

Structural and functional role of the magnesium ion in the human oxytocin receptor

Francesco Musiani^{1,*,#}, Samira Al Sadi^{1,2,#}, Alejandro Giorgetti³, Mercedes Alfonso-Prieto⁴
and Paolo Carloni^{4,5,6,7,*}

¹ Laboratory of Bioinorganic Chemistry, Department of Farmacy and Biotechnology, University of Bologna, Bologna, Italy.

² Present address: Department of Inorganic and Organic Chemistry & IQTCUB, University of Barcelona, Barcelona, Spain.

³ Department of Biotechnology, University of Verona, Verona, Italy.

⁴ Institute of Neuroscience and Medicine (INM-9), Computational Biomedicine, Forschungszentrum Jülich, Jülich, Germany.

⁵ JARA-HPC, Forschungszentrum Jülich, Jülich, Germany.

⁶ Department of Physics, RWTH Aachen University, Aachen, Germany.

⁷ JARA Institute Molecular Neuroscience and Neuroimaging (INM-11), Forschungszentrum Jülich, Jülich, Germany.

* FM: Via Piero Gobetti 87, 40129 Bologna, Italy. E-mail: francesco.musiani@unibo.it; PC: Wilhelm-Johnen-Straße, 52428 Jülich, Germany. E-mail: p.carloni@fz-juelich.de.

FM and SAS contributed equally to this work.

SUPPLEMENTARY MATERIAL

S1 – OXYTOCIN BINDING POCKET

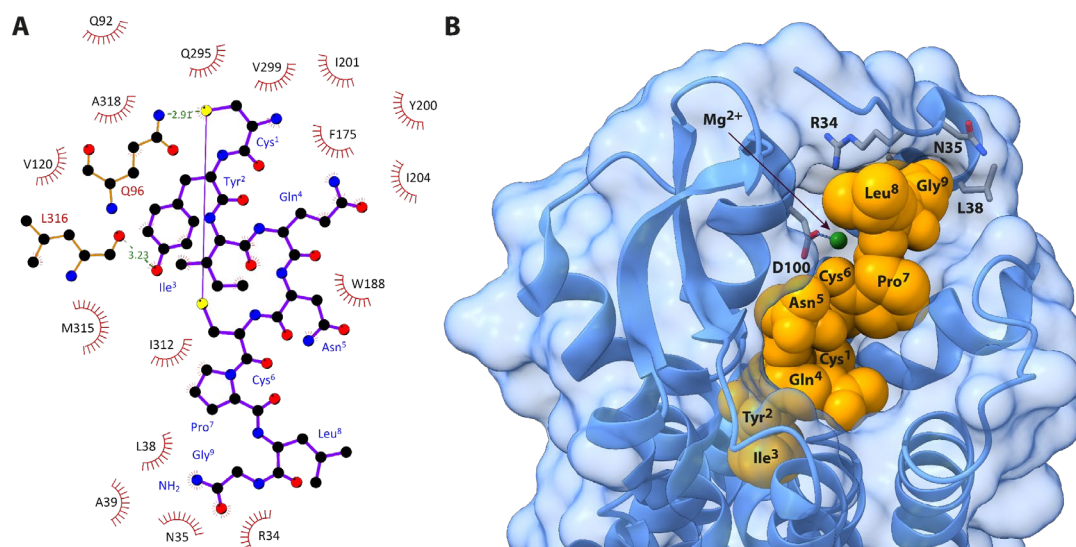


Figure S1. Detail of the interaction between oxytocin (OT) and the human oxytocin receptor (OTR) in the cryo-EM structure determined at a resolution of 0.290 nm (PDB ID 7RYC) in complex with a heterotrimeric G_q protein and bound to oxytocin (OT) and a single Mg²⁺ ion (AOM-OTR hereafter) [1]. **(A)** Schematic representation of OT binding pose calculated using LigPlot+ [2, 3]. The scheme represents H-bonds (green dashed lines) and van der Waals hydrophobic interactions (red spiked arcs). OT and OTR H-bonding residues are in ball-and-stick representation with the atoms coloured according to the atom type. **(B)** Detail of OT binding in the AOM-OTR cryo-EM structure. The OTR ribbon representation and the molecular surface are in light blue, while OT atoms are in orange. The three letters code with superscript numbering has been used for OT residues (e.g., Asn⁵), while the single letter code refers to OTR residues (e.g., R34).

S2 – MATERIAL AND METHODS

AOM-OTR homology modelling. The model on which the simulations of human oxytocin receptor (OTR) were based was the aforementioned AOM-OTR cryo-EM structure (PDB ID 7RYC) [1]. AOM-OTR presented a few unmodelled regions, namely, residues 1 to 29 at the N-terminus and 346 to 389 at the C-terminus. Additionally, a small loop portion (residue 68 in ICL2) and a 30-residue stretch (residues 237-265, TM5-ICL3-TM6) were also unresolved. To prepare the initial structure for this study, the receptor's missing regions were modelled using MODELLER 10.4 [4]. The modelling was performed on the receptor in complex with the G_q protein, to avoid generating conformations that overlapped with G_q. The unresolved N-terminal and C-terminal ends were not modelled, and an acetylation capping group was introduced in the structure at the N-terminus. Instead, the modelling focus was on the small ICL2 loop, and the 30-residue TM5-ICL3-TM6 stretch. Although allowing high-accuracy computational predictions, AlphaFold2 [5] only predicts a single state and exhibits a bias towards either the active or inactive conformation, based on the GPCR class and the availability of experimental structures in different activation states. Considering the importance of conformational states, and the identification of active-state and inactive-state configurations, the ICL2 loop and the 30-residue TM5-ICL3-TM6 stretch were modelled based on the active-state and inactive-state models generated with AlphaFold-Multistate [6], an extension of AlphaFold2 which employs a multistate prediction protocol to predict either active or inactive states at very high accuracy using state-annotated templated GPCR databases [7]. The inactive-state AlphaFold2 model structure featured the TM6 N-terminal portion divided into two helices, while the active one features an elongated TM6 helix (Figure 2A). Therefore, based on this configuration, the inactive AlphaFold2-based model was named as *TM6 broken* (OTRb), while the active AlphaFold-based model is the *TM6 elongated* (OTRe). The AOM-OTR structure together with the TM5-ICL3-TM6 portions obtained from AlphaFold-Multistate were used as templates to calculate 100 models of AOM-

OTR in both the OTRb and OTRe conformation using MODELLER. The best model was selected using DOPE potential [8] function included in MODELLER. A loop optimization routine was used to refine the regions that showed higher than average energy as calculated using the DOPE potential function. The stereochemical quality of the model structure was established using ProCheck [9]. The results of this analysis confirm the reliability of the OTRb and OTRe model structures.

System preparation and binding cavity hydration. The OTRb and OTRe model structures were oriented for membrane insertion using the PPM3 server [10] and then embedded into a lipid membrane using the CHARMM-GUI *Membrane Builder* functionality [11, 12]. In both cases, the membrane was composed by 70% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 30% cholesterol (CHL) for a total of 128 lipids/leaflet (Figure S2) [13]. The systems were then solvated in a rectangular simulation box and neutralized by adding 17 Cl^- ions. Analogously, additional Na^+ and Cl^- ions were placed in the water box to achieve physiological ionic strength (150 mM). For the simulations, the Amber14sb force field [14] and TIP3P water model [15] were used for the protein and water molecules, respectively. A nonbonded multi-site dummy atom model was used to parametrize the Mg^{2+} ion included in the simulated systems [16]. Considering the low resolution of the initial cryo-EM structure and the absence of water molecules in the first Mg^{2+} coordination shell [1], additional water molecules were included in the model. In particular, a water molecule was included in the Mg^{2+} cavity by using the CHARMM-GUI *Generate Pore Water* function [17]. Moreover, four additional water molecules were included in the first Mg^{2+} coordination shell by using the *Metal Geometry* tool included in UCSF Chimera 1.16 (Figure 2B in the main text) [18]. The final OTRb and OTRe systems comprised ca. 108,000 and 111,000 atoms, respectively.

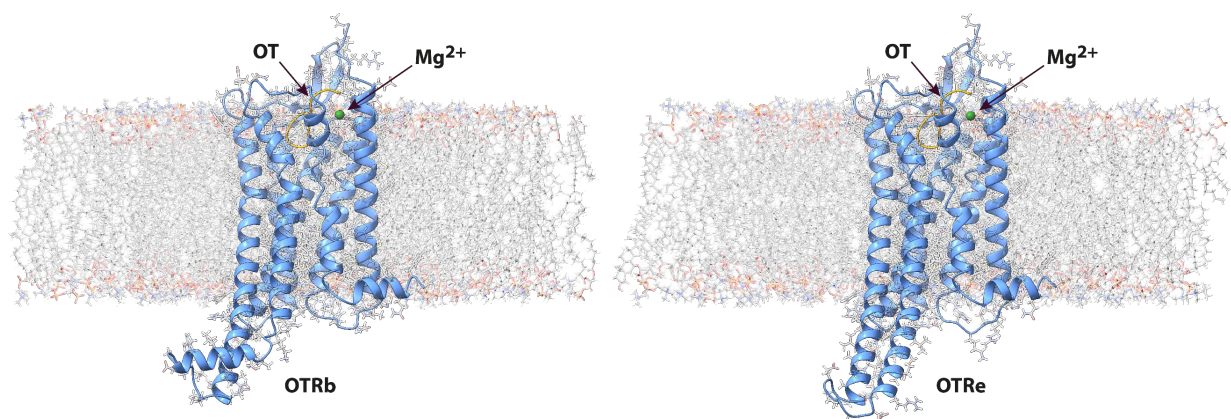


Figure S2. Ribbon schemes of the OTRb (left) and OTRe (right) model structure embedded in the membrane used for the simulations. The water molecules and the ions included in the final systems are not reported for the sake of clarity. OTRb and OTRe model structures are in light blue ribbons, OT is shown in orange ribbons and the Mg^{2+} ion is reported as a green sphere. The membrane constituents (POPC and CHL) as well as the protein atoms are in transparent sticks coloured according to the atom type.

Simulation details. All the molecular dynamics calculations were performed using GROMACS 2020.1 [19-21]. Periodic boundaries were applied to both systems and the electrostatic interactions were calculated through the Particles Mesh Ewald method [22]. The cut-off values for the real part of the electrostatic interactions and for the van der Waals interactions were set to 0.9 nm. The two systems were subjected to a single-stage energy minimization protocol, which involved 5000 steps of steepest descent with position restraints on the heavy atoms (see Table S1). The systems were then equilibrated with a six-stage protocol for a total of 1.25 ns (see Table S1). To assure gradual equilibration, positional restraints were applied to the protein and lipids, then slowly reduced as it progressed (see Table S1). A time step of 2 fs was adopted and the LINCS algorithm [23] was used for imposing constraints on the hydrogen-containing bonds. Temperature was kept constant at 303.15 K. To successfully equilibrate the systems, the NVT ensemble was employed in the first 0.5 ns, while

the NPT ensemble was adopted for the rest of the equilibrations. During equilibration, temperature and pressure were coupled to a Berendsen thermostat and barostat [24]. During production, a Nosé-Hoover thermostat [25], and a Parrinello-Rahman barostat [26] were employed. For each system, three independent, 1 μ s-long replicates were generated, allowing for robust and statistically relevant results [27, 28].

Table S1. Steps and restraints used in the minimization and in the six equilibration stages on each system.

Stage	Steps	Restraints force constants ($\text{J mol}^{-1} \text{ nm}^{-2}$) on:		
		Protein backbone heavy atoms	Protein side chain heavy atoms	Lipid heads heavy atoms
Minimization	5,000	4000	2000	1000
Equilibration 1	125,000	4000	2000	1000
Equilibration 2	125,000	2000	1000	400
Equilibration 3	125,000	1000	500	400
Equilibration 4	250,000	500	200	200
Equilibration 5	250,000	200	50	40
Equilibration 6	250,000	50	0	0

Simulation analysis. The trajectories were analyzed by using GROMACS tools and in-house Python3 scripts, making use of the MDTraj 1.9.8 [29] and mdciaio 1.0.0 [30] libraries. The A^{100} activation index was calculated as in Ref. [31]. The trajectories were visually inspected using UCSF ChimeraX 1.3 [32, 33].

S3 – SUPPLEMENTARY RESULTS

General behaviour of the OTRb and OTRe systems. The structural integrity of the two systems was evaluated throughout the simulations by assessing the fluctuation of individual residues relative to the average structure, as well as the deviation of each structure from its initial coordinates, following superimposition on the Ca atoms of the experimentally determined OTR regions. The residue-based root mean square fluctuation (RMSF) of the trajectories was calculated for the Ca atoms to measure the flexibility of the protein chain (Figure S3). The RMSF was calculated on each single replica and on the cumulative trajectory obtained summing up the three replicates in order to increase the sampling. The data shows that – with the exception of the N-terminal residues – the experimentally determined regions are quite rigid with some regions featuring a relatively high RMSF values, corresponding to the loops connecting the OTR transmembrane helices. On the other hand, residues 237-265, constituting the modelled TM5-ICL3-TM6 region, are particularly flexible. In particular, OTRb replica 3 (R#3) experiences large fluctuations in the latter region. This displacement causes a complete reorganization of the entire region, that takes on a conformation located below the receptor. Such conformation occupies the spatial region that would be occupied by the G_q protein and is not observed in the other replicates.

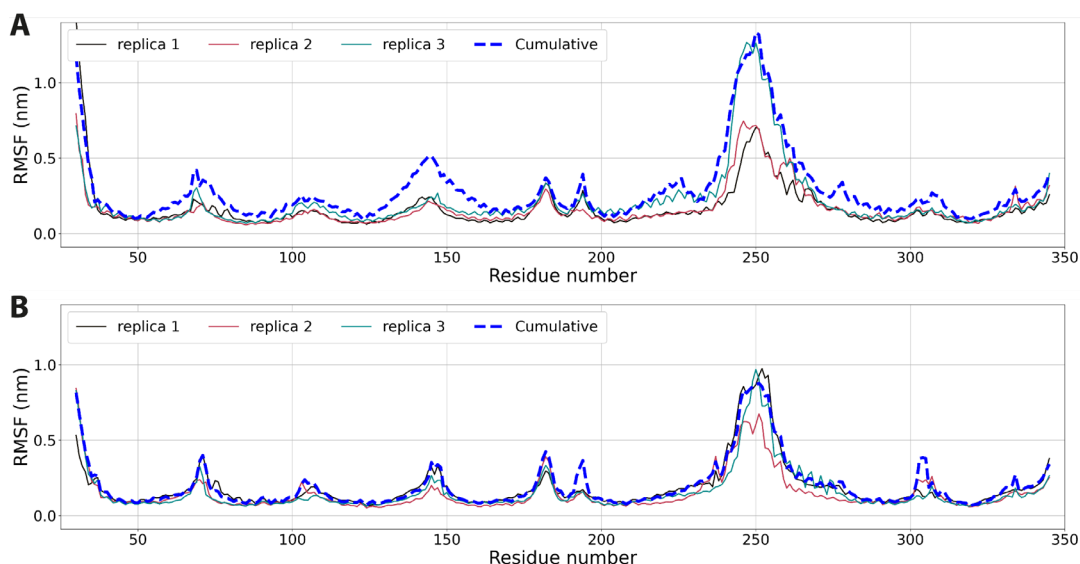


Figure S3. Values of the root mean square fluctuations for the Ca atoms as a function of the residue number for OTRb (A) and OTRe (B) systems. RMSF values for the three replicas simulated for each system are in black, dark red and cyan, respectively, while the RMSF calculated on the cumulative trajectory are reported using a blue bold dashed line.

The root mean square deviation (RMSD) for the Ca atoms of the whole protein with respect to the minimized structures was calculated and is depicted in Figure S4, while the same calculation performed by excluding the most flexible regions is reported in Figure S5. When considering the entire protein, the three replicas of the OTRe system experience less fluctuations in RMSD than the OTRb systems. By excluding the modelled TM5-ICL3-TM6 region, the behaviour of the RMSD as a function of the simulation time showed that each replica in both systems relaxes to conformations that differ by around 0.2 nm from the starting configuration and that are quite stable along the simulation time.

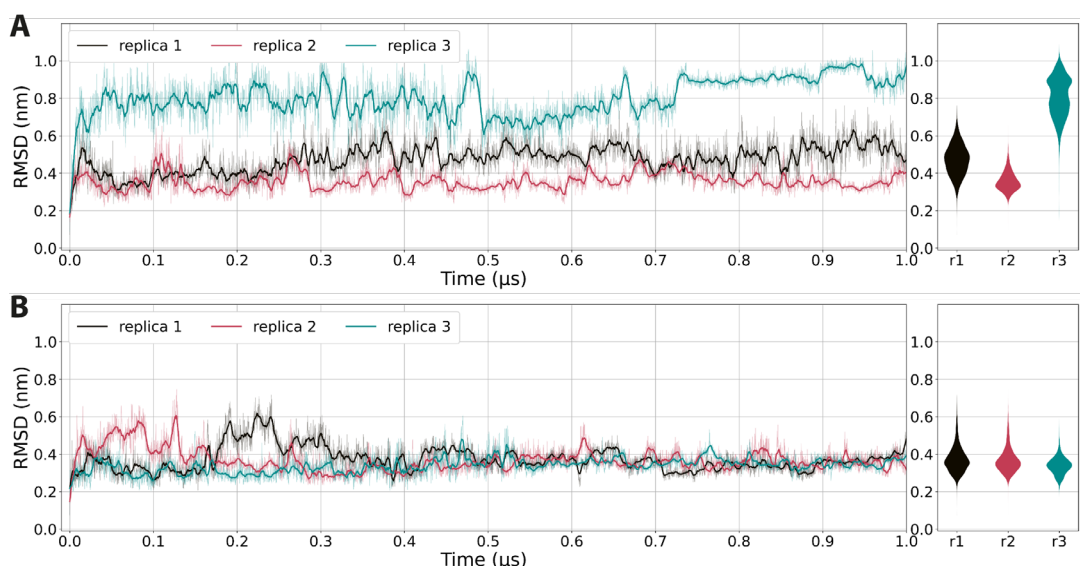


Figure S4. Values of the root mean square deviation from the minimized structures for the Ca atoms as a function of the simulation time for OTRb (A) and OTRe (B) systems. RMSD values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the RMSD values for the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

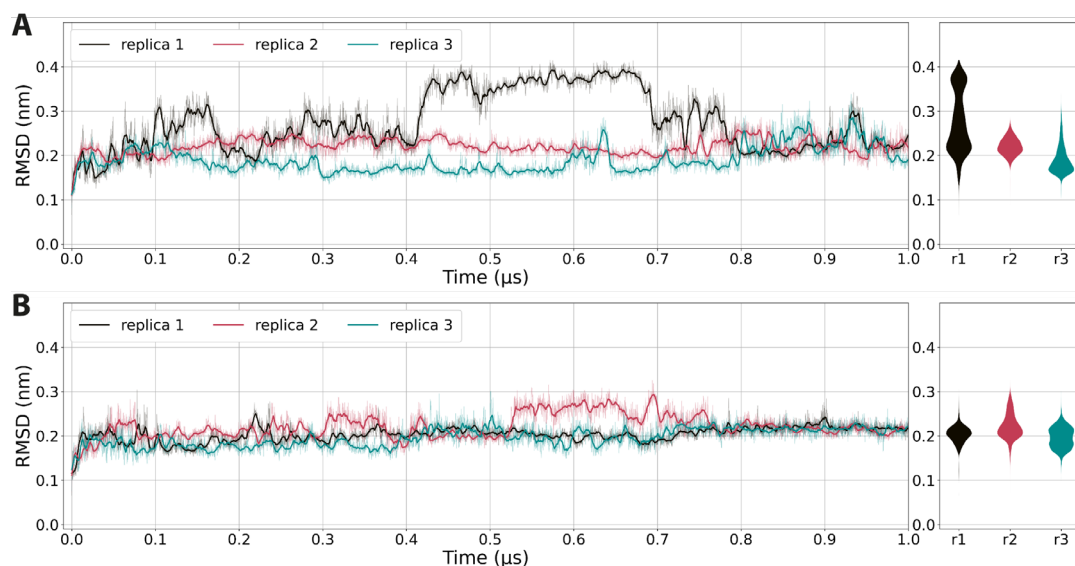


Figure S5. Values of the root mean square deviation from the minimized structures for the C α atoms as a function of the simulation time for OTRb (**A**) and OTRe (**B**) systems excluding the region comprising residues 237-265. RMSD values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the RMSD values for the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

E42 and D100 hydrogen bond network.

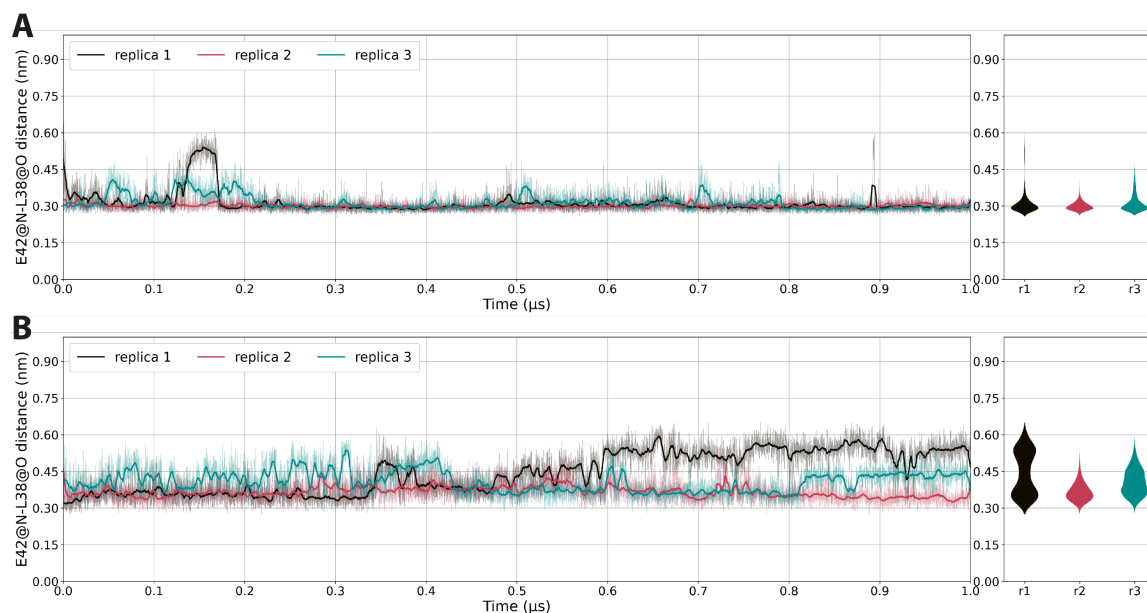


Figure S6. E42(N)-L38(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicates simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicates of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

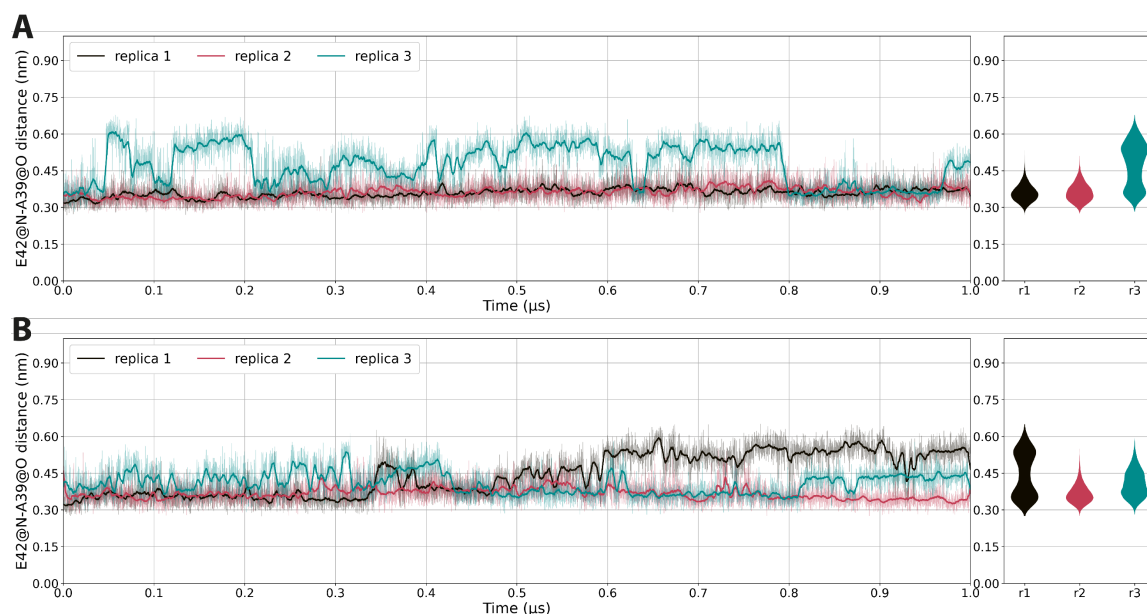


Figure S7. E42(N)-A39(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicates simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicates of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

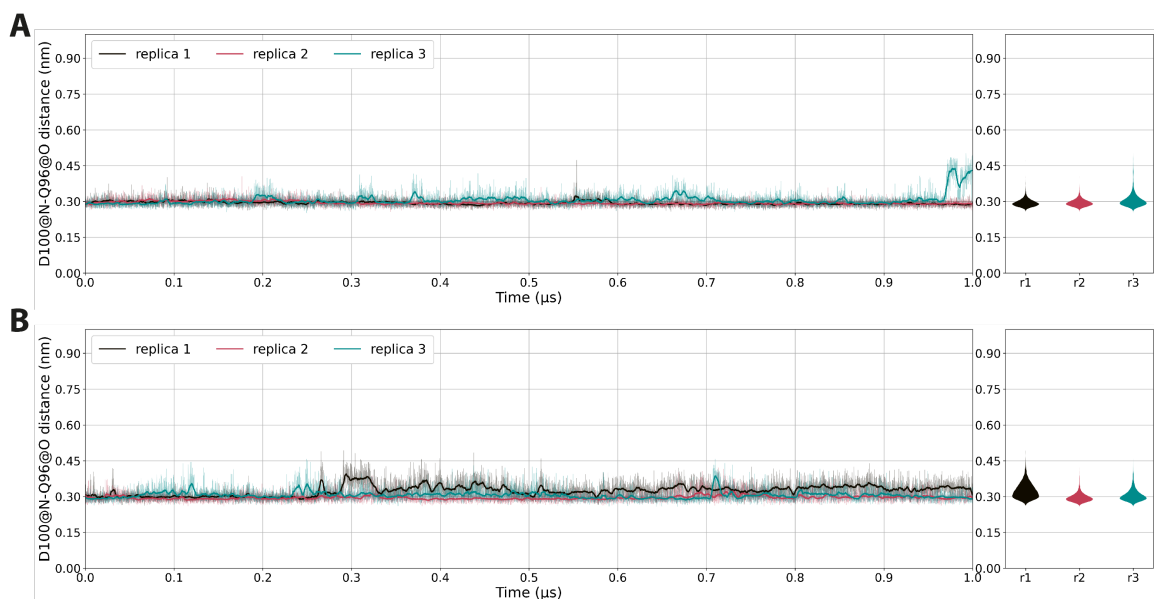


Figure S8. D100(N)-Q96(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

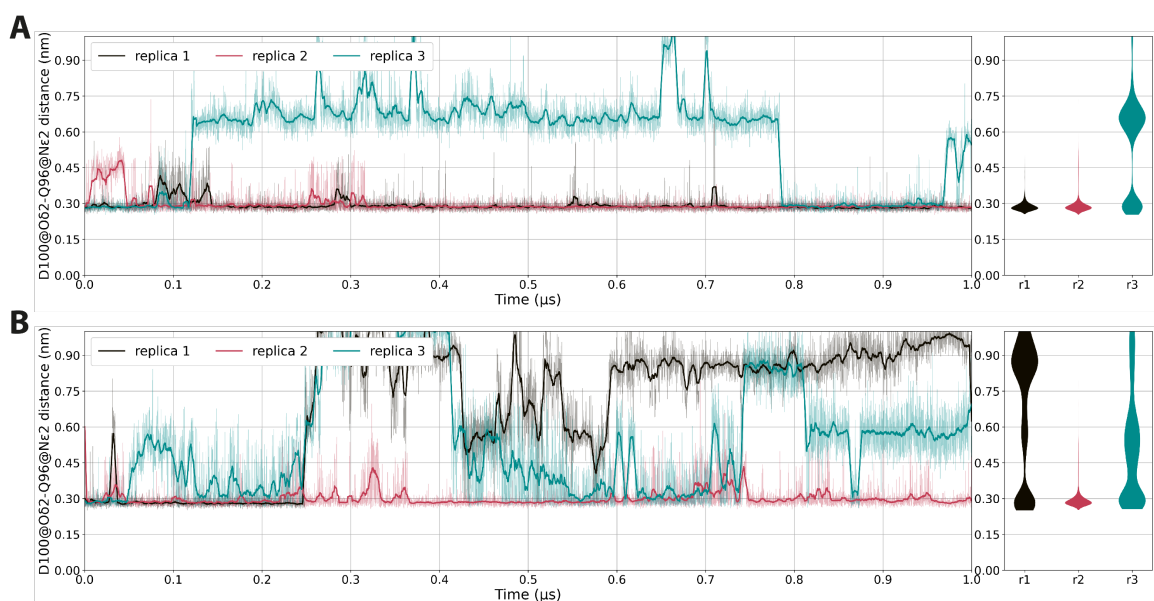


Figure S9. D100(O δ 2)-Q96(N ϵ 2) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

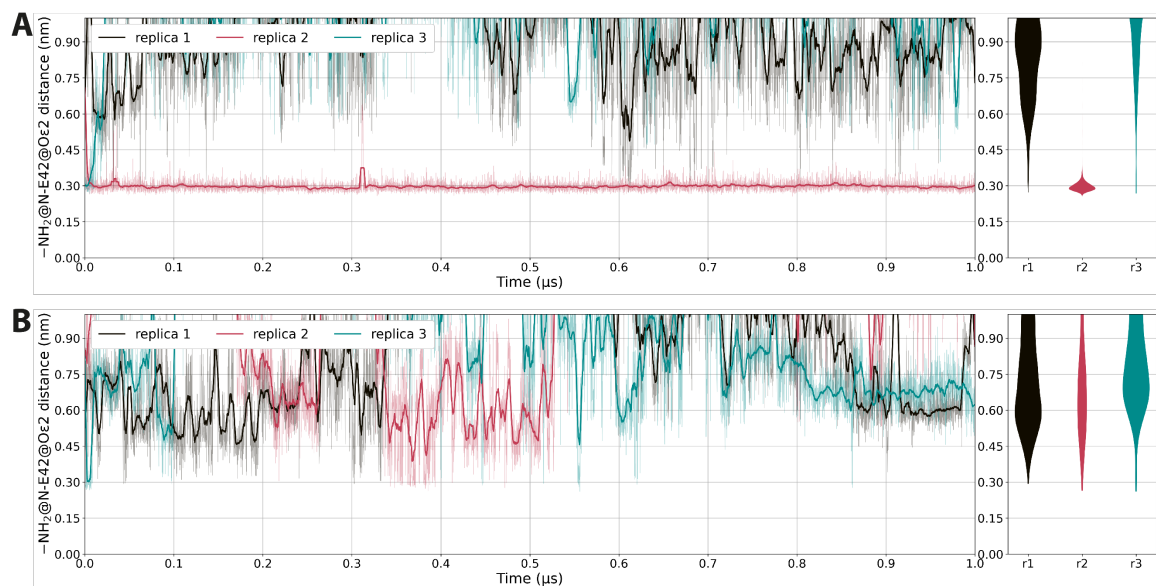


Figure S10. E42(Oε2) - OT -NH₂ nitrogen distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

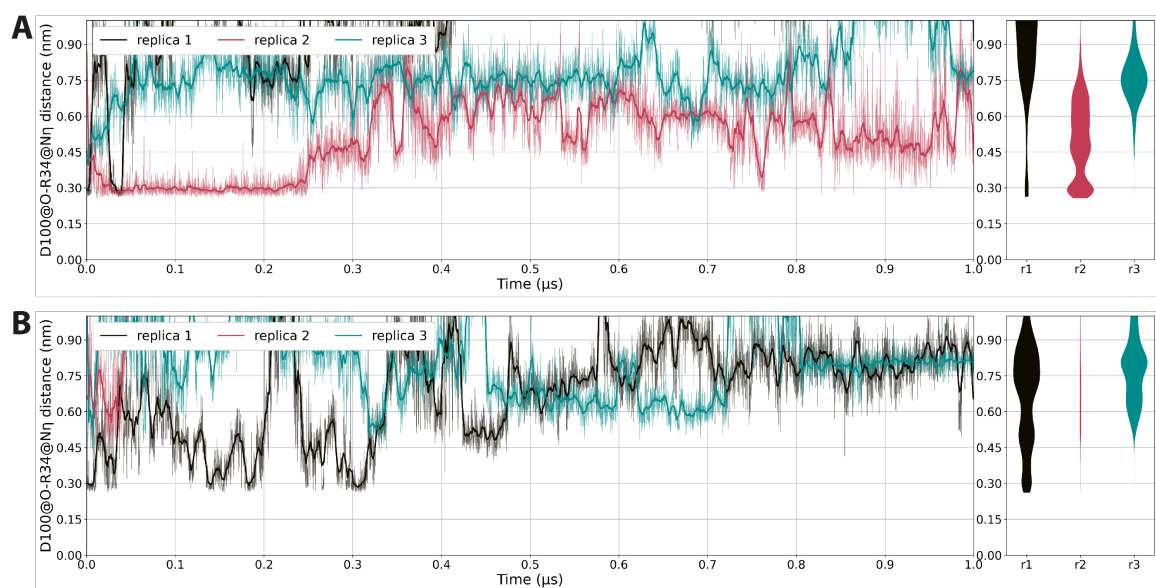


Figure S11. Minimum distance between D100(O) and R34(Nη1) or R34(Nη2) as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

Distance between Mg^{2+} and OT C-terminal residues.

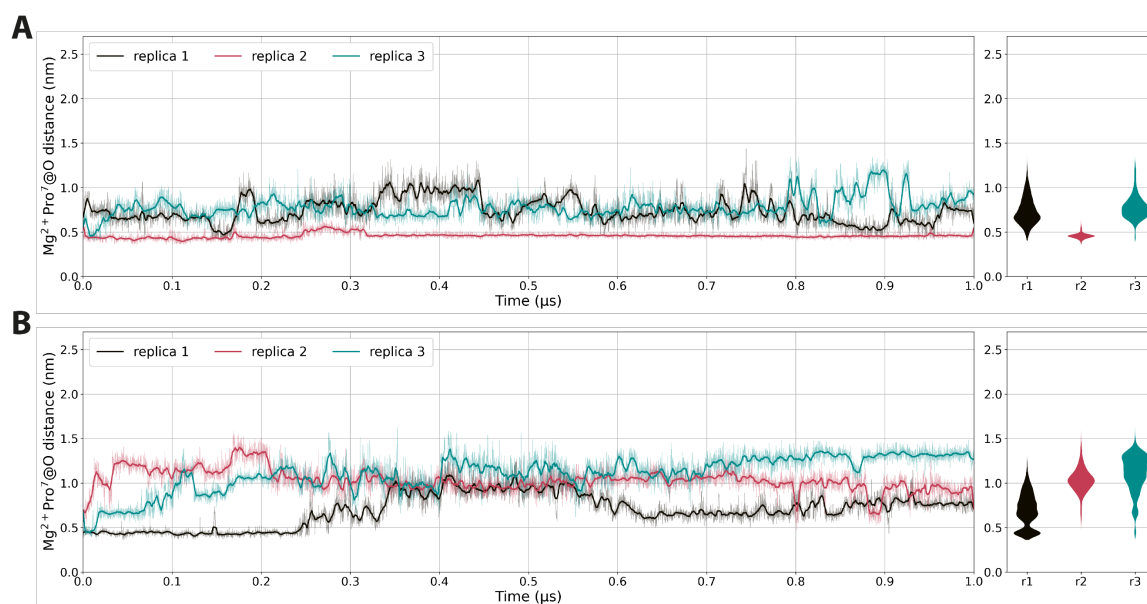


Figure S12. Mg^{2+} -Pro⁷(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

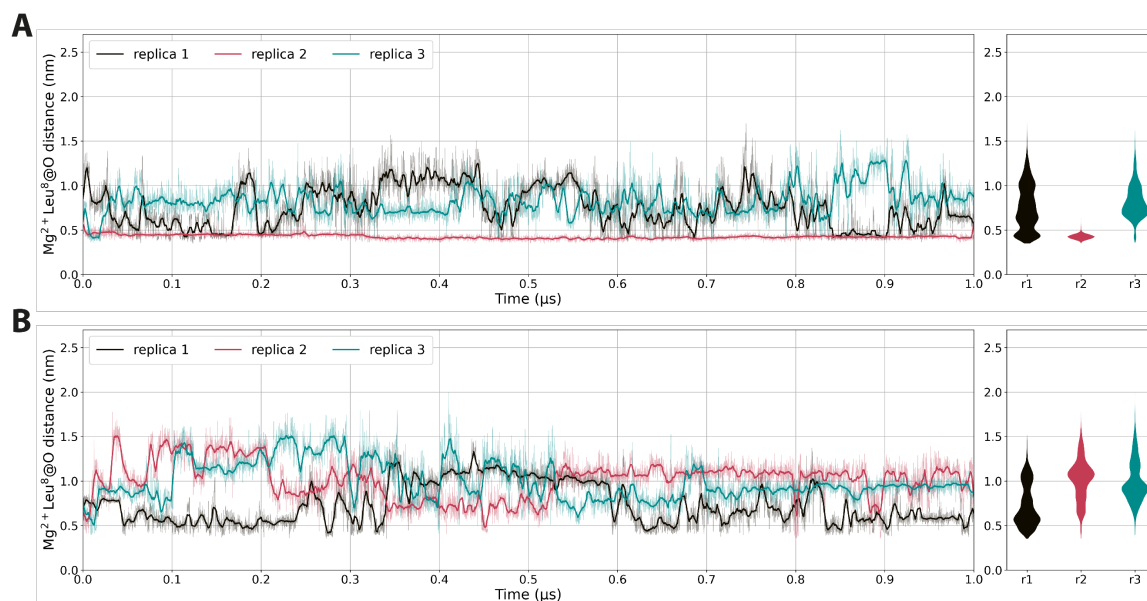


Figure S13. Mg^{2+} -Leu⁸(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

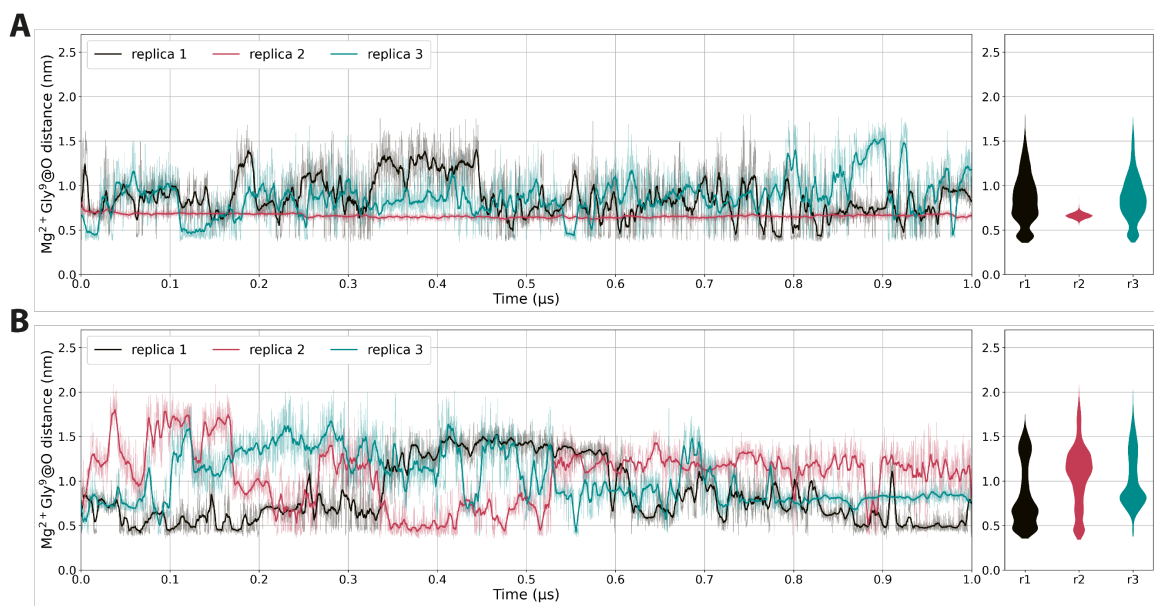


Figure S14. Mg^{2+} -Gly⁹(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky-Golay filter [34] in order to cut out the noise.

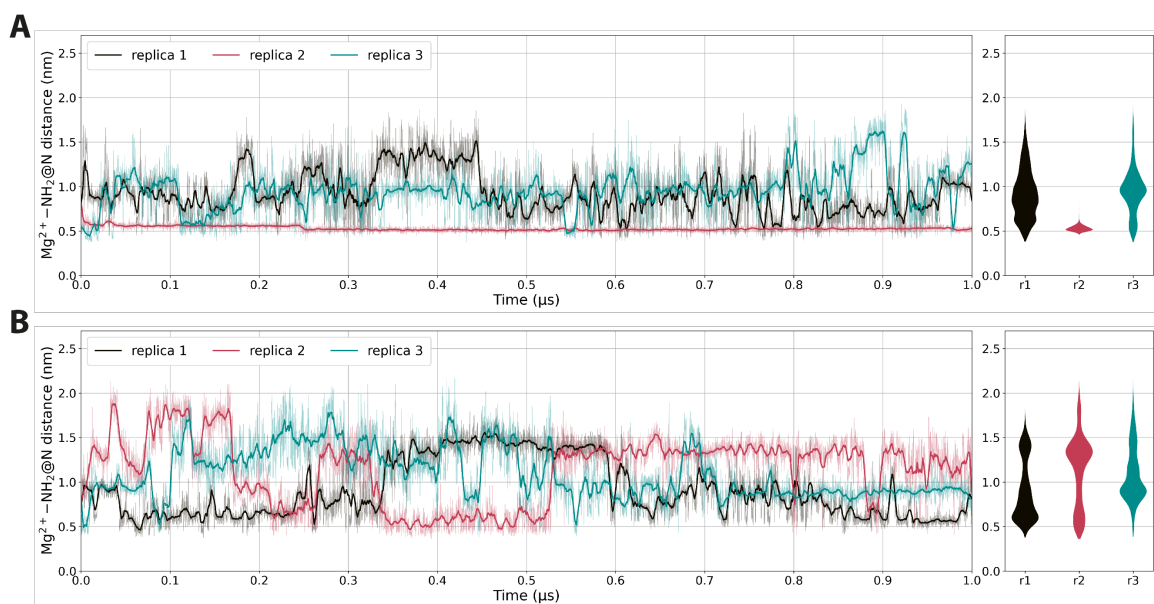


Figure S15. Mg^{2+} -NH₂(N) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky-Golay filter [34] in order to cut out the noise.

Hydrogen bond network formed by water molecules bound to Mg^{2+} .

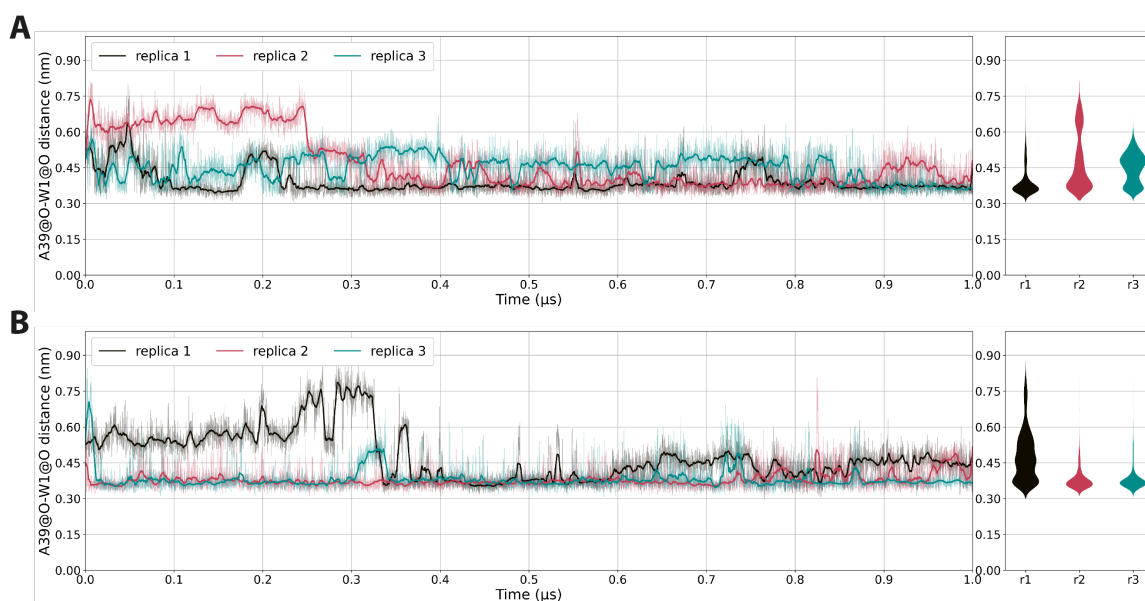


Figure S16. A39(O)-W1(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

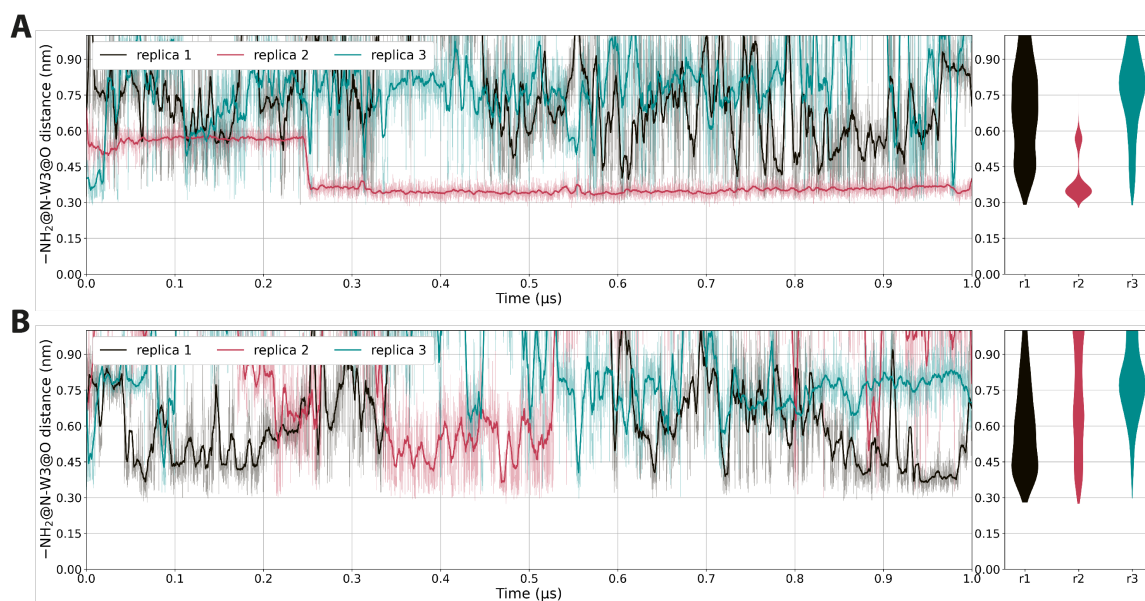


Figure S17. OT -NH₂(N)-W3(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

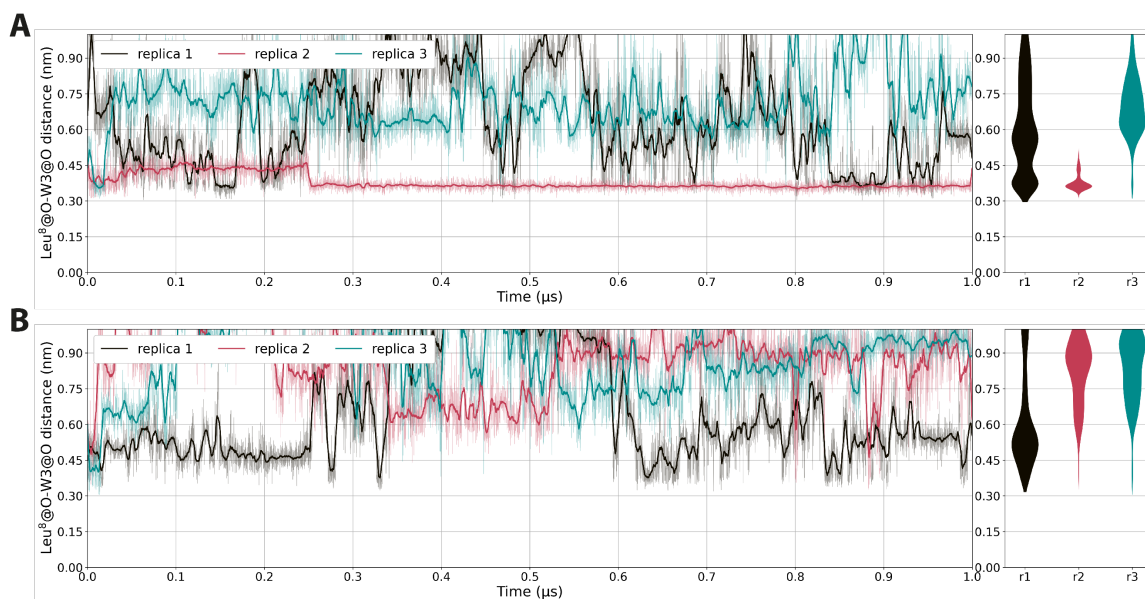


Figure S18. Leu⁸(O)-W3(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

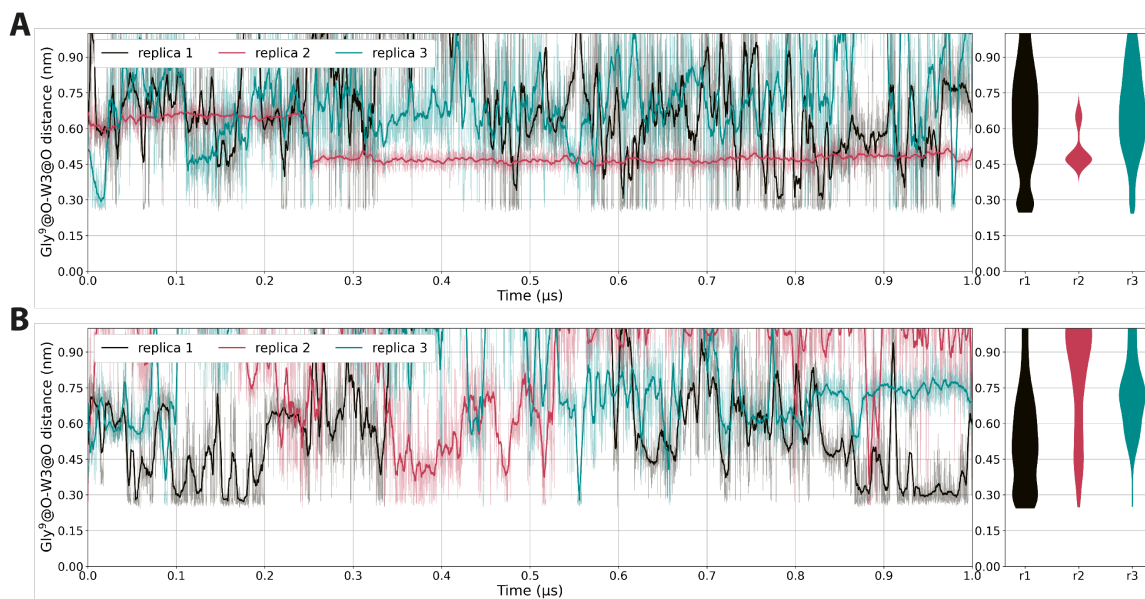


Figure S19. Gly⁹(O)-W3(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

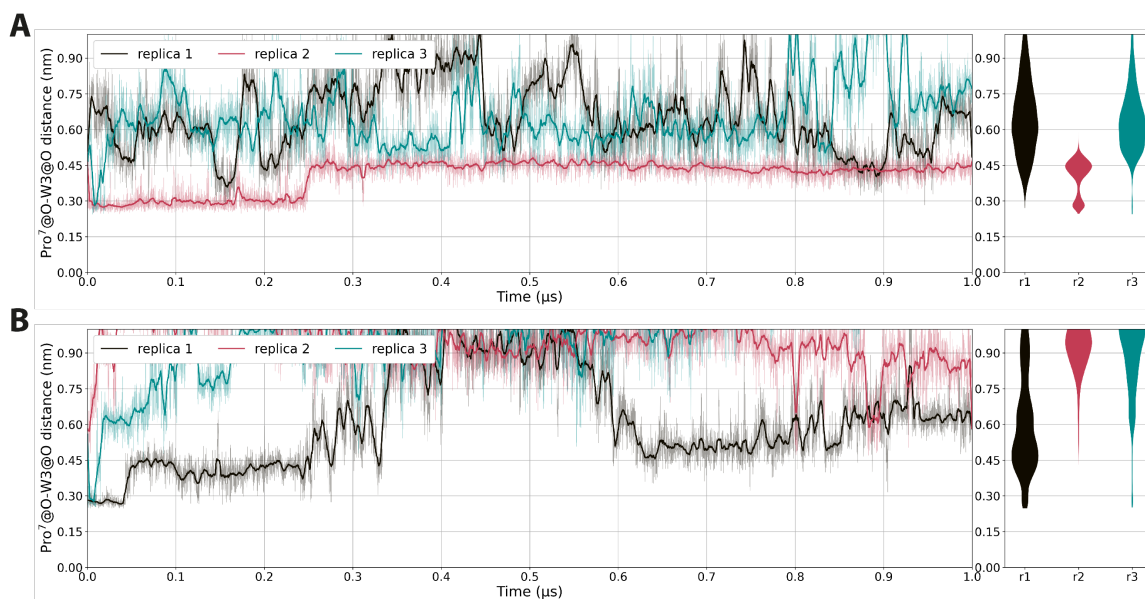


Figure S20. Pro⁷(O)-W3(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

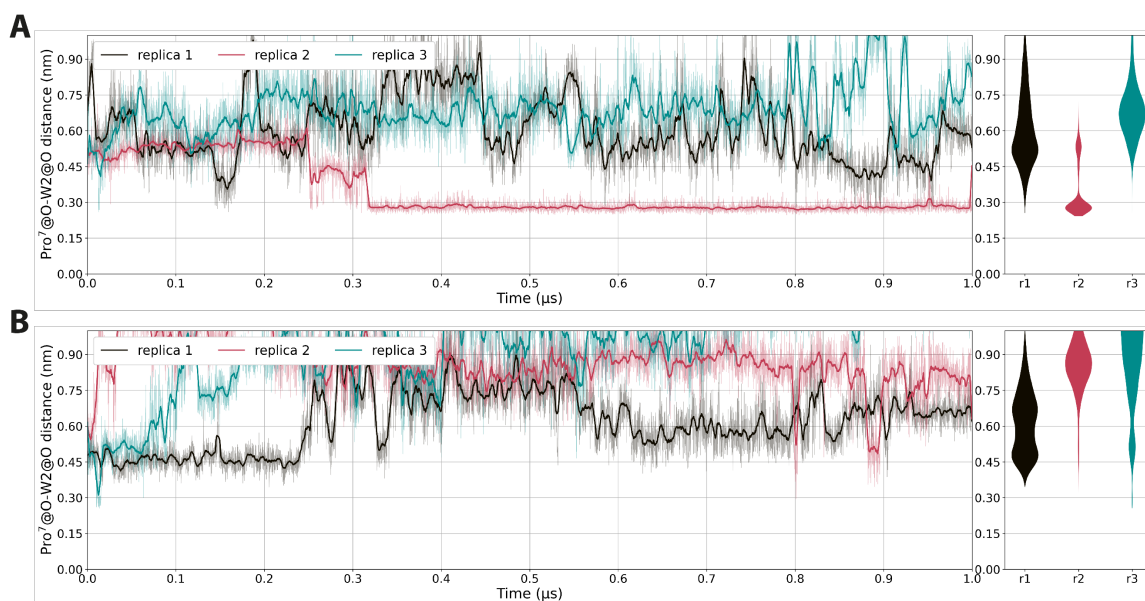


Figure S21. Pro⁷(O)-W2(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

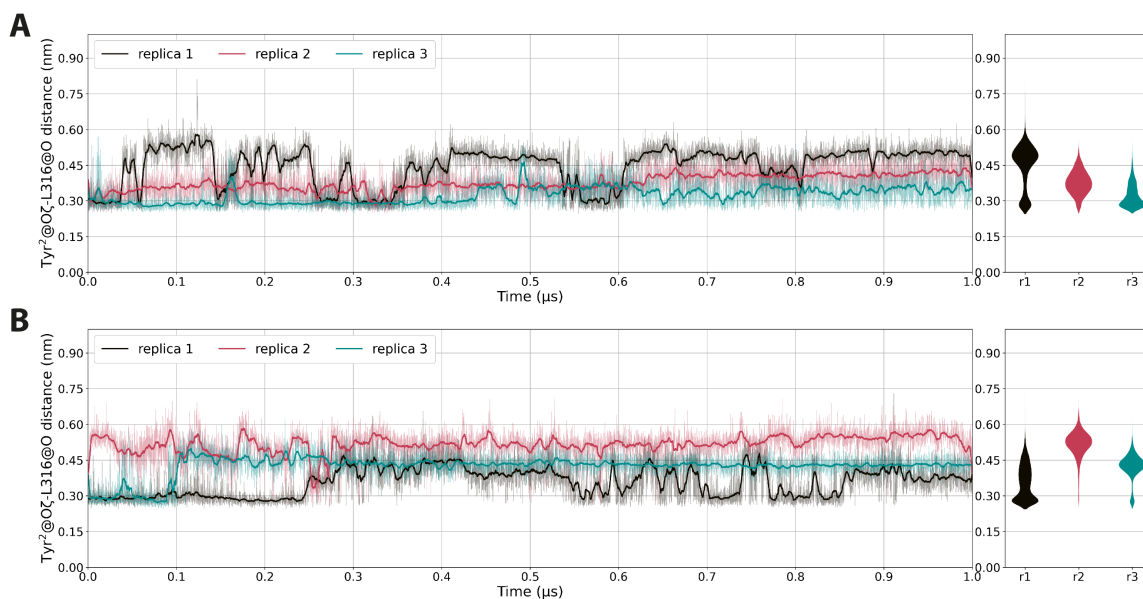


Figure S22. Tyr²(O ζ)-L316(O) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky-Golay filter [34] in order to cut out the noise.

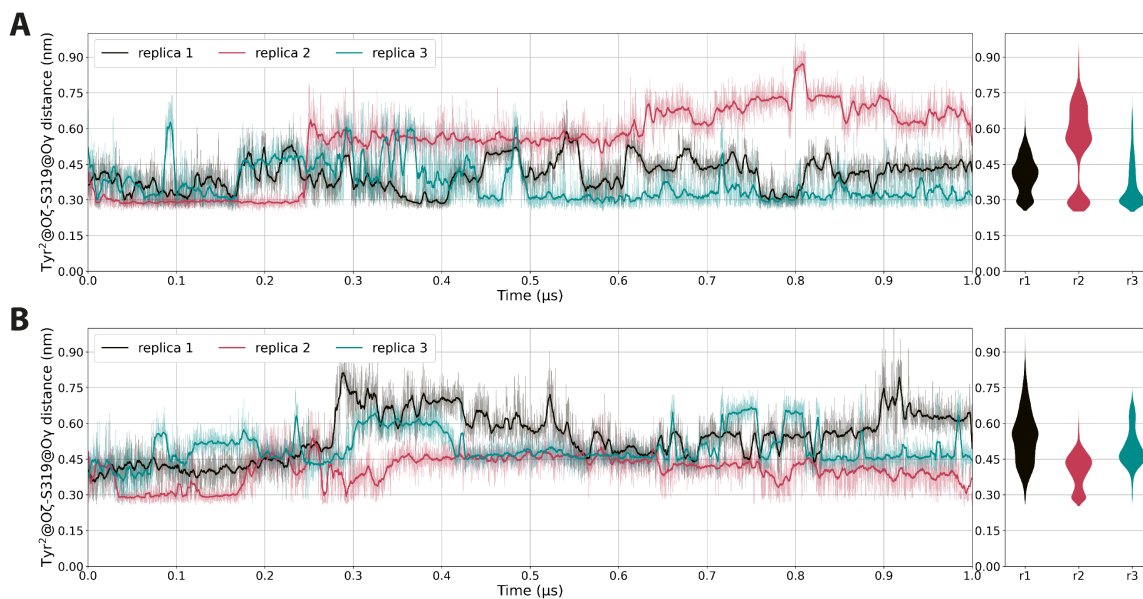


Figure S23. Tyr²(O ζ)-S319(O γ) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky-Golay filter [34] in order to cut out the noise.

R34 – F103 distance.

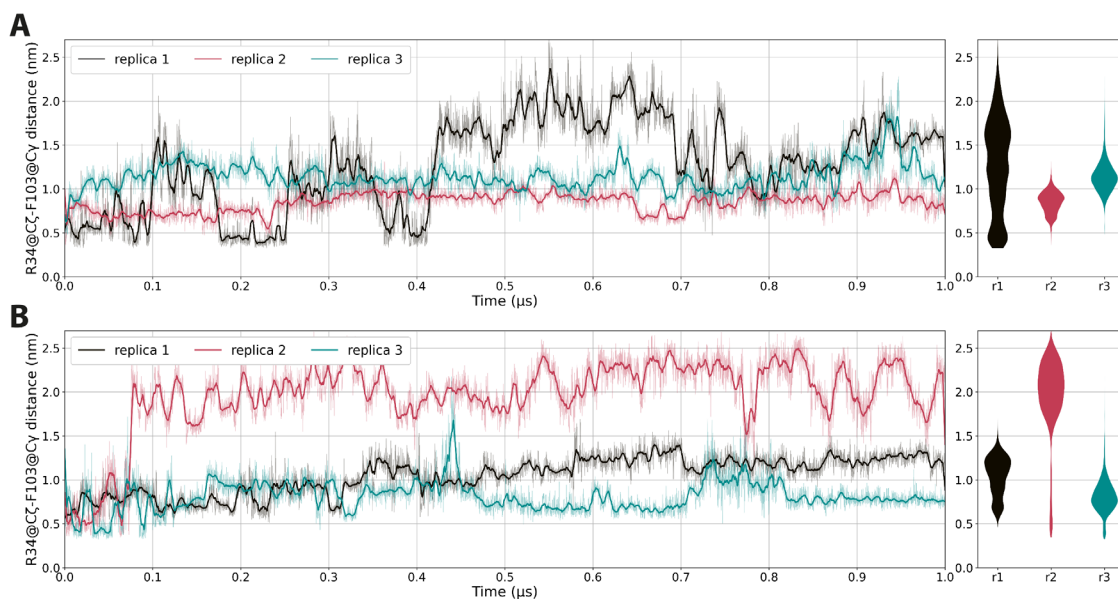


Figure S24. E45(C ζ)-F103(C γ) distance as a function of the simulation time for OTRb (A) and OTRe (B) systems. Values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the distance values in the three replicas of each system. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

Water molecules in the OT binding cavity. The water molecule added by the CHARMM-GUI Generate Pore Water function does not appear to have a relevant role and is replaced by other equally mobile water molecules after a few nanoseconds in all performed simulations.

A¹⁰⁰ activation index. In order to classify the activation status of the OTR along the simulations (carried out in the absence of the G_q protein), the A¹⁰⁰ activation index [31] was computed as a function of the simulation time and it was compared with the same parameter calculated the available OTR experimental structures. These are: *i*) the initial AOM-OTR structure (PDB ID 7RYC, A¹⁰⁰ = +42) [1]; *ii*) the cryo-EM structure of active OTR in complex with OT and a heterotrimeric G_q protein chimera (PDB ID 7QVM, AO-OTR hereafter, A¹⁰⁰ = +29) [35]; and *iii*) the X-ray OTR structure in the inactive form achieved in complex with retosiban, a nonpeptidic antagonist developed as an oral drug for the prevention of preterm labour (PDB ID 6TPK, IA-OTR hereafter, A¹⁰⁰ = -13) [36]. Assuming A¹⁰⁰ = +25 as the boundary between states (A¹⁰⁰ ≥ +25 = active), as proposed for other class A GPCRS [30], the analysis of the A¹⁰⁰ activation index (Figure S24) shows: *i*) All simulations show a decrease in the A¹⁰⁰ activation index compared to the starting structure on a timescale of a few nanoseconds if not already during the equilibration phase; and *ii*) only one replica among three started from the OTRb system (i.e., R#2) remains active along the simulation, while the three replicates started from the OTRe conformation remain active. This is in line with the TM5-ICL3-TM6 region having been modelled using an inactive template for the initial structure of OTRb, but an active template for OTRe.

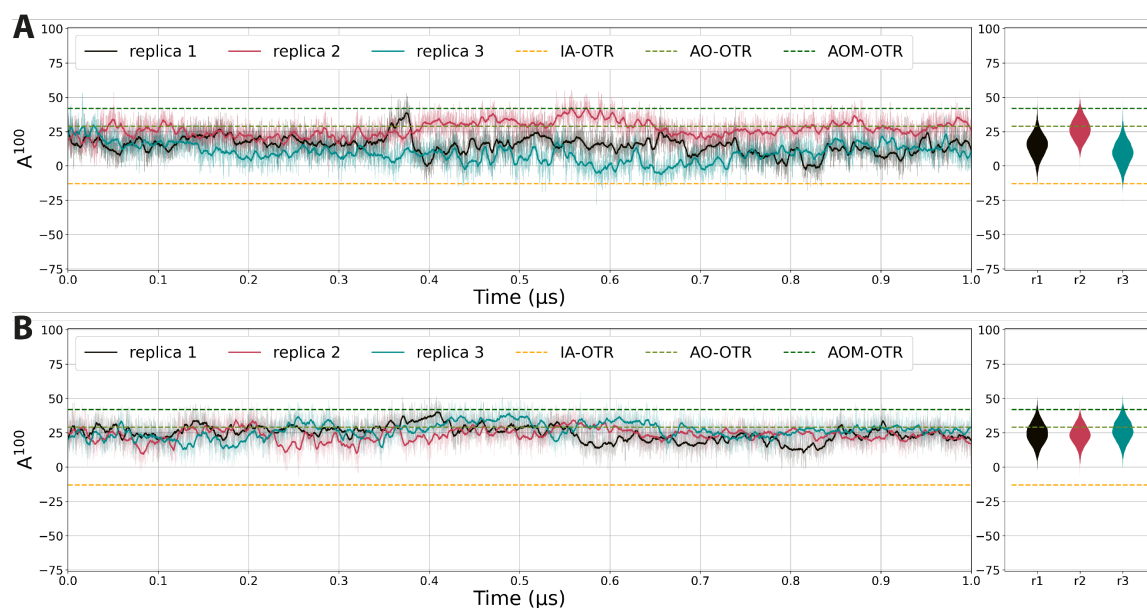


Figure S25. Values of the A¹⁰⁰ activation index as a function of the simulation time for OTRb (A) and OTRe (B) systems. A¹⁰⁰ values for the three replicas simulated for each system are in black, dark red and cyan, respectively. The right panels report the distribution of the A¹⁰⁰ values for the three replicas of each system. The A¹⁰⁰ values for the experimental active AOM-OTR structure as well as for the active AO-OTR and the inactive IA-OTR cryo-EM structures are indicated by horizontal dark green, light green and yellow dashed line, respectively. The bold lines were obtained by applying a Savitzky–Golay filter [34] in order to cut out the noise.

Cluster analysis. A cluster analysis was performed on each replica in accordance with the gromos algorithm [37] by using the *cluster* tool included in GROMACS. For each replica trajectory, 10,000 frames were analysed and clustered using a cut-off of 0.15 nm.

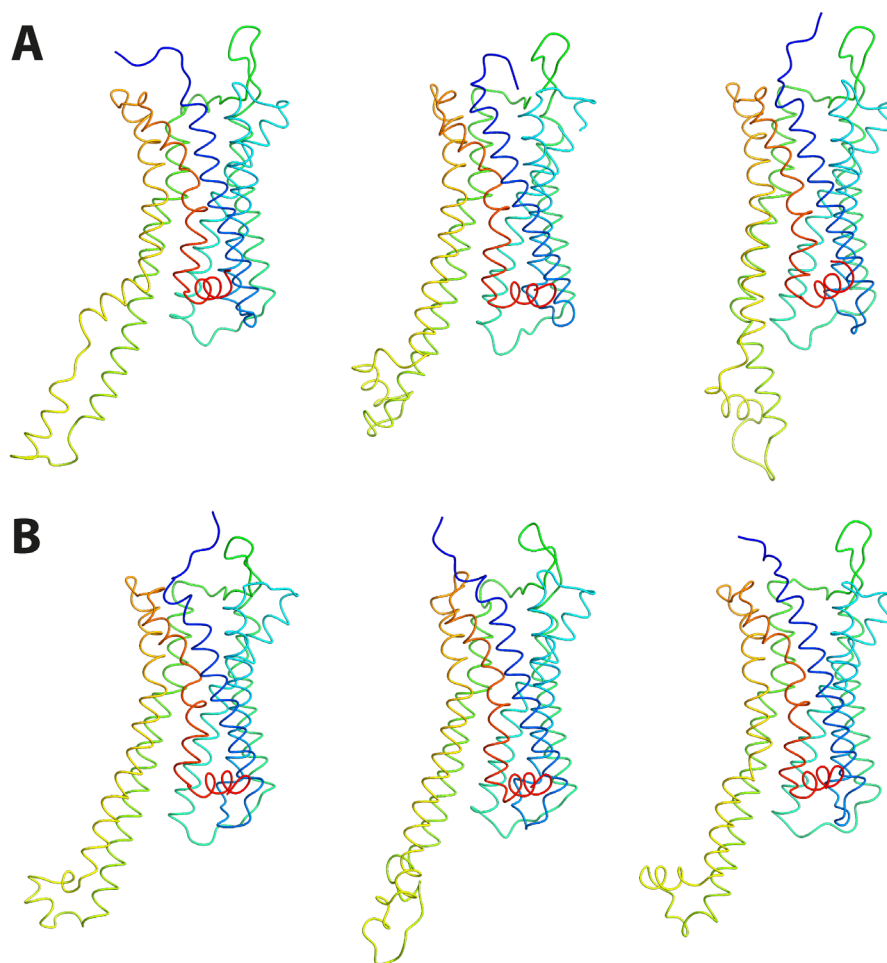


Figure S26. Cluster analysis on OTRb (A) and OTRe (B) simulations. Representative structures of the most populated clusters for R#1, R#2, and R#3 are reported in each panel from left to right, respectively. The ribbons are colored from blue to red going from the N- to the C-terminus.

OTR and OT interactions with the Mg^{2+} ion second coordination sphere. In the AOM-OTR structures, OT binds to OTR by burying its N-terminal end in a deep extracellular pocket of the receptor. At the same time, the C-terminal end of OT remains partially exposed to the solvent and interacts with the receptor's N-terminal region (Figure S1). During the simulations, this latter interaction was monitored by measuring the distance between the Mg^{2+} ion and R34 C ζ atom as a function of the simulation time (Figure 3B). Moreover, the Mg^{2+} -R34(C ζ) has been related to the A^{100} activation index for each simulation frame (Figure 4). From the analysis of the two OTRb system it appears that a movement of R34 – and thus of the entire OTR N-terminal fragment apart from the Mg^{2+} ion second coordination sphere – is a conformational property observed when the receptor is inactive. However, in one replica of the OTRe system, R34 appears to be able to move apart from the Mg^{2+} ion without causing OTR inactivation. A similar behaviour is observed monitoring the Mg^{2+} -Pro⁷(O) distance as a function of the simulation time (Figure S12) and of the A^{100} activation index (Figure 4). Moreover, visual inspection of the simulations and the analysis of selected distances (Figures 3A) revealed that, even if in some cases Pro⁷ goes away from the Mg^{2+} ion coordination sphere, it is replaced by other OT atoms that are part of the nonapeptide C-terminal region (i.e., Leu⁸(O), Gly⁹(O) and the nitrogen of the amide terminal group). In other words, the C-terminal region of the OT remains in the close proximity to the Mg^{2+} ion coordination sphere throughout most of the simulation time, albeit changing its specific interactions over time, as highlighted by the analysis of the minimum distance between the OT peptide and the Mg^{2+} ion (Figures 3A).

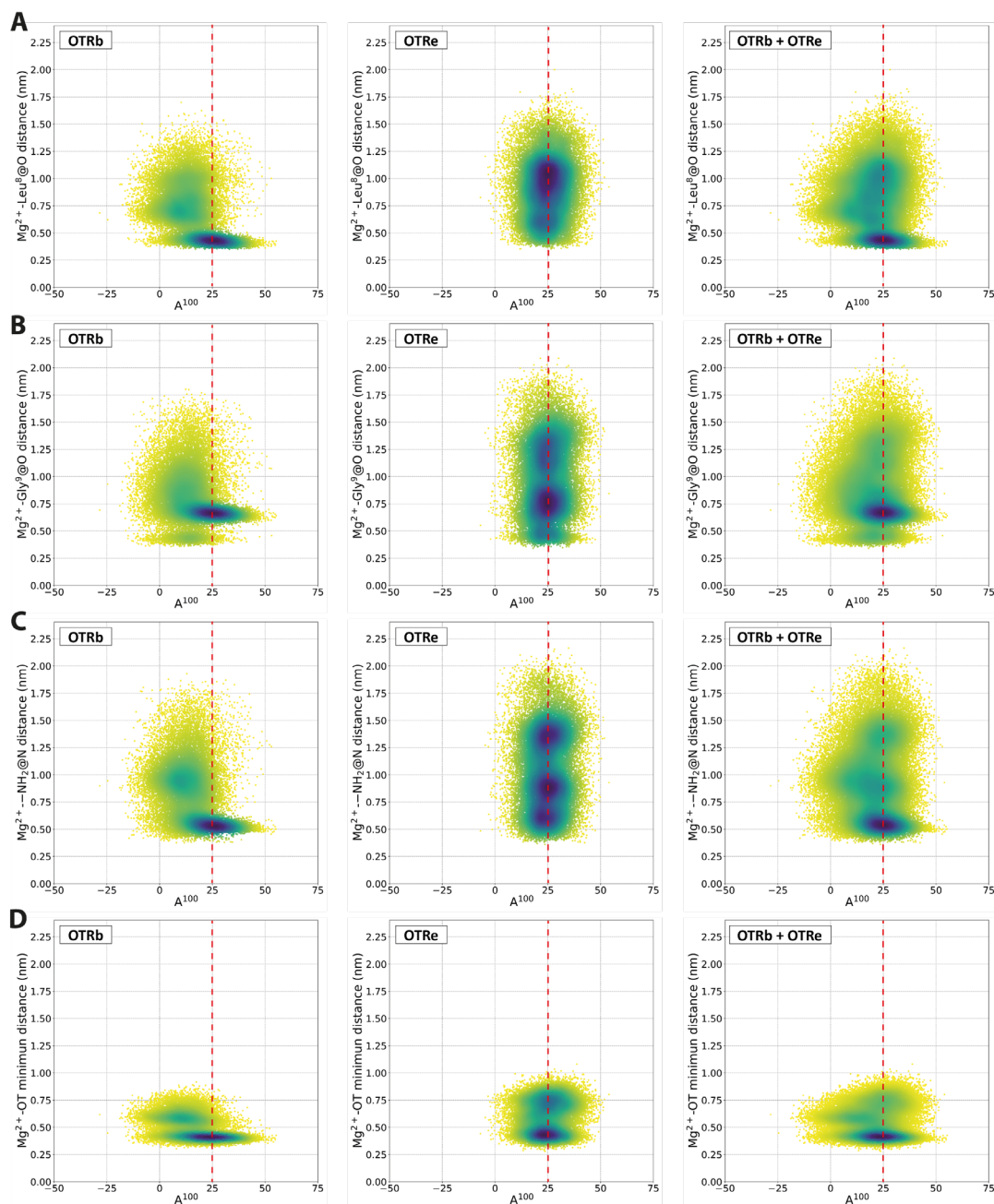


Figure S27. Distances between the Mg^{2+} ion and: $\text{Leu}^8(\text{O})$ (**A**), $\text{Gly}^9(\text{O})$ (**B**), the N in the OT amide terminal group (**C**), and OT (**D**), reported as a function of the A^{100} activation index for each frame of the OTRb and OTRe simulations (left and central panels) and by summing up the simulations on both system (right panel). The dot's colour changes from yellow to dark blue as their density increases. The vertical red dashed line highlights the A^{100} activation index boundary between states ($A^{100} \geq +25 = \text{active}$).

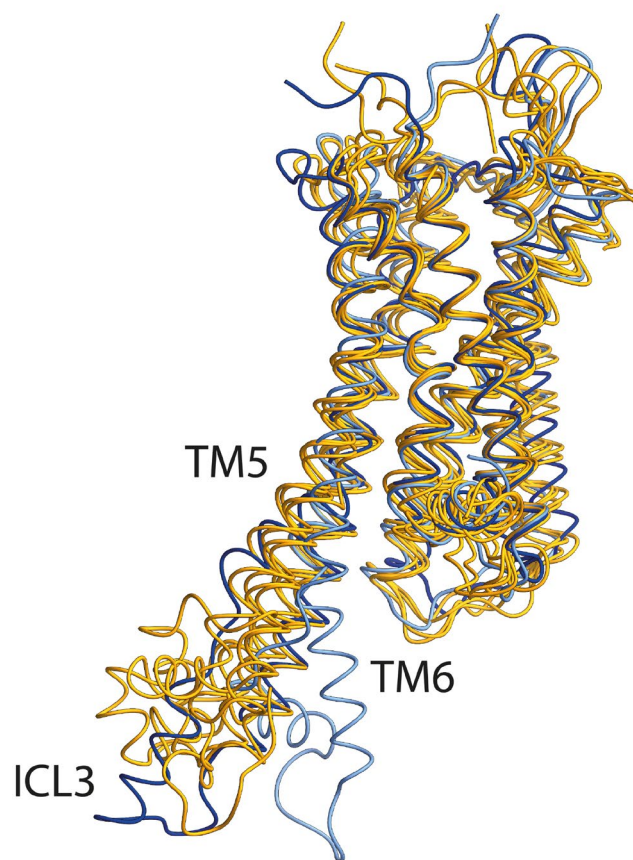


Figure S28. Superimposition between the most representative structures achieved in the cluster analysis on each replica. Active replicas according to the A^{100} activation index (OTRe R#1-3 and OTRb R#2) are in orange, while inactive replicas (OTRb R#1 and R#3) are in dark and light blue, respectively.

SUPPLEMENTARY REFERENCES

- [1] J.G. Meyerowitz, M.J. Robertson, X. Barros-Alvarez, O. Panova, R.M. Nwokonko, Y. Gao, G. Skinotis, *Nat Struct Mol Biol*, vol. 29, 2022, pp. 274-281.
- [2] A.C. Wallace, R.A. Laskowski, J.M. Thornton, *Protein Eng*, vol. 8, 1995, pp. 127-134.
- [3] R.A. Laskowski, M.B. Swindells, *J Chem Inf Model*, vol. 51, 2011, pp. 2778-2786.
- [4] B. Webb, A. Sali, *Curr Protoc Bioinformatics*, vol. 54, 2016, pp. 5 6 1-5 6 37.
- [5] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, 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. Hassabis, *Nature*, vol. 596, 2021, pp. 583-589.
- [6] L. Heo, M. Feig, *Proteins*, vol. 90, 2022, pp. 1873-1885.
- [7] G. Pandey-Szekeres, J. Caroli, A. Mamyrbekov, A.A. Kermani, G.M. Keseru, A.J. Kooistra, D.E. Gloriam, *Nucleic Acids Res*, vol. 51, 2023, pp. D395-D402.
- [8] M.Y. Shen, A. Sali, *Protein Sci*, vol. 15, 2006, pp. 2507-2524.
- [9] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, *Journal of Applied Crystallography*, vol. 26, International Union of Crystallography (IUCr), 1993, pp. 283-291.
- [10] A.L. Lomize, S.C. Todd, I.D. Pogozheva, *Protein Sci*, vol. 31, 2022, pp. 209-220.
- [11] S. Jo, T. Kim, W. Im, *PLoS One*, vol. 2, 2007, pp. e880.
- [12] S. Jo, T. Kim, V.G. Iyer, W. Im, *J Comput Chem*, vol. 29, 2008, pp. 1859-1865.
- [13] I.D. Pogozheva, G.A. Armstrong, L. Kong, T.J. Hartnagel, C.A. Carpino, S.E. Gee, D.M. Picarello, A.S. Rubin, J. Lee, S. Park, A.L. Lomize, W. Im, *J Chem Inf Model*, vol. 62, 2022, pp. 1036-1051.
- [14] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, *J Chem Theory Comput*, vol. 11, 2015, pp. 3696-3713.
- [15] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *The Journal of Chemical Physics*, vol. 79, 1983, pp. 926-935.
- [16] Y. Jiang, H. Zhang, W. Feng, T. Tan, *J Chem Inf Model*, vol. 55, 2015, pp. 2575-2586.
- [17] Y. Li, J. Liu, J.C. Gumbart, *Methods Mol Biol*, vol. 2302, 2021, pp. 237-251.
- [18] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J Comput Chem*, vol. 25, 2004, pp. 1605-1612.
- [19] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, *Computer Physics Communications*, vol. 91, 1995, pp. 43-56.
- [20] E. Lindahl, B. Hess, D. van der Spoel, *Molecular modeling annual*, vol. 7, 2001, pp. 306-317.
- [21] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, *J Comput Chem*, vol. 26, 2005, pp. 1701-1718.
- [22] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, *The Journal of Chemical Physics*, vol. 103, 1995, pp. 8577-8593.
- [23] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, *Journal of Computational Chemistry*, vol. 18, John Wiley & Sons, Ltd, 1997, pp. 1463-1472.
- [24] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, *The Journal of Chemical Physics*, vol. 81, 1984, pp. 3684-3690.
- [25] D.J. Evans, B.L. Holian, *The Journal of Chemical Physics*, vol. 83, 1985, pp. 4069-4074.
- [26] M. Parrinello, A. Rahman, *Journal of Applied Physics*, vol. 52, 1981, pp. 7182-7190.
- [27] B. Knapp, L. Ospina, C.M. Deane, *Journal of Chemical Theory and Computation*, vol. 14, American Chemical Society, 2018, pp. 6127-6138.
- [28] T.A. Soares, Z. Cournia, K. Naidoo, R. Amaro, H. Wahab, K.M. Merz, Jr., *Journal of Chemical Information and Modeling*, vol. 63, American Chemical Society, 2023, pp. 3227-3229.

- [29] R.T. McGibbon, K.A. Beauchamp, M.P. Harrigan, C. Klein, J.M. Swails, C.X. Hernandez, C.R. Schwantes, L.P. Wang, T.J. Lane, V.S. Pande, *Biophys J*, vol. 109, 2015, pp. 1528-1532.
- [30] G. Perez-Hernandez, P.W. Hildebrand, *PLoS Comput Biol*, vol. 21, 2025, pp. e1012837.
- [31] P. Ibrahim, D. Wifling, T. Clark, *J Chem Inf Model*, vol. 59, 2019, pp. 3938-3945.
- [32] T.D. Goddard, C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, T.E. Ferrin, *Protein Sci*, vol. 27, 2018, pp. 14-25.
- [33] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, *Protein Sci*, vol. 30, 2021, pp. 70-82.
- [34] A. Savitzky, M.J.E. Golay, *Analytical Chemistry*, vol. 36, American Chemical Society, 1964, pp. 1627-1639.
- [35] Y. Waltenspuhl, J. Ehrenmann, S. Vacca, C. Thom, O. Medalia, A. Pluckthun, *Nat Commun*, vol. 13, 2022, pp. 4153.
- [36] Y. Waltenspuhl, J. Schoppe, J. Ehrenmann, L. Kummer, A. Pluckthun, *Sci Adv*, vol. 6, 2020, pp. eabb5419.
- [37] X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *Angewandte Chemie International Edition*, vol. 38, John Wiley & Sons, Ltd, 1999, pp. 236-240.