J. Membrane Biol. 88, 233–247 (1985)

The Journal of
Membrane Biology

Stability of Transmembrane Regions in Bacteriorhodopsin Studied by Progressive Proteolysis

M.E. Dumont*, J. Trewhella**, D.M. Engelman, and F.M. Richards
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Summary. Proteinase K digestions of bacteriorhodopsin were carried out with the aim of characterizing the membrane-embedded regions of the protein. Products of digestions for two, eight or 24 hours were separated by high-pressure liquid chromotography. A computerized search procedure was used to compare the amino acid analyses of peptide-containing peaks with segments of the bacteriorhodopsin sequence. Molecular weight distributions of the products were determined by sodium dodecylsulfate-urea polyacrylamide gel electrophoresis. The structural integrity of the protein after digestion was monitored through the visible absorption spectrum, by X-ray diffraction of partially dried membranes, and by following release of biosynthetically-incorporated ^{3}H leucine from the digested membranes.

During mild proteolysis, bacteriorhodopsin was cleaved near the amino and carboxyl termini and at two internal regions previously identified as being accessible to the aqueous medium. Longer digestion resulted in cleavage at new sites. Under conditions where no fragments of bacteriorhodopsin larger than 9000 mol wt were observed, a significant proportion of the digested membranes retained diffraction patterns similar to those of native purple membranes. The harshest digestion conditions led to complete loss of the X-ray diffraction patterns and optical absorption and to release of half the hydrophobic segments of the protein from the membrane in the form of small soluble peptides. Upon cleavage of aqueous loop regions of the protein, isolated transmembrane segments may experience motion in a direction perpendicular to the plane of the membrane, allowing them access to protease.

Key Words bacteriorhodopsin · membrane proteins · proteolysis · purple membranes · proteinase K · transmembrane peptides

Introduction

Hydrophobic segments comprised predominantly of nonpolar amino acids have been found in the primary structures of many membrane proteins. In the absence of information about the three-dimensional folding of membrane proteins, the identification of hydrophobic sequences has been widely used as the basis for modeling the topologies, and in some cases the structures, of these proteins (e.g. *see* Argos, Rao & Hargrave, 1982; Engelman, Goldman & Steitz, 1982; Kyte & Doolittle, 1982). Such modeling usually assumes that hydrophobic sequences about 20 amino acids in length fold to form transmembrane alpha helices. The correspondence between nonpolar sequences and transmembrane helices has been based on theoretical considerations, such as the fact that alpha helical structures expose the minimum number of polar moieties in peptides to the surrounding medium (Von Heijne & Blomberg, 1979; Engelman & Steitz, 1981), on the spectroscopic studies showing a high percentage of alpha helix in the secondary structures of many membrane proteins (e.g., Moore, Holladay, Puett & Brady, 1974; Stubbs, Smith & Litman, 1976; Michel-Villaz, Saibil & Chabre, 1979; Mao & Wallace, 1984), and on the structure of bacteriorhodopsin as determined by Unwin and Henderson (1975).

Bacteriorhodopsin is a light-driven proton pump found in purple membranes from *Halobacterium halobium*. The fact that the protein forms a well-ordered two dimensional lattice in these membranes allowed the use of combined electron diffraction/electron microscope image processing to determine its structure (Henderson & Unwin, 1975; Unwin & Henderson, 1985; Michel, Oesterhelt & Henderson, 1980). At 7Å resolution, bacteriorhodopsin appears to be made up of seven rods of density, presumably alpha-helices, running back and forth across the membrane. The protein is predominantly alpha-helical based on spectroscopic (Huang et al., 1981; Mao & Wallace, 1984) and X-ray data (Blaurock, 1975; Henderson, 1975), but there remains a suggestion that some of the secondary structure in the membrane-embedded regions might

* *Present address:* Department of Biochemistry, University of Rochester Medical Center, Rochester, NY 14642.
** *Present address:* Life Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545.

be β-sheet (Jap, Maestre, Hayward & Glaeser, 1983), and it is difficult to entirely rule out the presence of some 3_{10} helix.

It would be useful to have a simple method for determining the topological distribution with respect to the membrane of a protein of known amino acid sequence. This has been attempted in the past by labeling proteins, either with membrane-impermeant reagents or with hydrophobic compounds. However, most of the reagents used in such studies, including photoactivatable compounds designed to be highly reactive and nonselective have, in fact, shown considerable specificity for some types of amino acid residues over others (Brunner & Richards, 1980; Ross, Radhakrishnan, Robson & Khorana, 1982). This makes them less useful than had been hoped for the general determination of protein topology. Another widely used approach to this problem has been that of modifying membrane proteins using enzymes which should not gain access to the membrane hydrophobic region. We report here the use of the relatively nonspecific enzyme proteinase K to study the topology of bacteriorhodopsin through exhaustive proteolysis. Aside from our interest in determining the details of the folding of bacteriorhodopsin in the membrane, this provided a useful test system for the enzymatic approach because of the availability of the sequence (Khorana et al., 1979; Ovchinnikov et al., 1979), and of topological information from previous experiments.

Of the various previous studies of the effects of proteases on bacteriorhodopsin, extensive proteolysis with a nonspecific enzyme was carried out by Gerber, Gray, Wildenauer and Khorana (1977), who reported 'complete degradation' of bacteriorhodopsin by proteinase K (as judged by SDS[1] polyacrylamide gel electrophoresis), and by Rosenheck et al. (1978), who found that pronase digested bacteriorhodopsin to a point where no bands could be detected on a 15% SDS polyacrylamide gel. In such experiments, it is of particular interest to determine the fate of the digested protein. The digestion products may correspond to transmembrane segments which are too small to show up in conventional SDS gel electrophoresis. In this case, if the fragments could be purified and characterized by some other technique, digestion could be used to obtain detailed information about the disposition of the polypeptide chain in the membrane. If, on the other hand, bacteriorhodopsin is hydrolyzed into peptides which are too small to span a membrane, this could call into question the usefulness of proteolytic enzymes for determining membrane topology, while, at the same time, providing information of the structural stability and dynamics of membrane-embedded regions of bacteriorhodopsin.

We have found that proteinase K digestion under mild conditions leaves bacteriorhodopsin in a form close to its native structure. However, the extensive proteolysis necessary to cleave the molecule into small (<5000 mol wt) peptides leads to a destabilization of the structure of the protein and the removal of up to half of its hydrophobic segments from the membrane.

Materials and Methods

Halobacterium halobium, strain S9, was grown on a defined medium as described previously (Engelman & Zaccai, 1980). Cultures were grown to stationary phase at 37° with controlled illumination and aeration. Purple membranes were isolated from washed cells by the method of Oesterhelt and Stoeckenius (1971). The isolated membranes were stored frozen in 45% sucrose. Prior to use, the membranes were thawed and the sucrose removed by repeated centrifugation, or by dialysis against three changes of 25 mM phosphate buffer or 25 mM TES, pH 6.9. Bacteriorhodopsin concentrations were determined from the absorbance at 568 nm (Rehorek & Heyn, 1979), with a correction for light scattering based on the optical density at 750 nm and assuming a dependence as the inverse square of wavelength (large particle scattering). Lipid phosphorus was measured by the method described by Dittmer and Wells (1969). Proteinase K was purchased from Boehringer Mannheim. Analytical or, where relevant, HPLC grade reagents were used throughout. Sequanol® grade trifluoracetic acid was obtained from Pierce Chemical.

Digestion Conditions

Washed purple membranes were diluted to 2.4 mg/ml in 25 mM phosphate buffer, pH 6.9. A freshly dissolved stock solution of proteinase K was then added so that the final bacteriorhodopsin concentration was 1.6 mg/ml and the final protease concentration was 0.8 mg/ml. This suspension was incubated at 37°C under room illumination and aliquots (usually several milligrams of bacteriorhodopsin) were removed after 2, 8 and 24 hours. A fresh solution of PMSF (0.1 M in ethanol) was added to each aliquot (final concentration 3 mM) immediately after incubation and left to stand 3 min at 37°C. The digested membranes were then washed once in the phosphate buffer, twice in 0.1 M NaCl and twice in H_2O by repeated centrifugation (45 min at 18,000 rpm in a Sorvall SS34 rotor: RCF_{max} = 35,000 × g). The pelleted membranes were lyophilized and stored at −20°C.

HPLC Conditions

Approximately 1 mg of the digested, lyophilized purple membrane was taken up in 20–30 μl of 90% formic acid. After 10 min with occasional vortex mixing, the still inhomogenous material

[1] *Abbreviations:* SDS, sodium dodecylsulfate; HPLC, high-pressure liquid chromatography; PMSF, phenylmethylsulphonylfluoride; TES, 2-[(tris-(hydroxymethyl)methyl)amino]ethanesulfonic acid.

was injected into the HPLC system consisting of a VYDAC® wide pore C^{18} column (25 cm × 10 mm i.d.) in series with a pellicular C^{18} guard column.

The columns were pre-equilibrated with 25% solvent B (0.1% trifluoracetic acid in 2 : 1 acetonitrile/2-propanol), 75% solvent A (0.1% trifluoracetic acid in water) at 37°C. Elution was controlled either by two Waters 510 pumps driven by a Waters model 680 gradient controller, or by two Altex 110A pumps and an Altex model 420 controller. Fractions containing absorbance peaks (at 210 nm) were dried by centrifugation under vacuum.

In order to obtain a smooth baseline, it was necessary to wash the column by running several conditioning gradient cycles between each separating gradient. Twenty microliters of formic acid were injected prior to each such conditioning gradient cycle.

Amino Acid Analysis

200 μl of 6 N HCl with 0.2% phenol was added to each of the dried fractions in tubes that were then sealed under vacuum (<25 mTorr) and maintained at 110°C for 20 hr. The hydrolysates were dried by centrifugation under vacuum and 60 μl of 0.2 N sodium citrate, pH 2.2, added to each in preparation for amino acid analysis, which was carried out on a Durrum D-500 analyzer.

Fitting of Amino Acid Compositions to the Bacteriorhodopsin Sequence

Amino acid analysis of peptides is a measurement of the relative abundances of amino acids of each type. The total number of amino acids of each type in a given fragment must be determined. A set of FORTRAN programs were therefore written to search for the best fit of a given amino acid analysis to any segment in a given protein sequence. The programs were run on a VAX 750.

The predicted number of amino acid residues of each type (C_i^{pred}) was calculated for each possible fragment of bacteriorhodopsin between 10 and 150 residues in length. The observed amino acid composition of the peptide to be fit (C_i^{obs}), expressed as the mole fractions of the total amino acid recovery present as each individual type of amino acid, was then compared with the predicted composition of each possible fragment. A residual (S_D) was computed as a measure of the quality of the fit.

$$S_D = \sum_{\substack{\text{all amino} \\ \text{acid types}}} [C_i^{\text{obs}} - (C_i^{\text{pred}}/N)]$$

where N is the numbers of amino acids in the fragment.

Because the residual tended to be larger for fits to smaller fragments, and because no initial assumptions were made about the expected lengths of peptides submitted to this analysis, the searches were carried out separately for different fragment lengths. The programs sought segments of the sequence giving best fits in length ranges 10–20, 20–50, and 50–150 residues. In cases where no portion of the sequence could be clearly identified from the composition, but good values of S_D were obtained for the fragment length range 50–150 residues, fits to even longer fragments were examined (up to 200 residues). Once the program identified the class of peptides with the lowest S_D, final selection of the best-fitting fragments was carried out by inspection of the actual expected and calculated compositions, the principal criterion being that a good uniformity of agreement should exist among all the amino acid types, allowing for systematic errors in analysis for certain types as described in Results.

Digestion of ^{3}H-Leucine Labeled Bacteriorhodopsin

Purple membranes labeled as described in Engelman and Zaccai (1980) were digested with proteinase K as above. Samples contained about 7×10^5 dpm per mg bacteriorhodopsin. During each of the washing steps, aliquots of the pellet and supernatant fractions were diluted in Aquasol® (New England Nuclear) and counted in a Packard 3255 scintillation counter. The lyophilized digested samples were subjected to HPLC as above. 0.7 ml-fractions were collected, of which 0.07 ml of each was removed and pooled for the determination of the total activity eluted during the run. The radioactivity in the remainder of each fraction was then determined.

The supernatant fraction from the first wash after each digestion was subjected to prolonged ultracentrifugation (6 hr) at 40,000 rpm in a Beckman type 70.1 Ti Rotor (150,000 RCF_{max}) in order to pellet broken membrane patches. This resulted in small, but visible pellets. Both the supernatant and pellet fractions were counted. All radioactivity determinations were corrected for quenching by comparing the ratio of counts in two energy channels with that seen in a standard quench curve.

In a separate series of digestions, ^{3}H leucine-labeled purple membranes were incubated with proteinase K as above, except that the buffer was 25 mM TES, pH 6.9, instead of phosphate. Following inhibition with PMSF, the membranes were subjected to centrifugation at 190,000 RCF_{max} for 2 hr at 4°C. An aliquot (0.8 ml.) of the supernatant from this spin was immediately loaded on a Biogel P2 size-exclusion column (Biorad Laboratories, −400 mesh, 0.9 × 25 cm) that had been equilibrated with the TES buffer. Fractions (0.5 ml) were diluted in Aquasol® and counted as above.

X-Ray Diffraction of Native and Digested Bacteriorhodopsin

Digested or native membranes were washed as described above, then pelleted once more in 15 mM NaCl, 3 mM sodium phosphate, pH 6.9, removing as much of the supernatant moisture as possible. Pellets were stirred, then transferred in small drops to thin acetate sheets and allowed to dehydrate under an atmosphere of 76% relative humidity over saturated NaCl. The plastic was mounted in the focussed beam of a Baird and Tatlock Searle toroid camera, and exposed by copper K_α X-rays from an Elliott GX-6 rotating anode source. The camera was purged with humidified helium for 30 min prior to recording each diffraction pattern for either one or 2 hr on Kodak DEF-5 X-ray film.

Measurement of relative diffraction intensities was carried out by digitizing the films on an Optronics International P-1000 microdensitometer, at a 25 μm raster. Circular averaging of the patterns was carried out over 90° sectors using a program written by Malcolm Capel. For each reflection of interest, a correction for background density was calculated by using a least-square routine to fit a parabolic curve to the intensity minima between the reflection and its neighbors. Integrated intensities from the four sectors were then averaged.

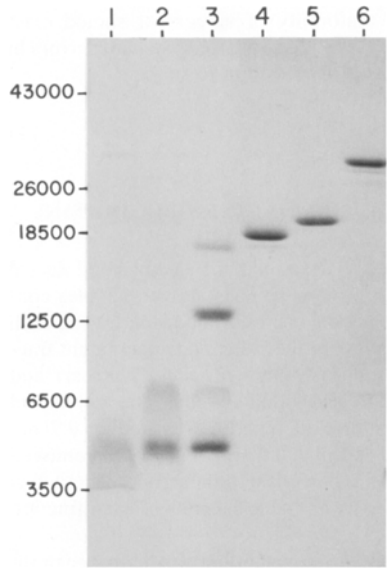

Fig. 1. SDS-urea polyacrylamide gradient gel electrophoresis of the products of proteinase K digestion of bacteriorhodopsin. Samples, either lyophilized or stored frozen after being washed free of proteinase K (*see* Materials and Methods) were made up to be 10% glycerol, 2% SDS, 0.063 M Tris/HCl at pH 6.8. β-mercaptoethanol was added to 0.5% final concentration just before placing the samples in a boiling water bath for 10 min. The system for polyacrylamide gel electrophoresis was as described by Laemmli (1970) except that the separating gel contained 6 M urea and was a gradient, 15–19% acrylamide, 0–10% sucrose, poured from the bottom up. Gels were stained with Coomassie Blue. *Lane 1:* Purple membranes digested for 24 hr. *Lane 2:* Purple membranes digested for 8 hr. *Lane 3:* Purple membranes digested for 2 hr. *Lane 4:* Purple membranes incubated with pre-inhibited proteinase K. The protease was incubated for 3 min at 37°C in a solution to which 0.1 M PMSF in ethanol had been added to a final concentration of 9 mM. The proteinase K was then immediately diluted with purple membranes and washed as for actual digestions. *Lane 5:* Undigested purple membranes. *Lane 6:* Proteinase K in the absence of purple membranes (PMSF inhibited)

Results

I. Extent of Digestion Monitored by SDS-Urea Gel Electrophoresis

To characterize the products of proteinase K digestion of bacteriorhodopsin, it was necessary to use polyacrylamide gradient gel electrophoresis in the presence of 6 M urea. Small peptides were resolved better on this system than on conventional Laemmli (1970) gels or on nongradient gels containing urea (Swank & Munkres, 1971). However, the logarithmic relationship between mobility and molecular weight is subject to considerable uncertainty for such small peptides. Also, intact bacteriorhodopsin exhibits an anomalous mobility on such gels, migrating with an apparent mol wt of about 21,000.

After two hours of digestion, peptides running at about 17,000, 13,000, 7,000 and 5,000 apparent mol wt were the main products, with no detectable intact bacteriorhodopsin and little proteinase K remaining in the washed samples (Fig. 1, lane 3). After longer digestion (8 hr) the 7,000 and 5,000 mol wt bands were the only ones still visible (lane 2). 24 hours of digestion resulted in destruction of most of the 7,000 mol wt band, broadening of the 5,000 mol wt band, and the appearance of material comigrating with what appeared to be an ion front at 3,500–4,000 mol wt.

Figure 1 (lane 4) shows the PMSF treatment of the membranes inhibits most, but not all, protease activity. The lack of complete inhibition may be due to partitioning of PMSF into the membrane patches. A short fragment, presumably the carboxyl terminus of bacteriorhodopsin, is removed by treatment with pre-inhibited proteinase K. However, the combination of inhibitor treatment with extensive washing leads to removal of all detectable protease activity, since samples stored after these procedures show little change in their electrophoretic patterns with time.

II. HPLC Separation of Bacteriorhodopsin Fragments and Identification of Peptides from Amino Acid Analyses

While HPLC has been used by other groups to purify bacteriorhodopsin peptides (Gerber et al., 1979: Lemke, Bergmeyer & Oesterhelt, 1982), the fragments left after proteinase K digestion in the membrane should be more hydrophobic than general cyanogen bromide or subtilisin peptides generated by cleaving delipidated material. We found that the acetonitrile/2-propanol (2 : 1) mix used by Tarr and Crabb (1983) gave sharper peaks than either of these solvents alone. The use of a completely volatile mobile phase was dictated by the low recovery expected in any additional steps which might be required to remove buffer from such small hydrophobic peptides. Chromatograms of lyophilized digested material resolubilized in 90% formic acid are shown in Fig. 2.

In attempting to recognize parts of the bR sequence from the amino acid compositions of the peptides in the most objective way, the computer program described in Materials and Methods was developed. By testing a wide range of possible fragment lengths, this program reduced the importance of subjective judgements in converting ratios of amino acid abundances from the analyses into ac-

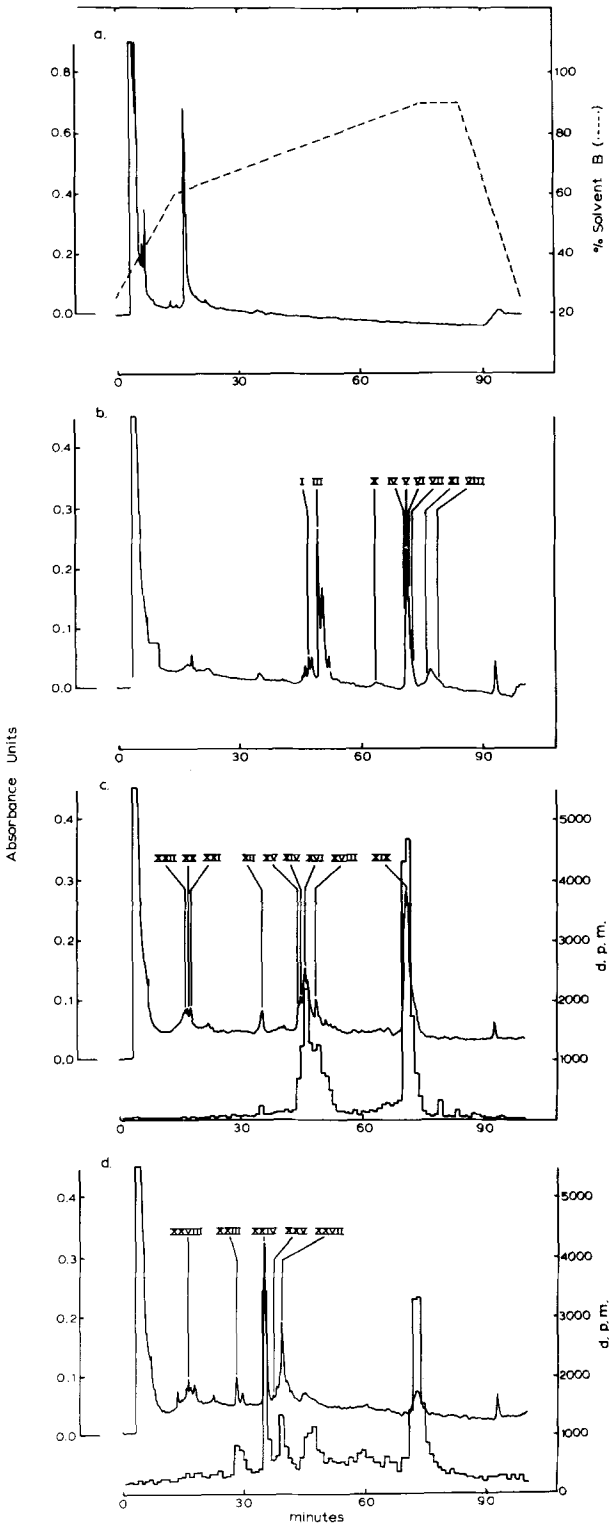

Fig. 2. Examples of HPLC chromatograms of proteolyzed purple membranes and self-digested proteinase K. Roman numerals identify the approximate elution positions of some of the fragments listed in Table 1. Most of the identified fragments were purified from chromatograms other than the particular ones shown here. Because elutions were carried out on different HPLC systems with slightly different gradient characteristics, the identification of these peaks is based on comparison of recognizable elements in the elution profiles, rather than absolute elution times. (*a*) Proteinase K that had been incubated in the absence of membranes at 37°C for 24 hr, then lyophilized. The dashed line shows the imposed gradient profile for this and following chromatograms. (*b*) Purple membranes digested for 2 hr with proteinase K, then washed and lyophilized as described in Materials and Methods. (*c*) Purple membranes digested for 8 hr with proteinase K. The histogram shows the radioactivity contained in each fraction (refer to right axis). (*d*) Purple membranes digested for 24 hr with proteinase K, with histogram showing the radioactivity in each fraction. Elution of fragments not identified in this figure: Fragment II eluted between I and III; fragment IX eluted between III and X; fragment XIII eluted at about the same time as XII; fragment XVII eluted at about the same time as XV; and fragment XXVI eluted at about the same time as XXVII. The peaks eluting at 47 and 73 min in *d* could not be identified as comprising a single peptide fragment. Based on a comparison of their elution times with those of fragments seen in milder digests, and on their amino acid compositions, they appear to be mixtures of peptides corresponding to remnants of fragments XIV and XIX, respectively, which were not completely digested in this run

tual numbers of residues of each type in a given fragment. The peaks in Fig. 2 that could be identified from their amino acid compositions are labeled.

Although most of the major peaks from each HPLC run could be identified from their amino acid compositions, some could not. The failure to identify peaks probably results in part from peak overlap—either two fragments running with the same mobility, or contamination eluting as a broad peak, contributing to analyses of sharp peaks. The latter situation is suggested by the rather high baseline levels of radioactivity in the chromatogram of the 24-hr-digested radioactive membranes (Fig. 2*d*).

In addition, the analyses from some peaks had high histidine levels which, since bacteriorhodopsin lacks histidine, suggests that these peaks contained proteinase K or fragments thereof which had stuck to the purple membranes during the washes. All peaks containing significant histidine were excluded from subsequent consideration.

As shown in Fig. 2*a*, all the peptides from self-digested proteinase K eluted early in HPLC chromatograms. Thus, early elution, coupled with low levels of radioactivity in digestions of tritium-labeled membranes, was taken as an indication that certain peptides might be fragments of proteinase K. Such fragments of uncertain origin are marked (+) in Table 1.

The most clearly identifiable HPLC peaks from a number of chromatographic runs of material digested for 2, 8, and 24 hr are shown in Fig. 3. A comparison of the expected and actual amino acid compositions for these peaks is shown in Table 1. Fits with high values for the measured glycine and

Table 1. Amino acid compositions of HPLC peaks[a]

Sample	Peptide	nM	ASP	THR	SER	GLU	PRO	GLY	ALA	VAL	MET	ILE	LEU	TYR	PHE	HIS	LYS	ARG
(2 hr digestion)																		
I	5,6–66,67	0.8	2.2(2)	5.8(6)	2.0(2)	1.2(1)	3.0(3)	8.1(8)	6.0(6)	3.1(3)	3.6(4)	3.0(3)	11.0(11)	3.8(4)	3.1(3)	0.0(0)	2.9(3)	1.3(1)
II	7–74	0.5	2.5(2)	5.9(6)	2.5(2)	2.0(2)	3.6(4)	9.3(9)	6.2(6)	4.0(4)	3.8(5)	3.5(3)	10.9(11)	3.7(4)	3.8(4)	0.2(0)	3.0(3)	1.3(1)
III	6,7–71,72	1.2	2.1(2)	6.1(6)	2.1(2)	1.2(1)	3.9(4)	8.0(8)	6.0(6)	4.1(4)	4.4(5)	3.4(3)	10.9(11)	3.8(4)	4.0(4)	0.0(0)	2.9(3)	1.1(1)
IV	72–161	1.1	6.2(6)	6.9(7)	3.0(3)	4.2(4)	2.2(2)	9.1(9)	11.3(11)	6.0(6)	1.8(2)	5.8(6)	14.6(15)	5.7(6)	5.1(5)	0.0(0)	2.1(2)	2.1(2)
V	72–160	0.7	5.9(6)	7.2(7)	3.1(3)	3.3(3)	2.4(2)	8.7(8)	10.9(11)	5.8(6)	1.7(2)	5.8(6)	14.7(15)	5.9(6)	5.2(5)	0.1(0)	2.1(2)	2.1(2)
VI	71–161*	0.6	6.4(6)	6.6(7)	3.0(3)	3.8(4)	1.9(2)	9.2(9)	11.0(11)	6.2(6)	1.8(2)	6.4(6)	15.0(15)	6.3(6)	5.6(6)*	0.0(0)	1.9(2)	1.9(2)
VII	72–164*	0.3	6.3(6)	7.1(7)	4.0(4)	4.4(4)*	2.4(2)*	9.2(9)	10.4(10)	6.6(6)*	1.7(3)*	5.8(6)	14.9(15)	5.9(6)	5.2(5)	0.2(0)	2.3(2)	2.6(3)*
VIII	72–234*	0.2	9.7(9)	10.3(10)	8.6(9)	8.6(9)*	5.3(5)	16.2(15)	17.1(17)*	15.7(17)	3.7(4)	10.3(11)	25.4(25)	6.8(7)	9.0(9)	0.2(0)	4.3(4)	6.1(6)
IX	170–233*	0.4	3.4(3)	3.3(3)	4.3(4)*	2.8(3)	2.3(2)	6.0(6)	5.0(5)	9.1(10)*	1.1(1)	4.3(5)	10.1(10)	1.4(1)	3.8(4)	0.1(0)	2.1(2)	3.0(3)
X	165–231*	0.2	3.3(3)	3.3(3)	5.0(5)	3.0(3)	3.0(3)	6.1(6)	5.5(5)	8.9(11)	1.1(1)	4.4(5)	10.5(10)	1.4(1)	3.9(4)	0.1(0)	2.0(2)	3.4(3)*
XI	6–162	0.1	9.0(9)	13.1(13)	6.8(6)	5.4(5)	5.8(6)	17.3(17)	17.3(17)	10.0(10)	4.7(7)	9.3(9)	26.4(24)	9.2(10)	8.6(9)	0.3(0)	4.8(5)	3.5(3)
(8 hr digestion)																		
XII	6,7–32,33*	0.9	0.2(0)	2.0(2)	0.4(0)*	1.0(1)	0.9(1)	4.6(5)	2.1(2)	1.3(1)*	1.4(2)	1.0(1)	5.6(6)	1.2(1)	1.3(1)	0.0(0)	0.9(1)	1.2(1)
XIII	6,7–32,33	2.0	0.2(0)	2.0(2)	0.3(0)	1.0(1)	1.0(1)	4.6(5)	2.1(2)	1.1(1)	1.7(2)	0.9(1)	5.7(6)	1.2(1)	1.3(1)	0.0(0)	0.9(1)	1.0(1)
XIV	5,6–66,67	0.7	2.0(2)	6.0(6)	2.0(2)	1.1(1)	2.9(3)	8.1(8)	5.9(6)	3.2(3)	3.8(4)	2.9(3)	11.0(11)	3.9(4)	3.0(3)	0.0(0)	3.0(3)	1.2(1)
XV	5,6–66,67*	0.6	2.1(2)	6.4(6)*	2.2(2)	1.2(1)	2.9(3)	7.9(8)	5.8(6)	3.1(3)	3.7(4)	3.2(3)	10.6(11)*	3.8(4)	2.9(3)	0.0(0)	2.8(3)	1.4(1)
XVI	5–65*	2.4	2.1(2)	6.2(6)*	2.0(2)	1.2(1)	2.9(3)	7.8(8)	5.7(6)	3.0(3)	4.3(4)*	3.3(3)	10.0(10)*	3.7(4)	2.9(3)	0.0(0)	2.9(3)	1.1(1)
XVII	6,7–64,65	0.4	2.3(2)	5.2(5)	2.4(2)	1.3(1)	2.9(3)	7.3(7)	5.8(6)	2.9(3)	3.0(4)	3.2(3)	10.1(10)	3.7(4)	3.1(3)	0.1(0)	2.7(3)	1.1(1)
XVIII	5–71*	0.9	2.2(2)	6.9(7)	2.1(2)	1.4(1)	3.8(4)	8.4(8)*	6.0(6)	4.0(4)	4.4(5)	3.3(3)	11.0(11)	3.8(4)	3.7(4)	0.1(0)	2.9(3)	1.2(1)
XIX	73–159*	1.2	6.0(6)	7.0(7)	3.2(3)	3.3(3)	2.2(2)	8.1(8)	10.3(10)*	6.0(6)	1.9(2)	5.6(6)	14.2(15)	5.8(6)	5.1(5)	0.1(0)	2.1(2)	2.1(2)
XX	168–197†	0.2	1.5(1)	1.7(2)	3.4(3)	1.4(1)	1.2(1)	3.2(3)	3.8(3)	3.6(6)	0.2(0)	1.2(1)	2.9(3)	1.0(1)	0.8(1)	0.2(0)	1.1(1)	1.1(1)
XXI	168–194*†	0.2	1.2(1)	1.8(2)	2.7(3)*	1.1(1)	1.2(1)	1.4(1)*	2.3(2)*	5.0(6)	0.1(0)	1.2(1)	3.2(3)	0.9(1)	0.8(1)	0.0(0)	1.0(1)	1.1(1)
XXII	174–194*†	0.3	1.1(1)	1.1(1)	2.1(2)	1.0(1)	1.2(1)	1.2(1)	1.3(1)	4.2(5)*	0.0(0)	1.0(1)	2.6(3)*	1.0(1)	0.2(0)	0.1(0)	0.5(0)*	0.6(1)*
(24 hr digestion)																		
XXIII	6,7–30,31	0.9	0.2(0)	1.9(2)	0.4(0)	1.2(1)	1.1(1)	4.0(4)	2.0(2)	1.1(1)	0.8(1)	0.7(1)	5.6(6)	1.0(1)	1.0(1)	0.1(0)	1.0(1)	1.1(1)
XXIV	6,7–32,33	3.4	0.0(0)	2.1(2)	0.0(0)	1.1(1)	1.1(1)	5.1(5)	2.1(2)	1.0(1)	1.9(2)	0.7(1)	5.8(6)	1.0(1)	1.0(1)	0.0(0)	1.0(1)	1.0(1)
XXV	6–37*	0.2	0.8(1)	2.1(2)	0.8(1)	1.6(1)	1.8(2)	5.6(6)*	2.4(2)	1.6(2)	1.3(2)	1.0(1)	5.9(6)	1.1(1)	1.3(1)	0.2(0)	1.1(1)	1.4(1)
XXVI	134–158*	0.3	0.5(0)	1.6(2)	1.9(2)	0.4(0)	0.2(0)	1.9(1)	2.7(3)	1.7(2)	0.6(1)	1.6(2)	2.8(3)	2.3(2)*	3.4(4)	0.1(0)	0.5(0)*	1.1(1)
XXVII	134–158*	1.1	0.2(0)	1.7(2)	1.8(2)	0.2(0)	0.2(0)	1.5(1)	2.9(3)	1.7(2)	0.8(1)	1.7(2)	3.0(3)	2.5(2)*	3.7(4)	0.0(0)	0.2(0)	1.0(1)
XXVIII	175–187*†	0.9	1.0(1)	0.8(1)	0.8(1)	0.3(0)	0.7(1)	0.6(0)	1.0(1)	2.9(4)*	0.1(1)	0.3(0)	1.1(1)	0.7(1)	0.1(0)	0.0(0)	0.3(0)	1.3(1)

[a] Expected amino acid compositions are shown in parentheses. Fragments shown having more than one set of end points refer to segments of the sequence for which different starting points lead to identical calculated amino acid compositions. Many fragments have additional ambiguity in their exact end points stemming from uncertainties in matching amino acid analyses to the sequence. Such fragments are indicated by asterisks (*), as are the particular amino acid residues for which the fit would be affected by the uncertainty. Fragments selected to be shown in this table were each the best-fitting of a class of peptides obtained over a series of digestions and subsequent HPLC analyses. Those identified by dagger (†) gave good agreement with portions of the bacteriorhodopsin sequence but cannot be definitely identified as fragments of bacteriorhodopsin, since they lack significant radioactivity in digestions of labeled purple membranes and elute at approximately the same times as proteinase K fragments in HPLC chromatograms.

Sequences most closely matching the amino acid compositions were identified using the computer program described in Materials and Methods. Measured amino acid compositions were calculated by multiplying the abundances of each type of amino acid by the ratio of the number of amino acids in the fragment to be fit (subtracting tryptophans, which do not show up to the analysis) to the sum of the measured abundances of all types of amino acids (except histidine, which is absent from the sequence).

serine contents were accepted, since these are common contaminants in the amino acid determinations. Methionine and valine tended to be low in all determinations because of oxidation during hydrolysis and incomplete hydrolysis, respectively. In particular, the valine-valine bonds at residues 179–180 and 187–188 tended to make fragments containing these residues reproducibly low by two valines in measured composition.

All the fragments shown in Table 1 were clearly preferred by the search program. The most probable fits identified by the program included the peptides listed in the table and closely correlated ones which differed by only a few amino acids on either end. In lists of different peptides ordered by their residuals, S_D, the first listed peptide from an entirely different part of the sequence had a much higher value of the residual than the fragment shown. For some fragments, the uncertainty in the exact starting and ending points could not be resolved, as indicated in Table 1 and Fig. 3. Such ambiguity may result from co-elution of similar fragments, from inexactness of the amino acid analyses, or from characteristics of the sequence. For several fragments, shifting the starting point by one residue does not alter the amino acid composition at all. In some cases, closely related, but not identical, peptides were identified eluting at similar times in separate HPLC runs. Such differences presumably stem from slight differences in fraction collecting or the detailed elution profile of independent separations.

III. Digestion and HPLC Separations of ^{3}H-Labeled Membranes

Digestion of ^{3}H-leucine labeled membranes was carried out to follow the fate of the bacteriorhodopsin during the initial incubation and subsequent treat-

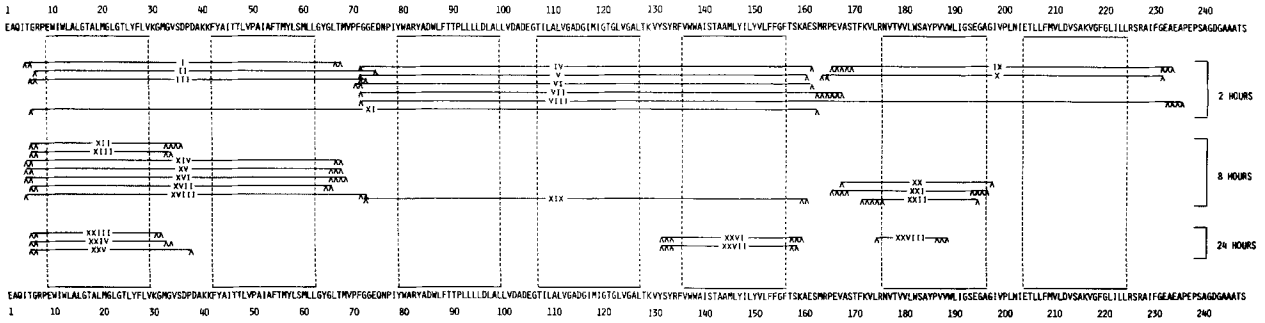

Fig. 3. Map of locations of best-fit fragments in the bacteriorhodopsin sequence (using one letter code). Roman numerals identify the fragments with respect to Table 1 and Fig. 2. Dashed boxes delimit the transmembrane segments predicted by Engelman et al. (1982). Where there is uncertainty about the exact end points of a fragment, arrowheads indicate possible cleavage sites. Fragments XX–XXII and XXVIII may be peptides from proteinase K, although they fit well to portions of the bacteriorhodopsin sequence (*see* text). Because of low recovery of peptide during HPLC and failure to identify some peaks, the peptides shown here do not necessarily represent the complete set of products of proteolysis

Table 2. ^{3}H leucine distribution following proteolysis[a]

Sample	Radioactivity in 1st supernatant (% of initial membrane suspension)	Radioactivity in 2nd supernatant (% of initial membrane suspension)	Radioactivity in ultracentrifuge pellet from 1st supernatant (% of 1st supernatant)
Purple membranes	2	—	44
Purple membranes + inhibited proteinase K[b]	4	2	43
Purple membranes + proteinase K (1 min)	5	1	43
Purple membranes + proteinase K (2 hr)	12	1	13
Purple membranes + proteinase K (8 hr)	20	1	6
Purple membranes + proteinase K (24 hr)	50	1	2

[a] Digestions of ^{3}H-labeled membranes and subsequent washings were carried out as described in Materials and Methods. Radioactivities of each sample were determined by resuspending pellets after centrifugation and counting this suspension. The radioactivity of each pellet is expressed here as the percent of the tritium in the total sample before centrifugation, determined by counting aliquots of the initial suspension. In the case of samples subjected to ultracentrifugation, the radioactivity in each pellet is expressed as a percent of the total radioactivity in the first supernatant from the low speed washes.

[b] Proteinase K was pre-inhibited with PMSF as described for Fig. 1.

ments, and to allow discrimination between bacteriorhodopsin and proteinase K peaks in HPLC runs. Table 2 gives the distribution of radioactivity during washing steps after digestion. No more than 1–2% of the total radioactivity in a sample was released to the supernatant in the absence of digestion, or in the second and subsequent washes after digestion. The slightly higher release of tritium to the supernatant in samples where proteinase K was present, either after inhibition or for a very brief time (4–5% *vs.* 1–2%), probably resulted from a loosening of the pellet on digestion of the carboxyl terminal tail of the molecule. Much of the radioactivity in the supernatants of the first washes of these samples was in the form of membrane patches too small to pellet under the conditions of the wash, since 45% of these supernatant counts could be pelleted at high speed in the ultracentrifuge.

At longer digestion times, significantly higher levels of radioactivity were released in a form with a small sedimentation coefficient (Table 2). By the time the digestion had proceeded for 24 hr, fully half the radioactivity was found in the first supernatant, yet only 2% of the radioactivity in this supernatant

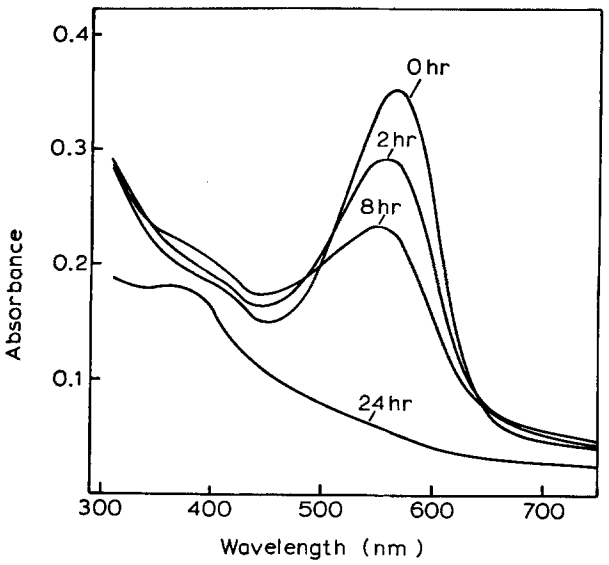

Fig. 4. Absorption spectra of proteolyzed purple membranes. Because PMSF addition in ethanol appeared to affect membrane color, spectra were obtained before inhibition and washing of samples by simply diluting the digestion mixture by a factor of 16.7 at the appropriate time. The spectrum of the 0 time sample was obtained within 5 min of adding proteinase K to the membranes

could be pelleted under more stringent conditions in the ultracentrifuge.

When the supernatant fraction from the first centrifugation following 2,8, or 24-hr digestion was subjected to gel chromatography on Biogel P2 (nominal fractionation range 100–1800 mol wt), all the radioactivity was at least partially included in the internal volume of the gel, with the major peak of radioactivity eluting at the approximate volume expected for a pentapeptide.

The inhomogenous nature of the formic acid suspension of the lyophilized membranes and the small volumes involved made it difficult to exactly quantitate the yield of radioactive peptides eluted during HPLC, relative to the number of counts injected onto the column. By comparing the radioactivity in small aliquots of the injected mixture with that of pooled aliquots of all the fractions collected during a run, we found this yield to be on the order of 50% for the eight and 24-hr digestions. Such low yields have been seen in HPLC of other hydrophobic peptides (Tarr & Crabb, 1983). Recovery of peptides during elution of the column must not, however, be a simple function of the hydrophobicity of the peptides, since we were able to elute with high yield some of the most hydrophobic segments of the molecule according to the hydrophobicity calculations of Engelman and Steitz (1984).

As much as 20% of the radioactivity in the injected sample could be eluted in the first gradient cycle after the actual run (and following the injection of 20 μl of 90% formic acid). Because the pattern of peaks eluted in the gradient cycles used to clean the column was roughly similar to that in the initial run, and because the bulk of the radioactivity could be eluted in a series of repeated runs, it seems likely that the low yield of radioactivity in the initial run reflected partial retention of the same peptides recovered in that run rather than retention of a wholly different set of peptides. The basis for the strong retention of peptides by the column is not understood, but it is possible that some of the loaded material was in an insoluble form, which bound strongly to the column and could only be solubilized and eluted by repeated injections of formic acid during conditioning gradients subsequent to the initial separation.

The profile of radioactivity in the fractions eluted in HPLC generally follows that of the absorbance at 210 nm, allowing for variation in the expected extinction coefficients of different peptides and for the different leucine compositions. This confirms that most of the peaks in the profile do, in fact, correspond to fragments of bacteriorhodopsin, rather than proteinase K that might have been carried along through the washes. However, some absorbance peaks, in particular those at 16 min in Figs. 2*c* and *d*, contain very little radioactivity in comparison to the absorbance. The measured leucine contents of these fragments are low and the expected radioactivity would be close to the baseline if these peptides did, in fact, come from bacteriorhodopsin. In addition, they contain little histidine. Thus, no definite decision can be made on the origin of these particular peaks.

IV. Optical Absorption Spectra of Digested Membranes

Absorption spectra of digested purple membranes are shown in Fig. 4. The 568-nm absorption peak of native, light-adapted bacteriorhodopsin decreases after the 2 and 8-hr digestions to 74 and 52%, respectively, of the native membrane absorbance and is completely lacking (~3% of native) in material digested for 24 hr. At the same time, the absorption peak shifts from 568 nm in the native membranes, to 560 and then to 550 nm in the 2 and 8-hr digestions. The loss of this absorption peaks coincides with the appearance of another peak, of much lower extinction coefficient, at 360 nm.

V. Lipid Determinations of Digested Membranes

Lipid phosphorus was followed as a function of digestion time by carrying out proteolysis in TES instead of phosphate. After digestion, all samples were inhibited with PMSF and pelleted at 190,000 RCF_{max} for 2 hr. The total phosphorus in the supernatant fraction from this spin never exceeded 1% of that in the pellet.

VI. X-Ray Diffraction of Digested Membranes

Because bacteriorhodopsin forms a two-dimensional crystal in the plane of the membrane, it is possible to obtain X-ray diffraction patterns of partially dried films of membranes which show rings characteristic of a circularly averaged P3 lattice (Blaurock, 1975; Henderson, 1975). We have used such diffraction patterns to monitor the organizational state of the digested material. From the diffraction patterns shown in Fig. 5, it is apparent that a significant fraction of the bacteriorhodopsin molecules are still arranged in a lattice after 2 or 8 hr digestion, but that by 24 hr there is little intensity in other than the lowest order reflections.

It is difficult to relate total diffraction intensity directly to the proportion of bacteriorhodopsin molecules which are organized as a lattice. Since the diffraction was carried out on partially dried drops of membranes, the thickness of each drop, the fraction of the drop exposed to the collimated beam, and the flatness with which the drop spreads on the acetate all affected the intensity of diffraction. We obtained three series of diffraction patterns from membranes digested for 0, 2, 8, and 24 hr. The integrated intensities of three of the strongest maxima in each pattern (the 2,1 reflection, the overlapping 2,2 and 3,1 reflections, and the overlapping 4,3 and 5,2 reflections) were compared within each series. The intensity of these maxima in the samples digested for 2 hr dropped to about 70% of the values for membranes treated with pre-inhibited proteinase K, and the maxima in the 8-hr digestions dropped to about 50% of these control values. Thus, bearing in mind the uncertainties in comparing diffraction patterns, the measurements indicate that roughly two thirds of the lattice is left after 2 hr digestion, and one half after 8 hr.

Relative intensities of the reflections measured were remarkably constant over the 0–8 hr digestions (*see* Fig. 5), given the amount of cleavage which had taken place. No changes greater than the ~15% uncertainty in the intensity measurements were detected. The lattice spacing decreased by 2% during the 8 hr digestion, a change of questionable significance compared with the observed 1% variation in the lattice spacing of undigested samples.

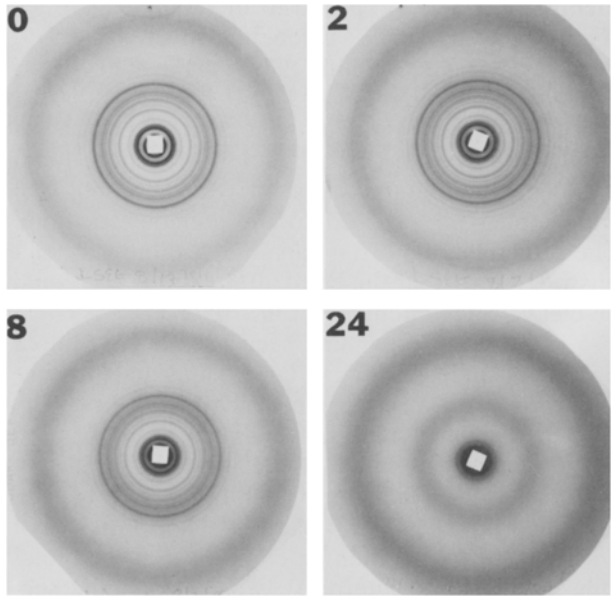

Fig. 5. X-ray diffraction patterns of proteolyzed purple membranes. Membranes digested for 0 time were exposed only to pre-inhibited proteinase K as described for Fig. 1. This sample and those digested for 2, 8, and 24 hr were inhibited and washed as described in Materials and Methods. Films were exposed for 2 hr. For purposes of comparison, all four films were re-photographed and printed as a group

Discussion

Extensive proteinase K digestion reduces bacteriorhodopsin to small fragments. In order to examine the process by which the proteolysis proceeds, we have used progressively increasing digestion times, characterizing the fragments produced, the state of the lattice, the fraction of labeled peptide pelleting with the membranes, and the optical absorption spectrum at each time.

Two Hour Digestion

After a 2-hr incubation with proteinase K, the general regions of cleavage which we have been able to identify occur at the NH_2 (residues 4–7) or the COOH (residues 231–236) ends of the molecule, or at one of two internal segments (66–75 and 160–170). All of these regions have previously been

identified as being exposed to the aqueous surface (Gerber et al., 1977; Ovchinnikov et al., 1979; Walker, Carne & Schmitt, 1979: Huang et al., 1981; Wallace & Henderson, 1982: Liao & Khorana, 1984). However, in one of these regions, proteinase K digestion, even for this relatively short time, identifies a minimum range of residues (160–166) which must be exposed to the aqueous medium. The previously reported papain digestion had demonstrated only that the bond after residue 162 could be cleaved (Ovchinnikov et al., 1979).

The largest fragment seen on the SDS-urea polyacrylamide gel (Fig. 1, lane 3) appears to correspond to bacteriorhodopsin which has lost only its NH_2 and COOH terminal ends (bearing in mind the anomalous mobility of the intact molecule on these gels) even though no such fragment has been identified among HPLC fractions. The 13,000 mol wt fragment would then be comprised of the molecules lacking the outermost two hydrophobic segments at one or the other end (fragments VIII or XI). Loss of all four of these segments probably results in the band seen at 7,000 mol wt (fragments IV–VII), while the paired outermost hydrophobic segments make up the smallest fragment at ~5000 mol wt (fragments I, II, III, IX, and X).

All the bacteriorhodopsin molecules subjected to 2 hr of digestion were cleaved at not less than two points in their peptide backbones. At the same time, the persistence of the optical absorption (which is quite sensitive to treatments such as detergent solubilization or alteration of the hydration state of the membrane) and the preservation of the diffraction pattern from these membranes demonstrate that they are made up of molecules with conformations resembling the native structure. This finding extends the results of Wallace and Henderson (1982) and Popot, Trewhella, Gerchman and Engelman (1984) to harsher digestion conditions.

Twelve percent of the ^3H leucine (corresponding to about four residues) is removed from the membrane after 2 hr proteolysis. This is consistent with either the destabilization of as much as one helical segment from a major fraction of the bacteriorhodopsin molecules or removal of a small fraction of the population of each of several membrane-embedded regions. In addition, some part of the solubilized counts may represent amino acids other than leucine that have been labeled through metabolic conversion during growth in the ^3H medium. The gas chromatography-mass spectrometry analysis described by Trewhella, Anderson, Fox, Gogol, Khan and Engelman (1983) showed that about 10% of the total isotopic label incorporated into purple membranes grown under similar conditions on deuterated leucine was found in amino acids other than leucine, with most of the converted label found in other hydrophobic amino acids.

Eight Hour Digestion

In addition to the sites seen in the 2-hr digestion, eight hours of proteolysis resulted in digestion at residues 32-36, not previously demonstrated to be accessible to the aqueous medium. Fragments XX, XXI, and XXII, encompassing residues 166–197 were included in Fig. 3 and Table 1 because of their agreement with that region of the bacteriorhodopsin sequence, but could be fragments of proteinase K based on their lack of radioactivity in separations of labeled samples (*see* Fig. 2*c* and Results, section V). The largest molecular weight band in the gel pattern of the 8-hr sample corresponds to the fragment 73 to 160. The ~5000 mol wt band should then contain the piece from 5–7 to 64–72, and some smaller fragments.

The 8-hr digestion led to removal of 20% of the tritiated leucine (about seven residues) from the membrane in the form of soluble peptides. Even allowing for the small metabolic carryover of tritium to amino acids other than leucine, this is less than the number of leucines not recovered in any of the identifiable bacteriorhodopsin HPLC peaks (10 leucine residues between residue 162 and the COOH terminal). The failure to clearly identify the sequence from residue 162 to the COOH terminal in the HPLC separations then must result from poor recovery of these peptides in HPLC and difficulty characterizing the peaks that contain them. Removal of the two hydrophobic segments in this region from all bacteriorhodopsin molecules would have been expected to lead to complete disruption of the native lattice and total loss of the native absorption spectrum, since the region contains the retinal binding site (lysine 216). These effects were not observed.

About 50% of the total intensity of the X-ray diffraction pattern was preserved after the 8-hr digestion (with little change in the relative intensities), even though no bacteriorhodopsin fragments greater than about 9000 mol wt could be detected by gel electrophoresis (Fig. 1). Based solely on this apparent molecular weight for the fragments, it is possible to state that a significant fraction of the bacteriorhodopsin molecules must have been maintained in a native-like conformation under digestion conditions where they all had at least one cleavage near the middle of the polypeptide chain. If, as seems likely, the first cleavage sites are at the extremities of the molecule and at the two internal sites identified by previous digestion experiments

and by HPLC in this study, the preservation of the diffraction pattern in the absence of fragments bigger than 9000 mol wt implies that molecules cut at least four times (at the ends plus two internal sites) maintain a native structure. The sharpness of the diffraction rings from the 2- and 8-hr digestions demonstrates that the cleaved molecules contributing to the diffraction pattern existed in an extended lattice, rather than as small oligomers.

Twenty Four Hour Digestion

Three types of fragments could be identified in the HPLC separations of purple membranes digested for 24 hr: peptides from amino acids 6–7 to 30–37, those from 132–135 to 155–159, and that from 175 to 186–188. This last fragment (XXVIII) may be a portion of proteinase K, as described above. The fragments recovered from the NH_2 terminus of bacteriorhodopsin are similar to some seen after 8 hr digestion. One of these (fragment XXV, Fig. 3) corresponds to a cleavage between proline 37 and aspartic acid 38, which has been reported by Liao et al. (1984) to be an acid-labile peptide bond. However, we do not believe fragment XXV was a product of acid hydrolysis since: (i) Our samples were subjected to formic acid for less than 10 min at room temperature (compared with 3–4 days at 37°C for the Liao et al. procedure). (ii) We saw other enzymatic cleavages in the region of the sequence surrounding these residues. (iii) We did not see this pro-asp cleavage under more mild conditions of enzymatic digestion where the same exposure to acid was involved.

The cleavages seen at residues 132–135 are in a region not previously identified as being protease-susceptible. Katre and Stroud (1981), however, reported lactoperoxidase-catalyzed iodination of tyrosines 131 and 133, indicating that these residues may be accessible to water-soluble enzymes in the intact protein.

By all criteria, bacteriorhodopsin that had been digested for 24 hr was greatly altered compared with the native protein. The absorbance at 568 nm disappeared, in agreement with the findings of Rosenheck et al. (1978) for extensive pronase digestion. All but the lowest resolution reflections were absent from the diffraction patterns. In addition, 50% of the labeled leucine could be washed free of the membrane following 24 hr proteolysis.

The preservation of the diffraction pattern and optical spectrum of bacteriorhodopsin after seen 8 hr digestion means that some of the additional cleavages detected between 8 and 24 hr must occur on molecules still retaining native-like structures. However, the eventual disruption of membrane structure raises the possibility that peptides recovered by HPLC after the longer time are not identical to transmembrane segments in the intact structure. Digestion may lead to a structural reorganization of the molecules resulting in cleavage at sites which are membrane-embedded in native bacteriorhodopsin, or protection of sites which are normally accessible. In addition, the large amount of material removed from the membrane raises the following issues:

1) The fate of peptides in the supernatant fractions: Failure of peptides to pellet could result from breaking up of the membranes with which they are associated into small pieces (perhaps mediated by the release of amphipathic peptides from the digested protein), or from complete dissociation of the peptides from membranes. Several factors support the second of these possibilities. (a) No significant amount of lipid could be detected in the supernatant fractions following digestion for any time interval. (b) All the ^3H leucine in the supernatant fractions from digestions of labeled bacteriorhodopsin eluted as small soluble peptides in gel chromatography. (c) The physical characteristics of the pellets from the low speed centrifugations are roughly similar for samples which have been digested for different amounts of time. (d) High speed centrifugation (about four times the relative centrifugal force of the low speed spins for eight times the duration) brought down a significant percentage of the radioactivity in the supernatant from undigested membranes, but very little of that from the most digested membranes (Table 2). This is consistent with the idea that the supernatants from undigested membranes consisted of patches which were too small to pellet at low speed, while those from digested membranes were small, soluble peptides.

2) The topology of the removed peptides with respect to the membrane: The peptides identified as remaining in the membrane after 24 hr of proteolysis do not include all seven transmembrane segments expected from structural analyses of bacteriorhodopsin. Even if the HPLC peaks we have identified do not represent the complete set of peptides remaining membrane-associated, the amount of ^3H leucine lost to the supernatants after digestion clearly demonstrates the removal of a significant fraction of the putative membrane-spanning sequences from the membrane. This finding is particularly striking in view of the distribution of label in the sequence, little of it being found outside of the hydrophobic segments. Thus, either some transmembrane segments are freed from the membrane by proteolysis, or the current view of the structure is incorrect.

In spite of some remaining controversy about the type of secondary structure that makes up the transmembrane segments (Jap et al., 1983), the three dimensional structural analysis of Unwin and Henderson (1975) clearly indicates the presence of seven transmembrane rods of density. In order to explain our results without requiring the displacement of any membrane-embedded portions of the protein, given the leucine distribution in the sequence, bacteriorhodopsin would have to be made up of 4–5 transmembrane segments. The difficulty of reconciling this possibility with the previously available structural information leads to the conclusion that proteolysis must result in the conversion of membrane-embedded sequences into soluble small peptides.

3) Mode of removal of hydrophobic peptides from the membrane: The energy required for the active site of a fully folded water-soluble hydrolytic enzyme to penetrate the hydrophobic core of a membrane is expected to be quite high. This is consistent with the widespread use of proteolytic enzymes as impermeant probes of the 'sidedness' of membrane preparations. Thus, digestion probably does not occur within the membrane. Rather, removal of hydrophobic peptides from the membrane occurs by destabilization of membrane-embedded segments.

The energy of partitioning of hydrophobic helices into membranes has been estimated to be 15–50 kcal/mole for the hydrophobic segments of bacteriorhodopsin (Engelman & Steitz, 1984). This means that these helices, as single transmembrane segments, spend a negligible fraction of their time outside of the membrane (e^{-25} or 10^{-11} in the case of the least stable helix). While considerable uncertainty exists in the estimations of the energies of transfer of amino acid residues from polar to nonpolar environments, in the estimation of the energies of the random-chain to helix transition of membrane proteins in solution, and in the extent of entropy gain on localization of a peptide in a membrane (Jahnig, 1983), such errors are not of the magnitude to allow hydrophobic segments of bacteriorhodopsin to spend significant amounts of time completely removed from the membrane. Protease binding could not decrease the energy barrier by more than 4–5 kcal/mole, the binding energy characteristic of the $\sim 10^{-3}$ M dissociation constant for serine protease-substrate binding (e.g. Hess, 1971).

The energy cost for removing a small number of residues at a time from the membrane could be much less than that for removing whole helices —as little as a kilocalorie or two per mole. A number of alternative mechanisms for exposing such short peptides can be envisioned: (i) Unfolding of the ends of transmembrane alpha helices. (ii) Local perturbation of the thickness of the bilayer. (iii) Translation, or 'bobbing' up and down, of transmembrane segments perpendicular to the plane of the membrane. Each of these mechanisms might be facilitated by binding to protease. If any of them occurred in a medium containing high concentrations of proteolytic enzyme, clipping could occur during excursions of peptide bonds out of the membrane. Shortening of the membrane-embedded fragment would render it less stable in the membrane, leading to progressively greater excursions, accelerated proteolysis, and eventually to total digestion of the peptide segment. Jahnig (1983) has considered the stability of membrane-inserted alpha helices as a function of helix length, concluding that segments just long enough to span half the bilayer are the shortest that would be expected to remain inserted.

Comparison of the peptides remaining after 24 hr of proteolysis with the hydrophobicity profiles of Engelman and Steitz (1984) is consistent with such mechanisms for the removal of the peptides from the membrane. The two sets of peptides reliably identified in the 24-hr digestion products (residues 6–7 to 30–37, fragments XXIII, XXIV, XXV and residues 132–135 to 155–159, fragments XXVI and XXVII) are two of the three deepest energy minima according to the calculations of these authors. They are also sharp minima. This means that there is an energy barrier to be surmounted before they can move perpendicular to the membrane, even by one or two residues. In addition, excursions of these fragments should be limited by the arginines and lysines at either end of the hydrophobic stretches (although lysine 159 at the COOH terminal end of one of these segments can eventually be removed). Basic residues may stabilize transmembrane segments through electrostatic interactions with acidic lipid head groups.

The region from residue 42 to residue 63 is a deep energy minimum according to the calculations of Engelman and Steitz. But, in comparison with the two recovered peptides, this minimum is not as sharp. The hydrophobic segment is bounded at its NH_2 terminus by two lysines, but the COOH terminal end trails off in a mixture of nonpolar and uncharged residues. This peptide may be capable of large excursions perpendicular to the membrane. Proteinase K cleavage would then generate new carboxyl and amino termini. Terminal amino groups are easily de-protonated and should be capable of entering the membrane in an uncharged form at neutral pH. The new NH_2 terminus thus presents less of a barrier to 'bobbing' and destabilization of the transmembrane peptide than did the original lysine side chains. Alternatively, the multiplicity of

possible membrane-inserted sequences in this portion of the protein could lead to multiple cutting sites and a broad distribution of digested peptides which would be difficult to characterize by HPLC. If motion of this segment perpendicular to the membrane occurs in native or lightly proteolyzed membranes, some of the previously reported papain sites at the COOH terminal end of the segment (which have been used in modeling the topology) might in fact represent conformations which are adopted only transiently by bacteriorhodopsin.

The four hydrophobic segments of the protein, which cannot be identified after extensive digestion, all contain at least one charged residue and form less deep and less sharp energy minima in the calculations of Engelman and Steitz (1984) than the three segments discussed above. In addition to simply having lower energy barriers to moving out of the membrane, some of these segments, when partially cleaved, may exhibit a tendency to orient along the membrane surface as amphipathic helices.

Since purple membranes are isolated in the form of flat patches which have a poorly understood structure at their edges, it is possible that digestion might occur through proteolysis at the patch margins. However, since the digested patches are easily pelleted and the released peptides are not associated with lipid, it seems that the patches are not breaking up. Even if digestion were initially occurring at the patch margins, once the margins were digested, the protease would have the problem of gaining access to molecules which are separated from the edge by a selectively protein-depleted, lipid-containing zone.

Any mechanism in which proteolysis occurs at patch margins would predict that the extent of proteolysis for a given patch would be a function of the patch size, since the ratio of peripheral to interior molecules would vary with size. Since membrane patches are heterogeneous in size, digestion would be expected to result in a broad distribution of patch densities. In equilibrium sucrose gradient centrifugations, digested membranes were seen to remain in tight bands of diminished density (results not shown), making edge-digestion an unlikely explanation for the observed proteolysis.

4) Ionic strength dependence of proteolysis: The digestions described here were carried out in solutions of low ionic strength. Gerber et al. (1977) proposed that bacteriorhodopsin undergoes a conformational change related to ionic strength. They found that bacteriorhodopsin was most susceptible to proteolysis at low salt concentrations (compared with the 4 M NaCl used to culture halobacteria). Much of the structural and functional characterization of bacteriorhodopsin has been carried out at low ionic strength, where the molecule retains a well defined structure (Blaurock, 1975: Henderson, 1975), as well as proton pumping capability, though the quantum efficiency of pumping is known to be ionic strength dependent (Racker & Stoeckenius, 1974; Ort & Parson, 1979; Govindjee, Ebrey & Crofts, 1980). The increase in protease susceptibility with decreasing ionic strength may result from a loosening of the folding of the aqueous parts of the protein, or may reflect a salt dependence in the activity of proteinase K. An additional possibility is a reduction in the solubility of the hydrophobic peptides in the aqueous medium in high salt, as occurs for small hydrophobic molecules (Tanford, 1980). This could decrease the extent of excursions of the peptides out of the plane of the membrane.

There has been considerable study of the rotational and translational motions of proteins in the plane of the membrane, but little is known about the motions of proteins in the perpendicular direction, even though such motion could have implications for the stability of insertion of membrane proteins. The mechanism of proteolysis described above may, for example, be related to the removal of cleaved signal peptides from the membrane after translocation of newly-synthesized secreted or membrane proteins. In *E. coli,* Zwizinski and Wickner (1980) have reported that a *soluble* protease performs this function. The general mechanisms of removal of unneeded membrane proteins from membranes are not known, and there is no information on whether any proteases are capable of activity in the hydrophobic membrane environment.

The destabilization of transmembrane segments described here raises questions about the use of proteolytic enzymes to characterize the topology of membrane proteins. In the absence of independent structural or spectroscopic evidence for retention of native structure by the protein being studied, proteolysis can lead to artifactual conclusions about protein topology. Detection of transmembrane segments protected against proteolysis demonstrates the presence of transmembrane portions of the protein under study, but the protected segments may not be the same as the transmembrane stretches of the native protein. Furthermore, failure to recover protected segments does not imply that none are present in the intact protein. We have recently reported another situation where the presence of proteolytic enzymes can alter the distribution of protein across membranes (Dumont & Richards, 1984).

This paper describes the identification of two regions of proteolytic cleavage not previously reported for bacteriorhodopsin in membranes. Although the new sites are consistent with proposed models for bacteriorhodopsin topology, the consid-

erable disruption of the membrane observed during digestion prevents us from being certain that these cleavages occur on bacteriorhodopsin molecules in their native conformation. X-ray diffraction patterns of digested membranes demonstrate that integrity of at least one, and probably two, of the presumed linking regions between hydrophobic transmembrane segments is not required for maintenance of the folded structure. HPLC followed by computer fitting of amino acid analyses to fragments of the protein sequence was found to provide a rapid and relatively inexpensive method for separating and identifying components of a diverse mixture of digestion products. Finally, we observed digestion of hydrophobic segments predicted to be stable as trans-bilayer helices. This has led to the proposal of mechanisms for the destabilization of membrane-embedded peptides which could play a role in the turnover of membrane proteins.

This research was supported by grants GM 21714 to F.M.R. from the National Institute of General Medical Sciences and AI 20466 to D.M.E. from the National Institute of Arthritis and Metabolic Diseases. M.E.D. was a Helen Hay Whitney Foundation postdoctoral fellow during some of the time this work was carried out.

We thank Jean-Luc Popot and Malcolm Capel for valuable discussions. Sue Ellen Gerchman provided able technical assistance in growing halobacteria and in X-ray diffraction. Amino acid analyses were performed by Tony Lanzetti, Gary Davis, and David Speicher, of the Protein Chemistry Laboratory of the Yale Membrane Center. Art Perlo assisted with computer programming. We are grateful for the assistance of Johnnie and Thelma Mouning in assembling this paper.

References

Argos, P., Rao, J.K.M., Hargrave, P.A. 1982. Structural prediction of membrane-bound proteins. *Eur. J. Biochem.* **128**:565–575

Blaurock, A.E. 1975. The structure of purple membrane from *Halobacterium halobium:* Analysis of the X-ray diffraction patterns. *J. Mol. Biol.* **93**:139–158

Brunner, J., Richards, F.M. 1980. Analysis of membranes photolabeled with lipid analogs. *J. Biol. Chem.* **255**:3319–3329

Dittmer, J.C., Wells, M.A. 1969. Quantitative and qualitative analysis of lipids and lipid components. *Methods Enzymol.* **14**:482–530

Dumont, M.E., Richards, F.M. 1984. Insertion of apocytochrome *c* into lipid vesicles. *J. Biol. Chem.* **259**:4147–4156

Engelman, D.M., Goldman, A., Steitz, T. 1982. The identification of helical segments in the polypeptide chain of bacteriorhodopsin *Methods Enzymol.* **88**:81–88

Engelman, D.M., Steitz, T.A. 1981. The spontaneous insertion of proteins into membranes. *Cell* **23**:411–422

Engelman, D.M., Steitz, T.A. 1984. On the folding and insertion of globular membrane proteins. *In:* The Protein Folding Problem. D.B. Wetlaufer, editor. pp. 87–113. AAAS Selected Symposium 89. Westview Press, Boulder, Colo.

Engelman, D.M., Zaccai, G. 1980. Bacteriorhodopsin is an inside out protein. *Proc. Nat'l. Acad. Sci. USA* **77**:5894–5898

Gerber, G.E., Anderegg, R.J., Herlihy, W.C., Gray, C.P., Biemann, K., Khorana, H.G. 1979. Partial primary structure of bacteriorhodopsin: Sequencing methods for membrane proteins. *Proc. Nat'l. Acad. Sci. USA* **76**:227–231

Gerber, G.E., Gray, C.P., Wildenauer, D., Khorana, H.G. 1977. Orientation of bacteriorhodopsin in *Halobacterium halobium* as studied by selective proteolysis. *Proc. Nat'l. Acad. Sci. USA* **74**:5426–5430

Govindjee, R., Ebrey, T.G., Crofts, A.R. 1980. The quantum efficiency of proton pumping of the purple membrane of *Halobacterium Halobium. Biophys. J.* **30**:231–242

Henderson, R. 1975. The structure of the purple membrane from *Halobacterium halobium:* Analysis of the diffraction pattern. *J. Mol. Biol.* **93**:123–138

Henderson, R., Unwin, P.N.T. 1975. Three dimensional model of purple membrane obtained by electroﬕ microscopy. *Nature (London)* **257**:28–32

Hess, G.P. 1971. Chymotrypsin-chemical properties and catalysis. *In:* The Enzymes (3rd Ed.) Vol. 3, pp. 213–248. Paul D. Boyer, editor. Academic, New York

Huang, K.-S., Bayley, H., Liao, M.-J., London, E., Khorana, H.G. 1981. Refolding of an integral membrane protein: Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin. *J. Biol. Chem.* **256**:3802–3809

Jahnig, F. 1983. Thermodynamics and kinetics of protein incorporation into membranes. *Proc. Nat'l. Acad. Sci. USA* **80**:3691–3695

Jap, B.K., Maestre, M.F., Hayward, S.B., Glaeser, R.M. 1983. Peptide chain secondary structure of bacteriorhodopsin. *Biophys. J.* **43**:81–89

Katre, N.V., Stroud, R.M. 1981. A probable linking sequence between two transmembrane components of bacteriorhodopsin. *FEBS Lett.* **136**:170–174

Khorana, H.G., Gerber, G.E., Herlihy, W.C., Gray, C.P., Anderegg, R.J., Nihei, K., Biemann, K. 1979. Amino acid sequence of bacteriorhodopsin. *Proc. Nat'l. Acad. Sci. USA* **76**:5046–5050

Kyte, J., Doolittle, R.F. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105–132

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685

Lemke, H.-D., Bergmeyer, J., Oesterhelt, D. 1982. Determination of modified positions in the polypeptide chain of bacteriorhodopsin. *Methods Enzymol.* **88**:89–98

Liao, M.-J., Huang, K.-S., Khorana, H.G. 1984. Regeneration of native bacteriorhodopsin structure from fragments. *J. Biol. Chem.* **259**:4200–4204

Liao, M.-J., Khorana, H.G. 1984. Removal of the carboxyl-terminal peptide does not affect refolding or function of bacteriorhodopsin as a light-dependent proton pump. *J. Biol. Chem.* **259**:4194–4199

Mao, D., Wallace, B.A. 1984. Differential light scattering and absorption flattening effects are minimal in the circular dichroism spectra of small unilamellar vesicles. *Biochemistry* **23**:2667–2673

Michel-Villaz, M., Saibil, H.R., Chabre, M. 1979. Orientation of rhodopsin alpha helices in retinal rod outer segment membranes. *Proc. Nat'l. Acad. Sci. USA* **76**:4405–4408

Michel, H., Oesterhelt, D., Henderson, R. 1980. Orthorhombic two-dimensional crystal form of purple membrane. *Proc. Nat'l. Acad. Sci. USA* **77**:338–342

Moore, W.M., Holladay, L.A., Puett, D., Brady, R.N. 1974. On

the conformation of the acetylcholine receptor protein from *Torpedo nobiliana*. *FEBS Lett.* **45**:145–149

Oesterhelt, D., Stoeckenius, W. 1971. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biol.* **233**:149–152

Ort, D.R., Parson, W.W. 1979. Enthalpy changes during the photo-chemical cycle of bacteriorhodopsin. *Biophys.* **25**:341–354

Ovchinnikov, Yu.A., Abdulaev, N.G., Feigina, M.Y., Kiselev, A.V., Lobanov, N.A. 1979. The structural basis of the functioning of bacteriorhodopsion: An overview. *FEBS Lett.* **100**:219–224

Popot, J.-L., Trewhella, J., Gerchman, S.E., Engelman, D.M. 1984. Two-dimensional lattice of hybrid bacteriorhodopsin molecules. *In:* Proceedings of the 8th International Biophysics Congress, Bristol, U.K. p. 51

Racker, E., Stoeckenius, W. 1974. Reconstitution of purple membrane vesicles catalyzing light-driven proton uptake and adenosine triphosphate formation. *J. Biol. Chem.* **249**:662–663

Rohorek, M., Heyn, M.P. 1979. Binding of all-trans retinal to the purple membrane: Evidence for cooperativity and determination of the extinction coefficient. *Biochemistry* **18**:4977–4983

Rosenheck, K., Brith-Lindner, M., Lindner, P., Zakaria, A., Caplan, S.R. 1978. Proteolysis and flash photolysis of bacteriorhodopsin in purple membrane fragments. *Biophys. Struct. Mechan.* **4**:301–313

Ross, A.H., Radhakrishnan, R., Robson, R.J., Khorana 1982. Glycophorin as studied using photoactivatable phospholipids. *J. Biol. Chem.* **257**:4152–4161

Stubbs, G.W., Smith, H.G., Jr., Litman, B.J. 1976. Alkyl glucosides as effective solubilizing agents for bovine rhodopsin: A comparison with several commonly used detergents. *Biochim. Biophys. Acta* **425**:46–56

Swank, R.T., Munkres, K.D. 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium docecyl sulfate. *Anal. Biochem.* **39**:462–477

Tanford, C. 1980. The Hydrophobic Effect, John Wiley & Sons, New York

Tarr, G.E., Crabb, J.W. 1983. Reverse phase high performance liquid chromatography of hydrophobic proteins and fragments thereof. *Anal. Biochem.* **39**:99–107

Trewhella, J., Anderson, S., Fox, R., Gogol, E., Khan, S., Engelman, D.M. 1983. Assignment of segments of the bacteriorhodopsin sequence to positions in the structural map. *Biophys. J.* **42**:233–241

Unwin, P.N.T., Henderson, R. 1975. Molecular structure determination by electron microscopy of unstained crystalline specimens. *J. Mol. Biol.* **94**:425–440

Von Heijne, G., Blomberg, C. 1979. Trans-membrane translocation of proteins: The direct transfer model. *Eur. J. Biochem.* **97**:175–181

Walker, J.E., Carne, A.F., Schmitt, H.W. 1979. Topography of the purple membrane. *Nature (London)* **278**:653–654

Wallace, B.A., Henderson, R. 1982. Location of the carboxyl terminal of bacteriorhodopsin in purple membrane. *Biophys. J.* **39**:233–239

Zwizinski, C., Wickner, W. 1980. Purification and characterization of the leader (signal) peptidase from *Escherichia coli*. *J. Biol. Chem.* **255**:7973–7977

Received 15 August 1985