Supporting Information

Molecular Mechanism of HIV-1 Tat Interacting with Human Dopamine Transporter

Yaxia Yuan,[a,b,†] Xiaoqin Huang,[b,†] Narasimha M. Midde,[c] Pamela M. Quizon,[c] Wei-Lun Sun,[c] Jun Zhu,[c] and Chang-Guo Zhan[a,b,*]

[a]Molecular Modeling and Biopharmaceutical Center and [b]Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 789 South Limestone Street, Lexington, KY 40536, and [c]Department of Drug Discovery and Biomedical Sciences, South Carolina College of Pharmacy, University of South Carolina, Columbia, SC 29208

Running title: Molecular Mechanism of hDAT-Tat Interaction

Correspondence:

Chang-Guo Zhan, Ph.D.

Director, *Molecular Modeling and Biopharmaceutical Center*

Professor, Department of Pharmaceutical Sciences

College of Pharmacy

University of Kentucky

789 South Limestone Street

Lexington, KY 40513

TEL: 859-323-3943

E-mail: zhan@uky.edu

* Corresponding author. E-mail: zhan@uky.edu

† These authors contributed equally to the work.

Computational Details – Refinement of the 3,000 candidates through MD simulations and binding energy calculations

As the rigid-body protein-protein docking through ZDOCK 3.0.2 program were not able to account for the conformational flexibility of the ligand and receptor proteins, and for the purpose to select more reasonable hDAT-Tat binding structures from the 3,000 candidates, we performed structural optimization and binding energy calculations by using the Amber 12 software package[1] and the MM-PBSA method.[2] To reduce the possibility of missing any reasonable candidates of the hDAT-Tat binding structure, our computational screening of the 3,000 candidates were performed for four steps as described below.

Step 1. To relax the candidates of the binding complexes from the initial rigid docking, and to perform a rough but quick screening for HIV-1 Tat binding with hDAT, energy minimization was performed with implicit solvent model (OBC2 model,[3] igb=5) implemented in the PMEMD module of the Amber 12 package,[1] with geometrical constraints using a force constant of 30 kcal/mol on the backbone atoms of HIV-1 Tat and all the atoms of hDAT except for the atoms on the side chains of residues within the extracellular region of hDAT. The energy minimization was performed first for 5,000 steps by using the steepest descent method and then for 5,000 steps by using the conjugated gradient method. Next, the geometrical constraints were kept on all atoms of hDAT except for the atoms of residues within the extracellular region of hDAT, and 10,000 steps of energy minimization were performed. After these steps of energy minimization, MM-PBSA module of the Amber 12 was used to calculate the binding energies for all the candidates of the hDAT-Tat binding structure. This led to a selection of top-256 candidates of the hDAT-Tat binding structures according to the ranking of the binding energies, geometrical matching quality, and whether HIV-1 Tat was in unfavorable contacts with any residues inside the vestibule near the extracellular end of hDAT. Among these 256 candidates of the hDAT-Tat binding structures, 139 are associated with the outward-open state of hDAT, 102 associated with the outward-occluded state of hDAT, and 15 associated with the inward-open state of hDAT.

Step 2. Further MD simulations with the implicit solvent model were performed on the top-256 candidates of the hDAT-Tat binding structures to optimize the protein-protein binding structures and then probe stable binding mode of the hDAT-Tat complex. These protein-protein binding systems were heated to 300 K by applying Langevin dynamics for 2 ns (for equilibration). The lengths of covalent bonds involving hydrogen atoms were fixed with the SHAKE algorithm,[4]

enabling the use of a 2-fs time step to numerically integrate the equations of motion. A geometrical constraint with a force constant of 30 kcal/mol was applied on all residues of hDAT except those within the extracellular region of hDAT during the simulation. The second-round of calculations were performed to calculate the binding energies by using the MM-PBSA module of the Amber 12. After this round of MM-PBSA calculations, top-32 candidates of the hDAT-Tat binding structures were selected according to the ranking of the hDAT-Tat binding energies, the visual checking of geometrical matching quality, and whether HIV-1 Tat had unfavorable contacts with residues within the vestibule near the extracellular end of hDAT. As visually checked, top-32 candidates of the hDAT-Tat binding structures are all associated with the outward-open state of hDAT. Thus, all candidates for HIV-1 Tat binding with either the outward-occluded state of hDAT or the inward-open state of hDAT were ignored in the subsequent steps of screening for the most favorable hDAT-Tat binding structure.

Step 3. In order to explore the stability of the top-32 candidates of the hDAT-Tat binding structures under the physiological environment, we inserted each of these 32 candidates into the pre-equilibrated palmitoyloleoyl phosphatidylcholine (POPC) phospholipid bilayers and solvated by solvent water (TIP3P) on each side of the bilayers at physiological pH 7.4. The force-field parameters for atoms of phospholipid molecules were directly adopted from the Amber Lipid11 force field included in Amber 12 software package. For each system containing the hDAT-Tat binding complex + phospholipid molecules + solvent water molecules, a set of 12 Cl^- ions were added to the solvent as counter ions to neutralize the system. The finally formed system for each of top-32 candidates of the hDAT-Tat binding structures reached a typical size of 120 × 117 × 109 Å. The total number of atoms for each system was about 151,124 atoms, including 350 POPC molecules and about 31,422 water molecules. The whole system was subjected to multiple steps of energy minimization by using the Amber 12 program package. The energy minimization was performed first for 30,000 steps using the steepest descent method, and then 30,000 steps using the conjugated gradient method implemented in the Amber 12 program package. In order to prevent possible steric hindrance between the protein and the environment (including the phospholipid molecules, counter ions, and water molecules), a force constant of 30 kcal/mol was first used to constrain all atoms of the two proteins and the phospholipid molecules while the positions of the water molecules and counter ions were energy-minimized. The constraints were then applied only to the atoms of HIV-1 Tat and hDAT, while the remaining atoms were energy

minimized. Then, the constraints were applied only to the backbone atoms of HIV-1 Tat and hDAT, while the atoms on the side chains of the proteins and the atoms of the environment were energy minimized. Finally, the whole system was energy minimized without any constraints, and the simulations reached the convergence criteria of energy gradient of 10^{-4} kcal/mole $Å^{-1}$.

After the energy minimizations, MD simulations were performed for each of the top-32 systems by using the PMEMD module of the Amber 12 program. Each system was gradually heated to 300 K by applying Langevin dynamics to be equilibrated for 1 ns. During the MD simulations, a cutoff of 10.0 Å non-bonded interaction and a cutoff of 2.0 Å non-bonded list updating were used. The motion for the mass center of the system was removed every 1,000 steps. The particle-mesh Ewald (PME) method was used to treat long-range electrostatic interactions. The lengths of covalent bonds involving a hydrogen atom were fixed with the SHAKE algorithm, enabling the use of a 2-fs time step to numerically integrate the equations of motion. For MD simulations on each of the top-32 membrane-containing systems, a constant surface tension (15 dyne/cm) with interfaces in the X-Y plane was applied. The production MD simulations were kept running for 5 ns with a periodic boundary condition in the NTP (constant temperature and pressure) ensemble at T = 300 K. Based on the trajectories of the 5-ns MD simulations for each of the top-32 systems, top-10 candidates were selected according to the ranking of third-round calculations on the binding energies calculated using the same MM-PBSA method as described above. The selection of top-10 candidates of the hDAT-Tat binding structures was also checked for the geometrical matching quality and whether HIV-1 Tat had unfavorable contacts with residues inside the vestibule near the extracellular end of hDAT. Interestingly, the top-10 candidates of the hDAT-Tat binding structures actually reflect the same binding mode.

Stage 4. The selected top-10 candidates for the hDAT-Tat binding structures were subjected to further MD simulations for a longer time, in order to further relax the binding interface between the two proteins and the whole binding structures. The MD simulations were performed in the same way as described in Step 3. The production MD simulations for each of these 10 binding structures were kept running for a total length of 10 ns. Finally, all the MD trajectories were analyzed, and the last snapshots of MD trajectories for all these 10 systems were energy-minimized and the binding energies were evaluated in the similar ways as described in the above steps. Based on the finally calculated binding energies, the simulated hDAT-Tat binding structure with the lowest binding energy was ranked #1. In fact, all of the 10 MD trajectories represent the same

hDAT-Tat binding mode, without significant difference in consideration of the reasonable dynamic fluctuations.

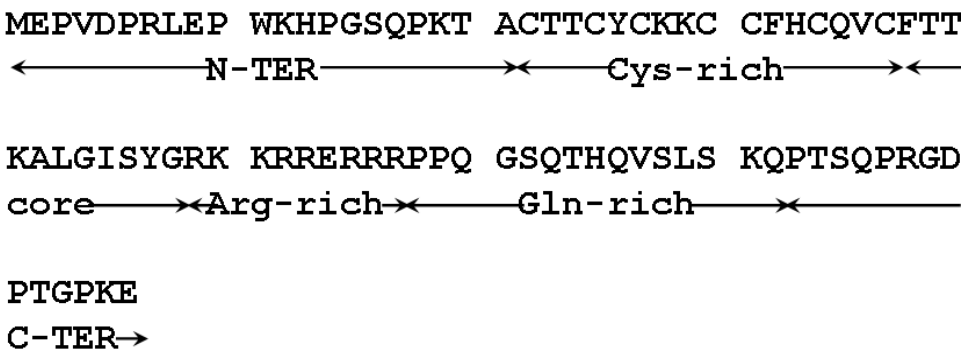

Figure S1. Amino acid sequence for B-type HIV-1 Tat (Tat Bru) protein. Different regions of Tat Bru are labeled under the sequence.

References

1. Case, D. A., Darden, T. A., Cheatham Iii, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M. J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) AMBER 12, *University of California, San Francisco*.
2. Miller III, B. R., McGee Jr, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA. py: An efficient program for end-state free energy calculations, *Journal of Chemical Theory and Computation 8*, 3314-3321.
3. Onufriev, A., Bashford, D., and Case, D. A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model, *Proteins: Structure, Function, and Bioinformatics 55*, 383-394.
4. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics 23*, 327-341.