JCTC Journal of Chemical Theory and Computation

Milestones in the Activation of a G Protein-Coupled Receptor. Insights from Molecular-Dynamics Simulations into the Human Cholecystokinin Receptor-1

Christophe Chipot*

Equipe de dynamique des assemblages membranaires, UMR No 7565, Nancy Université BP 239, 54506 Vandœuvre-lès-Nancy cedex, France

Received August 1, 2008

Abstract: Activation of G protein-coupled receptors (GPCRs) obeys an allosteric mechanism triggered by ligand binding. To understand how the signal is transduced in the cell, identification of the milestones paving the pathway between the active and the inactive states of the receptor is necessary. A model of the human cholecystokinin receptor-1 (CCK1R) has been proposed recently. The complex formed by CCK1R and an agonist ligand will serve as a paradigm of an active conformation to capture milestones in GPCR activation. To reach this goal, assuming microreversibility, the initial step toward the inactivation of CCK1R was modeled using free energy calculations, whereby the ligand is removed from the binding pocket. However accurate the reproduction of the experimental affinity constant, this simulation only represents an embryonic stage of the inactivation process. Starting from the *apo* receptor, an unprecedented 0.1-μs molecular dynamics trajectory was generated, bereft of experimental biases, bringing into the light key events in the inactivation of CCK1R, chief among which the hydration of its internal cavity, concomitant with the spatial rearrangement of the transmembrane helical segments. Hydration is intimately related to the isomerization of the highly conserved residue W326 of helix VI, acting as a two-state toggle switch, and of residue M121 of helix III. In the active state, the former residue obstructs the crevice, thereby preventing water leakage, which would otherwise trigger the disruption of an ionic lock between helices II and III involving the signature E/DRY motif ubiquitous to GPCRs.

Introduction

G protein-coupled receptors (GPCRs) constitute the largest family of membrane proteins responsible for signal transduction across the biological membrane, mediating the cellular response to a host of environmental stimuli.[1,2] A variety of pathologies are rooted in the malfunction of these proteins, which, not too surprisingly, have rapidly become privileged targets for drug discovery.[3] Activation of GPCRs follows an allosteric mechanism actuated by ligand binding and resulting in a conformational modification of its seven-helix transmembrane (TM) domain.[4] A complete understanding of how the signal is transduced across the cell membrane requires at its core the identification of the milestones that pave the pathway connecting the active state of the receptor to its inactive counterpart. A convenient framework for apprehending the activation of GPCRs is provided by the extended ternary complex model, which assumes that active and inactive conformations of the receptor coexist in an equilibrium.[5] The complexity of the activation process is further magnified by the possibility of the receptor to adopt a variety of active conformations,[6,7] to the extent that different agonist ligands can bind distinct active forms of the same GPCR, triggering in turn different signaling pathways.[8]

The dearth of structural information readily available for GPCRs has imparted a new momentum to the computational investigation of these biological systems.[9] Access to the

* Corresponding author phone: +33.3.83.68.40.97; e-mail: Christophe.Chipot@edam.uhp-nancy.fr.

10.1021/ct800313k CCC: $40.75
Published on Web 10/31/2008

three-dimensional structure of bovine rhodopsin,[10,11] the paradigm of family-A GPCRs, has opened new vistas for the modeling of structurally related membrane proteins. Within the past ten years, considerable effort has been invested to predict the structure and the function of GPCRs, based on an arsenal of modeling tools ranging from first principles[12] to knowledge-based approaches.[13,14] Models built following the latter route can be used profitably for the design of novel site-directed mutagenesis experiments, which, in turn, can serve to refine the model of the receptor and decipher the mechanisms whereby it accomplishes its cellular function. Combining synergistically theory and experiment, a model of the human cholecystokinin receptor-1 (CCK1R)—a GPCR pertaining to family A, was put forth recently[15] and probed by means of large-scale molecular dynamics (MD) simulations and free energy calculations targeted at reproducing the binding affinities from whence the model had been constructed.[16] Cholecystokinin[17] (CCK) is a hormone ubiquitous to the gastrointestinal system, where it mediates digestion, while in the central nervous systems, it acts as a neurotransmitter, which, among others, stimulates satiety. The proposed model[15] consists of a complex, henceforth referred to as CCK1R*:CCK9, and formed by the receptor, presumably in an active state, and the agonist ligand CCK9, a nonapeptide analogue of the endogenous ligand. It is the fruit of several years of a relentless exchange between theory and experiment, confronting hypotheses and inferences to obtain a self-consistent picture of the membrane receptor (see refs 18−21).

The number of theoretical studies aimed at characterizing the activation of GPCRs has remained hitherto limited, primarily due to the paucity of experimentally validated three-dimensional structures, and, to a somewhat lesser extent, the formidable computational effort involved in their simulation. A common feature shared by these investigations is the use of experimental biases, viz. usually in the form of harmonic restraints, to convert the inactive conformation into an active one and *vice versa*.[22−25] In the present work, a different strategy is followed to dissect the milestones of GPCR activation, whereby, assuming microreversibility, the inactivation of the receptor will be modeled, bereft of external biases, starting from the structure of the CCK1R*:CCK9 complex. Inactivation of the receptor is an intricate process, anticipated to span time scales that are not routinely amenable to classical, all-atom MD simulations. This process subsumes the dissociation of the agonist ligand from the native binding pocket, assumed here to be a rapid event, and the ensuing internal relaxation of the receptor, which constitutes the rate-limiting step of the overall transformation. The first step, consisting of removing the agonist ligand CCK9 from the binding pocket of the receptor, will be examined employing free energy perturbation (FEP) calculations. As will be seen in the following sections, provided that appropriate sampling is performed, accurate reproduction of experimental binding affinities can be achieved, even for very large, supramolecular assemblies. The second step, during which the constitutively active *apo* receptor, CCK1R*, relaxes toward the inactive conformation, CCK1R° will be tackled in a 0.1-μs MD simulation.

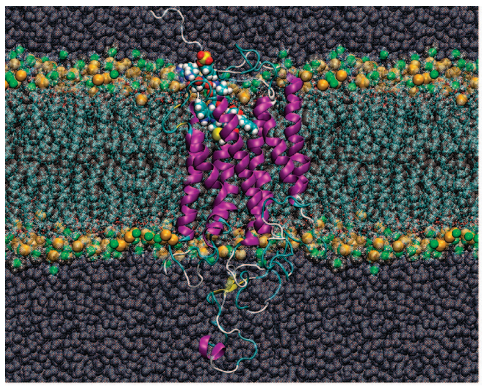
Computational investigations of GPCR activation based on three-dimensional models remain admittedly scarce[24] and are burdened by the pervasive dogmatic assumption that, in general, no relevant qualitative, let alone quantitative information can be inferred from calculations that rest on an unknown parameter: The structure. To a large extent, this assertion explains why efforts to disentangle the intricate activation mechanism of GPCRs[22,25−27] have remained focused on structures solved experimentally, namely the bovine rhodopsin[10,11] and the human β_2-adrenergic[28] receptors. Does it mean that attempts to gain new insights from numerical simulations of GPCR models are necessarily vain as long as the structure of the membrane protein is not solved? Whereas simulations over appreciably long time scales evidently cannot validate a three-dimensional receptor model, they may, nonetheless, be used fruitfully to probe hypotheses drawn from experiments.[16] In the particular instance of CCK1R, accepting the view that refined homology- or template-based constructs are inevitably wrong would be tantamount to ignoring the wealth of data accrued in the past ten years,[15,18−21,29−34] and during which the structural model of the CCK1R*:CCK9 complex has been built, progressively improved, and thoroughly tested. On the other hand, it is fair to recognize that the aforementioned tenet can be fueled and, thus, reinforced by controversies relying upon conflicting models—e.g. the hypothesis of a reverted agonist ligand interacting at its C-terminus end with residue W39[35] of the receptor, challenged by NMR studies of CCK bound to a fragment of CCK1R.[36]

The primary thrust of this work is not an unrealistic claim that an inactive conformation of CCK1R has been isolated, but rather the demonstration that unrestrained, all-atom MD can be used beneficially to delineate how the structural features of the active receptor evolve as its conformation slowly interconverts into an inactive one. Shedding new light on the molecular mechanisms that underlie activation, large-scale MD simulations complement experimental advances by rationalizing the lower affinity of the ligand for the receptor upon conformational transition. As will be seen in the present contribution, central to this structural modification is the role played by the hydration of the internal cavity triggering a variety of events, which can evidently only be captured in a biologically realistic environment.

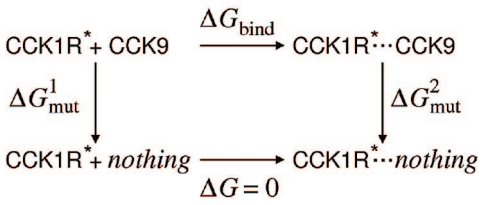
Methods

Three-Dimensional Model and Molecular Assay. The initial three-dimensional molecular assembly consisting of the complex formed by CCK1R* and agonist ligand CCK9,[15] inserted in a fully hydrated POPC bilayer, which altogether represents a total of 72,255 atoms, was equilibrated in a 31 ns MD simulation.[16] Over that period, it was observed that the distance rmsd computed over the TM α-helices never exceeded *ca.* 2.5 Å, revealing no alteration of their secondary structure.[16] The essential of the key protein-ligand interactions brought to light experimentally were preserved throughout the trajectory (see refs 18−21).

Free Energy Calculation. The free energy change associated with the binding of CCK9 to CCK1R was estimated



A



B

Figure 1. **A**. CCK1R*:CCK9 assembly in a fully hydrated POPC bilayer after 31 ns.[16] TM α-helices are shown as purple ribbons. The agonist ligand CCK9 is depicted as a space-filling molecular model. A semitransparent representation is used for water and lipid units, the phosphate and choline groups of which are featured as orange and green van der Waals spheres, respectively. The image rendering was done with Vmd.[65] **B**. Thermodynamic cycle used for the double annihilation of agonist ligand CCK9 in the free, hydrated state (left leg) and in the receptor (right leg). The binding free energy, ΔG_{bind}, which is determined experimentally, corresponds to the difference $\Delta G^2_{mut} - \Delta G^1_{mut}$.

by means of a double annihilation[37] of the ligand, as described in Figure 1. Transformations in bulk water and in the receptor were performed employing the free energy perturbation (FEP) method,[38] wherein the Gibbs free energy difference between two thermodynamic states connected by $M-2$ intermediate, nonphysical states is expressed as

$$\Delta G = -\frac{1}{\beta}\sum_{i=1}^{M-1} \ln\langle exp\{-\beta[U(x;\lambda_{i+1}) - U(x;\lambda_i)]\}\rangle_{\lambda_i} \quad (1)$$

where $\beta = 1/k_BT$, k_B is the Boltzmann constant, T is the temperature, and $U(x;\lambda_i)$ is the potential energy function that depends upon the Cartesian coordinates of the system $\{x\}$ and the coupling parameter, λ_i, which connects the initial and the final states of the annihilation. $\langle \cdots \rangle_{\lambda_i}$ denotes an ensemble average over configurations representative of intermediate state i. Because in the course of the alchemical transformations charged amino acids vanish, the reaction path was stratified into 110 stages of uneven widths.[37] Narrow intermediate states were defined toward the end point of the simulation to avoid singularities. Each transformation, whether in the aqueous medium or in the receptor, involved 10 ps of equilibration followed by 100 ps of data collection, corresponding to a total of 12.1 ns. The associated error was estimated from a first-order expansion of the free energy (see the Supporting Information).

MD Simulation. All simulations were performed using the Namd simulation package[39] in the isobaric-isothermal ensemble. The pressure and the temperature were fixed at 1 bar and 310 K, respectively, employing the Langevin piston algorithm[40] and softly damped Langevin dynamics. The molecular assays were replicated in the three directions of Cartesian space by means of periodic boundary conditions. The particle-mesh Ewald method[41] was employed to compute electrostatic interactions. The *r*-RESPA multiple time-step integrator[42] was used with a time step of 2 and 4 fs for for short- and long-range forces, respectively. Covalent bonds involving a hydrogen atom were constrained to their equilibrium length. The CCK1R*:CCK9 complex and its environment were described by the all-atom Charmm27 force field.[43,44] The main MD simulation, wherein the initial conformation of the *apo* receptor corresponds to the end point of the FEP calculation during which the agonist ligand was annihilated in the CCK1R*:CCK9 complex, was conducted over a period of 0.1 μs. In addition, a shorter, 50-ns control simulation was performed, starting with the equilibrated structure of the complex,[16] from whence the nonapeptide, CCK9, was removed abruptly. Running on an array of twelve 2.40-GHz Intel Xeon processors communicating via a gigabit network, the wall clock time was equal to 37.6 h per nanosecond.

Results and Discussion

Binding of the Agonist Ligand to the Receptor. Direct measurement of the free energy difference associated with the inactivation of the receptor goes far beyond the current scope of classical, all-atom MD simulations. Under the assumption that the conformation of the membrane protein is preserved upon dissociation of the ligand—i.e. the structural features of the constitutively active *apo* receptor, CCK1R*, and the ligand-bound receptor, CCK1R*:CCK9, remain unperturbed—this early event in the inactivation process can be tackled using free energy methods. Microreversibility further suggests that the binding free energy of the agonist ligand to the receptor can be estimated by simulating the dissociative process, whereby CCK9 is annihilated in the CCK1R*:CCK9 complex depicted in Figure 1 and in a bulk aqueous environment, following the thermodynamic cycle[37] in the same figure.

This so-called alchemical transformation is considerably more challenging than the simple *in silico* point mutations of selected amino-acid residues reported in ref 16, due to the larger perturbation involved.[37] Convergence of the ensemble average featured in eq 1 imposes that all accessible configurations of the solvent—whether the membrane protein or the aqueous medium—around the vanishing ligand be sampled thoroughly. As can be seen in Figure 2, appropriate overlap of the thermodynamic ensembles embodied in their density of states indicates that this requirement is likely to be fulfilled[45] (see the Supporting Information). Noteworthily, the binding free energy, ΔG_{bind}, consists of a difference of two large terms, the magnitude of which is dictated primarily by the ionic solvation of the participating charged residues. Based on two 12.1-ns FEP calculations, viz. in CCK1R* and in water, and employing the acceptance ratio method of Bennett that combines forward and reverse transformations,[46,45] ΔG_{bind} is found to be equal to -11.5 ± 1.2 kcal/mol. This

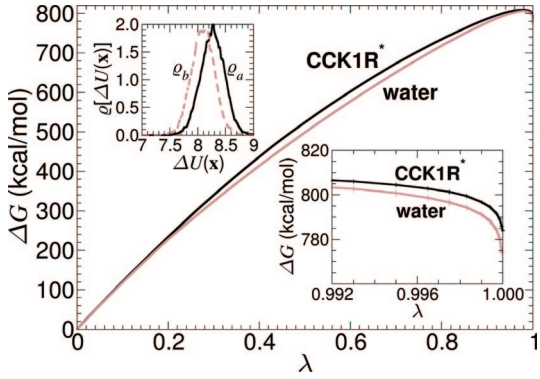


Figure 2. Evolution of the Gibbs free energy as a function of the general extent parameter, λ, utilized to annihilate CCK9. Upper inset: Overlapping configurational ensembles embodied in the density of states, $\varrho[\Delta U(x)]$, of contiguous intermediate states (here, at $\lambda = 0.5$)—where $\Delta U(x)$ denotes the variation of the potential energy between state *a* and state *b*, which is a function of Cartesian coordinates $\{x\}$ (see the Supporting Information). Lower inset: Closeup on the ultimate intermediate states of the transformation. ΔG_{bind} corresponds to the difference between the dark curve (bound state) and the light curve (free state).

estimate compares reasonably well with the experimental value of -12.6 ± 1.4 kcal/mol, which relies upon repeated radio-activity measurements[16,19] of transfected cells expressing the receptor after exposure to the radio-iodinated agonist ligand.[47]

The binding free energy of -11.5 ± 1.2 kcal/mol can be interpreted as a sum of contributions of different natures. Association of CCK9 to CCK1R proceeds through strong electrostatic interactions involving (i) down in the crevice, aspartate D8 of the agonist ligand and residue R336 of the receptor,[19] and (ii) above, in the upper region of the binding pocket, sulfated tyrosine sY3 of CCK9 and amino acids M195 and R197 of CCK1R. As was shown earlier, point mutations of residues D8 and sY3 into alanine yield a loss in the binding free energy equal to 3.2 ± 0.3 and 2.7 ± 0.1 kcal/mol, respectively[16,19]—i.e. roughly speaking, one-half of the total binding free energy. The affinity of CCK9 toward CCK1R can be further explained in terms of robust van der Waals contacts[21] formed between residues M7 and F9 of the agonist ligand and a network of hydrophobic residues of the receptor, which includes L53, V125, I329, and I352. Additional interactions of lesser strength involve the N-terminus of nonapeptide CCK9 and residues W39 and Q40 of CCK1R[48] and appear to be complemented by ancillary, water-mediated interactions that embrace amino acids T4, G5, W6, and F9 of the agonist ligand and residues F107, T117, T118, and A341 of the receptor.[16]

The present free energy calculation raises, however, several key issues, chief among which is the level of confidence that can be assigned to the estimate of the binding affinity. From a dogmatic point of view, such a computation represents a numerically challenging endeavor, based on a seemingly fragile construct, and, hence likely to yield a series of numbers that have unaccountable reliability and validity. Even for smaller ligands binding a structure determined experimentally,[49,50] accurate predictions of free energy differences by means of FEP calculations, which are notoriously plagued by convergence issues,[37] has proven to remain a difficult undertaking. Yet, although the present simulations rely upon a model of the receptor that cannot be indisputably validated until the three-dimensional structure of the latter is solved, it ought to be reminded that a host of experimental data[15,18−21,29−34] lends support to the putative placement of the ligand in the binding pocket. Moreover, precise reproduction of relative binding affinities for the CCK1R*:CCK9 complex[16] not only strengthens this contention but also suggests that the proposed methodology and sampling strategy are sufficiently robust to measure the free energy characterizing the association of the agonist ligand to the receptor with appreciable confidence.

Structural Features of the *apo* Receptor over the 0.1-μs Time Scale. It is interesting to observe that as the agonist ligand vanishes in the binding pocket, water molecules from the bulk environment rapidly seep in the receptor. Figure 3 reveals that after 0.1 μs of MD sampling, the *apo* structure is essentially flooded, which, as will be seen shortly, necessarily bears some implications on the conformational equilibrium of the membrane protein. The density profiles characterizing the location of the different components of the molecular assembly are shown in Figure 3. The nonzero probability to find water molecules across the protein is suggestive of a possible communication between the periplasmic and the cytoplasmic sides of the model membrane.

The density profiles also shed light on the organization of the receptor in its membrane environment. The width of the density characterizing the transmembrane (TM) domain of CCK1R roughly coincides with that of the acyl chains of the lipid bilayer. The marked mobility of the extracellular loops is mirrored in the broad distribution of their position along the normal to water-membrane interface.[16] This is especially true for the cytoplasmic loops, which, in the absence of the G protein subunits, drift without much hindrance. The distance root mean-square deviation (rmsd) of the receptor highlighted in Figure 3 increases by approximately 1 Å within the first 50 ns of the simulation and stays roughly constant at *ca.* 2 Å over the remaining 50 ns. The same figure indicates that the secondary structure of the receptor remains unaltered on the 0.1-μs time scale, thereby suggesting that the moderate deviation monitored with respect to the initial conformation, viz. CCK1R*, should be ascribed preferentially to internal motions within the TM scaffold.

Spatial Rearrangement of the Receptor toward Its Inactive Conformation. A distance rmsd represents a global, average measure of the deformations undergone by the membrane protein. It, therefore, does not indicate where the structural changes are localized. As underlined previously, the time scale spanned by classical, all-atom MD is *a priori* too short to capture the complete inactivation process. On the other hand, the only known structure of an activated GPCR, the bovine rhodopsin, reveals that the modifications with respect to its dark, inactivated form[10] are much smaller than hitherto expected.[11] Intricate local movements within the TM α-helix bundle can be disentangled in terms of

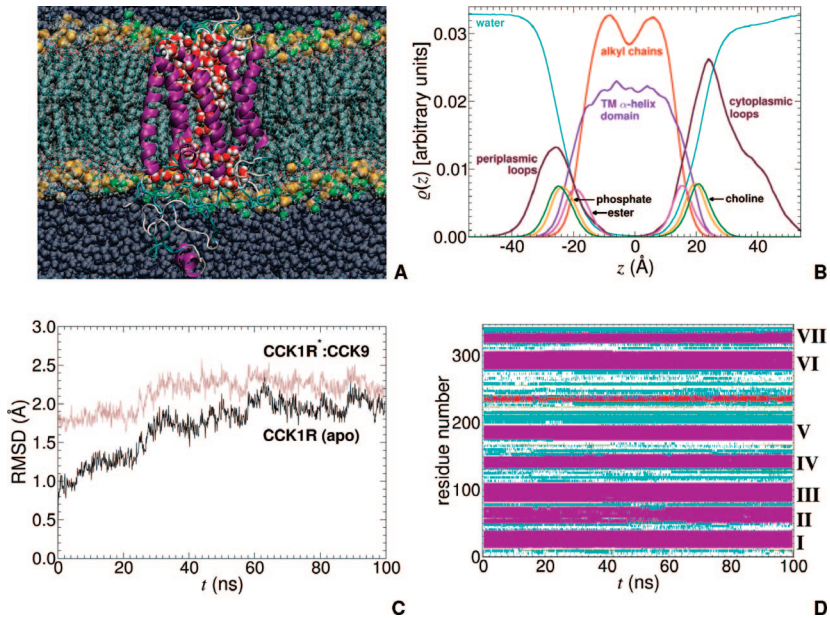


Figure 3. **A**. *apo* CCK1R in a fully hydrated POPC bilayer after 100 ns. Color coding is identical to that of Figure 1. Hydration water molecules flooding into the receptor are highlighted as van der Waals spheres (oxygen atoms in red and hydrogen atoms in white). **B**. Number density profiles of the *apo* CCK1R assembly averaged over the entire trajectory. **C**. Time-evolution of the distance rmsd over backbone atoms of *apo* CCK1R with respect to the initial structure at $t = 0$ (dark curve) and the CCK1R*:CCK9 complex (light curve). **D**. Time-evolution of the secondary structure of *apo* CCK1R. α-, 3_{10}-, and π-helices are shown in purple, pink, and red, respectively. Coils and turns are depicted as white and cyan tubes.

translational and rotational motions of the individual α-helical segments. A quantification of these movements is provided in Figure 4, distinguishing between interhelical separations, Δr, and rotation, Ω, of individual TM α-helices about their longitudinal axis as a function of time.

Using the CCK1R*:CCK9 complex as a reference, the geometrical analysis of the 0.1-μs trajectory illuminates a sequence of events in the reorganization of the receptor deprived of its agonist ligand. The most striking feature of the latter is the lateral translation of TM III, TM IV, TM V, and TM VI, while TM I, TM II, and TM VII remain virtually fixed in Cartesian space. In a seemingly concerted fashion, within the first 30 ns of the simulation, the interhelical separation of TM IV and TM V and of TM IV and TM VI increases, while TM V and TM VI move closer to each other. During the same period, TM II and TM III are initially pulled apart, prior to returning to their original separation, and eventually diminishing the latter. The distance between the centers of mass of TM I and TM VII is virtually unchanged over the entire simulation, in line with the site-directed spin labeling experiments on bovine rhodopsin.[51] Past the first 50 ns of the MD trajectory, only marginal fluctuations of Δr are measured. It is noteworthy that on the 0.1-μs time scale, variations of the different interhelical separations remain moderate, but in the most glaring examples of the TM II and TM V and the TM IV and TM VI pairs of α-helices depicted in Figure 4, they can be as large as *ca.* 2 Å. The continuous translation of TM III toward TM VI by *ca.* 1.5 Å is of somewhat lesser amplitude yet reflects previous measurements on the dark state of rhodopsin.[52] There is evidently a structural rationale for these broader displacements, as will be seen in the following sections.

Rotation of the α-helical segments is in general confined within a ±25° range. In reference with the CCK1R*:CCK9 complex, it is apparent that preliminary rotation of the individual TM α-helices occurred in the course of the alchemical transformation, whereby the agonist ligand was annihilated in the binding pocket. In the second half of the simulation, TM II rotates clockwise, albeit intermittently, congruent with the approach of TM III and TM V. No perceptible rotation is monitored for TM III, in line with previous observations for another cholecystokinin receptor, CCK2R, in light of biased MD simulations.[24] In contrast, TM IV rotates rapidly, yet in a transient fashion, within the first 40 ns, before returning to its original state. Rotation of TM IV upon activation has been previously described, for instance, in the case of the β_2 adrenergic receptor.[53,54] Reminiscent of bovine rhodopsin,[55] variations of Ω is significant for TM VI, undergoing an anticlockwise writhing motion as it moves closer to TM V and away from TM IV. A symmetrical behavior has been witnessed in the activation process of CCK2R.[24]

Hydration of the Binding Pocket Is an Early Event in the Inactivation of the Receptor. Figure 3 sheds light on the internal hydration of the receptor, conducive to a possible communication between the aqueous phases on the cytoplasmic and the periplasmic sides of the lipid bilayer. How can this communication be established? Initial simulations of the CCK1R*:CCK9 complex confirmed that the conduit formed by the binding pocket is too narrow to allow water molecules to crawl through when the agonist ligand is locked in.[16] They also revealed higher in the binding pocket the existence of a cluster of water molecules with residence times spanning the nanosecond time scale and communicating with the extracellular environment. It be-

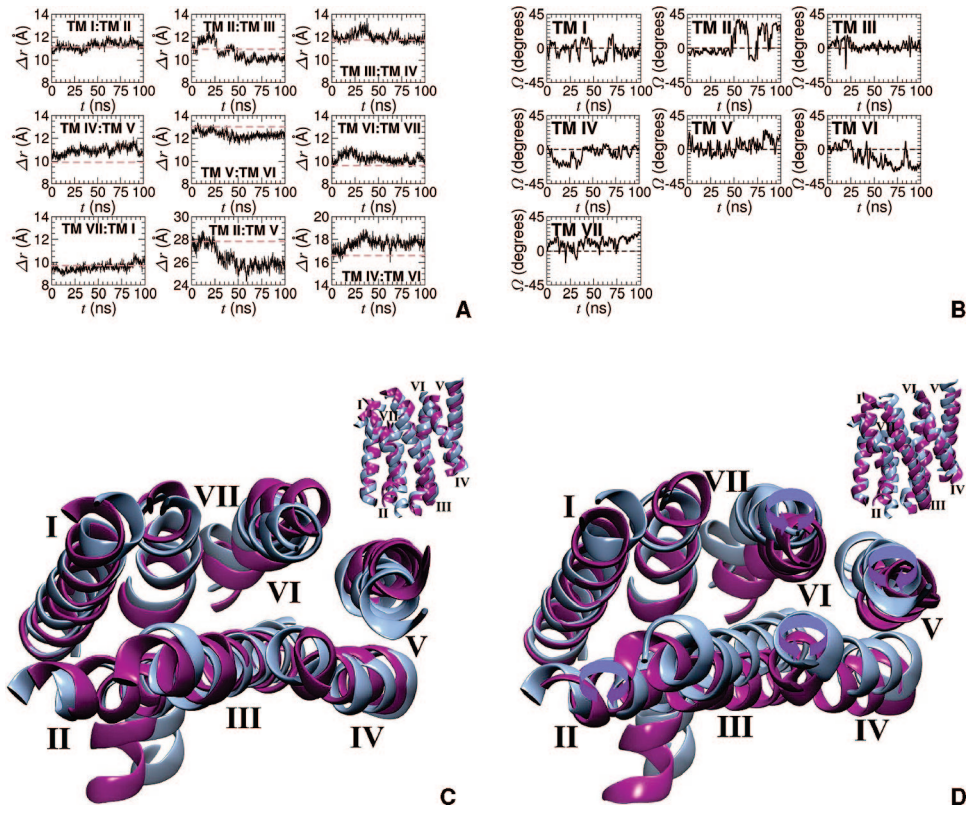


Figure 4. **A**. Time-evolution of the distance separating the center of mass of the α-helices forming the TM domain of *apo* CCKR1. Comparison with the corresponding separations in the CCK1R*:CCK9 complex (dashed line). **B**. Rotation of the TM helices of *apo* CCKR1 as a function of time. **C**. Top and side views of the TM region of *apo* CCK1R at $t = 0$ (purple). Comparison with the CCK1R*:CCK9 complex (pale blue). **D**. Top and side views of the TM region of *apo* CCK1R at $t = 100$ ns (purple).

comes obvious that in the absence of CCK9 blocking the access to the crevice, these water molecules can diffuse freely and accumulate at the bottom of the pocket.

As depicted in Figure 5, a well-conserved tryptophan residue, W326, is located underneath the binding pocket of CCK1R. Its position, within the simulation time, obeys a two-state regime driven by torsional angle χ_2, which is either equal to *ca.* 0° in the constitutively active receptor or to *ca.* 90° otherwise. Remarkably enough, an abrupt, first-order transition of this dihedral angle is observed around 27 ns, coinciding with the maximum separation of both the TM III and TM IV and the TM IV and TM VI α-helix pairs. Reorientation of the indole ring is accompanied by a concerted rearrangement of TM V and TM VI, leading to double π-π stacking, W326 being intercalated between F187 and F330. On the basis of a model of phenylalanine, a π-π interaction consisting of two optimally stacked aromatic rings has been shown to amount to *ca.* −3.4 kcal/mol.[56] This π-π interaction is expected to be reinforced in the case of tryptophan, as a result of an increased dispersion contribution. In addition, cooperative, Axilrod-Teller-like effects are envisioned to tighten even further the interaction of the two α-helical segments, thereby contributing to the overall stability of the TM scaffold.

This toggle-switch[57] isomerization in the side chain of W326 corresponds to virtually the maximum hydration of the receptor, whereby an excess of 300 water molecules occupy the interior of the α-helix bundle. As a basis of comparison, using the same criteria, the number of water molecules in the CCK1R*:CCK9 complex was found to be *ca.* three times less (see Figure 5). That the internal hydration of GPCRs is modified substantially in the course of their activation appears to be a ubiquitarian property, as was recently underlined by Grossfield et al. in the light of extended simulations of bovine rhodopsin[58]—albeit in the latter, the marked flow of water molecules follows the isomerization of the Schiff base, which is distinct in essence from the actual removal of a noncovalently bound ligand from its designated binding pocket. The present result suggests that flooding of the receptor occurred in the course of the free energy calculation, concomitantly as the agonist ligand vanishes. It is also reflected in the initial translational and rotational motions of the TM α-helices. It is probable that reorientation of the conserved tryptophan side chain of TM VI is a signature mechanism ubiquitous to the activation or inactivation of GPCRs, as has been observed previously in the case of bovine rhodopsin, both at the experimental[59,60] and the theoretical level.[23] Yet, it is worth noting that in metarhodopsin II, isomerization of W265[61,27] occurs through its χ_1 torsional angle, and not χ_2, as described here for CCK1R.

The Role of the E/DRY and NPxxY Motifs in the Activation of CCK1R. In light of site-directed mutagenesis experiments, it has been suggested that one of the important events in the activation of class A GPCRs is the protonation

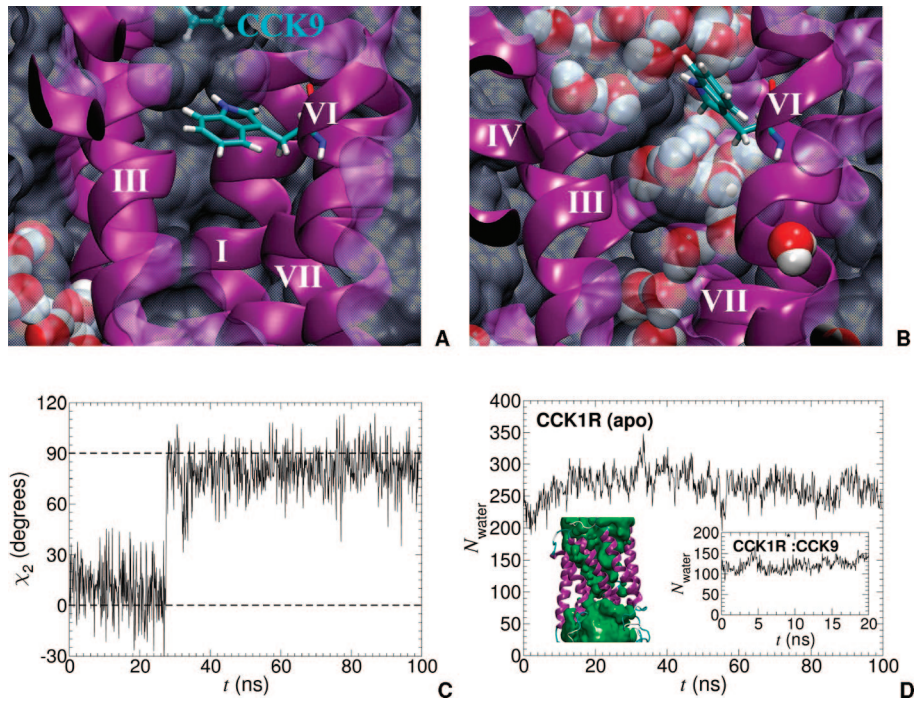


Figure 5. **A**. Cavity accommodating agonist ligand CCK9 in CCK1R*:CCK9 complex. The toggle switch W326 residue of TM helix VI is down. **B**. After 100 ns, the cavity in *apo* CCK1R is filled with water. W326 is up. **C**. Time-evolution of the χ_2 torsional angle of W326 in *apo* CCK1R. **D**. Instantaneous number of water molecules in the cavity of *apo* CCK1R. Inset: mass-weighted density of water in *apo* CCK1R measured over a period of 20 ns and highlighting a possible conduction between the cytoplasmic and the periplasmic face of the membrane (left). Number of water molecules in the cavity of CCK1R*:CCK9 (right). The density was computed with the VolMap module of Vmd.[65]

of the aspartic acid pertaining to the highly conserved E/DRY motif located at the bottom of TM III, on the cytoplasmic side of the membrane.[62] Computational studies have lent support to this conjecture, proposing, for instance, that during activation of the gonadotropin-releasing hormone receptor, an aspartate residue of TM III is protonated and subsequently replaced by another one in TM II to form a salt bridge with an arginine residue of TM III.[63] Experiment has also demonstrated that D83 of rhodopsin is hydrogen bonded upon activation, suggestive of a potential interaction with a neighboring residue in the active state of the protein.[64]

In the CCK1R*:CCK9 complex, E138 and R139 of TM III point in the same direction, albeit interact only mildly on account of their promiscuity. On the other hand, R139 forms a steady salt bridge with D87 of TM II, following a canonical C_{2v} symmetry pattern,[15] as highlighted in Figure 6. In the *apo* receptor, this ionic lock is preserved for approximately 15 ns, seemingly unaffected by the early separation of TM II and TM III. Its strength, however, subsides as the internal cavity of *apo* CCK1R reaches its maximum hydration, within roughly speaking 40 ns. In the meantime, the salt bridge breaks and reforms intermittently, until the two residues are sufficiently screened by the aqueous environment to no longer interact.

Another region of paramount importance for the activation of GPCRs is the NPxxY motif of TM VII. In the particular instance of CCK1R, mutation of N366 has been shown to abolish activation.[15] Over the 0.1-μs time scale, the present simulation reveals only marginal fluctuations about the original position of N366. Moreover, the hydrogen bond formed between N366 and R310 of the third intracellular loop is fully preserved. The same, however, cannot be said for residue M121 of TM III, which has been recognized to be involved indirectly in the activation of CCK1R through its hydrophobic environment.[29] Pointing toward the periplasm in the CCK1R*:CCK9 complex, M121 undergoes a conformational transition around 27 ns from *ca.* −50 to − 180°. Interestingly enough, this transition occurs concomitantly with the flip of W326 and the approach of TM II and TM III or TM V highlighted in Figures 5 and 4, respectively.

Relevance of the Molecular Dynamics Trajectories. At this stage of the study, is it legitimate to claim that after 0.1 μs, CCK1R has fully relaxed toward its inactive state? Although the receptor has undergone structural modifications, conversion from CCK1R* to CCK1R° is anticipated to span a commensurably longer time scale, not amenable to classical, all-atom MD simulations. While ascertaining without ambiguity that conformations of the inactive receptor have been sampled goes evidently beyond the scope of the present work, it is, nonetheless, possible to determine how the occupation of the conformational space has evolved over 0.1 μs. In addition to inferring displacements along anharmonic modes, essential dynamics (ED) will be employed to measure the overlap of the essential subspaces formed by the TM domain of CCK1R, explored at the beginning and at the end of the simulation, viz. in its first and last 20-ns leg (see the Supporting Information). Noteworthily, this principal component analysis reveals that modes corresponding to collective motions of large amplitude are suggestive of a functional movement, rather than a random diffusion. Of particular interest, projection onto the first essential mode of the 0.1-

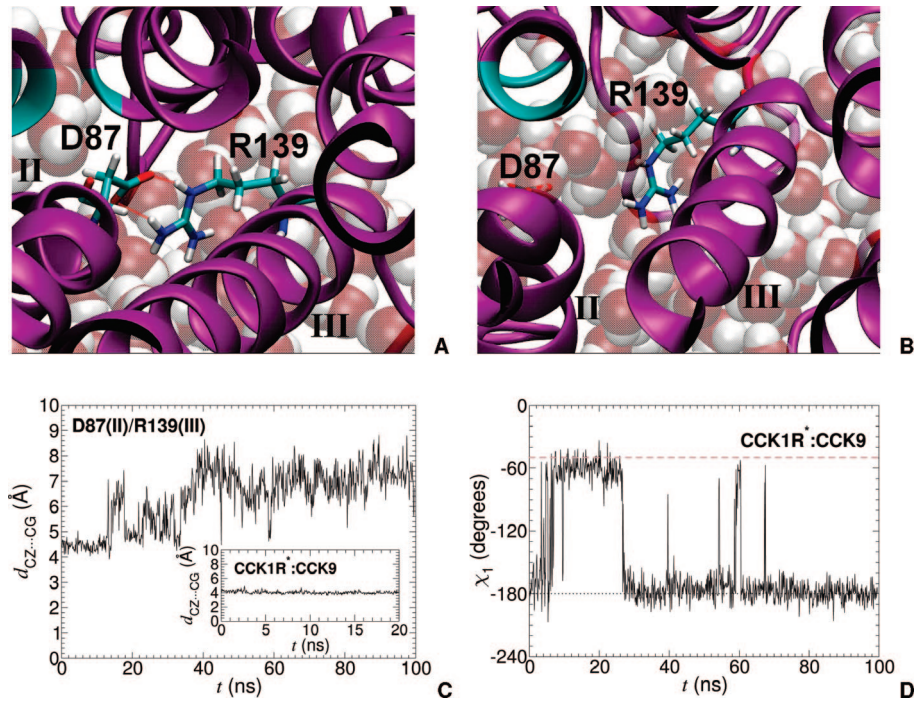


Figure 6. **A**. Salt bridge formed between residues D87 of TM helix II and R139 of TM helix III in the CCK1R*:CCK9 complex. **B**. Structural rearrangement in *apo* CCK1R and hydration of its internal cavity cause the rupture of the D87···R139 salt bridge. **C**. Time-evolution of the distance separating residues D87 and R139 in *apo* CCK1R. **D**. Time-evolution of the χ_1 torsional angle of M121 in *apo* CCK1R. The dashed light line depicts the value of χ_1 in the native CCK1R*:CCK9 complex.

μs trajectory suggests a possible tilting motion at the top of TM VI—i.e. seen from the extracellular side, consistent with previous observations in CCK2R[24] and seemingly concomitant with the displacement of TM III. This rationalizes the marked differences monitored in the interhelical separations when the latter are measured as distances between centers of mass or between selected residues.[25] More importantly, the overlap of the chosen essential subspaces—equal to 0.47, indicates that although these subspaces share analogous features, CCK1R can occupy markedly distinct regions of the conformational space over a 0.1-μs period.

Of equal concern is the choice of an initial conformation of CCK1R* to investigate by means of numerical simulations how the membrane protein evolves toward an inactive form. In other words, is the relaxation of the receptor in the course of an unphysical alchemical transformation relevant for the present study? To address this question, a second molecular assay was considered, wherein the agonist ligand, CCK9, bound to the hypothesized active receptor was removed abruptly from the latter. The additional, control, 50-ns MD trajectory highlights noteworthy differences in how the structural properties of CCK1R change as a function of time, compared to the reference, 0.1-μs simulation. Extraction of nonapeptide CCK9 from the binding pocket of CCK1R* results in the rapid reorganization of receptor, as suggested by the variations of the different interhelical separations, Δr (see the Supporting Information). In a number of instances, displacement of the TM segments appear to be at variance with the conclusions drawn from the 0.1-μs trajectory—e.g. whereas in the reference simulation, TM IV shifts away from TM VI, these two segments move toward each other in the shorter, control run. Similarly, rotation of the α-helices are somewhat different in the reference and in the control trajectories, barring TM IV, which in both cases, rapidly undergoes a temporary writhing motion. Abrupt removal of the agonist ligand perturbs significantly the structure of the crevice and is, thus, anticipated to drive the entire molecular assembly out of equilibrium. In sharp contrast with the long 0.1-μs simulation, hydration of the membrane protein is only partial (see the Supporting Information), viz. about 30% higher than in the CCK1R*:CCK9 complex. Yet, as has been seen previously, flooding of the internal cavity of CCK1R* constitutes an early event in its inactivation. Among the milestones paving the road toward CCK1R°, concerted reorganization of the TM scaffold triggers the signature toggle-switch isomerization of the side chain of W326 (see Figure 5), an event that is not observed in the 50-ns control simulation. From the present set of results, it may be legitimately contended that smooth removal of the agonist ligand ought to be preferred for investigating the inactivation of the receptor, allowing the latter to relax and adapt as annihilation of CCK9 proceeds in the binding pocket. That the key events recorded within *ca.* 27 ns of the reference trajectory remain undetected in the shorter, 50-ns control simulation does not necessarily imply that they will never be captured, but rather that this simulation is too short to allow a substantial perturbation—i.e. the extraction of the ligand, to be absorbed entirely by the molecular assembly.

Conclusions

On the long and winding road toward GPCR inactivation, FEP calculations have been carried out as a preliminary step, during which agonist ligand CCK9 was removed progressively from the binding pocket of CCK1R*, the human cholecystokinin receptor-1 in its putative activated conforma-

tion. Whereas the noteworthy accuracy of these unprecedented calculations does not validate *per se* the model of the receptor, the present results are, nonetheless, suggestive of an appropriate positioning of the ligand in the TM α-helix bundle, which further reinforces previous computations of relative binding affinities.[16] By shedding light onto the amino-acid residues participating in the membrane protein-ligand association, they also congrue with the wealth of experimental data accrued hitherto for CCK1R[15,18−21,29−34] thus challenging the pervading dogma that essentially no information of practical interest can be drawn from numerical simulations relying upon three-dimensional models of GPCRs. The additional MD simulation performed over the 0.1-μs time scale unveils early events of the inactivation process, sampling conformations of the receptor distinct from those adopted by the initial, native CCK1R*:CCK9 complex. The interplay of hydration and structural rearrangement of the TM scaffold is underscored by signature events, like the toggle-switch action of the well-conserved tryptophan residue of TM VI, the isomerization of the methionine residue of TM III, or the rupture of the salt bridge connecting TM II and TM III and involving the E/DRY motif ubiquitous to GPCRs. The latter are not random events but, on the contrary, constitute milestones toward the inactive conformation of the receptor, which are envisioned to be triggered by the flow of water molecules upon removal of the agonist ligand from the binding pocket. The subtle topological changes embodied in the fluctuations of the interhelical separations or the rotation of the TM segments appear to be sufficient to perturb durably the internal organization of the receptor. Intermediate conformations populating the path that leads to the inactive receptor, CCK1R°, constitute alternative targets of potential interest for rational drug design, which should help explain differences in binding affinities. As sizable as the assembly formed by *apo* CCK1R* and its membrane environment may be from the perspective of numerical simulations, it is still an incomplete model, from which the heterotrimeric G protein is clearly absent. Whether or not the latter component can be legitimately ignored in the simulation of GPCR activation remains an intriguing question likely to guide future modeling studies. The present contribution, nevertheless, provides valuable information on the important events that lead to activation, emphasizing the key role played by the environment and the necessity to explore these phenomena over sufficiently long time scales.

Acknowledgment. The author is indebted to Jérôme Hénin, Daniel Fourmy, François Dehez, Andrew Pohorille, and Bernard Maigret for fruitful discussions. The Cines, Montpellier, France, is gratefully acknowledged for provision of generous amounts of computational time.

Supporting Information Available: Details of the free energy calculation and convergence criteria thereof, the essential dynamics post-treatment, and the measure of key observables characterizing the spatial rearrangement of the receptor toward its inactive conformation together with the analyses performed from the control simulation. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Bockaert, J.; Pin, J. P. *EMBO J.* **1999**, *18*, 1723–1728.

(2) Lefkowitz, R. J. *Nat. Cell Biol.* **2000**, *2*, E133−E136.

(3) Sautel, M.; Milligan, G. *Curr. Med. Chem.* **2000**, *7*, 889–896.

(4) Gether, U. *Endocr. Rev.* **2000**, *21*, 90–113.

(5) Lefkowitz, R. J.; Cotecchia, S.; Samama, P.; Costa, T. *Trends Pharmacol. Sci.* **1993**, *14*, 303–307.

(6) Kenakin, T. *Trends Pharmacol. Sci.* **1997**, *18*, 416–417.

(7) Kenakin, T. *Trends Pharmacol. Sci.* **2004**, *25*, 186–192.

(8) Swaminath, G.; Deupi, X.; Lee, T. W.; Zhu, W.; Thian, F. S.; Kobilka, T. S.; Kobilka, E. *J. Biol. Chem.* **2005**, *280*, 22165–22171.

(9) Fanelli, F.; Benedetti, P. G. D. *Chem. Rev.* **2005**, *105*, 3297–3351.

(10) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739–745.

(11) Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Trong, I. L.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16123–16128.

(12) Vaidehi, N.; Floriano, W. B.; Trabanino, R.; Hall, S. E.; Freddolino, P.; Choi, E. J.; Zamanakos, G.; Goddard, W. A., III *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12622–12627.

(13) Ballesteros, J.; Palczewski, K. *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 561–574.

(14) Becker, O. M.; Shacham, S.; Marantz, Y.; Noiman, S. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 353–361.

(15) Archer-Lahlou, E.; Tikhonova, I.; Escrieut, C.; Dufresne, M.; Seva, C.; Clerc, P.; Pradayrol, L.; Moroder, L.; Maigret, B.; Fourmy, D. *J. Med. Chem.* **2005**, *48*, 180–191.

(16) Hénin, J.; Maigret, B.; Tarek, M.; Escrieut, C.; Fourmy, D.; Chipot, C. *Biophys. J.* **2006**, *90*, 1232–1240.

(17) Villanueva, M. L.; Collins, S. M.; Jensen, R. T.; Gardner, J. D. *Am. J. Physiol.* **1982**, *242*, G416–G422.

(18) Gigoux, V.; Escrieut, C.; Silvente-Poirot, S.; Maigret, B.; Gouilleux, L.; Fehrentz, J. A.; Gully, D.; Moroder, L.; Vaysse, N.; Fourmy, D. *J. Biol. Chem.* **1998**, *273*, 14380–14386.

(19) Gigoux, V.; Escrieut, C.; Fehrentz, J. A.; Poirot, S.; Maigret, B.; Moroder, L.; Gully, D.; Martinez, J.; Vaysse, N.; Fourmy, D. *J. Biol. Chem.* **1999**, *274*, 20457–20464.

(20) Gigoux, V.; Maigret, B.; Escrieut, C.; Silvente-Poirot, S.; Bouisson, M.; Fehrentz, J. A.; Moroder, L.; Gully, D.; Martinez, J.; Vaysse, N.; Fourmy, D. *Protein Sci.* **1999**, *8*, 2347–2354.

(21) Escrieut, C.; Gigoux, V.; Archer, E.; Verrier, S.; Maigret, B.; Behrendt, R.; Moroder, L.; Bignon, E.; Silvente-Poirot, S.; Pradayrol, L.; Fourmy, D. *J. Biol. Chem.* **2002**, *277*, 7546–7555.

(22) Gouldson, P. R.; Kidley, N. J.; Bywater, R. P.; Psaroudakis, G.; Brooks, H. D.; Diaz, C.; Shire, D.; Reynolds, C. A. *Proteins* **2004**, *56*, 67–84.

(23) Niv, M. Y.; Skrabanek, L.; Filizola, M.; Weinstein, H. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 437–448.

(24) Marco, E.; Foucaud, M.; Langer, I.; Escrieut, C.; Tikhonova, I. G.; Fourmy, D. *J. Biol. Chem.* **2007**,

(25) Tikhonova, I.; Best, R.; Engel, S.; Gershengorn, M.; Hummer, G.; Costanzi, S. *J. Am. Chem. Soc.* **2008**, *130*, 10141–10149.

(26) Martínez-Mayorga, K.; Pitman, M. C.; Grossfield, A.; Feller, S. E.; Brown, M. F. *J. Am. Chem. Soc.* **2006**, *128*, 16502–16503.

(27) Crozier, P. S.; Stevens, M. J.; Woolf, T. B. *Proteins* **2007**, *66*, 559–574.

(28) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. *Nov* **2007**, *318* (5854), 1258–1265.

(29) Fourmy, D.; Escrieut, C.; Archer, E.; Galès, C.; Gigoux, V.; Maigret, B.; Moroder, L.; Silvente-Poirot, S.; Martinez, J.; Fehrentz, J. A.; Pradayrol, L. *Mol. Pharmacol.* **2002**, *91*, 313–320.

(30) Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. *Trends Pharmacol. Sci.* **2003**, *24*, 36–40.

(31) Archer-Lahlou, E.; Escrieut, C.; Clerc, P.; Martinez, J.; Moroder, L.; Logsdon, C.; Kopin, A.; Seva, C.; Dufresne, M.; Pradayrol, L.; Maigret, B.; Fourmy, D. *J. Biol. Chem.* **2005**, *280*, 10664–10674.

(32) Mart Jourdan, M.; Escrieut, C.; Archer, E.; González-Muniz, R.; García-López, M. T.; Maigret, B.; Herranz, R.; Fourmy, D. *J. Med. Chem.* **2005**, *48*, 4842–4850.

(33) Tikhonova, I. G.; Marco, E.; Lahlou-Archer, E.; Langer, I.; Foucaud, M.; Maigret, B.; Fourmy, D. *Curr. Top. Med. Chem.* **2007**, *7* (12), 1243–1247.

(34) Foucaud, M.; Archer-Lahlou, E.; Marco, E.; Tikhonova, I. G.; Maigret, B.; Escrieut, C.; Langer, I.; Fourmy, D. *Regul. Pept.* **2008**, *145*, 17–23.

(35) Dawson, E. S.; Henne, R. M.; Miller, L. J.; Lybrand, T. P. *Pharmacol. Toxicol.* **2002**, *91*, 290–296.

(36) Giragossian, C.; Mierke, D. F. *Biochemistry* **2001**, *40*, 3804–3809.

(37) Chipot, C.; Pohorille, A., Eds., *Free energy calculations. Theory and applications in chemistry and biology*; Springer Verlag: 2007.

(38) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(39) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(40) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

(41) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(42) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Phys. Chem. B* **1992**, *97*, 1990–2001.

(43) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*3586–3616.

(44) Feller, S. E.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2000**, *104*, 7510–7515.

(45) Lu, N.; Kofke, D. A.; Woolf, T. B. *J. Comput. Chem.* **2004**, *25*, 28–39.

(46) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

(47) Fourmy, D.; Lopez, P.; Poirot, S.; Jimenez, J.; Dufresne, M.; Moroder, L.; Powers, S. P.; Vaysse, N. *Eur. J. Biochem.* **1989**, *185*, 397–403.

(48) Kennedy, K.; Gigoux, V.; Escrieut, C.; Maigret, B.; Martinez, J.; Moroder, L.; Fréhel, D.; Gully, D.; Vaysse, N.; Fourmy, D. *J. Biol. Chem.* **1997**, *272*, 2920–2926.

(49) Miyamoto, S.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402–8406.

(50) Dixit, S. B.; Chipot, C. *J. Phys. Chem. A* **2001**, *105*, 9795–9799.

(51) Altenbach, C.; Cai, K.; Klein-Seetharaman, J.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **2001**, *40*, 15483–15492.

(52) Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G. *Science* **1996**, *274*, 768–770.

(53) Ghanouni, P.; Steenhuis, J. J.; Farrens, D. L.; Kobilka, B. K. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5997–6002.

(54) Jensen, A. D.; Guarnieri, F.; Rasmussen, S. G.; Asmar, F.; Ballesteros, J. A.; Gether, U. *J. Biol. Chem.* **2001**, *276*, 9279–9290.

(55) Gether, U.; Kobilka, B. K. *J. Biol. Chem.* **1998**, *273*, 17979–17982.

(56) Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217–11224.

(57) Visiers, I.; Ballesteros, J. A.; Weinstein, H. *Methods Enzymol.* **2002**, *343*, 329–371.

(58) Grossfield, A.; Pitman, M. C.; Feller, S. E.; Soubias, O.; Gawrisch, K. *J. Mol. Biol.* **2008**, *381*, 478–486.

(59) Lin, S. W.; Sakmar, T. P. *Biochemistry* **1996**, *35*, 11149–11159.

(60) Patel, A. B.; Crocker, E.; Reeves, P. J.; Getmanova, E. V.; Eilers, M.; Khorana, H. G.; Smith, S. O. *J. Mol. Biol.* **2005**, *347*, 803–812.

(61) Crocker, E.; Eilers, M.; Ahuja, S.; Hornak, V.; Hirshfeld, A.; Sheves, M.; Smith, S. O. *J. Mol. Biol.* **2006**, *357*, 163–172.

(62) Arnis, S.; Fahmy, K.; Hofmann, K. P.; Sakmar, T. P. *J. Biol. Chem.* **1994**, *269*, 23879–23881.

(63) Ballesteros, J.; Kitanovic, S.; Guarnieri, F.; Davies, P.; Fromme, B. J.; Konvicka, K.; Chi, L.; Millar, R. P.; Davidson, J. S.; Weinstein, H.; Sealfon, S. C. *J. Biol. Chem.* **1998**, *273*, 10445–10453.

(64) Rath, P.; DeCaluwé, L. L.; Bovee-Geurts, P. H.; DeGrip, W. J.; Rothschild, K. J. *Biochemistry* **1993**, *32*, 10277–10282.

(65) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

CT800313K