JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

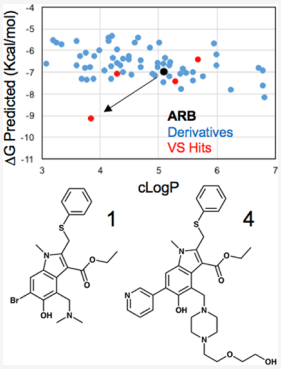
Modeling the Structure−Activity Relationship of Arbidol Derivatives and Other SARS-CoV-2 Fusion Inhibitors Targeting the S2 Segment of the Spike Protein

Matthew R. Freidel and Roger S. Armen*

Cite This: J. Chem. Inf. Model. 2021, 61, 5906−5922 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Umifenovir (Arbidol) has been reported to exhibit some degree of efficacy in multiple clinical trials for the treatment of COVID-19 as a monotherapy. It has also demonstrated synergistic inhibition of SARS-CoV-2 with other direct-acting antivirals such as Remdesivir. A computational approach was used to identify the most favorable binding site to the SARS-CoV-2 Spike S2 segment and to perform virtual screening. Compounds selected from modeling were evaluated in a live SARS-CoV-2 infection assay. An Arbidol (ARB) derivative with substitutions at both the C-4 and C-6 positions was found to exhibit a modest improvement in activity and solubility properties in comparison to ARB. However, all of the derivatives were found to only be partial inhibitors, rather than full inhibitors in a virus-induced cytopathic effect-based assay. The binding mode is also corroborated by parallel modeling of a series of oleanolic acid trisaccharide saponin fusion inhibitors shown to bind to the S2 segment. Recently determined experimental structures of the Spike protein allowed atomic resolution modeling of fusion inhibitor binding as a function of pH, and the implications for the molecular mechanism of direct-acting fusion inhibitors targeting the S2 segment are discussed.



1. INTRODUCTION

The 2019 emergence of a novel Coronavirus in Wuhan, China, highlighted the need to identify drugs with a known tolerability profile and clinical data, which showed efficacy in inhibiting coronaviruses. Even with previous coronavirus outbreaks, such as that of SARS in 2003 and MERS in 2012, there were not particularly effective antiviral agents that could be used clinically when the 2019 SARS-CoV-2 outbreak occurred.[1] Drug repurposing remains a critical way to help combat the current outbreak by providing clinical options for patients who present with severe cases of COVID-19 but may be unable to receive first-line treatment. Clinical drugs identified to exhibit antiviral activity for SARS, MERS, or SARS-CoV-2 may reasonably be expected to have activity against new emerging viruses from the nidovirales family.[2] In addition, repurposing drugs that are already approved for other indications allows clinicians to use them with a greater body of literature regarding their safety profile. One potential avenue that has not been as deeply explored in clinical research has been the use of antiviral medications which act as entry inhibitors by preventing fusion via the Spike protein. Small-molecule fusion inhibitors may also be candidates for future antiviral combination therapies using Remdesivir or other antiviral replication (Nsp12) inhibitors or in combination with (MPro) protease inhibitors.

Arbidol (ARB) is a commonly used drug in Russia and China for both prophylaxis and treatment of influenza A and B, as well as other respiratory virus infections. ARB exhibits antiviral activity against the influenza A/Aichi2/68 (H3N2) reference prototype strain that is sensitive to amantadine,[3] as well as highly pathogenic influenza A H5N1 subtype[4,5] and the pandemic 2009 H1N1 subtype.[6−9] ARB has also been shown to demonstrate weak in vitro inhibitory activity for several other enveloped viruses.[10−13] ARB is considered to be a broad-spectrum antiviral that has been used clinically in the past for outbreaks of unknown acute respiratory diseases. Several clinical trials in Russia[14−17] and China[18] corroborate efficacy and very good tolerability of treatment as a monotherapy, as well as in combination with interferons.[19−21] In summary, acute treatments of ARB ranging from 5 to 20 days have been shown to be well tolerated with minimal adverse events.[22−25]

Prior to the outbreak in Wuhan, ARB had been shown to display antiviral activity against the 2003 SARS-CoV virus;[26] thus, there was prior rationale for using it as a clinical agent for treatment. Early in the pandemic, the in vitro antiviral activity of ARB against SARS-CoV-2 was published[27] and then verified[28] supporting rationale for clinical usage to treat COVID-19, particularly in combination with other direct-acting antivirals.[29,30] The clinical efficacy of ARB for the treatment of COVID-19 remains the subject of controversy and conflicting clinical results with relatively low numbers of

Received: September 2, 2021
Published: December 13, 2021

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

enrolled patients. One of the earliest published small clinical trials in China reported improvements in outcomes in COVID-19 treatment with ARB as a monotherapy.[31] The subsequent published clinical data have been varied, with several studies showing improved outcomes,[32−35] while other studies have shown decreased efficacy.[36]

As ARB exhibits weak broad-spectrum activity, solution-nuclear magnetic resonance experiments have previously demonstrated that ARB binds to the surface of membranes via direct interactions with the phospholipid headgroup and the glycerol backbone.[37] Thus, nonspecific interactions (in the 10 μM range) with both membranes and aromatic amino acids in membranes may partially explain inhibition of membrane fusion and broad-spectrum fusion inhibitor activity and the local concentration of drug in proximity to fusion events.[37] However, in the case of influenza, it has been shown that ARB binds directly to hemagglutinin (HA2) from several complementary experimental approaches including biophysical experiments,[38] a mass-spectrometry proteomics approach,[39] and X-ray crystallography of the ARB bound cocrystal complex.[40] Thus, in the case of influenza, drug binding directly to the HA2 subunit is an important contribution to direct drug action, in addition to other mechanisms proposed to inhibit membrane fusion associated with membrane partitioning effects. In vitro studies of ARB-resistant influenza strains[41] demonstrate that ARB increases the stability of the HA and hinders low pH-mediated conformational changes requisite for formation of a conformation that is more favorable for membrane fusion.[41] Following this rationale and the report that ARB was an inhibitor of SARS-CoV-2,[27] we hypothesized that ARB may exhibit direct action as a fusion inhibitor by binding to the S2 segment of the Spike protein. Our previous work using a pharmacophore procedure to map SARS-CoV-2 drug targets identified the most favorable binding site for ARB (Figure 1)

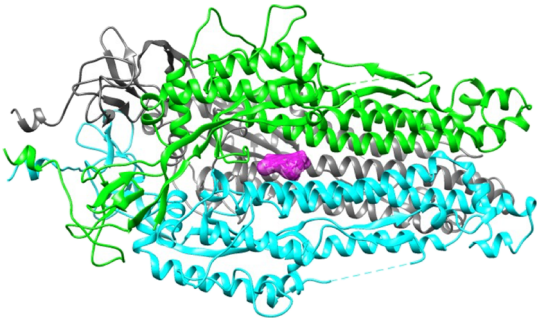


Figure 1. Structure of the SARS-CoV-2 Spike S2 segment with the Arbidol binding site predicted from pharmacophore mapping and CHARMM-based molecular docking. ARB is shown in magenta as a molecular surface bound to the trimeric S2 segment in a prefusion conformation.

on the S2 segment.[42] This site was also corroborated by structural knowledge using a structural alignment of the influenza HA2 subunit bound to ARB and the SARS-CoV-2 Spike protein S2 segment[42] shown in Figure 2. These binding sites exhibit minimal sequence similarity and only remote structural similarity in positioning and location of functional charged residues (Asp, Glu, His, Lys, and Arg) that form salt bridges and participate in pH-mediated conformational changes requisite for membrane fusion. This model of ARB bound to the SARS-CoV-2 binding site (July 2020) was utilized for virtual screening and selection of compounds

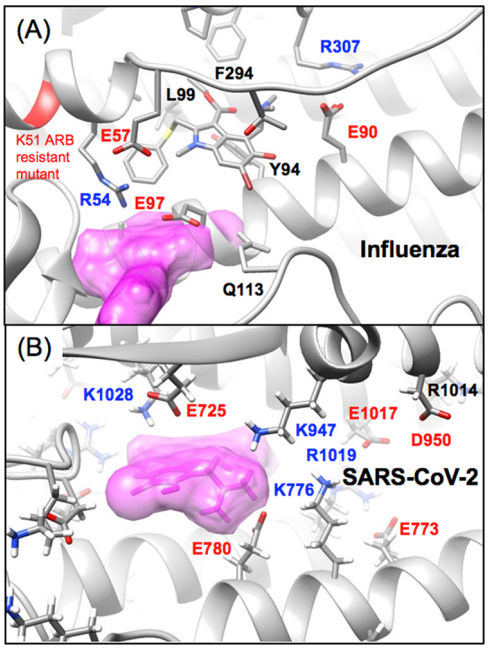


Figure 2. Structure of the predicted S2 segment Arbidol binding site compared to the experimental Influenza HA structure. (A) Experimental structure of ARB bound to the Influenza HA2 subunit (56TS.pdb) showing functional glutamic acid residues E57, E90, and E97 where the basic amine compound ARB binds in proximity of these residues. The magenta shown represents the real atom density from the Arbidol site in the superimposed SARS-CoV-2 S2 segment ARB binding site. (B) Predicted Arbidol binding site on the S2 segment shown highlighting analogous glutamic acid residues (E725, E773, and E780) where the basic amine compound ARB binds in proximity of these residues.

predicted to bind with a higher affinity to the S2 binding site (Figure 2B). Compounds selected (December 2020) were evaluated in a live SARS-CoV-2 infection assay (March 2021). Subsequent rounds of modeling were performed to additionally characterize the predicted structural implications of the structure−activity relationship (SAR) data and compared to other literature-reported fusion inhibitors targeting the S2 segment.

2. METHODS

2.1. Computational Methods. All molecular docking and free energy calculations were performed using the program CHARMM[43] and previously described methods.[44,45] The CHARMM-based methods used were shown to have a high "discriminative power" to correctly predict binding geometries over diverse classes of protein−ligand interactions compared to other common scoring functions.[44] Molecular docking utilized the LPDB CHARMm force field for modeling small-molecule potential functions and protein−ligand interactions.[46,47]

A 3-D grid is used to describe the static conformation of an individual protein binding site, where the interaction energy of 20 types of probe atoms is calculated at every point on a grid of 1.0 Å granularity. All-atom flexible models of a small molecule then interact with the potential energy of the grid. The docking approach includes a series of independent docking "trials" that are composed of a large number of individual docking attempts. Each docking "trial" is initiated from either 10 to 40 nonidentical geometric conformations of the input ligand geometry as described below. For each initiated "trial",

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

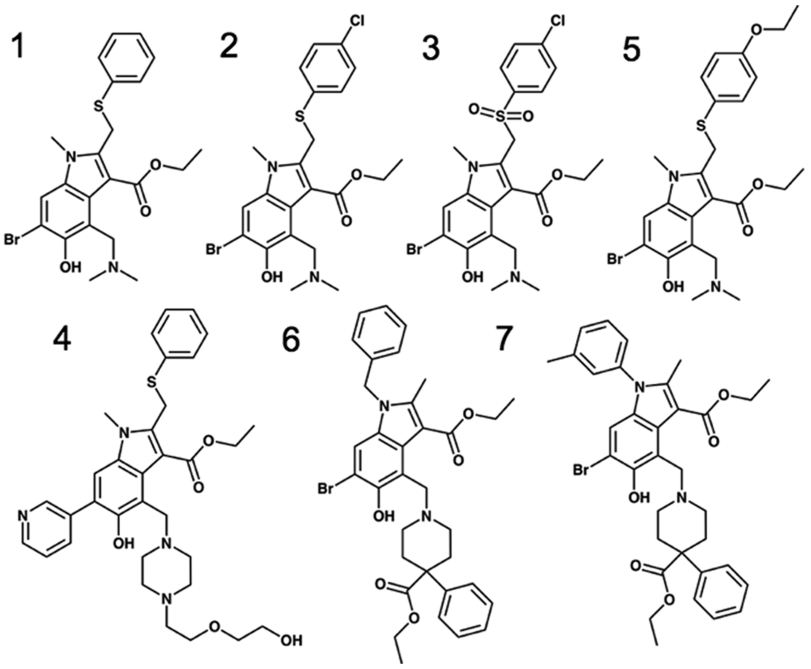


Figure 3. Structure of Arbidol derivatives used in this study.

thousands of individual docking attempts are performed, and a two-step scoring approach is utilized to rank the final TOP5 poses from any docking attempt.[44,45] First, the potential energy of the all-atom complex is minimized for 1000 steps using the standard hard-core repulsion potentials for both the VDW and electrostatic interaction and a distance-dependent dielectric function. Using a scoring function derived from these potential energies,[44] thousands of poses are evaluated, and then only the final TOP5 docking poses for a single docking attempt are then subjected to a second, much more computationally expensive scoring step.

For evaluation of the TOP5 docking poses, a final energy minimization of the protein−ligand complex is performed using one of the most accurate CHARMM-based implicit solvent models, the generalized Born using molecular volume method.[48,49] Starting from the minimized complex, minimizations of the bound and free states are performed where potential energy components (VDW and ELEC) and solvation (SOLV) are calculated to approximate the free energy of binding (ΔG_{bind}) using a linear interaction energy scoring approach with previously determined empirical generalized parameters for $\alpha = 0.20$ and $\beta = 0.05$ using the formalism for $\Delta G_{bind} = [\alpha(\Delta VDW) + \beta(\Delta(ELEC + SOLV)]$.[44] Thus, the scoring strategy utilized here represents an a priori physics-based approach, not a scoring approach that has been optimized using a multivariable optimization with experimental data for this specific protein target structure or compound series examined here. Results for comparing to SAR data are also compared to previous critical assessment benchmarks for acceptable levels of agreement between datasets of (predicted) and (observed) binding or activity data and this exact CHARMM-based scoring function.[44,45] Results using the predicted ΔG_{bind} values for the TOP5 poses of each individual docking "trial" are pooled and sorted by ΔG_{bind}, where the top-ranked members of a geometric cluster (RMSD <2.0 Å) are identified. Statistics for ΔG_{bind} are calculated from the average and standard deviation from the three top-ranked members of a geometric cluster (RMSD <2.0 Å) or a triplicate representing the geometric cluster.

For all the work performed in this study, independent docking "trials" were initiated from either 10 or 40 generated conformations of a given small-molecule ligand, such that the initial geometry is entirely independent of any CHARMM-based procedure. MarvinSketch version 15.8.31 is a publicly available 3D conformation generator that was used to generate nonidentical low-energy conformations.[50] Thus, for most docking performed in this study at a single binding site (as described below), 40 randomly generated conformations are used to initiate 40 independent docking "trials" that evaluate 1000 molecular docking "attempts" each. Thus, the final TOP5 conformations from 40 independent trials (200 low-energy complexes) are then used to identify the top-ranked members of a geometric cluster (RMSD <2.0 Å). For representative calculations, analysis of the entire energy distribution of calculated ΔG_{bind} values, calculated from these 200 low-energy complexes, are shown in the Results section to demonstrate statistical significance using standard calculated Z scores (Z_{score}) and associated P-values at the 99% confidence intervals for specific compounds binding to one site compared to other alternative sites.[51,52]

Our laboratory previously used a pharmacophore procedure to identify the most favorable TOP50 binding sites for an aromatic pharmacophore on the SARS-CoV-2 Spike protein S2 segment.[42] This structural analysis had been performed using the 3.2 Å CryoEM structure (6vyb.pdb) of the full-length Spike protein and where the ectodomain was in the "closed" state (6vxx.pdb).[53] From this model (6vxx.pdb), molecular docking was performed in a hierarchical approach, such that 10 conformations of ARB were initially docked to all 50 sites on the Spike S2 segment. Then after identification of the TOP5 most favorable sites from this first step, more extensive sampling was used to refine the ranking of the TOP5 sites, and 40 conformations of ARB and other derivatives were docked to the TOP5 sites on the Spike S2 segment. From this model (6vxx.pdb), the consensus binding mode for ARB, **1** binding to

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

Table 1. Compound Antiviral Activity and Predicted Physiochemical Properties

	compound information			SARS-CoV-2 CPE			cytotoxicity
	vendor	vendor ID	cLogP	% inhibition at 7.5 μM	EC50 (μM)	CC50 (μM)	safety ratio
1	ChemBridge	5,674,919	5.1	17.5	5.6	13.1	2.3
2	ChemBridge	5,137,758	5.7				
3	ChemBridge	8,885,134	4.3	11.7	29.5	>25	
4	ChemDiv	H027−0070	3.9	23.0	4.3	17.7	4.1
5	Pharmeks	PHAR008065	5.3				
6	Vitas M	STK689502	7.3				
7	Vitas M	STK688615	7.7				

the SARS-CoV-2 Spike S2 segment (Figure 2B) was used for virtual screening with 78 available chemical derivatives of ARB. Available derivatives of interest for virtual screening were identified by searching PUBCHEM using a substructure search with the following smiles string query:

CN1C(C)=C(C(O)=O)C2=C(CN)C=CC=C12

From screening 78 derivatives, six chemical derivatives were selected for experimental characterization shown in Figure 3 and Table 1. Data describing the chemical structures (smiles strings), physiochemical properties, and predicted ΔG_{bind} values for all 79 compounds evaluated and ranked by virtual screening data are available in Table S1 Supporting Information.

To better understand the structural implications of the experimental data from ARB derivatives, this study was extended to a series of oleanolic acid (OA) trisaccharide saponin derivatives[54] that are fusion inhibitors shown to directly bind to the S2 segment by a biophysical surface plasmon resonance (SPR) binding assay. OA saponin derivative **12a** was selected as a reference molecule to dock to all 50 sites, as it had one less flexible degree of freedom in the R-group in comparison to **12f**. Saponin derivative **12a** for example has a total of eight rotatable bonds (N_{rot} = 8), where the trisaccharide groups have a larger number of flexible degrees of freedom (N_{rot} = 5) compared to the larger hydrophobic OA group (N_{rot} = 2). In a hierarchical approach, in the first round of docking only the hydrophobic OA substituent was docked, leaving a hydroxyl (OH) group at the position of attachment and thus removing the trisaccharide group (without trisaccharide). Using this approach, the majority of binding modes were found to bury the hydrophobic surface of the OA group of the saponins as reasonably expected, leaving the OH group exposed to the solvent. Unphysical binding modes that buried the OH group in hydrophobic surfaces were discarded. In this way, 10 conformations of the saponin **12a** (without trisaccharide, N_{rot} = 2) were initially docked to all 50 sites on the Spike S2 segment. Then after identification of the TOP5 most favorable sites from this first step, a more extensive sampling strategy was used to refine the ranking of the TOP5 sites. Using the entire correct covalent structure of saponin **12a** (with trisaccharide, N_{rot} = 8), 40 conformations were docked to the TOP5 sites on the Spike S2 segment. As there was good agreement (RMSD < 2.0 Å) between the top-ranked geometries of saponin **12a** (without trisaccharide) initiated from 10 random conformations and saponin **12a** (with trisaccharide) initiated from 40 random conformations at the TOP2 ranked sites, docking studies were initiated from 10 random conformations of saponin derivatives (without trisaccharide) to compare to experimental activity data. OA saponin derivatives[54] (**12a**, **12b**, **12c**, **12d**, **12e**, **12f**, **12i**, **12j**, **12k**, **12l**, **12m**, **12n**, **12o**, and **12p**) were modeled in this way at the TOP2 binding sites on the S2 segment where the calculated ΔG_{bind} values are analyzed for correlation with experimental log IC50 values. Pearson's R and (R^2) values are calculated as well as root-mean-squared error (RMSE).

The robustness of this correlation analysis was examined using a cross-validation approach. Given 14 saponin derivatives modeled for comparison of predicted ones to those observed, a "leave 4 out" strategy was used where 14 random datasets are developed and used for cross-validation analysis. With these cross-validation datasets, linear correlation analysis is performed for only a "test" set of 10 derivatives, while four random derivatives are removed for a "validation" set. While this cross-validation may be exhaustively enumerated from a thousand datasets, it may also be appropriately modeled from averaging over an "unbiased" selection of 14 datasets. This was accomplished by creating 14 datasets where the first saponin derivative N is withheld (of 4) for each of the N "test" sets, such that each "test" set starts by removing one systematic compound. Then N datasets are randomly selected in composition, such that three additional random derivatives are removed from each "test" set leaving 4 in each "validation" set, under the constraint that of the 14 constructed datasets, an individual derivative is only removed 4 times, for all 14 derivatives. This strategy samples in an unbiased way to evenly sample the effects of removing individual derivatives. The randomly selected datasets constructed to satisfy this constraint and used for analysis are shown in Supporting Information Table S2.

All fusion inhibitors, both ARB and saponin derivatives, were initially modeled using (6vxx.pdb) as a reference neutral pH conformation of the Spike protein. These studies were extended using the experimentally determined structures at pH 7.0, pH 5.5 (6xmo.pdb), and pH 4.5 (7jwy.pdb).[55] Using these structures, CHARMM-based molecular docking methods were used to calculate the approximate trend in ΔG_{bind} as a function of pH for the compounds using standard CHARMM-based potential functions for the charged and neutral states of titratable residues of His, Asp, and Glu. In particular, for (6vxx.pdb) modeled at neutral pH, standard neutral pH residues were used for the neutral histidine and ionized (ASP and GLU) residues. For modeling (6xmo.pdb) at pH 5.5, standard protonated and ionized (HSP) and ionized (ASP and GLU) residues were used, while modeling (7jwy.pdb)[55] at pH 4.5, standard ionized (HSP) and neutral (ASPH and GLUH) residues were used to model the unionized neutral end-point of a pH titration of individual Asp and Glu residues, which more realistically models the pH range of 3.5−4.0. Statistics for ΔG_{bind} as a function of pH are also calculated from the average and standard deviation of the top-three ranked members of a geometric cluster (RMSD < 2.0 Å). For illustration of the

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

approximate trend of predicted ΔG_{bind} as a function of pH, sigmoidal functions were calculated with a five-parameter logistic equation[56] to best fit the predicted ΔG_{bind} values, using the predicted value from the pH 4.5 structure as an end point of the titration. All molecular graphics images of protein structures and protein−ligand interactions are generated with UCSF Chimera.[57]

2.2. Experimental Methods. Antiviral assays against live SARS-CoV-2 were performed by RetroVirox using the MEX-BC2/2020 strain [GISAID database ID: EPI_ISL_747242], which contains the D614G mutation in the Spike protein. A virus-induced cytopathic effect (CPE) antiviral assay was performed by infecting Vero E6 cells in the presence or absence of inhibitors. In this assay, as infection of cells leads to a significant CPE and cell death after 4 days of infection, a reduction of CPE in the presence of inhibitors is used as a surrogate marker for antiviral activity. Cell viability assays to determine inhibitor loss of cell viability were also monitored in parallel using the same readout (neutral red), but utilizing uninfected cells incubated with the compounds.

Vero E6 cells were maintained in DMEM with 10% fetal bovine serum (FBS), which is also abbreviated as (DMEM10). Twenty-four hours after cell seeding, test samples were subjected to serial dilutions with DMEM2 in a different plate. Then, media were removed from cells, and serial dilutions of inhibitors were added to the cells and incubated for 1 h at 37 °C in a humidified incubator. After cells were preincubated with inhibitors, then cultures were challenged with SARS-CoV-2 resuspended in DMEM with 2% FBS (DMEM2). The amount of viral inoculum was previously titrated to result in a linear response inhibited by antivirals with known activity against SARS-CoV-2. Cell culture media with the virus inoculum were not removed after virus adsorption, and inhibitors and virus were maintained in the media for the duration of the assay (96 h). After 96 h, the extent of cell viability was monitored with the neutral red (NR) uptake assay. Cells were stained with NR, where viable cells incorporate NR in their lysosomes. After a 3 h incubation with NR (0.017%), the extra dye is washed away, and the NR is extracted from lysosomes by incubating cells for 15 minutes with a solution containing 50% ethanol and 1% acetic acid. The amount of NR is estimated by measuring absorbance at 540 nm. The average absorbance at 540 nm (A540) observed in infected cells (in the presence of vehicle alone) was calculated and then subtracted from all samples to determine the inhibition of the virus-induced CPE. Data points were then normalized to the average A540 signal observed in uninfected cells ("mock") after subtraction of the absorbance signal observed in infected cells. In the NR CPE-based assay, uninfected cells remained viable and take up the dye at higher levels than nonviable cells. In the absence of antiviral agents, the virus-induced CPE kills infected cells and leads to lower A540 (this value equals 0% inhibition). In contrast, incubation with the positive control antiviral agent (GS-441524) prevents the virus-induced CPE and leads to absorbance levels similar to those observed in uninfected cells. GS-441524 is the main metabolite of Remdesivir, a broad-spectrum antiviral that is a potent inhibitor of the SARS-CoV-2 RNA polymerase.[58,59] Full recovery of cell viability in infected cells represents 100% inhibition of virus replication. Parallel cell viability assays were performed to assess inhibitor-induced cytotoxicity of uninfected cells. Absorbance readout values were given as a percentage of the average signal observed in uninfected cells treated with vehicle alone. The average signal obtained in wells with no cells was subtracted from all samples. Readout values were given as a percentage of the average signal observed in uninfected cells treated with vehicle alone. The signal-to-background (S/B) ratio obtained in two separate plates was 8.1-fold and 7.2-fold. DMSO was used as a cytotoxic compound control in the viability assays. DMSO blocked cell viability by more than 98% when tested at 10%.

All of the compounds were initially tested in duplicate at 5 μM and 25 μM, where controls included uninfected cells ("mock-infected)," and infected cells to which only vehicle was added. Cells were also treated with positive control antiviral GS-441524 (in a full dose−response curve on the same plate). Subsequent full dose−response studies of **1**, **3**, and **4** were performed in duplicates on a separate day, using serial 2-fold dilutions of eight concentration points starting at 15 μM. Controls included uninfected cells ("mock-infected)," and infected cells to which only vehicle was added. When possible, EC50 (antiviral) and CC50 (inhibition of viability) values of the compounds were determined by fitting the concentration series data to a sigmoidal function according to global nonlinear fit solutions to the five-parameter logistic equation.[56]

Quality controls for the infectivity assays were performed on each plate to determine: S/B values, inhibition by known inhibitors of SARS-CoV-2, and variation of the assay, as measured by the coefficient of variation of all data points. All controls worked as anticipated for each assay. GS-441524, a known inhibitor of SARS-CoV-2 infection, prevented completely the virus-induced CPE of the infected cells. The EC50 values obtained for GS-441524 were 0.91 and ∼0.27 μM. Overall variation of duplicates in the antiviral assay was 5.0 and 4.8% across two plates. Overall variation in the viability assays was 3.0 and 1.4%. The *S/B* ratio for the antiviral assay was 5.4-fold and 3.4-fold for plates 1 and 2, determined by comparing the A540nm values in uninfected ("mock") cells with that observed in cells challenged with SARS-CoV-2 in the presence of vehicle alone. When comparing the signal in uninfected cells to the signal in "no-cells" background wells, the S/B ratio of the antiviral assay was 10.0-fold and 9.4-fold. For the viability assay, the *S/B* ratio ("no cells" value) was 8.1-fold and 7.2-fold.

3. RESULTS AND DISCUSSION

3.1. Virtual Screening of Available Derivatives. Available derivatives of interest for virtual screening were identified by searching PUBCHEM[60] using a substructure search query, where the 2D chemical substructure is shown in Figure 4A. This substructure query strategy allows structural modifications but requires the compounds to contain a C-4 (methylamino), a C-3 carboxylic acid, and a 1,2-dimethyl group. This query is able to provide a limited screening dataset aimed particularly at sampling available structural modifications at the C-4 position that require the exact topological structure of the basic amine of ARB. Virtual screening was performed at Site 1 (Figure 1) where the set of 78 available chemical derivatives are compared to ARB by molecular docking and subsequent ranking of predicted ΔG_{bind}. Data describing the chemical structures (smiles strings), physiochemical properties, and predicted ΔG_{bind} values for all 79 compounds ranked by virtual screening data are available in Table S1 Supporting Information. From screening 78 derivatives, six chemical derivatives were selected for experimental characterization shown in Figure 3 and Table 1. ARB derivatives **3**, **4**, **5**, **6**, and **7** were selected as virtual

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

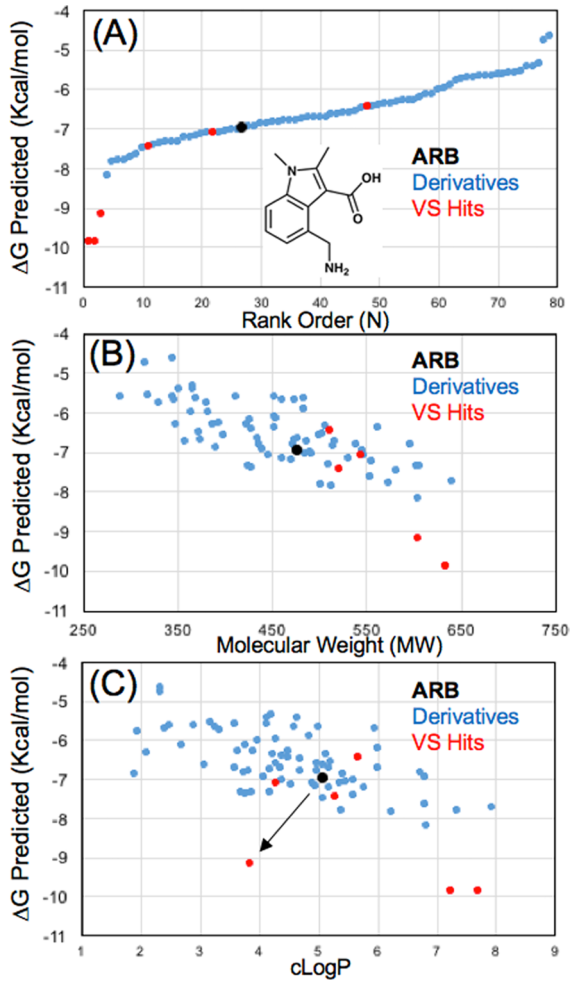


Figure 4. Virtual screening data for Arbidol derivatives. (A) A ranked hit list of sorted predicted ΔG_{bind} values where reference ARB is shown in black, VS hits are shown in red, and the rest of the derivatives are shown in blue. (B) Predicted ΔG_{bind} values as a function of MW. (C) Predicted ΔG_{bind} values as a function of cLogP. Derivative 4 is highlighted as a favorable VS hit with a lower cLogP value than the majority of the derivatives examined.

screening hits where ΔG_{bind} was predicted to be more favorable than ARB. Interestingly, in the ranked list of 79 compounds, ARB was ranked as 27th of 79, such that 26 derivatives were predicted to be more favorable than ARB (Figure 4A). In addition to the favorable predicted ΔG_{bind} values, derivatives **4**, **6**, and **7** were of significant interest as they exhibited more favorable C-4 substituents and were also shown to dock in the same binding mode (RMSD <2.0 Å, for the atoms of the substructure query) compared to the reference model of ARB. Derivatives **6** and **7** contained larger aromatic substituents attached to the indole nitrogen, where **4** retains the methyl group of ARB. Derivatives **3** and **5** were also found to have more favorable substituents and were also shown to dock in the same binding mode (RMSD < 2.0 Å) compared to ARB. Derivative **2** was initially predicted to be less favorable than ARB from the virtual screening data but was selected for a more straightforward experimental comparison between the (p-chloro) substituents of **2** and **3** compared to **1**. One reason why the predicted (ΔG_{bind}) values of **4**, **6**, and **7** are larger than those of other compounds in the dataset is due to effects from "artificial enrichment" because of the higher molecular weights (MWs) of these compounds compared to the rest of the dataset (Figure 4B).[44,61] Despite this concern, which is also a result of the limitation of the available chemical space, **4**, **6**, and **7** were still selected for characterization. Derivative **4** was of interest for reasons of experimental solubility as it had a much lower cLogP compared to all of the VS hits (Figure 4C).

3.2. Compound Solubility and Physiochemical Properties. Six chemical derivatives were selected for experimental characterization using a CPE-based antiviral assay (Figure 3). When these compounds were solubilized in either DMSO or buffer (for salts), cmps **5**, **6**, and **7** exhibited poor aqueous solubility. Cmps **5** and **6** exhibited visible precipitation while **7** exhibited a milky coloration when solubilized at 10 mM in DMSO. Microphotographs of infected cell monolayers were also able to visualize the extent of compound precipitation or recrystallization at 25 μM during the 96 h incubation experiments as shown in Figure 5, where **6** showed the greatest extent of recrystallization during the incubation (Figure 5J). These observations made it clear that cmps **5**, **6**, and **7** exhibited unacceptable aqueous solubility for successful

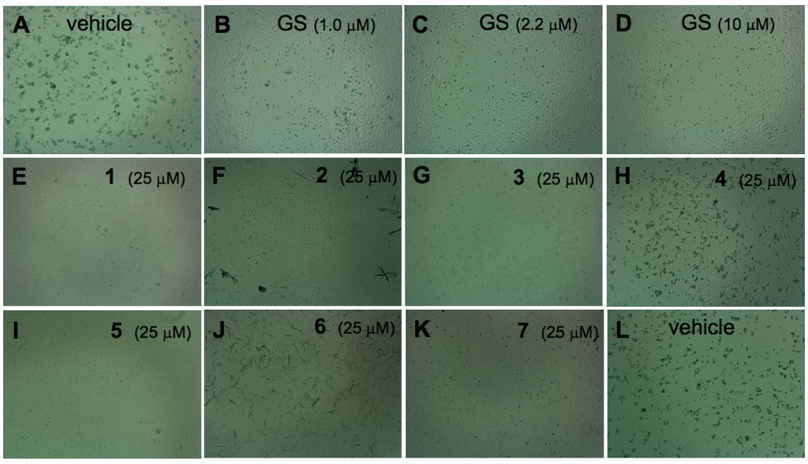


Figure 5. Microscopic evaluation of screening data at 25 μM. Images from infected cells (A) infection in the presence of vehicle alone; (B) through (D) infection with 1.0, 2.2, and 10 μM GS-441524; (E) infection with 1 (25 μM); (F) infection with 2 (25 μM); (G) infection with 3 (25 μM); (H) infection with 4 (25 μM); (I) infection with 5 (25 μM); (J) infection with 6 (25 μM); (K) infection with 7 (25 μM); and (L) infection in the presence of vehicle alone (duplicate).

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

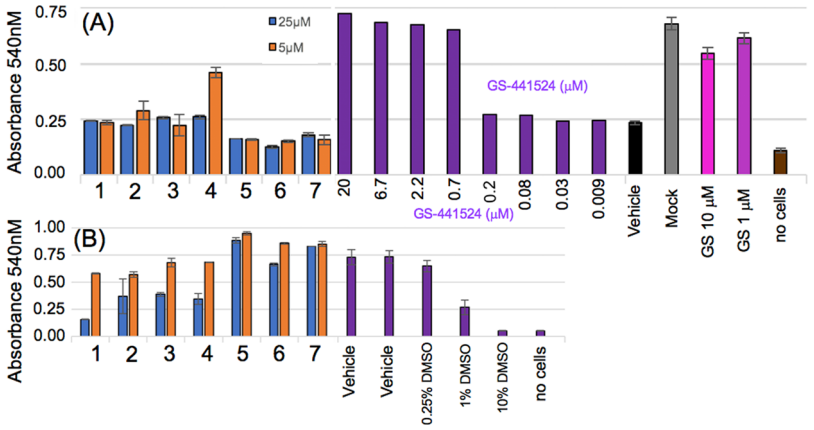


Figure 6. Screening data at two concentrations 5 and 25 μM against SARS-CoV-2. (A) Antiviral activity of 1−7 in duplicate compared to GS-441524. Error bars shown are standard deviation values from duplicate measurements. (B) Cell viability in uninfected cells compared to 1 or 10% DMSO.

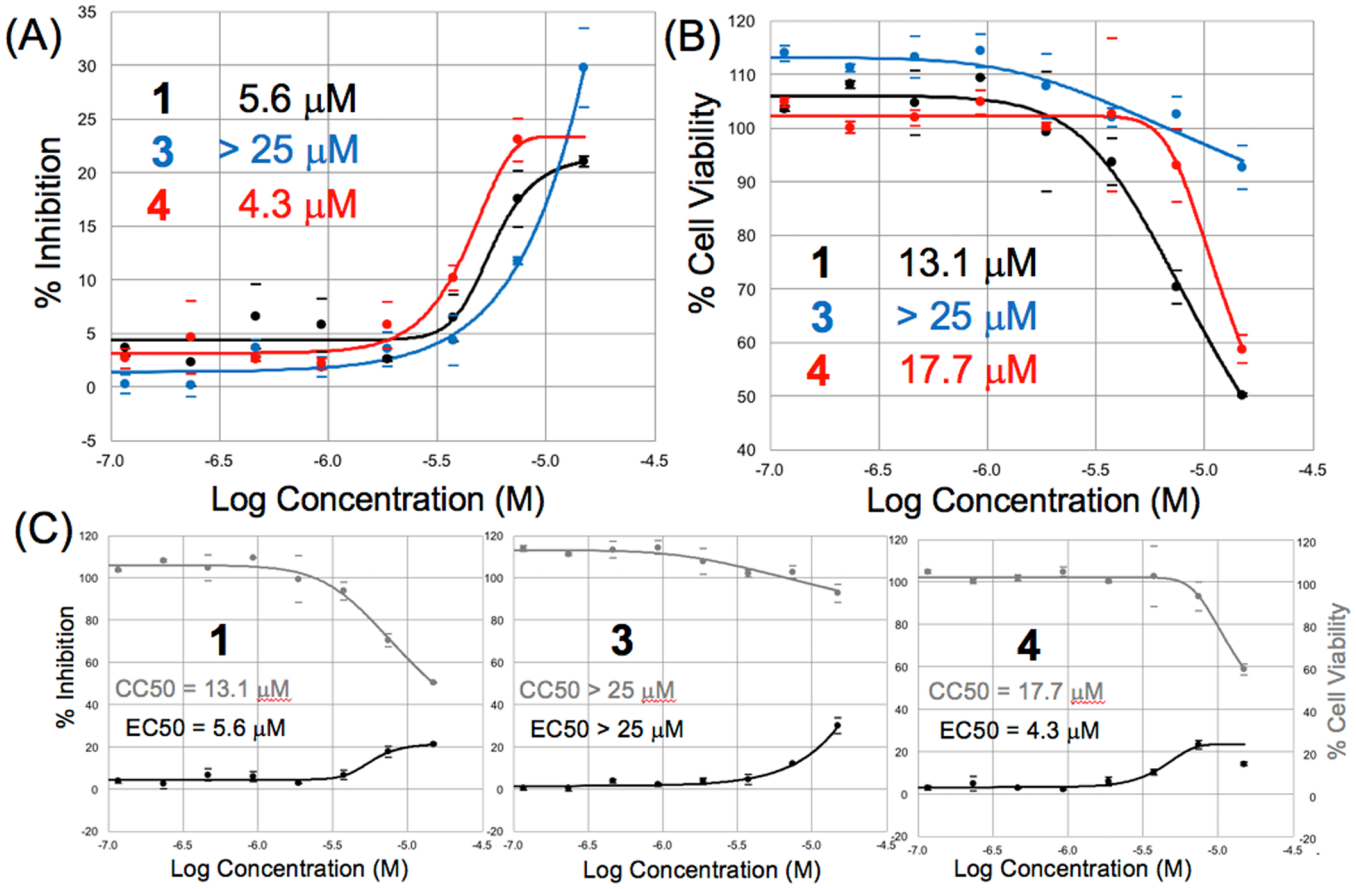


Figure 7. Dose−response curves and EC50 values for derivatives. (A) % antiviral activity against SARS-CoV-2. (B) % Cell Viability. (C) Plots of both data on the same scale for **1**, **3**, and **4** where determined EC50 and CC50 values are shown.

characterization of activity. Compounds **1**, **3**, and **4** exhibited good aqueous solubility with no observed compound precipitation or recrystallization following 96 h infection experiments at 25 μM. Cmp **2**, which was well solubilized from DMSO prior to incubation, was also shown to exhibit some degree of recrystallization in micrographs following incubation at 25 μM (Figure 5F). These collective observations correlate with the predicted higher cLogP of the compounds as expected (Table 1). Thus, because of the higher cLogP values and limited aqueous solubility of cmps **5**−**7**, the antiviral activity data with the greatest reliability are for cmps **1**−**4**. This also highlights to other researchers the challenges of working in this compound class and the importance of carefully balanced physiochemical properties. We were also aware of other small-molecule solubilizing strategies that are proposed in the literature for solubilizing derivatives in this structural class, such as solubilizing from ethanol or glycerol.[3,30] Derivatives with improved aqueous solubility as well as activity for SARS-CoV-2 will have advantages as more reliable literature reference compounds

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

or preclinical candidates compared to compounds with cLogP >5.0. A high cLogP value of the clinical drug ARB also limits the current clinical dosing beyond 200 mg three times a day for multiple days. Previous pharmacokinetics studies in human subjects have demonstrated minimal drug accumulation in multiple-day dosing studies.[62] However, as pointed out by several authors,[27] this multiple-day dosing regimen may not achieve a sufficient C_{max} for full therapeutic efficacy, such that a single oral administration of 800 mg is able to achieve an efficacious C_{max} of 4.1 μM,[27] but multiple-day dosing schemes are not able to achieve this concentration.

3.3. Inhibition of Live SARS-CoV-2 Infection. All compounds were assessed in parallel for antiviral and viability assays using Vero E6 cells to evaluate the antiviral activity against SARS-CoV-2. Compounds were preincubated first with target cells for 1 h at 37 °C before infection with live SARS-CoV-2. Following preincubation, cells were challenged with viral inoculum, and the compounds remain present in the cell culture for the duration of the infection (96 h), at which time a NR uptake assay is used to determine the virus-induced CPE. Thus, prevention of the virus-induced CPE is used as a surrogate marker to determine antiviral activity against SARS-CoV-2.

In the first round of screening in the infection assay, compounds were tested at two concentrations of 5 and 25 μM in duplicate shown in Figure 6. Cmp **4** demonstrated the highest percentage (%) of inhibition when tested in duplicate at 5 μM (Figure 6) which was statistically distinguishable to cmps **1**−**3** at 5 μM. The maximum % inhibition observed for **4** was calculated to be (46.5 ± 6.6%) inhibition compared to 50% or 100% inhibition as defined by dose responses of GS-441524 on the same plate. Thus, the inhibition of **4** was limited at a concentration of 25 μM because of decreased cell viability, but **4** exhibited more potent inhibition below 15 μM as corroborated by a full dose−response shown in (Figure 7C). Thus, from the initial round of screening, the SAR interpreted at 5 μM provided an unambiguous rank of activity from **4** (46.5%) > > **2** (9%) > **1** and **3**, where both **1** and **3** exhibited less than 5% inhibition at 5 μM.

Even though the maximum % inhibition observed for **4** was calculated to be ∼47% inhibition compared to 100% inhibition for GS-441524, such a fractional percent of inhibition is also in reasonable agreement with other independent literature reports of partial inhibition.[63] In similar CPE-based assays in Vero E6 cells,[63] Cepharanthine (92%) was found to be a full fusion inhibitor as measured by CPE (% efficacy) compared to several other partial inhibitors including NKH477 (45%), trimipramine (48%), and ingenol (38.2%). The magnitudes of fractional % inhibition for these three partial inhibitors are in a narrow range and quite similar in magnitude to that observed for **4** (46.5%). Higher levels of inhibition for ARB have been reported using other quantitative assay formats; for example, ARB was reported to achieve a maximum of 70% inhibition of virus entry in Vero E6 cells as quantitated by real-time polymerase chain reaction.[27] It is important to note that not all reported small-molecule fusion inhibitors of the Spike protein are full inhibitors, and numerous inhibitors with full inhibition in one assay format may demonstrate partial inhibition in CPE-based assays.[63]

Dose−response curves were recorded to determine EC50 values of derivatives as shown in Figure 7, confirming that cmps **1**, **3**, and **4** partially prevented the virus-induced CPE in a dose-dependent manner, reaching a maximum inhibition of approximately 25%. Although **3** displayed the highest levels of inhibition when tested at 15 μM, the two other cmps **1** and **4** showed greater levels of anti-SARS-CoV-2 activity as greater prevention of the viral CPE at 7.5 μM. Of note, **1** and **4** also displayed a dose-dependent trend of cytotoxicity in viability assays with uninfected cells, and this effect may have partially overcome the antiviral activity of these two compounds at concentrations at or above 7.5 μM. As expected from screening data results at 5 μM, cmp **4** was shown to have a modest improvement in inhibition activity compared to **1** and **3** from analysis of full dose responses. The half maximal inhibition was determined to be 4.3 μM for **4** and 5.6 μM for **1**, where **3** was found to be >25 μM (well modeled as 29.5 ± 1.5 μM). Increased activity was also evident for **4** as observed as the fractional percentage of inhibition, both in the first round of screening at 5.0 μM and as well as in full dose responses where **4** (23.0% inhibition at 7.5 μM) > **1** (17.5% inhibition at 7.5 μM) > **3** (11.7% inhibition at 7.5 μM). Micrographs of monolayers of cells following 96 h infections utilized in dose−response studies shown in Figure 8 illustrate the effects of

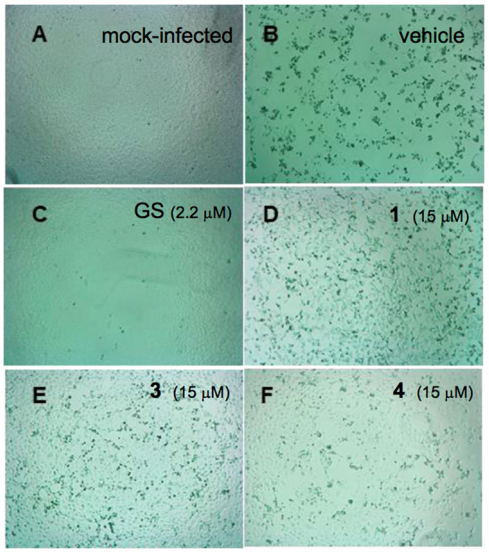


Figure 8. Microscopic evaluation of dose−response data. Microscopic evaluation of monolayers of Vero E6 cells after 96 h infection with SARS-CoV-2. Images from infected cells (B−F), or mock-infected cells (A) are shown after infection for 4 days with SARS-CoV-2 in the presence of inhibitors. (A) Mock-infected cells; (B) infection in the presence of vehicle alone; (C) infection with 2.2 μM GS-441524; (D) infection with 1 (15 μM); (E) infection with 3 (15 μM); and (F) infection with 4 (15 μM).

treatment at 15 μM compared to control 2.2 μM GS-441524. The antiviral effect of **1**, **3**, and **4** is evident in the micrographs of monolayers as viable cells were still visible in infections with these compounds, as compared to infections in the presence of vehicle alone (Figure 8B). GS-441524 at concentrations of 2.2 μM or greater completely prevented the virus-induced CPE as expected (Figure 8C).

In summary, cmps **1**, **3**, and **4** displayed a dose-dependent inhibition of the viral induced CPE, suggesting antiviral activity against SARS-CoV-2. While the level of inhibition achieved was not 100% as other literature-reported full inhibitors such as positive control GS-441524, this observation is also in good agreement with several other literature reports that have reported similar levels of partial inhibition ranging from 30 to

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

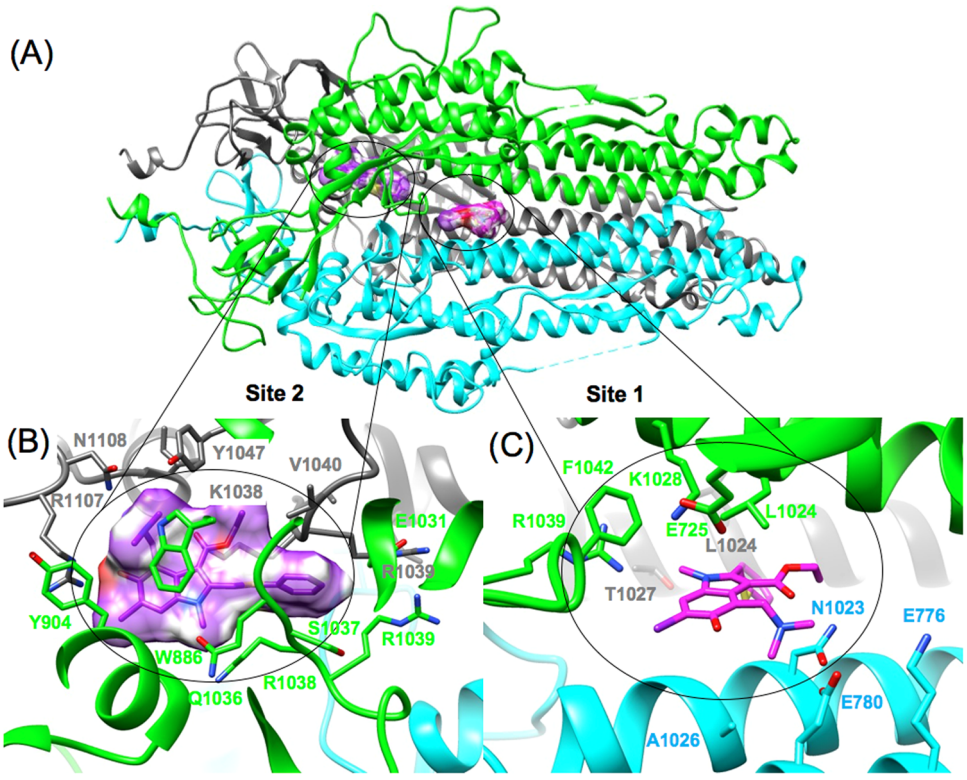


Figure 9. Two possible Arbidol binding sites on the S2 segment. (A) Shown is the full-length structure of the SARS-CoV-2 Spike protein colored as a trimer. (B) Predicted ARB binding site on the S2 segment (Site 2) is highlighted in magenta as the molecular surface of bound ARB. (C) Predicted ARB binding site on the S2 segment (Site 1) shows ARB in magenta. At Site 1, numerous residues (K776 and K947) that participate directly in binding ARB and derivatives are involved in pH-mediated conformational change.

50% inhibition.[27,63] With regard to changes in the compound structure, the substituents of **4** at the C-4 and C-6 positions were found to exhibit a modest improvement in the antiviral activity of **4** (4.3 μM) in comparison to reference cmp **1** (5.6 μM).

3.4. Comparison of Binding at Other Predicted Sites.

Our previous work using a pharmacophore procedure to map SARS-CoV-2 drug targets identified the most favorable binding site for ARB on the S2 segment,[42] and this model Site 1 (Figure 9C) was utilized for virtual screening and selection of compounds. The predicted S2 site identified from pharmacophore mapping is similar to the influenza HA2 ARB binding site (Figure 2A) according to our structural alignment[45] and as predicted by docking by Vankadari.[64] In one of the earliest publications on this topic early in the COVID-19 pandemic (submitted March 30, 2020), using molecular docking techniques Vankadari predicted that ARB bound to the S2 segment.[56] Our model binding site for ARB on the S2 segment (Figure 9C) is quite similar to that of Vankadari,[64] other than the fact that our binding site is formed in the trimeric prefusion conformation. ARB specifically was shown to stabilize the semistable prefusion conformational state preventing influenza HA2 conformational changes associated with membrane fusion,[40] so by analogy the putative location of the binding site for ARB is also seemingly consistent with the proposed mechanism of direct drug action. Based on the distribution of pH-titratable groups in the binding sites and salt-bridge residue pairs experimentally demonstrated to be involved in pH-mediated changes, Site 1 provides a structurally plausible mechanism as to how positively charged basic amine compounds such as ARB may bind and inhibit pH-mediated conformational changes. At Site 1, specific acidic residues are easily rationalized as being responsible for binding to the positively charged (+) basic amine. At Site 1, the pH-titratable residue E780 closely associates (N...O distance < 4.0 Å) with binding the positively charged (+) basic amine of ARB at neutral pH. Another close acidic residue E725 is also in reasonable proximity (N...O distance < 7.0 Å) to the positively charged (+) basic amine.

However, to date there has not yet been any published experimental structure of a small-molecule fusion inhibitor bound to the Spike S2 protein solved by either X-ray crystallography or CryoEM techniques. Thus, although there is a strong argument for drug binding at the putative S2 site by analogy (Figure 2), there remains to be any experimental conformation of drug binding at this specific site on the Spike protein. From docking, **1** or **4** into all of the TOP50 binding sites predicted on the S2 segment, the top two favorable sites were identified and are shown in Figure 9. Site 2 shown in Figure 9B is the only other reasonably feasible binding site for **1** or **4** according to our modeling data. From the analysis of molecular docking and calculated ΔG_{bind} values at all 50 sites, Site 2 is easily identified as being favorable, also from the identification of two other structurally related threefold symmetric sites, which are hydrophobic and buried near the C-terminus of the trimer. However, molecular docking and calculated ΔG_{bind} values alone are not able to statistically distinguish between Site 1 and Site 2, such that it could be thermodynamically possible for ARB and **4** to bind to both Site 1 and Site 2. In terms of predicted ΔG_{bind} values, this means that both the statistics for the top-ranked cluster (as a triplicate) and analysis of the corresponding underlying

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

distribution of predicted ΔG_{bind} values, Site 1 and Site 2 are statistically indistinguishable but are unambiguously the most favorable two sites of the TOP50.

Interestingly, both the predicted binding sites contain important clusters of highly conserved residues. According to the sequence and druggability analysis of the Spike protein published by Triguerio-Louro et al.,[65] Site 1 contains seven highly conserved residues (E725, K947, S1021, L1024, K1028, R1039, and F1042) that are involved in predicted protein−ligand interactions. Similarly, Site 2 also contains seven highly conserved residues (W886, Y904, E1031, Q1036, S1037, K1038, and R1039) that are involved in predicted protein−ligand interactions. The sequence conservation of key residues in both of these putative binding sites highlights an important potential advantage for developing S2 targeted fusion inhibitors. Inhibitors binding to these residues may be expected to have higher barriers to resistance, particularly in comparison to fusion inhibitors targeting S1 that exhibits greater variability in sequence. Many mutations to Spike are found in the S1 segment. Current mutations of interest on the Spike protein that have been observed in variants or variants of concern include [H69R, N74K, N99K, N185K, D215H, H245R, S247R, T259K, G261R, L452, Y453, E484K, N401Y, D614G, H655Y, Q677P, P681H, and R682L S686G] and were analyzed in comparison to our binding models. None of these point-mutations were found to map back to any residue that participates in ligand binding at Site 1 or Site 2.

From detailed structural analysis of the two sites, it is clear that both **1** and **4** are able to be modeled to bind at Site 2 with regard to overall "shape-fit" and hydrophobic complementarity considerations, particularly because of interactions with W886. This observation is also nicely illustrated in molecular docking data from consensus clustering of a 3-fold symmetric, similar binding mode conformation for **1** and **4** at the two other 3-fold symmetric sites. However, there are some structural features of the protein−ligand interactions at Site 2 that are less convincing. In sharp contrast, Site 2 does not provide any complementary favorable acidic residues (Asp or Glu) to form favorable polar interactions with the positively charged (+) basic amine of either **1** or **4**. The side chain of N1108 does provide close complementary hydrophilic, but nonionic interactions. At Site 2, the basic amine of **1** or **4** binds in close proximity (<6 Å) to two other positively charged residues, including a guanidino group of Arg (distance R1107@NE = 5.08 Å) and a basic amine of Lys (distance K1038@NZ = 5.66 Å), resulting in an unfavorable positively charged local environment for binding a basic amine. For this reason, if **1** or **4** is found to bind experimentally to Site 2, it would be likely that the small molecule would become deprotonated and neutral upon association with this site. We docked several reasonable protonation states of **1** and **4**, as well as neutral molecules using the same docking protocols. The same similar lowest energy binding modes were identified regardless of protonation states, and the side chain of residue N1108 was shown to form stabilizing interactions with both the protonated and neutral forms of the basic amine. Thus, it remains possible that Site 2 is a more favorable binding site than Site 1, particularly if **1** or **4** is able to bind in a deprotonated neutral form to Site 2. All other docking results within are calculated assuming ionized small-molecule basic amines (+) at neutral pH in solution.

However, in sharp contrast to Site 2, Site 1 has much better electrostatic complementarity with three important acidic residues in close proximity to bind the basic amine (E725, E1031, and D1041). E725 is highly conserved and has been implicated in mechanisms of pH-mediated conformational change requisite for fusion. Along these lines, as the structural hypothesis for the Site 1 is supported by structural knowledge and analogy to the previous influenza HA site, historical knowledge of prior structural efforts also argues against Site 2. Previous molecular docking work to predict the influenza HA2 binding site for ARB also predicted a putative site that is roughly structurally analogous to Site 2 on influenza HA2, but was later shown to not be correct by the X-ray crystal complex.[40,41] Another argument supporting Site 1 as the most plausible binding site is that it was utilized to discover derivative **4** through a structural hypothesis and virtual screening. However, in our retrospective structural analysis, derivative **4** would have been discovered by screening the same set of compounds at Site 1 or Site 2.

3.5. Modeling OA Trisaccharide Saponin Fusion Inhibitors Binding to the S2 Segment. Another independent line of evidence from modeling also strongly corroborates Site 1 as a putative site for Spike fusion inhibitors. Recently, a series of OA trisaccharide saponins (Figure 10A) was characterized with improved fusion inhibitor activity against SARS-CoV-2 Spike.[54] This study also showed good-quality SPR binding data demonstrating that small-molecule OA saponin **12f** specifically bound to the S2 segment, but not to the S1 segment. To our knowledge, this is the first literature report of a specific compound binding to the S2 segment.[54] According to our pharmacophore map and subsequent molecular docking results using that map, this series of OA saponin fusion inhibitors are predicted to bind to the S2 segment at our predicted Site 1 (Figure 11).

Using a hierarchical approach, conformations of saponin **12a** (without trisaccharide) were docked to all of the TOP50 sites of the Spike S2 segment. Then, after initial identification of the TOP5 most favorable sites from this first step, a more extensive sampling strategy was used to refine the ranking of the TOP5 sites docking with the entire covalent structure of saponin **12a** (with trisaccharide). There was good geometric agreement (RMSD < 2.0 Å) between the top-ranked binding modes predicted from saponin **12a** (with and without trisaccharide) for both Site 1 and Site 2. The statistical significance of these results is shown in Figure 11, not only from analysis of predicted (ΔG_{bind}) values from a top-ranked cluster (triplicate), but also from statistical Z_{score} analysis of the entire energy distribution of predicted (ΔG_{bind}) values compared to the rest of the TOP50 sites. For saponin **12a** (without trisaccharide) docking into the TOP50 (Figure 11B), Site 1 (ΔG_{bind} = −8.0, Z_{score} = −4.8) and Site 2 (ΔG_{bind} = −8.0, Z_{score} = −4.8) are both statistically significant compared to the distribution of the other TOP50 sites (99% confidence, *P*-value < 0.00001). In a comparison of the results for saponin **12a** (with trisaccharide) docking into the TOP5 (Figure 11D), Site 1 (ΔG_{bind} = −15.3, Z_{score} = −3.3) was found to be statistically significant from the distribution of the other TOP5 sites (99% confidence, *P*-value = 0.0005) including the data from Site 2. Thus, Site 1 is the most probable binding site on the entire S2 segment according to not only a single top-ranked binding mode (ΔG_{bind}) (*P*-value = 0.0005), but also from a comparison of the entire (ΔG_{bind}) energy distribution for Site 1 to the other 4 of the TOP5 sites. Thus, while analysis of docking data

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

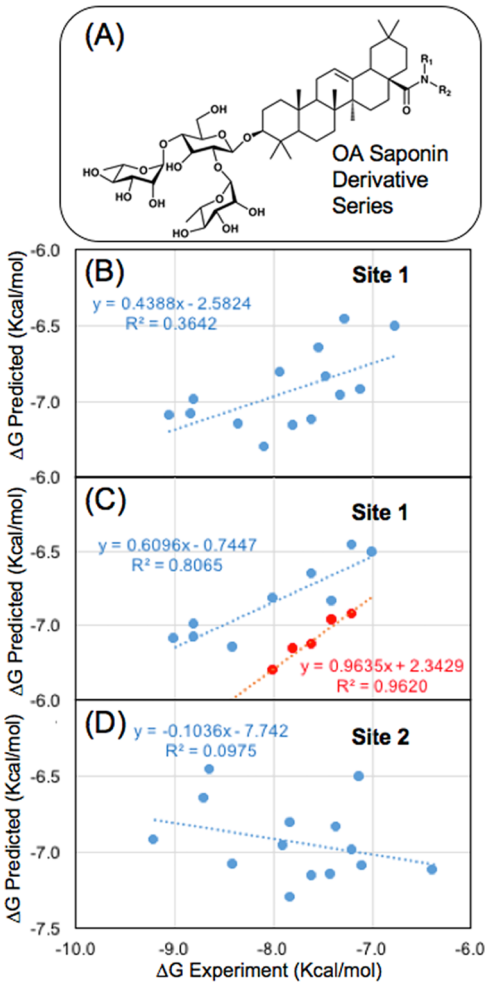


Figure 10. Predicted ΔG_{bind} values from Site 1 exhibit correlation with experimental SAR data. (A) Series of OA saponin derivatives[54] were well modeled binding to Site 1, where sufficient linear correlation is achieved either comparing (B) all 14 derivatives (blue) or (C) 14 OA saponin derivatives modeled as two separate groups of linear series 1 (blue) or series 2 (red). (D) Predicted ΔG_{bind} values for all 14 compound modeled at Site 2 exhibit lower correlation (R^2 = 0.097).

is not able to discriminate as to whether ARB, **1**, or **4** binds more favorably to Site 1 or Site 2, this parallel analysis of OA saponins definitively predicts that Site 1 is the most probable binding site in comparison to Site 2.

Next, a series of 14 derivatives of OA saponin fusion inhibitors were modeled binding to Site 1 and Site 2. The molecular geometry of the rigid hydrophobic structure of the OA natural product scaffold allowed a geometric search with minimal flexible degrees of freedom and statistically rigorous determination of binding geometry. This allowed for a robust determination of a consensus binding mode for the majority of the OA saponin derivatives. In modeling the series of 14 derivatives of saponins **12a** and **12f**, the "untrained" predicted ΔG_{bind} values exhibited some correlation with the experimental IC50 values (Table 2). R^2 correlation values range from R^2 = 0.364 for all 14 compounds (Figure 10B) modeled at Site 1 where high correlations of R^2 = 0.80 to 0.96 were achieved modeling the data set as two separate series of "untrained" predicted ΔG_{bind} rankings (Figure 10C). In comparison, 14 compounds (Figure 10D) modeled at Site 2 exhibited a negative correlation (Pearson's R), with a calculated R^2 = 0.098. Thus, from modeling all 14 compounds, the "untrained" predicted ΔG_{bind} values had much greater correlation at Site 1 (R^2 = 0.364) compared to Site 2 (R^2 = 0.098).

The robustness of this correlation analysis was examined using a cross-validation approach as described using 14 cross-validation datasets (Table S2 Supporting Information). Linear correlation analysis is performed for 14 cross-validation "test" sets of 10 derivatives and four derivative "validation" sets. Over the 14 cross-validation datasets (Table S3 Supporting Information), for Site 1 the average correlation for the "test" set was found to be (R^2 = 0.374 ± 0.111) compared to Site 2 (R^2 = 0.131 ± 0.120). The average correlation for the "validation" set was found to be (R^2 = 0.458 ± 0.318) compared to Site 2 (R^2 = 0.303 ± 0.245), where the average Pearson R value had a negative correlation for Site 2. This cross-validation confirms the conclusion that the data have greater correlation modeled at Site 1 compared to Site 2. Compared to previous benchmark studies characterizing this scoring function method and performance, these levels of R^2 correlation are adequate to establish confidence in the binding model as reflecting experimental SAR data.[44,45]

3.6. Structural Implications for Fusion Inhibitors from SARs. From the round of screening ARB derivatives at 5 μM, the ranking of the compounds from greatest activity to least activity remained (**4**, ≫**2**, >**1**, >**3**), and dose−response studies confirmed the same ranking of compounds from greatest activity to least (**4**, >**1**, ≫**3**). In terms of differences in the compound structure, compared to **1**, derivative **4** has important differences in C-4 substituents that participate in additional specific protein−ligand interactions. Previous SAR data for ARB derivatives for influenza A and B have also shown that the C-4 basic amine functional group and the C-5 hydroxyl were important for activity,[66,67] where the C-6 substituent (Br) had little effect. Modification of the C-5 position to (CH2OH) improved antiviral potency and was directly shown to directly bind to HA2 with a greater affinity than ARB. The C-6 substituent (Br) was also shown to have little effect on HBV or HCV antiviral activity, while substitution of the C-4 basic amine to heterocyclic groups increased antiviral activity for HBV but not HCV.[66,67] As shown in Figure 2A, extending the C-4 basic amine group with bulky substituents should reduce antiviral activity for influenza A; however, in the model of binding to the SARS-CoV-2, the C-4 substituent of derivative **4** forms complementary interactions with additional residues.

A key difference between the model of **1** and **4** is that the C-4 substituent 2-[2-(4-ethylpiperazin-1-yl)ethoxy]ethan-1-ol specifically forms interactions that stabilize the prefusion conformation of several residues involved in pH-mediated conformational change, including the highly conserved residue K947. In various structures of the prefusion core, four residues (L945, K947, L948, and V951) from HR1 make key contacts to the rest of the S2 NDT domain (E725, L727, and P728). As shown in Figure 12C, derivative **4** makes key hydrophobic contacts with both residues (L727 and P728) from NTD and residues K947 and V951 from HR1. The terminal polar (ethoxy)ethan-1-ol hydroxyl group is able to form favorable electrostatic interactions with the primary amine (NZ) of K776, while forming primarily favorable hydrophobic interactions with the hydrophobic side chain of K947. From this model of the protein−ligand interaction in the neutral pH prefusion conformation (6vxx.pdb),[53] our working hypothesis

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

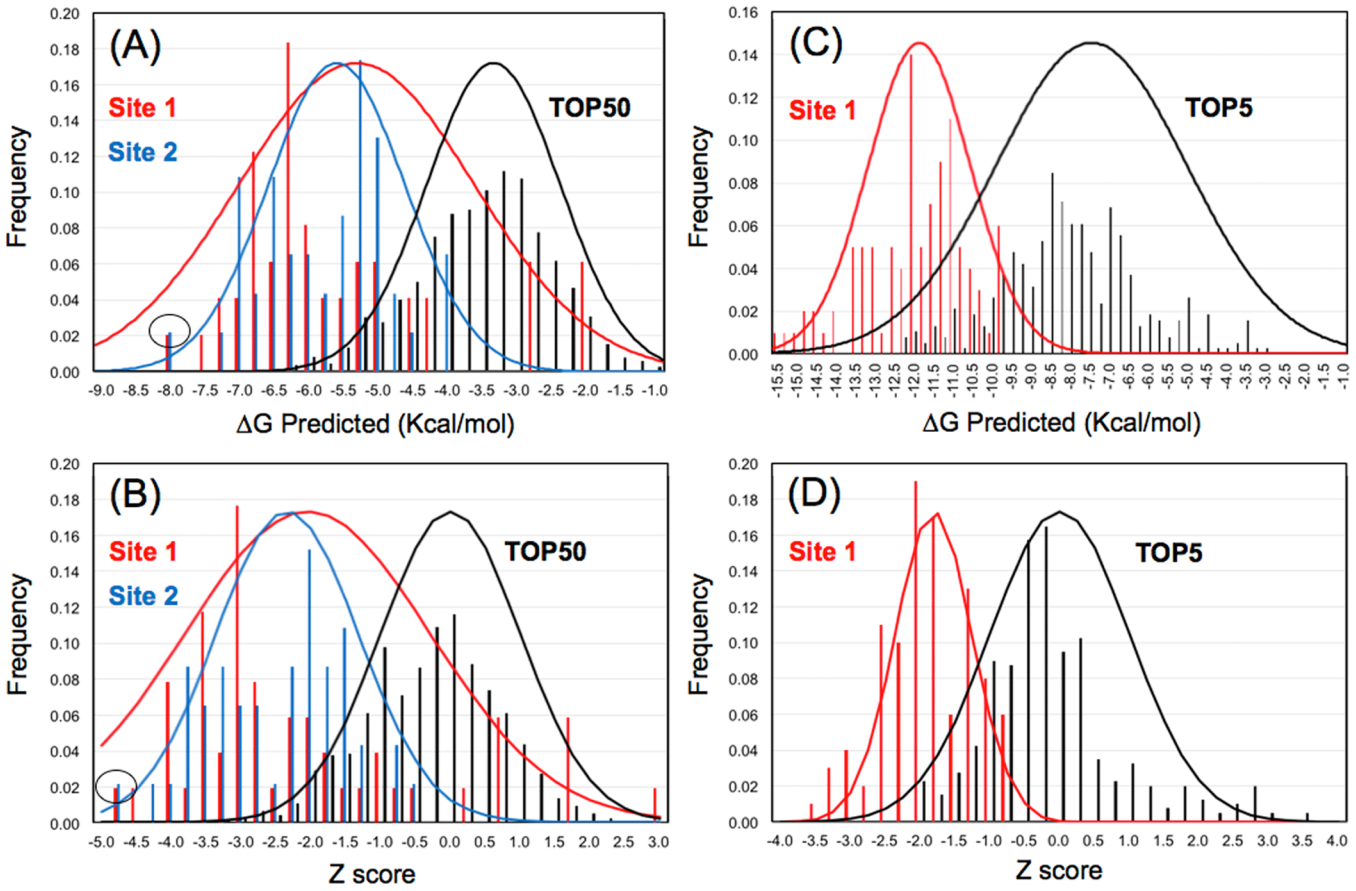


Figure 11. Predicted ΔG_{bind} energy distributions for OA saponin **12a** binding to S2. (A) Predicted ΔG_{bind} energy distributions for saponin **12a** (without trisaccharide) docked to all of the TOP50 sites of the Spike S2 segment. (B) Statistical Z_{score} values of the energy distribution showing that the lowest energy cluster of both Site 1 and Site 2 is (Z_{score} = −4.8). (C) Predicted ΔG_{bind} energy distributions for saponin **12a** (with trisaccharide) docked to all of the TOP5 sites of the Spike S2 segment, comparing the distribution of Site 1 (red) to the rest of the distribution of the other TOP4 sites, which includes Site 2. (D) Statistical Z_{score} values of the energy distribution showing that the lowest energy cluster of Site 1 is (Z_{score} = −3.3) compared to the rest of the distribution which includes Site 2.

Table 2. Comparison of OA Saponin Inhibitor Predicted Binding Free Energy at Site 1 and Site 2 to Those Approximated from Experimental IC50 Values

	compound	observed	observed	predicted	predicted
			Calc.	Site 1	Site 2
	OA saponin ID	IC50 (μM)	ΔG (kcal/mol)	ΔG (kcal/mol)	ΔG (kcal/mol)
1	12a	18.1	−6.47	−9.64	−8.64
2	12b	5.5	−7.17	−9.90	−7.61
3	12c	8.3	−6.93	−9.56	−9.20
4	12d	10.0	−6.82	−9.97	−7.82
5	12e	6.3	−7.09	−10.53	−7.10
6	12f	4.4	−7.31	−10.05	−7.82
7	12i	5.9	−7.13	−9.81	−6.38
8	12j	7.8	−6.97	−9.66	−7.90
9	12k	5.7	−7.15	−10.18	−7.42
10	12l	6.3	−7.09	−10.41	−8.41
11	12m	9.6	−6.84	−9.73	−7.35
12	12n	16.8	−6.51	−9.38	−7.13
13	12o	7.4	−7.00	−10.40	−7.20
14	12p	13.1	−6.66	−9.77	−8.70

was that the additional interactions provide additional stabilization to the prefusion conformation, providing additional stability to the residues (L945, K947, L948, and V951) from HR1 that make key contacts to the N-terminal S2 NTD domain, which undergoes major conformational rearrangement to form the postfusion conformation.

In modeling the series of OA saponins, the consensus binding mode for the trisaccharide of saponin **12a** or **12f** is modeled to form similar interactions to the C-4 substituent 2-[2-(4-ethylpiperazin-1-yl)ethoxy]ethan-1-ol of **4** as shown in Figure 12D. This observation is interesting in light of SAR observations for OA saponins, showing that removal of the branched trisaccharide scaffold (α-L-rhamnopyranosyl-(1 → 2)-[α-L-rhamnopyranosyl-(1 → 4)]-β-D-glucopyranosyl) or removal of (α-L-rhamnosyl) or (β-D-glucopyranosyl) substituents resulted in significant reduction in % inhibition.[54] From the model of OA saponins, the interactions of the (β-D-glucopyranosyl) substituents are structurally equivalent to the interactions formed from the C-4 substituent 2-[2-(4-ethylpiperazin-1-yl)ethoxy]ethan-1-ol of **4**, primarily involving interactions with residues (P728, K776, K947, and E1017). The superposition of these structural features of **4** and saponin **12f** highlights the importance of efficient protein−ligand interactions with these residues for fusion inhibitors targeting Site 1.

3.7. Modeling Small-Molecule Binding to Spike as a Function of pH. Recent electron microscopy studies have experimentally characterized snapshots of the full-length Spike

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

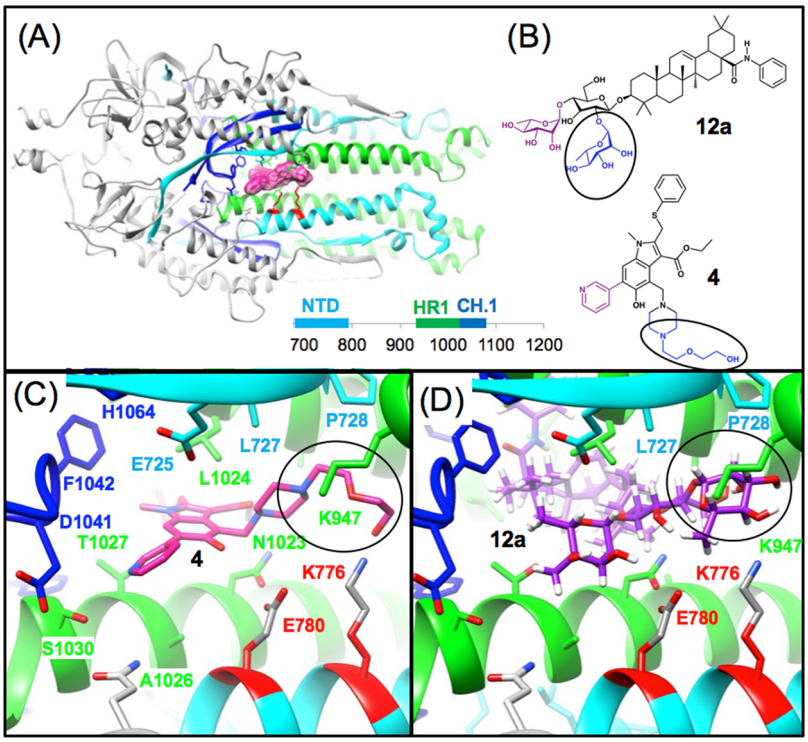


Figure 12. Common structural features of predicted ARB derivative 4 and OA saponins bound to Site 1. (A) Ribbon model of the S2 segment, using color to highlight the NTD (cyan), HR1 (green), and CH.1 (blue) segments. (B) 2D structures of 4 and saponin **12a** highlighting in blue and purple regions of overlap between the two 3D models. (C) Lowest energy binding mode of ARB derivative **4** shown in magenta highlighting residues E780 and K776 in red from the NTD segment which undergoes major pH-mediated conformational changes. (D) Predicted Structure of OA saponin **12a**, where the common overlapping structural features of **4** and **12a** are highlighted with a circle.

protein in the prefusion conformation from neutral pH to pH 5.5, 4.5, and 4.0, making it possible to model fusion inhibitor binding as a function of pH.[55] The immediate availability of these experimental structures provides an unprecedented opportunity to understand pH-mediated changes in the conformations of these binding sites using molecular modeling. At Site 1, residues K1045, K947, and K776 undergo minor conformational changes at pH 4.5 compared to pH 5.5 as illustrated in Figure 13A by showing only rearrangements of the side chain conformations of lysine residues such as K947, which was predicted to form more extensive interactions with the C-4 substituent of **4** compared to **1**. ARB derivative **4** forms stabilizing interactions with K947 and K776, where **1** does not. Thus, various structural differences in conformation as a function of pH made it possible to model the effects of the conformational changes on predicted small-molecule binding. Although more advanced and rigorous molecular simulation and conformational sampling techniques may be used to study this problem, we present results using the same consistent CHARMM-based molecular docking methods and ΔG_{bind} calculations as they are able to predict the expected trend as a function of pH using standard CHARMM-based potential functions for the charged and neutral states of titratable residues of His, Asp, and Glu.

Small-molecule binding was modeled at Site 1 as a function of pH, including compounds **1** and **4** and representative OA saponins **12a** and **12f**. The observed trend in predicted ΔG_{bind} values is consistent with the hypothesis that these small-molecule agents bind with a higher affinity at neutral pH prefusion conformations, specifically stabilizing those (neutral pH) conformations and inhibiting conformational changes that are required to adopt lower pH prefusion conformations. OA saponins **12a** and **12f** showed distinctively less favorable binding at lower pH values as shown in Figure 13B, where the greatest energetic contribution is directly from the rigid body of the hydrophobic OA core which was less favorable in the more collapsed binding pocket at lower pH. OA saponins **12a** and **12f** were well modeled in the same binding mode for calculations of neutral pH, pH 5.5, and pH 4.5 allowing a comparison of predicted ΔG_{bind} values for the same binding mode. A comparison of the same small-molecule binding mode makes it more straightforward to interpret the effects of structural changes to the binding site on the resulting calculation of (ΔG_{bind}).

For ARB derivatives, the distribution of pH-titratable groups at Site 1 provides a plausible mechanism as to how fusion inhibitors may bind and inhibit pH-mediated conformational changes of Spike. At Site 1, residue E780 forms the closest electrostatic interactions with the basic amine of ARB at neutral pH, where E725 is also in close proximity (N...O distance <7.0 Å). Particularly in going from pH 5.5 to pH 4.0, as Asp and Glu residues become protonated and neutral, the greater positive charge electrostatic environment of the binding pocket should be less favorable for binding the positively charged (+) basic amine groups in the fusion inhibitors, providing a biophysical explanation as to why these agents may preferentially stabilize the neutral pH conformations of Spike. As shown in Figure 13B, the compounds studied showed that that the lowest calculated (ΔG_{bind}) was from the neutral pH conformations and that predicted ΔG_{bind} became less favorable as a function of pH over the pH range of 5.5−4.0 as might be expected. Structural analysis and scoring function data suggest

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

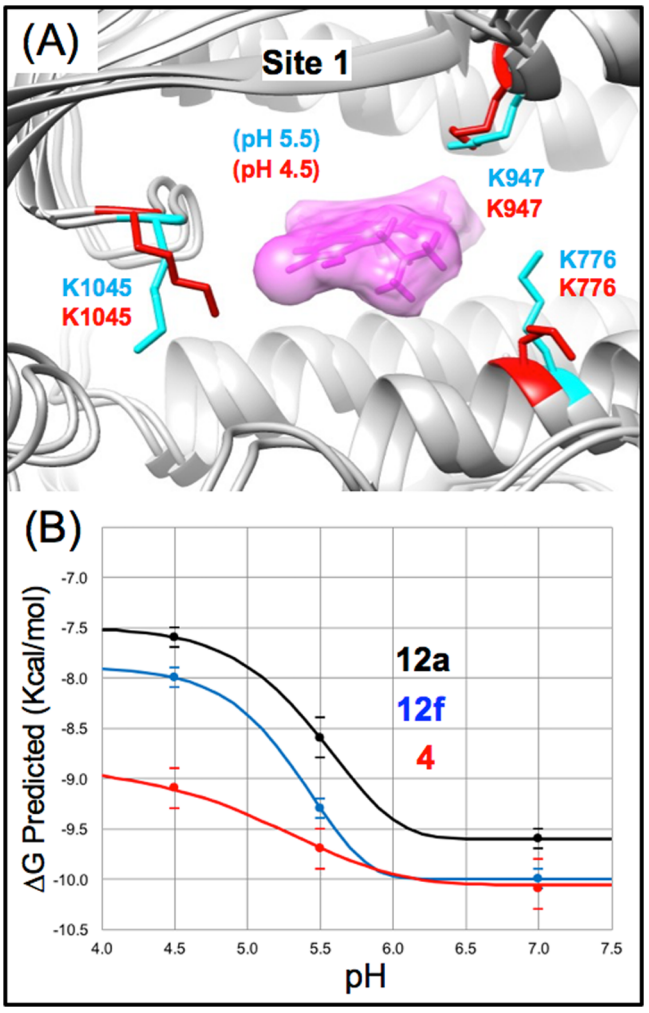


Figure 13. Fusion inhibitor predicted ΔG_{bind} as a function of pH. (A) Superimposed are two recently determined structures of the Spike protein that experimentally characterize pH-mediated conformational changes in structure.[55] At Site 1, numerous residues (K776 and K947) that participate directly in binding ARB and derivatives are involved in pH-mediated conformational change. (B) Predicted ΔG_{bind} as a function of pH at Site 1 for **12a**, **12f**, and **4**. To illustrate the trend of ΔG_{bind} as a function of pH, sigmoidal functions were calculated with the five-parameter logistic equation to best fit the data.

that this is most likely due to the subtle conformational changes in the binding site nonpolar interaction surfaces captured in the experimental structures. In summary, predicted ΔG_{bind} as a function of pH became less favorable as a function of pH, as might be expected from fusion inhibitors that act to specifically bind to and stabilize (neutral pH) conformations and inhibit conformational changes that are required to adopt lower pH prefusion conformations.

4. CONCLUSIONS

We hope that modeling these ARB derivatives in comparison to a series of OA saponin fusion inhibitors is helpful in deciphering various complex mechanisms of action for Spike fusion inhibitors. Docking alone makes it appear possible that **1** or **4** may bind to both Sites 1 and 2, and our data were unable to exclude this. From parallel modeling of a series of OA saponins binding to Sites 1 and 2, OA saponins are found to most likely bind to Site 1. The superposition of structural features for our binding models of **4** and saponin **12f** at Site 1 highlights the importance of efficient protein−ligand interactions with residues (P728, K776, K947, and E1017) for fusion inhibitors targeting Site 1 and provides a specific hypothesis for the structure-based design of improved inhibitors. Modeling small-molecule binding as a function of pH was able to tentatively confirm the working hypothesis that fusion inhibitors function by specifically binding to the neutral pH prefusion conformation and preventing pH-mediated conformational change.

As an agent with modest broad-spectrum antiviral activity and several proposed mechanisms of action, characterization of ARB derivatives with improved activity remains important for advancing the preclinical development of fusion inhibitors and selecting optimal candidates for combination therapy. The strategy here is to find multiple vectors of attack against SARS-CoV-2, to most effectively use drugs with different mechanisms of action to achieve a synergistic effect, while also minimizing side effects. As ARB is a clinical drug that has demonstrated synergistic inhibition with other direct-acting antivirals such as Remdesivir,[29] we hope that this candidate **4** with improved properties may be a promising reference compound for the ongoing development of improved inhibitors for synergistic multidrug combination therapies.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c01061.

Virtual screening data for ARB derivatives (Table S1), cross-validation datasets (Table S2), and cross-validation data (Table S3) (PDF)

SUPP_INFO.zip Summary PDB coordinate files for models of the Spike protein and lowest energy binding modes for compounds **1** and **4** and saponin **12a** at both Site 1 and Site 2 and example CHARMM input and output directories for docking saponin **12a** at both Site 1 and Site 2, including all required topology, parameter and input files, and example output files (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Roger S. Armen − *Department of Pharmaceutical Sciences, College of Pharmacy, Thomas Jefferson University, Philadelphia, Pennsylvania 19170, United States;* orcid.org/0000-0003-3830-1450; Email: roger.armen@jefferson.edu

Author

Matthew R. Freidel − *Department of Pharmaceutical Sciences, College of Pharmacy, Thomas Jefferson University, Philadelphia, Pennsylvania 19170, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c01061

Notes

The authors declare no competing financial interest.

The program CHARMM is publicly available under academic license for research (https://www.charmm.org). MarvinSketch is publicly available under academic license for research (http://www.chemaxon.com). UCSF Chimera is publicly available under academic license for research (https://www.

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

cgl.ucsf.edu/chimera). All relevant data are shown in figures, listed in tables, or included in the Supporting Information. The docking ΔG_{bind} data and 2D compound information for all compounds from virtual screening may be found in Table S1. PDB files and example CHARMM input and output files are provided in the Supporting Information as a ZIP file. All data required to reproduce cross-validation are available in Table [2], in cross-validation datasets and cross-validation data Tables S2 and S3, respectively.

■ ACKNOWLEDGMENTS

R.S.A. was partially supported by grants 1P01HL114471-06 and R01 AR077666 from the National Institutes of Health and an award from the Dr. Ralph and Marian Falk Medical Research Trust (to J. Benovic, A. Aplin, P. Wedegaertner, T. Sato, and J. Aguirre-Ghiso). There was no additional external funding received for this study.

■ ABBREVIATIONS:

ARB, Arbidol; CPE, cytopathic effect; MD, molecular dynamics; SAR, structure−activity relationship

■ REFERENCES

(1) Zhong, H.; Wang, Y.; Zhang, Z.-L.; Liu, Y.-X.; Le, K.-J.; Cui, M.; Yu, Y.-T.; Gu, Z.-C.; Gao, Y.; Lin, H.-W. Efficacy and safety of current therapeutic options for COVID-19 - lessons to be learnt from SARS and MERS epidemic: A systematic review and meta-analysis. *Pharmacol. Res.* **2020**, *157*, No. 104872.

(2) Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **2020**, *15*, 907−932.

(3) Brooks, M. J.; Burtseva, E. I.; Ellery, P. J.; Marsh, G. A.; Lew, A. M.; Slepushkin, A. N.; Crowe, S. M.; Tannock, G. A. Antiviral activity of arbidol, a broad-spectrum drug for use against respiratory viruses, varies according to test conditions. *J. Med. Virol.* **2012**, *84*, 170−181.

(4) Fediakina, I. T.; Leneva, I. A.; Iamnikova, S. S.; Livov, D. K.; Glushkov, R. G.; Shuster, A. M. Sensitivity of influenza A/H5 viruses isolated from wild birds on the territory of Russia to arbidol in the cultured MDCK cells. *Vopr. Virusol.* **2005**, *50*, 32−35.

(5) Leneva, I. A.; Shuster, A. M. Antiviral etiotropic chemicals: efficacy against influenza A viruses A subtype H5N1. *Vopr. Virusol.* **2006**, *51*, 4−7.

(6) Wang, Y.; Ding, Y.; Yang, C.; Li, R.; Du, Q.; Hao, Y.; Li, Z.; Jiang, H.; Zhao, J.; Chen, Q.; Yang, Z.; He, Z. Inhibition of the infectivity and inflammatory response of influenza virus by Arbidol hydrochloride in vitro and in vivo (mice and ferret). *Biomed. Pharmacother.* **2017**, *91*, 393−401.

(7) Leneva, I. A.; Burtseva, E. I.; Yatsyshina, S. B.; Fedyakina, I. T.; Kirillova, E. S.; Selkova, E. P.; Osipova, E.; Maleev, V. V. Virus susceptibility and clinical effectiveness of anti-influenza drugs during the 2010-2011 influenza season in Russia. *Int. J. Infect. Dis.* **2016**, *43*, 77−84.

(8) Liu, Q.; Xiong, H.-R.; Lu, L.; Liu, Y.-Y.; Luo, F.; Hou, W.; Yang, Z.-Q. Antiviral and anti-inflammatory activity of arbidol hydrochloride in influenza A (H1N1) virus infection. *Acta Pharmacol. Sin.* **2013**, *34*, 1075−1083.

(9) Fediakina, I. T.; Shchelkanov, M. I.; Deriabin, P. G.; Leneva, I. A.; Gudova, N. V.; Kondrat'eva, T. V.; L'vov, D. K. Susceptibility of pandemic influenza virus A H1N1 and highly pathogenic avian influenza vwZirus A H5N1 to antiinfluenza agents in cell culture. *Antibiot. Khimioter.* **2009**, *2011*, 3−9.

(10) Shi, L.; Xiong, H.; He, J.; Deng, H.; Li, Q.; Zhong, Q.; Hou, W.; Cheng, L.; Xiao, H.; Yang, Z. Antiviral activity of arbidol against influenza A virus, respiratory syncytial virus, rhinovirus, coxsackie virus and adenovirus in vitro and in vivo. *Arch. Virol.* **2007**, *152*, 1447−1455.

(11) Deng, H.-Y.; Luo, F.; Shi, L.-Q.; Zhong, Q.; Liu, Y.-J.; Yang, Z.-Q. Efficacy of arbidol on lethal hantaan virus infections in suckling mice and in vitro. *Acta Pharmacol. Sin.* **2009**, *30*, 1015−1024.

(12) Haid, S.; Pietschmann, T.; Pécheur, E.-I. Low pH-dependent hepatitis C virus membrane fusion depends on E2 integrity, target lipid composition, and density of virus particles. *J. Biol. Chem.* **2009**, *284*, 17657−17667.

(13) Pécheur, E.-I.; Lavillette, D.; Alcaras, F.; Molle, J.; Boriskin, Y. S.; Roberts, M.; Cosset, F.-L.; Polyak, S. J. Biochemical mechanism of hepatitis C virus inhibition by the broad-spectrum antiviral arbidol. *Biochemistry* **2007**, *46*, 6050−6059.

(14) Gagarinova, V. M.; Ignat'eva, G. S.; Sinitskaia, L. V.; Ivanova, A. M.; Rodina, M. A.; Tur'eva, A. V. The new chemical preparation arbidol: its prophylactic efficacy during influenza epidemics. *Zh. Mikrobiol. Epidemiol. Immunobiol.* **1993**, *5*, 40−43.

(15) Obrosova-Serova, N. P.; Burtseva, E. I.; Nevskiĭ, I. M.; Karmanova, R. I.; Nazarov, V. I.; Pitkenen, A. A.; Slepushkin, A. N. The protective action of arbidol during a rise in respiratory diseases in 1990. *Vopr. Virusol.* **1991**, *36*, 380−381.

(16) Shumilov, V. I.; Shuster, A. M.; Lobastov, S. P.; Shevtsov, V. A.; Mednikov, B. L.; Piiavskiĭ, S. A.; Litus, V. I. Efficacy of arbidol in prophylaxis and treatment of acute respiratory viral infections in servicemen. *Voen. Med. Zh.* **2002**, *323*, 51−53.

(17) Shuster, A. M.; Shumilov, V. I.; Shevtsov, V. A.; Mar'in, G. G.; Kozlov, V. N. Arbidol used in the prophylaxis of acute respiratory viral infections and their complications in servicemen. *Voen. Med. Zh.* **2004**, *325*, 44−45.

(18) Wang, M.-Z.; Cai, B.-Q.; Li, L.-Y.; Lin, J.-T.; Su, N.; Yu, H.-X.; Gao, H.; Zhao, J. Z.; Liu, L. Efficacy and safety of arbidol in treatment of naturally acquired influenza. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao* **2004**, *26*, 289−293.

(19) Gatich, R. Z.; Kolobukhina, L. V.; Vasil'ev, A. N.; Isaeva, E. I.; Burtseva, E. I.; Orlova, T. G.; Voronina, F. V.; Kol'tsov, V. D.; Malinovskaia, V. V. Viferon suppositories in the treatment of influenza in adults. *Antibiot. Khimioter.* **2008**, *53*, 13−17.

(20) Kolobukhina, L. V.; Malinovskaia, V. V.; Gatich, R. Z.; Merkulova, L. N.; Burtseva, E. I.; Isaeva, E. I.; Parshina, O. V.; Guseva, T. S.; Orlova, T. G.; Voronina, F. V. Evaluation of the efficacy of wiferon and arbidol in adult influenz. *Vopr. Virusol.* **2008**, *53*, 31−33.

(21) Kolobukhina, L. V.; Merkulova, L. M.; Shchelkanov, M. I.; Burtseva, E. I.; Isaeva, E. I.; Malyshev, N. A.; L'vov, D. K. Efficacy of ingavirin in adults with influenza. *Ter. Arkh.* **2009**, *81*, 51−54.

(22) Beliaev, A. L.; Burtseva, E. I.; Slepushkin, A. N.; Beliaeva, N. A.; Nosik, N. N.; Lavrukhina, L. A.; Maksimova, L. N.; Feklisova, L. V.; Shibekova, V. M.; Danina, N. F.; Tselipanova, E. E.; Bylinkina, S. A.; Voĭtsekhovskaia, E. M.; Sominina, A. A. Arbidole—a new drug for prevention of influenza and acute viral respiratory infections in children. *Vestn. Ross. Akad. Med. Nauk.* **1996**, *8*, 34−37.

(23) Drinevskiĭ, V. P.; Osidak, L. V.; Natsina, V. K.; Afanas'eva, O. I.; Mil'kint, K. K.; Danini, G. V.; Ispolatova, A. V.; Koreniako, I. E.; Karelina, N. N.; Marinich, I. G.; Boldasov, V. K. Chemotherapeutics for treatment of influenza and other viral respiratory tract infections in children. *Antibiot. Khimioter.* **1998**, *43*, 29−34.

(24) Blaising, J.; Polyak, S. J.; Pécheur, E.-I. Arbidol as a broad-spectrum antiviral: an update. *Antiviral Res.* **2014**, *107*, 84−94.

(25) Javorac, D.; Grahovac, L.; Manić, L.; Stojilković, N.; Anđelković, M.; Bulat, Z.; Đukić-Ćosić, D.; Curcic, M.; Djordjevic, A. B. An overview of the safety assessment of medicines currently used in the COVID-19 disease treatment. *Food Chem. Toxicol.* **2020**, *144*, No. 111639.

(26) Khamitov, R. A.; Loginova, S. I.; Shchukina, V. N.; Borisevich, S. V.; Maksimov, V. A.; Shuster, A. M. Antiviral activity of arbidol and its derivatives against the pathogen of severe acute respiratory syndrome in the cell cultures. *Vopr. Virusol.* **2008**, *53*, 9−13.

(27) Wang, X.; Cao, R.; Zhang, H.; Liu, J.; Xu, M.; Hu, H.; Li, Y.; Zhao, L.; Li, W.; Sun, X.; Yang, X.; Shi, Z.; Deng, F.; Hu, Z.; Zhong,

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

W.; Wang, M. The anti-influenza virus drug, arbidol is an efficient inhibitor of SARS-CoV-2 in vitro. *Cell Discov.* **2020**, *6*, 28.
(28) Xiao, X.; Wang, C.; Chang, D.; Wang, Y.; Dong, X.; Jiao, T.; Zhao, Z.; Ren, L.; Dela Cruz, C. S.; Sharma, L.; Lei, X.; Wang, J. Identification of Potent and Safe Antiviral Therapeutic Candidates Against SARS-CoV-2. *Front. Immunol.* **2020**, *11*, No. 586572.
(29) Bobrowski, T.; Chen, L.; Eastman, R. T.; Itkin, Z.; Shinn, P.; Chen, C. Z.; Guo, H.; Zheng, W.; Michael, S.; Simeonov, A.; Hall, M. D.; Zakharov, A. V.; Muratov, E. N. Synergistic and Antagonistic Drug Combinations against SARS-CoV-2. *Mol. Ther.* **2021**, *29*, 873−885.
(30) Pizzorno, A.; Padey, B.; Dubois, J.; Julien, T.; Traversier, A.; Dulière, V.; Brun, P.; Lina, B.; Rosa-Calatrava, M.; Terrier, O. In vitro evaluation of antiviral activity of single and combined repurposable drugs against SARS-CoV-2. *Antiviral Res.* **2020**, *181*, No. 104878.
(31) Zhu, Z.; Lu, Z.; Xu, T.; Chen, C.; Yang, G.; Zha, T.; Lu, J.; Xue, Y. Arbidol monotherapy is superior to lopinavir/ritonavir in treating COVID-19. *J. Infect.* **2020**, *81*, e21−e23.
(32) Costanzo, M.; Rachele De Giglio, M. A.; Roviello, G. N. SARS-CoV-2: Recent Reports on Antiviral Therapies Based on Lopinavir/Ritonavir, Darunavir/Umifenovir, Hydroxychloroquine, Remdesivir, Favipiravir and other Drugs for the Treatment of the New Coronavirus. *Curr. Med. Chem.* **2020**, *27*, 4536−4541.
(33) Li, Y.; Xie, Z.; Lin, W.; Cai, W.; Wen, C.; Guan, Y.; Mo, X.; Wang, J.; Wang, Y.; Peng, P.; Chen, X.; Hong, W.; Xiao, G.; Liu, J.; Zhang, L.; Hu, F.; Li, F.; Zhang, F.; Deng, X.; Li, L. Efficacy and Safety of Lopinavir/Ritonavir or Arbidol in Adult Patients with Mild/Moderate COVID-19: An Exploratory Randomized Controlled Trial. *Med* **2020**, *1*, 105−113.e4.
(34) Nojomi, M.; Yassin, Z.; Keyvani, H.; Makiani, M. J.; Roham, M.; Laali, A.; Dehghan, N.; Navaei, M.; Ranjbar, M. Effect of Arbidol (Umifenovir) on COVID-19: a randomized controlled trial. *BMC Infect. Dis.* **2020**, *20*, 954.
(35) Wen, C. Y.; Xie, Z. W.; Li, Y. P.; Deng, X. L.; Chen, X. T.; Cao, Y.; Ou, X.; Lin, W. Y.; Li, F.; Cai, W. P.; Li, L. H. Real-world efficacy and safety of lopinavir/ritonavir and arbidol in treating with COVID-19 : an observational cohort study. *Zhonghua Nei Ke Za Zhi.* **2020**, *59*, E012.
(36) Lian, N.; Xie, H.; Lin, S.; Huang, J.; Zhao, J.; Lin, Q. Umifenovir treatment is not associated with improved outcomes in patients with coronavirus disease 2019: a retrospective study. *Clin. Microbiol. Infect.* **2020**, *26*, 917−921.
(37) Teissier, E.; Zandomeneghi, G.; Loquet, A.; Lavillette, D.; Lavergne, J.-P.; Montserret, R.; Cosset, F. L.; Böckmann, A.; Meier, B. H.; Penin, F.; Pécheur, E.-I. Mechanism of inhibition of enveloped virus membrane fusion by the antiviral drug arbidol. *PLoS One* **2011**, *6*, No. e15874.
(38) Brancato, V.; Peduto, A.; Wharton, S.; Martin, S.; More, V.; Di Mola, A.; Massa, A.; Perfetto, B.; Donnarumma, G.; Schiraldi, C.; Tufano, M.-A.; de Rosa, M.; Filosa, R.; Hay, A. Design of inhibitors of influenza virus membrane fusion: synthesis, structure-activity relationship and in vitro antiviral activity of a novel indole series. *Antiviral Res.* **2013**, *99*, 125−135.
(39) Nasser, Z. H.; Swaminathan, K.; Müller, P.; Downard, K. M. Inhibition of influenza hemagglutinin with the antiviral inhibitor arbidol using a proteomics based approach and mass spectrometry. *Antiviral Res.* **2013**, *100*, 399−406.
(40) Kadam, R. U.; Wilson, I. A. Structural basis of influenza virus fusion inhibition by the antiviral drug Arbidol. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 206−214.
(41) Leneva, I. A.; Russell, R. J.; Boriskin, Y. S.; Hay, A. J. Characteristics of arbidol-resistant mutants of influenza virus: implications for the mechanism of anti-influenza action of arbidol. *Antiviral Res.* **2009**, *81*, 132−140.
(42) Freidel, M. R.; Armen, R. S. Mapping major SARS-CoV-2 drug targets and assessment of druggability using computational fragment screening: Identification of an allosteric small-molecule binding site on the Nsp13 helicase. *PLoS One* **2021**, *16*, No. e0246181.
(43) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(44) Rahaman, O.; Estrada, T. P.; Doren, D. J.; Taufer, M.; Brooks, C. B., 3rd; Armen, R. S. Evaluation of several two-step scoring functions based on linear interaction energy, effective ligand size, and empirical pair potentials for prediction of protein-ligand binding geometry and free energy. *J. Chem. Inf. Model.* **2011**, *51*, 2047−2065.
(45) Armen, R. S.; Chen, J.; Brooks, C. B., 3rd An Evaluation of Explicit Receptor Flexibility in Molecular Docking Using Molecular Dynamics and Torsion Angle Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2909−2923.
(46) Momany, F. A.; Rone, R. Validation of the general purpose QUANTA® 3.2/CHARMm® force field. *J. Comput. Chem.* **1992**, *13*, 888−900.
(47) Roche, O.; Kiyama, R.; Brooks, C. L., 3rd. Ligand-protein database: linking protein-ligand complex structures to binding data. *J. Med. Chem.* **2001**, *44*, 3592−3598.
(48) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L. 3rd.; New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. *J. Comput. Chem.* **2003**, *24*, 1348.
(49) Feig, M.; Im, W.; Brooks, C. L., 3rd Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Chem. Phys.* **2004**, *2*, 903.
(50) Available at: http://www.chemaxon.com
(51) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032−3047.
(52) Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L. Assessing energy functions for fexible docking. *J. Comput. Chem.* **1998**, *19*, 1612−1622.
(53) 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*, 281−292.e6.
(54) Li, H.; Cheng, C.; Li, S.; Wu, Y.; Liu, Z.; Liu, M.; Chen, J.; Zhong, Q.; Zhang, X.; Liu, S.; Song, G. Discovery and structural optimization of 3-O-β-chacotriosyl oleanane-type triterpenoids as potent entry inhibitors of SARS-CoV-2 virus infections. *Eur. J. Med. Chem.* **2021**, *215*, No. 113242.
(55) Zhou, T.; Tsybovsky, Y.; Olia, A. S.; Gorman, J.; Rapp, M. A.; Cerutti, G.; Chuang, G.-Y.; Katsamba, P. S.; Nazzari, A.; Sampson, J. M.; Schon, A.; Wang, P. D.; Bimela, J.; Shi, W.; Teng, I.-T.; Zhang, B.; Boyington, J. C.; Sastry, M.; Stephens, T.; Stuckey, J.; Wang, S.; Friesner, R. A.; Ho, D. D.; Mascola, J. R.; Shapiro, L.; Kwong, P. D. A pH-dependent switch mediates conformational masking of SARS-CoV-2 spike. *bioRxiv.* **2020**, *7*, No. 187989.
(56) Gottschalk, P. G.; Dunn, J. R. The five-parameter logistic: a characterization and comparison with the four-parameter logistic. *Anal. Biochem.* **2005**, *343*, 54−65.
(57) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenbtttlatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(58) Li, Y.; Cao, L.; Li, G.; Cong, F.; Li, Y.; Sun, J.; Luo, Y.; Chen, G.; Li, G.; Wang, P.; Xing, F.; Ji, X.; Zhao, J.; Zhang, Y.; Guo, D.; Zhang, X. Remdesivir Metabolite GS-441524 Effectively Inhibits SARS-CoV-2 Infection in Mouse Models. *J. Med. Chem.* **2021**, No. 0c01929.
(59) Ni, X.; Schröder, M.; Olieric, V.; Sharpe, M. E.; Hernandez-Olmos, V.; Proschak, E.; Merk, D.; Knapp, S.; Chaikuad, A. Structural Insights into Plasticity and Discovery of Remdesivir Metabolite GS-441524 Binding in SARS-CoV-2 Macrodomain. *ACS Med. Chem. Lett.* **2021**, *12*, 603−609.

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922

(60) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **2019**, *47*, D1102−D1109.
(61) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein- ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−806.
(62) Sun, Y.; He, X.; Qiu, F.; Zhu, X.; Zhao, M.; Li-Ling, J.; Su, X.; Zhao, L. Pharmacokinetics of single and multiple oral doses of arbidol in healthy Chinese volunteers. *Int. J. Clin. Pharmacol. Ther.* **2013**, *51*, 423−432.
(63) 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*, 1165−1175.
(64) Vankadari, N. Arbidol: A potential antiviral drug for the treatment of SARS-CoV-2 by blocking trimerization of the spike glycoprotein. *Int. J. Antimicrob. Agents* **2020**, *56*, No. 105998.
(65) Trigueiro-Louro, J.; Correia, V.; Figueiredo-Nunes, I.; Gíria, M.; Rebelo-de-Andrade, H. Unlocking COVID therapeutic targets: A structure-based rationale against SARS-CoV-2, SARS-CoV and MERS-CoV Spike. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2117−2131.
(66) Sellitto, G.; Faruolo, A.; de Caprariis, P.; Altamura, S.; Paonessa, G.; Ciliberto, G. Synthesis and anti-hepatitis C virus activity of novel ethyl 1H-indole-3-carboxylates in vitro. *Bioorg. Med. Chem.* **2010**, *18*, 6143−6148.
(67) Zhao, C.; Zhao, Y.; Chai, H.; Gong, P. Synthesis and in vitro anti-hepatitis B virus activities of some ethyl 5-hydroxy-1H-indole-3-carboxylates. *Bioorg. Med. Chem.* **2006**, *14*, 2552−2558.

https://doi.org/10.1021/acs.jcim.1c01061
J. Chem. Inf. Model. 2021, 61, 5906−5922