



# Advances in covalent drug discovery

Lydia Boike<sup>1,2,3,4</sup>, Nathaniel J. Henning<sup>1,2,3,4</sup> and Daniel K. Nomura<sup>1,2,3</sup>✉

**Abstract** | Covalent drugs have been used to treat diseases for more than a century, but tools that facilitate the rational design of covalent drugs have emerged more recently. The purposeful addition of reactive functional groups to existing ligands can enable potent and selective inhibition of target proteins, as demonstrated by the covalent epidermal growth factor receptor (EGFR) and Bruton's tyrosine kinase (BTK) inhibitors used to treat various cancers. Moreover, the identification of covalent ligands through 'electrophile-first' approaches has also led to the discovery of covalent drugs, such as covalent inhibitors for KRAS(G12C) and SARS-CoV-2 main protease. In particular, the discovery of KRAS(G12C) inhibitors validates the use of covalent screening technologies, which have become more powerful and widespread over the past decade. Chemoproteomics platforms have emerged to complement covalent ligand screening and assist in ligand discovery, selectivity profiling and target identification. This Review showcases covalent drug discovery milestones with emphasis on the lessons learned from these programmes and how an evolving toolbox of covalent drug discovery techniques facilitates success in this field.

## Selectivity profiling

With respect to covalent drugs, this refers to the process of mapping amino acid sites within the proteome that react covalently with a ligand of interest with the purpose of identifying the targets and off-targets of the ligand in a biological context.

## Chemoproteomics

Collection of methods used to study small molecule–protein interactions on a proteome-wide scale.

Covalent drugs incorporate a mildly reactive functional group that forms a covalent bond with protein targets to confer additional affinity beyond the non-covalent interactions involved in drug binding<sup>1</sup>. Historically, concerns about the interference of these reactive molecules with biological assays and potential lack of selectivity often discouraged further investigation<sup>2,3</sup>. Many early covalent drugs were discovered serendipitously and bind active sites to inhibit enzymatic activity<sup>4</sup>. These drugs often mimic a substrate transition state to enable covalent modification of a catalytic amino acid residue. Over the past 30 years, the rational design of covalent drugs has garnered increased interest, and covalently targeting non-conserved amino acids to increase selectivity has become commonplace<sup>2,5</sup>. The prolonged target engagement of covalent drugs can provide distinct pharmacodynamic profiles and exceptional potency<sup>6</sup>.

The potential benefits of covalency have inspired medicinal chemists to explore the covalent drug space despite concerns about reactivity. In many cases, compromises between reactivity, selectivity and potency have produced safe and effective drugs. Key examples that we discuss here (FIG. 1 and TABLE 1) include the Bruton's tyrosine kinase (BTK) inhibitor ibrutinib (AbbVie) and the epidermal growth factor receptor (EGFR) inhibitor osimertinib (AstraZeneca), with sales totalling US\$8.43 billion and \$4.33 billion in 2020, respectively<sup>7,8</sup>. Moreover, potent inhibition through covalent modification has enabled targeting of traditionally 'undruggable' proteins, exemplified by the approval of sotorasib (Amgen), which is an inhibitor of mutant KRAS(G12C), a GTPase that resisted decades of drug discovery efforts<sup>9,10</sup> (FIG. 1).

At the same time, more traditional covalent targeting of protease active sites has continued to yield valuable drugs, such as nirmatrelvir (Pfizer), which inhibits the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (M<sup>Pro</sup>)<sup>11</sup> (FIG. 1).

Targeted covalent inhibitors are often discovered through structure-guided design by incorporating an electrophile into a ligand that would otherwise reversibly bind the target protein. The incorporated electrophile binds irreversibly to an amino acid on the target protein, introducing a covalent interaction in addition to the reversible interactions already at play.

Covalent ligand screening is another ligand discovery approach that is becoming more common, whereby various methods are used to discover covalent ligands from libraries of electrophilic compounds<sup>7,12,13</sup>. This 'electrophile-first' approach is partly facilitated by the development of chemoproteomic platforms that enable rapid target identification and selectivity profiling of covalent ligands<sup>14–18</sup>. Combining the advances in covalent ligand screening and chemoproteomics with structural biology to empower medicinal chemistry has the potential to generate molecules that selectively bind challenging protein targets.

In this Review, we start by briefly highlighting historical examples of covalent drugs and their mechanisms of action. We then elaborate on milestones in covalent drug discovery over the past decade, categorizing our discussion on the basis of the discovery approach taken. Finally, we summarize the toolbox of emerging covalent drug discovery techniques, with emphasis on screening strategies and selectivity profiling.

<sup>1</sup>Department of Chemistry, University of California, Berkeley, Berkeley, CA, USA.

<sup>2</sup>Novartis-Berkeley Center for Proteomics and Chemistry Technologies, Berkeley, CA, USA.

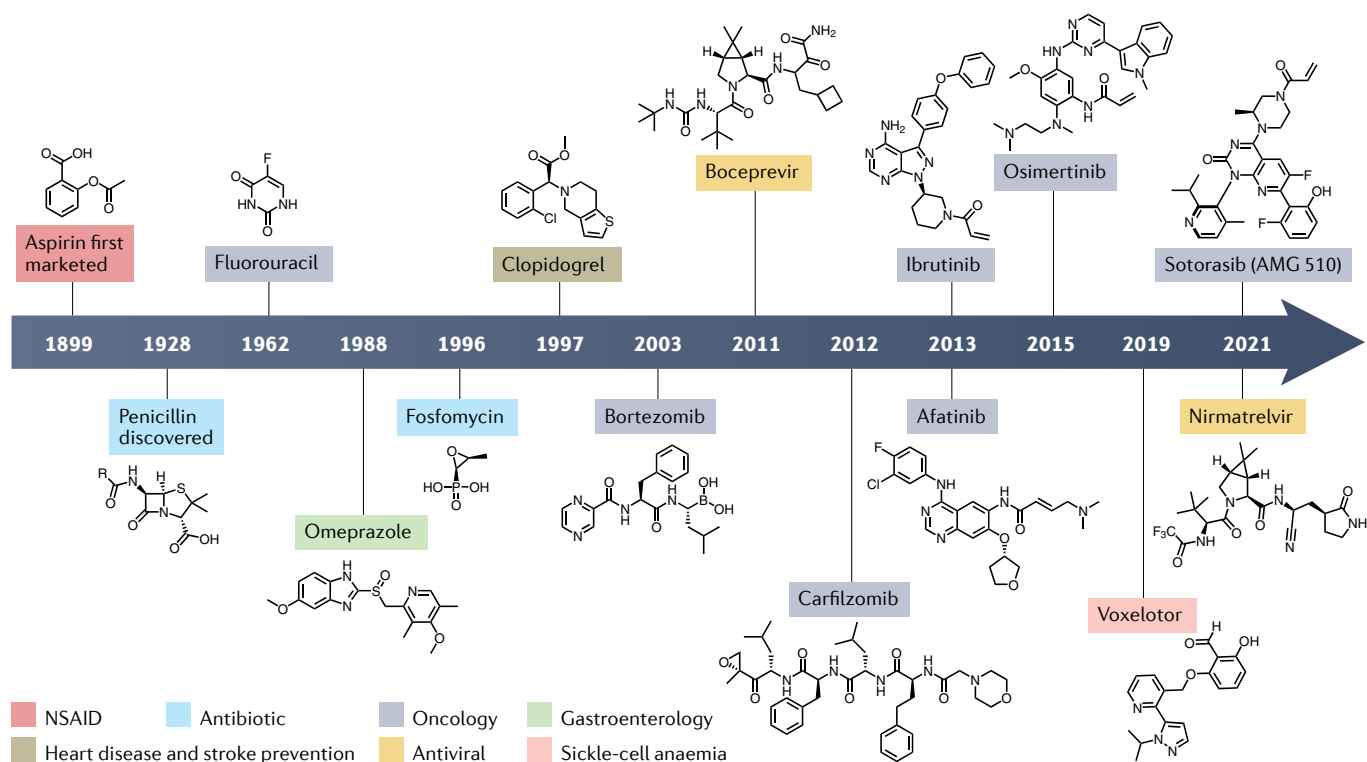
<sup>3</sup>Innovative Genomics Institute, Berkeley, CA, USA.

<sup>4</sup>These authors contributed equally: Lydia Boike, Nathaniel J. Henning.

✉e-mail:

[dnomura@berkeley.edu](mailto:dnomura@berkeley.edu)

<https://doi.org/10.1038/s41573-022-00542-z>



**Fig. 1 | Timeline of the development of major covalent drugs.** Each covalent drug is classified according to the drug type or type of disease it treats. Unless otherwise indicated, the date refers to the first approval by the US Food and Drug Administration. NSAID, non-steroidal anti-inflammatory drug.

### History of covalent drugs

Compounds that contain protein-reactive functional groups have often been avoided in medicinal chemistry and excluded from compound screening collections owing to their potential for assay interference and off-target promiscuity. Many historical examples of covalent drugs were discovered to act through covalent mechanisms after their use was already widespread. One of the most prominent among these is the non-steroidal anti-inflammatory drug (NSAID) aspirin, which has been marketed since 1899 (REF.<sup>19</sup>) (FIG. 1). Aspirin's mechanism of action was unknown until 1971, when it was discovered to exert its anti-inflammatory effects by acetylating Ser529 in the substrate-binding channel of cyclooxygenase 1, preventing conversion of the substrate arachidonic acid into prostaglandins<sup>20</sup>.

Early covalent drugs also tend to be derived from or inspired by natural sources.  $\beta$ -Lactam antibiotics such as penicillin (FIG. 1), produced by *Penicillium* fungi, bind to penicillin-binding proteins (PBPs), which are involved in bacterial cell wall synthesis<sup>21</sup>. All PBPs contain active-site serine residues that can be acylated by penicillin, inhibiting PBP activity and leading to cell membrane rupture<sup>21</sup>. Another covalent antibiotic is the epoxide-containing fosfomycin (FIG. 1), which is produced by some *Streptomyces* bacteria and acts by reacting with the catalytic cysteine of UDP-*N*-acetylglucosamine-enolpyruvyl transferase (MurA) to disrupt peptidoglycan synthesis and induce membrane rupture<sup>22–24</sup>.

Some covalent drugs are prodrugs with thiol-containing metabolites that form disulfide bonds to inactivate their

targets<sup>25</sup>. The proton pump inhibitor omeprazole (FIG. 1), approved by the US Food and Drug Administration (FDA) in 1988 to treat gastro-oesophageal reflux disease, is an example of this and is also a drug that was brought to market before its mechanism of action was understood to be covalent. Both omeprazole and clopidogrel (FIG. 1), an antiplatelet medication used to prevent strokes and heart attacks, are activated by cytochrome P450 enzymes in the liver to produce bioactive thiol metabolites<sup>26</sup>.

Covalent drugs have also been historically significant in cancer therapy. The pyrimidine nucleoside analogues 5-fluorouracil<sup>27,28</sup> and gemcitabine<sup>29</sup> are prodrugs used to inhibit thymidylate synthase and ribonucleotide reductase I, respectively, to treat a wide range of cancers (FIG. 1). Bortezomib (FIG. 1), a dipeptide boronic acid that covalently binds to and inhibits a catalytic threonine residue of the 26S proteasome, was approved by the FDA in 2003 to treat patients with multiple myeloma<sup>30</sup>.

Covalent drugs have been used to treat a variety of diseases. However, focusing on covalency from the outset of a project, instead of discovering a covalent mechanism of action after the fact, provides opportunities to improve drug design. Recent work in this field showcases how covalent drug discovery tools present solutions to otherwise intractable drug discovery challenges.

### Discoveries by ligand-first approaches

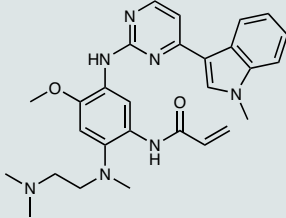
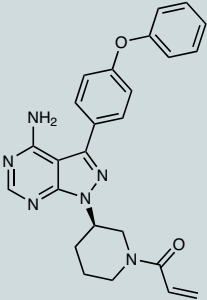
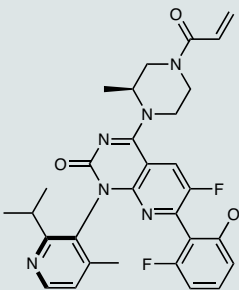
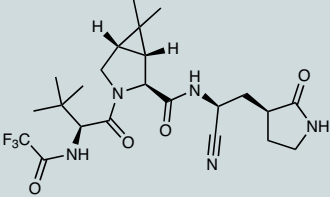
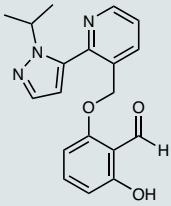
Major milestones of covalent drug discovery have been reached over the past decade, including the FDA approval of the first covalent EGFR inhibitor, afatinib,

in 2013, the BTK inhibitor, ibrutinib, in 2013 and the discovery of other kinase inhibitors. To discover these compounds, mildly reactive electrophilic functional groups were incorporated into known reversible ligands to enhance the inhibition of protein function. These examples offer lessons for future programmes, as each compound must balance reactivity, potency and selectivity.

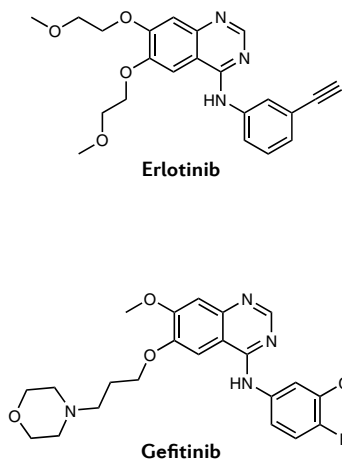
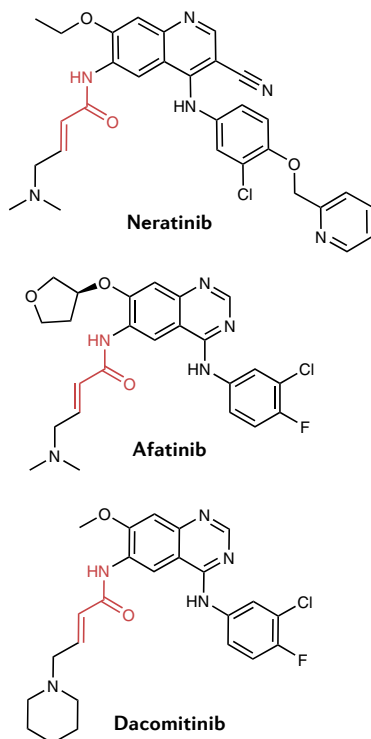
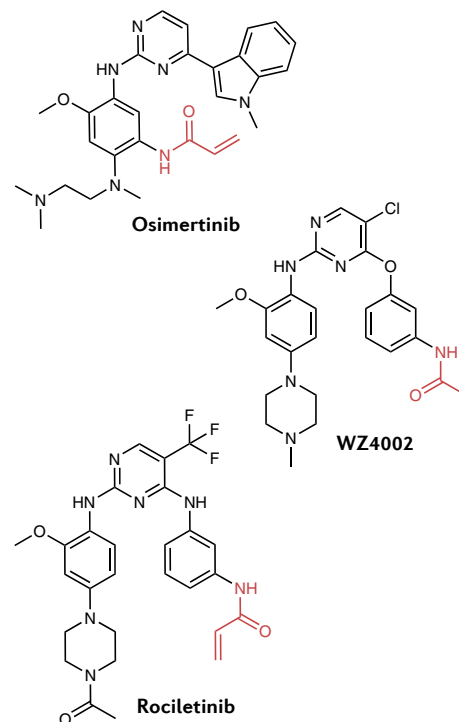
### Covalent EGFR inhibitors

Overactivity of the receptor tyrosine kinase EGFR drives the progression of non-small-cell lung cancer (NSCLC), making EGFR a key drug target in oncology<sup>31</sup>. During clinical development in the early 2000s, the reversible, first-generation EGFR inhibitors gefitinib and erlotinib (FIG. 2a) were discovered to be effective against tumours harbouring somatic activating mutations in *EGFR*,

Table 1 | Key examples of covalent drugs

Drug (company; former compound name)	Structure	Target	Approval <sup>a</sup>	Refs.
Osimertinib (AstraZeneca; AZD9291)		Mutant-selective EGFR inhibitor	2015 for treatment of NSCLC	44,45,48,49
Ibrutinib (AbbVie; PCI-32765)		BTK inhibitor	2013 for treatment of mantle-cell lymphoma and, subsequently, many other B cell malignancies	55,56,60
Sotorasib (Amgen; AMG-510)		KRAS(G12C) inhibitor	2021 for treatment of NSCLC with KRAS <sup>G12C</sup> mutation	93–95
Nirmatrelvir (Pfizer; PF-07321332)		SARS-CoV-2 main protease inhibitor	Authorized for emergency use by FDA in 2021 for treatment of COVID-19 <sup>b</sup>	103,105
Voxelotor (Global Blood Therapeutics; GBT-440)		Mutant-haemoglobin modulator	2019 for treatment of sickle cell anaemia	142,143,145

BTK, Bruton's tyrosine kinase; COVID-19, coronavirus disease 2019; EGFR, epidermal growth factor receptor; NSCLC, non-small-cell lung cancer; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2. <sup>a</sup>Unless otherwise indicated, the date of the first approval by one of the major regulatory agencies is provided, which was the FDA in each case. <sup>b</sup>Nirmatrelvir is approved for use in combination with ritonavir (Paxlovid). The FDA's emergency use authorization preceded approval by other agencies, but Paxlovid has not yet been fully approved by the FDA.

**a Reversible first-generation EGFR inhibitors**

**b Covalent second-generation EGFR inhibitors**

**c Covalent third-generation T790M mutant-selective EGFR inhibitors**


**Fig. 2 | Progression of EGFR inhibitor structures.** Progression from first-generation (part **a**) to second-generation (part **b**) epidermal growth factor receptor (EGFR) inhibitors involved the addition of a reactive acrylamide electrophile (highlighted in red) to covalently bind a cysteine residue (Cys797) in EGFR. In the progression from second-generation to third-generation (part **c**) EGFR inhibitors, the quinazoline moiety is replaced with a pyrimidine unit to provide selectivity for the T790M mutant.

either deletions in exon 19 or the L858R point mutation, which occur in 10–30% of patients with NSCLC<sup>31–33</sup>. However, the disease in these patients eventually still progressed; in 60% of cases this was due to the acquisition of the T790M ‘gatekeeper mutation’<sup>34,35</sup>. This mutation of the gatekeeper residue in the ATP-binding site of EGFR not only decreases the binding affinity of many reversible inhibitors for EGFR but also increases the binding affinity of EGFR for ATP<sup>36</sup>.

To overcome this problem, covalent second-generation inhibitors were strategically designed with acrylamide Michael acceptors to react with a cysteine residue (Cys797) in EGFR (FIG. 2b). Cys797 is located adjacent to the ATP-binding site, and irreversible binding of EGFR ligands to EGFR partially restores activity against the T790M gatekeeper mutant<sup>33</sup>. In addition to modest activity against T790M, covalent second-generation inhibitors provided prolonged suppression of EGFR signalling, suggesting that these covalent EGFR inhibitors could be more efficacious than reversible first-generation inhibitors such as erlotinib<sup>33</sup>. Afatinib (Boehringer Ingelheim) (FIGS. 1 and 2b) was approved by the FDA in 2013 as a first-line treatment for patients with metastatic NSCLC with activating mutations in *EGFR*<sup>37,38</sup>. Despite the increased potency that covalent engagement brought against the disease target, the dose-limiting toxicity caused by inhibition of wild-type EGFR likely prevented afatinib from increasing overall survival when compared head-to-head with platinum-based

chemotherapy in treating cancers bearing the T790M gatekeeper mutation<sup>39</sup>. Other second-generation inhibitors include neratinib (Puma), which potently inhibits HER2 by covalently binding Cys805 (a cysteine residue homologous to Cys797 on EGFR) and was approved by the FDA for treatment of HER2<sup>+</sup> breast cancer in 2017, and dacomitinib (Pfizer) (FIG. 2b), which was approved by the FDA to treat NSCLC in 2018 (REFS. 40–42).

A third generation of EGFR inhibitors followed afatinib; these covalent inhibitors selectively target the T790M mutant over wild-type EGFR, and include WZ4002 (Dana–Farber Cancer Institute)<sup>43</sup>, osimertinib<sup>44,45</sup> and rociletinib (Clovis Oncology; CO-1686)<sup>46</sup> (FIG. 2c). These compounds maintain the acrylamide group to covalently bind Cys797 but exchange the quinazoline moiety of first-generation and second-generation compounds for a pyrimidine to promote selectivity for mutant EGFR(T790M) (REF. 47). Higher affinity for T790M over wild-type EGFR not only results in efficacy in cancers with the EGFR gatekeeper mutation but also contributes to an improved safety profile and enables a higher recommended dose for osimertinib than for afatinib<sup>48</sup>. Osimertinib was granted accelerated approval by the FDA in 2015 as a second-line treatment for NSCLC, and was approved as a first-line treatment for metastatic NSCLC in 2018 (REF. 49). However, osimertinib depends on Cys797 for covalent binding, and C797X mutations account for 15% of cases of resistance to second-line osimertinib<sup>50–52</sup>. Generally, drugs whose efficacy relies

**Gatekeeper mutation**

Mutation that alters an amino acid residue at a specific location in the ATP-binding site of a kinase, known as the gatekeeper residue; alterations at this position have a particularly strong effect on binding of certain kinase inhibitors.

**Michael acceptors**

Electrophiles that can react with a nucleophile through Michael addition, often containing an  $\alpha$ ,  $\beta$ -unsaturated carbonyl and an electron-withdrawing group.

on covalent binding to a specific nucleophilic amino acid are vulnerable to mutations at that site, which could lead to drug resistance.

The success of covalent EGFR inhibitors has validated the approach of covalently engaging non-catalytic, non-conserved cysteines adjacent to kinase active sites to increase the potency and modulate the pharmacodynamics of initially reversible ligands. Development of these drugs has shown that the acrylamide electrophile is reactive enough to engage a cysteine adjacent to an ATP-binding site but not so reactive as to induce haptization and an adverse immune response. Incorporating covalent binding in EGFR inhibitors also enables selectivity between kinases through an interaction with a non-conserved cysteine instead of highly conserved active-site residues that typically interact with ATP.

### Covalent BTK inhibitors

The discovery of covalent BTK inhibitors shares several themes with the discovery of covalent EGFR inhibitors, including the ligand-first approach and the use of Michael acceptor electrophiles. BTK became a target of interest in chronic lymphocytic leukaemia owing to its crucial role downstream of the B cell receptor<sup>53</sup>. Activation of the B cell receptor induces phosphorylation of BTK through Lyn and Syk kinases, and eventually activates transcription factors related to B cell proliferation, differentiation, cell migration and adhesion<sup>54</sup>. This key role in B cell development indicated that BTK was a relevant target for B cell malignancies.

In the early 2000s, scientists at Celera Genomics who were interested in using BTK inhibitors to treat rheumatoid arthritis used a structure-based approach to discover an acrylamide-containing inhibitor of the BTK kinase domain that could be used as a tool compound to fluorescently label BTK<sup>55</sup>. It was subsequently discovered that the tool compound itself, later named ibrutinib (TABLE 1), had sufficient activity and suitable physicochemical properties to advance into clinical studies<sup>56–58</sup>. Ibrutinib was approved by the FDA for the treatment of mantle-cell lymphoma in 2013 and subsequently for chronic lymphocytic leukaemia (CLL), Waldenstrom's macroglobulinaemia and chronic graft versus host disease<sup>59–63</sup>.

Similarly to EGFR inhibitors, ibrutinib binds to a cysteine residue (Cys481) adjacent to the ATP-binding site in BTK, and because only a few kinases have a homologous cysteine, ibrutinib should exhibit a degree of selectivity for BTK over other kinases<sup>64</sup>. The rapid clearance of ibrutinib (which has a half-life of 2–3 h) could also enable kinase selectivity; ibrutinib should maintain activity against BTK owing to prolonged covalent engagement, while the reversible inhibition of off-targets is minimized<sup>65</sup>. This combination of fast covalent engagement of BTK with rapid clearance might allow for selectivity *in vivo* despite the off-target kinase inhibition observed in biochemical assays.

Several other covalent BTK inhibitors have been approved or are currently in clinical trials and some of these highlight the variety of Michael acceptors that can be used as alternative electrophiles to acrylamides<sup>66–68</sup>.

Most prominent among these is acalabrutinib (AstraZeneca), approved by the FDA in 2019 to treat CLL, which contains a butyramide electrophile instead of an acrylamide<sup>69</sup>. The butyramide electrophile is less reactive than an acrylamide, which, in addition to other substitutions, is hypothesized to account for the superior selectivity of acalabrutinib compared with ibrutinib for BTK and could be responsible for the reduced number of adverse cardiovascular events<sup>70,71</sup>. Further work has examined the use of cyanoacrylamides as electrophiles to design reversible covalent BTK inhibitors, which would ideally show increased potency and lower covalent off-target reactivity than irreversible covalent BTK inhibitors<sup>72–74</sup>. The long, tunable off-rates of reversible covalent inhibitors highlights the grey area that exists between reversible inhibition and irreversible covalent mechanisms.

Overcoming historical concerns relating to the potential toxicity of covalent drugs, the success of ibrutinib demonstrates that rationally designed covalent drugs can achieve acceptable safety profiles and blockbuster status. Ibrutinib, and covalent EGFR inhibitors, demonstrate that kinase inhibitors that target non-conserved cysteines adjacent to the ATP-binding site can be developed into selective and potent drugs. The pharmacokinetic and pharmacodynamic properties of ibrutinib allow for prolonged BTK blockade while reducing off-target kinase inhibition through rapid clearance *in vivo*. Notably, the performance of ibrutinib in treating B cell malignancies emphasizes that molecules once considered chemical biology tool compounds can become effective drugs.

### Other covalent kinase inhibitors

Covalent inhibitors have been used to selectively target kinases other than EGFR and BTK with non-conserved cysteine residues adjacent to their ATP-binding sites<sup>75,76</sup>. One example is Janus kinase 3 (JAK3), a non-receptor tyrosine kinase primarily expressed in leukocytes and involved in cytokine signalling<sup>77</sup>. Covalent targeting of the non-conserved Cys909 of JAK3 has yielded inhibitors selective for JAK3 over other JAK family members for the treatment of autoimmune diseases<sup>77–81</sup>. One of these inhibitors, ritlecitinib (Pfizer; PF-06651600), has shown promising results for patients with rheumatoid arthritis in a phase II clinical trial<sup>82</sup>. Several covalent inhibitors of fibroblast growth factor receptor 4 (FGFR4) target the non-conserved Cys552 residue in FGFR4 to confer selectivity over FGFR1, FGFR2 and FGFR3, as well as to overcome mutations that confer resistance to reversible FGFR inhibitors in hepatocellular carcinoma (HCC)<sup>83,84</sup>. The acrylamide-containing FGFR4 inhibitor fisogatinib (Blueprint Medicines; BLU-554) is currently the subject of a phase II clinical trial (NCT04194801). Aldehyde-containing roblitinib (Novartis; FGF401), which is a reversible covalent FGFR4 inhibitor that also reacts with Cys552, is also under clinical investigation (NCT02325739)<sup>83,85</sup>. Overall, the rational design of covalent kinase inhibitors that target non-conserved cysteines adjacent to the ATP-binding site has become a routine approach to enhancing the potency and selectivity of kinase inhibitors.



### Discoveries by electrophile-first approaches

Covalent drugs are also discovered through electrophile-first approaches, meaning that the initial discovery process is rooted in finding a covalent ligand from the outset, instead of incorporating covalency into a known reversible ligand. Key examples of drugs discovered through this approach include the KRAS(G12C) inhibitor sotorasib and the SARS-CoV-2 M<sup>Pro</sup> inhibitor nirmatrelvir (FIG. 1).

### Covalent KRAS(G12C) inhibitors

The discovery and development of covalent KRAS(G12C) inhibitors is one of the most exciting discovery-to-clinic stories featuring covalent drugs. *KRAS* is a GTPase-encoding oncogene that is mutated in about 25% of all cancers, most notably in pancreatic, colorectal and lung cancers<sup>86</sup>. Wild-type *KRAS* is carefully regulated between the active GTP-bound state and inactive GDP-bound state, but many *KRAS* mutations attenuate GTPase activity, leading to low rates of GTP hydrolysis and elevated RAS signalling, driving tumorigenesis<sup>87</sup>. Since the discovery of the role of *KRAS* in cancer nearly 30 years ago, attempts to drug it directly using traditional drug discovery methods have been unsuccessful<sup>9,10</sup>. *KRAS* does not have accessible pockets for reversible inhibitors to bind to, competitive inhibitors would need to overcome the picomolar binding affinities of GTP and GDP, and inhibitors active against wild-type *KRAS* could show on-target toxicity<sup>86,88</sup>.

Covalent *KRAS* inhibitors against the G12C mutant are appealing for several reasons. First, targeting mutant *KRAS* could allow for selective cytotoxicity to cancer cells. Second, the affinity enabled by covalent binding would be advantageous as *KRAS* lacks easily ligandable pockets. Third, 12–14% of *KRAS* mutations in NSCLC are *KRAS*<sup>G12C</sup>, presenting a promising patient group that would directly benefit from *KRAS*(G12C) inhibition<sup>89</sup>. Finally, position 12 in *KRAS* sits closely beneath the effector-binding region and the nucleotide-binding pocket, suggesting that covalent *KRAS*(G12C) ligands might affect *KRAS* function<sup>87</sup>.

In 2013, researchers at the University of California, San Francisco reported the first mutant-selective covalent *KRAS*(G12C) inhibitor. The inhibitor (compound 12 in their study) was discovered through a disulfide-fragment screening approach known as tethering, whereby a library of 480 disulfide fragments was screened against *KRAS*(G12C) in the GDP-bound state using intact protein mass spectrometry (MS)<sup>88,90</sup>. Co-crystal structures of *KRAS*(G12C) showed that hit compounds bound to the switch II region, and subsequent medicinal chemistry efforts to exchange the disulfide moiety for acrylamide and vinyl sulfonamide electrophiles yielded *KRAS*(G12C) inhibitors that were active in vitro, including compound 12. Binding of compound 12 to the switch II pocket impaired *KRAS* signalling by shifting nucleotide affinity from favouring GTP to GDP and led to the accumulation of *KRAS* in its inactive state<sup>91</sup>.

This novel mechanism for selective *KRAS*(G12C) inhibition set the stage for the development of clinical covalent *KRAS*(G12C) inhibitors. In 2016, WellSpring

Biosciences disclosed ARS-853, which is a selective covalent inhibitor of *KRAS*(G12C) with in cellulo efficacy in the low micromolar range<sup>92</sup>. Structure-guided optimization of compound 12 and use of a cellular liquid chromatography with tandem mass spectrometry (LC-MS/MS)-based assay to determine the degree of *KRAS*(G12C) engagement in H358 cells, yielded ARS-853 (REF.<sup>92</sup>). ARS-853 treatment in *KRAS*(G12C)-dependent cell lines decreased the amount of active *KRAS*(G12C), inhibited downstream RAS signalling and induced apoptosis<sup>92</sup>. Although *KRAS*(G12C) had been thought to be constitutively active, the selective binding of ARS-853 to GDP-bound, inactive *KRAS*(G12C) provided evidence that *KRAS* mutants cycle between GTP-bound and GDP-bound states<sup>92</sup>.

The discovery of clinical *KRAS*(G12C) inhibitors continued with ARS-1620, which was the result of an effort to overcome metabolic stability and bioavailability limitations of ARS-853 to facilitate in vivo studies of *KRAS*(G12C) inhibition<sup>87</sup>. ARS-1620 is based on a novel quinazoline core scaffold, designed to better occupy the switch II pocket and, thus rigidify a more favourable conformation for covalent reaction between the acrylamide electrophile and cysteine<sup>87</sup>. Ultimately, ARS-1620 was identified as the first *KRAS*(G12C) inhibitor suitable for in vivo studies and showed efficacy in *KRAS*(G12C) patient-derived xenograft models treated at 200 mg kg<sup>-1</sup> once per day or twice per day<sup>87</sup>. The increased potency of this series of *KRAS*(G12C) inhibitors and success in in vivo models indicated that it might be possible to design clinically efficacious drugs.

Sotorasib (AMG-510) (TABLE 1) was the first selective *KRAS*(G12C) inhibitor to enter clinical trials in 2018 and was developed by Amgen, building on discoveries from a partnership with Carmot Therapeutics in which a custom library of small molecules that covalently bind cysteine were screened against *KRAS*(G12C)<sup>93</sup>. Molecules identified through this collaboration led to the discovery of a previously unknown pocket on *KRAS* (a cryptic pocket), which Amgen scientists exploited to discover sotorasib through structure-based design<sup>94</sup>. Sotorasib was designed to occupy the cryptic pocket by interacting with His95, Tyr96 and Gln99 (REF.<sup>94</sup>) (FIG. 3). A phase II clinical trial investigating sotorasib was successfully completed in 2020, and was followed by FDA approval for the treatment of adults with *KRAS*<sup>G12C</sup>-mutated locally advanced or metastatic NSCLC in May 2021 (REF.<sup>95</sup>). Other covalent *KRAS*(G12C) inhibitors are entering clinical trials. Adagrasib (MRTX849) emerged from a joint drug discovery collaboration between Mirati Therapeutics and Array BioPharma, in which irreversible covalent inhibitors of *KRAS*(G12C) were identified; Mirati Therapeutics subsequently used structure-based design approaches to optimize adagrasib, which entered clinical trials in January 2019 (REFS.<sup>96,97</sup>) (FIG. 3). JNJ-74699157 (ARS-3248; J&J and WellSpring Biosciences) was being investigated in patients with several types of advanced solid tumour that express *KRAS*<sup>G12C</sup>, including NSCLC and colorectal cancer, but its clinical trials have been terminated<sup>98</sup>.

#### Cryptic pocket

A 3D pocket that forms only when a ligand is bound to a protein; these pockets are hidden when the ligand is not bound to the protein.

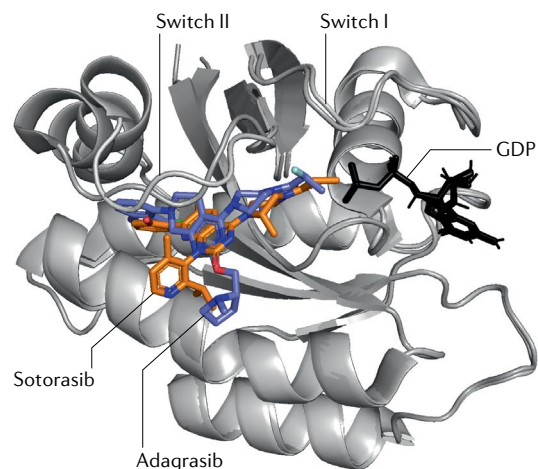
Designing small-molecule covalent KRAS(G12C)-selective inhibitors provides an elegant solution to drugging an undruggable cancer target. Before KRAS(G12C) inhibitors, recently discovered targeted covalent inhibitors in oncology were mostly identified using ligand-first approaches. The success of covalent KRAS(G12C) inhibitors validates an electrophile-first approach to covalent drug discovery and affirms the importance of covalent fragment screening techniques (discussed below). Optimization of initial hit compounds that emerge from covalent screening platforms, such as compound 12, can subsequently lead to programmes that produce potent covalent inhibitors such as sotorasib. In addition, the sotorasib story suggests that in other diseases in which a key protein target undergoes substitution of an amino acid to a cysteine residue, covalent inhibitors present an increasingly validated method to potentially provide precision therapy for patients.

#### SARS-CoV-2 main protease inhibitors

Vaccines against coronavirus disease 2019 (COVID-19) were developed at unprecedented speeds, and similar research momentum has led to the development of therapeutics that will benefit patients with COVID-19. In December 2021, the FDA issued an Emergency Use Authorization for Pfizer's Paxlovid (a combination of nirmatrelvir and ritonavir) to treat mild-to-moderate COVID-19 (caused by SARS-CoV-2) in adults and some paediatric patients, marking the first approved oral treatment for the disease<sup>99</sup>. Nirmatrelvir (TABLE 1) covalently inhibits the M<sup>Pro</sup> of SARS-CoV-2 (REF.<sup>3</sup>). This programme highlights how the adaptation of relatively inactive peptidomimetics into potent and selective cysteine protease inhibitors can be accomplished by the addition of cysteine-reactive covalent functional groups to target the protease active site and by structurally informed medicinal chemistry efforts.

SARS-CoV-2 is a virus with a single-stranded RNA genome that encodes two polyproteins (pp1a and pp1ab) as well as structural and accessory proteins<sup>100</sup>. Viral replication depends on successful cleavage of pp1a and pp1ab by the M<sup>Pro</sup> (also referred to as 3CLpro), which is a cysteine protease, into functional viral proteins<sup>100</sup>. The discovery of covalent inhibitors against SARS-CoV-2 M<sup>Pro</sup> emerged from extensive work on protease inhibitors for SARS-CoV-1, which is the causative virus for severe acute respiratory syndrome coronavirus 1 (SARS1)<sup>100</sup>. During the 2002–2003 SARS1 outbreak, researchers used a crystal structure of the homologous porcine transmissible gastroenteritis coronavirus (TGEV) M<sup>Pro</sup> bound to a hexapeptidyl chloromethylketone covalent cysteine-reactive inhibitor to provide a foundation for the design of covalent inhibitors against SARS-CoV-1 M<sup>Pro</sup> (REF.<sup>101</sup>). Because of the homology across all M<sup>Pro</sup>, this work enabled the discovery of rupintrivir (AG7088), which is a mechanism-based inhibitor of the human rhinovirus (HRV) M<sup>Pro</sup> (REF.<sup>101</sup>).

Because the SARS-CoV-1 outbreak subsided, work into developing coronavirus M<sup>Pro</sup> inhibitors slowed until the emergence of SARS-CoV-2 in 2019. SARS-CoV-2 M<sup>Pro</sup> shares 96% sequence identity with SARS-CoV-1 M<sup>Pro</sup>, and there is 100% sequence overlap



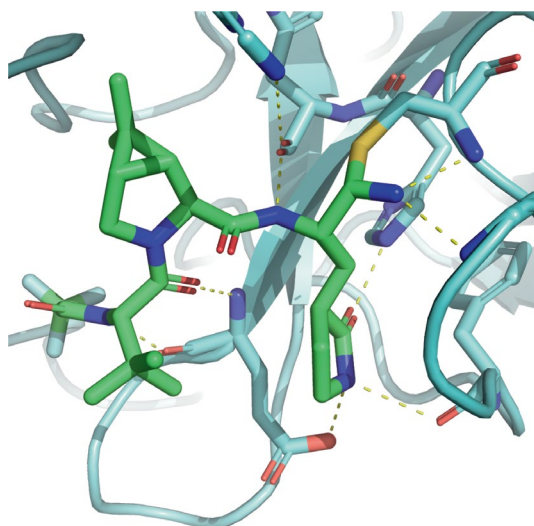
**Fig. 3 | Aligned structures of KRAS(G12C) co-crystallized with adagrasib (MRTX849) and sotorasib (AMG-510).** The covalent inhibitors adagrasib (PDB ID: 6UT0) and sotorasib (PDB ID: 6OIM) are bound to the switch II pocket, which is adjacent to the GDP-binding pocket<sup>93,96</sup>.

of the catalytic sites<sup>102</sup>. Renewed interest in improving on previous chloromethylketone inhibitors motivated researchers to adapt rupintrivir into a SARS-CoV-2 M<sup>Pro</sup> inhibitor potent enough to obtain a co-crystal structure. This discovery in turn enabled identification of the  $\alpha$ -hydroxymethylketone-containing antiviral PF-00835231, which demonstrated potent SARS-CoV-2 M<sup>Pro</sup> inhibition in an activity assay based on fluorescence resonance energy transfer (FRET), activity in antiviral cell-based assays, stability in plasma and low clearance in vivo<sup>103</sup>.

In subsequent studies, the oral bioavailability of PF-00835231 was improved by replacing the  $\alpha$ -hydroxymethylketone moiety with a nitrile group, which can also act as an electrophile<sup>103</sup>. Nitriles can covalently bind particularly reactive nucleophiles; however, the ease of thiol elimination from the thioimidate adduct makes nitriles more reversible than some other electrophiles, such as acrylamides<sup>104</sup>. Optimization from PF-00835231 eventually yielded PF-07321332, named nirmatrelvir (FIG. 4 and TABLE 1), which is a highly potent SARS-CoV-2 reversible covalent inhibitor that displayed potent inhibition in a FRET-based assay across all human coronaviruses, while no inhibitory effects were seen against human cysteine or serine proteases. The in vivo efficacy of nirmatrelvir was demonstrated in a mouse-adapted SARS-CoV-2 (SARS-CoV-2 MA10) model.

In phase II/III clinical trial data released in November 2021, Paxlovid was shown to be highly effective at preventing progression to severe COVID-19 in symptomatic patients<sup>105</sup>. The emergence of this orally bioavailable drug for COVID-19 will help to ameliorate illness for non-hospitalized patients in high-risk groups<sup>106</sup>. Overall, SARS-CoV-2 M<sup>Pro</sup> covalent inhibitors provide a promising avenue for the treatment of coronavirus infections either as monotherapies or in combination with other antiviral drugs.

The quick adaptation of previous protease inhibitors to selectively target SARS-CoV-2 M<sup>Pro</sup> is an example of



**Fig. 4 | Nirmatrelvir in complex with SARS-CoV-2 main protease.** The nitrile group of nirmatrelvir (shown in green) reacts with Cys145 of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (shown in light blue) to form a covalent thioimidate adduct. Extensive hydrogen-bond interactions (yellow dashed lines) occur throughout the pocket (PDB ID: 7RFW)<sup>103</sup>.

using structure-based design while taking into account the valuable properties of covalent drugs. Researchers started with an electrophile-containing peptide and optimized both the peptide and electrophile to obtain a highly potent covalent inhibitor. In some ways, this story mirrors that of the discovery and optimization of JAK3 kinase inhibitors, in which EGFR inhibitors were adapted to target Cys909 of JAK3 (REF.<sup>79</sup>). Both JAK3 and mutant EGFR(T790M) contain a methionine gate-keeper, as well as a homologous cysteine adjacent to the ATP-binding site. These shared features enabled the discovery of potent JAK3-selective covalent inhibitors using an EGFR inhibitor as a starting point, similar to how rupintrivir provided the starting structure for the discovery of nirmatrelvir. Generally, when reactive cysteine residues are shared across a protein family, structurally guided adaptation of previously studied cysteine-reactive covalent inhibitors can lead to selective and potent drugs for other proteins with this feature.

#### HCV NS3/4a protease inhibitors

$\alpha$ -Ketoamide-based covalent inhibitors have been developed to treat hepatitis C virus (HCV) infection through inhibiting NS3/4a — a serine protease that cleaves the HCV polyprotein into multiple non-structural proteins required for replication<sup>107</sup>. Although HCV had been treated with a combination of PEGylated interferon- $\alpha$  and ribavirin, modest response rates and notable adverse events prompted studies that resulted in the discovery of NS3/4a protease inhibitors to treat HCV<sup>108</sup>. On the basis of initial observations that hexapeptide cleavage products could inhibit NS3/4a, the linear peptidomimetic inhibitors boceprevir (Merck)<sup>109,110</sup> (FIG. 1) and telaprevir (Vertex)<sup>111,112</sup> were designed. These compounds, along with narlaprevir<sup>113</sup>, use a ketoamide electrophile to covalently engage the catalytic serine of NS3/4a. This

covalent interaction is relatively reversible owing to elimination of the serine alcohol group from the protein–inhibitor adduct<sup>114</sup>. Boceprevir and telaprevir were effective in treating HCV and were approved by the FDA in 2011 after successful trials<sup>115,116</sup>. However, telaprevir was withdrawn in 2014 owing to adverse events, and boceprevir was discontinued by Merck in 2015 owing to the superiority of newer direct-acting antivirals, in particular the ledipasvir–sofosbuvir combination (Gilead), which targets the HCV polymerases NS5a and NS5b<sup>117–119</sup>. Nevertheless, the success of ketoamide-based NS3/4a inhibitors in increasing the efficacy of interferon–ribavirin therapy emphasizes the utility of the ketoamide group as a serine-reactive electrophile in designing covalent antivirals.

#### Covalent proteasome inhibitors

Bortezomib (Takeda) was the first boron-containing drug to be approved by the FDA, and was approved for treatment of multiple myeloma in 2003 (REF.<sup>120</sup>). Bortezomib was discovered through optimization of an aldehyde-containing proteasome substrate peptide, and set a precedent for the discovery of HCV NS3/4a inhibitors such as telaprevir<sup>121–124</sup>. Exchange of the aldehyde electrophile for a boronic acid substantially increased the potency of proteasome inhibition<sup>121</sup>. The effect of bortezomib comes from the boronic acid covalently binding to the hydroxy group of the  $\beta 5$  subunit N-terminal threonine of the 20S proteasome, leading to inhibition of the proteasome's chymotrypsin-like activity<sup>125</sup>. Bortezomib takes advantage of the increased sensitivity of haematological cancers such as multiple myeloma and mantle-cell lymphoma to proteasome inhibition<sup>123</sup>.

The development of bortezomib validated the proteasome as a cancer target, encouraging the discovery of other proteasome inhibitors. Medicinal chemistry efforts transformed the natural product epoxomicin into carfilzomib (Amgen), which was approved by the FDA in 2012 (REF.<sup>126</sup>) (FIG. 1). The epoxyketone moiety in carfilzomib forms a morpholino ring with the catalytic N-terminal threonine of the 20S proteasome, and this mechanism has been proposed to confer selectivity because most proteases do not have nucleophilic side chains at their N terminus<sup>127</sup>. Although the epoxyketone could have additional off-targets that bortezomib does not, this proposed mechanism of carfilzomib illustrates how covalency can help to drive selectivity through highly specific mechanisms.

Further work has been done to identify orally bioavailable proteasome inhibitors to improve upon bortezomib, which is administered intravenously<sup>128</sup>. Ixazomib (Takeda) is a second-generation proteasome inhibitor that also contains a boronic acid group<sup>129</sup>. This drug is administered orally as the prodrug ixazomib citrate, a boronic ester that hydrolyses upon exposure to aqueous media or plasma<sup>130,131</sup>. Oprozomib (Amgen), a second-generation epoxyketone-containing proteasome inhibitor, can also be orally administered<sup>132</sup>.

The discovery of these covalent proteasome inhibitors illustrates the utility of boronic acid electrophiles for targeting protease active sites and demonstrates how a covalent drug discovery project can use an



unoptimized electrophile as a starting point and subsequently introduce an alternative electrophile. Just as a chloromethylketone group in a SARS-CoV-2 ligand was optimized to the nitrile in nirmatrelvir, bortezomib was discovered through optimization from an aldehyde. Although targeting especially nucleophilic protease active sites might provide greater flexibility in terms of electrophile choice, using highly reactive electrophiles to gain an initial foothold can be beneficial.

#### Other boron-containing drugs

Several additional boron-containing drugs inhibit serine hydrolases beyond the proteasome through formation of covalent adducts between catalytic serine residues and boron<sup>133,134</sup>. Most of these drugs contain benzoxaborole groups and have been discovered through screening of compound collections that contain boron-based electrophiles<sup>135</sup>. For example, the antifungal tavaborole (Pfizer) was discovered through focused screening of boron-containing compounds previously investigated as antibacterials, and was approved by the FDA to treat onychomycosis in 2014 (REFS.<sup>136,137</sup>). Tavaborole covalently binds to the 2' and 3' hydroxy groups on the 3' terminus of leucyl-tRNA, trapping the tRNA-tavaborole adduct in the editing site of leucyl-tRNA synthetase to block protein synthesis<sup>138</sup>. Crisaborole (Pfizer), a phosphodiesterase 4 inhibitor, was approved by the FDA to treat psoriasis in 2016, and the  $\beta$ -lactamase inhibitor vaborbactam (Rempex) was approved to treat various bacterial infections in 2017 (REFS.<sup>139,140</sup>). These boron-based drugs target serine hydrolases, and the weak boron-sulfur bond potentially provides selectivity for serine over cysteine hydrolases<sup>121</sup>. The reversibility of the serine-boron bond hinders chemoproteomic profiling experiments commonly used to characterize the selectivity of covalent ligands. But as the serine hydrolase family is rich with potential drug targets, the numerous boron-containing drugs in clinical trials hold promise across a wide variety of disease types<sup>141</sup>.

#### Mutant-haemoglobin modulators

Voxelotor (Global Blood Therapeutics) is a lysine-targeting covalent drug used to treat sickle cell anaemia, and the discovery of voxelotor (TABLE 1) was dependent on knowledge of heightened lysine side chain reactivity. Sickle cell anaemia is caused by a single mutation in the gene encoding the  $\beta$ -haemoglobin chain that induces polymerization of mutant haemoglobin (HbS) under hypoxic conditions<sup>142</sup>. An aldehyde-containing natural product and several synthetic aldehyde analogues were found to prevent polymerization by increasing the affinity of HbS for oxygen<sup>143</sup>. These aldehydes bind in a reversible covalent manner, forming a Schiff base with the N-terminal valine of the  $\alpha$ -haemoglobin chain<sup>143</sup>. Almost 50 years ago, this N-terminal amine was discovered to have a particularly low  $pK_a$  of 6.9, indicating that it is primarily unprotonated under physiological conditions and, thus, more nucleophilic<sup>144</sup>. Based on earlier aldehydes, voxelotor was discovered through a structure-guided effort to discover compounds that increase the oxygen affinity of HbS, and was designed to bind the HbS tetramer in a 1:1 stoichiometry, unlike the

2:1 ratio of earlier compounds<sup>143,145</sup>. With a remarkable red blood cell to plasma ratio of  $\sim 150$  that likely reduces off-target effects, voxelotor was approved by the FDA in 2019 with a recommended dose of 1.5 g daily, which is an unusually high dose for a covalent drug<sup>142,146</sup>. The success of voxelotor is similar to that of ibrutinib in demonstrating that covalent drugs can be dosed at high amounts given favourable absorption, distribution, metabolism and excretion (ADME) properties. Furthermore, the discovery of voxelotor shows how the identification of unusually reactive amino acid residues, such as the  $\alpha$ -haemoglobin N-terminus, provides opportunities for drug discovery.

#### The covalent drug discovery toolbox

Many covalent drug discovery programmes, including those for covalent EGFR and BTK inhibitors, have involved the addition of reactive functional groups to previously identified ligands. However, emerging technologies make it possible to approach covalent ligand discovery from an electrophile-first perspective, in which covalent ligands against protein targets of interest are identified before structure-based optimization. For example, this type of approach proved successful in drugging KRAS(G12C) and has facilitated the rapid discovery of E3 ligase ligands for targeted protein degradation applications. Activity-based protein profiling (ABPP) approaches have transformed the characterization of electrophilic compounds, facilitating selectivity profiling and target identification experiments. Special considerations must also be made in evaluating the binding affinity and reactivity of covalent ligands (BOX 1). In this section, we discuss the toolbox of emerging techniques that covalent drug discovery relies on.

#### Screening platforms

Many emerging covalent ligand screening platforms involve MS-based detection, enabled by covalent bond formation between the compound and protein. Other phenotypic, DNA-encoded or computational approaches have also been used, which are then paired with MS-based validation, ABPP-based experiments to inform selectivity and structural biology to enable medicinal chemistry. The most prominent methods for covalent screening are summarized in TABLE 2. The growth of commercial libraries of electrophilic fragment-like compounds is a key factor contributing to the rise of these electrophile-first discovery strategies.

**Intact protein MS.** Originally, MS-based compound screening grew out of 'tethering', which is a technique employed since 2000 that uses libraries of compounds linked to disulfides to identify fragments that bind cysteine-adjacent pockets<sup>90,147</sup>. Molecules that bind undergo disulfide exchange with the cysteine (which could be endogenous or engineered) to form an adduct with the protein, and pooled screening with MS detection can identify the bound compound. Binding fragments can be combined or grown in a fragment-based approach to identify high-affinity ligands. This strategy was originally designed to identify reversible ligands for

#### Activity-based protein profiling

A chemical proteomics technique used to characterize enzyme function and amino acid reactivity in native biological systems that uses chemical probes to react with specific classes of proteins and amino acids.

**Box 1 | Characterizing the potency and reactivity of covalent ligands**

Although half-maximal inhibitory concentration ( $IC_{50}$ ) values are often used to measure the potency of covalent inhibitors, the  $k_{inact}/K_i$  second-order rate constant (where  $K_i$  is the inhibition constant and  $k_{inact}$  is the rate of enzyme inactivation) is preferred because it describes inhibitor potency in a time-independent manner<sup>212,213</sup>. The ratio  $k_{inact}/K_i$  can be determined by measuring total occupancy over time in a binding assay, calculating the first-order rate constant of inhibition,  $k_{obs}$ , and in turn  $k_{inact}/K_i$  (REF. 212). One example of the direct measurement of  $k_{inact}/K_i$  involved a time-resolved fluorescence energy transfer assay in which competitive fluorescent probes were used to determine  $k_{inact}/K_i$  values versus epidermal growth factor receptor (EGFR)<sup>214</sup>. Because this kinetic parameter can be resource-intensive to measure, researchers have formulated models that relate  $IC_{50}$  values to  $k_{inact}/K_i$  (REFS. 215–217). For example, a model has been developed to estimate  $K_i$  and  $k_{inact}$  from time-dependent  $IC_{50}$  data, from work on CYP450 enzyme inhibitors<sup>216</sup>. Moreover, a method has been designed that uses a competitive covalent probe with a known  $k_{inact}/K_i$  value to facilitate the rapid calculation of inhibitor  $k_{inact}/K_i$  values based on standard end point  $IC_{50}$  data<sup>217</sup>. Scientists working on a covalent JAK3 inhibitor programme at Pfizer, however, suggested that fixed-time-point  $IC_{50}$  values can serve as a valuable surrogate for  $k_{inact}/K_i$  (REF. 218). When using  $IC_{50}$  as a surrogate, extra care should be taken to confirm that target engagement is covalent (such as by measuring off-rates). However, with an appropriate time-point choice,  $IC_{50}$  values can be a useful tool to rapidly assess potency for covalent inhibitors. Regardless, determination of  $k_{inact}/K_i$  provides the most complete information to medicinal chemists about the binding affinity and kinetics of covalent target engagement.

To ensure that potent covalent inhibitors are not promiscuous ligands that lack selectivity, electrophilic compounds are frequently tested for reactivity, usually with respect to thiols. This measurement is often carried out using a glutathione (GSH) reactivity assay, whereby a compound is incubated with excess GSH and consumption of the compound is observed using liquid chromatography–mass spectrometry to determine a rate constant<sup>99,219,220</sup>. Alternatively, plate-based 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB; also known as Ellman's reagent) and NMR-based GSH assays have also been used<sup>156,221</sup>. In a helpful resource for acrylamide-focused projects, two related studies used an NMR-based assay to explore the effects of various substitutions on acrylamide GSH reactivity<sup>222,223</sup>. Assaying for reactivity is a crucial step, particularly for key molecules and for compound series containing novel electrophiles.

challenging targets, but covalent binding can be maintained by replacing the disulfide with an electrophile such as an acrylamide, as in the case of the initial covalent KRAS(G12C) inhibitors<sup>88</sup>. This approach has been used to discover compounds that modulate protein–protein interactions<sup>148,149</sup>, and a more recent study employed a similar tethering strategy that uses aldehydes to form imines with lysine residues<sup>150</sup>.

Over the past decade, covalent ligand discovery has shifted towards screening more drug-like electrophilic fragments<sup>13</sup>. In 2012, researchers curated 177 electrophilic compounds from the Pfizer compound collection and used an MS-based primary screening strategy to identify covalent inhibitors of the interaction between hypoxia-inducible factor 1 $\alpha$  (HIF1 $\alpha$ ) and aryl hydrocarbon receptor nuclear translocator (ARNT)<sup>151</sup>. This approach was informed by determining an X-ray crystal structure of the HIF1 $\alpha$ –ARNT complex<sup>151</sup>. Around the same time, the concept of tethering was expanded by using a small set of acrylamides and an MS-based assay to identify thymidylate synthase inhibitors, an approach termed kinetic template-guided tethering<sup>152</sup>. Building on these studies, an acrylate functionality was appended to 100 fragments to identify non-peptidic inhibitors of the cysteine protease papain<sup>153</sup>. Electrophilic fragments were pooled, and screening through electrospray ionization (ESI) MS led to the identification of hit compounds out of pooled experiments. Use of a similar library of

acrylates enabled the discovery of covalent inhibitors for the HECT E3 ligase NEDD4-1 (REF. 154). As with the HIF1 $\alpha$ –ARNT protein–protein interaction inhibitors discovered at Pfizer, co-crystal structures were crucial to understanding the mechanism of these NEDD4-1 inhibitors, which prevent association of ubiquitin with the E3 ligase and thus induce a switch from a processive to a distributive mechanism. In another example, acrylate-based inhibitors of the RBR E3 HOIP were also discovered using MS-based screening, highlighting how covalent fragment screening approaches can be useful for protein classes that have been challenging to discover ligands against, such as E3 ligases<sup>155</sup>.

More recently, a commercial library of 993 acrylamides and chloroacetamides was screened using intact MS to identify ligands of the deubiquitinase OTUB2 and the pyrophosphatase NUDT7 (REF. 156). The authors used co-crystal structures with OTUB2 and NUDT7 in complex with the hit compounds to inform fragment growing to increase potency. Although previous studies had paired MS-based screening with structural information to identify cysteine residues in functional sites or to understand the mechanism of inhibition, in this case, the pairing of MS-based screening and structure-guided fragment-based drug discovery supported optimization of the potency of the hit compounds.

The same compound collection was also used to screen against the peptidyl-proline cis–trans isomerase Pin1, which is overexpressed or activated in several tumour types but has been challenging to target selectively<sup>157</sup>. The resulting chloroacetamide sulfopin was shown to be selective for Pin1 in a covalent inhibitor target-site identification (CITE-Id) chemoproteomics experiment and was effective in regressing neuroblastoma growth in mice<sup>157</sup>. This result suggests that although chloroacetamides have disadvantages, including rapid metabolism, they can be valuable tool compounds with which to assess target relevance in various disease models.

One particularly powerful example of covalent ligand screening is the discovery of the initial compounds in the series that led to the first approved KRAS(G12C) inhibitor, sotorasib. A library of 3,300 acrylamides was screened in three assays: a thiol reactivity assay, a RAF-coupled nucleotide exchange assay and an intact MS assay<sup>158</sup>. Combined with crystallographic data that showed how ligand binding revealed the presence of previously closed sub-pockets, this effort provided the basis for the rapid discovery of KRAS(G12C) inhibitors discussed above.

**Covalent DNA-encoded libraries.** DNA-encoded libraries (DELs) present an alternative approach that enables the screening of massive libraries of small molecules. Unlike MS or ABPP approaches, there is no specialized advantage of covalency in enabling DEL screening, but the throughput of DELs allows for the screening of much larger covalent libraries through a workflow of immobilization, enrichment, amplification and sequencing. The first reports of electrophilic protein–nucleic acid-encoded libraries described the targeting of protease active sites<sup>159</sup>, but over the past 5 years, cysteine-targeted

**Fragment growing**

The medicinal chemistry process of optimizing a fragment hit to a higher-affinity ligand through increasing the size and complexity of the molecule.

**DNA-encoded libraries**

Collections of molecules linked to DNA sequences that serve as amplifiable barcodes, carrying information about each library member; this technology enables screening of massive libraries of compounds.

**Covalent docking**

A group of methods used to predict the orientation of a covalent ligand when bound to a protein target.

DNA-encoded or protein–nucleic acid-encoded libraries have been used to identify covalent ligands for bromodomains, including PCAF and BRD4, as well as for JNK1, MEK2 and HER2 (REFS.<sup>160–162</sup>). Further work has explored improvements in the enrichment step of the covalent DEL screening workflow, which differs because covalent engagement prevents elution of DNA-tagged molecules<sup>163</sup>. A covalent ligand of mitogen-activated protein kinase kinase 6 (MAP2K6) was also identified serendipitously through screening of a DEL against a DNA-encoded protein library<sup>164</sup>. More recently, even larger covalent DELs (with approximately 100,000,000 members) have been developed and used to identify acrylamide-based and epoxide-based BTK inhibitors with novel scaffolds<sup>165</sup>. In this study, similar screening results were obtained after storing the library at  $-80^{\circ}\text{C}$  for several years, suggesting that the electrophilic compounds are sufficiently stable in this context. Expansion of the covalent DEL library size and the increasing commercial availability of DELs represent an exciting development, and although unique considerations with respect to enrichment workflow and compound stability must be kept in mind, DELs that contain electrophilic molecules may become more widespread in covalent ligand discovery.

**Covalent docking.** The advantages of various covalent docking methods in different covalent docking scenarios have been described elsewhere<sup>166,167</sup>. Most computational programs for covalent docking rely on directly linking models to model conformations under the constraint of a predefined bond between a ligand and a corresponding amino acid site. The covalent docking platform GOLD relies on this assumption, and an example of its use was in the virtual screening for covalent inhibitors of

the NEDD8-activating enzyme (NAE), in which three of the hits were confirmed as novel NAE inhibitors<sup>168</sup>. Development of the DOCKoValent method for virtual screening facilitated the discovery of boronic acid AmpC  $\beta$ -lactamase inhibitors, as well as cyanoacrylamide inhibitors of JAK3 (REF.<sup>169</sup>). DOCKoValent uses non-covalent docking methods to pre-generate conformations and states for an electrophilic virtual library and then samples each state against the target nucleophile. The same method has been used to identify new covalent inhibitors of the kinase MKK7 (REF.<sup>170</sup>), as well as compounds that bind KRAS(G12C) to destabilize the protein and accelerate nucleic acid exchange<sup>171</sup>.

Apart from screening, covalent docking is a useful tool for investigating binding modes of known covalent ligands. For example, GOLD has been used to model the binding modes of the aldehyde-containing proteasome inhibitor MG132 (REF.<sup>172</sup>). AutoDock uses a flexible side chain approach, whereby the covalent ligand is treated as an amino acid side chain and poses of this flexible ‘side chain’ are scored, using a physics-based scoring function that evaluates the energetics of ligand–protein interactions, as the remainder of the protein is held rigid<sup>166</sup>. Another method, called CovDock, is based on the Schrödinger Glide docking algorithm and Prime structure refinement methodology. CovDock uses traditional non-covalent docking approaches to dock a ligand to a protein target, and then models the covalent bond attachment and refines the complex<sup>166</sup>. This approach does not consider the reactivity of the electrophile, which can limit the ability to virtually study the differences in docking ligands with different electrophilic functional groups<sup>167</sup>.

Overall, covalent docking software provides useful information on ligand–protein interactions when key

Table 2 | Comparison of screening methods for covalent drug discovery

Method	Key features	Limitations	Refs.
Intact protein MS	MS detection; compounds can be pooled; provides information on binding stoichiometry; high-throughput	Requires a large amount of purified protein; electrophile must be swapped when using disulfide tethering approaches	88,156
MS-based ABPP	MS detection; provides information on the selectivity of the ligand; can be done in complex biological systems (such as lysate or tissue)	Lower throughput than intact protein MS; sensitivity affected by peptide abundance and digestion conditions	180,183
Gel- or plate-based ABPP	Easy detection of hits through fluorescence of competitive reactive probe; can be paired with ABPP proteomics to assess selectivity of hit ligands	Lower throughput compared with assays that are not gel based; complicated by the presence of multiple reactive amino acid residues within a protein; requires purified protein	189
Phenotypic	ABPP enables rapid identification of covalent hits; potentially high-throughput; can use identical assays for discovering reversible ligands	Does not take advantage of known reactive hotspots identified in chemoproteomics data; requires control counter screens to eliminate untenable (for example, promiscuous or toxic) hit ligands	186,211
Covalent docking	Advanced software is widely available (such as DOCKoValent); scoring assesses non-covalent interactions	Difficult to model covalent bond formation (relies on assumptions about ligand–target interactions)	167,169
Covalent DNA-encoded libraries	Affords a massive library size	Enrichment complicated by covalent binding; limited library availability	165

ABPP, activity-based protein profiling; MS, mass spectrometry.

assumptions can be made — namely, when the reactivity of the electrophile and the site of modification is known.

#### **Chemoproteomics-enabled discovery**

Chemoproteomic platforms enable the identification of covalent compounds and their corresponding ligandable sites on target proteins directly in complex biological systems. Advances in chemoproteomics have facilitated the discovery of covalent ligands against undruggable disease targets and enabled the selectivity profiling of covalent ligands across the proteome to identify targets and off-targets of these ligands. Here, we discuss the chemoproteomics profiling of reactive ligandable hotspots, ABPP screening platforms and target identification within the context of recent work relevant to drug discovery. Chemoproteomics experiments can provide key information on ligand selectivity and give early guidance for selecting targets for covalent drug discovery programmes.

**Chemoproteomics profiling of reactive ligandable hotspots.** ABPP facilitates the discovery of covalently ligandable sites and the corresponding ligands in complex biological samples. ABPP was pioneered by Cravatt and Bogoy using active-site-directed chemical probes that covalently target catalytic residues of various enzyme classes, including hydrolases, proteases and kinases<sup>173,174</sup>. This technique, often using gel-based assays, was employed to gain functional readouts of active enzymes in biological contexts<sup>14,173,175</sup>. ABPP probes contain a warhead that covalently reacts with nucleophilic amino acids (such as cysteine) and a reporter handle to monitor probe binding, such as a fluorophore, biotin or alkyne moiety for subsequent click chemistry-enabled applications<sup>173</sup> (FIG. 5a).

Instead of focusing on active sites, more recent ABPP approaches use MS and broadly reactive chemical probes to also map allosteric sites<sup>176</sup>. In the first reports of the isoTOP-ABPP (isotopic tandem orthogonal proteolysis–activity-based protein profiling) approach, an iodoacetamide probe functionalized with an alkyne handle was used to identify hyper-reactive cysteines across the proteome<sup>176,177</sup> (FIG. 5b). The alkyne handle can be used to link the probe-modified protein to a tobacco etch virus (TEV) protease-cleavable tag that contains an azide group and biotin moiety separated by either an isotopically light or heavy valine (FIG. 5b). These functionalities enable enrichment of probe-modified peptides and tandem analysis of the light and heavy samples with MS, which, after controlling for run-to-run variability, allows for quantitative comparisons between samples, including competitive analysis of covalent compounds (FIG. 5b). Using this approach, it was discovered that the hyper-reactivity of cysteines predicts their functionality in catalysis and at sites of post-translational modifications<sup>177</sup>. Recent adaptations of isoTOP-ABPP have been developed to increase the coverage of cysteines across the proteome and to increase the throughput. For example, optimization of the sample preparation steps (namely single-pot, solid-phase-enhanced sample preparation (SP3)) and combination of this workflow with off-line fractionation and field asymmetrical

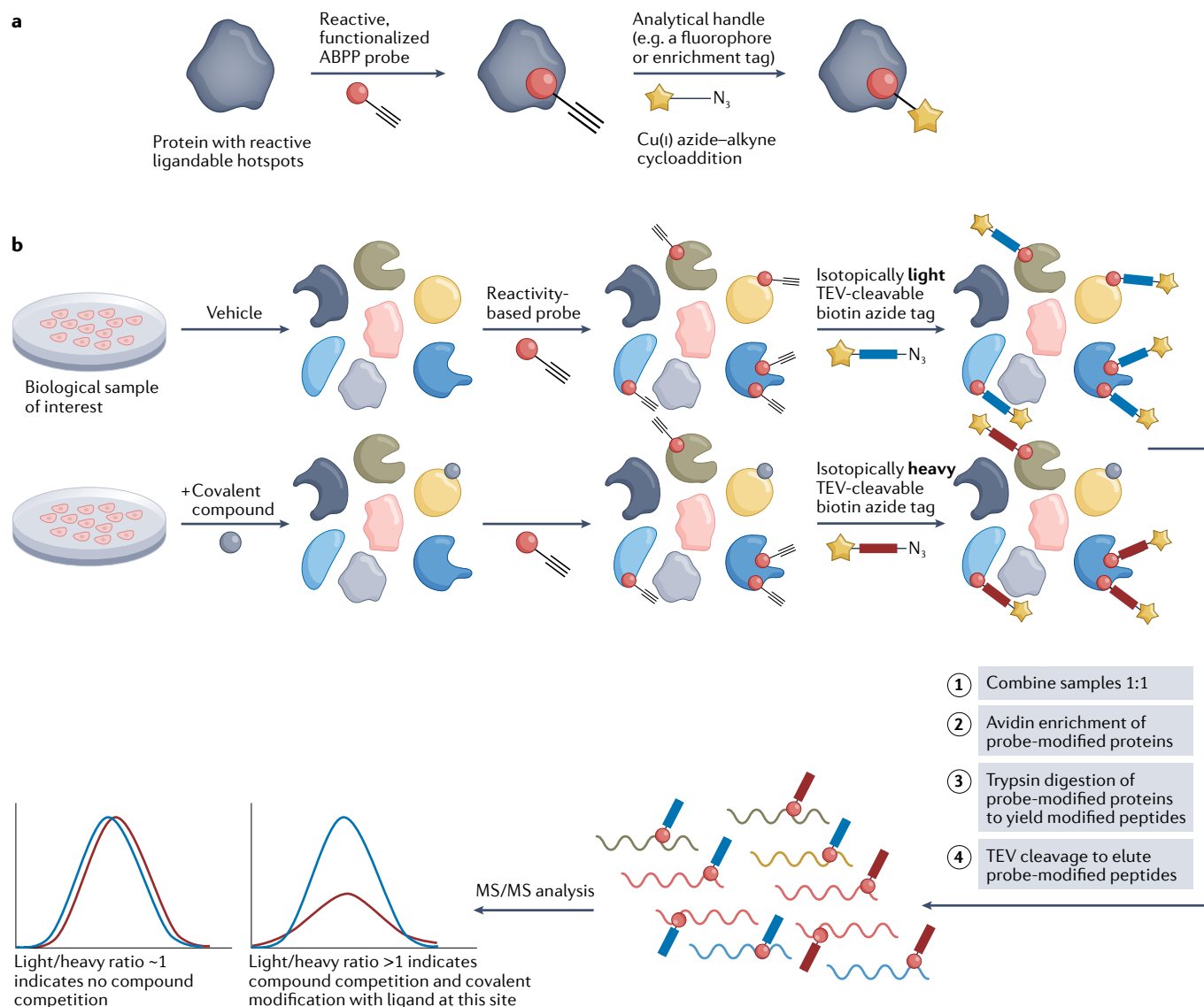
waveform ion mobility spectrometry (FAIMS) allow for additional separation before MS detection, and enabled the identification of more than 30,000 reactive cysteines across a panel of tumour cell lines<sup>178</sup>. Additionally, as new probes are developed for other nucleophilic amino acids, reactivity profiling of other amino acid hotspots (such as lysines) across the proteome will allow for the expansion of this technology beyond cysteine.

Overall, reactivity profiling generates large quantities of information across thousands of proteins, ultimately providing a relatively unbiased picture of nucleophilic (usually cysteine) amino acid reactivity. This information can be used to either select for appropriate protein targets in drug discovery programmes or to identify allosteric sites on proteins of interest that may have previously been considered un-ligandable or undruggable. As an example of targeting a traditionally undruggable protein with a covalent molecule, ABPP was paired with a MYC transcription factor activity assay to identify a covalent MYC ligand, EN4 (REF. 179). EN4 targets Cys171, which is located within a predicted intrinsically disordered region of MYC, and showed selectivity on a proteome-wide scale in profiling of more than 1,500 cysteines using competitive isoTOP-ABPP. Cys171 was initially identified as a ligandable hotspot on MYC through analysis of compiled cysteine-reactive chemoproteomics data, and this information spurred the subsequent search for a selective covalent ligand against that cysteine.

#### **Activity-based protein profiling screening platforms.**

The use of competitive isoTOP-ABPP was expanded to identify proteome-wide targets of a small covalent fragment library by competing individual acrylamide and chloroacetamide fragments against an iodoacetamide–alkyne probe<sup>180</sup>. In this study, more than 700 ligandable cysteines were identified, and information was provided about the proteome-wide selectivity of each covalent fragment in the library. The covalent ligands discovered with this approach and their corresponding ligandable sites were used to help elucidate the role of caspase 8 and caspase 10 in extrinsic apoptosis in T cells, showing that this approach can rapidly identify compounds that target proteins of biological interest. Several studies have built on this work, using isoTOP-ABPP to find new covalent ligands. For example, cysteine reactivity and ligandable sites were mapped in mutant Kelch-like ECH-associated protein 1 (KEAP1) and compared with those in wild-type KEAP1 NSCLC lines. The authors discovered compounds that bind to a ligandable cysteine on the nuclear receptor NR0B1, which is regulated by NRF2, a substrate of KEAP1 (REF. 181). Cysteine ligandability has also been explored in activated T cells through the use of promiscuous fragment-like compounds, termed ‘scout fragments’, to map ligandability and functional assays to identify more structurally complex electrophilic compounds that suppress T cell activity<sup>182</sup>. With exciting implications for drug discovery, this approach was also used to identify several proteins that could be targeted covalently to impair T cell activity, including BIRC2 and BIRC3, the nucleosome remodelling deacetylase (NuRD) complex, and the kinases ITK and CYTIP.





**Fig. 5 | Isotopic tandem orthogonal proteolysis–activity-based protein profiling.** **a** | Example of a reactive probe for activity-based protein profiling (ABPP) designed with a broadly reactive electrophilic warhead linked to an analytical handle. **b** | Schematic of the competitive isoTOP-ABPP (isotopic tandem orthogonal proteolysis–activity-based protein profiling) methodology. Treatment of cells or lysate with a protein-reactive compound prevents subsequent binding of the pan-reactive probe, and this competitive ligand binding can be detected by tandem mass spectrometry (MS/MS) after an enrichment step, to indirectly identify covalent protein targets for the compound of interest. TEV, tobacco etch virus.

To dramatically increase sample throughput, a tandem mass tag (TMT)-based streamlined cysteine activity-based protein profiling (SLC-ABPP) methodology was designed and used to profile an electrophilic fragment library at an impressive depth of more than 8,000 reactive cysteine sites with a total instrument time of 18 min per compound<sup>183</sup>. As competitive isoTOP-ABPP becomes more high throughput, comprehensive selectivity and reactivity information will rapidly become available for diverse covalent reactive libraries within a wide context of biological disease states. This information will enable rapid identification of covalent ligands against reactive hotspots, along with providing selectivity information on each ligand. When paired with a parallel phenotypic screen against a desired outcome (such as cancer cell death), this methodology facilitates the identification of functional

covalent ligands and their corresponding protein targets in a high-throughput manner.

**Chemoproteomics platforms for target identification.** IsoTOP-ABPP can also be used to identify the protein targets of known electrophilic drugs. As an example, this approach was applied to dimethyl fumarate, which is used to treat autoimmune disease. Although dimethyl fumarate has been used for three decades to treat psoriasis and was approved by the FDA in 2013 for the treatment of multiple sclerosis, the direct covalent targets of dimethyl fumarate remained unclear until more recently. In separate studies, chemoproteomic approaches were used to identify protein kinase C $\theta$  (PKC $\theta$ ) and IRAK4 as targets of dimethyl fumarate<sup>184,185</sup>. In both cases, covalent engagement of a cysteine residue disrupted a

protein–protein interaction to modulate immune cell function. Disrupting the interaction of PKC $\theta$  with the costimulatory receptor CD28 reduced T cell activation, and disrupting the IRAK4–MYD88 interaction suppressed the production of interferon- $\alpha$  in plasmacytoid dendritic cells<sup>184,185</sup>.

ABPP can also be used to identify off-targets and generally assess the selectivity of covalent molecules. For example, SLC-ABPP was used to analyse spleen tissue extracted from C57BL/6 mice treated with ibrutinib<sup>183</sup>. Of ~9,200 cysteine sites identified, BTK Cys481 was one of the cysteines most liganded by ibrutinib. Cys313 on B lymphocyte kinase (BLK), which, analogous to BTK, contains a cysteine within the ATP-binding pocket, was also identified as an off-target of ibrutinib<sup>183</sup>.

Novel screening platforms are also often easily paired with isoTOP-ABPP target identification experiments. For example, a multiplexed *in vivo* screening platform was developed in which barcoded pancreatic ductal adenocarcinoma lines were pretreated with electrophilic compounds and injected into mice to observe the compound-dependent decrease in metastatic potential<sup>186</sup>. IsoTOP-ABPP experiments also enabled the identification of the lipase ABHD6 as the target of hit compounds from this screen, even though ABHD6 was not previously known to have a role in metastasis or cancer progression. Beyond identifying the lipase ABHD6 as crucial for metastatic fitness, this approach enabled screening in a biological context more relevant to the disease state through adaptation of covalent ligand screening to a multiplexed *in vivo* phenotypic assay. In general, target identification experiments using chemoproteomic platforms are crucial in investigating the mechanism of action of electrophilic compounds discovered through phenotypic assays.

**Covalent ligand discovery for induced proximity modalities.** Covalent ligands are not only useful as functional inhibitors, but also have important roles in emerging induced proximity modalities. Covalent drug discovery platforms have facilitated the expansion of targeted protein degradation approaches by enabling the discovery of covalent recruiters against E3 ubiquitin ligases<sup>14,187–193</sup>. Although most bifunctional degrader molecules (proteolysis-targeting chimeras (PROTACs)) recruit the E3 ligases cereblon (CRBN) or von Hippel–Lindau (VHL) protein to degrade target proteins, there are more than 600 E3 ligases with varying substrate scopes. Since 2019, covalent recruiters have been used to validate a large proportion of the E3 ligases that have been harnessed for targeted protein degradation, including RING finger protein 114 (RNF114), RNF4, DDB1 and CUL4-associated factor 16 (DCAF16), DCAF11, KEAP1 and, most recently, fem-1 homologue B (FEM1B)<sup>187–193</sup>. In 2019, isoTOP-ABPP was used to identify RNF114 as the target of the enone-containing natural product nimbolide, which was used to make bifunctional degraders of bromodomain-containing protein 4 (BRD4) and BCR–ABL<sup>187</sup>. In a separate study, scout fragments were used to construct bifunctional FKBP12 and BRD4 degraders, and the authors identified DCAF16 as the target E3 ligase responsible for degradation<sup>190</sup>. These discoveries led to the variety of covalent E3 recruiters now available,

which have been reviewed elsewhere<sup>194</sup>. On the basis of analyses of chemoproteomic data sets assessing cysteine reactivity, 97% of E3 ligases possess reactive cysteines, suggesting that covalent approaches to harness more E3 ligases could continue to be successful<sup>194</sup>.

Beyond degradation, the identification of non-inhibitory covalent ligands has the potential to contribute to the discovery of novel induced proximity modalities. For example, a targeted protein stabilization platform, termed deubiquitinase-targeting chimeras (DUBTACs) has been developed using a covalent deubiquitinase recruiter<sup>195</sup>. Through analysis of chemoproteomic data and an ABPP-based screen, a covalent OTUB1 recruiter was discovered that could be incorporated into bifunctional compounds that stabilize mutant cystic fibrosis transmembrane conductance regulator (CFTR), the degradation of which drives cystic fibrosis. In general, the identification of non-inhibitory, allosteric ligands through covalent ligand screening has the potential to facilitate recruitment of other enzymes (such as kinases and deacetylases) to target proteins and direct protein function to neosubstrates for therapeutic benefit.

### Nucleophilic covalent ligands

Most protein-reactive covalent drugs tend to be electrophilic, to enable reactions with nucleophilic amino acids. By contrast, nucleophilic drugs can react with electrophilic cofactors and post-translational modifications. Several hydrazine-containing compounds act as mechanism-based inhibitors of monoamine oxidase (MAO) A and B, whereby activation by MAO enables alkylation of a flavin cofactor, inhibiting the enzyme<sup>196</sup>. ABPP principles and ‘reverse-polarity’ probes have been employed to study electrophilic post-translational modifications, such as N-terminal pyruvoyl and glyoxylyl modifications, that can react with hydrazine-containing compounds<sup>197</sup>. Hydrazine-based probes have also been used to help identify and characterize ligands for proteins with electrophilic cofactors or post-translational modifications<sup>198,199</sup>.

### Lysine-directed covalent ligands

The low abundance of cysteine enables selectivity but limits opportunities for covalently targeting specific proteins of interest. This problem has driven scientists to investigate the targeting of other nucleophilic amino acids, particularly lysine. Owing to the low nucleophilicity of the  $\epsilon$ -amino group of lysine under physiological conditions, the discovery of efficient lysine-targeting covalent ligands requires identification of unusually reactive lysines. IsoTOP-ABPP experiments using several lysine-directed probes have proved powerful in profiling lysine reactivity across the proteome<sup>200,201</sup>. Through isoTOP-ABPP experiments, more than 9,000 ligandable lysines were identified and more elaborated pentafluorophenol-containing or *N*-hydroxysuccinimide-ester-containing compounds could selectively label specific proteins of interest<sup>200,201</sup>. Building on these experiments, a library of approximately 180 electrophiles was assembled and isoTOP-ABPP experiments then performed to assess the selectivity of different chemotypes and identify lysines ligandable with small molecules<sup>202</sup>. This study yielded more broadly

reactive electrophiles, such as dicarboxaldehydes, that could be used for further lysine profiling experiments, but also identified less-reactive electrophiles, including *N*-acyl-*N*-alkyl sulfonamides, which had been previously used as tools for bioconjugation in cells<sup>202,203</sup>.

Aside from voxelotor, which was discovered through optimization from fragment-like aldehydes, most lysine-targeting ligands have been designed using structure-based methods from an existing ligand, often through rationally placing an electrophilic sulfonyl fluoride, fluorosulfate or vinyl sulfone in an appropriate orientation to react with an  $\epsilon$ -amino group of a lysine adjacent to an established binding site. Such ligands include a kinetic transthyretin stabilizer<sup>204</sup>, an isoform-selective PI3K $\delta$  inhibitor<sup>205</sup> and inhibitors of cyclin-dependent kinase 2 (CDK2)<sup>206</sup> and Hsp90 (REF.<sup>207</sup>). A sulfonyl fluoride-bearing promiscuous kinase inhibitor that targets a conserved lysine in the ATP-binding site was also used as a probe to profile kinase inhibitor selectivity in live cells<sup>208</sup>. However, sulfonyl fluoride-based probes are not completely selective for lysine and have also been used to target tyrosine residues<sup>209</sup>. A recent preprint reported the use of a chemoproteomic approach to profile the amino acid reactivity preference of 54 different electrophiles, which will prove to be a great resource for covalent ligand discovery<sup>210</sup>. Combining comprehensive electrophile profiling, lysine-directed chemoproteomics and structure-guided approaches will enable scientists to leverage the abundance of lysine residues adjacent to ligand binding sites to enhance covalent drug discovery.

## Outlook

Over the past decade, advances in covalent drug discovery have led to successful drugs, including inhibitors of EGFR, BTK, KRAS(G12C) and SARS-CoV-2 M<sup>pro</sup>. The approvals of these drugs represent milestones that

showcase the evolution of covalent drug discovery from a serendipitous effort to a field with established roadmaps for success.

Adoption of electrophile-first discovery strategies represents a notable shift in the field. Ligand-first strategies will continue to be highly applicable for designing covalent drugs against proteins when existing reversible ligands are already known to bind near a nucleophilic amino acid such as cysteine. However, we anticipate that electrophile-first approaches will be increasingly employed, especially when the discovery of reversible ligands proves challenging. Electrophile-first approaches will be facilitated, in part, by chemoproteomics experiments that profile amino acid reactivity across the proteome, leading to the identification of novel ligandable cysteines, for example, that could be targeted with electrophilic compounds.

Additionally, we expect to see an increase in research that explores the use of reversible covalent mechanisms that strike a balance between potency and selectivity in various contexts. Reversible covalent compounds with long off-rates might achieve a desired therapeutic effect while minimizing covalent off-target effects, and the use of more varied electrophiles will allow for increasingly tailored reactivity. We also look forward to further improvements in chemoproteomic workflows that enable additional multiplexing in MS experiments, which will be valuable for assessing compound selectivity and target engagement.

Covalent drug discovery overcomes obstacles in designing ligands against otherwise undruggable protein targets. We expect that the unique features of covalent ligands will continue to spur the discovery of covalent drugs.

Published online: 25 August 2022

- Tuley, A. & Fast, W. The taxonomy of covalent inhibitors. *Biochemistry* **57**, 3326–3337 (2018). **This review summarizes the mechanisms of covalent inhibition of enzymes and comprehensively covers examples for each mechanism.**
- Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011). **This review illustrates the key features of covalent drug discovery with emphasis on targeted covalent inhibitors.**
- Ghosh, A. K., Samanta, I., Mondal, A. & Liu, W. R. Covalent inhibition in drug discovery. *ChemMedChem* **14**, 889–906 (2019).
- De Cesco, S., Kurian, J., Dufresne, C., Mittermaier, A. K. & Moitessier, N. Covalent inhibitors design and discovery. *Eur. J. Med. Chem.* **138**, 96–114 (2017).
- Sutanto, F., Konstantinidou, M. & Dömling, A. Covalent inhibitors: a rational approach to drug discovery. *RSC Med. Chem.* **11**, 876–884 (2020).
- Zhang, T., Hatcher, J. M., Teng, M., Gray, N. S. & Kostic, M. Recent advances in selective and irreversible covalent ligand development and validation. *Cell Chem. Biol.* **26**, 1486–1500 (2019).
- Lonsdale, R. & Ward, R. A. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* **47**, 3816–3830 (2018).
- Sagonowsky, E. The top 20 drugs by worldwide sales in 2020. *Fierce Pharma* <https://www.fiercepharma.com/special-report/top-20-drugs-by-2020-sales> (2021).
- Moore, A. R., Rosenberg, S. C., McCormick, F. & Malek, S. RAS-targeted therapies: is the undruggable drugged? *Nat. Rev. Drug Discov.* **19**, 533–552 (2020).
- Huang, L., Guo, Z., Wang, F. & Fu, L. KRAS mutation: from undruggable to druggable in cancer. *Signal Transduct. Target. Ther.* **6**, 386 (2021).
- Vandyck, K. & Deval, J. Considerations for the discovery and development of 5-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Curr. Opin. Virol.* **49**, 36–40 (2021).
- Lagoutte, R., Patouret, R. & Winssinger, N. Covalent inhibitors: an opportunity for rational target selectivity. *Curr. Opin. Chem. Biol.* **39**, 54–63 (2017).
- Lu, W. et al. Fragment-based covalent ligand discovery. *RSC Chem. Biol.* **2**, 354–367 (2021).
- Spradlin, J. N., Zhang, E. & Nomura, D. K. Reimagining druggability using chemoproteomic platforms. *Acc. Chem. Res.* **54**, 1801–1813 (2021).
- Roberts, A. M., Ward, C. C. & Nomura, D. K. Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots. *Curr. Opin. Biotechnol.* **43**, 25–33 (2017).
- Drewes, G. & Knapp, S. Chemoproteomics and chemical probes for target discovery. *Trends Biotechnol.* **36**, 1275–1286 (2018).
- Moellering, R. E. & Cravatt, B. F. How chemoproteomics can enable drug discovery and development. *Chem. Biol.* **19**, 11–22 (2012).
- Maurais, A. J. & Weerapana, E. Reactive-cysteine profiling for drug discovery. *Curr. Opin. Chem. Biol.* **50**, 29–36 (2019).
- Chandrasekharan, N. & Simmons, D. L. The cyclooxygenases. *Genome Biol.* **5**, 241 (2004).
- Vane, J. R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nature* **231**, 232–235 (1971).
- Nicola, G., Tomberg, J., Pratt, R. F., Nicholas, R. A. & Davies, C. Crystal structures of covalent complexes of  $\beta$ -lactam antibiotics with *Escherichia coli* penicillin-binding protein 5: toward an understanding of antibiotic specificity. *Biochemistry* **49**, 8094–8104 (2010).
- Dijkman, A. C. et al. Fosfomycin: pharmacological, clinical and future perspectives. *Antibiotics* **6**, 24 (2017).
- Hendlin, D. et al. Phosphonomycin, a new antibiotic produced by strains of *Streptomyces*. *Science* **166**, 122–123 (1969).
- Kahan, F. M., Kahan, J. S., Cassidy, P. J. & Kropp, H. The mechanism of action of fosfomycin (phosphonomycin). *Ann. NY Acad. Sci.* **235**, 364–386 (1974).
- Olbe, L., Carlsson, E. & Lindberg, P. A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole. *Nat. Rev. Drug Discov.* **2**, 132–139 (2003).
- Savi, P. et al. Identification and biological activity of the active metabolite of clopidogrel. *Thromb. Haemost.* **84**, 891–896 (2000).
- Thomas, D. & Zalcberg, J. 5-Fluorouracil: a pharmacological paradigm in the use of cytotoxics. *Clin. Exp. Pharmacol. Physiol.* **25**, 887–895 (1998).
- Danenberg, P. V., Langenbach, R. J. & Heidelberger, C. Fluorinated pyrimidines. Structures of reversible and irreversible complexes of thymidylate synthetase and fluorinated pyrimidine nucleotides. *Biochemistry* **13**, 926–933 (1974).
- Xu, H., Faber, C., Uchiki, T., Racca, J. & Dealwis, C. Structures of eukaryotic ribonucleotide reductase I define gemcitabine diphosphate binding and subunit assembly. *Proc. Natl Acad. Sci. USA* **103**, 4028–4033 (2006).
- Curran, M. P. & McKeage, K. Bortezomib: a review of its use in patients with multiple myeloma. *Drugs* **69**, 859–888 (2009).
- Lynch, T. J., Okimoto, R. A., Supko, J. G. & Settleman, J. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).

32. Paez, J. G. et al. *EGFR* mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
33. Ou, S.-H. I. Second-generation irreversible epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs): a better mousetrap? A review of the clinical evidence. *Crit. Rev. Oncol. Hematol.* **83**, 407–421 (2012).
34. Yu, H. A. et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with *EGFR*-mutant lung cancers. *Clin. Cancer Res.* **19**, 2240–2247 (2013).
35. Recondo, G., Facchinetti, F., Olaussen, K. A., Besse, B. & Friboulet, L. Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI? *Nat. Rev. Clin. Oncol.* **15**, 694–708 (2018).
36. Yun, C.-H. et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl Acad. Sci. USA* **105**, 2070–2075 (2008).
37. Soria, J.-C. et al. Afatinib versus erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-Lung 8): an open-label randomised controlled phase 3 trial. *Lancet Oncol.* **16**, 897–907 (2015).
38. Yu, H. A. & Pao, W. Afatinib — new therapy option for EGFR-mutant lung cancer. *Nat. Rev. Clin. Oncol.* **10**, 551–552 (2013).
39. Harvey, R. D., Adams, V. R., Beardslee, T. & Medina, P. Afatinib for the treatment of *EGFR* mutation-positive NSCLC: a review of clinical findings. *J. Oncol. Pharm. Pract.* **26**, 1461–1474 (2020).
40. Tsou, H.-R. et al. Optimization of 6,7-disubstituted-4-(arylamino)quinoline-3-carbonitriles as orally active, irreversible inhibitors of human epidermal growth factor receptor-2 kinase activity. *J. Med. Chem.* **48**, 1107–1131 (2005).
41. Deeks, E. D. Neratinib: first global approval. *Drugs* **77**, 1695–1704 (2017).
42. Shirley, M. Dacomitinib: first global approval. *Drugs* **78**, 1947–1953 (2018).
43. Zhou, W. et al. Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070–1074 (2009).  
**This paper reports the discovery of covalent EGFR inhibitors selective for the T790M gatekeeper mutation.**
44. Cross, D. A. E. et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov.* **4**, 1046–1061 (2014).
45. Finlay, M. R. V. et al. Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor. *J. Med. Chem.* **57**, 8249–8267 (2014).
46. Walter, A. O. et al. Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. *Cancer Discov.* **3**, 1404–1415 (2013).
47. Butterworth, S., Cross, D. A. E., Finlay, M. R. V., Ward, R. A. & Waring, M. J. The structure-guided discovery of osimertinib: the first U.S. FDA approved mutant selective inhibitor of EGFR T790M. *Med. Chem. Commun.* **8**, 820–822 (2017).
48. Jänne, P. A. et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 1689–1699 (2015).
49. Mok, T. S. et al. Osimertinib or platinum–pemetrexed in EGFR T790M-positive lung cancer. *N. Engl. J. Med.* **376**, 629–640 (2016).
50. Passaro, A., Jänne, P. A., Mok, T. & Peters, S. Overcoming therapy resistance in EGFR-mutant lung cancer. *Nat. Cancer* **2**, 377–391 (2021).
51. Leonetti, A. et al. Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *Br. J. Cancer* **121**, 725–737 (2019).
52. Thress, K. S. et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat. Med.* **21**, 560–562 (2015).
53. Pal Singh, S., Dammeijer, F. & Hendriks, R. W. Role of Bruton's tyrosine kinase in B cells and malignancies. *Mol. Cancer* **17**, 57 (2018).
54. Satterthwaite, A. B. & Witte, O. N. The role of Bruton's tyrosine kinase in B-cell development and function: a genetic perspective. *Immunol. Rev.* **175**, 120–127 (2009).
55. Pan, Z. et al. Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. *ChemMedChem* **2**, 58–61 (2007).  
**This work describes the discovery of ibrutinib.**
56. Honigberg, L. A. et al. The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. *Proc. Natl Acad. Sci. USA* **107**, 13075–13080 (2010).
57. Herman, S. E. M. et al. Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood* **117**, 6287–6296 (2011).
58. Burger, J. A. & Buggy, J. J. Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765). *Leuk. Lymphoma* **54**, 2385–2391 (2013).
59. Wang, M. L. et al. Targeting BTK with Ibrutinib in relapsed or refractory mantle-cell lymphoma. *N. Engl. J. Med.* **369**, 507–516 (2013).
60. Byrd, J. C. et al. Targeting BTK with Ibrutinib in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **369**, 32–42 (2013).
61. Byrd, J. C. et al. Ibrutinib treatment for first-line and relapsed/refractory chronic lymphocytic leukemia: final analysis of the pivotal phase Ib/II PCYC-1102 study. *Clin. Cancer Res.* **26**, 3918–3927 (2020).
62. Treon, S. P. et al. Ibrutinib in previously treated Waldenström's macroglobulinemia. *N. Engl. J. Med.* **372**, 1430–1440 (2015).
63. Miklos, D. et al. Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood* **130**, 2243–2250 (2017).
64. Davids, M. S. & Brown, J. R. Ibrutinib: a first in class covalent inhibitor of Bruton's tyrosine kinase. *Future Oncol.* **10**, 957–967 (2014).
65. Advani, R. H. et al. Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765) has significant activity in patients with relapsed/refractory B-cell malignancies. *J. Clin. Oncol.* **31**, 88–94 (2013).
66. Narita, Y. et al. Phase I/II study of tirabrutinib, a second-generation Bruton's tyrosine kinase inhibitor, in relapsed/refractory primary central nervous system lymphoma. *Neuro-Oncol.* **23**, 122–133 (2021).
67. Guo, Y. et al. Discovery of zanubrutinib (GBG-3111), a novel, potent, and selective covalent inhibitor of Bruton's tyrosine kinase. *J. Med. Chem.* **62**, 7923–7940 (2019).
68. Caldwell, R. D. et al. Discovery of evobrutinib: an oral, potent, and highly selective, covalent Bruton's tyrosine kinase (BTK) inhibitor for the treatment of immunological diseases. *J. Med. Chem.* **62**, 7643–7655 (2019).
69. Barf, T. et al. Acalabrutinib (ACP-196): a covalent Bruton tyrosine kinase inhibitor with a differentiated selectivity and in vivo potency profile. *J. Pharmacol. Exp. Ther.* **363**, 240–252 (2017).
70. Byrd, J. C. et al. Acalabrutinib (ACP-196) in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **374**, 323–332 (2016).
71. Byrd, J. C. et al. Acalabrutinib versus ibrutinib in previously treated chronic lymphocytic leukemia: results of the first randomized phase III trial. *J. Clin. Oncol.* **39**, 3441–3452 (2021).
72. Serafimova, I. M. et al. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat. Chem. Biol.* **8**, 471–476 (2012).
73. Miller, R. M., Paavilainen, V. O., Krishnan, S., Serafimova, I. M. & Taunton, J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J. Am. Chem. Soc.* **135**, 5298–5301 (2013).
74. Bradshaw, J. M. et al. Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat. Chem. Biol.* **11**, 525–531 (2015).
75. Zhao, Z. & Bourne, P. E. Progress with covalent small-molecule kinase inhibitors. *Drug Discov. Today* **23**, 727–735 (2018).
76. Liu, Q. et al. Developing irreversible inhibitors of the protein kinase cysteinome. *Chem. Biol.* **20**, 146–159 (2013).
77. Thorarensen, A. et al. Design of a Janus kinase 3 (JAK3) specific inhibitor 1-((2S,5R)-5-((7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)-2-methylpiperidin-1-yl)prop-2-en-1-one (PF-06651600) allowing for the interrogation of JAK3 signaling in humans. *J. Med. Chem.* **60**, 1971–1993 (2017).
78. Goedken, E. R. et al. Tricyclic covalent inhibitors selectively target Jak3 through an active site thiol. *J. Biol. Chem.* **290**, 4573–4589 (2015).
79. Tan, L. et al. Development of selective covalent Janus kinase 3 inhibitors. *J. Med. Chem.* **58**, 6589–6606 (2015).
80. Kempson, J. et al. Discovery of highly potent, selective, covalent inhibitors of JAK3. *Bioorg. Med. Chem. Lett.* **27**, 4622–4625 (2017).
81. Telliez, J.-B. et al. Discovery of a JAK3-selective inhibitor: functional differentiation of JAK3-selective inhibition over pan-JAK or JAK1-selective inhibition. *ACS Chem. Biol.* **11**, 3442–3451 (2016).
82. Robinson, M. F. et al. Efficacy and safety of PF-06651600 (ritlecitinib), a novel JAK3/TEC inhibitor, in patients with moderate-to-severe rheumatoid arthritis and an inadequate response to methotrexate. *Arthritis Rheumatol.* **72**, 1621–1631 (2020).
83. Kim, R. D. et al. First-in-human phase I study of fisogatinib (BLU-554) validates aberrant FGF19 signaling as a driver event in hepatocellular carcinoma. *Cancer Discov.* **9**, 1696–1707 (2019).
84. Tan, L. et al. Development of covalent inhibitors that can overcome resistance to first-generation FGFR kinase inhibitors. *Proc. Natl Acad. Sci. USA* **111**, E4869–E4877 (2014).
85. Weiss, A. et al. FGF401, a first-in-class highly selective and potent FGFR4 inhibitor for the treatment of FGF19-driven hepatocellular cancer. *Mol. Cancer Ther.* **18**, 2194–2206 (2019).
86. Goebel, L., Müller, M. P., Goody, R. S. & Rauh, D. KRASG12C inhibitors in clinical trials: a short historical perspective. *RSC Med. Chem.* **11**, 760–770 (2020).
87. Janes, M. R. et al. Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor. *Cell* **172**, 578–589.e17 (2018).
88. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).  
**This work showcases the use of tethering as a screening methodology for discovering covalent ligands against KRAS(G12C) and further identifies ligands against this oncogenic protein.**
89. Hofmann, M. H., Gerlach, D., Misale, S., Petronczki, M. & Kraut, N. Expanding the reach of precision oncology by drugging all KRAS mutants. *Cancer Discov.* **12**, 924–937 (2022).
90. Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
91. Lito, P., Solomon, M., Li, L.-S., Hansen, R. & Rosen, N. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science* **351**, 604–608 (2016).
92. Patricelli, M. P. et al. Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer Discov.* **6**, 316–329 (2016).
93. Canon, J. et al. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223 (2019).
94. Lanman, B. A. et al. Discovery of a covalent inhibitor of KRAS<sup>G12C</sup> (AMG 510) for the treatment of solid tumors. *J. Med. Chem.* **63**, 52–65 (2020).
95. Skoulidis, F. et al. Sotorasib for lung cancers with KRAS p.G12C mutation. *N. Engl. J. Med.* **384**, 2371–2381 (2021).
96. Fell, J. B. et al. Identification of the clinical development candidate MRTX849, a covalent KRAS<sup>G12C</sup> inhibitor for the treatment of cancer. *J. Med. Chem.* **63**, 6679–6693 (2020).
97. Hallin, J. et al. The KRAS<sup>G12C</sup> inhibitor MRTX849 provides insight toward therapeutic susceptibility of KRAS-mutant cancers in mouse models and patients. *Cancer Discov.* **10**, 54–71 (2020).
98. Wang, J. et al. Phase I study of JNJ-74699157 in patients with advanced solid tumors harboring the KRAS G12C mutation. *Oncologist* **27**, 536–e553 (2022).
99. US FDA. *Coronavirus (COVID-19) Update: FDA Authorizes First Oral Antiviral for Treatment of COVID-19*. FDA <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-oral-antiviral-treatment-covid-19> (2021).
100. Qiao, J. et al. SARS-CoV-2 M<sup>pro</sup> inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374–1378 (2021).
101. Anand, K., Ziebuhr, J., Wadhvani, P., Mesters, J. R. & Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763–1767 (2003).
102. Hoffman, R. L. et al. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J. Med. Chem.* **63**, 12725–12747 (2020).
103. Owen, D. R. et al. An oral SARS-CoV-2 M<sup>pro</sup> inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586–1593 (2021).  
**This work describes the discovery of PF-07321332, later named nirmatrelvir, an orally bioavailable SARS-CoV-2 M<sup>pro</sup> inhibitor.**



104. Bandyopadhyay, A. & Gao, J. Targeting biomolecules with reversible covalent chemistry. *Curr. Opin. Chem. Biol.* **34**, 110–116 (2016).
105. Hammond, J. et al. Oral nirmatrelvir for high-risk, nonhospitalized adults with Covid-19. *N. Engl. J. Med.* **386**, 1397–1408 (2022).
106. Mahase, E. Covid-19: Pfizer's paxlovid is 89% effective in patients at risk of serious illness, company reports. *BMJ* **375**, n2713 (2021).
107. de Leuw, P. & Stephan, C. Protease inhibitor therapy for hepatitis C virus infection. *Expert. Opin. Pharmacother.* **19**, 577–587 (2018).
108. Hézode, C. et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N. Engl. J. Med.* **360**, 1839–1850 (2009).
109. Venkatraman, S. et al. Discovery of (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(S)-carboxamide (SCH 503034), a selective, potent, orally bioavailable hepatitis C virus NS3 protease inhibitor: a potential therapeutic agent for the treatment of hepatitis C infection. *J. Med. Chem.* **49**, 6074–6086 (2006).
110. Venkatraman, S. Discovery of boceprevir, a direct-acting NS3/4A protease inhibitor for treatment of chronic hepatitis C infections. *Trends Pharmacol. Sci.* **33**, 289–294 (2012).
111. Lin, C., Kwong, A. & Perni, R. Discovery and development of VX-950, a novel, covalent, and reversible inhibitor of hepatitis C virus NS3/4A serine protease. *Infect. Disord. Drug Targets* **6**, 3–16 (2006).
112. Kwong, A. D., Kauffman, R. S., Hurter, P. & Mueller, P. Discovery and development of telaprevir: an NS3/4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nat. Biotechnol.* **29**, 993–1003 (2011).
113. Arasappan, A. et al. Discovery of narlaprevir (SCH 900518): a potent, second generation HCV NS3 serine protease inhibitor. *ACS Med. Chem. Lett.* **1**, 64–69 (2010).
114. McCauley, J. A. & Rudd, M. T. Hepatitis C virus NS3/4A protease inhibitors. *Curr. Opin. Pharmacol.* **30**, 84–92 (2016).
115. Kwo, P. Y. et al. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alpha-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *Lancet* **376**, 705–716 (2010).
116. Sherman, K. E. et al. Response-guided telaprevir combination treatment for hepatitis C virus infection. *N. Engl. J. Med.* **365**, 1014–1024 (2011).
117. Ramsay, I. D., Lestner, J. M. & Barker, C. I. S. in *Side Effects of Drugs Annual* (ed. Aronson, J. K.) vol. 35 503–550 (Elsevier, 2014).
118. Merck & Co., Inc. Merck Voluntarily Discontinuing VICTRELIS. <https://www.fda.gov/files/drugs/published/Dear-Pharmacy-Professional-Letter.pdf> (2015).
119. Lawitz, E. et al. Sofosbuvir and ledipasvir fixed-dose combination with and without ribavirin in treatment-naive and previously treated patients with genotype 1 hepatitis C virus infection (LONESTAR): an open-label, randomised, phase 2 trial. *Lancet* **383**, 515–523 (2014).
120. Bross, P. F. et al. Approval summary for bortezomib for injection in the treatment of multiple myeloma. *Clin. Cancer Res.* **10**, 3954–3964 (2004).
121. Adams, J. et al. Potent and selective inhibitors of the proteasome: dipeptidyl boronic acids. *Bioorg. Med. Chem. Lett.* **8**, 333–338 (1998).
122. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071–3076 (2001).
123. Richardson, P. G., Hideshima, T. & Anderson, K. C. Bortezomib (PS-341): a novel, first-in-class proteasome inhibitor for the treatment of multiple myeloma and other cancers. *Cancer Control.* **10**, 361–369 (2003).
124. Arkwright, R., Pham, T. M., Zonder, J. A. & Dou, Q. P. The preclinical discovery and development of bortezomib for the treatment of mantle cell lymphoma. *Expert. Opin. Drug Discov.* **12**, 225–235 (2017).
125. Cromm, P. M. & Crews, C. M. The proteasome in modern drug discovery: second life of a highly valuable drug target. *ACS Cent. Sci.* **3**, 830–838 (2017).
126. Kim, K. B. & Crews, C. M. From epoxomicin to carfilzomib: chemistry, biology, and medical outcomes. *Nat. Prod. Rep.* **30**, 600 (2013).
127. Groll, M., Kim, K. B., Kairies, N., Huber, R. & Crews, C. M. Crystal structure of epoxomicin:20S proteasome reveals a molecular basis for selectivity of  $\alpha^1$ ,  $\beta^1$ -epoxyketone proteasome inhibitors. *J. Am. Chem. Soc.* **122**, 1237–1238 (2000).
128. Kirk, C. J. Discovery and development of second-generation proteasome inhibitors. *Semin. Hematol.* **49**, 207–214 (2012).
129. Chauhan, D. et al. *In vitro* and *In vivo* selective antitumor activity of a novel orally bioavailable proteasome inhibitor MLN9708 against multiple myeloma cells. *Clin. Cancer Res.* **17**, 5311–5321 (2011).
130. Kupperman, E. et al. Evaluation of the proteasome inhibitor MLN9708 in preclinical models of human cancer. *Cancer Res.* **70**, 1970–1980 (2010).
131. Smolewski, P. & Rydygier, D. Ixazomib: an investigational drug for the treatment of lymphoproliferative disorders. *Expert. Opin. Invest. Drugs* **28**, 421–433 (2019).
132. Zhou, H.-J. et al. Design and synthesis of an orally bioavailable and selective peptide epoxyketone proteasome inhibitor (PR-047). *J. Med. Chem.* **52**, 3028–3038 (2009).
133. Diaz, D. B. & Yudin, A. K. The versatility of boron in biological target engagement. *Nat. Chem.* **9**, 731–742 (2017).
134. Plescia, J. & Moitessier, N. Design and discovery of boronic acid drugs. *Eur. J. Med. Chem.* **195**, 112270 (2020).
135. Adamczyk-Woźniak, A., Borys, K. M. & Sporzynski, A. Recent developments in the chemistry and biological applications of benzoxaboroles. *Chem. Rev.* **115**, 5224–5247 (2015).
136. Baker, S. J. et al. Discovery of a new boron-containing antifungal agent, 5-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborole (AN2690), for the potential treatment of onychomycosis. *J. Med. Chem.* **49**, 4447–4450 (2006).
137. Benkovic, S. J. et al. Identification of borinic esters as inhibitors of bacterial cell growth and bacterial methyltransferases, CcrM and MenH. *J. Med. Chem.* **48**, 7468–7476 (2005).
138. Rock, F. L. et al. An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science* **316**, 1759–1761 (2007).
139. Akama, T. et al. Discovery and structure-activity study of a novel benzoxaborole anti-inflammatory agent (AN2728) for the potential topical treatment of psoriasis and atopic dermatitis. *Bioorg. Med. Chem. Lett.* **19**, 2129–2132 (2009).
140. Hecker, S. J. et al. Discovery of a cyclic boronic acid  $\beta$ -lactamase inhibitor (RPX7009) with utility vs class A serine carboxenemases. *J. Med. Chem.* **58**, 3682–3692 (2015).
141. Das, B. C. Boron-containing heterocycles as promising pharmacological agents. *Bioorg. Med. Chem.* **63**, 116748 (2022).
142. Vichinsky, E. et al. A phase 3 randomized trial of voxelotor in sickle cell disease. *N. Engl. J. Med.* **381**, 509–519 (2019).
143. Metcalf, B. et al. Discovery of GBT440, an orally bioavailable r-state stabilizer of sickle cell hemoglobin. *ACS Med. Chem. Lett.* **8**, 321–326 (2017).
144. Garner, M. H., Bogardt, R. A. & Gurd, F. R. Determination of the pK values for the alpha-amino groups of human hemoglobin. *J. Biol. Chem.* **250**, 4398–4404 (1975).
145. Oksenberg, D. et al. GBT 440 increases haemoglobin oxygen affinity, reduces sickling and prolongs RBC half-life in a murine model of sickle cell disease. *Br. J. Haematol.* **175**, 141–153 (2016).
146. Blair, H. A. Voxelotor: first approval. *Drugs* **80**, 209–215 (2020).
147. Erlanson, D. A. et al. Site-directed ligand discovery. *Proc. Natl Acad. Sci. USA* **97**, 9367–9372 (2000).
148. Arkin, M. R. et al. Binding of small molecules to an adaptive protein-protein interface. *Proc. Natl Acad. Sci. USA* **100**, 1603–1608 (2003).
149. Keedy, D. A. et al. An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**, e36307 (2018).
150. Wolter, M. et al. Fragment-based stabilizers of protein-protein interactions through imine-based tethering. *Angew. Chem. Int. Ed.* **59**, 21520–21524 (2020).
151. Cardoso, R. et al. Identification of Cys255 in HIF-1 $\alpha$  as a novel site for development of covalent inhibitors of HIF-1 $\alpha$ /ARNT PasB domain protein-protein interaction: covalent allosteric inhibitors of HIF-1 $\alpha$ /ARNT PasB PPI. *Prot. Sci.* **21**, 1885–1896 (2012).
152. Nonoo, R. H., Armstrong, A. & Mann, D. J. Kinetic template-guided tethering of fragments. *ChemMedChem* **7**, 2082–2086 (2012).
153. Kathman, S. G., Xu, Z. & Statsyuk, A. V. A fragment-based method to discover irreversible covalent inhibitors of cysteine proteases. *J. Med. Chem.* **57**, 4969–4974 (2014).
154. Kathman, S. G. et al. A small molecule that switches a ubiquitin ligase from a processive to a distributive enzymatic mechanism. *J. Am. Chem. Soc.* **137**, 12442–12445 (2015).
155. Johansson, H. et al. Fragment-based covalent ligand screening enables rapid discovery of inhibitors for the RBR E3 ubiquitin ligase HOIP. *J. Am. Chem. Soc.* **141**, 2703–2712 (2019).
156. Resnick, E. et al. Rapid covalent-probe discovery by electrophile-fragment screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).
- This paper highlights the use of intact MS screening methods paired with structure-guided design to identify covalent fragment ligands against protein targets.**
157. Dubiella, C. et al. Sulfopin is a covalent inhibitor of Pin1 that blocks Myc-driven tumors in vivo. *Nat. Chem. Biol.* **17**, 954–963 (2021).
158. Shin, Y. et al. Discovery of N-(1-acryloylazetid-3-yl)-2-(1H-indol-1-yl)acetamides as covalent inhibitors of KRAS<sup>G12C</sup>. *ACS Med. Chem. Lett.* **10**, 1302–1308 (2019).
- This study details the discovery of covalent KRAS(G12C) inhibitors through screening an acrylamide library, which laid the groundwork for the discovery of sotorasib.**
159. Debaene, F., Da Silva, J. A., Pianowski, Z., Duran, F. J. & Wissing, N. Expanding the scope of PNA-encoded libraries: divergent synthesis of libraries targeting cysteine, serine and metallo-proteases as well as tyrosine phosphatases. *Tetrahedron* **63**, 6577–6586 (2007).
160. Zimmermann, G. et al. A specific and covalent JNK-1 ligand selected from an encoded self-assembling chemical library. *Chem. Eur. J.* **23**, 8152–8155 (2017).
161. Zambaldo, C., Daguer, J.-P., Saarbach, J., Barluenga, S. & Wissing, N. Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *Med. Chem. Commun.* **7**, 1340–1351 (2016).
162. Daguer, J.-P. et al. Identification of covalent bromodomain binders through DNA display of small molecules. *Angew. Chem. Int. Ed.* **54**, 6057–6061 (2015).
163. Zhu, Z. et al. Development of a selection method for discovering irreversible (covalent) binders from a DNA-encoded library. *SLAS Discov.* **24**, 169–174 (2019).
164. Chan, A. I., McGregor, L. M., Jain, T. & Liu, D. R. Discovery of a covalent kinase inhibitor from a DNA-encoded small-molecule library: protein library selection. *J. Am. Chem. Soc.* **139**, 10192–10195 (2017).
165. Guilinger, J. P. et al. Novel irreversible covalent BTK inhibitors discovered using DNA-encoded chemistry. *Bioorg. Med. Chem.* **42**, 116223 (2021).
166. Sottriffer, C. Docking of covalent ligands: challenges and approaches. *Mol. Inf.* **37**, 1800062 (2018).
167. Scarpino, A., Ferenczy, G. G. & Keserü, G. M. Comparative evaluation of covalent docking tools. *J. Chem. Inf. Model.* **58**, 1441–1458 (2018).
- This article compares the reproducibility of known binding data for covalent complexes across AutoDock4, CovDock, FITTED, GOLD, ICM-Pro and MOE docking programs.**
168. Zhang, S. et al. Effective virtual screening strategy toward covalent ligands: identification of novel NEDD8-activating enzyme inhibitors. *J. Chem. Inf. Model.* **54**, 1785–1797 (2014).
169. London, N. et al. Covalent docking of large libraries for the discovery of chemical probes. *Nat. Chem. Biol.* **10**, 1066–1072 (2014).
170. Shraga, A. et al. Covalent docking identifies a potent and selective MKK7 inhibitor. *Cell Chem. Biol.* **26**, 98–108.e5 (2019).
171. Nnadi, C. I. et al. Novel K-Ras G12C switch-II covalent binders destabilize Ras and accelerate nucleotide exchange. *J. Chem. Inf. Model.* **58**, 464–471 (2018).
172. Zhang, S. et al. Covalent complexes of proteasome model with peptide aldehyde inhibitors MG132 and MG101: docking and molecular dynamics study. *J. Mol. Model.* **15**, 1481–1490 (2009).
173. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
- This review highlights the tools involved in ABPP and explains their corresponding applications.**

174. Edgington, L. E., Verdoes, M. & Bogoy, M. Functional imaging of proteases: recent advances in the design and application of substrate-based and activity-based probes. *Curr. Opin. Chem. Biol.* **15**, 798–805 (2011).
175. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc. Natl Acad. Sci. USA* **96**, 14694–14699 (1999).
176. Weerapana, E., Speers, A. E. & Cravatt, B. F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP) — a general method for mapping sites of probe modification in proteomes. *Nat. Protoc.* **2**, 1414–1425 (2007).
177. Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
178. Yan, T. et al. SP3-FAIMS chemoproteomics for high-coverage profiling of the human cysteineome. *ChemBioChem* **22**, 1841–1851 (2021).
179. Boike, L. et al. Discovery of a functional covalent ligand targeting an intrinsically disordered cysteine within MYC. *Cell Chem. Biol.* **28**, 4–13 (2020).
180. Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
- This work demonstrates how chemoproteomics can be paired with covalent ligand screening to identify the protein targets of covalent ligands in complex biological samples.**
181. Bar-Peled, L. et al. Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell* **171**, 696–709.e23 (2017).
182. Vinogradova, E. V. et al. An activity-guided map of electrophile-cysteine interactions in primary human T cells. *Cell* **182**, 1009–1026.e29 (2020).
183. Kuljanin, M. et al. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **39**, 630–641 (2021).
184. Blewett, M. M. et al. Chemical proteomic map of dimethyl fumarate-sensitive cysteines in primary human T cells. *Sci. Signal.* **9**, rs10 (2016).
185. Zaro, B. W. et al. Dimethyl fumarate disrupts human innate immune signaling by targeting the IRAK4–MyD88 complex. *J. Immunol.* **202**, 2737–2746 (2019).
186. Grüner, B. M. et al. An in vivo multiplexed small-molecule screening platform. *Nat. Methods* **13**, 883–889 (2016).
187. Spradlin, J. N. et al. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat. Chem. Biol.* **15**, 747–755 (2019).
188. Luo, M. et al. Chemoproteomics-enabled discovery of covalent RNF114-based degraders that mimic natural product function. *Cell Chem. Biol.* **28**, 559–566.e15 (2021).
189. Ward, C. C. et al. Covalent ligand screening uncovers a RNF4 E3 ligase recruiter for targeted protein degradation applications. *ACS Chem. Biol.* **14**, 2430–2440 (2019).
190. Zhang, X., Crowley, V. M., Wucherpfennig, T. C., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat. Chem. Biol.* **15**, 737–746 (2019).
191. Zhang, X. et al. DCAF11 supports targeted protein degradation by electrophilic proteolysis-targeting chimeras. *J. Am. Chem. Soc.* **143**, 5141–5149 (2021).
192. Tong, B. et al. Bardoxolone conjugation enables targeted protein degradation of BRD4. *Sci. Rep.* **10**, 15543 (2020).
193. Henning, N. J. et al. Discovery of a covalent FEM1B recruiter for targeted protein degradation applications. *J. Am. Chem. Soc.* **144**, 701–708 (2022).
194. Belcher, B. P., Ward, C. C. & Nomura, D. K. Ligandability of E3 ligases for targeted protein degradation applications. *Biochemistry* <https://doi.org/10.1021/acs.biochem.1c00464> (2021).
195. Henning, N. J. et al. Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat. Chem. Biol.* **18**, 412–421 (2022).
196. Binda, C. et al. Structural and mechanistic studies of arylalkylhydrazine inhibition of human monoamine oxidases A and B. *Biochemistry* **47**, 5616–5625 (2008).
197. Matthews, M. L. et al. Chemoproteomic profiling and discovery of protein electrophiles in human cells. *Nat. Chem.* **9**, 234–243 (2017).
198. Wang, X. et al. Discovery of potent and selective inhibitors against protein-derived electrophilic cofactors. *J. Am. Chem. Soc.* **144**, 5377–5388 (2022).
199. Lin, Z. et al. Activity-based hydrazine probes for protein profiling of electrophilic functionality in therapeutic targets. *ACS Cent. Sci.* **7**, 1524–1534 (2021).
200. Hacker, S. M. et al. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
201. Ward, C. C., Kleinman, J. I. & Nomura, D. K. NHS-esters as versatile reactivity-based probes for mapping proteome-wide ligandable hotspots. *ACS Chem. Biol.* **12**, 1478–1483 (2017).
202. Abbasov, M. E. et al. A proteome-wide atlas of lysine-reactive chemistry. *Nat. Chem.* **13**, 1081–1092 (2021).
203. Tamura, T. et al. Rap labeling and covalent inhibition of intracellular native proteins using ligand-directed *N*-acyl-*N*-alkyl sulfonamide. *Nat. Commun.* **9**, 1870 (2018).
204. Grimster, N. P. et al. Aromatic sulfonyl fluorides covalently kinetically stabilize transthyretin to prevent amyloidogenesis while affording a fluorescent conjugate. *J. Am. Chem. Soc.* **135**, 5656–5668 (2013).
205. Dalton, S. E. et al. Selectively targeting the kinase-conserved lysine of PI3K $\delta$  as a general approach to covalent kinase inhibition. *J. Am. Chem. Soc.* **140**, 932–939 (2018).
206. Ansbome, E. et al. Identification and characterization of an irreversible inhibitor of CDK2. *Chem. Biol.* **22**, 1159–1164 (2015).
207. Cuesta, A., Wan, X., Burlingame, A. L. & Taunton, J. Ligand conformational bias drives enantioselective modification of a surface-exposed lysine on Hsp90. *J. Am. Chem. Soc.* **142**, 3392–3400 (2020).
208. Zhao, Q. et al. Broad-spectrum kinase profiling in live cells with lysine-targeted sulfonyl fluoride probes. *J. Am. Chem. Soc.* **139**, 680–685 (2017).
209. Hett, E. C. et al. Rational targeting of active-site tyrosine residues using sulfonyl fluoride probes. *ACS Chem. Biol.* **10**, 1094–1098 (2015).
210. Zanon, P. R. A. et al. Profiling the proteome-wide selectivity of diverse electrophiles. *ChemRxiv* <https://doi.org/10.26434/chemrxiv.14186561.v1> (2021).
211. Chung, C. Y.-S. et al. Covalent targeting of the vacuolar H<sup>+</sup>-ATPase activates autophagy via mTORC1 inhibition. *Nat. Chem. Biol.* **15**, 776–785 (2019).
212. Strelow, J. M. A perspective on the kinetics of covalent and irreversible inhibition. *SLAS Discov.* **22**, 3–20 (2017).
213. McWhirter, C. in *Annual Reports in Medicinal Chemistry* (eds Ward, R. A. & Grimster, N. P.) vol. 56 1–31 (Elsevier, 2021).
214. Schnute, M. E. et al. Aminopyrazole carboxamide Bruton's tyrosine kinase inhibitors. Irreversible to reversible covalent reactive group tuning. *ACS Med. Chem. Lett.* **10**, 80–85 (2019).
215. Maurer, T. S., Tabrizi-Fard, M. A. & Fung, H. Impact of mechanism-based enzyme inactivation on inhibitor potency: implications for rational drug discovery. *J. Pharm. Sci.* **89**, 1404–1414 (2000).
216. Krippendorff, B.-F., Neuhaus, R., Lienau, P., Reichel, A. & Huisinga, W. Mechanism-based inhibition: deriving  $K_i$  and  $k_{inact}$  directly from time-dependent IC<sub>50</sub> values. *J. Biomol. Screen.* **14**, 913–923 (2009).
217. Miyahisa, I., Sameshima, T. & Hixon, M. S. Rapid determination of the specificity constant of irreversible inhibitors ( $k_{inact}/K_i$ ) by means of an endpoint competition assay. *Angew. Chem. Int. Ed.* **54**, 14099–14102 (2015).
218. Thorarensen, A. et al. The advantages of describing covalent inhibitor in vitro potencies by IC<sub>50</sub> at a fixed time point. IC<sub>50</sub> determination of covalent inhibitors provides meaningful data to medicinal chemistry for SAR optimization. *Bioorg. Med. Chem.* **29**, 115865 (2021).
219. Flanagan, M. E. et al. Chemical and computational methods for the characterization of covalent reactive groups for the prospective design of irreversible inhibitors. *J. Med. Chem.* **57**, 10072–10079 (2014).
220. Ward, R. A. et al. Structure- and reactivity-based development of covalent inhibitors of the activating and gatekeeper mutant forms of the epidermal growth factor receptor (EGFR). *J. Med. Chem.* **56**, 7025–7048 (2013).
221. Keeley, A., Ábrányi-Balogh, P. & Keserű, G. M. Design and characterization of a heterocyclic electrophilic fragment library for the discovery of cysteine-targeted covalent inhibitors. *Med. Chem. Commun.* **10**, 263–267 (2019).
222. Cee, V. J. et al. Systematic study of the glutathione (GSH) reactivity of *N*-arylacrylamides: 1. Effects of aryl substitution. *J. Med. Chem.* **58**, 9171–9178 (2015).
223. Birkholz, A. et al. Systematic study of the glutathione reactivity of *N*-phenylacrylamides: 2. Effects of acrylamide substitution. *J. Med. Chem.* **63**, 11602–11614 (2020).

## Acknowledgements

The authors thank the members of the Nomura Research Group for critical reading of the manuscript. This work was supported by the Nomura Research Group and the Mark Foundation for Cancer Research ASPIRE Award. This work was also supported by grants from the US NIH (award no. R01CA240981 for D.K.N. and no. F31CA268278 for N.J.H.), and the US National Science Foundation Graduate Fellowship (for L.B.).

## Author contributions

The authors contributed equally to all aspects of the article.

## Competing interests

D.K.N. is a co-founder, shareholder and adviser for Frontier Medicines and Vicinitas Therapeutics. D.K.N. is on the board of directors of Vicinitas Therapeutics. D.K.N. is also on the scientific advisory board of The Mark Foundation for Cancer Research and Photys Therapeutics. D.K.N. is a consultant for Droia Ventures. L.B. and N.J.H. are employees of Vicinitas Therapeutics.

## Peer review information

*Nature Reviews Drug Discovery* thanks Megan Matthews, who co-reviewed with Xie Wang, and the other, anonymous, reviewers for their contribution to the peer review of this work.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2022