JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Article

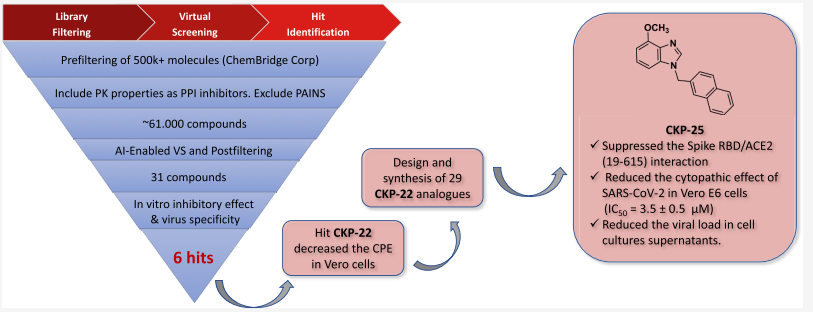
AI Promoted Virtual Screening, Structure-Based Hit Optimization, and Synthesis of Novel COVID-19 S-RBD Domain Inhibitors

Ioannis Gkekas, Sotirios Katsamakas, Stelios Mylonas, Theano Fotopoulou, George E. Magoulas, Alia Cristina Tenchiu, Marios Dimitriou, Apostolos Axenopoulos, Nafsika Rossopoulou, Simona Kostova, Erich E. Wanker, Theodora Katsila, Demetris Papahatjis, Vassilis G. Gorgoulis, Maria Koufaki,* Ioannis Karakasiliotis, Theodora Calogeropoulou,* Petros Daras, and Spyros Petrakis*

Cite This: https://doi.org/10.1021/acs.jcim.4c01110

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Coronavirus disease 2019 (COVID-19) is caused by a new, highly pathogenic severe-acute-respiratory syndrome coronavirus 2 (SARS-CoV-2) that infects human cells through its transmembrane spike (S) glycoprotein. The receptor-binding domain (RBD) of the S protein interacts with the angiotensin-converting enzyme II (ACE2) receptor of the host cells. Therefore, pharmacological targeting of this interaction might prevent infection or spread of the virus. Here, we performed a virtual screening to identify small molecules that block S-ACE2 interaction. Large compound libraries were filtered for drug-like properties, promiscuity and protein−protein interaction-targeting ability based on their ADME-Tox descriptors and also to exclude pan-assay interfering compounds. A properly designed AI-based virtual screening pipeline was applied to the remaining compounds, comprising approximately 10% of the starting data sets, searching for molecules that could bind to the RBD of the S protein. All molecules were sorted according to their screening score, grouped based on their structure and postfiltered for possible interaction patterns with the ACE2 receptor, yielding 31 hits. These hit molecules were further tested for their inhibitory effect on Spike RBD/ACE2 (19−615) interaction. Six compounds inhibited the S-ACE2 interaction in a dose-dependent manner while two of them also prevented infection of human cells from a pseudotyped virus whose entry is mediated by the S protein of SARS-CoV-2. Of the two compounds, the benzimidazole derivative **CKP-22** protected Vero E6 cells from infection with SARS-CoV-2, as well. Subsequent, hit-to-lead optimization of **CKP-22** was effected through the synthesis of 29 new derivatives of which compound **CKP-25** suppressed the Spike RBD/ACE2 (19−615) interaction, reduced the cytopathic effect of SARS-CoV-2 in Vero E6 cells (IC_{50} = 3.5 μM) and reduced the viral load in cell culture supernatants. Early in vitro ADME-Tox studies showed that **CKP-25** does not possess biodegradation or liver metabolism issues, while isozyme-specific CYP450 experiments revealed that **CKP-25** was a weak inhibitor of the CYP450 system. Moreover, **CKP-25** does not elicit mutagenic effect on *Escherichia coli* WP2 uvrA strain. Thus, **CKP-25** is considered a lead compound against COVID-19 infection.

INTRODUCTION

In March 2020, the coronavirus disease 2019 (COVID-19) caused by a novel, highly pathogenic severe-acute-respiratory syndrome coronavirus 2 (SARS-CoV-2) was declared a global pandemic by the World Health Organization (WHO).[1] COVID-19 is characterized as a serious respiratory disease that affects also the gastrointestinal system, heart, central nervous system and kidneys, with 771,679,618 confirmed cases, including 6,977,023 deaths globally, reported to WHO until November 2023.[2] The accelerated development of effective COVID-19 vaccines and other therapeutics has offered transient solutions to resolve the pandemic.[3,4]

Received: July 1, 2024
Revised: October 25, 2024
Accepted: October 29, 2024

However, the SARS-CoV-2 statistical data for the period of January 2020−January 2022 show a rapid evolution with a 4-fold increase in daily cases worldwide,[5] affirming that mankind is far from eradicating this pandemic.

SARS-CoV-2 belongs to the β-genus of the coronaviruses and its genome comprises 14 open reading frames encoding 16 non-structural proteins (nsp1−16), nine accessory proteins and four structural proteins [spike (S), envelope (E), membrane (M) and nucleocapsid (N)].[6] The human infection by SARS-CoV-2 involves the binding of the transmembrane homotrimeric spike (S) glycoprotein through its receptor binding domain (RBD) of its S1 subunit to the host protein angiotensin-converting enzyme 2 (ACE2). Then, the S2 subunit of spike fuses the host and viral membranes, enabling the entry of SARS CoV-2 genome into human cells. In addition, the transmembrane serine protease 2 (TMPRSS2), acting as major host protease is required for the priming of the S protein. During the intracellular life cycle, the virus expresses and replicates its genome RNA leading to the production of coding proteins. Then, the virion is assembled and the newly synthesized SARS-CoV-2 particles are released, surrounded by the membrane of the host cells.[7−11] Therefore, the prevention of the viral entry step represents the first point of interfering with viral transmission.

Since the beginning of the pandemic, drug repurposing[12] was used as a first option for the discovery of new therapeutics, due to shorter clinical trials, reduced risks and lower cost for development. Thus, the combination of in silico and biological strategies allowed the effective and fast selection of promising repurposed candidates for COVID-19 treatment.[13] Remdesivir, originally developed for the treatment of Ebola is a repurposed inhibitor of the SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) and was the first drug authorized by the FDA against COVID-19.[14] Other available treatment options for adults with COVID-19, involving repurposed drugs are Lagevrio[15,16] and Paxlovid.[17] Although drug repurposing may have several advantages, contrariwise it has some non-negligible disadvantages like dosing safety, specificity and delivery capability.[18] Recently, Ensitrelvir fumaric acid a 3C-like protease inhibitor developed through joint research between Hokkaido University and Shionogi Ltd. received emergency approval in Japan in November 2022 (Xocova).[19] In April 2023 the FDA has granted Fast Track designation for Ensitrelvir fumaric acid, while the NIH has initiated a multisite clinical trial for the evaluation in adults hospitalized with COVID-19.[20]

The spike glycoprotein is the antigen used in many COVID-19 vaccines[21] while the specific interaction of S(RBD) with hACE2 is the current pharmacological target of therapeutic antibodies.[22,23] However, it is subgenus and strain specific, and thus promotes the emergence of new mutated variants that in turn, result in decreased efficiency of vaccines and neutralizing antibodies.[24−30] Furthermore, recently published data highlight that vaccine effectiveness declines significantly between three and six months post-vaccination.[31] Moreover, there is a significant percentage of the population who for a variety of reasons refuse to or cannot be vaccinated.[32]

The interaction between S(RBD) and hACE2 is of outmost importance regulating the cross-species and human-to-human SARS-CoV-2 transmission.[33] Hence developing drugs that could disrupt the RBD-ACE2 interaction is of great interest and significance. Despite the difficulties in inhibiting protein−protein interactions (PPIs) with small molecules, they are more advantageous than antibodies. These molecules are less sensitive to different viral strains and mutations, can be administered as oral or inhalable formulations and result in lower immunogenic side effects.[34] Several in silico-based and in vitro screening approaches have been employed to identify potential small molecule and peptide-based inhibitors of the S(RBD)/ACE2 interaction.[34−58] Despite the absence of potent in vitro and in vivo inhibitors of this interaction to date, these results support the feasibility of targeting this PPI by small molecules.

In the current manuscript we performed an AI-enabled virtual screening to identify compounds which would block the S(RBD)-ACE2 interaction. The identified hits were validated experimentally and the most potent (**CKP-22**) was optimized through structure-based drug design (SBDD) followed by the synthesis of the proposed derivatives. The new compounds were evaluated for the inhibition of S(RBD)/ACE2 interaction and reduction of the cytopathic effect of SARS-CoV-2 in Vero E6 cells. The most potent were further assessed for preventing infection from the original SARS-CoV-2 in vitro.

RESULTS AND DISCUSSION

AI-Enabled Virtual Screening and Design. Shortly after the emergence of the COVID-19 pandemic in 2020 the crystal structure of the S(RBD)-ACE2 complex became available.[10] This shed light on important RBD domain structural features like the interface definition enclosed by the amino acids Lys417, Tyr449 and Asn487. These three pillar residues are of high significance for ACE2 binding (Figure 1).

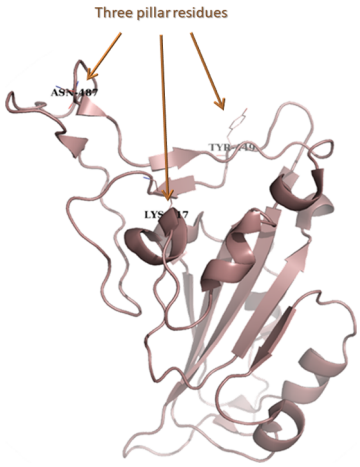


Figure 1. S(RBD) illustration as brown cartoon with the three pillar residues shown as sticks and labeled.

The lysine residue is the only amino acid sequentially belonging outside the receptor binding motif (RBM) but is recognized as highly important for the RBD-ACE2 complex formation with the contribution of two salt bridges and an equal number of hydrogen bonds. In contrast, the equivalent valine residue found in the RBD domain of SARS-CoV-1 fails to participate in the ACE2 binding interaction network, possessing a 9-fold difference in the dissociation constant.

Inspired by the example of the development of the antiviral agent Fostemsavir as a protein−protein interaction drug (PPID)[59] we set out to identify through a Virtual-Screening (VS) campaign potent PPI compounds that could disrupt the formation of the S(RBD)-ACE2 complex.

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Our team has been working on the training of a new AI-enabled pipeline[60] combining the docking result of Smina[61] with an AI-based rescoring function that is able to distinguish more efficiently potential binding complexes.[60] To alleviate possible inconsistencies originating from the unknown 3D structure of the input compounds used for screening, the docking step is applied on two different conformations of these compounds, created by the OpenBabel toolbox,[62] while 50 docked poses are generated for each compound conformation. The AI-enabled pipeline progresses in three distinct steps. First, the resulted protein−substrate solutions are rescored with our custom AI-based rescoring function,[60] which comprises a 3D convolutional neural network (CNN) that was trained on the DUD-E benchmark[63] with the aim to automatically capture the underlying mechanism responsible for the protein−ligand binding. The 103 proteins present in the DUD-E data set were split to algebraic 3 folds based on their sequence similarity, and 3-fold training was performed for hyperparameters selection (Figure 2A). A subsequent retraining on the whole DUD-E data set provided the final model which is used to rescore the predicted protein-compound complexes. Second, in an attempt to further enhance the CNN performance, a fusion approach is applied which combines under a unified scheme the AI-scores with the Smina predicted affinities, taking into account both conformations and multiple docked poses for each compound. The third and final step comprises cluster grouping of the potential hits based on their structural similarity utilizing a 3D shape descriptor[64] to ensure diversity among them.

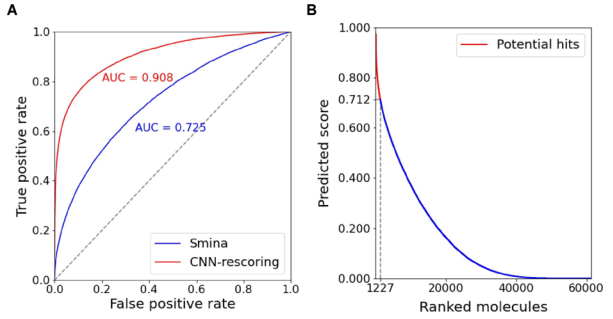


Figure 2. Virtual screening. (A) Comparison between the adopted CNN-rescoring approach and Smina on the DUD-E benchmark data set. (B) Molecule ranking scores after applying our VS pipeline on the initially filtered library (~60K molecules) and selection of the top-2% (1227).

With this newly trained AI promoted algorithm in hand we performed a virtual screening (VS) approach on the complete RBD surface against three commercial libraries (i.e., macrocycle, EXPRESS-Pick Stock & Coronavirus dedicated libraries) from ChemBridge Corporation, counting more than half a million compounds in total. All libraries were filtered prior to their use according to a drug-like subset of pharmacokinetic (PK) properties to exclude non-favorable compounds from the process and minimize the load. Filtering included descriptors like Lipinski's rule of 5,[65] Veber[66] and Egan[66] bioavailability scores. The two first take into consideration the number of rotatable bonds and tPSA, while Egan correlates tPSA with lipophilicity. Additional filters included phospholipidosis,[67] GSK 4/400 rule,[67] the Pfizer 3/75 rule[68] and Lilly Medchem rule predictions.[69] The latter can identify potential interference with biological assays, such as substrate chemical reactivity, instability and/or lack of druggability. In detail, the libraries were filtered utilizing FAF-drugs4[70] to eliminate pan-assay interfering compounds (PAINS) and unwanted metabolites and, afterward, the Eli Lilly MedChem set of rules[69] along with favorable PPI profile were applied.[71] PPIs peculiar features do not match traditional binding pockets which have in general larger, flatter and more hydrophobic surfaces. Hence, they bind to narrow PK windows for several descriptors such as molecular weight around 421, log P values in the range of 3.56 and topological polar surface area (tPSA) of 89 Å^2. These values are higher than usual binding pocket properties encountered in drugs.[72] The above filtering resulted in the selection of 61,349 compounds that met the required criteria and acted as our screening library for the upcoming VS stage.

In the VS step as target protein we used the crystal structure of the spike protein S1 subunit bound to the extracellular domain of the ACE2 receptor (PDB ID: 6M0J), solved by cryo-EM at a 2.45 Å resolution.[10] From this complex, the coronavirus's spike protein was cleaned, and blind docking was applied on the RBD domain of the S-protein with the ACE2. Our CNN scoring scheme assigned to the vast majority of the screened compounds a very low (near zero) score index, and only few of them were distinguished (Figure 2B). Furthermore, we decided to qualify as potential hits the top-2% which corresponds to 1,227 compounds with a unified score higher than 0.712. The structural-oriented grouping generated 248 clusters which were subsequently subjected to a secondary filtering applying a cutoff score above 0.85 of accepted compounds based on expert decision-making, including visualization of VS solutions. This revealed a scattering of all compounds around three distinct regions on the RBD domain

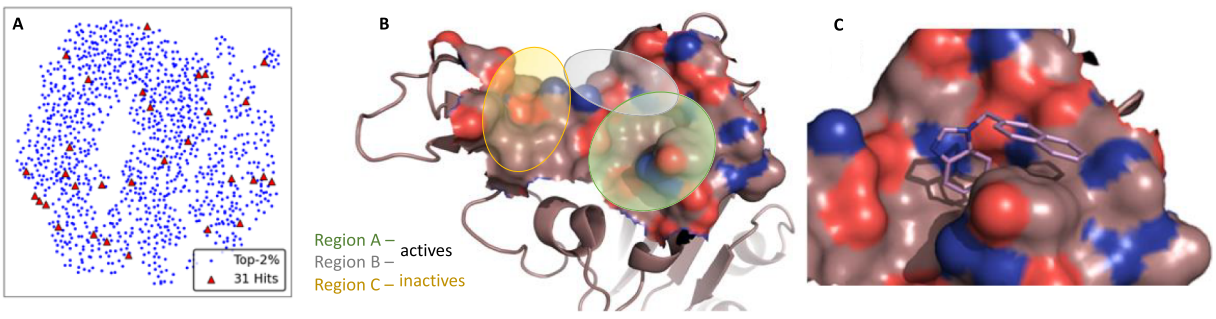


Figure 3. (A) UMAP visualization of the finally selected hits (n = 31) compared to the top-2%, on a reduced feature space of shape descriptors. (B) Crystal structure of wt-Spike protein (PDB ID: 6M0J) colored in brown as surface representation for the RBD domain. (C) Docking solution of hit compound **CKP-22** (Chembridge code: 5979349) shown as magenta-colored sticks residing in region A of the RBD domain.

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

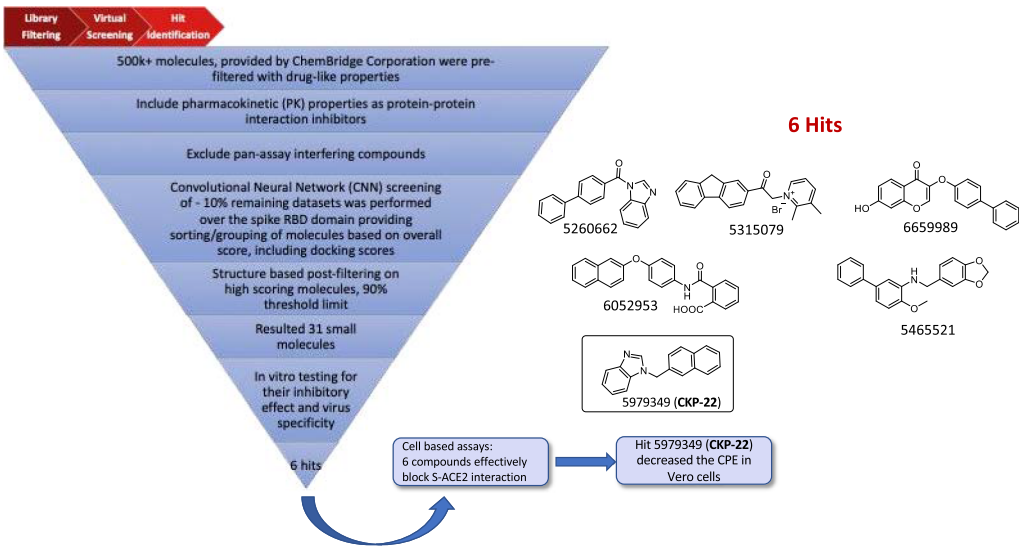


Figure 4. Virtual screening workflow, structures of identified hit compounds inhibiting S(RBD)-ACE2 interaction in a dose-dependent manner and of inhibitor of SARS-CoV-2 entry hit 5979349 (**CKP-22**).

(Figure 3), centering around the previously described pillar residues important for the ACE2 interactions. Hence, this post-filtering step resulted in the selection of the final 31 potential hits (see Table S1) that are quite diverse in shape (Figure 3A) and representatively cover all highlighted RBD regions (Figure 3B). Their biological evaluation, as described below (see Biological Evaluation), revealed that compounds fitting Region C are inactive, while compounds residing solely in Region A or dually in Regions A and B showed inhibitory activity (Figure 3B). Regions A and B are formed from the amino acids Arg403, Glu406, Lys417, Tyr449, Tyr453, Leu455, Gln493, Tyr495, Gly496, Phe497, Gln498, Asn501, Tyr505 and Gln506.

Among the 31 evaluated hit compounds six effectively blocked the S-ACE2 interaction and two prevented the infection of human cells from a pseudotyped virus whose entry is mediated by the S protein of SARS-CoV-2. In particular, the 1-substituted benzimidazole derivative Chembridge code: 5979349 (**CKP-22**) strongly prevented infection from the original SARS-CoV-2 in vitro at a 50 μM concentration. Docking of **CKP-22** revealed that the only witnessed interaction with the protein was a pi-pi stacking between Tyr505 residue and the naphthyl substituent (Figure 3C).

A summary of the virtual screening workflow leading to the hit compound **CKP-22** is presented in Figure 4.

The benzimidazole privileged scaffold has been extensively explored among other heterocycles in drug discovery.[73,74] In particular, compounds possessing the benzimidazole moiety were previously identified through VS as binders at RBD/ACE2 interface.[75] Furthermore, Mudi et al. synthesized 5-membered heterocycle-substituted benzimidazole derivatives with in vitro anti SARS-CoV-2 activity. Molecular docking and MD simulations proposed that this activity is mediated by interaction with main protease (M^{pro}) and non-structural proteins.[76,77]

Following our validation of **CKP-22** as an anti-SARS-CoV-2 compound inhibiting RBD/ACE2 interaction, a SBDD approach was implemented in order to introduce additional interactions with the S-protein and increase potency. As summarized in Figure 5, we varied the substituents at positions

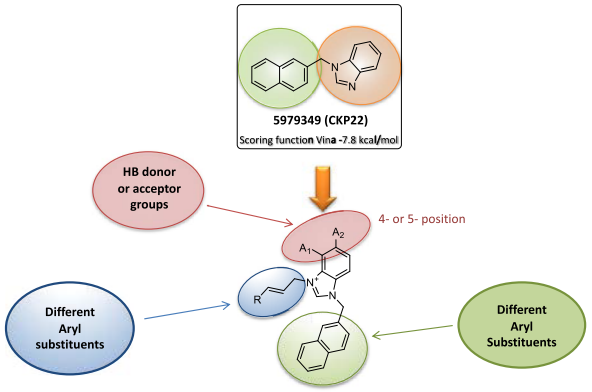


Figure 5. Design strategy of 5979349 (**CKP-22**) derivatives.

C-4 and C-5 and incorporated different aryls replacing the naphthyl group at N^1 in **CKP-22.** In an effort to extend the interaction network toward region B of the RBD domain we introduced substituents on N^3 bearing a trans vinyl group stereochemistry in order to reach suitable residues (i.e., Tyr449) of this site.

The S-protein has been continuously and rapidly mutating providing numerous variants during the evolution of the COVID-19 pandemic. However, the three pillar residues were largely maintained across variants. Our design aiming to keep the Tyr505 interaction proved crucial since this residue was one of the most conserved. The only alteration was found in the omicron variant mutations where Tyr505 was replaced by a histidine residue. Moreover, when comparing the wt with other variants it was found that Tyr449 and Asn487 are conserved, while the 417 amino acid is replaced by a Asn in beta,[78] Thr in gamma,[78] delta[79] and kappa[80] variants maintaining wt relevance, Val in omicron BA.1 variant[81] and a Asn mutation for BA.2/3/4/5.[82−84] All the above represent targeting/interacting residues based in our docking studies. Taking into account the epidemiological manifestation of each variant it can be understood that the most severe forms among the variants had Lys417 and Tyr505 present. On regard to regions A and B identified from our study mutation-wise (Figure 6 and Table S2), wt exhibits no difference with delta and kappa

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

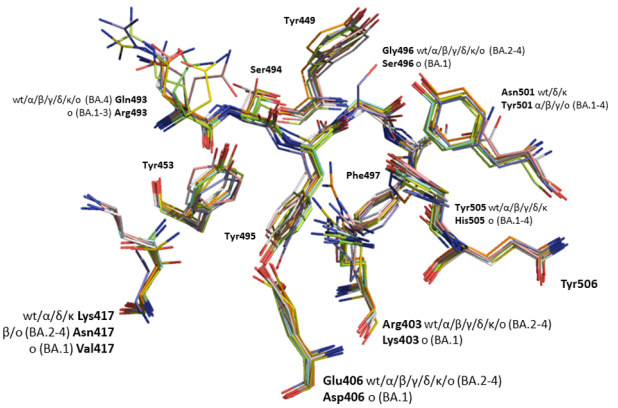


Figure 6. Overlay of COVID-19 line representation for S1 spike glycoprotein residues consisting described regions A and B for wt (PDB entry 6m0j), alpha (PDB entry 7ekf), beta (PDB entry 7ekg), gamma (PDB entry 7ekc), delta (PDB entry 7wbq), kappa (PDB entry 7tez), omicron BA.1 (PDB entry 7u0n), omicron BA.2 (PDB entry 7xo9), omicron BA.3 (PDB entry 7xb1) and omicron BA.4/5 (PDB entry 7zxu).

variants, in contrary to alpha, beta and kappa variants in which the Asn501 mutated to a tyrosine residue (i.e., N501Y). Omicron variants were the first to swap changes vastly on the RBD domain with the most important being the Q493R in BA.1/2/3, G496S only in BA.1, N501Y and Y505H.

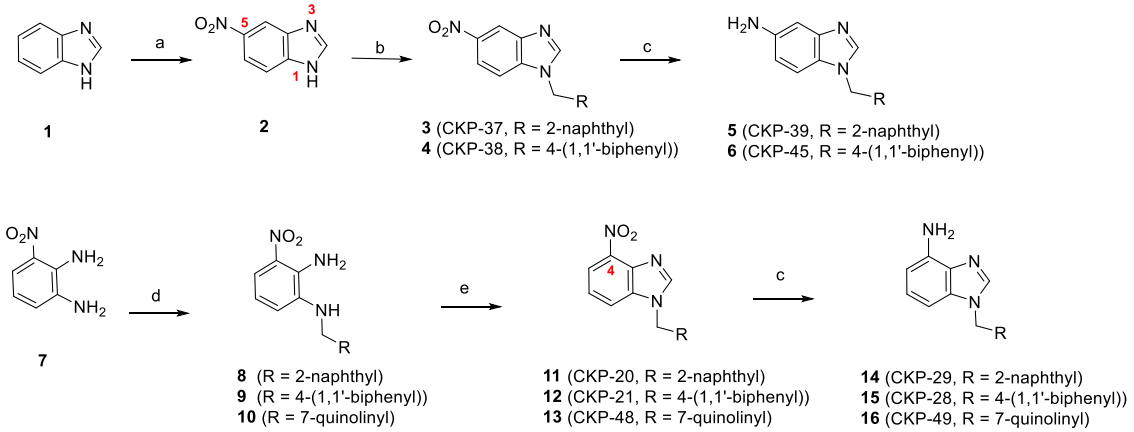
Uniformly, in all docking experiments using the different variants **CKP-22** and its synthetic derivatives showed no difference in their binding mode since the mutated amino acids maintained the same or similar chemical nature.

Chemistry. The first series of **CKP22** derivatives was obtained through variation of the *N*-aryl group and the substituents on the phenyl moiety of the benzimidazole scaffold. The synthesis of N^1-substituted benzimidazoles bearing a nitro-substituent at the C-5 position was realized using 1*H*-benzo[*d*]imidazole (**1**) as starting material. Initially, **1** was treated with a mixture of HNO_3/H_2SO_4 to give 5-nitro-1*H*-benzo[*d*]imidazole (**2**) quantitatively (Scheme 1). The latter was subjected to an alkylation reaction using the commercially available 2-(bromomethyl)naphthalene or 4-biphenylmethyl bromide as alkylating agents leading to compounds **3** or **4**, respectively. However, as anticipated, in both cases almost equimolar mixtures of regioisomers were obtained. As exemplified by compound **4**, the ^{1}H NMR of the crude mixture showed two *N*-C$\underline{H}_2$-biphenyl signals at 5.55 and 5.50 ppm. The regioisomers of **3** and **4** were isolated by FCC and the desired configuration was established based on the observation that for benzimidazoles bearing electron-withdrawing substituents at the C-5 position, the N^1-C$\underline{H}_2$-alkyl and H-2 signals resonate more upfield (lower δ) compared to the corresponding N^3-C$\underline{H}_2$-alkyl signals.[85] Finally, compounds **5** and **6** were obtained in high yields upon $SnCl_2$-mediated reduction of nitro-compounds **3** and **4**, respectively.

For the synthesis of N^1-substituted benzimidazoles substituted at C-4 a different synthetic strategy was implemented. The commercially available 3-nitrobenzene-1,2-diamine (**7**) was selectively alkylated with either 2-(bromomethyl)-naphthalene or 4-biphenylmethyl bromide or 7-(bromomethyl)quinoline hydrobromide salt to give rise to compounds **8**−**10**, respectively (Scheme 1). For the formation of the benzimidazole core, trimethyl orthoformate was used as the CH source under mild conditions to give nitro-compounds **11**−**13**, respectively in excellent yields. Then, compounds **11**−**13** were converted to their corresponding amino-congeners **14**−**16**, upon treatment with $SnCl_2{\cdot}2H_2O$.

The targeted N^1-alkylated benzimidazole derivatives **28**−**30** bearing a methoxy group at the C-4 position, were synthesized using the commercially available 3-methoxy-2-nitrobenzoic acid (**17**) as starting material (Scheme 2). Benzoic acid **17** was subjected to a DPPA-mediated Curtius rearrangement in refluxing *tert*-butanol to afford *N*-Boc aniline **18**.[86] The latter was then alkylated with either 2-(bromomethyl)naphthalene or 4-biphenylmethyl bromide or 7-(bromomethyl)quinoline hydrobromide salt in the presence of NaH in DMF to give compounds **19**−**21**, respectively, which in turn, were subjected to TFA-mediated removal of the Boc-protection to afford **22**−**24**. The nitro-compounds **22**−**24** were then converted to the corresponding anilines **25**−**27** with anhydrous $SnCl_2$ in MeOH at 50 °C in the presence of conc. HCl. Subsequently, anilines **25**−**27** upon treatment with trimethyl orthoformate in toluene at 80 °C using catalytic amount of PTSA gave the

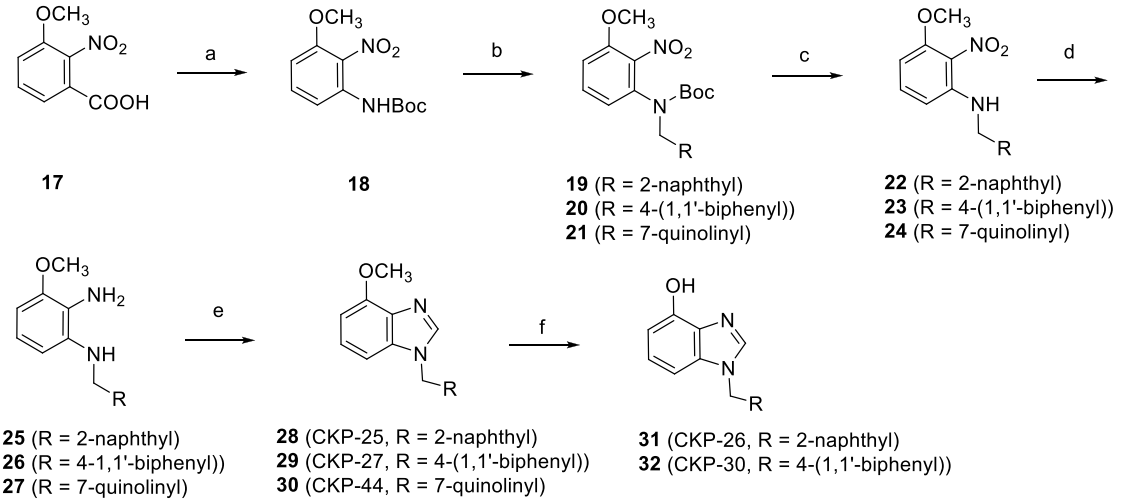
Scheme 1. Synthesis of 4-Nitro- and 4-Amino-N^1-Substituted Benzimidazoles[a]



[a]*Reagents and conditions*: (a) HNO_3/H_2SO_4, 0 °C, 3 h; (b) 2-(bromomethyl)naphthalene (for **3**) or 4-biphenylmethyl bromide (for **4**), K_2CO_3, DMF, 90 °C, 2 h; (c) $SnCl_2{\cdot}2H_2O$, conc. HCl, gl. AcOH, rt, 1 h; (d) 2-(bromomethyl)naphthalene (for **8**) or 4-biphenylmethyl bromide (for **9**) or 7-(bromomethyl)quinoline hydrobromide salt (for **10**), K_2CO_3, DMF, rt, 3−18 h; (e) trimethyl orthoformate, cat. PTSA, PhMe, 80 °C, 2−18 h or trimethyl orthoformate, $BF_3{\cdot}Et_2O$, DCM, rt, 5 h.

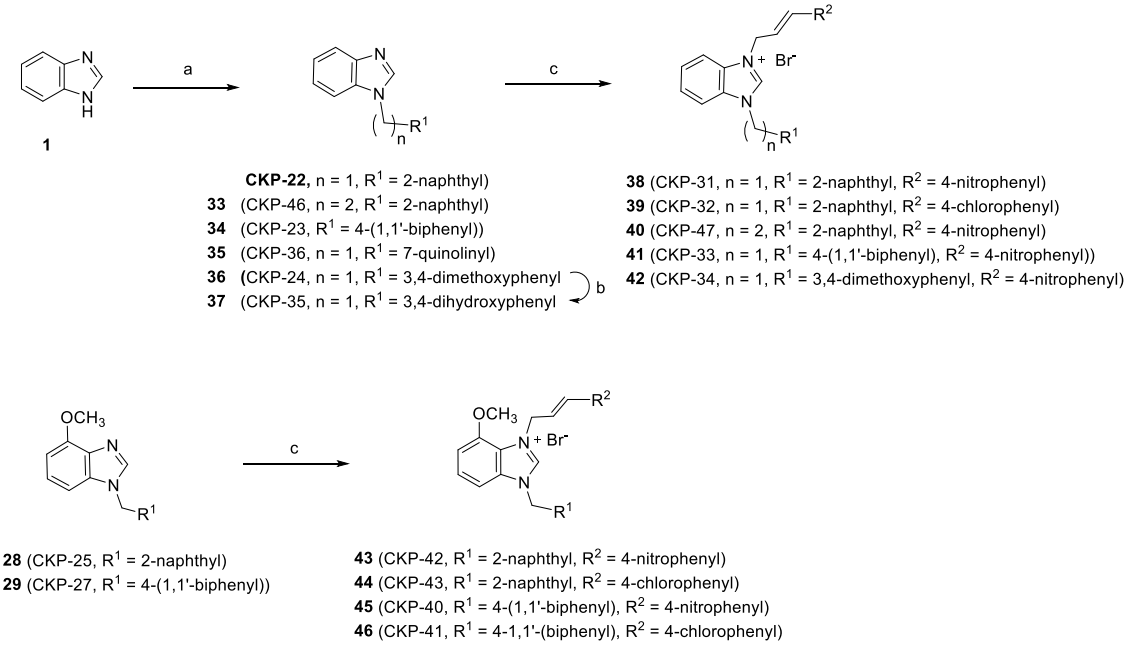
https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Scheme 2. Synthesis of 4-Methoxy and 4-Hydroxy-N^1-Substituted Benzimidazoles[a]



[a]*Reagents and conditions*: (a) DPPA/TEA, *t*-BuOH, 100 °C, 4 h; (b) 2-(bromomethyl)naphthalene (for **19**) or 4-biphenylmethyl bromide (for **20**) or 7-(bromomethyl)quinoline hydrobromide salt (for **21**), NaH, DMF, rt, 3 h; (c) TFA, DCM, 0 °C, 1 h; (d) $SnCl_2$, conc. HCl, MeOH, 50 °C, 18 h; (e) trimethyl orthoformate, cat. PTSA, PhMe, 80 °C, 2-18 h; (f) $BF_3{\cdot}S(CH_3)_2$, DCM, rt, 24−36 h.

Scheme 3. Synthesis of Benzimidazolium Salts[a]



[a]*Reagents and conditions*: (a) 2-(bromomethyl)naphthalene (for **CKP22**) or 2-(2-bromoethyl)naphthalene (for **33**) or 4-biphenylmethyl bromide (for **34**) or 7-(bromomethyl)quinoline hydrobromide salt (for **35**), 3,4-dimethoxybenzyl methanesulfonate (for **36**), NaH, THF, rt, 1−3 h; (b) $BF_3{\cdot}S(CH_3)_2$, DCM, rt, 24 h; (c) (*E*)-4-nitro-cinnamyl bromide (for **38**, **40-43**, **45**) or (*E*)-4-chloro-cinnamyl bromide (for **39**, **44**, **46**), 1,4-Dioxane, 100 °C, 24 h.

substituted benzimidazoles **28**−**30**, respectively Finally, compounds **28** and **29**, were treated with $BF_3{\cdot}S(CH_3)_2$ in DCM to give rise to the corresponding phenols **31** and **32**.[87]

Aiming to expand our compound library for SAR purposes, the naphthyl group of the hit compound was replaced by other aryl substituents (compounds **33**−**37**). Thus, the impact of the nature of the aryl substituent on the activity could be assessed. Hit compound **CKP-22** and derivatives **33**−**36** were easily obtained upon reaction of 1*H*-benzo[*d*]imidazole (**1**) with the appropriate bromides or mesylates in the presence of NaH in THF. Furthermore, compound **36** was treated with $BF_3{\cdot}S(CH_3)_2$ in DCM to give **37** bearing a catechol group (Scheme 3). Subsequently, **CKP-22** and **33**−**36** were further reacted with either (*E*)-4-nitro-cinnamyl bromide or (*E*)-4-chloro-cinnamyl bromide in refluxing dioxane to afford the quaternized benzimidazoles **38**−**42**.[88] Following the same quaternization procedure, compounds **28** and **29** bearing a methoxy-substituent at the C-4 position, gave rise to compounds **43**−**46** (Scheme 3).

Validation of Computationally Predicted Spike/ACE2 Inhibitors. The activity of the qualified S(RBD)-ACE2 (19−615) inhibitors was evaluated in cell-based assays. First,

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

expression clones of PA-mCitrine-Spike RBD and cmyc-NL-ACE2 (19−615) were generated, and their expression efficiency was assessed by immunoblotting in extracts of HEK293T cells transiently transfected with PA-mCitrine-Spike RBD or cmyc-NL-ACE2 (19−615) cDNAs. Bands corresponding to the recombinant PA-mCitrine-Spike RBD and cmyc-NL-ACE2(19−615) proteins were detected at 65 kDa and 84 kDa, respectively suggesting that expression plasmids are suitable for the study of the interaction between Spike RBD and ACE2 (19−615) (Figure 7A). Then, a cell-based LuTHy

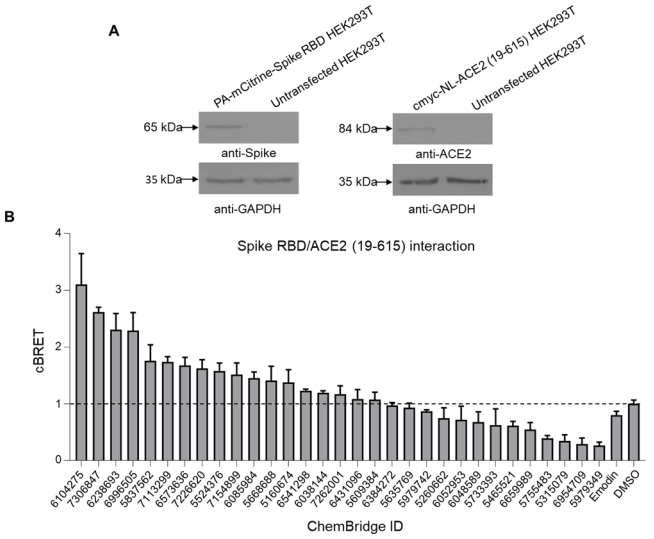


Figure 7. In vitro screening of computationally predicted compounds for Spike/ACE2 inhibitors. (A) Immunoblot analysis of HEK293T cells transiently transfected with expression clones encoding the PA-mCitrine-Spike RBD and cmyc-NL-ACE2 (19−615) proteins generated for the quantification of the Spike RBD/ACE2 (19−615) interaction. Immunoblots were detected with an anti-Spike antibody or an anti-ACE2 antibody. GAPDH was used as a loading control. (B) Effects of computationally predicted compounds on Spike RBD/ACE2 (19−615) interaction in LuTHy assay. The interacting proteins PA-mCitrine-Spike RBD and cmyc-NL-ACE2 (19−615) were co-produced in HEK293T cells in the presence of compounds (100 μM). The cBRET signal corresponding to Spike RBD/ACE2 (19−615) interaction was quantified 48 h post-treatment. Eleven compounds significantly reduced the interaction as compared to the cells treated with DMSO or the positive control Emodin (100 μM).

assay was set up for the quantification of Spike RBD/ACE2 (aa19−615) PPI. HEK293T cells were transiently co-transfected with PA-mCitrine-Spike RBD and cmyc-NL-ACE2 (19−615) plasmids. Control experiments were performed using combinations of PA-mCitrine-Spike RBD/NL or PA-mCitrine/cmyc-NL-ACE2 (19−615). The co-production of recombinant proteins resulted in significantly higher BRET ratios compared to control interactions (Figure S1A), suggesting that this method can be applied for the quantification of Spike RBD/ACE2 (19−615) interaction.

Then, the inhibitory effect of the 31 computationally selected hit molecules on Spike RBD/ACE2 (19−615) interaction was assessed using the previously established LuTHy assay. Compounds were tested in a non-cytotoxic concentration (100 μM) as indicated by MTT assay (Figure S1B) and those reducing Spike RBD/ACE2 (19−615) interaction were considered as positive hits. In total, 11 compounds [5979742, 5260662, 6052953, 6048589, 5733393, 5465521, 6659989, 5755483, 5315079, 6954709, 5979349 (**CKP-22**)] suppressed the interaction in vitro compared to treatment with the solvent DMSO (Figure 7B). Interestingly, the positive control Emodin[89−91] indeed suppressed Spike/ACE2 interaction but to a lesser extent (~27%) compared to most positive hits and only in a very high non-cytotoxic concentration (100 μM) (Figures S1C and S2).

Subsequently these compounds were further assessed for their inhibitory effect in dose-dependent LuTHy assays. These experiments indicated that when tested in a non-cytotoxic concentration, 6 of the 11 selected compounds [compounds 5260662, 6052953, 5465521, 6659989, 5315079, 5979349 (**CKP-22**)] suppressed Spike RBD/ACE2 (19−615) interaction in a concentration-dependent manner, confirming their activity (Figure 8A for **CKP-22** and Figure S2).

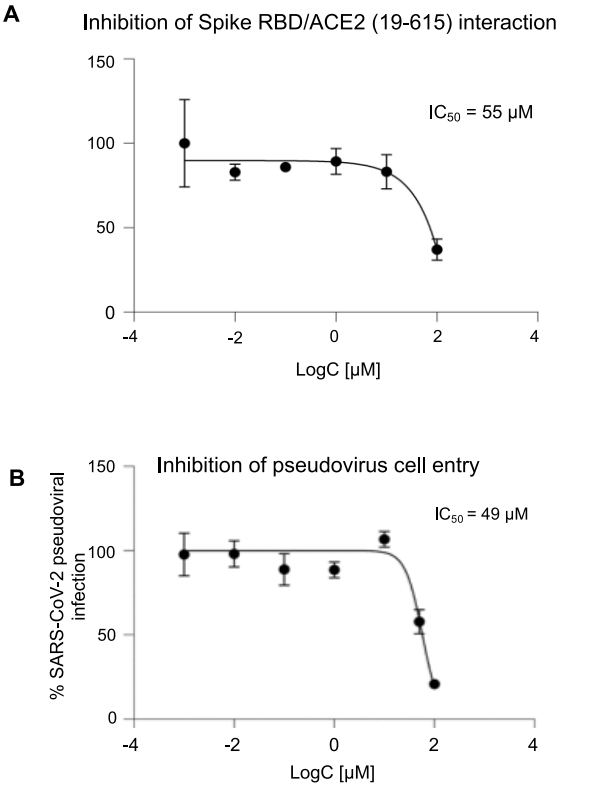


Figure 8. Concentration−response curves obtained for the inhibitory effect of compound 5979349 (**CKP-22**) in cell-based assays. Inhibition of (A) Spike RDB/ACE2 (aa19−615) interaction in LuTHy assay and (B) SARS-CoV-2 pseudovirus entry. Data are mean ± SD and were fitted with standard sigmoid curves for IC_{50} determination.

The specificity of these compounds was further corroborated in LuTHy assays for a previously established interaction between ataxin-1 and MED-15 proteins.[92] Interestingly, none of the compounds suppressed ATXN1-MED15 interaction (Figure S3), confirming their specific inhibitory effects on the Spike RBD/ACE2 (19−615) interaction.

Next, we tested whether the compounds suppressing the Spike RBD/ACE2 (19−615) interaction, would prevent infection with SARS-CoV-2. To this end, the six hits were evaluated for their ability to block the entry of a green fluorescent SARS-CoV-2 pseudovirus into human SH-SY5Y cells transduced with red fluorescent ACE2. The blockade of pseudovirus entry would result in a concentration-dependent

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

reduction of the green fluorescence in infected cells. Compared to LuTHy assays, pseudoviral assays depend on the number of viable cells. Initially, the toxicity of the six hits on SH-SY5Y cells was assessed at a concentration range of 100−0.001 μM (Figure S4A). Subsequently, the six S(RBD)-ACE2 inhibitors were tested at non-cytotoxic concentrations for inhibition of pseudovirus cell entry. In order to remove any false-positives, results were normalized to the number of viable cells producing red ACE2 and are prone to infection with the green SARS-COV-2 pseudovirus. Two out of the 6 tested compounds, namely 1-(2-naphthylmethyl)-1*H*-benzimidazole [5979349 (**CKP-22**)] and 3-(4-biphenylyloxy)-7-hydroxy-4H-chromen-4-one (6659989) protected human cells from infection in a dose-dependent manner possessing IC_{50} values of 49 ± 2.5 and 3.2 ± 0.2 μM, respectively (Figure 8B for **CKP-22** and Figure S4B for 6659989).

Finally, we tested whether the two inhibitors of the SARS-CoV-2 pseudoviral entry protected Vero cells from infection with SARS-CoV-2. Only 1-(2-naphthylmethyl)-1*H*-benzimidazole [5979349 (**CKP-22**)] at a non-cytotoxic concentration (Figure S5A), significantly and reproducibly reduced the cytopathic effect (CPE) of the virus by approximately 90% compared to control cells treated only with the solvent (Figure 9A,B). These results were further verified by a qPCR for the

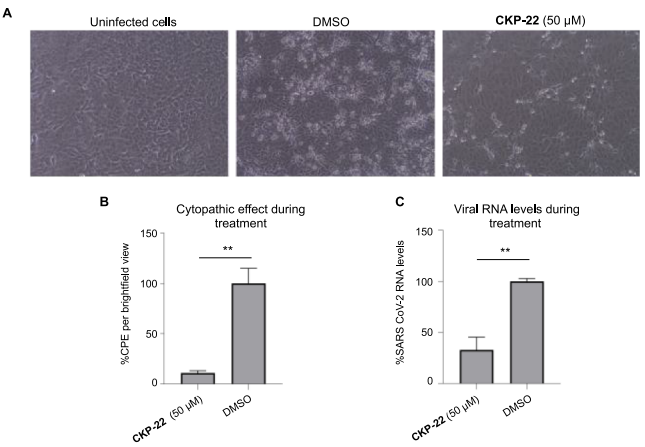


Figure 9. Antiviral activity of 5979349 (**CKP-22**) against SARS-CoV-2 in Vero E6 cells. (A) Brightfield microscopy of Vero E6 cells. Vero E6 cells were infected with SARS-CoV-2 in the absence and presence of compound **CKP-22** (50 μM). Two days post-infection, cells were monitored by microscopy. (B) CPE activity for **CKP-22** in cells. Cells infected with SARS-CoV-2 were cultured in the absence or presence of **CKP-22** (50 μM) for 48 h. After incubation, CPE was recorded using an inverted microscope with phase contrast. **CKP-22** decreases the cytopathic effect of SARS-CoV-2 compared to cells treated only with DMSO. (C) Quantification of SARS-CoV-2 viral load in the supernatant of Vero E6 cells. Viral RNA measured by RT-qPCR in cell supernatant infected with SARS-CoV-2 (MOI of 0.01) and treated with **CKP-22** or DMSO (positive control) for 48 h. Treatment with **CKP-22** leads to significant reduction in viral load compared to cells treated only with DMSO.

quantification of the viral load in the cell culture supernatant. This analysis showed that SARS-CoV-2 levels were reduced by 68% (Figure 9C) indicating that **CKP-22** efficiently blocks Spike-ACE2 interaction and strongly prevents infection and spread of SARS-CoV-2 in vitro.

Biological Evaluation and Structure Activity Relations of Synthesized CKP-22 Derivatives. The 29 synthesized **CKP-22** derivatives were initially tested for their cytotoxicity against HEK293T cells (Figure 10A). Subsequently their effect on Spike RBD/ACE2 (19−615) interaction was assessed using LuTHy assay (Figure 10B and Figure 11A). Each analogue was tested at the highest non-cytotoxic concentration against HEK293T cells. **CKP-22** was included for comparison. Ten derivatives (**CKP-20**, **CKP-25**, **CKP-27**, **CKP-32**, **CKP-40**, **CKP-42**, **CKP-46**, **CKP-47**, **CKP-48** and **CKP-49**) suppressed the Spike RBD/ACE2 (19−615) interaction. In particular, **CKP-20** at 50 μM by 51%, **CKP-25** at 25 μM by 34%, **CKP-27** at 25 μM by 72%, **CKP-32** at 10 μM by 70%, **CKP-40** at 25 μM by 14%, **CKP-42** at 10 μM by 74%, **CKP-46** at 1 μM by 20%, **CKP-47** at 10 μM by 34%, **CKP-48** at 50 μM by 36% and **CKP-49** at 50 μM by 30% in comparison to **CKP-22** at 50 μM by 81%.

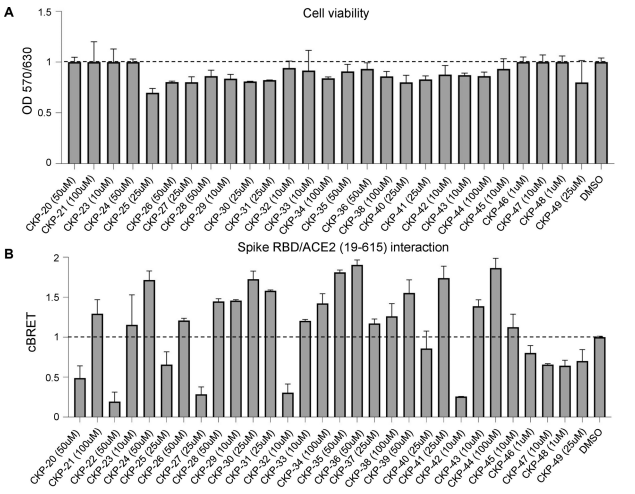


Figure 10. Cytotoxicity and effects of the synthesized compounds on the Spike RBD/ACE2 (aa19−615) interaction. (A) Viability of HEK293T cells treated with the new compounds. Cell viability was measured 48 h later using a MTT assay. Cells treated with DMSO served as control. (B) Effect of the new compounds on Spike RBD/ACE2 (aa19−615) interaction using a cell-based quantification assay. Interaction of PA-mCitrine-Spike RBD and cmyc-NL-ACE2 (aa19−615) proteins was quantified in HEK293T cells in the presence of the new compounds used at the highest non-cytotoxic concentration. The cBRET signal corresponding to Spike RBD/ACE2 (aa19−615) interaction was quantified 48 h post treatment. Data are presented as a mean of triplicates ± SD.

Among the 4-nitro-N^1-substituted benzimidazoles (compounds **11**−**13**) compounds **11** (**CKP-20**) and **13** (**CKP-48**) substituted by a naphthyl or a 7-quinolinyl group showed activity. Conversely, among their 4-amino congeners the N^1-7-quinolinylmethyl-substituted **16** (**CKP-49**) suppressed the Spike RBD/ACE2 (19−615) interaction by 30% compared to the control (DMSO). Gratifyingly the 4-methoxy-substituted benzimidazole derivatives **28** (**CKP-25**) and **29** (**CKP-27**) suppressed Spike RBD/ACE2 (19−615) interaction by 34 and 72%, respectively compared to the control (DMSO). However, none of the 4-hydroxy congeners of **28** and **29**, compounds **31** (**CKP-26**) and **32** (**CKP-30**) maintained inhibitory activity. Interestingly, in this series the 4-methoxy congener of the active amino compound **16** (**CKP-49**) [derivative **30** (**CKP-44**)] did not exhibit any effect. This points out to the fact that both the C-4 and N^1 substituents are key for activity. Modifying the N^1−naphthyl substituent in the hit compound **CKP-22**, by other aryl moieties, namely

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

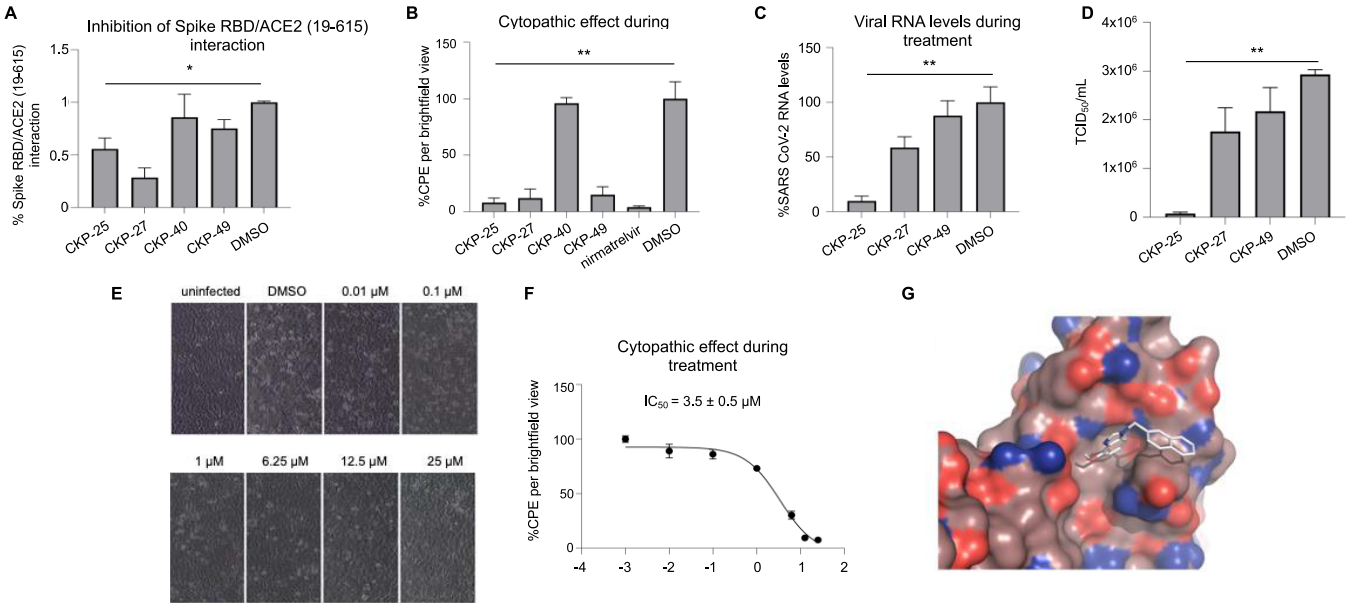


Figure 11. Derivatives of 1-(2-naphthylmethyl)-1H-benzimidazole (**CKP-22**) protect from infection with SARS-CoV-2. (A) Quantification (cBRET signal) of Spike-RBD/ACE2 (19−615) interaction in the presence of **CKP-22** derivatives (**CKP-25**, **CKP-27**, **CKP-40** and **CKP-49**, 25 μM final concentration) or the solvent DMSO. (B) SARS-CoV-2 CPE in Vero E6 cells treated with the new compounds (12.5−25 μM final concentration) compared to the solvent DMSO. The commercial SARS-CoV-2 inhibitor nirmatrelvir was used as a positive control. (C) SARS-CoV-2 load in the supernatant of infected Vero E6 cells treated with derivative compounds or the solvent DMSO. (D) $TCID_{50}$ of SARS-CoV-2 after treatment with compounds **CKP-25**, **CKP-27**, **CKP-49** or the solvent DMSO. (E) Brightfield microscopy of SARS-CoV-2 infected Vero E6 cells treated with various varying concentrations (ranging from 0.01 to 25 μM) of compound **CKP-25**. (F) Concentration-dependent effect of compound **CKP-25** inhibition in SARS-CoV-2 CPE (IC_{50} = 3.5 ± 0.5 μM). Treatment with DMSO was arbitrarily set to 100%. Data in bar graphs are shown as mean ± SD (*p-value < 0.05, **p-value < 0.0.01). (G) Docking solution of compound **CKP-25** shown as white colored sticks residing in region A of the RBD domain.

biphenyl, 7-quinolinyl, 3,4-dimethoxyphenyl or 3,4-dihydroxyphenyl (compounds **34**−**37**) resulted in loss of activity. In contrast when the methylene spacer between the naphthyl group and the N^1 nitrogen in **CKP-22** was extended by one carbon, compound **33** (**CKP-46**) activity was retained. Finally, we proceeded in quaternization of N^3 in the hit compound **CKP-22** and its ethylnaphthyl, biphenyl and 3,4-dimethoxyphenyl congeners compounds **33** (**CKP-46**), **34** (**CKP-23**) and **36** (**CKP-24**), respectively. Derivatives **39** (**CKP-32**) and **40** (**CKP-47**) resulted in 70 and 35% reduction, respectively. In the case of the N^1-naphthylmethyl-substituted salt the N^3-4-chlorophenylethylene group is preferred (**CKP-32**), while in the case of the N^1-naphthylethyl-substituted salt the N^3-4-nitrophenylethylene group (**CKP-47**) results in inhibitory activity. Furthermore, among the compounds resulting from the quaternization of the active derivatives **CKP-25** and **CKP-27**, with 4-nitrophenylethylene or 4-chlorophenyethylene groups, derivatives **43**−**45**, the 4-nitrophenylethylene derivatives were active, namely **43** (**CKP-42**) and **45** (**CKP-40**).

Subsequently, we set out to evaluate the antiviral activity of the active compounds against SARS-CoV-2. As a first step their toxicity against Vero E6 cells was assessed (Figure S5B). Derivatives **CKP-32** and **CKP-47** were toxic thus, they were not further tested. Out of the remaining 8 compounds only **CKP-25**, **CKP-27** and **CKP-49** consistently reduced the cytopathic effect of the virus with the most potent being **CKP-25.** This effect was comparable to the effect of the commercial inhibitor of SARS-CoV-2 nirmatrelvir (Paxlovid, Pfizer) (Figure 11B). As expected, derivative **CKP40**, which was used as a negative control and did not significantly suppress the Spike RBD/ACE2(19−615) interaction, was also not protective against SARS-CoV-2. Furthermore, compounds **CKP-25, CKP-27** and **CKP-49** also reduced the viral load in the cell culture supernatant (Figure 11C). In particular, calculation of the half maximal tissue culture infectious dose ($TCID_{50}$) indicated that compound **CKP-25** reduced by approximately 90% the viral titter (Figure 11D), while, the IC_{50} was found equal to 3.5 ± 0.5 μM (Figure 11E,F). Collectively, these results suggest that compound **CKP-25** protects mammalian cells from infection with SARS-CoV-2 and is a lead compound against COVID-19. Based on our docking studies compounds **CKP-22** and **CKP-25** reside in region A of the Spike RBD domain (Figures 3C and 11G, respectively). Moreover, the methoxy substituent in **CKP-25** serves as the HB acceptor resulting in an additional interaction, which turn positively influences the activity. MD simulations were performed to study the interaction behavior of **CKP-25** in parallel with any protein conformational changes that may occur. The data collected during 100 ns of simulation time (Charts S1−S4) showcase that compound **CKP-25** resides for 77.5% of the time with median RMSD of 0.3 nm in a similar manner as predicted in Figure 11G (region A of the S(RBD) domain, while for the remaining 22.5% (RMSD > 0.5 nm) it flips mainly with the naphthyl substituent toward region B. The pi-pi stacking of Tyr505 indeed plays an important role on maintaining this binding mode for **CKP-25** since beside this interaction the compound forms only one HB for the 99.9% of the simulation. Moreover, for the ∼2/3 of the 77.5% that have an RMSD above 0.2 nm we noticed a configuration that showed loss of planarity within either of the naphthyl substituent with Tyr505, the Tyr505-Arg403 or both and modifying the RBD domain. In all cases the methoxy

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

substituent served as the HB acceptor between the ligand and the protein with adjacent residues like Tyr496 or Arg403 during our simulation (Chart S4).

In Vitro ADME-Tox Studies. *Calculated Water Solubility.* A predicted solubility value (log S_w) of compound **CKP-25** was available to us since the initial filtering over FAF4Drugs server.[70] The method used for calculation by the server is a structure only method based on Delaney et al. ESOL method[93] resulting to a **−4.48** log S_w. Although, when the melting point (M pt) is available the most efficient method is using the "General Solubility Equation" (see eq 1).[94]

Equation 1 General Solubility Equation

$$\log S = -\log P - 0.01 \times (\text{M pt} - 25) + 0.5 \quad (1)$$

Therefore, the GSE based solubility for **CKP-25** according to eq 1 is log S = **−4.25** − 0.01(163.5 − 25) + 0.5 = **−5.13.** It is should be noted that the lipophilicity (log P) used in the GSE above is the calculated value using *X*log *P*3[95]) and not an experimental partition coefficient over octanol/water.

Microbial Mutagenicity Assay: Ames Test. To evaluate the mutagenic potential of **CKP-25**, a bacterial reverse mutation assay was performed using the *E. coli* WP2 uvrA strain. The genetic background of the bacterial strains was verified prior to testing to ensure the accuracy and reliability of the results. Positive control experiments were conducted to confirm the sensitivity of the test system and the efficacy of the S9 metabolic activation mix. Methylmethanesulfonate (without metabolic activation) and 2-aminoanthracene (with metabolic activation) were used as positive controls, both significantly increasing the number of revertant colonies grown in selective agar plates. The results, summarized in Table 1, demonstrate that **CKP-25**, tested at concentrations ranging from 12.5 to 200 μg/plate, did not induce a significant increase in the number of revertant colonies of the *E. coli* WP2 uvrA strain, when compared to the negative control (DMSO). This lack of increase was observed in the presence and absence of the S9-mix and was consistent across all tested concentrations. Thus, **CKP-25** does not exhibit mutagenic activity in AMES testing.

Table 1. Assessment of the Mutagenic Activity of Compound CKP-25

	−S9		+S9	
compound	mean revertants	STDEV(∓)	mean revertants	STDEV (∓)
	CKP-25 (μg/plate)			
200	5.5	0.7	16	2.8
100	11.5	2.1	13	0
50	13	2.8	15	5.6
25	12.5	2.1	19	1.4
2.5	7.5	0.7	13	2.8
DMSO	18.5	0.7	16.5	0.7
methyl methanesulfonate	>300			
2-aminoanthracene			>300	

Metabolic Stability (Residual % of Time Zero). **CKP-25** was studied for its metabolic stability. In particular, the metabolic stability of a compound can be interpreted using several approaches, as it can be ranked in terms of its intrinsic clearance (CL_{int}) and in vitro $t_{1/2}$ values or based on parent structure loss during metabolic reactions[96,97] CL_{in} describes the maximum activity of liver (microsomal proteins or hepatocytes) toward a compound not influenced by other physiological determinants, such as hepatic blood flow and drug binding within the blood matrix and in vitro $t_{1/2}$ expresses the time for 50% disappearance of the compound. Parent structure loss was our option for metabolic stability data interpretation of **CKP-25**. Parent structure loss is classified as very slow (<5%), slow (5−19%), moderate (20−50%), fast (50−80%) or very fast (>80%). Such categories have been defined according to set criteria, namely, high metabolism ($t_{1/2}$ value of <30 min), moderate metabolism (30 min < $t_{1/2}$ value of <60 min) and low metabolism ($t_{1/2}$ value of >60 min).[98−100]

CKP-25 shows slow substrate depletion, suggesting a low intrinsic clearance classification band (Table 2). For humans, a CL_{int} <8.6 μL/min/mg protein defines a low intrinsic clearance classification band, whereas a CL_{int} >47.0 μL/min/mg protein defines a high intrinsic clearance classification band. Low clearance compounds are characterized by reduced doses, enhanced exposure and prolonged half-life, being suitable for once-daily dosing.

Isozyme-Specific CYP450 Metabolism. Human CYP450 enzymes are crucial for xenobiotic biodegradation, metabolism, and toxicity as well as xenobiotic-host and/or xenobiotic-xenobiotic interactions.[101]

Upon linear velocity conditions in vitro, the depletion rate of test-compounds may be extrapolated to (a) in vivo hepatic clearance, (b) extraction ratio, and (c) the effect of hepatic first-pass metabolism to total oral bioavailability. Biodegradation, metabolic, and toxicity liabilities can be identified early on and thus, inform structure−activity relationships.[98−100]

For this, the activity of CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 was assessed after the administration of **CKP-25** at 1 μM to determine (a) the oxidative (CYP-mediated) metabolic stability profile in question and (b) the enzyme metabolizing isoforms responsible (the test system consists of recombinant human CYP450 and CYP450 reductase; cytochrome b5 may also be present). Herein, no concentration-dependent effects were reported. CYP450 substrates can alter enzyme activity by blocking the enzyme active site, changing enzyme conformation, disrupting enzyme structure and/or functioning.[102,103] **CKP-25** did not decrease the enzyme (catalytic) activity of the CYP450 system tested herein. CYP450 enzyme inhibition may lead to unexpectedly high exposure of coadministered xenobiotics and hence, increase the risk for adverse effects. No product inhibition or mechanism-based inactivation of the CYP450

Table 2. Metabolic Stability and CYP450-Mediated Metabolism Data of Compound CKP-25

metabolic stability in human liver microsomes (residual % of time zero)						
94						
CYP450s (% inhibition)						
CYP1A2	CYP2A6	CYP2B6	CYP2C9	CYP2C19	CYP2D6	CYP3A4
0	6	0	11	20	1	3

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

isoenzymes in question was obtained. Test-compounds with poor solubility can show artificially low CYP450 inhibition and thus, chemical entities with potential drug−drug interaction toxicities may be overlooked. No solubility issues were observed. Overall, **CKP-25** was a weak inhibitor of the CYP450 system (see Table 2 and Figure 12).

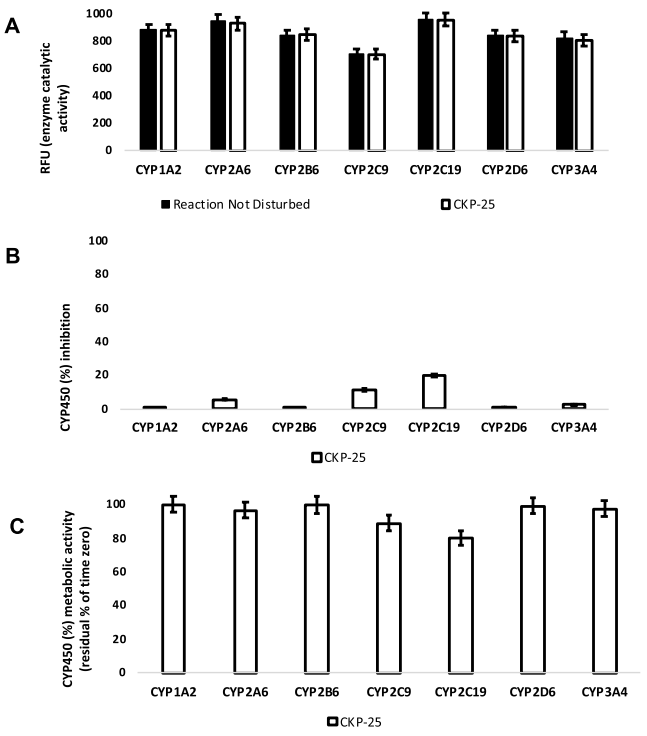


Figure 12. (A) Enzyme (catalytic) activity of the CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 isoenzymes upon the administration of **CKP-25** at 1 μM (60 min). RFU, relative fluorescence units. Reaction Not Disturbed, reaction without test-compound. (B) CYP450 (%) inhibition of CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 isoenzymes upon the administration of **CKP-25** at 1 μM (60 min). (C) CYP450 (%) metabolic activity of CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 isoenzymes upon the administration of **CKP-25** at 1 μM (60 min).

CONCLUSIONS

In conclusion through AI enabled VS and subsequent in vitro screening we were able to identify benzimidazole derivative **CKP-22** as a hit compound protecting Vero E6 cells from infection with SARS-CoV-2 mediated through the inhibition of S(RBD)-ACE2 interaction. Hit to lead optimization studies revealed that 4-methoxy-1-(naphthalen-2-ylmethyl)-benzimidazole (**CKP-25**) suppressed the Spike RBD/ACE2 (19−615) interaction, reduced the cytopathic effect of SARS-CoV-2 in Vero E6 cells (IC_{50} = 3.5 ± 0.5 μM) and reduced the viral load in cell cultures supernatants. In addition, the isozyme-specific CYP450 study revealed that **CKP-25** was a weak inhibitor of the CYP450 system, while it did not show biodegradation or liver metabolism (≥94% residual of time zero at t = 60 min) or implied safety issues. Therefore, **CKP-25** can be considered as a lead compound against COVID-19 infection.

EXPERIMENTAL SECTION

Pre-filtering Profiling. ChemBridge corporation precompiled macrocycle, core and dedicated SARS-CoV-2 libraries in SDF formatted files were downloaded from their Hit2Lead site. Libraries were cleaned from any salts' counterions being present using OpenBabel v2.4.1[62] and introduced as obtained to FAF-Drugs4 server[70] for filtering applying drug-like properties,[65,104−108] promiscuity and protein−protein interaction (PPI)[71] options based on their ADME-Tox descriptors (i.e., XLOGP3,[109] PKs, bioavailability and Lilly Medchem relaxed rules.[109] Accepted molecules meeting all requirements were compiled in SDF formatted files and used to perform the virtual screening (VS).

AI-Based Virtual Screening. Following the Prefiltering step, accepted molecules in SDF format were given as input to Virtual Screening against the crystal structure of the spike protein S1 subunit bound to the extracellular domain of the ACE2 receptor (PDB ID: 6M0J). The AI-enabled Virtual Screening pipeline combines the docking result of Smina[61] with an AI-based rescoring function which operates upon Smina's docking results and is able to distinguish more efficiently potential binding complexes.[60] Docking was constrained on the interaction site of the S-protein with the ACE2, while 50 docked poses were generated for each compound conformation. The resulted protein-compound docked complexes were rescored with our custom AI-based rescoring function, a properly designed 3D convolutional neural network (3D-CNN), which aims to automatically capture the underlying mechanism responsible for the protein−ligand binding. A fusion approach is proposed to achieve the optimal performance, which combines the Smina predicted affinities with the new AI-score. The already qualified molecules were further grouped into clusters of structural similarity using a 3D shape descriptor.[60]

Post-filtering. Decision-making for selection of compounds in VS data sets were based on three methods; molecular modeling, molecular properties and protein−ligand interactions formed.[110,111]

Molecular Modeling. Hit validation of the VS results and modeling of the benzimidazole analogues **CKP-22** compound library was performed with the program PyRx[112] using Vina[113,114] functionality, running in Genetic Algorithm. Crystal structures of the SARS-CoV-2 virus spike protein variants; wild-type,[10] alpha,[78] beta,[78] gamma,[78] delta,[79] kappa,[80] and omicron BA1[81]/BA2[82]/BA3[84]/BA4[83] (PDB entries 6m0j, 7ekf, 7ekg, 7ekc, 7wbq, 7tez and 7u0n/7xo9/7xb1/7zxu, respectively) were obtained from the protein databank. All of them were cleaned from the respective ACE2 protein parts and prepared accordingly for docking after being aligned over the wild-type entry 6m0j[83] with the use of software PyMol molecular graphic system v1.5.0.3 (see Table S2). All compounds were designed using PerkinElmer ChemDraw v20.0 and underwent structure minimization in MM2 force field[115] found within PerkinElmer Chem3D v22.0.0.0, where they were eventually saved as *.mol2 extension files. Ligand preparation was done within PyRx[112] and used thereof for molecular modeling. Based on VS the resulting search space was centered in X: −32.1814, Y: 22.2131, Z: 2.9869 with respective volumes X: 22.5547, Y: 24.6395, Z: 21.6589, generating a grid space of 12,036.6 Å.[116] No residues modifications were made on the protein structures and the docking was performed without any constraints. While the

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

protein is treated as rigid and the solely parameter adjustment was observed on the "exhaustiveness", which value was set to 20 instead of the default 8.

MD Simulations. Vina derived S(RBD) domain docking complex with inhibitor **CKP-25** was used as the input for molecular dynamics simulation using GROMACS 2023.3[117] on a Hyper-V virtual machine running Ubuntu 24.04.1 over a typical Dell XPS desktop. A cubic box was created with 1 nm dimension, and TIP3P water molecules were added to solvate the system. The system was then neutralized by adding chloride ions which replaced equivalent solvent molecules. Parametrization of the protein−ligand complex was performed with the CHARMM36 force field July 2022 version[118] and energy minimized. Equilibration stage was performed for NVT/NPT 100 ps with hydrogen-bond constraints. Following the equilibration we performed the MD simulation for 100 ns at 300 K using modified Berendsen thermostat with 2 fs integration time and VDW cutoff distance at 1.2 nm. Molecular dynamics trajectory was analyzed using Grace 5.1.25[119] and VMD 1.9.4a53.[120,121]

Chemistry. *General.* Melting points were determined with a Shanghai Micro Melting Point apparatus and are uncorrected. ^{1}H NMR spectra were recorded on Varian spectrometers operating at 300 or 600 MHz and ^{13}C NMR spectra were recorded at 75 or 150 MHz, using $CDCl_3$, CD_3OD, DMSO-d_6 or acetone-d_6 as solvents. Chemical shifts are reported in δ units, parts per million (ppm) downfield from TMS. Electro-spray ionization (ESI) mass spectra were recorded on a LC-MSn Fleet, Thermo spectrometer using MeOH as solvent. HRMS spectra were recorded in the ESI mode, on UPLC-MSn Orbitrap Velos-Thermo. Flash column chromatography (FCC) was performed on Merck silica gel 60 (230−400 mesh) and TLC on Merck 60 F254 films (0.2 mm) precoated glass plates. Spots were visualized with UV light at 254 nm and PMA stain. All solvents were dried and/or purified according to standard procedures prior to use. All reagents employed in the present work were purchased from commercial suppliers and used without further purification. Reactions were run in flame-dried glassware under an atmosphere of argon. The purity of final compounds was determined by high-performance liquid chromatography (HPLC) (Thermo Scientific HPLC SPECTRASYSTEM with a Thermo Scientific SPECTRASYSTEM UV2000 detector and a EC 250/4.6 NUCLEOSIL 100−5 C18 HD column) using the following analytical method: 20% H_2O_0.1%TFA/80% CH_3CN, isocratic, flow rate 1 mL/min, injection volume 20 μL, UV detection at 254 nm; The purity of all final compounds was >97%.

General Procedure for the Alkylation of 5-Nitrobenzimidazole. To a suspension of 5-nitrobenzimidazole (0.16 g, 1.0 mmol) and K_2CO_3 (0.14 g, 1.0 mmol) in DMF (2 mL), the appropriate bromide (1.0 mmol) was added. The resulting mixture was heated at 90 °C for 2 h, it was left to attain RT, diluted with water and the aqueous phase was extracted twice with EtOAc. The combined organic layers were washed with brine, dried over Na_2SO_4 and evaporated to dryness to give the projected compounds as a mixture of regioisomers. The desired regioisomers were isolated after FCC purification.

*1-(Naphthalen-2-ylmethyl)-5-nitro-1H-benzo[d]imidazole (**3**) (**CKP-37**).* Compound **3** was prepared using 2-(bromomethyl)naphthalene (0.22 g, 1.0 mmol) according to the general procedure described above. White solid, 0.15 g (50% yield); R_f = 0.29 (Hexanes/EtOAc 1:1 v/v); ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.78 (d, J = 1.8 Hz, 1H), 8.36 (s, 1H), 8.19 (dd, J = 8.4 and 1.8 Hz, 1H), 7.88−7.82 (m, 2H), 7.80−7.78 (m, 1H), 7.68−7.65 (m, 1H), 7.54−7.50 (m, 2H), 7.41 (d, J = 8.4 Hz, 1H), 7.28 (dd, J = 8.4 and 1.8 Hz, 1H), 5.61 (s, 2H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 146.6, 144.3, 142.4, 137.8, 133.3 (two C), 131.4, 129.7, 128.0 (two C), 127.2, 127.1, 126.7, 124.5, 119.4, 117.0, 110.6, 50.0; ESI-HRMS (m/z): calcd. for $C_{18}H_{14}N_3O_2$ $[M + H]^+$ 304.1086; found 304.1075. HPLC analysis: t_R = 5.44 min, purity = 99.2%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-5-nitro-1H-benzo[d]-imidazole (**4**) (**CKP-38**).* Compound **4** was prepared using 4-biphenylmethyl bromide (0.25 g, 1.0 mmol) according to the general procedure described above. White solid, 0.17 g (52% yield); R_f = 0.27 (Hexanes/EtOAc 3:7 v/v); ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.77 (s, 1H), 8.31 (s, 1H), 8.23 (d, J = 8.4 Hz, 1H), 7.60 (d, J = 7.8 Hz, 2H), 7.55 (d, J = 7.8 Hz, 2H), 7.46−7.40 (m, 4H), 7.38−7.34 (m, 1H), 7.28 (d, J = 8.4 Hz, 1H), 5.49 (s, 2H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 146.5, 144.2, 142.1, 140.1, 137.9 (two C), 133.1, 129.1 (two C), 128.2 (two C), 128.0, 127.8 (two C), 127.2 (two C), 119.4, 117.2, 110.5, 49.4; ESI-HRMS (m/z): calcd. for $C_{20}H_{16}N_3O_2$ $[M + H]^+$ 330.1243; found 330.1233. HPLC analysis: t_R = 7.14 min, purity = 99.5%.

*General Procedure for N-Alkylation of 3-Nitro-1,2-phenyl-enediamine (**7**).* To a solution of 3-nitro-1,2-phenylenediamine (**7**) (0.15 g, 1 mmol) in DMF (5 mL), K_2CO_3 (0.28 g, 2 mmol) was added followed by the addition of the appropriate aryl bromide (1.3 mmol). The reaction mixture was stirred at RT for 6 h, diluted with water and extracted twice with EtOAc. The combined organic layers washed with water, dried over Na_2SO_4 and evaporated to dryness. The desired compounds were obtained in pure form after FCC purification.

*N^1-(Naphthalen-2-ylmethyl)-3-nitrobenzene-1,2-diamine (**8**).* Compound **8** was prepared using 2-(bromomethyl)-naphthalene (0.29 g, 1.3 mmol) according to the general procedure described above. Red solid, 0.22 g (75% yield); R_f = 0.11 (Hexanes/EtOAc 9:1 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 7.93 (s, 1H), 7.91−7.88 (m, 2H), 7.87−7.85 (m, 1H), 7.59 (dd, J = 8.4 and 1.2 Hz 1H), 7.51−7.47 (m, 3H), 6.82 (d, J = 7.8 Hz, 2H), 6.57 (t, J = 7.8 Hz, 1H), 5.20 (unresolved t, 1H), 4.61 (d, J = 5.4 Hz, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 138.7, 137.6, 137.1, 134.5, 133.8, 133.0, 129.0, 128.51, 128.48, 127.0, 126.9 (two C), 126.6, 116.9, 116.0, 115.0, 49.1; ESI-HRMS (m/z): calcd. for $C_{17}H_{14}N_3O_2$ $[M - H]^+$ 292.1086; found 292.1083.

*N^1-([1,1′-Biphenyl]-4-ylmethyl)-3-nitrobenzene-1,2-dia-mine (**9**).* Compound **9** was prepared using 4-biphenylmethyl bromide (0.32 g, 1.3 mmol) according to the general procedure described above. Red solid, 0.20 g (63% yield); R_f = 0.13 (Hexanes/EtOAc 9:1 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 7.67−7.61 (m, 4H), 7.56−7.52 (m, 2H), 7.50 (dd, J = 8.7 and 1.3 Hz, 1H), 7.45 (t, J = 7.8 Hz, 2H), 7.38−7.33 (m, 1H), 6.93−6.67 (m, 3H), 6.60 (dd, J = 8.8 and 7.6 Hz, 1H), 5.14 (t, J = 5.4 Hz, 1H), 4.49 (d, J = 5.4 Hz, 1H); ^{13}C NMR (150 MHz, acetone-d_6): δ 141.5, 140.7, 139.2, 138.7, 137.1, 133.0, 129.7 (two C), 129.0 (two C), 128.1, 127.8 (two C), 127.6 (two C), 116.9, 115.9, 114.9, 48.6; ESI-HRMS (m/z): calcd. for $C_{19}H_{16}N_3O_2$ $[M - H]^+$ 318.1243; found 318.1245.

*3-Nitro-N^1-(quinolin-7-ylmethyl)benzene-1,2-diamine (**10**).* Compound **10** was prepared using 7-(bromomethyl)-quinoline hydrobromide salt (0.39 g, 1.3 mmol) according to the general procedure described above. Red solid, 0.19 g (66%

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

yield); R_f = 0.10 (Hexanes/EtOAc 1:1 v/v); ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.86 (dd, J = 4.3 and 1.7 Hz, 1H), 8.33 (dd, J = 8.2 and 1.7 Hz, 1H), 7.97 (s, 1H), 7.95 (d, J = 8.2 Hz, 1H), 7.63 (dd, J = 8.5 and 1.7 Hz, 1H), 7.49 (dd, J = 8.2 and 4.2 Hz, 1H), 7.31 (dd, J = 8.5 and 1.7 Hz, 1H), 7.24−7.19 (m, 2H), 6.59 (d, J = 8.2 Hz, 1H), 6.49−6.41 (m, 1H), 4.61 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6): δ 150.6, 147.8, 140.9, 137.1, 135.8, 135.8, 130.9, 128.2, 127.0, 126.4, 126.3, 121.1, 115.7, 113.4, 112.8, 46.8; ESI-HRMS (m/z): calcd. for $C_{16}H_{14}N_4O_2Na$ $[M + Na]^+$ 317.1009; found 317.1009; calcd. for $C_{16}H_{15}N_4O_2$ $[M + H]^+$ 295.1190; found 295.1189.

*General Procedure for the Cyclization of Diamines **8** and **9**.* To an ice-cold solution of trimethyl orthoformate (0.13 mL, 1.2 mmol) in DCM (4 mL), $BF_3{\cdot}Et_2O$ (0.1 mL, 0.8 mmol) was added dropwise, followed by the addition of **8** or **9** (0.4 mmol). The resulting mixture was stirred at RT for 5 h and quenched with sat. aqueous $NaHCO_3$. The aqueous phase was extracted twice with DCM. The combined organic layers dried over Na_2SO_4 and evaporated to dryness. Compounds **11** and **12** were obtained in pure form after FCC purification.

*1-(Naphthalen-2-ylmethyl)-4-nitro-1H-benzo[d]imidazole (**11**) (**CKP-20**).* Compound **11** was prepared starting from **8** (0.12 g, 0.4 mmol) according to the general procedure described above. White solid, 0.114 g (94% yield); R_f = 0.14 (Hexanes/EtOAc 1:1 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.64 (s, 1H), 8.03 (dd, J = 7.8 and 1.2 Hz, 2H), 7.96 (dd, J = 8.4 and 1.2 Hz, 1H), 7.92−7.86 (m, 3H), 7.58−7.49 (m, 2H), 7.48 (dd, J = 8.5 and 1.9 Hz, 1H), 7.39 (t, J = 8.1 Hz, 1H), 5.87 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 148.3, 140.5, 138.2, 137.8, 134.5, 134.3, 133.9, 129.7, 128.7, 128.6, 127.4, 127.3, 127.2, 126.0, 122.8, 119.2, 118.1, 49.7; ESI-HRMS (m/z): calcd. for $C_{18}H_{14}N_3O_2$ $[M + H]^+$ 304.1086; found 304.1076. HPLC analysis: t_R = 6.70 min, purity = 99.7%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-4-nitro-1H-benzo[d]-imidazole (**12**) (**CKP-21**).* Compound **12** was prepared starting from **9** (0.13 g, 0.4 mmol) according to the general procedure A described above. White solid, 0.121 g (92% yield); R_f = 0.15 (Hexanes/EtOAc 1:1 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.61 (s, 1H), 8.05 (d, J = 8.4 Hz, 1H), 7.98 (d, J = 8.4 Hz, 1H), 7.67−7.64 (m, 2H), 7.63−7.60 (m, 2H), 7.48−7.42 (m, 5H), 7.37−7.33 (m, 1H), 5.75 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 148.2, 141.7, 141.0, 140.6, 138.2, 137.7, 136.2, 129.7 (two C), 128.9 (two C), 128.4, 128.2 (two C), 127.7 (two C), 122.8, 119.2, 118.0, 49.1; ESI-HRMS (m/z): calcd. for $C_{20}H_{16}N_3O_2$ $[M + H]^+$ 330.1243; found 330.1232. HPLC analysis: t_R = 7.41 min, purity = 100%.

*7-((4-Nitro-1H-benzo[d]imidazol-1-yl)methyl)quinoline (**13**) (**CKP-48**).* A mixture of **10** (0.12 g, 0.4 mmol), trimethyl orthoformate (1.2 mmol) and catalytic amount of p-TsOH·H_2O (0.08 mmol) in toluene (0.1 M) was stirred at 80 °C for 2 h. White solid, 0.106 g (87% yield); R_f = 0.25 (DCM/EtOAc 5:3 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.90 (s, 1H), 8.68 (s, 1H), 8.32 (d, J = 8.5 Hz, 1H), 8.05 (dd, J = 8.1 and 2.6 Hz, 1H), 8.01 (d, J = 4.3 Hz, 2H), 7.96 (dd, J = 8.5 and 2.6 Hz, 1H), 7.56 (d, J = 8.1 Hz, 1H), 7.51 (dt, J = 8.1 and 3.4 Hz, 1H), 7.43 (dt, J = 8.1 and 4.3 Hz, 1H), 5.96 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 152.0, 149.1, 148.3, 140.6, 138.4, 138.2, 137.7, 136.6, 129.8, 128.7, 128.6, 126.4, 122.9, 122.6, 119.3, 118.1, 49.4; ESI-HRMS (m/z): calcd. for $C_{17}H_{12}N_4O_2Na$ $[M + Na]^+$ 327.0852; found 327.0853; calcd. for $C_{17}H_{13}N_4O_2$ $[M + H]^+$ 305.1033; found 305.1032. HPLC analysis: t_R = 5.51 min, purity = 98%.

General Procedure A for Nitro Group Reduction. To a suspension of **3** or **4** or **11** or **12** or **13** (0.3 mmol) in gl. CH_3CO_2H (1.1 mL), conc. HCl (0.44 mL) was added followed by the addition of $SnCl_2{\cdot}2H_2O$ (0.45 g, 2 mmol). The reaction mixture stirred at RT for 30 min. Solvents were azeotropically removed with PhMe and to the residue, an aq. solution NaOH (10 M) and brine were added. The aqueous phase was extracted three times with DCM and the combined organic layers dried over Na_2SO_4 and evaporated to dryness to afford the desired compounds.

*1-(Naphthalen-2-ylmethyl)-1H-benzo[d]imidazol-5-amine (**5**) (**CKP-39**).* Compound **5** was prepared starting from **3** (0.09 g, 0.3 mmol) according to the general procedure A described above. White solid, 0.071 g (86% yield); R_f = 0.26 (DCM/MeOH 95:5 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.08 (s, 1H), 7.90−7.84 (m, 4H), 7.81 (s, 1H), 7.52−7.48 (m, 2H), 7.40 (dd, J = 8.4 and 1.8 Hz, 1H), 7.14 (d, J = 8.4 Hz, 1H), 6.95 (d, J = 2.4 Hz, 1H), 5.59 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 145.9, 145.0, 144.1, 135.4, 134.3, 133.9, 129.5, 129.4, 128.7, 128.6, 128.5, 127.3, 127.2, 127.1, 116.7, 111.2, 110.3, 49.3; ESI-HRMS (m/z): calcd. for $C_{18}H_{16}N_3$ $[M + H]^+$ 274.1344; found 274.1334. HPLC analysis: t_R = 4.48 min, purity = 97.2%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-1H-benzo[d]imidazol-5-amine (**6**) (**CKP-45**).* Compound **6** was prepared starting from **4** (0.1 g, 0.3 mmol) according to the general procedure A described above. White solid, 0.074 g (90% yield); R_f = 0.28 (DCM/MeOH 95:5 v/v); ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.17 (s, 1H), 7.62 (d, J = 7.7 Hz, 4H), 7.44 (t, J = 7.5 Hz, 2H), 7.35 (d, J = 7.7 Hz, 3H), 7.18 (d, J = 8.6 Hz, 1H), 6.79 (d, J = 2.0 Hz, 1H), 6.56 (dd, J = 8.6, 2.0 Hz, 1H), 5.40 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6): δ 144.9, 144.1, 143.3, 139.7, 139.5, 136.5, 128.9 (two C), 127.9 (two C), 127.5, 126.9 (two C), 126.7 (two C), 126.2, 112.3, 110.5, 102.6, 47.3; ESI-HRMS (m/z): calcd. for $C_{20}H_{18}N_3$ $[M + H]^+$ 300.1501; found 300.1489. HPLC analysis: t_R = 5.62 min, purity = 99.3%.

*1-(Naphthalen-2-ylmethyl)-1H-benzo[d]imidazol-4-amine (**14**) (**CKP-29**).* Compound **14** was prepared starting from **11** (0.09,0.3 mmol) according to the general procedure A described above. White solid, 0.073 g (89% yield); R_f = 0.25 (DCM/MeOH 97:3 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.11 (s, 1H), 7.89−7.84 (m, 4H), 7.81 (s, 1H), 7.51−7.47 (m, 3H), 7.41 (dd, J = 8.4 and 1.8 Hz, 1H), 6.90 (t, J = 7.8 Hz, 1H), 6.69 (dd, J = 8.4 and 1.2 Hz, 1H), 6.44 (dd, J = 7.8 and 1.2 Hz, 1H), 5.60 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 143.8, 142.0, 141.2, 135.7, 134.3, 133.8, 129.4, 128.7, 128.5, 127.2, 127.0, 126.9, 126.1, 124.6, 123.8, 105.6, 99.6, 49.2; ESI-HRMS (m/z): calcd. for $C_{18}H_{16}N_3$ $[M + H]^+$ 274.1344; found 274.1333. HPLC analysis: t_R = 6.46 min, purity = 99%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-1H-benzo[d]imidazol-4-amine (**15**) (**CKP-28**).* Compound **15** was prepared starting from **12** (0.1 g, 0.3 mmol) according to the general procedure A described above. White solid, 0.082 g (91% yield); R_f = 0.27 (EtOAc); ^{1}H NMR (600 MHz, acetone-d_6): δ 7.97 (s, 1H), 7.53−7.50 (m, 4H), 7.35−7.30 (m, 2H), 7.29−7.21 (m, 3H), 6.82 (t, J = 7.8 Hz, 1H), 6.60 (dd, J = 8.4 and 1.2 Hz, 1H), 6.34 (dd, J = 8.4 and 1.2 Hz, 1H), 5.39 (s, 2H), 4.85 (br.s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 141.9, 141.3, 141.2, 137.4 (two C), 135.6, 134.0, 129.7 (two C), 128.7 (two C), 128.3, 128.0 (two C), 127.7 (two C), 124.6, 105.6, 99.6, 48.6; ESI-HRMS (m/z): calcd. for $C_{20}H_{18}N_3$ $[M + H]^+$ 300.1501; found 300.1489. HPLC analysis: t_R = 7.36 min, purity = 99.9%.

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

*1-(Quinolin-7-ylmethyl)-1H-benzo[d]imidazol-4-amine (**16**) (**CKP-49**).* Compound **16** was prepared starting from **13** (0.09 g, 0.3 mmol) according to the general procedure A described above. White solid, 0.07 g (89% yield); R_f = 0.16 (DCM/Acetone 5:3 v/v); ^{1}H NMR (600 MHz, acetone-d_6): δ 8.88−8.86 (m, 1H), 8.29 (d, J = 8.5 Hz, 1H), 8.23 (s, 1H), 7.93 (d, J = 4.3 Hz, 1H), 7.91 (d, J = 8.5 Hz, 1H), 7.53−7.45 (m, 2H), 6.91 (t, J = 7.8 Hz, 1H), 6.72 (d, J = 8.1 Hz, 1H), 6.46 (d, J = 7.8 Hz, 1H), 5.72 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 151.8, 149.2, 142.1, 139.5, 136.5 (two C), 129.5 (two C), 128.5, 128.2, 126.5 (two C), 124.7, 122.3, 105.7, 99.5, 48.9; ESI-HRMS (m/z): calcd. for $C_{17}H_{15}N_4$ $[M + H]^+$ 275.1291; found 275.1287. HPLC analysis: t_R = 5.24 min, purity = 99.4%.

General Procedure for N-Alkylation of tert-Butyl (3-Methoxy-2-nitrophenyl)carbamate. NaH (60% w/w dispersion in mineral oil, 1.2 equiv) was added at 0 °C to a solution of *tert*-butyl (3-methoxy-2-nitrophenyl)carbamate **18** (1 equiv) in DMF (0.1 M). The resulting suspension was stirred at the same temperature for 20 min and then the appropriate bromide (1.2 equiv) was added. The reaction mixture was stirred at room temperature for 3 h. Upon completion, the reaction was cooled to 0 °C, quenched with sat. aq. NH_4Cl and extracted with ethyl acetate. The organic layer was washed with brine, dried over Na_2SO_4, filtered and the solvent was evaporated under reduced pressure. The crude residue was used in the next step without further purification

*tert-Butyl (3-Methoxy-2-nitrophenyl)(naphthalen-2-ylmethyl)carbamate (**19**).* Compound **19** was prepared using **18** (0.15 g, 0.56 mmol) and 2-(bromomethyl)-naphthalene (0.15 g, 0.67 mmol) according to the above general procedure above. The compound was used in the next step without further purification, R_f = 0.35 (Hexanes/EtOAc 80:20 v/v).

*tert-Butyl ([1,1′-Biphenyl]-4-ylmethyl)(3-methoxy-2-nitrophenyl)carbamate (**20**).* Compound **20** was prepared using **18** (0.15 g, 0.56 mmol) and 4-biphenylmethyl bromide (0.17 g, 0.67 mmol) according to the above general procedure above. The compound was used in the next step without further purification, R_f = 0.29 (Hexanes/EtOAc 80:20 v/v).

*tert-Butyl (3-Methoxy-2-nitrophenyl)(quinolin-7-ylmethyl)carbamate (**21**).* Compound **21** was prepared using **18** (0.08 g, 0.30 mmol) and 7-(bromomethyl)quinoline hydrobromide (0.11 g, 0.36 mmol) according to the above general procedure above. The compound was used in the next step without further purification, R_f = 0.10 (Hexanes/EtOAc 60:40 v/v).

General Procedure for N-Boc Deprotection of N-Aryl-tert-butyl(3-methoxy-2-nitrophenyl)carbamate. To a solution of *tert*-butyl carbamate **19** or **20** or **21** (1 equiv) in dry DCM (0.1M) was added TFA (20 equiv) at 0 °C and the reaction mixture was stirred at RT for 1 h. The solvent was evaporated in vacuo and the residue was diluted with sat. aq. $NaHCO_3$ and extracted with EtOAc. The organic phase was washed with brine, dried over Na_2SO_4, filtered and evaporated in vacuo. The residue was purified by FCC to afford the desired compound.

*3-Methoxy-N-(naphthalen-2-ylmethyl)-2-nitroaniline (**22**).* The title compound **22** was synthesized from **19** (0.23 g, 0.56 mmol) following the above general procedure. The desired product was obtained after FCC (Hexanes/EtOAc 8:2 v/v) as a yellow solid, 0.14 g, (79% yield over 2 steps from **18**); mp: 111−112 °C; ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.86−7.72 (m, 4H), 7.55−7.37 (m, 3H), 7.16 (t, J = 8.4 Hz, 1H), 6.37 (d, J = 8.6 Hz, 1H), 6.29 (d, J = 8.3 Hz, 1H), 4.59 (s, 2H), 3.88 (s, 3H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 155.2, 144.0, 135.4, 133.7, 133.6, 133.0, 128.8, 127.9, 127.8, 126.5, 126.1, 125.7, 125.2, 105.8, 100.3, 56.6, 47.8; ESI-HRMS (m/z): calcd. for $C_{18}H_{17}N_2O_3$ $[M + H]^+$ 309.1234; found 309.1228.

*N-([1,1′-Biphenyl]-4-ylmethyl)-3-methoxy-2-nitroaniline (**23**).* The title compound **23** was synthesized from **20** (0.24 g, 0.56 mmol) following the above general procedure. The desired product was obtained after FCC (Hexanes/EtOAc 8:2 v/v) as an orange solid, 0.14 g (77% over 2 steps from **18**); mp: 129−130 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 7.73−7.60 (m, 4H), 7.50−7.42 (m, 4H), 7.37−7.32 (m, 1H), 7.22 (t, J = 8.4 Hz, 1H), 6.45 (t, J = 7.7 Hz, 3H), 4.59 (d, J = 5.9 Hz, 2H), 3.86 (s, 3H); ^{13}C NMR (150 MHz, acetone-d_6): δ 154.7, 142.9, 141.5, 140.7, 139.2, 133.5, 129.7, 128.4, 128.1, 127.9, 127.6, 106.5, 100.8, 56.7, 47.1; ESI-MS (m/z): 357.34 $(M+Na)^+$, 690.80 $(2M+Na)^+$.

*3-Methoxy-2-nitro-N-(quinolin-7-ylmethyl)aniline (**24**).* The title compound **24** was synthesized from **21** (0.11 g, 0.27 mmol) following the above general procedure. The desired product was obtained as an orange solid, 0.07 g, (80% over 2 steps from **18**); mp: 176−177 °C; ^{1}H NMR (600 MHz, $CDCl_3$): δ 9.01−8.79 (m, 1H), 8.18 (d, J = 8.3 Hz, 1H), 8.08 (s, 1H), 7.83 (d, J = 8.5 Hz, 1H), 7.54 (d, J = 8.4 Hz, 1H), 7.42 (dd, J = 8.5 and 4.1 Hz, 1H), 7.15 (t, J = 8.4 Hz, 1H), 6.72 (s, 1H), 6.34 (d, J = 8.6 Hz, 1H), 6.29 (d, J = 8.3 Hz, 1H), 4.67 (d, J = 5.8 Hz, 2H), 3.88 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.8, 149.7, 143.3, 136.6, 133.3, 128.3, 127.4, 126.2, 125.7, 120.8, 105.2, 100.2, 56.2, 47.1; ESI-HRMS (m/z): calcd. for $C_{17}H_{16}N_3O_3$ $[M + H]^+$ 310.1186; found 310.1177.

General Procedure B for Nitro Group Reduction. $SnCl_2$ (5 equiv) was added to a mixture of the nitro compound **22** or **23** or **24** (1 equiv) in MeOH (0.1M) and conc. HCl (70 equiv) and stirred at 50 °C overnight. The mixture was concentrated in vacuo. The residue was taken up in EtOAc and washed with sat. aq. $NaHCO_3$, brine, dried over Na_2SO_4, filtered and evaporated in vacuo. The crude was purified by FCC to give the desired compound.

*3-Methoxy-N^1-(naphthalen-2-ylmethyl)benzene-1,2-diamine (**25**).* The title compound **25** was prepared starting from **22** (0.12 g, 0.39 mmol) according to the above general procedure B. Purification of the compound with FCC (Hexanes/EtOAc 8:2 v/v) gave a brown oil, 0.089 g (82% yield); R_f = 0.25 (Hexanes/EtOAc 8:2 v/v); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.89−7.74 (m, 4H), 7.57−7.35 (m, 3H), 6.79−6.75 (m, 1H), 6.45 (t, J = 8.1 Hz, 2H), 4.51 (s, 2H), 3.87 (s, 3H), 3.70 (bs, 3H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 148.4, 137.1, 133.6, 132.9, 128.4, 127.9, 127.8, 126.2, 126.2, 126.1, 125.8, 122.9, 119.8, 105.8, 102.1, 55.9, 49.1; ESI-MS (m/z): 279.22 $(M+H)^+$.

*N^1-([1,1′-Biphenyl]-4-ylmethyl)-3-methoxybenzene-1,2-diamine (**26**).* The title compound **26** was prepared starting from **23** (0.120 g, 0.36 mmol) according to the above general procedure B. Purification of the compound with FCC (Hexanes/EtOAc 8:2 v/v) gave a brown oil, 0.082 g (75% yield); R_f = 0.20 (Hexanes/EtOAc 8:2 v/v); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.61 (dd, J = 11.8 and 8.0 Hz, 4H), 7.48 (dd, J = 16.5 and 8.0 Hz, 4H), 7.38 (t, J = 7.4 Hz, 1H), 6.80 (t, J = 8.1 Hz, 1H), 6.47 (dd, J = 14.8 and 8.1 Hz, 2H), 4.41 (s, 2H), 3.88 (s, 3H), 3.68 (bs, 3H); ^{13}C NMR (150 MHz, $CDCl_3$): δ

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

148.4, 141.0, 140.3, 138.5, 138.0, 128.9, 128.3, 127.4, 127.3, 127.2, 122.9, 119.8, 105.8, 102.2, 55.9, 48.6; ESI-HRMS (m/z): calcd. for $C_{20}H_{21}N_2O$ $[M + H]^+$ 305.1648; found 305.1649.

*3-Methoxy-N¹-(quinolin-7-ylmethyl)benzene-1,2-diamine (**27**).* The title compound **27** was prepared starting from **24** (0.058 g, 0.19 mmol) according to the above general procedure B. Purification of the compound with FCC (Hexanes/EtOAc 7:3 to 6:4 v/v) gave a brown oil, 0.043 g (83% yield); R_f = 0.08 (Hexanes/EtOAc 6:4 v/v); ^{1}H NMR (600 MHz, $CDCl_3$): δ 9.02−8.83 (m, 1H), 8.24−8.05 (m, 2H), 7.80 (d, J = 8.4 Hz, 1H), 7.61 (d, J = 8.1 Hz, 1H), 7.40 (dd, J = 8.2 and 4.3 Hz, 1H), 6.72 (t, J = 8.2 Hz, 1H), 6.41 (dd, J = 22.6 and 8.2 Hz, 2H), 4.57 (s, 2H), 3.85 (s, 3H), 3.54 (bs, 3H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 150.2, 148.4, 147.9, 141.6, 137.9, 136.2, 128.0, 127.4, 127.0, 126.6, 122.7, 120.8, 119.7, 105.5, 102.0, 55.7, 48.5; ESI-HRMS (m/z): calcd. for $C_{17}H_{18}N_3O$ $[M + H]^+$ 280.1444; found 280.1437.

*General Procedure B for Cyclization Diamines **25**−**27**.* A mixture of the diamine **25** or **26** or **27** (1 equiv), trimethyl orthoformate (3 equiv) and catalytic amount of p-TsOH·H_2O (0.2 equiv) in toluene (0.1 M) was stirred at 80 °C for 2−18 h. After completion of the reaction the solvent was removed in vacuo. The residue was purified by FCC to afford the desired compound.

*4-Methoxy-1-(naphthalen-2-ylmethyl)-1H-benzo[d]-imidazole (**28**) (**CKP-25**).* Compound **28** was prepared using **25** (0.088 g, 0.31 mmol) according to the general procedure B described above. Purification with FCC (DCM/Acetone 9:1 v/v) gave a white solid, 0.061 g (67% yield); R_f = 0.40 (DCM/Acetone 9:1 v/v); mp: 163−164 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 8.21 (s, 1H), 7.90−7.81 (m, 4H), 7.49 (dd, J = 6.5 and 2.9 Hz, 2H), 7.42 (d, J = 8.4 Hz, 1H), 7.11−7.05 (m, 2H), 6.69 (d, J = 7.5 Hz, 1H), 5.68 (s, 2H), 4.00 (s, 3H); ^{13}C NMR (150 MHz, acetone-d_6): δ 153.0, 143.3, 137.0, 135.6, 134.4, 134.0, 129.6, 128.8, 128.7, 127.4, 127.1, 127.1, 126.1, 124.3, 104.6, 104.2, 56.5, 49.3; ESI-HRMS (m/z): calcd for $C_{19}H_{17}N_2O$ $[M + H]^+$ 289.1335, found 289.1336. HPLC analysis: t_R = 5.89 min, purity = 100%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-4-methoxy-1H-benzo[d]-imidazole (**29**) (**CKP-27**).* Compound **29** was prepared using **26** (0.080 g, 0.26 mmol) according to the general procedure B described above. Purification with FCC (DCM/Acetone 8:2 v/v) gave a white solid, 0.062 g (76% yield); R_f = 0.35 (DCM/Acetone 9:1 v/v); mp: 150−151 °C. ^{1}H NMR (600 MHz, acetone-d_6): δ 8.18 (s, 1H), 7.63 (dd, J = 7.9 and 4.1 Hz, 4H), 7.44 (t, J = 7.6 Hz, 2H), 7.39 (d, J = 7.9 Hz, 2H), 7.35 (t, J = 7.3 Hz, 2H), 7.13 (t, J = 7.9 Hz, 1H), 7.08 (d, J = 7.8 Hz, 1H), 6.71 (d, J = 7.8 Hz, 1H), 5.57 (s, 3H), 4.01 (s, 4H); ^{13}C NMR (150 MHz, acetone-d_6): δ 152.0, 142.2, 140.3, 136.2, 128.8, 127.8, 127.4, 127.2, 126.8, 123.3, 103.6, 103.2, 55.5, 47.8; ESI-HRMS (m/z): calcd for $C_{21}H_{19}N_2O$ $[M + H]^+$ 315.1492, found 315.1494. HPLC analysis: t_R = 5.98 min, purity = 99.3%.

*7-((4-Methoxy-1H-benzo[d]imidazol-1-yl)methyl)-quinoline (**30**) (**CKP-44**).* Compound **30** was prepared using **27** (0.026 g, 0.093 mmol) according to the general procedure B described above. Purification with FCC (DCM/Acetone 8:2 v/v) gave a white solid, 0.020 g (74% yield); R_f = 0.25 (DCM/Acetone 9:1 v/v); mp: 161−162 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 8.88 (dd, J = 4.3 and 1.8 Hz, 1H), 8.32−8.23 (m, 2H), 7.98−7.91 (m, 2H), 7.50 (ddd, J = 14.1, 8.4, and 3.0 Hz, 2H), 7.16−7.07 (m, 2H), 6.72 (dd, J = 7.5 and 1.3 Hz, 1H), 5.78 (s, 2H), 4.01 (s, 3H); ^{13}C NMR (150 MHz, acetone-d_6): δ 152.9, 151.8, 149.1, 143.2, 139.3, 136.8, 136.6, 129.6, 128.6, 128.2, 126.5, 124.4, 122.4, 104.6, 104.1, 56.4, 49.0; ESI-HRMS (m/z): calcd for $C_{18}H_{16}N_3O$ $[M + H]^+$ 290.1288, found 290.1292. HPLC analysis: t_R = 5.17 min, purity = 99.8%.

General Procedure for Methoxy Group Removal. To an ice-cold solution of the **28** or **29** (1 equiv) in DCM (0.05 M), $BF_3S(CH_3)_2$ (10 equiv, per methoxy group) was added and the reaction mixture was stirred at RT for 2 d. After completion of the reaction, the solvent and excess reagent were evaporated under argon stream. To the residue sat. aq. $NaHCO_3$ was added and extracted with ethyl acetate. The organic layer was washed with sat. aq. NaCl, dried over Na_2SO_4 and the solvent was evaporated in vacuo. The residue was purified by FCC to obtain the desired compound.

*1-(Naphthalen-2-ylmethyl)-1H-benzo[d]imidazol-4-ol (**31**) (**CKP-26**).* Compound **31** was synthesized from **28** (0.026 g, 0.09 mmol) following the above general method. Purification with FCC (DCM/Acetone 9:1 v/v) gave a white solid, 0.015 g (60% yield); R_f = 0.28 (DCM/Acetone 8:2 v/v); mp: 235−236 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 8.19 (s, 1H), 7.85−7.78 (m, 3H), 7.72 (s, 2H), 7.47 (dd, J = 5.9 and 3.3 Hz, 1H), 7.35 (d, J = 8.5 Hz, 1H), 7.04 (t, J = 7.9 Hz, 1H), 6.91 (d, J = 8.2 Hz, 1H), 6.63 (d, J = 7.7 Hz, 1H), 5.61 (s, 2H); ^{13}C NMR (150 MHz, CD_3OD): δ 150.4, 143.4, 137.1, 135.2, 134.8, 134.4, 133.9, 129.8, 128.9, 128.7, 127.5, 127.3, 127.2, 125.9, 125.3, 108.2, 103.0, 49.8; ESI-HRMS (m/z): calcd for $C_{18}H_{15}ON_2$ $[M + H]^+$ 275.1179, found 275.1182.; HPLC analysis: t_R = 5.53 min, purity = 99.4%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-1H-benzo[d]imidazol-4-ol (**32**) (**CKP-30**).* Compound **32** was synthesized from **29** (0.039 g, 0.12 mmol) following the above general method. Purification with FCC (DCM/Acetone 9:1 v/v) gave an off-white solid, 0.023 g, (62% yield); R_f = 0.20 (DCM/Acetone 8:2 v/v); mp: 258−259 °C; ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.31 (s, 1H), 7.55 (dd, J = 20.7 and 7.8 Hz, 4H), 7.42 (t, J = 7.6 Hz, 2H), 7.36−7.32 (m, 1H), 7.29 (d, J = 7.8 Hz, 2H), 7.23−7.19 (m, 2H), 6.88 (dd, J = 22.3 and 8.0 Hz, 2H), 5.42 (s, 2H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 147.6, 141.6, 139.9, 132.8, 128.8, 127.8, 127.8, 127.6, 126.9, 125.9, 109.7, 101.8, 49.4; ESI-HRMS (m/z): calcd for $C_{20}H_{17}ON_2$ $[M + H]^+$ 301.1335, found 301.1339. HPLC analysis: t_R = 5.55 min, purity = 100%.

General Procedure for N-Alkylation of 1H-Benzo[d]-imidazole. To an ice-cold solution of 1*H*-benzo[*d*]imidazole (**1**) (1 equiv) in THF (0.3 M), NaH (60% w/w dispersion in mineral oil, 1.2 equiv) was added portion wise. The reaction was warmed to RT and stirred at the same temperature for 30 min followed by the addition of the appropriate bromide or mesylate (1.05 equiv). The reaction mixture was stirred at RT for 17 h, diluted with water and extracted twice with EtOAc. The combined organic layers were washed with water, dried over Na_2SO_4 and evaporated to dryness. Pure compound was obtained after FCC.

*1-(Naphthalen-2-ylmethyl)-1H-benzo[d]imidazole[122] (**CKP-22**).* Compound **CKP-22** was prepared according to the above general procedure using 1 (0.08 g, 0.66 mmol) and 2-(bromomethyl)naphthalene (0.15 g, 0.69 mmol). The desired product was obtained after FCC (Hexanes/EtOAc 1:1 v/v) as a white solid 0.17 g (99% yield); mp: 139−140 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 8.32 (s, 1H), 7.95−7.80 (m, 4H), 7.73−7.65 (m, 1H), 7.54−7.47 (m, 3H), 7.44 (dd, J = 8.5 and 1.8 Hz, 1H), 7.19 (tt, J = 7.2 and 5.5 Hz, 2H), 5.72 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 145.3, 144.8, 135.4,

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

135.1, 134.3, 133.9, 129.5, 128.7, 128.5 127.3, 127.1, 126.1, 123.3, 122.5, 120.7, 111.3, 49.2.

*1-(2-(Naphthalen-2-yl)ethyl)-1H-benzo[d]imidazole (**33**) (**CKP-46**).* Compound **33** was prepared according to the above general procedure using **1** (0.08 g, 0.66 mmol) and 2-(2-bromoethyl)naphthalene (0.19 g, 0.79 mmol). The desired product was obtained after FCC (Hexanes/EtOAc 2:1 v/v) as an off-white solid, 0.04 g (24% yield); mp: 139−142 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 7.90 (s, 1H), 7.88−7.83 (m, 1H), 7.82 (d, J = 8.4 Hz, 1H), 7.78−7.74 (m, 1H), 7.68−7.62 (m, 3H), 7.49−7.42 (m, 2H), 7.38 (dd, J = 8.4 and 1.9 Hz, 1H), 7.27−7.24 (m, 1H), 7.23−7.19 (m, 1H), 4.68 (t, J = 7.3 Hz, 2H), 3.38 (t, J = 7.3 Hz, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 144.4, 136.8, 133.3, 129.0, 128.4, 128.3, 128.2, 128.0, 126.9, 126.4, 123.2, 122.3, 120.6, 111.0, 46.7, 36.9; ESI-HRMS (m/z): calcd. for $C_{19}H_{17}N_2$ $[M + H]^+$ 273.1386; found 273.1384. HPLC analysis: t_R = 4.16 min, purity = 99.6%.

*1-([1,1′-Biphenyl]-4-ylmethyl)-1H-benzo[d]imidazole (**34**) (**CKP-23**).*[123] Compound **34** was prepared according to the above general procedure using **1** (0.07 g, 0.60 mmol) and 4-biphenylmethyl bromide (0.15 g, 0.63 mmol). The desired product was obtained after FCC (Hexanes/EtOAc 1:1 v/v) as a white solid, 0.16 g (95% yield); mp: 184−186 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 8.30 (s, 1H), 7.73−7.67 (m, 1H), 7.66−7.60 (m, 4H), 7.56−7.48 (m, 1H), 7.48−7.39 (m, 4H), 7.39−7.29 (m, 1H), 7.22 (dd, J = 6.5 and 2.7 Hz, 2H), 5.61 (s, 2H); ^{13}C NMR (150 MHz, acetone-d_6): δ 145.3, 144.4, 141.4, 141.2, 137.0, 135.0, 129.7, 128.7, 128.3, 128.1, 127.7, 123.3, 122.5, 120.7, 111.3, 48.6.[123]

*7-((1H-Benzo[d]imidazol-1-yl)methyl)quinoline (**35**) (**CKP-36**).* Compound **35** was prepared according to the above general procedure using **1** (0.08 g, 0.66 mmol) and 7-(bromomethyl)quinoline hydrobromide (0.22 g, 0.72 mmol). The desired product was obtained after FCC (Hexanes/EtOAc 1:1 v/v) as a beige solid, 0.05 g (27% yield); mp: 166−168 °C decompose; ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.92 (dd, J = 4.3 and 1.8 Hz, 1H), 8.19 (s, 1H), 8.14 (dd, J = 8.4 and 1.8 Hz, 1H), 8.00 (s, 1H), 7.85 (d, J = 8.0 Hz, 1H), 7.79 (d, J = 8.4 Hz, 1H), 7.42 (dd, J = 8.3 and 4.2 Hz, 1H), 7.33 (td, J = 5.2 and 2.5 Hz, 2H), 7.31−7.18 (m, 2H), 5.60 (s, 2H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 151.3, 148.2, 143.2, 136.9, 136.1, 129.2, 128.1, 127.9, 125.4, 123.7, 122.9, 121.8, 120.4, 110.3, 49.2; ESI-HRMS (m/z): calcd. for $C_{17}H_{14}N_3$ $[M + H]^+$ 260.1182; found 260.1178. HPLC analysis: t_R = 3.05 min, purity = 100%.

*1-(3,4-Dimethoxybenzyl)-1H-benzo[d]imidazole (**36**) (**CKP-24**).*[124] Compound 36 was prepared according to the above general procedure using **1** (0.15 g, 1.30 mmol) and 3,4-dimethoxybenzyl methanesulfonate (0.34 g, 1.37 mmol). The desired product was obtained after FCC (Hexanes/EtOAc 1:1 v/v) as a white solid, 0.17 g (49% yield); mp: 131−132 °C; ^{1}H NMR (600 MHz, acetone-d_6): δ 8.21 (s, 1H), 7.72−7.60 (m, 1H), 7.56−7.44 (m, 1H), 7.26−7.14 (m, 2H), 7.05 (s, 1H), 6.95−6.82 (m, 2H), 5.44 (s, 2H), 3.76 (d, J = 6.6 Hz, 6H); ^{13}C NMR (150 MHz, acetone-d_6): δ 150.5, 150.1, 145.3, 144.5, 135.0, 130.1, 123.1, 122.3, 120.8, 120.6, 112.6, 112.4, 111.4, 56.0, 55.9, 48.8.

*4-((1H-Benzo[d]imidazol-1-yl)methyl)benzene-1,2-diol (**37**) (**CKP-35**).* Compound **37** was synthesized from **36** (0.050 g, 0.19 mmol) following the above general method. Purification with FCC (DCM/MeOH 9:1 v/v) gave a white solid, 0.017 g (37% yield); mp: 188−190 °C (dec.); ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.92 (d, J = 10.7 Hz, 2H), 8.48 (s, 1H), 7.66 (d, J = 7.6 Hz, 1H), 7.54 (d, J = 7.6 Hz, 1H), 7.34−7.18 (m, 2H), 6.51−6.89 (m, 3H), 5.32 (s, 2H); ^{13}C NMR (150 MHz, DMSO-d_6): δ 145.4, 145.0, 143.8, 133.4, 127.4, 122.6, 122.0, 118.9, 118.7, 115.5, 114.9, 111.1, 47.7; ESI-HRMS (m/z): calcd. for $C_{14}H_{13}N_2O_2$ $[M + H]^+$ 241.0972; found 241.0971. HPLC analysis: t_R = 3.63 min, purity = 99.6%.

General Procedure for the Quaternization. To a solution of the appropriate N^1-substituted benzimidazole **28** or **29** or **CKP-22** or **33**, or **34** or **36** (1 equiv) in 1,4-dioxane (0.03 M) the appropriate alkylating agent (1.3 equiv) was added and the resulting mixture was stirred at 100 °C for 24 h. Then, the reaction was cooled to RT. The precipitate was filtered and the solid was washed with dioxane and diethyl ether. The resulting compound was pure as obtained or was recrystallized from acetone when necessary.

*(E)-1-(Naphthalen-2-ylmethyl)-3-(3-(4-nitrophenyl)allyl)-1H-benzo[d]imidazol-3-ium Bromide (**38**) (**CKP-31**).* The title compound **38** was prepared according to the above general procedure using 1-(naphthalen-2-ylmethyl)-1*H*-benzo[*d*]imidazole **CKP-22** (0.022 g, 0.082 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.026 g, 0.107 mmol). Pure compound was obtained as off-white solid, 0.031 g (75% yield); mp: 189−191 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 9.76 (s, 1H), 8.21 (d, J = 8.9 Hz, 2H), 8.10−8.00 (m, 2H), 7.95 (dd, J = 8.4 and 4.2 Hz, 2H), 7.89 (dd, J = 6.5 and 3.4 Hz, 2H), 7.73−7.66 (m, 3H), 7.56 (ddd, J = 13.6, 7.4, and 2.6 Hz, 2H), 7.04 (d, J = 15.9 Hz, 1H), 6.81 (dt, J = 15.9 and 6.4 Hz, 1H), 5.95 (s, 2H), 5.42 (d, J = 5.9 Hz, 2H); ^{13}C NMR (150 MHz, CD_3OD): δ 149.0, 143.4, 135.3, 134.9, 134.8, 133.2, 133.1, 131.8, 130.5, 129.2, 129.1, 128.9, 128.8, 128.53, 128.50, 128.1, 128.0, 126.9, 126.3, 125.0, 115.1, 114.9, 52.3, 50.2; ESI-HRMS (m/z): calcd for $C_{27}H_{22}N_3O_2$ $(M - Br)^+$ 420.1706; found 420.1704; HPLC analysis: t_R = 5.57 min, purity = 98.9%.

*(E)-3-(3-(4-Chlorophenyl)allyl)-1-(naphthalen-2-ylmethyl)-1H-benzo[d]imidazol-3-ium Bromide (**39**) (**CKP-32**).* The title compound **39** was prepared according to the above general procedure using 1-(naphthalen-2-ylmethyl)-1*H*-benzo[*d*]imidazole **CKP-22** (0.021 g, 0.081 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-chlorobenzene (0.025 g, 0.106 mmol). Pure compound was obtained as white solid, 0.03 g (74% yield); Solid purification by recrystallization with acetone; mp: 166−167 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 8.03 (d, J = 6.1 Hz, 1H), 7.94 (t, J = 7.0 Hz, 1H), 7.89 (t, J = 4.9 Hz, 1H), 7.69 (dt, J = 7.7 and 24.2 Hz, 1H), 7.52−7.59 (m, 2H), 7.47 (d, J = 8.1 Hz, 1H), 7.35 (d, J = 8.6 Hz, 1H), 6.93 (d, J = 15.9 Hz, 1H), 6.57 (dt, J = 6.5 and 14.1 Hz, 1H), 5.93 (s, 1H), 5.49 (s, 2H), 5.34 (d, J = 6.6 Hz, 2H); ^{13}C NMR (150 MHz, CD_3OD): δ 136.7, 135.7, 135.5, 134.9, 134.8, 133.2, 133.1, 131.8, 130.5, 129.9, 129.4, 129.1, 128.9, 128.5, 128.1, 128.0, 126.2, 122.6, 115.0, 114.9, 52.2, 50.4; ESI-HRMS (m/z): calcd for $C_{27}H_{22}ClN_2$ $(M - Br)^+$ 409.1466; found 409.1461; HPLC analysis: t_R = 5.84 min, purity = 100%.

*(E)-1-(2-(Naphthalen-2-yl)ethyl)-3-(3-(4-nitrophenyl)-allyl)-1H-benzo[d]imidazol-3-ium Bromide (**40**) (**CKP-47**).* The title compound **40** was prepared according to the above general procedure using 1-(2-(naphthalen-2-yl)ethyl)-1*H*-benzo[*d*]imidazole **33** (**CKP-46**) (0.020 g, 0.073 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.023 g, 0.095 mmol). Pure compound was obtained as off-white solid, 0.021 g (54% yield); mp: 176−177 °C decomposition; ^{1}H NMR (600 MHz, CD_3OD): δ 8.20 (d, J = 8.8 Hz, 2H), 8.07−8.03 (m, 1H), 8.00−7.97 (m, 1H), 7.81−7.70 (m, 5H), 7.65 (d, J = 7.9 Hz, 1H), 7.53−7.48 (m, 2H), 7.46 (s, 1H), 7.40 (dddd, J = 18.1 and 8.2 and 6.9 and 1.4 Hz, 2H), 7.28 (dd, J = 8.4 and 1.8

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Hz, 1H), 6.75 (d, J = 15.9 Hz, 1H), 6.42 (dt, J = 15.8, 6.4 Hz, 1H), 5.22 (d, J = 6.4 Hz, 2H), 4.95 (t, J = 6.8 Hz, 2H), 3.49 (t, J = 6.7 Hz, 2H); ^{13}C NMR (150 MHz, CD_3OD): δ 149.0, 143.1, 135.3, 135.2, 134.9, 134.0, 132.82, 132.79, 129.9, 128.8, 128.7, 128.6, 128.52, 128.50, 127.7, 127.5, 127.1, 126.4, 124.9, 114.8, 114.7, 49.8, 49.6, 36.1; ESI-HRMS (m/z): calcd for $C_{28}H_{24}N_3O_2$ $(M - Br)^+$ 434.1863; found 434.1859; HPLC analysis: t_R = 4.28 min, purity = 99.5%.

*(E)-1-([1,1′-Biphenyl]-4-ylmethyl)-3-(3-(4-nitrophenyl)-allyl)-1H-benzo[d]imidazol-3-ium Bromide (**41**) (**CKP-33**).* The title compound **42** was prepared according to the above general procedure using 1-([1,1′-biphenyl]-4-ylmethyl)-1*H*-benzo[*d*]imidazole **34** (**CKP-23**) (0.020 g, 0.071 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.022 g, 0.092 mmol). Pure compound was obtained as white solid, 0.014 g (36.9% yield); mp: 236.0−238.0 °C decomposition; δ 8.21 (d, J = 8.3 Hz, 2H), 8.04 (d, J = 7.9 Hz, 1H), 7.96 (d, J = 7.8 Hz, 1H), 7.71 (t, J = 7.5 Hz, 7H), 7.60 (t, J = 7.2 Hz, 4H), 7.44 (t, J = 7.6 Hz, 2H), 7.36 (t, J = 7.5 Hz, 1H), 7.04 (d, J = 15.8 Hz, 1H), 6.81 (dt, J = 16.0, 6.5 Hz, 1H), 5.83 (s, 2H), 5.42 (d, J = 6.4 Hz, 2H); ^{13}C NMR (150 MHz, CD_3OD): δ 149.0, 143.6, 143.4, 141.3, 135.3, 133.4, 133.2, 133.0, 130.1, 130.0, 129.0, 128.8, 128.6, 128.5, 128.0, 126.9, 125.0, 115.1, 114.9, 51.8, 50.2; ESI-HRMS (m/z): calcd for $C_{29}H_{24}N_3O_2$ $(M - Br)^+$ 446.1863; found 446.1859; HPLC analysis: t_R = 4.66 min, purity = 99.8%.

*(E)-1-(3,4-Dimethoxybenzyl)-3-(3-(4-nitrophenyl)allyl)-1H-benzo[d]imidazol-3-ium Bromide (**42**) (**CKP-34**).* The title compound **41** was prepared according to the above general procedure using 1-(3,4-dimethoxybenzyl)-1*H*-benzo-[*d*]imidazole **36** (0.021 g, 0.076 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.024 g, 0.099 mmol). Pure product was obtained as off-white solid, 0.018 g (55% yield); mp: 146−148 °C (dec.); ^{1}H NMR (600 MHz, CD_3OD): δ 9.62 (s, 1H), 8.21 (d, J = 8.8 Hz, 2H), 7.99 (dd, J = 29.6 and 7.7 Hz, 2H), 7.77−7.67 (m, 4H), 7.17 (s, 1H), 7.13−6.98 (m, 3H), 6.85−6.75 (m, 1H), 5.68 (s, 2H), 5.40 (d, J = 6.4 Hz, 2H), 3.83 (s, 6H);^{13}C NMR (150 MHz, CD_3OD): δ 152.7, 152.4, 150.2, 144.62, 144.60, 144.2, 136.6, 136.4, 134.4, 134.3, 130.1, 130.0, 129.75, 129.71, 129.1, 128.24, 128.23, 128.1, 127.78, 127.77, 126.2, 124.1, 116.3, 116.12, 116.06, 114.8, 114.4, 57.9, 57.7, 53.3, 51.4. ESI-HRMS (m/z): calcd for $C_{25}H_{24}N_3O_4$ $(M - Br)^+$ 430.1761; found 430.1759. HPLC analysis: t_R = 3.78 min, purity = 100%.

*(E)-4-Methoxy-1-(naphthalen-2-ylmethyl)-3-(3-(4-nitrophenyl)allyl)-1H-benzo[d]imidazol-3-ium Bromide (**43**) (**CKP-42**).* The title compound **43** was prepared according to the above general procedure using 4-methoxy-1-(naphthalen-2-ylmethyl)-1*H*-benzo[*d*]imidazole **28** (**CKP-25**) (0.017 g, 0.059 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.019 g, 0.077 mmol). Pure compound was obtained as yellowish solid, 0.026 g (84% yield); mp: 227−228 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 9.63 (s, 1H), 8.21 (d, J = 8.7 Hz, 2H), 8.02 (s, 1H), 7.94 (d, J = 8.5 Hz, 1H), 7.91−7.87 (m, 2H), 7.69 (d, J = 8.7 Hz, 2H), 7.59−7.52 (m, 4H), 7.47 (d, J = 8.5 Hz, 1H), 7.22 (d, J = 8.2 Hz, 1H), 6.90 (d, J = 15.9 Hz, 1H), 6.84−6.79 (m, 1H), 5.88 (s, 2H), 5.52 (d, J = 6.2 Hz, 2H), 4.09 (s, 3H); ^{13}C NMR (150 MHz, CD_3OD): δ 150.1, 148.9, 143.6, 142.9, 135.0, 134.8, 134.7, 134.5, 131.8, 130.4, 129.8, 129.1, 129.06, 128.8, 128.7, 128.4, 128.1, 127.9, 126.2, 124.9, 123.0, 109.4, 106.7, 57.1, 52.5, 52.3; ESI-HRMS (m/z): calcd for $C_{28}H_{24}O_3N_3$ $[M - Br]^+$ 450.1812, found 450.1808; HPLC analysis: t_R = 4.11 min, purity = 97.8%.

*(E)-3-(3-(4-Chlorophenyl)allyl)-4-methoxy-1-(naphthalen-2-ylmethyl)-1H-benzo[d]imidazol-3-ium Bromide (**44**) (**CKP-43**).* The title compound **44** was prepared according to the above general procedure using 4-methoxy-1-(naphthalen-2-ylmethyl)-1*H*-benzo[*d*]imidazole **28** (**CKP-27**) (0.020 g, 0.069 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-chlorobenzene (0.021 g, 0.090 mmol). Pure compound was obtained as white solid, 0.013 g (36% yield); mp: 228−229 °C; ^{1}H NMR (600 MHz, $CDCl_3$): δ 11.73 (s, 1H), 7.97 (s, 1H), 7.85−7.77 (m, 3H), 7.54 (d, J = 8.5 Hz, 1H), 7.48 (dd, J = 6.2 and 3.2 Hz, 2H), 7.38 (d, J = 8.1 Hz, 3H), 7.25 (d, J = 8.4 Hz, 2H), 7.14 (d, J = 8.4 Hz, 1H), 6.99−6.92 (m, 2H), 6.55 (dt, J = 15.1 and 6.8 Hz, 1H), 5.96 (s, 2H), 5.55 (d, J = 6.8 Hz, 2H), 4.04 (s, 3H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 148.5, 135.9, 134.2, 133.8, 133.18, 133.16, 132.9, 129.9, 129.4, 128.7, 128.3, 128.2, 128.0, 127.8, 127.7, 126.8, 126.7, 125.1, 122.0, 121.2, 107.5, 105.5, 56.3, 51.7; ESI-HRMS (m/z): calcd for $C_{28}H_{24}ON_2Cl$ $[M - Br]^+$ 439.1572, found 439.1566; HPLC analysis: t_R = 5.08 min, purity = 99.4%.

*(E)-1-([1,1′-Biphenyl]-4-ylmethyl)-4-methoxy-3-(3-(4-nitrophenyl)allyl)-1H-benzo[d]imidazol-3-ium Bromide (**45**) (**CKP-40**).* The title compound **45** was prepared according to the above general procedure using 1-([1,1′-biphenyl]-4-ylmethyl)-4-methoxy-1*H*-benzo[*d*]imidazole **29** (**CKP-27**) (0.020 g, 0.064 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-nitrobenzene (0.020 g, 0.083 mmol). Pure compound was obtained as off-white solid, 0.033 g (92% yield); mp: 203−204 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 8.21 (d, J = 8.4 Hz, 2H), 7.70 (d, J = 8.2 Hz, 4H), 7.62−7.58 (m, 4H), 7.55 (d, J = 7.9 Hz, 1H), 7.50−7.41 (m, 3H), 7.39−7.34 (m, 1H), 7.24 (d, J = 8.2 Hz, 1H), 6.90 (d, J = 16.1 Hz, 1H), 6.81 (dt, J = 16.0 and 6.2 Hz, 1H), 5.75 (s, 2H), 5.52 (d, J = 6.3 Hz, 2H), 4.09 (s, 3H); ^{13}C NMR (150 MHz, CD_3OD): δ 150.0, 148.8, 143.5, 141.2, 134.8, 134.4, 133.3, 129.9, 129.7, 128.8, 128.7, 128.6, 128.2, 127.8, 124.8, 109.3, 106.5, 57.0, 52.4, 51.7; ESI-HRMS (m/z): calcd for $C_{30}H_{26}O_3N_3$ $[M - Br]^+$ 476.1969, found 476.1966. HPLC analysis: t_R = 4.65 min, purity = 100%.

*(E)-1-([1,1′-Biphenyl]-4-ylmethyl)-3-(3-(4-chlorophenyl)-allyl)-4-methoxy-1H-benzo[d]imidazol-3-ium Bromide (**46**) (**CKP-41**).* The title compound **46** was prepared according to the above general procedure using 1-([1,1′-biphenyl]-4-ylmethyl)-4-methoxy-1*H*-benzo[*d*]imidazole **29** (**CKP-27**) (0.014, 0.062 mmol) and (*E*)-1-(3-bromoprop-1-en-1-yl)-4-chlorobenzene (0.015 g, 0.048 mmol). Pure product was obtained as white solid, 0.018 g (50% yield); mp: 205−206 °C; ^{1}H NMR (600 MHz, CD_3OD): δ 7.68 (d, J = 7.8 Hz, 2H), 7.62−7.53 (m, 6H), 7.50−7.41 (m, 5H), 7.34 (dd, J = 12.2 and 7.8 Hz, 3H), 7.22 (d, J = 8.2 Hz, 1H), 6.82 (d, J = 15.8 Hz, 1H), 6.59 (dt, J = 13.4 and 6.5 Hz, 1H), 5.75 (s, 2H), 5.46 (d, J = 6.5 Hz, 2H), 4.09 (s, 3H); ^{13}C NMR (150 MHz, CD_3OD): δ 150.2, 143.5, 141.4, 136.0, 135.9, 134.9, 133.5, 130.0, 129.95, 129.90, 129.8, 129.4, 128.9, 128.8, 128.0, 124.0, 109.3, 106.7, 57.1, 52.8, 51.8; ESI-HRMS (m/z): calcd for $C_{30}H_{26}ON_2Cl$ $[M - Br]^+$ 465.1728, found 465.1722; HPLC analysis: t_R = 5.75 min, purity = 98.9%.

Biological Evalution. *Generation of Expression Clones of S Protein RBD and ACE2 Receptor (aa19−615).* A 672bp DNA fragment corresponding to the S protein RBD (aa residues 319−541) of SARS-CoV-2 and a 1,797bp DNA fragment corresponding to its interacting region of (aa residues 19−615) ACE2 receptor were PCR amplified from pDONR223 SARS-CoV-2 S (Addgene #149329) and pENTR223 ACE2 (Addgene #149719) plasmids, respectively

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

using KAPA HiFi DNA Polymerase (KAPA biosystems). Amplified DNA fragments were purified from a 1% agarose gel and combined with the pDONR221 plasmid (1:3 ratio) in the presence of BP clonase enzyme mix (ThermoFisher). Reaction mixtures were incubated overnight at 25 °C before used for the transformation of competent Mach1 *E. coli* cells. Transformants were grown on LB agar plates supplemented with Kanamycin. Extraction of plasmid DNA was carried out using the NucleoSpin Plasmid Kit (Macherey-Nagel). The identity of generated entry clones was verified by BsrGI restriction digestion. The cDNAs encoding the S protein RBD and ACE2 receptor (aa19−615) (Addgene #167013 and #167012) were shuttled into pcDNA3.1 PA-mCitrine-GW and pcDNA3.1 c-myc-NL-GW destination vectors, respectively[125] in the presence of LR clonase enzyme mix (ThermoFisher). Reaction mixtures were incubated overnight at 25 °C and used for the transformation of Mach1 *E. coli* cells. Transformants were grown on LB agar plates supplemented with 100 μg/mL Ampicillin. Plasmid extraction was carried out using the Nucleospin Plasmid Kit (Macherey-Nagel) and the identity of plasmids was verified by BsrGI restriction digestion.

Immunoblotting. HEK293T cells (2×10^4) were seeded in a 96-well plate and transiently transfected with 0.2 ug plasmid DNA using Polyethylenimine (PEI) (1:3 ratio). Cells were lysed 48 h post transfection in RIPA buffer containing protease and phosphatase inhibitors (Thermofisher) and benzonase (Calbiochem-Novagen). Production of mCitrine-S protein RBD and NL-ACE2 (aa19−615) was verified by Western Blotting using an anti-Spike (69323, CST) and an anti-ACE2 antibody (4355, CST), respectively. GAPDH was used as loading control and was detected with a rabbit anti-GAPDH antibody (5174, CST). Finally, blots were incubated for 2 h with the appropriate secondary alkaline phosphatase-conjugated antibodies and protein bands were visualized with NBT/BCIP (Applichem).

Cytotoxicity Assay. Chemical compounds were dissolved and serially diluted in DMSO. Their cytotoxicity was assessed using an MTT assay. Briefly, HEK293T, SH-SY5Y or Vero E6 cells were cultured in 96-well plates (4×10^4 cells/well) at 37 °C in a 5% CO_2 atmosphere. Cells were treated for 48 h with varying concentrations of the compounds (100−0.1 μM) or solvent. The final concentration of DMSO in culture did not exceed 1%. Treated cells were incubated for 4 h with 0,5 μg/mL MTT (Applichem). After incubation, the MTT-containing medium was removed and formazan crystals were dissolved in DMSO. The optical density (OD) of formazan solution at 570 and 630 nm was measured in a SPARK plate reader (Tecan).

LuTHy Assay for Spike RBD and ACE2 Receptor (aa19−615). The effect of chemical compounds in the interaction of Spike RBD and ACE2 receptor (aa19−615) was assessed using LuTHy assay, as previously described.[125] In brief, HEK293T cells were transfected with pcDNA3.1 PA-mCitrine-S RBD and pcDNA3.1 c-myc-NL- ACE2 receptor (aa19−615). The next day, compounds were added in varying noncytotoxic concentrations and BRET/cBRET signals corresponding to Spike RBD/ACE2 (aa19−615) protein interaction was quantified 48 h later. The nonspecific/autofluorescence signal of each compound was calculated and subtracted from all relevant interactions. For dose-dependent LuTHy assays, HEK293T were treated with decreasing concentrations of compounds for 8 h. The % inhibitory effect of compounds was calculated compared to control cBRET signal in cells treated with the solvent.

SARS-CoV-2 Pseudovirus Assay. The effect of chemical compounds in SH-SY5Y cells was assessed using a pseudoviral assay (Montana Molecular). In brief, cells were seeded in a 96-well plate (2×10^4 cells per well) and treated for 24 h with a transduction mixture containing red ACE2 BacMam (3.3×10^8 VG/mL) and 2 mM sodium butyrate in culture medium. Chemical compounds at a noncytotoxic concentration were preincubated with a transduction mixture containing green SARS-CoV-2 pseudovirus (3.3×10^8 VG/mL) and 2 mM sodium butyrate for 30 min at 37 °C under shaking, before added in cell cultures. After incubation for 24 h, medium was removed, cells were washed with PBS and green/red fluorescence were measured at 488/535 nm and 558/603 nm excitation/emission wavelengths, respectively. The background fluorescence of chemical compounds in green and red wavelengths was measured in treated cells producing only red ACE2 BacMam or uninfected cells, respectively (normalized green and red fluorescence) and was subtracted from the total fluorescence. The inhibitory effect of compounds was calculated by dividing the normalized green fluorescence signal (cells infected with the SARS-CoV-2 pseudovirus) to the normalized red fluorescence signal (viable cells producing BacMam).

SARS-CoV-2 Culture. SARS-CoV-2 (isolate 30−287) was obtained through culture in Vero E6 cells (ATCC CRL-1586) from a COVID-19 patient in Alexandroupolis, Greece. The virus was recovered from a nasopharyngeal swab, after rinsing with 1 mL saline and double filtering through a 0.22 nm filter. Viral stock was prepared by infecting fully confluent Vero E6 cells in DMEM, supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics (penicillin-streptomycin) at 37 °C and 5% CO_2. Four days after inoculation, the supernatant was frozen at −80 °C until use.

Treatment of Vero E6 Cells. Infections were carried out in 24-well plates, using SARS-CoV-2 (M.O.I. of 0.01) on Vero E6 cells. Before cell treatment, the virus was preincubated for 30 min at 37 °C and 5% CO_2 with chemical substances (final concentration ranging from 50 to 0.01 μM) and nirmatrelvir (Paxlovid, Pfizer) (final concentration of 300 nM in DMEM). After the initial incubation, the medium containing the corresponding compound and the virus was used to treat Vero E6 cells for 48 h. As controls, cells were treated only with SARS-CoV-2 or were left untreated for 48 h. Cell morphology was observed by phase contrast in an inverted microscope in order to record cytopathic effect (CPE) after treatment. CPE values were calculated as percentile percentages of the respective values of cells treated only with SARS-CoV-2 and were used for the calculation of half maximal inhibitory concentration (IC_{50}) of chemical substances. For the calculation of half maximal tissue culture infectious dose ($TCID_{50}$), Vero E6 cells were infected with serial dilutions (1/10−$1/10^8$) of supernatants collected from the infection of Vero E6 cells in the presence or absence of chemical substances, followed by incubation at 37 °C and 5% CO_2 for 5 days. Cells incubated only with DMEM were used as control. To determine the SARS-CoV-2 viral load, RNA was extracted from supernatants (100 μL) using NucleoSpin Dx Virus kit (Macherey Nagel). Multitarget real-time PCR was performed using COVID-19 SARS-CoV-2 Real-TM kit according to the manufacturer's instructions (Sacace Biotechnologies).

Statistics and Data Fitting. All assays were performed in technical triplicates. Data are presented as mean ± standard deviation and p-values were calculated in a standard *t* test. For

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

the calculation of the half-maximal inhibitory concentration (IC_{50}), inhibition data were converted to percent inhibition and fitted with standard log inhibitor vs normalized response model, using nonlinear regression in GraphPad Prism 9 software.

Microbial Mutagenicity Assay: Ames Test. The mutagenicity assay for **CKP-25** was conducted following the methodology outlined by Maron and Ames.[126] All required media and the bacterial strain *E. coli* WP2 uvrA were obtained from TrinovaBiochem. The plate incorporation method was used for this assay both in the absence or presence of the liver microsomal metabolic activation (S9) mixture at a concentration of 10% v/v. Bacterial cultures were initially grown in nutrient broth under shaking conditions (150 rpm) at 37 °C for 15 h, until reaching a density of (1−2) × 10^9 bacteria/mL (OD 650 nm approximately 1.0−1.4). For the plate incorporation procedure, 100 μL of the bacterial suspension was mixed with 100 μL of either **CKP-25** or the negative control (DMSO) in 2 mL of tryptophan-supplemented top agar, maintained at 50 °C. This mixture was then poured onto selective agar plates and incubated at 37 °C for 3 days. Five concentrations of **CKP-25** (12.5, 25, 50, 100, and 200 μg/plate) were tested. Positive control substances included 2.5 μL/plate of methylmethanesulfonate (for tests without metabolic activation) and 10 μg/plate of 2-aminoanthracene (for tests with metabolic activation). Each concentration of **CKP-25** and the controls were tested in duplicate. Genotoxicity was determined by assessing the fold increase in the number of revertant colonies. A biologically relevant increase was defined as a 2-fold increase in the mean number of revertant colonies in the *WP2 uvrA* strain, compared to the negative control.

Metabolic Stability. Metabolic stability was evaluated using pooled human liver microsomes at a 0.5 mg/mL concentration. The incubation conditions were optimized to ensure linear formation of metabolites, taking into account both reaction time and protein concentration. A 1 mM concentration of NADPH was used as a cofactor. Each test-compound was assessed at a concentration of 1 μM. To differentiate between low and rapid clearance rates, appropriate negative and positive controls were included. The reactions were conducted in triplicate at a temperature of 37 °C and were terminated after 1 h. The depletion of test compounds was then profiled, and the residual percentage compared to time zero was determined. The readouts were recorded using a Lionheart FX system from Agilent BioTek (Winooski, VT, USA).

Isozyme-Specific CYP450-Metabolism. Human cytochrome P450 (CYP450) isoenzymes were expressed in Baculosomes obtained from Thermo Fisher Scientific (Waltham, MA, USA). Reagents were handled and prepared following the manufacturer's guidelines. Compound **CKP-25** was evaluated at a concentration of 1 μM according to our established standard operating procedures and protocols as described previously.[127,128] The enzymatic activity of various CYP450 isoforms (CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, and CYP3A4) was measured, considering the kinetic model for each isoform. The reactions were conducted in triplicate at 20 °C. NADP+ (10 mM in 100 mM potassium phosphate, pH 8.0) was converted to NADPH using a regeneration system containing glucose-6-phosphate (333 mM) and glucose-6-phosphate dehydrogenase (30 U/mL in 100 mM potassium phosphate, pH 8.0). Upon the addition of the fluorescent substrate, signal monitoring was performed immediately (within 2 min) using appropriate excitation and emission wavelengths, recorded by an Agilent BioTek Lionheart FX (Winooski, VT, USA). CYP450 inhibition (%) was calculated based on the reaction rates (fluorescence intensity changes per unit time). In total, n = 60 meaurements per minute were recorded over a 60 min period.

$$\%\text{Inhibition} = (1 - X/A) \times 100\%$$

where X, the rate observed in the presence of a test-compound; A, the rate observed in the presence of negative (solvent, DMSO) control.

■ ASSOCIATED CONTENT

Data Availability Statement

The source code and the trained model utilized to run the virtual screening can be found in https://github.com/stemylonas/ResNetVS. The model was trained on the DUD-E data set (https://dude.docking.org/). The following third-party software was used. The structures were drawn using ChemDraw 20.0 (PerkinElmer, https://perkinelmerinformatics.com). NMR spectra were. analyzed with MestReNova v14.1.0 (Mestrelab Research; https://mestrelab.com). HPLC chromatograms were analyzed with ChromQuest 5.0. (Thermo Fisher Scientific; https://tools.thermofisher.com). HRMS spectra were analyzed with XCALIBURTM 2.1W/FOUNDATION 1.0.1 (Thermo Fisher Scientific; https://tools.thermofisher.com). Compound filtering: OpenBabel (https://sourceforge.net/projects/openbabel/), PyRx (https://sourceforge.net/projects/pyrx/), PyMol (https://pymol.org/).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.4c01110.

Supplementary Tables S1 and S2, Supplementary Figures S1−S5, ^{1}H NMR, ^{13}C NMR and HR-MS spectra, HPLC traces and Supplementary Charts S1-S4 of the final compounds (PDF)

Molecular formula strings (CSV)

■ AUTHOR INFORMATION

Corresponding Authors

Maria Koufaki − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0000-0001-9351-8279; Email: mkoufa@eie.gr

Theodora Calogeropoulou − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0000-0002-3369-5393; Email: tcalog@eie.gr

Spyros Petrakis − *Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece;* orcid.org/0000-0001-9094-3480; Email: spetrak@certh.gr

Authors

Ioannis Gkekas − *Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece;* orcid.org/0009-0006-7840-6344

Sotirios Katsamakas − *Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece; Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;*

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Present Address: Laboratory of Pharmaceutical Chemistry, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, Campus, Thessaloniki 54124, Greece; orcid.org/0000-0001-8715-0359

Stelios Mylonas − *Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece;* orcid.org/0000-0003-2110-3618

Theano Fotopoulou − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0000-0002-1315-1524

George E. Magoulas − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0000-0001-8944-1635

Alia Cristina Tenchiu − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece*

Marios Dimitriou − *Laboratory of Biology, Department of Medicine, Democritus University of Thrace, Alexandroupolis 68100, Greece;* orcid.org/0000-0001-5699-2289

Apostolos Axenopoulos − *Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece;* orcid.org/0000-0001-8314-4242

Nafsika Rossopoulou − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0009-0001-5755-4341

Simona Kostova − *Max-Delbrueck-Center for Molecular Medicine in the Helmholtz Association, Berlin 13125, Germany*

Erich E. Wanker − *Max-Delbrueck-Center for Molecular Medicine in the Helmholtz Association, Berlin 13125, Germany*

Theodora Katsila − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece;* orcid.org/0000-0002-6263-4231

Demetris Papahatjis − *Institute of Chemical Biology, National Hellenic Research Foundation, Athens 11635, Greece*

Vassilis G. Gorgoulis − *Molecular Carcinogenesis Group, Department of Histology and Embryology, Medical School, National and Kapodistrian University of Athens, Athens 11635, Greece; Ninewells Hospital and Medical School, University of Dundee, DD19SY Dundee, U.K.*

Ioannis Karakasiliotis − *Laboratory of Biology, Department of Medicine, Democritus University of Thrace, Alexandroupolis 68100, Greece*

Petros Daras − *Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki 57001, Greece;* orcid.org/0000-0003-3814-6710

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.4c01110

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. I.G., S.K., and S.M. contributed equally. Conceived and designed the experiments: I.G., S.K., S.M., A.A., V.G.G., D.P., M.K., T.C., S.P.; Performed the experiments: I.G., S.K., S.M., A.C.T., G.E.M., T.F., N.R., and T.K.; Analyzed the data: I.G., S.K., S.M., A.A., N.R., T.K., D.P., M.K., T.C., P.D., S.P.; Contributed material: E.E.W.; Wrote the paper: I.G., S.K., S.M., T.F., G.E.M., T.K., M.K., T.C., S.P.; Edited the manuscript: S.K., M.K., T.C., P.D.; Designed the study: S.K., A.A., M.K., T.C., S.P.

Funding

T.C., M.K., D.P., T.F., G.E.M., and A.C.T. acknowledge financial support of this work by the project by "OPEN-SCREEN-GR: An Open-Access Research Infrastructure of Chemical Biology and Target-Based Screening Technologies for Human and Animal Health Agriculture and the Environment" "(2018−2020)" (MIS) 5002691 which is implemented under the Action "Reinforcement of the Research and Innovation Infrastructure", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014−2020) and cofinanced by Greece and the European Union (European Regional Development Fund). This work was also funded by the National Public Investment Program of the Ministry of Development and Investment/General Secretariat for Research and Technology, in the framework of the Flagship Initiative to address SARS-CoV-2 (2020ΣΕ01300001) and the Operational Program "Competitiveness, Entrepreneurship and Innovation" of the Greek National Strategic Reference Framework 2014-2020 in the context of the "Strengthening the Technology Transfer Unit of CERTH for the Greek Research Area - "SPIRA" Action" with MIS code 5150060. The open access publishing of this article is financially supported by HEAL-Link.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank Dr. Kostas Stamatopoulos for critical suggestions during manuscript preparation.

■ ABBREVIATIONS

ACE2	angiotensin-converting enzyme 2
ADME-Tox	absorption, distribution, metabolism, excretion and toxicity
CCL2	CC chemokine ligand 2
CCR2	CC chemokine receptor 2
CCR5	CC chemokine receptor 5
CNN	convolutional neural network
CPE	cytopathic effect
DPPA	diphenyl phosphorylazide
FCC	Flash column chromatography
Mpro	main protease
PAINS	pan-assay interfering compounds
PK	pharmacokinetic
PPID	protein−protein interaction drug
RBD	receptor binding domain
RBM	receptor binding motif
RdRp	RNA-dependent RNA polymerase
PTSA	p-Toluenesulfonic acid
qPCR	quantitative real-time polymerase chain reaction
SARS-CoV-2	severe-acute-respiratory syndrome coronavirus 2
SBDD	structure-based drug design
TCID50	tissue culture infectious dose
TLC	thin layer chromatography
TMPRSS2	transmembrane serine protease 2
tPSA	topological polar surface area
VS	virtual-screening
WHO	World Health Organization

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

REFERENCES

(1) WHO, *COVID-19 declared a pandemic 11 March 2020*. Accessed online at www.who.int.

(2) https://covid19.who.int/ Accessed online on 2 November 2023.

(3) Kushwaha, N. D.; Mohan, J.; Kushwaha, B.; Ghazi, T.; Nwabuife, J. C.; Koorbanally, N.; Chuturgoon, A. A. A Comprehensive Review on the Global Efforts on Vaccines and Repurposed Drugs for Combating COVID-19. *Eur. J. Med. Chem.* **2023**, *260*, No. 115719.

(4) Valipour, M.; Irannejad, H.; Keyvani, H. An Overview on Anti-COVID-19 Drug Achievements and Challenges Ahead. *ACS Pharmacol. Transl. Sci.* **2023**, *6* (9), 1248−1265.

(5) Worldometer. *COVID-19 pandemic statistical data*. Accessed online at www.worldometers.

(6) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W. Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding. *Lancet* **2020**, *395* (10224), 565−574.

(7) Jackson, C. B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 Entry into Cells. *Nat. Rev. Mol. Cell Biol.* **2022**, *23* (1), 3−20.

(8) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature* **2020**, *581* (7807), 221−224.

(9) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281−292.e6.

(10) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581* (7807), 215−220.

(11) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181* (4), 894−904.e9.

(12) GNS, H. S.; GR, S.; Murahari, M.; Krishnamurthy, M. An Update on Drug Repurposing: Re-Written Saga of the Drug's Fate. *Biomed. Pharmacother.* **2019**, *110*, 700−716.

(13) Rodrigues, L.; Bento Cunha, R.; Vassilevskaia, T.; Viveiros, M.; Cunha, C. Drug Repurposing for COVID-19: A Review and a Novel Strategy to Identify New Targets and Potential Drug Candidates. *Molecules* **2022**, *27* (9), 2723.

(14) Pardo, J.; Shukla, A. M.; Chamarthi, G.; Gupte, A. The Journey of Remdesivir: From Ebola to COVID-19. *Drugs Context* **2020**, *9*, 1−9.

(15) Cave, J. A.; Phizackerley, D. Molnupiravir: Evidence by Press Release. *Drug Ther. Bull.* **2022**, *60* (1), 2−2.

(16) McIntosh, J. A.; Benkovics, T.; Silverman, S. M.; Huffman, M. A.; Kong, J.; Maligres, P. E.; Itoh, T.; Yang, H.; Verma, D.; Pan, W.; Ho, H.-I.; Vroom, J.; Knight, A. M.; Hurtak, J. A.; Klapars, A.; Fryszkowska, A.; Morris, W. J.; Strotman, N. A.; Murphy, G. S.; Maloney, K. M.; Fier, P. S. Engineered Ribosyl-1-Kinase Enables Concise Synthesis of Molnupiravir, an Antiviral for COVID-19. *ACS Cent. Sci.* **2021**, *7* (12), 1980−1985.

(17) Lamb, Y. N. Nirmatrelvir Plus Ritonavir: First Approval. *Drugs* **2022**, *82* (5), 585−591.

(18) Sultana, J.; Crisafulli, S.; Gabbay, F.; Lynn, E.; Shakir, S.; Trifirò, G. Challenges for Drug Repurposing in the COVID-19 Pandemic Era. *Front. Pharmacol.* **2020**, *11*, No. 588654.

(19) Unoh, Y.; Uehara, S.; Nakahara, K.; Nobori, H.; Yamatsu, Y.; Yamamoto, S.; Maruyama, Y.; Taoda, Y.; Kasamatsu, K.; Suto, T.; Kouki, K.; Nakahashi, A.; Kawashima, S.; Sanaki, T.; Toba, S.; Uemura, K.; Mizutare, T.; Ando, S.; Sasaki, M.; Orba, Y.; Sawa, H.; Sato, A.; Sato, T.; Kato, T.; Tachibana, Y. Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *J. Med. Chem.* **2022**, *65* (9), 6499−6512.

(20) https://www.nih.gov/news-events/news-releases/nih-trial-evaluate-shionogi-antiviral-adults-hospitalized-covid-19.

(21) Rahman, M. M.; Masum, M. H. U.; Wajed, S.; Talukder, A. A Comprehensive Review on COVID-19 Vaccines: Development, Effectiveness, Adverse Effects, Distribution and Challenges. *VirusDisease* **2022**, *33* (1), 1−22.

(22) Sun, X.; Yi, C.; Zhu, Y.; Ding, L.; Xia, S.; Chen, X.; Liu, M.; Gu, C.; Lu, X.; Fu, Y.; Chen, S.; Zhang, T.; Zhang, Y.; Yang, Z.; Ma, L.; Gu, W.; Hu, G.; Du, S.; Yan, R.; Fu, W.; Yuan, S.; Qiu, C.; Zhao, C.; Zhang, X.; He, Y.; Qu, A.; Zhou, X.; Li, X.; Wong, G.; Deng, Q.; Zhou, Q.; Lu, H.; Ling, Z.; Ding, J.; Lu, L.; Xu, J.; Xie, Y.; Sun, B. Neutralization Mechanism of a Human Antibody with Pan-Coronavirus Reactivity Including SARS-CoV-2. *Nat. Microbiol.* **2022**, *7* (7), 1063−1074.

(23) Liu, M.; Gan, H.; Liang, Z.; Liu, L.; Liu, Q.; Mai, Y.; Chen, H.; Lei, B.; Yu, S.; Chen, H.; Zheng, P.; Sun, B. Review of Therapeutic Mechanisms and Applications Based on SARS-CoV-2 Neutralizing Antibodies. *Front. Microbiol.* **2023**, *14*, No. 1122868.

(24) Thakur, S.; Sasi, S.; Pillai, S. G.; Nag, A.; Shukla, D.; Singhal, R.; Phalke, S.; Velu, G. S. K. SARS-CoV-2 Mutations and Their Impact on Diagnostics, Therapeutics and Vaccines. *Front. Med.* **2022**, *9*, No. 815389.

(25) Tada, T.; Zhou, H.; Dcosta, B. M.; Samanovic, M. I.; Chivukula, V.; Herati, R. S.; Hubbard, S. R.; Mulligan, M. J.; Landau, N. R. Increased Resistance of SARS-CoV-2 Omicron Variant to Neutralization by Vaccine-Elicited and Therapeutic Antibodies. *eBioMedicine* **2022**, *78*, No. 103944.

(26) Cao, Y.; Yisimayi, A.; Jian, F.; Song, W.; Xiao, T.; Wang, L.; Du, S.; Wang, J.; Li, Q.; Chen, X.; Yu, Y.; Wang, P.; Zhang, Z.; Liu, P.; An, R.; Hao, X.; Wang, Y.; Wang, J.; Feng, R.; Sun, H.; Zhao, L.; Zhang, W.; Zhao, D.; Zheng, J.; Yu, L.; Li, C.; Zhang, N.; Wang, R.; Niu, X.; Yang, S.; Song, X.; Chai, Y.; Hu, Y.; Shi, Y.; Zheng, L.; Li, Z.; Gu, Q.; Shao, F.; Huang, W.; Jin, R.; Shen, Z.; Wang, Y.; Wang, X.; Xiao, J.; Xie, X. S. BA.2.12.1, BA.4 and BA.5 Escape Antibodies Elicited by Omicron Infection. *Nature* **2022**, *608* (7923), 593−602.

(27) Wang, Q.; Guo, Y.; Iketani, S.; Nair, M. S.; Li, Z.; Mohri, H.; Wang, M.; Yu, J.; Bowen, A. D.; Chang, J. Y.; Shah, J. G.; Nguyen, N.; Chen, Z.; Meyers, K.; Yin, M. T.; Sobieszczyk, M. E.; Sheng, Z.; Huang, Y.; Liu, L.; Ho, D. D. Antibody Evasion by SARS-CoV-2 Omicron Subvariants BA.2.12.1, BA.4 and BA.5. *Nature* **2022**, *608* (7923), 603−608.

(28) Mannar, D.; Saville, J. W.; Zhu, X.; Srivastava, S. S.; Berezuk, A. M.; Tuttle, K. S.; Marquez, A. C.; Sekirov, I.; Subramaniam, S. SARS-CoV-2 Omicron Variant: Antibody Evasion and Cryo-EM Structure of Spike Protein−ACE2 Complex. *Science* **2022**, *375* (6582), 760−764.

(29) Cao, Y.; Wang, J.; Jian, F.; Xiao, T.; Song, W.; Yisimayi, A.; Huang, W.; Li, Q.; Wang, P.; An, R.; Wang, J.; Wang, Y.; Niu, X.; Yang, S.; Liang, H.; Sun, H.; Li, T.; Yu, Y.; Cui, Q.; Liu, S.; Yang, X.; Du, S.; Zhang, Z.; Hao, X.; Shao, F.; Jin, R.; Wang, X.; Xiao, J.; Wang, Y.; Xie, X. S. Omicron Escapes the Majority of Existing SARS-CoV-2 Neutralizing Antibodies. *Nature* **2022**, *602* (7898), 657−663.

(30) Lopez Bernal, J.; Andrews, N.; Gower, C.; Gallagher, E.; Simmons, R.; Thelwall, S.; Stowe, J.; Tessier, E.; Groves, N.; Dabrera, G.; Myers, R.; Campbell, C. N. J.; Amirthalingam, G.; Edmunds, M.; Zambon, M.; Brown, K. E.; Hopkins, S.; Chand, M.; Ramsay, M. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N. Engl. J. Med.* **2021**, *385* (7), 585−594.

(31) Zhuang, C.; Liu, X.; Chen, Q.; Sun, Y.; Su, Y.; Huang, S.; Wu, T.; Xia, N. Protection Duration of COVID-19 Vaccines: Waning Effectiveness and Future Perspective. *Front. Microbiol.* **2022**, *13*, No. 828806.

(32) Marcec, R.; Majta, M.; Likic, R. Will Vaccination Refusal Prolong the War on SARS-CoV-2? *Postgrad. Med. J.* **2021**, *97* (1145), 143−149.

(33) Rodrigues, J. P. G. L. M.; Barrera-Vilarmau, S.; Teixeira, M. C. J.; Sorokina, M.; Seckel, E.; Kastritis, P. L.; Levitt, M. Insights on

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Cross-Species Transmission of SARS-CoV-2 from Structural Modeling. *PLoS Comput. Biol.* **2020**, *16* (12), No. e1008449.
(34) Bojadzic, D.; Alcazar, O.; Chen, J.; Chuang, S.-T.; Condor Capcha, J. M.; Shehadeh, L. A.; Buchwald, P. Small-Molecule Inhibitors of the Coronavirus Spike: ACE2 Protein−Protein Interaction as Blockers of Viral Attachment and Entry for SARS-CoV-2. *ACS Infect. Dis.* **2021**, *7* (6), 1519−1534.
(35) Zhang, D.; Hamdoun, S.; Chen, R.; Yang, L.; Ip, C. K.; Qu, Y.; Li, R.; Jiang, H.; Yang, Z.; Chung, S. K.; Liu, L.; Wong, V. K. W. Identification of Natural Compounds as SARS-CoV-2 Entry Inhibitors by Molecular Docking-Based Virtual Screening with Bio-Layer Interferometry. *Pharmacol. Res.* **2021**, *172*, No. 105820.
(36) Chen, C. Z.; Xu, M.; Pradhan, M.; Gorshkov, K.; Petersen, J. D.; Straus, M. R.; Zhu, W.; Shinn, P.; Guo, H.; Shen, M.; Klumpp-Thomas, C.; Michael, S. G.; Zimmerberg, J.; Zheng, W.; Whittaker, G. R. Identifying SARS-CoV-2 Entry Inhibitors through Drug Repurposing Screens of SARS-S and MERS-S Pseudotyped Particles. *ACS Pharmacol. Transl. Sci.* **2020**, *3* (6), 1165−1175.
(37) Sun, C.; Zhang, J.; Wei, J.; Zheng, X.; Zhao, X.; Fang, Z.; Xu, D.; Yuan, H.; Liu, Y. Screening, Simulation, and Optimization Design of Small Molecule Inhibitors of the SARS-CoV-2 Spike Glycoprotein. *PLoS One* **2021**, *16* (1), No. e0245975.
(38) Xiu, S.; Dick, A.; Ju, H.; Mirzaie, S.; Abdi, F.; Cocklin, S.; Zhan, P.; Liu, X. Inhibitors of SARS-CoV-2 Entry: Current and Future Opportunities. *J. Med. Chem.* **2020**, *63* (21), 12256−12274.
(39) Wang, C.; He, F.; Sun, K.; Guo, K.; Lu, S.; Wu, T.; Gao, X.; Fang, M. Identification and Characterization of 7-Azaindole Derivatives as Inhibitors of the SARS-CoV-2 Spike-HACE2 Protein Interaction. *Int. J. Biol. Macromol.* **2023**, *244*, No. 125182.
(40) He, Z.; Yuan, J.; Zhang, Y.; Li, R.; Mo, M.; Wang, Y.; Ti, H. Recent Advances towards Natural Plants as Potential Inhibitors of SARS-Cov-2 Targets. *Pharm. Biol.* **2023**, *61* (1), 1186−1210.
(41) Carvalhal, F.; Magalhães, A. C.; Rebelo, R.; Palmeira, A.; Resende, D. I. S. P.; Durães, F.; Maia, M.; Xavier, C. P. R.; Pereira, L.; Sousa, E.; Correia-da-Silva, M.; Vasconcelos, M. H. Evaluation of the Cytotoxic and Antiviral Effects of Small Molecules Selected by In Silico Studies as Inhibitors of SARS-CoV-2 Cell Entry. *Molecules* **2023**, *28* (20), 7204.
(42) Palla, S. R.; Li, C.-W.; Chao, T.-L.; Lo, H.-L. V.; Liu, J.-J.; Pan, M. Y.-C.; Chiu, Y.-T.; Lin, W.-C.; Hu, C.-W.; Yang, C.-M.; Chen, Y.-Y.; Fang, J.-T.; Lin, S.-W.; Lin, Y.-T.; Lin, H.-C.; Kuo, C.-J.; Wang, L. H.-C.; Chang, S.-Y.; Liang, P.-H. Synthesis, Evaluation, and Mechanism of 1-(4-(Arylethylenylcarbonyl)Phenyl)-4-Carboxy-2-Pyrrolidinones as Potent Reversible SARS-CoV-2 Entry Inhibitors. *Antiviral Res.* **2023**, *219*, No. 105735.
(43) Thakkar, R.; Agarwal, D. K.; Ranaweera, C. B.; Ishiguro, S.; Conda-Sheridan, M.; Gaudreault, N. N.; Richt, J. A.; Tamura, M.; Comer, J. *De Novo* Design of a Stapled Peptide Targeting SARS-CoV-2 Spike Protein Receptor-Binding Domain. *RSC Med. Chem.* **2023**, *14* (9), 1722−1733.
(44) Rodríguez, Y.; Cardoze, S. M.; Obineche, O. W.; Melo, C.; Persaud, A.; Fernández Romero, J. A. Small Molecules Targeting SARS-CoV-2 Spike Glycoprotein Receptor-Binding Domain. *ACS Omega* **2022**, *7* (33), 28779−28789.
(45) Lü, X.; Feng, C.; Lü, R.; Wei, X.; Fan, S.; Yan, M.; Zhu, X.; Zhang, Z.; Yang, Z. Identification of Potential Inhibitors of Omicron Variant of SARS-Cov-2 RBD Based Virtual Screening, MD Simulation, and DFT. *Front. Chem.* **2022**, *10*, No. 1063374.
(46) Carino, A.; Moraca, F.; Fiorillo, B.; Marchianò, S.; Sepe, V.; Biagioli, M.; Finamore, C.; Bozza, S.; Francisci, D.; Distrutti, E.; Catalanotti, B.; Zampella, A.; Fiorucci, S. Hijacking SARS-CoV-2/ACE2 Receptor Interaction by Natural and Semi-Synthetic Steroidal Agents Acting on Functional Pockets on the Receptor Binding Domain. *Front. Chem.* **2020**, *8*, No. 572885.
(47) Sandholtz, S. H.; Drocco, J. A.; Zemla, A. T.; Torres, M. W.; Silva, M. S.; Allen, J. E. A Computational Pipeline to Identify and Characterize Binding Sites and Interacting Chemotypes in SARS-CoV-2. *ACS Omega* **2023**, *8* (24), 21871−21884.
(48) Wang, Q.; Meng, F.; Xie, Y.; Wang, W.; Meng, Y.; Li, L.; Liu, T.; Qi, J.; Ni, X.; Zheng, S.; Huang, J.; Huang, N. In Silico Discovery of Small Molecule Modulators Targeting the Achilles' Heel of SARS-CoV-2 Spike Protein. *ACS Cent. Sci.* **2023**, *9* (2), 252−265.
(49) Lapaillerie, D.; Charlier, C.; Guyonnet-Dupérat, V.; Murigneux, E.; Fernandes, H. S.; Martins, F. G.; Magalhães, R. P.; Vieira, T. F.; Richetta, C.; Subra, F.; Lebourgeois, S.; Charpentier, C.; Descamps, D.; Visseaux, B.; Weigel, P.; Favereaux, A.; Beauvineau, C.; Buron, F.; Teulade-Fichou, M.-P.; Routier, S.; Gallois-Montbrun, S.; Meertens, L.; Delelis, O.; Sousa, S. F.; Parissi, V. Selection of Bis-Indolyl Pyridines and Triphenylamines as New Inhibitors of SARS-CoV-2 Cellular Entry by Modulating the Spike Protein/ACE2 Interfaces. *Antimicrob. Agents Chemother.* **2022**, *66* (8), No. e0008322.
(50) Buchwald, P. Developing Small-Molecule Inhibitors of Protein-Protein Interactions Involved in Viral Entry as Potential Antivirals for COVID-19. *Front. Drug Discovery* **2022**, *2*, No. 898035.
(51) Khan, A.; Randhawa, A. W.; Balouch, A. R.; Mukhtar, N.; Sayaf, A. M.; Suleman, M.; Khan, T.; Ali, S.; Ali, S. S.; Wang, Y.; Mohammad, A.; Wei, D.-Q. Blocking Key Mutated Hotspot Residues in the RBD of the Omicron Variant (B.1.1.529) with Medicinal Compounds to Disrupt the RBD-HACE2 Complex Using Molecular Screening and Simulation Approaches. *RSC Adv.* **2022**, *12* (12), 7318−7327.
(52) Pomplun, S.; Jbara, M.; Quartararo, A. J.; Zhang, G.; Brown, J. S.; Lee, Y.-C.; Ye, X.; Hanna, S.; Pentelute, B. L. *De Novo* Discovery of High-Affinity Peptide Binders for the SARS-CoV-2 Spike Protein. *ACS Cent. Sci.* **2021**, *7* (1), 156−163.
(53) Marković, V.; Shaik, J. B.; Ożga, K.; Ciesiołkiewicz, A.; Lizandra Perez, J.; Rudzińska-Szostak, E.; Berlicki, Ł. Peptide Foldamer-Based Inhibitors of the SARS-CoV-2 S Protein−Human ACE2 Interaction. *J. Enzyme Inhib. Med. Chem.* **2023**, *38* (1), No. 2244693.
(54) Xiong, J.; Xiang, Y.; Huang, Z.; Liu, X.; Wang, M.; Ge, G.; Chen, H.; Xu, J.; Zheng, M.; Chen, L. Structure-Based Virtual Screening and Identification of Potential Inhibitors of SARS-CoV-2 S-RBD and ACE2 Interaction. *Front. Chem.* **2021**, *9*, No. 740702.
(55) Wang, L.; Wu, Y.; Yao, S.; Ge, H.; Zhu, Y.; Chen, K.; Chen, W.; Zhang, Y.; Zhu, W.; Wang, H.; Guo, Y.; Ma, P.; Ren, P.; Zhang, X.; Li, H.; Ali, M. A.; Xu, W.; Jiang, H.; Zhang, L.; Zhu, L.; Ye, Y.; Shang, W.; Bai, F. Discovery of Potential Small Molecular SARS-CoV-2 Entry Blockers Targeting the Spike Protein. *Acta Pharmacol. Sin.* **2022**, *43* (4), 788−796.
(56) Acharya, A.; Pandey, K.; Thurman, M.; Klug, E.; Trivedi, J.; Sharma, K.; Lorson, C. L.; Singh, K.; Byrareddy, S. N. Discovery and Evaluation of Entry Inhibitors for SARS-CoV-2 and Its Emerging Variants. *J. Virol.* **2021**, *95* (24), No. e0143721.
(57) Hu, X.; Chen, C. Z.; Xu, M.; Hu, Z.; Guo, H.; Itkin, Z.; Shinn, P.; Ivin, P.; Leek, M.; Liang, T. J.; Shen, M.; Zheng, W.; Hall, M. D. Discovery of Small Molecule Entry Inhibitors Targeting the Fusion Peptide of SARS-CoV-2 Spike Protein. *ACS Med. Chem. Lett.* **2021**, *12* (8), 1267−1274.
(58) Sun, H.; Wang, Y.; Chen, C. Z.; Xu, M.; Guo, H.; Itkin, M.; Zheng, W.; Shen, M. Identification of SARS-CoV-2 Viral Entry Inhibitors Using Machine Learning and Cell-Based Pseudotyped Particle Assay. *Bioorg. Med. Chem.* **2021**, *38*, No. 116119.
(59) Meanwell, N. A.; Krystal, M. R.; Nowicka-Sans, B.; Langley, D. R.; Conlon, D. A.; Eastgate, M. D.; Grasela, D. M.; Timmins, P.; Wang, T.; Kadow, J. F. Inhibitors of HIV-1 Attachment: The Discovery and Development of Temsavir and Its Prodrug Fostemsavir. *J. Med. Chem.* **2018**, *61* (1), 62−80.
(60) Mylonas, S.; Axenopoulos, A.; Katsamakas, S.; Gkekas, I.; Stamatopoulos, K.; Petrakis, S.; Daras, P. Deep Learning-Assisted Pipeline for Virtual Screening of Ligand Compound Databases: Application on Inhibiting the Entry of SARS-CoV-2 into Human Cells. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*; IEEE: 2020; pp 132−139.
(61) Koes, D. R.; Baumgartner, M. P.; Camacho, C. J. Lessons Learned in Empirical Scoring with Smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf. Model.* **2013**, *53* (8), 1893−1904.

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(62) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3* (1), 33.
(63) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55* (14), 6582−6594.
(64) Axenopoulos, A.; Rafailidis, D.; Papadopoulos, G.; Houstis, E. N.; Daras, P. Similarity Search of Flexible 3D Molecules Combining Local and Global Shape Descriptors. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **2016**, *13* (5), 954−970.
(65) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23* (1−3), 3−25.
(66) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45* (12), 2615−2623.
(67) Przybylak, K. R.; Alzahrani, A. R.; Cronin, M. T. D. How Does the Quality of Phospholipidosis Data Influence the Predictivity of Structural Alerts? *J. Chem. Inf. Model.* **2014**, *54* (8), 2224−2232.
(68) Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y. Physiochemical Drug Properties Associated with in Vivo Toxicological Outcomes. *Bioorg. Med. Chem. Lett.* **2008**, *18* (17), 4872−4875.
(69) Bruns, R. F.; Watson, I. A. Rules for Identifying Potentially Reactive or Promiscuous Compounds. *J. Med. Chem.* **2012**, *55* (22), 9763−9772.
(70) Lagorce, D.; Bouslama, L.; Becot, J.; Miteva, M. A.; Villoutreix, B. O. FAF-Drugs4: Free ADME-Tox Filtering Computations for Chemical Biology and Early Stages Drug Discovery. *Bioinformatics* **2017**, *33* (22), 3658−3660.
(71) Sperandio, O.; Reynès, C. H.; Camproux, A.-C.; Villoutreix, B. O. Rationalizing the Chemical Space of Protein−Protein Interaction Inhibitors. *Drug Discovery Today* **2010**, *15* (5−6), 220−229.
(72) Shin, W.-H.; Kumazawa, K.; Imai, K.; Hirokawa, T.; Kihara, D. Current Challenges and Opportunities in Designing Protein−Protein Interaction Targeted Drugs. *Adv. Appl. Bioinforma. Chem.* **2020**, *13*, 11−25.
(73) Basha, N. J. Therapeutic Efficacy of Benzimidazole and Its Analogs: An Update. *Polycycl. Aromat. Compd.* **2023**, *43* (7), 6549−6569.
(74) Mahurkar, N. D.; Gawhale, N. D.; Lokhande, M. N.; Uke, S. J.; Kodape, M. M. Benzimidazole: A Versatile Scaffold for Drug Discovery and beyond−A Comprehensive Review of Synthetic Approaches and Recent Advancements in Medicinal Chemistry. *Results Chem.* **2023**, *6*, No. 101139.
(75) Omotuyi, O.; Olatunji, O. M.; Nash, O.; Oyinloye, B.; Soremekun, O.; Ijagbuji, A.; Fatumo, S. Benzimidazole Compound Abrogates SARS-COV-2 Receptor-Binding Domain (RBD)/ACE2 Interaction In Vitro. *Microb. Pathog.* **2023**, *176*, No. 105994.
(76) Mudi, P. K.; Mahanty, A. K.; Kotakonda, M.; Prasad, S.; Bhattacharyya, S.; Biswas, B. A Benzimidazole Scaffold as a Promising Inhibitor against SARS-CoV-2. *J. Biomol. Struct. Dyn.* **2023**, *41* (5), 1798−1810.
(77) Mudi, P. K.; Mahato, R. K.; Verma, H.; Panda, S. J.; Purohit, C. S.; Silakari, O.; Biswas, B. In Silico Anti-SARS-CoV-2 Activities of Five-Membered Heterocycle-Substituted Benzimidazoles. *J. Mol. Struct.* **2022**, *1261*, No. 132869.
(78) Han, P.; Su, C.; Zhang, Y.; Bai, C.; Zheng, A.; Qiao, C.; Wang, Q.; Niu, S.; Chen, Q.; Zhang, Y.; Li, W.; Liao, H.; Li, J.; Zhang, Z.; Cho, H.; Yang, M.; Rong, X.; Hu, Y.; Huang, N.; Yan, J.; Wang, Q.; Zhao, X.; Gao, G. F.; Qi, J. Molecular Insights into Receptor Binding of Recent Emerging SARS-CoV-2 Variants. *Nat. Commun.* **2021**, *12* (1), 6103.
(79) Han, P.; Li, L.; Liu, S.; Wang, Q.; Zhang, D.; Xu, Z.; Han, P.; Li, X.; Peng, Q.; Su, C.; Huang, B.; Li, D.; Zhang, R.; Tian, M.; Fu, L.; Gao, Y.; Zhao, X.; Liu, K.; Qi, J.; Gao, G. F.; Wang, P. Receptor Binding and Complex Structures of Human ACE2 to Spike RBD from Omicron and Delta SARS-CoV-2. *Cell* **2022**, *185* (4), 630−640.e10.
(80) Saville, J. W.; Mannar, D.; Zhu, X.; Srivastava, S. S.; Berezuk, A. M.; Demers, J.-P.; Zhou, S.; Tuttle, K. S.; Sekirov, I.; Kim, A.; Li, W.; Dimitrov, D. S.; Subramaniam, S. Structural and Biochemical Rationale for Enhanced Spike Protein Fitness in Delta and Kappa SARS-CoV-2 Variants. *Nat. Commun.* **2022**, *13* (1), 742.
(81) Geng, Q.; Shi, K.; Ye, G.; Zhang, W.; Aihara, H.; Li, F. Structural Basis for Human Receptor Recognition by SARS-CoV-2 Omicron Variant BA.1. *J. Virol.* **2022**, *96* (8), No. e0024922.
(82) Xu, Y.; Wu, C.; Cao, X.; Gu, C.; Liu, H.; Jiang, M.; Wang, X.; Yuan, Q.; Wu, K.; Liu, J.; Wang, D.; He, X.; Wang, X.; Deng, S.-J.; Xu, H. E.; Yin, W. Structural and Biochemical Mechanism for Increased Infectivity and Immune Evasion of Omicron BA.2 Variant Compared to BA.1 and Their Possible Mouse Origins. *Cell Res.* **2022**, *32* (7), 609−620.
(83) Tuekprakhon, A.; Nutalai, R.; Dijokaite-Guraliuc, A.; Zhou, D.; Ginn, H. M.; Selvaraj, M.; Liu, C.; Mentzer, A. J.; Supasa, P.; Duyvesteyn, H. M. E.; Das, R.; Skelly, D.; Ritter, T. G.; Amini, A.; Bibi, S.; Adele, S.; Johnson, S. A.; Constantinides, B.; Webster, H.; Temperton, N.; Klenerman, P.; Barnes, E.; Dunachie, S. J.; Crook, D.; Pollard, A. J.; Lambe, T.; Goulder, P.; Paterson, N. G.; Williams, M. A.; Hall, D. R.; Fry, E. E.; Huo, J.; Mongkolsapaya, J.; Ren, J.; Stuart, D. I.; Screaton, G. R.; Conlon, C.; Deeks, A.; Frater, J.; Frending, L.; Gardiner, S.; Jämsén, A.; Jeffery, K.; Malone, T.; Phillips, E.; Rothwell, L.; Stafford, L. Antibody Escape of SARS-CoV-2 Omicron BA.4 and BA.5 from Vaccine and BA.1 Serum. *Cell* **2022**, *185* (14), 2422−2433.e13.
(84) Li, L.; Liao, H.; Meng, Y.; Li, W.; Han, P.; Liu, K.; Wang, Q.; Li, D.; Zhang, Y.; Wang, L.; Fan, Z.; Zhang, Y.; Wang, Q.; Zhao, X.; Sun, Y.; Huang, N.; Qi, J.; Gao, G. F. Structural Basis of Human ACE2 Higher Binding Affinity to Currently Circulating Omicron SARS-CoV-2 Sub-Variants BA.2 and BA.1.1. *Cell* **2022**, *185* (16), 2952−2960.e10.
(85) Howell, J.; Rasmussen, M. Heterocyclic Ambident Nucleophiles. V. Alkylation of Benzimidazoles. *Aust. J. Chem.* **1993**, *46* (8), 1177.
(86) Ochiai, K.; Takita, S.; Eiraku, T.; Kojima, A.; Iwase, K.; Kishi, T.; Fukuchi, K.; Yasue, T.; Adams, D. R.; Allcock, R. W.; Jiang, Z.; Kohno, Y. Phosphodiesterase Inhibitors. Part 3: Design, Synthesis and Structure−Activity Relationships of Dual PDE3/4-Inhibitory Fused Bicyclic Heteroaromatic-Dihydropyridazinones with Anti-Inflammatory and Bronchodilatory Activity. *Bioorg. Med. Chem.* **2012**, *20* (5), 1644−1658.
(87) Koufaki, M.; Fotopoulou, T.; Kapetanou, M.; Heropoulos, G. A.; Gonos, E. S.; Chondrogianni, N. Microwave-Assisted Synthesis of 3,5-Disubstituted Isoxazoles and Evaluation of Their Anti-Ageing Activity. *Eur. J. Med. Chem.* **2014**, *83*, 508−515.
(88) Haque, R. A.; Choo, S. Y.; Budagumpi, S.; Iqbal, M. A.; Al-Ashraf Abdullah, A. Silver(I) Complexes of Mono- and Bidentate N-Heterocyclic Carbene Ligands: Synthesis, Crystal Structures, and in Vitro Antibacterial and Anticancer Studies. *Eur. J. Med. Chem.* **2015**, *90*, 82−92.
(89) Ho, T.; Wu, S.; Chen, J.; Li, C.; Hsiang, C. Emodin Blocks the SARS Coronavirus Spike Protein and Angiotensin-Converting Enzyme 2 Interaction. *Antiviral Res.* **2007**, *74* (2), 92−101.
(90) Dellafiora, L.; Dorne, J. L. C. M.; Galaverna, G.; Dall'Asta, C. Preventing the Interaction between Coronaviruses Spike Protein and Angiotensin I Converting Enzyme 2: An In Silico Mechanistic Case Study on Emodin as a Potential Model Compound. *Appl. Sci.* **2020**, *10* (18), 6358.
(91) Li, C.; Zhou, H.; Guo, L.; Xie, D.; He, H.; Zhang, H.; Liu, Y.; Peng, L.; Zheng, L.; Lu, W.; Mei, Y.; Liu, Z.; Huang, J.; Wang, M.; Shu, D.; Ding, L.; Lang, Y.; Luo, F.; Wang, J.; Huang, B.; Huang, P.; Gao, S.; Chen, J.; Qian, C.-N. Potential Inhibitors for Blocking the

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Interaction of the Coronavirus SARS-CoV-2 Spike Protein and Its Host Cell Receptor ACE2. *J. Transl. Med.* **2022**, *20* (1), 314.
(92) Petrakis, S.; Raskó, T.; Russ, J.; Friedrich, R. P.; Stroedicke, M.; Riechers, S. P.; Muehlenberg, K.; Möller, A.; Reinhardt, A.; Vinayagam, A.; Schaefer, M. H.; Boutros, M.; Tricoire, H.; Andrade-Navarro, M. A.; Wanker, E. E. Identification of Human Proteins That Modify Misfolding and Proteotoxicity of Pathogenic Ataxin-1. *PLoS Genet.* **2012**, *8*, No. e1002897.
(93) Delaney, J. S. ESOL: Estimating Aqueous Solubility Directly from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 1000−1005.
(94) Jain, N.; Yalkowsky, S. H. Estimation of the Aqueous Solubility I: Application to Organic Nonelectrolytes. *J. Pharm. Sci.* **2001**, *90* (2), 234−252.
(95) Cheng, T.; Zhao, Y.; Li, X.; Lin, F.; Xu, Y.; Zhang, X.; Li, Y.; Wang, R.; Lai, L. Computation of Octanol−Water Partition Coefficients by Guiding an Additive Model with Knowledge. *J. Chem. Inf. Model.* **2007**, *47* (6), 2140−2148.
(96) Barter, Z.; Bayliss, M.; Beaune, P.; Boobis, A.; Carlile, D.; Edwards, R.; Brian Houston, J.; Lake, B.; Lipscomb, J.; Pelkonen, O.; Tucke, G.; Rostami-Hodjegan, A. Scaling Factors for the Extrapolation of In Vivo Metabolic Drug Clearance From In Vitro Data: Reaching a Consensus on Values of Human Micro-Somal Protein and Hepatocellularity Per Gram of Liver. *Curr. Drug Metab.* **2007**, *8* (1), 33−45.
(97) Masimirembwa, C. M.; Bredberg, U.; Andersson, T. B. Metabolic Stability for Drug Discovery and Development. *Clin. Pharmacokinet.* **2003**, *42* (6), 515−528.
(98) EMA. *Guideline on the Investigation of Drug Interactions*; Guid Doc: London, UK, 2013.
(99) EMA. *Guideline on the Investigation of Drug Interactions*; Guid Doc: London, UK, 2012.
(100) FDA. *Drug Development and Drug Interactions*; Guid Doc: London, UK, 2020.
(101) McGinnity, D. F.; Parker, A. J.; Soars, M.; Riley, R. J. Automated Definition of the Enzymology of Drug Oxidation by the Major Human Drug Metabolizing Cytochrome P450s. *Drug Metab. Dispos.* **2000**, *28* (11), 1327−1334.
(102) Pelkonen, O.; Turpeinen, M. In Vitro-in Vivo Extrapolation of Hepatic Clearance: Biological Tools, Scaling Factors, Model Assumptions and Correct Concentrations. *Xenobiotica* **2007**, *37* (10−11), 1066−1089.
(103) Zanger, U. M.; Schwab, M. Cytochrome P450 Enzymes in Drug Metabolism: Regulation of Gene Expression, Enzyme Activities, and Impact of Genetic Variation. *Pharmacol. Ther.* **2013**, *138* (1), 103−141.
(104) Irwin, J. J.; Shoichet, B. K. ZINC−a Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177−182.
(105) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41* (5), 1308−1315.
(106) Oprea, T. I. Property Distribution of Drug-Related Chemical Databases. *J. Comput. Aided. Mol. Des.* **2000**, *14* (3), 251−264.
(107) Park, B. K.; Boobis, A.; Clarke, S.; Goldring, C. E. P.; Jones, D.; Kenna, J. G.; Lambert, C.; Laverty, H. G.; Naisbitt, D. J.; Nelson, S.; Nicoll-Griffith, D. A.; Obach, R. S.; Routledge, P.; Smith, D. A.; Tweedie, D. J.; Vermeulen, N.; Williams, D. P.; Wilson, I. D.; Baillie, T. A. Managing the Challenge of Chemically Reactive Metabolites in Drug Development. *Nat. Rev. Drug Discovery* **2011**, *10* (4), 292−306.
(108) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53* (7), 2719−2740.
(109) Mannhold, R.; Poda, G. I.; Ostermann, C.; Tetko, I. V. Calculation of Molecular Lipophilicity: State-of-the-Art and Comparison of LogP Methods on More than 96,000 Compounds. *J. Pharm. Sci.* **2009**, *98* (3), 861−893.
(110) Kohlbacher, S.; Schmid, M.; Seidel, T.; Langer, T. Applications of the Novel Quantitative Pharmacophore Activity Relationship Method QPhAR in Virtual Screening and Lead-Optimisation. *Pharmaceuticals* **2022**, *15* (9), 1122.
(111) Amarasinghe, K. N.; De Maria, L.; Tyrchan, C.; Eriksson, L. A.; Sadowski, J.; Petrović, D. Virtual Screening Expands the Non-Natural Amino Acid Palette for Peptide Optimization. *J. Chem. Inf. Model.* **2022**, *62* (12), 2999−3007.
(112) Dallakyan, S.; Olson, A. J. Small-Molecule Library Screening by Docking with PyRx. *Methods Mol. Biol.* **2015**, *1263*, 243−250.
(113) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455−461.
(114) Eberhardt, J.; Santos-Martins, D.; Tillack, A. F.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, *61* (8), 3891−3898.
(115) Allinger, N. L. Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V1 and V2 Torsional Terms. *J. Am. Chem. Soc.* **1977**, *99* (25), 8127−8134.
(116) Egan, W. J.; Merz; Kenneth, M.; Baldwin, J. J. Prediction of Drug Absorption Using Multivariate Statistics. *J. Med. Chem.* **2000**, *43* (21), 3867−3877.
(117) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(118) Wacha, A. F.; Lemkul, J. A. Charmm2gmx: An Automated Method to Port the CHARMM Additive Force Field to GROMACS. *J. Chem. Inf. Model.* **2023**, *63* (14), 4246−4252.
(119) https://plasma-gate.weizmann.ac.il/Grace/.
(120) http://www.ks.uiuc.edu/Research/vmd/.
(121) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33−38.
(122) Li, H.; Zhang, Y.; Yan, Z.; Lai, Z.; Yang, R.; Peng, M.; Sun, Y.; An, J. Methanol as the C1 Source: Redox Coupling of Nitrobenzenes and Alcohols for the Synthesis of Benzimidazoles. *Green Chem.* **2022**, *24* (2), 748−753.
(123) Tessier, J.; Schmitzer, A. R. Benzimidazolium Salts Prevent and Disrupt Methicillin-Resistant *Staphylococcus Aureus* Biofilms. *RSC Adv.* **2020**, *10* (16), 9420−9430.
(124) Iddon, B.; Petersen, A. K.; Becher, J.; Christensen, N. J. Azoles. Part II. Synthesis of Imidazole-2(and −5)-Carbaldehydes and Derivatives of Imidazo[1,2-b]Isoquinoline; Transmetallation of Imidazol-5-Yllithium Compounds. *J. Chem. Soc. Perkin Trans. 1* **1995**, *12*, 1475.
(125) Trepte, P.; Kruse, S.; Kostova, S.; Hoffmann, S.; Buntru, A.; Tempelmeier, A.; Secker, C.; Diez, L.; Schulz, A.; Klockmeier, K.; Zenkner, M.; Golusik, S.; Rau, K.; Schnoegl, S.; Garner, C. C.; Wanker, E. E. LuTHy: A Double-readout Bioluminescence-based Two-hybrid Technology for Quantitative Mapping of Protein−Protein Interactions in Mammalian Cells. *Mol. Syst. Biol.* **2018**, *14* (7), No. e8071.
(126) Maron, D. M.; Ames, B. N. Revised Methods for the Salmonella Mutagenicity Test. *Mutat. Res. Mutagen. Relat. Subj.* **1983**, *113* (3−4), 173−215.
(127) Yilmaz, C.; Rogdakis, T.; Latorrata, A.; Thanou, E.; Karadima, E.; Papadimitriou, E.; Siapi, E.; Li, K. W.; Katsila, T.; Calogeropoulou, T.; Charalampopoulos, I.; Alexaki, V. I. ENT-A010, a Novel Steroid Derivative, Displays Neuroprotective Functions and Modulates Microglial Responses. *Biomolecules* **2022**, *12* (3), 424.
(128) Rogdakis, T.; Charou, D.; Latorrata, A.; Papadimitriou, E.; Tsengenes, A.; Athanasiou, C.; Papadopoulou, M.; Chalikiopoulou, C.; Katsila, T.; Ramos, I.; Prousis, K. C.; Wade, R. C.; Sidiropoulou, K.; Calogeropoulou, T.; Gravanis, A.; Charalampopoulos, I. Development and Biological Characterization of a Novel Selective TrkA Agonist with Neuroprotective Properties against Amyloid Toxicity. *Biomedicines* **2022**, *10* (3), 614.

https://doi.org/10.1021/acs.jcim.4c01110
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX