Linkers in fragment-based drug design: an overview of the literature

Dylan Grenier, Solène Audebert, Jordane Preto, Jean-François Guichou, Isabelle Krimm

Linkers in Fragment-based drug design: an overview of the literature

Dylan Grenier[1#], Solène Audebert[2#], Jordane Preto[1], Jean-François Guichou[2], and Isabelle Krimm[2*]

[#] equally contributed to this work

*corresponding author: isabelle.krimm@univ-lyon1.fr

Author's names and addresses

Solène Audebert

Centre de Biologie Structurale, CNRS, INSERM, Univ. Montpellier, Montpellier 34090, France

Dylan Grenier

Team Small Molecules for Biological Targets, Centre de Recherche en Cancérologie (CRCL) - INSERM 1052 - CNRS 5286 - Centre Léon Bérard - Université Claude Bernard Lyon 1, Institut Convergence Plascan, Lyon, France

Professor Jean-François Guichou

Centre de Biologie Structurale, CNRS, INSERM, Univ. Montpellier, Montpellier 34090, France

Dr Jordane Preto

Team Small Molecules for Biological Targets, Centre de Recherche en Cancérologie (CRCL) - INSERM 1052 - CNRS 5286 - Centre Léon Bérard - Université Claude Bernard Lyon 1, Institut Convergence Plascan, Lyon, France

Dr Isabelle Krimm

Team Small Molecules for Biological Targets, Centre de Recherche en Cancérologie (CRCL) - INSERM 1052 - CNRS 5286 - Centre Léon Bérard - Université Claude Bernard Lyon 1, Institut Convergence Plascan, Lyon, France

Funding

We thank Institut Convergence PLAsCAN (ANR-17-CONV-002), as well as ANR-21-CE18-0014-01 for project ANR-CK2COV, ANR-21-CE29-0012 for project ANR-NanoWAC and ANR-20-CE44-0012 for project ANR-PALUMET for financial support for D.G, S.A and J.P.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

ORCID

Jean-François Guichou https://orcid.org/0000-0002-7699-3235
Jordane Preto https://orcid.org/0000-0002-6041-4008
Isabelle Krimm https://orcid.org/0000-0002-5981-109X

Linkers in Fragment-based drug design: an overview of the literature

Abstract

Introduction: In fragment-based drug design, fragment linking is a popular strategy where two fragments binding to different sub-pockets of a target are linked together. This attractive method remains challenging especially due to the design of ideal linkers.

Areas covered: We review the types of linkers and chemical reactions commonly used to the synthesis of linkers, including those utilized in protein-templated fragment self-assembly, where fragments are directly linked in the presence of the protein. Finally, we detail computational workflows and software including generative models that have been developed for fragment linking.

Expert opinion: We believe that fragment linking offers key advantages for compound design, particularly for the design of bivalent inhibitors linking two distinct pockets of the same or different subunits. On the other hand, more studies are needed to increase the potential of protein-templated approaches in FBDD. Important computational tools such as structure-based de novo software are emerging to select suitable linkers. Fragment linking will undoubtedly benefit from developments in computational approaches and machine learning models.

Keywords

click chemistry; computational chemistry; de novo design software; docking; drug design; fragment; linker; linking; medicinal chemistry; self-assembling.

Article highlights:

- **We review 58 published cases of fragment linking including 40 cases of standard linking and 18 cases of protein-templated fragment self-assembly.**
- **Most represented linkers are flexible and less than 5-atom long, usually made of aliphatic chains or amid functions.**
- **Target-guided-synthesis studies involving dynamic combinatorial chemistry and kinetic target-guided synthesis are based on the knowledge of known protein inhibitors or ligands. Triazoles are among the most common linkers.**

- **Various computational workflows have been suggested based on molecular docking to assist fragment linking strategies.**
- **Structure-based *de novo* design programs like LigBuilder can be utilized to generate linked compounds from pre-docked fragments.**
- **An increasing number of machine learning methods have been proposed to perform fragment linking that now include 3D constraints.**

1. Introduction

Fragment based drug design (FBDD) is a well-established strategy aimed to design leads from fragment-like compounds (1,2). The method was reported for the first time for the generation of nanomolar ligands targeting the FK506-binding protein (FKBP) (3). The ligands were designed using the SAR by NMR approach, where fragments binding proximal to each other in the protein active site were identified by NMR. Since then, FBDD has been successfully used in the development of six FDA-approved drugs and more than 50 molecules currently in clinic (4) (Table S1). FBDD has also been widely used in academic research to generate leads as the method starts by screening a moderately sized library (thousands of molecules) using biochemical or biophysical techniques such as thermal shift assay, NMR, X-ray crystallography, surface plasmon resonance (SPR) among others (5). The main advantage of the FBDD approach is the high hit rate, which leads to multiple hits for the development of new compounds, even for challenging targets. On the other hand, such a method generally requires a considerable effort to determine the structures of protein-small molecule complexes.

While a first proof of FBDD concept was provided for the generation of protein inhibitors by linking two fragments weakly binding into two sub-pockets (**Figure 1**), review of the literature clearly shows that fragment growing is more commonly used than the other two main FBDD strategies, namely, fragment linking and fragment merging (1,2,6,7). Nevertheless, fragment linking remains an approach with high potential as binding affinities might dramatically increase when superadditivity is achieved (8). Superadditivity means that the binding free energy of the linked compound exceeds the simple sum of the binding energies of the individual fragments (9). In a recent review, Bancet *et al.* (6) discussed cases of fragment linking up to 2018, focusing on the strategies required to identify the initial fragments. In half of the cases, a second screening campaign was performed to identify fragments binding to a second sub-pocket. Furthermore, by analyzing the gain in affinity upon fragment linking, very few examples of superadditivity were found in the literature. In a more recent review, Bedwell *et al.* discussed the strengths and drawbacks of structure-based approaches used in fragment linking (7).

The authors also reviewed target-guided approaches where the fragments were assembled *in situ* in the presence of the protein.

In the present manuscript, we focus on two main aspects. First, we review the chemical synthesis strategies used in fragment linking, with special attention to the most common types of linkers and chemical reactions to link fragments. We also discuss fragment self-assembly studies, *i.e.,* where linking is directly achieved by a chemical reaction thermodynamically or kinetically favored by the protein that acts as a template for the reaction. Second, we detail important studies that incorporated computational techniques to identify novel compounds by fragment linking. Structure-based *de novo* design methods and generative models are also presented as valuable tools for compound and linker design. Finally, fragment and linker libraries as well as filtering methods are briefly discussed.

2. Linkers used in published studies

While fragment linking may be viewed as a powerful method to increase binding affinity (8), it faces two major challenges. The first one is the determination of the distance and the relative orientation of the two fragment hits once they are shown to bind into distinct sub-pockets of the protein. This is generally done using experimental techniques, preferably X-ray crystallography, although *in silico* methods can provide insights into fragment binding modes (section 4). The second challenge is to design a linker that preserves the initial protein-fragment interactions. In the following paragraphs, we will review the types of linkers and the chemical reactions used to link fragments in 40 examples reported in the literature (Table 1; SMILES are provided in Table S2). Notably, cases where protein-templated fragment self-assembly was used were excluded as they will be the main focus of section 3 (Table 2; SMILES are provided in Table S3).

2.1. Aliphatic chains

Aliphatic chain linkers have the advantage of being non-cleavable with increased stability in a cellular context as compared to cleavable linkers. These linkers usually require the formation of a carbon-

carbon bond using chemical reactions such as the Suzuki, Sonogashira and Grubbs reactions, among others.

The length of a linker is an important feature for the success of fragment linking. In several reports, fragments hits could be linked by short aliphatic chains containing only one or two carbons thanks to the proximity and ideal orientation of the fragments. For example, Hajduk *et al.* (10) successfully used a two-carbon aliphatic linker to obtain of a 15 nM matrix metalloproteinase stromelysin (MMP-3) inhibitor by alkylating a phenolic fragment with β-propiolactone. Jordan *et al.* (11) also used a two-carbon linker to obtain of a 0.8 nM inhibitor of BACE-1 from two initially identified fragments. The chlorine-containing fragment was modified into an iodine-containing fragment, whereas an alkyne function was incorporated into the second fragment. A Sonogashira coupling of the alkyne and iodine derivatives led to an alkyne compound, which was then reduced to an aliphatic chain. Another example with a short single-carbon linker was published for a thrombin inhibitor (12). The authors introduced an aldehyde function on the aromatic ring of one of the fragments, which could react with the amine moiety of the other fragment, resulting in a potent 1.4 nM thrombin inhibitor.

Other short aliphatic linkers were reported in the case of fragment dimerization, *i.e.*, where two identical fragments binding to separate chains of an oligomeric target are linked, taking advantage of the spatial proximity of the binding pockets. Fragment dimerization was successfully applied to inhibitors targeting the oligomeric form of the pyruvate kinase M2 (13). Owing to the proximity of the amide moieties of the fragments, linking of the fragments required a simple methylene group. The dimerization reaction was carried out by activating DMSO with DCMT. The best activator molecule showed an AC_{50} of 3 nM. Similarly, the dimerization of a fragment bound to the AMPA receptor was published (14). The authors used a Grubbs reaction to join the two identical fragments. The alkene-containing molecule was then reduced to a two-carbon aliphatic chain, resulting in a 13.4 nM ligand.

In the pioneering FBDD work reported on the FKPB target, the Fesik group used a three-carbon aliphatic chain to link two fragments bound to two sub-pockets of the protein (3). The authors combined carboxylic acid and alcohol moieties to generate potent FKPB ligands with a 1000-10000-

times greater affinity than the starting fragments. Möbitz *et al.* (15) also proposed a three-carbon aliphatic chain to link fragments bound to Dot1L kinase. A fragment with an amide function was modified by the addition of an aliphatic chain containing an aldehyde moiety. The modified fragment could then react with the second fragment containing an amine function, resulting in a highly potent Dot1L inhibitor (K_i = 2 pM). Iannelli *et al.* (16) reported an example with a linker containing three atoms of carbons to target Protein Arginine Methyl Transferase (PRMT4). The fragments were linked by the reaction of an amine with a protected thiourea moiety to generate potent inhibitor 3 nM.

A four-carbon chain was proposed for HSP90 inhibitors by Barker *et al.* (17). Interestingly, the authors observed that the second fragment exhibited different binding modes depending on the presence or absence of the first fragment. Fragment linking required two main steps, the formation of a carbon-carbon bond by the reaction of an enolate with a halide and an aromatic nucleophilic substitution on a chloro-purine cycle.

Finally, the longest aliphatic chain linker was reported for the design of acetylcholinesterase inhibitors (18), where a chain of seven carbon atoms led to 1.4 nM compounds.

2.2. Aliphatic chains with an ether function

The ether function is also widely used in medicinal chemistry due to its high stability. Two main reactions give access to ether derivatives, the Mitsunobu and the Williamson reactions. Szczepankiewicz *et al.* (19) linked two fragments with amide and ether functions, leading to the discovery of protein PTP1B inhibitors. One of the fragments carried a carboxylic acid function while the second, a naphthoic acid, had to be functionalized. A Mitsunobu reaction was used to obtain the ether linker containing a primary amine. Finally, the combination of the acid and amine fragments resulted in a 22 nM PTP1B inhibitor. For the same target, Liu *et al.* (20) combined two fragments with a more rigid linker (allyl ether) than usual. One fragment was modified by a Still coupling reaction to introduce an allyl alcohol function. The modified fragment was linked to the second fragment containing a phenolic function using a Mitsunobu reaction. A lead compound with an IC_{50} of 6.9 µM was obtained.

Other examples based on Mitsunobu coupling are shown in Table 1, for inhibitors of the Blc-2 family of proteins (21), Mycobacterium Tuberculosis Phospho-pantetheine Adenylyl-Transferase (22) and the human N-myristoyltransferase (23).

The Williamson reaction has been used in several cases reported in Table 1 (24–28). For example, Frank *et al.* (24) successfully linked two fragments that were shown to bind in two sub-pockets of the protein RPA70N. Initial fragments were modified in order to introduce a phenol and a benzyl-bromide moieties. Gelin *et al.* (25) used a propargyl ether to link two AMP molecules shown to bind to two adjacent pockets of the NADK protein. To link the 8-position of adenine and 5'-position of D-ribose, adenosine was modified (bromination at the 8-position and alkylation of D-ribose at 5'-position using a Williamson reaction). The bromide-AMP and the alkyne-AMP were linked by a Sonagashira reaction. The Williamson reaction was also used to generate an ether-containing linker for Bcl-xL inhibitors (26), BACE-1 inhibitors (27) and NADPH inhibitors (28).

2.3. Amide functions

The amide function is quite common in medicinal chemistry and is easily obtained by condensation of acid and amine functions. Several examples of fragment linking using amide-containing linkers are shown in Table 1.

A pioneering example was reported by Burgess *et al.* (29) with the dimerization of a benzamidine fragment bound to two adjacent active sites of the tryptase enzyme. The initial benzamidine fragment was substituted in the para position to introduce an acidic function that could then react with an amine-containing linker. A long linker (21 bonds) also containing aromatic rings and an ether function was required to bridge the two fragments and obtain picomolar inhibitors of the enzyme.

In a number of cases, amide linkers were obtained by the reaction of acidic and amine functions that have been introduced on the initially discovered fragments (30–37). In terms of chemical synthesis, this strategy generally requires steps of protection and deprotection of other acidic and amine functions if present on the initial fragments. As shown in several reports (31–33), linkers can also contain two amide functions, especially when the linked inhibitor requires more than four bonds. In

particular, the Hyvönen group reported a bivalent inhibitor of the CK2 kinase. The authors combined a fragment bound to the ATP site with a fragment bound to an allosteric pocket (33). A case of fragment dimerization was reported by Trapero *et al.* (38) for inhibitors of the IMPDH protein from *Mycobacterium tuberculosis*. The fragment was shown to bind simultaneously to two sub-pockets of the protein. A Mitsunobu reaction was used to introduce a carboxylic acid function to the fragment. At the same time, an amine function was added by the cyclization of bromo-acetone and acetyl-guanine followed by an acetyl-deprotection. The best compound showed an IC_{50} of 270 nM.

Remarkably, amide linkers were also used for RNA targets (39). Swayze *et al.* selected fragment hits using the SAR by mass spectrometry method. One of the fragments contained a carboxylic acid function while the second fragment was modified by a Suzuki reaction to introduce an aniline moiety. The fragments were then combined using different types of amide linkers (aromatic, aliphatic and alkyne), resulting in RNA ligands with micromolar affinities (K_d = 6.5 µM).

2.4 Other functions

A sulfonamide-containing linker was used for the discovery of fragment-based HSP90 inhibitors (40). Based on the orientation and distance between the two fragments, the sulfonamide function was proposed as the linker. The originally identified fragments were modified to introduce an amine and an activated sulfonic acid (sulfonyl chloride). The linked compound showed a K_i of 1.9 µM.

Acyl-sulfonamide linkers have also been exemplified (41,42). Hung *et al.* (41) linked fragments bound to pantothenate synthetase using an acylation reaction of a sulfonamide moiety. This moiety required carboxylic acid activation followed by nucleophilic substitution on the carbonyl group by the nitrogen of the sulfonamide. A similar linker has been described by Rega *et al.* (42). The authors obtained a protein-protein inhibitor targeting the Bxl-xL and MCl1 proteins with affinities in the sub-micromolar range.

A polyalcohol linker was used by Kohlmann *et al.* (43) to generate inhibitors of lactate dehydrogenase. A first polyethylene glycol chain led to a linked compound exhibiting an IC_{50} of 59 µM. The authors then modified the linker to an aliphatic chain substituted with four alcohol functions using a Mitsunobu

reaction and reductive amination. The final compound showed improved inhibition activity (IC_{50} = 120 nM).

A propargyl linker was chosen by Villemagne *et al.* (44) for the optimization of inhibitors targeting the EthR protein of *Mycobacterium tuberculosis* using a Sonagashira coupling reaction while the group of Blundell explored a disulfide bridge for linking fragments targeting the same enzyme (45).

Urea linkers were reported by two groups (46,47). Belkacem *et al.* (46) designed amide- and urea-containing linkers inspired by natural ligands, for inhibitors of human cyclophilins. Thanks to the optimized interactions of the urea function with the cyclophilins, a cyclophilin inhibitor was generated with IC_{50} from 200 to 80 nM. Fragment linking involved the reaction of one amine with carbonyldiimidazole, followed by the reaction of the other amine. Benmansour *et al.* (47) linked fragments that bind to the NS5 methyltransferase of dengue virus. The authors chose the urea moiety to link the two amine functions of the fragments.

Recently, Zeller *et al.* used an aromatic nucleophilic substitution to obtain RNA binder with a one chemical-bond linker (48).

3. Self-assembly of fragments as a linking strategy

3.1. Overview of the self-assembly strategies

Another strategy for fragment linking is fragment self-assembly, a method called target-guided-synthesis (TGS) (49). This method uses the protein binding site as a template to promote the chemical reaction *in situ* between two reactive fragments. Two strategies of TGS have been reported in the literature: dynamic combinatorial chemistry (DCC) and kinetic target-guided synthesis (KTGS) (**Figure 2**). In DCC, fragments with complementary reactive groups undergo reversible reactions in the absence of the target, resulting in a dynamic combinatorial library (DCL). These interconvertible assemblages are thermodynamically stabilized and amplified upon complex formation with the protein target (50,51). KTGS also relies on the protein as a template. However, unlike DCC, the final ligand is generated by the formation of an irreversible bond formed between functionalized fragments that are

positioned in close proximity upon protein binding (49,52,53). Self-assembly can also take place inside protein crystals (the method is referred to as dynamic combinatorial X-ray crystallography) (54). One key advantage is that the complex structure can be directly observed, providing insights into the protein-inhibitor interactions. To our knowledge, although this is an attractive method (54), very few examples have been reported in the literature (55). Below, we report published cases of fragment linking using DCC and KTGS (Table 2 and Table S2). The chemical reactions involved in these reports are shown in **Figure 3**.

3.2. Amines (DCC)

DCC proof of concept was first reported by Huc and Lehn for the generation of carbonic anhydrase II (CAII) inhibitors. The authors combined aldehyde and amine moieties to form imines, which were then reduced in amine linkers (**Figure 3a**) (56). Amine linkers were also exemplified for a neuraminidase inhibitor using ketone and amine moieties (57). Although these examples are not true fragment-linking cases, the strategy may be useful in FBDD.

3.3. Hydrazones (DCC)

Congreve *et al.* (54) reported the reaction of hydrazines and isatins in the protein crystal of the kinase CDK2. Crystals of CDK2 were soaked in reaction solutions containing seven hydrazines and six isatin monomers generated from the deconstruction of a known 8 nM CDK2 inhibitor. The resulting hydrazone-containing compounds (**Figure 3b**), which were observed to bind to the ATP site in the crystal, were all shown to inhibit CDK2 activity (IC_{50} = 30 nM).

3.4. Acyl-hydrazones (DCC)

Using modelling and structural information from co-crystal structures of endothiapepsin-fragment complexes, acyl-hydrazone-based inhibitors were designed and selected. Then, five hydrazides and five aldehydes inspired by the deconstruction of the putative endothiapepsin inhibitors were then either synthesized or purchased. Out of twenty-five potential acyl-hydrazone inhibitors (**Figure 3c**), the formation of eight acyl-hydrazones bound to endothiapepsin was detected by Saturation Transfer Difference (STD) NMR experiments. All the compounds were shown to inhibit endothiapepsin with

IC_{50} values ranging from 13 to 365 µM (58). An elegant example has also been reported for the generation of inhibitors of the glutathione-S-transferase enzyme family (59). To bridge the two GSH sites of the GST homodimer, three aldehydes derived from a known chloro-2,4-dinitrobenzene inhibitor were mixed with four bis-hydrazide linkers also based on a known inhibitor. The DCL was tested against different GST isoenzymes (mouse, human, *Schistosoma japonicum*). The approach yielded selective bivalent inhibitors with nanomolar affinity targeting the dimeric protein (59). To our knowledge, this is the only example where DCC has been used to generate an inhibitor that binds to two subunits of a dimeric protein.

3.5. Triazoles (KTGS)

One of the most popular KTGS reactions is the 1,3 dipolar cycloaddition between azides and alkynes, also called protein templated click chemistry (PTCC) (60,61) (**Figure 3d**). As reported in Table 2, one strategy is to modify the two starting fragments into an alkyne-containing library and an azide-containing library. This strategy was illustrated in the work of Mondal *et al.* (62), with the design of inhibitors of endothiapepsin. KTGS was also used to generate stabilizers of c-MYC G-quadruplex. Pomeislova *et al.* (63) identified the best analogue of a previously identified fragment using FRET experiments. They derived the fragment into an alcyne-containing library of eight compounds that were directly coupled to four azides inspired from a previous *in situ* click-chemistry study. The study led to the discovery of newly synthesized bis-amino derivatives of 4-(4-methylpiperazin1-yl)aniline as potent stabilizers of the c-MYC G-quadruplex.

Using natural product derivatives of galantamine, cryspine A and melatonin Molęda *et al.* (64) used PTCC to generate a bivalent inhibitor of human acetylcholine esterase (*h*AChE) and butyrylcholinesterase (*h*BChE). Five alkynes were derived from galantamine and cryspine while seven azides were derived from melatonin. Several inhibitors were obtained in the presence of the proteins, including a potent inhibitor derived from galantamine (IC_{50} = 0.018 µM for *h*AChE and IC_{50} = 0.963 µM for *h*BChE).

Another KTGS approach consists in deconstructing a known inhibitor into two fragments, which are then chemically modified to create azide- and alkyne-containing libraries. This strategy was applied for inhibitors of histone deacetylases (65). A library of 15 azide-containing compounds and two alkyne-containing compounds was obtained from the deconstruction of vorinostat, a 0.28 µM IC_{50} inhibitor. The best histone deacetylase inhibitor with a triazole linker was shown to display an IC_{50} of 0.51 µM (65). Similarly, Mocharla *et al.* (66) obtained a 0.9 nM inhibitor of the bovine carbonic anhydrase, starting from a 37 nM alkyne fragment coupled to a library of 24 azide-containing fragments. A similar KTGS proof of concept has been reported for the HIV-1 protease (67).

In a study by Grimster *et al.* (68), an azide-containing library (10 compounds) was inspired by the natural protein substrate acetylcholine of acetylcholine binding protein AChBP. Five alkynes were used to react with the azide moieties and one 0.96 nM triazole-based acetylcholine mimetic was successfully produced.

Finally, Peruzzotti *et al.* (69) linked two fragments with a triazole moiety to give a 0.9 µM inhibitor of the Abl tyrpsone kinase.

3.6. Amides (KTGS)

As revealed by an analysis of the World Drug Index (70), amide bonds are found in many bioactive ligands for diverse protein targets. In this context, Jaegle *et al* (71) reported a successful protein-templated irreversible amide-based linkage (**Figure 3e**) for Blood Coagulation Factor Xa (BcF Xa). An activated carboxylic acid derivative and a nucleophilic amine were derived from the deconstruction of a known inhibitor, resulting in the generation a BcF Xa amide-based inhibitor (K_i = 29 nM).

3.7. Thioethers (KTGS)

To improve the selectivity of a c-Src kinase inhibitor (K_i = 0.36 µM), Kwarcinski *et al.* (72) used enzyme-templated Michael addition (**Figure 3f**) to link the thiol-decorated inhibitor to acrylamide compounds. The best inhibitor showed an improved affinity for the target (K_i = 0.09 µM) as well as an improved selectivity. One limitation of this approach is that the cysteine residues of the protein that are accessible on the surface have to be mutated to prevent side reactions. Another biocompatible

reaction was conducted as a proof of concept by Loussouarn *et al.* (73) for the synthesis of carbonic anhydrase inhibitors using thiol-containing fragments with alkyne-containing fragments (**Figure 3g**).

3.8. Acylsulfonamides (KTGS)

In a recent publication (74), acylsulfonalmide inhibitors of the Bcl-xL protein, a well-known target in oncology, were obtained using the sulfo-click reaction (**Figure 3h**) as previously reported for the same target (75). Nacheva *et al.* (74) significantly increased the throughput of KTGS, allowing the screening of 1710 combinations of fragments. Two libraries of 38 sulfonyl-azides and 45 thioacids, respectively, were inspired by the deconstruction of the known Bcl-xL inhibitors ABT-737 and ABT-263. Thioacids were generated from fluorenylmethylthioester just prior to screening due to their relative instability. They used LCMS/MS with triple quadrupole mass spectrometry to detect the synthesized acylsulfonamides. A limitation of this method is that the protein must be stable at 37°C for several hours.

3.9. Multicomponent linking (KTGS)

As proposed by Mancini *et al.* (76), multicomponent reactions (MCRs) represent another KTGS method where "*one-pot reactions, in which more than two starting materials are incorporated to form a new adduct comprising most of the structural motifs of the starting building block* ". Gladysz *et al.* (77) performed a Groebke–Blackburn-Bienaymé reaction templated by Urokinase plasminogen activator (uPA). An isocyanide moiety reacted with three aldehydes and two amines in the presence of uPA to generate an imidazopyridine amine inhibitor (IC_{50} = 0.184 µM). Recently, Mancini *et al.* (76) also used an example of a multicomponent reaction (Ugi 4-Component Reaction : 4-CR) using endothiapepsin as a protein model. The fragments contained either carboxylic acid, amine, aldehyde, or isocyanide moieties. The best compound showed an improved activity (IC_{50} = 1.3 µM) compared to the original inhibitor (IC_{50} = 12.8 µM).

4. *In silico* approaches and fragment-linking software

4.1. Docking and virtual-screening-based strategies

Over the past decades, many computational workflows have been developed to support fragment linking. Approaches involving docking and virtual screening (VS) are usually conducted to prioritize fragments for experimental testing or to select most promising linked candidates to be synthesized. For example, researchers from AstraZeneca (78) used docking to identify novel inhibitors of the activated factor XI through the linking of fragments with substructures of known inhibitors. First, VS of an in-house library made of 65000 fragments was performed to highlight fragments targeting the sub-pocket (S1) in the catalytic site. X-ray crystallography was used to probe the binding mode of hits validated by NMR and SPR. Second, linking of fragments with inhibitor substructures known to bind to adjacent sub-pockets (S1'-S2') was considered. Various linkers (amide and acetamide derivatives) were proposed from visual inspection of crystal structures while linked compounds were virtually tested by docking. Compounds displaying preserved binding modes for both moieties before and after docking were chosen for synthesis. The best molecule showed nanomolar activity, making it one of the most potent and selective FXIa inhibitors (78).

Yu *et al.* (79) introduced *in silico* site-directed FBDD, a computational strategy to identify potent ligands based on docking and deconstruction-reconstruction approach. The method was applied to the discovery of novel STAT3 inhibitors. As a first step, nine known STAT3 inhibitors were docked and deconstructed into fragments and linkers. Reconstruction was then carried out by randomly linking fragments following a combinatorial approach. To link the fragments, linkers obtained from the deconstruction step were considered as well as bioisosteres picked for their flexibility and synthesis feasibility (amide and alkyl amide). Virtually designed compounds were docked and selected for synthesis when the fragments could reposition well inside the pockets according to docking simulations. Five compounds were synthesized; four of them exhibited an IC_{50} below 5 μM.

In another study, docking was utilized to guide the design of new inhibitors of histone methyltransferase Dot1L via fragment linking (15). This work was supported by a previous experimental screening by the same authors that revealed fragments hits targeting a pocket adjacent

to the cofactor S-adenosylmethionine (SAM) binding site (80). In the new study (15), VS was first run on a set of fragments satisfying an adenosine-based pharmacophore to obtain hits binding in the SAM pocket. The position of the fragments in the SAM pocket was confirmed by X-ray crystallography. Fragments from each pocket were selected when linking via a three or four-carbon chain appeared feasible. Notably, only carbon chains were considered as linkers, i.e., no further investigation into linker optimization was reported. Finally, synthesized molecules showed a pIC_{50} of 8.4 while pIC_{50} of fragments were 3.6 and 5.1, respectively, making the process almost superadditive.

Another important study was conducted by Benmansour *et al.* (47) who used molecular modelling to design new inhibitors of the NS5 methyltransferase of the Dengue virus. Using fragments previously reported to bind into two adjacent pockets, linking, mainly through amide and carbamide functions was virtually performed using the MOE 2011 (81) program. Linked compounds were designed from MOE's Link Multiple Fragments utility (now known as Scaffold Replacement tool) by considering linkers from MOE's internal library (22 603 segments in the 2011 version). Generated compounds were docked using pharmacophore constraints to preserve the initial placement of the fragments. From docking predictions, several urea derivatives were chosen for synthesis. Despite particular care taken with *in-silico* design, associated cocrystal structures revealed different binding modes than those of the initial fragments and of the docking poses. The best compounds showed a 100-fold improvement in IC_{50} compared to the initial fragments.

4.2. De novo design software

Although docking may help prioritize the synthesis of linked molecules, e.g., by screening a large set of possible candidates designed from fragment hits and a linker library, random or combinatorial linking approaches may lead to many incompatible molecules to be docked which can slow down the identification of suitable ligands. To address this issue, *de novo* design programs were introduced to virtually link (or grow) fragments from a list of structural constraints such as the receptor structure and/or the original binding modes of the fragments. In this way, promising candidates can be designed

right away while docking might be needed in a second step to confirm the high affinity of the compounds and that original binding modes are preserved upon linking.

One of the first proposed *de novo* design programs, LUDI (82), uses a three-step approach including *(i)* the identification of interaction sites as protein regions made of hydrophobic cavities or containing well suited patterns to form hydrogen bonds *(ii)* the fitting of fragments into interaction sites based on distance criteria *(iii)* the linking of fragments by fitting a bridge fragment (linker) between well-positioned fragments in vicinity.

Many popular programs for ligand design are based on optimization algorithms, especially genetic algorithms (GA), which enable to build molecules while improving the binding energy. This is the case of the GANDI software (83) which employs a GA to join pre-docked fragments from a user-supplied list of linkers. One of the main benefits of GANDI is that the compounds are designed to simultaneously optimize the binding energy (force field term) and the similarity to known inhibitors (based on fingerprints and/or the 3D overlap with known binding modes) thereby improving the chances of generating drug-like molecules.

Another popular GA-based program is LigBuilder (84) which offers to build linked compounds by stacking chemical building-blocks one by one starting from pre-docked fragments. Building blocks include around 60 of the smallest and most common chemical groups including CH3, C2H6…, as well as single/multiple rings. Therefore, LigBuilder has the advantage of not testing many pre-built linkers from a library but to construct the linker on the fly considering the geometry of the binding pocket, which increases specificity. Over the years, two major upgrades have been made to the program. First, LigBuilder2 (85) which incorporates the same features as the first version while offering to analyze the synthetic accessibility of designed compounds as well as to generate by itself starting fragments to be linked (exploring mode). Second, LigBuilder3 (86) which applies the above strategies on multiple protein structures simultaneously, supporting the design of multi-target drugs. In addition, an "ensemble-linking" strategy derived from the original method can now be used to facilitate the linking of two fragments and increase ligand efficiency. As a recent application, Pascoini *et al.* (87) used

LigBuilder to identify potential inhibitors of acetylcholinesterase (AChE). To design the compounds, several approaches were considered including the linking of fragments obtained from deconstruction of three known AChE inhibitors as well as the linking of fragments sampled via the exploring mode. Among the 2.5 million generated molecules, only those displaying satisfactory ADME properties and non-similar structures were kept. Docking of selected molecules suggested that three of them could possibly show similar or better binding energy than known drugs.

4.3 Machine learning and generative models

As artificial intelligence (AI) continues to evolve at a rapid pace, many research groups are proposing machine-learning (ML) methods as an alternative to structure-based *de novo* approaches like those mentioned in the previous section. Many *de novo* design programs involving ML are based on generative models, a class of unsupervised methods aimed at learning the underlying probability distribution of a dataset. In this way, models can learn how the data were generated in the first place and sample new data points that are similar but different from the training data. Generative models for *de novo* design often rely on simplified molecular input line entry system (SMILES) representation of molecules. Models are trained against large databases (CheMBL, ZINC, etc.) to learn the syntax of SMILES language and the associated distribution representing the chemical space. Eventually, the model is able to sample new strings that are new SMILES corresponding to real molecules (88,89).

In the case of linking, generative approaches should be done by incorporating user-supplied fragments as substructures while ideally preserving original binding modes. Yang *et al*. introduced SyntaLinker (90), a generative model based on SMILES representation which tackles the linking problem in a similar way to sentence completion in natural language processing. Besides terminal fragments, SyntaLinker allows the user to specify a list of constraints on the linker like the shortest bond distance or the presence of hydrogen donors, acceptors, rotatable bonds, etc. In this way, fragment binding modes can be more easily preserved despite no structure-based approach is used. Similarly, DeLinker by Imrie *et al.* (91) can generate linked structures from fragments and 3D structural constraints like the distance and the orientation between fragments. The model relies on a graph representation of the molecules

and a particular type of architecture for generative models called variational auto-encoder (VAE) which is not detailed here. We refer the interested reader to the original paper or (92) for more information. As compared to simple database lookup, DeLinker was shown to design 60% more molecules resembling the original binders in terms of binding mode.

Finally, we would like to mention DRlinker (93) as another recent program for the design of linked molecules. The method is based on *reinforcement learning*, another class of ML where models interact with an environment and learn from feedback in the form of rewards or penalties. DRlinker can perform fragment linking while ensuring specific properties of the final compounds such as the linker length, log P or synthetic accessibility. Importantly, DRlinker requires the user to define a scoring function beforehand so the model can automatically sample compounds with expected properties.

4.4 Databases for linkers and fragments

High-quality chemical libraries largely contribute to the success of linking strategies involving experimental or virtual screenings. Whether it is about fragments (for the identification initial hits) or linkers (for the design of final compounds), diversity remains a criterion of choice to the design of a good library. Regarding fragments, several libraries made of commercial compounds exist such as the Maybridge library (about 30k fragments), Enamine (about 172k fragments) or ChemDiv (about 11k fragments) among the most popular (94). Once fragment hits are identified, virtual linker libraries like Molinspiration (95) or MOE's internal linker library (81) may be used to support the design of linked molecules either from combinatorial approaches or *de novo* software. As mentioned above, it is recommended to dock the designed compounds to confirm their high affinity and/or to verify that binding modes may be preserved upon linking. It should be noted that computational fragment-based strategies can also be applied to speed up the virtual screening of regular compounds. This is the case of the recent virtual synthon hierarchical enumeration screening (V-SYNTHES) approach (96) which allows the screening of the Enamine REAL Space (97) (11 billion compounds) in a reasonable amount time, taking advantage of the building-block architecture of the library.

Apart from popular libraries, fragments or linkers derived from moieties of known ligands for the target of interest should be considered, as they may increase the chance of designing potent molecules, especially in FBDD. In this context, *in-silico* tools can be used to deconstruct known binders into fragments and linkers to make a custom library. Such programs comprise the Retrosynthetic Combinatorial Analysis Procedure (RECAP) that recognizes biologically meaningful building blocks from standard molecules. A successful computational study involving RECAP was performed for the discovery of allosteric inhibitors of metabotropic glutamate receptor 5 (mGlu5), an excitatory G Protein Coupled Receptor (GPCR) (98). As an improvement of the RECAP procedure, the Breaking of Retrosynthetically Interesting Chemical Substructures (BRICS) approach was proposed based on a more advanced set of rules for fragment generation (99).

After the initial fragment library has been chosen or linked compounds have been designed, filtering strategies can be applied, *e.g.*, to improve diversity in the dataset or to reduce the number of molecules to be screened. This is usually done by generating 2D fingerprints of the compounds, like Morgan fingerprints, and by clustering fingerprints based on Tanimoto similarity. Importantly, other types of filtering may be considered, *e.g.*, discarding molecules that do not satisfy the Rule of Three (fragments) or Lipinski's Rule of Five (linked molecules). Other approaches include filtering the dataset for synthetic accessibility (SA) or ADME properties as predicted by computational programs or web servers. Examples of popular tools are SYLVIA (100) for SA prediction and, SwissADME (101) and the vNN web server (102) for ADME properties while we refer the reader to the recent review by de Souza Neto *et al.* (88) for a more exhaustive list.

5. Conclusion

In this review, 58 cases of fragment linking involving two or more fragments to generate protein or nucleic acid ligands have been examined (Tables 1 & 2). In all the cases, fragments were successfully linked either without the target (Table 1) – as a standard structure-based design approach – or with the target that served as a template to facilitate reactions between the fragments upon binding (Table

2). Popular linkers for standard fragment linking appeared to be flexible and short (less than 4 carbons in 75% of cases) where aliphatic chains (with or without an ether function) and amide functions are among the most represented types. Chemical reactions compatible with DCC and KTGS conditions yielded amine, hydrazone, acyl-hydrazone, triazole, amide, thioether, ene-thioether, and acylsulfonamide linkers, where triazole and amide are the most common functions. Importantly, DCC and KTGS strategies are usually applied as proofs of concept based on deconstruction and target-guided synthesis reconstruction of known protein inhibitors. In addition, except for a recent report based on acylsulfonamide linkers, the throughput of target-guided synthesis has been rather low, involving less than ten moieties derived from each initial fragment. Finally, we have detailed important studies which illustrate how computational tools can assist the discovery of new compounds by fragment linking. Strategies based on docking and/or structure-based *de novo* design programs are generally used to highlight most interesting fragments or linkers to be tested. Most popular *de novo* design programs were also described including ML-based models for fragment linking which have recently gained popularity as AI techniques continue to improve.

6. Expert opinion

In contrast to fragment growing, determination of the protein-fragment complex structure is a central step in fragment linking as it provides crucial information on the distance and the relative orientation of the fragments to be linked. This information is used to determine whether and how possible linkers can connect the two fragments and to estimate the length of the linker. Compounds made with linkers of various types and lengths must be synthesized in order to identify molecules exhibiting the best activity. Moreover, the identification of two fragments binding to two distinct pockets may be quite challenging. For all these reasons, it is no wonder that fragment-linking studies are less represented than their fragment-growing counterparts. However, we believe that fragment linking is particularly convenient for the discovery of bivalent molecules targeting either two distinct pockets in a protein

structure (*e.g.*, an active and an allosteric site) or a pocket involving different subunits of the same oligomeric complex.

Protein-templated chemical reactions like DCC and KTGS offer a great opportunity to link fragments in a direct manner. First, 3D structures of the protein or the protein-fragment complexes are not essential for the success of the method. Second, chemical synthesis efforts can be more easily streamlined by focusing on generating libraries of fragments containing the reactive functions that will lead to the final compound. However, DCC and KTGS may be limited as both methods use a large amount of target that need to be stable at 25°C or 37°C. As other limitations, hit identification from complex mixtures is often not straightforward whereas other biocompatible reactions might be needed to increase linker diversity. At last, structural knowledge of the target might be required in order to validate the approach with a first proof of concept, before deciding to invest in the synthesis of fragment libraries. As all published cases were based on already known inhibitors, it will be interesting to keep an eye on future strategies involving protein-templated methods applied to *de novo* lead design.

Regarding in silico techniques, molecular docking is usually presented as a popular and valuable tool to assist fragment linking as illustrated by the successful studies mentioned in section 4.1. However, special care should be taken as docking predictions may be inexact, *i.e.*, docking poses may be far away from experimental binding modes (RMSD > 2 Å). From our own experience, false positives in docking may occur in more than 50% of the cases irrespective of the program used. Although docking can provide some guidance on the choice of the fragments or linkers to be tested experimentally, we recommend to verify predicted binding modes by NMR or X-ray crystallography when possible and to select docking hits with diverse structures and functions to increase the chance of identifying true binders.

In order to determine suitable linkers, structure-based *de novo* software should be preferred over combinatorial approaches where many fragment-linker-fragment combinations, virtually designed from fragment hits and a linker library, are docked. Indeed, *de novo* programs have the benefit of constructing linked molecules from the 3D structure of the complex by optimizing the binding energy

in a similar way to docking algorithms. Programs like LigBuilder use a building-block approach that enable to generate "tailor-made" linkers that may not be accessible in a preset library. Again, linked molecules built from *de novo* software should be re-docked to investigate whether fragment binding modes have a chance to be preserved upon linking. In addition, molecular dynamics simulations can be run over dozens or hundreds of nanoseconds to check the stability of the linked molecule. In this case, one should verify if the molecule remains well anchored in the binding pocket over the simulation or if it has a tendency to unbind that may be indicative of a sub-optimal linking process.

As reported in section 4.3, many ML-based models have been introduced recently to address the linking problem offering to design linked compounds from fragments. Whereas earliest models were substantially ligand-based, new ML models are able to include 3D constraints such as the distance and/or the relative orientation between fragments. Although standard structure-based programs may still be more efficient in designing high-affinity compounds, *e.g.*, by taking care of the complexity of protein-ligand interactions, increasingly powerful ML models are being developed every day as the amount of available data and computational resources continue to grow. Therefore, it would not be surprising if ML models were able to drastically improve the design of superadditive compounds in the near future.

Figure legends

Figure 1. Schematic representation of the fragment-linking approach in FBDD.

Figure 2. Schematic representation of target-guided-synthesis (TGS) in FBDD: dynamic combinatorial chemistry (DCC) and kinetic target-guided synthesis (KTGS). A/C and B/D represent complementary reactive groups: A and B undergo reversible reactions thermodynamically favored by the protein while C and D undergo irreversible reactions kinetically favored in the presence of the protein.

Figure 3. Chemical reactions used in DCC and KTGS. Linkers include amine (a), hydrazone (b) acyl-hydrazone (c) 1,4- and 1,5-triazole (d), amide (e), ene-thioether (f), thioether (g) and acylsulfonamide (h).

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Walsh L, Erlanson DA, de Esch IJP, Jahnke W, Woodhead A, Wren E. Fragment-to-Lead Medicinal Chemistry Publications in 2021. J Med Chem. 2023;66:1137–1156. doi: 10.1021/acs.jmedchem.2c01827.

2. de Esch IJP, Erlanson DA, Jahnke W, Johnson CN, Walsh L. Fragment-to-Lead Medicinal Chemistry Publications in 2020. J Med Chem. 2022;65:84–99. doi: 10.1021/acs.jmedchem.1c01803.

3. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering High-Affinity Ligands for Proteins: SAR by NMR. Science. 1996;274:1531–1534. doi: 10.1126/science.274.5292.1531.

•• Pioneering work demonstrating the fragment linking feasibility in FBDD

4. Rees DC, Hirsch AKH, Erlanson DA. Introduction to the themed collection on fragment-based drug discovery. RSC Med Chem. 2022;13:1439–1439. doi: 10.1039/D2MD90037H.

5. Kirsch P, Hartman AM, Hirsch AKH, Empting M. Concepts and Core Principles of Fragment-Based Drug Design. Molecules. 2019;24:4309. doi: 10.3390/molecules24234309.

6. Bancet A, Raingeval C, Lomberget T, Le Borgne M, Guichou J-F, Krimm I. Fragment Linking Strategies for Structure-Based Drug Design. J Med Chem. 2020;63:11420–11435. doi: 10.1021/acs.jmedchem.0c00242.

7. Bedwell EV, McCarthy WJ, Coyne AG, Abell C. Development of potent inhibitors by fragment-linking strategies. Chem Biol Drug Des. 2022;100:469–486. doi: 10.1111/cbdd.14120.

•• Important review discussing fragment linking strategies

8. Ichihara O, Barker J, Law RJ, Whittaker M. Compound Design by Fragment-Linking. Mol Inform. 2011;30:298–306. doi: 10.1002/minf.201000174.

9. Jencks WP. On the attribution and additivity of binding energies. Proc Natl Acad Sci. 1981;78:4046–4050. doi: 10.1073/pnas.78.7.4046.

10. Hajduk PJ, Sheppard G, Nettesheim DG, Olejniczak ET, Shuker SB, Meadows RP, Steinman DH, Carrera GM, Marcotte PA, Severin J, et al. Discovery of Potent Nonpeptide Inhibitors of Stromelysin Using SAR by NMR. J Am Chem Soc. 1997;119:5818–5827. doi: 10.1021/ja9702778.

11. Jordan JB, Whittington DA, Bartberger MD, Sickmier EA, Chen K, Cheng Y, Judd T. Fragment-Linking Approach Using 19F NMR Spectroscopy To Obtain Highly Potent and Selective Inhibitors of β-Secretase. J Med Chem. 2016;59:3732–3749. doi: 10.1021/acs.jmedchem.5b01917.

••Very interesting example of a successful fragment-linking approach for a challenging target

12. Howard N, Abell C, Blakemore W, Chessari G, Congreve M, Howard S, Jhoti H, Murray CW, Seavers LCA, van Montfort RLM. Application of Fragment Screening and Fragment Linking to the Discovery of Novel Thrombin Inhibitors. J Med Chem. 2006;49:1346–1355. doi: 10.1021/jm050850v.

13. Matsui Y, Yasumatsu I, Asahi T, Kitamura T, Kanai K, Ubukata O, Hayasaka H, Takaishi S, Hanzawa H, Katakura S. Discovery and structure-guided fragment-linking of 4-(2,3-dichlorobenzoyl)-1-methyl-pyrrole-2-carboxamide as a pyruvate kinase M2 activator. Bioorg Med Chem. 2017;25:3540–3546. doi: 10.1016/j.bmc.2017.05.004.

14. Drapier T, Geubelle P, Bouckaert C, Nielsen L, Laulumaa S, Goffin E, Dilly S, Francotte P, Hanson J, Pochet L, et al. Enhancing Action of Positive Allosteric Modulators through the Design of Dimeric Compounds. J Med Chem. 2018;61:5279–5291. doi: 10.1021/acs.jmedchem.8b00250.

15. Möbitz H, Machauer R, Holzer P, Vaupel A, Stauffer F, Ragot C, Caravatti G, Scheufler C, Fernandez C, Hommel U, et al. Discovery of Potent, Selective, and Structurally Novel Dot1L Inhibitors by a Fragment Linking Approach. ACS Med Chem Lett. 2017;8:338–343. doi: 10.1021/acsmedchemlett.6b00519.

••An important research article showing an example of superadditivity

16. Iannelli G, Milite C, Marechal N, Cura V, Bonnefond L, Troffer-Charlier N, Feoli A, Rescigno D, Wang Y, Cipriano A, et al. Turning Nonselective Inhibitors of Type I Protein Arginine Methyltransferases into Potent and Selective Inhibitors of Protein Arginine Methyltransferase 4 through a Deconstruction–Reconstruction and Fragment-Growing Approach. J Med Chem. 2022;65:11574–11606. doi: 10.1021/acs.jmedchem.2c00252.

17. Barker JJ, Barker O, Courtney SM, Gardiner M, Hesterkamp T, Ichihara O, Mather O, Montalbetti CAGN, Müller A, Varasi M, et al. Discovery of a Novel Hsp90 Inhibitor by Fragment Linking. ChemMedChem. 2010;5:1697–1700. doi: 10.1002/cmdc.201000219.

18. Pang Y-P, Quiram P, Jelacic T, Hong F, Brimijoin S. Highly Potent, Selective, and Low Cost Bis-tetrahydroaminacrine Inhibitors of Acetylcholinesterase: STEPS TOWARD NOVEL DRUGS FOR TREATING ALZHEIMER'S DISEASE. J Biol Chem. 1996;271:23646–23649. doi: 10.1074/jbc.271.39.23646.

19. Szczepankiewicz BG, Liu G, Hajduk PJ, Abad-Zapatero C, Pei Z, Xin Z, Lubben TH, Trevillyan JM, Stashko MA, Ballaron SJ, et al. Discovery of a Potent, Selective Protein Tyrosine Phosphatase 1B Inhibitor Using a Linked-Fragment Strategy. J Am Chem Soc. 2003;125:4087–4096. doi: 10.1021/ja0296733.

20. Liu G, Xin Z, Liang H, Abad-Zapatero C, Hajduk PJ, Janowick DA, Szczepankiewicz BG, Pei Z, Hutchins CW, Ballaron SJ, et al. Selective Protein Tyrosine Phosphatase 1B Inhibitors: Targeting the Second Phosphotyrosine Binding Site with Non-Carboxylic Acid-Containing Ligands. J Med Chem. 2003;46:3437–3440. doi: 10.1021/jm034088d.

21. Petros AM, Huth JR, Oost T, Park C-M, Ding H, Wang X, Zhang H, Nimmer P, Mendoza R, Sun C, et al. Discovery of a potent and selective Bcl-2 inhibitor using SAR by NMR. Bioorg Med Chem Lett. 2010;20:6587–6591. doi: 10.1016/j.bmcl.2010.09.033.

22. El Bakali J, Blaszczyk M, Evans JC, Boland JA, McCarthy WJ, Fathoni I, Dias MVB, Johnson EO, Coyne AG, Mizrahi V, et al. Chemical Validation of Mycobacterium tuberculosis Phosphopantetheine Adenylyltransferase Using Fragment Linking and CRISPR Interference. Angew Chem. 2023;135:e202300221. doi: 10.1002/ange.202300221.

23. Mousnier A, Bell AS, Swieboda DP, Morales-Sanfrutos J, Pérez-Dorado I, Brannigan JA, Newman J, Ritzefeld M, Hutton JA, Guedán A, et al. Fragment-derived inhibitors of human N-

myristoyltransferase block capsid assembly and replication of the common cold virus. Nat Chem. 2018;10:599–606. doi: 10.1038/s41557-018-0039-2.

•Example of inhibitor designed via fragment reconstruction and linking leading to a 100 000-fold improvement in potency

24. Frank AO, Feldkamp MD, Kennedy JP, Waterson AG, Pelz NF, Patrone JD, Vangamudi B, Camper DV, Rossanese OW, Chazin WJ, et al. Discovery of a Potent Inhibitor of Replication Protein A Protein–Protein Interactions Using a Fragment-Linking Approach. J Med Chem. 2013;56:9242–9250. doi: 10.1021/jm401333u.

25. Gelin M, Paoletti J, Nahori M-A, Huteau V, Leseigneur C, Jouvion G, Dugué L, Clément D, Pons J-L, Assairi L, et al. From Substrate to Fragments to Inhibitor Active In Vivo against Staphylococcus aureus. ACS Infect Dis. 2020;6:422–435. doi: 10.1021/acsinfecdis.9b00368.

26. Tao Z-F, Hasvold L, Wang L, Wang X, Petros AM, Park CH, Boghaert ER, Catron ND, Chen J, Colman PM, et al. Discovery of a Potent and Selective BCL-XL Inhibitor with in Vivo Activity. ACS Med Chem Lett. 2014;5:1088–1093. doi: 10.1021/ml5001867.

27. Fang W-S, Sun D, Yang S, Cheng C, Moschke K, Li T, Sun S, Lichtenthaler SF, Huang J, Wang Y. Discovery of a series of selective and cell permeable beta-secretase (BACE1) inhibitors by fragment linking with the assistance of STD-NMR. Bioorganic Chem. 2019;92:103253. doi: 10.1016/j.bioorg.2019.103253.

28. Solbak SMØ, Zang J, Narayanan D, Høj LJ, Bucciarelli S, Softley C, Meier S, Langkilde AE, Gotfredsen CH, Sattler M, et al. Developing Inhibitors of the p47phox–p22phox Protein–Protein Interaction by Fragment-Based Drug Discovery. J Med Chem. 2020;63:1156–1177. doi: 10.1021/acs.jmedchem.9b01492.

29. Burgess LE, Newhouse BJ, Ibrahim P, Rizzi J, Kashem MA, Hartman A, Brandhuber BJ, Wright CD, Thomson DS, Vigers GPA, et al. Potent selective nonpeptidic inhibitors of human lung tryptase. Proc Natl Acad Sci. 1999;96:8348–8352. doi: 10.1073/pnas.96.15.8348.

•A paper describing an example of fragment dimerization

30. Sledz P, Silvestre HL, Hung AW, Ciulli A, Blundell TL, Abell C. Optimization of the Interligand Overhauser Effect for Fragment Linking: Application to Inhibitor Discovery against Mycobacterium tuberculosis Pantothenate Synthetase. J Am Chem Soc. 2010;132:4544–4545. doi: 10.1021/ja100595u.

31. Becattini B, Culmsee C, Leone M, Zhai D, Zhang X, Crowell KJ, Rega MF, Landshamer S, Reed JC, Plesnila N, et al. Structure–activity relationships by interligand NOE-based design and synthesis of antiapoptotic compounds targeting Bid. Proc Natl Acad Sci. 2006;103:12602–12606. doi: 10.1073/pnas.0603460103.

32. Ward RA, Brassington C, Breeze AL, Caputo A, Critchlow S, Davies G, Goodwin L, Hassall G, Greenwood R, Holdgate GA, et al. Design and Synthesis of Novel Lactate Dehydrogenase A Inhibitors by Fragment-Based Lead Generation. J Med Chem. 2012;55:3285–3306. doi: 10.1021/jm201734r.

33. Brear P, Fusco CD, Georgiou KH, J. Francis-Newton N, J. Stubbs C, F. Sore H, R. Venkitaraman A, Abell C, R. Spring D, Hyvönen M. Specific inhibition of CK2α from an anchor outside the active site. Chem Sci. 2016;7:6839–6845. doi: 10.1039/C6SC02335E.

34. Wang P-F, Wang Z-F, Qiu H-Y, Huang Y, Hu H-M, Wang Z-C, Zhu H-L. Identification and Biological Evaluation of Novel Type II B-RafV600E Inhibitors. ChemMedChem. 2018;13:2558–2566. doi: 10.1002/cmdc.201800574.

35. Grädler U, Schwarz D, Blaesse M, Leuthner B, Johnson TL, Bernard F, Jiang X, Marx A, Gilardone M, Lemoine H, et al. Discovery of novel Cyclophilin D inhibitors starting from three dimensional fragments with millimolar potencies. Bioorg Med Chem Lett. 2019;29:126717. doi: 10.1016/j.bmcl.2019.126717.

36. Floresta G, Fallica AN, Patamia V, Sorrenti V, Greish K, Rescifina A, Pittalà V. From Far West to East: Joining the Molecular Architecture of Imidazole-like Ligands in HO-1 Complexes. Pharmaceuticals. 2021;14:1289. doi: 10.3390/ph14121289.

37. Wu X, Ramesh R, Wang J, Zheng Y, Armaly AM, Wei L, Xing M, Roy S, Lan L, Gao FP, et al. Small Molecules Targeting the RNA-Binding Protein HuR Inhibit Tumor Growth in Xenografts. J Med Chem. 2023;66:2032–2053. doi: 10.1021/acs.jmedchem.2c01723.

38. Trapero A, Pacitto A, Singh V, Sabbah M, Coyne AG, Mizrahi V, Blundell TL, Ascher DB, Abell C. Fragment-Based Approach to Targeting Inosine-5'-monophosphate Dehydrogenase (IMPDH) from Mycobacterium tuberculosis. J Med Chem. 2018;61:2806–2822. doi: 10.1021/acs.jmedchem.7b01622.

39. Swayze EE, Jefferson EA, Sannes-Lowery KA, Blyn LB, Risen LM, Arakawa S, Osgood SA, Hofstadler SA, Griffey RH. SAR by MS: A Ligand Based Technique for Drug Lead Discovery Against Structured RNA Targets. J Med Chem. 2002;45:3816–3819. doi: 10.1021/jm0255466.

40. Huth JR, Park C, Petros AM, Kunzer AR, Wendt MD, Wang X, Lynch CL, Mack JC, Swift KM, Judge RA, et al. Discovery and Design of Novel HSP90 Inhibitors Using Multiple Fragment-based Design Strategies. Chem Biol Drug Des. 2007;70:1–12. doi: 10.1111/j.1747-0285.2007.00535.x.

41. Hung AW, Silvestre HL, Wen S, Ciulli A, Blundell TL, Abell C. Application of Fragment Growing and Fragment Linking to the Discovery of Inhibitors of *Mycobacterium tuberculosis* Pantothenate Synthetase. Angew Chem Int Ed. 2009;48:8452–8456. doi: 10.1002/anie.200903821.

42. Rega MF, Wu B, Wei J, Zhang Z, Cellitti JF, Pellecchia M. SAR by Interligand Nuclear Overhauser Effects (ILOEs) Based Discovery of Acylsulfonamide Compounds Active against Bcl-xL and Mcl-1. J Med Chem. 2011;54:6000–6013. doi: 10.1021/jm200826s.

43. Kohlmann A, Zech SG, Li F, Zhou T, Squillace RM, Commodore L, Greenfield MT, Lu X, Miller DP, Huang W-S, et al. Fragment Growing and Linking Lead to Novel Nanomolar Lactate Dehydrogenase Inhibitors. J Med Chem. 2013;56:1023–1040. doi: 10.1021/jm3014844.

44. Villemagne B, Flipo M, Blondiaux N, Crauste C, Malaquin S, Leroux F, Piveteau C, Villeret V, Brodin P, Villoutreix BO, et al. Ligand Efficiency Driven Design of New Inhibitors of Mycobacterium tuberculosis Transcriptional Repressor EthR Using Fragment Growing, Merging, and Linking Approaches. J Med Chem. 2014;57:4876–4888. doi: 10.1021/jm500422b.

45. Surade S, Ty N, Hengrung N, Lechartier B, Cole ST, Abell C, Blundell TL. A structure-guided fragment-based approach for the discovery of allosteric inhibitors targeting the lipophilic binding site of transcription factor EthR. Biochem J. 2014;458:387–394. doi: 10.1042/BJ20131127.

46. Ahmed-Belkacem A, Colliandre L, Ahnou N, Nevers Q, Gelin M, Bessin Y, Brillet R, Cala O, Douguet D, Bourguet W, et al. Fragment-based discovery of a new family of non-peptidic small-molecule

cyclophilin inhibitors with potent antiviral activities. Nat Commun. 2016;7:12777. doi: 10.1038/ncomms12777.

•A study demonstrating how protein-linker interactions can improve inhibitor potency

47. Benmansour F, Trist I, Coutard B, Decroly E, Querat G, Brancale A, Barral K. Discovery of novel dengue virus NS5 methyltransferase non-nucleoside inhibitors by fragment-based drug design. Eur J Med Chem. 2017;125:865–880. doi: 10.1016/j.ejmech.2016.10.007.

48. Zeller MJ, Favorov O, Li K, Nuthanakanti A, Hussein D, Michaud A, Lafontaine DA, Busan S, Serganov A, Aubé J, et al. SHAPE-enabled fragment-based ligand discovery for RNA. Proc Natl Acad Sci. 2022;119:e2122660119. doi: 10.1073/pnas.2122660119.

49. Bosc D, Camberlein V, Gealageas R, Castillo-Aguilera O, Deprez B, Deprez-Poulain R. Kinetic Target-Guided Synthesis: Reaching the Age of Maturity: Miniperspective. J Med Chem. 2020;63:3817–3833. doi: 10.1021/acs.jmedchem.9b01183.

• A significant review on the applications of KTGS strategies in medicinal chemistry

50. Frei P, Hevey R, Ernst B. Dynamic Combinatorial Chemistry: A New Methodology Comes of Age. Chem – Eur J. 2019;25:60–73. doi: 10.1002/chem.201803365.

• A review on the applications of DCC strategies in medicinal chemistry

51. Mondal M, Hirsch AKH. Dynamic combinatorial chemistry: a tool to facilitate the identification of inhibitors for protein targets. Chem Soc Rev. 2015;44:2455–2488. doi: 10.1039/C4CS00493K.

52. Nguyen R, Huc I. Using an Enzyme's Active Site To Template Inhibitors. Angew Chem Int Ed. 2001;40:1774–1776. doi: 10.1002/1521-3773(20010504)40:9<1774::AID-ANIE17740>3.0.CO;2-G.

53. Oueis E, Sabot C, Renard P-Y. New insights into the kinetic target-guided synthesis of protein ligands. Chem Commun. 2015;51:12158–12169. doi: 10.1039/C5CC04183J.

54. Congreve MS, Davis DJ, Devine L, Granata C, O'Reilly M, Wyatt PG, Jhoti H. Detection of Ligands from a Dynamic Combinatorial Library by X-ray Crystallography. Angew Chem Int Ed. 2003;42:4479–4482. doi: 10.1002/anie.200351951.

• A paper reporting an example of DCC performed in protein crystals

55. Gelin M, Poncet-Montange G, Assairi L, Morellato L, Huteau V, Dugué L, Dussurget O, Pochet S, Labesse G. Screening and In Situ Synthesis Using Crystals of a NAD Kinase Lead to a Potent Antistaphylococcal Compound. Structure. 2012;20:1107–1117. doi: 10.1016/j.str.2012.03.024.

56. Huc I, Lehn J-M. Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly. Proc Natl Acad Sci. 1997;94:2106–2110. doi: 10.1073/pnas.94.6.2106.

• The original paper reporting the first proof of concept of DCC

57. Hochgürtel M, Biesinger R, Kroth H, Piecha D, Hofmann MW, Krause S, Schaaf O, Nicolau C, Eliseev AV. Ketones as Building Blocks for Dynamic Combinatorial Libraries: Highly Active

Neuraminidase Inhibitors Generated via Selection Pressure of the Biological Target. J Med Chem. 2003;46:356–358. doi: 10.1021/jm025589m.

58. Mondal M, Radeva N, Köster H, Park A, Potamitis C, Zervou M, Klebe G, Hirsch AKH. Structure-Based Design of Inhibitors of the Aspartic Protease Endothiapepsin by Exploiting Dynamic Combinatorial Chemistry. Angew Chem Int Ed. 2014;53:3259–3263. doi: 10.1002/anie.201309682.

59. Clipson AJ, Bhat VT, McNae I, Caniard AM, Campopiano DJ, Greaney MF. Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry. Chem – Eur J. 2012;18:10562–10570. doi: 10.1002/chem.201201507.

• An example of fragment dimerization using DCC

60. Lewis WG, Green LG, Grynszpan F, Radić Z, Carlier PR, Taylor P, Finn MG, Sharpless KB. Click Chemistry In Situ: Acetylcholinesterase as a Reaction Vessel for the Selective Assembly of a Femtomolar Inhibitor from an Array of Building Blocks. Angew Chem Int Ed. 2002;41:1053–1057. doi: 10.1002/1521-3773(20020315)41:6<1053::AID-ANIE1053>3.0.CO;2-4.

61. Kolb HC, Finn MG, Sharpless KB. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew Chem Int Ed. 2001;40:2004–2021. doi: 10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

62. Mondal M, Unver MY, Pal A, Bakker M, Berrier SP, Hirsch AKH. Fragment-Based Drug Design Facilitated by Protein-Templated Click Chemistry: Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin. Chem – Eur J. 2016;22:14826–14830. doi: 10.1002/chem.201603001.

• An interesting example of click chemistry in KTGS

63. Pomeislová A, Vrzal L, Kozák J, Dobiaš J, Hubálek M, Dvořáková H, Reyes-Gutiérrez PE, Teplý F, Veverka V. Kinetic Target-Guided Synthesis of Small-Molecule G-Quadruplex Stabilizers. ChemistryOpen. 2020;9:1236–1250. doi: 10.1002/open.202000261.

64. Molęda Z, Zawadzka A, Czarnocki Z, Monjas L, Hirsch AKH, Budzianowski A, Maurin JK. “Clicking“ fragment leads to novel dual-binding cholinesterase inhibitors. Bioorg Med Chem. 2021;42:116269. doi: 10.1016/j.bmc.2021.116269.

65. Suzuki T, Ota Y, Kasuya Y, Mutsuga M, Kawamura Y, Tsumoto H, Nakagawa H, Finn MG, Miyata N. An Unexpected Example of Protein-Templated Click Chemistry. Angew Chem Int Ed. 2010;49:6817–6820. doi: 10.1002/anie.201002205.

66. Mocharla VP, Colasson B, Lee LV, Röper S, Sharpless KB, Wong C-H, Kolb HC. In Situ Click Chemistry: Enzyme-Generated Inhibitors of Carbonic Anhydrase II. Angew Chem. 2005;117:118–122. doi: 10.1002/ange.200461580.

67. Whiting M, Muldoon J, Lin Y-C, Silverman SM, Lindstrom W, Olson AJ, Kolb HC, Finn MG, Sharpless KB, Elder JH, et al. Inhibitors of HIV-1 Protease by Using In Situ Click Chemistry. Angew Chem Int Ed. 2006;45:1435–1439. doi: 10.1002/anie.200502161.

68. Grimster NP, Stump B, Fotsing JR, Weide T, Talley TT, Yamauchi JG, Nemecz Á, Kim C, Ho K-Y, Sharpless KB, et al. Generation of Candidate Ligands for Nicotinic Acetylcholine Receptors via in

situ Click Chemistry with a Soluble Acetylcholine Binding Protein Template. J Am Chem Soc. 2012;134:6732–6740. doi: 10.1021/ja3001858.

69. Peruzzotti C, Borrelli S, Ventura M, Pantano R, Fumagalli G, Christodoulou MS, Monticelli D, Luzzani M, Fallacara AL, Tintori C, et al. Probing the Binding Site of Abl Tyrosine Kinase Using in Situ Click Chemistry. ACS Med Chem Lett. 2013;4:274–277. doi: 10.1021/ml300394w.

70. Lisurek M, Rupp B, Wichard J, Neuenschwander M, von Kries JP, Frank R, Rademann J, Kühne R. Design of chemical libraries with potentially bioactive molecules applying a maximum common substructure concept. Mol Divers. 2010;14:401–408. doi: 10.1007/s11030-009-9187-z.

71. Jaegle M, Steinmetzer T, Rademann J. Protein-Templated Formation of an Inhibitor of the Blood Coagulation Factor Xa through a Background-Free Amidation Reaction. Angew Chem Int Ed. 2017;56:3718–3722. doi: 10.1002/anie.201611547.

72. Kwarcinski FE, Steffey ME, Fox CC, Soellner MB. Discovery of Bivalent Kinase Inhibitors via Enzyme-Templated Fragment Elaboration. ACS Med Chem Lett. 2015;6:898–901. doi: 10.1021/acsmedchemlett.5b00167.

73. Lossouarn A, Puteaux C, Bailly L, Tognetti V, Joubert L, Renard P-Y, Sabot C. Metalloenzyme-Mediated Thiol-Yne Addition Towards Photoisomerizable Fluorescent Dyes. Chem – Eur J. 2022;28:e202202180. doi: 10.1002/chem.202202180.

74. Nacheva K, Kulkarni SS, Kassu M, Flanigan D, Monastyrskyi A, Iyamu ID, Doi K, Barber M, Namelikonda N, Tipton JD, et al. Going beyond Binary: Rapid Identification of Protein–Protein Interaction Modulators Using a Multifragment Kinetic Target-Guided Synthesis Approach. J Med Chem. 2023;66:5196–5207. doi: 10.1021/acs.jmedchem.3c00108.

•• A very interesting example of high throughput sulfo-click to obtain inhibitors of the Bcl-2 family

75. Kulkarni SS, Hu X, Doi K, Wang H-G, Manetsch R. Screening of Protein–Protein Interaction Modulators via Sulfo-Click Kinetic Target-Guided Synthesis. ACS Chem Biol. 2011;6:724–732. doi: 10.1021/cb200085q.

76. Mancini F, Unver MY, Elgaher WAM, Jumde VR, Alhayek A, Lukat P, Herrmann J, Witte MD, Köck M, Blankenfeldt W, et al. Protein-Templated Hit Identification through an Ugi Four-Component Reaction. Chem – Eur J. 2020;26:14585–14593. doi: 10.1002/chem.202002250.

•• A study reporting a protein-templated 4-component reaction

77. Gladysz R, Vrijdag J, Van Rompaey D, Lambeir A-M, Augustyns K, De Winter H, Van der Veken P. Efforts towards an On-Target Version of the Groebke–Blackburn–Bienaymé (GBB) Reaction for Discovery of Druglike Urokinase (uPA) Inhibitors. Chem – Eur J. 2019;25:12380–12393. doi: 10.1002/chem.201901917.

78. Fjellström O, Akkaya S, Beisel H-G, Eriksson P-O, Erixon K, Gustafsson D, Jurva U, Kang D, Karis D, Knecht W, et al. Creating Novel Activated Factor XI Inhibitors through Fragment Based Lead Generation and Structure Aided Drug Design. PLOS ONE. 2015;10:e0113705. doi: 10.1371/journal.pone.0113705.

• Molecular docking applied to the design of a superadditive compound

79. Yu W, Xiao H, Lin J, Li C. Discovery of Novel STAT3 Small Molecule Inhibitors via in Silico Site-Directed Fragment-Based Drug Design. J Med Chem. 2013;56:4402–4412. doi: 10.1021/jm400080c.

80. Scheufler C, Möbitz H, Gaul C, Ragot C, Be C, Fernández C, Beyer KS, Tiedt R, Stauffer F. Optimization of a Fragment-Based Screening Hit toward Potent DOT1L Inhibitors Interacting in an Induced Binding Pocket. ACS Med Chem Lett. 2016;7:730–734. doi: 10.1021/acsmedchemlett.6b00168.

81. Molecular Operating Environment (MOE) | MOEsaic | PSILO [Internet]. [cited 2023 May 3]. Available from: https://www.chemcomp.com/Products.htm.

82. Böhm H-J. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. J Comput Aided Mol Des. 1992;6:61–78. doi: 10.1007/BF00124387.

83. Dey F, Caflisch A. Fragment-Based de Novo Ligand Design by Multiobjective Evolutionary Optimization. J Chem Inf Model. 2008;48:679–690. doi: 10.1021/ci700424b.

84. Wang R, Gao Y, Lai L. LigBuilder: A Multi-Purpose Program for Structure-Based Drug Design. Mol Model Annu. 2000;6:498–516. doi: 10.1007/s0089400060498.

85. Yuan Y, Pei J, Lai L. LigBuilder 2: A Practical de Novo Drug Design Approach. J Chem Inf Model. 2011;51:1083–1091. doi: 10.1021/ci100350u.

86. Yuan Y, Pei J, Lai L. LigBuilder V3: A Multi-Target de novo Drug Design Approach. Front Chem [Internet]. 2020 [cited 2023 Apr 27];8.

87. Pascoini AL, Federico LB, Arêas ALF, Verde BA, Freitas PG, Camps I. In silico development of new acetylcholinesterase inhibitors. J Biomol Struct Dyn. 2019;37:1007–1021. doi: 10.1080/07391102.2018.1447513. Cited: in: : PMID: 29607738.

88. de Souza Neto LR, Moreira-Filho JT, Neves BJ, Maidana RLBR, Guimarães ACR, Furnham N, Andrade CH, Silva FP. In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. Front Chem [Internet]. 2020 [cited 2023 Apr 28];8.

•• A detailed review containing many references on fragment databases and computational tools for ADME and synthetic accessibility predictions

89. Segler MHS, Kogej T, Tyrchan C, Waller MP. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. ACS Cent Sci. 2018;4:120–131. doi: 10.1021/acscentsci.7b00512.

90. Yang Y, Zheng S, Su S, Zhao C, Xu J, Chen H. SyntaLinker: automatic fragment linking with deep conditional transformer neural networks. Chem Sci. 2020;11:8312–8322. doi: 10.1039/D0SC03126G.

91. Imrie F, Bradley AR, van der Schaar M, Deane CM. Deep Generative Models for 3D Linker Design. J Chem Inf Model. 2020;60:1983–1995. doi: 10.1021/acs.jcim.9b01120.

92. Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, Aguilera-Iparraguirre J, Hirzel TD, Adams RP, Aspuru-Guzik A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. ACS Cent Sci. 2018;4:268–276. doi: 10.1021/acscentsci.7b00572.

93. Tan Y, Dai L, Huang W, Guo Y, Zheng S, Lei J, Chen H, Yang Y. DRlinker: Deep Reinforcement Learning for Optimization in Fragment Linking Design. J Chem Inf Model. 2022;62:5907–5917. doi: 10.1021/acs.jcim.2c00982.

94. Bon M, Bilsland A, Bower J, McAulay K. Fragment-based drug discovery—the importance of high-quality molecule libraries. Mol Oncol. 2022;16:3761–3777. doi: 10.1002/1878-0261.13277.

95. Molinspiration Cheminformatics [Internet]. [cited 2023 May 3]. Available from: https://www.molinspiration.com/.

96. Sadybekov AA, Sadybekov AV, Liu Y, Iliopoulos-Tsoutsouvas C, Huang X-P, Pickett J, Houser B, Patel N, Tran NK, Tong F, et al. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. Nature. 2022;601:452–459. doi: 10.1038/s41586-021-04220-9.

97. Lyu J, Wang S, Balius TE, Singh I, Levit A, Moroz YS, O'Meara MJ, Che T, Algaa E, Tolmachova K, et al. Ultra-large library docking for discovering new chemotypes. Nature. 2019;566:224–229. doi: 10.1038/s41586-019-0917-9.

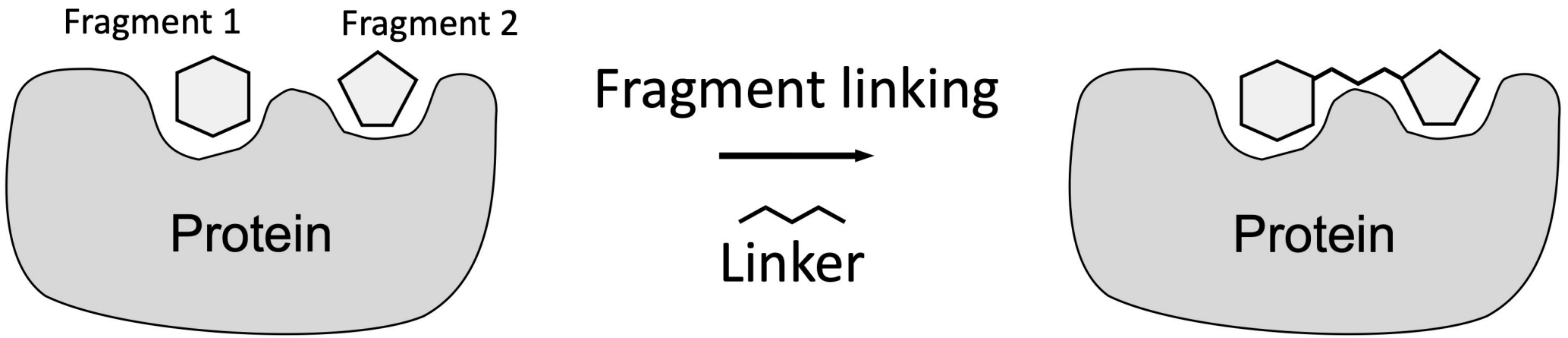
98. Bian Y, Feng Z, Yang P, Xie X-Q. Integrated In Silico Fragment-Based Drug Design: Case Study with Allosteric Modulators on Metabotropic Glutamate Receptor 5. AAPS J. 2017;19:1235–1248. doi: 10.1208/s12248-017-0093-5.

99. Degen J, Wegscheid-Gerlach C, Zaliani A, Rarey M. On the Art of Compiling and Using "Drug-Like" Chemical Fragment Spaces. ChemMedChem. 2008;3:1503–1507. doi: 10.1002/cmdc.200800178.

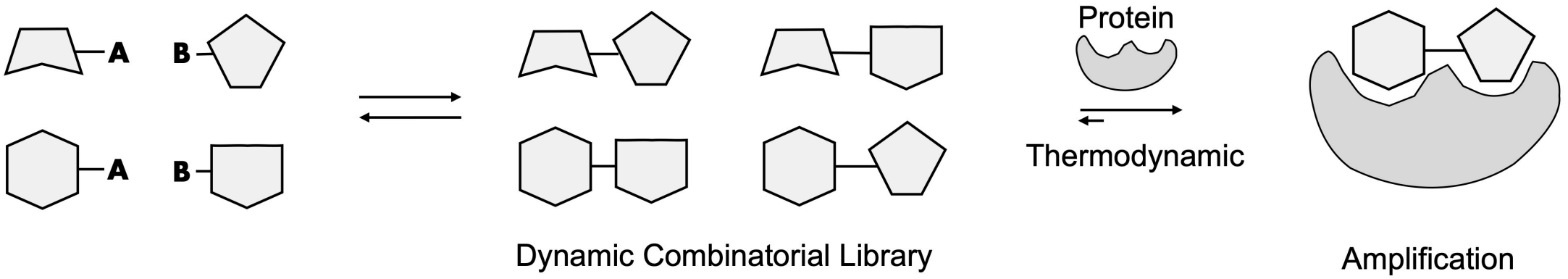
100. Boda K, Seidel T, Gasteiger J. Structure and reaction based evaluation of synthetic accessibility. J Comput Aided Mol Des. 2007;21:311–325. doi: 10.1007/s10822-006-9099-2.

101. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi: 10.1038/srep42717.

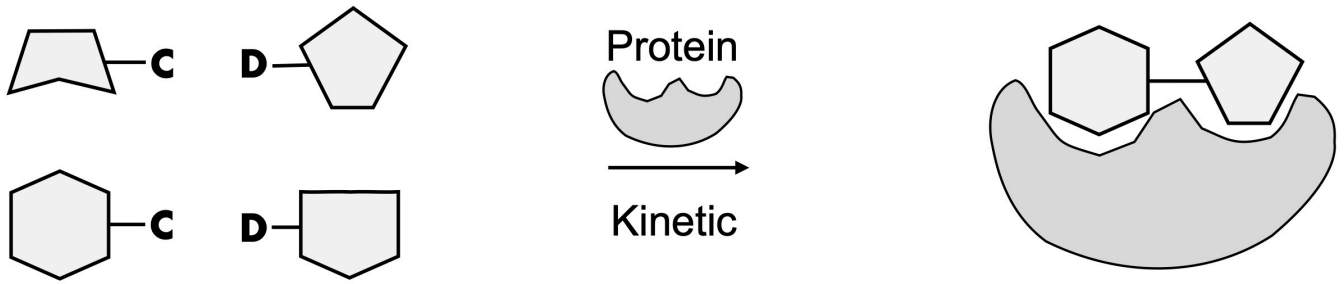
102. Schyman P, Liu R, Desai V, Wallqvist A. vNN Web Server for ADMET Predictions. Front Pharmacol [Internet]. 2017 [cited 2023 May 3];8.


Fragment 1
Fragment 2
Protein
Fragment linking
Linker
Protein

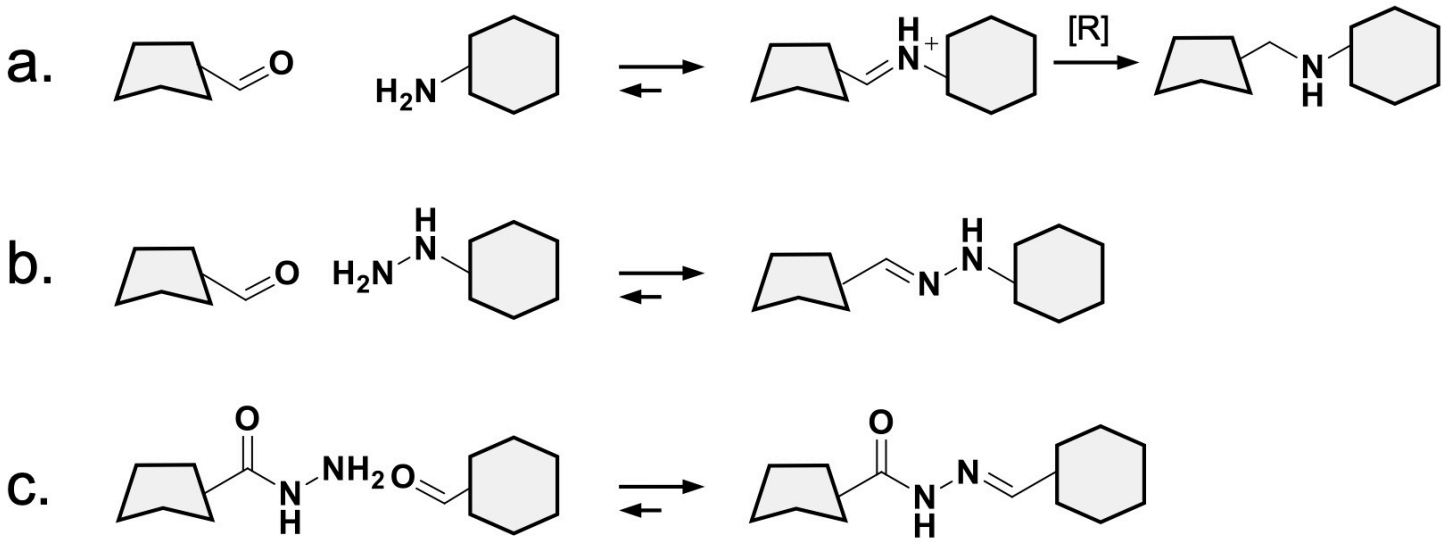
Dynamic Combinatorial Chemistry



Kinetic Target-Guided Synthesis



Dynamic Combinatorial Chemistry



Kinetic Target-Guided Synthesis

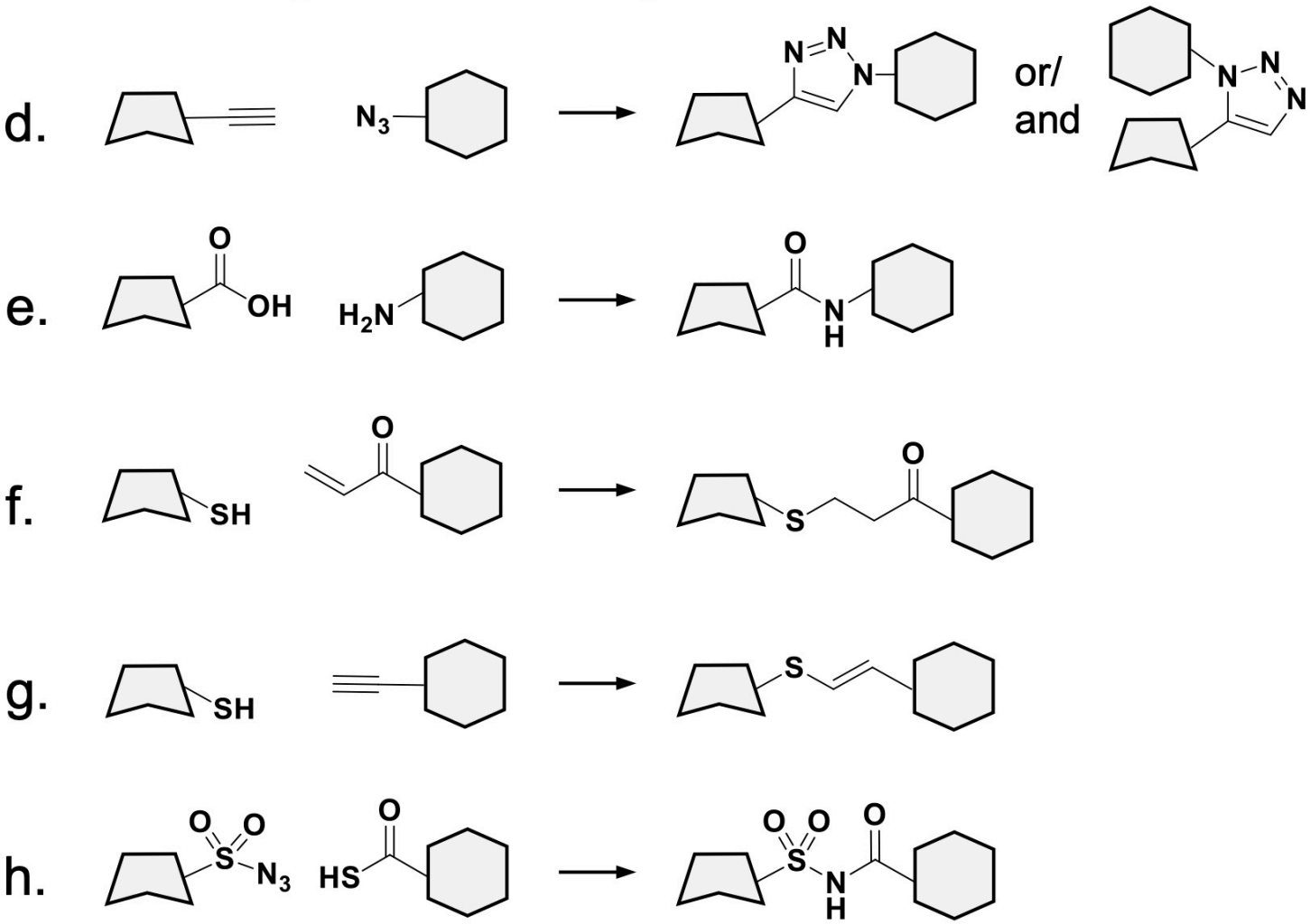
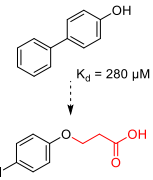
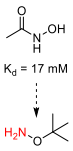
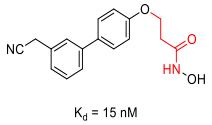
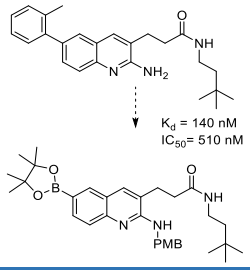
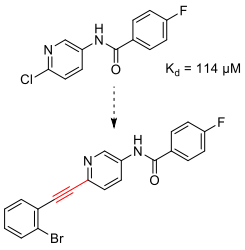
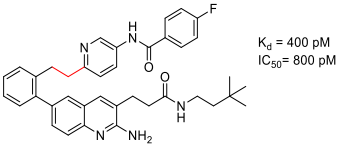
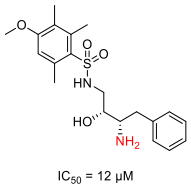
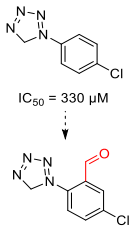
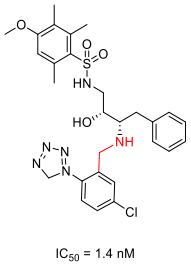
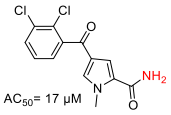
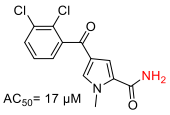
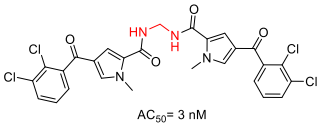
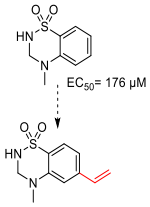
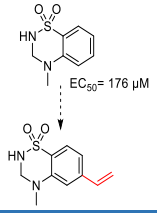
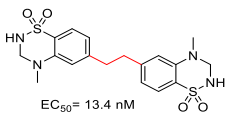
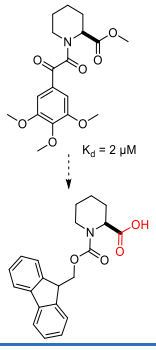
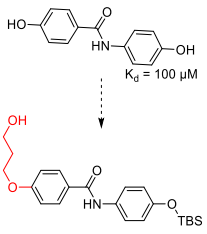
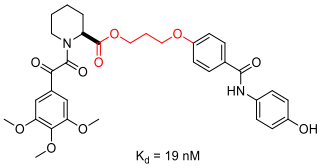
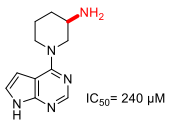
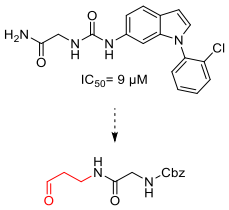
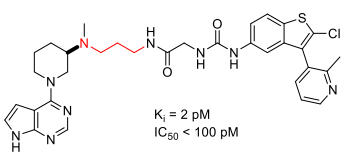
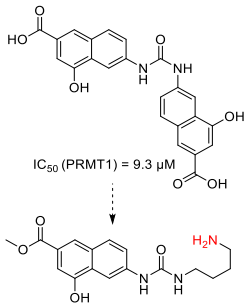
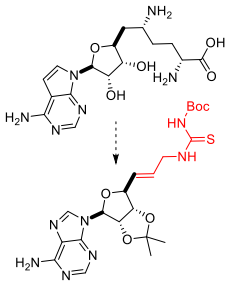
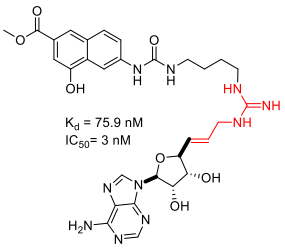
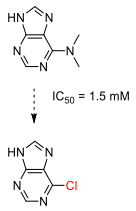
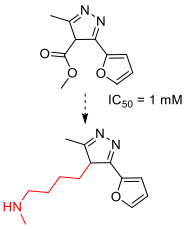
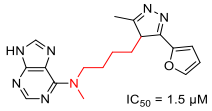
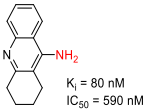
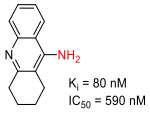
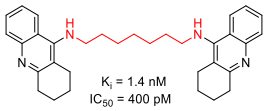
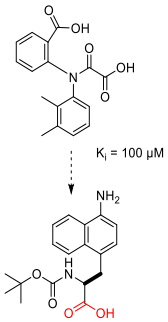
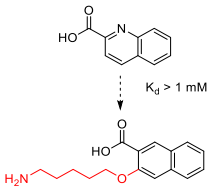
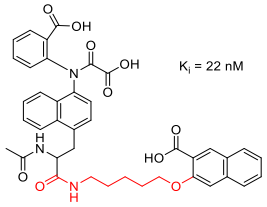
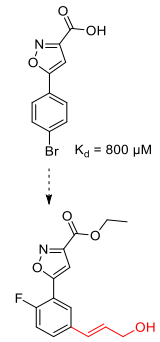
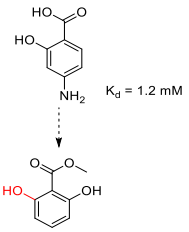
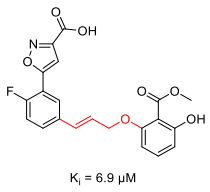
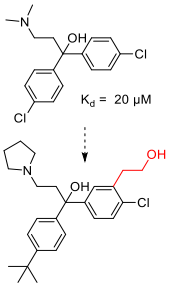
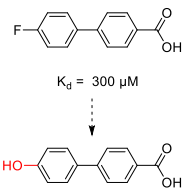
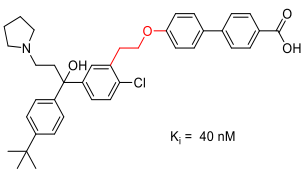
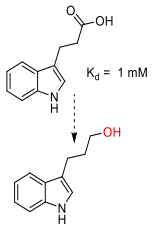
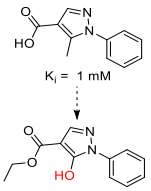
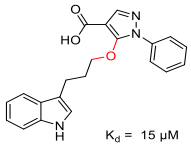
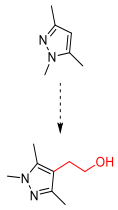
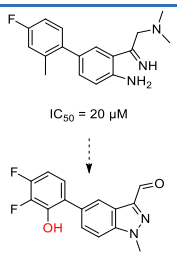
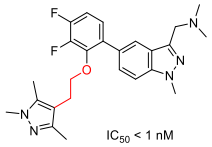
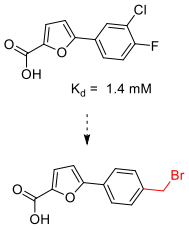
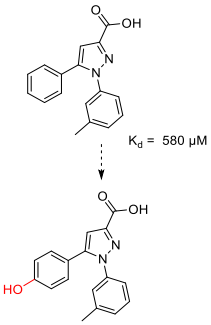
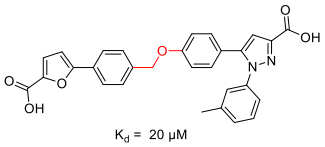
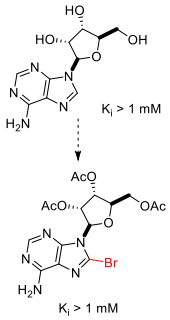
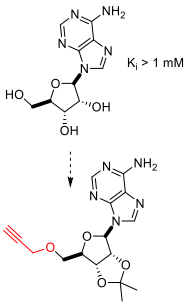
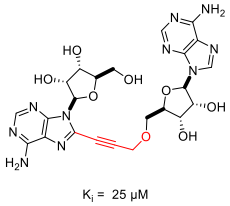
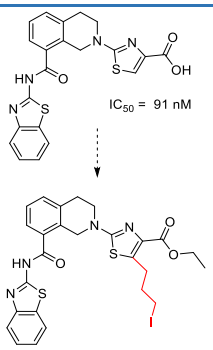
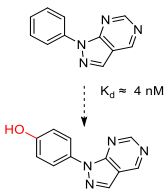
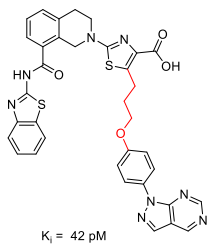
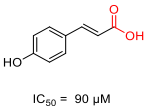
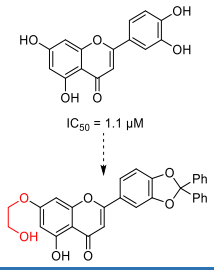
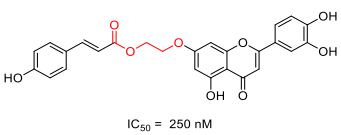


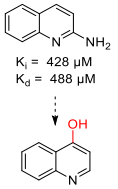
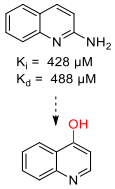
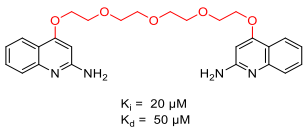
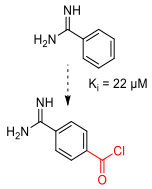
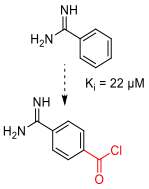
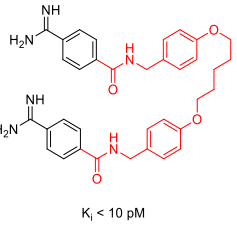
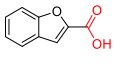
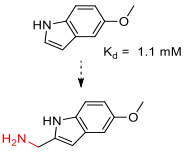
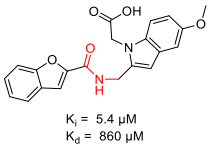
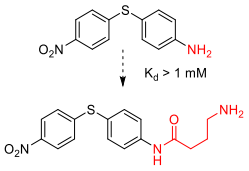
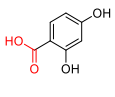
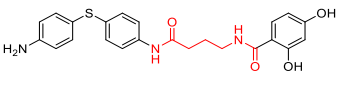
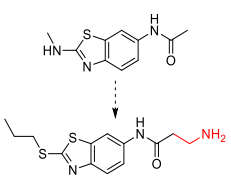
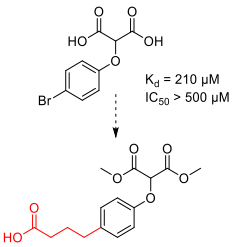
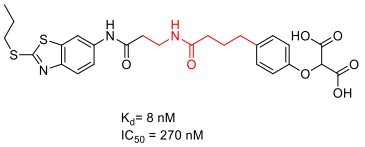
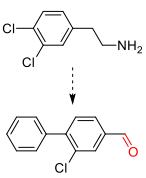
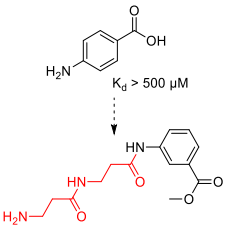
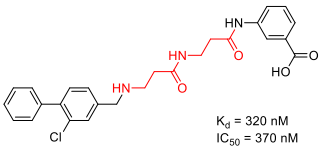
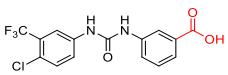
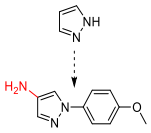
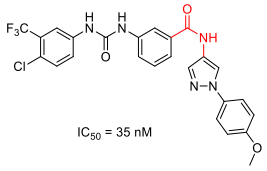
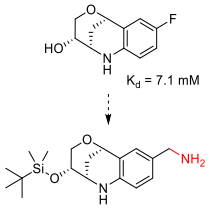
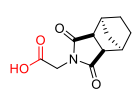
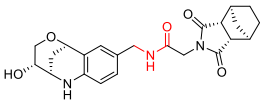
Table 11. Reported compounds generated by fragment linking

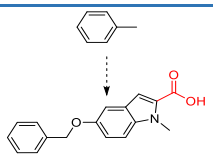
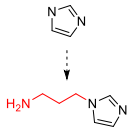
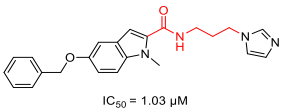
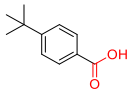
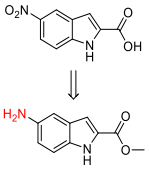
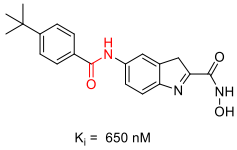
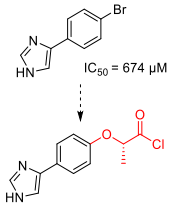
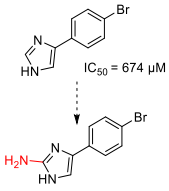
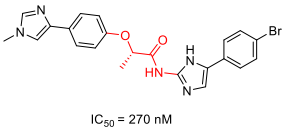
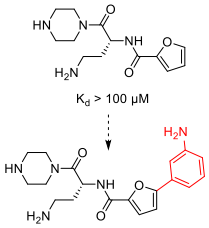
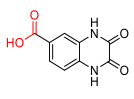
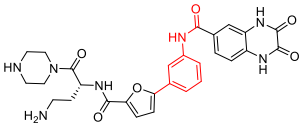
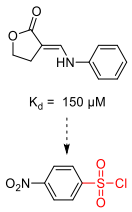
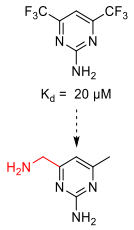
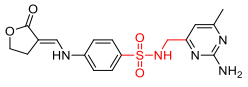
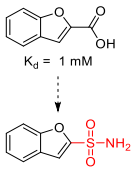
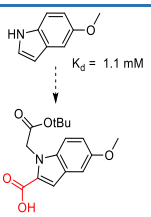
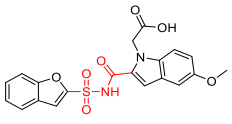
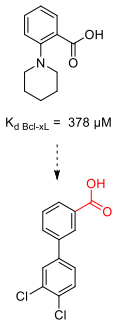
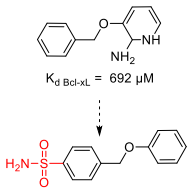
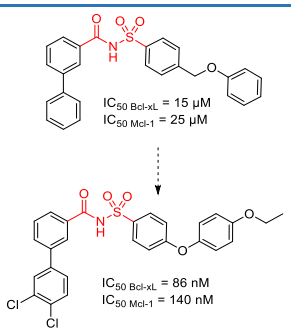
The target, the linker chemistry, the fragment modification, and the final lead are described for each case.

Target	Linker chemistry	Fragment 1	Fragment 2	Lead from fragment linking	Ref.
MMP-3 (Matrix Metallo-Proteinase Stromelysin)	Alkylation followed by acylation and Stille coupling				(10) 1997
BACE-1 (Beta-site APP Cleaving Enzyme)	Sonogashira couplings followed by hydrogenation				(11) 2016
Thrombin	Reductive amination				(12) 2006
PKM2 (Protein Kinase M2)	Alkylation of amide (dimerization)				(13) 2017
AMPA receptor (Amino-3-hydroxy-5-Methyl-4-isoxazole-Propionic Acid)	Heck reaction followed by metathesis and hydrogenation				(14) 2018
FKBP (FK506 Binding Protein)	Esterification reaction				(3) 1996
Dot1L (Disruptor Of Telomeric silencing 1-Like)	Reductive amination				(15) 2017

PRMT4 (Protein Arginine Methyl-Transferase)	Thioacylation and thiourea-to-guanidine conversion				(16) 2022
Hsp90 (Heat Shock Protein)	Enolate alkylation followed by a nucleophilic aromatic substitution				(17) 2010
AChE (Acetyl-Choline-Esterase)	Nucleophilic substitution				(18) 1996
PTP1B (Protein Tyrosine Phosphatase 1B)	Mitsunobu followed by an amidation				(19) 2003
PTP1B (Protein Tyrosine Phosphatase 1B)	Stille coupling followed by Mitsunobu				(20) 2003
Bcl-2 (B-cell lymphoma)	Mitsunobu coupling				(21) 2010

***Mtb*PPAT (*Mycobacterium Tuberculosis* Phospho-pantetheine Adenylyl-Transferase)**	Mitsunobu coupling				(22) 2023
HsNMT1 (Human N-myristoyl-transferase 1)	Mitsunobu coupling				(23) 2018
RPA (Replication Protein A)	Williamson ether synthesis				(24) 2013
NADK (Nicotinamide Adenine Dinucleotide Kinase)	Williamson ether synthesis followed by Sonogashira coupling				(25) 2020
Bcl-X_L (B-cell lymphoma extra large)	Heterocyclic synthesis followed by Williamson ether synthesis				(26) 2014
BACE1 (Beta-secretase 1)	Williamson ether synthesis followed by esterification				(27) 2019

NADPH (Nicotinamide Adenine Dinucleotide Phosphate)	Williamson ether synthesis				(28) 2020
Tryptase	Mitsunobu reaction followed by Schotten-Baumann reaction				(29) 1999
PtS (Pantothenate Synthetase)	Schotten-Baumann reaction				(30) 2010
Bid (BH3 Interacting Domain death agonist)	Schotten-Baumann reaction				(31) 2006
LDHA (Lactate De-Hydrogenase A)	HATU coupling				(32) 2012
CK2α (Casein Kinase 2)	Schotten-Baumann reaction followed by HATU coupling and a reductive amination				(33) 2016
B-Raf^{V600E} (B-Rapidly Accelerated Fibro-sarcoma)	Amidic coupling				(34) 2018
CypD (Cyclophilin D)	EDCI/HOPO-mediated amide coupling reaction				(35) 2019

HO-1 (Heme Oxygenase-1)	Peptide coupling reaction				(36) 2021
HuR (Human antigen R)	Schotten-Baumann reaction				(37) 2023
IMPDH (Inosine Monophosphate Dehydrogenase	Mitsunobu reaction followed by Schotten-Baumann reaction				(38) 2018
23S rRNA	Suzuki coupling followed by HATU mediated coupling				(39) 2002
HSP90 (Heat Shock Protein)	Negishi coupling followed by reduction and sulfonation				(40) 2007
PS (Pantothenate Synthetase)	Acylation of sulfonamide				(41) 2009
Bcl-X_L (B-cell lymphoma extra large) Mcl-1 (Myeloid Cell Leukemia)	Amidation of sulfonyl chloride followed by acylation reaction				(42) 2011

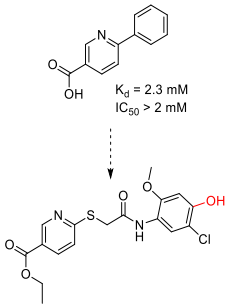
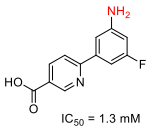
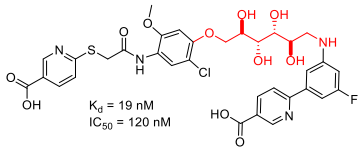
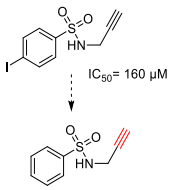
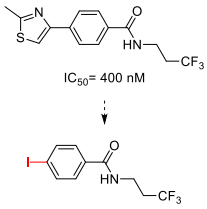
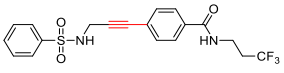
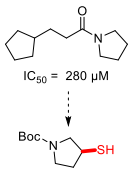
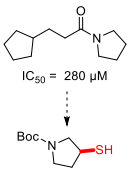
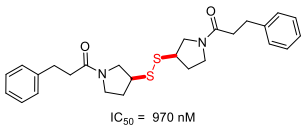
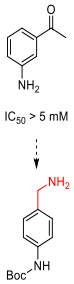
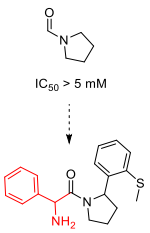
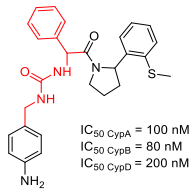
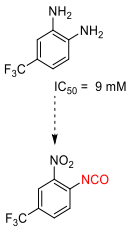
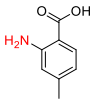
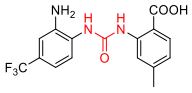
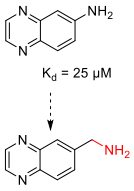
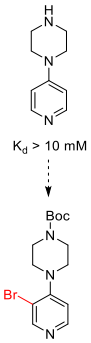
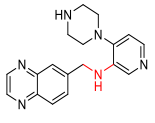
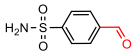
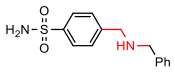
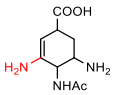
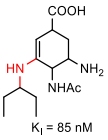
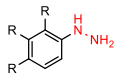
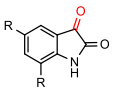
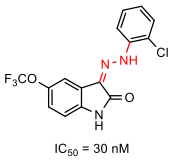
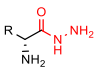
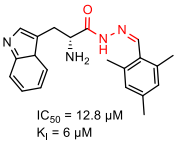
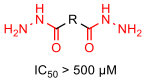
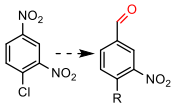
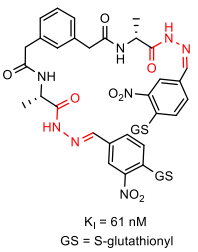
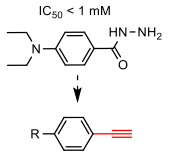
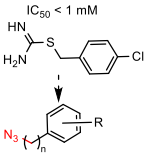
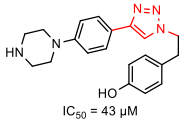
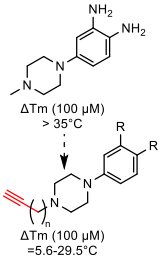
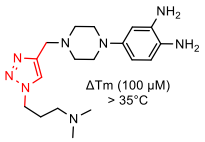
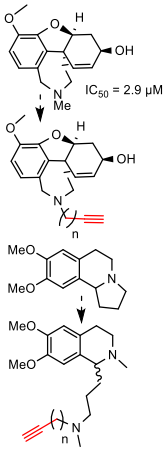
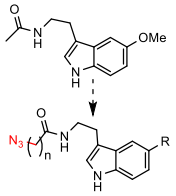
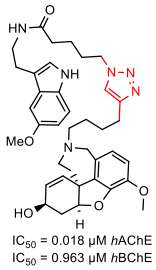
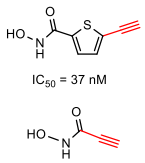
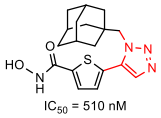
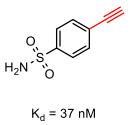
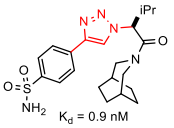
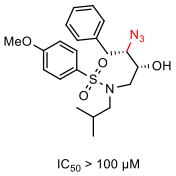
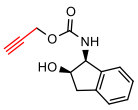
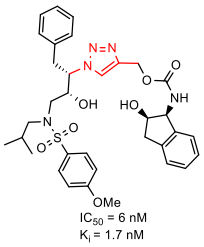
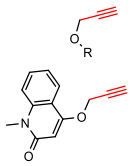
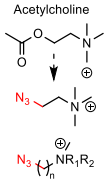
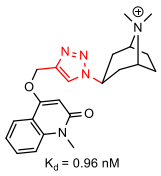
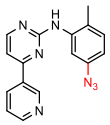
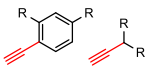
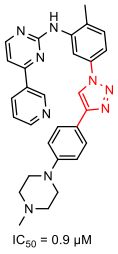
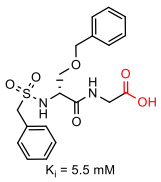
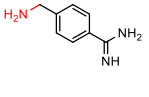
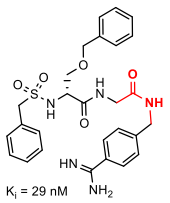
LDH-A (Lactate Dehydrogenase A)	Mitsunobu coupling followed by reductive amination				(43) 2013
EthR (Ethionamide Resistance)	Sonogashira coupling				(44) 2014
EthR (Ethionamide Resistance)	Thiol oxidation				(45) 2014
CypD (Cylophilin D)	Urea synthesis CDI mediated coupling				(46) 2016
DENV-MTase (Dengue Virus Methyl Transferase)	Urea synthesis reaction by isocyanate route				(47) 2017
TPP (Thiamine Pyro-phosphate)	Aromatic nucleophilic substitutions				(48) 2022

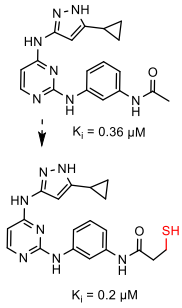
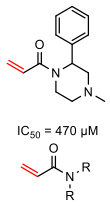
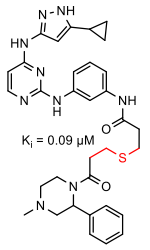
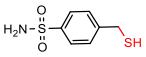
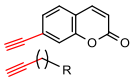
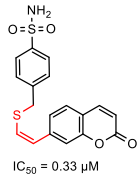
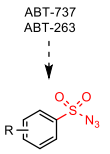
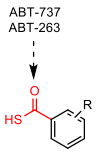
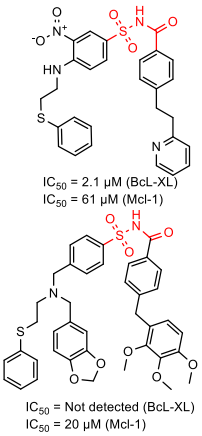
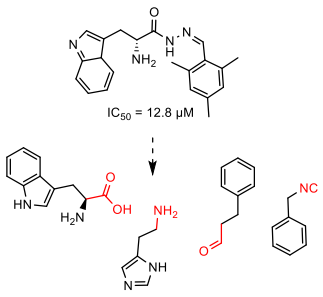
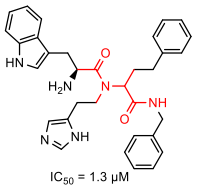
Table 21. Self-assembly of fragments as a linking strategy.

The target, the method (DCC or KTGS), the chemical reaction as well as the linker type are indicated for each case. In addition, the method used to generate fragments 1 and 2 (as well as the number of fragments in the libraries) are shown. R substituents and n number of carbon are not described (further information can be found in the corresponding references).

Target	Method and chemical reaction	Linker	Fragment 1	Fragment 2	Best linked compound	Ref.
Carbonic Anhydrase II	DCC Reductive amination	Amine	 Deconstruction of known inhibitor Library of 3 compounds	 Deconstruction of known inhibitor Library of 4 compounds		(56) 1997
Neuramidase	DCC Reductive amination	Amine	 Deconstruction of known inhibitor 1 compound	 Deconstruction of known inhibitor Library of 22 compounds	 K_I = 85 nM	(57) 2003
Cyclin-dependent kinase 2 (CDk2)	DCC in crystal Imine formation	Hydrazone	 Deconstruction of known inhibitor Library of 6 compounds	 Deconstruction of known inhibitor Library of 5 compounds	 IC_{50} = 30 nM	(54) 2003
Endothiapepsin	DCC Imine formation	Acyl-Hydrazone	 Deconstruction of potential docking based inhibitor Library of 5 compounds	 Deconstruction of potential docking based inhibitor Library of 5 compounds	 IC_{50} = 12.8 µM K_I = 6 µM	(58) 2014
Glutathione-S-transferase (GST)	DCC Imine formation	Acyl-Hydrazone	 IC_{50} > 500 µM Binding motif from known inhibitor Library of 4 compounds	 IC_{50} = 342 µM Fragment screening followed by chemical modification Library of 3 compounds	 K_I = 61 nM GS = S-glutathionyl	(59) 2012

Endothiapepsin	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 Fragment screening followed by chemical modification Library of 9 compounds	 Fragment screening followed by chemical modification Library of 4 compounds		(62) 2016
G-quadruplexe (for c-MYC gene promoter)	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 Fragment screening followed by chemical modification Library of 8 compounds	 Fragment screening Library of 4 compounds		(63) 2020
human acetyl and butyryl choline esterase (*h*AChE) (*h*BChE)	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 Fragment modified from known natural product binding galantamine and cryspine Library of 5 compounds	 Fragment modified from known natural product binding melatonin Library of 7 compounds		(64) 2021
Histone deacetylase (HDAC)	KTGS Protein templated click chemistry (PTCC)	Triazole 1,5	 Deconstruction of known inhibitor vorinostat Library of 2 compounds	 Deconstruction of known inhibitor vorinostat Library of 4 compounds		(65) 2010

Carbonic Anhydrase II	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 K_d = 37 nM Deconstruction of known inhibitor 1 compound	 Library of 24 compounds	 K_d = 0.9 nM	(66) 2005
HIV-1 Protease	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 IC_{50} > 100 µM Deconstruction of known inhibitor 1 Compound	 IC_{50} = 4.2 µM Deconstruction of known inhibitor 1 Compound	 IC_{50} = 6 nM K_i = 1.7 nM	(67) 2006
Acetylcholine binding protein	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 Library of 5 compounds	 Acetylcholine Fragments inspired from acetylcholine Library of 10 compounds	 K_d = 0.96 nM	(68) 2012
Abl Tyrosine Kinase	KTGS Protein templated click chemistry (PTCC)	Triazole 1,4	 Fragment docking Library of 2 compounds	 Fragments docked in the presence of Fragment 1 Library of 5 compounds	 IC_{50} = 0.9 µM	(69) 2013
Coagulation factor Xa	KTGS Amidation	Amide	 K_i = 5.5 mM Deconstruction of known inhibitor 1 compound	 K_i = 0.68 mM Deconstruction of known inhibitor 1 compound	 K_i = 29 nM	(71) 2017

c-Src kinase	KTGS Thiol-Michael Click	Thioether				(72) 2015
			Modification of known inhibitor	Fragment screening		
			1 compound	Library of 110 compounds		
Carbonic Anhydrase II	KTGS Thiol-yne Click	Ene-thioether				(73) 2022
			Fragment modified from 4-carboxybenzenesulfonamide			
			Library of 4 compounds	Library of 8 compounds		
Bcl-XL and Mcl-1 kinases	KTGS Sulfpho-Click	Acyl-sulfon-amide				(74) 2023
			Deconstruction of known inhibitors ABT-737 and ABT-263	Deconstruction of known inhibitors ABT-737 and ABT-263	Examples of linked molecules selective against one of the two targets	
			Library of 38 compounds	Library of 45 compounds		
Endothiapepsin	KTGS MCRs Ugi	Peptide				(76) 2020
			Deconstruction of known inhibitor acyl hydrazone			
			Library of 3 acids, amines aldehydes, isocyanate			

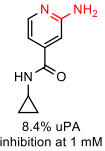
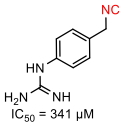
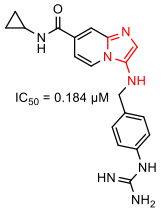
Urokinase plasminogen activator (uPA)	KTGS MCRs Groebke–Blackburn–Bienaymé	imidazo-pyridine	 8.4% uPA inhibition at 1 mM Fragment from molecular modelling Library of 2 amines	 IC_{50} = 341 µM Known binding fragment	 Fragment from molecular modelling Library of 2 aldehydes	 IC_{50} = 0.184 µM	(77) 2019

Table S11. Fragment-based compounds in clinical trials or FDA-approved in 2022.

The drug, the biological target, the clinical stage and the company are indicated.

Drug	Target(s)	Clinical stage	Company
Asciminib	BCR-ABL1	Approved	Novartis
Erdafitinib	FGFR1-4	Approved	Astex/J&J
Pexidartinib	CSF1R, KIT	Approved	Plexxikon
Sotorasib	$KRAS^{G12C}$	Approved	Amgen
Vemurafenib	$B\text{-}RAF^{V600E}$	Approved	Plexxikon
Venetoclax	BCL-2	Approved	AbbVie/Genentech
Capivasertib	AKT	Phase 3	AstraZeneca/Astex/CR-UK
Lanabecestat	BACE1	Phase 3	Astex/AstraZeneca/Lilly
Navitoclax (ABT-263)	$BCL\text{-}2/BCL_{xL}$	Phase 2	Abbott
Pelabresib (CP-0610)	BET	Phase 2	Constellation
Verubecestat	BACE1	Phase 2	Merck
ASTX029	ERK1,2	Phase 2	Astex
ASTX660	XIAP/cIAP1	Phase 2	Astex
AT7519	CDK1,2,4,5,9	Phase 2	Astex
AT9283	Aurora, JAK2	Phase 2	Astex
AUY-922	HSP90	Phase 2	Vernalis/Novartis
AZD5991	MCL1	Phase 2	AstraZeneca
DG-051	LTA4H	Phase 2	deCODE
eFT508	MNK1/2	Phase 2	eFFECTOR
Indeglitazar	pan-PPAR agonist	Phase 2	Plexxikon
LY2886721	BACE1	Phase 2	Lilly
LY3202626	BACE1	Phase 2	Lilly
LY3372689	OGA	Phase 2	Lilly
LY517717	FXa	Phase 2	Lilly/Protherics
LYS006	LTA4H	Phase 2	Novartis
MAK683	PRC2 EED	Phase 2	Novartis
Onalespib	HSP90	Phase 2	Astex
PF-06650833	IRAK4	Phase 2	Pfizer
PF-06835919	KHK	Phase 2	Pfizer
PLX51107	BET	Phase 2	Plexxikon
S64315	MCL1	Phase 2	Vernalis/Servier/Novartis
VK-2019	EBNA1	Phase 2	Cullinan Oncology / Wistar
AG-270	MAT2A	Phase 1	Agios/Servier
ABBV-744	BD2-selective BET	Phase 1	Abbott
ABT-518	MMP-2 & 9	Phase 1	Abbott
ABT-737	$BCL\text{-}2/BCL_{xL}$	Phase 1	Abbott
AT13148	AKT, p70S6K, ROCK	Phase 1	Astex
AZD3839	BACE1	Phase 1	AstraZeneca

AZD5099	Bacterial topoisomerase II	Phase 1	AstraZeneca
BI 1823911	$KRAS^{G12C}$	Phase 1	Boehringer Ingelheim
BI 691751	LTA4H	Phase 1	Boehringer Ingelheim
CFTX-1554	AT_2 receptor	Phase 1	Confo Therapeutics
ETC-206	D3	Phase 1	MNK1/2
GDC-0994	ERK2	Phase 1	Genentech/Array
HTL0014242	mGlu5 NAM	Phase 1	Sosei Heptares
HTL0018318	M1-receptor partial agonist	Phase 1	Sosei Heptares
HTL9936	M1-receptor partial agonist	Phase 1	Sosei Heptares
IC-776	LFA-1	Phase 1	Lilly/ICOS
LP-261	Tubulin	Phase 1	Locus
LY2811376	BACE1	Phase 1	Lilly
Mivebresib	BRD2-4	Phase 1	AbbVie
MRTX1719	PRMT5/MTA	Phase 1	Mirati
Navoximod	IDO1	Phase 1	New Link/Genentech
PLX5568	RAF	Phase 1	Plexxikon
SGX-393	BCR-ABL	Phase 1	SGX
SGX-523	MET	Phase 1	SGX
SNS-314	Aurora	Phase 1	Sunesis
TAK-020	BTK	Phase 1	Takeda

Table S21. Smiles of molecules listed in Table 1

SMILES strings of fragments 1 and 2 and of the linked compound are reported. For each case, the linker as colored in red in Table 1 is also shown.

Target	Fragment 1	Fragment 2	Linker	Lead from fragment linking
MMP-3 (Matrix Metallo-Proteinase Stromelysin) (10) 1997	IC1=CC=C(OCCC(O)=O)C=C1	NOC(C)(C)C	O=C(CC[R])N[R]	O=C(NO)CCOC(C=C1)=CC=C1C2=CC=CC(CC#N)=C2
BACE-1 (Beta-site APP Cleaving Enzyme) (11) 2016	O=C(NCCC(C)(C)C)CCC1=CC2=CC(B(OC3(C)C)OC3(C)C)=CC=C2N=C1NCC4=CC=C(OC)C=C4.C.C	BrC1=C(C=CC=C1)C#CC2=NC=C(C=C2)NC(C3=CC=C(C=C3)F)=O	[R]C[R]	NC1=NC2=CC=C(C3=C(CCC4=NC=C(NC(C5=CC=C(F)C=C5)=O)C=C4)C=CC=C3)C=C2C=C1CCC(NCCC(C)(C)C)=O
Thrombin (12) 2006	O=S(NC[C@@H](O)[C@@H](N)CC1=CC=CC=C1)(C2=C(C)C(C)=C(OC)C=C2C)=O	ClC1=CC=C([N]2=NN=NC2)C(C=O)=C1	[R]CN[R]	O=S(NC[C@@H](O)[C@@H](NCC1=CC(Cl)=CC=C1[N]2=NN=NC2)CC3=CC=CC=C3)(C4=C(C)C(C)=C(OC)C=C4C)=O
PKM2 (Protein Kinase M2) (13) 2017	O=C(C1=CN(C)C(C(N)=O)=C1)C2=CC=CC(Cl)=C2Cl	O=C(C1=CN(C)C(C(N)=O)=C1)C2=CC=CC(Cl)=C2Cl	[R]NCN[R]	O=C(C1=CN(C)C(C(NCNC(C2=CC(C(C3=CC=CC(Cl)=C3Cl)=O)=CN2C)=O)=O)=C1)C4=CC=CC(Cl)=C4Cl
AMPA receptor (Amino-3-hydroxy-5-Methyl-4-isoxazole-Propionic Acid) (14) 2018	O=S1(C2=CC=C(C=C)C=C2N(C)CN1)=O	O=S1(C2=CC=C(C=C)C=C2N(C)CN1)=O	[R]CC[R]	O=S1(C2=CC=C(CCC3=CC=C4C(N(C)CNS4(=O)=O)=C3)C=C2N(C)CN1)=O
FKBP (FK506 Binding Protein) (3) 1996	COC1=C(OC)C(OC)=CC(C(C(N2[C@H](C(O)=O)CCCC2)=O)=O)=C1	O=C(NC1=CC=C(O[Si](C)(C)C(C)(C)C)C=C1)C2=CC=C(OCCCO)C=C2	O=C([R])OCCCO[R]	COC1=C(OC)C(OC)=CC(C(C(N2[C@H](C(OCCCOC3=CC=C(C(NC4=CC=C(O)C=C4)=O)C=C3)=O)CCCC2)=O)=O)=C1
Dot1L (Disruptor Of Telomeric silencing 1-Like) (15) 2017	N[C@H](CCC1)CN1C2=NC=NC3=C2C=CN3	ClC1=CC=CC=C1N2C3=CC(NC(NCC(N)=O)=O)=CC=C3C=C2	[R]CCCN[R]	CC1=NC=CC=C1C2=C(Cl)SC3=CC=C(NC(NCC(NCCCN(C)[C@H](CCC4)CN4C5=NC=NC6=C5C=CN6)=O)=O)C=C32
PRMT4 (Protein Arginine Methyl-Transferase) (16) 2022	O=C(OC)C1=CC(O)=C(C=C(NC(NCCCCN)=O)C=C2)C2=C1	S=C(NC/C=C/[C@@H]1O[C@H](N2C=NC3=C(N)N=CN=C23)[C@@H]4[C@H]1OC(C)(C)O4)NC(OC(C)(C)C)=O	N=C(NC/C=C/[R])N[R]	O=C(OC)C1=CC(O)=C2C=C(NC(NCCCCNC(NC/C=C/[C@@H]3O[C@H](N4C=NC5=C(N)N=CN=C45)[C@@H](O)[C@H]3O)=N)=O)C=CC2=C1
Hsp90 (Heat Shock Protein) (17) 2010	CN(C)C1=NC=NC2=C1N=CN2	CC(C1CCCCNC)=NN=C1C2=CC=CO2	CN([R])CCCC[R]	CN(CCCCC1C(C)=NN=C1C2=CC=CO2)C3=NC=NC4=C3N=CN4
AChE (Acetyl-Choline-Esterase) (18) 1996	NC1=C2C(CCCC2)=NC3=CC=CC=C31	NC1=C2C(CCCC2)=NC3=CC=CC=C31	[R]NCCCCCCCN[R]	C12=CC=CC=C1C(NCCCCCCCNC3=C4CCCCC4=NC5=CC=CC=C35)=C6C(CCCC6)=N2
PTP1B (Protein Tyrosine Phosphatase 1B) (19) 2003	NC1=CC=C(C[C@@H](NC(OC(C)(C)C)=O)C(O)=O)C2=C1C=CC=C2	OC(C(C=C1)=NC2=C1C=CC=C2)=O	O=C([R])NCCCCCO[R]	OC(C(N(C1=C(C(O)=O)C=CC=C1)C2=CC=C(CC(NC(C)=O)C(NCCCCCOC3=CC(C=CC=C4)=C4C=C3C(O)=O)=O)=O)C5=C2C=CC=C5)=O)=O

PTP1B (Protein Tyrosine Phosphatase 1B) (20) 2003	O=C(OCC)C1=NOC(C2=CC(/C=C/CO)=CC=C2F)=C1	O=C(OC)C1=C(O)C=CC=C1O	[R]/C=C/CO[R]	O=C(O)C1=NOC(C2=CC(/C=C/COC3=CC=CC(O)=C3C(OC)=O)=CC=C2F)=C1
Bcl-2 (B-cell lymphoma) (21) 2010	OC(C1=CC=C(Cl)C(CCO)=C1)(C2=CC=C(C(C)(C)C)C=C2)CCN3CCCC3	O=C(O)C(C=C1)=CC=C1C2=CC=C(O)C=C2	[R]CCO[R]	OC(C1=CC=C(Cl)C(CCOC2=CC=C(C3=CC=C(C(O)=O)C=C3)C=C2)=C1)(C4=CC=C(C(C)(C)C)C=C4)CCN5CCCC5
***Mtb*PPAT (*Mycobacterium Tuberculosis* Phospho-pantetheine Adenylyl-Transferase) (22) 2023**	OCCCC1=CNC2=CC=CC=C21	O=C(OCC)C1=C(O)N(N=C1)C2=CC=CC=C2	[R]CCCO[R]	O=C(O)C1=C(OCCCC2=CNC3=CC=CC=C32)N(N=C1)C4=CC=CC=C4
HsNMT1 (Human N-myristoyl-transferase 1) (23) 2018	OCCC1=C(C)N(C)N=C1C	O=CC1=NN(C)C2=C1C=C(C3=CC=C(F)C(F)=C3O)C=C2	[R]CCO[R]	CN(C)CC1=NN(C)C2=C1C=C(C3=CC=C(F)C(F)=C3OCCC4=C(C)N(C)N=C4C)C=C2
RPA (Replication Protein A) (24) 2013	O=C(O)C1=CC=C(C2=CC=C(CBr)C=C2)O1	OC1=CC(C(O)=O)=NN1C2=CC(C)=CC=C2	[R]CO[R]	O=C(O)C1=CC=C(C2=CC=C(COC3=CC=C(C4=CC(C(O)=O)=NN4C5=CC(C)=CC=C5)C=C3)C=C2)O1
NADK (Nicotinamide Adenine Dinucleotide Kinase) (25) 2020	NC1=NC=NC2=C1N=C(Br)N2[C@H]3[C@H](OC(C)=O)[C@H](OC(C)=O)[C@@H](COC(C)=O)O3	NC1=NC=NC2=C1N=CN2[C@H]3[C@H]4[C@H](OC(C)(C)O4)[C@@H](COCC#C)O3	[R]C#CCO[R]	O[C@@H]1[C@@H](COCC#CC2=NC3=C(N=CN=C3N)N2[C@H]4[C@H](O)[C@H](O)[C@@H](CO)O4)O[C@@H](N5C(N=CN=C6N)=C6N=C5)[C@@H]1O
Bcl-X_L (B-cell lymphoma extra large) (26) 2014	O=C(C1=C(CN(C2=NC(C(OCC)=O)=C(CCCl)S2)CC3)C3=CC=C1)NC4=NC5=C(S4)C=CC=C5	OC1=CC=C(N2N=CC3=C2N=CN=C3)C=C1	[R]CCCO[R]	O=C(C1=C(CN(C2=NC(C(O)=O)=O)=C(CCCOC3=CC=C(N4N=CC5=C4N=CN=C5)C=C3)S2)CC6)C6=CC=C1)NC7=NC8=C(S7)C=CC=C8
BACE1 (Beta-secretase 1) (27) 2019	OC1=CC=C(/C=C/C(O)=O)C=C1	O=C1C2=C(O)C=C(OCCO)C=C2OC(C3=CC4=C(OC(C5=CC=CC=C5)(C6=CC=CC=C6)O4)C=C3)=C1	O=C([R])OCCO[R]	OC(C=C1)=C(O)C=C1C2=CC(C3=C(O)C=C(OCCOC(/C=C/C4=CC=C(O)C=C4)=O)C=C3O2)=O
NADPH (Nicotinamide Adenine Dinucleotide Phosphate) (28) 2020	OC1=CC=NC2=CC=CC=C21	OC1=CC=NC2=CC=CC=C21	[R]OCCOCCOCCOCCO[R]	NC1=NC2=CC=CC=C2C(OCCOCCOCCOCCOC3=CC(N)=NC4=C3C=CC=C4)=C1
Tryptase (29) 1999	NC(C1=CC=C(C(Cl)=O)C=C1)=N	NC(C1=CC=C(C(Cl)=O)C=C1)=N	O=C([R])NCC(C=C1)=CC=C1OCCCCCOC(C=C2)=CC=C2CNC([R])=O	N=C(N)C1=CC=C(C(NCC2=CC=C(OCCCCCOC3=CC=C(CNC(C4=CC=C(C(N)=N)C=C4)=O)C=C3)C=C2)=O)C=C1
PtS (Pantothenate Synthetase) (30) 2010	O=C(O)C1=CC2=CC=CC=C2O1	COC1=CC(C=C(CN)N2)=C2C=C1	O=C([R])NC[R]	O=C(NCC(N1CC(O)=O)=O)=CC2=C1C=CC(OC)=C2)C3=CC4=CC=CC=C4O3

Bid (BH3 Interacting Domain death agonist) (31) 2006	O=C(CCCN)NC(C=C1)=CC=C1SC2=CC=C([N+]([O-])=O)C=C2	OC1=CC=C(C(O)=O)C(O)=C1	[R]C(NCCCC(N[R])=O)=O	OC1=CC=C(C(NCCCC(NC(C=C2)=CC=C2SC3=CC=C(N)C=C3)=O)=O)C(O)=C1
LDHA (Lactate De-Hydrogenase A) (32) 2012	O=C(CCN)NC1=CC2=C(C=C1)N=C(SCCC)S2	O=C(OC)C(C(OC)=O)OC1=CC=C(CCCC(O)=O)C=C1	[R]CCCC(NCC[R])=O	O=C(O)C(C(O)=O)OC1=CC=C(CCCC(NCCC(NC2=CC3=C(C=C2)N=C(SCCC)S3)=O)=O)C=C1
CK2α (Casein Kinase 2) (33) 2016	O=CC1=CC(Cl)=C(C2=CC=CC=C2)C=C1	O=C(CCNC(CCN)=O)NC1=CC(C(OC)=O)=CC=C1	[R]NCCC(NCCC([R])=O)=O	ClC1=C(C2=CC=CC=C2)C=CC(CNCCC(NCCC(NC3=CC=CC(C(O)=O)=C3)=O)=O)=C1
B-Raf^{V600E} (B-Rapidly Accelerated Fibro-sarcoma) (34) 2018	O=C(NC1=CC=CC(C(O)=O)=C1)NC2=CC(C(F)(F)F)=C(Cl)C=C2	O=C(NC1=CC(C(F)(F)F)=C(C=C1)Cl)NC2=CC=CC(C(O)=O)=C2	[R]NC([R])=O	COC(C=C1)=CC=C1N2N=CC(NC(C3=CC(NC(NC4=CC(C(F)(F)F)=C(Cl)C=C4)=O)=CC=C3)=O)=C2
CypD (Cyclophilin D) (35) 2019	C[Si](C(C)(C)C)(C)O[C@@H]1CO[C@H]2C3=C(C=CC(CN)=C3)N[C@@H]1C2	O=C1N(C([C@H]2[C@@H]3C[C@@H](CC3)[C@@H]12)=O)CC(O)=O	[R]NC([R])=O	O[C@@H]1CO[C@H]2C3=C(C=CC(CNC(CN(C4=O)C([C@@H]5[C@@H]6C[C@@H](CC6)[C@H]45)=O)=O)=C3)N[C@@H]1C2
HO-1 (Heme Oxygenase-1) (36) 2021	O=C(O)C1=CC2=CC(OCC3=CC=CC=C3)=CC=C2N1C	NCCCN1C=CN=C1	O=C(NCCC[R])[R]	O=C(C(N1C)=CC2=C1C=CC(OCC3=CC=CC=C3)=C2)NCCCN4C=CN=C4
HuR (Human antigen R) (37) 2023	CC(C)(C)C1=CC=C(C(O)=O)C=C1	O=C(OC)C1=CC2=CC(N)=CC=C2N1	[R]NC([R])=O	CC(C)(C)C1=CC=C(C(NC2=CC=C(N=C(C(NO)=O)C3)C3=C2)=O)C=C1
IMPDH (Inosine Monophosphate Dehydrogenase (38) 2018	ClC([C@H](C)OC1=CC=C(C2=CNC=N2)C=C1)=O	BrC1=CC=C(C2=CNC(N)=N2)C=C1	[R]O[C@H](C(N[R])=O)C	CN1C=NC(C2=CC=C(O[C@@H](C)C(NC3=NC=C(C4=CC=C(Br)C=C4)N3)=O)C=C2)=C1
23S rRNA (39) 2002	O=C([C@H](NC(C1=CC=C(C2=CC=CC(N)=C2)O1)=O)CCN)N3CCNCC3	OC(C1=CC=C(NC(C(N2)=O)=O)=O)C2=C1)=O	[R]C1=CC=CC(NC([R])=O)=C1	O=C([C@H](NC(C1=CC=C(C2=CC=CC(NC(C3=CC=C(NC(C(N4)=O)=O)C4=C3)=O)=C2)O1)=O)CCN)N5CCNCC5
HSP90 (Heat Shock Protein) (40) 2007	O=S(C(C=C1)=CC=C1[N+]([O-])=O)(Cl)=O	CC1=CC(CN)=NC(N)=N1	[R]S(=O)(NC[R])=O	O=C(OCC/1)C1=C\NC2=CC=C(S(=O)(NCC3=NC(N)=NC(C)=C3)=O)C=C2
PS (Pantothenate Synthetase) (41) 2009	O=S(C1=CC2=CC=CC=C2O1)(N)=O	COC1=CC(C=C(C(O)=O)N2CC(OC(C)(C)C)=O)=C2C=C1	[R]C(NS([R])(=O)=O)=O	COC1=CC(C=C(C(NS(C2=CC3=CC=CC=C3O2)(=O)=O)=O)=O)N4CC(O)=O)=C4C=C1
Bcl-X$_L$ (B-cell lymphoma extra large)	O=C(C1=CC(C2=CC=C(Cl)C(Cl)=C2)=CC=C1)O	O=S(C1=CC=C(COC2=CC=CC=C2)C=C1)(N)=O	O=C([R])NS([R])(=O)=O	O=C(NS(C1=CC=C(OC2=CC=C(OCC)C=C2)C=C1)(=O)=O)C3=CC=CC(C4=CC(Cl)=C(Cl)C=C4)=C3

Mcl-1 (Myeloid Cell Leukemia) (42) 2011				
LDH-A (Lactate Dehydrogenase A) (43) 2013	O=C(OCC)C1=CN=C(SCC(NC2=CC(Cl)=C(O)C=C2OC)=O)C=C1	NC1=CC(F)=CC(C2=CC=C(C(O)=O)C=N2)=C1	[R]OC[C@H]([C@H]([C@@H]([C@@H](CN[R])O)O)O)O	O=C(O)C1=CN=C(SCC(NC2=CC(Cl)=C(OC[C@@H](O)[C@@H](O)[C@H](O)[C@H](O)CNC3=CC(C4=CC=C(C(O)=O)C=N4)=CC(F)=C3)C=C2OC)=O)C=C1
EthR (Ethionamide Resistance) (44) 2014	O=S(C1=CC=CC=C1)(NCC#C)=O	IC1=CC=C(C(NCCC(F)(F)F)=O)C=C1	[R]C#CC[R]	O=S(C1=CC=CC=C1)(NCC#CC2=CC=C(C(NCCC(F)(F)F)=O)C=C2)=O
EthR (Ethionamide Resistance) (45) 2014	S[C@H]1CCN(C(OC(C)(C)C)=O)C1	S[C@H]1CCN(C(OC(C)(C)C)=O)C1	[R]SS[R]	O=C(N1C[C@H](SS[C@H]2CCN(C(CCC3=CC=CC=C3)=O)C2)CC1)CCC4=CC=CC=C4
CypD (Cylophilin D) (46) 2016	NCC1=CC=C(NC(OC(C)(C)C)=O)C=C1	O=C(N1C(C2=CC=CC=C2SC)CCC1)C(N)C3=CC=CC=C3	[R]CNC(NC([R])C1=CC=CC=C1)=O	NC1=CC=C(CNC(NC(C2=CC=CC=C2)C(N3CCCC3C4=C(SC)C=CC=C4)=O)=O)C=C1
DENV-MTase (Dengue Virus Methyl Transferase) (47) 2017	O=[N+](C1=C(N=C=O)C=CC(C(F)(F)F)=C1)[O-]	O=C(O)C1=CC=C(C)C=C1N	[R]NC(N[R])=O	NC1=C(NC(NC2=CC(C)=CC=C2C(O)=O)=O)C=CC(C(F)(F)F)=C1
TPP (Thiamine Pyrophosphate) (48) 2022	NCC1=CC2=NC=CN=C2C=C1	BrC1=CN=CC=C1N2CCN(C(OC(C)(C)C)=O)CC2	[R]NC[R]	C12=NC=CN=C1C=C(CNC3=CN=CC=C3N4CCNCC4)C=C2

Table S31. Smiles of molecules listed in Table 2.

SMILES strings of fragments 1 and 2 and of the linked compound are reported.

Target	Fragment 1	Fragment 2	Best linked compound
Carbonic Anhydrase II (56) 1997	O=CC1=CC=C(S(=O)(N)=O)C=C1	NCC1=CC=CC=C1	)=S(C1=CC=C(CNCC2=CC=CC=C2)C=C1)(N)=O
Neuramidase (57) 2003	NC1C(NC(C)=O)C(N)=CC(C(O)=O)C1	CCC(CC)=O	C1C(NC(C)=O)C(NC(CC)CC)=CC(C(O)=O)C1
Cyclin-dependent kinase 2 (CDk2) (54) 2003	ClC1=C(NN)C=CC=C1	O=C(C1=CC(OC(F)(F)F)=CC=C1N2)C2=O	ClC1=C(N/N=C(C2=CC(OC(F)(F)F)=CC=C2N3)\C3=O)C=CC=C1 ClC1=C(N/N=C(C2=CC(OC(F)(F)F)=CC=C2N3)\C3=O)C=CC=C1
Endothiapepn (58) 2014	NNC([C@H](N)CC1=CN=C2C=CC=CC12)=O	O=CC(C(C)=CC(C)=C1)=C1C	O=C([C@H](N)CC1=CN=C2C=CC=CC12)N/N=C\C(C(C)=CC(C)=C3)=C3C
Glutathione-S-transferase (GST) (59) 2012	O=C([C@H](C)NC(CC1=CC(CC(N[C@H](C)C(NN)=O)=O)=CC=C1)=O)NN	ClC1=C([N+]([O-])=O)C=C([N+]([O-])=O)C=C1	O=C([C@H](C)NC(CC1=CC(CC(N[C@H](C)C(N/N=C\C2=CC([N+]([O-])=O)=C(Cl)C=C2)=O)=O)=CC=C1)=O)N/N=C/C3=CC([N+]([O-])=O)=C(Cl)C=C3
Endothiapepsin (62) 2016	C#CC1=CC=C(N2CCNCC2)C=C1	[N-]=[N+]=NCCC1=CC=C(O)C=C1	OC(C=C1)=CC=C1CCN(N=N2)C=C2C3=CC=C(N4CCNCC4)C=C3
G-quadruplexe (for c-MYC gene promoter) (63) 2020	C#CCN1CCN(C2=CC=C(N)C(N)=C2)CC1	N=[N]=NCCCN(C)C	NC1=CC(N2CCN(CC3=CN(CCCN(C)C)N=N3)CC2)=CC=C1N
human acetyl and butyryl choline esterase (*h*AChE) (*h*BChE) (64) 2021	[H][C@]12C[C@H](C=C[C@]13CCN(CCCCC#C)CC4=C3C(O2)=C(C=C4)OC)O	O=C(CCCCN=[N+]=[N-])NCCC1=CNC2=CC=C(OC)C=C21	[H][C@]12C[C@H](C=C[C@]13CCN(CCCCC4=CN(CCCCC(NCCC5=CNC6=CC=C(OC)C=C65)=O)N=N4)CC7=C3C(O2)=C(C=C7)OC)O
Histone deacetylase (HDAC) (65) 2010	O=C(NO)C(S1)=CC=C1C#C	[N-]=[N+]=NCC1(C2)C[C@@H](C[C@@H]2C3)C[C@@H]3C1	O=C(NO)C(S1)=CC=C1C2=CN=NN2CC3(C4)C[C@@H](C[C@@H]4C5)C[C@@H]5C3

Carbonic Anhydrase II (66) 2005	C#CC1=CC=C(S(=O)(N)=O)C=C1	O=C([C@H](C(C)C)N=[N+]=[N-])N(C1)C[C@H]2CC[C@@H]1CC2	)=S(C1=CC=C(C2=CN([C@@H](C(C)C)C(N(:3)C[C@H]4CC[C@@H]3CC4)=O)N=N2)C=C1)(N)=O
HIV-1 Protease (67) 2006	O=S(C1=CC=C(OC)C=C1)(N(CC(C)C)C[C@@H](O)[C@H](CC2=CC=CC=C2)N=[N+]=[N-])=O	O[C@H](C1)[C@@H](NC(OCC#C)=O)C2=C1C=CC=C2	[C@H](C1)[C@@H](NC(OCC#C)=O)C2=C1C=CC=C2
Acetylcholine binding protein (68) 2012	CN1C2=C(C=CC=C2)C(OCC#C)=CC1=O	C[N+]1(C)[C@@H](C2)CC[C@H]1C[C@@H]2N=[N+]=[N-]	O=C(N1C)C=C(OCC2=CN(N=N2)[C@H]3C[C@@H]4CC[C@H](C3)[N+]4(C)C)C5=C1C=CC=C5
Abl Tyrosine Kinase (69) 2013	CC1=CC=C(N=[N+]=[N-])C=C1NC2=NC=CC(C3=CN=CC=C3)=N2	C#CC1=CC=C(N2CCN(C)CC2)C=C1	CC1=CC=C(N(N=N2)C=C2C3=CC=C(N4CCN(C)CC4)C=C3)C=C1NC5=NC=CC(C6=CN=CC=C6)=N5
Coagulation factor Xa (71) 2017	O=S(CC1=CC=CC=C1)(N[C@H](COCC2=CC=CC=C2)C(NCC(O)=O)=O)=O	NCC1=CC=C(C(N)=N)C=C1	O=S(CC1=CC=CC=C1)(N[C@H](COCC2=CC=CC=C2)C(NCC(NCC3=CC=C(C(N)=N)C=C3)=O)=O)=O
c-Src kinase (72) 2015	O=C(CCS)NC1=CC=CC(NC2=NC=CC(NC3=NNC(C4CC4)=C3)=N2)=C1	C=CC(N1CCN(C)CC1C2=CC=CC=C2)=O	O=C(CCSCCC(N1CCN(C)CC1C2=CC=CC=C2)=O)NC3=CC=CC(NC4=NC=CC(NC5=NNC(C6CC6)=C5)=N4)=C3
Carbonic Anhydrase II (73) 2022	O=S(C1=CC=C(CS)C=C1)(N)=O	O=C(C=C1)OC2=C1C=CC(C#C)=C2	O=S(C1=CC=C(CS/C=C\C2=CC(OC(C=C3)=O)=C3C=C2)C=C1)(N)=O
Bcl-XL and Mcl-1 kinases (74) 2023	[O-][N+](C1=CC(S(=O)(N=[N+]=[N-])=O)=CC=C1NCCSC2=CC=CC=C2)=O	O=C(S)C1=CC=C(C=C1)CCC2=NC=CC=C2	O=C(C1=CC=C(C=C1)CCC2=NC=CC=C2)NS(C3=CC=C(C([N+]([O-])=O)=C3)NCCSC4=CC=CC=C4)(=O)=O
Endothiapepsin (76) 2020	N[C@H](C(O)=O)CC1=CNC2=C1C=CC=C2.NCCC3=CN=CN3.[C-]#[N+]CC4=CC=CC=C4.O=CCCC5=CC=CC=C5		N[C@H](C(N(C(C(NCC1=CC=CC=C1)=O)CCC2=CC=CC=C2)CCC3=CN=CN3)=O)CCC4=CNC5=C4C=CC=C5
Urokinase plasminogen activator (uPA) (77) 2019	NC1=NC=CC(C(NC2CC2)=O)=C1.[H]C=O.N=C(N)NC(C=C3)=CC=C3C[N+]#[C-]		O=C(NC1CC1)C2=CC3=NC=C(NCC4=CC=C(NC(N)=N)C=C4)N3C=C2