Journal of Computer-Aided Molecular Design, **12:** 275–287, 1998.
KLUWER/ESCOM

Molecular modeling of the human vasopressin V2 receptor/agonist complex

Cezary Czaplewski, Rajmund Kaźmierkiewicz and Jerzy Ciarkowski*
Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80–952 Gdańsk, Poland

Received 23 May 1997; Accepted 11 November 1997

Key words: AMBER 4.1, constrained simulated annealing, GPCR, molecular dynamics, receptor-agonist interaction, vasopressin V2 receptor

Summary

The V2 vasopressin renal receptor (V2R), which controls antidiuresis in mammals, is a member of the large family of heptahelical transmembrane (7TM) G protein-coupled receptors (GPCRs). Using the automated GPCR modeling facility available via Internet (http://expasy.hcuge.ch/swissmod/SWISS-MODEL.html) for construction of the 7TM domain in accord with the bovine rhodopsin (RD) footprint, and the SYBYL software for addition of the intra- and extracellular domains, the human V2R was modeled. The structure was further refined and its conformational variability tested by the use of a version of the Constrained Simulated Annealing (CSA) protocol developed in this laboratory. An inspection of the resulting structure reveals that the V2R (likewise any GPCR modeled this way) is much thicker and accordingly forms a more spacious TM cavity than most of the hitherto modeled GPCR constructs do, typically based on the structure of bacteriorhodopsin (BRD). Moreover, in this model the 7TM helices are arranged differently than they are in any BRD-based model. Thus, the topology and geometry of the TM cavity, potentially capable of receiving ligands, is in this model quite different than it is in the earlier models. In the subsequent step, two ligands, the native [arginine^{8}]vasopressin (AVP) and the selective agonist [D-arginine^{8}]vasopressin (DAVP) were inserted, each in two topologically non-equivalent ways, into the TM cavity and the resulting structures were equilibrated and their conformational variabilities tested using CSA as above. The best docking was selected and justified upon consideration of ligand-receptor interactions and structure-activity data. Finally, the amino acid residues were indicated, mainly in TM helices 3-7, as potentially important in both AVP and DAVP docking. Among those Cys^{112}, Val^{115}-Lys^{116}, Gln^{119}, Met^{123} in helix 3; Glu^{174} in helix 4; Val^{206}, Ala^{210}, Val^{213}-Phe^{214} in helix 5; Trp^{284}, Phe^{287}-Phe^{288}, Gln^{291} in helix 6; and Phe^{307}, Leu^{310}, Ala^{314} and Asn^{317} in helix 7 appeared to be the most important ones. Many of these residues are invariant for either the GPCR superfamily or the neurophyseal (vasopressin V2R, V1aR and V1bR and oxytocin OR) subfamily of receptors. Moreover, some of the equivalent residues in V1aR have already been found critical for the ligand affinity [Mouillac et al., J. Biol. Chem, 270 (1995) 25771].

Introduction

The nonapeptide hormone vasopressin (CYFQNCPRG-NH_2, AVP) regulates water absorption in kidneys via the interaction with the V2 receptor (V2R). A disorder of this interaction causes a serious ailment viz. nephrogenic diabetes insipidus. V2R is closely related to three other neurohypophyseal receptors, i.e. the vascular V1a and the pituitary V1b receptors of AVP (V1aR and V1bR, respectively), and to the oxytocin (CYIQNCPLG-NH_2, OT) receptor (OTR). This family of four receptor subtypes belongs to the large group of G protein-coupled receptors (GPCRs), which during the past decade have been recognized as the most important superfamily of transmembrane proteins. GPCRs mediate, in tandem with the intracellular heterotrimeric GTP-binding proteins (G proteins), the relay of information from extracellular stimuli (neu-

*To whom correspondence should be addressed.

Table 1. Identity matrix for the V2, V1a, V1b and OT human receptors

	V2R	V1aR	V1bR
V1aR	0.35		
V1bR	0.40	0.47	
OTR	0.40	0.44	0.48

rotransmitters, hormones, odorants, light) to intracellular second messenger systems. Even though some 1000 sequences of GPCRs have been determined thus far, no details on their native architecture are known.

Milestones towards approaching this task have been (i) a low resolution (9 Å) structure of bovine [1, 2] and frog (6 Å) [3] rhodopsin followed by (ii) the sequence analysis of Baldwin [4]. They provided for the first time a reliable basis for (iii) convincing GPCR architecture predictions implementing most of the known experimental constraints [5], and (iv) resulted eventually in an automated modeling facility [6]. These studies, while confirming the GPCR to belong to the seven-helical transmembrane (7TM) proteins, simultaneously proved that spatial organisation of a GPCR is different than that of bacteriorhodopsin (BRD) [7, 8] used hitherto as a template in most of the GPCR modeling [9]. Moreover, our comparison of the interiors of the TM bundles in RD and BRD reveals that the former is much more capacious from the extracellular side than the latter. Precisely, the entrance to the ligand cleft in RD is some 15 Å wide and 20 Å long while that of BRD is only ca. 10 Å wide and 20 Å long, see Figure 1.

All four neurophyseal receptor subtypes share a high degree (35–50%) of sequence identity [10, 11]. Yet, the V1a, V1b and OT receptors are selectively coupled to G proteins of the $G_{q/11}$ family [12] which mediate the activation of distinct isoforms of phospholipase Cβ (PLCβ), resulting in the breakdown of phosphoinositide lipids (PI hydrolysis). The V2 receptor, on the other hand, preferentially activates the G_s protein [12], resulting in the activation of adenylyl cyclase. This is to an extent reflected in the sequence similarities, which are higher among the V1a, V1b and OT receptors than between any of the former and the V2 receptor, see Table 1. Looking faint enough in Table 1, this difference becomes quite profound when the comparison is limited to the intracellular loops 2 (IL2) and 3 (IL3) only.

These loop sequences are very well conserved among V1aR, V1bR and OTR (50 to 78% identity) but diverge with the V2R subtype (23 to 30% identity) [11]. Interestingly, in accord with this property, Liu and Wess have recently provided evidence that IL3 of V2R plays a decisive role in the G_s protein-adenylyl cyclase coupling while IL2 of V1aR, V1bR and OTR play a decisive role in the coupling of the $G_{q/11}$ protein to phospholipase C [13].

In contrast to the differences on the intracellular side, the extracellular 'halves' of the subtypes do not differ significantly among themselves. In fact the extracellular loop 1 (EL1) and the C-terminal part of the extracellular loop 2 (EL2) belong, together with the helices TM2, TM3, TM6 and TM7, to the most conserved parts among the four subtypes [10, 11] with identity reaching 87%. In this respect, it seems a reasonable assumption that more or less common recognition and binding modes occur for both OT and VP in their host receptors, even if scarce residues diverge in the sequences while some of them, i.e., Asp^{103} in EL1 of V2R, Phe in an equivalent position of OTR and Tyr in equivalent positions of V1aR and V1bR, are postulated to be responsible for ligand selectivity [14, 15]. Perhaps the most notable feature on the extracellular side, common for a majority of the members of the GPCR superfamily, is a putative disulphide bond, keeping EL1 and EL2 together and made of two invariant cysteines: one from the EL1/TM3 edge and the other from the center of EL2 [16].

Our interest has been the development of a reliable model of the vasopressin V2 receptor. We believe that this would be helpful not only for detailed studies of the V2R/bioligand (hormone and/or antagonist) interactions but also for extentions of this model to other neurophyseal receptor/ bioligand systems.

Methods

Protein modeling, computing and display

The 7TM domain of human V2R was obtained from the Swiss-Model protein modeling server [6], given the alignment and helix definitions in agreement with those proposed by Baldwin [4, 17]. This would correspond to 26 residues per TM helix with the specific TM sequences defined as follows. TM1: R38-A63, TM2: A72-L97, TM3: R113-H138, TM4: A154-A179, TM5: T204-F229, TM6: T269-A294 and TM7: A305-S330. These sequences are distinguished within

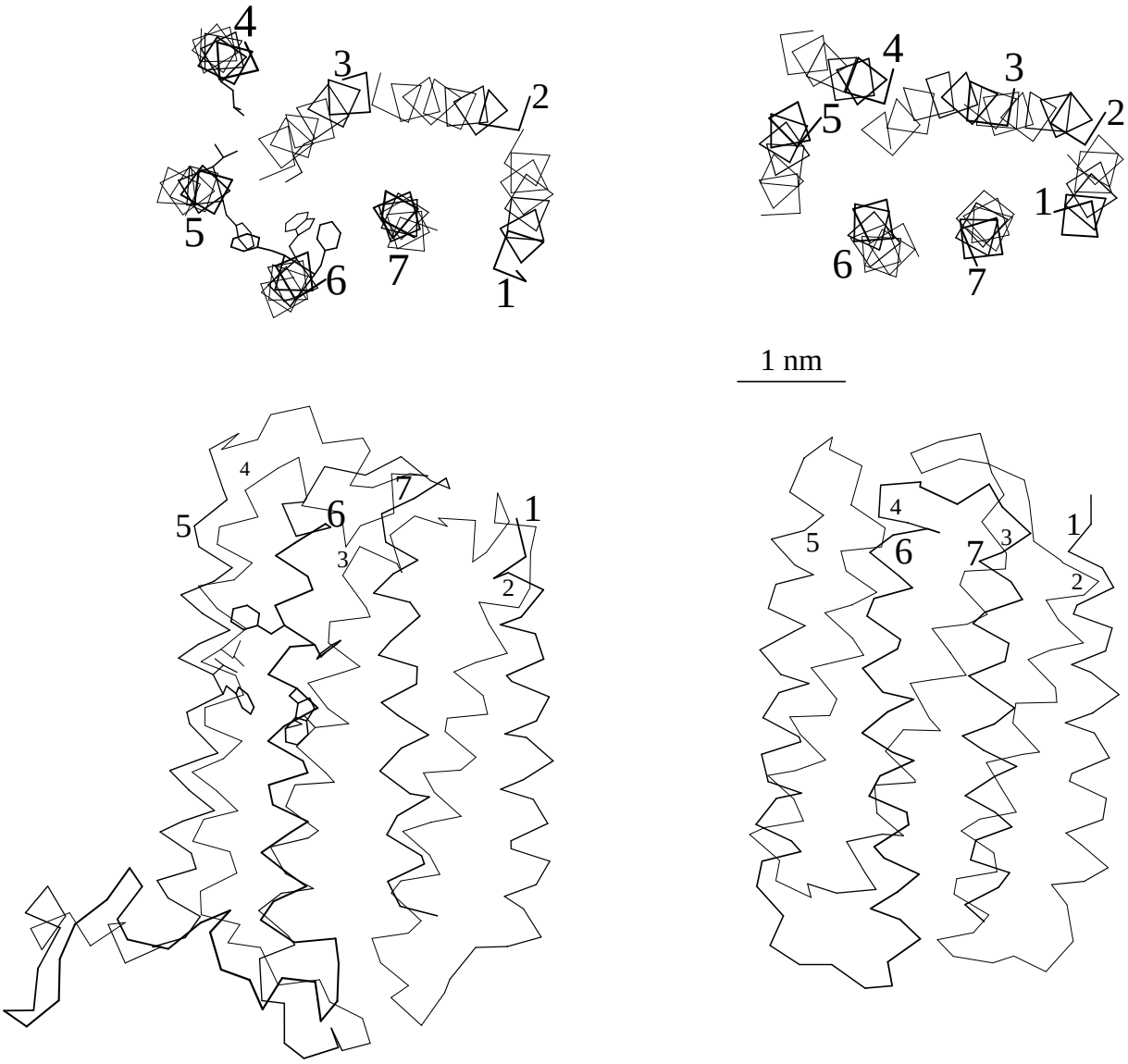

Figure 1. The comparison (C^{α} trace) of spatial TM domain organization in V2R (typical of bovine RD [5,6]; left) with that in BRD [8] (right). Both 7TM proteins are drawn to the common scale given by the 1 nm (10 Å) bar. The digits mark consecutive TM helices and are attached to their extracellular tips. The digit's size correlates to the depth-cueing signified by the thickness of the lines. Top: extracellular views roughly perpendicular to the membrane normal. The EL loops are removed for clarity. Bottom: lateral views, parallel to the membrane normal. It is seen that the extracellular cavity opens in V2R (top left) with a nearly regular pentagonal rim, flanked by TM3-TM7, which narrows to the interior. The corresponding cavity in BRD (top right) opens with a longitudinal heptagonal hole widening to the interior and flanked by all seven TM helices. The amino acid residues making the hydrophobic floor in the V2R cavity (see text) are shown by their explicit side chains.

the total V2R sequence [18] by using bold italic bordered with dots: MLMASTTSAV PGHPSLPSLP SNSSQERPLD TRDPLLA***.RAE LALLSIVFVA VAL SNGLVLA ALA.***RRGRRGH W***.APIHVFIGH LCLA DLAVAL FQVLPQL.*** AWK ATDRFRGPDA LC***.RA VKYLQM VGMYASSYMI LAMTLDRH.***RA ICRP MLAYRH GSG***.AHWNRPV LVAWAFSLLL SLPQ LFIFA.***Q RNVEGGSGVT DCWACFAEPW GRR ***.TYVTWIA LMVFVAPTLG IAACQVLIF.***R EIHA SLVPGP SERPGGRRRG RRTGSPGEGA HVSAA VAK.***TV RMTLVIVVVY VLCWAPFFLV QLWA.***AW DPEA PLEG***.APFVLL MLLASLNSCT NPWIYASF SS.*** SVSSELRSLL CCARGRTPPS LGPQDESCTT ASSSLAKDTS S (total 371 residues). The missing

loops and the amino domain were adjoined, using the Biopolymer module of the SYBYL suite of programs [19] on either a Sun or an IBM workstation. Briefly, the algorithm searches for a specified fragment through a dedicated protein-segment database constructed from the Brookhaven Protein Data Bank [20], to fulfill the distances and coordinates involving the end residues of the loop; in this case identical with the end residues of the respective two helices. When the retrieved fragment is inserted into the target molecule small adjustments to the torsion angles at the fragment are performed using the TWEAK algorithm [21, 22] to avoid discontinuities. No carboxyl domain was added so that our model receptor, missing 41 C-terminal residues, eventually consisted of 330 amino acids. The putative Cys^{112}-Cys^{192} disulphide, coupling EL1 with EL2 [16], was set manually. Initial ligand docking was accomplished using SYBYL as well; this is described in more detail in a subsequent paragraph. All high-power computing, involving molecular mechanics and/or dynamics, was executed using the AMBER v. 4.1 suite of programs [23] on an SGI Power Challenge 8× R10000 or an IBM SP2 15× POWER2 at the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, or on a CRAY Y-MP/EL-98 at the Interdisciplinary Center for Math and Computer Modeling at the University of Warsaw (ICM UW). The images for presentation were prepared using either RasMol [24] or MOLSCRIPT [25].

Neurophyseal receptors multiple sequence alignment

The multiple sequence alignment we refer to or otherwise imply in this work concerns 11 neurophyseal receptors from various species, viz: four V2R (pig, rat, human and bovine), three V1aR (rat, human and sheep), two V1bR (human and rat) and two OTR (pig and human). They all have been compiled, aligned and provided by Vriend et al. [26].

Initial 3D models of V2R/bioligand complexes

A visual inspection confirmed that the arrangement of the TM helices in the crude V2R structure was, as expected, virtually identical with the arrangement of equivalent helices in bovine rhodopsin (RD) [1], as seen in its high resolution refinement [5]. This arrangement includes a cavity on the extracellular side of the receptor, fringed with TM2-TM7 and capable of fitting the pressinoic ring together with a good part of the C-terminal tail of AVP, see Figure 1 and Results below. For the purpose of docking, the solid-state structure of [desamino^1]oxytocin (dOT) [27, 20] was converted to AVP using SYBYL. Given subsequent extensive conformational changes of an initial ligand architecture (vide infra), any reasonable AVP structure would be good at this point. The reason for using this structure [20] was that it provided a convenient set of initial coordinates in the proper pdb format.

Just like the cavity has provided a prerequisite for *where* one might possibly dock AVP, the structure-activity data, emphasizing the role of the bulky and hydrophobic residues 1–3 in binding [28, 29], provided a prerequisite for *how* one could possibly dock the bioligand. An additional hint was offered by an observation of Fahrenholz and co-workers [15, 30] that Asp^{103} in EL1 interacts with the ligand's Arg^8 while effecting affinity. Thus, we attempted an initial docking so that the AVP N-terminal hydrophobic triad Cys^1-Tyr^2-Phe^3 could sit on the floor of the V2R cavity, while AVP Arg^8 could be proximal to Asp^{103} in V2R EL1 at the same time. Allowing some molecular editing, mainly consisting of rotations of the C-terminal tail against the pressinoic ring, this attempt could have been completed in two alternative ways distinguished by an opposite facing (termed Dock I and Dock II from here on) of the ligand cyclic part against the receptor cavity. Similar molecular editing could also lead to two starting complexes of V2R with [D-Arg^8]AVP and [desamino^1, D-Arg^8]AVP, known as as V2R agonist and superagonist, respectively, of increased selectivity [31]. Prior to further elaboration, any model was locally relaxed by 1000–2000 steps of AMBER 4.1 minimization. The minimization was brought about under the following constraints: (i) all TM C^α carbons of V2R could not move at all, (ii) the ligand was anchored to the V2R at three points by harmonic constraints. The distances, d, selected for these constraints were, for Dock I: C^γ(TM6:Phe^{287})-C^γ(AVP:Tyr^2), C^γ(TM6:Phe^{288})-C^γ(AVP:Phe^3) and C^γ(IL1:Asp^{103})-C^ζ(AVP:Arg^8); for Dock 2: C^γ(TM6:Phe^{287})-C^γ(AVP:Phe^3) and C^γ(TM6:Phe^{288})-C^γ(AVP:Tyr^2), i.e., the first two selections inverted and the last one unchanged C^γ(IL1:Asp^{103})-C^ζ(AVP:Arg^8). The harmonic constants were $k = 200$ kcal mol^{-1} Å^{-2} and $d =$ 4–5 Å.

Table 2. Constrained simulated annealing (CSA)[a]

Stage	Steps (fs)	Duration (ps)	T_0^b (K)	TAUTP[c]
1. Fast heating	0–1000	1	10–600 (linear)	0.2
2. Hot equilibration	1001–3000	2	600	0.2
3. Slow cooling	3001–10000	7	0–300 (linear)	4.0–0.2 (linear)

[a] Constraints: positional for all C^α carbons contributing to the 7TM helices, harmonic ($k = 50$ kcal mol^{-1} rad^{-2}) for all peptides, harmonic ($k = 10$ kcal mol^{-1} rad^{-2}) for all improper dihedrals forming chirality centers; dielectrics: distance-dependent, no water and lipid bilayer included.

[b] This is the temperature of a fictitious environment serving, together with another parameter TAUTP, as controls during the simulated heating (cooling) of the object.

[c] TAUTP is a temperature relaxation time, inversely related to the heat conductivity between the object and the generic environment.

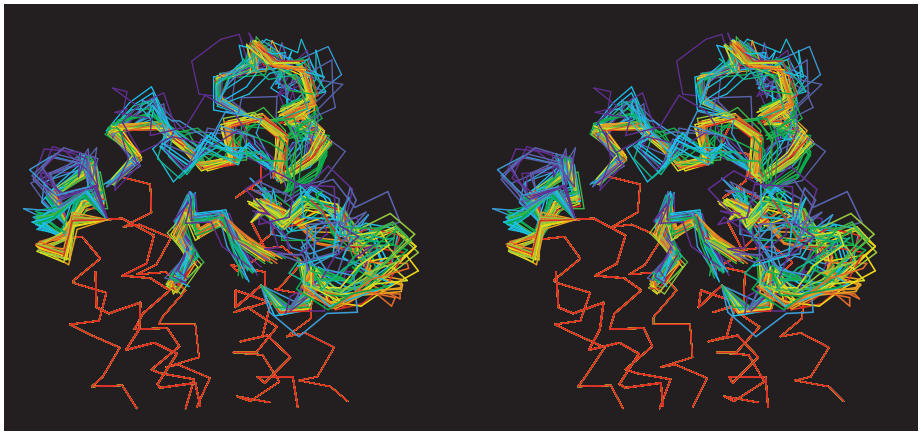

Figure 2. The convergence of 40 consecutive CSA runs of the V2R/DAVP complex in stereo. Each C^α trace represents the ultimate snapshot of a respective CSA run. The view is roughly from the right side of the left-bottom projection in Figure 1. Color code: the rainbow from blue to red traces the structures from the initial to the final CSA, respectively. The loops and the ligand as the only unconstrained components of the complex produce the rainbow while the constrained 7TM helices are uniform. EL1 is on the right, EL2 in the top middle, EL3 on the left and DAVP (disconnected) in the interior. A decay of dispersion over the rainbow from blue to red is clearly seen for all moving chains, although the DAVP chain as located within the constrained cavity moves quite moderately. Similar features (not shown) were observed for both Dock I and Dock II V2R/AVP models.

Constrained simulated annealing of V2R in vacuo

A subsequent step consisted of constrained simulated annealing (CSA), which is a variant of molecular dynamics (MD) with variable temperature. Although our ultimate goal is an implementation of lipid and water within an entirely unconstrained simulation under periodic boundary conditions, until we will have the lipid parameterization ready, we will perform CSA in vacuo with positional constraints imposed on all C^α carbon atoms belonging to the 7TM helices, to prevent the 7TM template from deformations or even a total breakdown. Although physically unjustified, these constraints were a necessary compromise between a need for an as free as possible optimization of ligand packing and a simultaneous requirement for the receptor to conform as much as possible to its putative image under otherwise demolishing conditions typical of unconstrained SA in vacuo, see Table 2. In addition, harmonic constraints have been imposed on peptide bonds and chiral centers to preserve geometrical and chiral integrity of the system. Generally, no constraints were imposed on a ligand during CSA, although sometimes, especially at initial steps, the constraints anchoring a ligand (as described in the former paragraph) were applied. Distance-dependent dielectric permittivity was chosen as most appropriate for the simulation in vacuo. A typical protocol for a single CSA run is given in Table 2. T_0 and TAUTP are the parameters controlling the simulation of thermal coupling of the object to the environment [32]. Higher temperatures than 600 K (up to 1200 K: see Step 1 in

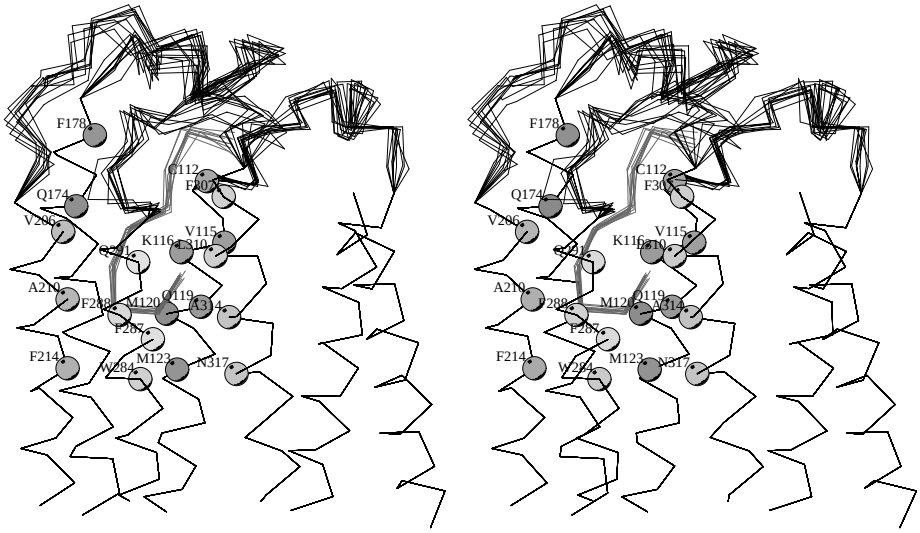

Figure 3. V2R/AVP:DockI in stereo, an overview. This is a collection of the snapshots (C^{α} trace) terminating the final ten converging CSA runs. The view is roughly the same as that in Figure 1, bottom-left. The amino acid residues from TM4-TM7 which interact with the ligand are represented by balls. The ligand is so oriented that its N-terminal end is closer to the hydrophobic floor.

Table 2) were also applied, however, 600 K was found to be of the most universal use. A CSA cycle as defined in Table 2 is very fast. In fact, for a V2R/bioligand complex it can be performed some 50 times overnight on a supercomputer. Thus, a typical CSA simulation consisted of a sequence of 30–40 runs so correlated that the output geometry of the *i*-th run was the input geometry for the $(i + 1)$th run.

Results

The receptor

A landmark feature of any 7TM arrangement based on the structure of RD [1–3] is a deep cleft on the extracellular side. The cleft opens with a pentagonal rim of ca. 14 Å internal diameter, made of the extracellular tips of TM3-TM7. The cavity is extended laterally with an aisle some 8 Å long and 7 Å wide, formed by TM2 adhering the TM3-TM7 plane. The cavity narrows into its interior, due to a tilt of TM3 toward the pentagon axis [2, 5]. This is just opposite to the respective feature in BRD, whose cavity widens toward the interior, due to expanding tilts of TM4 and TM5, see Figure 1. The aisled well ends up some 21 Å below the rim with a floor made of mainly hydrophobic residues (TM3: M123, TM4: L170, TM5: V213, F214 and TM6: W284, F287, F288), see Figure 1. Although TM1 remains outside and does not contribute to the cleft, there is more than enough space to fit a cyclic part of vasopressin within its cavity. The extracellular view of our V2R is similar to the view of V1aR in a recent model of the receptor/AVP complex [33], also based on the low resolution structure of bovine RD [1]. A major difference between the shape of the latter [33] and our V2R model consists of an apparently widened pentagonal rim in the V1aR model, which probably results from a 2000 step unconstrained minimization of the whole V1aR/AVP complex [33]. There is another not so apparent although possibly significant difference between both vasopressin receptor models. It results from differences in the specific TM1-TM7 sequence assignments. While our assignments exploit conclusions based upon the analysis of alignments of 204 GPCR sequences [4] (see Methods), the ones for V1aR use hydrophobicity profiles for biogenic amine receptors [34], combined with a subsequent sequence alignment of V1aR with the latter [33]. Major disagreement arising from the two approaches consists of extracellular extensions of the TM2, TM3 and TM7 helices by at least one helical turn (4–5 residues) in the V1aR model [33] which, as we argue in the Discussion (*vide infra*), may be not realistic.

Constrained simulated annealing (CSA)

The repetitive CSA in vacuo allows unrestrained motion for the loops and the docked ligand only. An analysis of the trajectories arising from the collection

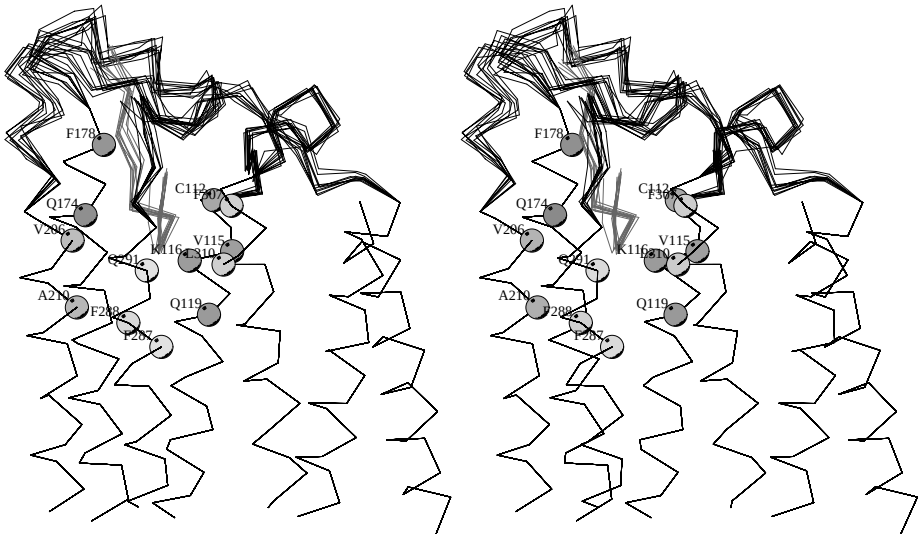

Figure 4. V2R/AVP:DockII in stereo, an overview. See legend to Figure 3.

of the final snapshots from some 40 CSA runs indicates that the CSA procedure as defined in Table 2 appears convergent with respect to the motion of the unrestrained parts of the V2R/ligand complexes, see Figure 2.

Clearly, as the motional convergence of the loops (both EL and IL) is achieved during a simulation in vacuo, the final orientations of the extracellular loops are of no biological relevance, however, the property of convergence implicates a potential power of CSA for future simulations, including water and lipid at the membrane-extracellular interface. In contrast, the final orientations of the ligand may appear biologically relevant since they are deeply buried within the V2R interior and thus sufficiently shielded from the environment.

Receptor/ligand interactions

The CSA results for the V2R/DAVP are given in Figure 2. A similar convergence although to different structures was observed for both V2R/AVP complexes defined as Dock I and Dock II (see Methods). The assemblies of 10 structures, each completing one of 10 final converging CSA runs, are given schematically as C^{α} carbon traces in Figures 3, 4 and 5, for V2R/AVP:Dock I, V2R/AVP:Dock II and V2R/DAVP (Dock I only), respectively.

The figures include only the extracellular V2R sides. The receptor residues involved in the strongest interactions with the ligands are represented with balls. A more detailed illustration of V2R/bioligand interactions is given in the interaction maps in Figures 6, 7 and 8, for V2R/AVP:Dock I, V2R/AVP:Dok II and V2R/DAVP, respectively.

A comparison of Figure 6 with Figure 7 reveals that DockII allows for an unhampered flexible insertion of the pressinoic ring into the cavity by approximately one helical turn shallower than DockI does. Hence, it is clear that the AVP docking in accordance with DockII does not exploit the potential hydrophobic interactions of the AVP Cys^{1}-Tyr^{2}-Phe^{3} *N*-terminal triad with the hydrophobic floor on the bottom of the cavity, provided by TM3: Met^{123}, TM4: Leu^{170}, TM5: Val^{213}, Phe^{214} and TM6: Trp^{284}, $Phe^{287}$and Phe^{288}, see Figure 1, while DockI does. As such, the alternative way of docking referred to as DockII was abandoned and not used for docking the V2R selective agonist DAVP.

A more detailed image of AVP embedded in the V2R cavity is shown in Figure 9. A simultaneous comparison of Figure 6 with Figure 8 warrants an inference that, despite somewhat different ligands' conformations and orientations, both agonists develop very similar sets of interactions with the receptor cavity. These are (see Figures 6 and 8 for the AVP and DAVP comparison and Figure 9 for details) TM3: C112, V115-K116, Q119, M123; TM4: Q174; TM5: V206, A210, V213; TM6: W284, F287, Q291; and TM7: F307, L310, A314, N317. They provide a wall of balanced polarity for the V2R extracellular well. In addition, DAVP intensively engages residues from receptor's TM2, in particular F91, L94, P95 and L97

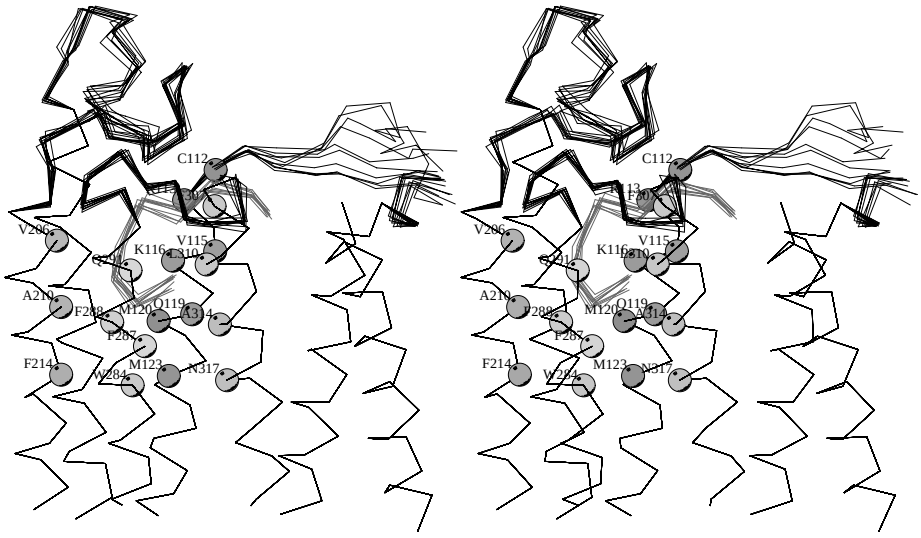

Figure 5. V2R/DAVP in stereo, an overview. See legend to Figure 3.

(compare Figures 6 and 8) due to an entirely different orientation of its *C*-terminal tail. Unfortunately, these are not the residues (Q92 and Q96) equivalent to those developing interactions with AVP in the V1aR model [33] (Q214 and 218 in the authors' notation). Interestingly, however, in TM3 through TM7 most ligand/receptor interactions in both our models find their counterparts in the V1aR/AVP model [33], see below.

Discussion

Any biological response to a drug requires at least (i) an affinity, measured by the concentration required for the half-maximal stimulation, and (ii) an intrinsic activity, measured by the maximal stimulation achievable [35]. Both features, although typically linked in a consecutive sequence, may be even totally independent of each other as in point-mutated GPCR analogs displaying high agonist affinity with no response or, on the other hand, a constitutional activity [36] with no affinity necessary. There are indications that the AVP analogs devoid of the *C*-terminal tail, i.e., pressinamide and desaminopressinamide, behave halfway: despite a few orders lower affinity to V2R than the mother AVP, they evoke a full biological response as measured by the adenylate cyclase activation [37, 38]. This result indicates that no *C*-terminal tail may be necessary for the intrinsic V2R activity. On the other hand, a role of the extracellular domain, and in particular EL1, in specific recognition of neurohypophyseal agonists is getting better and better documented [15, 39, 40]. There is also an indication of a role of the *N*-terminal domain and EL1 of the neurophyseal receptor in the recognition of the *C*-terminal tail of an agonist and a similar role of EL2 in the recognition of the cyclic part of the hormone. All this information has pointed out that an affinity (recognition) and intrinsic activity (response) may be separate events and may be accomplished in albeit consecutive yet independent steps.

Given this hypothesis and the fact that there is enough space for the pressinoic ring within the V2R cavity (see above), we assumed that the response may be elicited due to a complete immersion of this ring within the receptor. It is well documented that small GPCR agonists (biogenic amines, neurotransmitters) interact with the TM region while peptide agonists are usually believed to interact both with TM and extracellular regions of the receptor to trigger the response [16, 42], although several peptide hormones like endothelin [43], Substance P [44], angiotensin II [45], and thyrotropin-releasing hormone (TRH) [46] have been modeled entirely immersed within the TM bundle. Regarding TRH, a complete immersion has indeed been proved to be the case [47]. Moreover, while the TRH/receptor model was based on the bovine RD structure [1], all other agonists were fitted to the BRD-based [7] receptors, which are considerably less capacious than the former, see above.

Another argument for docking the pressinoic ring immersed within the V2R TM cavity is provided by the bundles of hydrophobic residues in the *N*-terminal

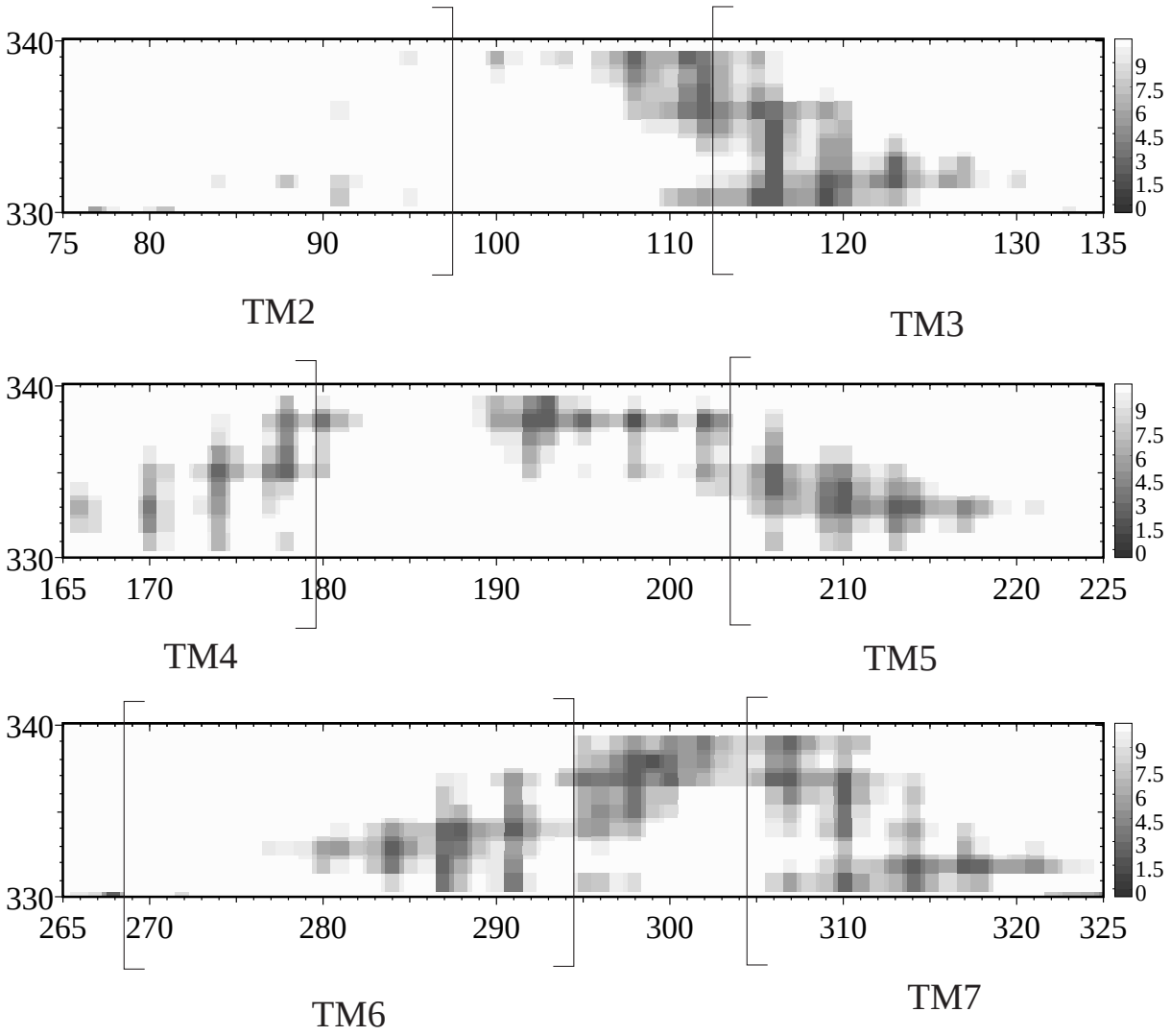

Figure 6. V2R/AVP:DockI. The map of receptor-ligand contacts, typical of the converged and relaxed structure. The contacts are represented by the closest distances between pairs of residues. The distances in Å are defined by the scale of variable shading on the far right. Horizontal axis: successive residues of V2R, with the sequences interrupted where the receptor-ligand contacts exceed 9 Å. Once the TM helices are marked, it is clear that EL1-EL3 are located in the central sections of each band from the top to the bottom, respectively. As argued in the text, we consider the involvement of the loops irrelevant at the present stage. Vertical axis: residue numbers 331–339 correspond to the sequence of either AVP or DAVP ligand.

part of AVP (the C1-Y2-F3 sequence) and on the floor of the V2R cavity, thus tempting to bring them together within an interaction. Last but not least, hypotheses on a common signal transduction mechanism, involving several invariant residues of the GPCR receptors in an allosteric signal transmission from the agonist to the G protein on the cytosolic side [48] argue in favor of an eventual fairly deep penetration of the receptor by the agonist. We believe that in view of all these facts our choice for docking AVP and DAVP inside the V2R cavity is thus warranted.

It is worth mentioning that our assumptions (see above) concerning the selection of possible targets for docking are coincident with those guiding the only to date docking attempt of AVP to another although homological (V1aR) receptor [33]. In the latter work, however, the authors used different TM sequence assignments and length than we did. Thus, their TM2, TM3 and TM7 helices extend on the extracellular side

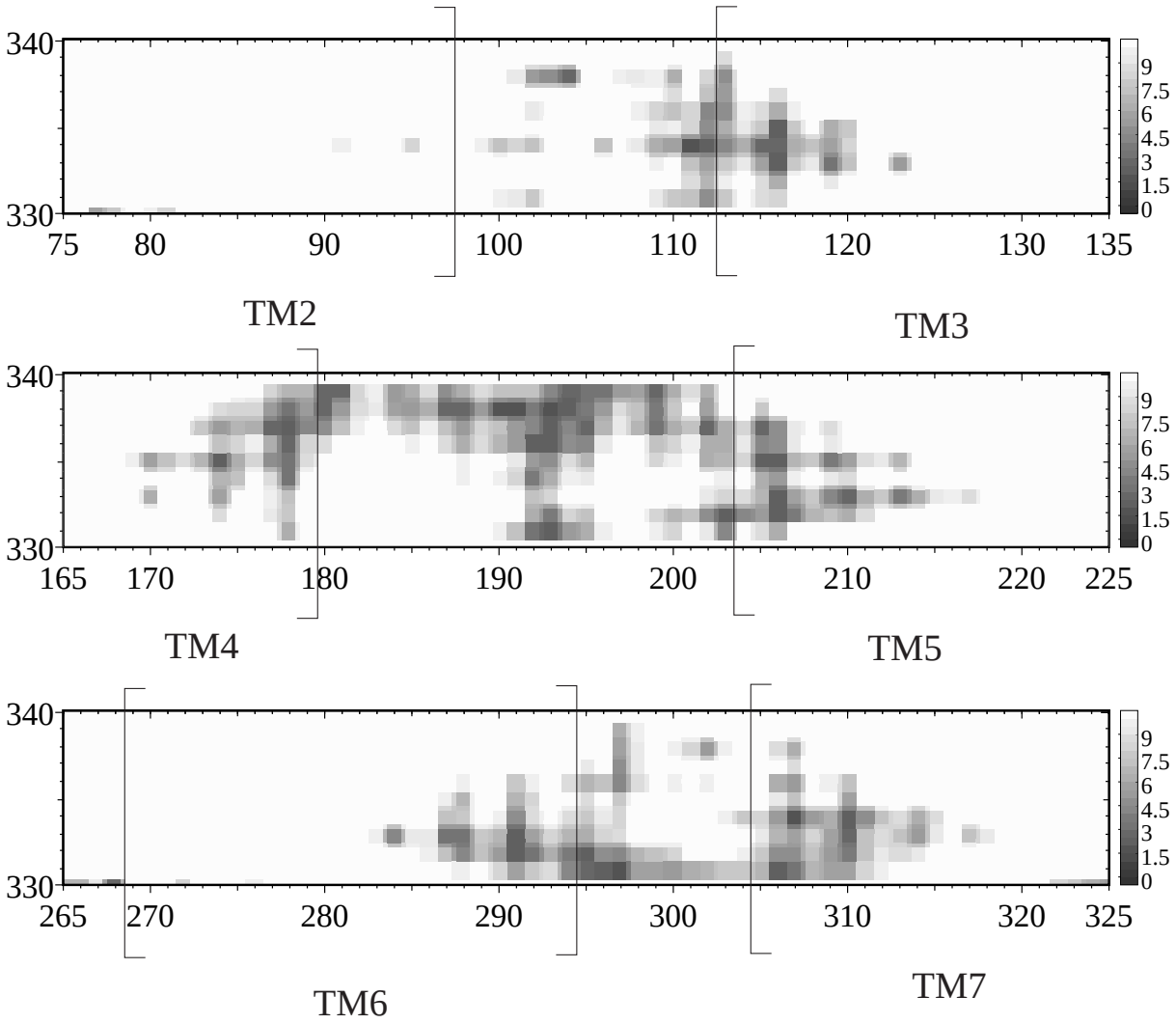

Figure 7. V2R/AVP:DockII. See legend to Figure 6. Notice that the ligand is shifted approximately by one helical turn towards the V2R exterior, relative to its position in DockI (compare Figure 6).

by at least one helical turn (4–5 residues) further than our TM2, TM3 and TM7 helices do. In view of the balanced and convincing multi-sequence assignments of Baldwin [4], the V1aR assignments [33, 34] would not suffice for the accomplishment of the EL1 and/or EL3 loops in some sequences where these loops would appear particularly short [5]. Moreover, TM2 in accordance with the V1aR model [33] imposes in a majority of GPCR sequences a Pro or occasionally even a double Pro some 5 residues away of the TM2 helix carboxyl end. While well in the interior of TM helices Pro residues are tolerable and some perhaps indispensable as helix hinges of possible role in allosteric message transmission from the first messenger to the cytosolic side of the receptor, a Pro residue(s) one turn away of the helix terminus is very unlikely in view of its strong helix-breaking property.

Interestingly, in spite of a somewhat different 7TM helix constitution and arrangement, major interacting residues in both V2R (this work) and V1aR [14, 33] appear the same or equivalent ones (given the receptor differences). This argues in favor of some tolerance of GPCR modeling results against specific choice of helix assignments. In particular, in spite of the differences mentioned above, the possible importance of TM3: K116, Q119 (K308 and Q311, respectively, in the notation used in V1aR [14,33]), TM4: Q174 (Q413 in V1aR) and TM6: Q291 (Q620 in V1aR) for AVP

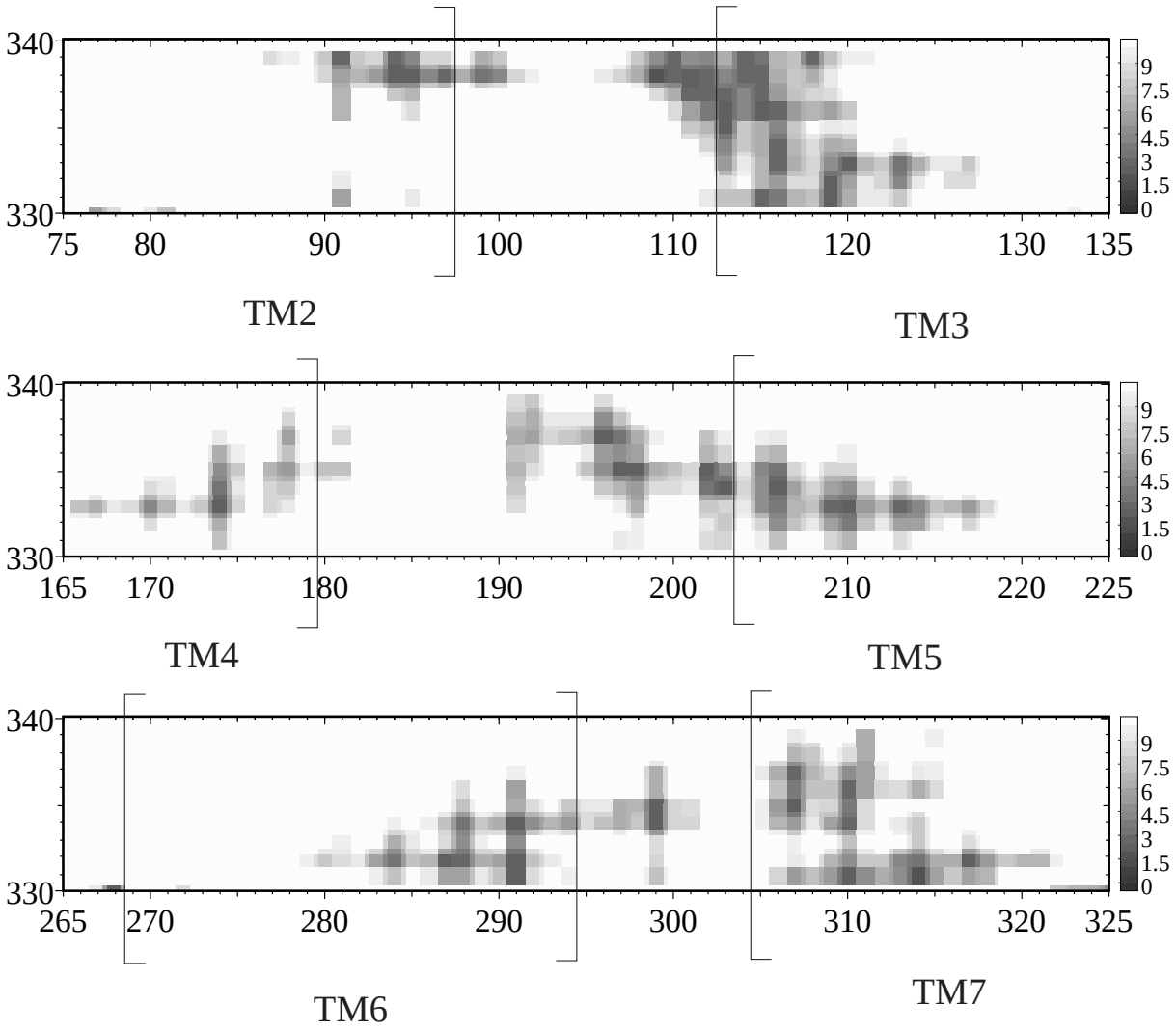

Figure 8. V2R/DAVP. See legend to Figure 6. Notice that the ligand, despite having a somewhat different conformation than that of AVP:DockI (see Figures 3 and 5), develops a similar set of interactions with the receptor (compare Figure 6).

agonist binding was reinforced. The role, sometimes dramatical (as for Q413 in V1aR), of these residues for binding has been experimentally confirmed [33].

Of other residues possibly involved in V2R agonist binding TM3: V115, M123; TM5: V206, A210, V213, F214; TM6: W284, F287, F288 and TM7: F307, L310, A314, N317 deserve in our opinion site directed mutagenesis studies. In particular, it would be interesting to see: (i) an effect of perturbing the floor (TM3:M123, TM5:V213,F214, TM6:W284, F287, F288, TM7:N317; (ii) an effect of tightening the cavity (TM7:A314).

Of features not related to the current modeling, it might be interesting to see if point mutations on TM1:N55, TM3:D136 (DRY sequence) and TM6:V270 could evoke constitutional activity as is the case for equivalent residues (N63, D142 and A293) in the $\alpha_{1\beta}$-adrenergic receptor [36]. Also, an explanation of the V1aR IL2 and V2R IL3 selective roles in transmitting the primary signals to $G_{q/11}$ and G_s proteins, respectively [13], remains a challenge.

Clearly, this work is naive in that it provides only a fragmentary picture of the interactions in question due to constraints imposed (see Methods) and neglects the lipid and the solvent. It appears even more fragmentary if one considers that ligand binding is only a step in a chain of events, which subsequently may involve an allosteric conformational change of a GPCR from

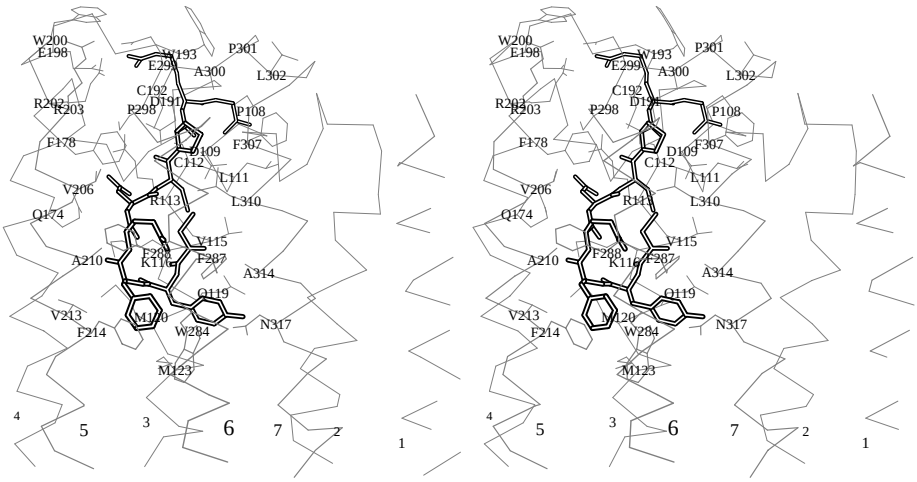

Figure 9. Stereoview of the converged and relaxed structure of AVP docked in the V2R cavity. Details of the interactions indicated in the overview (Figure 3) are seen explicitly.

its idle to its active form capable of interacting with the G protein. Thus, our current studies are aimed at the development of a model including the GPCR environment, i.e. the lateral lipid and the upright water, to produce a rectangular box for imposing periodic boundary conditions, which eventually would allow a simulation of an unconstrained GPCR model, possibly permitting for a realistic modeling of events on GPCR/extracellular and/or GPCR/intracellular interfaces.

Acknowledgements

We acknowledge the use of the GPCRDB data bases (http://swift.embl-heidelberg.de/7tm/), and of the Swiss-Model: an automated knowledge-based protein modeling server at Glaxo Wellcome Research and Development S.A. in Geneva (http://expasy.hcuge.ch/swissmod/SWISS_MODEL.html). The work has been supported by the Polish Scientific Research Committee (KBN), grant no. 3 T09A 027 11, by the Interdisciplinary Center for Modelling at the University of Warsaw (ICM UW), and by the Academic Computer Center in Gdańsk TASK. All referee's remarks are greatly appreciated.

References

1. Schertler, G.F.X., Villa, C. and Henderson, R., Nature, 362 (1993) 770.
2. Unger, V.M. and Schertler, G.F.X., Biophys. J., 68 (1995) 1776.
3. Schertler, G.F.X. and Hargrave, P.A., Proc. Natl. Acad. Sci. USA, 92 (1995) 11578.
4. Baldwin, J.M., EMBO J., 12 (1993) 1693.
5. Herzyk, P. and Hubbard, R.E., Biophys. J., 69 (1995) 2419.
6. Swiss-Model, an automated knowledge-based protein modeling server at Glaxo Wellcome Research and Development S.A. in Geneva (http://expasy.hcuge.ch/swissmod/SWISS-MODEL.html).
7. Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H., J. Mol. Biol., 213 (1990) 899.
8. Grigorieff, N., Ceska, T.A., Downing, K.H., Baldwin, J.M. and Henderson, R., Mol. Biol., 259 (1996) 393.
9. Strader, C.D., Fong, T.M., Tota, M.R., Underwood, D. and Dixon, R.A.F., Annu. Rev. Biochem., 63 (1994) 101.
10. Sugimoto, T., Saito, M., Mochizuki, S., Watanabe, Y., Hashimoto, S. and Kawashima, H., J. Biol. Chem., 269 (1994) 27088.
11. De Keyzer, Y., Auzan, C., Lenne, F., Beldjord, C., Thibonnier, M., Bertagna, X. and Clauser, E., FEBS Lett., 356 (1994) 215.
12. Laszlo, F.A., Laszlo, F., Jr. and De Wied, D., Pharmacol. Rev., 43 (1991) 73.
13. Liu, J. and Wess, J., J. Biol. Chem., 271 (1996) 8772.
14. Chini, B., Mouillac, B., Ala, Y., Balestre, M.-N., Trumpp-Kallmeyer, S., Hoflack, J., Elands, J., Hibert, M., Manning, M., Jard, S. and Barberis, C., EMBO J., 14 (1995) 2176.
15. Ufer, E., Postina, R., Gorbulev, V. and Fahrenholz, F., FEBS Lett., 362 (1995) 19.
16. Iismaa, T.P., Biden, T.J. and Shine, J., G Protein-Coupled Receptors, Springer-Verlag, Heidelberg, Germany, 1995, Chapter 1, pp. 16–22.
17. Baldwin, J.M., Curr. Opin. Cell. Biol., 6 (1994) 180.
18. Swiss-Prot accession code P30518 (http://expasy.hcuge.ch/sprot-top.html)
19. SYBYL, v. 6.1, Tripos Inc., St. Louis, MO, U.S.A., 1994.
20. Brookhaven Protein Data Bank accession code 1xy1.
21. Fine, R.M., Wang, H., Shenkin, P.S., Yarmush, D.L. and Levinthal, C., Proteins, 1 (1987) 342.
22. Shenkin, P.S., Yarmush, D.L., Fine, R.M., Wang, H. and Levinthal, C., Biopolymers, 26 (1987) 2053.

23. AMBER 4.1, Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., Ferguson, D.M., Seibel, G.L., Singh, U.C., Weiner, P.K. and Kollman, P.A., University of California, San Francisco, CA, U.S.A., 1995.
24. RasMol, v. 2.6, Molecular Visualisation Program, Sayle, R., Glaxo Wellcome Research and Development, Stevenage, Hertfordshire, U.K., 1996.
25. Kraulis, P., J. Appl. Crystallogr., 24 (1991) 946.
26. Vriend, G., GPCRDB © 1996 (http://swift.embl-heidelberg.de/7tm/).
27. Wood, S.P., Tickle, I.J., Treharne, A.M., Pitts, J.E., Mascarenhas, Y., Li, J.Y., Husain, J., Cooper, S., Blundell, T.L., Hruby, V.J., Buku, A., Fischman, J.F. and Wyssbrod, H.R., Science, 232 (1986) 633.
28. Tarnowska, M., Liwo, A., Kasprzykowski, F., Łankiewicz, L., Grzonka, Z. and Ciarkowski, J., Curr. Top. Med. Chem., 1 (1993) 145.
29. Hruby, V.J., Chow, M.-S. and Smith D.D., Annu. Rev. Pharmacol. Toxicol., 30 (1990) 501.
30. Kojro, E., Eich, P., Gimpl, G. and Fahrenholz, F., Biochemistry, 32 (1993) 13537.
31. Brtník, F., In Jošt, K., Lebl, M. and Brtník, F. (Eds.), CRC Handbook of Neurophyseal Hormone Analogs, Vol. II, part 1, CRC Press, Inc., Boca Raton, FL, USA, 1987, pp. 126–154.
32. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.
33. Mouillac, B., Chini, B., Balestre, M.-N., Elands, J., Trumpp-Kallmeyer, S., Hoflack, J., Hibert, M., Jard, S. and Barberis, C., J. Biol. Chem., 270 (1995) 25771.
34. Trumpp-Kallmeyer, S., Hoflack, J., Bruinvels, A. and Hibert, M., J. Med. Chem., 35 (1992) 3448.
35. Brtník, F., In Jošt, K., Lebl, M. and Brtník, F. (Eds.), CRC Handbook of Neurohypophyseal Hormone Analogs, Vol II, part 1, CRC Press, Inc., Boca Raton, FL, USA, 1987, pp. 131–134.
36. Scheer, A., Fanelli, F., Costa, T., De Benedetti, P. and Cotecchia, S., EMBO J., 15 (1996) 3566.
37. Hechter, O., Terada, S., Nakahara, T. and Flouret, G., J. Biol. Chem., 253 (1978) 3219.
38. Hechter, O., Terada, S., Spitsberg, V., Nakahara, T., Nakagawaga, S.H. and Flouret, G., J. Biol. Chem., 253 (1978) 3230.
39. Chini, B., Mouillac, B., Ala, Y., Balestre, M.-N., Trumpp-Kallmeyer, S., Hoflack, J., Elands, J., Hibert, M., Manning, M., Jard, S. and Barberis, C., EMBO J., 14 (1995) 2176.
40. Howl, J. and Wheatley, M. In Kaumaya, P.T.P. and Hodges,R.S. (Eds.), Peptides: Chemistry, Structure and Biology, Mayflower Scientific, Kingswinford, U.K., 1996, pp. 400–402.
41. Postina, R., Kojro, E. and Fahrenholz, F., J. Biol. Chem., 271 (1996) 31593.
42. Birnbaumer, M., J. Receptor Sign. Transduct. Res., 15 (1995) 131.
43. Krystek, S.R., Hunt, J.T., Jr, Stein, P.D. and Stouch, T.R., J. Med. Chem., 38 (1995) 659.
44. Trumpp-Kallmeyer, S., Hoflack, J. and Hibert, M., In Buck, S.H; (Ed.), The Tachykinin Receptor, Humana Press, Totowa, NJ, USA, 1994, pp. 237–255.
45. Underwood, D.J., Strader, C.D., Rivero, R., Patchett, A.A., Greenlee, W. and Prendergast, K., Chem. Biol., 1 (1994) 211.
46. Laakkonen, L.J., Guarnieri, F., Pearlman, J.H., Gershengorn, M.C. and Osman, R., Biochemistry, 35 (1996) 7651.
47. Pearlman, J.H., Laakkonen, L.J., Guarnieri, F., Osman, R. and Gershengorn, M.C., Biochemistry, 35 (1996) 7643.
48. Scheer, A. and Cotecchia, S., J. Receptor Sign. Transduct. Res., 17 (1997) 57.