Supporting Information

Conformational dynamics underlying Atypical Chemokine Receptor 3 activation

Omolade Otun[1], Christelle Aljamous[1], Elise Del Nero[1], Marta Arimont-Segura[3], Reggie Bosma[3], Barbara Zarzycka[3], Tristan Girbau[1], Cédric Leyrat[1], Chris de Graaf[3≠], Rob Leurs[3], Thierry Durroux[1], Sebastien Granier[1], Xiaojing Cong[1*] and Cherine Bechara[1,2*]

Supplementary Methods

Reservoir generation for R-REMD

Typically, the reservoir in R-REMD is generated by simulations at a very high temperature to achieve rapid conformational sampling. The conformations in the reservoir need not follow the Boltzmann distribution (1). During R-REMD, these conformations are thermally reweighted through exchanges among the replicas, eventually resulting in converged Boltzmann-weighted ensembles. For membrane proteins, however, very high temperature simulations should be avoided. Therefore, we generated the reservoir by MD simulations using a modified Hamiltonian, similar to those used in the REST2 MD scheme. The "solute" atoms (ACKR3 and the ligands) experienced weaker non-bonded, electrostatic and dihedral interactions to facilitate their conformational changes, whereas the membrane and the solvent maintained their real Hamiltonian. Several trials were performed with differential Hamiltonian modifications until obtaining MD trajectories that satisfied 3 conditions: 1) the reservoir generated thereof exhibited no apparent artefacts, e.g. protein unfolding, deformation, ligand destabilization or unbinding; 2) the potential energy of the reservoir was comparable to that of the highest-temperature replica (450 K) in the R-REMD, to ensure exchange from the reservoir to the replica; and 3) the reservoir sampled ACKR3 activation (e.g. TM6 outward movements) in the agonist-bound form. The same conditions were used to generate the reservoirs for the *apo* the ligand-bound forms, starting from the inactive model and the ligand binding poses determined with the REST2 MD until convergence (Fig. S10A). Only 1 isomer per ligand (the most favourable in the REST2 MD) was considered here. Discarding the first 10 ns for equilibration, 50 ns of the MD trajectory (1000 frames) were collected to create each reservoir.

Pros and cons of REST2 and R-REMD

REST2 and R-REMD are two variants of replica-exchange simulations which enhance conformational sampling by frequent interchanges of conformations among "cold" and "hot" replicas. In REST2, the "hot" effects are simulated by modifying the Hamiltonian of the “solute” (protein and ligands) to facilitate conformational changes (Fig. S8A). Without heating the membrane lipids or the solution, RETS2 avoids membrane phase transition at high temperatures and also significantly reduces the computing costs (2). In R-REMD, a "reservoir" of conformations is pre-computed to quickly sample the possible conformations. The reservoir is then coupled to the hottest replica of REMD to populate the conformational space of all replicas during the simulations (Fig. S8B). This accelerates conformational sampling by at least 5-20 folds (1, 3).

REST2 has the strength to require little *a priori* knowledge of the system. It explores the conformational space efficiently and generates Boltzmann-weighted ensembles. This is particularly beneficial for predicting ligand binding pose. While in standard MD simulations, ligands tend to be stuck at a local energy minimum, REST2 MD can explore multiple ligand positions and numerous receptor conformations in one run. Here, we performed several runs of REST2 MD with different docking poses to identify the most populated binding pose. Nevertheless, achieving convergence remains a challenge for REST2 MD (and for replica-exchange MD in general). In the early phase of the simulations, the "cold" replicas do not properly sample all the energy minima. Although the "hot" replicas can easily travel across emery barriers, we are only interested in the rate-limiting "coldest" replica. Until all the replicas travel through the "cold" and "hot" replica space multiple times, the simulations are not converged and the early phase should be discarded. It is often non-trivial to determine which part of the simulations should be discarded and whether all the minima are properly sampled. Here, this was determined by comparison with the HDX-MS data in a trial-and-error manner.

R-REMD tackles the convergence issue by regularly introducing the precomputed conformations into the REMD (Fig. S8B). The major challenge in R-REMD, however, is to generate a reservoir that covers the relevant conformational space. Here, we already obtained prior knowledge of the system by the REST2 MD/HDX-MS results, in terms of ligand binding pose and receptor conformational changes. Therefore, we could rapidly generate a relevant reservoir for each system. The potential of relying solely on R-REMD to identify the ligand binding pose (which governs the subsequent receptor conformational sampling) is a subject for future studies.

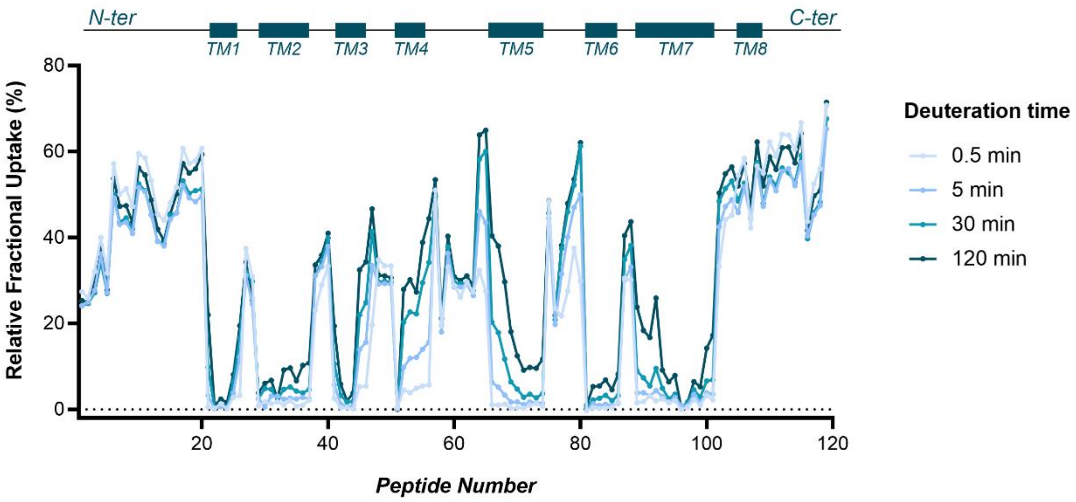


Figure S1: Relative fractional uptake for apo ACKR3 peptides.

Plot of representative peptides identified in HDX experiments and their total % relative fractional uptake across tested deuteration times (0.5, 5, 30 and 120 min). Graph shows uptake for *apo* ACKR3, detected peptide regions are according to the included scheme.

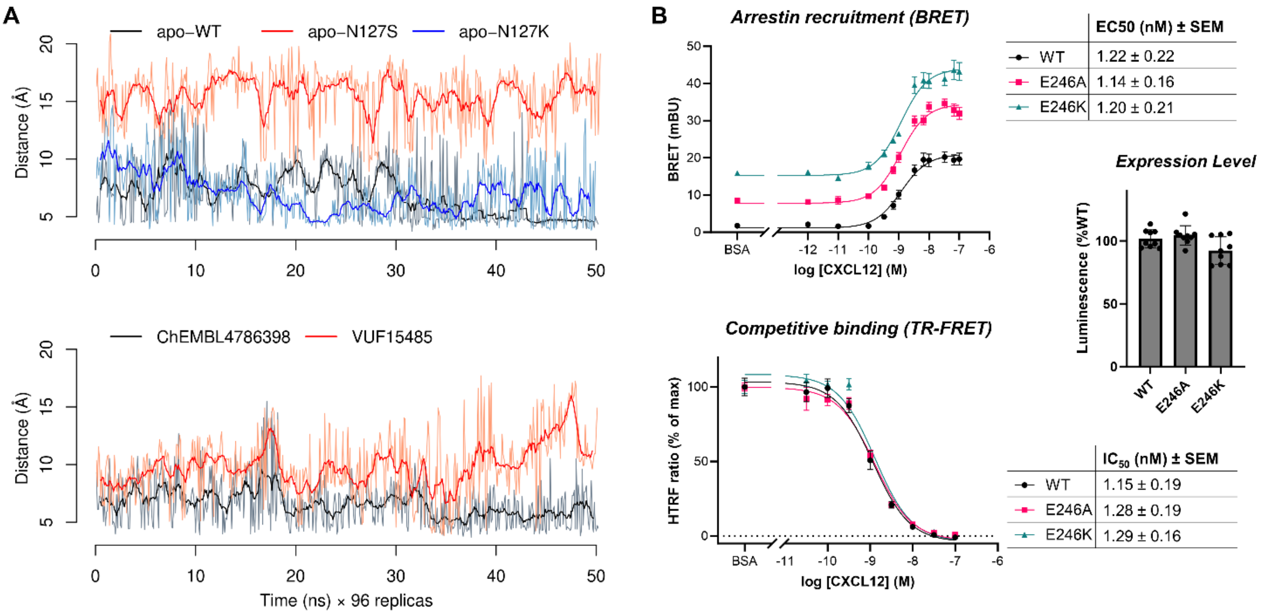


Figure S2: The conserved ionic lock in ACKR3 inactive state

A) Plots of the ionic-lock distance (between R142$^{3.50}$-E246$^{6.29}$ charged groups) in each simulation system. **B)** Recruitment of β-arrestin 1 to ACKR3 ionic-lock mutants followed across a concentration range of agonist CXCL12 (top panel) and TR-FRET competitive binding assay followed between receptor and CXCL12-d2 in the presence of a concentration range of unlabelled CXCL12 (bottom panel). Histogram shows the nLuc luminescence for ACKR3 ionic-lock mutants normalised to WT levels. Data is a composite of three separate experiments measured in triplicate. Error bars represent mean ± SEM across the three experiments.

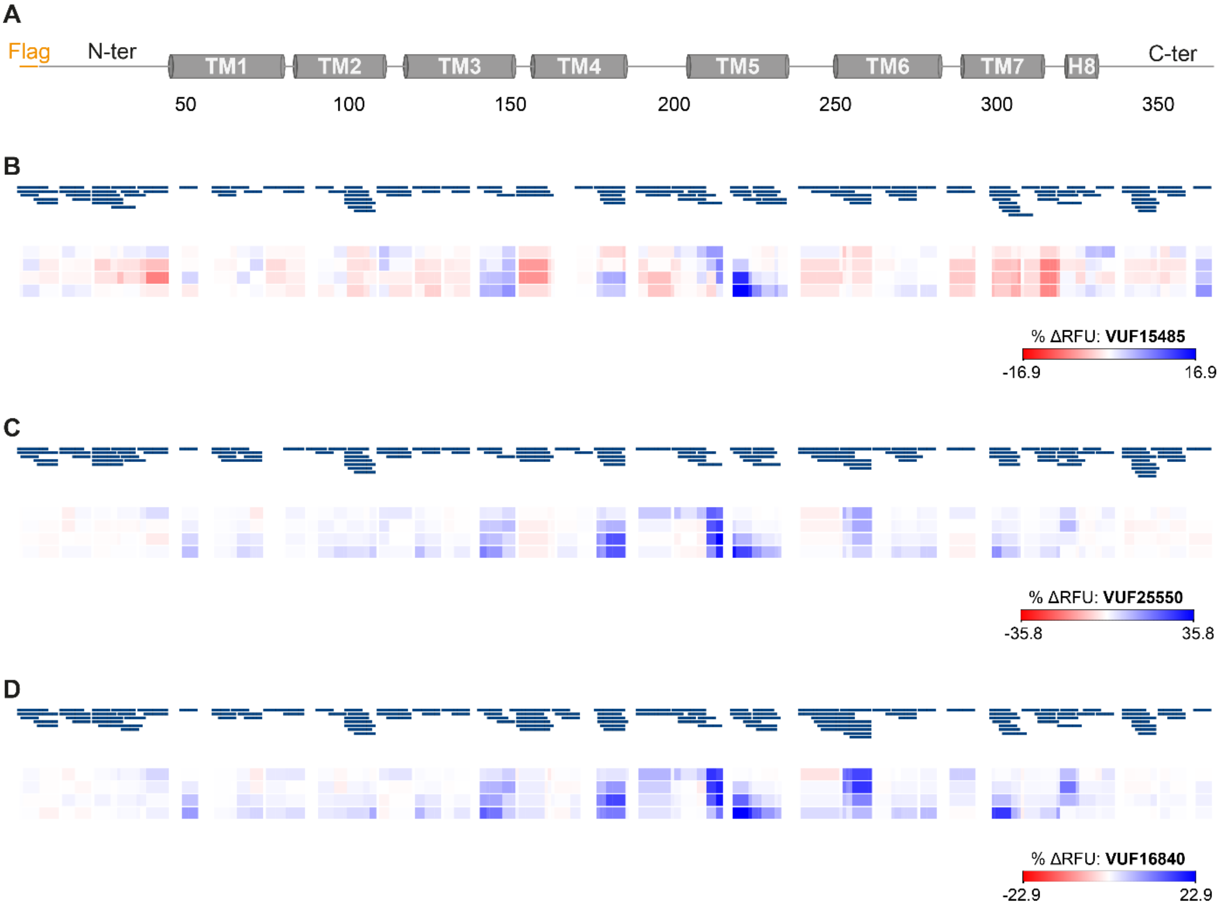


Figure S3: Coverage maps and representative heat maps for tested ligands.

A) Scheme representing transmembrane helices in correspondence to residue numbers. **B-D)** Representative relative fractional uptake differences (%) and coverage maps for HDX analysis of ACKR3 in the presence of all tested ligands (apo – ligand-bound ACKR3). Deuteration was performed at 0.5, 5, 30 and 120 min for all states. For additional information, see supported Data S1 and S2 (xls files).

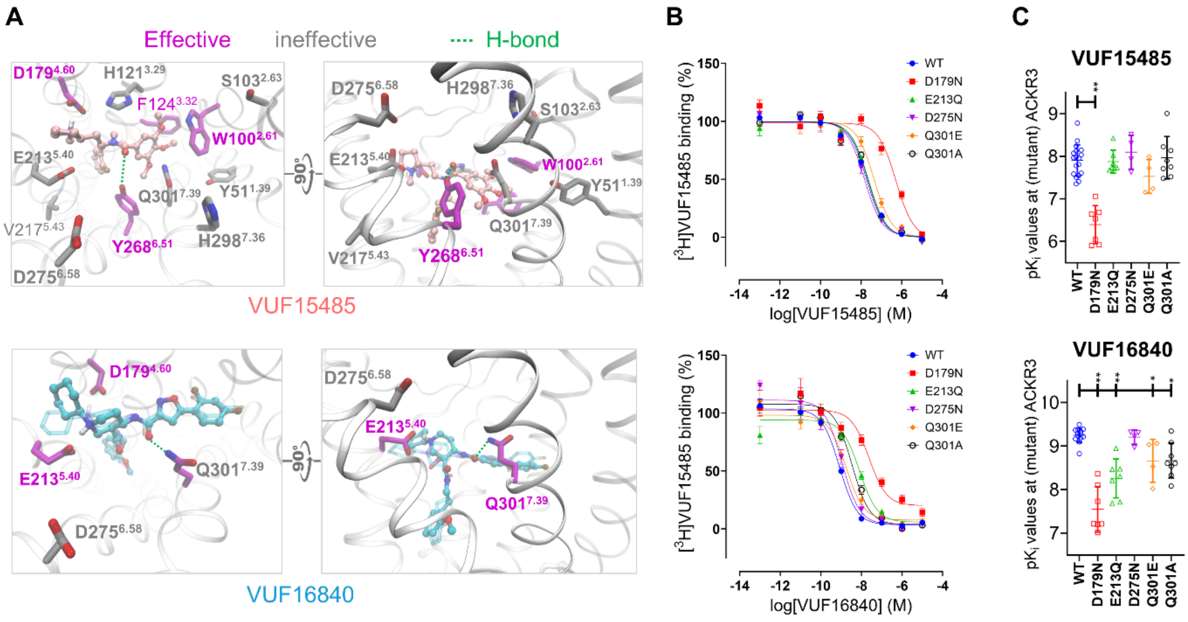


Figure S4: Site-directed mutagenesis data validated the ligand binding poses.

A) Binding-site residues mutated in this study and in ref. (4) are illustrated in sticks. Residues that showed significant impact on ligand binding upon mutation are coloured in magenta (effective). The other mutation sites are coloured in grey (ineffective). Note that mutations H121A$^{3.29}$ and Q301A/E$^{7.39}$ had negligible impacts on VUF15485 binding despite their close proximity to the ligands. Both residues displayed high side-chain flexibility during the MD simulations of the ACKR3-VUF15485 complex, suggesting limited contributions to the ligand binding. **B)** Dose-dependent displacement of radioligand [^{3}H]VUF15485 by VUF15485 (top) and VUF16840 (bottom) in HEK293T membranes transiently expressing WT or mutant ACKR3. The percentage of residual probe binding was normalized to the top and bottom plateau of VUF15485. Data are mean ± SEM of ≥ 3 experiments performed in triplicates. **C)** pK$_i$ values with the respective means (see table S1 for exact values). Statistical significance was analysed with one-way ANOVA using Dunnett multiple comparison test (all were compared to WT). * $P < 0.05$; ** $P < 0.001$.

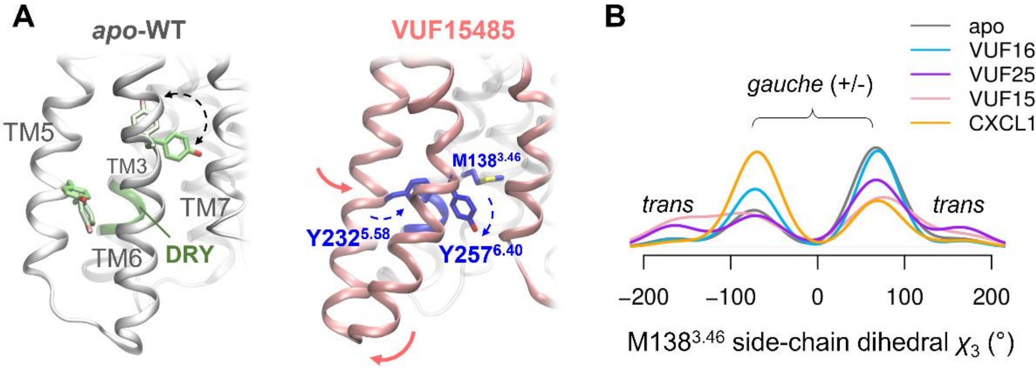


Figure S5: Protection of the DRY motif in VUF15485-bound ACKR3 during MD simulations.

A) MD simulations of ACKR3 in it *apo* and VUF15485-bound forms, showing that the DRY motif is shielded from the solvent by Y232$^{5.58}$ and Y257$^{6.40}$ upon agonist binding despite the TM6 opening. While TM6 moved outward, TM5 moved inward and provided additional protection to the DRY motif through intramolecular contacts. In the *apo* form, Y232$^{5.58}$ pointed outside the 7TM bundle, while Y257$^{6.40}$ pointed upward or laid above the DRY motif. In each form, Y257$^{6.40}$ showed two orientations during the simulations. **B)** Density distribution of M138$^{3.46}$ side-chain dihedral angle χ_3 during the MD simulations.

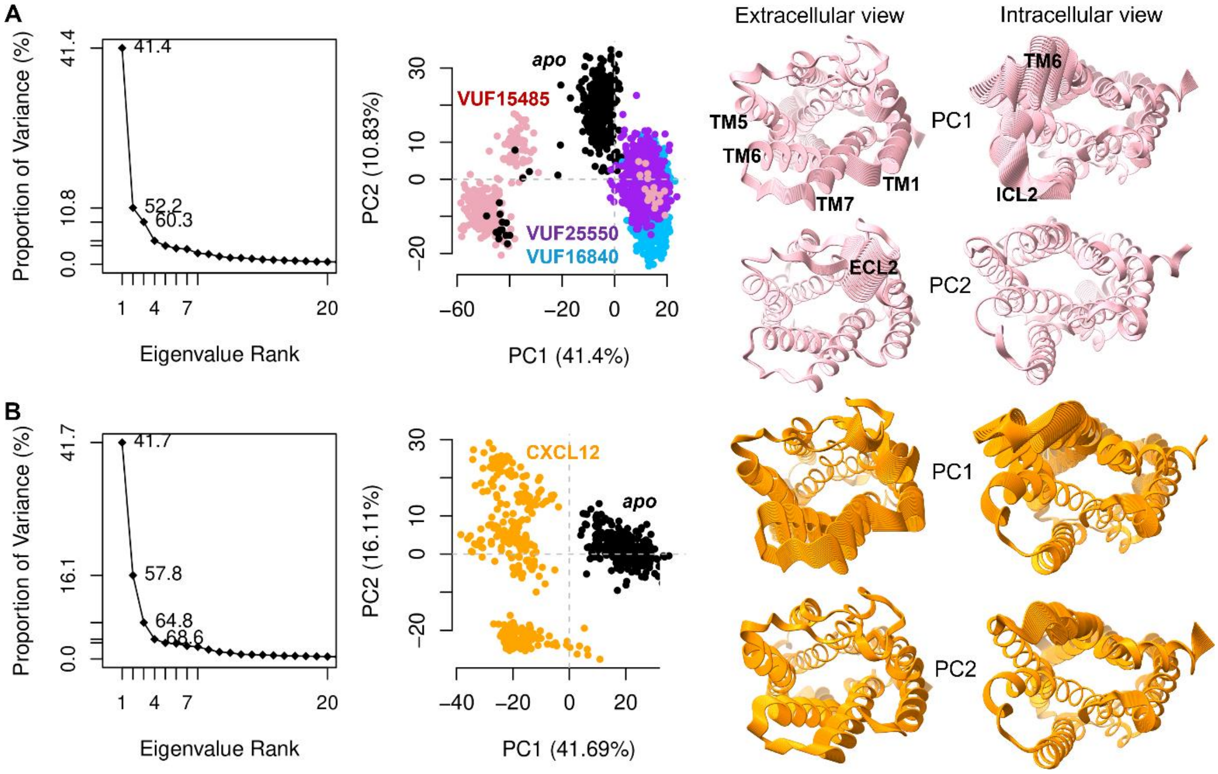
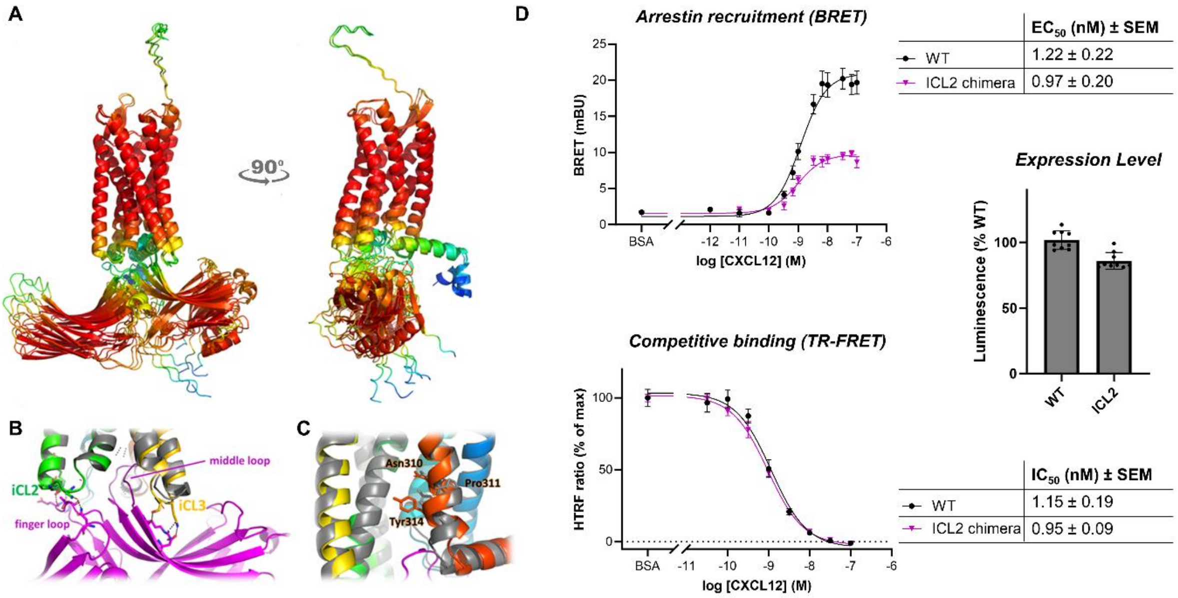


Figure S6: Principal component analysis (PCA) of the MD trajectories.

A) Grouped PCA of 4 MD trajectories (*apo* ACKR3 and bound with ligands), which allows direct comparison of the movements in the 4 systems. Principal component (PC) 1 corresponds to VUF15484-induced 7TM twist around the pocket and the receptor conformational changes on the intracellular side. PC1 corresponds to ECL2 movements in the systems, which may be artefacts due to the truncated N-terminus in the MD simulations. **B)** Grouped PCA of *apo* ACKR3 *versus* the CXCL12-bound form. Both PC1 and PC2 both exhibit CXCL2-induced 7TM twist and intracellular conformational changes, similar to the PC1 in (A). PCA was performed on the Cα atoms of the 7TM domain. The same trajectories were used to calculate the lnPF in Figs. 2 and 3.



	EC_{50} (nM) ± SEM
WT	1.22 ± 0.22
ICL2 chimera	0.97 ± 0.20

	IC_{50} (nM) ± SEM
WT	1.15 ± 0.19
ICL2 chimera	0.95 ± 0.09

Figure S7: Alphafold models of the ACKR3 – β-arrestin 1 complex and impact of ICL2 in β-arrestin 1 recruitment.

A) The top three AF models of ACKR3– β-arrestin 1 complex in two views rotated by 90°. Models were superimposed by alignment on the ACKR3 chain, and are shown in cartoon representation and coloured by AF predicted local distance difference test (pLDDT) score. **B)** Close up view of the interaction interface between ACKR3 and β-arrestin 1, highlighting the polar contacts formed between β-arrestin 1 and ACKR3 ICL2 and ICL3. The β-arrestin 1 is coloured in magenta while ACKR3 is coloured from blue (N-terminus) to red (C-terminus). The ACKR3 monomeric model is superimposed onto the model of the complex and coloured in grey. **C)** Superimposed ACKR3– β-arrestin 1 and ACKR3 models, highlighting the motion of the NPxxY motif. The colour code is the same as in B. **D)** Recruitment of β-arrestin 1-YFP to nLuc-tagged ACKR3 ICL2 chimera followed across a concentration range of agonist CXCL12 (top panel) and TR-FRET competitive binding assay followed between receptor and CXCL12-d2 in the presence of a concentration range of unlabelled CXCL12 (bottom panel). In this chimera, residues 144^{ICL2}-149^{ICL2} of ACKR3 were replaced with corresponding residues of CXCR4. Histogram shows nLuc luminescence for ACKR3 ICL2 chimera mutant normalised to WT levels. Data is a composite of three separate experiments measured in triplicate. Error bars represent mean ± SEM across the three experiments.

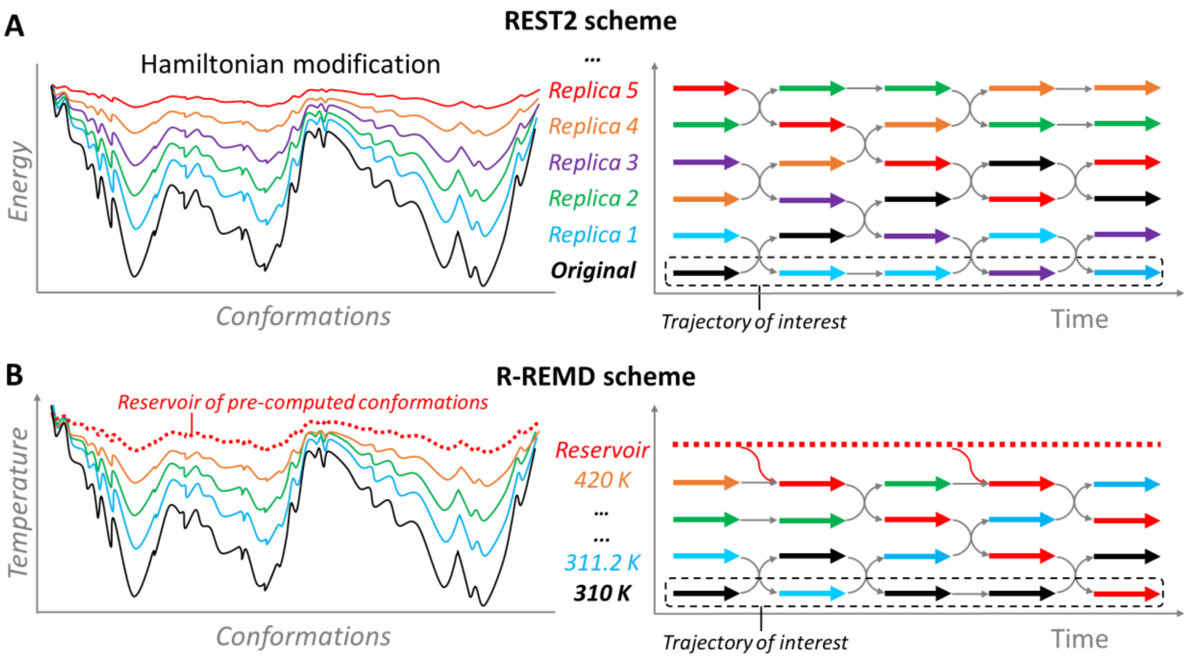


Figure S8: Scheme of REST2 and R-REMD simulations.

A) In REST2 MD, the Hamiltonians of the replicas are modified to facilitate conformational changes of the "solute" (protein and ligands), whereas the membrane and solvent are intact. All the replicas start from the same initial point and run at the same temperature (310 K). The conformational sampling rate is mostly limited by the slowest replicas in terms of conformational changes (the bottom ones). **B)** In R-REMD, a reservoir of pre-computed conformations of the system is coupled to the temperature REMD simulations to drive rapid conformational sampling. Thus, the low-temperature replicas impose fewer restrictions on the sampling rate.

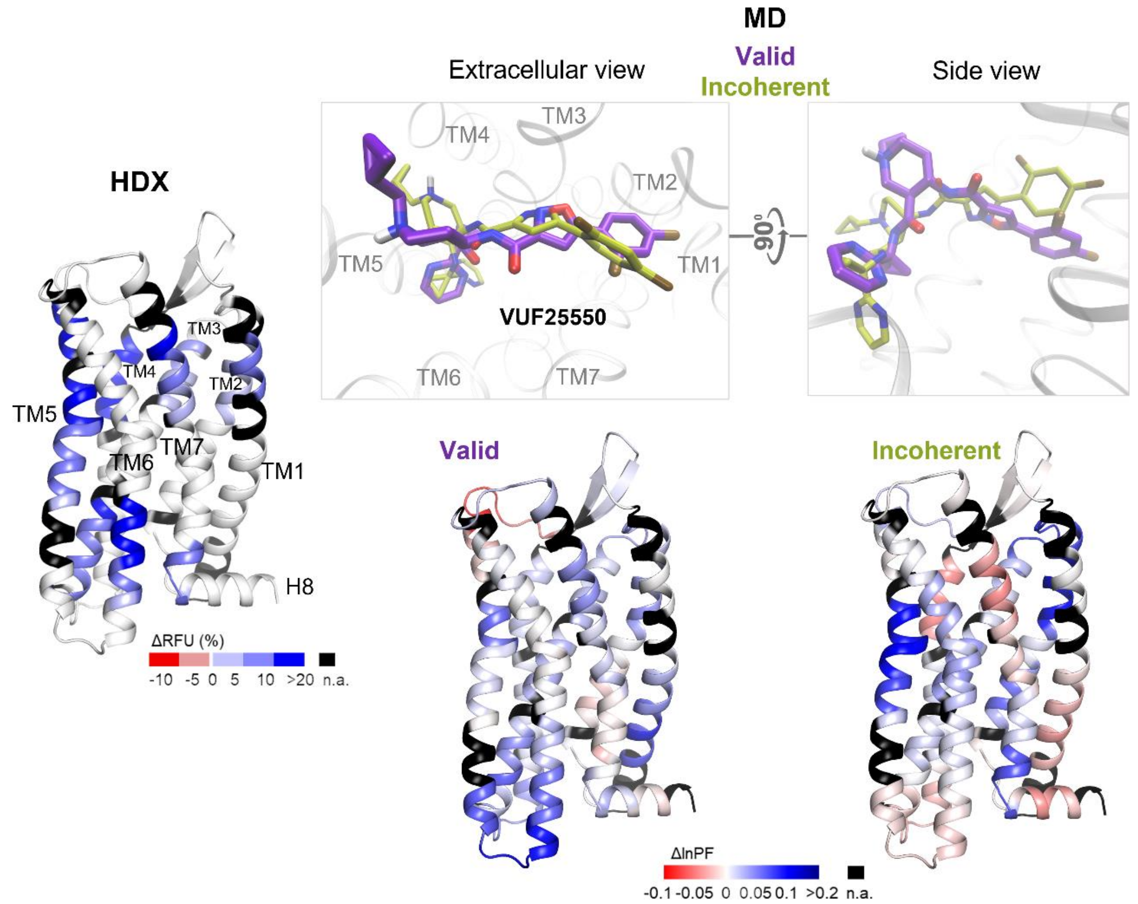


Figure S9: Evaluating REST2 MD prediction accuracy with HDX-MS data.

An example of REST2 MD predicted binding poses for VUF25550. The valid binding pose (violet) resulted in ΔlnPF values that were coherent with the ΔRFU from HDX-MS for this ligand (left panel), namely protections at the orthosteric pocket and on the intracellular side of TM3, TM5, TM6 and the TM7-H8 linker. A different binding pose (green) gave incoherent ΔlnPF values, namely the deprotections at the pocket (in TM4 and TM7) and on the intracellular side of TM1, TM3-TM6 and H8. Thus, this pose was discarded.

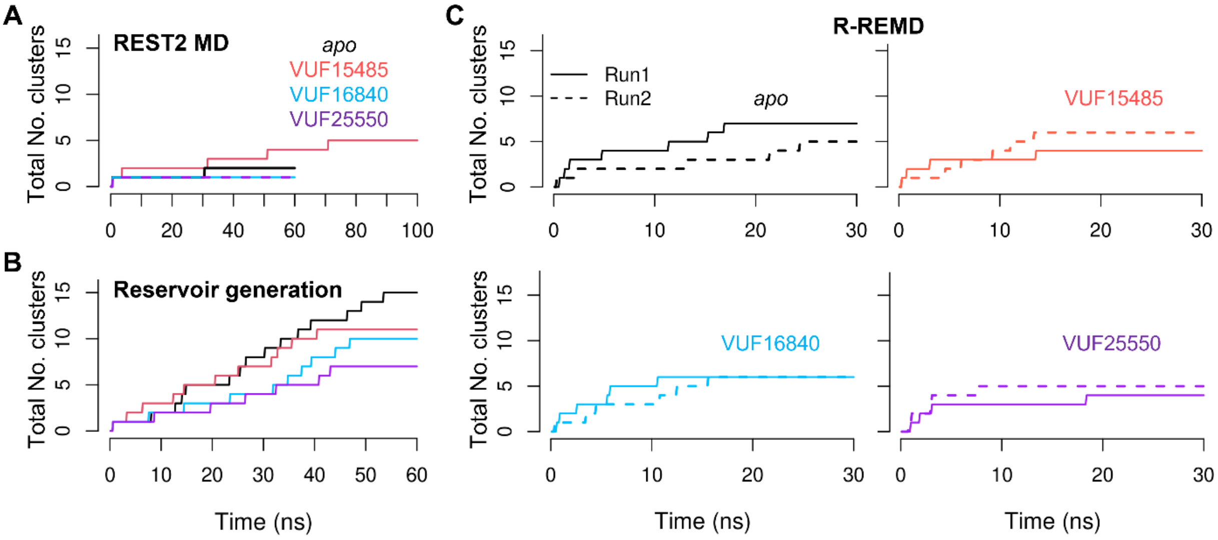


Figure S10: Convergence of MD simulations evaluated by a clustering analysis.

A) Convergence of REST2 MD simulations. The simulations were considered converged when the total number of clusters became constant (no more appearance of new clusters). **B)** Reservoir generation by MD simulations with modified Hamiltonians to facilitate conformational change. **C)** Lowest-temperature (310 K) replica of the R-REMD simulations.

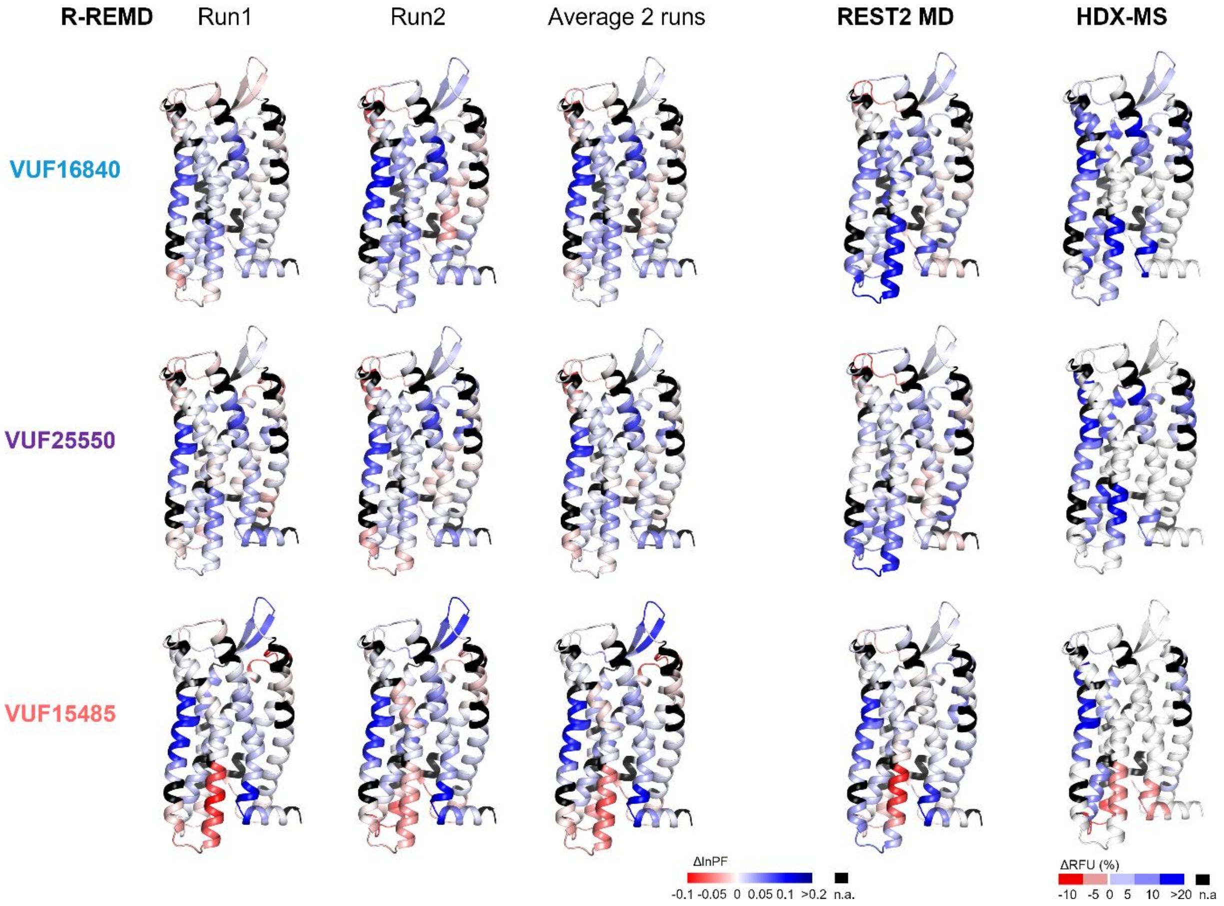


Figure S11: Comparison of R-REMD, REST2 MD predictions with HDX-MS data.

Table S1: Binding affinity of VUF15485 and VUF16840 for ACKR3 and selected binding site mutants receptors as measured by [^{3}H]VUF15485 binding. Binding affinities (K_i) were determined from pIC_{50} using the Cheng-Prusoff equation. All values depict the mean ± SEM of (N) experiments.

pK_i ± SEM (N)	**WT**	**D179N$^{4.60}$**	**E213Q$^{5.39}$**	**D275N$^{6.58}$**	**Q301E$^{7.39}$**	**Q301A$^{7.39}$**
VUF15485	7.9 ± 0.1 (18)	6.4 ± 0.2 (7)**	7.9 ± 0.1 (7)	8.1 ± 0.2 (4)	7.5 ± 0.2 (4)	8.0 ± 0.2 (8)
VUF16840	9.3 ± 0.0 (13)	7.5 ± 0.2 (7)**	8.3 ± 0.2 (7)**	9.2 ± 0.1 (4)	8.7 ± 0.3 (4)*	8.7 ± 0.1 (8)*

* $P < 0.05$ compared to WT
** $P < 0.001$ compared to WT

SI References:

1. A. E. Roitberg, A. Okur, C. Simmerling, Coupling of Replica Exchange Simulations to a Non-Boltzmann Structure Reservoir. *J Phys Chem B* **111**, 2415–2418 (2007).

2. L. Wang, R. A. Friesner, B. J. Berne, Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2). *J Phys Chem B* **115**, 9431–9438 (2011).

3. K. Kasavajhala, K. Lam, C. Simmerling, Exploring Protocols to Build Reservoirs to Accelerate Temperature Replica Exchange MD Simulations. *J Chem Theory Comput* **16**, 7776–7799 (2020).

4. A. Zarca, *et al.*, Pharmacological characterization and radiolabeling of VUF15485, a high-affinity small-molecule agonist for the atypical chemokine receptor ACKR3. [Preprint] (2023). Available at: https://www.biorxiv.org/content/10.1101/2023.07.12.548622v1 [Accessed 13 July 2023].