Structure of Bovine Rhodopsin in a Trigonal Crystal Form

Jade Li*, Patricia C. Edwards1, Manfred Burghammer2, Claudio Villa1 and Gebhard F. X. Schertler1

1Medical Research Council Laboratory of Molecular Biology Hills Road Cambridge CB2 2QH, UK
2European Synchrotron Radiation Facility, BP 220 8043 Grenoble, France

We have determined the structure of bovine rhodopsin at 2.65 Å resolution using untwinned native crystals in the space group P31, by molecular replacement from the 2.8 Å model (1F88) solved in space group P41. The new structure reveals mechanistically important details unresolved previously, which are considered in the membrane context by docking the structure into a cryo-electron microscopy map of 2D crystals.

Kinks in the transmembrane helices facilitate inter-helical polar interactions. Ordered water molecules extend the hydrogen bonding networks, linking Trp265 in the retinal binding pocket to the NPxxY motif near the cytoplasmic boundary, and the Glu113 counterion for the protonated Schiff base to the extracellular surface. Glu113 forms a complex with a water molecule hydrogen bonded between its main chain and side-chain oxygen atoms. This can be expected to stabilise the salt-bridge with the protonated Schiff base linking the 11-cis-retinal to Lys296.

The cytoplasmic ends of helices H5 and H6 have been extended by one turn. The G-protein interaction sites mapped to the cytoplasmic ends of H5 and H6 and a spiral extension of H5 are elevated above the bilayer. There is a surface cavity next to the conserved Glu134-Arg135 ion pair. The cytoplasmic loops have the highest temperature factors in the structure, indicative of their flexibility when not interacting with G protein or regulatory proteins. An ordered detergent molecule is seen wrapped around the kink in H6, stabilising the structure around the potential hinge in H6.

These findings provide further explanation for the stability of the dark state structure. They support a mechanism for the activation, initiated by photo-isomerisation of the chromophore to its all-trans form, that involves pivoting movements of kinked helices, which, while maintaining hydrophobic contacts in the membrane interior, can be coupled to amplified translation of the helix ends near the membrane surfaces.

Keywords: G protein-coupled receptor; G protein activation; ligand binding pocket; membrane protein structure; visual pigments

Abbreviations used: C8E4, n-octyltetraoxyethylene; C1, cytoplasmic loop connecting helices 1 and 2; C2, cytoplasmic loop connecting helices 3 and 4; C3, cytoplasmic loop connecting helices 5 and 6; cGMP, 3',5'-cyclic guanosine monophosphate; cryo-EM, electron cryomicroscopy; E1, extracellular loop connecting helices 2 and 3; E2, extracellular loop connecting helices 4 and 5; E3, extracellular loop connecting helices 6 and 7; EM, electron microscopy; EMTS, ethyl mercurithiosalicylate; FTIR, Fourier transform infrared spectroscopy; GPCR, G protein-coupled receptor; G-protein, heterotrimeric guanine nucleotide binding protein; Gt, transducin; Gtα, transducin α-subunit; Gtα340-350, C-terminal peptide from transducin α-subunit; LDAO, N,N-dimethyldecylamine-N-oxide; M1 and MII, metarhodopsin I and II; NCS, non-crystallographic symmetry; PDE, phosphodiesterase.

© 2004 Elsevier Ltd. All rights reserved.
**Introduction**

G protein-coupled receptors (GPCRs) constitute the largest superfamily of transmembrane signalling proteins in the eukaryotic kingdom. They share a common core structure comprising seven transmembrane helices, as implied by a conserved pattern of amino acids at key positions along each of the seven hydrophobic sequence segments. These proteins capture external signals, including light, odorants, hormones and neurotransmitters, and transmit the stimuli across the plasma membrane by selectively activating heterotrimeric guanine nucleotide binding proteins (G proteins) on the cytoplasmic surface. Activation of G protein amplifies the signal and leads to activation of effector enzymes or molecules, which elicit the cell's signalling response indirectly by altering second messenger concentrations or by direct control of ion channel activities. Thus, the GPCRs play a central role in regulating many physiological processes. They are consequently major targets for drug design.

Rhodopsin, the photoreceptor protein in retina rod cells, is a prototypical GPCR. It contains a light-sensitive ligand, the 11-cis-retinal chromophore, bound covalently to the apoprotein opsin via a protonated Schiff base with Lys296 on helix 7. Absorption of a photon at a wavelength of about 500 nm isomerises the retinal to all-trans within picoseconds and with a quantum efficiency of 0.67. This event initiates the formation of a series of photointermediates with conformational changes in the opsin. The biochemically active conformation R* is attained within one millisecond at physiological temperature. It is spectroscopically identified with the metarhodopsin II (MII) intermediate, which has an absorption maximum at 380 nm due to deprotonation of the Schiff base of all-trans-retinal. R* binds and activates the heterotrimeric G-protein transducin (Gt) at the cytoplasmic surface, to catalyse GDP/GTP exchange on the Gt α-subunit and dissociation of the Gt heterotrimer. The GTP-bound α-subunit then activates the effector enzyme, phosphodiesterase (PDE). Hydrolysis of cyclic guanine monophosphate (cGMP), the second messenger, by PDE leads to closure of the cGMP-gated cation channel in the plasma membrane, causing hyperpolarisation and initiation of nerve impulse in the retina.

In contrast to the efficient and rapid photoisomerisation and activation, the rate of thermal isomerisation is very low, about one per 400 years per rhodopsin molecule at 37 °C. The low thermal rate is coupled with large amplification via transducin activation, underpins the sensitivity of single photon detection operating in dim light vision.

An atomic structure of rhodopsin provides a model for all the visual pigments as well as the majority of the GPCRs. Crystallisation in two-dimensional (2D) lattices with endogenous membrane lipids, followed by electron cryo-microscopy (cryo-EM) at resolution limits up to 5 Å in the membrane plane and 13.5 Å normal to it, have produced images of bovine, frog and squid rhodopsins in a membrane-like environment, which are in essence similar. Using packing constraints derived by extensive sequence comparisons across the GPCR superfamily, the density peaks in cryo-EM maps were assigned to hydrophobic sequences, leading eventually to a Cα model for the seven transmembrane helices in the rhodopsin-like GPCR family. Remarkably, this model came within 2.3 Å rms deviation of Cα coordinates determined by X-ray crystallography subsequently and it provided a framework for mutagenesis and biophysical studies in the absence of an atomic structure.

To determine the atomic structure of rhodopsin, we have obtained untwinned three-dimensional crystals of bovine rhodopsin in the trigonal space group P3₁ that diffract X-rays to 2.65 Å resolution, and prepared an ethylmercury derivative that showed anomalous scattering. However, significant non-isomorphism prevented experimental phasing. While this work was in progress, a structure of bovine rhodopsin in a tetragonal space group at 2.8 Å resolution was reported. Using these coordinates as search model, we have determined the structure of bovine rhodopsin to 2.65 Å resolution in the P3₁ lattice by molecular replacement and molecular replacement followed by multiple crystal averaging showed anomalous scattering. However, significant non-isomorphism prevented experimental phasing. While this work was in progress, a structure of bovine rhodopsin in the trigonal crystal form, and compare it with previous structural reports of the tetragonal crystal form. In addition, we relate the crystal structures to the membrane environment by docking them into a cryo-EM map of 2D crystals. Structural implications for the stability of the "dark" state and the potential for light-induced conformational change, leading to transducin activation on the cytoplasmic surface, are discussed.

**Results and Discussion**

**Structure determination**

Bovine rhodopsin was crystallised in the trigonal space group P3₁ with two protein molecules per asymmetric unit, from a detergent mixture of ChE₄ and LDAO as described in the accompanying paper. Data sets were obtained from native and mercury-derivatised crystals prepared and frozen under dim red light; however, they showed significant non-isomorphism with one another. Data collection and refinement statistics are given in Table 1.

Using the protein part only of the rhodopsin coordinates solved in a tetragonal space group (PDB code 1F88) at 2.8 Å resolution as search model, molecular replacement solutions were calculated in parallel for three sets of amplitudes (Native-2, EMTS-1 and EMTS-2). The resulting 2Fo−Fc maps showed no density in three cytoplasmic regions of 1F88 coordinates, and the
Table 1. Data collection and final refinement statistics: space group $P_3_1$

### A. Data collection

<table>
<thead>
<tr>
<th></th>
<th>Native-2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>EMTS-1</th>
<th>EMTS-2</th>
<th>Native-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X-ray source</strong></td>
<td>ESRF ID13</td>
<td>ESRF ID13</td>
<td>ESRF ID14-4</td>
<td>ESRF ID13</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.782</td>
<td>1.005</td>
<td>1.0065</td>
<td>0.782</td>
</tr>
<tr>
<td><strong>Unit cell a, c (Å)</strong></td>
<td>104.2, 77.1</td>
<td>113.9, 78.4</td>
<td>109.3, 77.6</td>
<td>103.8, 76.6</td>
</tr>
<tr>
<td><strong>No. of crystals</strong></td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><strong>Mosaicity (deg.)</strong></td>
<td>1.0</td>
<td>1.1</td>
<td>1.1</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>Twin fraction</strong></td>
<td>0</td>
<td>0</td>
<td>0.31</td>
<td>0</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>3.2 (3.37–3.20)</td>
<td>3.6 (3.71–3.60)</td>
<td>3.4 (3.45–3.40)</td>
<td>2.65 (2.79–2.65)</td>
</tr>
<tr>
<td><strong>R&lt;sub&gt;merge&lt;/sub&gt;</strong></td>
<td>0.127 (0.322)</td>
<td>0.169 (0.464)</td>
<td>0.139 (0.426)</td>
<td>0.119 (0.434)</td>
</tr>
<tr>
<td><strong>I/&lt;I&gt;</strong></td>
<td>7.5 (2.3)</td>
<td>7.6 (2.2)</td>
<td>17.2 (4.7)</td>
<td>11.0 (1.4)</td>
</tr>
<tr>
<td><strong>Unique reflections</strong></td>
<td>15087</td>
<td>11481</td>
<td>13320</td>
<td>26026</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>0.982 (0.894)</td>
<td>0.868 (0.180)</td>
<td>0.834 (0.172)</td>
<td>0.970 (0.861)</td>
</tr>
<tr>
<td><strong>Multiplicity</strong></td>
<td>3.1 (1.6)</td>
<td>5.3 (2.9)</td>
<td>11.3 (9.5)</td>
<td>4.4 (1.6)</td>
</tr>
<tr>
<td><strong>Wilson B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</strong></td>
<td>84.6</td>
<td>59.7</td>
<td>54.2</td>
<td>58.2</td>
</tr>
</tbody>
</table>

### B. Refinement against Native-3

| Reflections in working set | 24704 (2165) | Protein chains |
| Reflections in test set    | 1322 (130)   | Protein residues |
| Resolution range (Å)       | 46–2.65 (2.74–2.65) | Palmitoyl chains |
| $R_{cryst}$                | 0.202 (0.312) | N-linked carbohydrate chains |
| $R_{free}$                 | 0.235 (0.315) | Carbohydrate residues |
| Luzzati coordinate error (5.0–2.65 Å) | 0.31 Å | LDAO |
| SigmaA coordinate error (5.0–2.65 Å) | 0.42 Å | CREA |
| rms-deviation from ideal geometry | Bond lengths (Å) | Water |
| Bond angles (deg.)         | 1.293 | Ions |
| Dihedral angles (deg.)     | 18.7 | 2 |
| Improper rotations (deg.)  | 0.876 | 2 |
| Ramachandran plot          | Residues in most favoured regions (%) | 90.6 |
|                           | Residues in additional allowed regions (%) | 7.1 |
|                           | Residues in generously allowed regions (%) | 2.4 |
|                           | Residues in disallowed regions (%) | 0 |
| Average B-factor (Å<sup>2</sup>) | 56.0 |
| $B$ rmsd for bonded main-chain atoms (Å<sup>2</sup>) | 1.501 |
| $B$ rmsd for bonded side-chain atoms (Å<sup>2</sup>) | 1.996 |
| $B$ rmsd for angle main-chain atoms (Å<sup>2</sup>) | 2.624 |
| $B$ rmsd for angle side-chain atoms (Å<sup>2</sup>) | 3.134 |

Values in parentheses apply to the outer resolution shell.

<sup>a</sup> Names of data sets conform to Table I of the accompanying paper. Native-1 is not referred to here, because it was not used in the structure determination.

<sup>b</sup> $R_{merge} = \frac{\sum_{hkl} \sum_{i} |I_i - \langle I \rangle|}{\sum_{hkl} \sum_{i} I_i}$, where $I_i$ is the intensity of an individual reflection and $<I>$ is the mean intensity obtained from multiple observations of symmetry-related reflections taken from one or more crystals.

<sup>c</sup> $R_{cryst} = \frac{\sum_{hkl} |F_o(hkl)-F_c(hkl)|}{\sum_{hkl} F_o(hkl)}$, calculated for reflections in the working set.

<sup>d</sup> $R_{free} = \frac{\sum_{hkl} |F_o(hkl)-F_c(hkl)|}{\sum_{hkl} F_o(hkl)}$, calculated for a random 5% of the data placed in the test set.

corresponding $F_o - F_c$ maps showed negative peaks in the same regions. These regions were: the C2 loop connecting transmembrane helices H3 and H4; the C3 loop connecting H5 and H6; and the C-terminal tail of residues 334–348. By contrast, density was present in the 2F<sub>r</sub> – F<sub>c</sub> maps over positions of the retinal chromophore, even though, as a control, we had omitted the retinal from the 1F88 coordinates used for molecular replacement. No negative peaks were seen in the retinal-binding pocket. Therefore our amplitude data disagreed with the starting model in the cytoplasmic regions. Exploiting the large non-isomorphism, we carried out cross-crystal averaging among these data sets. The averaged map for Native-2 showed densities for the C2 and C3 loops at positions different from the starting model, enabling them to be rebuilt and the cytoplasmic end of helix H6 to be extended from residue 247 in the 1F88 coordinates to residue 244. The middle section of the C3 loop (residues 233–239 in molecule A and 233–237 in molecule B) and the C terminus still had no interpretable densities, so they were left out of the model. The Schiff base linked 11-cis-retinal chromophore was built from the small molecule crystal structures by adjusting torsion angles along the polyene according to the averaged density.

A new native data set (Native-3) was then collected from four untwinned crystals to 2.65 Å resolution, and used to refine the atomic model, using CNS under 2-fold non-crystallographic symmetry restraints with simulated annealing and restrained individual B-factors. Bound water molecules were located using the water_pick task of CNS and confirmed in maps calculated using EDEN. Tightly bound phospholipids and
detergent molecules were added. Including the bound water and amphiphile molecules in the refinement further improved the map, so that the cytoplasmic ends of both H5 and H6 were extended and the C3 loop was completed. The most significant differences between our structure and all the published coordinates from the P41 crystal are found in this region.

Quality of the structure

The final crystallographic R-factor is 0.202 with a free-R-factor of 0.235, and there is close agreement with ideal geometry (Table 1B). The Ramachandran plot contains 90.6% of the protein residues in the most favoured regions, compared with values of 80.9–82.5% for the published rhodopsin structures in the tetragonal crystal form. The A-weighted\(^4\) \(2F_o - F_c\) map showed an average real-space correlation of 0.95 for all residues including the non-protein residues. These statistics demonstrate that the structure was well determined, and bias toward the starting model was minimal.

The refined model (Table 1B) comprises amino acid residues 1–326 in both molecules, compared with the complete sequence of 348 residues. Post-translational modifications have been defined on both molecules, including N-acetylation of Met1, N-glycosylation of Asn2 and Asn15, palmitoylation of Cys322 and Cys323, and the Schiff base linkage of 11-cis-retinal to Lys296. In addition, there are 20 ordered water molecules bound to each protein chain and a number of tightly bound lipid and detergent molecules. The covalent structure is shown as a ribbon diagram in Figure 1(a). The residue ranges of the seven transmembrane helices are: H1, 34–64; H2, 71–100; H3, 106–140; H4, 150–173; H5, 200–230; H6, 241–276 and H7, 286–309. Residues in the peripheral helix H8 are 311–321, and those making up the B-strands are: B1, 286–309. Portions of the transmembrane helices, H4 between residues 170 and 173 and H7 between residues 296 and 299, are in the \(\beta\)-conformation.

Comparison with the tetragonal crystal form

Packing differences

Both the P3\(_1\) and P4\(_1\) crystal forms contain two rhodopsin molecules per asymmetric unit. In the P4\(_1\) form, they are related by a non-crystallographic rotation of 172.5° about an axis nearest to a and are oriented obliquely from the P4\(_1\) axis. They make contacts in the middle of helix 1 but mainly over the hydrophilic extra-membrane segments. In particular, the C-terminus of molecule A from the cytoplasmic domain forms a contact with the extracellular domain of molecule A in another asymmetric unit.

In the P3\(_1\) form, the non-crystallographic symmetry (NCS) is a nearly perfect 2-fold about an axis parallel with the \(\alpha\) axis and passing through the midplane of molecule A. Both helix bundles are tilted by 108° from c and make antiparallel contacts along the entire length of helix 5. The antiparallel dimers are stacked according to the \(\alpha\) screw symmetry, forming two protein columns in the unit cell, which are centred on two of the three 3-fold positions of the trigonal lattice, and leaving a solvent channel centred on the remaining 3-fold position. Between layers of stacked dimers the contacts involve mainly the sides of helices 1 and 4. The other helices participate in crystal contacts to a lesser extent or indirectly via interleaved palmitoylate chains or the C\(_E\)E molecule (see subsection titled Bound lipid and detergent molecules). Thus, uncommonly among crystals of membrane proteins, the inter-molecular contacts in the P3\(_1\) form are predominantly hydrophobic. The hydrophilic extra-membrane segments point into the solvent channel and are not responsible for ordered crystal contacts (Figure 1(b)). Consequently, the P3\(_1\) crystal packing has retained an amphipathic molecular environment characteristic of the native membrane, permitting the hydrophilic segments to display conformations present in a membrane environment unhindered by packing interactions.

Coordinate differences between NCS copies and between crystal forms

The 1I9H coordinates\(^9\) at 2.6 Å resolution are used to represent the P4\(_1\) form. Here, the rms C\(^\beta\) distance between NCS copies is 0.470 Å for 301 matched residues, and the rms B-factor difference is 9.530 Å\(^2\) with molecule A being the more ordered copy. In the P3\(_1\) form with closer NCS, the two molecules show rms differences of only 0.13 Å and 2.216 Å\(^2\) for 326 matched residues. Therefore, molecules A from the 1I9H coordinates and the present structure are used to compare the crystal forms.

Figure 2(a) shows the coordinate difference per residue as a function of sequence position, between NCS copies in each crystal form and between the molecules A of different crystal forms. Both crystal forms showed significant main chain NCS differences, above 1 Å, in the C2 and C3 loops and the C-terminal tail following helix 8. These regions also have the highest B-factors and disorder in each crystal form (Figure 2(b)). Part of the C3 loop is unresolved in both molecules of the P4\(_1\) form. The C-terminal tail in the P3\(_1\) form is located in the solvent channel and disordered, while in the P4\(_1\) form it is partially resolved only in molecule A, where it is stabilised through contacts with the extracellular domain of the neighbouring molecule. Finding significant NCS differences in the same surface segments indicates an intrinsic conformational variability or flexibility of those segments, and the high B-factors of the segments support this. These segments are known to contain residues that interact with transducin (Gt) and the regulatory proteins rhodopsin kinase and arrestin\(^{22,31}\), their...
flexibility in the absence of the binding partners may have a functional role.

If main-chain atoms in the molecules A of the two crystal forms are superimposed by aligning the more ordered regions selected by a B-factor cut-off of 90 Å², then the rms coordinate difference is 0.36 Å, which is slightly less than the mean coordinate error for all atoms in either structure. Therefore, the two structures are in good agreement in the ordered core of the molecule, which comprised residues 1–137, 154–223 and 250–321, and excluded only the neighbourhood of the C2 and C3 loops, and the C-terminal tail. The backbone differences are concentrated in the same regions shown by the NCS comparisons to be intrinsically more flexible. Figure 2(a) shows that, for the C2 loop, the difference between crystal forms is on the same scale as the NCS difference within the P4₁ form, so they are accountable by local flexibility and packing contacts. However, around the C3 loop the differences between the two crystal forms are much larger than the NCS differences in either form. Therefore, they indicate statistically significant disagreement.

**Location of major coordinate differences between crystal forms**

Figure 2(b) shows the Cα trace of the molecules A from the two crystal forms side-by-side, coloured
Figure 2. (legend opposite)
by the mean B-factor in each residue. The colour scheme shows that the new structure from the P31 crystal form has the lower B-factors everywhere except in the C3 loop, which is absent in the other structure, and the major differences lie in the regions of higher B-factors.

The C2 loop has a similar “L-shape” in both structures but with a different orientation relative to the helix bundle. The difference may be described as a hinge movement about the junctions with the cytoplasmic ends of H3 and H4.

The C3 loop and the cytoplasmic ends of H5 and H6 show the most striking difference between the two structures. In molecule A of the P41 form, H5 terminates at residue 226, following which the polypeptide chain re-enters the bilayer and appears disordered between residues 236 and 240, while H6 terminates at residue 244. In the P31 form, H5 extends to residue 230, then the C trace continues in a helix-like spiral path away from the membrane to residue 236, where it changes direction to run parallel with the membrane and joins the cytoplasmic end of H6 at residue 241. The cytoplasmic tip of H6 is slightly bent at residue 244 towards H5.

The C-terminal tail between residues 334–348 in the P41 structure adopts an extended conformation supported by contacts with a neighbouring molecule. In the P31 structure this segment is disordered, except for a di-peptide suggested by the density in contact with the cytoplasmic loop 1 (loop C1). The disorder can be expected, since the fragment, without disulphide bridges, is too small to form an independently folded domain, and it is located in a solvent channel in this crystal form.

The structure of the cytoplasmic domain of the P31 form is discussed at greater length in the context of explaining the surface accessibility and mobility of the polypeptide chain re-enters the bilayer ~nd appears to form an independently folded domain, and it is located in a solvent channel in this crystal form.

The structure of the cytoplasmic domain of the P31 form is discussed at greater length in the context of explaining the surface accessibility and mobility of the protein. The eight N-terminal residues of the P31 structure, except for H4, which terminates within the hydrophobic zone, are contained within the hydrophobic zone. They form the external domain. However only the residues 1–5 and 13–21, which carry the N-linked carbohydrate chains on Asn2 and Asn15, reach into the intra-discal space.

Placing the X-ray structure in the membrane context

Before coordinates from the tetragonal 3D crystals became available, we were able to determine the orientation of rhodopsin molecules in the P31 crystals by rotation search at 5 Å resolution,16 using as search model the density for a single molecule masked from a cryo-EM map of 2D crystals10 (see Methods). Conversely, by docking the refined X-ray coordinates into this cryo-EM map,10 we placed the atomic structure in the membrane context of the 2D crystal. The height of rhodopsin relative to the bilayer was therefore ascertained, since the p2z symmetry of those 2D crystals requires symmetrical insertion of molecules from both faces of the bilayer and therefore the plane defined by ε = 0 is equivalent to the midplane of the bilayer.

Figure 3(a) shows the mean residue height in the bilayer as a function of sequence position, overlaid by the mean residue B-factors of the refined P31 structure. The sequence ranges of the α and β secondary structure elements are shown as vertical bars. Small-angle X-ray and neutron scattering have shown that phosphate peaks in the retinal disc membranes are separated by 40 Å across the bilayer,40,41 and lipid head groups in model membranes occupy a layer about a 10 Å thick centred on the phosphate peaks.42 Therefore in Figure 3 the hydrophobic zone extends from approximately −15 Å to +15 Å, and the lipid head group layers extend from there to approximately −25 Å and +25 Å. On this scale, the 33 N-terminal residues preceding H1 fall entirely outside the hydrophobic zone. They form the extracellular domain. However only the residues 1–5 and 13–21, which carry the N-linked carbohydrate chains on Asn2 and Asn15, reach into the intra-discal space.

On the extracellular face all transmembrane helices terminate in the lipid head group layer, except H4, which terminates within the hydrophobic zone. The strand β3 lies at the interface between hydrophobic zone and head group layer, while β4 is contained within the hydrophobic zone. On the cytoplasmic face, helices H1, H2, H4 and H7

Figure 2. Comparison between the P31 and P41 crystal forms. (a) Coordinate difference as a function of sequence position. The main chain and side-chain differences between the molecules A from different crystal forms are superimposed on the main chain difference between NCS copies within each crystal form. Significant differences exist only around the C2 loop (residues 141–149) connecting H3 and H4 and the C3 loop (residues 231–240) connecting H5 and H6, and in the C-terminal tail following H8. (b) C trace of the molecules A from the two crystal forms coloured by their mean residue B-factors. The two structures differ in the orientation of the C2 loop, the lengths of H5 and H6, the conformation of the C3 loop and the C-terminal fragment. Regions of maximal structural differences also show the highest B-factors.
Figure 3. Relation of the P3₁ crystal structure to the membrane environment. (a) Mean residue height relative to the bilayer derived by docking the X-ray coordinates in the cryo-EM map of 2D crystals, with mean residue B-factors superimposed. The mid-plane of the bilayer is defined by the c = 0 plane of the p22₁2₁ 2D lattice. The hydrophobic zone is estimated to lie in a 30 Å band about this plane, and the lipid headgroups in two 10 Å wide layers outside that. The secondary structure assignments are represented by vertical bars, with notches in the bars for H₄ and H₇ indicating the location of 3₁₀ segments. The curve of residue heights follows a straight path over sections of H₄ and H₆ that are perpendicular to the bilayer, but shows a zigzag pattern for the tilted sections. Superimposed on this curve is the mean B-factor per residue as a function of the sequence. (b) Mean residue B-factor as a function of transmembrane position. The B-factors are well clustered between the lipid phosphate positions at ±20 Å. They increase from the membrane interior towards the surface, but the lowest B-factors are found just to the extracellular side of the bilayer mid-plane, at the depth where the retinal is bound. The sharp rise of B-factors on the cytoplasmic surface is due to conformational variability in several regions.
Table 2. Tilt and kinks in the transmembrane helices

<table>
<thead>
<tr>
<th>Helix</th>
<th>Straight segments</th>
<th>Residue range</th>
<th>Tilt from membrane normal</th>
<th>Azimuthal orientation</th>
<th>Kink angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>N-segment</td>
<td>35–54</td>
<td>23</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>55–65</td>
<td>30</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>N-segment</td>
<td>70–87</td>
<td>-24</td>
<td>-85</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>88–100</td>
<td>-33</td>
<td>-14</td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>N-segment</td>
<td>106–113</td>
<td>34</td>
<td>55</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Midsegment</td>
<td>113–127</td>
<td>32</td>
<td>77</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>127–139</td>
<td>24</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>H4</td>
<td>N-segment</td>
<td>150–169</td>
<td>-4</td>
<td>67</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>169–173</td>
<td>-33</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>H5</td>
<td>N-segment</td>
<td>200–211</td>
<td>20</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>211–231</td>
<td>26</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>H6</td>
<td>N-segment</td>
<td>244–265</td>
<td>-3</td>
<td>9</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>265–279</td>
<td>-34</td>
<td>-53</td>
<td></td>
</tr>
<tr>
<td>H7</td>
<td>N-segment</td>
<td>286–296</td>
<td>25</td>
<td>62</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Midsegment</td>
<td>297–302</td>
<td>4</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>C-segment</td>
<td>303–309</td>
<td>18</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>H8</td>
<td></td>
<td>311–321</td>
<td>-86</td>
<td>89</td>
<td></td>
</tr>
</tbody>
</table>

- *Residue range of the straight segments were determined by fitting cylinders of 2.8 Å radius over the Cα positions. All segments have α-helical conformation, except for the C-segment of helix 4 and residues 296–299 within the mid-segment of helix 7, which have the 310 conformation.
- *The membrane normal was taken from the c-axis of the 2D crystal in the two-sided plane group p22121. Positive tilt angle means the helix runs from the extracellular side towards the cytoplasmic side as the sequence number increases.
- *Angle from the a-axis of the p22121 lattice to the projection of helix segments in the plane of the 2D crystals, by a right-handed rotation about the membrane normal.

terminate in the head group layer, whereas H3, H5 and H6 reach into the cytoplasmic space. Residues 66, 138–141, 144–147 and 229–248 from the three cytoplasmic loops are well exposed above the lipid head group layer (Figure 3(a)) and available for interactions with cytoplasmic proteins. The peripheral helix H8 is partitioned between the head group layer and the cytoplasm and is slightly inclined with its N terminus more exposed to the cytoplasm than its C terminus. The polypeptide following H8 turns to the cytoplasm, but the palmitoylate chains on Cys322 and Cys323 are inserted into the hydrophobic zone.

Transmembrane helices in rhodopsin are kinked (see subsection Transmembrane domain), and only the N-terminal segments of H4 and H6 are oriented roughly normal to the membrane (Table 2). Figure 3(b) shows that the mean residue B-factors are strongly correlated with the depth in the bilayer, irrespective of the helix tilt angles. The lowest B-factors, on average, are found not at the midplane of the bilayer midplane but in a 4 Å band to the extracellular side of it. Therefore the dark state rhodopsin structure is most ordered at that depth in the membrane, which coincides with where the ionone ring of the 11-cis-retinal chromophore is located. Beyond the phosphate head group positions (± 20 Å), the B-factors rise distinctly more rapidly on the cytoplasmic surface than in the extracellular domain, which is due to conformational variability in several regions.

**Extracellular domain**

The extracellular domain includes the short β1-β2 hairpin, which together with the β3-β4 hairpin in the transmembrane domain (Figure 1(a)) has been described as forming a "plug" for the chromophore-binding pocket. The two β-hairpins are cross-linked by water-mediated H-bonds between main-chain atoms of Tyr10 in β2 and Pro180 in β3 (Figure 4). Three water molecules (3, 5 and 13) are responsible for this, of which one (13) was identified in the P41 structure. These water molecules form additional H-bonds to Gly182 and Gln184 in the β3-β4 hairpin and to Tyr192 near β4, further linking the β1-β2 hairpin to the transmembrane domain and securing the plug. Between Pro12 at the end of β2 and Pro23, the polypeptide forms an outward "horn" (Figures 1(a) and 4), which is rigid with a turn (residues 15–18) at the apex and several side-chain to main-chain H-bonds. The horn was suggested to act as a spacer between opposite faces of the retinal disc, but the N-linked oligosaccharides on Asn2 and Asn15 may also contribute in this role (Figure 1(a)). In our structure there is density for only three residues in each of the oligosaccharide chains on Asn2 and Asn15: GlcNAc-(β1,4)-GlcNAc-(β1,4)-mannose. The distal carbohydrate residues cannot be resolved, probably because they adopt heterogeneous orientations within the solvent channel. The orientations of the oligosaccharide chains relative to the protein differ from those in the P41 structure, in which they were found to form intermolecular contacts. The remainder of this domain contains two turns (residues 22–25, 28–32) in the lipid head group layer adjacent to the transmembrane domain.

Many H-bonds and van der Waals interactions exist between the extracellular domain and the base of the transmembrane domain, lending credence to the view that the N-terminal domain may fold first
Figure 4. Interface between the extracellular and transmembrane domains. Loops in the two domains are coloured blue and yellow, respectively. Note the water-mediated H-bonds linking β2 and β3, the stacking of Pro12 with Pro285 and Pro27 with Tyr102, and the environment of Pro23. Amino acids are designated by the one-letter code for clarity in Figures 4-10.

and provide a template to assist the assembly of the helical domain. Figure 4 shows that Asn2 is in contact with Gly280 and Asp282 in the E3 (H6-H7) loop, and Pro12 is stacked with Pro285 in the same loop. Pro27 is stacked with Tyr102 in the E1 (H2-H3) loop, which is completely conserved among vertebrate opsins. Pro23 and Gln28 form main-chain H-bonds with Gly101 and Tyr102. Pro23, the most conserved (99%) residue in the extracellular domain of vertebrate opsins, is in a buried turn and in contact with Phe9 and Val11 in β2, and Gln28, as well as Tyr102 and Phe103 in the E1 loop and Pro180 and Gln184 in the E2 loop. Mutations of Pro23 can prevent correct folding of both domains. Pro23His, the most frequent mutation in North American patients with retinitis pigmentosa, has been shown to cause aggregation of misfolded rhodopsin, which is targeted for degradation by the ubiquitin proteosome system in cultured cells but also impairs the degradation pathway, and that impairment may underlie the neurodegenerative phenotype.

Sequences of visual pigments are highly conserved and, relative to vertebrate opsins, the invertebrate opsins show few insertions (GPCRDB data base†). The longest insertion, of 13 residues, is located in the turn of the β1-β2 hairpin; the others, of two, one and five residues each, are found after residue 332. Within the β1 and β2 strands visual pigments share around 30% sequence identity. At this level of similarity the invertebrate opsins can be expected to contain a similar β-hairpin and accommodate the insertion on the extracellular surface (Figure 3(a)).

Transmembrane domain

Fitting cylinders of 2.8 Å radius to the Cα trace in Cδ2 divided the transmembrane helices into two or three straight segments each, with kink angles of 8° to 36° in-between (Table 2). Having kinks allows the tilted transmembrane helices to pack more intimately than otherwise possible and to enfold the chromophore in a tight binding pocket. The kinks in H1, H4, H5, H6 and H7 are caused by proline

† http://www.gpcr.org/7tm/
residues at positions 53, 170, 171, 215, 267 and 302, respectively. Except for Pro53 these proline residues are among the most conserved residues in rhodopsin-like GPCRs, so this set of kinks is expected at equivalent sequence positions throughout this family, although the kink angles may vary in a sequence-dependent manner. On the other hand, H2 is distorted at the consecutive glycine residues Gly89-Gly90, which allows the side-chain of Thr92 to H-bond to the main chain at residue 88. In H3 the kinks occur at Glu113, the counterion for the protonated Schiff base in the dark state, and next to the ionone ring of the chromophore. In H7 a second kink occurs near Lys296, which is at the start of the $\beta$10 segment and the site of chromophore attachment. These kinks in H2, H3 and H7 are in contact with the chromophore and would be shared by the visual pigments. Helix kinks have been noted in rhodopsin structures determined in the tetragonal form of 3D crystals and in the $\beta2,2$ form of 2D crystals, our definition (Table 2) differs slightly from earlier descriptions regarding location and angle of the kinks.

The kinks are also important as sites of lateral polar interactions, where rotation of the peptide planes enables formation of inter-helical H-bonds and water-mediated H-bonds. Thus the helices are linked by three-dimensional H-bonding networks, which in our refined structure are more extensive, laterally and perpendicular to the membrane, than previously reported.

Figure 5 shows one of the H-bonding networks that involves Asn55 at the kink in H1, and also Asp83 in H2, Gly120 in H3, Met257 and Trp265 in H6, Ser298, Ala299, Tyr301 and Asn302 in H7, and four water molecules including the hitherto unidentified Wat10. This network extends from Trp265 in the chromophore binding pocket via Wat10 all the way to the highly conserved NPxxY motif (Asn302-Tyr306 in bovine rhodopsin) in H7 close to the cytoplasmic limit of the hydrophobic zone (see Figure 3(a)). Asn55 and Asp83 in this network are respectively the most conserved residues in H1 and H2 of rhodopsin-like GPCRs. Trp265, Ser298, Ala299 and Asn302 are all widely conserved across this family, and Met257 is conserved in vertebrate opsins. The proximity and interactions of these residues are critical to the stability and function of GPCRs and of rhodopsin in particular.

The H-bonds between Asn55, Asp83, between Asn55 and the peptide O of Ala299 in the kink of H7 at the C terminus of its $\beta10$ segment, and the water-mediated H-bonds between Asp83 and Asn302 on the other side of this H7 kink can be expected to stabilise the packing of H1, H2 and H7 in the dark state. H-bonding between these conserved residues underlies the apparent interaction among corresponding residues in related GPCRs. For example, the thyrotropin-releasing hormone receptor was completely inactivated by mutation to a non-H-bonding residue at the position of Asp83 or at both positions of Asn55 and Asn302. In several receptors an Asp-Asn pair at positions of Asp83 and Asn302 was required for activation of G proteins, but their positions can be interchanged without loss of this activity. The examples suggest that different H-bonding interactions of these residues are required for stabilising the dark state and for attaining the activated state.

In bovine rhodopsin Asp83 is not in contact with the chromophore (Figure 5), but its replacement by Asn causes a slight blue shift of the absorption maximum, indicating an alteration of the Schiff base environment. Low-temperature Fourier transform infrared (FTIR) difference spectroscopy showed structural changes in two internal water molecules and peptide groups during the conversion to bathorhodopsin, the first intermediate state following photo-isomerisation of the retinal, which would have been reached within picoseconds of light absorption had it occurred at physiological temperature. The spectral change due to one of the water molecules is abolished when the Glu113 counterion is replaced by Gln, but the change due to the second water molecule persists with a frequency shift. The latter was thought to be vibrationally linked to the Schiff base region through the peptide backbone and at a location affected by replacement of Asp83 and Gly120. This interpretation is borne out by the water-mediated H-bonds between Asp83 and the peptide backbone of H3 and H7 (Figure 5), and Wat12 can be identified as that second water molecule. Other FTIR studies showed that Asp83 is protonated in the dark state and remains so in the biochemically active metarhodopsin II (MII) state, but it underwent a change or increase of H-bonding on entering MII. Note that the internal water molecules in the surrounding of Asp83 (Figure 5) could, by small movements, facilitate rapid exchange of H-bonds without requiring concerted movement of the protein.

Trp265 in this H-bonding network was one of the first chromophore contacts identified. Its side-chain is in a U-shaped bend formed by the 11-cis-retinal and the side-chain of Lys296, with the indole plane perpendicular to the membrane and in contact with the ionone ring (Figures 5 and 6). Resonance Raman data suggest that its micro-environment becomes less hydrophobic within picoseconds of photon absorption, which was attributed to changing interaction with the ionone ring. On MII formation, UV absorbance changes also indicate a weakening of the indole H-bond in Trp265 and Trp126. In the dark state structure the side-chain of Trp265 is oriented with the NH group pointing to the cytoplasmic side and forming an H-bond with Wat10 (Figure 5), which is located in-line with the NH group. The UV difference in MII might indicate a net displacement of the Trp265 side-chain towards the extracellular side. The newly identified water molecule Wat10 is in close contact with the side-chain of Ala124 in H3 (Figure 5). In rhodopsin-like GPCRs the residue at position 124 is highly conserved, but the predominant amino acid is Ser. From this we infer that
Figure 5. A transmembrane slice showing H-bonding networks and hydrophobic contacts between the retinal-binding pocket and the cytoplasmic surface.

Wat10 is a surrogate for the Ser OH in providing an H-bond to the indole NH of Trp265, and furthermore that H-bonding between this pair of side-chains in H3 and H6 is a conserved feature of GPCR structure in the inactive state.

The side-chain of Ser298 is adjacent to Trp265, but it is H-bonded instead to the main chain O at residue 295 (Figure 5). This side-chain to main-chain H-bond compensates for the elongation of helix pitch at the start of the 310 segment (residues 296–299). In rhodopsin-like GPCRs, position 298 is highly conserved as Ser or Asn, both being residues capable of forming side-chain to main chain H-bonds. Therefore it is likely, that H7 in the wider family also contains a 310 segment at this level in the membrane.

Asp83 and Trp265 are both linked by H-bonds through ordered water molecules to Asn302 and Met257. This extended H-bonding network is sealed from the cytoplasmic compartment by a hydrophobic barrier made of six residues from helices H2 (Leu76, Leu79), H3 (Leu128, Leu131) and H6 (Met253, Met257). Five of those residues are highly conserved in rhodopsin-like GPCRs, and the sixth, Met253, is 100% conserved in vertebrate visual pigments. They are arranged with methionine residues on one side and leucine residues on the other (Figure 5) and could facilitate the relative movement of H3 and H6 following photoactivation while retaining van der Waals contacts. Replacement of Met257 by all other amino acids except Leu caused bovine opsin to activate Gt
Rhodopsin Structure in Trigonal Crystals

1421

Figure 6. Environment of the 11-cis-retinal. (a) Environment of the protonated Schiff base and its counterion Glu113. (b) Environment of the ionone ring and the kinks in H6 and H7 cross-linked by a water molecule.

The kinks in H3 and H5 are intimately associated with the lining of the chromophore pocket. At Glu113 in H3, a water molecule (Wat16) is H-bonded between the peptide carbonyl and side-chain carboxyl oxygen atoms. Consequently, the peptide oxygen atom of Glu113 is not H-bonded along the helix axis, hence the kink (Figure 6(a)). Wat16 is most likely the water molecule responsible for that low-temperature FTIR difference feature between the dark and batho states, that was abolished when Glu113 was replaced by Gln.61

OE1 of Glu113 is H-bonded to the peptide N of Cys187 in β4 (Figure 6(a)), so the Schiff base counterion is in contact with an H-bonding network that leads from the chromophore-binding pocket, through the plug, to the extracellular surface. This network encompasses Tyr268 in H6, residues Glu181, Gly182, Gly184, Tyr191 and Tyr192 from the β3-β4 and β4-H5 loops (Figure 6(a)), as well as Tyr10 from β2 and five water molecules. The H-bonds that cross-link the β1-β2 and β3-β4 hairpins (Figure 4) belong to this network. Tyr268, whose OH group is 3.6 Å from the 11,12-cis bond of the retinal in the dark state (see Table 4 below) was implicated by FTIR and Raman studies to participate in the initial conformational change at the batho stage and in the transition to the MII state.68-70

The second kink in H3 is in contact with the ionone ring of the chromophore, where H3 and H5 are linked due to a triangular set of H-bonds on the side of the ionone ring away from Trp265 (Figure 6(b)). The Trp126 indole NH forms a pair of bifurcated H-bonds with OE1 of Glu122 and ND1 of His211, while OE2 of Glu122 is H-bonded to the backbone O of His211, which is so oriented because of the kink in H5 at His211 due to Pro215. Glu122 is known to be protonated (neutral) in the dark state and in MII.65 These interactions account for the

constitutively. The level of constitutive activity is well correlated with the affinity of mutant opsins for all-trans-retinal, suggesting that the mutations disposed the apoprotein towards the active conformation normally attained only following light-induced chromophore isomerisation.70 By implication, Met257 in wild-type rhodopsin contributes a critical packing interaction. It is at the junction between the H-bonding network and the hydrophobic barrier, is in contact with Asn302 of the NPxxY motif in H7, and prevents the helix movement until this is triggered by chromophore isomerisation.

Tyr306, at the cytoplasmic end of the NPxxY motif in H7, is at the height of the lipid headgroup layer (Figure 3(a)) where it is H-bonded via Wat7 to Thr62 in H1 and directly to Asn73 in H2 (Figure 5), two residues which are conserved in vertebrate visual pigments and rhodopsin-like GPCRs, respectively. Tyr306 is stacked against Phe313 in H8, and similarly Ile307 in H7 is in contact with Met317. Such interactions can relay movements of transmembrane helices to parts of the cytoplasmic surface, for example realigning H8 after the photoisomerisation. Spin-label data on Tyr306, Phe313 and Cys316 showed light-dark structural differences,71 and double spin labelling showed movements by 2–4 Å of the cytoplasmic end of H7 relative to that of H1,72 and of H2 relative to H8.73

The proper realignment of H8 appears to be important in the interaction between the cytoplasmic peptide of rhodopsin and the C-terminal peptide of Gzα, because the inhibitory mutations either map to the structurally sensitive connecting region between the transmembrane H7 and peripheral H8 (Asn310, Lys311 and Cln312), or form buried contacts between H7 and H8 (Tyr306, Ile307, Phe313 and Met317).39,74-77

OE1 of Glu113 is H-bonded to the peptide N of Cys187 in β4 (Figure 6(a)), so the Schiff base counterion is in contact with an H-bonding network that leads from the chromophore-binding pocket, through the plug, to the extracellular surface. This network encompasses Tyr268 in H6, and residues Glu181, Gly182, Gly184, Tyr191 and Tyr192 from the β3-β4 and β4-H5 loops (Figure 6(a)), as well as Tyr10 from β2 and five water molecules. The H-bonds that cross-link the β1-β2 and β3-β4 hairpins (Figure 4) belong to this network. Tyr268, whose OH group is 3.6 Å from the 11,12-cis bond of the retinal in the dark state (see Table 4 below) was implicated by FTIR and Raman studies to participate in the initial conformational change at the batho stage and in the transition to the MII state.68-70

The second kink in H3 is in contact with the ionone ring of the chromophore, where H3 and H5 are linked due to a triangular set of H-bonds on the side of the ionone ring away from Trp265 (Figure 6(b)). The Trp126 indole NH forms a pair of bifurcated H-bonds with OE1 of Glu122 and ND1 of His211, while OE2 of Glu122 is H-bonded to the backbone O of His211, which is so oriented because of the kink in H5 at His211 due to Pro215. Glu122 is known to be protonated (neutral) in the dark state and in MII.65 These interactions account for the
FTIR observation that mutations of His211 influence Glu122 through a third residue, which is itself affected by mutation of Glu122 to Ala but not to Cln.\cite{80} Figure 6(b) suggests that Trp126 is the third residue. These residues can be expected to show sensitivity to movements of the ionone ring, with which Glu122 is in direct contact. FTIR difference suggested stronger H-bonding of Glu122 on transition to M\textsubscript{II} \cite{80,81} whereas UV absorption difference indicated reduced hydrophobicity of the microenvironment for Trp126 as well as weaker H-bonding of its indole NH caused by a general conformational change in the formation of M\textsubscript{II} \cite{69}.

The kinks in H\textsubscript{6} and H\textsubscript{7} are only 4.8 Å apart between C\textsuperscript{a} atoms of Tyr268 and Pro291. A water molecule (Wat1), H-bonded to the backbone at Tyr268, Pro291 and Ala295, bridges the kinks (Figure 6(b)). H\textsubscript{6} shows the most pronounced kink among the rhodopsin helices (Table 2) and contains the conserved Pro267-Tyr268 pair at this kink. Similarity in the light versus dark FTIR difference between [\textsuperscript{2}H]tyrosine-labelled rhodopsin and bacteriorhodopsin has led to the proposal\cite{79} that during photoactivation Pro267-Tyr268 in H\textsubscript{6} of rhodopsin may play an analogous role to Tyr185-Pro186 in helix F of bacteriorhodopsin.\cite{82} That role is to provide a pivot for the outward tilt of the cytoplasmic segment of H\textsubscript{6} during the transition to M\textsubscript{II} as indicated by spin-label studies.\cite{4} The structures of dark-state bacteriorhodopsin and a triple mutant mimic of its M-state (analogous to the M\textsubscript{II} state in rhodopsin) were shown to differ mainly by an outward tilt of the mutant’s helix F by a maximum of 3.5 Å at the cytoplasmic end, which can be described as a hinge movement extrapolating to a pivot around the Tyr-Pro pair.\cite{82} If a similar movement occurs in the transition of rhodopsin to its M\textsubscript{II} state, then the water molecule between the H\textsubscript{6} and H\textsubscript{7} kinks would appear to be securing the potential pivot in the configuration of the dark state.

At the kink in H\textsubscript{4} near its extracellular end, the 2F\textsubscript{o} – F\textsubscript{c} map (Figure 7(a)) reveals a non-conventional H-bond, with an aromatic ring hydrogen atom of Phe203 in H\textsubscript{5} acting as donor\cite{94} to the carbonyl O atom of Cys167 as acceptor (Figure 7(b)). The donor-acceptor distance of 3.10 Å indicates a relatively strong H-bond. Residue 203 is Tyr in the consensus sequence of rhodopsin-like GPCRs, and it is Tyr in 76% of the visual pigment sequences but Phe in the rest\cite{84} such as bovine rhodopsin. While
Rhodopsin Structure in Trigonal Crystals

Table 3. Dihedral angles along the conjugated double bond system of the 11-cis-retinal chromophore

<table>
<thead>
<tr>
<th>Dihedral angle</th>
<th>Observed values (deg.)</th>
<th>Refinement restraints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Molecule A</td>
<td>Molecule B</td>
</tr>
<tr>
<td>C4, C5, C6, C7</td>
<td>179</td>
<td>177</td>
</tr>
<tr>
<td>C5, C6, C7, C8</td>
<td>-57</td>
<td>-53</td>
</tr>
<tr>
<td>C6, C7, C8, C9</td>
<td>-179</td>
<td>-176</td>
</tr>
<tr>
<td>C7, C8, C9, C10</td>
<td>-178</td>
<td>-178</td>
</tr>
<tr>
<td>C8, C9, C10, C11</td>
<td>179</td>
<td>179</td>
</tr>
<tr>
<td>C9, C10, C11, C12</td>
<td>176</td>
<td>173</td>
</tr>
<tr>
<td>C10, C11, C12, C13</td>
<td>171</td>
<td>171</td>
</tr>
<tr>
<td>C11, C12, C13, C14</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>C12, C13, C14, C15</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>C13, C14, C15, N5</td>
<td>177</td>
<td>178</td>
</tr>
<tr>
<td>C14, C15, NE, Cc</td>
<td>179</td>
<td>-179</td>
</tr>
</tbody>
</table>

\(^a\) Error: deviation of observed values from their mean in the absence of noncrystallographic symmetry restraints on the retinal moiety.

Tyr can be expected to form an H-bond at this position, our observation demonstrates that in the low dielectric medium of the membrane interior, Phe can replace Tyr in donating a stabilising H-bond to the conserved kink in H4.

From the Pro170 at the kink in H4 to Asp190 at the end of the β3-β4 hairpin, visual pigments contain nine highly conserved residues, attesting to the importance of this region to the structure and function of rhodopsin.\(^8\) Figure 7(b) shows some of the H-bonds from the conserved Pro170, Pro171, Gly174, Trp175, Tyr178 and Asp190, which fasten the β-hairpin to the helix bundle. The hairpin is anchored to H3 by the disulphide bond between Cys110 and Cys187, and its first and last residues, Arg177 and Asp190, form a salt-bridge (Figure 7(b)) that helps to position β4 at the “floor” of the retinal-binding pocket.

Conformation and environment of the chromophore

Table 3 lists the torsion angles along the conjugated double bonds of the 11-cis-retinal and its imine linkage to Lys296. Uncertainties were estimated by comparing values in the two NCS copies after one cycle of mock refinement in the absence of NCS restraints on the retinal moiety. The chromophore is confirmed as containing three planar segments,\(^8\) with deviations from the planarity of the π-bonding system about the 6s-cis, and 11-cis, 12s-trans bonds because of the steric interactions between 5-CH\(_3\) and 8-H, and between 13-CH\(_3\) and 10-H, respectively. The conformation of the chromophore in our structure is similar to that in the 1F88 coordinates for the tetragonal crystal form,\(^13\) and disagrees with the 1HZX\(^15\) and 11LH\(_{9}^9\) coordinates.\(^13\) For example, 11LH\(_{9}\) shows a planar 11-cis bond but significantly non-planar C8-C9 and C14-C15 bonds. It should be pointed out that when we used the 1F88 coordinates for the molecular replacement model, we excluded the retinal as a test against model bias, and we have built the initial model of the chromophore from the small molecule crystal structure of 11-cis-retinal\(^20\) by adjusting its torsion angles to fit the cross-crystal averaged electron density map (see Methods), which was consistent with three planar segments. Refinement was carried out with planarity restraints imposed on the chromophore, as was appropriate at the medium resolution. However, the weights of the restraints (Table 3) were set to be 5–12 times smaller than those on peptide bonds or aromatic rings, so they could not have induced a planar conformation had it not been rooted in the X-ray data. Thus our structure provides a credible description of the retinal-protein interactions in the dark state.

The torsion angle about the 11-cis bond is determined to be -13(±2)° (Table 3 and Figure 6). This value is compatible with the calculation, based on analysis of resonance Raman intensities, that the C11=C12 torsion angle increases, within 50 fs after excitation, from -15° in the dark state towards the 90° transition state for the cis-to-trans isomerisation.\(^86\) Both the rapidity and stereospecificity of the initial torsional dynamics are driven by the strain on this non-planar segment.\(^86,87\) The chirality and pucker of the ionone ring are also in agreement with NMR data.\(^88\)

Figure 8(a) shows the electron density map around the retinal. Figure 8(b) is a schematic diagram of the retinal-binding pocket, showing H-bonds of Lys296 and Glu113, and the mostly hydrophobic contacts with the retinal within a distance of 4 Å. Those contact distances are given in Table 4. The elongated 11-cis-retinal buries about 220 Å\(^2\) of surface area on a total of 22 residues that form the binding pocket. Roughly one-third of this area is contributed by Trp265 and Tyr266 in H6, another third is from Thr118, Glu122 in H3 and Met207, Phe212 in H5, and the rest is made up by 16 residues from helices H3 to H7 and the plug\(^14\) (Ala117, Lys296, Cys187, Phe261, Ala269, Phe208, Ala292, Tyr191, Gly121, His211, Gly188, Cys167, Ile189, Glu181, Ser186 and Ala272 in descending order of contact area). The large number of contact residues in addition to its covalent linkage...
Figure 8. The retinal-binding pocket. (a) Electron density map around the 11-cis-retinal. (b) Schematic drawing of the binding pocket showing all residues within 4 Å of the retinal and Lys296. This drawing was modified from the output of LIGPLOT.148
**Table 4. Retinal contact distances**

<table>
<thead>
<tr>
<th>PROTEIN ATOMS</th>
<th>CHROMOPHORE ATOMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contacts up to 4Å</strong></td>
<td>C2</td>
</tr>
<tr>
<td>H1</td>
<td>Met 44 CE</td>
</tr>
<tr>
<td></td>
<td>Leu 47 CD1</td>
</tr>
<tr>
<td>H2</td>
<td>Phe 91 CE1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>Glu 113 CD</td>
</tr>
<tr>
<td></td>
<td>Ala 117 CB</td>
</tr>
<tr>
<td></td>
<td>Thr 118 OG1</td>
</tr>
<tr>
<td></td>
<td>Gly 121 CA</td>
</tr>
<tr>
<td></td>
<td>Glu 122 N</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>B3-β4</td>
<td>Ser 186 CB</td>
</tr>
<tr>
<td></td>
<td>Cys 187 O</td>
</tr>
<tr>
<td></td>
<td>Gly 188 CA</td>
</tr>
<tr>
<td></td>
<td>Ile 189 N</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>B4-H5</td>
<td>Tyr 191 OH</td>
</tr>
<tr>
<td>H5</td>
<td>Met 207 CG</td>
</tr>
<tr>
<td></td>
<td>Phe 212 CB</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe 261 C2</td>
</tr>
<tr>
<td>H6</td>
<td>Trp 265 CD2</td>
</tr>
<tr>
<td></td>
<td>Tyr 268 CD2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>H7</td>
<td>Ala 292 O</td>
</tr>
<tr>
<td></td>
<td>Phe 293 CA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
contribute to a favourable binding enthalpy of \(-11 \text{ kcal mol}^{-1}\). In our structure, the seven residues with the lowest \(B\)-factors, Trp265, Ala295, Ala269, Thr118, Phe91, Ala117 and Tyr268, are all located around the retinal-binding pocket (Figure 8), which attests to the stabilising interaction between the chromophore and its binding pocket. This stabilisation is responsible for the "inverse agonist" effect of the 11-cis-retinal that keeps the basal activity of rhodopsin towards Gt in the dark very low, below that of the ligand-free opsin, which is itself 10\(^6\)-fold less active than the MII state with all-trans-retinal bound.\(^9\) The three aromatic residues showing the largest chromophore contact areas, Trp265, Tyr268 and Phe212, are highly conserved throughout rhodopsin-like GPCRs,\(^3\) suggesting common elements in the mechanism for stabilising the inactive receptor conformation from the boundaries of the ligand-binding pocket.

Electrostatic interactions of the binding pocket with the chromophore affect the visible absorption maximum.\(^{91,94}\) Three residues near the ionone ring correspond to the major colour tuning positions in red/green cone pigments identified by primate genetics.\(^95\) Among these Phe261 and Ala269 are direct contacts of the ionone ring (Figure 6(b)), and mutagenesis to Tyr and Thr, respectively, reproduced the predicted red shift,\(^91\) which is consistent with a dipolar stabilisation of the positive charge redistribution in the excited state chromophore towards the ionone ring.\(^94,96\) The third residue, Ala164, is a second shell contact of the ionone ring via Glu122. In rhodopsin, its substitution by Ser produced only a minor red shift,\(^91\) but this can be explained by the dipole contribution of Ser being overshadowed by the intervening Glu122,\(^97\) since replacing Glu122 by Asp was known to cause a blue shift.\(^91\) In the red/green cone pigments, where the conserved residue at the position 122 is Ile, a strong tuning effect can be expected.

The structure of the binding pocket is illustrated in Figure 6 with the Schiff base (Figure 6(a)) or the ionone ring (Figure 6(b)) in the foreground. In the Schiff base region, the Lys296 side-chain is fully extended and held in a channel between Met44, Leu47 and Phe91 from H1 and H2 and Phe293 from H7, as previously reported.\(^14\) The Glu113 side-chain is in an unusual rotamer because Wat16 is H-bonded between its peptide O and carboxyl OE2 atoms. The OE2 atom is thus placed in-line with the NH bond of the protonated Schiff base. Thr94 and Gly90 from H2 make close contact with Glu113 and Wat16 bound to it (see below). OE1 of Glu113 is H-bonded to the backbone of Cys187, which is in a disulphide bond with Cys110 in H3. Thus the geometry of the Lys296, Glu113 ion pair is stringently defined. The structure suggests that a salt-bridge between them can also exist in the apoprotein, where it is required for maintaining the inactive state.\(^98\)

In Figure 8(b), Wat16 forms a shorter H-bond to the negatively charged carboxyl OE2 than to the neutral peptide O, showing that it is strongly bound to the anion. The Wat16-mediated H-bonds delocalise the negative charge on the Glu113 carboxylate group and lower its \(pK_a\). Stabilisation of the anion, aided by the precise geometry of the ion pair, in turn stabilises the proton on the Schiff base, elevating its \(pK_a\) in the dark state to >16.\(^99\) Therefore spontaneous deprotonation is kept at a low rate, which is essential for dim light vision.\(^100\) Change of geometry in the Schiff base region after the photoisomerisation can destabilise the proton and pave the way for the deprotonation of Schiff base in MI.

The distance between OE2 of Glu113 and the Schiff base nitrogen atom is 3.2 Å (Table 4) on both NCS copies, and is in agreement with the 3.1 Å distance in the P4\(_1\) crystal form.\(^19\) This is less than the 4.3 Å distance implied by a similarity of \(^{15}\)N chemical shift of the protonated Schiff base in rhodopsin and that in model retinylidene with iodide as counterion.\(^101\) However, delocalisation of the negative charge in Glu113 effectively moves the centre of the counterion further away from the protonated Schiff base by about 1 Å, thus weakening their electrostatic coupling to permit greater delocalisation of the positive charge from the Schiff base into the polyene of the retinal,\(^102\) which was indicated by the nitrogen chemical shift. Therefore, Glu113 and Wat16 form a complex counterion in the dark state, which both stabilises the proton on the Schiff base and determines the wavelength of the visible absorption maximum. Mutating Glu113 to the smaller Asp results in a red shift without dependence on external chloride ions is also consistent with an increase of the ion pair distance.\(^103\)

Recently, because the Glu181 to Gln mutation was found to lower the \(pK_a\) of the Schiff base under MI-like conditions, the counterion was proposed to switch in MI to Glu181,\(^104,105\) accompanied by reorganisation of the E2 loop coupled to movements of H3.\(^105\) The dark state structure suggests that the proposed backbone movements are not required for the Glu181 side-chain to get close enough to the Schiff base to measureably influence its electrostatic environment. The Schiff base NH bond in the all-trans retinal may be tilted more towards Ser186 and Glu181 on the H6, H7 side of the polyene backbone (out of the page in Figure 6(a)), instead of tilting towards Glu113 on the H3 side as in the dark state structure.\(^106\) The NMR result, that the ionone ring and the adjacent segment of the retinal in MI despite being more relaxed remain largely in the dark state conformation,\(^107\) is consistent with the retinal isomerisation up to this stage involving rotations mainly on the Schiff base side of the isomerised bond. There is also evidence that the retinal isomerisation impinges on Tyr268.\(^65,78\) Its effect on the H-bonding network (Figure 6(b)) might allow the Glu181 side-chain to rotate towards the Schiff base. Such localised rearrangements from the known dark state structure may apply to the invertebrate rhodopsins, which lack an acidic residue at position 113 and use a conserved residue
corresponding to Glu181 as the counterion, and to UV-absorbing pigments, which have a neutral Glu113 in the dark state and switch to Glu181 for the counterion in MI.

Steric interactions with the binding pocket allow relaxation of the photo-isomerised retinal to trigger conformational changes that result in Gt activation in MI. The ionone ring is under steric restraints from residues in helices H3, H5 and H6 (Table 4; Figure 6(b)) up to the MI state. Evidence for a change in the microenvironment of adjacent Trp265 in the formation of MI may indicate a movement of the ionone ring. A ring movement will impinge on Trp265, which is the most ordered residue in the dark state structure, the residue with the largest contact area with the retinal, and is widely conserved among GPCRs. It may also affect Phe321, a conserved aromatic residue in visual pigments which is in contact with the C3, C4 edge of the ring and located in the cytoplasmic segment of H6 one helical turn above Trp265 and below Met257 (Figure 5). The residue at position 261 in the glycoprotein hormone receptors has been shown to play a similar role to Met257 in rhodopsin in forming a stabilising helix packing interaction with Asn302 in H7.

Thus retinal interactions with the aromatic residues following photo-isomerisation could precipitate a release of the packing interactions between H3 and H6 (see previous subsection on Transmembrane domain), and between H6 and H7, that stabilise the inactive state. As shown in Figure 5, and in Figure 9(a), when the dark state structure is placed in the membrane frame, the ionone ring appears vertically below the conserved Glu134, Arg135 charge pair at the cytoplasmic border of H3, where light-induced protonation of Glu134 is a prerequisite for rhodopsin to bind Gt. The H-bonding networks and van der Waals contacts described in the transmembrane domain serve as conduits for vectorial transmission of the conformational changes across this distance, in details that are still unknown. Deletion of the 19-CH3 attached to C9 of the polyene chain was shown to result in a change in the microenvironment of adjacent residues in H1, H2 and H8; the other residues point outward to form a basic patch. The 2Fo-Fc map next to the C1 loop shows a strong density feature with the flatness of a peptide plane in contact with the side-chains of His65 and Lys67, and the main chain from His65 to Lys67. Its aqueous location and shortness make it rather unlikely to belong to a detergent molecule. We have interpreted this density as a di-peptide from the disordered C-terminal tail and tentatively assigned it to Asp330 and Asp331, because of the basic contact site and because the distance from the last ordered residue, Asn326, can be spanned by three or more residues with reasonable main chain torsion angles. The C-terminus of the putative di-peptide points toward H6. No other interpretable density was found for the C-terminal tail in the P31 crystal form. The presence of a thermolysin cleavage site between Pro327 and Leu328 and an Asp-N endoproteinase cleavage site between Gly329 and Asp330 is consistent with solvent exposure of the unobserved 327-329 segment. Site-directed spin labelling of rhodopsin in dodecylmaltoside micelles showed the C-terminal tail to be in equilibrium between an ordered and disordered conformation, with the disordered component increasing dramatically beyond Asp331. The density adjacent to loop C probably corresponds to the partially ordered component. In the P41 structure, loop C also forms contact with the C-terminal tail, but the chain orientations (Figure 2(b)) and residue assignments differ. Our tentative assignment of Asp330 and Asp331 pointing towards H6 is reconcilable with the report that the Cys mutant at position 338 in the C-terminal tail does not form a disulphide with the Cys mutant at position 65 in the C1 loop, but it does cross-link with Cys mutants at positions 242, 245 and 246 in H6, rapidly in the dark and slowly in the light.

**Cytoplasmic domain**

Having positioned the X-ray coordinates in the membrane frame (Figure 3(a)), we regard all residues above the mean level of phosphate groups (beyond 20 Å from the mid-plane of the bilayer) as belonging to the cytoplasmic domain. They include the three cytoplasmic loops (C1, C2, C3) and adjoining ends of transmembrane helices (residues Glu64-Pro71, Glu134-His152 and Glu225-Arg252), the peripheral helix H8 and adjoining loops (residues Asn310-Asn326), and a di-peptide fragment from the C-terminal tail (Figure 9).

**C1 loop and C-terminal fragment**

The C1 loop (residues 65-69) is very similar to that in the P41 form. It is well ordered (Figure 3(a)) and comprised of basic and hydrophobic residues (His65, Lys66, Lys67, Leu68, Arg69), which except for His65 are more than 97% conserved among vertebrate opsins, more conserved than the ends of H1 and H2. Leu68 and Arg69 point inward, making van der Waals and H-bonding contacts with residues in H1, H2 and H8; the other residues point outward to form a basic patch. The 2Fo-Fc map next to the C1 loop shows a strong density feature with the flatness of a peptide plane in contact with the side-chains of His65 and Lys67, and the main chain from His65 to Lys67. Its aqueous location and shortness make it rather unlikely to belong to a detergent molecule. We have interpreted this density as a di-peptide from the disordered C-terminal tail and tentatively assigned it to Asp330 and Asp331, because of the basic contact site and because the distance from the last ordered residue, Asn326, can be spanned by three or more residues with reasonable main chain torsion angles. The C-terminus of the putative di-peptide points toward H6. No other interpretable density was found for the C-terminal tail in the P31 crystal form. The presence of a thermolysin cleavage site between Pro327 and Leu328 and an Asp-N endoproteinase cleavage site between Gly329 and Asp330 is consistent with solvent exposure of the unobserved 327-329 segment. Site-directed spin labelling of rhodopsin in dodecylmaltoside micelles showed the C-terminal tail to be in equilibrium between an ordered and disordered conformation, with the disordered component increasing dramatically beyond Asp331. The density adjacent to loop C probably corresponds to the partially ordered component. In the P41 structure, loop C also forms contact with the C-terminal tail, but the chain orientations (Figure 2(b)) and residue assignments differ. Our tentative assignment of Asp330 and Asp331 pointing towards H6 is reconcilable with the report that the Cys mutant at position 338 in the C-terminal tail does not form a disulphide with the Cys mutant at position 65 in the C1 loop, but it does cross-link with Cys mutants at positions 242, 245 and 246 in H6, rapidly in the dark and slowly in the light.
Figure 9. The cytoplasmic domain. (a) Residues implicated in interaction with the heterotrimeric G protein transducin. The main chain is coloured in the four segments that project above the lipid phosphate groups and form the cytoplasmic domain: C1 loop region (residues 64–71), C2 loop region (134–152), C3 loop region (225–252), and H8 and C-terminal tail (310–326). The side-chain colours indicate: green, mutations causing constitutive Gt activation; red, residues cross-linked to Gtα; pink, mutations inhibiting stabilisation of MII by Gtα 340–350; yellow, mutations reducing Gt activation; greenish-yellow, gain-of-function for Gt activation upon reverse mutation from Ala; blue, Arg135 of the conserved E(D)RY motif; blue-grey, other residues discussed in the text. Residues falling in more than one category are coloured according to the former category. Long broken chains joining pairs of Cα atoms marked by backbone spheres indicate inter-helical cross-links that inhibit activation. The footprint of the 11-cis-retinal in the membrane plane (light pink) is shown in the background to illustrate its vertical alignment with key residues on the cytoplasmic surface. (b) Topography of the cytoplasmic surface showing a cavity adjacent to H3, with Leu131 at its bottom, and the Glu134-Arg135 pair forming a salt-bridge in its walls. Note the close contact of Arg135 with H6. Part (b) was prepared using PYMOL (http://pymol.sourceforge.net/).
were related through a hinge movement, and comparisons with the NCS differences showed that it is most likely caused by different packing contacts in the two crystal forms. The coordinate differences are above background level for residues 136–150 (Figure 2(a)), starting at the point where the highly tilted H3 becomes accessible from the surface (Figure 9). Thus the cytoplasmic tip of H3 appears to bend with C2 as the loop is re-oriented. The middle of C2, where B-factors exceed 100 Å² in both crystal forms, is probably rather flexible.

Peptides corresponding to the C2 and C3 loops, and H8, have been shown to interact with Gt. Residues important for Gt activation (Figure 9(a)) are found in the C terminus of H3 and in the C2 loop. Glu134, Arg135 and Tyr136 form the highly conserved E(D)RY motif found in H3 of all rhodopsin-like GPCRs. In the dark state, Glu134 is negatively charged and forms a salt-bridge with Arg135. The latter faces H6 and forms van der Waals contacts with Arg135. The former faces H6 and forms a salt-bridge with Glu134. The latter faces H6 and forms van der Waals contacts with Va1250, Thr251 and Va1254 in H6. Protonation of Glu134, in addition to deprotonation of the Schiff base, is required for rhodopsin to activate Gt. The cytoplasmic surface of rhodopsin shows a cavity next to H3 (Figure 9(b)). Leu131 from the hydrophobic cluster that seals the inter-helical H-bonding network (Figure 5) from the cytoplasm is visible from the opening. But the cavity extends like a tunnel under Pro71 and Arg147 to reach Phe148 and Ala153, and may provide a site for transducin to interact with rhodopsin. The Glu134, Arg135 charge pair is sequestered in the wall of this cavity in the dark state, so that they can form a salt-bridge, but Glu134 also has access to the aqueous phase to take up a proton. Protonation is thought to cause Glu134 to move to a non-polar environment and favour binding to Gt. Loss of the salt-bridge partner might also cause Arg135 to transfer to the aqueous phase, thus breaking its van der Waals contacts with the residues along H6 (Figure 9(a)). This would favour a parting of the cytoplasmic ends of H3 and H6, which is believed to occur in the activation of rhodopsin as indicated by studies using spin labelling, metal chelation, and disulphide cross-linking (Figure 9(a)). The Glu134 mutant, which mimics the protonation, favours binding of rhodopsin to Gt and causes constitutive activity of opsin. This mutant shows increased mobility of spin label attached to Cys140 in H3 or Cys316 in H8 in the dark, which is otherwise associated with phototransduction. Eliminating the charge on Arg135, the other half of the salt-bridge, does not cause the same mobility change.

The binding of Gt or the C-terminal peptide of Gt-stabilises MII (Gt340-350) and causes partial loss of Gt activation. Mutating individual residues of Tyr136–Val139 to cysteine also caused partial loss of Gt activation. A spin label attached to Lys140 showed a rise and a fall of mobility with time constants corresponding to the formation and decay of the MII state. Therefore, the cytoplasmic end of H3 is involved in the formation of the active state and binding of Gt.

In the C2 loop (residues 141–149), mutating Lys141 to Cys reduced Gt activation. Reagents attached to this residue cross-linked photoactivated rhodopsin to Gt. Replacing the sequence Lys141–Arg147 caused partial loss of Gt activation but did not prevent binding of Gt340-350. After mutating Lys141–Phe148 to polyalanine activation, restoring only two residues from each end, namely Cys140, Lys141, Arg147 and Phe148, restored 50% of the activity, but the middle residues also contribute to activation, possibly by influencing the conformation of the loop. The flexible middle section might become more ordered when MII binds to Gt and help to increase the affinity of Gt binding.

C3 loop and cytoplasmic ends of H5 and H6

Helices H5 (residues 200–230) and H6 (residues 241–276) in our structure are both longer by one helical turn at the cytoplasmic boundary in the P4₁ form. Therefore the C3 loop is elevated above the membrane surface (Figure 2(b)). This loop has the highest B-factors in the whole rhodopsin molecule (Figure 3(a)). However, these B-factors are authentic indicators of main chain flexibilities, since the relative scale of B-factors in the C3 loop compared to those in the C2 loop is similar to the relative mobilities of these surface loops measured by spin-label studies in detergent micelles. During model building, main chain fragments selected from a data base of well-refined, high-resolution structures by the program O for their similarity to the densities of the cytoplasmic ends of H5 and H6 and the loop were tightly clustered and all showed helix-like sections. After refinement with the best fitting fragment inserted into the C3 loop density, the real space correlation was 0.72 in the C3 loop, whereas the average for the whole structure was 0.95. The density was poor for residues Ala233–Ala235 and Gln238, which have the highest B-factors. Coordinates for the C3 loop represent the most probable conformation in a population of overlapping, variable conformations. However, the extra helical turn at the ends of H5 and H6 compared with the P4₁ form were related through a hinge movement, and comparisons with the NCS differences showed that it is most likely caused by different packing contacts in the two crystal forms. The coordinate differences are above background level for residues 136–150, starting at the point where the highly tilted H3 becomes accessible from the surface (Figure 9). Thus the cytoplasmic tip of H3 appears to bend with C2 as the loop is re-oriented. The middle of C2, where B-factors exceed 100 Å² in both crystal forms, is probably rather flexible.

Peptides corresponding to the C2 and C3 loops, and H8, have been shown to interact with Gt. Residues important for Gt activation (Figure 9(a)) are found in the C terminus of H3 and in the C2 loop. Glu134, Arg135 and Tyr136 form the highly conserved E(D)RY motif found in H3 of all rhodopsin-like GPCRs. In the dark state, Glu134 is negatively charged and forms a salt-bridge with Arg135. The latter faces H6 and forms van der Waals contacts with Va1250, Thr251 and Va1254 in H6. Protonation of Glu134, in addition to deprotonation of the Schiff base, is required for rhodopsin to activate Gt. The cytoplasmic surface of rhodopsin shows a cavity next to H3 (Figure 9(b)). Leu131 from the hydrophobic cluster that seals the inter-helical H-bonding network (Figure 5) from the cytoplasm is visible from the opening. But the cavity extends like a tunnel under Pro71 and Arg147 to reach Phe148 and Ala153, and may provide a site for transducin to interact with rhodopsin. The Glu134, Arg135 charge pair is sequestered in the wall of this cavity in the dark state, so that they can form a salt-bridge, but Glu134 also has access to the aqueous phase to take up a proton. Protonation is thought to cause Glu134 to move to a non-polar environment and favour binding to Gt. Loss of the salt-bridge partner might also cause Arg135 to transfer to the aqueous phase, thus breaking its van der Waals contacts with the residues along H6 (Figure 9(a)). This would favour a parting of the cytoplasmic ends of H3 and H6, which is believed to occur in the activation of rhodopsin as indicated by studies using spin labelling, metal chelation, and disulphide cross-linking (Figure 9(a)). The Gln134 mutant, which mimics the protonation, favours binding of rhodopsin to Gt and causes constitutive activity of opsin. This mutant shows increased mobility of spin label attached to Cys140 in H3 or Cys316 in H8 in the dark, which is otherwise associated with phototransduction. Eliminating the charge on Arg135, the other half of the salt-bridge, does not cause the same mobility change.

The binding of Gt or the C-terminal peptide of Gt (Gt340–350) stabilises MII (Gt124, Gt125). Replacing the Tyr136–Val139 sequence reduced Gt activation by ~80% and abolished stabilisation of MII by Gt340–350. Tyr136, which points outwards and does not interact with the rest of the protein, is an important determinant for the stabilisation of MII by Gt340–350, but Val137–Val139 may also contribute stabilising interactions for Gt binding. Val138 is adjacent to Glu134 and Arg135 (Figure 9(a)). Mutating individual residues of Tyr136–Val139 to cysteine also caused partial loss of Gt activation. A spin label attached to Lys140 showed a rise and a fall of mobility with time constants corresponding to the formation and decay of the MII state. Therefore, the cytoplasmic end of H3 is involved in the formation of the active state and binding of Gt.

In the C2 loop (residues 141–149), mutating Lys141 to Cys reduced Gt activation. Reagents attached to this residue cross-linked photoactivated rhodopsin to Gt. Replacing the sequence Lys141–Arg147 caused partial loss of Gt activation but did not prevent binding of Gt340-350. After mutating Lys141–Phe148 to polyalanine activation, restoring only two residues from each end, namely Cys140, Lys141, Arg147 and Phe148, restored 50% of the activity, but the middle residues also contribute to activation, possibly by influencing the conformation of the loop. The flexible middle section might become more ordered when MII binds to Gt and help to increase the affinity of Gt binding.
The structure of H8 (residues 311–321) is very similar in both crystal forms. Asn310 is the only non-helical residue between H7 and H8. Several side-chain to main chain H-bonds stabilise the H7/H8 corner, such as from Glu249 to Met309 and Lys311, from Asn310 to Phe313 and from Arg314 to Ile307. In addition, H8 is anchored at the membrane–aqueous interface by buried side-chain contacts with residues in H1 and H7 underneath, such as the contacts of Phe313 with Thr58, Val61, Thr62 and Tyr306, and of Met317 with Leu57, Val61 and Ile307. The presence of Asn310 and Lys311 are required for correct folding of the rhodopsin. If these two residues are kept, then replacing all other residues in H8 except Phe313 and Met317 by Ala, including replacing all outward-pointing residues, produced a mutant pigment with a potency for Gt activation equal to the wild-type rhodopsin. The contacts of these four residues just described suggest that they are required for the correct tertiary structure and orientation of H8. Deleterious effects of mutations in the N-terminal part of H8 are all interpretable in terms of perturbation of the interaction between H8 and the transmembrane domain. Therefore, although H8 participates in binding of Gt, its role in activation of Gt is secondary to the C2 and C3 loop regions.

*Expected location of G protein interaction sites in other GPCRs*

In the C2 loop, residues important to activation have been mapped in the conserved E(D)RY motif or just C-terminal to this. The C3 loop varies widely in length among GPCRs. As seen in the bovine rhodopsin structure, however, residues believed to participate in Gt binding and activation are clustered around the cytoplasmic ends of H5 and H6, and the spiral extension immediately above H5, while residues shown by alanine scanning mutagenesis to have no effect on Gt activation map to the C-terminal section of C3, which runs across the gap between H5 and H6. We may expect that, in the C3 region of other GPCRs, the Gt interaction sites will be similarly clustered with respect to the transmembrane helices. In the cannabinoid receptor, the C3 loop is about 30 residues long and a small helix was detected by NMR adjacent to the cytoplasmic end of H5, which is analogous to the H-bonded spiral extension of H5 in bovine rhodopsin, even though there is no sequence homology in this region. In the muscarinic acetylcholine receptors, the C3 loops are very large, but deletions of over 100 residues to reduce it to the length of C3 in bovine rhodopsin affected only the desensitisation process and left the G protein-dependent functions intact, such as agonist and antagonist binding and phospholipase activation. Therefore it is likely that these residues mapped in the C2 and C3 regions of bovine rhodopsin are the common G protein-binding sites in all GPCRs.

*Bound lipid and detergent molecules*

The presence of a bound phospholipid molecule near the cytoplasmic ends of H6 and H7...
Figure 10. Bound lipid and detergent molecules. (a) Two acyl chains of a phospholipid molecule bound to H6 and H7 in the cytoplasmic half of rhodopsin. The cytoplasmic segment of H6 is orientated roughly along the membrane normal. The acyl chains are inclined relative to it. Met309 is the C-terminal residue of H7. (b) An ordered LDAO molecule bound to H5, H6 and H7 just to the extracellular side of the kinks in H6 and H7. Pro267 is one of the contact residues. (c) A C8E4 molecule bound in the crevice between stacked helix bundles of the P31 crystals.
Cys323 it fills this crevice (Figure 10(c)). There are hydrophobic-to-hydrophobic and hydrophilic-to-hydrophilic contacts between the CρE4 and two protein molecules, one of them presenting the cytoplasmic end of H7 and H8 and the other presenting the extracellular end of H4 and H5 (Figure 10(c)). The choice of CρE4 as the detergent and its interactions in this crevice most likely precipitated the P31 crystal form.

Methods

Data collection and processing

Crystalisation and preparation of heavy atom derivatives are described in the accompanying paper. Complete data sets were obtained by combining wedges of 20–30° collected from different positions on one or more crystals. Data were processed using programs in the CCP4 suite. Intensities were integrated to anisotropic scattering. The atomic coordinates were calculated using SCALA. A consistent indexing regime was maintained, along with solvent flattening, 2-fold NCS averaging. The intensities were scaled in two rounds. The first round was done in the “batch” mode, with the batch showing the least negative B-factor serving as reference for B-factor normalisation; batches found with relative B-factors above 13, indicating significant radiation damage, were rejected from further scaling. The second round used smooth scale and B-factors, in which cumulative intensity distribution given by TRUNCATE. A “tails” correction for contributions from diffuse scattering was applied during both batch and smooth scaling rounds. Twinning was detected by inspecting the cumulative intensity distribution given by TRUNCATE. When it was present, the twin fraction was estimated and intensities were de-twinned using the program DETWIN.

Molecular replacement from the 1F88 coordinates and refinement

Molecular replacement was done using the program AMoRe and the rhodopsin coordinates in the P41 crystal form at 2.8 Å resolution (PDB code 1F88) as search model. The two rhodopsin molecules in the P41 asymmetric unit were superimposed to form one search model, but only the more ordered and more complete of the two molecules (chain A) was used to calculate initial phases. Non-protein atoms were omitted from the search model and initial phase calculation. In the calculated maps, the density for the retinal moiety, which was omitted from the model, was used to assess the quality of the phasing. Molecular replacement was done in parallel against each data set in Table 1A, and each solution was subjected to rigid body refinement, using first one molecule, then one helix, per rigid group. Cross-crystal averaging was carried out among Native-2, EMTS-1 and EMTS-2, together with solvent flattening, 2-fold NCS averaging and histogram matching, in DMMULTI.
model was rebuilt using O\textsuperscript{48} into the averaged map of Native-2. The model was refined using CNS\textsuperscript{22} with simulated annealing and group B-factors. NCS restraints were applied between the two molecules in the asymmetric unit, but the restraints were relaxed in the loop regions. This was because in simulated annealing trials, the loop residues showed systematically larger inter-molecular coordinate differences than residues in the helices, regardless of whether the NCS restraints have been imposed over the loops, whereas excluding the loop regions from NCS restraints was found to reduce the free-R-factor and improve overall stereochemistry.

Native-3 was not averaged with the other data sets, but was subjected to 20 cycles of solvent flattening (46% solvent) and 2-fold NCS averaging, during which the averaging radius was reduced and the resolution extended from 2.9 Å to 2.65 Å. The resulting map showed clear densities for the retinal and carbohydrate residues, the palmitoyl modifications to CysS22 and CysS23, as well as a few tightly bound water and detergent molecules. The 1F88 model used for molecular replacement was extensively rebuilt, especially the side-chain conformations. Coordinates for the retinal and carbohydrate moieties and cytoplasmic loops were taken from the partially refined Native-2 model and rebuilt. Refinement against Native-3 data was done using simulated annealing and individual B-factors, under NCS restraints imposed over the helical segments. When the free-R-factor decreased to 24.5%, ordered water molecules were located in the difference map using the water_pick task file in CNS. After another round of refinement with individual B-factors, NCS copies of the bound water molecules were located in the difference map using the water, F\textsubscript{o} - F\textsubscript{c} map and confirmed in the solvent-flattened EDEN\textsuperscript{40} map (kindly calculated by Dr Luca Jovine). Detergent and phospholipid molecules were added during refinement.

Geometric parameters for the 11-cis-retinal were taken from the small molecule structure,\textsuperscript{20} except the 12-s-cis bond was made trans, and parameters for the protonated Schiff base were taken from the retinylidene structure.\textsuperscript{21} In early rounds of the refinement, dihedral angles in the polyene section of retinal were tightly constrained to the plane using energy constants for the tropothen ring (500 kcal/mol), but for dihedrals about the C6-C7, C11-C12 and C12-C13 bonds the energy constants were reduced tenfold. After addition of water molecules, the tight dihedral restraints along the polyene were relaxed to 100 kcal/mol.

Initial models for the detergent molecules LDAO and C\textsubscript{12}E\textsubscript{4}, and the hydrocarbon chains of phospholipids were taken from structures in the Protein Data Bank as provided by the HIC-UP web site. Topology parameters for these groups and for the triester linkage of palmitoylated cysteine were derived from structures in the Cambridge Small Molecule Database.

Molecular replacement from an EM model

From 2D crystals of bovine rhodopsin imaged by electron cryomicroscopy, a three-dimensional map has been calculated to 5 Å resolution in the plane and 13.5 Å normal to it.\textsuperscript{10} The seven transmembrane helices appear as density rods of unknown azimuthal orientations. To reduce the distortion caused by the missing cone in data collected from 2D crystals, this map was subjected to one cycle of solvent-flattening using the program dm\textsuperscript{139} and a solvent content of 75%, which was estimated from the volume per molecule in the P\textsubscript{3} crystal\textsuperscript{110} and a unit cell thickness of 200 Å in the 2D crystal.

The solvent-flattened EM map was re-calculated to 6 Å resolution for skeletonisation.\textsuperscript{76} The bones within the boundary of a single rhodopsin molecule were edited in O to follow the high-density ridges along the helix axes, and they were annotated in O for the helices represented, in accordance with the model of helix arrangement proposed by Baldwin et al.\textsuperscript{12} Using the bone_mask option in O and a 6 Å radius for the bones atoms, a molecular mask was formed. The mask was edited using MAMA\textsuperscript{143} to remove internal cavities, and converted to the CCP4 format on a 1 Å sampling grid using MAMA2cpp4.\textsuperscript{136}

The solvent-flattened EM map was interpolated onto the 1 Å grid inside the molecular mask using MAPROT,\textsuperscript{142} to yield the masked density of a single rhodopsin molecule. The centre of gravity for the masked region was shifted to the origin of a P\textsubscript{1} unit cell that has 150 Å cell edges and 90° cell angles, by editing the map header using the program IMEDIT in the MRC image processing suite.\textsuperscript{145} The asymmetric unit was filled with zeros outside the mask, using the program MAPMASK.\textsuperscript{136} This density of a single rhodopsin molecule, cut out from the EM map, and placed in a large P\textsubscript{1} cell was the "EM model" for molecular replacement.

Structure factors for the EM model were calculated using SFALL.\textsuperscript{139} Cross-rotation function was calculated using GLRF,\textsuperscript{144} in which the best signal-to-noise ratio was obtained by using a 28 Å radius of Patterson integration and a resolution limit of 9-5 Å. Using MAPROT\textsuperscript{142} again, the EM model was rotated about its centre of gravity to the cross-rotation angles, and re-interpolated into the actual P\textsubscript{3} unit cell at 1 Å sampling intervals. Translation searches were conducted using TFFC\textsuperscript{145} in both P\textsubscript{3} and P\textsubscript{3}2\textsubscript{1}2\textsubscript{1}2\textsubscript{1} unit cells, but the restraints were relaxed in the loop regions.

The solvent-flattened EM map was re-calculated to 6 Å resolution for skeletonisation,\textsuperscript{76} The bones within the boundary of a single rhodopsin molecule were edited in O to follow the high-density ridges along the helix axes, and they were annotated in O for the helices represented, in accordance with the model of helix arrangement proposed by Baldwin et al.\textsuperscript{12} Using the bone_mask option in O and a 6 Å radius for the bones atoms, a molecular mask was formed. The mask was edited using MAMA\textsuperscript{143} to remove internal cavities, and converted to the CCP4 format on a 1 Å sampling grid using MAMA2cpp4.\textsuperscript{136}

The solvent-flattened EM map was interpolated onto the 1 Å grid inside the molecular mask using MAPROT,\textsuperscript{142} to yield the masked density of a single rhodopsin molecule. The centre of gravity for the masked region was shifted to the origin of a P\textsubscript{1} unit cell that has 150 Å cell edges and 90° cell angles, by editing the map header using the program IMEDIT in the MRC image processing suite.\textsuperscript{145} The asymmetric unit was filled with zeros outside the mask, using the program MAPMASK.\textsuperscript{136} This density of a single rhodopsin molecule, cut out from the EM map, and placed in a large P\textsubscript{1} cell was the "EM model" for molecular replacement.

Docking X-ray coordinates to the EM map

The C\textsuperscript{\alpha} coordinates of the 1F88 model and of the refined structure from this work were first manually aligned with the "bones" for the EM map, which trace out the helix axes. Then using the RSR\_rigid option in O, in the convolution mode, the fit of all atoms of the protein molecule to the EM density was optimised.

Protein Data Bank accession numbers

Coordinates and structure factors have been released through the RCSB Protein Data Bank (code 1GZM, 1GZM-SP).

Acknowledgements

We thank the ESRF, SPring-8, APS and Daresbury Laboratory for providing synchrotron radiation facilities and staff support during data collection,

\textsuperscript{†} http://alpha2.bmc.uu.se/hicup/
and in particular the ESRF for a long-term award to G.F.X. for use of the microfocus beamline ID13. We are grateful to Drs Richard Henderson, Paul Hargrave, Joyce Baldwin, Harry Powell and Alexei Murzin for helpful discussions; Luca Jovine for calculating the EDEN map from partially refined coordinates; Richard Henderson, Phil Evans, Andrew Leslie and Dan Opriam for comments on the manuscript; colleagues at the MRC Laboratory of Molecular Biology for encouragement. We acknowledge the use of the GPCRDB data base.

References

32. McDowell, J. H., Ubel, A., Brown, R. A. & Hargrave,
Rhodopsin Structure in Trigonal Crystals


120. Kim, J. M., Altenbach, C., Thurmond, R. L., Khorana,


Edited by R. Huber