

Path of the polypeptide in bacteriorhodopsin

(purple membrane/diffraction/protein folding)

D. M. ENGELMAN*, R. HENDERSON, A. D. MCLACHLAN, AND B. A. WALLACE†

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Communicated by Frederic M. Richards, January 7, 1980

ABSTRACT An attempt has been made to fit the amino acid sequence of bacteriorhodopsin to the three-dimensional density map of the molecule. First, seven segments of the sequence were selected as being probable transmembrane α helices. Then each of the 5040 possible ways of fitting these seven segments into the seven regions of helical density in the map were evaluated based on the criteria of connectivity of the nonhelical link regions, charge neutralization, and total scattering density per helix. A single model that may be experimentally tested emerged as the most probable.

Bacteriorhodopsin is a transmembrane protein found in *Halobacter halobium*. Light energy absorbed by retinal bound to the protein is used to pump protons across the membrane, and the proton gradient is then used as an energy source (for review see ref. 1). The molecules of this protein exist naturally in a highly ordered two-dimensional array. A low-resolution map (2) of the electron-scattering density of purple membrane shows that each bacteriorhodopsin molecule is composed of seven rods of density oriented perpendicular to the membrane surface. These are believed to represent α -helical segments of the polypeptide that cross the membrane. The molecular boundary of a single bacteriorhodopsin molecule is now known from observation of the same molecule in a different crystal environment (3).

The complete amino acid sequence of bacteriorhodopsin has been determined (4) and portions of it have been independently confirmed (5, 6). Ovchinnikov *et al.* (4) used the proteolytic cleavage points of the protein in the native membrane to propose an arrangement of the polypeptide in the membrane. This consisted of seven α -helical segments of polypeptide, each of which contained between 26 and 32 amino acids, with short nonhelical segments of up to 8 amino acids linking them. Adjacent helices in the sequence had opposite orientations in the membrane. Additional cleavage sites at the NH_2 and COOH termini determined by Walker *et al.* (6) provide further constraints on the polypeptide arrangement.

A useful step towards a more detailed description of the structure of bacteriorhodopsin would result from a determination of the way in which the sequence fits into the density map. There are, *a priori*, 5040 (7!) ways of doing this. There would be 10,080 ways if the orientation of the sequence and density map were not known. However, the correlation of electron diffraction with freeze fracture (7, 8) has established the orientation of the map with respect to the outer surface of the cell. The top of the model in our convention is the cytoplasmic surface. Similarly, digestion experiments on inside-out vesicles have demonstrated that the carboxyl end of the sequence is on the cytoplasmic surface of the membrane (4, 9).

In this paper we examine the 5040 possible models by using three criteria: the lengths of links between helix ends, the formation of ion pairs in the protein interior, and the electron

scattering power of each helix. We reject a large number of models as unrealistic and propose a single model as the most probable.

RESULTS AND DISCUSSION

Arrangement of polypeptide in membrane

Any considerations limiting the ways in which a model can be constructed depend on having a reliable idea of the arrangement of the polypeptide segments across the membrane. We must, therefore, make an estimate of the exact lengths of helical segments and link regions connecting them before examining ways of fitting the polypeptide segments of the sequence into the density map.

Fig. 1 shows a choice of helical segments rather different from that proposed by Ovchinnikov *et al.* (4). In deciding the exact positions of the ends of the helices, we have endeavored to construct an arrangement conservatively, where the only residues included in the helical segments are those that fit strong criteria of hydrophobicity and inaccessibility to proteolytic cleavage. The accessibility to proteolytic cleavage is clearly a property of surface residues; the hydrophobicity in the helical segments is preferred on energetic grounds for residues that are either buried inside the protein or below the surface of the lipid bilayer. Correspondingly, the occurrence of a high density of charged and polar residues is taken as a strong indication of a nonhelical link region. We have also tried to avoid assumptions that result in either unusually short or long helices because the known membrane thickness, the absence of substantial surface projections, and the three-dimensional density map suggest that the helices should be comparable in length. As a result, the lengths of our nonhelical link regions are rather longer than may eventually be found. For example, they are on average three residues longer than those proposed by Ovchinnikov *et al.* (4).

Several features of Fig. 1 are important, some of which have already been noted (4). First, most of the charged residues (Asp, Glu, Lys, and Arg) are on one or the other surface. On the outer surface of the membrane, there are six charged residues that are either entirely accessible or are close enough to the end of a helix so that their side chains can reach the solvent. Similarly, there are 19 charged residues at the cytoplasmic surface.

Second, there are nine charged residues in Fig. 1 that are sufficiently far from either membrane surface to make direct interaction with water unlikely. These charged side chains would be energetically very difficult to bury except as self-neutralizing ion pairs. Of these nine, four may naturally form ion pairs within the same helix because they are separated by either three or four residues in the sequence. These are (Arg-82 and Asp-85) and (Asp-211 and Lys-215). Other ion pairs may

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

* Present address: Department of Molecular Biophysics and Biochemistry, Yale University, Box 1937 Yale Station, New Haven, CT 06520.

† Present address: Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, NY 10032.

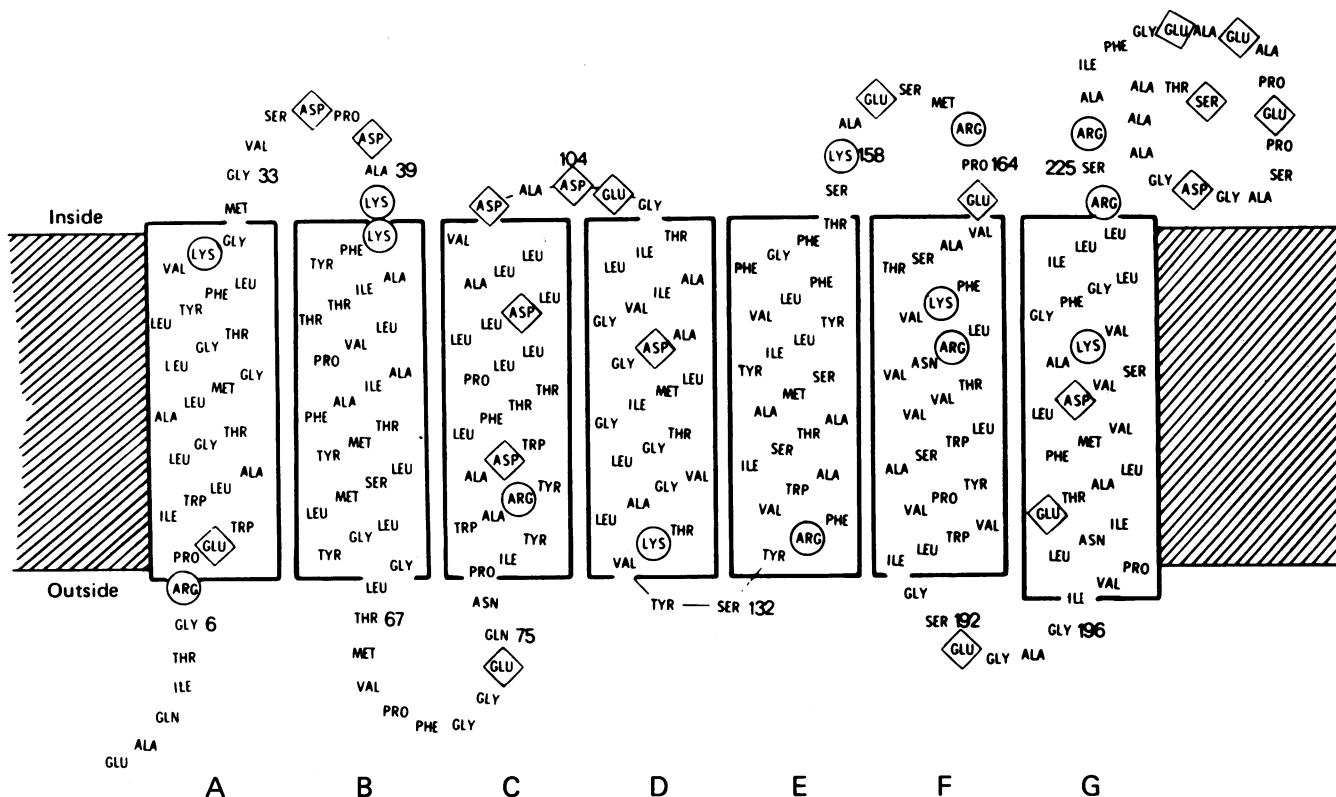


FIG. 1. Suggested arrangement of the polypeptide of bacteriorhodopsin across the membrane. The seven helices are labeled A-G, starting from the amino terminus. The sequence is from Ovchinnikov *et al.* (4). The hatched marks indicate the approximate location of the lipid hydrocarbon regions. O and ◇ indicate positive and negative charges, respectively.

exist between the positively charged residues of helix F (Lys-171 and Arg-174) and the otherwise isolated negative charges on Asp-96 of helix C and Asp-115 of helix D. These pairings could neutralize four more buried charges and will be considered later.

Thus, Glu-203 is the only other deeply buried single charge, and this may even form a second salt bridge with Arg-82. In fact, Arg-82 is the only residue that Glu-203 can reach. Another notable feature of the arrangement of the polypeptide in Fig. 1 is that the α carbon of lysine 41 is near the surface, approximately 19 Å from the membrane center. The retinal attached to lysine 41 (ref. 10) will extend no closer to the center than 9 Å and is nearer to the inner membrane surface than proposed previously (2). This position is in conflict with reports from neutron analysis (11).

Application of criteria for correlation of sequence with three-dimensional density map

Connectivity of Polypeptide Must Be Possible. The first criterion that we use to eliminate unrealistic fits of the helices to the densities is that the lengths of the polypeptides linking the ends of the α helices must be sufficient to connect the ends of the features in the three-dimensional map. To apply this criterion systematically, we use the numbering shown in Fig. 2 to refer to the rods of helix density in the map, together with the lettering of Fig. 1 to refer to the helix assigned from the sequence, giving the letters in order of the numbered helices 1-7.

The connectivity criterion is then applied by considering each link in each of the 5040 possible models and eliminating any model in which one or more impossible connections occur. Table 1 gives the estimated number of residues allowed in each link between connected helices in the sequences and the maximal length of the link if the chain is in a fully extended con-

formation (3.6 Å per residue). Alternate links lie on the inside and outside surfaces. Table 2 gives the measured center-to-center distance between the ends of the rods in the three-dimensional map. Note that the distances between the ends of the pairs of rods are different on the two surfaces of the membrane because of the different tilts of the helices. If the connecting path is obstructed by an intervening helix, we add 5 Å to the

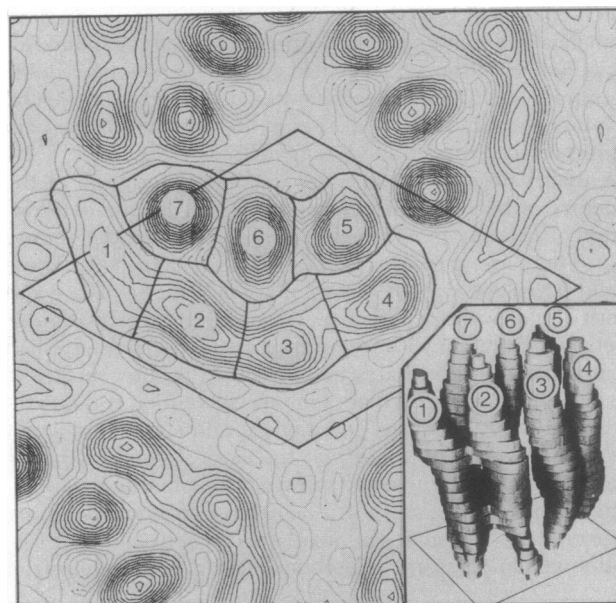


FIG. 2. Convention used for numbering the seven regions of presumed helical density in the projection (from ref. 12) and (Inset) the three-dimensional map (from ref. 2). The areas used in the density calculation are shown.

Table 1. Numbers of residues in links between helices

Helix pair	Side of membrane	Residues	Maximum length of link, Å*
A-B	In	10†	36.0
B-C	Out	9†	32.4
C-D	In	5	18.0
D-E	Out	4†	14.4
E-F	In	8	28.8
F-G	Out	6	21.6

* Maximum length assumed as 3.6 Å per residue fully extended.

† The numbers of residues in these links are slightly different from those in Fig. 1, where A-B = 9, B-C = 11, and D-E = 2. The four residues between D and E are computationally necessary because we have assumed that links must span the center-to-center distance between helices. The two links A-B and B-C are very long and changes in them only affect structures containing links greater than 32 Å.

measured lengths in Table 2 to allow for the longer linkages required; if the path is obstructed by several intervening helices and is extra long, we add 10 Å to the measured length of the linkage. The connectivity test is applied by a simple computer program and reduces the number of feasible models to 405. These 405 are then ranked in order of increasing total link length, models with longer link lengths being considered less likely. At this stage the number of possible models is intentionally large. If we used shorter link lengths (e.g., ref. 4), the number would be substantially smaller. Also, in some of the models (144), the polypeptides forming two different links cross over one another on the same side of the membrane. Such a topography has not been found in any of the five known protein structures that contain groups of four parallel helices, so we consider these crossed models less likely. In any crossover, at least one of the links would need to take a longer path.

Buried Charges Must Form Ion Pairs. As mentioned above, the isolated charges of Asp-96, Asp-115, Lys-171, and Arg-174 are considered to form ion pairs with charges on other helices. These may neutralize one another directly in a simple C-F and

D-F interaction as described earlier (and see Table 3) or may form pairs with other charges from the intrahelix pairs (Arg-82 and Asp-85) or (Asp-211 and Lys-215). There are therefore many possible combinations of neutralizing pairs among these eight buried charges.

A hydrogen-bonded Asp-Arg or Asp-Lys ion pair between two helices can only form if the axes of the two helices in the middle depth of the membrane lie closer than 18.5 Å and 17.0 Å, respectively. Table 2 indicates the pairs of helices that satisfy this condition and allow an unobstructed bridge between ion pairs (pairings between helices 1 and 6, 2 and 4, 2 and 5, 3 and 7, and 5 and 7 are considered to be obstructed by other helices).

The vertical separation between charges, along the helix axes, is also an important factor. We considered that the nine buried residues could take part in ion pairs between two compensating charges on helices with axes within 18.5 Å of one another, provided that the paired amino acids were less than eight residues apart vertically. Pairings between residues more than eight apart vertically were allowed only for helix axis pairs within 13.5 Å of one another. Thus, pairs 3-5 and 4-6 were forbidden for the interactions Asp-85 and Lys-171 and Asp-115 and Arg-82, which require a long vertical jump. Pairings between residues more than 11 apart vertically were considered impossible.

A computer program to explore possible ion pairs between different helices was used to set up a table for each model. The pairings within helices between Arg-82 and Asp-85 and Asp-211 and Lys-215 were assumed to be feasible in every model. We then applied the additional strong condition that an acceptable model must allow every one of the four isolated charges Asp-96, Asp-115, Lys-172, and Arg-174 to be neutralized by pairing it off with a unique exclusive partner from among the nine buried charges. This requirement reduced the number of acceptable models from 405 to 172.

All of the surviving models allow the neutralization of at least eight buried charges. Some models also allow the pairing of Glu-203 (to Arg-82), making possible the neutralization of all nine buried charges.

Table 3 lists the best 35 of these models, which come from the top quarter of the original ranked list of 405 models. Models 36-41 have been added for special reasons, described later.

The simplest way to satisfy the strong condition is to place helix F close to both C and D (indicated by f in Table 3), but this condition is not essential. Models with helix C near G (g in Table 3) allow an Arg-82-Glu-203 pair. Some models can only satisfy the strong condition by pairing Asp-115 with Arg-82 because helix D is close to C but not to F and G. This pairing is unfavorable and may indeed be impossible, depending on the precise vertical registration between helices, so that such structures (d in Table 3) have been placed in a special group at the bottom of the table (models 30-35).

Correlation of Electron Scattering Cross Section with Density in Map. The projection map of purple membrane (12) is very accurately determined, as seen objectively by the entirely independent map in the orthorhombic form of purple membrane (3). Thus, it should be possible to use the differences in integrated scattering density in each helix, together with the differences in the chemical composition of each helix, to provide additional constraints on possible fits of the sequence to the map. This has been done in Tables 4 and 5. Table 4 gives the integrated density of the regions of the projection map of the p3 structure (12) corresponding to each helix. Fig. 2 shows the areas used. The boundary of the protein was chosen to have an area of 68% of the asymmetric unit. With this area, the total mass of the protein would be 75% of the whole, assuming that

Table 2. Measured distances (Å) between helices in structure

Helix pair	Inside surface (top)	Outside surface (bottom)	Middle
1-2	11.5	10.0	11.5 I
2-3	10.0	8.0	12.0 I
3-4	10.0	11.0	12.5 I
4-5	12.0	9.0	10.5 I
5-6	11.5	8.5	10.0 I
6-7	9.5	10.0	10.0 I
1-7	11.0	13.5	9.5 I
2-6	11.5	14.0	12.0 I
2-7	9.5	14.0	10.5 I
3-6	10.5	13.0	13.0 I
4-6	12.0	14.5	15.0 I
1-3	20.5 (+5)	17.0 (+5)	23.0
1-6	19.0 (+5)	17.0 (+5)	18.0
2-4	20.0 (+5)	18.5 (+5)	21.5
2-5	23.5 (+5)	19.5 (+5)	21.0
3-5	17.5 (+5)	12.5	17.0 I
3-7	15.5 (+5)	21.5 (+5)	19.0
4-7	21.5 (+5)	24.0 (+10)	25.0
5-7	19.5 (+5)	20.0 (+5)	19.0
1-4	30.0 (+10)	27.0 (+10)	32.0
1-5	30.0 (+10)	25.0 (+10)	28.0

I indicates the pairs of helices that may form ion pairs. The (+5 Å) and (+10 Å) distances, where indicated, are used to augment the measured distances as described in the text.

Table 3. The 41 best models

Model	Helix position	Total link length, Å	Cross-over	Link types	Ion pairing	Density correlation
1	AGFEDCB	59.5		v ⁶	fg	DA
2	ABCDEFG	60.0		v ⁶	f	
3	BCDEGFA	62.5		v ⁶	f	
4	BCDEFGA	62.5		v ⁶	g	
5	GFCBADE	63.0		v ⁶	f	
6	EFGABCD	63.0		v ⁵ s	fg	
7	BCGFEDA	64.0		v ⁶	g	
8	EDCBAGF	64.0		v ⁶	g	
9	GFCABDE	64.5		v ⁵ s	f	
10	EDCABGF	65.5		v ⁵ s	g	
11	ACDEGFB	66.5		v ⁵ s	f	
12	ACDEFGB	66.5		v ⁵ s	g	
13	ACGFEDB	68.0		v ⁵ s	g	
14	EFGBACD	69.0		v ⁴ sm	fg	
15	BCDGEFA	69.5	X	v ⁴ sl	f	
16	BCDFEGA	70.0	X	v ⁴ sm	g	
17	BGFEDCA	71.5		v ⁵ m	fg	D
18	AGFDECB	72.5	X	v ⁵ m	fg	
19	ECBAGFD	73.5	X	v ⁴ sl	f	
20	AEDFGCB	73.5		v ⁵ l	fg	
21	ACDGEFB	73.5	X	v ³ s ² l	f	
22	ABCGFDE	73.5		v ⁵ l	fg	
23	GDEABCF	74.0	X	v ⁴ sm	f	
24	ACDFEGB	74.0	X	v ³ s ² m	g	
25	BCFGEDA	74.5		v ⁵ l	f	
26	BACDEFG	74.5		v ⁵ l	f	
27	AFGEDCB	74.5		v ⁵ l	g	DA
28	ABCFGDE	75.5	X	v ⁵ l	fg	
29	GCB AEDF	76.0		v ⁴ sl	fg	
30	ADEFGCB	59.5		v ⁶	gd	
31	DCBAGFE	61.0		v ⁵ s	d	
32	GEDABCF	62.0		v ⁵ s	d	
33	GEDBACF	68.0		v ⁴ sm	d	
34	BDEFGCA	71.5		v ⁵ m	gd	
35	ADEGFCB	72.0		v ⁵ m	gd	
36	DCABGFE	76.5		v ⁴ sl	d	D
37	DEBAGFC	77.5		v ⁵ l	d	D
38	DCBGAFE	79.5		v ³ sml	d	DA
39	ABECDFG	80.0		v ⁴ sl	f	DA
40	DCFGABE	82.0		v ³ s ² m	d	DA
41	DCGBAFE	83.0		v ³ s ² l	gd	DA

The later columns of the table give some detailed description of the features of different models. Total link length was computed from the augmented distances in Table 2. X indicates that two links cross over on the same side of the membrane. The link types describe the length of each of the link regions: v means that a connection is very short (8–12 Å), s stands for short (12.5–14.0 Å), m for medium (14.5–17.5 Å), and l for long (18.0–24.0 Å). In the ion-pairing column, f indicates that helix F is close to both C and D, so that the four isolated charges could all neutralize one another directly; g means that helices C and G are close, allowing an Arg-82–Glu-203 pair; d means that helix D is only close to C. Models in which helices D and A are in positions of low scattering density are indicated in the last column. Models 1–29 come from the top quarter of the ranked list of 405 possible connected models. Models 30–35 are also in the top quarter, but have helix D distant from F and G and can only neutralize Asp-115 by making an unusually long ion-pair connection to Arg-82. Models 36–39 have D in low density and lie between the top 25% and top 35% in the ranked list. Models 40 and 41 have both A and D in low density, but lie only between the top 35% and top 50% in the ranked list. Thus, v, f, g, D, and A are all favorable descriptions; model 1 is the only one with all of these.

Table 4. Helix densities in map

Helix	Area on map, Å ²	Integrated density above lipid, contours × Å ²	Absolute integrated density, contours × Å ²	Relative scattering strength
1	126	215	1106	0.955
2	111	302	1193	1.030
3	107	310	1200	1.036
4	109	294	1185	1.023
5	94	224	1115	0.963
6	104	255	1146	0.989
7	102	272	1162	1.004

The total area within each contour level was measured and assigned to the separate helices by using the boundaries shown in Fig. 2. The absolute zero level of –9.98 contours was obtained by comparing the mean projected protein and lipid densities with values calculated from their known atomic compositions. The calculated protein/lipid electron-scattering density ratio is 1.30. Absolute integrated densities (column 4) were obtained by adding a constant amount to each helix. This was estimated to be 1/7 of the scattering expected from an area of lipid equal to the total projected area of the protein (753 Å²). The total area of the asymmetric unit (lipid plus protein) is 1102 Å².

protein has a density 1.4 times that of lipid and that the thickness of the protein and lipid regions are the same (13). The helix areas reflect the larger areas subtended by tilted helices. The relative densities given in Table 4 are corrected for the absence of the F(0,0) term in the Fourier map. This was estimated to be 9.98 contours (from the scale of contour interval of Fig. 2) by using the calculated electron-scattering density ratio of protein and lipid of 1.30. The densities for helices 1 and 5 are significantly weaker than those for all other features. This is also obvious by inspection of the map (Fig. 2). Helix 5 is smaller than helix 6 or 7 both in peak height and in diameter, even when it is assigned a maximum area. Helix 1 is by far the smallest in peak height but has a bigger area than helix 5. The lower peak heights of helices 2, 3, and 4 compared to helices 6 and 7 are compensated for by their larger areas due to the greater tilting of the helices when viewed in three-dimensions (Fig. 2 *Inset* and ref. 2).

Table 5 gives the calculated electron-scattering density of the atoms in each helix, from the theoretical scattering cross sections (14) for H, C, N, O, and S atoms. It is clear that helix D is significantly weaker than the others. This is not due to any arbitrary decision about the number of residues in each helix, but mainly to the intrinsic amino acid composition of the helices. Helix D contains five glycine residues and no large hydrophobic side chains at all, in striking contrast to many of the others, resulting in a low scattering power per amino acid. The points at which helix D begins and ends are also well defined, and it would be very difficult to make the helix much longer.

Table 5. Electron-scattering density in each helix from atomic composition

Helix	No. of amino acids	Total scattering power	Scattering power per amino acid
A	24	236.5	9.85
B	25	247.3	9.89
C	25	267.3	10.69
D	24	204.0	8.50
E	24	254.3	10.60
F	25	257.5	10.30
G	27	259.5	9.61

From the helical assignments from Fig. 1, the total scattering power of each helix was calculated. The weights given the different atoms were: C = N = O = 1, H = 0.25, and S = 1.5.

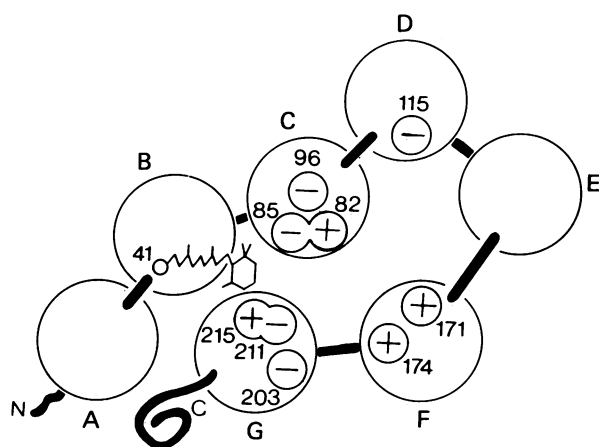


FIG. 3. The preferred model viewed from the inside of the cell. Retinal is attached to Lys-41 of helix B near the inside surface, but the orientation shown is unproven. The buried charges are drawn facing towards the interior of the protein. The thick connecting lines between helices indicate the path of the polypeptide.

To equal the average scattering of the other helices, D would have to be lengthened by five amino acids; helix D is the best candidate for a weakly scattering helix. Also lower than average, however, is helix A, which we made as long as possible in an effort to minimize the differences in the lengths of each helix, which might bias our comparison. The remaining five helices have significantly higher scattering power compared to these two, and none stands out as being the strongest.

We therefore apply a third criterion to our models. Helix D should not occupy positions with strong features in the map. Models 1, 17, and 27 in Table 3 all satisfy this criterion. In addition, models 36–41 are included because they have helix D in weak density in spite of a poorer connectivity or charge distribution. If helix A is also constrained to be in weak density, then models 17, 36, and 37 are eliminated. Thus, the most probable model, which has the shortest total link length of all 5040, is model 1, shown schematically in Fig. 3. The choice seems to be independent of the finer details of our criteria; model 1 always ranked best when we tried a wide variety of other weighting schemes for both connectivity and charge.

The preferred model

This model possesses a number of features we believe desirable for a membrane-embedded proton pump. The charged amino acids in helices C, D, F, and G line up along single surfaces that face each other towards the center of the molecule, possibly forming a hydrophilic proton channel, whereas the hydrophobic, uncharged faces of these helices face outward towards the lipid molecules. A tentative atomic model of this structure

suggests that many of the potential ion pairs are stereochemically possible and that they form one or more networks within the core of the protein through which protons could jump. The retinal is likely to lie near the surface in the region between helices B, C, and G and, therefore, the retinal ring may be close to the buried negative charge of Asp-96. Undesirable features, such as a crossing-over of two chains forming link regions between different helix pairs, also do not occur in model 1. Finally, the preferred model possesses a topology that would allow folding during biosynthesis into a globular molecule. All covalently connected helices are adjacent directly after synthesis, and no helix later intervenes between two previously existing helices so that no structural rearrangement would be necessary in the later stages of polypeptide synthesis.

Thus, we believe model 1 to be preferred on several grounds. On the other hand, all our arguments do not constitute proof, and it remains only the most probable of a number of possible models for the folding of the bacteriorhodopsin molecule in the membrane.

B.A.W. is a Fellow of the Jane Coffin Childs Memorial Fund for Medical Research. This investigation was aided by a grant from the Jane Coffin Childs Memorial Fund for Medical Research. D.M.E. is grateful to the John Simon Guggenheim Foundation for their support and to the National Science Foundation (PCM 78-10361) and the National Institutes of Health (GM 22778 and HL 14111).

1. Stoeckenius, W., Lozier, R. & Bogomolni, R. A. (1979) *Biochim. Biophys. Acta* **505**, 215–278.
2. Henderson, R. & Unwin, P. N. T. (1975) *Nature (London)* **257**, 28–32.
3. Michel, H., Oesterhelt, D. & Henderson, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 338–342.
4. Ovchinnikov, Y., Adbulaev, N., Feigira, M., Kiselev, A. & Lobanov, N. (1979) *FEBS Lett.* **100**, 219–224.
5. Gerber, G. E., Anderegg, R. J., Herlihy, W. C., Gray, C. P., Biemann, K. & Khorana, H. G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 227–231.
6. Walker, J. E., Carne, A. F. & Schmitt, H. (1979) *Nature (London)* **278**, 653–654.
7. Henderson, R., Jubb, J. S. & Whytock, S. (1978) *J. Mol. Biol.* **123**, 259–274.
8. Hayward, S. B., Grano, D. A., Glaeser, R. M. & Fisher, K. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4320–4324.
9. Gerber, G., Gray, C., Wildernauer, D. & Khorana, H. G. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5426–5430.
10. Bridgen, J. & Walker, I. D. (1976) *Biochemistry* **15**, 792–798.
11. King, G. I., Stoeckenius, W., Crespi, H. & Schoenbom, P. P. (1979) *J. Mol. Biol.* **130**, 395–404.
12. Unwin, P. N. T. & Henderson, R. (1975) *J. Mol. Biol.* **94**, 425–440.
13. Henderson, R. (1975) *J. Mol. Biol.* **93**, 125–141.
14. Lonsdale, K., ed. (1968) *International Tables for X-ray Crystallography* (Kynoch, Birmingham, England), Vol. 3, pp. 216–226.