Three-Dimensional Model for Meta-II Rhodopsin, an Activated G-Protein-Coupled Receptor[†]

Gregory V. Nikiforovich* and Garland R. Marshall

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

Received April 11, 2003; Revised Manuscript Received June 1, 2003

ABSTRACT: A novel approach that iteratively combined the results of energy calculations and experimental data was used to generate a three-dimensional (3D) model of the photoactivated state (R*) of bovine rhodopsin (Rh). The approach started with simplified energy calculations in an effort to find a set of sterically and energetically reasonable options for transmembrane (TM) helix arrangements with *all-trans*-retinal. Various 3D models of TM helix packing found by computations were then compared to limited site-directed spin-label experimental data regarding the transition of the TM helices of Rh in the inactive state (R) to those in the R* state to identify the most plausible model of the TM helical bundle. At the next step, all non-TM structural elements, such as the non-TM helix 8, the N- and C-terminal fragments, and the loops connecting TM helices, were reconstructed, and after the entire R* structure had been relaxed, all other currently available additional experimental data, both mutational and spectroscopic, on the structure of the meta-II state of rhodopsin were used to validate the resulting 3D model.

Rhodopsin (Rh),[1] the 348-residue seven-transmembrane α-helical photoreceptor of the visual system, is the prototype for a vast subclass of G-protein-coupled receptors (GPCRs). The chromophore for this receptor is retinal that is covalently attached to K296. In the dark-adapted state, retinal has the 11-*cis* isomer that stabilizes the inactive dark-adapted conformation of Rh (termed hereafter R). After being exposed to light, 11-*cis*-retinal isomerizes into *all-trans*-retinal in less than 200 fs. The photoactivation cycle may be briefly described as follows (*1*). The resulting PHOTO state where *all-trans*-retinal is highly distorted transforms into the BATHO state in picoseconds, then to the LUMI state in nanoseconds (at this stage, *all-trans*-retinal is presumably no longer distorted), then to the meta-I (MI) state in microseconds, and, finally, to the meta-II (MII) state (termed here R*) in milliseconds. R* is the state where the Schiff base of retinal is unprotonated. At this stage, photoactivated Rh acts like other GPCRs; i.e., it activates its G-protein transducin, and triggers the visual transduction process.

Several hundred GPCRs belong to the so-called rhodopsin family [sometimes called family A (*2*)], the largest known family of GPCRs that are distinctly sequentially homologous to Rh. GPCRs, in general, serve to transduce an extracellular event, typically binding of an extracellular message (neurotransmitter, peptide hormone, etc.) to an intracellular signal (cAMP production, cGMP hydrolysis, or activation of phosphoinositol-specific phospholipase C) to generate IP_3 and other second messengers, etc. It has been estimated that almost 50% of the therapeutic compounds in use act on GPCRs (*3*). Knowledge of the detailed three-dimensional (3D) structures of activated GPCRs would, therefore, be extremely relevant to wide areas of biochemistry, biophysics, and medicinal chemistry.

At the same time, dark-adapted Rh (R) is the only GPCR whose 3D structure has been determined by X-ray crystallography (*4−6*). Despite obvious differences between photoreceptors and other GPCRs (see the discussion on the similarity of Rh and other GPCRs in refs *7* and *8*), 3D models for Rh have been used as templates for building 3D structures of other GPCRs in their inactive states (see, e.g., ref *9* for a minireview). Templates based on Rh may be also extremely useful for understanding processes of activation of specific GPCRs (*2*, *8−10*), provided that 3D structures corresponding to R* are available or, at least, are hypothesized as working 3D models based on experimental data and/or computational simulations.

Experimental observations of light-induced conformational changes occurring during the transition from R to R* were recently reviewed (*10−12*). Roughly, they can be divided into several main categories: direct distance measurements between site-directed spin-labels (SLs) inserted into Rh by site-directed mutagenesis (*13−15*), estimations of the changes in mobility for SLs in various mutants of Rh (*16−21*), estimations of the changes in accessibility of the Cys residues in the Rh mutants to various sulfhydryl-reacting reagents (*22*, *23*), and studies of photoactivation of the Rh mutants that are cross-linked by various additional disulfide bonds (*13*, *24−28*) or by employing Zn^{2+} ions to form bridges between histidines in Rh mutants (*29*). Roughly, these results were interpreted as an indication of concerted movements of several transmembrane (TM) helices of Rh during the transition from R to R* (see ref *11* for a recent review). Specifically, it was suggested that the C-terminal part of TM6

---

[1] Abbreviations: Rh, rhodopsin; GPCR, G-protein-coupled receptor; R, inactive state; R*, activated state; TM, transmembrane or transmembrane helix; 3D, three-dimensional; MI, meta-I state; MII, meta-II state; cAMP, cycloadenosine monophosphate; cGMP, cycloguanine monophosphate; IP, inositol phosphate; SL, spin-label; SDSL, site-directed spin labeling; MD, molecular dynamics; IC, intracellular; NMR, nuclear magnetic resonance; PDB, Protein Data Bank.

rotates clockwise and moves away from the core TM region of Rh by up to 8 Å, accompanied by less significant movements of TM3, TM7, and TM2 away from the core TM region (see, e.g., Figure 8 in ref *14*). However, detailed structural interpretation of the available experimental data is much more difficult. The most direct distance measurements available from site-directed spin labeling (SDSL) can only determine rather wide ranges of possible distances between two Cys residues because of significant internal flexibility of the Cys−SL constructs that were used (five rotatable bonds between the nitroxide SL and the C_α atom of the attached Cys) as well as due to intrinsic difficulties in direct structural interpretation of ESR spectra (*11*, *14*, *15*, *30*). Structural interpretation is also difficult for the Rh mutants with disulfide bond cross-links permitting or inhibiting structural changes from R to R* (see, e.g., ref *28*). Here the major problem is that conformational changes associated with a particular disulfide cross-link are not necessarily the same as those associated with a different cross-link, though each of them may allow activation of R to R*; see also the Results and Discussion.

Nevertheless, experimental data have been used for 3D modeling of R* by applying as many experimental constraints as were available to the NMR-derived structure of the dark-adapted Rh [which is in good agreement with the X-ray structure (*4*)] with subsequent molecular dynamics (MD) simulation for 200 ps (*31*, *32*). The resulting model of R* (PDB entry 1LN6) significantly differs from the X-ray structure of R. TM helices in 1LN6 are not simply moved and rotated compared to the R structure, but in fact, their very helical structure in some fragments of TM5−TM7 is destroyed. Also, *all-trans*-retinal in 1LN6 is detached from residue K296 instead of being covalently linked (the valence bond $N_\epsilon^{K296}-C_{15}^{ret}$ distance is 4.76 Å in 1LN6). These limitations make it difficult to use 1LN6 as a working 3D model of R*.

On the other hand, current computational studies by direct molecular dynamics (MD) simulations that accurately account for the complex membrane environment of Rh are unable to follow the entire transition from R to R* that requires milliseconds. The first attempts featured an MD trajectory for R* (with *all-trans*-retinal) of ca. 400 ps, and no noticeable changes in the 3D structure of Rh were observed (*33*). More extensive MD simulations that accounted for the modeled lipid/water environment focused on the effects of protonation/deprotonation of the E181 residue in the vicinity of retinal (*34*). The *cis−trans* isomerization of retinal was successfully achieved in picoseconds by applying an additional harmonic potential, and one of the simulations with the *all-trans*-retinal has been prolonged to 7 ns. The main difference in the structure resulting from R was that the ionone ring of retinal moved from residue Y268 toward residue A169, in accordance with experimental data (*1*). Some structural rearrangements in spatial positions of TM6, TM5, and TM4 helices were also observed (*34*). These observations were confirmed by even more extensive MD simulations (for 10 ns, each nanosecond requiring ca. 4.5 days on 128 processors of the Cray T3E supercomputer), where the greatest deviations from the R structure were observed in TM6, TM5, and TM1 helices (*35*). Interestingly, the authors did not observe rotation of TM6 to the extent that was postulated by site-directed spin labeling (*13*);

instead, this rotation of TM6 in R* was examined in a separate study by a forcing potential with steered molecular dynamics (*35*). The study also suggested folding of the intracellular loop connecting TM5 and TM6 helices (loop IC3) inward toward the molecule core (*35*); however, experimental observations, on the other hand, suggest creation of the binding site for transducin in R* by movement of IC3 and IC2 (the loop between TM3 and TM4) outward (*36*; see also ref *37* and references therein).

In summary, a working 3D model of R*, which may have important general mechanistic implications for other GPCRs, is difficult, if not impossible, to deduce either from direct MD simulations that accurately account for the complex environment of Rh* (trajectories of milliseconds are required) or from available experimental data (insufficient precise experimental constraints for direct structural interpretation). At the same time, the urgent need for a working 3D model of R* is evident. Therefore, this study presents a different approach that focuses not on the process of transition from R to R* but on the final structural features of the MII state of rhodopsin R* with *all-trans*-retinal. The approach iteratively combines the results of energy calculations and experimental data. The approach starts with simplified energy calculations in an effort to find a set of sterically and energetically reasonable options of TM helix packing with *all-trans*-retinal. Various 3D models of TM helix packing found by calculations are then compared to limited SDSL experimental data related to the TM region of Rh in the R* state to select the most plausible models. In the next step, non-TM helix 8 as well as the N- and C-terminal fragments and the interhelical loops is added to the candidate helical packing. This assumes that the non-TM fragments may retain in the R* state the same local energy minima they have possessed in the R state. After the entire R* structure had been relaxed, all other currently available additional experimental data, both mutational and spectroscopic, on the structure of the meta-II state of rhodopsin are used for validation of the resulting 3D model.

METHODS

Generally, 3D models for GPCRs can be built in a "block-by-block" manner by performing several distinct steps: (i) finding possible TM helices in a sequence, (ii) assembling of a TM helical bundle, and (iii) restoration of interhelical loops (*38−41*). For Rh, the first step is unnecessary, since boundaries of TM helices can be determined directly from the X-ray structure.

Energy Calculations for Obtaining 3D Models of the TM Region of R.* Packing of the seven-helix bundle for the starting 3D model(s) of the TM region of Rh in the R* state was performed according to a previously described procedure (*42*). It consisted of minimization of the sum of all intra- and interhelical interatomic energies [the ECEPP/2 force field with rigid valence geometry (*43*, *44*) was used] in the multidimensional space of parameters that included the "global" parameters (those related to movements of individual helices like rigid bodies, namely, translations along the coordinate axes *X*, *Y*, and *Z* and rotations T_x, T_y, and T_z, respectively, around these axes) and the "local" parameters [the dihedral angles of the side chains for all helices; the starting values of those angles were optimized prior to energy

minimization with an algorithm described previously (*45*)]. At this stage, the dihedral angles of the peptide backbones were frozen at the values corresponding to the X-ray structure of Rh [PDB entry 1F88 (*4*); the more recent entries, 1HZX (*5*) and 1L9H (*6*), are essentially identical to 1F88 in their TM regions]; also, the interatomic electrostatic interactions were excluded. The coordinate system for the global parameters was selected as follows. The long axial *X* coordinate axis for each TM helix has been directed from the first to the last C_α atom; the *Y* axis was perpendicular to the *X* axis and went through the C_α atom of the "middle" residue of each helix, and the *Z* axis was built perpendicular to the *X* and *Y* axes to maintain the right-handed coordinate system. For this purpose, boundaries of TM helices in Rh were defined as follows: TM1, W35−M49−Q64 (the first, middle, and last residues, respectively); TM2, L72−F85−L99; TM3, T108−I123−V139; TM4, E150−W161−L172; TM5, N200−F212−Q225; TM6, A246−F261−T277; and TM7, I286−K296−Y306. All carbohydrate moieties, water molecules, ions, detergents, and additive molecules present in 1F88 have been removed; chain A has been selected as a template. Unprotonated retinal was attached to the side chain of K296 in the *all-trans* conformation; the valence structure of 11-*cis*-retinal was borrowed from 1F88. Arg, Lys, Glu, and Asp residues were present as charged species. The dihedral angles of retinal and K296 were allowed to change during energy minimization, but they were not optimized prior to minimization (except the last $C_5−C_6−C_7−C_8$ angle in retinal). One computer run for each combination of global parameters required ca. 20−30 h at a single PC 500−1000 MHz processor under the Linux operational system.

Simplified Energy Calculations in the Space of Global Parameters. The term "simplified energy calculations" refers to helical packing where energy minimization has been performed only in the space of global parameters (though the values of the dihedral angles of side chains were still optimized prior to energy minimization). Accordingly, the minimized energy consisted only of the interhelical, but not intrahelical, interatomic interactions. In this case, one run of energy minimization for a given combination of global parameters required up to only 1−1.5 h at a single PC 500−1000 MHz processor.

Restoring Structural Elements Missing in the X-ray Structures. Some fragments of the peptide chain are missing in all available X-ray structures of Rh. In 1F88, they are fragments A235−S240 of loop IC3 and P327−S334 that connects non-TM helix 8 to the C-terminal tail (*4*). In 1HZX and 1L9H, they are fragments A235−A241 and D330−S334 (*5, 6*). Also, there is an ambiguity between 1F88, on one hand, and 1HZX and 1L9H, on the other, in spatial positions of some residues belonging to loop IC3. Namely, fragment S240−T243 in 1F88 is oriented inward with respect to the core molecule, whereas fragment A241−T243 is turned outward in 1HZX and 1L9H. To accommodate this structural ambiguity, the intracellular fragments A234−T243 and G324−S334 have been independently restored according to the previously described computational procedure (*39*). Specifically, the procedure generated conformations of the peptide backbone obtained by all possible combinations of local energetic minima on the Ramachandran map for residues A234−T243 and G324−S334 within the immobile 1F88 structure that lacks these particular fragments. Frag-

ments were attached to the residues of their origin, i.e., to A234 and G324, respectively. Two filters were used to select conformations of the fragments for further consideration. First, to avoid steric collisions, only conformations where all $C_\alpha−C_\alpha$ distances within the fragment as well as between any C_α atom of the fragment and any C_α atom in the rest of the molecule were less than 4.0 ± 1.0 Å were selected. Second, only those conformations were selected where $C_\alpha−C_\alpha$ distances between the end residue of a generated conformation and the corresponding immobile residue in the rest of the molecule (i.e., residues T243 and S334, respectively) were less than [number of residues in a fragment] × (1.5 ± 2) Å (this empirical dependence was deduced from analysis of protein loops in the Protein Data Bank). Selected conformers of both fragments (171 and 788 conformers, respectively) were subjected to energy minimization within the larger fragments K231−A246 and L321−V337. In that step, the dihedral angles of the peptide backbone for the "additional" N- and C-terminal tripeptides were fixed in their values in 1F88, and additional parabolic potentials were added to the conformational energy to keep $C_\alpha−C_\alpha$ distances between the additional tripeptides exactly as they are in 1F88. The obtained low-energy structures of both fragments (those with a relative energy of <15 kcal/mol; 44 and 66 conformers, respectively) were then clustered by geometrical similarity (in the same cluster, the rms values for all C_α atoms in the fragment were less than 2.0 Å), and checked again for possible steric collisions with the rest of the molecule. Six clusters were found for each of the fragments 234−243 and 324−334. Finally, the single "mean" 3D structures for these fragments were obtained by averaging over positions of C_α atoms in representatives of all the clusters without steric collisions with the immobile remainder of 1F88.

Building and Relaxing the Entire 3D Structure of R*. When the most plausible 3D model of the TM region of R* had been selected, the starting 3D structure for the entire molecule of R* was obtained by manually adding to the TM region the remaining fragments, i.e., the N-terminal tail, extra- and intracellular loops, and the C-terminal tails. These fragments were added using the dihedral angle values calculated for their atomic Cartesian coordinates in 1F88 (or, for fragments missing in the X-ray structures, using the dihedral angle values for the mean structures described in the previous subsection). During each subsequent addition, the starting dihedral angles were slightly adjusted to the Cartesian coordinates in 1F88 to avoid discrepancies between the rigid valence geometry used in the ECEPP force field and the deviations of valence geometry in 1F88. Finally, energy minimization was performed for the entire molecule involving all dihedral angles (including those in the peptide backbone of the TM region that were frozen at the earlier stages) and including electrostatic interactions.

RESULTS AND DISCUSSION

TM Region of R*. The starting 3D model of the TM region for R* was built by energy calculations as described in Methods. The starting values for the 42 global parameters were the same as in 1F88 (the R state). The resulting 3D model differed from that of R by the rms value of 2.23 Å (C_α atoms only).

Then, for each TM helix in R*, independent rotations around the T_x axis on a grid of 30° were considered (in total,
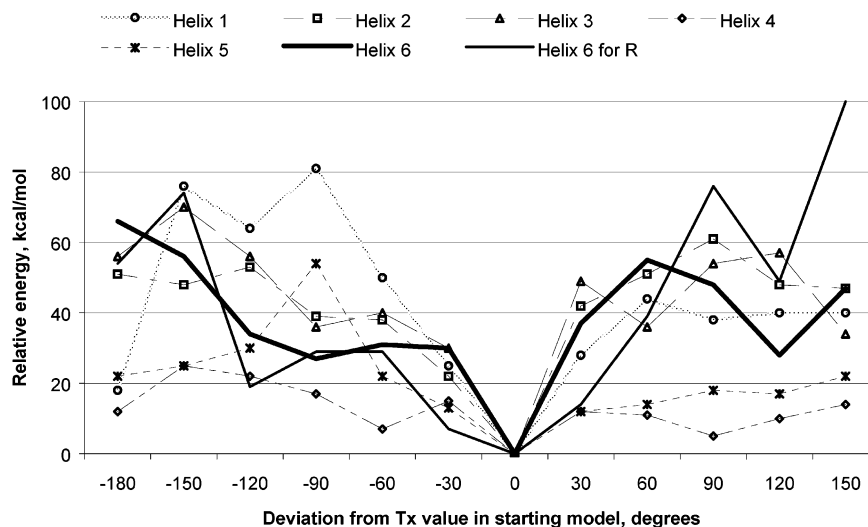
FIGURE 1: Energy profiles of rotation of TM helices in Rh around the T_x axes.

12 positions for each helix); all other starting values for global parameters were the same as in the starting model of R*. Simplified energy calculations were performed for each position of T_x, yielding energetic profiles for each helix (see Figure 1; the energetic profile for TM7 is absent in Figure 1, since any possible rotation of TM7 is severely hindered by the *all-trans*-retinal linked to K296). The energetic profile for rotation of TM6 in R also is presented in Figure 1 for comparison. Deviations in T_x values are shown in increments relative to zero, which is the starting position for a given T_x rotation in R.

Energy profiles in Figure 1 provide a rough general indication as to which TM helix can rotate around the T_x axis within the TM bundle of R*, and which rotational positions are sterically and energetically more reasonable. The energy profiles in Figure 1 are, in fact, the upper bound estimates, since actual transitions in helical packing seen with photoactivation involve concerted movements of the helices; studies of the exact low-energy path from R to R* are beyond the limits of this paper. It is evident, for instance, that the profile for TM4 is essentially flat, with no rotational position really preferred. On the other hand, helices TM2 and TM5 possess a preferred rotational position at a Δ_{T_x} of 0°. As for the other three helices, one can identify several preferred rotational positions with relative energies of <40 kcal/mol (an arbitrary energy cutoff), namely, $\Delta_{T_x} = -180°$, 0°, and 90° for TM1, −90°, 0°, 60°, and 150° for TM3, and −90°, 0°, and 120° for TM6, suggesting 36 possible low-resolution options of TM helix packing for further consideration. Interestingly, the energy profile for TM6 in R (thin solid line in Figure 1) suggests two preferred rotational positions, namely, $\Delta_{T_x} = -120°$ and 0° being different from that of TM6 in R* (thick solid line in Figure 1).

Simplified energy calculations were performed for each of 36 combinations of TM helix packing. The lowest energy was found for TM helix packing with all Δ_{T_x} values equal to 0°, i.e., for the starting 3D model corresponding to R. Seven other combinations, however, possessed relative energy values of <40 kcal/mol, and were selected as possibilities for TM helix packing alternatives to that seen with R. To determine the most plausible R* state, the $C_\alpha-C_\alpha$ distances between residues 139 and 248−252 were calculated for all eight models of TM helical packing. These

particular inter-residue distances have been estimated for R and R* by SDSL (*13*). The experimental estimations cannot be regarded as exact measurements of $C_\alpha-C_\alpha$ distances, but rather as indications of general trends for structural changes in the transition from R to R*. Namely, the experimental data suggest that distances between the nitroxide spin-labels attached to residues 139 and 248 or to residues 139 and 251 increase from 12−14 to 23−25 Å after the transition from R to R*, which is more significant than the distance from residue 139 to 252 (from 15−20 to 23−25 Å); that distance from residue 139 to 249 remains basically the same (15−20 Å), and that distance from residue 139 to 250 decreases (from 15−20 to 12−14 Å) (*13*). Previously, we have rescaled these experimental distance ranges for the nitroxide spin-labels in both R and R* to be consistent with the actual $C_\alpha-C_\alpha$ distances in 1F88 (*37*); these rescaled experimental estimated $C_\alpha-C_\alpha$ distances in R* were compared to those calculated for all eight possible options of TM helix packing in the R* state.

Only one option, that of Δ_{T_x} values of 0°, 0°, and 120° for TM1, TM3, and TM6, respectively, correlated with experimental observations ($R^2 = 0.94$; see Figure 2); all other options produced negative correlations with low correlation coefficients. This model of TM helical packing was, therefore, selected as the most probable candidate for an energetically reasonable 3D model of the R* state compatible with limited SDSL data.

The selected type of TM region of R* differs from that of R, in fact, only by rotation of TM6 around the T_x helical axis, as proposed previously (*13*) based on SDSL data. The quality of the correlation seen in Figure 2 suggests that the simplified energetic approach to helical packing has correctly identified the orientation of helices 3 and 6 in accord with experimental data, and that this helical packing model can be readily selected as the best starting point for further refinement of a 3D model of R*. To account for possible structural changes associated with tilts of TM6 around helical axes T_y and T_z, as well as with small translational movements along coordinate axes *X*, *Y*, and *Z*, the same simplified energy calculations were performed for the selected 3D model of the TM region of R* employing the following complete grid: $\Delta_{T_y,T_z} = 0°$, ±10°, ±20° and $\Delta_{X,Y,Z} = 0$, ±1 Å for TM6 positioned at $\Delta_{T_x} = 120°$. However, any deviations from
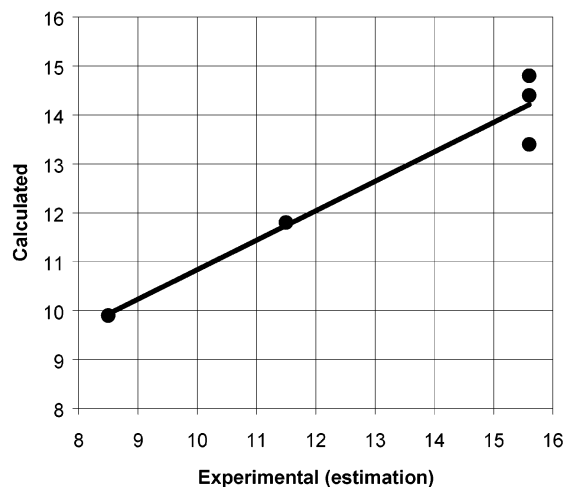
FIGURE 2: Correlation between $C_\alpha-C_\alpha$ distances from residue 139 to residues 248−252 (angstroms) calculated for the selected model of TM helix packing for R* and the rescaled distances estimated from the experimental SDSL data (*13*) for R*.
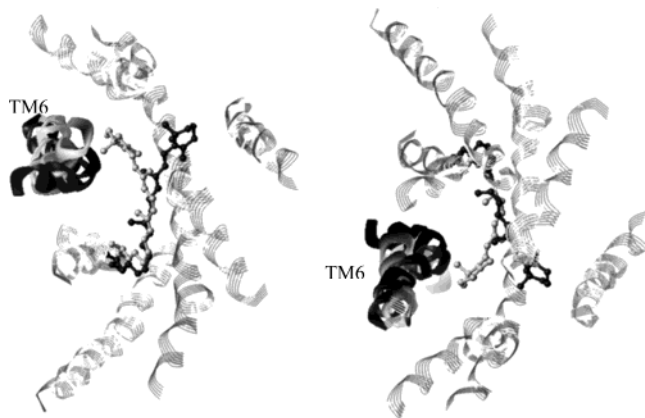

FIGURE 3: Intracellular (left) and extracellular (right) views on the TM helical region of R (light gray) overlapped over the TM helical region of R* (dark gray). TM6 helices are shown as shaded ribbons; retinal is shown as a ball-and-stick model. TM helices 1−5 and 7, but not TM6, have been overlapped above.

$\Delta_{T_y,T_z} = 0°$ and/or $\Delta_{X,Y,Z} = 0$ Å resulted either in much higher relative energies or in worse agreement with the experimental data (data not shown). Therefore, the selected 3D model for the TM region of R* remained as follows: $\Delta_{T_x} = 120°$, $\Delta_{T_y,T_z} = 0°$, and $\Delta_{X,Y,Z} = 0$ Å for TM6. The model is shown in Figure 3; it clearly shows that rotation of the kinked TM6 helix fills the cavity left by retinal after isomerization.

Interestingly, Figures 1 and 3 imply that the role of 11-*cis*-retinal in rhodopsin may simply be stabilization of TM6 in its dark-adapted rotational state with the intracellular loops occluding the binding site for transducin. Photoisomerization to *all-trans*-retinal simply acts as a latchkey allowing the different rotational state ($\Delta_{T_x} = 120°$, $\Delta_{T_y,T_z} = 0°$, and $\Delta_{X,Y,Z} = 0$ Å) of TM6 to be energetically accessible. Rotation of TM6 primarily imposes a conformational change on the IC3 loop that exposes the binding site for the C-terminal segment of the α-subunit of transducin. These three steps (photoisomerization, rotation of TM6, and movement of the IC3 loop to expose a binding site for transducin) are the essential components for photoactivation of rhodopsin.

3D Model of the Entire R Molecule.* The starting 3D model for the complete 3D structure of R* was built by step-by-step addition of the N-terminal fragment, the intra- and extracellular loops, and the C-terminal fragment to the 3D model of the TM region of R*, i.e., as sequential elongation of the Rh sequence from residues 1−64 to 1−99 to 1−139 to 1−172 to 1−225 to 1−234 to 1−277 to 1−306 to 1−348. The single mean conformations for fragments 234−243 and 324−334 (see Methods), absent in 1F88, were used at this step. The fundamental assumption was that the non-TM fragments of Rh do not significantly change their conformations as a result of the transition from R to R*, or they remain, at least, in the same local energetic minima they possessed in R. This assumption seems reasonable, since retinal remains covalently bound to K296 during the transition from R to R*; only the further photocyclic transition from R* to opsin might require significant movements of the extracellular parts of rhodopsin to release *all-trans*-retinal. On the other hand, conformational changes in the intracellular parts of Rh, necessary to bind transducin and/or rhodopsin kinase, may be limited to rigid body movements and, in any case, could be reproduced correctly only in 3D model(s) of the complex of Rh−transducin or Rh−rhodopsin kinase, far beyond the scope of this study. In fact, this study aims to provide a working 3D model of R* that, in turn, should be a starting point (refined as necessary by additional experimental data) for 3D modeling of these complexes.

Energy minimization for the 3D model of R* was initially performed with additional constraints, keeping the distances between the ends of the TM helices at the same values they possessed in the TM helical bundle of R* to prevent possible disruption of the TM bundle. These additional restraints were removed during the second run of energy minimization. Then, representatives of all clusters of conformations for fragments 234−243 and 324−334 (see Methods) were readjusted to the final 3D model of R* displayed in Figure 4a. The model is deposited to the Protein Data Bank as entry 1OV0.

Figure 4b displays comparisons between R and R*; stereoviews in panels a and b of Figure 4 orient the TM bundle to show TM5−TM7 and TM1 (from left to right) in the front. It is noteworthy that the IC3 loop (the bottom left loop in the front) is oriented more inward toward the molecule core in R and more outward from the molecule core in R*, in agreement with suggestions deduced from experimental data (see ref *37* and references therein).

Validation of the 3D Model of R.* It is not surprising that the final 3D model of R* is consistent with the initial set of SDSL measurements on the R to R* transition as they were used to select the TM helix bundle as a basis for the model (though the correlation between the calculated and experimentally estimated $C_\alpha-C_\alpha$ distances becomes less pronounced, R^2 being 0.61 for the final 3D model instead of 0.94 seen in the starting TM packing option; see Figure 2). There are a variety of additional experimental constraints, both spectroscopic and mutational, not used in derivation of the 3D model of R* that can be used to validate the proposed model. It should be recognized, however, that experimental data also come with their own caveats, not the least of which is the fact that alternative three-dimensional structures may allow the correct orientation of the intracellular loops of R*, consistent with activation of its G-protein transducin. Only in the case of simultaneous multiple constraints (e.g., in the "straitjacket" mutant containing four disulfides; see below) can one be confident that constraints apply to a single active
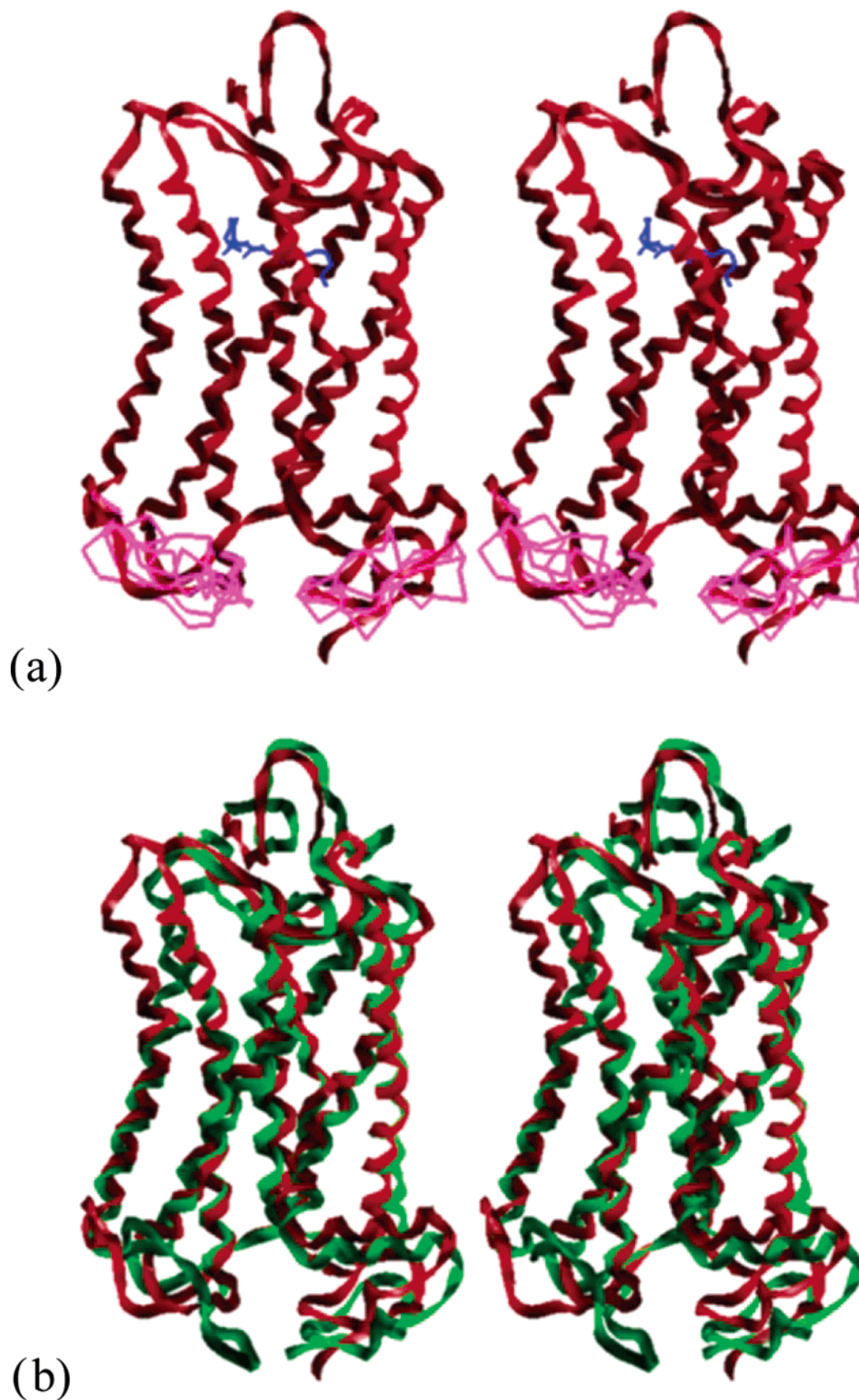
FIGURE 4: (a) Stereoview of the final 3D model of R* featuring the mean structures of residues 234−243 and 324−334 (shown as a red shadow ribbon) and representatives of all clusters for these fragments (shown as magenta lines). Retinal is shown in blue. (b) Stereoview of R (green shadow ribbon) overlapped with R* (red shadow ribbon).

conformational state for R*. Nevertheless, a viable 3D model for R* should be consistent with all experimental constraints,

and/or provide a plausible platform for rationalizing such constraints.
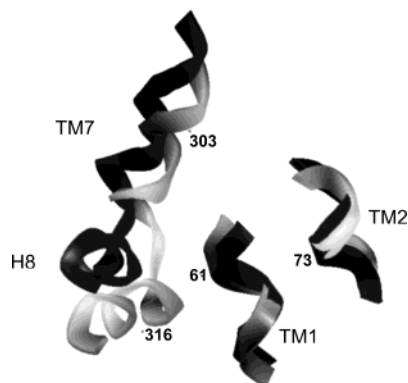
FIGURE 5: Changes between R (light gray) and R* (dark gray) in positions of the cytoplasmic portion of TM7 and H8 relative to cytoplasmic portions of TM1 and TM2. Positions mentioned in the text are labeled.

Distances between Spin-Labels. The recent results of two more studies mapping light-induced changes between SLs in position 65 and in fragment 306−319 (*15*) and between SLs in position 316 and in fragment 60−75 (*14*) were not included in selection of the model and, therefore, were used for validation.

As mentioned before, estimation of the exact distances between SLs is a complex and, sometimes, ambiguous procedure (*30*), especially in cases when light-induced changes in distances are small (*14, 15*). Accordingly, the cautious general conclusions of the recent experimental studies were that the experimentally observed changes suggest "a movement of the cytoplasmic portion of TM7 away from TM1 by 2−4 Å" (*15*) and "an outward displacement of TM2 relative to H8 by ≈3 Å" (*14*). Figure 5 depicts the cytoplasmic portions of TM1 (residues 60−65), TM2 (residues 71−76), TM7 (residues 302−310), and H8 (residues 311−319) for both R and R*, where the portions of TM1 and TM2 in R and R* are overlapped. One can see that the small light-induced changes in the relative spatial positions of the structural elements in question are in good agreement with the new experimental observations. For instance, the C_α−C_α distance between positions 61 in TM1 and 303 in the cytoplasmic portion of TM7 changed from 11.3 Å in R to 13.8 Å in R*, and the distance between positions 73 in TM2 and 316 in H8 changed from 11.8 Å in R to 13.4 Å in R*. These data are consistent with the proposed 3D model of R*.

Accessibilities of Cysteine Residues and Mobilities of Spin-Labels. The experimental technique of SDSL also allows estimation of the light-induced changes in the mobilities of SLs. The SL dynamics of spin-labeled rhodopsin are dominated by the local motions of the peptide backbone and the rotatable nitroxide side chains (*17*). It is difficult, therefore, to determine a single calculated parameter to represent such complex characteristics of the SLs. However, since the mobility of the SL is related to its accessibility to an external probe (*17*), it seemed reasonable to correlate the observed light-induced changes in mobilities of SLs in various positions of Rh with the changes in accessibilities of the corresponding C_α atoms to an external probe with a suitable radius. Generally, a decrease in SL mobility should correspond to the decreased accessibility of a certain position in a protein, whereas an increase in SL mobility should be accompanied by an increased accessibility.

The average calculated distance between the C_α atom of the Cys−SL construct and the oxygen atom of SL carrying the unpaired electron is ∼9 Å. Figure 6 shows differences in the accessible surface areas calculated for several positions in the C_α traces of R and the 3D model of R* with a probe with a radius of 4.5 Å. The readily available program GETAREA [http://www.scsb.utmb.edu/cgi-bin/get_a_form.tcl (*46*)] was used for the calculations. The positions selected for accessibility calculations were those with the experimentally observed light-induced changes in local mobilities as summarized in Figure 20 of ref *11*, and those with light-induced changes in accessibilities of the Cys residues to sulfhydryl-reacting reagents in the Rh mutants (*22, 23*).

Though approximate, the results depicted in Figure 6 demonstrate a rather high degree of correlation with the experimentally observed local mobilities of SLs (*11*); those in general agreement with experimental data are shown as white bars, and those in disagreement are shown as gray bars. For instance, the observed decreased mobilities of SLs in positions 70, 147, 149, 227, and 231 as well as the increased mobilities of SLs in positions 136, 140, 248, and 251 correspond to the calculated decreased or increased accessibilities in Figure 6. In fact, the only significant disagreement between the experimental and calculated estimations is in position 244 where the experimentally estimated SL mobility increases, whereas the calculated accessibility decreases. Also, the data depicted in Figure 6 correctly reflect the increased accessibility of the Cys residues in positions 68 and 72 (*22*) and, though slightly, the increased accessibility in positions 250 and 253 (*23*), but fail to reflect the decreased accessibility of the Cys residues in positions 69 and 73 (*22*).

Light-Induced Conformational Changes in Disulfide-Linked Mutants. Experimental data on effects of photoactivation for mutants of Rh containing the additional disulfide links (in addition to the native Cys110−Cys187 link present in the wild type as well as in all mutants of Rh and most GPCRs) divide the mutants into those permitting interaction with transducin after photoactivation, and those inhibiting this interaction. These qualitative experimental observations are even more difficult to interpret by a simple quantitative parameter than SL mobilities. C_α−C_α distances between positions in R, where the additional links were inserted, do not correlate with the ability of the mutant to interact with transducin. Figure 7 displays the distribution of such distances; one can see that permissive disulfides may be inserted in positions that are located as far away as ca. 20 Å in R [generally, in globular proteins, C_α−C_α distances for cystine disulfides are between 4.0 and 7.5 Å (*47, 48*)]. Therefore, the most immediate interpretation, that a distance between positions connected by a permissive disulfide link should decrease in R* compared to that in R and increase in the case of an inhibitory disulfide link, is too simplistic. For instance, of six disulfide-linked mutants, namely, those with disulfide links between positions 139 and 247−252, only one (139−252) was permissive, and all others are inhibitory (*13*). However, the same authors showed that distances between SLs in these positions increased upon photoactivation (139−248, 139−251, and 139−252), decreased (139−250), or remained unchanged (139−249) (*13*). The same reasoning is true for interpretation of the inhibitory Zn^{2+} bridge inserted between residues 138 and 251 (*29*); despite the C_α^{138}−C_α^{251} distance increasing on going from R to R*
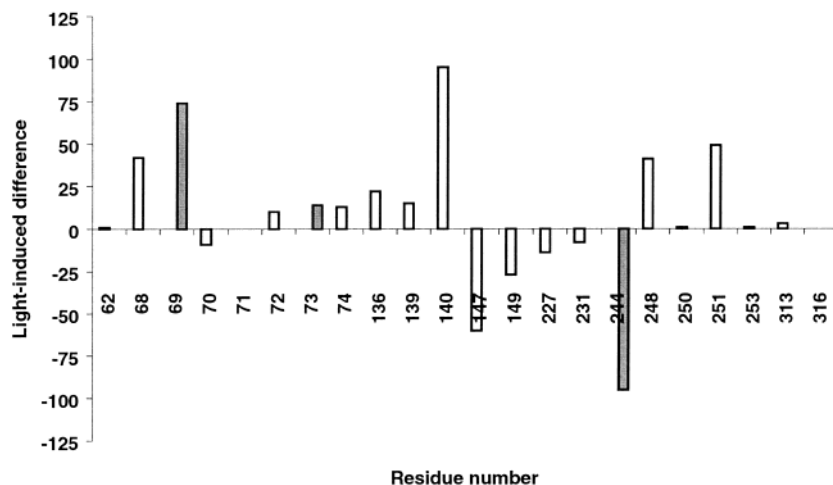
FIGURE 6: Differences in the surface accessible areas (in square angstroms) calculated for several positions in C_α traces of R and R* to the probe with a radius of 4.5 Å. Data for those residues from the model in general agreement with the experimental data are shown with white bars, and data for those in disagreement are shown with gray bars.
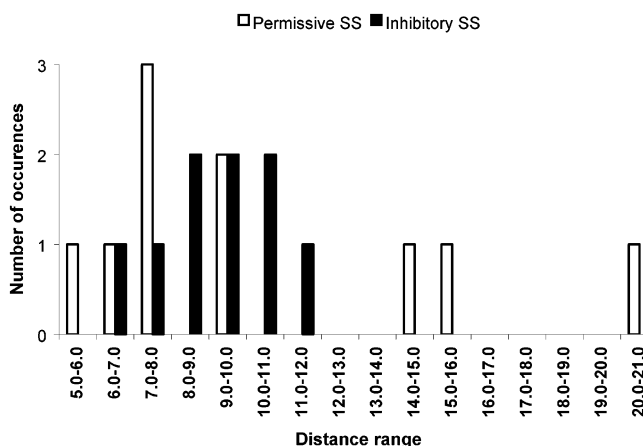

FIGURE 7: Distribution of $C_\alpha-C_\alpha$ distances for Cys substitutions in the known disulfide-linked mutants of Rh (10) based on the crystal structure of R.

(from 9.7 to 11.3 Å), the actual inhibitory effect may be due to steric hindrance of rotation of TM6.

Since no direct correlation between C_α distances for a disulfide with permissiveness was obvious (Figure 7), it was necessary to study the effect of each individual disulfide constraint imposed on the 3D model of R*. Toward this aim, energy minimization was performed, starting from the R* model, for three mutants, namely, Cys136−Cys222 and Cys136−Cys225 [the former inhibitive and the latter permissive (27)], as well as for the permissive straitjacketed mutant containing four disulfide links, namely, Cys110−Cys187 (native disulfide), Cys65−Cys316, Cys140−Cys225, and Cys204−Cys276 (25). The resulting structures of the three mutants are depicted in Figure 8a. The R* structure of the straitjacketed mutant has also been deposited in the Protein Data Bank as entry 1OV1.

Figure 8a clearly shows the main conformational difference between the permissive Cys136−Cys225 mutant and the inhibitory Cys136−Cys222 link. Namely, whereas in the permissive mutant the disulfide link does not disturb the 3D structure of R*, leaving the functionally important IC3 loop in the same spatial position, the inhibitory Cys136−Cys222 link forces TM5 (the far left helix in Figure 8a) as a whole to move down toward the cytoplasmic side of the molecule,

changing the orientation of the IC3 loop more inward to the molecule core, which is characteristic of the inactive R state (cf. Figure 4b). In straitjacketed Rh, also permissive, the spatial position of the IC3 loop is very close to that in R*. In other words, permission or inhibition of interaction with transducin in the disulfide-linked mutants of Rh can be rationalized by changes in the spatial orientation of the IC3 loop within the general frame of the current 3D model of the R* state.

In straitjacketed Rh, the endoplasmic end of TM6 is located closer to TM5 than in the 3D model of R* because of the slightly increased kink in TM6 (see Figure 8a). This is achieved by a concerted gradual changes of the dihedral angles of the TM6 peptide backbone, accompanied by rotation of the W265 side chain from a χ_1 of approximately 120° to a χ_1 of approximately −120° (data not shown). Interestingly, in our model, the W265 side chain experiences rotation from a χ_1 of approximately −60° to a χ_1 of approximately 120° with the transition from R to R*, in agreement with recent computational (49, 50) as well as experimental data (UV absorption spectroscopy) (51). On the other hand, our data support neither the suggestion of a pronounced additional kink of TM6 in Rh upon activation (11) nor the suggestion of straightening the kink in TM6 in some activated GPCRs (50).

The proposed 3D model of R* may also explain a somewhat surprising experimental observation that the disulfide-linked Cys245−Cys338 mutant permits activation of transducin (but not interaction with rhodopsin kinase) despite large $C_\alpha^{245}-C_\alpha^{338}$ distances in both R and the R* model (20.2 and 15.5 Å, respectively). Figure 8b presents the 3D structure of the mutant obtained by energy minimization starting from R*. One can see that the functionally important IC3 loop does not change its spatial position compared to R*; the conformational changes are limited to the spatial positions of helix 8 and the C-terminal fragment. Therefore, it is reasonable to assume that IC3 (residues 225−246) in R* is directly involved in the interaction with transducin [in good agreement with direct experimental data (36, 52); possible interaction of transducin with both IC2 and IC3 has been observed (53)], but not with rhodopsin kinase. Experimental data also suggest that the side chain carboxyl of E134 may
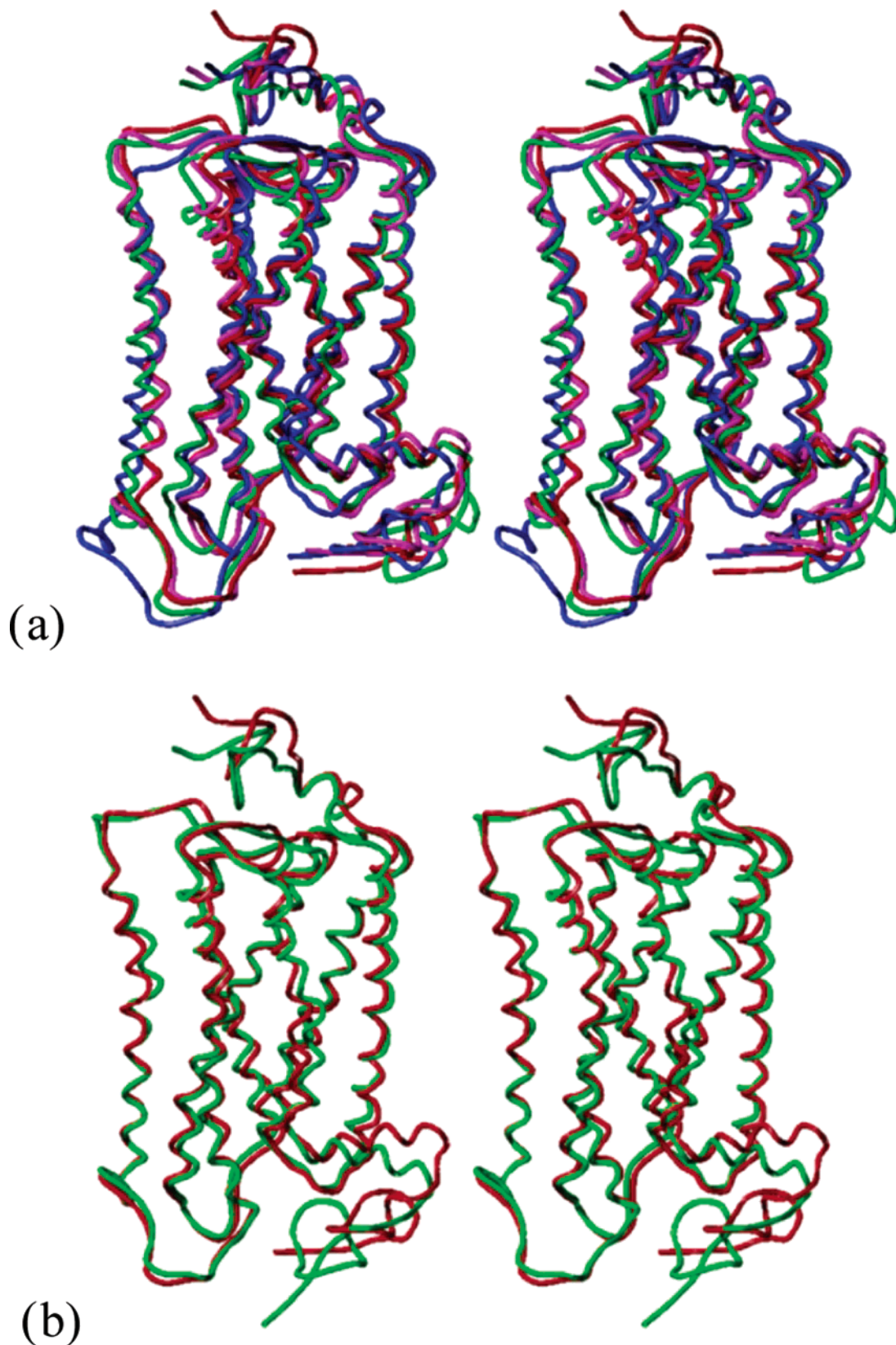
FIGURE 8: (a) Stereoview of the 3D model of R* (red tube) overlapped over TM regions with 3D models of the permissive Cys136−Cys225 mutants (magenta) and the straitjacketed Rh (green), and with that of the inhibitory Cys136−Cys222 mutant (blue). (b) Stereoview of 3D model of the Cys245−Cys338 mutant (green) overlapped with R* (red).

be protonated in R*, since mutations in this position influenced the interaction of Rh with transducin (see ref *12*). One possible result is disruption of the salt bridge between the side chains of E134 and neighboring R135 (*12*) that exists in the R state. Indeed, the $C_\delta^{134}−C_\zeta^{135}$ distance changed from 4.7 Å in R to 10.9 Å in our model of R*. Also, the recent mutational and disulfide cross-linking data suggested that eliminating interactions between Y306 located in the conserved NPXXY motif connecting TM7 and non-TM helix 8 and F313 (helix 8) allows the transition from R to R*,

whereas the Cys306−Cys313 disulfide link prevents formation of R* (*54*). According to our model, the C_α−C_α distance from Y306 to F313 increases from 7.3 Å in R to 7.9 Å in R*, which is not very significant; however, the C_α−C_α distance from N302 to C316 that characterizes the kink between TM7 and helix 8 changes from 16.6 to 18.3 Å, which, again, is in general agreement with the experimental data.

CONCLUDING REMARKS

This study presents a working 3D model of the photoactivated state (R*) of rhodopsin, a G-protein-coupled receptor. The model was obtained by a simple and straightforward computational procedure that suggested several possible options of energetically and sterically reasonable packing of the TM helices with *all-trans*-retinal. Limited experimental data (the light-induced changes in distances between spin-labels in positions 139 and 248−252) were used to select the most plausible 3D model from a set of energetically possible TM bundles. After addition of the non-TM fragments of the molecule, the entire R* structure was relaxed by energy minimization. The resulting 3D structure of R* was validated by confronting additional experimental data, both spectroscopic and mutational, not used in the process of model building.

The proposed model of R* shows good agreement with available experimental data related to light-induced conformational changes in Rh, such as changes in distances between spin-labels or changes in their relative local mobilities and in accessibilities of the Cys residues in Rh mutants. The model was also instrumental in the interpretation of experimental data on disulfide-linked mutants of Rh emphasizing, in agreement with experiment observations, the functional importance of the proper spatial position of the IC3 loop for interaction with transducin. Therefore, we believe that this model can serve as an excellent starting point for rationalizing future experimental measurements on R* as well as for the rational design of mutants of Rh selectively interacting with transducin and/or rhodopsin kinase.

The 3D model of R* can be used also as a template in studies of the other GPCRs. However, any direct use of the model in this case should be limited by certain caveats, most importantly by the assumption that the non-TM fragments of the molecule may retain similar conformations in both R and R*. While quite reasonable for Rh, where the internal chromophore retinal remains inside the molecule during the transfer from R to R* and, therefore, the extracellular fragments of the molecule do not necessarily change conformations to accommodate its binding or unbinding, this assumption may be not true for many other GPCRs, especially for those interacting with large extracellular ligands, such as peptides. On the other hand, the TM region of the proposed model of R* may be a more reliable template for the corresponding regions of the other activated GPCRs (see also the recent review in ref *55*). Most importantly, this study demonstrates that 3D models of the activated states of GPCRs can evolve by iteratively selecting those closest to the available experimental data from 3D structures suggested by energy calculation/modeling followed by subsequent refinement as additional experimental data become available.

REFERENCES

1. Borhan, B., Souto, M. L., Imai, H., Shichida, Y., and Nakanishi, K. (2000) *Science 288*, 2209−2212.
2. Gether, U. (2000) *Endocr. Rev. 21*, 90−113.
3. Drews, J. (2000) *Science 287*, 1960−1964.
4. 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., and Miyano, M. (2000) *Science 289*, 739−745.
5. Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K., and Stenkamp, R. E. (2001) *Biochemistry 40*, 7761−7772.
6. Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E. M., and Shichida, Y. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 5982−5987.
7. Sakmar, T. P. (2002) *Curr. Opin. Cell Biol. 14*, 189−195.
8. Okada, T., and Palczewski, K. (2001) *Curr. Opin. Struct. Biol. 11*, 420−426.
9. Ballesteros, J. A., Shi, L., and Javitch, J. A. (2001) *Mol. Pharmacol. 60*, 1−19.
10. Meng, E. C., and Bourne, H. R. (2001) *Trends Pharmacol. Sci. 22*, 587−593.
11. Hubbell, W. L., Altenbach, C., and Khorana, H. G. (2003) *Adv. Protein Chem. 63*, 243−290.
12. Sakmar, T. P., Menon, S. T., Marin, E. P., and Awad, E. S. (2002) *Annu. Rev. Biophys. Biomol. Struct. 31*, 443−484.
13. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L., and Khorana, H. G. (1996) *Science 274*, 768−770.
14. Altenbach, C., Klein-Seetharaman, J., Cai, K., Khorana, H. G., and Hubbell, W. L. (2001) *Biochemistry 40*, 15493−15500.
15. Altenbach, C., Cai, K., Klein-Seetharaman, J., Khorana, H. G., and Hubbell, W. L. (2001) *Biochemistry 40*, 15483−15492.
16. Faranbakhsh, Z. T., Ridge, K. D., Khorana, H. G., and Hubbel, W. L. (1995) *Biochemistry 34*, 8812−8819.
17. Altenbach, C., Yang, K., Farrens, D. L., Farahbakhsh, Z. T., Khorana, H. G., and Hubbell, W. L. (1996) *Biochemistry 35*, 12470−12478.
18. Altenbach, C., Klein-Seetharaman, J., Hwa, J., Khorana, H. G., and Hubbell, W. L. (1999) *Biochemistry 38*, 7945−7949.
19. Altenbach, C., Cai, K., Khorana, H. G., and Hubbell, W. L. (1999) *Biochemistry 38*, 7931−7937.
20. Yang, K., Farrens, D. L., Altenbach, C., Farahbakhsh, Z. T., Hubbell, W. L., and Khorana, H. G. (1996) *Biochemistry 35*, 14040−14046.
21. Resek, J. F., Faranbakhsh, Z. T., Hubbel, W. L., and Khorana, H. G. (1993) *Biochemistry 32*, 12025−12032.
22. Klein-Seetharaman, J., Hwa, J., Cai, K., Altenbach, C., Hubbell, W. L., and Khorana, H. G. (1999) *Biochemistry 38*, 7938−7944.
23. Dunham, T. D., and Farrens, D. L. (1999) *J. Biol. Chem. 274*, 1683−1690.
24. Struthers, M., Yu, H., Kono, M., and Oprian, D. D. (1999) *Biochemistry 38*, 6597−6603.
25. Struthers, M., Yu, H., and Oprian, D. D. (2000) *Biochemistry 39*, 7938−7942.
26. Struthers, M., and Oprian, D. D. (2000) *Methods Enzymol. 315*, 130−143.
27. Yu, H., and Oprian, D. D. (1999) *Biochemistry 38*, 12033−12040.
28. Cai, K., Klein-Seetharaman, J., Hwa, J., Hubbell, W. L., and Khorana, H. G. (1999) *Biochemistry 38*, 12893−12898.
29. Sheikh, S. P., Zvyaga, T. A., Lichtarge, O., Sakmar, T. P., and Bourne, H. R. (1996) *Nature 383*, 347−350.
30. Hubbell, W. L., Cafiso, D. S., and Altenbach, C. (2000) *Nat. Struct. Biol. 7*, 735−739.
31. Choi, G., Landin, J., Galan, J. F., Birge, R. R., Albert, A. D., and Yeagle, P. L. (2002) *Biochemistry 41*, 7318−7324.
32. Yeagle, P. L., and Albert, A. D. (2003) *Biochemistry 42*, 1365−1368.
33. Ciarkowski, J., Drabik, P., Gieldon, A., Kazmierkewicz, R., and Slusarz, R. (2001) in *Peptides: The Wave of the Future.*

Proceedings of the Second International and the Seventeenth American Peptide Symposium, pp 896−897, American Peptide Society, San Diego.

34. Rohrig, U. F., Guidoni, L., and Rothlisberger, U. (2002) *Biochemistry 41*, 10799−10809.
35. Saam, J., Tajkhorshid, E., Hayashi, S., and Schulten, K. (2002) *Biophys. J. 83*, 3097−3112.
36. Acharya, S., Saad, Y., and Karnik, S. S. (1997) *J. Biol. Chem. 272*, 6519−6524.
37. Arimoto, R., Kisselev, O. G., Makara, G. M., and Marshall, G. R. (2001) *Biophys. J. 81*, 3285−3293.
38. Nikiforovich, G. V., Galaktionov, S., Balodis, J., and Marshall, G. R. (2001) *Acta Biochim. Pol. 48*, 53−64.
39. Galaktionov, S., Nikiforovich, G. V., and Marshall, G. R. (2001) *Biopolymers 60*, 153−168.
40. Visiers, I., Ballesteros, J. A., and Weinstein, H. (2002) *Methods Enzymol. 343*, 329−371.
41. Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., and Goddard, W. A., III (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 12622−12627.
42. Nikiforovich, G. V., Galaktionov, S., Balodis, J., and Marshall, G. R. (2001) *Acta Biochim. Pol. 48*, 53−64.
43. Dunfield, L. G., Burgess, A. W., and Scheraga, H. A. (1978) *J. Phys. Chem. 82*, 2609−2616.
44. Nemethy, G., Pottle, M. S., and Scheraga, H. A. (1983) *J. Phys. Chem. 87*, 1883−1887.
45. Nikiforovich, G. V., Hruby, V. J., Prakash, O., and Gehrig, C. A. (1991) *Biopolymers 31*, 941−955.
46. Fraczkewicz, R., and Braun, W. (1998) *J. Comput. Chem. 19*, 319−333.
47. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−336.
48. Goldblum, A., and Rein, R. (1987) *Int. J. Pept. Protein Res. 30*, 784−793.
49. Shi, L., Liapakis, G., Xu, R., Guarneri, F., Ballesteros, J. A., and Javitch, J. A. (2002) *J. Biol. Chem. 277*, 40989−40996.
50. Singh, R., Hurst, D. P., Barnett-Norris, J., Lynch, D. L., Reggio, P. H., and Guarnieri, F. (2002) *J. Pept. Res. 60*, 357−370.
51. Lin, S. W., and Sakmar, T. P. (1996) *Biochemistry 35*, 11149−11159.
52. Cai, K., Itoh, Y., and Khorana, F. C. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 4877−4882.
53. Itoh, Y., Cai, K., and Khorana, H. G. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 4883−4887.
54. Fritze, O., Filipek, S., Kuksa, V., Palczewski, K., Hofmann, K. P., and Ernst, O. P. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 2290−2295.
55. Mirzadegan, T., Benko, G., Filipek, S., and Palczewski, K. (2003) *Biochemistry 42*, 2769−2767.