Published in final edited form as:

J Chem Inf Model. 2019 September 23; 59(9): 3744–3754. doi:10.1021/acs.jcim.9b00175.

Prediction of Conformation Specific Thermostabilizing Mutations for Class A GPCRs

Suvamay Jana^{1,#}, Soumadwip Ghosh^{1,#}, Sanychen Muk¹, Benjamin Levy¹, Nagarajan Vaidehi^{1,*}

¹Department of Computation and Quantitative Medicine, Beckman Research Institute of the City of Hope, 1500 E. Duarte Road, Duarte, CA 91010.

Abstract

G protein-coupled receptors (GPCRs) are highly flexible and prone to denaturation during protein extraction in detergents, and purification. This poses a huge challenge to purify conformationally homogenous solution of GPCRs. Thermostabilizing mutations have been used widely to purify and obtain crystal structures of several GPCRs. However, identifying thermostabilizing mutations for GPCRs remains a tedious and expensive task as the thermostabilizing mutations from one GPCR are not transferable among closely related other GPCRs. Additionally, the mutations stabilizing one conformational state of a GPCR do not always stabilize other conformational state(s) of the same GPCR. Previously we developed a computational method, LiticonDesign, for rapid prediction of thermostabilizing mutations for a specific GPCR conformation. In this study, we have used LiticonDesign to predict thermostabilizing mutations for the agonist bound active-intermediate state of the human adenosine receptor (A_{2A}R) using the structure of the inactive state of the same GPCR and vice versa. Our study shows that the thermostable mutation predictions using LiticonDesign, for an active-intermediate state of a GPCR (A_{2A}R in our case) requires a homology model that is derived from an active/active-intermediate state GPCR structure as template. Similarly, the homology models derived from inactive state GPCR conformations are better in predicting the thermostable mutations for the inactive state of A_{2A}R. Overall, our LiticonDesign method is not only efficient in predicting thermostabilizing mutations for a given GPCR sequence, but also can recover conformation specific mutations for a state of interest, if a suitable starting structure of desired conformation is chosen.

Graphical Abstract

*To whom correspondence should be addressed. NVaidehi@coh.org.

#Contributed equally to this work

Supporting Information

Calculation of percentage recovery of thermostable mutants, performance of homology models derived using various class A GPCRs with varying sequence identity to A_{2A}R, LiticonDesign predictions of mutations destabilizing human A_{2A}R (A) active-intermediate and (B) inactive conformational states by starting from ensemble of crystal structures of and homology models of the respective and the other functional states. This information is available free of charge via the internet at <http://pubs.acs.org>.

Conflict of interest

The authors declare no competing financial interest.

Keywords

GPCR; Thermal stability; Conformational sampling; LiticonDesign; A_{2A}R

1. Introduction

G-protein coupled receptors (GPCRs) constitute a superfamily of signaling proteins of critical importance to cellular function and are the largest group of drug targets. The three dimensional structures of GPCRs greatly aid structure based design of drugs.^{1–3} However, purification and crystallization of GPCRs in various conformational states are particularly challenging, since they are highly dynamic proteins that exhibit conformational heterogeneity, such as antagonist bound inactive state, agonist bound active-intermediate state, and agonist and G-protein bound fully active state.⁴ These different conformational states are characterized by the extent of the outward movement of transmembrane (TM) helix TM6 compared to TM3, as shown in Figure 1A. In the inactive state TM6 is packed close to TM3, while in the active-intermediate state TM6 undergoes an outward movement away from TM3 and this distance increases markedly in the G-protein bound fully active state.⁵ Many of the GPCRs with published three dimensional structures have been stabilized in a specific ligand bound conformation by mutating multiple positions that yielded thermostable mutant GPCRs. The thermostable mutants facilitated the formation of well-diffracting crystals.^{6–7} The thermostable mutations in GPCRs are derived by systematic mutagenesis of all the positions in a given GPCR sequence or by directed evolution techniques.^{8–12} The systematic mutagenesis involves multiple steps: (i) systematic alanine mutation of every residue in the receptor to identify residue positions leading to thermostabilization, and (ii) combining the positions that showed thermal stability to make a synergistic combinatorial mutant with higher thermostability that will aid crystallization. This strategy has been successful for multiple GPCRs in various conformational states.^{13–16} However, the process of deriving an optimal combination of thermostable mutations gets expensive especially in the second step when the number of combinations to be tested gets larger. In addition to the tedious process of deriving thermostabilizing mutations by systematic alanine scanning mutation, the thermostabilizing mutations of the inactive state of a GPCR for example, are not transferable to stabilizing other conformational states of the

same GPCR. For example, the thermostabilizing mutations of the inactive state of the antagonist bound human adenosine A_{2A}R receptor⁵ are not the same as those that stabilize the agonist bound active-intermediate state of A_{2A}R.²¹

Computational prediction has been proven helpful in identifying thermostable point mutations that aid purification/crystallization of GPCRs. Computational methods based on analysis on void space in GPCR structures,¹⁷ sequence conservation and structural properties combined with support vector machine technique^{18, 19} were used successfully to predict thermostabilizing mutations in GPCRs. Yasuda et al used a modified free energy function in identifying stabilizing mutations.²⁰

To expedite the identification of thermostable mutations in GPCRs, previously we developed a computational method LiticonDesign (Ligand Induced Transmembrane Conformational Change Design) and validated the method for mutations of class A GPCRs.^{22–23}

LiticonDesign generates an ensemble of conformations starting from an initial structure by systematically rotating all of the TM helices along their helical axis, followed by single point mutations of the hydrophobic residues of the transmembrane region of the receptor to alanine, and if alanine to leucine. Previously, we showed that an ensemble of conformations generated by LiticonDesign improved the overall predictability compared to the predictions made from a single structure.²¹ However, we did not address in detail which starting structure would be preferred to generate LiticonDesign ensemble that can maximize the predictability for identifying conformation specific thermostabilizing mutations. Here, we have studied the application of LiticonDesign computational method in detail for predicting thermostable mutations for a specific conformation of a GPCR using the crystal structures of the other conformational states of the same GPCR, and also from the homology models of similar conformational states. We chose the human Adenosine receptor A_{2A}R as model for this study since the experimental data on the conformation specific thermostabilizing mutations are reported for the inactive state and the active-intermediate states of this receptor¹⁴ and the crystal structures of different (inactive, active-intermediate and fully active) conformational states are also available for A_{2A}R in the protein data bank.^{5, 21} Our study showed that the prediction of thermostablizing mutants for the active-intermediate conformational state is sensitive to the starting conformation of the structural model used in LiticonDesign. This is because the thermostable mutations for the active-intermediate state are mostly clustered in the intracellular part of TM5, TM6, and TM7 of the receptor. Thus homology models of A_{2A}R derived using active-intermediate or fully active state conformation structures of other GPCRs as templates recover more of the positives thermostable mutants for the active states compared to using the crystal structures of the inactive state of the same receptor, A_{2A}R. In contrast, the percentage recovery of the thermostable mutants of the inactive state using crystal structures of active-intermediate state of A_{2A}R is comparable to that recovered by the inactive state homology models. Overall, our computational LiticonDesign method is useful for not only predicting thermostabilizing GPCR mutations, but also for recovering conformation specific mutations, if a suitable starting structural model is used.

2. Methods

2.1. LiticonDesign:

We have described the LiticonDesign method in detail in our previous work.^{22–25} Briefly, the LiticonDesign method for predicting thermostable GPCR mutants involves two steps: (i) using a starting structural model of the GPCR, the method generates a small ensemble of conformations that allows for the perturbations in the GPCR conformation caused by mutations and, (ii) an all-atom energy function to calculate the stability of the conformations that takes into account the difference in the structural stability of the mutants and the wild type to score the positive thermostable mutants. Starting from an initial receptor structural model, a homology model of the GPCR for example, all of the seven transmembrane (TM) helices are rotated simultaneously about their respective helical axes by $+5^\circ$ and -5° . Thus, $2^7=128$ conformations are generated for the wild type receptor. Then we perform mutation of each residue to alanine (alanine in the wild type is mutated to leucine) in each of the 128 conformations and repack the side chain conformations of all of the residues using SCWRL4.0,²⁶ followed by steepest descent (SD) energy minimization using CHARMM27 force field for 1000 steps.²⁷ The van der Waals energy component of the total potential energy (E_{vdw}) of the wild type and the mutant are then extracted for each of the 128 conformations, and averaged over the 128 conformations for the wild type and the mutant separately. The stability score is calculated as the difference in the average mutant energy to the average wild type energy. Stability score = $\langle E_{vdw, \text{mutant}} \rangle - \langle E_{vdw, \text{wild type}} \rangle$ for a given single mutant where the respective energies are averaged over the 128 conformations. To give equal weight to every amino acid type in our scoring, we calculated the z score from the stability score, for every amino acid type. We will refer to this z score as the stability score from here afterwards in the paper. We calculate the stability score for all the hydrophobic residue mutations in the TM region. This is because our previous studies have shown that most of the stabilizing inter-helical interactions in the TM region of GPCRs come from van der Waals interactions.^{9, 28–30} Additionally, since most of the mutations tested here are to alanine we observed that van der Waals energy captures the differences between mutants and the wild type. Hence, we considered only the differences in hydrophobic interactions between the mutant and the wild type.

2.2. Systems used for the study

We tested the accuracy of predictions of thermostabilizing single point mutations in A_{2A}R in the inactive state starting from each of the crystal structures of the inactive state (PDBs 3EML, 3PWH, 3REY, 3RFM, 4EIY, and 5NM4), the active-intermediate state (PDBs 2YDV, 3QAK, and 4UG2), and the fully active state conformations of A_{2A}R (PDB 5G53).^{5, 21, 31–36} We also predicted the single point thermostabilizing mutations for the active-intermediate state of A_{2A}R, starting from the crystal structure of the fully active state, the active-intermediate state, and also from the inactive state using the same crystal structures, mentioned above. We used LiticonDesign for all the predictions. The goal behind these tests was to assess the accuracy of thermostabilizing mutation predictions of a GPCR, given its crystal structure in another conformational state. We also constructed several homology models of human A_{2A}R in different conformational states (fully active, active-intermediate, and inactive) using templates with varying sequence identity to A_{2A}R to predict the

thermostabilizing mutations. The objective here was to examine the performance of a homology model with LiticonDesign in the case where there are no crystal structures available. For predictions of the thermostable mutations of the active-intermediate conformation of A_{2A}R, we used homology models derived using the templates of fully active state conformations of human β 2 adrenergic receptor (PDB 3SN6,³⁷ sequence identity 31%), human muscarinic acetylcholine receptor (PDB 4MQS,³⁸ sequence identity 26%), mouse μ opioid receptor (PDB 5C1M,³⁹ sequence identity 24%), and active-intermediate conformation of rat neurotensin type 1 receptor (PDB 4XES,⁴⁰ sequence identity 22%). Here the GPCRs that are co-crystallized with either G protein or G-protein mimetic nano body are defined as the fully active state. For prediction of the thermostable mutations of the inactive state A_{2A}R, we used homology models derived using the crystal structures of the inactive states of human adenosine A₁ receptor (PDB 5UEN,⁴¹ sequence identity 54%), turkey β ₁ adrenergic receptor (PDB 2YCX,⁴² sequence identity 34%), and mouse μ -opioid receptor (PDB 4DKL,⁴³ sequence identity 24%) as templates.

2.3. Preparation of protein structures for thermostability predictions:

In this work, the LiticonDesign predictions were done for the ligand free (apo) state of the receptor in all the cases. If the crystal structures of GPCRs contained fusion proteins (e.g. T4 lysozyme, cytochrome b562), the fusion proteins were removed before using LiticonDesign predictions. The residues mutated in the crystal structures were all mutated back to the wild type receptor sequence using PyMOL mutagenesis tool.⁴⁴ The coordinates of the ligand, crystal waters, ions, lipid molecules, if present, were also removed from the crystal structures. The homology models of wild type A_{2A}R were constructed in different conformational states from several templates, as listed earlier using MODELLER software package.^{45–46} We chose three top scoring structural models out of 20 constructed for each template by the DOPE score for further process. Finally, after initial system preparation, 5000 steps of steepest descent minimization were performed in vacuum, for each of the crystal structure and the homology model using CHARMM27 force field.²⁷ The minimized structures, thus obtained, were used as the starting conformations for running LiticonDesign.

3. Results

3.1. The thermostable mutants of the active-intermediate state of A_{2A}R are clustered on TM5, TM6 and TM7 while the inactive state thermostable mutations are spread across the receptor:

The conformation specific thermostabilizing single point mutations for human A_{2A}R receptor were identified previously using a radio-labeled ligand binding assay in presence of either agonist 5'-N-ethylcarboxamidoadenosine (NECA) or antagonist ZM-241385 (ZMA).¹⁵ The mutations with stability quotients higher than the wild type (WT) were defined as thermostable mutants and used as benchmarking dataset in our study. These thermostable mutations are shown in Figure 1B.

The mutations that stabilize the agonist NECA bound active-intermediate state of A_{2A}R (shown in yellow in Figure 1B) are mostly clustered in the transmembrane helices TM5, TM6 and TM7. This is likely to reduce the flexibility of the TM5, TM6, and TM7 helices in

the agonist bound active-intermediate state of the receptor. Figure 1C shows the difference in root mean square deviation (RMSD) in backbone coordinates calculated between the inactive and the active-intermediate states of A_{2A}R crystal structures (PDBs 3PWH and 2YDV, respectively) for each residue. We observed that the backbone RMSD is close to 2 Å across TM1 to TM4, and above 3 Å near the intracellular part of TM5, TM6 and TM7. Thus, the point mutations located on TM5, TM6 and TM7 shown in Figure 1B may individually reduce the flexibility of TM helices to stabilize the active-intermediate state conformation of the receptor. Interestingly, the inactive state specific mutations (shown in magenta in Figure 1B) are spread across all the TM helices, excluding TM7. These mutations might lead to the reduction of the overall flexibility of the receptor.³⁰ Point mutants T88A, G114A, G118A, A203L, A204L, and A231L in the TM stabilize both the active-intermediate and the inactive state conformations (shown in green in Figure 1C).

3.2. Prediction of thermostable mutations for the active-intermediate and inactive conformations of A_{2A}R:

In this section, we used LiticonDesign to compute the stability scores for mutating each hydrophobic residue position in the TM region to Ala (and to Leu when the wild type is Ala) starting from several inactive (PDBs 3EML, 3PWH, 3REY, 3RFM, 4EIY, and 5NM4), active-intermediate (PDBs 2YDV, 3QAK, and 4UG2), and fully active state (PDB 5G53) crystal structures of human A_{2A}R receptor (see Methods section for details). These tests were done to assess the performance of LiticonDesign to identify the thermostable mutations specific to active-intermediate state and inactive state conformations of A_{2A}R starting from the crystal structure of other conformational states. Here, we defined the TM domain of A_{2A}R with residue ranges: 7 to 37 as TM1, 40 to 67 as TM2, 78 to 108 as TM3, 119 to 141 as TM4, 174 to 208 as TM5, 225 to 258 as TM6, and 267 to 292 as TM7, thus containing 160 hydrophobic residues, in total. Out of the 160 TM hydrophobic residues 10 mutations were experimentally found to thermostabilize only the active-intermediate conformation, 10 were found to thermostabilize only the inactive conformation and three were found to stabilize both conformational states. It is worth mentioning here that the mutation G118A is not considered by LiticonDesign since it is located in the junction of TM4 and ICL2.

Using LiticonDesign, the stability scores calculated for mutation of the hydrophobic residues in the TM region starting from various crystal structures were compared to the experimental thermostability scores specific to the active-intermediate and the inactive state conformation. Since, the goal of using the computational method to predict thermostabilizing mutations is to reduce the number of experimental trials, we have used the percentage recovery of thermostabilizing mutations as a function of different number of mutational trials to assess the prediction quality (shown in Figure 2). The percentage recovery was calculated as the ratio of the true positives recovered based on the predicted stability scores to the total number of true positives. There are 10 thermostable mutants each for active-intermediate and inactive states and 3 common to both states as shown in Figure 1B. The percentage recovery has been shown up to 50 mutational trials (nearly 30% of the total TM hydrophobic residues) since this number of mutations is testable for many academic laboratories. The percentage recovery of the thermostable mutations stabilizing both active-intermediate and inactive states is shown in Figures 2A and 2C respectively. The

percentage recovery of each conformation specific thermostabilizing mutations is shown in Figures 2B and 2D.

As observed in Figure 2, the percentage recovery varies from 10% to 65% (top 10 to top 50 mutation trials) for recovering conformation specific mutations. The stability scores computed using LiticonDesign method starting from the structure of the targeted conformational state yields better results than the random prediction (solid blue line). However, the quality of prediction varies significantly with the starting conformation used for thermostable mutation predictions. While identifying the thermostable mutations specific for the active-intermediate state using the crystal structures of different conformational states, the stability scores computed starting from the active-intermediate crystal structures, as expected, yields better percentage recovery than starting from the inactive state crystal structures of A_{2A}R (Figures 2A and 2B). However, it was interesting to note that the fully active state crystal structure of A_{2A}R (PDB: 5G53) had slightly better predictability than the inactive state crystal structures in recovering the active-intermediate state conformation specific mutations (Figure 2B), at least within top 30 predicted mutations.

Similarly, for predicting the stabilizing mutations that are specific to the inactive state, we observe that starting from the inactive state crystal structures could recover more mutations than the active-intermediate state structures, especially within top 30 mutation trials. In addition, we also observe that the active-intermediate crystals performed poorer than the random prediction to identify inactive state conformation specific mutations during the early phase of the mutation trials (top 30), but the predictability improves for higher mutational trials (Figure 2D). Interestingly, we observe that the predictions from the fully active state crystal structure are even worse compared to the active-intermediate state crystal structures for recovering conformation specific mutations corresponding to the inactive state (Figure 2D). Taken together, these results show that the starting conformation is indeed crucial, especially in identifying thermostabilizing mutations specific to a given state.

3.3. Crystal structures of a GPCR in the same conformational state show conformational heterogeneity:

The percentage recovery varies when using different inactive or active state crystal structures for LiticonDesign predictions as shown in Figures 2. To understand the origin of this variability, we compared the root mean square deviation (RMSD) in coordinates of the backbone atoms of the residues in the TM regions across all of the active-intermediate crystal structures and the inactive state crystal structures, as shown in Figure 3. As seen in Figure 3A, although all of the active-intermediate crystals represent a specific conformation, there are high local structural deviations indicated by RMSD of 6 Å. Similar trend was observed while comparing the backbone RMSD in the TM regions of the crystal structures of the inactive state (Figure 3B). This structural comparison suggests that the conformational space corresponding to the active-intermediate and the inactive state is broad, and the crystal structures used for prediction in the current study only represent a subset of the conformation of interest. Although our LiticonDesign method improves the conformation sampling, it may not be sufficient to sample the entire conformational space corresponding to a specific state. In this regard, it may be useful to consider multiple starting structures

representing the same conformational state to generate LiticonDesign ensemble to generate a wider ensemble of conformations to improve the predictability of thermostable mutations. In the next section we used an ensemble of crystal structures of A_{2A}R in the active-intermediate and inactive states to test the predictability starting from using multiple crystal structures representing the same conformation state.

3.4. Prediction of thermostabilizing mutations of A_{2A}R from an ensemble of starting structures and homology models of similar conformational state:

In a realistic scenario, for a blind prediction of thermostable mutations, a crystal structure of the target GPCR may not be available in public domain. Therefore, in this section we have used homology models of inactive and active state conformations of A_{2A}R to assess the performance of LiticonDesign. We also used multiple templates to derive multiple homology models of A_{2A}R in the active-intermediate state and in the inactive states to test the performance using an enriched ensemble of states. An additional objective was to investigate whether sequence similarity to the target or conformation specificity of the starting model or both are critical in predicting conformation specific thermostabilizing mutations.

The thermostable mutation predictions were done using LiticonDesign starting from an ensemble of conformations generated from multiple starting structures using two approaches. In the first approach, we used a combination of all the crystal structures representing a specific conformational state (active-intermediate or inactive) of A_{2A}R to generate an ensemble of structures to sample the conformational space of a specific state better compared to a single structure in that particular state. Thus, for the prediction of thermostable mutations of the active-intermediate state of A_{2A}R, we used three crystal structures of A_{2A}R in the active-intermediate states (PDB ID: 2YDV, 3QAK, and 4UG2) and generated 384 conformations (128 conformations for each starting structure) instead of only 128 to calculate the stability score. Similarly, for the inactive state predictions we used six crystal structures of the inactive state of A_{2A}R (PDB ID: 3EML, 3PWH, 3REY, 3RFM, 4EIY, and 5NM4) and generated 768 conformations (128 × 6) for the stability score calculations.

The second approach was to use an ensemble of conformations generated from multiple homology models of a given conformation state of A_{2A}R from a combination of homology models derived from multiple templates and retaining three homology model from each template. The homology models of A_{2A}R in the active-intermediate state were generated using the crystal structures of the fully active state conformations of β₂-adrenergic receptor (β₂AR), M₂ Muscarinic receptor, μ Opioid receptor, and active-intermediate conformation of the neurotensin receptor 1, NTSR1. The inactive state homology models of A_{2A}R were generated using the crystal structures of the inactive state conformations of adenosine receptor A₁R, β₁ adrenergic receptor β₁AR, and μ Opioid receptor with varying sequence identity (details of the structural templates used for homology modeling is shown in Table S1, Supporting Information) to human A_{2A}R receptor. The homology models were generated using MODELLER and the models with the top three most favorable DOPE scores for a given template were chosen to generate LiticonDesign ensemble of conformations (128×3 = 384 total conformations instead of 128 conformations when using

only one model) for each template for stability score calculations. We combined all of the LiticonDesign conformations generated from all of the homology models of the inactive state to calculate the stability score. We did the same for the active-intermediate state predictions of thermostable mutations. To compute the stability score here, the difference in vdw energy between the mutant and the wild type is averaged over all the conformations (instead of 128), followed by conversion to the z-score (see Methods section for details). The percentage recovery of the thermostable mutations is shown in Figure 4. In this figure, the percentage of thermostable mutants recovered using LiticonDesign starting from multiple crystal structures are shown in solid bars and in open bars for predictions made starting from multiple homology models. The percentage of thermostable mutants recovered from top 10 to top 50 predicted mutants are comparable starting from either multiple crystal structures or from multiple homology models. However, the predictions are sensitive to conformations used in the starting models. The thermostable mutations for the active-intermediate state are better recovered by using the homology models of the active-intermediate or active states rather than the inactive state homology models. The predictions for the inactive state thermostable mutations are not as sensitive to the conformation of the starting models as the predictions for the active-intermediate states. However, a slight improvement was observed in recovering the inactive state stabilizing mutations starting from inactive state homology models compared to the active-intermediate crystals, at least within top 20 of the predictions (Figure 4C and 4D). The percentage recovery for the inactive state is 30 to 40% in the top 50 prioritized mutant list. The rate of recovery for the inactive state is poor due to the large spread in the TM3-TM6 distance in the homology models, that shows the conformational heterogeneity in the receptor models which may not capture the heterogeneity seen in experiments. We also calculated the percentage recovery of true negatives (destabilizing mutants). The destabilizing mutations have negative stability scores (implying lower stability than the WT). We observed that the true negative mutations could also be recovered by the ensemble of homology models of the appropriate conformational state (Figure S1, Supporting Information). The data in Fig. S1 was calculated using the mutant list was sorted by the negative z-scores and we identified how many of the experimentally verified non-thermostable (147 each for active-intermediate and inactive state specific) mutations were being recovered within top 50 trials. We found that the recovery of these true negatives specific for active-intermediate states was highest using the active-intermediate crystals, followed by the active-intermediate homology models (57 and 51% respectively shown in Figure S1A, of the Supporting Information). On the other hand, mutations destabilizing the inactive state were best recovered (51%) using the ensemble generated by We have calculated the percentage recovery of thermostable mutations of various states as a function of the sequence identity of the structural template of the same conformational state used for building homology models. These values are shown in Table S1. There is a 60% recovery of the active-intermediate state mutations in the top 50 mutation trials using the homology model of A_{2A}R generated using the fully active state μ OR as template, although its sequence identity to A_{2A}R is only 24% (Table S1). The percentage recovery of active-intermediate state specific mutations afforded by closer homologs of A_{2A}R, namely, β_2 AR is lower than that of μ OR. The inactive state crystal structure of adenosine receptor 1, A₁R has the closest homology to A_{2A}R and shows the highest recovery of 54% in the top 50 mutation trial among all structural templates considered in this study enables the best recovery of inactive

state stabilizing mutations (Table S1). Thus, we conclude that sequence identity of the structural template used for building homology models is not as critical as its starting conformation in determining the percent recovery of mutations representing a particular functional state.

3.5. Assessing the performance of ensembles based on crystals/homology models in recovering conformation-specific thermostable mutations:

We analyzed how well an ensemble of conformations derived from homology models using different template structures other than A_{2A}R performs in comparison to an ensemble of crystal structures of A_{2A}R. We used 3 fully active (PDB IDs: 3SN6, 4MQS and 5C1M) and 1 active-intermediate state crystals (PDB ID: 4XES) of other class A GPCRs for generating four A_{2A}R active-intermediate state homology models. Similarly, 3 inactive states crystals (PDB IDs: 5UEN, 2YCX and 4DKL) were used for generating three homology models for A_{2A}R in the inactive state. Details of the GPCR used as structural templates and their percentage homology to A_{2A}R are given in Table S1. We compared these recovery rates to those derived from two ensembles (representing the respective and the other functional state) of crystal structures. The crystal structures used here for making an ensemble representing inactive state of A_{2A}R have PDB IDs 3EML, 3PWH, 3REY, 3RFM, 4EIY and 5NM4. The ones used for generating an ensemble for the active-intermediate state are 2YDV, 3QAK, 4UG2. The prediction of active-intermediate state specific mutations afforded by the fully active state crystal structure (PDB ID 5G53) is also shown as gray histogram in Figure 5A for comparison. Figures 5A and 5B show the mutation trial number at which each of the thermostable mutant is recovered for both active-intermediate and inactive states by different ensembles used in Fig. 4.

It is evident that most of the mutations specific to the active-intermediate state are retrieved earlier by the A_{2A}R active-intermediate crystal structures (Figure 5A). On the other hand, both inactive as well as active-intermediate crystal structures recover the thermostable mutations for the inactive state in early mutation trials (Figure 5B). Interestingly, the ensemble derived from homology models recovers the mutations stabilizing both the conformational states much earlier (within the 50th mutation trial) than the ensemble generated using the crystal structures of either state (open histograms for mutations with green background in Figure 5A & B). Furthermore, the ensemble of homology models recovers some of state specific mutations earlier than the crystal structures representing that state. For instance, the mutations L208A^{5.69}, L225A^{6.27} and L244A^{6.46} stabilizing the active-intermediate state and the mutations W029A^{1.55} and V239A^{6.41} stabilizing the inactive state are retrieved much earlier than the crystal structures of respective states (Figure 5A and 5B). Interestingly, 5 out of 13 active-intermediate state specific mutations (A081L^{3.29}, A203L^{5.64}, A231L^{6.33}, L244A^{6.46} and V282A^{7.47}) are recovered much earlier by the fully active crystal than the ensemble of active-intermediate crystal structures (Figure 5A). We speculate that this is because the crystal structures of the active-intermediate state of A_{2A}R show more active state characteristics than the inactive state characteristics. Thus, the consideration of multiple homology models of a particular receptor derived from various functional states of other receptors results in a better outcome for a blind prediction using LiticonDesign. This enrichment is possibly due to wider set of conformations derived from

multiple homology models based on structurally diverse templates. Such an ensemble could be more reflective of the active state of the GPCR than more constrained agonist bound crystal structures of A_{2A}R.

4. Discussion

Determining GPCR structure in a specific conformation remains challenging, as GPCRs display conformational heterogeneity in the fully active, active-intermediate and inactive states. The inter-residue distances between TM3 and TM6 and TM3 and TM7 are often used as indicators of these three distinct functional conformation states of GPCRs. Figure 6 shows a plot of the distances between residues 3.50 and 6.30 (TM3-TM6 distance) and between residues 3.50 and 7.53 (TM3-TM7 distance). Here we have used the Ballesteros-Weinstein GPCR residue numbering system.⁴⁵ Figure 6A shows the spread of the TM3-TM6 and TM3-TM7 distances for the crystal structures of A_{2A}R and the homology models of various conformational states used in this study. The template structures used for deriving each homology model is also indicated in this figure. Figure 6B shows the spread of these distances in other distinct class A GPCRs in the active-intermediate and inactive states. As can be seen in Figures 6A and 6B, the spread in TM3-TM6 distance among the inactive state crystal structures is overlapping with the TM3-TM6 distance in the active-intermediate states of A_{2A}R and other GPCRs. The spread in TM3-TM7 distance separates the inactive and active-intermediate structures better within A_{2A}R (Figure 6A).

LiticonDesign method distinguishes these structural changes between different conformational states of A_{2A}R as the results clearly show that the conformation of starting structural model used for predicting thermostable mutations using LiticonDesign method is critical to performance of the method in recovering the thermostable mutants in the top 50 predicted mutations. The notable conclusions from this study are: (1) an ensemble of conformations shows better predictions of thermostability than a single conformation even if it is a crystal structure of the same conformation. This is because crystal structures of inactive and active-intermediate states show overlapping TM3-TM6 distance that could be the result of crystallizing conditions. (2) The predictions are sensitive to the starting conformation of the model. The combination of fully active state or the active-intermediate state based homology models are far better starting points for predicting the thermostable mutations for the active-intermediate state than the inactive state crystal structures or the homology models derived from inactive state (Figures 2 and 4). Perhaps this is reflective of the ensemble of conformations that the receptor samples in the membrane lysates used for measuring thermostability. On the contrary, we find that the inactive state homology models are better in predicting the inactive state stabilizing mutations than the fully active state homology models, at least in the top 10 and top 20 trials. Our analysis on the true negatives indicates that using LiticonDesign the recovery of stabilizing mutations is achieved through a concurrent enrichment of destabilizing mutations known experimentally (Figure S1, Supporting Information).

Our analysis however raised two important questions. First, why do the fully active state crystal structures or the active state homology models predict the thermostable mutations for the active-intermediate state better than the inactive state crystal structures or the homology

models in spite of the TM3-TM6 distance of the inactive state being comparatively closer to the active-intermediate crystals than the fully active state models (Figure 1A)? Second, why do inactive state homology models and the active-intermediate crystals have similar predictability in recovering the inactive state specific mutations (after top 20, Figure 4C and 4D)? To answer these two questions, we focused on the TM3-TM6 and the TM3-TM7 distance distributions in Figure 6A. One interesting observation we made is that the TM3-TM6 distance overlaps across all the active-intermediate and the inactive state crystals, with no clear overlap for the TM3-TM7 distance. For example, the inactive state crystal structure of A_{2A}R, 3EML, has very similar TM3-TM6 distance as the active-intermediate state crystal structures (nearly 9.6 Å), though the TM3-TM7 distance of the former is distinctly different from the latter. Since both the TM3-TM6 and TM3-TM7 distances of the fully active state is closer to that of active-intermediate states compared to the inactive states, these structures are better predictive of the conformation sensitive thermostable mutation predictions. As seen in Figure 1B, the active-intermediate state stabilizing mutations are not only clustered in TM6 but also in TM7 which further strengthens our hypothesis that TM3-TM7 distance is also a hallmark in characterizing the active-intermediate state. Similarly, we also observed that the predictions made from inactive 3EML crystal (Figure 2) clearly fails to recover the active-intermediate state stabilizing mutations though it is efficient in predicting the inactive state stabilizing mutations. On the contrary, for the inactive state stabilizing mutation predictions, inactive state homology models and the active-intermediate crystals predict to a similar extent beyond top 20 trials because there are no thermostabilizing mutations on TM7 for A_{2A}R. However, for any other receptor, probably the inactive state homology models would have had a better predictability over the active-intermediate crystal structures if the inactive state stabilizing mutations were present in TM7 (e.g. turkey β₁AR receptor).¹⁵

It is worth mentioning here that the thermostable mutations common to both the states are also equally important as they can stabilize a structure without showing a bias towards a certain state. For example, the thermostabilized active-intermediate and the inactive state crystals of A_{2A}R (2YDV and 3PWH) had one residue common (A54L) which stabilizes both the conformational states. Similar trend was also observed for rat NTSR1 receptor where two mutations (A86L and F258A) were found common between the active-intermediate and the inactive state crystals (4GRV and 4BUO). Our study suggests that in order to recover such non-specific mutations the starting structural model can be representative of either conformational state. Furthermore, an ensemble of homology models representing a certain functional state of the receptor derived using templates of that state is more effective in retrieving such non-specific mutations than a single crystal structure. This is presumably due wider representation of the conformational space represented when using an ensemble of models compared to using a single structure.

In conclusion, our LiticonDesign method is very useful for predictions of thermostable mutations specific to a particular conformation state of a class A GPCR. The predictions are comparatively more sensitive (especially within top 10 to top 20 trials) to the target conformation state under consideration than the sequence similarity of the template structure used to build the structural model of the target GPCR. Using an ensemble of conformations around the targeted conformational state enriches the predictions compared to using a single structure of the same conformational state.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was funded by the National Institutes of Health NIH-RO1GM097261 to N.V.

References

1. Lundstrom K An Overview on GPCRs and Drug Discovery: Structure-based Drug Design and Structural Biology on GPCRs. *Methods Mol. Biol* 2009, 552, 51–66. [PubMed: 19513641]
2. Jazayeri A; Dias JM; Marshall FH From G Protein-coupled Receptor Structure Resolution to Rational Drug Design. *J. Biol. Chem* 2015, 290, 19489–19495. [PubMed: 26100628]
3. Roth BL; Irwin JJ; Shoichet BK Discovery of new GPCR Ligands to Illuminate New Biology. *Nat. Chem. Biol* 2017, 13, 1143–1151. [PubMed: 29045379]
4. Latorraca NR; Venkatakrishnan AJ; Dror RO GPCR Dynamics: Structures in Motion. *Chem. Rev* 2017, 117, 139–155. [PubMed: 27622975]
5. Dore AS; Robertson N; Errey JC; Ng I; Hollenstein K; Tehan B; Hurrell E; Bennett K; Congreve M; Magnani F; Tate CG; Weir M; Marshall FH Structure of the Adenosine A(2A) Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure* 2011, 19, 1283–1293. [PubMed: 21885291]
6. Rosenbaum DM; Cherezov V; Hanson MA; Rasmussen SG; Thian FS; Kobilka TS; Choi HJ; Yao XJ; Weis WI; Stevens RC; Kobilka BK GPCR Engineering yields High-resolution Structural Insights into Beta2-Adrenergic Receptor Function. *Science* 2007, 318, 1266–1273. [PubMed: 17962519]
7. Steyaert J; Kobilka BK Nanobody Stabilization of G protein-coupled Receptor Conformational States. *Curr. Opin. Struct. Biol* 2011, 21, 567–572. [PubMed: 21782416]
8. Tate CG A Crystal Clear Solution for Determining G-protein-coupled Receptor Structures. *Trends Biochem. Sci* 2012, 37, 343–352. [PubMed: 22784935]
9. Vaidehi N; Grisshammer R; Tate CG How Can Mutations Thermostabilize G-Protein-Coupled Receptors? *Trends Pharmacol. Sci* 2016, 37, 37–46. [PubMed: 26547284]
10. Sarkar CA; Dodevski I; Kenig M; Dudli S; Mohr A; Hermans E; Pluckthun A Directed Evolution of a G protein-coupled receptor for Expression, Stability, and Binding Selectivity. *Proc. Natl. Acad. Sci. U. S. A* 2008, 105, 14808–14813. [PubMed: 18812512]
11. Dodevski I; Pluckthun A Evolution of Three Human GPCRs for Higher Expression and Stability. *J. Mol. Biol* 2011, 408, 599–615. [PubMed: 21376730]
12. Egloff P; Hillenbrand M; Klenk C; Batyuk A; Heine P; Balada S; Schlinkmann KM; Scott DJ; Schutz M; Pluckthun A Structure of Signaling-Competent Neurotensin Receptor 1 Obtained by Directed Evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* 2014, 111, E655–E662. [PubMed: 24453215]
13. Shibata Y; White JF; Serrano-Vega MJ; Magnani F; Aloia AL; Grisshammer R; Tate CG Thermostabilization of the Neurotensin receptor NTS1. *J. Mol. Biol* 2009, 390, 262–277. [PubMed: 19422831]
14. Magnani F; Shibata Y; Serrano-Vega MJ; Tate CG Co-evolving Stability and Conformational Homogeneity of the Human Adenosine A2a Receptor. *Proc. Natl. Acad. Sci. U. S. A* 2008, 105, 10744–10749. [PubMed: 18664584]
15. Serrano-Vega MJ; Magnani F; Shibata Y; Tate CG Conformational Thermostabilization of the Beta1-Adrenergic Receptor in a Detergent-Resistant Form. *Proc. Natl. Acad. Sci. U. S. A* 2008, 105, 877–882. [PubMed: 18192400]
16. Lebon G; Bennett K; Jazayeri A; Tate CG Thermostabilisation of an Agonist-bound Conformation of the Human Adenosine A(2A) Receptor. *J. Mol. Biol* 2011, 409, 298–310. [PubMed: 21501622]

17. Chen KY; Zhou F; Fryszczyn BG; Barth P Naturally Evolved G protein-coupled Receptors Adopt Metastable Conformations. *Proc. Natl. Acad. Sci. U. S. A* 2012, 109, 13284–13289. [PubMed: 22847407]
18. Popov P; Kozlovskii I; Katritch V Computational Design for Thermostabilization of GPCRs. *Curr. Opin. Struct. Biol* 2019, 55, 25–33. [PubMed: 30909106]
19. Popov P; Peng Y; Shen L; Stevens RC; Cherezov V; Liu ZJ; Katritch V Computational Design of Thermostabilizing Point Mutations for G Protein-Coupled Receptors. *eLife* 2018, 7 (1 – 22).
20. Yasuda S; Kajiwara Y; Takamuku Y; Suzuki N; Murata T; Kinoshita M Identification of Thermostabilizing Mutations for Membrane Proteins: Rapid Method Based on Statistical Thermodynamics. *J. Phys. Chem. B* 2016, 120, 3833–3843. [PubMed: 27056055]
21. Lebon G; Warne T; Edwards PC; Bennett K; Langmead CJ; Leslie AG Tate CG, Agonist-bound Adenosine A2A Receptor Structures Reveal Common Features of GPCR Activation. *Nature* 2011, 474, 521–525. [PubMed: 21593763]
22. Balaraman GS; Bhattacharya S; Vaidehi N Structural Insights into Conformational Stability of Wild-type and Mutant Beta1-adrenergic Receptor. *Biophys. J* 2010, 99, 568–577. [PubMed: 20643076]
23. Bhattacharya S; Lee S; Grisshammer R; Tate CG; Vaidehi N Rapid Computational Prediction of Thermostabilizing Mutations for G Protein-Coupled Receptors. *J. Chem. Theory Comput* 2014, 10, 5149–5160. [PubMed: 25400524]
24. Bhattacharya S; Vaidehi N LITiCon: A Discrete Conformational Sampling Computational Method for Mapping Various Functionally Selective Conformational States of Transmembrane Helical Proteins. *Methods Mol. Biol* 2012, 914, 167–178. [PubMed: 22976028]
25. Bhattacharya S; Hall SE; Li H; Vaidehi N Ligand-Stabilized Conformational States of Human Beta(2) Adrenergic Receptor: Insight into G-protein-coupled Receptor Activation. *Biophys. J* 2008, 94, 2027–2042. [PubMed: 18065472]
26. Krivov GG; Shapovalov MV; Dunbrack RL Jr. Improved Prediction of Protein Side-Chain Conformations with SCWRL4. *Proteins* 2009, 77, 778–795. [PubMed: 19603484]
27. Brooks BR; Brooks CL 3rd; Mackerell AD Jr.; Nilsson L; Petrella RJ; Roux B; Won Y; Archontis G; Bartels C; Boresch S; Caflisch A; Caves L; Cui Q; Dinner AR; Feig M; Fischer S; Gao J; Hodoseck M; Im W; Kuczera K; Lazaridis T; Ma J; Ovchinnikov V; Paci E; Pastor RW; Post CB; Pu JZ; Schaefer M; Tidor M; Venable RM; Woodcock HL; Wu X; Yang W; York DM; Karplus M CHARMM: the Biomolecular Simulation Program. *J. Chem. Theory Comput* 2009, 30, 1545–1614.
28. Krumm BE; Lee S; Bhattacharya S; Botos I; White CF; Du H; Vaidehi N; Grisshammer R Structure and Dynamics of a Constitutively Active Neurotensin Receptor. *Sci. Rep* 2016, 6, 38564. [PubMed: 27924846]
29. Lee S; Bhattacharya S; Tate CG; Grisshammer R; Vaidehi N Structural Dynamics and Thermostabilization of Neurotensin Receptor 1. *J. Phys. Chem. B* 2015, 119, 4917–4928. [PubMed: 25807267]
30. Lee S; Bhattacharya S; Grisshammer R; Tate C; Vaidehi N Dynamic Behavior of the Active and Inactive states of the Adenosine A(2A) Receptor. *J. Phys. Chem. B* 2014, 118, 3355–3365. [PubMed: 24579769]
31. Jaakola VP; Griffith MT; Hanson MA; Cherezov V; Chien EYT; Lane JR; IJerman AP; Stevens RC The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science* 2008, 322, 1211–1217. [PubMed: 18832607]
32. Liu W; Chun E; Thompson AA; Chubukov P; Xu F; Katritch V; Han GW; Roth CB; Heitman LH; IJerman AP; Cherezov V; Stevens RC Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. *Science* 2012, 337, 232–236. [PubMed: 22798613]
33. Weinert T; Olieric N; Cheng R; Brunle S; James D; Ozerov D; Gashi D; Vera L; Marsh M; Jaeger K.; Dworkowski F; Panepucci E; Basu S; Skopintsev P; Dore AS; Geng T; Cooke RM; Liang M; Protá A; Paneels V; Nogly P; Ermler U; Schertler G; Hennig M; Steinmetz O; Wang M; Standfuss J Serial Millisecond Crystallography for Routine Room-Temperature Structure Determination at Synchrotrons. *Nature Commun.* 2017, 8 (1 – 11).

34. Xu F; Wu HX; Katritch V; Han GW; Jacobson KA; Gao ZG; Cherezov V; Stevens RC Structure of an Agonist-Bound Human A(2A) Adenosine Receptor. *Science* 2011, 332, 322–327. [PubMed: 21393508]
35. Lebon G; Edwards PC; Leslie AGW; Tate CG Molecular Determinants of CGS21680 Binding to the Human Adenosine A(2A) Receptor. *Mol. Pharmacol* 2015, 87, 907–915. [PubMed: 25762024]
36. Carpenter B; Nehme R; Warne T; Leslie AGW; Tate CG Structure of the Adenosine A(2A) Receptor Bound to an Engineered G protein. *Nature* 2016, 538, 7626.
37. Rasmussen SGF; DeVree BT; Zou YZ; Kruse AC; Chung KY; Kobilka TS; Thian FS; Chae PS; Pardon E; Calinski D; Mathiesen JM; Shah ST; Lyons JA; Caffrey M; Gellman SH; Skiniotis G; Weis WI; Sunahara RK; Kobilka BK Crystal Structure of the Beta(2) Adrenergic Receptor-Gs Protein Complex. *Nature* 2011, 477, 549–555. [PubMed: 21772288]
38. Kruse AC; Ring AM; Manglik A; Hu JX; Hu K; Eitel K; Hubner H; Pardon E; Valant C; Sexton PM; Christopoulos A; Felder CC; Gmeiner P; Steyaert J; Weis WI; Garcia KC; Wess J; Kobilka BK Activation and Allosteric Modulation of a Muscarinic Acetylcholine Receptor. *Nature* 2013, 504, 101–106. [PubMed: 24256733]
39. Huang WJ; Manglik A; Venkatakrishnan AJ; Laeremans T; Feinberg EN; Sanborn AL; Kato HE; Livingston KE; Thorsen TS; Kling RC Granier S; Gmeiner P; Husbands SM; Traynor JR; Weis WI; Steyaert J; Dror RO; Kobilka BK Structural Insights into Mu-opioid Receptor Activation. *Nature* 2015, 524, 315–321. [PubMed: 26245379]
40. Krumm BE; White JF; Shah P; Grishammer R Structural Prerequisites for G-protein Activation by the Neurotensin Receptor. *Nature Commun.* 2015, 6, 7895. [PubMed: 26205105]
41. Glukhova A; Thal DM; Nguyen AT; Vecchio EA; Jorg M; Scammells PJ; May LT; Sexton PM; Christopoulos A Structure of the Adenosine A(1) Receptor Reveals the Basis for Subtype Selectivity. *Cell* 2017, 168, 867–877. [PubMed: 28235198]
42. Moukhametzianov R; Warne T; Edwards PC; Serrano-Vega MJ; Leslie AGW; Tate CG; Schertler GFX Two Distinct Conformations of Helix 6 Observed in Antagonist Bound Structures of a Beta(1)-Adrenergic Receptor. *Proc. Natl. Acad. Sci. U. S. A* 2011, 108, 8228–8232. [PubMed: 21540331]
43. Manglik A; Kruse AC; Kobilka TS; Thian FS; Mathiesen JM; Sunahara RK; Pardo L; Weis WI; Kobilka BK; Granier S Crystal Structure of the Mu-opioid Receptor Bound to a Morphinan Antagonist. *Nature* 2012, 485, 321–326. [PubMed: 22437502]
44. Schrödinger L The PyMOL Molecular Graphics System. 2010, (1.5.0.4).
45. Sali A; Blundell TL Comparative Protein Modeling by Satisfaction of Spatial Restraints. *J. Mol. Biol* 1993, 234, 779–815. [PubMed: 8254673]
46. Fiser A; Do RKG; Sali A Modeling of Loops in Protein Structures. *Prot. Sci* 2000, 9, 1753–1773.
47. Isberg V; Mordalski S; Munk C; Rataj K; Harpoe K; Hauser AS; Vroiling B; Bojarski AJ; Vriend G; Gloriam DE GPCRdb: An Information System for G protein-coupled Receptors. *Nucleic Acids Res.* 2016, 44, D356–D364. [PubMed: 26582914]
48. 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* 2008, 4, 435–447. [PubMed: 26620784]

Figure 1:

Analysis of the locations of the thermostable mutations from experimental measurements of A_{2A}R, and the mutations stabilizing each of the conformational states. **(A)** Intracellular view of the crystal structures of the inactive state (magenta, PDB 3PWH), active-intermediate (shown in yellow, PDB 2YDV), and fully active state (shown in grey, PDB 5G53). This shows the movement of the intracellular part of TM6 away from TM3, as indicated by Ca (cyan sphere) distances between R102^{3.50} and E228^{6.30} (dashed line). **(B)** The snakeplot of A_{2A}R receptor sequence highlighting the mutation sites specific to the stabilization of the active-intermediate state conformation (yellow), and the inactive state conformation

(magenta), respectively. The mutation sites that are thermostabilizing both of the inactive and active-intermediate states are shown in green. The thermostabilizing mutations are reported previously by Magnani et al.¹⁴ The snake plot was generated using the GPCRdb web server.⁴⁷ (C) Comparison of TM Backbone RMSD per residue of A_{2A}R active-intermediate state (PDB 2YDV) with respect to inactive (PDB 3PWH) state conformation. RMSD was calculated using *gmx rmsfutility* with *-od* option as implemented in the GROMACS package.⁴⁸ The residues in the loops are omitted to focus the differences in the TM region residues.

Figure 2:

LiticonDesign predictions of thermostable mutations for the active-intermediate and inactive state of human A_{2A}R receptor starting from the crystal structures of different conformational states of A_{2A}R. The A_{2A}R crystal conformations used for predictions are: fully active state (color code: black; PDB 5G53), active-intermediate state (color code: yellow; PDBs 2YDV, 3QAK, and 4UG2), and inactive state conformations (color code: magenta; PDBs 3EML, 3PWH, 3REY, 3RFM, 4EIY, and 5NM4). **(A)** and **(C)** Percentage recovery of total number of thermostable mutations (combination of both conformation specific and non-specific) recovered for active-intermediate state and the inactive state, respectively, for different number of mutational trials (top 10, top 20, top 30, top 40, and top 50 based on stability score ranking). **(B)** Percentage recovery of thermostable mutations specific to stabilizing the active-intermediate state and not the inactive state **(D)** Percentage recovery of conformation specific thermostable mutations stabilizing the inactive state. The solid blue line in all the figures corresponds to the random prediction.

Figure 3:

Comparison of the backbone RMSD per residue amongst the active-intermediate and the inactive crystal structures of A_{2A}R. **(A)** The RMSD of active-intermediate state crystal structures of A_{2A}R: PDB ID 3QAK, and 4UG2 calculated with respect to active-intermediate 2YDV crystal structure. **(B)** The RMSD of the inactive crystal structures of A_{2A}R: PDB ID 3EML, 3REY, 3RFM, 4EIY, and 5NM4 calculated with respect to the inactive state 3PWH crystal structure. The backbone RMSD was calculated for the TM helices of A_{2A}R using *gmx rmsf* utility with *-od* option as implemented in GROMACS.⁴⁸

Figure 4:

Percentage recovery of thermostable mutations using LiticonDesign predictions for the active-intermediate and inactive states of A_{2A}R. **(A)** Percentage recovery of total number of thermostable mutations for active-intermediate state for different number of mutational trials. **(B)** Percentage recovery of thermostable mutations specific to the active-intermediate state of A_{2A}R. **(C)** Percentage recovery for predicting the total number of thermostable mutations for the inactive state of A_{2A}R. **(D)** Percentage recovery of conformation specific thermostable mutations for the inactive state of A_{2A}R. The solid blue line denotes random prediction. The solid bars are the predictions starting from an ensemble of crystal structures. The open bars are the predictions made from an ensemble of homology models. Three fully active, one active-intermediate and three inactive state crystal structures of class A GPCRs were used to generate the homology models of the active-intermediate and the inactive state homology models of A_{2A}R respectively. The percentage sequence identity and the individual percentage recovery using each of these templates are provided in Table S1 of the Supporting Information. The solid bar in yellow represents predictions from active-intermediate crystals of A_{2A}R. The solid bar in magenta represents predictions from inactive state crystals of A_{2A}R. The hollow bar in black represents predictions from active/active-intermediate state homology models. The hollow bar in magenta represents predictions from inactive state homology models.

Figure 5:

Mutation trial at which each of the positive thermostable mutant was recovered using crystal structures or homology models. (A) thermostable mutants for the active-intermediate state shown in the Y-axis and the mutation trial number ordered by calculated stability scores shown in the X-axis (B) the same data as (A) for the inactive state of A_{2A}R. Thermostable mutants exclusive for active-intermediate and inactive states are colored in yellow and magenta respectively on the Y-axes of the figures. Thermostable mutants that are common to both the conformational states are colored in green. The amino acid residue numbers and their corresponding Ballesteros-Weinstein GPCR residue numbering are shown for the thermostable mutations. The X-axes shows how early these mutations are recovered by each ensemble using LiticonDesign. The 50th mutation trial has been shown as dotted vertical lines.

Figure 6:

Inter-residue distances between the C α atoms of the residues with Ballesteros-Weinstein numbering 3.50 and 6.30 (TM3-TM6) and between 3.50 and 7.53 (TM3-TM7) measured in several class A GPCR crystal structures.⁴⁵ **(A)** The spread of the distances for all of the A_2 AR crystal structures and the homology models used in the current study. **(B)** The spread of the distances between the active-intermediate and the inactive state crystals of class A GPCRs other than the A_2 AR extracted from GPCRdb database.⁴⁷ We used several crystals of active-intermediate state (PDB ID: 5TZR, 5TVN, 5NDZ, 4XNW), and the inactive state (PDB ID: 5V54, 5O9H, 5WS3, 5WIU, 5XPR, 5X7D, 5UIW, 5T1A). The crystal structures are shown in solid circle and the homology models are shown in dashed circles. The fully

active state conformation is shown in black, the active-intermediate conformations are shown in yellow, and the inactive state conformations are highlighted in magenta.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript