. Y.; Kalocsay, M.; Gygi, S. P.; Forli, S.; Vite, G.; Lawrence, R. M.; Parker, C. G.; Cravatt, B. F. Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nat. Chem.* **2019**, *11* (12), 1113−1123.
(101) Zanon, P. R. A.; Lewald, L.; Hacker, S. M. Isotopically Labeled Desthiobiotin Azide (isoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome. *Angew. Chem., Int. Ed. Engl.* **2020**, *59* (7), 2829−2836.
(102) Kozakov, D.; Brenke, R.; Comeau, S. R.; Vajda, S. PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins* **2006**, *65* (2), 392−406.
(103) Marze, N. A.; Roy Burman, S. S.; Sheffler, W.; Gray, J. J. Efficient flexible backbone protein-protein docking for challenging targets. *Bioinformatics* **2018**, *34* (20), 3461−3469.
(104) Gainza, P.; Sverrisson, F.; Monti, F.; Rodola, E.; Boscaini, D.; Bronstein, M. M.; Correia, B. E. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat. Methods* **2020**, *17* (2), 184−192.
(105) Du, B. X.; Qin, Y.; Jiang, Y. F.; Xu, Y.; Yiu, S. M.; Yu, H.; Shi, J. Y. Compound-protein interaction prediction by deep learning: Databases, descriptors and models. *Drug Discov Today* **2022**, *27* (5), 1350−1366.
(106) Jaeger, S.; Fulle, S.; Turk, S. Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition. *J. Chem. Inf Model* **2018**, *58* (1), 27−35.
(107) Shen, W. X.; Zeng, X.; Zhu, F.; Wang, Y. l.; Qin, C.; Tan, Y.; Jiang, Y. Y.; Chen, Y. Z. Out-of-the-box deep learning prediction of pharmaceutical properties by broadly learned knowledge-based molecular representations. *Nature Machine Intelligence* **2021**, *3* (4), 334−343.
(108) Rong, Y.; Bian, Y.; Xu, T.; Xie, W.; Wei, Y.; Huang, W.; Huang, J. Self-Supervised Graph Transformer on Large-Scale Molecular Data. *arXiv* **2020**, DOI: 10.48550/arXiv.2007.02835.
(109) Krenn, M.; Ai, Q.; Barthel, S.; Carson, N.; Frei, A.; Frey, N. C.; Friederich, P.; Gaudin, T.; Gayle, A. A.; Jablonka, K. M.; Lameiro, R. F.; Lemm, D.; Lo, A.; Moosavi, S. M.; Nápoles-Duarte, J. M.; Nigam, A.; Pollice, R.; Rajan, K.; Schatzschneider, U.; Schwaller, P.; Skreta, M.; Smit, B.; Strieth-Kalthoff, F.; Sun, C.; Tom, G.; Falk von Rudorff, G.; Wang, A.; White, A. D.; Young, A.; Yu, R.; Aspuru-Guzik, A. SELFIES and the future of molecular string representations. *Patterns* **2022**, *3* (10), 100588.
(110) Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.; Pande, V. MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.* **2018**, *9* (2), 513−530.
(111) Hansen, J. D.; Condroski, K.; Correa, M.; Muller, G.; Man, H. W.; Ruchelman, A.; Zhang, W.; Vocanson, F.; Crea, T.; Liu, W.; Lu, G.; Baculi, F.; LeBrun, L.; Mahmoudi, A.; Carmel, G.; Hickman, M.; Lu, C. C. Protein Degradation via CRL4(CRBN) Ubiquitin Ligase: Discovery and Structure-Activity Relationships of Novel Glutarimide Analogs That Promote Degradation of Aiolos and/or GSPT1. *J. Med. Chem.* **2018**, *61* (2), 492−503.
(112) Altae-Tran, H.; Ramsundar, B.; Pappu, A. S.; Pande, V. Low Data Drug Discovery with One-Shot Learning. *ACS Cent Sci.* **2017**, *3* (4), 283−293.
(113) Bongers, B. J.; IJzerman, A. P.; Van Westen, G. J. P. Proteochemometrics - recent developments in bioactivity and selectivity modeling. *Drug Discov Today Technol.* **2019**, *32*−*33*, 89−98.
(114) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrian-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magarinos, M. P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R. The

ChEMBL database in 2017. *Nucleic Acids Res.* **2017**, *45* (D1), D945−D954.
(115) Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res.* **2016**, *44* (D1), D1045−53.
(116) Wang, Y.; Bryant, S. H.; Cheng, T.; Wang, J.; Gindulyte, A.; Shoemaker, B. A.; Thiessen, P. A.; He, S.; Zhang, J. PubChem BioAssay: 2017 update. *Nucleic Acids Res.* **2017**, *45* (D1), D955−D963.
(117) Alley, E. C.; Khimulya, G.; Biswas, S.; AlQuraishi, M.; Church, G. M. Unified rational protein engineering with sequence-based deep representation learning. *Nat. Methods* **2019**, *16* (12), 1315−1322.
(118) Rives, A.; Meier, J.; Sercu, T.; Goyal, S.; Lin, Z.; Liu, J.; Guo, D.; Ott, M.; Zitnick, C. L.; Ma, J.; Fergus, R. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (15), e2016239118.
(119) Bileschi, M. L.; Belanger, D.; Bryant, D. H.; Sanderson, T.; Carter, B.; Sculley, D.; Bateman, A.; DePristo, M. A.; Colwell, L. J. Using deep learning to annotate the protein universe. *Nat. Biotechnol.* **2022**, *40* (6), 932−937.
(120) Wan, X.; Wu, X.; Wang, D.; Tan, X.; Liu, X.; Fu, Z.; Jiang, H.; Zheng, M.; Li, X. An inductive graph neural network model for compound-protein interaction prediction based on a homogeneous graph. *Brief Bioinform* **2022**, *23* (3), bbac073.
(121) Piazza, I.; Beaton, N.; Bruderer, R.; Knobloch, T.; Barbisan, C.; Chandat, L.; Sudau, A.; Siepe, I.; Rinner, O.; de Souza, N.; Picotti, P.; Reiter, L. A machine learning-based chemoproteomic approach to identify drug targets and binding sites in complex proteomes. *Nat. Commun.* **2020**, *11* (1), 4200.
(122) Bender, B. J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balius, T. E.; Carlsson, J.; Irwin, J. J.; Shoichet, B. K. A practical guide to large-scale docking. *Nat. Protoc* **2021**, *16* (10), 4799−4832.
(123) Combs, S. A.; Deluca, S. L.; Deluca, S. H.; Lemmon, G. H.; Nannemann, D. P.; Nguyen, E. D.; Willis, J. R.; Sheehan, J. H.; Meiler, J. Small-molecule ligand docking into comparative models with Rosetta. *Nat. Protoc* **2013**, *8* (7), 1277−98.
(124) Mosca, R.; Ceol, A.; Aloy, P. Interactome3D: adding structural details to protein networks. *Nat. Methods* **2013**, *10* (1), 47−53.
(125) Lohning, A. E.; Levonis, S. M.; Williams-Noonan, B.; Schweiker, S. S. A Practical Guide to Molecular Docking and Homology Modelling for Medicinal Chemists. *Curr. Top Med. Chem.* **2017**, *17* (18), 2023−2040.
(126) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296−W303.
(127) Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Zidek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; Velankar, S.; Kleywegt, G. J.; Bateman, A.; Evans, R.; Pritzel, A.; Figurnov, M.; Ronneberger, O.; Bates, R.; Kohl, S. A. A.; Potapenko, A.; Ballard, A. J.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Clancy, E.; Reiman, D.; Petersen, S.; Senior, A. W.; Kavukcuoglu, K.; Birney, E.; Kohli, P.; Jumper, J.; Hassabis, D. Highly accurate protein structure prediction for the human proteome. *Nature* **2021**, *596* (7873), 590−596.
(128) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millan, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373* (6557), 871−876.
(129) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Language models of protein sequences at the scale of evolution enable accurate structure prediction. *bioRxiv* **2022**, DOI: 10.1101/2022.07.20.500902.
(130) Wu, R.; Ding, F.; Wang, R.; Shen, R.; Zhang, X.; Luo, S.; Su, C.; Wu, Z.; Xie, Q.; Berger, B.; Ma, J.; Peng, J. High-resolution < em > de novo structure prediction from primary sequence. *bioRxiv* **2022**, DOI: 10.1101/2022.07.21.500999.
(131) Porta-Pardo, E.; Ruiz-Serra, V.; Valentini, S.; Valencia, A. The structural coverage of the human proteome before and after AlphaFold. *PLoS Comput. Biol.* **2022**, *18* (1), e1009818.
(132) Wong, F.; Krishnan, A.; Zheng, E. J.; Stark, H.; Manson, A. L.; Earl, A. M.; Jaakkola, T.; Collins, J. J. Benchmarking AlphaFold-enabled molecular docking predictions for antibiotic discovery. *Mol. Syst. Biol.* **2022**, *18* (9), e11081.
(133) Mullard, A. What does AlphaFold mean for drug discovery? *Nat. Rev. Drug Discov* **2021**, *20* (10), 725−727.
(134) Binder, J. L.; Berendzen, J.; Stevens, A. O.; He, Y.; Wang, J.; Dokholyan, N. V.; Oprea, T. I. AlphaFold illuminates half of the dark human proteins. *Curr. Opin Struct Biol.* **2022**, *74*, 102372.
(135) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **2009**, *10*, 168.
(136) Zhao, J.; Cao, Y.; Zhang, L. Exploring the computational methods for protein-ligand binding site prediction. *Comput. Struct Biotechnol J.* **2020**, *18*, 417−426.
(137) Aggarwal, R.; Gupta, A.; Chelur, V.; Jawahar, C. V.; Priyakumar, U. D. DeepPocket: Ligand Binding Site Detection and Segmentation using 3D Convolutional Neural Networks. *J. Chem. Inf Model* **2022**, *62* (21), 5069−5079.
(138) Pu, L.; Govindaraj, R. G.; Lemoine, J. M.; Wu, H. C.; Brylinski, M. DeepDrug3D: Classification of ligand-binding pockets in proteins with a convolutional neural network. *PLoS Comput. Biol.* **2019**, *15* (2), e1006718.
(139) Ouyang, X.; Zhou, S.; Su, C. T.; Ge, Z.; Li, R.; Kwoh, C. K. CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J. Comput. Chem.* **2013**, *34* (4), 326−36.
(140) London, N.; Miller, R. M.; Krishnan, S.; Uchida, K.; Irwin, J. J.; Eidam, O.; Gibold, L.; Cimermancic, P.; Bonnet, R.; Shoichet, B. K.; Taunton, J. Covalent docking of large libraries for the discovery of chemical probes. *Nat. Chem. Biol.* **2014**, *10* (12), 1066−72.
(141) Bianco, G.; Forli, S.; Goodsell, D. S.; Olson, A. J. Covalent docking using autodock: Two-point attractor and flexible side chain methods. *Protein Sci.* **2016**, *25* (1), 295−301.
(142) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; Tolmachev, A. A.; Shoichet, B. K.; Roth, B. L.; Irwin, J. J. Ultra-large library docking for discovering new chemotypes. *Nature* **2019**, *566* (7743), 224−229.
(143) Gorgulla, C.; Boeszoermenyi, A.; Wang, Z. F.; Fischer, P. D.; Coote, P. W.; Padmanabha Das, K. M.; Malets, Y. S.; Radchenko, D. S.; Moroz, Y. S.; Scott, D. A.; Fackeldey, K.; Hoffmann, M.; Iavniuk, I.; Wagner, G.; Arthanari, H. An open-source drug discovery platform enables ultra-large virtual screens. *Nature* **2020**, *580* (7805), 663−668.
(144) Sadybekov, A. A.; Sadybekov, A. V.; Liu, Y.; Iliopoulos-Tsoutsouvas, C.; Huang, X. P.; Pickett, J.; Houser, B.; Patel, N.; Tran, N. K.; Tong, F.; Zvonok, N.; Jain, M. K.; Savych, O.; Radchenko, D. S.; Nikas, S. P.; Petasis, N. A.; Moroz, Y. S.; Roth, B. L.; Makriyannis, A.; Katritch, V. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2022**, *601* (7893), 452−459.
(145) Popova, M.; Isayev, O.; Tropsha, A. Deep reinforcement learning for de novo drug design. *Sci. Adv.* **2018**, *4* (7), eaap7885.
(146) Vanhaelen, Q.; Lin, Y. C.; Zhavoronkov, A. The Advent of Generative Chemistry. *ACS Med. Chem. Lett.* **2020**, *11* (8), 1496−1505.

(147) Cheng, Y.; Gong, Y.; Liu, Y.; Song, B.; Zou, Q. Molecular design in drug discovery: a comprehensive review of deep generative models. *Brief Bioinform* **2021**, *22* (6), bbab344.
(148) Meyers, J.; Fabian, B.; Brown, N. De novo molecular design and generative models. *Drug Discov Today* **2021**, *26* (11), 2707−2715.
(149) Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Cent Sci.* **2018**, *4* (1), 120−131.
(150) Skalic, M.; Jimenez, J.; Sabbadin, D.; De Fabritiis, G. Shape-Based Generative Modeling for de Novo Drug Design. *J. Chem. Inf Model* **2019**, *59* (3), 1205−1214.
(151) Wang, M.; Hsieh, C. Y.; Wang, J.; Wang, D.; Weng, G.; Shen, C.; Yao, X.; Bing, Z.; Li, H.; Cao, D.; Hou, T. RELATION: A Deep Generative Model for Structure-Based De Novo Drug Design. *J. Med. Chem.* **2022**, *65* (13), 9478−9492.
(152) Maziarka, L.; Pocha, A.; Kaczmarczyk, J.; Rataj, K.; Danel, T.; Warchol, M. Mol-CycleGAN: a generative model for molecular optimization. *J. Cheminform* **2020**, *12* (1), 2.
(153) Li, Y.; Zhang, L.; Liu, Z. Multi-objective de novo drug design with conditional graph generative model. *J. Cheminform* **2018**, *10* (1), 33.
(154) Blaschke, T.; Arus-Pous, J.; Chen, H.; Margreitter, C.; Tyrchan, C.; Engkvist, O.; Papadopoulos, K.; Patronov, A. REINVENT 2.0: An AI Tool for De Novo Drug Design. *J. Chem. Inf Model* **2020**, *60* (12), 5918−5922.
(155) Xu, M.; Ran, T.; Chen, H. De Novo Molecule Design Through the Molecular Generative Model Conditioned by 3D Information of Protein Binding Sites. *J. Chem. Inf Model* **2021**, *61* (7), 3240−3254.
(156) Thomas, M.; Smith, R. T.; O'Boyle, N. M.; de Graaf, C.; Bender, A. Comparison of structure- and ligand-based scoring functions for deep generative models: a GPCR case study. *J. Cheminform* **2021**, *13* (1), 39.
(157) Moret, M.; Friedrich, L.; Grisoni, F.; Merk, D.; Schneider, G. Generative molecular design in low data regimes. *Nature Machine Intelligence* **2020**, *2* (3), 171−180.
(158) Brown, N.; Fiscato, M.; Segler, M. H. S.; Vaucher, A. C. GuacaMol: Benchmarking Models for de Novo Molecular Design. *J. Chem. Inf Model* **2019**, *59* (3), 1096−1108.
(159) Polykovskiy, D.; Zhebrak, A.; Sanchez-Lengeling, B.; Golovanov, S.; Tatanov, O.; Belyaev, S.; Kurbanov, R.; Artamonov, A.; Aladinskiy, V.; Veselov, M.; Kadurin, A.; Johansson, S.; Chen, H.; Nikolenko, S.; Aspuru-Guzik, A.; Zhavoronkov, A. Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models. *Front Pharmacol* **2020**, *11*, 565644.
(160) Chen, H. Can Generative-Model-Based Drug Design Become a New Normal in Drug Discovery? *J. Med. Chem.* **2022**, *65* (1), 100−102.
(161) Grisoni, F.; Huisman, B. J. H.; Button, A. L.; Moret, M.; Atz, K.; Merk, D.; Schneider, G. Combining generative artificial intelligence and on-chip synthesis for de novo drug design. *Sci. Adv.* **2021**, *7* (24), eabg3338.
(162) Hadfield, T. E.; Imrie, F.; Merritt, A.; Birchall, K.; Deane, C. M. Incorporating Target-Specific Pharmacophoric Information into Deep Generative Models for Fragment Elaboration. *J. Chem. Inf Model* **2022**, *62* (10), 2280−2292.
(163) Guo, J.; Knuth, F.; Margreitter, C.; Janet, J. P.; Papadopoulos, K.; Engkvist, O.; Patronov, A. Link-INVENT: Generative Linker Design with Reinforcement Learning. *ChemRxiv* **2022**, DOI: 10.26434/chemrxiv-2022-qkx9f.
(164) Wang, J.; Chu, Y.; Mao, J.; Jeon, H. N.; Jin, H.; Zeb, A.; Jang, Y.; Cho, K. H.; Song, T.; No, K. T. De novo molecular design with deep molecular generative models for PPI inhibitors. *Brief Bioinform* **2022**, *23* (4), bbac285.
(165) Flam-Shepherd, D.; Zhu, K.; Aspuru-Guzik, A. Language models can learn complex molecular distributions. *Nat. Commun.* **2022**, *13* (1), 3293.
(166) Mitsiades, N.; Mitsiades, C. S.; Poulaki, V.; Chauhan, D.; Richardson, P. G.; Hideshima, T.; Munshi, N. C.; Treon, S. P.; Anderson, K. C. Apoptotic signaling induced by immunomodulatory thalidomide analogs in human multiple myeloma cells: therapeutic implications. *Blood* **2002**, *99* (12), 4525−30.
(167) Hansen, J. D.; Correa, M.; Nagy, M. A.; Alexander, M.; Plantevin, V.; Grant, V.; Whitefield, B.; Huang, D.; Kercher, T.; Harris, R.; Narla, R. K.; Leisten, J.; Tang, Y.; Moghaddam, M.; Ebinger, K.; Piccotti, J.; Havens, C. G.; Cathers, B.; Carmichael, J.; Daniel, T.; Vessey, R.; Hamann, L. G.; Leftheris, K.; Mendy, D.; Baculi, F.; LeBrun, L. A.; Khambatta, G.; Lopez-Girona, A. Discovery of CRBN E3 Ligase Modulator CC-92480 for the Treatment of Relapsed and Refractory Multiple Myeloma. *J. Med. Chem.* **2020**, *63* (13), 6648−6676.
(168) Surka, C.; Jin, L.; Mbong, N.; Lu, C. C.; Jang, I. S.; Rychak, E.; Mendy, D.; Clayton, T.; Tindall, E.; Hsu, C.; Fontanillo, C.; Tran, E.; Contreras, A.; Ng, S. W. K.; Matyskiela, M.; Wang, K.; Chamberlain, P.; Cathers, B.; Carmichael, J.; Hansen, J.; Wang, J. C. Y.; Minden, M. D.; Fan, J.; Pierce, D. W.; Pourdehnad, M.; Rolfe, M.; Lopez-Girona, A.; Dick, J. E.; Lu, G. CC-90009, a novel cereblon E3 ligase modulator, targets acute myeloid leukemia blasts and leukemia stem cells. *Blood* **2021**, *137* (5), 661−677.
(169) Hansen, J. D.; Correa, M.; Alexander, M.; Nagy, M.; Huang, D.; Sapienza, J.; Lu, G.; LeBrun, L. A.; Cathers, B. E.; Zhang, W.; Tang, Y.; Ammirante, M.; Narla, R. K.; Piccotti, J. R.; Pourdehnad, M.; Lopez-Girona, A. CC-90009: A Cereblon E3 Ligase Modulating Drug That Promotes Selective Degradation of GSPT1 for the Treatment of Acute Myeloid Leukemia. *J. Med. Chem.* **2021**, *64* (4), 1835−1843.
(170) Srivatsan, S. R.; McFaline-Figueroa, J. L.; Ramani, V.; Saunders, L.; Cao, J.; Packer, J.; Pliner, H. A.; Jackson, D. L.; Daza, R. M.; Christiansen, L.; Zhang, F.; Steemers, F.; Shendure, J.; Trapnell, C. Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **2020**, *367* (6473), 45−51.
(171) Bray, M. A.; Singh, S.; Han, H.; Davis, C. T.; Borgeson, B.; Hartland, C.; Kost-Alimova, M.; Gustafsdottir, S. M.; Gibson, C. C.; Carpenter, A. E. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc* **2016**, *11* (9), 1757−74.
(172) Feldman, D.; Singh, A.; Schmid-Burgk, J. L.; Carlson, R. J.; Mezger, A.; Garrity, A. J.; Zhang, F.; Blainey, P. C. Optical Pooled Screens in Human Cells. *Cell* **2019**, *179* (3), 787−799.e17.
(173) Chandrasekaran, S. N.; Ceulemans, H.; Boyd, J. D.; Carpenter, A. E. Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat. Rev. Drug Discov* **2021**, *20* (2), 145−159.
(174) Rohban, M. H.; Fuller, A. M.; Tan, C.; Goldstein, J. T.; Syangtan, D.; Gutnick, A.; DeVine, A.; Nijsure, M. P.; Rigby, M.; Sacher, J. R.; Corsello, S. M.; Peppler, G. B.; Bogaczynska, M.; Boghossian, A.; Ciotti, G. E.; Hands, A. T.; Mekareeya, A.; Doan, M.; Gale, J. P.; Derynck, R.; Turbyville, T.; Boerckel, J. D.; Singh, S.; Kiessling, L. L.; Schwarz, T. L.; Varelas, X.; Wagner, F. F.; Kafri, R.; Eisinger-Mathason, T. S. K.; Carpenter, A. E. Virtual screening for small-molecule pathway regulators by image-profile matching. *Cell Syst* **2022**, *13* (9), 724−736.e9.
(175) Lukonin, I.; Serra, D.; Challet Meylan, L.; Volkmann, K.; Baaten, J.; Zhao, R.; Meeusen, S.; Colman, K.; Maurer, F.; Stadler, M. B.; Jenkins, J.; Liberali, P. Phenotypic landscape of intestinal organoid regeneration. *Nature* **2020**, *586* (7828), 275−280.
(176) Lukonin, I.; Zinner, M.; Liberali, P. Organoids in image-based phenotypic chemical screens. *Exp Mol. Med.* **2021**, *53* (10), 1495−1502.
(177) Vladimer, G. I.; Snijder, B.; Krall, N.; Bigenzahn, J. W.; Huber, K. V. M.; Lardeau, C. H.; Sanjiv, K.; Ringler, A.; Berglund, U. W.; Sabler, M.; de la Fuente, O. L.; Knobl, P.; Kubicek, S.; Helleday, T.; Jager, U.; Superti-Furga, G. Global survey of the immunomodulatory potential of common drugs. *Nat. Chem. Biol.* **2017**, *13* (6), 681−690.
(178) Schmidl, C.; Vladimer, G. I.; Rendeiro, A. F.; Schnabl, S.; Krausgruber, T.; Taubert, C.; Krall, N.; Pemovska, T.; Araghi, M.; Snijder, B.; Hubmann, R.; Ringler, A.; Runggatscher, K.; Demirtas, D.;

de la Fuente, O. L.; Hilgarth, M.; Skrabs, C.; Porpaczy, E.; Gruber, M.; Hoermann, G.; Kubicek, S.; Staber, P. B.; Shehata, M.; Superti-Furga, G.; Jager, U.; Bock, C. Combined chemosensitivity and chromatin profiling prioritizes drug combinations in CLL. *Nat. Chem. Biol.* **2019**, *15* (3), 232−240.
(179) Heinemann, T.; Kornauth, C.; Severin, Y.; Vladimer, G. I.; Pemovska, T.; Hadzijusufovic, E.; Agis, H.; Krauth, M. T.; Sperr, W. R.; Valent, P.; Jager, U.; Simonitsch-Klupp, I.; Superti-Furga, G.; Staber, P. B.; Snijder, B. Deep Morphology Learning Enhances Ex Vivo Drug Profiling-Based Precision Medicine. *Blood Cancer Discov* **2022**, *3* (6), 502−515.
(180) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **2003**, *21* (1), 86−9.
(181) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **2008**, *3* (6), 373−82.
(182) Yu, M. K.; Ma, J.; Fisher, J.; Kreisberg, J. F.; Raphael, B. J.; Ideker, T. Visible Machine Learning for Biomedicine. *Cell* **2018**, *173* (7), 1562−1565.
(183) Schaffer, L. V.; Ideker, T. Mapping the multiscale structure of biological systems. *Cell Syst* **2021**, *12* (6), 622−635.
(184) Ma, J.; Yu, M. K.; Fong, S.; Ono, K.; Sage, E.; Demchak, B.; Sharan, R.; Ideker, T. Using deep learning to model the hierarchical structure and function of a cell. *Nat. Methods* **2018**, *15* (4), 290−298.
(185) Elmarakeby, H. A.; Hwang, J.; Arafeh, R.; Crowdis, J.; Gang, S.; Liu, D.; AlDubayan, S. H.; Salari, K.; Kregel, S.; Richter, C.; Arnoff, T. E.; Park, J.; Hahn, W. C.; Van Allen, E. M. Biologically informed deep neural network for prostate cancer discovery. *Nature* **2021**, *598* (7880), 348−352.
(186) Trapotsi, M. A.; Mouchet, E.; Williams, G.; Monteverde, T.; Juhani, K.; Turkki, R.; Miljkovic, F.; Martinsson, A.; Mervin, L.; Pryde, K. R.; Mullers, E.; Barrett, I.; Engkvist, O.; Bender, A.; Moreau, K. Cell Morphological Profiling Enables High-Throughput Screening for Proteolysis TArgeting Chimera (PROTAC) Phenotypic Signature. *ACS Chem. Biol.* **2022**, *17* (7), 1733−1744.
(187) Iorio, F.; Knijnenburg, T. A.; Vis, D. J.; Bignell, G. R.; Menden, M. P.; Schubert, M.; Aben, N.; Goncalves, E.; Barthorpe, S.; Lightfoot, H.; Cokelaer, T.; Greninger, P.; van Dyk, E.; Chang, H.; de Silva, H.; Heyn, H.; Deng, X.; Egan, R. K.; Liu, Q.; Mironenko, T.; Mitropoulos, X.; Richardson, L.; Wang, J.; Zhang, T.; Moran, S.; Sayols, S.; Soleimani, M.; Tamborero, D.; Lopez-Bigas, N.; Ross-Macdonald, P.; Esteller, M.; Gray, N. S.; Haber, D. A.; Stratton, M. R.; Benes, C. H.; Wessels, L. F. A.; Saez-Rodriguez, J.; McDermott, U.; Garnett, M. J. A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **2016**, *166* (3), 740−754.
(188) Corsello, S. M.; Nagari, R. T.; Spangler, R. D.; Rossen, J.; Kocak, M.; Bryan, J. G.; Humeidi, R.; Peck, D.; Wu, X.; Tang, A. A.; Wang, V. M.; Bender, S. A.; Lemire, E.; Narayan, R.; Montgomery, P.; Ben-David, U.; Garvie, C. W.; Chen, Y.; Rees, M. G.; Lyons, N. J.; McFarland, J. M.; Wong, B. T.; Wang, L.; Dumont, N.; O'Hearn, P. J.; Stefan, E.; Doench, J. G.; Harrington, C. N.; Greulich, H.; Meyerson, M.; Vazquez, F.; Subramanian, A.; Roth, J. A.; Bittker, J. A.; Boehm, J. S.; Mader, C. C.; Tsherniak, A.; Golub, T. R. Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat. Cancer* **2020**, *1* (2), 235−248.
(189) Echle, A.; Rindtorff, N. T.; Brinker, T. J.; Luedde, T.; Pearson, A. T.; Kather, J. N. Deep learning in cancer pathology: a new generation of clinical biomarkers. *Br. J. Cancer* **2021**, *124* (4), 686−696.
(190) Chawla, S.; Rockstroh, A.; Lehman, M.; Ratther, E.; Jain, A.; Anand, A.; Gupta, A.; Bhattacharya, N.; Poonia, S.; Rai, P.; Das, N.; Majumdar, A.; Jayadeva; Ahuja, G.; Hollier, B. G.; Nelson, C. C.; Sengupta, D. Gene expression based inference of cancer drug sensitivity. *Nat. Commun.* **2022**, *13* (1), 5680.
(191) Kuenzi, B. M.; Park, J.; Fong, S. H.; Sanchez, K. S.; Lee, J.; Kreisberg, J. F.; Ma, J.; Ideker, T. Predicting Drug Response and Synergy Using a Deep Learning Model of Human Cancer Cells. *Cancer Cell* **2020**, *38* (5), 672−684.e6.
(192) Duran-Frigola, M.; Pauls, E.; Guitart-Pla, O.; Bertoni, M.; Alcalde, V.; Amat, D.; Juan-Blanco, T.; Aloy, P. Extending the small-molecule similarity principle to all levels of biology with the Chemical Checker. *Nat. Biotechnol.* **2020**, *38* (9), 1087−1096.
(193) Bertoni, M.; Duran-Frigola, M.; Badia-i-Mompel, P.; Pauls, E.; Orozco-Ruiz, M.; Guitart-Pla, O.; Alcalde, V.; Diaz, V. M.; Berenguer-Llergo, A.; Brun-Heath, I.; Villegas, N.; de Herreros, A. G.; Aloy, P. Bioactivity descriptors for uncharacterized chemical compounds. *Nat. Commun.* **2021**, *12* (1), 3932.
(194) Himmelstein, D. S.; Lizee, A.; Hessler, C.; Brueggeman, L.; Chen, S. L.; Hadley, D.; Green, A.; Khankhanian, P.; Baranzini, S. E. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *Elife* **2017**, *6*, e26726.
(195) Santos, A.; Colaco, A. R.; Nielsen, A. B.; Niu, L.; Strauss, M.; Geyer, P. E.; Coscia, F.; Albrechtsen, N. J. W.; Mundt, F.; Jensen, L. J.; Mann, M. A knowledge graph to interpret clinical proteomics data. *Nat. Biotechnol.* **2022**, *40* (5), 692−702.
(196) Fernandez-Torras, A.; Duran-Frigola, M.; Bertoni, M.; Locatelli, M.; Aloy, P. Integrating and formatting biomedical data as pre-calculated knowledge graph embeddings in the Bioteque. *Nat. Commun.* **2022**, *13* (1), 5304.
(197) Fernandez-Torras, A.; Comajuncosa-Creus, A.; Duran-Frigola, M.; Aloy, P. Connecting chemistry and biology through molecular descriptors. *Curr. Opin Chem. Biol.* **2022**, *66*, 102090.
(198) BIONIC: discovering new biology through deep learning-based network integration. *Nat. Methods* **2022**, *19* (10), 1185−1186..
(199) Garnett, M. J.; Edelman, E. J.; Heidorn, S. J.; Greenman, C. D.; Dastur, A.; Lau, K. W.; Greninger, P.; Thompson, I. R.; Luo, X.; Soares, J.; Liu, Q.; Iorio, F.; Surdez, D.; Chen, L.; Milano, R. J.; Bignell, G. R.; Tam, A. T.; Davies, H.; Stevenson, J. A.; Barthorpe, S.; Lutz, S. R.; Kogera, F.; Lawrence, K.; McLaren-Douglas, A.; Mitropoulos, X.; Mironenko, T.; Thi, H.; Richardson, L.; Zhou, W.; Jewitt, F.; Zhang, T.; O'Brien, P.; Boisvert, J. L.; Price, S.; Hur, W.; Yang, W.; Deng, X.; Butler, A.; Choi, H. G.; Chang, J. W.; Baselga, J.; Stamenkovic, I.; Engelman, J. A.; Sharma, S. V.; Delattre, O.; Saez-Rodriguez, J.; Gray, N. S.; Settleman, J.; Futreal, P. A.; Haber, D. A.; Stratton, M. R.; Ramaswamy, S.; McDermott, U.; Benes, C. H. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **2012**, *483* (7391), 570−5.
(200) Deng, L.; Cai, Y.; Zhang, W.; Yang, W.; Gao, B.; Liu, H. Pathway-Guided Deep Neural Network toward Interpretable and Predictive Modeling of Drug Sensitivity. *J. Chem. Inf Model* **2020**, *60* (10), 4497−4505.
(201) Ma, J.; Fong, S. H.; Luo, Y.; Bakkenist, C. J.; Shen, J. P.; Mourragui, S.; Wessels, L. F. A.; Hafner, M.; Sharan, R.; Peng, J.; Ideker, T. Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nat. Cancer* **2021**, *2* (2), 233−244.
(202) Kurimchak, A. M.; Herrera-Montavez, C.; Montserrat-Sangra, S.; Araiza-Olivera, D.; Hu, J.; Neumann-Domer, R.; Kuruvilla, M.; Bellacosa, A.; Testa, J. R.; Jin, J.; Duncan, J. S. The drug efflux pump MDR1 promotes intrinsic and acquired resistance to PROTACs in cancer cells. *Sci. Signal* **2022**, *15* (749), eabn2707.
(203) Ziegler, S.; Sievers, S.; Waldmann, H. Morphological profiling of small molecules. *Cell Chem. Biol.* **2021**, *28* (3), 300−319.
(204) Schubert, M.; Klinger, B.; Klunemann, M.; Sieber, A.; Uhlitz, F.; Sauer, S.; Garnett, M. J.; Bluthgen, N.; Saez-Rodriguez, J. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **2018**, *9* (1), 20.
(205) Subramanian, A.; Narayan, R.; Corsello, S. M.; Peck, D. D.; Natoli, T. E.; Lu, X.; Gould, J.; Davis, J. F.; Tubelli, A. A.; Asiedu, J. K.; Lahr, D. L.; Hirschman, J. E.; Liu, Z.; Donahue, M.; Julian, B.; Khan, M.; Wadden, D.; Smith, I. C.; Lam, D.; Liberzon, A.; Toder, C.; Bagul, M.; Orzechowski, M.; Enache, O. M.; Piccioni, F.; Johnson, S. A.; Lyons, N. J.; Berger, A. H.; Shamji, A. F.; Brooks, A. N.; Vrcic, A.;

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

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

(206) Musa, A.; Ghoraie, L. S.; Zhang, S. D.; Glazko, G.; Yli-Harja, O.; Dehmer, M.; Haibe-Kains, B.; Emmert-Streib, F. A review of connectivity map and computational approaches in pharmacogenomics. *Brief Bioinform* **2017**, *19* (3), 506−523.

(207) Lamb, J.; Crawford, E. D.; Peck, D.; Modell, J. W.; Blat, I. C.; Wrobel, M. J.; Lerner, J.; Brunet, J. P.; Subramanian, A.; Ross, K. N.; Reich, M.; Hieronymus, H.; Wei, G.; Armstrong, S. A.; Haggarty, S. J.; Clemons, P. A.; Wei, R.; Carr, S. A.; Lander, E. S.; Golub, T. R. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313* (5795), 1929−35.

(208) McFarland, J. M.; Paolella, B. R.; Warren, A.; Geiger-Schuller, K.; Shibue, T.; Rothberg, M.; Kuksenko, O.; Colgan, W. N.; Jones, A.; Chambers, E.; Dionne, D.; Bender, S.; Wolpin, B. M.; Ghandi, M.; Tirosh, I.; Rozenblatt-Rosen, O.; Roth, J. A.; Golub, T. R.; Regev, A.; Aguirre, A. J.; Vazquez, F.; Tsherniak, A. Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and therapeutic mechanism of action. *Nat. Commun.* **2020**, *11* (1), 4296.

(209) Douglass, E. F., Jr; Allaway, R. J.; Szalai, B.; Wang, W.; Tian, T.; Fernandez-Torras, A.; Realubit, R.; Karan, C.; Zheng, S.; Pessia, A.; Tanoli, Z.; Jafari, M.; Wan, F.; Li, S.; Xiong, Y.; Duran-Frigola, M.; Bertoni, M.; Badia, I. M. P.; Mateo, L.; Guitart-Pla, O.; Chung, V.; Consortium, D.; C.-s, P. D. A. C.; Tang, J.; Zeng, J.; Aloy, P.; Saez-Rodriguez, J.; Guinney, J.; Gerhard, D. S.; Califano, A. A community challenge for a pancancer drug mechanism of action inference from perturbational profile data. *Cell Rep. Med.* **2022**, *3* (1), 100492.

(210) Filzen, T. M.; Kutchukian, P. S.; Hermes, J. D.; Li, J.; Tudor, M. Representing high throughput expression profiles via perturbation barcodes reveals compound targets. *PLoS Comput. Biol.* **2017**, *13* (2), e1005335.

(211) Replogle, J. M.; Saunders, R. A.; Pogson, A. N.; Hussmann, J. A.; Lenail, A.; Guna, A.; Mascibroda, L.; Wagner, E. J.; Adelman, K.; Lithwick-Yanai, G.; Iremadze, N.; Oberstrass, F.; Lipson, D.; Bonnar, J. L.; Jost, M.; Norman, T. M.; Weissman, J. S. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell* **2022**, *185* (14), 2559−2575.e28.

(212) Funk, L.; Su, K. C.; Ly, J.; Feldman, D.; Singh, A.; Moodie, B.; Blainey, P. C.; Cheeseman, I. M. The phenotypic landscape of essential human genes. *Cell* **2022**, *185* (24), 4634−4653.e22.

(213) Bock, C.; Datlinger, P.; Chardon, F.; Coelho, M. A.; Dong, M. T. B.; Lawson, K. A.; Lu, T.; Maroc, L.; Norman, T. M.; Song, B. A.; Stanley, G.; Chen, S. D.; Garnett, M.; Li, W.; Moffat, J.; Qi, L. S.; Shapiro, R. S.; Shendure, J.; Weissman, J. S.; Zhuang, X. W. High-content CRISPR screening. *Nature Reviews Methods Primers* **2022**, *2* (1), 8.

(214) Qin, Y.; Huttlin, E. L.; Winsnes, C. F.; Gosztyla, M. L.; Wacheul, L.; Kelly, M. R.; Blue, S. M.; Zheng, F.; Chen, M.; Schaffer, L. V.; Licon, K.; Backstrom, A.; Vaites, L. P.; Lee, J. J.; Ouyang, W.; Liu, S. N.; Zhang, T.; Silva, E.; Park, J.; Pitea, A.; Kreisberg, J. F.; Gygi, S. P.; Ma, J.; Harper, J. W.; Yeo, G. W.; Lafontaine, D. L. J.; Lundberg, E.; Ideker, T. A multi-scale map of cell structure fusing protein images and interactions. *Nature* **2021**, *600* (7889), 536−542.

(215) Garrido-Rodriguez, M.; Zirngibl, K.; Ivanova, O.; Lobentanzer, S.; Saez-Rodriguez, J. Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Mol. Syst. Biol.* **2022**, *18* (7), e11036.

(216) Kelly, R. T. Single-cell Proteomics: Progress and Prospects. *Mol. Cell Proteomics* **2020**, *19* (11), 1739−1748.

(217) Labib, M.; Kelley, S. O. Single-cell analysis targeting the proteome. *Nature Reviews Chemistry* **2020**, *4* (3), 143−158.

(218) Schoof, E. M.; Furtwangler, B.; Uresin, N.; Rapin, N.; Savickas, S.; Gentil, C.; Lechman, E.; Keller, U. A. D.; Dick, J. E.; Porse, B. T. Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **2021**, *12* (1), 3341.

(219) Alexandrov, T. Spatial Metabolomics and Imaging Mass Spectrometry in the Age of Artificial Intelligence. *Annu. Rev. Biomed Data Sci.* **2020**, *3*, 61−87.

(220) Rappez, L.; Stadler, M.; Triana, S.; Gathungu, R. M.; Ovchinnikova, K.; Phapale, P.; Heikenwalder, M.; Alexandrov, T. SpaceM reveals metabolic states of single cells. *Nat. Methods* **2021**, *18* (7), 799−805.

(221) Dixit, A.; Parnas, O.; Li, B.; Chen, J.; Fulco, C. P.; Jerby-Arnon, L.; Marjanovic, N. D.; Dionne, D.; Burks, T.; Raychowdhury, R.; Adamson, B.; Norman, T. M.; Lander, E. S.; Weissman, J. S.; Friedman, N.; Regev, A. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **2016**, *167* (7), 1853−1866 e17.

(222) Datlinger, P.; Rendeiro, A. F.; Schmidl, C.; Krausgruber, T.; Traxler, P.; Klughammer, J.; Schuster, L. C.; Kuchler, A.; Alpar, D.; Bock, C. Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **2017**, *14* (3), 297−301.

(223) Schraivogel, D.; Gschwind, A. R.; Milbank, J. H.; Leonce, D. R.; Jakob, P.; Mathur, L.; Korbel, J. O.; Merten, C. A.; Velten, L.; Steinmetz, L. M. Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat. Methods* **2020**, *17* (6), 629−635.

(224) Savitski, M. M.; Reinhard, F. B.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Martinez Molina, D.; Jafari, R.; Dovega, R. B.; Klaeger, S.; Kuster, B.; Nordlund, P.; Bantscheff, M.; Drewes, G. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **2014**, *346* (6205), 1255784.

(225) Kurzawa, N.; Becher, I.; Sridharan, S.; Franken, H.; Mateus, A.; Anders, S.; Bantscheff, M.; Huber, W.; Savitski, M. M. A computational method for detection of ligand-binding proteins from dose range thermal proteome profiles. *Nat. Commun.* **2020**, *11* (1), 5783.

(226) Scholes, N. S.; Mayor-Ruiz, C.; Winter, G. E. Identification and selectivity profiling of small-molecule degraders via multi-omics approaches. *Cell Chem. Biol.* **2021**, *28* (7), 1048−1060.

(227) Gaetani, M.; Sabatier, P.; Saei, A. A.; Beusch, C. M.; Yang, Z.; Lundstrom, S. L.; Zubarev, R. A. Proteome Integral Solubility Alteration: A High-Throughput Proteomics Assay for Target Deconvolution. *J. Proteome Res.* **2019**, *18* (11), 4027−4037.

(228) Mateus, A.; Kurzawa, N.; Becher, I.; Sridharan, S.; Helm, D.; Stein, F.; Typas, A.; Savitski, M. M. Thermal proteome profiling for interrogating protein interactions. *Mol. Syst. Biol.* **2020**, *16* (3), e9232.

(229) Brunner, A. D.; Thielert, M.; Vasilopoulou, C.; Ammar, C.; Coscia, F.; Mund, A.; Hoerning, O. B.; Bache, N.; Apalategui, A.; Lubeck, M.; Richter, S.; Fischer, D. S.; Raether, O.; Park, M. A.; Meier, F.; Theis, F. J.; Mann, M. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **2022**, *18* (3), e10798.

(230) Mund, A.; Coscia, F.; Kriston, A.; Hollandi, R.; Kovacs, F.; Brunner, A. D.; Migh, E.; Schweizer, L.; Santos, A.; Bzorek, M.; Naimy, S.; Rahbek-Gjerdrum, L. M.; Dyring-Andersen, B.; Bulkescher, J.; Lukas, C.; Eckert, M. A.; Lengyel, E.; Gnann, C.; Lundberg, E.; Horvath, P.; Mann, M. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat. Biotechnol.* **2022**, *40* (8), 1231−1240.

(231) Vegvari, A.; Rodriguez, J. E.; Zubarev, R. A. Single-Cell Chemical Proteomics (SCCP) Interrogates the Timing and Heterogeneity of Cancer Cell Commitment to Death. *Anal. Chem.* **2022**, *94* (26), 9261−9269.

(232) Sievers, Q. L.; Gasser, J. A.; Cowley, G. S.; Fischer, E. S.; Ebert, B. L. Genome-wide screen identifies cullin-RING ligase machinery required for lenalidomide-dependent CRL4(CRBN) activity. *Blood* **2018**, *132* (12), 1293−1303.

(233) Mayor-Ruiz, C.; Jaeger, M. G.; Bauer, S.; Brand, M.; Sin, C.; Hanzl, A.; Mueller, A. C.; Menche, J.; Winter, G. E. Plasticity of the Cullin-RING Ligase Repertoire Shapes Sensitivity to Ligand-Induced Protein Degradation. *Mol. Cell* **2019**, *75* (4), 849−858 e8.

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732

(234) de Almeida, M.; Hinterndorfer, M.; Brunner, H.; Grishkovskaya, I.; Singh, K.; Schleiffer, A.; Jude, J.; Deswal, S.; Kalis, R.; Vunjak, M.; Lendl, T.; Imre, R.; Roitinger, E.; Neumann, T.; Kandolf, S.; Schutzbier, M.; Mechtler, K.; Versteeg, G. A.; Haselbach, D.; Zuber, J. AKIRIN2 controls the nuclear import of proteasomes in vertebrates. *Nature* **2021**, *599* (7885), 491−496.
(235) Brockmann, M.; Blomen, V. A.; Nieuwenhuis, J.; Stickel, E.; Raaben, M.; Bleijerveld, O. B.; Altelaar, A. F. M.; Jae, L. T.; Brummelkamp, T. R. Genetic wiring maps of single-cell protein states reveal an off-switch for GPCR signalling. *Nature* **2017**, *546* (7657), 307−311.
(236) Schraivogel, D.; Kuhn, T. M.; Rauscher, B.; Rodriguez-Martinez, M.; Paulsen, M.; Owsley, K.; Middlebrook, A.; Tischer, C.; Ramasz, B.; Ordonez-Rueda, D.; Dees, M.; Cuylen-Haering, S.; Diebold, E.; Steinmetz, L. M. High-speed fluorescence image-enabled cell sorting. *Science* **2022**, *375* (6578), 315−320.
(237) Hanzl, A.; Casement, R.; Imrichova, H.; Hughes, S. J.; Barone, E.; Testa, A.; Bauer, S.; Wright, J.; Brand, M.; Ciulli, A.; Winter, G. E. Functional E3 ligase hotspots and resistance mechanisms to small-molecule degraders. *Nat. Chem. Biol.* **2022**, 1−11.
(238) Gosavi, P. M.; Ngan, K. C.; Yeo, M. J. R.; Su, C.; Li, J.; Lue, N. Z.; Hoenig, S. M.; Liau, B. B. Profiling the Landscape of Drug Resistance Mutations in Neosubstrates to Molecular Glue Degraders. *ACS Cent Sci.* **2022**, *8* (4), 417−429.
(239) Maniaci, C.; Ciulli, A. Bifunctional chemical probes inducing protein-protein interactions. *Curr. Opin Chem. Biol.* **2019**, *52*, 145−156.
(240) Vakser, I. A. Challenges in protein docking. *Curr. Opin Struct Biol.* **2020**, *64*, 160−165.
(241) Negroni, J.; Mosca, R.; Aloy, P. Assessing the applicability of template-based protein docking in the twilight zone. *Structure* **2014**, *22* (9), 1356−1362.
(242) Desta, I. T.; Porter, K. A.; Xia, B.; Kozakov, D.; Vajda, S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure* **2020**, *28* (9), 1071−1081.e3.
(243) Kong, R.; Wang, F.; Zhang, J.; Wang, F.; Chang, S. CoDockPP: A Multistage Approach for Global and Site-Specific Protein-Protein Docking. *J. Chem. Inf Model* **2019**, *59* (8), 3556−3564.
(244) Orasch, O.; Weber, N.; Muller, M.; Amanzadi, A.; Gasbarri, C.; Trummer, C. Protein-Protein Interaction Prediction for Targeted Protein Degradation. *Int. J. Mol. Sci.* **2022**, *23* (13), 7033.
(245) Schneider, M.; Radoux, C. J.; Hercules, A.; Ochoa, D.; Dunham, I.; Zalmas, L. P.; Hessler, G.; Ruf, S.; Shanmugasundaram, V.; Hann, M. M.; Thomas, P. J.; Queisser, M. A.; Benowitz, A. B.; Brown, K.; Leach, A. R. The PROTACtable genome. *Nat. Rev. Drug Discov* **2021**, *20* (10), 789−797.
(246) Hu, L.; Wang, X.; Huang, Y. A.; Hu, P.; You, Z. H. A survey on computational models for predicting protein-protein interactions. *Brief Bioinform* **2021**, *22* (5), bbab036.
(247) Gao, M.; Nakajima, A. D.; Parks, J. M.; Skolnick, J. AF2Complex predicts direct physical interactions in multimeric proteins with deep learning. *Nat. Commun.* **2022**, *13* (1), 1744.
(248) Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; Ronneberger, O.; Bodenstein, S.; Zielinski, M.; Bridgland, A.; Potapenko, A.; Cowie, A.; Tunyasuvunakool, K.; Jain, R.; Clancy, E.; Kohli, P.; Jumper, J.; Hassabis, D. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* **2022**, DOI: 10.1101/2021.10.04.463034.
(249) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **2022**, *378* (6615), 49−56.
(250) Rao, A.; Tunjic, T. M.; Brunsteiner, M.; Müller, M.; Fooladi, H.; Weber, N. Bayesian Optimization for Ternary Complex Prediction (BOTCP). *bioRxiv* **2022**, DOI: 10.1101/2022.06.03.494737.
(251) Kozicka, Z.; Thoma, N. H. Haven't got a glue: Protein surface variation for the design of molecular glue degraders. *Cell Chem. Biol.* **2021**, *28* (7), 1032−1047.
(252) Domostegui, A.; Nieto-Barrado, L.; Perez-Lopez, C.; Mayor-Ruiz, C. Chasing molecular glue degraders: screening approaches. *Chem. Soc. Rev.* **2022**, *51* (13), 5498−5517.
(253) St-Cyr, D.; Ceccarelli, D. F.; Orlicky, S.; van der Sloot, A. M.; Tang, X.; Kelso, S.; Moore, S.; James, C.; Posternak, G.; Coulombe-Huntington, J.; Bertomeu, T.; Marinier, A.; Sicheri, F.; Tyers, M. Identification and optimization of molecular glue compounds that inhibit a noncovalent E2 enzyme-ubiquitin complex. *Sci. Adv.* **2021**, *7* (44), eabi5797.
(254) Garcia-Seisdedos, H.; Empereur-Mot, C.; Elad, N.; Levy, E. D. Proteins evolve on the edge of supramolecular self-assembly. *Nature* **2017**, *548* (7666), 244−247.
(255) Nishiguchi, G.; Keramatnia, F.; Min, J.; Chang, Y.; Jonchere, B.; Das, S.; Actis, M.; Price, J.; Chepyala, D.; Young, B.; McGowan, K.; Slavish, P. J.; Mayasundari, A.; Jarusiewicz, J. A.; Yang, L.; Li, Y.; Fu, X.; Garrett, S. H.; Papizan, J. B.; Kodali, K.; Peng, J.; Pruett Miller, S. M.; Roussel, M. F.; Mullighan, C.; Fischer, M.; Rankovic, Z. Identification of Potent, Selective, and Orally Bioavailable Small-Molecule GSPT1/2 Degraders from a Focused Library of Cereblon Modulators. *J. Med. Chem.* **2021**, *64* (11), 7296−7311.
(256) Ran, X.; Gestwicki, J. E. Inhibitors of protein-protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Curr. Opin Chem. Biol.* **2018**, *44*, 75−86.
(257) Mabonga, L.; Kappo, A. P. Protein-protein interaction modulators: advances, successes and remaining challenges. *Biophys Rev.* **2019**, *11* (4), 559−581.
(258) Martinez-Jimenez, F.; Muinos, F.; Lopez-Arribillaga, E.; Lopez-Bigas, N.; Gonzalez-Perez, A. Systematic analysis of alterations in the ubiquitin proteolysis system reveals its contribution to driver mutations in cancer. *Nat. Cancer* **2020**, *1* (1), 122−135.
(259) Sijbesma, E.; Visser, E.; Plitzko, K.; Thiel, P.; Milroy, L. G.; Kaiser, M.; Brunsveld, L.; Ottmann, C. Structure-based evolution of a promiscuous inhibitor to a selective stabilizer of protein-protein interactions. *Nat. Commun.* **2020**, *11* (1), 3954.
(260) Watson, E. R.; Novick, S. J.; Matyskiela, M. E.; Chamberlain, P. P.; de la Peña, A. H.; Zhu, J.-Y.; Tran, E.; Griffin, P. R.; Wertz, I. E.; Lander, G. C. Molecular glue CELMoD compounds are allosteric regulators of cereblon conformation. *bioRxiv* **2022**, *378*, 549.
(261) Ni, D.; Lu, S.; Zhang, J. Emerging roles of allosteric modulators in the regulation of protein-protein interactions (PPIs): A new paradigm for PPI drug discovery. *Med. Res. Rev.* **2019**, *39* (6), 2314−2342.
(262) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508* (7496), 331−9.
(263) Kozlovskii, I.; Popov, P. Spatiotemporal identification of druggable binding sites using deep learning. *Commun. Biol.* **2020**, *3* (1), 618.
(264) Lazar, T.; Martinez-Perez, E.; Quaglia, F.; Hatos, A.; Chemes, L. B.; Iserte, J. A.; Mendez, N. A.; Garrone, N. A.; Saldano, T. E.; Marchetti, J.; Rueda, A. J. V.; Bernado, P.; Blackledge, M.; Cordeiro, T. N.; Fagerberg, E.; Forman-Kay, J. D.; Fornasari, M. S.; Gibson, T. J.; Gomes, G. W.; Gradinaru, C. C.; Head-Gordon, T.; Jensen, M. R.; Lemke, E. A.; Longhi, S.; Marino-Buslje, C.; Minervini, G.; Mittag, T.; Monzon, A. M.; Pappu, R. V.; Parisi, G.; Ricard-Blum, S.; Ruff, K. M.; Salladini, E.; Skepo, M.; Svergun, D.; Vallet, S. D.; Varadi, M.; Tompa, P.; Tosatto, S. C. E.; Piovesan, D. PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res.* **2021**, *49* (D1), D404−D411.
(265) Nussinov, R.; Zhang, M.; Liu, Y.; Jang, H. AlphaFold, Artificial Intelligence (AI), and Allostery. *J. Phys. Chem. B* **2022**, *126* (34), 6372−6383.
(266) Stein, R. A.; McHaourab, H. S. SPEACH_AF: Sampling protein ensembles and conformational heterogeneity with Alphafold2. *PLoS Comput. Biol.* **2022**, *18* (8), e1010483.
(267) Tan, Z. W.; Tee, W. V.; Guarnera, E.; Berezovsky, I. N. AlloMAPS 2: allosteric fingerprints of the AlphaFold and Pfam-

trRosetta predicted structures for engineering and design. *Nucleic Acids Res.* **2023**, *51* (D1), D345−D351.
(268) Wong, J. H.; Alfatah, M.; Sin, M. F.; Sim, H. M.; Verma, C. S.; Lane, D. P.; Arumugam, P. A yeast two-hybrid system for the screening and characterization of small-molecule inhibitors of protein-protein interactions identifies a novel putative Mdm2-binding site in p53. *BMC Biol.* **2017**, *15* (1), 108.
(269) Li, F.; Hu, Q.; Zhang, X.; Sun, R.; Liu, Z.; Wu, S.; Tian, S.; Ma, X.; Dai, Z.; Yang, X.; Gao, S.; Bai, F. DeepPROTACs is a deep learning-based targeted degradation predictor for PROTACs. *Nat. Commun.* **2022**, *13* (1), 7133.
(270) Douglass, E. F., Jr; Miller, C. J. A comprehensive kinetic model for ternary complex catalysis. *ChemRxiv* **2022**, DOI: 10.26434/chemrxiv-2021-wgn82-v2.
(271) Schneider, P.; Walters, W. P.; Plowright, A. T.; Sieroka, N.; Listgarten, J.; Goodnow, R. A., Jr.; Fisher, J.; Jansen, J. M.; Duca, J. S.; Rush, T. S.; Zentgraf, M.; Hill, J. E.; Krutoholow, E.; Kohler, M.; Blaney, J.; Funatsu, K.; Luebkemann, C.; Schneider, G. Rethinking drug design in the artificial intelligence era. *Nat. Rev. Drug Discov* **2020**, *19* (5), 353−364.
(272) Jimenez-Luna, J.; Skalic, M.; Weskamp, N.; Schneider, G. Coloring Molecules with Explainable Artificial Intelligence for Preclinical Relevance Assessment. *J. Chem. Inf Model* **2021**, *61* (3), 1083−1094.
(273) Rao, J.; Zheng, S.; Lu, Y.; Yang, Y. Quantitative evaluation of explainable graph neural networks for molecular property prediction. *Patterns (N Y)* **2022**, *3* (12), 100628.
(274) Jiménez-Luna, J.; Grisoni, F.; Schneider, G. Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence* **2020**, *2* (10), 573−584.
(275) Rodriguez-Perez, R.; Bajorath, J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *J. Comput. Aided Mol. Des* **2020**, *34* (10), 1013−1026.
(276) Wellawatte, G. P.; Seshadri, A.; White, A. D. Model agnostic generation of counterfactual explanations for molecules. *Chem. Sci.* **2022**, *13* (13), 3697−3705.
(277) Yu, J.; Wang, D.; Zheng, M. Uncertainty quantification: Can. we trust artificial intelligence in drug discovery? *iScience* **2022**, *25* (8), 104814.
(278) Alvarsson, J.; Arvidsson McShane, S.; Norinder, U.; Spjuth, O. Predicting With Confidence: Using Conformal Prediction in Drug Discovery. *J. Pharm. Sci.* **2021**, *110* (1), 42−49.
(279) Wang, Z.; Lachmann, A.; Keenan, A. B.; Ma'ayan, A. L1000FWD: fireworks visualization of drug-induced transcriptomic signatures. *Bioinformatics* **2018**, *34* (12), 2150−2152.
(280) Hie, B.; Bryson, B. D.; Berger, B. Leveraging Uncertainty in Machine Learning Accelerates Biological Discovery and Design. *Cell Syst* **2020**, *11* (5), 461−477 e9.
(281) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Zidek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50* (D1), D439−D444.
(282) Callaway, E. 'The entire protein universe': AI predicts shape of nearly every known protein. *Nature* **2022**, *608* (7921), 15−16.
(283) Haibe-Kains, B.; Adam, G. A.; Hosny, A.; Khodakarami, F.; Massive Analysis Quality Control Society Board of, D.; Waldron, L.; Wang, B.; McIntosh, C.; Goldenberg, A.; Kundaje, A.; Greene, C. S.; Broderick, T.; Hoffman, M. M.; Leek, J. T.; Korthauer, K.; Huber, W.; Brazma, A.; Pineau, J.; Tibshirani, R.; Hastie, T.; Ioannidis, J. P. A.; Quackenbush, J.; Aerts, H. Transparency and reproducibility in artificial intelligence. *Nature* **2020**, *586* (7829), E14−E16.
(284) Bansal, M.; Yang, J.; Karan, C.; Menden, M. P.; Costello, J. C.; Tang, H.; Xiao, G.; Li, Y.; Allen, J.; Zhong, R.; Chen, B.; Kim, M.; Wang, T.; Heiser, L. M.; Realubit, R.; Mattioli, M.; Alvarez, M. J.; Shen, Y.; Community, N.-D.; Gallahan, D.; Singer, D.; Saez-Rodriguez, J.; Xie, Y.; Stolovitzky, G.; Califano, A.; Community, N.-D. A community computational challenge to predict the activity of pairs of compounds. *Nat. Biotechnol.* **2014**, *32* (12), 1213−22.
(285) Huang, K.; Fu, T.; Gao, W.; Zhao, Y.; Roohani, Y.; Leskovec, J.; Coley, C. W.; Xiao, C.; Sun, J.; Zitnik, M. Artificial intelligence foundation for therapeutic science. *Nat. Chem. Biol.* **2022**, *18* (10), 1033−1036.

https://doi.org/10.1021/jacs.2c11098
J. Am. Chem. Soc. 2023, 145, 2711−2732