Supplementary Information for

Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2

Mohammad M. Ghahremanpour,[†] Julian Tirado-Rives,[†] Maya Deshmukh,[‡] Joseph A. Ippolito,[†‡] Chun-Hui Zhang,[†] Israel Cabeza de Vaca,[†] Maria-Elena Liosi, [†] Karen S. Anderson,[‡,#,*] and William L. Jorgensen[†,*]

[†]*Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107,*

[‡]*Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520-8066, and* [#]*Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06520-8066*

Contents

Details for the Docking Calculations

Glide. Schrödinger's Protein Prep wizard utility was used for preparing the protein. A 20-Å grid was then generated and centered on the co-crystallized ligand, which was subsequently removed. The drug library members were neutralized and/or ionized via Schrödinger's LigPrep.[1] The Epik program[2] was used for estimating the p*K*a values of each compound. Plausible tautomers and stereoisomers within the pH range of 7 ± 1 were generated for each compound using the OPLS3 force field.[3] These conditions resulted in a total of 16000 structures, which were then docked into M^{pro} using Schrödinger's standard-precision (SP) Glide.[4,5]

AutoDock. **T**he AutoDockTools (ADT) software[6] was used for creating PDBQT files from SDF and PDB files of compounds and the protein, respectively. Non-polar hydrogen atoms were removed and Gasteiger–Marsili charges were assigned for both the protein and the ligands using ADT. The AutoGrid 4.2 program[6] was used for generating affinity grids with a spacing of 0.375 Å and with a box size of 74 × 80 × 62 Å. The affinity grids were centered at two different points of the active site for performing two sets of runs. In the first run, the grid box was centered at C_β of Cys145 of monomer A. In the second run, the grid center was displaced toward the geometric center of the active site. The AutoDock 4.2 program[6] was applied for docking the ligands into M^{pro}. The Lamarckian genetic algorithm (LGA) was used for ligand conformational searching. LGA was iterated 15 and 50 times in the first and the second run, respectively, for each compound. The maximum RMS tolerance for conformational cluster analysis was 2.0 and 0.5 Å in the first and second runs, respectively. The number of generations was set to 27000 with 300 individuals in each population in both runs. The maximum number of energy evaluations was 30×10^6 for all compounds and 40×10^6 for re-docking of the selected consensus compounds. Other parameters were set to their default values.

AutoDockVina. The PDBQT files generated by ADT for the protein and library compounds also used for running AutoDock Vina.[7] Non-polar hydrogen atoms were removed. An affinity grid box with a size of 18 × 21 × 18 Å was generated and centered on the active site. The default docking parameters were used, except for the number of modes that was set to 9.

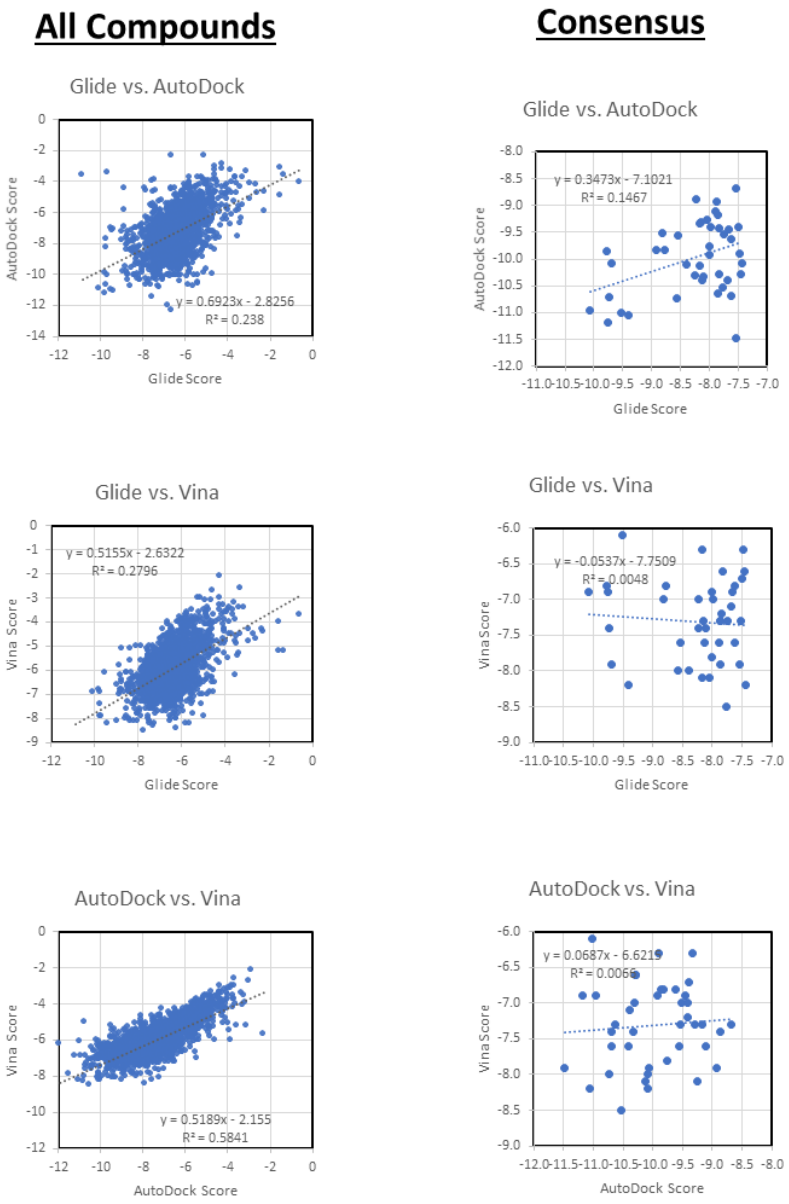

Figure S2. Correlation of docking scores for all drugs (left) and for the 42 consensus compounds (right).

Molecular Dynamics Simulations. The GROMACS software, version 2018a compiled in double precision, was used for performing all molecular dynamics (MD) simulations.[8] The protonated M^{pro} dimer, with a net charge of -8 *e*, was represented by the OPLS-AA/M force field.[9] TIP4P water was used as the solvent.[10] Sodium counterions were added to neutralize the net charge of each system. The selected ligand candidates were represented by the OPLS/CM1A force field,[11] as assigned by the BOSS software[12] (version 4.9) and the LigParGen Python code.[13] The parameters were converted to GROMACS format using LigParGen.[13] For neutral ligands, the CM1A partial atomic charges were scaled by a factor of 1.14.[11]

Each M^{pro}-ligand complex was put at the center of a triclinic simulation box with 10-Å padding. An energy minimization was then performed until the steepest descent algorithm converged to a maximum force smaller than 2.4 kcal mol^{-1} $Å^{-1}$. A cutoff radius of 12 Å was used to explicitly calculate non-bonded interactions. Long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) algorithm.[14] The PME was used with an interpolation order of 4, a Fourier spacing of 1.2 Å, and a relative tolerance of 10^{-6}. The van der Waals forces were smoothly switched to zero between 10 and 12 Å. Analytical corrections to the long-range effect of dispersion interactions were applied to both energy and pressure. All covalent bonds to hydrogen atoms were constrained at their equilibrium lengths using the LINCS algorithm.[15] Each system was initially simulated for 1 ns in the canonical ensemble (*NVT*) in order for the solvent to relax and the temperature to equilibrate. Initial velocities were sampled from a Maxwell-Boltzmann distribution at 310 K. The V-rescale thermostat with a stochastic term[16] was used for keeping the temperature at 310 K. The coupling constant of the thermostat was set to 2.0 ps. The system was then equilibrated for 1.5 ns in the isothermal-isobaric ensemble (*NPT*) for obtaining a density consistent with the reference pressure. The pressure was kept at 1 bar by the Berendsen

barostat[17] with a coupling constant of 4.0 ps and a compressibility factor of 4.5×10^{-5} bar. A harmonic position restraint with a force constant of 2.4 kcal mol^{-1} $Å^{-2}$ was applied to the protein backbone and to all solute heavy atoms during the equilibration steps. A 70 ns unrestrained run was then performed in the NPT ensemble with the Parrinello−Rahman barostat[18] using a coupling time of 4.0 ps.

MD Analyses for the M^{pro}-ligand Complexes. Before the assaying was carried out, the 70-ns MD simulations were run for complexes of 14 of the promising compounds starting from the Glide poses. The idea was to obtain insight on which compounds gave more stable complexes and were, therefore expected to be more active inhibitors. In addition to visualization of the evolving structures, the all-atom RMSD of each ligand was computed over the course of the simulation time with and without least-square (LS) fitting of the ligand's atoms onto the initial structure of the ligand in the complex (Figure S3). The LS-fit RMSD monitors only the ligand's conformational changes, whereas the no-fit RMSD also reflects rotational and translational movements away from the starting structure. The LS-fit RMSD converged relatively quickly to 2-3 Å for all ligands, except for carindacillin, which converged to 4 Å. However, as expected, no-fit RMSD values are larger than the LS-fit RMSD values for all ligands, demonstrating the contribution of rotational and translational movements. The no-fit RMSD value converged for bedaquiline, idarubicin, indinavir, and perampanel after about 10 ns, while it converged for efonidpine after 60 ns. The no-fit RMSD values steadily fluctuated about an average value of 4 Å for lapatinib and periciazine. In general, most ligands showed some displacement from their initial position, whereas they remained close to their initial conformation. Differences between initial and final structures were considered in selecting compounds for purchase and assay.

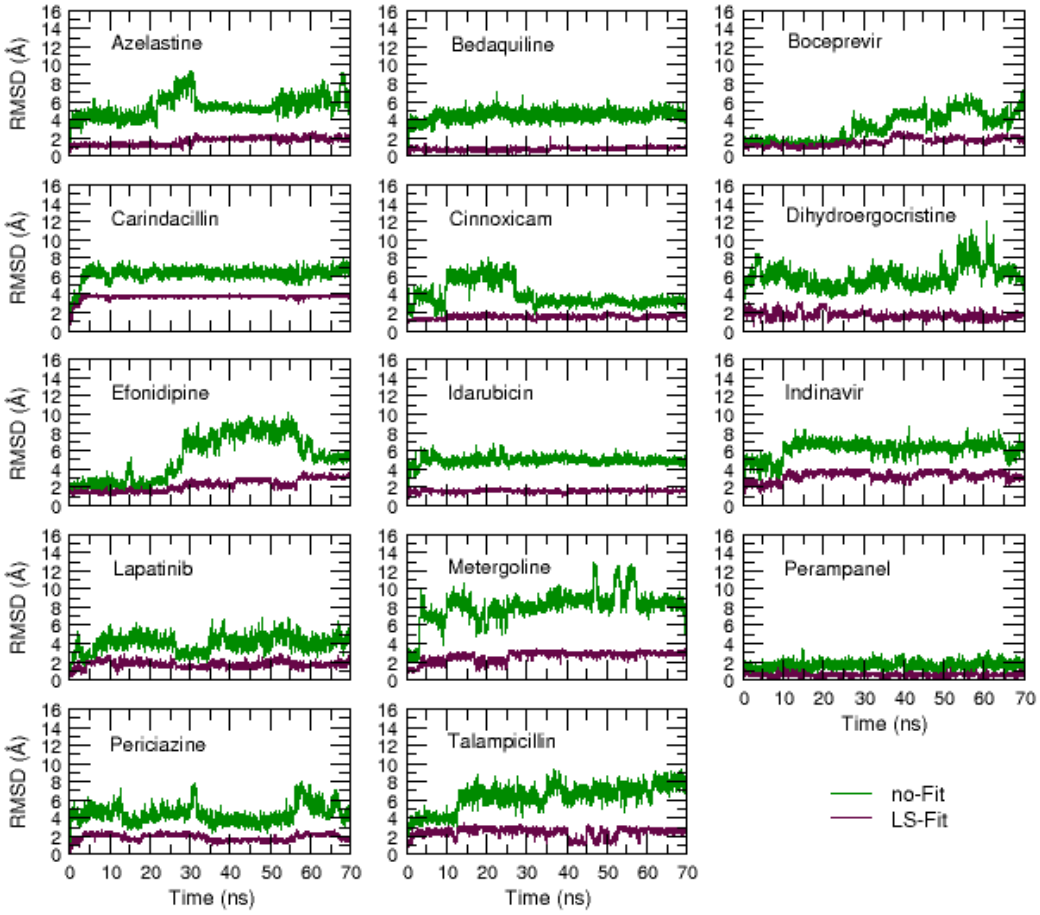

Figure S2. RMSD in Å of all ligand atoms with and without Least-Square fitting to the original complex structure during the course of 70-ns MD simulations.

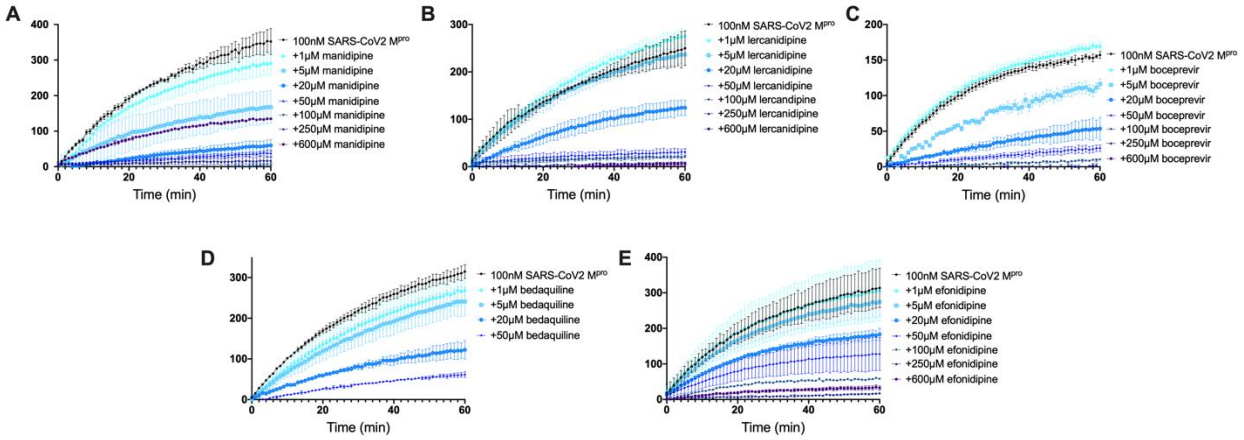

Figure S3. Kinetic data measuring activity of SARS-CoV2 M^{pro} in the presence of A) manidipine, B) lercanidipine, C) boceprevir, D) bedaquiline, and E) efonidipine. All measurements were performed in triplicate, averaged, and plotted with standard deviation. Kinetic data for inhibition of M^{pro} by bedaquiline was collected at concentrations up to 50 μM due to solubility issues.

References

(1) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. Towards the Comprehensive, Rapid, and Accurate Prediction of the Favorable Tautomeric States of Drug-like Molecules in Aqueous Solution. *J. Comput. Aided Mol. Des.* **2010**, *24*, 591–604.

(2) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: A Software Program for pK_a Prediction and Protonation State Generation for Drug-like Molecules. *J. Comput. Aided Mol. Des.* **2007**, *21*, 681–691.

(3) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296.

(4) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.

(5) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.

(6) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(7) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, 31, 455-460.

(8) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(9) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499–3509.

(10) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(11) Jorgensen, W. L.; Tirado-Rives, J. Potential Energy Functions for Atomic-Level Simulations of Water and Organic and Biomolecular Systems. *Proc. Natl. Acad. Sci.* **2005**, *102*, 6665–6670.

(12) Jorgensen, W. L.; Tirado-Rives, J. Molecular Modeling of Organic and Biomolecular Systems Using BOSS and MCPRO. *J. Comput. Chem.* **2005**, *26*, 1689–1700.

(13) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands. *Nucleic Acids Res.* **2017**, *45*, 331–336.

(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(15) Hess, B.; Bekker, H.; Berendsen, H. J. C. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.

(16) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(17) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(18) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.