

# Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008

Mayako Michino\*, Enrique Abola\*, GPCR Dock 2008 participants, Charles L. Brooks III†, J. Scott Dixon§, John Moulth|| and Raymond C. Stevens¶

**Abstract** | Recent breakthroughs in the determination of the crystal structures of G protein-coupled receptors (GPCRs) have provided new opportunities for structure-based drug design strategies targeting this protein family. With the aim of evaluating the current status of GPCR structure prediction and ligand docking, a community-wide, blind prediction assessment — GPCR Dock 2008 — was conducted in coordination with the publication of the crystal structure of the human adenosine A<sub>2A</sub> receptor bound to the ligand ZM241385. Twenty-nine groups submitted 206 structural models before the release of the experimental structure, which were evaluated for the accuracy of the ligand binding mode and the overall receptor model compared with the crystal structure. This analysis highlights important aspects for success and future development, such as accurate modelling of structurally divergent regions and use of additional biochemical insight such as disulphide bridges in the extracellular loops.

Molecular modelling has an important role in rational drug design<sup>1,2</sup>. Reliable three-dimensional models can provide valuable insights into basic principles of molecular recognition and aid in structure-based approaches to lead discovery and optimization<sup>3</sup>. G protein-coupled receptors (GPCRs) are membrane proteins involved in signal transduction pathways and are important therapeutic targets for numerous diseases<sup>4,5</sup>. As such, significant structure prediction efforts using methods ranging from *de novo* to homology-based approaches have been applied to members of the GPCR family<sup>6,7</sup>.

Until recently, most GPCR homology modelling efforts have been based on the templates of bovine rhodopsin and bacteriorhodopsin, with refinement of the models achieved through molecular dynamics simulations, ligand docking and incorporation of additional biochemical and biophysical data<sup>8–12</sup>. The refinement step is necessary in building accurate models, especially around the ligand-binding site, owing to the expected structural differences among members of the family. These differences result from the generally low sequence identity and the large diversity of ligands accommodated within the family<sup>7,13–15</sup>, and from the various conformational states that are associated with different levels of ligand efficacy<sup>16–18</sup>.

The most recently solved GPCR structure is the 2.6 Å crystal structure of the human adenosine A<sub>2A</sub> receptor bound to an antagonist<sup>19</sup>. Adenosine receptors belong to the class A rhodopsin-like GPCR family and represent promising therapeutic targets in a wide range of conditions, including cerebral and cardiac ischaemic diseases, sleep disorders, immune and inflammatory disorders, and cancer<sup>20</sup>. The A<sub>2A</sub> receptor structure shows an overall seven transmembrane (TM) helix architecture similar to that of the rhodopsin and adrenergic receptor structures, but with shifts in the positions and orientations of the helices and a markedly different structure of the extracellular loops<sup>19</sup>.

To evaluate current progress in GPCR structure prediction and the docking of potential ligands, as well as highlight areas for future efforts in method development, we carried out a community-wide, blind prediction assessment — GPCR Dock 2008 — in coordination with the publication of the human adenosine A<sub>2A</sub> receptor structure in October 2008 (REF 19). GPCR Dock 2008 was organized in a similar manner to the previous CASP (Critical Assessment of methods of Protein Structure) and CAPRI (Critical Assessment of PRediction of Interactions) studies<sup>21,22</sup>. In this paper, we report the

\*Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, USA.

†Department of Chemistry and Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109, USA.

§Daylight Chemical Information Systems Inc., Aliso Viejo, California 92656, USA.

||Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850, USA.

¶Departments of Molecular Biology and Chemistry, The Scripps Research Institute, La Jolla, California 92037, USA

Correspondence to R.C.S. e-mail: [stevens@scripps.edu](mailto:stevens@scripps.edu)

doi:10.1038/nrd2877

Published online 22 May 2009

## Rhodopsin and bacteriorhodopsin

These two light-activated membrane proteins have a seven transmembrane alpha-helical bundle architecture that is similar to the general structure of the larger GPCR family.

outcome of the assessment together with our analysis of the current status of GPCR structure and ligand docking predictions.

## GPCR Dock 2008

In August 2008, before the publication of the human adenosine A<sub>2A</sub> receptor structure in October 2008 (REF. 19) and public release of the three-dimensional coordinates, participants were asked to predict and submit up to ten ranked models of the human A<sub>2A</sub> receptor in complex

with the ligand ZM241385, starting from the amino acid sequence of the receptor and a two-dimensional structure of the ligand (see BOX 1 for list of GPCR Dock 2008 participants). A total of 63 different groups initially registered, with 206 models submitted by 29 different groups in the final data set (see [Supplementary information S1 \(box\)](#) for details). Of the 206 submitted models, 37 were either missing the ligand or had incorrect bond connectivity for the ligand. We assessed the remaining 169 models for the prediction accuracy of the ligand binding mode, and all 206 models were assessed for the prediction accuracy of the receptor alone.

## Box 1 | GPCR assessment participants

- Arthur Olson: Department of Molecular Biology, The Scripps Research Institute, USA
- Wiktor Jurkowski and Arne Elofsson: Center of Biomembrane Research, Department of Biochemistry & Biophysics, Stockholm University, Sweden
- Sławomir Filipek: Laboratory of Biomodelling, International Institute of Molecular and Cell Biology, Poland
- Irina Pogozheva and Andrei Lomize: Peptide Synthesis and Molecular Recognition Laboratory, University of Michigan, USA
- Bernard Maigret: Orpailleur team, LORIA, Nancy University, France
- Jeremy Horst, Brady Bernard, Shyamala Iyer and Ram Samudrala: Computational Biology Group, University of Washington, USA; Ambrish Roy and Yang Zhang: Department of Molecular Biosciences, Center for Bioinformatics, University of Kansas, USA
- Osman Ugur Sezerman: Biological Science and Bioengineering, Sabanci University, Turkey
- Gregory V. Nikiforovich: MolLife Design LLC, USA; Christina M. Taylor: Department of Biochemistry and Molecular Biophysics, Washington University, USA
- Stefano Costanzi: Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, USA
- Y. Vorobjev, N. Bakulina, and V. Solovyev: Department of Computer Science, Royal Holloway, University of London and Softberry Inc., UK
- Kazuhiko Kanou, Daisuke Takaya, Genki Terashi, Mayuko Takeda-Shitaka and Hideaki Umeyama: School of Pharmacy, Kitasato University and RIKEN Systems and Structural Biology Centre, Japan
- William A. Goddard III, Youyong Li, Soo-Kyung Kim, Bartosz Trzaskowski, Ravinder Abrol and Adam Griffith: Materials and Process Simulation Center, California Institute of Technology, USA
- Vsevolod Katritch, Manuel Rueda and Ruben Abagyan: Molsoft LLC, USA
- Ian Davis, Patrick Barth and David Baker: Department of Biochemistry, University of Washington, USA
- Michael Feig: Department of Biochemistry and Molecular Biology, Michigan State University, USA
- Michal Brylinski, Hongyi Zhou, Seung Yup Lee and Jeffrey Skolnick: Center for the Study of Systems Biology, Georgia Institute of Technology, USA
- Liliana Ostopovici-Halip and Cristian Bologa: Division of Biocomputing, University of New Mexico, USA
- Polo Lam and Ruben Abagyan: Department of Molecular Biology, The Scripps Research Institute, USA
- Eric S. Dawson, Kristian Kaufmann, Nils Woetzel and Jens Meiler: Center for Structural Biology, Vanderbilt University, USA
- Feng Ding, Adrian Serohijos, Shuangye Yin and Nikolay V. Dokholyan: Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, USA
- David Rodríguez and Hugo Gutiérrez-de-Terán: Fundación Pública Galega de Medicina Xenómica, Complejo Hospitalario Universitario de Santiago de Compostela, Spain
- Henri Xhaard: Center for Drug Research, Faculty of Pharmacy, University of Helsinki, Finland

For full details, see [Supplementary information S1 \(box\)](#).

**Assessment criteria.** Assessment criteria are dependent on the purpose of the generated models. Given the value of the GPCR structural models in expanding our knowledge in basic molecular recognition and their potential use in the design and development of new small molecules, the quality of the models was primarily assessed by the accuracy of the ligand binding mode. Particular attention was given to the fact that the crystal structure is a static structure with positional errors, and the value of modelling is ultimately to guide drug discovery and provide biological insight. Our numerical measure of accuracy for the ligand binding mode was based on two metrics: ligand root mean square deviation (RMSD) and the number of correct receptor–ligand contacts. Neither metric alone was sufficient to capture the accuracy of prediction around the ligand binding site; hence, both were used and combined into a z-score to rank the models.

The ligand RMSD between the model and the crystal structure was calculated as the coordinate RMSD for the 25 non-hydrogen atoms of ZM241385 after superimposing the C $\alpha$  atoms of the protein in the model and the crystal structure. In addition, the ligand RMSD is also calculated excluding the phenoxy group of ZM241385 that has high B-factor values. The number of correct contacts is counted as the number of correctly predicted native contacts observed between protein atoms and the ligand. A native contact is defined as any interatomic distance within 4 Å of the ligand in the crystal structure. There are 75 such receptor–ligand contacts, and an additional 15 contacts formed with water.

The models were ranked by assigning a combined mixed z-score to each model. The combined z-score was calculated as the average of z-scores for ligand RMSD and the number of correct contacts:

$$Z_{\text{combined}} = (-Z_{\text{ligand RMSD}} + Z_{\text{Number of correct contacts}})/2.$$

The z-scores for ligand RMSD and the number of correct contacts were computed by the following steps. First, a z-score was assigned to each model using the average and standard deviation (SD) values from all models. Second, the average and SD was re-computed excluding models with z-scores that were more than two SDs above (for ligand RMSD) or below (for the number of correct contacts) the average. Third, a z-score was reassigned to each model using the revised average and SD values obtained in step two. The best model — that is, the model with the highest combined z-score — from each group was analysed.

### Molecular dynamics simulation

This molecular modelling approach uses numerical integration to solve the equations of motion based on the forces arising from interatomic interactions. The dynamic behaviour of atoms in a macromolecular system, such as that in a membrane protein, can be understood by running a molecular dynamics (MD) simulation. MD simulation can also be used to refine structural models of proteins and protein–ligand complexes.

### Ligand docking

A molecular modelling approach that predicts the ligand binding mode within a targeted binding site. In this approach, the known or predicted three-dimensional structure of a protein is probed using computationally generated energy landscapes to identify the most favourable binding pose for the ligand.

### RMSD (root mean square deviation)

RMSD is used as a quantitative measure of the similarity between two superimposed atomic coordinates. RMSD values (units of Å) can be calculated for any type and subset of atoms; for example, Ca atoms of proteins (Ca RMSD) for all residues, for residues in the transmembrane helices or the loops; heavy atoms of small-molecule ligands (ligand RMSD).

### Z-score

A standard dimensionless score that normalizes a value with respect to the sample mean and standard deviation.

### Ca atoms

The chiral carbon atoms to which the primary amine, the carboxylic group and the side chain are attached to in an amino acid. Comparison of three-dimensional structures of proteins is sometimes carried out by superimposing the Ca atoms of proteins as this provides a simple estimate of the similarity of their skeleton or backbone structure.

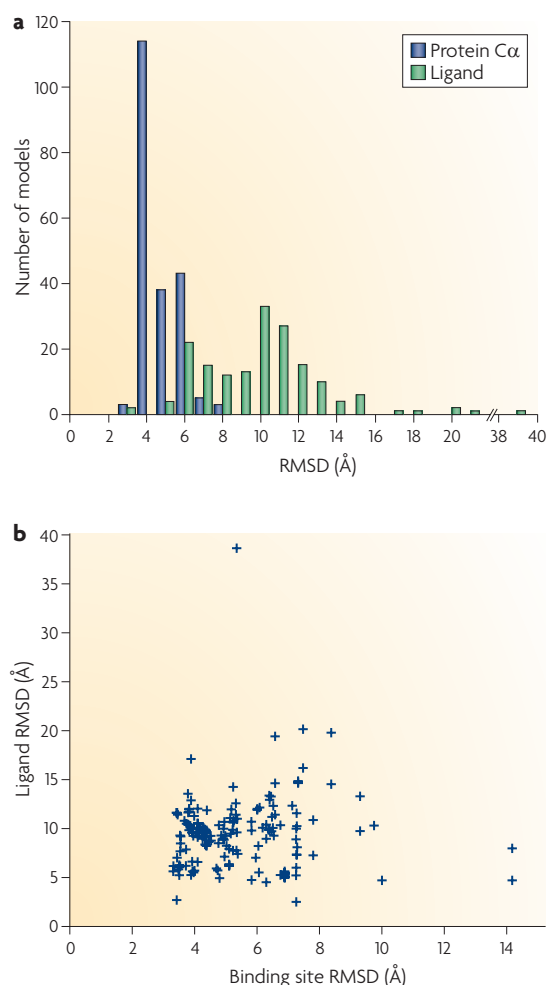
**Overall outcome of analysis.** The submitted models showed a wide distribution in prediction accuracy of the ligand binding mode, with average values of 9.5 Å (SD 3.8 Å) for ligand RMSD (FIG. 1a) and 4 (SD 7) for the number of correct contacts. These statistics indicate that the majority of the submitted models did not predict the ligand position and the binding interactions very accurately. The lack of a strong correlation between ligand RMSD and binding site RMSD (FIG. 1b) (for example, models with a binding site RMSD of less than 4.0 Å have a range of 2.8 to 17.2 Å ligand RMSD), suggests that the performance of some ligand docking methods can be improved.

Very few models score well in both ligand RMSD and the number of correct contacts (only 13 out of the 169 total receptor–ligand models have a combined z-score greater than 1, compared with 40 models that score well solely in ligand RMSD ( $Z_{\text{ligand RMSD}}$  less than  $-1$  Å). For models with relatively low ligand RMSD values but a small number of correct contacts, the inaccuracy in binding interactions could be attributed to errors in the side chain placement of the ligand binding residues. Although nearly a third of the models capture the hydrogen bonding interaction between the N253<sup>6.55</sup> side chain and the exocyclic N15 atom of the ligand (44 out of 169 models have a N253 OD1–ZM241385 N15 interaction distance of less than 4 Å), other key receptor–ligand interactions, such as the aromatic stacking interaction between the F168<sup>5.29</sup> side chain and the bicyclic ring of the ligand, are not captured well in most models (FIG. 2).

Although the overall outcome clearly shows that there are remaining challenges in accurately predicting the ligand binding mode, the quality of the predictions for the receptor alone seem relatively good:  $4.2 \pm 0.9$  Å for the receptor Ca RMSD, and  $2.8 \pm 0.5$  Å for the TM helices Ca RMSD. Not surprisingly, loop regions, with the exception of the short intracellular loop 1 (ICL1), are not modelled very well in most of the models (FIG. 3a,b and FIG. 4a,b). It is notable that some groups that accurately predicted the TM region of the receptor did not predict the ligand binding mode very well (for example, the TM Ca RMSD is 2.0 Å for the model submitted by I. Pogozheva and A. Lomize, and 2.1 Å for the model submitted by J. Horst and A. Roy), indicating that the methods for modelling the receptor and docking of the ligand can be generally considered as distinct steps in the generation of models for the receptor–ligand complex.

### Analysis of the best models

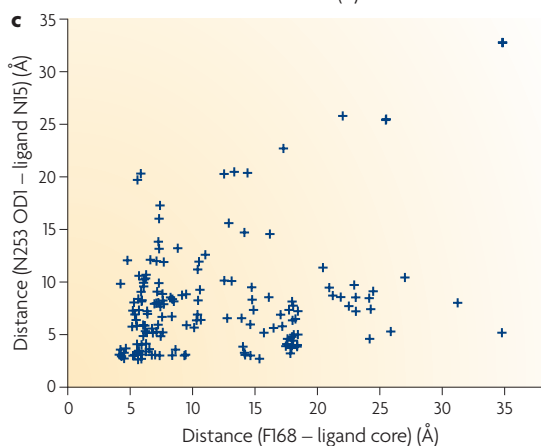
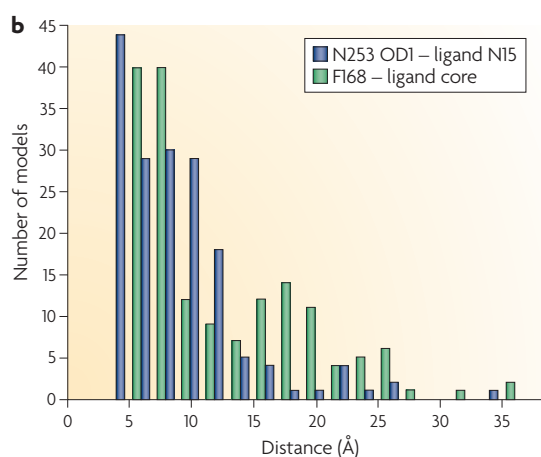
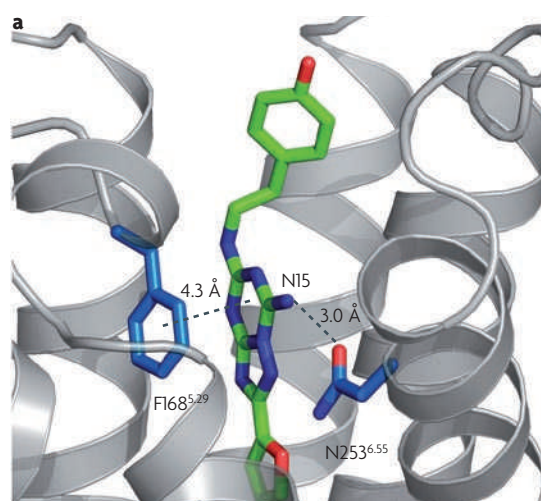
Despite the challenges in accurately predicting the receptor–ligand interactions, some models had consistent features with the crystal structure, although model ranking continues to be one of the most challenging areas of development. Here, we focus on the predictions from the top ten groups, ranked according to the combined z-score, and assess the model quality in greater detail (FIG. 4c). Note that, with predictions for only one target, the statistical significance of the group ranking cannot be judged as is typically done in CASP experiments by a head-to-head comparison of common targets between



**Figure 1 | Root mean square deviation (RMSD) of submitted models.** **a** | Distribution of ligand RMSD (green bars) and protein Ca RMSD (blue bars) for all models. **b** | A scatterplot of ligand RMSD (y axis) versus binding site RMSD (x axis) for all models. The binding site RMSD values are calculated for heavy atoms of the binding site residues (F168<sup>5.29</sup>, E169<sup>5.30</sup>, M177<sup>5.38</sup>, W246<sup>6.48</sup>, L249<sup>6.51</sup>, H250<sup>6.52</sup>, N253<sup>6.55</sup>, H264<sup>6.66</sup>, M270<sup>7.35</sup>) after the models were superimposed to the crystal structure using the Ca atoms of the protein.

the top groups<sup>23</sup>. To support our selection of the best predictions, we ranked all models using an alternative metric — binding site contact RMSD — which gives all ligand binding residues equal weight and is an RMSD of receptor–ligand contact distance for all ligand-binding residues. We found that both the z-score ranking and the contact RMSD ranking agree on the selection of the best model.

The best model overall (submitted by S. Costanzi) has a ligand RMSD of 2.8 Å and 34 of 75 correct contacts (FIG. 5a and TABLE 1). The ligand is modelled in a native-like binding pose, with an extended conformation and a nearly perpendicular orientation to the membrane plane. The model accurately predicts some of the key receptor–ligand interactions: it captures the hydrogen



**Figure 2 | Statistics of the two key receptor–ligand interactions in all models.** **a** | The hydrogen bonding interaction with N253<sup>6.55</sup> and the aromatic stacking interaction with F168<sup>5.29</sup> are shown by dashed lines with the distance measurements from the crystal structure. **b** | Distribution of the distance for the interaction between the side chain carbonyl oxygen OD1 atom in N253<sup>6.55</sup> and the exocyclic N15 atom of the ligand (ZM241385), and the average interatomic distance for the aromatic stacking interaction between the heavy atoms in the F168<sup>5.29</sup> side chain and the bicyclic ring (atoms C11, N12, N13, C14, N15, N16, N17, C18, N19, C20) of the ligand. **c** | A scatterplot of the distances for the hydrogen bonding interaction (y axis) versus the aromatic stacking interaction (x axis).

bonding interaction between the N253<sup>6.55</sup> side chain and the exocyclic amino group (N15 atom) of the ligand, and the aromatic stacking interaction between the F168<sup>5.29</sup> side chain and the bicyclic triazolotriazine core of the ligand. Compared with the crystal structure, the ligand in the model is positioned deeper in the binding pocket, bringing the furan ring closer to TM helices III and V. The inaccuracy in the ligand position is most probably due to errors in the side chain positions of the two crucial ligand binding residues (F168<sup>5.29</sup> and E169<sup>5.30</sup>) in extracellular loop 2 and the side chain orientation of M177<sup>5.38</sup> at the extracellular end of TM helix V. The aromatic ring of F168<sup>5.29</sup>, which interacts with the bicyclic ring, is positioned too deeply; the adjacent E169<sup>5.30</sup> forms a hydrogen bonding interaction with the hydroxyl group in the phenolic substituent, instead of the exocyclic N15 atom near the bicyclic ring; and the side chain of M177<sup>5.38</sup> is not oriented towards the binding cavity. In addition, the family conserved disulphide bond between C77<sup>3.25</sup>–C166<sup>5.27</sup> is predicted accurately, but the disulphide bond in extracellular loop 3 between C259<sup>6.61</sup>–C262<sup>6.64</sup> is not, presumably contributing to the inaccuracy in the side chain orientation of H264<sup>6.66</sup>, which is not pointed towards the binding site.

The best predictions from the top six groups (S. Costanzi; V. Katritch and R. Abagyan; P. Lam and R. Abagyan; I. Davis, P. Barth and D. Baker; B. Maigret; W. Jurkowski and A. Elofsson) highlight the successes and challenges in accurately predicting the ligand binding pose and receptor–ligand interactions (FIG. 5B,C,D and TABLE 1). The extended ligand conformation is accurately predicted in all six models, and the nearly perpendicular orientation is captured in four of the six models. The hydrogen bonding interaction between the N253<sup>6.55</sup> side chain and the exocyclic N15 atom of the ligand is correctly modelled in four models; however, in one of the four, the ligand makes no interaction with residues in extracellular loop 2. The aromatic stacking interaction between the F168<sup>5.29</sup> side chain and the bicyclic ring of the ligand is correctly modelled in four models; however, in all four models, the ligand is positioned too deeply in the binding pocket, and the M177<sup>5.38</sup> side chain is not oriented towards the binding cavity. There is one model that does not accurately capture either the hydrogen bonding interaction with N253<sup>6.55</sup> or the aromatic stacking interaction with F168<sup>5.29</sup>, whereas five of the six models accurately predict the family conserved disulphide bond between C77<sup>3.25</sup>–C166<sup>5.27</sup>. None of the six models capture the hydrogen bonding interaction between E169<sup>5.30</sup> in extracellular loop 2 and the exocyclic N15 atom of the ligand.

Other models that ranked near the top (those submitted by K. Kanou, W. A. Goddard, C. Bologna and A. Olson) are slightly less accurate, but show similar trends to the top six models in their ability to accurately predict the ligand binding mode (TABLE 1). The ligand is modelled in a native-like extended conformation in three of the four models. The hydrogen bonding interaction between the N253<sup>6.55</sup> side chain and the exocyclic N15 atom of the ligand is modelled accurately in three of the four models, whereas the aromatic stacking interaction

#### B-factor

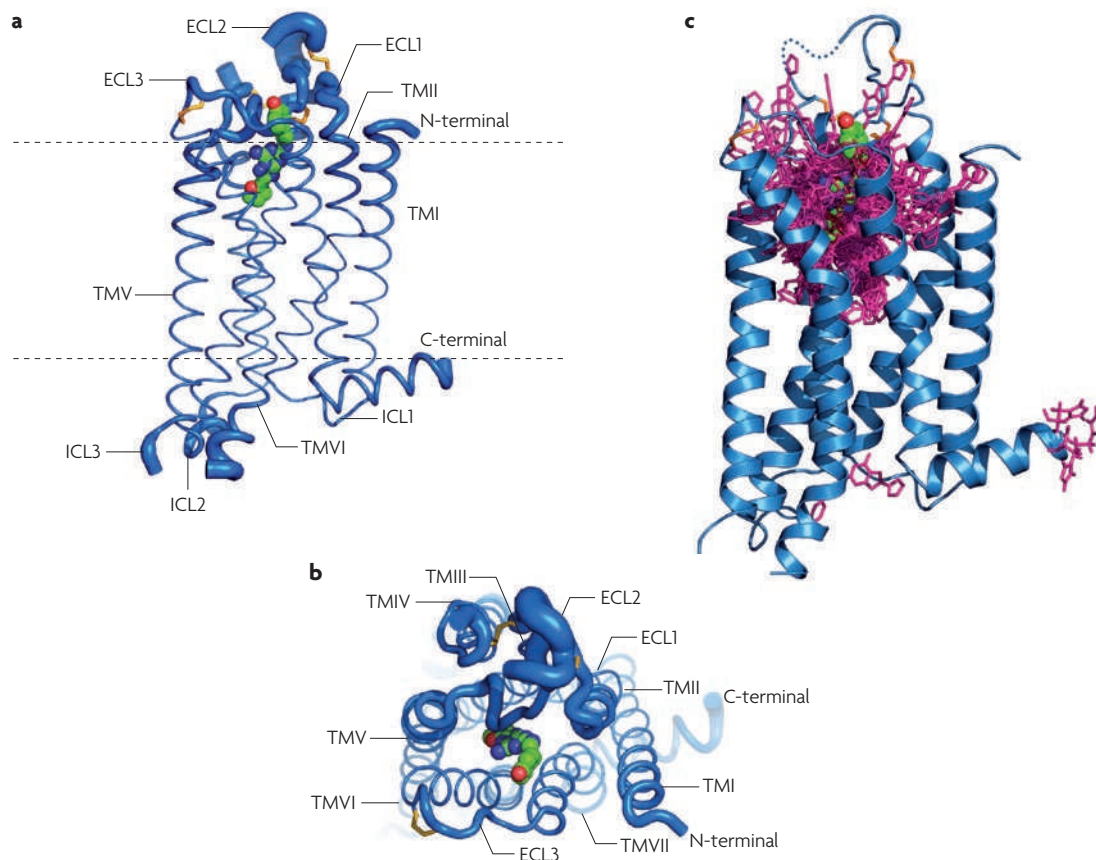
A descriptor that reflects the fluctuation of atomic position from an atom's average position and provides important insight into a protein's potential dynamic behaviour.

#### Hydrogen bond

Attractive interaction between one electronegative atom and a hydrogen covalently bonded to another electronegative atom such as nitrogen or oxygen.

#### Aromatic stacking

Attractive interactions between the aromatic rings of amino acids. Overlapping of p-orbitals of  $\pi$ -conjugated systems result in the rings arranging themselves in preferred orientations.



**Figure 3 | Superposition of all 206 submitted models to the crystal structure of the human adenosine A<sub>2A</sub> receptor.** Protein Ca atom superposition between each model and the crystal structure was done using the align command in PyMOL (version 1.0r2, [www.pymol.org](http://www.pymol.org)) (Protein Data Bank ID code: 3EML without the T4-lysozyme). **a,b** | The receptor is shown as two orthogonal views of Ca traces, with tube thickness being proportional to the root mean square deviation (RMSD) of each Ca position, showing how well the transmembrane (TM) regions were modelled and how much uncertainty there is in the loop regions. **c** | A superposition of stick diagrams of the ligand (ZM241385) from 169 models; a CPK model is used to delineate the observed position in the crystal structure. The carboxy-terminus (residue numbers greater than 306) is removed from all models. ECL, extracellular loop; ICL, intracellular loop.

between the F168<sup>5,29</sup> side chain and the bicyclic ring of the ligand is modelled accurately in only one of the four models. The family conserved disulphide bond between C77<sup>3,25</sup>–C166<sup>5,27</sup> is captured in two models. Remarkably, one of the models (submitted by W. A. Goddard) accurately places the E169<sup>5,30</sup> side chain proximal to the exocyclic N15 atom of the ligand, and almost captures the hydrogen bonding interaction, even though the overall conformation of extracellular loop 2 is inaccurate.

The best predictions were generally not ranked as the best models by the predictors at the time of model submission (before the release of the crystal structure) (TABLE 1). Only two of the six best models were ranked first, and three of the six groups show a weak correlation between their model ranking and the model quality as assessed by the combined z-score for the accuracy around the ligand-binding site. Furthermore, the additional models submitted by the six groups are generally of lower quality than the best predictions (TABLE 1). Only one of the six best models has a z-score that is within one SD of the group average z-score.

### Status of GPCR structure modelling and docking

The assessment of the submitted models showed that the best participating methods have the ability to predict close, native-like ligand binding, but have limitations in capturing all of the key receptor–ligand interactions and correctly estimating model quality by ranking. The majority of the submitted models are quite far from predicting a native-like ligand binding pose. The most challenging aspect of GPCR structure prediction highlighted in this assessment seems to be in accurately modelling the ligand interactions with residues in the extracellular loop regions. This result is not surprising given the lack of structural homology in the loops among the known GPCR structures<sup>24</sup>, and the general difficulties in modelling loop regions<sup>25,26</sup>.

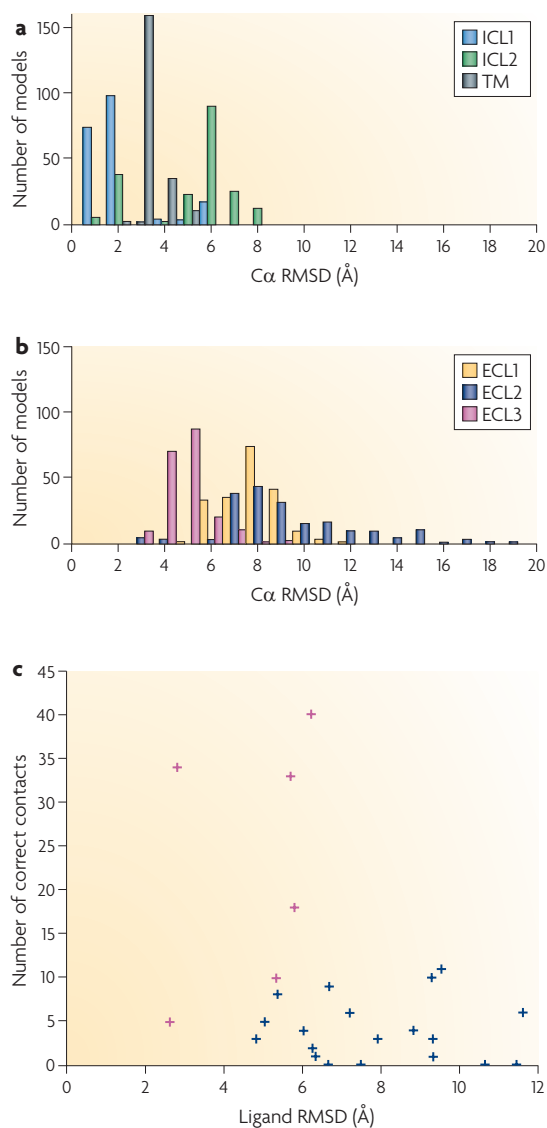
The most successful prediction methods relied on homology modelling approaches based on the template structures of  $\beta$ -adrenergic receptors, and in some cases with the additional template structures of rhodopsin (Protein Data Bank ID code: 2RH1 ( $\beta_2$ AR), 2VT4 ( $\beta_1$ AR), 1U19 (bovine rhodopsin), 2Z73 (squid rhodopsin)) to

generate models of the receptor, followed by docking of the ligand to one or more receptor models using small-molecule docking programmes such as Glide<sup>27</sup>, ICM<sup>28</sup>, GOLD<sup>29</sup> and AutoDock<sup>30</sup> (see [Supplementary information S1 \(box\)](#) for description of prediction methods). The alignment of the human A<sub>2A</sub> receptor sequence to the

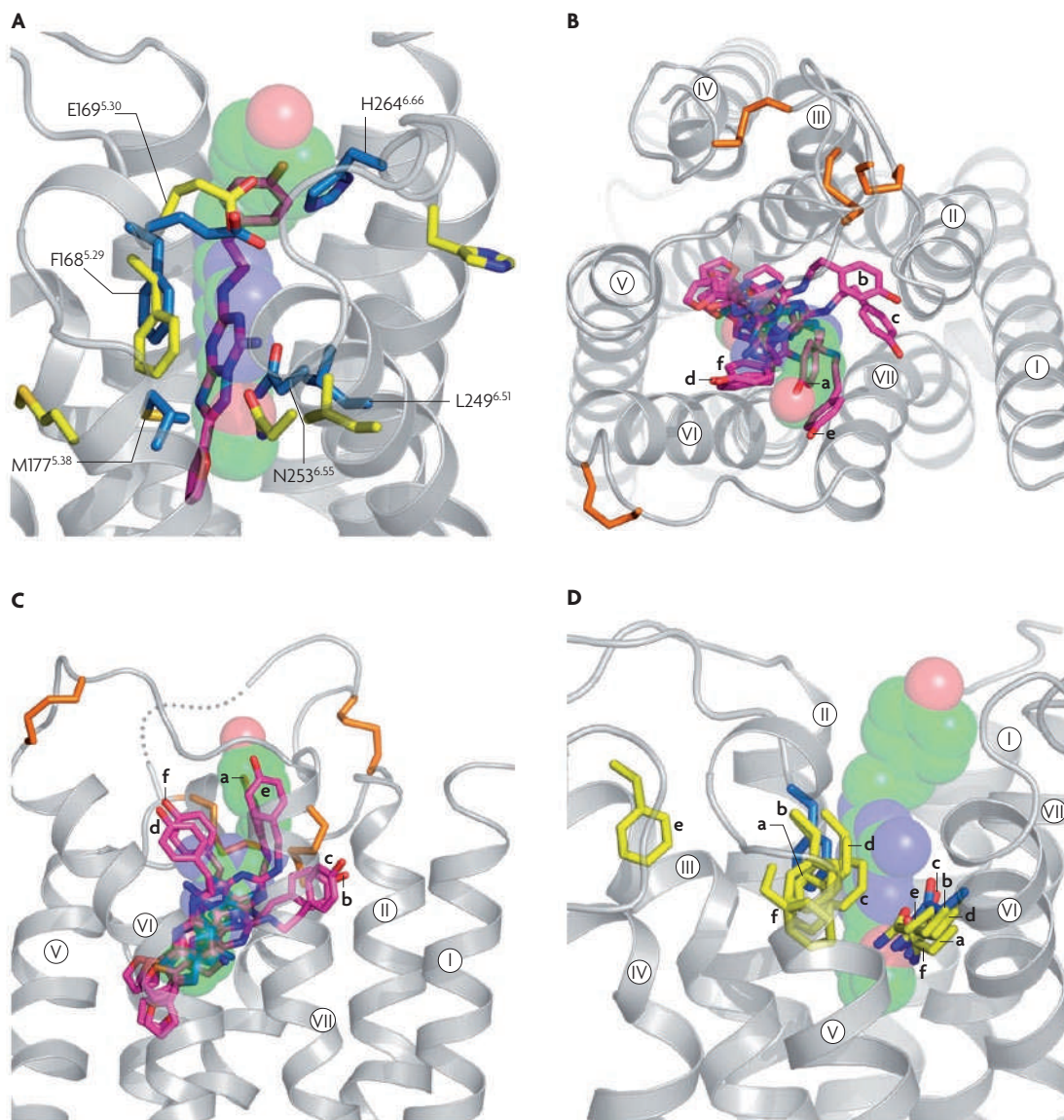
template structure seemed to have been straightforward, given the family conserved motifs and residues in the TM helices<sup>31</sup>. The extracellular loop 2 was modelled by *de novo* approaches in many of the top predictions (V. Katritch and R. Abagyan; P. Lam and R. Abagyan; I. Davis, P. Barth and D. Baker; W. Jurkowski and A. Elofsson; W. A. Goddard), but only partially modelled in the best prediction (S. Costanzi) for a short segment of eight residues, located amino-terminal to TM helix V, which includes the disulphide bond forming C166<sup>5,27</sup>. Some of the criteria used to select and rank the final receptor–ligand complex models were: docking scores, conformational energy of the complex, agreement with mutagenesis and structure–activity relationship data, and binding selectivity studied by virtual ligand screening or by modelling other subtypes of adenosine receptor.

The reliability of the homology modelling approach depends on the availability of suitable templates<sup>32</sup>. The results of the current assessment show that the structures of  $\beta$ -adrenergic receptors alone or together with rhodopsin were suitable transmembrane templates in predicting the general structure of the adenosine A<sub>2A</sub> receptor. However, given the expected structural diversity in class A GPCRs, it is unclear whether the current set of techniques applied to the structure prediction of the A<sub>2A</sub>–ZM241385 complex would result in a similar level of accuracy for the prediction of other GPCRs, especially for those belonging to subfamilies that are phylogenetically distant from the amine and the opsin receptor clusters<sup>33</sup>. We believe the database of GPCR structures needs to expand further to provide suitable templates for accurate modelling of those other receptors.

The inaccuracies in homology models can arise from errors in side chain packing, main chain shifts in aligned regions, errors in unaligned loop regions, misalignments and incorrect templates<sup>34</sup>. These errors relate to the issue of ‘adding value’ to the template structure, which was addressed in the recent CASP experiment<sup>35</sup>, and also seems to be applicable to GPCR modelling. Indeed, ligand interactions with residues located in structurally divergent regions from the templates are consistently not modelled accurately in all of the six best predictions: the hydrogen bonding interaction between E169<sup>5,30</sup> in extracellular loop 2 and the exocyclic N15 atom of the ligand is not captured, and the side chains of H264<sup>6,66</sup> in extracellular loop 3 and M177<sup>5,38</sup> in the extended bulge structure unique to the A<sub>2A</sub> receptor at the extracellular end of TM helix V are not oriented towards the binding site. An exception is the aromatic stacking interaction between F168<sup>5,29</sup> in extracellular loop 2 and the bicyclic ring of the ligand, which is correctly modelled in some of the predictions. F168<sup>5,29</sup> is located in the loop, but it is structurally homologous to F193<sup>5,32</sup>, which interacts with the carbazole heterocycle of the ligand carazolol in the  $\beta_2$ AR structure; hence modelling of this interaction may have been guided by homology. Interestingly, F168<sup>5,29</sup> is modelled more accurately than E169<sup>5,30</sup> even though mutagenesis data showed that mutation of E169<sup>5,30</sup> to alanine reduces the affinity for both antagonists and agonists<sup>36</sup>, and no data is available for F168<sup>5,29</sup>.



**Figure 4 | Model analysis. a,b** | Distribution of C $\alpha$  root mean square deviations (RMSDs) for adenosine A<sub>2A</sub> receptor domains. Panel **a** shows C $\alpha$  RMSDs for transmembrane (TM) helices and intracellular loop (ICL) regions: helix I: 6–34; helix II: 40–67; helix III: 73–107; helix IV: 117–142; helix V: 173–205; helix VI: 222–258; helix VII: 266–291; ICL1: 35–39; ICL2: 108–116. Panel **b** shows C $\alpha$  RMSDs for extracellular loop (ECL) regions: ECL1: 68–72; ECL2: 143–172 (excluding 149–155 that are missing in the crystal structure); ECL3: 259–265. **c** | Scatterplot of the number of correct contacts versus ligand RMSDs for the best predictions from all groups. The best predictions from the top six groups are marked as mauve crosses.



**Figure 5 | Comparison between the best models and the crystal structure around the ligand-binding site.** The ligand and the ligand-binding residues F168<sup>5,29</sup>, E169<sup>5,30</sup>, M177<sup>5,38</sup>, L249<sup>6,51</sup>, N253<sup>6,55</sup> and H264<sup>6,66</sup> are shown for the best model (S. Costanzi) and the crystal structure (A). The ligand is shown as magenta sticks for the model, and as green semitransparent spheres for the crystal structure; the ligand-binding residues are shown as yellow sticks for the model, and blue sticks for the crystal structure. Extracellular (B) and side views (C) of the ligand in the binding pocket for the best predictions from the top six groups (magenta sticks for models and green spheres for the crystal structure). The receptor crystal structure is shown as grey ribbons. The disulphide bonds are shown as orange sticks. D | The ligand-binding residues F168<sup>5,29</sup> and N253<sup>6,55</sup> are shown as sticks for the best predictions from the top six groups (yellow for models and blue for the crystal structure). In B–D the models are labelled as: a, S. Costanzi; b, V. Katritch; c, P. Lam; d, I. Davis; e, B. Maigret; f, W. Jurkowski.

The inaccuracy in the orientation of the ligand binding pose — for example, the parallel orientation with the phenolic substituent positioned close to TM helices II and III — may in part be due to the inaccurate modelling of the helical shifts in TM helices I, II and III. The helical shifts alter the location of the binding pocket and redefine the pocket size and shape<sup>19</sup>; thus, it is expected that accurately modelling the helical shifts would contribute to a better prediction of the

ligand binding pose. The helical shifts were most accurately modelled by an effective use of multiple template structures of rhodopsin and  $\beta$ -adrenergic receptors (I. Pogozheva and A. Lomize), or an all-atom refinement approach implemented by the ROSETTA programme using a physically realistic model that recapitulated protein interatomic and protein–solvent interactions in the membrane environment<sup>37</sup> (I. Davis, P. Barth and D. Baker).

Table 1 | Summary of results for the best models from the top ranking groups

| Group name           | Rank (total number of models) | Ligand RMSD (Å) | Ligand RMSD without phenoxy group (Å) | Number of correct contacts | Binding site residues RMSD (Å) | Protein C $\alpha$ RMSD (Å) | TM I–VII C $\alpha$ RMSD (Å) | ECL2 C $\alpha$ RMSD (Å) | Combined z-score (average $\pm$ SD) |
|----------------------|-------------------------------|-----------------|---------------------------------------|----------------------------|--------------------------------|-----------------------------|------------------------------|--------------------------|-------------------------------------|
| Costanzi             | 2 (4)                         | 2.8             | 2.7                                   | 34                         | 3.4                            | 3.0 (266)                   | 2.5 (212)                    | 3.8 (8)                  | 3.02 (0.86 $\pm$ 1.48)              |
| Katritch & Abagyan   | 1 (10)                        | 6.2             | 4.0                                   | 40                         | 3.5                            | 4.0 (283)                   | 2.7 (214)                    | 8.9 (23)                 | 2.76 (1.89 $\pm$ 1.13)              |
| Lam & Abagyan        | 1 (3)                         | 5.7             | 3.6                                   | 33                         | 3.3                            | 4.1 (283)                   | 3.6 (214)                    | 7.3 (23)                 | 2.42 (0.88 $\pm$ 1.34)              |
| Davis, Barth & Baker | 4 (5)                         | 5.8             | 5.4                                   | 18                         | 4.0                            | 3.5 (283)                   | 2.1 (214)                    | 8.4 (23)                 | 1.46 (0.16 $\pm$ 0.86)              |
| Maigret              | 8 (10)                        | 2.6             | 2.1                                   | 5                          | 7.3                            | 5.1 (283)                   | 4.1 (214)                    | 9.1 (23)                 | 1.23 (0.05 $\pm$ 0.57)              |
| Jurkowski & Elofsson | 2 (8)                         | 5.3             | 5.2                                   | 10                         | 3.9                            | 6.2 (283)                   | 2.9 (214)                    | 12.7 (23)                | 1.04 (–0.02 $\pm$ 0.98)             |
| Kanou                | 7 (10)                        | 5.4             | 5.5                                   | 8                          | 6.9                            | 3.5 (279)                   | 2.8 (214)                    | 7.1 (23)                 | 0.91 (0.66 $\pm$ 0.11)              |
| Goddard              | 8 (10)                        | 5.0             | 3.9                                   | 5                          | 4.8                            | 4.3 (284)                   | 2.5 (214)                    | 10.7 (23)                | 0.78 (0.16 $\pm$ 0.37)              |
| Bologa               | 3 (10)                        | 6.7             | 2.8                                   | 9                          | 3.9                            | 3.4 (278)                   | 2.5 (213)                    | 7.2 (19)                 | 0.72 (–0.14 $\pm$ 0.39)             |
| Olson                | 1 (9)                         | 4.8             | 4.7                                   | 3                          | 5.8                            | 3.5 (284)                   | 2.3 (214)                    | 7.5 (23)                 | 0.69 (–0.14 $\pm$ 0.58)             |

Participants were allowed to submit up to 10 models. Rank indicates the ranking that the participant assigned to their best model as determined in the GPCR Dock 2008 study with the total number of models submitted by that participant in parentheses. The root mean square deviation (RMSD) values were calculated for the heavy atoms of the ligand ZM241385 (all 25 atoms and partially without the phenoxy group), heavy atoms of the binding site residues (F168<sup>5,29</sup>, E169<sup>5,30</sup>, M177<sup>5,38</sup>, W246<sup>6,48</sup>, L249<sup>6,51</sup>, H250<sup>6,52</sup>, N253<sup>6,55</sup>, H264<sup>6,66</sup>, M270<sup>7,35</sup>), C $\alpha$  atoms of all residues, C $\alpha$  atoms of residues in the transmembrane helices (TM) I to VII (helix I: 6–34; helix II: 40–67; helix III: 73–107; helix IV: 117–142; helix V: 173–205; helix VI: 222–258; helix VII: 266–291), and C $\alpha$  atoms of residues in extracellular loop 2 (ECL2) (143–172 excluding 149–155 that are missing in the crystal structure). All RMSD values were obtained after the models were superimposed to the crystal structure, using the protein C $\alpha$  atoms in PyMOL (version 1.0r2, www.pymol.org). The assignment of residues in the ligand-binding site and the secondary structure elements is from the Protein Data Bank header section (PDB ID: 3EML). The number of residues used in the RMSD calculation is in brackets. The combined z-score value for the best model, as well as the average and standard deviation (SD) values for all models submitted by each group, are shown.

Other sources of error include not modelling the water molecules that are either structurally important or directly involved in ligand binding interactions<sup>3</sup>. The ligand binding cavity in the A<sub>2A</sub>-ZM241385 structure has four ordered water molecules<sup>19</sup>, yet none of the submitted predictions included water molecules. We tried re-docking the ligand to the crystal structure using ICM<sup>28</sup> and found that a native-like binding pose (within 1 Å heavy atom RMSD for the bicyclic ring and the furanyl substituent of the ligand, and less than 3 Å overall ligand RMSD) can be recovered without any water molecules, which suggests that water may not be critical for accurately predicting the ligand interactions. However, modelling water molecules together with the ligand might contribute to a better prediction of the ligand binding pose or affinity. Additional re-docking studies with the docking protocols used by the participating methods would help assess the effect of the water molecules, and the accuracy of the docking methods separately from that of the receptor modelling methods.

Finally, it is interesting that the best model was from the S. Costanzi group, which has previously worked on adenosine receptor modelling and docking. Their

domain knowledge on the adenosine receptor is likely to have been crucial for the evaluation and interpretation of the mutagenesis and ligand interaction data.

## Conclusions

Accurate prediction of GPCR structure and ligand interactions remains a challenge, and the approach will improve with the recent availability of experimentally solved GPCRs. Assessment of these predictions highlights similar issues addressed by the CASP predictions for template-based modelling targets; that is, the difficulty in loop modelling, refinement and improvement over the best available template and model ranking. Accurate modelling of the structurally divergent regions (such as the extracellular loops that form defined architectures), and disulphide bond formation affecting helix residue registry and helical shifts in the TM region seem to be crucial for accurately predicting the key ligand interactions in GPCRs, and this area is perhaps the most in need of technological development. Progress in GPCR modelling and docking will require further improvements in the current prediction methods to enhance the best available templates and generate models that will be more useful for applications in structure-based drug design.

1. Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **303**, 1813–1818 (2004).
2. Richon, A. B. Current status and future direction of the molecular modeling industry. *Drug Discov. Today* **13**, 665–669 (2008).
3. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Rev. Drug Discov.* **3**, 935–949 (2004).
4. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–1964 (2000).
5. Klabunde, T. & Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem* **3**, 928–944 (2002).
6. Becker, O. M. *et al.* G protein-coupled receptors: *in silico* drug discovery in 3D. *Proc. Natl Acad. Sci. USA* **101**, 11304–11309 (2004).
7. Ballesteros, J. & Palczewski, K. G protein-coupled receptor drug discovery: implications from the crystal structure of rhodopsin. *Curr. Opin. Drug Discov. Devel.* **4**, 561–574 (2001).
8. Bu, L., Michino, M., Wolf, R. M. & Brooks, C. L. III. Improved model building and assessment of the Calcium-sensing receptor transmembrane domain. *Proteins* **71**, 215–226 (2008).
9. Henin, J. *et al.* Probing a model of a GPCR/ligand complex in an explicit membrane environment: the human cholecystokinin-1 receptor. *Biophys. J.* **90**, 1232–1240 (2006).
10. Fowler, C. B., Pogozheva, I. D., LeVine, H., 3rd & Mosberg, H. I. Refinement of a homology model of the mu-opioid receptor using distance constraints from intrinsic and engineered zinc-binding sites. *Biochemistry* **43**, 8700–8710 (2004).
11. Evers, A. & Klabunde, T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* **48**, 1088–1097 (2005).
12. Manivet, P. *et al.* The serotonin binding site of human and murine 5-HT<sub>2B</sub> receptors: molecular modeling and site-directed mutagenesis. *J. Biol. Chem.* **277**, 17170–17178 (2002).
13. Archer, E., Maigret, B., Escricout, C., Pradayrol, L. & Fourmy, D. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **24**, 36–40 (2003).
14. Gershengorn, M. C. & Osman, R. Minireview: Insights into G protein-coupled receptor function using molecular models. *Endocrinology* **142**, 2–10 (2001).
15. Ballesteros, J. A., Shi, L. & Javitch, J. A. Structural mimicry in G protein-coupled receptors: implications of the high-resolution structure of rhodopsin for structure-function analysis of rhodopsin-like receptors. *Mol. Pharmacol.* **60**, 1–19 (2001).
16. Kobilka, B. K. & Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **28**, 397–406 (2007).
17. Bhattacharya, S., Hall, S. E., Li, H. & Vaidehi, N. Ligand-stabilized conformational states of human beta(2) adrenergic receptor: insight into G-protein-coupled receptor activation. *Biophys. J.* **94**, 2027–2042 (2008).
18. Kenakin, T. Efficacy at G-protein-coupled receptors. *Nature Rev. Drug Discov.* **1**, 103–110 (2002).
19. Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A<sub>2A</sub> adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217 (2008). **The human adenosine A<sub>2A</sub> receptor crystal structure served as the experimental template for comparison for this modelling and docking assessment. This is the second human GPCR structure to be experimentally determined.**
20. Jacobson, K. A. & Gao, Z. G. Adenosine receptors as therapeutic targets. *Nature Rev. Drug Discov.* **5**, 247–264 (2006).
21. Moul, J. *et al.* Critical assessment of methods of protein structure prediction-Round VII. *Proteins* **69** (Suppl. 8), 3–9 (2007). **The very successful CASP (Critical Assessment of Protein Structure) project started in 1994 and served as the model to conduct the reported GPCR Dock 2008 modelling and docking assessment.**
22. Lensink, M. F., Mendez, R. & Wodak, S. J. Docking and scoring protein complexes: CAPRI 3rd Edition. *Proteins* **69**, 704–718 (2007).
23. Kopp, J., Bordoli, L., Battey, J. N., Kiefer, F. & Schwede, T. Assessment of CASP7 predictions for template-based modeling targets. *Proteins* **69** (Suppl. 8), 38–56 (2007).
24. Kobilka, B. & Schertler, G. F. New G-protein-coupled receptor crystal structures: insights and limitations. *Trends Pharmacol. Sci.* **29**, 79–83 (2008).
25. Jacobson, M. P. *et al.* A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351–367 (2004).
26. Rohl, C. A., Strauss, C. E., Chivian, D. & Baker, D. Modeling structurally variable regions in homologous proteins with rosetta. *Proteins* **55**, 656–677 (2004).
27. 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).
28. Totrov, M. & Abagyan, R. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins* **29** (Suppl. 1), 215–220 (1997).
29. 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).
30. Morris, G. *et al.* Automated docking using a lamarkian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
31. Mirzadegan, T., Benko, G., Filipek, S. & Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **42**, 2759–2767 (2003).
32. Baker, D. & Sali, A. Protein structure prediction and structural genomics. *Science* **294**, 93–96 (2001). **The use of protein models and docking is dependent on how such data will be used. In this paper, Baker and Sali provide an excellent presentation of where models are useful, in particular as hypothesis generators with the application being dependent on the resolution of the structure.**
33. Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).
34. Marti-Renom, M. A. *et al.* Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325 (2000).
35. Read, R. J. & Chavali, G. Assessment of CASP7 predictions in the high accuracy template-based modeling category. *Proteins* **69** (Suppl. 8), 27–37 (2007).
36. Kim, J. *et al.* Glutamate residues in the second extracellular loop of the human A<sub>2A</sub> adenosine receptor are required for ligand recognition. *Mol. Pharmacol.* **49**, 683–691 (1996).
37. Barth, P., Schonbrun, J. & Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc. Natl Acad. Sci. USA* **104**, 15682–15687 (2007).

#### Acknowledgements

We thank M. Hanson, V.-P. Jaakola, C. Roth and V. Cherezov for help with the analysis and comments on the manuscript, and K. Kadyshchevskaya and V. Cherezov for figure preparation. We are grateful to the Goddard group for providing the script to calculate the binding site contact RMSD. We thank A. Walker for data tracking and assistance with the manuscript and J. Kunken for IT help during the assessment. This work was supported in part by the Protein Structure Initiative grant U54 GM074961 (ATCG3D), the NIH Roadmap grant P50 GM073197 (JCIMPT), and the Multiscale Modeling Tools for Structural Biology NCCR via grant P41 RR012255.

#### SUPPLEMENTARY INFORMATION

See online article: [S1 \(box\)](#)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF