Analogous amino acid sequences in myelin proteolipid and viral proteins

Shyh-Yu Shaw, Richard A. Laursen* and Marjorie B. Lees[1]

Department of Chemistry, Boston University, Boston, MA 02215 and [1]*E.K. Shriver Center, Waltham, MA 02254, USA*

Received 22 August 1986; revised version received 3 September 1986

Computer analysis of the intrinsic membrane protein, myelin proteolipid, shows strong sequence similarities between the putative extramembrane segments of the proteolipid protein and a number of viral proteins, several of which infect humans. These similarities are even more striking than those reported previously between viral proteins and the encephalitogenic myelin basic protein (MBP). These findings, along with other reports of molecular mimicry by viruses, suggest that immunological cross-reactions between virus-induced antibodies or T-cells and analogous antigenic determinants (epitopes) in myelin proteolipid could be involved in the pathophysiology of multiple sclerosis or post-infectious demyelinating syndromes.

Proteolipid protein *Multiple sclerosis* *Viral protein* *Myelin basic protein* *Sequence homology* *Amino acid sequence*

1. INTRODUCTION

Current evidence suggests that the etiology of multiple sclerosis and other demyelinating diseases involves a combination of viral and autoimmune factors [1]. A particularly well studied model for multiple sclerosis is EAE which can be induced by immunization of laboratory animals with CNS tissue or with MBP, one of the major myelin proteins [2]. Recently it has been reported that MBP shows homology with certain viral proteins [3]. Furthermore, immunization with peptides having regions common to MBP and hepatitis B virus DNA polymerase showed histological EAE in rabbits [4]. However, it has long been recognized that components other than MBP contribute to the encephalitogenic response to CNS tissue [5].

* To whom correspondence should be addressed

Abbreviations: MBP, myelin basic protein; EAE, experimental allergic encephalomyelitis; CNS, central nervous system; MLV, murine leukemia virus; AIDS, acquired immune deficiency syndrome; HTLV-III/LAV, human T-lymphotropic retrovirus

The most abundant protein of CNS myelin is myelin proteolipid protein, which is embedded in the myelin membrane. Its amino acid sequence has been reported recently [6,7], and two different models have been proposed for its conformation in the myelin membrane [8,9]. Immunization of rabbits with the proteolipid apoprotein leads to the development of a chronic, progressive or relapsing form of EAE [10]. However, the epitopes in myelin proteolipid which cause the encephalitogenic response are not known.

In the present study we compared decapeptide sequences from proposed [8,9] extramembrane hydrophilic regions of myelin proteolipid with various viral proteins by computer analysis, with the aim of searching for similarities, such as found between MBP and viral proteins, that might shed light on the origin of demyelinating diseases.

2. METHODS

Decapeptides were chosen for comparison with the extramembrane segments of the proteolipid protein, the rationale being that decapeptides are

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

more than sufficient to induce an immune response [11] and that epitopes of membrane proteins are generally located at extramembrane hydrophilic regions. Comparisons and statistical scorings were made using the database (release 6.8 of 28 August 1985, containing 3309 sequences) and programs of the Protein Identification Resource of the National Biomedical Research Foundation. The search program compared every decapeptide in bovine brain myelin proteolipid with every decapeptide in the database. The calculation of similarity scores was accomplished using the Mutation Data Matrix [12] program. For each proteolipid decapeptide, 709253 comparisons were made and the highest scoring segments and their standard deviations were printed out. We selected those viral decapeptides, similar to extramembrane hydrophilic segments of myelin proteolipid, with scores higher than 30, and which also exceed the average by 5 standard deviations; i.e., the probability that the match is due to random chance is $<3 \times 10^{-7}$. A VAX 11/750 computer was used for searching and calculations.

3. RESULTS AND DISCUSSION

Sequence alignments with scores of 30 or greater were observed between proteolipid and 249 viral decapeptides, excluding identities seen with proteins from similar strains of virus. These include decapeptides from a variety of types of viruses, such as adenoviruses and Epstein-Barr, influenza and measles viruses. Only those with a score of 38 or greater, or which show at least 60% identity are shown in table 1. The degree of similarity between proteolipid and viral proteins is significantly greater than that observed [3] between MBP and viral proteins. Furthermore, although only 10% of the proteins in the protein sequence database are viral proteins, they account for over 30% of the similarities.

It is interesting and perhaps significant that many of the strong similarities lie between residues 135 and 150 in the proteolipid. This region is predicted (Shaw, S.-Y., unpublished) to have a strong tendency to form an amphipathic helix (hydrophobic on one side and polar on the other), and thus might be regarded as a prime candidate for an antigenic site. Likewise, many similarities are found with the N-terminus of the proteolipid, a region which has also been predicted to be an amphipathic helix [8].

The two proposed models [8,9] for the organization of the proteolipid in the myelin membrane are

Table 1

Comparison of myelin proteolipid decapeptide sequences with those of certain viral proteins

Protein	Virus	Residues	Sequence	Score
Proteolipid		**1– 14**	**G L L E C C A R C L V G A P**	
Late 100 kDa	Adenovirus	561– 570	G L L E C H C R C N	(43)
Structural polyprotein	Sindbis virus	782– 791	V L M R C C S C C L	(47)
Coat protein gp37	Rous sarcoma	201– 210	L L V V C L P C L L	(38)
Middle T antigen	Mouse polyomavirus	137– 146	K C D A R C L V L G	(34)
Probable DNA polymerase	Woodchuck hepatitis	767– 776	A C L A R C W T G A	(31)
Ribonucleotide reductase	Epstein-Barr	409– 418	C L P R C L V N A P	(43)
Hypothetical BHLF1	Epstein-Barr	19– 28	C P P P C L P G A P	(40)
Proteolipid		**31– 57**	**F C G C G H E A L T G T E K L I E T Y F S K N Y Q D Y**	
Probable *gag-pol* polyprotein	Human adult T-cell leukemia	1038–1047	F T H C G Q T A L T	(36)
P3	Human influenza A	522– 531	Q G T E K L T I T Y	(33)

Table 1 (continued)

Protein	Virus	Residues	Sequence	Score
Hypothetical protein A-198	Mouse mammary tumor	183– 192	L I E H Y S A K T Y	(37)
Genome protein	Human rhinovirus	917– 926	V Y Y S T Y Y R K Y	(38)
Proteolipid		**96– 108**	**V R Q I F G D Y K T T I C**	
E1-glycoprotein	Murine coronavirus	136– 145	V R P I I E D Y H T	(32)
Envelope	Human adult T-cell leukemia	18 – 27	I F G D Y S P S C C	(44)
Proteolipid		**118– 127**	**G G Q K G R G S R G Q H**	
Probable nuclear antigen	Epstein-Barr	326– 335	G G G R G R G G S G	(34)
Hypothetical ORF-1	AIDS	178– 187	Q K T K G H R G S H	(32)
Proteolipid		**135– 153**	**E R V C H C L G K W L G H P D K F V G**	
Hypothetical BGLF1	Epstein-Barr	34– 43	N R V P N C E G A W	(40)
Neuraminidase	Human influenza A (strain A/PR/8/34)	289– 298	M C V C R D N W H G	(43)
Neuraminidase	Human influenza A (strain A/NT/60/68)	289– 298	C I C R D N W K G S	(39)
Middle T antigen	Hamster polyomavirus	150– 159	D C F A L W F G L P	(41)
Small T antigen	Mouse polyomavirus	152– 161	F C Y M Q W F G T P	(39)
Probable L1 protein	Cottontail rabbit papillomavirus	2– 11	A V W L S T Q N K F	(38)
Probable DNA polymerase	Squirrel hepatitis	623– 632	K W W G H T L H F M	(38)
Genome polyprotein	Foot and mouth disease	1140–1149	W I A S E E K F V T	(39)
Core antigen	Duck hepatitis B	214– 223	W L S T P E K Y R G	(47)
Proteolipid		**224– 233**	**L S I C K T A E F Q M T**	
34 kDa	Simian 11 rotavirus	282– 291	L E I C K K L L F Q	(36)
33 kDa	Bovine rotavirus	282– 291	L D I C K K L L F Q	(36)
Kinase-related transforming protein	Rous sarcoma	399– 408	V C K V A D F G L A	(39)

All of the peptides shown have alignment scores of 38 or greater and/or have at least 6 identities. Amino acids which are identical with the corresponding residues in the proteolipid are underlined

similar in that they depict the protein as threading through the lipid bilayer several times with the polar segments external to the bilayer; they differ, however, in the sidedness of some of the extramembrane segments. While both models put residues 1–10 and 220–235 on the extracellular face of the membrane, the model of Laursen et al. [8] puts residues 35–60 on the extracellular face and 90–150 on the cytoplasmic face, whereas the model of Stoffel et al. [9] predicts the opposite orientation for the latter two segments. Regardless of which (if either) model is correct, similarities between viral proteins and proposed extracellular and cytoplasmic segments of the proteolipid are seen.

It has been proposed [3,4] that EAE-like diseases can arise when virus-evoked antibodies and/or sensitized T-cells cross-react with homologous amino acid sequences (epitopes) in MBP. However, since MBP is located entirely on the cytoplasmic face of the membrane and is therefore within the oligodendroglial cell, it would be expected to be relatively inaccessible to immune surveillance. On the other hand, the proteolipid has potential antigenic sites not only on the cytoplasmic face but also on the outer surface of the membrane. Recognition of these sites by antiviral antibodies or sensitized cells could induce a cascade of immunological events leading to cell destruction. In the process, MBP would be released, which might result in further inflammatory consequences.

The similarities between MBP and virus proteins have been discussed [3,4] previously in terms of 'homology', which implies a common ancestor for the proteins. When comparing myelin proteins with viral proteins, it is difficult to make a compelling case for homology, since the similarities fall off after a few residues. While it is not uncommon to encounter five or six identical residues in a stretch of ten (see table 1), extended regions of similar sequence are encountered only rarely. For example, although a 55% identity is seen between the first 20 residues of the proteolipid and the late 100 kDa protein of adenovirus 2 or 5, i.e.,

```
Proteolipid (1–20)      GLLECCARCLVGAPLASLVA
                        | | | | |  ||   |  | | |
Adenovirus (561–580) GLLECHCRCNLCTPHRSLVC,
```

the similarity does not extend further. It seems more likely that the virus and myelin proteins are 'analogous', i.e., that similar portions arose by convergent evolution to a common sequence or secondary structure. This idea is consistent with the well-known propensity of viruses to undergo rapid mutation, particularly of their surface antigenic sites [13].

The tryptophan containing region of MBP has long been recognized as being highly encephalitogenic [2,5]. In particular, the nonapeptide, Phe-Ser-Trp-Gly-Ala-Glu-Gly-Gln-Lys, has been shown to be highly effective in inducing EAE in guinea pigs, and studies with peptide analogs indicate that the essential features of an encephalitogenic determinant are X-X-Trp-X-X-X-X-Gln/Asn-Lys/Arg. The proteolipid contains two regions that have a Trp and a Lys residue separated by 5 residues, but otherwise they are not remarkably similar to the encephalitogenic MBP peptide. It is interesting, however, that one of the tryptophans (Trp-144) occurs in a region that shows several similarities with viral proteins (table 1). In the course of our analysis, we also found a strong similarity, not previously reported [3], between the encephalitogenic MBP protein and the *gag-pol* polyprotein of murine leukemia virus (MLV):

```
Myelin basic protein (114–123)     FSWGAEGQKP
                                   | | |    ||
gag-pol protein (MLV) (992–1001) FNWGPDQQKA.
```

It would not be surprising if antibodies against the MLV peptide were found to cross-react with MBP.

There are now at least two reports of mimicry of normal human proteins by viruses: MBP by hepatitis B virus polymerase [4] and α_1-thymosin by the AIDS virus HTLV-III/LAV [14]. In addition Pruijn et al. [15] have reported that sera from patients suffering from autoimmune diseases can inhibit adenovirus DNA replication, suggesting a similar phenomenon. The similarities between viral proteins and myelin proteolipid may represent yet another case. In this instance, it seems likely that portions of the protein which are potential antigenic determinants are located on the outer surface of the cell membrane, making it more understandable how the cell membrane could come under immunological attack.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Institutes of Health (NS-13649 and NS 16945) and National Science Foundation (DMB 85-03940).

REFERENCES

[1] Johnson, R.T. (1984) Viral Infections of the Nervous System, p.263 ff, Raven, New York.
[2] Alvord, E.C. jr, Kies, M.W. and Suckling, A.J. (1984) Experimental Allergic Encephalomyelitis: A Useful Model for Multiple Sclerosis, Liss, New York.
[3] Jahnke, U., Fischer, E.H. and Alvord, E.C. jr (1985) Science 229, 282–284.
[4] Fujinami, R.S. and Oldstone, M.B.S. (1985) Science 230, 1043–1045.
[5] Brostoff, S.W. (1984) in: Myelin (Morell, P. ed.) pp.405–439, Plenum, New York.
[6] Lees, M.B., Chao, B.H., Lin, L.-F., Samiullah, M. and Laursen, R.A. (1983) Arch. Biochem. Biophys. 226, 643–656.
[7] Stoffel, W., Hillen, H., Schroeder, W. and Deutzman, R. (1983) Hoppe Seyler's Z. Physiol. Chem. 364, 1455–1466.
[8] Laursen, R.A., Samiullah, M. and Lees, M.B. (1984) Proc. Natl. Acad. Sci. USA 81, 2912–2916.
[9] Stoffel, W., Hillen, H. and Giersiefen, H. (1984) Proc. Natl. Acad. Sci. USA 81, 5012–5016.
[10] Cambi, F., Lees, M.B., Williams, R.M. and Macklin, W.B. (1983) Ann. Neurol. 13, 303–308.
[11] Hopp, T.P. and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824–3828.
[12] Dayhoff, M.O. (1978) Atlas of Protein Sequence and Structure, vol.5, suppl.3, pp.353–358, Natl. Biomed. Res. Foundation, Silver Springs, MD.
[13] Wiley, D.C., Wilson, I.A. and Skehel, J.J. (1981) Nature 289, 373–378.
[14] Sarin, P.S., Sun, D.K., Thornton, A.H., Naylor, P.H. and Goldstein, A.L. (1986) Science 232, 1135–1137.
[15] Pruijn, G.J.M., Kusters, H.G., Gmelig Meyling F.H.J. and Van der Vliet, P.C. (1986) Eur. J. Biochem. 154, 363–370.