

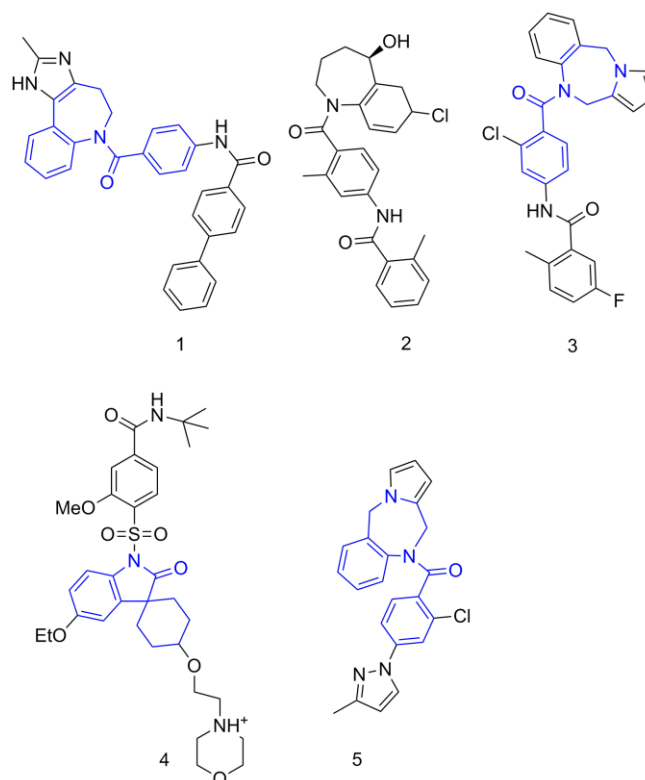
A Three-Site Mechanism for Agonist/Antagonist Selective Binding to Vasopressin Receptors

Noureldin Saleh,^[a] Giorgio Saladino,^[b] Francesco L. Gervasio,^[b] Elke Haensele,^[c] Lee Banting,^[c] David C. Whitley,^[c] Jana Sopkova-de Oliveira Santos,^[d] Ronan Bureau,^[d] Timothy Clark^{[a,c]*}

Abstract: Molecular-dynamics simulations with metadynamics enhanced sampling reveal three distinct binding sites for arginine vasopressin (AVP) within its V_2 -receptor (V_2R). Two of these, the vestibule and intermediate sites, block (antagonize) the receptor and the third is the orthosteric activation (agonist) site. The contacts found for the orthosteric site satisfy all the requirements deduced from mutagenesis experiments. Metadynamics simulations for V_2R and its $V_{1a}R$ -analog give an excellent correlation with experimental binding free energies by assuming that the most stable binding site in the simulations corresponds to the experimental binding free energy in each case. The resulting three-site mechanism separates agonists from antagonists and explains subtype selectivity.

G-protein coupled receptors (GPCRs) represent one of nature's most successful designs. They not only control many biological processes but are also remarkably versatile. One receptor can exert control mechanisms *via* one or more G-proteins and/or the arrestin pathway. These pathways involve intracellular binding partners that have been shown^[1] to affect ligand binding in accord with the ternary-complex model.^[2] GPCR ligands are also diverse in function; they may be agonists, antagonists, inverse or partial agonists and act differently on the alternative pathways (biased agonists). They account for approximately 40% of all marketed drugs.^[3] Much remains unknown about the mechanisms for these diverse functions. Even though high-resolution crystal structures have contributed substantially to clarifying the structure and function of GPCRs, they capture only static snapshots that occur within very complex molecular mechanisms. Molecular-dynamics (MD) simulations represent one of the few possibilities to extend the information provided by crystallographic structures to reveal the atomistic details of GPCR-activation.^[4]

The vasopressin receptors are GPCR drug targets, in particular for antidiuresis and vasopression.^[5] Their natural ligand in mammals is arginine vasopressin (AVP, Cys-Tyr-Phe-Gln-Asn-



Scheme 1. Conivaptan **1**, Tolvaptan **2**, Lixivaptan **3**, Satavaptan **4** and MCF18, **5**.

Cys-Pro-Arg-Gly-NH₂). AVP is closely related to the non-mammalian lysine-vasopressin, oxytocin and urotensin, which all exhibit six-residue rings closed by a disulfide bond. The ring can exhibit different conformations, both in different crystal structures^[6] and solution.^[7]

Three different AVP-receptor-subtypes are known: $V_{1a}R$, $V_{1b}R$ and V_2 .^[8] V_2R is localized in the renal collecting duct^[9] and associated with nephrogenic diabetes insipidus (NDI).^[10] Only two non-peptide ligands for this target, the $V_2R/V_{1a}R$ -antagonist conivaptan **1** and the V_2R -antagonist tolvaptan **2** (Scheme 1) have received FDA approval, although the selective V_2R -agonist peptide desmopressin remains the drug of choice for NDI.^[11]

No crystal structures are available for the vasopressin-receptors but abundant mutagenesis data exist.^[8a] The N-terminus and first extracellular loop (ECL1) of $V_{1a}R$ are essential for ligand binding. The hydrophilic residues Gln108^{2,61}, Lys128^{3,29}, Gln131^{3,32} and Gln185^{4,60} in $V_{1a}R$ contribute to both AVP binding and its ability to activate vasopressin receptors.^[12] No binding pose that satisfies all these contacts has been proposed for AVP in its receptors.

- [a] Noureldin Saleh, Prof. Dr. Timothy Clark*
Computer-Chemie-Centrum
Friedrich-Alexander-Universität Erlangen-Nürnberg,
Nägelsbachstr. 25, 91052 Erlangen, Germany.
E-mail: Tim.Clark@fau.de
- [b] Dr. Giorgio Saladino, Prof. Francesco L. Gervasio
Department of Chemistry and Institute of Structural and Molecular
Biology,
University College London
London WC1E 6BT, United Kingdom.
- [c] Elke Haensele, Dr. Lee Banting, Dr. David C. Whitley, Prof. Dr.
Timothy Clark
School of Pharmacy and Biomedical Sciences,
University of Portsmouth, Portsmouth PO1 2DT, United Kingdom.
- [d] Prof. Jana Sopkova-de Oliveira Santos, Prof. Ronan Bureau
Centre d'Etudes et de Recherche sur le Médicament de Normandie,
UPRES EA 4258 - FR CNRS 3038 INC3M
Boulevard Becquerel, 14032 CAEN Cedex, France

We now report a computational study of the AVP/V₂R system. It not only reveals a binding pose compatible with all the V_{1a}R

Adaptive biasing force simulations^[20] provided an initial guess for possible binding pathways for metadynamics simulations on the

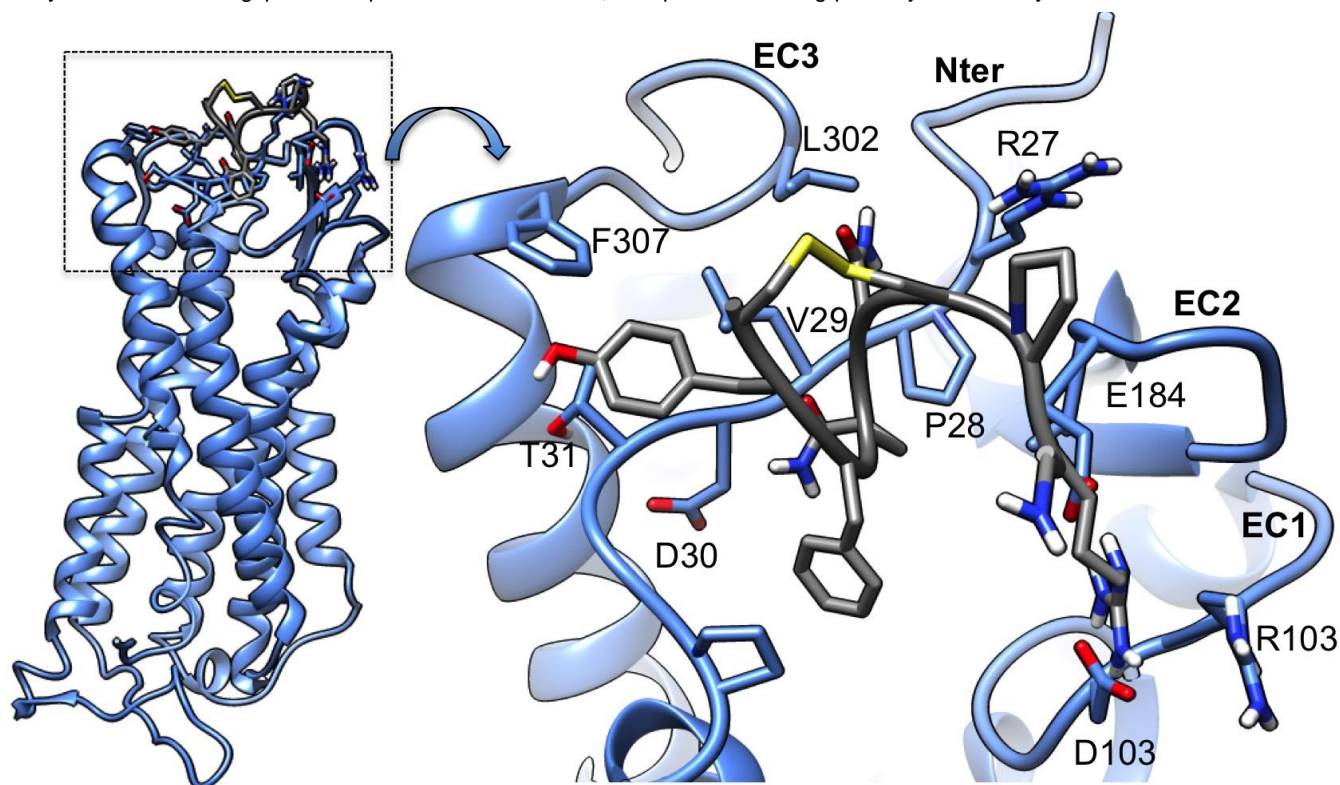


Fig. 1. A representative average structure of the stable AVP-V₂R extracellular vestibule complex obtained from the unbiased 6.5 μ s MD simulation.

mutation data and a quantitative model for ligand binding in V_{1a}R and V₂R but also proposes a new mechanism for the modes of action of different ligand types.

A homology model (see the Supporting Materials, SM) was refined using a 1 μ s MD-simulation, above all to equilibrate the loops and the modeled N-terminus. This subdomain folded to form hydrogen bonds with the receptor surface and remained stable for the last 900 ns of the simulation. Although we cannot be sure, the simulation suggests that the receptor has reached a stable conformation. The N-terminal insertion quickly adopted a conformation resembling those of the neurotensin 1 and μ -opioid receptors.^[1c,13] (see Fig. S1 of the SM). Otherwise, the simulation changed little from the homology model.

Two 6.5 μ s MD-simulations in which two AVP molecules were initially placed in the extracellular solvent above the receptor surface were performed. One resulted in an AVP molecule docking into a vestibule complex that remains stable for the final six μ s of the simulation (Fig. 1). These simulations indicate an energy barrier between the vestibule complex and deeper binding pockets. This barrier was first reported by Dror *et al.* using unbiased MD simulations on the β 2-adrenergic receptor^[14] and extended to the muscarinic M2 and M3 receptors using both unbiased^[15] and accelerated MD (enhanced sampling).^[16] The ligand paused in a vestibule position (between ECL2 and ECL3) before moving into multiple deeper binding sites, including intermediate and orthosteric sites.^[14,17] However, the magnitude of the barriers between the sites and their relative free-energy differences were not quantified. Vestibule complexes have been found, for instance, for the muscarinic M2^[18] and M3^[15] receptors and the extracellular loops are known to play important roles in the biology of GPCRs.^[19] As these simulations suggested the existence of at least two binding sites (vestibule and orthosteric), we investigated the binding path using metadynamics simulations.

binary AVP-V₂R complex (see the SM and Fig. S2 for details). We used a combination of funnel metadynamics^[21] in its well-tempered variant^[22] and a path collective variable (PCV),^[23] a combination that has proven to be very effective for investigating protein-ligand interactions.^[20,24] Three walkers were placed along the path using the multiple walker technique.^[25] Fig. 2 shows a projection of the free-energy landscape onto the path variables and the main minima and barriers involved in binding. The simulation converged after more than 3.9 μ s to reveal three energy minima (see Fig. S3) that correspond to:

1. The vestibule complex (binding free energy, calculated $\Delta G_B = -11.6$ kcal mol⁻¹),
2. An intermediate binding site (calculated $\Delta G_B = -10.3$ kcal mol⁻¹) separated by a low barrier from the vestibule binding site and a higher one (calculated barrier 15 kcal mol⁻¹) from the orthosteric site, and
3. An orthosteric-binding site that satisfies all interactions required by the mutagenesis studies for AVP to interact and activate the vasopressin receptors. This site is the most stable of the three with a calculated ΔG_B of -12.6 kcal mol⁻¹ (experimental $K_I = 0.78$ nM = -12.9 kcal mol⁻¹). Figure S3 of the SM shows the contacts between ligand and receptor found for these three complexes.

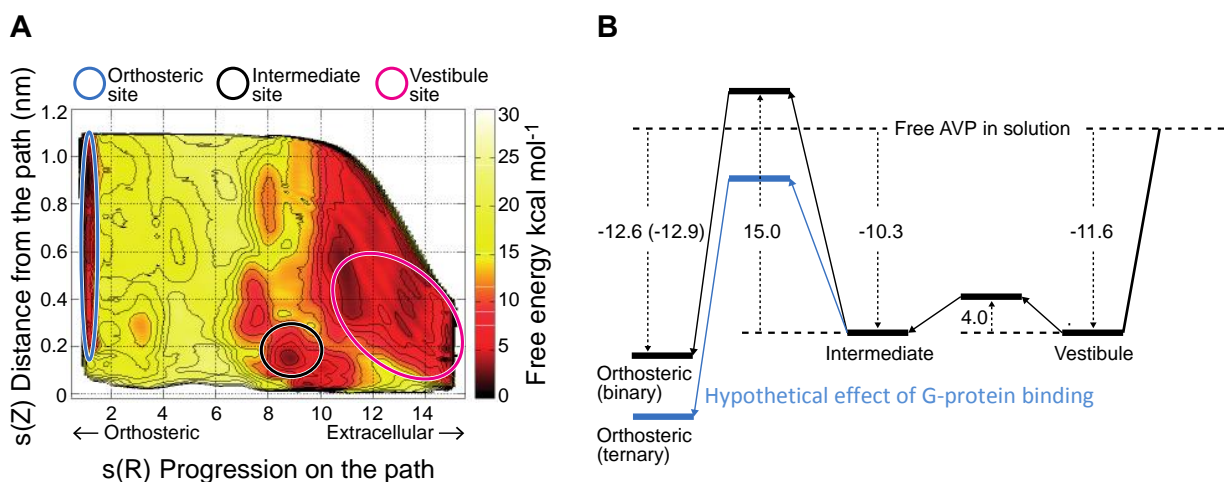


Figure 2. A) The free energy hypersurface obtained from the metadynamics simulation using the path PCV as a function of the progression on the path $s(R)$ (a dimensionless number representing the relative position on the path) and the distance from the path $s(Z)$ in nm (see the SM for more information). B) A schematic diagram of the energy levels found for AVP bound to V_2R . The dotted arrows give energy differences (binding free energies or free energies of activation) in kcal mol^{-1} . The blue arrows and energy levels show the hypothetical effect of binding a G-protein or β -arrestin. The experimental binding free energy is shown in parentheses.

The calculated ΔG_B is most negative for the orthosteric site and agrees very well with experiment. However, both the vestibule and intermediate sites correspond to metastable minima in the free energy landscape and might be druggable. The large barrier between the intermediate and orthosteric sites (the “activation barrier”) strongly suggests a cooperative mechanism between the G-protein (or β -arrestin) and a ligand for receptor activation. We hypothesize that ligand binding in the intermediate site of the receptor in the absence of an intracellular protein binding partner cannot lead to activation because the intermediate site blocks the orthosteric one, which is known^[2] to lead to activation (black profile in Fig. 2B). Furthermore, a plausible cooperative activation mechanism is that the activation barrier is lower in the ternary complex than in the binary one, so that the orthosteric site can be populated in the presence of the G-protein and receptor activation can occur. This hypothetical barrier lowering (blue profile in Fig. 2B) results from the extra stabilization of the ligand binding to the orthosteric site, as proposed by the ternary complex model.^[1,20] In the ternary complex model, the ligand and the G-protein cooperate to reinforce binding to the orthosteric site. This stabilization should also result in a lowering of the activation barrier, as indicated in Figure 2B.

In order to test this interpretation, we simulated four non-peptide ligands to investigate the subtype and functional selectivity toward $V_{1a}R$ and V_2R . The V_2R structure after 7.5 μs MD simulation was used as a template to construct a model for $V_{1a}R$, which was then refined by 0.5 μs of MD simulation. The ligands were a non-selective V_2/V_{1a} antagonist **1**,^[26] a selective V_2 antagonist **3**,^[27] a functionally selective V_2 modulator (G_s antagonist/arrestin partial agonist) **4**,^[28] and a functionally selective V_2 modulator (G_s agonist) **5**.^[29] The experimental free binding energies were derived from the corresponding binding constants, which are shown in Table 1.

The free-energy profiles obtained from metadynamics simulations (Fig. 3) all indicate that V_2R always favors the vestibule, and $V_{1a}R$ the intermediate site. Note that the three-site model leads to the unusual prediction that competitive antagonists need not bind to the same site as the agonists that they displace. This is because ligands in both the vestibule and the intermediate sites block access to the orthosteric site so that, once it has left the receptor, an agonist can no longer bind in the presence of an antagonist in one of the non-activating sites. It is known^[30] that the important residues for non-peptide antagonist binding to $V_{1a}R$ are different to those involved in agonist binding.^[31]

Table 1. Binding affinities (K_i in nM) of the non-peptide ligands used in this study, the experimental free energy of binding, ΔG_B is calculated using $\Delta G_B = -RT \ln(K_i)$ ($T = 298\text{K}$) and the calculated one is taken from the depths of the corresponding minima in the metadynamics profiles shown in Figure 3 (both in kcal mol^{-1}). The assignments of the binding sites refer to the simulations and errors to the funnel sampling next to the global minima according to the scheme reported by Saladino et al.^[32] The underlined simulation data are used in the correlation shown in Figure 4.

Ligand	Effect	Site	$V_{1a}R$		V_2R		
			Experiment	Simulation	Site	Experiment	Simulation
Conivaptan, 1 ^[26]	Non-selective antagonist	Intermediate	0.43	-13.0			
Lixivaptan, 3 ^[27]	Selective V_2 -antagonist	Intermediate	44	-10.1		0.36	-12.9
Satavaptan, 4 ^[28]	Selective V_2 -biased selective G_s -antagonist and partial arrestin agonist	Orthosteric	460	-8.7		2.3	-11.9
		Intermediate					-10.7 \pm 0.2
							-11.0 \pm 0.7
MCF18, 5 ^[29]	V_2 -Selective G_s agonist	Intermediate	106	-9.6		0.54	-12.7
		Orthosteric					-10.1 \pm 2.3
		Vestibule				20	-10.5
							-10.6 \pm 2.3

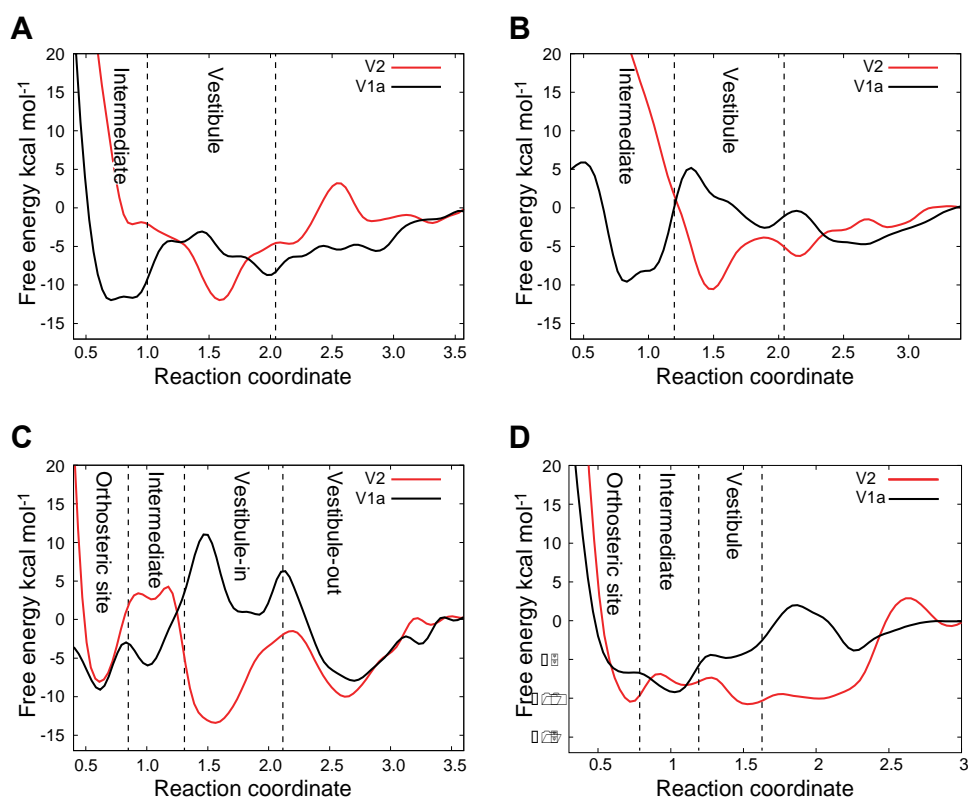


Figure 3. Free-energy profiles from the metadynamics simulations of A) Conivaptan **1**, B) Lixivaptan **3**, C) Satavaptan **4** and D) MCF18 **5**. The approximate value-ranges for the path coordinate for the three sites are: orthosteric 0.5-0.7 nm, intermediate 0.7-1.0 nm and vestibule 1.4-1.6 nm. However, the reaction coordinate does not distinguish well between orthosteric and intermediate around 0.7 nm, so that visual inspection and contacts to the receptor were also used to assign the sites (See Fig. S4). The simulation binding energies listed in Table 1 refer to the depths of the corresponding minima in these curves. The free energy scale is relative to the separated ligand and receptor.

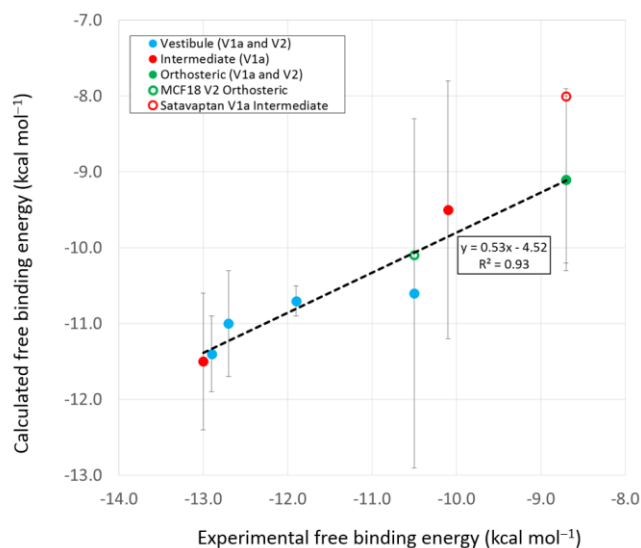


Figure 4. Correlation of the calculated and experimental free binding energies for four $V_{1a}R$ and V_2R ligands (data from Table 1 and binding-site assignments from Fig. 3). The two open circles indicate less stable alternative binding sites for Satavaptan and MCF18 that are not included in the correlation.

Is the three-site model compatible with the known facts? Fig. 4 shows a correlation of the experimental binding free energies against those found for the most stable binding site in Figures 3A-D (the calculated ΔG_B). This correlation is good ($R^2 = 0.93$), even though it involves three different types of binding site. This agreement can be seen to support the three-site model. The

simplest hypothesis is that any ligand that binds most strongly to the orthosteric binding site will be an agonist. Partial agonists most likely bind in equilibrium between the orthosteric and either the intermediate or vestibule sites. Antagonists may bind to either the vestibule or intermediate site (or both). The observed behavior of ligands **1**, **3** and **4** is consistent with this picture. MCF18, **5**, gives the least clear picture. It activates less strongly than AVP but binds only weakly, so that its experimental classification is ambiguous.

The simulations have provided a wealth of information and a new hypothesis that explains many features of the complex interplay between GPCRs, their ligands and their intracellular binding partners. Previous free-energy perturbation simulations have been used successfully to predict relative binding-free-energy differences ($\Delta\Delta G$) between GPCR-ligands.^[33] However, as far as we are aware, the current simulations are unique because they provide absolute binding energies with unprecedented accuracy, classify unknown ligands as agonists, antagonists or partial agonists and reveal unsuspected sources of subtype-specificity such as, for instance, the fact that antagonists bind to a different site (vestibule) in V_2R than in $V_{1a}R$ (intermediate). It now remains to test this hypothesis for other GPCRs and to demonstrate that simulations can reproduce the effect of the intracellular binding partner on ligand binding and *vice versa*. Above all, however, this work has demonstrated the power of modern hard- and software combinations in GPCR research.

Acknowledgements

This work was supported by the *Deutsche Forschungsgemeinschaft* as part of GK1910 "Medicinal Chemistry of Selective GPCR Ligands", by the European Union project "Peptide Research Network of Excellence" PeRENE as

part of the Interreg IVA France (Channel) – England 2007-2014 program (Interreg EU) and by a generous grant of computer time on SuperMUC at the *Leibniz Rechenzentrum*, Munich. GS and FLG acknowledge the EPSRC [grant: EP/M013898/1] for financial support

Keywords: Molecular Dynamics, Metadynamics, Receptors, Hormones, GPCR.

- [1] a) S. G. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Tian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara, B. K. Kobilka, *Nature* **2011**, *477*, 549-555; b) A. M. Ring, A. Manglik, A. C. Kruse, M. D. Enos, W. I. Weis, K. C. Garcia, B. K. Kobilka, *Nature* **2013**, *502*, 575-579; c) W. Huang, A. Manglik, A. J. Venkatakrishnan, T. Laeremans, E. N. Feinberg, A. L. Sanborn, H. E. Kato, K. E. Livingston, T. S. Thorsen, R. C. Kling, S. Granier, P. Gmeiner, S. M. Husbands, J. R. Traynor, W. I. Weis, J. Steyaert, R. O. Dror, B. K. Kobilka, *Nature* **2015**, *524*, 315-321.
- [2] A. De Lean, J. M. Stadel, R. J. Lefkowitz, *J. Biol. Chem.* **1980**, *255*, 7108-7117.
- [3] D. Filmore, *Modern Drug Disc.* **2004**, *7*, 24-28.
- [4] R. O. Dror, D. H. Arlow, P. Maragakis, T. J. Mildorf, A. C. Pan, H. F. Xu, D. W. Borhani, D. E. Shaw, *Proc. Nat. Acad. Sci. USA* **2011**, *108*, 18684-18689.
- [5] J. Howl, M. Wheatley, *Gen. Pharmacol.* **1995**, *26*, 1143-1152.
- [6] a) C. K. Wu, B. Hu, J. P. Rose, Z. J. Liu, T. L. Nguyen, C. Zheng, E. Breslow, B. C. Wang, *Prot. Sci.* **2001**, *10*, 1869-1880; b) J. P. Rose, C. K. Wu, C. D. Hsiao, E. Breslow, B. C. Wang, *Nature Struct. Biol.* **1996**, *3*, 163-169.
- [7] E. Haensele, L. Banting, D. C. Whitley, T. Clark, *J. Mol. Model.* **2014**, *20*, 2485.
- [8] a) S. R. Hawtin, V. J. Wesley, R. A. Parslow, S. Patel, M. Wheatley, *Biochem.* **2000**, *39*, 13524-13533; b) C. W. Gruber, M. Muttenthaler, M. Freissmuth, *Curr. Pharmaceut. Des.* **2010**, *16*, 3071-3088.
- [9] E. Kojro, P. Eich, G. Gimpl, F. Fahrenholz, *Biochem.* **1993**, *32*, 13537-13544.
- [10] Y. Pan, P. Wilson, J. Gitschier, *J. Biol. Chem.* **1994**, *269*, 31933-31937.
- [11] M. Manning, A. Misicka, A. Olma, K. Bankowski, S. Stoev, B. Chini, T. Durroux, B. Mouillac, M. Corbani, G. Guillon, *J. Neuroendocrin.* **2012**, *24*, 609-628.
- [12] a) M. Hibert, J. Hoflack, S. Trumpp-Kallmeyer, B. Mouillac, B. Chini, E. Mahe, N. Cotte, S. Jard, M. Manning, C. Barberis, *J. Recept. Sig. Transd. Res.* **1999**, *19*, 589-596; b) N. Cotte, M. N. Balestre, A. Aumelas, E. Mahe, S. Phalipou, D. Morin, M. Hibert, M. Manning, T. Durroux, C. Barberis, B. Mouillac, *Eur. J. Biochem.* **2000**, *267*, 4253-4263.
- [13] P. Eglhoff, M. Hillenbrand, C. Klenk, A. Batyuk, P. Heine, S. Balada, K. M. Schlinkmann, D. J. Scott, M. Schutz, A. Pluckthun, *Proc. Nat. Acad. Sci. USA* **2014**, *111*, E655-662.
- [14] R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. Shan, H. Xu, D. E. Shaw, *Proc. Nat. Acad. Sci. USA* **2011**, *108*, 13118-13123.
- [15] A. C. Kruse, J. Hu, A. C. Pan, D. H. Arlow, D. M. Rosenbaum, E. Rosemond, H. F. Green, T. Liu, P. S. Chae, R. O. Dror, D. E. Shaw, W. I. Weis, J. Wess, B. K. Kobilka, *Nature* **2012**, *482*, 552-556.
- [16] K. Kappel, Y. Miao, J. A. McCammon, *Quart. Rev. Biophys.* **2015**, *48*, 479-487.
- [17] R. O. Dror, H. F. Green, C. Valant, D. W. Borhani, J. R. Valcourt, A. C. Pan, D. H. Arlow, M. Canals, J. R. Lane, R. Rahmani, J. B. Baell, P. M. Sexton, A. Christopoulos, D. E. Shaw, *Nature* **2013**, *503*, 295-299.
- [18] K. Haga, A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga, T. Kobayashi, *Nature* **2012**, *482*, 547-551.
- [19] M. Wheatley, D. Wootton, M. T. Conner, J. Simms, R. Kendrick, R. T. Logan, D. R. Poyner, J. Barwell, *Brit. J. Pharmacol.* **2012**, *165*, 1688-1703.
- [20] E. Darve, D. Rodriguez-Gomez, A. Pohorille, *J. Chem. Phys.* **2008**, *128*, 144120.
- [21] V. Limongelli, M. Bonomi, M. Parrinello, *Proc. Nat. Acad. Sci. USA* **2013**, *110*, 6358-6363.
- [22] a) A. Laio, M. Parrinello, *Proc. Nat. Acad. Sci. USA* **2002**, *99*, 12562-12566; b) A. Barducci, G. Bussi, M. Parrinello, *Phys. Rev. Lett.* **2008**, *100*, 020603.
- [23] a) J. Juraszek, G. Saladino, T. S. van Erp, F. L. Gervasio, *Phys. Rev. Lett.* **2013**, *110*, 108106; b) D. Branduardi, F. L. Gervasio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 054103.
- [24] a) V. Limongelli, M. Bonomi, L. Marinelli, F. L. Gervasio, A. Cavalli, E. Novellino, M. Parrinello, *Proc. Nat. Acad. Sci. USA* **2010**, *107*, 5411-5416; b) L. Troussicot, F. Guilliere, V. Limongelli, O. Walker, J. M. Lancelin, *J. Am. Chem. Soc.* **2015**; c) G. Grazioso, V. Limongelli, D. Branduardi, E. Novellino, C. De Micheli, A. Cavalli, M. Parrinello, *J. Am. Chem. Soc.* **2012**, *134*, 453-463.
- [25] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, M. Parrinello, *J. Phys. Chem. B* **2006**, *110*, 3533-3539.
- [26] A. L. Crombie, T. M. Antrilli, B. A. Campbell, D. L. Crandall, A. A. Failli, Y. He, J. C. Kern, W. J. Moore, L. M. Nogle, E. J. Trybulski, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3742-3745.
- [27] J. M. Matthews, W. J. Hoekstra, A. B. Dyatkin, L. R. Hecker, D. J. Hlasta, B. L. Poulter, P. Andrade-Gordon, L. de Garavilla, K. T. Demarest, E. Ericson, J. W. Gunnet, W. Hageman, R. Look, J. B. Moore, C. H. Reynolds, B. E. Maryanoff, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2747-2752.
- [28] C. Serradeil-Le Gal, *Cardiovasc. Drug Rev.* **2001**, *19*, 201-214.
- [29] F. Jean-Alphonse, S. Perkovska, M. C. Frantz, T. Durroux, C. Mejean, D. Morin, S. Loison, D. Bonnet, M. Hibert, B. Mouillac, C. Mendre, *J. Am. Society Nephrol.* **2009**, *20*, 2190-2203.
- [30] R. Lemmens-Gruber, M. Kamyar, *Cell. Mol. Life Sci.* **2006**, *63*, 1766-1779.
- [31] M. Thibonnier, P. Coles, A. Thibonnier, M. Shoham, M., *Prog. Brain Res.* **2002**, *139*, 179-196.
- [32] G. Saladino, L. Gauthier, M. Bianciotto, F. L. Gervasio, *J. Chem. Theory. Comput.* **2012**, *8*, 1165-1170.
- [33] L. Boukharta, H. Gutierrez-de-Teran, J. Åqvist, *PLoS Comput. Biol.* **2014**, *10*, e1003585.