Impact of pathogenic mutations of the GLUT1 glucose transporter on channel dynamics using ConsDYN enhanced sampling

Supporting Methods

Halima Mouhib[1], Akiko Higuchi[2], Sanne Abeln[3], Kei Yura[4,5], K. Anton Feenstra[3]*

[1)] Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, Marne-la-Vallée, France
[2)] Graduate School of Frontier Sciences, The University of Tokyo, Japan
[3)] Dept. Computer Science, Integrative Bioinformatics, Vrije Universiteit, Amsterdam, The Netherlands
[4)] Graduate School of Humanities and Sciences, Ochanomizu University, Japan,
[5)] School of Advanced Science and Engineering, Waseda University, Japan

*Corresponding author: k.a.feenstra@vu.nl

1 Methods

1.1 Software

The mutate model script in Modeller (Sali and Blundell 1993) was used to produce mutant structures. DSSP (Kabsch and Sander 1983) and JOY (Mizuguchi et al. 1998) were used to calculate the absolute and relative solvent accessibility of residues, respectively. Chimera (Pettersen et al. 2004) was used to visualise the structures.

1.2 MD Simulation System

We used GROMACS 4.0.5 (Hess et al. 2008) for all MD simulations. Atomistic simulations were performed using the CHARMM36 all-atom forcefield (Huang and MacKerell 2013), which we here label 'AT'. The coarse-grained ('CG') simulations were performed using the MARTINI force field with the default time step (Δt=20 fs) (Marrink et al. 2007). Coarse-graining (CG-ing) was performed as previously described for the MARTINI model (Monticelli et al. 2008; Marrink et al. 2004), where the 20 amino acids were mapped into four different bead types with respect to their physicochemical properties. The non-bonded interactions between the CG solvent and solute particles were modelled by truncated and shifted Lennard-Jones pair-potential with a cutoff radius of 1.2 nm (Marrink et al. 2007; Monticelli et al. 2008).
To enable sampling of larger conformational changes, we developed the following protocol. For each of the elastic network restraints generated during set-up of the CG simulations (see next section), we compared them with the equivalent restraints for the other protein in the other state. We then excluded all elastic network restraints where the reference distances differed by more than 1Å between both states, or where the restraint was present in only one. We label this approach 'conserved dynamics' or ConsDYN for short.

1.3 System setup, equilibration and production MD

The wild-type (WT) X-ray structures of the inward-open GLUT1 (Kapoor et al. 2016) and outward-occluded GLUT3 (Deng et al. 2015) structures were obtained from the Protein Data Bank entries 5EQI (resolved at 3.0Å resolution) and 4ZW9 (1.5Å), respectively. Construction of the simulation system, setup of the membrane and solvent, and equilibration were performed using Charmm-gui v1.8 (Jo et al. 2014; Lee et al. 2016), using the membrane builder (Wu et al. 2014), and interface (input generator) for GROMACS (Jo et al. 2008). Coarse-graining was done using the martinizy.py function in the Charm-gui PDB manipulator (Jo et al. 2014). The membrane was built out of POPC lipid molecules. Charges were neutralized using potassium (K^+) or sodium (Na^+) and chlorine (Cl^-) ions, to a salt concentration of 0.1M. Final system composition is summarized in Table S1.
Equilibration procedure consisted of the following steps: first energy minimization on the assembled system; then short simulations position restraints, first 3 cycles of 25 ps at 1 fs timestep, with 10, 5 and 2.5 kCal/mol force constant on the protein backbone, 5.0, 2.5 and 1.0 on the sidechains, 2.5, 2.5 and 1.0 on the lipids, and 10, 0 and 0 on the ions, and then another three cycles of 100 ps at 2 fs timestep with 1.0, 0.5 and 0.1 kCal/mol force constant on the protein backbone, and 0.5, 0.1 and 0.0 on the sidechain and lipids. For the AT equilibration, the first three cycles were performed with a 1 fs timestep, the last three with 2fs; the CG equilibration cycles were performed with 1 fs, 2 fs, 5 fs, 10 fs, 15 fs, and 20 fs, respectively.
Per protein and simulation setup (AT, CG, and CG-CON), the equilibration procedure was performed once, and then five production runs for each simulations were performed starting from the equilibrated conformation with different random starting velocities. For AT, production was 10

ns at 2fs timestep, for CG 1ms at 20 fs time step, and for CG-CON 1 ms at 20 fs time step. Simulations of the mutants were performed using the ConsDYN setup. All analysis were done on the ensemble of the five production runs, unless stated otherwise.

1.4 Overall dynamics using Essential Dynamics analysis

Essential dynamics analysis was performed using GROMACS on the GLUT1 and GLUT3 simulations. To allow this comparison between these two homologous proteins, and allow to focus on overall motions of the channel region, we selected the structurally conserved helical segments, as summarized in Table S2. Then, the covariance and eigenvalue calculation was performed on the ensemble of both wildtype systems, using the AT, CG and ConsDYN simulations.

1.5 TM Helix motions

For each of the ring of six central helices that make up the channel, TM2, TM1, TM5, TM8, TM7, TM11 (and back to TM2), we defined an ‘inside’ and ‘outside’ segment of ten residues (See Figure 1 and Figure S3). Between each of the six pairs of adjacent helices in the ring, we calculated the shortest atomic distances between the ‘inside’ segments, and between the ‘outside’ segments. We did likewise for the three helix pairs ‘across’ the ring: TM1-TM7, TM5-TM11, TM8-TM2. This yields a total of 9 x 2 = 18 distances. For analysis of the simulations, we extracted snapshots at 10 ns intervals, and calculated all 18 distances for each of the snapshots.
We compared distance distributions between mutants and wildtype based on histograms with a bin width of 0.005 nm, by calculating the overlap integral as fraction of the histogram with the smallest integral. This we used to estimate the magnitude of the shift in distances between the two conditions (wildtype and respective mutant). To estimate the direction of the shift, we took the distance between the maxima of both distributions: a positive value meaning the maximum at a larger distance in the mutant, than in the wildtype (i.e., an ‘opening’ shift, see Table S3).
For further representation and comparison of these distances, 2-dimensional distribution plots were made for each pair of ‘in’ and ‘out’ distances, for each of the 9 pairs of helices described above (see Figure S4). Contour plots were made using the KDE function of the stats module in SciPy, with a Gaussian kernel and default parameters. Contouring was done on a logarithmic scale, to enable visualization of broad, low-density outlying areas, as well as details of high-density peaks.

Bibliography

Deng, D., Sun, P., Yan, C., et al. 2015. Molecular basis of ligand recognition and transport by glucose transporters. *Nature* 526(7573), pp. 391–396.
Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E. 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* 4(3), pp. 435–447.
Huang, J. and MacKerell, A.D. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry* 34(25), pp. 2135–2145.
Jo, S., Cheng, X., Islam, S.M., et al. 2014. CHARMM-GUI PDB manipulator for advanced modeling and simulations of proteins containing nonstandard residues. *Advances in protein chemistry and structural biology* 96, pp. 235–265.
Jo, S., Kim, T., Iyer, V.G. and Im, W. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* 29(11), pp. 1859–1865.
Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: pattern recognition of

hydrogen-bonded and geometrical features. *Biopolymers* 22(12), pp. 2577–2637.
Kapoor, K., Finer-Moore, J.S., Pedersen, B.P., et al. 2016. Mechanism of inhibition of human glucose transporter GLUT1 is conserved between cytochalasin B and phenylalanine amides. *Proceedings of the National Academy of Sciences of the United States of America* 113(17), pp. 4711–4716.
Lee, J., Cheng, X., Swails, J.M., et al. 2016. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* 12(1), pp. 405–413.
Marrink, S.J., de Vries, A.H. and Mark, A.E. 2004. Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B* 108(2), pp. 750–760.
Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P. and de Vries, A.H. 2007. The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry. B* 111(27), pp. 7812–7824.
Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. 1998. JOY: protein sequence-structure representation and analysis. *Bioinformatics* 14(7), pp. 617–623.
Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P. and Marrink, S.-J. 2008. The MARTINI coarse-grained force field: Extension to proteins. *Journal of Chemical Theory and Computation* 4(5), pp. 819–834.
Pettersen, E.F., Goddard, T.D., Huang, C.C., et al. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13), pp. 1605–1612.
Sali, A. and Blundell, T.L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* 234(3), pp. 779–815.
Wu, E.L., Cheng, X., Jo, S., et al. 2014. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *Journal of Computational Chemistry* 35(27), pp. 1997–2004.