Journal of Molecular Graphics and Modelling 66 (2016) 115–122

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

# Homology modeling and molecular docking studies of *Drosophila* and *Aedes* sex peptide receptors

Jeong-hyun Kim[a,1], Soo-Kyung Kim[b,1], Jae-Hyuk Lee[a], Young-Joon Kim[a], William A. Goddard III[b,*], Yong-Chul Kim[a,*]

[a] School of Life Sciences, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Republic of Korea
[b] Materials and Process Simulation Center (MC139-74), California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, United States

ARTICLE INFO

*Article history:*
Received 21 October 2015
Received in revised form 7 March 2016
Accepted 24 March 2016
Available online 25 March 2016

*Keywords:*
*Drosophila* sex peptide receptor
G protein-coupled receptor
GPCR modeling
Docking
Binding site prediction
BiHelix/SuperBiHelix
GEnSeMBLE
Pyrazolodiazepine

ABSTRACT

The *Drosophila melanogaster* sex peptide receptor (*Drm*SPR), which is a G protein-coupled receptor (GPCR), is known as the specific receptor for sex peptide (SP). It is responsible for the reproductive behavior in the *Drosophila* model system; in particular, it is involved in the post-mating responses such as the increase in egg-laying ability and decrease in receptivity in females. In a previous study, we discovered a small molecule agonist of *Drm*SPR for the first time, which could not, however, activate *Aedes aegypti* SPR (*Aedes*SPR). To investigate the binding mechanism of the small molecule agonist of *Drm*SPR, the ensemble structures of low-lying packing structures of *Drm*SPR and *Aedes*SPR were assembled using the GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) method. The generated homology models exhibited the typical pattern of inter-helical interactions of the class A GPCRs. The docking experiments of the small molecule agonist suggest that $Tyr^{5.35}$ and $Phe^{2.67}$ residues may be involved in a hydrophobic interaction and that $Ser^{3.25}$ forms a hydrogen bond with the agonist. Additionally, we found that the docking results were consistent with the experimental data of the reference compounds with variable agonistic activities. Moreover, a potential distinction of the putative binding sites in two GPCR models of *Drm*SPR and *Aedes*SPR, which was determined in this study, can explain the selective action of the agonist for *Drm*SPR but not for *Aedes*SPR.



## 1. Introduction

The *Drosophila melanogaster* sex peptide receptor (*Drm*SPR), which belongs to the class A G protein-coupled receptor (GPCR) super family, has been reported as a specific receptor for sex peptide (SP) as a natural ligand. SP is responsible for the reproductive behavior of *Drosophila* model systems [1–5]. The activation of *Drm*SPR by the peptide ligand SP, which is transferred from the seminal fluid of a male organ, results in post-mating responses (PMRs) such as the suppression of mating receptivity and increase of the egg-laying ability in female drosophila [6–14].

Sex peptide receptor (SPR) is detected in many species, such as lophotrochozoa, ecdysozoa, *Aedes aegypti* (mosquito), *Bombyx mori* (moth) and *Aplysia californica* (sea slug). Interestingly, most of these species have their own genes encoding SP-like peptides, which usually interact with only the SPRs in the same species and not in other species [4]. For example, the *Drosophila* SP activates only *Drm*SPR, but it does not activate mosquito *Aedes*SPR, although the sequence identity of the transmembrane (TM) domains between the two species is as high as 68%.

Despite the importance of SPR and SP in *Drosophila* research, no report has been published regarding the structural insights of SPR and the binding mechanism of SP, mainly because of the limited resources of alternative ligands such as small molecule agonists. However, our group recently discovered small molecule agonists. Our new agonist selectively binds to *Drm*SPR with an $EC_{50}$ value of 3 μM but does not activate *Aedes*SPR at all (Supplementary information Fig. S1) [15]. This result suggests that specific residues may exist for the agonist activation and species selectivity of SPR.

Because the 3-D X-ray structures of SPRs have not been determined, homology modeling of SPRs is a good starting point to generate the 3D structure of SPR. Currently, 26 experimental structures of GPCRs have been reported, including that of bacteriorhodopsin, using X-ray crystallography and electron diffraction [16]. Although the information of solved 3D structures of GPCRs are not good enough to study the molecular mechanism and applied to

* Corresponding author.
*E-mail addresses:* wag@wag.caltech.edu (W.A. Goddard III), yongchul@gist.ac.kr (Y.-C. Kim).
[1] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.jmgm.2016.03.014
1093-3263/

drug discovery, in silico homology modeling has become an alternative solution by aid of advances of technologies to challenge the protein modeling research [17–19].

To provide accurate structural information, we report an ensemble of stable 3D structures that were predicted using the GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) Complete Sample Monte Carlo (CS-MC) method for *Drm*SPR and *Aedes*SPR. This GEnSeMBLE CS-MC technique has been validated to predict structures with close to X-ray accuracy [20–23]. Here, we report the predicted 3D structures for *Drm*SPR and *Aedes*SPR and find a specific interaction that may provide the subtype selectivity for *Drosophila* over *mosquito*. This study will provide new insights into the species selectivity between *Drm*SPR and *Aedes*SPR in SPR. These detailed atomistic models will provide a number of notably specific suggestions for experimental validation, in which various specific mutations are applied and various other ligands are tested, including ones that are designed based on the structural models.

## 2. Methods

### 2.1. PredicTM

PredicTM was used to perform a multiple sequence alignment of 1726 GPCR sequences using the MAFFT program and to predict the TM domains for the target GPCR [24]. These TM domains were extended using capping rules, and the secondary structure was predicted using PORTER, PSIPRED, JPRED, MINNOU and TMHMM [25–29].

### 2.2. Helix generation

To predict the shape of the TM domains, each TM region was initially generated using OptHelix, which calculates helical conformations by the sequences, resulting in inappropriate structures (data not shown). Thus, the 12 initial helix bundle structures were generated from X-ray structures using Homologize Helices method [20]. We considered as templates the X-ray structures for human adenosine $A_{2A}$ receptor (*h*$AA_{2A}R$), human Chemokine $CXCR_4$ (*h*$CXCR_4$), human sphingosine 1 phosphate receptor 1 (*h*S1PR1), *Meleagris gallopavo* Beta-1 adrenergic receptor (Melga_$ADRB_1$), human $D_3$ dopamine receptor (*h*$D_3$DR), mouse mu opioid receptor (*m*OPRM), human Beta-2 adrenergic receptor (*hb*$_2$AR), *bovine* Rhodopsin (*b*OPSD), human Histamine H1 receptor (*h*HRH1), rat Muscarinic acetylcholine receptor M3 (*r*$ACM_3$) and human Muscarinic acetylcholine receptor M2 (*r*$ACM_2$) [30–40].

### 2.3. BiHelix optimization of helix rotations

Because the sequence identity in the TM region (22–36%) was high, we sampled a 60° rotation about each TM axis with 15° increments. The BiHelix method, which uses the interactions within the seven-helix bundle by partitioning the interactions into 12 sets of pairwise helix interactions, was used to evaluate the energies practical for all samples. For each pairwise interaction, we used SCREAM to optimize the side chains. The best 2000 structures were selected in this step. Then, we built the full helix bundle for each of these 2000 structures, optimized the side chains for each structure using SCREAM, and neutralized the charged residues for more accurate energy scoring.

### 2.4. SuperBiHelix

For the lowest energy predicted structures from the BiHelix step, the optimal rotation angles ($\theta$, $\varphi$, $\eta$) were determined for the packing of the seven helices. For the $\theta$ tilt angle and $\varphi$ and $\eta$ angles, $\alpha \pm 10°$, $\alpha \pm 30°$ and $\alpha \pm 30°$ were considered, respectively.

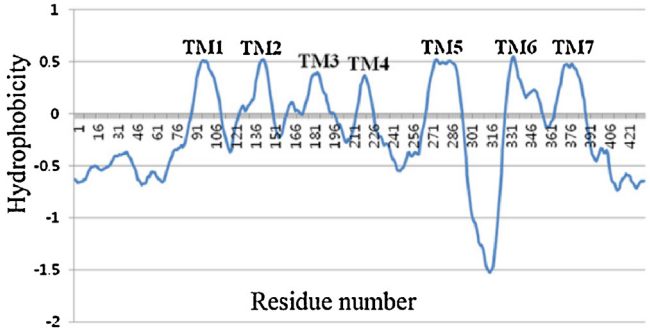


**Fig. 1.** The transmembrane region of *Drm*SPR was predicted using the hydropathy plot.

The total energies for each of these $12 \times 5625$ helix pair combinations were used to estimate the energy for all 7-TM helix bundle conformational combinations. Finally, the top 1000 best energy predicted structures were selected.

### 2.5. Pocket prediction and ligand docking: CDOCKER

The pocket prediction and molecular docking procedure was performed using CDOCKER, which was a docking program for rigid receptor and flexible ligands, interfaced with Accelrys DiscoveryStudio3.5 [41]. For the receptor preparation, the active site for docking was generated using the cavity search tools from the predicted 3D structure of *Drm*SPR, and the CHARMm force field was applied. For the ligand preparation, the 3D structures of the compounds were generated using DiscoveryStudio3.5, and minimized by the smart minimizer algorithm in minimization protocol using CHARMm force field. From the docking study of the compounds, the top 20 hit conformations were generated by sorting the values of CDOCKER energy and CDOCKER interaction energy using CDOCKER protocol in DiscoveryStudio3.5.

## 3. Results and discussion

### 3.1. Structure predictions of DrmSPR and AedesSPR

#### 3.1.1. TM predictions

The seven TM regions of *D. melanogaster* (*Drm*SPR) were first predicted by hydropathy analysis (Fig. 1) and multiple sequence alignments. To verify the sequence alignment results, eta residues (the most conserved residues in each TM, $N^{1.50}$, $D^{2.50}$, $R^{3.50}$ $W^{4.50}$, $P^{5.50}$, $P^{6.50}$, and $P^{7.50}$) were compared with the standard eta residues of general class A GPCRs. As shown in Table 1, all eta residues of *Drm*SPR were well matched, except with the standard residues $Trp^{4.50}$ of TM4 and $Tyr^{5.58}$ of TM5. Although the conserved $W^{4.50}$ and $Y^{5.58}$ in class A GPCRs are not conserved in SPR, $A^{4.50}$ and $N^{5.58}$ are conserved in SPR [42]. Thus, the TM prediction analysis in this study was acceptable. In SPR species, there are some discrepancies with conserved motif in class A GPCRs.

The sequence alignment shows the QRY motif in TM3 instead of the D/ERY motif in class A GPCRs. In addition, AExP in TM6 and NFxxY motifs in TM7 were replaced by CWxP and NPxxY, respectively, in the class A family. The final $\alpha$-helix region was determined from PredicTM. Other web-based TM prediction servers (PORTER, PSIPRED, JPRED, MINNOU and TMHMM) were used to extend the secondary structure of *Drm*SPR at the terminal end of each TM (Fig. 2).