

Summary of integrative structure determination of The integrative model of the TRV026-stabilized conformation of AT1R was obtained from experimental DEER distance restraints using a protocol combining AlphaLink, AlphaFold2-Multimer, and short unbiased molecular dynamics simulations in GROMACS. (PDB ID: 9AAB | [pdb_00009aab](#))

1. Model Composition	
1.1. Entry composition	- Type-1 angiotensin II receptor: chain(s) A (319 residues) - TRV026, beta-arrestin biased peptide: chain(s) B (8 residues)
1.2. Datasets used for modeling	- EPR data, Zenodo: 10.5281/zenodo.20259483 - EPR data, Zenodo: 10.5281/zenodo.20259483 - Experimental model, PDB: pdb_00004yay - Experimental model, PDB: pdb_00006os2
2. Representation	
2.1. Number of representations	1
2.2. Scale	Atomic
2.3. Number of rigid and flexible segments	0, 2
3. Restraints	
3.1. Physical principles	Information about physical principles was not provided
3.2. Experimental data	- 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 30.871-38.852 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 32.844-36.864 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 23.375-33.31 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 19.252-25.675 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 29.424-34.66 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 17.855-25.813 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 25.044-32.648 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 35.176-41.433 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 28.035-35.661 - 1 unique DerivedDistanceRestraint: Lower Upper Bound Distance: 22.627-29.004
4. Validation	
4.2. Number of ensembles	0
4.3. Number of models in ensembles	Not applicable

<i>4.4. Number of deposited models</i>	3
<i>4.5. Model precision</i>	Not available
<i>4.6. Data quality</i>	Data quality has not been assessed
<i>4.7. Model quality: assessment of atomic segments</i>	- Clashscore: 0.74-1.29 - Ramachandran outliers: 1-8 - Sidechain outliers: 3-11
<i>4.8. Fit to data used for modeling</i>	Fit of model to information used to compute it has not been determined
<i>4.9. Fit to data used for validation</i>	Fit of model to information not used to compute it has not been determined
5. Methodology and Software	
1. <i>5.1. Method name</i>	Sampling
<i>5.2. Method type</i>	AlphaLink
<i>5.3. Method description</i>	AlphaLink builds AT1R models using ten experimental DEER restraints and the AT1R sequence. A total of 2000 models were generated: 1000 with -Neff 5 and 1000 with -Neff 10.
<i>5.4. Number of computed models</i>	2000
2. <i>5.1. Method name</i>	Filtering
<i>5.2. Method type</i>	Filtering: adequacy of models
<i>5.3. Method description</i>	AlphaLink models were filtered based on per-helix RMSD relative to inactive (PDB: 4YAY) and active (PDB: 6OS2) experimental AT1R structures. Each model was aligned to both references using backbone heavy atoms of TM1-TM7, and RMSD was computed separately for all seven helices. Models were discarded if any TM helix exceeded 4.5 Å RMSD from both reference states, ensuring removal of distorted conformations.
<i>5.4. Number of computed models</i>	1734
3. <i>5.1. Method name</i>	Clustering
<i>5.2. Method type</i>	Clustering of AlphaLink models
<i>5.3. Method description</i>	AlphaLink models were clustered using three non-correlated Ca-Ca distances: F1.59(55)-R34.56(139), F1.59(55)-D6.31(236), and F1.59(55)-R8.52(311). Clustering was performed with GaussianMixture model, and one representative structure was selected from each cluster.
<i>5.4. Number of computed models</i>	8

4. 5.1. Method name	Sampling
5.2. Method type	AlphaFold2-Multimer
5.3. Method description	AlphaFold2-Multimer (ColabFold version) was used to place TRV026 onto each representative AlphaLink model. Each AlphaFold2-Multimer run used the AlphaLink representative as the main receptor template and the experimental AT1R-TRV026 structure (PDB: 6OS2) as a hypothesis for the peptide pose.
5.4. Number of computed models	8
5. 5.1. Method name	Sampling
5.2. Method type	MD simulations in GROMACS
5.3. Method description	Molecular dynamics simulations were performed in GROMACS. 80 trajectories total, each 500 ns long: 8 systems 5 replicas 2 D3.49(125) protonation variants. Details are available in the manuscript.
5.4. Number of computed models	80
6. 5.1. Method name	Filtering
5.2. Method type	RMSD-based filtering: ligand and receptor stability
5.3. Method description	RMSD-based filtering was applied to ensure (i) stability of the AT1R-TRV026 complex and (ii) absence of rapid, non-physical conformational changes in the receptor. Ligand stability: TRV026 heavy-atom RMSD between consecutive frames ≤ 2.0 in at least 95% of frames. Receptor stability: maximum absolute RMSD change of TM6 intracellular part (D6.31-W6.48) within any 50-ns window ≤ 2.0 . Equilibration period was identified using PELT changepoint detection, capped at 75 ns. 39 trajectories passed both filters.
5.4. Number of computed models	39
7. 5.1. Method name	Analysis
5.2. Method type	Identification of plateau regions in TM1-TM6 distance
5.3. Method description	For the 39 filtered trajectories, we identified plateau regions where TM6 adopted stable conformations. DEER distance distributions were simulated for each frame using Off-Rotamer Sampling (chiLife library, V1 spin label, 5000 iterations). Plateaus were identified from TM1-TM6 peak distance time series: smoothed with 25-frame rolling average, local slope calculated over 50-frame windows, slopes $<$ threshold classified as plateaus. Plateaus classified as inactive-like (≤ 32.7 Å), intermediate-like (32.7-38.3 Å), or active-like (≥ 38.3 Å) based on experimental TM1-TM6 distance.

<i>5.4. Number of computed models</i>	39
8. <i>5.1. Method name</i>	Analysis
<i>5.2. Method type</i>	Computing AT1R-TRV026 interactions
<i>5.3. Method description</i>	Receptor-ligand interactions, hydrogen bonds, salt bridges and pi- or T-stacking, were identified in each frame of the 39 filtered trajectories using MDAnalysis.
<i>5.4. Number of computed models</i>	39
9. <i>5.1. Method name</i>	Clustering
<i>5.2. Method type</i>	Clustering of interactions into binding modes
<i>5.3. Method description</i>	Interactions were classified as scaffold (constant frequency) or BM-defining (concentrated in specific intervals). Time-lagged Independent Component Analysis (TICA) was applied to reduce dimensionality. K-means clustering (scikit-learn) was performed on TICA-transformed coordinates. Binding mode segments were defined as continuous stretches. Similar BMs were merged using hierarchical clustering. As a result, we obtained 76 binding modes.
<i>5.4. Number of computed models</i>	76
10. <i>5.1. Method name</i>	Classification
<i>5.2. Method type</i>	Classification of binding modes by TM6 conformation
<i>5.3. Method description</i>	TM6 conformation (inactive/intermediate/active) was assigned to each BM by overlapping with plateau regions. The procedure resulted in seven inactive-like, nine intermediate-like, and 60 active-like BMs. Nine intermediate BMs (INTs) were selected for linear combination analysis.
<i>5.4. Number of computed models</i>	9
11. <i>5.1. Method name</i>	Validation
<i>5.2. Method type</i>	Linear combination analysis of intermediate binding modes
<i>5.3. Method description</i>	For each of the nine intermediate BMs, representative DEER distributions were obtained by averaging over a subset of frames. The optimal percentage (8%) was selected using five-fold leave-one-out cross-validation (LOO-CV). Linear combination coefficients were fitted using SLSQP optimizer (SciPy) with non-negative constraint, sum-to-one constraint, constant offset term. Coefficient uncertainties were estimated by bootstrapping (500 iterations). Final model retained three INTs: INT_2_p, INT_3_p, INT_4_d, representative structures of which were added to the current model.

<i>5.4. Number of computed models</i>	3
<i>5.5. Software</i>	- AlphaLink (version 1.0) - AlphaFold2-Multimer (version 3) - GROMACS (version 2023) - LongDistances (version 590)