Supporting Information

Investigation of the Drug Resistance Mechanism of M2-S31N Channel Blockers through Biomolecular Simulations and Viral Passage Experiments

Rami Musharrafieh,[†, §,#] Panagiotis Lagarias,[‡,#] Chunlong Ma,[†] Raymond Hau,[†] Alex Romano,[†] George Lambrinidis, [‡] Antonios Kolocouris,[*, ‡] Jun Wang[*, †]

[†]Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, Arizona 85721, United States

[§]Department of Chemistry and Biochemistry, The University of Arizona, Tucson, Arizona 85721, United States

[‡]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens, Greece

Supplementary table of contents:

1. Protein Preparation - Docking Calculations. Structure of **3** was built with Schrodinger 2017-1 platform [1] and minimized by the conjugate gradient method using the MMFF94 force field [2] and a distance-dependent dielectric constant of 4.0 until a convergence threshold of 0.0001 kJ mol^{-1} $Å^{-1}$ was reached. The PDB ID 2L0J was used to model mutant M2TMC (22-62) apo proteins and the "protein preparation" tool of Maestro was applied. N- and C-termini of the PDB ID 2L0J was capped by acetyl and methylamino groups. After applying the protein preparation module of Maestro, [1, 3] all hydrogens of the protein complex were minimized with the AMBER* force field by means of Maestro/Macromodel 9.6 using a distance-dependent dielectric constant of 4.0. The molecular mechanics minimizations were performed with a conjugate gradient (CG) method and a threshold value of 0.0001 kJ $Å^{-1}$ mol^{-1}as the convergence criterion. Mutant M2TMC-S31N protein was obtained by manual mutation of S31 to N31. Similarly, the five double-mutant apo proteins M2TMC-S31N/V27A, M2TMC-S31N/V27F, M2TMC-S31N/I32N, M2TMC-S31N/G34E, M2TMC-S31N/R45H protein were obtained by applying the S31N mutation and additionally mutation V27 to A, V27 to F, I32 to N, G34 to E, R45 to H respectively.[4] MD simulations of the five mutant proteins M2TMC-S31N/V27A, M2TMC-S31N/V27F, M2TMC-S31N/I32N, M2TMC-S31N/G34E, M2TMC-S31N/R45H in hydrated POPC for 300 ns produced well equilibrated *apo* proteins.

The simulated M2TMC-S31N *apo* protein was superimposed with M2TM (19-49) S31N in complex with **M2WJ332** (**2**) PDB ID 2LY0 [5] which after deletion of M2TM (19-49) S31N resulted in complex of M2-S31N with **M2WJ332** (**2**). The structure of the protein M2-S31N and ligand **M2WJ332** (**2**) were saved separately and were used for the subsequent docking calculations of ligand **3** to M2TMC-S31N with GOLD 5.2, [6] using the GoldScore scoring function. [7] The region of interest used by GOLD was defined to contain the atoms that were within ~15 Å of **3** binding

site in the structure. For all the calculations with GOLD 5.2 default values were used, apart from the "allow early termination" that was not applied. Ligand **3** was submitted to 150 genetic algorithm runs. Ten docking poses were produced which were visually inspected using the UCSF Chimera package and Schrodinger Maestro 2017-1. [1, 8] Finally, the docking pose with the best GoldScore score was used for the subsequent MD simulations.

The five double-mutant complexes M2TMC-S31N/V27A, M2TMC-S31N/V27F, M2TMC-S31N/I32N, M2TMC-S31N /G34E, M2TMC-S31N/R45H protein with **M2WJ332** (**2**) were obtained using M2TMC-S31N with **M2WJ332** (**2**) and applying manual mutation of V27 to A, V27 to F, I32 to N, G34 to E, R45 to H using the more likely side chain rotamer according to the x-ray structure of M2TM complexes with ligands. [4] The complexes of **3** with the six simulated apo proteins M2TMC-S31N/V27A, M2TMC-S31N/V27F, M2TMC-S31N/I32N, M2TMC-S31N/G34E, M2TMC-S31N/R45H were prepared by superimposing them with M2TM (19-49) S31N in complex with **M2WJ332** (**2**) (PDB ID 2LY0). [5] After deletion of M2TM (19-49) S31N, structures of the six mutated M2-S31N proteins in complex with **M2WJ332** (**2**) were obtained. The complexes of **3** with the six M2TMC-S31N proteins were prepared from docking calculations as described above.

2. MD Simulations

Systems Setup. The apo proteins M2TMC-S31N/V27A, M2TMC-S31N/V27F, M2TMC-S31N/I32N, M2TMC-S31N/G34E, M2TMC-S31N/R45H or their complexes with **3** were embedded in a hydrated 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) bilayer extending 10 x 10 x 15 $Å^3$ in XYZ directions from M2. For the membrane insertion, neutralization, the System Builder module of Desmond was used.[9-10] Periodic boundary conditions were applied

in 75 x 75 x 95 $Å^3$, including 107 POPC lipids. A 15 Å area above and below the protein included 8960 waters. The total number of atoms included the protein inside the hydrated bilayer was ~ 42000 atoms. The systems were neutralized by adding Na^+ ions. A 0.150 M NaCl solution was added adjusted using the System Builder utility of Desmond.[11-13] Ligand **3** was positively charged (+1), using Schrodinger 2017-1 Maestro platform. The H37 residues of M2TM were protonated at Nε2 (HIE atom type) because this form was found to be most populated in M2TM and four uncharged H37 residues were applied.[14-15]

3. MD Simulations Protocol. The MD simulations were performed using Desmond v11.1 software [11, 16] using the OPLS2005 force field[17-18] and the TIP3P [13] as the water model. The orthorhombic periodic box boundaries were set 15 Å away from the protein atoms.

For the MD simulations with amber99sb force field Desmond Viparr tool was used to assign amber99sb force field parameters [19-20] to protein and lipid, and GAFF[21] force field parameters for the ligand. Initial ligand parameters were constructed with the antechamber module of AMBER14 software package.[22] Desmond Viparr tool was also used to assign CHARMM36 parameters in protein, lipids and ions. [23,24] For simulations with CHARMM36 the corresponding implementation of TIP3P was used through Viparr and ligand parameters were assigned with the SwissParam online utility, i.e., the dihedral angle terms, harmonic part of the bond, angle and improper terms, and charges were taken from the MMFF [2] and van der Waals parameters were taken from the closest atom type in CHARMM22.[25] According to the literature,[18] very recent unpublished work and analysis of the results in this work we concluded that OPLS2005 is more adequate for the calculations of ligand-protein interactions and the results with this force field are included in the Results and Discussion.

Particle mesh Ewald (PME) was employed to calculate long-range electrostatic interactions [26-27] with a grid spacing of 0.8 Å. The Shake method was used to keep all bonds with hydrogen rigid, at ideal lengths and angles.[28] Van der Waals and short-range electrostatic interactions were smoothly truncated at 9.0 Å. Langevin dynamics were utilized for temperature control establishing a Langevin collision frequency of 1.0 ps^{-1}.[29] Pressure to $p = 1$ bar was applied by a pressure relaxation time of 1.0 ps. The Martyna-Tobias-Klein method [30] was used for pressure control. The equations of motion were integrated using the multistep RESPA integrator [31] with an inner time step of 2.0 fs for bonded interactions and non-bonded interactions within the cutoff. An outer time step of 6.0 fs was used for non-bonded interactions beyond the cut-off.

For all MD simulations the same relaxation protocol was used. In the first run, 250 steps steepest descent and 750 steps conjugate gradient minimization were carried out in the presence of a harmonic restraint with a force constant of 500 kcal mol^{-1} $Å^{-2}$ on all protein, ions and membrane atoms. In the second run a minimization of the membrane was conducted by 2500 steps steepest descent and 7500 steps conjugate gradient minimization in the presence of a harmonic restraint with a force constant of 5 kcal mol^{-1} $Å^{-2}$ on all protein and ligand atoms. MD simulations were performed at 310 K in order to ensure that the membrane state is above the main phase transition temperature of 298 K for POPC bilayers.[32] Then, a series of MD simulations was applied. The first simulation was performed for 50 ps starting at 100 K with gradual heating until 308 K in the NVT ensemble with the solute heavy atoms restrained with a force constant of 50 kcal mol^{-1} $Å^{-2}$. The heating was followed by two NPT MD 1 ns equilibration runs. One for 1 ns simulation in the NPT ensemble with the solute heavy atoms restrained with a force constant of 10 kcal mol^{-1} $Å^{-2}$ to equilibrate solvent and lipids. In the next 1 ns MD simulation the harmonic constraints were gradually decreased from 10.0 to 2.0 kcal mol^{-1} $Å^{-2}$ on solute heavy atoms following with removal

of all harmonic constraints except a 2.0 kcal mol^{-1} $Å^{-2}$ set on protein C_α atoms. This equilibration protocol was followed by 100 ns NPT simulation without restraints. The replicas of the system were saved every 15 ps. Within this simulation time, the total energy and RMSD of the protein backbone Cα carbons[33] reached a plateau, and the systems were considered equilibrated and suitable for statistical analysis. For the calculation of protein-lipid hydrogen bonds, a cutoff angle of 20º between the donor-hydrogen-acceptor atoms and a cutoff distance of 3.2 Å between the donor and acceptor atoms were applied. The replicas were created with Maestro's implementation of PyMol and VMD [1, 34] and UCSF Chimera package.[8] MD simulations were run in workstations and ARIS supercomputer using the GPU implementation and parallel CPU algorithms of MD simulations codes as provided by Desmond.[35]

4. Analysis of MD Simulations Trajectories. The visualization of produced trajectories was performed using the GUI of Maestro [16] and the protein-ligand interaction analysis was done with the Simulation Interaction Diagram (SID) tool, available with Schrodinger Desmond v. 11.1.[11, 16] For hydrogen bond interactions, a distance of 2.5 Å between donor and acceptor heavy atoms, and an angle ≥120º between donor-hydrogen-acceptor atoms and ≥ 90º between hydrogen-acceptor-bonded atom were considered. Non-specific hydrophobic contacts were identified when the side chain of a hydrophobic residue fall within 3.6 Å from a ligand's aromatic or aliphatic carbon. Water-mediated interactions were characterized by a distance of 2.7 Å between donor and acceptor atoms, as well as an angle ≥ 110º between donor-hydrogen-acceptor atoms and ≥ 80º between hydrogen-acceptor-bonded atom. The visualization of produced trajectories and structures was performed using the programs Chimera[8] and VMD.[34]

5. MM-GBSA Calculations. The effective binding free energies (ΔG_{eff}) of the complexes between compound **3** and M2-S31N proteins were computed using the OPLS2005 force field considering the gas phase energy and solvation free energy contributions to binding [36] using the 1-trajectory MM-GBSA approach. [37-39] For this, structural ensembles for each complex were extracted every 50 ps from the last 50 ns of the production simulations. Prior to the calculations all water molecules, ions, and lipids were removed, and the structures were positioned such that the geometric center of each complex was located at the coordinate origin. The module in Schrodinger Suite was used, i.e. the thermal_mmgbsa.py script that takes individual trajectory snapshots and calculates ΔG_{eff}.

References

1. Schrödinger Release 2017-1: Maestro, Schrödinger, LLC, New York, NY, 2017.

2. Halgren, T. A., Merck Molecular Force Field. 1. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *Journal of Computational Chemistry* **1996,** *17* (5-6), 490-519.

3. Maestro, Version 10.5; Schrodinger, Inc.: New York, NY, 2015.

4. Thomaston, J. L.; DeGrado, W. F., Crystal Structure of the Drug-Resistant S31N Influenza M2 Proton Channel. *Protein Science* **2016,** *25* (8), 1551-1554.

5. Wang, J.; Wu, Y. B.; Ma, C. L.; Fiorin, G.; Wang, J. Z.; Pinto, L. H.; Lamb, R. A.; Klein, M. L.; DeGrado, W. F., Structure and Inhibition of the Drug-Resistant S31N Mutant of the M2 Ion Channel of Influenza A Virus. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (4), 1315-1320.

6. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and Validation of a Genetic Algorithm for Flexible Docking. *Journal of Molecular Biology* **1997,** *267* (3), 727-748.

7. Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R., Modeling Water Molecules in Protein-Ligand Docking Using GOLD. *Journal of Medicinal Chemistry* **2005,** *48* (20), 6504-6515.

8. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry* **2004,** *25* (13), 1605-1612.

9. Schrödinger Release 2018-1: Maestro, Schrödinger, LLC, New York, NY, 2018.

10. Desmond Molecular Dynamics System, Version 3.0, D.E. Shaw Research, New York, NY, 2011.

11. Desmond Molecular Dynamics System, Version 3.0, D.E. Shaw Research: New York, 2011. *Desmond Mol. Dyn. Syst. version 3.0; D.E. Shaw Res. New York, 2011; Maest. Interoperability Tools, 3.1; Schrodinger Res. New York, 2012*.

12. Bowers, K. J. S., F. D. Salmon, J. K. Shan, Y.; Shaw, D. E. C., E. Xu, H. Dror, R. O. Eastwood, M. P. Gregersen, B. A. Klepeis, J. L. Kolossvary, I. Moraes, M. A. In *Molecular Dynamics---Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters.* , Proceedings of the 2006 ACM/IEEE conference on Supercomputing, SC, ACM Press: New York, New York, USA: SC, 2006; p 84.

13. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* **1983,** *79* (2), 926-935.

14. Sharma, M.; Yi, M. G.; Dong, H.; Qin, H. J.; Peterson, E.; Busath, D. D.; Zhou, H. X.; Cross, T. A., Insight into the Mechanism of the Influenza A Proton Channel from a Structure in a Lipid Bilayer. *Science* **2010,** *330* (6003), 509-512.

15. Hu, F. H.; Luo, W. B.; Hong, M., Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science* **2010,** *330* (6003), 505-508.

16. Maestro-Desmond Interoperability Tools, Version 3.1; Schrodinger: New York, 2012.

17. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *Journal of Physical Chemistry B* **2001,** *105* (28), 6474-6487.

18. Gumbart, J. C.; Roux, B.; Chipot, C., Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *Journal of Chemical Theory and Computation* **2013,** *9* (1), 794-802.

19. Wang, J. M.; Cieplak, P.; Kollman, P. A., How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *Journal of Computational Chemistry* **2000,** *21* (12), 1049-1074.

20. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins-Structure Function and Bioinformatics* **2006,** *65* (3), 712-725.

21. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.

22. D.A. Case J.T. Berryman, R. M. B., Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M, V. B. *The Amber Molecular Dynamics Package*, AMBER: 2014.

23. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. *Journal of Chemical Theory and Computation* **2012,** *8* (9), 3257-3273.

24. Huang, J.; MacKerell, A. D., CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *Journal of Computational Chemistry* **2013,** *34* (25), 2135-2145.

25. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O., SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *Journal of Computational Chemistry* **2011,** *32* (11), 2359-2368.

26. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **1993,** *98* (12), 10089-10092.

27. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.

28. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *Journal of Computational Physics* **1977,** *23* (3), 327-341.

29. Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R., Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *Journal of Chemical Physics* **1995,** *103* (11), 4613-4621.

30. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant Pressure Molecular Dynamics Algorithms. *Journal of Chemical Physics* **1994,** *101* (5), 4177-4189.

31. Humphreys, D. D.; Friesner, R. A.; Berne, B. J., A Multiple-Time-Step Molecular Dynamics Algorithm for Macromolecules. *Journal of Physical Chemistry* **1994,** *98* (27), 6885-6892.

32. Koynova, R.; Caffrey, M., Phases and Phase Transitions of the Phosphatidylcholines. *Biochimica Et Biophysica Acta-Biomembranes* **1998,** *1376* (1), 91.

33. Lyman, E.; Zuckerman, D. M., Ensemble-Based Convergence Analysis of Biomolecular Trajectories. *Biophysical Journal* **2006,** *91* (1), 164-172.

34. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics & Modelling* **1996,** *14* (1), 33-38.

35. Schrödinger Release 2017-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2017. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2017.

36. Gohlke, H.; Case, D. A., Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein-Protein Complex Ras-Raf. *Journal of Computational Chemistry* **2004,** *25* (2), 238-250.

37. Homeyer, N.; Gohlke, H., Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Molecular Informatics* **2012,** *31* (2), 114-122.

38. Homeyer, N.; Stoll, F.; Hillisch, A.; Gohlke, H., Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *Journal of Chemical Theory and Computation* **2014,** *10* (8), 3331-3344.

39. Massova, I.; Kollman, P. A., Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspectives in Drug Discovery and Design* **2000,** *18*, 113-135.