Biophys Struct Mech (1983) 9:235–244

Biophysics of Structure and Mechanism


# The Structure of Bovine Rhodopsin*

P. A. Hargrave[1,2], J. H. McDowell[1], Donna R. Curtis[1], Janet K. Wang[1], Elizabeth Juszczak[1], Shao-Ling Fong[1], J. K. Mohana Rao[3], and P. Argos[3]

[1] Department of Medical Biochemistry, School of Medicine, Southern Illinois University, Carbondale, IL 62901, USA
[2] Department of Chemistry and Biochemistry, College of Science, Southern Illinois University, Carbondale, IL 62901, USA
[3] Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

**Abstract.** We have isolated 16 peptides from a cyanogen bromide digest of rhodopsin. These cyanogen bromide peptides account for the complete composition of the protein. Methionine-containing peptides from other chemical and enzymatic digests of rhodopsin have allowed us to place the cyanogen bromide peptides in order, yielding the sequence of the protein. We have completed the sequence of most of the cyanogen bromide peptides. This information, in conjunction with that from other laboratories, forms the basis for our prediction of the secondary structure of the protein and how it may be arranged in the disk membrane.

**Key words:** Rhodopsin – Amino acid sequence – Secondary structure – Topography – Disk membrane

## Introduction

A principal goal of vision research is elucidation of the molecular events of the visual transduction process. Rhodopsin initiates this processing by absorbing a photon – an event which leads to a change in conformation of the protein. It is this change in protein conformation which appears to express rhodopsin's mechanism of action as a light receptor, initiating the light-dependent biochemical events of the rod cell. An understanding of rhodopsin structure, therefore, is central to an understanding of the visual process, and will establish the framework for understanding the mechanisms by which rhodopsin functions as a visual receptor.

Studies in our laboratory have been directed toward the determination of the covalent sequence of rhodopsin and its disposition within the disk membrane lipid bilayer. We have further sought to locate sites within the structure of

* Based on material presented at the Fifth International Congress of Eye Research, Eindhoven, October 1982

rhodopsin which are important for various functions of the receptor protein. In this report we present new results and summarize our current level of understanding of rhodopsin's structure.

## Strategy for the Sequence of Rhodopsin

Rhodopsin's polypeptide chain has a molecular weight of approximately 39,000 daltons and contains nearly 350 amino acids. In order to obtain the amino acid sequence of a protein of this size it is first necessary to fragment the protein into a number of smaller peptides. Each peptide must then be purified and its sequence determined.

Determining the sequence of soluble globular proteins is now routine. Applying the same simple strategy to determining the structure of hydrophobic membrane proteins such as rhodopsin should, in theory, work equally as well. In practice it has proven difficult to obtain many of rhodopsin's hydrophobic peptides in pure form and in reasonable yield. Over a period of years, however, better methods for handling of hydrophobic peptides have been developed (e.g., Gerber et al. 1979) and new techniques have proven useful in separation of peptides from rhodopsin (Pellicone et al. 1980; Findlay et al. 1981; Hargrave et al. 1982a).

## Cyanogen Bromide Cleavage of Rhodopsin Yields Valuable Peptides

We have found cleavage with cyanogen bromide to be the most useful approach to obtain reasonable yields of rhodopsin peptides of moderate size (Hargrave et al. 1982a). Either the entire protein is cleaved or the large proteolytic fragments F1 and F2 are prepared first and then submitted to CNBr cleavage. We have been able to separate 16 such peptides and have obtained the complete sequence of 14 of them. Our approach has been to first partially fractionate the peptide mixture by size on Sephadex G50 in 20% formic acid. Purified peptides were then obtained following further steps such as gel filtration in formic acid/ethanol, ion exchange chromatography or paper electrophoresis. Several of the purified CNBr peptides had to be further fragmented by either chemical or enzymatic methods before their complete sequence could be obtained. Sequence determination was performed manually, or by solid phase sequencer, or by spinning cup automated sequencer, depending upon the peptide to be sequenced. The sequences of these peptides are shown in Table 1.

## Placing the Cyanogen Bromide Peptides in Order

The order in which the CNBr peptides are arranged in the rhodopsin polypeptide chain had to be determined from the results of complementary types of fragmentation (digestion) experiments. Peptides which contained methionine internally in their sequence were obtained from limited proteolysis [e.g., F1 and

**Table 1.** The cyanogen bromide peptides of bovine rhodopsin. Designations for the peptides derive from their gel filtration column elution positions (Hargrave et al. 1982a). CB-V peptides elute in the void volume. Peptides are listed in the table in their order of location in the covalent sequence of rhodopsin. Amino acids whose position in the sequence remain to be resolved are shown in brackets. Dots indicate a region of the amino acid sequence which is not yet completed

| Peptide | Tentative sequence |
|---|---|
| CB-11 | Ac-Met |
| CB-1 | CHO CHO<br>\| \|<br>Asn-Gly-Thr-Glu-Gly-Pro-Asn-Phe-Tyr-Val-Pro-Phe-Ser-Asn-Lys-Thr-Gly-Val-Val-Arg-Ser-Pro-Phe-Glu-Ala-Pro-Gln-Tyr-Tyr-Leu-Ala-Glu-Pro-(Trp, Glx)-Phe-Ser-Met |
| CB-7 | Leu-Ala-Ala-Tyr-Met |
| CB-8 | Phe-Leu-Leu-Ile-Met |
| CB-2 | Leu-Gly-Phe-Pro-Ile-Asn-Phe-Leu-Thr-Leu-Tyr-Val-Thr-Val-Gln-His-Lys-Lys-Leu-Arg-Thr-Pro-Leu-Asn-Tyr-Ile-Leu-Leu-Asn-Leu-Ala-Val-Ala-Asp-Leu-Phe-Met |
| CB-V1a | Val-Phe-Gly-Gly-Phe-Thr-Thr-Thr-Leu-Tyr-Thr-Ser-Leu-His-Gly-Tyr-Phe-Val-Phe-. . . Met |
| CB-5 | Ser-Asn-Phe-Arg-Phe-Gly-Glu-Asn-His-Ala-Ile-Met |
| CB-4c | Gly-Val-Ala-Phe-Thr-Xxx-Val-Met |
| CB-4a | Ala-Leu-Ala-Cys-Ala-Ala-Pro-Pro-Leu-Val-Gly-Trp-Ser-Arg-Tyr-Ile-Pro-Glu-Gly-Met |
| CB-4b | Gln-(Cys, Ser)-Cys-Gly-Ile-Asp-Tyr-Tyr-Thr-Pro-His-Glu-Glu-Thr-Asn-Asn-Glu-Ser-Phe-Val-Ile-Tyr-Met |
| CB-V1b | Phe-Val-Val-His-Phe-Ile-Ile-Pro-Leu-Ile-Val-Ile-Phe-Phe-Cys-Tyr-Gly-Gln-Leu-Val- . . . Ser-Ala-Thr-Thr-Gln-Lys-Ala-Glu-Lys-Glu-Val-Thr-Arg-Met |
| CB-9 | Val-Ile-Ile-Met |
| CB-V2 | Val-Ile-Ala-Phe-Leu-Ile-Cys-Trp-Leu-Pro-Tyr-Ala-Gly-Val-Ala-Phe-Tyr-Ile-Phe-Thr-His-Gln-Gly-Ser-Asp-Phe-Gly-Pro-Ile-Phe-Met-Thr-Ile-Pro-Ala-Phe-Phe-Ala-Lys-Thr-Ser-Ala-Val-Tyr-Asn-Pro-Val-Ile-Tyr-Ile-Met |
| CB-10 | Met |
| CB-6 | Asn-Lys-Gln-Phe-Arg-Asn-Cys-Met |
| CB-3 | Val-Thr-Thr-Leu-Cys-Cys-Gly-Lys-Asn-Pro-Leu-Gly-Asp-Asp-Glu-Ala-Ser-Thr-Thr-Val-Ser-Lys-Thr-Glu-Thr-Ser-Gln-Val-Ala-Pro-Ala |

F2; Hargrave et al. 1980] and from complete proteolysis of rhodopsin with pepsin (Curtis et al. 1982), thermolysin or trypsin. Peptides which establish overlaps of the CNBr peptides were obtained as shown in Table 2. The results reported were obtained from six different types of experiments. Many of the overlapping sequences were observed again and again from peptides isolated from several different digestion experiments. These data led us to propose the following order for the CNBr peptides of rhodopsin (Hargrave et al. 1982a).

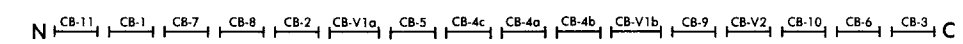

**Table 2.** Peptides providing overlap of rhodopsin CNBr peptides

| Cyanogen bromide peptides overlapped | Type of cleavage | Rhodopsin derivative | Peptide sequence |
|---|---|---|---|
| [CB-11]–[CB-1] | Trypsin | RAE[a] | Ac-Met-Asn-Gly[b] . . . |
| [CB-1]–[CB-7] | Pepsin | – | Ser-Met-Leu[c] |
| [CB-7]–[CB-8] | Pepsin | OX[a] | Ala-Ala-Tyr-Met-Phe[d] |
| [CB-8]–[CB-2] | CNBr | RAE | Phe-Leu-Leu-Ile-Met-Leu-Gly[e] . . . |
| [CB-2]–[CB-V1a] | Pepsin | OX | Phe-Met-Val-Phe-Gly-Gly-Phe[d] |
| [CB-V1a]–[CB-5] | Pepsin | OX | Arg-Tyr-Val-Val-Val (Cys, Lys, Pro, Met, Ser, Asx, Phe)[d] |
| [CB-5]–[CB-4c] | Pepsin | OX | His-Ala-Ile-Met-Gly-Val-Ala[d] |
| [CB-4c]–[CB-4a] | Thermolysin | OX | Val-Met-Ala[c] |
| [CB-4a]–[CB-4b] | Trypsin | RAE | Tyr-Ile-Pro-Glx-Gly-Met-Glx (Ser, Cys)[c] |
| [CB-4b]–[CB-V1b][f] | – | – | – |
| [CB-V1b]–[CB-9] | Thermolysin | – | Amino-terminal sequence fragment F2[g, h] |
| [CB-9]–[CB-V2] | Thermolysin | – | Amino-terminal sequence fragment F2[g, h] |
| [CB-V2]–[CB-10] | Pepsin | OX | Ile-Met-Met-Asx-Lys-Glx[d] |
| [CB-10]–[CB-6] | Pepsin | OX | Ile-Met-Met-Asx-Lys-Glx[d] |
| [CB-6]–[CB-3] | Trypsin | RAE; cit.[a] | Asn-Cys-Met-Val-Thr-Thr[j] . . . |

[a] RAE, reduced and aminoethylated; OX, performic acid oxidized; cit., reacted with citraconic anhydride.
[b] Hargrave (1977)
[c] Hargrave (unpublished)
[d] Curtis, McDowell and Hargrave (1982)
[e] This peptide is found in low yield due to incomplete cleavage by CNBr (McDowell and Hargrave, unpublished)
[f] This overlap within the sequence of F1 is obtained by difference since the positions of all other CNBr peptides from F1 have been uniquely determined.
[g] Siemiatkowski-Juszczak (1981)
[h] Hargrave et al. (1982b)
[j] McDowell and Griffith (1978)

Completion of the sequence of peptides CB-V1a and CB-V1b will complete our sequence determination of rhodopsin.

## The Complete Amino Acid Sequence of Bovine Rhodopsin

The efforts of many laboratories have contributed to current knowledge of the structure of rhodopsin. Progress in achieving the sequence of rhodopsin was aided by isolation and characterization of an 83-amino acid hydrophobic peptide from the carboxyl-terminal region of rhodopsin (Pellicone et al. 1980). The sequence was extended to include the entire carboxyl-terminal one-third of rhodopsin for both the sheep (Findlay et al. 1981) and cattle rhodopsins (Siematkowski-Juszczak 1981). Knowledge of the structure of rhodopsin has been summarized (Hargrave 1982). While this manuscript was in preparation,

the complete amino acid sequence of bovine rhodopsin was published (Ovchinnikov et al. 1982). The results which we report (Table 1) are in generally good agreement. The proposed sequence of Ovchinnikov and coworkers completes the gaps in our peptides CB-V1a and CB-V1b and shows the (Trp, Glx) in our peptides CB-1 to be Trp-Gln, and the (Cys, Ser) in our peptide CB-4b to be Cys-Ser. Also, the amino acid in peptide CB-4c which we have not positively identified, the Soviet group reports to be tryptophan.

Comparison of the two sequences does reveal one significant difference. We find the sequence . . . Ile-Asp-Tyr-Tyr-Thr-Pro . . . in our peptide CB-4b, whereas the sequence by Ovchinnikov et al. (1982) reports only a single tyrosyl residue. Our amino acid analysis of peptide CB-4b (Hargrave 1982; Table 2), however, shows the three tyrosines which we find by sequence analysis. Sequence of this peptide was complicated by acid-catalyzed cyclization of the N-terminal glutamine which blocked the peptide amino-terminus, thereby preventing it from being sequenced directly. We prepared the peptide from carboxymethylated rhodopsin and obtained fragments from it both by limited alkaline hydrolysis and by pepsin digestion (Wang, McDowell and Hargrave, unpublished experiments). A peptic peptide (DF) obtained from CB-4b had the amino acid sequence Ile-Asp-Tyr-Tyr-Thr-Pro-His-Glu-Glu-Thr-Asn-Asn-Glu(Ser,Phe). [Peptide sequencing was performed on a Beckman Protein Sequencer, using polybrene. Amino acid residues were identified both by HPLC and following HI hydrolysis.] Other peptic peptides which overlapped peptide DF also gave analyses and/or sequences supportive of the Tyr-Tyr linkage, so we feel that this assignment is secure. If Ovchinnikov and coworkers relied solely on the sequence of their peptide XT-4 (obtained by cleavage at tyrosyl residues) for the protein sequence in this region, it would be easy to understand how a Tyr-Tyr sequence could escape detection. If there are no other revisions necessary to the sequence, then bovine rhodopsin would contain 18 tyrosyl residues and a total of 348 amino acids.

## Rhodopsin's Sequence Contains Alternate Hydrophilic and Hydrophobic Segments

Inspection of the complete sequence of rhodopsin (Table 1, and Ovchinnikov et al. 1982) shows that there are regions which can be recognized to be predominantly polar and hydrophilic in composition, and other regions which are predominantly non-polar and hydrophobic in their amino acid content. The predominantly hydrophobic regions are from 19 to 28 amino acids in length and are rich in the amino acids valine, leucine, isoleucine, and phenylalanine. The alternating hydrophilic regions are variable in length and contain the majority of the charged amino acids of the protein. We observed this type of pattern earlier in examination of the sequence of the carboxyl-terminal one-third of rhodopsin (Hargrave et al. 1982b). Numerous topographical studies using both chemical and enzymatic probes (reviewed in Hargrave 1982) showed that the hydrophilic regions were located at the membrane aqueous surfaces and that the hydrophobic regions were embedded in the membrane lipid bilayer. This led to

development of the topographic model in which rhodopsin's polypeptide chain traversed the lipid bilayer, exposed a hydrophilic loop region and then reentered and traversed the lipid bilayer (Hargrave et al. 1982b).

## Seven Hydrophobic Helices are Predicted for Rhodopsin

It would be reasonable to think of the transmembrane segments of the polypeptide chain of bovine rhodopsin as $\alpha$-helices by analogy to what is know about other integral membrane proteins. However, when we applied a typical helix-prediction scheme (which uses soluble globular proteins as its data base) to the sequence of the purple membrane protein "bacteriorhodopsin" (which contains seven transmembrane helices), it yielded a random correlation of 0.01. This prompted us to develop a helix prediction algorithm with a multiple-parameter data base designed to predict the location of hydrophobic helices in membrane proteins (Argos et al. 1982). When this algorithm is applied to the rhodopsin sequence, seven helices are predicted (Fig. 1). This number of helices had been previously postulated on the basis of the physical measurements made on rhodopsin by many laboratories (Hargrave and McDowell 1981). The predicted helices vary from 21 to 28 amino acids ($\bar{x} \cong 24$), a length which is quite adequate to span the distances of the fatty acid side chains in the lipid bilayer.

## A Model for Rhodopsin in the Membrane

The carboxyl-terminal region of rhodopsin has been demonstrated to be exposed at the disk membrane cytoplasmic surface (Hargrave and Fong 1977). Given this location of the carboxyl-terminus, and the presence of seven transmembrane helices, the path of the polypeptide chain is defined as shown in Fig. 2. This uniquely places the helix-connecting segments at alternating membrane surfaces. The amino-terminus is thereby positioned at the intradiskal surface, in accord with experimental data. To our knowledge the surfaces assigned to the helix-connecting regions are in accord with all chemical and enzymatic topographic studies.

In drawing the helices in Fig. 2 we have made only minor adjustments in helical lengths compared to that predicted by the algorithm in Fig. 1. For example, when the parametric sum changed from negative to positive at the amino-terminal end (top) of helix 6, the arginine was actually included in the helix. We chose, in our model, to exclude the charged arginine from the "top" of helix 6 and to include phenylalanine at the "bottom" of the helix-choices which represent only minor deviations from the arbitrary zero point of the parametric sum (Fig. 1). The model predicts that the lysine to which retinal is bound is located near the midpoint of the lipid bilayer (in helix 7). Other charged groups are predicted to be buried, including an aspartic acid in helix 2, a glutamic acid in helix 3, and a potentially charged histidine in helix 5. It is gratifying to find two buried negatively charged groups since such groups have been predicted on

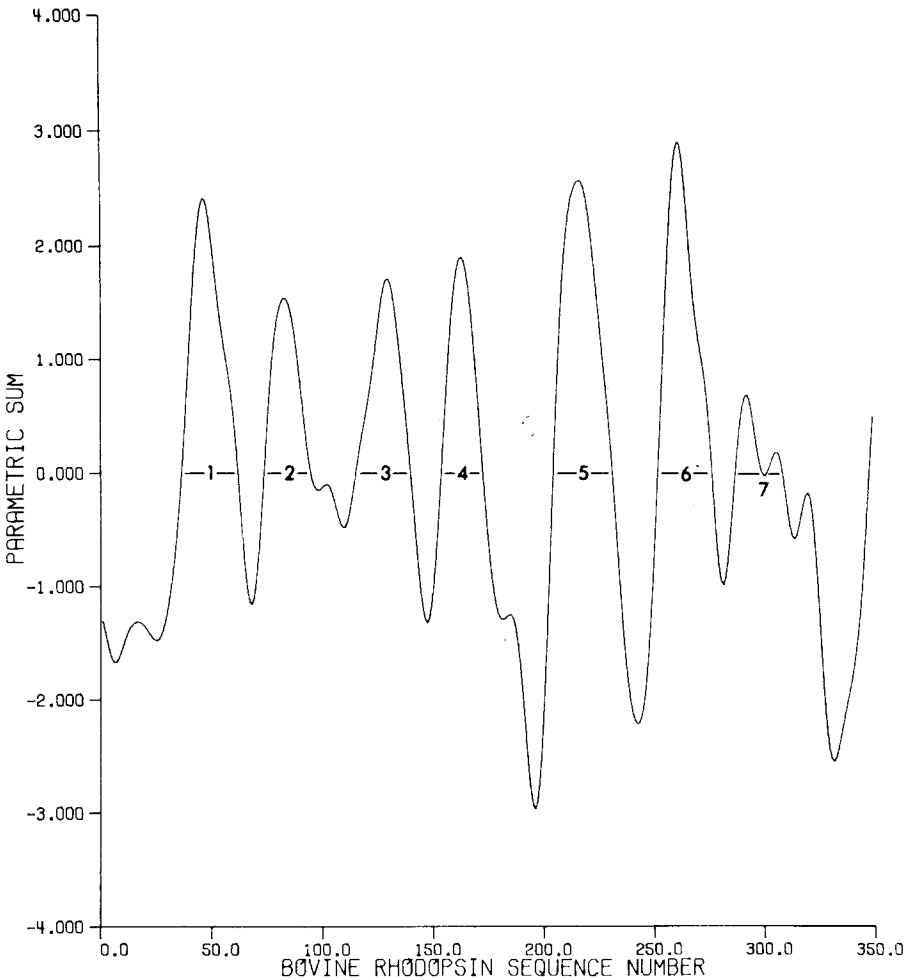


**Fig. 1.** Prediction of transmembrane helices of bovine rhodopsin. The sum of nine physical parameters is plotted vs amino acid residue number of the protein sequence. Residues having positive values of the parametric sum are predicted to occur in helical conformation

theoretical grounds to be necessary to explain the spectral properties of rhodopsin (see Honig et al. 1979, and references reviewed in Hargrave 1982). The aspartic and glutamic acid residues might be expected to serve as counterions to the histidine and retinyl-lysine. In addition, a Glu-Arg sequence and its adjacent Tyr-Val-Val-Val-Cys are predicted to exist in the lipid bilayer rather than to be exposed at the aqueous interface. These and other predictions are clearly testable and should lead investigators to perform further topographic labelling experiments to seek evidence for the location of specific groups in the molecule.

The model places one-half of rhodopsin's mass in the bilayer and distributes the remainder equally at each membrane surface. We would expect rhodopsin's

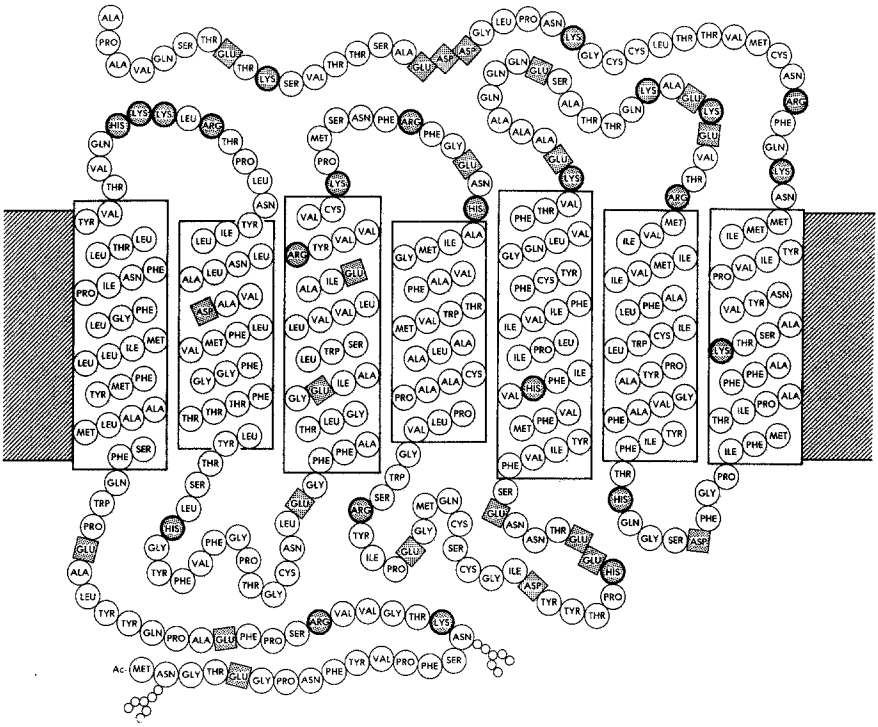

**Fig. 2.** A model for the organization of rhodopsin's polypeptide chain with respect to the disk membrane lipid bilayer. Striped areas at each side of the drawing indicate the low polarity portion of the lipid bilayer which includes the fatty acid side chains of the phospholipids and the lower half of the phospholipid polar head groups. The amino-terminal Ac-Met is at the bottom of the drawing, exposed to the aqueous interior of the disk membrane. The branched arrays of circles attached to $Asn^2$ and $Asn^{15}$ represents oligosaccharides. The carboxyl-terminal Ala at the top of the drawing is exposed to the rod outer segment cytoplasm. Positively charged amino acids are shown as shaded heavy circles and negatively charged amino acids as shaded squares

helices to be arranged in a bundle such as that shown in Fig. 3. When rhodopsin is compared with the small number of well characterized membrane proteins, only a few common features are evident. For example, both the purple membrane protein of *Halobacterium halobium* (Henderson 1977) and bovine rhodopsin are integral membrane proteins which appear to consist of seven transmembrane helices and which contain a molecule of bound retinal. Details of the structure and function of these two proteins, however, are quite different. We have compared their amino acid sequences and find no statistically significant homology.

Current structural, theoretical and topographic studies on rhodopsin form the basis for construction of our working model. In the future we would hope to come closer to an understanding of how the structure of the protein is related to various aspects of rhodopsin function including wavelength regulation in visual pigments and interaction with other components in the visual system.

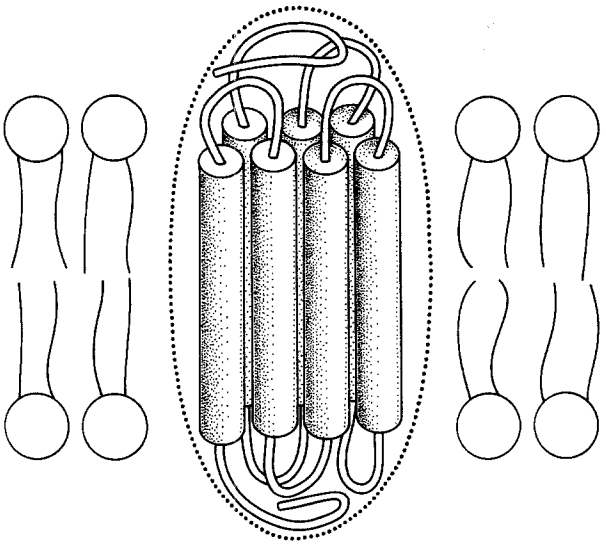

**Fig. 3.** A model for the location and structure of rhodopsin in the disk membrane. Rhodopsin is shown as an elongated bundle of helices. Half of the mass of the protein is embedded in the lipid bilayer, and approximately equal portions of the molecule are exposed at each membrane surface

*Acknowledgements.* This work has been supported by grants from the National Institutes of Health (EY1275 and EY2875, to P.A.H.; GM27682 to P.A.), the National Science Foundation (equipment grant PCM77-17808 to P.A.H.), and the American Cancer Society (FRA173 to P.A.).

## References

1. Argos P, Mohana Rao JK, Hargrave PA (1982) Structural prediction of membrane-bound proteins. Europ J Biochem 128: 565–575
2. Curtis DR, McDowell JH, Hargrave PA (1982) Isolation of peptides containing any desired amino acid: methionyl peptides of bovine rhodopsin. (Submitted for publication)
3. Findlay JBC, Brett M, Pappin DJC (1981) Primary structure of C-terminal functional sites in ovine rhodopsin. Nature (London) 293: 314–316
4. Gerber GE, Anderegg RJ, Herlihy WC, Gray CP, Biemann K, Khorana HG (1979) Partial primary structure of bacteriorhodopsin: sequencing methods for membrane proteins. Proc Natl Acad Sci USA 76: 227–231
5. Hargrave PA, Fong SL (1977) The amino- and carboxyl-terminal sequence of bovine rhodopsin. J Supramol Struct 6: 559–570
6. Hargrave PA, Fong S-L. McDowell JH, Mas MT, Curtis DR, Wang JK, Juszczak E, Smith DP (1980) The partial primary structure of bovine rhodopsin and its topography in the retinal rod cell disc membrane. Neurochem Int 1: 231–244
7. Hargrave PA, McDowell JH (1981) Highly speculative model for rhodopsin in the disc membrane. In: Abrahamson EW, Ostroy SE (eds) Molecular processes in vision. Hutchinson Ross, Stroudsburg, Pennsylvania, pp 186–189
8. Hargrave PA, McDowell JH, Curtis DR, Fong, S-L, Juszczak E (1982a) Preparation of peptides from bovine rhodopsin. In: Packer L (ed) Methods in enzymology, vol 81. Academic Press, New York, pp 251–256
9. Hargrave PA, McDowell JH, Curtis DR, Wang JK, Juszczak EC (1982c) Structure of bovine rhodopsin. Invest Ophthalmol Vis Sci 22: 229

10. Hargrave PA, McDowell JH, Siemiatkowski-Juszczak EC, Fong S-L, Kühn H, Wang JK, Curtis DR, Mohana Rao JK, Argos P, Feldmann RJ (1982b) The carboxyl-terminal one-third of bovine rhodopsin: its structure and function. Vis Res 22: 1429–1436
11. Hargrave PA (1982) Rhodopsin chemistry, structure and topography. Prog Retinal Res 1: 1–51
12. Henderson R (1977) The purple membrane from *Halobacterium halobium*. Annu Rev Biophys Bioeng 6: 87–109
13. Honig B, Dinur U, Nakanishi K, Balogh-Nair V, Gawinowicz MA, Arnaboldi M, Motto MG (1979) An external point-charge model for wavelength regulation in visual pigments. J Am Chem Soc 101: 7084–7086
14. McDowell JH, Griffith K (1978) Localization of a dark reactive cysteine in rhodopsin. Invest Ophthamol Vis Sci [Suppl 1] 17: 125
15. Ovchinnikov YA, Abdulaev NG, Feigina MY, Artamonov ID, Zolotarev AS, Kostina MB, Bogachuk AS, Miroshnikov AI, Martinov VI, Kudelin AB (1982) The complete amino acid sequence of visual rhodopsin. Bioorg Khim 8: 1011–1014
16. Pellicone C, Bouillon P, Virmaux N (1980) Purification et séquence partielle d'un polypeptide hydrophobe de la rhodopsine bovine fragmentée par le BNPS-scatole. CR Acad Sci Paris 290: 567-569
17. Siemiatkowski-Juszczak EC (1981) Limited proteolysis of bovine rhodopsin and sequence determination of the F2 fragment. Ph.D. dissertation, Southern Illinois University

Received November 12, 1982