RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Cite this: *RSC Chem. Biol.*, 2021, 2, 354

Fragment-based covalent ligand discovery

Wenchao Lu, †[ab] Milka Kostic, †[a] Tinghu Zhang,[ab] Jianwei Che,[abc] Matthew P. Patricelli,[d] Lyn H. Jones,[c] Edward T. Chouchani[ae] and Nathanael S. Gray *[ab]

Received 4th December 2020,
Accepted 20th January 2021

DOI: 10.1039/d0cb00222d

rsc.li/rsc-chembio

Targeted covalent inhibitors have regained widespread attention in drug discovery and have emerged as powerful tools for basic biomedical research. Fueled by considerable improvements in mass spectrometry sensitivity and sample processing, chemoproteomic strategies have revealed thousands of proteins that can be covalently modified by reactive small molecules. Fragment-based drug discovery, which has traditionally been used in a target-centric fashion, is now being deployed on a proteome-wide scale thereby expanding its utility to both the discovery of novel covalent ligands and their cognate protein targets. This powerful approach is allowing 'high-throughput' serendipitous discovery of cryptic pockets leading to the identification of pharmacological modulators of proteins previously viewed as "undruggable". The reactive fragment toolkit has been enabled by recent advances in the development of new chemistries that target residues other than cysteine including lysine and tyrosine. Here, we review the emerging area of covalent fragment-based ligand discovery, which integrates the benefits of covalent targeting and fragment-based medicinal chemistry. We discuss how the two strategies synergize to facilitate the efficient discovery of new pharmacological modulators of established and new therapeutic target proteins.

Introduction

Inspired by recent approval of covalent kinase inhibitors (TKIs) for cancer treatment including inhibitors targeting EGFR: afatinib (Gilotrif) and osimertinib (Tagrisso) or BTK: acalabrutinib (Calquence) and ibrutinib (Imbruvica), the development of covalent probes and drugs has undergone a renaissance and now attracts intense interest from both industry and academia.[1–4] Unlike noncovalent small molecules that target conserved substrate and/or allosteric binding site transiently, covalent inhibitors often exhibit differentiated pharmacology in terms of potency, selectivity, pharmacokinetics and pharmacodynamics as a consequence of their ability to form irreversible, covalent bonds with their target proteins.[5,6] Despite these advantages, many viewed covalent inhibitors skeptically due to their ability to result in protein adducts capable of triggering idiosyncratic immune responses and allergic/hypersensitivity reactions.[7,8] Historically the discovery of covalent drugs was often serendipitous based on natural products such as the beta-lactam class of antibiotics exemplified by the discovery of penicillin or based on rational structure-based design from non-covalent inhibitors as exemplified by the development of covalent kinase inhibitors targeting EGFR or BTK. However, there has been a resurgence of interest in covalent inhibitors that has resulted in a host of new approaches for discovering and developing covalent inhibitors, resulting in 7 covalent drugs (from the total of 161 small molecule drugs) approved by the FDA from 2015 to 2019.[9]

Rational covalent inhibitor design usually starts with a known noncovalent binder and explores strategies to incorporate an appropriate electrophilic warhead to achieve desired target selectivity and efficacy ("binder-first" approaches).[2,3,10] While this approach has been successful, it is limited to targets that: (1) have existing ligands amenable to further derivatization with a reactive warhead; and (2) include a suitably reactive residue within or near the noncovalent ligand binding site.[11] Moreover, although recent chemoproteomic studies have identified additional protein targets that feature at least one reactive residue suitable for covalent strategies, many of these sites are located within cryptic ligand pockets, for which reversible ligand discovery remains difficult.

Fragment-based drug discovery (FBDD) has been widely applied as an alternative to large chemical library screening for ligand discovery, especially in the context of intractable biological targets.[12–14] The FBDD approach for discovery of noncovalent binders involves assembling a library of low

[a] *Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA. E-mail: nathanael_gray@dfci.harvard.edu*
[b] *Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02215, USA*
[c] *Center for Protein Degradation, Dana-Farber Cancer Institute, Boston, MA 02215, USA*
[d] *Vividion Therapeutics, La Jolla, CA 92121, USA*
[e] *Department of Cell Biology, Harvard Medical School, Boston, MA 02215, USA*
† These authors contributed equally.

molecular weight ligands (typical MW of less than 300 Da) called fragments, which are screened against a purified target of interest to identify low affinity binders with high ligand efficiency (LE). The fragments are subsequently elaborated and optimized to yield higher affinity ligands.[15] Typically, FBDD screening can yield a lead from a limited size library (few thousand fragments *vs.* hundred thousand and larger libraries typically used for high-throughput screening of drug-like libraries). Another advantage of fragments is their ability to access cryptic sites beyond substrate pockets, which holds great promise for allosteric drug discovery and discovery of ligands for targets that lack well-defined binding pockets.[16] However, one of the liabilities of fragment-based strategies is the low intrinsic binding affinity of the fragments, making these interactions sometimes difficult to detect. Additionally, fragment optimization into higher affinity, selective ligands is difficult, and places significant emphasis on the need for structural biology information. Lastly, validation of fragment binding and on-target mechanism of action *via* genetic methods is currently not possible.

One way to mitigate challenges related to both FBDD and the ''binder-first'' approaches for covalent ligand discovery is to use covalent fragments in FBDD. Covalent fragments have a potential to achieve excellent target engagement *via* covalent bond formation, and selectivity by reacting with distinct target residues. Additionally, due to covalent nature of their interaction with the target, target engagement and on-target mechanism of action can be unequivocally established using mass spectrometry and confirmed through target residue mutagenesis. Besides, optimization of covalent fragments could be more straightforward since their binding mode does not change easily during fragment merging/growing. These benefits of covalent fragments in drug discovery has resulted in increased interest in this area. Therefore, this review aims to provide an overview of the recent efforts to discover, develop and validate covalent fragments as a promising strategy for covalent ligand discovery. We will discuss ways to discover new covalent fragments, as well as how covalent fragments are used in chemoproteomics for: (1) mapping proteomic reactivity of the targets; (2) identifying new hotspots for inhibitor/ligand development; and (3) mapping compound selectivity. We will also provide a more comprehensive overview of the recent advances in expanding the repertoire of electrophiles towards protein nucleophiles other than cysteine and highlight some important chemical biology applications of covalent fragments (Fig. 1). We would like to note that several excellent reviews focusing on covalent fragment library design have recently been published,[1,2,17–19] and we will not comment on this here.

Strategies for covalent fragment discovery

Target-based covalent fragment discovery

A widely applied covalent fragment screening strategy is called site-directed disulfide tethering, also known as disulfide trapping.[19–21] The tethering compound usually contains a cysteamine group to increase compound solubility and a disulfide group to capture naturally occurring or proteins with a cysteine introduced by site-directed mutagenesis through rapid thiol exchange under reducing conditions. The disulfide part of hit compounds is typically replaced by a more stable moiety or other cysteine-directing warheads like acrylamides in follow-up studies in order to improve drug-like properties. Given its convenience and versatility, the covalent tethering method has been widely adopted in drug discovery programs for different protein classes including enzymes, G protein coupled receptors (GPCRs) and protein–protein interactions (PPIs).[22–28] Perhaps the most successful application of disulfide tethering was the identification of fragments targeting allele-specific K-Ras^{G12C} oncoprotein.[29] Genetic alterations of Ras family proteins are the most prevalent oncogenic mutations in cancer cells. However, Ras oncoproteins has been considered to be undruggable due to the picomolar binding affinity for GTP substrate.[30] In 2013, Ostrem *et al.* screened a set of 480 tethering fragments against K-Ras^{G12C} *in vitro* by intact protein mass spectrometry.[31] The screen led to the identification of fragment 6H05 that could bind to the allosteric site of K-Ras^{G12C} protein, in a ligand-induced pocket named the switch-II pocket with great selectivity over wildtype K-Ras due to the unique G12C residue. Iterative chemistry optimization, including replacement of the disulfide group with more stable carbon-based electrophiles like acrylamides, resulted in promising lead compounds such as compound 12. Importantly, these studies served as a proof-of-ligandability concept for K-Ras^{G12C}, which inspired numerous groups to develop further optimized inhibitors,[32–36] including AMG-510 (proposed INN sotorasib), which is the first drug candidate entering clinical stage for the treatment of advanced/metastatic solid tumors with KRAS^{G12C} mutation (NCT03600883, NCT04380753), and MRTX849 (adagrasib) (Fig. 2A). These studies have contributed significantly to the increased interests in using covalent fragments for drug discovery and development.

Although covalent tethering has been successfully used for fragment discovery, disulfide-containing fragment libraries are not generally available from commercial vendors. Additionally, considering the metabolic liability of disulfide group and complex redox environment *in vivo*, this strategy requires significant chemistry efforts post hit identification to replace the disulfide group with more suitable warheads.[41] Therefore, additional electrophilic fragment screening strategies have been developed in recent years.[42–47] For example, Johansson *et al.* recently designed and synthesized a 104-fragment library of α,β-unsaturated methyl ester warheads, which have been demonstrated to exhibit narrower reactivity profile compared with other commonly used electrophiles, such as acrylamides.[37] The fragments were screened against the E3 ligase HOIP RBR domain in 22 pools by LC-MS assays. Among the hits, compound 5 (Fig. 2B) was identified as a covalent hit that targets catalytic Cys885 with k_{inact}/K_I value of 0.97 ± 0.01 M^{-1} s^{-1} (k_{inact}/K_I value is an important measure of covalent inhibitor performance, and we refer readers interested in learning more to a recent publication

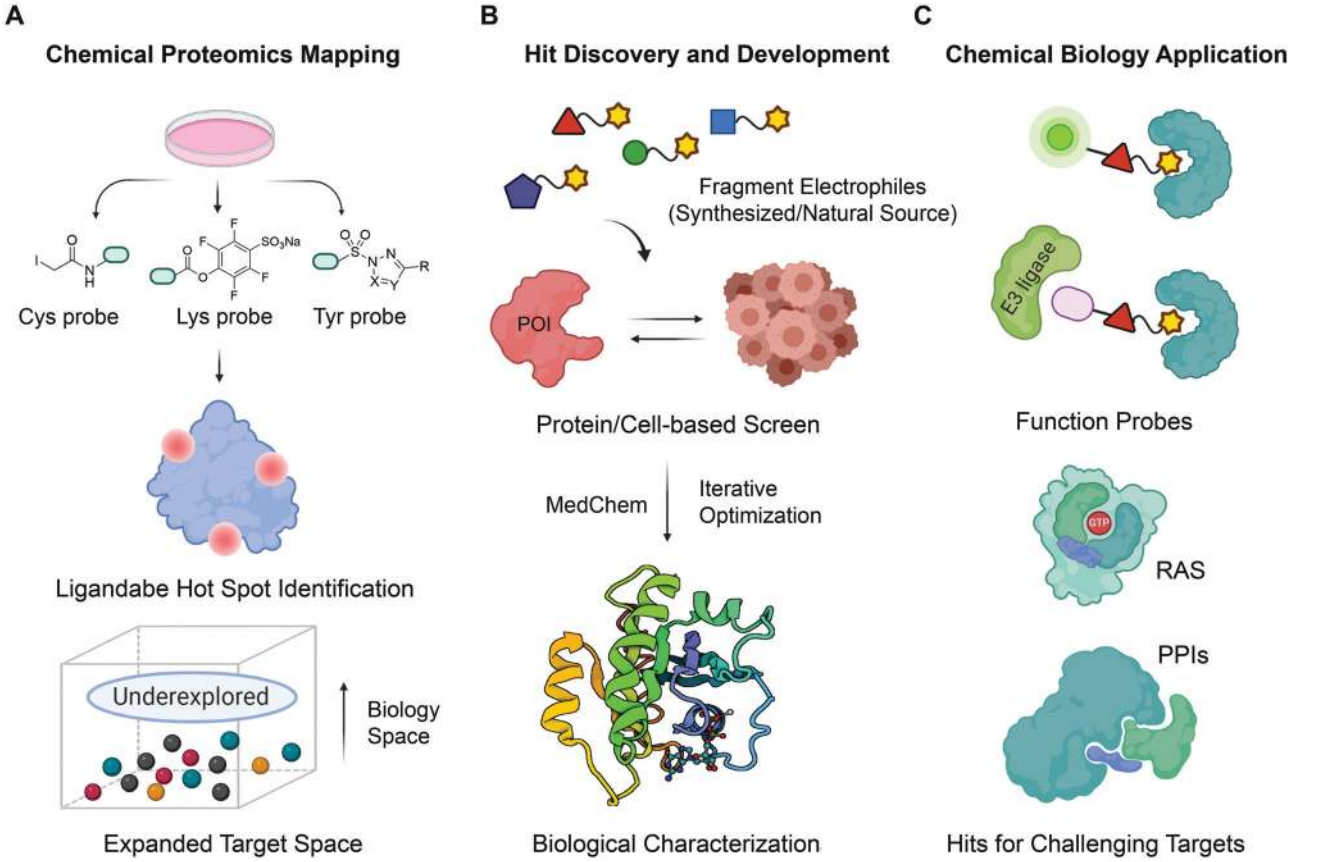

Fig. 1 The roadmap for fragment-based covalent ligand discovery. (A) State-of-the-art chemoproteomic strategies helps to systematically unveil potential ligandable sites in disease-associated targets. Cells are treated with reactive fragments (biased towards thiols (Cys), amines (lysine) and phenols (tyrosine)) and then chemoproteomics allows identification of proteins and reaction sites. (B) Fragment-based covalent ligand screening identified covalent fragment hits, which can be evolved into a more potent, selective, biocompatible and drug-like ligands by iterative elaboration and optimization. (C) Fragment-based ligand discovery pipeline holds great promise as an initial ligand-discovery approach that can be elaborated to make bivalent molecules that can recruit other enzymes including E3s for PROTACs *etc*.

on this topic[48,49]). The compounds showed great selectivity against a panel of RBR E3 ligases, HECT E3 ligases, E1/E2 enzymes and deubiquitinating enzymes. Further activity-based protein profiling (ABPP; Fig. 3A) study using a slightly more potent derivative demonstrated the on-target effect in cells. This work shows the potential of fragment-based covalent ligand screening for first-in-class drug discovery against targets like E3 ligases, which have traditionally been challenging. Moreover, given the importance of novel E3 ligase ligands for targeted degrader development,[50] strategies like this one could substantially expand the toolbox of E3 ubiquitin ligases that can be exploited in this context.

Another example aimed at finding covalent fragments for difficult-to-drug cysteine-containing proteins, used a larger fragment library of 993 covalent building blocks containing acrylamides or chloroacetamides to screen against 10 proteins, including bovine serum albumin (BSA) as a negative control.[38] Among those targets, the intact protein MS screen led to discovery of 37 strictly nonpromiscuous hits with >50% labeling efficiency. Streamlined high throughput crystallographic studies resulted in the successful determination of 15 inhibitor-protein complex structures. Further fragment analogue evaluation yielded the best inhibitor OTUB2-COV-1 (Fig. 2C) with the k_{inact}/K_I value of 3.75 M^{-1} s^{-1}. IsoTOP-ABPP demonstrated that OTUB2-COV-1 (10 μM compound, 2 h incubation) labeled <1% probe-accessible cysteines (Heavy/Light ratio >4) in HEK293T cells suggesting promising selectivity profile suitable for further development. Additionally, screening against pyrophosphohydrolase NUDT7 yielded 20 strictly nonpromiscuous hits with a total hit rate of 2%, which did not label two or more proteins by more than 30%. Fragment merging with previously identified noncovalent hit led to the discovery of NUDT7-COV-1 (Fig. 2C) with the k_{inact}/K_I value of 757 M^{-1} s^{-1}. Cellular target engagement was further validated by cellular thermal shift assay, providing additional support that these types of screens of limited covalent fragment libraries can yield high quality leads for development of selective ligands for challenging targets.

Very recently, screening the same library of 993 fragments with acrylamides or chloroacetamides warheads followed with rational chemistry optimization, led to the discovery of a potent and selective covalent Pin1 inhibitor, called Sulfopin with the k_{inact}/K_I value of 84 M^{-1} s^{-1} (Fig. 2D).[39] Pin1, a peptidyl-prolyl *cis–trans* isomerase, has been a challenging target and Sulfopin therefore represents an important addition to the currently

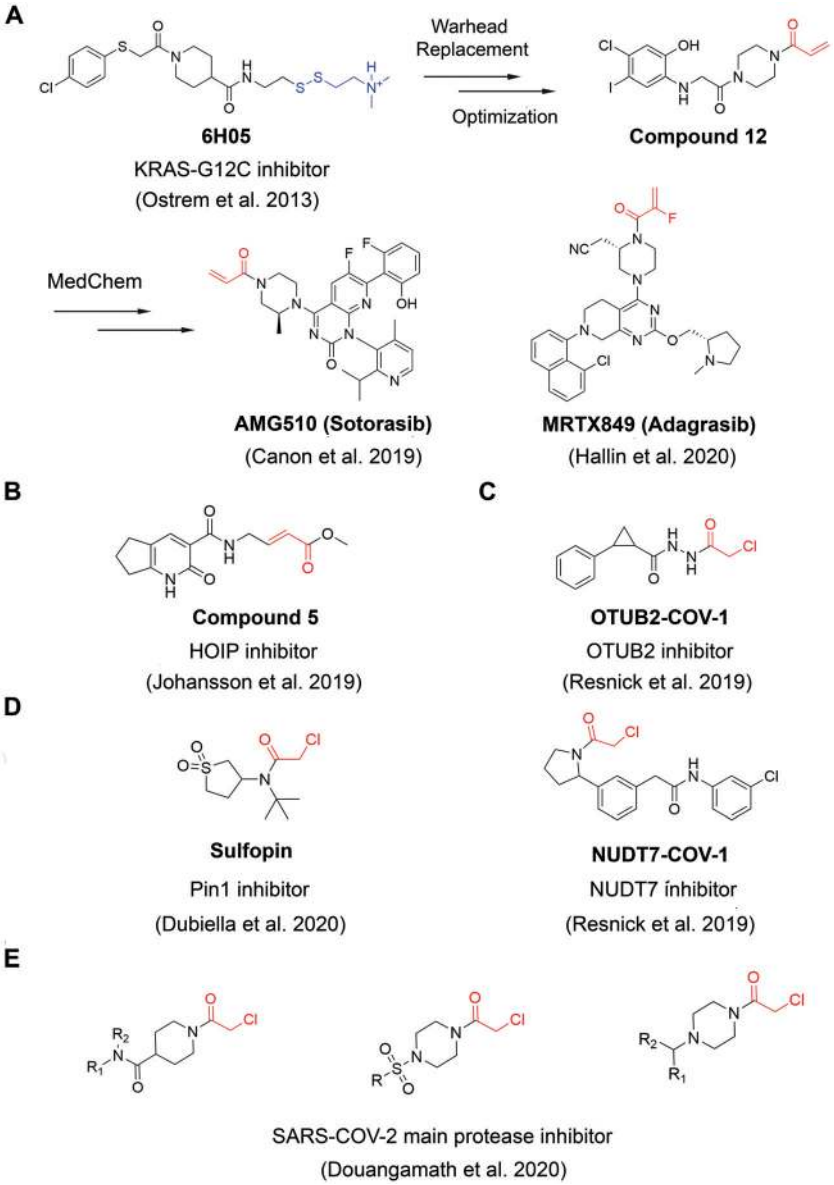

Fig. 2 The structures of representative well-characterized electrophilic fragments identified from target-based screening strategies in recent years. (A) KRAS-G12C allele-specific covalent fragment (6H05) identified from tethering screen, which was further elaborated to compound 12.[31] This inspired numerous groups to develop further optimized inhibitors, within which AMG510[33] and MRTX849[36] successfully entered clinical trials. (B) Compound 5 targets the active cysteine (C885) of HOIP.[37] (C) OTUB2-COV-1 targets the active cysteine (C51) of OTUB2 and NUDT7-COV-1 target C73 of NUDT7.[38] (D) Sulfopin targets the active cysteine of Pin1 (C113).[39] (E) Representative covalent fragment scaffolds target the active cysteine (C145) of SARS-COV-2 main protease (M^{pro}).[40]

available Pin1 ligands. To confirm the mode of binding, crystallographic studies clearly demonstrated that Sulfopin binds to Pin1 covalently to Cys113. Despite the relatively small size of the fragment and the typically more reactive chloroacetamide electrophile, Sulfopin exhibited impressive selectivity in chemoproteomic experiments and significantly retarded neuroblastoma initiation in zebrafish tumor model.

Using a compound library with over 1250 fragments, Douangamath *et al.* conducted a combined mass spectrometry and high throughput crystallographic fragment screen against SARS-CoV-2 main protease (M^{pro}), which is an essential target for viral replication. This screen finally yielded 48 high-value covalent fragments co-crystallized with M^{pro} featuring a variety of scaffolds, which offered unprecedented structural resources for the follow-up structure-based anti-viral drug discovery[40] (Fig. 2E).

Besides the hit identification, covalent fragment-based approaches are also suitable for mapping the protein hot-spots. Very recently, Petri *et al.* described a library of covalent fragments with identical scaffold but chemically diverse electrophilic warheads. They screened this collection against a panel of kinases with at least one targetable cysteine (BTK, ERK2, RSK2 and MAP2K6), and used JAK3 and MELK for further validation studies. This study demonstrated that covalent fragments can be used to

map cysteine tractability across a wide range of targets[51,52] Collectively, these recent examples illustrate the robustness and efficiency of electrophilic fragment screening, as well as the potential for identifying non-promiscuous covalent hits and evaluating the accessibility and reactivity of targeted cysteines.

Cell-based covalent fragments discovery

Although target-based electrophilic library screens described above are effective and relatively easy to implement, they require production and isolation of stable recombinant target protein. For many targets of interest, this is either not feasible on a scale needed for these experiments or impossible given intrinsic instability of the target. Additionally, in many cases selectivity against isolated protein does not directly translate into global chemoselectivity in cellular environment. To address this, several cell-based screens of electrophilic fragment libraries using state-of-art chemoproteomic strategies have recently been conducted (Fig. 3).[53]

In 2016, Cravatt *et al.* used cysteine-reactive ABPP probes (Fig. 4A) to conduct proteome-wide screening of 56 cysteine-directing fragments containing chloroacetamide or acrylamide electrophiles. The screen demonstrated that ~20% of quantified proteins harbored ligandable cysteine including transcription factors, adaptor and scaffold proteins. One of the hits identified both *in vitro* and *in situ* is pro-caspase 8/10 fragment that could selectively modify pro-caspase 8 and not the active form, caspase 8. Further chemistry efforts led to the development of selective caspase 8 inhibitor, thus demonstrating that covalent fragment hits identified in this manner are readily optimizable.[54]

Following on this work, Cravatt and colleagues have expanded the scope of their studies further. For example, by using broadly cysteine-reactive fragments as ''scout fragments'',[54] they examined proteomic cysteine ligandability in *KEAP1*-mutant and *KEAP1*-wild-type (WT) human non-small cell lung cancer (NSCLC) cell lines. Among more than 1000 cysteine sites that exhibited sensitivity to ''scout fragments'', the

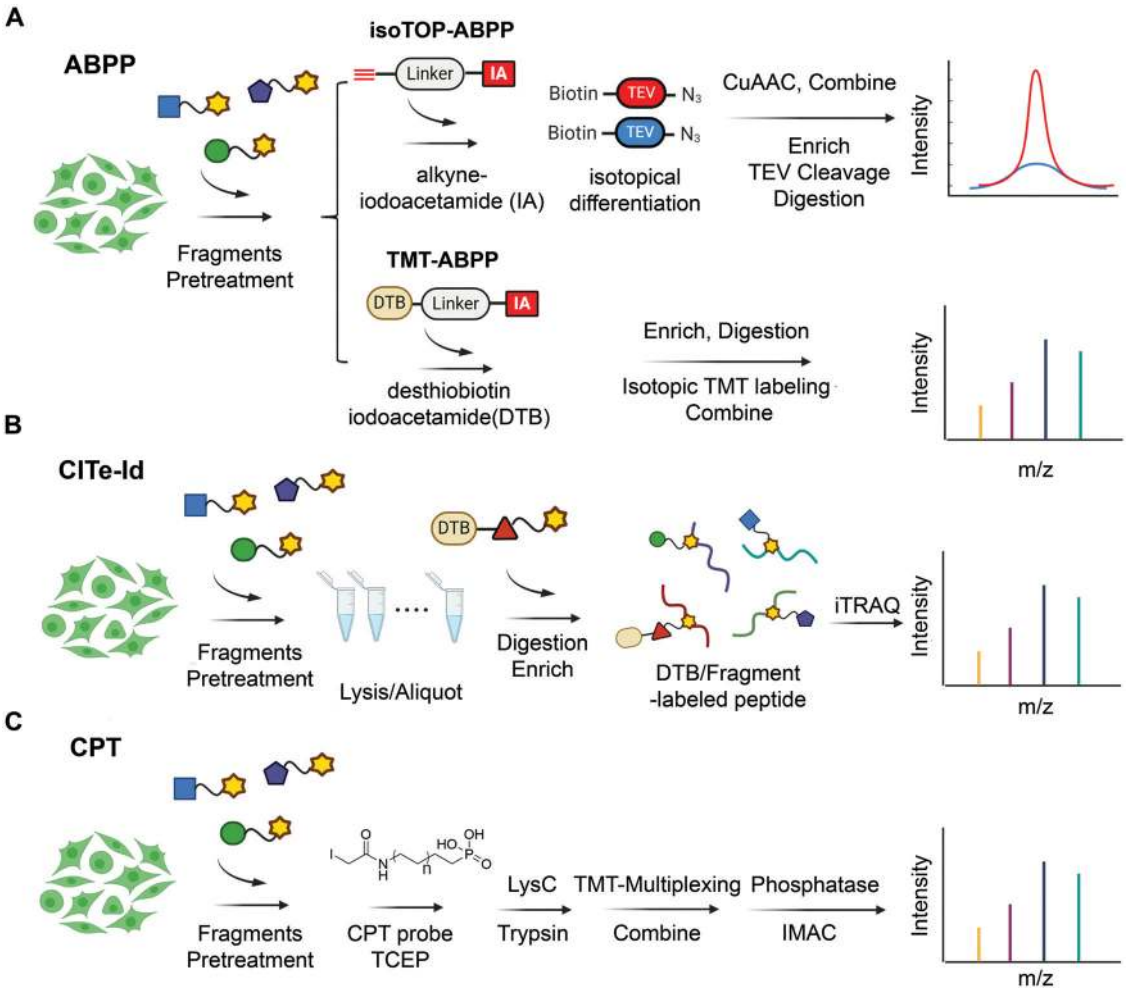

Fig. 3 State-of-the-art chemoproteomic approaches for reimaging druggable proteome and selectivity profiling of covalent fragments. (A) Activity-based protein profiling (IsoTOP-ABPP and TMT-ABPP) using isotope-labeled probes or isotopic TMT labeling agents for multiplex quantitative chemoproteomics.[39,54–57] (B) Covalent Inhibitor Target-site Identification (CITe-Id), another complementary chemoproteomic platform to understand the proteome-wide on/off-target effect in covalent fragment development program, in which desthiobiotinylated covalent inhibitor is used in lieu of non-selective iodoacetamide probe to directly monitor target engagement.[58] (C) Cysteine-Reactive Phosphate Tags (CPTs) developed recently can be applied as a chemoproteomic using phosphate-tagged iodoacetamide for global cysteine profiling with high coverage of the cysteine proteome.[59]

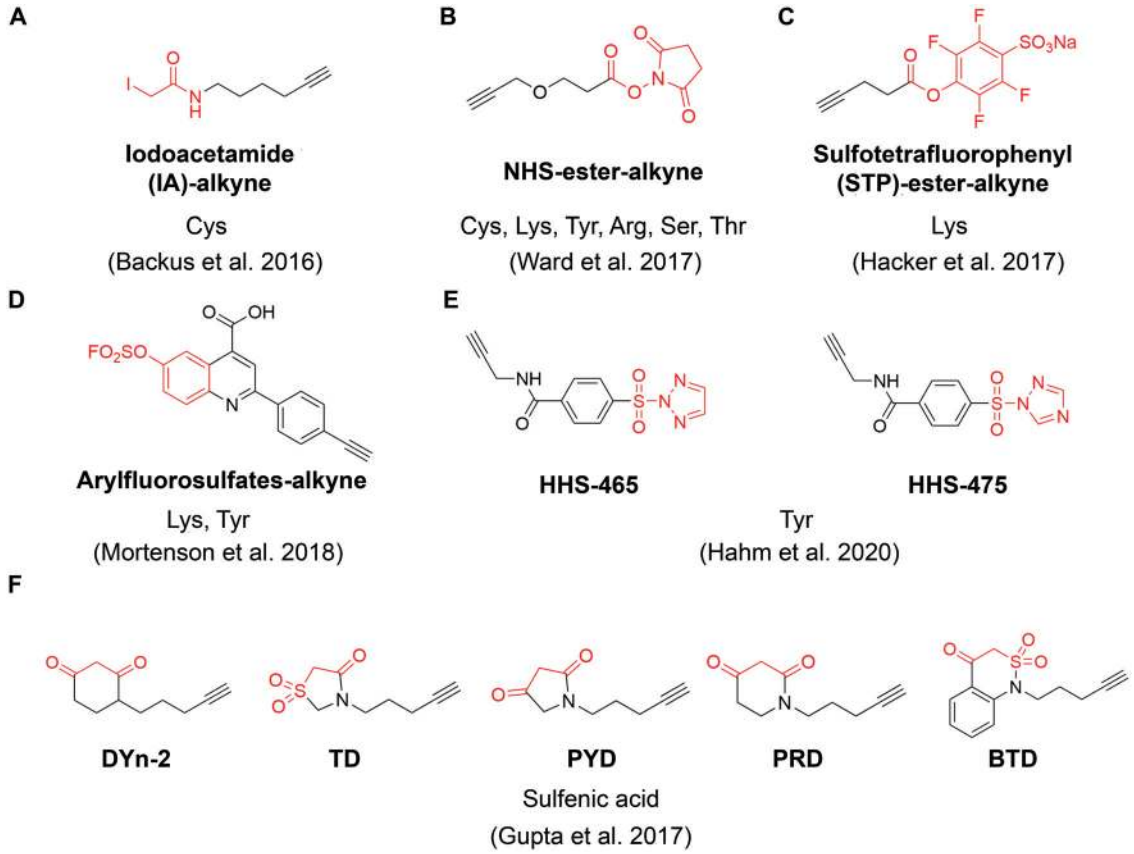

Fig. 4 The structures of representative ABPP probes applied in state-of-the-art chemoproteomic strategies. (A) Iodoacetamide-alkyne is a widely applied cysteine-reactive ABPP probes.[54] (B) NHS-ester-alkyne is a versatile covalent ABPP probe for protein nucleophiles (cysteines, lysines, tyrosines, serines, threonines, and arginines).[65] (C) STP-alkyne is an amine-reactive covalent ABPP probe.[67] (D) Latent electrophiles as exemplified by SuFEx-based arylfluorosulfates for lysine/tyrosine targeting.[72] (E) HHS-465 and HHS-475 are SuTEx-chemistry based ABPP probes for tyrosine profiling.[74] (F) Carbon nucleophiles help to explore ligandability of sulfenylated (oxidized) cysteines.[75]

authors detected orphan nuclear hormone receptor NR0B1 (modified on Cys247). Follow up analysis demonstrated that NR0B1 expression was associated with KEAP1 mutational status, and that covalent inhibitor targeting Cys247 inhibited anchorage-independent NSCLC cell growth.[60] Very recently, they further investigated ligandability of cysteines in primary human T cells using a similar electrophilic "scout fragment" strategy.[61] This work generated a view of cell-state dependent ligandable cysteines in human T cells, and the study further suggested that immunomodulatory compounds can be obtained rapidly *via* elaboration of electrophilic compounds. Together, these recent results demonstrate the great potential of phenotypic electrophilic fragment screening for identifying lead fragments for further optimization as well as additional ligandable targets suitable for covalent inhibitor development.

Although covalent targeting cysteine has achieved great success, cysteine is one of the least abundant amino acids with ~2% occurrence in human proteome, limiting the number of potentially targetable sites.[62,63] Furthermore, cancer cells are able to acquire resistance through simple oxidation/mutation of the target cysteine, thus making the drug ineffective.[64] Therefore, in addition to cysteine side chains, lysines, threonines, serines, tyrosines and potentially other side chains may represent an opportunity for screening covalent fragments.

For example, Ward *et al.* deployed *N*-hydroxysuccinimide-ester (NHS-esters) (Fig. 4B) as versatile ABPP probes to profile ligandable protein nucleophiles (primarily lysines) in mouse liver proteomes.[65] This work led to the discovery of more than 3000 potential hotspots within the proteome that could potentially be explored for development of covalent ligands. Based on the competitive platform, the authors then screened NHS-ester containing fragments leading to the identification of selective covalent probes for dihydropyrimidine dehydrogenase (DPYD), aldehyde dehydrogenase (ALDH2), and glutathione S transferase theta 1 (GSTT1). Although the authors documented potential targeting of residues beyond cysteines and lysines (such as serines, threonines, tyrosines and arginines), it is important to note that NHS-ester reactivity strongly favors lysines, thus the potential for this group to be used for chemical probe development towards residues other than lysine is unclear. Nonetheless, the work further supports a view that side chains beyond cysteine can be used in this context, which would certainly expand the potential of covalent fragment based approaches.[66]

In another example, a lysine-specific, amine-reactive sulfotetrafluorophenyl (STP) alkyne probe (Fig. 4C) was used to screen ~30 lysine-reactive electrophilic fragments.[67] The study identified a number of ligandable lysine residues, as well as resulted in the identification of lysine-directed covalent ligands targeting enzymatic sites (PNPO-K100, NUDT2-K89), allosteric pockets (PFKP-K688) and protein–protein interactions (PPIs) (SIN3A-K115). Very recently, Wolter *et al.* described a concept of imine tethering, which uses fragments featuring aldehyde groups to react with lysine side chains.[68] The authors applied this strategy for discovery of PPI stabilizers (''molecular glues'') of 14-3-3/NF-κB interaction. Targeting of an interface lysine with unique (low) pK_a (Lys122) was confirmed using X-ray crystallography, and further optimization of the initial fragment lead resulted in two new aldehyde-containing stabilizers of this PPI, as validated in biochemical assays. Overall, this strategy may open additional opportunities for discovery of PPI stabilizers (''molecular glue'' compounds[69]) of other complexes.

Beyond lysine targeting, the development of SuFEx (Sulfur Fluoride Exchange) chemistry enables the covalent fragment screening against tyrosine in human proteome.[70,71] In 2018, Kelly group employed ''inverse drug discovery'' strategy using latent electrophiles as exemplified by aryl fluorosulfates fragments (Fig. 4D), to survey the ligandable sites in HEK293T cells.[72] This led to the identification and validation of covalent ligands for 11 important proteins. Very recently, they developed 16 structurally diverse sulfuramidimidoyl fluoride (SAF) functionalized probes to expand the SuFEx toolbox for inverse drug discovery in proteome. 72% of the protein targets identified by SAFs have not been previously identified in previous SuFEx-based chemoproteomics.[73]

Hsu *et al.* have recently adapted sulfur-triazole exchange (SuTEx) chemistry (Fig. 4E) for fragment-based ligand discovery of Tyr-directed binders.[74] By fine-tuning the SuTEx reactivity, they identified more than 10 000 SuTEx-reactive tyrosines and phosphotyrosines in human proteome. This suggests that SuTEx chemistry is suitable for developing Tyr-reactive covalent inhibitors as well as covalent fragments.

In addition to electrophilic covalent fragments, the Carroll group has recently reported a novel class of nucleophilic covalent fragments targeted at sulfenylated (oxidized) cysteines.[75] Partial oxidation of cysteine to the sulfenic acid is a physiological process that has been exploited to develop chemical probes that only recognize a particular oxidation state of a protein. These sulfenic acid-reactive covalent ABPP probes (Fig. 4F) has resulted in the identification of more than 1280 sulfenylated cysteines in the colon carcinoma cell line RKO, which shed light on the development of covalent fragments therapeutically targeting redox-active cysteines. Overall, in a way similar to covalent inhibitors, covalent fragments directed at different residues are achievable, and their use as chemoproteomic profiling probes in cell-based assays offers a promising strategy for identifying new targetable sites within the proteome.

We would like to note that so far one of the biggest barriers to all of the chemoproteomic approaches is throughput and the current published chemoproteomic methodologies are limited to very small libraries (usually <100 molecules). Another barrier is the stochastic nature on MS data acquisition that makes it difficult to interpret structure–activity relationship (SAR) trends when screening compound derivates collections. The above challenges are due to the fact that most chemoproteomic approaches to date rely on pairwise labeling strategies, which limit both throughput and consistent identification and quantification of target engagement proteome wide. The development of multiplexing-based chemoproteomic methodologies is an important step towards addressing the throughput issues. The most advanced of these technologies makes use of tandem mass tag (TMT) reagents, isobaric amine-reactive molecules that label N termini of peptides and the ε-amino groups of lysines. With both 10plex, and recently developed 16plex varieties, TMT labeling strategies provide an opportunity to greatly increase throughput and reproducible proteome coverage in chemoproteomic workflows.[76] One additional strategy to overcome these limitations is to employ a gene family approach, where screening is done against a complete (or nearly complete) gene family. One well-stablished example of this strategy is KiNativ™ platform for *in situ* kinase profiling, in which users can obtain complete coverage for target of interest in each run when screening a library, and achieve higher quantitative accuracy.[77]

As with noncovalent FBDD, many covalent fragments that emerge from these efforts remain far from ideal in terms of potency, thus requiring systematic medicinal chemistry optimization. We expect that continued advancement and implementation of chemoproteomic methods in terms of library size and associated throughput and closer collaboration between mass spectrometry professionals and medicinal chemists might reveal many more opportunities in drug discovery.

Photoactivation-assisted fragment discovery

The major limitation of chemoproteomic studies described thus far is that they can only identify targets featuring suitably reactive nucleophilic (primarily cysteine) residues, which may significantly narrow the target space. To circumvent this, several research groups developed fragment screening platforms that use photoactivatable groups, such as diazirine. In this case, a fragment remains inactive until activated by UV irradiation, which in the case of diazirines generates a reactive carbene intermediate that can react with any amino acid side chain or peptide backbone. Therefore, once bound to the protein pocket, the photoreactive fragment could crosslink to adjacent residues, resulting in an adduct that can be identified using MS (Fig. 5). Parker *et al.* used 14 photoactivatable (diazirine group-containing) fragment probes featuring different chemical structures to profile fragment target space in HEK293T cells.[78] Using quantitative MS-based proteomics, they identified more than 2000 protein targets, including many proteins for which no ligand has previously been reported in DrugBank. Using a similar strategy, researchers from GSK developed a fragment screen platform named Photo-Affinity Bits (PhABit) for rapid fragment hit identification.[79]

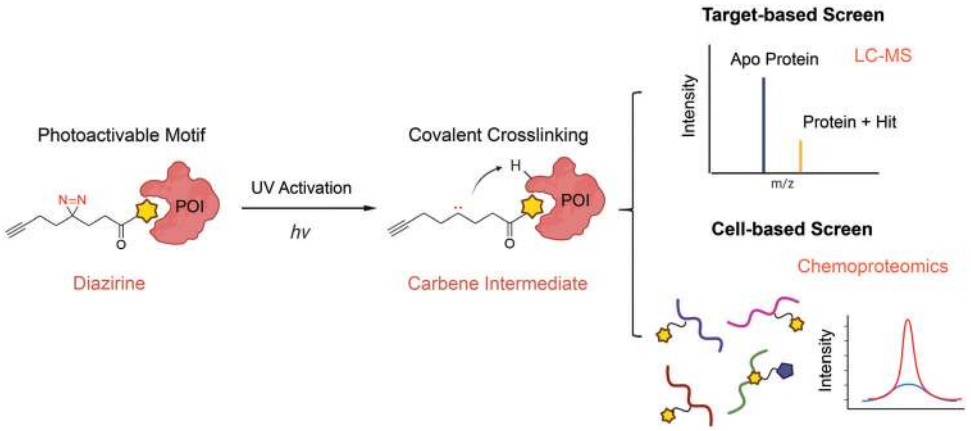

Fig. 5 Schematic of photoactivation-assisted fragment discovery. Photoreactive fragments contains a photoactivable diazirine, which could generate the carbene intermediate upon UV activation that can crosslink to proximal protein residues. Photoactivation-assisted fragment discovery has broad applicability in both target-based[79] and cell-based screening[78] campaigns.

The resulting fragment-protein adducts could be easily detected by LC-MS and MS/MS approaches result in the identification of the sites of covalent modification. As proof-of-concept, they established a compound collection of 556 PhABits and screened against a variety of targets including BRD4-BD1, BCL6 and KRas4B^{G12D}. This enables rapid identification of several selective hits that have been successfully verified by follow-up biophysical studies.

Overall, one major advantage of using photoactivatable groups in FBDD screens is that the fragment interactions with targets are driven by LE, while photoactivation and subsequent covalent adduct formation overcomes the low K_d values associated with noncovalent fragments. Therefore, we expect that these types of strategies will continue to play an important role in fragment screening and discovery. However, it should be noted that fine tuning of the photolabeling method is required to ensure specificity and avoid the formation of putative long-lived species, which may lead to diffusion-based, non-specific crosslinking.

Computer-aided covalent fragment design

In addition to methods described above, computational methodologies and software have been recently tailored towards covalent fragment-based ligand discovery. Designing covalent fragments with proper warhead reactivity is of utmost importance in covalent ligand discovery. In recent years, a variety of computational approaches have been investigated to estimate warhead reactivity based on quantum mechanics (QM) calculation and experimental calculated intrinsic warhead reactivity. Some of the warheads analyzed in this manner include nitrile-carrying compounds,[80] Michael acceptors,[81] acrylamides and 2-chloroacetamides.[82]

Besides predicting the reactivity of a warhead, understanding reactivity of protein nucleophiles is also of paramount importance. The pK_a of a typical lysine side chain is around 10.4, which means under physiological conditions, they are protonated and cannot readily react with electrophiles.[83] However, the pK_a of lysine could be perturbed by its environment, thus changing its reactivity towards covalent fragments.[84] In 2019, Liu *et al.* developed a continuous constant pH molecular dynamics (CpHMD) for lysine nucleophilicities prediction.[85] This new strategy successfully identified catalytic lysines in 8 human kinases, which hold great promise for the development of lysine-targeting covalent kinase inhibitors. Overall, computational reactivity prediction methodologies could be potentially useful for virtual compound library design as well as for guiding covalent fragment library design in general.

Free energy calculation is considered to be a robust computational approach for binding affinity prediction, and this strategy has been successfully extended to the development of covalent inhibitors. For example, Chatterjee *et al.* established a two-state model for selectivity prediction of reversible covalent inhibitors among protein homologues.[86] Very recently, Mihalovits *et al.* also extended free energy calculation for the affinity and selectivity evaluation of irreversible binders as exemplified by covalent inhibitors targeting K-Ras ^{G12C} and EGFR.[87]

Another important advancement has been the development of covalent docking approaches.[88–90] Structure-based virtual screening has been successfully adopted to enrich active compounds from large compound libraries.[91] In response to growing interest in targeted covalent inhibitors, many computational platforms have integrated covalent docking modules to assist drug discovery campaigns.[88,92–100] For example, London *et al.* described a method called DOCKovalent for large-scale covalent fragment virtual screening campaigns (http://covalent.docking.org/).[88] In this method, the receptor nucleophiles were set as rigid while the fragment conformations and possible stereoisomers were exhaustively sampled with constraints on the distance and angle of the covalent bond. The scoring was calculated based on non-covalent interactions to evaluate structural complementarity without direct calculation of covalent bond energy. This method has been successfully applied to target different nucleophiles and discover reversible covalent inhibitors targeting AmpC β-lactamase, RSK2, MSK1 and JAK3 kinases by

screening electrophile fragment libraries. Very recently, Tang *et al.* developed an artificial intelligence (AI)-based design strategy namely advanced deep Q-learning network (ADQN) for covalent ligand discovery targeting SARS-CoV-2 $3CL^{pro}$ which is also a nice example of how artificial intelligence make the hunt quickly and efficiently.[101]

In another example, Robert *et al.* generated a covalent 3D pharmacophore collection based on protein–ligand interactions, which was subsequently applied for covalent fragment screen against enteroviral 3C protease.[102] By screening against a compound collection of approximately 3000 fragments, scaffold hopping and follow-up biological evaluation, they identified phenylthiomethyl ketone-based fragments as 3C protease inhibitors.

Despite these successes, one of the major bottlenecks in this field is that the docking performance varies when it comes to different targets making it challenging to develop the most suitable and applicable scoring function for hit enrichment or accurately rank the hits. Additionally, most current covalent docking tools fail to calculate covalent bond energy, which relies on computationally expensive QM calculations. Proper consideration for covalency, stability, reaction rate, and potential reversibility add layers of complexity to the calculations and it remains to be seen whether those methodologies could be generally applied against a broader range of targets. We believe that based on vastly increasing amount chemoproteomics data, the community efforts will enable the development and validation of new computational methods including modern machine learning algorithms for *in silico* screening of covalent virtual library and rational design of covalent ligand for targets of interest.

Advancing challenging pharmacological modalities

Inhibition of protein–protein interactions

Fragment-based ligand and drug discovery approaches have a great potential to drug some of the more challenging targets, such as protein–protein interactions (PPIs).[103] There are estimated 650,000 PPIs in human proteome.[104] However, due to the large and flat nature of most PPI interfaces, it remains quite challenging to develop small molecules to target them.[105,106] Due to the covalent nature, covalent fragments have the ability to bind when the partners are apart, which could avoid the huge loss of potency due to competition with the complex. This could be of special importance when targeting signaling pathways where interactions are regulated and/or transient. Thus, leveraging the covalency may help address shallow pockets and offer additional opportunities for chemical intervention against PPIs by irreversible attachment.

In 2015, by screening covalent fragments, Statsyuk *et al.* identified a covalent inhibitor targeting non-catalytic Cys627 located at the HECT domain of Nedd4-1, an E3 ubiquitin ligase.[45] The covalent mechanism of action (MOA) was validated by mass spectrometry and crystallographic study. Further chemistry optimization led to the development of more potent inhibitor which could disrupt Nedd4-1:Ub interaction with k_{inact}/K_I value of 1.98 M^{-1} s^{-1}. In another example, starting from non-covalent fragment hit, researchers from AstraZeneca used sulfonyl fluoride warhead to rationally design covalent fragments targeting Tyr101 of antiapoptotic protein Bcl-xl.[107] Subsequent structure-based elaboration led to more potent compound with time-dependent biochemical potency against Bcl-2 family proteins-BIM/BAK interactions. These are important examples of how covalent fragment screens combined with chemistry optimization could generate novel chemical tools to target challenging PPIs. Over the coming decade, we expect to see fragment-based covalent ligand discovery to rise to challenges posed by PPIs.

Covalent fragments for allosteric regulator development

Although covalent allosteric inhibitors have recently gained interest in drug discovery for challenging target proteins,[108,109] the systematic discovery of those hidden low-occupancy allosteric sites remains highly challenging. Due to their improved ability to access cryptic sites, fragments could be used for discovery of allosteric binders as well as identification of functionally relevant secondary binding sites on proteins.[16] In the context of covalent fragments, Keedy *et al.* screened a disulfide-capped fragment library containing around 1600 fragments against the engineered cysteine mutant (Lys197Cys) of tyrosine phosphatase 1B (PTP1B), a target that has been difficult to selectively inhibit *via* active site directed route.[110] After a combined labeling study and a follow-up optimization, fragment 2 was identified as a noncompetitive allosteric inhibitor in an enzymatic assay. Crystallographic studies clearly demonstrated that fragment 2 was covalently tethered to Lys197Cys instead of the catalytic cysteine in the active site, with protein conformational changes similar to those induced by endogenous regulatory mechanism. Additionally, Keedy *et al.* identified additional fragments that were found to bind to the same location in wild type (WT) PTP1B, suggesting that identified site could be exploited for allosteric inhibitor development.[110] Overall, this work highlights the utility of covalent fragments for allosteric site binder discovery.

Development of proteolysis targeting chimeras (PROTACs)

Another compelling application is the use of covalent fragments to develop new modalities, such as proteolysis targeting chimeras (PROTACs). Recently, PROTACs emerged as an appealing strategy in drug discovery. These heterobifunctional small molecules recruit E3 ligases to the target of interest, subsequently resulting in target ubiquitination and degradation.[111] Although there are more than 600 putative E3 ligases available in human proteome, only a handful of them (CRBN, VHL, MDM2 and cIAP1) has been successfully employed to develop degraders due to the lack of well-characterized E3 ligase ligands.[112] Thus, identifying additional E3 recruiters by covalent fragment screen may offer great opportunity for the development of novel electrophilic PROTACs with distinct pharmacological profiles. For example, ''scout fragments'' with chloroacetamide or acrylamide warheads have recently been coupled to selective ligands to screen bifunctional covalent degrader molecules.[113] This effort led to

the discovery of nuclear-localized FKBP12 degrader KB02-SLF and BRD4 degrader KB02-JQ1, which hijacked E3 ligase DCAF16 for nuclear protein degradation. This is an exciting example that demonstrated the feasibility of chemoproteomics-based covalent fragment screen in electrophilic PROTACs development.

Additionally, a recent ABPP-based covalent screen against purified RNF4, a RING-domain E3 ligase identified new covalent fragments that target this protein.[114,115] The most promising hit TRH 1-23 was shown to label either Cys132 or Cys135, two cysteine residues that form one of zinc-coordinating sites within the RING domain. Although disrupting either one of these sites through mutagenesis was previously shown to inhibit RNF4 function, Ward *et al.* included evidence to suggest that RNF4 retains *in vitro* self-ubiquitination activity, suggesting that TRH 1-23 did not disrupt RNF4 E3 ligase function. Further work is needed to characterize this biding and explain the lack of impact on E3 ligase activity, despite the targeting of the two cysteines essential for zinc binding. The same study also demonstrated that these hits can be used to generate PROTACs featuring covalent E3 recruiting arm. The PROTACs were validated based on targeted protein degradation of BRD4.

Overall, these examples highlight the potential utility of covalent fragments to discover ''hidden'' druggable sites and serve as a starting point for developing ligands, including PROTACs. Although covalent PROTACs have initially been somewhat controversial, there is now broader acceptance that PROTACs featuring a covalent handle on E3 ligase recruiting arm may have benefits. Very recently, Bond *et al.* reported an effective degrader of Kras-G12C by coupling a covalent Kras-G12C ligand with a VHL E3 recruiter. While there have been very few examples of using covalent target ligands for bifunctional degraders, going forward it would be interesting to realize utility of covalent ligands to degrade high-value targets that have been recalcitrant to ligand discovery.

Outlook and future directions

Covalent fragment-based ligand discovery has already yielded a significant number of useful and synthetically accessible covalent fragments for drug discovery. These binders provide a solid starting point for developing probes for the ligandable hot spot of the protein targets. However, transforming a covalent fragment hit into an optimized molecule still requires a considerable medicinal chemistry effort. Here, strategies such as structure-guided drug design, covalent SAR (covSAR) interpretation,[116] fragment merging[117] and similar can help speed up the process and we expect that these efforts will continue to be merged into unified campaigns towards novel leads and drugs.

Chemoproteomic methods for characterizing covalent fragments can play an important role towards the identification of diverse electrophiles with suitable properties for future covalent fragment library expansions. For instance, the incorporation of masked thiol-reactive fragments into commonly used cysteine-directing library that could be triggered under precisely defined physiological conditions has recently been shown as a viable option to diversify the cysteine-directing library.[118,119] We expect that further expansion of the covalent fragment chemical space will occur as the efforts move towards targeting alternative residues beyond cysteines, as we discussed above. Projecting forward, the identification of covalent fragments modifying those protein nucleophiles (lysine, serine, histidine, tyrosine, methionine, glutamic acid *etc.*) would help expand the scope of covalent fragment and covalent inhibitor use, as well as the target space for small molecule ligand discovery and development. We also expect that the growing amount of chemoproteomic data will further feed into development of new computational methods, including machine learning algorithms for virtual screening of covalent fragments for a target of interest.

One of the main concerns related to the use of covalent fragments is their potential promiscuous nature, which may not directly correlate with intrinsic reactivity. In this context, frequent hitters, such as aminothiazole chloroacetamides that are not significantly more reactive on average, have been recognized as promiscuous binders.[38] However, we should still be careful about covalent warheads with high reactivity ($k > 1 \times 10^{-7}$ M^{-1} s^{-1}) with the caveat about their potential non-specific cell toxicity. Some compounds with non-specific reactivity, such as quinones, maleimides and *N*-hydroxysuccinimide (NHS)-esters, have been flagged as pan-assay interference compounds (PAINS)[120,121] and should also be avoided in the construction of covalent fragment library.

We believe that in the foreseeable future, the covalent fragments will continue to help us assess biologically relevant chemical space and systematically identify more chemical-tractable targets in the native biological systems. Additionally, covalent fragments and the chemical probes and lead compounds they spawn will be of essential importance for facilitating broader study of biological processes and how they contribute to normal human physiology as well as disease. In summary, we expect that covalent fragments will be of growing interest to chemical biologist, as well as those interested in drug discovery and development.

The fact that covalent fragment discovery has already succeeded in enabling actual drug candidates against KRAS, the notorious and most sought-after target in oncology also leads to a big question for the field. Was this a lucky, one-off case where covalent targeting just happened to match this particular challenging field or is this a preview, of continued and increased impacts that these approaches will have on drug discovery as the improved methods, larger libraries, and increased focus start to bear fruit? We eagerly await further forthcoming efforts in this field and hope we are truly entering a new era for targeted therapy.

Conflicts of interest

N. S. G. is a founder, science advisory board member (SAB) and equity holder in Gatekeeper, Syros, Inception, Jengu, C4, B2S, Aduro and Soltego. The Gray lab receives or has received

research funding from Novartis, Takeda, Astellas, Taiho, Janssen, Kinogen, Voronoi, Her2llc, Epiphanes, Deerfield and Sanofi. L. H. J. receives research funding from Deerfield. M. K. is a paid consulting editor at Life Science Editors. M. P. P. is an employee and have equity in Vividion Therapeutics. J. C. is a consultant for Jengu, Allorion, and Soltego, and an equity holder in Allorion and M3 bioinformatics & technology.

Acknowledgements

Figures created with Biorender.com.

References

1 A. K. Ghosh, I. Samanta, A. Mondal and W. R. Liu, *ChemMedChem*, 2019, **14**, 889–906.
2 T. Zhang, J. M. Hatcher, M. Teng, N. S. Gray and M. Kostic, *Cell Chem. Biol.*, 2019, **26**, 1486–1500.
3 R. Lonsdale and R. A. Ward, *Chem. Soc. Rev.*, 2018, **47**, 3816–3830.
4 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307–317.
5 M. Forster, A. Chaikuad, S. M. Bauer, J. Holstein, M. B. Robers, C. R. Corona, M. Gehringer, E. Pfaffenrot, K. Ghoreschi, S. Knapp and S. A. Laufer, *Cell Chem. Biol.*, 2016, **23**, 1335–1340.
6 T. Zhang, N. Kwiatkowski, C. M. Olson, S. E. Dixon-Clarke, B. J. Abraham, A. K. Greifenberg, S. B. Ficarro, J. M. Elkins, Y. Liang, N. M. Hannett, T. Manz, M. Hao, B. Bartkowiak, A. L. Greenleaf, J. A. Marto, M. Geyer, A. N. Bullock, R. A. Young and N. S. Gray, *Nat. Chem. Biol.*, 2016, **12**, 876–884.
7 B. K. Park, D. J. Naisbitt, S. F. Gordon, N. R. Kitteringham and M. Pirmohamed, *Toxicology*, 2001, **158**, 11–23.
8 B. K. Park, A. Boobis, S. Clarke, C. E. Goldring, D. Jones, J. G. Kenna, C. Lambert, H. G. Laverty, D. J. Naisbitt, S. Nelson, D. A. Nicoll-Griffith, R. S. Obach, P. Routledge, D. A. Smith, D. J. Tweedie, N. Vermeulen, D. P. Williams, I. D. Wilson and T. A. Baillie, *Nat. Rev. Drug Discovery*, 2011, **10**, 292–306.
9 F. Sutanto, M. Konstantinidou and A. Dömling, *RSC Med. Chem.*, 2020, **11**, 876–884.
10 R. Lagoutte, R. Patouret and N. Winssinger, *Curr. Opin. Chem. Biol.*, 2017, **39**, 54–63.
11 C. V. Dang, E. P. Reddy, K. M. Shokat and L. Soucek, *Nat. Rev. Cancer*, 2017, **17**, 502–508.
12 D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke and H. Jhoti, *Nat. Rev. Drug Discovery*, 2016, **15**, 605–619.
13 F. N. Edfeldt, R. H. Folmer and A. L. Breeze, *Drug Discovery Today*, 2011, **16**, 284–287.
14 D. R. Hall, D. Kozakov, A. Whitty and S. Vajda, *Trends Pharmacol. Sci.*, 2015, **36**, 724–736.
15 D. C. Rees, M. Congreve, C. W. Murray and R. Carr, *Nat. Rev. Drug Discovery*, 2004, **3**, 660–672.
16 R. F. Ludlow, M. L. Verdonk, H. K. Saini, I. J. Tickle and H. Jhoti, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 15910–15915.
17 A. Keeley, L. Petri, P. Abranyi-Balogh and G. M. Keseru, *Drug Discovery Today*, 2020, **25**, 983–996.
18 N. S. Troelsen and M. H. Clausen, *Chemistry*, 2020, **26**, 11391–11403.
19 S. G. Kathman and A. V. Statsyuk, *MedChemComm*, 2016, **7**, 576–585.
20 D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 9367–9372.
21 R. H. Nonoo, A. Armstrong and D. J. Mann, *ChemMedChem*, 2012, **7**, 2082–2086.
22 E. Buck and J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 2719–2724.
23 W. Yang, R. V. Fucini, B. T. Fahr, M. Randal, K. E. Lind, M. B. Lam, W. Lu, Y. Lu, D. R. Cary, M. J. Romanowski, D. Colussi, B. Pietrak, T. J. Allison, S. K. Munshi, D. M. Penny, P. Pham, J. Sun, A. E. Thomas, J. M. Wilkinson, J. W. Jacobs, R. S. McDowell and M. D. Ballinger, *Biochemistry*, 2009, **48**, 4488–4496.
24 D. A. Erlanson, J. W. Arndt, M. T. Cancilla, K. Cao, R. A. Elling, N. English, J. Friedman, S. K. Hansen, C. Hession, I. Joseph, G. Kumaravel, W. C. Lee, K. E. Lind, R. S. McDowell, K. Miatkowski, C. Nguyen, T. B. Nguyen, S. Park, N. Pathan, D. M. Penny, M. J. Romanowski, D. Scott, L. Silvian, R. L. Simmons, B. T. Tangonan, W. Yang and L. Sun, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 3078–3083.
25 J. Gao and J. A. Wells, *Chem. Biol. Drug Des.*, 2012, **79**, 209–215.
26 J. M. Scheer, M. J. Romanowski and J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 7595–7600.
27 J. A. Hardy, J. Lam, J. T. Nguyen, T. O'Brien and J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12461–12466.
28 M. R. Arkin, M. Randal, W. L. DeLano, J. Hyde, T. N. Luong, J. D. Oslob, D. R. Raphael, L. Taylor, J. Wang, R. S. McDowell, J. A. Wells and A. C. Braisted, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1603–1608.
29 D. A. Erlanson, J. A. Wells and A. C. Braisted, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 199–223.
30 A. R. Moore, S. C. Rosenberg, F. McCormick and S. Malek, *Nat. Rev. Drug Discovery*, 2020, **19**, 533–552.
31 J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells and K. M. Shokat, *Nature*, 2013, **503**, 548–551.
32 M. R. Janes, J. Zhang, L. S. Li, R. Hansen, U. Peters, X. Guo, Y. Chen, A. Babbar, S. J. Firdaus, L. Darjania, J. Feng, J. H. Chen, S. Li, S. Li, Y. O. Long, C. Thach, Y. Liu, A. Zarieh, T. Ely, J. M. Kucharski, L. V. Kessler, T. Wu, K. Yu, Y. Wang, Y. Yao, X. Deng, P. P. Zarrinkar, D. Brehmer, D. Dhanak, M. V. Lorenzi, D. Hu-Lowe, M. P. Patricelli, P. Ren and Y. Liu, *Cell*, 2018, **172**, 578–589 e517.
33 J. Canon, K. Rex, A. Y. Saiki, C. Mohr, K. Cooke, D. Bagal, K. Gaida, T. Holt, C. G. Knutson, N. Koppada,

B. A. Lanman, J. Werner, A. S. Rapaport, T. San Miguel, R. Ortiz, T. Osgood, J. R. Sun, X. Zhu, J. D. McCarter, L. P. Volak, B. E. Houk, M. G. Fakih, B. H. O'Neil, T. J. Price, G. S. Falchook, J. Desai, J. Kuo, R. Govindan, D. S. Hong, W. Ouyang, H. Henary, T. Arvedson, V. J. Cee and J. R. Lipford, *Nature*, 2019, **575**, 217–223.

34 M. P. Patricelli, M. R. Janes, L. S. Li, R. Hansen, U. Peters, L. V. Kessler, Y. Chen, J. M. Kucharski, J. Feng, T. Ely, J. H. Chen, S. J. Firdaus, A. Babbar, P. Ren and Y. Liu, *Cancer Discovery*, 2016, **6**, 316–329.

35 J. B. Fell, J. P. Fischer, B. R. Baer, J. Ballard, J. F. Blake, K. Bouhana, B. J. Brandhuber, D. M. Briere, L. E. Burgess, M. R. Burkard, H. Chiang, M. J. Chicarelli, K. Davidson, J. J. Gaudino, J. Hallin, L. Hanson, K. Hee, E. J. Hicken, R. J. Hinklin, M. A. Marx, M. J. Mejia, P. Olson, P. Savechenkov, N. Sudhakar, T. P. Tang, G. P. Vigers, H. Zecca and J. G. Christensen, *ACS Med. Chem. Lett.*, 2018, **9**, 1230–1234.

36 J. Hallin, L. D. Engstrom, L. Hargis, A. Calinisan, R. Aranda, D. M. Briere, N. Sudhakar, V. Bowcut, B. R. Baer, J. A. Ballard, M. R. Burkard, J. B. Fell, J. P. Fischer, G. P. Vigers, Y. Xue, S. Gatto, J. Fernandez-Banet, A. Pavlicek, K. Velastagui, R. C. Chao, J. Barton, M. Pierobon, E. Baldelli, E. F. Patricoin, 3rd, D. P. Cassidy, M. A. Marx, Rybkin, II, M. L. Johnson, S. I. Ou, P. Lito, K. P. Papadopoulos, P. A. Janne, P. Olson and J. G. Christensen, *Cancer Discovery*, 2020, **10**, 54–71.

37 H. Johansson, Y. C. Isabella Tsai, K. Fantom, C. W. Chung, S. Kumper, L. Martino, D. A. Thomas, H. C. Eberl, M. Muelbaier, D. House and K. Rittinger, *J. Am. Chem. Soc.*, 2019, **141**, 2703–2712.

38 E. Resnick, A. Bradley, J. Gan, A. Douangamath, T. Krojer, R. Sethi, P. P. Geurink, A. Aimon, G. Amitai, D. Bellini, J. Bennett, M. Fairhead, O. Fedorov, R. Gabizon, J. Gan, J. Guo, A. Plotnikov, N. Reznik, G. F. Ruda, L. Diaz-Saez, V. M. Straub, T. Szommer, S. Velupillai, D. Zaidman, Y. Zhang, A. R. Coker, C. G. Dowson, H. M. Barr, C. Wang, K. V. M. Huber, P. E. Brennan, H. Ovaa, F. von Delft and N. London, *J. Am. Chem. Soc.*, 2019, **141**, 8951–8968.

39 C. Dubiella, B. J. Pinch, D. Zaidman, T. D. Manz, E. Poon, S. He, E. Resnick, E. M. Langer, C. J. Daniel, H. Seo, Y. Chen, S. B. Ficarro, Y. Jamin, X. Lian, S. Kibe, S. Kozono, K. Koikawa, Z. M. Doctor, B. Nabet, C. M. Browne, A. Yang, L. Stoler-Barak, R. B. Shah, N. E. Vangos, E. A. Geffken, R. Oren, S. Sidi, Z. Shulman, C. Wang, J. A. Marto, S. Dhe-Paganon, T. Look, X. Zhou, K. Lu, R. C. Sears, L. Chesler, N. S. Gray and N. London, *bioRxiv*, 2020, DOI: 10.1101/2020.03.20.998443.

40 A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Abranyi-Balogh, J. Brandao-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keseru, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft and M. A. Walsh, *Nat. Commun.*, 2020, **11**, 5047.

41 Y. M. Go and D. P. Jones, *Crit. Rev. Biochem. Mol. Biol.*, 2013, **48**, 173–181.

42 R. Cardoso, R. Love, C. L. Nilsson, S. Bergqvist, D. Nowlin, J. Yan, K. K. Liu, J. Zhu, P. Chen, Y. L. Deng, H. J. Dyson, M. J. Greig and A. Brooun, *Protein Sci.*, 2012, **21**, 1885–1896.

43 S. G. Kathman, Z. Xu and A. V. Statsyuk, *J. Med. Chem.*, 2014, **57**, 4969–4974.

44 G. B. Craven, D. P. Affron, C. E. Allen, S. Matthies, J. G. Greener, R. M. L. Morgan, E. W. Tate, A. Armstrong and D. J. Mann, *Angew. Chem., Int. Ed.*, 2018, **57**, 5257–5261.

45 S. G. Kathman, I. Span, A. T. Smith, Z. Xu, J. Zhan, A. C. Rosenzweig and A. V. Statsyuk, *J. Am. Chem. Soc.*, 2015, **137**, 12442–12445.

46 C. Jost, C. Nitsche, T. Scholz, L. Roux and C. D. Klein, *J. Med. Chem.*, 2014, **57**, 7590–7599.

47 Q. Zheng, J. L. Woehl, S. Kitamura, D. Santos-Martins, C. J. Smedley, G. Li, S. Forli, J. E. Moses, D. W. Wolan and K. B. Sharpless, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 18808–18814.

48 I. Miyahisa, T. Sameshima and M. S. Hixon, *Angew. Chem., Int. Ed.*, 2015, **54**, 14099–14102.

49 J. M. Strelow, *SLAS Discovery*, 2017, **22**, 3–20.

50 M. Kostic and L. H. Jones, *Trends Pharmacol. Sci.*, 2020, **41**, 305–317.

51 L. Petri, P. Abranyi-Balogh, I. Timea, G. Palfy, A. Perczel, D. Knez, M. Hrast, M. Gobec, I. Sosic, K. Nyiri, B. G. Vertessy, N. Jansch, C. Desczyk, F. J. Meyer-Almes, I. Ogris, S. Golic Grdadolnik, L. G. Iacovino, C. Binda, S. Gobec and G. M. Keseru, *ChemBioChem*, 2020, DOI: 10.1002/cbic.202000700.

52 L. Petri, A. Egyed, D. Bajusz, T. Imre, A. Hetenyi, T. Martinek, P. Abranyi-Balogh and G. M. Keseru, *Eur. J. Med. Chem.*, 2020, **207**, 112836.

53 O. Plettenburg, *Angew. Chem., Int. Ed.*, 2017, **56**, 446–448.

54 K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. Gonzalez-Paez, S. Chatterjee, B. R. Lanning, J. R. Teijaro, A. J. Olson, D. W. Wolan and B. F. Cravatt, *Nature*, 2016, **534**, 570–574.

55 E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. Dillon, D. A. Bachovchin, K. Mowen, D. Baker and B. F. Cravatt, *Nature*, 2010, **468**, 790–795.

56 F. Yang, J. Gao, J. Che, G. Jia and C. Wang, *Anal. Chem.*, 2018, **90**, 9576–9582.

57 Y. Wang, M. M. Dix, G. Bianco, J. R. Remsberg, H. Y. Lee, M. Kalocsay, S. P. Gygi, S. Forli, G. Vite, R. M. Lawrence, C. G. Parker and B. F. Cravatt, *Nat. Chem.*, 2019, **11**, 1113–1123.

58 C. M. Browne, B. Jiang, S. B. Ficarro, Z. M. Doctor, J. L. Johnson, J. D. Card, S. C. Sivakumaren, W. M. Alexander, T. M. Yaron, C. J. Murphy, N. P. Kwiatkowski, T. Zhang, L. C. Cantley, N. S. Gray and J. A. Marto, *J. Am. Chem. Soc.*, 2019, **141**, 191–203.

59 H. Xiao, M. P. Jedrychowski, D. K. Schweppe, E. L. Huttlin, Q. Yu, D. E. Heppner, J. Li, J. Long, E. L. Mills, J. Szpyt, Z. He, G. Du, R. Garrity, A. Reddy, L. P. Vaites, J. A. Paulo, T. Zhang, N. S. Gray, S. P. Gygi and E. T. Chouchani, *Cell*, 2020, **180**(968-983), e924.

60 L. Bar-Peled, E. K. Kemper, R. M. Suciu, E. V. Vinogradova, K. M. Backus, B. D. Horning, T. A. Paul, T. A. Ichu, R. U. Svensson, J. Olucha, M. W. Chang, B. P. Kok, Z. Zhu, N. T. Ihle, M. M. Dix, P. Jiang, M. M. Hayward, E. Saez, R. J. Shaw and B. F. Cravatt, *Cell*, 2017, **171**(696-709), e623.

61 E. V. Vinogradova, X. Zhang, D. Remillard, D. C. Lazar, R. M. Suciu, Y. Wang, G. Bianco, Y. Yamashita, V. M. Crowley, M. A. Schafroth, M. Yokoyama, D. B. Konrad, K. M. Lum, G. M. Simon, E. K. Kemper, M. R. Lazear, S. Yin, M. M. Blewett, M. M. Dix, N. Nguyen, M. N. Shokhirev, E. N. Chin, L. L. Lairson, B. Melillo, S. L. Schreiber, S. Forli, J. R. Teijaro and B. F. Cravatt, *Cell*, 2020, **182**(1009-1026), e1029.

62 L. B. Poole, *Free Radicals Biol. Med.*, 2015, **80**, 148–157.

63 A. Miseta and P. Csutora, *Mol. Biol. Evol.*, 2000, **17**, 1232–1239.

64 T. H. Truong and K. S. Carroll, *Biochemistry*, 2012, **51**, 9954–9965.

65 C. C. Ward, J. I. Kleinman and D. K. Nomura, *ACS Chem. Biol.*, 2017, **12**, 1478–1483.

66 H. Mukherjee and N. P. Grimster, *Curr. Opin. Chem. Biol.*, 2018, **44**, 30–38.

67 S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia and B. F. Cravatt, *Nat. Chem.*, 2017, **9**, 1181–1190.

68 M. Wolter, D. Valenti, P. J. Cossar, L. M. Levy, S. Hristeva, T. Genski, T. Hoffmann, L. Brunsveld, D. Tzalis and C. Ottmann, *Angew. Chem., Int. Ed.*, 2020, **59**, 21520–21524.

69 B. Z. Stanton, E. J. Chory and G. R. Crabtree, *Science*, 2018, **359**, eaao5902.

70 L. H. Jones, *ACS Med. Chem. Lett.*, 2018, **9**, 584–586.

71 A. S. Barrow, C. J. Smedley, Q. Zheng, S. Li, J. Dong and J. E. Moses, *Chem. Soc. Rev.*, 2019, **48**, 4731–4758.

72 D. E. Mortenson, G. J. Brighty, L. Plate, G. Bare, W. Chen, S. Li, H. Wang, B. F. Cravatt, S. Forli, E. T. Powers, K. B. Sharpless, I. A. Wilson and J. W. Kelly, *J. Am. Chem. Soc.*, 2018, **140**, 200–210.

73 G. J. Brighty, R. C. Botham, S. Li, L. Nelson, D. E. Mortenson, G. Li, C. Morisseau, H. Wang, B. D. Hammock, K. B. Sharpless and J. W. Kelly, *Nat. Chem.*, 2020, **12**, 906–913.

74 H. S. Hahm, E. K. Toroitich, A. L. Borne, J. W. Brulet, A. H. Libby, K. Yuan, T. B. Ware, R. L. McCloud, A. M. Ciancone and K. L. Hsu, *Nat. Chem. Biol.*, 2020, **16**, 150–159.

75 V. Gupta, J. Yang, D. C. Liebler and K. S. Carroll, *J. Am. Chem. Soc.*, 2017, **139**, 5588–5595.

76 J. Li, J. G. Van Vranken, L. Pontano Vaites, D. K. Schweppe, E. L. Huttlin, C. Etienne, P. Nandhikonda, R. Viner, A. M. Robitaille, A. H. Thompson, K. Kuhn, I. Pike, R. D. Bomgarden, J. C. Rogers, S. P. Gygi and J. A. Paulo, *Nat. Methods*, 2020, **17**, 399–404.

77 M. P. Patricelli, T. K. Nomanbhoy, J. Wu, H. Brown, D. Zhou, J. Zhang, S. Jagannathan, A. Aban, E. Okerberg, C. Herring, B. Nordin, H. Weissig, Q. Yang, J. D. Lee, N. S. Gray and J. W. Kozarich, *Chem. Biol.*, 2011, **18**, 699–710.

78 C. G. Parker, A. Galmozzi, Y. Wang, B. E. Correia, K. Sasaki, C. M. Joslyn, A. S. Kim, C. L. Cavallaro, R. M. Lawrence, S. R. Johnson, I. Narvaiza, E. Saez and B. F. Cravatt, *Cell*, 2017, **168**(527-541), e529.

79 E. K. Grant, D. J. Fallon, M. M. Hann, K. G. M. Fantom, C. Quinn, F. Zappacosta, R. S. Annan, C. W. Chung, P. Bamborough, D. P. Dixon, P. Stacey, D. House, V. K. Patel, N. C. O. Tomkinson and J. T. Bush, *Angew. Chem., Int. Ed.*, 2020, **59**, 21096–21105.

80 A. Berteotti, F. Vacondio, A. Lodola, M. Bassi, C. Silva, M. Mor and A. Cavalli, *ACS Med. Chem. Lett.*, 2014, **5**, 501–505.

81 J. A. Schwobel, D. Wondrousch, Y. K. Koleva, J. C. Madden, M. T. Cronin and G. Schuurmann, *Chem. Res. Toxicol.*, 2010, **23**, 1576–1585.

82 F. Palazzesi, M. R. Hermann, M. A. Grundl, A. Pautsch, D. Seeliger, C. S. Tautermann and A. Weber, *J. Chem. Inf. Model.*, 2020, **60**, 2915–2923.

83 J. Pettinger, K. Jones and M. D. Cheeseman, *Angew. Chem., Int. Ed.*, 2017, **56**, 15200–15209.

84 H. Ishikita, *FEBS Lett.*, 2010, **584**, 3464–3468.

85 R. Liu, Z. Yue, C. C. Tsai and J. Shen, *J. Am. Chem. Soc.*, 2019, **141**, 6553–6560.

86 P. Chatterjee, W. M. Botello-Smith, H. Zhang, L. Qian, A. Alsamarah, D. Kent, J. J. Lacroix, M. Baudry and Y. Luo, *J. Am. Chem. Soc.*, 2017, **139**, 17945–17952.

87 L. M. Mihalovits, G. G. Ferenczy and G. M. Keseru, *J. Chem. Inf. Model.*, 2020, **60**, 6579–6594.

88 N. London, R. M. Miller, S. Krishnan, K. Uchida, J. J. Irwin, O. Eidam, L. Gibold, P. Cimermancic, R. Bonnet, B. K. Shoichet and J. Taunton, *Nat. Chem. Biol.*, 2014, **10**, 1066–1072.

89 C. Sotriffer, *Mol. Inf.*, 2018, **37**, e1800062.

90 H. M. Kumalo, S. Bhakat and M. E. Soliman, *Molecules*, 2015, **20**, 1984–2000.

91 D. B. Kitchen, H. Decornez, J. R. Furr and J. Bajorath, *Nat. Rev. Drug Discovery*, 2004, **3**, 935–949.

92 A. Scarpino, G. G. Ferenczy and G. M. Keseru, *J. Chem. Inf. Model.*, 2018, **58**, 1441–1458.

93 G. Bianco, S. Forli, D. S. Goodsell and A. J. Olson, *Protein Sci.*, 2016, **25**, 295–301.

94 C. Scholz, S. Knorr, K. Hamacher and B. Schmidt, *J. Chem. Inf. Model.*, 2015, **55**, 398–406.

95 J. Schroder, A. Klinger, F. Oellien, R. J. Marhofer, M. Duszenko and P. M. Selzer, *J. Med. Chem.*, 2013, **56**, 1478–1490.

96 S. De Cesco, S. Deslandes, E. Therrien, D. Levan, M. Cueto, R. Schmidt, L. D. Cantin, A. Mittermaier, L. Juillerat-Jeanneret and N. Moitessier, *J. Med. Chem.*, 2012, **55**, 6306–6315.

97 M. Rachman, A. Scarpino, D. Bajusz, G. Palfy, I. Vida, A. Perczel, X. Barril and G. M. Keseru, *ChemMedChem*, 2019, **14**, 1011–1021.
98 K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid and E. Harder, *J. Chem. Inf. Model.*, 2014, **54**, 1932–1940.
99 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.
100 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.
101 B. Tang, F. He, D. Liu, M. Fang, Z. Wu and D. Xu, *bioRxiv*, 2020, DOI: 10.1101/2020.03.03.972133.
102 R. Schulz, A. Atef, D. Becker, F. Gottschalk, C. Tauber, S. Wagner, C. Arkona, A. A. Abdel-Hafez, H. H. Farag, J. Rademann and G. Wolber, *J. Med. Chem.*, 2018, **61**, 1218–1230.
103 D. Valenti, S. Hristeva, D. Tzalis and C. Ottmann, *Eur. J. Med. Chem.*, 2019, **167**, 76–95.
104 M. P. Stumpf, T. Thorne, E. de Silva, R. Stewart, H. J. An, M. Lappe and C. Wiuf, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6959–6964.
105 V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173.
106 D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discovery*, 2016, **15**, 533–550.
107 H. Mukherjee, N. Su, M. A. Belmonte, D. Hargreaves, J. Patel, S. Tentarelli, B. Aquila and N. P. Grimster, *Bioorg. Med. Chem. Lett.*, 2019, **29**, 126682.
108 R. Nussinov and C. J. Tsai, *Annu. Rev. Pharmacol. Toxicol.*, 2015, **55**, 249–267.
109 Y. Bian, J. J. Jun, J. Cuyler and X. Q. Xie, *Eur. J. Med. Chem.*, 2020, **206**, 112690.
110 D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandao-Neto, N. M. Pearce, F. von Delft, J. A. Wells and J. S. Fraser, *eLife*, 2018, **7**, e36307.
111 A. C. Lai and C. M. Crews, *Nat. Rev. Drug Discovery*, 2017, **16**, 101–114.
112 X. Sun, H. Gao, Y. Yang, M. He, Y. Wu, Y. Song, Y. Tong and Y. Rao, *Signal Transduction Targeted Ther.*, 2019, **4**, 64.
113 X. Zhang, V. M. Crowley, T. G. Wucherpfennig, M. M. Dix and B. F. Cravatt, *Nat. Chem. Biol.*, 2019, **15**, 737–746.
114 C. C. Ward, J. I. Kleinman, S. M. Brittain, P. S. Lee, C. Y. S. Chung, K. Kim, Y. Petri, J. R. Thomas, J. A. Tallarico, J. M. McKenna, M. Schirle and D. K. Nomura, *ACS Chem. Biol.*, 2019, **14**, 2430–2440.
115 M. Luo, J. N. Spradlin, S. M. Brittain, J. M. McKenna, J. A. Tallarico, M. Schirle, T. J. Maimone and D. K. Nomura, *bioRxiv*, 2020, DOI: 10.1101/2020.07.12.198150.
116 G. B. Craven, D. P. Affron, T. Kosel, T. L. M. Wong, Z. H. Jukes, C. T. Liu, R. M. L. Morgan, A. Armstrong and D. J. Mann, *ChemBioChem*, 2020, **21**, 3417–3422.
117 M. Jaegle, E. L. Wong, C. Tauber, E. Nawrotzky, C. Arkona and J. Rademann, *Angew. Chem., Int. Ed.*, 2017, **56**, 7358–7378.
118 Z. V. Boskovic, M. M. Kemp, A. M. Freedy, V. S. Viswanathan, M. S. Pop, J. H. Fuller, N. M. Martinez, S. O. Figueroa Lazu, J. A. Hong, T. A. Lewis, D. Calarese, J. D. Love, A. Vetere, S. C. Almo, S. L. Schreiber and A. N. Koehler, *ACS Chem. Biol.*, 2016, **11**, 1844–1851.
119 R. Paxman, L. Plate, E. A. Blackwood, C. Glembotski, E. T. Powers, R. L. Wiseman and J. W. Kelly, *eLife*, 2018, **7**, e37168.
120 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719–2740.
121 J. Baell and M. A. Walters, *Nature*, 2014, **513**, 481–483.