the interaction of profilin with PIP_2 (Goldschmidt-Clermont et al., 1991). Such modulation of the PI-MAP2 interactions may constitute a link by which extracellular factors influence the microtubule cytoskeleton.

REFERENCES

Bhattacharyya, B., & Wolff, J. (1975) *J. Biol. Chem. 250*, 7639-7646.

Blitz, A. L., & Fine, R. E. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4472-4476.

Bryan, J., Nagle, B. W., & Doenges, K. H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 723*, 3570-3574.

Burns, R. G. (1990) *Cell Motil. Cytoskeleton 17*, 167-173.

Burns, R. G. (1991) *Biochem. J. 277*, 231-238.

Burns, R. G., & Islam, K. (1981) *Eur. J. Biochem. 117*, 515-519.

Burns, R. G., & Islam, K. (1984) *Eur. J. Biochem. 141*, 599-608.

Caron, J. M., & Berlin, R. D. (1979) *J. Cell Biol. 81*, 665-671.

Daleo, G. R., Piras, M. M., & Piras, R. (1974) *Biochem. Biophys. Res. Commun. 61*, 1043-1050.

Daleo, G. R., Piras, M. M., & Piras, R. A. (1977) *Arch. Biochem. Biophys. 180*, 288-297.

Feit, H., & Barondes, S. (1970) *J. Neurochem. 17*, 1355-1364.

Feit, H., Dutton, G. R., Barondes, S. H., & Shelanski, M. L. (1971) *J. Cell Biol. 51*, 138-147.

Goldschmidt-Clermont, P. J., Kim, J. W., Machesky, L. M., Rhee, S. G., & Pollard, T. D. (1991) *Science 251*, 1231-1233.

Hargreaves, A. J., & Avila, J. (1985) *J. Neurochem. 45*, 490-496.

Hargreaves, A. J., & Avila, J. (1986) *Biosci. Rep. 6*, 913-919.

Hargreaves, A. J., & McLean, W. G. (1988) *Int. J. Biochem. 20*, 1133-1138.

Hartree, E. F. (1972) *Anal. Biochem. 48*, 422-427.

Kelly, P. T., & Cotman, C. W. (1978) *J. Cell Biol. 79*, 173-183.

Klausner, R. D., Kumar, N., Weinstein, J. N., Blumenthal, R., & Flavin, M. (1981) *J. Biol. Chem. 256*, 5879-5885.

Kornguth, S. E., & Sunderland, E. (1975) *Biochem. Biophys. Acta 393*, 100-114.

Lagnado, J. R., & Kirazov, E. P. (1975) in *Microtubules and Microtubule Inhibitors* (Borgers, M., & De Brabander, M., Eds.) pp 127-140, Elsevier, London.

Matus, A., Bernhardt, R., Bodmer, R., & Alaimo, D. (1986) *Neuroscience 17*, 371-389.

Murthy, A. S. N., Bramblett, G. T., & Flavin, M. (1985) *J. Biol. Chem. 260*, 4364-4370.

Reaven, E., & Azhar, S. (1981) *J. Cell Biol. 89*, 300-308.

Smith, R., & Tanford, C. (1972) *J. Mol. Biol. 67*, 75-83.

Walters, B. B., & Matus, A. (1975) *Biochem. Soc. Trans. 3*, 109-112.

Yamauchi, P. S., & Purich, D. L. (1987) *J. Biol. Chem. 262*, 3369-3375.

Bacteriorhodopsin Can Be Refolded from Two Independently Stable Transmembrane Helices and the Complementary Five-Helix Fragment†

Theodore W. Kahn and Donald M. Engelman*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Received January 10, 1992; Revised Manuscript Received April 7, 1992

ABSTRACT: This paper describes experimental tests of the hypothesis that bacteriorhodopsin (BR) can fold by the association of independently stable transmembrane helices. Peptides containing the first and second helical segments of BR were chemically synthesized. These two peptides and the complementary five-helix fragment of BR were reconstituted in three separate populations of native-lipid vesicles which were then mixed and fused to allow the fragments to interact. After addition of retinal, absorption spectroscopy of the reconstituted BR and X-ray diffraction of two-dimensional crystals of this material showed that the native structure of BR was regenerated. The first two helices of BR can therefore be considered as independent folding domains, and covalent connections in the loops connecting the helices to each other and to the rest of the molecule are not essential for the appropriate association of the helices.

Predicting the folded structure of a protein on the basis of its amino acid sequence is a major goal of modern biology. Achieving this goal with membrane proteins may prove to be easier than with water-soluble proteins because of the constraints on secondary structure imposed by the hydrophobic environment of the membrane. Although β structure is present in some instances (Weiss et al., 1990; 1991), in the majority of integral membrane proteins the membrane-spanning regions appear to be α-helical (Popot & de Vitry, 1990). These segments are localized in the membrane because of the hydrophobicity of their amino acid side chains, and they are helical because of the strong energetic favorability of backbone hydrogen bonding in the nonpolar environment (Engelman et al., 1986).

Bacteriorhodopsin (BR)[1] is an integral membrane protein that is composed of seven transmembrane helices connected by short extramembranous loops (Henderson et al., 1990). It has been hypothesized that the secondary structure of the

† This research was supported by grants from the NIH (GM22778 and GM39546), the NSF (DMB8805587), and the National Foundation for Cancer Research.

* To whom correspondence should be addressed.

[1] Abbreviations: BR, bacteriorhodopsin; BSA, bovine serum albumin; KB, 30 mM potassium phosphate, 150 mM KCl, 0.025% NaN_3, pH 6; KB/5, 6 mM potassium phosphate, 30 mM KCl, 0.005% NaN_3, pH 6; SDS, sodium dodecyl sulfate; SDS buffer, 50 mM sodium phosphate, 5% SDS, 0.025% NaN_3, pH 6, 7, or 8; UV-CD, ultraviolet circular dichroism.

0006-2960/92/0431-6144$03.00/0

helices does not depend on the tertiary structure of BR and that the structure of the protein therefore can be thought of as arising from the side-to-side association of independently stable helices (Popot & Engelman, 1990). To test this hypothesis, BR has previously been regenerated from two fragments containing two and five helices each (Popot et al., 1987). These fragments were incorporated independently into separate populations of lipid vesicles, in which their secondary structures were determined to be largely helical. They were then allowed to interact by mixing the two populations and fusing the vesicles, under which conditions the fragments associated to reform the native structure of BR. To date, however, this kind of experiment has not been performed under conditions that stringently test the idea that each helix acts as an independently stable folding domain, separately adopting its secondary structure, and then associating with the remaining helices to form the tertiary structure of the protein.

In this paper, we show that BR can be reformed from peptides containing the first helix, the second helix, and the remaining five helices, which have been independently refolded in lipid bilayers, and then allowed to interact with each other. Each of the fragments was reconstituted separately into vesicles. Other workers have shown that the five-helix fragment (Popot et al., 1987) and each of the single-helix fragments (Hunt et al., 1991) are largely helical when independently reconstituted in vesicles. The vesicles were then mixed and fused to allow the fragments to interact with each other in the same bilayer, and retinal was added, resulting in the regeneration of the native BR absorption spectrum. Further, the reconstituted protein could be induced to form the crystal lattice characteristic of native BR in purple membrane, and X-ray diffraction demonstrates that the structure of the reconstituted molecule is similar to that of the native molecule. These results are discussed in the context of previous studies of BR folding, and they contribute the most direct evidence yet obtained that individual helices in a polytopic membrane protein can act as separate folding domains.

Nomenclature

A new nomenclature will be used which is designed to show in a clear way which helices are present in a sample and whether any covalent connections between the helices have been cut. The helices are lettered according to Engelman et al. (1980). A dot (·) between two letters indicates that no covalent bond connects those two helices, either because the sample has been cleaved chemically or proteolytically or because the sample has been reconstituted from synthetic peptides corresponding to individual helices. If the sample contains retinal, the name of the sample will be prefixed with the letter R. Subscripts indicate the state of the sample: S indicates a sample in the form of membrane sheets as obtained from *Halobacterium halobium*; V indicates a sample reconstituted into lipid vesicles; O indicates a sample dissolved in organic solvents; D indicates a sample solubilized in detergent; and M indicates a sample in mixed micelles (detergent plus lipid). If the sample is being referred to in an abstract sense, no subscripted letter will be used. Some examples:

$R(ABCDEFG)_S$	purple membrane
$(ABCDEFG)_V$	bleached BR that has been reconstituted into vesicles
$(AB)_D$	the fragment containing the first two helices, solubilized in detergent
$R(A{\cdot}B{\cdot}CDEFG)_V$	BR reconstituted in vesicles from fragments containing the first helix, the second helix, and the last five helices, with retinal added

The phrase "reconstituted" will be used to refer to samples in which BR or fragments of BR have been incorporated into vesicles. The phrase "regenerated" will be used to refer to samples which have regained the native chromophore of BR after addition of retinal to bleached material.

Materials and Methods

Materials. All aqueous solutions were made using water purified by a Milli-Q water system (Millipore). Baker PCS reagent-grade 90% formic acid was used. Ethanol (200 proof) was from Quantum Chemical Corporation. *all-trans*-Retinal, ethanolamine, and TLCK-treated chymotrypsin were from Sigma. Retinal was dissolved in ethanol just before use, and the concentration was determined from the absorbance at 380 nm using an extinction coefficient of 42 800 cm^{-1} M^{-1}. Sodium taurocholate was from Calbiochem. Electrophoresis-purity SDS was from Bio-Rad Laboratories.

Buffers. SDS buffer: 50 mM sodium phosphate; 5% SDS; 0.025% sodium azide; adjusted with HCl or NaOH to desired pH (either 8, 7, or 6). KB: 30 mM potassium phosphate; 0.15 M KCl; 0.025% sodium azide; pH 6.0. KB/5: same as KB, but diluted 5-fold.

Scintillation Counting. Typically, aliquots of 5, 10, 15, and 20 μL were counted, and linear regression was used to determine the number of counts per minute per microliter. Seven-milliliter glass vials were used containing 6.5 mL of Packard Opti-fluor liquid scintillation counter cocktail. After the addition of aliquots to the vials, the samples were stored in the dark for 8 h to allow chemiluminescence to decay. Then, the samples were counted with a Packard 2000CA Tri-carb liquid scintillation analyzer.

Amino Acid Analysis. Quantitative amino acid analysis was performed by M. Crawford of the Protein and Nucleic Acid Chemistry Facility of the Yale University School of Medicine as described in Popot et al. (1987). A known quantity of norleucine was added to each sample and was used to calibrate the quantitation of each amino acid. Usually, three or four aliquots of the same sample were analyzed and linear regression was used to determine the protein concentration.

Preparation of Purple Membranes. H. halobium was grown and purple membranes were purified as described in Popot et al. (1987). For ease of protein concentration determination, radioactively labeled purple membranes were also prepared by adding [^{3}H]leucine to defined growth medium as described in Popot et al. (1987). The labeled membranes were mixed with unlabeled membranes to give a specific activity of about 10–20 cpm/nmol of leucine. The specific activity was determined by scintillation counting and quantitative amino acid analysis. The protein concentration of samples made from the labeled material could then be easily measured by scintillation counting.

Preparation of (CDEFG). (CDEFG) was prepared essentially by using the methods described in Popot et al. (1987), which in turn are based on the methods given in Liao et al. (1984). This procedure also produces (AB), which was not used for the experiments described in this paper. The formic acid was neutralized with ethanolamine instead of ammonium hydroxide. Typical protein concentrations in SDS buffer were 0.1–0.5 mg/mL.

Preparation of Peptides (A) and (B). (A) and (B) were synthesized using the solid-phase method and were purified by reversed-phase HPLC by J. Elliott of the Protein and Nucleic Acid Chemistry Facility of the Yale University School of Medicine. After purification, analytical HPLC showed one main peak with only very small contaminating peaks, amino acid analysis showed the expected numbers of each type of residue, and mass spectrometry showed the expected molecular

weights. The N-terminus of (A) was acetylated; the N-terminus of (B) was not. The C-termini of both peptides were amidated.

The lyophilized peptides were dissolved in organic solvents and then transferred to SDS buffer solution as follows: About 0.5 mL of formic acid/mg of peptide was added to the dry peptide, and the solution was vortexed for 1–2 min. A total of 3.5 volumes of ethanol were added followed by 20 mg of SDS/mg of peptide. The solution was vortexed for about 5 min until the SDS had all dissolved. The solution was cooled on ice, and the pH was brought to about 7 (as measured by pH paper) by adding dropwise with mixing a volume of ethanolamine approximately equal to the volume of formic acid. The solution was transferred to Spectra/Por-6 (Spectrum Medical Industries) dialysis tubing and was dialyzed at room temperature against SDS buffer at pH 8. After a few hours, the dialysate was replaced with fresh buffer at pH 8. After 12 h, the buffer was again replaced. Twelve hours later, the buffer was replaced with fresh buffer at pH 7, and 12 h later with fresh buffer at pH 6. After another 12 h, the peptide solution was removed from the bag. The concentration was determined by the Pierce BCA method, using γ-globulin and BSA as standards.

Preparation of H. halobium Lipids. Lipids were prepared as described in Popot et al. (1987). Thin-layer chromatography as described in Popot et al. (1987) showed no variation in lipid content between preparations.

Co-reconstitutions. Reconstitutions with all three fragments mixed together were performed essentially as described in Popot et al. (1987). Retinal was added at a 1.5:1 retinal to (CDEFG) molar ratio. In some cases, taurocholate was added at a 1:1 taurocholate to protein mass ratio in order to increase the recovery of protein in the reconstitution process.

Reconstitution by Vesicle Fusion. (A), (B), and (CDEFG) were independently reconstituted in native lipid vesicles at a 2:1 lipid to protein mass ratio with a 1:1 taurocholate to protein mass ratio using the procedure described above. The three populations of vesicles were then mixed so as to give a 2:2:1 molar ratio of (A) to (B) to (CDEFG). The vesicles were fused using the freeze/thaw method: The sample was frozen by submersion in liquid nitrogen. It was allowed to thaw in ice water, and it was then sonicated using a Heat Systems-Ultrasonics, Inc. sonifier equipped with a microtip probe at a power level between 4 and 7. The tube was submerged in room temperature water, and six bursts of 10 s each separated by 1.5 min of cooling were used. The entire freeze/thaw/sonication procedure was performed three times. Retinal was added either before or after fusion.

Measurement of Chromophore Regeneration. The extent of chromophore regeneration was determined by absorption spectroscopy using 1-cm-path length semimicro self-masking cuvettes (Hellma) in a Perkin Elmer Lambda 6 UV/vis spectrophotometer. The reference cuvette contained buffer. Samples were sonicated with a microtip probe sonicator, using six 10-s bursts separated by 1.5 min of cooling. The sample tube was submerged in room temperature water throughout. The sonication reduced light scattering by reducing the size of the vesicles. Nevertheless, some light scattering was always present in the spectra, so a straight line drawn tangential to the tails of the peak was used as a baseline. The peak height was taken to be the distance from the line to the point in the peak farthest from the line. This method actually underestimates the peak height because the flat baseline is only an approximation; however, this method has been used consistently for all samples, and was used in previous work to determine extinction coefficients for similar samples. The extinction coefficient was assumed to be 45 000 M^{-1} cm^{-1}, the same as determined for R(AB-CDEFG)$_v$ (Popot et al., 1987). The concentration of regenerated BR was then divided by the concentration of (CDEFG) as determined by scintillation counting to give the extent of regeneration.

Preparation of Samples for X-ray Diffraction. Between 0.5 and 1 mg of protein, either in sheets or reconstituted into vesicles, was spun down in an SW60 rotor at 60K rpm (370000g at r_{avg}) for 70 min, usually from KB. The pellet was resuspended in KB/5 and was spun down again, the supernatant was carefully removed with a drawn-out Pasteur pipet, and the inside of the tube was dried with a tissue. One to two microliters of KB/5 was added back to the pellet, which was homogenized with a stir bar (1 × 2 mm) for about 5 min. The consistency of the sample was as thick as possible, but without any lumps. The tube was then spun briefly in a table-top centrifuge to coalesce the sample. Approximately 2 μL of the sample was deposited on a piece of clear, biaxially oriented polystyrene 0.0012 in thick (Kama Corporation), using care not to allow any bubbles to form. If a larger sample was desired, another 2 μL was added to the spot. The sample was allowed to equilibrate overnight in a sealed box containing a saturated solution of $CaCl_2$, which produces a relative humidity of about 30–35% (Spencer, 1926). Sometimes a saturated solution of K_2SO_4 was used instead to produce a relative humidity of about 97%. The result was a dry disk of membranes about 1.5–2 mm in diameter and about 0.1–0.2 mm thick. The sample was stored in the sealed box when not being used for measurements.

Collection of X-ray Patterns. The X-ray source was an Elliott GX6 rotating anode X-ray generator producing copper K_α X-rays, operating at 45 kV and 50 mA, employing a 0.2 × 2 mm focal spot. The take-off angle was approximately 6.5°, giving a 0.2 × 0.2 mm beam source. The diffraction patterns were collected using a Searle toroidal X-ray diffraction camera made by Baird & Tatlock, in which a gold-plated ellipsoidal mirror focuses the X-ray beam. The sample-to-film distance was 7.3 cm.

The camera was sealed and flushed with helium bubbled through a saturated solution of $CaCl_2$ for 20 min prior to beginning the exposure and throughout the exposure. This serves the dual purpose of maintaining the sample at 30–35% relative humidity and minimizing air scattering of X-rays. Exposure times were typically 30–45 min. At the conclusion of the exposure, the position of the X-ray beam was recorded on the film by attenuating the beam with aluminum and then briefly exposing the film with the beam stop displaced from the path of the beam.

Patterns were recorded on Kodak DEF 5 direct exposure diagnostic film. The film was developed in Kodak GBX developer following the manufacturer's instructions. After drying, films were digitized using an Optronics International Inc. Photoscan System P-1000 film scanner, at a 50 × 50 μm pixel size. The films were then radially averaged using a FORTRAN program running on a VAX 8800 computer (Digital Equipment Corporation). The center of the pattern was located using the position of the beam spot recorded on the film.

RESULTS

If BR is composed of independently stable helices that associate to form the tertiary structure, and if therefore the loops between the helices are not required for folding, then it might be possible to reconstitute BR from individual helices. To test this hypothesis, two peptides were chemically synthesized, one containing the helix A sequence, and the other the helix B

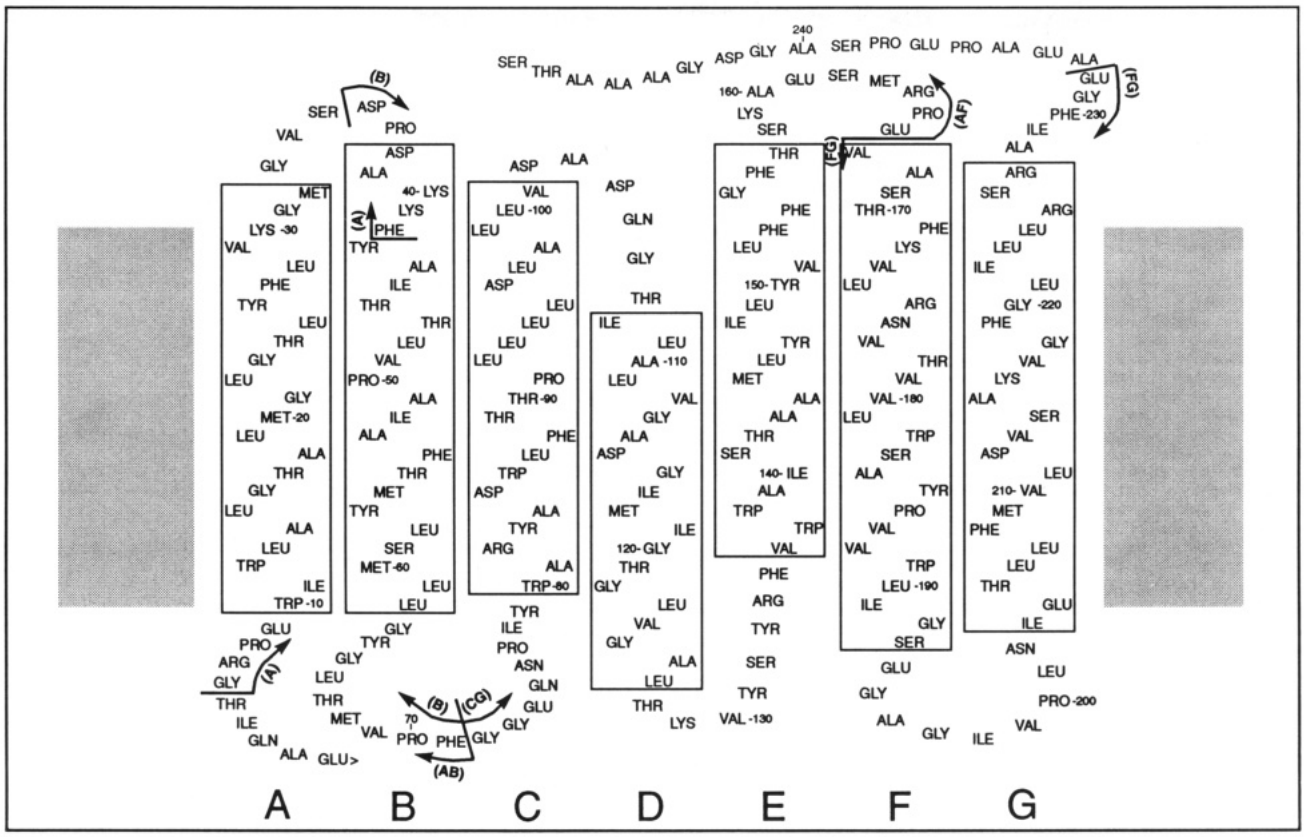

FIGURE 1: The sequence of BR, showing the secondary structure determined by Henderson et al. (1990). The boxed-in regions are helical. The shaded areas show the approximate location of the hydrophobic region of the membrane. The beginning and ending points of (A) and (B) are indicated. Other fragments described in the text are also shown.

sequence. These helices are the two helices of the (AB) fragment. Thus, a reconstitution of BR from (A) + (B) + (CDEFG) + R is conceptually analogous to a reconstitution from (AB) + (CDEFG) + R. However, the reconstitution using (A) and (B) is a direct test of whether these two individual peptides are separately capable of folding and associating with (CDEFG) to form the complete BR structure.

As synthesized, (A) extends from residue 6 to 42, and (B) extends from 36 to 71 (see Figure 1, which shows the entire sequence of BR). (A) and (B) overlap by 7 residues, the overlap being entirely within the polar region of the sequence between the first and second helices. (A) ends within the second helix in the Henderson et al. (1990) structure, but before the beginning of the hydrophobic and therefore membrane-spanning region of that helix. (B) ends at the chymotryptic cleavage site used to make (AB) and (CDEGF). The overlap between (A) and (B) was included so as to give each of these extremely hydrophobic peptides a significant number of polar residues so as to promote solubility in mixed organic/aqueous solvent systems and in aqueous detergent solutions.

Co-reconstitution. Reconstitutions were performed using the SDS precipitation method of Popot et al. (1987). Unlike R(AB·CDEFG)$_v$, which begins to turn purple in less than 30 min after precipitation of the SDS, R(A·B·CDEFG)$_v$ took about 12 h before a purple color became apparent. Regeneration of the chromophore was complete after about 2 days. When (CDEFG) was reconstituted with only (A) or only (B), and retinal was added, no purple chromophore recovery was observed, showing that both helices are required to reform the native structure of BR.

The concentration of protein in a reconstituted sample was determined by scintillation counting. Since only (CDEFG) was labeled with tritium, it was assumed that the recovery of (A) and (B) was proportional to the recovery of (CDEFG).

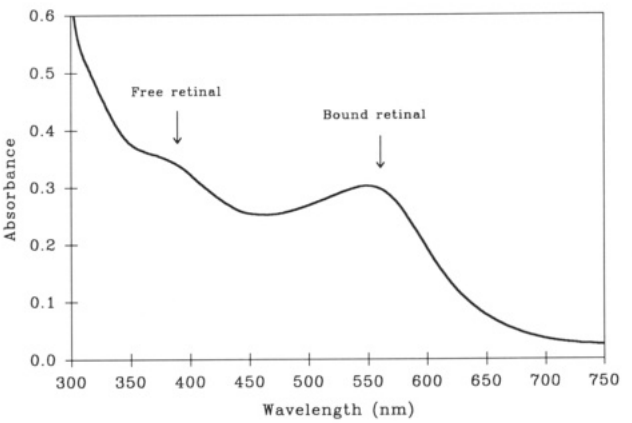

FIGURE 2: The absorption spectrum of R(A·B·CDEFG)$_v$. The peaks of bound and free retinal are indicated.

This is probably a reasonable assumption because Popot et al. (1987) found that in reconstitutions of (AB) plus (CDEFG), the two fragments were recovered equally. When no taurocholate was used in the reconstitution, the recovery of protein varied from about 50% to about 90%, with samples of larger volumes giving higher recoveries. With taurocholate, the recovery was typically 90–100%, regardless of sample volume.

The spectrum of free retinal has an absorption maximum at about 380 nm. When retinal is attached to a lysine by a Schiff base, the absorbance peak is shifted to about 450 nm. When retinal binds to native BR, the peak is further shifted to about 560 nm. This spectral shift is a sensitive assay for the native structure of BR. Figure 2 shows the spectrum of R(A·B·CDEFG)$_v$. Because retinal is added in excess relative to (CDEFG), some free retinal is present in the sample. The peaks for both free and bound retinal can be seen in the spectrum. The height of the bound retinal peak above the

baseline was used as the absorbance. The maximum was in the range of 560–570 nm. Typical regeneration levels were in the range of 20–50% relative to the amount of (CDEFG) present in the sample.

For most reconstitutions, $(A)_D$ and $(B)_D$ were added at a 2-fold molar excess compared to $(CDEFG)_D$. Lower ratios gave less regeneration, while higher ratios did not increase the regeneration when fresh $(A)_D$ and $(B)_D$ were used. When $(A)_D$ and $(B)_D$ were stored in SDS buffer for long periods of time, the extent of regeneration decreased. The same has also been found to be true for $(AB)_D$ and $(CDEFG)_D$ (Liao et al., 1983; Popot et al., 1987).

The regeneration of $R(A{\cdot}B{\cdot}CDEFG)_v$ was dependent on the lipid to protein ratio. The optimum mass ratio was 1:1 or 2:1. Reducing the ratio to 0.75:1 or 0.5:1 caused the level of regeneration to drop by about two-thirds. When the reconstitution was done at a 10:1 lipid to protein ratio, the absorbance maximum was at 455 nm instead of the usual 560–570 nm. Apparently the fragments were associating and binding retinal, but the environment of the retinal was not native. The retinal had approximately the absorbance maximum expected if it were connected to the protein by a protonated Schiff base but were not otherwise interacting with charged groups in the protein (Honig & Ebrey, 1982). Adding lipid to a sample which had already regenerated caused the retinal to dissociate from the protein. A spectrum was taken of a regenerated sample at a 0.5:1 lipid to protein ratio. Then sonicated pure lipid vesicles were added to the sample so as to raise the lipid to protein ratio to 10:1. Before the vesicles were fused the spectrum was found to be unchanged, but after the vesicles were fused by the freeze/thaw method, the peak at 560 nm had dropped by about 50% and a free retinal peak at 380 nm had appeared. When the lipid to protein ratio was raised to 100:1 by adding more lipid vesicles and fusing them, the BR chromophore was completely lost. In a sample that consists entirely of 100% regenerated BR, a 1:1 mass ratio of lipid to protein would correspond to approximately a 30:1 molar ratio.

Reconstitution by Vesicle Fusion. (A), (B), and (CDEFG) were independently reconstituted in native lipid vesicles at a 2:1 lipid to protein mass ratio with a 1:1 taurocholate to protein mass ratio. This lipid to protein ratio was used because vesicle fusion becomes more efficient as the lipid to protein ratio is raised. However, the ratio could not be much higher since at a 10:1 ratio BR reconstituted from the three fragments mixed together did not regenerate the native spectrum (see previous section). The three populations of vesicles were then fused so as to yield a 2:2:1 molar ratio of (A) to (B) to (CDEFG). Retinal was added either before or after fusion. The extent of chromophore regeneration was comparable to that obtained when (A), (B), and (CDEFG) were reconstituted together.

Use of X-ray Diffraction to Compare Reconstituted Bacteriorhodopsin to Native Bacteriorhodopsin. One way to test whether reconstituted BR has a structure similar to that of native BR is by X-ray diffraction. Purple membrane occurs naturally as a two-dimensional crystal of BR with a small amount of lipid (the lipid to protein mass ratio is 1:3). $R(ABCDEFG)_v$ and $R(AB{\cdot}CDEFG)_v$ can also be induced to crystallize by drying down the vesicles on a flat substrate (Popot et al., 1987). X-ray diffraction of the membranes produces a characteristic diffraction pattern, with strong, sharp reflections to about 6.5 Å and weaker reflections to at least 3.5 Å (Henderson, 1975). To obtain such a pattern, a sample is pelleted, and the pellet is then homogenized and transferred to a flat substrate transparent to X-rays on which it is allowed to dry slowly at a controlled relative humidity. The result is a stack of membranes that are well ordered parallel to the substrate but that are rotationally randomized about the normal to the membrane plane. Each membrane layer in the stack contains a two-dimensional crystal of BR. When the X-ray beam is passed through the sample perpendicular to the plane of the membranes, each crystal produces a hexagonal array of reflections. These reflections are averaged into concentric circles because of the rotational disorder of the membranes. A Fourier transform calculated using the appropriately processed integrated intensities of these X-ray reflections plus phases obtained from electron microscopy would give a projection of the electron density of BR into the plane of the membrane [see, for example, Engelman and Zaccai (1980) and Koch et al. (1991)].

$R(A{\cdot}B{\cdot}CDEFG)_v$ was reconstituted with a 2:2:1 molar ratio of (A) to (B) to (CDEFG) and a 1:1 mass ratio of lipid to protein. The material was dried down on a flat substrate at a controlled relative humidity of 35% (saturated calcium chloride). Figure 3 shows a diffraction pattern of crystallized $R(A{\cdot}B{\cdot}CDEFG)_v$, with a pattern of $R(AB{\cdot}CDEFG)_v$ shown for comparison. Figure 4 shows radial averages of diffraction patterns of intact BR and of BR reconstituted from fragments. Comparison of the pattern of $R(A{\cdot}B{\cdot}CDEFG)_v$ with that of $R(AB{\cdot}CDEFG)_v$ or $R(ABCDEFG)_v$ shows that all of the reflections are present at their correct locations and that their relative intensities are approximately correct. The unit cell dimension is 62.2 ± 0.6 Å. The crystallization was repeated several times, and it gave substantially the same results each time. $R(A{\cdot}B{\cdot}CDEFG)_v$ dried at 97% relative humidity gave a weaker pattern (not shown) than material dried at 35% relative humidity. Also, the strength of the pattern obtained from a sample dried at 97% relative humidity could be improved by transferring the sample to 35% relative humidity. The relatively low signal-to-noise ratio is presumably due to the presence of denatured material that does not crystallize. In an attempt to drive the protein molecules closer together in the hope of increasing the amount of crystalline material, samples were made at lower lipid to protein mass ratios of 0.75:1 and 0.5:1. These samples did not regenerate as well, and the intensity of the lattice was lower. Because $(AB{\cdot}CDEFG)_v$ does not crystallize in the absence of retinal (Popot et al., 1987), and because (CDEFG) reconstituted with only (A) or only (B) did not bind retinal (see above), crystallization of $(A{\cdot}CDEFG)_v$ or $(B{\cdot}CDEFG)_v$ was not attempted.

Despite the relatively low level of crystallization, it is clear that $R(A{\cdot}B{\cdot}CDEFG)_v$ is capable of crystallizing in the same way as $R(ABCDEFG)_v$. It is also clear from the resulting diffraction pattern that the structure of $R(A{\cdot}B{\cdot}CDEFG)_v$ is very similar to that of $R(ABCDEFG)_v$. All of the reflections, extending out to the (7,1) reflection, are in the same locations with an error of less than 1% in S, showing that the material crystallizes in the same lattice. The relative sizes of the peaks are roughly the same by inspection, showing that the structure is similar to a resolution of about 7 Å.

Discussion

Previous studies on the folding of BR have shown that BR can be regenerated from fragments containing two or more helices each. By reconstituting BR from (AB), (CDEFG), and retinal in mixed micelles containing lipids and detergents, Huang et al. (1981) showed that two pieces of BR can associate and refold properly after being denatured, separated, and mixed together again. Liao et al. (1984) and Sigrist et al. (1988) have obtained similar results with (ABCDE) and (FG). Liao et al. (1983) found that $(AB)_M$ and $(CDEFG)_M$ each contain as great a percentage of helix as $(ABCDEFG)_s$,

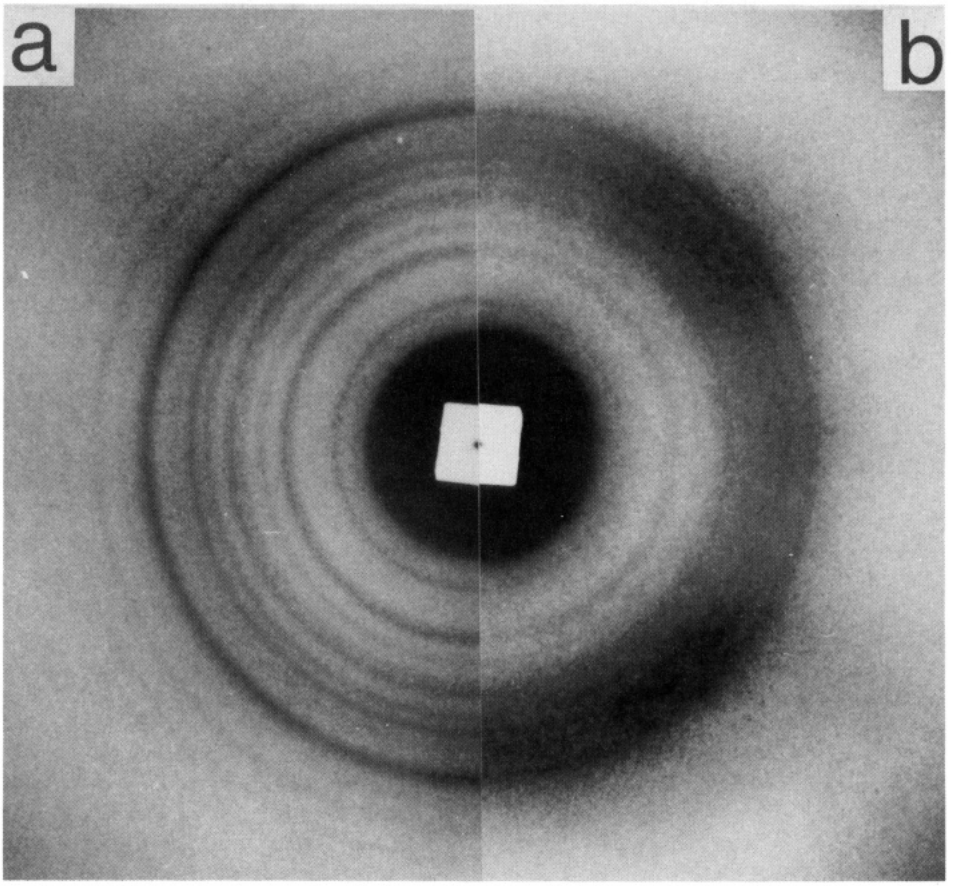

FIGURE 3: X-ray diffraction patterns of (a) R(AB·CDEFG)$_v$ and (b) R(A·B·CDEFG)$_v$.

showing that the fragments need not associate in order to adopt their native secondary structure.

In order to determine whether all seven helices of BR are required for regeneration, Liao et al. (1984) cleaved (AB) into single helices using formic acid. A 1:1:1 molar ratio of (A) to (B) to (CDEFG) in mixed micelles only regenerated the chromophore to a level of 5–10%. Liao et al. concluded that cleaving (AB) has a profound effect on the ability of BR to regenerate, but still considered it theoretically possible that (BCDEFG) might be capable of binding retinal to reform the native chromophore.

Popot et al. (1987) extended the studies of (AB) and (CDEFG) and regeneration of BR from those fragments by reconstituting them independently into vesicles. Each fragment showed a UV-CD spectrum quite similar to that of R-(ABCDEFG)$_v$, indicating that the fragments fold individually into highly helical structures in lipid bilayers. Neither fragment alone bound retinal. When the vesicles containing the two fragments were mixed and fused so that the fragments were free to interact and retinal was added, the native chromophore of BR was regenerated. The UV-CD spectrum of the associated fragments with or without retinal was virtually indistinguishable from that of R(ABCDEFG)$_v$. When these vesicles were dried down on a smooth substrate, and the dried sample was placed in an X-ray camera, it was found that R(AB·CDEFG)$_v$ formed a two-dimensional crystal lattice identical to that of R(ABCDEFG)$_s$ [see also Popot et al. (1986)].

Gilles-Gonzalez et al. (1991) have shown by preparing deletion mutants of BR that significant portions of the loops connecting the second to the third helices and the fifth to the sixth helices can be eliminated without abolishing the protein's ability to fold properly.

The idea that the helices of BR are independently stable has been tested in the present study by examining the structure of BR reconstituted from two individual helices plus a five-helix fragment, reconstituted together or each reconstituted separately and then allowed to interact. The fact that hydrophobicity analyses such as that described by Engelman et al. (1986) are able to locate membrane-spanning α-helices within the sequences of proteins without considering interhelix interactions or the structure of the extramembranous parts of the proteins leads to the hypothesis that the α-helices may be independently stable (Popot & Engelman, 1990). Interhelix interactions and extramembranous structure may be important in determining the exact arrangements of helices in a protein and may contribute to the stability of the protein but may not be needed to specify the secondary structure within the membrane. According to this hypothesis, in the case of globular membrane proteins (those that span the membrane several times and have a significant fraction of their mass within the membrane), the energetics leading to the stable structure of a protein can be thought of as consisting of two components: (a) the formation of membrane-spanning α-helices, dictated by the low dielectric constant of the lipid environment, each of which would be stable on its own; and (b) the energies that drive the helices together in their appropriate tertiary arrangement, including interhelix interactions, the lipophobic and packing effects, and the constraints imposed by the structure of the extramembranous portions of the protein. The association of the helices to form the tertiary structure probably does not significantly alter the secondary structure, since the prediction routines correctly find the secondary structure without considering tertiary effects. This two-stage model for the formation of membrane protein structure is reviewed by Popot and Engelman (1990).

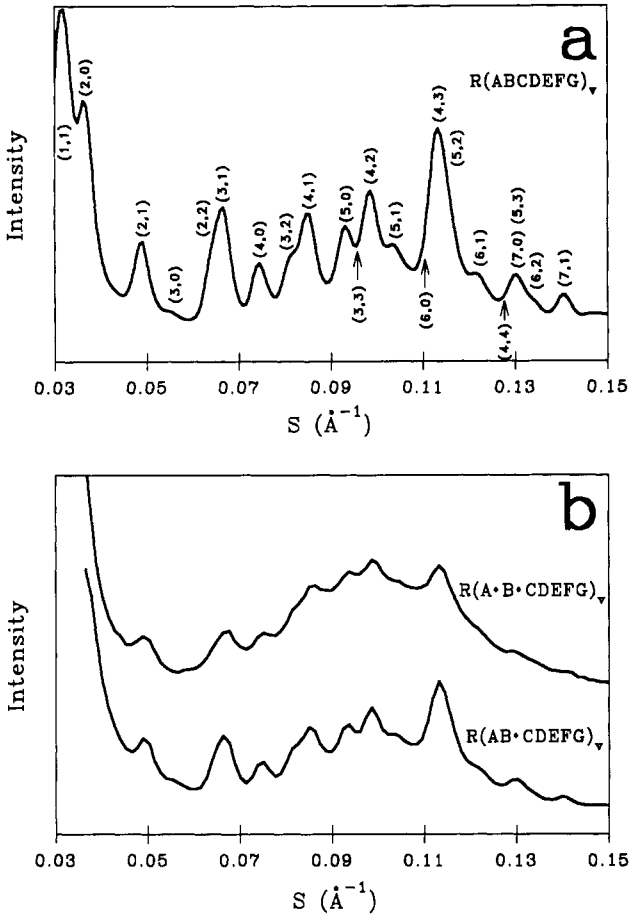

FIGURE 4: Radial averages of X-ray diffraction patterns from dried-down samples. (a) Purple membrane, with the indices of the reflections indicated. The positions of reflections (3,3), (6,0), and (4,4), which have intensities of zero, have also been indicated. (b) R(A·B·CDEFG)$_v$, with R(AB·CDEFG)$_v$ shown for comparison. The radial average of a R(A·B·CDEFG)$_v$ sample that did not crystallize has been used as a background and subtracted from the R(A·B·CDEFG)$_v$ pattern in order to increase the signal-to-noise ratio in the figure. The vertical positions of the patterns have been arbitrarily shifted for clearer presentation.

The results presented in this paper provide direct experimental evidence to support the idea that BR is composed of independently stable helices that can associate in a lipid bilayer to form the structure of the molecule. Previous work has shown that a two-helix fragment and a five-helix fragment can associate properly, but this is the first time BR has been reassembled from fragments corresponding to individual helices that have been refolded in bilayers and then allowed to interact. The first two individual helices are able to fold and associate with the (CDEFG) fragment, even though no covalent bonds connect the three pieces together. This is true even when the fragments are reconstituted into three separate populations of vesicles and are then brought together by vesicle fusion. The three fragments independently adopt structures in the membranes which allow them to associate properly to reform BR. Popot et al. (1987) have shown by UV-CD that (CDEFG) adopts its native secondary structure when reconstituted in vesicles, and Hunt et al. (1991) have shown by UV-CD and Fourier transform infrared spectroscopy studies that (A) and (B) each forms helices when reconstituted independently into lipid vesicles and that the helices are perpendicular to the membrane. The covalent connections in the loops connecting the individual helices are not needed to specify the correct structure of the helices, nor are they necessary for the proper association of the helices. On the basis of absorption spectroscopy and X-ray diffraction, R(A·B·CDEFG)$_v$ is very similar in structure to R(ABCDEFG)$_s$. Although the ability of R(A·B·CDEFG)$_v$ to pump protons was not assayed in this study, Khorana et al. (1988) and Marti et al. (1991) have found that all examples of genetically mutated versions of BR that can regenerate the native chromophore can also pump protons. Kataoka et al. (1992) have shown that R(A·B·CDEFG)$_v$ undergoes a photocycle, indicating that it is probably functional.

Although the covalent connections in the two loops studied are not essential for the folding of BR, they do contribute to the protein's stability. At a lipid to protein ratio of 10:1, retinal did not interact properly with the protein. At a 100:1 ratio, retinal did not bind at all. Apparently the protein must be at a fairly high concentration in order for the three fragments to remain associated. When the protein is diluted by raising the lipid to protein ratio, the fragments dissolve in the lipid and the chromophore is lost. The fact that a three-fragment reconstitution takes significantly longer than a two-fragment reconstitution to regenerate the native chromophore presumably is due to the need for three molecules to associate, perhaps in a particular order.

In a few other cases, the extramembranous portions of membrane proteins have been shown not to be essential for the association of transmembrane helices. When rhodopsin is cleaved into three fragments using a water-soluble protease, the fragments remain associated and retain their native secondary structure (Litman, 1979); this may however be due to interactions between the extramembranous portions of the protein. The dimerization of glycophorin A depends only on the transmembrane region of the protein (Furthmayr & Marchesi, 1976; Bormann et al., 1989). This dimerization depends on the sequence of the transmembrane region, since conservative amino acid substitutions on one face of the helix disrupt dimerization, while similar substitutions on the opposite face have no effect (Lemmon et al., 1992). Kurosaki et al. (1991) showed that a substitution of isoleucine for leucine at one position within the membrane-spanning portion of one subunit of an IgG Fc receptor interferes with the ability of the subunits to associate to form the complete receptor. Some synthetic peptides that are just long enough to span a membrane in the form of α-helices are able to associate within a membrane and form ion channels (Lear et al., 1988; Oiki et al., 1988).

If the extramembranous loops are not essential for the association of helices in a membrane protein, it is interesting to speculate about what other factors might be involved. Helices may tend to be driven together by what Jähnig (1983) calls the lipophobic effect. A helix inserted into a membrane above the phase-transition temperature of the lipid will order the fatty acyl chains near it, and therefore helices may associate in order to minimize the surface area of protein exposed to lipid, much as the hydrophobic effect causes the nonpolar residues of a water-soluble protein to associate in the center of the protein in order to minimize the surface area of nonpolar residues exposed to water. Engelman and Steitz (1984) and Popot and Engelman (1990) point out that the rough surface of a helix contains many cavities that can be filled well when helices pack together with each other, but not as well when helices pack with lipids, which are comparatively straight. The free energy cost of leaving a cavity unfilled is high because of the unsatisfied van der Waals interactions and the presence of vacuum in the cavity. Rees et al. (1989), Komiya et al. (1988), and Yeates et al. (1987) compared the membrane-spanning sequences of photosynthetic reaction centers from different

species and found that the residues that point toward the interfaces between the helices tend to be conserved, while those that point toward the lipids tend not to be conserved. The result is a periodicity in the positions of conserved residues in the sequences, with the conserved residues mostly on one face of each helix. They concluded that the packing in the center of the membrane-spanning portion of the molecule is quite specific, while the packing between helices and lipids is not specific. Water-soluble proteins show a similar pattern, with conserved residues in the center where packing is tight, and more sequence variability at the exterior. Yeates et al. (1987) found that the packing of residues at the interfaces of the helices of the reaction center is as tight as in the centers of water-soluble proteins. Helices may also associate in order to form hydrogen bonds or salt bridges in the interior of a protein, although Yeates et al. found no salt bridges and very few hydrogen bonds between the helices of the reaction center. Although covalent connections in the loops are not required for the association of at least some of the helices of BR, the secondary and tertiary structures of the loops may still play a role in the arrangement of the helices. However, the work of Gilles-Gonzales et al. (1991), in which the loops connecting helix B to helix C and helix E to helix F were shortened, shows that in these two cases, the structures of the loops are not important.

We have shown that peptides containing helix A and helix B of BR can associate with a peptide containing the remaining five helices, C through G, and with retinal, to form the native structure of BR. The fact that the peptides can be reconstituted separately in membrane bilayers and then allowed to interact supports the idea that single helices can be regarded as folding domains. For at least two of the helices of BR, covalent connections in the loops are not essential for the association of the helices.

Acknowledgments

We thank Jean-Luc Popot for helpful discussions, Gerald Johnson for building and maintaining much of the equipment used for the X-ray diffraction experiments, Arthur Perlo for help with computer programming, and Mark Lemmon for comments on the manuscript.

Registry No. Retinal, 116-31-4.

References

Bormann, B.-J., Knowles, W., & Marchesi, V. (1989) *J. Biol. Chem. 264*, 4033–4037.

Engelman, D., & Zaccai, G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5894–5898.

Engelman, D., & Steitz, T. (1984) in *The Protein Folding Problem* (Wetlaufer, D. B., Ed.) pp 87–113, Westview Press, Boulder, CO.

Engelman, D., Henderson, R., McLachlan, A., & Wallace, B. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2023–2027.

Engelman, D., Steitz, T., & Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 321–353.

Furthmayr, H., & Marchesi, V. (1976) *Biochemistry 15*, 1137–1144.

Gilles-Gonzalez, M., Engelman, D., & Khorana, H. (1991) *J. Biol. Chem. 266*, 8545–8550.

Henderson, R. (1975) *J. Mol. Biol. 93*, 123–138.

Henderson, R., Baldwin, J., Ceska, T., Zemlin, F., Beckmann, E., & Downing, K. (1990) *J. Mol. Biol. 213*, 899–929.

Honig, B., & Ebrey, T. (1982) *Methods Enzymol. 88*, 462–470.

Huang, K.-S., Bayley, H., & Khorana, H. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 323–327.

Huang, K.-S., Bayley, H., Liao, M.-J., London, E., & Khorana, H. (1981) *J. Biol. Chem. 256*, 3802–3809.

Hunt, J., Bousche, O., Meyers, K., Rothschild, K., & Engelman, D. (1991) *Biophys. J. 59*, 400a.

Jähnig, F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3691–3695.

Kataoka, M., Kahn, T., Tsujiuchi, Y., Engelman, D., & Tokunaga, F. (1992) *Photochem. Photobiol.* (in press).

Khorana, H., Braiman, M., Chao, B., Doi, T., Flitsch, S., Gilles-Gonzalez, M., Hackett, N., Jones, S., Karnik, S., Lo, K.-M., Marti, T., Mogi, T., Nassal, M., Stern, L., & Subramanian, S. (1988) in *Structure and Expression. Volume 1: From Proteins to Ribosomes* (Sarma, R. H., & Sarma, M. H., Eds.) pp 1–23, Adenine Press, Guilderland, NY.

Koch, M., Dencher, N., Oesterhelt, D., Plöhn, H.-J., Rapp, G., & Büldt, G. (1991) *EMBO J. 10*, 521–526.

Komiya, H., Yeates, T., Rees, D., Allen, J., & Feher, G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9012–9016.

Kurosaki, T., Gander, I., & Ravetch, J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3837–3841.

Lear, J., Wasserman, Z., & DeGrado, W. (1988) *Science 240*, 1177–1181.

Lemmon, M., Flanagan, J., Hunt, J., Adair, B., Bormann, B.-J., Dempsey, C., & Engelman, D. (1992) *J. Biol. Chem. 267*, 7683–7689.

Liao, M.-J., London, E., & Khorana, H. (1983) *J. Biol. Chem. 258*, 9949–9955.

Liao, M.-J., Huang, K.-S., & Khorana, H. (1984) *J. Biol. Chem. 259*, 4200–4204.

Litman, B. (1979) *Photochem. Photobiol. 29*, 671–677.

Marti, T., Otto, H., Mogi, T., Rösselet, S., Heyn, M., & Khorana, H. (1991) *J. Biol. Chem. 266*, 6919–6927.

Morrissey, J. (1981) *Anal. Biochem. 117*, 307–310.

Oiki, S., Danho, W., & Montal, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2393–2397.

Popot, J.-L., & de Vitry, C. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 369–403.

Popot, J.-L., & Engelman, D. (1990) *Biochemistry 29*, 4031–4037.

Popot, J.-L., Trewhella, J., & Engelman, D. (1986) *EMBO J. 5*, 3039–3044.

Popot, J.-L., Gerchman, S., & Engelman, D. (1987) *J. Mol. Biol. 198*, 655–676.

Rees, D., Komiya, H., Yeates, T., Allen, J., & Feher, G. (1989) *Annu. Rev. Biochem. 58*, 607–633.

Sigrist, H., Wenger, R., Kislig, E., & Wüthrich, M. (1988) *Eur. J. Biochem. 177*, 125–133.

Spencer, H. (1926) in *International Critical Tables of Numerical Data, Physics, Chemistry and Technology. Volume 1* (Washburn, E. W., Ed.) pp 67–68, McGraw-Hill, New York.

Weiss, M., Wacker, T., Weckesser, J., Welte, W., & Schulz, G. (1990) *FEBS Lett. 267*, 268–272.

Weiss, M., Kreusch, S., Schiltz, E., Nestel, U., Welte, W., Weckesser, J., & Schulz, G. (1991) *FEBS Lett. 280*, 379–382.

Yeates, T., Komiya, H., Rees, D., Allen, J., & Feher, G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6438–6442.