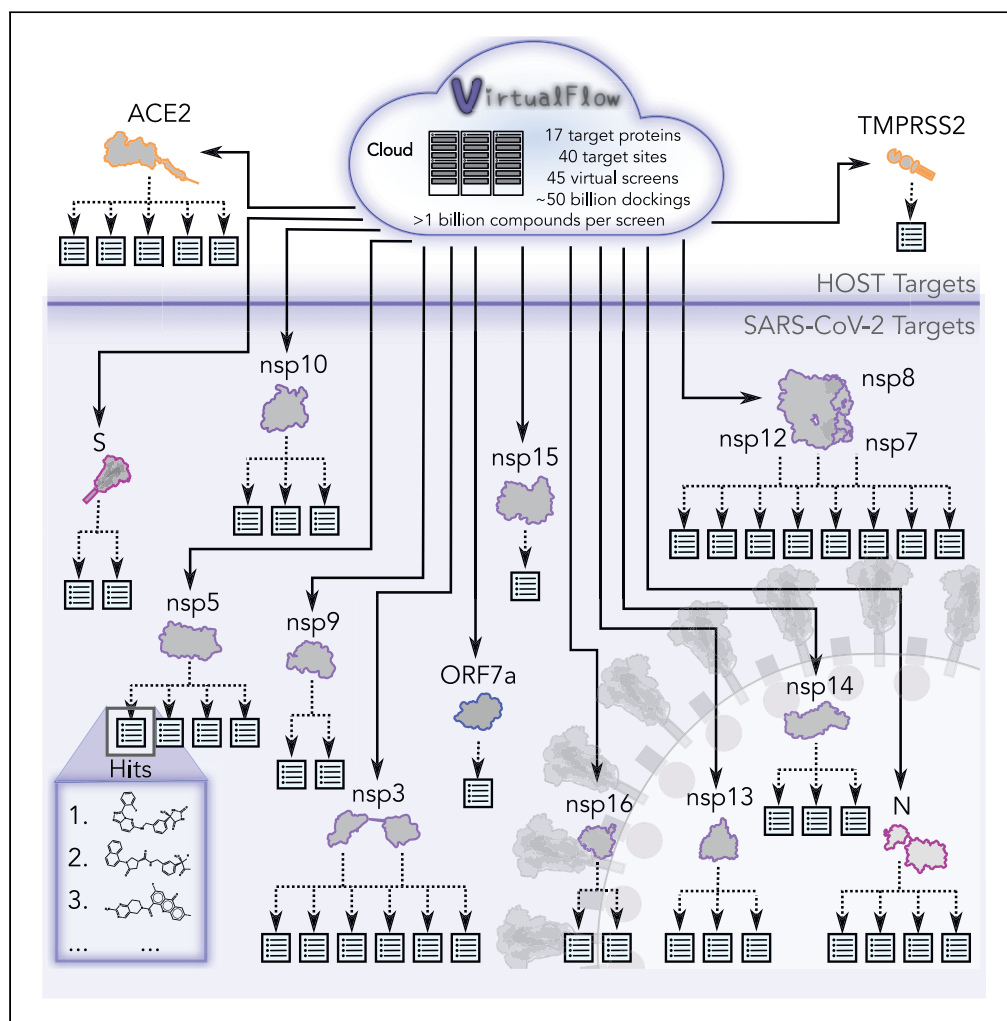


Article

A multi-pronged approach targeting SARS-CoV-2 proteins using ultra-large virtual screening



Christoph Gorgulla, Krishna M. Padmanabha Das, Kendra E. Leigh, ..., Gerhard Wagner, Jamie Kinney, Haribabu Arthanari

cgorgulla@g.harvard.edu (C.G.)
hari@hms.harvard.edu (H.A.)

HIGHLIGHTS

SARS-CoV-2 related proteins were targeted in ultra-large in silico screens.

Multiple functional sites on individual target proteins were screened.

17 virus-related targets, 45 screens, and ~50 billion docking instances were covered.

Conservation in some target sites means hits could exhibit pan-coronavirus function.

Screening results are available as an interactive web resource and for download.

Gorgulla et al., iScience 24, 102021
February 19, 2021 © 2020 The Author(s).
<https://doi.org/10.1016/j.isci.2020.102021>

Article

A multi-pronged approach targeting SARS-CoV-2 proteins using ultra-large virtual screening

Christoph Gorgulla,^{1,2,3,*} Krishna M. Padmanabha Das,^{1,3} Kendra E. Leigh,^{11,20,21} Marco Cespugli,¹³ Patrick D. Fischer,^{1,3,9} Zi-Fu Wang,¹ Guilhem Tesseyre,⁴ Shreya Pandita,⁴ Alec Shnapir,⁴ Anthony Calderaio,¹⁷ Minko Gechev,⁴ Alexander Rose,¹⁶ Noam Lewis,⁴ Colin Hutcheson,⁴ Erez Yaffe,⁴ Roni Luxenburg,⁴ Henry D. Herce,^{1,3} Vedat Durmaz,¹³ Thanos D. Halazonetis,¹² Konstantin Fackeldey,^{8,10} J.J. Patten,¹⁵ Alexander Chuprina,⁵ Igor Dziuba,¹⁹ Alla Plekhova,⁶ Yurii Moroz,^{6,7} Dmytro Radchenko,^{5,7} Olga Tarkhanova,⁶ Irina Yavnyuk,⁵ Christian Gruber,^{13,14} Ryan Yust,⁴ Dave Payne,⁴ Anders M. Näär,¹⁸ Mark N. Namchuk,¹ Robert A. Davey,¹⁵ Gerhard Wagner,¹ Jamie Kinney,⁴ and Haribabu Arthanari^{1,3,22,*}

SUMMARY

The unparalleled global effort to combat the continuing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic over the last year has resulted in promising prophylactic measures. However, a need still exists for cheap, effective therapeutics, and targeting multiple points in the viral life cycle could help tackle the current, as well as future, coronaviruses. Here, we leverage our recently developed, ultra-large-scale *in silico* screening platform, VirtualFlow, to search for inhibitors that target SARS-CoV-2. In this unprecedented structure-based virtual campaign, we screened roughly 1 billion molecules against each of 40 different target sites on 17 different potential viral and host targets. In addition to targeting the active sites of viral enzymes, we also targeted critical auxiliary sites such as functionally important protein-protein interactions.

INTRODUCTION

At the end of 2019, cases of pneumonia with an initially unknown etiology were identified in Wuhan city, in the Hubei Province of China (Zhou et al., 2020a; Wu et al., 2020; Andersen et al., 2020). The cause was determined to be a novel coronavirus (CoV) (Zhu et al., 2020), and by early July 2020, there were over 12 million confirmed cases worldwide (Dong et al., 2020a; Dong et al., 2020b) of what is now designated severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020). A *Betacoronavirus* (lineage B) related to severe acute respiratory syndrome CoV (SARS-CoV), SARS-CoV-2 commonly causes fever, cough, myalgia, and/or fatigue (Huang et al., 2020; Roman et al., 2020). While even a year later, our clinical knowledge is still developing, in addition to asymptomatic and mild cases, dyspnea, lymphopenia, and anosmia, with or without dysgeusia, have also been reported as clinical features (Huang et al., 2020; Xydakis et al., 2020; Rothe et al., 2020; Zhou et al., 2020b), and complications can include acute respiratory distress syndrome, acute cardiac injury, and secondary infections (Huang et al., 2020). As of July 8th, 2020, over half a million deaths have been attributed to coronavirus disease 2019 (COVID-19) (Dong et al., 2020a; Dong et al., 2020b), and the rapid expansion in case number in combination with severe symptoms requiring hospitalization has resulted in an unprecedented strain on the global healthcare system.

Coronaviridae comprises a family of large positive-sense, single-stranded RNA viruses that derive their name from the “corona” that fringes the virions in electron micrographs (Almedia et al., 1968; Masters, 2006). CoV virions are composed of a lipid envelope decorated with spike (S) protein, which facilitates entry and causes their corona-like appearance (Neuman and Buchmeier, 2016). Envelope (E) protein, which contributes to virion assembly and viral pathogenesis, and membrane (M) protein, which also facilitates virion assembly, are also both embedded in this bilayer (Neuman and Buchmeier, 2016), and the viral genome, in close association with nucleoprotein (N), is encapsulated within.

To initiate entry, the receptor-binding domain of S must engage with its receptor on the surface of its target cell, and several studies have already identified the SARS-CoV receptor, angiotensin-converting enzyme 2

¹Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Harvard University, Boston, MA 02115, USA

²Department of Physics, Faculty of Arts and Sciences, Harvard University, Cambridge, MA 02138, USA

³Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

⁴Google, Mountain View, CA 94043, USA

⁵Enamine, Kyiv 02094, Ukraine

⁶Chemspace, Kyiv 02094, Ukraine

⁷Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine

⁸Zuse Institute Berlin (ZIB), Berlin 14195, Germany

⁹Department of Pharmacy, Pharmaceutical and Medicinal Chemistry, Saarland University, Saarbrücken, Saarland 66123, Germany

¹⁰Institute of Mathematics, Technical University Berlin, Berlin 10587, Germany

¹¹Max Planck Institute of Biophysics, Frankfurt am Main, Hessen 60438, Germany

¹²Department of Molecular Biology, University of Geneva, Geneva 1205, Switzerland

¹³Innophore GmbH, Graz 8010, Austria

¹⁴Institute of Molecular Biosciences, University of Graz, Graz 8010, Austria

Continued



(ACE2), as a receptor for SARS-CoV-2 (Zhou et al., 2020a; Yan et al., 2020a; Walls et al., 2020). While engagement with the receptor initiates conformational rearrangements in S, the spike protein must also be cleaved at its S2' site as part of the entry process. Unlike the S1/S2 cleavage event, which can occur at any point from viral assembly to entry, S2' cleavage likely only occurs during entry and involves host proteases at the cell surface, such as transmembrane protease, serine 2 (TMPRSS2), or in endosomes, such as cathepsins (Millet and Whittaker, 2015). Further conformational rearrangements in S result in membrane fusion, allowing the release of the nucleocapsid into the cytoplasm.

As the genome is positive-sense, replication starts with the expression of ORF1a and ORF1ab. The resulting polyproteins (pp1a and pp1ab) are further processed into sixteen non-structural proteins (nsp1-16; Figure 1) that form, in conjunction with host proteins, membrane-associated replication and transcription complexes (Snijder et al., 2016). The genome is replicated via an intermediate negative-sense copy of the genome, and both structural and accessory proteins are expressed from 3'-co-terminal sub-genomic RNAs (de Wilde et al., 2017). Assembly occurs on membranes between the endoplasmic reticulum (ER) and the trans-Golgi network, with the virions budding into vesicular compartments that then fuse with the plasma membrane, releasing their cargo (de Wilde et al., 2017) (Figure 1).

While low-pathogenicity human CoVs such as HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 (Su et al., 2016) continually circulate and generally cause mild disease, the emergence of SARS-CoV (Ksiazek et al., 2003; Kuiken et al., 2003; Peiris et al., 2003; Drosten et al., 2003) in 2002 and Middle East respiratory syndrome CoV (MERS-CoV) (de Groot et al., 2013; Sander van et al., 2012; Zaki et al., 2012) in 2012 as the result of zoonotic jumps into the human population (Cui et al., 2018) demonstrated that CoVs could cause serious morbidity and mortality. Their importance made them the subject of extensive research and attractive therapeutic targets. While research is only just beginning on SARS-CoV-2, a comparison of the SARS-CoV-2 sequence early in the outbreak to previously identified CoV sequences revealed it to be most related to previously isolated bat SARS-CoV-like coronaviruses (Zhou et al., 2020a; Wu et al., 2020; Lu et al., 2020), and comparison to both a bat SARS-CoV-like virus (bat SL-CoVZC45), as well as a previous human SARS-CoV isolate (SARS-CoV Tor2), predicted that SARS-CoV-2 likely has a similar genomic organization and expresses similar proteins (Wu et al., 2020) (Figure 1). The assignment to the *Betacoronavirus* genus and in particular the similarity to SARS-CoV allows us to also draw on a field of already published work to infer promising targets within the SARS-CoV-2 life cycle for small molecule intervention.

Therapeutics against viruses generally fall into two broad categories: (i) vaccines, which prime the host immune system to mount a targeted defense against infection by the virus, and (ii) other agents with antiviral effects, which include, but are not limited to, small molecule inhibitors, peptides, and biologics. Vaccine development traditionally takes several years or more, and there are many viruses for which there is still no vaccine, even after decades of research. However, the goal of developing new vaccines to target threats such as human immunodeficiency virus (HIV), influenza, and ebolaviruses has resulted in a great deal of research into next-generation vaccine platforms (Delany et al., 2014; Antrobus et al., 2014; Meyer et al., 2018). Before this year, there were no approved vaccines for any of the human CoVs, including SARS-CoV and MERS-CoV, but during the current outbreak, scientists have been able to rapidly leverage previous platform work into prophylactic vaccines for SARS-CoV-2 (Merryn Voysey et al., 2020; Jackson et al., 2020; Polack et al., 2020).

In addition to vaccines, viral infection can be limited by the use of other therapeutics with antiviral properties. These therapeutics can be broadly divided into two categories: (i) drugs that target viral proteins and (ii) drugs that target host proteins, including the viral receptor, cellular proteins essential to the viral life cycle, and components of the immune system.

Currently, available viral therapeutics include protease inhibitors, reverse transcriptase and polymerase inhibitors (both nucleoside analogs and non-nucleoside inhibitors), integrase inhibitors, and viral entry blockers (Figure S1). Although there are no drugs approved by the United States Food and Drug Administration (FDA) specifically for SARS-CoV, MERS-CoV, or SARS-CoV-2, there have been a number of efforts to target proteins of the viral machinery, including the proteases, helicase, and RNA-dependent RNA polymerase (RdRP). A partial list of these published small molecules is shown in Table S2. A well-publicized example, remdesivir, is an RdRP inhibitor originally developed for hepatitis C that was also previously tested against both SARS-CoV and MERS-CoV. It should be noted that subsequent studies on its use in

¹⁵Department of Microbiology, Boston University Medical School, Boston University, Boston, MA 02118, USA

¹⁶Mol* Consortium, <https://molstar.org>, San Diego, CA 92109, USA

¹⁷VirtualFlow Organization, <https://virtual-flow.org/>, Boston, MA 02115, USA

¹⁸Department of Nutritional Sciences & Toxicology, University of California Berkeley, Berkeley, CA 94720, USA

¹⁹UkrOrgSyntez Ltd., Kyiv 01133, Ukraine

²⁰Present address: Molecular Immunity Unit, Department of Medicine, University of Cambridge, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK

²¹Present address: Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Department of Medicine, University of Cambridge, Cambridge CB2 0AW, UK

²²Lead contact

*Correspondence: cgorgulla@g.harvard.edu (C.G.), hari@hms.harvard.edu (H.A.) <https://doi.org/10.1016/j.isci.2020.102021>

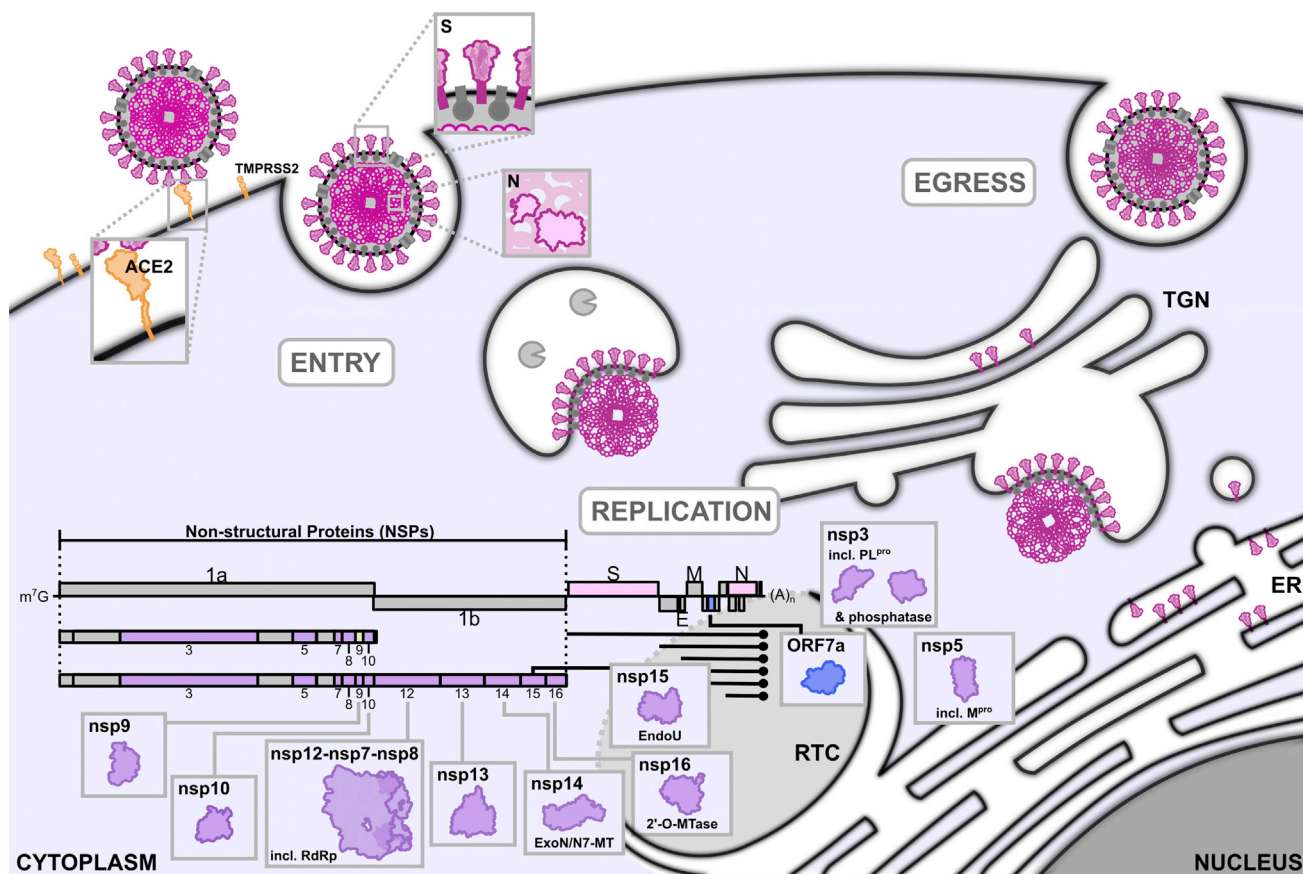


Figure 1. Schematic of the viral life cycle of SARS-CoV-2

The genome organization is based on other coronaviruses and published predictions (Snijder et al., 2016; Wu et al., 2020). ACE2: angiotensin-converting enzyme 2; TMPRSS2: transmembrane protease, serine 2; RdRp: RNA-dependent RNA polymerase; ExoN: exonuclease; N7-MT N7-methyltransferase; 2'-O-MTase: 2'-O-methyltransferase; EndoU: uridylylate-specific endonuclease; RTC: replication and transcription complex; ER: endoplasmic reticulum; TGN: trans-Golgi network.

the treatment of COVID-19 have yielded conflicting results, with some suggesting that remdesivir may reduce the time to recovery in patients with SARS-CoV-2 (Wang et al., 2020a; Beigel et al., 2020), while others showed no appreciable effect (W.H.O. Solidarity Trial Consortium, 2020).

For every virus, there are targets such as the entry protein or the viral polymerase that have direct effects on the ability of the virus to replicate, even under ideal conditions for the virus (e.g., in the absence of an immune system). There are also other targets that, if inhibited, would reduce the severity of the viral infection, often referred to as the viral virulence. Some of these targets may only have an effect when inhibitors are evaluated in the context of a whole organism. For example, several viruses paralyze the host immune system by hijacking or sequestering key signaling proteins, and targeting the viral proteins that mediate this effect would reduce the virulence of the virus. In this study, we targeted an array of SARS-CoV-2 proteins in addition to two human proteins, and we posited that inhibitors, which target proteins that affect the viability and/or the virulence, could synergistically combine for more efficacious combination therapies.

In silico screening methods enabled us to identify lead drug candidates for SARS-CoV-2 in an expedited manner, fueled by the availability of high-resolution structures of many of the SARS-CoV-2 proteins. Previous work has shown that the number of compounds screened in a structure-based *in silico* screen affects the quality of the resulting hits, and the potency of the hits derived increases with the number of compounds screened (Lyu et al., 2019). A flurry of recent papers has demonstrated that by using *in silico* screening to expand the chemical space (screening ~100 million compounds), one can identify picomolar affinity inhibitors (Lyu et al., 2019; Wan et al., 2020; Stein et al., 2020). To become capable of screening

billions of molecules in a relatively short period (weeks), we previously developed VirtualFlow, an open-source computational drug discovery platform (Gorgulla et al., 2020a). VirtualFlow, with its linear scaling feature, leverages the power of computing clusters to perform ultra-large-scale screens. Here, we used the computational power of the Google Cloud Platform (GCP) to identify molecules that could inhibit SARS-CoV-2. We screened over one billion molecules against each of fifteen SARS-CoV-2 proteins and two human proteins, ACE2 and TMPRSS2. For some of the proteins, we targeted multiple functional sites on the protein surface; an approach which provides the following two distinct advantages: (i) the possibility of increasing efficacy using a cocktail approach and (ii) the identification of alternative ways to target the same protein, a necessity when faced with the possibility of resistance mutations. Whenever available, we used recently determined high-resolution structures of SARS-CoV-2 proteins in our screens. A complete list of the target proteins along with the models used can be found in Table 1.

When performing *in silico* screening, the probability of finding a true hit (a hit that is validated experimentally) increases with the number of molecules screened and the accuracy of the chosen structure relative to what exists in solution inside the cell. Small changes in structure or dynamics can be accommodated by flexible docking, but large conformational changes decrease the ability of *in silico* screening methods to identify potent binders. Large conformational changes can be tackled by ensemble docking methods (molecular dynamics simulations), but the trade-off for these approaches is additional computational time. Targeting protein-protein interactions (PPIs) is also challenging for *in silico* screening. Typically, with PPIs, the off-rate between the two interacting proteins is slow, and the interaction interface is large compared to a small molecule inhibitor. Thus, targeting PPIs was approached in this study with the rationale that a small molecule engaging the monomer at the hot spot of the PPI interface could prevent complex formation, particularly if driven by an excess of the small molecule. In the absence of a solved dimeric/multimeric structure, we made the assumption that the structure of the monomer is similar to that of the protein in complex, a fact which may not always hold true. In light of such caveats, we describe any potential limitations that we anticipate could affect the experimental outcome for each of the targets that were screened, as well as generally for the study as a whole.

RESULTS

Preventing coronavirus entry

The spike protein forms the highly glycosylated trimeric receptor-binding protein that decorates the virion surface and facilitates entry into the host cell through interaction with its receptor ACE2 (Hoffmann et al., 2020). The Binding of the receptor-binding domain (RBD) of S to ACE2 induces large conformational changes in the S2 domain. These changes are followed by processing by host proteases such as TMPRSS2 and cathepsins and result in the formation of a stable six-helix bundle (6-HB) by heptad repeats 1 and 2 (HR1 and HR2) culminating in the fusion of the viral and cellular membranes (Huang et al., 2020). In order to develop compounds capable of preventing SARS-CoV-2 from entering cells, we focused on the following three targets: (a) the interaction interface of the RBD and ACE2, (b) the host protease TMPRSS2, and (c) the HR2 hydrophobic binding groove of HR1.

ACE2-spike-RBD interaction

Recent studies have revealed molecular details of ACE2 binding with spike-RBD, and multiple high-resolution structures of the RBD in complex with ACE2 have been released (including Protein Data Bank (PDB): 6m17, 6m0j, 6vw1) (Yan et al., 2020a; Lan et al., 2020; Shang et al., 2020a). As the primary mediator of host cell attachment, the spike trimer is a clear potential target for therapeutic intervention, and in order to find an effective inhibitor of viral entry into host cells, we targeted both ACE2 and the RBD of the spike in parallel.

Targeting the spike protein interaction interface on ACE2. The primary physiological function of the receptor, ACE2, is the maturation of angiotensin, a signaling peptide with cardiovascular regulatory effects (Donoghue et al., 2000), making any targeting of the catalytic site of ACE2 unfavorable. Fortunately, the structure of the ACE2-RBD complex shows that the peptidase (PD) domain of ACE2 has two distinct lobes that are essential for substrate engagement, and while the RBD-ACE2 binding interface includes one of these lobes, it does not obstruct the PD active site (Yan et al., 2020a; Lan et al., 2020) (Figures 2 and S3–S6). This suggests that targeting the RBD interaction interface should not inhibit ACE2 functionality and therefore not affect angiotensin homeostasis. In fact, previous biochemical studies with SARS-CoV have

Table 1. Overview of the performed virtual screens

Protein index	Protein name	Alternative names	Screen ID	Target site	Structure used
1	ACE2	Angiotensin-converting enzyme 2	1	Spike RDB binding region – site 1	6m17
			2	Spike RDB binding region – site 1	6m18
			3	Spike RDB binding region – site 2	6m17
			4	Dynamic pocket 1 besides spike RDB binding region	DE Shaw MD simulation 10875754 frame 2715
			5	Dynamic pocket 2 besides spike RDB binding region	DE Shaw MD simulation 10875754 frame 5273
2	TMPRSS2	Transmembrane protease serine 2	6	Active site	SWISS-MODEL of TMPRSS2 – model 1
3	Spike	S-protein, S	7	Spike RDB – ACE2 interface	6w41
			8	Spike HR1 domain	6lxt
4	ORF7a	Protein 7a	9	Blind docking	6w37
5	nsp3-macrodomein	Phosphatase, (macro) X domein	10	Active site	6w6y chain A (closed active site)
			11	Active site	6w6y chain B (open active site)
6	nsp3-PL ^{pro}	PL ^{pro} , PLP, papain-like protease	12	Active site	6w9c
			13	Accessory pocket	6w9c
			14	DUB binding site	6w9c
			15	Active site and accessory pocket	6wx4*
7	nsp5	M ^{pro} , main protease	16	Active site	6lu7
			17	Active site	6m0k*
			18	Dimerization site	6wqf
			19	α -helix 5 attachment site	Hybrid in-house model
8	nsp7	Replicase polyprotein 1ab	20	Blind docking (nsp8 PPI, nsp12 PPI)	6wiq
9	nsp8	Primase complex	21	nsp7 PPI	6wiq
			22	nsp12 PPI	7bv1
10	nsp9	Replicase	23	Dimerization interface – site 1	6w4b
			24	Dimerization interface – site 2	6w4b
11	nsp10		25	nsp16 PPI	6w4h
			26	nsp16 PPI	SWISS-MODEL of nsp10/14 - model 3
			27	nsp14 PPI	SWISS-MODEL of nsp10/14 - model 3

(Continued on next page)

Table 1. Continued

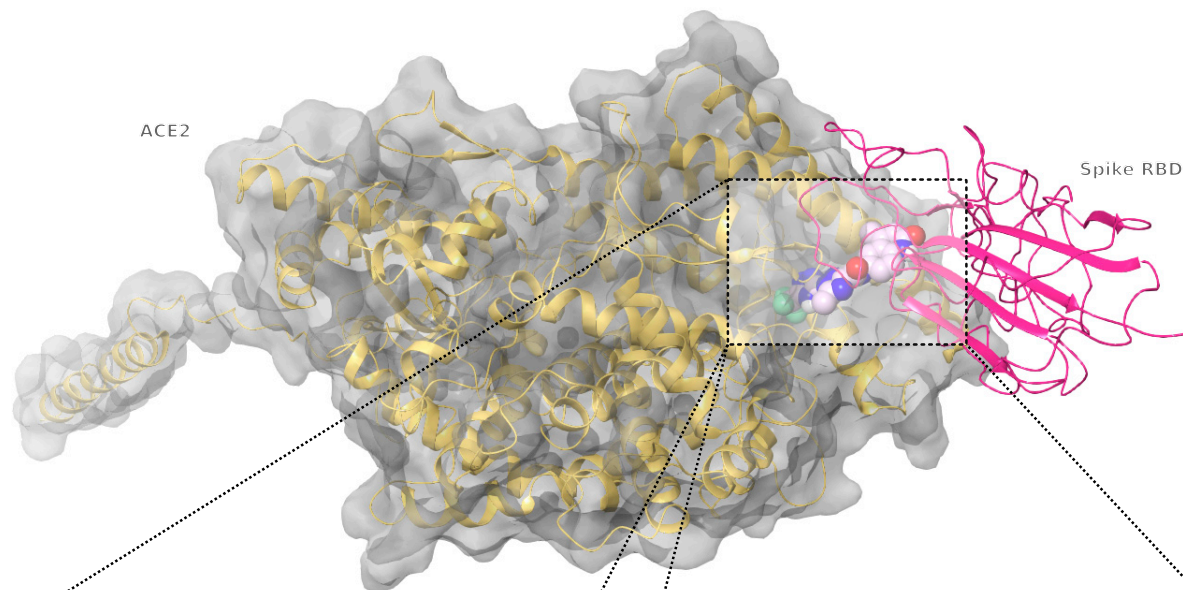
Protein index	Protein name	Alternative names	Screen ID	Target site	Structure used
12	nsp12	RNA-dependent RNA polymerase (RdRP)	28	RNA binding interface – site 1	7bv1
			29	RNA binding interface – site 2	7bv1
			30	Nucleotide binding site	7bv1
			31	nsp8 PPI	6m71
			32	nsp7/8 PPI	7bv1
13	nsp13	Helicase	33	Active site	SWISS-MODEL nsp13 – model 1
			34	RNA binding interface – site 1	SWISS-MODEL nsp13 – model 1
			35	RNA binding interface – site 2	SWISS-MODEL nsp13 – model 1
14	nsp14	Exoribonuclease, N7 methyltransferase, N7-MTase	36	nsp10 PPI	SWISS-MODEL of nsp10/14 – model 3
			37	Active site (ExoN)	SWISS-MODEL of nsp10/14 – model 3
			38	Active site (N7-MT)	SWISS-MODEL of nsp10/14 – model 3
15	nsp15	Endoribonuclease, XendoU	39	Active site	6w01
16	nsp16	2'-O-MTase, 2'-O methyltransferase	40	nsp10 PPI	6w4b
			41	Active site (2'-O MT)	6w4b
17	nucleoprotein	N, NC, NP, RNP, ribonucleocapsid protein	42	NTD – RNA binding site	6yi3
			43	NTD – oligomerization site	6yi3
			44	CTD – dimerization interface	6wji
			45	CTD – oligomerization site	6wji

A total of 45 virtual screens were performed, involving 17 different target proteins and 40 unique target sites. In each screen, approximately 1 billion molecules from the Enamine REAL library and approximately 10 million compounds from the in-stock compounds of the ZINC 15 database were screened. 6wx4*: The side chain rotamer of Leu162 in PDB 6wx4 was rotated to open up the end of the tunnel that forms the active site (for more details, please refer to the main text). 6m0k*: The side chain rotamers of the residues Met49, Ser46, and Cys145 of the PDB 6m0k crystal structure were modified to give increased access to the binding pocket (for more details, please refer to the main text and [Figure S12](#)).

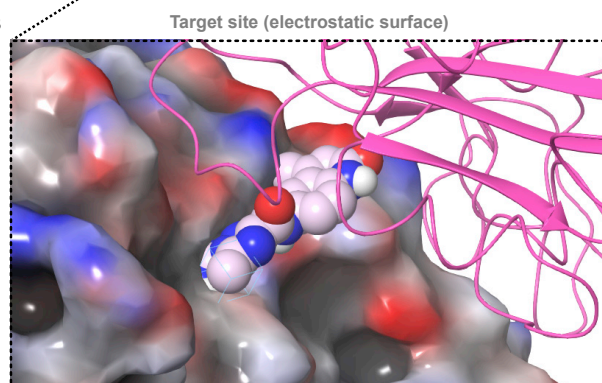
shown that the ACE2 inhibitor, MLN-4760, which favors the closed state, does not interfere with S protein binding and vice versa, also arguing for the independence of ACE2 function and spike binding ([Li et al., 2005](#)).

We have targeted four different sites on the ACE2 surface. Two of them are located on the spike interaction interface and would directly disrupt the ACE2-spike interaction ([Figures 2, S3, and S4](#)). For the first of these two sites (centered around Glu37), two independent virtual screens were carried out (screen IDs: 1 and 2): one using an ACE2 structure where the spike protein was bound (PDB: 6m17; screen ID: 1) and one using an ACE2 structure to which the RBD domain of the spike protein was bound (PDB: 6m18; screen ID: 2). The difference between these two conformations at the site of binding is relatively small (Root-mean-square deviation [RMSD]: 0.355 Å in the region of the first two N-terminal α -helices), yet still significant for

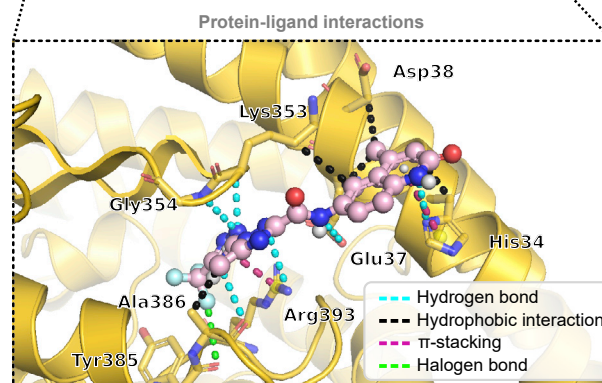
A



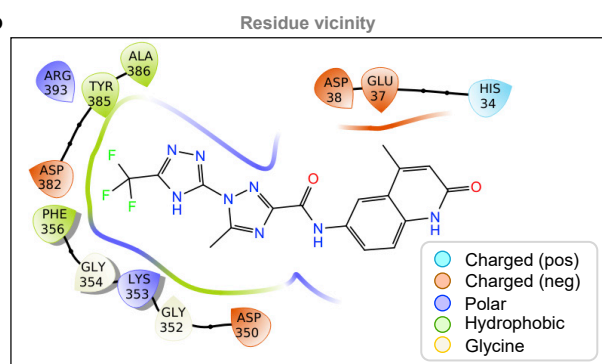
B



C



D



E

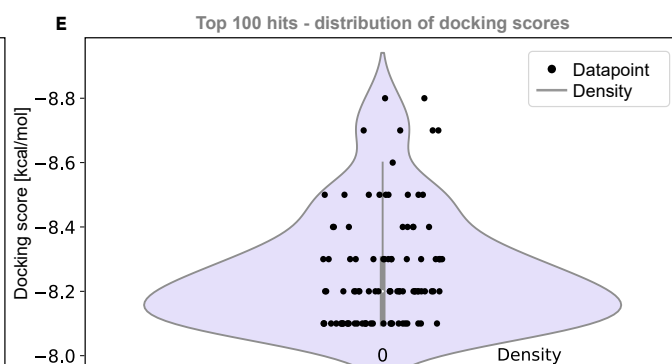


Figure 2. The ACE2 receptor and an example compound from the top 0.0001% of screened compounds bound at the spike interaction interface (site1)

(A) The target protein ACE2 (gold) bound to the RBD of the spike protein (magenta) and an example compound (light pink) from the virtual screen bound to the spike interaction interface (site 1, around Glu37) (screen ID: 1).

(B) The electrostatic surface of the target protein (ACE2) bound to the RBD domain of the spike protein (magenta) and an example compound (light pink).

(C) An overview of the interactions between the ligand and the receptor structure.

(D) Receptor residues within 4 Å of the ligand.

(E) Distribution of the docking scores of the top 100 virtual screening hits.

structure-based virtual screening. The second site that we targeted on the spike binding interface (screen ID: 3) is centered around residue Gly354 and was chosen because there is a shallow hydrophobic pocket. The other two sites targeted were two dynamic pockets adjacent to the spike interface on ACE2, potentially capable of accommodating tighter binders (screen IDs: 4 and 5; [Figures S5 and S6](#)). In addition, compounds that bind to the two dynamic pockets could be chemically linked to each other in such a way that the linker, which would then pass over the spike interaction interface, could potentially block the binding of spike protein. In these latter two screens, we used conformations of ACE2 derived from molecular dynamics (MD) simulations conducted by D.E. Shaw Research ([Shaw Research, 2020](#)). In these MD simulations, ACE2 samples an open state, revealing the two dynamic/cryptic pockets that we targeted. These conformations have not been confirmed by experimental methods and thus are of an exploratory nature, and hits from these screens can be considered high-risk, high reward.

Targeting the ACE2 interaction interface on the spike RBD. *In vitro* measurements suggest that SARS-CoV-2 RBD binds to ACE2 with nanomolar affinity ([Walls et al., 2020](#); [Wrapp et al., 2020](#)). Each PD domain of ACE2 can accommodate one RBD, and the interaction is mediated through polar interactions. Furthermore, it has previously been shown that the “up” conformation of the RBD is required for receptor binding and that two spike protein trimers can bind to a single ACE2 dimer ([Yan et al., 2020a](#)) ([Figure S7](#)). The RBD itself consists of a core containing an antiparallel β -sheet (β 1 to β 4 and β 7), with three short α -helices and an extended loop. This extended loop engages the α 1 helix of the ACE2-PD, and it is the center of this interface on the RBD that we have targeted by virtual screening (screen ID: 7) ([Figure S7](#)).

The spike interaction interface on ACE2 is relatively flat and polar, which makes it challenging to identify molecules that bind to it with high affinity. This difficulty is reflected in the docking scores resulting from the virtual screen of this binding interface: the average docking score of the top 100 virtual hit compounds from this screen was approximately -8 kcal/mol, which is among the weakest docking score averages of all the target sites screened ([Figure S8](#)).

The SARS-CoV-2 RBD has a significantly higher affinity for ACE2 as compared to that of the SARS-CoV RBD ([Shang et al., 2020a](#); [Wrapp et al., 2020](#)); however, the full-length spike protein from SARS-CoV-2 shows an either similar or weaker affinity for ACE2 ([Walls et al., 2020](#); [Shang et al., 2020b](#)) when compared to that of the full-length SARS-CoV spike. This contrast in binding affinities between the RBD alone and the full-length spike protein could be a consequence of RBD dynamics where the protein switches between RBD “up” and “down” states, as is also seen in the case of SARS-CoV ([Yuan et al., 2017](#); [Gui et al., 2016](#)). There is limited information available as to whether the SARS-CoV-2 spike favors a “lying down” binding inactive state or a “standing up” state capable of receptor binding ([Walls et al., 2020](#); [Wrapp et al., 2020](#)). Furthermore, while structural studies have looked at conformational flexibility in the spike protein ([Turoňová et al., 2020](#); [Ke et al., 2020](#)), there is little information on the inherent dynamics within the RBD domain itself, which is where the targeted molecules are expected to bind. This means that it is possible that local protein dynamics could adversely affect the binding of some of the hits in ways that the screen is unable to predict.

It should be noted that while targeting the flat surface of the spike-binding interface with ACE2 (screen IDs: 1-3) resulted in compounds with an average binding energy of -8.3 kcal/mol, targeting the dynamic pockets resulted in hits with the potential for higher affinity binding. The binding energy of hits in the dynamic pockets was as high as -10.3 kcal/mol. However, experimental validation of these hits would be required to confirm this apparent difference in affinity.

TMPRSS2

The essential priming of S during entry can be executed by the host serine protease TMPRSS2 in the case of SARS-CoV-2, making it a potential therapeutic target. In addition, recent research has shown that the TMPRSS2 inhibitor, camostat mesylate, can block viral entry in cell-based assays ([Hoffmann et al., 2020](#)). Effective inhibition of TMPRSS2 could have multiple applications since processing of ACE2 by TMPRSS2 has also been shown to be crucial for activation of S in SARS-CoV ([Heurich et al., 2014](#)). Although TMPRSS2 has been shown to activate protease-activated receptor 2 and hepatocyte growth factor, giving it possible roles in tumor metastasis and epithelial-mesenchymal transition ([Lucas et al., 2014](#)), the exact physiological function of this protein is still unknown. TMPRSS2-deficient mice showed no significant phenotypic changes ([Kim et al., 2006](#)), and thus, TMPRSS2 is considered to be dispensable for normal cellular function, making a TMPRSS2-specific inhibitor less likely to have side effects.

TMPRSS2 itself contains a cytoplasmic domain, a transmembrane domain, an extracellular low-density lipoprotein receptor domain, a scavenger receptor cysteine-rich domain, and a C-terminal serine protease domain. The TMPRSS2 precursor consists of 492 amino acids and auto-catalytic cleavage at Arg255 generates the active protease (Afar et al., 2001). As a chymotrypsinogen-like protease, TMPRSS2 contains a catalytic triad consisting of residues His296, Asp345, and Ser441. As there are currently no experimental structures available for TMPRSS2, we used a homology model from SWISS-MODEL that is based on the human serine protease hepsin (PDB: 5ce1), which shares 33.82% sequence identity with TMPRSS2 (O15393, n.d.). Approximately 1 billion compounds were screened against the active site of this homology model (Figure S9).

Since a homology model was used to screen for inhibitors, success would depend on how closely the model represents the actual structure. The active site of TMPRSS2 is well conserved when compared to the template, the serine protease hepsin, and also contains a catalytic triad consisting of Ser, His, and Asp residues (Figure S10). The residues around the binding pocket are also well conserved, making it more likely that the predicted structure is sufficiently similar to the true structure in the targeted region around the active site binding pocket (screen ID: 6). Our top 10 hits had an average binding energy of -11.3 kcal/mol, which should ideally correspond to nanomolar binding affinities. It must be noted that experiments with the TMPRSS2 inhibitor camostat mesylate in cultured cells showed that complete inhibition of viral entry is achieved only in conjunction with an inhibitor of cathepsins B and L (CatB/L), E-64d (Hoffmann et al., 2020); however, the impact of this residual priming of S by CatB/L on viral pathogenesis *in vivo* is yet to be determined.

HR1-HR2 interface

While the S1 subunit of the spike protein is crucial for host cell attachment, the S2 subunit plays an essential role in membrane fusion and subsequent entry into the host cell. Heptad repeats HR1 (residues 912–984) and HR2 (residues 1163–1213) in S2 interact to form the classic 6-HB of a class I entry protein (Xia et al., 2020a; White et al., 2008). The HR domains of the spike protein and their mode of interaction are known to be highly conserved across CoVs, making them an attractive target for the development of pan-CoV fusion inhibitors (Liu et al., 2004).

The central trimeric coiled coil of HR1 has three hydrophobic grooves (Figure S11) that are proposed to be potential drug binding sites for inhibition of viral fusion. The recent crystal structure of HR1-L6-HR2 (PDB: 6lxt) displays a 6-HB structure with the HR1 domains forming a parallel trimeric coiled coil around the anti-parallel HR2 domain. The hydrophobic groove formed by the HR1 helices is also the binding interface for the HR2 domain (Xia et al., 2020b). The hydrophobic interaction between HR1 and HR2 is in the area of the fusion core, and previous studies have shown that both HR1 and HR2 components of this fusion core interface could be effectively targeted using complementary peptides in SARS-CoV-2 and MERS-CoV (Xia et al., 2020a; Lu et al., 2014).

We performed a blind docking using QuickVina W against the entire surface of the trimeric alpha helix (screen ID: 8; Figure S11). The resulting hits could not only potentially have high inhibitory activity against multiples CoVs but also have both prophylactic and therapeutic applications, as has been previously observed in the case of a peptide inhibitor (Xia et al., 2019). Ligands binding to each of the three different hydrophobic grooves could be identified and linked together in order to enhance the overall affinity and thereby increase the potency.

Preventing coronavirus replication and curtailing viral load

Once CoV enters the cell, translation of the replicase gene gives rise to two large polyproteins, PP1a and PP1ab, composed of nsp1-11 and nsp1-16, respectively. M^{pro} and PL^{pro} cleave these polyproteins into individual nsps, which facilitate genomic replication and transcription of sub-genomic RNAs. In order to develop small molecules capable of inhibiting replication, we targeted the following five proteins and protein complexes, selected based on the availability of high-resolution structures and available information about their molecular mechanisms: (a) the CoV primary protease (M^{pro} , nsp5), (b) the papain-like protease (PL^{pro} , nsp3), (c) the RNA replication complex, composed of the RNA-dependent RNA polymerase (RdRp) (nsp12) and its co-factors nsp7 and nsp8, (d) the helicase (nsp13), and (e) the ssRNA-binding protein nsp9.

M^{pro} (3CL pro)

M^{pro} (3CL pro), the main protease of CoV, cleaves PP1a and PP1ab into many of their constituent nsps (11 cleavage sites in PP1ab). The inhibition of M^{pro} would not only inhibit the protease itself but also hinder

downstream processes by preventing the production of key viral proteins through inhibition of their proteolytic processing. Proteolytic cleavage by M^{Pro} via the catalytic dyad (Cys-His) occurs at the P1 glutamine residue in the cleavage motif (Anand, 2003). Dimerization is thought to create a substrate binding cleft (Shi et al., 2008), and simulations suggest that only one monomer is active at a time (Chen et al., 2006). M^{Pro} is highly conserved not only in SARS-CoV and SARS-CoV-2 but also among other viruses of the *Nidovirales* order (enveloped, positive-sense, single-stranded RNA viruses that produce nested 3'-co-terminal genomic RNAs), making it a potential target for the development of antiviral drugs with broad efficacy (Nukoolkarn et al., 2008).

Using the same strategies that allowed crystallization of SARS-CoV M^{Pro} (Yang et al., 2003), several groups determined the three-dimensional structure of SARS-CoV-2 M^{Pro} in its apo form or bound to various inhibitors (Zhang et al., 2020; Jin et al., 2020a; Dai et al., 2020; Jin et al., 2020b; Su et al., 2020; and deposited only, PDB: 6y84). These structures revealed that the N-terminus of SARS-CoV-2 M^{Pro} contains two domains (domain I, residues 10–99; domain II, residues 100–184) that adopt a chymotrypsin-like fold, whereas the M^{Pro} C-terminal domain III (residues 201–303) is composed of five α -helices and is responsible for the intermolecular interactions critical for the dimerization. The protease active site of the picornavirus 3C-protease-like domain is located at the cleft formed between domains I and II and contains a Cys145-His41 catalytic dyad. A total of five protein segments comprising residues 25–27 and 44–50 from domain I and residues 140–143, 165–168, and 188–190 from domain II form the walls of the active site (Figure S12a).

It is important to note that the M^{Pro} active site seems to have significant conformational flexibility (Bzówka et al., 2020). Comparison of three apo structures and five structures in complex with inhibitors reveals differences in the conformation and position of the Gln189-containing loop and the short Ser46-containing α -helix (Figure S12b). Specifically, the cleft between domains I and II is narrower in the apo structures than in the structures of M^{Pro} bound to inhibitors, consistent with an induced fit or conformational selection upon ligand binding.

Given the plasticity of the active site, we used two different conformations of M^{Pro} for *in silico* screening (screen IDs: 16 and 17; Figures S13 and S14). The first conformation corresponded to the structure of M^{Pro} bound to the peptide-like inhibitor N3 (Jin et al., 2020a). For this search, the protein model used was identical to the protein structure reported in the PDB file, after removal of all water molecules and bound ligands (screen ID: 16; Figure S12c). The second conformation corresponded to the structure of M^{Pro} bound to inhibitor 11b, an inhibitor with peptide-like features (Dai et al., 2020). However, the model used for screening differed slightly from the conformation reported in the PDB file. In the crystal structure, the side chain of Met49 projects into the active site and thus would potentially interfere with compound docking. To avoid this, a different Met49 side chain rotamer was used that did not project as much into the active site. This change resulted in steric clashes with the side chain of Ser46, for which selection of a different side chain rotamer resolved these incompatibilities. Finally, a rotamer where the side chain points toward the core of the protein was selected for Cys145. This rotamer was selected to allow the possible docking of compounds that could form covalent bonds with the Cys145 thiol group (Figure S12d).

Screens against these two conformations, hereafter referred to as M^{Pro}-N3 (screen ID: 16) and 6m0k* (screen ID: 17), respectively (Figure S12), identified compounds that docked with high binding energies. Interestingly, the 6m0k* conformation resulted in predicted binding free energies that were approximately 0.5 kcal/mol lower than the binding free energies (docking scores) obtained from screening the M^{Pro}-N3 conformation. Moreover, among the top 1000 hits identified by screening each of the two conformations (with the application of the standard filtering criteria of logP less than 6, molecular weight less than 600 daltons, and no reactive groups), only 12 hits overlapped. These results illustrate how screening closely related protein conformations can lead to the identification of a diverse set of docked compounds, and this may be a generalizable strategy for *in silico* screening of proteins that exhibit conformational flexibility, such as SARS-CoV-2 M^{Pro} (Bzówka et al., 2020).

Since the dimerization of M^{Pro} is known to be essential for its catalytic activity (Xia and Kang, 2011; Kuo et al., 2004; Hsu et al., 2005), we also targeted the dimerization interface, which relies primarily on the α 1 helix of the C-terminal domain (screen ID: 18). Previous studies have shown that the inter-conversion of M^{Pro} from monomer to dimer is mediated by an order-to-disorder transition of α 5, and this aids domain

swapping without exposure of the protein's hydrophobic core (Kang et al., 2012). Since the C-terminal region of M^{pro} is fully conserved between SARS-CoV and SARS-CoV-2, we used an NMR ensemble of the SARS-CoV M^{pro} C-terminal domain (PDB: 2liz) to build a hybrid model of a disordered C-terminus. We subsequently used this model and screened against the attachment site of the $\alpha 5$ helix (screen ID: 19). Potential binders at this position would be expected to trap the protein in its inactive monomeric form by preventing dimerization.

PL^{pro}

SARS-CoV-2 nsp3 is a multi-domain protein containing an N-terminal ubiquitin-like fold, a glutamic-acid-rich acidic domain, a phosphatase domain, a SARS unique domain, a catalytically active PL^{pro} domain, a marker domain (G2M), and two transmembrane domains (Báez-Santos et al., 2015). In addition to the protease activity, PL^{pro} also possesses deubiquitination and deISGylation activities that aid in the disruption of the interferon regulatory factor 3 pathway and host innate immune responses (Barretto et al., 2005). The central α -helical "thumb" domain of PL^{pro} harbors the catalytic Cys112, while the C-terminal domain is mostly β -sheet and includes the "palm" domain, which harbors the other two residues of the catalytic triad, His275 and Asp287, and the "fingers", which coordinate a Zn^{2+} ion (PDB: 6w9c). PL^{pro} also has two blocking loops (BL1 and BL2), targeting of which could prevent the substrate from entering the active site.

Comparison of the crystal structures of PL^{pro} in its unbound and peptide inhibitor-bound states (PDBs: 6w9c and 6wx4, respectively) revealed conformational changes in both the "finger" domain, as well as the BL2 loop. We therefore targeted the active site in the unliganded and the bound states, which correspond to the open and closed conformations of the active site (screen IDs: 12 and 15, respectively; Figures S15 and S16). The side chain rotamer of Leu162 of structure 6wx4 was rotated to open one end of the tunnel that forms the active site. Without this modification, the end of the tunnel would not be accessible with the docking programs we used. However, the co-crystal structure 6wx4 has a ligand extending through that tunnel, thus justifying the possibility of a small molecule passing through the tunnel. In addition to this, we also targeted the accessory cleft leading to the active site, which would prevent substrate access (screen ID: 13; Figure S17). In addition to the three sites that are described here, the ubiquitin contact site in PL^{pro} , which is critical for the deubiquitination and deISGylation functions, was also targeted and is described in detail in a later section.

Replication complex: nsp7, 8, and 12

nsp7, nsp8, and nsp12 form a minimal 160-kDa complex capable of nucleotide polymerization with additional nsps playing important roles in RNA modification (Lehmann et al., 2015; Sevajol et al., 194). nsp12 encodes for the RdRp, while nsp7 and nsp8 are cofactors that may form a hexadecameric supercomplex (Smith et al., 2013). When compared to the SARS-CoV replicase, key functional residues are found to be fully conserved in SARS-CoV-2; however, substitutions have been found on the surface of the complex. The polymerase domain of nsp12 has an N-terminal nidovirus-unique N-terminal extension in addition to the "finger", "palm", and "thumb" domains common to polymerases (Gao et al., 2020) (Figures S18–S22). The seven catalytic motifs found in other known viral polymerases and which are involved in template and nucleotide binding and in catalysis are also well conserved in the SARS-CoV-2 polymerase (Gao et al., 2020; Kirchoerfer and Ward, 2019).

The SARS-CoV-2 RNA synthesis machinery is mediated largely by PPIs, and this provides us with an opportunity to target these interfaces using small molecules. The C-terminal head domain of nsp8 wraps around the nsp7 helical bundle, forming a hexadecameric complex. nsp8 has been suggested to possess RdRp activity of its own and produce the primers utilized by the primer-dependent nsp12 RdRp. This nsp12 activity requires the formation of a large oligomeric complex that brings active site residues into proximity with each other (Imbert et al., 2006; te Velthuis et al., 2011). A previously determined crystal structure (PDB: 2ahm) shows nsp7 and nsp8 capable of interacting in at least two ways: (i) residues on the C-terminal end of the nsp8 shaft form a hydrophobic core with residues in helix 1 of nsp7, and (ii) helix 1 and the short helix of nsp7 interact with one of the helices of the nsp8 head domain (Figure S23). The heterodimer formed by nsp7-nsp8 interaction binds to nsp12 on the polymerase "thumb" domain facing the nucleotide triphosphate (NTP) entry channel, thus placing the nsp12 polymerase "index finger" loop between nsp7-nsp8 and the polymerase "thumb" domain (Figure S23). In addition to this interaction, a second nsp8 monomer binds to the nsp12 interface domain in the vicinity of Leu117. We targeted the nsp12 binding interface on nsp8 (screen ID: 22; Figure S24), the nsp7, binding interface on nsp8 (screen ID: 21; Figure S25), the entire alpha-

helical surface of nsp7, which includes the nsp8 and nsp12 binding interfaces (screen ID: 20; [Figure S23](#)), the nsp8 binding interface on nsp12 (screen ID: 31; [Figure S18](#)), two sites in the RNA binding region on nsp12 (screen IDs: 28 and 29; [Figures S19 and S21](#)), the nucleotide binding site (screen ID: 30; [Figure S20](#)), and the nsp7 binding interface on nsp12 (screen ID: 32; [Figure S22](#)).

The nsp7 binding interface on nsp12 is relatively shallow, and to a lesser degree, the RNA binding sites are as well. One of the principal challenges of targeting an RNA binding interface is the strong avidity effect, frequently in the nanomolar range, that restricts the ability of ligands to displace the RNA. Moreover, biological processes involving RNA-protein interactions are also known to often be dynamics dependent. The average binding free energies of the top 100 hits from the screens of the RNA binding interfaces of nsp12 (screen IDs: 28 and 29) were -9.9 and -9.3 kcal/mol, whereas the docking score for the nucleotide binding site (screen ID: 30) was -9.7 kcal/mol. The screens targeting the PPIs (screen IDs: 31 and 32) resulted in hits with potential for much higher affinity binding with an average binding energy of -11.1 and -10.4 kcal/mol ([Figure S8](#)).

nsp13: helicase

nsp13, which is one of the few proteins that are fully conserved in SARS-CoV and SARS-CoV-2, has known activity as an RNA helicase, NTPase, dNTPase, RTPase, and DNA helicase ([Ivanov et al., 2004](#)). It is also known to interact with nsp12 ([Adedeji et al., 2012](#)). The helicase is 603 amino acids long, contains an N-terminal zinc binding domain, a stalk domain, RecA-like domains 1A, 1B, and 2A, and forms a triangular-shaped structure ([Figure S26](#)) ([Jia et al., 2019](#)). This protein has multiple potential target sites. Of these, we chose to target the active (ATP binding) site (screen ID: 33; [Figure S26](#)) and two sites on the RNA interaction interface (Screen IDs: 34 and 35; [Figures S27 and 3](#)).

We made use of a homology model derived from SWISS-MODEL repository ([N2ngu3, n.d.](#)) for these screens ([Bienert et al., 2016](#)); this model is expected to be reliable because the helicase of SARS-CoV-2 shares 99.8% sequence identity with the chosen template (PDB: 6jyt, structure of SARS-CoV nsp13). In fact, the only difference between the helicases of SARS-CoV and SARS-CoV-2 is a valine to isoleucine missense mutation in the C-terminus that is distant from our targeted sites. While this work was under review, additional structures of the helicase as a monomer and in a heterooligomeric complex with nsp12 and RNA were deposited (PDB: 6zsl; [Yan et al., 2020b](#), PDB: 7xcm, 7xcn). The backbone RMSD of the newly deposited apo structure (PDB: 6zsl) when compared to the model we used is 1.46 Å.

Previously reported hydrogen-deuterium exchange experiments suggest that the 1A and 2A domains of the helicase undergo a series of conformational changes during nucleotide binding and ATP hydrolysis ([Jia et al., 2019](#)). This potentially inherent flexibility could impact the plasticity of the screened binding pockets and therefore the reliability of the model screened. Our two virtual screens for the RNA binding interface (screen IDs: 34 and 35) resulted in top 100 hits with an average binding free energy of -10.2 and -10.7 kcal/mol, respectively, and the screen for the active site inhibitors (screen ID: 33) resulted in an average binding energy of -10.2 kcal/mol.

nsp9: ssRNA-binding protein

nsp9 is known to co-localize with nsp7, nsp8, and nsp10 within the replication complex and is presumed to play a role in RNA replication ([Yang et al., 2014](#)). A study on SARS-CoV nsp9 has shown that nsp9 is an ssRNA-binding protein indispensable for viral RNA replication in mouse models ([Frieman et al., 2011](#)). There are several high-resolution crystal structures available for nsp9 ([Sutton et al., 2004](#); [Hu et al., 2017](#)), including those of the protein from SARS-CoV-2 (PDB: 6w4b and 6w9q). Since dimerization has been found to be essential for viral replication ([Miknis et al., 2009](#)), we chose to target the dimerization interface as a method of inhibiting viral replication.

nsp9 contains seven antiparallel β -strands and an α -helix ([Figure S28](#)). The dimer interface is formed mainly by the parallel association of the C-terminal α -helices ([Figure S29](#)), and this surface, in particular the GXXXG motif, was targeted using VirtualFlow (screen ID: 23). We also targeted a shallow cavity around Phe75 that partially overlaps with the first targeted site centered on the α -helix (screen ID: 24). This second site interacts with the N-terminal RNA-binding domain (NTD) of the other nsp9 monomer and is more concave in shape than the first site, making it more pocket-like. The average docking score of the top 100 compounds of the second targeted site was -9.2 kcal/mol, which is a slight improvement over that of the first site with an average of -9.1 kcal/mol ([Figure S8](#)).

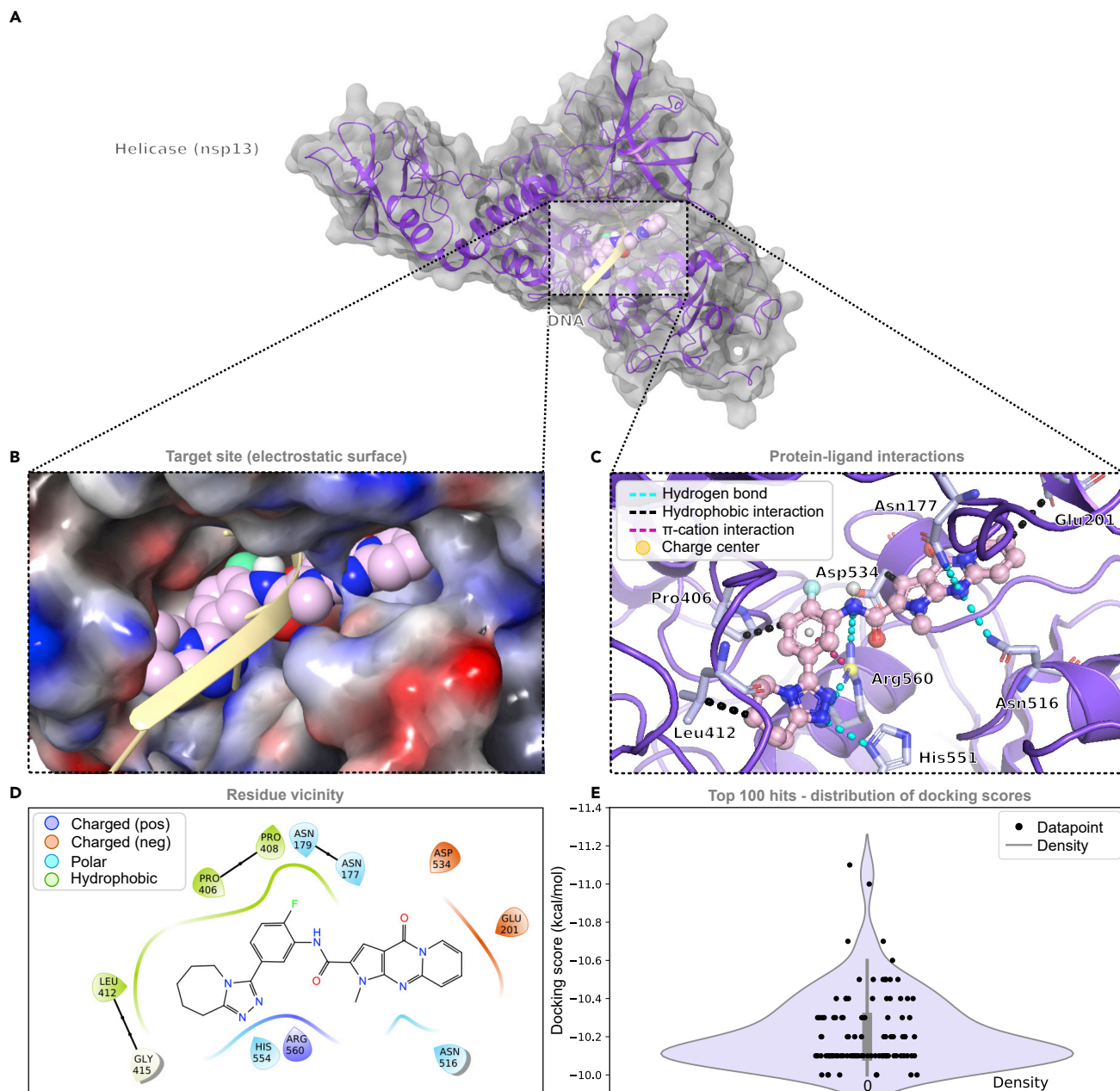


Figure 3. The helicase (nsp13) and an example compound from the top 0.0001% of screened compounds bound at site 2 of the RNA binding interface

(A) The helicase (violet) bound to an example compound (light pink) at region 2 of the RNA binding interface. A docked DNA strand is shown in light gold (screen ID: 35).

(B) Electrostatic surface of the helicase to which an example compound (light pink) and a docked DNA strand (light gold) are bound.

(C) An overview of the interactions between the compound and the helicase structure.

(D) Residues within 4 Å of the inhibitor.

(E) Distribution of the docking scores of the top 100 virtual screening hits.

The C-terminal α -helix, which is responsible for dimerization, has a high hydrophobic amino acid content, and dimer formation is mainly facilitated by hydrophobic interactions. Targeting a region with these characteristics could result in hits that are also hydrophobic and that interact in a non-specific manner, something which would need to be addressed during follow-up experimental validation. The interface also lacks deep pockets, making it additionally challenging to find specific, tight binding compounds. Finally,

targeting the dimerization interface is made even more difficult because it is not clear if the protein adopts different conformations between its monomeric form, which is to what an inhibitor might bind, and its dimeric form, which is the form that was used for the screen.

Disrupting the ability of coronaviruses to evade or subvert host defenses

Initial studies on SARS-CoV-2 in *ex vivo* systems and on clinical samples suggest that infection produces lower levels of interferon and differential cytokine/chemokine profiles, even when compared to the closely related SARS-CoV (Daniel Blanco-Melo et al., 2020; Chu et al., 2020). This suggests that SARS-CoV-2 is able to escape, at least initially, host immune surveillance, and when combined with robust early replication (Chu et al., 2020; To et al., 2020), it allows the virus to establish a strong foothold in the host, forcing the immune system to play catch-up. In light of this, targeting SARS-CoV-2 immune evasion factors, particularly in combination with other inhibitors, could have clinical benefits.

CoVs have an RNA genome and during their life cycle produce RNA species normally absent in the host cell's cytoplasm, such as dsRNA and RNA with a 5'-triphosphate. Sensing these types of RNA intermediates is one way in which the host cell can detect viral infection and initiate an antiviral immune program (Barbalat et al., 2011). An effective way in which CoVs avoid recognition by these intrinsic immune sensors is by using methyltransferases to add a cap structure to viral RNA, thus mimicking the cell's own mRNA (Daffis et al., November 2010). CoVs also encode an endoribonuclease, which prevents the accumulation of viral RNA that could activate dsRNA sensors in the cell (Deng et al., 2018). The deubiquitination and deISGylation activities of PL^{Pro} further interfere with the host antiviral response by antagonizing the induction of type I interferon (IFN) pathways (Clementz et al., 2010). With the goal of preventing SARS-CoV-2 from being able to evade or subvert host defense mechanisms, we targeted the following four proteins: (a) the guanine-N7 methyltransferase/ExoN (nsp10-14) complex, (b) the 2'-O-ribose methyltransferase (nsp16), (c) the uridylate-specific endoribonuclease (NendoU/nsp15), and (d) the deubiquitination and deISGylation activities of PL^{Pro}.

Guanine-N7 methyltransferase/ExoN (nsp10-14 complex)

nsp14 is critical for viral replication and transcription, as it plays dual roles in proofreading and mRNA capping (Eckerle et al., 2007; Denison et al., 2011; Chen et al., 2009). Disruption of nsp14 exonuclease activity (Minskaia et al., 2006) has been shown to result in increased sensitivity of the virus to ribavirin (RBV) and 5-fluorouracil (Smith et al., 2013), demonstrating the importance of this 3'-mismatched dsRNA excision activity (Bouvet et al., 2012). The ExoN activity relies on heterodimer formation of nsp14 with nsp10, and disruption of this interaction has been demonstrated to result in lowered replication fidelity (Smith et al., 2015). Unlike many other RNA viruses, CoVs have a low mutation rate by virtue of this ExoN activity, which supported the expansion of CoV genomes (Ogando et al., 2019) and is directly linked to its virulence (Graham et al., 2012). Previous work with SARS-CoV has suggested that the limited efficacy of RBV *in vivo* may be due to excision by this ExoN activity (Ferron et al., 2018), and work in mouse hepatitis virus (MHV) has shown that a mutant strain lacking ExoN activity was more sensitive to remdesivir (Agostini et al., 2018). These data suggest that use of a nucleoside analog might be more effective if combined with an ExoN inhibitor (Ogando et al., 2019). nsp14 also functions as an S-adenosyl-methionine (SAM)-dependent guanine-N7 methyltransferase (N7-MTase) (Chen et al., 2009). Mutation studies in a replicon system have shown that this N7-MTase activity is critical for viral replication and transcription (Chen et al., 2009). This activity is used to add a cap 0 structure at the 5' end of viral mRNA, thus mimicking a defining structural feature of the host mRNAs and assisting translation (Joseph et al., 1997), as well as evasion of host defenses (Decroly et al., 2011a). For nsp14, we targeted three sites: (i) the active site of the ExoN domain (screen ID: 37; Figure S30), (ii) the active site of the methyltransferase domain (screen ID: 38; Figure S31), and (iii) the PPI interface with nsp10 (screen ID: 36; Figure S32). In addition to this, we also targeted the nsp14 interaction interface on nsp10 (screen ID: 27; Figure S33).

In the bimodular nsp14, the N-terminal domain of nsp14 up to Lys288 interacts solely with nsp10 and forms the ExoN domain, whereas the C-terminal domain possesses the N7-MTase activity. The ExoN domain is composed of a twisted parallel β -sheet made up of five β -strands bordered by α -helices on both sides (Figures S30–32) (Ma et al., 2015). The N7-MTase domain of nsp14 is made up of five β -strands, and the ligand binding pocket is situated between the β 1 and β 2 sheets (Ma et al., 2015). At the time of screening, no experimental structures were yet available for the SARS-CoV-2 nsp10-14 complex, so we used a high confidence homology model available in the SWISS MODEL repository (Bienert et al., 2016) constructed from

the SARS-CoV ExoN/N7-MTase structure (PDB: 5c8s) (7bgwu, n.d.) for our docking experiments. While nsp10 is 97% conserved between the two viruses, nsp14 has a slightly lower 95.1% identity.

The nsp10 interaction site on nsp14 has a fairly hydrophobic shallow pocket, resulting in relatively high docking scores for the top scoring compounds. However, due to the sparsity of polar atoms at this site, it is challenging to find compounds that form sufficient electrostatic interactions to provide the compounds with specificity. By comparison, the ExoN active site is more polar in nature, forming a wide shallow pocket. The polarity should lead to a larger proportion of hits that bind with high specificity, but the absence of a deep pocket could affect the binding affinity. The third target site, the active site of the N7-MTase domain of nsp14, also harbors a relatively large hydrophobic cavity but still has polar elements. The suitability of each of these nsp14 pockets to accommodate ligands is reflected in their mean docking scores (for the top 100 hits for each target site): -10.4 kcal/mol for the PPI site, -13.4 kcal/mol for the N7-MTase active site, and -9.8 kcal/mol for the ExoN active site.

2'-O-ribose methyltransferase (nsp16)

RNA cap modifications are known to play a role in the host cell's identification of self-RNA. For example, foreign RNA which lacks 2'-O methylation is inhibited by IFIT1 (Hyde et al., 2014). Pathogenic viruses such as CoVs that replicate in the cytoplasm have developed tactics to escape recognition by the host innate immune system. One such coronaviral mechanism is 2'-O-methyl capping of viral RNA by the nsp10-nsp16 complex. nsp16 is an m⁷GpppA-specific, SAM-dependent, 2'-O-MTase that must form a complex with nsp10 for its function (Decroly et al., 2011b; Chen and Guo, 2016). We therefore targeted the nsp10 binding interface on nsp16 (Figure S34).

nsp16 has a catalytic core comprising a β -sheet flanked by eleven α -helices (Figures S34 and S35). nsp10 has a small antiparallel β -sheet sandwiched between several α -helices and a large loop. The interaction surface of nsp10 with nsp16 is a mixture of hydrophobic, polar, and positively charged residues and helps in the stabilization of the SAM binding site (PDB: 6w75). However, since the interface is relatively flat and polar, it is challenging to find small molecules that bind with sufficient strength.

Earlier structural studies showed that stabilization of the SAM binding pocket by the nsp10-16 interaction also expands the RNA-binding pocket of nsp16 (Chen et al., 2011), and short peptides derived from the interaction interface have been shown to inhibit 2'-O-methyltransferase activity of the nsp10-16 complex (Min et al., 2012). Therefore, in addition to targeting the interaction interface of nsp10 and nsp16 (screen ID: 40, Figure S34), we also targeted the putative SAM binding site (screen ID: 41, Figure S35). SAM binding has been shown to be essential for nsp10-16 complex formation and catalytic function in another highly pathogenic CoV, MERS-CoV (Aouadi et al., 2016). The screen targeting the nsp10 PPI and the SAM binding site resulted in top 100 hits with average docking scores of -9.6 kcal/mol and -11.4 kcal/mol, respectively. Besides these target sites on nsp16, we targeted the nsp16 interaction interface on nsp10 using two conformations of the protein (screen IDs: 25 and 26; Figures S36 and S37).

Uridylate-specific endoribonuclease

nsp15 is a uridylate-specific endoribonuclease (NendoU) carrying a C-terminal catalytic EndoU domain that has been described as having different roles in immune evasion in different coronaviruses. A recent study in MHV, a murine CoV, showed that nsp15 was critical for evasion from host dsRNA sensors in macrophages (Deng et al., 2017), and in porcine deltacoronavirus, nsp15 was found to inhibit IFN- β production (Liu et al., 2019). The SARS-CoV-2 nsp15 monomer is composed of three distinct domains: (a) an N-terminal domain containing an antiparallel β -sheet and two α -helices, (b) a central domain composed of 10 β -strands and three short helices, and (c) a C-terminal NendoU domain composed of two antiparallel β -sheets and five α -helices (PDB: 6vww) (Kim et al., 2020). The SARS-CoV-2 nsp15 monomer is capable of forming dimeric, trimeric, and double-ring hexameric assemblies. It has been shown that the hexameric form is essential for enzymatic activity and that the hexamer is stabilized by interactions of the N-terminal oligomerization domain (Guarino et al., 2005). The NendoU active site itself is located in a groove between the two β -sheets, and conserved, crucial residues have been previously identified: His235, His250, Lys290, Thr341, Tyr343, and Ser294. With the aim of disrupting the endoribonuclease activity of nsp15, the active site was targeted in an *in silico* screen (screen ID: 39; Figure S38), and this resulted in a top 100 hits with an average docking score of -10.9 kcal/mol (Figure S8), which theoretically corresponds to nanomolar inhibitors; however, experimental validation is necessary to confirm this.

Deubiquitination and deISGylation activities of PL^{Pro}

PL^{Pro} (nsp3) cleaves at a consensus cleavage motif, LXGG, which is also the consensus sequence recognized by cellular deubiquitinating enzymes (DUBs). CoVs attenuate the host anti-viral response by deubiquitinating key components of the interferon-mediated immune response. While PL^{Pro} uses its catalytic site for its deubiquitination function, ubiquitin itself makes contacts with the “palm” and “fingers” regions of PL^{Pro} (Ratia et al., 2014). This protein-protein interface is well conserved in SARS-CoV-2 and could be targeted in order to inhibit the deubiquitination function of nsp3. SARS-CoV has also been shown to possess deISGylation activity in which ISG-15, a protein modifier consisting of two tandem ubiquitin-like domains involved in modulating the innate immune response, is cleaved (Lindner et al., 2007). We therefore targeted the proximal “S1” ubiquitin contact site of the “palm” of nsp3 (Figure S39), inhibition of which could effectively block both the deubiquitination, as well as deISGylation, activities of PL^{Pro} (screen ID: 14). The binding interface is moderately hydrophobic and concave, making it an attractive drug target. This particular screen resulted in the top 100 ranking compounds exhibiting an average docking score of -9.2 kcal/mol (Figure S8).

nsp3 macrodomain

The macrodomain, sometimes called the X domain, is a highly conserved region of ~ 180 amino acids that binds to ADP-ribose (Karras et al., 2005) and which has been shown to be dispensable in RNA replication, at least in the context of an RNA replicon (Kusov et al., 2015). However, it has also been shown to have possible roles in evading the host innate immune response (Kuri et al., 2011; Eriksson et al., 2008). Viral macrodomains are known to have two major enzymatic activities: (i) ADP-ribose 1''-phosphatase activity where the phosphate is removed from ADP-ribose 1''-phosphate, and (ii) ADP-ribose hydrolase activity where ADP-ribose is removed from mono- or poly-ADP-ribosylated proteins (Li et al., 2016).

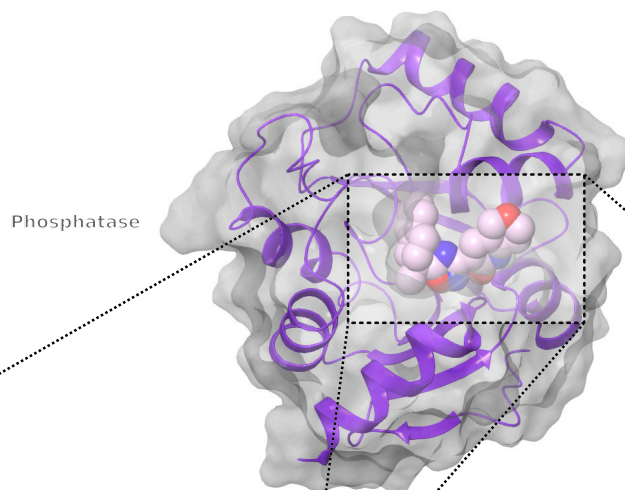
Previous biochemical studies showed that the nsp3 macrodomain is responsible for processing of ADP-ribose 1''-phosphate, which is a byproduct of pre-tRNA splicing (Snijder et al., 2003; Putics et al., 2005). This phosphatase activity has been found to be specific for ADP-ribose 1''-phosphate; however, the turnover was found to be very low (Putics et al., 2006; Singh Saikatendu et al., Joseph, 2005) with a k_{cat} of $1.7-20$ min⁻¹, making it unlikely that this activity has a strong direct effect on viral virulence.

ADP-ribosylation is a critical post-translational modification catalyzed mainly by poly (ADP-ribose) polymerases (PARPs) in which ADP-ribose is added to a protein (Michael, 2015). Several PARPs are known to be interferon-stimulated genes, and some PARPs have been shown to have antiviral activities. Studies in SARS-CoV and MERS-CoV have shown that these viruses were less virulent in the absence of the macrodomain and that this directly associated with changes in pro-inflammatory cytokine expression (Fehr et al., 2016; Fehr et al., 2014). Additional studies in HCoV-299E and SARS-CoV suggested that the nsp3 macrodomain may mediate the resistance to antiviral interferon responses (Kuri et al., 2011). Mutation of the macrodomain in SARS-CoV also induced a strong interferon and inflammatory cytokine response in the lungs of infected mice (Fehr et al., 2016). However, the exact mechanism linking the hydrolase/phosphatase activity of nsp3 and the observed cytokine production is unknown. The fact that (a) the macrodomain participates in reversing antiviral ADP-ribosylation (Li et al., 2016; Fehr et al., July 2018) and (b) mutation of a highly conserved asparagine residue abolishes phosphatase and hydrolase activities and mitigates viral virulence (Fehr et al., 2016; McPherson et al., 2017) suggests that targeting the ADP-ribose binding pocket could be an effective therapeutic strategy, particularly in combination with other targets. We observed two different conformations for the active site in the same X-ray structure, which indicated protein dynamics that could affect the ability of the docking to accurately recapitulate the solution-state structure. We, therefore, targeted both conformations of the active site of the macrodomain: one in which the active site is more closed (screen ID: 10; Figure 4) and a second conformation in which the active site is more open (screen ID: 11; Figure S40). It is also important to note that the active site itself is very hydrophobic, which increases the risk that the resulting hits could bind non-specifically to membranes and hydrophobic pockets in other proteins, and this should be accounted for in any follow-up screening. The top 100 ranking compounds from screens 10 and 11 have average docking scores of -10.8 kcal/mol and -12.7 kcal/mol, respectively (Figure S8).

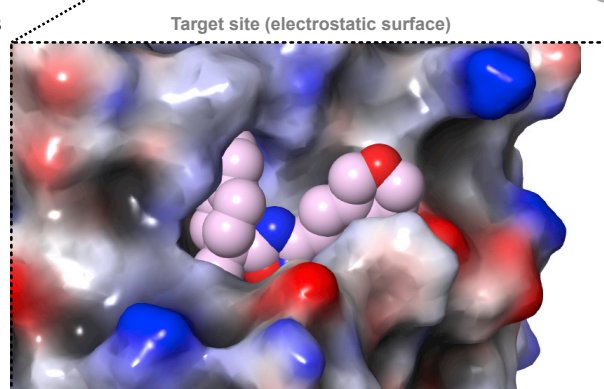
ORF7a

ORF7a is a non-essential accessory protein with a transmembrane helix at the C-terminus that has been shown to localize to the ER, Golgi, and cell surface (Nelson et al., 2005; Taylor et al., 2015). The assembly of ORF7a into viral particles suggests that the protein is important in the viral replication cycle and that it

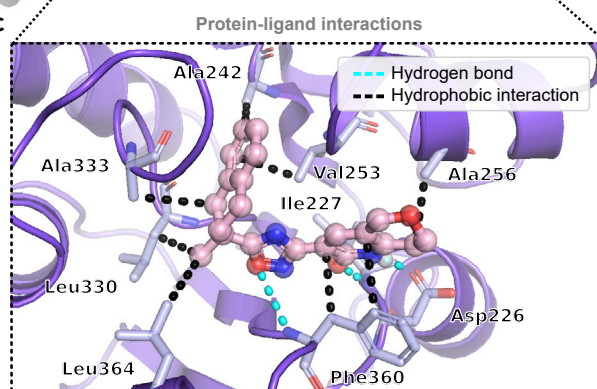
A



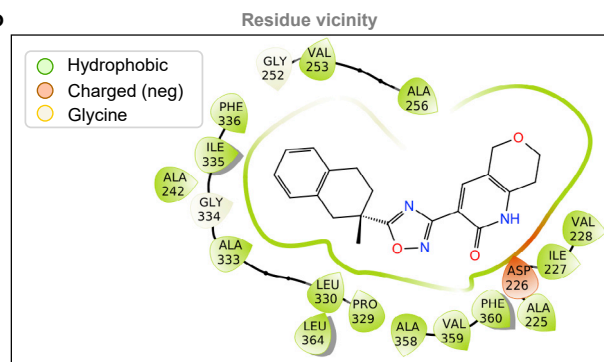
B



C



D



E

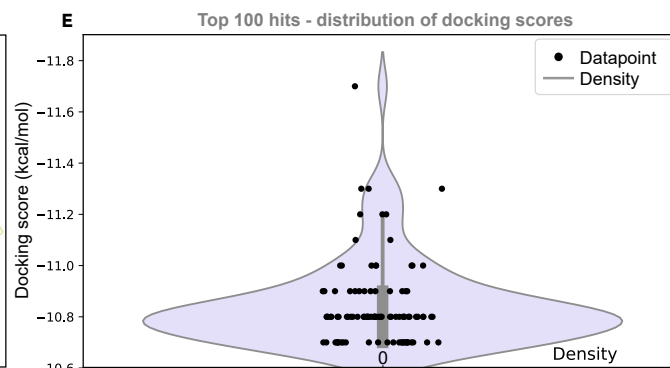


Figure 4. The phosphatase (closed conformation) and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site

- (A) The phosphatase (violet), which is part of nsp3, and an example compound (light pink) from the virtual screen bound to the active site (closed conformation) (screen ID: 10).
 (B) Electrostatic surface of the target protein to which an example compound (light pink) is bound.
 (C) An overview of the interactions between the inhibitor and the phosphatase structure.
 (D) Residues within 4 Å of the inhibitor.
 (E) Distribution of the docking scores of the top 100 virtual screening hits.

might have a function early in infection (Huang et al., 2006a). In SARS-CoV, ORF7a is known to interact with a “host restriction factor”, bone marrow stromal antigen 2 (BST-2) (Taylor et al., 2015), which inhibits viral replication by preventing virus budding from the plasma membrane (Sauter et al., 2010). ORF7a has been shown to localize to the ER when co-expressed with BST-2. Moreover, ORF7a interferes with the glycosylation of BST-2, which is essential for its tethering/restriction function (Taylor et al., 2015). There

are also indications that ORF7a may play a role in virus-induced apoptosis (Schaecher et al., 2007); however, further studies are needed to understand the mechanism of this role.

Structurally, ORF7a has an N-terminal signal peptide, an 80-amino-acid-long luminal domain, a transmembrane domain, and a short C-terminal cytoplasmic tail (Nelson et al., 2005). In the recent experimental structure of the SARS-CoV-2 ORF7a (PDB: 6w37 (Nelson et al., 2020)), the luminal domain displays structural characteristics similar to those of previously published ORF7a structures and is composed of a seven-stranded β -sandwich (Figure S41) with topological similarities to the Ig superfamily. Intriguingly, recent genomic analysis of a sample picked up during sentinel surveillance in the state of Arizona (USA) revealed a 27-amino-acid in-frame deletion in ORF7a, in a region corresponding to the N-terminal signal peptide (Holland et al., 2020). The functional consequences of this deletion, however, are not known. Since the molecular mechanism of the ORF7a interaction is largely unknown, we pursued a blind docking strategy for this protein, covering the entire protein structure (screen ID: 9; Figure S41). Interestingly, almost all top scoring compounds bound at a shallow cavity around Phe46, while a few bound to the opposite side near Leu31.

During experimental validation, it would need to be verified whether these hits can inhibit the interaction of ORF7a with small glutamine-rich tetratricopeptide repeat-containing protein (SGT) (Fielding et al., 2006) or lymphocyte-associated antigen 1 (LFA-1) (Hänel and Willbold, 2007), both of which are thought to play a role in virus-host interactions and host cell cycle modulation. The presence of ORF7a on the viral surface has been suggested to mean that LFA-1 could be a potential viral attachment factor in leukocytes (Huang et al., 2006a; Hänel and Willbold, 2007); however currently, insufficient data are available about the possible residues and structural dynamics that would be involved in this interaction. Moreover, the rather flat surface of ORF7a, devoid of pockets or clefts, makes it challenging to find small molecules with selective and potent binding.

Disrupting the viral assembly and packaging

As the viral life cycle progresses, the structural proteins (S, E, M) are also expressed. These structural proteins are initially inserted into the ER but then transit to the ER-Golgi intermediate compartment, where they are incorporated into mature virions (de Haan and Rottier, 2005). Nucleocapsid protein (N) forms large oligomeric complexes with the replicated viral genome. The resulting ribonucleoprotein (RNP) complexes associate with M, facilitating packaging of the genome into a complete virion assembly. With limited structural information available about M and E, our efforts to find compounds that can disrupt the viral assembly and packaging were centered on the N.

Nucleocapsid protein

N packages the viral RNA into a helically ordered RNP (Stohman et al., 1988) and plays a critical role in virion assembly by interacting with M. N is also known to play key roles in the regulation of viral RNA synthesis and modification of the host cell metabolism (Cong et al., 2019). N is composed of an NTD (Figures S42 and S43), an intrinsically disordered SR-rich linker, and a C-terminal dimerization domain (CTD, Figures S44 and S45). The published crystal structure of the RNA-binding domain reveals a β -sheet core sandwiched between two loops. The β -sheet core has five antiparallel β -strands plus a short helix and a longer β -hairpin between strands β 2 and β 5 (PDB: 6m3m) (Kang et al., 2020). The central β -strands of the β -sheet core contain a purine/pyrimidine monophosphate binding site, and this ribonucleotide binding mechanism is essential for virulence, making it an attractive drug docking site (Lin et al., 2014). The NTD is also expected to interact with the CTD during RNA packaging. We have therefore targeted two minimally overlapping regions of the NTD, which in combination cover the entire surface of the NTD (screen IDs: 42 and 43) (Figures S42 and S43). The first NTD site includes the ribonucleotide-binding site reported in (Lin et al., 2014).

The CTD of N aids dimerization of the protein, as well as RNA binding. The monomer consists of eight α -helices and a β -hairpin. The dimerization interface itself is composed of three helices (α 5, α 6, and α 7) and one β -hairpin from each monomer and involves a combination of hydrogen bonds and hydrophobic interactions, which result in a very stable dimer (PDB: 6wji). The fact that the C-terminal domain has also been found to bind RNA (Hsieh et al., 2005) highlights the critical role of this domain in the overall helical packaging of the viral genome (Chen et al., 2007). We targeted the dimerization interface of the CTD by removing one monomer from the crystallized dimer structure (screen ID: 44) (Figure S44), and we also screened the dimerization interface via a blind docking procedure against the entire dimer surface (screen

ID: 45) (Figure S45). The goal of these screens was to identify a small molecule that will bind to the monomer at the dimer interface and disrupt the formation of the nucleocapsid. However, the structure of the monomer alone has not yet been solved, and given the extent of the interaction at the dimer interface, we anticipate considerable dynamics in the monomer. Therefore, the structure extracted from the dimer may not adequately describe the monomer structure, yielding imperfect hits. Another challenge of this target is that each virion has ~65 spike trimers and with a ratio of S trimer to N of 1:4 (Neuman et al., 2006; Beniac et al., 2006), this amounts to ~260 monomers or ~130 dimer interfaces. The abundance of contacts that would need to be disrupted means that a higher concentration of the inhibitor would be required for efficacy. Despite these reservations, it is worth noting that the HIV nucleocapsid inhibitor lenacapvir (GS-6207/GS-CA1), which targets the dimerization interface of the capsid protein has shown significant efficacy and has made it to clinical trials recently (Yant et al., 2019; Daar et al., 2019; Sager et al., 2019). The screens targeting the RNA-binding site (screen ID: 42) and the N-terminal oligomerization site (screen ID: 43) resulted in hits with an average docking score of -10.3 kcal/mol and -9.78 kcal/mol, respectively. The screens targeting the C-terminal dimerization interface (screen ID: 44) and the oligomerization site (screen ID: 45) resulted in average docking scores of -12.7 and -10.4 kcal/mol, respectively (Figure S8).

DISCUSSION

We screened against 17 proteins, 15 proteins of SARS-CoV-2 and two human proteins, targeting a total of 40 sites across these proteins in 45 screens with approximately 1.1 billion molecules per screen. Due to the unprecedented global situation, we wanted to make the screening results available to researchers prior to experimental validation in the hope that others may also be able to take advantage of our large multi-target data set. The average docking score of the top 100 filtered hits as ranked by the docking score for each target site can be seen in Figure S8. The top 1000 hits per virtual screen, along with additional information such as the docking scores, molecular properties, and docking poses, can be found online at <https://vf4covid19.hms.harvard.edu/>. In addition, the top million hits for each target site are available for download from the same website in DataWarrior format and from Mendeley Data (see the Resource Availability section for more details).

As discussed earlier, not all the functional surfaces that were targeted had similar potential to accommodate a small molecule. The targeted site that yielded the weakest virtual hit compounds was the first (site 1) of the two targeted sites located on the spike interaction interface of ACE2. The reason for the relatively weak predicted binding affinities of the top scoring compounds is the flat and polar nature of this target site. These characteristics make it very challenging for small molecules to bind sufficiently well to the surface so as to disrupt the interaction with spike protein. The targeted site that yielded the virtual hits with the highest predicted binding affinities was the active site of the N7-MT domain of nsp14. The high affinities are likely a product of the remarkably deep, buried pocket of this active site and the relatively high number of hydrophobic sites within that pocket. If we consider the top 1000 filtered hits, the weakest screen (ACE2, screen ID: 2) had a mean docking score of -7.6 kcal/mol (2.3 μ M) while nsp14 (screen ID: 37) had the best, with a mean docking score of -12.9 kcal/mol (0.3 nM). Although the trend was the same, the mean docking scores were better for the top 100 hits: -8.2 kcal/mol (1.1 μ M) for ACE2 (screen ID: 2) and -13.4 kcal/mol (0.14 nM) for nsp14 (screen ID: 37). Finally, as expected, the mean docking scores were even better for the top 10 hits: -8.5 kcal/mol (0.5 μ M) for ACE2 (screen ID: 2) and -14 kcal/mol (0.03 nM) for nsp14 (screen ID: 37). The mean docking scores of the top 100 hits for each screen are shown in Figure S8. Out of the 45 virtual screens, 41 had a mean docking score for the top 1000 compounds higher than -8.2 kcal/mol. These scores indicate that sub-micromolar binders should be identifiable from the hits for each target site.

The hits derived from the rigid docking procedures described here could be further improved for potency by a second-stage screen, in which the target proteins are allowed to be flexible. This could be achieved either by ensemble docking or allowing flexible side chains, for example, by using GWOVina (Wong et al., 2020; Gorgulla et al., 2020b). Docking results could also be improved in second-stage screens by using alternative docking programs, such as AutoDock Vina, Smina, or Gnina (Trott and Olson, 2010; Koes et al., 2013; Ragoza et al., 2017). For special types of ligands such as compounds containing carbohydrates or halogens, specialized docking programs such as VinaXB or Vina-Carb could be utilized to further increase the accuracy of the results (Koebel et al., 2016; Nivedha et al., 2016).

Small-molecule drug candidates that have been previously pre-approved by the FDA or other regulatory bodies for other indications and molecules that have been vetted for use in humans including preclinical

candidates, nutraceuticals, metabolites, and select natural products are particularly desirable hits in any inhibitor screen (Kiplin Guy et al., 2020). This is because these compounds can be fast-tracked to clinical trials on the basis of previous data. Here, we mined for such compounds in the top 1% of our hits from the “in stock” compounds of the ZINC library. We found 161 drugs (with 491 occurrences across all the screens) that are “world approved” in our top 1%. We further filtered these drug candidates to only those with a docking score better than -8 kcal/mol and show these 80 approved drugs along with their screening target and docking score in Table S46. Of these 80 drugs, 16 of them are being considered in clinical trials for COVID-19. The complete list of “world approved” hits along with their binding targets and docking scores is available at <https://vf4covid19.hms.harvard.edu/world-approved-drugs>.

Several hits from the “world approved” drugs are from the steroid family. Corticosteroids such as betamethasone, ciclesonide, flumetasone, and meprednisone acetate are already being investigated as COVID-19 treatments (Recovery Collaborative Group et al., 2020; Iwabuchi et al., 2020; Wang et al., 2020b). For example, betamethasone was identified as a hit for nsp14, the bifunctional enzyme that acts as a methyltransferase and an exoribonuclease. Betamethasone bound in the simulation to the active site of the exoribonuclease of nsp14 (Figure S47). While the usefulness of corticosteroids is largely attributed to their anti-inflammatory effects in the host, our screening results suggest that additional mechanisms of action that involve viral proteins may be possible. Other steroid hormones in clinical trial such as estradiol or dutasteride also came up as hits in our screens.

Of the non-steroid approved drug hits, tyrosine kinase inhibitors occupy a significant fraction. Drugs such as imatinib (Glivec) are being tested in clinical trials due to their reported *in vitro* efficacy against SARS-CoV and MERS-CoV (Coleman et al., 2016; Dyall et al., 2014); however, their mode of action is not yet fully understood. In our screens, imatinib binds to two targets: (i) nsp14, at the active site of the methyltransferase, and (ii) the RNA-binding interface of the helicase, nsp13 (Figure S47).

Of the 3897 “investigational” drugs in the ZINC 15 database, 137 were hits (500 occurrences), and of the 101378 “in-man” compounds, 401 were hits (1061 occurrences). A full list of these molecules, along with their respective target sites and docking scores, is available at (<https://vf4covid19.hms.harvard.edu/investigational-drugs>). Multiple occurrences of a hit molecule partly stem from the different docking scenarios that we used for the same target site. In addition, some of the compounds may be represented more than once in the list because of differences in tautomeric, isomeric, or protonation states, any of which earns the molecule a unique ID in the ZINC database. Despite the abundance of approved hits, some of which are already involved in or related to drugs in clinical trials, experimental confirmation of the specificity of hits from our screen against SARS-CoV-2 proteins is still needed to determine an antiviral mechanism of action.

Comparison to previously identified inhibitors

SARS-CoV-2 is not the first CoV to cause serious disease in the human population, and these other CoVs (SARS-CoV and MERS-CoV) have also been the subject of research into small-molecule inhibitors. There have been several efforts in the past, both experimental and computational, to target specific viral proteins from these viruses. The proteins that were most often targeted in these efforts were the proteases, M^{Pro} (Jin et al., 2020a; Park et al., 2012; Wang et al., 2017; Yang et al., 2005) and PL^{Pro} (Park et al., 2012; Ghosh et al., 2010; Chou et al., 2008; Cheng et al., 2015), the polymerase (Agostini et al., 2019; Sheahan et al., 2017), the spike protein (Xia et al., 2019; Liao et al., 2015; Kao et al., 2004), the endoribonuclease (Ortiz-Alcantara et al., 2010), and the helicase (Lee et al., 2017; Cho et al., 2015; Alnazawi et al., 2017). Although direct binding affinities were not available for most of the molecules, the inhibitory potential, IC_{50} , or the effective potential, EC_{50} , were reported.

In order to determine if previously published SARS-CoV and MERS-CoV hits for which an exact binding site had been reported were also top hits against SARS-CoV-2, we calculated docking scores using the same computational methods used in the screens described here, allowing direct comparison with our hits. One of the best inhibitors for SARS-CoV M^{Pro} is an asymmetric aromatic disulfide molecule with an IC_{50} of 0.52 ± 0.06 μ M (Wang et al., 2017). When we docked this molecule against SARS-CoV-2 M^{Pro} , we got a docking score of -7.7 kcal/mol, whereas the top 100 hits from two of our recent screens described here, which targeted the active site of M^{Pro} (screen IDs: 16 and 17), resulted in mean docking scores of -10.8 and -11.4 kcal/mol (Figure S8). While promising, this apparent improvement in binding affinity

must still be validated experimentally. Similarly, in the case of PL^{Pro}, the best previously described active site inhibitor (Ghosh *et al.*, 2010) had a binding free energy of -7.2 kcal/mol when re-screened, while the top 100 compounds from the two corresponding VirtualFlow screens (screen IDs: 12 and 15) have mean binding scores of -8.9 and -10.0 kcal/mol, respectively (Figure S8). Lee, *et al.* described a novel helicase inhibitor that can block both the dSDNA unwinding and ATP hydrolysis activities of the SARS-CoV helicase (Lee *et al.*, 2017). Docking of this molecule to the SARS-CoV-2 helicase gave a score of -7.9 kcal/mol. We also targeted the active site of the SARS-CoV-2 helicase (screen ID: 33) in our own study and that screen resulted in a top 100 hits with a mean docking score of -10.2 kcal/mol. There have been multiple attempts to develop small molecule inhibitors that target the active site of the RdRp, considered a promising target, from both SARS-CoV and MERS-CoV (Agostini *et al.*, 2019; Sheahan *et al.*, June 2017). Docking of the best molecules from these previously published studies against the SARS-CoV-2 RdRp resulted in docking scores of -6.1 and -6.8 kcal/mol, while our screens (screen IDs: 28 and 29) resulted in compounds with mean docking scores of -11.1 and -9.7 kcal/mol, respectively, for the top 100 hits. It must be noted that, as stated before, the comparisons provided here are based on docking scores, and therefore any apparent improvements in binding still need to be experimentally validated.

Multiple attempts have already been made to employ virtual screening of varying library sizes to identify approved drugs, as well as discover new molecules, that can effectively inhibit SARS-CoV-2 viral infection (Jin *et al.*, 2020a; Strodel *et al.*, 2020; Zhou *et al.*, 2020c; Naik *et al.*, 2020; Fischer *et al.*, 2020; Ton *et al.*, 2020; Panda *et al.*, 2020; Yu *et al.*, 2020; Singh and Florez, 2020; Kadioglu *et al.*, 2020), with more than one screen is focused solely on M^{Pro} (Strodel *et al.*, 2020; Fischer *et al.*, 2020; Yu *et al.*, 2020; Singh and Florez, 2020). It is important to note that in the case of M^{Pro}, the conformational plasticity of the active site means that screening against a single conformation, as is done in these studies, may be less effective at identifying functional hits. One work of note is that by Ton, *et al.*, in which 1.3 billion compounds from the ZINC 15 library were screened against M^{Pro} using a Deep Docking algorithm (Ton *et al.*, 2020). While Deep Docking significantly accelerates the screening of ultra-large compound libraries, it relies heavily on using the docking scores of a small random subset of compounds to eliminate a large fraction of the compounds early in the docking workflow, and this can be detrimental to finding an absolute docking score for individual compounds.

A recent mass-spectrometry-based study has identified several host proteins that interact with SARS-CoV-2 viral proteins, leading to the identification of multiple druggable host targets and a series of known modulating compounds for these targets that were then shown to have antiviral activity in cell-culture-based assays (Gordon *et al.*, 2020). The efficacy of these already approved host-directed compounds lends support to the development of these or similar drugs as therapeutic interventions for SARS-CoV-2 infection. Coincidentally, some of the compounds from this mass spectrometry study were also identified as hits in our screens against diverse viral proteins. For example, "nafamostat", an approved anticoagulant identified in Gordon, *et al.* as having a possible impact on cell entry, was also identified in our screens as a hit with good scores against three viral proteins: ORF7a, PL^{Pro}, and nucleoprotein. This compound is already undergoing clinical trials for COVID-19 (Ref:NCT04352400, NCT04418128). Similarly, "progesterone", a common birth control drug that is also known to have anti-inflammatory properties, was identified in the proteomics study as a ligand with antiviral activity that targets sigma-1 or sigma-2 receptors. In our screens, this compound was identified as an inhibitor for nsp16, nsp7, and spike. Progesterone is also in clinical trials for the treatment of COVID-19 (NCT04365127). Silmitasertib (CX-4945), an inhibitor of casein kinase II, was identified both in the proteomics study as well as in another recent study that mapped the host phosphorylation landscape in the context of SARS-CoV-2 infection (Gordon *et al.*, 2020; Bouhaddou *et al.*, 2020). Silmitasertib was originally developed to be used in combination with chemotherapy. In our screen, silmitasertib showed potential binding to nsp16. If some of these host-protein-targeted compounds can also directly inhibit some viral proteins, as suggested by the docking scores derived from our screens, this could form the basis for a unique regimen of multi-target drugs to treat COVID-19 that acts on both viral as well as host proteins simultaneously. A list of compounds that are common to both the interactome study and our virtual screens, along with their potential targets, is provided in Table S48.

To verify if the docking scenarios and procedures that were used in this study worked in principle, we carried out two types of evaluation studies with compounds that were previously identified as inhibitors and/or binders. These verification studies should be interpreted on a qualitative level as they are not able to recapitulate the unique advantages of the ultra-large scale of the screens presented here. First, we completed

re-dockings of ligand-protein complexes for the macrodomain, M^{Pro}, PL^{Pro}, nsp15, and nsp16. Here, the similarity of the binding geometry of the ligand, derived from re-dock, is compared with that of the available high-resolution experimental structure for the target protein and ligand. Approximately, 35% of the 29 compounds that we re-docked have an RMSD below 2.5 Å when compared to the experimental structure. As a relative benchmark of our performance, in the original AutoDock Vina publication, an average re-docking success rate of around 50% was reported when using rigid receptors (Trott and Olson, 2010). However, the docking accuracy in the aforementioned publication was set to a much higher value (i.e., exhaustiveness parameter set to 8), while we used QuickVina 2 in its fastest setting (exhaustiveness of 1) in our study due to the computational costs of screening a billion compounds per target. The key idea of our approach is to use the scale of the screen to increase the quality of the resulting top scoring compounds (Lyu et al., 2019; Gorgulla et al., 2020a; Gorgulla, 2018). One of the successfully re-docked compounds for nsp16 can be found in Figure S49.

In addition, we carried out enrichment studies for M^{Pro} and PL^{Pro}, as these were the only two targets with sufficient non-covalent binders reported. Enrichment studies assess the ability of the docking routines to re-identify and enrich previously identified binders when mixed with a pool of random compounds or decoys. Two different flavors of enrichment studies were performed: (i) using experimentally identified molecules and (ii) using the virtual hits identified in this study. The Directory of Useful Decoys, Enhanced (DUD-E) was used for generating decoy molecules based on the active compounds (Huang et al., 2006b; Mysinger et al., 2012). The complete details of these studies can be found in the Supplemental Information under "Enrichment and re-docking verification studies". The experimentally active compounds that were used in the enrichment studies are listed in Table S50, and the virtual hits that were used as active compounds are listed in Table S51. The enrichment factors at 2%, EF₂, that were obtained for the experimental binders are 20 and 25 for M^{Pro} and PL^{Pro}, respectively. The maximum possible EF₂ is 50 for each of the two target proteins. The results obtained here are typical of similar docking procedures (Huang et al., 2006b; Mysinger et al., 2012). When the enrichment studies were performed with ten virtual hits, the enrichment factors at 2% are equal to the theoretical maximum value of 50, which reflects the effect of ultra-large-scale screening.

Our ultra-large library size and the multi-pronged screening approach we present here is not only unprecedented in terms of scale, significantly enhancing the chances of finding an effective antiviral, but also provides a pool of potential antiviral compounds that could be developed further as pan-CoV drugs, as well as combination therapies. Combination regimens of reverse transcriptase inhibitors and protease inhibitors have been found to be effective as anti-retroviral therapy for HIV and have led to a dramatic improvement in HIV-infected patient health and survival worldwide (Vella et al., 2012; Kuritzkes et al., April 1999; Palella et al., 1998). Clinical studies have also shown that in cases of severe influenza, a combination of multiple antiviral drugs is much more effective than a monotherapy (Dunning et al., 2014; Hung et al., 2017). Therefore, it could also be beneficial to develop a combination therapy to treat severely affected patients with COVID-19. Considering that vaccine deployment will take time and that the long-term efficacy of any SARS-CoV-2 vaccines is still unknown, combination therapeutic strategies for the control and treatment of CoV infections should continue to be developed in parallel.

The mutating nature of CoVs and the potential for the emergence of novel CoVs in the future make the development of broad spectrum antivirals desirable. For a protein to qualify as a candidate target for pan-CoV drug development, it should be (i) critical for the viral replication or virulence, and (ii) highly conserved across all known CoVs and considered likely to be conserved in emerging CoVs. While the structural proteins (S, N, M, and E) are not as well conserved, the accessory proteins are conserved to an even lesser extent across different CoV species (ProViz (Peter et al., 2016)). Highly conserved non-structural proteins and conserved host proteins would therefore be ideal targets for the development of broad spectrum CoV drugs (Allison and Bavari, 2019). While studies (Benvenuto et al., 2020) have suggested that the S and N genes of SARS-CoV-2 are undergoing episodic selection during human transmission, recent analyses of genome samples from across the world suggest that the greatest genetic diversity is occurring in S, ORF3a, and ORF8 (James, 2020). Structural genomic analyses across multiple virus classes, including SARS-CoV-2, have shown that intra-viral interactions are more conserved than viral host interactions (Srinivasan et al., 2020; Warren et al., 2013; Wang et al., 2016; Voitenko et al., 2016). The PPI interfaces involved in the heteromeric RNA polymerase complex (composed of nsp7, nsp8, and nsp12) are evolutionarily conserved (Srinivasan et al., 2020), and our top hits against these crucial PPIs represent promising candidates for the development of pan-CoV inhibitors. Similarly, structural genomic analysis of the methyltransferase complexes (nsp10-nsp14 and nsp10-nsp16) also reveals PPIs that are fully

conserved with mutations limited to the surface (Srinivasan et al., 2020). Moreover, the substrate binding regions we used for PL^{pro}, M^{pro}, nsp14, nsp16, and the RNA binding site of the N have all been shown to be highly conserved (Voitenko et al., 2016), and the correspondingly identified small molecules could also potentially be developed as pan-CoV therapeutics. Since most human CoVs differ in their use of host cell receptors and priming proteases (Millet and Whittaker, 2015), identifying a cell-based assay or functional animal model for developing pan-CoV therapeutics could be challenging. The use of normal human bronchial epithelia cells as a potential universal screening platform for multiple classes of CoVs shows promise (Pyrce et al., 2010; Chan et al., 2014; Sheahan et al., 2020), and the general need for high-throughput *ex vivo* testing models will hopefully drive development in this area.

It is important to understand that the work described here is only the first step in the drug discovery process. Virtual screening seeks to use structure and modeling-guided approaches to identify higher quality starting hits. Hits from these screens still need to be experimentally validated for binding, inhibition, and activity in the context of viral infection. Some of these targets, which affect viral immune evasion rather than replication, may need more complex evaluation in cell culture or in animal models to determine efficacy. While our own experimental validation of hits from these screens has begun, we believe that making these results available to a broader audience for experimentation is important in light of the current pandemic.

The alarming global spread of SARS-CoV-2 and the dearth of effective prophylaxis or treatment methods have resulted in a pressing need for accelerated drug discovery. Here, we make use of the recent chemical space revolution that resulted in ultra-large make-on-demand libraries to prioritize lead molecules and minimize the transition time from discovery to clinic. We have used the available-to-date SARS-CoV-2 structure to carry out detailed structure-based virtual screening against 17 different target proteins, frequently at multiple target sites, by leveraging the versatility of our recently developed *in silico* screening platform, VirtualFlow. The set of small molecule hits that we present here can aid drug development happening at multiple institutions worldwide, and the set of prospective drug molecules derived from this ultra-large multi-target screening campaign will potentially contribute to the building of a future arsenal of anti-coronaviral drugs, ready to tackle future outbreaks.

Limitations of the study

We present here results from our ultra-large-scale computational effort, and the hits have not been experimentally validated for their efficacy. Given the multi-pronged approach targeting 40 different sites on 17 proteins, it is a challenge for any one research group to validate all the targets in the timely manner that the urgency of the current pandemic warrants. We hope that the results provided here can be leveraged collectively by the research community to identify potent inhibitors of SARS-CoV-2. In addition, we have also discussed the potential and plausible challenges, if any, for each target site from a docking perspective that researchers should bear in mind while validating hits.

When performing *in silico* screening, we generally choose a 3D surface on the target protein, referred to as the “docking box”, and evaluate the energetics of a molecule bound to that chosen surface. The docking box is judiciously chosen to target a known functional domain, either an active site, an allosteric site, or a PPI interface. In practice, the docking box is typically extended to accommodate secondary interactions with the small molecule, especially when competing with a substrate. However, occasionally, this results in identifying a tight binder that does not harbor any inhibitory function. It is important to take into account this possibility during follow-up experiments and test for binding using a direct binding assay such as isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), microscale thermophoresis (MST), or nuclear magnetic resonance (NMR). In particular, NMR and X-ray crystallography structures would provide detailed information on the binding site and how the small molecule engages the target. It could still be possible to functionalize these non-inhibitory binders with warheads to degrade the targeted protein by established methods, referred to as proteolysis-targeting chimeras (PROTACs) (Lai & Crews, 2016; Morgan et al., 2017). However, the effectiveness of these PROTACs would depend on the concentration, rate of turnover, and accessibility of the target viral proteins.

As noted before, when testing these compounds in viral assays, one should be cognizant of the nature of the target and the possible differential effects on viral viability versus viral virulence. It may also be useful to correlate “pre-approved” hits from the virtual screens with patient data, for example, to determine if patients’ regular medication and supplement regimens reveal any hits that positively correlate with either

a milder disease course or a faster time to recovery. This could help narrow down targets and provide justification for a more detailed and time-consuming experimental analysis of a hit.

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Dr. Haribabu Arthanari, hari@hms.harvard.edu.

Material availability

This study did not generate nor use any new or unique reagents.

Data and code availability

The virtual screening results and special subsets (e.g., the hits which are natural products) are available on the project homepage at <https://vf4covid19.hms.harvard.edu/>. The homepage includes the top 200 hits (partially processed/filtered) in csv format, the top ~1000 hits (partially processed/filtered) in an interactive table with 3D visualization capabilities, as well as the top ~1 million hits in csv format (raw data, unprocessed/unfiltered). The top 1 million hits (raw data, unprocessed/unfiltered) for each screen has been deposited on Mendeley Data, available at <https://doi.org/10.17632/knnpnsdpyn.1>. VirtualFlow, the open-source drug discovery platform which was used in this work for the virtual screenings, is available on GitHub at <https://github.com/VirtualFlow> and via the primary homepage of the project at <https://virtual-flow.org/>.

METHODS

All methods can be found in the accompanying [Transparent methods supplemental file](#).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.102021>.

ACKNOWLEDGMENTS

This work was supported by a Google Cloud COVID-19 Research Grant, and we thank Anna-Bettis Padgett for help with formal issues related to the grant and appropriate attributions of credit. We thank Ankita Chaudhary for help with curating the data. We thank Brigitte Klein for proofreading the manuscript. We would like to thank Arthur Jaffe, Mareike Weller and Meng Zhang for their support. We thank Jerry Parks, Jeremy Smith, and Jerome Baudry for helpful discussions. Funding for this project was provided by the Fondation Aclon to T.D.H. H.A. acknowledges funding from NIH (GM136859) and the Claudia Adams Barr Program for Innovative Cancer Research. K.F. acknowledges the support of Math+ Berlin Mathematics Research Center. This research was supported in part by grant TRT 0159 from the Templeton Religion Trust and by ARO Grant W911NF1910302 to Arthur Jaffe. JJP and RAD were supported in part by NIH grants P01AI120943, 5R01AI128364-02 and MassCPR. The authors also acknowledge the fellowship for K.M.P. from Max Kade Foundation, New York, and the Austrian Academy of Sciences. We thank ChemAxon for a free academic license for the JChem package. P.D.F. would like to acknowledge the Chleck Family Foundation Inc for funding.

AUTHOR CONTRIBUTIONS

C.G. and H.A. conceived and managed the project. C.G. carried out the *in silico* screens on the GCP, analyzed, and curated the data. K.M.P.D., K.E.L., Z.F.W., C.G., and H.A. selected the target proteins and the target sites with help from the other authors. G.T., S.P., R.Y., and J.K. assisted C.G. with the execution of the ultra-large-scale virtual screens on the GCP. M.G., A.C., A.R., H.A., and C.G. designed and implemented the web interface. A.S., N.L., C.H., E.Y., D.P., and R.L., as part of the Elastifile team at Google, worked with C.G. on the used cluster file systems. M.C., V.D., and C.C.G. prepared some of the protein conformations used in this study. T.D.H. prepared the structure 6m0k* of M^{P19}. H.D.H. advised on the choice of target sites on the spike protein and ACE2. P.D.F. helped with filtering the hits and identifying experimentally validated hits for comparison. D.R., Y.M., A.P., I.Y., I.D., O.T., and A.C. helped with evaluating, preparing, and visualizing the hit compounds in the manuscript. J.J.P., A.M.N., M.N.N., R.A.D., and G.W. advised and contributed to vital discussions on the choice of targets and target sites. C.G. and P.D.F. performed the enrichment studies and redocking analysis with input from H.A. C.G., K.E.L., K.M.P.D., K.F., T.D.H., and H.A.

wrote the manuscript with input from all the authors. All the authors edited the manuscript to its final form. G.W., K.F., and H.A. supervised the project.

DECLARATION OF INTERESTS

M.C. and V.D. report working for Innophore. C.C.G. reports being a shareholder and CEO of Innophore, an enzyme and drug discovery company. G.W. and C.G. report being co-founders of the company Virtual Discovery, Inc., which provides virtual screening services. G.W. reports serving as the director of this company. G.W. reports being a co-founder of PIC Therapeutics, Cellmig Biolabs, and Skinap Therapeutics. H.A. reports being an equity holder in PIC Therapeutics. I.I. and D.R. report working for Enamine, a company that is involved in the synthesis and distribution of chemical building blocks, fragments, and screening compounds. Y.M. reports being a scientific advisor for Enamine. Y.M., O.T., and A.P. report working for Chem-space, a company that is involved in the distribution of chemical building blocks, fragments, and screening compounds. I.D. reports working for UkrOrgSyntez Ltd. (UORSY), a company that is involved in the synthesis of chemical building blocks, fragments, and screening compounds. G.T., S.P., A.S., M.G., N.L., C.H., E.Y., R.L., R.Y., D.P., and J.K. report working for Google, a company that also provides cloud computing services. The research described here is scientifically and financially independent of the efforts in any of the above mentioned companies.

Received: August 3, 2020

Revised: October 28, 2020

Accepted: December 29, 2020

Published: February 19, 2021

REFERENCES

- Adedeji, A.O., Marchand, B., te Velhuis, A.J.W., Snijder, E.J., Weiss, S., Eoff, R.L., Singh, K., and Sarafianos, S.G. (2012). Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLoS One* 7, e36521, <https://doi.org/10.1371/journal.pone.0036521>.
- Afar, D.E., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B., and Jakobovits, A. (2001). Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res.* 61, 1686–1692.
- Agostini, M.L., Andres, E.L., Sims, A.C., Graham, R.L., Sheahan, T.P., Lu, X., Smith, E.C., Case, J.B., Feng, J.Y., Jordan, R., et al. (2018). Coronavirus susceptibility to the antiviral Remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. *mBio* 9, <https://doi.org/10.1128/mBio.00221-18>.
- Agostini, M.L., Puijssers, A.J., Chappell, J.D., Gribble, J., Lu, X., Erica, L.A., Bluemling, G.R., Lockwood, M.A., Sheahan, T.P., Sims, A.C., Natchus, M.G., et al. (2019). Small-molecule antiviral β -d-n4-hydroxycytidine inhibits a proofreading-intact coronavirus with a high genetic barrier to resistance. *J. Virol.* 93, <https://doi.org/10.1128/jvi.01348-19>.
- Allison, L.T., and Bavari, S. (2019). Broad-spectrum coronavirus antiviral drug discovery. *Expert Opin. Drug Discov.* 14, 397–412, <https://doi.org/10.1080/17460441.2019.1581171>.
- Almedia, J.D., Berry, D.M., Cunningham, C.H., Hamre, D., Hofstad, M.S., Mallucci, L., McIntosh, K., and Tyrell, D.A.J. (1968). Coronaviruses. *Nature* 220, 650.
- Anand, K. (2003). Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300, 1763–1767, <https://doi.org/10.1126/science.1085658>.
- Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., and Garry, R.F. (2020). The proximal origin of SARS-CoV-2. *Nat. Med.* 26, 450–452, <https://doi.org/10.1038/s41591-020-0820-9>.
- Antrobus, R.D., Coughlan, L., Berthoud, T.K., Dicks, M.D., Hill, A.V.S., Lambe, T., and Gilbert, S.C. (2014). Clinical assessment of a novel recombinant simian adenovirus ChAdOx1 as a vectored vaccine expressing conserved influenza A antigens. *Mol. Ther.* 22, 668–674, <https://doi.org/10.1038/mt.2013.284>.
- Aouadi, W., Blanjoie, A., Vasseur, J.-J., Debart, F., Canard, B., and Decroly, E. (2016). Binding of the methyl donor S-adenosyl-L-methionine to Middle East respiratory syndrome coronavirus 2'-O-methyltransferase nsp16 promotes recruitment of the allosteric activator nsp10. *J. Virol.* 91, <https://doi.org/10.1128/jvi.02217-16>.
- Báez-Santos, Y.M., St John, S.E., and Mesecar, A.D. (2015). The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res.* 115, 21–38, <https://doi.org/10.1016/j.antiviral.2014.12.015>.
- Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Uhl, S., Hoagland, D., Møller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* 181, 1036–1045.e9, <https://doi.org/10.1016/j.cell.2020.04.026>.
- Barbalat, R., Ewald, S.E., Mouchess, M.L., and Barton, G.M. (2011). Nucleic acid recognition by the innate immune system. *Annu. Rev. Immunol.* 29, 185–214, <https://doi.org/10.1146/annurev-immunol-031210-101340>.
- Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A.D., and Baker, S.C. (2005). The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *J. Virol.* 79, 15189–15198, <https://doi.org/10.1128/jvi.79.24.15189-15198.2005>.
- Beigel, J.H., Tomashek, K.M., Dodd, L.E., Mehta, A.K., Zingman, B.S., Kalil, A.C., Hohmann, E., Chu, H.Y., Luetkemeyer, A., Kline, S., Lopez de Castilla, D., et al. (2020). Remdesivir for the treatment of Covid-19—final report. *N. Engl. J. Med.* 383, 1813–1826, <https://doi.org/10.1056/NEJMoa2007764>.
- Beniac, D.R., Andonov, A., Grudski, E., and Booth, T.F. (2006). Architecture of the SARS coronavirus prefusion spike. *Nat. Struct. Mol. Biol.* 13, 751–752, <https://doi.org/10.1038/nsmb1123>.
- Benvenuto, D., Giovanetti, M., Ciccozzi, A., Spoto, S., Angeletti, S., and Ciccozzi, M. (2020). The 2019-new coronavirus epidemic: Evidence for virus evolution. *J. Med. Virol.* 92, 455–459, <https://doi.org/10.1002/jmv.25688>.
- 7bgwu. (n.d.). SWISS-MODEL Repository, ExoNSARSCoV2, <https://swissmodel.expasy.org/interactive/7BgWuu/>.
- Bienert, S., Waterhouse, A., de Beer, T.A., Tauriello, G., Studer, G., Bordoli, L., and Schwede, T. (2016). The SWISS-MODEL repository—new features and functionality. *Nucleic Acids Res.* 45, D313–D319, <https://doi.org/10.1093/nar/gkw1132>.
- Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezeli, V.V., Correa Marrero, M., Polacco, B.J., Melnyk, J.E., Ulferts, S., Kaake, R.M., Batra, J., et al. (2020). The global phosphorylation landscape of SARS-CoV-2 infection. *Cell*. <https://doi.org/10.1016/j.cell.2020.06.034>.

- Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012). 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* 109, 9372–9377, <https://doi.org/10.1073/pnas.1201130109>.
- Bzówka, M., Mitusińska, K., Raczynska, A., Samol, A., Tuszyński, J.A., and Góra, A. (2020). Structural and evolutionary analysis indicate that the SARS-CoV-2 mpro is a challenging target for small-molecule inhibitor design. *Int. J. Mol. Sci.* 21, 3099, <https://doi.org/10.3390/ijms21093099>.
- Chan, R.W.Y., Hemida, M.G., Kayali, G., Chu, D.K.W., Poon, L.L.M., Alnaeem, A., Ali, M.A., Tao, K.P., Ng, H.Y., Chan, M.C.W., Guan, Yi, et al. (2014). Tropism and replication of middle east respiratory syndrome coronavirus from dromedary camels in the human respiratory tract: an in-vitro and ex-vivo study. *Lancet Respir. Med.* 2, 813–822, [https://doi.org/10.1016/s2213-2600\(14\)70158-4](https://doi.org/10.1016/s2213-2600(14)70158-4).
- Chen, Y., and Guo, D. (2016). Molecular mechanisms of coronavirus RNA capping and methylation. *Viro. Sin.* 31, 3–11, <https://doi.org/10.1007/s12250-016-3726-4>.
- Chen, H., Wei, P., Huang, C., Tan, L., Liu, Y., and Lai, L. (2006). Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* 281, 13894–13898, <https://doi.org/10.1074/jbc.m510745200>.
- Chen, C.-Y., Chung, K.C., Chang, Y.-W., Sue, S.-C., Bai, H.-I., Riang, L., Hsiao, C.-D., and Huang, T. (2007). Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J. Mol. Biol.* 368, 1075–1086, <https://doi.org/10.1016/j.jmb.2007.02.069>.
- Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. U S A* 106, 3484–3489, <https://doi.org/10.1073/pnas.0808790106>.
- Chen, Y., Su, C., Min, K., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Po, T., et al. (2011). Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog.* 7, e1002294, <https://doi.org/10.1371/journal.ppat.1002294>.
- Cheng, K.-W., Cheng, S.-C., Chen, W.-Y., Lin, M.-H., Chuang, S.-J., Cheng, I.-H., Sun, C.-Y., and Chou, C.-Y. (2015). Thiopurine analogs and mycophenolic acid synergistically inhibit the papain-like protease of middle east respiratory syndrome coronavirus. *Antiviral Res.* 115, 9–16, <https://doi.org/10.1016/j.antiviral.2014.12.011>.
- Cho, J.-B., Lee, J.-M., Ahn, H.-C., and Jeong, Y.-J. (2015). Identification of a novel small molecule inhibitor against SARS coronavirus helicase. *J. Microbiol. Biotechnol.* 25, 2007–2010, <https://doi.org/10.4014/jmb.1507.07078>.
- Chou, C.-Y., Chien, C.-H., Han, Y.-S., Prebanda, M.T., Hsieh, H.-P., Turk, B., Chang, G.-G., and Chen, X. (2008). Thiopurine analogues inhibit papain-like protease of severe acute respiratory syndrome coronavirus. *Biochem. Pharmacol.* 75, 1601–1609, <https://doi.org/10.1016/j.bcp.2008.01.005>.
- Chu, H., Chan, J.F.-W., Wang, Y., Yuen, T.T.-T., Chai, Y., Hou, Y., Shuai, H., Dong, Y., Hu, B., Huang, X., Zhang, Xi, et al. (2020). Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an Ex Vivo study with implications for the pathogenesis of COVID-19. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/cia410>.
- Clementz, M.A., Chen, Z., Banach, B.S., Wang, Y., Sun, L., Ratia, K., Baez-Santos, Y.M., Wang, J., Takayama, J., Ghosh, A.K., et al. (2010). Deubiquitinating and interferon antagonism activities of coronavirus papain-like proteases. *J. Virol.* 84, 4619–4629, <https://doi.org/10.1128/jvi.02406-09>.
- Coleman, C.M., Sisk, J.M., Mingo, R.M., Nelson, E.A., White, J.M., and Frieman, M.B. (2016). Abelson kinase inhibitors are potent inhibitors of severe acute respiratory syndrome coronavirus and middle east respiratory syndrome coronavirus fusion. *J. Virol.* 90, 8924–8933, <https://doi.org/10.1128/jvi.01429-16>.
- Cong, Y., Ulasli, M., Hein, S., Mauthe, M., V'kovski, P., Kriegenburg, F., Thiel, V., Cornelis, A., de Haan, M., and Reggiori, F. (2019). Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. *J. Virol.* 94, <https://doi.org/10.1128/jvi.01925-19>.
- Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020). The species *severe acute respiratory syndrome-related coronavirus*: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* 5, 536–544, <https://doi.org/10.1038/s41564-020-0695-z>.
- Cui, J., Li, F., and Shi, Z.L. (2018). Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192, <https://doi.org/10.1038/s41579-018-0118-9>.
- E.S. Daar, C. McDonald, G. Crofoot, et al. Safety and antiviral activity over 10 days following a single dose of subcutaneous gs-6207, a first-in-class, long-acting hiv capsid inhibitor in people living with hiv. In Tenth IAS Conference on HIV Science, Mexico City, Mexico, 2019.
- Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Stewart, S., Zust, R., Dong, H., et al. (2010). Diamond. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452–456, <https://doi.org/10.1038/nature09489>.
- Dai, W., Zhang, B., Su, H., Li, J., Zhao, Y., Xie, X., Jin, Z., Liu, F., Li, C., You, L., et al. (2020). Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science*, eabb4489, <https://doi.org/10.1126/science.abb4489>.
- Decroly, E., Ferron, F., Lescar, J., and Bruno, C. (2011a). Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* 10, 51–65, <https://doi.org/10.1038/nrmicro2675>.
- Decroly, E., Debarnot, C., Ferron, F., Bouvet, M., Coutard, B., Imbert, I., Gluais, L., Papageorgiou, N., Sharff, A., Bricogne, G., et al. (2011b). Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog.* 7, e1002059, <https://doi.org/10.1371/journal.ppat.1002059>.
- Delany, I., Rappuoli, R., and De Gregorio, E. (2014). Vaccines for the 21st century. *EMBO Mol. Med.* 6, 708–720, <https://doi.org/10.1002/emmm.201403876>.
- Deng, X., Hackbart, M., Mettelman, R.C., O'Brien, A., Mielech, A.M., Yi, G., Cheng Kao, C., Susan, C., and Baker. (2017). Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc. Natl. Acad. Sci. U S A* 114, E4251–E4260, <https://doi.org/10.1073/pnas.1618310114>.
- Deng, X., Susan, C., and Baker. (2018). An “old” protein with a new story: coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology* 517, 157–163, <https://doi.org/10.1016/j.virol.2017.12.024>.
- Denison, M.R., Graham, R.L., Donaldson, E.F., Eckerle, L.D., and Ralph, S.B. (2011). Coronaviruses. *RNA Biol.* 8, 270–279, <https://doi.org/10.4161/rna.8.2.15013>.
- Dong, E., Du, H., and Gardner, L. (2020a). An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* 20, 533–534, [https://doi.org/10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1).
- Dong, E., Du, H., and Gardner, L. (2020b). COVID-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). <https://coronavirus.jhu.edu/map.html>.
- Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M., Stagliano, N., Donovan, M., Woolf, B., Keith, R., Raju, J., et al. (2000). A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ. Res.* 87, <https://doi.org/10.1161/01.RES.87.5.e1>.
- Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.-R., Becker, S., Rabenau, H., Marcus, P., Kolesnikova, L., Ron, A., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1967–1976, <https://doi.org/10.1056/NEJMoa030747>.
- Dunning, J., Baillie, J.K., Cao, B., and Hayden, F.G. (2014). Antiviral combinations for severe influenza. *Lancet Infect. Dis.* 14, 1259–1270, [https://doi.org/10.1016/s1473-3099\(14\)70821-7](https://doi.org/10.1016/s1473-3099(14)70821-7).
- Dyall, J., Coleman, C.M., Hart, B.J., Venkataraman, T., Holbrook, M.R., Kindrachuk, J., Johnson, R.F., Olinger, G.G., Jahrling, P.B., Laidlaw, M., et al. (2014). Repurposing of clinically developed drugs for treatment of middle east respiratory syndrome coronavirus infection. *Antimicrob. Agents Chemother.* 58, 4885–4893, <https://doi.org/10.1128/aac.03036-14>.
- Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., and Denison, M.R. (2007). High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 81, 12135–12144, <https://doi.org/10.1128/jvi.01296-07>.
- Eriksson, K.K., Cervantes-Barragán, L., Ludewig, B., and Thiel, V. (2008). Mouse hepatitis virus liver

- pathology is dependent on ADP-ribose-1"-phosphatase, a viral function conserved in the alpha-like supergroup. *J. Virol.* 82, 12325–12334, <https://doi.org/10.1128/jvi.02082-08>.
- Fehr, A.R., Athmer, J., Channappanavar, R., Phillips, J.M., Meyerholz, D.K., and Perlman, S. (2014). The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. *J. Virol.* 89, 1523–1536, <https://doi.org/10.1128/jvi.02596-14>.
- Fehr, A.R., Channappanavar, R., Jankevicius, G., Craig, F., Zhao, J., Athmer, J., Meyerholz, D.K., Ahel, I., and Perlman, S. (2016). The conserved coronavirus macrodomain promotes virulence and suppresses the innate immune response during severe acute respiratory syndrome coronavirus infection. *mBio* 7, December, <https://doi.org/10.1128/mbio.01721-16>.
- Fehr, A.R., Jankevicius, G., Ahel, I., and Perlman, S. (2018). Viral macrodomains: unique mediators of viral replication and pathogenesis. *Trends Microbiol.* 26, 598–610, <https://doi.org/10.1016/j.tim.2017.11.011>.
- Ferron, F., Subissi, L., Silveira De Moraes, A.T., Thi Tuyet Le, N., Sevajol, M., Gluais, L., Decroly, E., Vonrhein, C., Bricogne, G., Bruno, C., and Imbert, I. (2018). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U S A* 115, E162–E171, <https://doi.org/10.1073/pnas.1718806115>.
- Fielding, B.C., Gunalan, V., Timothy, H., Tan, P., Tuet Le, N., Shen, S., Khan, S., Lim, S.G., Hong, W., and Tan, Y.-J. (2006). Severe acute respiratory syndrome coronavirus protein 7a interacts with hSGT. *Biochem. Biophysical Res. Commun.* 343, 1201–1208, <https://doi.org/10.1016/j.bbrc.2006.03.091>.
- Fischer, A., Sellner, M., Neranjan, S., Martin, S., and Lill, M.A. (2020). Potential inhibitors for novel coronavirus protease identified by virtual screening of 606 million compounds. *Int. J. Mol. Sci.* 21, 3626, <https://doi.org/10.3390/ijms21103626>.
- Frieman, M., Yount, B., Agnihothram, S., Page, C., Donaldson, E., Roberts, A., Vogel, L., Woodruff, B., Scarpio, D., Subbarao, K., and Baric, R.S. (2011). Molecular determinants of severe acute respiratory syndrome coronavirus pathogenesis and virulence in young and aged mouse models of human disease. *J. Virol.* 86, 884–897, <https://doi.org/10.1128/jvi.05957-11>.
- Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., Wang, T., Sun, Q., Ming, Z., Zhang, L., Ge, J., et al. (2020). Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, eabb7498, <https://doi.org/10.1126/science.abb7498>.
- Ghosh, A.K., Takayama, J., Venkateswara Rao, K., Ratia, K., Chaudhuri, R., Mulhearn, D.C., Lee, H., Nichols, D.B., Baliji, S., Baker, S.C., et al. (2010). Severe acute respiratory syndrome coronavirus papain-like novel protease inhibitors: design, synthesis, protein-ligand x-ray structure and biological evaluation. *J. Med. Chem.* 53, 4968–4979, <https://doi.org/10.1021/jm1004489>.
- Gordon, D.E., Jang, G.M., Bouhadou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., Tummino, T.A., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. <https://doi.org/10.1038/s41586-020-2286-9>.
- Gorgulla, C. (2018). Free energy methods involving Quantum Physics, Path Integrals, and virtual screenings: development, implementation and application in drug discovery. Dissertation. <https://doi.org/10.17169/refubium-11597>.
- Gorgulla, C., Boeszoermenyi, A., Wang, Z.F., Fischer, P.D., Coote, P.W., Krishna, M., Das, P., Malets, Y.S., Radchenko, D.S., Moroz, Y.S., et al. (2020a). An open-source drug discovery platform enables ultra-large virtual screens. *Nature* 580, 663–668, <https://doi.org/10.1038/s41586-020-2117-z>.
- Gorgulla, C., Fackeldey, K., Wagner, G., and Arthanari, H. (2020b). Accounting of receptor flexibility in ultra-large virtual screens with VirtualFlow using a grey wolf optimization method. *Supercomputing Front. Innov.* 7, 4–12, <https://doi.org/10.14529/sf1200301>.
- Graham, R.L., Becker, M.M., Eckerle, L.D., Bolles, M., Denison, M.R., and Baric, R.S. (2012). A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* 18, 1820–1826, <https://doi.org/10.1038/nm.2972>.
- de Groot, R.J., Baker, S.C., Baric, R.S., Brown, C.S., Drosten, C., Enjuanes, L., Fouchier, R.A.M., Galiano, M., Gorbalenya, A.E., Memish, Z.A., Perlman, S., et al. (2013). Middle East respiratory syndrome coronavirus (MERS-CoV): Announcement of the coronavirus study group. *J. Virol.* 87, 7790–7792, <https://doi.org/10.1128/jvi.01244-13>.
- Guarino, L.A., Bhardwaj, K., Wen, D., Sun, J., Holzenburg, A., and Kao, C. (2005). Mutational analysis of the SARS virus nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J. Mol. Biol.* 353, 1106–1117, <https://doi.org/10.1016/j.jmb.2005.09.007>.
- Gui, M., Song, W., Zhou, H., Xu, J., Chen, S., Xiang, Y., and Wang, X. (2016). Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* 27, 119–129, <https://doi.org/10.1038/cr.2016.152>.
- de Haan, C.A.M., and Rottier, P.J.M. (2005). Molecular interactions in the assembly of coronaviruses. *Adv. Virus Res.* 165–230, [https://doi.org/10.1016/s0065-3527\(05\)64006-7](https://doi.org/10.1016/s0065-3527(05)64006-7).
- Hänel, K., and Willbold, D. (2007). SARS-CoV accessory protein 7a directly interacts with human LFA-1. *Biol. Chem.* 388, <https://doi.org/10.1515/bc.2007.157>.
- Heurich, A., Hofmann-Winkler, H., Gierer, S., Liepold, T., Jahn, O., and Pohlmann, S. (2014). TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.* 88, 1293–1307, <https://doi.org/10.1128/jvi.02202-13>.
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., and Wu, N.-H. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280.e8, <https://doi.org/10.1016/j.cell.2020.02.052>.
- Holland, L.A., Kaelin, E.A., Maqsood, R., Estifanos, B., Wu, L.I., Varsani, A., Halden, R.U., Hogue, B.G., Scotch, M., and Lim, E.S. (2020). An 81 nucleotide deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in Arizona (jan-mar 2020). *J. Virol.* <https://doi.org/10.1128/jvi.00711-20>.
- Hsieh, P.-K., Chang, S.C., Huang, C.-C., Lee, T.-T., Hsiao, C.-W., Kou, Y.-H., Chen, I.-Y., Chang, C.-K., Huang, T.-H., and Chang, M.-F. (2005). Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J. Virol.* 79, 13848–13855, <https://doi.org/10.1128/jvi.79.22.13848-13855.2005>.
- Hsu, W.-C., Chang, H.-C., and Chou, C.-Y. (2005). Pui-Jen Tsai, Pei-In Lin, and Gu-Gang Chang. Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. *J. Biol. Chem.* 280, 22741–22748, <https://doi.org/10.1074/jbc.m502556200>.
- Hu, T., Chen, C., Li, H., Dou, Y., Zhou, M., Lu, D., Qi, Z., Li, Y., Cheng, Y., Zhong, Z., Singh, N., et al. (2017). Structural basis for dimerization and RNA binding of avian infectious bronchitis virus nsp9. *Protein Sci.* 26, 1037–1048, <https://doi.org/10.1002/pro.3150>.
- Huang, C., Ito, N., Tseng, C.-T.K., and Makino, S. (2006a). Severe acute respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. *J. Virol.* 80, 7287–7294, <https://doi.org/10.1128/jvi.00414-06>.
- Huang, N., Shoichet, B.K., and Irwin, J.J. (2006b). Benchmarking sets for molecular docking. *J. Med. Chem.* 49, 6789–6801, <https://doi.org/10.1021/jm0608356>.
- Hung, I.F.N., To, K.K.W., Chan, J.F.W., Cheng, V.C.C., Liu, K.S.H., Tam, A., Chan, T.-C., Zhang, A.J., Li, P., Wong, T.-L., Zhang, R., et al. (2017). Efficacy of clarithromycin-naproxen-oseltamivir combination in the treatment of patients hospitalized for influenza a(h3n2) infection. *Chest* 151, 1069–1080, <https://doi.org/10.1016/j.chest.2016.11.012>.
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in wuhan, China. *Lancet* 395, 497–506, [https://doi.org/10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5).
- Hyde, J.L., Gardner, C.L., Kimura, T., White, J.P., Liu, G., Trobaugh, D.W., Huang, C., Tonelli, M., Paessler, S., Takeda, K., Klimstra, W.B., et al. (2014). A viral RNA structural element alters host recognition of nonself RNA. *Science* 343, 783–787, <https://doi.org/10.1126/science.1248465>.
- Imbert, I., Guillemot, J.-C., Bourhis, J.-M., Bussetta, C., Coutard, B., Egloff, M.-P., Ferron, F., Gorbalenya, A.E., and Bruno, C. (2006). A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 25, 4933–4942, <https://doi.org/10.1038/sj.emboj.7601368>.

- Iwabuchi, K., Yoshie, K., Kurakami, Y., Takahashi, K., Kato, Y., and Morishima, T. (2020). Therapeutic potential of ciclesonide inhalation for COVID-19 pneumonia: report of three cases. *J. Infect. Chemother.* 26, 625–632, <https://doi.org/10.1016/j.jiac.2020.04.007>.
- Ivanov, K.A., Thiel, V., Dobbe, J.C., van der Meer, Y., Snijder, E.J., and Ziebuhr, J. (2004). Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J. Virol.* 78, 5619–5632, <https://doi.org/10.1128/jvi.78.11.5619-5632.2004>.
- Jackson, L.A., Anderson, E.J., Roupael, N.G., Roberts, P.C., Makhene, M., Coler, R.N., McCullough, M.P., Chappell, J.D., Denison, M.R., Stevens, L.J., Pruijssers, A.J., et al. (2020). An mRNA vaccine against SARS-CoV-2 — preliminary report. *N. Engl. J. Med.* 383, 1920–1931, <https://doi.org/10.1056/NEJMoa2022483>.
- James, H. (2020). Genomic epidemiology of novel coronavirus - global subsampling. <http://nextstrain.org/ncov/global>.
- Jia, Z., Yan, L., Ren, Z., Wu, L., Wang, J., Guo, J., Zheng, L., Ming, Z., Zhang, L., Lou, Z., and Rao, Z. (2019). Delicate structural coordination of the severe acute respiratory syndrome coronavirus nsp13 upon ATP hydrolysis. *Nucleic Acids Res.* 47, 6538–6550, <https://doi.org/10.1093/nar/gkz409>.
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., et al. (2020a). Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature*. <https://doi.org/10.1038/s41586-020-2223-y>.
- Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., Duan, Y., et al. (2020b). Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-020-0440-6>.
- Joseph, M., Gingras, A.-C., Nahum, S., and Burley, S.K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4e) bound to 7-methyl-GDP. *Cell* 89, 951–961, [https://doi.org/10.1016/s0092-8674\(00\)80280-9](https://doi.org/10.1016/s0092-8674(00)80280-9).
- Kadioglu, O., Saeed, M., Greten, H.J., and Efferth, T. (2020). Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. *Bull. World Health Organ.* <https://doi.org/10.2471/blt.20.255943>.
- Kang, X., Zhong, N., Zou, P., Zhang, S., Jin, C., and Xia, B. (2012). Foldon unfolding mediates the interconversion between Mpro-C monomer and 3D domain-swapped dimer. *Proc. Natl. Acad. Sci. U S A* 109, 14900–14905, <https://doi.org/10.1073/pnas.1205241109>.
- Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., Yan, Y., et al. (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm. Sin. B*. <https://doi.org/10.1016/j.apsb.2020.04.009>.
- Kao, R.Y., Wayne, H., Tsui, W., Lee, T.S.W., Tanner, J.A., Watt, R.M., Huang, J.-D., Hu, L., Chen, G., Chen, Zhiwei, Zhang, L., et al. (2004). Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics. *Chem. Biol.* 11, 1293–1299, <https://doi.org/10.1016/j.chembiol.2004.07.013>.
- Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. *EMBO J.* 24, 1911–1920, <https://doi.org/10.1038/sj.emboj.7600664>.
- Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C.J., Cerikan, B., Lu, J.M., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* 588, 498–502, <https://doi.org/10.1038/s41586-020-2665-2>.
- Kim, T.S., Heinlein, C., Hackman, R.C., and Nelson, P.S. (2006). Phenotypic analysis of mice lacking the Tmprss2-encoded protease. *Mol. Cell Biol.* 26, 965–975, <https://doi.org/10.1128/mcb.26.3.965-975.2006>.
- Kim, Y., Jedrzejczak, R., Maltseva, N.I., Wilamowski, M., Endres, M., Adam, G., Michalska, K., and Joachimiak, A. (2020). Crystal structure of nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci.* <https://doi.org/10.1002/pro.3873>.
- Kiplin Guy, R., DiPaola, R.S., Romanelli, F., and Dutch, R.E. (2020). Rapid repurposing of drugs for COVID-19. *Science* 368, 829–830, <https://doi.org/10.1126/science.abb9332>.
- Kirchdoerfer, R.N., and Ward, A.B. (2019). Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* 10, <https://doi.org/10.1038/s41467-019-10280-3>.
- Koebel, M.R., Grant, S., Posner, R.G., et al. (2016). AutoDock VinaXB: implementation of xbsf, new empirical halogen bond scoring function, into AutoDock Vina. *J. Cheminformatics* 8, 27, <https://doi.org/10.1186/s13321-016-0139-1>.
- Koes, D.R., Baumgartner, M.P., and Camacho, C.J. (2013). Lessons Learned in empirical scoring with smina from the CSAR 2011 benchmarking Exercise. *J. Chem. Inf. Model.* 53, 1893–1904, <https://doi.org/10.1021/ci300604z>.
- Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., et al.; SARS Working Group (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1953–1966, <https://doi.org/10.1056/NEJMoa030781>.
- Kuiken, T., Ron, A., Fouchier, M., Martin, S., Rimmelzwaan, G.F., van Amerongen, G., van Riel, D., Laman, J.D., de Jong, T., van Doornum, G., Lim, W., Ai Ee Ling, P.K., et al. (2003). Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* 362, 263–270, [https://doi.org/10.1016/s0140-6736\(03\)13967-0](https://doi.org/10.1016/s0140-6736(03)13967-0).
- Kuo, C.-J., Chi, Y.-H., Hsu, J.T.-A., and Liang, P.-H. (2004). Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate. *Biochem. Biophysical Res. Commun.* 318, 862–867, <https://doi.org/10.1016/j.bbrc.2004.04.098>.
- Kuri, T., Eriksson, K.K., Putics, A., Züst, R., Snijder, E.J., Davidson, A.D., Siddell, S.G., Thiel, V., Ziebuhr, J., and Weber, F. (2011). The ADP-ribose-1''-monophosphatase domains of severe acute respiratory syndrome coronavirus and human coronavirus 229E mediate resistance to antiviral interferon responses. *J. Gen. Virol.* 92, 1899–1905, <https://doi.org/10.1099/vir.0.031856-0>.
- Kuritzkes, D.R., Marschner, I., Johnson, V.A., Bassett, R., Eron, J.J., Fischl, M.A., Murphy, R.L., Fife, K., Maenza, J., and Rosandich, M.E. (1999). Lamivudine in combination with zidovudine, stavudine, or didanosine in patients with HIV-1 infection. a randomized, double-blind, placebo-controlled trial. *AIDS* 13, 685–694, <https://doi.org/10.1097/00002030-199904160-00009>.
- Kusov, Y., Tan, J., Alvarez, E., Enjuanes, L., and Hilgenfeld, R. (2015). A G-quadruplex-binding macrodomain within the “SARS-unique domain” is essential for the activity of the SARS-coronavirus replication–transcription complex. *Virology* 484, 313–322, <https://doi.org/10.1016/j.virol.2015.06.016>.
- Lai, A.C., and Crews, C.M. (2016). Induced protein degradation: an emerging drug discovery paradigm. *Nat. Rev. Drug Discov.* 16, 101–114, <https://doi.org/10.1038/nrd.2016.211>.
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., and Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. <https://doi.org/10.1038/s41586-020-2180-5>.
- Lee, J.-M., Cho, J.-B., Ahn, H.-C., Jung, W., and Jeong, Y.-J. (2017). A novel chemical compound for inhibition of SARS coronavirus helicase. *J. Microbiol. Biotechnol.* 27, 2070–2073, <https://doi.org/10.4014/jmb.1707.07073>.
- Lehmann, K.C., Gulyaeva, A., Zevenhoven-Dobbe, J.C., George, M., Janssen, C., Ruben, M., Overkleef, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res.* 43, 8416–8434, <https://doi.org/10.1093/nar/gkv838>.
- Li, C., Debing, Y., Jankevicius, G., Neyts, J., Ahel, I., Coutard, B., and Bruno, C. (2016). Viral macro domains reverse protein ADP-ribosylation. *J. Virol.* 90, 8478–8486, <https://doi.org/10.1128/jvi.00705-16>.
- Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.-K., Huang, I.-C., Xu, K., Vasilieva, N., Murakami, A., et al. (2005). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J.* 24, 1634–1643, <https://doi.org/10.1038/sj.emboj.7600640>.
- Liao, Y., Zhang, S.M., Neo, T.L., and Tam, J.P. (2015). Tryptophan-dependent membrane interaction and heteromerization with the internal fusion peptide by the membrane proximal external region of SARS-CoV spike protein. *Biochemistry* 54, 1819–1830, <https://doi.org/10.1021/bi5011352u>.
- Lin, S.-Y., Liu, C.-L., Chang, Y.-M., Zhao, J., Perlman, S., and Hou, M.-H. (2014). Structural

- basis for the identification of the N-terminal domain of coronavirus nucleocapsid protein as an antiviral target. *J. Med. Chem.* 57, 2247–2257, <https://doi.org/10.1021/jm500089r>.
- Lindner, H.A., Lytvyn, V., Qi, H., Lachance, P., Ziomek, E., and Ménard, R. (2007). Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease. *Arch. Biochem. Biophys.* 466, 8–14, <https://doi.org/10.1016/j.abb.2007.07.006>.
- Liu, S., Xiao, G., Chen, Y., He, Y., Niu, J., Escalante, C.R., Xiong, H., Farmar, J., Debnath, A.K., Po, T., and Jiang, S. (2004). Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. *The Lancet* 363, 938–947, [https://doi.org/10.1016/s0140-6736\(04\)15788-7](https://doi.org/10.1016/s0140-6736(04)15788-7).
- Liu, X., Fang, P., Fang, L., Hong, Y., Zhu, X., Wang, D., Peng, G., and Xiao, S. (2019). Porcine deltacoronavirus nsp15 antagonizes interferon β production independently of its endoribonuclease activity. *Mol. Immunol.* 114, 100–107, <https://doi.org/10.1016/j.molimm.2019.07.003>.
- Lu, L., Qi, L., Zhu, Y., Chan, K.-H., Qin, L., Yuan, L., Wang, Q., Chan, J.F.-W., Du, L., Fei, Y., et al. (2014). Structure-based discovery of Middle East respiratory syndrome coronavirus fusion inhibitor. *Nat. Commun.* 5, <https://doi.org/10.1038/ncomms4067>.
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., et al. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395, 565–574, [https://doi.org/10.1016/s0140-6736\(20\)30251-8](https://doi.org/10.1016/s0140-6736(20)30251-8).
- Lucas, J.M., Heinlein, C., Kim, T., Hernandez, S.A., Malik, M.S., True, L.D., Morrissey, C., Corey, E., Montgomery, B., Mostaghel, E., Clegg, N., et al. (2014). The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discov.* 4, 1310–1325, <https://doi.org/10.1158/2159-8290.cd-13-1010>.
- Lyu, J., Wang, S., Balus, T.E., Singh, I., Levit, A., Moroz, Y.S., O'Meara, M.J., Che, T., Alga, E., Tolmachova, K., et al. (2019). Ultra-large library docking for discovering new chemotypes. *Nature* 566, 224–229, <https://doi.org/10.1038/s41586-019-0917-9>.
- Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc. Natl. Acad. Sci.* 112, 9436–9441, <https://doi.org/10.1073/pnas.1508686112>.
- Masters, P.S. (2006). The molecular biology of coronaviruses. *Adv. Virus Res.* 66, 193–292, [https://doi.org/10.1016/s0065-3527\(06\)66005-3](https://doi.org/10.1016/s0065-3527(06)66005-3).
- McPherson, R.L., Abraham, R., Sreekumar, E., Ong, S.-E., Cheng, S.-J., Baxter, V.K., Kistemaker, H.A.V., Filippov, D.V., Griffin, D.E., and Leung, A.K.L. (2017). ADP-ribosylhydrolase activity of chikungunya virus macrodomain is critical for virus replication and virulence. *Proc. Natl. Acad. Sci. U S A* 114, 1666–1671, <https://doi.org/10.1073/pnas.1621485114>.
- Meyer, M., Huang, E., Yuzhakov, O., Ramanathan, P., Ciaramella, G., and Alexander, B. (2018). Modified mRNA-based vaccines elicit robust immune responses and protect Guinea pigs from ebola virus disease. *J. Infect. Dis.* 217, 451–455, <https://doi.org/10.1093/infdis/jix592>.
- Michael, O. (2015). Hottiger. SnapShot: ADP-ribosylation signaling. *Mol. Cell* 58, 1134–1134.e1, <https://doi.org/10.1016/j.molcel.2015.06.001>.
- Miknis, Z.J., Donaldson, E.F., Umland, T.C., Rimmer, R.A., Baric, R.S., and Wayne Schultz, L. (2009). Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J. Virol.* 83, 3007–3018, <https://doi.org/10.1128/jvi.01505-08>.
- Millet, J.K., and Whittaker, G.R. (2015). Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* 202, 120–134, <https://doi.org/10.1016/j.virusres.2014.11.021>.
- Min, K., Chen, Y., Wu, A., Sun, Y., Su, C., Wu, H., Jin, X., Tao, J., Wang, Y., Xiao, M., et al. (2012). Short peptides derived from the interaction domain of SARS coronavirus nonstructural protein nsp10 can suppress the 2'-O-methyltransferase activity of nsp10/nsp16 complex. *Virus Res.* 167, 322–328, <https://doi.org/10.1016/j.virusres.2012.05.017>.
- Minskaia, E., Hertzog, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., and Ziebuhr, J. (2006). Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U S A* 103, 5108–5113, <https://doi.org/10.1073/pnas.0508200103>.
- Alnazawi, M., Abdallah, A., and Kandeel, M. (2017). Comparative genomic analysis MERS CoV isolated from humans and camels with special reference to virus encoded helicase. *Biol. Pharm. Bull.* 40, 1289–1298, <https://doi.org/10.1248/bpb.17-00241>.
- Morgan, S.G., Testa, A., Lucas, X., Chan, K.-Ho, Chen, W., Lamont, D.J., Zengerle, M., and Ciulli, A. (2017). Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat. Chem. Biol.* 13, 514–521, <https://doi.org/10.1038/nchembio.2329>.
- Mysinger, M.M., Carchia, M., Irwin, J.J., and Shoichet, B.K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55, 6582–6594, <https://doi.org/10.1021/jm300687e>.
- N2ngu3. (n.d.). SWISS-MODEL Repository, Helicase SARSCoV2, <https://swissmodel.expasy.org/interactive/N2NgU3/>.
- Naik, B., Gupta, N., Ojha, R., Singh, S., Kumar Prajapati, V., and Prusty, D. (2020). High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment. *Int. J. Biol. Macromolecules* 160, 1–17, <https://doi.org/10.1016/j.ijbiomac.2020.05.184>.
- Nelson, C.A., Pekosz, A., Lee, C.A., Diamond, M.S., and Fremont, D.H. (2005). Structure and intracellular targeting of the SARS-coronavirus orf7a accessory protein. *Structure* 13, 75–85, <https://doi.org/10.1016/j.str.2004.10.010>.
- Nelson, C.A., Minasov, G., Shuvalova, L., and Fremont, D.H. (2020). Structure of the Sars-Cov-2 ORF7a Encoded Accessory Protein.
- Neuman, B.W., and Buchmeier, M.J. (2016). Supramolecular architecture of the coronavirus particle. *Adv. Virus Res.* 96, 1–27, <https://doi.org/10.1016/bs.avir.2016.08.005>.
- Neuman, B.W., Adair, B.D., Yoshioka, C., Quispe, J.D., Orca, G., Kuhn, P., Milligan, R.A., Yeager, M., and Buchmeier, M.J. (2006). Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. *J. Virol.* 80, 7918–7928, <https://doi.org/10.1128/jvi.00645-06>.
- Nivedha, A.K., Thieker, D.F., Makeneni, S., Hu, H., and Woods, R.J. (2016). Vina-Carb: improving glycosidic angles during carbohydrate docking. *J. Chem. Theor. Comput.* 12, 892–901, <https://doi.org/10.1021/acs.jctc.5b00834>.
- Nukoolkarn, V., Lee, V.S., Malaisree, M., Aruksakulwong, O., and Hannongbua, S. (2008). Molecular dynamic simulations analysis of ritonavir and lopinavir as SARS-CoV 3CLpro inhibitors. *J. Theor. Biol.* 254, 861–867, <https://doi.org/10.1016/j.jtbi.2008.07.030>.
- O15393. (n.d.) SWISS-MODEL Repository, TMPRSS2Human, <https://swissmodel.expasy.org/repository/uniprot/O15393>.
- Ogando, N.S., Ferron, F., Decroly, E., Bruno, C., Posthuma, C.C., and Snijder, E.J. (2019). The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. *Front. Microbiol.* 10, <https://doi.org/10.3389/fmicb.2019.01813>.
- Ortiz-Alcantara, J., Bhardwaj, K., Palaninathan, S., Frieman, M., Baric, R., and Kao, C. (2010). Small molecule inhibitors of the SARS-CoV Nsp15 endoribonuclease. *Virus Adapt. Treat* 2, 125–133, <https://doi.org/10.2147/VAAT.S12733>.
- Palella, F.J., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J., and Holmberg, S.D. (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N. Engl. J. Med.* 338, 853–860, <https://doi.org/10.1056/nejm199803263381301>.
- Panda, P.K., Arul, M.N., Patel, P., Verma, S.K., Luo, W., Rubahn, H.-G., Mishra, Y.K., Suar, M., and Ahuja, R. (2020). Structure-based drug designing and immunoinformatics approach for SARS-CoV-2. *Sci. Adv.* eabb8097, <https://doi.org/10.1126/sciadv.abb8097>.
- Park, J.-Y., Kim, J.H., Young, M.K., Jeong, H.J., Kim, D.W., Park, K.H., Kwon, H.-J., Park, S.-J., Lee, W.S., and Young, B.R. (2012). Tanshinones as selective and slow-binding inhibitors for SARS-CoV cysteine proteases. *Bioorg. Med. Chem.* 20, 5928–5935, <https://doi.org/10.1016/j.bmc.2012.07.038>.
- Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T., Cheng, V.C.C., et al. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *The Lancet* 361,

1319–1325. [https://doi.org/10.1016/s0140-6736\(03\)13077-2](https://doi.org/10.1016/s0140-6736(03)13077-2).

Peter, J., Jean, M., Shields, D.C., Higgins, D.G., and Davey, N.E. (2016). ProViz—a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Res.* 44, W11–W15. <https://doi.org/10.1093/nar/gkw265>.

Polack, F.P., Thomas, S.J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J.L., Pérez Marc, G., Moreira, E.D., Zerbini, C., et al. (2020). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2034577>.

Putics, Á., Filipowicz, W., Hall, J., Gorbalenya, A.E., and Ziebuhr, J. (2005). ADP-ribose-1^o-monophosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J. Virol.* 79, 12721–12731. <https://doi.org/10.1128/jvi.79.20.12721-12731.2005>.

Putics, Á., Slaby, J., Filipowicz, W., Gorbalenya, A.E., and Ziebuhr, J. (2006). Adp-ribose-1^o-phosphatase activities of the human coronavirus 229e and sars coronavirus x domains. In *Advances in Experimental Medicine and Biology* (Springer US), pp. 93–96. <https://doi.org/10.1007/978-0-387-33012-915>.

Pyrk, K., Sims, A.C., Dijkman, R., Jebbink, M., Long, C., Deming, D., Donaldson, E., Vabret, A., Baric, R., van der Hoek, L., and Pickles, R. (2010). Culturing the unculturable: human coronavirus HKU1 infects, replicates, and produces progeny virions in human ciliated airway epithelial cell cultures. *J. Virol.* 84, 11255–11263. <https://doi.org/10.1128/jvi.00947-10>.

Ragoza, M., Hochuli, J., Idrobo, E., Sunseri, J., and Koes, D.R. (2017). Protein–ligand scoring with convolutional neural networks. *J. Chem. Inf. Model.* 57, 942–957. <https://doi.org/10.1021/acs.jcim.6b00740>. PMID: 28368587.

Ratia, K., Kilianski, A., Baez-Santos, Y.M., Baker, S.C., and Mesecar, A. (2014). Structural basis for the ubiquitin-linkage specificity and delSGylating activity of SARS-CoV papain-like protease. *PLoS Pathog.* 10, e1004113. <https://doi.org/10.1371/journal.ppat.1004113>.

Recovery Collaborative Group, Horby, P., Lim, W.S., Emberson, J.R., Mafham, M., Bell, J.L., Linsell, L., Staplin, N., Brightling, C., Ustianowski, A., et al. (2020). Dexamethasone in hospitalized patients with covid-19 - preliminary report. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2021436>.

Roman, W., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., et al. (2020). Virological assessment of hospitalized patients with COVID-2019. *Nature.* <https://doi.org/10.1038/s41586-020-2196-x>.

Rothe, C., Schunk, M., Peter, S., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., et al. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N. Engl. J. Med.* 382, 970–971. <https://doi.org/10.1056/NEJMc2001468>.

Sager, J., Begley, R., Rhee, M., West, S., Ling, J., Escobar, J., and Mathias, A. (2019). Safety and pk

of subcutaneous gs-6207, a novel hiv-1 capsid inhibitor. *HIV MEDICINE, volume 20* (WILEY), p. 39.

Sander van, B., de Graaf, M., Lauber, C., Bestebroer, T.M., Stalin Raj, V., Zaki, A.M., Albert, D., Osterhaus, M.E., Haagmans, B.L., Gorbalenya, A.E., et al. (2012). Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *mBio* 3. <https://doi.org/10.1128/mBio.00473-12>.

Sauter, D., Specht, A., and Frank, K. (2010). Tetherin: Holding on and letting go. *Cell* 141, 392–398. <https://doi.org/10.1016/j.cell.2010.04.022>.

Schaecher, S.R., Touchette, E., Schriewer, J., Mark Buller, R., and Pekosz, A. (2007). Severe acute respiratory syndrome coronavirus gene 7 products contribute to virus-induced apoptosis. *J. Virol.* 81, 11054–11068. <https://doi.org/10.1128/jvi.01266-07>.

Sevajol, M., Subissi, L., Decroly, E., Bruno, C., and Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Res.* 194, 90–99. <https://doi.org/10.1016/j.virusres.2014.10.008>.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., and Li, F. (2020a). Structural basis of receptor recognition by SARS-CoV-2. *Nature* 581, 221–224. <https://doi.org/10.1038/s41586-020-2179-y>.

Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., and Li, F. (2020b). Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U S A* 117, 11727–11734. <https://doi.org/10.1073/pnas.2003138117>.

Shaw Research, D.E. (2020). Molecular Dynamics Simulations Related to SARS-CoV-2. https://www.deshawresearch.com/downloads/download_trajectory_sarscov2.cgi.

Sheahan, T.P., Sims, A.C., Zhou, S., Graham, R.L., Puijssers, A.J., Agostini, M.L., Leist, S.R., Schäfer, A., Dinnon, K.H., Stevens, L.J., et al. (2020). An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci. Translational Med.* 12, eabb5883. <https://doi.org/10.1126/scitransmed.abb5883>.

Sheahan, T.P., Sims, A.C., Graham, R.L., Menachery, V.D., Gralinski, L.E., Case, J.B., Leist, S.R., Pyrk, K., Feng, J.Y., Trantcheva, I., et al. (2017). Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci. Translational Med.* 9, eaal3653. <https://doi.org/10.1126/scitransmed.aal3653>.

Shi, J., Sivaraman, J., and Song, J. (2008). Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. *J. Virol.* 82, 4620–4629. <https://doi.org/10.1128/jvi.02680-07>.

Singh, S., and Florez, H. (2020). Coronavirus disease 2019 drug discovery through molecular docking. *F1000Research* 9, 502. <https://doi.org/10.12688/f1000research.24218.1>.

Singh Saikatendu, K., Joseph, J.S., Subramanian, V., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B.W., Buchmeier, M.J., and Stevens,

R.C. (2005). Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1^o-phosphate dephosphorylation by a conserved domain of nsP3. *Structure* 13, 1665–1675. <https://doi.org/10.1016/j.str.2005.07.022>.

Smith, E.C., Blanc, H., Vignuzzi, M., and Denison, M.R. (2013). Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathog.* 9, e1003565. <https://doi.org/10.1371/journal.ppat.1003565>.

Smith, E.C., Brett Case, J., Blanc, H., Isakov, O., Shomron, N., Vignuzzi, M., and Denison, M.R. (2015). Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. *J. Virol.* 89, 6418–6426. <https://doi.org/10.1128/jvi.00110-15>.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Leo, L., Poon, M., Guan, Y., Rozanov, M., Spaan, W.J., and Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* 331, 991–1004. [https://doi.org/10.1016/s0022-2836\(03\)00865-9](https://doi.org/10.1016/s0022-2836(03)00865-9).

Snijder, E.J., Decroly, E., and Ziebuhr, J. (2016). The nonstructural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.* 96, 59–126. <https://doi.org/10.1016/bs.aivir.2016.08.008>.

W.H.O. Solidarity Trial Consortium, Pan, H., Peto, R., Henaio-Restrepo, A.M., Preziosi, M.P., Sathiyamoorthy, V., Abdool Karim, Q., Alejandria, M.M., Hernandez Garcia, C., Kieny, M.P., Malekzadeh, R., et al. (2020). Repurposed antiviral drugs for Covid-19 - interim WHO Solidarity Trial results. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2023184>.

Srinivasan, S., Cui, H., Gao, Z., Liu, M., Lu, S., Mkandawire, W., Narykov, O., Sun, M., and Korkin, D. (2020). Structural genomics of SARS-CoV-2 indicates evolutionary conserved functional regions of viral proteins. *Viruses* 12, 360. <https://doi.org/10.3390/v12040360>.

Stein, R.M., Kang, H.J., McConvy, J.D., Glatfelter, G.C., Jones, A.J., Che, T., Slocum, S., Huang, X.-P., Savych, O., Moroz, Y.S., et al. (2020). Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. *Nature* 579, 609–614. <https://doi.org/10.1038/s41586-020-2027-0>.

Stohlman, S.A., Baric, R.S., Nelson, G.N., Soe, L.H., Welter, L.M., and Deans, R.J. (1988). Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* 62, 4288–4295.

Strodel, B., Olubiyi, O., Olagunju, M., Keutmann, M., and Loschwitz, J. (2020). High throughput virtual screening to discover inhibitors of the main protease of the coronavirus SARS-CoV-2. Preprints. <https://doi.org/10.20944/preprints202004.0161.v1>.

Su, S., Wong, G., Shi, W., Liu, J., Alexander, C., Lai, K., Zhou, J., Liu, W., Bi, Y., and Gao, G.F. (2016). Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* 24, 490–502. <https://doi.org/10.1016/j.tim.2016.03.003>.

- Su, Hx., Yao, S., Zhao, Wf., et al. (2020). Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. *Acta Pharmacologica Sinica* 41, 1167–1177, <https://doi.org/10.1038/s41401-020-0483-6>.
- Sutton, G., Fry, E., Carter, L., Sainsbury, S., Walter, T., Nettleship, J., Berrow, N., Owens, R., Gilbert, R., Davidson, A., et al. (2004). The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* 12, 341–353, <https://doi.org/10.1016/j.str.2004.01.016>.
- Taylor, J.K., Coleman, C.M., Postel, S., Sisk, J.M., Bernbaum, J.G., Venkataraman, T., Sundberg, E.J., and Frieman, M.B. (2015). Severe acute respiratory syndrome coronavirus ORF7a inhibits bone marrow stromal antigen 2 virion tethering through a novel mechanism of glycosylation interference. *J. Virol.* 89, 11820–11833, <https://doi.org/10.1128/jvi.02274-15>.
- te Velthuis, A.J.W., van den Worm, S.H.E., and Snijder, E.J. (2011). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* 40, 1737–1747, <https://doi.org/10.1093/nar/gkr893>.
- To, K.K.-W., Tsang, O.T.-Y., Leung, W.-S., Tam, A.R., Wu, T.-C., Lung, D.C., Yip, C.C.-Y., Cai, J.-P., Chan, J.M.-C., Chik, T.S.-H., Lau, D.P.-L., et al. (2020). Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect. Dis.* 20, 565–574, [https://doi.org/10.1016/s1473-3099\(20\)30196-1](https://doi.org/10.1016/s1473-3099(20)30196-1).
- Ton, A.-T., Gentile, F., Hsing, M., Ban, F., and Cherkasov, A. (2020). Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds. *Mol. Inform.* <https://doi.org/10.1002/minf.202000028>.
- Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461, <https://doi.org/10.1002/jcc.21334>.
- Turoňová, B., Sikora, M., Schürmann, C., Wim, J., Hagen, H., Welsch, S., Florian, E., Blanc, C., von Bülow, S., Gecht, M., Bagola, K., et al. (2020). In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* 370, 203–208, <https://doi.org/10.1126/science.abd5223>.
- Vella, S., Bernard, S., Papa Sow, S., Paul Eholie, S., and Murphy, R.L. (2012). The history of antiretroviral therapy and of its implementation in resource-limited areas of the world. *AIDS* 26, 1231–1241, <https://doi.org/10.1097/qad.0b013e32835521a3>.
- Voitenko, O.S., Dhroso, A., Feldmann, A., Korkin, D., and Kalinina, O.V. (2016). Patterns of amino acid conservation in human and animal immunodeficiency viruses. *Bioinformatics* 32, i685–i692, <https://doi.org/10.1093/bioinformatics/btw441>.
- Voysey, M., Costa Clemens, S.A., Madhi, S.A., Weckx, L.Y., Folegatti, P.M., Aley, P.K., Angus, B., Baillie, V.L., Barnabas, Shaun L., Bhorat, Q.E., Bibi, S., et al. (2020). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet.* [https://doi.org/10.1016/s0140-6736\(20\)32661-1](https://doi.org/10.1016/s0140-6736(20)32661-1).
- Walls, A.C., Park, Y.-J., Alejandra Tortorici, M., Wall, A., McGuire, A.T., and Velesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181, 281–292.e6, <https://doi.org/10.1016/j.cell.2020.02.058>.
- Wan, X., Yang, T., Cuesta, A., Pang, X., Balias, T.E., Irwin, J.J., Shoichet, B.K., and Jack, T. (2020). Discovery of lysine-targeted eIF4E inhibitors through covalent docking. *J. Am. Chem. Soc.* 142, 4960–4964, <https://doi.org/10.1021/jacs.9b10377>.
- Wang, L., Valderramos, S.G., Wu, A., Ouyang, S., Li, C., Brasil, P., Bonaldo, M., Coates, T., Nielsen-Saines, K., Jiang, T., et al. (2016). From mosquitos to humans: genetic evolution of zika virus. *Cell Host & Microbe* 19, 561–565, <https://doi.org/10.1016/j.chom.2016.04.006>.
- Wang, L., Bao, B.-B., Song, G.-Q., Chen, C., Zhang, X.-M., Lu, W., Wang, Z., Cai, Y., Li, S., Fu, S., et al. (2017). Discovery of unsymmetrical aromatic disulfides as novel inhibitors of SARS-CoV main protease: chemical synthesis, biological evaluation, molecular docking and 3D-QSAR study. *Eur. J. Med. Chem.* 137, 450–461, <https://doi.org/10.1016/j.ejmech.2017.05.045>.
- Wang, Y., Zhang, D., Du, G., Du, R., Zhao, J., Jin, Y., Fu, S., Gao, L., Cheng, Z., Lu, Q., Hu, Y., Luo, G., et al. (2020a). Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *The Lancet* 395, 1569–1578, [https://doi.org/10.1016/s0140-6736\(20\)31022-9](https://doi.org/10.1016/s0140-6736(20)31022-9).
- Wang, Y., Jiang, W., Qi, H., Wang, C., Wang, B., Pan, Z., Dong, N., and Tong, Q. (2020b). A retrospective cohort study of methylprednisolone therapy in severe patients with COVID-19 pneumonia. *Signal. Transduction Targeted Ther.* 5, <https://doi.org/10.1038/s41392-020-0158-2>.
- Warren, S., Wan, X.-F., Conant, G., and Korkin, D. (2013). Extreme evolutionary conservation of functionally important regions in H1N1 influenza proteome. *PLoS One* 8, e81027, <https://doi.org/10.1371/journal.pone.0081027>.
- White, J.M., Delos, S.E., Brecher, M., and Schornberg, K. (2008). Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.* 43, 189–219, <https://doi.org/10.1080/10409230802058320>.
- de Wilde, A.H., Snijder, E.J., Kikkert, M., and van Hemert, M.J. (2017). Host factors in coronavirus replication. In *Roles of Host Gene and Non-coding RNA Expression in Virus Infection*, volume 419, Ralph A. Tripp and S. Mark Tompkins, eds. (SpringerNATURE), pp. 1–42, <https://doi.org/10.1007/82201725>.
- Wong, K.M., Tai, H.K., and Siu, S.W. (2020). GWOVina: a grey wolf optimization approach to rigid and flexible receptor docking. *Chem. Biol. Drug Des.* [cbdd.13764](https://doi.org/10.1111/cbdd.13764), <https://doi.org/10.1111/cbdd.13764>.
- Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S., and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367, 1260–1263, <https://doi.org/10.1126/science.abb2507>.
- Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., Yuan, M.L., and Zhang, Y.L. (2020). A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269, <https://doi.org/10.1038/s41586-020-2008-3>.
- Xia, B., and Kang, X. (2011). Activation and maturation of SARS-CoV main protease. *Protein Cell* 2, 282–290, <https://doi.org/10.1007/s13238-011-1034-1>.
- Xia, S., Yan, L., Xu, W., Agrawal, A.S., Abdullah, A., Tseng, C.-T.K., Wang, Q., Du, L., Tan, W., Wilson, I.A., et al. (2019). A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Sci. Adv.* 5, eaav4580, <https://doi.org/10.1126/sciadv.aav4580>.
- Xia, S., Zhu, Y., Liu, M., Lan, Q., Xu, W., Wu, Y., Ying, T., Liu, S., Shi, Z., Jiang, S., and Lu, L. (2020a). Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell Mol. Immunol.* <https://doi.org/10.1038/s41423-020-0374-2>.
- Xia, S., Liu, M., Wang, C., Xu, W., Lan, Q., Feng, S., Qi, F., Bao, L., Du, L., Liu, S., et al. (2020b). Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res.* 30, 343–355, <https://doi.org/10.1038/s41422-020-0305-x>.
- Xydakis, M.S., Dehghani-Mobaraki, P., Holbrook, E.H., Geisthoff, U.W., Bauer, C., Hautefort, C., Herman, P., Manley, G.T., Lyon, D.M., and Hopkins, C. (2020). Smell and taste dysfunction in patients with COVID-19. *Lancet Infect. Dis.* [https://doi.org/10.1016/s1473-3099\(20\)30293-0](https://doi.org/10.1016/s1473-3099(20)30293-0).
- Yan, R., Zhang, Y., Li, Y., Lu, X., Guo, Y., and Zhou, Q. (2020a). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 367, 1444–1448, <https://doi.org/10.1126/science.abb2762>.
- Yan, L., Zhang, Y., Ge, J., Zheng, L., Gao, Y., Wang, T., Jia, Z., Wang, H., Huang, Y., Li, M., Wang, Q., Rao, Z., and Lou, Z. (2020b). Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat. Commun.* 11, <https://doi.org/10.1038/s41467-020-19770-1>.
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G.F., et al. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci.* 100, 13190–13195, <https://doi.org/10.1073/pnas.1835675100>.
- Yang, H., Xie, W., Xue, X., Yang, K., Ma, J., Liang, W., Qi, Z., Zhou, Z., Pei, D., Ziebuhr, J., Hilgenfeld, R., et al. (2005). Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* 3, e324, <https://doi.org/10.1371/journal.pbio.0030324>.
- Yang, X., Chen, X., Bian, G., Tu, J., Xing, Y., Wang, Y., and Chen, Z. (2014). Proteolytic processing, deubiquitinase and interferon antagonist activities of Middle East respiratory syndrome

coronavirus papain-like protease. *J. Gen. Virol.* 95, 614–626, <https://doi.org/10.1099/vir.0.059014-0>.

Yant, S.R., Mulato, A., Hansen, D., Tse, W.C., Niedziela-Majka, A., Zhang, J.R., Stepan, G.J., Jin, D., Wong, M.H., Perreira, J.M., et al. (2019). A highly potent long-acting small-molecule HIV-1 capsid inhibitor with efficacy in a humanized mouse model. *Nat. Med.* 25, 1377–1384, <https://doi.org/10.1038/s41591-019-0560-x>.

Yu, R., Chen, L., Lan, R., Shen, R., and Peng, L. (2020). Computational screening of antagonists against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking. *Int. J. Antimicrob. Agents*, 106012, <https://doi.org/10.1016/j.ijantimicag.2020.106012>.

Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor

binding domains. *Nat. Commun.* 8, <https://doi.org/10.1038/ncomms15092>.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Albert, D., Osterhaus, M.E., and Fouchier, R.A. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367, 1814–1820, <https://doi.org/10.1056/NEJMoa1211721>.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., and Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science*, eabb3405, <https://doi.org/10.1126/science.abb3405>.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., et al. (2020a). A pneumonia outbreak associated with a new coronavirus of probable bat origin.

Nature 579, 270–273, <https://doi.org/10.1038/s41586-020-2012-7>.

Zhou, X., Li, Y., Li, T., and Zhang, W. (2020b). Follow-up of asymptomatic patients with SARS-CoV-2 infection. *Clin. Microbiol. Infect.* 26, 957–959, <https://doi.org/10.1016/j.cmi.2020.03.024>.

Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., and Cheng, F. (2020c). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* 6, <https://doi.org/10.1038/s41421-020-0153-3>.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733, <https://doi.org/10.1056/NEJMoa2001017>.

Supplemental Information

A multi-pronged approach targeting SARS-CoV-2

proteins using ultra-large virtual screening

Christoph Gorgulla, Krishna M. Padmanabha Das, Kendra E. Leigh, Marco Cespugli, Patrick D. Fischer, Zi-Fu Wang, Guilhem Tesseyre, Shreya Pandita, Alec Shnapir, Anthony Calderaio, Minko Gechev, Alexander Rose, Noam Lewis, Colin Hutcheson, Erez Yaffe, Roni Luxenburg, Henry D. Herce, Vedat Durmaz, Thanos D. Halazonetis, Konstantin Fackeldey, J.J. Patten, Alexander Chuprina, Igor Dziuba, Alla Plekhova, Yurii Moroz, Dmytro Radchenko, Olga Tarkhanova, Irina Yavnyuk, Christian Gruber, Ryan Yust, Dave Payne, Anders M. Näär, Mark N. Namchuk, Robert A. Davey, Gerhard Wagner, Jamie Kinney, and Haribabu Arthanari

Transparent Methods.

***In silico* screening.** VirtualFlow was used to screen around 1 billion on-demand (synthesizable) compounds from the Enamine REAL library and around 10 million in-stock compounds from the ZINC 15 library (Sterling and Irwin, 2015) for each targeted site. Both libraries were previously prepared with 'VirtualFlow for Ligand Preparation' (VFLP) (Gorgulla et al., 2020a). For each virtual screen a single receptor structure was used, and the protein was held rigid. QuickVina W was used to perform a blind docking procedure (Hassan et al., 2017) for the HR1 domain of the spike protein (screen ID: 8), the RNA binding interface of the nucleoprotein (screen ID: 42), the RNA binding site of nsp12 (screen IDs: 28 and 29), as well as for nsp7 (screen ID: 20) and ORF7a (screen ID: 9). For all the other docking routines, we used QuickVina 2 (Alhossary et al., 2015). Both docking programs are based on AutoDock Vina (Trott and Olson, 2010), and the exhaustiveness parameter was set to 1 during all virtual screens. This setting minimizes the search of the conformational space within the docking box and maximizes the computational efficiency. The receptor structures were prepared with AutoDockTools (Morris et al., 2009) from the PDB format to the PDBQT format.

To conduct VirtualFlow screens we needed to leverage a significant amount of high performance computing (HPC) time and infrastructure in order to complete multiple virtual screens in parallel. For this purpose we have taken advantage of Slurm (<https://slurm.schedmd.com>), a highly scalable and fault-tolerant workload manager capable of handling multiple jobs in parallel. In addition, the Elastifile Cloud File Storage (ECFS) was used, which is a POSIX-compliant and highly scalable NFS file service mounted to the Slurm cluster compute nodes. ECFS provides a full suite of cloud-native NAS features (including snapshots and multi-zone accessibility). The Google Cloud Platform (GCP, <https://cloud.google.com>), which combines all of these components, including Google Compute Engine preemptible virtual machines, was used to provide the scale of computation required for the ultra-large virtual screens reported here. Up to 160,000 CPUs (Intel Xeon or AMD EPYC) were used in parallel per virtual screen. Google Cloud Storage was used to durably store the generated data.

We operated within a Google Cloud project that has a single Virtual Private Cloud to provide networking functionality to all the virtual instances. To set up this infrastructure we leveraged automated deployment tools like Terraform (<https://www.terraform.io/>) and the Google Cloud Deployment Manager to replicate this architecture across 3 different zones belonging to two different geographic regions. This resulted in three different clusters (Figure S52) that were used for virtual screens against different targets in parallel. This globally distributed architecture allowed us to run hundreds of thousands of cores in parallel, consuming a total of approximately 100 million CPU hours in a total of about 3 weeks for all the virtual screens reported here. The types of nodes used are listed in Table S53.

Filtering the hits. The virtual hits resulting from these screens were filtered using the open-source software DataWarrior to remove compounds that have reactive functional groups, a molecular weight greater than 600 daltons, a cLogP greater than 6, and/or more than 10 hydrogen bond acceptors (Sander et al., 2015). A smaller number of non-druglike molecules from the ZINC library were also removed after visual inspection.

In addition, for each target site, we showcase one of the top hits. This hit was chosen based on consideration of a number of parameters including, but not limited to, docking score, 'drug-likeness' (which includes criteria such as Lipinski's rule of five and the tendency to aggregate, as estimated by a fraction of sp³-hybridized carbon atoms between 0.25 and 1), and the presence of PAINS (pan-assay interference compounds) or other potentially toxic moieties. The hits depicted were also chosen to highlight particularly plausible small molecule-protein interactions that can be accommodated by the docking site. Due to these filtering criteria, it is therefore important to note that although the example hit molecule displayed in each figure is one of the top hits for the represented screen, it is often not the hit with the highest docking score.

Docking score and binding affinity. In *in silico* screening efforts the docking score is the metric that informs on how well a molecule is predicted to bind to the target protein. The docking score of the AutoDock Vina scoring function is expressed in kcal/mol and represents the free energy of binding ΔG , which can be converted into the binding affinity (dissociation constant, K_d) using the equivalence provided by the Gibbs equation $\Delta G = -RT \ln K_d$, where R is the ideal gas constant and T is the temperature (Keserü and Swinney, 2015). Docking programs like the ones we have used (QuickVina 2, QuickVina-W) are calibrated with respect to the free energy of binding based on experimental binding data, as they use the original AutoDock Vina scoring function (Alhossary et al., 2015; Hassan et al., 2017; Trott and Olson, 2010). This implies that a docking score of -10.0 represents a predicted binding free energy of -10.0 kcal/mol, which can be converted into a dissociation constant using the above equation. We used the docking score to rank the hits and associated it with a corresponding binding affinity, a metric commonly used by experimental researchers. However, it should be noted that there are inconsistencies between computationally calculated binding affinities and experimentally derived ones. This is partly due to the fact that the computational method, especially in the least exhaustive mode required to ensure high-throughput, cannot capture all the factors that contribute to binding and therefore can only deliver an estimation of the

affinity. The docking score and associated binding constant should only be taken as comparative metrics in the absence of further experimental validation.

Enrichment and re-docking verification studies. We carried out re-docking studies of known ligands to some of the SARS-CoV-2 proteins that were targeted in this work to verify whether the docking scenarios we used worked in principle. At the time of this work, only a few identified ligands with experimental ligand-complex structures were publicly available. Furthermore, it is important to note that the published studies do not match the scale of the screens that we have carried out. This is an important point as the scale improves the quality of the top scoring compounds (Gorgulla, 2018; Gorgulla et al., 2020a; Lyu et al., 2019).

In the re-dockings, the same docking program (QuickVina 2) and docking settings were used as in the primary virtual screens (e.g., exhaustiveness 1). A docked compound was classified as successfully reproducing the experimental binding geometry when the RMSD in comparison to the experimental ligand binding mode was below 2.5 Å. Based on this definition, 10 out of the 29 re-docked compounds were successful, which corresponds to a success rate of roughly 35%.

PL^{pro}. For PL^{pro} a total of 10 ligands were re-docked, based on the following co-crystal structures: PDBs 7cjm, 7d7t, 7jir, 7jit, 7jiv, 7jiw, 7jn2, 7jrn, 7koj, 7kol. All of the ligands were reported to bind to the accessory pocket, so the docking scenario used in screen ID 13 was used for the re-docking. The obtained RMSD values were as follows: 7cjm: 8.64 Å, 7d7t: 3.24 Å, 7jir: 9.10 Å, 7jit: 0.92 Å, 7jiv: 2.22 Å, 7jiw: 7.61 Å, 7jn2: 10.15 Å, 7jrn: 0.35 Å, 7koj: 6.13 Å, 7kol: 1.45 Å.

M^{pro}. For M^{pro} a total of 10 ligands were re-docked, based on the following co-crystal structures: PDBs: 5ret, 5reu, 5rev, 5rf7, 7c7p, 7c8t, 7cx9, 7d3i, 7jkv, and 7k40. The obtained RMSD values were as follows: 5ret: 4.64 Å, 5reu: 3.63 Å, 5rev: 6.81 Å, 5rf7: 1.87 Å, 7c7p: 5.06 Å, 7c8t: 6.75 Å, 7cx9: 4.33 Å, 7d3i: 4.42 Å, 7jkv: 2.41 Å, and 7k40: 2.37 Å.

Macrodomain. For the macrodomain a total of 2 ligands were re-docked, based on the following co-crystal structures: PDBs 6wcf and 7jme. The obtained RMSD values were as follows: 6wcf: 1.57 Å, and 7jme: 7.78 Å.

nsp15. For nsp15 a total of 5 ligands were re-docked, based on the following co-crystal structures: PDBs 6w01, 6wlc, 6x1b, and 6wxc. The obtained RMSD values were as follows: 6w01: 2.33 Å, 6wlc: 3.67 Å, 6x1b: 4.66 Å, and 6wxc: 2.95 Å.

nsp16. For nsp16 a total of 4 ligands were re-docked, based on the following co-crystal structures: PDBs 7jhe, 6wvn, and 6yz1. The obtained RMSD values are as follows: 7jhe: 2.25 Å, 6wvn: 3.26 Å, and 6yz1: 3.63 Å. The re-docking results related to PDB 7jhe are illustrated in Figure S49.

We carried out enrichment studies for M^{pro} and PL^{pro}, in which 10 experimentally confirmed compounds from the literature for each protein were used as positive hits. For each of these positive hit compounds, approximately 50 decoys were obtained from the DUD-E database (Huang et al., 2006b; Mysinger et al., 2012), resulting in a total of 500 decoys for each of the two target proteins. The experimentally confirmed compounds used are listed in Table S50. Docking of the decoys and of the original positive hit compounds was carried out with the same settings as were used in the original virtual screens (screen IDs: 13 and 16). For M^{pro} an enrichment factor at 2% (EF₂) of 20.0 was obtained, and for PL_{pro} an EF₂ of 25.0 was obtained. The maximum possible EF₂ is approximately 50 for each of the two target proteins.

For comparison we also carried out enrichment studies for both M_{pro} (screen ID: 16) and PL_{pro} (screen ID: 13) where the positive/active ligands were the top 10 virtual screening hits for each of the targets (see Table S51). Again, for each ligand we obtained 50 decoys from the DUD-E database, resulting in a total of 500 decoys per target site. In both cases the EF₂ equaled 50.0, which is the theoretical maximum value possible for these simulations. Although limited in their scope for interpretation, these results demonstrate the effect of scale on the quality of the virtual screening results. Similar results would be expected for the other target sites due to the extraordinarily high docking scores of the top scoring compounds, which stem from the ultra-large scale of the screens.

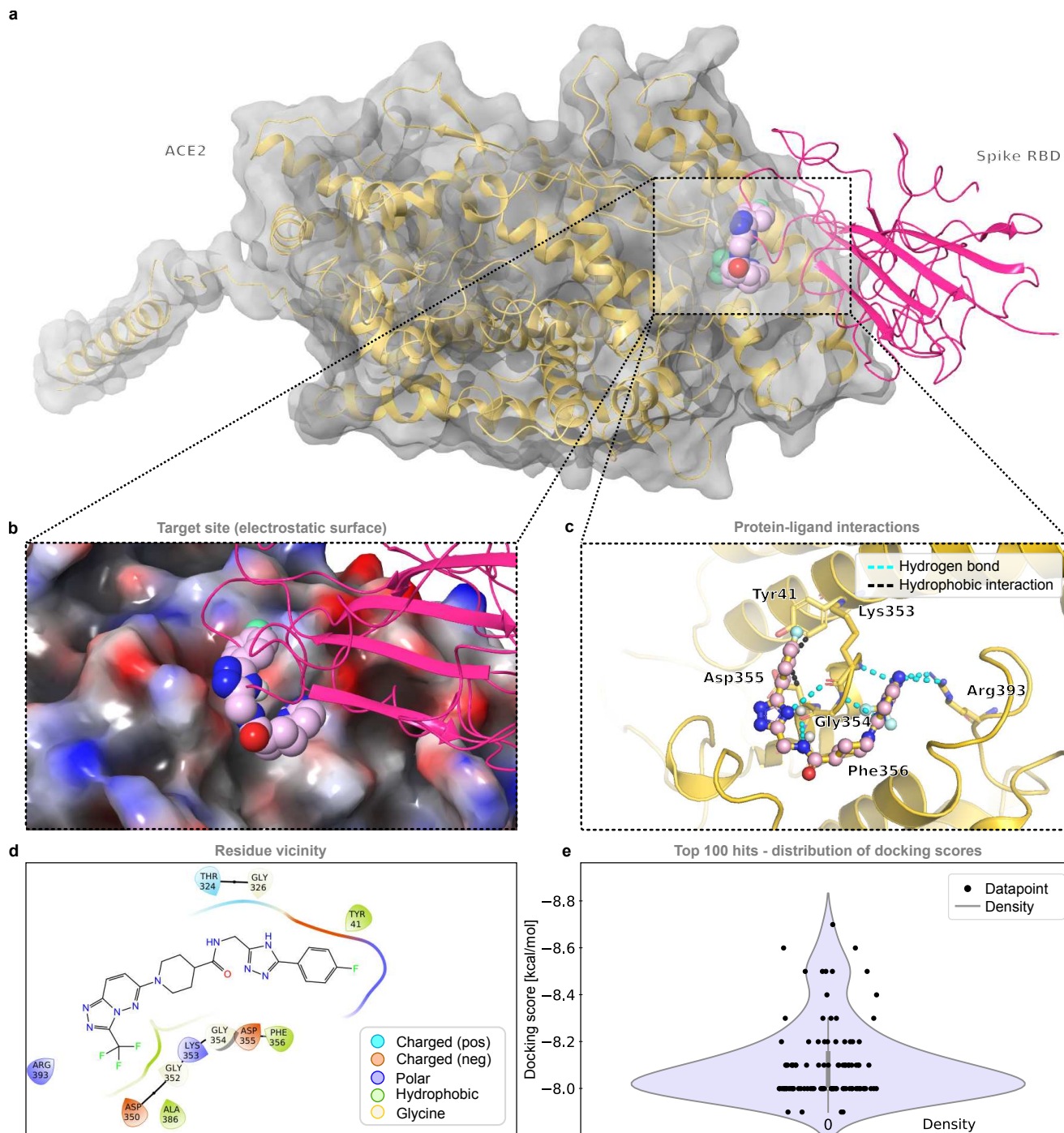


Figure S3. The ACE2 receptor (PDB 6m18) and an example compound from the top 0.0001% of screened compounds bound at the spike protein interaction interface (site 1). **a**, The target protein ACE2 (gold) bound to an example compound (light pink) from the virtual screen at the interaction interface with the RBD of the spike protein (magenta, screen ID: 2, site 1, around Glu37). **b**, Electrostatic surface of ACE2 to which the RBD domain of the spike protein (magenta) and an example compound (light pink) are bound. **c**, An overview of the interactions between the compound and the ACE2 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

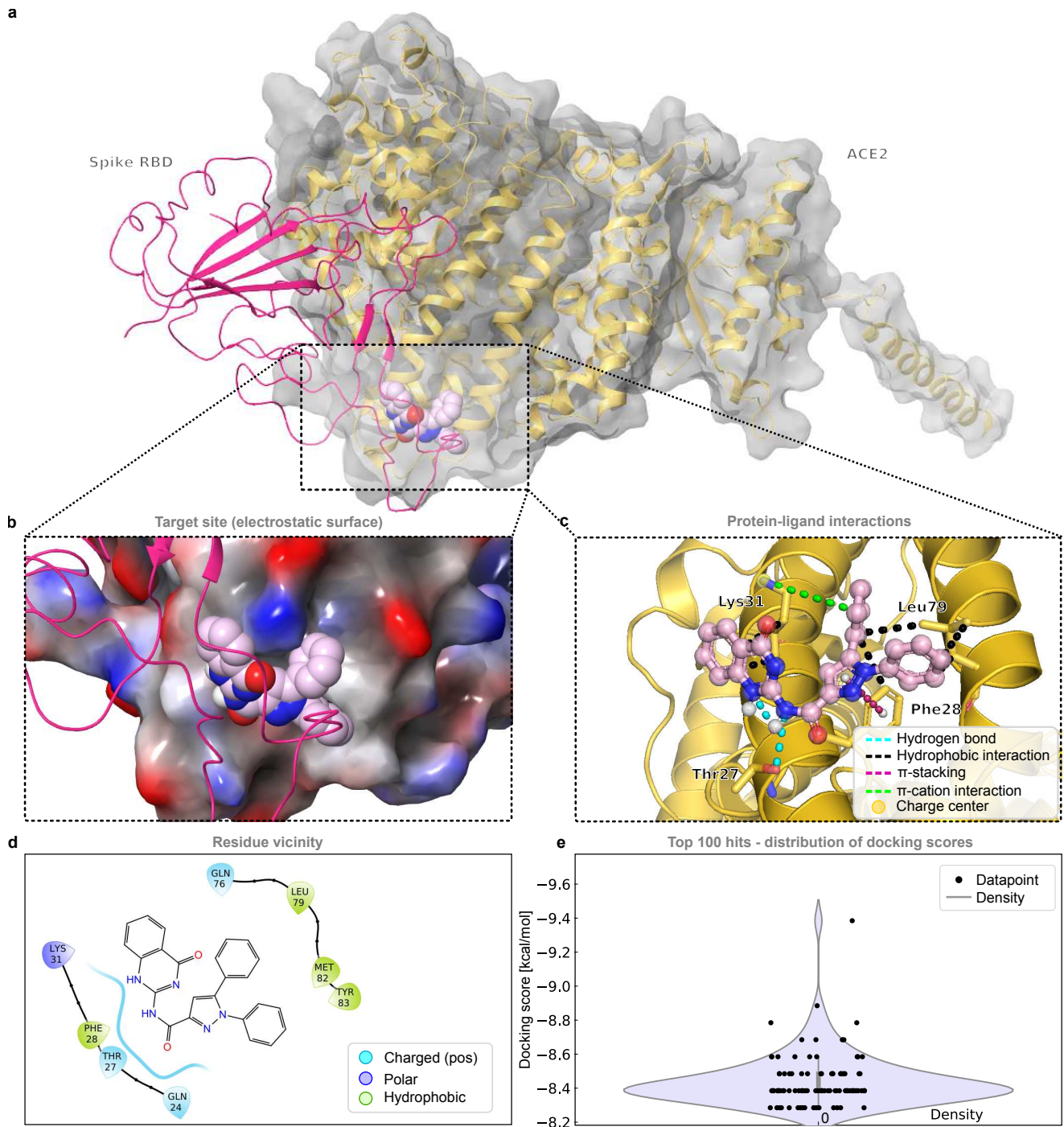


Figure S4. The ACE2 receptor (PDB 6m17) and an example compound from the top 0.0001% of screened compounds bound at the spike protein interaction interface (site 2). Related to Table 1. **a**, The target protein ACE2 (gold) bound to an example compound (light pink) from the virtual screen at the interaction interface with the RBD of the spike protein (magenta, screen ID: 3, site 2, around Phe28). **b**, Electrostatic surface of ACE2 bound to the RBD domain of the spike protein (magenta) and an example compound (light pink). **c**, An overview of the interactions between the compound and the ACE2 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

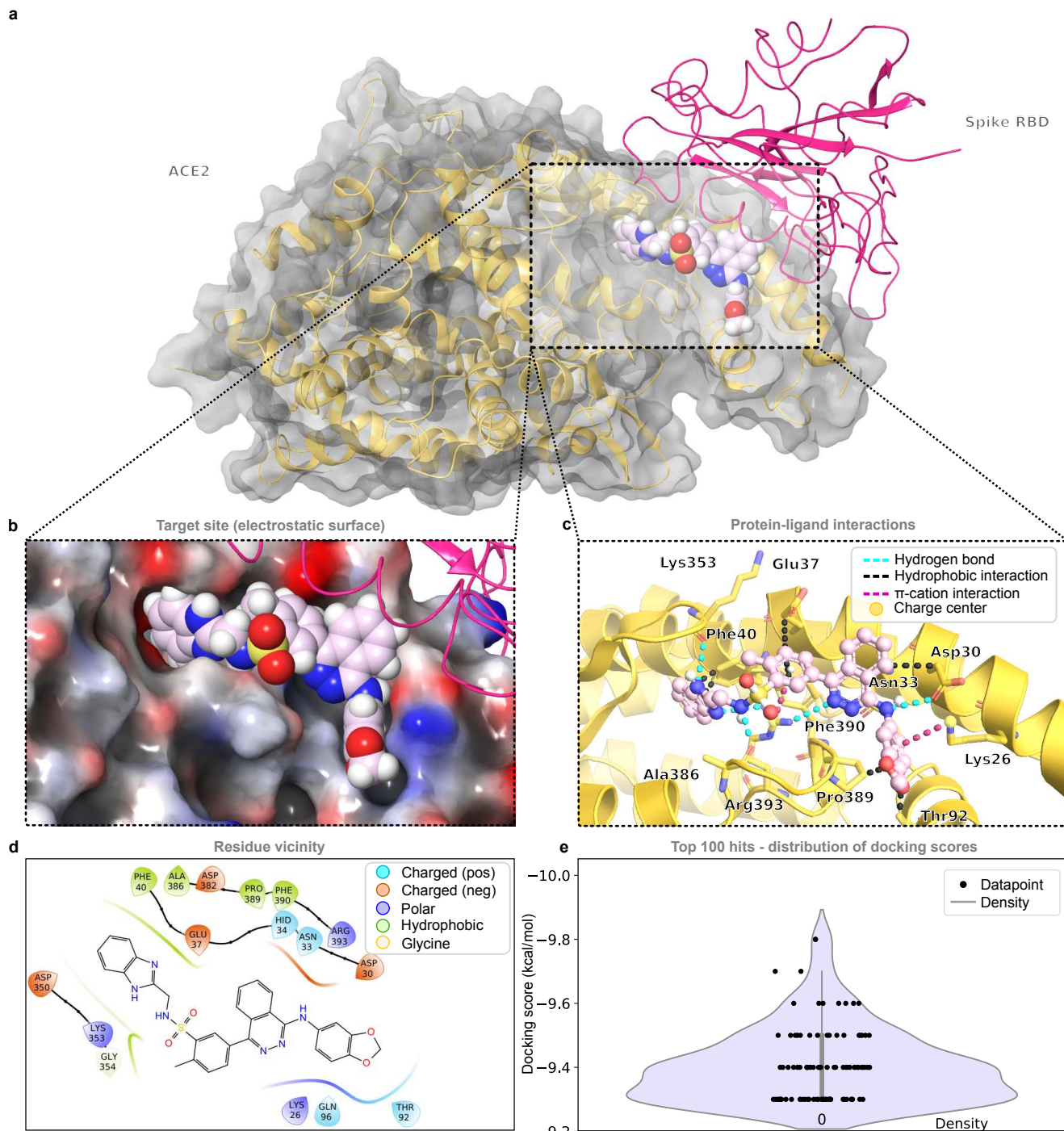


Figure S5. The ACE2 receptor and an example compound from the top 0.0001% of screened compounds bound near the spike interaction interface (dynamic pocket 1). Related to Table 1. **a**, The target protein ACE2 (gold) bound to the RBD of the spike protein (magenta) with an example compound (light pink) from the virtual screen bound at dynamic pocket 1, located adjacent to the RBD interface (around Phe40, screen ID: 4). **b**, Electrostatic surface of the target protein (ACE2) bound to the RBD domain of the spike protein (magenta) and an example compound (light pink). **c**, An overview of the interactions between the compound and the ACE2 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

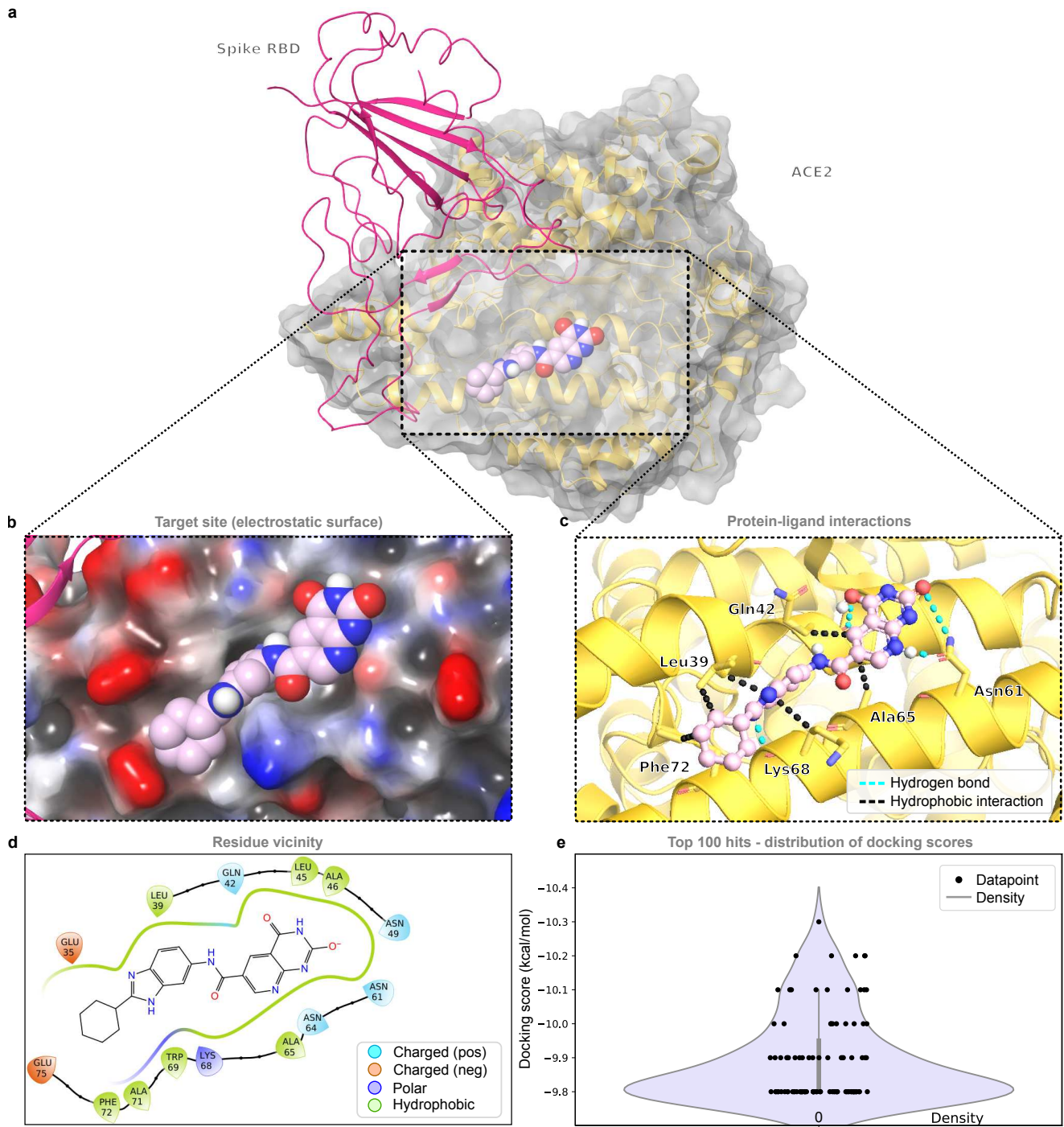


Figure S6. The ACE2 receptor and an example compound from the top 0.0001% of screened compounds bound near the spike interaction interface (dynamic pocket 2). Related to Table 1. **a**, The target protein ACE2 (gold) bound to the RBD of the spike protein (magenta) with an example compound (light pink) from the virtual screen bound to dynamic pocket 2, which is adjacent to the RBD binding interface (around Leu39, screen ID: 5). **b**, Electrostatic surface of the target protein (ACE2) bound to the RBD domain of the spike protein (magenta) and an example compound (light pink). **c**, An overview of the interactions between the compound and the ACE2 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

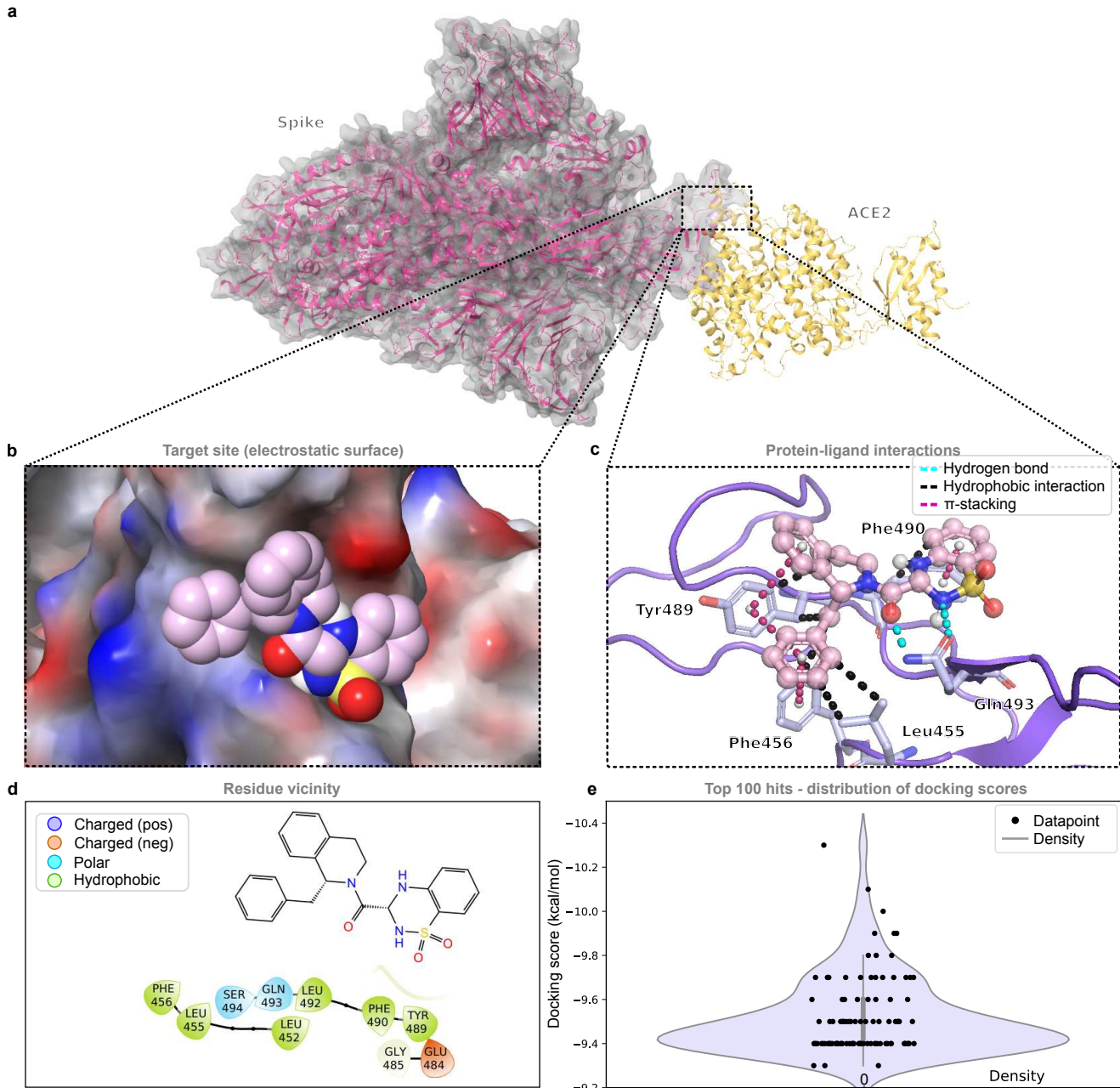


Figure S7. The RBD of the spike protein and an example compound from the top 0.0001% of screened compounds bound at the ACE2 binding interface. Related to Table 1. a, The targeted spike protein (magenta) and an example compound (light pink) from the virtual screen bound to the ACE2 (gold) binding interface (screen ID: 7). **b,** Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c,** An overview of the interactions between the inhibitor and the receptor structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

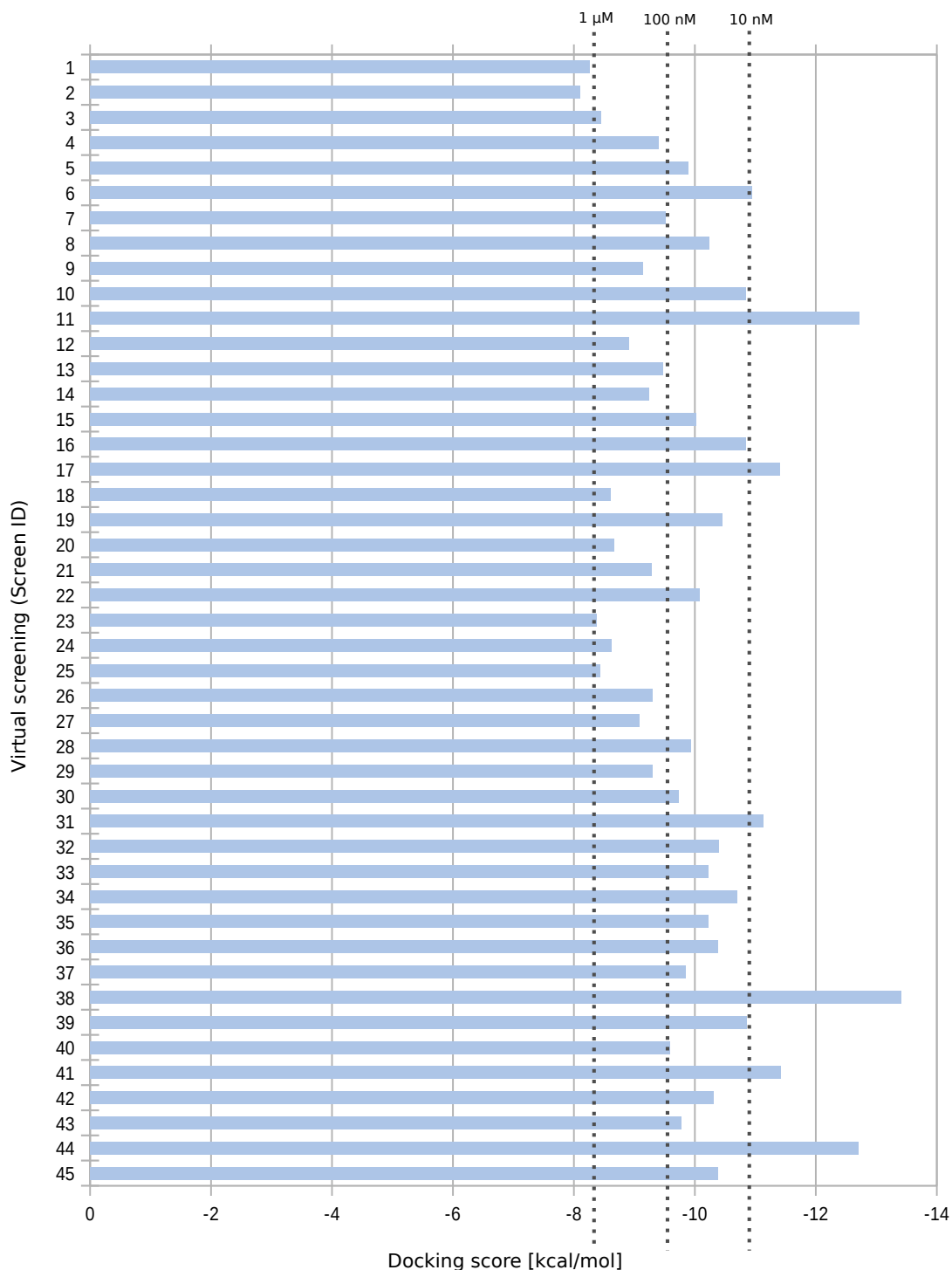


Figure S8. Average docking scores of the top hits. Related to Table 1. Average docking scores of the top 100 compounds of each virtual screen that was carried out. Only compounds with cLogP less than 6, molecular weight less than 600 dalton, no reactive groups, and at most ten hydrogen bond acceptors were considered (a smaller number of non-druglike molecules from the ZINC library were also removed by visual inspection). The docking scores estimate the free energy of binding in kcal/mol. These are computational values that can diverge significantly from the experimental binding free energies. The dashed lines indicate the corresponding K_d values of 1 μ M, 100 nM, and 10 nM under standard conditions.

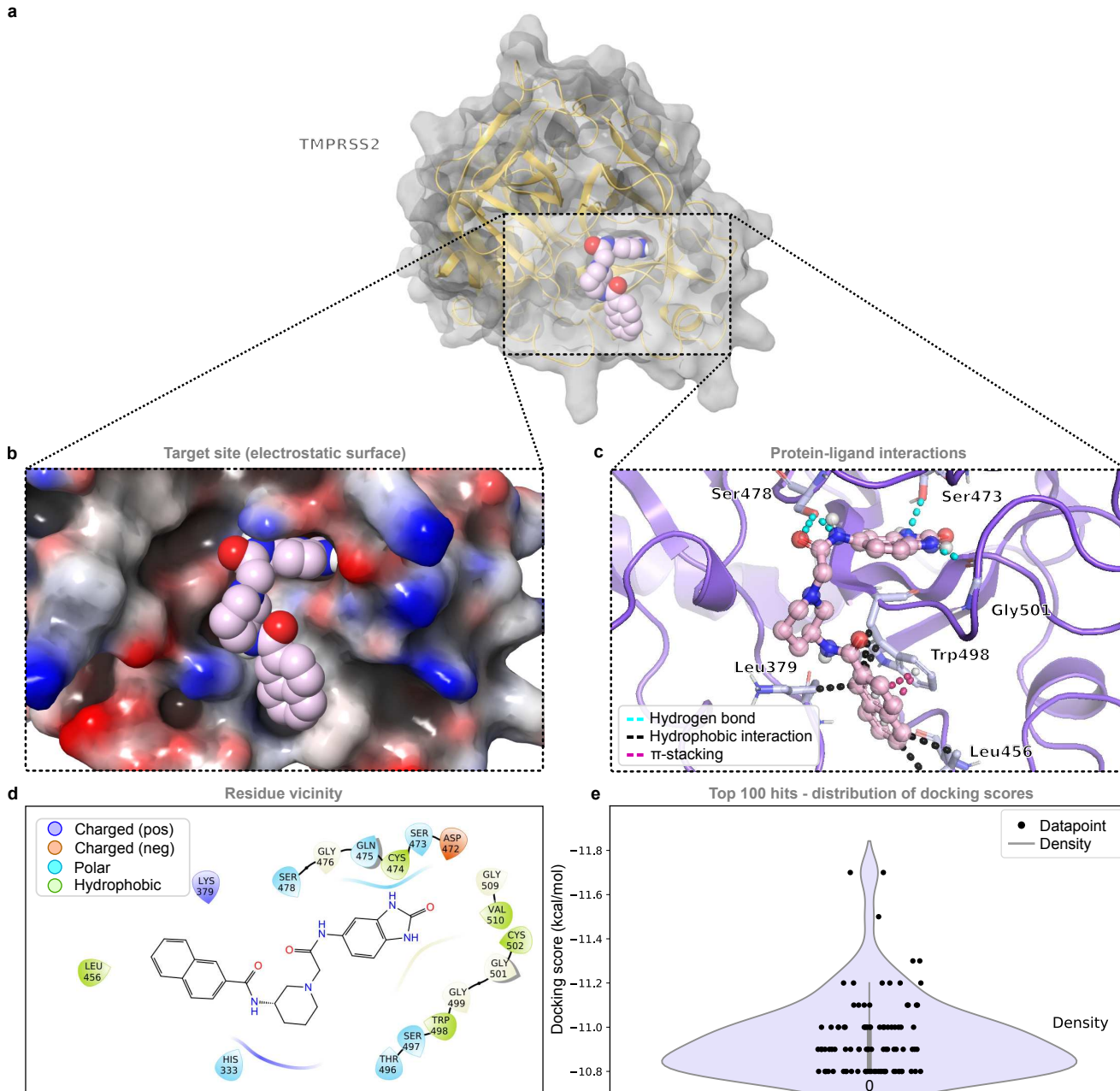


Figure S9. TMPRSS2 and an example compound from the top 0.0001% of screened compounds bound at the active site of the serine protease domain. Related to Table 1. a. The target protein TMPRSS2 (gold) and an example compound (light pink) from the virtual screen bound at the active site of the serine protease domain (screen ID: 6). **b.** Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c.** An overview of the interactions between the inhibitor and protease structure. **d.** Residues within 4 Å of the ligand. **e.** Distribution of the docking scores of the top 100 virtual screening hits.

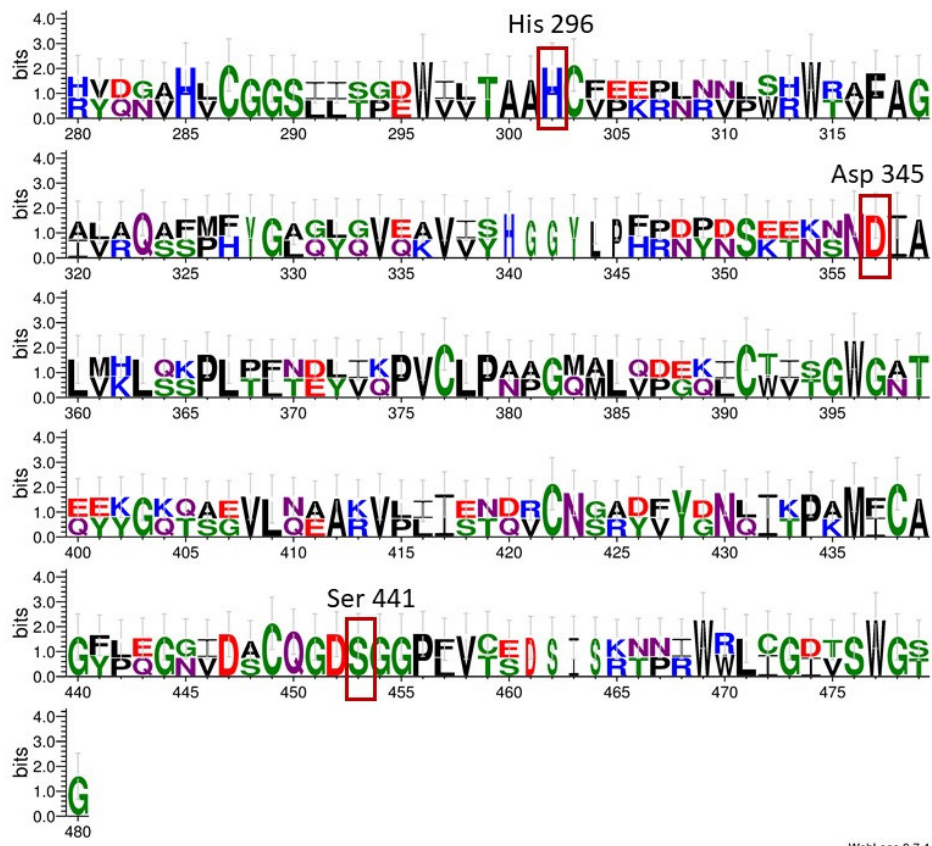


Figure S10. Alignment of the extracellular domain of TMPRSS2 against the extracellular region of the transmembrane serine protease hepsin as represented by WebLogo images (<https://weblogo.berkeley.edu/logo.cgi>). The structure of hepsin (PDB: 1z8g) was used as a template for making the homology model used in Screen 6. The residues which comprise the catalytic triad are marked, and are fully conserved.

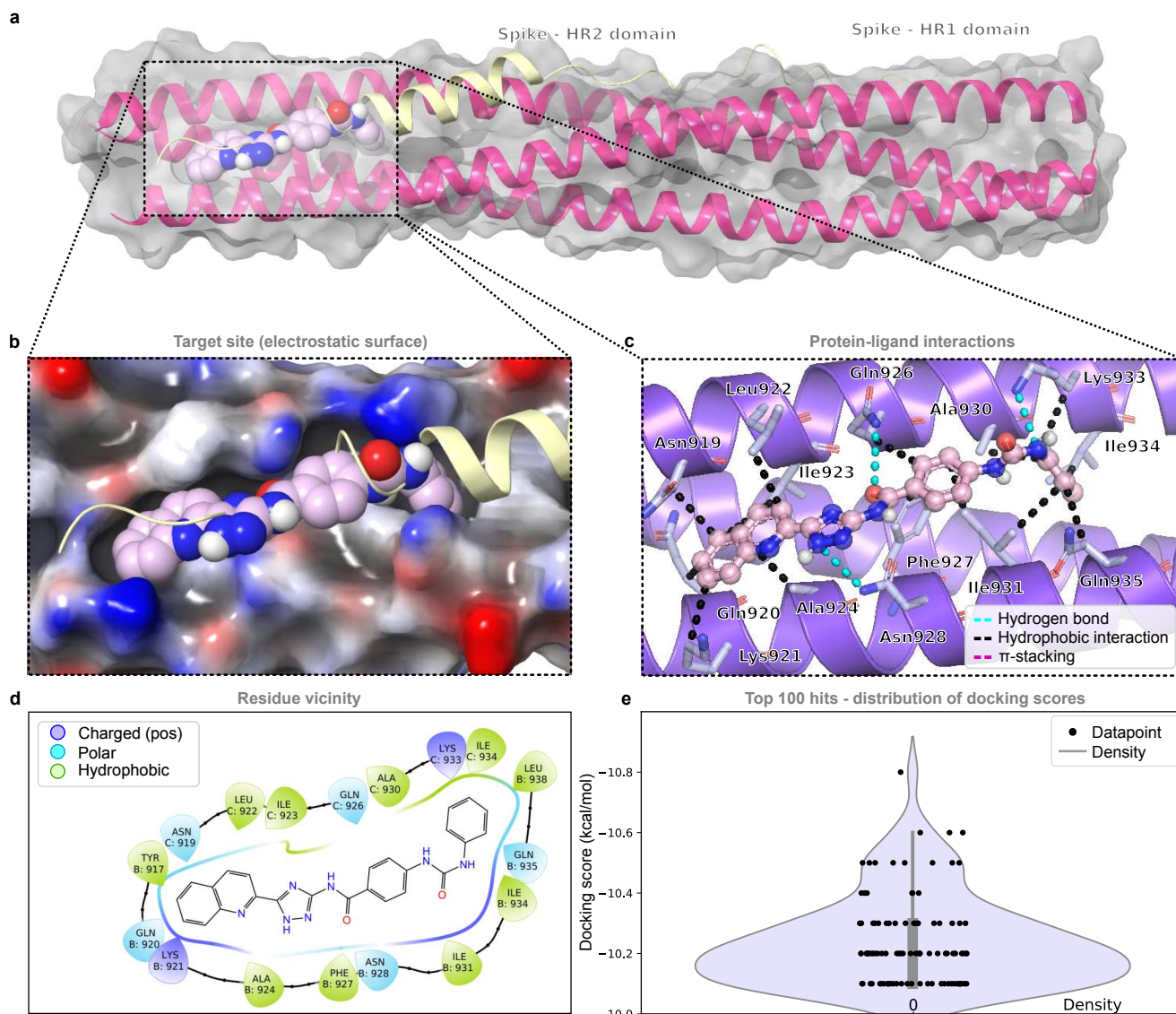


Figure S11. The HR1 domain and an example compound from the top 0.0001% of screened compounds bound at the interaction interface to the HR2 domain of the spike protein. Related to Table 1. a, The HR1 domain of the targeted spike protein (magenta) and an example compound (light pink) from the virtual screen bound to the HR2 interaction interface. The HR2 domain is shown in light gold (screen ID: 8). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the HR1 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

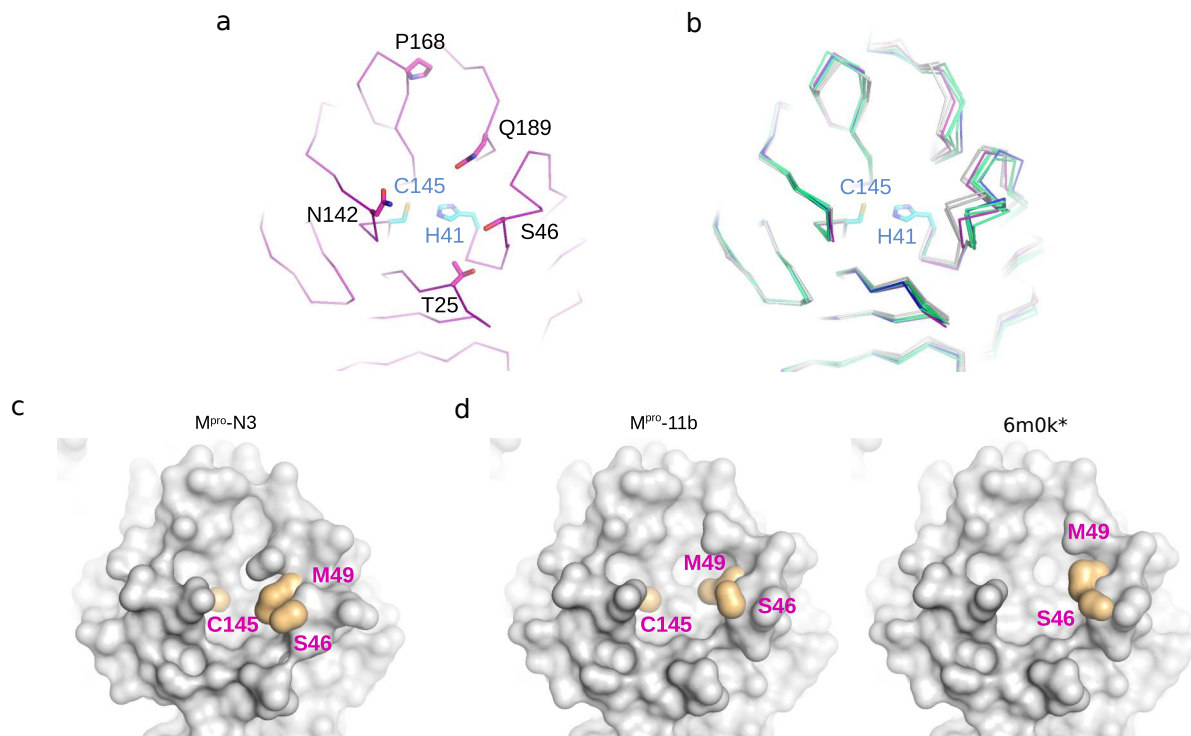


Figure S12. Conformations of the active site of M^{pro} and models for *in silico* screening. Related to Table 1. **a**, Secondary structural elements that form the active site of M^{pro} . The catalytic dyad Cys145 (C145)-His41 (H41) at the base of the active site is surrounded by secondary structural elements contributed by domains I and II of the protease. The side chains of selected residues are shown: Thr25 (T25), Ser46 (S46), Asn142 (N142), Pro168 (P168) and Gln189 (Q189). The structure of the active site shown here has been taken from the structure of M^{pro} bound to the inhibitor N3 (PDB: 6lu7). **b**, Overlay of M^{pro} active site structures showing that binding of inhibitors induces conformational changes in two secondary structural elements that form the walls of the active site: the loop harboring Gln189 (Q189) and the short α -helix harboring Ser46 (S46, see panel **a** for reference). The main chain is colored gray for the three apo structures of M^{pro} (PDB IDs: 6y2e, 6yb7, 6m2q), purple for the structure of M^{pro} bound to the inhibitor N3 (PDB: 6lu7), blue for the structure of M^{pro} bound to the inhibitor 11b (PDB: 6m0k) and green for three other structures of M^{pro} bound to inhibitors (PDBs: 6m2n, 6y2g, 7buy). **c**, Van der Waals surface of the active site of M^{pro} bound to the inhibitor N3 (PDB: 6lu7), after removing all water and ligand molecules. The side chains of Met49 (M49), Ser46 (S46) and Cys145 (C145) are highlighted in yellow. **d**, Van der Waals surface of the active site of M^{pro} bound to the inhibitor 11b (PDB ID: 6m0k), after removing all water and ligand molecules (left) and the same structure after selecting different rotamers for the side chains of Met49, Ser46 and Cys145 (right). The side chains of Met49 (M49), Ser46 (S46) and Cys145 (C145) are highlighted in yellow. The images in panels **a-d** all show the active site of M^{pro} from the same orientation and were generated using PyMOL.

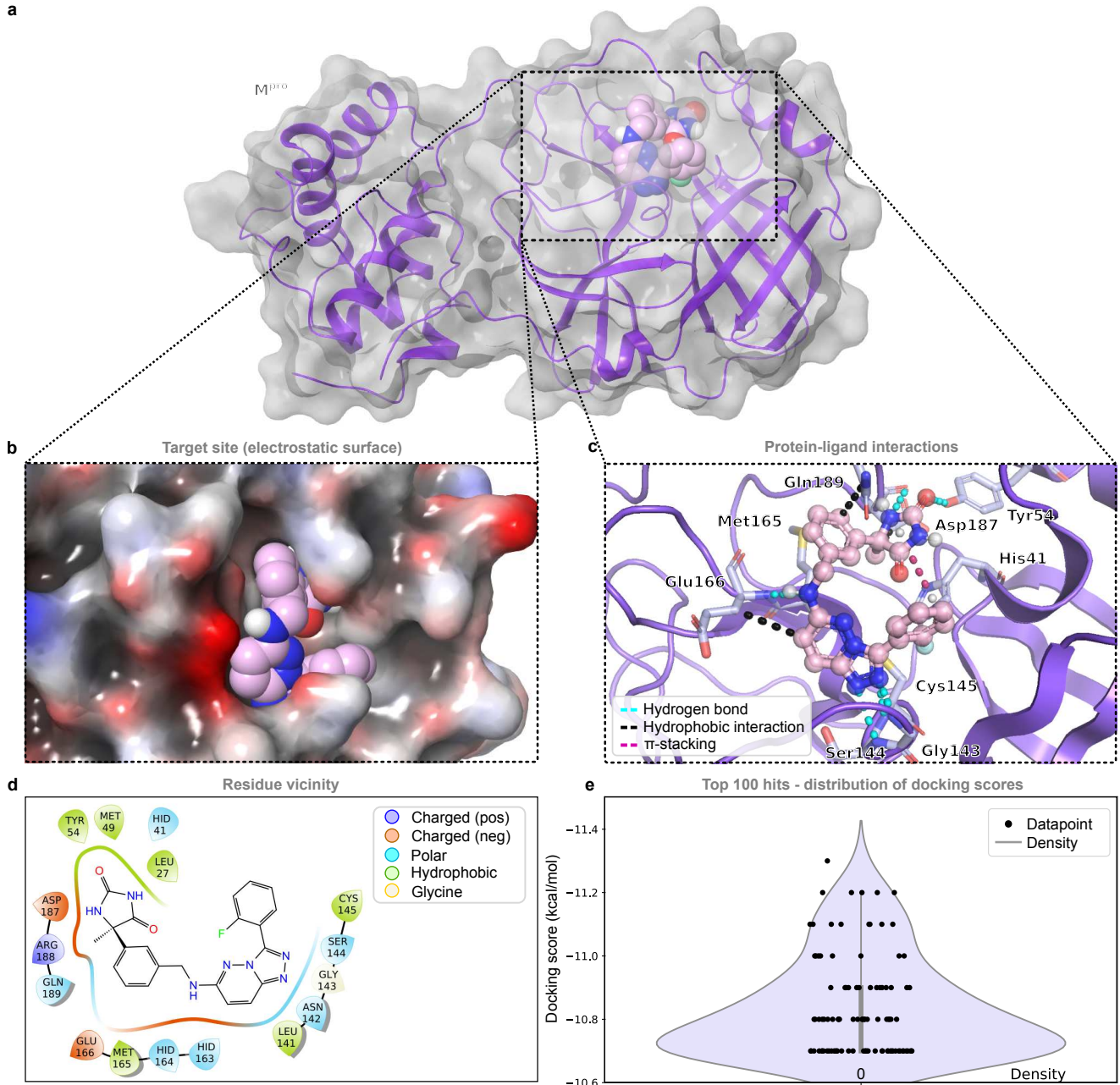


Figure S13. M^{pro} and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site. Related to Table 1. **a**, M^{pro} (violet) and an example compound (light pink) from the virtual screen bound to the active site. Here, the crystal structure deposited as PBD 6lu7 was used (screen ID: 16). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the protease structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

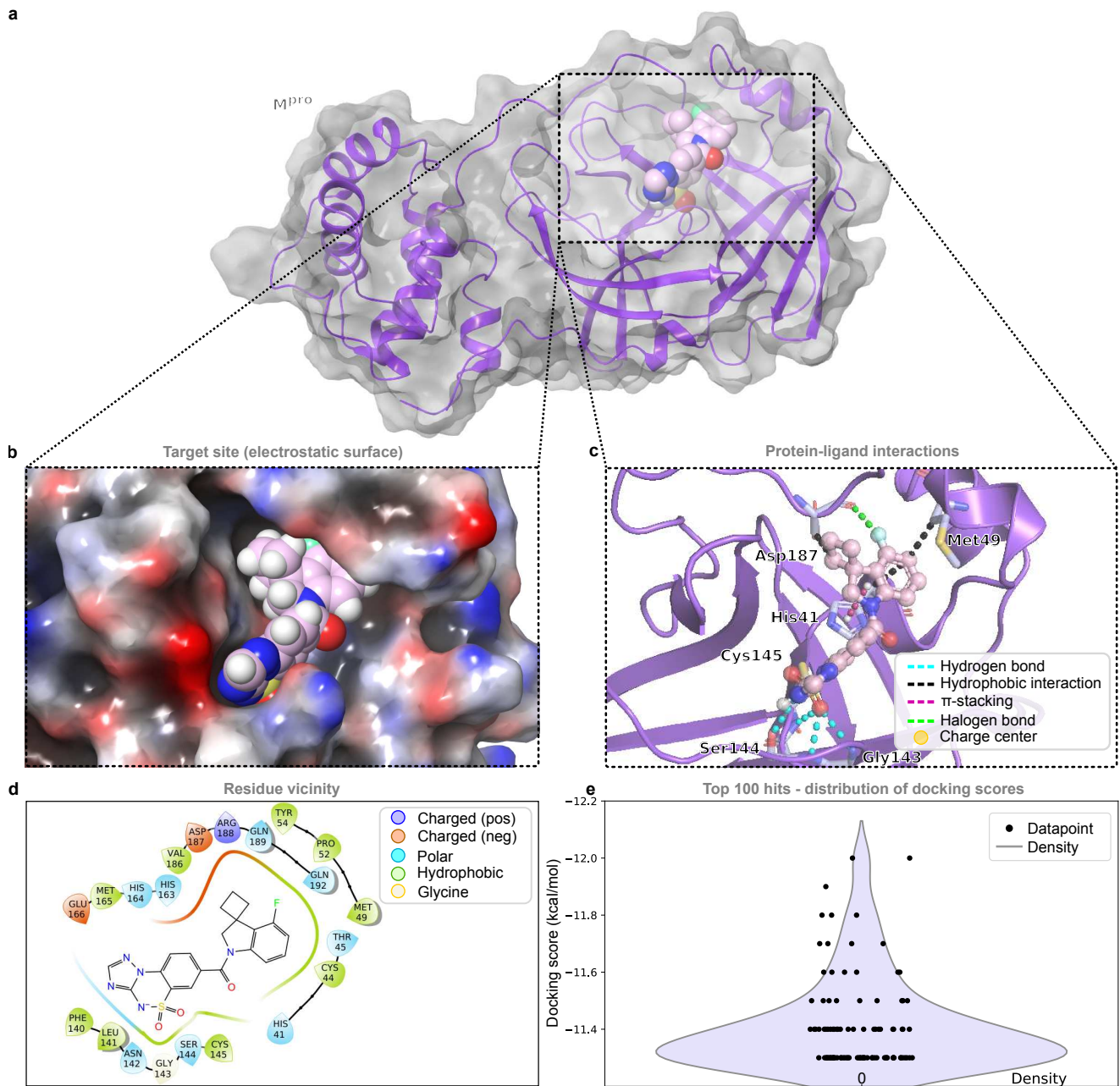


Figure S14. M^{pro} and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site. Related to Table 1. **a**, M^{pro} (violet) and an example compound (light pink) from the virtual screen bound to the active site. Here, a modified version of the PDB 6m0k structure was used (details described in the main text) (screen ID: 17). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the protease structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

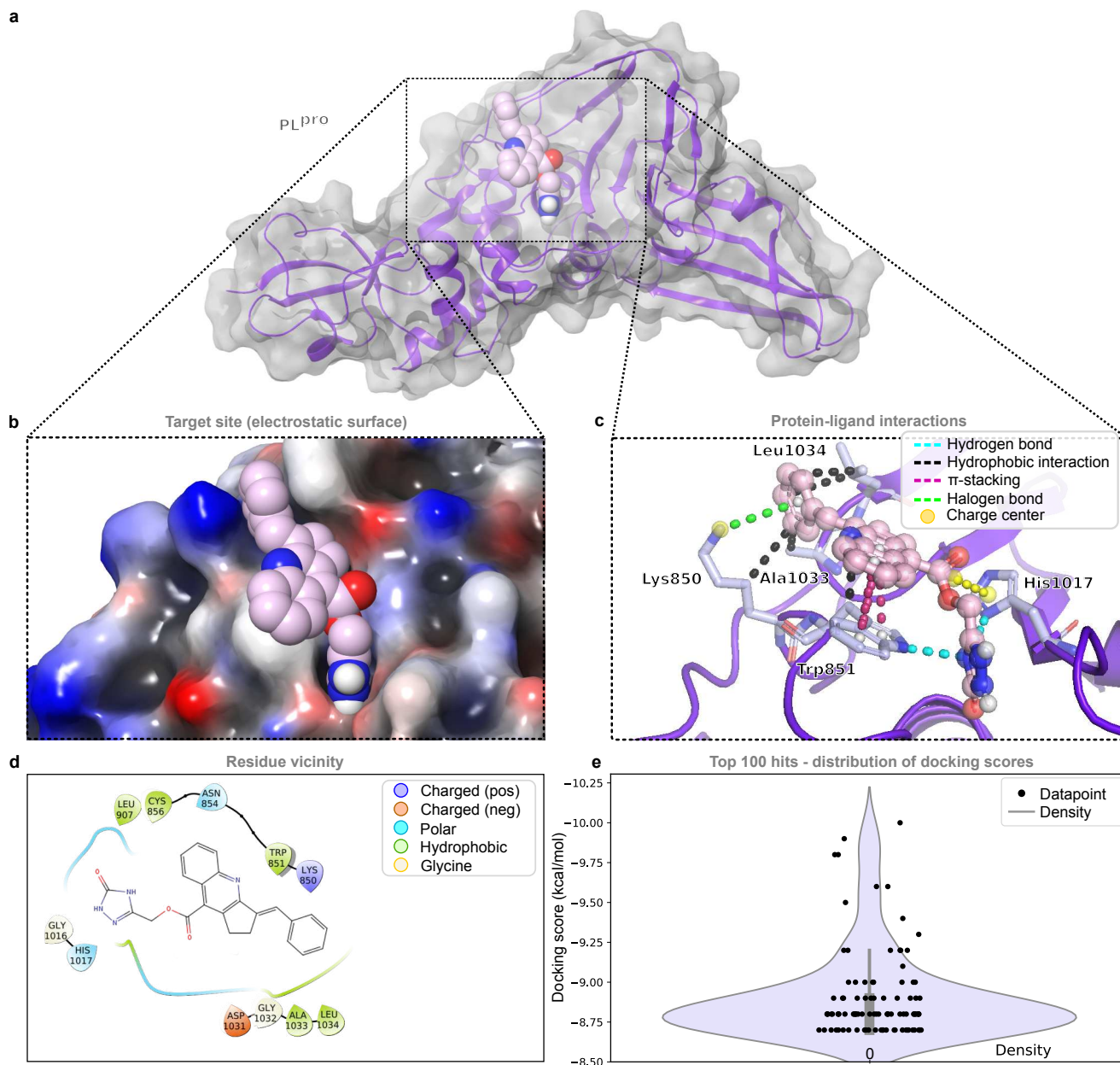


Figure S15. PL^{PRO} and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site. Related to Table 1. **a**, PL^{PRO} (violet) and an example compound (light pink) from the virtual screen bound to the active site (screen ID: 12). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the protease structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

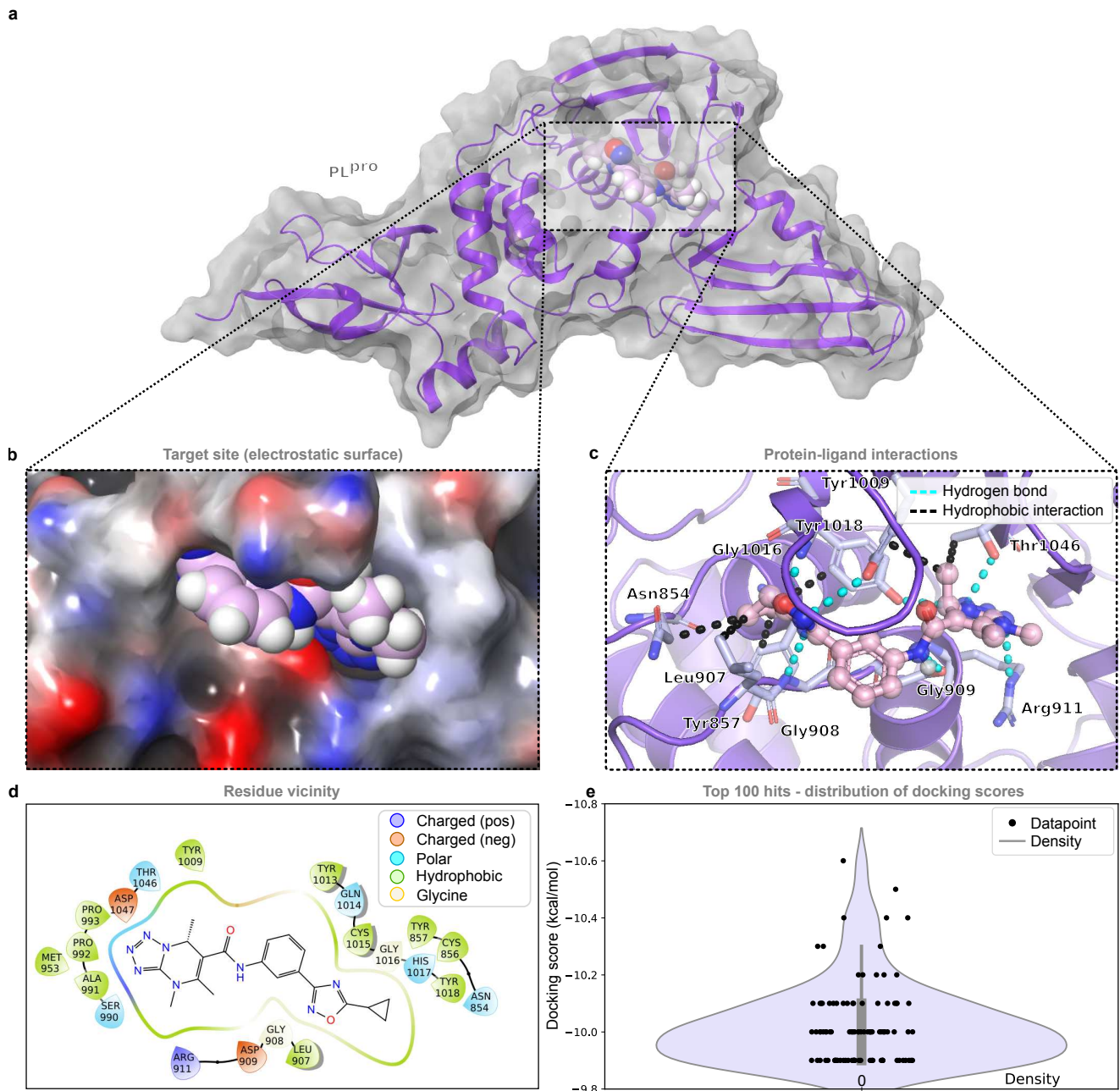


Figure S16. PL^{Pro} and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site (tunnel region). Related to Table 1. **a**, PL^{Pro} (violet) and an example compound (light pink) from the virtual screen bound to the active site (tunnel region) (screen ID: 15). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the protease structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

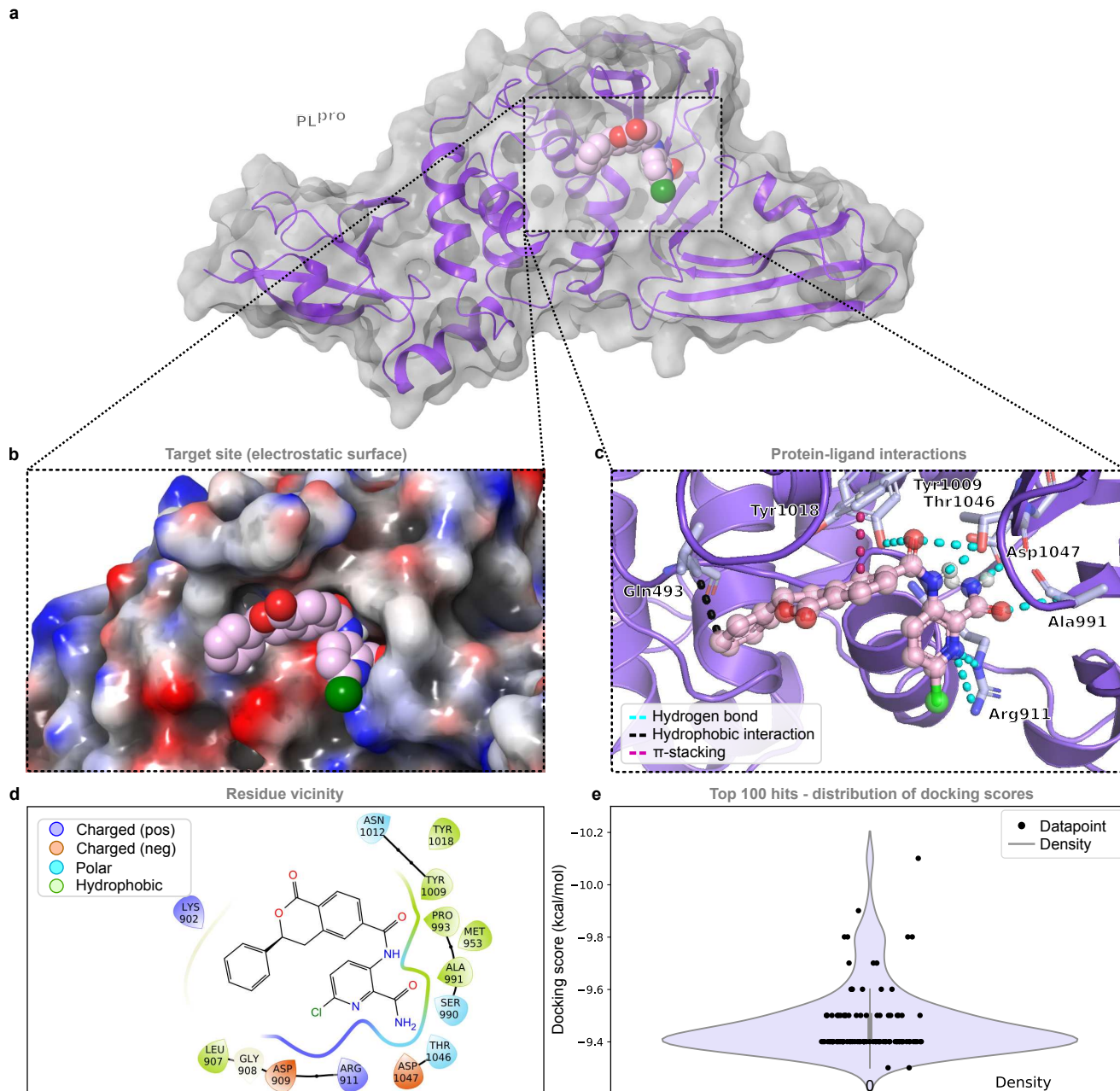


Figure S17. PL^{Pro} and an example compound from the top 0.0001% of screened compounds bound at the accessory pocket near the active site. Related to Table 1. **a**, PL^{Pro} (violet) and an example compound (light pink) from the virtual screen bound at the accessory pocket near the active site (screen ID: 13). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the protease structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

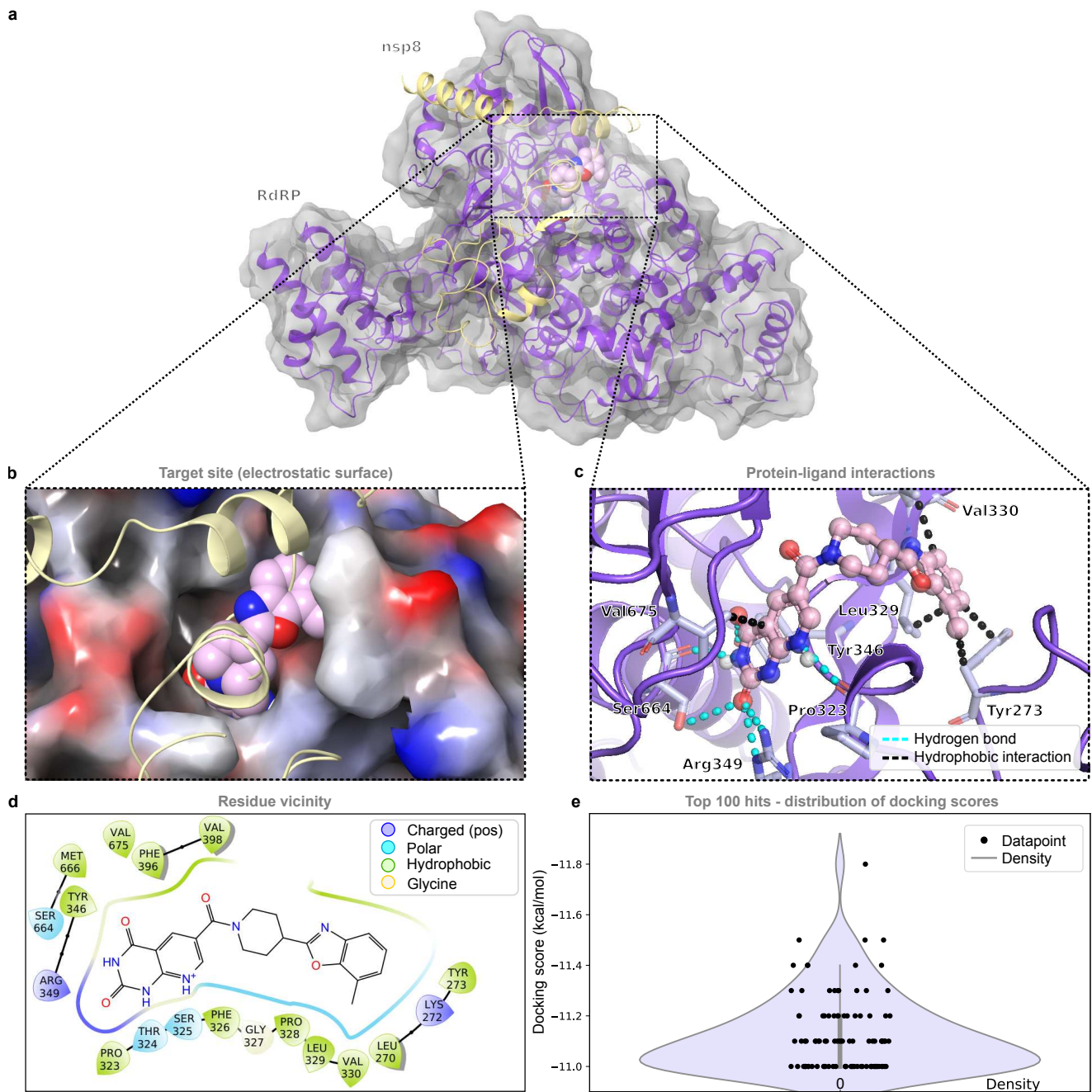


Figure S18. nsp12 and an example compound from the top 0.0001% of screened compounds bound at the nsp8 binding interface. Related to Table 1. **a**, nsp12 (violet) bound to nsp8 (light gold) and an example compound (light pink) at the nsp8 binding interface of nsp12 (screen ID: 31). **b**, Electrostatic surface of nsp12 to which an example compound (light pink) as well as nsp8 (light gold) are bound. **c**, An overview of the interactions between the inhibitor and the nsp12 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

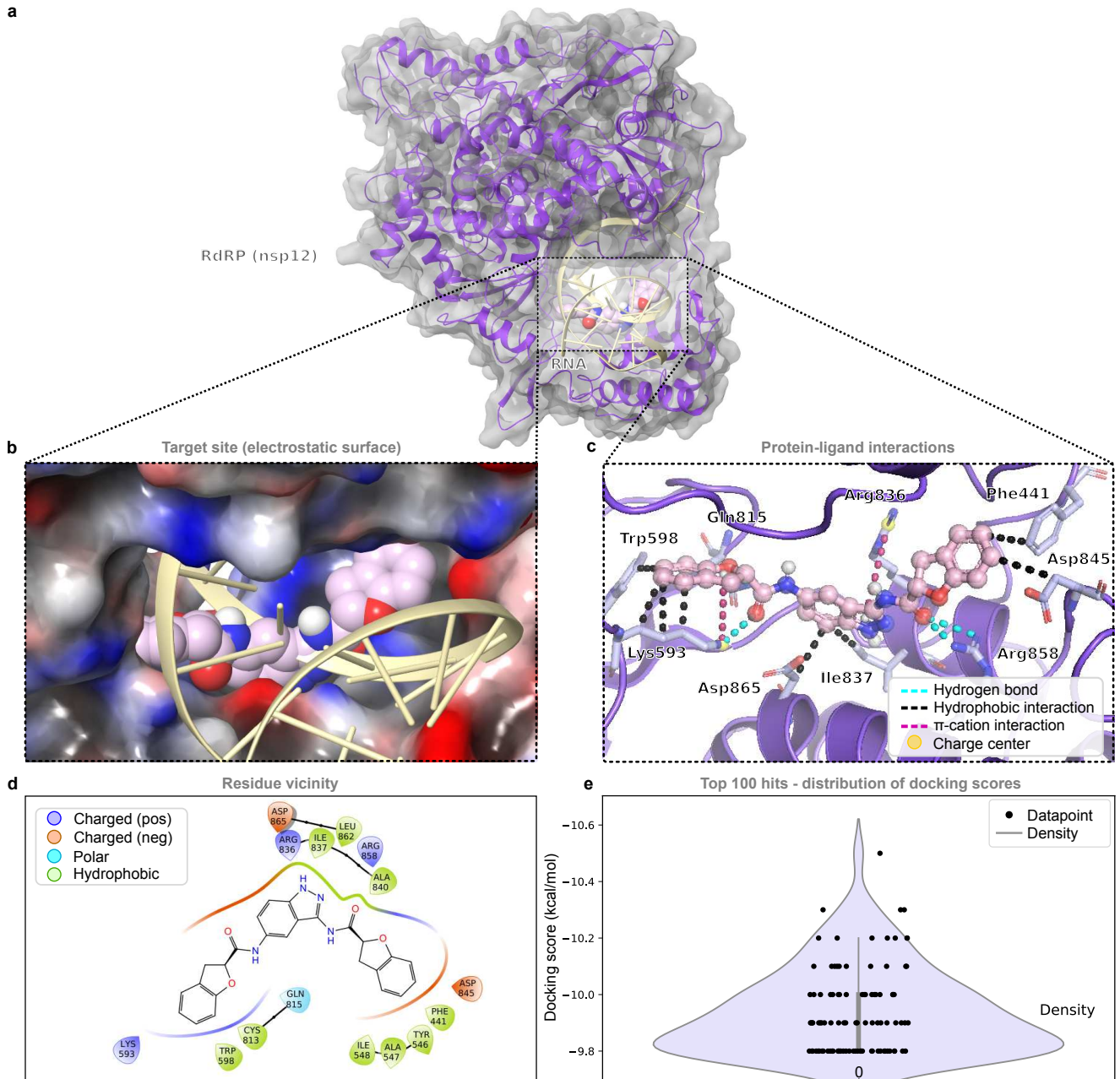


Figure S19. nsp12 and an example compound from the top 0.0001% of screened compounds bound at site 1 of the RNA binding interface. Related to Table 1. **a**, nsp12 (violet) bound to RNA (light gold) and an example compound (light pink) at site 1 of the RNA binding interface of nsp12 (screen ID: 28). **b**, Electrostatic surface of nsp12 to which an example compound (light pink) is bound at site 1 of the RNA (light gold) binding interface. **c**, An overview of the interactions between the inhibitor and the nsp12 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

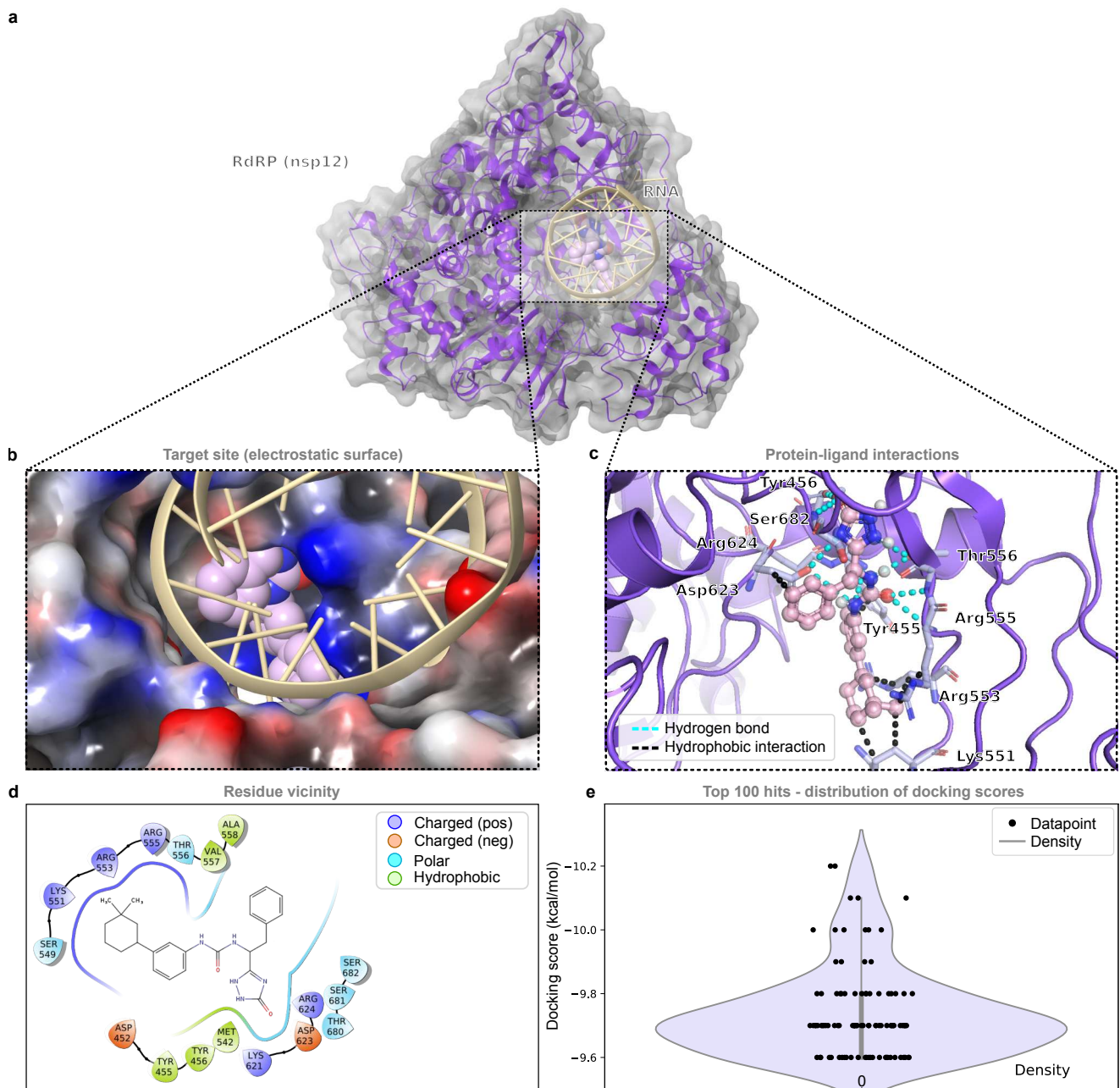


Figure S20. nsp12 and an example compound from the top 0.0001% of screened compounds bound at the nucleotide binding site. Related to Table 1. **a**, nsp12 (violet) bound to RNA (light gold) and an example compound (light pink) at the nucleotide binding site (screen ID: 30). **b**, Electrostatic surface of nsp12 to which an example compound (light pink) is bound at the nucleotide binding site. RNA is depicted in light gold. **c**, An overview of the interactions between the inhibitor and the nsp12 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

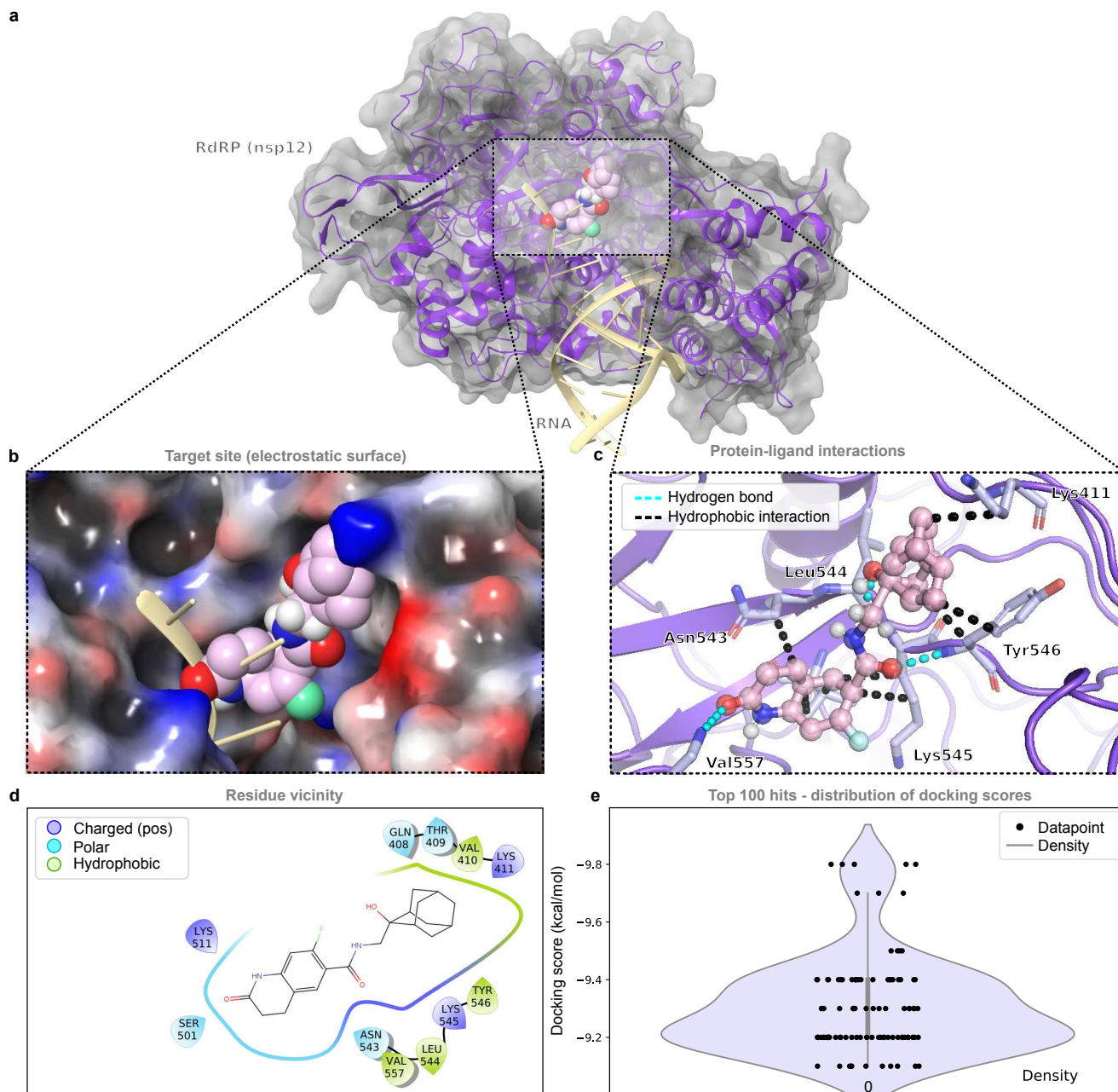


Figure S21. nsp12 and an example compound from the top 0.0001% of screened compounds bound at site 2 of the RNA binding interface. Related to Table 1. **a**, nsp12 (violet) bound to RNA (light gold) and an example compound (light pink) at site 2 of the RNA binding region (screen ID: 29). **b**, Electrostatic surface of nsp12 to which an example compound (light pink) at site 2 of the RNA binding region. RNA is shown in light gold. **c**, An overview of the interactions between the inhibitor and the nsp12 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

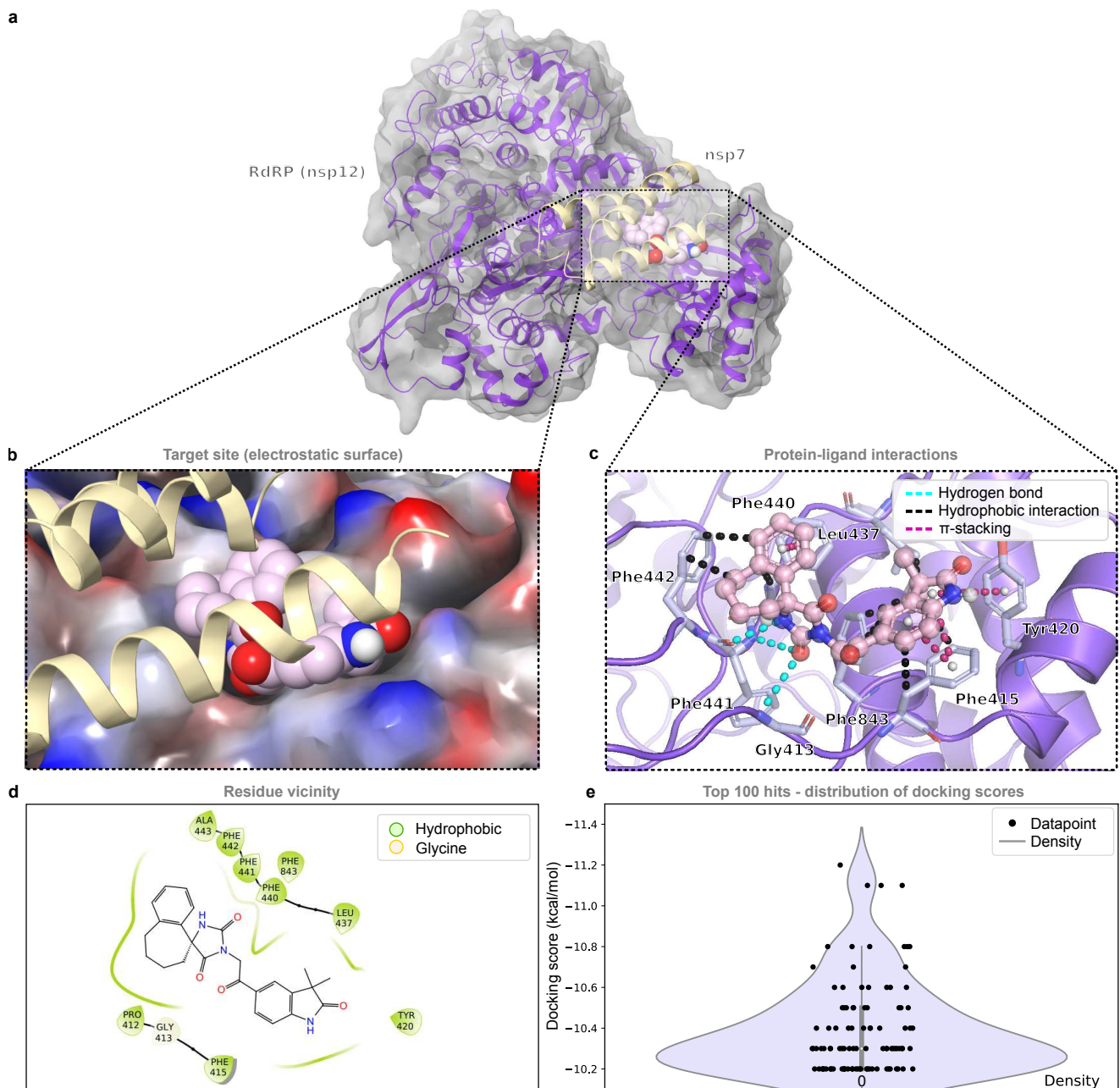


Figure S22. nsp12 and an example compound from the top 0.0001% of screened compounds bound at the nsp7 binding interface. Related to Table 1. **a**, nsp12 (violet) bound to nsp7 (light gold) and an example compound (light pink) at the nsp7 binding interface (screen ID: 32). **b**, Electrostatic surface of nsp12 to which an example compound (light pink) is bound to the nsp7 binding site. nsp7 is shown in light gold. **c**, An overview of the interactions between the inhibitor and the nsp12 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

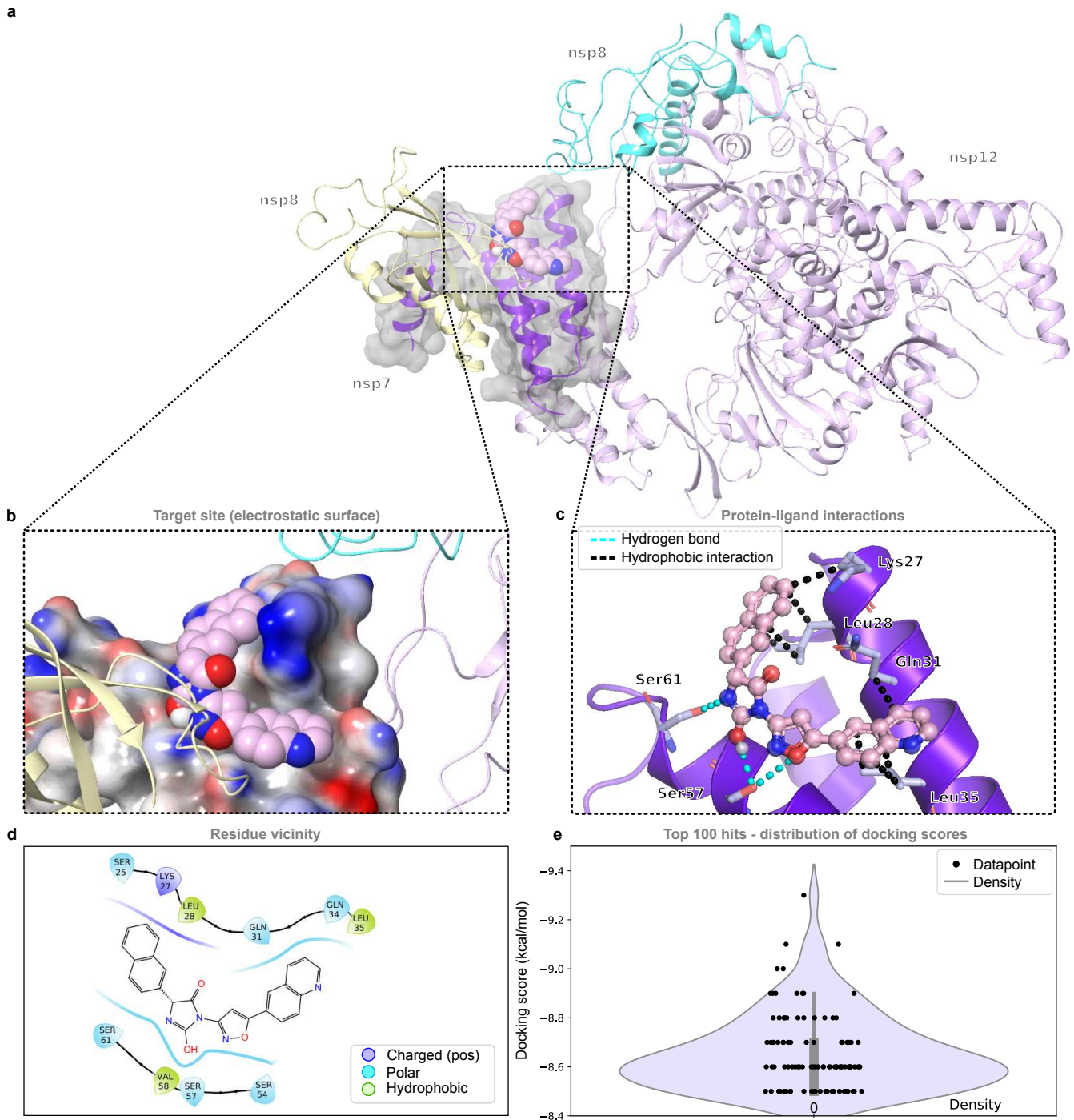


Figure S23. nsp7 and an example compound from the top 0.0001% of screened compounds bound to the surface. Related to Table 1. **a**, nsp7 (violet), where an example compound (light pink) from the virtual screen is bound at the nsp8 (light gold) interface. In the virtual screen, a blind docking was carried out for each screened compound over the entire surrounding helical surface. This helical surface includes large parts of the nsp8 (light gold, cyan) and nsp12 (lavender) binding interfaces (screen ID: 20). **b**, Electrostatic surface of nsp7 to which an example compound (light pink) is bound. Also shown are the protein-protein interaction partners nsp8 (light gold, cyan) and nsp12 (lavender). **c**, An overview of the interactions between the inhibitor and the nsp7 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

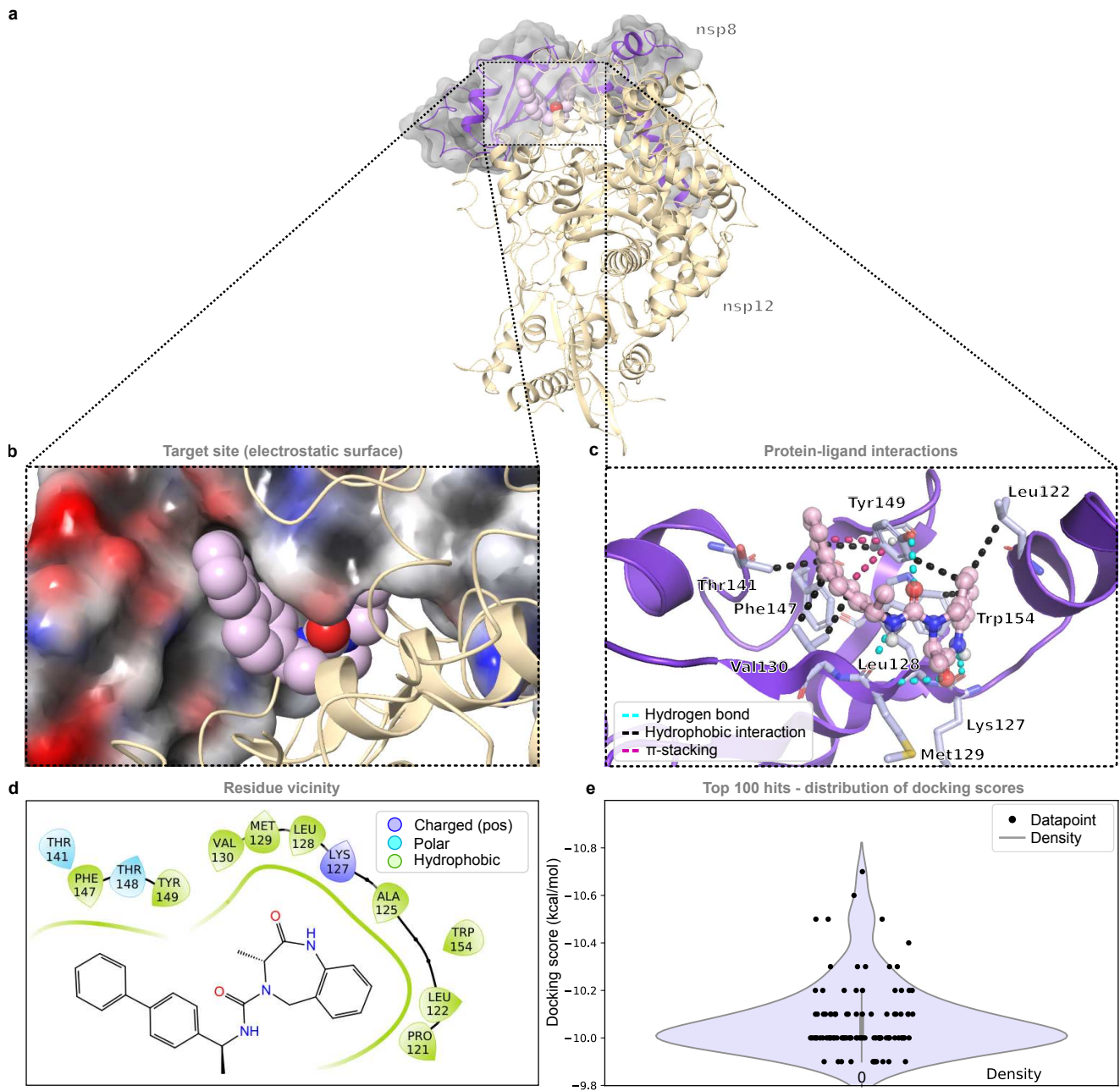


Figure S24. nsp8 and an example compound from the top 0.0001% of screened compounds bound at the nsp12 binding interface. Related to Table 1. a, nsp8 (violet) bound to nsp12 (in light gold) and an example compound (light pink) at the nsp12 binding interface (screen ID: 22). **b,** Electrostatic surface of nsp8 to which an example compound (light pink) and nsp12 (light gold) is bound. **c,** An overview of the interactions between the inhibitor and the nsp8 structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

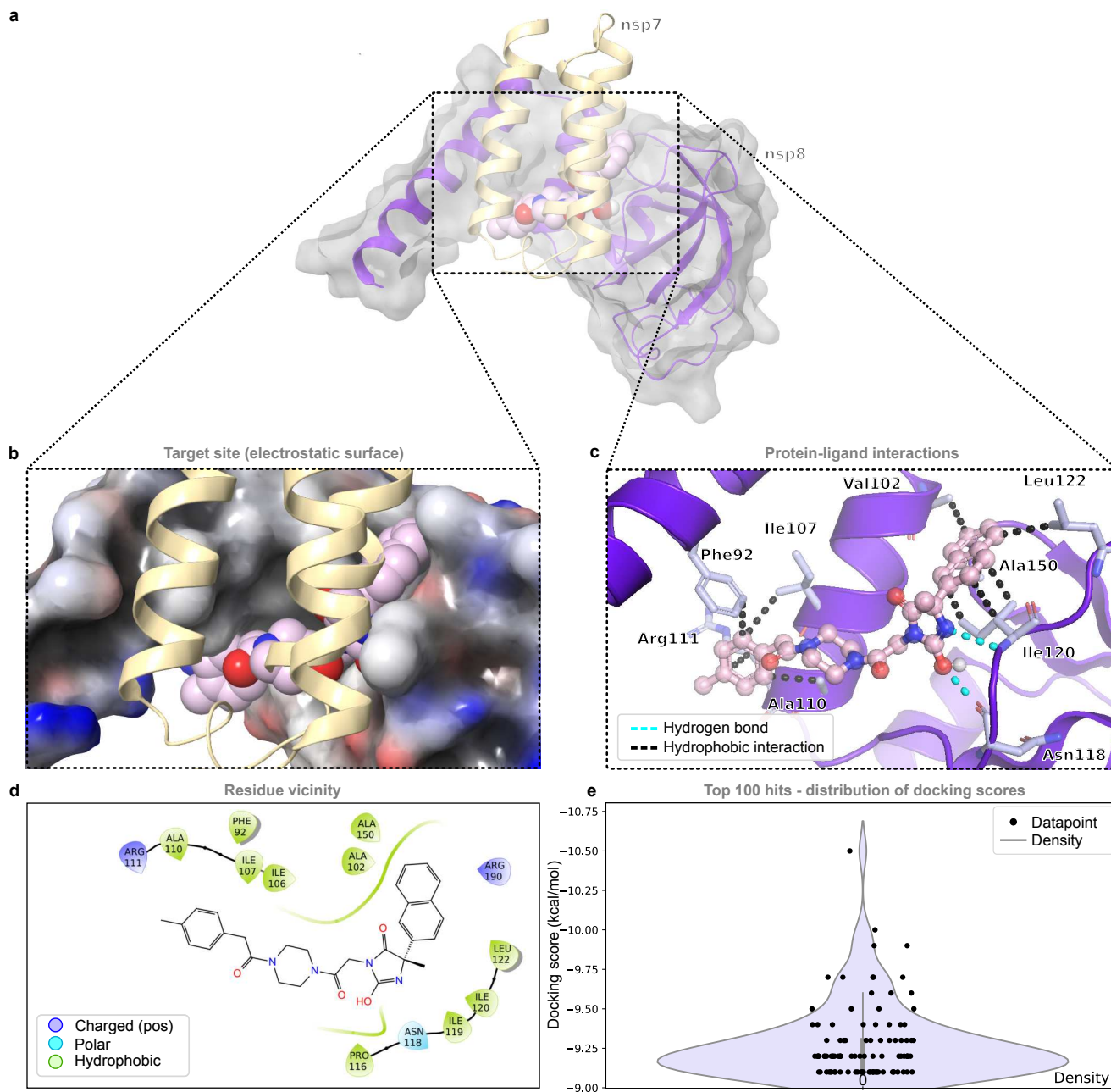


Figure S25. nsp8 and an example compound from the top 0.0001% of screened compounds bound at the nsp7 binding interface. Related to Table 1. **a**, nsp8 (violet) bound to nsp7 (light gold) and an example compound (light pink) (screen ID: 21). **b**, Electrostatic surface of the target protein to which an example compound (light pink) and nsp7 (light gold) bound. **c**, An overview of the interactions between the inhibitor and the nsp8 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

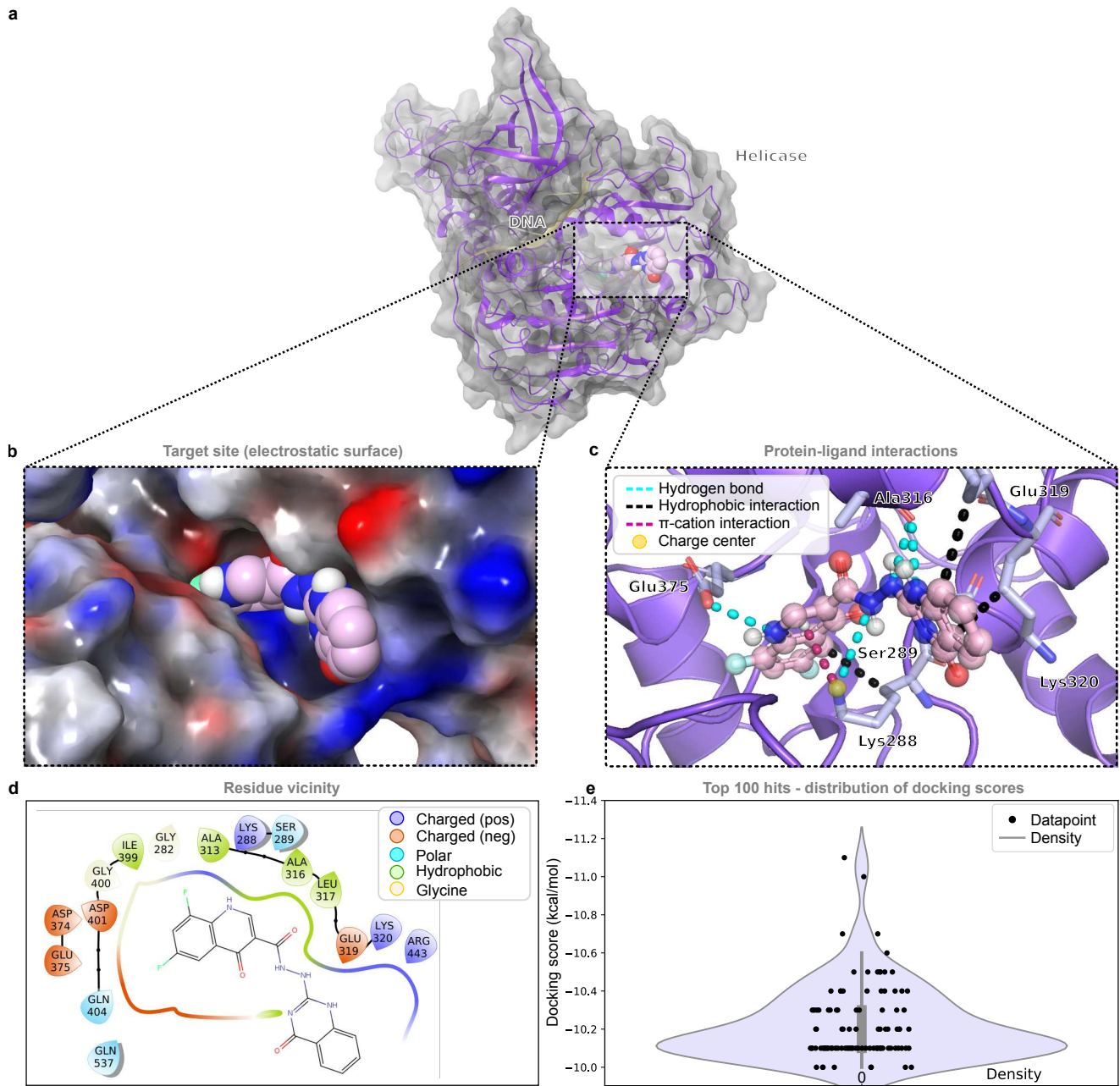


Figure S26. The helicase (nsp13) and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site. Related to Table 1. a, The helicase (violet) bound to an example compound (light pink) at the enzymatic active site. A docked DNA strand is shown in light gold (screen ID: 33). **b,** Electrostatic surface of the helicase to which an example compound (light pink) is bound. **c,** An overview of the interactions between the compound and the helicase structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

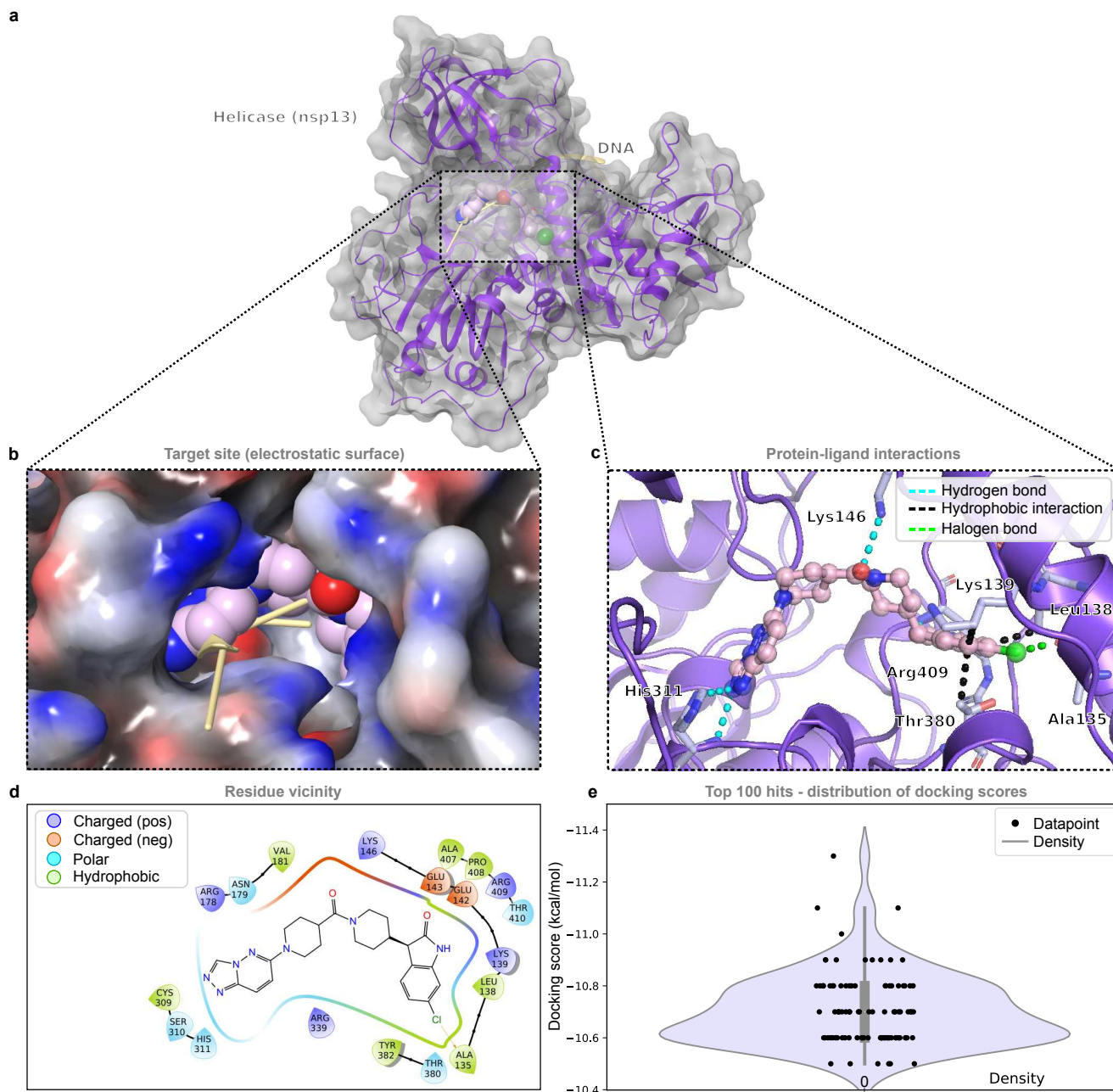


Figure S27. The helicase (nsp13) and an example compound from the top 0.0001% of screened compounds bound at site 1 of the RNA-binding interface. Related to Table 1. a. The helicase (violet) bound to an example compound (light pink) at site 1 of the RNA-binding interface. A docked DNA strand is shown in light gold (screen ID: 34). **b.** Electrostatic surface of the helicase to which an example compound (light pink) and a docked DNA strand (light gold) are bound. **c.** An overview of the interactions between the compound and the helicase structure. **d.** Residues within 4 Å of the ligand. **e.** Distribution of the docking scores of the top 100 virtual screening hits.

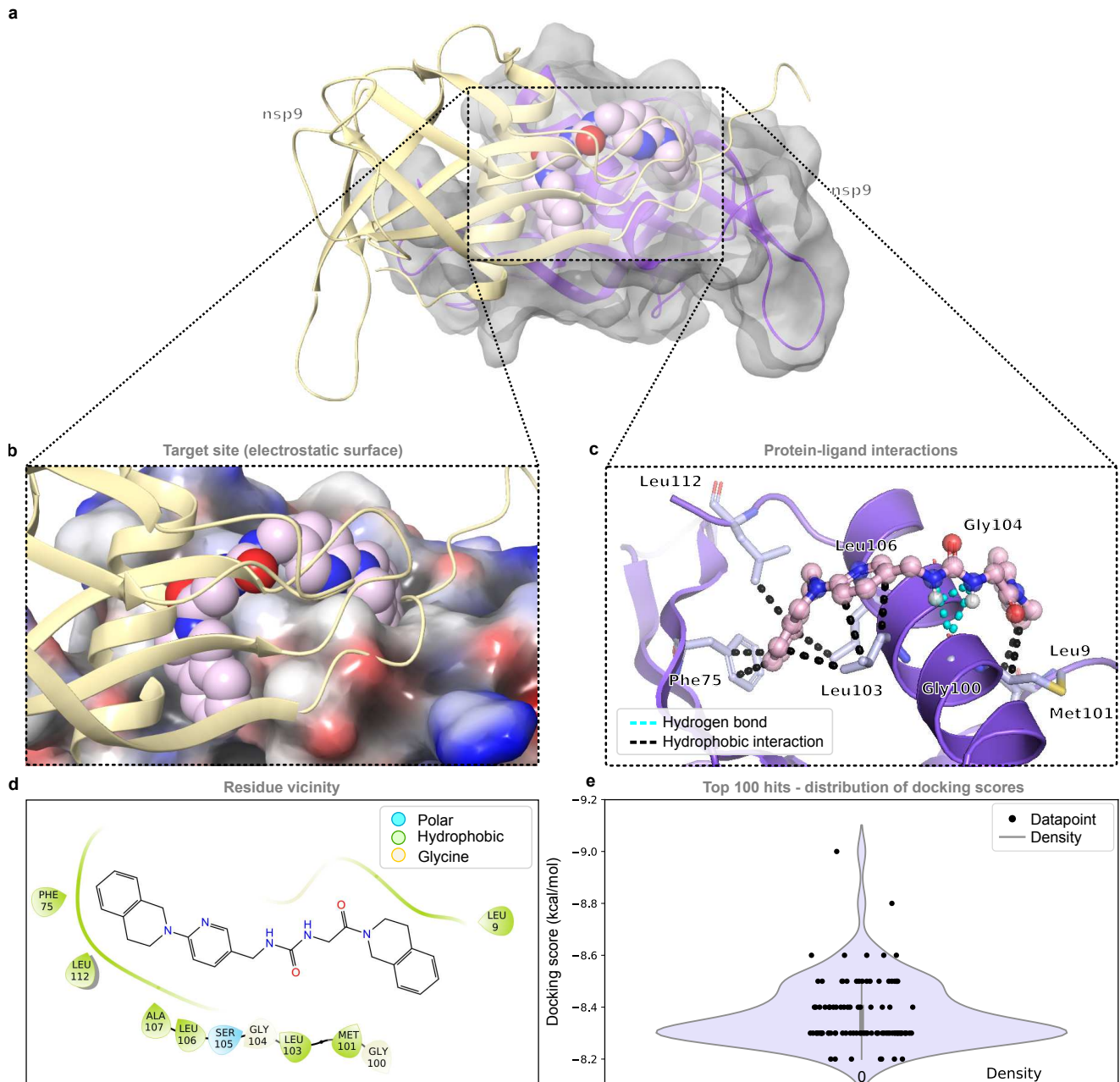


Figure S28. nsp9 and an example compound from the top 0.0001% of screened compounds bound at the dimerization interface (site 1). Related to Table 1. **a**, nsp9 dimer (one monomer in violet, one monomer in gold), where an example compound (light pink) from the virtual screen is bound to the dimerization interface (site 1) of the nsp9 monomer in violet (screen ID: 23). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the nsp9 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

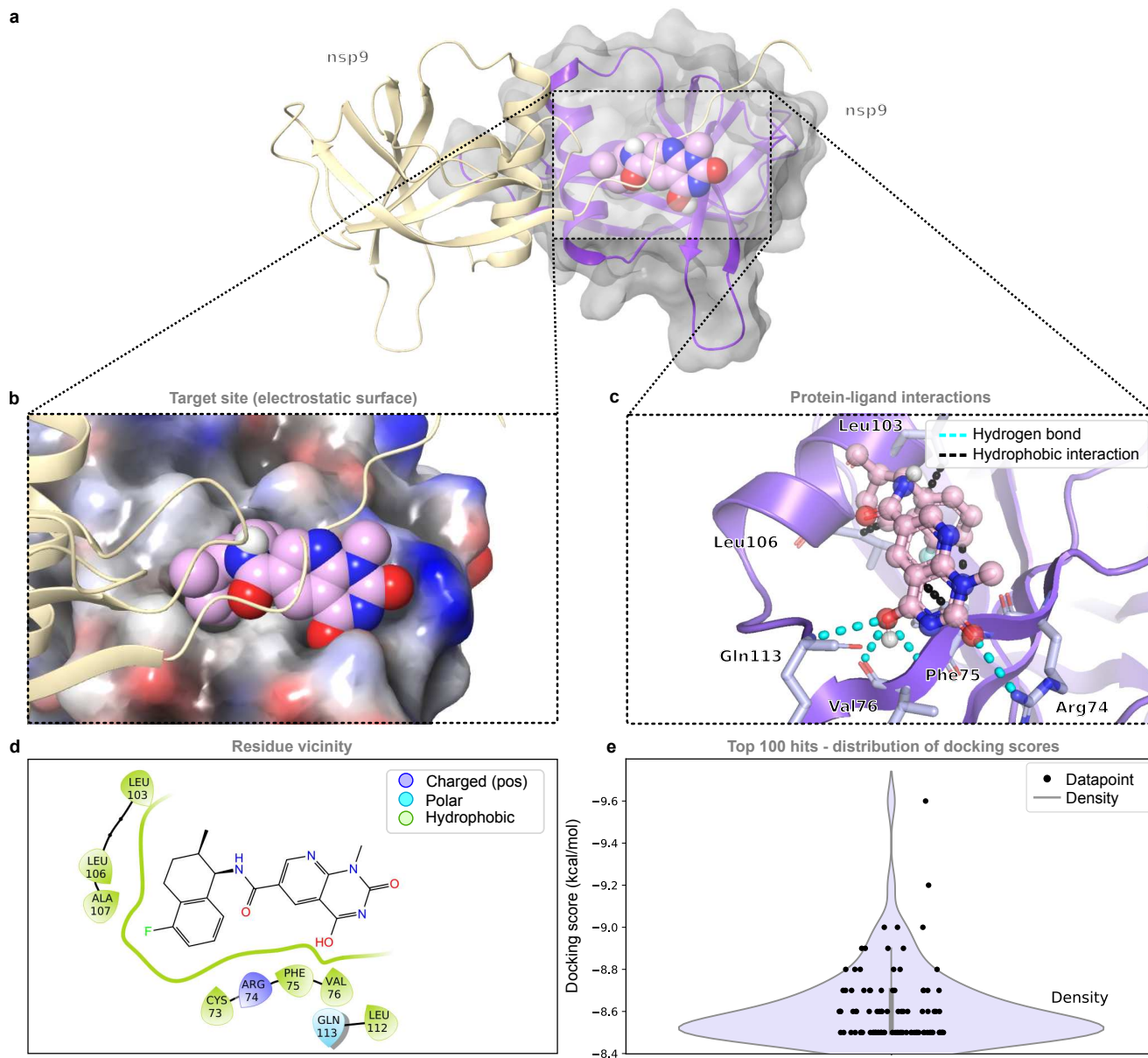


Figure S29. nsp9 and an example compound from the top 0.0001% of screened compounds bound at the dimerization interface (site 2). Related to Table 1. **a**, The nsp9 dimer (one monomer in violet, one monomer in gold), where an example compound (light pink) from the virtual screen is bound to the dimerization interface (site 2) of the nsp9 monomer in violet (screen ID: 24). **b**, Electrostatic surface of the target protein to which an example compound (light pink) and a second nsp9 monomer (light gold) is bound. **c**, An overview of the interactions between the inhibitor and the nsp9 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

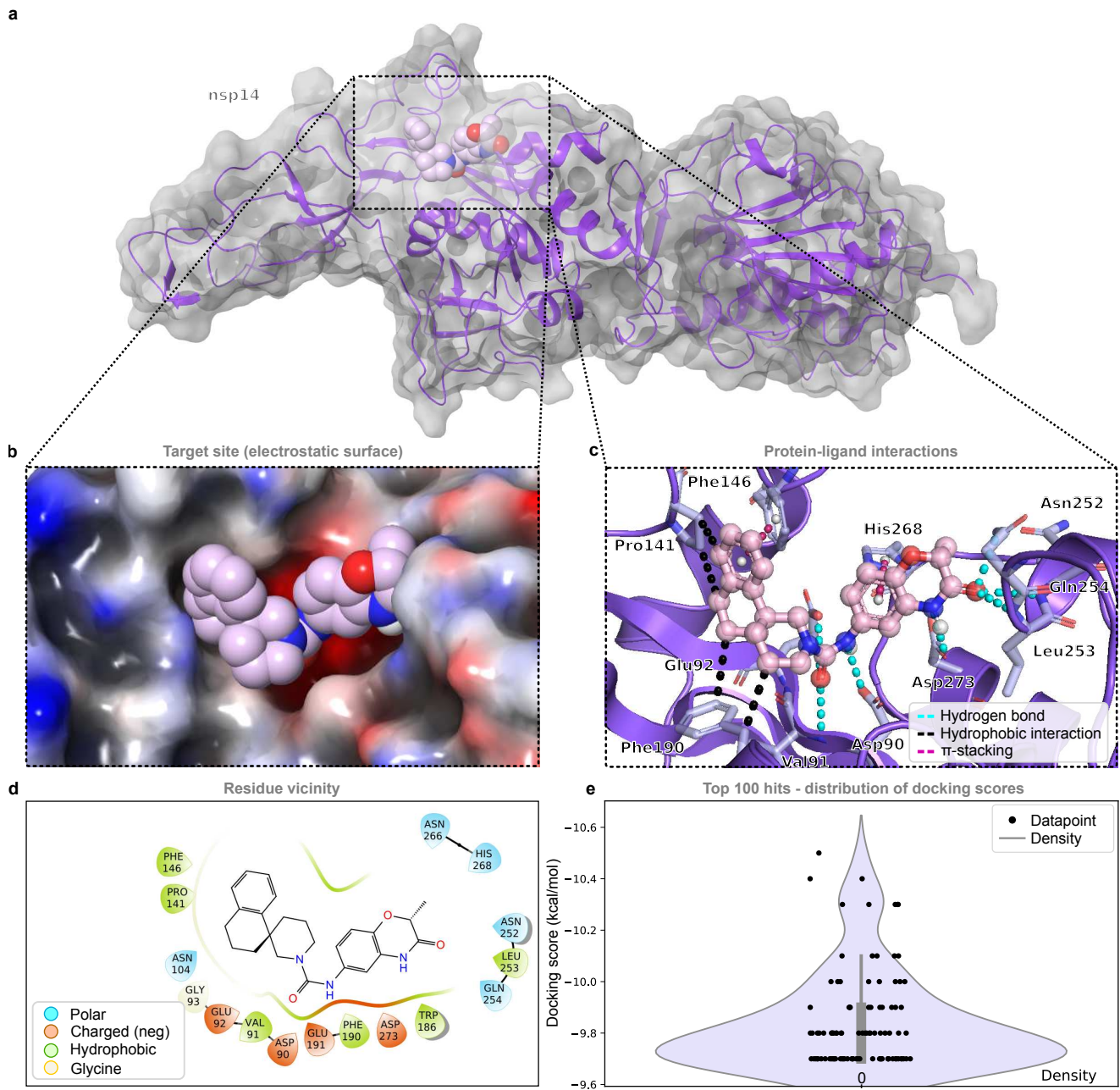


Figure S30. nsp14 and an example compound from the top 0.0001% of screened compounds bound at the active site of the exoribonuclease domain. Related to Table 1. a. nsp14 (violet) bound to an example compound (light pink) at the active site of the exoribonuclease domain (screen ID: 37). **b.** Electrostatic surface of nsp14 to which an example compound (light pink) is bound. **c.** An overview of the interactions between the compound and the nsp14 structure. **d.** Residues within 4 Å of the ligand. **e.** Distribution of the docking scores of the top 100 virtual screening hits.

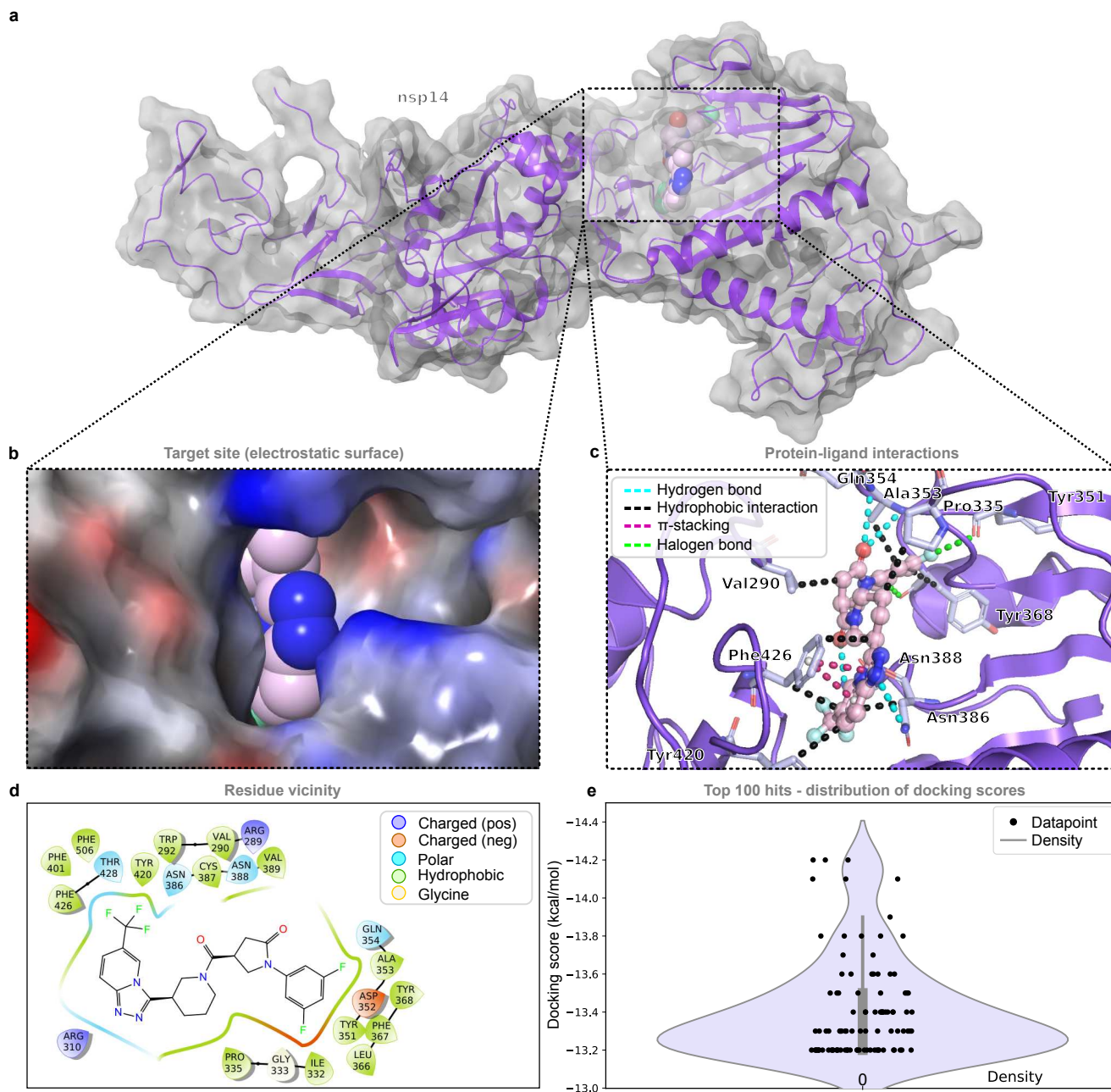


Figure S31. nsp14 and an example compound from the top 0.0001% of screened compounds bound at the active site of the N7 methyltransferase domain. Related to Table 1. **a**, nsp14 (violet) bound to an example compound (light pink) at the active site of the N7 methyltransferase domain (screen ID: 38). **b**, Electrostatic surface of nsp14 to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nsp14 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

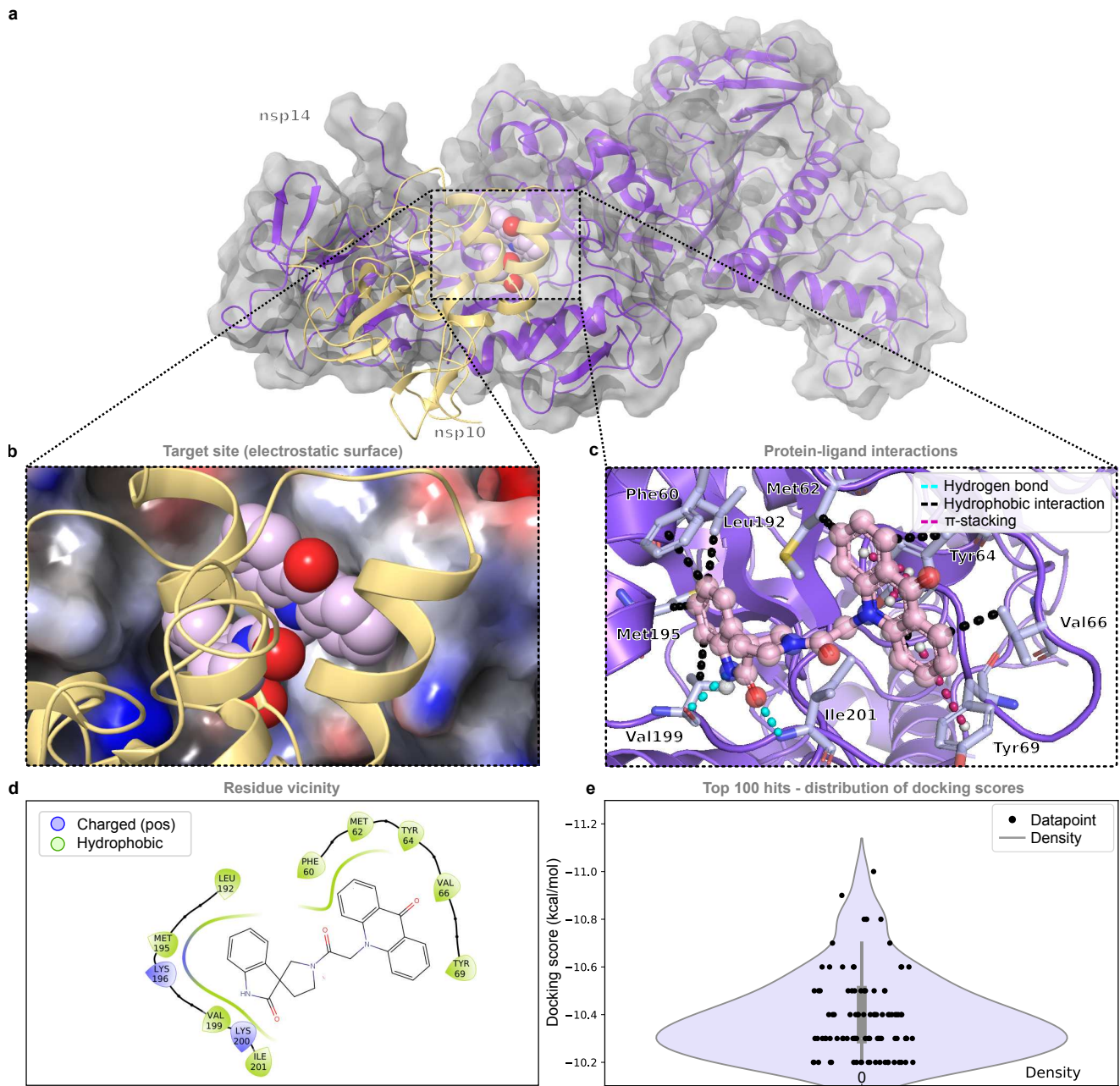


Figure S32. nsp14 and an example compound from the top 0.0001% of screened compounds bound at the nsp10 protein-protein interface. Related to Table 1. **a**, nsp14 (violet) bound to an example compound (light pink) at the nsp10 (light gold) binding interface (screen ID: 36). **b**, Electrostatic surface of nsp14 to which an example compound (light pink) and nsp10 (light gold) are bound. **c**, An overview of the interactions between the compound and the nsp14 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

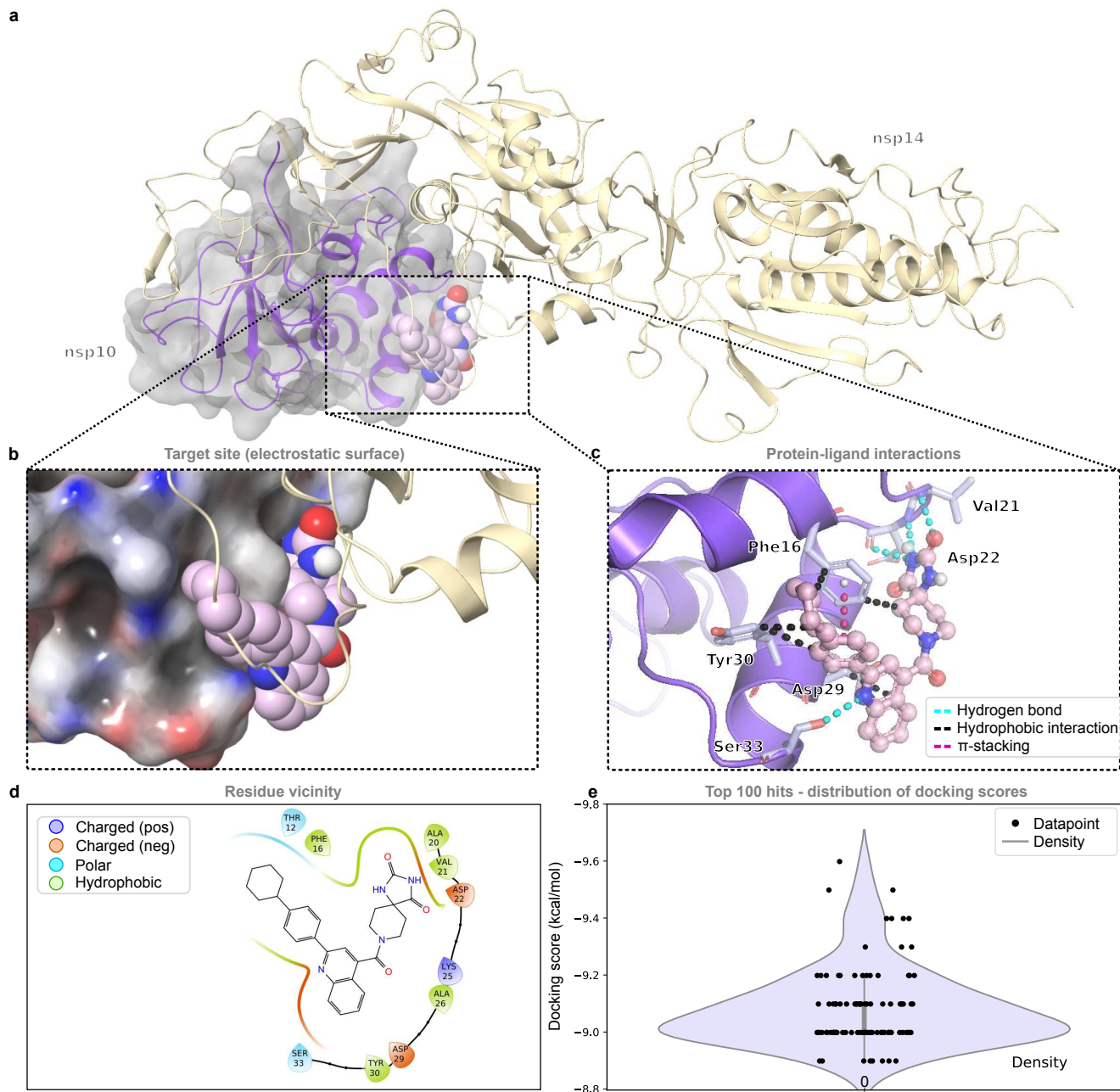


Figure S33. nsp10 and an example compound from the top 0.0001% of screened compounds bound at the nsp14 binding interface. Related to Table 1. a, nsp10 (violet) bound to nsp14 (light gold) and an example compound (light pink) at the nsp14 binding interface of nsp10 (screen ID: 27). **b,** Electrostatic surface of nsp10 to which an example compound (light pink) and nsp14 (light gold) are bound. **c,** An overview of the interactions between the inhibitor and the nsp10 structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

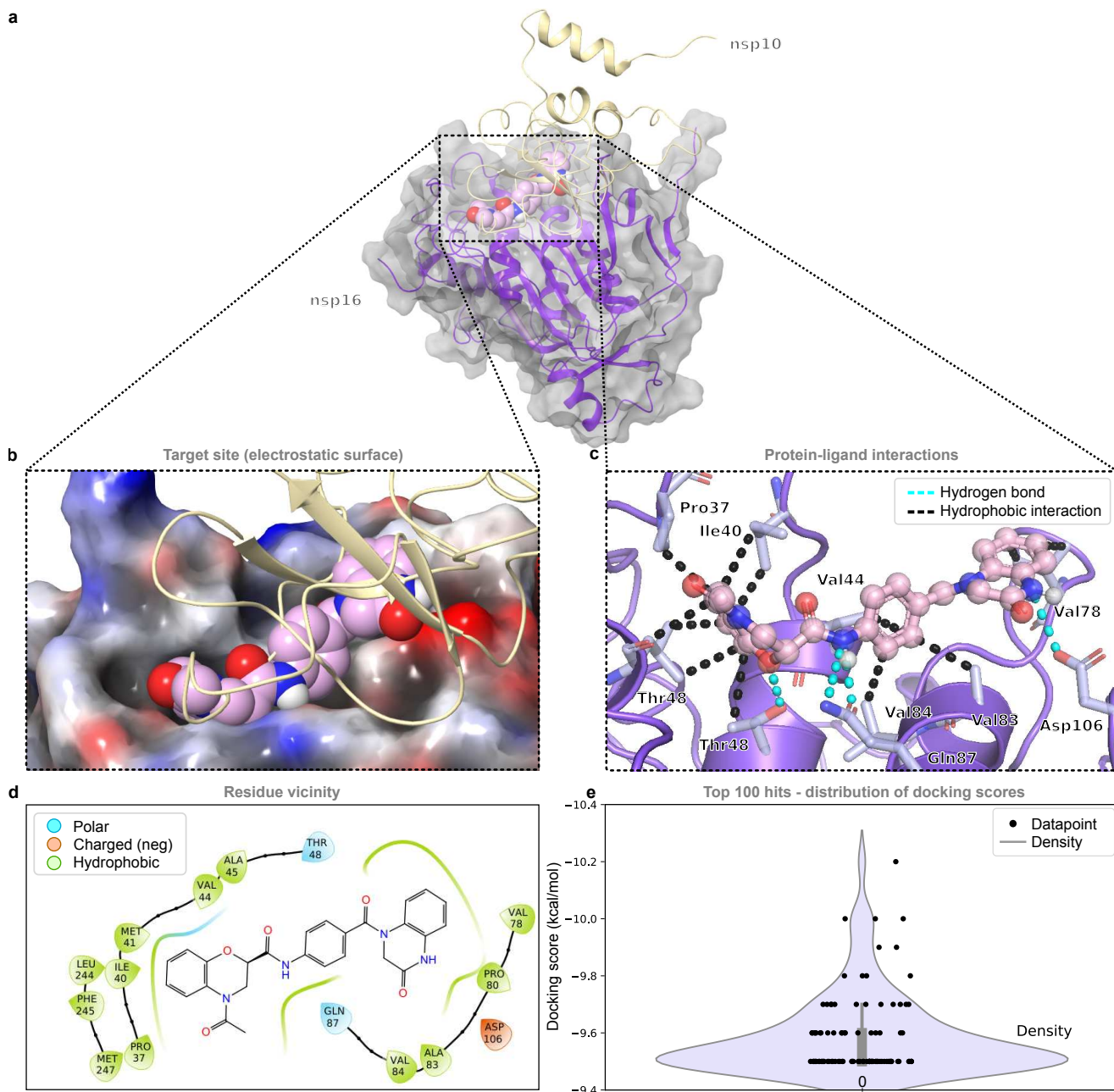
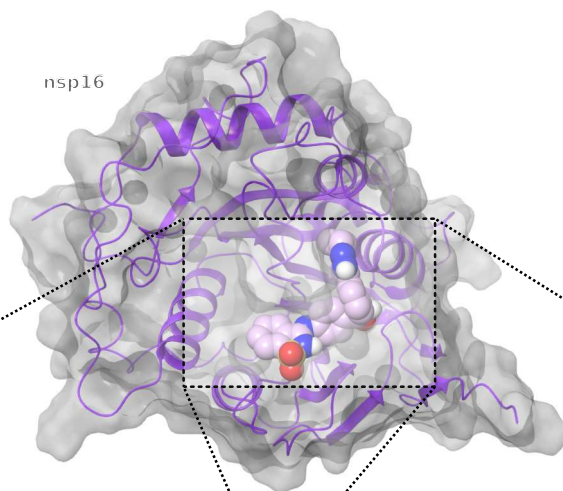
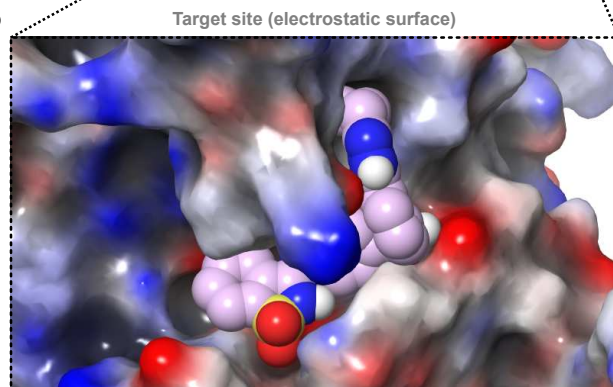


Figure S34. nsp16 and an example compound from the top 0.0001% of screened compounds bound at the nsp10 protein-protein interaction interface. Related to Table 1. **a**, nsp16 (violet) bound to an example compound (light pink) at the nsp10 (light gold) protein-protein interface (screen ID: 40). **b**, Electrostatic surface of nsp16 to which an example compound (light pink) and nsp10 (light gold) are bound. **c**, An overview of the interactions between the compound and the nsp16 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

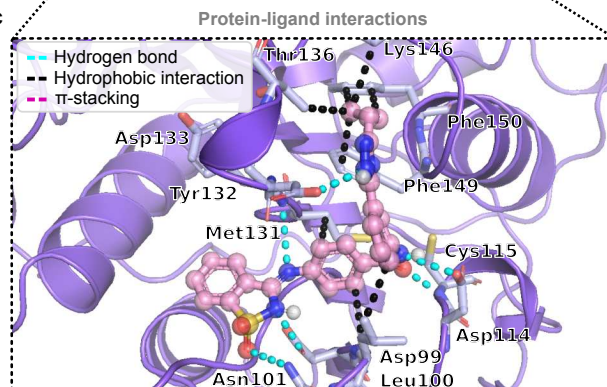
a



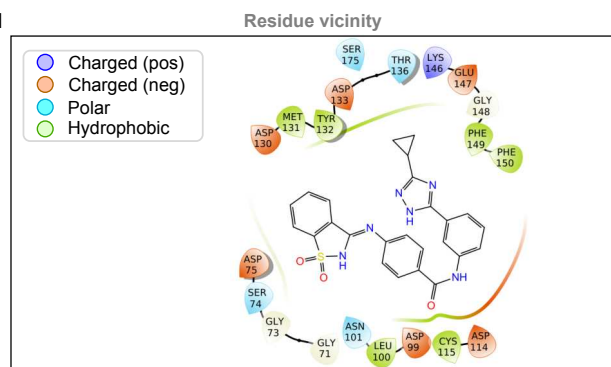
b



c



d



e

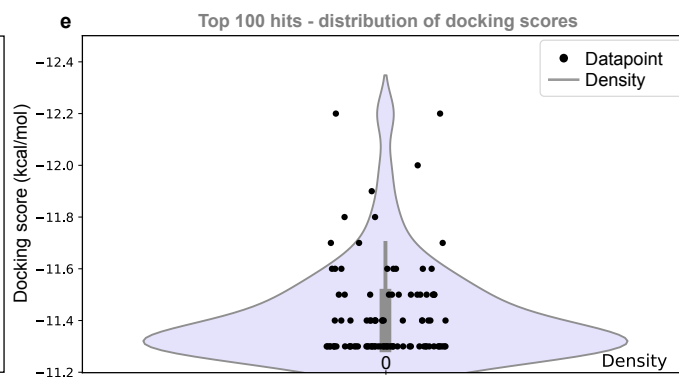


Figure S35. nsp16 and an example compound from the top 0.0001% of screened compounds bound at the active site. Related to Table 1. **a**, nsp16 (violet) bound to an example compound (light pink) at the active site (screen ID: 41). **b**, Electrostatic surface of nsp16 to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nsp16 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

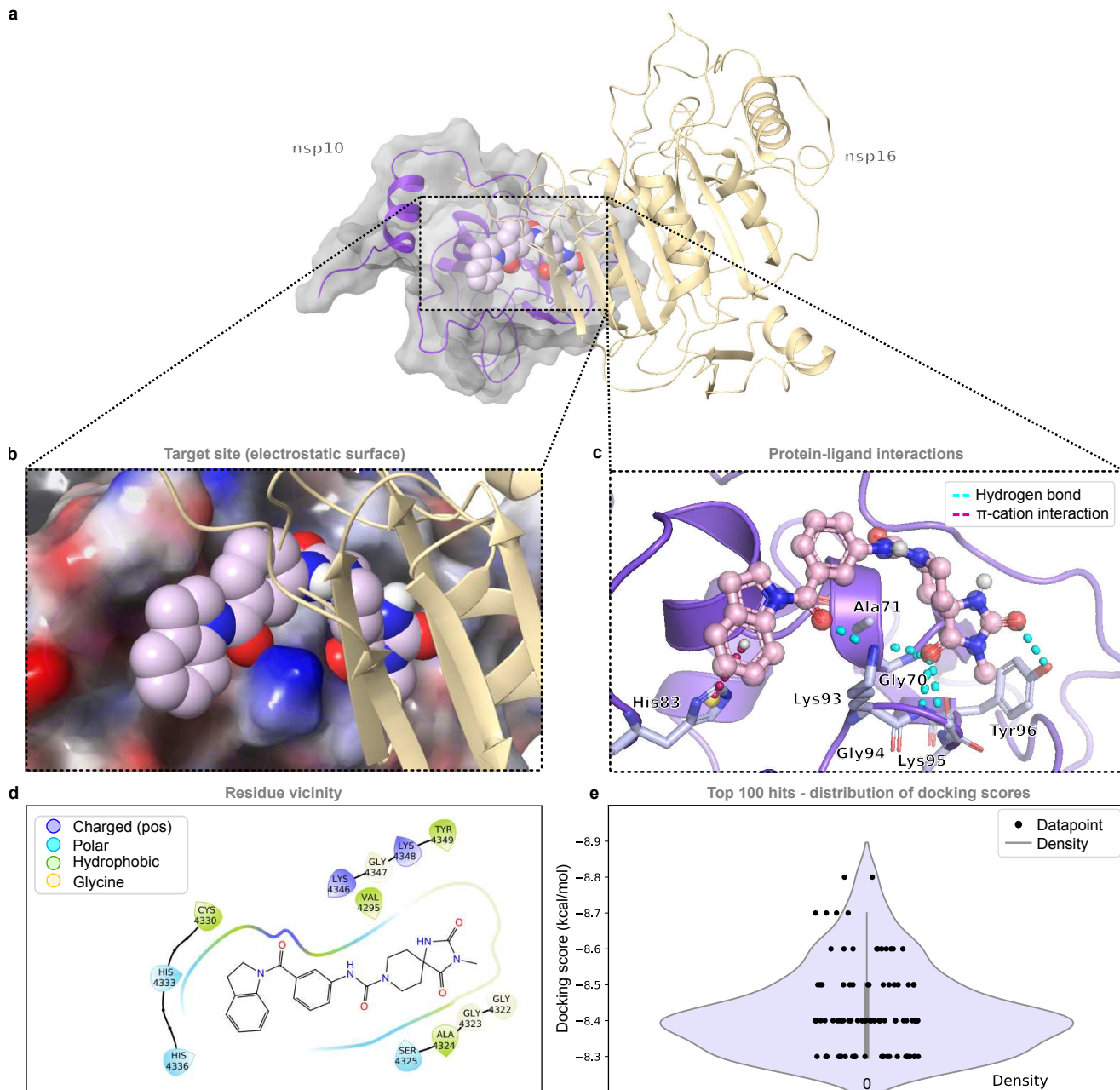


Figure S36. nsp10 and an example compound from the top 0.0001% of screened compounds bound at the protein-protein interaction interface with nsp16. Related to Table 1. **a**, nsp10 (violet) bound to nsp16 (light gold) and an example compound (light pink), at the nsp16 binding interface of nsp10 (screen ID: 25). The region of the nsp14 interface which was targeted here is overlapping with the nsp16 binding interface. **b**, Electrostatic surface of nsp10 to which an example compound (light pink) as well nsp16 (light gold) are bound. **c**, An overview of the interactions between the inhibitor and the nsp10 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

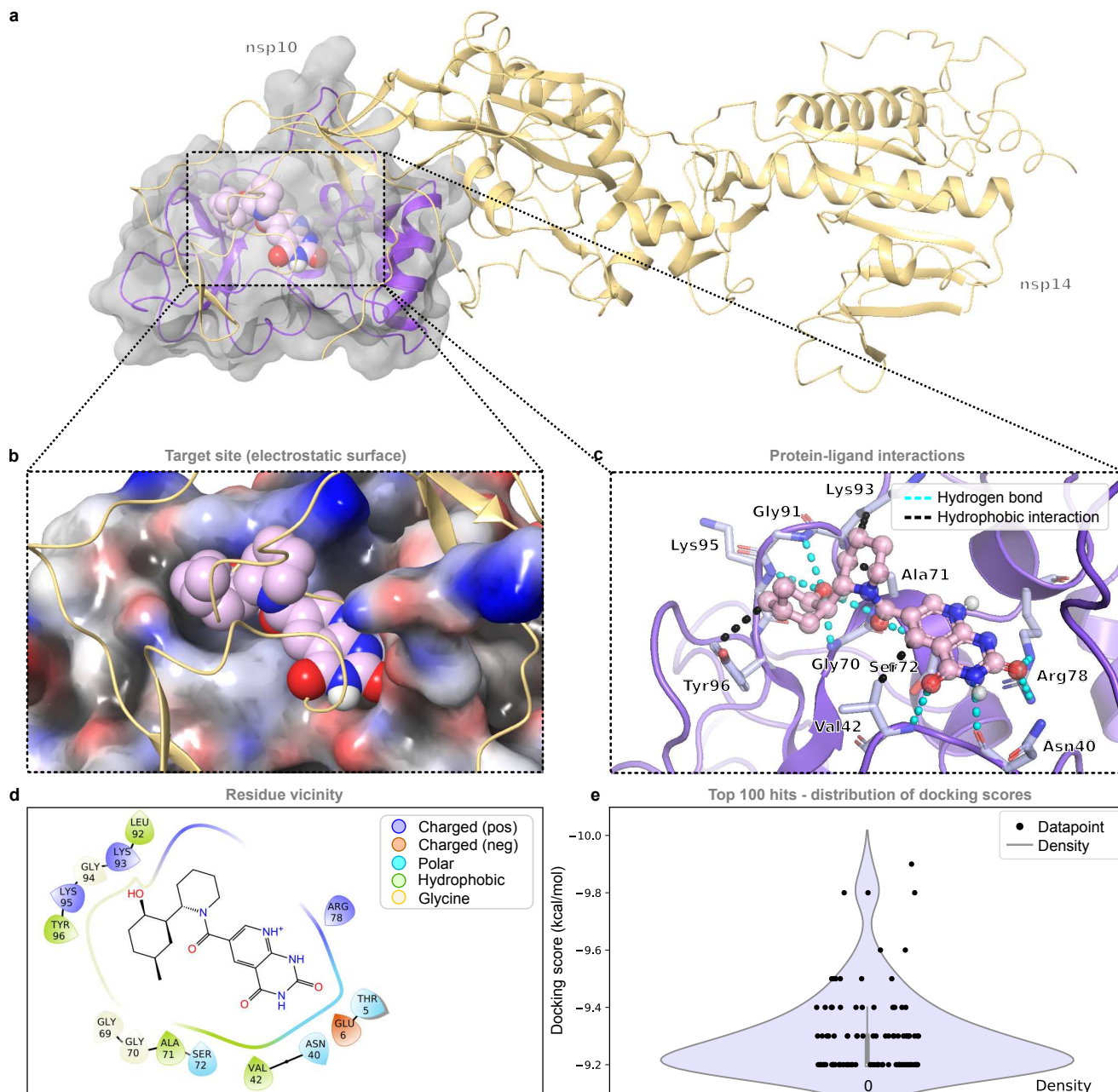


Figure S37. nsp10 and an example compound from the top 0.0001% of screened compounds bound at the binding interface to nsp14/16. Related to Table 1. **a**, nsp10 (violet) bound to nsp14 (light gold) and an example compound (light pink) at the binding interface of nsp10 to nsp14/16 (screen ID: 26). The region of the nsp14 interface which was targeted here is overlapping with the nsp16 binding interface. **b**, Electrostatic surface of nsp10 to which an example compound (light pink) and nsp14 (light gold cartoon) are bound. **c**, An overview of the interactions between the inhibitor and the nsp10 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

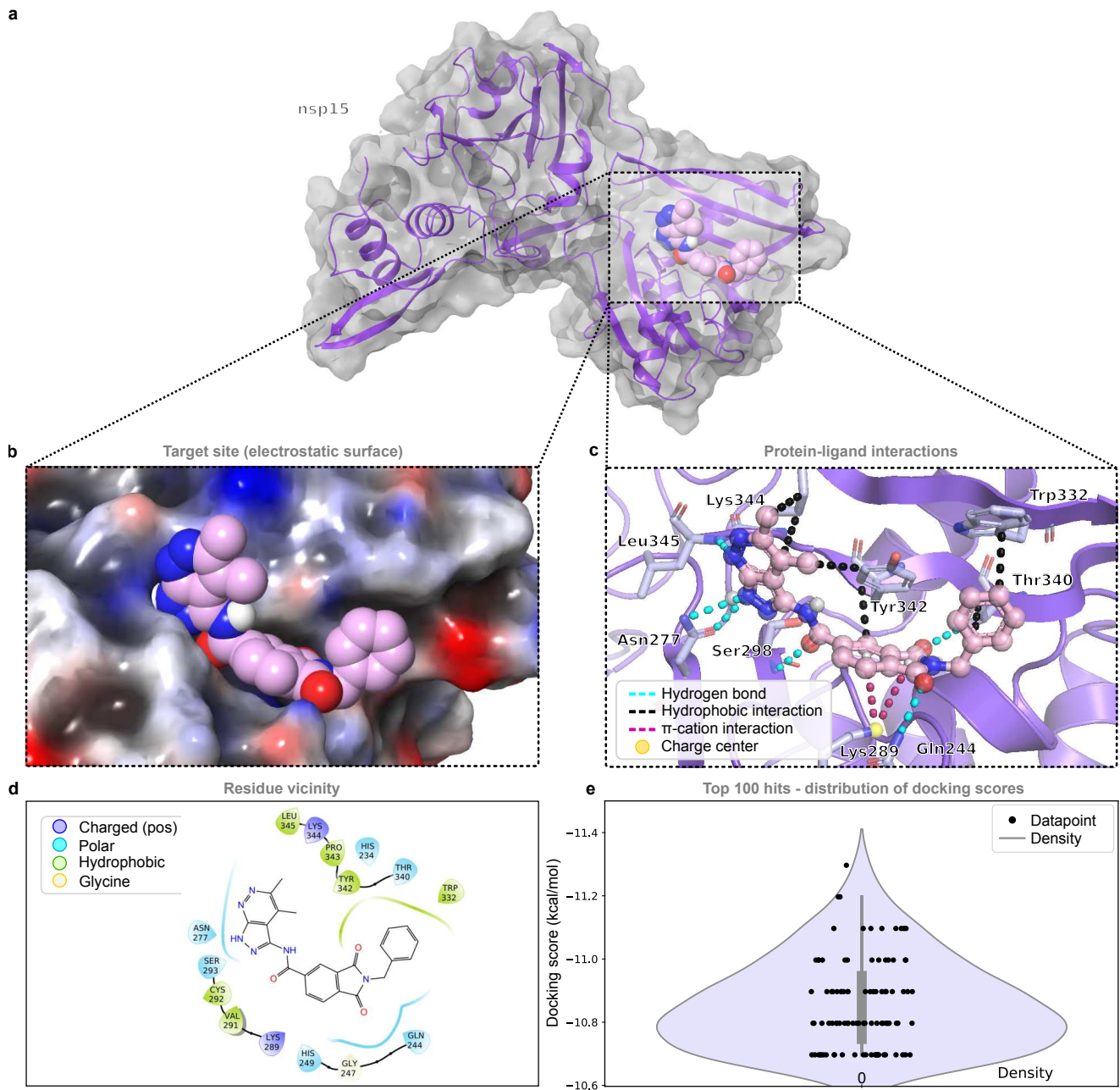


Figure S38. nsp15 and an example compound from the top 0.0001% of screened compounds bound at the active site. Related to Table 1. **a**, nsp15 (violet) bound to an example compound (light pink) at the active site (screen ID: 39). **b**, Electrostatic surface of nsp15 to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nsp15 structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

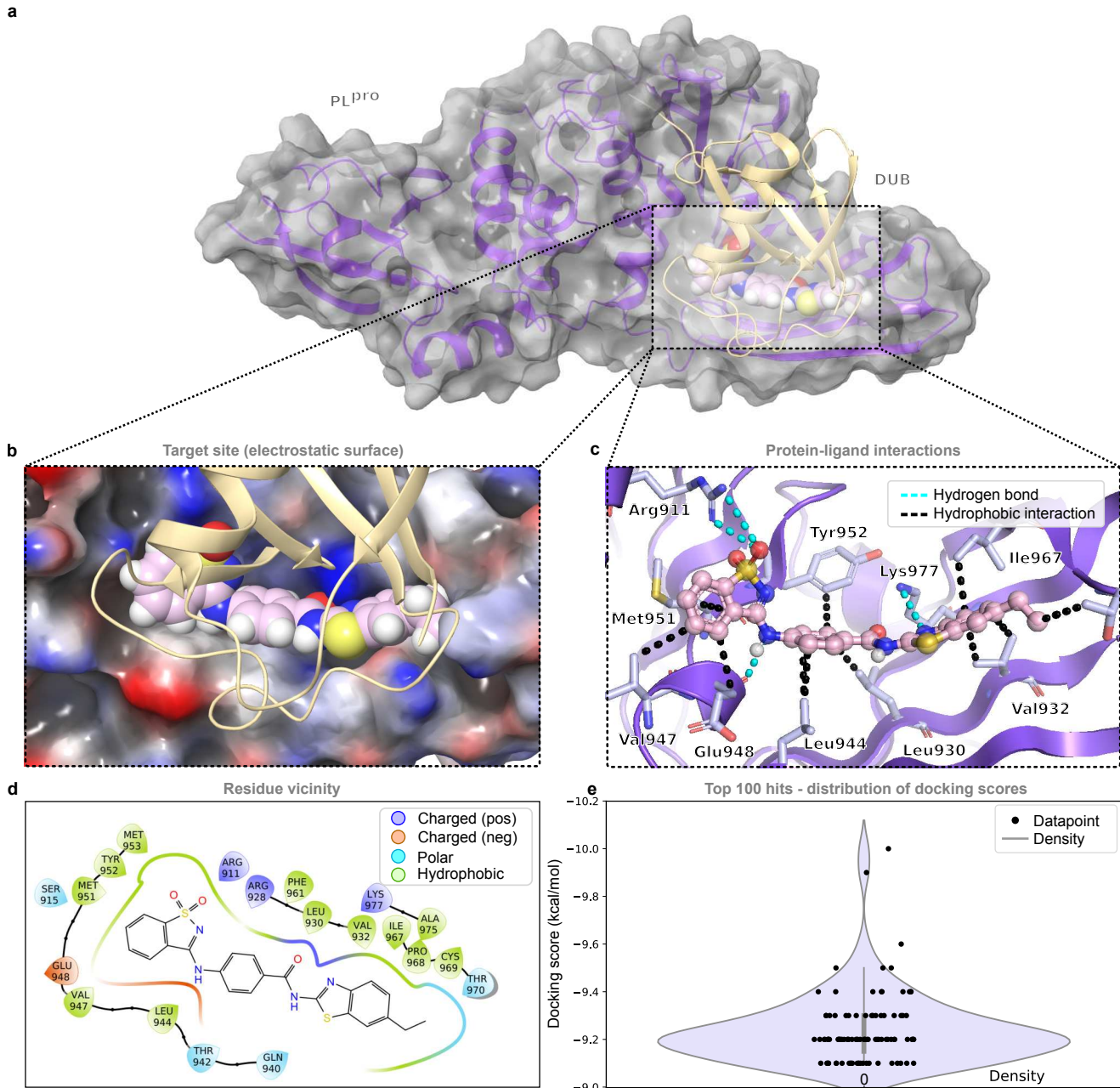


Figure S39. PL^{pro} and an example compound from the top 0.0001% of screened compounds bound at the DUB binding site. Related to Table 1. **a**, PL^{pro} (violet) and an example compound (light pink) from the virtual screen bound at the DUB binding site (screen ID: 14). **b**, Electrostatic surface of the target protein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the inhibitor and the receptor structure. **d**, Receptor Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

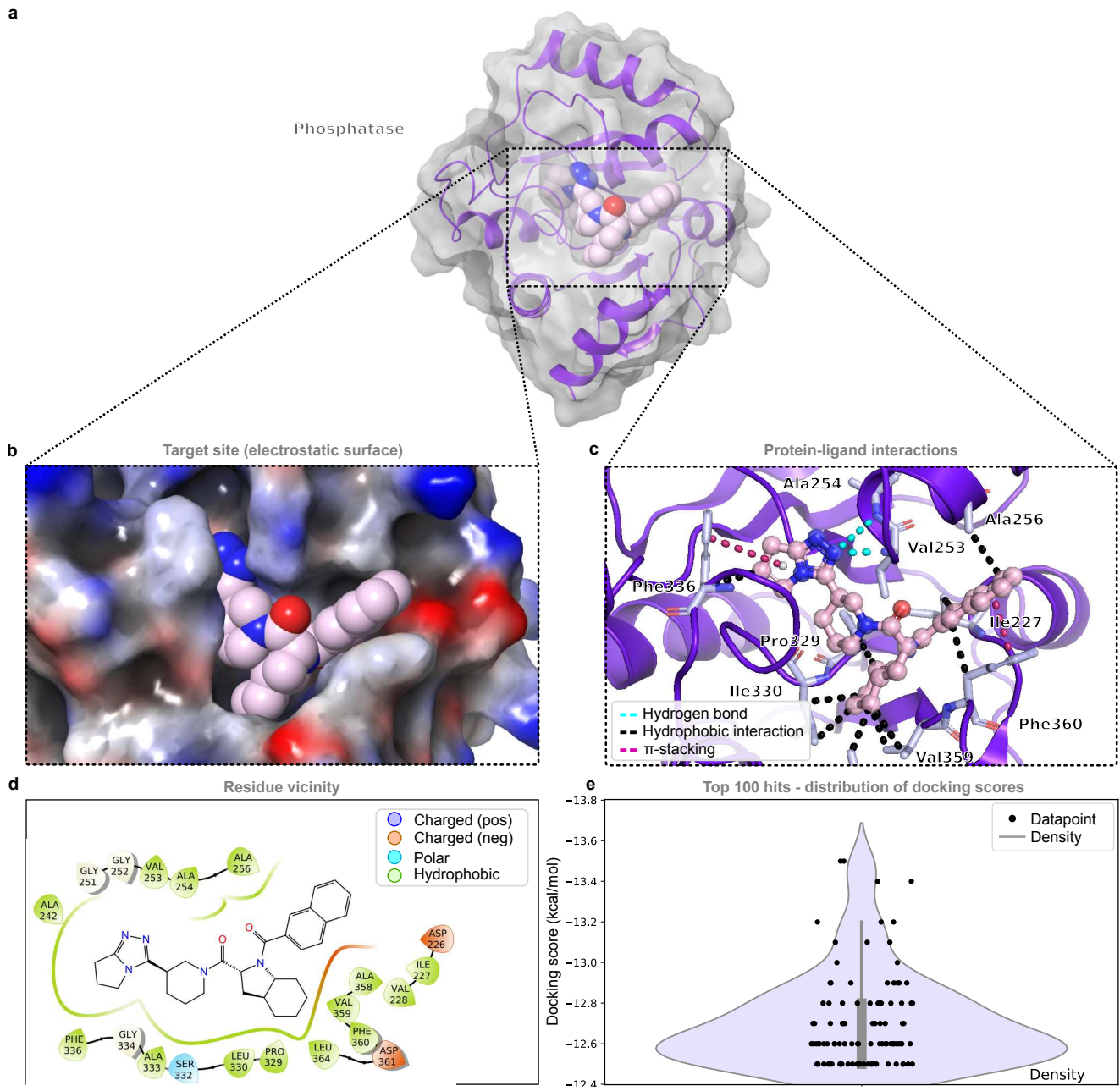


Figure S40. The phosphatase (open conformation) and an example compound from the top 0.0001% of screened compounds bound at the enzymatic active site. Related to Table 1. a, The phosphatase (violet) which is part of nsp3, and an example compound (light pink) from the virtual screen bound to the active site (open conformation)(screen ID: 11). **b,** Electrostatic surface of the phosphatase to which an example compound (light pink) is bound. **c,** An overview of the interactions between the inhibitor and the phosphatase structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

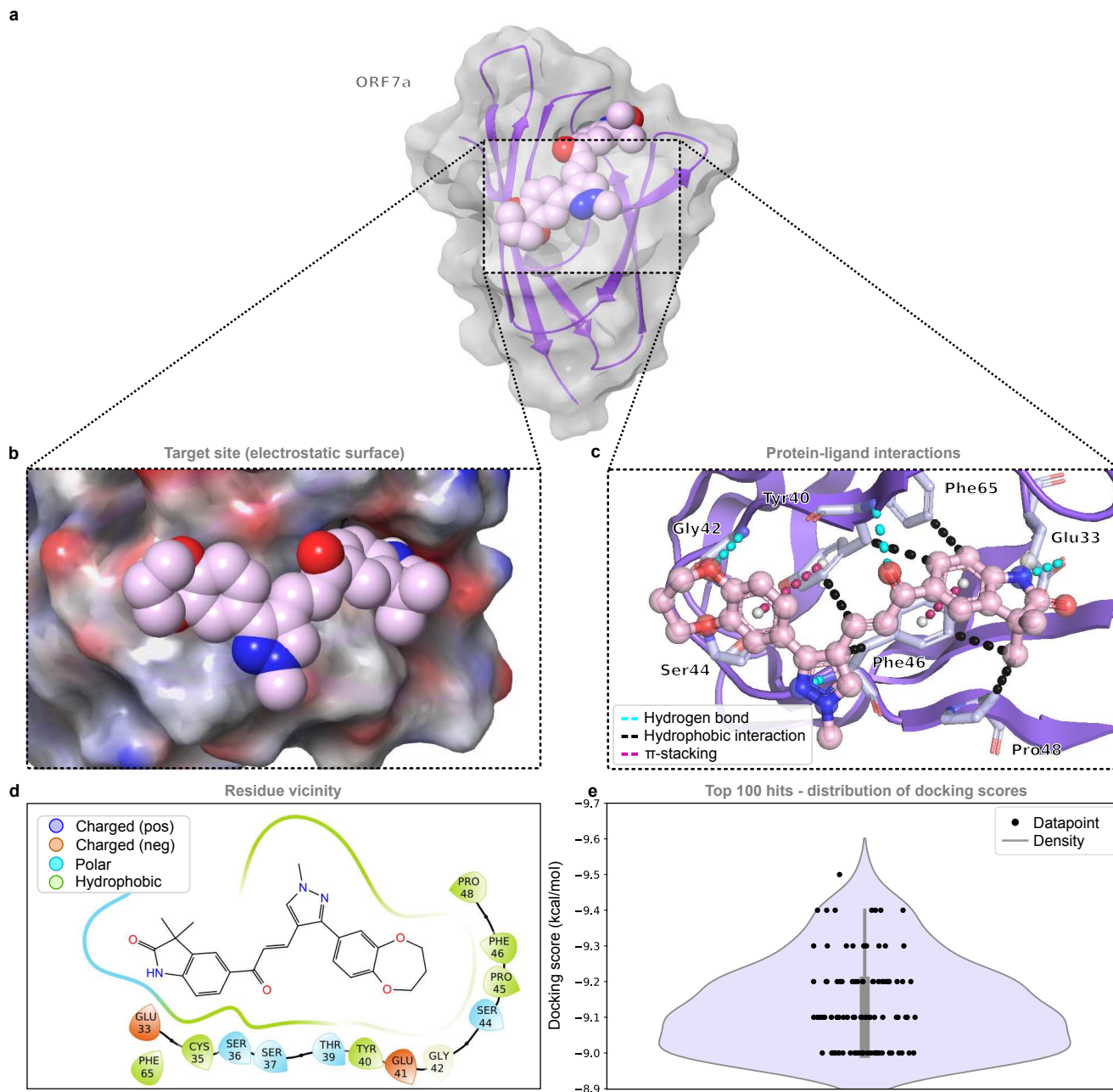


Figure S41. ORF7a and an example compound from the top 0.0001% of screened compounds bound. Related to Table 1. a, ORF7a (violet) bound to an example compound (light pink) at the surface (screen ID: 9). **b,** Electrostatic surface of ORF7a to which an example compound (light pink) is bound. **c,** An overview of the interactions between the compound and the ORF7a structure. **d,** Residues within 4 Å of the ligand. **e,** Distribution of the docking scores of the top 100 virtual screening hits.

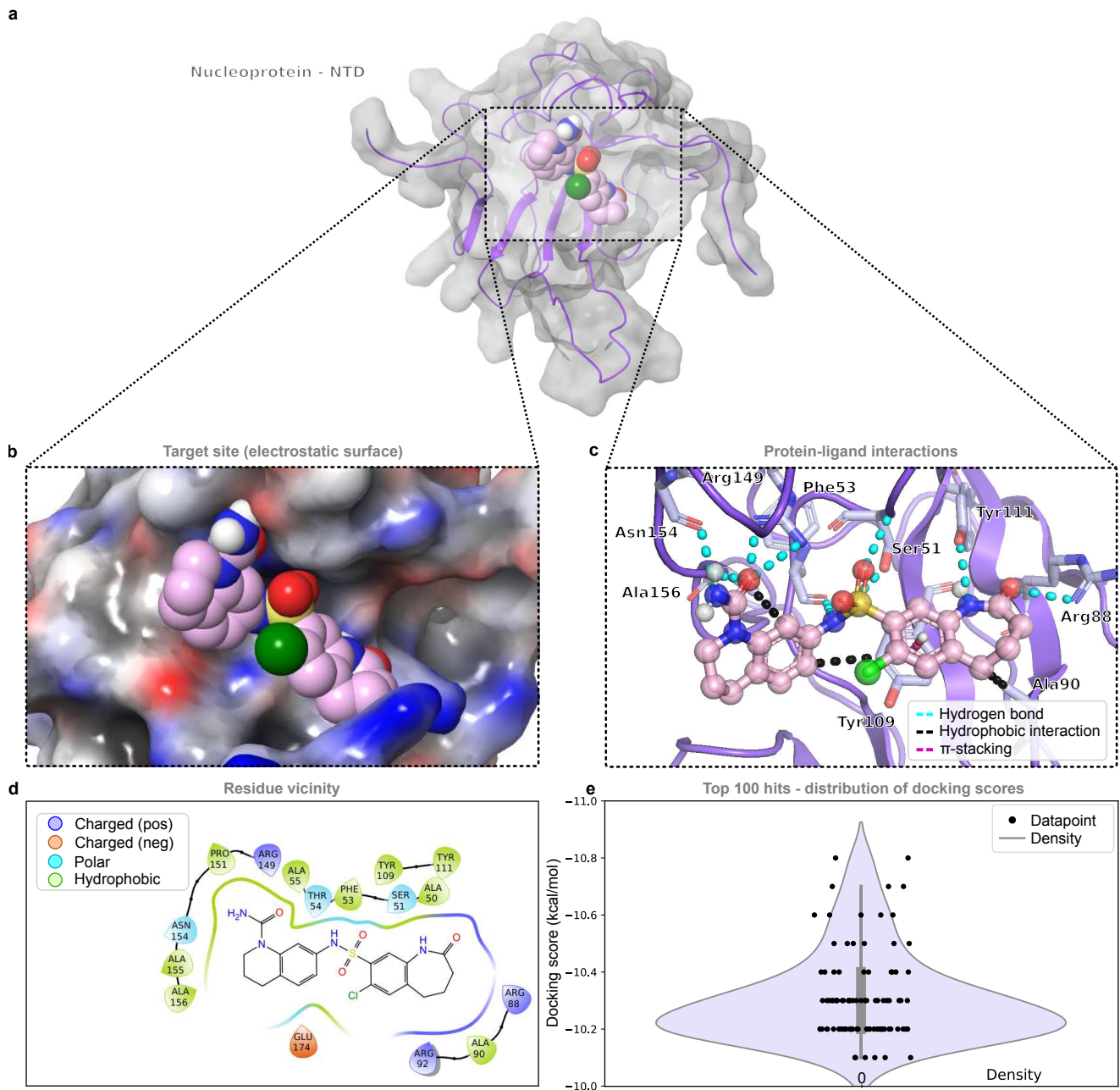


Figure S42. N-terminal domain of the nucleoprotein and an example compound from the top 0.0001% of screened compounds bound at the RNA-binding interface. Related to Table 1. **a**, The N-terminal domain (NTD) of nucleoprotein (violet) bound to an example compound (light pink) at the RNA binding interface (screen ID: 42). **b**, Electrostatic surface of nucleoprotein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nucleoprotein NTD structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

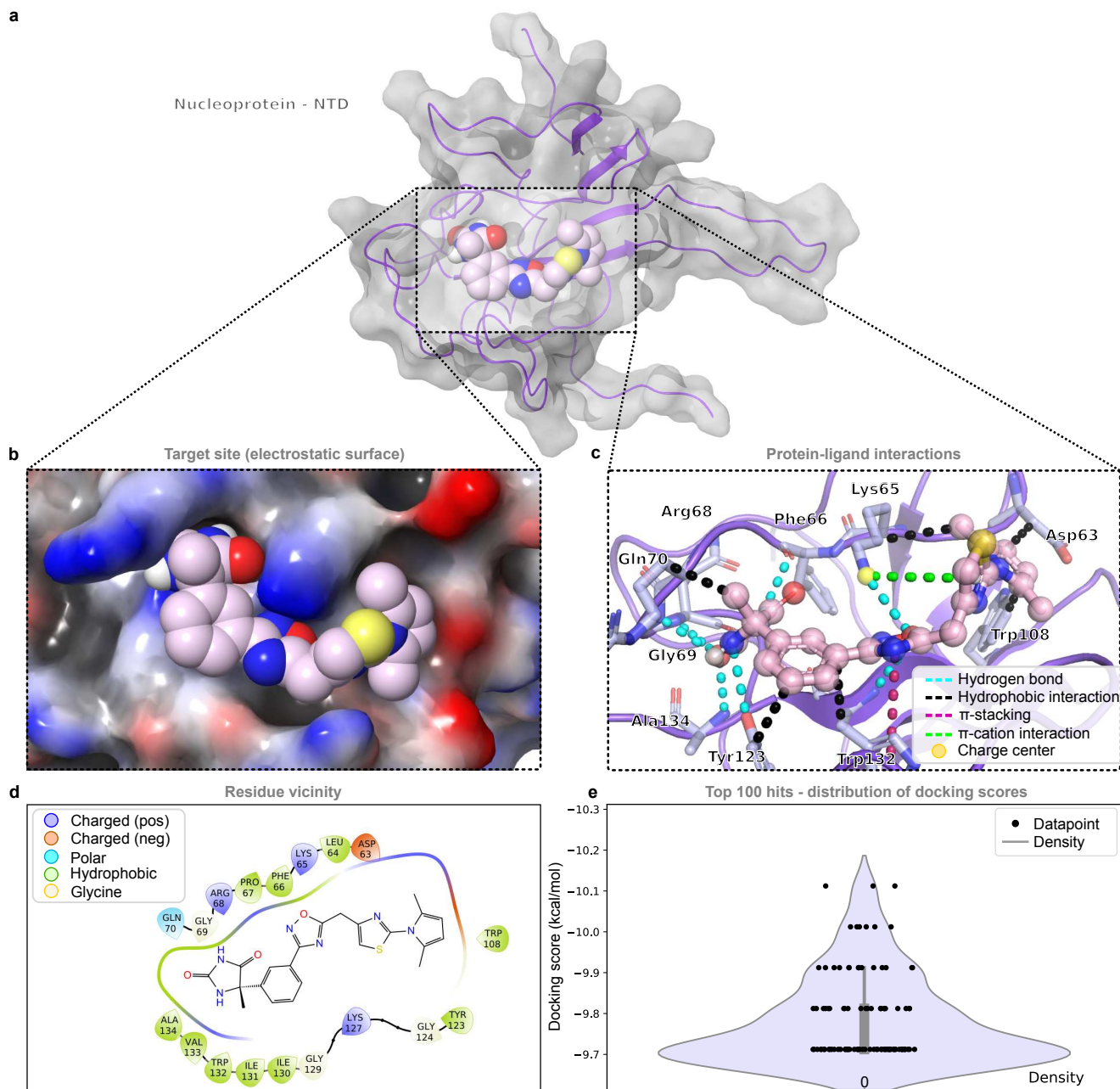


Figure S43. N-terminal domain of the nucleoprotein and an example compound from the top 0.0001% of screened compounds bound at the oligomerization interface. Related to Table 1. a, N-terminal domain (NTD) of the nucleoprotein (violet) bound to an example compound (light pink) at the oligomerization interface (screen ID: 43). **b**, Electrostatic surface of the NTD of the nucleoprotein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nucleoprotein NTD structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

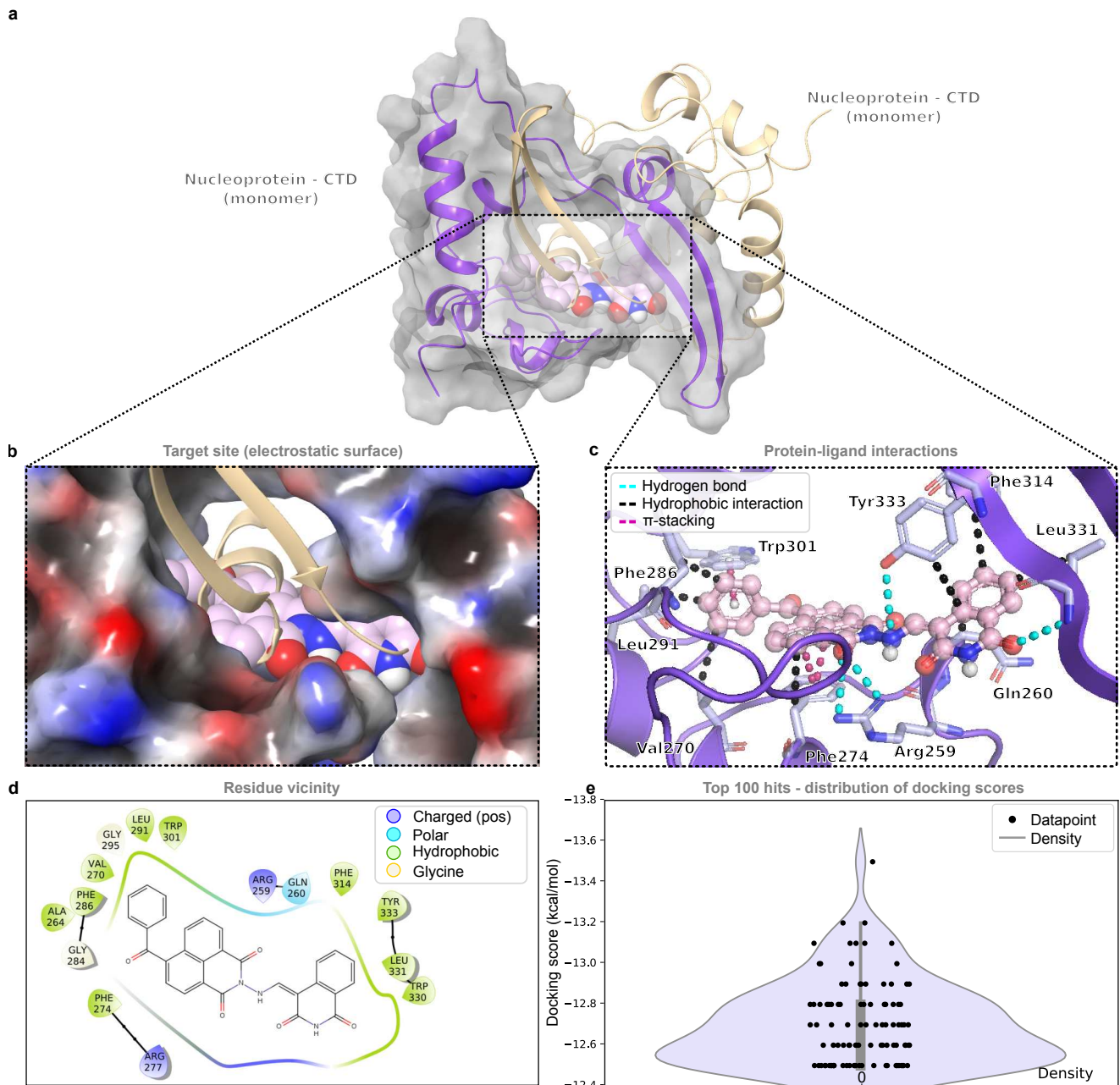


Figure S44. C-terminal domain of the nucleoprotein and an example compound from the top 0.0001% of screened compounds bound at the dimerization interface. Related to Table 1. **a**, C-terminal domain (CTD) of the nucleoprotein (violet) bound to an example compound (light pink) at the dimerization interface. A second CTD of the nucleoprotein is shown in light gold (screen ID: 44). **b**, Electrostatic surface of the CTD of the nucleoprotein to which an example compound (light pink) and a second nucleoprotein CTD (light gold) are bound. **c**, An overview of the interactions between the compound and the nucleoprotein CTD structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits.

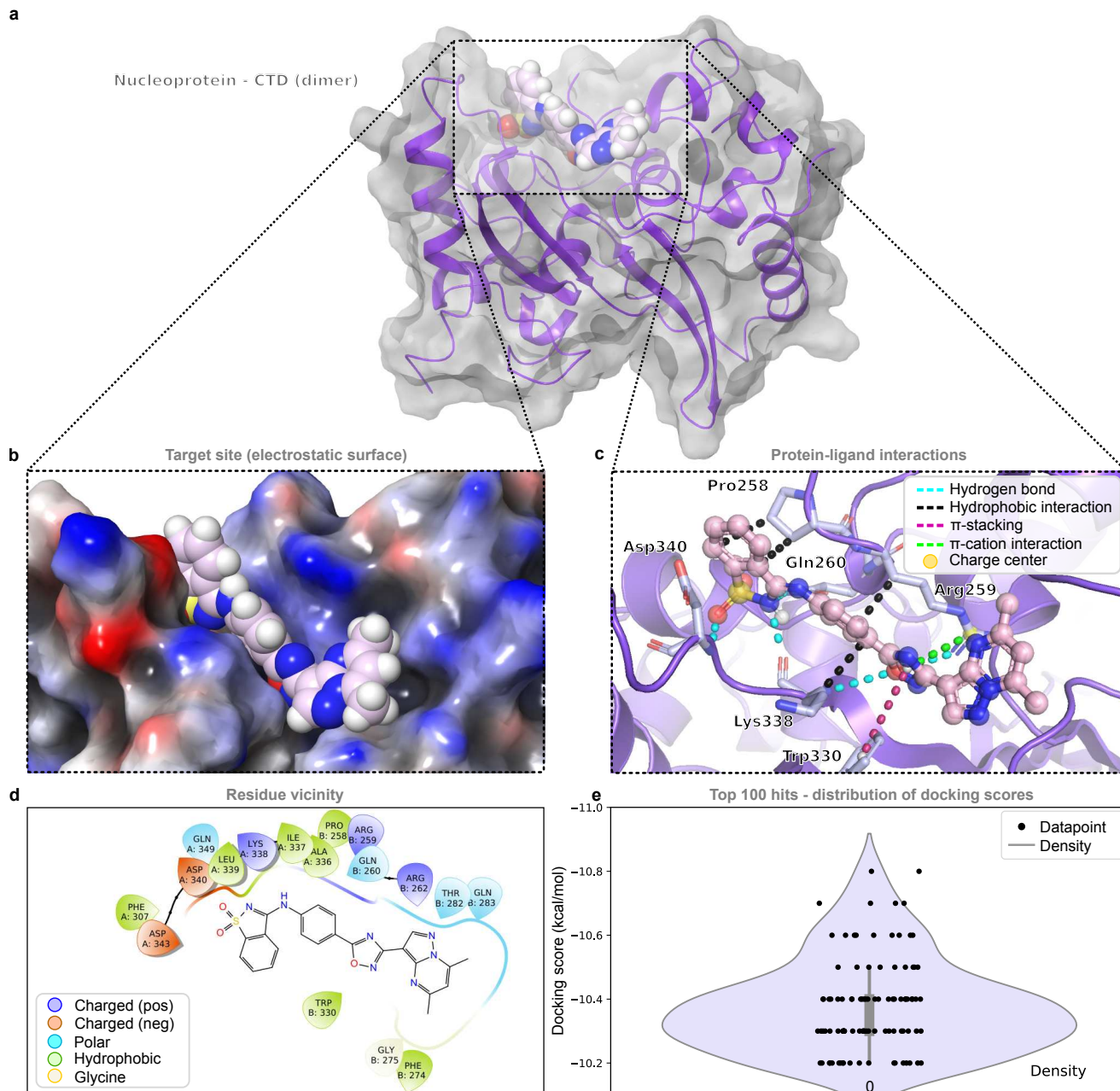


Figure S45. C-terminal domain of the nucleoprotein and an example compound from the top 0.0001% of screened compounds bound at the oligomerization interface. Related to Table 1. **a**, The C-terminal domain (CTD) dimer of nucleoprotein (violet) bound to an example compound (light pink) at the oligomerization interface (screen ID: 45). **b**, Electrostatic surface of the CTD dimer of nucleoprotein to which an example compound (light pink) is bound. **c**, An overview of the interactions between the compound and the nucleoprotein CTD structure. **d**, Residues within 4 Å of the ligand. **e**, Distribution of the docking scores of the top 100 virtual screening hits. Related to Table 1.

ZINCID	Generic name	COVID clinical trial	Targets (screen IDs)
ZINC000040165249	11-beta-Hydroxyandosterone-3-glucuronide	No	nsp16 (41), TMPRSS2 (6)
ZINC000013540864	3'-Azido-3'-deoxy-5'-O-beta-D-glucopyranuronosylthymidine	No	nsp13 (33)
ZINC000040164488	5-alpha-Dihydrotestosterone glucuronide	No	ORF7a (9), nucleoprotein (44)
ZINC0000268774715	Acalabrutinib	Yes	nucleoprotein (45)
ZINC000003891599	Aldosterone	Yes	nsp14 (37)
ZINC000066166864	Alectinib	No	nsp12 (32)
ZINC000014261579	Alvesco	Yes	nsp13 (34)
ZINC000014768568	Aprepitant	No	nsp14 (38)
ZINC000003973334	Aromasin	No	nucleoprotein (44)
ZINC000012504271*	Atovaquone	No	nsp15 (39), nsp16 (40, 41), nucleoprotein (44, 45)
ZINC0000116473771*	Atovaquone	Yes	nsp12 (31), nsp13 (33), macrodomain (10, 11), nucleoprotein (45)
ZINC000003872931	Avapro	No	macrodomain (11)
ZINC000003932831*	Avodart	Yes	nsp13 (34)
ZINC000009232411	Benperidol	No	macrodomain (10)
ZINC000004097384	Betamethasone	Yes	nsp14 (38)
ZINC000019203128	Biosone	No	M ^{pro} (19), nsp10 (26), nsp13 (34)
ZINC000003977868	Biosone	No	nsp10 (17)
ZINC000019203134	Biosone	No	nsp10 (26), nsp12 (29), nsp13 (34)
ZINC000019203131	Biosone	No	nsp13 (34)
ZINC000011677911	Bisectrizole	No	nsp12 (32), nsp13 (34), nsp8 (21), nucleoprotein (43)
ZINC0000148723177	Bimatoprost	No	nsp12 (29), nsp13 (35)
ZINC000018364228	Bucifazine	No	nsp12 (31)
ZINC000003977981	Capex	No	M ^{pro} (17)
ZINC000043195321*	Capmatinib	No	M ^{pro} (16), nsp12 (28), nsp15 (39), macrodomain (11)
ZINC000004215257	Cefpiramide	No	nsp13 (34, 35)
ZINC000021981222	cap74588	No	nsp12 (32), nsp14 (37)
ZINC000003993865	Calcitriol	No	M ^{pro} (16), TMPRSS2 (6)
ZINC000004098610	Daidzin	No	nucleoprotein(44)
ZINC000043710727	Dehydroepiandrosterone	No	macrodomain (10)
ZINC000003784182	Diflufen	No	nsp8 (21, 22), nsp12 (28), nucleoprotein (45)
ZINC000011677837	Eliquis	No	TMPSR2 (6)
ZINC000011679756	Eltrombopag	No	nsp14 (37), nucleoprotein (43), nsp16 (40), M ^{pro} (19)
ZINC000022751806	Enasidenib	No	ACE2 (4), M ^{pro} (16, 17)
ZINC0000245204943	Eplerenone	No	nucleoprotein (41)
ZINC0000238850852	Eprifibatid	No	nsp13 (34)
ZINC0000238850855	Eprifibatid	No	M ^{pro} (17), nsp12 (29), nucleoprotein (42)
ZINC0000238850853	Eprifibatid	No	nsp12 (28, 29, 32), nucleoprotein (43)
ZINC0000238850854	Eprifibatid	No	nsp12 (29, 31), nsp14 (38)
ZINC000052955754*	Ergolamine	No	M ^{pro} (16, 17), nsp13 (34, 35), nsp8 (22)
ZINC000003876078	Estradiol	Yes	nsp14 (38), spike (8)
ZINC000004099035	Estradiol 17b glucuronide	No	nsp10 (26), nsp13 (33), nucleoprotein (45)
ZINC000003881345	Estradiol benzoate	Yes	spike (8)
ZINC000004096817	Estradiol glucuronide	No	nsp13 (33)
ZINC0000102136963	Estril-3-glucuronide	No	nsp13 (33)
ZINC0000102136971	Estril-3-glucuronide	No	nucleoprotein (44)
ZINC000008234405	Estril-3-glucuronide*	No	nsp10 (26), nsp12 (31), nsp13 (33), ORF7a (9), nucleoprotein (44)
ZINC000004099004	Estrore glucuronide	No	nsp12 (31), nsp13 (33, 35), nucleoprotein (44)
ZINC000001481815	Exjade	No	M ^{pro} (16), nsp12 (32), RNP (44), TMPRSS2 (6)
ZINC0000038215434	FAD	No	M ^{pro} (17), nsp13 (35), RNP (42)
ZINC000003977985	Flumetasone	No	nsp15 (39)
ZINC000004097305	Flunisolide	No	nsp14 (38)
ZINC000001543181	Ibutamoren	No	nsp14 (41)
ZINC000013986658	Idelalisib	No	M ^{pro} (17), nsp13 (33)
ZINC000019632618	Imatinib	Yes	nsp13 (35), nsp14 (37)
ZINC000010005073	Indigo carmine	No	nucleoprotein (44)
ZINC0000261527149	Neiramezin	Yes	nsp12 (29), nucleoprotein (45)
ZINC00000537877	Ketanserin	No	TMPSR2 (6)
ZINC000031417974	Ketoprofen glucuronide	No	nsp13 (33)
ZINC000003781738*	Lestaurtinib	No	M ^{pro} (16)
ZINC000001290884	Lorazepam glucuronide	No	nsp13 (33)
ZINC00000603452*	Lumacaftor	No	M ^{pro} (17), nsp12 (26), nsp13 (33), nsp14 (37), tmprss2 (6)
ZINC0000118912516	Meprednisone acetate	No	M ^{pro} (19)
ZINC0000100017856*	Mepron	Yes	nsp15 (39), nsp16 (40), RNP (43, 44, 45)
ZINC000013831810	Mizolastine	No	M ^{pro} (16), nsp12 (30)
ZINC000003874467	Nafamostat	Yes	ORF7a (9), p1pro (13), nucleoprotein (44)
ZINC000003881613	Nandrolone phenpropionate	No	nsp12 (31), spike (8)
ZINC000011681563	Netupitant	No	nsp12 (31), nsp14 (36), RNP (42, 43)
ZINC000006716957*	Nilotinib	No	M ^{pro} (16, 17), nsp14 (37), RNP (43), tmprss2 (6)
ZINC000031460590	Oxazepam glucuronide	No	nsp13 (33)
ZINC000031460595	Oxazepam glucuronide	No	M ^{pro} (19), TMPSR2 (6)
ZINC0000095626782	Oxytetracycline	Yes	nucleoprotein (44)
ZINC000001481956	Paliperidone	No	macrodomain (11), RNP (44), tmprss2 (6)
ZINC000004214700	Paliperidone	No	nucleoprotein (43)
ZINC00000538152	Pamoic acid	No	macrodomain (10)
ZINC000011617039	Pazopanib	No	nucleoprotein (44)
ZINC000004217252	Penfluridol	No	ace2 (5), nsp8 (22), RNP (43)
ZINC000034138264	Pilafutrin	No	M ^{pro} (17), nsp12 (29, 32)
ZINC000023849850	Riluzin	No	nsp12 (29, 32)
ZINC000003816514	Relapitant	No	nsp14 (37), nsp16 (41)
ZINC00000598829	Rupatadine	No	nucleoprotein (43)
ZINC00000898237	Rutacarpine	No	macrodomain (11)
ZINC000004096846	Rutacarpine	No	M ^{pro} (17), nsp12 (29)
ZINC000017209342	Rutin	No	nsp12 (28)
ZINC000010074262	Solobvir	Yes	p1pro (15)
ZINC00000601275*	Talinflumate	No	ace2 (5), M ^{pro} (16, 19), nsp15 (39), nsp8 (22)
ZINC000001844627	Talinflumate	No	ace2 (5), M ^{pro} (16), nsp13 (33), nsp15 (39)
ZINC00000538415	Tamibarotene	No	nsp12 (28)
ZINC000011525121	Tetracycline hydrochloride	Yes	nsp14 (38), nsp16 (40), nucleoprotein (45)
ZINC000003812889	Tibolone	No	nsp12 (31)
ZINC000003938684	Toposar	Yes	nsp16 (40)
ZINC000003925861	Vorapaxar	No	ACE2 (4)
ZINC0000218037687	Voxilaprevir	No	nsp12 (29, 30, 32), nucleoprotein (43)
ZINC000000577115*	Xaliproden	No	nsp12 (28), nsp14 (36), macrodomain (11), tmprss2 (6)
ZINC000003927200	Yaz, Drospirone	No	nsp12 (31), nsp15 (39), macrodomain (10), RNP (43)
ZINC00005293666	Zalcitabine	No	M ^{pro} (17)
ZINC000003797541	Zylgta	No	nsp12 (31)

Table S46. Approved drugs amongst the virtual screening hits. Related to Table 1. A table of "now approved" drugs that bind to SARS-CoV-2 proteins as identified by VirtualFlow. These drugs are mined from the top 1% of the hits from "in stock" compounds of the ZINC library. The drugs presented here are further filtered from the top 1% of hits using a docking score threshold of better than -8 kcal/mol. Some of the drugs are predicted to bind to multiple protein targets, and for some of the drugs (denoted by *) only the targets of the top five virtual screens as judged by the docking score are shown. Some of the drugs shown here are currently in clinical trials for the treatment of COVID-19 (highlighted in light green).

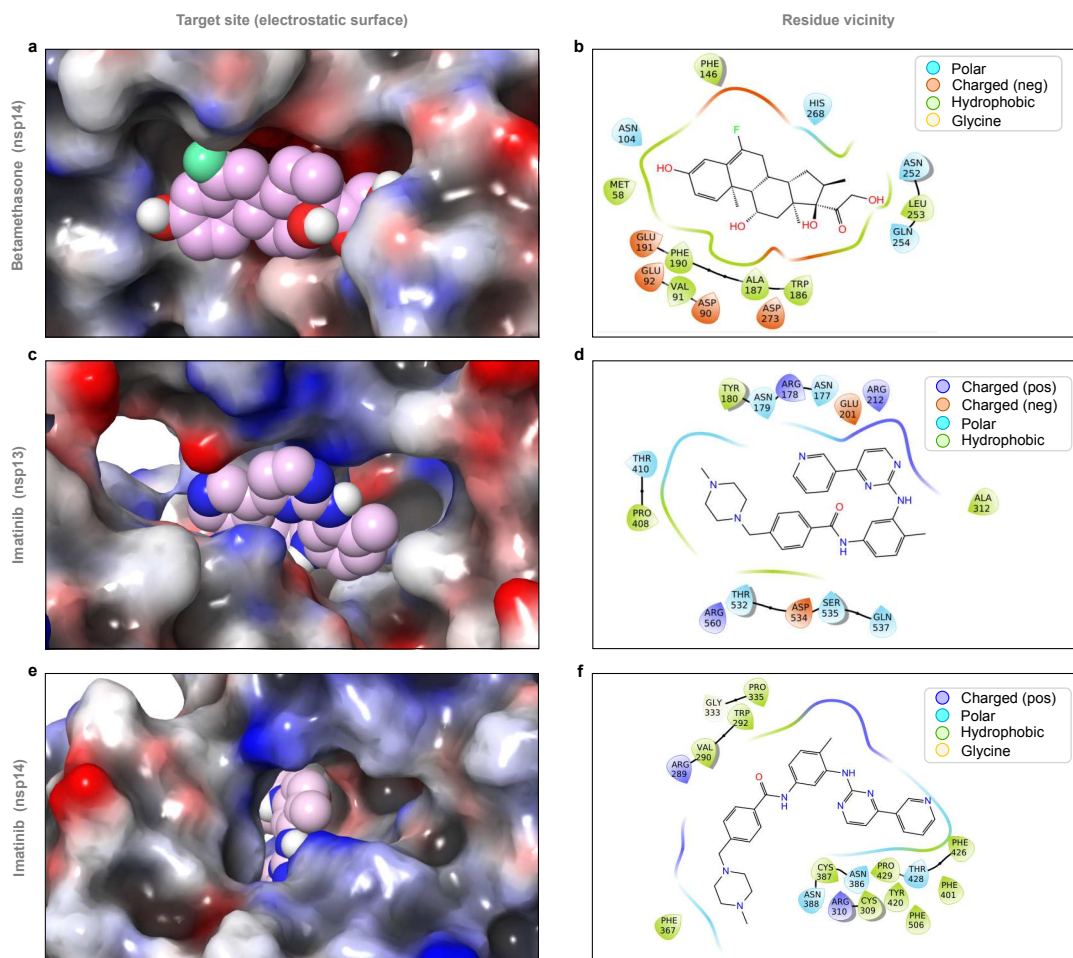


Figure S47. Docking poses of two example approved drug hits that came up in the top 1% of the hits from “in stock” compounds of the ZINC library, bound to their respective targets. a, Electrostatic surface of the nsp14 N7 methyltransferase active site (N7-MTase, screen ID: 38) with betamethasone **b,** An overview of the interactions between betamethasone and the N7-MTase active site pocket of nsp14 **c,** Electrostatic surface of the RNA binding pocket of the nsp 13 helicase (screen ID: 35) with imatinib bound **d,** An overview of interactions between imatinib and the RNA binding pocket of the helicase, nsp13 **e,** Electrostatic surface representation of the nsp14 exoribonuclease (ExoN) active site (screen ID: 37) with imatinib bound **f,** An overview of the interactions between imatinib and the nsp14 ExoN active site.

Compound name	ZINC_ID	Human gene	Viral bait	Indication	Clinical Trial identifier	Target 1 (screen ID)	Docking score	Target 2 (screen ID)	Docking score	Target 3 (screen ID)	Docking score
Nafamostat	ZINC3874467	Cell Entry	-	Anticoagulant	NCT04352400, NCT04418128	orf7a (9)	-8	plpro(13)	-8.5	RNP(43)	-8.5
Pazopanib	ZINC11617039	MARK2	orf9b	Cancer	-	RNP (43)	-8.3				
progesterone	ZINC4428529	SIGMAR1/SIGMAR2	Nsp6/Orf9c	birth control	NCT04365127	plpro (12)	-7.1				
Pimozide	ZINC4175630	SIGMAR1/SIGMAR2	Nsp6/Orf9c	Tourette syndrome	-	nsp16 (41)	-7.8	nsp7(20)	-7	spike(7)	-7.9
Silmitasertib	ZINC58638454	CSNK2A2	-	Cancer	under consideration	nsp16(40)	-9.6				

Table S48. A list of drug hits from our virtual screens found in common with a previously published SARS-COV-2 interactome study (Gordon et al., 2020). These drugs, upon experimental validation, could potentially be developed into multi-target drugs capable of modulating both host and viral targets simultaneously. Three of the drugs listed here are currently undergoing clinical trials for the treatment of COVID-19 (highlighted in light green).

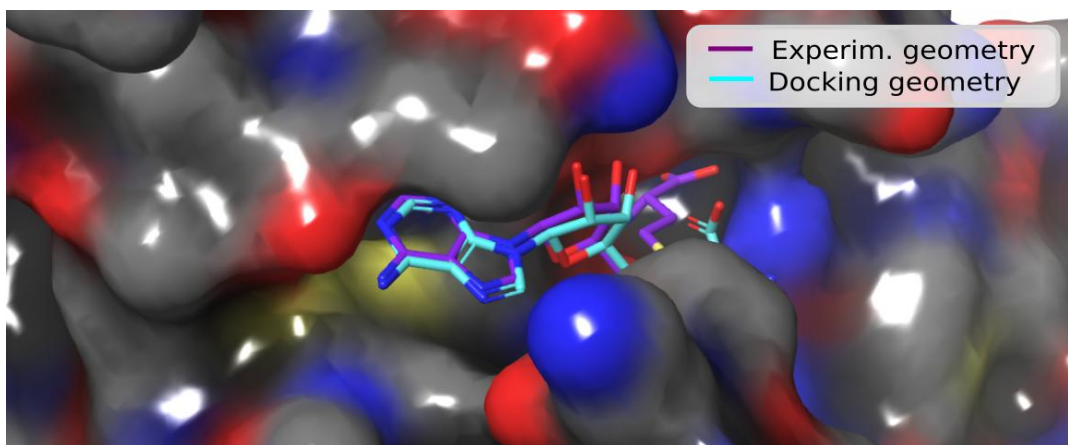


Figure S49. Redocking geometries for PDB structure 7jhe (nsp16). Electrostatic surface of SARS-CoV-2 nsp16 is shown. Redocking of the ligand S-Adenosyl-L-homocysteine (SAH) from the PDB structure 7jhe was performed using the same docking settings as in the primary virtual screening for this target. The docking program used was QuickVina 2 and the exhaustiveness parameter was set to 1. The experimentally reported binding geometry is shown in violet, while the docked compound is shown in cyan. The RMSD between the two geometries is 2.25 Å.

Compound SMILES	Receptor	Source
<chem>C[C@@H](NC(C1=CC(NC(C)=O)=CC=C1C)=O)C2=CC=CC3=CC=CC=C32</chem>	PLpro	Ghosh et al., 2009
<chem>C[C@@H](N1CC(C)(NCC2=CC(OCC3)=C3C=C2)=O)CC1)C4=CC=CC5=CC=CC=C54</chem>	PLpro	Ghosh et al., 2009
<chem>NC1=NC=NC2=C1N=C(C(F)F)N2</chem>	PLpro	Lee et al., 2015
<chem>CC1(C)CCC2=C1C=CC(C)(OCC3C)=C3C4=O=C2C4=O</chem>	PLpro	Park et al., 2012
<chem>O=C(C=C/C/C=C1=CC=C(O)(C)=C1)CCC2=CC(O)=C(O)C=C2</chem>	PLpro	Park et al., 2012
<chem>OC1=C(O)C=C(C=C1)[C@@H](O2)[C@@H](O)C(C)(C2=C3)=C(O)C4=C3O[C@@](C(C)(C)(C)O)(C)CC4=O</chem>	PLpro	Park et al., 2017
<chem>C[C@@H](NC(C1=CC=CC=C1C)=O)C2=CC(C=C=C3)=C3C=C2</chem>	PLpro	Ghosh et al., 2009
<chem>C[C@@H](NC(C1=CC=CC=C1C)=O)C2=CC=CC3=CC=CC=C32</chem>	PLpro	Ghosh et al., 2009
<chem>C[C@@H](NC(C1=CC(N)=CC=C1C)=O)C2=C(C=C=C3)C3=CC=C2</chem>	PLpro	Ratia et al., 2008
<chem>C[C@H](C1=CC=CC2=CC=CC=C2)N3CCC(CC3)C(NCC4=CC=CC(F)=C4)=O</chem>	PLpro	Pillaiyar et al., 2020
<chem>O=C(C1=CC(C=CN2)=C2C=C1)ON3N=NC4=CC=CC=C43</chem>	Mpro	Wu et al., 2006
<chem>C1C1=CC(OC(C2=CC=CC2)=O)=CN=C1</chem>	Mpro	Zhang et al., 2007
<chem>C1C1=CC(OC(C2=CC=CC2=C2C=CN3)=O)=CN=C1</chem>	Mpro	Ghosh et al., 2008
<chem>O=C(C1=CC=CC(Br)=C1N2CC3=CC(C=CC=C4)=C4S3)C2=O</chem>	Mpro	Chen et al., 2005
<chem>IC1=CC=C(N(CC2=CC(C=CC=C3)=C3S2)C(C4=O)=O)C4=C1</chem>	Mpro	Chen et al., 2005
<chem>O=C(C1=CC(C(N)=O)=CC=C1N2CC3=CC(C=CC=C4)=C4C=C3)C2=O</chem>	Mpro	Zhou et al., 2006
<chem>C1C1=CC=C(C(O)=O)C(C2=CC=C(C=C3C(N(C4=CC=C(C(C)C)C=C4)N=C3C5=CC=CC=C5)=O)O2)=C1</chem>	Mpro	Kumar et al., 2016
<chem>O=C(N(C1=CC=C(C(O)=O)C=C1)N=C2)C2=C(C3=CC=C(F)C=C3</chem>	Mpro	Ramajayam et al., 2010
<chem>Fc1cc(F)cc(CCC(=O)N2C[CH]3CCC[CH]3[CH]2C(=O)N[CH][C][CH]4CCNC4=O)C=O)c1</chem>	Mpro	PDB ID: 7d3i
<chem>C(C(NC(=O)C(NC(c2cc1c(O)cccc1n2)=O)CC(C)C)CC3CCNC3=O)(c5nc4c(ccc4)s5)O</chem>	Mpro	PDB ID: 7jkv

Table S50. Experimentally confirmed positive compounds reported in the literature used for the enrichment studies with M^{pro} and PL^{pro}.

Compound SMILES	Receptor	Source
<chem>C[C@]1(c2cccc(CNc3ccc4nnc(-c5ccccc5F)n4n3)c2)NC(=O)NC1=O</chem>	Mpro	screen ID 16
<chem>Cc1cc(-c2cc(C(=O)NCc3cccc([C@@]4(C)NC(=O)NC4=O)c3)c3c(C)noc3n2)c(C)c1</chem>	Mpro	screen ID 16
<chem>Nc1ncc2c(n1)CCN(C(=O)c1cc(F)cc3c(=O)c4cc(F)ccc4[nH]c13)C2</chem>	Mpro	screen ID 16
<chem>C[C@@]1(c2cccc(CNC(=O)[C@H]3CC(=O)N(c4cccc5ccccc45)C3)c2)NC(=O)NC1=O</chem>	Mpro	screen ID 16
<chem>O=C1NC(=O)/C(=N/c2cccc3c(O)nnc(O)c23)[C@@H](c2nc3ccccc3s2)C1=O</chem>	Mpro	screen ID 16
<chem>Cc1noc2nc(-c3ccc(F)cc3)cc(C(=O)NCc3cccc([C@@]4(C)NC(=O)NC4=O)c3)c12</chem>	Mpro	screen ID 16
<chem>O=C([C@@H]1[C@@H]2CCCC[C@H]2N1C(=O)c1ccc(F)cc1F)N1CC=C(c2c[nH]c3ccccc23)CC1</chem>	Mpro	screen ID 16
<chem>C[C@@]1(c2cccc(C(=O)N3CCC[C@H](c4n[nH]c(-c5ccccc5)n4)C3)c2)NC(=O)NC1=O</chem>	Mpro	screen ID 16
<chem>C[C@@H](OC(=O)c1cc(-c2ccncc2)nc2ccccc12)C(=O)NC(=O)NC12CC3CC(C(C3)C1)C2</chem>	Mpro	screen ID 16
<chem>Cc1ccc(-c2cc(C(=O)NCc3cccc([C@@]4(C)NC(=O)NC4=O)c3)c3c(C)noc3n2)cc1</chem>	Mpro	screen ID 16
<chem>NC(=O)c1nc(Cl)ccc1NC(=O)c1ccc2c(c1)C[C@@H](c1ccccc1)OC2=O</chem>	PLpro	screen ID 13
<chem>NC(=O)c1nc(Cl)ccc1NC(=O)c1ccc2c(=O)n3c(nc2c1)CCCCC3</chem>	PLpro	screen ID 13
<chem>CC(C)(C)OC(=O)Nc1ccc(NC(=O)N[C@H]2C[C@@H]3C[C@@H]2[C@H]3C[C@@H]2[C@@H]2CCCC[C@H]32)cc1</chem>	PLpro	screen ID 13
<chem>O=C(Nc1cccc2[nH]c(=O)oc12)[C@H]1CC(=O)N(c2ccc3c(c2)Cc2ccccc2-3)C1</chem>	PLpro	screen ID 13
<chem>O=C(CC12C[C@H]3C[C@@H](C1)CC(C(=O)Nc1nc4ccccc4[nH]1)(C3)C2)Nc1nc2ccccc2[nH]1</chem>	PLpro	screen ID 13
<chem>NC(=O)c1nc(Cl)ccc1NC(=O)c1ccc2c(c1)S(=O)(=O)c1ccccc1C2=O</chem>	PLpro	screen ID 13
<chem>O=C(NNc1ccc(-n2ccnc2)nn1)c1ccc2c(c1)C[C@@H](c1ccccc1)OC2=O</chem>	PLpro	screen ID 13
<chem>C[C@@H](OC(=O)c1ccc2c(c1)C(=O)N(C)C2=O)C(=O)Nc1ccc2c(c1)C(=O)c1ccccc1C2=O</chem>	PLpro	screen ID 13
<chem>C[C@@]1(c2cccc(CNC(=O)NNC(=O)c3cccc4ccccc34)c2)NC(=O)NC1=O</chem>	PLpro	screen ID 13
<chem>O=C(NNc1ccc(-c2ccccc2)nn1)c1ccc2c(c1)C[C@@H](c1ccccc1)OC2=O</chem>	PLpro	screen ID 13

Table S51. Ten top virtual screening hits for M^{pro} and PL^{pro} obtained from Screens 16 and 13, respectively.

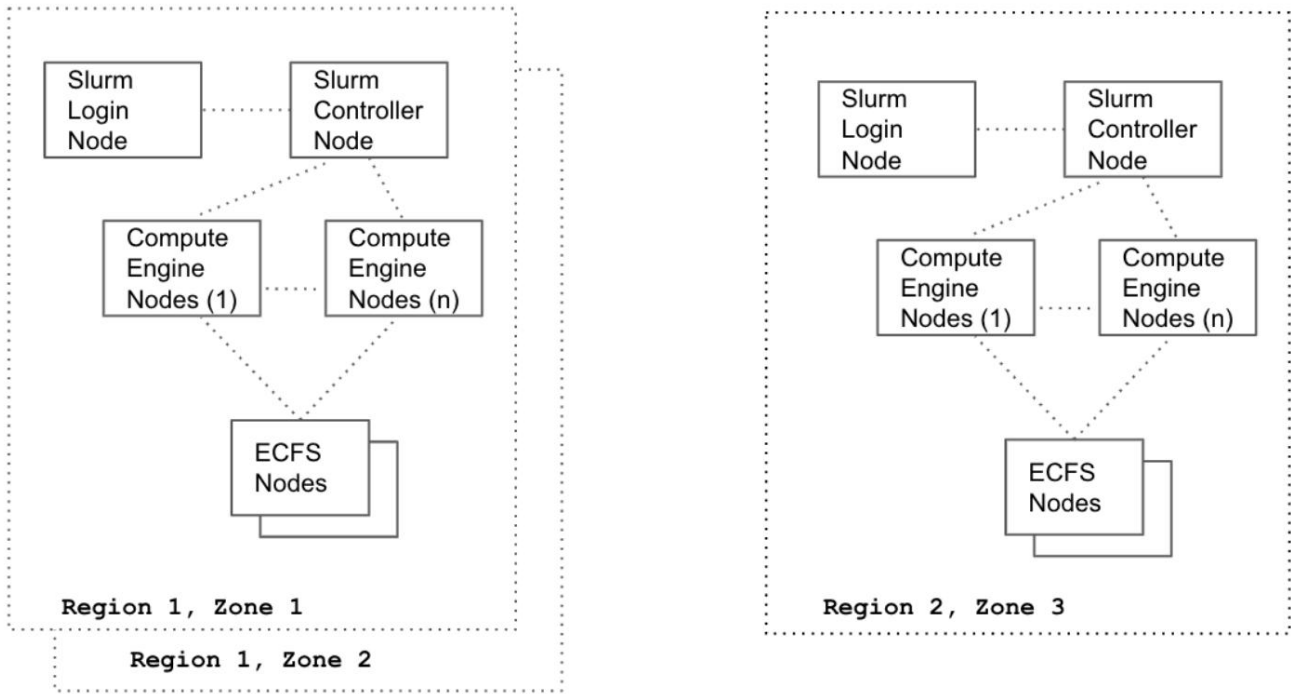


Figure S52. Google Cloud Platform (VPC-1)

Nodes	GCE instance type
Slurm controller node	n1-standard-16
Slurm login node	n1-standard-8
Slurm compute nodes	n2d-highcpu-64 (cluster 1 and 2), n1-highcpu-32 (cluster 3)
Elastifile management node	n1-standard-8
ECFS storage node	custom (16 vCPUs, 96 GB memory)

Table S53. Virtual machine types employed in the Google Cloud Platform.

References.

- Alhossary, A., Handoko, S. D., Mu, Y., & Kwoh, C.-K. (2015). Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, 31(13), 2214–2216. <https://doi.org/10.1093/bioinformatics/btv082>
- Chen, L.-R., Wang, Y.-C., Lin, Y. W., Chou, S.-Y., Chen, S.-F., Liu, L. T., Wu, Y.-T., Kuo, C.-J., Chen, T. S.-S., & Juang, S.-H. (2005). Synthesis and evaluation of isatin derivatives as effective SARS coronavirus 3CL protease inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 15(12), 3058–3062. <https://doi.org/10.1016/j.bmcl.2005.04.027>
- Ghosh, A. K., Gong, G., Grum-Tokars, V., Mulhearn, D. C., Baker, S. C., Coughlin, M., Prabhakar, B. S., Sleeman, K., Johnson, M. E., & Mesecar, A. D. (2008). Design, synthesis and antiviral efficacy of a series of potent chloropyridyl ester-derived SARS-CoV 3CLpro inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 18(20), 5684–5688. <https://doi.org/10.1016/j.bmcl.2008.08.082>
- Ghosh, A. K., Takayama, J., Aubin, Y., Ratia, K., Chaudhuri, R., Baez, Y., Sleeman, K., Coughlin, M., Nichols, D. B., Mulhearn, D. C., Prabhakar, B. S., Baker, S. C., Johnson, M. E., & Mesecar, A. D. (2009). Structure-Based Design, Synthesis, and Biological Evaluation of a Series of Novel and Reversible Inhibitors for the Severe Acute Respiratory Syndrome-Coronavirus Papain-Like Protease. *Journal of Medicinal Chemistry*, 52(16), 5228–5240. <https://doi.org/10.1021/jm900611t>
- Hassan, N. M., Alhossary, A. A., Mu, Y., & Kwoh, C.-K. (2017). Protein-Ligand Blind Docking Using QuickVina-W With Inter-Process Spatio-Temporal Integration. *Scientific Reports*, 7(1), 15451. <https://doi.org/10.1038/s41598-017-15571-7>
- Keserü, G. M., & Swinney, D. C. (2015). *Thermodynamics and Kinetics of Drug Binding* (G. M. Keserü & D. C. Swinney (eds.)). Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527673025>
- Kumar, V., Tan, K.-P., Wang, Y.-M., Lin, S.-W., & Liang, P.-H. (2016). Identification, synthesis and evaluation of SARS-CoV and MERS-CoV 3C-like protease inhibitors. *Bioorganic & Medicinal Chemistry*, 24(13), 3035–3042. <https://doi.org/10.1016/j.bmc.2016.05.013>
- Lee, H., Lei, H., Santarsiero, B. D., Gatuz, J. L., Cao, S., Rice, A. J., Patel, K., Szypulinski, M. Z., Ojeda, I., Ghosh, A. K., & Johnson, M. E. (2015). Inhibitor Recognition Specificity of MERS-CoV Papain-like Protease May Differ from That of SARS-CoV. *ACS Chemical Biology*, 10(6), 1456–1465. <https://doi.org/10.1021/cb500917m>
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/10.1002/jcc.21256>
- Park, J.-Y., Yuk, H. J., Ryu, H. W., Lim, S. H., Kim, K. S., Park, K. H., Ryu, Y. B., & Lee, W. S. (2017). Evaluation of polyphenols from *Broussonetia papyrifera* as coronavirus protease inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 32(1), 504–512. <https://doi.org/10.1080/14756366.2016.1265519>
- Pillaiyar, T., Wendt, L. L., Manickam, M., & Easwaran, M. (2021). The recent outbreaks of human coronaviruses: A medicinal chemistry perspective. *Medicinal Research Reviews*, 41(1), 72–135. <https://doi.org/10.1002/med.21724>
- Ramajayam, R., Tan, K.-P., Liu, H.-G., & Liang, P.-H. (2010). Synthesis and evaluation of pyrazolone compounds as SARS-coronavirus 3C-like protease inhibitors. *Bioorganic & Medicinal Chemistry*, 18(22), 7849–7854. <https://doi.org/10.1016/j.bmc.2010.09.050>
- Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S., Chaudhuri, R., Fu, W., Prabhakar, B. S., Johnson, M. E., Baker, S. C., Ghosh, A. K., & Mesecar, A. D. (2008). A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proceedings of the National Academy of Sciences*, 105(42), 16119–16124. <https://doi.org/10.1073/pnas.0805240105>
- Sander, T., Freyss, J., von Korff, M., & Rufener, C. (2015). DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *Journal of Chemical Information and Modeling*, 55(2), 460–473. <https://doi.org/10.1021/ci500588j>
- Sterling, T., & Irwin, J. J. (2015). ZINC 15 – Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling*, 55(11), 2324–2337. <https://doi.org/10.1021/acs.jcim.5b00559>
- Wu, C.-Y., King, K.-Y., Kuo, C.-J., Fang, J.-M., Wu, Y.-T., Ho, M.-Y., Liao, C.-L., Shie, J.-J., Liang, P.-H., & Wong, C.-H. (2006). Stable Benzotriazole Esters as Mechanism-Based Inactivators of the Severe Acute Respiratory Syndrome 3CL Protease. *Chemistry & Biology*, 13(3), 261–268. <https://doi.org/10.1016/j.chembiol.2005.12.008>

Zhang, J., Pettersson, H. I., Huitema, C., Niu, C., Yin, J., James, M. N. G., Eltis, L. D., & Vederas, J. C. (2007). Design, Synthesis, and Evaluation of Inhibitors for Severe Acute Respiratory Syndrome 3C-Like Protease Based on Phthalhydrazide Ketones or Heteroaromatic Esters. *Journal of Medicinal Chemistry*, 50(8), 1850–1864. <https://doi.org/10.1021/jm061425k>

Zhou, L., Liu, Y., Zhang, W., Wei, P., Huang, C., Pei, J., Yuan, Y., & Lai, L. (2006). Isatin Compounds as Noncovalent SARS Coronavirus 3C-like Protease Inhibitors. *Journal of Medicinal Chemistry*, 49(12), 3440–3443. <https://doi.org/10.1021/jm0602357>