Amino Acid Sequence of the Encephalitogenic Basic Protein from Human Myelin

By P. R. CARNEGIE*

Russell Grimwade School of Biochemistry, University of Melbourne, Parkville, Vic. 3052, Australia

(Received 29 December 1970)

Myelin from the central nervous system contains an unusual basic protein, which can induce experimental autoimmune encephalomyelitis. The basic protein from human brain was digested with trypsin and other enzymes and the sequence of the 170 amino acids was determined. The localization of the encephalitogenic determinants was described. Possible roles for the protein in the structure and function of myelin are discussed.

An unusual basic protein, which accounts for 30% of the total protein of myelin from human brain (Eng, Chao, Gerstl, Pratt & Tavaststjerna, 1968), is able to induce experimental autoimmune encephalomyelitis (Alvord, 1968). Its isolation has been studied in several laboratories (Caspary & Field, 1965; Einstein, Csejtey, Davis & Rauch, 1965; Oshiro & Eylar, 1970), and it is similar in its physical properties to the corresponding protein from bovine myelin (reviewed by Carnegie, 1971). The protein has a high isoelectric point and is similar in size, shape and lack of helical content to histones (Palmer & Dawson, 1969; Oshiro & Eylar, 1970; Chao & Einstein, 1970) but differs from histones in amino acid composition (Carnegie, Bencina & Lamoureux, 1967). Basic proteins of quite different composition from the brain protein were found in peripheral nervous tissue (Bencina, Carnegie, McPherson & Robson, 1969).

The present paper presents evidence for the amino acid sequence of the human basic protein. The sequences of the *N*-terminal region (Carnegie, 1969*a*) and the area around the sole tryptophan residue (Carnegie, 1969*b*) have been published. Two errors at residues 86 and 107 in a preliminary account of the sequence (Carnegie, 1971) are corrected in this paper.

MATERIALS

Encephalitogenic basic protein. The protein was obtained from human brain 12-16 h after death. The brain was defatted with chloroform-methanol (2:1, v/v) and the protein extracted as described by Kies (1965). It was further purified by ion-exchange chromatography on

carboxymethyl-Sephadex and gel filtration on Sephadex G-75 with 0.1M-HCl as solvent (Carnegie *et al.* 1967). Samples of basic protein from different brains were not pooled.

Proteolytic enzymes. Pepsin and trypsin were from Worthington Biochemical Corp., Freehold, N.J., U.S.A. The trypsin was treated with chloromethyl 2-phenyl-L-1-toluene-*p*-sulphonamidoethyl ketone (Wang & Carpenter, 1965) to inhibit chymotryptic activity. Thermolysin was from Calbiochem Inc., Los Angeles, Calif., U.S.A. Subtilisin, papain, chymotrypsin and carboxypeptidases A and B were from Sigma Chemical Co., St Louis, Mo., U.S.A. Aminopeptidase O was from Röhm and Hass, Darmstadt, Germany.

METHODS

General. The strategy and techniques used to determine the amino acid sequence of the basic protein were similar to those developed by Dr B. S. Hartley and colleagues in Cambridge (Hartley, 1970).

Enzymic digestion. (a) Trypsin. The protein (150 mg) was digested with trypsin (1.5 mg) for 1.5 h in a pH-stat at pH 8 or in 1% NH_4HCO_3 solution at 38°C. (b) Pepsin. The protein (250 mg) was digested with pepsin (4 mg) in 5% (v/v) formic acid for 3 h at 38°C. (c) Chymotrypsin. The protein (100 mg) was digested with chymotrypsin (2 mg) for 1 h in 1% NH_4HCO_3 solution at 38°C. (d) Thermolysin. The protein (200 mg) was digested with thermolysin (2 mg) for 30 min in a pH-stat at pH 7.8 at 37°C. (e) Peptides. Peptides (100-200 nmol) were further digested with 25 or 50 μg of subtilisin, thermolysin, chymotrypsin or aminopeptidase O in 1% NH_4HCO_3 for 1 h at 38°C. Digestion with carboxypeptidase A and carboxypeptidase B was as described by Ambler (1967).

Gel filtration. Peptic and thermolytic digests were fractionated on a column of Sephadex G-15 (11.5 cm $\times$ 5.2 cm diam.), which was superimposed on a column of Sephadex G-50 (133 cm $\times$ 5.2 cm diam.), with 1% acetic acid as solvent. The effluent was continuously scanned at 225 nm and appropriate fractions were pooled and submitted to electrophoresis and chromatography.

* Present address: Medical Research Council Demyelinating Diseases Unit, Westgate Road, Newcastle upon Tyne NE4 6BE, U.K.

Table 1. *Tryptic peptides from the*

Peptide ...	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
Asp	2.2	2.1	1.0	1.1	—	1.0	1.2	—	1.0	—	1.0	—	—	—	—
Thr	1.1	0.9	—	—	—	—	0.9	—	2.0	—	1.0	—	0.7	—	1.0
Ser	1.1	—	1.1	1.1	1.1	—	1.2	1.0	1.1	0.9	—	—	1.2	2.0	1.1
Glu	—	2.0	—	—	2.0	—	1.0	1.1	—	2.0	—	—	1.0	—	1.0
Pro	—	1.0	—	—	—	—	—	—	—	1.1	1.0	1.1	3.0	1.0	1.0
Gly	2.4	—	—	—	2.0	2.2	1.8	—	—	6.3	—	3.6	1.1	1.2	1.0
Ala	—	—	—	1.0	1.0	—	1.0	1.0	2.8	1.0	—	—	—	1.0	1.0
Val	—	0.9	—	—	—	—	0.9	—	—	—	0.5	—	—	—	—
Met	—	—	—	—	—	—	—	—	1.0	—	—	—	—	0.9	—
Ile	1.9	—	—	—	—	—	—	—	—	—	0.4	—	—	—	—
Leu	1.0	—	—	—	—	—	0.9	—	1.0	—	—	—	—	—	1.0
Tyr	—	—	—	1.0	—	—	—	—	1.0	0.9	—	0.8	—	—	0.8
Phe	—	2.0	—	—	0.8	2.0	—	—	—	2.0	—	1.0	—	—	—
Lys	—	1.0	—	0.9	—	—	1.1	0.9	—	—	—	—	0.8	—	1.1
His	—	1.0	—	—	—	—	—	—	1.1	—	—	—	—	—	0.9
Arg	1.0	—	0.9	—	0.9	0.9	—	—	1.1	2.2	1.1	1.0	—	1.1	—
Trp	—	—	—	—	+†	—	—	—	—	0.8	—	—	—	—	—
Yield (%)	20	18	20	24	4.3	22	23	31	20	17	15	7.5	20	20	20
m_{Asp} (pH 6.5)	-0.18	-0.07	0.0	0.0	0.0	0.0	0.0	0.0	0.20	0.21	0.39	0.32	0.38	0.31	0.38
R_{Leu} (BAWP)	0.79	0.68	0.26	0.45	0.70	0.85	0.45	0.45	0.57	0.83	0.70	0.73	0.23	0.48	0.23

* Including MeArg.

† By Ehrlich reagent.

Electrophoresis and chromatography. High-voltage electrophoresis of the digests was carried out essentially as described by Ambler (1963) or Milstein (1968) on Whatman no. 3MM paper or for smaller quantities on Whatman no. 54 paper. Mobilities (m) at pH 6.5 refer to the distance from the leading edge of the peptide to the centre of the neutral zone as a fraction of the distance moved (or calculated from markers to have moved) by aspartic acid ($m = -1.0$) (Offord, 1966). A butan-1-ol-acetic acid-water-pyridine (15:3:12:10, by vol.) system (BAWP) (Waley & Watson, 1953) was used for chromatography on Whatman 3MM paper or for smaller quantities on Whatman no. 54 paper. Peptides were located with ninhydrin, Pauly reagent, Ehrlich reagent and Sakaguchi reagent (Smith, 1960) and their position calculated relative to leucine (R_{Leu}). Peptide zones were eluted with 0.04 M-NH₃ and dried on a rotary evaporator.

Amino acid analysis. Peptides were hydrolysed with constant-boiling HCl containing 1% phenol for 22 h at 105°C and the hydrolysates analysed on a Technicon C2 resin with the Technicon '40 ml-gradient system' and with norleucine as internal standard. The yield of amino acids was computed on a Technicon integrator-calculator.

N-Terminus and sequence determination. The N-terminal residues of the peptides were determined by the dansyl procedure (Gray, 1967) and the derivatives identified as described by Pink, Buttery, De Vries & Milstein (1970), and by electrophoresis (Gray, 1967) where it was necessary to distinguish between arginine and histidine. The dansyl-Edman procedure (Gray, 1967) was used to obtain the sequence. Arrows under the peptides show the amino acids that were identified.

Amide residues. Amide residues were assigned on the basis of electrophoretic mobilities of the peptides at pH 6.5 (Offord, 1966) and by digestion with aminopeptidase O.

RESULTS

The results are presented in two sections. The isolation and amino acid sequence of 30 peptides

from a tryptic digest of the encephalitogenic protein are presented in the first section. Evidence for the overlap of these peptides is presented in the second section from the products of digestion of the protein with various enzymes.

Amino acid sequences of tryptic peptides

The composition of the tryptic peptides that were isolated by electrophoresis at pH 6.5, chromatography in solvent BAWP, and, where necessary by electrophoresis at pH 3.5 is given in Table 1. Their sequences were determined by the dansyl-Edman procedure and are shown in Fig. 1. The larger peptides were further digested to yield smaller fragments whose composition and sequence were obtained as described in the Tables, Figures and the text below. A peptide 'map' of the tryptic digest has been published (Carnegie, McPherson & Robson, 1970).

Peptide T1 was an acidic peptide with a mobility consistent with there being two aspartic acid residues and one arginine residue. Additional evidence comes from a thermolysin digest of the protein which yielded an acidic peptide with the sequence Ile-Leu-Asp-Ser.

A subtilisin digest of peptide T2 yielded peptides T2S1 ($m -0.30$, R_{Leu} 0.31, yield 31%), T2S2 (m 0.41, R_{Leu} 0.99, yield 29%), T2S3 (m 0.52, R_{Leu} 0.64, yield 25%) and free phenylalanine. A thermolysin digest of peptide T2 gave, among others, peptides T2L1 ($m -0.29$, yield 15%) and T2L2 (m 0.31, yield 15%). Hydrolysis of peptide T2 for 22 h yielded one valine residue; hydrolysis for 96 h was required to demonstrate the second valine residue. The composition and sequence of these fragments (Fig. 1) established the sequence of peptide T2. The position of the

Fig. 1. Tryptic peptides from encephalitogenic basic protein from human brain. Residue identified by dansyl-Edman method is shown by $\rightarrow$. Amide groups were assigned from the electrophoretic mobility. Evidence for the sequence of the larger peptides is discussed in the text or in the references indicated, * Carnegie (1969a), † Carnegie (1969b). Some peptides were digested with papain (Pa), subtilisin (S) or thermolysin (L).

Table 2. *Mobilities at pH 6.5 and N-terminal amino acids of Edman degradation products of peptide T2S1*

Edman steps completed	m_{Asp}	DNS-amino acid	Residue (from mobility)
0	-0.30	Thr	Thr
1	-0.35	Glu	Gln
2	-0.38	Asp	Asp
3	0.00	Glu	Gln
4	0.00	Asp	Asp
5	0.35	Pro	Pro
6	0.45	Val	Val

Digestion with aminopeptidase O yielded aspartic acid but no glutamic acid. From a thermolysin digest of the protein an acidic peptide Val-Asp-Ala-Gln-Gly-Thr was isolated and its sequence determined.

Peptide T9 contained methionine. The sequence of the first eight residues had been determined (Carnegie, 1969a) from a cyanogen bromide digest of the protein. The sequence of the C-terminal portion of this peptide was obtained from a chymotryptic peptide from the protein, with the composition Asx 1.0, Ala 1.0, His 1.8, Arg 1.0. Aminopeptidase O liberated aspartic acid. The sequence was shown by the dansyl-Edman procedure to be Asp-His-Ala-(Arg,His).

Peptide T11 gave, on analysis of an acid hydrolysate, a low yield of valine and isoleucine. Digestion with aminopeptidase O yielded approximately equimolar proportions of isoleucine, valine and asparagine. The resistance of Ile-Val bonds to acid hydrolysis is well known.

Peptide T13 was a basic fragment the sequence of which was determined by the dansyl-Edman procedure. Digestion with papain (300 μ g) with 0.5 mg of sodium cyanide in pyridine-acetate buffer, pH 6.5, at 38°C for 17 h yielded peptides T13Pa1 (mobility at pH 2.1 relative to lysine 0.44, R_{Leu} 0.20, yield 43%), T13Pa2 (mobility 0.85, R_{Leu} 0.25, yield 5%) and T13Pa3 (mobility 0.50, R_{Leu} 0.40, yield 36%). Peptide T13Pa3 from its mobility and lack of N-terminal residue was apparently pGlu*-Gly-Lys. The composition and sequence of these fragments (Fig. 1) established the sequence of peptide T13.

Peptide T14 was adjacent to the C-terminal of the basic protein. In the original studies on the products from treatment of the basic protein with cyanogen bromide a small peptide was overlooked (Carnegie, 1969a). Later this peptide was isolated and characterized as Ala-Arg-Arg (Carnegie *et al.* 1970). Digestion of the basic protein with a mixture of carboxypeptidases A and B liberated 2 mol of arginine/mol. Hashim & Eylar (1969) originally claimed a different sequence for the C-terminal end

of the bovine and human proteins, but this has been revised (Eylar, 1970) and there is now agreement. The sequence of peptide T14 was confirmed by studies on a chymotryptic peptide derived from peptic peptide P11 (Table 7 and Fig. 4). Free arginine, which presumably arises from the C-terminal end of the protein was found in the peptide labelled T30.

Peptide T15 was a basic peptide. Digestion of peptide T15 with subtilisin yielded peptides T15S1 (m 0.30, R_{Leu} 0.71, yield 33%) and T15S2 (m 0.44, R_{Leu} 0.51, yield 38%). The composition and sequence of these fragments (Fig. 1) established the sequence of peptide T15.

Peptide T17 was a basic peptide which when digested with aminopeptidase O yielded aspartic acid, serine and histidine. From the dansyl-Edman procedure the sequence was found to be Asp-Ser-His-His-Pro-Ala-Arg. From its composition peptide T18 appeared to have resulted from the partial fission of a Lys-Asp bond by trypsin. This bond is known to be resistant to digestion and a similar peptide was isolated from the bovine encephalogenic protein by Kibler *et al.* (1969).

Peptide T23 was a basic peptide containing two unusual amino acids which were identified as ω -N-monomethylarginine and ω -NN'-dimethylarginine (Baldwin & Carnegie, 1971b). The N-terminal sequence was established by Edman degradation and hydrolysis of the products. Digestion of peptide T23 with thermolysin yielded peptides T23L1 (mobility relative to lysine at pH 2.1, 0.58, R_{Leu} 0.88, yield 45%), T23L2 (mobility 0.80, R_{Leu} 0.67, yield 40%), T23L3a (mobility 0.91, R_{Leu} 0.38, yield 6.4%) and T23L3b (mobility 0.91, R_{Leu} 0.40, yield 11.3%). The composition and sequence of these fragments (Fig. 1) established the sequence of peptide T23.

Peptide T23 was obtained in a yield comparable with the other peptides with similar chromatographic and electrophoretic properties. There was no evidence for peptides with Gly-MeArg in the tryptic digest. Thus it would appear that peptide bonds involving methylarginine are resistant to digestion by trypsin. Baldwin & Carnegie (1971b)

* Abbreviation: pGlu, pyrrolidone-5-carboxylic acid.

have shown that free arginine can occur in place of the methylated arginines but because of the small proportion (6%) of the un-methylated arginine neither of the possible tryptic peptides, which would result from fission at this residue, were obtained. The arginine in this position was shown to act as a receptor for the transfer of [^{14}C]methyl groups from *S*-adenosylmethionine (Baldwin & Carnegie, 1971a).

Overlap of tryptic peptides

Digestion of the basic protein with pepsin, thermolysin, chymotrypsin and cyanogen bromide was used to obtain peptides that would overlap the tryptic peptides. The amino acid compositions of selected peptides are presented in Tables 3 and 4. Yields are not reported since variable losses resulted during the gel-filtration steps and homogeneity rather than quantitative recovery was the aim. Detailed analysis of the larger peptic peptides is given in Tables 5, 6 and 7 and in Fig. 2. In Fig. 3 the use of these peptides in overlapping the tryptic peptides is summarized. A peptide 'map' of the peptic digest was published by Lennon, Wilks & Carnegie (1970) who studied the immunological properties of peptides P1, P2 and P3.

The order of the tryptic peptides in the *N*-terminal region of the protein was previously

established from a polypeptide derived from digestion with cyanogen bromide (Carnegie, 1969a). This sequence was supported by the isolation of peptide P4, which had a blocked *N*-terminal residue and a composition consistent with the first 14 residues in the protein.

Evidence for the order of tryptic peptides T9, T16, T29, T1 and T6 is presented in Figs. 2 and 3 and in Tables 3, 4 and 5. Thermolytic peptide L1, which was isolated from a digest of the whole protein, was also obtained from digestion of peptic peptide P5 (P5L3 in Fig. 2).

A large peptic peptide P7 (PK in Fig. 1 of Lennon *et al.* 1970) gave a number of overlaps. This peptide of 45 residues is similar to a peptide of 45 residues which Kibler *et al.* (1969) isolated from the bovine protein. Their peptide was derived from digestion of the basic protein during acid extraction by an acid proteinase, which is present in acetone-dried powders of nervous tissue. The proposed sequence of peptide P7 (Fig. 2) is similar to the *N*-terminal half of the sequence given by Kibler *et al.* (1969). The *C*-terminal portion of their sequence was in error and their revised sequence is now similar to that for peptide P7 (Shapira, McKneally, Chou & Kibler, 1971). The brain acid proteinase appears to cleave a His-Phe bond whereas pepsin cleaved the Phe-Phe bond. In some batches of peptide P7 the Asp-Pro bond, which is particularly labile to

Table 3. Amino acid composition of selected peptic peptides from the human encephalitogenic basic protein

Purification stages: G, gel filtration; E, electrophoresis at pH6.5; C, paper chromatography in solvent BAWP; E', electrophoresis at pH3.5.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Asp	—	1.3	1.3	—	2.1	1.0	3.9	1.0	—	—	1.0
Thr	—	—	—	—	2.0	—	2.3	1.0	1.0	—	—
Ser	2.0	2.8	2.0	2.9	1.0	1.0	4.0	—	2.2	1.0	2.9
Glu	2.3	2.0	2.4	1.9	—	—	3.1	—	1.0	—	—
Pro	0.8	1.3	1.2	1.4	1.0	—	4.0	—	3.8	—	0.8
Gly	4.2	6.0	5.6	1.4	2.2	1.1	7.0	—	3.0	—	2.8
Ala	1.4	2.2	2.5	1.0	2.1	—	3.3	—	—	—	1.2
Val	—	—	—	—	—	—	0.5	0.5	—	—	—
Met	—	—	—	—	0.9	—	—	—	—	—	0.8
Ile	—	—	—	—	0.7	0.9	—	0.5	—	1.0	—
Leu	—	—	—	—	1.8	—	1.0	—	2.0	—	1.0
Tyr	—	1.0	1.0	0.8	—	—	0.8	—	—	—	—
Phe	1.9	1.9	1.1	—	0.9	1.0	1.8	1.0	—	—	0.9
Lys	—	—	—	2.0	—	—	2.9	1.0	1.1	1.1	1.3
His	—	—	—	1.0	2.8	—	4.9	—	—	—	—
Arg	2.2	2.8	1.9	2.4	2.8	1.1	3.8	—	2.2*	—	4.0
Trp	1†	1‡	1‡	—	—	—	—	—	—	—	—
<i>m</i> _{Asp}	0.20	0.17	0.0	0.67	0.23	0.0	0.43	0.30	0.62	0.53	0.62
<i>R</i> _{Leu}	0.50	0.40	0.45	0.12	0.11	0.55	0.01	0.65	0.32	0.51	0.17
Purification	GECE'	GECE'	GECE'	GEC	GEC	GEC	GECE'	GEC	GECE'	ECE'	GECE'

* Includes methylarginine.

† By aminopeptidase digestion.

‡ Detected by Ehrlich reagent.

Table 4. *Amino acid composition of selected thermolytic (L), chymotryptic (C) and cyanogen bromide (CNBrL) peptides from the encephalitogenic basic protein from human brain*

For abbreviations of purification procedures see Table 3.										
	L1	L2	L3	L4	L5	L6	L7	C1	C2	CNBrL
Asp	1.0	—	1.0	—	1.0	—	—	—	1.1	—
Thr	0.8	—	—	1.9	—	—	—	—	1.0	2.1
Ser	—	—	—	1.0	2.0	—	—	1.0	—	4.0
Glu	—	—	—	1.0	—	—	—	—	1.1	2.2
Pro	0.9	—	—	3.7	—	—	—	—	—	1.0
Gly	1.5	1.0	—	3.0	2.9	1.1	1.0	1.0	1.9	1.1
Ala	—	—	—	—	0.9	1.0	—	1.1	1.0	3.1
Val	—	—	—	0.4	—	—	—	—	1.0	—
Met	—	—	—	—	—	—	—	—	—	0.7†
Ile	—	0.9	—	0.3	—	—	—	—	—	—
Leu	1.0	—	—	—	—	—	—	—	0.9	1.0
Tyr	—	—	—	—	1.9	—	—	—	—	1.0
Phe	1.0	1.2	0.9	—	1.0	—	1.0	0.9	—	—
Lys	—	—	1.1	0.9	1.0	0.9	1.0	2.0	0.9	1.9
His	1.0	—	—	—	—	1.0	—	1.1	—	1.1
Arg	2.0	1.0	—	2.0*	0.9	—	—	—	—	1.9
<i>m</i> _{Asp}	0.43	0.38	0.49	0.62	0.23	0.72	0.65	0.75	0.0	0.60
<i>R</i> _{Leu}	0.37	0.84	0.39	0.23	0.42	0.19	0.49	0.39	0.68	0.01
Purification	GECE'	EC	EC	GECE'	EC	EC	EC	ECE'	ECE'	GE

* Includes methylarginine.

† As homoserine.

Table 5. *Peptides obtained after digestion of peptide P5 with thermolysin*

Peptides were isolated by electrophoresis at pH 6.5.

Peptide	<i>m</i> _{Asp}	Yield (%)	Composition
P5L1	0.0	50	Ile 1.0, Leu 1.0
P5L2	0.43	10	Asp 1.0, Thr 0.9, Ser 0.9, Ala 2.1, Met 0.8, His 2.4, Arg 1.2
P5L3	0.50	40	Asp 1.1, Thr 1.0, Pro 1.1, Gly 1.2, Leu 1.0, Phe 1.0, His 1.5, Arg 2.4

acid (see above), was partially cleaved during the isolation.

The tryptic peptides within peptide P7 were overlapped by digestion of peptide P7 with thermolysin for 45 min at 38°C. The compositions and sequences of selected thermolytic peptides are presented in Table 6 and Fig. 2. Hydrolysis occurred at the *N*-terminal side of some bonds which are not normally split by this enzyme. Ambler & Meadway (1968) found that bonds where glycine and histidine supply the amino group can be split by thermolysin. The sequence was confirmed by isolation of chymotryptic peptides from the whole protein which corresponded to residues 45–53 and 54–64 inclusive.

The next region of the protein was overlapped with peptide P9 and peptide L4 which contain the unusual amino acid methylarginine. The sequences of the *N*-terminal regions of peptides L4 and P9 were determined with the dansyl-Edman procedure. Aminopeptidase O released isoleucine and valine in approximately equimolar amounts

from peptide L4. Carboxypeptidase A liberated leucine from peptide P9. Digestion of peptide L4 with trypsin was used to show the presence of a small proportion of unmethylated arginine at position 107 (Baldwin & Carnegie, 1971b).

Peptides P1, P2 and P3 all came from the region around the sole tryptophan residue in the protein. Their immunological properties have been studied (Lennon *et al.* 1970) and peptides P1 and P2 were almost as encephalitogenic as the whole protein. The *N*-terminal sequences of these peptides were determined by the dansyl-Edman procedure and supported by digestion with aminopeptidase O. The amino acid sequence of this region of the molecule was previously studied in detail (Carnegie, 1969b). The sequence of peptides P1 and P2 is summarized in Fig. 3. Peptide P3 arises from the fission of a Phe-Ser bond at residues 114–115 and an Asp-Tyr bond at 133–134.

Evidence for the order of the remaining tryptic peptides was obtained from thermolytic, chymo-

Fig. 2. Amino acid sequence of larger peptic peptides which were digested with thermolysin (L) or chymotrypsin (C).

Table 6. *Selected peptides obtained after digestion of peptide P7 with thermolysin*

Peptides were isolated by electrophoresis at pH 6.5 and in some cases (*) further purified by chromatography in solvent BAWP.

Peptide	m_{Asp}	R_{Leu}	Yield (%)	Composition
P7L1*	0.50	0.32	29	Asp 1.0, Pro 1.0, Gly 3.0, Ala 1.0, Phe 0.9, Lys 1.0, Arg 2.1
P7L2*	0.59	0.07	37	Asp 1.1, Thr 1.0, Ser 1.9, Pro 1.2, Gly 2.1, Ala 0.9, Lys 1.1, His 1.8, Arg 1.0
P7L3*	0.42	0.77	25	Ala 1.1, Tyr 1.0, His 1.0
P7L4	0.20	0.10	12	Asp 2.2, Thr 1.0, Ser 2.2, Glu 3.4, Pro 2.1, Gly 2.2, Leu 0.9, Lys 0.8, His 1.2, Arg 1.0
P7L5	-0.12	0.10	12	Asp 2.2, Thr 1.0, Glu 2.2, Pro 1.0, Gly 1.4, His 0.8, Arg 1.0
P7L6*	0.36	1.10	40	Val 1.1, Phe 1.1, His 1.0

tryptic and peptic peptides, and the results are presented in Tables 3, 4 and 7 and in Figs. 2 and 3.

The proposed sequence for the encephalitogenic basic protein from human brain is shown in Fig. 4.

DISCUSSION

The amino acid sequence of the encephalitogenic basic protein from human myelin is the first sequence of a protein which is unique to brain and which is a component of membranes. There was no evidence in the sequence for heterogeneity of the type associated with immunoglobulins, nor was there any variation in the protein isolated from different adult humans. Microheterogeneity as a result of methylation of an arginine residue was found and is discussed in the following paper (Baldwin & Carnegie, 1971b). Eylar (1970) has published a preliminary account of the sequence of the bovine and human proteins and Shapira *et al.* (1971) have compared the region from residues 45 to 89 in several species. The differences in sequence are typical of those found between species for enzymic and respiratory proteins. The rat is unusual in having two basic proteins in myelin from its central nervous system (Martenson, Deibler & Kies, 1970); one is the

same size as the human and bovine proteins and the second appears to contain approximately 30 fewer amino acids.

There is disagreement between Eylar's (1970) sequence of the human protein and that presented here (Fig. 4) at five residues. He has an additional alanine residue between Ser-2 and Gln-3 and a different sequence (Ala-Gln-Gly-His-) with an additional glutamine between Lys-75 and Arg-79. Shapira *et al.* (1971) find that glutamine is present in the latter position in the bovine protein and that there is a deletion in this position in the human, rabbit and monkey proteins. Shapira *et al.* (1971) claimed that the sequence in this region is Ser-Gly-His but their evidence came from digestion of peptides with leucine aminopeptidase. In this paper the assignment His-Gly rather than Gly-His was based initially on the dansyl-Edman procedure and was confirmed by the subtractive Edman technique on peptides T25 and P7L5 (Fig. 2).

At residues 83 and 84 there is disagreement on the assignment of the amide group, which Eylar (1970) and Shapira *et al.* (1971) assign to asparagine and which is shown in this paper to be in glutamine. The observation on the susceptibility of the bond between residues 84 and 85 to acid hydrolysis (see

Fig. 3. Thermolytic (L), peptic (P), chymotryptic (C) and cyanogen bromide (CNBrL) peptides used to overlap tryptic (T) peptides. Broken lines show compositions, solid lines indicate established sequence. Residue identified by dansyl-Edman method is shown by $\rightarrow$. See Fig. 2 for characterization of peptides P5, P7 and P11.

Table 7. *Selected peptides obtained after digestion of peptide P11 with chymotrypsin*

Peptides were isolated by electrophoresis at pH 3.5 and in some cases (*) further purified by chromatography in solvent BAWP.

Peptide	m_{Lys}	R_{Leu}	Yield (%)	Composition
P11C1	0.36	0.43	37	Ser 2.0, Pro 1.0, Gly 1.0, Met 1.0
P11C2*	0.80	1.03	8	Leu 1.3, Phe 0.9, Lys 1.0
P11C3*	0.87	0.34	11	Asp 1.2, Ser 1.3, Gly 2.5, Leu 1.0, Lys 1.0, Arg 2.2
P11C4	1.08	0.31	30	Ala 1.0, Arg 2.0

above) supports the assignment of aspartate rather than asparagine to residue 84. Otherwise there is now complete agreement in the remainder of the sequence, as Eylar (1970) has now revised the earlier sequences for the tryptophan region (Eylar &

Hashim, 1968) and for the C-terminal end of the protein (Hashim & Eylar, 1969).

In earlier studies on the encephalitogenic protein it was shown that polypeptides retained encephalitogenic activity (Lumsden, Robertson & Blight,

Fig. 4. Proposed amino acid sequence of the encephalitogenic basic protein from myelin of human brain. The main encephalitogenic determinant is underlined.

1966; Carnegie & Lumsden, 1967). These polypeptides were shown to be derived from the protein as a consequence of hydrolysis by a proteinase during acid extraction of acetone-dried powders (Nakao, Davis & Einstein, 1966; Carnegie *et al.* 1967). Lennon *et al.* (1970) have isolated the main encephalitogenic determinant from the human protein. It was shown to be present in a peptide of 15 amino acids from around the sole tryptophan residue (residues 112 to 126) in the protein. Westall, Robinson, Caccam, Jackson & Eylar (1971) have synthesized the corresponding determinant. This peptide of only eleven amino acids (Fig. 4) is almost as potent in inducing experimental autoimmune encephalomyelitis as the whole protein (Lennon *et al.* 1970). The juxtaposition of the phenylalanine and the tryptophan residues is reminiscent of the hydrophobic binding site for aromatic anions in serum albumin, the sequence of which has been determined by Swaney & Klotz (1970). Carnegie (1971) has drawn attention to the resemblance in structure between the encephalitogenic determinant and the requirements for a receptor site for 5-hydroxytryptamine in a neuronal membrane in brain which was calculated by Smythies, Benington & Morin (1970) from a study of antagonists of 5-hydroxytryptamine. An immune response to 5-hydroxytryptamine receptors in certain regions of the central nervous system could explain the localized clinical symptoms found in experimental autoimmune encephalomyelitis (Carnegie, 1971). Also oligodendroglial cells are known to respond to 5-hydroxytryptamine, and thus probably contain a receptor, since 5-hydroxytryptamine was shown by Woolley (1962) to affect contractability of the cells in culture. E. J. Field & E. A. Caspary (personal communication) showed that 5-hydroxytryptamine inhibited the reaction of the encephalitogenic basic protein with sensitized lymphocytes (Field & Caspary, 1970).

When compared on a molar basis with the tryptophan region, other parts of the basic protein such as the *N*-terminal region (Carnegie *et al.* 1970) and the region peptide (P7) studied by Kibler *et al.* (1969), are much weaker encephalitogens in guinea pigs (Lennon *et al.* 1970).

What is the function of the basic protein in myelin? Is it merely part of the structural cement that binds the myelin sheath or has it a more active role? Experiments by Dickinson, Jones, Aparicio & Lumsden (1970) suggest that the protein is located in the intraperiod line in the lamellar structure of myelin, i.e. between the outer surfaces of the apposed plasma membranes from the oligodendroglial cell (Le Baron, 1969). Yet when the basic protein disappears from myelin the lamellar structure is retained except that the intraperiod line takes on a dense appearance in the electron microscope. In addition in the developing rat brain there appears to be a second myelin-like fraction lacking the basic protein (Agrawal *et al.* 1970). Thus a form of myelin can exist without the basic protein; yet it appears from studies by Hallpike, Adams & Bayliss (1970) that damage to the protein is accompanied by demyelination. They found that there was extensive digestion of the basic protein by acid proteinase in the active plaques of multiple sclerosis and concluded that the basic protein is the most vulnerable component of myelin.

Does the basic protein have a more active role in the maintenance of the oligodendroglial cell-neurone complex? One problem in myelinogenesis is: what is the signal from the neuron which causes the glial cell to wrap around its axon (Lumsden, 1968; Field, Raine & Hughes, 1968)? Perhaps the basic protein is synthesized and extruded by the neuron and acts as an 'organizer molecule' which controls the wrapping process. As Kibler *et al.* (1969) and Eylar (1970) have suggested, the clustering of basic residues in the *N*-terminal half of

the molecule would facilitate interaction with acidic phospholipids in the plasma membrane of the glial cell.

Myelin is often thought of as an insulator analogous to the rubber round an electric cable. However, the finding of the basic protein in the intra-period line (Dickinson *et al.* 1970) suggests that a more appropriate analogy would be that of a coaxial, waxed-paper, spiral-wound type of condenser where the basic protein would represent one foil and the remaining cytoplasm of the glial cell the other foil, and with the lipids and proteolipid of the glial plasma membrane representing the waxed paper. Although it is well known that myelinated peripheral nerve axons have a small capacitance this property is usually ignored in discussions on the structure and function of the central nervous system. Since the basic protein of myelin of the central nervous system is quite different from that of peripheral myelin (Bencina *et al.* 1969) and since in the central nervous system a single glial cell can wrap around more than one axon (Bunge, 1968) it should not be assumed that myelinate axons in the central nervous system will function in a manner identical with peripheral axons. Within the non-polar environment of myelin the numerous histidine, serine and aspartic acid residues of the protein could be involved in a 'charge-relay system' analogous to that in the active site of chymotrypsin (Blow, Birktoft & Hartley, 1969). In addition Fernandez-Moran (1959) has suggested that lamellar membranes such as myelin may function as semi-conductors. A knowledge of the role of the basic protein in myelin could be important in understanding the relationship between neurons and oligodendroglial cells in normal and diseased brain.

I acknowledge financial support from the National Multiple Sclerosis Society, New York, The Australian Research Grants Committee and the Multiple Sclerosis Society of Victoria. Mr S. Modrich, Mrs M. Poutet, Mrs B. Taylor and Mr G. G. Luxford are thanked for technical assistance at various stages in the project.

REFERENCES

- Agrawal, H. C., Banik, N. L., Bone, A. H., Davison, A. N., Mitchell, R. F. & Spohn, M. (1970). *Biochem. J.* **120**, 635.
- Alvord, E. C. (1968). In *The Central Nervous System*, p. 52. Ed. by Bailey, O. T. R. & Smith, D. E. Baltimore: Williams and Wilkins Co.
- Ambler, R. P. (1963). *Biochem. J.* **89**, 349.
- Ambler, R. P. (1967). In *Methods in Enzymology*, vol. 11, p. 436. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
- Ambler, R. P. & Meadway, R. J. (1968). *Biochem. J.* **108**, 893.
- Baldwin, G. S. & Carnegie, P. R. (1971a). *Science, N.Y.*, **171**, 579.
- Baldwin, G. S. & Carnegie, P. R. (1971b). *Biochem. J.* **123**, 69.
- Bencina, B., Carnegie, P. R., McPherson, T. A. & Robson, G. S. M. (1969). *FEBS Lett.* **4**, 9.
- Blow, D. M., Birktoft, J. J. & Hartley, B. S. (1969). *Nature, Lond.*, **221**, 337.
- Bunge, R. P. (1968). *Physiol. Rev.* **48**, 197.
- Carnegie, P. R. (1969a). *Biochem. J.* **111**, 240.
- Carnegie, P. R. (1969b). *Nature, Lond.*, **223**, 958.
- Carnegie, P. R. (1971). *Nature, Lond.*, **229**, 25.
- Carnegie, P. R., Bencina, B. & Lamoureux, G. (1967). *Biochem. J.* **105**, 559.
- Carnegie, P. R. & Lumsden, C. E. (1967). *Immunology*, **12**, 133.
- Carnegie, P. R., McPherson, T. A. & Robson, G. S. M. (1970). *Immunology*, **19**, 55.
- Caspary, E. A. & Field, E. J. (1965). *Ann. N.Y. Acad. Sci.* **122**, 182.
- Chao, L.-P. & Einstein, E. R. (1970). *J. Neurochem.* **17**, 1121.
- Dickinson, J. P., Jones, K. M., Aparicio, S. R. & Lumsden, C. E. (1970). *Nature, Lond.*, **227**, 1133.
- Eng, L. F., Chao, F. C., Gerstl, B., Pratt, D. & Tavaststjerna, M. G. (1968). *Biochemistry, Easton*, **7**, 4455.
- Einstein, E. R., Csejtey, J., Davis, W. & Rauch, H. C. (1965). *Proc. 8th int. Cong. Neurology*, vol. 4, p. 137.
- Eylar, E. H. (1970). *Proc. natn. Acad. Sci. U.S.A.* **67**, 1425.
- Eylar, E. H. & Hashim, G. A. (1968). *Proc. natn. Acad. Sci. U.S.A.* **61**, 644.
- Fernandez-Moran, H. (1959). In *Biophysical Science—A Study Program*, p. 319. Ed. by Oncley, J. L. New York: John Wiley and Sons Inc.
- Field, E. J., Raine, C. S. & Hughes, D. (1968). *J. neurol. Sci.* **8**, 129.
- Field, E. J. & Caspary, E. A. (1970). *Lancet*, ii, 1337.
- Gray, W. R. (1967). In *Methods in Enzymology*, vol. 11, p. 469. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
- Hallpike, J. F., Adams, C. W. M. & Bayliss, O. B. (1970). *Histochem. J.* **2**, 323.
- Hartley, B. S. (1970). *Biochem. J.* **119**, 805.
- Hashim, G. A. & Eylar, E. H. (1969). *Archs Biochem. Biophys.* **135**, 324.
- Kibler, R. F., Shapira, R., McKneally, S., Jenkins, J., Selden, P. & Chou, F. (1969). *Science, N.Y.*, **164**, 577.
- Kies, M. W. (1965). *Ann. N.Y. Acad. Sci.* **122**, 161.
- LeBaron, F. N. (1969). In *Structural and Functional Aspects of Lipoproteins in Living Systems*, p. 201. Ed. by Tria, E. & Scanu, A. M. New York: Academic Press Inc.
- Lennon, V. A., Wilks, A. V. & Carnegie, P. R. (1970). *J. Immun.* **105**, 1223.
- Lumsden, C. E. (1968). In *The Structure and Function of Nervous Tissue*, vol. 1, p. 68. Ed. by Bourne, G. H. New York: Academic Press Inc.
- Lumsden, C. E., Robertson, D. M. & Blight, R. (1966). *J. Neurochem.* **13**, 127.
- Martenson, R. E., Deibler, G. E. & Kies, M. W. (1970). *Biochim. biophys. Acta*, **200**, 353.

- Milstein, C. (1968). *Biochem. J.* **110**, 652.
- Nakao, A., Davis, W. J. & Einstein, E. R. (1966). *Biochim. biophys. Acta*, **130**, 163.
- Offord, R. E. (1966). *Nature, Lond.*, **211**, 591.
- Oshiro, Y. & Eylar, E. H. (1970). *Archs Biochem. Biophys.* **138**, 606.
- Palmer, F. B. & Dawson, R. M. C. (1969). *Biochem. J.* **111**, 629.
- Pink, J. R. L., Buttery, S. H., De Vries, G. M. & Milstein, C. (1970). *Biochem. J.* **117**, 33.
- Piszkiewicz, D., Landon, M. & Smith, E. L. (1970). *Biochem. biophys. Res. Commun.* **40**, 1173.
- Shapira, R., McKneally, S. S., Chou, F. & Kibler, R. F. (1971). *J. biol. Chem.* (in the Press).
- Smith, I. (1960). *Chromatographic and Electrophoretic Techniques*, p. 95. London: W. Heinemann (Medical Books) Ltd.
- Smythies, J. R., Benington, F. & Morin, R. D. (1970). *Neurosci. Res. Prog. Bull.* **8**, 117.
- Swaney, J. B. & Klotz, I. M. (1970). *Biochemistry, Easton*, **9**, 2570.
- Waley, S. G. & Watson, J. (1953). *Biochem. J.* **55**, 328.
- Wang, S.S. & Carpenter, F. H. (1965). *J. biol. Chem.* **240**, 1619.
- Westall, F. C., Robinson, A. B., Caccam, J., Jackson, J. & Eylar, E. H. (1971). *Nature, Lond.*, **229**, 22.
- Woolley, D. W. (1962). *The Biochemical Bases of Psychoses*. New York: John Wiley and Sons Inc.