Biochem. J. (1986) 236, 389–395 (Printed in Great Britain)

Identification of the sites in opsin modified by photoactivated azido[^{125}I]iodobenzene

Matthew D. DAVISON and John B. C. FINDLAY*
Department of Biochemistry, University of Leeds, Leeds LS2 9JT, U.K.

Opsin labelled with photoactivated 1-azido-4-[^{125}I]iodobenzene was proteolysed *in situ* with *Staphylococcus aureus* V8 proteinase to yield two radioactive membrane-bound fragments. These were separated, cleaved with CNBr and the resultant peptides sequenced in order to locate the radiolabelled residues. In the whole molecule, there was clear evidence for modification of at least 20 sites, identified as derivatives of cysteine, tryptophan, tyrosine, histidine and lysine residues. The probe primary reacted, therefore, with nucleophilic substituents. The positions of the modified sites relative to the confines of the phospholipid bilayer were consistent with all other studies on the disposition of the polypeptide chain. The location of these sites substantiated an earlier suggestion that not all the transmembrane segments should be regarded as continuous regular α-helices.

INTRODUCTION

Although it is useful to identify which segments of a polypeptide chain penetrate the hydrophobic phase of the bilayer, much more meaningful information can come from a knowledge of exactly which residues are exposed to the fatty acid milieu. The approach of choice has involved the use of photosensitive hydrophobic probes of various kinds, and in a few cases it has been possible to identify those sites modified by the photoactivated reagent (Brunner & Richards, 1980; Richards & Brunner, 1980; Ross *et al.*, 1982; Hoppe *et al.*, 1983*a*,*b*, 1984). In a previous paper (Davison & Findlay, 1986), the critical optimal conditions for the labelling of opsin with 1-azido-4-[^{125}I]iodobenzene were established and used to determine the exposed residues in the retinal-binding domain of the protein. In the present study the remaining two-thirds of the polypeptide have been similarly analysed in order to obtain a more complete picture of the three-dimensional organization of the molecule.

EXPERIMENTAL

All the materials and methods used in the present study have been described in detail in a previous paper (Davison & Findlay, 1986). In summary, photoreceptor disc membranes were incubated with 0.65 mM-1-azido-4-[^{125}I]iodobenzene at 4 °C and the suspension very rapidly frozen as a thin film in liquid N_2 before u.v. radiation for 60 min at −100 °C. The thawed membranes were subjected in the dark to proteolysis with *Staphylococus aureus* V8 proteinase, solubilized, reduced and carboxymethylated, and the two membrane-bound fragments isolated, all as described previously (Brett & Findlay, 1983). Cleavage of the larger V8L fragment with a 100-fold molar excess of CNBr, the purification of the various peptides and their sequences were also reported.

Sequence analysis

Sequence analysis was carried out by solid-phase methods using the Hsl (Horn & Laursen, 1973) and TFA anhydride (Davison & Findlay, 1986) coupling methods. In order to ensure absolute purity of the peptides being sequenced, a blocking step involving *o*PA was employed at that cycle in the automatic sequencing when a proline residue was exposed at the *N*-terminal of the principal peptide (Machleidt *et al.*, 1982). At the end of the appropriate sequencer cycle, 5 ml of a solution of *o*PA (1 mg/ml) in 10% (v/v) *N*-methylmorpholine in methanol mixed 1:1 (v/v) with aq. 5% (v/v) propan-1-ol, was pumped through the glass support at 56 °C, in five portions over a period of 20 min. After extensive washing of the resin with methanol, the automatic sequencing program was restarted with a double TFA step for the proline residues. Since *o*PA reacts covalently, efficiently and essentially irreversibly with the *N*-terminus of all amino acids but proline (Brauer *et al.*, 1984), this procedure ensured that any contaminating material was irreversibly blocked and hence did not contribute in any way to the sequence obtained from the main peptide. Owing to the modest size of most of these peptides, reappearance of the background was not a significant factor.

In some instances, contaminating peptides containing Hsl at their *C*-termini could be rendered refractory to attachment to the glass support by allowing the TFA-activated lactone moiety to react with 15% (v/v) *n*-propylamine in DMF. This was particularly useful if the TFA anhydride method for glutamic acid residues was subsequently to be used.

Residue identification

Amino acid PTH derivatives were identified and quantified by h.p.l.c. using a modification of the procedure reported by Zimmerman *et al.* (1977).

Abbreviations used: AEAP, aminoethylaminopropyl; AIB, 1-azido-4-[^{125}I]iodobenzene; *o*PA, *o*-phthalaldehyde; Hsl, homoserine lactone; PTH, phenylthiohydantoin; TFA, trifluoroacetic acid; DMF, dimethylformamide.

* To whom correspondence and reprint requests should be addressed.

Table 1. Specific radioactivities of modified V8L residues

Values are calculated from the corrected recovery of radioactivity at the relevant round of sequencing and from the quantity of amino acid obtained using the average repetitive yield over the stretch of the peptide.

Residue	Sp. radioactivity (d.p.m./nmol)
Trp-35	244
Tyr-43	3049
Tyr-60	195
His-65	198
Lys-66	207
Tyr-74	318
His-152	328
Trp-161	610
Cys-167	1460
Trp-175	207
Cys-222	2115
Tyr-223	613
Lys-231	253

In order to confirm the identity of any modified residue (i.e. a round releasing radioactivity), samples were also subjected to C_{18} reverse-phase h.p.l.c. (μBondapak) using a linear gradient (1 ml/min) of 20–100% (v/v) acetonitrile in 0.05% TFA over 50 min. Although the chemical natures of the amino acid derivatives are not known, they were eluted at very different times which were characteristic of the particular modified amino acid. In this way it was possible to positively identify the residue modified and hence guard against misassignment due to contamination by peptides present in amounts too small (i.e. less than 5%) to be detected chemically.

RESULTS

The purification of the peptides obtained as a result of treatment of the V8L fragment with CNBr was carried out exactly as described elsewhere, except that in some cases gel-filtration steps utilizing Sephadex LH-60, LH-20 and G-50 were repeated to improve purity (Brett & Findlay, 1983). The numbering of peptides reflects their elution order in the initial fractionation on Sephadex LH-60. The results are presented in *N*- to *C*-terminal order, however, and the order of the peptides in the sequence is given in Fig. 9 (below).

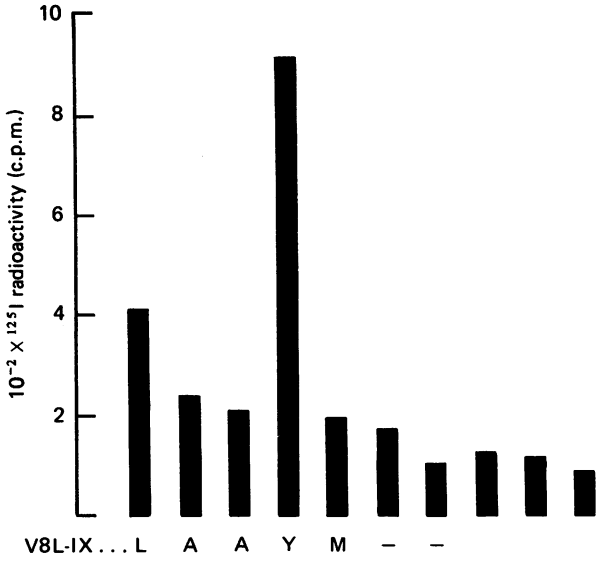

Fig. 2. Sequencing of V8L-IX

The peptide (30 nmol; 230000 c.p.m.) was coupled to AEAP-glass by the Hsl method in 11% yield. The initial sequencing yield was 80% and the repetitive yield 99%.

The figures show the radioactivity eluted at each cycle of the automated sequencing procedure and are not corrected for repetitive yield. The specific radioactivities of the various residues, as calculated from the radioactivity recovered and the corresponding PTH yield at that particular round, are shown in Table 1.

V8L-III (Fig. 1)

The blocking procedure using *o*PA was carried out during the sequencing of this peptide (38 residues long) at rounds denoted by the arrows. The only significant release of radioactivity occurred with the tryptophan residue at round 34.

V8L-IX (Fig. 2)

The radiolabel released at round 4 was attributed to a modified tyrosine residue. This was confirmed by further analysis by h.p.l.c., which revealed the double peak of

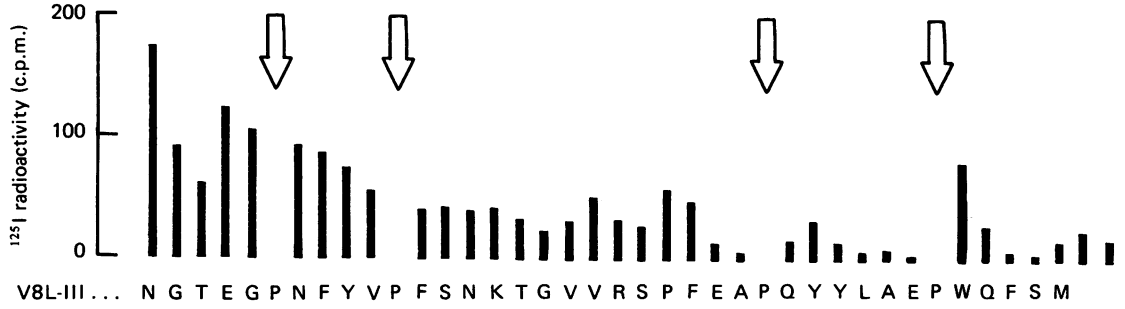

Fig. 1. Sequencing of V8L-III

The peptide (72 nmol; 11200 c.p.m.) was coupled to AEAP-glass in 60% yield. The initial sequencing yield was 38% and the repetitive yield 94% (Glu-4 to Glu-32). The arrows represent the positions at which the *o*PA blocking procedure was carried out. In this Figure and those following the single-letter code for amino acids is used.

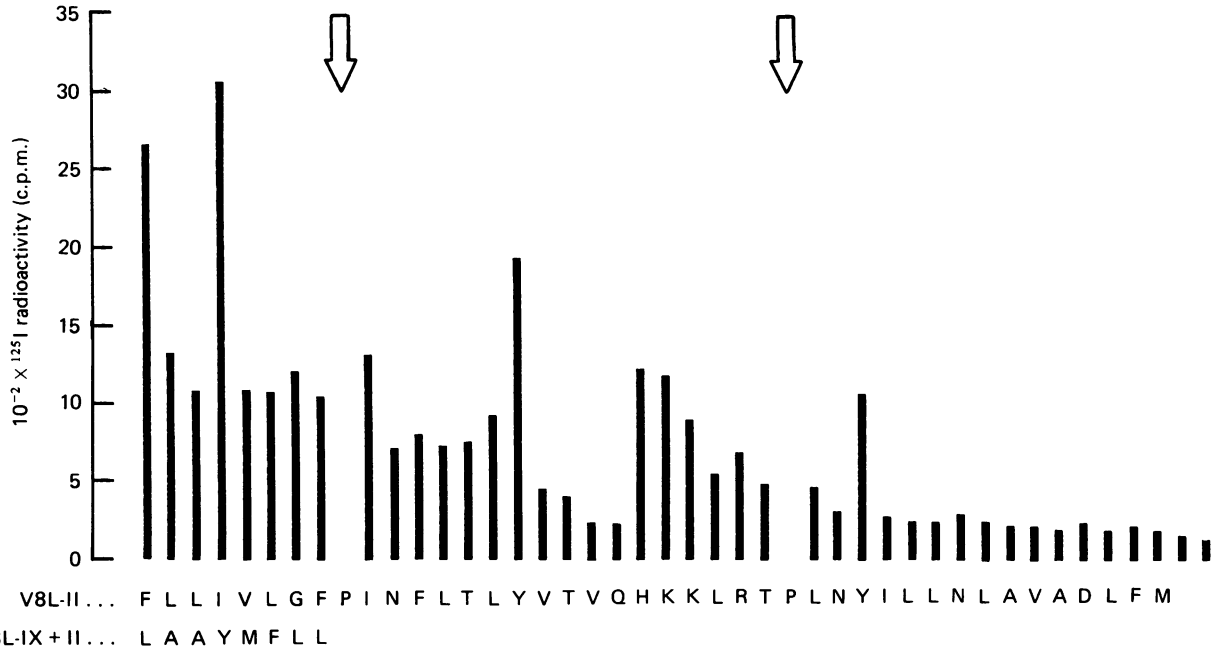

Fig. 3. Sequencing of V8L-II

The peptide (120 nmol; 180000 c.p.m.) was coupled to AEAP-glass in 59% yield. The initial sequencing yield was 32%. The *o*PA blocking procedure was performed at rounds 9 and 27 (⇩). The repetitive yield was 97% (Phe-1 to Phe-41).

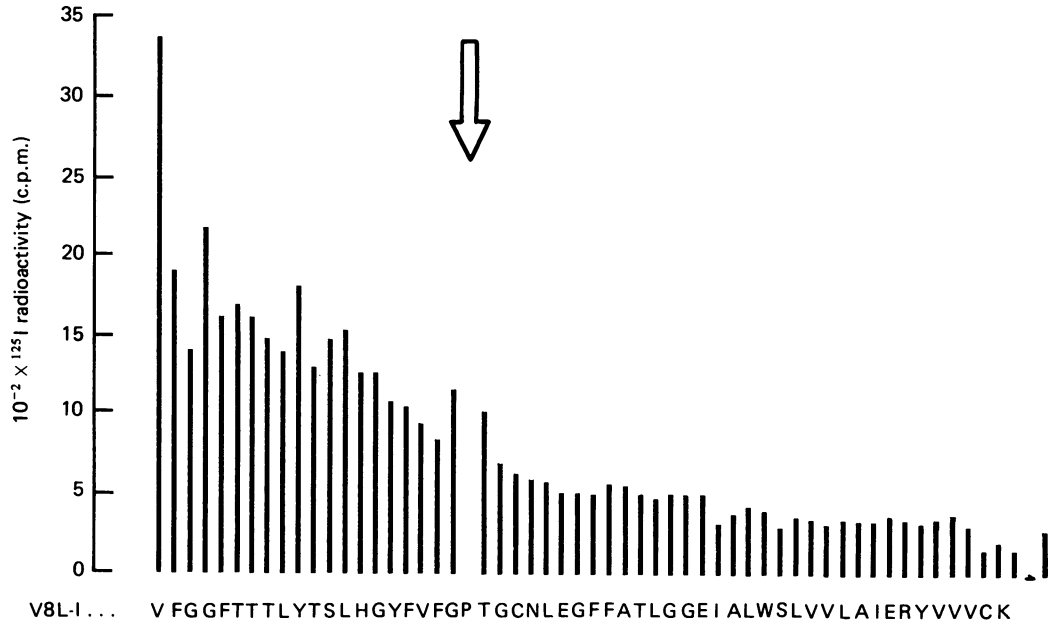

Fig. 4. Sequencing of V8L-I

The material (0.9×10^6 c.p.m.) coupled to AEAP-glass by the Hsl method in 14% yield. The initial sequencing yield was 28%. At the end of round 20 an *o*PA blocking treatment was performed. The repetitive yield was 96% (Val-1 to Val-53).

radioactivity characteristically associated with a modified tyrosine PTH derivative.

V8L-II (Fig. 3)

This peptide (42 residues) was isolated in 75–80% purity, the only contaminant being a partial cleavage product consisting of V8L-IX plus V8L-II. The radiolabel in round 4 was therefore ascribed to, and confirmed by h.p.l.c. as, the modified tyrosine residue in V8L-IX. The *o*PA-blocking reaction at rounds 9 and 27 ensured that the radioactivity released in rounds 16, 21, 22, and 30 was unambiguously associated with modified tyrosine, histidine, lysine and tyrosine residues (respectively) found at these positions in peptide V8L-II. Although lysine (23)

may also be a modification site, it is conceivable that at least some of the label at this position could be accounted for by carry-over.

V8L-I (Fig. 4)

Although this peptide (57 residues) was normally obtained in aggregated form and coupled in low yield, it nevertheless sequenced well (greater than 96% repetitive yield). Only one of the potential sites for modification, Tyr-10, corresponded with a release of radioactivity above background level, but this was not statistically significant. Since the peptide was obtained with high specific radioactivity, the background level of radioactivity was high, resulting from the usual release rate of 1–2% of peptide per round due to cleavage of the peptide–glass linkage by TFA. It is possible that this might obscure sites further into the peptide; however, it seems likely that potential sites, namely His-14, Tyr-16 and particularly Cys-22, were not available for reaction.

V8L-VII (Fig. 5)

V8L-VII (12 residues) was sequenced along with V8L-VI (20 residues) and V8L-VIII (eight residues) which were present at 10% and 7% respectively of the total peptide material. The release of radioactivity in rounds 4, 6 and 9 is consistent with modification of Cys-4 (V8L-VI), Trp-6 (V8L-VIII) and His-9 (V8L-VII). These results were supported by further results reported below. The sequence of V8L-VIII terminated at round 8 (i.e. before His-9).

V8L-VIII (Fig. 6)

This peptide (eight residues) was sequenced together with V8L-VII at a 10:3 molar ratio. The results support those presented above, suggesting that Trp-6 (V8L-VIII) was modified by the probe. Reverse-phase h.p.l.c. analysis of these rounds using the TFA/acetonitrile system confirmed the assignment of Trp-6 and also showed that any radioactivity associated with round 4 was associated with the particularly reactive Tyr-4 residue in any contaminating V8L-IX (see Fig. 2).

Fig. 5. Sequencing of V8L-VII

The peptide pool (35000 c.p.m.) containing V8L-VII, V8L-VI and V8L-VIII was coupled to AEAP-glass by using the Hsl method in 36% yield (based on radioactivity). Initial sequencing yield indicated that the peptides were present in the proportion 83:10:7 and sequenced with repetitive yields of 95%, 90% and 99% respectively.

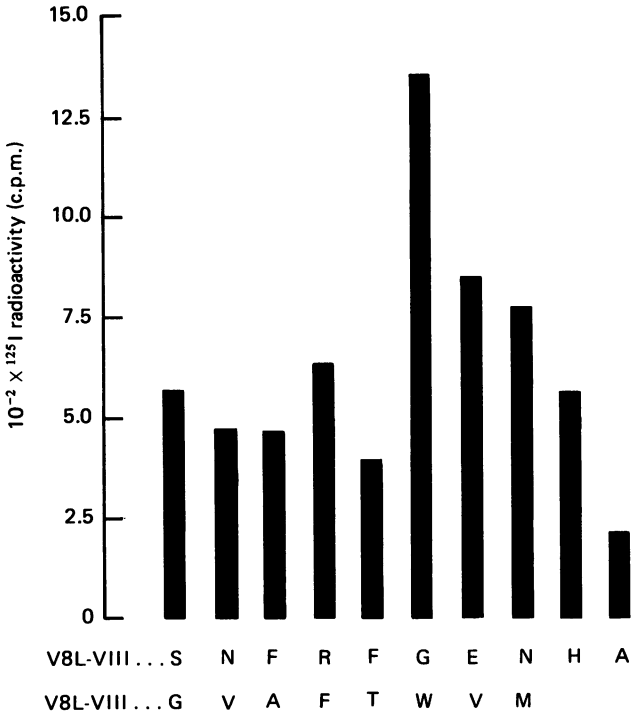

Fig. 6. Sequencing of V8L-VIII

The peptide (68000 c.p.m.) was coupled to AEAP-glass by the Hsl method in 44% yield. The initial sequencing yield was 35% and the repetitive yield 82%.

V8L-VI (Fig. 7)

The only contaminant seen on sequencing the major fraction of this peptide (20 residues) was V8L-VII, which could be irreversibly blocked at positions 7 and 8 by using the *o*PA procedure. Significant release of radioactivity at rounds 4 and 12 is consistent with modification of cysteine and tryptophan residues respectively. Interestingly, tyrosine (position 15), normally a very reactive residue, does not in this instance appear to have been modified. V8L-V, which can be purified by using its unique insolubility in 20% (v/v) formic acid followed by gel filtration with Sephadex LH-60 in formic acid/ethanol (3:7, v/v), could not be sequenced, presumably owing to a pyrrolidonecarboxylic acid residue at the *N*-terminus (Brett & Findlay, 1983).

V8L-IV (Figure 8)

This peptide (32 residues) was obtained in 80% purity, the principal contaminant being V8L-VI. Purity could be improved by first using the Hsl coupling procedure, which will effectively couple out those peptides containing a homoserine residue at their *C*-terminus and secondly by *C*-terminal blocking of any unbound Hsl peptides with *n*-propylamine. The residual uncoupled material (V8L-IV), which possessed a *C*-terminal glutamic acid residue, was then attached to a new batch of AEAP-glass by using the TFA anhydride method. Sequencing of this material gave release of radioactivity principally at rounds 15, 16 and 24, consistent with modified cysteine, tyrosine and lysine residues. These assignments were confirmed by h.p.l.c. analysis and ruled out problems of carry-over. The radioactivity at round 4 was shown to be characteristic of the cysteine PTH derivative rather than

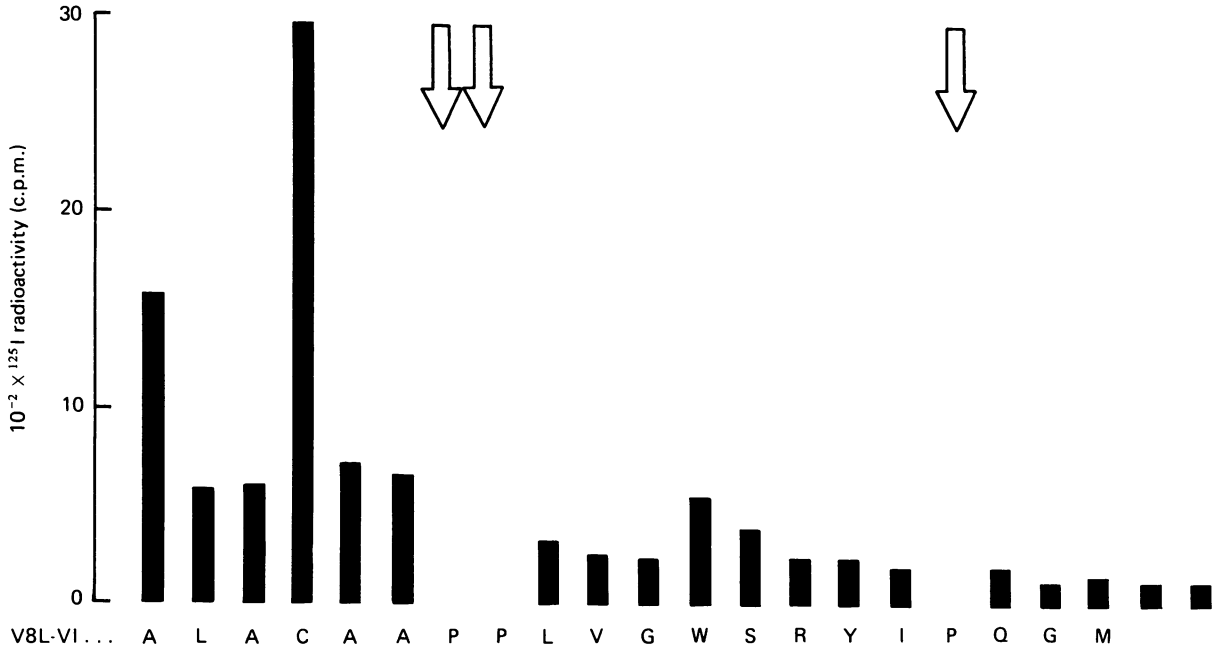

Fig. 7. Sequencing of V8L-VI

The peptide (112000 c.p.m.) was coupled to triethylenetetramine resin by the Hsl method in 63% yield. The repetitive yield was 92%. The *o*PA blocking procedure was carried out at rounds 7, 8 and 17 (⇩).

the histidine PTH derivative, and that at round 12 of a modified tryptophan PTH derivative, both suggesting low-level contamination with residual V8L-VI.

DISCUSSION

A broad spectrum of reactivity was thought desirable in hydrophobic photoactivated probes because it was believed that the intramembranous regions of proteins would consist almost exclusively of chemically unreactive hydrophobic residues with relatively few nucleophilic groups. Early experiments examined the use of the photogenerated nitrenes, very reactive species potentially capable of inserting into carbon–hydrogen bonds (Knowles 1972). From both this and other work, however, it must be concluded that the nitrene-generating aryl azides exhibit considerable selectivity towards nucleophiles, with little insertion into saturated hydrocarbons in homologous systems and even less in heterogenous systems (Reiser & Leyshon, 1971; Bayley & Knowles, 1978; Gupta *et al.*, 1979).

Assuming equal accessibility, these studies with opsin suggest that the overall chemical selectivity of photoactivated AIB was, in decreasing order, cysteine/tyrosine, tryptophan, histidine and lysine. It is possible that methionine residues are potential targets, but the approach followed here did not allow us to verify this. It must be remembered, however, that the peptides of opsin were generated and purified under very acidic conditions. It is possible, therefore, that other residues were also modified, but via bonds labile under these harsh conditions. Interestingly, this relatively narrow specificity of AIB has proved to be of significant practical advantage, since it simplified the problems of peptide purification and allows unambiguous identification of the modified sites. As a result, both the general and specific information gained on the structure of opsin has been of considerable value.

The exact chemical nature of the products of amino-acid-chain modification are uncertain, particularly since there is evidence from this and other work for the generation of reactive species from photolysed aryl azides which have considerably longer lifetimes than those characteristic of the nitrene radical (Mas *et al.*, 1980; Vandest *et al.*, 1980).

The first general point to be made is that the probe did not appear to significantly modify potential target sites that by other criteria are thought to exist in aqueous environments removed from the confines of the bilayer. The intradiscal *N*-terminal glycopeptide V8L-III is one example of this observation, for only Trp-35 (numbering now relative to the intact protein; see Fig. 9) was modified, despite the proximity of more reactive sites (e.g. tyrosine) and the absence of added lysine as scavenger in the sealed aqueous compartment.

Used in conjunction with hydrophilic probes, this approach seems to identify those residues which are in locations loosely described as the 'phospholipid-headgroup' region of the bilayer. Lys-66 (207 d.p.m./nmol) and Lys-67 (120 d.p.m./nmol), for example, are also modified by [^{14}C]succinic anhydride and diazo[^{125}I]diiodosulphanilic acid (Barclay & Findlay, 1984), but (and this may be significant) with reversed specific radioactivities (16.8 and 57.1 d.p.m./nmol respectively). It may be just this type of residue whose labelling is prevented by the addition of glutathione to the aqueous medium.

The general picture which emerges is almost entirely consistent with the model derived from structure-prediction studies (Pappin *et al.*, 1984 and confirmed by the data from hydrophilic labelling (Barclay & Findlay, 1984) and proteolytic cleavage (Ovchinnikov *et al.*, 1982). The only possible modifications are that Trp-35 and

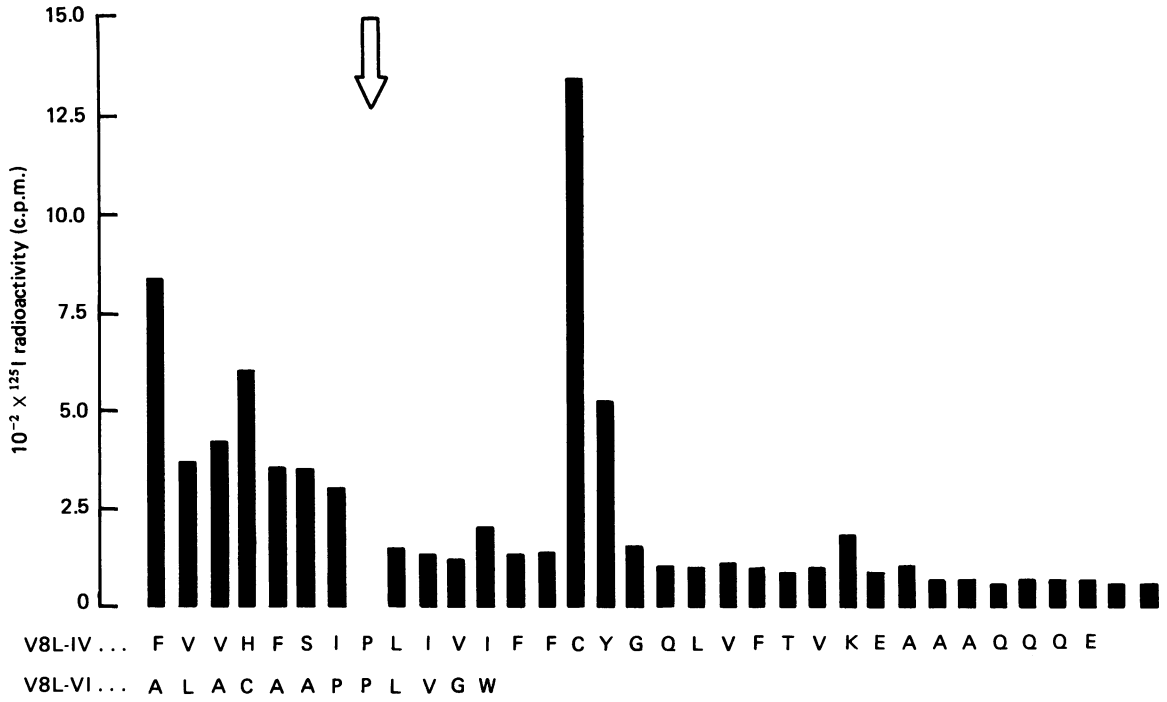

Fig. 8. Sequencing of V8L-IV

The peptide material (10^5 c.p.m.) was coupled to AEAP-glass by the Hsl method in 55% yield. The uncoupled fraction (68000 c.p.m.) attached to AEAP-glass in 34% yield using the TFA anhydride method. The repetitive yield was 96%.

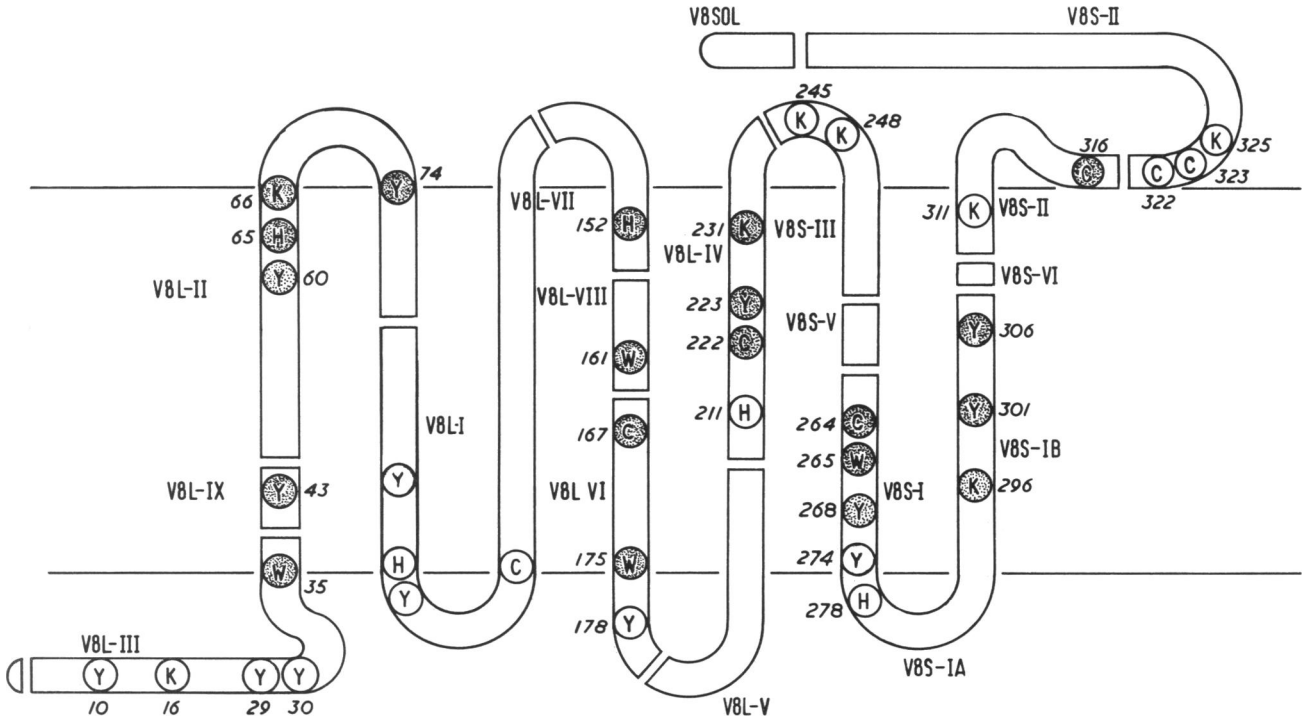

Fig. 9. Structural model of opsin in relation to the lipid bilayer as indicated from the location of photolysed AIB-modified residues

The CNBr peptides of opsin are shown in the order in which they appear in the sequence of the protein. They are arranged with respect to the bilayer/aqueous interface (horizontal lines). Modified residues are shown by stippled circles and unmodified residues that were potential labelling sites are shown by open circles.

Tyr-74 ought perhaps to be placed slightly closer to the bilayer, Lys-252 and Lys-255 further away.

In general, most of the tryptophan, tyrosine and cysteine residues are accessible to the probe, suggesting that these residues project outward from the protein surface into the lipid environment. Failure of any particular residue to react with the reagent may be indicative of a protected environment (e.g. an 'internal' orientation) or a special location (e.g. chromophore-binding site). Histidine residues at positions 100 and 211 are interesting examples of this. Similar conclusions can also be drawn as regards glutamic acid residues at positions 122 and 134, which were not reactive towards [^{14}C]dicyclohexyl carbodi-imide (P. Thompson & J. B. C. Findlay, unpublished work) under conditions that gave rise to labelling of the proteolipid protein of myelin (Lin & Lees, 1982). Finally, the ability to unambiguously demonstrate ready reaction with all but one (Cys-110) of the intramembranous cysteine residues indicates that these residues are probably not involved in disulphide-bond formation and therefore do not play a major role in the structural cohesion of opsin. This assumes, of course, that the probe does not react with disulphide bonds but we have no evidence to support this. It might be relevant, however, that these peptides are not purified (under acidic/non-reducing conditions) in oligomeric forms consistent with linkage via disulphide bonds.

It is likely that most, and probably all, of the transmembrane segments of opsin have one face exposed to the lipid milieu while the rest is involved in interactions with other regions of the polypeptide chain. Reactivity towards the hydrophobic probe under these experimental conditions can reasonably be assumed to reflect intralipid disposition and so be indicative of an exposed face. Assuming that the transmembrane segments are α-helical, the relative orientation of the amino acid side chains can be portrayed diagrammatically by using the 'helical-wheel' concept. Identification of the modified residues should then reveal the minimum limits of the exposed surface and whether the distribution of the accessible and unlabelled residues are consistent with the concept of regular transmembrane helices. The picture that emerges from this exercise suggests that at least one (helix 7) and possibly two others (helices 1 and 4) are likely to contain significant distortions from a regular α-helical backbone. This lends a measure of experimental support to the conclusions of the structure prediction work (Pappin *et al.*, 1984). Drawing those various threads together it should be possible to construct a crude model for opsin with the use of computer graphic techniques.

The help and advice of Dr. M. Brett, Dr. D. J. C. Pappin and Mr T. Thackrah are gratefully acknowledged. This work was supported by grants from the Medical and Science and Engineering Research Councils to J. B. C. F.

REFERENCES

Barclay, P. L. & Findlay, J. B. C. (1984) Biochem. J. **220**, 75–84
Bayley, H. & Knowles, J. R. (1978) Biochemistry **17**, 2420–2423
Brauer, A. W., Oman, C. L. & Margolies, M. N. (1984) Anal. Biochem. **137**, 134–142
Brett, M. & Findlay, J. B. C. (1983) Biochem. J. **211**, 661–670
Brunner, J. & Richards, F. M. (1980) J. Biol. Chem. **255**, 3319–3329
Davison, M. D. & Findlay, J. B. C. (1986) Biochem. J. **234**, 413–420
Gupta, C. M., Radhakrishnan, R., Gerber, G. E., Olsen, W. L., Quay, S. C. & Khorana, H. G. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 2595–2599
Hoppe, J., Montecucco, C. & Friedl, P. (1983*a*) J. Biol. Chem. **258**, 2882–2885
Hoppe, J., Friedl, P. & Jorgensen, B. B. (1983*b*) FEBS Lett. **60**, 239–242
Hoppe, J., Brunner, J. & Jorgensen, B. B. (1984) Biochemistry **23**, 5610–5616
Horn, M. & Laursen, R. A. (1973) FEBS Lett. **36**, 285–289
Knowles, J. R. (1972) Acc. Chem. Res. **5**, 155–160
Lin, L.-F. M. & Lees, M. B. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 941–945
Machleidt, W. & Hoffner, H. (1982) in Methods in Protein Sequence Analysis (Elzinga, M., ed.), pp. 189–203, Humana Press, Clifton, NJ
Mas, M. T., Wang, J. K. & Hargrave, P. A. (1980) Biochemistry **19**, 684–692
Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Artamonov, I. D., Zolotarev, A. S., Miroshinikov, A. I., Martynov, V. I., Kostina, M. B., Kudelin, A. B. & Bogachuk, A. S. (1982) Bioorg. Khim. **8**, 1424–1427
Pappin, D. J. C., Eliopoulos, E., Brett, M. & Findlay, J. B. C. (1984) Int. J. Biol. Macromol. **6**, 73–76
Reiser, A. & Leyshon, J. (1971) J. Am. Chem. Soc. **93**, 4051–4052
Richards, F. M. & Brunner, J. (1980) Ann. N.Y. Acad. Sci. **346**, 144–163
Ross, A. H., Radhakrishnan, R., Robson, R. J. & Khorana, H. G. (1982) J. Biol. Chem. **257**, 4152–4161
Vandest, P., Labbe, J. P. & Kassab, R. (1980) Eur. J. Biochem. **104**, 433–442
Zimmerman, C. L., Apella, E. & Pisano, J. J. (1977) Anal. Biochem. **77**, 569–573

Received 10 September 1985/6 January 1986; accepted 22 January 1986