Journal of Molecular Graphics and Modelling 129 (2024) 108725

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

The full-length TSH receptor is stabilized by TSH ligand

Mihaly Mezei [a,b,*], Rauf Latif [b,c], Terry F. Davies [b,c]

[a] Department of Pharmacological Sciences, New York, NY, USA
[b] Thyroid Research Unit, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[c] James J. Peters VA Medical Center, New York, NY, USA

ARTICLE INFO

Keywords:
TSHR
Molecular dynamics
Simulation
Transmembrane domain
Hydrogen bond
Hinge region
Linker region
TSH

ABSTRACT

The receptor for thyroid stimulating hormone (TSHR), a GPCR, is the primary antigen in autoimmune hyperthyroidism (Graves' disease) caused by stimulating TSHR antibodies. While we have previously published a full length model of the TSHR, including its leucine rich domain (LRD), linker region (LR) and transmembrane domain (TMD), to date, only a partial LRD (aa 21–261) stabilized with TSHR autoantibodies has been crystallized. Recently, however, cryo-EM structures of the full-length TSHR have been published but they include only an incomplete LR. We have now utilized the cryo-EM models, added disulfide bonds to the LR and performed longer (3000 ns) molecular dynamic (MD) simulations to update our previous model of the entire full-length TSHR, with and without the presence of TSH ligand.

As in our earlier work, the new model was embedded in a lipid membrane and was solvated with water and counterions. We found that the 3000 ns Molecular Dynamic simulations showed that the structure of the LRD and TMD were remarkably constant while the LR, known more commonly as the "hinge region", again showed significant flexibility, forming several transient secondary structural elements. Analysis of the new simulations permitted a detailed examination of the effect of TSH binding on the structure of the TSHR. We found a structure-stabilizing effect of TSH, including increased stability of the LR, which was clearly demonstrated by analyzing several intrinsic receptor properties including hydrogen bonding, fluctuation of the LRD orientation, and radius of gyration.

In conclusion, we were able to quantify the flexibility of the TSHR and show its increased stability after TSH binding. These data indicated the important role of ligands in directing the signaling structure of a receptor.

1. Introduction

The thyroid-stimulating hormone receptor (TSHR) is the major factor for thyroid growth and development and the major regulator of thyroid hormone synthesis and secretion. We now know that the TSHR is also the target of autoantibodies in Graves' hyperthyroid disease and thyroid eye disease [1]. The TSHR has a large ectodomain (ECD) and membrane-bound signal transducing transmembrane domain (TMD) [2,3]. The ECD can be divided into a curved leucine rich domain (LRD) attached to the TMD by a 130 amino acid (AA) linker region (LR) more commonly called the "hinge region" (AA280-410). The TSHR has a large 50 amino acid cleavage region (AA316-366) within the LR that is cleaved from the ectodomain which then attaches to the TMD via 3 cysteine bonds [4,5].

We recently generated a full length model of this complex TSHR structure [6] that combined, using a heuristic procedure, the extracellular part of the TSHR generated with the AI-based Alphafold2 [7] program together with our recently reported model of the TSHR TMD [8]. After a 1000 ns Molecular Dynamics (MD) simulation we were able to conclude that the LR was an intrinsically disordered protein (IDP) explaining its difficulty to crystallize. We also generated a model for the TSHR with TSH ligand bound to its ectodomain, which suggested that introducing TSH into the system might have induced more stability to the LR.

Shortly after the completion of our model, cryo-EM structures involving the full-length TSHR were reported [9–11]. A comparison of our model with the cryo-EM structures showed that our extracellular domain was rotated abnormally by about 180°. Therefore, we generated a new version of the TSHR, guided by the cryo-EM data, after combining the Alphafold2 model for the LRD and LR with our model of the TMD. In addition, we then obtained an analogous new model for the TSHR-TSH complex.

The most important conclusion from our earlier simulation of the

* Corresponding author. Icahn School of Medicine at Mount Sinai, New York, NY, USA.
E-mail address: mihaly.mezei@mssm.edu (M. Mezei).

https://doi.org/10.1016/j.jmgm.2024.108725
Received 15 September 2023; Received in revised form 24 January 2024; Accepted 8 February 2024
Available online 11 February 2024

Abbreviations

AA	amino acid
cryo-EM	cryo electron microscopy
DPPC	Dipalmitoylphosphatidylcholine
ECD	ectodomain
GCE	grand-canonical ensemble
GPCR	G protein coupled receptor
LR	linker region
LRD	leucine-rich domain
MD	molecular dynamics
SSE	secondary structure element
TMD	trans-membrane domain
TSH	thyroid stimulating hormone
TSHR	TSH receptor

full-length TSHR was that the LR was an intrinsically disordered, protein. One of the major aims of the present work, in addition to correcting the orientation of the ECD was, therefore, to test the hypothesis that TSH ligand would indeed stabilize the disordered LR structure. We compared the simulated trajectories (3000 ns long each) of the TSHR and TSHR-TSH complex under identical conditions and quantified the effect of adding TSH to the TSHR, which we found to result in improved stability as hypothesized.

2. Results

2.1. The new full-length TSHR model

The initial structure of the TSHR as assembled from the Alphafold2 derivative, the cryo-EM available data and our TMD model (called TRIO) [8] and embedded in a DPCC lipid bilayer membrane is shown in Fig. 1. Our previous TSHR model is shown in Fig. 2A and compared with

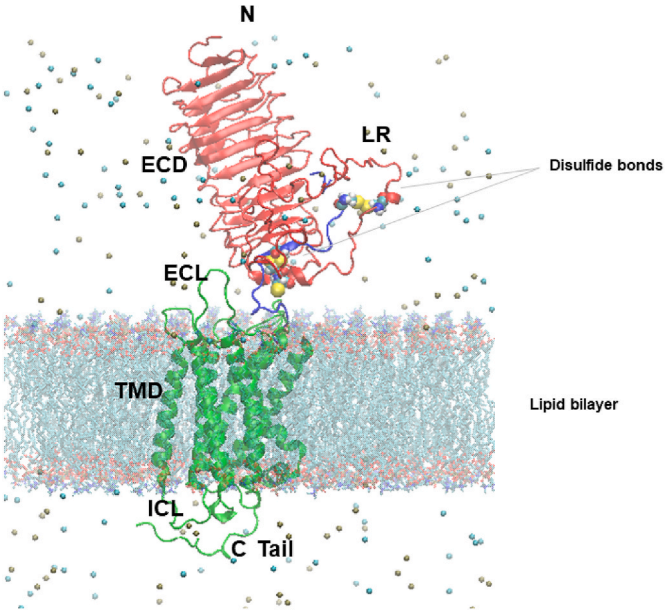

Fig. 1. The structure of the TSHR in the context of the lipid bilayer (transparent sticks) and counterions (spheres). TSHR LRD and LR, residues 24–381 from Alphafold 2: red; TSHR LR residues 382–413 from Alphafold2 for which cryo-EM structure exists: blue; TSHR TMD, residues 414–717: green. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the new corrected version followed by the result of docking TSH to this new model (Fig. 2B and C). While our previous TSHR model was built without the help of the cryo-EM data, these diagrams show the different parts of the new model from Alphafold2 (LRD and most of the LR), cryo-EM (part of the LR) and molecular dynamics (TMD) using different colors and can be compared with the original.

The models of the TSHR and the TSHR-TSH complex, which included Monte Carlo-generated internal waters in the TMD, were sent to the Charmm-gui [12,13] server where they were embedded in a DPCC lipid bilayer and immersed in water with counterions. Charrmm-gui added 178 and 177 DPPC molecules in the upper and lower layer, respectively, to both systems. The total number of waters was 15,074 and 52,043 for the TSHR and TSHR-TSH systems, respectively. 150 K^+ ions were added to both systems while the number of Cl^- ions was 157 and 151 for the TSHR and TSHR-TSH systems, respectively. While the structure was simulated for 3000 ns; some of the analyses were performed only on the last 2000 ns of the trajectory since the structures submitted to the Charmm-gui server to initiate the calculation of the conformations was the same as in the longer simulation. In addition, as our earlier work presented atomic-level modeling of the LRD-TSH complex [8], this report was focused more on describing the changes in the LR and in the TMD upon TSH binding to the TSHR.

2.2. Integrity and fluctuation of the structure

The preliminary results for the TSHR-TSH complex in our earlier work showed that the inclusion of the TSH put a strain on the LR-TMD interface; In fact, we found that even though in our earlier ligand-free run of 1000 ns two of the cysteine pairs in the LR stayed close without forming a disulfide bond. But the addition of TSH then pulled the LR away from the TMD, which was unsatisfactory since without LR-TMD contact there may be no signal transduction. In the current version, and with a three times longer simulation, and now with all three disulfide bonds included in the model, the LR-TMD interface stayed unchanged, as observed both by animating the trajectory (using VMD) and from the hydrogen-bond histories discussed below. Animation of the trajectories also showed that both the LRD and the TMD structures maintained their fold but the LR again displayed large structural variations. However, the relative orientation of the LRD to the TMD also showed significant variations. Fig. 3 show the radius of gyration (R_g, calculated) of the LR in the TSHR and compared to the TSHR-TSH complex. The effect of TSH was a clear reduction in the $R_{g.}$, both in its magnitude and in its fluctuation as indicated by red trace line.

2.3. Structural variations

The 2D RMSD maps of just the LR backbones are shown on Fig. 4 for the last 2000 ns of the simulations with and without TSH. The trajectories are clustered into 5 clusters and representative structures were extracted from each. While the maximum RMSD among the TSHR conformations was 17.7 Å, in the TSHR-TSH system the maximum was only 9.5 Å, showing that the introduction of TSH **resulted in the reduction of the disorder in the LR.** The five representative structures of the LR frozen at different time points of simulation are shown in Fig. 5 for both the TSHR and TSHR-TSH systems.

2.4. Fluctuation of the LRD orientation

For the calculation of the circular variances of the LRD the local coordinate system was defined so that the X–Y plane of the coordinate system is essentially the plane of the leucine-rich beta sheets with the Y axis being roughly parallel with the beta sheets. The circular variances and the normal distribution equivalent standard deviations are given in Table 1 for the last 2000 ns of the simulations. The largest fluctuations are seen around the Y axis. Furthermore, the fluctuations are uniformly smaller in the TSHR-TSH system, providing another indication of the

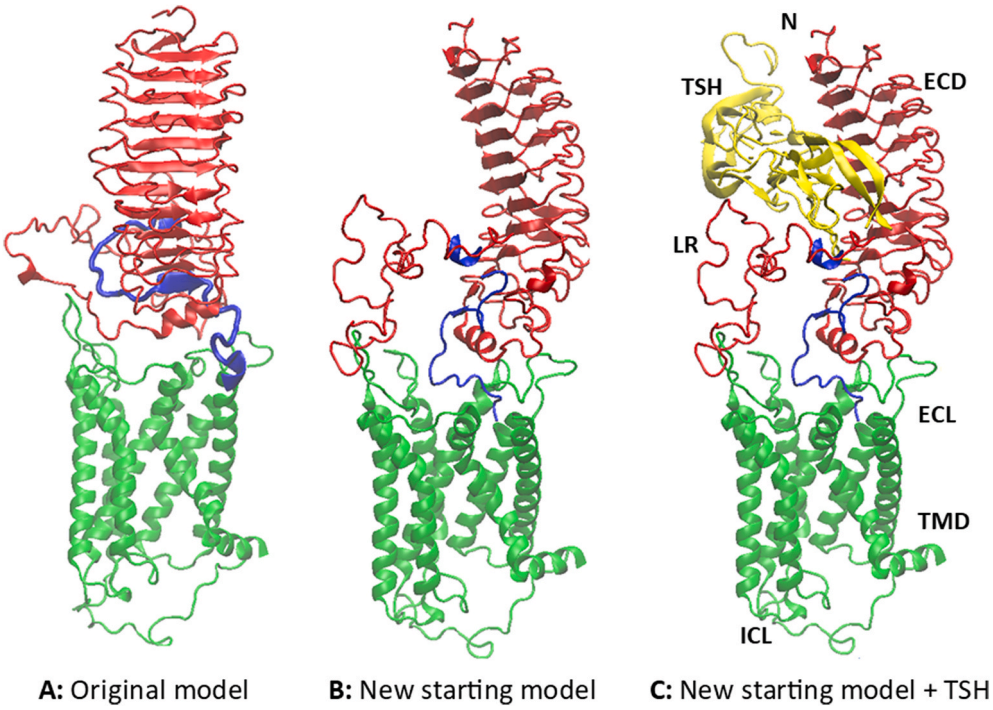

Fig. 2. (A) The structure of our original TSHR model (from 6); (B) The initial structure of the TSHR built with the help of cryo-EM data; (C) The initial structure of the TSHR-TSH complex. TSH: yellow; TSHR LRD and LR, residues 24–381 from Alphafold 2: red; TSHR LR residues 382–413 from cryo-EM: blue; TSHR TMD, residues 414–717: green. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

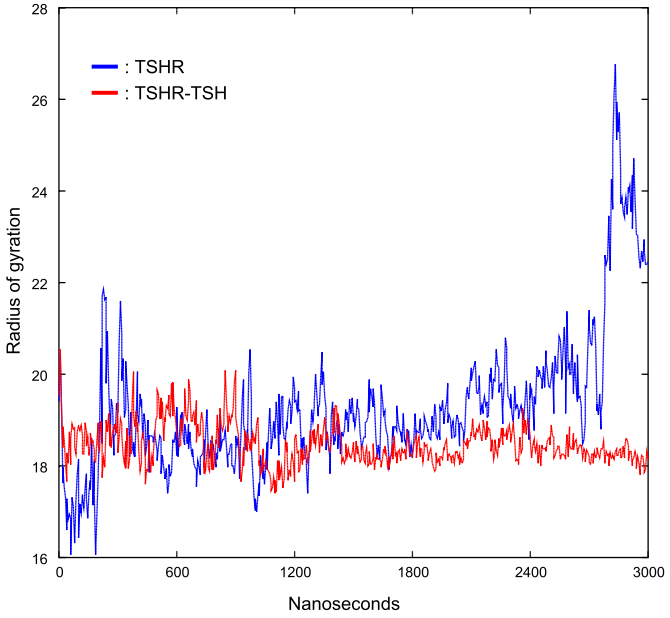

Fig. 3. The radius of gyration (in Å) of the LR during the simulation for the TSHR (blue) and the TSHR-TSH complex (red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

stabilizing effect of TSH on the TSHR. The evolution of the angle between the membrane normal and the first principal axis of the LRD (calculated on the backbone atoms only) is shown on Fig. 6. While both systems show significant fluctuations, the TSHR-TSH complex appears to settle at a somewhat larger angle than the average of the TSHR only system.

2.5. Hydrogen-bond history

The histories of residue pairs that were hydrogen bonded during more than 20 % of the time for the TSHR and for the TSHR-TSH complex are shown in Fig. 7A and B. The residue pairs represented by each line of the plots are given in Supplementary Tables S1 and S2 for the TSHR and the TSHR-TSH complex respectively. The number of similarly filtered hydrogen bonded residue pairs over the last 2000 ns of the simulation, categorized by the domains the residues are in, are shown in Table 2. The number of LR-TMD hydrogen bonds in the TSHR-TSH complex had more than twice the number of hydrogen bonded residue pairs than in the TSHR alone while the number of LR-TMD hydrogen-bonded residue pairs were the same in both systems. The sum of the percentages was 157 for the TSHR and 206 for the TSHR-TSH complex, which showed a strengthening of the interface. Two of the LR residues (LYS290 and ASN 288) were involved in hydrogen bonding with the TMD in both systems while the TMD residues involved were all different suggesting a rearrangement of the LR-TMD interface upon TSH binding and activation. The intra-helix hydrogen bonds were excluded so these plots could be considered to complement the DSSP plots discussed below. These plots also assessed the convergence of the simulations since unconverged runs keep forming new hydrogen bonds.

2.6. DSSP plots

The DSSP plot is a compact representation of the history of secondary structure element formation and unraveling during the simulation. Fig. 8 shows these plots for the TSHR and the TSHR-TSH complex. The DSSP plot of the TSHR (Fig. 8A) showed that the residue range 345–408 (outlined) is an intrinsically disordered protein. Comparison of the two plots shows that, as suggested in our earlier work, the introduction of TSH induced a significant amount of structure in this region.

2.7. Domain-domain distances

The distances between the three domains of the TSHR (LRD, LR and

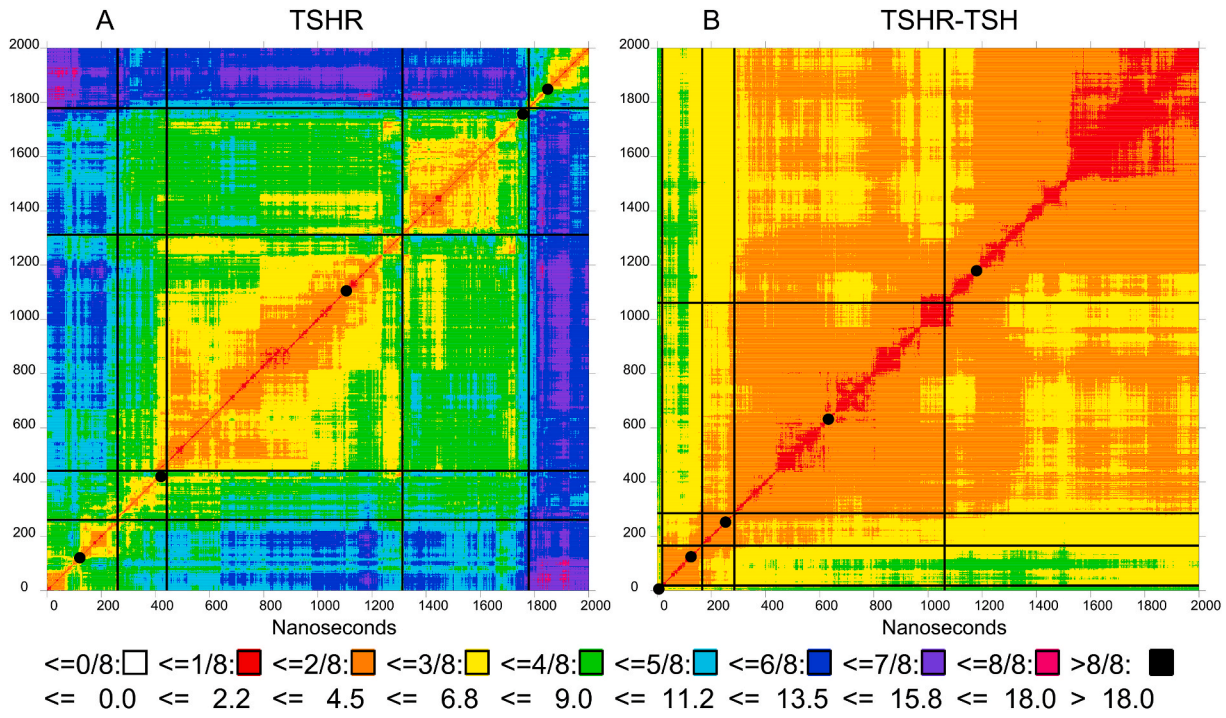

Fig. 4. The 2D RMSD maps of the LR (without the N-terminal tail and using the backbones only) during the simulation for the TSHR (4A) and the TSHR-TSH complex (4B).

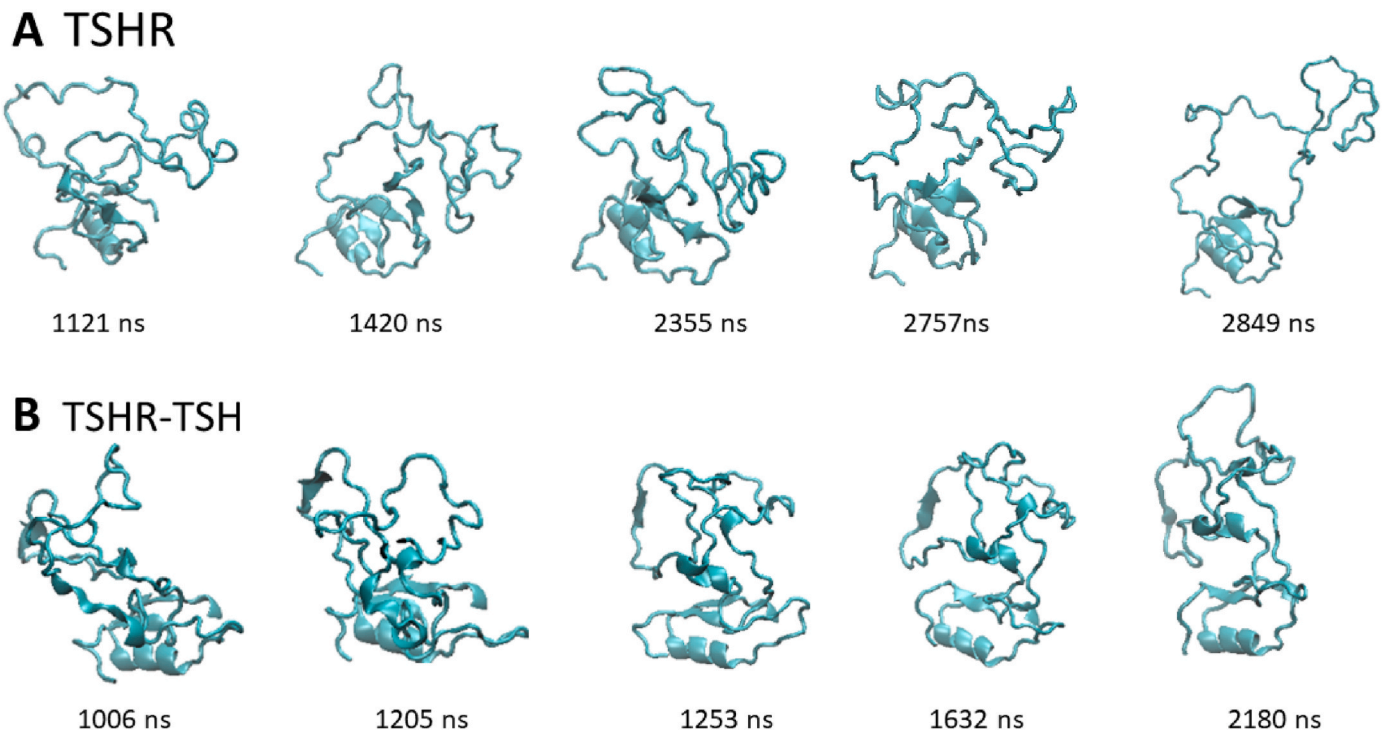

Fig. 5. The representative structures of the five clusters of the LR backbone obtained from the last 2000 ns of the simulations. A: TSHR, B: TSHR-TSH.

Table 1
Fluctuation of LRD orientation.

	CV_x	CV_y	CV_z	SD_x	SDy	SD_z
TSHR	0.037	0.108	0.107	15.8°	27.4°	27.3°
TSHR-TSH	0.031	0.043	0.041	14.4°	16.9°	16.6°

TMD) were also examined for both the TSHR and the TSHR-TSH complex. Table 3 shows the average domain-domain distances and the corresponding standard deviations. The data suggested that the addition of TSH increased the LR-TMD distance by 3.5, a difference that was significant at $p < 0.0001$.

The effect of increasing the overall LR – TMD distance was also examined by calculating the contacts between the LR and TMD. Two

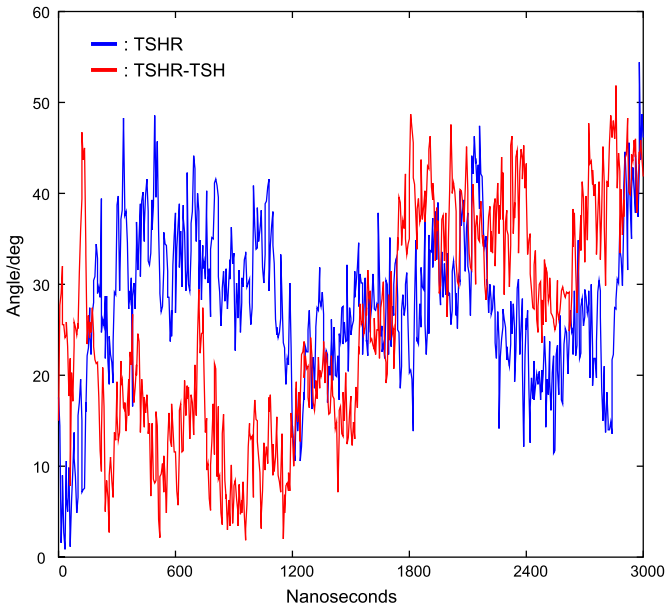

Fig. 6. Evolution of the angle between the first principal axis of the LRD and the membrane normal. Red: TSHR; Blue: TSHR-TSH. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(heavy) atoms A_{LR} and B_{TMD} are considered to be in contact if they are mutually proximal; i.e., A_{LR} is the nearest LR atom to B_{TMD} and B_{TMD} is the nearest TMD atom to A_{LR}. The average number of contact pairs was found to be 17.6 and 12.7 for the TSHR and TSHR-TSH complex, respectively and the average contact distances were 3.52 ± 0.28 and 3.66 ± 0.37 (± indicates one S.D.) for the TSHR and TSHR-TSH complex, respectively **(p=<0.0001)** based on 8000 structures.

2.8. The TMD helix bundle

The helix geometry of the two systems were compared based on 4000 evenly spaced structures from the last 2000 ns of the MD simulations. First, the 2D RMSD maps of the 4000 structures were generated and clustered (Fig. 9). The RMSDs were calculated, after superimposition, only on the backbones of the TMD, without the C-terminal tail. The inertia plots [14] suggested 4 and 3 clusters for TSHR and TSHR-TSH. The clustering also extracted representative structures from the trajectories as the structure whose largest distance from the rest of the cluster members is the smallest. Fig. 9 shows the 2D RMSD maps for the TMD of the TSHR and the TSHR-TSH complex. The helices of the representative structures were identified with the TRAJELIX [15] module of Simulaid. This program is based on the geometry of the C_α atoms, defining the helix axis. Helices with proline are broken up into sub helices. Data in Table 4 shows the average change upon TSH binding (positive number indicates increase) and their significance in the helix length (measured as the end to end distance) and in the radius of the circle fitted to the C_α atoms. This latter is a measure of the bend of the helix - the smaller it is, the more bent is the helix. The largest changes were observed in helices 3 and 8. In both cases the change in length is largely the result in the change in the curvature.

Changes in the distance between the helix axis centers and the change in the closest approach of the helix axes are shown in Table 5. Changes of similar magnitude indicated that the helix axes moved in the membrane plane. When the center-center distance change was larger than the change in the axis-axis distance, the helix movement is interpreted as largely in the direction perpendicular to the membrane plane.

Upper triangle: distances between the helix axis centers; lower triangle: distances between the helix axes.

The changes upon TSH binding in the helix-helix angles are shown in Table 6. **The largest changes were seen for helices 7 and 8**. Helix 7 previously displayed the largest changes when our earlier TSHR model was compared with simulation of the TMD only [8]. However, the large changes in the angles that helix 8 formed with the rest of the bundle were not seen earlier.

2.9. Adequacy of sampling

The adequacy of the sampling is always of concern as most processes occur at time scales far exceeding the simulation times accessible. Figs. 7 and 8 provided some indication that our sampling was rather extensive:

Table 2
Number of hydrogen-bonded residues by domains.

	LR-LR	LR_LRD	LR-TMD	LR-TSH
TSHR	25	26	4	0
TSHR-TSH	55	12	4	15

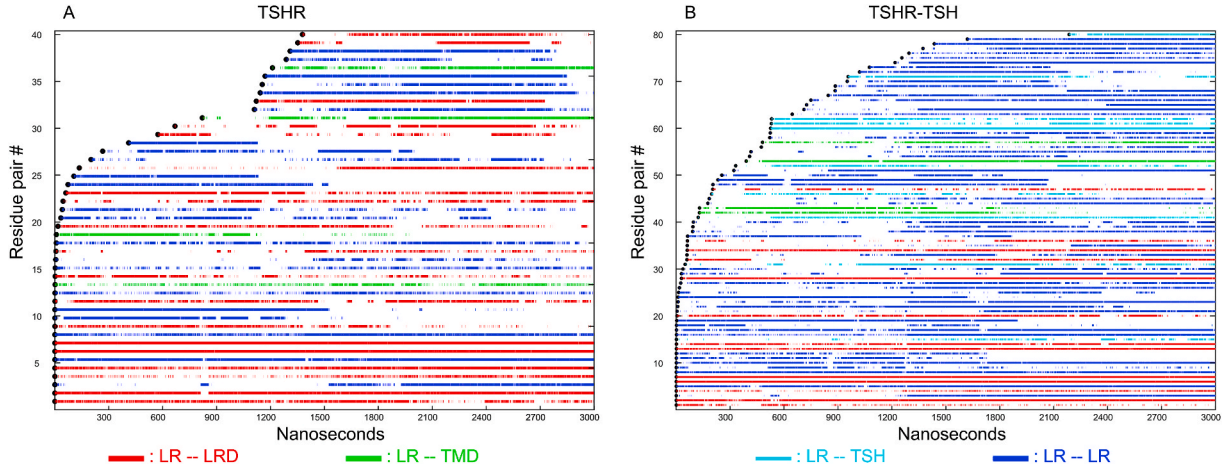

Fig. 7. Plot of the residue pairs involving the LR that were hydrogen bonded at some parts of the simulation. The lines are only present when the residue pair was hydrogen bonded. Blue represents residue pairs within the LR, red represents hydrogen bonds between residues in the LR and the LRD, green represents hydrogen bonds between the LR and the TMD, and cyan represents hydrogen bonds between the LR and the TSH. Residue pairs have to be at least five residues apart (to exclude the many intra-helix hydrogen bonds) and be hydrogen-bonded at least 20 % of simulation time to be represented. A: TSHR, B: TSHR-TSH. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

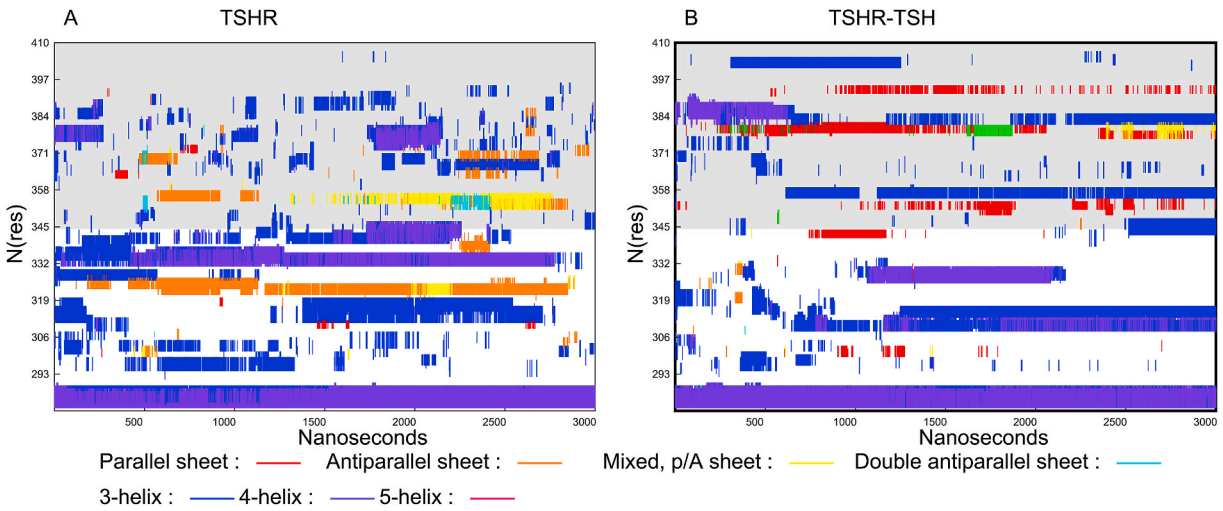

Fig. 8. DSSP plot showing the secondary structure elements formed in the LR during the simulation of the TSHR with and without TSH. The X axis is the simulation time and the Y axis is the residue number. A: TSHR; B: TSHR-TSH.

Table 3
Domain-domain distances.

TSHR			TSHR-TSH		
d(LRD-LR)	d(LRD-TMD)	d(LR-TMD)	d(LRD-LR)	d(LRD-TMD)	d(LR-TMD)
32.1 ± 7.1	65.2 ± 2.4	**51.9** ± 3.3	33.8 ± 1.4	62.9 ± 3.3	**55.4** ± 1.7

Center-of-mass – center-of mass distances in Å averaged over 2000 ns MD runs. ± indicates one S.D.

not only hydrogen bonds were formed, broken and reformed several times, no new significant hydrogen bonds appeared after the first ~1000 ns of simulations. Similarly, and perhaps more importantly, secondary structures (that take much longer to form) were seen to form, unravel and reform.

3. Discussion

We now know that the TSHR consists of a large extracellular ligand binding domain incorporating 11 leucine rich repeats (LRD) and a transmembrane domain (TMD), linked via a 130 AA linker region (LR). As in most GPCRs, the TMD is made up of eight helices joined by extracellular and intracellular loops and a C-terminal cytoplasmic tail. Importantly, the TMD is embedded in a phospholipid bilayer and transmits signals by engaging various G proteins [16] as well as β arrestins [17,18]. Mapping of TSH binding sites and other interaction partners was revealed by homology modeling [19] and then crystallization of the partial ectodomain bound to TSHR autoantibodies [20,21]. This information suggested several possible mechanisms by which

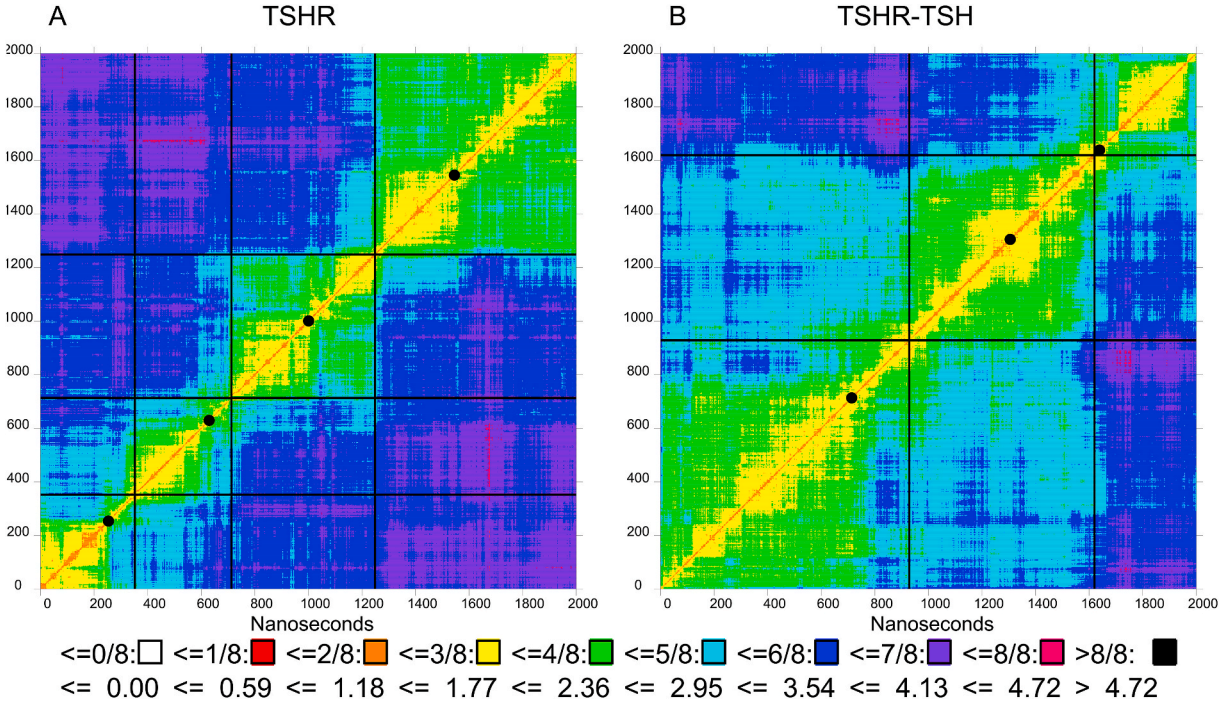

Fig. 9. 2D RMSD (in Å) map of the TMD during the last 2000 ns simulation. Black lines delineate the clusters identified and the black discs on the diagonal indicate the representative structure.

Table 4
Changes in TMD helix length and radius of curvature upon TSH binding.

Helix #:	1		2		3		4		5		6.1		6.2		7.1		7.3		8	
Length:	−0.6	0	−0.2	0	−3.6	2	0.3	1	−0.7	2	**1.2**	**3**	**−0.3**	**3**	0.7	2	0.3	1	**2.7**	**3**
Radius:	−1.1	2	−0.8	1	−3.0	2	0.2	2	**−0.5**	**3**	**0.7**	**3**	**−0.3**	**1**	0.3	2	0.0	0	**2.1**	**3**

A positive number indicates an increase upon TSH binding. The integers 0–3 indicate the number of TSHR-TSH representative structures that were outside the range of the TSHR values. Changes where the TSHR value range was outside all TSHR-TSH values are shown in bold face. The labels of the proline-separated segments of helices 6 and 7 have 0.1 and 0.3 added.

Table 5
Helix-helix distance changes.

HX	1		2		3		4		5		6.1		6.2		7.1		7.3		8	
1			2.0	2	**1.6**	**3**	**2.2**	**3**	−0.7	0	**−2.6**	**3**	−3.2	1	−2.4	2	**3.6**	**3**	−1.3	0
2	1.2	2			**1.3**	**3**	0.6	2	**1.8**	**3**	−0.8	2	−1.1	2	**2.7**	**3**	**4.5**	**3**	**1.2**	**3**
3	0.2	1	**1.1**	**3**			0.6	1	1.5	1	−0.7	1	−2.1	1	**3.3**	**3**	**3.3**	**3**	3.0	2
4	**−9.0**	**3**	0.8	1	−0.3	2			2.4	2	**−1.3**	**3**	−0.7	1	**3.9**	**3**	2.2	0	3.9	2
5	−2.7	1	**0.8**	**3**	1.5	1	2.5	2			**−0.4**	**3**	−0.0	0	**1.4**	**3**	−1.8	1	1.2	2
6.1	**−5.0**	**3**	−2.0	2	−2.7	1	**4.0**	**3**	0.7	0			0.2	1	0.1	0	−3.6	2	**−1.5**	**3**
6.2	−5.9	1	−0.7	0	−1.2	1	0.6	0	−0.0	0	−0.3	1			0.8	1	−0.7	2	**−1.9**	**3**
7.1	0.0	2	4.1	2	**5.8**	**3**	8.1	2	**4.1**	**3**	1.7	2	−0.0	1			0.7	0	**−2.8**	**3**
7.3	−0.1	2	**6.2**	**3**	**11.4**	**3**	**6.4**	**3**	7.8	1	−2.7	2	−1.4	1	−2.6	1			**0.2**	**3**
8	1.9	2	4.8	2	**5.2**	**3**	5.2	2	**−0.9**	**3**	**7.4**	**3**	−1.0	2	0.4	0	**−1.0**	**3**		

Table 6
Helix-helix angle changes.

HX	1		2		3		4		5		6.1		6.2		7.1		7.3		
2	4.0	2																	
3	**−5.4**	**3**	**−0.2**	**3**															
4	**0.8**	**3**	3.0	2	2.8	1													
5	1.9	0	**−5.9**	**3**	−0.4	0	−3.3	2											
6.1	**7.6**	**3**	−1.4	1	−0.5	0	**−7.5**	**3**	**−10.6**	**3**									
6.2	−0.1	0	1.2	1	3.0	0	−1.1	0	−0.5	0	−0.6	0							
7.1	18.1	2	**−18.6**	**3**	**15.9**	**3**	**−15.0**	**3**	**17.1**	**3**	−7.5	2	2.1	0					
7.3	**−22.6**	**3**	**19.4**	**3**	**−16.0**	**3**	**20.8**	**3**	−2.4	0	14.2	1	−2.2	1	−4.6	0			
8	**31.5**	**3**	**−27.9**	**3**	**26.9**	**3**	**−30.6**	**3**	**30.1**	**3**	**−25.9**	**3**	**−25.3**	**3**	11.9	2	**25.2**	***3***	

receptor activation may occur [22–24]. However, a full-length model of the receptor remained incomplete due to the lack of a defined structure for the large TSHR linker region (LR).

Our present studies developed a revised full-length model of the TSHR and the TSHR-TSH complex that was built on the ectodomain structure generated by the artificial intelligence (AI) based protein folding program Alphafold2 [7] together with our MD-refined homology model of the TSHR TMD [8]. The amalgamated full-length structure was further refined with 3000 ns molecular dynamic (MD) simulations in a DPCC membrane environment. In our earlier study of the TSHR we concluded that the LR was an intrinsically disordered protein and we speculated that the highly flexible nature of the LR allowed the TSHR to accommodate either the TSH ligand or autoantibodies to the LRD which act as TSH agonists. We also presented preliminary data supporting the notion that the natural ligand of TSHR, the hormone TSH, may induce more order into the LR.

The current work has improved the TSHR model by **(a)** using the cryo-EM structures that appeared in the meantime [10,25] to correct the TMD-LR interface, **(b)** including three important disulfide bonds into the LR which were previously absent and **(c)** performing longer MD simulations on both the TSHR alone and the TSHR-TSH complex.

Despite the improvements listed above, it is important to point out the limitations of our work that is inherent in (a) the limited accuracy of the force field employed and (b) the (still) limited extent of sampling. While the force field used has been significantly improved from its original version, it is important to remember that multibody effects are only included in an average manner, limiting its accuracy, especially when charged species are involved. As for the adequacy of the sampling, we noted that the hydrogen-bond histories indicated saturation, i.e., no new hydrogen bonds were seen in the second half on the simulations. On the other hand, the 2D RMSD plots, that are more sensitive to conformational changes, showed them to be reaching a significantly different conformational basin near the end of the run of the TSHR system suggesting that the runs are not long enough to quantitate the equilibrium constant between the two states.

However, analysis of these simulations permitted a more detailed examination of the effect of TSH binding on the structure of the TSHR. The structure stabilizing effect of TSH on the TSHR was then demonstrated by analyzing several intrinsic properties:

(a) The radius of gyration of the LR in the TSHR-TSH complex was significantly smaller than in the TSHR alone. Furthermore, the fluctuations in the LR were essentially eliminated in the TSHR-TSH complex while they persisted in the TSHR simulation.
(b) The number of hydrogen bonds (even without counting the intra-helix hydrogen bonds) was more than doubled in the TSHR-TSH complex than in the TSHR alone.
(c) Unlike the TSHR-TSH complex, the residue range 345–408 only contained short-lived transient secondary structures in the TSHR simulation.
(d) The fluctuation of the LRD with respect to the TMD was also reduced by the inclusion of TSH.
(e) TSH binding increased the distance between the centers of mass of the LR and the TMD. While the number of hydrogen-bonded residues did not change upon TSH binding, the number of atom-atom contacts were significantly reduced upon TSH binding.

The significant change in the average LR-TMD distance between the centers of mass was likely to have an electrostatic origin since the formal charges, defined as the difference between the number of basic residues (i.e., ARG and LYS) and acidic residues (i.e., ASP and GLU) were 0, -11 and + 12 for the LRD, LR and TMD, respectively. In contrast, the formal charges on the TSH_α and TSH_β were +2 and + 4. The inclusion of TSH would weaken the electrostatic interaction between the LR and the TMD as found. Since inclusion of TSH activates the TSHR, it can be concluded that the movement of the LR from the TMD is a significant step in receptor activation.

The comparison of the TMD helices in the two simulations also showed a number of significant changes. The reduction in the number of clusters upon TSH binding suggested a reduction in the conformation variability of the TMD, somewhat similarly to its effect on the LR. Helix 3, that was seen to bend and shorten in the earlier TSHR model when the ectodomain was included in the system, was now seen to straighten, thus lengthening upon TSH binding. This would suggest the possibility that the TMD, even without its ectodomain, could be constitutively active. The observation about the LR movement upon TSH binding also leads to this same conclusion.

The structures of the LRD and the TMD of the TSHR have been described in earlier studies [26,27] and here we have focused on the structure of the LR, its intramolecular and molecular bonding dynamics and its structural variations. Based on these studies we can conclude that the LR in the TSHR is highly flexible and without a well-defined tertiary structure although contacts between the LR and the LRD persisted throughout the simulations. These observations, together with the earlier difficulties in obtaining its crystal structure [7], indicate that the LR protein is intrinsically disordered and allows it to accommodate both the ligand or the autoantibodies to the LRD.

In conclusion, our simulation of the full length TSHR embedded in a lipid membrane, solvated with water containing counterions, i.e., in a biologically relevant environment, suggested:

(a) the LRD and the TMD continue to maintain their fold after TSH binding;
(b) the LR is flexible, but maintains protein-like behavior forming secondary structure elements that are, however, transient, and
(c) the relative orientation of the LRD is also variable.

Both the orientation of the LRD and the structural flexibility of the LR suggests that these features are likely to be important for the TSHR to accommodate the diverse ligands such as TSH and autoantibodies that are known to bind to its extracellular region. These aspects add to the confirmation that TSH adds structural stability to the TSHR-TSH complex, which then effectuates a complex signaling procedure.

4. Materials and methods

4.1. Formation of a full-length TSH model by combination of the Alphafold2 LRD and LR model with the TRIO TMD model guided by a cryo-EM structure

The TSHR model was generated in a three-step process guided by the cryo-EM structure obtained from the PDB (PDB ID: 7XW5). First, the Alphafold2 model of the extracellular domain (residues 24–408, downloaded from the Swissprot database [28]; Uniprot #: P16473 and Swissprot file/P1/64/73) was aligned to the cryo-EM structure, using only the LRD for the alignment. Next, the TMD model with the internal waters was aligned to the TMD of the cryo-EM model. This, however, resulted in a large distance between the C terminal of the ectodomain and the N terminal of the TMD. To bridge this gap, we replaced residues 381–413 of the aligned parts with the corresponding coordinates of the cryo-EM structure. This range was chosen because the cryo-EM structure of the LR contained only residues >380 and residue 413 of the cryo-EM model was found to be the nearest to the N terminal of the TMD. We also incorporated all three disulfide bonds since our preliminary work showed that they are important in maintaining the integrity of the structure once the TSH is bonded to it.

The model for the TSHR-TSH complex was generated by aligning the LRD-TSH complex from our earlier model to the LRD of the new model. The coordinates of the aligned structure than were added to the new TSHR model.

The TMD model used was obtained from previous work [8] that generated an the MD trajectory of the TMD. We used our previous model for the TMD (instead of the cry-EM structure) to minimize the need for equilibration as the full structure was to be run with the same MD setup as the earlier model. The MD trajectory was clustered into three clusters using k-medoid clustering [29], performed by the program Simulaid [30]. The largest cluster formed during the second half of the MD trajectory and its representative structure was chosen for our model. The positions of internal waters were determined using grand-canonical ensemble Monte Carlo simulation [31], followed by circular variance [32] filtering [33] and derivation of generic sites [34]. The Monte Carlo simulation, as well as the circular variance and generic site calculations, were performed with the program MMC [35].

4.2. Immersion in bilayer

Both the full model of TSHR, and the TSHR-TSH complex, including the internal waters, were uploaded to the Charmm-gui server [12,13] to (a) immerse it into a bilayer of DPPC molecules; (b) add a water layer; and (c) add counterions (K^+ and Cl^- ions), both to ensure electroneutrality and an ionic strength of 0.15 M to best represent physiological conditions. Hexagonal prism periodic boundary conditions were applied. A six-step protocol [12] equilibrated the system by progressively releasing constraints, as set up by the Charmm-gui server. All simulations were run using the program NAMD [36].

4.3. Molecular dynamics simulation

The simulations used the default parameters set by Charmm-gui. The protein and the ions were represented by the (pairwise additive) Charmm36 m force field [37] and the TIP3P model [38] was used to represent water. The Ewald method was used to treat the long-range electrostatics and the VdW interactions were smoothly cut off to zero at 12 Å, starting at 10 Å. Two fs time step was used and the simulations were run in the (*T,P,N*) ensemble.

Analyses - The program Simulaid [30] performed most analyses on the trajectories.

Hydrogen bonds were defined as X···H–Y where X and Y are polar heavy atoms, the X···H–Y angle is above 120° and the X–H distance is below threshold. The thresholds used for N–H, O–H, P–H, and S–H were 2.52, 2.52, 3.24, and 3.15 Å, respectively. Note, that this definition ignores the actual charges thus it includes salt bridges as well; several of the hydrogen bonds thus defined indeed qualified for being a salt bridge. **The variation of the shape of the LR** was tracked by calculating the radius of gyration. **The formation and unraveling of secondary structure elements in the LR** was tracked with the DSSP algorithm [39].

The orientation fluctuation of the LRD was characterized with two properties: (a) the circular variance of the three coordinate axes attached to the LRD and (b) the angle between the membrane plane normal and the first principal axis of the LRD. For the circular variance calculations, a local coordinate system was attached to the LRD and for each axis the circular variance of the angles between the axes in the initial orientation and in the trajectory frames analyzed was calculated. The local X axis was defined as the unit vector connecting the average (over 20 frames) of the C_α atoms of residues 48 and 220, an approximate Y′ axis was defined as the unit vector connecting the C_α atoms of residues 115 and 126. The Z axis was defined as the vector product of the X and Y' unit vectors and the exact Y axis was defined as the vector product of the

Z and X unit vectors. The standard deviation corresponding to a circular variance value was estimated by assuming normal distribution of the angles. The principal axis calculation used the backbone atoms only.

The **clustering** of the LR and TMD conformations were performed using k-medoid clustering. The number of clusters requested was obtained by generating an inertia plot. The representative structure of a cluster was obtained as the member of the cluster whose largest RMSD w.r.t. the rest of the cluster members is the smallest.

The **changes in the helix properties** were calculated using the representative structures as follows. First, each property was averaged over the four representative TMD structures of the TSHR. Then, difference between these averages and each of the three representative structure of the TSHR-TSH complex was calculated and their average was given in the tables of change. Furthermore, as a measure of significance, for each property it was checked how many of the TSHR-TSH complex value was outside the range of the TSHR value.

CRediT authorship contribution statement

Mihaly Mezei: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Rauf Latif:** Conceptualization, Data curation, Visualization, Writing – review & editing. **Terry F. Davies:** Conceptualization, Formal analysis, Funding acquisition, Project administration, Visualization, Writing – review & editing.

Declaration of competing interest

Terry F. Davies is member of the Board of Kronus Inc (Starr, ID, USA); Mihaly Mezei, and Rauf Latif have nothing to disclose.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by a VA Merit Award BX000800 (to TFD), the Icahn School of Medicine at Mount Sinai and additional anonymous donors. It was also supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmgm.2024.108725.

References

[1] T.F. Davies, S. Andersen, R. Latif, Y. Nagayama, G. Barbesino, M. Brito, A. K. Eckstein, A. Stagnaro-Green, G.J. Kahaly, Graves' disease, Nat. Rev. Dis. Prim. 6 (1) (2020) 52.
[2] B. Rapoport, G.D. Chazenbalk, J.C. Jaume, S.M. McLachlan, The thyrotropin (TSH) receptor: interaction with TSH and autoantibodies, Endocr. Rev. 19 (6) (1998) 673–716.
[3] T.F. Davies, T. Ando, R.Y. Lin, Y. Tomer, R. Latif, Thyrotropin receptor-associated diseases: from adenomata to Graves disease, J. Clin. Invest. 115 (8) (2005) 1972–1983.
[4] K. Tanaka, G.D. Chazenbalk, S.M. McLachlan, B. Rapoport, Subunit structure of thyrotropin receptors expressed on the cell surface, J. Biol. Chem. 274 (48) (1999) 33979–33984.
[5] K. Tanaka, G.D. Chazenbalk, B. Rapoport, S.M. McLachlan, Reassessment of the location of the thyrotropin receptor 50 amino acid "insertion" provides evidence in favor of a second downstream cleavage site, Thyroid 11 (2) (2001) 111–114.
[6] M. Mezei, R. Latif, T.F. Davies, Computational model of the full-length TSH receptor, Elife 11 (2022) e81415.
[7] J. Jumper Re, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabi, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589.
[8] M. Mezei, R. Latif, B. Das, T.F. Davies, Implications of an improved model of the TSH receptor transmembrane domain (TSHR-TMD -TRIO), Endocrinology (2021) 162, bqab051.
[9] R.N. Miguel, P. Sanders, L. Allen, M. Evans, M. Holly, W. Johnson, A. Sullivan, J. Sanders, J. Furmaniak, B.R. Smith, Structure of full-length TSH receptor in complex with antibody K1-70, J. Mol. Endocrinol. 70 (2022) e220120.
[10] J. Duan, P. Xu, X. Luan, Y. Ji, X. He, N. Song, Q. Yuan, Y. Jin, X. Cheng, H. Jiang, J. Zheng, S. Zhang, Y. Jiang, H.E. Xu, Hormone- and antibody-mediated activation of the thyrotropin receptor, Nature 609 (2022) 854–859.
[11] B. Faust, C.B. Billesbølle, C.-M. Suomivuori, I. Singh, K. Zhang, N. Hoppe, A.F. M. Pinto, J.K. Diedrich, Y. Muftuoglu, M.W. Szkudlinski, A. Saghatelian, R.O. Dror, Y. Cheng, A. Manglik, Autoantibody mimicry of hormone action at the thyrotropin receptor, Nature 609 (2022) 846–860.
[12] E.L. Wu, X. Cheng, S. Jo, H. Rui, K.C. Song, E.M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R.M. Venable, J.B. Klauda, W. Im, CHARMM-GUI membrane builder toward realistic biological membrane simulations, J. Comput. Chem. 35 (2014) 1997–2004.
[13] S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM, J. Comput. Chem. 29 (11) (2008) 1859–1865.
[14] R.L. Thorndike, Who belongs in the family? Psychometrika 18 (1953) 267–276.
[15] M. Mezei, M. Filizola, TRAJELIX: a computational tool for the geometric characterization of protein helices during molecular dynamics simulations, J. Comput. Aided Mol. Des. 20 (2006) 97–107.
[16] K.L. Laugwitz, A. Allgeier, S. Offermanns, K. Spicher, J. Van Sande, J.E. Dumont, G. Schultz, The human thyrotropin receptor: a heptahelical receptor capable of stimulating members of all four G protein families, Proc. Natl. Acad. Sci. U.S.A. 93 (1) (1996) 116–120.
[17] A. Boutin, E. Eliseeva, M.C. Gershengorn, S. Neumann, beta-Arrestin-1 mediates thyrotropin-enhanced osteoblast differentiation, Faseb. J. 28 (8) (2014) 3446–3455.
[18] R. Frenzel, C. Voigt, R. Paschke, The human thyrotropin receptor is predominantly internalized by beta-arrestin 2, Endocrinology 147 (6) (2006) 3114–3122.
[19] R. Nunez Miguel, J. Sanders, J. Jeffreys, H. Depraetere, M. Evans, T. Richards, T. L. Blundell, B. Rees Smith, J. Furmaniak, Analysis of the thyrotropin receptor-thyrotropin interaction by comparative modeling, Thyroid 14 (12) (2004) 991–1011.
[20] J. Sanders, D.Y. Chirgadze, P. Sanders, S. Baker, A. Sullivan, A. Bhardwaja, J. Bolton, M. Reeve, N. Nakatake, M. Evans, T. Richards, M. Powell, R.N. Miguel, T. L. Blundell, J. Furmaniak, B.R. Smith, Crystal structure of the TSH receptor in complex with a thyroid-stimulating autoantibody, Thyroid 17 (5) (2007) 395–410.
[21] P. Sanders, S. Young, J. Sanders, K. Kabelis, S. Baker, A. Sullivan, M. Evans, J. Clark, J. Wilmot, X. Hu, E. Roberts, M. Powell, R. Nunez Miguel, J. Furmaniak, B. Rees Smith, Crystal structure of the TSH receptor (TSHR) bound to a blocking-type TSHR autoantibody, J. Mol. Endocrinol. 46 (2) (2011) 81–99.
[22] X. Jiang, D. Fischer, X. Chen, S.D. McKenna, H. Liu, V. Sriraman, H.N. Yu, A. Goutopoulos, S. Arkinstall, X. He, Evidence for follicle-stimulating hormone receptor as a functional trimer, J. Biol. Chem. 289 (20) (2014) 14273–14282.
[23] G. Kleinau, G. Krause, Thyrotropin and homologous glycoprotein hormone receptors: structural and functional aspects of extracellular signaling mechanisms, Endocr. Rev. 30 (2) (2009) 133–151.
[24] G. Krause, A. Kreuchwig, G. Kleinau, Extended and structurally supported insights into extracellular hormone binding, signal transduction and organization of the thyrotropin receptor, PLoS One 7 (12) (2012) e52920.
[25] J. Duan Px, Xi Cheng, C. Mao, T. Croll, X. He, J. Shi, X. Luan, W. Yin, E. You, Q. Liu, S. Zhang, H. Jiang, Y. Zhang, Yi Jiang, H. Xu, Structures of full-length glycoprotein hormone receptor signalling complexes, Nature 598 (2021) 688–692.
[26] M.R. Ali, R. Latif, T.F. Davies, M. Mezei, Monte Carlo loop refinement and virtual screening of the thyroid-stimulating hormone receptor transmembrane domain, J. Biomol. Struct. Dyn. (2014) 1–13.
[27] R. Latif, M.R. Ali, M. Mezei, T.F. Davies, Transmembrane domains of attraction in the TSH receptor, Endocrinology 156 (2014) 488–489.
[28] A. Bairocha. R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, Nucleic Acids Res 28 (2000) 45–48.
[29] L. Kaufman, P.J. Rousseeuw, in: Y. Dodge (Ed.), Clustering by Means of Medoids, in Statistical Data Analysis Based on the L1 - Norm and Related Method, North Holland, Amsterdam, 1987, pp. 405–416.
[30] M. Mezei, Simulaid: a simulation facilitator and analysis program, J. Comput. Chem. 31 (14) (2010) 2658–2668.
[31] M. Mezei, Grand-canonical ensemble Monte Carlo simulation of dense fluids: Lennard-Jones, soft spheres and water, Mol. Phys. 61 (1987) 565–582.
[32] KV Mardia, PE Jupp. Directional Statistics, John Wiley & Sons, Ltd, Chichester, 1999.
[33] M. Mezei, A new method for mapping macromolecular topography, J. Mol. Graph. Model. 21 (5) (2003) 463–472.
[34] M. Mezei, D.L. Beveridge, Generic solvation sites in a crystal, J. Comput. Chem. 6 (1984) 523–527.
[35] Mezei M. MMC: Monte Carlo program for molecular assemblies. URL: https://mezeim01.dmz.hpc.mssm.edu/mmc.
[36] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.

[37] S.R.J. Huang, G. Nawrocki, T. Ran, Mi Feig, B.L. de Groot, H. Grubmüller, A. D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat, Methods 14 (2017) 71–73.
[38] W.L. Jorgensen Jc, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[39] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.