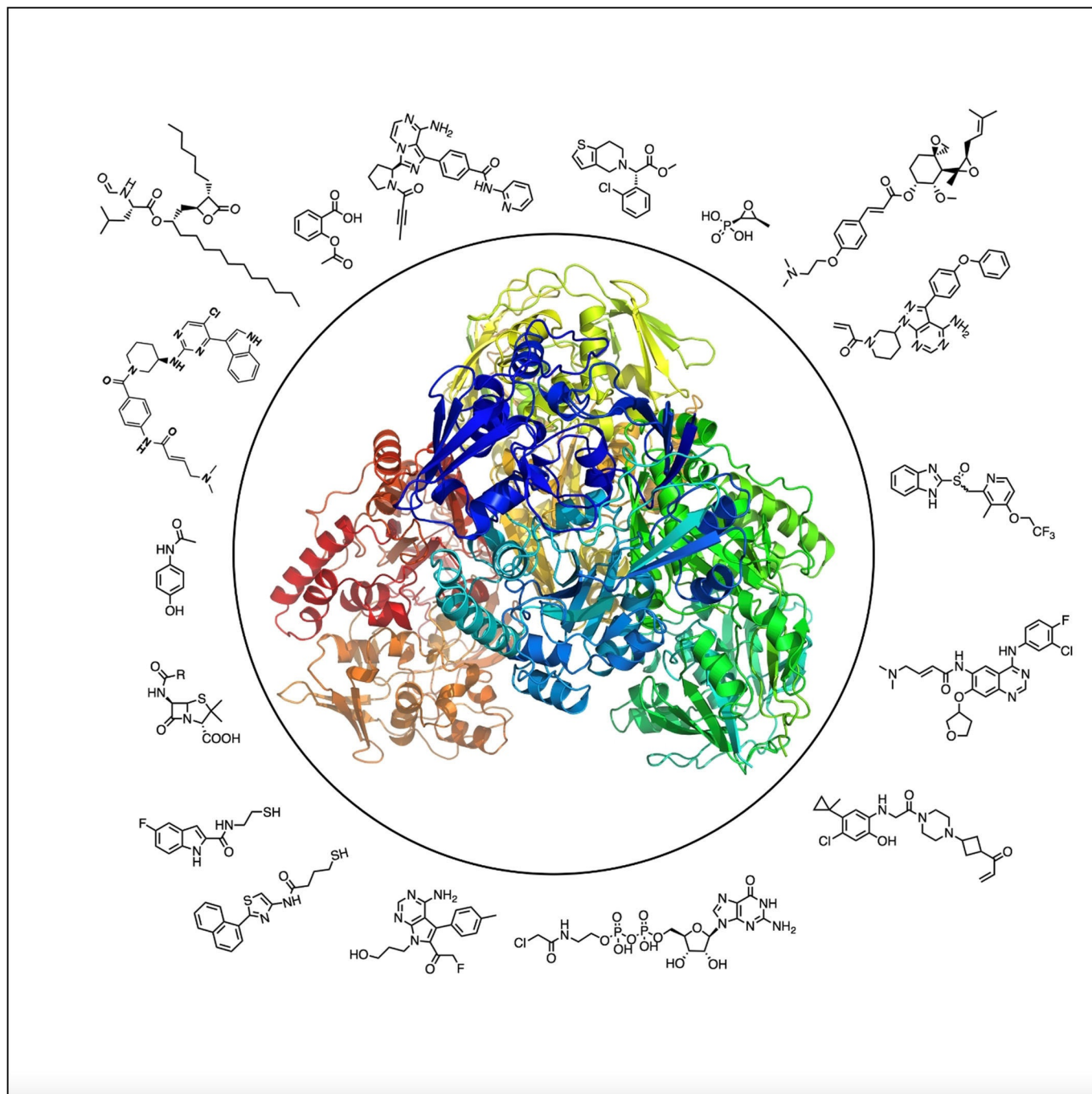


Covalent Inhibition in Drug Discovery

Avick Kumar Ghosh⁺, Indranil Samanta⁺, Anushree Mondal⁺, and Wenshe Ray Liu^{*[a]}



Although covalent inhibitors have been used as therapeutics for more than a century, there has been general resistance in the pharmaceutical industry against their further development due to safety concerns. This inclination has recently been reverted after the development of a wide variety of covalent inhibitors to address human health conditions along with the US Food and Drug Administration (FDA) approval of several covalent therapeutics for use in humans. Along with this exciting

resurrection of an old drug discovery concept, this review surveys enzymes that can be targeted by covalent inhibitors for the treatment of human diseases. We focus on protein kinases, RAS proteins, and a few other enzymes that have been studied extensively as targets for covalent inhibition, with the aim to address challenges in designing effective covalent drugs and to provide suggestions in the area that have yet to be explored.

1. Introduction

Covalent inhibitors are usually small molecules that bind to enzymes and inactivate them temporarily or permanently. In general, covalent inhibition is a two-step process.^[1] First, an inhibitor reversibly associates with the target enzyme, thus bringing the chemical warhead of the inhibitor within close proximity of a targeted reactive residue in the enzyme. In the second step, reaction occurs between the two reactive entities of the inhibitor and the enzyme to form a covalent bond. Reversible inhibitors differ from covalent inhibitors in that they do not involve the second step. A covalently bound inhibitor may undergo further chemical transformation(s) to get released from the

target enzyme after a certain period of time. It may also permanently bind to the target, causing the enzyme to be locked in an inactive state. The use of small molecules as covalent inhibitors to target functionally critical enzymes in cells has been implemented since the late 19th Century, when Bayer started manufacturing acetylsalicylic acid (ASA, aspirin) as a painkiller and anti-inflammatory drug.^[2] Although it has been on the market since the beginning of the 20th Century, its mechanism of action was not revealed until the 1970s when Roth et al. showed that aspirin irreversibly inhibits cyclooxygenase-1 (COX-1), an enzyme that plays an instrumental role in the biosynthesis of prostaglandin.^[3] Upon interacting with COX-1, aspirin acetylates an active site serine residue and thus inacti-

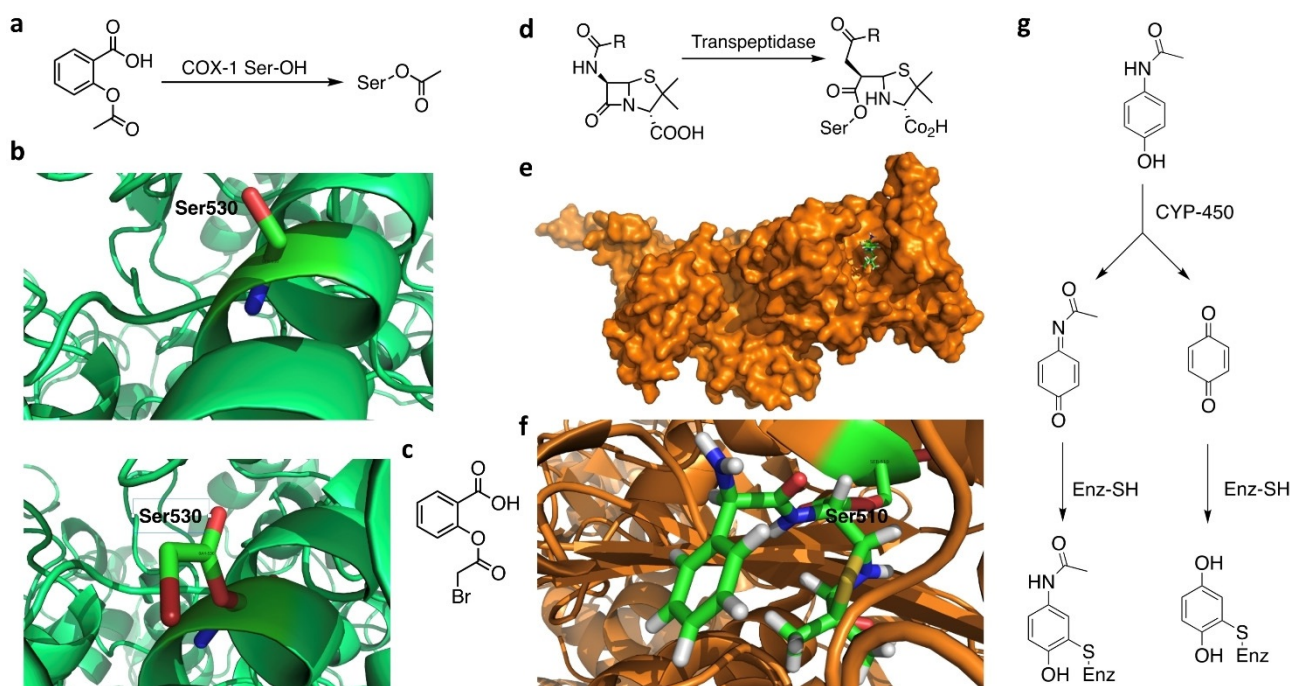


Figure 1. Early covalent drugs. a) The mechanism of action of aspirin. b) COX-1 active site serine and its covalent adduct after its reaction with c) a bromo-derivative of aspirin. d) The mechanism of action of penicillin. e) Co-crystal structure of DD-transpeptidase bound covalently with ampicillin (PDB ID: 5HL9). f) The ampicillin–serine complex in the active site of DD-transpeptidase. g) The mechanism of action of acetaminophen.

[a] A. K. Ghosh,[†] I. Samanta,[†] A. Mondal,[†] Prof. Dr. W. R. Liu
Department of Chemistry, Texas A&M University, Corner of Ross and Spence
Streets, College Station, TX 77843 (USA)
E-mail: wliu@chem.tamu.edu

[[†]] These authors contributed equally to this work.

ⓑ The ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/cmdc.201900107>.

vates COX-1 (Figure 1a–c).^[4] Besides aspirin, acetaminophen was also discovered later in the 19th Century and was quickly introduced into medical practice as a painkiller. Although its mechanism of action has not been clearly defined, the electron-rich characteristics of acetaminophen makes it prone to oxidation, giving rise to quinone-like structures. These qui-

none-like molecules are susceptible to attack by nucleophilic protein/enzyme residues that may result in inhibition of proteins/enzymes.^[5] Therefore, acetaminophen can also be considered as a covalent inhibitor. Another early covalent drug is

Avick K. Ghosh was born in West Bengal, India in 1993. He completed his BS degree from Hooghly Mohsin College with honors in chemistry, and his MS degree from the Indian Institute of Technology Kharagpur. He is currently pursuing his PhD at Texas A&M University under the supervision of Prof. Tadhg P. Begley, where his research focuses on the identification and mechanistic enzymology of novel cofactor catabolism pathways.



Indranil Samanta received his BS degree in chemistry in 2014 from Ramakrishna Mission Vidyamandira, West Bengal, India, and obtained his MS degree in chemistry in 2016 from the Indian Institute of Technology Kanpur. He is currently pursuing his PhD at Texas A&M University under the guidance of Prof. Tadhg P. Begley, where his research focuses on thiamin-dependent proteins and flavin-dependent enzymes.



Anushree Mondal was born in West Bengal, India in 1994. She completed her BS degree with honors in chemistry at Bankura Christian College, West Bengal, India, and obtained her MS degree from the Indian Institute of Technology Kharagpur. She is currently pursuing her PhD at Texas A&M University in the chemical biology division under the supervision of Prof. Tadhg P. Begley, where her work focuses on vitamin biosynthesis pathways.



Wenshe Ray Liu was born in Anhui Province, China in 1977. He completed his BS degree at Peking University and his PhD at the University of California, Davis. After postdoctoral work at the Scripps Research Institute, he started his independent career at Texas A&M University in 2007. He is now a full professor, the current holder of the Gradipore Chair in chemistry, and the director of the Texas A&M Drug Discovery Center.



penicillin. The serendipitous identification of penicillin as an antibiotic can be considered as one of the most significant findings in the history of drug discovery. To date, a number of penicillin analogues have been approved for use in human patients. Together with penicillin, they all share a similar mechanism of action and contain a β -lactam as the chemical warhead. This β -lactam reacts with an active site serine residue in D-Ala-D-Ala transpeptidase, which is involved in bacterial cell wall biosynthesis. Its inactivation leads to disruption of bacterial cell wall biosynthesis and consequently bacterial cell lysis (Figure 1 d–f).^[6]

Other covalent antibiotics include some β -lactamase inhibitors such as avibactam,^[7] thienamycin,^[8] and cephalosporin.^[9] Although covalent inhibitors have long been used to treat human health conditions, the idea of covalent inhibition was not very popular until the 1990s, as many covalent drug metabolites were found to have negative effects. For example, cellular metabolites of acetaminophen were found to be hepatotoxic.^[10] During the metabolism of acetaminophen, it is oxidized by cytochrome P450 to highly reactive quinone intermediates (NAPQI and benzoquinone), which react with either glutathione (GSH) or the sulfhydryl group of cysteine residues present in proteins (Figure 1 g) for covalent modifications.^[5] Nonspecific covalent drug–protein adducts could lead to unwanted immunogenic responses in patients. Despite these drawbacks, several factors have revived interest in the development of covalent drugs as therapeutic agents. First are the cases of successful covalent drugs on the market, such as aspirin and penicillin. Secondly, not every covalent drug becomes toxic after undergoing metabolic activation. Moreover, many natural products are covalent inhibitors. Covalent inhibitors also display advantages over reversible inhibitors, such as strong target affinity and prolonged activity in patients. Therefore, undesirable pharmacokinetic properties can often be tolerated, as the pharmacodynamic properties of these inhibitors can outlast measurable inhibitor concentrations in the plasma. Based on these observations, it has been concluded that if the reactivity of the warhead of a covalent inhibitor can be controlled, critical concerns of its use as a therapeutic can be minimized. For these reasons, the development of covalent inhibitors has thrived in the past decade. A brief overall history of covalent inhibitor drug discovery is outlined in Figure 2.^[1,2,5]

In general, covalent inhibitors contain electrophiles that react with nucleophilic residues in enzymes. To date, molecules with a variety of electrophilic warheads including epoxide, aziridine, ester, ketone, α,β -unsaturated carbonyl, and nitrile have been identified as covalent inhibitors.^[11] A new inhibitor class known as “sulfur tethers” have also caught attention due to their ability to covalently conjugate cysteines in enzymes. To design a covalent inhibitor for a given enzyme target, three steps are typically involved: First, structural analysis of the target provides crucial information regarding which nucleophile (e.g., cysteine, serine) is present at or around a potential binding pocket. The nucleophile needs to be unique in that protein family, otherwise there will be low selectivity. Second, a reversible inhibitor with some potency (IC_{50} in the micro- to millimolar range) is found for which the binding mode and in-

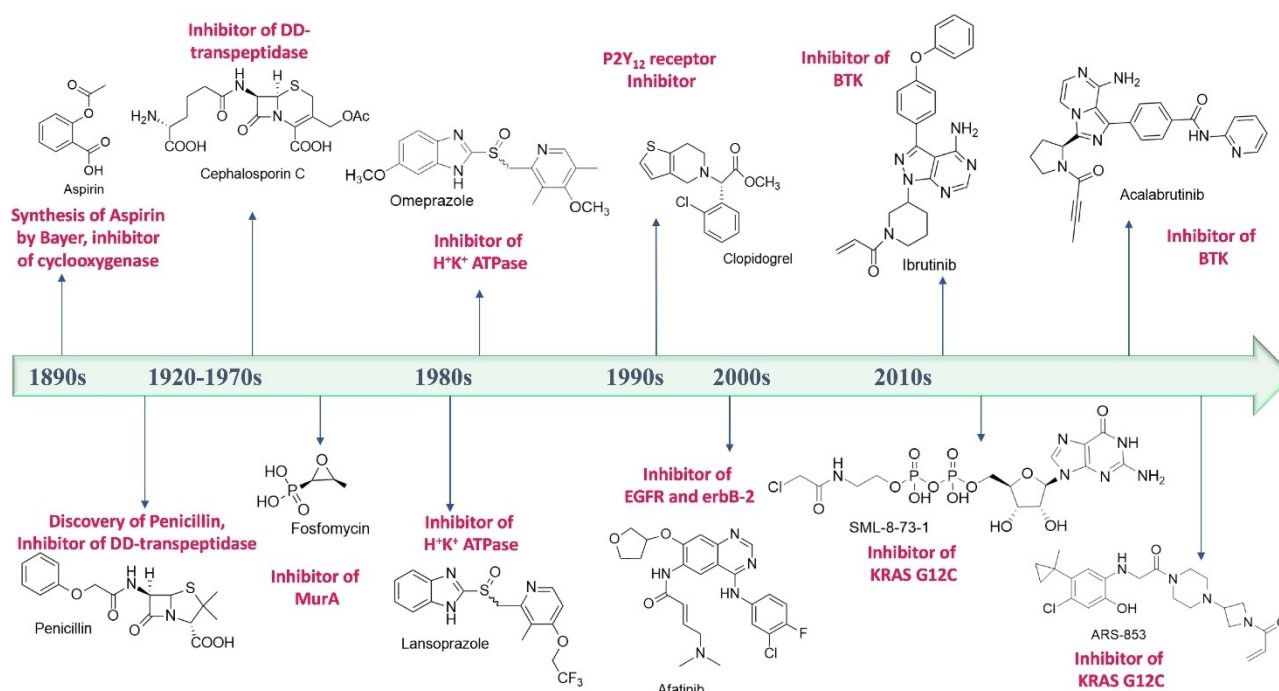


Figure 2. A brief timeline of covalent drug discovery. Structures of covalent inhibitors are provided along with the enzymes/proteins they inhibit.

teractions are known. This inhibitor can come from archived inhibitors that were developed for related enzymes. Finally, an electrophilic “warhead” is positioned in a selected reversible inhibitor to react specifically with the chosen nucleophile in the enzyme target. Isosteric replacement and analogue synthesis are typically used to obtain an active covalent inhibitor candidate. Although a co-crystal structure of an initial noncovalent scaffold bound to an enzyme target will provide more insight for designing a covalent inhibitor, it is not absolutely essential. In many cases, docking the initial noncovalent scaffold to a homology model has proven to be sufficient. Structure–activity relationship studies can be further applied to optimize a lead covalent inhibitor for the identification of more potent compounds. Applications of computational chemistry in the design of covalent inhibitors has been recently reviewed by Awoonor-Williams et al.^[12] In the following sections, we describe advances in targeted enzyme families including protein kinases and RAS proteins for the development of covalent inhibitors and briefly touch on other enzymes. Given the scope of this review, we apologize for not being able to cite other critical work in the covalent drug field.

2. Covalent Inhibitors of Protein Kinases

Although conceptually eye-catching, it is difficult to design a covalent inhibitor because of the challenge in maintaining the right balance of selectivity, reactivity, and efficacy. This is even more critical when targeting protein kinases. Kinases are enzymes that transfer the γ -phosphate group from ATP to small molecules or amino acid residues in proteins. So far, 518 human kinases have been discovered, not to mention that over 900 genes in humans potentially encode kinases.^[13] Al-

though highly diversified in their amino acid sequences, the three-dimensional structures of human kinases are strikingly similar, especially in the catalytically active ATP binding domain. Careful optimization of noncovalent binding affinity and reactivity of electrophiles are crucial for designing covalent drugs for protein kinases, as differences in the active site environment across the kinome are very subtle. On the basis of previously mentioned strategies of drug design, attempts have been made to design targeted covalent inhibitors (TCIs) for kinases. Some of these efforts are summarized below.

2.1. Irreversible covalent inhibitors of epidermal growth factor receptor (EGFR)

EGFR belongs to the receptor tyrosine kinase family which catalyzes protein tyrosine phosphorylation to control signal transduction. EGFR is a cell-surface protein that binds to its natural ligand, epidermal growth factor (EGF), to induce tyrosine autophosphorylation and signals cell proliferation.^[14] Specific mutations in this gene are linked to non-small-cell lung cancer (NSCLC).^[15] The most common are in-frame deletion of exon 19 ($\Delta E746$ – $A750$) and a single point mutation, L858R, in exon 21. These are called “activating mutations”, as they lead to ligand-independent tyrosine kinase activity.^[15] This elevated basal level kinase activity can be inhibited by ATP-competitive reversible drugs such as gefitinib^[16] and erlotinib^[17] (Figure 3a). However, another active site mutation, T790M, decreases their efficacy by at least 50%.^[18] It is believed that this mutation reduces the K_M values for ATP and its analogues, thus increasing the total amount of drug needed to inhibit EGFR T790M efficiently.

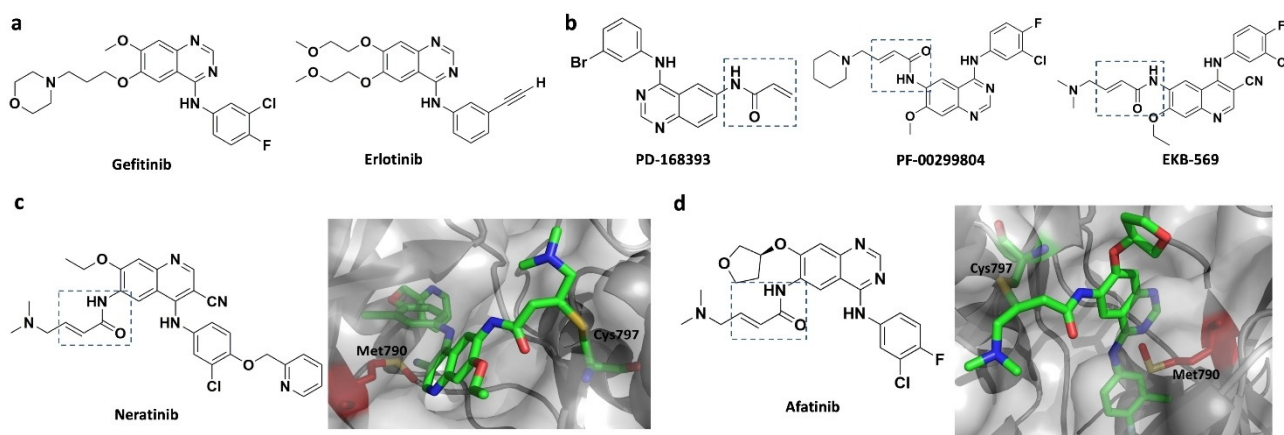


Figure 3. a) Two reversible EGFR inhibitors, gefitinib and erlotinib. b) First-generation irreversible covalent inhibitors of EGFR, PD168393, PF00299804 and EKB569; all contain an acrylamide moiety as an electrophilic warhead (highlighted by a dashed box). c) Neratinib and its complex with EGFR T790M (PDB ID: 2JIV); the gatekeeper residue (Met790) is shown in red. d) Afatinib and its complex with EGFR T790M (PDB ID: 4G5P).

The development of covalent inhibitors for the T790M mutant started in the early 1990s at Parke-Davis and Wyeth (now Pfizer).^[19] Several irreversible inhibitors like PD168393,^[20] PF00299804 (dacomitinib),^[21] and EKB569 (pelitinib)^[22] (Figure 3b) were reported, but in the long run none succeeded in overcoming drug efficacy problems related to the T790M mutation. Unlike erlotinib and gefitinib, these covalent EGFR inhibitors contain the 4-anilinoquinazoline scaffold, which is equipped with an electrophilic warhead, the acrylamide moiety that undergoes Michael addition with the conserved C797 residue in the EGFR active site. The covalent attachment increases the drug action time to inhibit EGFR T790M.^[21] Neratinib (Figure 3c) is another covalent EGFR inhibitor that was thoroughly investigated for counteracting the T790M mutation. Despite showing great promise initially, neratinib exhibits low potency both in TCI-naïve patients and those who have taken TCIs before, probably due to insufficient bioavailability from a diarrhea-imposed dose limit.^[23] Interestingly, afatinib (Figure 3d), an ErbB family blocker, was the first FDA-approved (2013) covalent EGFR inhibitor. Although afatinib alone elongates the progression free survival (PFS) time by almost three-fold relative to patients treated with placebo, combination therapy with cetuximab, a human–murine monoclonal antibody, produced far more convincing results, even in mice with L858R/T790M erlotinib-resistant tumors. Afatinib binds covalently to the ATP binding pocket of EGFR and partially hinders its tyrosine kinase activity. On the other hand, cetuximab induces receptor degradation by blocking ligand binding, but it is not effective enough to fully shut down ligand-independent activity of receptors. Only in the presence of both agents is the depletion level of EGFR so high that it compels mutant EGFRs to decrease the amount of signaling below a certain threshold required for cell survival.^[24] Besides inhibiting the T790M mutant, traditional quinazoline-based inhibitors display similar inhibition of wild-type EGFR, causing side effects such as skin rash and diarrhea.^[23]

To alleviate side effects, Zhou et al. developed a third-generation covalent inhibitor, WZ-4002 (Figure 4a), which shows

high selectivity against EGFR T790M.^[25] The crystal structure (PDB ID: 3IKA) shows that WZ-4002 approaches C797 from a unique angle. Intriguingly, the authors described an interaction between chlorine on the pyrimidine ring and the –SMe group of methionine as a “halogen bond”.^[25,26] It was claimed that this interaction makes the inhibitor more selective toward EGFR (T790M). In November of 2015, the FDA approved TAGRISSTO™ (osimertinib, formerly known as AZD9291 shown in Figure 4b), developed by AstraZeneca, to target EGFR for advanced NSCLC. Preclinical studies indicated IC₅₀ values of 12 nM against the L858R mutant and 1 nM against the L858R/T790M double mutant of EGFR. This drug showed approximately 200-fold greater potency against EGFR L858R/T790M than wild-type EGFR. In vitro cellular phosphorylation along with phenotypic studies and biochemical profiling indicate that osimertinib is highly potent against EGFR^{m+} (EGFR containing “activating mutations”) and T790M positive EGFR mutants with a significantly broad margin of selectivity against wild-type EGFR.^[27] Although Cross et al.^[27] modeled osimertinib in the T790M mutant of EGFR (PDB ID: 4ZAU) and showed that C797 can form an irreversible adduct with it, an actual co-crystal structure of wild-type EGFR with osimertinib does not exhibit a covalent linkage between C797 and the acrylamide moiety.^[28] However, C797 is positioned close enough to the electrophilic warhead so that a slight movement of the loop containing C797 could induce formation of the covalent linkage. Similar to previous-generation drugs, secondary acquired resistance was reported for osimertinib, usually observed after eight or nine months of treatment.^[29] After analyzing the cell-free plasma DNA (cfDNA) collected from patients suffering from osimertinib-resistant advanced lung cancer with EGFR^{m+}, a new C797S mutation was identified.^[30] Although the C797S mutation is the most common mechanism of drug resistance,^[31] there are other mechanisms as well, for example: ErbB2 receptor tyrosine kinase 2 gene (HER2) amplification,^[32] MNG HOS transforming gene (MET) amplification,^[33] an acquired BRAF V600E mutation,^[34] and the EGFR G796D mutation.^[35]

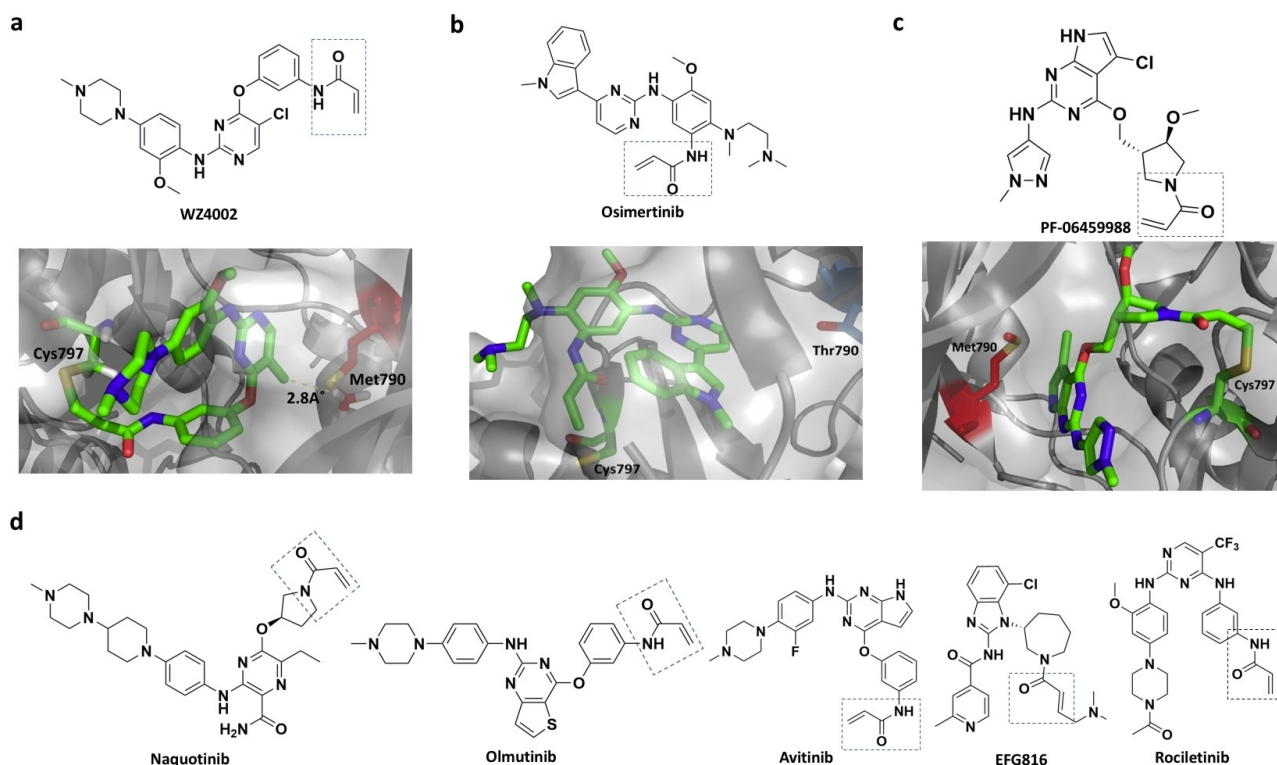


Figure 4. Third-generation irreversible covalent EGFR inhibitors. a) WZ4002 and its complex with EGFR T790M (PDB ID: 3IKA). b) Osimertinib and its complex with wild-type EGFR (PDB ID: 4ZAU). c) PF-06459988 and its complex with EGFR L858R/T790M (PDB ID: 5HG7). d) Other third-generation irreversible covalent EGFR inhibitors.

There are several other third-generation EGFR inhibitors under development by pharmaceutical companies, including PF-06459988 (developed by Pfizer; Figure 4c),^[36] nazartinib (EFG 816; developed by Novartis),^[36,37] naquotinib (ASP8273; developed by Astellas),^[38] olmutinib (HM61713; developed by Hanmi),^[39] and avitinib (AC 0010; developed by ACEA Bioscience).^[40] Most of these compounds are shown in Figure 4d. A highly potent ($IC_{50} = 8$ nM) inhibitor against the osimertinib-resistant L858R/T790M/C797S mutant has also been developed.^[41] Notably, rociletinib,^[42] another third-generation EGFR irreversible covalent inhibitor for EGFR^{m+} developed by Clovis Oncology, exhibited good oral bioavailability (65%) as well as antitumor activity similar to erlotinib and afatinib, while having more than 20-fold greater selectivity over wild-type EGFR. Unfortunately, its development was terminated due to some inconsistency in the published data. Out of 130 patients who were enrolled for phase I/II trials, the response rate, according to RECIST 1.1 (Response Evaluation Criteria in Solid Tumors)^[43] was 59% [95% confidence interval (CI): 45–73];^[44] 20 of 92 patients who were treated with rociletinib suffered from hyperglycemia due to undesirable inhibition of the type I insulin-like growth factor receptor (IGF-IR) and insulin receptor kinases by its metabolite, which could be suppressed by dose reduction. However, in November 2015, Clovis Oncology announced that the preliminary response rate of 59% was based mainly on unconfirmed responses, and the confirmed response rate dropped to 34% in 170 patients treated with an active dose of

625 mg b.i.d. (twice a day) and 28% among 79 patients with an active dose of 500 mg b.i.d.^[45]

2.2. Irreversible covalent inhibitors of Bruton's tyrosine kinase (BTK)

BTK is a non-receptor protein tyrosine kinase that belongs to the Tec family. It plays a key role in the maturation of B cells. Given this unique role, targeting BTK is an established treatment for B cell lymphoma and leukemia.^[46] Mutations of BTK are also implicated in X-linked agammaglobulinemia (XLA), an inherited immunodeficiency disease.^[47] XLA patients show defects strictly in their B cells but not in other immune cells, which makes BTK a potential target for treating XLA.^[48] Ibrutinib (formerly known as PCI32765, shown in Figure 5a), a covalent BTK inhibitor, was first developed by Pharmacyclics LLC and approved by the FDA as a second covalent kinase inhibitor drug. It was approved for the treatment of chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL), and Waldenström's macroglobulinemia.^[49] Ibrutinib inhibits BTK by diminishing autophosphorylation at its Y223 residue. Its acrylamide moiety reacts with C481 to form a covalent bond in the BTK kinase domain, which locks BTK in an allosteric inhibitory state (Figure 5b). Although it succeeds in treating B cell lymphoma and leukemia, ibrutinib does lead to side effects such as bleeding, rash, diarrhea, and atrial fibrillation due to off-target binding to EGFR and other Tec family proteins^[50] that also possess

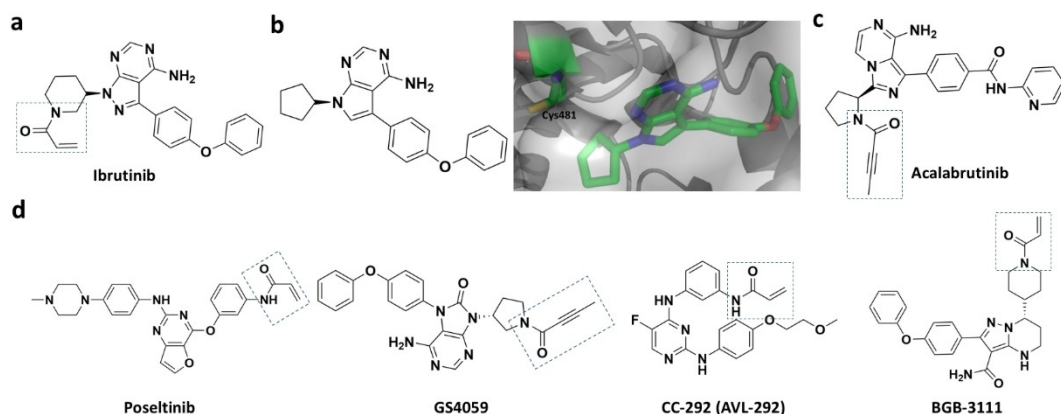


Figure 5. Covalent BTK inhibitors. a) Ibrutinib. b) A pyrrolopyrimidine precursor of ibrutinib and its complex with BTK (PDB ID: 3GEN). c) Acalabrutinib. d) Second-generation irreversible covalent BTK inhibitors.

a similar cysteine residue at the active site.^[51] Resistance to ibrutinib was also observed.^[52] To overcome side effects of ibrutinib, several second-generation BTK inhibitors have been developed.^[53] One of these inhibitors, acalabrutinib (ACP-196, Figure 5c), exhibits high potency, rapid oral absorption, a short half-life as well as decreased binding to EGFR and other Tec family proteins.^[54] Combination of acalabrutinib with other agents (such as obinutuzumab, an anti-CD20 monoclonal antibody) looks promising for the treatment of B cell malignancies.^[55] Wu and co-workers developed CHMFL-BTK-11, another covalent BTK inhibitor, which suppresses the activation of B cells effectively and also blocks the secretion of various cytokines like IgG1, IgG2, and IL-6. It can be potentially used as a drug for rheumatoid arthritis.^[56] Other second-generation inhibitors such as ONO/GS-4059,^[57] BGB-3111,^[53] PRN1008,^[58] and CC-292^[59] (Figure 5d) have also been developed. These inhibitors are in their early clinical testing stage.

2.3. Irreversible covalent inhibitors of p90 ribosomal S6 kinases (RSKs)

The RSK family is a group of highly conserved Ser/Thr kinases that promote cell proliferation, growth, motility, and survival. As they are almost exclusively activated downstream of extracellular signal-regulated kinases 1 and 2 (ERK1/2), therapeutic intervention by RSK inhibition is less likely to produce severe side effects such as those observed following inhibition of the upstream master regulators Raf (rapidly accelerated fibrosarcoma), MEK or MAPK (mitogen-activated protein kinase), and ERK1/2.^[60] The RSK family consists of four closely related paralogues (RSK1–4).^[61] Structurally, RSKs consist of two kinase domains, a C-terminal kinase domain (CTD) that belongs to the Ca^{2+} /calmodulin-dependent protein kinase family and an N-terminal kinase domain (NTD) that belongs to the AGC kinase group [mainly consisting of protein kinase A (PKA), protein kinase G (PKG), and protein kinase C (PKC)]. The two domains are connected by a short linker with several regulatory phosphorylation sites. RSKs are well documented for phosphorylating proteins related to the Na^+/H^+ exchanger NHE1,^[62] the tumor suppressor kinase LKB1,^[63] the translation initiation factor

e1F4B,^[64] etc. Abnormal RSK activity has been implicated in tumor cell invasion,^[65] colonic epithelia,^[66] and endothelial dysfunction.^[42] Sequence alignment of the human kinome has led to the identification of a poorly conserved noncatalytic cysteine in the ATP binding site in RSK1–4 and seven other kinases.^[67] It was also observed that RSK1, 2, and 4 have a “gatekeeper” threonine in the ATP binding pocket.^[68] Inspired by the co-crystal structure of a pyrazolopyrimidine molecule (PP1 in Figure 6a) bound to hematopoietic cell kinase (HCK; Figure 6c), a Src-family kinase with a “gatekeeper” threonine, a fluoromethylketone (FMK)-containing PP1 analogue (Figure 6b) was designed to target RSKs.^[69] Although FMK was supposed to inhibit CTD of RSK1, RSK2, and RSK4, both biochemical ($\text{IC}_{50} = 15 \text{ nM}$) and cellular assays ($\text{EC}_{50} = 300 \text{ nM}$) indicated that it only inhibits RSK2 significantly.^[70] With the aid of computer simulations, it was expected that the poorly conserved C436 in RSK2 would take a similar position as that of V281 in HCK and react with FMK (Figure 6c). When the “gatekeeper” threonine residue T493 in RSK2 is mutated to methionine or C436 is mutated to valine, it confers resistance. Quite unexpectedly, FMK was found to be extremely potent against a large panel of kinases.^[71] In 2018, Gothelf, Nissen, et al. demonstrated that dimethyl fumarate acts as an allosteric covalent inhibitor of RSK2/mitogen- and stress-activated kinases (MSKs), and further

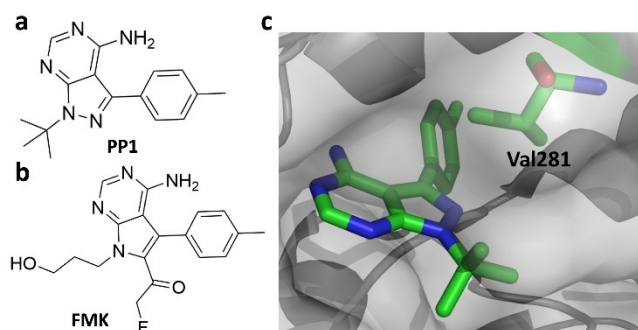


Figure 6. a) PP1, a Src family kinase inhibitor. b) FMK, an irreversible covalent inhibitor of RSK2. c) Co-crystal structure of the Src family kinase HCK bound to PP1 (PDB ID: 1QCF) with N2 properly oriented to Val281 (in green) which corresponds to C436 in RSK2.

studies are being pursued to improve the selectivity toward RSK2.^[72] To the best of our knowledge, no other irreversible inhibitor of RSK1 and RSK4 has been developed yet.

2.4. Irreversible covalent inhibitors of other kinases

In accordance with the increased abundance of structural information of different kinase families, there has been an exponential growth of interest in developing irreversible covalent kinase inhibitors, although the scope of research has been confined mainly in receptor tyrosine kinases. A review by Zhao and Bourne^[73] summarized small-molecule covalent inhibitors against kinases. Not surprisingly, EGFR inhibitors top the list. However, attempts have also been made to develop covalent inhibitors against other kinases. The fibroblast growth factor receptor (FGFR) is a kinase essential for cell proliferation and differentiation. It has been implicated in the development of colorectal, lung, and renal cell cancers and hepatocellular carcinoma.^[74] Human FGFRs consist of four members (FGFR1–4) which can bind to a diverse set of 18 FGF ligands. Zhou et al. developed a first-generation covalent FGFR inhibitor, FIIN-1 (Figure 7 a), that inhibits FGFR1–4 selectively against a library of 402 different kinases, along with some off-target kinases like Flt1, Flt4, and VEGFR.^[75] FIIN-1 features an acrylamide warhead which targets a conserved cysteine residue in all four FGFRs and shows EC_{50} values against FGFR1–3 in a range (≈ 10 nM) similar to those of BGJ398^[76] and AZD4547,^[77] two potent reversible FGFR inhibitors. As mutations at the gatekeeper valine residue (V550 in FGFR4) in FGFRs induced strong resistance to AZD4547 and FIIN-1, Tan et al. developed FIIN-2 and FIIN-3 to

overcome these mutations (Figure 7 a).^[78] FIIN-3 shows potent inhibition against EGFR as well. Using structure-guided design, Brameld et al. developed another irreversible FGFR inhibitor, PRN1371 (Figure 7 b). This inhibitor is undergoing phase I clinical trials for the treatment of solid tumors.^[79] Recently, Novartis developed FGF401, an aldehyde-based covalent inhibitor of FGFR4 which can target a different non-conserved cysteine, Cys552 of FGFR4.^[80]

Covalent inhibitors that target cyclin-dependent kinases (CDKs)^[81] have also been reported recently. Anscombe et al. developed NU6300 (Figure 7 c), an irreversible covalent inhibitor of CDK2, from a known reversible inhibitor NU6102^[82] by positioning a unique vinyl sulfone electrophile proximal to the residue K89, which is situated just outside the ATP binding cleft.^[83] It achieves high selectivity, as K89 is not conserved in closely related kinases. Kwiatkowski et al. were able to target a remote cysteine residue located outside of the canonical kinase domain of CDK7 by THZ1 (Figure 7 d), which bears a cysteine-reactive acrylamide moiety.^[84] Sequence alignment, crystal structure, and biochemical assays indicated that CDK12 and CDK13 also possess C-terminal cysteines that can be modified by THZ1. Derived from THZ1, Zhang et al. developed THZ531, a potent covalent inhibitor of CDK12 and CDK13 (Figure 7 d) that has been demonstrated not only to decrease the expression of genes associated with DNA damage response and super-enhancer-associated transcription factors, but also to cause substantial loss of elongating and hyperphosphorylated RNA polymerase II.^[85] Another covalent kinase inhibitor worth mentioning is CNX-1351 (Figure 7 e), which inhibits lipid phosphoinositide 3-kinase- α (PI3K α) with $EC_{50} < 100$ nM. It targets

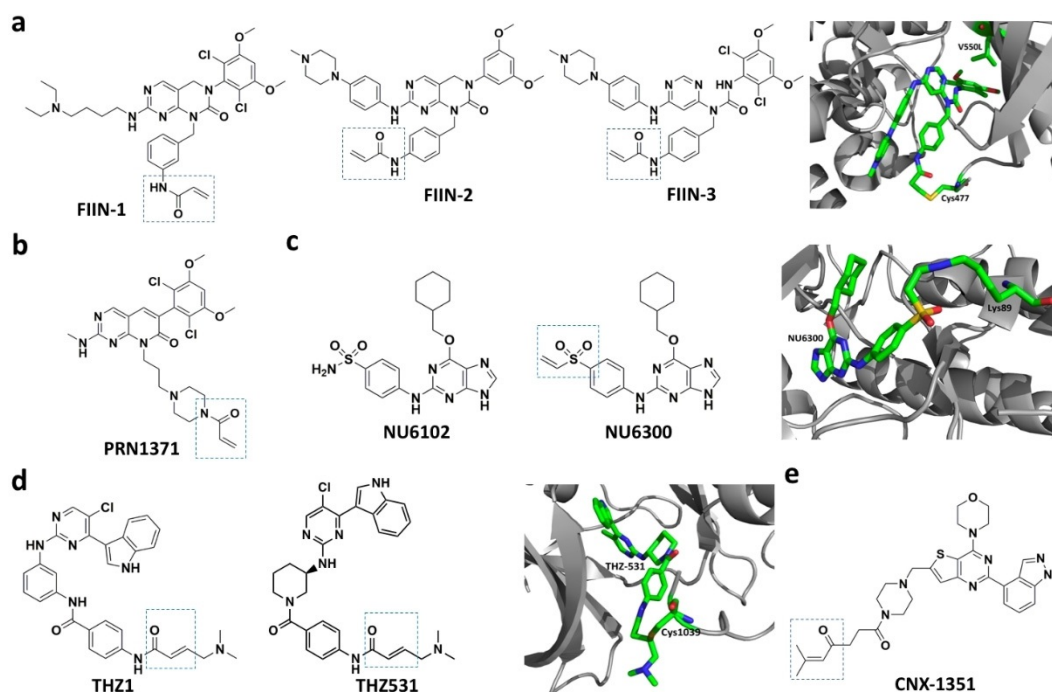


Figure 7. a) The FIIN series inhibitors of FGFR and the co-crystal structure of FIIN3 complexed with FGFR V550L (PDB ID: 4R6V). b) PRN1371, a covalent FGFR inhibitor. c) Noncovalent CDK2 inhibitor NU6102, its corresponding covalent CDK2 inhibitor NU6300, and a co-crystal structure of CDK2 bound covalently with NU6300 (PDB ID: 5CYI) via K89. d) Covalent CDK7 inhibitor THZ1, THZ531, noncovalent CDK12 and CDK13 inhibitor THZ531, and a co-crystal structure of CDK12 bound covalently with THZ531 (PDB ID: 5ACB). e) PI3K α inhibitor CNX-1351.

the residue C862 that is unique to the α isoform, thus providing specificity over PI3K β , $-\gamma$, and $-\delta$ isoforms.^[86]

2.5. Reversible covalent inhibitors of protein kinases

Given the difficulty in predicting the long-term effects of locking a kinase permanently in an inactive state, the development of inhibitors that covalently but reversibly bind to the kinase undoubtedly bears significant merit. It was previously known that the Michael addition reaction between a thiol and a 2-cyanoacrylate is reversible at neutral pH.^[87] Based on this observation, Taunton and co-workers successfully developed highly reactive 2-cyanoacrylate-based reversible covalent inhibitors of RSK2 (Figure 8a).^[71] Based on analysis of the binding mode of

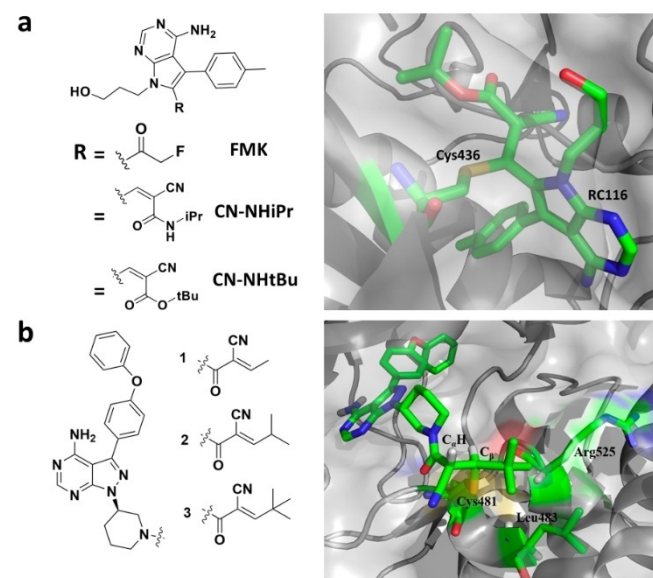


Figure 8. a) Reversible covalent inhibitors CN-NH*i*Pr and CN-NH*t*Bu of RSK2 that were derived from FMN and a co-crystal structure of RSK2 bound covalently with CN-NH*t*Bu at C436 (PDB ID: 4D9U). b) Cyanoacrylamide-based reversible covalent inhibitors of BTK and a co-crystal structure of BTK bound covalently with compound **3** at C481.

FMK, the pyrrolopyrimidine scaffold was modified so that C436 of the CTD of RSK2 can easily reach the electrophile to form a reversible Michael adduct. Several inhibitors have been made; these include CN-NH*i*Pr and CN-NH*t*Bu, which link covalently to C436 in its co-crystal form with RSK2. Both inhibitors target only C436 of RSK2, although there are many solvent-exposed cysteine residues on the kinase. Furthermore, water-soluble thiols like glutathione, even at millimolar concentrations, do not affect the inhibitory action of both reversible covalent inhibitors on RSK2. In contrast, a mutation at C436 easily decimates the inhibitory effect of both inhibitors on the enzyme.^[88] Reversible covalent inhibitors have also been developed for BTK. Bradshaw and co-workers explored the feasibility of designing reversible covalent inhibitors by introducing the cyanoacrylamide moiety in lieu of acrylamide and keeping the core structure of ibrutinib intact.^[89] They noticed that having a unique *tert*-butyl group as a β -substituent not only increases

the residence time of the inhibitor, it also makes it more durable under physiological conditions. The $C\alpha$ proton, which must be abstracted during the elimination of Cys481 thiol (Figure 8b) in the backward reaction, is not oriented properly inside the active site in terms of stereoelectronics to have sufficient kinetic and thermodynamic acidity. Also, both the capping *tert*-butyl group and the linker piperidine shield the $C\alpha$ proton from any nearby base in the active site.^[89] Therefore, a covalent complex form is strongly favored. Although a reversible covalent inhibitor with residence time similar to irreversible inhibitors has several advantages over the latter, including lower toxicity attributed to less permanent off-target adducts, this strategy cannot be generalized, as the branched alkyl capping groups impart high steric demand, which may not be suitable for relatively unhindered and solvent-exposed cysteines.

3. Covalent Inhibitors for Proteins of the RAS Family

For three decades since their discovery, RAS oncogenes have probably been the most studied genes in cancer research. RAS proteins are also among the most mutated oncogenes ever known. To date, no efficient treatment is available for RAS oncogenic mutations. However, recent reports of the development of potent small covalent inhibitors of RAS proteins have revived the hope that these so-called “undruggable” proteins can at last be tuned to be druggable. RAS proteins are a family of related small GTPases, which are expressed in all animal cell lineages and organs. These proteins are involved in transmitting signals within cells. Mutations may permanently activate RAS superfamily proteins, leading to unintended and overactive signaling inside the cell, even if there are no incoming signals. As these signals result in cell growth and division, overactive RAS signaling can ultimately lead to cancer.^[90] In spite of efficacious developments in the ATP-dependent protein kinase family, very similar strategies with respect to RAS proteins have been unsuccessful. This is mainly due to the high picomolar binding interactions of RAS with GTP. In addition, the smooth spatial arrangements of RAS protein structures have hindered the search for small-molecule covalent inhibitors, thus rendering RAS proteins as “undruggable”.^[91] There are three RAS genes in humans: HRAS (Harvey rat sarcoma viral oncogene homologue), KRAS (Kristen rat sarcoma viral oncogene homologue), and NRAS (neuroblastoma RAS viral oncogene homologue). They are the most common oncogenes in human cancer.^[92] In total, there are four RAS isoforms including HRAS, NRAS, KRAS4A, and KRAS4B due to alternative splicing of KRAS during transcription.^[93] The N-terminal domains of all four RAS proteins share 92–98% sequence similarity. The C-terminal regions that consist of 23–24 residues vary dramatically and are therefore called hypervariable regions (HVRs). HVRs contain a C-terminal CAAX box, which, for all four RAS proteins, undergoes farnesyltransferase-catalyzed prenylation at the cysteine residue.^[92] This is also the key post-translational modification of RAS proteins for RAS activation.^[94] Much effort in previous anti-RAS drug discovery was the development of strategies

that prevent RAS activation to associate with plasma membrane. This led to exhaustive efforts in the 1990s to build libraries of farnesyltransferase inhibitors (FTIs).^[95] Although there were some promising results in preclinical studies, the results of clinical experiments with FTIs have not been encouraging. Some of these inhibitors, including lonafarnib and tipifarnib, work well for HRAS but not for KRAS and NRAS. This disappointing outcome dampened the interest in inhibiting RAS proteins based on the intervention of plasma membrane association, though lonafarnib is still under development as a drug for the genetic disorder progeria.^[94] Very recently Shokat et al. reported a series of substrates for farnesyltransferase that can stop the alternative prenylation by geranylgeranyltransferase and mis-localize oncogenic KRAS in cells.^[96] The development of other strategies for RAS inhibition is essential.

3.1. Oncogenic mutations in RAS proteins

The RAS proteins share a common mechanism of activation and downstream signaling which have been well characterized. There are four main regions surrounding the nucleotide binding pocket of RAS proteins: the phosphate binding loop (P loop; residues 10–17), switch I (residues 30–38), switch II (residues 60–67), and the base binding loops (residues 116–120 and 145–147). All these regions contribute to nucleotide-dependent interactions. The arrangements of these regions play critical roles in RAS oncogenic mutations.^[92] The most common oncogenic mutations occurring in RAS proteins are at residues 12 and 13 (belonging to the P loop) and residue 61 (switch II). Key residues for mutations in KRAS and NRAS are residue 12 and residue 61, respectively, whereas for HRAS it is almost equally distributed between residues 12 and 61. KRAS mutations are the most common overall and occur mostly in pancreatic, colorectal, and lung cancer. HRAS mutations occur highly in bladder cancer, and mutations of NRAS are mostly associated with melanoma and myeloid leukemia.^[92] Most of the current drug discovery strategies are based on targeting these mutations in RAS subfamily proteins. One method of achieving mutant specificity is through covalent attachment of an inhibitor to the mutated residue itself.^[92]

3.2. Covalent inhibitors for KRAS mutants

Mutations of the two KRAS isoforms occur in 60% of pancreatic, 34% of colorectal, and 16% of lung cancers.^[97] There are three most common sites for mutation in KRAS (residues 12, 13, and 61) which show minimized GTPase-activating protein (GAP)-dependent GTPase activity. Except the mutation G12C, other mutations at residues 12, 13, and 61 decrease the affinity of RAS for its downstream kinase RAF.^[98] Some of these mutants are G12A, G12D, G13D, Q61L, and G12V.^[92] As one of the most common KRAS mutations leading to cancer, the G12C mutation is particularly interesting in the sense that it has an active non-native cysteine residue that can be easily targeted for covalent inhibition without affecting wild-type KRAS. Recently, Shokat and colleagues reported a set of cysteine-reactive inhibitors that interact with the KRAS G12C mutant and

subsequently form covalent adducts with G12C.^[99] Using a disulfide-tethering approach, they first screened a library of disulfide compounds to identify a series of compounds that selectively bind to KRAS G12C. In their experiments, fragments 2E07 and 6H05 showed the greatest degrees of modification (Figure 9a). They also screened carbon-based electrophiles such as acrylamides and vinyl sulfonamides that form irreversible covalent bonds with the G12C residue (Figure 9c,e). Crystal

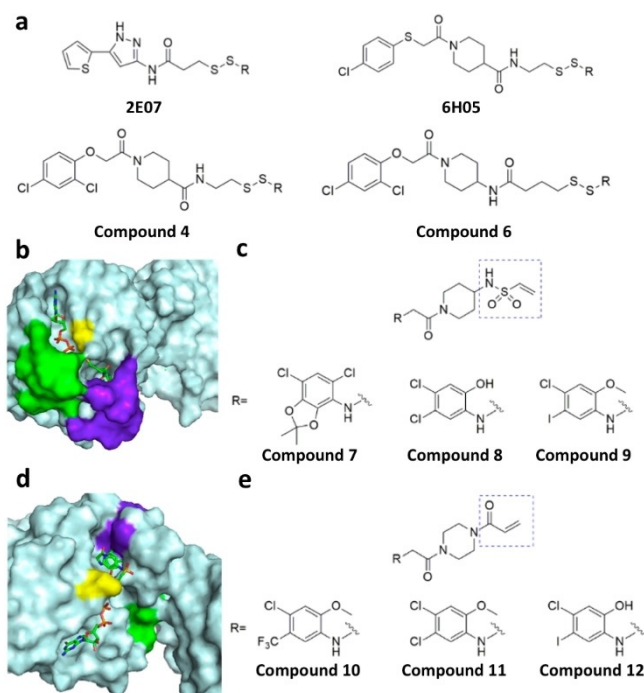


Figure 9. Tethering and electrophilic compounds that selectively bind to oncogenic KRAS G12C. a) Tethering compounds that bind covalently to KRAS G12C. b) Crystal structure of the KRAS G12C complex with compound 6 (PDB ID: 4LUC). c) Electrophiles with the vinyl sulfonamide moiety that bind covalently to KRAS G12C. d) Crystal structure of the KRAS G12C complex with compound 8. e) Electrophiles with the acrylamide moiety that bind covalently to KRAS G12C (PDB ID: 4LYF).

structures of KRAS G12C complexed with these identified molecules showed that these compounds bind to a newly exposed pocket on KRAS G12C, below the effector-binding switch II region. Binding of these compounds to KRAS G12C leads to the nucleotide preference for GDP over GTP and thus results in blocked KRAS signaling. When the structures of KRAS G12C complexed with two kinds of covalent inhibitors are compared, it can be stated that disulfide compounds induce a small shift in the switch II region, leading to slight conformational change in switch I. On the other hand, carbon-based electrophiles showed more prominent effect by displacing the switch II region, which in turn creates greater disorder in switch I.^[99] A series of such compounds along with some of their co-crystal structures with KRAS G12C are shown in Figure 9b,d.^[98]

Gray and colleagues recently recognized that a GTP mimic with a reactive functional group can target the guanine nu-

cleotide (GN) binding pocket, as the natural content of this pocket dictates the signaling state of KRAS.^[100] Using this structure-based design, they synthesized SML-8-73-1 (Figure 10a), a GDP analogue with an electrophilic α -chloroacetamide moiety.

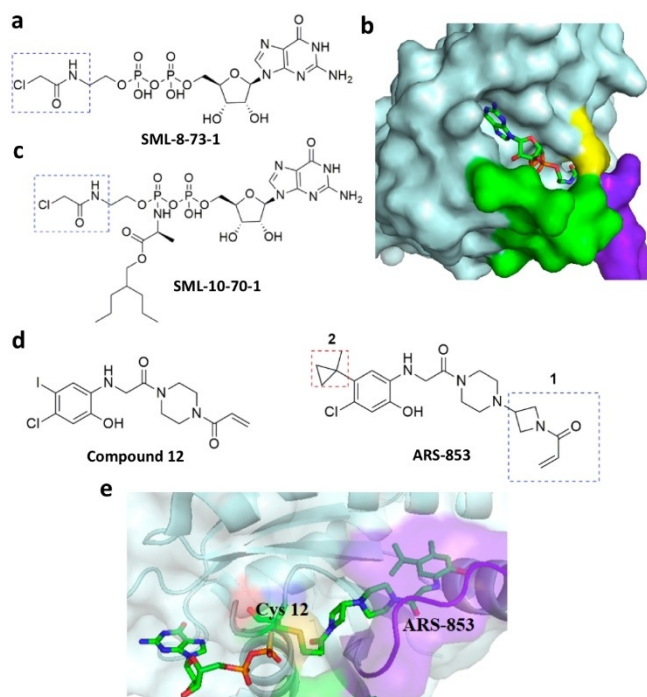


Figure 10. Some other covalent inhibitors of KRAS G12C. a) SML-8-73-1. b) The structure of KRAS G12C bound with SML-8-73-1 (PDB ID: 4NMM). c) SML-10-70-1. d) ARS-853. e) The structure of KRAS G12C bound with ARS-853 (PDB ID: 5F2E).

With help from mass spectrometry analysis, they demonstrated SML-8-73-1 binds covalently at G12C. Hydrogen-exchange mass spectrometry analysis showed that this compound stabilizes the inactive form of KRAS G12C, leading to low affinity for effector proteins. Hunter et al. solved the crystal structure of KRAS G12C bound with this covalent inhibitor (PDB ID: 4NMM).^[99] They showed that SML-8-73-1 binds to KRAS G12C in a manner similar to that of GDP, forming a conformation that is predicted to result in a non-productive state for activating downstream effectors (Figure 10b).^[100] Due to the presence of two negatively charged phosphate groups, SML-8-73-1 cannot pass through the cell membrane. Therefore, it was modified by capping the β -phosphate unit with an alanine ester phosphoramidate (SML-10-70-1; Figure 10c) which penetrates through cell membranes and binds covalently with KRAS G12C.^[32] Patrecelli et al. reported the compound ARS-853 which is a modified version of compound 12 in Shokat's previous study (Figure 10d).^[101] ARS-853 shows robust cellular activity against KRAS G12C in the low micromolar range. Structural and iterative SAR studies based on compound 12 and some of its modified forms showed that the position of the 5-chloro group on the benzene ring is very important for binding. Based on this information, multiple compound 12 derivatives

were synthesized and screened against KRAS G12C to reveal ARS-853 as the most potent inhibitor. While compound 12 has an IC_{50} value around 100 μ M, the IC_{50} value of the much-improved ARS-853 is about 1.6 μ M. The structure of KRAS G12C complexed with ARS-853 is shown in Figure 10e.^[101] Another very recent report by Patrecelli et al. demonstrated the structure-based design and identification of ARS-1620, which is a covalent-mechanism-based inhibitor for KRAS G12C with high potency and selectivity. This compound can rapidly achieve sustained in vivo target occupancy to induce tumor regression. This is an important step toward proving that KRAS can be selectively targeted in vivo by ARS-1620 and also has the scope for promising therapeutic potential.^[102]

New concepts are always emerging in discovering and designing covalent inhibitors for KRAS on a proteome-wide scale. Recent studies by Hansen et al. showed that ARS-853 and ARS-1620 are activated by KRAS-mediated catalysis of the chemical reaction in human KRAS G12C. This biochemical mechanism operates while the reversible binding affinity is weak. This mechanism of action is very similar to how enzymes activate their substrates.^[103]

3.3. Covalent inhibitors of HRAS and NRAS mutants

In comparison with KRAS, whose mutations occur in 21.6% of human cancer, mutations of NRAS and HRAS have relatively low occurrence, with mutations in NRAS associated with 8.0% of human cancer and mutations in HRAS associated with 3.3% of human cancer.^[104] Despite being historically the most investigated, HRAS genes are the least mutated in human cancer among all three RAS genes. As mentioned above, HRAS mutations occur predominantly in bladder cancer, with the key residues for mutation being G12 and Q61. Mutation of G12 in HRAS to any other amino acid except proline induces colony formation and anchorage-independent growth in rat fibroblasts. Although potent HRAS inhibitors have been developed, those that form covalent adducts with HRAS have yet to be identified.^[105]

As of today, NRAS-associated cancers are mostly found in melanoma and in some other cases in lung cancer and T-cell lymphoma. Major mutations associated with NRAS occur at residue 61 along with a few occurring at residues 12 and 13.^[105] All these mutations are known to activate the RAS-RAF-MEK-ERK pathway. NRAS mutants are also reported to activate the PI3K/mTOR signaling cascade.^[105] Posch et al. reported that MEK inhibitors such as JTP-74057 and PD325901 show good inhibition toward the growth of NRAS mutant cells.^[106] Again, in this case the use of combined MEK inhibitors with PI3K or mTOR inhibitors also showed greater efficiency toward decreasing the growth of NRAS mutant cells than using either of them individually. This combination of inhibitors reduces cell viability in vitro and decreases tumor size in vivo in a large panel of human melanoma NRAS cells.^[106] Potent covalent inhibitors for NRAS mutants have also yet to be identified.

4. Covalent Inhibitors of Acetylcholinesterase (AChE)

Acetylcholine is a neurotransmitter that stimulates cholinergic receptors at chemical synapses in the central nervous system. Patients suffering from Alzheimer's disease have decreased levels of these receptors.^[107] One strategy that has been used to combat Alzheimer's disease is to increase the level of acetylcholine at neuronal synapses by inhibiting acetylcholinesterase (AChE), which regulates the level of endogenous acetylcholine. Two reversible drugs, tacrine and donepezil, have been approved by the FDA to treat Alzheimer's disease.^[107] However, they both inhibit AChE for a short time. A novel inhibitor, rivastigmine (Figure 11 a), was recently approved for the treatment of Alzheimer's disease, given its prolonged inhibition of AChE

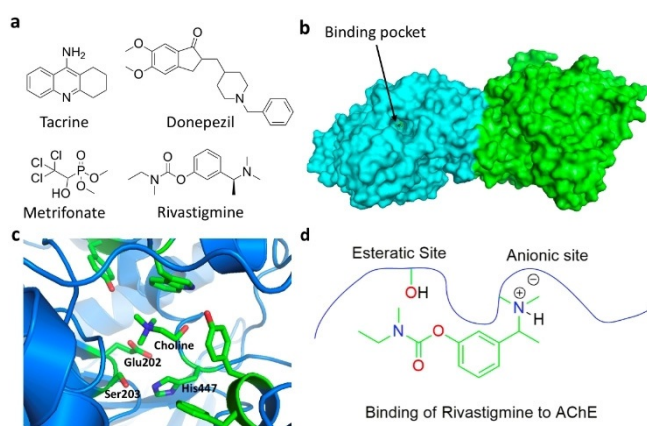


Figure 11. a) AChE inhibitors. b) Structure of mouse AChE. c) AChE with choline bound at the active site (PDB ID: 2HA3). d) Interactions of rivastigmine with AChE in the active site.

for about 10 h.^[108] When it binds to the AChE active site, rivastigmine covalently modifies the protein.^[107] From the crystal structure of AChE bound with choline, as reported by Bourne et al., there are two key residues in the active site, S203 (esteratic site) and E202 (anionic site) (Figure 11 b,c).^[109] The anionic site plays an important role in stabilizing the positively charged nitrogen atom in the native substrate or in rivastigmine, which helps the binding of either the substrate or rivastigmine at the active site. Once rivastigmine binds to AChE, its carbamoyl group is located proximal to the active site serine that exchanges the carbamoyl group from rivastigmine to release a hydrolyzed phenolic derivative (Figure 11 d).^[109] This results in a carbamoyl serine at the active site that slowly hydrolyzes to recover the activity of AChE. Another covalent inhibitor of AChE is metrifonate, which shows even stronger inhibition of AChE.^[110] Metrifonate is a prodrug that is hydrolyzed non-enzymatically to 2,2-dichlorovinyl dimethyl phosphate (DDVP).^[110] DDVP acts as the active agent that binds at the esteratic site of AChE to phosphorylate the active site serine. Hydrolysis of the O–P bond in the modified serine residue is extremely slow, leading to inactivation of AChE for a couple of weeks.

5. Covalent Inhibitors of Cathepsins

Cathepsins are a group of cysteine proteases involved in proteolysis in the lysosome and control various cell signaling pathways.^[111] There are 11 cathepsins coded in humans (B, C, F, H, K, L, O, S, V, W, and X). Hyperactivity of these enzymes is often related to disease development.^[112] Among the different cathepsins, CatK has been of high interest. It is abundant in osteoclasts and plays an important role in resorption and remodeling of bones. As such, it has been a drug target for the treatment of osteoporosis in which bones decrease their density significantly and become fragile. Several covalent inhibitors were designed to target CatK for the treatment of osteoporosis, but have been discontinued owing to their side effects. Among problems during the development of CatK inhibitors, one was that inhibitors identified to show good results on rodent CatK were not so efficient for human CatK.^[113,114] This could be due to species variation of CatK. Another problem arose from attaching basic functional groups to inhibitors. It increased the specificity for CatK but also lysosomal accumulation. One such example is balicatib^[115] (Figure 12 a). Balicatib

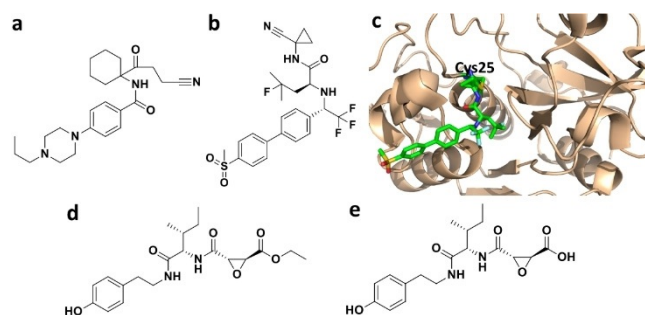


Figure 12. Covalent cathepsin inhibitors. a) Balicatib. b) Odonocatib. c) The covalent adduct between odonocatib and Cys25 in the CatK active site (PDB ID: 5TDI). d) JPM-OEt. e) JPM-565.

showed great selectivity for CatK with respect to other cathepsins, but was discontinued after phase II clinical studies as it led to morphea-like skin lesions in some patients.^[116] The lysosomal accumulation of balicatib gave rise to off-target activities in osteoclasts as well as in skin fibroblasts.^[117] Among many covalent inhibitors of CatK, odonocatib, which is not lysosomotropic, reached phase III trials (Figure 12 b).^[118,119] Odonocatib has a nitrile group that reacts with C25 of CatK to form an iminothioester adduct (Figure 12 c; PDB ID: 5TDI).^[120] Although odonocatib was established as being efficient in increasing bone mineral density and reducing hip or vertebrae fractures, its phase III trial was prematurely terminated. Evidence indicated that it increases the likelihood of cardiovascular complications such as stroke.^[121] As of today, one CatK inhibitor that is undergoing clinical trials is MIV-711.^[122] It is at the phase II clinical trial stage for osteoporosis and osteoarthritis.^[116]

Among other cathepsins, CatS has been found to play a unique role in mediating the immune response in dendritic

and B cells. Hence, the inhibition of CatS can be a useful strategy to combat hyperactivation of immune systems against host antigens in several autoimmune diseases such as rheumatoid arthritis and bronchial asthma.^[123] Moreover, CatS was found to be overexpressed in psoriatic skin, where involvement of chronic antigens in dermal dendritic cells might influence psoriasis.^[116] Efforts have been made to identify noncovalent CatS inhibitors for therapeutic intervention. In 2005, clinical trials of the CatS inhibitor CRA-028129^[124] was launched for the treatment of psoriasis, but was discontinued after phase I. Another CatS inhibitor, RWJ-445380,^[116] which was launched for treating psoriasis and rheumatoid arthritis, suffered the same fate due to its lack of efficacy. Another CatS inhibitor, RO5461111, was used as an orally available drug successfully to suppress clinically advanced lupus nephritis in a mouse model.^[125] The development of SAR113137 as a CatS/CatK inhibitor for pain management was initially discontinued after carrying out safety studies, but is now being re-evaluated for treatment of Chagas disease, as it probably cross-reacts with cathepsin-like proteases in parasites.^[126] Because the upregulated activities of cathepsins are related to cancer development, there has also been effort to use cathepsin inhibitors for treating cancer. Some of these inhibitors contain an epoxide functionality that potentially forms covalent adducts with cathepsins. An epoxide-based inhibitor JPM-OEt (Figure 12d) was tested successfully in treating pancreatic neuroendocrine cancer in a mouse model, but failed when tested against breast-to-bone metastasis and breast cancer in mammary gland.^[116] The reason was found to be the poor bioavailability of the drug. An improved version JPM-565^[127] (Figure 12e) led to antitumor effects similar to gene ablation studies in the same breast cancer model.

6. Covalent Inhibitors of Caspases

Caspases are a family of cysteine proteases that regulate protein cellular homeostasis.^[128] They play essential roles in programmed cell death and inflammation. More than 10 caspases have been identified in humans. Malfunctions of these caspases have been implicated in the development a number of diseases and therefore targeting caspases have been an establishment for disease intervention.^[129] Due to extensive studies of these enzymes, substrate preferences of different caspases have been known for some time, which has been used to design covalent inhibitors for caspases. These inhibitors generally contain a tetrapeptidyl moiety with an electrophilic warhead to react with the active site cysteine of a particular caspase. One covalent inhibitor of caspase-8 is presented in Figure 13a that also shows how it interacts with the active site C360 residue of caspase-8.^[130] This inhibitor has a 2-oxoalkyl *O,O*-dichlorobenzoate moiety that reacts with C360 to form a covalent adduct. A similar strategy has also been applied to design covalent inhibitors for caspase-3 and caspase-6 (Figure 13b). These inhibitors have a 2-oxoalkyl tetrafluorophenyl ether moiety that reacts with a nucleophilic cysteine for releasing tetrafluorophenol, leading to a covalent 2-oxoalkyl cysteinyl ether adduct. The three molecules are under consideration

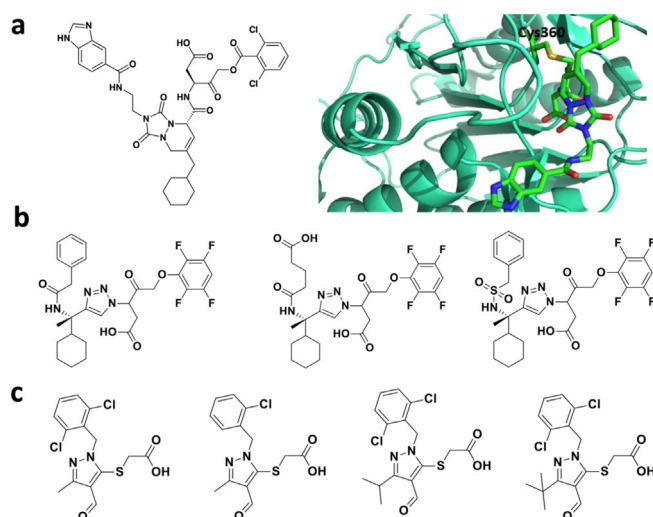


Figure 13. Covalent caspase inhibitors. a) A covalent caspase-8 inhibitor and its caspase-8 adduct structure (PDB ID: 3KJN). b) Three covalent caspase-3 and caspase-6 inhibitors that contain the tetrafluorophenyl ether moiety. c) Caspase-3 inhibitors that are 2-acetic acid derivatives.

for the treatment of Huntington's disease.^[131] Various 2-acetic acid derivatives have also been synthesized and have shown caspase-3 inhibition with micromolar IC_{50} values (Figure 13c). Many other caspase inhibitors are well known in the literature.^[132] Interestingly, a group of thiol-containing compounds were identified to inhibit caspase activities.^[133] Instead of targeting the active site cysteine residue, these compounds, known as "disulfide tethers" bind to an allosteric site. The binding of these "sulfur tethers" usually traps enzymes in their zymogen state, thereby preventing their activation. Compound 1 in Figure 14a is one such "sulfur tether" which binds at the dimeric interface of caspase-7 to form a disulfide linkage with the C290 residue (Figure 14b,c); compound 2 inactivates caspase-3, compound 3 traps caspase-1, and compound 4 inhibits caspase-5.^[134]

7. Covalent Inhibitors of Several Other Enzymes for Treating Obesity and Diabetes

There are also many reports about covalent inhibitors for other enzymes. Inhibitors of three enzymes for treating obesity and diabetes are discussed here. Pancreatic lipase is an enzyme that hydrolyses triacylglycerol fatty acids. It has been implicated as one therapeutic target for treating obesity.^[135] Pancreatic lipase has an active site serine that directly participates in the native hydrolysis reaction. One covalent inhibitor that targets this active site serine is orlistat, which has a β -lactone moiety.^[136] Orlistat (Figure 15) has been approved by the FDA as an oral drug for obesity. Another well-known target to treat obesity is MetAP2. MetAP2 is responsible for cleaving the N-terminal methionine from the newly synthesized protein, which may further undergo post-translational modifications. Although initially it was thought to be a target for cancer treatment, inhibition of MetAP2 was found to be useful to treat obesity. MetAP2 inhibition suppresses sterol regulatory el-

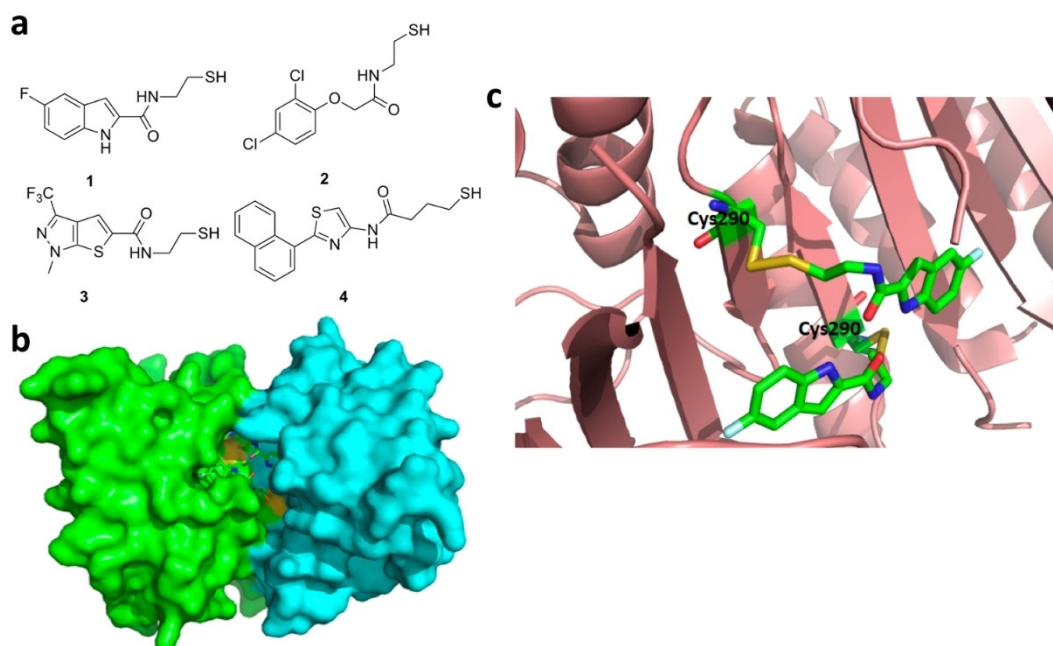


Figure 14. a) Thiol-containing allosteric inhibitors of caspase enzymes. b) Crystal structure of caspase-7 bound with compound 1 (PDB ID: 1SHL). c) An enlarged view at the active site for the structure shown in panel b.

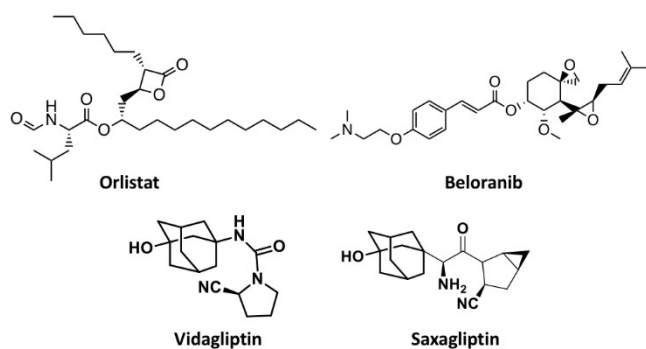


Figure 15. Structures of orlistat, beloranib, vildagliptin, and saxagliptin.

ment binding protein activity, leading to lower levels of lipid and cholesterol biosynthesis.^[137] Beloranib (Figure 15) is a covalent inhibitor that is currently under clinical trials for treating obesity. Mechanistically the less hindered spiro-epoxide of beloranib reacts with residue H231 of MetAP2 at the active site, resulting in inhibition of the enzyme.^[137] Another enzyme that has been targeted by covalent inhibition is dipeptidyl peptidase-4 (DPP4). DPP4 is a serine protease that enhances the body's own ability to control blood glucose by increasing incretin levels and thus regulating insulin and glucagon secretion. As such, DPP4 has been researched as a potent target to treat diabetes mellitus.^[138] Covalent inhibitors of DPP4 has been used to treat type 2 diabetes mellitus (T2DM) by enabling patients to produce their own insulin and control the glucose level in their blood.^[139] Vidagliptin and saxagliptin (Figure 15) are two DPP4 covalent inhibitors that are FDA-approved orally available drugs to treat T2DM and manage glucose levels in diabetic patients.^[139] The warhead electrophilic

center in both these molecules is a nitrile group, which forms a covalent adduct with the active site residue S630 of DPP4.

8. Conclusions

Despite a long history in the treatment of human health conditions, the direct use of covalent inhibitors for human enzymes was not popular in the pharmaceutical industry. Besides potential toxicity and off-target binding, another disadvantage of a covalent inhibitor is the over-dependence of its efficacy on a single residue in the target protein or receptor that can undergo mutation to acquire resistance. The targeted residues for both EGFR and BTK inhibitors have been shown to undergo mutation as a strategy to escape inhibition. For example, resistance was acquired against osimertinib generally within 9–13 months of treatment.^[134c] This acquired resistance raises serious concerns, given the huge investments made during the development of a therapeutic drug. To relieve this potential problem, one strategy is to develop dual inhibitors that target two different enzymes that function in the same disease. For example, a patient suffering from NSCLC can be treated with inhibitors for both anaplastic lymphoma kinase and EGFR.^[140] This strategy may be applied to other diseases. To improve the selectivity and efficacy of a covalent drug, bivalent inhibitors that can target both the active site and an allosteric site might also be considered. It should be noted that allosteric inhibitors may also play a crucial role in combating covalent drug resistance. Several allosteric inhibitors for EGFR and BTK mutants have been developed.^[141] Despite potential problems, the development of covalent inhibitors for human kinases has achieved an excellent start. Greater success is anticipated.

Although researched for more than three decades, targeting RAS proteins for drug development remains a huge challenge.

RAS proteins have very high affinities for GTP, unique protein–protein binding modes, and high dynamic structures, all of which contribute to the difficulty of developing efficient RAS inhibitors. However, recent progress in developing covalent RAS inhibitors has provided some hope to make these traditionally “undruggable” proteins druggable. One driving force for this change is the development of covalent inhibitors for KRAS G12C, although efficient covalent inhibitors for other RAS proteins have yet to be seen.

In summary, although some concerns related to covalent inhibitors remain, the past decade has witnessed a reversal of this trend with a number of covalent inhibitors developed for human enzymes. Several of these covalent inhibitors have already been approved by the FDA for use in human patients. The development of covalent inhibitors for KRAS G12C is especially exciting. Improved potency is a clear advantage of using covalent inhibitors, which could improve the therapeutic window by permitting a decreased dose. Using covalent inhibitors may also lower the risk of idiosyncratic toxicity. Of course, caution still needs to be taken when applying covalent inhibitors as drugs. Clinical trials of several covalent inhibitors have resulted in patient deaths.^[1] These include belorinib. Research teams working on covalent inhibitors must strike a balance between the efficiency of the adduct formation of an inhibitor with its target enzyme and the dose with limited toxicity. Given the increased number of covalent therapeutics approved by the FDA for the treatment of human diseases, we expect more covalent inhibitors will be developed and enter clinical trials in the coming years.

Acknowledgements

Research support in the Liu research group is provided from the US National Institutes of Health (grants R01GM121584 and R01GM127575), the Cancer Prevention & Research Institute of Texas (grant RP170797), and the Welch Foundation (grant A-1715). This review is the term paper of three leading student authors required for fulfilling the TAMU graduate level course Chem 630: Chemical Biology.

Conflict of interest

The authors declare no conflict of interest.

Keywords: acetylcholinesterase • caspase • cathepsin • covalent inhibitors • protein kinases • RAS proteins

- [1] T. A. Baillie, *Angew. Chem. Int. Ed.* **2016**, *55*, 13408–13421; *Angew. Chem.* **2016**, *128*, 13606–13619.
 [2] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discovery* **2011**, *10*, 307–317.
 [3] G. J. Roth, N. Stanford, P. W. Majerus, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 3073–3076.
 [4] a) D. Picot, P. J. Loll, R. M. Garavito, *Nature* **1994**, *367*, 243–249; b) P. J. Loll, D. Picot, R. M. Garavito, *Nat. Struct. Biol.* **1995**, *2*, 637–643.
 [5] T. A. Baillie, *Drug Metab. Rev.* **2015**, *47*, 4–11.

- [6] R. R. Yocum, J. R. Rasmussen, J. L. Strominger, *J. Biol. Chem.* **1980**, *255*, 3977–3986.
 [7] D. E. Ehmann, H. Jahic, P. L. Ross, R. F. Gu, J. Hu, G. Kern, G. K. Walkup, S. L. Fisher, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11663–11668.
 [8] J. S. Kahan, F. M. Kahan, R. Goegelman, S. A. Currie, M. Jackson, E. O. Stapley, T. W. Miller, A. K. Miller, D. Hendlin, S. Mochales, S. Hernandez, H. B. Woodruff, J. Birnbaum, *J. Antibiot.* **1979**, *32*, 1–12.
 [9] A. Yotsuji, J. Mitsuyama, R. Hori, T. Yasuda, I. Saikawa, M. Inoue, S. Mitsushashi, *Antimicrob. Agents Chemother.* **1988**, *32*, 1848–1853.
 [10] D. J. Jollow, J. R. Mitchell, W. Z. Potter, D. C. Davis, J. R. Gillette, B. B. Brodie, *J. Pharmacol. Exp. Ther.* **1973**, *187*, 195–202.
 [11] R. A. Bauer, *Drug Discovery Today* **2015**, *20*, 1061–1073.
 [12] E. Awoonor-Williams, A. G. Walsh, C. N. Rowley, *Biochim. Biophys. Acta Proteins Proteomics* **2017**, *1865*, 1664–1675.
 [13] J. Zhang, P. L. Yang, N. S. Gray, *Nat. Rev. Cancer* **2009**, *9*, 28–39.
 [14] J. Stamos, M. X. Sliwkowski, C. Eigenbrot, *J. Biol. Chem.* **2002**, *277*, 46265–46272.
 [15] A. F. Gazdar, *Oncogene* **2009**, *28*, S24–S31.
 [16] Y. Nakamura, M. Oka, H. Soda, K. Shiozawa, M. Yoshikawa, A. Itoh, Y. Ikegami, J. Tsurutani, K. Nakatomi, T. Kitazaki, S. Doi, H. Yoshida, S. Kohno, *Cancer Res.* **2005**, *65*, 1541–1546.
 [17] C. Schettino, M. A. Bareschino, V. Ricci, F. Ciardiello, *Expert Rev. Respir. Med.* **2008**, *2*, 167–178.
 [18] P. A. Schwartz, P. Kuzmic, J. Solowiej, S. Bergqvist, B. Bolanos, C. Almaden, A. Nagata, K. Ryan, J. Feng, D. Dalvie, J. C. Kath, M. Xu, R. Wani, B. W. Murray, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 173–178.
 [19] J. Singh, E. M. Dobrusin, D. W. Fry, T. Haske, A. Whitty, D. J. McNamara, *J. Med. Chem.* **1997**, *40*, 1130–1135.
 [20] D. W. Fry, A. J. Bridges, W. A. Denny, A. Doherty, K. D. Greis, J. L. Hicks, K. E. Hook, P. R. Keller, W. R. Leopold, J. A. Loo, D. J. McNamara, J. M. Nelson, V. Sherwood, J. B. Smail, S. Trumpp-Kallmeyer, E. M. Dobrusin, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12022–12027.
 [21] J. A. Engelman, K. Zejnullahu, C. M. Gale, E. Lifshits, A. J. Gonzales, T. Shimamura, F. Zhao, P. W. Vincent, G. N. Naumov, J. E. Bradner, I. W. Althaus, L. Gandhi, G. I. Shapiro, J. M. Nelson, J. V. Heymach, M. Meyerson, K. K. Wong, P. A. Janne, *Cancer Res.* **2007**, *67*, 11924–11932.
 [22] N. Yoshimura, S. Kudoh, T. Kimura, S. Mitsuoka, K. Matsuura, K. Hirata, K. Matsui, S. Negoro, K. Nakagawa, M. Fukuoka, *Lung Cancer* **2006**, *51*, 363–368.
 [23] L. V. Sequist, B. Besse, T. J. Lynch, V. A. Miller, K. K. Wong, B. Gitlitz, K. Eaton, C. Zacharchuk, A. Freyman, C. Powell, R. Ananthkrishnan, S. Quinn, J. C. Soria, *J. Clin. Oncol.* **2010**, *28*, 3076–3083.
 [24] Y. Y. Janjigian, E. F. Smit, H. J. Groen, L. Horn, S. Gettinger, D. R. Camidge, G. J. Riely, B. Wang, Y. Fu, V. K. Chand, V. A. Miller, W. Pao, *Cancer Discovery* **2014**, *4*, 1036–1045.
 [25] W. Zhou, D. Ercan, L. Chen, C. H. Yun, D. Li, M. Capelletti, A. B. Cortot, L. Chirieac, R. E. Iacob, R. Padera, J. R. Engen, K. K. Wong, M. J. Eck, N. S. Gray, P. A. Janne, *Nature* **2009**, *462*, 1070–1074.
 [26] M. Z. Hernandez, S. M. Cavalcanti, D. R. Moreira, W. F. de Azevedo, Jr., A. C. Leite, *Curr. Drug Targets* **2010**, *11*, 303–314.
 [27] D. A. Cross, S. E. Ashton, S. Ghiorghiu, C. Eberlein, C. A. Nebhan, P. J. Spitzler, J. P. Orme, M. R. Finlay, R. A. Ward, M. J. Mellor, G. Hughes, A. Rahi, V. N. Jacobs, M. Red Brewer, E. Ichihara, J. Sun, H. Jin, P. Ballard, K. Al-Kadhimi, R. Rowlinson, T. Klinowska, G. H. Richmond, M. Cantarini, D. W. Kim, M. R. Ranson, W. Pao, *Cancer Discovery* **2014**, *4*, 1046–1061.
 [28] Y. Yosaatmadja, S. Silva, J. M. Dickson, A. V. Patterson, J. B. Smail, J. U. Flanagan, M. J. McKeage, C. J. Squire, *J. Struct. Biol.* **2015**, *192*, 539–544.
 [29] T. S. Mok, Y. L. Wu, M. J. Ahn, M. C. Garassino, H. R. Kim, S. S. Ramalingam, F. A. Shepherd, Y. He, H. Akamatsu, W. S. Theelen, C. K. Lee, M. Sebastian, A. Templeton, H. Mann, M. Marotti, S. Ghiorghiu, V. A. Papadimitrakopoulou, A. Investigators, *N. Engl. J. Med.* **2017**, *376*, 629–640.
 [30] K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Janne, G. R. Oxnard, *Nat. Med.* **2015**, *21*, 560–562.
 [31] a) J. J. Chabon, A. D. Simmons, A. F. Lovejoy, M. S. Esfahani, A. M. Newman, H. J. Haringsma, D. M. Kurtz, H. Stehr, F. Scherer, C. A. Karlovich, T. C. Harding, K. A. Durkin, G. A. Otterson, W. T. Purcell, D. R. Camidge, J. W. Goldman, L. V. Sequist, Z. Piotrowska, H. A. Wakelee, J. W. Neal, A. A. Alizadeh, M. Diehn, *Nat. Commun.* **2016**, *7*, 11815; b) H. N.

- [127] G. Mikhaylov, U. Mikac, A. A. Magaeva, V. I. Itin, E. P. Naiden, I. Psakhye, L. Babes, T. Reinheckel, C. Peters, R. Zeiser, M. Bogyo, V. Turk, S. G. Psakhye, B. Turk, O. Vasiljeva, *Nat. Nanotechnol.* **2011**, *6*, 594–602.
- [128] Z. Wang, W. Watt, N. A. Brooks, M. S. Harris, J. Urban, D. Boatman, M. McMillan, M. Kahn, R. L. Heinrikson, B. C. Finzel, A. J. Wittwer, J. Blinn, S. Kamtekar, A. G. Tomasselli, *Biochim. Biophys. Acta Proteins Proteomics* **2010**, *1804*, 1817–1831.
- [129] S. H. MacKenzie, J. L. Schipper, A. C. Clark, *Curr. Opin. Drug Discovery Devel.* **2010**, *13*, 568–576.
- [130] Y. Shi, *Protein Sci.* **2004**, *13*, 1979–1987.
- [131] B. A. Callus, D. L. Vaux, *Cell Death Differ.* **2007**, *14*, 73–78.
- [132] a) H. G. Häcker, M. T. Sisay, M. Gutschow, *Pharmacol. Ther.* **2011**, *132*, 180–195; b) I. N. Lavrik, A. Golks, P. H. Krammer, *J. Clin. Invest.* **2005**, *115*, 2665–2672.
- [133] J. A. Hardy, J. Lam, J. T. Nguyen, T. O'Brien, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12461–12466.
- [134] a) J. Gao, J. A. Wells, *Chem. Biol. Drug Des.* **2012**, *79*, 209–215; b) J. M. Scheer, M. J. Romanowski, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7595–7600; c) X. Gao, X. Le, D. B. Costa, *Expert Rev. Anticancer Ther.* **2016**, *16*, 383–390.
- [135] C. Drahl, B. F. Cravatt, E. J. Sorensen, *Angew. Chem. Int. Ed.* **2005**, *44*, 5788–5809; *Angew. Chem.* **2005**, *117*, 5936–5958.
- [136] R. Guercioli, *Int. J. Obes. Relat. Metab. Disord.* **1997**, *21*, S12–S23.
- [137] A. A. Joharapurkar, N. A. Dhanesha, M. R. Jain, *Diabetes Metab. Syndr. Obes.* **2014**, *7*, 73–84.
- [138] P. R. Flatt, C. J. Bailey, B. D. Green, *Front. Biosci.* **2008**, *13*, 3648–3660.
- [139] A. Barnett, *Int. J. Clin. Pract.* **2006**, *60*, 1454–1470.
- [140] a) J. Jang, J. B. Son, C. To, M. Bahcall, S. Y. Kim, S. Y. Kang, M. Mushajiang, Y. Lee, P. A. Janne, H. G. Choi, N. S. Gray, *Eur. J. Med. Chem.* **2017**, *136*, 497–510; b) A. Rossi, P. Maione, P. C. Sacco, A. Sgambato, F. Casaluce, M. L. Ferrara, G. Palazzolo, F. Ciardiello, C. Gridelli, *Int. J. Oncol.* **2014**, *45*, 499–508; c) N. Karachaliou, R. Rosell, *Cancer Biol. Med.* **2014**, *11*, 173–181.
- [141] a) Y. Jia, C. H. Yun, E. Park, D. Ercan, M. Manuia, J. Juarez, C. Xu, K. Rhee, T. Chen, H. Zhang, S. Palakurthi, J. Jang, G. Lelais, M. DiDonato, B. Bursulaya, P. Y. Michellys, R. Eppele, T. H. Marsilje, M. McNeill, W. Lu, J. Harris, S. Bender, K. K. Wong, P. A. Janne, M. J. Eck, *Nature* **2016**, *534*, 129–132; b) H. Patel, R. Pawara, A. Ansari, S. Surana, *Eur. J. Med. Chem.* **2017**, *142*, 32–47; c) A. R. Johnson, P. B. Kohli, A. Katewa, E. Gogol, L. D. Belmont, R. Choy, E. Penuel, L. Burton, C. Eigenbrot, C. Yu, D. F. Ortwine, K. Bowman, Y. Franke, C. Tam, A. Estevez, K. Mortara, J. Wu, H. Li, M. Lin, P. Bergeron, J. J. Crawford, W. B. Young, *ACS Chem. Biol.* **2016**, *11*, 2897–2907.

Manuscript received: February 18, 2019

Accepted manuscript online: February 27, 2019

Version of record online: March 26, 2019