Title

7×7 RMSD matrix: A new method for quantitative comparison of the transmembrane domain structures in the G-protein coupled receptors

Journal

Journal of Structural Biology, 199(2)

ISSN

1047-8477

Authors

Wang, Ting
Wang, Yang
Tang, Leihan
et al.

Publication Date

2017-08-01

DOI

10.1016/j.jsb.2017.02.005

Peer reviewed

HHS Public Access

Author manuscript

J Struct Biol. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:

J Struct Biol. 2017 August ; 199(2): 87–101. doi:10.1016/j.jsb.2017.02.005.

7x7 RMSD Matrix: A New Method for Quantitative Comparison of the Transmembrane Domain Structures in the G-protein Coupled Receptors

Ting Wang[1,2,*], **Yang Wang**[1], **Leihan Tang**[1], **Yong Duan**[2,*], and **Haiguang Liu**[1,*]

[1]Complex Systems Division, Beijing Computational Science Research Center, 10 W. Dongbeiwang Rd, Haidian District, Beijing 100193, China

[2]Genome Center, 451 East Health Science Drive, University of California, Davis, California 95616, United States

Abstract

The G-protein coupled receptors (GPCRs) share a conserved heptahelical fold in the transmembrane (TM) region, but the exact arrangements of the seven TM helices vary with receptors and their activation states. The differences or the changes have been observed in the experimentally solved structures, but have not been systematically and quantitatively investigated due to lack of suitable methods. In this work, we describe a novel method, called 7x7 RMSD matrix that is proposed specifically for comparing the characteristic 7TM bundle structures of GPCRs. Compared to the commonly used overall TM bundle RMSD as a single parameter, a 7x7 RMSD matrix contains 49 parameters, which reveal changes of the relative orientations of the seven TMs. We demonstrate the novelty and advantages of this method by tackling two problems that are challenging for the existing methods. With this method, we are able to identify and quantify the helix movements in the activated receptor structures and reveal structural conservation and divergence as well as the structural relationships of different GPCRs in terms of the relative orientations of the seven TMs.

Keywords

Membrane proteins; GPCR; transmembrane; helix movement; RMSD; classification

Introduction

The G-protein coupled receptors (GPCRs) comprise the largest membrane protein family in the human genome with over 800 members (Fredriksson et al., 2003). As signaling

*Corresponding author: Wang T., Duan Y. and Liu H. twang@ucdavis.edu, duan@ucdavis.edu, hgliu@csrc.ac.cn.

Author contributions

Conceived and designed research: TW and HL. Performed analysis: TW, YW and HL. Wrote the manuscript: TW, LT, HL and YD.

molecules that transduce extracellular signals into cells, the GPCR superfamily controls a wide variety of biological processes including metabolism, development and aging, and has served as targets of about 40% drugs on the market. Despite the large diversity of functions and low sequence identities (< 25%) across families (Lagerstrom and Schioth, 2008), GPCRs are believed to share a common heptahelical fold in the transmembrane (TM) region (Kobilka, 2007; Lagerstrom and Schioth, 2008). This hypothesis is supported by the fact that the GPCR structures solved to date all demonstrate a 7TM bundle as the core structure. Nevertheless, structural differences in the 7TM cores have been observed in the experimentally solved structures of different GPCRs, which are often qualitatively described as helix shifts and tilts. In addition, rigid-body helix movements have been observed in the agonist-bound activated structures, which have been considered as the major conformational changes during receptor activation. The typical arrangement of the seven TMs in a GPCR structure is shown in Figure S1.

The method conventionally used for structural comparison is to superpose the structures and calculate the root-mean-square-deviation (RMSD). For GPCRs, calculation of RMSDs typically generates a value in the range of 2 Å to 3 Å for the 7TM bundles. Even for the recently solved non-class A receptors, the overall 7TM RMSD is less than 4 Å (Hollenstein et al., 2013). Such a small single parameter is unable to interpret the structural differences of diverse families and functions. In the case of receptor activation, the RMSDs of the 7TM bundles between fully activated and the corresponding inactive structures are also as small as 2 Å (Huang et al., 2015; Rasmussen et al., 2011; Scheerer et al., 2008). Other measures include individual RMSD for each TM (Hollenstein et al., 2013) and the distances between the tips of TMs (Siu et al., 2013). There are also a few studies that specifically analyzed the 7TM structures of the GPCRs available at the time. Dalton and coworkers (Dalton et al., 2015) investigated inter-helical angles and distances in 25 GPCR structures and computed their changes in the inactive and active structures for five GPCRs. Tehan and coworkers (Tehan et al., 2014) described TM movements in the active structures of rhodopsin and β2-adrenoceptor (β2AR) by superposing on the inactive structures. Kinoshita and Okada (Kinoshita and Okada, 2015) investigated structural conservation of the 7TM bundles among 19 receptors by computing intramolecular Cα-Cα distances. However, those measures or analysis methods are not able to identify helix movements that change the relative orientations of the seven TMs. Moreover, GPCR structures are being solved in an unprecedented speed nowadays, with more than 5 structures appearing in a year. There needs a simple and efficient method for GPCR structural comparison.

In this work, we present a novel method, called 7x7 RMSD matrix, which is proposed specifically for comparing the 7TM bundle structures of GPCRs. We demonstrate the application of 7x7 RMSD matrix as a metric of helix movements and a 49-parameter structural similarity index. We applied this method to the analysis of the GPCR structures solved to date and addressed two problems that have been difficult for the existing methods. The first problem is how to identify and measure helix movements in the activated GPCR structures, which is important for understanding receptor activation mechanisms. The second problem is structural comparison and clustering of different GPCR receptors by the relative orientations of seven TMs. We at first present a systematic view of the structural

conservation and divergence of the 7TM cores and then reveal the structural relationships of different GPCRs by their 7TM core structures.

Material and Methods

Definition of the 7x7 RMSD Matrix

As the 7TM bundle is confined in the membrane environment and tethered by inter-helical loops, large translation of a whole helix is unlikely, which is evident in the small RMSDs of superposing different GPCR structures; instead, rotations of TMs and local conformational distortions are more likely to happen. Such movements are not reflected in the overall RMSD of the 7TM bundle. We propose a 7x7 RMSD matrix that is calculated by successive fitting corresponding TMs of two structures and recording the RMSD for each of all seven TM pairs. A 7x7 RMSD matrix contains seven rows and seven columns, thus 7 x 7 = 49 RMSDs. Specifically, the seven elements in the *i*-th row are the RMSDs calculated for TM1 through TM7 when fitting only the *i*-th TM. If the *i*-th TM transforms relative to the reference structure, especially by rotation, large RMSDs may appear for other TMs. A schematic view explaining how rotating one TM results in large positional displacements of other TMs, therefore large RMSDs, is shown in Figure S2. Compared to a single parameter of the overall RMSD, the 7x7 RMSD matrix decomposes and magnifies the structural differences between the reference and target structures into 49 parameters. Two types of important information can be obtained from the matrix: 1) the seven diagonal elements serve as an indicator of conformational changes or conservations within the helices themselves; 2) the off-diagonal elements at each row reveal whether the corresponding TM moves from the reference structure relatively to the other TMs, thus changing its orientation in the 7TM bundle.

Structures of Seven GPCR Receptors in Inactive and Active States

To demonstrate the ability of the 7x7 RMSD matrix quantifying helix movements in the active receptor structures, we analyze a set of seven GPCR receptors whose crystal structures have been solved in both inactive and active states. These receptors are the bovine light receptor rhodopsin, the human β2-adrenoceptor (β2AR), the human M2 muscarinic receptor (M2), the murine μ–opioid receptor (μOR) and, the turkey β1-adrenoceptor (β1AR), the human A2A adenosine receptor (A2A). Rhodopsin has been crystallized to capture several different activated states with and without G-protein binding, and thus nine intermediate and active structures are used (PDB ids: 1U19(Okada et al., 2004), 2G87(Nakamichi and Okada, 2006a), 3PQR(Choe et al., 2011), 3PXO(Choe et al., 2011), 3DQB(Scheerer et al., 2008), 3CAP(Park et al., 2008), 4PXF(Szczepek et al., 2014), 4ZWJ(Kang et al., 2015) and 4X1H(Blankenship et al., 2015)). Other non-rhodopsin receptors are co-crystallized in complexes with different agonists and with or without G-protein binding, and thus several active and active-like structures are used (PDB ids of β2AR: 2RH1(Cherezov et al., 2007), 3SN6(Rasmussen et al., 2011), 3PDS(Rosenbaum et al., 2011) and 4LDE(Ring et al., 2013); PDB ids of M2: 3UON(Haga et al., 2012), 4MQS(Kruse et al., 2013) and 4MQT(Kruse et al., 2013); PDB ids of μOR: 4DKL(Manglik et al., 2012) and 5C1M(Huang et al., 2015); PDB ids of β1AR: 2VT4(Warne et al., 2008a), 2Y00(Warne et al., 2011) and 2Y02(Warne et al., 2011); PDB ids of A2A: 3EML(Jaakola et

al., 2008), 3QAK(Xu et al., 2011), 2YDO(Lebon et al., 2011), and 2YDV). In addition, we include the rat neurotensin type 1 receptor (NTS1), for which, two agonist-bound structures have been solved (PDB ids: 4BUO(Egloff et al., 2014) and 4GRV(White et al., 2012)) and the thermostabilized mutant structure (PDB id 4BUO(Egloff et al., 2014)) is treated as inactive. The details of this set of structures are listed in Table 1.

Structures of 33 Unique GPCR Receptors

To demonstrate the ability of the 7x7 RMSD matrix quantifying structural differences of the 7TM cores among different GPCR receptors, we analyze the structures of the GPCR receptors whose structures are available to date. The GPCRdb database (Isberg et al., 2014; Isberg et al., 2016) contains a complete list of the receptors whose crystal structures have been solved. We study all of them excluding the viral chemokine receptor U28, in total 33 receptors. In this analysis, we focus on the inactive states of the receptors and use one inactive structure for each receptor. In case that several crystal structures are available, we choose the one with the highest resolution. The A2A receptor is an exception. The structure with the highest resolution (PDBid: 4EIY) is not chosen because TM5 and TM6 are significantly straightened and extended in that structure, which is the result of formation of a chimeric protein. The receptors and the corresponding crystal structures are listed in Table 2 and Table S1. There are 28 receptors in class A, 2 in class B, 2 in class C, and 1 in class F. These receptors represent 21 families and subfamilies, including rhodopsin (RHO: 1U19 (Okada et al., 2004)), adrenoceptor (β1AR: 2VT4 (Warne et al., 2008b) and β2AR: 2RH1 (Cherezov et al., 2007)), muscarinic (M1: 5CXV (Thal et al., 2016), M2: 3UON (Haga et al., 2012), M3: 4U15 (Thorsen et al., 2014) and M4: 5DSG (Thal et al., 2016)), dopamine (D3: 3PBL (Chien et al., 2010)), histaminine (H1: 3RZE (Shimamura et al., 2011)), 5-hydroxytryptamine(5-HT) (5-HT1B: 4IAR (Wang et al., 2013b) and 5-HT2B: 4IB4 (Wacker et al., 2013)), adenosine (A2A: 3EML (Jaakola et al., 2008)), purinoreceptor (P2Y1: 4XNV (Zhang et al., 2015a) and P2Y12: 4NTJ (Zhang et al., 2014)), neurotensin (NTS1: 4BUO (Egloff et al., 2014)), angiotensin (AT1: 4YAY (Zhang et al., 2015b)), opioid (NOP: 4EA3 (Thompson et al., 2012), δOR: 4N6H (Fenalti et al., 2014), κOR: 4DJH (Wu et al., 2012) and μOR: 4DKL (Manglik et al., 2012)), chemokine (CXCR4: 3ODU (Wu et al., 2010) and CCR5: 4MBS (Tan et al., 2013)), orexin (OX1: 4ZJ8 (Yin et al., 2016) and OX2: 4S0V (Yin et al., 2015)), free fatty acid (FFR1: 4PHU (Srivastava et al., 2014)), proteinase-activated (PAR1: 3VW7 (Zhang et al., 2012)), lysophosphatidic acid (LPA1: 4Z35 (Chrencik et al., 2015)), sphingosine 1-phosphate (S1P1: 3V2Y (Hanson et al., 2012)), cortincotropin-releasing factor (CRF1: 4K5Y (Hollenstein et al., 2013)), glugacon (glucagon: 4L6R (Siu et al., 2013)), metabotropic glutamate (mGlu1: 4OR2 (Wu et al., 2014) and mGlu5: 5CGD (Christopher et al., 2015)) and smoothened (SMO: 4JKV (Wang et al., 2013a)). Several families have two or more subtypes solved. The opioid receptors and the muscarinic receptors are the most comprehensively covered subfamilies, each having four subtype structures.

Our selection of the structures is focused on the inactive states of the receptors. In the majority of the structures, the receptors are co-crystallized in complexes with antagonists, inverse agonists or negative allosteric modulators. A few receptors are bound with agonists, which are the neruotensin receptor (PDB id 4BUO), the free fatty acid receptor (PDB id

4PHU) and two 5-HT receptors (PDB ids 4IAR and 4IB4). The former two are believed to be in the inactive states due to a number of thermostabilization mutations (Egloff et al., 2014; Srivastava et al., 2014). The two 5-HT receptors are thought to exhibit some active features (Wacker et al., 2013; Wang et al., 2013b).

Structural Preparation

For each crystal structure, only the GPCR protein part is used for analysis. Other parts are discarded such as ligands, water molecules, fusion proteins, intracellular binding partners and solvent molecules. In case that multiple copies of GPCR molecules are present, chain A is used except the β1AR structure (PDB id: 2VT4), in which TM1 of chain A is sharply bended and thus chain B is used instead.

The Ballesteros-Weinstein (BW) numbers (Ballesteros and Weinstein, 1995) of class A GPCRs are assigned using the GPCRdb (Isberg et al., 2014; Isberg et al., 2016) web service (http://gpcrdb.org). For the five structures of classes B, C and F, we follow the references (Christopher et al., 2015; Hollenstein et al., 2013; Siu et al., 2013; Wang et al., 2013a; Wu et al., 2014) that originally reported the x-ray structures, in which the authors obtained the BW numbers based on structural superposition with class A GPCRs. To determine the start and stop residues of each TM in the crystal structures, we perform visual inspection on each structure with the VMD software (Humphrey et al., 1996). Table S1 lists the BW .50 residues, the start and stop residues of each TM of the crystal structures of 33 receptors.

The class A GPCRs share the conserved BW .50 residues, N1.50, D2.50, R3.50, W4.50, P5.50, P6.50 and P7.50, with a few exceptions being L167 at position 1.50 in 4BUO, N201 at position 5.50 in 4NTJ, L213 at position 5.50 in 3V2Y, V319 at position 6.50 in 4S0V. The non-class A GPCRs have completely different BW X.50 residues. The lengths of TMs vary with receptors and the maximal common TM residues that are solved and present in all 33 crystal structures are 1.36 – 1.57, 2.40–2.60, 3.23–3.52, 4.45–4.58, 5.40–5.61, 6.34–6.51 and 7.35–7.53 as defined by the BW numbers. These 146 residues thus comprise the core of the TM region with the BW numbers common to all structures. The sequence alignment of the TM1–TM7 residues is shown in Figure S5.

For the analysis of the inactive-active structural pairs, the common core TM residues of each structural pair are defined as the TM residues that are present in both structures (Table S2).

7x7 RMSD Matrix Calculations

Structural superposition and RMSD calculations are performed with the VMD software (Humphrey et al., 1996). The RMSD values are calculated for the backbone atoms of each TM, i.e. the CA, C, O, N atoms. A 7 x 7 RMSD matrix is obtained by successively fitting one of the seven TMs and recording RMSDs for all seven TMs. For the comparison between inactive and active structures, there are 21 active or active-like structures and thus 21 matrices are generated for the 21 inactive-active structural pairs. For all-against-all structural comparison of the 33 receptors, each of the 33 receptor structures serves as a reference structure to compute the 7x7 RMSD matrices of all 33 receptors, resulted in 33 x 33 = 1089 7x7 RMSD matrices. As the RMSD of structure A against structure B is identical to that of

structure B against structure A, there are actually 33 x (33 + 1) / 2 = 561 different 7x7 RMSD matrices.

Rotational Angle Calculations

Rotations of the TMs are defined as around the principal axes of the reference structure. It is worth noting that the principal axes of the reference structure are not the helical axis of any TM. To compute the rotational angles, we first need to align the reference and target structures into a homologous coordinate system. For inactive-active structural comparison, each inactive structure serves as the reference structure and the active structures as the target structures. For the comparison of 33 receptors against the rhodopsin structure, the rhodopsin structure 1U19 serves as the reference structure and the remaining 32 receptors as the target structures.

The alignment is done in two steps: 1) the coordinates of the reference structure are reset by aligning the principal axes of the molecule in the X, Y, and Z-directions. As a result, the Z-axis is proximately perpendicular to the membrane, pointing from the intracellular to extracellular side; 2) the coordinates of the target structure are reset by being superposed onto the re-oriented inactive structure using all TMs. Figure S3 shows the principal axes of rhodopsin structure 1U19 and β2AR structure 2RH1, as examples.

After the reference and target structures are aligned into a homologous coordinate system, structural superposition is performed for each of the reference-target structural pair using one of the seven TMs at a time. The VMD software (Humphrey et al., 1996) automatically generated a 4 x 4 transformation matrix for each structural superposition. A python script has been written to convert the transformation matrix into rotational angles of the corresponding TM. Figure S4 shows two sample transformation matrices: superposing rhodopsin inactive and active pair 1U19-3CAP using TM6 and superposing 1U19 and 2RH1 using TM4.

Clustering Analysis

We perform all-against-all structural comparison for the 7TM core structures in the 33 receptors. Four different structural similarity indices are used, which are the 7x7 RMSD matrix we propose here, overall 7TM bundle RMSD, Dali Z-score (Holm and Laakso, 2016) and TM-align score (Zhang and Skolnick, 2005).

For the all-against-all structural comparion using the 7x7 RMSD matrix, each of the 33 receptors serves as a reference structure to compute the 7x7 RMSD matrices of all 33 receptors, which generates a similarity matrix that has a dimension of 33 rows and 33 x 49 = 1617 columns. The *33 x 1617* matrix of pairwise similarities is then input into the R program (R-Core-Team, 2014) to perform an average linkage clustering by calling the hclust function.

For the all-against-all analysis of using the overall 7TM bundle RMSD, we compute the pairwise overall 7TM RMSD values using the VMD software (Humphrey et al., 1996). For the all-against-all structural comparison of using the Dali Z-score, we obtain a *33 x 33* matrix of pairwise similarities by using the Dali web server (Holm and Laakso, 2016) at

http://ekhidna2.biocenter.helsinki.fi/dali. For the all-against-all structural comparison of using the TM-align score (Zhang and Skolnick, 2005), we compute the pairwise TM-align scores by using the stand alone version of the TM-align program (downloaded from http://zhanglab.ccmb.med.umich.edu/TM-align/). Because overall 7TM RMSD, Dali Z-score and TM-align score all are single-parameter indices, the resulted pairwise similarity matrices all have a dimension of 33 rows and 33 columns.

It is important to note that same as in the 7x7 RMSD matrix calculations, only the 7TM core structures are used as input for similarity calculations, including here the calculations of overall 7TM RMSDs, Dali Z-scores, and TM-align scores. It is also important to note that the same average linkage clustering algorithm is used for clustering the pairwise similarity matrices generated by all four similarity indices, i.e. calling the hclust function in the R program (R-Core-Team, 2014).

Results

We have described the 7x7 RMSD matrix in the Material and Methods section and here we apply the method to analysis of the GPCR structures available to date to address two problems: quantitative description of helix movements in activated GPCR structures and structural comparison and clustering of the 7TM cores in different GPCR receptors.

Helix Movements in Active Structures

7x7 RMSD Matrices of Active Structures—As described in the Material and Methods section, we analyze a set of seven GPCR receptors whose crystal structures are available in both inactive and active states. The bovine rhodopsin is the best characterized GPCR and more than 30 crystal structurers have been solved presenting the receptor in various states and conditions. We choose the dark-state structure 1U19 (Okada et al., 2004) as the inactive structure for analysis because of its high resolution of 2.2 Å. For the active structures, we select nine structures, 2G87(Nakamichi and Okada, 2006b), 2HPY(Nakamichi and Okada, 2006a), 3PQR(Choe et al., 2011), 3PXO(Choe et al., 2011), 3DQB(Scheerer et al., 2008), 3CAP(Park et al., 2008), 4PXF(Szczepek et al., 2014), 4ZWJ(Kang et al., 2015), 4X1H(Blankenship et al., 2015), which represent the intermediate states of the light-activation cycle of rhodopsin as well as the activated states with different binding partners at the extracellular and intracellular sides. 2G87 represents the earliest intermediate state, the bathorhodopsin state, in which the retinal chromophore has isomerized from *11*-cis to *all*-trans (Nakamichi and Okada, 2006b). 2HPY represents the lumirhodopsin state, in which the highly distorted retinal in bathorhodopsin has changed to a more relaxed conformation (Nakamichi and Okada, 2006a). 3PQR and 3PXO represent the metarhodopsin-II state with and without peptide GαCT binding at the intracellular side, respectively (Choe et al., 2011). In 3DQB, 3CAP, 4PXF, 4ZWJ and 4X1H, rhodopsin is in the retinal-free, i.e. the active Ops* state. The major differences among these structures are the intracellular binding partners, which are GαCT in 3DQB, nothing in 3CAP, the finger loop of arrestin (ArrFL) and the full length arrestin in 4ZWJ. 4X1H is essentially same as 3DQB but solved with pretty high resolution (2.3 Å). (See Table 1 for the details)

In Figure 1 and Figure S6, we show the 7x7 RMSD matrices of the nine rhodopsin active structures. All RMSDs are close to zero in the 2G87 bathorhodopsin state, and start to increase, but still smaller than 1.0 Å in the 2HPY lumirhodopsin state, which indicate that global motions of the helices have not started in these early intermediate states.

Large RMSDs with a maximum approaching 14.0 Å appear in the meta-II structures 3PQR and 3PXO, as well as the retinal-free Ops* structures 3DQB, 3CAP, 4PXF, 4ZWJ and 4X1H. The rows labeled by “S-TM6” show the red-like color most and brightest, indicating that when fitting structures using TM6, all TMs (except TM6 itself) move far from their inactive positions. However, the RMSD at position of row 6 and column 6 is smaller than 1.0 Å, indicating negligible conformational change **within** the TM6 itself. The small RMSDs at other diagonal elements also indicate no significant conformational changes within individual TMs, which ensure the conservation of the 7TM fold. Here, it is worth to point out that TM5 is rather short in the inactive (1U19), bathorhodopsin (2G87) and lumirhodopsin (2HPY) structures with the last helical turn ending at residue Gly224. In the active structures, the helical structure of TM5 ends at Gln236, i.e. 12 more residues in TM5. Those 12 residues are not included in the 7x7 RMSD matrix calculations and the matrices therefore do not reflect the movements of the extended segment of TM5.

In addition to the large RMSDs in row “S-TM6”, yellow and orange colors appear in four columns at rows labeled by “S-TM1”, “S-TM3” “S-TM4” “S-TM7” in the matrices of the fully activated structures, indicating notable movements of those TMs. The very similar RMSDs in those matrices also suggest that absence or presence of the intracellular binding partners does not make much difference in terms of rigid-body helix movements. However, there are minor differences in rows “S-TM3” and “S-TM4” between the matrices of the GαCT-bound structure 4X1H and the arrestin-bound structure 4ZWJ.

In Figure 2 and Figure S6, we show the 7x7 RMSD matrices of the non-rhodopsin GPCRs. Despite that agonists are bound with the receptors, only those bound with a G-protein or a nanobody at the intracellular side show large RMSDs in the matrices. Among these structures, the G_s protein-bound structure of β2AR (3SN6) shows the largest RMSD of 14.2 Å at row “S-TM6”. In contrast to the results of rhodopsin, the absence or presence of an extracellular binding partner does make notable differences for the non-rhodopsin structures in terms of helix movements. As experiments and molecular dynamics simulations have demonstrated that agonists alone cannot fully stabilize receptors in active conformations (Dror et al., 2011; Manglik et al., 2015; Nygaard et al., 2013). In the case of β2AR, while the G_s protein-bound structure (3SN6) shows the largest RMSD of 14.2 Å, the structure with covalently bound agonist (3PDS) in the absence of any extracellular binding partner shows RMSDs mostly smaller than 1.0 Å, indicating the latter structure essentially remains in the inactive state.

The similar heat maps presented by the matrices of the fully activated structures of rhodopsin, β2AR, M2, and μOR indicates a similar pattern of helix movements in those structures. The minor differences in individual RMSDs may reflect minor receptor-specific conformational differences. It is possible that the minor structural differences may determine the specificity of intracellular signaling pathways of the receptors (Rose et al., 2014).

The matrices of receptors NTS1, β1AR and A2A show patterns of helix movements different from those of the fully active structures. These structures are agonist-bound but without G-protein binding. The matrix of the NTS1 receptor reveals moderate movements of TM6 and TM7. As both inactive and active structures of this receptor contain a number of mutations (10 in 4GRV and 11 in 4BUO) at different positions, the 7x7 RMSD matrix may also reveal 7TM rearrangements resulted from the large number of mutations. The three A2A structures contain different mutations (two in 3QAK and five in 3YDO and 2YDV) and different agonists and have been solved in two laboratories, but all show notable movements at TM5 and TM7, which may be considered as activation features of this receptor. There are 36 mutations in the β1AR structures (2Y00 and 2Y02), which essentially lock the receptor in the inactive state and result in rather small RMSDs.

Rotations of TMs in the Active Structures—The origin of the large RMSDs shown in the 7x7 RMSD matrices is the rotation of TMs. As described in the Material and Methods section, rotations of TMs are defined as around the principal axes of the inactive reference structure. The inactive structure is re-oriented so that its three principal axes are aligned along the X, Y, and Z directions. For all structures, the Z-axis is proximately perpendicular to the membrane, pointing from the intracellular to the extracellular side. The directions of the X and Y-axes vary more with structures. For instance, the X-axis points toward the cleft between TM6 and TM7 and the Y-axis points toward the cleft between TM1 and TM2 in the inactive rhodopsin structure 1U19. The two axes point toward the clefts between TM2 and TM4 and between TM3 and TM5, respectively, in the β2AR inactive structure 2RH1 (Figure S3). In Figure 3, we show the rotational angles of TMs around the principal axes for the 21 inactive-active structure pairs in the X, Y and Z-directions.

As shown in Figure 3, TM6 rotates by almost same amount of 30 degrees around the Z-axis in the fully activated rhodopsin structures (3PQR, 3PXO, 3DQB, 3CAP, 4PXF, 4ZWJ and 4X1H), with an overall of 50 degrees rotation around three axes. Such large scale rotations of TM6 change the distances and angles between TM6 and other TMs. One of the significant consequences is that TM6 looks like tilting outwards at the intracellular side, as has been often mentioned (Choe et al., 2011; Park et al., 2008; Scheerer et al., 2008), that creates a crevice for a G-protein to bind there. The rotational angles of TM1, TM3, TM4, and TM7 are also notable, while those of TM2 and TM5 are negligible. The concerted rotations of these TMs might contribute to creating the two inter-helical openings between TM1 and TM7 and between TM5 and TM6 that have been suggested to serve as the retinal uptake and release gates (Piechnick et al., 2012; Wang and Duan, 2007; Wang and Duan, 2011).

In the five fully active non-rhodopsin structures (β2AR: 2RH1-3SN6, β2AR: 2RH1-4LDE, M2: 3UON-4MQS, M2: 3UON-4MQT, μOR: 4DKL-5C1M), TM6 also shows the largest rotations. Same as in rhodopsin, the large scale rotations of TM6 lead to the often-mentioned observation that TM6 looks like tilting outwards at the intracellular side. In addition, TM5 rotated notably in four out of the five fully active structures (except in μOR: 4DKL-5C1M). In line of this, the backbone motion of TM5 in active β1AR receptor was recently detected in an NMR experiment (Isogai et al., 2016). These results may suggest that receptor activation involves large scale rotations of both TM6 and TM5. In contrast to the fully active structures, the three semi-active structures of the A2A receptor (3EML-3QAK,

3EML-2YDO and 3EML-2YDV) show negligible rotations at TM6. Instead, TM3, TM4, TM5 and TM7 all show rotational angles up to 20 degrees. Moderate rotations of these TMs might be an activation feature of this receptor.

Altogether, as shown in Figures 1, 2 and 3, using 7x7 RMSD matrix, we are able to quantitatively describe helix movements so that we are able to identify a pattern of helix movements in an active structure such as which TMs move significantly and which not and how much TM-TM RMSDs the movements result in. The fully active structures exhibit a highly similar pattern of helix movements, whereas the semi-active structures show distinct features. Helix movements are characterized by rotations of TMs around the principal axes of the molecules, which are more complex conformational changes than simple tilt. These results suggest that helix movements induced by receptor activation are not as simple and direct as has often been suggested.

Structural Comparison of the 7TM Cores in 33 Receptors

7x7 RMSD matrices of 32 receptors against the rhodopsin structure—We analyze the crystal structures of 33 unique GPCR receptors (Table 2). As described in section Material and Methods, the start and stop residues of each TM in every structure has been carefully examined and the 7TM core residues that are present in all crystal structures are determined and defined as 1.36 – 1.57, 2.40 – 2.60, 3.23 – 3.52, 4.45 – 4.58, 5.40 – 5.61, 6.34 – 6.51 and 7.35 –7.53 by the Ballesteros-Weinstein numbers (Ballesteros and Weinstein, 1995).

As rhodopsin is the precursor of GPCR evolution and the prototype of GPCR structures (Ernst et al., 2014; Wolf and Grünewald, 2015), we use it as a reference structure to probe the structural variations among the rest 32 receptor structures. In Table 2, we show the overall 7TM RMSDs of 32 structures against the rhodopsin structure 1U19. Most of class A structures show less than 2.0 Å RMSDs and the five non-class A structures have larger RMSDs of between 2.55 Å to 3.15 Å. Superposition of the 7TM cores of the 33 structures is shown in Figure 4. We can see that although the 7TM bundles can superpose well (Figure 4a), when only one TM is used for superposing, the rest six TMs cannot fit well, which generate fuzzy images (Figure 4b–4h). This observation indicates that although the overall topology of the 7TM core is well conserved, the relative orientations of the TMs are rather different in different receptors, implying rigid-body TM movements.

We then quantify the TM movements by computing the 7x7 RMSD matrices for the 32 receptors, as described in section Material and Methods. Figure 5 shows the matrix of the β2-adrenoceptor structure (2RH1) as an example. We can see that the RMSD values at the diagonal positions are small between 0.4 Å to 1.0 Å, indicating that the helical conformations of the TMs are well conserved between rhodopsin and β2AR. However, the off-diagonal RMSDs are much larger, reaching 10 Å in the rows labeled by “S-TM4” and “S-TM5”, indicating that fitting TM4 or TM5 alone lead to large displacements of other TMs. This matrix thus reveals that TM4 and TM5 in β2-adrenoceptor moved notably relative to those in rhodopsin. The structural consequence of such movements is that inter-helical angles and distances involving TM4 and TM5 are significantly different from those in rhodopsin.

In Figure 6, we show the matrices of 32 receptors against the rhodopsin structure. To facilitate comparison of all matrices, we place the seven rows of each matrix into one line for one receptor and align the 32 lines together. We then obtain a systematic view on structural divergence and conservation within the 7TM cores. As shown in Figure 6, the RMSDs in the blocks labelled by “Superpose-TM1”, “Superpose-TM4”, “Superpose-TM5”, “Superpose-TM6” and “Superpose-TM7” are significantly larger than in the blocks labelled by “Superpose-TM2” and “Superpose-TM3”. This indicates that TM movements mainly occur on TM1, TM4, TM5, TM and TM7 while TM2 and TM3 are conserved as in rhodopsin. Within class A, TM1 and TM7 are also relatively well conserved.

Contrast to the small range of the overall 7TM RMSDs (1.57 Å to 3.15 Å in Table 2), the maximal RMSDs approach 28 Å in the 7x7 RMSD matrices (Figure 6). In addition, distributions of large RMSDs and small RMSDs significantly vary with receptors, indicating various patterns of TM movements. For instances, TM4, TM5 and TM6 show large movements in purinoreceptors P2Y1 and P2Y12; TM4 and TM5 (not TM6) move significantly in the adenosine receptor A2A, adrenoceptors β1 and β2, and muscarinic receptors M1–M4. There is no significant movement in the dopamine receptor D3 (maximal RMSD is only 5.6 Å), indicating that its 7TM core structure is highly similar to rhodopsin.

As helix movements result in changes in inter-helical angles and distances and TMs constitute orthosteric and allosteric ligand binding sites for most receptors, the diversity of the helix movements revealed here suggest that even considering only the backbones of the TMs, the ligand binding pockets are rather diverse in shape and size among different GPCRs.

Rotations of TMs in 32 Receptors—The RMSDs of fitting two structures come from two components of transformation: translation and rotation. As the overall 7TM RMSDs are rather small (Table 2), the translation of each TM should be minor. The major contributor of the large RMSDs in the 7x7 RMSD matrices is rotation of the TMs. In Figure 7, we show the rotational angles of TMs around the principal axes of the rhodopsin structure. As described in section Material and Methods, the rhodopsin structure is re-oriented so that its three principal axes are aligned in the X, Y, and Z directions. The Z-axis is proximately perpendicular to the membrane, pointing from the intracellular to extracellular side (Figure S3).

Comparing Figure 7 and Figure 6, we can see that the TMs exhibiting large rotational angles correspond to the rows with large RMSD values, confirming that TM movements revealed in the 7x7 RMSD matrices are indeed due to rotations of the TMs. In addition, Figure 7 reveals that rotations are mainly around the Z-axis, with minor rotations around X-axis or Y-axis. Also, the signs of the rotational angles indicate the rotational directions. Viewing from the extracellular to intracellular side, an angle with a positive sign means counter-clockwise rotation around the Z-axis. Thus, Rotation of TM4 is mainly clockwise whereas TM5 is counter-clockwise. In the five non-class A receptor structures rotations of TM1 and TM7 are also through the opposite directions (counter-clockwise and clockwise, respectively) around the Z-axis. TM6 rotates counter-clockwise in P2Y1, P2Y12, FFR1, PAR1 but clockwise in SMO, mGlu1, mGlu2 and glucagon.

Clustering of 7TM Core Structures in 33 Receptors—In addition to a measure of helix movements, a 7x7 RMSD matrix can be used as a 49-parameter structural similarity index for multiple structural comparison. This is so-called all-against-all comparison. In Figure 8, we show clustering of the 33 receptors by 7x7 RMSD matrix, compared with three other structural similarity indices, which are overall 7TM bundle RMSD, Dali Z-score (Holm and Laakso, 2016) and TM-align score (Zhang and Skolnick, 2005) (see Material and Methods). The sequence-based clustering of the 33 receptors is shown in Figure S7.

Dali and TM-align are two often-used structural comparison tools, which in principle can be applied to compare structures of any fold and any size. Here, we apply them to compare only the 7TM core structures with the same fold and same number of residues. In these two methods similarity scores are measured by computing intra-molecular residue-residue distances and finally summed up to a single parameter called Dali Z-score and TM-align score, respectively. So, same as using overall 7TM RMSD, each structural pair is represented by a single-parameter similarity index in Dali and TM-align. The pairwise similarity matrices computed with all four methods are provided in Supplemental Material (Tables S3, S4, S5 and S6).

Figure 8A shows the dendrogram derived by using 7x7 RMSD matrix as similarity index. We can see that the class F receptor SMO stands out and forms a single cluster, indicating that this receptor has no close relationship with any of other receptors in terms of 7TM arrangements. The rest of the receptors are grouped into two major clusters. One cluster contains all class A receptors except two lipid receptors S1P1 and LPA1. The other cluster contains non-class A receptors including the two class B receptors (CRF1 and Glugacon), two class C receptors (Mglu1 and Mglu5) and two class A lipid receptors S1P1 and LPA1. It thus appears that the 7TM domains of S1P1 and LPA1 are arranged in a manner dissimilar to other class A receptors. This is likely due to two features in the structures. One is that they do not have the highly conserved proline residue at position 5.50 (Figure S5) and thus do not have a proline-induced kink in TM5. The other feature is that the extracellular segment of TM1 is more towards outside compared to other class A receptors (Chrencik et al., 2015; Hanson et al., 2012), especially in S1P1, forming an opening between TM1 and TM7 (Hanson et al., 2012). Indeed, fitting TM1 and TM5 against rhodopsin resulted in large RMSDs (Figure 6).

The class A receptors are classified into several sub-clusters. The muscarinic receptors (M1, M2, M3 and M4), the adrenoceptors (β1 and β2), and the adenosine receptor (A2A) are close relatives despite involving three different species (human, rat, and turkey), implying their similar 7TM structures. Compared to the adrenoceptors (β1 and β2), the dopamine receptor D3 and histamine receptor H1 are closer to rhodopsin. The free fatty acid receptor FFR1 and peptide receptor PAR1 are grouped into a same clade, which is consistent with sequence-based clustering (Figure S7) as well as with the structural analysis by Okada and colleagues (Srivastava et al., 2014). These two receptors are the closest relatives of the purinoreceptors P2Y1 and P2Y12.

For almost all receptors, their subtypes are the closest relatives, indicating conservation of TM arrangements within a same subfamily. Exceptions are the 5HT1B and 5HT2B subtypes

of the 5-HT receptor, which are placed far from each other. However, this is consistent with the fact that both structures are bound with an agonist and they are in different activation states (Wacker et al., 2013; Wang et al., 2013b).

The major differences between clustering by 7x7 RMSD matrix and clustering by the three other methods (Figure 8B–8D) are on the placements of class F receptor SMO, class A lipid receptors S1P1 and LPA1, and subtypes 5HT1B and 5HT2B. SMO is placed close to Class B receptor Glugacon; S1P1 and LPA1 are grouped with the adenosine receptor A2A; subtypes 5HT1B and 5HT2B are placed in a same clade. In addition, the histamine receptor H1 is placed close to the muscarinic receptors (M1, M2, M3 and M4) and the dopamine receptor D3 close to the adrenoceptors (β1 and β2) (Figure 8B–8D). These results are similar to sequence-based clustering (Figure S7). The differences of the clustering results between 7x7 RMSD matrix and other methods indicate that 7x7 RMSD matrix is more sensitive to variations in the 7TM arrangements than other methods as it explicitly measures the orientation of each TM pair. On the other hand, if two receptors are placed into a same clade by 7x7 RMSD matrix, it indicates that the two receptors share highly similar 7TM arrangements. Such cases are common for subtypes. A non-subtype case revealed in Figure 8A is the dopamine receptor D3 and the 5HT1B receptor. As shown in the 7x7 RMSD matrix of these two receptors (Figure S8), they differ only in TM6 with a maximal RMSD of 5.9 Å.

Discussion

The transmembrane domain of a GPCR plays important roles in both structure and function. At the extracellular side, it constitutes the orthosteric endogenous ligand binding sites for class A receptors, part of the orthosteric ligand binding sites for class B receptors, and allosteric binding sites for class C and class F receptors. At the intracellular side, it couples with the G-proteins. Given the large variety of ligands and low sequence identities across families, structural variations in the TM regions are expected and have been observed in the experimentally solved structures. However, the highly similar heptahelical fold presents a challenge for the existing methods to quantify the structural differences. In this work we demonstrate that our 7x7 RMSD matrix provides a simple and efficient measure of the structural differences in the 7TM domains. With this measure, movements of all seven TMs or changes in 7TM arrangements can be quantitatively identified and visualized in one picture. It can be used to compare any structural pair to reveal detailed structural differences; the matrix itself can be used as a structural similarity index for multiple-structure comparison to reveal structural relationships.

We have compared 21 inactive-active structural pairs of seven GPCR receptors. Their 7x7 RMSD matrices reveal a highly similar pattern of helix movements in fully active structures but distinct features in semi-active structures. Rather than simply tilting, as has often been suggested, the helix movements are characterized by rotations of TMs around the principal axes of the molecules in the X, Y, and Z-directions. We have also compared each of 32 receptors of known structure with the rhodopsin structure. The results reveal high conservation of the orientations of TM2 and TM3 across four major classes while TM1, TM4, TM5, TM6 and TM7 exhibit various patterns of movements relative to the rhodopsin

structure. Such information is useful for evaluating homology models constructed for a receptor of unknown structure and predicting conformational changes in an activation state.

We have also used 7x7 RMSD matrix as a structural similarity index to classify 33 unique receptors, which reveals family-specific arrangements of the 7TM structures. Moreover, it highlights several structural relationships that are not identified by other methods, including the unique 7TM structure in SMO, separation of lipid receptors (LAP1 and S1P1) from other class A receptors, and the different 7TM arrangements in agonist-bound subtypes 5HT1B and 5HT2B. These results provide valuable insights into relations between 7TM structures, receptor functions, and receptor activation states. To obtain a complete classification of the 7TM core structures, more structures would be needed. According to Stevens and colleagues(Katritch et al., 2013), at least 100 representative receptor structures are needed to cover the GPCR superfamily with a cutoff of 35% sequence identity.

In the structural comparison and classification of 33 receptors, we used one structure for one receptor. Certainly, a 7x7 RMSD matrix could be different if another structure was used. But, the conformational differences in multiple inactive-state structures of a same receptor have been found to be minor and negligible especially in the 7TM domains (Katritch et al., 2013). As a test, we add another structure of receptor CXCR4 into the all-against-all comparison data set, in which receptor CXCR4 is bound with a cyclic peptide antagonist (PDB id: 3OE0 (Wu et al., 2010)). As shown in Figure S9 and Figure S10, the two structures (3ODU and 3OE0) are rather similar, and adding 3OE0 into the data set does not change the structural clustering at all and the two CXCR4 structures are placed as the closest relatives.

Although we develop and demonstrate here the 7x7 RMSD matrix for structural comparison of GPCRs, the method is directly applicable to other 7TM proteins such as the microbial rhodopsin family. Moreover, the idea used in developing the 7x7 RMSD matrix may be adopted to derive similar methods for structural comparison of other protein families with conserved folds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project is supported in part by Natural Science Foundation of China (Project ID: U1430237) and National Institute of Health (GM079383 to YD). The 1000 talent award to HL is also acknowledged.

References

Ballesteros, JA., Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In: Stuart, CS., editor. Methods in Neurosciences. Academic Press; 1995. p. 366-428.

Blankenship E, Vahedi-Faridi A, Lodowski David T. The High-Resolution Structure of Activated Opsin Reveals a Conserved Solvent Network in the Transmembrane Region Essential for Activation. Structure. 2015; 23:2358–2364. [PubMed: 26526852]

Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-Resolution Crystal Structure of an Engineered Human {beta}2-Adrenergic G Protein Coupled Receptor. Science. 2007; 25:1258–1265.

Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science. 2010; 330:1091–1095. [PubMed: 21097933]

Choe HW, Kim YJ, Park JH, Morizumi T, Pai EF, Krausz N, Hofmann KP, Scheerer P, Ernst OP. Crystal structure of metarhodopsin II. Nature. 2011; 471:651–655. [PubMed: 21389988]

Chrencik JE, Roth CB, Terakado M, Kurata H, Omi R, Kihara Y, Warshaviak D, Nakade S, Asmar-Rovira G, Mileni M, Mizuno H, Griffith MT, Rodgers C, Han GW, Velasquez J, Chun J, Stevens RC, Hanson MA. Crystal Structure of Antagonist Bound Human Lysophosphatidic Acid Receptor 1. Cell. 2015; 161:1633–1643. [PubMed: 26091040]

Christopher JA, Aves SJ, Bennett KA, Dore AS, Errey JC, Jazayeri A, Marshall FH, Okrasa K, Serrano-Vega MJ, Tehan BG, Wiggin GR, Congreve M. Fragment and Structure-Based Drug Discovery for a Class C GPCR: Discovery of the mGlu5 Negative Allosteric Modulator HTL14242 (3-Chloro-5-[6-(5-fluoropyridin-2-yl)pyrimidin-4-yl]benzonitrile). J Med Chem. 2015; 58:6653–6664. [PubMed: 26225459]

Dalton JA, Lans I, Giraldo J. Quantifying conformational changes in GPCRs: glimpse of a common functional mechanism. BMC Bioinformatics. 2015; 16:124. [PubMed: 25902715]

Dror RO, Arlow DH, Maragakis P, Mildorf TJ, Pan AC, Xu H, Borhani DW, Shaw DE. Activation mechanism of the b2-adrenergic receptor. Proceedings of the National Academy of Sciences. 2011; 108:18684–18689.

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:E655–662. [PubMed: 24453215]

Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H. Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. Chem Rev. 2014; 114:126–163. [PubMed: 24364740]

Fenalti G, Giguere PM, Katritch V, Huang XP, Thompson AA, Cherezov V, Roth BL, Stevens RC. Molecular control of [dgr]-opioid receptor signalling. Nature. 2014; 506:191–196. [PubMed: 24413399]

Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 2003; 63:1256–1272. [PubMed: 12761335]

Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. Nature. 2012; 482:547–551. [PubMed: 22278061]

Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, Sanna MG, Han GW, Kuhn P, Rosen H, Stevens RC. Crystal structure of a lipid G protein-coupled receptor. Science. 2012; 335:851. [PubMed: 22344443]

Hollenstein K, Kean J, Bortolato A, Cheng RKY, Dore AS, Jazayeri A, Cooke RM, Weir M, Marshall FH. Structure of class B GPCR corticotropin-releasing factor receptor 1. Nature. 2013; 499:438–443. [PubMed: 23863939]

Holm L, Laakso LM. Dali server update. Nucleic Acids Research. 2016; 44:W351–W355. [PubMed: 27131377]

Huang W, 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 micro-opioid receptor activation. Nature. 2015; 524:315–321. [PubMed: 26245379]

Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996; 14:33–38. [PubMed: 8744570]

Isberg V, Vroling B, van der Kant R, Li K, Vriend G, Gloriam D. GPCRDB: an information system for G protein-coupled receptors. Nucleic Acids Res. 2014; 42:D422–425. [PubMed: 24304901]

Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, Vroling B, Bojarski AJ, Vriend G, Gloriam DE. GPCRdb: an information system for G protein-coupled receptors. Nucleic Acids Research. 2016; 44:D356–D364. [PubMed: 26582914]

Isogai S, Deupi X, Opitz C, Heydenreich FM, Tsai CJ, Brueckner F, Schertler GFX, Veprintsev DB, Grzesiek S. Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. Nature. 2016; 530:237–241. [PubMed: 26840483]

Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman 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]

Kang Y, Zhou XE, Gao X, He Y, Liu W, Ishchenko A, Barty A, White TA, Yefanov O, Han GW, Xu Q, de Waal PW, Ke J, Tan MHE, Zhang C, Moeller A, West GM, Pascal BD, Van Eps N, Caro LN, Vishnivetskiy SA, Lee RJ, Suino-Powell KM, Gu X, Pal K, Ma J, Zhi X, Boutet S, Williams GJ, Messerschmidt M, Gati C, Zatsepin NA, Wang D, James D, Basu S, Roy-Chowdhury S, Conrad CE, Coe J, Liu H, Lisova S, Kupitz C, Grotjohann I, Fromme R, Jiang Y, Tan M, Yang H, Li J, Wang M, Zheng Z, Li D, Howe N, Zhao Y, Standfuss J, Diederichs K, Dong Y, Potter CS, Carragher B, Caffrey M, Jiang H, Chapman HN, Spence JCH, Fromme P, Weierstall U, Ernst OP, Katritch V, Gurevich VV, Griffin PR, Hubbell WL, Stevens RC, Cherezov V, Melcher K, Xu HE. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. Nature. 2015; 523:561–567. [PubMed: 26200343]

Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol. 2013; 53:531–556. [PubMed: 23140243]

Kinoshita M, Okada T. Structural conservation among the rhodopsin-like and other G protein-coupled receptors. Sci Rep. 2015; 5:9176. [PubMed: 25775952]

Kobilka BK. G Protein Coupled Receptor Structure and Activation. Biochimica et biophysica acta. 2007; 1768:794–807. [PubMed: 17188232]

Kruse AC, Ring AM, Manglik A, Hu J, 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]

Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat Rev Drug Discov. 2008; 7:339–357. [PubMed: 18382464]

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]

Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the u-opioid receptor bound to a morphinan antagonist. Nature. 2012; 485:321–326. [PubMed: 22437502]

Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, Lerch MT, Kobilka TS, Thian FS, Hubbell WL, Prosser RS, Kobilka BK. Structural Insights into the Dynamic Process of beta2-Adrenergic Receptor Signaling. Cell. 2015; 161:1101–1111. [PubMed: 25981665]

Nakamichi H, Okada T. Local peptide movement in the photoreaction intermediate of rhodopsin. Proc Natl Acad Sci U S A. 2006a; 103:12729–12734. [PubMed: 16908857]

Nakamichi H, Okada T. Crystallographic analysis of primary visual photochemistry. Angew Chem Int Ed Engl. 2006b; 45:4270–4273. [PubMed: 16586416]

Nygaard R, Zou Y, Dror Ron O, Mildorf Thomas J, Arlow Daniel H, Manglik A, Pan Albert C, Liu Corey W, Fung Juan J, Bokoch Michael P, Thian Foon S, Kobilka Tong S, Shaw David E, Mueller L, Prosser RS, Kobilka Brian K. The Dynamic Process of β2-Adrenergic Receptor Activation. Cell. 2013; 152:532–542. [PubMed: 23374348]

Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. J Mol Biol. 2004; 342:571–583. [PubMed: 15327956]

Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP. Crystal structure of the ligand-free G-protein-coupled receptor opsin. Nature. 2008; 454:183–187. [PubMed: 18563085]

Piechnick R, Ritter E, Hildebrand PW, Ernst OP, Scheerer P, Hofmann KP, Heck M. Effect of channel mutations on the uptake and release of the retinal ligand in opsin. Proc Natl Acad Sci U S A. 2012; 109:5247–5252. [PubMed: 22431612]

R-Core-Team. R: A Language and Environment for Statistical Computing. 2014. http://www.Rproject.org/

Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. Nature. 2011; 477:549–555. [PubMed: 21772288]

Ring AM, Manglik A, Kruse AC, Enos MD, Weis WI, Garcia KC, Kobilka BK. Adrenaline-activated structure of beta2-adrenoceptor stabilized by an engineered nanobody. Nature. 2013; 502:575–579. [PubMed: 24056936]

Rose AS, Elgeti M, Zachariae U, Grubmüller H, Hofmann KP, Scheerer P, Hildebrand PW. Position of Transmembrane Helix 6 Determines Receptor G Protein Coupling Specificity. Journal of the American Chemical Society. 2014; 136:11244–11247. [PubMed: 25046433]

Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SGF, Choi HJ, DeVree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK. Structure and function of an irreversible agonist-[bgr]2 adrenoceptor complex. Nature. 2011; 469:236–240. [PubMed: 21228876]

Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP. Crystal structure of opsin in its G-protein-interacting conformation. Nature. 2008; 455:497–502. [PubMed: 18818650]

Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S. Structure of the human histamine H1 receptor complex with doxepin. Nature. 2011; 475:65–70. [PubMed: 21697825]

Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, Liu W, Lau J, Cherezov V, Katritch V, Wang MW, Stevens RC. Structure of the human glucagon class B G-protein-coupled receptor. Nature. 2013; 499:444–449. [PubMed: 23863937]

Srivastava A, Yano J, Hirozane Y, Kefala G, Gruswitz F, Snell G, Lane W, Ivetac A, Aertgeerts K, Nguyen J, Jennings A, Okada K. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. Nature. 2014; 513:124–127. [PubMed: 25043059]

Szczepek M, Beyriere F, Hofmann KP, Elgeti M, Kazmin R, Rose A, Bartl FJ, von Stetten D, Heck M, Sommer ME, Hildebrand PW, Scheerer P. Crystal structure of a common GPCR-binding interface for G protein and arrestin. Nat Commun. 2014; 5:4801. [PubMed: 25205354]

Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q, Wu B. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. Science. 2013; 341:1387–1390. [PubMed: 24030490]

Tehan BG, Bortolato A, Blaney FE, Weir MP, Mason JS. Unifying family A GPCR theories of activation. Pharmacol Ther. 2014; 143:51–60. [PubMed: 24561131]

Thal DM, Sun B, Feng D, Nawaratne V, Leach K, Felder CC, Bures MG, Evans DA, Weis WI, Bachhawat P, Kobilka TS, Sexton PM, Kobilka BK, Christopoulos A. Crystal structures of the M1 and M4 muscarinic acetylcholine receptors. Nature. 2016; 531:335–340. [PubMed: 26958838]

Thompson AA, Liu W, Chun E, Katritch V, Wu H, Vardy E, Huang XP, Trapella C, Guerrini R, Calo G, Roth BL, Cherezov V, Stevens RC. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. Nature. 2012; 485:395–399. [PubMed: 22596163]

Thorsen TS, Matt R, Weis WI, Kobilka BK. Modified T4 Lysozyme Fusion Proteins Facilitate G Protein-Coupled Receptor Crystallogenesis. Structure. 2014; 22:1657–1664. [PubMed: 25450769]

Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, Liu W, Xu HE, Cherezov V, Roth BL, Stevens RC. Structural features for functional selectivity at serotonin receptors. Science. 2013; 340:615–619. [PubMed: 23519215]

Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC. Structure of the human smoothened receptor bound to an antitumour agent. Nature. 2013a; 497:338–343. [PubMed: 23636324]

Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang XP, Vardy E, McCorvy JD, Gao X, Zhou XE, Melcher K, Zhang C, Bai F, Yang H, Yang L, Jiang H, Roth BL, Cherezov V, Stevens RC, Xu HE. Structural basis for molecular recognition at serotonin receptors. Science. 2013b; 340:610–614. [PubMed: 23519210]

Wang T, Duan Y. Chromophore channeling in the G-protein coupled receptor rhodopsin. J Am Chem Soc. 2007; 129:6970–6971. [PubMed: 17500517]

Wang T, Duan Y. Retinal release from opsin in molecular dynamics simulations. Journal of Molecular Recognition. 2011; 24:350–358. [PubMed: 21360617]

Warne T, Moukhametzianov R, Baker JG, Nehme R, Edwards PC, Leslie AG, Schertler GF, Tate CG. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. Nature. 2011; 469:241–244. [PubMed: 21228877]

Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Structure of a beta1-adrenergic G-protein-coupled receptor. Nature. 2008a; 454:486–491. [PubMed: 18594507]

Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Structure of a beta1-adrenergic G-protein-coupled receptor. Nature. 2008b; 454:486–491. Epub 2008 Jun 2025. [PubMed: 18594507]

White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R. Structure of the agonist-bound neurotensin receptor. Nature. 2012; 490:508–513. [PubMed: 23051748]

Wolf S, Grünewald S. Sequence, Structure and Ligand Binding Evolution of Rhodopsin-Like G Protein-Coupled Receptors: A Crystal Structure-Based Phylogenetic Analysis. PLoS ONE. 2015; 10:e0123533. [PubMed: 25881057]

Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science. 2010; 330:1066–1071. [PubMed: 20929726]

Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, Conn PJ, Stevens RC. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. Science. 2014; 344:58–64. [PubMed: 24603153]

Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V, Stevens RC. Structure of the human kappa-opioid receptor in complex with JDTic. Nature. 2012; 485:327–332. [PubMed: 22437504]

Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC. Structure of an agonist-bound human A2A adenosine receptor. Science. 2011; 332:322–327. [PubMed: 21393508]

Yin J, Mobarec JC, Kolb P, Rosenbaum DM. Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. Nature. 2015; 519:247–250. [PubMed: 25533960]

Yin J, Babaoglu K, Brautigam CA, Clark L, Shao Z, Scheuermann TH, Harrell CM, Gotter AL, Roecker AJ, Winrow CJ, Renger JJ, Coleman PJ, Rosenbaum DM. Structure and ligand-binding mechanism of the human OX1 and OX2 orexin receptors. Nat Struct Mol Biol. 2016; 23:293–299. [PubMed: 26950369]

Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, Green HF, Pandey A, Dror RO, Shaw DE, Weis WI, Coughlin SR, Kobilka BK. High-resolution crystal structure of human protease-activated receptor 1. Nature. 2012; 492:387–392. [PubMed: 23222541]

Zhang D, Gao ZG, Zhang K, Kiselev E, Crane S, Wang J, Paoletta S, Yi C, Ma L, Zhang W, Han GW, Liu H, Cherezov V, Katritch V, Jiang H, Stevens RC, Jacobson KA, Zhao Q, Wu B. Two disparate ligand-binding sites in the human P2Y1 receptor. Nature. 2015a; 520:317–321. [PubMed: 25822790]

Zhang H, Unal H, Gati C, Han GW, Liu W, Zatsepin NA, James D, Wang D, Nelson G, Weierstall U, Sawaya MR, Xu Q, Messerschmidt M, Williams GJ, Boutet S, Yefanov OM, White TA, Wang C, Ishchenko A, Tirupula KC, Desnoyer R, Coe J, Conrad CE, Fromme P, Stevens RC, Katritch V, Karnik SS, Cherezov V. Structure of the Angiotensin receptor revealed by serial femtosecond crystallography. Cell. 2015b; 161:833–844. [PubMed: 25913193]

Zhang K, Zhang J, Gao ZG, Zhang D, Zhu L, Han GW, Moss SM, Paoletta S, Kiselev E, Lu W, Fenalti G, Zhang W, Muller CE, Yang H, Jiang H, Cherezov V, Katritch V, Jacobson KA, Stevens RC, Wu B, Zhao Q. Structure of the human P2Y12 receptor in complex with an antithrombotic drug. Nature. 2014; 509:115–118. [PubMed: 24670650]

Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Research. 2005; 33:2302–2309. [PubMed: 15849316]

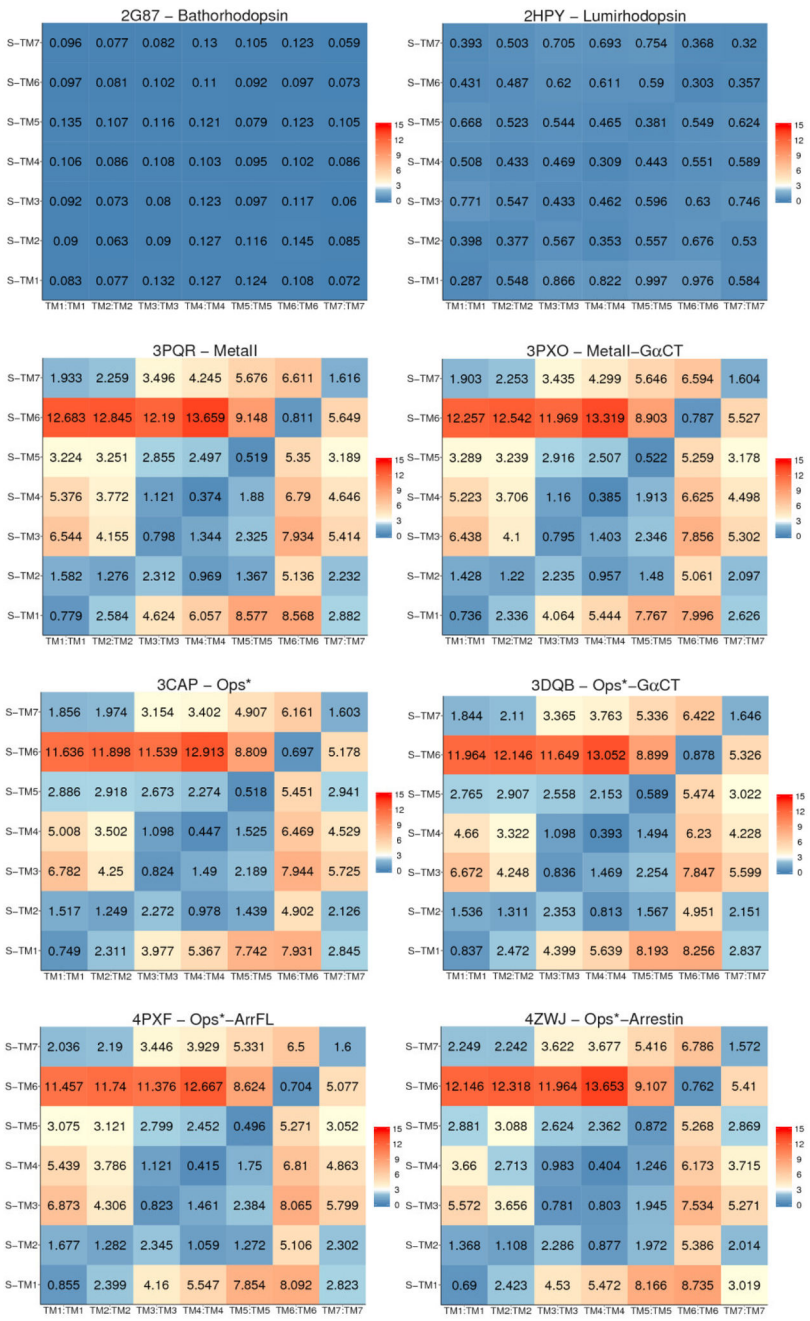

Figure 1.

7x7 RMSD matrices of bovine rhodopsin active structures against the inactive structure (1U19). Each matrix is labeled by the PDB id of the active structure, the activated state, and the intracellular binding partner, if present in the x-ray structure. Also see Figure S6.

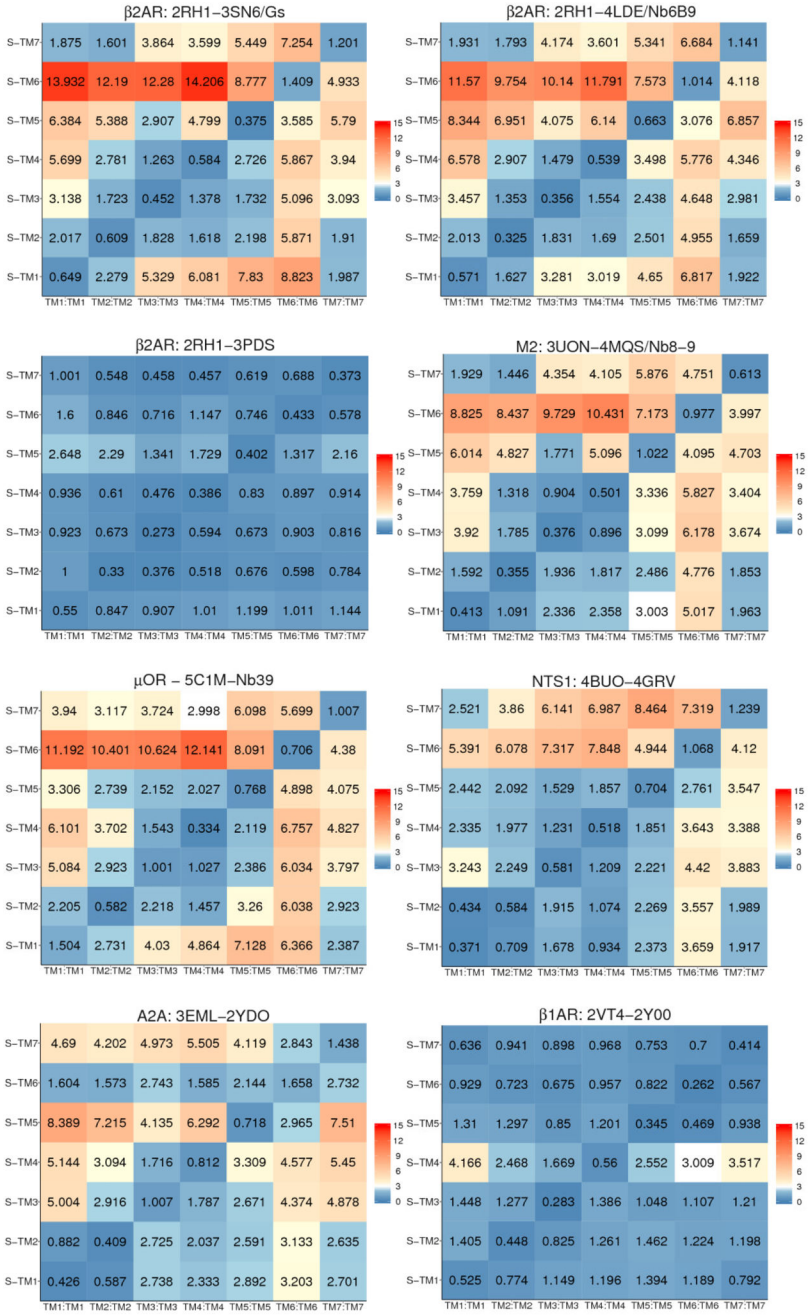

Figure 2.

7x7 RMSD matrices of non-rhodopsin GPCR active structures including receptors β2AR, M2, μOR, β1AR, A2A and NTS1. Each matrix is labeled by the receptor name, the PDB id of the inactive structure, the PDB id of the active structure, and the intracellular binding partner, if present in the x-ray structure. Also see Figure S6.

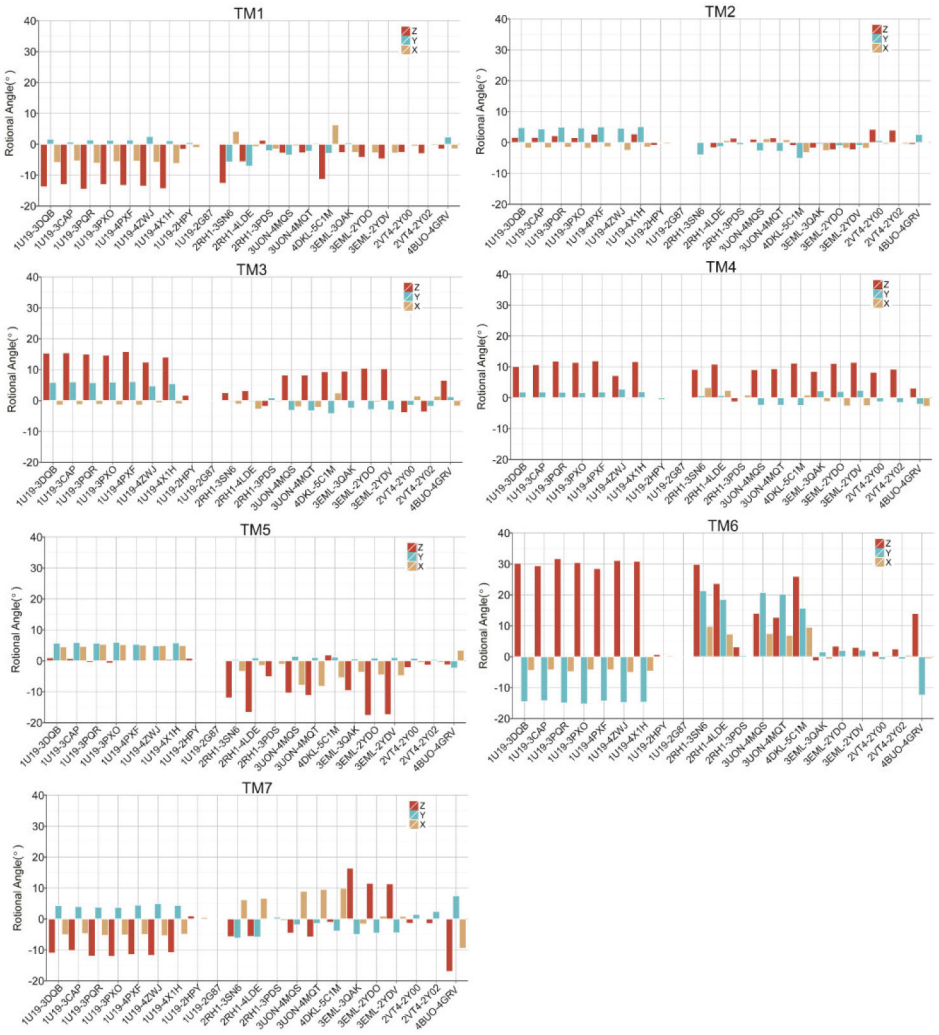

Figure 3.

Rotational angles of TM1 – TM7 computed for the 21 inactive-active structure pairs around the principal axes in the X, Y, and Z-directions. Z-axis is approximately perpendicular to the membrane.

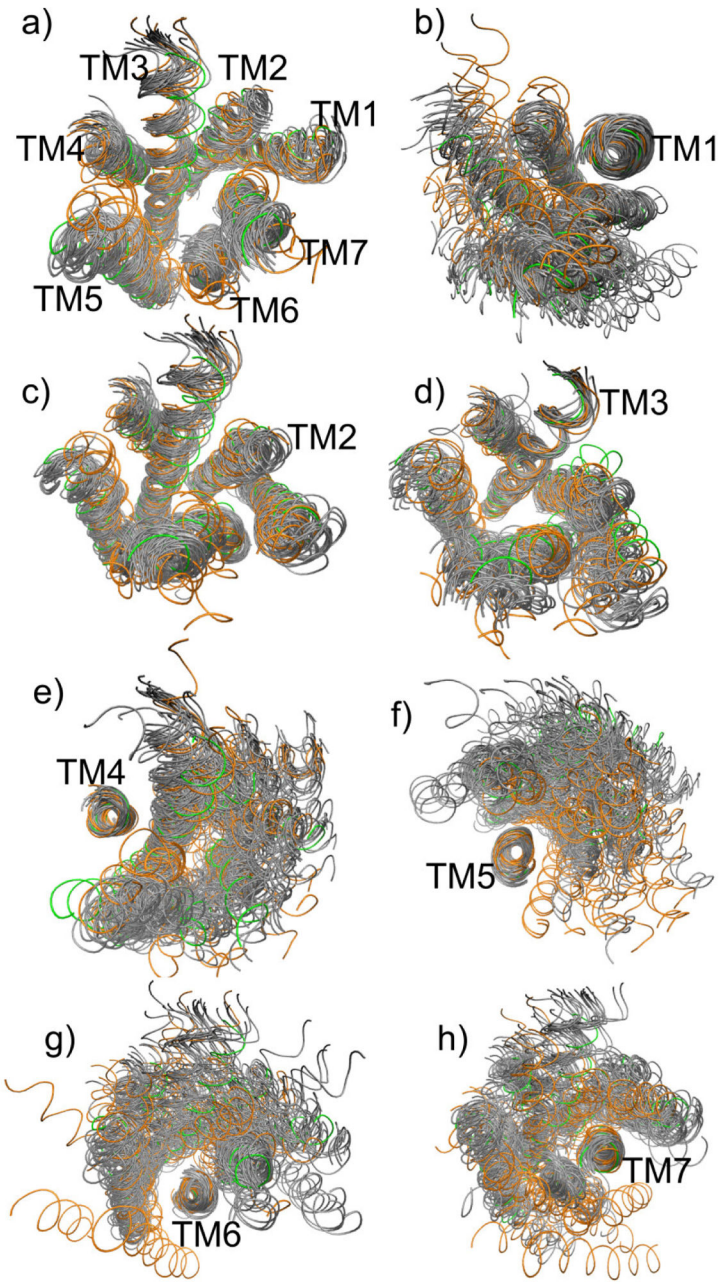

Figure 4.

Superposition of the 7TM cores of 33 GPCR structures using the rhodopsin structure (green) 1U19 as a reference. a) all 7 TMs are fitted; b) only TM1 is fitted; c) only TM2 is fitted; d) only TM3 is fitted; e) only TM4 is fitted; f) only TM5 is fitted; g) only TM6 is fitted; h) only TM7 is fitted. The class A receptors are shown in grey and non-class A in orange. 24

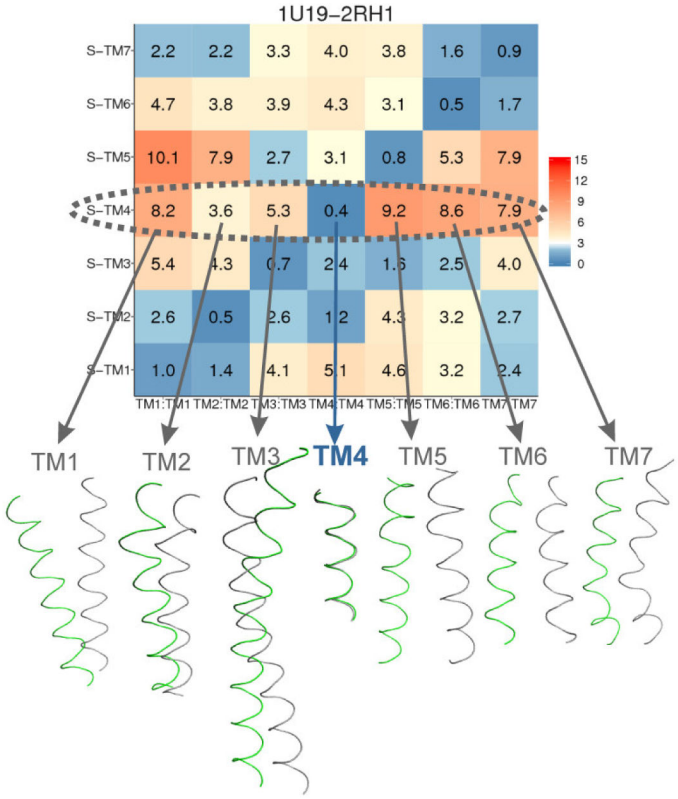

Figure 5.

7x7 RMSD matrix of β2-adrenoceptor (β2AR) structure 2RH1 against the rhodopsin structure 1U19. Structural superposition is shown for the row of fitting TM4 (labelled by "S-TM4"). TMs of rhodopsin are in green and those of β2AR in silver.

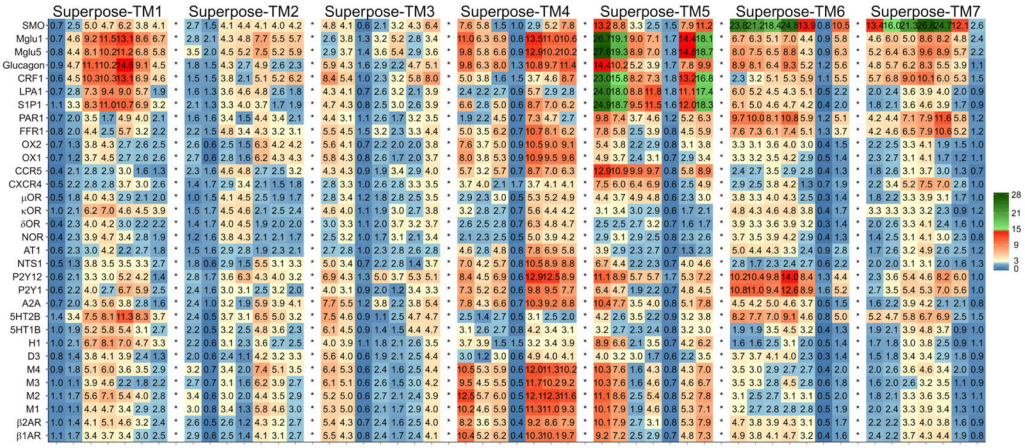

Figure 6.

7x7 RMSD matrices of 32 GPCR structures against the rhodopsin structure 1U19. For easy comparison of all matrices, the seven rows of each matrix are placed into one line for each receptor.

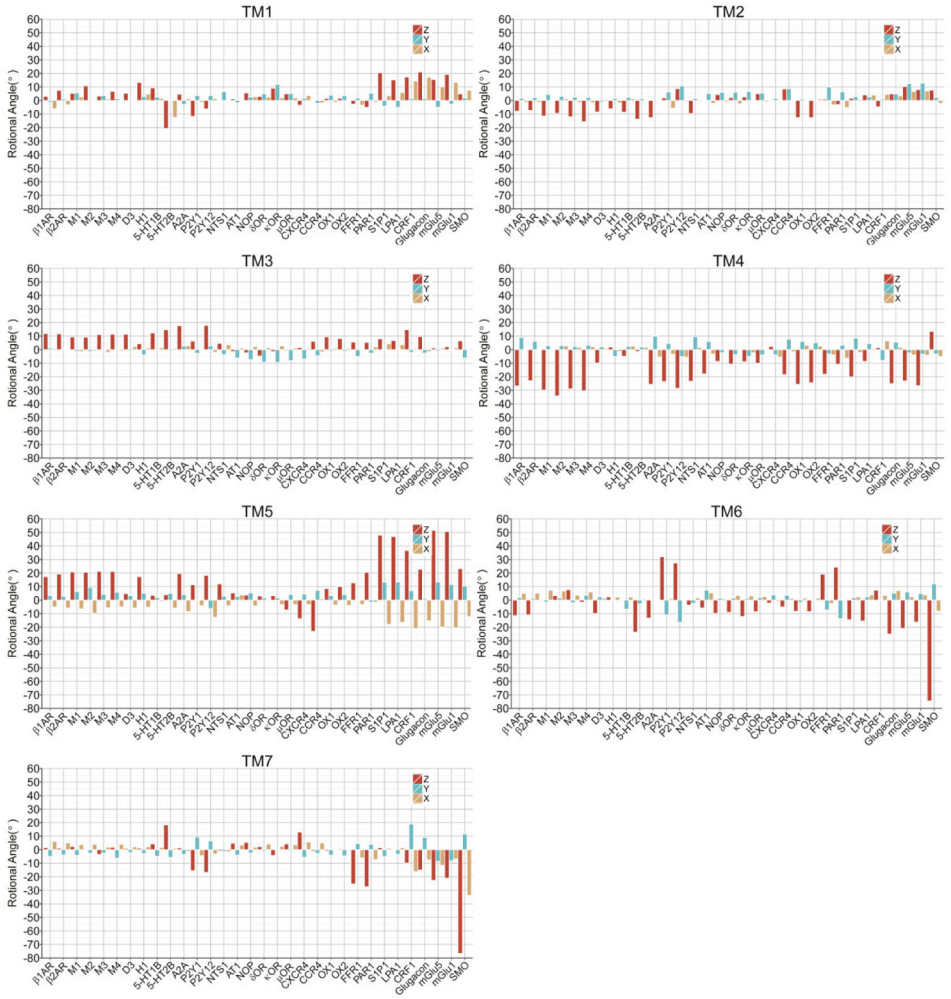

Figure 7.

Rotational angles of each TM in 32 receptors relative to the rhodopsin structure 1U19. Rotations are computed around the X, Y and Z-directions, which are aligned along the principal axes of the rhodopsin structure 1U19. The Z-axis is proximately perpendicular to the membrane, pointing from the intracellular to extracellular side.

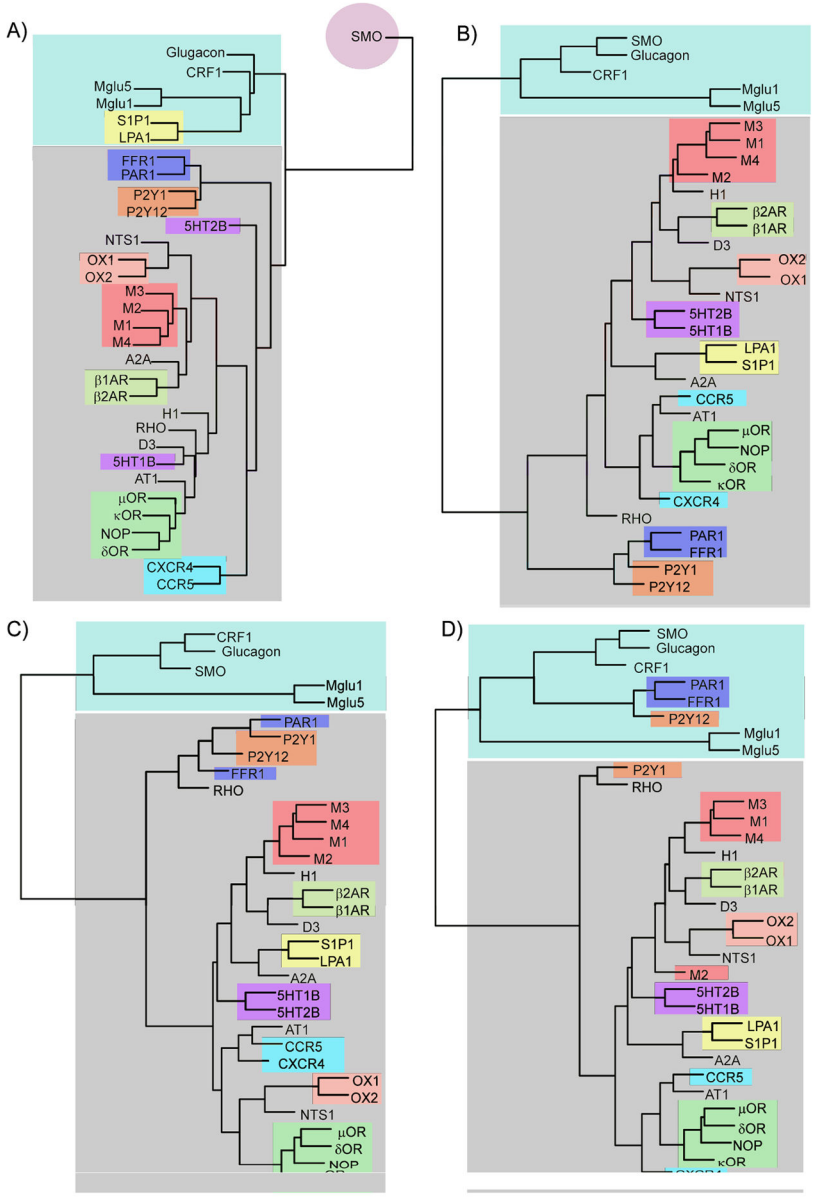

Figure 8.

Clustering of 7TM core structures in 33 receptors. A) by 7x7 RMSD matrix; B) by overall 7TM RMSD; C) by Dali Z-score; D) by TM-align score. See Figure S7 for sequence-based clustering result.

J Struct Biol. Author manuscript; available in PMC 2018 August 01.

Table 1

Structures of seven GPCR receptors in inactive and active states

Receptor[a]	**Inactive Structure**		**Active Structure**			
rhodopsin	PDB id[a]	Resolution (Å)	PDB id[b]	Resolution (Å)	Extracellular ligand	Intracellular binding partner
	1U19	2.2	2G87	2.6	Retinal	No
			2HPY	2.8	Retinal	No
			3PQR	2.85	Retinal	GαCT
			3PXO	3.0	Retinal	No
			3DQB	3.2	No	GαCT
			3CAP	2.9	No	No
			4PXF	2.75	β-D-octylglucoside	Arrestin fingerloop
			4ZWJ	3.3	No	Arrestin
			4X1H	2.3	Nonyl-glucoside	GαCT
β2AR	2RH1	2.4	3SN6	3.2	Agonist BI-167107	G_s protein
			3PDS	3.5	Covalent agonist FAUC50	No
			4LDE	2.8	Agonist BI-167107	Nanobody 6B9
M2	3UON	3.0	4MQS	3.5	Agonist iperoxo	Nanobody 8–9
			4MQT	3.7	Agonist iperoxo & Allosteric modulator LY2119620	Nanobody 8–9
μOR	4DKL	2.8	5C1M	2.1	Agonist BU72	Nanobody 39
β1AR	2VT4	2.7	2Y00	2.5	Partial agonist dobutamine	No
			2Y02	2.6	Agonist carmoterol	No
A2A	3EML	2.6	3QAK	2.7	Agonist UK-432097	No
			2YDO	3.0	Agonist adenosine	No
			2YDV	2.6	Agonist NECA	No
NTS1	4BUO	2.75	4GRV	2.8	Agonist neurotensin C-terminal	No

[a]Abbreviations of receptor names

[b]Due to space limit in table, references of the structures are only provided in the text.

Table 2

Structures of 33 unique GPCR receptors, including sequence identities and overall 7TM RMSDs against the bovine rhodopsin structure (1U19).

Receptor	Abb[a]	Class	Species	PDB id[b]	Seq.%[c]	7TM RMSD[d] (Å)
rhodopsin	RHO	A	bovine	1U19	100.0	0.0
Adrenoceptor β1	β1AR	A	turkey	2VT4	25.3	1.72
Adrenoceptor β2	β2AR	A	human	2RH1	25.3	1.70
Muscarinic receptor M1	M1	A	human	5CXV	23.3	1.90
Muscarinic receptor M2	M2	A	human	3UON	28.1	1.93
Muscarinic receptor M3	M3	A	rat	4U15	24.0	1.90
Muscarinic receptor M4	M4	A	human	5DSG	24.0	1.97
Dopamine receptor D3	D3	A	human	3PBL	31.5	1.57
Histamine receptor H1	H1	A	human	3RZE	20.5	1.78
5-Hydroxytryptamine receptor 1B	5HT1B	A	human	4IAR	25.3	1.81
5-Hydroxytryptamine receptor 2B	5HT2B	A	human	4IB4	24.7	1.94
Adenosine receptor A2A	A2A	A	human	3EML	27.4	1.94
P2Y purinoreceptor 1	P2Y1	A	human	4XNV	24.0	1.92
P2Y purinoreceptor 12	P2Y12	A	human	4NTJ	22.6	2.51
Neurotensin receptor 1	NTS1	A	rat	4BUO	26.0	1.97
Angiotensin receptor 1	AT1	A	human	4YAY	23.3	1.79
Nociception/orphan FQ receptor	NOP	A	human	4EA3	26.0	1.68
δ-opioid receptor	δOR	A	human	4N6H	26.7	1.92
κ-opioid receptor	κOR	A	human	4DJH	24.7	2.14
μ-opioid receptor	μOR	A	mouse	4DKL	28.8	1.79
C-X-C chemokine receptor 4	CXCR4	A	human	3ODU	28.1	1.98
C-C chemokine receptor 5	CCR5	A	human	4MBS	28.1	1.98
Orexin receptor 1	OX1	A	human	4ZJ8	22.6	1.96
Orexin receptor 2	OX2	A	human	4S0V	23.3	1.90
Free fatty acid receptor 1	FFR1	A	human	4PHU	21.9	2.58
Proteinase-activated receptor 1	PAR1	A	human	3VW7	21.2	2.21

Receptor	Abb[a]	Class	Species	PDB id[b]	Seq.%[c]	7TM RMSD[d] (Å)
Lysophosphatidic acid receptor 1	LPA1	A	human	4Z35	24.0	2.08
Sphingosine 1-phosphate receptor 1	S1P1	A	human	3V2Y	25.3	2.07
Cortincotropin-releasing factor receptor 1	CRF1	B	human	4K5Y	12.3	3.15
Glucagon receptor	Glucagon	B	human	4L6R	13.0	2.55
Metabotropic glutamate receptor 1	Mglu1	C	human	4OR2	11.0	2.99
Metabotropic glutamate receptor 5	Mglu5	C	human	5CGD	11.6	3.06
Smoothened receptor	SMO	F	human	4JKV	11.0	2.79

[a]Abbreviations of receptor names

[b]Due to space limit in table, references of the structures are provided only in the text.

[b,c]7TM sequence identity compared to rhodopsin 1U19 and backbone RMSDs were calculated for the core TM region. The 7TM residues are defined by Ballesteros-Weinstein numbers 1.36–1.57, 2.40–2.60, 3.23–3.52, 4.45–4.58, 5.40–5.61, 6.34–6.51 and 7.35–7.53. See Figure S5 for sequence alignment of the core TM residues.