Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

# Transmembrane helix 6 observed at the interface of $\beta_2$AR homodimers in blind docking studies

Ayca Koroglu & Ebru Demet Akten

**To cite this article:** Ayca Koroglu & Ebru Demet Akten (2015) Transmembrane helix 6 observed at the interface of $\beta_2$AR homodimers in blind docking studies, Journal of Biomolecular Structure and Dynamics, 33:7, 1503-1515, DOI: 10.1080/07391102.2014.962094

**To link to this article:** https://doi.org/10.1080/07391102.2014.962094

*Journal of Biomolecular Structure and Dynamics*, 2015
Vol. 33, No. 7, 1503–1515, http://dx.doi.org/10.1080/07391102.2014.962094

Taylor & Francis
Taylor & Francis Group

# Transmembrane helix 6 observed at the interface of $\beta_2$AR homodimers in blind docking studies

Ayca Koroglu[a] and Ebru Demet Akten[b]*

[a]*Graduate School of Computational Biology and Bioinformatics, Kadir Has University, Cibali, 34083 Istanbul, Turkey;* [b]*Faculty of Engineering and Natural Sciences, Department of Bioinformatics and Genetics, Kadir Has University, Cibali, 34083 Istanbul, Turkey*

Communicated by Ramaswamy H. Sarma

(*Received 5 June 2014; accepted 2 September 2014*)

Peptide– and protein–protein dockings were carried out on $\beta_2$-adrenergic receptor ($\beta_2$AR) to confirm the presence of transmembrane helix 6 (TM6) at the interface region between two $\beta_2$AR monomers, thereby its possible role in dimerization as suggested in numerous experimental and computational studies. Initially, a portion of TM6 was modeled as a peptide consisting of 23 residues and blindly docked to $\beta_2$AR monomer using a rigid body approach. Interestingly, all highest score conformations preferred to be near TM5 and TM6 regions of the receptor. Furthermore, longer peptides generated from a whole TM region were blindly docked to $\beta_2$AR using the same rigid body approach. This yielded a total of seven docked peptides, each derived from one TM helix. Most interestingly, for each peptide, TM6 was among the most preferred binding site region in the receptor. Besides the peptide dockings, two $\beta_2$AR monomers were blindly docked to each other using a full rigid-body search of docking orientations, which yielded a total of 16,000 dimer conformations. Each dimer was then filtered according to a fitness value based on the membrane topology. Among 149 complexes that met the topology requirements, 102 conformers were composed of two monomers oriented in opposite directions, whereas in the remaining 47, the monomers were arranged in parallel. Lastly, all 149 conformers were clustered based on a root mean-squared distance value of 6 Å. In agreement with the peptide results, the clustering yielded the largest population of conformers with the highest Z-score value having TM6 at the interface region.

**Keywords:** $\beta_2$-adrenergic receptor; transmembrane helix 6; protein–protein docking; homodimer; peptide–protein docking; interface; $\beta_2$AR dimer

## 1. Introduction

$\beta_2$-Adrenergic receptor ($\beta_2$AR) is a G-protein coupled receptor (GPCR), which is the member of the largest family of seven-transmembrane-helix receptors that are responsible for transmitting a wide variety of signals such as odorants, photons, hormones, and neurotransmitters (Strader, Fong, Tota, Underwood, & Dixon, 1994). A well-known feature of any GPCR-mediated signaling is that it is initiated by the formation of a ternary complex consisting of the active GPCR, the G protein, and the signaling molecule (Fung, Hurley, & Stryer, 1981). In this ternary model, $\beta_2$AR is monomeric while activating G proteins (Whorton et al., 2007). However, there exist several experimental observations, which show $\beta_2$AR forming dimeric and even oligomeric associations to perform certain functions (Angers et al., 2000; Bouvier, 2001; Hebert et al., 1996; Milligan, 2007).

In this respect, using differential epitope tagging and co-immunoprecipitation, Hebert and his coworkers (1996) demonstrated that $\beta_2$ARs formed homodimers, and that the receptor was found in a dynamic equilibrium between monomeric and dimeric forms under basal conditions. While the addition of an agonist stabilized the $\beta_2$AR dimer, the presence of an inverse agonist shifted the equilibrium toward the monomer formation. Furthermore, a peptide derived from the transmembrane domain 6 (TM6) inhibited the dimerization to different extents depending on the ratio of dimer to monomer in the system. This finding suggested that TM6 may represent part of an interface or may play a role in receptor dimerization. Their study also supported the functional importance of dimerization since the peptide that inhibited dimerization also inhibited β-adrenergic agonist-promoted stimulation of adenylyl cyclase activity.

Following the peptide experiments that provided a direct biochemical evidence for the existence of $\beta_2$AR homodimers, Angers and his coworkers demonstrated the constitutive homodimerization of human $\beta_2$AR in living cells using a biophysical approach; bioluminescence resonance energy transfer (BRET) (Angers et al., 2000). A few years later, the same group carried out BRET experiments to study $\beta_1$ and $\beta_2$AR homo- and heterodimerization (Lavoie et al., 2002; Mercier, Salahpour, Angers, Breit, & Bouvier, 2002). The potential heterodimerization was

*Corresponding author. Email: demet.akten@khas.edu.tr

investigated due to their high level of sequence identity. Their results clearly showed a similar propensity of $\beta_1$ and $\beta_2$AR to form homo- and heterodimers, and a relatively high proportion of constitutive dimeric receptors. On a functional level, the adenylyl cyclase activity was not affected by heterodimerization. However, the agonist-promoted internalization of the $\beta_2$AR and its ability to activate the ERK1/2 MAPK signaling pathway was inhibited by the presence of $\beta_1$AR and $\beta_1$AR/$\beta_2$AR heterodimerization. The same group also exposed the functional significance of $\beta_2$AR dimerization as being a prerequisite for cell surface targeting (Salahpour et al., 2004). Inhibition of the dimerization motif $^{276}$GXXXGXXXL$^{284}$ on TM6 prevented transport of the receptor to the plasma membrane.

The oligomerization of $\beta_2$AR was first suggested by fluorescence recovery after photobleaching studies (Dorsch, Klotz, Engelhardt, Lohse, & Bünemann, 2009) conducted on $\beta_2$AR, which revealed extensive and stable oligomers (probably 4–5 receptors) both in HEK 293T cells and in cardiac myocytes. In the same year, Fung and his coworkers conducted fluorescence resonance energy transfer (FRET) and suggested that $\beta_2$AR is capable of forming tetrameric assemblies in a lipid bilayer and that the binding of inverse agonists promotes the oligomerization by reducing conformational fluctuations in individual protomers (Fung et al., 2009).

Despite the growing amount of experimental evidence about the existence of dimerization/oligomerization, the information about the interface region is still lacking. Hebert and his coworkers suggested TM6 as the most plausible interface region, however with no direct confirmation (Hebert et al., 1996). Computer simulations conducted by Gouldson and his coworkers on 5,6-domain swapped dimers where the dimer interface lies between TM5 and TM6, proposed them as the active, high affinity form induced by agonists (Gouldson, Snell, Bywater, Higgs, & Reynolds, 1998). In the same study, the correlated mutation analysis identified the mutations mainly amongst the external residues found at the TM5-TM6 region. The presence of correlated mutations is known to indicate a functionally important interface region, also supported by a large amount of computational studies (Gouldson, Bywater, & Reynolds, 1997; Gouldson & Reynolds, 1997; Gouldson et al., 2001; Nilsson et al., 1999; Oliveira, Paiva, & Vriend, 1993; Pazos, Helmer-Citterich, Ausiello, & Valencia, 1997).

One computational study based on spatial distribution of highly conserved residues on the surface of the receptor, predicted TM6 as the most plausible interface region in agreement with Hebert et al.'s experimental result (Nemoto & Toh, 2005). A sequence alignment study conducted on 700 aligned GPCR sequences using a data-mining approach (Livingstone & Barton, 1993) determined the external faces of TM5 and TM6 with significant levels of conservation and thus suggested them as potential functional sites for each family of receptors (Dean et al., 2001; Gkoutos, Higgs, Bywater, Gouldson, & Reynolds, 1999). Immunological studies conducted on $A_1$ adenosine receptor (Ciruela et al., 1995), suggested that intracellular loop 3 (ICL3) connecting TM5 to TM6 becomes less exposed in the dimer, suggesting that TM5 and TM6 might play a role in dimerization.

In a review study by Filizola and Weinstein (2005), it is noted that some TM segments appeared more often than others in the prediction studies of rhodopsin-like GPCR interfaces. By collecting the results of different studies conducted so far, the authors showed that most of the lipid-exposed residues predicted most frequently were within TM4, TM5, and TM6 regions. Among the residues identified within each of these three helices, 4.58, 5.48, and 6.42 had the greatest number of occurrences. Especially, 6.42 had almost the same number of occurrence as 6.30, which is located at the boundary between TM6 and the cytoplasmic loop ICL3. This was explained by a possible involvement of ICL3 in dimerization/oligomerization, also suggested by the atomic force microscopy studies (Liang et al., 2003) conducted on rhodopsin indicating that ICL3 facilitates the formation of rows of dimers.

Considering both experimental and computational findings regarding the location of the interface region, the computational docking study presented here used short and long peptides derived from different TM regions of $\beta_2$AR that were rigidly and blindly docked to the monomeric form of the receptor. Our computational approach for peptide docking was fully inspired by the peptide experiments of Hebert and his coworkers (Hebert et al., 1996). The goal here was to determine the TM regions of the receptor for which the docked peptide would have the highest affinity and consequently suggest a plausible interface region. In the second part of our study, we used a protein–protein blind docking approach using two exactly similar $\beta_2$AR monomers to propose a dimeric form of the receptor and consequently its corresponding interface. The receptor conformation that was used as a target structure in all docking runs represents an inactive state with a relatively wrapped up ICL3 that blocks the intracellular G-protein binding site. Such an ICL3 alignment is selected, since it would interfere the least in any kind of dimeric or oligomeric association. Also, an inactive state presents a relatively constrained structure with fewer conformational fluctuations, which favors the closer packing of dimeric/oligomeric associations, as suggested by FRET experiments (Fung et al., 1981).

## 2. Methods

### *2.1. System preparation for AutoDock*

The peptide molecule to be docked to $\beta_2$AR monomer was derived from a part of TM6 region consisting of

residues from 276 to 296, GIIMGTFTLCWLPFFIVNIVH, as it is used in the activity measurements conducted by Hebert and his group (Hebert et al., 1996). By adopting the same conformation as in the receptor, the peptide was blindly docked using a rigid body approach via AutoDock v4.0 software tool (Morris et al., 2009). Docked conformations were then reevaluated with a knowledge-based scoring function called DSX$^{online}$ v0.88 (Neudert & Klebe, 2011). The selection of a rigid body approach was adopted since the flexibility of the whole peptide would be computationally too demanding and partial flexibility would only bias the docking results. Furthermore, a rigid peptide would better represent the stable helical segment that preserves its secondary structure, especially when embedded in the transmembrane region.

As the docking tool, AutoDock was especially selected based on the satisfactory accuracy rates reported on membrane proteins previously (Liu, Perez Aguilar, Liang, & Saven, 2012; Jois & Siahaan, 2003; Song et al., 2009). Especially, its robust and efficient algorithm was shown in the blind docking study of peptides conducted by Hetenyi and his coworkers (Hetenyi & Spoel, 2002). Regarding the membrane environment, the proposed dimer structure has been predicted based on the fact that homodimerization is constitutive (hormone independent) and that it occurs in the endoplasmic reticulum prior to being transported to the cell membrane (Salahpour et al., 2004). Since the lipid contribution to dimerization is not required at all times, its effect was simply discarded from our docking runs.

The monomeric conformation of $\beta_2$AR was selected among 5608 snapshots obtained from four independent molecular dynamics simulation trajectories (one 1 μs long MD and three 100 ns long MD runs) (Ozcan, Uyar, Doruker, & Akten, 2013). These simulations were carried out to understand the effect of the long ICL3 on the intrinsic dynamics of $\beta_2$AR and yielded a 'very inactive' conformation of the receptor, which was characterized by the close packing of ICL3 underneath the membrane (See Figure S1).

The starting crystal structure for the MD study was the inactive state of $\beta_2$AR (PDB id:2RH1), which was the first reported crystal structure in 2007 by Cherezov and his coworkers (Cherezov et al., 2007). The ICL3 region was removed from the receptor and replaced by an anchor protein T4-lysozyme to facilitate crystallization. Prior to our MD study, T4-lysozyme was removed and the missing ICL3 region was generated as an unstructured loop of 32-residues length via MODELLER homology tool (Narayanan et al., 2003). The simulated system consisted of a protein embedded into a POPC membrane and surrounded by water molecules. The preparation of the system prior to MD consisted of three stages; melting of lipids, protein under restraints, and finally protein released. Each of these preparation steps consisted of 1000 steps of energy minimization followed by .5 ns MD runs. The equilibrated system was then subjected to a total of 1000 ns MD run and three separate 100 ns MD runs, which produced a total of 5608 different snapshots.

In the end, all 5608 conformers were clustered based on the root mean-squared distance value of ICL3 using k-clust algorithm of MMTSB toolset (Feig, Karanicolas, & Brooks, 2001). Next, a conformer incorporating an ICL3 well packed underneath the receptor was selected considering the fact that an ICL3 with such an orientation would interfere the least with the dimerization (See Figure 1(a)). The conformational variation in the transmembrane region, which is the target site of blind

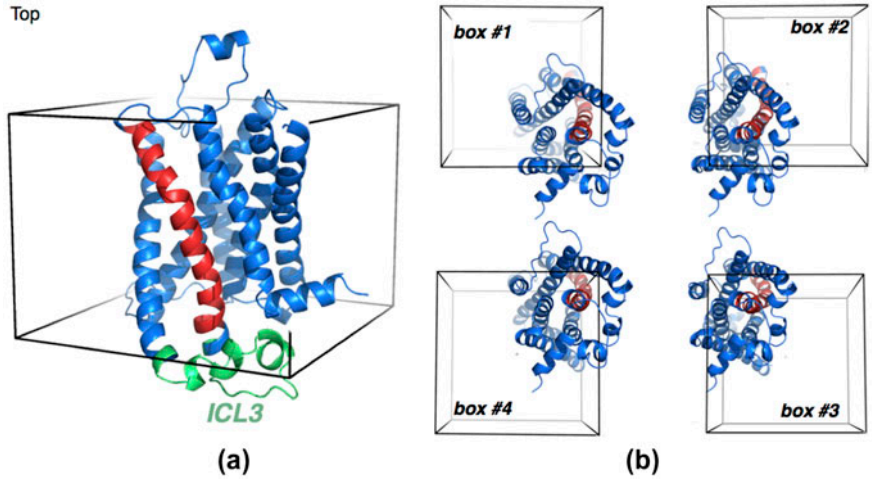


Figure 1. $\beta_2$AR monomer inside (a) a single box with .45 Å grid spacing and 56.25 × 56.25 × 42.3 Å dimensions, (b) four smaller boxes each with .375 Å grid spacing and 46.88 × 46.88 × 46.88 Å dimensions. The transmembrane helix VI is colored in red and the ICL3 is colored in green.

docking in our study, was below 2 Å among the snapshots, thus using one single conformer of the receptor was sufficient.

A pre-calculated three-dimensional energy grid of equally spaced discrete points was generated before docking for a rapid energy evaluation using the program AutoGrid (Morris et al., 2009). In the first docking experiment, a grid box that is large enough to cover the receptor yet excluding the extracellular loops was built using a grid spacing of .45 Å (See Figure 1(a)). This is simply a 'blind docking' experiment where the peptide is free to move in a conformational space covering all transmembrane regions of the receptor. In the second docking, four distinct grid boxes with a smaller grid spacing of .375 Å were generated such that each box only covers approximately one quarter of the receptor's transmembrane region (See Figure 1(b)). For each five docking experiment, a total of 200 docking runs were employed using the Lamarckian genetic algorithm (Morris et al., 1998) as the conformational search. Finally, all 200 docked poses were clustered based on a root mean square distance (RMSD) value of 2 Å.

Additional blind dockings were performed on $\beta_2$AR monomer, using a longer peptide generated from the whole helical region of each TM helix. Here, the goal was to reveal the pair of TM helices with the highest binding affinity that might also indicate a likely interface region. For seven TM helices, seven independent blind docking experiments with 200 runs each, were performed using a single box with a .45 Å grid spacing which covers the entire transmembrane region as in the blind docking experiments of the short peptide.

For both short and long peptide dockings, a total of $2.5 \times 10^6$ energy evaluations and 27,000 generations were assigned to each run, which makes a total of $5 \times 10^8$ energy evaluations for one peptide only. In addition, population size, mutation, and crossover rates were set to 150, .02, and .8, respectively. The energy convergence was reached early in the run, after about $1.0 \times 10^6$ energy evaluations. Each docking run takes an average of 3.5–4 h for short peptides and 4–6 h for long peptides, on a 2.67 GHz/8 MB IntelR Xeon processor.

### 2.2. System preparation and evaluation steps for ZDOCK

The protein–protein docking tool ZDOCK v2.3 (Chen, Li & Weng, 2003) was used to estimate the dimer structure of $\beta_2$AR. The docking program, which is based on a Fast Fourier Transform, takes two exactly similar $\beta_2$AR monomers and searches all possible binding modes in the translational and rotational space. Moreover, it takes into account surface complementarity, electrostatics, and desolvation to find the optimal fit between two proteins.

The input parameters were provided via a web-based interface. Default parameters cannot be modified, however, one can select certain residues that are not likely to appear in the protein–protein interface (in our system, all intracellular and extracellular loop regions) and these residues will simply be blocked from the conformational search space. For the default parameters, cutoff values for RMSD and Interface were defined as 6.0 and 9.0, respectively. Also, the maximum number of clusters was set to 60 as a default. The computing time scales as $1/\Delta 3$, where $\Delta$ is the rotational sampling interval. For our run, a value of 6° was used since it provides more accurate predictions by sampling 54,000 poses. The average computing time with a 6° rotational sampling interval and a $128 \times 128 \times 128$ grid is 30 min on a 16-processor IBM SP3.

By default, the program can generate up to 2000 complexes. To obtain a large variety of dimer structures, two different conformations of the receptor, one with ICL3 and one without ICL3, were used as the initial conformation. Furthermore, certain residues such as intra- and extracellular loop regions were blocked from being in the binding site during conformational search to sample only biologically relevant dimer association. Table 1 summarizes eight different layouts used in protein–protein docking experiments, each generating 2000 complexes, which amounts to a total of 16,000 conformers to be analyzed.

All 16,000 complexes were filtered based on their 'membrane topology' feature consisting of the tilt angle about the *z*-axis that is perpendicular to the surface of the membrane and the *z*-offset taken as the displacement of the geometrical center of the monomer along the *z*-axis. The tilt angle is calculated as the angle between the first principle coordinate axis of the receptor and the *z*-axis. .4 radians and 6 Å were employed as maximum allowable values for tilt angle and *z*-offset, respectively. The complexes conforming to the topological requirements were clustered based on an RMSD value of 6 Å,

Table 1. Set of docking experiments with different conformers and blocked residues.

| Set # | Presence of ICL3 | Blocked residues | # of complexes |
|---|---|---|---|
| 1 | Yes | NONE | 2000 |
| 2 | Yes | ECL2 and ICL3 on one monomer | 2000 |
| 3 | Yes | All loops on one monomer | 2000 |
| 4 | Yes | All loops on both monomers | 2000 |
| 5 | No | NONE | 2000 |
| 6 | No | ECL2 and ICL3 on one monomer | 2000 |
| 7 | No | All loops on one monomer | 2000 |
| 8 | No | All loops on both monomers | 2000 |

using *k*-means clustering algorithm (Feig, Karanicolas, & Brooks, 2001). Two conformations representing the most populated clusters for parallel and antiparallel dimers were selected and further analyzed for the presence of TM6 at the interface region. Finally, all complexes with an acceptable membrane topology were reevaluated using a ZRANK scoring function, which consists of detailed electrostatics, van der Waals, and desolvation terms (Pierce & Weng, 2007). The complexes with the highest ZRANK score values were proposed as the most plausible conformations and similarly inspected for the presence of TM6 at the interface region. The whole procedure described so far was summarized in a flowchart as illustrated in Figure 2.

To determine the interface region between two monomers in a dimer, a quantitative approach was used. First, a representative complex from each cluster was selected as being the closest conformer to the cluster's average structure (or cluster's centroid). Next, the two monomers in the complex were set apart without being subject to any local conformational change, leaving them completely disconnected. Finally, the amount of change in solvent accessible surface area (SASA) of each residue $i$ upon complex formation was determined using the following equation,

$$\mathrm{dSASA}_i = \frac{\mathrm{SASA}_{u,i} - \mathrm{SASA}_{c,i}}{\mathrm{SASA}_{u,i}}$$

where $\mathrm{SASA}_{u,i}$ and $\mathrm{SASA}_{c,i}$ are the SASA values of residue $i$ in the monomer before and after complex formation, respectively. The value of dSASA varies between 0 (apart from the interface) and 1 (completely buried at the interface). If a residue is not located at the interface, its dSASA is expected to be zero since no change in its SASA value is observed upon dimerization, whereas, if a residue $i$ is completely buried at the interface, the change

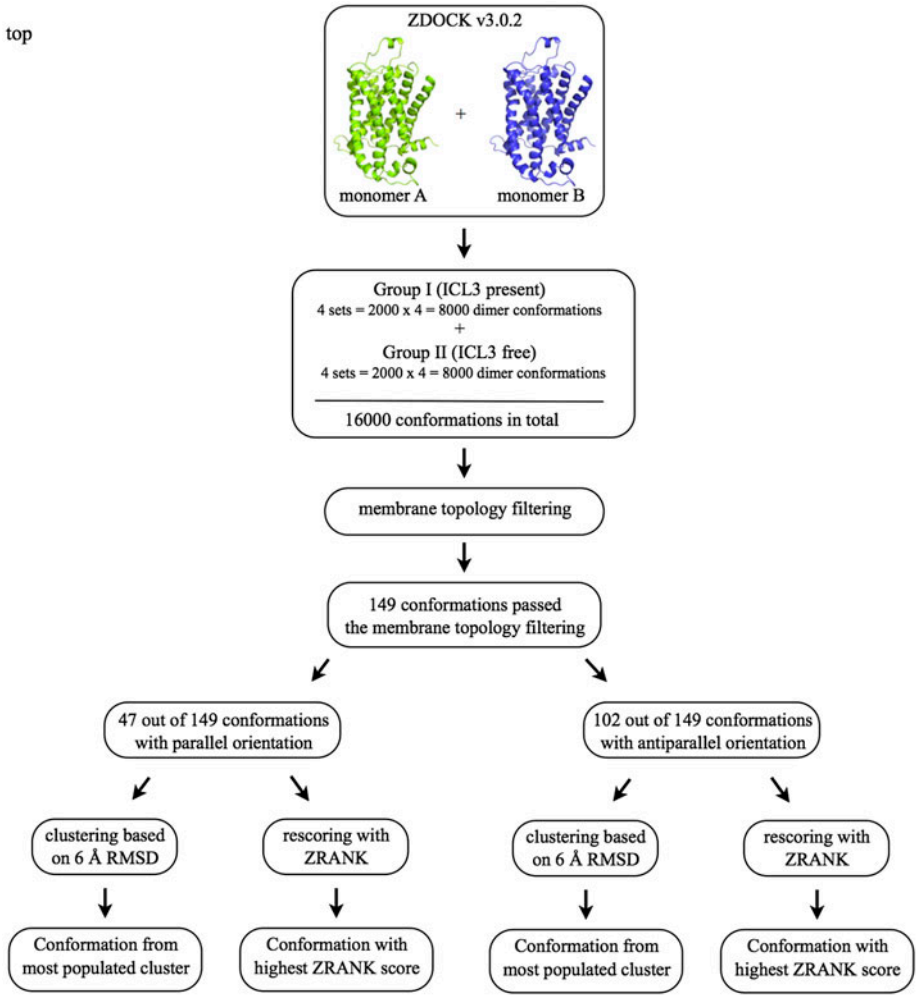


Figure 2. The flowchart that describes the procedure for determining the most plausible dimer conformation among the docked poses generated by ZDOCK (Chen et al., 2003).

in SASA value is expected to be close to one. All SASA calculations were conducted in VMD using the *measure SASA* module with a probe radius of 1.4 Å larger than the van der Waals radius (Humphrey, Dalke, & Schulten, 1996).

## 3. Results and discussion

### 3.1. Docking results for the short peptide extracted from TM6 region

Docking results for the peptide consisting of residues of TM6 from 276 to 296 in one single large box and four smaller boxes are listed in Table 2. For each run, the transmembrane regions of the target receptor that are confined in the specified grid box are provided in the second column. Docking results are listed for the first cluster corresponding to the cluster with the highest negative AutoDock (AD) score and the most-populated cluster that holds the highest number of docked conformers designated by the cluster's size. Alongside the AD score, the corresponding DSX/CSD (derived from Cambridge Structural Database, (Allen, 2002)), DSX/PDB (derived from Protein Data Bank, (Berman, 2002)), and X-Score (Wang, Lai, & Wang, 2002) values are provided as well.

As listed in the first row of Table 2, the peptide preferred to bind near TM5 and TM6 regions in 66 out of 200 docked conformers in blind docking experiments for which a large box covering all TM regions was used. Furthermore, four scoring functions ranked in the top three in all 8 clusters (5 *first* and 3 *most-populated*) are typed in bold and underlined in Table 2. Accordingly, the blind docking using all TM helices (first row) yielded a peptide pose with all four score values ranked in the top three. The same pose was also illustrated in Figure 3(a), where the peptide nicely fits alongside TM5 and TM6 (colored in red) making interactions through residues shown on the interaction profile in Figure 3(b). Additionally, a representative pose taken from the most-populated cluster with 84 elements shows that the peptide was oriented towards TM1 and TM7, but also making interactions with the intracellular part of TM5 and TM6 (See Figure 3(c) and (d)). Likewise, the peptide was found to be oriented upside down.

The highest number of conformers observed in a cluster was achieved in docking experiments confined to selected regions which included a significant portion of TM6, as in boxes #2 and #3 with 199 and 197 poses in the first cluster, respectively (See Table 2). Besides, their corresponding DSX/CSD, DSX/PDB, and X-score values were considerably higher than those of the two docking runs conducted in boxes #1 and #4 that exclude TM6.

Overall, considering the score values and the number of docked poses in the first cluster, the preference of TM6 over the other regions is noticeable in all peptide dockings. The representative snapshots of the peptide's pose in the first and most-populated clusters for other short peptide dockings listed in Table 2 are provided alongside their interaction profiles as supplementary materials (See Figure S2).

Table 2. Docking results for the short peptide consisting of residues from 276 to 296 in TM6.

| Run Type | Box #/ (Covered TM Regions) | First cluster | | | | | | Most-populated cluster | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Size[a] | TM position[b] | AD score[c] | DSX/ CSD score[d] | DSX/ PDB score[e] | X-Score[f] | Size[a] | TM position[b] | AD score[c] | DSX/ CSD scored | DSX/ PDB score[d] | X-Score[f] |
| Blind docking | All TM helices | 66 | V, VI | **−10.5** | **−102.6** | **−114.2** | **−7.92** | 84 | I, V, VI, VII | −10.2 | −97.7 | **−115.9** | −7.89 |
| Docking with selected binding sites | 1/(2–5) | 38 | II, III, IV | **−10.6** | −89.9 | −100.2 | −7.76 | 109 | III, IV, V | −9.1 | −77.1 | −98.1 | −6.89 |
| | 2/(3, 5–7) | 199 | V, VI | **−10.6** | **−101.0** | −111.5 | −7.81 | – | – | – | – | – | |
| | 3/(1, 5–8) | 197 | I, V, VI, VII | **−10.4** | −93.3 | −112.3 | **−7.90** | – | – | – | – | – | |
| | 4/(1–4, 7, 8) | 43 | II, III, IV | **−10.6** | −88.2 | −98.4 | −7.76 | 73 | I, II, III, IV | **−10.6** | **−114.2** | **−127.9** | **−8.92** |

Notes: Four scoring functions ranked in the top three in all 8 clusters (5 *first* and 3 *most-populated*), were typed in bold and underlined.
[a]Number of docked poses in the cluster.
[b]Transmembrane region to which the docked pose is within a distance of 5 Å at most.
[c]Highest negative AutoDock score in the first cluster (kcal/mol).
[d]DSX/CSD score of the best pose of AutoDock (unitless).
[e]DSX/PDB score of the best pose of AutoDock (unitless).
[f]X-Score's predicted binding energy for the best pose of AutoDock (kcal/mol).

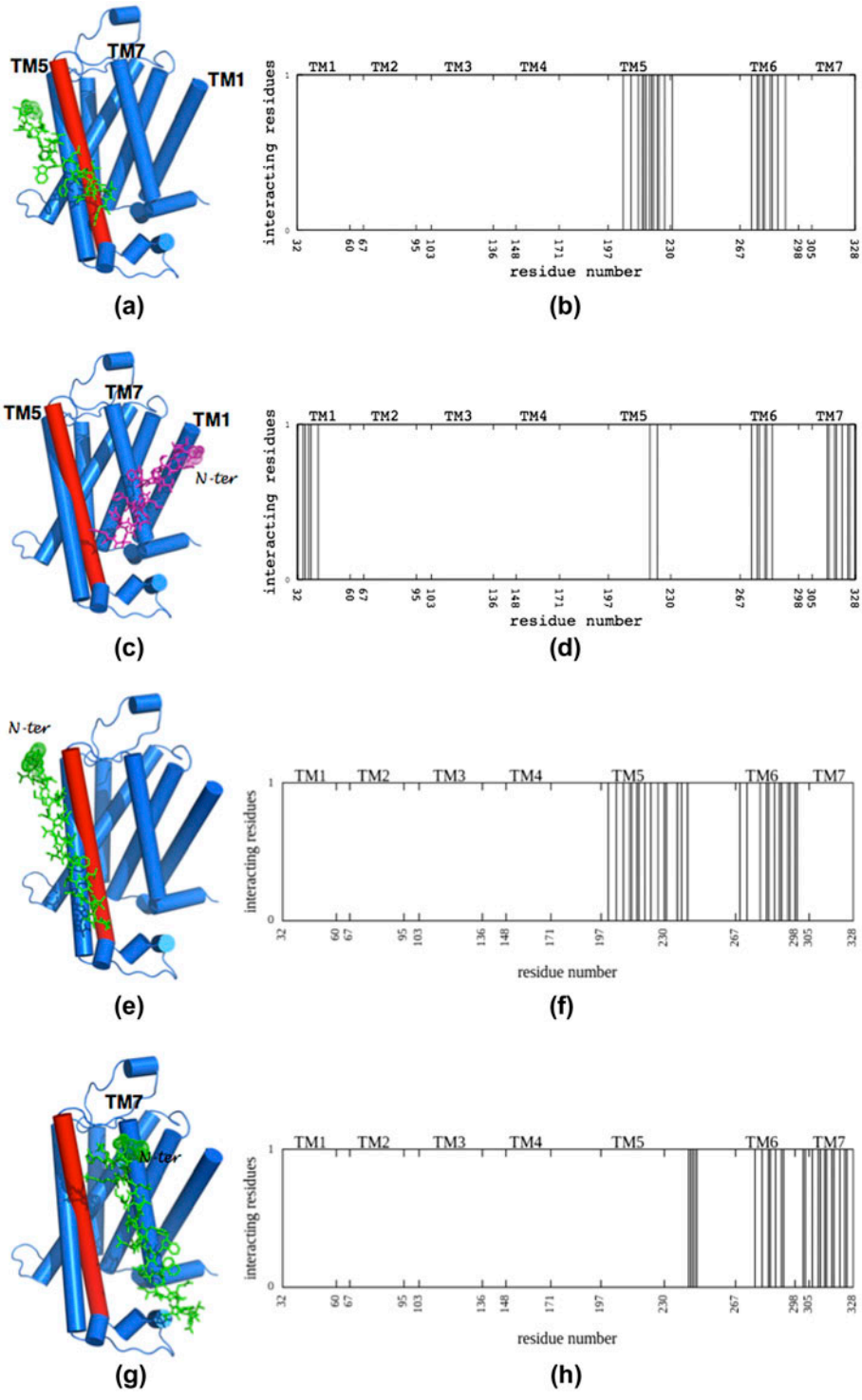


Figure 3. Snapshots of the short peptide's pose in blind docking runs from the (a) *first* cluster (green) and the (c) *most-populated* cluster (magenta), and their corresponding profiles of the neighboring residues within a radius of 5 Å in (b) and (d), respectively. Snapshots of the (e) TM1-derived and (g) TM6-derived peptide's poses from the first cluster (green) and, their corresponding profiles of the neighboring residues within a radius of 5 Å in (f) and (h), respectively. The dots on the peptide correspond to amino terminal end. TM6 region of the receptor is illustrated in red.

### *3.2. Docking results for the long peptides generated from each TM helices*

Each docking experiment corresponds to a peptide generated from the whole region of one of the seven TM helices, blindly docked to $\beta_2$AR monomer. Surprisingly, for all seven docking runs, the best pose in the first cluster and the pose in the most-populated cluster were persistently located near TM5, TM6, and TM7 regions (See Table 3). The score values ranked in the top three for all 10 clusters (7 *first* and 3 *most-populated*) are typed in bold and underlined as in Table 2. Accordingly, the TM1-, TM2-, and TM6-derived peptides were found to dock with three high score values, which are all ranked in the top three and their best poses are located near TMV, TMVI, and TMVII regions of the receptor. Similar to short-peptide docking results, there is a strong accordance between DSX and X-Score values. Moreover, in the first four dockings (TM1 through TM4), the *first* cluster is observed to be the *most-populated* cluster with a number of peptide conformations varying between 104 and 178, and additionally, their corresponding best pose has a high negative AutoDock score value varying between −10.6 and −9.5 kcal/mol.

As listed in Table 3, the highest negative AutoDock score value of −10.64 belongs to TM3-derived peptide that was bound to the monomer near TM5, TM6, and TM7 regions. Moreover, the highest negative DSX/CSD score value of −103.6 belongs to TM2-derived peptide, whereas the highest negative DSX/PDB score value of −121.5 belongs to TM6-derived peptide. Figure 3(e) and (g) illustrate the position of the best pose of TM1- and TM6-derived peptides interacting with the monomer's TM6 region, as well as an interaction profile showing all the residues, which are found within a radius of 5 Å of the best pose (See Figure 3(f) and (h)). Similar snapshots alongside an interaction profile for the other six TM-derived peptides are provided as supplementary material (See Figure S3). These results suggest that in all possible interface regions in a dimer formation, either TMV-TMVI, or TMV-TMVI-TMVII, or TMVI-TMVII helices of one monomer would be observed with any region of a second monomer. This persistent binding of all seven helices might also indicate that TM5-TM6-TM7 presents an energetically more favorable binding cavity for any kind of helical structure than any other site in the receptor.

### *3.3. Protein–Protein docking results indicates TM6 region as the most plausible site for an interface*

Out of 16,000 homodimeric complexes generated from ZDOCK, 149 conformers have passed the membrane topology filtering. Further evaluation of these complexes revealed that 47 of these conformers had their monomers oriented in a parallel arrangement, while the remaining 102 conformers had their monomers with an antiparallel orientation. Although the antiparallel association is considered as a rare and nonphysiological state, it is observed in several crystal packing of GPCRs (Hanson et al., 2008; Li, Edwards, Burghammer, Villa, & Schertler, 2004; Palczewski et al., 2000; Warne et al., 2008, 2011). Next, 149 complexes were separated into two groups based on their orientations, either parallel or antiparallel. Each group was further clustered based on an RMSD value of 6 Å, which produced 20 and 56 clusters for parallel and antiparallel dimers, respectively. As a quantitative assessment for the existence of TM6 at the interface region, $dSASA_i$ values were calculated for

Table 3. Docking results for the long peptide that consists of the whole part of one transmembrane helix.

| | First cluster | | | | | | Most-populated cluster | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Docked Peptide | Size[a] | TM location[b] | AD Score[c] | DSX/CSD score[d] | DSX/PDB score[e] | X-Score[f] | Size[a] | TM location[b] | AD Score[c] | DSX/CSD score[d] | DSX/PDB score[e] | X-Score[f] |
| TM1 | 104 | V, VI | **−9.5** | −95.7 | **−115.2** | **−6.61** | – | – | – | – | – | |
| TM2 | 178 | V, VI, VII | **−9.8** | **−103.6** | −71.8 | **−7.42** | – | – | – | – | – | |
| TM3 | 136 | V, VI, VII | **−10.6** | −86.3 | −106.7 | −5.83 | – | – | – | – | – | |
| TM4 | 162 | V, VI, VII | **−9.8** | −82.0 | −99.2 | −6.06 | – | – | – | – | – | |
| TM5 | 14 | I, V, VI, VII | −8.3 | −79.7 | −95.8 | −4.70 | 52 | V, VI, VII | −7.9 | **−100.3** | **−114.7** | −5.48 |
| TM6 | 62 | VI, VII | −9.1 | **−101.6** | **−121.5** | **−6.30** | 73 | V, VI, VII | −8.8 | −90.7 | −111.0 | −5.79 |
| TM7 | 9 | VI, VII | −9.2 | −78.2 | −77.9 | −5.85 | 58 | V, VI, VII | −8.8 | −97.4 | −110.2 | −6.01 |

Notes: Four scoring functions ranked in the top three in all 10 clusters (7 *first* and 3 *most-populated*), were typed in bold and underlined. See footnotes in Table 2.

all residues in every representative complex structure that was selected for each 76 clusters. Figure 4(b) and (d) illustrate two of these $dSASA_i$ profiles that represent the most-populated clusters obtained for parallel and antiparallel dimers, respectively. Accordingly, a TM region was considered to be at the interface, if any one of its residue's $dSASA_i$ was greater than 0. For parallel dimers, both monomers have their TM6 at the interface, whereas, only one of the monomer's TM6 region exists at the interface for antiparallel dimers, in their *most-populated* cluster.

Furthermore, based on the presence of TM6 at the interface evaluated by $dSASA_i$ values, each cluster was assigned to one of three states colored by red, green, or blue, as illustrated in Figure 5(a) and (b) where each vertical bar indicates the existence of TM6 at the interface as well as the number of elements in that cluster. According to the color code, the green clusters would indicate a dimer formation where TM6 regions of both monomers are found at the interface, whereas in the blue clusters, only one of the monomer's TM6 would be present at the interface. Finally, the red clusters would suggest that no TM6 is located at the interface.

In parallel associations, the most-populated cluster with 7 dimers corresponded to the 'green state' where both TM6 were detected at the interface. Furthermore, in 11 out of 20 clusters, TM6 was not observed at the interface. On the other hand, the total number of conformers in the remaining 9 clusters in which at least one TM6 is present is 26, which is greater than the total number of conformers in 11 clusters, which is determined as 21.

In antiparallel associations, the most-populated cluster had 8 dimers and was colored in blue suggesting the presence of one of the monomer's TM6 at the interface. There exist a total of 26 clusters with 48 dimers in which one TM6 exist at the interface. About 18 dimers found in 9 clusters have both TM6 at the interface,

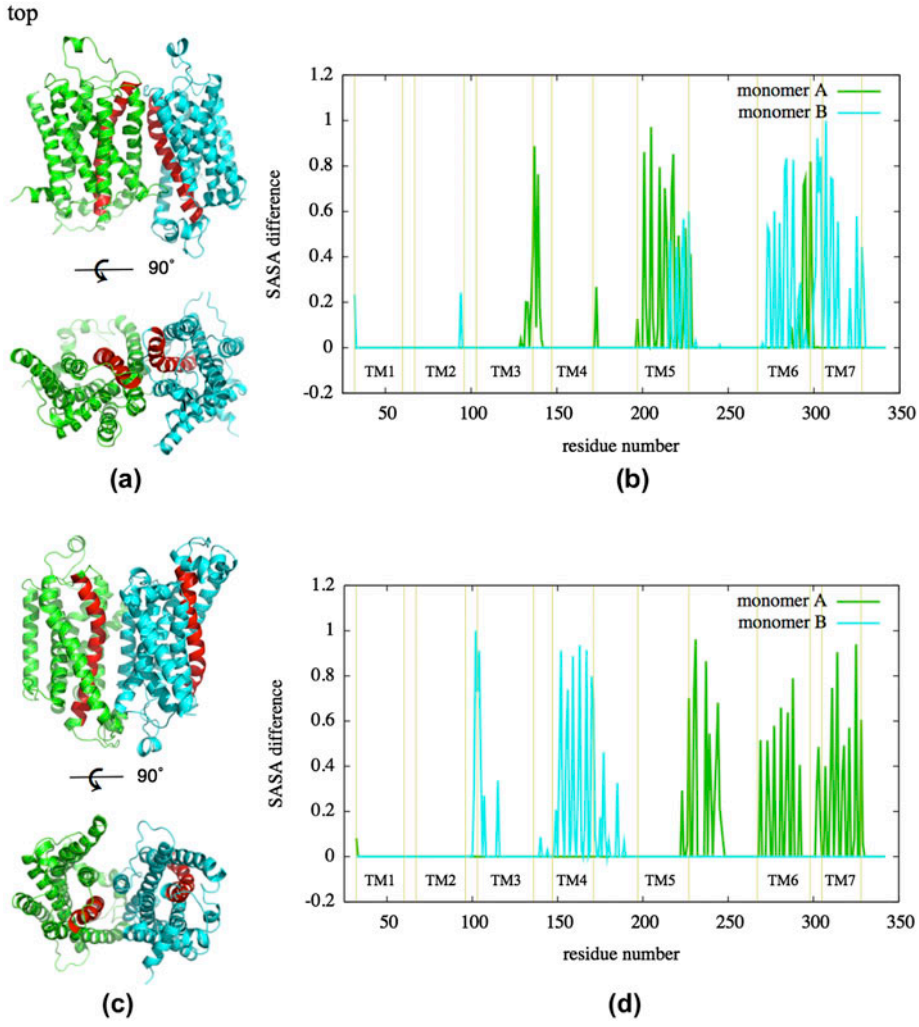


Figure 4. Representative snapshots of the *most-populated* cluster for (a) parallel and (c) antiparallel dimers and their $dSASA_i$ profiles illustrated in (b) and (d), respectively.

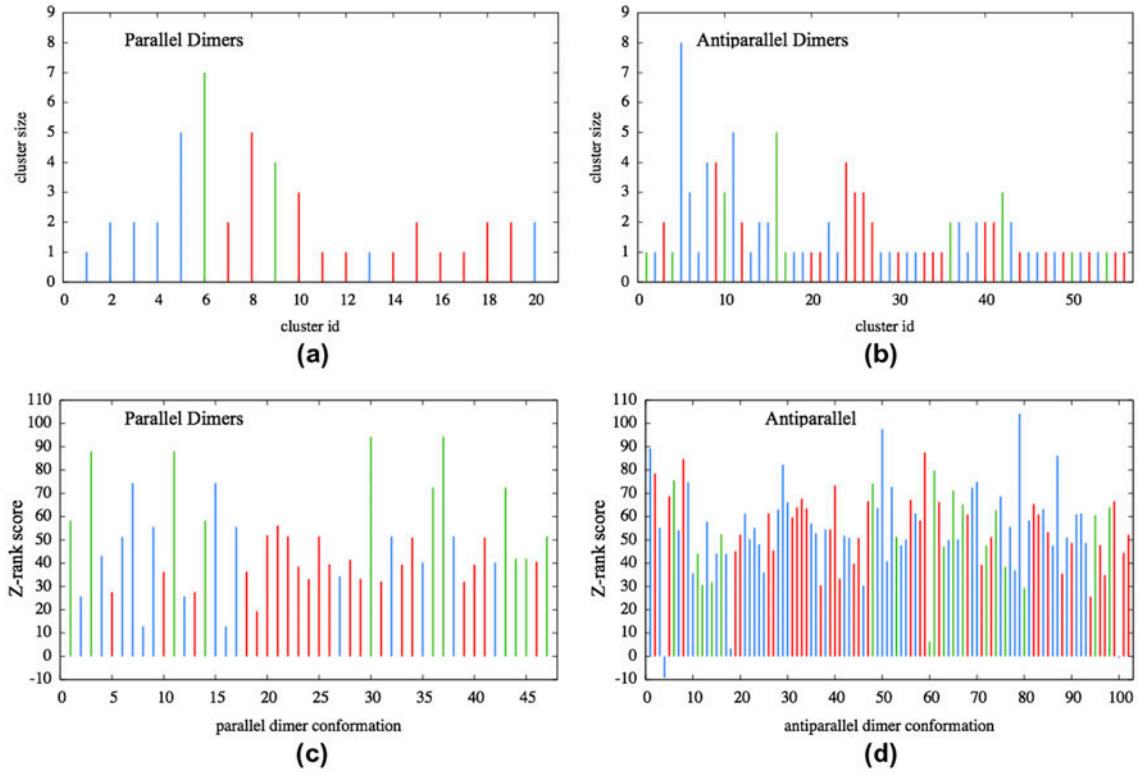


Figure 5. Cluster profile and Z-rank score values for (a), (c) parallel and (b), (d) antiparallel dimers colored based on the presence of TM6. Green: both monomers' TM6 at the interface, Blue: one of the monomer's TM6 at the interface, Red: no TM6 at the interface.

whereas 21 clusters with 36 dimers did not have a TM6 region at the interface. Overall, the number of dimers with at least one TM6 at the interface is greater than the number of dimers with no TM6, in both parallel and antiparallel associations.

Finally, each 47 parallel and 102 antiparallel dimer conformations were reevaluated using their Z-rank score value, which is a positive number for plausible protein–protein associations. As illustrated in Figure 5(c) and (d), each conformer's Z-rank score is designated with a vertical bar colored based on the presence of TM6 at the interface as in Figure 5(a) and (b). It is noteworthy that for parallel dimers, the conformers having two TM6 regions at the interface (green cluster) also had significantly high score values on the average. For parallel dimers, the average Z-rank score values for the green, blue, and red clusters are determined as 69.3 ± 20.1, 43.3 ± 18.9, and 39.5 ± 9.7, respectively. On the other hand, the average Z-rank values for antiparallel dimers are calculated as 51.8 ± 19.5, 54.8 ± 21.4, and 55.7 ± 14.9 for the green, blue, and red clusters, respectively. Clearly, all three values are found to be close to each other and near 50. Thus, the antiparallel dimers are not as distinguishable as parallel dimers with respect to their score values.

## 4. Conclusions

Extensive computational and experimental studies have been conducted to reveal the interface region in $\beta_2$AR dimers and oligomers (Gouldson et al., 1997, 2001; Gouldson & Reynolds, 1997; Nilsson et al., 1999; Oliveira et al., 1993; Pazos et al., 1997). The majority of these studies suggest TM6 as the most plausible interface region. However, there is still no direct evidence to support this implication. Inspired by the peptide experiments conducted by Hebert and his coworkers (Hebert et al., 1996), we carried out docking calculations using peptides that were generated from different portions of the transmembrane region in the receptor. First, the very same peptide of 23 residues in length used in Hebert et al.'s peptide experiments was docked to $\beta_2$AR monomer using a blind docking approach. Here, the peptide was free to bind anywhere in the transmembrane region, except the intracellular and extracellular loop regions. In agreement with the previous findings, the results of 200 distinct docking runs showed that the peptide preferred to bind to TM5 and TM6 regions with considerably high AD, DSX/CSD, DSX/PDB, and X-Score score values in both the *first* and the *most-populated* clusters.

Furthermore, the same peptide was docked 200 times while it was confined to various transmembrane regions

of the receptor. Especially, when the grid box was covering the TM6 region, almost 99% of all docked poses preferred to bind to TM5 and TM6 regions with relatively high score values. Interestingly, when the grid box did not incorporate TM6, then all the score values and the number of conformations in the first cluster reduced significantly. Overall, the peptide strongly preferred TM6 region almost all the time when TM6 was inside the grid box, and interacted with other TM sites with less affinity and less frequently when TM6 was left outside the grid box.

Along with the short peptide docking experiments, each TM helix was turned into a single peptide (relatively longer than the previous peptide), and was docked to $\beta_2$AR monomer using the same blind docking approach. In agreement with the previous docking results, all seven TM-derived peptides preferred TM6 regions persistently with relatively high score values. Especially, it is noteworthy that the best pose of TM6-derived peptide interacted specifically with TM6 and TM7, and with a DSX/PDB score value of −121.5, which was the highest negative one observed in all peptide docking experiments. Overall, TM6 region was the most preferred site in all peptide dockings, and this might indicate TM6 as a possible contact point when another $\beta_2$AR monomer was nearby.

In the second part of our study, we focused on protein–protein dockings where two copies of $\beta_2$AR monomer were rigidly docked to each other. Exceptionally, the analysis of 16,000 different dimer structures provided by these docking runs revealed TM6 at the interface region of the dimer conformations having the highest Z-rank score values and in the most populated clusters. In addition, both parallel and antiparallel orientations have been observed in these dimer structures where TM6 was persistently seen at the interface. The average Z-rank score for parallel dimer conformations was calculated as 69.3 ± 20.1, which was the highest average value observed in our docking results and remarkably, in these conformers, the TM6 region of one monomer was found to be near the TM6 of the second monomer at the interface.

From another perspective, the TM6 in $\beta_2$AR is known to play a key role in the receptor activation, which is characterized by an outward shift of the intracellular end of helix 6. Therefore, this outward motion might also be involved in the dimerization by acting as a reaching branch to the partner monomer. Consequently, the potential involvement of TM6 in dimerization might increase the likelihood of having TM6 at the interface region, appearing at least in one monomer. Our protein–protein docking results strongly support the presence of at least one TM6 at the interface. In the near future, the predicted model of $\beta_2$AR dimer which has the highest Z-rank score and the two TM6 at the interface obtained from our study will be used as a starting conformation to investigate the intrinsic dynamics of the dimer structure via molecular dynamics simulation.

## Supplementary material

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2014.962094.

## Funding

This work has been partially supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant # 109M281.

## References

Allen, F. H. (2002). The Cambridge Structural Database: A quarter of a million crystal structures and rising. *Acta Crystallographica Section B: Structural Science, 58*, 380–388.

Angers, S., Salahpour, A., Joly, E., Hilairet, S., Chelsky, D., Dennis, M., & Bouvier, M. (2000). Detection of beta 2-adrenergic receptor dimerization in living cells using bioluminescence resonance energy transfer (BRET). *Proceedings of the National academy of Sciences of the United States of America, 97*, 3684–3689.

Berman, H. M. (2002). The Protein Data Bank. *Acta Crystallographica Section D Biological Crystallography, 58*, 899–907.

Bouvier, M. (2001). Oligomerization of G-protein-coupled transmitter receptors. *Nature Reviews Neuroscience, 2*, 274–286.

Chen, R., Li, L., & Weng, Z. (2003). ZDOCK: An initial-stage protein-docking algorithm. *Proteins: Structure, Function, and Genetics, 52*, 80–87.

Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., … Stevens, R. C. (2007). High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. *Science, 318*, 1258–1265.

Ciruela, F., Casado, V., Mallol, J., Canela, E. I., Lluis, C., & Franco, R. (1995). Immunological identification of A1 adenosine receptors in brain cortex. *Journal of Neuroscience Research, 42*, 818–828.

Dean, M. K., Higgs, C., Smith, R. E., Bywater, R. P., Snell, C. R., Scott, P. D., … Reynolds, C. A. (2001). Dimerization of G-protein-coupled receptors. *Journal of Medicinal Chemistry, 44*, 4595–4614.

Dorsch, S., Klotz, K. N., Engelhardt, S., Lohse, M. J., & Bünemann, M. (2009). Analysis of receptor oligomerization by FRAP microscopy. *Nature Methods, 6*, 225–230.

Feig, M., Karanicolas, J., & Brooks, C. L. (2001). *MMTSB Tool Set (2001)*. San Diego, CA: MMTSB NIH Research Resource, The Scripps Research Institute.

Filizola, M., & Weinstein, H. (2005). The study of G-protein coupled receptor oligomerization with computational modeling and bioinformatics. *FEBS Journal, 272*, 2926–2938.

Fung, B. K., Hurley, J. B., & Stryer, L. (1981). Flow of information in the light-triggered cyclic nucleotide cascade of vision. *Proceedings of the National Academy of Sciences, 78*, 152–156.

Fung, J. J., Deupi, X., Pardo, L., Yao, X. J., Velez-Ruiz, G. A., DeVree, B. T., … Kobilka, B. K. (2009). Ligand-regulated oligomerization of β2-adrenoceptors in a model lipid bilayer. *EMBO Journal, 28*, 3315–3328.

Gkoutos, G. V., Higgs, C., Bywater, R. P., Gouldson, P. R., & Reynolds, C. A. (1999). Evidence for dimerization in the β2-adrenergic receptor from the evolutionary trace method. *International Journal of Quantum Chemistry, 74*, 371–379.

Gouldson, P. R., Bywater, R. P., & Reynolds, C. A. (1997). Correlated mutations amongst the external residues of G protein-coupled receptors. *Biochemical Society Transactions, 25*, 529S.

Gouldson, P. R., Dean, M. K., Snell, C. R., Bywater, R. P., Gkoutos, G., & Reynolds, C. A. (2001). Lipid-facing correlated mutations and dimerization in G-protein coupled receptors. *Protein Engineering Design and Selection, 14*, 759–767.

Gouldson, P. R., & Reynolds, C. A. (1997). Simulations on dimeric peptides: Evidence for domain swapping in G protein-coupled receptors? *Biochemical Society Transactions, 25*, 1066–1071.

Gouldson, P. R., Snell, C. R., Bywater, R. P., Higgs, C., & Reynolds, C. A. (1998). Domain swapping in G protein-coupled receptor dimers. *Protein Engineering Design and Selection, 11*, 1181–1193.

Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V. P., Chien, E. Y., … Stevens, R. C. (2008). A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure, 16*, 897–905.

Hebert, T. E., Moffett, S., Morello, J. P., Loisel, T. P., Bichet, D. G., Barret, C., & Bouvier, M. (1996). A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *Journal of Biological Chemistry, 271*, 16384–16392.

Hetenyi, C., & Spoel, D. V. D. (2002). Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Science, 11*, 1737–1737.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics, 14*, 33–38.

Jois, S. D. S., & Siahaan, T. J. (2003). A peptide derived from LFA-1 protein that modulates T-cell adhesion binds to soluble ICAM-1 protein. *Journal of Biomolecular Structure and Dynamics, 20*, 635–644.

Lavoie, C., Mercier, J. F., Salahpour, A., Umapathy, D., Breit, A., Villeneuve, L. R., … Hebert, T. E. (2002). β1/β2-Adrenergic receptor heterodimerization Regulates β2-Adrenergic receptor internalization and *ERK* Signaling Efficacy. *Journal of Biological Chemistry, 277*, 35402–35410.

Li, J., Edwards, P. C., Burghammer, M., Villa, C., & Schertler, G. F. (2004). Structure of bovine rhodopsin in a trigonal crystal form. *Journal of Molecular Biology, 343*, 1409–1438.

Liang, Y., Fotiadis, D., Filipek, S., Saperstein, D. A., Palczewski, K., & Engel, A. (2003). Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *Journal of Biological Chemistry, 278*, 21655–21662.

Liu, R., Perez-Aguilar, J. M., Liang, D., & Saven, J. G. (2012). Binding site and affinity prediction of general anesthetics to protein targets using docking. *Anesthesia and Analgesia, 114*, 947–955.

Livingstone, C. D., & Barton, G. J. (1993). Protein-sequence alignments: A strategy for the hierarchical analysis of residue conservation. *Computer Applications in the Biosciences, 9*, 745–756.

Mercier, J. F., Salahpour, A., Angers, S., Breit, A., & Bouvier, M. (2002). Quantitative assessment of beta 1- and beta 2-adrenergic receptor homo- and heterodimerization by bioluminescence resonance energy transfer. *Journal of Biological Chemistry, 277*, 44925–44931.

Milligan, G. (2007). G protein-coupled receptor dimerisation: Molecular basis and relevance to function. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1768*, 825–835.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry, 19*, 1639–1662.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry, 30*, 2785–2791.

Narayanan, E., John, B., Mirkovic, N., Fiser, A., Ilyin, V. A., Pieper, U., … Sali, A. (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research, 31*, 3375–3380.

Nemoto, W., & Toh, H. (2005). Prediction of interfaces for oligomerizations of G-protein coupled receptors. *Proteins, 58*, 644–660.

Neudert, G., & Klebe, G. (2011). DSX: A knowledge-based scoring function for the assessment of protein-ligand complexes. *Journal of Chemical Information and Modeling, 51*, 2731–2745.

Nilsson, A. M., Wijaywardene, M., Gkoutos, G., Wilson, K. M., Fernandez, N., & Reynolds, C. A. (1999). Correlated mutations in the HLA class II molecule. *International Journal of Quantum Chemistry, 73*, 85–96.

Oliveira, L., Paiva, A. C. M., & Vriend, G. (1993). A common motif in G protein-coupled seven transmembrane helix receptors. *Journal of Computer-Aided Molecular Design, 7*, 649–658.

Ozcan, O., Uyar, A., Doruker, P., & Akten, E. D. (2013). Effect of intracellular loop 3 on intrinsic dynamics of human β2-adrenergic receptor. *BMC Structural Biology, 13*(29), 1–16.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., … Stenkamp, R. E. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science, 289*, 739–745.

Pazos, F., Helmer-Citterich, M., Ausiello, G., & Valencia, A. (1997). Protein docking with sequences. Correlated mutations contain information about protein-protein interaction. *Journal of Molecular Biology, 271*, 511–523.

Pierce, B., & Weng, Z. (2007). ZRANK: Reranking protein docking predictions with an optimized energy function. *Proteins: Structure, Function, and Bioinformatics, 67*, 1078–1086.

Salahpour, A., Angers, S., Mercier, J. F., Lagace, M., Marullo, S., & Bouvier, M. (2004). Homomerization of the beta 2-adrenergic receptor as a pre-requisite for cell surface targeting. *Journal of Biological Chemistry, 279*, 33390–33397.

Song, S. X., Liu, D., Peng, J. L., Deng, H. W., Guo, Y., Xu, L. X., … Xu, Y. H. (2009). Novel peptide ligand directs liposomes toward EGF-R high-expressing cancer cells *in vitro* and *in vivo*. *The FASEB Journal, 23*, 1396–1404.

Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., & Dixon, R. A. (1994). Structure and function of G protein-coupled receptors. *Annual Review of Biochemistry, 63*, 101–132.

Wang, R., Lai, L., & Wang, S. (2002). Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design, 16*, 11–26.

Warne, T., Moukhametzianov, R., Baker, J. G., Nehmé, R., Edwards, P. C., Leslie, A. G. W., … Tate, C. G. (2011). The structural basis for agonist and partial agonist action on a β1-adrenergic receptor. *Nature, 469*, 241–244.

Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., … Schertler, G. F. (2008). Structure of a β1-adrenergic G-protein-coupled receptor. *Nature, 454*, 486–491.

Whorton, M. R., Bokoch, M. P., Rasmussen, S. G., Huang, B., Zare, R. N., Kobilka, B., & Sunahara, R. K. (2007). A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proceedings of the National Academy of Sciences, 104*, 7682–7687.