# Supporting Information for:

## Identification of Zika Virus Inhibitors Using Homology Modeling and Similarity-based Screening to Target Glycoprotein E

Stephen M. Telehany,[a,§] Monica S. Humby,[b,§] T. Dwight McGee Jr.,[c] Sean P. Riley,[b] Amy Jacobs,[b,*] and Robert C. Rizzo[c,d,e,*]

[a] Department of Chemistry, Stony Brook University, Stony Brook, NY 11794

[b] Department of Microbiology and Immunology, School of Medicine and Biomedical Sciences, State University of New York (SUNY) at Buffalo, Buffalo, NY

[c] Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook, NY 11794, USA

[d] Institute of Chemical Biology & Drug Discovery, Stony Brook University, Stony Brook, NY 11794

[e] Laufer Center for Physical & Quantitative Biology, Stony Brook University, Stony Brook, NY 11794

§ These authors contributed equally to this work

* Corresponding author e-mails: ajacobs2@buffalo.edu (AJ), rizzorc@gmail.com (RCR)

**Molecular Dynamics Simulations.** Molecular Dynamics (MD) simulations employed the program AMBER[1] in conjunction with ff14SB[2] (protein), TIP3P[3] (solvent) and GAFF[4] plus AM1BCC[5] (ligand) force fields. Sodium counter ions were added to each simulation setup to neutralize the charge, and each protein-ligand complex was equilibrated using a nine-step

procedure described as follows: (1) 5,000 steps of energy minimization with 5.0 kcal $mol^{-1}$ $Å^{-2}$ restraints, (2) 50,000 steps of MD with 5.0 kcal $mol^{-1}$ $Å^{-2}$ restraints on the non-hydrogen atoms of the ligand and protein, (3) 1,000 steps of minimization with 2.0 kcal $mol^{-1}$ $Å^{-2}$ restraints, (4) 1,000 steps of minimization with 0.1 kcal $mol^{-1}$ $Å^{-2}$ restraints, (5) 1,000 steps of minimization with 0.05 kcal $mol^{-1}$ $Å^{-2}$ restraints, (6) 50,000 steps of MD with 1.0 kcal $mol^{-1}$ $Å^{-2}$ restraints on the non-hydrogen atoms of the ligand and protein, (7) 50,000 steps of MD with 0.5 kcal $mol^{-1}$ $Å^{-2}$ restraints on the non-hydrogen atoms of the ligand and protein, (8) 50,000 steps of MD with 0.5 kcal $mol^{-1}$ $Å^{-2}$ restraints on just the backbone carbons and nitrogens of the protein, and (9) 50,000 steps of MD with 0.1 kcal $mol^{-1}$ $Å^{-2}$ restraints on just the backbone nitrogens and carbons. All minimization steps placed restraints on non-hydrogen atoms of the ligand and protein. All MD simulations were performed at constant pressure, using a timestep of 1fs and a temperature of 298.15K. Production simulations were comprised of 4 replicates for each ligand (20 ns each), employed a 2-femtosecond time step, and were run with relatively weak (0.1 kcal $mol^{-1}$ $Å^{-2}$) restraints on the protein backbone. MM-GBSA[6] calculations (single trajectory method without additional entropy terms) were employed to estimate relative binding affinities for hits, and RMSDs were used to gauge their geometric stability over the course of the simulations. The MD trajectories were visualized using the program VMD.[7]

**Free Energy Calculations.** Table S1 shows MD-derived results from replicate simulations of docked ligands complexed with ZIKV E at various block sizes. Data tabulated includes heavy atom (non-hydrogen) root mean square deviation (RMSD) for ligands relative to their original DOCK poses, free energies of binding ($\Delta G_{bind}$) estimated using the single trajectory MM-GBSA method without additional entropic terms,[8, 9] autocorrelation function in percent correlated data

(ACF %),[10, 11] and block-averaged standard error of the mean (BASEM)[10, 11] for ligands complexed with ZIKV E at various block sizes. For each complex, four independent MD simulations labeled #1-#4 were initiated using different random seeds. We have previously used the ACF/BASEM technique to estimate errors computed by the MM-GBSA method for MD simulations of ligands complexed with HIV gp41, HER2, and FABP5.[12-14] Based on the present results, the data (on average) is reasonable uncorrelated (~82%) at a block size of 200 frames which corresponds to N=20 blocks. At this block size, the "replicate-averaged" error in $\Delta G_{bind}$ ranges from 0.993 to 1.418 kcal/mol while individual MD trajectories yield uncertainties from 0.249 to 0.914 kcal/mol.

**Table S1.** Ligand root mean square deviation (RMSD), free energies of binding ($\Delta G_{bind}$), autocorrelation function in percent correlated data (ACF %), and block-averaged standard error of the mean (BASEM) for ligands complexed with ZIKV E at various block sizes.

| | | | 5 ps[a] block size 1 frame N blocks = 4000 | | 250 ps block size 50 frame N blocks = 80 | | 1000 ps block size 200 frame N blocks = 20 | | 2000 ps block size 400 frame N blocks = 10 | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Lig, MD rep #** | **RMSD[a]** | **$\Delta G_{bind}$ [b]** | **ACF %[c]** | **BASEM[d]** | **ACF %** | **BASEM** | **ACF %** | **BASEM** | **ACF %** | **BASEM** |
| BOG, #1 | 2.48 | -36.26 | 48.33 | 0.069 | 15.73 | 0.267 | 2.00 | 0.445 | 6.91 | 0.504 |
| BOG, #2 | 1.93 | -37.68 | 63.37 | 0.075 | 44.02 | 0.389 | 19.79 | 0.702 | 17.81 | 0.863 |
| BOG, #3 | 1.77 | -40.36 | 62.59 | 0.073 | 47.90 | 0.390 | 13.54 | 0.693 | 0.08 | 0.905 |
| BOG, #4 | 1.70 | -38.36 | 49.93 | 0.064 | 24.81 | 0.284 | 10.32 | 0.456 | 10.29 | 0.585 |
| BOG, average | 1.97 | -38.17 | 56.06 | 0.141 | 33.12 | 0.675 | 11.41 | 1.174 | 8.77 | 1.470 |
| | | | | | | | | | | |
| 8, #1 | 2.69 | -37.94 | 73.18 | 0.083 | 60.76 | 0.485 | 47.82 | 0.914 | 31.10 | 1.215 |
| 8,# 2 | 4.12 | -30.96 | 54.19 | 0.061 | 43.05 | 0.303 | 25.05 | 0.552 | 16.92 | 0.710 |
| 8, #3 | 5.50 | -26.23 | 81.95 | 0.073 | 61.81 | 0.450 | 32.40 | 0.852 | 8.02 | 1.078 |
| 8, #4 | 2.59 | -38.69 | 40.20 | 0.055 | 22.36 | 0.218 | 11.19 | 0.379 | -1.03 | 0.483 |
| 8, average | 3.81 | -33.46 | 62.38 | 0.138 | 47.00 | 0.760 | 29.12 | 1.418 | 13.75 | 1.837 |
| | | | | | | | | | | |
| 15, #1 | 3.93 | -27.24 | 70.04 | 0.068 | 43.05 | 0.368 | 23.06 | 0.646 | 26.93 | 0.800 |
| 15, #2 | 3.62 | -30.55 | 72.21 | 0.069 | 49.51 | 0.388 | 28.98 | 0.764 | 23.02 | 1.098 |
| 15, #3 | 3.21 | -34.95 | 43.23 | 0.056 | 28.01 | 0.237 | 11.87 | 0.419 | -10.09 | 0.551 |
| 15, #4 | 4.43 | -25.83 | 79.58 | 0.074 | 59.72 | 0.444 | 36.03 | 0.875 | 4.42 | 1.234 |
| 15, average | 3.80 | -29.64 | 66.27 | 0.134 | 45.07 | 0.734 | 24.99 | 1.394 | 11.07 | 1.916 |
| | | | | | | | | | | |
| 30, #1 | 5.98 | -27.24 | 60.29 | 0.062 | 26.15 | 0.298 | 13.19 | 0.462 | 6.26 | 0.582 |
| 30, #2 | 3.96 | -29.05 | 48.23 | 0.055 | 24.33 | 0.226 | 8.02 | 0.382 | -1.32 | 0.482 |
| 30, #3 | 3.80 | -23.90 | 60.98 | 0.060 | 38.81 | 0.299 | 18.59 | 0.541 | 6.50 | 0.709 |
| 30, #4 | 5.07 | -23.60 | 68.56 | 0.062 | 49.37 | 0.340 | 32.06 | 0.626 | 12.08 | 0.806 |
| 30, average | 4.67 | -25.96 | 59.52 | 0.120 | 34.67 | 0.587 | 17.97 | 1.022 | 5.88 | 1.313 |
| | | | | | | | | | | |
| 36, #1 | 5.51 | -24.76 | 67.07 | 0.058 | 44.04 | 0.307 | 14.32 | 0.550 | -7.39 | 0.540 |
| 36, #2 | 5.77 | -21.49 | 76.84 | 0.068 | 49.12 | 0.394 | 13.96 | 0.683 | 14.83 | 0.786 |
| 36, #3 | 5.86 | -23.30 | 55.92 | 0.048 | 33.08 | 0.236 | 8.38 | 0.395 | -10.28 | 0.441 |
| 36, #4 | 4.67 | -26.52 | 38.87 | 0.054 | 8.55 | 0.184 | 6.56 | 0.249 | -0.34 | 0.344 |
| 36, average | 5.46 | -24.02 | 59.68 | 0.115 | 33.70 | 0.582 | 10.81 | 0.993 | -0.80 | 1.106 |
| | | | | | | | | | | |
| 41, #1 | 5.54 | -22.97 | 50.78 | 0.063 | 24.02 | 0.275 | 5.13 | 0.436 | 7.37 | 0.541 |
| 41, #2 | 4.91 | -24.43 | 73.51 | 0.064 | 55.13 | 0.368 | 38.80 | 0.680 | 30.77 | 0.847 |
| 41, #3 | 4.61 | -26.24 | 64.86 | 0.056 | 42.90 | 0.293 | 22.10 | 0.518 | 5.41 | 0.721 |
| 41, #4 | 5.41 | -24.00 | 50.79 | 0.047 | 21.65 | 0.203 | 0.58 | 0.340 | -4.76 | 0.396 |
| 41, average | 5.10 | -24.41 | 59.99 | 0.116 | 35.93 | 0.581 | 16.65 | 1.018 | 9.70 | 1.299 |
| | | | | | | | | | | |
| 43, #1 | 3.39 | -29.86 | 63.38 | 0.068 | 43.70 | 0.354 | 24.58 | 0.670 | 3.96 | 0.856 |
| 43, #2 | 3.71 | -29.75 | 39.03 | 0.051 | 18.44 | 0.195 | 6.99 | 0.329 | 1.02 | 0.442 |
| 43, #3 | 2.57 | -29.65 | 77.42 | 0.081 | 58.83 | 0.477 | 30.34 | 0.885 | 9.03 | 1.157 |
| 43, #4 | 3.88 | -30.32 | 50.52 | 0.055 | 25.32 | 0.244 | 9.31 | 0.434 | -4.12 | 0.615 |
| 43, average | 3.37 | -29.90 | 57.59 | 0.130 | 36.57 | 0.671 | 17.81 | 1.236 | 2.47 | 1.626 |

[a]Ligand RMSD values in Å relative to the original DOCK pose (hits) or modeled pose (BOG) from four independent MD simulations each. [b]$\Delta G_{bind}$ (MM-GBSA method) in kcal/mol recorded every 5 ps which corresponds to one frame over 20 ns of production time (4000 frames total). [c]ACF in % of correlated data. [d]BASEM energies in kcal/mol with average values computed as square root of the sum of the squares.

**References**

[1] Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., 3rd , Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Ghoreishi, D., Gilson, M. K., Gohlke, H., Goetz, A. W., Greene, D., Harris, R., Homeyer, N., Izadi, S., Kovalenko, A., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Mermelstein, D. J., Merz, K. M., Miao, Y., Monard, G., Nguyen, C., Nguyen, H., Omelyan, I., Onufriev, A., Pan, F., Qi, R., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Smith, J., Salomon-Ferrer, R., Swails, J., Walker, R. C., Wang, J., Wei, H., Wolf, R. M., Wu, X., Xiao, L., York, D. M., and Kollman, P. A. (2018) AMBER 2018, University of California, San Francisco.

[2] Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J Chem Theory Comput 11*, 3696-3713.

[3] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J Chem Phys 79*, 926-935.

[4] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field, *J Comput Chem 25*, 1157-1174.

[5] Jakalian, A., Jack, D. B., and Bayly, C. I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J Comput Chem 23*, 1623-1641.

[6] Miller, B. R., 3rd, McGee, T. D., Jr., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations, *J Chem Theory Comput 8*, 3314-3321.

[7] Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J Mol Graph 14*, 33-38.

[8] Srinivasan, J., Cheatham, T. E., 3rd, Cieplak, P., Kollman, P. A., and Case, D. A. (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices, *JACS 120*, 9401-9409.

[9] Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E., 3rd. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc Chem Res 33*, 889-897.

[10] Hess, B. (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations, *J Chem Phys 116*, 209-217.

[11] Grossfield, A., and Zuckerman, D. M. (2009) Quantifying uncertainty and sampling quality in biomolecular simulations, *Annu Rep Comput Chem 5*, 23-48.

[12] Huang, Y. L., and Rizzo, R. C. (2012) A Water-Based Mechanism of Specificity and Resistance for Lapatinib with ErbB Family Kinases, *Biochemistry 51*, 2390-2406.

[13] McGillick, B. E., Balius, T. E., Mukherjee, S., and Rizzo, R. C. (2010) Origins of resistance to the HIVgp41 viral entry inhibitor T20, *Biochemistry 49*, 3575-3592.

[14] Zhou, Y., Elmes, M. W., Sweeney, J. M., Joseph, O. M., Che, J., Hsu, H. C., Li, H., Deutsch, D. G., Ojima, I., Kaczocha, M., and Rizzo, R. C. (2019) Identification of Fatty Acid Binding Protein 5 Inhibitors Through Similarity-Based Screening, *Biochemistry 58*, 4304-4316.