Supplementary Information

Protein Kinase C mediated Sodium Glucose Transporter 1 activation in precondition induced cardioprotection

Abhinav Kanwal[1,2], Sujatha Kasetti[3], Uday Kumar Putcha[4], Shailendra Asthana[2*], Sanjay K Banerjee[1,2*]

Affiliations

[1]Division of Medicinal Chemistry and Pharmacology, Indian Institute of Chemical Technology, Hyderabad, India.

[2] Drug Discovery Research Center (DDRC), *Translational Health Science and Technology Institute* (*THSTI*), Faridabad, Haryana. India.

[3]Department of Pharmacology, National Institute of Pharmaceutical Education and Research, Hyderabad, India.

[4]Department of Pathology, National Institute of Nutrition, Hyderabad, India.

Running Title: PKC and SGLT1 in cardioprotection.

*To whom the correspondence should be addressed:

Dr. Shailendra Asthana
Drug Discovery Research Center (DDRC)
Translational Health Science and Technology Institute (THSTI),
NCR Biotech Science Cluster, 3rd Milestone
Faridabad – Gurgaon Expressway, Haryana-121001, India
E Mail: sasthana@thsti.res.in

Dr. Sanjay K Banerjee
Drug Discovery Research Center (DDRC)
Translational Health Science and Technology Institute (THSTI),
NCR Biotech Science Cluster, 3rd Milestone
Faridabad – Gurgaon Expressway, Haryana-121001, India
E Mail: skbanerjee@thsti.res.in

Molecular modeling: A complete set of PKC structure which constitute four domain i.e C1, C2 and

catalytic C3 and C4, is not available, therefore a complete homology model of PKC were generated. The sequence P13866 for SGLT1, and P17252, Q05655 for PKC was retrieved from UniProt. In this regard the multiple-sequence alignment with all template sequences were constructed with clustalW[1]. Four PKC structures (separately for different domains PDB-IDs 2I0E (C3+C4), 1A25 (C2), 3PFQ (C1), and 3IW4 (C3+C4)) and two SGLT1 structures (PDB codes 3DH4 and 2XQ2) are used for modelling templates and they are structurally aligned with VMD. The same was carried out for rat SGLT1. All templates were structurally aligned with secondary structure matching. Individual pairwise a sequence-to-structure alignment between the PKC and SGLT1 sub unites and all templates subunits were obtained from the Fugue server[2]. From these data, alignment variants of variable segments were constructed, and homology models were built using modeller[3]. Separately, 300 models were generated for proteins and among them (300 each) the best models quality based on Z-Score, dope score and ramachandran plot were chosen for protein-protein docking. Since PKC has multi-domain structure (C1 to C4), therefore threading approach was also applied by using stand-alone I-tasser program[4]. The I-tasser was used to generate high-quality predictions of 3D structure and biological function of protein molecule from their amino acid sequence without any template structure. The more detail description of the method is available in the references[4,5]

Analysis and validation of homology model: The assessment of reliability of our modeled structure of protein was further carried out using various programs such as WHAT-IF, PROCHECK, QMEAN and ProSA[6-9]. The What if tool helped to determine the Ramachandran z-score value which signifies the overall quality of the modeled structure[8]. Determination of phi and psi torsion angles using Ramachandran plot available through PROCHECK helped to calculate the backbone conformation of the modeled structure[7]. The main–chain parameter file generated through PROCHECK also suggested overall quality of model in terms of overall G-factor and bad contacts per 100 residues[7]. The overall quality of model was also confirmed from the score obtained from QMEAN server[6]. The comparative analysis of modeled proteins was carried out, by superposing both the structures using VMD to obtain root mean square deviation (RMSD)[10]. The RMSD value showed close relation between both the structures.

Molecular dynamics Simulation: AMBER99SB[11], AMBER-modified[12] and TIP3P[13] force fields were used for model proteins of PKC and SGLT1, ions and water, respectively. State-of-the-art all-atom MD simulations were carried out with the NAMD2.8[14] package for the modeled proteins. The solute was placed within a cubic box ensuring a minimum distance of 16 Å between any protein atom and the edge of the box filled with explicit water molecules (TIP3P) and counter-ions. Briefly, geometry

optimizations were carried out with a two-step protocol: (i) up to 10000 cycles (2000 of steepest descent plus 8000 of conjugate gradient) with harmonic restraint (k = 1 kcal mol-1 Å-2) on non-hydrogen atoms of the solute; (ii) up to 10000 conjugate gradient cycles with no restraints. Next, heating up to 310 °K was achieved by linearly increasing the temperature within 100 ps of NVT MD, while imposing restraints of 1 kcal mol-1 Å-2 on non-hydrogen atoms of solute. Restraints were then released for 100 ps and, as a last step preceding the productive dynamics, 1 ns of NPT MD was carried out in order to relax the simulation box. Finally, an MD simulation of 10 ns duration for protein in explicit water solution under the NPT ensemble was performed. Temperature and pressure were regulated at 310 °K and 1.013 bar using a Langevin thermostat (damping constant 5 ps–1)[15] and the Nosé-Hoover-Langevin piston pressure control[16] Electrostatic interactions were evaluated using Soft Particle Mesh Ewald schemes with 1 Å grid spacing and a cut-off of 12 Å, i.e. the same used for Lennard-Jones interactions.

Generation of docking poses: Two different approaches were applied to generate reliable complex of PKC and SGLT1. The pyDock was used for rigid docking, while SwarmDock [17] was used for flexible docking. We scored the docking models generated by the above described methods with our pyDock protocol[18] based on energy terms previously optimized for rigid-body docking. The binding energy is basically composed of accessible surface area-based desolvation, Coulombic electrostatics and van der Waals energy (with a weighting factor of 0.1 to reduce the noise of the scoring function). Electrostatics and van der Waals were limited to ±1.0 and 1.0 kcal/mol for each interatomic energy value, respectively, to avoid excessive penalization from possible clashes in the structures generated by the rigid-body approach. The same protocol was used in the scoring experiment to score all the docking models that were proposed.

Removal of redundant docking poses: After scoring, we eliminated redundant predictions to increase the variability of the predictions and maximize the success chances using a simple clustering algorithm with a distance cutoff of 4.0 Å, as previously described[19].

Minimization of final models: The final best complex docking poses was minimized to improve the quality of the docking model and reduce the number of interatomic clashes. The AMBER10 with AMBER99SB force field was used for minimization protocol consisted of a 500-cycle steepest descent minimization with harmonic restraints applied at a force constant of 25 kcal/ (mol·$Å^2$) to all the backbone atoms to optimize the side chains, followed by another 500-cycle conjugate gradient

minimization without restraints. This minimization step was performed after ranking, solely to remove clashes.

Analysis of structure and energetics: An inventory of structural and energetic features of the complexes was obtained by analysing in terms of hydrogen bonds (HB), and hydrophobic contacts (HpH). The HB's between proteins were counted applying cut-offs of 3.5 Å for the donor-acceptor distance and 150° for the donor-hydrogen-acceptor angle. A HpH was counted when non-polar atoms were separated by a distance of at most 4 Å. π-π interactions were considered to be formed when the short inter-atomic carbon-carbon distance (SICD) was smaller than 4.8 Å.

Supplementary Figures

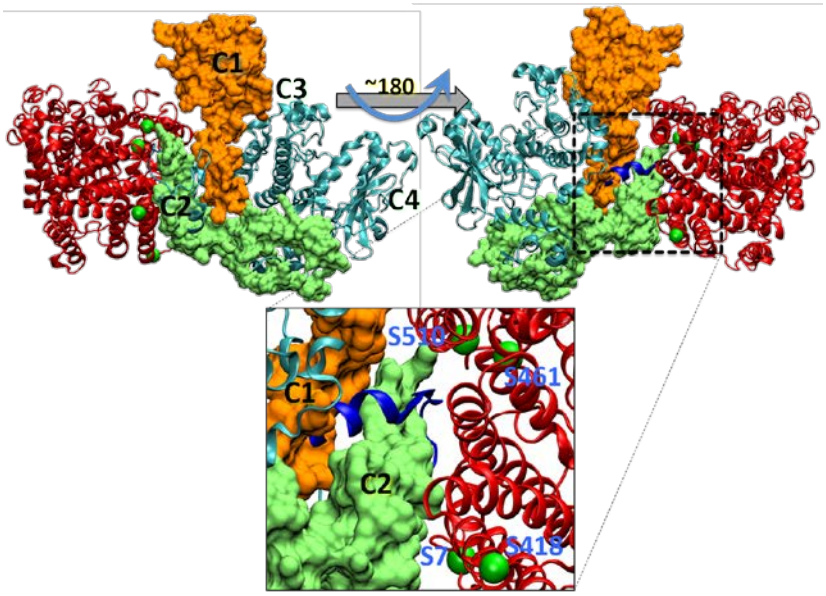

Figure S1: PKC-SGLT1 complex showing the interface site in surface view. The interacting domain of PKC shown in orange (C1) and lime (C2), while SGLT1 has shown in red color. The inset view highlighting the serine residues in green beads, arranged at interface site, and linker in blue.

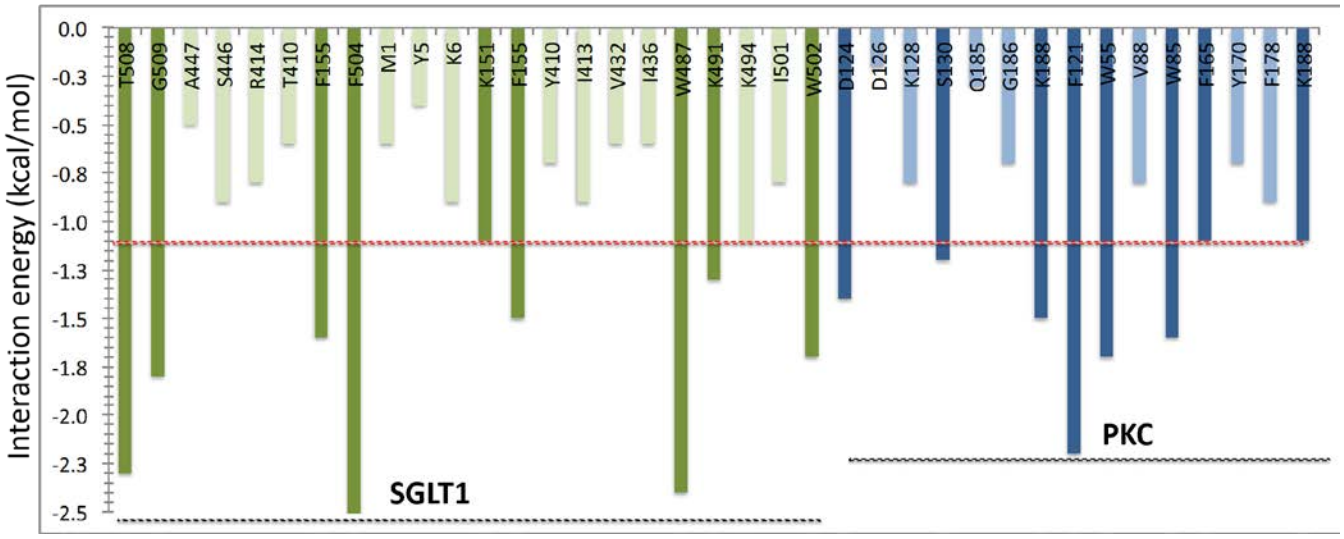

Figure S2: Residue wise interaction map of PKC-SGLT1 complex. The green and blue bars showing the SGLT1 and PKC residues, respectively. The residues contributing less than -1.2 (cutoff value shown by dotted red line) kcal/mol are shown by same transparent color.

References

1. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).

2. Shi, J., Blundell, T. L. & Mizuguchi, K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol***. 310**, 243-257, doi:10.1006/jmbi.2001.4762 (2001).

3. Eswar, N. *et al.* Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics. / editoral board, Andreas D. Baxevanis ... [et al.]* **Chapter 5**, Unit 5.6, doi:10.1002/0471250953.bi0506s15 (2006).

4. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738, doi:10.1038/nprot.2010.5 (2010).

5. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC bioinform.* **9**, 40, doi:10.1186/1471-2105-9-40 (2008).

6. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)* **27**, 343-350, doi:10.1093/bioinformatics/btq662 (2011).

7. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291, doi:doi:10.1107/S0021889892009944 (1993).

8. Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graphics.* **8**, 52-56, 29 (1990).

9. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nuc. Acids Res.* **35**, W407-410, doi:10.1093/nar/gkm290 (2007).

10. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

11. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712-725, doi:10.1002/prot.21123 (2006).

12. Aytenfisu, A. H., Spasic, A., Seetin, M. G., Serafini, J. & Mathews, D. H. Modified Amber Force Field Correctly Models the Conformational Preference for Tandem GA pairs in RNA. *J. Chem. Theory. Comput.* **10**, 1292-1301, doi:10.1021/ct400861g (2014).

13. Mark, P. & Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **105**, 9954-9960, doi:10.1021/jp003020w (2001).

14. Wang, Y., Harrison, C. B., Schulten, K. & McCammon, J. A. Implementation of accelerated molecular dynamics in NAMD. *Comput. Sci. Discov.* **4**, 015002 (2011).

15. Arkhipov, A., Yin, Y. & Schulten, K. Membrane-Bending Mechanism of Amphiphysin N-BAR Domains. *Biophys. J.* **97**, 2727-2735, doi:10.1016/j.bpj.2009.08.051 (2009).

16. Freites, J. A., Tobias, D. J., von Heijne, G. & White, S. H. Interface connections of a transmembrane voltage sensor. *Proceedings of the Proc. Natl. Acad. Sci. U.S.A.* **102**, 15059-15064 (2005).

17. Torchala, M., Moal, I. H., Chaleil, R. A., Fernandez-Recio, J. & Bates, P. A. SwarmDock: a server for flexible protein-protein docking. *Bioinformatics (Oxford, England)* **29**, 807-809, doi:10.1093/bioinformatics/btt038 (2013).

18. Cheng, T. M., Blundell, T. L. & Fernandez-Recio, J. pyDock: electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. *Proteins* **68**, 503-515, doi:10.1002/prot.21419 (2007).

19. Pons, C., Solernou, A., Perez-Cano, L., Grosdidier, S. & Fernandez-Recio, J. Optimization of pyDock for the new CAPRI challenges: Docking of homology-based models, domain-domain assembly and protein-RNA binding. *Proteins* **78**, 3182-3188, doi:10.1002/prot.22773 (2010).