

RESEARCH ARTICLE OPEN ACCESS

Fragment-Based Design of Targeted Covalent Inhibitors: The Scope and Limitation of Linking Approaches

Levente Kollár^{1,2}  | Levente M. Mihalovits¹  | Dávid Bajusz¹  | Damijan Knez³  | József Simon^{1,4}  | Blake H. Balcomb^{5,6}  | Daren Fearon⁵  | Stanislav Gobec³  | György M. Keserű^{1,2} 

¹Medicinal Chemistry Research Group and Drug Innovation Centre, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary | ²Department of Organic Chemistry and Technology, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Budapest, Hungary | ³Department of Medicinal Chemistry, Faculty of Pharmacy, University of Ljubljana, Ljubljana, Slovenia | ⁴MS Metabolomics Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary | ⁵Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK | ⁶Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, UK

Correspondence: György M. Keserű (keseru.gyorgy@ttk.hu)

Received: 21 December 2025 | **Revised:** 26 March 2026 | **Accepted:** 30 April 2026

Keywords: 3CL^{PRO} inhibitors | covalent inhibitors | fragment linking | fragment-based drug discovery | SARS-CoV-2

ABSTRACT

Linking of fragments in neighboring binding sites is one of the optimization strategies in fragment-based drug discovery, where additive or even more substantial bioactivity improvements can be realized. However, such efforts present a considerable challenge when one fragment binds covalently to the target protein, as small modifications can influence the correct positioning of the covalent warhead toward the targeted nucleophilic residue. Here, we present a case study of fragment linking that yielded single-digit micromolar, covalent inhibitors of the SARS-CoV-2 main protease, starting from fragments that were inactive in the biochemical assay. Using structural information from a recent, high-throughput crystallographic fragment screen, we show that the success of fragment linking in the design of targeted covalent inhibitors is heavily impacted by several factors, including the warhead type, the labeling chemistry, and even subtle changes in the designed linker. Notably, we observe that induced fit effects might override the original fragment orientations in the linked molecule, highlighting the need for reliable structure verification, especially in consecutive rounds of fragment elaboration.

1 | Introduction

The first successes of fragment-based drug discovery (FBDD) date back to the late 1990s [1–3]. The potential of FBDD was quickly recognized by the pharmaceutical industry and has been used widely as a tool for hit discovery over the past two decades [4]. Fragments are compounds with small molecular weight and relative chemical simplicity. To describe their general physicochemical parameters, Congreve and colleagues coined the

Rule of 3 (Ro3), which includes the molecular weight (≤ 300 Da), the number of hydrogen bond donors and acceptors (≤ 3 for both separately), and the calculated logP (≤ 3) [5]. The concept has been widely adopted; however, subsequent fragment screening experiments revealed that additional molecular descriptors can be considered to increase the success rate of FBDD. Although compounds that match Ro3 are not necessarily useful fragments, Ro3 provides a preliminary tool to eliminate molecules with undesired properties [6–8]. It is worth noting that

Abbreviations: 3CL^{PRO}, 3-chymotrypsin-like protease; DABCO, 1,4-diazabicyclo[2.2.2]octane; DCC, N,N'-dicyclohexylcarbodiimide; DCM, dichloromethane; DIPEA, N,N-diisopropylethylamine; DMF, N,N-dimethylformamide; DMSO, dimethylsulfoxide; FBDD, fragment-based drug discovery; HATU, Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium; IFD, induced fit docking; M^{PRO}, main protease; MS, mass spectrometry; NHS, N-hydroxysuccinimide; NMR, nuclear magnetic resonance; PDA, photodiode array; PDB, Protein Data Bank; RA, remaining activity; RMSD, Root Mean Square Deviation; SAR, structure-activity relationship; SPhos, dicyclohexyl[2',6'-dimethoxy[1,1'-biphenyl]-2-yl]phosphane; TFA, trifluoroacetic acid; THF, tetrahydrofuran.

Levente Kollár and Levente M. Mihalovits contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *ChemMedChem* published by Wiley-VCH GmbH.

other, at times simpler definitions have also been used to describe fragment-like chemical space (i.e., compounds between 10 and 16 heavy atoms) [9].

The first step in an FBDD campaign is to identify fragment-sized hits. The small size of the molecules under investigation means that they typically have a weak affinity for the target due to the limited possibilities of protein-ligand interactions and the limited complementarity with the binding site. Therefore, the screening methods used have been adapted accordingly to overcome this shortcoming. Techniques that are sufficiently sensitive and thus most commonly used for screening include nuclear magnetic resonance (NMR) [10], X-ray crystallography [11], and surface plasmon resonance (SPR) [12]. Additionally, other approaches, such as differential scanning fluorimetry (DSF), isothermal titration calorimetry (ITC), and computational methods offer useful alternatives [13, 14]. The main advantage of the fragment-based approach comes from the core idea behind the method: by screening molecules of limited size, we can increase the probability of a binding event and cover the chemical space more efficiently with a carefully assembled library [4, 13, 15, 16]. This — together with recent, highly developed screening techniques and a wealth of accumulated experience — significantly improves the success and efficiency of hit-finding attempts. However, the maturation of hits into lead molecules is far more difficult and time-consuming, as illustrated by the small (but growing) share of FBDD-related lead optimization strategies in literature data [17].

There are three general strategies for the optimization process, namely fragment growing, fragment merging, and fragment linking (Figure 1) [4, 18]. Fragment growing is by far the most widely used technique and involves the optimization of a hit molecule, including the insertion and/or substitution of various chemical groups to improve affinity. Fragment merging can be applied when two identified hits have an overlapping binding site, while fragment linking addresses the case when two hits are in close proximity and have adjacent binding sites [14].

As of January 2020, 46 molecules have progressed to at least Phase I clinical trials, where a fragment-based strategy played a key role in their discovery. Out of these, 41 are the result of growing, but only 4 of linking, and 1 of merging strategies [19]. It is also noteworthy that dozens of successful fragment-to-lead research projects have been published in recent years, and the number of FBDD programs in the clinical stage has reached 59 by 2024 [20], including 7 that already provided approved drugs. It has been applied to a wide range of targets such as kinases, proteases, G-protein coupled receptors (GPCRs), protein–protein interactions (PPIs), and other targets [21–28].

The concept of fragment linking is elegant and simple, but it is considered the most challenging strategy to implement. It holds the promise of a leap in affinity; however, designing an appropriate linker that does not disrupt the original binding modes of fragments is particularly challenging [18]. This difficulty is even greater when the distance between the two fragments is large (>3 Å) [19]. In addition, linking can lead to an altered binding mode, protonation/tautomeric state of the fragments, internal strain energy of the linked molecule, or unfavorable interactions between the target and the linker. Moreover, the linkage might decrease the relative configurational entropy of the fragments, which was assumed to be the main reason for the failure of most fragment linking strategies [29]. Despite all discussed difficulties, there are success stories, such as the discovery of venetoclax, a selective Bcl-2 inhibitor, which was approved to treat chronic lymphocytic leukaemia [30, 31]. Besides that, a number of leads have been generated by fragment linking [19].

At the dawn of fragment-based drug discovery, fragment libraries were specifically designed to exclude molecules that could potentially form covalent bonds with the target proteins [14]. However, around the same time, the potential of covalent inhibitors was recognized, which in turn has influenced fragment-based drug discovery. Nowadays, covalent fragments are an integral part of the drug discovery toolbox and play an important role in target identification and validation, as well as hit identification [32–34].

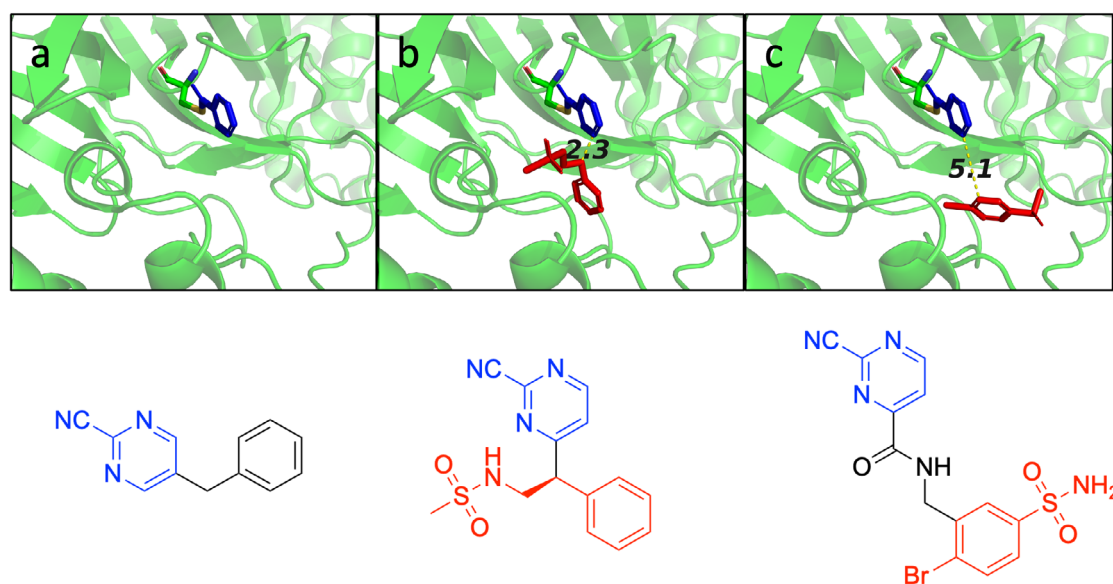


FIGURE 1 | Strategies for fragment development: (a) growing, (b) merging, and (c) linking. The examples involve 3CL^{pro} versus fragment structures investigated in this study, with interatomic distances between the linkage points displayed in Angstroms.

Although rarely selective, they can possess relatively high affinity compared to non-covalent fragments. During a covalent fragment-to-lead campaign, affinity and selectivity should be carefully optimized together. One of the greatest advantages of the application of covalent FBDD is that it provides hits for targets previously thought to be “undruggable”. An example of that is the discovery of sotorasib—a selective KRas^{G12C} inhibitor—and its preceding hit [35–37].

Covalent drug discovery projects can follow two different strategies, widely termed as “ligand-first” and “electrophile-first” approaches. The former means the purposeful addition of reactive functional groups to existing non-covalent ligands, while the latter consists of screening covalent ligands and optimizing the non-covalent core or even by replacing the warhead with another [34]. Both can be effective, as demonstrated by examples: the ligand-first approach afforded epidermal growth factor receptor (EGFR) inhibitors (e.g., osimertinib, rociletinib) [38, 39] and the Bruton’s tyrosine kinase (BTK) inhibitor ibrutinib [40, 41], while the electrophile-first approach resulted in the aforementioned KRas^{G12C} inhibitor sotorasib [35–37] and a SARS-CoV-2 main protease (M^{pro} or 3CL^{pro}) inhibitor, nirmatrelvir [42].

We note that, as in non-covalent fragment-to-lead optimization, growing is the most common strategy for covalent fragment hit maturation, with merging sometimes being justified [43]. Examples of fragment linking are scarce; however, in a recent study, an electrophilic fragment screen of two hundred acrylate methyl esters resulted in a selective binder of BRD4-BD2 that was linked via a PEG linker to a known non-covalent BRD4-BD1 inhibitor, which resulted in a covalent BRD4 inhibitor with cellular activity [44]. Considering that the linked motifs bind at different domains of the target and one of them is not a fragment at all, this study cannot be considered as a typical fragment linking experience.

Therefore, we were prompted to gain insight into covalent fragment linking based on a recent fragment screening campaign against the SARS-CoV-2 main protease, 3CL^{pro} [45], a primary target for small-molecule drug discovery against SARS-CoV-2. The main protease cleaves its substrate at eleven sites to release non-structural proteins (nsp4–16), which form the replicase complex responsible for replication and transcription of the viral genome [45, 46]. 3CL^{pro} is a highly conserved cysteine protease [47–50], and contains a catalytic His41-Cys145 dyad at the active site. Therefore, it holds great potential for the development of covalent ligands targeting Cys145 [51], as demonstrated by its only approved inhibitor, nirmatrelvir. It is important to add that multiple mutations have recently been identified in 3CL^{pro} from naturally occurring viruses that are resistant to nirmatrelvir [52]. For this reason alone, it is important to identify a new generation of 3CL^{pro} inhibitors that can help us defend against current and future coronavirus infections [53, 54]. 3CL^{pro} inhibitors will continue to attract significant interest, as evidenced by the clinical progress of leritrelvir [55], simnotrelvir [56], and ensitrelvir [57]. The XChem facility at Diamond Light source screened over 1250 fragments against the 3CL^{pro} enzyme of SARS-CoV-2 using a combination of mass spectrometry and X-ray. As a result, 74 high-value fragment hits were found, including 23 non-covalent and 48 covalent hits in the active site, and 3 hits at the vital dimerization interface [45].

2 | Results and Discussion

2.1 | Design and Synthesis

The XChem screening campaign included our Covalent MiniFrag [58] library of five- and six-membered nitrogen heterocycles with various substituents (nitrile, ethynyl, and vinyl groups and halogens) [59]. Two hits were found, pyrimidine-2-carbonitrile and imidazole-2-carbonitrile (PDB codes: 5RHB and 5RHC, respectively) [45], both of which were able to bind to Cys145 by forming a thioimide (Figure 2D). With the structural information provided by the X-ray diffraction study in hand, we started a fragment-based drug design effort to investigate whether these covalent fragments could be linked to other, non-covalent fragment hits that were identified.

First, we extracted the non-covalent fragments from their respective X-ray structures after structural alignment. Hits that occupy feasible positions for linking with the covalent heterocycles were selected based on various criteria and by visual inspection (Figure S1 and 2). To design linkers between the selected covalent and non-covalent fragments, DeLinker — a graph-based deep generative method — was utilized [60]. The main advantage of DeLinker is that it incorporates 3D structural information directly into the design process; therefore, the structural alignment of the linked molecular parts is conserved, or at least the linking is forced to fulfill the positional constraints defined by the X-ray structures. Also, applying a dedicated tool for linker design ensures an optimal exploration of the available linker space in a more time-efficient and reproducible way than designing linkers by hand. Next, covalent docking was performed for the linked molecules, and compounds with the best docking scores and RMSD values calculated for the non-covalent fragment parts were selected for visual inspection and synthetic feasibility assessment (Figure S3). Note that the RMSD of the non-covalent fragment was selected as the primary evaluation metric, as correct positioning of the non-covalent core plays an essential role in the appropriate orientation of the warhead toward its binding partner. Therefore, suboptimal binding poses of the non-covalent core result in unproductive warhead orientations. After exploring the linker space and filtering by fragment pose conservation, candidates for synthesis were selected by an assessment of synthetic feasibility. As such, at the end of the fragment-based design workflow (Figure 2A), three compounds were chosen for synthesis (Figure 2C). Note that all three compounds consisted of the fragments extracted from the PDB entries 5RHB and 5RF1 (Figure 2E) [45].

Before synthesizing new molecules, the activities of the separate fragment molecules, pyrimidine-2-carbonitrile (**1**) and 4-bromobenzenesulfonamide (**2**), were determined against 3CL^{pro} in vitro at concentrations of 500 and 100 μM with 30 min of pre-incubation. Both were found practically inactive under the assay conditions (Figure 2B).

To obtain **3** and **4**, the appropriate building blocks were synthesized first (Scheme 1A). 1-Bromo-2-nitrobenzene (**6**) was treated with chlorosulfonic acid at 90°C for 5 h [61], and the **7** obtained was reacted with ammonia (0.4 M solution in THF) to furnish the sulfonamide **8**. To reduce the nitro group, tin(II) chloride dihydrate was utilized, and the amine (**9**) was obtained in a good

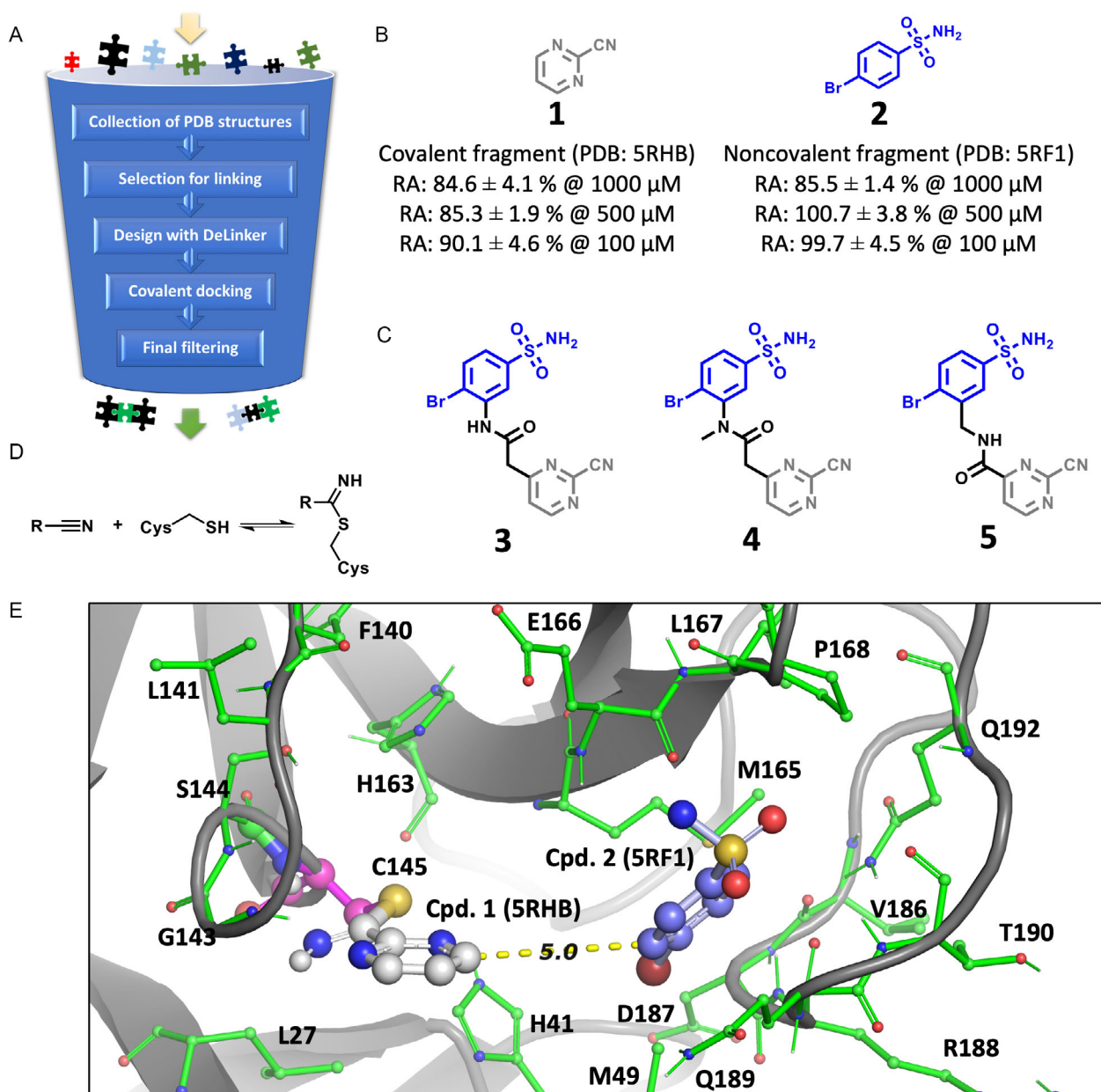


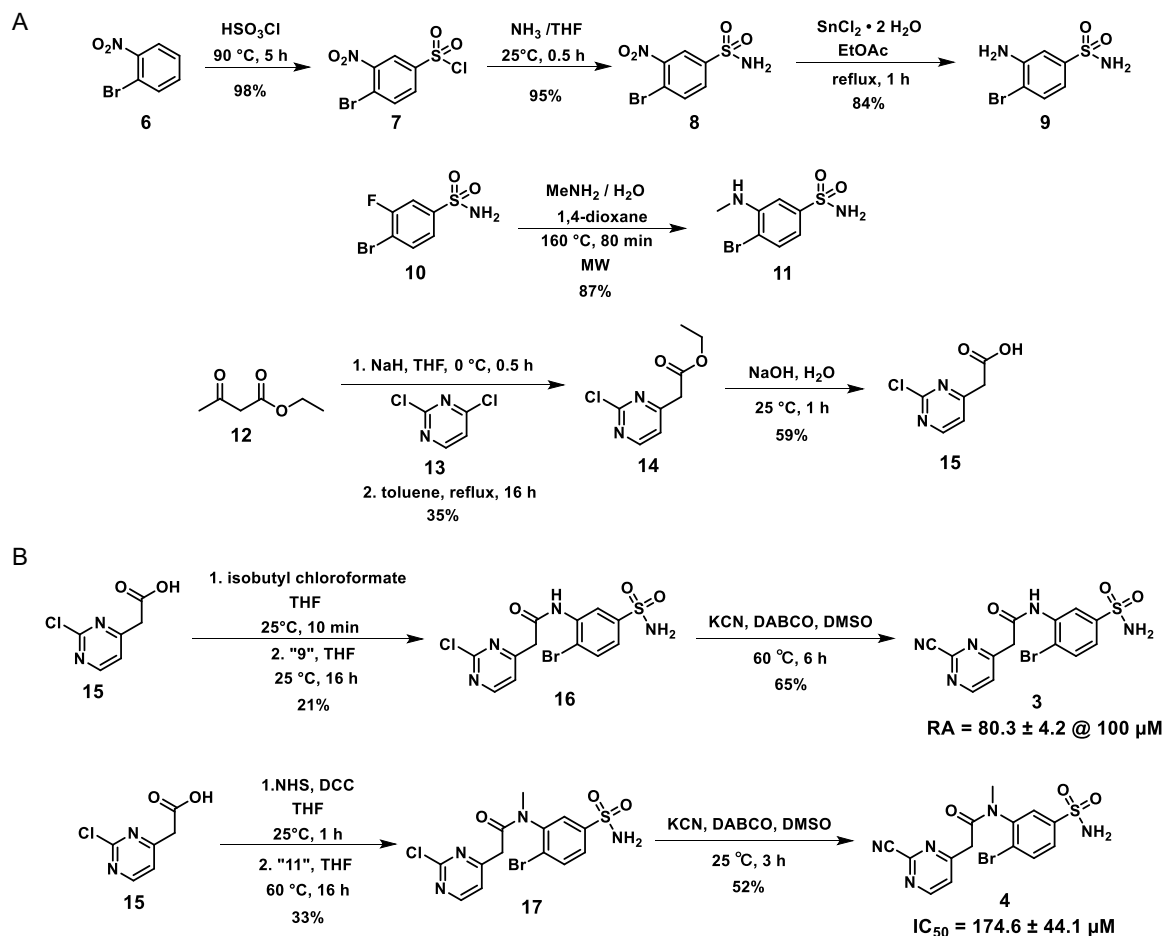
FIGURE 2 | (A) Applied fragment-based virtual screening workflow. (B) Residual activity (RA) values of the initial fragment hits, linked in the selected compounds. (C) Compounds selected for synthesis. (D) Reversible thioimide formation in the reaction of a nitrile and a cysteine residue. (E) Structural overlay of the fragments found in PDB structures 5RHB [45] and 5RF1 [45] (protein structure included from 5RHB).

yield. To get the *N*-methylated analog of the latter, a different approach was used. 3-Fluoro-4-bromobenzenesulfonamide (**10**) was reacted with methylamine in a solvent mixture of water and 1,4-dioxane. The reaction was conducted at 160°C in a microwave reactor to get **11**. To construct the other part of the target molecules, ethyl acetoacetate (**12**) was deprotonated with NaH, and then an aromatic nucleophilic substitution reaction was carried out with 2,4-dichloropyrimidine (**13**) [62]. Alkaline hydrolysis of the product (**14**) led to the appropriate acid (**15**).

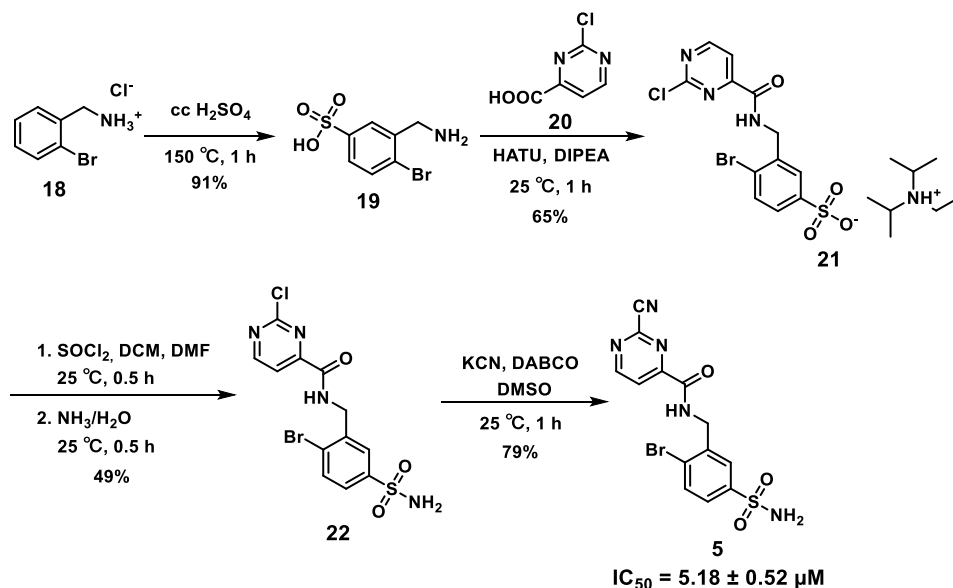
To acylate **9**, carboxylic acid **15** was activated with isobutyl chloroformate, and the reaction resulted in the formation of **16** with a low yield. In the last step, a nucleophilic substitution was done at 60°C with KCN and DABCO was used as a base to get one of the target molecules (**3**). For the *N*-acylation of **11**, compound **15** was

pre-activated as an NHS-ester, and the acylation reaction was done at 60°C overnight, yielding the amide (**17**). The last step was again a substitution with KCN to get **4** (Scheme 1B).

The third target molecule was prepared on a different synthetic route (Scheme 2). 2-Bromobenzylamine hydrochloride (**18**) was sulfonated at 150°C in 96% sulfuric acid to obtain compound **19**, which was purified by crystallization from 1,4-dioxane, and then reacted with 2-chloropyrimidine-4-carboxylic acid (**20**) using HATU as the coupling agent. Following chromatography, the resulting amide was obtained as a diisopropyl-ethylammonium salt (**21**). It was first converted to the appropriate sulfonyl chloride, which was not isolated, it was reacted with ammonia(aq) to furnish the respective sulfonamide (**22**). It was then converted to the desired compound (**5**) with a nucleophilic substitution



SCHEME 1 | Synthesis of **3** and **4**.



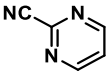
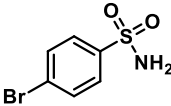
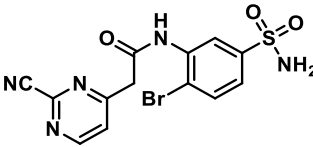
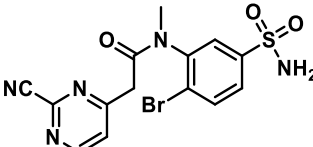
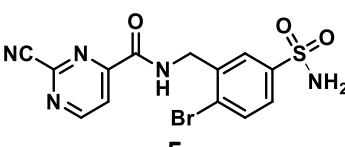
SCHEME 2 | Synthesis of **5**.

reaction with KCN; the reaction was much faster compared to the preparation of **3** and **4**.

The activity of the linked molecules is highlighted in Schemes 1 and 2, and the results are summarized in Table 1. Compound **3**

was found to be inactive, while **4** showed very weak activity. Compound **5** was the most potent analog with an IC_{50} value of $5.18 \pm 0.52 \mu M$. Interestingly, linking of **11** with another non-covalent fragment hit (**23**, $96.0 \pm 4.6\%$ residual activity at $100 \mu M$) using the same linker also resulted in an inactive

TABLE 1 | Summary of the measured biological activities against 3CL^{pro}, intact protein MS, and jump dilution assay results.

Compound	3CL ^{pro} RA, % or IC ₅₀ , μM	Intact protein MS	Jump dilution assay
 1	90.1 ± 4.6% @ 100 μM	labeling confirmed	n. d.
 2	99.7 ± 4.5% @ 100 μM	n. d.	n. d.
 3	80.3 ± 4.2% @ 100 μM	n. d.	n. d.
 4	174.6 ± 44.1 μM	n. d.	reversible inhibitor
 5	5.18 ± 0.52 μM	labeling confirmed	reversible inhibitor

n. d.: not determined.

compound (**24**, 89.3 ± 0.5% residual activity at 100 μM, Table 2, Supplementary Note 1).

Although non-peptidic covalent 3CL^{pro} inhibitors with double-digit nanomolar IC₅₀ are known [63], it is important to note that the optimization leading to these inhibitors has typically been based on a hit of similar or lower potency than **5**. It should be highlighted that the original fragment hits (**1** and **2**) were completely inactive. The compounds with IC₅₀s were evaluated in a jump dilution assay, and both of them were found to be reversible inhibitors, just as expected of the pyrimidine-2-carbonitrile warhead [64]. When incubated together with 3CL^{pro}, **5** covalently labelled the protein, **1** was used as a control (Figures S4 and 5). The results are summarized in Table 1.

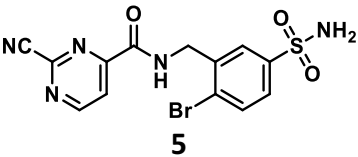
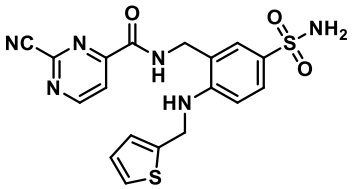
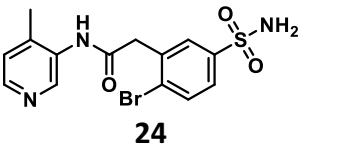
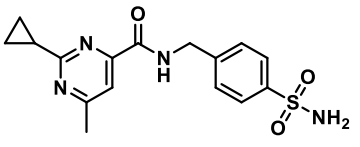
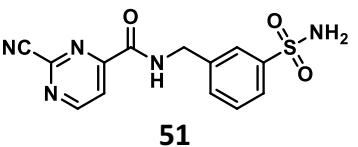
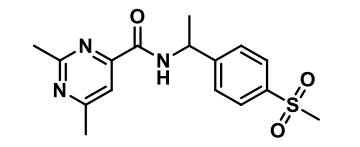
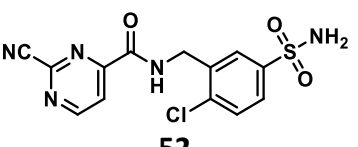
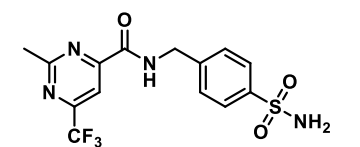
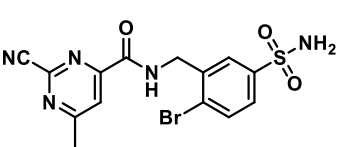
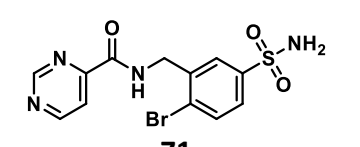
2.2 | Crystal Structure of the 5-3CL^{pro} Complex

With **5** in hand, we received a reversible covalent inhibitor with notable activity compared to the fragment hits. As a follow-up, we aimed to improve the activity of the core compound by subsequent design attempts. First, the crystal structure of the 5-3CL^{pro} complex was acquired (Figure 3A). Interestingly, the moiety adopted from the fragment of 5RF1 (**2**) adopts a “flipped” position

compared to its initial pose at the 3CL^{pro} binding site and the previously modeled covalent docking pose of **5** (Figure 3B), i.e., linking the covalent and non-covalent fragments did not result in a compound adopting the same orientation as its building blocks. However, the initial position of the sulfonamide moiety of the non-covalent fragment is occupied by a DMSO molecule in the experimental crystal structure, suggesting that the placement of a sulfoxide group is, in fact, still feasible in that position (Figure 3C).

Visual inspection of the linked molecules and fragments' crystal structure revealed the position of two residues, namely M49 and Q189, as important factors on the relative position of the non-covalent moiety. The position of the sidechains pointing towards the binding pocket is marked as “in”, the ones that point outside the pocket are denoted as “out” (Figure 3D). With this nomenclature, the M49(in)-Q189(out) position aids the orientation of the initial 5RF1 fragment and clashes with the orientation found for the crystallographic structure of the non-covalent moiety of **5**; whereas the M49(out)-Q189(in) promotes the latter orientation and clashes with the fragment of 5RF1. To test this hypothesis, we have obtained a more recent X-ray (PDB: 7RNK) [65] structure of 3CL^{pro} from the Fragalysis library [66] that captures the M49(out)-Q189(in) orientation, and we docked **5** into that particular structure covalently (notably, this sidechain orientation was

TABLE 2 | Summary of the measured biological activities of compounds obtained in the frame of “classic” SAR exploration.

Compound	3CL ^{pro} RA, % or IC ₅₀ , μM	Compound	3CL ^{pro} RA, % or IC ₅₀ , μM
 5 “Core compound”	5.18 ± 0.52 μM	 63	75.9 ± 7.6 μM
 24	89.3 ± 0.5% @ 100 μM	 68	97.9 ± 1.9% @ 100 μM
 51	226.3 ± 36.5 μM	 69	89.3 ± 0.5% @ 100 μM
 52	4.26 ± 0.22 μM	 70	90.4 ± 0.6% @ 100 μM
 59	114.2 ± 11.2 μM	 71	93.0 ± 6.1% @ 100 μM

not captured in any structures of the original fragment screen). The experimental binding pose of the linked molecule was reproduced with an RMSD of 1.74 Å by docking into a M49(out)-Q189(in) structure (Figure S6) instead of the crystal structure of the covalent fragment, which initially resulted in an RMSD of 7.95 Å. This corroborates our observation that there is a considerable induced-fit effect during ligand binding that involves sidechain movements of M49 and Q189.

2.3 | Structure-Activity Exploration Around the Core Compound 5

After the exhaustive structural analysis of the experimental crystal structure, we aimed to improve our core compound based on the abovementioned observations and further X-ray structures extracted from the Fragalysis 3CL^{pro} library. We wondered what kind of modification this structure could tolerate without a loss of activity, and whether we could improve its activity by making substitutions in the mapped regions. According to the crystal structure of **5**, two possible sites of modification were revealed: at the aromatic bromine substituent and the region of the sulfonamide group.

First, we have performed a “classic” SAR exploration of the identified regions; the measured activities of derivatives are summarized in Table 2. The compound tolerates the substitution of bromine to chlorine (**52**), but not the removal of the bromine atom (**51**) (Supplementary Note 2). Installing a larger apolar moiety (**63**) or supplementing the pyrimidine ring with a methyl group (**59**) (to accommodate a small cavity of the binding site) was also not tolerated (Supplementary Note 3). Commercial compounds selected and purchased (**68–70**) based on shape similarity to **5** did not yield any further hits, either (Supplementary Note 4).

Next, we investigated the effect of warhead modification. Here, we found that the removal of the warhead (**71**) resulted in a complete loss of activity, highlighting that the secondary interactions between the protein and ligand are not strong enough to achieve measurable potency, but are crucial for warhead positioning during covalent inhibition (Supplementary Note 5). For warhead substitution, we investigated the easily accessible 2-chloropyrimidine warhead. Since 2-chloropyrimidine derivatives were key intermediates in all syntheses of the compounds described so far, we investigated whether they react as irreversible covalent warheads [67]. We hypothesized that the electron-deficient aromatic ring activates the halogen for aromatic nucleophilic substitution

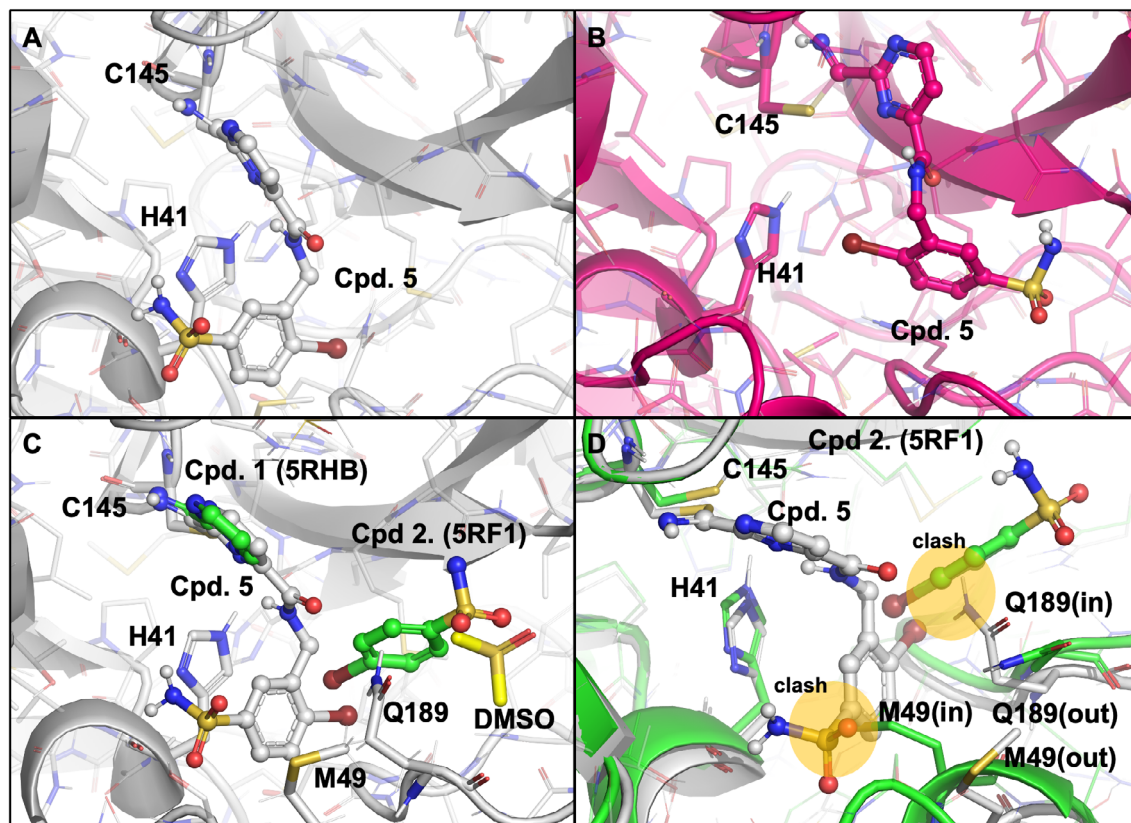


FIGURE 3 | (A) Crystal structure of the 5-3CL^{pro} complex. (B) Modeled covalent docking pose of 5. (C) Superposition of the crystal structure of the 5(gray)-3CL^{pro} complex and the two fragments extracted from PDB: 5RHB and 5RF1 (green). (D) Superposition of the 5-3CL^{pro} (gray) complex and PDB: 5RF1 (green) structures highlighting the alternative positions of M49 and Q189 (clashes between alternative ligand poses and enzyme environments are highlighted with orange). Throughout the figure, colors correspond to experimental fragment structures (green), experimental structure of cpd. 5 (white) and modeled structures (pink).

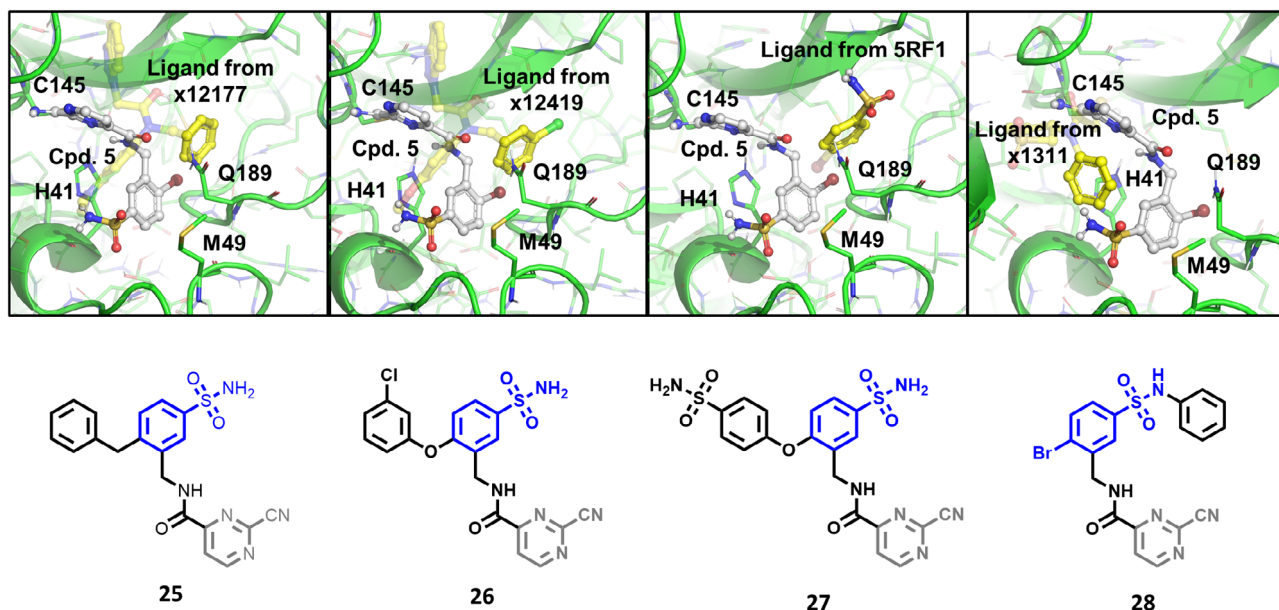


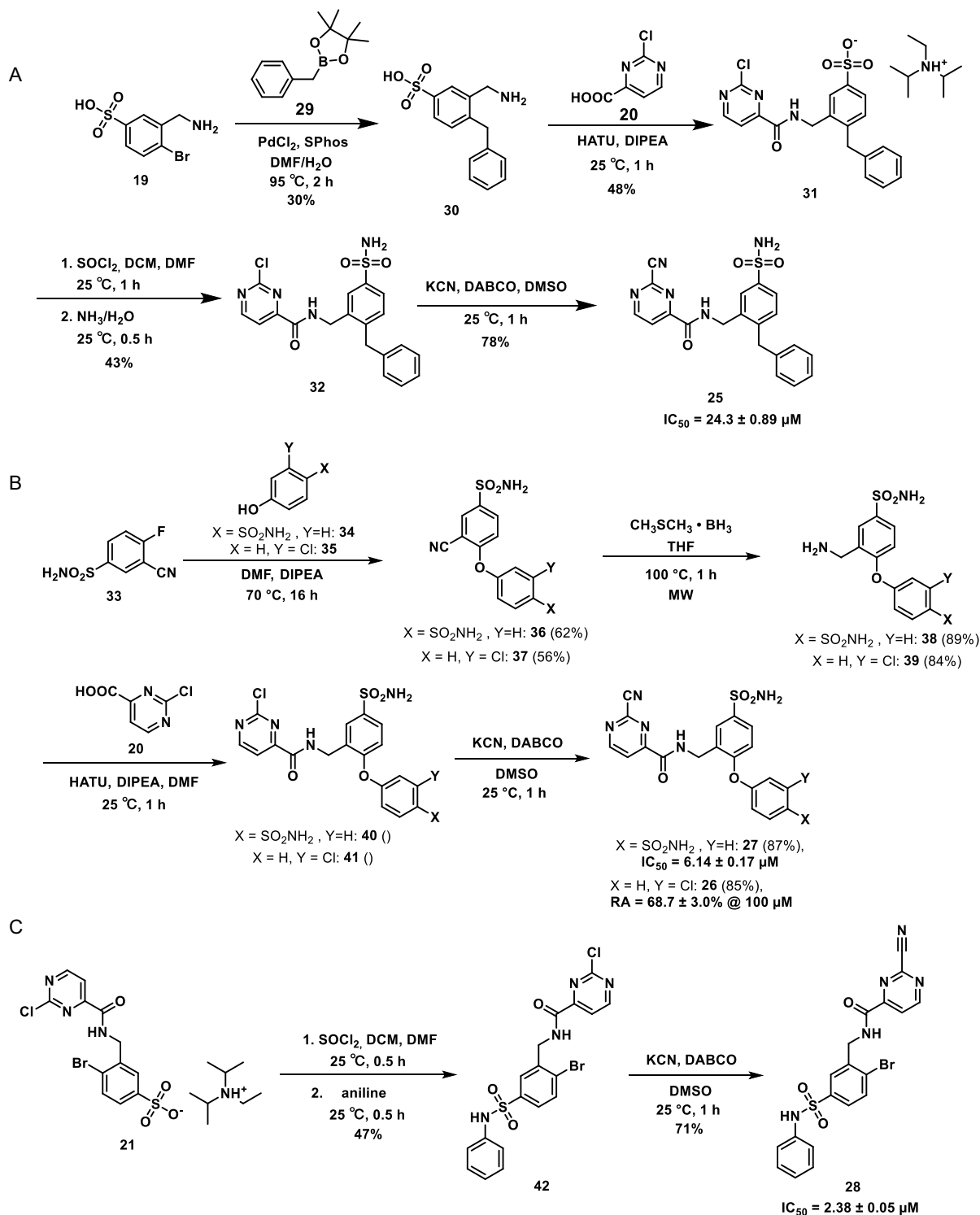
FIGURE 4 | Structure of designed analogs of 5 and the basis of the respective designs where the additional substituents are highlighted in full yellow ball and stick representation (labels refer to PDB IDs, where available, or Fragalysis IDs – see MPro among “Legacy targets” at <https://fragalysis.diamond.ac.uk/viewer/react/landing>).

that provides a covalent warhead with differences in reactivity, positioning, and labeling mechanism. Here, we identified two derivatives as irreversible covalent inhibitors, but their activities were two orders of magnitude weaker than **5**, due to the suboptimal orientation of the warhead (Supplementary Note 6).

Turning once again to our structure-based design protocol (this time for fragment merging, rather than linking), suitable substituents were selected by clustering, constrained docking

and visual inspection of the collected X-ray ligands. Here, instead of the complete molecules, only those moieties that can be reasonably attached to the core structure were taken into account, in a manner analogous to fragment merging. Three of the newly designed compounds were selected for synthesis (Figure 4).

To get **25** (Scheme 3A), the key step was the cross-coupling of **19** with benzyl boronic acid pinacol ester (**29**), using PdCl₂ as a



SCHEME 3 | Synthesis of **25** (A), **27**, **26** (B), and **28** (C).

catalyst and SPhos as a ligand, which afforded the desired reaction product (**30**). Cross-coupling was followed by acylation with 2-chloropyrimidine carboxylic acid (**20**), sulfonamide formation, and nucleophilic substitution with KCN to obtain **25**, similarly to the synthesis of **5**.

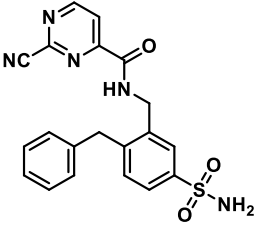
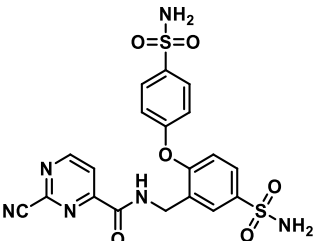
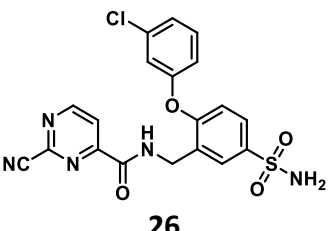
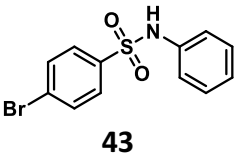
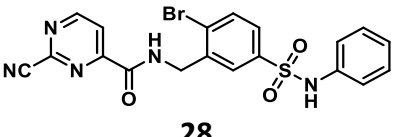
A new route was developed for the synthesis of **27** and **26** (Scheme 3B). In the first step, 3-cyano-4-fluorobenzensulfonamide (**33**) was reacted with the appropriate phenol derivatives (**34** and **35**). The S_NAr reaction needed 70°C and 16 h to get a high conversion. After chromatography, **36** and **37** were obtained with medium yields. Their nitrile groups were reduced using borane dimethyl sulfide complex under MW conditions; the resulting benzylamine derivatives (**38** and **39**) were used

immediately in the next reaction. The last two steps were straightforward—acylation with 2-chloropyrimidine carboxylic acid (**20**) to produce **40** and **41**, and finally nucleophilic substitution with KCN to give the desired nitrile **26** and **27**.

In the synthesis of **28** (Scheme 3C), the previously described intermediate **21** was utilized. Again, it was converted to the sulfonyl chloride using $SOCl_2$, then aniline was acylated to get sulfonamide **42**. At last, it was converted with a nucleophilic substitution using KCN to the desired compound (**28**).

The results of the biochemical evaluation are summarized in Table 3. Regarding the designed compounds: if the bromine substituent of **5** ($IC_{50} = 5.18 \pm 0.52 \mu M$) was replaced with a benzyl

TABLE 3 | Second round design: summary of the measured biological activities against 3CL^{pro} and jump dilution assay results.

Compound	3CL ^{pro} RA, % or IC_{50} , μM	Jump dilution assay
 25	$24.3 \pm 0.89 \mu M$	n. d.
 27	$6.14 \pm 0.17 \mu M$	reversible inhibitor
 26	$68.7 \pm 3.0\% @ 100 \mu M^a$	n. d.
 43	$96.7 \pm 0.6\% @ 100 \mu M$	n. d.
 28	$2.38 \pm 0.05 \mu M$	reversible inhibitor

n. d.: not determined.

^a: poor-dose response, IC_{50} could not be determined.

group (**25**, $IC_{50} = 24.3 \pm 0.89 \mu M$), the activity has dropped by an order of magnitude. The incorporation of the *m*-chlorophenol moiety (**26**, $RA = 68.7 \pm 3.0\%$ @ $100 \mu M$) led to a significant deterioration in biochemical activity. If a *p*-H₂NSO₂Ph unit was present at this position (**27**, $IC_{50} = 6.14 \pm 0.17 \mu M$), a bit less active inhibitor was received comparing to the parent compound. Jump dilution assay confirmed its reversible covalent mode of action. However, when the sulfonamide moiety of **5** was modified by the incorporation of a phenyl substituent (**28**, $IC_{50} = 2.38 \pm 0.05 \mu M$), an improvement in activity was observed. It was also characterized as a reversible covalent inhibitor. Interestingly, if we also consider this molecule as the result of two linked fragments, we can see that they are inactive on their own [**1** ($90.1 \pm 4.6\%$ @ $100 \mu M$) and **43** ($96.7 \pm 0.6\%$ @ $100 \mu M$)] while a simple linker between them yields an inhibitor with a low micromolar IC_{50} value in the single digits.

3 | Conclusion

In this study, we aimed to investigate the feasibility of fragment linking in the optimization of covalent fragment hits obtained with the electrophile-first strategy. We used 3CL^{Pro} as a model system, and subjected fragments previously identified by X-ray crystallography to fragment linking. We started with fragments **1** and **2**, which are themselves inactive but have been shown to bind to the protein. After designing appropriate linkers, three molecules were synthesized, one of which had a low micromolar activity (**5**) and a reversible, covalent mode of action. However, the linking effort resulted in an unexpected binding mode of this compound, where the non-covalent fragment adapted a flipped binding pose to accommodate the alternative orientation of two side chains in the binding site. This highlights the necessity of robust structural verification, especially when multiple consecutive rounds of fragment elaboration are performed. It is also interesting to note that the designed molecules (**3**, **4**, and **5**) show a remarkable similarity in structure, with only the linker showing some differences; however, there is a major difference in their biological activities. We attribute this to the general sensitivity of the resulting binding modes to even minor structural changes, especially when one of the fragments is covalent and needs correct warhead positioning. Another factor could be the suboptimal performance of the computational linker design approach (DeLinker), but this was not investigated in detail and could not be reliably concluded without sampling and evaluating a more robust and diverse selection of linkers.

We wanted to assess the robustness of fragment linking, therefore the crystal structure of the **5**-3CL^{Pro} complex was determined and further compounds were designed and synthesized based on the modification of the core compound (**5**) and its experimental binding position. The results highlighted the above-mentioned vulnerability of the correct binding pose during fragment linking efforts. Maturation of hit compounds derived from covalent fragment linking faces multiple difficulties; structural modifications alter not only the overall binding pose but also the positioning of the warhead towards the targeted sidechain. This, in turn, can exert a significant effect on the energy profile of the covalent reaction. This finding was reinforced during our warhead replacement efforts: while we have managed to identify two

2-chloropyrimidine derivatives as irreversible covalent inhibitors, their activities were two orders of magnitude weaker than **5**, due to the suboptimal orientation of the warhead.

At the same time, the covalent mode of action was found to be crucial: without a warhead, the molecule (**71**) lost its activity completely. It is important to underline that using the “ligand-first” approach would not have led to our most active molecule (**5**), as in the absence of activity, we would not have identified the warhead-free core as a hit.

Finally, we have performed a round of fragment merging based on further fragment positions in the neighboring areas. These have further confirmed the general sensitivity of fragment merging/linking approaches to small structural modifications (cf. significant activity loss for two closely related analogs), but also yielded comparably active 3CL^{Pro} inhibitors, providing further examples for the utility of the electrophile-first approach, as detailed above.

In summary, we have conducted a fragment-linking study using the “electrophile-first” approach for targeted covalent inhibitors. Most importantly, our effort resulted in an inhibitor (**5**) with an IC_{50} value of $5.2 \mu M$ by linking practically inactive fragments, which was further optimized using a merging approach to get an inhibitor (**28**) with an IC_{50} value of $2.38 \mu M$. Crucially, early detection of the unexpected binding pose of **5** was necessary to guide further elaboration. Furthermore, our results show the importance of maintaining the correct binding pose for the compound to act as a covalent inhibitor. This is very fragile during fragment linking and optimization, but high-quality structural information significantly increases the chances of success. Finally, we should note that the *in vivo* utility of the compounds resulting from fragment-based approaches (or any other drug discovery effort) is influenced by a number of orthogonal factors (permeability, cytotoxicity, etc.) that were not addressed in detail, in light of the methodological nature of the present study.

4 | Experimental Section

4.1 | General Chemistry Methods

Reagents and solvents were purchased from commercial sources (Sigma Aldrich, Enamine, Combi-Blocks, Fluorochem) and used as received. Flash column chromatography was performed using a CombiFlash Rf 200 instrument (Teledyne ISCO, Lincoln, NE, USA). RediSep Rf reversed-phase C18 columns (26 g, 43 g, 86 g) were used for reversed-phase chromatography. Normal-phase flash column chromatography was performed on Merck Silica Gel 60 (particle size 0.040–0.063 mm; Merck, Germany). Melting points were determined using a Reichelt hot stage apparatus and are uncorrected. ¹H and ¹³C NMR spectra were recorded at 295 K on a Varian System 500 NMR spectrometer (Varian, Palo Alto, CA, USA) or Varian System 300 NMR spectrometer operating at frequencies for ¹H NMR at 500 MHz or 300 MHz, and for ¹³C NMR at 126 or 75 MHz, respectively. Chemical shifts (δ) are in parts per million (ppm) and are referenced to the deuterated solvent used. The coupling constants (*J*) are given in Hz, and the splitting patterns are designated as follows: s, singlet;

br s, broad singlet; d, doublet; app d, apparent doublet; dd, doublet of doublets; ddd, doublet of doublets of doublets; t, triplet; dt, doublet of triplets; td, triplet of doublets; m, multiplet. All ^{13}C NMR spectra were proton decoupled. HPLC-MS measurements were performed using a Shimadzu LC-MS-2020 instrument (Shimadzu Corporation, Kyoto, Japan) equipped with a Reprospher 100 C18 (5 mm, 100×3 mm) column and a positive-negative double ion source (DUIS $\pm$) with a quadrupole mass spectrometer in the range of 50–1000 m/z. Samples were eluted by gradient elution using eluent A (0.1% HCOOH in H_2O) and eluent B (0.1% HCOOH in MeCN). The flow rate was set to 1.5 mL/min. The initial condition was 0% B eluent, followed by a linear gradient to 100% B eluent in 2 min, from 2 to 3.75 min, 100% B eluent was maintained, and from 3.75 to 4.5 min, the initial condition was restored and maintained until 5 min. The column temperature was maintained at 30°C, and the injection volume was 1 μL . For purity determination, the PDA detection wavelength was set at 254 nm, unless stated otherwise.

4.2 | Cloning, Expression, and Purification of Recombinant 3CL^{pro}

Codon optimized gene encoding the SARS-CoV-2 3CL^{pro} gene (Integrated DNA Technologies, USA) was cloned into the pET-28c(+) plasmid, and used to transform *E. coli* NiCo21(DE3) (New England Biolabs, USA). Transformed bacteria were grown at 37°C and 250 rpm in LB broth supplemented with kanamycin (50 $\mu\text{g}/\text{mL}$) to an $\text{OD}_{600} \sim 1.8$. Cultures were cooled on ice for 10 min and expression was induced with isopropyl β -D-1-thiogalactopyranoside – IPTG (200 μM). Expression was continued at 16°C, and 250 rpm for 24 h. Cells were harvested by centrifugation (2×10 min, $3000 \times g$ at 4°C). Cell pellets were resuspended in buffer A (20 mM Tris, pH 7.5, 0.05 mM EDTA, 2.5 mM DTT, 10% glycerol) and lysed on ice by sonication. Cell debris was removed by centrifugation (2×30 min, $16,000 \times g$ at 4°C). Clear lysate was filtered through a 100-kDa MWCO unit (Amicon Ultra-15 centrifugal filter units; Merck, Germany). $(\text{NH}_4)_2\text{SO}_4$ (0.5 M final concentration) was slowly added to the filtrate, which was then loaded onto a 1 mL HiTrap Phenyl HP column (Cytiva, USA) pre-equilibrated with buffer B (50 mM Tris, pH 7.5, 0.5 M $(\text{NH}_4)_2\text{SO}_4$, 0.05 mM EDTA, 2.5 mM DTT, 10% glycerol). The column was washed with 20 volumes of buffer B and the 3CL^{pro} was eluted with a linear gradient into buffer A. The eluted protein was concentrated with a 30-kDa MWCO unit (Ultra-4 centrifugal filter units; Amicon), snap-frozen in liquid nitrogen, and stored at -80°C . Protein concentration was determined by measuring absorbance at 280 nm and considering the extinction coefficient of $34,380 \text{ M}^{-1} \text{ cm}^{-1}$. Purity was determined by SDS-PAGE.

4.3 | Enzyme Activity Assay

The enzymatic activity of 3CL^{pro} was measured by kinetic assay using FRET fluorogenic substrate HilyteFluor488-ESATLQSGLRKAKQXL520 (HilyteFluor488, HPLC peptide purity 93%, Product #: AS-65 599, Anaspec, USA). Experiments were performed in assay buffer 50 mM Tris-HCl, pH 7.3, 1 mM EDTA, 0.05% Triton X-114. For the screening, compounds were pre-incubated at a concentration of 500 μM and 100 μM

with 3CL^{pro} for 30 min at 30°C. The reaction was started by adding substrate, and the increase in fluorescence intensity was measured using Synergy H4 microplate reader (BioTek Instruments, Inc., USA) at $\lambda_{\text{ex}} = 485/20$ and $\lambda_{\text{em}} = 528/20$ nm. The final concentrations were as follows: substrate, HilyteFluor488 (2 μM); 3CL^{pro}, 50 nM; DMSO, 10% [v/v]. In control experiments, the compound was replaced by DMSO. For the determination of b (blank), the enzyme was replaced with Tris-HCl buffer solution. Initial velocities (v) were calculated from the linear trends obtained, with each measurement performed in duplicate. Inhibitory potencies were expressed as residual activities ($\text{RA} = (v_i - b)/(v_o - b)$), where v_i is the velocity in the presence of the test compound, and v_o is the control velocity in the presence of DMSO. At least seven different concentrations were used for IC_{50} determinations for each compound. The IC_{50} values were calculated from experimental data fitted to a four-parameter logistic function (GraphPad Prism 9.4).

4.4 | Jump Dilution Assay

For 100-fold jump dilution assay, compounds were preincubated with 3CL^{pro} (5 μM) for 30 min at the concentration equal to $10 \times \text{IC}_{50}$ value, and then diluted 100-fold into buffer containing DABCYL-EDANS substrate (DABCYL-EDANS, HPLC peptide purity 97%, Product #: COVD-001, CPC Scientific, CA, USA, 20 μM). The fluorescence increase was followed using Synergy H4 microplate reader (BioTek Instruments, Inc., USA) at $\lambda_{\text{ex}} = 360/40$ and $\lambda_{\text{em}} = 460/40$ nm. Control experiment was carried out in the same manner, and the inhibitor solution was replaced by DMSO, and GC376 was used as a positive control.

4.5 | Crystallization and Data Collection of 3CL^{pro}-Small Molecule Complexes

Purified protein at a concentration of 5 mg/mL in 10 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol buffer was used in performing crystallization using sitting-drop vapor diffusion method using Swissci 3 lens crystallisation plates. The reservoir solution (35 μL added to each reservoir well) used in crystallization contained 11% (w/v) PEG 4000, 5% DMSO, 0.1 M MES, pH 6.7. Crystals of 3CL^{pro} orthorhombic crystal form ($\text{P}2_12_12_1$) were grown with drop ratios of 0.15 μL protein, 0.15 μL reservoir solution, and 0.05 μL seeds prepared from previously produced crystals of the same crystal form [68]. Plates were incubated in a Formulatrix Rock Imager (20°C) with the first images taken after 12 h and the imaging schedule following a Fibonacci sequence of days for further collections. Crystal formation was observed in 95% of drops within 24 h and reached their maximum size after 48 h. Further details can be found here [68, 69].

For 3CL^{pro}-small molecule complexes, all experiments were performed at Diamond Light Source XChem facility [70]. Compound was dispensed into crystal drops by an ECHO 550 Liquid Handler (Beckman Coulter) with 17.64 (v/v) % DMSO, with final concentration of 17.64 mM compound soaked into each crystal drop. Thereafter, crystallisation plates were incubated for 1 h at 20°C. Crystals were then harvested using Crystal Shifter [70] (Oxford Lab Technologies) and flash frozen in liquid nitrogen. Diffraction data were collected at Diamond Light Source on the beamline

I04–1 at 100 K and processed with the fully automated pipelines at Diamond, including XDS27, Autoproc28, Xia2 29, and DIALS30 [71–76]. All further analysis was performed using XChemExplorer [77]. Electron density map were generated by Dimple [78] and ligand restraints were generated with GRADE [79]. Model building and refinement were carried out in COOT [80] using Buster [81] via the XChemExplorer platform.

4.6 | Intact Protein LC-MS Measurements

The protein stock solution (10.0 μ L, 37.6 μ M) was diluted with buffer solution (10.0 μ L; pH 7.3; 50 mM Tris $\times$ HCl, 1 mM EDTA, 0.05% Triton X-114), then the DMSO stock solution of the inhibitor (0.66 μ L, 50 mM; 87-fold reagent excess) was added. The sample was incubated at 37°C for 90 min, and then subjected to protein LC-MS measurements. The molecular weights of the conjugates of 3CL^{Pro} were identified using a Triple TOF 5600+ hybrid Quadrupole-TOF LC/MS/MS system (Sciex, Singapore, Woodlands) equipped with a DuoSpray IonSource coupled with a Shimadzu Prominence LC20 UFLC (Shimadzu, Japan) system consisting of binary pump, an autosampler, and a thermostated column compartment. Data acquisition and processing were performed using Analyst TF software version 1.7.1 (AB Sciex Instruments, CA, USA). Chromatographic separation was achieved on a Thermo Beta Basic C8 (50 mm $\times$ 3 mm, 3 μ m, 150 Å) HPLC column. The sample was eluted in gradient elution mode using solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in ACN). The initial condition was 15% B for 1 min, followed by a linear gradient to 55% B by 12 min, from 12 to 15 min to 95%, from 15 to 17.5 min B was retained; and from 17.5 to 18 min back to initial condition with 10% eluent B and retained from 18 to 20 min. The flow rate was set to 0.5 ml/min. The column temperature was 40°C, and the injection volume was 10 μ L. Nitrogen was used as the nebulizer gas (GS1), heater gas (GS2), and curtain gas with the optimum values set at 40, 40, and 45 (arbitrary units), respectively. Data were acquired in positive electrospray mode in the mass range of m/z = 250 to 2500, with 1 s accumulation time. The source temperature was 350°C, and the spray voltage was set to 5000 V. Declustering potential value was set to 80 V. Peak View Software V.2.2 (version 2.2, Sciex, Redwood City, CA, USA) was used for deconvoluting the raw electrospray data to obtain the neutral molecular masses.

4.7 | Computational Methods

4.7.1 | Structure Alignment, Fragment Selection, and Fragment Linking

First, the respective PDB structures, containing non-covalent and covalent fragments, were collected using Schrödinger's Maestro [82], and protein structures were prepared by the Protein Preparation Wizard [83]. Then, protein structure alignment was carried out based on the amino acid sequence of SARS-CoV-2 3CL^{Pro}. Non-covalent ligands were extracted, and molecules that overlap with the covalent warheads or bind to another part of the protein were filtered out. The remaining structures were visually inspected, and non-covalent fragments that occupy a position that is feasible for linking were kept for the next phase. At the end of the selection phase, two covalent and 17 non-covalent fragments

were selected for linking. Based on the feasibility of linking (i.e. suitable distance and orientation), 17 covalent – non-covalent fragment pairs were created (see Supporting Information) and were linked using DeLinker [60]. Additionally, fragment pairs that contain atoms within bounding distance were merged based on visual inspection. This approach resulted in 20 further directly merged compounds. The design process resulted in a total of 2780 compounds.

4.7.2 | Covalent Docking and Compound Filtering

The designed compounds were prepared by LigPrep [84] and docked covalently into the catalytic cysteine of SARS-CoV-2 3CL^{Pro} using Maestro's Covdock [85]. The previously prepared PDB structure of 5RHB [45] was used for the docking, with the covalent ligand set as the center of the grid and Cys145 selected as the targeted sidechain. The volume of the inner grid was set to 10 $\times$ 10 $\times$ 10 Å, while the outer box was set to 30 $\times$ 30 $\times$ 30 Å. The reaction type was set to “Nucleophilic Addition to a Triple Bond” and the virtual screening precision was used to dock the compounds. Based on the docking score the best 1000 compounds and docking poses were written out. The RMSD of the molecular parts derived from the non-covalent fragments was evaluated for the docked molecules compared to the original pose of the non-covalent fragments. Based on the docking scores, RMSD values, visual inspection, and synthetic feasibility, three compounds were selected to be synthesized and tested.

Covalent docking of **5** into the crystal structure of PDB 7RNK [65] possessing M41(out) Q189(in) orientation was carried out similarly as described previously. The protein was pre-processed with the Protein Preparation Wizard. The crystallographic ligand was selected as the center of the grid, Cys145 was selected as the targeted residue, while the reaction type was set to “Nucleophilic Addition to a Triple Bond” with pose prediction (PP) precision.

4.7.3 | Linking of Non-Covalent Fragments

For the linking of non-covalent fragments, the previously collected and prepared PDB structures were used. After visual inspection, four fragments were found to be feasible to be linked. Based on their crystallographic positions, four possible fragment pairs were created and linked using DeLinker [60]. Finally, 481 new non-covalent inhibitors were designed. The molecules created by DeLinker were prepared with LigPrep [84] and were docked with Glide [86] using single precision (SP). The protein structure of 5RHB [45] was used as a docking structure in which the covalent bond was deleted between the ligand and Cys145, Cys145 was present in its neutral form. The restored nitrile derivative was selected as the center of the grid; further parameters were set to their default values. The designed compounds were evaluated by docking scores, RMSD values to their original fragment positions, and synthetic feasibility.

Author Contributions

Levente Kollár synthesized and characterized the compounds, performed protein labeling experiments; **Levente M. Mihalovits** performed the computational work; **József Simon** performed analytical characterization of the synthesized compounds; **Damijan Knez** and **Stanislav**

Gobec performed bioactivity measurements; Blake H. Balcomb and Daren Fearon performed crystallization and X-ray diffraction measurements; Dávid Bajusz and György M. Keserű supervised the project. All authors contributed to writing the manuscript.

Acknowledgments

The authors thank Krištof Bozovičar (Univ. Ljubljana), Peter G. Marples (Diamond Light Source), Anett Kuczog, and Gábor Kemenesi (National Laboratory of Virology, University of Pécs) for their contributions to the experimental work. We thank Dr. Tímea Imre for her contribution to protein LC-MS/MS measurements and Dorottya Csányi for her contribution to NMR measurements. This work was supported by the National Research Development and Innovation Office of Hungary [FK146063, SNN135335, 2020-1.1.2-PIACI-KFI-2020-00039 and PharmaLab (RRF-2.3.1-21-2022-00015)], the National Brain Research Program (NAP 3.0) of the Hungarian Academy of Sciences, and the Slovenian Research and Innovation Agency – ARIS (bilateral grant BI-HU/21-22-003 and core financing P1-0208). The work of D. B. was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. We acknowledge the HPC time and support of the Governmental Information Technology Development Agency, Hungary.

Funding

This study was supported by Nemzeti Kutatási Fejlesztési és Innovációs Hivatal (Grants FK146063, SNN135335, 2020-1.1.2-PIACI-KFI-2020-00039 and RRF-2.3.1-21-2022-00015), Hungarian Academy of Sciences (NAP3.0) and Slovenian Research and Innovation Agency - ARIS (grants BI-HU/21-22-003 and P1-0208).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The crystallographic data referenced in this paper are openly available at the [RCSB Protein Data Bank](#) under entries 5RHB, 5RHC, 5RF1, 7RNK and 29OK (for the structure of compound **5** vs. 3CLPro), or at the [Fragalysis](#) service of the Diamond Light Source, UK, under Legacy targets/MPro, entries ×1311, ×12177, ×12419. The crystallographic data table of the structure of compound **5** vs. 3CLPro is included in the Supplementary Information (Table S2).

References

1. S. B. Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, "Discovering High-Affinity Ligands for Proteins: SAR by NMR," *Science* (80-) 274 (1996): 1531–1534.
2. P. J. Hajduk, R. P. Meadows, and S. W. Fesik, "Discovering High-Affinity Ligands for Proteins," *Science* (80-) 278 (1997): 497–499.
3. V. L. Nienaber, P. L. Richardson, V. Klighofer, J. J. Bouska, V. L. Giranda, and J. Greer, "Discovering Novel Ligands for Macromolecules Using X-Ray Crystallographic Screening," *Nature Biotechnology* 18 (2000): 1105–1108.
4. B. Lamoree and R. E. Hubbard, "Current Perspectives in Fragment-Based Lead Discovery (FBLD)," *Essays in Biochemistry* 61 (2017): 453–464.
5. M. Congreve, R. Carr, C. Murray, and H. Jhoti, "A Rule of Three for Fragment-Based Lead Discovery? Recent," *Drug Discovery Today* 8 (2003): 876–877.
6. G. M. Keserű, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray, and S. D. Pickett, "Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia," *Journal of Medicinal Chemistry* 59 (2016): 8189–8206.

7. H. Jhoti, G. Williams, D. C. Rees, and C. W. Murray, "The Rule of Three for Fragment-Based Drug Discovery: Where Are We now?," *Nature Reviews Drug Discovery* 12 (2013): 644.
8. H. Köster, T. Craan, S. Brass, et al., "A Small Nonrule of 3 compatible Fragment Library Provides High Hit Rate of Endothiapepsin Crystal Structures with Various Fragment Chemotypes," *Journal of Medicinal Chemistry* 54 (2011): 7784–7796.
9. C. W. Murray and D. C. Rees, "Opportunity Knocks: Organic Chemistry for Fragment-Based Drug Discovery (FBDD)," *Angewandte Chemie International Edition* 55 (2016): 488–492.
10. A. D. Gossert and W. Jahnke, "NMR in Drug Discovery: A Practical Guide to Identification and Validation of Ligands Interacting with Biological Macromolecules," *Progress in Nuclear Magnetic Resonance Spectroscopy* 97 (2016): 82–125.
11. D. R. Davies, D. W. Begley, R. C. Hartley, B. L. Staker, and L. J. Stewart, "Predicting the Success of Fragment Screening by X-Ray Crystallography," *Methods in Enzymology* 493 (2011): 91–114.
12. A. M. Giannetti, "From Experimental Design to Validated Hits: A Comprehensive Walk-through of Fragment Lead Identification Using Surface Plasmon Resonance," *Methods in Enzymology* 493 (2011): 169–218.
13. P. Kirsch, A. M. Hartman, A. K. H. Hirsch, and M. Empting, "Concepts and Core Principles of Fragment-Based Drug Design," *Molecules* 24 (2019): 4309.
14. Q. Li, "Application of Fragment-Based Drug Discovery to Versatile Targets," *Frontiers in Molecular Biosciences* 7 (2020): 1–13.
15. D. A. Erlanson, R. S. McDowell, and T. O'Brien, "Fragment-Based Drug Discovery," *Journal of Medicinal Chemistry* 47 (2004): 3463–3482.
16. C. W. Murray and D. C. Rees, "The Rise of Fragment-Based Drug Discovery," *Nature Chemistry* 1 (2009): 187–192.
17. A. Racz, L. Mihalovits, M. Beckers, et al., "Changing Landscape of Medicinal Chemistry Optimizations," *Nature Reviews Drug Discovery* 24 (2025): 870–887.
18. L. R. de Souza Neto, J. T. Moreira-Filho, B. J. Neves, et al., "In Silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery," *Frontiers in Chemistry* 8 (2020): 1–18.
19. A. Bancet, C. Raingeval, T. Lomberget, M. Le Borgne, J. F. Guichou, and I. Krimm, "Fragment Linking Strategies for Structure-Based Drug Design," *Journal of Medicinal Chemistry* 63 (2020): 11420–11435.
20. D. Erlanson, "Fragments in the Clinic: 2024 Edition," Practical Fragments blog, accessed may 12, 2026, <https://practicalfragments.blogspot.com/2024/02/fragments-in-clinic-2024-edition.html>.
21. C. N. Johnson, D. A. Erlanson, C. W. Murray, and D. C. Rees, "Fragment-to-Lead Medicinal Chemistry Publications in 2015," *Journal of Medicinal Chemistry* 60 (2017): 89–99.
22. C. N. Johnson, D. A. Erlanson, W. Jahnke, P. N. Mortenson, and D. C. Rees, "Fragment-to-Lead Medicinal Chemistry Publications in 2016," *Journal of Medicinal Chemistry* 61 (2018): 1774–1784.
23. P. N. Mortenson, D. A. Erlanson, I. J. P. De Esch, W. Jahnke, and C. N. Johnson, "Fragment-to-Lead Medicinal Chemistry Publications in 2017," *Journal of Medicinal Chemistry* 62 (2019): 3857–3872.
24. D. A. Erlanson, I. J. P. De Esch, W. Jahnke, C. N. Johnson, and P. N. Mortenson, "Fragment-to-Lead Medicinal Chemistry Publications in 2018," *Journal of Medicinal Chemistry* 63 (2020): 4430–4444.
25. W. Jahnke, D. A. Erlanson, I. J. P. De Esch, et al., "Fragment-to-Lead Medicinal Chemistry Publications in 2019," *Journal of Medicinal Chemistry* 63 (2020): 15494–15507.
26. I. J. P. De Esch, D. A. Erlanson, W. Jahnke, C. N. Johnson, and L. Walsh, "Fragment-to-Lead Medicinal Chemistry Publications in 2020," *Journal of Medicinal Chemistry* 65 (2022): 84–99.

27. L. Walsh, D. A. Erlanson, I. J. P. de Esch, W. Jahnke, A. Woodhead, and E. Wren, "Fragment-to-Lead Medicinal Chemistry Publications in 2021," *Journal of Medicinal Chemistry* 66 (2023): 1137–1156.
28. A. J. Woodhead, D. A. Erlanson, I. J. P. de Esch, R. S. Holvey, W. Jahnke, and P. Pathuri, "Fragment-to-Lead Medicinal Chemistry Publications in 2022," *Journal of Medicinal Chemistry* 67 (2024): 2287–2304.
29. H. S. Yu, K. Modugula, O. Ichihara, et al., "General Theory of Fragment Linking in Molecular Design: Why Fragment Linking Rarely Succeeds and How to Improve Outcomes," *Journal of Chemical Theory and Computation* 17 (2021): 450–462.
30. A. J. Souers, J. D. Levenson, E. R. Boghaert, et al., "ABT-199, a Potent and Selective BCL-2 Inhibitor, Achieves Antitumor Activity while Sparing Platelets," *Nature Medicine* 19 (2013): 202–208.
31. T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, et al., "An Inhibitor of Bcl-2 Family Proteins Induces Regression of Solid Tumours," *Nature* 435 (2005): 677–681.
32. A. Keeley, L. Petri, P. Ábrányi-Balogh, and G. M. Keserű, "Covalent Fragment Libraries in Drug Discovery," *Drug Discovery Today* 25 (2020): 983–996.
33. K. McAulay, A. Bilsland, and M. Bon, "Reactivity of Covalent Fragments and Their Role in Fragment Based Drug Discovery," *Pharmaceuticals* 15 (2022): 1–22.
34. L. Boike, N. J. Henning, and D. K. Nomura, "Advances in Covalent Drug Discovery," *Nature Reviews. Drug Discovery* 21 (2022): 881–898.
35. J. Canon, K. Rex, A. Y. Saiki, et al., "The Clinical KRAS(G12C) Inhibitor AMG. 510 Drives Anti-Tumour Immunity," *Nature* 575 (2019): 217–223.
36. B. A. Lanman, J. R. Allen, J. G. Allen, et al., "Discovery of a Covalent Inhibitor of KRASG12C (AMG. 510) for the Treatment of Solid Tumors," *Journal of Medicinal Chemistry* 63 (2020): 52–65.
37. F. Skoulidis, B. T. Li, G. K. Dy, et al., "Sotorasib for Lung Cancers with KRAS p.G12C Mutation," *The New England Journal of Medicine* 384 (2021): 2371–2381.
38. A. O. Walter, R. T. T. Sjin, H. J. Haringsma, et al., "Discovery of a Mutant-Selective Covalent Inhibitor of EGFR that Overcomes T790M Mediated Resistance in NSCLC," *Cancer Discovery* 3 (2013): 1404–1415.
39. M. R. V. Finlay, M. Anderton, S. Ashton, et al., "Discovery of a Potent and Selective EGFR Inhibitor (AZD9291) of Both Sensitizing and T790M Resistance Mutations that Spares the Wild Type Form of the Receptor," *Journal of Medicinal Chemistry* 57 (2014): 8249–8267.
40. R. H. Advani, J. J. Buggy, J. P. Sharman, et al., "Bruton Tyrosine Kinase Inhibitor Ibrutinib (PCI-32765) Has Significant Activity in Patients with Relapsed/Refractory B-Cell malignancies," *Journal of Clinical Oncology* 31 (2013): 88–94.
41. M. S. Davids and J. R. Brown, "Ibrutinib: A First in Class Covalent Inhibitor of Bruton's Tyrosine Kinase," *Future Oncology* 10 (2014): 957–967.
42. K. Vandeyck and J. Deval, "Considerations for the Discovery and Development of 3-Chymotrypsin-Like Cysteine Protease Inhibitors Targeting SARS-CoV-2 Infection," *Current Opinion in Virology* 49 (2021): 36–40.
43. E. Resnick, A. Bradley, J. Gan, et al., "Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening," *Journal of the American Chemical Society* 141 (2019): 8951–8968.
44. M. D. Olp, D. J. Sprague, C. J. Goetz, et al., "Covalent-Fragment Screening of BRD4 Identifies a Ligandable Site Orthogonal to the Acetyl-Lysine Binding Sites," *Acs Chemical Biology* 15 (2020): 1036–1049.
45. A. Douangamath, D. Fearon, P. Gehrtz, et al., "Crystallographic and Electrophilic Fragment Screening of the SARS-CoV-2 Main Protease," *Nature Communications* 11 (2020): 5047.
46. R. Hilgenfeld, "From SARS to MERS: Crystallographic Studies on Coronaviral Proteases Enable Antiviral Drug Design," *The Febs Journal* 281 (2014): 4085–4096.
47. C. A. Ramos-Guzmán, J. J. Ruiz-Pernía, and I. Tuñón, "Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale Methods," *ACS Catalysis* 10 (2020): 12544–12554.
48. Z. Jin, X. Du, Y. Xu, et al., "Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors," *Nature* 582 (2020): 289–293.
49. W. Dai, B. Zhang, X. M. Jiang, et al., "Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease," *Science* (80-.) 368 (2020): 1331–1335.
50. L. Zhang, D. Lin, X. Sun, et al., "Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved a-Ketoamide Inhibitors," *Science* (80-.) 368 (2020): 409–412.
51. D. W. Kneller, G. Phillips, H. M. O'Neill, et al., "Structural Plasticity of SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-Ray Crystallography," *Nature Communications* 11 (2020): 7–12.
52. Y. Hu, E. M. Lewandowski, H. Tan, et al., "Naturally Occurring Mutations of SARS-CoV-2 Main Protease Confer Drug Resistance to Nirmatrelvir," *ACS Central Science* 9 (2023): 1658–1669.
53. B. Tan, M. Sacco, H. Tan, et al., "Exploring Diverse Reactive Warheads for the Design of SARS-CoV-2 Main Protease Inhibitors," *European Journal of Medicinal Chemistry* 259 (2023): 115667.
54. B. Tan, R. Joyce, H. Tan, Y. Hu, and J. Wang, "SARS-CoV-2 Main Protease Drug Design, Assay Development, and Drug Resistance Studies," *Accounts of Chemical Research* 56 (2023): 157–168.
55. Y. Zhan, Z. Lin, J. Liang, et al., "Leritrelvir for the Treatment of Mild or Moderate COVID-19 without Co-Administered Ritonavir: A Multicentre Randomised, Double-Blind, Placebo-Controlled phase 3 trial," *EclinicalMedicine* 67 (2024): 1–11.
56. B. Cao, Y. Wang, H. Lu, et al., "Oral Simnotrelvir for Adult Patients with Mild-to-Moderate Covid-19," *The New England Journal of Medicine* 390 (2024): 230–241.
57. H. Yotsuyanagi, N. Ohmagari, Y. M. Yamato, et al., "Efficacy and Safety of 5-Day Oral Ensitrelvir for Patients With Mild to Moderate COVID-19," *Jama Network Open* 7 (2024): e2354991.
58. M. O'Reilly, A. Cleasby, T. G. Davies, et al., "Crystallographic Screening Using Ultra-Low-Molecular-Weight Ligands to Guide Drug Design," *Drug Discovery Today* 24 (2019): 1081–1086.
59. A. Keeley, P. Ábrányi-Balogh, and G. M. Keserű, "Design and Characterization of a Heterocyclic Electrophilic Fragment Library for the Discovery of Cysteine-Targeted Covalent Inhibitors," *Medchemcomm* 10 (2019): 263–267.
60. F. Imrie, A. R. Bradley, M. Van Der Schaar, and C. M. Deane, "Deep Generative Models for 3D Linker Design," *Journal of Chemical Information and Modeling* 60 (2020): 1983–1995.
61. J. Popovici-Muller, F. G. Salituro, J. O. Saunders, J. Travins, and Y. Shungi, *Therapeutic Compounds and Compositions* (2014), intl. patent WO2014062511A1.
62. K. Kyung-Hee, C. Jung-Wook, L. Ji-Hoon, et al., *2-Cyanopyrimidin-4-Yl Carbamate or Urea Derivative or Salt Thereof, and Pharmaceutical Composition Including Same* (2019), intl. patent WO2019190117.
63. Y. Q. Xiao, J. Long, S. S. Zhang, Y. Y. Zhu, and S.-X. Gu, "Non-Peptidic Inhibitors Targeting SARS-CoV-2 Main Protease: A Review," *Bioorganic Chemistry* 147 (2024): 107380.
64. S. Brogi, R. Ibba, S. Rossi, et al., "Covalent Reversible Inhibitors of Cysteine Proteases Containing the Nitrile Warhead: Recent Advancement in the Field of Viral and Parasitic Diseases," *Molecules* 27 (2022): 2561.
65. D. W. Kneller, H. Li, S. Galanie, et al., "Structural, Electronic, and Electrostatic Determinants for Inhibitor Binding to Subsites S1 and S2

in SARS-CoV-2 Main Protease,” *Journal of Medicinal Chemistry* 64 (2021): 17366–17383.

66. “FRAGALYSIS: MPRO | LB18145-1 (<https://fragalysis-legacy.xchem.diamond.ac.uk/viewer/react/preview/target/Mpro/tas/lb18145-1>)”

67. A. Hall, J. Abendroth, M. J. Bolejack, et al., “Discovery and Characterization of a Novel Series of Chloropyrimidines as Covalent Inhibitors of the Kinase MSK1,” *ACS Medicinal Chemistry Letters* 13 (2022): 1099–1108.

68. B. H. Balcomb, P. Marples, L. Koekemoer, D. Fearon, and C. Tomlinson, *Crystallization of SARS-CoV-2 Mpro* (2024), <https://www.protocols.io/view/crystallization-of-sars-cov-2-mpro-kqdg326y7v25/v1>, <https://doi.org/10.17504/protocols.io.kqdg326y7v25/v1>.

69. A. Douangamath, A. Powell, D. Fearon, et al., “Achieving Efficient Fragment Screening at XChem Facility at Diamond Light Source,” *Journal of Visualized Experiments* 171 (2021): e62414.

70. N. D. Wright, P. Collins, L. Koekemoer, et al., “The Low-Cost Shifter Microscope Stage Transforms the Speed and Robustness of Protein Crystal Harvesting,” *Acta Crystallographica Section D, Structural Biology* 77 (2021): 62–74.

71. G. Winter, D. G. Waterman, J. M. Parkhurst, et al., “DIALS: Implementation and Evaluation of a New Integration Package,” *Acta Crystallographica Section D, Structural Biology* 74 (2018): 85–97.

72. G. Winter, R. J. Gildea, N. G. Paterson, et al., “How Best to use Photons,” *Acta Crystallographica Section D, Structural Biology* 75 (2019): 242–261.

73. G. Winter and K. E. McAuley, “Automated Data Collection for Macromolecular Crystallography,” *Methods* 55 (2011): 81–93.

74. W. Kabsch, “Integration, Scaling, Space-Group Assignment and Post-Refinement,” *Acta Crystallographica Section D Biological Crystallography* 66 (2010): 133–144.

75. G. Winter, C. M. C. Lobley, and S. M. Prince, “Decision Making in Xia 2,” *Acta Crystallographica Section D Biological Crystallography* 69 (2013): 1260–1273.

76. C. Vonrhein, C. Flensburg, P. Keller, et al., “Data Processing and Analysis with the autoPROC Toolbox,” *Acta Crystallographica Section D Biological Crystallography* 67 (2011): 293–302.

77. T. Krojer, R. Talon, N. Pearce, et al., “The XChemExplorer Graphical Workflow Tool for Routine or Large-Scale Protein-Ligand Structure Determination,” *Acta Crystallographica Section D Biological Crystallography* 73 (2017): 267–278.

78. R. Keegan, M. Wojdyr, G. Winter, and A. Ashton, “DIMPLE : A Difference Map Pipeline for the Rapid Screening of Crystals on the Beamline,” *Acta Crystallographica Section A Foundations and Advances* 71 (2015): s18–s18.

79. O. S. Smart, T. O. Womack, A. Sharff, et al., *Grade, version 1.2.20* (Global Phasing Ltd., 2011), n.d, <https://www.globalphasing.com>.

80. P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, “Features and Development of Coot,” *Acta Crystallographica Section D Biological Crystallography* 66 (2010): 486–501.

81. O. S. Smart, T. O. Womack, C. Flensburg, et al., “Exploiting Structure Similarity in Refinement: Automated NCS and Target-Structure Restraints in BUSTER,” *Acta Crystallographica Section D Biological Crystallography* 68 (2012): 368–380.

82. Schrödinger Release 2023-3: Maestro (Schrödinger, LLC, 2023),

83. Schrödinger Release 2023-3: Protein Preparation Wizard (Impact, Schrödinger, LLC, 2023).

84. Schrödinger Release 2023-3: LigPrep (Schrödinger, LLC, 2023).

85. K. Zhu, K. W. Borrelli, J. R. Greenwood, et al., “Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and

Scoring,” *Journal of Chemical Information and Modeling* 54 (2014): 1932–1940.

86. Schrödinger Release 2023-3: Glide, (Schrödinger, LLC, 2023).

Supporting Information

Additional supporting information can be found online in the Supporting Information section.