Proceedings of the National Academy of Sciences
Vol. 67, No. 3, pp. 1425–1431, November 1970

Amino Acid Sequence of the Basic Protein of the Myelin Membrane*

E. H. Eylar†

THE SALK INSTITUTE, SAN DIEGO, CALIFORNIA 92112

Communicated by Martin D. Kamen, August 10, 1970

Abstract. The amino acid sequences of the encephalitogenic basic protein, A1, from bovine and human myelin are similar, differing by only 11 residues. The sequence reveals that while basic residues are spread randomly over most of the polypeptide chain, several regions (8–10 residues) exist that are nonpolar in character. The bovine protein has 170 residues with molecular weight 18,400. The human protein, which has an additional His-Gly sequence, contains 172 residues. The major encephalitogenic determinant (tryptophan region) of the bovine protein differs from the human only by a lysine to arginine substitution. The structural features of the A1 protein are discussed, with special reference to its role in stabilization of the myelin membrane, and its relation to multiple sclerosis.

Myelin of the central nervous system appears to be derived from the plasma membrane of the oligodendroglial cell[1]; as such it represents a relatively simple prototype of a biological membrane since it contains only two major protein components, proteolipid (50%) and a basic protein[2] (30%) referred to as the A1 protein. When injected at low levels in complete Freund's adjuvant, the A1 protein[2–4] induces experimental allergic encephalomyelitis (EAE), a demyelinating, paralytic disease[5] which has been induced in many animal species including man, monkey, rabbit, guinea pig, and rat. Over the last 20 years,[6,7] the similarity between EAE and human demyelinating diseases, such as multiple sclerosis, has stimulated interest in EAE as a model for elucidation of delayed hypersensitive phenomena,[8] particularly as it relates to human autoimmune diseases.

Previous work from this laboratory[2,9,10] has focused on the isolation and characterization[11,12] of the A1 protein from the brain and spinal cord of several species. It has been demonstrated that the induction of EAE in guinea pigs is due to a unique segment of the A1 molecule, the nine residues surrounding the single tryptophan residue.[13–16] The 9-residue peptide and numerous active and inactive analogues have been synthesized[16,17] in an effort to specify the essential residues. This report extends previous studies by describing the complete amino acid sequences of the bovine and human A1 proteins.

Experimental Procedure. Preparation of peptides: The peptides used to evaluate the amino acid sequence were obtained from tryptic and peptic digestion of the bovine A1 protein. For pepsin digestion, 3.5 g of A1 protein were incubated for 2 hr at 37°C and separated into peptide families by Cellex-P chromatography, as described.[14,18] Tryptic peptides[18] were separated into groups by Dowex 50-X2 chromatography, using a gradient of ammonium acetate. In both cases, the peptide groups were further re-

solved into individual peptides by gel filtration followed by paper electrophoresis at pH 4.6 and/or paper chromatography in butanol–acetic acid–water (4:1:5). The peptide representing the carboxy-terminal portion of the A1 protein (54 residues) was obtained as described[12] by oxidation and cleavage of the carboxy-tryptophyl bond with *N*-bromsuccinimide. In all cases, the peptides were mapped as described above. When only a single spot was observed with ninhydrin, the peptide was considered homogeneous. When peptides were to be tested for encephalitogenic activity, 1–3 mg was used for peptide mapping; the peptide was then eluted from the paper with 5% NH_4OH. Most peptides could be resolved into a single component by this procedure. The composition of each peptide was determined with a Beckman amino acid analyzer after hydrolysis in 6 N HCl for 24 and 48 hr. The sequence of the purified peptides from bovine A1 protein was determined by the direct and indirect Edman procedures.[19,20]

Carboxypeptidases A and B (Worthington) and leucine aminopeptidase were also used. Amidation of glutamic and aspartic acid residues was evaluated from the direct Edman procedure and the electrophoretic mobility of the peptide on paper.

For the human protein, the sequence was determined from peptides eluted from the peptide map of the tryptic, chymotryptic, or thermolysin digestion mixtures. The sequence was assumed to match that of the analogous bovine peptide if the composition was identical or differed by only one residue. If the substitution was not obvious, the correct sequence was determined as described above.

Results and Discussion. The complete amino acid sequence of the A1 protein from bovine spinal cord myelin is shown in Fig. 1. The sequence was established with peptides obtained by tryptic and peptic hydrolysis at positions shown by the arrows. The determination of the sequence was aided by isolating the 54-residue peptide[12] from the carboxyterminal region after cleavage of the C-tryptophyl linkage with *N*-bromsuccinimide. It is of interest that most of the glutamic residues are amidated, whereas the aspartic are not. The single tryptophan residue is located at position no. 116 from the NH_2-terminus; the two methionine residues are near the respective ends. The basic residues are fairly randomly distributed over the entire polypeptide chain, and in this regard the sequence differs from that of histone IV where the basic residues are restricted to only half of the molecule.[21] No similarity in sequence exists between this protein and histone IV. However, in both cases the NH_2-terminal residue is acetylated; in this case as *N*-acetylalanine and in the histone as *N*-acetylserine.

Microheterogeneity: The amidation of glutamic acid contributes to the basic character of the A1 protein. Martenson *et al.*[22] have observed a microheterogeneity in the basic protein on electrophoresis at pH 10.5. We find, however, that approximately 90% of the A1 protein migrates as a single component at pH 10.5; this phenomenon, therefore, appears trivial, probably arising from deamination of glutamine residues during extraction under acidic conditions. However, for histone IV[21], lysine residues modified by methylation and acetylation account for polymorphism. Although none of these derivatives were observed in the A1 protein, a modification at arginine residue no. 107 could possibly account for some microheterogeneity. The peptic peptides containing this residue showed a peak between arginine and NH_3 on the basic column of the amino acid analyzer, and were trypsin resistant.

Phylogenetic variation: The sequences of the A1 proteins from human and bovine sources are shown in Fig. 1, and confirm earier studies that showed the human and bovine proteins to be very similar on the basis of physical, chemical,

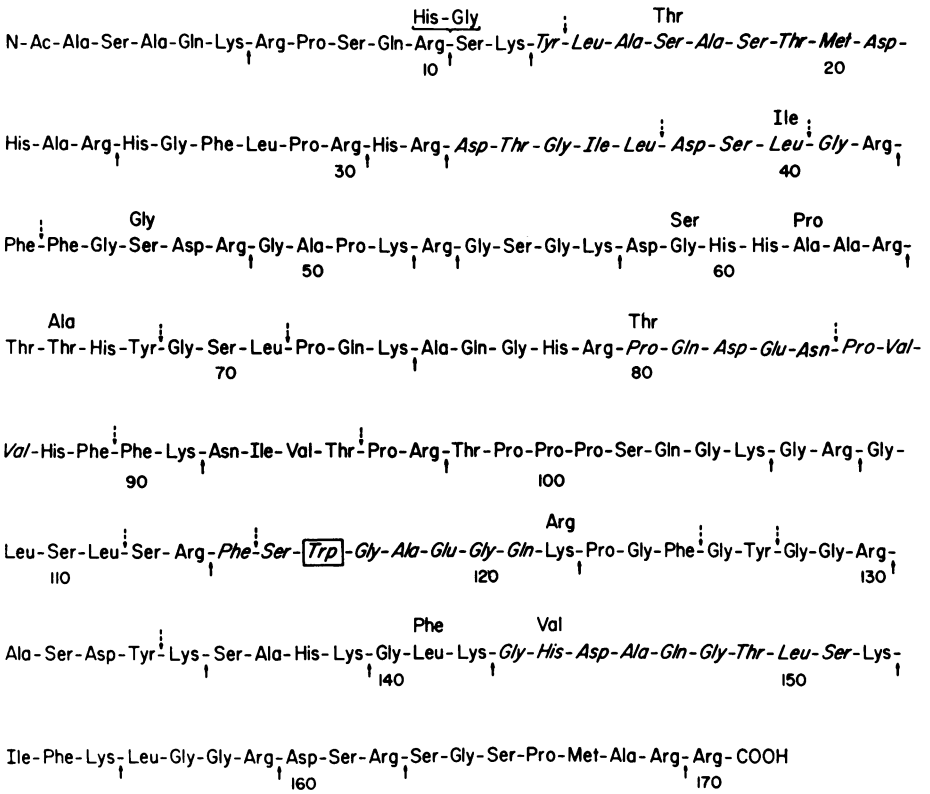

FIG. 1. The amino acid sequence of the bovine A1 protein. Substitutions for the human A1 protein are shown above the appropriate position of the bovine sequence. Linkages sensitive to trypsin and pepsin are shown by dark and dotted arrows, respectively. The five main sections lacking basic residues are shown with slanted letters. Between positions 10 and 11, the His-Gly dipeptide occurs in the human protein.

and immunologic properties.[10] There are 11 substitutions between the bovine and human proteins; in addition, a His-Gly dipeptide is present in the human protein[20,23] between residues 10–11. Thus, in spite of the static, structural role probably played by the A1 protein as part of the myelin ultrastructure, relatively few changes are found between the bovine and human proteins. All changes between these proteins could result from single-base changes in DNA except for a valine for histidine change at no. 144 where two base changes would be required.

Because of the possibility that induction of human demyelinating disease may involve the Al protein, the amino acid sequence of the human protein assumes special interest. In view of the unfolded conformation of the A1 protein,[11] knowledge of the sequence, combined with peptide synthesis, should provide new approaches for not only clinical studies but for further elucidation of those sites involved in induction, blocking, and suppression of the disease process.[7,8]

Properties of the A1 protein: According to the amino acid analyses, there are 170 and 172 amino acid residues, respectively, in the bovine and human A1 proteins; the molecular weights are 18,395 and 18,625 respectively. Molecular

TABLE 1. *Properties of the A1 protein from myelin.*

Molecular weight		Method
Bovine	Human	
18,600–20,100	19,400–19,600	Ultracentrifugation[10,11,26]
18,200	17,400	Tryptophan analysis[10,12]
19,500	—	Gel filtration[27]
18,395	18,625	Sequence (this report)
Conformation		
Open (Axial ratio 10:1)		Viscosity,[11] proteolysis[14]
No α-helical or β-structure		Optical rotatory dispersion[10,26], circular dichroism[18]

weight values from sedimentation and chemical methods, 18,000–20,000 are in good agreement (Table 1). Earlier estimates[4,24,25] from several laboratories differed markedly, primarily because of heterogeneity caused by proteolysis, which may occur *in situ*[2] or during acid extraction.[24] One of the most interesting features of the A1 protein is its open conformation and lack of secondary structure,[10,11,26] doubtless an important factor in its immunologic interactions and *in vivo* role as a structural protein of myelin.

Disease-inducing site: Lumsden *et al.*[25] presented evidence that EAE could be induced with low molecular weight peptide fragments derived from degradation of central nervous system tissue. From many studies[14,16,17] performed on encephalitogenic peptides, both derived and synthesized, it has been established that the region of the A1 molecule surrounding the single tryptophan is the major disease-inducing site in guinea pigs:

NH_2-Phe-Ser-*Trp*-Gly-Ala-Glu-Gly-*Gln-Lys*(Arg)-COOH

The sequence prevails in the rabbit and human A1 proteins as well (Fig. 1), except that arginine replaces lysine in the human peptide.[17,28] A study[17] of several synthetic peptides, differing by only one residue, showed that the Trp——Gln–Lys relationship is essential for encephalitogenic activity, although the arginine for lysine substitution is equally active. Thus, the requirements for disease induction are highly specific; at least three residues are necessary for the immunologic response characteristic of EAE.

Number of disease-inducing site: It is of interest that the encephalitogenic determinant defined by the tryptophan region, which has a molar activity approximately equal to the A1 protein,[16] appears to be at least several hundred times more active in guinea pigs than any other region of the A1 molecule. Modification of the essential tryptophan residue with 2-hydroxy-5-nitrobenzyl bromide inactivates the molecule.[12,16] Since this reagent is specific for tryptophan,[29] it is evident that other regions of the molecule are relatively inactive. It can be concluded that only this one small region of the polypeptide chain is significantly active in disease induction in the guinea pig.

Chao and Einstein[30] obtained a large encephalitogenic peptide from the Al protein with an acid protease from brain.[24] Since this peptide was approximately half as large as the original A1 protein, and had *N*-terminal phenylalanine, it is likely that the Phe-Phe linkage (no. 90–91), a linkage susceptible to pepsin-like proteolysis,[18] was cleaved. This peptide contained the tryptophan region, and

was found to be as encephalitogenic in guinea pigs as the A1 protein. This result adds support to the contention that the tryptophan region defines the predominant encephalitogenic site.

Carnegie[23] has reported that the peptide having the first 20 residues, from *N*-acetylalanine to methionine, is slightly encephalitogenic in guinea pigs. However, on a molar basis, it is 600 times less active than the A1 protein or the tryptophan region and, thus, plays a very minor role in guinea pigs.

Kibler *et al.*[31] have isolated a peptide from spinal cord under conditions which promote autolysis by the acid protease.[24] In our laboratory, this same peptide was derived[18] from the A1 protein with pepsin, thus demonstrating its *in vivo* origin and the similarity of the acid protease and pepsin. This peptide contains 45 residues and occupies that part of the A1 protein from Phe-Phe (no. 43–44) to Phe-Phe (no. 90–91). While this peptide is encephalitogenic in rabbits, it is inactive in guinea pigs.[18] A distinction, therefore, in active regions can be made according to species; the tryptophan region being active in both guinea pigs and rabbits,[18] whereas the peptide composed of residues 43–88 is active only in rabbits. Thus, rabbits respond to two distinct regions of the polypeptide chain; guinea pigs only to one.

The sequence reported (without data) by Kibler *et al.*[32] for their peptide appears markedly different from that given here (residues 65–88). The discrepancies may result from their use of acid hydrolysis to obtain fragments; in addition, the original peptide material was reported[30] to contain 10% contamination. A slight contamination could give erroneous results in the EAE assay.

Relation to myelin structure: Evidence from numerous sources, such as electron microscopy, x-ray diffraction, and freeze-etching, suggests that the myelin substructure consists primarily of a bimolecular leaflet along the lines of the classical Davson-Danielli model.[33] Since the basic A1 protein is a major component of myelin, it is reasonable to assume it plays a major structural role. Because of its open conformation, it is ideally designed for maximal interaction with other myelin constituents, most notably the phospholipids, where it could engage in both electrostatic and nonpolar interactions. In this sense, the A1 protein fits the role conceived for proteins in the early Davson-Danielli models. It appears that the A1 molecule could stabilize the leaflet structure if spread tangentially over the negatively-charged phosphate groups of the phospholipids, particularly triphosphoinositide which has been shown to form complexes with the basic protein *in vitro*.[26] As seen from Fig. 1, the distribution of basic residues in the A1 molecule is compatible with this arrangement since the lysine, histidine, and arginine are rather randomly distributed over most of the polypeptide chain. Generally, 3–4 residues are located between the basic residues. There are five regions (Fig. 1), however, where basic residues are absent over spans of 8–10 residues; these regions contain nonpolar residues, and at least one negative charge. It is tempting to suggest that these regions more readily participate primarily in nonpolar, rather than electrostatic, interactions. It is one of these regions that contains the encephalitogenic determinant.

Although this picture seems plausible for myelin, it cannot be generalized at this time to other biological membranes because (*a*) myelin has a high lipid con-

tent and differs markedly from other membranes, such as the mitochondria and plasma membrane, where subunit structures are likely[33]; and (*b*) basic proteins are not generally found in cellular membranes. Thus, the role of the basic protein in myelin may be highly specialized, reflecting primarily the static role of myelin as an insulator of the axon in contrast to other, more dynamic, cellular membranes.

Relationship to multiple sclerosis: Although no evidence has been presented directly implicating the basic A1 protein in multiple sclerosis, EAE has long been regarded as an animal model for this disease.[5] Since the basic protein is the exclusive agent in EAE induction, it follows that the basic protein could be instrumental in induction of multiple sclerosis, possibly in combination with a virus. In this sense then, and in spite of species variations, the tryptophan region assumes special significance, since it could represent a focus for immunologic events in persons afflicted with multiple sclerosis just as it has in guinea pigs with EAE.

We thank Drs. F. Westall, S. Brostoff, J. Caccam, G. Hashim, and Mr. J. Jackson, who have made this work possible; the advice and encouragement of Dr. R. Doolittle; and the help of Mr. I. Goldman and Mrs. R. Goldman with the amino acid analyses.

* This investigation was supported in part by U.S. Public Health Service grant 1RO1 NB 08268-02, The Salk Institute, and grants to Professor M. D. Kamen, HD-10262(NIH) and GB-7033X(NSF).

† U.S. Public Health Service Career Development Awardee (1-K4-A1-0848-02).

Abbreviation: EAE, experimental allergic encephalomyelitis.

[1] Bunge, M., R. Bunge, and G. Pappas, *J. Cell Biol.*, **12**, 448 (1962).

[2] Eylar, E. H., J. Salk, G. Beveridge, and L. Brown, *Arch. Biochem. Biophys.*, **132**, 34 (1969).

[3] Roboz-Einstein, E., D. Robertson, J. DiCaprio, and W. Moore, *J. Neurochem.*, **9**, 353 (1962).

[4] Kies, M., *Ann. N.Y. Acad. Sci.*, **122**, 161 (1965).

[5] Paterson, P., *Advan. Immunol.*, **5**, 131 (1966).

[6] Kabat, E. A., A. Wolf, and A. Bezer, *J. Exp. Med.*, **85**, 117 (1947).

[7] Adams, C., and S. Leibowitz, in *The Structure and Function of Nervous Tissue*, ed. G. Bourne (Academic Press 1969), vol. 3, p. 309.

[8] Waksman, B., *Hosp. Practice*, **3**, 22 (1968).

[9] Oshiro, Y., and E. H. Eylar, *Arch. Biochem. Biophys.*, **138**, 392 (1970).

[10] Oshiro, Y., and E. H. Eylar, *Arch. Biochem. Biophys.*, **138**, 606 (1970.

[11] Eylar, E. H., and M. Thompson, *Arch. Biochem. Biophys.*, **129**, 468 (1969).

[12] Eylar, E. H., and G. Hashim, *Arch. Biochem. Biophys.*, **131**, 215 (1969).

[13] Hashim, G., and E. H. Eylar, *Arch. Biochem. Biophys.*, **129**, 635 (1969).

[14] Hashim, G., and E. H. Eylar, *Arch. Biochen. Biophys.*, **129**, 645 (1969).

[15] Eylar, E. H., and G. Hashim, *Proc. Nat. Acad. Sci. USA*, **61**, 644 (1968).

[16] Eylar, E. H., J. Caccam, J. Jackson, F. Westall, and A. Robinson, *Science*, **168**, 1220 (1970).

[17] Westall, F., A. Robinson, J. Caccam, J. Jackson, and E. H. Eylar, *Nature*, in press.

[18] Eylar, E. H., F. Westall, J. Caccam, S. Brostoff, and G. Hashim, *J. Biol. Chem.*, submitted.

[19] Blomback, B., M. Blomback, P. Edman, and B. Hessel, *Biochim. Biophys. Acta*, **115**, 371 (1966).

[20] Hashim, G., and E. H. Eylar, *Arch. Biochem. Biophys.*, **135**, 324 (1969).

[21] DeLange, R., D. Fambrough, E. Smith, and J. Bonner, *J. Biol. Chem.*, **25**, 319 (1969).

[22] Martenson, R , G. Diebler, and M. Kies, *J. Biol. Chem.*, **244**, 4261 (1969).

[23] Carnegie, P., *Biochem. J.*, **111**, 240 (1969).

[24] Nakao, A., W. Davis, and E. Roboz-Einstein, *Biochim. Biophys. Acta*, **130**, 163 (1966).

[25] Lumsden, C., D. Robertson, and R. Blight, *J. Neurochem.*, **13**, 127 (1966).

[26] Palmer, F., and R. Dawson, *Biochem. J.*, **111**, 629 (1969).

[27] Chao, L., and E. R. Einstein *J. Chromatogr.*, **42**, 485 (1969).
[28] Carnegie, P., *Nature*, **223**, 958 (1969).
[29] Barman, T., and D. Koshland, *J. Biol. Chem.*, **242**, 5771 (1967).
[30] Chao, L., and E. Einstein, *J. Biol. Chem.*, **243**, 6050 (1968).
[31] Kibler, R., and R. Shapira, *J. Biol. Chem.*, **243**, 281 (1968).
[32] Kibler, R., R. Shapira, S. McKneally, J. Jenkins, P. Selden, and F. Chow, *Science*, **164**, 577 (1969).
[33] Green, D. E., and A. Tzagoloff, *J. Lipid Res.*, **7**, 587 (1966).