Title

Structure-Based Discovery of Inhibitors of the SARS-CoV–2 Nsp14 N7-Methyltransferase

Authors

Singh, Isha

Li, Fengling

Fink, Elissa A

et al.

Peer reviewed

Structure-Based Discovery of Inhibitors of the SARS-CoV-2 Nsp14 N7-Methyltransferase

Isha Singh[○○],

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

Fengling Li[○○],

Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

Elissa A. Fink,

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

Corresponding Authors: John J. Irwin – Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States; QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States; jir322@gmail.com, **Masoud Vedadi** – Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada; Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada; Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada; QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States; m.vedadi@utoronto.ca, **Brian K. Shoichet** – Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States; QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States; bshoichet@gmail.com.
[○○]**Author Contributions** I.S., F.L., and E.A.F. contributed equally.

Graduate Program in Biophysics, University of California San Francisco, San Francisco, California 94143, United States

Irene Chau[○○],
Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

Alice Li,
Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada

Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Annía Rodriguez-Hernández,
Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94158, United States

Isabella Glenn,
Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

Francisco J. Zapatero-Belinchón,
Gladstone Institutes, San Francisco, California 94158, United States

M. Luis Rodriguez,
Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Kanchan Devkota,
Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

Zhijie Deng,
Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences, Oncological Sciences and Neuroscience, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Kris White,
Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Xiaobo Wan,
Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

Nataliya A. Tolmachova,
Enamine Ltd, Kyïv 02094, Ukraine; Institute of Bioorganic Chemistry and Petrochemistry, National Ukrainian Academy of Science, Kyïv 02660, Ukraine

Yurii S. Moroz,

National Taras Shevchenko University of Kyïv, Kyïv 01601, Ukraine; Chemspace, Riga LV-1082, Latvia

H. Ümit Kaniskan,

Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences, Oncological Sciences and Neuroscience, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Melanie Ott,

Gladstone Institutes, San Francisco, California 94158, United States; QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Department of Medicine, University of California, San Francisco, San Francisco, California 94158, United States

Chan Zuckerberg Biohub, San Francisco, California 94158, United States

Adolfo García-Sastre,

Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Global Health and Emerging Pathogens Institute, Department of Medicine, Division of Infectious Diseases, and The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Jian Jin,

Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences, Oncological Sciences and Neuroscience, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Danica Galoni   Fujimori,

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94158, United States

QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

John J. Irwin,

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States; QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Masoud Vedadi,

Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada; Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Brian K. Shoichet

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, United States

QBI COVID-19 Research Group (QCRG), San Francisco, California 94158, United States

Abstract

An under-explored target for SARS-CoV-2 is the *S*-adenosyl methionine (SAM)-dependent methyltransferase Nsp14, which methylates the N7-guanosine of viral RNA at the 5′-end, allowing the virus to evade host immune response. We sought new Nsp14 inhibitors with three large library docking strategies. First, up to 1.1 billion lead-like molecules were docked against the enzyme's SAM site, leading to three inhibitors with IC_{50} values from 6 to 50 μM. Second, docking a library of 16 million fragments revealed 9 new inhibitors with IC_{50} values from 12 to 341 μM. Third, docking a library of 25 million electrophiles to covalently modify Cys387 revealed 7 inhibitors with IC_{50} values from 3.5 to 39 μM. Overall, 32 inhibitors encompassing 11 chemotypes had IC_{50} values < 50 μM and 5 inhibitors in 4 chemotypes had IC_{50} values < 10 μM. These molecules are among the first non-SAM-like inhibitors of Nsp14, providing starting points for future optimization.

Graphical Abstract
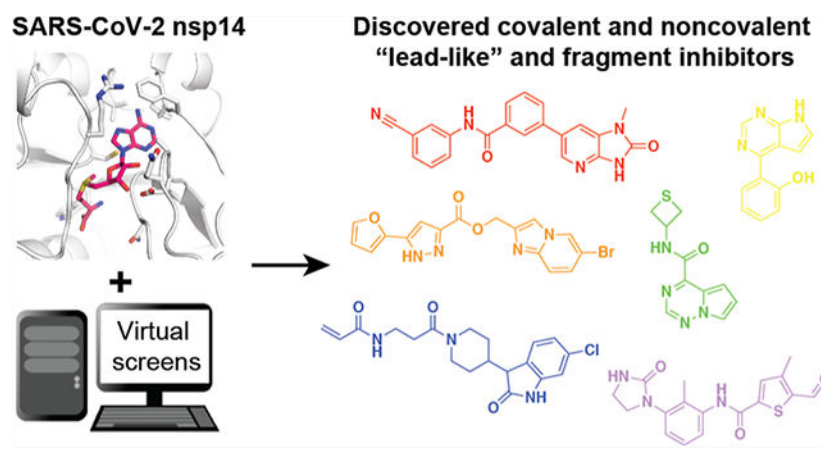
INTRODUCTION

The Covid-19 pandemic has inspired a search for targets whose inhibition would combat the virus. Fruits of such efforts have been the development of Paxlovid,[1] an inhibitor of the major protease (M[pro]) of SARS-CoV-2, of Molnupiravir,[2,3] a disruptor of viral RNA polymerization, and the introduction of Remdesivir,[4,5] an RNA-dependent RNA polymerase (RdRp) inhibitor first developed to treat Ebola virus. These targets are well-precedented in antiviral research, with successful drugs treating analogous enzymes for

HIV, HCV, RSV, HBV, HCMV, HSV, HPV, and human influenza virus,[6–9] among others, and these SARS-CoV-2 enzymes have been the focus of enormous efforts among many groups.[1,4,5,10–13] Other SARS-CoV-2 enzymes have attracted less attention, likely because there is less precedence for their targeting as antivirals. Nevertheless, enzymes like the macrodomain[14,15] and the papain-like protease of Nsp3[16], and the MTases Nsp10–Nsp16 complex[17] and Nsp14 play key roles in the virulence of SARS-CoV-2.[18–21] While they have little precedence as antiviral drug targets, they seem attractive as novel enzymes for antiviral drug discovery.

The Nsp14 SAM-dependent MTase catalyzes the methylation of the N7 position of the terminal guanine of viral RNA, forming a cap-0 structure similar to those in eukaryotic mRNA, which are required for translation.[18,22–26] Subsequently, Nsp10–Nsp16 methylates the 2′-O of the cap ribose to form cap-1 on the 5′ end. The capping of viral RNA by Nsp14 evades the host innate immune response to viral RNA, while ensuring efficient ribosome binding and engagement of the host–translational complex. Deletion of Nsp14 is thought to eliminate viral virulence, confirming its importance and potential status as a SARS-CoV-2 drug target.[18,19] Similarly, methylation by Nsp16 shields SARS-CoV-2 from recognition by the RNA sensor MDA5 and protects the virus from the interferon-induced antiviral response.[27] The critical role of RNA methylation in enabling translation and conferring immune protection suggests that methyltransferase inhibitors may work through multiple mechanisms to inhibit the virus.

Both cap-forming methyltransferases are highly conserved across seven coronaviruses that infect humans—SARS-CoV, SARS-CoV-2, MERS-CoV, 229E, HKU1, OC43, and NL63. Sequence alignment of Nsp16 from pathogenic coronaviruses indicates that SARS-CoV-2 Nsp16 shows 57, 58, 66, 64, 66, and 93% sequence identity to the corresponding C2′-O methylating enzymes in 299E, NL63, OC43, HKU1, MERS-CoV, and SARS-CoV-1, respectively.[28] Similarly, SARS-CoV-2 Nsp14 shows 53, 55, 62, 61, 63, and 96% sequence identity to its orthologs in 299E, NL63, OC43, HKU1, MERS-CoV, and SARS-CoV-1. Such a trend of Nsp14–Nsp16 conservation has been shown for a much larger number of coronaviruses.[29] This may indicate that viruses with inactivating Nsp14–Nsp16 mutations may have less chance of having an efficient replication and survival. Therefore, targeting these methyltransferases could lead to development of pan-coronavirus therapeutics, a novel path perhaps in development of antivirals. The essential role of RNA capping in coronaviruses suggests that small molecules that target RNA methyltransferases could have pan–viral activity against coronaviruses. In this regard, MTase domains of Nsp14 and Nsp16 are particularly appealing targets due to a high degree of overall fold similarity and a strong conservation of residues in SAM binding sites across all seven coronaviruses.[30,31]

If SAM-dependent MTases have little precedence in antiviral development, they have long been targeted in cancer chemotherapy.[32–34] Methyltransferases are druggable, and more than 20 chemical probes (potent, selective, and cell-active inhibitors/ligands) have been reported for human protein lysine (PKMT) and arginine (PRMT) methyltransferases within the last decade.[35] These chemical probes cover a range of SAM, peptide, or SAM-dependent peptide competitive mechanism of actions. Contribution of SAM binding in such ligand interactions has been previously reviewed.[36] This clearly supports the likelihood of successful Nsp14

and/or Nsp16 potent and selective inhibitor development. The binding determinants of these enzymes have been explored, especially in the SAM site, and several inhibitory analogs of the co-factor are available.[36] This has supported the determination of the crystal structure of Nsp14 complexed with *S*-adenosyl homocysteine (SAH),[23] along with other structures,[23,37] and the development of an enzyme inhibition assay.[38,39] The latter revealed relatively potent SAM-like inhibitors as others also have,[40,41] some even with antiviral activity. However, most of these were relatively large and charged, likely reducing permeability and bioavailability, as SAM analogs are likely to have activities against other MTases, especially the human class I MTases.[39] This makes them problematic as leads to new antivirals.

Building off the structural and enzymatic work, we sought to discover novel scaffolds, dissimilar to that of the SAM-like inhibitors previously investigated, which would complement the Nsp14 structure with better physical properties than the SAM analogs. We adopted a structure-based docking approach, where large libraries of "tangible", make-on-demand molecules were fit into the SAM binding site of Nsp14. Those that fit well[42–46] were prioritized for synthesis and testing. Three libraries were docked: one composed of up to 1.1 billion lead-like molecules,[47] one composed of 16 million fragment-like molecules, and a newly constructed library of 25 million electrophiles, which might covalently modify the active-site Cys387 (Figure 1). Inhibitors emerged from all three campaigns, and subsequent structure-based optimization led to several classes of low μM inhibitors. Methodologically, it was interesting that the fragment screen revealed perhaps the most diverse set of compounds, and the set most useful for characterizing the binding site. This has also been seen for other SARS-CoV-2 targets such as M[pro48] and Mac1[14,15] and is a point to which we will return.

RESULTS

Ultra-Large Library Docking against Nsp14 Identifies Novel Inhibitors.

Due to the lack of SARS-CoV-2 Nsp14 protein structure when this project began, we initially used the N7-MTase domain of SARS-CoV-1 Nsp14 (PDB ID 5C8S)[23] for retrospective control calculations, which helped us to validate the recognition of known ligands. The SAM binding site of SARS-CoV-1 Nsp14 was used without any modifications as the active site residues are conserved in both SARS-CoV-1 and SARS-CoV-2 N7-MTase domains (Figure 1A). These control calculations confirmed that we could preferentially dock known MTase adenosyl-containing compounds (SAM, SAH, and Sinefungin) and other known MTase inhibitors including LLY283,[49] BMS-compd7f,[50] and Epz04777[49–53] in favorable geometries with high ranks versus 300 property matched decoys.[54,55] Results of these retrospective calculations supported an ability to capture MTase adenosyl-containing compounds in sensible geometries relative to property-matched decoys (Figure S1). The structure used in these retrospective control calculations was subsequently supported by the recently determined SAM binding pocket of the N7-MTase domain of Nsp14 of SARS-CoV-2 (PDB ID 7N0B);[37] the C_a carbons of the two domains superpose with rmsd of 0.938 Å.

Seeking non-covalent inhibitors, we first docked over 680 million molecules, mostly in the "lead-like" range of the ZINC20 database (e.g., molecular weight 350 amu, $c \log P$

values ≥ 3.5). Each library molecule was sampled for complementarity, in an average of 3438 orientations and for each of these about 187 conformations—over 3.6×10^{14} ligand configurations were sampled in the site in 121,018 core hours (about 5 days on 1000 cores). Seeking novel chemotypes, molecules topologically similar to SAM analogs were discarded. Compounds remaining were clustered based on ECFP4 fingerprints to identify unique chemotypes. Most cluster representatives were prioritized for interactions with Trp292, Gly333, Asn334, Asp352, Ala353, Phe367, Tyr368, and Val389 using LUNA,[56] where Trp292, Phe367, and Tyr368 engage in π–π stacking with the ligands and Gly333, Asn334, Asp352, Ala353, and Val389 participate in polar interactions in the active site. Molecules with strained conformations were deprioritized.[46] Of the remaining molecules, the best scoring 5000 were visually inspected for key interactions and for unfavorable features, such as uncomplemented polar groups buried in the active site, using Chimera.[57] Ultimately, 93 molecules, each in a different scaffold, were de novo synthesized and tested for enzyme inhibition at 30 and 50 μM, measuring the transfer of [^{3}H]-methyl from the SAM methyl donor onto the cap structure of an RNA substrate (GpppAC$_4$) (Table S7). Of the 93 molecules tested, only **ZINC475239213** ('**9213**) inhibited by more than 50% and was considered active. This molecule had an IC$_{50}$ of 20 μM in concentration–response (Figure 2A, middle panel). In its docked pose, the base-like moiety of '**9213** hydrogen-bonds with backbone amides of Ala353, Phe367, and Tyr368, while more distal parts of the molecule hydrogen-bond with Gln354 and Lys336 (Figure 2A, right panel). Van der Waals and stacking interactions are also apparent in the docked pose; overall, these interactions resemble those observed among SAM and established SAM-like inhibitors but are made with different ligand groups.

With the determination of the cryo-EM structure of the SARS-CoV-2 Nsp10-Nsp14 complex (PDB ID 7N0B),[23] and the development of a larger "tangible" ZINC22[58] library of 1.1 billion molecules, we launched a second docking screen. The 680 million molecules previously docked were a subset of this 1.1 billion molecule set; the key differences in docking came from the change in orientation of the side chain of Asp352 in the cryo-EM structure of SARS-CoV-2 Nsp14 (PDB ID 7N0B).[23] In the SARS-CoV-2 Nsp14 cryo-EM structure, the side chain of Asp352 faced away from the pocket not engaging the ligands, changing the charge preference in the binding pocket in addition to the conformation of the control ligands in the docked pose. The same retrospective control calculations were performed to optimize docking parameters. Following the same prioritization strategy as before, but seeking different chemotypes, 72 diverse molecules were synthesized and experimentally tested for enzyme inhibition (Table S7). All molecules were filtered for novelty using the ECFP4-based Tanimoto coefficient (Tc < 0.35) against known inhibitors of methyltransferases and annotated aminergic ligands. The remaining molecules were further clustered by an ECFP4-based Tc of 0.5. The resulting 72 molecules come from both the original 680 million docked previously (though they ranked differently owing to the different parameterization of this new docking screen) and from the new molecules making up to the 1.1 billion now docked. Out of the 72, two inhibitors were found, **ZINC730084824** ('**4824**) with an IC$_{50}$ of 50 μM and **ZINC61142882** ('**2882**) with an IC$_{50}$ of 6 μM (Figure 2B,C), a hit rate of 2.7%. The origins of the low hit rates for these two initial screens, and strategies to improve upon them, will be considered below.

Optimization of the Lead-Like Compounds.

To improve the affinity of the three docking actives, we sought analogs among the 20 billion tangible molecules that have been enumerated in a version of the REAL database (http://enamine.net/compound-collections/real-compounds/real-space-navigator), using substructure and similarity searches in the SmallWorld (http://sw.docking.org) and Arthor (http://arthor.docking.org) search engines (NextMove Software, Cambridge UK).[59] Conservative analogs were prioritized, their structures and physical properties were calculated, and these were then docked into the Nsp14 SAM site. Analogs that docked to interact via π–π stacking with Phe367 or Trp292 and that appeared to hydrogen-bond with Tyr368, Ala353, Asp352, and Asp388 were prioritized. Overall, 12, 20, and 36 analogs of '**9213**, '**4824**, and '**2882**, respectively, were synthesized and tested for enzyme inhibition (Table S7, Figure 3). The affinities of one of the analogs of '**2882** were similar to the initial hit, with an IC_{50} of 2.2 μM for compound '1988 (Figure 3C, Table S1). The docked pose of '**1988** suggests a new hydrogen bond with Asp352, while conserving previous hydrogen bonds with Tyr368 and Ala353 (Figure 3C). Two '**9213** analogs, **Z5347169163** ('9163, IC_{50} 15 μM) and **ZINC001342858621** ('8621, IC_{50} 14 μM), also had similar affinities as the lead molecule (Figure 3A, Table S2). For '**4824**, two-fold improvement was observed for analogs **ZINC000916131631** ('1631, IC_{50} 25 μM) and **Z5347186947** ('6947, IC_{50} 19 μM) (Figure 3B, Table S3). Even though improvements were modest, the SAR was revealing. While little improvement was seen over the parent '**4824** or '**9213**, for instance, many of the analogs tested remained relatively potent, with IC_{50} values often below 40 μM (Tables S1–S3). Moreover, replacement of the Michael acceptor vinyl group of '**4824** in analogs '**6947** and **Z5347186943** ('6943, IC_{50} 27 μM) and removing the nitrile in analogs '**1631** and '**6947** left molecules that remained active. In '**9213**, this was also the case with the removal of the nitrile group in analogs '**9163** and '**8621**. These analogs support the idea that docking hits '**9213** and '**4824** are acting via a non-covalent mechanism, as modeled. Assessment of reversibility by jump dilution also suggests that inhibition by '**4824** is not due to the nitrile electrophile (Figure S2). '**1988** also appeared to be a noncovalent inhibitor in rapid dilution experiments (Figure S2). In addition, '**1988** and '**4824** showed a SAM- and RNA-competitive pattern of inhibition (Figure S3). Taken together, these results support the non-covalent nature of binding of these families of molecules.

We tested the docking hits and analogs for colloidal aggregation, perhaps the dominant mechanism of artifactual activity in early discovery[60–62] (Figures S4–S8, Table S4). As a first line of defense, all Nsp14 assays were conducted in the presence of 0.01% v/v Triton-X 100, a non-ionic detergent that disrupts colloidal aggregates and right-shifts their potency.[63,64] We also conducted follow-up assays for actives looking for particle formation by dynamic light scattering (DLS) and for activity against the widely used counter-screening enzymes malate dehydrogenase (MDH) and AmpC β-lactamase (AmpC), both with and without detergent. The '**9213** and its analog, '**3888**, did form colloid-like particles by DLS, with an apparent critical aggregation concentration (CAC) in the 10 to 30 μM range. Like most of the inhibitors studied here, however, the scattering intensity was relatively modest, and the molecule did not inhibit the counter-screening enzymes at concentrations substantially higher than the IC_{50} for Nsp14, even in the absence of detergent. While '**3888** may form particles, we do not believe these are relevant for its inhibition of Nsp14.

Compound '**4824** did not form detectable particles by DLS, but did inhibit the counter-screening enzymes MDH and AmpC at relevant concentrations. However, this inhibition can be at least partly attributed to strong absorbance at assay wavelengths (such absorbance was not an issue for the Nsp14 radioligand assay). Moreover, when we imitated the conditions of the Nsp14 assay by the addition of 0.01% Triton to the MDH and AmpC assays, inhibition was largely or entirely eliminated. Meanwhile, the '**4824** analogs ('**6947**, '**6953**, and '**6943**) did form colloid-like particles by DLS, though again with relatively modest scattering intensities. All three inhibited MDH or AmpC at relevant concentrations, but here too inhibition was largely eliminated by the addition of 0.01% of detergent. Compound '**1988** formed colloid-like particles by DLS, but with a CAC 10-fold higher than its Nsp14 IC_{50}. While the compound's inhibition of MDH was in a relevant range, it's activity against AmpC was not, and the inhibition of both enzymes disappeared when we copied the Nsp14 conditions by the addition of Triton. We conclude that many of these inhibitors do aggregate, but this does not appear to be relevant for their inhibition of Nsp14, for which the inclusion of detergent appears to be prophylactic. These studies do support the usefulness of including detergents like Triton or Tween in enzyme and receptor inhibition assays.

Docking 16 Million Fragment-Like Molecules.

With only three inhibitor scaffolds discovered by lead-like docking, we stepped back to interrogate the site with fragment-based docking. Fragment screens explore more chemical space than a larger lead-like library,[14,65,66] which may be helpful for an under-explored site where warheads and key residue interactions have not been characterized. With the proviso that they have lower affinities, fragments also have higher hit rates in empirical[67] and docking screens[65,66,68] than do lead-like molecules, providing a richer tiling of the binding site by ligand functional groups. Indeed, a strategy of fragment-docking was effective against another under-studied SARS-CoV-2 enzyme, Mac1,[14,15] and fragment-based discovery nucleated a successful drug-discovery campaign against the M^{pro} enzyme.[10] Accordingly, from the 16 million molecule fragment-like set (e.g., molecular weight 250, $c \log P$ 2.5) in ZINC22, we targeted the full SARS-CoV-2 (PDB 7N0B) SAM site, the adenine portion of that site, and the SAM-tail region in three independent campaigns (Figure 4A) (Methods).[69] Overall, 14,406,946, 14,124,978, and 14,908,652 million molecules were scored, respectively. For each, the top-ranked 300,000 fragments were filtered as above, and the remaining fragments were clustered by topological similarity. Top-ranking cluster heads were visually inspected in Chimera[57] for favorable interactions, prioritizing those in the adenine site campaign for hydrogen bonds to Tyr368 and Ala353, and hydrophobic interactions with Phe367.[56] For the SAM-tail docking screen, interactions with Gly333 were prioritized, and additional interactions were selected for such as Gln313 and Asn386. For fragments docked against the entire SAM binding site, a combination of these interaction criteria were used. Ultimately 69 fragments were prioritized, of which 54 were successfully synthesized (78% fulfilment rate) (Table S1).

Of these 53, 9 fragments were hits with at least 50% inhibition at 300 μM and had IC_{50} values ranging from 12 to 341 μM (ligand efficiencies (LE) 0.29 to 0.42 kcal/heavy atoms). The most potent fragment, '**0683**, had an IC_{50} of 12 μM and an LE of 0.42 kcal/heavy atom (Figures 5 and S9). As with the larger lead-like inhibitors (above), '**0683** was a

competitive and presumably reversible inhibitor of both SAM and RNA binding (Figure S3). In their docked poses, the fragment inhibitors were modeled to dock in the adenine-site forming hydrogen bonds with Tyr368, Phe367, or Ala353, often mimicking interactions of the adenine of SAM/SAH but with different functional groups and with diverse chemotypes (Figures 4 and S10). One of the active fragments, '**6066**, emerged from docking to the SAM-tail site, with hydrogen bonds to Gly333 and Gln313; we note that its activity was lower than most of those in the more tightly defined adenine site, with an IC_{50} of 341 μM (Figures 4B,C,E and S9).

Fragment inhibitors were also evaluated for colloidal aggregation, as described above (Figures S4–S8, Table S4). While several did have CACs at relevant concentrations ('**0222, '5604, '5763, '9744**, and '**6066**), or inhibited MDH ('**0741** and '**5604**) or both MDH and AmpC ('**0683**) in the absence of detergent, most did not inhibit AmpC under any measured concentration, and on addition of detergent, mimicking the Nsp14 assay, MDH inhibition was largely or entirely eliminated (this effect was smaller for '**0741**). Here too we believe that while under some conditions these fragments can form colloidal particles, such aggregation is not relevant for the Nsp14 inhibition we observe.

Curation of 25 Million Aldehyde and Acrylamide Electrophiles for Covalent Docking.

In a final strategy, we sought potential covalent electrophiles that could react with the enzyme's active site Cys387. Such covalent docking has been successful in campaigns that targeted catalytic serine and non-catalytic, active site cysteine and lysine residues in enzymes such as β-lactamase, Jak kinases,[70] eIF4e,[71] M[pro],[72] and targets such as RSK2 and MSK1.[73] These earlier campaigns had been limited to several hundred thousand electrophiles, largely from "in-stock" libraries. With the advent of the ultra-large tangible libraries, we thought to curate a larger set of electrophiles, focusing on aldehydes and acrylamides. Searching smart patterns allowed us to build databases of 7.3 million aldehydes and 17.7 million acrylamides. We compared our aldehyde and acrylamide libraries to those that can be found in other in-stock or physical screening libraries, including the UCSF Small Molecule Discovery Center (SMDC),[74] Molecular Libraries Small Molecule Repository of the NIH (MLSMR),[75] and the in-stock set curated in ZINC20.[59] By total numbers, the aldehyde library is 196- to 10,000-fold larger than the number of aldehydes in the other libraries, while the number of core scaffolds represented by these electrophiles is 252- to 3600-fold larger that those sampled in the previous libraries (Table 1). For acrylamides, there are 811- to 465,000-fold more electrophiles in the new library than in the public and in-stock sets, encompassing 250- to 58,000-fold more scaffolds. For the acrylamides, there are on average 13 molecules per scaffold in the new libraries, compared to 2–4 per scaffold in the other libraries. For aldehydes, the databases are comparable with our new library averaging 7 molecules per scaffold and 2–9 molecules per scaffold in the other databases. Given the rising interest in covalent-based inhibitors,[73,76] we have made these 25 million electrophiles openly available via http://covalent2022.docking.org in both 2D and DOCKovalent 3D format (Methods), along with ZINC and Enamine codes for ready acquisition.

Covalent Inhibitors from 25 Million Docking Screen against Cys387.

17.7 million acrylamides and 6.2 million aldehydes were docked against the SAM site adjacent to Cys387, using DOCKovalent.[70,71] Molecules were docked to form a covalent adduct with Cys387. Those with non-covalent DOCK3.7 scores <0 kcal/mol were further filtered for internal strain,[46] stranded hydrogen bond donors and acceptors, and modeled hydrogen bonds with either Tyr368, Ala353, or Gly333.[56] Lastly, 33,156 molecules were clustered for topological similarity, and 9591 molecules were prioritized for visual inspection in Chimera.[57] From these, 92 molecules were selected for de novo synthesis. Of 61 aldehydes and 31 acrylamides, 47 and 26 were successfully synthesized, respectively, a 79% fulfilment rate (Table S7). On experimental testing, hits were defined as having at least 50% inhibition at 100 μM. For the aldehydes, five compounds were active of 51 tested (a hit rate of 10%) and had IC_{50} values ranging from 3.5 to 25 μM (Figures 5 and S11 and S12). The most potent were '**4975** and '**1911** with IC_{50} values of 3.5 μM and 8, respectively (Figure 5A). These aldehydes were selected to hydrogen-bond with Ala 353 or Gly333 in their docked poses and form additional hydrogen bonds in the pocket, including with Asn388, Arg310, Gln313, or Trp385 (Figures 5D and S12). Of the acrylamides, 2 of 26 tested had >50% inhibition at 100 μM, for a hit rate of 8% (Figures 5C and S11). Inhibitor **acryl42** and **acryl41** had an IC_{50} values of 32 and 39 μM, respectively. Acryl42 hydrogen-bonds with Ala353 and Gly333 in its docked pose, while acryl41 is a longer molecule extending farther than SAM/SAH, making hydrogen bonds with Gly333 and Gly313 in its docked pose (Figure S12).

In early optimization of **acryl42**, analog **acryl42_10** was 4.5-fold more potent at 7 μM with the addition of a methoxy (Figure 5C). Adding a hydroxyl in the same place in analog **acryl42_11** resulted in an inactive analog, indicating that the methoxy could be adding hydrophobic contacts, as opposed to additional hydrogen bonds with the protein (Table S5). We tested the importance of the free amide of the acrylamide warhead with methylation of analog **acryl42_5**; the analog was inactive, perhaps reflecting the loss of a modeled hydrogen bond with the mainchain of Ala353.

We evaluated the covalent mechanism for the most potent covalent docking hits, aldehyde '**1911** and acrylamides **acryl41** and **acryl42**, and the **acryl42** analog, **acryl42_10**, first by mass spectrometry analysis. '**1911** did not increase the molecular mass of Nsp14 (Figure S13), likely reflecting the reversible binding of aldehydes to cysteines (Cys387 of Nsp14). In rapid dilution enzymatic experiments, '**1911**, incubated at high concentrations, showed little residual inhibition when diluted below its IC_{50}, further supporting a reversible covalent mechanism (Figure S2). While acryl41 did not form a measurable adduct by mass spectrometry, **acryl42** and its analog **acryl42_10** did do so, supporting a covalent inhibition mechanism (Figure S13). We also changed the acrylamide warhead to the saturated propanamide group in compound **ZD160–68**, resulting in no enzymatic inhibition, which furthered support for **acryl42** acting through covalent inhibition (Figure 5C). Overall, **acryl_42** and its analog, **acryl42_10**, appear to be irreversible covalent inhibitors, while '1911 appears to be a reversible covalent inhibitor. We expect that **acryl_41** is also acting as a covalent inhibitor, but note that further mechanistic study of these classes is warranted.

The covalent inhibitors were evaluated for colloidal aggregation (Figures S4–S8 and Table S4). The 12 μM aldehyde Z5185631889 ('**1889**) (Figure 5) had a CAC five-fold higher than its Nsp14 IC$_{50}$ and did not inhibit either counter-screening enzyme under any measured condition—if this compound aggregates, it is not relevant for its Nsp14 activity. While the 3.8 μM aldehyde '**1911** (Figure 5) did form colloid-like particles by DLS at a relevant concentration and did inhibit AmpC and MDH in the absence of Triton, this inhibition disappeared on the addition of the same amount of detergent (0.01% v/v) used in the Nsp14 assay. While this molecule likely is an aggregator, its aggregation is unlikely to be relevant to its Nsp14 inhibition.

Selectivity against Human Protein and RNA Methyl-transferases.

Three of the most potent inhibitors were counter-screened against a panel of 30 human SAM-dependent MTases. Compounds were tested for inhibition of the enzymes at 10 μM and then selected for IC$_{50}$ determination if higher than 50% inhibition was observed. The non-covalent, 6 μM lead-like inhibitor, '**1988**, showed only modest selectivity, inhibiting nine enzymes more than 50% with IC$_{50}$ values ranging from 4 to 26 μM (Figure 6). The apparently reversible covalent, 8 μM, inhibitor, '**1911** had much better selectivity, inhibiting only two histone MTases G9a and G9a-like protein (GLP) with IC$_{50}$ values of 30 and 14 μM, respectively. The most selective compound was '**4824** with no inhibition greater than 50% for any enzyme in the panel.

DISCUSSION AND CONCLUSIONS

This study revealed the first Nsp14 inhibitors unrelated to SAM, either topologically or by physical properties. Overall, 23 non-covalent, lead-like inhibitors across three scaffolds were found with IC$_{50}$ values less than 50 μM, providing SAR for additional optimization (Figures 2 and 3, Tables S1–S3). Additional characterization and structure-based optimization demonstrated their competitive, non-covalent mechanism of action against Nsp14 (Figures S2 and S3). The most active covalent inhibitors were the initial aldehyde docking hits, with IC$_{50}$ values ranging 3.5 to 12 μM, and the acrylamide analog **acryl42_10** with an IC$_{50}$ of 7 μM, all modeled to modify Cys387 of Nsp14 (Figure 5). Considering that these depended on developing new tangible libraries of 25 million electrophiles, these have been made publicly available for community use (https://covalent2022.docking.org) (Table 1). Another eight families of inhibitors were revealed from docking a library of 16 million tangible fragments (Figure 4). While affinities were naturally lower than the best of the lead-like inhibitors, several fragments had mid-μM IC$_{50}$ values, and the four most potent had LEs 0.32 to 0.42 kcal/HAC. Taken together, 19 new chemotypes were found; of these, 11 had members with IC$_{50}$ values < 50 μM.

SARS-CoV-2 Nsp14 inhibitors described to date are SAM analogs[39–41,77,78] or fragments with extensive water networks.[79] While the SAM analogs are widely studied, they typically suffer from both low cell permeability, owing to their size and ionization state, and from low selectivity, owing to their high similarity to the shared co-factor of this large family of MTases. Conversely, the new molecules described here are smaller and mostly uncharged and topologically unrelated to SAM (Table S6). These properties may support optimization

for cell permeability and bioavailability and for selectivity. Consistent with this idea, in counter-screens of 30 SAM-dependent human protein and RNA MTases, '**4824** was Nsp14-selective, and '**1911** only hit two very closely related MTases (G9a and GLP).

Apart from the particular inhibitors discovered, lessons from this work may be useful for the other less-studied SARS-CoV-2 targets, which are the great majority of the viral genome. Unlike enzymes for which an investigational drug had been developed via similarities with other viruses, such as M[pro] or RdRp,[2–5,80] the MTase of Nsp14 had little inhibitor precedence on which to draw. As a SAM-dependent enzyme with many related human enzymes, chemical novelty was important. Thus, as may be true with many SARS-CoV-2 targets, we could not leverage knowledge from previous chemical series, other than from SAM analogs (and these we needed to avoid). Meanwhile, the methyltransferase active site itself is large with multiple key interacting residues. The lack of knowledge of critical residues that determine ligand potency and selectivity, the relative lack of known inhibitors with which to do control calculations, the need for chemical diversity from SAMs, and the large size of the pocket combine to make Nsp14 target challenging for docking.

Initial modeling lacked chemical precedence, and a small number of successes in early docking campaigns enabled us to optimize subsequent ones, contributing to improved hit-rates and affinities. We do note that our most informative screens—against the 16 million tangible fragments—occurred late in the campaign. Whereas there may still be skepticism about fragment docking, our own experience, not only here but also against the SARS-2 enzyme Mac1[14,15] and in earlier studies against β-lactamases,[65,66,68] is that fragment docking can reveal multiple chemotypes with high-ligand efficiency and fidelity to subsequently determined crystal structures. Were we to begin again, we might have started with the fragment screen, leveraging the interactions it revealed for campaigns against the larger, lead-like libraries. Such an approach may be useful against other understudied viral targets.

There are certain caveats. Our most potent inhibitors are low-μM, weaker than the most potent of the SAM analogs previously characterized for Nsp14, the best of which inhibited in the 100 nM range.[39,77,78] '**1911** needs additional characterization of its reversible covalent mechanism of inhibition, limited here by its reversibility in mass spectrometry analysis, and low-μM activity in the rapid dilution experiments. Some of the more potent compounds, including **acryl42_10**, '**4975**, '**0683**, '**1988**, and '**1911**, were tested for antiviral activity, but for none of them could we distinguish anti-viral activity from cell toxicity—improving affinity and reducing general cellular toxicity are clearly goals for subsequent optimization of these series. At the in vitro level, we note that many of the inhibitors form colloidal aggregates, which would ordinarily be a concern for selectivity and artifactual activity. Control experiments suggest that such aggregation is not relevant for Nsp14 inhibition. Still, it remains true that this activity must be controlled for in subsequent optimization, and is a general hazard to navigation in early discovery.

These caveats should not obscure the key observation of this study, the discovery of 19 new families of Nsp14 inhibitors. These new inhibitors are not only diverse, but they do not resemble the SAM-related molecules previously described for Nsp14 either topologically or

by physical properties. They represent both non-covalent and covalent families, as well as fragments that tile the binding site. With the ongoing pandemic, they are being made openly available without restriction in the hopes that they may support a broad attack on this key but understudied target for antiviral drug discovery against Covid-19.

EXPERIMENTAL SECTION

Non-covalent Ultra-Large Scale Docking.

The N7-MTase domain of SARS-CoV Nsp14 (PDB ID 5C8S) (12) and the N7-MTase domain of SARS-CoV-2 Nsp14 from cryo-EM structure of the Nsp10-Nsp14 complex (PDB ID 7N0B) were used in two docking campaigns of >680 million "lead-like" molecules from the ZINC20 database (http://zinc20.docking.org)[59] and the ZINC22 > 1.1 billion "lead-like" molecules (http://files.docking.org/zinc22), respectively, using DOCK3.7.[45] Forty-five matching spheres or local hot-spots generated from the crystal pose of SAM/SAH were used in the binding site for superimposing pre-generated flexible ligands, and the different poses were scored by summing the different energies including; van der Waals interaction energies, Poisson–Boltzmann-based electrostatic interaction, and Generalized-Born/surface area-based ligand desolvation energies.[43,44] Receptor atoms were protonated with reduction,[81] and partial atomic charges were calculated using united-atom AMBER force field.[82] AMBER atom-types were also used for Poisson–Boltzmann electrostatic potential energy grids using QNIFFT,[83] CHEMGRID[84] was used for calculating van der Waals potential grids, and SOLVMAP[43] was used to calculate the Generalized-Born/surface area grids for ligand desolvation.

The docking setup was optimized for its ability to enrich knows MTase adenosyl group-containing compounds (SAM, SAH, and Sinefungin) and other known MTase inhibitors including Lly283, BMS-compd7f, and Epz04777,[49–53] in favorable geometries with high complementarity versus a set of property matched decoys.[54,55] About 50 decoys were generated for each ligand that had similar chemical properties to known ligands but were different topologically. The best optimized docking setup was evaluated for enrichment of ligands over decoys using the log-adjusted area under the curve (log AUC values).[54,55] All docked ligands were protonated with Marvin (version 15.11.23.0, ChemAxon, 2015; https://www.chemaxon.com) at pH 7.4, rendered into 3D with Corina (v.3.6.0026, Molecular Networks GmbH; https://mn-am.com/products/corina/), and conformationally sampled using Omega (v.2.5.1.4, OpenEye Scientific Software; https://www.eyesopen.com/omega). Before docking the lead-like libraries, an "extrema set"[54,85] of 61,687 molecules was docked in the optimized system to ensure that the molecules with correct physical properties were enriched.

Overall, in the prospective screen, each library molecule was sampled in about 3438 orientations; on average, about 187 conformations were sampled over 5 days on 1000 cores. The top-ranking 300,000 molecules were filtered for novelty using the ECFP4-based Tanimoto coefficient (Tc < 0.35) against known inhibitors of MTases. The remaining molecules were then clustered into related groups using an ECFP4-based Tc of 0.4. From the top 10,000 novel chemotypes, molecules with >2 kcal mol^{-1} internal strain[46] were excluded and the remaining candidates were visually inspected for the best docked poses

with favorable interactions with the MTase active site. Ultimately, overall 165 molecules were selected for de novo synthesis and testing.

Non-covalent Optimization.

Analogs for docking hits '**2882**, '**9213**, and '**4824** were queried in Arthor and SmallWorld 1.4 and 12 billion make-on-demand libraries (http://sw.docking.org, http://arthor.docking.org), the latter primarily containing Enamine REAL compounds (http://enamine.net/compound-collections/real-compounds/real-space-navigator). The resulting analogs were further filtered based on Tc > 0.4 and docked to the N7-MTase domain of SARS-CoV-2 Nsp14. Compounds were also designed by modifying the 2D structure and custom synthesis by Enamine Ltd. (Kyïv, Ukraine). The docked poses were visually inspected for compatibility with the site, and prioritized analogs were synthesized and tested for each series (Table S7).

Fragment Docking.

The optimized docking setup from the SARS-CoV-2 second non-covalent lead-like screen described above was used. Three different screens were run with different matching spheres[69]—those in the adenine-site, SAM-tail site, or all matching spheres (Figure 4A) —15,738,235 docked and 14,406,946 scored, 15,738,278 docked and 14,124,978 scored, and 16,299,173 docked and 14,908,652 scored, respectively. Each setup was analyzed separately until visualization in Chimera[57]—the top 300,000 ranked poses were filtered for having torsional strain less than 7 TEU total, single strain of 2.5 TEU,[46] less than 2 stranded hydrogen bond donors, less than 4 stranded hydrogen bond acceptors, and greater than 1 hydrogen bond to Tyr368, Ala353, or Gly333.[56] Remaining molecules were visually inspected for having favorable interactions. In total, 65 compounds were selected for purchasing, 50 from Enamine and 19 from WuXi, and overall, 53 were successfully synthesized for a fulfilment rate of 82%.

Covalent Database Curation.

SMARTS patterns for aldehydes or acrylamides ([CX3H1](=O)[#6] and ([CD1]=[CD2]–C(=O)–[NX3]), respectively) were searched in Enamine REAL databases, finding 20 million acrylamides and 6 million aldehydes. The DOCKovalent 3D files were generated as previously described.[71–73] Briefly, the electrophiles were converted to their transition state product and a dummy atom was placed, indicating to the docking algorithm which atom should be modeled covalently bound to the sulfur of the cysteine. Both 2D structures and 3D DOCKovalent files are now publicly available at http://covalent2022.docking.org. To compare with other public molecule databases, we used the ZINC20 in-stock set,[59] the MLSMR library,[75] and the UCSF SMDC library[74] and searched the same SMARTS patterns for acrylamides and aldehydes. The number of chemotypes were determined by Bemis–Murcko clustering.[86]

Covalent Docking and Compound Optimization.

The optimized docking setup from the first SARS-CoV-1 lead-like screen described above was used, with differences being which residues have been hyper-polarized[85] (Tyr368,

Tyr368, and Ala353 or Tyr368, Ala353, and Gly333, referred to as 1-HP, 2-HP, and 3-HP, respectively). For the acrylamide screen against 1-HP, molecules with docked scores less than 0 were selected for filtering (top 341,000); those with internal torsional strain less than total strain of 6.5 TEU and single strain of 2 TEU,[46] molecules with less than 2 stranded hydrogen bond donors, and less than 4 stranded hydrogen bond acceptors were prioritized. Molecules were also selected that formed at least one hydrogen bond to Tyr368, Ala353, or Gly333 using LUNA[56] leaving 2423. After clustering for chemical similarity, 533 were visually inspected in Chimera.[57] For the 2-HP setup, molecules with scores less than 0 (top 440,661) were filtered using the same criteria with 2961 molecules remaining, comprising of 622 clusters that were visually inspected. For the 3-HP setup, no molecules passed the strain, IFP, and hydrogen bond filter and were not considered further. Visual inspection prioritized molecules with the same criterion as above. Lastly, selected compounds from both 1-HP and 2-HP setups were clustered to select unique chemotypes, and 31 were purchased. Synthesis was successful for 26 for a fulfilment rate of 84%.

For the aldehydes in the 1-HP setup, the top 894,979 compounds (dock score less than 0) were filtered to be prioritized as the acrylamides were above, with clustering for chemical similarity leaving 1340 for visual filtering. For the 2-HP setup, the top 1,494,350 were filtered to 3548, and 3-HP setup of top 1,494,345 to 3548 for visual inspection. Compounds were prioritized for the same interactions as the acrylamides, and finally 61 aldehydes were selected. Synthesis was successful for 47 of these for a fulfilment rate of 77% and an overall covalent fulfillment rate of 79%.

Acryl42 analogs acryl42_5, acryl42_11, and acryl42_10 were designed off the acryl42 2D chemical structure and synthesized by Enamine; ZD160–68 was designed to test the activity of the acrylamide warhead. The modeled pose of acryl42_10 was performed in Maestro (version 2021–2, Schrödinger, Inc.) by manually changing the acryl42 docked pose to acryl42_10 and minimizing the Nsp14-acryl42_10 complex using the Protein Preparation Wizard protocol.

Chemistry.

Make-on-Demand Synthesis.—Compounds were purchased from Enamine Ltd. (Kyiv, Ukraine) or WuXi Appetec (Shanghai, China). Purities of active molecules were >95% (assessed by LC/MS) and inactives typically >90% (Table S7 and Figure S14).

General Synthetic Procedures for Enamine Compounds.—All solvents were treated according to standard methods. ^{1}H NMR spectra were recorded at 400, 500, or 600 MHz (Varian or Bruker spectrometers). ^{1}H chemical shifts are calibrated using residual undeuterated solvent DMSO: δ = 2.50 ppm. Coupling constants are given in Hz. LC/MS analysis was performed utilizing the Agilent 1200 Series LC/MSD system with DAD/ELSD (column Zorbax SB-C18 1.8 μm, 4.6 × 15 mm); solvent A (water, 0.1% formic acid) and solvent B (acetonitrile, 0.1% formic acid); gradient 0–100% solvent B, run time, 1.8 min; flow rate, 3 mL/min; and Agilent LC/MSD SL (G6130A), SL (G6140A) mass-spectrometer (APCI mode). All the LC/MS data were obtained using positive/negative mode switching. Method 1. An amine (100 mg) and DIPEA (1.1 mol equiv to the amine; additional

equivalents were added when amine used was in salt form) were dissolved in 0.5 mL of DMSO. The vial was then shaken for 20 min at room temperature (RT). Alkyl chloride (1 mol equiv to the amine) was then added. The vial was sealed and stirred for 1 h. Then the solution was heated for 8 h at 100 °C. After cooling down, the mixture was filtered; the solvent and volatile components were evaporated under reduced pressure to give the crude product. The product was further purified by HPLC. **Method 2.** Methylene-active compound (100 mg), aldehyde (1 mol equiv to the methylene-active compound), DMF (0.5 mL), and trimethylchlorosilane (2.1 mol equiv to the methylene-active compound) were placed in the vial. The vial was then placed in the thermostat (set to 100 °C) for 48 h. After cooling the reaction mixture down to RT, DIPEA (0.2 mL) was added to the vial, and it was stirred for 30 min. The solvents were evaporated under reduced pressure to give the crude product. The product was further purified by HPLC. **Method 3.** An amine (100 mg), an acid (1.1 mol equiv to the amine), and 0.5 mL of DMSO were placed into a 4 mL capped glass vial, and the mixture was stirred for 30 min at RT. Then 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 1.2 mol equiv to the amine) was added, and the mixture was stirred for 1 h. If the solution was transparent, the mixture was left overnight at room temperature as is; otherwise, the vial was placed in the ultrasonic bath and left overnight. The solution was filtered, and the solvent and volatile components were evaporated under reduced pressure to give the crude product. The product was further purified by HPLC. **Method 4.** An amine (100 mg), DIPEA (1.2 mol equiv to the amine), and DMSO (0.5 mL) were placed into a 4 mL capped glass vial and stirred for 30 min at RT. After the addition of an aryl halide (1.2 mol equiv to the amine), the mixture was stirred for 1 h at RT. Then the vial was placed on a thermostat (set to 100 °C) for 9 h. After cooling down, the mixture was filtered; the solvent and volatile components were evaporated under reduced pressure to give the crude product. The product was further purified by HPLC. **Method 5.** An aryl halide (100 mg), boronic derivative (1 mol equiv to the aryl halide), sodium carbonate (1.5 mol equiv to the aryl halide), 1,1′-[bis(diphenylphosphino)ferrocene]-dichloropalladium(II) (catalytic amount), and water-dioxane mixture 1:3 (1 mL) were placed in a vial. The vial was then placed into a thermostat and stirred for 24 h at 95 °C. After cooling down the mixture, the solvents were evaporated under reduced pressure. The residue was dissolved in chloroform (3 mL) and washed with water (3 × 1 mL). The catalyst was then filtered, and solvents were evaporated under reduced pressure. The product was further purified by HPLC.

(6-Bromoimidazo[1,2-a]pyridin-2-yl)methyl 5-(Furan-2-yl)-1H-pyrazole-3-carboxylate—Z793205438 (Method 1).—Yield: 15%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 14.06 (s, 1H), 8.89 (s, 1H), 8.01 (s, 1H), 7.76 (s, 1H), 7.52 (d, J = 9.5 Hz, 1H), 7.36 (dd, J = 9.6, 1.9 Hz, 1H), 7.01 (s, 1H), 6.88 (s, 1H), 6.59 (s, 1H), 5.42 (s, 2H). LC/MS (APSI) m/z: [M + H] calcd for $C_{16}H_{12}BrN_4O_3$, 389.0; found, 389.0.

N-(3-Ethynylphenyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-6-yl)benzamide—Z1723431234 (Method 3).—Yield: 36%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.48 (s, 1H), 10.22 (s, 1H), 8.26 (d, J = 2.0 Hz, 1H), 8.23 (d, J = 1.9 Hz, 1H), 7.95–7.89 (m, 2H), 7.83 (dd, J = 14.5, 8.2 Hz, 1H), 7.68

(d, J = 1.9 Hz, 1H), 7.56 (t, J = 7.7 Hz, 1H), 7.29 (t, J = 7.9 Hz, 1H), 7.15 (d, J = 7.5 Hz, 1H), 3.68 (s, 1H), 2.99 (s, 4H). LC/MS (APSI) m/z: [M + H] calcd for C22H17N4O2, 369.1; found, 369.0.

2-((4-Bromophenyl)amino)-2-oxoethyl 3-(Furan-3-yl)-1H-pyra-zole-5-carboxylate—Z795161988 (Method 1).—Yield: 38%; purity, >95% (assessed by LC/MS). 1H NMR (500 MHz, DMSO-*d6*): δ 13.84 (s, 1H), 10.33 (s, 1H), 7.78 (s, 1H), 7.72 (s, 0H), 7.55 (d, J = 8.6 Hz, 2H), 7.49 (d, J = 8.7 Hz, 3H), 7.06 (s, 1H), 6.96 (s, 1H), 4.86 (s, 1H). LC/MS (APSI) m/z: [M + H] calcd for C16H13BrN3O4, 392.0; found, 392.0.

(E)-6-(3-(4-Hydroxy-3-nitrophenyl)acryloyl)-5-methyl-2,3-dihydro-1H-pyrrolizine-7-carbonitrile—Z1143247121 (Method 2).—Yield: 29%; purity, >95% (assessed by LC/MS). ¹H NMR (400 MHz, DMSO-d_6): δ 11.54 (s, 1H), 8.22 (t, J = 2.4 Hz, 1H), 7.90 (dt, J = 8.6, 2.5 Hz, 1H), 7.56 (dd, J = 15.7, 2.4 Hz, 1H), 7.42 (dd, J = 15.6, 2.4 Hz, 1H), 7.19 (dd, J = 8.7, 2.4 Hz, 1H), 4.06–3.96 (m, 2H), 3.01–2.92 (m, 2H), 2.46 (d, J = 2.5 Hz, 3H). LC/MS (APSI) m/z: [M + H] calcd for C18H16N3O4, 338.1; found, 338.2.

N-(3-(4-(6-Chloro-2-oxoindolin-3-yl)piperidin-1-yl)-3-oxopropyl)-acrylamide—Z1479200718 (Method 3).—Yield: 35%; purity, >95% (assessed by LC/MS). ¹H NMR (500 MHz, DMSO-d_6): δ 10.50 (s, 1H), 8.11–8.01 (m, 1H), 7.22 (dd, J = 8.0, 1.0 Hz, 1H), 6.96 (dt, J = 8.0, 1.7 Hz, 1H), 6.80 (d, J = 2.0 Hz, 1H), 6.02 (ddd, J = 16.8, 13.7, 2.3 Hz, 1H), 4.43 (dd, J = 21.4, 13.5 Hz, 1H), 3.81 (t, J = 14.6 Hz, 1H), 3.46–3.41 (m, 1H), 3.27 (q, J = 6.6 Hz, 2H), 2.90 (s, 1H), 2.43 (dq, J = 12.8, 9.3, 8.7 Hz, 3H), 2.21 (td, J = 12.3, 3.6 Hz, 1H), 1.56 (d, J = 13.1 Hz, 1H), 1.43 (d, J = 13.1 Hz, 1H). LC/MS (APSI) m/z: [M + H] calcd for C19H23ClN3O3, 376.1; found, 376.0.

N-(3-Cyanophenyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo-[4,5-b]pyridin-6-yl)benzamide—Z1723430981 (Method 3).—Yield: 21%; purity, >95% (assessed by LC/MS). ¹H NMR (500 MHz, DMSO-d_6): δ 11.59 (s, 1H), 10.63 (s, 1H), 8.32 (d, J = 2.0 Hz, 1H), 8.25 (dt, J = 3.7, 1.8 Hz, 2H), 8.06 (dt, J = 7.3, 2.2 Hz, 1H), 7.97–7.89 (m, 2H), 7.83 (d, J = 2.0 Hz, 1H), 7.64 (t, J = 7.7 Hz, 1H), 7.62–7.54 (m, 2H), 3.36 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C21H16N5O2, 370.1; found, 370.2.

5-(1H-Indol-2-yl)-1-methyl-1H-pyrazol-4-amine-Z2732986066 (Method 5).—Yield: 28%; purity, >95% (assessed by LC/MS). LC/MS (APSI) m/z: [M + H] calcd for C12H13N4, 213.1; found, 213.2.

6-((Isopropylthio)methyl)-1-methyl-1H-pyrazolo[4,3-d]-pyrimidin-7(6H)-One—Z3343635604 (Method 1).—Yield: 29%; purity, >95% (assessed by LC/MS). ¹H NMR (500 MHz, DMSO-d_6): δ 8.26 (s, 1H), 7.98 (s, 1H), 5.16 (s, 2H), 4.19 (s, 3H), 3.09 (p, J = 6.7 Hz, 1H), 1.21 (d, J = 6.7 Hz, 6H). LC/MS (APSI) m/z: [M + H] calcd for C10H15N4OS, 239.1; found, 239.0.

N - (4 - (Imidazo[2, 1-b]thiazol - 6 - yl) phenyl) - 4 - (N-methylacrylamido)benzamide—Z3756698609 (Method 3).—Yield: 51%; purity, >95% (assessed by LC/MS). ¹H NMR (500 MHz, DMSO-d_6): δ 10.33 (s, 1H), 8.16 (d,

J = 2.7 Hz, 1H), 8.01 (dd, *J* = 8.7, 2.4 Hz, 2H), 7.92 (dd, *J* = 4.6, 2.6 Hz, 1H), 7.80 (s, 3H), 7.79 (d, *J* = 6.1 Hz, 1H), 7.47–7.40 (m, 2H), 7.23 (dd, *J* = 4.6, 2.5 Hz, 1H), 6.22–6.07 (m, 2H), 5.61 (dd, *J* = 9.6, 3.0 Hz, 1H), 3.30 (d, *J* = 4.3 Hz, 4H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C22H19N4O2S, 403.1; found, 403.0.

N-(Thietan-3-yl)pyrrolo[2,1-f][1,2,4]triazine-4-carboxamide—Z4324535763

(Method 3).—Yield: 26%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 9.70 (d, *J* = 8.8 Hz, 1H), 8.65 (d, *J* = 2.9 Hz, 1H), 8.24 (s, 1H), 7.44 (t, *J* = 3.8 Hz, 1H), 7.17 (q, *J* = 3.3 Hz, 1H), 3.71–3.62 (m, 2H), 3.30 (s, 1H), 3.19 (t, *J* = 8.1 Hz, 2H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C10H11N4OS, 235.1; found, 235.0.

3-(Oxetan-3-ylidenemethyl)imidazo[1,2-a]pyridine-2-carbaldehyde—

Z5185631875 (Method 5).—Yield: 38%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 10.06 (d, *J* = 1.3 Hz, 1H), 8.40 (dd, *J* = 6.9, 1.4 Hz, 1H), 7.67 (dt, *J* = 9.1, 1.3 Hz, 1H), 7.42 (ddt, *J* = 9.2, 6.7, 1.4 Hz, 1H), 7.08 (td, *J* = 6.9, 1.4 Hz, 1H), 6.71 (t, *J* = 2.5 Hz, 1H), 5.33 (q, *J* = 3.2 Hz, 2H), 5.10 (h, *J* = 2.6, 1.9 Hz, 2H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C12H11N2O2, 215.1; found, 215.0.

6-Bromo-N-(1-(5-(thiophen-2-yl)-1H-1,2,4-triazol-3-yl)ethyl)-indolizine-2-carboxamide—Z5347191013 (Method 3).—

Yield: 53%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 13.88 (s, 1H), 8.78 (s, 1H), 8.64 (s, 1H), 8.02 (s, 1H), 7.67 (s, 1H), 7.53 (s, 1H), 7.44 (d, *J* = 9.5 Hz, 1H), 7.10 (s, 1H), 6.96 (d, *J* = 1.6 Hz, 1H), 6.83–6.77 (m, 1H), 5.29 (s, 1H), 1.57–1.53 (m, 3H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C17H15BrN5OS, 416.0; found, 415.8.

2-(2-Bromo-4-chlorophenoxy)ethyl 5-(Furan-2-yl)-1H-pyrazole-3-carboxylate—Z5347190975 (Method 1).—

Yield: 17%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 14.15–14.03 (s, 1H), 7.79 (s, 1H), 7.67 (t, *J* = 2.3 Hz, 1H), 7.41 (dt, *J* = 8.9, 2.2 Hz, 1H), 7.22 (dd, *J* = 9.0, 1.8 Hz, 1H), 6.98–6.91 (m, 1H), 6.79–6.62 (s, 1H), 6.56 (s, 1H), 4.59 (s, 2H), 4.39 (t, *J* = 4.5 Hz, 2H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C16H13BrClN2O4, 413.0; found, 412.8.

2-((5-Chloropyridin-2-yl)amino)-2-oxoethyl 3-(Furan-3-yl)-1H-pyrazole-5-carboxylate—Z795161408 (Method 1).—

Yield: 19%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 14.06–13.85 (s, 1H), 10.96 (d, *J* = 17.9 Hz, 1H), 8.38 (d, *J* = 2.7 Hz, 1H), 8.19 (s, 1H), 8.14–8.04 (d, *J* = 8.8 Hz, 1H), 7.91 (dd, *J* = 8.8, 2.7 Hz, 1H), 7.78–7.72 (m, 1H), 7.18–7.06 (m, 1H), 6.96–6.87 (m, 1H), 4.98 (s, 1H), 4.94 (s, 1H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C15H12ClN4O4, 347.0; found, 347.0.

N-(5-(Dimethylamino)pyridin-3-yl)-4-formyl-2-hydroxybenza-mide—Z5185631889 (Method 3).—

Yield: 58%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*6): δ 11.44 (s, 1H), 9.95 (s, 1H), 8.18 (s, 1H), 7.96 (d, *J* = 7.9 Hz, 1H), 7.86 (d, *J* = 2.8 Hz, 1H), 7.54 (t, *J* = 2.4 Hz, 1H), 7.34 (d, *J* = 1.6 Hz, 1H), 7.33–7.27 (m, 1H), 2.92 (s, 7H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C15H16N3O3, 286.1; found, 286.2.

2-Methyl-6-(1H-1,2,3-triazol-4-yl)-1H-imidazo[4,5-b]pyridine—Z5348530222 (Method 5).—Yield: 34%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.83 (s, 1H), 8.33 (s, 1H), 2.58 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C9H9N6, 201.1; found, 201.2.

6-(((4-Bromo-1H-pyrrol-2-yl)methyl)amino)pyrazine-2-carboni-trile— Z5348530626 (Method 4).—Yield: 48%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.08 (s, 1H), 8.19 (d, J = 6.6 Hz, 2H), 8.02 (t, J = 5.3 Hz, 1H), 6.79 (dd, J = 2.8, 1.8 Hz, 1H), 6.06–6.00 (m, 1H), 4.33 (d, J = 5.2 Hz, 2H). LC/MS (APSI) m/z: [M + H] calcd for C10H9BrN5, 278.0; found, 278.0.

2-(7H-Pyrrolo[2,3-d]pyrimidin-4-yl)phenolv—Z5348530683 (Method 5).—Yield: 33%; purity, >95% (assessed by LC/MS). ^{1}H NMR (600 MHz, DMSO-d_6): δ 14.43 (s, 1H), 12.56 (s, 1H), 8.83 (s, 1H), 7.74 (d, J = 3.6 Hz, 1H), 7.43–7.37 (m, 1H), 7.12 (d, J = 3.6 Hz, 1H), 7.04–6.96 (m, 2H). LC/MS (APSI) m/z: [M + H] calcd for C12H10N3O, 212.1; found, 212.2.

2-Chloroallyl 1H-pyrrolo[2,3-b]pyridine-5-carboxylate—Z5348530741 (Method 1).—Yield: 50%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.12 (s, 1H), 8.81 (d, J = 2.1 Hz, 1H), 8.57 (d, J = 2.3 Hz, 1H), 7.61 (dd, J = 3.5, 2.5 Hz, 1H), 6.63 (dd, J = 3.5, 1.9 Hz, 1H), 5.55 (d, J = 2.1 Hz, 1H), 4.96 (s, 2H). LC/MS (APSI) m/z: [M + H] calcd for C11H10ClN2O2, 237.0; found, 237.0.

6-(4-Chlorothiophen-3-yl)imidazo[1,2-a]pyrazin-2-amine—Z5348530772 (Method 5).—Yield: 23%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.72 (t, J = 1.2 Hz, 1H), 8.57 (s, 1H), 7.94 (dd, J = 3.7, 1.1 Hz, 1H), 7.72 (dd, J = 3.6, 1.1 Hz, 1H), 7.26 (s, 1H), 5.62 (s, 2H). LC/MS (APSI) m/z: [M + H] calcd for C10H8ClN4S, 251.0; found, 251.0.

(6-Bromoimidazo[1,2-a]pyridin-2-yl)methyl 4-(1H-Pyrazol-1-yl)-benzoate— Z791504634 (Method 1).—Yield: 48%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.89 (d, J = 2.0 Hz, 1H), 8.63 (d, J = 2.6 Hz, 1H), 8.10–8.04 (m, 2H), 8.04–7.97 (m, 3H), 7.81 (d, J = 1.7 Hz, 1H), 7.53 (d, J = 9.6 Hz, 1H), 7.36 (dd, J = 9.5, 2.0 Hz, 1H), 6.59 (t, J = 2.1 Hz, 1H), 5.44 (s, 2H). LC/MS (APSI) m/z: [M + H] calcd for C18H14BrN4O2, 399.0; found, 399.0.

3-Bromo-N-(4-hydroxy-3-nitrobenzyl)-1-methyl-1H-indole-2-carboxamide— Z5347186947 (Method 3).—Yield: 78%; purity, >95% (assessed by LC/ MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.04 (s, 1H), 9.16–9.10 (m, 1H), 7.95 (d, J = 2.3 Hz, 1H), 7.58 (tt, J = 5.7, 2.3 Hz, 2H), 7.48 (d, J = 7.9 Hz, 1H), 7.34 (t, J = 7.5 Hz, 1H), 7.24–7.17 (m, 1H), 7.12 (dd, J = 8.7, 2.1 Hz, 1H), 4.48 (dd, J = 6.2, 2.0 Hz, 2H), 3.78 (d, J = 2.2 Hz, 3H). LC/MS (APSI) m/z: [M + H] calcd for C17H15BrN3O4, 406.0; found, 405.9.

(E)-6-(3-(5-Bromo-2-fluoro-4-hydroxyphenyl)acryloyl)-5-methyl-2,3-dihydro-1H-pyrrolizine-7-carbonitrile—Z5347186953 (Method 2).—Yield: 13%; purity, >95% (assessed by LC/MS). ^{1}H NMR (600 MHz, DMSO-d_6): δ 7.95 (d, J = 8.0 Hz, 1H), 7.50 (d,

J = 15.7 Hz, 1H), 7.44 (d, *J* = 15.7 Hz, 1H), 6.80 (d, *J* = 12.2 Hz, 1H), 4.00 (t, *J* = 7.2 Hz, 2H), 2.95 (t, *J* = 7.5 Hz, 2H), 2.47 (d, *J* = 7.2 Hz, 1H), 2.45 (s, 3H), 1.21 (s, 1H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C18H15BrFN2O2, 389.0; found, 389.0.

3-(1-Methyl-2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-6-yl)-N-phenylbenzamide—Z1723429659 (Method 3).—Yield: 16%;

purity, >95% (assessed by LC/MS). [1]H
NMR (500 MHz, DMSO-*d*₆): δ 11.64 (s, 1H), 10.33 (s, 1H), 8.33 (d, *J* = 2.1 Hz, 1H), 8.24 (d, *J* = 2.8 Hz, 1H), 7.91 (dd, *J* = 7.6, 1.9 Hz, 2H), 7.86 (d, *J* = 2.1 Hz, 1H), 7.79 (d, *J* = 7.9 Hz, 2H), 7.62 (t, *J* = 7.7 Hz, 1H), 7.36 (t, *J* = 7.9 Hz, 2H), 7.10 (t, *J* = 7.4 Hz, 1H), 3.36 (s, 3H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C20H17N4O2, 345.1; found, 345.2.

N-(8-Chloro-2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-2,5-difluoro-4-formylbenzamide—Z5185631903 (Method 3).—Yield: 15%;

purity, >95% (assessed by LC/MS). [1]H
NMR (500 MHz, DMSO-*d*₆): δ 10.59 (s, 1H), 10.17 (d, *J* = 2.1 Hz, 1H), 7.75 (dt, *J* = 9.0, 5.7 Hz, 2H), 7.39 (d, *J* = 2.4 Hz, 1H), 7.19 (d, *J* = 2.5 Hz, 1H), 4.30 (ddd, *J* = 20.5, 6.2, 3.0 Hz, 4H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C16H11ClF2NO4, 354.0; found, 354.0.

(6-Methoxybenzo[d]thiazol-2-yl)methyl 5-(Furan-2-yl)-1H-pyra-zole-3-carboxylate—Z1724303092 (Method 1).—Yield: 18%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*₆): δ 14.28–14.17 (s, 1H), 7.89 (d, *J* = 8.9 Hz, 1H), 7.80 (s, 1H), 7.68 (d, *J* = 2.6 Hz, 1H), 7.12 (dd, *J* = 9.0, 2.6 Hz, 1H), 7.06 (s, 1H), 6.93 (s, 1H), 6.62 (s, 1H), 5.68 (s, 2H), 3.81 (s, 3H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C17H14N3O4S, 356.0; found, 356.1.

N-(4-Cyano-7-hydroxynaphthalen-1-yl)-3-formyl-5,6,7,8-tetra-hydroindolizine-1-carboxamide—Z5185631911 (Method 3).—Yield: 15%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-*d*₆): δ 9.53 (s, 1H), 8.02 (d, *J* = 2.3 Hz, 1H), 7.97 (d, *J* = 8.9 Hz, 1H), 7.60 (d, *J* = 8.1 Hz, 1H), 7.54 (s, 1H), 7.40 (dd, *J* = 9.0, 2.2 Hz, 1H), 6.70–6.63 (m, 2H), 4.38 (t, *J* = 6.0 Hz, 2H), 3.17 (t, *J* = 6.4 Hz, 2H), 2.98 (s, 1H), 2.05–1.97 (m, 2H), 1.90 (dd, *J* = 8.0, 4.0 Hz, 2H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C21H18N3O3, 360.1; found, 360.0.

N-((2S)-3-(4-(6-Chloro-2-oxoindolin-3-yl)piperidin-1-yl)-2-methoxy-3-oxopropyl)acrylamide—Z5472168416 (Method 3).—Yield: 15%;

purity, >95% (assessed by LC/MS). ^{1}H NMR
(500 MHz, DMSO-*d*₆): δ 10.51 (s, 1H), 8.21 (d, *J* = 12.4 Hz, 1H), 7.23 (dd, *J* = 8.1, 4.5 Hz, 1H), 7.00–6.94 (m, 1H), 6.80 (d, *J* = 2.0 Hz, 1H), 6.22 (dd, *J* = 10.7, 5.7 Hz, 1H), 6.17–5.99 (m, 1H), 5.62–5.47 (m, 1H), 4.45 (d, *J* = 16.4 Hz, 1H), 4.38 (d, *J* = 14.2 Hz, 1H), 4.22–4.14 (m, 2H), 4.02 (d, *J* = 16.0 Hz, 1H), 3.45 (s, 1H), 3.22 (d, *J* = 7.0 Hz, 1H), 3.16 (dd, *J* = 13.8, 6.5 Hz, 5H), 2.94 (d, *J* = 13.6 Hz, 1H), 2.24 (d, *J* = 11.9 Hz, 2H), 1.59 (d, *J* = 11.4 Hz, 1H), 1.46 (d, *J* = 15.6 Hz, 1H), 1.38 (d, *J* = 11.7 Hz, 1H), 1.25 (dd, *J* = 27.4, 13.9 Hz, 2H). LC/MS (APSI) *m*/*z*: [M + H] calcd for C20H25ClN3O4, 406.2; found, 406.0.

(5-Methylbenzo[d]thiazol-2-yl)methyl 5-(Furan-2-yl)-1H-pyra-zole-3-carboxylate —Z5347190944 (Method 1).—Yield: 21%; purity, >95% (assessed by LC/MS). ^{1}H NMR (600 MHz, DMSO-d_6): δ 14.18 (s, 1H), 7.98 (d, J = 8.1 Hz, 1H), 7.82 (s, 1H), 7.79 (s, 1H), 7.30 (d, J = 8.2 Hz, 1H), 7.08 (s, 1H), 6.92 (s, 1H), 6.62 (s, 1H), 5.73 (s, 2H), 2.45 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C17H14N3O3S, 340.1; found, 340.0.

N-(3-Fluoro-5-methoxyphenyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-6-yl)benzamide—Z1723442714 (Method 3).—Yield: 11%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.65 (s, 1H), 10.43 (s, 1H), 8.33 (d, J = 2.0 Hz, 1H), 8.21 (t, J = 1.9 Hz, 1H), 7.95–7.87 (m, 2H), 7.84 (d, J = 1.9 Hz, 1H), 7.63 (t, J = 7.7 Hz, 1H), 7.35 (dt, J = 11.3, 2.1 Hz, 1H), 7.26 (d, J = 2.1 Hz, 1H), 6.58 (dt, J = 10.9, 2.4 Hz, 1H), 3.76 (s, 3H), 3.36 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C21H18FN4O3, 393.1; found, 393.0.

N-(4-Chloro-3-(hydroxymethyl)phenyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-6-yl)benzamide—Z5347169163 (Method 3).—Yield: 13%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.65 (s, 1H), 10.45 (s, 1H), 8.33 (d, J = 2.0 Hz, 1H), 8.25 (t, J = 1.8 Hz, 1H), 7.98 (d, J = 2.6 Hz, 1H), 7.95–7.89 (m, 2H), 7.85 (d, J = 2.0 Hz, 1H), 7.80 (dd, J = 8.7, 2.7 Hz, 1H), 7.62 (t, J = 7.7 Hz, 1H), 7.37 (d, J = 8.7 Hz, 1H), 5.47 (t, J = 5.4 Hz, 1H), 4.55 (d, J = 4.5 Hz, 2H), 3.36 (s, 3H). LC/MS (APSI) m/z: [M – H] calcd for C21H18ClN4O3, 409.1; found, 409.2.

3-Cyano-N-(4-hydroxy-3-nitrobenzyl)-1H-indole-2-carboxamide—Z5347186943 (Method 3).—Yield: 18%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.68 (s, 1H), 10.92 (s, 1H), 9.09 (s, 1H), 7.92 (d, J = 2.6 Hz, 1H), 7.68 (d, J = 7.9 Hz, 1H), 7.57 (dd, J = 8.5, 2.5 Hz, 2H), 7.42–7.35 (m, 1H), 7.29 (t, J = 7.5 Hz, 1H), 7.14–7.09 (m, 1H), 4.49 (d, J = 5.6 Hz, 2H). LC/MS (APSI) m/z: [M + H] calcd for C17H11N4O4, 335.1; found, 335.2.

N-(3-(Methoxymethyl)phenyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-6-yl)benzamide—Z1723432553 (Method 3).—Yield: 57%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.64 (s, 1H), 10.32 (s, 1H), 8.33 (d, J = 1.9 Hz, 1H), 8.25 (t, J = 1.9 Hz, 1H), 7.91 (dt, J = 7.9, 1.8 Hz, 2H), 7.84 (d, J = 2.0 Hz, 1H), 7.77 (d, J = 2.1 Hz, 1H), 7.75–7.70 (m, 1H), 7.62 (t, J = 7.7 Hz, 1H), 7.33 (t, J = 7.8 Hz, 1H), 7.04 (d, J = 7.6 Hz, 1H), 4.41 (s, 2H), 3.36 (s, 3H), 3.30 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C22H21N4O3, 389.2; found, 389.2.

5-Formyl-4-methyl-N-(2-methyl-3-(2-oxoimidazolidin-1-yl)-phenyl)thiophene-2-carboxamide—Z1907784975 (Method 3).—Yield: 66%; purity, >95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.19 (s, 1H), 10.09 (s, 1H), 7.89 (s, 1H), 7.25–7.20 (m, 2H), 7.19–7.13 (m, 1H), 6.66 (s, 1H), 3.71 (t, J = 7.8 Hz, 2H), 3.42 (t, J = 7.7 Hz, 2H), 2.58 (s, 3H), 2.08 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C17H18N3O3S, 344.1; found, 344.2.

N-(Benzo[b]thiophen-2-ylmethyl)-3-(thiophen-2-yl)-1H-pyra-zole-5-carboxamide—Z5347191754 (Method 3).—Yield: 44%; purity,

>95% (assessed by LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.69
(s, 1H), 9.27 (s, 0.5H), 8.94 (s, 0.5H), 7.87 (d, J = 7.8 Hz, 1H), 7.77 (d, J = 7.5 Hz, 1H), 7.50
(s, 1H), 7.46 (s, 1H), 7.30 (dt, J = 21.2, 7.0 Hz, 3H), 7.15 (s, 1H), 7.09 (s, 1H), 4.70 (d, J =
20.7 Hz, 2H). LC/MS (APSI) m/z: [M + H] calcd for C17H14N3OS2, 340.1; found, 340.0.

N-((5-Fluoro-3-methylbenzofuran-2-yl)methyl)-3-(thiophen-2-yl)-1H-pyrazole-5-carboxamide—Z5347191764 (Method 3).—Yield: 36%; purity, >95% (assessed by
LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.68 (d, J = 29.0 Hz, 1H), 9.10 (d, J = 7.3
Hz, 0.5H), 8.76 (s, 0.5H), 7.51 (d, J = 4.1 Hz, 1H), 7.45 (d, J = 5.1 Hz, 1H), 7.41–7.32 (m,
1H), 7.32–7.28 (m, 1H), 7.14 (s, 1H), 7.11–7.05 (m, 2H), 4.61–4.52 (m, 2H), 2.22 (d, J =
8.4 Hz, 3H). LC/MS (APSI) m/z: [M + H] calcd for C18H15FN3O2S, 356.1; found, 356.0.

(E)-1-(3-Ethylbenzofuran-2-yl)-3-(4-hydroxy-3-nitrophenyl)prop-2-en-1-one—Z46008250 (Method 2).—Yield: 28%; purity, >95% (assessed by LC/MS).
^{1}H NMR (500 MHz, DMSO-d_6): δ 11.67 (s, 1H), 8.35 (d, J = 2.2 Hz,
1H), 8.02 (dd, J = 8.8, 2.3 Hz, 1H), 7.87 (d, J = 7.9 Hz, 1H), 7.78
(d, J = 15.8 Hz, 1H), 7.71 (d, J = 8.3 Hz, 1H), 7.68 (d, J = 15.8 Hz, 1H), 7.61–7.53
(m, 1H), 7.38 (t, J = 7.5 Hz, 1H), 7.15 (d, J = 8.7 Hz, 1H), 3.14 (q, J = 7.5 Hz, 2H), 1.24 (t,
J = 7.5 Hz, 3H). LC/MS (APSI) m/z: [M + H] calcd for C19H16NO5, 338.1; found, 338.1.

N-(3-Chlorobenzyl)-3-(1-methyl-2-oxo-2,3-dihydro-1H-imidazo-[4,5-b]pyridin-6-yl)benzamide—Z1723430059 (Method 3).—Yield: 29%; purity, >95% (assessed by
LC/MS). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.62 (s, 1H), 9.18 (t, J = 6.0 Hz, 1H), 8.29
(d, J = 2.0 Hz, 1H), 8.19 (t, J = 1.8 Hz, 1H), 7.87 (dd, J = 7.7, 1.8 Hz, 2H), 7.79 (d, J =
2.0 Hz, 1H), 7.57 (t, J = 7.7 Hz, 1H), 7.40–7.32 (m, 2H), 7.30 (dd, J = 7.0, 1.8 Hz, 2H), 4.51
(d, J = 5.9 Hz, 2H), 3.35 (s, 3H). LC/MS (APSI) m/z: [M + H] calcd for C21H18ClN4O2,
393.1; found, 393.0.

**N-(3-(4-(6-Chloro-2-oxoindolin-3-yl)piperidin-1-yl)-3-oxopropyl)-propionamide
(ZD160–68).—**To a solution of compound **I** (5.6 mg, 0.015 mmol) in THF (2 mL) was
added Pd/C (10%, 5 mg). The mixture was stirred under a H$_2$ atmosphere for 2 h followed
by filtering. The filtrate was collected and purified by prep-HPLC to yield ZD160–68, a
white solid (4 mg, 70% yield). ^{1}HNMR (400 MHz, methanol-d_4): δ 7.32–7.21 (m, 1H),
7.03 (dt, J = 8.0, 2.2 Hz, 1H), 6.91 (d, J = 2.2 Hz, 1H), 4.61 (t, J = 15.8 Hz, 1H), 4.00 (t,
J = 13.9 Hz, 1H), 3.50 (q, J = 2.6 Hz, 1H), 3.42 (q, J = 6.3, 5.8 Hz, 2H), 3.11–3.00 (m,
1H), 2.63–2.52 (m, 3H), 2.36 (d, J = 12.7 Hz, 1H), 2.24–2.10 (m, 2H), 1.73–1.59 (m, 2H),
1.49 (dq, J = 17.2, 5.7, 4.3 Hz, 2H), 1.14–1.08 (m, 3H). MS (ESI) m/z: [M + H]$^+$ calcd for
C$_{19}$H$_{24}$ClN$_3$O$_3$, 378.8; found, 378.3.

Nsp14 Expression and Purification.—For activity assays, Nsp14 was expressed and
purified as previously described.[39] The codon-optimized gene was also sub-cloned in a
modified pET28b with 6×-histidine and SUMO tag at the N-terminus. Nsp14 was expressed
in *E. coli* Rosetta2(DE3) PlysS cells, growing in terrific broth at 37 °C, induced at 18 °C
with 0.4 mM IPTG at OD 600 nm of 1.2 for 18 h. Cell pellets were recovered and stored at
−80 °C.

For purification, cells were suspended in lysis buffer containing 50 mM HEPES, 500 mM NaCl, 10 mM imidazole, 10% v/v glycerol, 5 mM MgCl$_2$, and 1 mM TCEP pH 8.1 supplemented with EDTA-free protease cocktail inhibitor tablets (Thermo Scientific). Cells were disrupted by sonication and lysate centrifuged at 16,000 rpm for 30 min. Nsp14 was purified using a 5 mL HisTrap HP column, washed with 20 column volumes of lysis buffer with additional 20 and 30 mM imidazole, and eluted with buffer containing 500 mM imidazole. Protein fractions were exchanged to 50 mM HEPES, 150 mM NaCl, 5 mM MgCl$_2$, 10% glycerol, and 1 mM TCEP pH 8.0 and incubated overnight at 4 °C with SenP1 protease at a 1:100 mass ratio. SUMO tag was removed using a MonoQ 10/100 column and pre-equilibrated in 50 mM HEPES, 20 mM NaCl, 5 mM MgCl$_2$, and 1 mM TCEP pH 8.0. Nsp14 was in the unbound fraction. As a final step, the protein was purified using a size exclusion column s200 16/600 in the same buffer for SenP1 digestion. Purest fractions were pulled together, flash-frozen, and stored at −80 °C until needed.

Enzyme Inhibition.—The inhibitory effect of compounds on the methyltransferase activity of SARS-CoV-2 Nsp14 was assessed using a previously developed radiometric assay.[39]

Jump Dilution.—The recovery of Nsp14 activity after incubation with each inhibitor and rapid dilution was monitored. Nsp14 at 100-fold higher concentration than what was required for activity measurement (1.5 nM) was incubated with each compound at 10-fold the IC$_{50}$ value concentration for 1 h at room temperature. Reaction mixtures were then rapidly diluted 100-fold into the assay buffer containing substrate RNA and SAM, and recovery of the Nsp14 activity was monitored.

Mechanism of Action.—IC$_{50}$ values were determined at a fixed concentration of RNA substrate (0.25 μM; 5× K_m) and varying concentrations of SAM (up to 2.5 μM/10× K_m) and varying concentrations of RNA (up to 0.5 μM; 10× K_m) at fixed 1.25 μM (5× K_m) of SAM. Linear increase in IC$_{50}$ values as the concentration of substrate is increased indicated a competitive pattern of inhibition as described by.[87]

Assessment of Covalent Binding by LC–MS.—To form the protein–ligand ('**1911**) complex, Nsp14 was incubated with 20 molar excess of compound for 2 h at room temperature (20 °C) before adding MS running buffer (0.1% FA). The resulting samples were separated on a HPLC column with 5–95% acetonitrile in water as eluent. The MS data were analyzed using an Agilent LC/MSD time-of-flight mass spectrometer equipped with an electrospray ionization source.

For compounds acryl42 and acryl42–10, an aliquot of pure Nsp14 enzyme was thawed on ice, centrifuged for 10 min at 14,000 rpm, diafiltrated, and concentrated to 40 μM in 50 mM HEPES pH 8.0, 20 mM NaCl, 5 mM MgCl$_2$, 1 mM TCEP. Then, 1 μM Nsp14 was incubated alone or in the presence of 500 μM acryl42 or 50 μM acryl42–10 at room temperature in 100 μL aliquots. Time points were taken at 0, 2, 4, 8, and 18 h. For each time point, 1 μL of the mix was injected into a Xevo G2-XS QTof Quadrupole time of flight mass spectrometer (Waters) using a solution of 0.05% formic acid at room temperature. Collected spectra from

700 to 1400 m/z were used to determine protein mass using MaxEnt with a 58,500 to 62,000 Da range and 1 Da/channel resolution.

Compound Selectivity.—Selectivity assays were performed as previously described.[35] Compounds were tested at 10 μM in triplicate using radiometric assays. Enzymes with >50% inhibition were prioritized for concentration–response curves for IC$_{50}$ determination.

Aggregation.

Dynamic Light Scattering.—Samples were prepared as 8-point half-log dilutions in filtered 50 mM KPi buffer, pH 7.0 with final DMSO concentration at 1% (v/v). Colloidal particle formation was detected using DynaPro Plate Reader II (Wyatt Technologies). All compounds were screened in triplicate at each concentration. For compounds that formed colloidal-like particles, the critical aggregation concentration (CAC) was determined by splitting the data into two data sets based on aggregating (i.e. >106 scattering intensity) and non-aggregating (i.e. <106 scattering intensity) and were fitted with separate nonlinear regression curves, and the point of intersection was determined using GraphPad Prism software version 9.1.1 (San Diego, CA).

Enzyme Inhibition Assays.—Enzyme inhibition assays were performed at room temperature using a CLARIOstar Plate Reader (BMG Labtech). Samples were prepared in 50 mM KPi buffer, pH 7.0 with a final DMSO concentration at 1% (v/v). Compounds were incubated with 2 nM AmpC β-lactamase (AmpC) or malate dehydrogenase (MDH) for 5 min. AmpC reactions were initiated by the addition of the 50 μM CENTA chromogenic substrate or 50 μM nitrocefin. The change in absorbance was monitored at 405 nm for CENTA (219475, Calbiochem) or 490 for nitrocefin (484400, Sigma-Aldrich) for 1 min 30 s. MDH reactions were initiated by the addition of 200 μM nicotinamide adenine dinucleotide (NADH) (54839, Sigma-Aldrich) and 200 μM oxaloacetic acid (324427, Sigma-Aldrich). The change in absorbance was monitored at 340 nm for 1 min 30 s. Initial rates were divided by the DMSO control rate to determine % enzyme activity. Each compound was screened at top concentration in triplicate. For compounds that showed greater than 40% inhibition against either enzyme, eight-point half-log concentration–response curves were performed in triplicate. Data was analyzed using GraphPad Prism software version 9.1.1 (San Diego, CA).

For detergent reversibility experiments, compounds that showed greater than 40% inhibition were screened again as 8-point half-log concentration–response curves in the presence of 0.01% (v/v) Triton X-100 in triplicates. Enzymatic reactions were performed/monitored as previously described.

UV–vis Spectroscopy.—Samples were prepared in 50 mM KPi buffer, pH 7 with a final DMSO concentration at 1% (v/v). Samples were loaded into 1.5 mL methacrylate cuvette (14955128, FisherBrand) and measured using Cary UV–vis Multicell Peltier (Agilent) from 200 to 800 nm with a spectral bandwidth of 2 nm. Data were analyzed using GraphPad Prism software version 9.1.1 (San Diego, CA).

Statistical Analyses.—Data were analyzed using Prism 8.0 or 9.1.1 (GraphPad, San Diego, CA). For Nsp14 dose response curves, data were fitted to the four-parameter logistic equation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

ABBREVIATIONS

SAM	S-adenosyl methionine
RNA	ribonucleic acid
IC$_{50}$	half maximal inhibitory concentration
M^{pro}	main protease
SARS	severe acute respiratory syndrome
CoV	coronavirus
HIV	human immunodeficiency virus
HCV	hepatitis C virus
RSV	respiratory syncytial virus
HBV	hepatitis B virus
HCMV	human cytomegalovirus
HSV	herpes simplex virus
HPV	human papillomavirus
MTase	methyltransferase
MDA5	melanoma differentiation-associated protein 5

SAH	*S*-adenosyl-L-homocysteine
MERS	middle east respiratory syndrome
PKMT	protein lysine methyltransferases
PRMT	protein arginine methyltransferases
PDB	protein databank
rmsd	root mean square deviation
ECFP	extended-connectivity fingerprints
cryo-EM	cryogenic electron microscopy
SAR	structure activity relationship
DLS	dynamic light scattering
MDH	malate dehydrogenase
AmpC	ampicillin β-lactamase
CAC	critical aggregation concentration
LE	ligand efficiencies
eIF4E	eukaryotic translation initiation factor 4E
UCSF	University of California San Francisco
SMDC	Small Molecule Discovery Center
MLSMR	Molecular Libraries Small Molecule Repository
BM	Bemis–Murcko
GLP	G9a-like protein
HAC	heavy atom count
RdRp	RNA-dependent RNA polymerase
AUC	area under curve
TEU	Torsion energy units
SMARTS	SMILES arbitrary target specification
HP	hyper polarized
LC–MS	liquid chromatography–mass spectrometry
MHz	megahertz
DMSO	dimethyl sulfoxide

RT	room temperature
DIPEA	*N*,*N*-diisopropylethylamine
DMF	dimethylformamide
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
HPLC	high performance liquid chromatography
NMR	nuclear magnetic resonance
THF	tetrahydrofuran
SUMO	small ubiquitin-like modifier
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
TCEP	tris(2-carboxyethyl)phosphine
KPi	potassium phosphate
MOA	mechanism of action

REFERENCES

(1). Owen DR; Allerton CMN; Anderson AS; Aschenbrenner L; Avery M; Berritt S; Boras B; Cardin RD; Carlo A; Coffman KJ; et al. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. Science 2021, 374, 1586–1593. [PubMed: 34726479]

(2). Sheahan TP; Sims AC; Zhou S; Graham RL; Pruijssers AJ; Agostini ML; Leist SR; Schäfer A; Dinnon KH; Stevens LJ; et al. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. Sci. Transl. Med. 2020, 12, No. eabb5883. [PubMed: 32253226]

(3). Wahl A; Gralinski LE; Johnson CE; Yao W; Kovarova M; Dinnon KH; Liu H; Madden VJ; Krzystek HM; De C; et al. SARS-CoV-2 infection is effectively treated and prevented by EIDD-2801. Nature 2021, 591, 451–457. [PubMed: 33561864]

(4). Siegel D; Hui HC; Doerffler E; Clarke MO; Chun K; Zhang L; Neville S; Carra E; Lew W; Ross B; et al. Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1-f] [triazin-4-amino] Adenine C-Nucleoside (GS-5734) for the Treatment of Ebola and Emerging Viruses. J. Med. Chem. 2017, 60, 1648–1661. [PubMed: 28124907]

(5). Gordon CJ; Tchesnokov EP; Woolner E; Perry JK; Feng JY; Porter DP; Götte M Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. J. Biol. Chem. 2020, 295, 6785–6797. [PubMed: 32284326]

(6). Rehman S; Ashfaq UA; Javed T Antiviral drugs against hepatitis C virus. Genet. Vaccines Ther. 2011, 9, 11. [PubMed: 21699699]

(7). Clercq ED The design of drugs for HIV and HCV. Nat. Rev. Drug Discovery 2007, 6, 1001–1018. [PubMed: 18049474]

(8). De Clercq E The history of antiretrovirals: key discoveries over the past 25 years. Rev. Med. Virol. 2009, 19, 287–299. [PubMed: 19714702]

(9). De Clercq E; Li G Approved Antiviral Drugs over the Past 50 Years. Clin. Microbiol. Rev. 2016, 29, 695–747. [PubMed: 27281742]

(10). Luttens A; Gullberg H; Abdurakhmanov E; Vo DD; Akaberi D; Talibov VO; Nekhotiaeva N; Vangeel L; De Jonghe S; Jochmans D; et al. Ultralarge Virtual Screening Identifies SARS-CoV-2

Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses. J. Am. Chem. Soc. 2022, 144, 2905–2920. [PubMed: 35142215]

(11). Ma C; Sacco MD; Hurst B; Townsend JA; Hu Y; Szeto T; Zhang X; Tarbet B; Marty MT; Chen Y; et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. Cell Res. 2020, 30, 678–692. [PubMed: 32541865]

(12). Zhang CH; Spasov KA; Reilly RA; Hollander K; Stone EA; Ippolito JA; Liosi ME; Deshmukh MG; Tirado-Rives J; Zhang S; et al. Optimization of Triarylpyridinone Inhibitors of the Main Protease of SARS-CoV-2 to Low-Nanomolar Antiviral Potency. ACS Med. Chem. Lett. 2021, 12, 1325–1332. [PubMed: 34408808]

(13). Unoh Y; Uehara S; Nakahara K; Nobori H; Yamatsu Y; Yamamoto S; Maruyama Y; Taoda Y; Kasamatsu K; Suto T; et al. Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. J. Med. Chem. 2022, 65, 6499–6512. [PubMed: 35352927]

(14). Schuller M; Correy GJ; Gahbauer S; Fearon D; Wu T; Díaz RE; Young ID; Carvalho Martins L; Smith DH; Schulze-Gahmen U; et al. Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking. Sci. Adv. 2021, 7, 7.

(15). Gahbauer S; Correy GJ; Schuller M; Ferla MP; Umay Doruk Y; Rachman M; Wu T; Diolaiti M; Wang S; Neitz RJ; et al. Structure-based Inhibitor Optimization for the Nsp3 Macro-domain of SARS-CoV-2. bioRxiv 2022, DOI: 10.1101/2022.06.27.497816.

(16). Shan H; Liu J; Shen J; Dai J; Xu G; Lu K; Han C; Wang Y; Xu X; Tong Y; et al. Development of potent and selective inhibitors targeting the papain-like protease of SARS-CoV-2. Cell Chem. Biol. 2021, 28, 855–865.e9.

(17). Klima M; Khalili Yazdi A; Li F; Chau I; Hajian T; Bolotokova A; Kaniskan H; Han Y; Wang K; Li D; et al. Crystal structure of SARS-CoV-2 nsp10-nsp16 in complex with small molecule inhibitors, SS148 and WZ16. Protein Sci. 2022, 31, No. e4395. [PubMed: 36040262]

(18). Chen Y; Cai H; Pan J; Xiang N; Tien P; Ahola T; Guo D Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 3484–3489. [PubMed: 19208801]

(19). Jin Y; Ouyang M; Yu T; Zhuang J; Wang W; Liu X; Duan F; Guo D; Peng X; Pan JA Genome-Wide Analysis of the Indispensable Role of Non-structural Proteins in the Replication of SARS-CoV-2. Front. Microbiol. 2022, 13, 907422. [PubMed: 35722274]

(20). Eckerle LD; Becker MM; Halpin RA; Li K; Venter E; Lu X; Scherbakova S; Graham RL; Baric RS; Stockwell TB; et al. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog. 2010, 6, No. e1000896. [PubMed: 20463816]

(21). Bouvet M; Imbert I; Subissi L; Gluais L; Canard B; Decroly E RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc. Natl. Acad. Sci. U. S. A. 2012, 109, 9372–9377. [PubMed: 22635272]

(22). Ogando NS; Ferron F; Decroly E; Canard B; Posthuma CC; Snijder EJ The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. Front. Microbiol. 2019, 10, 1813. [PubMed: 31440227]

(23). Ma Y; Wu L; Shaw N; Gao Y; Wang J; Sun Y; Lou Z; Yan L; Zhang R; Rao Z Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc. Natl. Acad. Sci. U. S. A. 2015, 112, 9436–9441. [PubMed: 26159422]

(24). Bouvet M; Debarnot C; Imbert I; Selisko B; Snijder EJ; Canard B; Decroly E In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog. 2010, 6, No. e1000863. [PubMed: 20421945]

(25). Marcotrigiano J; Gingras AC; Sonenberg N; Burley SK Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 1997, 89, 951–961. [PubMed: 9200613]

(26). Decroly E; Ferron F; Lescar J; Canard B Conventional and unconventional mechanisms for capping viral mRNA. Nat. Rev. Microbiol. 2011, 10, 51–65. [PubMed: 22138959]

(27). Russ A; Wittmann S; Tsukamoto Y; Herrmann A; Deutschmann J; Lagisquet J; Ensser A; Kato H; Gramberg T Nsp16 shields SARS-CoV-2 from efficient MDA5 sensing and IFIT1-mediated restriction. EMBO Rep. 2022, 23, No. e55648. [PubMed: 36285486]

(28). Khalili Yazdi A; Li F; Devkota K; Perveen S; Ghiabi P; Hajian T; Bolotokova A; Vedadi M A High-Throughput Radioactivity-Based Assay for Screening SARS-CoV-2 nsp10-nsp16 Complex. SLAS Discovery 2021, 26, 757–765. [PubMed: 33874769]

(29). Li F; Ghiabi P; Hajian T; Klima M; Li ASM; Khalili Yazdi A; Chau I; Loppnau P; Kutera M; Seitova A; et al. SS148 and WZ16 inhibit the activities of nsp10-nsp16 complexes from all seven human pathogenic coronaviruses. Biochim. Biophys. Acta, Gen. Subj. 2023, 1867, 130319. [PubMed: 36764586]

(30). Chang LJ; Chen TH NSP16 2'-O-MTase in Coronavirus Pathogenesis: Possible Prevention and Treatments Strategies. Viruses 2021, 13, 538. [PubMed: 33804957]

(31). Romano M; Ruggiero A; Squeglia F; Maga G; Berisio R A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. Cells 2020, 9, 1267. [PubMed: 32443810]

(32). Talukdar A; Mukherjee A; Bhattacharya D Fascinating Transformation of SAM-Competitive Protein Methyltransferase Inhibitors from Nucleoside Analogues to Non-Nucleoside Analogues. J. Med. Chem. 2022, 65, 1662–1684. [PubMed: 35014841]

(33). Spannhoff A; Sippl W; Jung M Cancer treatment of the future: inhibitors of histone methyltransferases. Int. J. Biochem. Cell Biol. 2009, 41, 4–11. [PubMed: 18773966]

(34). Ham MS; Lee JK; Kim KC S-adenosyl methionine specifically protects the anticancer effect of 5-FU via DNMTs expression in human A549 lung cancer cells. Mol. Clin. Oncol. 2013, 1, 373–378. [PubMed: 24649178]

(35). Scheer S; Ackloo S; Medina TS; Schapira M; Li F; Ward JA; Lewis AM; Northrop JP; Richardson PL; Kaniskan H; et al. A chemical biology toolbox to study protein methyltransferases and epigenetic signaling. Nat. Commun. 2019, 10, 19. [PubMed: 30604761]

(36). Ferreira de Freitas R; Ivanochko D; Schapira M Methyltransferase Inhibitors: Competing with, or Exploiting the Bound Cofactor. Molecules 2019, 24, 4492. [PubMed: 31817960]

(37). Liu C; Shi W; Becker ST; Schatz DG; Liu B; Yang Y Structural basis of mismatch recognition by a SARS-CoV-2 proof-reading enzyme. Science 2021, 373, 1142–1146. [PubMed: 34315827]

(38). Li ASM; Li F; Eram MS; Bolotokova A; Dela Seña CC; Vedadi M. Chemical probes for protein arginine methyltransferases. Methods 2020, 175, 30–43. [PubMed: 31809836]

(39). Devkota K; Schapira M; Perveen S; Khalili Yazdi A; Li F; Chau I; Ghiabi P; Hajian T; Loppnau P; Bolotokova A; et al. Probing the SAM Binding Site of SARS-CoV-2 Nsp14 In Vitro Using SAM Competitive Inhibitors Guides Developing Selective Bisubstrate Inhibitors. SLAS Discovery 2021, 26, 1200–1211. [PubMed: 34192965]

(40). Bobiļeva O; Bobrovs R; Kaņepe I; Patetko L; Kalniņš G; Šišovs M; Bula AL; Griņberga S; Borodušķis MR; Ramata-Stunda A; et al. Potent SARS-CoV-2 mRNA Cap Methyltransferase Inhibitors by Bioisosteric Replacement of Methionine in SAM Cosubstrate. ACS Med. Chem. Lett. 2021, 12, 1102–1107. [PubMed: 34257831]

(41). Jung E; Soto-Acosta R; Xie J; Wilson DJ; Dreis CD; Majima R; Edwards TC; Geraghty RJ; Chen L Bisubstrate Inhibitors of Severe Acute Respiratory Syndrome Coronavirus-2 Nsp14 Methyltransferase. ACS Med. Chem. Lett. 2022, 13, 1477–1484. [PubMed: 36097498]

(42). Meng EC; Gschwend DA; Blaney JM; Kuntz ID Orientational sampling and rigid-body minimization in molecular docking. Proteins 1993, 17, 266–278. [PubMed: 8272425]

(43). Mysinger MM; Shoichet BK Rapid context-dependent ligand desolvation in molecular docking. J. Chem. Inf. Model. 2010, 50, 1561–1573. [PubMed: 20735049]

(44). Wei BQ; Baase WA; Weaver LH; Matthews BW; Shoichet BK A model binding site for testing scoring functions in molecular docking. J. Mol. Biol. 2002, 322, 339–355. [PubMed: 12217695]

(45). Coleman RG; Carchia M; Sterling T; Irwin JJ; Shoichet BK Ligand pose and orientational sampling in molecular docking. PLoS One 2013, 8, No. e75992. [PubMed: 24098414]

(46). Gu S; Smith MS; Yang Y; Irwin JJ; Shoichet BK Ligand Strain Energy in Large Library Docking. J. Chem. Inf. Model. 2021, 61, 4331–4341. [PubMed: 34467754]

(47). Oprea TI Chemical space navigation in lead discovery. Curr. Opin. Chem. Biol. 2002, 6, 384–389. [PubMed: 12023120]

(48). Douangamath A; Fearon D; Gehrtz P; Krojer T; Lukacik P; Owen CD; Resnick E; Strain-Damerell C; Aimon A; Ábrányi-Balogh P; et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nat. Commun. 2020, 11, 5047. [PubMed: 33028810]

(49). Bonday ZQ; Cortez GS; Grogan MJ; Antonysamy S; Weichert K; Bocchinfuso WP; Li F; Kennedy S; Li B; Mader MM; et al. LLY-283, a Potent and Selective Inhibitor of Arginine Methyltransferase 5, PRMT5, with Antitumor Activity. ACS Med. Chem. Lett. 2018, 9, 612–617. [PubMed: 30034588]

(50). Huynh T; Chen Z; Pang S; Geng J; Bandiera T; Bindi S; Vianello P; Roletto F; Thieffine S; Galvani A; et al. Optimization of pyrazole inhibitors of Coactivator Associated Arginine Methyltransferase 1 (CARM1). Bioorg. Med. Chem. Lett. 2009, 19, 2924–2927. [PubMed: 19419866]

(51). Daigle SR; Olhava EJ; Therkelsen CA; Majer CR; Sneeringer CJ; Song J; Johnston LD; Scott MP; Smith JJ; Xiao Y; et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. Cancer Cell 2011, 20, 53–65. [PubMed: 21741596]

(52). Chory EJ; Calarco JP; Hathaway NA; Bell O; Neel DS; Crabtree GR Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome. Mol. Cell 2019, 73, 61–72.e3. [PubMed: 30472189]

(53). Bonham K; Hemmers S; Lim YH; Hill DM; Finn MG; Mowen KA Effects of a novel arginine methyltransferase inhibitor on T-helper cell cytokine production. FEBS J. 2010, 277, 2096–2108. [PubMed: 20345902]

(54). Stein RM; Yang Y; Balius TE; O'Meara MJ; Lyu J; Young J; Tang K; Shoichet BK; Irwin J Property-Unmatched Decoys in Docking Benchmarks. J. Chem. Inf. Model. 2021, 61, 699–714. [PubMed: 33494610]

(55). Mysinger MM; Carchia M; Irwin JJ; Shoichet BK Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J. Med. Chem. 2012, 55, 6582–6594. [PubMed: 22716043]

(56). Fassio AV; Shub L; Ponzoni L; McKinley J; O'Meara MJ; Ferreira RS; Keiser MJ; de Melo Minardi RC Prioritizing Virtual Screening with Interpretable Interaction Fingerprints. J. Chem. Inf. Model. 2022, 62, 4300–4318. [PubMed: 36102784]

(57). Pettersen EF; Goddard TD; Huang CC; Couch GS; Greenblatt DM; Meng EC; Ferrin TE UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 2004, 25, 1605–1612. [PubMed: 15264254]

(58). Tingle B; Tang K; Castanon J; Gutierrez J; Khurelbaatar M; Dandarchuluun C; Moroz Y; Irwin J ZINC-22-A Free Multi-Billion-Scale Database of Tangible Compounds for Ligand Discovery. ChemRxiv 2022, DOI: 10.26434/chemrxiv-2022-82cz.

(59). Irwin JJ; Tang KG; Young J; Dandarchuluun C; Wong BR; Khurelbaatar M; Moroz YS; Mayfield J; Sayle RA ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. J. Chem. Inf. Model. 2020, 60, 6065–6073. [PubMed: 33118813]

(60). McGovern SL; Caselli E; Grigorieff N; Shoichet BK A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. J. Med. Chem. 2002, 45, 1712–1722. [PubMed: 11931626]

(61). Shoichet BK Screening in a spirit haunted world. Drug Discovery Today 2006, 11, 607–615. [PubMed: 16793529]

(62). Irwin JJ; Duan D; Torosyan H; Doak AK; Ziebart KT; Sterling T; Tumanian G; Shoichet BK An Aggregation Advisor for Ligand Discovery. J. Med. Chem. 2015, 58, 7076–7087. [PubMed: 26295373]

(63). McGovern SL; Helfand B; Feng B; Shoichet BK A Specific Mechanism for Non-Specific Inhibition. J. Med. Chem. 2003, 46, 4265–4272. [PubMed: 13678405]

(64). Feng BY; Simeonov A; Jadhav A; Babaoglu K; Inglese J; Shoichet BK; Austin CP A high-throughput screen for aggregation-based inhibition in a large compound library. J. Med. Chem. 2007, 50, 2385–2390. [PubMed: 17447748]

(65). Chen Y; Shoichet BK Molecular docking and ligand specificity in fragment-based inhibitor discovery. Nat. Chem. Biol. 2009, 5, 358–364. [PubMed: 19305397]

(66). Barelier S; Eidam O; Fish I; Hollander J; Figaroa F; Nachane R; Irwin JJ; Shoichet BK; Siegal G Increasing chemical space coverage by combining empirical and computational fragment screens. ACS Chem. Biol. 2014, 9, 1528–1535. [PubMed: 24807704]

(67). Schuffenhauer A; Ruedisser S; Marzinzik AL; Jahnke W; Selzer P; Jacoby E; Jacoby E Library design for fragment based screening. Curr. Top. Med. Chem. 2005, 5, 751–762. [PubMed: 16101415]

(68). Teotico DG; Babaoglu K; Rocklin GJ; Ferreira RS; Giannetti AM; Shoichet BK Docking for fragment inhibitors of AmpC beta-lactamase. Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 7455–7460. [PubMed: 19416920]

(69). Kuntz ID; Blaney JM; Oatley SJ; Langridge R; Ferrin TE A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 1982, 161, 269–288. [PubMed: 7154081]

(70). Zautra AJ; Maxwell BM; Reich JW Relationship among physical impairment, distress, and well-being in older adults. J. Behav. Med. 1989, 12, 543–557. [PubMed: 2534485]

(71). Wan X; Yang T; Cuesta A; Pang X; Balius TE; Irwin JJ; Shoichet BK; Taunton J Discovery of Lysine-Targeted eIF4E Inhibitors through Covalent Docking. J. Am. Chem. Soc. 2020, 142, 4960–4964. [PubMed: 32105459]

(72). Fink EA; Bardine C; Gahbauer S; Singh I; White K; Gu S; Wan X; Ary B; Glenn I; O'Connell J; et al. Large Library Docking for Novel SARS-CoV-2 Main Protease Non-covalent and Covalent Inhibitors. bioRxiv 2022, DOI: 10.1101/2022.07.05.498881.

(73). London N; Miller RM; Krishnan S; Uchida K; Irwin JJ; Eidam O; Gibold L; Cimerman i P; Bonnet R; Shoichet BK; et al. Covalent docking of large libraries for the discovery of chemical probes. Nat. Chem. Biol. 2014, 10, 1066–1072. [PubMed: 25344815]

(74). Arkin MR; Ang KK; Chen S; Davies J; Merron C; Tang Y; Wilson CG; Renslo AR UCSF Small Molecule Discovery Center: innovation, collaboration and chemical biology in the Bay Area. Comb. Chem. High Throughput Screening 2014, 17, 333–342.

(75). PubChem Molecular Libraries Small Molecule Repository. http://pubchem.ncbi.nlm.nih.gov.

(76). Boike L; Henning NJ; Nomura DK Advances in covalent drug discovery. Nat. Rev. Drug Discovery 2022, 21, 881–898. [PubMed: 36008483]

(77). Ahmed-Belkacem R; Hausdorff M; Delpal A; Sutto-Ortiz P; Colmant AMG; Touret F; Ogando NS; Snijder EJ; Canard B; Coutard B; et al. Potent Inhibition of SARS-CoV-2 nsp14 N7-Methyltransferase by Sulfonamide-Based Bisubstrate Analogues. J. Med. Chem. 2022, 65, 6231–6249. [PubMed: 35439007]

(78). Otava T; ála M; Li F; Fanfrlík J; Devkota K; Perveen S; Chau I; Pakarian P; Hobza P; Vedadi M; et al. The Structure-Based Design of SARS-CoV-2 nsp14 Methyltransferase Ligands Yields Nanomolar Inhibitors. ACS Infect. Dis. 2021, 7, 2214–2220. [PubMed: 34152728]

(79). Imprachim N; Yosaatmadja Y; Newman JA Crystal Structures and Fragment Screening of SARS-CoV-2 NSP14 Reveal Details of Exoribonuclease Activation and mRNA Capping and Provide Starting Points for Antiviral Drug Development. bioRxiv 2022, DOI: 10.1101/2022.03.11.483836.

(80). Owen DR; Allerton CMN; Anderson AS; Aschenbrenner L; Avery M; Berritt S; Boras B; Cardin RD; Carlo A; Coffman KJ; et al. An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. Science 2021, 374, 1586–1593. [PubMed: 34726479]

(81). Word JM; Lovell SC; Richardson JS; Richardson DC Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. 1999, 285, 1735–1747. [PubMed: 9917408]

(82). Pearlman DA; Case DA; Caldwell JW; Ross WS; Cheatham TE III; DeBolt S; Ferguson D; Seibel G; Kollman P AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput. Phys. Commun. 1995, 91, 1–41.

(83). Gallagher K; Sharp K Electrostatic contributions to heat capacity changes of DNA-ligand binding. Biophys. J. 1998, 75, 769–776. [PubMed: 9675178]

(84). Meng EC; Shoichet BK; Kuntz ID Automated docking with grid-based energy evaluation. J. Comput. Chem. 1992, 13, 505–524.

(85). Bender BJ; Gahbauer S; Luttens A; Lyu J; Webb CM; Stein RM; Fink EA; Balius TE; Carlsson J; Irwin JJ; et al. A practical guide to large-scale docking. Nat. Protoc. 2021, 16, 4799–4832. [PubMed: 34561691]

(86). Bemis GW; Murcko MA The properties of known drugs. 1. Molecular frameworks. J. Med. Chem. 1996, 39, 2887–2893. [PubMed: 8709122]

(87). Copeland RA Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis; Wiley-VCH, Inc., 2000.
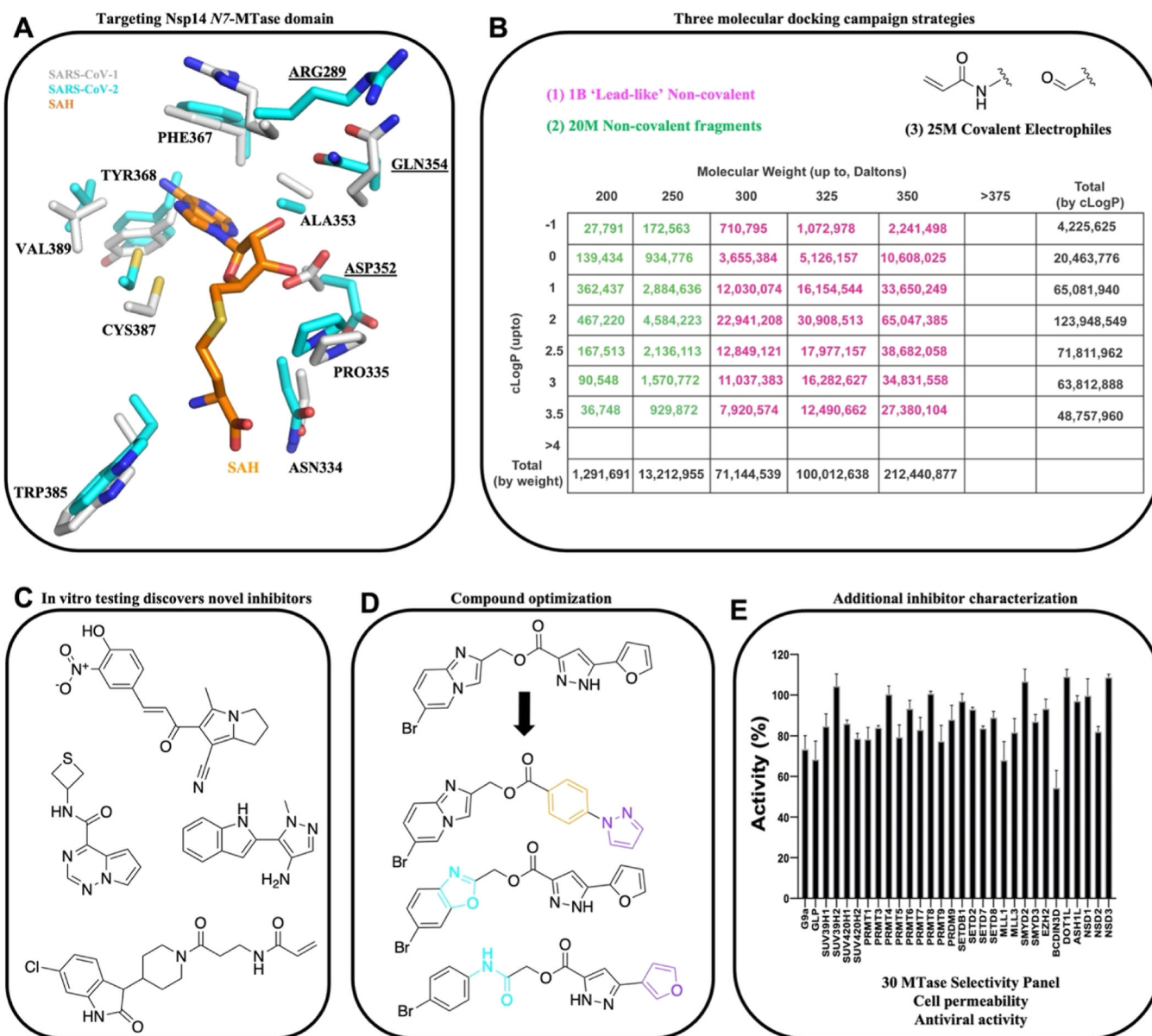
Figure 1.
Workflow for inhibitor discovery against the N7-MTase domain of Nsp14 using molecular docking. (A) SARS-CoV-1 (PDB ID: 5C8S) and −2 Nsp14 (PDB ID: 7N0B) MTase domains targeted with (B) three molecule subsets in molecular docking: lead-like non-covalent, fragment non-covalent, and acrylamide and aldehyde covalent electrophiles. (C) Diverse inhibitors discovered from each docking strategy, followed by (D) compound optimization to improve potencies. (E) Best inhibitors evaluated for additional properties including MTase selectivity and antiviral efficacy.
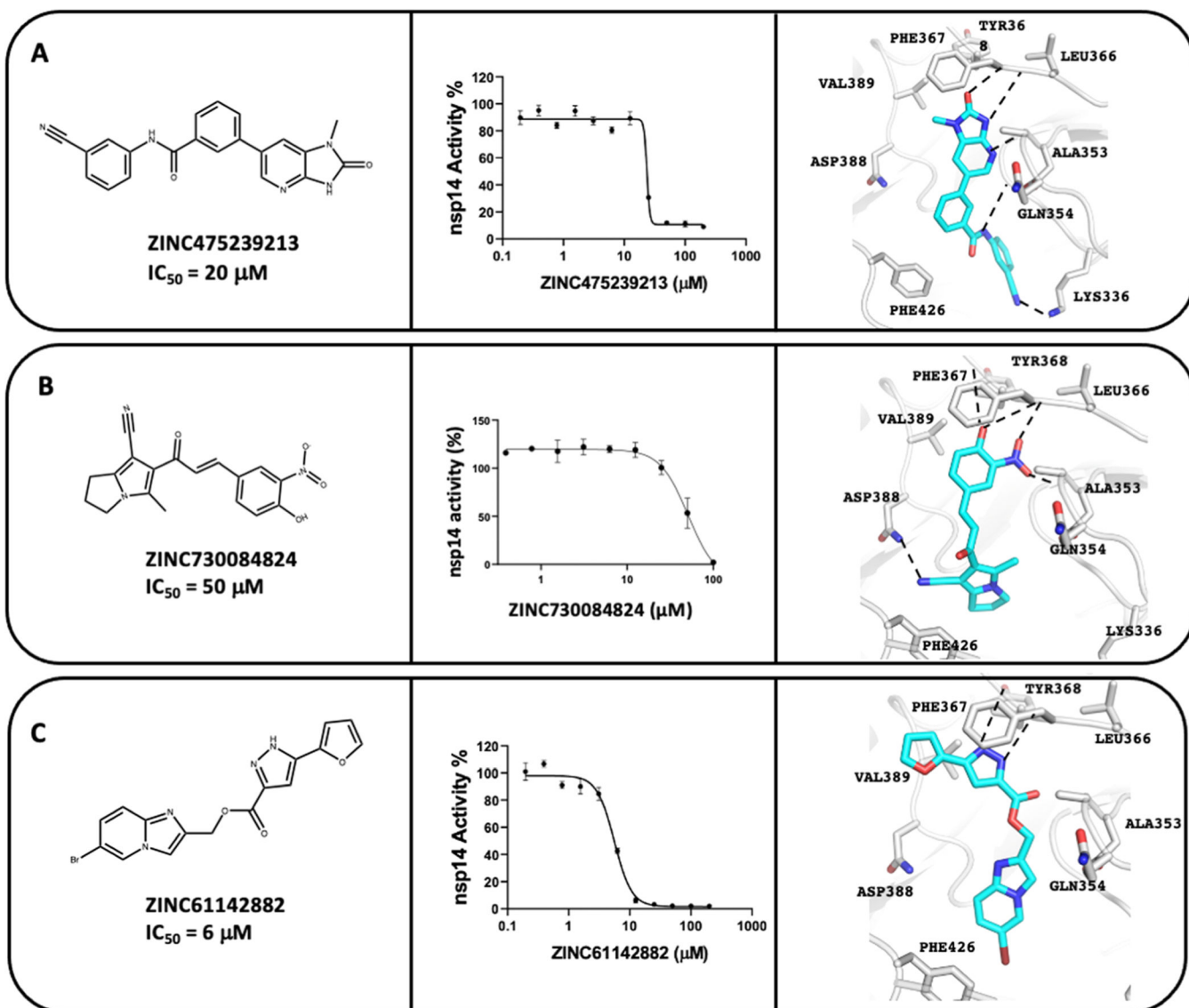
Figure 2.
Ultra-large scale docking identifies three **Nsp14** inhibitors with novel chemical scaffolds.
2D chemical structures, concentration-dependent Nsp14 MTase inhibition, and docked poses
are represented for compounds ZINC475239213, ZINC730084824, and ZINC61142882 in
panels (A–C), respectively. SARS-CoV-2 Nsp14 (PDB ID: 7N0B) and inhibitors are shown
in gray and cyan carbons, respectively, and hydrogen bonds are shown as black dashed lines.
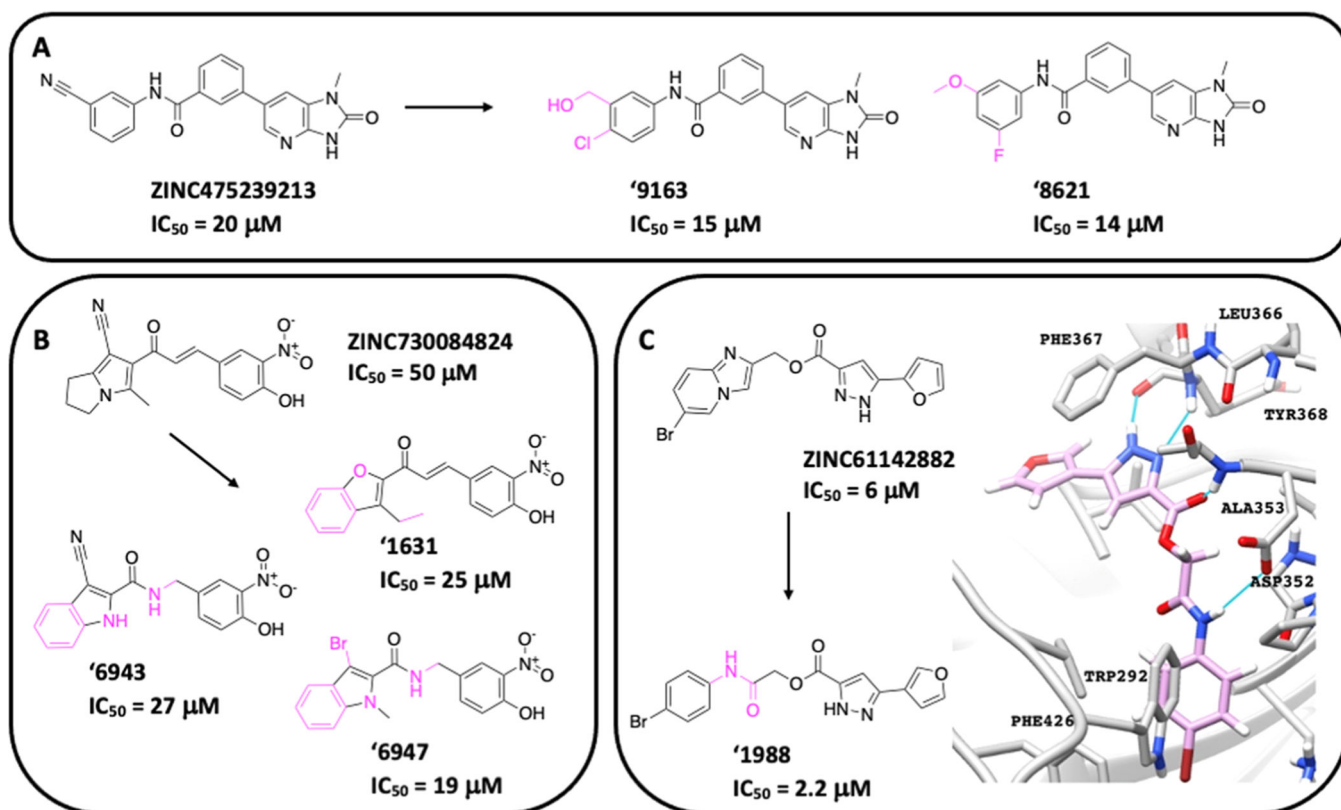The experiments were performed in triplicate.

Figure 3.
Hit optimization of the non-covalent compounds '**9213**, '**4824**, and '**2882**. 2D chemical structures of the parent hit and corresponding analogs with chemical changes represented in pink. (A) '**9213** analogs with the nitrile removed have similar IC_{50} values indicating the non-covalent mechanism of action. (B) '**4824** analogs with the nitrile or vinyl group removed have similar or more potent IC_{50} values. (C) '**2882** analog '**1988** is just as potent with opening of the bicyclic group. The '**1988** docked pose (magenta carbons) is shown in SARS-CoV-2 (PDB ID: 7N0B) Nsp14 (gray carbons). The experiments were performed in triplicate.
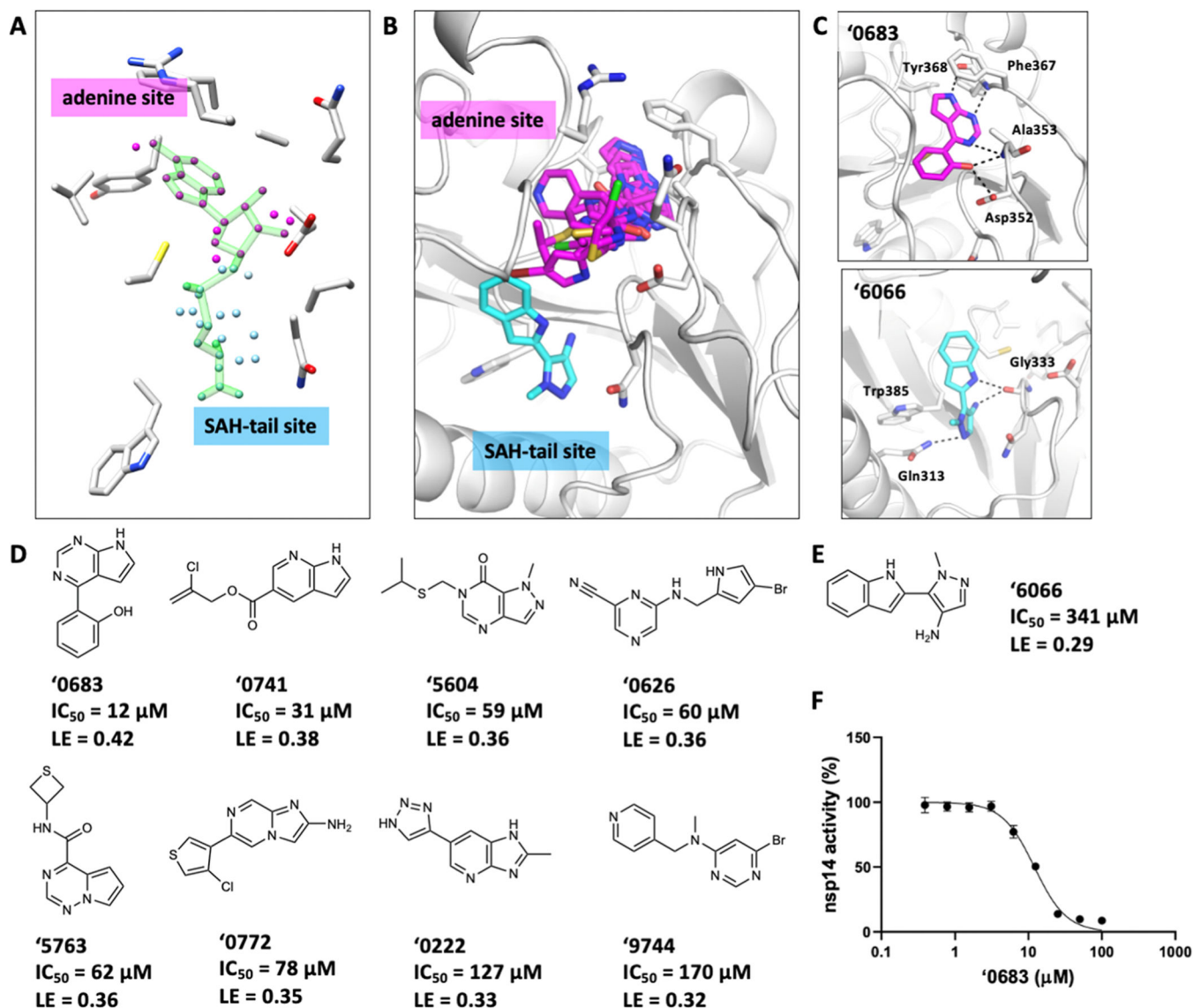
Figure 4.
Fragment inhibitors from the 16M docking screen. (A) Three sets of pseuodo atoms, which define where ligands, are sampled in the binding site ("spheres") used in the docking screens[69] including the adenine-site spheres (pink), SAM-tail site spheres (blue), and a superset of both. SARS-CoV-2 (PDB ID: 7N0B) Nsp14 (gray carbons) with SAH (green carbons). (B) Overlay of all fragment docking hits in the SAM binding site of SARS-CoV-2 (tan carbons). Docked poses for adenine-site inhibitors shown (pink carbons) and SAM-tail site inhibitors (cyan carbons). (C) Docked poses of the best adenine-site fragment '**0683** and the SAM-tail site fragment '**6066**. (D) Eight adenine-site fragment hits shown with their respective IC_{50} values. (E) SAM-tail site fragment hit '**6066**. (F) Concentration response curve of '**0683** in the N7-MTase inhibitory activity assay. For (D,E), IC_{50} values derived from concentration–response curves are shown in Figure S9. The experiments were performed in triplicate.
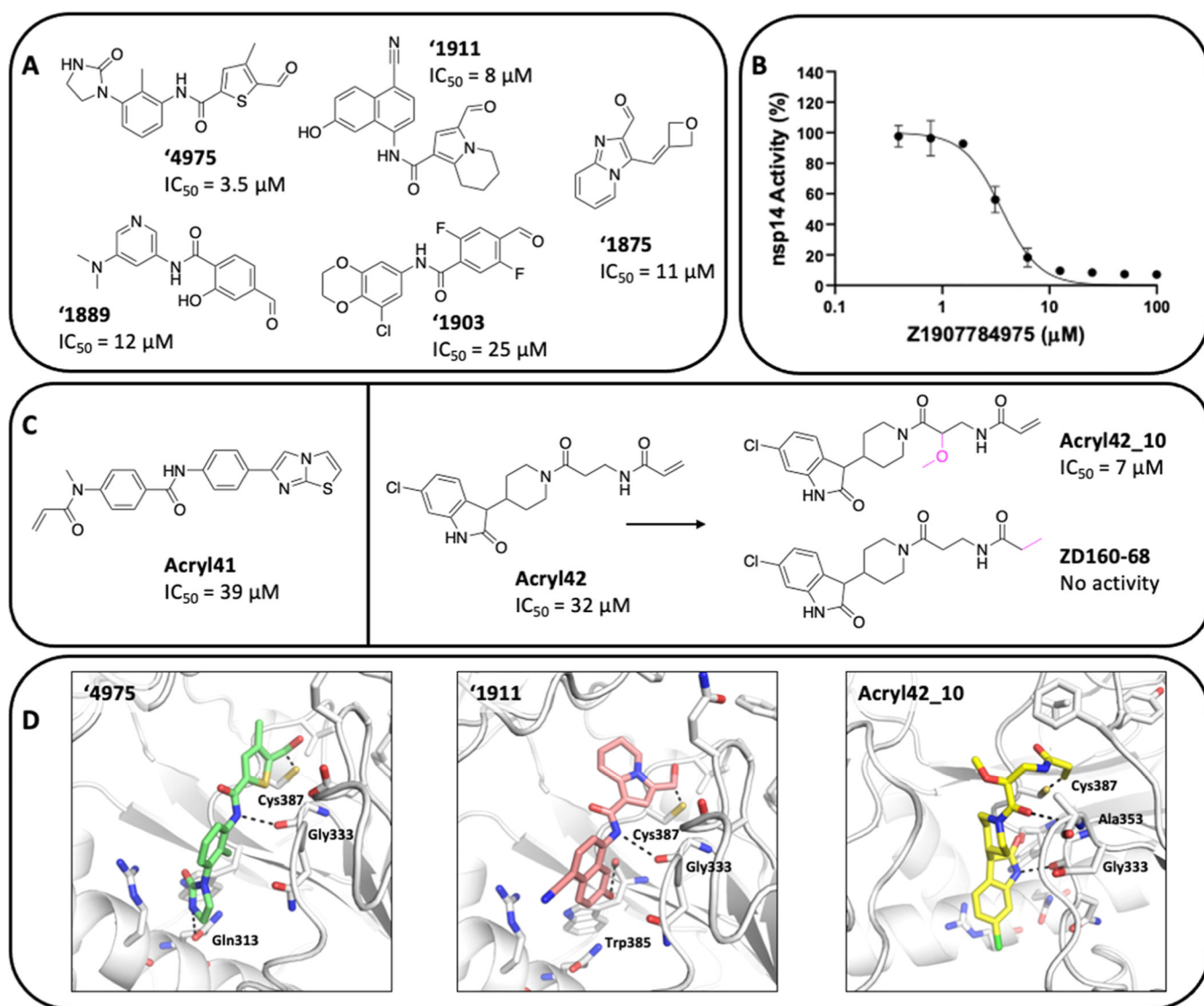
Figure 5.
Docking 25 million electrophiles reveals aldehyde and acrylamide inhibitors. (A) Aldehyde docking hits. (B) Concentration–response curve for the most potent aldehyde '**4975** in the N7-MTase inhibitory activity assay. (C) Acrylamide docking hits **acryl41** and **acryl42**, with analog **acryl42_10** and inactive analog **ZD160–68**. (D) Docked poses of '**4975** and '**1911**, and modeled pose of analog **acryl42_10** (PDB ID: 7N0B). The experiments were performed in triplicate and additional concentration–response curves found in Figure S11.
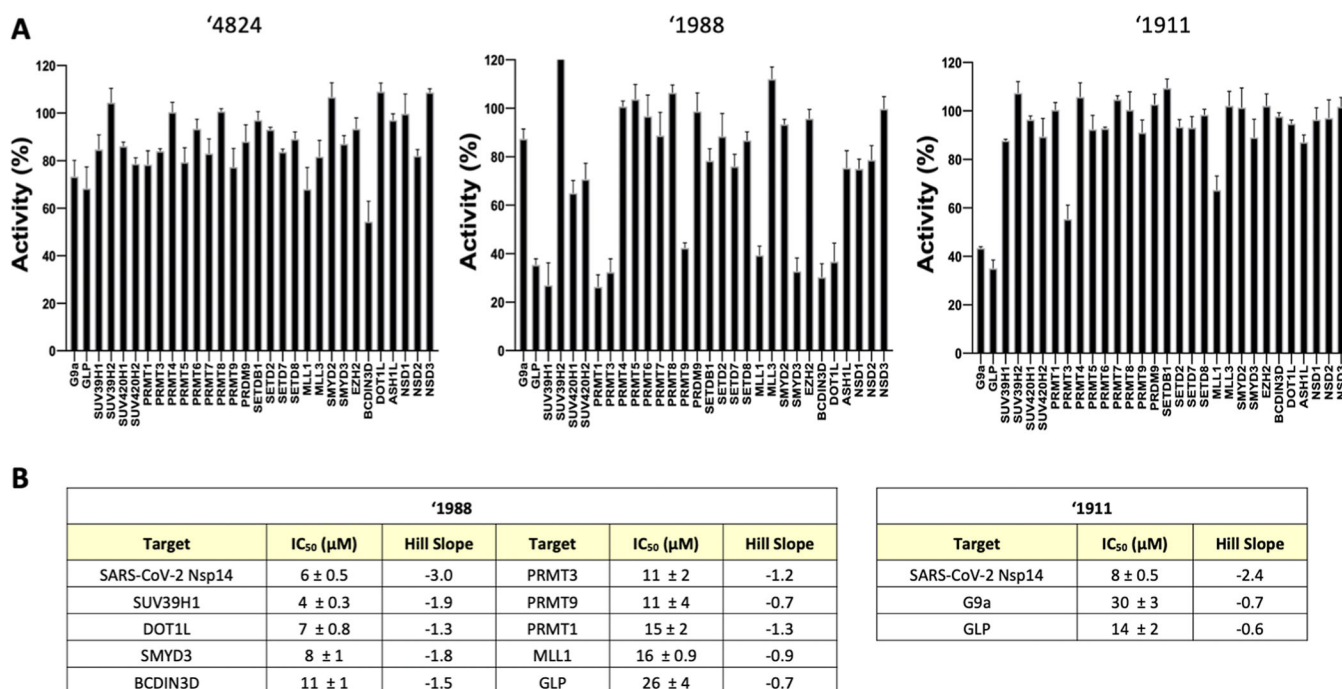
Figure 6.
MTase selectivity of docking-derived inhibitors. (A) Compounds were tested against a panel of 30 SAM-dependent human protein and RNA MTases. Those with >50% inhibition were prioritized for (B) IC$_{50}$ determination. The experiments were performed in triplicate.

Table 1.

Expanded DOCKovalent Electrophile Databases

database		aldehydes		acrylamides	
name	size	# molecules	# BM scaffolds	# molecules	# BM scaffolds
Covalent2022.docking.org	31,000,000,000	6,197,526	848,830	17,680,357	1,404,874
ZINC20 in-stock	7,517,254	31,554	3373	21,798	5612
UCSF SMDC	690,125	615	235	38	24
MLSMR	406,098	908	407	63	27