Title

Targeting Non-Catalytic Cysteine Residues Through Structure-Guided Drug Discovery.

Journal

Current Topics in Medicinal Chemistry, 17(1)

ISSN

1568-0266

Authors

Hallenbeck, Kenneth K

Turner, David M

Renslo, Adam R

et al.

Publication Date

2017

DOI

10.2174/1568026616666160719163839

Peer reviewed

Published in final edited form as:

Curr Top Med Chem. 2017 ; 17(1): 4–15.

Targeting Non-Catalytic Cysteine Residues Through Structure-Guided Drug Discovery

Kenneth K. Hallenbeck[a], **David M. Turner**[a], **Adam R. Renslo**[a], and **Michelle R. Arkin**[a,*]

[a]Small Molecule Discovery Center and Department of Pharmaceutical Chemistry, University of California, San Francisco, California, USA

Abstract

The targeting of non-catalytic cysteine residues with small molecules is drawing increased attention from drug discovery scientists and chemical biologists. From a biological perspective, genomic and proteomic studies have revealed the presence of cysteine mutations in several oncogenic proteins, suggesting both a functional role for these residues and also a strategy for targeting them in an 'allele specific' manner. For the medicinal chemist, the structure-guided design of cysteine-reactive molecules is an appealing strategy to realize improved selectivity and pharmacodynamic properties in drug leads. Finally, for chemical biologists, the modification of cysteine residues provides a unique means to probe protein structure and allosteric regulation. Here, we review three applications of cysteine-modifying small molecules: 1) the optimization of existing drug leads, 2) the discovery of new lead compounds, and 3) the use of cysteine-reactive molecules as probes of protein dynamics. In each case, structure-guided design plays a key role in determining which cysteine residue(s) to target and in designing compounds with the proper geometry to enable both covalent interaction with the targeted cysteine and productive non-covalent interactions with nearby protein residues.

Keywords

Non-catalytic cysteine; Covalent drugs; Structure-based design; Chemical probes; disulfide Tethering; Lead optimization; Protein dynamics; Protein allostery

1. INTRODUCTION

Cysteine is an underrepresented residue in protein sequence (3.3% frequency [1]) but is disproportionately involved in protein function, with >50% of cysteine residues being solvent exposed and implicated in a myriad of biochemical processes [2]. Cysteine serves as the reactive nucleophile in many hydrolases (such as cysteine proteases) and can mediate redox reactions (e.g., protein disulfide isomerase). Oxidized forms of cysteine with sulfenic acid or nitrosothiol functionality are increasingly appreciated as playing a role in cellular signaling, and this suggests the possibility of targeting such oxidized forms with specific

*Address correspondence to this author at the Department of Pharmaceutical Chemistry, University of California, San Francisco, Box 2552, San Francisco, CA 94158, USA; Tel/Fax: ++1-415-514-4313, +1-415-514-4504; michelle.arkin@ucsf.edu.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflict of interest.

small molecules [3–4]. Finally, disulfide bond formation between two cysteine residues has been long recognized as contributing to protein tertiary structure. Taken together, these features make cysteine an attractive target for modification by small molecules. Here, we focus on approaches to target non-catalytic cysteine residues.

The concept behind covalent modification of cysteine residues is schematized in Fig. (**1**). An initial non-covalent complex (E*I) positions the electrophilic group within range of the nucleophilic thiol moiety and facilitates bond formation (k_2). For truly irreversible inhibitors, the resulting covalent complex (E-I) remains intact; however, for reversible electrophiles (e.g. disulfides), the ligand-bound complex dissociates (k_{-2}) over time to reform the initial non-covalent complex (E*I). Thus, cysteine-modifying drugs rely on two binding interactions – covalent and non-covalent – that can be independently and iteratively optimized to obtain the necessary selectivity and potency to be useful chemical probes or drug leads.

Cysteine-modifying compounds have been directed at both catalytic and non-catalytic residues. Modifying catalytic cysteines, such as those found in deubiquitinases and caspases, have an obvious impact on enzyme function. However, catalytic residues in enzyme active sites are generally highly conserved within families, and isoform selectivity can be difficult to achieve. Non-catalytic cysteines are generally less conserved, making them attractive for selective target modulation. Chemical proteomic studies employing activity-based probes have identified various reactive, functional, and non-catalytic cysteine residues whose functions could be probed and modulated with drug-like covalent molecules. These studies have revealed that inherent thiol reactivity spans six orders of magnitude [5], an observation that is gergermane in any effort to develop highly selective cysteine-targeted compounds.

As appreciation for the targetable nature of cysteine residues has grown, covalent approaches to drug discovery are also resurgent. However, covalent pharmacology inevitably raises the concern that reactive drugs or drug metabolites can induce organ damage or evoke an immune response through off-target (nonselective) binding to other proteins [6–7]. A related concern is that an electrophilic drug can lose activity because it is rapidly inactivated and eliminated via reaction with native nucleophiles (e.g. GSH) [6, 8].

These arguments are often countered by noting that many safe and well tolerated drugs in use for decades, such as aspirin and penicillin, act via covalent modification of their targets [9–10] and that the intentional targeting of nucleophilic sites with appropriately tuned electrophiles can mitigate the risk of covalent pharmacology. Giving appropriate attention to these potential issues is likely a factor in the recent clinical successes of covalent drug candidates [11].

Taking the advantages and challenges into account, the design of selective cysteine-modifying molecules as drug leads or as chemical probes for biomolecules has proven an attractive approach for the following distinct applications:

1. Lead optimization. Covalent pharmacology can enhance potency and selectivity of lead compounds, most compellingly by increasing target residence time. Covalent pharmacology is typically associated with ‘durable’ target inhibition *in*

vivo, and often very different pharmacokinetic/ pharmacodynamic (PK/PD) relationships as compared to drugs exhibiting reversible, fast-off inhibition kinetics [6]. Selectivity for protein isoforms containing the targeted residue is another benefit of this approach for lead optimization. Nearly any cysteine proximal to a known drug-binding site is a potential candidate. Knowledge of the structure, to support computational-guided design of the cysteine-reactive analog, is also highly desirable.

2. Chemical handles to identify new lead scaffolds. In targets where structural data indicates a binding pocket is available near a solvent-exposed cysteine, screening libraries of diverse molecules containing a cysteine-reactive moiety is an effective strategy for lead discovery. Computational approaches help triage promising target sites and identify scaffolds around which to build electrophile libraries for screening. One particularly interesting application for new drug discovery is targeting cysteine mutations found in oncogenic proteins; cysteine reactive molecules could first validate the function of these mutations in disease, then serve as lead compounds for therapeutic development.

3. Site-specific study of protein allostery, dynamics, and structure-function relationships. Native or engineered surface cysteines can be used to find molecules that bind at known sites of allosteric regulation, or can uncover previously undetected (‘cryptic’) binding pockets.

2. IMPROVING DRUG PROPERTIES FOR KNOWN SCAFFOLDS

2.1. Kinases

It is startling to recall that twenty-five years ago, kinases were considered ‘undruggable’ targets. The central importance of kinases and the high structural homology within the family suggested that imperfect selectivity would lead to unacceptable levels of toxicity. Through the creative efforts of many laboratories, kinase inhibitors are now a well-established class of cancer therapeutics. However, the development of drug resistance during kinase-inhibitor therapy is also common. First-generation inhibitors of epidermal growth factor receptor (EGFR) were effective in treating certain subtypes of lung carcinoma, but a mutation in the gatekeeper residue (T790M) resulted in a steric clash in the binding site (Fig. **2A**) ultimately leading to clinical relapse [12–13]. Walter and colleagues demonstrated that EGFR Cys797, which sits at the edge of the ATP-binding pocket and is present in just 2% of kinases, could be targeted to improve potency and recover function in the presence of T790M [14]. Selectively targeting a rare cysteine to increase drug residence time was expected to result in an improved clinical outcome (Fig. **2A, B**). However, the effectiveness of second-generation EGFR inhibitors was limited by on-target toxicity, since the drugs inhibited both mutant and WT EGFR [15–16]. To selectively target T790M EGFR, the 1st and 2nd generation quinazoline scaffold was replaced by other heterocyclic scaffolds into which an acrylamide electrophile could be readily introduced [17]. Several pyrimidine-based molecules were identified that exhibited 30 to 100-fold selectivity for T790M over WT EGFR. One of these pyrimidines, Osimertinib [18] was approved by the FDA in November 2015, while and Rociletinib [19], was not approved for treatment of non-small-cell lung

cancer [20–21]. The search for the next generation of covalent EGFR inhibitors continues, guided by rational design and lessons learned from the clinic [22].

As these examples illustrate, the Michael reaction between cysteine thiol as nucleophile and an alpha-beta unsaturated carbonyl (e.g. acrylamide) has figured prominently in the design of covalent kinase inhibitors. Michael ‘acceptors’ are attractive for these applications because their reactivity can be tuned by changing the nature of the electron-withdrawing carbonyl (or related) function and/or by altering the steric environment surrounding the electrophilic beta carbon atom.

The early success of covalent EGFR inhibitors motivated use of the approach in many other kinases. Ibrutinib, which received a breakthrough drug designation in 2013 for mantle cell lymphoma, del17p chronic lymphocytic leukemia and Waldenstrom's macroglobulinemia [23], contains an acrylamide warhead that irreversibly modifies Cys481 in Bruton’s tyrosine kinase (BTK). Ibrutinib’s scaffold was identified in a screen and was prioritized because it showed selectivity for a small group of Tec and Src-family kinases. Sequence comparison and structural homology modeling suggested that BTK contained a nucleophilic cysteine in the position analogous to Cys797 in EGFR [24]. This observation motivated a structure-guided medicinal chemistry effort that sampled three potential Michael acceptors – propiolamide, vinyl sulfonamide, and acrylamide – and found the last had the best activity (0.5 nM against BTK) and selectivity profile.

The Janus kinase (JAK) family have >80% amino acid identity in their ATP-binding site, exemplifying the kinase selectivity problem. Because of their importance to cytokine signaling pathways and potential as autoimmune disease therapeutics, many pan-JAK inhibitors have been described [25]. However, no reversible, isoform-specific inhibitors exist. JAK3 is the only JAK that contains a cysteine (Cys909) at the EGFR and BTK site, and was therefore targeted with tricyclic JAK inhibitors that included a terminal electrophile designed to irreversibly react with JAK3 Cys909 [26]. These compounds inhibited JAK3 with <100 nM potency in cells and selectively ablated JAK3-dependent signaling pathways, with little to no JAK2 activity up to 50 μM. However, Goedken and colleagues reported poor pharmacokinetic profiles for these JAK3 compounds [26], and more optimization is necessary before JAK3 inhibitors catch up to their EGFR or BTK counterparts.

As it happens, covalent kinase inhibition is not solely a product of human ingenuity. The fungal natural product hypothemycin [27] and related macrocycles [28] are known to covalently inhibit a subset of human kinases with a cysteine preceding the kinase DXG motif (Fig. **2C,D**). These socalled CDXG kinases comprise 48 of 518 human kinases and include important cancer drug targets such as MEK, ERK, PDGFR, VEGFR2, and FLT3 [29–30]. The macrocyclic structure of these compounds contains a *cis*-enone that serves as the cysteine-reactive moiety. The epoxide present in some family members is remarkably unreactive due to a macrocyclic conformation that blocks the approach of potential nucleophiles. The wholly synthetic drug candidate E6201, described as a dual MEK1 and FLT3 inhibitor, is in early clinical trials for advanced hematologic malignancies with documented FLT3 mutation [31]. The pharmacokinetics of E6201 in preclinical species and in humans is characterized by moderate to high distribution but rapid clearance [32]. This

PK profile might be regarded as appropriate for a drug exhibiting covalent pharmacology, since high distribution allows the compound to get to its target, but rapid clearance means that unbound, potentially reactive molecules are removed from circulation.

2.2. Beyond Kinases

These kinase examples illustrate a general principle: if a non-covalent scaffold for a target pocket is already available, adding a suitably positioned electrophilic group to form a covalent bond with a nearby non-catalytic cysteine residue provides large gains in potency. To achieve this, analogs of the inhibitor are prepared in which an electrophilic function, typically a Michael acceptor, is placed in a position and orientation informed by structural information about ligand binding, if available. Recent studies have extended this approach to non-kinase enzymes. For example, in 2011, Avila Therapeutics reported selective peptidomimetic inhibitors targeting a cysteine in Hepatitis C virus (HCV) protease [33]. Using sequence and structural alignment, they identified non-catalytic Cys159, which does not occur in human proteases but was conserved across all 919 HCV NS3 sequences known at the time. To target this cysteine, an acrylamide electrophile was introduced in a chemical scaffold based on the protease inhibitor telaprevir, realizing gains in IC_{50} from 2,500 nM for telaprevir to 2 nM for the electrophilic analog (Fig. **2E,F**). The addition of the acrylamide also improved selectivity vs off-target mammalian proteases. Whereas covalent protease inhibitors targeting active-site nucleophiles are common, to our knowledge, HCV NS3 is the only published example where a non-catalytic cysteine was leveraged to develop an irreversible protease inhibitor.

Daniel et al, recently reported a similar strategy for inhibiting histone deacetylases (HDACs). The HDAC inhibitor suberoylanilide hydroxamic acid (SAHA) was modified to include a covalent moiety that reacted with a conserved cysteine 5.6A from the enzyme active site [34]. This dual-action inhibitor did not achieve isoform selectivity or increased potency, but represented an interesting example of using structure-guided design and cysteine modification to develop a novel drug-targeting scheme.

3. IDENTIFYING NEW SCAFFOLDS

Incorporating cysteine reactivity into de novo drug (and probe) discovery efforts is a newer concept that builds on the successes described above. This hit-discovery strategy might be adopted because non-covalent approaches failed to give validated chemical starting points, as is often the case for challenging targets like protein-protein interactions and some classes of proteases, or because the targeted cysteine is hypothesized to be important for the protein’s pathological function. For instance, a recent survey of oncogenic mutations found that mutations to cysteine were 2.6-fold more likely than would be expected by chance; of the fifteen most commonly found cysteine mutants, fourteen were surface exposed and several were known to affect protein function [35]. Oncogenic cysteine mutations in the extracellular domain of fibroblast growth factor receptor (FGFR3), for instance, are found in 88% of mutated FGFR3, and have been shown to induce ligand-independent activation of the receptor through disulfide bond formation. Conceptually, these mutations could be strong candidates for a cysteine-targeted discovery approach.

Of surface-available, non-catalytic cysteines, only a subset is proximal to a site capable of interacting favorably with a small molecule. When structural data are available, computational methods for detecting potential binding pockets are therefore useful for assessing potential druggability. Several programs, including FTmap [36] and WaterMap [37], predict binding pockets on a protein surface. As an example, FTmap is a simple, DOCKING-based algorithm that scans a protein surface for propensity to bind very small organic molecules. Clusters of molecules indicate hotspots for ligand binding and suggest whether a tractable pocket is available [36]. FTMap has been applied to regions of FGFR3 near two of the known cysteine mutations, and the hotspot analysis suggests that one residue has more of a pocket than the other [35]. Neither pocket was high scoring, but it is important to note that proteins have regions of structural flexibility, and even small changes in side chain orientation or secondary structure can create cryptic pockets not seen in the unbound protein structure [38].

Of equal importance to the selection of a druggable cysteine/ binding site, is the selection of a suitable compound library to screen. As in traditional high-throughput screening, there are three general approaches: targeted libraries designed to bind conserved features of target class (e.g., kinases), diversity libraries meant for screening a wide range of targets, and virtual screens.

3.1. Computational Library Design

Computational methods have recently been described for virtual screening of covalent small molecules [39–40]. For example, *in silico* screening using DOCKovalent, an adaptation of DOCK3.6, led to experimentally validated inhibitors of the β-lactamase AmpC and kinases RSK2 and MSK1 [40]. In this method, a large virtual library based on commercially available compounds was built with a range of electrophiles; DOCKing was evaluated and the top 1% were manually prioritized for experimental validation. For AmpC, a library of boronic acids was screened for covalent modifiers of catalytic Ser64. Six diverse hits were tested for AmpC inhibition and three had a K_i < 1 μM. The most potent inhibitor (K_i = 40 nM) was crystalized to confirm the predicted docking pose. Comparisons with the original screen motivated the purchase of seven additional compounds, ultimately yielding a 10 nM inhibitor with a similar binding pose. This compound represented a novel AmpC inhibitor with good selectivity (>104-fold) over common serine proteases that bind boronic acids.

3.2. Experimental Library Design

Assembly of cysteine-reactive small molecule libraries for experimental screening has tended to use a fragment-based philosophy [41–42]. Fragment-based drug discovery (FBDD) seeks to identify low molecular weight fragments (typically <300 Da) that bind with high ligand efficiency (defined as binding energy/heavy atom [43]) to sub-pockets within a binding site. An attractive feature of FBDD is the ability to efficiently sample chemical space with a relatively small number of compounds (often ≤ 2000 fragments) [44]. Once fragments are identified, hits are further evolved into more complex and optimized leads [45–46]. Since a cysteine-reactive library generally needs to be synthesized from scratch, the focus on small libraries of fragment-sized molecules is appealing. Furthermore, the use of covalent elements in the fragment library serves to increase the initial potency,

making fragments easier to find in a binding- or activity-based screen. Finally, the use of engineered or native cysteine residues makes these methods site-directed, allowing the chemical biologist to evaluate the ligandability of a given site on a protein.

In designing a cysteine-reactive fragment library, one must consider the nature/reactivity of the electrophile, the structure of the non-covalent ‘diversity’ elements, and the linker between them. The distance between the electrophilic warhead and the diversity element is important because an effective hit must make productive interactions with the protein whilst also displaying the electrophile at the correct distance to react with the cysteine residue. Thus, linker lengths and geometries provide another opportunity for diversification. The composition of the linker can also have important effects on the chemical reactivity of the electrophile, as described below.

In selecting diversity elements for a library, one can take cues from a large literature on fragment library design [47– 48]. For instance, researchers at Astex proposed guidelines for constructing fragment libraries with desirable physiochemical properties [49]. This rule of thumb, dubbed the "Rule of three" [50] recommends a molecular weight <300 Da, number of H-bond acceptors ≤3, number H-bond donors ≤3, and cLogP ≤3. Further considerations include limiting the number of rotatable bonds to ≤3 and the polar surface area to ≤60 $Å^2$ [51]. Another recent trend favors selecting fragments with greater shape diversity, including more sp^3-rich structures to complement the generally ‘flat’ aromatic heterocycles commonly included in fragment libraries. While such fragments are underrepresented in commercial libraries, they can be accessed through bespoke synthesis or diversity-oriented synthesis [52–53].

Another noteworthy trend in library design is the consideration of ‘pan-acting interference’ (PAINs) compounds, also called ‘promiscuous,’ or ‘bad apples’ [54–58]. Several groups have identified chemical structures that tend to bind nonspecifically to proteins through covalent modification, aggregation, or redox reactions. These kinds of nonspecific covalent reactions are contrasted with the selective and low-reactivity electrophiles we describe here; however, it is important to be aware of (and perhaps remove) known ‘bad actors’ from the libraries during design [58].

Finally, electrophilic warheads must be selected. Electrophiles come in three flavors, irreversible, reversible, and slowly reversible, which will be considered separately.

3.3. Irreversible Warheads

Traditional cysteine-reactive compounds are irreversible electrophiles, and 100 – 200-member libraries of acrylates [41] and acrylamides [59–60] have recently been reported. Given the wide range of cysteine nucleophilicity in proteins, it is desirable to include a range of electrophilic warheads with differing reactivity during library construction. This concept was illustrated by Flanagan and coworkers at Pfizer using LCMS and NMR based kinetic studies to measure thiol reactivity for a number of irreversible electrophiles. This study revealed a 450-fold range in reaction rates, from minutes to days [61]. To further guide warhead design, computational methods are available to identify those electrophiles likely to exhibit undesirable non-selective protein reactivity [62–63]. As seen above, acrylamides are

commonly used warheads, since they often show low rates of reaction in solution. Non-covalent affinity for the binding site, however, increases the local concentration and residency time enough to allow reaction with the cysteine of interest.

Identifying the targets and off-targets of irreversibly covalent inhibitors can be significantly facilitated by the covalent nature of inhibition. In a recent example, the natural product hypothemycin (Fig. **2D**) was found to inhibit growth of the eukaryotic parasite T. brucei, the causative agent of Human African Trypanosomiasis (sleeping sickness). Though the parasite target was unknown, Nishino and Choy et al suspected a CDXG kinase, given the known reactivity of hypothemycin with mammalian CDXG kinases. They used the X-ray structure of hypothemycin bound to ERK2 (Fig. **2C**) to design a propargyl analog of hypothemycin for labeling and pull-down applications [64]. Application of this probe to T. brucei lysates, followed by 'click' conjugation of a fluorescent dye, allowed putative protein targets to be visualized by SDS-PAGE. Specific, 'saturable' binding interactions could be distinguished from non-specific labeling by co-incubation with hypothemycin, which competed with the probe for labeling of saturable (specific) targets. The same probe was also employed for pull-down and quantitative MS analysis, leading to the identification of TbGSK3short and TbCLK1/2 as bona fide, saturable targets of hypothemycin in T. brucei. The probe was then employed to demonstrate that hypothemycin treatment at cytotoxic concentrations fully inhibited TbCLK1 but only marginally inhibited TbGSK3short, suggesting that TbCLK1 was the central target. While hypothemycin proved to be a relatively selective electrophile in the parasite proteome, the same approach can be applied to identify off-targets for less selective covalent inhibitors [65].

3.4. Reversible Warheads

In contrast to irreversible fragment library screening, where compounds are selected through a combination of kinetic trapping and binding thermodynamics, disulfide trapping (tethering) uses readily reversible disulfide bonds to screen for fragments based primarily on thermodynamic stabilization [66].

In disulfide tethering, a library of disulfide-containing fragments is assayed against a cysteine-containing protein under reducing conditions. Mass spectrometry or functional assays are used to screen for fragments that form disulfide bonds with the desired cysteine thiolate [67]. Using reversible disulfide-exchange chemistry allows the screening assay to reach thermodynamic equilibrium; this equilibrium (and the stringency of the screen) is also controlled by the reduction potential of the buffer, which is varied based on the chemical reactivity of the target cysteine [68]. The library design also favors hit-selection based on non-covalent binding interactions; the disulfide moieties are separated from the diversity elements by 2–3 carbon aliphatic linkers, which serve to separate the diversity element from the reacting thiol and lend similar intrinsic nucleophilicity to each library member. Disulfide hits are then developed into leads through replacement of the disulfide bond with electrophiles such as acrylamide [69], or they can be converted to non-reactive ligands through removal of the thiol and structure-guided optimization [70]. For example, in developing inhibitors of the interleukin-2 (IL-2)/IL-2 receptor interactions, disulfide

trapping identified a pharmacophore that bound in a site proximal to that of a known inhibitor; linking the two compounds provided a 30-fold enhancement in affinity [71].

3.5. Converting Reversible to Irreversible Warheads

The GTPase K-Ras (Ki-ras2 Kirsten rat sarcoma viral oncogene homolog) is an oncogene with many cancer-associated activating mutations [72]. Pharmacological inhibition of K-Ras has been the goal of many drug discovery efforts but the target has proved mostly intractable [73]. One K-Ras mutation, G12C, introduces a cysteine that sits proximal to the Switch-I and -II regions involved in K-Ras nucleotide binding. A disulfide tethering screen of 480 fragments identified two fragments that labeled the G12C mutant [69]. Crystal structures of K-Ras bound to an optimized screening hit (compound 6; Fig. **3A,B**) informed the replacement of the reversible thiol with irreversible electrophiles such as vinyl sulfonamide and acrylamide. The less reactive acrylamides were selected for further optimization because these electrophiles have precedence in modern drugs. The acrylamide-containing compound bound to a previously unidentified allosteric pocket that formed under the switch-II pocket, stabilizing the inactive GDP-bound state and also disrupting the binding of effector proteins. Despite this novel mechanism of inhibition, the compound was ultimately incapable of engaging intracellular K-Ras, motivating efforts to find more potent G12C-specific molecules. Wellspring Biosciences recently reported ARS-853 (Fig. **3A,C**), which was optimized from the original acrylamides through iterative structure-guided design and crystallography [74]. ARS-853 follows a different trajectory in the binding site (Fig. **3A**), but binds to the same switch-II pocket and ablates downstream signaling in cells with low micromolar potency. ARS-853 is the first reported K-Ras inhibitor to reach the potency range of a quality drug lead.

The K-Ras G12C mutation is also close enough to the nucleotide binding site that it can be used in conjunction with a GDP analog to covalently inhibit at the active site. This approach yielded SML-8-73-1, a nucleotide analog that forms a stable thioether with G12C and competes for binding with GDP and GTP [75–76]. Though SML-8-73-1 bears two ionizable functions and is cell impermeable, cell-permeable analogs showed EC_{50} values of 25–45 μM against three K-Ras G12C cell lines. Achieving sub-micromolar cell-based activity with a GDP analog will be challenging and may require a permeable prodrug strategy. Nevertheless, the general strategy of targeting disease-specific cysteine mutants is a novel and highly attractive one that might well be applied to other targets with surface-exposed cysteine mutations [35].

3.6. Slowly Reversible Warheads

The most recent development in cysteine targeting has been the development of slowly reversible electrophiles. These warheads offer the benefits of irreversible and rapidly reversible chemistries, while mitigating their limitations. Since slowly reversible inhibitors can be selected based on thermodynamic binding, it should be possible to optimize their non-covalent interactions with the binding site. Additionally, reversibly covalent ligands could provide the high affinity and long residency time of covalent warheads while reducing the risks of immunogenicity arising from truly irreversible binding. Modulating the

residency time of the E-I complex (Fig. **1**) can enable on-target action *in vivo* long after free drug has been eliminated from circulation [77].

The concept of slowly reversible electrophiles was described by Taunton and coworkers in 2012 [78]. In this work, a Michael acceptor based on the α-cyanoacrylamide function was found to react with cysteine thiols in a slowly reversible reaction. The labeling of C436 on the kinase RSK2 was evaluated with a series of congeneric pyrrolopyrimidine inhibitors bearing various electrophilic side chains positioned to react with C436. Acrylate and acrylamide-based inhibitors showed irreversible inhibition, while the cyanoacrylamide analog exhibited reversible inhibition. Interestingly, the kinetically stable covalent bond between C436 and the cyanoacrylamide inhibitor was rapidly broken upon unfolding of the kinase domain of RSK2 with detergent. This result revealed that non-covalent interactions between the inhibitor and surrounding residues of the binding site served to stabilize the covalent adduct.

The slowly reversible reactivity of cyanoacrylamides motivated a subsequent effort to explore tuning of target residence time via systematic modifications to the cyanoacrylamide warhead [79]. Bradshaw *et al.* targeted C481 in BTK, the same cysteine targeted by the irreversible acrylamide Ibrutinib. Unlike C436 in RSK2, C481 in BTK lies outside the immediate confines of the ATP binding site. Accessing this surface-exposed residue thus necessitated the introduction of a piperidine spacer between the hinge-binding pyrazolopyrimidine core and the cyanoacrylamide. Additionally, the cyanoacrylamide warhead in the BTK inhibitor was ‘reversed’, with the linkage formed via the amide function of the cyanoacrylamide rather than at the beta carbon as in the RSK2 inhibitors. Thus, structure-based design was crucial in successfully engaging cysteines C436 and C481 in RSK2 and BTK, respectively.

Reversing the orientation of the cyanoacrylamide also allowed the kinetics of compound dissociation to be altered via introduction of various substituents at the electrophilic beta carbon. Increasing the steric demands of this substituent correlated with increasing target residence time (reduced off rate); 20 hours after removing unbound compound, ~60% of the *t*-Bu compound was still bound to BTK (Fig. **4**, compound **3**), compared to ~30% for *i*-Pr (**2**) and ~10% for the Me-containing compound (**1**). The authors hypothesize that the increased residence time was due to changes in binding orientation and shielding of the Cα proton, which slowed elimination of Cys481. Further exploration of diverse beta substituents led to a series of inhibitors exhibiting a range of target residence times, from minutes to several days. Importantly, the intrinsic potency of these inhibitors in biochemical assays was not correlated with the durability of target engagement, highlighting the potential of this approach to enable separate optimization of potency and target residence times to meet the requirements of a specific therapeutic application.

One of the most promising BTK inhibitors, exhibiting a residence time of >1 week *in vitro*, was tested for kinase selectivity across a panel of kinases. Only 6 of 254 tested kinases were inhibited >90% at 1 μM, and each of these sensitive kinases possessed the analogous C481. Importantly, other C481 kinases, including EGFR and JAK3 that were regarded as possible off-targets, were relatively unaffected (IC_{50} > 3 μM). Finally, this compound was evaluated

for BTK inhibition *in vivo* and was found to retain target engagement >24 hours after oral dosing, by which time free inhibitor had been cleared from circulation. Overall, the tunability of slowly reversible warheads and the corresponding benefits in terms of selectivity and *in vivo* PD properties make a convincing case for the wider application of this approach in drug discovery [79].

4. EXPLORING ALLOSTERY

Cysteine-targeted agents have also served as chemical-biology tools to explore protein conformation and allostery. Conformational changes, e.g., caused by posttranslational modification, protein-protein interactions, and the binding of signaling molecules, play a critical role in protein function. Designing molecules that affect a specific protein state could make for highly selective drugs. However, compounds that act at allosteric sites on proteins are often found serendipitously. Allosteric sites can also be cryptic sites, in that they are not apparent in crystal structures of the unbound protein and are therefore difficult to discover and model computationally. On the other hand, some cryptic sites might not bind endogenous ligands, yet might have nascent allosteric potential that synthetic ligands could harness. For these reasons, predicting and evaluating the functional relevance of cryptic sites are active areas of research, and the potential to proactively identify and target allosteric and/or cryptic sites remains an important challenge for structure-based drug discovery.

4.1. Native Cysteine Residues

In the case of caspases, disulfide trapping identified a previously unknown allosteric site. Caspase-7 is a dimeric cysteine protease and a potent effector of cell death by apoptosis. The enzyme is expressed as an inactive dimeric zymogen that is cleaved under apoptotic conditions to the active form; a series of loop movements causes a significant change in the overall conformation of the active sites and the dimer interface. Disulfide tethering identified two compounds called FICA and DICA (Fig. **3D,E**) that bound selectively to the Cys290 residue at the dimer interface and stabilized a zymogen-like, inactive conformation of the ‘active’ caspase-7 [80]. Caspase-1, a caspase involved in proinflammatory processes, was also found to have a cysteine residue (Cys331) at the dimer interface that could be targeted by a disulfide-containing compound (Fig. **3F,G**) [81]. Each of these disulfide-trapped compounds bound with two molecules tethered to symmetry-related cysteines at the interface, but the compounds bound with different orientations; for instance, compound 34 made compound-compound interactions in the caspase-1 site (Fig. **3G**), while the two bound DICA molecules did not interact (Fig. **3D**). For both proteins, analysis of the structures of the compound-trapped inactive state and the active conformation uncovered an allosteric network over the 15A between the allosteric site and the enzyme active site [82–83]. In an interesting and potentially generalizable application, compound 34-bound caspase-1 was used to select anti-caspase-1 antibodies that preferentially bound to the inactive conformation [84].

4.2. Engineered Cysteine Residues

Naturally, many allosteric or cryptic sites will not have native cysteine residues nearby. In these cases, engineering cysteine residues and probing them with cysteine-reactive

compounds can provide deep insight into protein structure and dynamics. The surface of the protein hormone IL-2 was probed with a series of cysteine residues followed by disulfide trapping. IL-2 is a four-helix bundle that binds to a trimeric receptor. Eleven cysteine mutations were made, one-at-a-time, along the surface of the alpha-chain binding site of IL-2 [38]. This face of IL-2 was found to be amphiphilic, with one side being hydrophilic, flat, and structurally stable, while the other side was more hydrophobic and structurally dynamic. Disulfide screening identified many more compounds that bound to the dynamic portion of the interface, in some cases trapping conformations not seen in structures of the apo protein, including conformations induced by the binding of ligands at other locations on the surface of IL-2. Thus, protein-protein interfaces can have regions of structural adaptivity, which might have a role in binding multiple protein partners [85] and/or may be exploited by small-molecule ligands [38, 86].

Protein-protein interactions can also allosterically regulate enzyme activity, as is the case with different classes of kinases. The AGC kinases, for instance, have a common allosteric site in the N-lobe of the kinase domain, where substrate proteins or regulatory domains of the kinase itself can bind [87]. Binding to this allosteric site co-localizes the substrate and enzyme and also allosterically activates catalysis by positioning the regulatory ‘C helix’ into an active position. The AGC kinase PDK1 is a well-studied example of this allosteric regulation, and multiple small-molecule modulators that bind to the allosteric site have been designed. Cysteine mutagenesis followed by disulfide trapping identified several compounds that allosterically activated or inhibited kinase activity [88]. Importantly, both inhibitors and activators could be selected at the same cysteine residue; hence, the details of the non-covalent binding interactions and molecular shape determined allosteric outcome, not structural changes in the protein due to cysteine mutagenesis per se. X-ray structures of an activator and inhibitor bound to PDK1 (Fig. **3H**) highlighted the mechanism of allostery; the smaller compound (Fig. **3J**) stabilized a conformation in which the regulatory C-helix was pulled away from the active site, while the larger compound (Fig. **3I**) pushed the Chelix down and into the active conformation [88]. These studies underscore the subtlety between binding and allostery that makes small-molecule design of allosteric modulators both fascinating and complicated.

Bishop and coworkers have taken a protein engineering approach to develop selective allosteric inhibitors of protein tyrosine phosphatase (PTP) enzymes [89]. This important class of signaling enzymes has encountered significant challenges for drug discovery, given the similarity of the active site across the family and the challenge with obtaining cell-permeable active-site inhibitors. The Bishop lab found that the PTP Shp2 was sensitive to inhibition by the cysteine-reactive dye FLAsH, which binds to 2-, 3-, or 4 cysteine residues. They identified two nearby cysteine residues, Cys333 and Cys367 that were buried in the apo-structure of Shp2, but became surface accessible in the presence of FLAsH [89]. Cys333 is unique to Shp2 among PTPs, potentially providing a novel therapeutic strategy for targeting Shp2. However, FLAsH itself did not bind tightly enough to wild type Shp2 to bind selectively in cell lysates, so the investigators engineered an additional Cys368 to provide trivalent coordination of FLAsH [90]. Intriguingly, they were able to create similar allosteric sites by engineering cysteine residues at the analogous position in several PTPs. Thus, they

discovered a novel, cryptic allosteric site that can be used to probe the functions of specific PTPs in cells.

Cysteine mutagenesis/reactivity has also been used to experimentally validate potential hidden/cryptic allosteric sites identified computationally. Bowman, *et al.* utilized a Markov state model that evaluated protein structural changes on the microsecond to millisecond timescale [91]. Using a drug-resistant mutant of TEM-1 beta-lactamase as a model system, they collected an ensemble of protein structures and looked for transient pockets that a) were fragment-to-lead sized, b) correlated with motions at the active site, and c) included residues that changed from buried to surface-exposed upon pocket formation. They then mutated these residues to cysteine and used the thiol-detection reagent 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) to demonstrate that they could be labeled. The rate of DTNB reaction with cysteine mutants supported the hypothesis that pockets opened and closed transiently. In three cases, TNB-labeled enzyme had a reduced catalytic efficiency, suggesting that trapping these pockets did have an allosteric effect on the active site. The authors envisaged a pipeline in which cryptic pockets would be identified computationally, then validated through cysteine mutation and screening with cysteine-reactive libraries. Structure-guided design could then be used to optimize the compounds so that they allosterically inhibit the non-cysteine containing protein.

CONCLUSION

The post-translational modification (PTM) of reactive amino acid side chains is well recognized as an essential mechanism by which biology regulates cell signaling, protein structure, and the epigenetic control of gene expression. The various examples and approaches to cysteine modification described herein might be considered as examples of unnatural PTM leveraging the vastly greater access to chemical space enabled by synthetic organic chemistry. While chemical biologists and drug discovery scientists have yet to equal the exquisite selectivity of biochemical PTM, structure-guided design and ever improving computational tools for predicting chemical reactivity and ligand binding portend a bright future for cysteine-reactive small molecules in chemical biology and drug discovery.

Acknowledgments

The authors thank the reviewers for insightful comments and suggestions, the National Science Foundation for a Predoctoral Fellowship (KKH) and Pfizer Inc. for postdoctoral support (DMT).

Biography

Michelle R. Arkin

LIST OF ABBREVIATIONS

BTK	Bruton's Tyrosine Kinase
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
EGFR	Epidermal Growth Factor Receptor
ERK	Extracellular-Signal-Regulated Kinase
FBDD	Fragment Based Drug Discovery
FGFR	Fibroblast Growth Factor Receptor
GSH	Glutathione
HCV	Hepatitis C Virus
HDAC	Histone Deacetylase
IL-2	Interleukin-2
JAK	Janus Kinase
LCMS	Liquid Chromatography Mass Spectrometry
MOA	Mechanism of Action
NMR	Nuclear Magnetic Resonance
PD	Pharmacodynamics
PDGFR	Platelet-Derived Growth Factor Receptor
PDK1	Phosphoinositide-Dependent Kinase-1
PK	Pharmacokinetics
PTM	Post-Translational Modification
PTP	Protein Tyrosine Phosphatase
RSK	Ribosomal s6 Kinase
SAHA	Suberoylanilide Hydroxamic Acid
VEGFR	Vascular Endothelial Growth Factor Recepotor

REFERENCES

1. King JL, Jukes TH. Non-Darwinian evolution. Science (New York, N.Y.). 1969; 164(3881):788–798.
2. Requejo R, Hurd TR, Costa NJ, Murphy MP. Cysteine residues exposed on protein surfaces are the dominant intramito-chondrial thiol and may protect against oxidative damage. Febs J. 2010; 277(6): 1465–1480. [PubMed: 20148960]

3. Leonard SE, Reddie KG, Carroll KS. Mining the Thiol Proteome for Sulfenic Acid Modifications Reveals New Targets for Oxidation in Cells. ACS Chem. Biol. 2009; 4(9):783–799. [PubMed: 19645509]
4. Majmudar JD, Martin BR. Strategies for profiling native S-nitrosylation. Biopolymers. 2014; 101(2):173–179. [PubMed: 23828013]
5. Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MBD, Bachovchin DA, Mowen K, Baker D, Cravatt BF. Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature. 2010; 468(7325):790–795. [PubMed: 21085121]
6. Johnson DS, Weerapana E, Cravatt BF. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. Fut. Med. Chem. 2010; 2(6):949–964.
7. Mah R, Thomas JR, Shafer CM. Drug discovery considerations in the development of covalent inhibitors. Bioorg. Med. Chem. Lett. 2014; 24(1):33–39. [PubMed: 24314671]
8. David-Cordonnier M-H, Laine W, Joubert A, Tardy C, Goossens J-F, Kouach M, Briand G, Thi Mai HD, Michel S, Tillequin F, Koch M, Leonce S, Pierre A, Bailly C. Covalent binding to glutathione of the DNA-alkylating antitumor agent, S23906-1. Euro. J. Biochem. 2003; 270(13):2848–2859.
9. Bauer RA. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. Drug Discov. Today. 2015; 20(9):1061–1073. [PubMed: 26002380]
10. Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. Nat. Rev. Drug Discov. 2011; 10(4):307–317. [PubMed: 21455239]
11. Moghaddam MF, Tang Y, O'Brien Z, Richardson SJ, Bacolod M, Chaturedi P, Apuy J, Kulkarni A. A Proposed Screening Paradigm for Discovery of Covalent Inhibitor Drugs. Drug Metab. Lett. 2014; 8(1):19–30. [PubMed: 24628405]
12. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, Kris MG, Varmus H. Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. PLoS. Med. 2005; 2(3):e73. [PubMed: 15737014]
13. Yun C-H, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong K-K, Meyerson M, Eck MJ. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc. Natl. Acad. Sci. U.S.A. 2008; 105(6):2070–2075. [PubMed: 18227510]
14. Walter AO, Sjin RTT, Haringsma HJ, Ohashi K, Sun J, Lee K, Dubrovskiy A, Labenski M, Zhu Z, Wang Z, Sheets M, Martin TS, Karp R, van Kalken D, Chaturvedi P, Niu D, Nacht M, Petter RC, Westlin W, Lin K, Jaw-Tsai S, Raponi M, Dyke TV, Etter J, Weaver Z, Pao W, Singh J, Simmons AD, Harding TC, Allen A. Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. Cancer Discov. 2013; 3(12):1404–1415. [PubMed: 24065731]
15. Katakami N, Atagi S, Goto K, Hida T, Horai T, Inoue A, Ichinose Y, Koboyashi K, Takeda K, Kiura K, Nishio K, Seki Y, Ebisawa R, Shahidi M, Yamamoto N. LUX-Lung 4: A Phase II Trial of Afatinib in Patients With Advanced Non–Small-Cell Lung Cancer Who Progressed During Prior Treatment With Erlotinib, Gefitinib, or Both. J. Clin. Oncol. 2013; 31(27):3335–3341. [PubMed: 23816963]
16. Sequist LV, Besse B, Lynch TJ, Miller VA, Wong KK, Gitlitz B, Eaton K, Zacharchuk C, Freyman A, Powell C, Ananthakrishnan R, Quinn S, Soria J-C. Neratinib, an Irreversible Pan-ErbB Receptor Tyrosine Kinase Inhibitor: Results of a Phase II Trial in Patients With Advanced Non–Small-Cell Lung Cancer. J. Clin. Oncol. 2010; 28(18):3076–3083. [PubMed: 20479403]
17. Zhou W, Ercan D, Chen L, Yun C-H, Li D, Capelletti M, Cortot AB, Chirieac L, Iacob RE, Padera R, Engen JR, Wong K-K, Eck MJ, Gray NS, Janne PA. Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. Nature. 2009; 462(7276):1070–1074. [PubMed: 20033049]
18. ClinicalTrials.gov Search: Osimertinib. https://www.clinicaltri-als.gov/ct2/results?term=Osimertinib
19. ClinicalTrials.gov Search: Rociletinib. https://www.clinicaltrials.gov/ct2/results?term=Rociletinib
20. Jänne PA, Yang JC-H, Kim D-W, Planchard D, Ohe Y, Ramalingam SS, Ahn M-J, Kim S-W, Su W-C, Horn L, Haggstrom D, Felip E, Kim J-H, Frewer P, Cantarini M, Brown KH, Dickinson PA, Ghiorghiu S, Ranson M. AZD9291 in EGFR Inhibitor–Resistant Non–Small-Cell Lung Cancer. N. Engl. J. Med. 2015; 372(18):1689–1699. [PubMed: 25923549]

21. Sequist LV, Soria J-C, Goldman JW, Wakelee HA, Gadgeel SM, Varga A, Papadimitrakopoulou V, Solomon BJ, Oxnard GR, Dziadziuszko R, Aisner DL, Doebele RC, Galasso C, Garon EB, Heist RS, Logan J, Neal JW, Mendenhall MA, Nichols S, Piotrowska Z, Wozniak AJ, Raponi M, Karlovich CA, Jaw-Tsai S, Isaacson J, Despain D, Matheny SL, Rolfe L, Allen AR, Camidge DR. Rociletinib in EGFR-Mutated Non–Small-Cell Lung Cancer. N. Engl. J. Med. 2015; 372(18): 1700–1709. [PubMed: 25923550]
22. Engel J, Richters A, Getlik M, Tomassi S, Keul M, Termathe M, Lategahn J, Becker C, Mayer-Wrangowski S, Grütter C, Uhlenbrock N, Krüll J, Schaumann N, Eppmann S, Kibies P, Hoffgaard F, Heil J, Menninger S, Ortiz-Cuaran S, Heuckmann JM, Tinnefeld V, Zahedi RP, Sos ML, Schultz-Fademrecht C, Thomas RK, Kast SM, Rauh D. Targeting Drug Resistance in EGFR with Covalent Inhibitors: A Structure-Based Design Approach. J. Med. Chem. 2015; 58(17):6844–6863. [PubMed: 26275028]
23. Gayko U, Fung M, Clow F, Sun S, Faust E, Price S, James D, Doyle M, Bari S, Zhuang SH. Development of the Bruton's tyrosine kinase inhibitor ibrutinib for B cell malignancies. Ann. N. Y. Acad. Sci. 2015; 1358(1):82–94. [PubMed: 26348626]
24. Pan Z, Scheerens H, Li S-J, Schultz BE, Sprengeler PA, Burrill LC, Mendonca RV, Sweeney MD, Scott KCK, Grothaus PG, Jeffery DA, Spoerke JM, Honigberg LA, Young PR, Dalrymple SA, Palmer JT. Discovery of Selective Irreversible Inhibitors for Bruton's Tyrosine Kinase. Chem Med Chem. 2007; 2(1):58–61. [PubMed: 17154430]
25. O'Shea JJ, Kontzias A, Yamaoka K, Tanaka Y, Laurence A. Janus kinase inhibitors in autoimmune diseases. Ann. Rheum. Dis. 2013; 72(suppl 2):ii111–ii115. [PubMed: 23532440]
26. Goedken ER, Argiriadi MA, Banach DL, Fiamengo BA, Foley SE, Frank KE, George JS, Harris CM, Hobson AD, Ihle DC, Marcotte D, Merta PJ, Michalak ME, Murdock SE, Tomlinson MJ, Voss JW. Tricyclic Covalent Inhibitors Selectively Target Jak3 through an Active Site Thiol. J. Biol. Chem. 2015; 290(8):4573–4589. [PubMed: 25552479]
27. Tanaka H, Nishida K, Sugita K, Yoshioka T. Antitumor efficacy of hypothemycin, a new Ras-signaling inhibitor. Jpn. J. Cancer. Res. 1999; 90(10):1139–1145. [PubMed: 10595743]
28. Barluenga S, Jogireddy R, Koripelly GK, Winssinger N. *In Vivo* Efficacy of Natural Product-Inspired Irreversible Kinase Inhibitors. Chem Bio Chem. 2010; 11(12):1692–1699.
29. Winssinger N, Barluenga S. Chemistry and biology of resorcylic acid lactones. Chem. Commun. 2007; (1):22–36.
30. Schirmer A, Kennedy J, Murli S, Reid R, Santi DV. Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. Proc. Natl. Acad. Sci. U.S.A. 2006; 103(11):4234–4239. [PubMed: 16537514]
31. ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT02418000
32. Kumar V, Schuck EL, Pelletier RD, Farah N, Condon KB, Ye M, Rowbottom C, King BM, Zhang Z-Y, Saxton PL, Wong YN. Pharmacokinetic characterization of a natural product–inspired novel MEK1 inhibitor E6201 in preclinical species. Cancer Chemother. Pharmacol. 2011; 69(1):229–237. [PubMed: 21698359]
33. Hagel M, Niu D, St Martin T, Sheets MP, Qiao L, Bernard H, Karp RM, Zhu Z, Labenski MT, Chaturvedi P, Nacht M, Westlin WF, Petter RC, Singh J. Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine. Nat, Chem. Biol. 2011; 7(1):22–24. [PubMed: 21113170]
34. Daniel KB, Sullivan ED, Chen Y, Chan JC, Jennings PA, Fierke CA, Cohen SM. Dual-Mode HDAC Prodrug for Covalent Modification and Subsequent Inhibitor Release. J. Med. Chem. 2015; 58(11):4812–4821. [PubMed: 25974739]
35. Visscher M, Arkin MR, Dansen TB. Covalent targeting of acquired cysteines in cancer. Curr. Opin. Chem. Biol. 2016; 30:61–67. [PubMed: 26629855]
36. Kozakov D, Grove LE, Hall DR, Bohnuud T, Mottarella SE, Luo L, Xia B, Beglov D, Vajda S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. Nat. Protocols. 2015; 10(5):733–755. [PubMed: 25855957]
37. Beuming T, Che Y, Abel R, Kim B, Shanmugasundaram V, Sherman W. Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. Proteins: Struct., Funct., Bioinf. 2012; 80(3):871–883.

38. Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC. Binding of small molecules to an adaptive protein–protein interface. Proc. Natl. Acad. Sci. U.S.A. 2003; 100(4):1603–1608. [PubMed: 12582206]
39. Dong GQ, Calhoun S, Fan H, Kalyanaraman C, Branch MC, Mashiyama ST, London N, Jacobson MP, Babbitt PC, Shoichet BK, Armstrong RN, Sali A. Prediction of Substrates for Glutathione Transferases by Covalent Docking. J. Chem. Inform. Modeling. 2014; 54(6):1687–1699.
40. London N, Miller RM, Krishnan S, Uchida K, Irwin JJ, Eidam O, Gibold L, Cimermančič P, Bonnet R, Shoichet BK, Taunton J. Covalent docking of large libraries for the discovery of chemical probes. Nat. Chem. Biol. 2014; 10(12):1066–1072. [PubMed: 25344815]
41. Kathman SG, Xu Z, Statsyuk AV. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. J. Med. Chem. 2014; 57(11):4969–4974. [PubMed: 24870364]
42. Miller RM, Paavilainen VO, Krishnan S, Serafimova IM, Taunton J. Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors. J. Am. Chem. Soc. 2013; 135(14):5298–5301. [PubMed: 23540679]
43. Hopkins, AL., Groom, CR., Alex, A. Drug Discov. Today. Vol. 9. England: 2004. Ligand efficiency: a useful metric for lead selection; p. 430-431.
44. Murray CW, Rees DC. The rise of fragment-based drug discovery. Nat. Chem. 2009; 1(3):187–192. [PubMed: 21378847]
45. Kumar A, Zhang AVAKYJ. Fragment Based Drug Design: From Experimental to Computational Approaches. Curr. Med. Chem. 2012; 19(30):5128–5147. [PubMed: 22934764]
46. Joseph-McCarthy D, Campbell AJ, Kern G, Moustakas D. Fragment-Based Lead Discovery and Design. J. Chem. Inform. Modeling. 2014; 54(3):693–704.
47. Doak BC, Morton CJ, Simpson JS, Scanlon MJ. Design and Evaluation of the Performance of an NMR Screening Fragment Library. Aust. J. Chem. 2013; 66(12):1465–1472.
48. Francis CL, Kenny PW, Dolezal O, Saubern S, Kruger M, Savage GP, Peat TS, Ryan JH. Construction of the CSIRO Fragment Library. Aust. J. Chem. 2013; 66(12):1473–1482.
49. Rahman, A. Frontiers in Drug Design and Discovery. Bentham Science Publishers; 2007.
50. Jhoti H, Williams G, Rees DC, Murray CW. The 'rule of three' for fragment-based drug discovery: where are we now? Nat. Rev. Drug Discov. 2013; 12(8):644–644. [PubMed: 23845999]
51. Congreve M, Carr R, Murray C, Jhoti H. A ‘Rule of Three’ for fragment-based lead discovery? Drug Discov. Today. 2003; 8(19):876–877.
52. Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discov. 2007; 6(3):211–219. [PubMed: 17290284]
53. Hung AW, Ramek A, Wang Y, Kaya T, Wilson JA, Clemons PA, Young DW. Route to three-dimensional fragments using diversity-oriented synthesis. Proc. Natl. Acad. Sci. U.S.A. 2011; 108(17):6799–6804. [PubMed: 21482811]
54. Dahlin JL, Nissink JWM, Strasser JM, Francis S, Higgins L, Zhou H, Zhang Z, Walters MA. PAINS in the Assay: Chemical Mechanisms of Assay Interference and Promiscuous Enzymatic Inhibition Observed during a Sulfhydryl-Scavenging HTS. J. Med. Chem. 2015; 58(5):2091–2113. [PubMed: 25634295]
55. Baell J, Walters MA. Chemistry: Chemical con artists foil drug discovery. Nature. 2014; 513(7519):481–483. [PubMed: 25254460]
56. McGovern SL, Caselli E, Grigorieff N, Shoichet BK. A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. J. Med. Chem. 2002; 45(8): 1712–1722. [PubMed: 11931626]
57. Oprea TI, Davis AM, Teague SJ, Leeson PD. Is There a Difference between Leads and Drugs? A Historical Perspective. J. Chem. Inform. Comput. Sci. 2001; 41(5):1308–1315.
58. Baell JB, Holloway GA. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. J. Med. Chem. 2010; 53(7):2719–2740. [PubMed: 20131845]
59. Kwarcinski FE, Steffey ME, Fox CC, Soellner MB. Discovery of Bivalent Kinase Inhibitors via Enzyme-Templated Fragment Elaboration. ACS. Med. Chem. Lett. 2015; 6(8):898–901. [PubMed: 26286460]

60. Allen CE, Curran PR, Brearley AS, Boissel V, Sviridenko L, Press NJ, Stonehouse JP, Armstrong A. Efficient and facile synthesis of acrylamide libraries for protein-guided tethering. Org. Lett. 2015; 17(3):458–460. [PubMed: 25582165]

61. Flanagan ME, Abramite JA, Anderson DP, Aulabaugh A, Dahal UP, Gilbert AM, Li C, Montgomery J, Oppenheimer SR, Ryder T, Schuff BP, Uccello DP, Walker GS, Wu Y, Brown MF, Chen JM, Hayward MM, Noe MC, Obach RS, Philippe L, Shanmugasundaram V, Shapiro MJ, Starr J, Stroh J, Che Y. Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. J. Med. Chem. 2014; 57(23): 10072–10079. [PubMed: 25375838]

62. Schwöbel JAH, Koleva YK, Enoch SJ, Bajot F, Hewitt M, Madden JC, Roberts DW, Schultz TW, Cronin MT. D. Measurement and Estimation of Electrophilic Reactivity for Predictive Toxicology. Chem. Rev. 2011; 111(4):2562–2596. [PubMed: 21401043]

63. Enoch SJ, Ellison CM, Schultz TW, Cronin MT. D. A review of the electrophilic reaction chemistry involved in covalent protein binding relevant to toxicity. Crit. Rev. Toxicol. 2011; 41(9): 783–802. [PubMed: 21809939]

64. Nishino M, Choy JW, Gushwa NN, Oses-Prieto JA, Koupparis K, Burlingame AL, Renslo AR, McKerrow JH, Taunton J. Hypothemicin, a fungal natural product, identifies therapeutic targets in Trypanosoma brucei. eLife. 2013; 2:e00712. [PubMed: 23853713]

65. Choy JW, Bryant C, Calvet CM, Doyle PS, Gunatilleke SS, Leung SSF, Ang KKH, Chen S, Gut J, Oses-Prieto JA, Johnston JB, Arkin MR, Burlingame AL, Taunton J, Jacobson MP, McKerrow JM, Podust LM, Renslo AR. Chemical–biological characterization of a cruzain inhibitor reveals a second target and a mammalian off-target. Beilstein. J. Org. Chem. 2013; 9:15–25. [PubMed: 23400640]

66. Erlanson DA, Braisted AC, Raphael DR, Randal M, Stroud RM, Gordon EM, Wells JA. Site-directed ligand discovery. Proc. Natl. Acad. Sci. U.S.A. 2000; 97(17):9367–9372. [PubMed: 10944209]

67. Lodge JM, Justin Rettenmaier T, Wells JA, Pomerantz WC, Mapp AK. FP tethering: a screening technique to rapidly identify compounds that disrupt protein-protein interactions. Med-ChemComm. 2014; 5(3):370–375.

68. Erlanson DA, Lam JW, Wiesmann C, Luong TN, Simmons RL, DeLano WL, Choong IC, Burdett MT, Flanagan WM, Lee D, Gordon EM, O'Brien T. In situ assembly of enzyme inhibitors using extended tethering. Nat. Biotech. 2003; 21(3):308–314.

69. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. Nature. 2013; 503(7477):548–551. [PubMed: 24256730]

70. Erlanson DA, Arndt JW, Cancilla MT, Cao K, Elling RA, English N, Friedman J, Hansen SK, Hession C, Joseph I, Kumaravel G, Lee W-C, Lind KE, McDowell RS, Miatkowski K, Nguyen C, Nguyen TB, Park S, Pathan N, Penny DM, Romanowski MJ, Scott D, Silvian L, Simmons RL, Tangonan BT, Yang W, Sun L. Discovery of a potent and highly selective PDK1 inhibitor via fragment-based drug discovery. Bioorg. Med. Chem. Lett. 2011; 21(10):3078–3083. [PubMed: 21459573]

71. Raimundo BC, Oslob JD, Braisted AC, Hyde J, McDowell RS, Randal M, Waal ND, Wilkinson J, Yu CH, Arkin MR. Integrating Fragment Assembly and Biophysical Methods in the Chemical Advancement of Small-Molecule Antagonists of IL-2: An Approach for Inhibiting Protein–Protein Interactions†. J. Med. Chem. 2004; 47(12):3111–3130. [PubMed: 15163192]

72. Adjei AA. Blocking Oncogenic Ras Signaling for Cancer Therapy. J. Natl. Cancer Inst. 2001; 93(14):1062–1074. [PubMed: 11459867]

73. Wang Y, Kaiser CE, Frett B, Li H-Y. Targeting Mutant KRAS for Anticancer Therapeutics: A Review of Novel Small Molecule Modulators. J. Med. Chem. 2013; 56(13):5219–5230. [PubMed: 23566315]

74. Patricelli MP, Janes MR, Li L-S, Hansen R, Peters U, Kessler LV, Chen Y, Kucharski JM, Feng J, Ely T, Chen JH, Firdaus SJ, Babbar A, Ren P, Liu Y. Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. Cancer Discov. 2016

75. Hunter JC, Gurbani D, Ficarro SB, Carrasco MA, Lim SM, Choi HG, Xie T, Marto JA, Chen Z, Gray NS, Westover KD. In situ selectivity profiling and crystal structure of SML-8-73-1, an active

site inhibitor of oncogenic K-Ras G12C. Proc. Natl. Acad. Sci. U.S.A. 2014; 111(24):8895–8900. [PubMed: 24889603]

76. Lim SM, Westover KD, Ficarro SB, Harrison RA, Choi HG, Pacold ME, Carrasco M, Hunter J, Kim ND, Xie T, Sim T, Jänne PA, Meyerson M, Marto JA, Engen JR, Gray NS. Therapeutic Targeting of Oncogenic K-Ras by a Covalent Catalytic Site Inhibitor. Angewandte Chemie Intl. Edition. 2014; 53(1):199–204.
77. Copeland RA, Pompliano DL, Meek TD. Drug-target residence time and its implications for lead optimization. Nat. Rev. Drug Discov. 2006; 5(9):730–739. [PubMed: 16888652]
78. Serafimova IM, Pufall MA, Krishnan S, Duda K, Cohen MS, Maglathlin RL, McFarland JM, Miller RM, Frodin M, Taunton J. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. Nat. Chem. Biol. 2012; 8(5):471–476. [PubMed: 22466421]
79. Bradshaw JM, McFarland JM, Paavilainen VO, Bisconte A, Tam D, Phan VT, Romanov S, Finkle D, Shu J, Patel V, Ton T, Li X, Loughhead DG, Nunn PA, Karr DE, Gerritsen ME, Funk JO, Owens TD, Verner E, Brameld KA, Hill RJ, Goldstein DM, Taunton J. Prolonged and tunable residence time using reversible covalent kinase inhibitors. Nat. Chem. Biol. 2015; 11(7):525–531. [PubMed: 26006010]
80. Hardy JA, Lam J, Nguyen JT, O'Brien T, Wells JA. Discovery of an allosteric site in the caspases. Proc. Natl. Acad. Sci. U.S.A. 2004; 101(34):12461–12466. [PubMed: 15314233]
81. Scheer JM, Romanowski MJ, Wells JA. A common allosteric site and mechanism in caspases. Proc. Natl. Acad. Sci. U.S.A. 2006; 103(20):7595–7600. [PubMed: 16682620]
82. Datta D, Scheer JM, Romanowski MJ, Wells JA. An Allosteric Circuit in Caspase-1. J. Mol. Biol. 2008; 381(5):1157–1167. [PubMed: 18590738]
83. Hardy JA, Wells JA. Dissecting an Allosteric Switch in Caspase-7 Using Chemical and Mutational Probes. J. Biol. Chem. 2009; 284(38):26063–26069. [PubMed: 19581639]
84. Gao J, Sidhu SS, Wells JA. Two-state selection of conformation- specific antibodies. Proc. Natl. Acad. Sci. U.S.A. 2009; 106(9):3071–3076. [PubMed: 19208804]
85. DeLano WL, Ultsch MH, de AM, Vos, Wells JA. Convergent Solutions to Binding at a Protein-Protein Interface. Science (New York, N.Y.). 2000; 287(5456):1279–1283.
86. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature. 2007; 450(7172):1001–1009. [PubMed: 18075579]
87. Gold MG, Barford D, Komander D. Lining the pockets of kinases and phosphatases. Curr. Opin. Structural Biol. 2006; 16(6):693–701.
88. Sadowsky JD, Burlingame MA, Wolan DW, McClendon CL, Jacobson MP, Wells JA. Turning a protein kinase on or off from a single allosteric site via disulfide trapping. Proc. Natl. Acad. Sci. U.S.A. 2011; 108(15):6056–6061. [PubMed: 21430264]
89. Chio CM, Lim CS, Bishop AC. Targeting a Cryptic Allosteric Site for Selective Inhibition of the Oncogenic Protein Tyrosine Phosphatase Shp2. Biochemistry. 2015; 54(2):497–504. [PubMed: 25519989]
90. Chio CM, Yu X, Bishop AC. Rational design of allosteric-inhibition sites in classical protein tyrosine phosphatases. Bioorg. Med. Chem. 2015; 23(12):2828–2838. [PubMed: 25828055]
91. Bowman GR, Bolin ER, Hart KM, Maguire BC, Marqusee S. Discovery of multiple hidden allosteric sites by combining Markov state models and experiments. Proc. Natl. Acad. Sci. U.S.A. 2015; 112(9):2734–2739. [PubMed: 25730859]

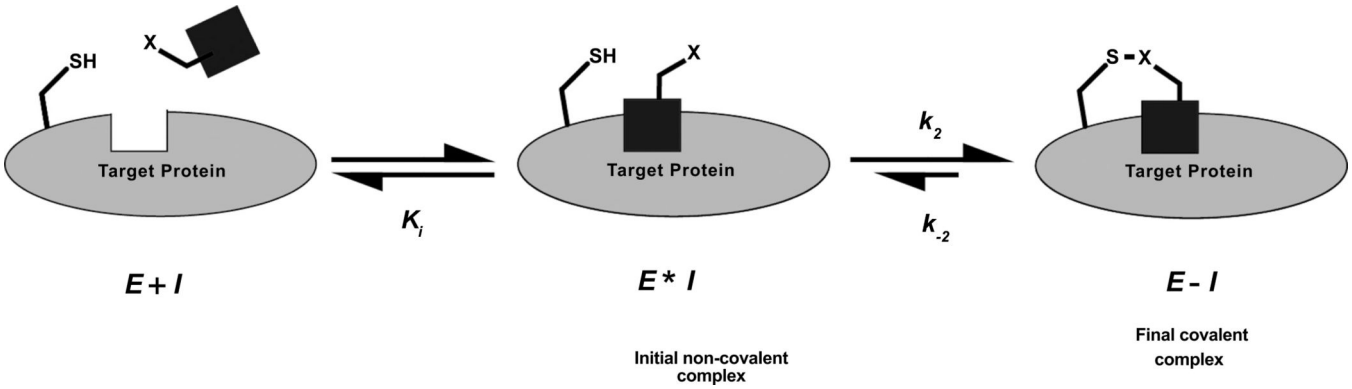

Fig. (1).

Schematic of Covalent Enzyme Inhibition. An initial non-covalent binding event brings the cysteine sulfhydryl group in proximity to the warhead X, driving covalent bond formation. The covalent bond can be irreversible (e.g vinyl sulfonamide), where k_{-2} is zero, reversible (e.g. another sulfhydryl group) where k_{-2} depends on reaction conditions or slowly reversible (e.g. cyanoacrylamide), where k_{-2} depends on warhead reactivity.

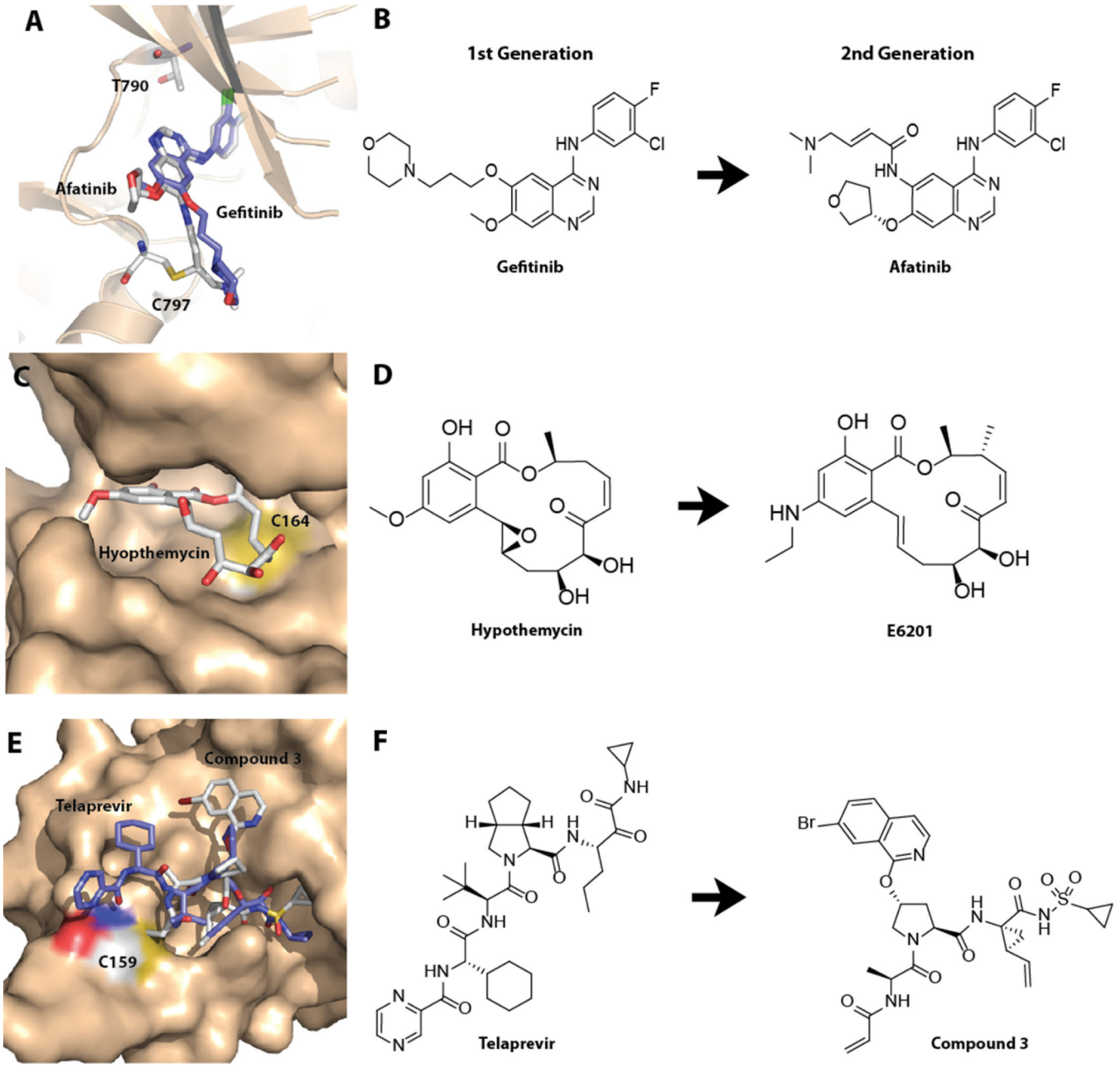

Fig. (2).

Deriving Covalent Inhibitors from Known Scaffolds. A–B) Non-covalent 1st generation EGFR inhibitor Gefitinib (purple) overlaid with covalent 2nd generation inhibitor Afatinib (white). T790, the eventual site of resistance mutation, is shown. B–C) Natural product Hypothemycin bound to ERK2 and synthetic drug candidate E6201. E–F) Telaprevir (purple) overlaid with covalent derivative Compound 3 (white), which binds to conserved Cys159.

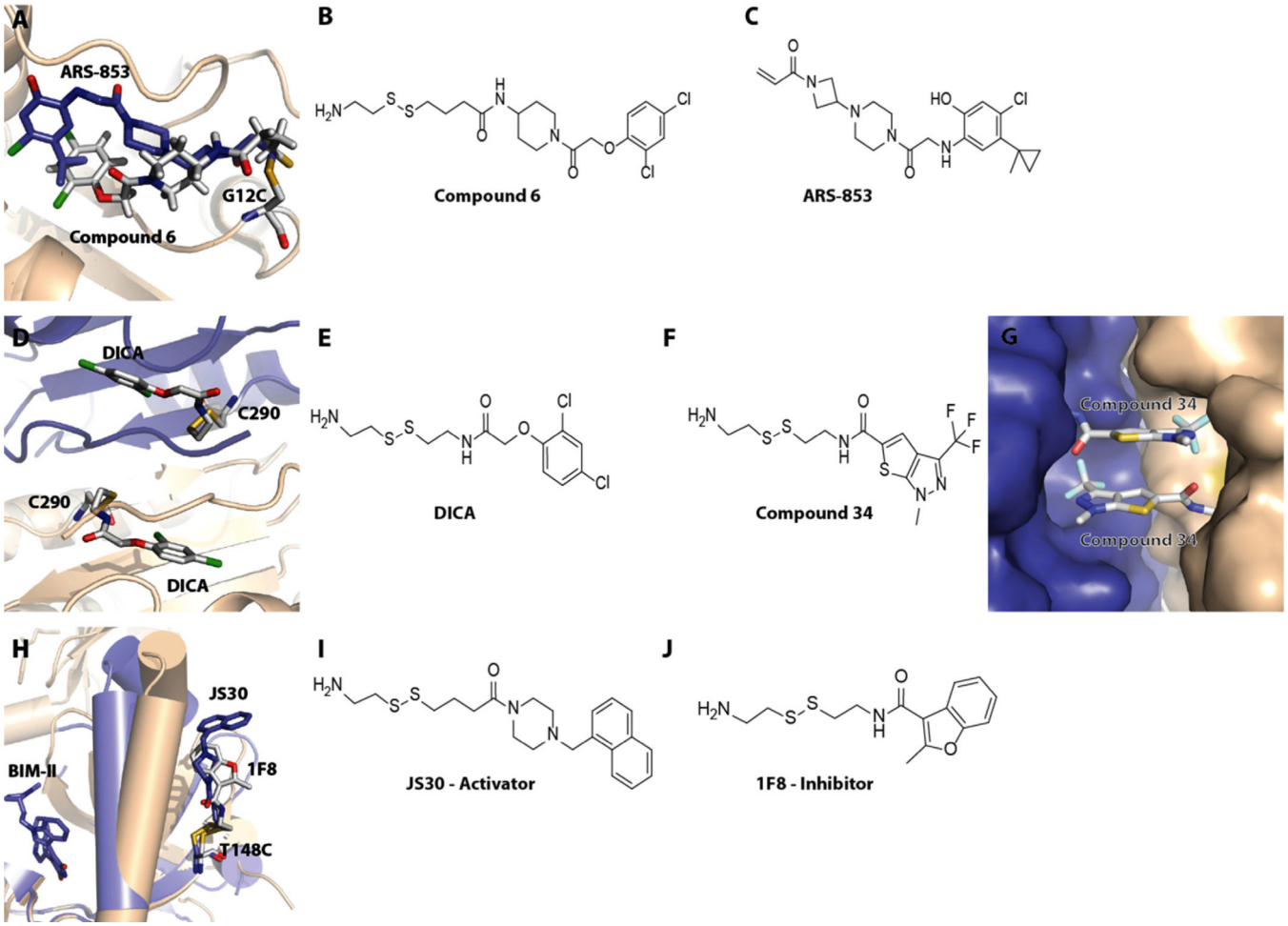

Fig. (3).

Applications of Disulfide Tethering. A–C) K-Ras G12C mutation targeted by an optimized tethering hit (white; B), validating a new pocket; chemical optimization of electrophilic analogs led to an irreversible inhibitor (purple; C) with nM potency. D,E) Caspase-1 zymogen dimer (monomers colored tan/purple) bound to two DICA molecules at Cys290. F,G) Caspase-7 dimer (monomer surface colored tan/purple) bound to two interacting copies of Compound 34 (F). H–J) PDK1 bound to activator JS30 (purple; I) and inhibitor 1F8 (white; J) at the PIF-pocket. The activator shifts the regulatory C-helix down toward the active site where an active-site inhibitor BIM-II is bound (purple).

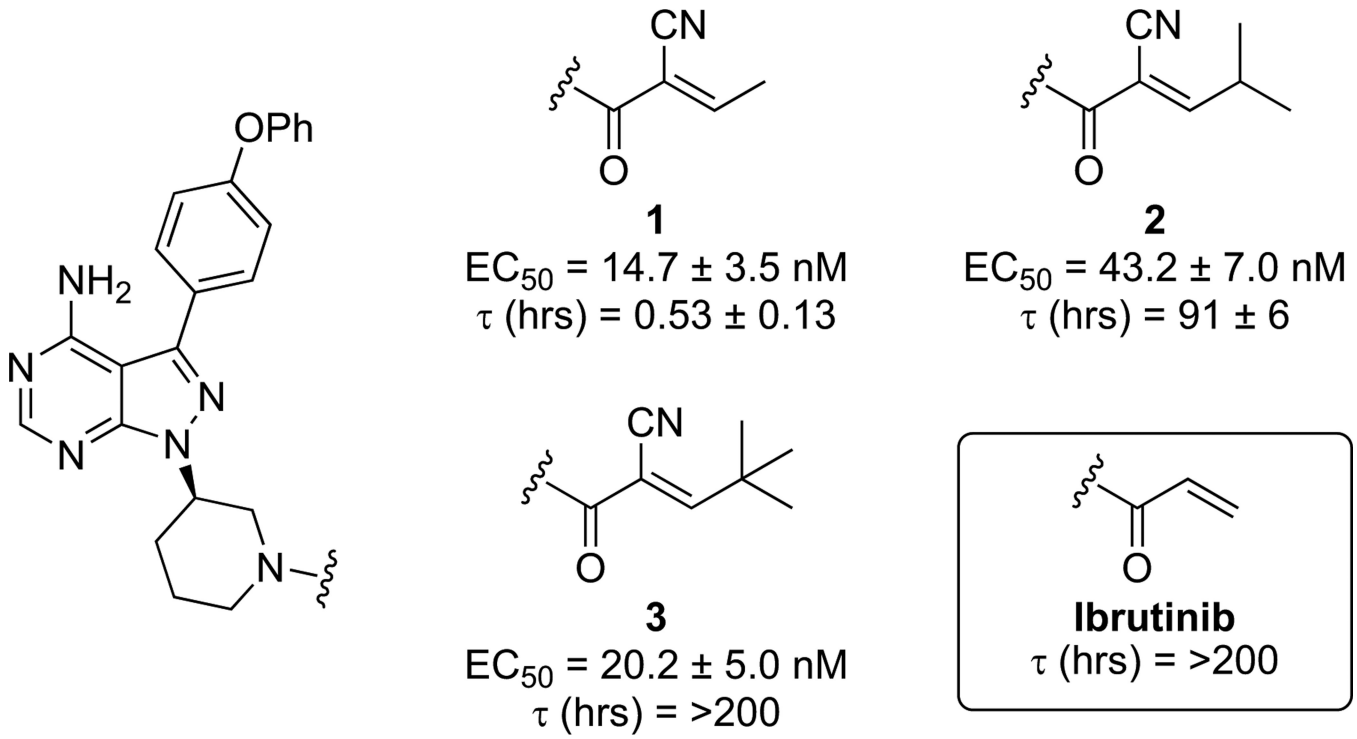

Fig. (4).

Slowly Reversible Cyanoacrylamide Inhibitors of BTK. The core Ibrutinib scaffold was systematically substituted with three cyanoacrylamide warheads: **1** Me, **2** *i*-Pr, and **3** *t*-Bu. Compounds **1–3** had similar EC_{50} values, but vastly differing residence times (**τ**) that correlated with increasing steric demand (*t*-Bu > *i*-Pr > Me).