Box S1 | Computational methods used by the GPCR Dock 2008 participants

This supplementary information contains descriptions of the computational methods used by GPCR Dock 2008 participants to predict the structure of the human adenosine A_{2A} receptor bound to ZM241385 from the amino acid sequence of the receptor and a 2D structure of the ligand. Each participant supplied a brief description of the following:

1) Information/data used for modelling
2) Algorithm description for modelling and ligand docking
3) Criteria for prediction analysis and ranking.

The information is provided to help understand how the models were created and ranked for future algorithm development and to improve our understanding of the current state of the art in methods for GPCR structure modelling and docking. Only formatting corrections were made to the content the participants submitted.

Arthur Olson
Department of Molecular Biology
The Scripps Research Institute
olson@scripps.edu

Sequence search

The sequences used for the multiple sequence alignment (human and mouse) were obtained from the UniProtKB/Swiss-Prot database.

Multiple sequence alignment

Human and murine sequences of AA1R, AA2AR, AA2BR, AA3R were aligned together with the sequences extracted from the PDB structures of turkey β1-Adrenergic Receptor (PDB:2VT4) and human β2-Adrenergic Receptor (PDB:2RH1). Sequences were aligned by means of ClustalX with following parameters: PAM250 matrix; gap open and gap elongation penalties of 10 and 0.05.

Homology modeling

The homology modeling procedure was performed with Modeller 9v5 using a multiple sequence alignment/multiple template protocol. The template structures used were the PDB entries 2VT4 and 2RH1. The loop refinement module "loopmodel" with the annealing MD mode as "slow" was used for modeling the EC loop.

Structure refinement and MD simulations

The most promising models were minimized and equilibrated for 5 ns at 300 K, after explicit solvation with ~35k water molecules in a cubic box of ~100 Angstrom side. The simulation was carried out using the GROMACS package.

Models evaluation and binding site analysis

Beside energy score from Modeller, the models were also ranked based on the disposition of key residues identified by mutagenesis studies and the EC loop with respect to the binding site. The principal residues used to drive the model selection studies were V84, E151, E169, F182, H250, N253, F257, I264, and H278. Models showing the best overall residue placement in proximity of the binding site were ranked better. The models were further evaluated by running the AutoLigand code (part of AutoDock Tools) to find optimal docking sites. The protein surface was scanned calculating binding affinity grids with a 1 Angstrom step size, and the grids were analyzed with the contiguous volume fill algorithm of AutoLigand to identify the most favorable binding pockets for ligand binding by energy density. AutoLigand results were then used to both drive and estimate the quality of the eventual docking process.

Docking and ranking

AutoDock was used to dock the fully flexible ligand to the ten best ranked receptor models, kept as rigid targets, to produce the submitted results. The ligand poses were evaluated by energy and clustering, according to the established AutoDock protocol. The agreement with the AutoLigand results and the interactions with the residues identified by mutagenesis studies constituted a further positive score.

Wiktor Jurkowski and Arne Elofsson
Center of Biomembrane Research
Stockholm University
wiktor.jurkowski@sbc.su.se

Information/data used for modeling

Receptor models were built using human β-2-Adrenergic receptor (2RH1) [1] and Squid Rhodopsin (2Z73) [2] structures. According to published data,[3,4] the following residues appear to be important for ligand binding: Phe182, His250, Asn253, Ile274, His278, Ser281, Glu151 and Glu169 and were chosen for filtering in the modeling pipeline.

Algorithm description for modeling/docking

Modeller 9v4[5] was used to build models using multiple alignments created with Blast [6], Pcons [7] and HHpred [8]. Alignment was assumed to be correct when it contained a minimal number of gaps in the TM regions. The final alignment was chosen on the basis of maximal target structure coverage and the presence of conserved TM residues. When required, alignments were manually adjusted to remove gaps from within the TM regions. Preliminary structure model building was followed by TM segment quality assessment which includes MQAP measures and side-chain orientation check. The structure with the best score was then used as the final template from which to build the receptor structure. All extra-cellular loops were rebuilt and refined in Modeller 9v4 using the DOPE force-field.

Conformational search, geometry optimization and partial charges of ligand were calculated with Jaguar [9] QM program (DFT/B3LYP). Ligand docking was done in Autodock4 [10] program. Calculations were repeated for different state of ligand and/or receptor side chains flexibility. To provide maximal conformational freedom of the ligand, calculation grid was not limited to any specific localization covering most of TM part and complete extracellular part of the molecule.

Criteria for prediction analysis and ranking

Receptor models were selected on the basis arrangement and accessibility for the ligand of selected residues in binding pocket and in ECL2 and ECL3 loops. Generated loop conformations were clustered and ranked by potential energy and assumed proper side chain orientation. Winning conformations were additionally rewarded for opening of receptor TM binding site.

First rigid docking was done for pre-calculated geometry of ligand. Extended conformation of ligand represented energy minimum and was found as most probable to fit TM binding pocket of receptor. If large clusters representing low energy conformers were found, they indicated preferred alternative binding sites, which were then additionally scored by presence of preferred contacts. In all of this, binding sites docking was repeated with flexible ligand and additionally allowing partial flexibility of selected side-chains. Final ranking based on consensus scoring consisting of: binding energy, cluster population, occurrence of measured contacts and ligand geometry preference.

References

1. Cherezov, V., *et al.*, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258-65 (2007).
2. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453**, 363-7 (2008).
3. Yuzlenko, O. & Kiec-Kononowicz, K. Molecular modeling of A(1) and A(2A) adenosine receptors: Comparison of rhodopsin- and beta(2)-adrenergic-based homology models through the docking studies, *J. Comp. Chem*. **30,** 14-32 (2008).
4. Kim, J., Wess, J., van Rhee, A. M., Schöneberg, T. & Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor. *J. Biol. Chem.* 270, 13987-97 (1995).
5. Eswar, N., *et al.* Comparative Protein Structure Modeling with MODELLER. Curr. Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 200.
6. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol*. **215**, 403-410 (1990).
7. Wallner, B. & Elofsson, A. Identification of correct regions in protein models using structural, alignment and consensus information. *Protein Sci*. **15,** 900-913 (2006).
8. Söding J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951-960 (2005).
9. Schrödinger, Inc.: Portland, OR, 1991-2000.
10. Morris, G. M. et al. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Comp. Chem*. **19**, 1639-1662 (1998).

Irina Pogozheva and Andrei Lomize
Peptide Synthesis and Molecular Recognition Laboratory
University of Michigan
irinap@umich.edu

Summary:
In order to assess the performance of docking methods in a blind prediction test, it is helpful to have a "null hypothesis model" that is constructed using standard molecular modeling tools but without any docking software and without serious analysis of experimental data published for a given receptor.

Here we tried to obtain such a model during one-day express modeling of the adenosine receptor-antagonist complex. The modeling was done using three experimental GPCR-templates (PDB files: 2U19, 2RH1, 2Z73), with in-house software for homology modeling, and manual docking of a ligand similar to other ligands in structural templates using Quanta visualization and modeling modules. Modeling was followed by a brief energy minimization using CHARMM to remove hindrances.

The comparison of the thus produced model with the experimental structure (PDB 3EML) lets us make a few conclusions:

1. The express-modeling of the membrane part of the receptor was relatively accurate (rmsd 1.28 Å for 218 Cα-atoms in TM α-helices).

2. Prediction of loop structures was poor but may be substantially improved by implementing some distance constraints (as disulfide bonds), if such information is experimentally available.

3. Small-sized unnatural ligands may only partially occupy the large water-filled binding cavity of a receptor in a position that may deviate from that of natural ligands.

Bernard Maigret
Orpailleur team, LORIA (http://www.loria.fr)
Nancy University
bernard.maigret@loria.fr

A2ar Receptor Model Building:

Firstly the available structures for β1- (PDB code 2VT4) and β2- (PDB codes 2RH1, 2R4R, 2R4S, 3D4S) adrenergic receptors were considered and compared in order to select the best possible template (figure $1_{a,b}$). Next the consensus β-adrenergic 3D structure identified this way was used for the homology model building of human adenosine A2ar receptor using for that the sequence alignments depicted on figure 2_a. MODELLER [1] was used at this stage and a preliminary full $3D^0$ A2ar receptor model was achieved (figure 2_b).

Figure 1a: alignment between all the available β1- and β2-adrenergic receptor 3D structures
(in black color the conserved residues between the β1- and the β2-receptors)

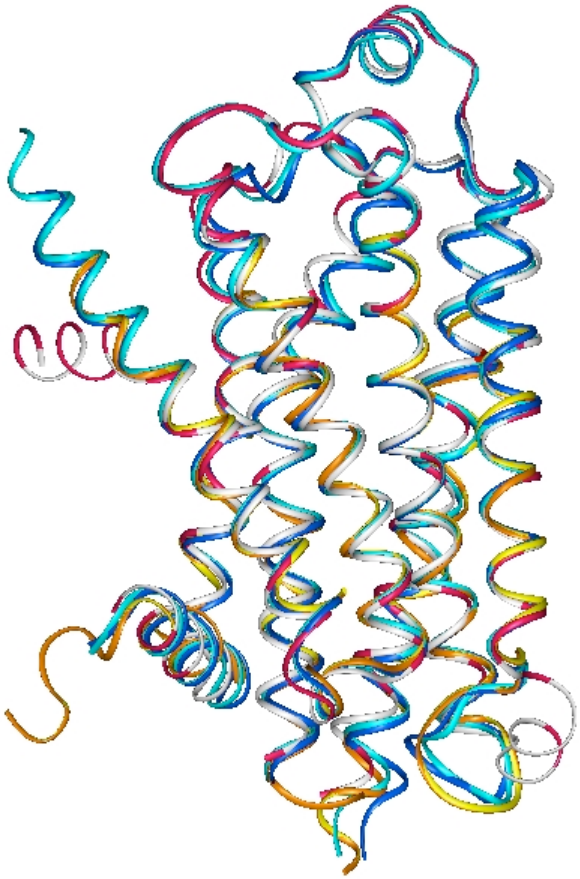


Figure 1b: superposition of the β1- and β2- adrenergic receptor structures

Figure 2a: alignement between the consensus β-adrenergic and A2a receptors
(in black color the conserved residues between the β2 and A2a- receptors)

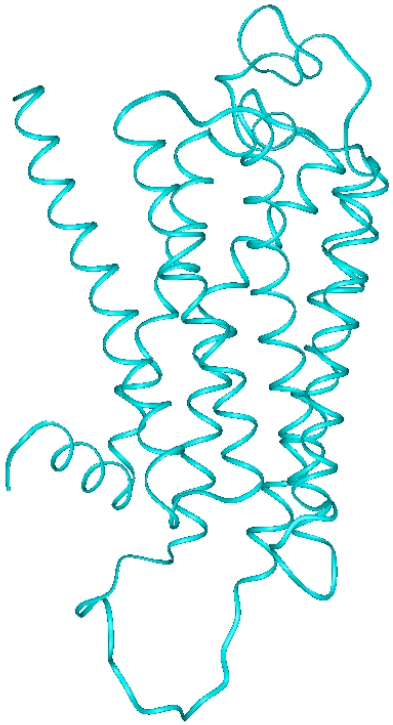


Figure 2b: the A2ar $3D^0$ model (the model was truncated on the C-terminal side after the last helix).

A2ar Receptor Model Refinement:

This preliminary $3D^0$ model was next embedded into a lipid + water surrounding using the VMD [2] plug-ins, giving this way the $3D^1$ model (figure 3). The box size of the system was 80x80x100 $Å^3$ for a total of 53,459 atoms including the necessary counter ions. This $3D^1$ model was then energy refined (64,000 steps of conjugate gradient minimizer), and its stability was checked by Molecular Dynamics (MD) simulations. For that purpose, a 5 ns MD was recorded after a preliminary equilibrium of 500 ps. The last MD snapshoot was energy refined using the same protocol as before, leading to the $3D^2$ model which will be used for the docking calculations. The program NAMD [3] was used for that purpose with the charmm27 force field.

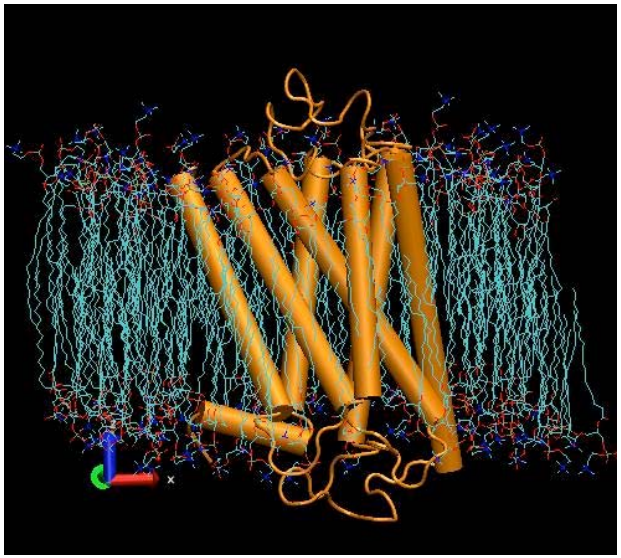


Figure 3: the 3D[1] A2ar model

Docking:

Docking of the ZM241385 ligand into the final 3D[2] model as obtained above was done using the docking GOLD software [4]. For that purpose, the protein A2ar structure alone, as extracted from the 3D[2] model, was used. The 10 best poses obtained this way were next shortly energy refined (protein+ligand only) giving the submitted models.

References:

1. MODELLER: Generation and refinement of homology-based protein structure models, FISER A, SALI A, *Meth. Enzymol.* **374,** 461-491 (2003).

2. http://www.ks.uiuc.edu/Research/vmd/

3. http://www.ks.uiuc.edu/Research/namd/

4. http://www.ccdc.cam.ac.uk/products/life_sciences/gold/

Jeremy Horst[1], Ambrish Roy [2], Brady Bernard[1], Shyamala Iyer[1], Yang Zhang[2], Ram Samudrala[1]
1 Computational Biology Group, University of Washington
2 Department of Molecular Biosciences, Center for Bioinformatics, University of Kansas
jhorst@compbio.washington.edu

Structure and function prediction of the alpha 2A adenosine receptor

Decoy generation:
Initial decoys for the A2A adenosine receptor (A2Ar) were generated by I-TASSER [1-3] which first threads the target sequence through the PDB library and then assembles the continuous fragments into full-length models by Monte Carlo simulations. The only differences from the standard I-TASSER programs was (a) secondary structures for membrane bound region was predicted by membrane protein specific tools of MEMSTAT2 [4] and TMHMM [5] while PSIPRED was used for predicting c3-loop region and cytosolic domain; (b) sequence-based C-alpha contact predictions by the machine-learning algorithm of SVMSEQ [6] was incorporated in the assembly of initial decoys.

Model selection:
The decoys were selected based on an empirical scoring function which is equal to the sum of number of H-bonds, I-TASSER energy for the model, TM-score of the model with the SPICKER [7] cluster centroid, and the average TM-score [8] of the model with the initial threading templates.

Consensus refinement [9]:
Consensus interatomic distances constraints were derived from five selected decoys. Three constraint sets with 12, 16, and 20 Ångstrom cutoffs were compiled with a single interatomic distance constraint for each residue pair, according to and weighted by the corresponding RAPDF score [10]. Each set was used in CYANA restrained torsion angle dynamics [11] simulations to produce many models, scored as above.

Ligand docking and functional analysis:
Binding modes were produced for the inhibitor ZM241385 using AUTODOCK [12], starting with a binding site mapped from the RET ligand present in PDBid 3c9m. The conformations were scored by our radial mean reduced Cambridge Structural Database [13] derived 6 Ångstrom cutoff Bayesian discriminatory function (RMR6) [14], followed by energy minimization with the ligand using GROMACS molecular mechanics software [15] and PRODRG ligand parameters [16]. For two submitted cases, we also docked the ligand by hand, following the functional mutation analysis compiled for A2Ar by Kim, et al [17]. Final selection was also guided by contacts with residues predicted for high functional significance by our meta-functional signature method (MFS)[18]. High MFS scoring residues occur in linear progression through the center of the A2Ar models, directly suggesting a signal transduction mechanism (MFS scores were applied as 0-1 values to the B factor / temperature column of submitted models).

References:
1. Wu, S., Skolnick, J. & Zhang Y. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol*. **5**, 17 (2007).
2. Zhang Y. Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins* **69 Supp 8**, 108-17 (2007).
3. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinform*. **9**, 40 (2008).
4. Jones, D. T., Taylor, W. R. & Thornton, J. M. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**, 3038-49 (1994).
5. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol*. **305**, 567-80 (2001).
6. Wu S & Zhang Y. A comprehensive assessment of sequence-based and template-based methods for protein contact prediction. *Bioinformatics* **24**, 924-931 (2008).
7. Zhang Y & Skolnick J. SPICKER: a clustering approach to identify near-native protein folds. *J Comput Chem*. **25**, 865-871 (2004).
8. Zhang Y & Skolnick J. Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702-710 (2004).
9. Liu, T., Horst, J. A., Samudrala, R. A novel method for predicting and using distance constraints of high accuracy for refining protein structure prediction. *Proteins* (in review).

10. Samudrala, R. & Moult, J. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *J. Mol. Biol.* **275**, 893-914 (1997).
11. Guntert P. Automated NMR structure calculation with CYANA. *Meth. Mol. Biol.* **278**, 353-378 (2004).
12. Morris, G. M. *et al.* Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Comp. Chem.* **19**, 1639-1662 (1998).
13. Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr.* **B58**, 380-388 (2002).
14. Bernard, B. & Samudrala, R. A generalized knowledge-based discriminatory function for biomolecular interactions. *Proteins*. EPub ahead of print 2008 Nov 11.
15. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435-447 (2008).
16. Schuettelkopf, A. W. & van Aalten, D. M. F. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.* **D60**, 1355-1363 (2004).
17. Kim, S. K. *et al.* Modeling the adenosine receptors: comparison of the binding domains of A2A agonists and antagonists. *J. Med. Chem.* **46**, 4847-59 (2003).
18. Wang, K., Horst, J. A., Cheng, G., Nickle, D. C. & Samudrala, R. Protein meta-functional signatures from combining sequence, structure, evolution, and amino acid property information. *PLoS Comput Biol.* **4**:e1000181 (2008).

Osman Ugur Sezerman
Sabanici University
ugur@sabanciuniv.edu

The functionally important motif sequences were found from multiple alignment and Correspondence Analysis.
These positions were then used as a constraint to our sequence alignment algorithm, keeping these positions fixed we obtained the alignment of the query sequence with the pdb sequence and submitted the model based on this alignment. We obtained the top 10 alignments depending on the alignment score and submitted the models accordingly.

Since we are trying to fix the motif positions there are multiple solutions to possible assignments of the key residues between two proteins, we run the alignment of all these and ranked them according to alignment scores.

Gregory V. Nikiforovich[1] and Christina M. Taylor[2]
[1]MolLife Design LLC, 751 Aramis Drive, St. Louis, MO 63141
[2]Department of Biochemistry and Molecular Biophysics, Washington University, St. Louis, 700 S. Euclid Ave, MO 63110
gnikiforovich@gmail.com

Modeling/docking:
Boundaries of TM helices were determined by sequential alignment with rhodopsin, ADRB2 and ADRB1 by Clustal-W.
Energies for individual TM helices of AA2AR were minimized with some restraints.
The structures of TM helices obtained were aligned to the 2RH1 PDB entry (ADRB2) and packed together; also, some additional sampling of TM arrangement was performed by small rotations of helices along the long axes. The two slightly different structures of the TM region were selected for loop restoration. Generally, the procedure of building the TM region was as described in [1].
The extracellular loops were restored completely *de novo* according to [2] in assumption that the only one conservative disulfide bridge, that of C77-C166, exists. The lowest-energy options of the loops (for two options of the TM regions) were selected for docking.
Docking was performed employing Autodock 3.0. The ligand was placed in the middle of the EC loop region, and a grid with 0.375 Å spacing and 126 x 126 x 126 box was placed around the ligand, encompassing nearly the entire EC loop region. With the exception of the number of GA runs being 250, the Autodock parameters used were essentially as described in [3]. The poses obtained from Autodock were clustered by Autodock3.0 at 2.0 Å.
The final models were selected according to lowest energies obtained by Autodock as well as to cover all different structures of AA2AR and/or distinct clusters of ZMA poses obtained by Autodock.

References:
1. Nikiforovich, G. V., Marshall, G. R. & Baranski, T. J. Modeling Molecular Mechanisms of Binding of the Anaphylotoxin C5a to the C5a Receptor, *Biochemistry* **47**, 3117-3130 (2008).
2. Nikiforovich, G. V. & Marshall, G. R. Modeling Flexible Loops in the Dark-Adapted and Activated States of Rhodopsin, a Prototypical G-Protein-Coupled Receptor, *Biophys. J.* **89**, 3780-3789 (2005).
3. Taylor, C.M., Barda, Y., Kisselev, O.G. & Marshall, G. R. Modulating G-Protein Coupled Receptor / G-Protein Signal Transduction by Small Molecules Suggested by Virtual Screening, *J. Med Chem*. **51**, 5297-5303 (2008).

Stefano Costanzi
Laboratory of Biological Modeling
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health.
stefanoc@mail.nih.gov

Information/data used for modeling

Crystal structure of the β_2-adrenergic receptor (β_2-AR) in complex with carazolol (PDB ID: 2rh1).[1,2] Available mutational data for the adenosine A_{2A} receptor[3,4,5] as interpreted in light of my previous modeling studies.[6,7]

Algorithm description for modeling/docking

Construction of the model. The homology modeling of the human A_{2A} receptor (residues 2-305) has been conducted with MOE,[8] using the CHARMM22 force field and the crystal structure of the β_2-AR in complex with carazolol as a template.[1,2] Before the procedure, the fused lysozyme T4 was cut out from the template. Additionally, the template was protonated with the Protonate_3D function of MOE.[9]

An initial sequence alignment was obtained with the Blosum62 substitution matrix, with penalties for gap insertions and extensions of 7 and 1, respectively. The proper alignment of the conserved motifs that characterize each of the transmembrane (TM) helices was checked and, when necessary, manually adjusted. All the gaps in the alignment of the TM helices were eliminated; gaps in the loops were consolidated into a single gap per loop, and positioned where insertions or deletions seemed compatible with the structure of the template. Due to the lack of structural conservation in the second extracellular loop (EL2), only a segment of five residues from this domain, including the conserved Cys that connects the loop with the extracellular end of the 3rd TM (TM3) via a disulfide bridge, were aligned. The rest of EL2 was built independently from the structure of the template. The final alignment used for the construction of the model is provided in **Figure 1**.

Ten models were built and scored on the basis of electrostatic solvation energy (GB/VI).[10] Intermediate and final models were subjected to energy minimizations with the CHARMM22 force field, until reaching a cutoff parameter on the potential energy gradient of 1 kcal/(mol Å) for the former and 0.5 kcal/(mol Å) for the latter.

Optimization of the model. Most of the residues in the second extracellular and third intracellular loop (specifically residues 142-164 and 209-222) were removed from the A_{2A} model, due lack of homology with the template. The portion of EL2 close to the junction with TM5 was retained, as it is likely to be an integral part of the ligand binding pocket. The model was subsequently optimized with the "protein preparation wizard" workflow, as implemented in the Schrödinger package[11] to: 1) add hydrogens; 2) add N-acetyl and N-methyl amide capping groups to the truncated N- and C-termini; 3) optimize the protonation state of His; 4) optimize the orientation of hydroxyl groups, Asn, Gln, and His; 5) perform a cycle of constrained minimizations allowing a maximum root mean square deviation (RMSD) of 0.30 Å from the original structure.

Molecular docking of ZM241385. The molecular docking of ZM241385 at the homology model of the A_{2A} receptor has been conducted with the Schrödinger package.[11] The ligand was sketched and subsequently docked at the receptor by means of the "Induced Fit" docking procedure,[12] which combines automatic docking of flexible ligands to a rigid representation of the receptor (Glide[13]) with reconstructions of the receptor binding site around the docked ligands (Prime[14]).

The docking box was centered on Leu85(3.33, Ballesteros and Weinstein residue indexing) and featured sides of 30 Å. The procedure started with a Glide docking run, conducted in "standard precision" mode, with a vdW scaling factor of 0.5 for non polar atoms. These were defined as the atoms with partial charge lower than 0.25 for the receptor and lower than 0.15 for the ligand. For each obtained docking pose, a Prime refinement was then performed on all the residues located within 5 Å from the ligand. In the final step, the ligand was extracted and redocked with Glide, in "extra precision" mode, at each of the refined binding sites obtained in the previous step, with a vdW scaling factor of 0.8 for the non polar atoms of the ligand only.

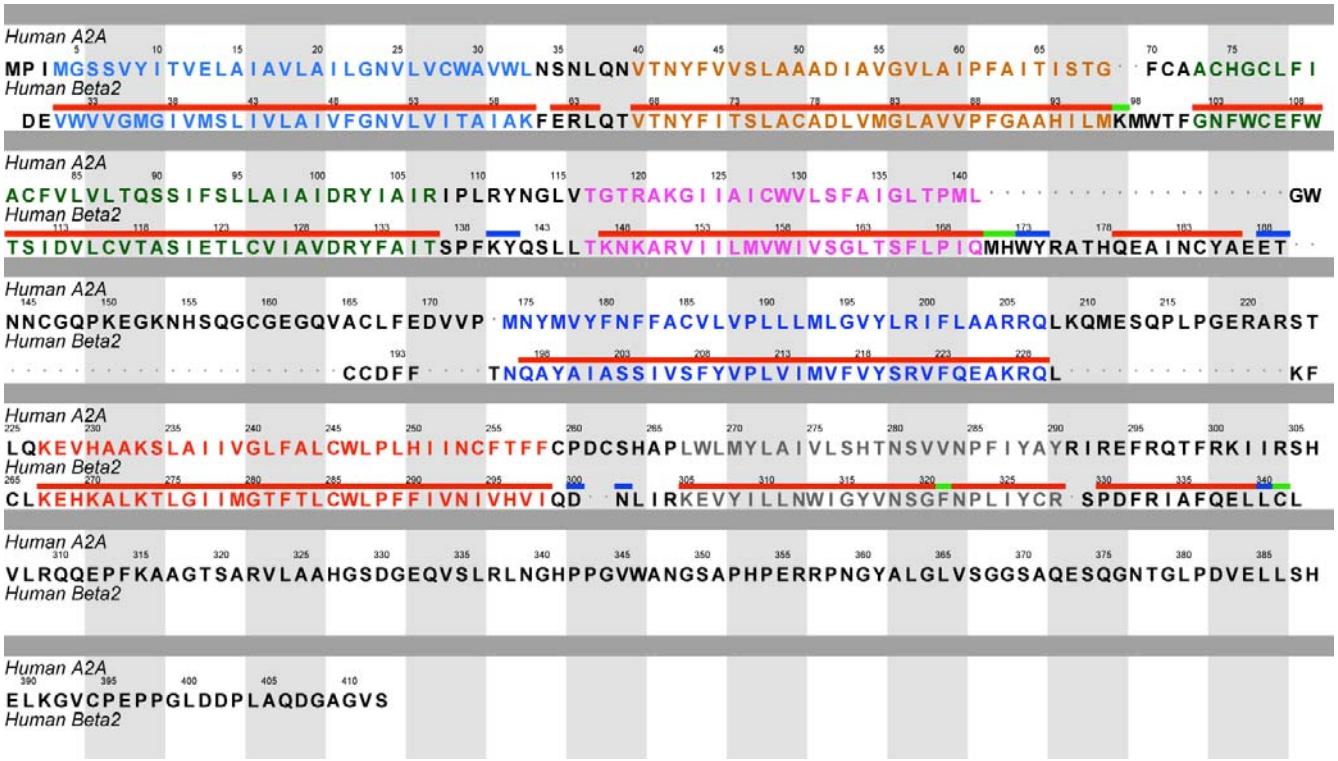
Criteria for prediction analysis and ranking

The poses were selected and ranked on the basis of their docking scores and their agreement with the available mutational data[3,4,5] as interpreted in light of my previous modeling studies.[6,7]

In particular, poses 1 and 2 were selected as representative of the poses with a hydrogen bond between the exocyclic amino group of the ligand and Asn253(6.55), while poses 3 and 4 were selected as representative of the poses with a hydrogen bond between the ligand and His278(7.43). Since mutation of Asn253(6.55) dramatically affects the binding of both agonists and

antagonists, while mutation of His278(7.43) seems to have a less pronounced effect on antagonist binding, the two poses in which ZM interacts with Asn253(6.55) were prioritized. Besides those represented by the four selected models, the other obtained binding modes were not consistent with the mutagenesis data.

Figure 1. Alignment of the whole amino acid sequence of the human adenosine A_{2A} receptor with the portion of the sequence of the β_2-AR solved in the X-ray structure (2rh1). The experimentally determined secondary structure of β_2-AR is represented as a color-coded bar above the sequence (red: helices, blue: 1-4 turns, green: 1-5 turns). The TM helices are color coded (TM1: cyan, TM2: orange, TM3: green, TM4: magenta, TM5: blue, TM6: red, TM7: gray).



References

1. Cherezov, V. *et al*. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258-1265 (2007).
2. Rosembaum, D.M. *et al*. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **318**, 1266-1273 (2007).
3. Kim, J. et al. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor. *J. Biol. Chem*. **270**, 13987-13997 (1995).

4. Jiang, Q. et al. Mutagenesis reveals structure-activity parallels between human A2A adenosine receptors and biogenic amine G protein-coupled receptors. *J. Med. Chem.* **40**, 2588-2595 (1997).
5. Gao, Z.G., Jiang, Q., Jacobson, K. A. & Ijzerman, A. P. Site-directed mutagenesis studies of human A(2A) adenosine receptors: involvement of glu(13) and his(278) in ligand binding and sodium modulation. *Biochem Pharmacol* **60**, 661-668 (2000).
6. Costanzi, S., Lambertucci, C., Vittori, S., Volpini, R. & Cristalli G. 2- and 8-alkynyladenosines: conformational studies and docking to human adenosine A3 receptor can explain their different biological behavior. *J. Mol. Graphics Modelling*, 21, 253-262 (2003).
7. Costanzi, S. et al., *Frontiers in Drug Design and Discovery*, Bentham, 63-79 (2007).
8. MOE, [2007.0902]; Chemical Computing Group Inc., 2007.
9. Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins*, doi: 10.1002/prot.22234 (2008).
10. Labute, P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J. Comput. Chem.* **29**, 1693-1698 (2008).
11. Maestro, [8.5]; Schrödinger, LLC, 2008.
12. Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **49**, 534 -553 (2006).
13. Glide, [5.0]; Schrödinger, LLC, 2008.
14. Prime, [5.0]; Schrödinger, LLC, 2008.

Blind Docking method: Analysis of Protein accessible surface and refinement of binding model by molecular dynamics.

Y. Vorobjev[2] , N. Bakulina[2],V. Solovyev [1,2]

[1]Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX,UK; [2]Softberry Inc., 116 Radio Circle, Suite 400 Mount Kisco, NY 10549, USA

victor@cs.rhul.ac.uk

Initial step in 3D modeling is the selection of a template structure for a query sequence by ***Ffold*** program with the subsequent structure of protein model building by the ***Getatoms*** program performing the template structure and query sequence alignment. Models built by ***Getatoms*** program are further refined by ***MDynSB*** program realizing optimization of a protein structure via molecular dynamics A hierarchical blind docking method consists of two steps. The first step is an analysis of protein accessible surface to identify pockets and clefts. The surface analysis is performed for the given rigid protein structure. The probe's sphere contacts are clustered and the centers of clusters represent the coordinates of binding sites. The number of contacts in cluster represents the contact score of the binding site candidate. The binding site candidates are ranked by the contact score value. The number of binding site candidates for protein of medium size 200-300 residues to be evaluated on the second stage of method does not exceed 10 sites. The second stage is a global optimization of the given ligand position and conformation for each binding site by all atom molecular dynamic simulated annealing. The global optimization is done for 144 starting orientation of ligand placed at the position of binding site candidate. The temperature of simulated annealing is combined with softening of the repulsive Van der Waals atom-atom interactions in all atom force field. Protein binding site can be rigid or flexible at the second stage. A scoring of the binding modes at the second stage is done by the minimum of total potential energy of interaction between ligand and protein. The ligand model with the lowest potential energy of proten-ligand interactions is selected as the best binding mode of the blind docking method.

Homology Modeling of 3-Dimensional GPCR Structure and Ligand Docking

Kazuhiko Kanou[1], Daisuke Takaya[1,2], Genki Terashi[1],
Mayuko Takeda-Shitaka[1,2] and Hideaki Umeyama[1,2]
1) School of Pharmacy, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641 Japan
2) RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
kanouk@pharm.kitasato-u.ac.jp

Information/data used for modeling;

The X-ray structure of the turkey β_1 adrenergic receptor with bound cyanopindolol (PDB code: 2VT4) was used as a template structure for homology modeling. The sequence alignment between AA2AR and the template was generated by our in-house programs FAMSD[1,2] and SKE-CHIMERA[3], and the sequence identity was 32.6%.

Algorithm description for modeling/docking;

First, algorithm description for modeling is described. Using FAMS Ligand&Complex[4], we constructed a 3-dimensional (3D) structure of AA2AR with two ligand molecules, which are cyanopindolol and (2S)-1-(9H-Carbazol-4-yloxy)-3-(isopropylamino)-propan-2-ol obtained from 2VT4 and 2RH1, respectively. The coordinates of the latter ligand were obtained by superimposing 2RH1 onto 2VT4 based upon a structural fitting between two 3D structures.

Secondly, algorithm description for docking is described. We generated an initial coordinates of the target ligand so that the coordinates calculated based on the simulation equation would interact with some significant amino acid residues, ASN 181, ASN 253 and TYR 271. TYR 271 seemed to form a hydrogen-bond from the location of in the binding site, referring 2VT4 and 2RH1. ASN 181 and ASN 253 seemed to pinch a functional group based on the positions of two amino acid residues. Indeed, the results searching for the PDB database including ligands showed that two asparagine residues can pinch various adenine-like rings. Next, we executed the docking calculations including the ligand-optimization process between the GPCR target protein and the ligand having a first arbitrary conformation using our original potential function. This potential function includes the potential energy of hydrogen bond, hydrophobic interaction, stacking, collision, internal energy and the protein-ligand interaction information such as atom-atom distances which seem to be functionally important.

Criteria for prediction analysis and ranking.

We performed clustering for docked ligand conformations since many simulation calculations of such a docking process were carried out. As the results, the biggest cluster was selected, and the top ten conformations to propose for the organizer of the contest were determined from the score values.

References:

1. CASP8 abstracts: http://www.predictioncenter.org/casp8/doc/CASP8_book.pdf
2. Ogata, K. & Umeyama, H. *J. Mol. Graph. Model.* **18**, 258-272, 305-306 (2000).
3. Takeda-Shitaka, M., Terashi, G., Takaya, D., Kanou, K., Iwadate, M. & Umeyama, H. Protein structure prediction in CASP6 using CHIMERA and FAMS. *Proteins* **61 Suppl 7**, 122-7 (2005).
4. Takeda-Shitaka, M., Terashi, G., Chiba, C., Takaya, D. & Umeyama, H. FAMS complex: a fully automated homology modeling system for protein complex structures. *Med Chem.* **2,** 191-201 (2006).

William A. Goddard III, Youyong Li, Soo-Kyung Kim, Bartosz Trzaskowski, Ravinder Abrol, Adam Griffith
Materials and Process Simulation Center (MC 139-74)
California Institute of Technology, Pasadena, CA 91125
wag@wag.caltech.edu

Information/data used for modeling:
An ensemble of 10 structures for human A_{2A} adenosine receptor (hA2AR) with ZM241385 docked, was generated using four different structures for the hA2AR protein. One was obtained based on a fully first-principles based method (GEnSeMBLE [1]) starting from the amino acid sequence. This is similar in concept to our earlier MembStruk method used for predictions of GPCR-ligand complexes {15-20}, but provides far more accurate predictions. Two were obtained by applying our BiHelix sampling method [1] to a homology model based on the human β2 adrenergic receptor (PDBid: 2rh1 [2]). One was obtained by an identical approach based on the turkey β1 adrenergic receptor (PDBid: 2vt4 [3]).

Algorithm description for modeling Protein Structures:

Gensemble/β2. The TM domain was obtained using a fully first-principles based method GEnSeMBLE [1], which uses TM predictions based on our PredicTM method [5] and based on the multiple sequence alignment of sequences related to hA2AR, helix optimization using molecular dynamics (OptHelix [6]), initial placement of optimized helices in a bundle based on β2 adrenergic template, followed by a full BiHelix sampling procedure that sampled rotations of all helices (~35 million based on a 30° grid) to identify the bundle with the lowest energy.

β2HomOp1 and β2HomOp2. The TM domain was based on Homology to the TM of human β2 AR structure optimized using our BiHelix sampling procedure [1] on TMs 5 and 6 (using a 10° grid in the ±90° range from starting positions) that optimizes the amino acid sidechains using our SCREAM method [4] for each rotational combination. Two low energy structures were identified [TM5-TM6 rotation angles of 10-0 (labeled β2HomOpt1 in Table 1 below) and 0-0 (β2HomOpt2)]

β1HomOpt: The TM domain was based on Homology to the TM of turkey β1 AR receptor structure optimized as for the β2 case [TM5-TM6 rotation angles of 0-0 (β1HomOpt)]

Loops. A common set of loops were built using our loop builder [7] based on continuous configuration Monte Carlo (CCB) sampling [21] (not on homology).

Algorithm description for Docking ZM241385 to predicted Protein Structures:
For the protein structures based on β2 template, we used the HierDock method [8], which involves a hierarchical strategy for examining ligand binding conformations and calculating their binding energies. The whole protein was partitioned into different regions and scanned to find the putative binding regions. DOCK [9] was used to generate 1000 conformations in these putative regions and top 1% (10) were selected for further optimization involving binding site minimization and complex minimization. For protein structures based on β1 template, Glide [10] was used followed by our GenMSCDock [11] optimization. The Glide SP docking was used on a partially alanized protein (I, L, V, F replaced by A) for selecting top 100 poses based on GlideScore. This was followed by SCREAM [4] to dealanize the 100 complexes for optimum side-chain conformations for each ligand pose, neutralization of charged residues, minimization of binding site, minimization of the whole complex, and final scoring.

Criteria for prediction analysis and ranking:
The 4 protein structures used for docking, were each the most stable by total energy. The final ligand poses in the predicted complexes were selected based on vertical binding energies or binding cavity energies. The same methods were also applied to predict the structures for the A_1, A_{2B}, and A_3 Adenosine receptor subtypes and to dock ZM241385 in order to explain selectivity [12] across Adenosine receptor subtypes. The following table includes the criteria for ranking of the 10 submitted structures. Also included is the 11th ranked structure in our analysis that was not selected because it had a very weak interaction with N6.55 (Asn253) implicated to be important in binding assays [13].

Our Analysis:
We analyzed our predicted structures against the crystal structure of hA2AR bound to ZM241385 [14] using both ligand RMSD and the energetic contributions of hA2AR residues in the binding site. Table 2 presents the residues (in order of decreasing contribution to binding) that contribute more than 0.5 kcal/mol to binding, for each of our submitted structures (along with Rank 11 that was excluded from the submission due to weak interaction with Asn253). In each column, the residues up to first red line contribute more than 3 kcal/mol, and those up to second

red line contribute more than 1 kcal/mol to binding (except Rank 4 structure which has no residues that contribute above 3 kcal/mol). The table also shows the number of contacts from the XTAL that are found in predicted complexes. Residues shaded blue are hydrophilic and those shaded orange are hydrophobic. Analysis of the data in this table (especially structures with Ranks 4, 5, 8, and 11) shows that RMSD alone does not capture the accuracy of the predictions very well. Structures with Ranks 4 and 8 have the best RMSD of the submitted structures, but share least number of binding site residues with the XTAL. A measure that also takes into account the number and binding energy of protein residues interacting with the ligand is a better measure of accuracy and more useful for drug design purposes.

TABLE 1: Summary of final structures

Final Rank	Protein Method	Docking Method	Angle (TM5, 6)	Ranking Criteria and Comments	RMSD (Å)
1	_2HomOpt 1	DOCK+HierDock	10,0	Best pose with alternate β2: good H -bonds, explains subtype selectivity	10.34
2	_2HomOpt 2	DOCK+HierDock	0,0	Best pose with β2 homology: good H -bonds, explains subtype selectivity	8.86
3	_1HomOpt	Glide/GenMSCDock	0,0	Best pose with β1 homology: good H -bonds, explains subtype selectivity except A3	8.53
4	_2HomOpt 2	DOCK+HierDock	0,0	Similar to Rank1 but rotated 180 °: good H -bonds	5.88
5	_1HomOpt	Glide/GenMSCDock	0,0	2nd Best pose with β1 homology: good H -bonds, explains subtype selectivity except A3	8.73
6	Gensemble/ β2	DOCK+HierDock	0,0	Pose with best gensemble structure using β2 template: good H -bonds	8.24
7	_1 HomOpt	Glide+GenMSCDock	0,0	Similar to Rank3 but rotated 180 °: good H -bonds	8.90
8	_2 HomOpt 2	DOCK+HierDock	0,0	Most vertical pose (perpendular to membrane) with β2 homology: H -bond with TM6 N253	5.58
9	_1 HomOpt	Glide/GenMSCDock	0,0	Similar to Rank5 but rotated about horizontal axis: loses H -bond with TM6 N253	8.58
10	_2 HomOpt 2	DOCK+HierDock	0,0	Similar as Rank1, but l igand's phenolic group is between TM5 and TM6 losing H -bond to N181	10.61
11	_1 HomOpt	Glide/GenMSCDock	0,0	Similar to Rank 3 but rotated 180 °: Rejected because of very weak interaction at N6.55	4.72

ZM:A2AR binding site residues in the order of energetic importance												
XTAL		Rank of Predicted Complexes										
Res	ResID	1	2	3	4*	5	6	7	8	9	10	11
ASN	253	253	253	253	63	168	253	168	169	168	253	250
GLU	169	169	169	168	169	181	257	181	270	182	169	168
PHE	168	85	85	145	85	182	182	85	168	181	274	85
LEU	249	182	274	84	182	85	62	274	182	85	167	167
MET	177	181	182	249	253	177	167	140	174	250	84	278
ILE	274	274	167	178	166	249	269	84	257	274	85	274
LEU	267	84	84	140	84	140	179	253	172	140	251	271
TYR	271	167	250	177	140	274	168	177	171	178	247	249
LEU	85	250	181	181	181	88	178	88	181	84	254	270
HIS	264	140	168	85	278	178	258	246	274	177	187	88
LEU	167	185	140	271	270	84	169	178	253	246	278	84
TRP	246	80	257	274		278	250	270	173	88	80	182
ASN	181	278	178	63		246	186	278	250	249		246
VAL	84	81	80	66		250	183	277		81		181
				278		145	272	250		170		185
				170		170	84	81		278		253
				250		169	273	273		145		267
						253		169		277		
								170				
RMSD	0.0	10.3	8.9	8.5	5.9	8.7	8.2	8.9	5.6	8.6	10.6	4.7
# Contacts	14	7	8	9	5	10	5	9	5	8	6	11

TABLE 2: Binding site analysis

5) References:

1. R. Abrol, J. Bray, A.R. Griffith, W.A.Goddard III; The GEnSeMBLE Method for First Principles Predictions of 3-dimentional structures of GPCRs" To be submitted (2009).
2. Cherezov, V. et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258-1265 (2007).
3. Warne, A. et al. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature 454*, 486 (2008).
4. Kam, V. W. T. & Goddard, W. A. *J. Chem. Theory Comput.*, In press (2008).
5. Abrol, R., Griffith, A. R. & Goddard, W. A. The PredicTM Method for Predicting the trans membrane domains of membrane bound proteins. To be submitted (2009).
6 Griffith, A.R., Abrol, R., Trzaskowski, B. & Goddard, W. A. The OptHelix method of predicting the structures of trans membrane helices", To be submitted (2009).
7. Li, Y. & Goddard, W. A. The CCB Continuous Sampling Monte Carlo Predictions of loops for transmembrane proteins, To be submitted (2009).
8. Trabanino R.J. et al. First principles predictions of the structure and function of g-protein-coupled receptors: validation for bovine rhodopsin. *Biophys. J.* **86**, 1904-1921 (2004).
9. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. & Ferrin, T.E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**, 269-288, (1982); Shoichet, B.K., Bodian, D.L. & Kuntz, I.D., *J. Comp. Chem.* **13**, 380-397, (1992).
10. Friesner, R. A., et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

11. A.R. Griffith, R. Abrol, J. Wendel, and W.A. Goddard, "GenMSCDock", To be submitted (2009).
12. Ongini, E., Dionisotti, S., Gessi, S., Irenius, E. & Fredholm, B.B. Comparison of CGS 15943, ZM 241385 and SCH 58261 as antagonists at human adenosine receptors. *Naunyn-Schmiedeberg's Arch Pharmacol*. **359**, 7-10 (1999).
13. Kim, J., Wess, J., van Rhee, A. M., Schoneberg, T., & Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor. *J Biol Chem* **270**, 13987 (1995).
14. Jaakola, V. P. et al. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science*. **322**, 1211-7 (2008).
15. Li, Y. Y., Zhu, F. Q., Vaidehi, N. & Goddard, W. A. Prediction of the 3D structure and dynamics of human DP G-protein coupled receptor bound to an agonist and an antagonist. *J. Am. Chem. Soc*. **129,** 10720-10731 (2007).
16. Heo, J. Y. et al. Prediction of the 3D structure of FMRF-amide neuropeptides bound to the mouse MrgC11 GPCR and experimental validation. *ChemBioChem* **8,** 1527-1539 (2007).
17. Peng, J. Y. C., Vaidehi, N., Hall, S.E. & Goddard, W.A. The Predicted 3D Structures of the Human M1 Muscarinic Acetylcholine Receptor with Agonist or Antagonist Bound. *ChemMedChem*, **1**, 878-890 (2006).
18. Vaidehi, N. et al. Predictions of CCR1 chemokine receptor structure and BX 471 antagonist binding followed by experimental validation; *J Biol Chem* **281**, 27613-27620 (2006).
19. Kalani, M.Y. et al. Predicted 3D Structure Of The Human D2 Dopamine Receptor And The Binding Site And Binding Affinities For Agonists And Antagonists. *Proc. Natl. Acad. Sci. USA* **101**, 3815 (2004).
20. Freddolino, P. et al. Predicted 3D structure for the human □2 adrenergic receptor and its binding site for agonists and antagonists. *Proc. Natl. Acad. Sci. USA* **101**, 2736-2741 (2004).
21. Sadanobu, J. & Goddard, W. A. The Continuous Configurational Boltzmann Biased Direct Monte Carlo Method for Free Energy Properties of Polymer Chains. *J. Chem. Phys*. **106**, 6722 (1997).

Vsevolod Katritch, Manuel Rueda and Ruben Abagyan
Molsoft L.L.C.
seva@molsoft.com

Our approach to modeling of the A2a Adenosine receptor (AA2AR) complex with antagonist ZM241385 was based on the high resolution crystal structure of the $\beta_2$adrenergi receptor β_2AR (PDB code 2rh1)[1, 2], as well as a published ligand affinity dataset from a recent study [3]. The combined homology modeling and ligand guided backbone ensemble receptor optimization (LiBERO) algorithm included the following steps, performed in the framework of ICM molecular modeling package (ICM version 3.6.-1b, Molsoft LLC):

1. Initial 3D models of the AA2AR were obtained with a standard homology modeling function BuildModel [4] using ZEGA alignment algorithm [5] and the 2rh1 structural template [1, 2]. The T4L fusion domain of the template was removed and the alignment was manually adjusted to eliminate minor gaps in TM1, TM5 and TM7 domains. The unaligned portion of the extracellular loop 2 (EL2), residues G142-A165, was not included into the initial AA2AR model.
2. After limited side chain optimization in ICM torsion coordinates, multiple conformations of the model were generated using Elastic Network Normal Modes analysis, as described in [6]. Overall, over 200 conformations were generated with average protein backbone deviations about 2Å from the initial model.
3. For each of these conformations, ZM241385 compound and its three high affinity analogues (compounds #2, #3, #7 from ref[3]) were docked into the all atom representation of the AA2AR model with flexible side chains. The analysis of the ligand affinity data in ref.[3] points to the importance of the H-bond donor in position N5 for the high affinity binding of pyrazolo-triazolo-pyrimidine (PTP) analogues. The potential contacts of the N5 donor with one of the five possible acceptors in the side chains (His250, Asn253, His278, Tyr271, Thr88) in the pocket found by ICM PocketFinder algorithm [7] , were therefore used as alternative harmonic distance restraints to guide initial placement of the ligand. Energy optimization was performed for the all-atom models with flexible ligand and flexible receptor side chains in the 8 Å proximity of the binding pocket, using ICM biased-probability Monte Carlo algorithm [8]. The resulting ligand-receptor models with acceptable conformational energy were clustered according to the conformations of the binding pocket residues, yielding ~150 non-redundant models.
4. These models of the AA2AR binding pocket were assessed in virtual ligand screening (VLS) benchmark [9] for their ability to discriminate AA2AR specific ligands from 500 decoys randomly selected from ChemDiv drug like compound database. In addition, the decoy set included AA3AR specific PTP analogues from ref.[3] dataset. No distance restraints were used in the ligand docking procedure at the model assessment stage. Three models with the best VLS performance were chosen (AUROC values 91., 89. and 86.), and the top scoring binding poses of the compound ZM241385 in each of these models were used in the final model submission.
5. Modeling of the non-conserved part of the extracellular loop EL2 (residues G142 to A165) was performed with ICM loop modeling algorithm based on global optimization of conformational energy with disulfide bonding restraints. Five possible disulfide bonding configurations were tested for the 6 cysteins in the loop and 7TM domain, only one of them (C71-C159, C74-146, C77-C166) resulting in an energetically feasible loop conformations. The two loop conformations with the best predicted conformational energy were further refined in the context of the three ligand binding models generated in the previous step.
6. The optimized binding pocket - EL2 loop combinations comprised the first six submitted models of AA2AR- ZM241385 complex, and were ranked according to their overall conformational energy. The remaining four models (#7 to #10) included possible alternative conformations of the ligand with the hydrogen bond donor N5 in contact with His250 side chain or A81 main chain acceptors, as well as an alternative (β_2AR-like) conformation of the EL2 loop.

References:

1. Cherezov, V., et al., High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258-65 (2007).
2. Rosenbaum, D.M., et al., GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science*, **318,** 1266-73 (2007).
3. Cacciari, B., et al., Pyrazolo-triazolo-pyrimidines as adenosine receptor antagonists: A complete structure-activity profile. *Purinergic Signal* **3,** 183-93 (2007).
4. Cardozo, T., Totrov, M. & Abagyan, R. Homology modeling by the ICM method. *Proteins* **23**, 403-14 (1995).
5. Abagyan, R.A. & Batalov, S. Do aligned sequences share the same fold? *J Mol Biol* **273**, 355-68 (1997).

6. Rueda, M., Bottegoni, G. & Abagyan, R. Consistent improvement of cross-docking results using binding site ensembles generated with Elastic Network Normal Modes. *J. Chem. Information and Modeling*, in press, 2008.
7. An, J., Totrov, M. & Abagyan, R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol Cell Proteomics* **4,** 752-61 (2005).
8. Totrov, M. & Abagyan, R. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins* **Suppl 1**, 215-20 (1997).
9. Reynolds, K.A., Katritch, V. & Abagyan, R. Identifying conformational changes of the beta(2) adrenoceptor that enable accurate prediction of ligand/receptor interactions and screening for GPCR modulators. *J Comput Aided Mol Des*, (2009).

Ian Davis, Patrick Barth, David Baker
University of Washington
iwd@u.washington.edu
barthp@u.washington.edu

Summary of submission

Models 3600_1 to 3600_5 are models with the antagonist docked on the receptor and should be primarily used to assess our ability to predict receptor/ligand interactions.
Models 3600_6 to 3600_10 are models of the receptor without docked antagonist and should be used to assess our ability to predict the overall receptor structure.

Modeling of the receptor

The adenosine receptor structure was modeled by a comparative modeling technique developed for membrane protein structures (Barth & Baker, unpublished results). The main information used for the modeling was the high-resolution X-ray structures of the beta1 and beta2 adrenergic receptors. These receptors were identified as best initial templates for the comparative modeling based on the quality of the sequence / structure alignments performed by the HHpred server. The regions of the adenosine receptor sequence that exhibited poor alignment with the template sequences were selected for complete *de novo* remodeling by Monte-Carlo-based peptide fragment insertion.[1] The well-aligned regions of the query sequence were threaded onto the structural templates. When the poorly aligned regions were refolded, the entire structure was subjected to all-atom refinement using a high-resolution potential developed for alpha-helical membrane proteins.[2] Around 10000 independent simulations were performed for each starting template. The resulting structures were clustered either based on the putative ligand-binding regions for docking the antagonist or based on the entire structure for selection of the receptor conformation without docked ligand. For the submission of the receptor alone models, the clusters were ranked based on their lowest energy models. Five among the top-ranked clustered models were selected from the modeling performed with either the beta1 or the beta2 adrenergic receptor as starting template. These models were submitted as blind predictions with the rank following their relative energies (3600_6 and 3600_10 had the lowest and the highest all-atom energy among the five submitted models, respectively).

Modeling of the antagonist

Receptor models were clustered based on the putative ligand-binding regions (residues 54-90, 133-185, and 242-280) and the centers of the 50 largest clusters for each template were used for docking. Ligand docking used a combination of Monte Carlo moves, sidechain repacking, and gradient-based minimization with the Rosetta all-atom potential.[3] Receptor backbone and sidechain torsions along with ligand torsions, position, and orientation were optimized simultaneously. The ligand conformer library was generated with Omega (OpenEye Inc.). 5000 docking trajectories were run for each receptor model. Results were pooled by template and the 5% lowest-energy models were clustered by ligand position. The 20 largest clusters and the 20 with lowest binding energy were inspected visually to select the submitted models, based on chemical intuition and fit to published mutational data.[4]

1. Qian, B., et al. High-resolution structure prediction and the crystallographic phase problem. *Nature* **450**, 259-64 (2007).
2. Barth, P., Schonbrun, J. & Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci U S A*. **104**, 15682-7 (2007).
3. Davis, I. W. & Baker, D. RosettaLigand docking with full ligand and receptor flexibility. *J Mol Biol*. **385**, 381-92 (2009).
4. Kim, S. K., et al. Modeling the adenosine receptors: comparison of the binding domains of A2A agonists and antagonists. *J Med Chem*. **46,** 4847-59 (2003).

Michael Feig
Biochemistry & Molecular Biology
Michigan State University
feig@msu.edu

Models of the human A2A adenosine receptor were built as follows: Initial templates were constructed from known structures of human β2 adrenergic receptor (PDB ID: 2RH1), turkey β1 adrenergic receptor (PDB ID: 2VT4), squid rhodopsin (PDB ID: 2Z73) and bovine rhodopsin (PDB ID: 1U19) using the residue mutation function of the MMTSB Tool Set.[1] These initial, incomplete models, were then submitted to as input to TASSER[2] to generate complete all-atom models. Structures closest to the resulting ten most populated clusters were then rescored using an MMGB/SA protocol with the HDGB implicit membrane model[3] after orientation along the membrane z-axis and further minimization of rigid body translation/rotation in the implicit membrane model. The ZM241385 inhibitor was then docked to the best receptor structure using the CHARMM CDOCKER module through DiscoveryStudio from Accelrys.

1. Feig, M., Karanicolas, J. & Brooks, C. L. 3rd. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J Mol Graph Model*. **22**, 377-95 (2004).
2. Pandit, S. B., Zhang, Y. & Skolnick, J. TASSER-Lite: an automated tool for protein comparative modeling. *Biophys J*. **91**, 4180-90 (2006).
3. Tanizaki, S. & Feig, M. A generalized Born formalism for heterogeneous dielectric environments: application to the implicit modeling of biological membranes. *J Chem Phys*. **122**, 124706 (2005).

Low-resolution modeling of protein-ligand complexes

Michal Brylinski, Hongyi Zhou, Seung Yup Lee and Jeffrey Skolnick
Center for the Study of Systems Biology, Georgia Institute of Technology
skolnick@gatech.edu

G-protein coupled receptors (GPCRs) are a large protein family of transmembrane receptors that mediate a variety of biological effects through pathways involving the activation of G-proteins. The numerous physiological processes controlled by GPCRs make them particularly important targets in the treatment of many diseases. Recently, we have developed a collection of tools to support ligand- and receptor-based virtual screening. Since the experimentally solved structures are available only for a small fraction of potential drug targets, including GPCRs, our goal was to develop a protocol that effectively utilizes low-quality protein structures as targets for ligand-binding site prediction and for flexible ligand docking. In this spirit, we developed FINDSITE, a threading-based method for ligand-binding pocket detection (1) and Q-Dock, a low-resolution flexible ligand docking approach that employs knowledge-based potentials (2).

In the CASP-GPCR experiment, the model of the human adenosine A_{2A} receptor was generated by pro-sp3-TASSER that uses PRO-SP3 threading method to identify structure templates and TASSER (3) to refine template models. PRO-SP3 threading employs 5 scoring functions from SP^3 (4) and PROSPECTOR_3 (5), where each threading score ranks templates independently. Top-ranked templates identified by each scoring function were used to derive contact and distant restraints for model refinement by TASSER simulations. Multiple short TASSER runs were carried out to generate an initial collection of full-length models. Subsequently, the top 20 models were selected by TASSER-QA (6) and used for the second round of TASSER modeling. Final models were selected from both rounds of TASSER runs by TASSER-QA. The top ranked model was then used to dock an antagonist, ZM241385, into the binding pocket using Q-Dock that employs the Replica Exchange Monte Carlo (REMC) protocol to optimize the binding mode of a small molecule inside the receptor pocket. The conformations of ZM241385 collected from REMC simulations were clustered using a 2 Å RMSD cutoff and the lowest energy pose was selected from each cluster. Ten complexes with the lowest receptor-ligand interaction energy were submitted as final models.

1. Brylinski, M. & Skolnick, J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci U S A* **105,** 129-34 (2008).
2. Brylinski, M. & Skolnick, J. Q-Dock: Low-resolution flexible ligand docking with pocket-specific threading restraints. *J Comput Chem* **29,** 1574-88 (2009).
3. Zhang, Y. & Skolnick, J. Automated structure prediction of weakly homologous proteins on a genomic scale. *Proc. Natl. Acad. Sci. (USA)* **101,** 7594-9 (2004).
4. Zhou, H. & Zhou, Y. Fold recognition by combining sequence profiles derived from evolution and from depth-dependent structural alignment of fragments. *Proteins* **58** 321-8 (2005).
5. Skolnick, J., Kihara, D. & Zhang, Y. Development and large scale benchmark testing of the PROSPECTOR_3 threading algorithm. *Proteins* **56,** 502-18 (2004).
6. Zhou, H. & Skolnick, J. Protein model quality assessment prediction by combining fragment comparisons and a consensus C(alpha) contact potential. *Proteins* **71,** 1211-8 (2007).

Liliana Ostopovici-Halip, Cristian Bologa
Division of Biocomputing, University of New Mexico
cbologa@salud.unm.edu

Homology modeling.

The sequence of the human A2A receptor was aligned with the sequences of the human beta-1 (PDB: 2VT4 [1]) and beta-2 (2RH1 [2]) adrenergic receptors, and the bovine (1U19 [3]) and squid (2Z73 [4]) rhodopsins. The alignments have been manually adjusted in order to keep the highly conserved residues aligned, to avoid insertions or deletions in the transmembranar domains (TM) and to allow the Cys residues in the TM3 and the third extracellular loop to form a sulfur bridge. The final alignments were submitted for automatic modeling to Swiss-Model [5,6] (4 models) and Modeller [7,8] (3 models) servers. The seven resulted models for the A2A receptor have been minimized using Biopolymer module from Sybyl software suite [9] to reduce the steric clashes of the side chains without changing the backbone of the receptor.

Docking.

Ligand docking was performed for each of the 7 models, using AutoDock 4.0 [10] and FRED 2.2.3 [11]. In both cases the grids were chosen to be sufficiently large to include not only the active site but also significant portions of the surrounding surface and to allow the ligand to rotate freely, even when it is in the most fully extended conformation. Except for the large grids, the default docking parameters implemented in FRED and Autodock were used. For docking with FRED, the conformers were generated by OMEGA [12], part of the OpenEye suite.

Ranking.

No specific ranking of the docking poses was employed, but rather a diverse set of 10 binding modes was selected and sent for evaluation.

1. Warne, A. et al. Structure of the Beta1-Adrenergic G Protein-Coupled Receptor *Nature* **454,** 486 (2008)
2. Cherezov, V.et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258-1265 (2007).
3. Okada, T. et al. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure *J.Mol.Biol.* **342:** 571-583 (2004)
4. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin *Nature* **453:** 363-367 (2008).
5. Arnold, K., Bordoli, L., Kopp, J. & Schwede T.. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201 (2006).
6. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381-3385 (2003).
7. Eswar, N. et al. Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., **Suppl 15**, 5.6.1-5.6.30, 200
8. Marti-Renom, M.A. et al. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291-325 (2000).
9. SYBYL 7.3, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA
10. Morris, G. M. et al. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function *J. Computational Chemistry* 19: 1639-1662 (1998).
11. OpenEye Scientific Software, Santa Fe, USA, FRED, http://www.eyesopen.com/products/applications/fred.html
12. OpenEye Scientific Software, Santa Fe, USA, Omega, http://www.eyesopen.com/products/applications/omega.html

Polo Chun-Hung Lam and Ruben Abagyan
The Scripps Research Institute
pololam@scripps.edu

All sequence alignment, molecular modeling, and virtual screening were carried out in ICM Pro[1] on 2.33 GHz Intel Dual Core XEON-EMT processors. The structure of the ligand ZM241385 was obtained from PubChem. Residues 1-305 of the SWISS-PROT entry AA2AR_HUMAN were retained and aligned with the sequence β1-adrenergic receptor. A crude homology model was generated using chain B of the X-ray model of β1-adrenergic receptor (PDB code: 2VT4) as template using ICM's homology modeling method[2]. This crude model was then used in a ligand-guided homology modeling method described in the following steps:
1. All side chain variables were sampled with a biased probability Monte Carlo procedure[3]. **2**. A hybrid full atom/grid model was generated: Loop residues 67-74, 155-165, 167-179, 206-229, 258-267 were represented by full atom model with flexible backbone and side chains, with its end residues tethered to a static copy. Residues within 7 Å of the loop residues were represented by full atom model with fixed backbone and flexible side chains. The rest of the receptor were represented by multiple grid maps. All the variables were sampled. The lowest energy conformation was used to generate a full atom model of the whole receptor. **3**. All backbone and conserved side chain atoms were tethered to the template. All side chain and backbone variables were minimized in iterative cycles with decreasing tether strength until the tether strength reached a pre-set minima and the energy converged. **4**. The ligand was docked to a grid representation of the receptor, a maximum of 40 ligand docking poses were kept. **5**. For each of the ligand docking poses, a sphere of residues within 5 Å of the ligand was carved out and represented by fully flexible atom model with backbone tethered to a static copy of itself, while the rest of receptor represented by grid map. It was then globally minimized by iterative cycles of Monte Carlo sampling of all side chain followed by minimization of all variables. A full atom model was then constructed.
A total of 12 independent runs were attempted, each taking approximately 140 hours, generating ~300 models. These models were then ranked by the ICM docking score and their enrichment factor in a set of 600 decoys seeded with 40 A2a antagonist at 1% scoring cutoff. The three best models were selected and submitted as the final models.

1. Abagyan, R., Totrov, M. & Kuznetsov, D. Icm: A New Method For Protein Modeling and Design: Applications To Docking and Structure Prediction From The Distorted Native Conformation. *J Comp Chem* **15**, 488-506 (1994).
2. Cardozo, T., Totrov, M. & Abagyan, R. Homology modeling by the ICM method. *Proteins* **23**, 403-14 (1995).
3. Abagyan, R. & Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* **235**, 983-1002 (1994).

Eric S. Dawson, Kristian Kaufmann, Nils Woetzel, Jens Meiler

Vanderbilt University, Center for Structural Biology

eric.dawson@vanderbilt.edu

Critical Assessment of GPCR structure – Prediction of adenosine A_{2a} receptor complex with inverse agonist

Models were generated from slightly perturbed 7TM bundles (0.0-0.3Å) of the β_2-adrenergic receptor 7TM protein backbone structure (PDB ID: 2RH1)[1] using rigid body translation of 7TM helices followed by low resolution scoring and then sidechain repacking in ROSETTA. The research group of Co-Investigator Jens Meiler recently reported application of the ROSETTA approach to generate models for serotonin transporter proteins in complex with substrate analogs[2]. The general protocol involves 1) alignment of the target sequence with the structural homolog template protein using a variable scoring function and the dynamic programming algorithm[3] to optimize the alignment of secondary structural elements and conserved regions in transmembrane (TM) spans; 2) generation of the backbone TM spans by copying coordinates from the template protein; 3) sampling of an ensemble of protein backbones via a combination of protein fragment replacement (including building of loop regions) with gradient energy-based refinements. The inverse agonist compound ZM241385 was docked into the 7TM bundles using ROSETTALIGAND[4] prior to attachment of extracellular loops.

Ligand docking into the 7TM bundle alone proved unsuccessful in identification of the native inverse agonist binding pose as our best scoring model from the largest binding pose cluster (Figure 1) had an 8.5 Å RMSD compared with the native ligand binding mode (PDB ID: 3EML). Our initial docking strategy resulted in omission of the ECL2/3 contribution to the inverse agonist ligand binding site and therefore from the docking pose scoring calculations leading us to prioritize poses from other local energy minima sampled by the calculations. This receptor system also features three unexpected non-conserved ECL disulfide bonds that contribute to energetic stabilization of an inverse agonist binding site with substantial contributions from the receptor ECL region. Future investigations of GPCR structure prediction should place a strong emphasis on loop sampling and scoring as well as prediction of disulfide bonds prior to ligand docking and model refinement calculations.

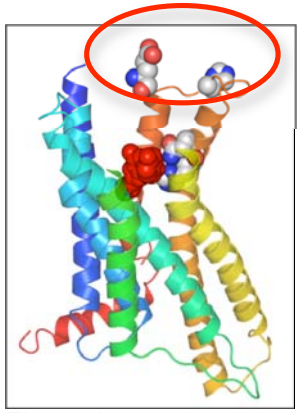


Figure 1. Model of A_{2a} adenosine receptor with top scoring binding pose from the largest cluster of inverse agonist ZM241385 (red spheres) docked into 7TM bundle with ECL added after docking calculation. Loop contributions of key residues to ligand binding site (red oval) are therefore missing from the energy score.

References:

1. Cherezov, V., et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258-65 (2007).
2. Kaufmann, K. W., et al.. Structural Determinants of human and Drosophila Serotonin Transporter Species Selective Substrate Recognition Revealed through Docking Studies. *Proteins: Structure, Function, and Bioinformatics* **74**, 630-642 (2008).
3. Dong, E., Smith, J., Heinze, S., Alexander, N. & Meiler, J. BCL::Align-sequence alignment and fold recognition with a custom scoring function online. *Gene* 422, 41-46 (2008).
4. Meiler, J. & Baker, D. ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility. *Proteins* 65, 538-548 (2006).

Feng Ding, Adrian Serohijos, Shuangye Yin, and Nikolay V. Dokholyan
University of North Carolina at Chapel Hill
dokh@med.unc.edu.

Homology Modeling

To obtain an initial alignment of the human A2A Adenosine Receptor (hA2A) with GPCR proteins with known crystal structures, we used CLUSTALW (www.ebi.ac.uk/clustalw) to perform a sequence alignment of the hA2A receptor, bovine rhodopsin (PDB ID: 1u19), squid rhodopsin (PDB ID: 2z73), human β1-adrenergic receptor (PDB ID: 2vt4) and human β2-adrenergic receptor (PDB ID: 2rhi). From the sequence similarity and using the fact that adrenergic receptor ligands are non-covalently bound as opposed to that of the rhodopsin, we chose the human β1-adrenergic receptor as the best template for constructing the hA2A model. To optimize the conformation of the intra- and extra-cellular loops, we performed discrete molecular dynamics simulations (DMD) (1) where the transmembrane helices were constrained while the loops were allowed to move. Finally, in the absence of homology to any proteins of known structure (as determined from 3DJury (http://meta.bioinfo.pl/)), we folded *ab initio* the C-terminal domain of hA2A using all-atom DMD (1). To validate the final model, we evaluated the surface charge distribution of the transmembrane region (i.e., packing of hydrophobic residues with the lipid bilayer and hurriedness of charged/polar residues in the protein core), the formation of conserved disulfide bonds and the agreement with known mutagenesis studies (compiled in (2)).

Ligand pose prediction

From the homology model of the hA2A protein, we identified the conserved binding pocket of small-molecule ligands. The target ligand was docked to the hA2A binding pocket using AutoDock(3, 4). In order to rank the ligand poses with our recently developed MedusaScore (6), we refined each AutoDock-derived poses using the MedusaScore force field. The ligand was subject to the rigid-body motions and the protein sidechain flexibility around the ligand-binding pocket was modeled by discrete rotamer libraries (5). For each ligand pose from AudoDock, we locally minimize the total energy with iterative protein sidechain repacking and ligand rigid-body motions. We ranked the energy-minimized docking poses using the MedusaScore program (6). In MedusaScore, the protein-ligand binding strength is evaluated using a linear combination of van der Waals, solvation and hydrogen bonding interactions. The highest-ranking models are selected as the putative binding poses.

References

1. Ding, F., Tsao, D., Nie, H. F. & Dokholyan, N. V. Ab initio folding of proteins with all-atom discrete molecular dynamics. *Structure* 16, 1010-1018 (2008).
2. Cristalli, G., Lambertucci, C., Marucci, G., Volpini, R. & Dal Ben, D. A(2A) adenosine receptor and its modulators: Overview on a druggable GPCR and on structure-activity relationship analysis and binding requirements of agonists and antagonists. *Curr Pharma Design* **14**, 1525-1552 (2008).
3. Goodsell, D. S. & Olson, A. J. Automated Docking of Substrates to Proteins by Simulated Annealing. *Proteins-* **8**, 195-202 (1990).
4. Morris. G. M., et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* **19**, 1639-1662 (1998).
5. Dunbrack, R. L. & Cohen, F. E. Bayesian Statistical Analysis of Protein Side-Chain Rotamer Preferences. *Protein Sci* **6**, 1661-1681 (1997).
6. Yin, S., Biedermannova, L., Vondrasek, J. & Dokholyan, N. V. MedusaScore: An Accurate Force-Field Based Scoring Function for Virtual Drug Screening. *J Chem Inform Modeling* **48**, 1656-1662 (2008).

David Rodriguez and Hugo Gutiérrez-de-Terán
Fundación Pública Galega de Medicina Xenómica. Complejo Hospitalario Universitario de Santiago, A Choupana, s/n. Santiago de Compostela (SPAIN).

HOMOLOGY MODELLING *of the* hA_{2A} *ADENOSINE RECEPTOR*. The target sequence was aligned with the sequences of the three candidate template sequences: human b2 adrenergic receptor (PDB code 2RH1),[1] turkey b1 adrenergic receptor (PDB code 2VT4) [2] and bovine rhodopsin (PDB code 1U19).[3] Multiple sequence alignment was performed by ClustalX,[4] using the PAM250 substitution matrix, with open and elongationon gap penalties of 10 and 0.05, respectively followed manual inspection in order to avoid gaps withing helices and proper alignment in loops.

Homology models were generated by Modeller v9.3.[5] Cysteines at positions 77 and 166 were identified as the ones involved in the well-known disulfide bridge between H3 and ECL2 in class A GPCRs, thus the condition was imposed in Modeller. Additionally, we detected a potential disulfide bridge between residues 146 and 159 (ECL2-ECL2). We explored the secondary structure features of ECL2 with three different servers: PSIPRED,[6] JPRED,[7] and APSSP2 [8] All of them agreed in a consensus beta-sheet secondary structure between positions 163 to 168 and the corresponding restraint was added. Separate modelling was done considering *i)* the use of one of the three crystallized GPCRs as a template, *ii)* combination of two templates *iii)* imposing or not the potential second disulphide bridge. The program generated 8 models each time, the quality of which was assessed by looking at the stereochemistry with Procheck [9] and visual inspection, considering available mutagenesis data [10]. Selected models were further refined with the aid of the Molprobity server[11].

LIGAND-PROTEIN DOCKING The ligand ZM241385 was draw in 2D with Marvin Sketch (Chemaxon, Inc) and a single 3D reliable conformation generated with OMEGA.[12] Automated docking was carried out by GOLD,[13] allowing full flexibility for the ligand, using a sphere of 12 or 15 Å centered on the Cε of HIS278 (necessary radius varied with the protein model used) and 20 GA runs per docking. Both Chemscore and Goldscore scoring functions were considered.[14]

The best complexes were obtained using the protein models generated considering the structure of hβ2 adrenergic receptor (2RH1) as template. 5 complexes from 3 different models were selected and energy minimized with MOE, using MMFF94x force field. Final solutions were ranked attending to the orientation of side chains of the residues that were reported to be important by mutagenesis data [10] in ZM241385 antagonism.

REFERENCES

1. Cherezov et al, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258-1265 (2007).
2. Warne et al, Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **454**, 486-491 (2008).
3. Okada et al, The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J. Mol. Biol.* **342**, 571-583 (2004).
4. Larkin M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
5. Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815 (1993).
6. Bryson, K. et al. Protein structure prediction servers at University College London. *Nucl. Acids Res.* **33**, W36-38 (2005).
7. Cole, C., Barber, J. D. & Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucl Acids Res.* **36**, W197-W201 (2008).
8. Raghava, G. P. S. APSSP2 : A combination method for protein secondary structure prediction based on neural network and example based learning. CASP5. A-132 (2002)
9. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. *J. App. Cryst.* **26,** 283 (1993)
10. Martinelli, A. & Tuccinardi, T., Molecular Modeling of Adenosine Receptors: New Results and Trends. *Med Res Rev* **28**, 247-277 (2008).
11. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucl Acids Res* **35**, W375-W383 (2007).
12. Kirchmair, J., Wolber, G., Laggner, C. & Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J Chem Inf Model.* **46**, 1848-61 (2006).
13. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking *J. Mol. Biol.*, **267**, 727-748, (1997)
14. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray C. W. & Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins*, **52**, 609-623, (2003)

Henri Xhaard
Centre for Drug Research
Faculty of Pharmacy, University of Helsinki
henri.xhaard@helsinki.fi

Adenosine 2A receptor modeling

Known three-dimensional structures of GPCRs (Turkey β1-Adrenergic receptor, human β2-Adrenergic receptor[1], squid and bovine rhodopsins[2]) were first retrieved and superimposed. A multiple sequence alignment was then constructed that contained the amino acid sequences of the human and rat AA1R, AA2R, AA3R sequences, as well as these sequences from known GPCRs. The known three-dimensional structures were used to guide the location of insertion/deletions in the sequence alignment.

The transmembrane core of the AA2R was modeled using the human β2-adrenergic receptor (PDB code 2RH1) as a structural template together with the pairwise sequence alignment described above and the software Modeller 8.0[3].

The second extracellular loop (EC2) was modeled independently and a single folding hypothesis was considered (due to time constraints). A disulphide bridge connecting C164 in EC2 to a cysteine residue on the top of TM3 was formed, similarly to other known GPCRs. A region of EC2 was remodeled using the rhodopsin 1U19 template: based on local sequence similarity, a fold of EC2 in the region equivalent to the beta-turn seen in the rhodopsin structures was considered, with E161 mimicking the structural role of E181 in rhodopsin in stabilizing EC2.

Rhodopsin I(179)P**E**(181)**GMQ**KSCG**C**(187)
AA2R C(159)G**E**(161)**G**-**Q**--*--VA**C**(164)

(*) corresponds to a deletion accommodated within the beta-turn

The remainer of the EC2 loop was freely modeled by Modeller 8.0 but for the region P(148)KEGKNHSQ(156) which was modeled similarly to the alpha-helix present in the human β2-adrenergic receptor structure.

Ligand docking

Ligand sketching and docking was performed manually using Sybyl 8.0[4]. The adenosine molecule was first docked and used to guide the docking of ZM241385. Three models were submitted where ZM241385 is positioned parallel to the plane of the membrane and interacting with W6.48; the second complex corresponds to a 180 degrees rotation of the ligand along its short axis, and the third to a lateral shift towards TM5. Due to the limited amount of time devoted to this docking study, external experimental data from the literature were not taken into account.

1. Cherezov, V. *et al*. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258-1265 (2007).
2. Okada, T. *et al*. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure *J.Mol.Biol*. **342:** 571-583 (2004)
3. Fiser A. and Sali A. MODELLER: Generation and refinement of homology-based protein structure models. *Meth. Enzymol*. **374,** 461-491 (2003).
4. SYBYL 7.3, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA