An Extended Metadynamics Protocol for Binding/Unbinding Free Energies of Peptide Ligands to Class A G-Protein Coupled Receptors

*Jacqueline C. Calderón,[a] Eva Plut,[b] Max Keller,[c] Chiara Cabrele,[b] Oliver Reiser,[b] Francesco L. Gervasio[d] and Timothy Clark[a]**

[a] Computer-Chemistry-Center, Department of Chemistry and Pharmacy, Friedrich-Alexander-University Erlangen-Nuernberg, Naegelsbachstr. 25, 91052 Erlangen, Germany.

[b] Institute of Organic Chemistry, Faculty of Chemistry and Pharmacy, University of Regensburg, 93040 Regensburg, Germany

[c] Institute of Pharmacy, Faculty of Chemistry and Pharmacy, University of Regensburg, D-93040 Regensburg, Germany.

[d] Pharmaceutical Sciences, University of Geneva, Geneva CH-1211, Switzerland.

Molecular Dynamics; Metadynamics; Free Energy of Binding; G-Protein Coupled Receptor; Peptide Hormones.

ABSTRACT

A metadynamics protocol is presented to characterize the binding and unbinding of peptide ligands to class A G-protein coupled receptors (GPCRs). The protocol expands on the one previously presented for binding and unbinding of small-molecule ligands to class A GPCRs and takes the more demanding nature of the peptide binding/unbinding process into account. It is applicable to almost all class A GPCRs. Exemplary simulations are described for subtypes Y_1R, Y_2R and Y_4R of the neuropeptide Y receptor family, vasopressin binding to the vasopressin V_2 receptor (V_2R), and oxytocin binding to the oxytocin receptor (OTR). Binding free energies and the positions of alternative binding sites are presented and, where possible, compared with experiment.

INTRODUCTION

Six years ago, we presented a standard metadynamics protocol for characterizing the binding/unbinding pathways of small-molecule ligands to class A G-protein coupled receptors[1] that has proved to be transferable to a wide variety of GPCRs and has been adopted by several groups.[2,3,4] However, our subsequent unpublished experience has shown that it can fail when simulating the binding/unbinding pathways of peptide ligands to the appropriate GPCRs. Our first simulations on the peptide hormone vasopressin[5] used metadynamics with two collective variables (CVs), which allowed adequate flexibility for the peptide ligand to sample all the relevant conformational space. This strategy was, however, replaced by our published single-CV protocol,[1] as using a single CV is computationally far more efficient than two-dimensional metadynamics simulations. Despite this simplification, the published protocol works well for small-molecule ligands and provides adequate sampling. In order to extend the utility of this "small-molecule" protocol to make it suitable for peptide ligands, we now report a modified protocol that can treat peptide ligands reliably while still being applicable to small-molecule ligands.

The main problems found using the "small-molecule" protocol for peptide ligands are that

- The original funnel restraint is too restrictive to allow adequate sampling of long, flexible peptide ligands

- The simulated pathway must allow the peptide to separate from the extracellular loops of the receptor, which are often long and can extend far into the extracellular medium

- Because of their potentially large number of interactions with the receptor, peptide ligands can exert larger pulling forces on the lower anchor point of the standard CV, $W^{6.48}$, than small molecules. These large forces can result in distortion of the receptor,

so that a mechanically more stable lower anchor point is needed for the CV for peptide binding.

The first point is particularly important. Ligand flexibility increases the complexity of the docking process immensely, so that, for instance, even though docking approaches that use conformational ensembles for ligands were developed 30 years ago,[6] peptide ligands remain challenging for docking techniques in general.[7] Although MW techniques are better able to sample ligand conformations than docking (at a corresponding cost in computer resources),[8] the flexibility of peptide ligands remains very challenging. Our experience suggests that the conformational problem is most serious in the extracellular medium, where sampling peptide conformations is a formidable task.[9,10] We emphasize here that the protocol described gives good results for free-energies of binding but may not reproduce the conformational equilibrium of the peptide ligand in solution. However, the peptide conformations given in the extracellular medium are apparently stable enough that they do not introduce significant errors in the free-energies of binding.

Natural peptide ligands for GPCRs include, among others, angiotensin,[11] neurotensin,[12] and vasopressin.[13] Synthetic peptide ligands may be derived from either combinatorial libraries or the natural ligands.[14,15,16] Peptide-like ligands that feature non-natural amino-acid residues[17,18,19] or peptidomimetics[20,21] often offer advantages over their natural analogs, including higher affinity,[22,23] enhanced proteolytic stability[24] or photo-switching.[25] There is therefore an increasing need for a generally applicable, reliable simulation protocol to simulate binding and unbinding of peptide ligands to GPCRs, especially to identify and characterize potential alternative binding sites to the orthosteric one. Such a protocol would not only allow standard and reproducible results to be obtained by different researchers worldwide but also would provide accurate predictions of free energies of binding in addition to structural details of often multiple binding sites within the receptor. An operational advantage in the case of

peptide ligands is that the same force field can be used for both receptor and ligand, removing possible inconsistencies between protein and ligand Hamiltonians.

COMPUTATIONAL PROTOCOL

As in the small-molecule protocol,[1] we use well-tempered metadynamics simulations[26] with a funnel restraint[27] to improve efficiency by limiting translation of the free ligand in the extracellular medium. Multiple walkers[28] (MW) are used to improve conformational sampling and to increase computational efficiency on highly parallel hardware. All simulations were performed on the CPU and GPU clusters of NHR@FAU.[29]

Collective variable:

Although metadynamics does not apply a pulling potential directly to the ligand in a binding/unbinding simulation, the addition of Gaussian hills to the potential hypersurface results indirectly in forces on the anchor point of the CV within the receptor and on the ligand. Because typical peptide ligands are larger than small-molecule ligands and offer many more potentially strong interactions with the receptor, these intrinsic forces can be large enough to distort the lower anchor point of the CV (i.e. induce strain in the receptor rather than binding the ligand). The choice of the lower anchor point within the receptor and of the upper one within the ligand therefore becomes critical when binding peptides. The former can be stabilized relative to that used in the small-molecule protocol (C_α of the highly conserved $W^{6.48}$) by including a residue from the neighboring helix in the definition of the anchor point. We thus use the midpoint of the vector between the α-carbons of the $W^{6.48}$ and $V^{3.36}$ residues (Figure 1A). This point is mechanically stable because moving it requires moving both helices 3 and 6. Residue 3.36 is, in contrast to $W^{6.48}$, not conserved (it is $V^{3.36}$ in Y_4R and Y_2R, $I^{3.36}$ in Y_1R, $M^{3.36}$ in V_2R, and OTR) but its position relative to $W^{6.48}$ is fairly constant, so that it can be used generally for class A GPCRs.

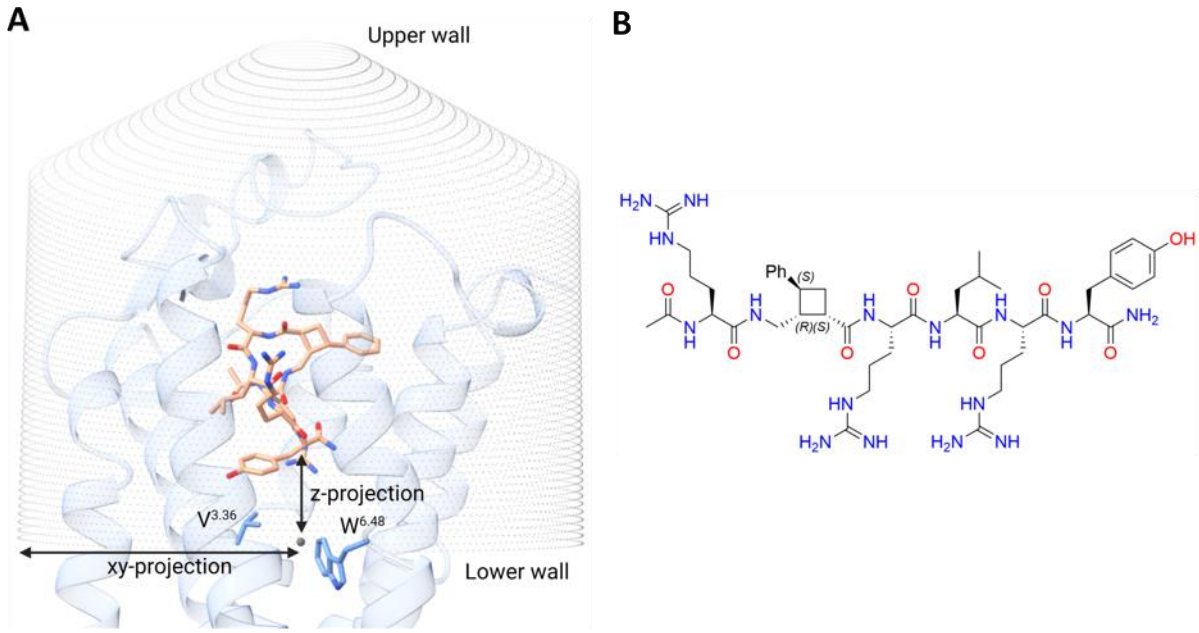

Figure 1. (A) Schematic representation of the funnel-shaped restraints applied to the metadynamics binding/unbinding simulations of peptide ligands. (B) Chemical structure of the α,γ-hexapeptide with (1*S*,2*R*,3*S*)-γ-cyclobutane amino acid (SRS)[30] used in the optimization of the protocol.

It is quite a common feature of peptide-binding paths at GPCRs that an initial salt bridge is formed early in the path and that this constrains one end of the peptide, allowing it to sample conformational space within the receptor with the other end more effectively.[5] The anchor point of the CV in the peptide ligand should then be at the other terminus of the peptide to that involved in a potential salt-bridge (often arginine). In the case of the synthetic peptide ligand for the neuropeptide Y receptor Y_4 (Y_4R) shown in Figure 1B, arginine is the N-terminal residue, so that we therefore used C_α of the C-terminal residue of the peptide. Tikhonova et al.[31] have provided guidance as to how peptide ligands bind to GPCRs. Peptide ligands most frequently bind with the C-terminus deep in the orthosteric binding site, except for the opioid and chemokine receptors. In these cases, the N-terminus binds in the orthosteric site. In both cases. The other peptide terminus interacts with extracellular loops and the receptor N-terminus. In cases where the correct choice of peptide terminus is not clear, both should be investigated. The appropriate terminus is that which binds deepest in the receptor binding site.

Funnel restraint:

Figure 2A shows the sampling obtained for a peptide-like ligand at Y_4R using the published[1] "small molecule" protocol with an increased upper wall value (7.0 nm, compared to 5.0 nm for the "small molecule" protocol).

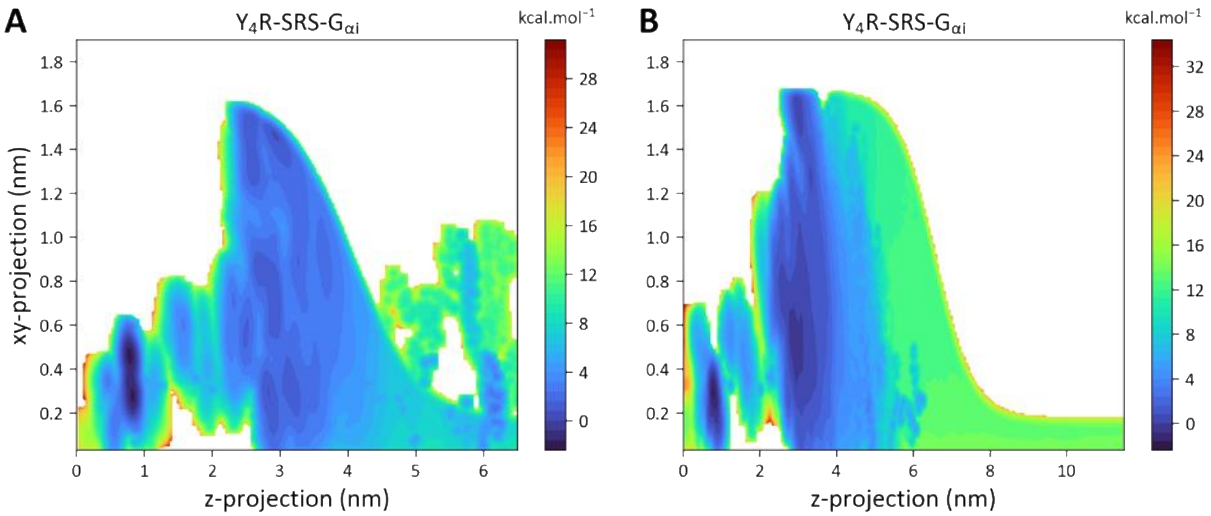

Figure 2. 2D free energy surfaces as a function of the xy-projection and z-projection obtained with: (A) funnel dimensions according to the binding/unbinding protocol of small molecules[1], (B) optimized dimensions to simulate binding/unbinding process of peptides.

The plot shows clearly that the funnel is preventing adequate sampling on the extracellular side of the receptor. The extracellular constriction of the funnel must therefore be moved further outwards for peptide-like ligands. Figure 2B shows the change made to the funnel in the new protocol. All parameters from the original protocol were retained except that the inflexion point was moved 2.5 nm towards the extracellular medium and the upper wall was set at 11.5 nm. Detailed parameters for the "small molecule" and "peptide" funnels are given in Figure S1 of the Supporting Information.

Path extent:

The three plots in Figure 3A-C show the time-evolution of the free-energy profiles for the above example using upper wall potentials set at 7.0, 9.5 and 11.5 nm from the lower anchor point of the CV, respectively. They demonstrate that the free energy of the ligand in the extracellular medium does not become constant until a CV value of approximately 8 nm. A horizontal free-energy profile in this region is necessary because it defines the reference (zero relative free energy) level for determining binding energies. Figure 3D shows a representative

structure at a CV value of 5.9 nm, illustrating that some interactions between the peptide and the receptor remain operative at this distance. We therefore adopt a standard upper wall distance of 11.5 nm for the "peptide" binding/unbinding protocol, although this value must be checked by visual inspection of the "dissociated" structures and from the convergence of the free energy to a constant value in the extracellular medium.

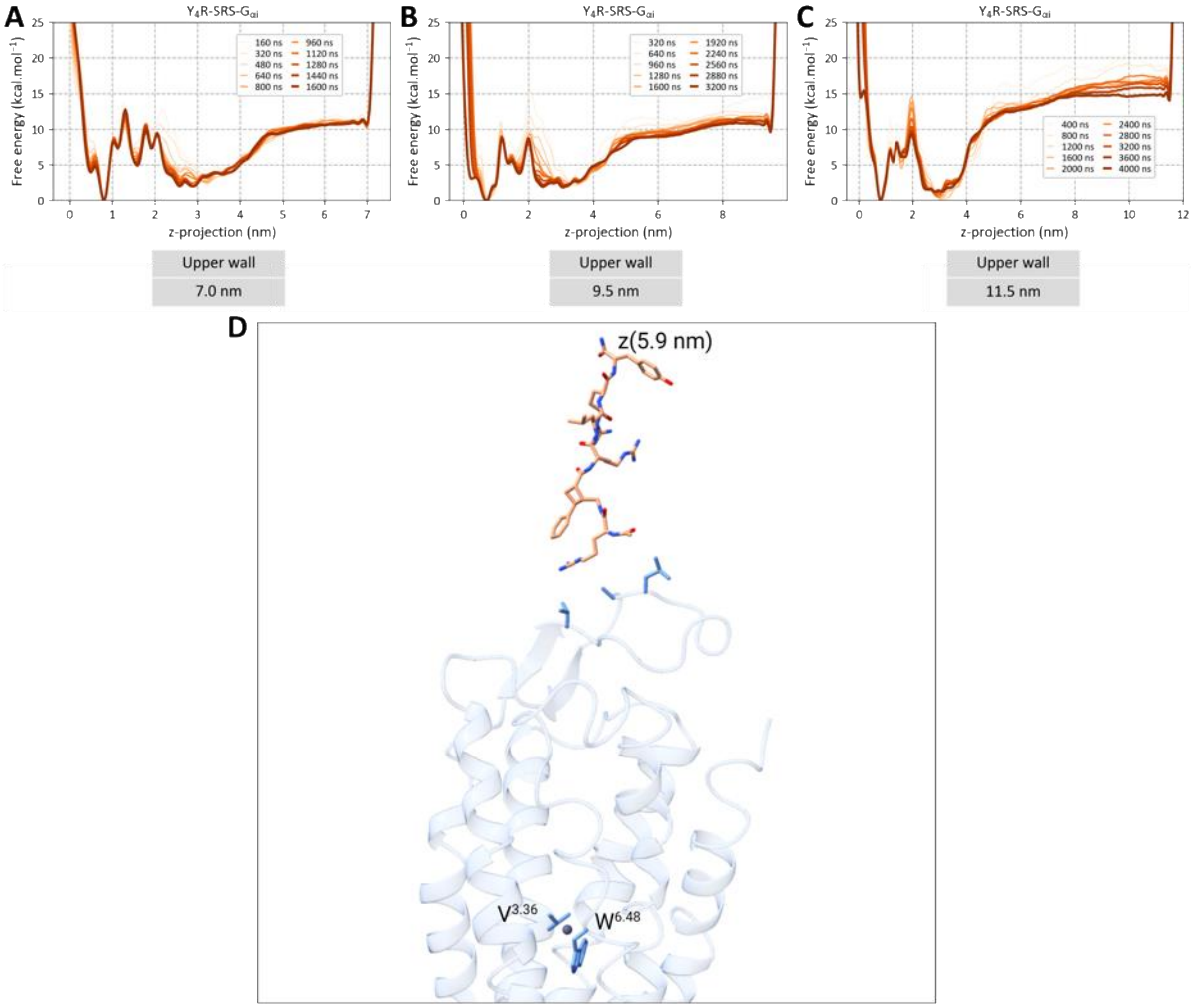

Figure 3. (A-C) Change in the reconstructed free-energy profiles as a function of the sampling time for different upper wall potentials: (A) 7.0, (B) 9.5 and (C) 11.5 nm. (D) Representative structure of the α,γ-hexapeptide SRS[30] at a CV value of 5.9 nm.

RESULTS AND DISCUSSION

Table 1 shows the calculated binding free energies (BFEs) for seven peptide-GPCR complexes consisting of five receptors and six peptide ligands ranging from 6 to 36 amino acids in length (Figures 1B and S2-6), and Figure 4 shows the correlation between the calculated and experimental BFEs.

Table 1. Complexes used for the MW-WT-metadynamics simulations and their experimental and calculated binding free energies (in kcal mol^{-1})[a]

Ligand	Receptor	Effect	System	K_i (nM)	ΔG_{exp}	ΔG_{calc}
• pancreatic polypeptide (hPP)	Y_4R	agonist	ternary-G_i	0.69[30]	−13.0	−12.7 ± 0.5
• α,γ-hexapeptide with (1*S*,2*R*,3*S*)-γ-CBAA (SRS)	Y_4R	partial agonist	ternary-G_i	12[30]	−11.2	−11.1 ± 0.5
• α,γ-hexapeptide with (1*R*,2*S*,3*R*)-γ-CBAA (RSR)	Y_4R	partial agonist	ternary-G_i	0.66[30]	−13.0	−12.9 ± 0.5
• neuropeptide Y (pNPY)	Y_1R	agonist	ternary-G_i	0.39[23]	−13.3	−13.4 ± 0.7
• neuropeptide Y (pNPY)	Y_2R	agonist	ternary-G_i	0.50[23]	−13.2	−12.2 ± 0.6
• arginine vasopressin (AVP)	V_2R	agonist	ternary-G_s	1.39[32]	−12.6	−12.8 ± 0.5
• oxytocin (OT)	OTR	agonist	ternary-G_q	0.79[33]	−12.7	−13.4 ± 0.5

[a] Experimental free energies were obtained from the relation $\Delta G = -RT\ln(K_i)$ at T = 310 K and the computed ones as described in Methods. All calculated ΔG-values have been corrected for the standard volume and funnel potential used, as described in references 34 and 35.

The mean unsigned error (MUE) between calculated and experimental binding energies is 0.39 kcal mol^{-1} and the root mean square deviation (RMSD) 0.49 kcal mol^{-1}. The slope of the least-squares fitted line is 0.86 and R^2 0.59. We consider these values acceptable because the range of binding energies is less than 3 kcal mol^{-1}, as is likely to be the case for peptide ligands in general. The uncertainties (± one standard deviation) in the calculated BFEs are shown as vertical bars, but Figure 4 does not show estimated experimental uncertainties. The calculated BFEs for five of the seven ligand/receptor combinations lie within 0.5 kcal mol^{-1} of the experimental values (shown by the dashed red lines); the exceptions being pNPY/Y_2R and OT/OTR. We conclude that the simulations reproduce experimental BFEs closely in an absolute sense, but we cannot judge whether trends are reproduced well because the range of BFEs is so small, and because the point for SRS/Y_4R dominates the regression. The trends for Y_4R are reproduced well, but simulations suggest a preference of pNPY for Y_1R over Y_2R that is not found experimentally.

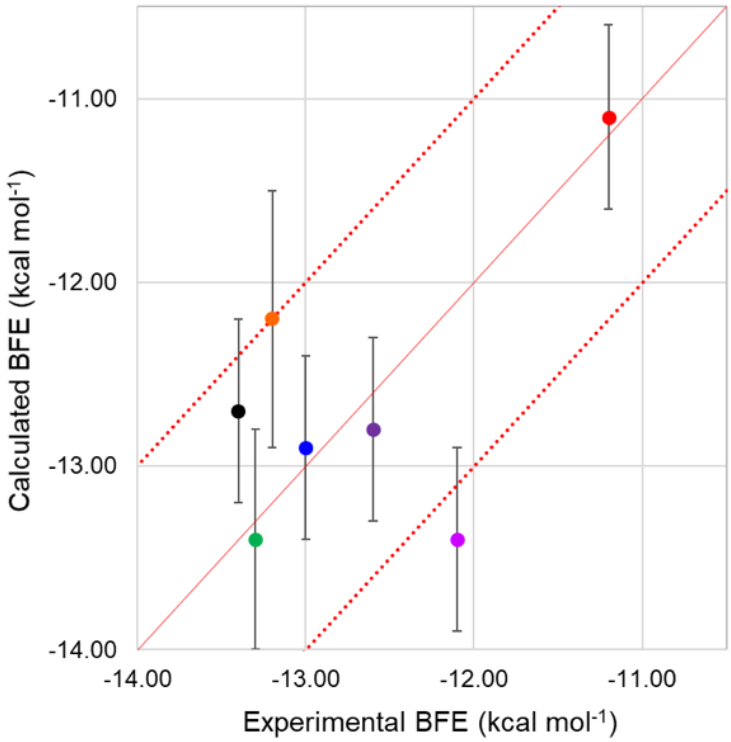

Figure 4. Correlation between the experimental and calculated binding free energies for the data shown in Table 1. The error bars represent one standard deviation. The solid red line shows perfect agreement, and the dashed red lines indicate an error of ± 0.5 kcal mol^{-1}. The dots are color coded as in Table 1.

Because the systems studied here entail a ligand-binding mechanism that requires more complex conformational changes of both the receptor and the ligand than the binding of small-molecule ligands, the following additional convergence criteria were adopted to ensure a balance of computational cost and accuracy:

- at least one re-crossing (rebinding) event must be observed, and
- the calculated error in the reconstructed free energy profiles must reach a plateau.

For this purpose, two reference distances were defined to track re-crossing events along the CV pathway. The first distance represents the lower limit of the unbound states ($z_{unbound}$). We have established that no interaction between the peptide ligand and the receptor is operative at CV distances greater than 6.5 nm ($z_{unbound} \geq 6.5$ nm). The second distance represents the lower limit at which the ligand exits the binding pocket (z_{BP}). For CV distances lower than z_{BP} the ligand is in the receptor binding pocket ($z_{BP} = 4.7$ nm for Y_4R, Y_2R and Y_1R receptors, and $z_{BP} = 3.3$ for V_2R and OTR).

Figure 5 shows the plots used to determine the convergence of the MW-WT-metadynamics simulations for the Y_4R-SRS-$G_{\alpha i}$ complex. Re-crossing events during the metadynamics run were tracked by inspecting the CV fluctuations of the individual trajectories for each walker (Figure 5B). It is not surprising that only rarely rebinding events are observed, whereby the ligand returns to the initial deep binding pose. This phenomenon is commonly observed[36] and is due to the fact that a simple geometry-based CV has difficulties in guiding the peptide all the way back to its crystallographic pose. Still, as it can be seen in Fig. 5B, a number of walkers remain in or close to the bound pose, while many others explore unbound states before rebinding. The fact that the range explored by different walkers fully overlaps and a number of walkers go from bound to unbound and back, leads to a good convergence of the reconstructed FES (in a way reminiscent of Umbrella Sampling). The average error of the MW-WT-metadynamics simulation was obtained through block analysis (Figure 5A). As expected, the error increases with the block size until it reaches a plateau with a value of ≈ 0.5 kcal mol^{-1}. The convergence behavior of the other systems is similar, as can be seen from Figures S7-12. Altogether, these results show that the free-energy profiles for the binding/unbinding of large and highly flexible peptide ligands converged, according to the chosen convergence criteria, between 3,200 and 4,800 ns of cumulative simulation time.

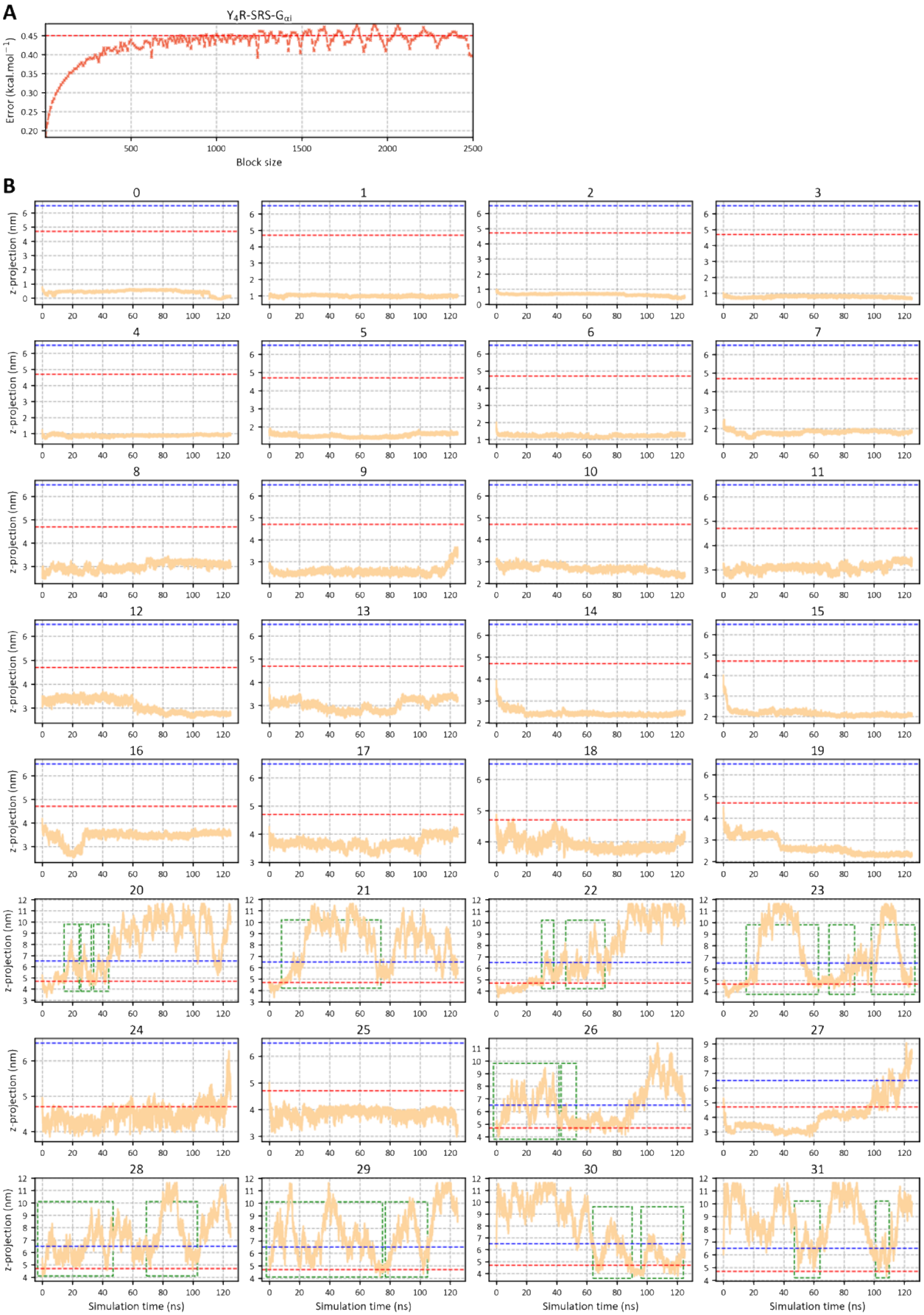

Figure 5. Results of binding/unbinding MW-WT-metadynamics simulation for Y_4R-SRS-$G_{\alpha i}$ (A) Block analysis of the metadynamics simulation. (B) Binding CV as a function of the sampling time for 32 walkers. The red (z_{BP}) and blue ($z_{unbound}$) dashed lines represent the reference distances used to track re-crossing events. For CV distances $\geq z_{unbound}$ and CV distances $\leq z_{BP}$ the ligand is in the extracellular medium and in the binding pocket, respectively. Re-crossing is highlighted by green dashed boxes.

BINDING SITES

Figure 6 shows the reconstructed binding free-energy profiles obtained through the reweighting procedure described in the Computational Methods section, with the estimated errors shown as shaded bands. An overlay of the representative minima structures obtained from the MW-WT-metadynamics simulations for the seven ligand/receptor combinations studied here is also shown. Representative structures for each system were obtained through a clustering analysis based on their global and local minima.

Comparison of the binding pose for global (orthosteric) and secondary (vestibule) minima indicates a significant role of the extracellular vestibule in driving the binding process. According to Figures 6B and 6C, orthosteric and vestibule binding poses found for peptide SRS in Y_4R are comparable. This is particularly noticeable for residues R31 and (1*S*,2*R*,3*S*)-γ32 which accommodate in the binding pocket in a similar manner to that of the global minimum. In addition, residues R33 and R35 occupy the extracellular vestibule maintaining their relative position in the receptor. The di-arginine motif R33-X-R35 is projected to the extracellular vestibule pointing toward transmembrane helix 6 (TM6), while Y36 faces ECL2. For details of the truncations used for the receptors in the simulations, please see the Computational Methods section. In the secondary minimum, RSR undergoes a more significant conformational rearrangement, rotating its core away from TM5 and TM6 and projecting its residues Y36 and R31 towards TM2, while R35 points in direction of the extracellular medium. In addition, similar to peptide RSR, significant conformational rearrangements were also observed for the C-terminus of the hPP and NPY in Y_4R and Y_1R respectively. For Y_1R, Y_2R, and Y_4R subtypes, the conformation found in the secondary minimum shows that neuropeptide ligands are closer to the ECL2 in comparison to the global minimum. The ECL2, in particular, is often the first point of contact between the ligand and the receptor.[37,38] Unlike receptors with

long ECL2, such as the neuropeptide Y receptors, V_2R-AVP-$G_{\alpha s}$ and OTR-OT-$G_{\alpha q}$ complexes show binding modes of the secondary minimum for AVP and OT in close proximity with the receptor N-terminus and TM1 (Figures 6Q,R and 6T,U respectively).

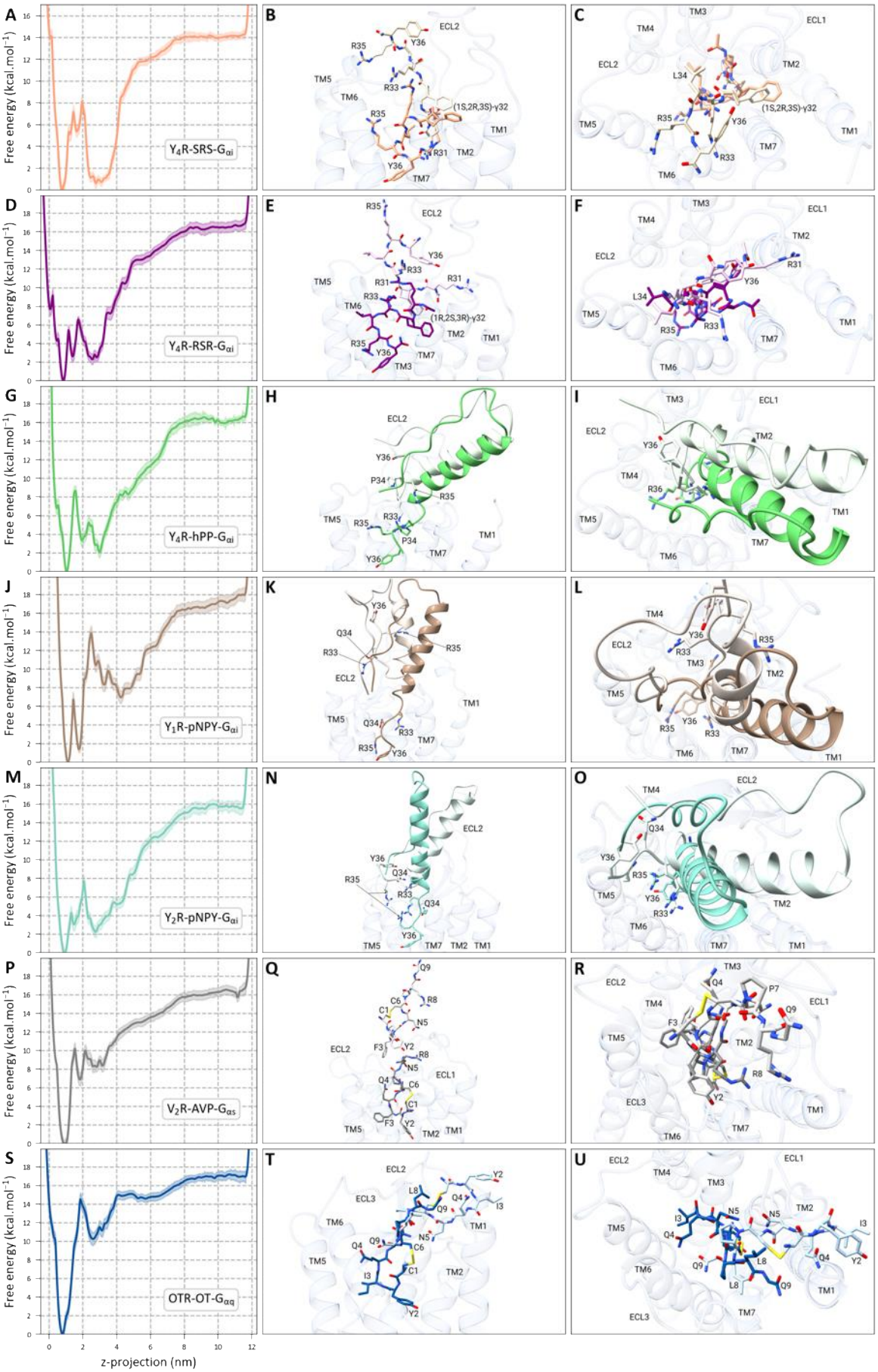

Figure 6. Binding free energy profiles and overlay of the representative minima structures obtained from the MW-WT-metadynamics simulations for the seven ligand/receptor combinations: (A-C) Y_4R-SRS-$G_{\alpha i}$ (D-F) Y_4R-RSR-$G_{\alpha i}$ (G-I) Y_4R-hPP-$G_{\alpha i}$ (J-L) Y_1R-pNPY-$G_{\alpha i}$ (M-O) Y_2R-pNPY-$G_{\alpha i}$ (P-R) V_2R-AVP-$G_{\alpha s}$ (S-U) OTR-OT-$G_{\alpha q}$. Receptor starting geometries were used as receptor reference structures. The second and third column represent side and top view, respectively. Light and dark colored ligand structures represent global and secondary minimum respectively.

A detailed description of the interactions established by peptide ligands with their receptors, in both global and local minima, is shown in Figure 7. For Y_4R and Y_1R subtypes, a hydrogen-bond interaction between Y36 and $Q^{5.46}$ anchors the C-terminus of the peptide ligands into the binding pocket (Figures 7A, C, E, and G). Except for the small α,γ-hexapeptides, common interactions among Y_nR subtypes include hydrogen-bond interactions between the C-terminus and $Q^{3.32}$, the salt bridges formed between residues R35 and $D^{6.59}$, and the H-bond between R33 and the residue at position 6.55 (asparagine in Y_4R and Y_1R, glutamine in Y_2R). Although interactions exhibited by the small α,γ-hexapeptides do not match those of the native ligand hPP, their residue R33 is involved in salt-bridge interactions with residues $D^{6.59}$ and $E203^{ECL3}$ in SRS (Figure 7A) and RSR (Figure 7C) respectively, mimicking the interactions of the conserved di-arginine motif R33-X-R35 of the NPY family. Note that these interactions do not agree with the docking poses that we reported previously.[30] This is not entirely surprising, as flexible peptide ligands present a formidable challenge for docking, especially when several alternative polar interactions are possible. We note that docking was also not able to reproduce the binding pose of vasopressin, but that the 2D-metadynamics simulations reproduced the experimentally observed contacts.[5] In this respect, the "anchoring" of one terminus of the peptide ligand by a salt bridge early in the binding path likely provides considerable sampling advantages for steered MD-methods such as metadynamics. We therefore consider the metadynamics docking/undocking protocol to be more reliable than the docking protocol used previously.[30]

As illustrated in Figures 7K and 7M, both AVP and OT exhibited comparable binding poses characterized by a deep location of the C-terminus oriented to TM7 and TM2, in which the hydroxyl group of Y2 forms an H-bond with the backbone carbonyl of $L^{7.40}$. While OT forms polar interactions with $Q^{2.61}$, $D^{2.65}$, $Q^{4.60}$, $Q^{6.55}$, $K^{7.30}$, $A^{7.32}$, and $S^{7.43}$, a more extended H-bond network is observed for arginine vasopressin. C1 of AVP forms a stabilizing H-bond network with residues $E^{1.35}$, $Q^{2.61}$ and $M^{7.39}$ of V_2R. F3 main chain CO group forms H-bonds with $Q^{4.60}$

and $K^{3.29}$. Q4 is H-bonded to $A194^{ECL2}$ and N5 makes polar interactions with $K^{3.29}$ and $C192^{ECL2}$. R8 forms salt bridges with $D^{1.28}$ and $E303^{ECL3}$. In addition, Q9 makes contacts with $W193^{ECL2}$ while forming H-bonds with the receptor N-terminus ($R32^{N\text{-}term}$). Overall, the binding pose found in the global minimum for the seven ligand/receptor combinations described here is similar to that of the binding modes previously reported in Cryo-EM structures and docking studies.[30,39,40,41]

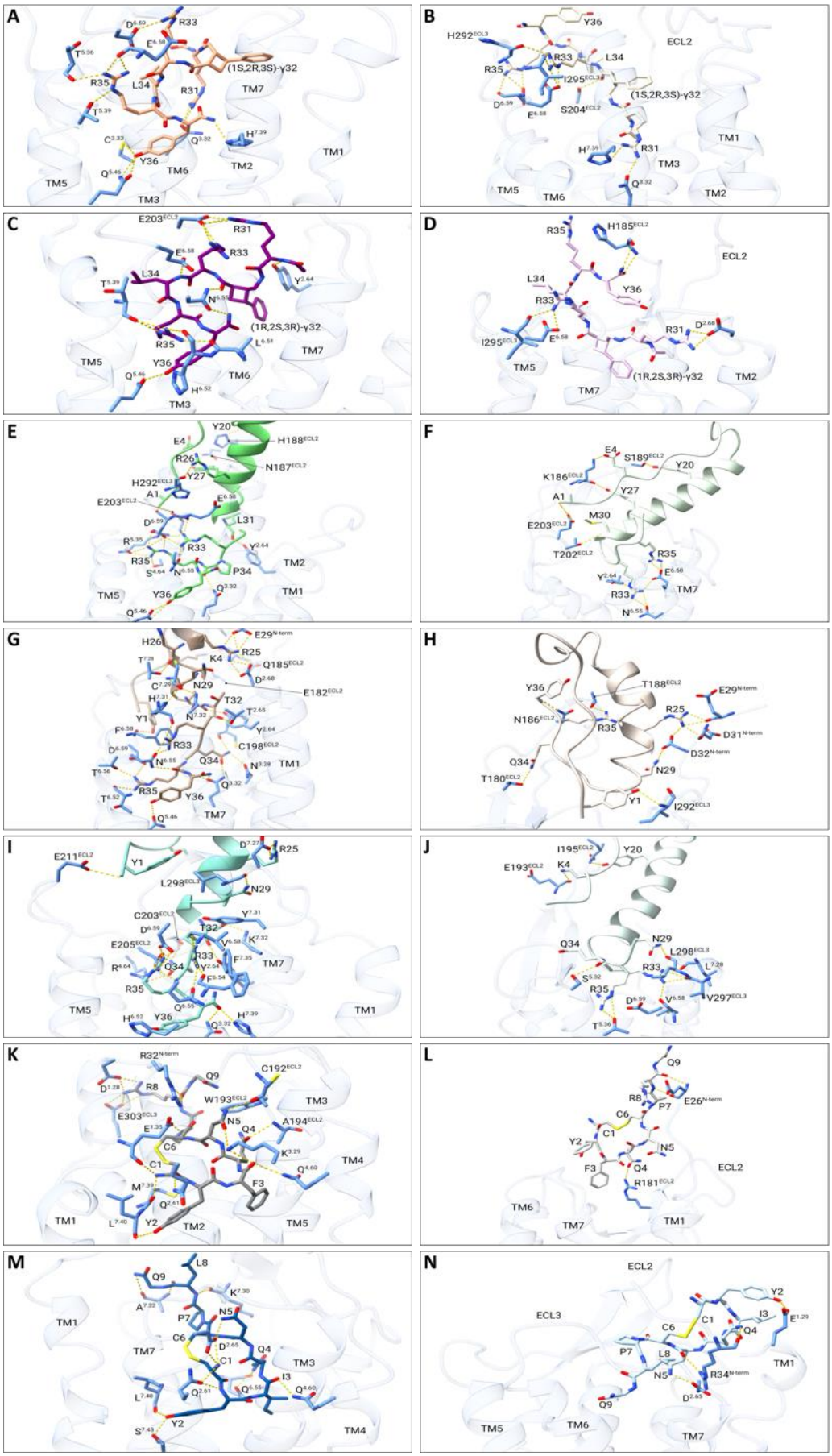

Figure 7. Interactions of the representative minima structures obtained from the MW-WT-metadynamics simulations for the seven ligand/receptor combinations: (A,B) Y_4R-SRS-$G_{\alpha i}$ (C,D) Y_4R-RSR-$G_{\alpha i}$ (E,F) Y_4R-hPP-$G_{\alpha i}$ (G,H) Y_1R-pNPY-$G_{\alpha i}$ (I,J) Y_2R-pNPY-$G_{\alpha i}$ (K,L) V_2R-AVP-$G_{\alpha s}$ (M,N) OTR-OT-$G_{\alpha q}$. Nitrogen, oxygen, and sulfur atoms are shown in blue, red and yellow, respectively. H-bonds are represented by yellow dashed lines. The left and right column represent global (dark colored) and secondary (light colored) minimum, respectively.

In addition to the global minima, Y_1R, Y_2R, and Y_4R exhibit comparably stable secondary minima closer to the extracellular medium than the global (orthosteric) binding minimum and often thermodynamically competitive at room temperature. We originally thought that these vestibule-like binding sites would be potentially interesting for designing peptide antagonists because a ligand bound in these sites blocks the orthosteric one. However, the A^{100} activation index[42] indicates that both minima correspond to active conformations of the receptor (see Table 3 below). We are not aware of any precedent in which a single ligand can take up two different binding modes, both of which lead to receptor activation, although we must point out that it is possible that the receptors have simply retained their starting activation state in the simulations. However, recently reported biphasic association and dissociation curves obtained from binding studies with labeled peptidic agonists at neuropeptide receptors could indeed be explained by the coexistence of two binding modes. The two minima are most competitive for the peptide SRS bound to Y_4R. The secondary binding site is calculated to be only 0.8 kcal mol^{-1} less stable than the global minimum, indicating that this site would be approximately 26% populated at equilibrium (compared to the orthosteric binding site). The secondary binding site for the hPP and RSR peptides are relatively less stable (2.0 kcal mol^{-1}, 3.5%) but nonetheless competitive.

Table 2. Complexes used for the MW-WT-metadynamics simulations and their experimental and calculated binding free energies (in kcal mol^{-1})[a]

Ligand	Receptor	Effect	$\Delta G_{calc(ortho.)}$	$A^{100}_{(ortho.)}$	$\Delta G_{calc(sec.)}$	$A^{100}_{(sec.)}$	$\Delta\Delta G$
pancreatic polypeptide	Y_4R	agonist	−12.7 ± 0.5	80.2	−10.7 ± 0.6	78.3	2.0
α,γ-hexapeptide SRS	Y_4R	partial agonist	−11.1 ± 0.5	114.0	−10.3 ± 0.3	91.3	0.8
α,γ-hexapeptide RSR	Y_4R	partial agonist	−12.9 ± 0.5	28.2	−10.9 ± 0.5	139.7	2.0
neuropeptide Y	Y_1R	agonist	−13.4 ± 0.7	68.2	−6.4 ± 0.6	66.0	7.0
neuropeptide Y	Y_2R	agonist	−12.2 ± 0.6	94.6	−10.1 ± 0.6	65.3	2.1
arginine vasopressin	V_2R	agonist	−12.8 ± 0.5	43.1	−5.8 ± 0.6	53.0	7.0
oxytocin	OTR	agonist	−13.4 ± 0.5	35.5	−3.2 ± 0.7	37.9	10.2

[a] $\Delta G_{calc(ortho.)}$ indicates the binding energy at the orthosteric site and $\Delta G_{calc(sec.)}$ that at the secondary binding site.

A detailed description of the interactions found in the secondary minima for the α,γ-hexapeptides SRS and RSR are shown in Figures 7B and 7D, respectively. Given the

stereochemistry of the constrained γ-residue of the α,γ-hexapeptides, residues R31 and γ32 are oppositely oriented in the Y_4R binding cavity. Residue γ32 in peptide SRS points toward TM2, while it is positioned between TM3 and TM7 in peptide RSR. Different interactions are observed also for arginine residues. Residue R31 in RSR is salt-bridged to $D^{2.68}$, whereas R31 in SRS buries in the receptor cavity making favorable interactions with $H^{7.39}$ and $Q^{3.32}$. These two residues constitute the C-terminal anchor site of both hPP in Y_4R and pNPY in the Y_1R and Y_2R subtypes.[39] Moreover, while R33 in RSR is H-bonded to $I295^{ECL3}$ and $E^{6.58}$, a more extensive polar interaction network is formed among extracellular residues ($E^{6.58}$, $D^{6.59}$, $H292^{ECL3}$ and $I295^{ECL3}$) of Y_4R and arginine residues R35 and R33 in SRS.

It was previously shown that partial agonist ligands can occupy minima energetically close to the global minimum.[5] Despite both α,γ-peptides acting as partial agonists, the peptide SRS was reported to show lower intrinsic activity (α) in comparison to its enantiomer RSR (α of RSR/SRS: 50%/39%).[30] Therefore, the antagonistic-like effect of SRS is likely explained by the favorable secondary minimum found in its binding free energy profile, which corresponds to a binding mode in which the C-terminal anchor site for neuropeptide ligands is blocked by residue R31 in SRS. Thus, the peptide SRS adopts a secondary binding pose that is incompatible with the ligand-receptor interactions reported for hPP in Y_4R and suggested to be important for agonistic activity.[39] Note that (see below) the A^{100} activation index[42] indicates that the receptor with the peptide bound in the secondary site is activated. This does not necessarily contradict the above rationalization of the partial agonism of the ligands, but it does require that the receptor with the ligand bound in the secondary minimum cannot recruit the G-protein effectively.

The binding modes corresponding to the secondary minima found for the neuropeptide ligands hPP and pNPY are shown in Figures 7F, 7H, and 7J. Different interactions can be observed between Y_nR subtypes and their native neuropeptide ligands. For hPP, main interactions with

Y_4R involve residues in TM6 and ECL2 (Figure 7F). Residue R33 of hPP is involved in H-bond interactions with $Y^{2.64}$, $N^{6.55}$ and $E^{6.58}$ while R35 forms a salt bridge with $E^{6.58}$. Additionally, Y20 forms an H-bond with $S189^{ECL2}$, E4 and Y27 are H-bonded to $K186^{ECL2}$, while A1 and M30 establish H-bond interactions with residues $E203^{ECL2}$ and $T202^{ECL2}$, respectively. Figure 7H shows that interactions between pNPY and the Y_1R subtype comprise H-bond interactions with extracellular residues of the receptor in ECL2 (residues Q34, R35 and Y36 of pNPY with $T180^{ECL2}$, $T188^{ECL2}$ and $N186^{ECL2}$ of Y_1R respectively), ECL3 (Y1 with $I292^{ECL3}$) and N-terminus (N29 with $D32^{N\text{-}term}$, R25 with $E29^{N\text{-}term}$, $D31^{N\text{-}term}$, and $D32^{N\text{-}term}$). In the other hand, the secondary minimum for pNPY in the Y_2R comprises a binding mode in which polar interactions are formed between pNPY and extracellular residues in TM5, TM6, TM7, ECL2, and ECL3. These interactions include H-bonds formed by K4 and Y20 with residues in ECL2 ($E193^{ECL2}$ and $I195^{ECL2}$ respectively). In addition, R35 in pNPY is involved in polar contacts and H-bond interactions with residues $D^{6.59}$ and $T^{5.36}$ respectively. While an H-bond network is formed by R33 with residues $V^{6.58}$, $V297^{ECL3}$, and $L^{7.28}$. Moreover, Q34 and N29 form H-bond interactions with residues $S^{5.32}$ and $L298^{ECL3}$ respectively.

Importantly, despite the different interaction patterns shown by the systems of the neuropeptide Y family studied here (Figure 7), we can identify common features driving the binding process. For the peptide ligands, important interactions involved residue R35, while interactions among Y_nR subtypes were mostly established with residues in ECL2.

As shown in Figures 7L and 7N, the binding poses of AVP and OT that represent secondary minima are comparable. Both AVP and OT bind their receptor approaching TM1 and establishing interactions with polar residues in the receptor N-terminus. While residues Q4 and Q9 in AVP are H-bonded to $R181^{ECL2}$ and $E26^{N\text{-}term}$ respectively (Figure 7L), N5 in OT is forming H-bond interactions with residues $D^{2.65}$ and $R34^{N\text{-}term}$, Q4 and Y2 are H-bonded to

R34$^{N\text{-term}}$ and Y$^{1.29}$ respectively (Figure 7N). All these interactions highlight the important role played by the N-terminus for agonist binding to V_2R and OTR.[43]

We also estimate the GPCR A^{100} activation index[42] to investigate the activation state of the peptide-GPCR complexes studied here. For all the systems considered, we obtained active A^{100}-values for global and secondary minima consistent with receptor active-like states stabilized by the G-protein.[44,45] This is not unexpected because partial agonists likely exert their effect in the binary ligand-receptor complex (without G-protein) and interfere with G-protein recruitment.[45]

CONCLUSIONS

The modified binding protocol gives good agreement with experimental free energies of binding, although the range of binding energies tested does not allow us to judge the accuracy of the relative ordering of the binding energies. The results obtained for both proteogenic and artificial peptides are reproducible and the protocol generally converges within 5 μs cumulative simulation time. We have described a conservative set of criteria to judge the convergence of the simulations. Although not described here, simulations using the present protocol should also be suitable and reliable, if computationally more expensive than that described previously for small-molecule ligands.[1] The peptide ligand-receptor combinations tested here give secondary binding minima closer to the extracellular medium than the primary, orthosteric binding site.

COMPUTATIONAL METHODS

General system preparation

The structural models for this study were based on cryo-EM structures of subtypes Y_1R, Y_2R and Y_4R of the neuropeptide Y receptor family, vasopressin V_2 receptor, and the oxytocin receptor. G_s, G_i and G_q α-subunit conformations were generated using chain A of PDB 3SN6[46]

as a template. All structural mutations were corrected to correspond to the naturally occurring human G_S, G_i and G_q sequence (protein identification codes P63092, P63096, P50148 respectively). Missing loops were modelled with the MODELLER[47] tool in UCSF Chimera[48] and the resulting G-protein α-subunit models were truncated at residue Thr9, Gly2, Ala8 for G_s, G_i and G_q respectively.

AmberTools18[49] was used to prepare all systems coordinates and topologies. With tleap, missing hydrogen atoms were added, disulfide bonds were set as specified in the PDB files and N-termini and C-termini charges were generated. Topologies for the receptor, peptides and G protein were generated using the AMBER99SB-ILDN force field.[50,51] A short energy minimization that comprised 500 steps of *steepest descent* followed by 4500 steps of the *conjugate gradient* algorithm, was conducted with sander. The geometry for the peptides were taken from PDB crystal structures. A pre-equilibrated dioleoylphosphatidylcholine (DOPC) lipid bilayer model was used.[52] The membrane was subjected to energy minimization (no restraints) and successively, equilibration runs were performed in the NVT (100 ps; weak harmonic restraints of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ applied on DOPC atoms) and NPT (10 ns; no restraints) ensembles. With the GROMACS tool gmx membed,[53] receptors were aligned and inserted into the membrane according to the orientation in the OPM database.[54] Overlapping lipids and water molecules were deleted. Peptides coordinates were transferred to the receptor and then the appropriate number of sodium and chloride ions was added to simulate a physiological salt concentration of 100mM. The resulting systems were energy minimized and then equilibrated for 300 ns.

Neuropeptide Y receptors. A model for the pancreatic polypeptide in complex with the neuropeptide Y_4R and G_i was built based on the cryo-EM structure of active-state Y_4R (PDB access code 7X9C[39]). Missing loops were modelled with the MODELLER[47] tool in UCSF Chimera.[48] The N and C termini were truncated at Gln35 and Gln341, respectively. Ternary

complex models for an γ-peptide and its enantiomer were based on the model described above, in which the pancreatic polypeptide was removed. Coordinates of the α,γ-hexapeptides were taken from a previous induced-fit docking study.[30] A model for the neuropeptide Y in complex with the neuropeptide Y_1R and G_i was built based on the cryo-EM structure of active-state Y_1R (PDB access code 7X9A[39]). The N and C termini were truncated at Phe29 and Asn336, respectively. A model for the neuropeptide Y in complex with the neuropeptide Y_2R and G_i was built based on the cryo-EM structure of active-state Y_2R (PDB access code 7X9B[39]). The N and C termini were truncated at Leu40 and Arg341, respectively.

Vasopressin V_2 receptor. A ternary vasopressin-V_2R-G_s complex model was built based on the cryo-EM structure of active-state V_2R (PDB access code 7DW9[40]). Missing loops were modelled with the MODELLER[47] tool in UCSF Chimera.[48] The N and C termini were truncated at Glu26 and Leu339, respectively.

Oxytocin receptor. A ternary oxytocin-OTR-G_q complex model was built based on the cryo-EM structure of active-state OTR (PDB access code 7RYC[41]). ICL3 was truncated between Glu242 and Ala258. The N and C termini were truncated at Pro31 and Leu345, respectively.

General setup of the simulations

All simulations were performed using GROMACS 2021.4[55] with the PLUMED 2.7.3 plug-in.[56,57] Periodic boundary conditions were applied in all directions. The simulations were run at constant pressure and temperature in the NPT ensemble. The Berendsen barostat[58] was applied to maintain a pressure of 1 bar. The temperature was held constant at 310 K with temperature coupling achieved by the V-rescale thermostat[58] in three separate coupling groups for (*i*) solvent and ions, (*ii*) protein and peptide, and (*iii*) the DOPC membrane. Bonds involving hydrogen atoms were constrained using the LINCS algorithm,[59] enabling a time step of 2 fs. A cut-off of 1.2 nm was used for short-range van der Waals interactions. Particle mesh Ewald

(PME)[60] was used to treat electrostatic interactions, using a cut-off distance of 1.2 nm. All simulations used the SPC/E water model[61] and included a box size of 9.6×9.7×28.0 nm^3 with approximately 78,000 water molecules, and 278 DOPC molecules.

Metadynamics parameters

Metadynamics simulations in the well-tempered variant (WT)[26] in combination with a funnel-shaped walls in the spirit of path collective variables (PCVs) and funnel metadynamics (FM)[27] were performed to obtain estimates of the binding free-energy profiles. A metadynamics history-dependent bias was applied along the *z* component of the distance between the geometric center of the Cα of residues 6.48 and 3.36 of the receptor and Cα of the deepest Tyr residue of the peptides. This distance was used as the single collective variable. Preliminary metadynamics simulations of 20ns with a bias factor of 60 were performed. Gaussian hills with initial height of 1.67 kcal mol^{-1} were applied every 1 ps. The hill width for the z projection was set to 0.1 nm. Representative structures for each receptor were extracted from the preliminary simulation along the binding path, spanning a range between the bound and the unbound state of the peptide. Prior to simulation, all the representative structures were energy minimized. Based on these structures, well-tempered metadynamics[26] simulations were run in the multiple-walkers[28] scheme using 32 walkers at 310 K. The multiple-walkers metadynamics simulations were conducted with a reduced bias factor of 10 and hills height of 0.47 kcal mol^{-1}. The calculation of the reweighting factor c(t) was enabled by using the keyword CALC_RCT during the metadynamics simulations. Free energies were calculated using the sum_hills function of the PLUMED 2.7.3 plug-in[56,57] and corrected for the loss of translational and rotational freedom of the unbound ligand due to the funnel-like boundaries using the following equations:

$$\Delta G_{binding} = \Delta G_{metaD} + C_{f/sv}$$

$$\Delta G_{metaD} = k_B T \log \left[\frac{\int_{bound} \exp\left(\frac{-F_{(s)}}{k_B T}\right) \mathrm{ds}}{\int_{unbound} \exp\left(\frac{-F_{(s)}}{k_B T}\right) \mathrm{ds}} \right]$$

where $C_{f/sv}$ is the funnel/standard volume correction. The bound states were defined by the position of the global minimum, and the unbound states were defined by values of the distance CV greater than 6.5 nm. An upper limit for the CV was set at 11.5 nm on the basis of the box size and available solvent phase and to avoid interactions with the extracellular side of the receptor. The correction for the standard volume and funnel restraint, $CV_{f/sv}$, was computed as described in refs 34 and 35 according to the formula:

$$C_{f/sv} = -RT \ln \left(\frac{\xi_{bulk}^{metaD} V_{bulk}^{metaD} V_{box}}{8\pi V_{bulk} V_0} \right)$$

where ξ_{bulk}^{metaD} is the fraction of the total possible orientations explored by the ligand in the unbound state, V_0 is the standard volume accessible to a ligand at 1 mol·dm^{-3} concentration, and V_{bulk} is the bulk volume (i.e., $V_{box} - V_{protein+membrane}$). The correction was found to be 1.7 kcal·mol^{-1}.

Reprojection of the free energy surfaces. Reweighting of the free energy surfaces as a function of the xy-projection and z-projection was performed using the Tiwary et al. algorithm.[62]

Convergence criteria. Convergence of the binding metadynamics simulations is based on two important statements, i.e., (*i*) the minimum number of re-crossing events and (*ii*) the calculated error in the reconstructed free energy profiles.

Re-crossing. In this case, a re-crossing is defined as one unbinding/rebinding event, where the ligand explores bound conformations, exits the binding pocket to the extracellular medium, and then returns to the pocket.

Error estimation. Error estimates were calculated according to the block averaging method described in ref 63.

Representative structures from each minimum (± 0.01 nm) were extracted and clustered with the CPPTRAJ module of module of AmberTools18.[49] The trajectory frames were clustered using a hierarchical agglomerative approach with a minimum distance between clusters of 0.4 nm. UCSF ChimeraX,[64,65] was used for data analysis and image preparation. H-bonds were displayed by using default values of Chimera X (distance tolerance of 0.04 nm and an angle tolerance of 20°). Contacts were displayed by setting a distance cut-off of 0.29 nm. Plots were created with Matplotlib.[66]

ASSOCIATED CONTENT

Supporting Information. Schematic representation of the funnel-shaped restraints applied to the metadynamics simulations; chemical structure of the (*i*) α,γ-hexapeptide with (1*R*,2*S*,3*R*)-γ-CBAA (RSR), (*ii*) human pancreatic polypeptide (hPP), (*iii*) porcine neuropeptide Y (pNPY) (*iv*) arginine vasopressin (AVP) and (*v*) oxytocin (OT); results of binding/unbinding MW-WT-metadynamics simulations: (*i*) time-dependent evolution of the calculated free-energy profiles as a function of z-projection, (*ii*) 2D free energy surfaces as a function of the xy-projection and z-projection, (*iii*) block analysis of the metadynamics simulations, (*iv*) binding CV as a function of the sampling time for 32 walkers.

DATA AND SOFTWARE AVAILABILITY

Input files to run binding/unbinding MW-WT-MetaD have been deposited to the public repository of the PLUMED consortium, PLUMED-NEST (plumID:23.035).

AUTHOR INFORMATION

Corresponding Author

Timothy Clark − *Computer-Chemistry-Center, Department of Chemistry and Pharmacy, Friedrich-Alexander-University Erlangen-Nuernberg, 91052 Erlangen, Germany;* orcid.org/0000-0001-7931-4659; Email: Tim.Clark@fau.de

Authors

Jacqueline C. Calderón − *Computer-Chemistry-Center, Department of Chemistry and Pharmacy, Friedrich-Alexander-University Erlangen-Nuernberg, 91052 Erlangen, Germany;* orcid.org/0000-0002-1505-0102, Email: Jacqueline.c.calderon@fau.de.

Eva Plut – *Institute of Organic Chemistry, Faculty of Chemistry and Pharmacy, University of Regensburg, D-93053 Regensburg, Germany;* orcid.org/0000-0002-2630-7284, Email: eva.plut@chemie.uni-regensburg.de.

Max Keller – *Institute of Pharmacy, Faculty of Chemistry and Pharmacy, University of Regensburg, D-93053 Regensburg, Germany;* orcid.org/0000-0002-8095-8627; Phone: (+49) 941-9433329; Email: max.keller@ur.de

Chiara Cabrele – *Institute of Organic Chemistry, Faculty of Chemistry and Pharmacy, University of Regensburg, D-93053 Regensburg, Germany*; orcid.org/0000-0003-1430-573X; Email: chiara.cabrele@chemie.uni-regensburg.de

Oliver Reiser – *Institute of Organic Chemistry, Faculty of Chemistry and Pharmacy, University of Regensburg, D-93053 Regensburg, Germany*; orcid.org/0000-0003-1430-573X; Email: oliver.reiser@chemie.uni-regensburg.de

Francesco Luigi Gervasio − *Pharmaceutical Sciences, University of Geneva, CH1206 Geneva, Switzerland; Institute of Pharmaceutical Sciences of Western Switzerland, CH1206 Geneva, Switzerland; Chemistry Department, University College London, WC1H 0AJ London, U.K.*; orcid.org/0000-0003-4831-5039; Email: Francesco.Gervasio@unige.ch

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This study was supported by the Deutsche Forschungsgemeinschaft as part of Research Training Group GRK 1910 “Medicinal Chemistry of Selective GPCR Ligands” (Fellowships to J.C.C. and E.P). Work in Erlangen was further supported by grants of supercomputer time from the Gauss Alliance on SuperMUC (Project Nr. Pr74su) and from the Erlangen National High Performance Computing Center (NHR@FAU) of the Friedrich-Alexander-University Erlangen-Nuernberg (FAU) (Project Nr b126dc). NHR funding was provided by federal and Bavarian state authorities. NHR@FAU hardware was partially funded by the DFG-Grant No. 440719683.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

MD, molecular dynamics; PDB, Protein Data Bank; RMSD, root mean squared deviation; TM, transmembrane; 1D, one-dimensional; 2D, two-dimensional; CV, collective variable; α,γ-hexapeptide with (1*S*,2*R*,3*S*)-γ-CBAA, SRS; α,γ-hexapeptide with (1*R*,2*S*,3*R*)-γ-CBAA, RSR; CBAA, cyclobutane amino acid; human pancreatic polypeptide, hPP; porcine neuropeptide Y, pNPY; vasopressin, AVP; oxytocin, OT.

REFERENCES

1 Saleh, N.; Ibrahim, P.; Saladino, G.; Gervasio, F. L.; Clark, T. An efficient Metadynamics-Based Protocol to Model the Binding Affinity and the Transition State Ensemble of GPCR-Ligands. *J. Chem. Inf. Model.* **2017**, *57*, 1210–1217.

2 Söldner, C. A.; Horn, A. H. C.; Sticht, H. Binding of Histamine to the H_1 Receptor–A Molecular Dynamics Study. *J. Mol. Model.* **2018**, *24*, 346.

3 Klösel, I.; Schmidt, M. F.; Kaindl, J.; Hübner, H.; Weikert, D.; Gmeiner, P. Discovery of Novel Nonpeptidic PAR2 Ligands. *ACS Med. Chem. Lett.* **2020**, *11*, 1316–1323.

4 Conrad, N.; Söldner, C. A.; Sticht, H. Effect of Ions and Sequence Variants on the Antagonist Binding Properties of the Histamine H1 Receptor. *Int. J. Mol. Sci.* **2022**, *23*, 1420.

5 Saleh, N.; Haensele, E.; Banting, L.; Sopkova-de Oliveira Santos, J.; Whitley, D. C.; Saladino, G.; Gervasio, F. L.; Bureau, R.; Clark, T. A Three-Site Mechanism for Agonist/Antagonist Action on the Vasopressin Receptors. *Angew. Chemie*, **2016**, *128*, 8140–8144; *Angew. Chemie Int. Ed.*, **2016**, *128*, 8008–8012.

6 Lorber, D. M.; Shoichet, B. K. Flexible ligand docking using conformational ensembles. Protein Sci. 1998, 7, 938-950.

7 Agrawal, P.; Singh, H.; Srivastava, H. K.; Singh, S.; Kishore, G.; Raghava, G. P. S. Benchmarking of different molecular docking methods for protein-peptide docking. BMC Bioinform. 2019, 19, 426.

8 Raniolo, S.; Limongelli, V. Ligand binding free-energy calculations with funnel metadynamics. Nature Prot. 2020, 15, 2837-2866.

9 Haensele, E.; Saleh, N.; Read, C. M.; Banting, L.; Whitley, D. C.; Clark, T. Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine-Vasopressin, *J. Chem. Inf. Model*., **2016**, *56*, 1798-1807

10 Haensele, E.; Mele, N.; Miljak, M.; Read, C.; Whitley, D. C.; Banting, L.; Delepee, C.; Sopková-de Oliveira Santos, J.; Lepailleur, A.; Bureau, R.; Essex, J. W.; Clark, T. Conformation and Dynamics of Human Urotensin II and Urotensin Related Peptide in Aqueous Solution, *J. Chem. Inf. Model*., **2017**, *57*, 298-310

11 Basso, N.; Terragno, N. A. History about the Discovery of the Renin-Angiotensin System. *Hypertension*. **2001**, *38*, 1246–1249.

12 Carraway, R.; Leeman, S. E. The Isolation of a New Hypotensive Peptide, Neurotensin, from Bovine Pypothalamic. *J. Biol. Chem.* **1973**, *248*, 6854–6861.

13 Turner, R. A.; Pierce, J. G.; Du Vigneaud, V. The Purification and the Amino Acid Content of Vasopressin Preparations. *J. Biol. Chem*. **1951**, *191*, 21–28.

14 Watkins, H. A.; Au, M.; Hay, D. L. The Structure of Secretin Family GPCR Peptide Ligands: Implications for Receptor Pharmacology and Drug Development. *Drug Disc. Today* **2012**, *17*, 1006–1014.

15 Gruber, C. W.; Muttenthaler, M.; Freissmuth, M. Ligand-Based Peptide Design and Combinatorial Peptide Libraries to Target G Protein-Coupled Receptors. *Curr. Pharmaceut. Des.* **2010**, *16*, 3071–3088.

16 Muratspahi, E.; Freissmuth, M.; Gruber, C. W. Nature-Derived Peptides: A Growing Niche for GPCR Ligand Discovery. *Trends Pharmacol. Sci.* **2019**, *40*, 309–326.

17 Keller, M.; Kuhn, K. K.; Einsiedel, J.; Hübner, H.; Biselli, S.; Mollereau, C.; Wifling, D.; Svobodová, J.; Bernhardt, G.; Cabrele, C.; Vanderheyden, P. M. L.; Gmeiner, P.; Buschauer, A. Mimicking of Arginine by Functionalized Nω-Carbamoylated Arginine as a New Broadly Applicable Approach to Labeled Bioactive Peptides: High Affinity Angiotensin, Neuropeptide Y, Neuropeptide FF and Neurotensin Receptor Ligands as Examples. *J. Med. Chem.* **2016**, *59*, 1925–1945.

18 Schindler, L.; Bernhardt, G.; Keller, M. Modifications at Arg and Ile Give Neurotensin(8-13) Derivatives with High Stability and Retained NTS1 Receptor Affinity. *ACS Med. Chem. Lett.* **2019**, *10*, 960–965.

19 Seebach, D.; Lukaszuk, A.; Patora-Komisarska, K.; Podwysocka, D.; Gardiner, J.; Ebert, M.-O.; Reubi, J. C.; Cescato, R.; Waser, B.; Gmeiner, P.; Hübner, H.; Rougeot, C. On the Terminal Homologation of Physiologically Active Peptides as a Means of Increasing Stability in Human Serum – Neurotensin, Opiorphin, B27-KK10 Epitope, NPY. *Chem. Biodivers.* **2011**, *8*, 711–739.

20 Einsiedel, J.; Hübner, H.; Hervet, M.; Härterich, S.; Koschatzky, S.; Gmeiner, P. Peptide Backbone Modifications on the C-Terminal Hexapeptide of Neurotensin. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2013–2018.

21 Einsiedel, J.; Held, C.; Hervet, M.; Plomer, M.; Tschammer, N.; Hübner, H.; Gmeiner, P. Discovery of Highly Potent and Neurotensin Receptor 2 Selective Neurotensin Mimetics. *J. Med. Chem.* **2011**, *54*, 2915–2923.

22 Konieczny, A.; Braun, D.; Wifling, D.; Bernhardt, G.; Keller, M. Oligopeptides as Neuropeptide Y Y_4 Receptor Ligands: Identification of a High-Affinity Tetrapeptide Agonist and a Hexapeptide Antagonist. *J. Med. Chem.* **2020**, *63*, 8198–8215.

23 Konieczny, A.; Conrad, M.; Ertl, F. J.; Gleixner, J.; Gattor, A. O.; Grätz, L.; Schmidt, M. F.; Neu, E.; Horn, A. H. C.; Wifling, D.; Gmeiner, P.; Clark, T.; Sticht, H.; Keller, M. N-Terminus to Arginine Side-Chain Cyclization of Linear Peptidic Neuropeptide Y Y_4 Receptor Ligands Results in Picomolar Binding Constants. *J. Med. Chem.* **2021**, *64*, 16746–16769.

24 Lai, X.; Tang, J.; ElSayed, M. E. H. Recent Advances in Proteolytic Stability for Peptide, Protein, and Antibody Drug Discovery. *Expert Opin. Drug. Discov.* **2021**, *16*, 1467–1482.

25 Lachmann, D.; Konieczny, A.; Keller, M.; König, B. Photochromic Peptidic NPY Y_4 Receptor Ligands. *Org. Biomol. Chem.* **2019**, *17*, 2467–2478.

26 Barducci, A.; Bussi, G.; Parrinello, *M.* Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett*. **2008**, *100*, 020603.

27 Limongelli, V.; Bonomi, N.; Parrinello, M. Funnel Metadynamics as Accurate Binding Free-Energy Method. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6358–6363.

28 Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *J. Phys. Chem. B* **2006**, *110*, 3533–3539.

29 https://hpc.fau.de, accessed on August 2nd 2022.

30 Plut, E.; Calderón, J. C; Stanojlović, V.; Gattor, A. O.; Höring, C.; Humphrys, L. J; Konieczny, A.; Kerres, S.; Schubert, M.; Keller, M.; Cabrele, C.; Clark, T.; Reiser, O. Stereochemistry-Driven Interactions of α,γ-Peptide Ligands with the Neuropeptide Y Y_4-Receptor. *J. Med. Chem*. **2023**, *66*, 9642−9657.

31 Tikhonova, I. G.; Gigoux, V.; Fourmy, D. Understanding Peptide Binding in Class A G Protein-Coupled Receptors. Mol. Pharmacol. **2019**, *96*, 550-561.

32 Wang, L.; Xu, J.; Cao, S.; Sun, D.; Liu, H.; Lu, Q.; Liu, Z.; Du, Y.; Zhang, C. Cryo-EM Structure of the AVP–Vasopressin Receptor 2–Gs Signaling Complex. *Cell Res.* **2021**, *31*, 932–934.

33 Chini, B.; Mouillac, B.; Ala, Y.; Balestre, M. N.; Trumpp-Kallmeyer, S.; Hoflack, J.; Elands, J.; Hilbert, M.; Manning M.; Jard, S. Tyr115 is the Key Residue for Determining Agonist Selectivity in the V1a Vasopressin Receptor. *EMBO J.* **1995**, *14*, 2176–2182.

34 Deng, Y.; Roux, B. Calculation of Standard Binding Free Energies: Aromatic Molecules in the T4 Lysozyme L99A Mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.

35 Deng, Y.; Roux, B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113*, 2234–2246.

36 Evans, R.; Hovan, L.; Tribello, G. A.; Cossins, B. P.; Estrarellas, C.; Gervasio, F. L. Combining Machine Learning and Enhanced Sampling Techniques for Efficient and Accurate Calculation of Absolute Binding Free Energies. *J. Chem. Theory Comput.* **2020**, *16*, 4641–4654.

37 Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. USA.* **2011**, *108*, 13118–13123.

38 Mattedi, G.; Deflorian, F.; Mason, J. S.; Graaf, C.; Gervasio, F. L. Understanding Ligand Binding Selectivity in a Prototypical GPCR Family. *J. Chem. Inf. Model.* **2019**, *59*, 2830–2836.

39 Tang, T.; Tan, Q.; Han, S.; Diemar, A.; Löbner, K.; Wang, H.; Schüß, C.; Behr, V.; Mörl, K.; Wang, M.; Chu, X.; Yi, C.; Keller, M.; Kofoed, J.; Reedtz-Runge, S.; Kaiser, A.; Beck-Sickinger, A.G.; Qiang Zhao, Q.; Wu, B. Receptor-Specific Recognition of NPY Peptides Revealed by Structures of NPY Receptors. *Sci. Adv.* **2022**, *8*, eabm1232.

40 Zhou, F.; Ye, C.; Ma, X.; Yin, W.; Croll, T. I.; Zhou, Q.; He, X.; Zhang X.; Yang, D.; Wang, P.; Xu, H. E.; Wang, M.-W.; Jiang, Y. Molecular Basis of Ligand Recognition and Activation of Human V_2 Vasopressin Receptor. *Cell Res.* **2021**, *31*, 929–931.

41 Meyerowitz, J. G.; Robertson, M. J.; Barros-Álvarez, X.; Panova, O.; Nwokonko, R. M.; Gao, Y.; Skiniotis, G. The Oxytocin Signaling Complex Reveals a Molecular Switch for Cation Dependence. *Nat. Struct. Mol. Biol.* **2022**, *29*, 274–281.

42 Ibrahim, P.; Wifling, D.; Clark, T. Universal Activation Index for Class A GPCRs. *J. Chem. Inf. Model.* **2019**, *59*, 3938–3945.

43 Hawtin, S. R.; Wesley, V. J.; Parslow, R. A.; Patel, S.; Wheatley, M. Critical Role of a Subdomain of the N-Terminus of the V1a Vasopressin Receptor for Binding Agonists but Not Antagonists; Functional Rescue by the Oxytocin Receptor N-Terminus. *Biochemistry* **2000**, *39*, 13524–13533.

44 Calderón, J. C.; Ibrahim, P.; Gobbo, D.; Gervasio, F. L.; Clark. T. A General Metadynamics Protocol to Simulate Activation/Deactivation of Class A GPCRs: Proof of Principle for the Serotonin Receptor. *J. Chem. Inf. Model.* **2023**, *63*, 3105–3117.

45 Calderón, J. C.; Ibrahim, P.; Gobbo, D.; Gervasio, F. L.; Clark. T. Activation/Deactivation Free-Energy Profiles for the β_2-Adrenergic Receptor: Ligand Modes of Action. Accepted.

46 Rasmussen, S. G. F.; De Vree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the β_2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2011**, *477*, 549–555.

47 Yang, Z.; Lasker, K.; Schneidman-Duhovny, D.; Webb, B.; Huang, C. C.; Pettersen, E. F.; Goddard, T. D.; Meng, E. C.; Sali, A.; Ferrin, T. E.; UCSF Chimera, MODELLER, and IMP: an Integrated Modeling System. *J. Struct. Biol.* **2012**, *179*, 269–278.

48 Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

49 Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. III.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2018, University of California, San Francisco, 2018.

50 Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. L. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712–725.

51 Lindorff-Larsen, K. P.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. I.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* **2010**, *78*, 1950–1958.

52 Siu, S. W. I.; Vácha, R.; Jungwirth, P.; Böckmann, R. A. Biomolecular Simulations of Membranes: Physical Properties from Different Force Fields. *J. Chem. Phys.* **2008**, *128*, 125103.

53 Wolf, M. G.; Hoefling, M.; Aponte-Santamaria, C.; Grubmüller, H.; Groenhof, G. g_membed: Efficient Insertion of a Membrane Protein into an Equilibrated Lipid Bilayer with Minimal Perturbation. *J. Comput. Chem.* **2010**, *31*, 2169–2174.

54 Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **2006**, *22*, 623–625.

55 Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

56 Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free Energy Calculations with Molecular Dynamics. *Comp. Phys. Comm.* **2009**, *180*, 1961–1972.

57 Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED2: New Feathers for an Old Bird. *Comp. Phys. Comm.* **2014**, *185*, 604–613.

58 Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

59 Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J Comput. Chem.* **1997**, *18*, 1463–1472.

60 Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys*. **1993**, *98*, 10089–10092.

61 Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem*. **1987**, *91*, 6269–6271.

62 Tiwary, P.; Parrinello, M. A Time-Independent Free Energy Estimator for Metadynamics. *J. Phys. Chem. B* **2015**, *119*, 736−742.

63 https://www.plumed.org/doc-v2.6/user-doc/html/master-_i_s_d_d-2.html, accessed in May 2023.

64 Goddard, T. D.; Huang, C. C.; Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Meeting Modern Challenges in Visualization and Analysis. *Protein Sci.* **2018**, *27*, 14−25.

65 Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin. T. E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* **2021**, *30*, 70−82.

66 Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.