Structure of the Interhelical Loops and Carboxyl Terminus of Bacteriorhodopsin by X-ray Diffraction Using Site-Directed Heavy-Atom Labeling†

Wolfgang Behrens,‡ Ulrike Alexiev,‡ Ramin Mollaaghababa,§ H. Gobind Khorana,§ and Maarten P. Heyn*,‡

Biophysics Group, Department of Physics, Freie Universität Berlin, D-14195 Berlin, Germany, and Departments of Biology and Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Received July 16, 1997; Revised Manuscript Received April 7, 1998

ABSTRACT: The positions of single amino acids in the interhelical loop regions and the C-terminal tail of bacteriorhodopsin (bR) were investigated by X-ray diffraction using site-directed heavy-atom labeling. Since wild-type bR does not contain any cysteines, appropriate cysteine mutants were produced with a unique sulfhydryl group at specific positions. These sites were then labeled with mercury using the sulfhydryl specific reagent *p*-chloromercuribenzoate (p-CMB). The cysteine mutants D96A/V101C, V130C, A160C, and G231C were derivatized with labeling stoichiometries of 0.93 ± 5%, 0.85 ± 5%, 0.79 ± 7%, and 0.77 ± 8%, respectively (Hg per bR). No incorporation was observed with wild-type bR under the same conditions. All mutants and heavy-atom derivatives were fully active as judged by the kinetics of the photocycle and of the proton release and uptake. Moreover, the unit cell dimensions of the two-dimensional P_3 lattice were unchanged by the mutations and the derivatization. This allowed the position of the mercury atoms, projected onto the plane of the membrane, to be calculated from the intensity differences in the X-ray diffraction pattern between labeled and unlabeled samples using Fourier difference methods. The X-ray diffraction data were collected at room temperature from oriented purple membrane films at 100% relative humidity without the use of dehydrating solvents. These native conditions of temperature, humidity, and solvent are expected to preserve the structure of the surface-exposed loops. Sharp maxima corresponding to a single mercury atom were found in the difference density maps for D96A/V101C and V130C. Residues 101 and 130 are in the short loops connecting helices C/D and D/E, respectively. No localized difference density was found for A160C and G231C. Residue 160 is in the longer loop connecting helices E and F, whereas residue 231 is in the C-terminal tail. Residues 160 and 231 are apparently in a more disordered and mobile part of the structure.

Investigations on the structure of the light-driven proton pump bacteriorhodopsin (bR)[1] by diffraction methods are facilitated by the organization of these proteins in a natural two-dimensional lattice, the so-called purple membrane. Three-dimensional structures of bR were obtained by electron cryocrystallography using tilted specimens consisting of single purple membranes (*1*−*3*). The refined structure of (*2*) has a resolution of 3.5 Å in the plane of the membrane and 4.3 Å perpendicular to the membrane. A higher resolution of 3 Å was recently achieved by using larger tilt angles (*3*). A three-dimensional structure with a resolution of 2.5 Å was also obtained by X-ray crystallography using microcrystals (*4*). These three structures (*2*−*4*) show a high degree of similarity in the transmembrane helical regions as judged by the root-mean-square deviations. The retinal chromophore with its protonated Schiff base linkage to the protein is the key functional element of this proton pump. The in-plane and transmembrane positions of the chromophore were determined by neutron diffraction with selectively deuterated retinals (*5*−*7*). The recent high-resolution structures (*2*−*4*) are in good agreement with the results of these earlier studies on the chromophore location. Detailed high-resolution structural information on the configuration of the chromophore was obtained from resonance Raman (*8*) and solid-state NMR spectroscopy (*9, 10*).

The interhelical loops and C- and N-terminal tails are not commonly believed to play an important role in the function of bR. However, protons are released and taken up on the extracellular and cytoplasmic surfaces, respectively. The release group is currently thought to be Glu 204 at the extracellular end of helix G near the FG loop (*11*). The uptake group is unknown, but prominent candidates are Asp 102 and Asp 104 in the CD loop on the cytoplasmic side just above the proton donor Asp 96 in helix C. Transient surface charge changes were recently reported with a pH-indicator dye at position 101 in the CD loop (*12*). Another residue involved in proton uptake is Asp 38 in the AB loop (*13*). The structures of the cytoplasmic half of bR and the loops on that side change in the M- and N-intermediates (*14, 15*) presumably as the protein opens up to let in water

† This work was supported by Grant GM 28289 from the National Institutes of Health (H.G.K.) and by Grant 03-HE4 FUB3 from the Bundesministerium für Bildung und Forschung (M.P.H.).

‡ Freie Universität Berlin.

§ Massachusetts Institute of Technology.

[1] Abbreviations: bR, bacteriorhodopsin; p-CMB, *p*-chloromercuribenzoate; V130C, G231C, and A160C, single mutants of bacteriorhodopsin with valine 130, glycine 231, and alanine 160 replaced by cysteine; D96A/V101C, double mutant of bacteriorhodopsin with aspartate 96 replaced by alanine and valine 101 by cysteine; V130C-MB, G231C-MB, A160C-MB, and D96A/V101C-MB, mercuribenzoate derivatives of the corresponding cysteine mutants.

S0006-2960(97)01735-2 CCC: $15.00
Published on Web 06/26/1998

molecules. Thus, an investigation of the loop structure in the unphotolyzed state and the intermediates is clearly of interest. Little is known with certainty about the structure of the interhelical loops and C- and N-terminal tails, however, since these surface-exposed regions are disordered, mobile, and susceptible to environmental perturbations. The three model structures (*2*−*4*) differ markedly in the structure of the loops, probably because the different specimen conditions employed perturb the surface structure to different extents. In (*3*), the membrane specimens were in the presence of trehalose and were rapidly frozen. This treatment may have prevented the specimens from being completely dehydrated (*3*). These authors observed a well-defined structure for each loop (*3*). In (*2*), completely dried specimens in the presence of glucose were used. These authors found very high temperature factors for the AB, BC, and EF loops, with corresponding uncertainty in the structure. For the CD, DE, and FG loops, the temperature factors were smaller and the atomic coordinates were more certain. There are major differences between (*2*) and (*3*) in the structures of the AB, BC, and EF loops. In (*4*), no density was obtained for the EF loop, but the structures of the AB, BC, and DE loops were again very different from those in (*2*). The X-ray diffraction work was performed with densely packed three-dimensional crystals with crystal contacts involving the AB and BC loops of adjacent layers of molecules and intratrimer contacts involving the AB and CD loops (*4*). The electron cryocrystallography data were collected at 153 K (*2*) and liquid helium temperature (*3*), respectively, with grids in the vacuum of the microscope. The X-ray diffraction data were obtained with flash-frozen crystals at 100 K. It is quite likely that the different sample conditions of solvent, water content, and temperature as well as the altered loop−loop contacts in the crystals will affect the soft surface structure to different degrees, thereby explaining the major differences in these studies. The excellent agreement for the transmembrane part of the structure is in full accordance with this argument, for this densely packed hydrophobic region is not expected to be perturbed by these external parameters. In view of these contradictory results for the loops, it is clearly desirable to carry out structural investigations under native conditions of temperature, solvent, humidity, and protein−protein contacts, as performed in this study.

BR is the prototype of the large class of transmembrane proteins with the structural motif of a seven α-helical bundle. The related retinal protein rhodopsin is a G-protein-coupled receptor. Activation and recognition by the G-protein are believed to involve the cytoplasmic loops and C-terminal tail of rhodopsin (*16*). The structure and dynamics of the surface loops are clearly of major significance for this large class of heptahelical receptors. It is thus appropriate to develop new methods designed to study the structure of these surface-exposed loops using bR as the model system.

In this study, we introduce a mercury atom by site-directed heavy-atom labeling using a cysteine mutation. The position of this label is then determined by Fourier difference methods using the phase information from electron crystallography. Since the labeling is site-specific, the interpretation of the Fourier difference map, based on the known transmembrane structure of bR, is very simple. The labeling strategy is particularly promising for studying structural changes in the various intermediates of the functional cycle.

The method of using localized heavy-atom labels is alternative and complementary to electron and X-ray crystallography in which the structure of the whole protein is determined. One significant advantage of our method, in view of the controversial results obtained for the loops so far (*2*−*4*), is that the experiments are performed at room temperature and 100% relative humidity. The samples consist of thin fully hydrated films of stacked purple membranes. From the lamellar diffraction of such films at 100% relative humidity, it is known (*7*) that there is a water gap of approximately 10 Å separating the adjacent membranes, thus preventing loop−loop contacts. The electron crystallography experiments, on the other hand, are carried out at liquid nitrogen or helium temperature with partially or fully dehydrated glucose- or trehalose-embedded membranes in a vacuum. The X-ray diffraction experiments are performed with flash-frozen densely packed crystals in which crystal contacts perturb the loop structure. These features of the experimental conditions and specimen preparation are not expected to have significant effects on the transmembrane part of the structure, but may well affect the structure and mobility of the surface-exposed loops.

In a previous communication, we already demonstrated the feasibility of this strategy by mercury-labeling A103C in the CD loop (*17*). Here we extend these investigations to residues V101, A160, and G231 on the cytoplasmic side and V130 on the extracellular side of the membrane. The labeled residues are indicated by circles in Figure 1. V101 and V130 are in the CD and DE loop, respectively. These loops are quite short (6 and 8 residues, respectively), and, accordingly well-defined positions may be expected. A160 is in the 8 residue long EF loop, which may be more mobile and disordered. G231 is in the beginning of the C-terminal tail.

MATERIALS AND METHODS

Construction of the bR Mutants V130C, A160C, G231C, and D96A/V101C. The bacteriorhodopsin mutant genes were constructed and expressed in *H. salinarium* as described (*17*, *18*). The preparation of the bR mutants A160C and G231C has been reported (*19*). The mutants V130C and D96A/V101C were prepared by using the synthetic restriction fragments *Bsp*HI−*Sph*I and Asp718−*Bsp*HI, respectively, for construction of the mutant genes.

Heavy-Atom Labeling. The single-cysteine mutants were derivatized with *p*-chloromercuribenzoate (p-CMB) as described (*20*, *17*), with minor changes. The bR mutants were isolated and purified in solutions containing 1 mM DTT. Prior to derivatization, the purple membranes were suspended in 30 mM NaP_i, 150 mM NaCl, pH 7, at a final concentration of 10−20 mM bR. The labeling reaction was performed with a 1.5−10-fold molar excess of p-CMB in the dark at room temperature. The maximum of the difference spectrum between the mercaptide and the unbound p-CMB is at 250 nm with $\Delta\epsilon = 7600\ M^{-1}\ cm^{-1}$. The time course of the labeling reaction was therefore monitored by recording the absorbance change at 250 nm. When the reaction reached completion, the excess p-CMB was removed by chromatography on a Sephadex G25 column preequilibrated and eluted with 30 mM NaP_i, pH 7.0. This was followed by two washes with MilliQ (Millipore) water at 4 °C. A control

sample was prepared in the same way except that no p-CMB was added. After removal of the excess unbound p-CMB, the final labeling stoichiometry was determined from the absorbance difference between the light-adapted labeled and control samples at 238 nm. This is the absorption maximum of the mercaptide. The stoichiometry (moles of mercuribenzoate per mole of bR) was calculated using

$$c_{Hg}/c_{BR} = (\Delta A_{238}/\epsilon_{238})(\epsilon_{568}/A_{568}) \quad (1)$$

c_{Hg} and c_{BR} are the molar concentrations of the bound mercuribenzoate and bR, respectively. ΔA_{238} is the absorbance difference at 238 nm ($\epsilon_{238} = 20\,700\ M^{-1}\ cm^{-1}$). A_{568} is the absorbance of the light-adapted bR sample at 568 nm ($\epsilon_{568} = 63\,000\ M^{-1}\ cm^{-1}$). The UV/VIS absorption spectra were recorded with a Perkin-Elmer or a Shimadzu 2105 spectrometer equipped with an integrating sphere.

Flash Spectroscopy. The kinetics of the photocycle and of proton release were measured in 150 mM KCl, pH 7.3, 22 °C, as described (*21*).

Sample Preparation. The labeled and control (unlabeled) samples were centrifuged, and each pellet was dissolved in a small volume of MilliQ water to form a concentrated bR suspension (~20 mg of bR/mL). To avoid differences in the diffraction patterns due to differences in sample treatment, all operations on the labeled and unlabeled samples were performed in parallel. A concentrated drop of the suspension was applied to a thin Mylar support ($d = 6\ \mu m$) which was stored in a sealed box at 86% relative humidity. Under these conditions, bR patches form oriented multilamellar films when the excess water evaporates. When the drop became flat, another drop was applied right on top of it. An oriented purple membrane film was thus produced covering an area of approximately 7 mm^2 and containing about 1.0 mg of bR.

After film formation, the sample was stored in a box with 100% relative humidity to prevent dehydration. The sample was mounted in the sample holder between two Mylar windows 24 h before X-ray measurements. The sample was kept in the sample holder at 100% relative humidity.

X-ray Diffraction. In-plane powder diffraction patterns were recorded with a one-dimensional position-sensitive detector (Braun OED-50-M) in the radial direction. A rotating anode X-ray generator (Marconi-Elliott GX21) at 1.542 Å was used as the X-ray source. The hydrated purple membrane samples were mounted with the film plane perpendicular to the X-ray beam. A helium chamber between the sample and the detector reduced the scattering of the primary beam in air. Each sample was allowed to equilibrate for 24 h in the constant humidity sample holder before measurement. For each sample, about 30 data collection runs of 1 h were recorded and divided into 2 data sets. The two data sets were analyzed independently to check for the reproducibility of the experiment. All measurements were performed at room temperature in the dark-adapted state.

Data Analysis. For both the labeled and unlabeled samples, the same number of individual 1 h runs were summed up to one spectrum. The individual runs were first checked for errors. Individual counts in every Bragg reflection were judged statistically.

Background subtraction was performed by fitting a 7 or 8 degree polynomial to the data. The intensities were derived from the Bragg reflections by fitting Gaussian line shapes to the diffraction peaks. Bragg reflections with the same value for $(h^2 + hk + k^2)$ occur at the same Bragg angle and contribute to the same powder Bragg reflection (Debye−Scherrer ring). These intensities were split using intensity ratios from electron microscopy (*22*). In the refinement procedure, this constraint was dropped. The justification for this procedure has been discussed in detail (*23*). The validity of this approach was confirmed experimentally (*5*, *24*). All intensities were corrected by a Lorentz factor of $(h^2 + hk + k^2)^{1/2}$.

The Fourier synthesis was calculated using the experimental structure factors and the phases from electron microscopy (*22*). The position found for the label was used as the starting point for the refinement procedure (*25*). To check whether the intensity differences are due to one mercury atom per molecule of bR, model calculations were performed.

RESULTS

Stoichiometry of Heavy-Atom Labeling. The four cysteine mutants V130C, A160C, G231C, and D96A/V101C were used for heavy-atom labeling and X-ray diffraction. The positions of the amino acid residues changed to cysteine are indicated in Figure 1 by circles. The isolated mutant purple membranes are similar to wild-type purple membranes both in density and in crystallinity as checked by density centrifugation and X-ray diffraction (see below).

The sulfhydryl group of cysteine is known to react highly specifically with mercury-containing reagents such as p-CMB, forming a mercaptide (*20*). p-Chloromercuribenzoate (p-CMB) undergoes a change in the ultraviolet absorption spectrum of the benzoate upon mercaptide formation with cysteine. The maximal difference absorption between unbound p-CMB and the mercaptide occurs at 250 nm. The absorbance change at this wavelength was used to follow the label incorporation during the reaction in the presence of a large excess of free p-CMB. The cysteine mercaptide has its maximal absorbance at 238 nm ($\epsilon = 20\,700\ M^{-1}\ cm^{-1}$), and the absorbance at this wavelength was used to calculate the labeling stoichiometry using eq 1 after removal of the excess free p-CMB. Figure 2A shows as an example the absorption spectra of purified A160C (control) and A160C-MB. The difference spectrum is plotted in Figure 2B and shows an absorption maximum between 230 and 240 nm as expected for the mercaptide. The labeling stoichiometries determined using eq 1 were 0.85 ± 5%, 0.79 ± 7%, 0.77 ± 8%, and 0.93 ± 5% for V130C-MB, A160C-MB, G231C-MB, and V101C-MB/D96A, respectively. No reaction was observed for wild-type bacteriorhodopsin under the same conditions.

The cysteine−mercuribenzoate complex is known to be unstable under certain conditions (*26*). The labeling stoichiometry was therefore determined both before the X-ray sample preparation (ca. 1 h after the reaction and the removal of the excess reagents were completed) and after the X-ray measurements on an aliquot that was not exposed to the X-ray radiation. Figure 2A shows the spectrum of A160C-MB directly after labeling. Figure 2C shows the spectrum of the A160C-MB sample before and after the X-ray experiments. From a comparison of these spectra, we

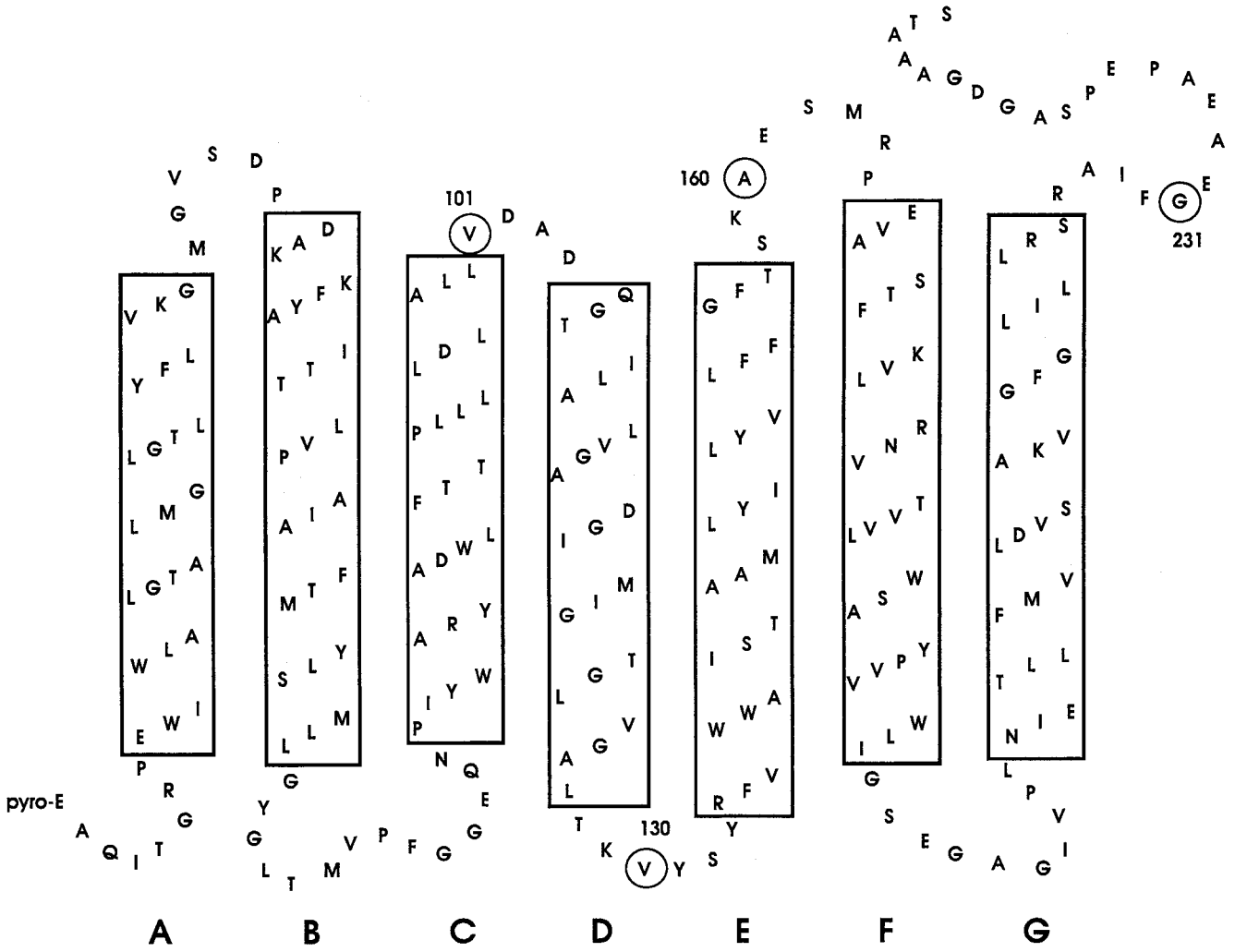

FIGURE 1: Secondary structure of bacteriorhodopsin based on the model of (*2*). Boxed regions labeled A−G correspond to the seven transmembrane α-helices. The circled amino acid residues 101, 130, 160, and 231 were changed to cysteine in this study.

conclude that no label loss occurs in the time interval between the labeling and the completion of the X-ray diffraction experiments.

Absorption Spectra, Photocycle, and Proton Pumping. The single-cysteine mutants have absorption spectra which are very similar to those of wild-type (see, e.g., Figure 2 for A160C). The kinetics of the photocycle and of proton release and uptake were measured for the mutants before and after derivatization to see if changes due to the labeling or with respect to wild-type occur. The kinetics of the p-CMB-labeled samples were very similar to those for the unlabeled mutants. For A160C, G231C, and V130C, these data have already been reported (*19*). As an example of the very minor effect of labeling, Figure 3 shows the time course of the absorbance changes at 410 and 650 nm for V130C and V130C-MB. The rise and decay of the M and O intermediates contribute mainly at these wavelengths. A global fit of the photocycle kinetics with seven exponentials led to photocycle times of 1.4 μs, 43.9 μs, 167 μs, 740 μs, 1.9 ms, 7.7 ms, and 47 ms for V130C-MB versus 1.2 μs, 36.2 μs, 129 μs, 680 μs, 1.9 ms, 6.5 ms, and 38 ms for V130C. These photocycles are identical within experimental error. Similar measurements were carried out with D96A/V101C and D96A/V101C-MB at pH 6.95. A global fit with six exponentials led to the following time constants for D96A/V101C-MB and the unlabeled double mutant (in parentheses): 0.88 μs (0.93 μs), 55 μs (52 μs), 410 μs (430 μs), 22 ms (18 ms), 1.0 s (1.1 s), and 2.0 s (2.2 s). Again the labeling has no significant effect on the photocycle kinetics. The kinetics of the decay of M are slowed with respect to wild-type, an effect well-known for the D96A single mutant at this pH (*21*).

The proton release times detected with the pH indicator dye pyranine in the aqueous bulk phase are also similar (830 and 785 ms for V130C and V130C-MB, respectively) and close to the value for wild-type (data not shown).

X-ray Diffraction. The diffraction patterns of the labeled and unlabeled samples were indexed on a hexagonal lattice of P_3 symmetry. The unit cell dimensions for D96A/V101C, D96A/V101C-MB, V130C, V130C-MB, A160C, and A160C-MB were 62.3 ± 0.2 Å. For G231C and G231C-MB, the lattice constants were 62.4 ± 0.2 Å. These values are, within experimental error, equal to the value of 62.4 Å established for wild-type purple membrane (*27*).

For each sample, Bragg reflections were recorded out to the (7,1) reflection which corresponds to a resolution of 7.2 Å. Since the amount of scattering material in the beam cannot be exactly the same for the labeled and unlabeled samples, the observed intensities first have to be scaled such that the sum of the intensities is the same. For D96A/V101C, V130C, G231C, and A160C, the scaling factors were 1.40, 1.13, 1.07, and 1.75, respectively. Clear-cut differences in

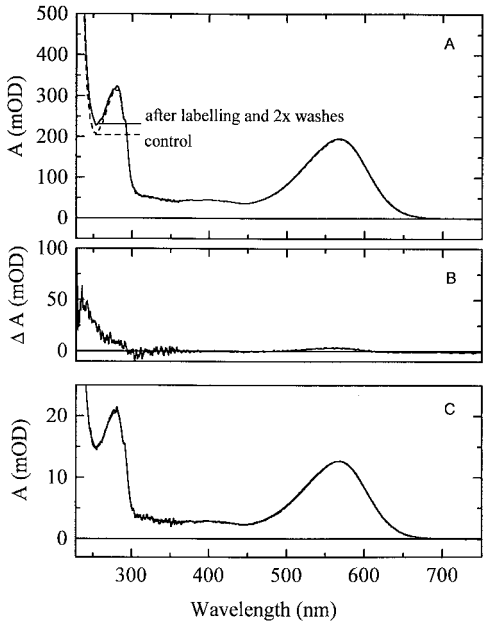

FIGURE 2: (A) Light-adapted absorption spectra of A160C (---) and A160C-MB (—). (B) Difference spectrum obtained by subtracting the two spectra in (A) (stoichiometry 0.78). (C) Light-adapted absorption spectra of the same A160C-MB aliquot taken just after labeling and at a time after completion of the X-ray measurements (two spectra superimposed). The aliquot was not exposed to the X-ray radiation.

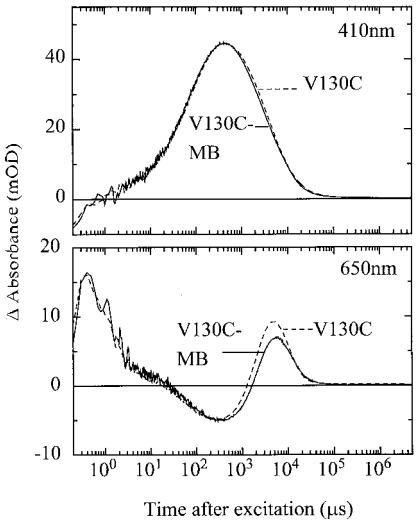

FIGURE 3: Photocycle kinetics of V130C (---) and V130C-MB (—) monitored at 410 nm (top panel) and 650 nm (bottom panel). Conditions: pH 7.0, 150 mM KCl, 22 °C.

intensity could be observed for the samples D96A/V101C-MB and V130C-MB. These differences are due to the mercury label which gives rise to positive and negative changes in the scattered intensities.

This can easily be seen by comparing the corresponding reflections of the D96A/V101C and D96A/V101C-MB samples in Figure 4. The most prominent positive change in intensity is detected in the (2,0) reflection whereas the most negative change was observed in the (4,0) reflection. The overall relative intensity changes $\sum|\Delta I|/\sum I$ after Lorentz correction were calculated to be 6.9% and 7.2% for the two independent experiments. The correlation of both experiments was judged by computing the correlation coefficient as described (*27*). The calculated correlation coefficient was

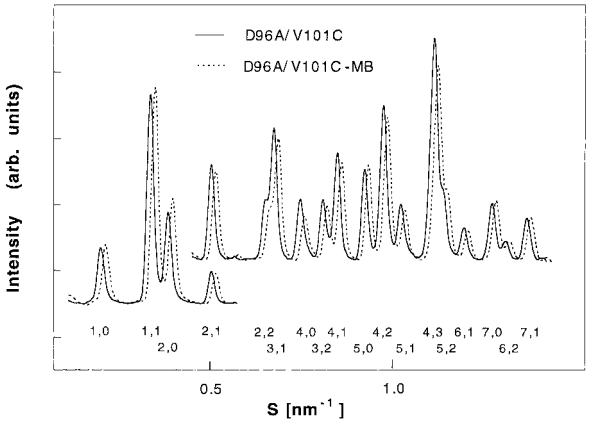

FIGURE 4: X-ray diffraction intensities of D96A/V101C (—) and D96A/V101C-MB (---) before Lorentz correction as a function of the length of the scattering vector $S = 2 \sin \theta/\lambda$. The Bragg reflections belong to the two-dimensional hexagonal lattice of P_3 symmetry and are labeled with indices (h,k). Beginning with the (2,1) reflection, the vertical scale has been expanded by a factor of 3. For a better presentation, the diffraction pattern of the labeled sample is slightly displaced to the right.

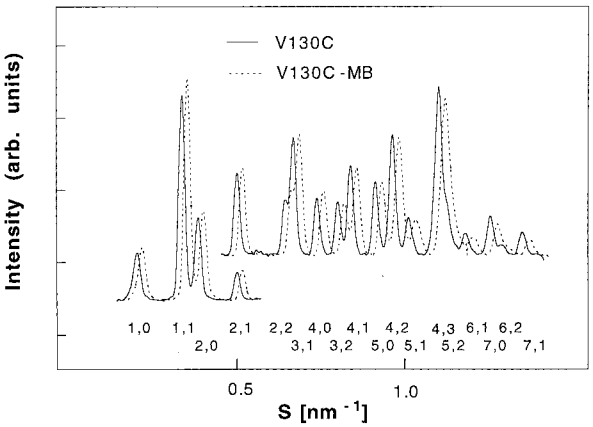

FIGURE 5: X-ray diffraction patterns of V130C (—) and V130C-MB (---). Refer to Figure 4 for conditions.

0.94 which indicated a high degree of similarity between the experiments.

The scaled diffraction patterns for the V130C and V130C-MB samples before Lorentz correction are shown in Figure 5. The most noticeable positive change in intensity occurs in the (4,0) reflection, whereas the most negative changes in scattered intensity were observed in reflections (7,0) and (7,1). The relative changes in intensity after Lorentz correction were 7.4% and 6.9% in two independent experiments. This resulted in a high sample correlation coefficient of 0.95 for both experiments.

For the samples A160C-MB and G231C-MB, no clear-cut differences in intensities were observed. Figure 6 shows the intensities for the samples G231C and G231C-MB. Only minor changes were observable. The difference spectrum calculated from the labeled and unlabeled samples is also plotted in Figure 6. Positive and negative changes can easily be recognized in this plot. Only small changes could be detected in both independent experiments. The overall relative changes in intensities were 2.9% and 3.1%, respectively, for both experiments. In experiments where only small changes are detectable, the modeled background has a larger effect on the result than in experiments where greater

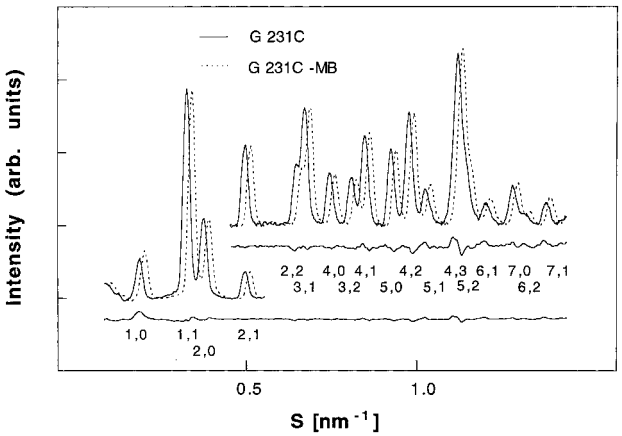

FIGURE 6: X-ray diffraction patterns of G231C (—) and G231C-MB (---). To help with recognition of the small intensity differences, the difference was calculated and plotted under the two patterns. Scale expansion at (2,1) is as in Figure 4.

changes are observable. Therefore, the background subtraction for these samples was performed more rigorously. The subtraction was done several times with slightly different backgrounds to check for the effect on the small changes in intensities. These various tests did not increase the overall intensity changes significantly. The calculated correlation coefficient for the two independently analyzed data sets for samples G231C and G231C-MB was 0.79. This indicates good reproducibility, even though it is not as high as reached for the D96A/V101C and V130 samples. The corresponding correlation coefficient for the A160C and A160C-MB samples was only 0.25, suggesting poor reproducibility.

Refinement of the p-CMB Label Position. The intensities of the Bragg reflections were calculated from each spectrum for all samples, as described under Materials and Methods. Fourier difference maps were computed from the structure factor differences, and the label positions found for D96A/V101C and V130 in these maps were then used as the starting points for the refinement procedure. Fractional coordinates are defined with the origin in the center of the unit cell and with x and y increasing along the oblique and horizontal axes from the top to bottom and from left to right, respectively (see Figure 7). In these fractional coordinates, the true x and y values are divided by the unit cell dimension (62.2 Å). The two corners of the unit cell shown in Figure 7 then have fractional coordinates of (1/2, 1/2) (lower right corner) and (−1/2, −1/2) (upper left corner). In these coordinates, the refined positions for the p-CMB label from sample D96A/V101C were (0.335, 0.178) and (0.337, 0.187) for the two independent experiments. The difference in position was only $\Delta d = 0.89$ Å and accounts for the excellent reproducibility of the experiments. The refined label position did not depend on the starting values (x_0, y_0) for the refinement. Points on a circle of 7 Å diameter around the position found in the Fourier difference map were chosen for refinement, and the final position was independent of the chosen starting position. Figure 7 shows the refined electron difference density peak and two contour lines outlining the projected density of the protein. The label position found is in the region between the cytoplasmic ends of helices C and E. The refinement procedure also yields a value for the temperature factor which is a measure of the mean square displacement of the label from its position. With

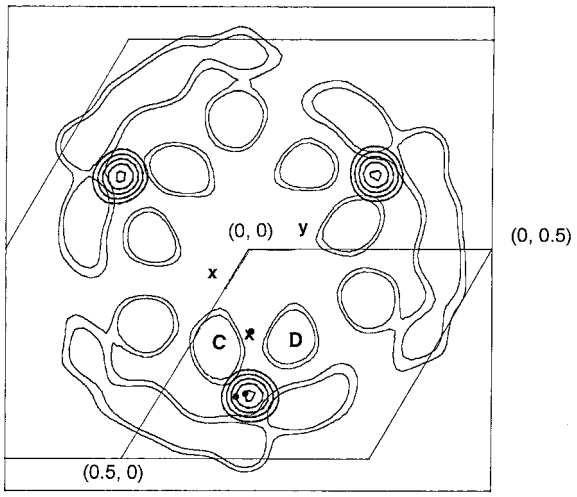

FIGURE 7: Two-dimensional difference density map showing the in-plane position of the mercuribenzoate label bound at Cys 101 of the bR mutant D96A/V101C after refinement. Negative contour lines were omitted. The five boldface contour lines range from 53% to 97% of the positive difference density. The two thin contour lines represent the outer boundary of the projected density of bR. Helices C and D are labeled. The origin of the oblique coordinate system is in the center of the unit cell with positive x and y as indicated. There are three bR molecules per unit cell. The two black circles (●) superimposed on the difference density mark the positions of the C_γ atoms of V101 from (*2*). The cross (×) marks the mercury position of A103C-MB obtained previously (*17*). The black circle right above it marks the C_β atom of A103 according to (*2*).

the label at residue 101, we find a temperature factor of 192 Å^2.

For the V130C-MB label, the coordinates found by the refinement are (0.279, 0.257) and (0.275, 0.252) for the two experiments. The small difference of only 0.27 Å indicates excellent experimental reproducibility. The coordinates found in the refinement were independent of the starting values. This was tested in the same way as described for the D96A/V101C-MB label. Points chosen on a circle of 7 Å diameter as starting values for the refinement led to the same final coordinates. Figure 8 shows the refined electron density map for the label together with two contour lines outlining the boundary of the projected protein density. The refined position is clearly in the region between the extracellular ends of helices D and E where it is expected to be found. The temperature factor for this label position was 50 Å^2.

The Fourier difference map for samples G231C and G231C-MB is shown in Figure 9. For a proper comparison, the difference in electronic density was brought to the same scale as for the corresponding maps for D96A/V101C-MB and V130C-MB. No single maximum can be discerned in this map. Most of the electron difference density is at the noise level, which is due to the cutoff of the Fourier series and due to the inherent assumptions of the Fourier difference method. We note that the largest difference density is in the lipid phase.

For the label at position 160, similar intensity differences and Fourier difference maps were obtained as for position 231 (data not shown). We conclude that the lack of a definite position is due to disorder or mobility of this part of the EF loop under our experimental conditions (room temperature, 100% relative humidity).

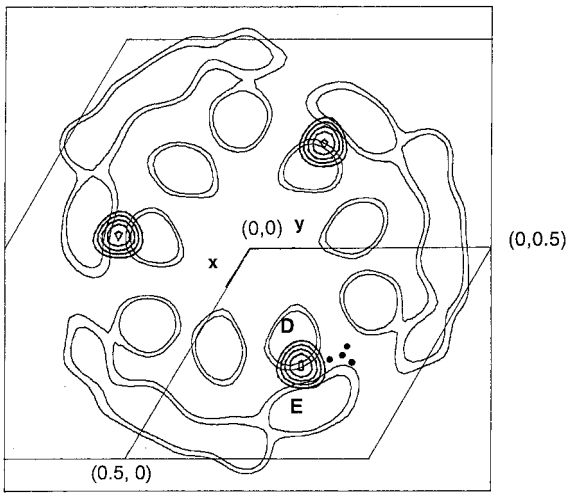

FIGURE 8: Two-dimensional difference density map showing the in-plane position of the mercuribenzoate label bound at Cys 130 after refinement. The five boldface contour lines range from 53% to 97% of the positive difference density. The two thin contour lines indicate the outer boundary of the projected density of bR. Helices D and E are labeled. The positions of the C_α, C_β, and C_γ atoms of V130 from (*2*) are marked by black circles (●), with the side chain pointing to the right.

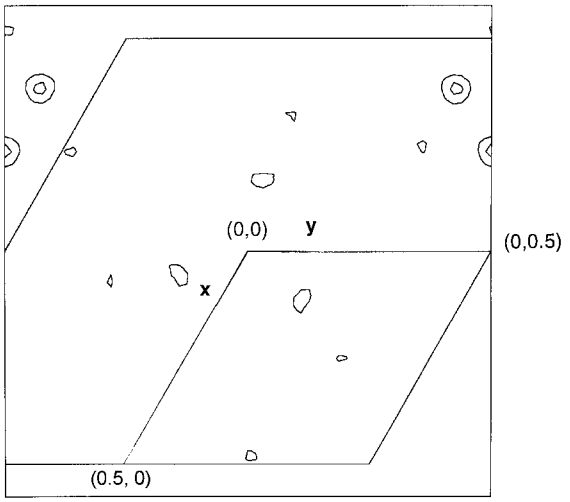

FIGURE 9: Two-dimensional Fourier difference map for the mercuribenzoate label bound at Cys 231. The four contour lines range from 53% to 97% of the positive difference density. To compare this map with those for the label at positions 101 (Figure 7) and 130 (Figure 8), the electron density difference was brought to the same absolute scale. The difference density is rather uniformly distributed throughout the unit cell, and no significant difference peak becomes apparent. The absolute maximum (upper right corner) is located in the lipid phase. Most of the difference density is at the noise level and is due to the cutoff error in the Fourier series and to the inherent limitations of the Fourier difference method.

DISCUSSION

In this study we have used site-directed heavy-atom labeling to incorporate mercury atoms at four defined positions in the interhelical loops and C-terminal tail of bacteriorhodopsin to determine their positions in the unphotolyzed state. The wild-type residues at positions 101, 130, 160, and 231 were replaced by cysteine, and the unique sulfhydryl groups were reacted with the reagent p-CMB to yield the corresponding *p*-mercuribenzoate derivatives. The mercury positions were then determined from the X-ray intensity differences between samples of the labeled and unlabeled cysteine mutants using the phases from electron cryomicroscopy.

Our experiments were carried out at room temperature, with samples that were fully hydrated and without using additives such as glucose (*2*) or trehalose (*3*). This is an important advantage over cryoelectron microscopy (*2*, *3*) which is carried out at low temperature (liquid nitrogen or helium) with air-dried glucose-embedded single membranes on a grid in vacuo. The specimen preparation employed in cryoelectron microscopy may well perturb the structure of the surface-exposed loops. Temperature, solvents (glucose, trehalose), and dehydration are expected to alter the dynamics and structure of the floppy and easily deformable loops that protrude from the surface. These fragile and flexible parts of the structure are expected to be very sensitive to these external parameters. The recent X-ray diffraction work was performed with flash-frozen densely packed three-dimensional crystals (*4*). The crystal contacts lead to some intramolecular loop−loop interactions which are absent in the purple membrane (*4*). In our sample with a 10 Å wide water layer between the membranes, these loop−loop contacts, which may well change the loop structure, are absent.

Another important advantage of our approach is that we determine the label position in a direct way. Despite the moderate resolution of our experiments, the position of the label density peak can be obtained with good precision. The limited resolution only leads to a broadening of the peak. In cryoelectron microscopy (*2*, *3*) and X-ray diffraction (*4*), on the other hand, the protein structure is modeled into the observed electron density in a global way. The protein *R*-factor, a measure of the refinement of the entire protein structure, is not very sensitive to the local structure of the loops, whose scattering density is rather weak.

For the labeling strategy to work, three criteria have to be met:

(1) The replacement has to be isomorphic. Since p-CMB is bulky and charged, it is important to check this point. We showed that neither the cysteine mutation nor the derivatization at these four surface-exposed positions had significant effects on the function of bR as a proton pump as judged by the kinetics of the photocycle and of proton release (e.g., see Figure 3).

(2) The unit cell dimensions before and after derivatization should be the same; otherwise, the amplitudes of the continuous Fourier transform of the density are sampled at different points. Moreover, the unit cell dimensions should be the same as for wild-type, so that the phase information from cryoelectron microscopy can be used. Our results show that this criterion is also met.

(3) The labeling stoichiometry should be as close as possible to 1, leading to a correspondingly high occupancy of the heavy-atom site. Low occupancy leads to positional disorder and a reduction in the amplitude of the Fourier difference peaks, making it more difficult to localize the label. We determined the labeling stoichiometries spectrophotometrically using the UV absorbance of p-CMB. The high stoichiometries of 0.93 ± 5% (D96A/V101C), 0.85 ± 5% (V130C), 0.79 ± 7% (A160C), and 0.77 ± 8% (G231C) are clearly adequate for these experiments.

We now discuss the results for each labeling site separately:

D96A/V101C. This position in the short cytoplasmic loop connecting helices C and D (101−104) is of particular interest, since we detected transient surface charge changes with a pH-indicator dye at this site which is close to the proton donor D96 (*12*). Moreover, experiments with a nitroxide spin-label attached at this position suggested movement of the CD loop associated with the decay of the M-intermediate (*15*). For this reason, the double mutant D96A/V101C was prepared which allows trapping the system easily in the M-intermediate. Figure 7 shows the refined position of the label. The CD loop apparently starts at the top of helix C on the inside of the helical bundle. The position found (0.336, 0.182) is actually closer to the end of helix E rather than to that of helix D. This will be discussed further below. Two independent experimental determinations led, within experimental error, to the same position. It is also in good agreement with our previous work in which we obtained the mercury position for A103C-MB (0.199, 0.105) using the same approach (*17*). Since we find well-defined and mutually consistent positions for two of the five amino acids of the CD loop (101, 103), we conclude that this short loop has a well-defined structure.

V130C. This residue is in the short extracellular DE loop (128−133). We conclude from Figure 8 that this loop label also has a definite position and is found precisely where expected, between helices D and E. The low-temperature factor of only 50 Å^2 is also consistent with an ordered structure for the DE loop.

A160C and G231C. No difference density was observed for these positions. Since the label stoichiometry was high, we conclude that these results are due to loop disorder or mobility. Position 160 is in the longer EF loop (158−165). Position 231 is in the long C-terminal tail (227−248) which is expected to be disordered.

It is of interest to compare our results obtained at room temperature and 100% relative humidity with those from cryoelectron microscopy and X-ray diffraction, which were obtained under very different conditions (*2*−*4*). In the introduction and the beginning of the Discussion, we have already argued that the large differences between the results of (*2*−*4*) for the loops are probably due to the different extents to which the surface structure is perturbed by dehydration, temperature, and loop−loop contacts. Agreement between our results and those of (*2*−*4*) is therefore not necessarily to be expected. In Figure 1, we have adopted the definition of the loop regions from (*2*). It was noted (*2*) that there remains an uncertainty of ±1 residue, and indeed there are differences of this magnitude in the starting and ending residues of the helices between the refined structure (*2*) and the previous structure (*1*) as well as (*4*). For all three structures (*2*−*4*), the coordinates have been deposited in the Protein Data Bank, but only for (*2*) are the coordinates currently available. For V101C-MB, the refined position for the mercury label (in reduced coordinates) is (0.336, 0.182). For the γ-carbon of V101, Grigorieff et al. (*2*) obtained coordinates of (0.324, 0.168). The agreement is excellent for this first residue of the short CD loop. The two C_γ positions from (*2*) are marked by black circles in Figure 7. Also shown in Figure 7 are the mercury positions for A103C-MB from (*17*) and the C_β of A103 from (*2*) marked by X and (●), respectively. The backbone chain connecting helices C and D starts on the interior of helix C and then loops out to the exterior side of helix D (*2*). The side chain of V101 points according to (*2*) toward helix E, and our label position is in complete accord with this finding.

For V130C-MB, we found a refined mercury position of (0.277, 0.255) which may be compared to (0.253, 0.339) from (*2*) for the valine γ-carbon. The positions of the V130 side chain carbons from (*2*) are marked by (●) in Figure 8. The agreement with the result of (*2*) is not as good for this position in the DE loop as for 101 and 103 in the CD loop. The distance between the label position and the C_α atom of V130 according to (*2*) is about 3 Å, i.e., well within the radius of 4.5 Å for an extended labeled cysteine side chain. The two positions are thus consistent as far as the main chain C_α atom is concerned, but the side chains appear to point to opposite sides of the main chain. There may thus be a true disagreement possibly due to differences in sample preparation and measuring conditions. One should keep in mind, however, that the valine side chain was replaced by the p-CMB-derivatized cysteine, which may take up a slightly different position. For A160C-MB, we found no density. In (*2*), defined coordinates for A160 were given, but the authors were unsure whether their very disordered structure for the EF loop was correct, observed no density for residues 164−166, and obtained a very high temperature factor of 350 Å^2. In (*4*), no density was observed for the complete EF loop (residues 157−166), in agreement with our result. In (*3*), on the other hand, excellent density was found for every residue of the EF loop. For G231C-MB in the C-terminal tail, we obtained no defined position. This is in agreement with (*2*) and (*4*), whereas in (*3*) residue 231 was the last one for which density was observed.

It is also of interest to compare the backbone temperature factors for the loops reported by Grigorieff et al. (*2*) with the ones we obtained for the labels. The lowest temperature factor of about 60 Å^2 was obtained for the DE loop. This value is only slightly outside the range for well-ordered helices, for which these authors found values between 10 and 50 Å^2. The temperature factor of 50 Å^2 reported here for the label attached at residue 130 of the DE loop is in excellent agreement with Grigorieff's results. The next lowest temperature factor of about 150 Å^2 and significant density were observed for the CD loop (*2*). This result is also consistent with our observation of a clear density peak for V101C-MB, the first residue of the short CD loop (101−104), for which we observed a temperature factor of 192 Å^2.

The dynamics of the C-terminal tail were studied by time-resolved fluorescence depolarization (*28*, *29*) and ^{2}H and ^{13}C NMR (*30*, *31*). Contradictory results were obtained concerning the mobility of the tail with some studies claiming isotropic reorientation on the available time scale and others claiming immobilization. These controversial results were recently reviewed (*32*). We point out here that the information obtained by X-ray or electron diffraction concerns the time-averaged structure. High mobility of loops or tail will lead to a smeared-out density and the lack of a clear-cut maximum in the Fourier-difference map.

The structure and dynamics of the cytoplasmic loops are of great functional significance for the large class of heptahelical receptor proteins. Upon activation, changes in the structure and dynamics of these loops occur exposing recognition sites for G-proteins in the G-protein-coupled

members of this family. Although not a receptor protein itself, it is interesting to note that the disorder of the loops and helices of bR is highest on the cytoplasmic side of the membrane (*2*). The cytoplasmic ends of helices B, F, and G are more disordered than the extracellular ends, with temperature factors as high as 150 Å^2. It was observed by cryoelectron microscopy that the cytoplasmic part of helix F moves outward in the M intermediate with a hinge, possibly at Pro 186 (*14*). This state of bR corresponds to the activated M_{II} form of rhodopsin. For such hinge motions of helices to occur, the loops on the cytoplasmic side of the membrane have to be flexible, as observed here for the EF loop. Labeled sites on the cytoplasmic ends of transmembrane helices at the interface with the loops may be most suitable to detect the rigid body tilt motion of helices. Position 101 at the end of helix C is an example.

REFERENCES

1. Henderson, R., Baldwin, J. M., Ceska, T., Zemlin, F., Beckmann, E., and Downing, K. H. (1990) *J. Mol. Biol. 213*, 899−929.
2. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M., and Henderson, R. (1996) *J. Mol. Biol. 259*, 393−421.
3. Kimura, Y., Vassylyev, D. G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T., and Fujiyoshi, Y. (1997) *Nature 389*, 206−211.
4. Pebay-Peroula, E., Rummel, G., Rosenbusch, J. P., and Landau, E. M. (1997) *Science 277*, 1676−1681.
5. Seiff, F., Wallat, I., Ermann, P., and Heyn, M. P. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3227−3231.
6. Heyn, M. P., Westerhausen, J., Wallat, I., and Seiff, F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2146−2150.
7. Hau*β*, T., Grzesiek, S., Otto, H., Westerhausen, J., and Heyn, M. P. (1990) *Biochemistry 29*, 4904−4913.
8. Althaus, T., Eisfeld, W., Lohrmann, R., and Stockburger, M. (1995) *Isr. J. Chem. 35*, 227−251.
9. Thompson, L. K., McDermott, A. E., Raap, J., van der Wielen, C. M., Lugtenburg, J., Herzfeld, J., and Griffin, R. G. (1992) *Biochemistry 31*, 7931−7938.
10. Ulrich, A. S., Watts, A., Wallat, I., and Heyn, M. P. (1994) *Biochemistry 33*, 5370−5375.
11. Brown, L. S., Sasaki, J., Kandori, H., Maeda, A., Needleman, R., and Lanyi, J. K. (1995) *J. Biol. Chem. 270*, 27122−27126.
12. Alexiev, U., Scherrer, P., Marti, T., Khorana, H. G., and Heyn, M. P. (1995) *FEBS Lett. 373*, 81−84.
13. Riesle, J., Oesterhelt, D., Dencher, N. A., and Heberle, J. (1996) *Biochemistry 35*, 6635−6643.
14. Subramaniam, S., Gerstein, M., Oesterhelt, D., and Henderson, R. (1993) *EMBO J. 12*, 1−8.
15. Steinhoff, H.-J., Mollaaghababa, R., Altenbach, C., Hideg, K., Krebs, M., Khorana, H. G., and Hubbell, W. L. (1994) *Science 266*, 105−107.
16. Hofmann, K. P., Jäger, S., and Ernst, O. P. (1995) *Isr. J. Chem. 35*, 339−356.
17. Krebs, M. P., Behrens, W., Mollaaghababa, R., Khorana, H. G., and Heyn, M. P. (1993) *Biochemistry 32*, 12830−12834.
18. Krebs, M. P., Mollaaghababa, R., and Khorana, H. G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1987−1991.
19. Alexiev, U., Mollaaghababa, R., Scherrer, P., Khorana, H. G., and Heyn, M. P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 372−376.
20. Boyer, P. D. (1954) *J. Am. Chem. Soc. 76*, 4331−4337.
21. Otto, H., Marti, T., Holz, M., Mogi, T., Stern, L. J., Engel, F., Khorana, H. G., and Heyn, M. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1018−1022.
22. Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J., and Zemlin, F. (1986) *Ultramicroscopy 19*, 147−178.
23. Plöhn, H.-J., and Büldt, G. (1986) *J. Appl. Crystallogr. 19*, 255−261.
24. Koch, M. H. J., Dencher, N. A., Oesterhelt, D., Plöhn, H.-J., Rapp, G., and Büldt, G. (1991) *EMBO J. 10*, 521−526.
25. Dickerson, R. E., Weinzierl, J. E., and Palmer, R. A. (1968) *Acta Crystallogr. B24*, 997−1003.
26. Epstein, C. C., and Datta, P. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4862−4866.
27. Glaeser, R. M., Jubb, J. S., and Henderson, R. (1985) *Biophys. J. 48*, 775−780.
28. Renthal, R., Dawson, H., Tuley, J., and Horowitz, P. (1983) *Biochemistry 22*, 5−12.
29. Marque, J., Kinosita, K., Govindjee, R., Ikegami, A., Ebrey, T. G., and Otomo, J. (1986) *Biochemistry 25*, 5555−5559.
30. Keniry, M. A., Gutowsky, H. S., and Oldfield, E. (1984) *Nature 307*, 383−386.
31. Bowers, J. L., and Oldfield, E. (1988) *Biochemistry 27*, 5156−5161.
32. Engelhard, M., and Bechinger, B. (1995) *Isr. J. Chem. 35*, 273−288.

BI971735S