Isolation and characterization of the CNBr peptides from the proteolytically derived N-terminal fragment of ovine opsin

Michael BRETT and John B. C. FINDLAY*  
Department of Biochemistry, University of Leeds, Leeds LS2 9JT, U.K.

(Received 5 November 1982/Accepted 24 February 1983)

Ovine rhodopsin may be cleaved in situ by Staphylococcus aureus V8 proteinase into two membrane-bound fragments designated V8-L (27000 mol.wt.) and V8-S (12000 mol.wt.). After purification of the proteolysed complex by affinity chromatography in detergent using concanavalin A immobilized on Sepharose 4B, the two polypeptide fragments may be separated by gel-permeation chromatography on Sephadex LH-60. Digestion of the N-terminal-derived V8-L fragment with CNBr in 70% (v/v) trifluoroacetic acid resulted in a peptide mixture that could be fractionated by procedures involving gel-permeation chromatography in organic and aqueous solvents and the use of differential solubility. The complete or partial sequences of all ten peptides are reported.

Rhodopsin, the photoreceptor molecule of vertebrate retinal rod cells, comprises an integral membrane glycoprotein, opsin, of mol.wt. 38000, to which is bound a molecule of 11-cis-retinal (for review, see Abrahamson & Ostroy, 1981). The primary role played by the protein in the processes of visual transduction is still the subject of much controversy. It would be useful, therefore, if the structure, organization and disposition of the polypeptide chain in the bilayer were known.

Unfortunately, progress in the sequence analysis of such integral membrane proteins has been hampered by their content of relatively long stretches of hydrophobic amino acids, which results in problems of insolubility or aggregation when classical methods for protein and peptide isolation are used. One helpful procedure could be proteolytic degradation of the polypeptide chains while they are still in the membrane, an approach which has been used for bovine rhodopsin (Pober & Stryer, 1975; Sale et al., 1977) and the anion-transport protein of human erythrocytes (Steck et al., 1976; Williams et al., 1979).

We have reported that incubation of ovine rod cell membranes with the proteinase from Staphylococcus aureus V8 resulted in the rapid and efficient proteolysis of rhodopsin into two large membrane-bound fragments, designated V8-L (27000 mol.wt.) and V8-S (12000 mol.wt.), together with a small seven-residue peptide, which was released into the supernatant (Findlay et al., 1981). The S. aureus V8-proteinase-digested rhodopsin was stable in the dark and in detergent and could be readily purified as the undissociated complex by affinity chromatography on concanavalin A–Sepharose 4B (van Breugel et al., 1977). The V8-L and V8-S polypeptides may then be separated by gel-permeation chromatography on Sephadex LH-60 in solvent A. We have exploited this strategy to purify and completely sequence the C-terminal region (109 residues) of ovine opsin (Findlay et al., 1981).

In the present paper we describe procedures for the purification of the peptides of V8-L after cleavage of this fragment with CNBr. A combination of gel-permeation chromatography in novel organic solvent mixtures together with differential solubility of peptides in aqueous solvents has allowed the preparation of the fragments in yields and purities required for sequence determination.

Experimental

Materials

Concanavalin A–Sepharose 4B and Sephadex
LH-60, LH-20 and G-50 were obtained from Pharmacia Fine Chemicals (G.B.) Ltd. (Hounslow, Middx., U.K.). *S. aureus* V8 proteinase was purchased from Miles Laboratories (Stoke Poges, Slough, Berks., U.K.); chymotrypsin and trypsin were obtained from Worthington (Croydon, Surrey, U.K.). *N,N*-Dimethylformamide (Sequanal grade, Hypovial) was purchased from Pierce and Warriner (Chester, U.K.). Other sequencing reagents and solvents, including trifluoroacetic acid, *N*-methylmorpholine, phenyl isothiocyanate, methanol and *p*-phenylene di-isothiocyanate, were sequencer grade from Rathburn (Walkerburn, Peeblesshire, Scotland, U.K.) or Fluka (via Fluorochem Ltd., Glossop, Derbyshire, U.K.) and 0.25 m-Quadrol exclusively from the latter company. CNBr, Polybren and p-phenthaldehide were obtained from the Aldrich Chemical Co. (Gillingham, Dorset, U.K.). Triethylamine, from BDH Chemicals (Poole, Dorset, U.K.), was redistilled and stored under *N*₂. Controlled-pore AEAP-glass (pore size 7.5 nm) was purchased from LKB (Cambridge, U.K.). Dodecyl trimethylammonium bromide was bought from Eastman Kodak (Kirkby, Liverpool, U.K.) and recrystallized before use. Spectrapor 3 dialysis tubing came from Spectrum Medical Industries (via Raven Scientific Ltd., Haverhill, Suffolk, U.K.). Keisegel 60F thin-layer plates were obtained from Merck via BDH. All other reagents and solvents were bought from BDH Ltd., the highest grades available being used.

**Methods**

**Preparation of rod outer segments (ROS).** Sheep eyes (obtained from the Leeds Municipal Abbatoir) were removed from freshly killed animals and stored in lightproof containers on ice. Retinas were excised under dim red light (15 W pigmy bulb; Kodak Wratten series 1 filter) within 3 h, frozen in liquid *N*₂ and stored at −20 °C in the dark. All subsequent protein-purification steps were performed in this dim red light at 4 °C unless otherwise indicated.

ROS membranes were prepared by mild homogenization and flotation on sucrose solutions as previously described (Brett & Findlay, 1979), except that a 67 mM-phosphate buffer, pH 7.4, was used.

**Purification and delipidation of *S. aureus*-V8-proteinase-cleaved rhodopsin.** ROS membranes containing about 1.5 μmol of rhodopsin were suspended to a protein concentration of 2 mg·ml⁻¹ in 67 mM-phosphate buffer, pH 7.4, and *S. aureus* V8 proteinase was added to 2% (w/w). Digestion was allowed to proceed for 3 h at 30 °C and terminated by rapidly cooling to 4 °C, pelleting the membranes (50000 g for 20 min) and washing them three times with 100 ml of phosphate buffer. The digested washed ROS membranes were solubilized in 20 ml of 0.1 M-sodium acetate buffer, pH 6.0, containing 3% (w/v) DTAB, 0.1 M-NaCl and 1 mM-each of CaCl₂, MgCl₂, MnCl₂ and dithiothreitol (the last added immediately before use). The solubilized membrane preparation was applied to a 1.0 cm × 5.0 cm column of concanavalin A-Sepharose 4B (10 mg of concanavalin A/ml of settled bed resin) equilibrated in the above buffer but containing only 1% (w/v) DTAB. The resin was washed with at least 2 column volumes of buffer and the *S. aureus*-V8-proteinase-cleaved rhodopsin specifically eluted with 0.2 M α-methyl D-mannoside. The eluted protein was dialysed at 4 °C against 3 litres of 0.25 M-Tris/HCl, pH 8.6, containing 0.5 mM-dithiothreitol and 0.25% (w/v) DTAB.

**Carboxymethylation.** Solid DTAB, dithiothreitol and guanidinium chloride were added to the dialysed material to give final concentrations of 1% (w/v), 5 mM and 6 M respectively. The solution was stirred for 1 h at room temperature and iodoacetic acid, dissolved in 0.5 ml of 1.0 M-NaOH, added to 10 mM. The reaction was allowed to proceed in the dark for 1 h and terminated by the addition of fresh dithiothreitol to 10 mM. The material was dialysed against 3 × 5 litres of water for 24 h at 4 °C. The precipitated carboxymethylated material was stored as a freeze-dried powder.

**Gel-permeation chromatography and organic solvents.** Chromatography using organic solvents was performed on Sephadex LH-60 or LH-20 resins. Two solvent systems, A and B, were used (see the ‘Abbreviations used’ footnote). Freeze-dried or rotary-evaporated protein and peptides were dissolved in aq. 90% (v/v) formic acid and the other components of the solvent mixture added sequentially to give the final volume ratios. Chromatography in aq. 20% (v/v) formic acid was performed with Sephadex G-50 equilibrated in this solvent.

**Cleavage with CNBr.** The purified V8-L fragment was dissolved in anhydrous trifluoroacetic acid (4 mg of V8-L/ml of trifluoroacetic acid) and the solution adjusted to 70% (v/v) by the addition of water. A 300-fold molar excess of CNBr over methionine was added and the reaction mixture left, stirring in the dark at room temperature for 24 h. The reaction was terminated by rotary-evaporating the sample to dryness.

**Isolation of CNBr peptides.** The peptide mixture resulting from treatment of V8-L with CNBr was dissolved in solvent A and subjected to gel-permeation chromatography on Sephadex LH-60 equilibrated with solvent A or, for peptides with less than 30 residues, on Sephadex LH-20 equilibrated with solvent B. Column eluates were monitored by their *A*₂₈₀ and the analysis of portions by the manual ninhydrin method of Hirs (1967).

Where described, further purification of hydrophobic peptides was carried out by dissolving the peptide mixture in 1 ml of 90% formic acid and adjusting to 20% by the addition of water. Fraction-
ation was then performed by chromatography on Sephadex G-50 equilibrated withaq. 20% formic acid.

**Amino acid analysis.** Peptides were hydrolysed in 6m-HCl (containing 3 mg of phenol/ml) at 110°C for 24 and 72 h in sealed evacuated tubes. Compositions were determined with a Chromaspek J180 amino acid analyser and peak integration performed automatically with a Digico micro 16V computer.

**Carbohydrate analysis.** Neutral sugars and hexosamines were assayed by g.l.c. of the trimethylsilyl derivatives of the O-methyl glycosides (Laine et al., 1972). Mannitol was used as the internal standard and methanolysis was carried out in methanolic 1m-HCl under N₂ at 90°C for 24 h. G.l.c. was performed on a Pye 204 gas–liquid chromatograph with an SE 30 2-metre capillary column with a temperature program of 140–200°C at 1°C/min.

**Amino acid-sequence analysis.** Solid-phase sequencing. Peptides (50–200nmol) were coupled to AEAP-glass by the conventional C-terminal lactone coupling procedure (Horn & Laursen, 1973) or by using DITC (Wachter et al., 1973). A Rank–Hilger APS-240 sequencer with standard delivery program was employed. The buffer system comprised 10% (v/v) N-methylmorpholine in methanol mixed 3 : 2 (v/v) withaq. 50% (v/v) propan-1-ol and the pH was adjusted to 8.5 with trifluoroacetic acid.

Thiazolinone derivatives were converted into their corresponding phenylthiohydantoin derivatives manually in 200µl of 20% (v/v) trifluoroacetic acid at 70°C for 15 min. Amino acid phenylthiohydantoin derivatives were identified by reverse-phase h.p.l.c. with 5µ Zorbax-ODS resin and a Du Pont 8800 pump system coupled to a WISP autosampler (Waters Associates). The program was an acetonitrile gradient (20–40%, v/v) in 12 mM-sodium acetate, pH 4.3, with a sample cycle time of 22 min (Zimmerman et al., 1977). Quantification of amino acid phenylthiohydantoin derivatives was achieved with a Spectra Physics SP4100 computing integrator.

Residue assignments were confirmed by t.l.c. of the amino acid phenylthiohydantoin derivatives (Bridgen et al., 1975), back-hydrolysis in 6m-HCl containing 0.1% (w/v) SnCl₂ at 150°C for 4h in vacuo to the parent amino acid (Mendes & Lai, 1975) and/or absorbance measurements at 313nm (serine and threonine) rather than 254nm.

Spinning-cup sequencing. Peptides (50–200nmol) were sequenced with a Beckman 890C sequencer using 0.25m-Quadruplex buffer and the delivery program described by Brauer et al. (1975). Peptides were solubilized and presented to the cup in 90% (v/v) formic acid and mixed thoroughly with 3 mg of Polybrene (Tarr et al., 1978). Thiazolinone conversion and analysis of phenylthiohydantoin derivatives were performed as described above.

**Blocking of amino groups before proline (background reduction).** To increase the length of sequence runs, some peptides were treated with o-phthalaldehydes at cycles where previous sequence determinations had established the presence of a proline residue (Machleidt & Hofner, 1982). A 5 mg portion of o-phthalaldehyde was dissolved in 5 ml of sequencing buffer and applied to the glass-coupled peptide mounted in the sequencer column at 30°C, in 1 ml samples at 5 min intervals. The glass was washed with methanol and the Edman degradation recommenced. For the spinning cup, fluorescamine was used instead of o-phthalaldehyde (Bhown et al., 1981).

**N-terminal-residue determinations.** The N-terminal amino acids of the V8-L fragment and its CNBr peptides were determined by a modified dansyl chloride procedure (Gray, 1967). Dansylation was carried out in 1% SDS (added to the peptide in 90% formic acid, followed by exhaustive drying over NaOH)/0.2 m NaHCO₃. The N-terminal dansyl-amino acids were extracted from subsequent acid hydrolysates by ethyl acetate oraq. 50% (v/v) pyridine. Dansylated amino acids were identified by two-dimensional ascending chromatography on polyamide layers [see Findlay & Brew (1972) and references therein].

**Enzyme digestions.** Hydrophobic peptides, which were refractory to digestion both as insoluble suspensions or when solubilized in detergent, were readily digested when coupled to AEAP-glass. The glass-coupled peptide (50–200nmol) was suspended in 400µl of 0.1 m-NH₄HCO₃ and digested with 50µg of trypsin or chymotrypsin for 3h at 37°C. The digestion was terminated by washing the glass successively with methanol, trifluoroacetic acid and methanol and drying the resin in vacuo.

**Results**

**Isolation of fragment V8-L.**

The purification by gel filtration using Sephadex LH-60 of the two membrane-bound polypeptide fragments (V8-L and V8-S) obtained by proteolysis in situ of rhodopsin with S. aureus V8 proteinase was described previously (Findlay et al., 1981). Solvent A gives complete solubility of the two fragments, good flow properties with the LH-60 resin and permits excellent separation of the V8-L fragment (apparent mol.wt. 27000) from the 102-residue V8-S fragment. Any lipid and detergent present are eluted in the column volume. Amino acid analysis of material in the eluate other than the two pools shown revealed no peptide material. Starting with 180–200 retinas the yield of the V8-L fragment was 0.8–1.0µmol. No free N-terminal residue could be detected by dansylation. The amino acid composition of the fragment is given in Table 1.
Table 1. Amino acid composition of fragment V8-L. Hydrolyses were carried out in vacuo for 24, 48, 72 and 96 h at 110°C in 6M-HCl containing 3 mg of phenol/ml. Serine and threonine values were corrected to zero-time hydrolysis and 96 h values were used for valine, isoleucine, and leucine residues. Abbreviations used: CmCys, carboxymethylcysteine; N.D., not determined.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Composition (mol/100 mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CmCys</td>
<td>2.0</td>
</tr>
<tr>
<td>Asx</td>
<td>6.3</td>
</tr>
<tr>
<td>Thr</td>
<td>6.4</td>
</tr>
<tr>
<td>Ser</td>
<td>4.5</td>
</tr>
<tr>
<td>Glx</td>
<td>8.8</td>
</tr>
<tr>
<td>Pro</td>
<td>7.2</td>
</tr>
<tr>
<td>Gly</td>
<td>7.1</td>
</tr>
<tr>
<td>Ala</td>
<td>7.1</td>
</tr>
<tr>
<td>Val</td>
<td>8.7</td>
</tr>
<tr>
<td>Met</td>
<td>3.6</td>
</tr>
<tr>
<td>Ile</td>
<td>5.1</td>
</tr>
<tr>
<td>Leu</td>
<td>10.5</td>
</tr>
<tr>
<td>Tyr</td>
<td>5.7</td>
</tr>
<tr>
<td>Phe</td>
<td>9.4</td>
</tr>
<tr>
<td>His</td>
<td>2.5</td>
</tr>
<tr>
<td>Trp</td>
<td>N.D.</td>
</tr>
<tr>
<td>Lys</td>
<td>2.4</td>
</tr>
<tr>
<td>Arg</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Separation of CNBr peptides

The elution profile of the peptides resulting from cleavage of V8-L with CNBr in 70% (v/v) aqueous trifluoroacetic acid is shown in Fig. 1. Solubilization of the digest in solvent A was complete, and subsequent fractionation on Sephadex LH-60 equilibrated with this solvent allowed the designation of seven pools (A–G) on the basis of the A_280. The whole elution volume was included in these seven pools.

Isolation and characterization of peptides from Sephadex LH-60 pools

The amino acid compositions of the peptides described below are given in Table 2; the sequences are presented in Fig. 4 (below). Tables containing the phenylthiohydantoin assignments and yields are available from J.B.C.F. on request.

Pools E–G

Examination of pool G by direct application of a sample hydrolysed with 6M-HCl to the amino acid analyser gave peaks eluting in the homoserine and homoserine lactone positions. Similar examination of an unhydrolysed sample gave no ninhydrin-positive peaks. This material (CNBr 1) therefore consisted of a single homoserine residue blocked at its α-amino group. Attempts to liberate the N-terminus by mild methanolic hydrolysis (Kawasaki & Itano, 1972) were unsuccessful. The indications are, therefore, that the protein contains a methionine residue at its N-terminal, blocked with an acetyl group as reported for the bovine opsins molecule (Tsunasawa et al., 1980). One can conclude, therefore, that no Met-Met sequences occur in the V8-L fragment.

Pool F, on re-fractionation using Sephadex LH-20 in solvent B, gave three distinct peptides (CNBr 2, 3 and 4) containing 5, 8 and 12 residues respectively. Pool E contained three peptides, two of which were recovered by dissolving the freeze-dried mixture in 90% (v/v) formic acid and adding water to give a final formic acid concentration of 20%. The soluble material from this procedure was rechromatographed on Sephadex LH-20 to yield a 20-residue (CNBr 5) and a 12-residue (CNBr 4) peptide. The precipitated fraction, after re-fractionation on Sephadex LH-60 in solvent A, had the composition given as CNBr 6.

From the yields, compositions and sequences of these peptides, it was clear that CNBr 4 eluted partly in peaks E and F. The complete spectrum of peptides could, therefore, be obtained more easily by combining all the material in the entire region, drying and redissolving in 90% (v/v) formic acid and adjusting to 20% formic acid. The precipitate which formed (CNBr 6) was removed by centrifugation and rechromatographed as above. The soluble material was dried and applied to Sephadex LH-20. The elution profile is reproduced in Fig. 2.

The four peptides (CNBr 2–CNBr 5) obtained by this fractionation were subjected to automated solid-phase sequence analysis after coupling to AEAP-glass. All the peptides coupled in 90% or better yield and were completely sequenced to give structures that were in excellent agreement with their
Table 2. Amino acid compositions, yields and N-termini of the CNBr peptides of fragment V8-L

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Peptide</th>
<th>CNBr 1</th>
<th>CNBr 2</th>
<th>CNBr 3</th>
<th>CNBr 4</th>
<th>CNBr 5</th>
<th>CNBr 6</th>
<th>CNBr 7</th>
<th>CNBr 8</th>
<th>CNBr 9</th>
<th>CNBr 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>CmCys*</td>
<td></td>
<td>0.4 (1)</td>
<td>0.7 (1–2)</td>
<td>0.4 (1)</td>
<td>0.4 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asx</td>
<td>1.8 (2)</td>
<td>1.8 (2)</td>
<td>0.4 (0)</td>
<td>3.0 (3)</td>
<td>3.5 (4)</td>
<td>1.5 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>0.8 (1)</td>
<td>1.0 (1)</td>
<td>1.3 (1)</td>
<td>2.0 (2)</td>
<td>2.5 (3)</td>
<td>5.4 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>0.8 (1)</td>
<td>0.9 (1)</td>
<td>1.3 (2)</td>
<td>0.7 (1)</td>
<td>2.6 (3)</td>
<td>0.2 (0)</td>
<td>2.1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gix†</td>
<td>1.2 (1)</td>
<td>1.1 (1)</td>
<td>2.7 (3)</td>
<td>4.9 (6)</td>
<td>4.9 (5)</td>
<td>1.4 (1)</td>
<td>2.7 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>3.0 (3)</td>
<td>1.4 (1)</td>
<td>0.8 (1)</td>
<td>5.1 (5)</td>
<td>1.6 (2)</td>
<td>2.0 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>1.0 (1)</td>
<td>1.2 (1)</td>
<td>1.9 (2)</td>
<td>1.2 (1)</td>
<td>1.4 (1)</td>
<td>3.5 (3)</td>
<td>1.1 (1)</td>
<td>8.2 (8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>2.0 (2)</td>
<td>1.0 (1)</td>
<td>1.2 (1)</td>
<td>4.1 (4)</td>
<td>0.8 (1)</td>
<td>2.7 (3)</td>
<td>2.1 (2)</td>
<td>2.4 (2)</td>
<td>3.0 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>1.7 (2)</td>
<td>1.2 (1)</td>
<td>1.4 (1)</td>
<td>4.3 (5)</td>
<td>3.0 (3)</td>
<td>4.1 (4)</td>
<td>6.0 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>0.8 (1)</td>
<td>0.9 (1)</td>
<td>1.9 (2)</td>
<td>2.4 (3)</td>
<td>3.2 (3)</td>
<td>2.1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>1.0 (1)</td>
<td>0.2 (0)</td>
<td>1.6 (2)</td>
<td>1.7 (2)</td>
<td>1.5 (2)</td>
<td>1.3 (1)</td>
<td>9.9 (11)</td>
<td>7.3 (7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>1.0 (1)</td>
<td>0.9 (1)</td>
<td>1.9 (2)</td>
<td>1.3 (1)</td>
<td>3.3 (3)</td>
<td>1.8 (2)</td>
<td>2.7 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>1.1 (1)</td>
<td>1.7 (2)</td>
<td>1.8 (2)</td>
<td>3.8 (5)</td>
<td>4.0 (4)</td>
<td>3.6 (4)</td>
<td>6.3 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His†</td>
<td>1.5 (1)</td>
<td>1.2 (1)</td>
<td>1.7 (1)</td>
<td>2.3 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>0.9 (1)</td>
<td>0.7 (1)</td>
<td>0.9 (1)</td>
<td>1.6 (2)</td>
<td>0.8 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>0.7 (1)</td>
<td>0.8 (1)</td>
<td>0.8 (1)</td>
<td>0.9 (1)</td>
<td>0.8 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hse/Hse&gt;‡</td>
<td>1.0 (1)</td>
<td>0.6 (1)</td>
<td>0.6 (1)</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp§</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td>+ (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrate residues</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>20 (23/24)</td>
<td>32</td>
<td>38</td>
<td>42 (56/58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-terminus</td>
<td>— Leu Gly Ser Ala — Phe Asn Phe Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield (%)‖</td>
<td>72</td>
<td>85</td>
<td>58</td>
<td>84</td>
<td>80</td>
<td>85</td>
<td>32</td>
<td>81</td>
<td>76</td>
<td>75</td>
<td></td>
</tr>
</tbody>
</table>

* Carboxymethylcysteine values were low. This may reflect the partial modification of cysteine residues under the conditions used.
† Histidine values are increased by partial coincidence with homoserine lactone and glutamic acid values by homoserine.
‡ The presence of glutamic acid and histidine in a peptide does not allow an accurate assessment of homoserine and homoserine lactone content, although the presence of these residues is clear.
§ The presence of tryptophan was inferred from the relatively high absorbance of the peptide at 280 nm.
‖ Yield is based on the amount of material subjected to cleavage with CNBr.

Fig. 2. Fractionation of peptides from pools E and F on Sephadex LH-20 (column size 1.8 cm × 120 cm) equilibrated with solvent B

Pools E and F were combined and solubilized in 2 ml of 90% (v/v) formic acid. The solution was adjusted to 20% (v/v) formic acid by addition of water and the precipitate which formed was removed by centrifugation. The soluble fraction was dried, redissolved in 5 ml of solvent B, and applied to the column. The eluate was collected in 4.5 ml fractions at 20 ml/h and the A280 monitored. Portions were subjected to the manual ninhydrin method (Hirs, 1967) and the A570 read. Fractions were pooled as shown.

Fig. 3. Separation of CNBr 8 from CNBr 9

Pool C, after rechromatography in solvent B (see the text), was rotary-evaporated to dryness and redissolved in 1 ml of 90% (v/v) formic acid. The solution was adjusted to 20% (v/v) formic acid by the addition of water and applied to a column (2.0 cm × 140 cm) of Sephadex G-50 equilibrated with 99% (v/v) formic acid. Fractions (6 ml) were collected at a flow rate of 20 ml/h and the eluate monitored for A280. The fractions were pooled as shown.
corresponding amino acid compositions. CNBr 5 was digested with trypsin, after coupling to AEAP glass, to give complete cleavage at the arginine residue (14), and the sequence of the resultant C-terminal portion of peptide 5 is also shown (Fig. 4). It is noteworthy that the yield of peptide CNBr 4 was much decreased when CNBr digestion was performed in 70% (v/v) formic acid rather than 70% (v/v) trifluoroacetic acid, presumably because of the presence of a Met–Ser sequence.

CNBr 6, present in high yield and with a unique composition, did not give an N-terminal residue by the dansyl method and was refractory to both forms of automated sequencing procedures even though it coupled to AEAP-glass with more than 90% efficiency. The blocked N-terminus was not freed by treatment for 3 h at 37°C with 1 M HCl in anhydrous methanol (Kawasaki & Itano, 1972).

Since the amino acid analysis indicated the presence of a single lysine residue, the AEAP-glass-coupled peptide was treated with trypsin. Subsequent sequencing gave the structure seen in Fig. 4. The data suggested 30% cleavage at an Xaa–Thr bond and the presence of a Lys-Pro sequence that was resistant to trypsin digestion. The remaining N-terminal sequence of this peptide requires further investigation. Since the composition of the peptide indicated only one lysine residue, we assume that the partial cleavage is actually at a chymotryptic site.

**Pools C and D**

These pools were subjected, separately, to rechromatography on Sephadex LH-60 equilibrated with solvent B rather than solvent A. One component of the peptide mixture, most of which occurred in pool C and lesser amounts in pool B, completely aggregated in solvent B and eluted in the void volume of the Sephadex LH-60. It could therefore be resolved from the remaining peptides of pool C. Amino acid analysis of this peptide (CNBr 10) gave a composition very similar to that of the material in pool A, and both components had the same N-terminal residue, valine, and the same subsequent sequences.

The included peptide material from each pool was then dried, redissolved in 90% (v/v) formic acid and adjusted to 20% (v/v) formic acid by the addition of water. Fig. 3 shows the elution profile obtained when the material from pool C was chromatographed on Sephadex G-50 equilibrated with 20% (v/v) formic acid. Two peptides were obtained, a soluble aggregated component, CNBr 9, which eluted in the void volume of the column, and an included component, CNBr 8. Alternatively, the two peptides could be separated by dropwise addition of 33% (w/v) NH₃ until a flocculated precipitate comprising CNBr 9 appeared. Carbohydrate analyses of peptide CNBr 8 revealed the presence of N-acetylgalactosamine and mannose (Table 2). A yield of approx. 6 mol of mannose and 3 mol of N-acetylgalactosamine/mol of CNBr 8 was obtained.

Gel filtration of the material in Pool D using Sephadex G-50 gave much smaller amounts of the included glycopeptide and a larger void-volume peak. Rechromatography of this 'void' material on Sephadex LH-60 in solvent B gave two incompletely resolved peaks, the first of which contained...
some CNBr 9 and the second a new peptide, CNBr 7. If necessary, this latter material could be further purified by rechromatography and was sequenced by using the spinning cup (Fig. 5).

Peptides CNBr 8 and 9 were separately coupled to AEAP-glass in more than 90% yield and completely sequenced. Cycles 1 and 14 of the Edman degradation of peptide CNBr 8 each gave only an aspartic acid phenylthiohydantoin derivative in low yield, but the degradation proceeded normally through each cycle. This behaviour is characteristic of N-glycosidically linked asparagine residues (Tomita & Marchesi, 1975) and allowed a tentative identification of these positions as the attachment sites for the carbohydrate side chains. The more hydrophobic peptide CNBr 9 was sequenced to completion.

Tryptic digestion of the glycopeptide CNBr 8 coupled to AEAP-glass resulted in complete cleavage of the peptide bond C-terminal to arginine-20. The hydrophobic peptide CNBr 9, when coupled to AEAP-glass, proved equally amenable to digestion with trypsin, the peptide bond C-terminal to arginine-25 of this peptide cleaving in 100% yield. Sequence analysis of these two tryptic fragments confirmed the structures of residues 21–38 of CNBr 8 and 26–42 of CNBr 9 (Fig. 5).

The sequence CNBr 7', a larger version of CNBr 7 isolated from intact rhodopsin and coupled to AEAP-glass by using the homoserine lactone method, is also shown. This peptide (CNBr 7') was obtained from a CNBr digest by rechromatography on Sephadex LH-60 of material from the front edge of peak C after removal of CNBr 10 and the glycopeptide by methods involving aggregation and gel filtration using Sephadex G-50 as described above.

**Pool B**

N-Terminal determinations on this material gave a mixture of the amino acids valine, phenylalanine, isoleucine and leucine. Sequence determination using the spinning cup indicated the presence of at least four peptides in no more than 5% yield. These peptides are assumed to be partial cleavage products and no further work was attempted with them.

**Pool A**

As described above, re-fractionation of pools B and C in solvent B resolved various amounts of an aggregated component in the void volume of the Sephadex LH-60 resin. Comparison of this material by amino acid analysis, N-terminal determination and sequencing, with the material contained in pool A of the initial fractionation in solvent A (Fig. 1), suggested that these pools contained a single peptide, designated CNBr 10, with a relatively high threonine and glycine content and a N-terminal valine residue (Table 2). Presumably partial disaggregation of CNBr 10 is achieved in solvent A, accounting for the appearance of about 20% of the peptide over the volume of pools B and C. Attempts to attach this peptide to AEAP-glass by using the homoserine lactone method gave very poor coupling (~5%), but the peptide did couple via its lysine residue(s) using the DITC method (Wachter et al., 1973). A partial sequence determination was also achieved by using the spinning cup. The triple threonine sequence found at positions 6–8 of this peptide proved partially refractory to Edman degradation, resulting in a poor repetitive yield thereafter. Further sequence work is required on this material.

**Discussion**

In the present work, proteolysis in situ of an integral membrane protein has been exploited to generate a few large polypeptide fragments whose preliminary resolution greatly simplified the problems of subsequent peptide purification. Our strategy has been to use the highly specific proteincase from *S. aureus* V8, which cleaves the rhodopsin molecule at two extradiscal sites to give two membrane-bound fragments, V8-L (mol.wt. 27000) and V8-S (mol.wt. 12000) (Findlay et al., 1981). The difference in molecular weight of these fragments allows their separation by gel-permeation chromatography on Sephadex LH-60 with organic-solvent mixtures that completely solubilize and disaggregate the polypeptides. The usefulness of this procedure has been demonstrated by the elucidation of the complete sequence of the V8-S fragment (Findlay et al., 1981) and the isolation and characterization of the CNBr peptides from V8-L reported here. Similar techniques have been described elsewhere for structural studies on bovine rhodopsin using other proteinases (Bonting et al., 1974; Pober & Stryer, 1975; van Breugel et al., 1975; Trayhurn et al., 1974a,b; Saari, 1974; Klip et al., 1976; Sale et al., 1977; Towner et al., 1977; Hargrave et al., 1980; Mullen & Akhtar, 1981).

Digestion of fragment V8-L with CNBr was routinely performed in aq. 70% (v/v) trifluoroacetic acid because early experiments indicated that partial cleavage and incomplete dissociation were significant problems when 70% (v/v) formic acid was used as the solvent. In particular, the yield of peptide CNBr 4 was less than 20% in 70% (v/v) formic acid. The N-terminal serine residue of this peptide suggested that its poor recovery was due to a Met–Ser peptide bond in the V8-L fragment (Schroeder et al., 1969).

The use of organic-solvent mixtures B and, particularly, A (Findlay et al., 1982) prevented the problems of insolubility and irreversible aggregation for most peptides during fractionation. We have also
found the latter solvent mixture particularly useful in the delipidation of membrane proteins. The success of these solvents in overcoming the hydrophobicity of the opsin molecule suggests their general applicability to membrane proteins.

The peptides described here, together with the C-terminal sequence previously published (Findlay et al., 1981), account for the entire opsin molecule. The V8-L peptides added to the 109-residue C-terminal sequence comprise nearly 350 residues, in good agreement with the apparent molecular weight of 38000 for vertebrate opsin.

Comparison of the structures of peptides CNBr 1 and 8 from ovine opsin with the reported bovine opsin N-terminus (Hargrave et al., 1980), reveals almost complete sequence homology. A point of discrepancy arises at residues 34 and 35 of the glycopeptide (CNBr 8), the ovine sequence being Trpâ-Gln35 in contrast with the Glx34-Trp35 sequence reported for bovine protein. The glycopeptide isolated from porcine opsin also contains the sequence Trpâ-Gln35 (M. Brett, unpublished work). In agreement with the finding of two carbohydrate-attachment sites at asparagine residues 1 and 14 of the bovine glycopeptide (Hargrave, 1977), our data suggest that both sites in the ovine molecule also contain N-glycosidically linked carbohydrate. Further, the structures of these carbohydrate moieties are probably identical in both proteins, since we detect only mannose and N-acetylglucosamine and in yields consistent with the structure reported for bovine opsin (Fukuda et al., 1979; Liang et al., 1979). These similarities contrast strikingly with variations in the sequences after the retinal-binding lysine residue (Eliopoulos et al., 1982; Findlay et al., 1981; Mullen & Akhtar, 1981; Pellicone et al., 1981).

CNBr peptides 1 and 8 can be assigned to the N-terminus of the opsin molecule, ordered NH3â–1–8– by homology with the known N-terminal sequence of the bovine protein (Hargrave, 1977; Schechter et al., 1979; Hargrave et al., 1980). The C-terminus of V8-L should be a glutamic acid residue (Houmard & Drapeau, 1972) and peptide CNBr 7 satisfies this condition. From the structure of CNBr 7', it can be seen that this peptide represents CNBr 7 plus the N-terminal peptide of V8-S. We would suggest, therefore, that this constitutes the linking region between V8-L and V8-S. Ordering of the remaining peptides cannot at present be confirmed and the establishment of overlaps is likely to prove a time-consuming task. Chemical cleavage at Trp–Xaa bonds is an obvious choice of strategy, although the reports of Pellicone et al. (1981) that several large fragments isolated from bovine opsin after cleavage with 2-(2-nitrophenylsulphenyl)-3-methyl-3'-bromoindolenine (BNPS-skatole) were partial-cleavage products with blocked N-termini make this procedure problematic. A more successful approach might be the generation of small peptides by partial acid hydrolysis and the isolation of the methionine-containing peptides.

CNBr 7' obviously loops out into the cytosol because of the presence of a S. aureus-V8-proteinase-sensitive site. The accessibility of cysteine and lysine residues in two other peptides to labelling when intact discs were treated with various hydrophilic probes (M. Brett, P. Barclay and J. B. C. Findlay, unpublished work) suggests that these two peptides contain regions of the molecule that also loop out into the cytosolic compartment. Taken together with the likelihood that V8-S contains two intramembranous segments (Eliopoulos et al., 1982), it is possible to identify seven potential transmembrane segments in the whole polypeptide. Importantly, from an analysis of the sequences, these intramembranous regions do not contain exclusively hydrophobic amino acids. Polar residues such as serine, threonine and tyrosine may be relatively common in these segments, and charged residues are present although more rarely than is the case for bacteriorhodopsin.

During the review of this manuscript our attention was drawn to the work of Ovchinnikov et al. (1982) on the primary structure of bovine opsin. Sequence homology between the bovine and ovine molecules is extensive, only seven amino acid differences being identified in the compared sequences of the equivalent V8-L regions. The comparative analysis confirms the peptide alignments discussed above and allows the complete ordering of the ovine CNBr peptides as:

\[ \text{N-1-8-2-9-10-4-3-5-6-7-C} \]

The ovine protein has one methionine residue (and therefore one CNBr peptide) less than bovine opsin, valine at position 6 of CNBr 9 replacing methionine in the bovine sequence. Of the remaining six amino acid differences, four require only a single base change in the respective codons to effect the observed substitution, but two adjacent residues in CNBr 6, namely Lys-Pro, are His-Glu in the bovine sequence and cannot be so explained. In fact there is an interesting clustering of the amino acid differences in CNBr 6 and the adjacent peptides that remains to be explained.

The availability of comparative sequence data for an integral membrane protein such as opsin opens the possibility of investigating the structure/function constraints on these molecules. The precise conservation of large hydrophobic segments that span the membrane may suggest restraints on the structure of the polypeptide which in turn may indicate that there is more to the phospholipid bilayer than simply the provision of a hydrophobic environment.
This research was supported by grants to J. B. C. F. from the Research Council (grant BR/B/25315) and Medical Research Council (grant G979/352N). We are grateful to Dr. M. Chaplin, who carried out the carbohydrate analyses, and Mr. A. Thackrah, who performed the amino acid analyses.

References


Brauer, A. W., Margolis, N. M. & Haber, E. (1975) *Biochemistry* 14, 3029–3035


