Volume 269, number 1, 145–147 FEBS 08748 August 1990

Prediction of arenavirus fusion peptides on the basis of computer analysis of envelope protein sequences

Svetlana E. Glushakova[1], Igor S. Lukashevich[1] and Ludmila A. Baratova[2]

[1]*Department of Special Pathogens, Byelorussian Research Institute of Epidemiology and Microbiology, Nogina st. 3, 220050 Minsk* and [2]*A.N. Belozersky Research Laboratory of Molecular Biology, Moscow State University, Lenin Hills, 119899 Moscow, U.S.S.R.*

Received 18 May 1990; revised version received 18 June 1990

Theoretical search and selection criteria for putative fusion peptides of enveloped viruses are proposed. Arenavirus fusion peptides are predicted on the basis of computer-assisted analysis of amino acid sequences of arenavirus envelope proteins and elements of their secondary and tertiary structure. Accordingly, two regions of GP2 surface protein from 5 viruses of Arenaviridae family have been detected with properties typical of fusion peptides of other enveloped viruses. One region, named peptide IV, located at the N-terminus of the GP2 protein, is followed by the other region or peptide V, more likely candidate for the arenavirus fusion peptide.

Arenavirus; Computer-assisted analysis of amino acid sequences; Secondary structure prediction; Fusion peptide

1. INTRODUCTION

The fusion of biological membranes, being a key process in the replication of enveloped viruses, has drawn attention from many investigators as the initial step of virus infections [1]. However, many molecular events of this fusion are yet poorly understood. Of special interest in this respect are the viral glycoprotein hydrophobic regions (named fusion peptides) which, owing to their properties, can interact with the cellular and probably with the viral membranes and cause their impairment [2,3]. Fusion peptides are highly conserved not only within one virus family but show homology with members of different virus families [4]. Since synthetic analogs of fusion peptides have been found to inhibit virus replication in vitro [5], this approach may be a useful way for virus infection chemotherapy if the virus is not neutralized by antibodies to virus antigens.

A computer-assisted analysis of amino acid sequences of precursor proteins of glycoproteins (GPC) from 5 Arenaviridae family viruses has been performed for the detection of the putative fusion peptide of these viruses, many of which are highly pathogenic for man.

2. MATERIALS AND METHODS

The published sequences of GPC proteins of the following arenaviruses were used for computer analysis: Lassa virus, LCM virus (ARM and WE strains), as well as Pichinde and Tacaribe viruses [6–10]. The alignment of the deduced amino acid sequences of GPC proteins was performed as described by Auperin and McCormick [11]. Computer programs from the package 'PC Gene' were used for analysis of the primary amino acid sequence and the secondary protein structure: NOVOTNY [12,13], SOAP [14,15], AASCALE [16], RAOARGOS [17], GARNIER [18]. Hydrophobic indexes were computed according to the SOAP program [14,15].

Correspondence address: S.E. Glushakova, Department of Special Pathogens, Byelorussian Research Institute of Epidemiology and Microbiology, Nogina st. 3, 220050 Minsk, USSR

3. RESULTS AND DISCUSSION

3.1. *Criteria for selection of putative fusion peptides of enveloped viruses*

A data analysis of amino acid sequences and functional activities of fusion peptides of enveloped viruses allowed to determine the main physicochemical and topological parameters of fusion peptides: (i) the length of about 20 amino acid residues; (ii) the hydrophobic index 0.5 and higher up by the scale of Kyte and Doolittle [15]; (iii) the high degree of conservation within the family; (iv) location predominantly at the 'extracytoplasmic' terminus of the envelope protein; (v) remoteness from potential glycosylation sites; (vi) the predominantly α-helical conformation; and (vii) location inside the protein globule or viral glycoprotein spikes.

3.2. *Search for putative fusion peptides of arenaviruses*

The amino acid sequences of GP1 and GP2 glycoproteins were analyzed. The distribution along the molecule of α-helical regions, β-sheets, turns, charged amino acid residues, potential glycosylation sites, as well as α-helical regions buried within the protein globule was characterized. Hydropathic profiles of proteins were computed. Six long hydrophobic regions,

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

Fig. 1. Location of hydrophobic peptides on the GPC-map of Lassa virus. The cleavage site of GPC is indicated by a vertical bar. Potential sites of glycosylation are marked by dots.

named peptides I–VI, were located on the linear GPC map of Lassa virus (Fig. 1): peptides I–III were in GP1 and peptides IV–VI in GP2 glycoprotein. Peptide IV was located at the N-termini of GP2. This position is rather typical of fusion peptides of other virus families.

3.3. *Properties of the selected peptides*

As seen in Table I, the length of the peptides selected for analysis varies from 15 to 24 amino acid residues. Such length is typical of fusion peptides of other viruses. A similar conclusion may be drawn with regard to their hydrophobic indexes (0.450–0.581). However, such characteristics as conservation within the family, proximity to potential glycosylation sites, percentage of α-helix structures in the three selected peptides from GP1 are far from conforming to requirements (see section 3.1), which makes their candidacy for arenaviral fusion peptides seem doubtful.

Analysis of the peptides from protein GP2, according to the characteristics presented in Tables I and II as well as to the degree of remoteness from the protein surface (data not shown), has revealed two distinct structures, peptides IV and V. The latter has the maximal degree of homology. Thus, a comparison of this peptide in Lassa and LCM viruses showed 23 out of 24 amino acid residues to be identical or to possess conserved amino acid substitutions (Fig. 2). Moreover, a comparison of this peptide in two serologically different arenaviruses, Tacaribe and LCM, demonstrated 21 conserved or analogous amino acid residues (Fig. 2).

The described peptide V of Lassa virus GP2 protein has a length of 24 amino acid residues, hydrophobic index 0.554, and 75% of all residues in the α-helical conformation (GARNIER program). The peptide is probably buried within the protein globule (RAOARGOS program, data not presented), though in the immediate proximity to the N-terminus of GP2.

The second candidate for the fusion peptide, peptide IV, is located at the N-terminus of GP2 and consists of 18 amino acid residues (Table I). It satisfies many requirements for these functional parts of viral proteins. In addition, the tripeptide Gly-X-Phe, found in peptide IV, was observed to be one of the frequent elements in the fusion peptides of enveloped virus proteins [4].

Thus, the computer-aided analysis of amino acid sequences of arenaviral glycoproteins performed with the aid of the software package 'PC Gene' and on the basis of summarised properties of the other enveloped

Table I

Physico-chemical properties of hydrophobic peptides of Lassa virus

Peptide no. (see Fig. 1)	Sequences, secondary structure[a] and % aa residues in α-helix conformation	Remoteness from glycosylation sites[b]	Hydrophobic index
I 128–147 aa	L Y D H A L M S I I S T F H L S I P N F H H H H H H H H E E E E E E E C C C T C 40%	9; 25	0.581
II 173–192 aa	A G D A A N H C G T V A N G V L Q T F M T T H H H E E E E E E E E E H E E E E 17%	6; 32	0.538
III 233–247 aa	F S R P S P I G Y L G L L S Q T T C C C T T E E E E E E E E –	9; 9	0.522
IV 258–275 aa	L L G T F T W T L S D S E G K D T P T C E E E E E C C T C T T T T C T T –	–; 90	0.450
V 276–299 aa	G G Y C L T R W M L I E A E L K C F G N T A V A T E E E E E H H H H H H H H H H H H H H H H 75%	–; 66	0.554
VI 328–345 aa	A E A Q M S I Q L I N K A V N A L I H H H H H H H H H H E H H H E H H H 89%	–; 20	0.575

[a] H, α-helix; E, β-sheet; T, turn; C, coil

[b] The nearest sites of glycosylation are indicated as no. of amino acid residues towards the polypeptide N- and C-ends

Table II

Homology between hydrophobic peptides of arenavirus glycoproteins

Peptide no. (see Fig. 1)	Percentage of homologous residues		
	Among all arenaviruses	Among 'Old World' arenaviruses	Among 'New World' arenaviruses
I	5	65	–
II	13	25	42
III	6	25	31
IV	50	72	56
V	67	96	67
VI	33	39	44

According to the alignment of the deduced amino acid sequences of GPC proteins (Auperin and McCormick [11]) the percentages of amino acid similarity (identity and position) between GP1 and GP2 proteins of Lassa, LCM, Tacaribe, and Pichinde viruses are about 11% (GP1) and 36% (GP2)

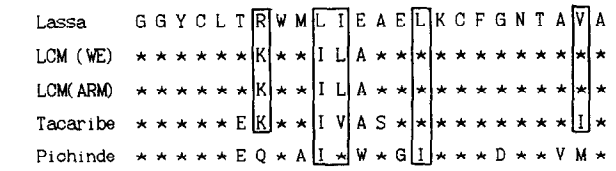

Fig. 2. Amino acid sequence comparison of the predicted fusion peptide of arenaviruses. Identical amino acids are marked by asterisks, the positions with conserved amino acid substitutions are framed.

viruses revealed two peptides, IV and V, in the GP2 glycoprotein (most conserved among arenavirus proteins [11]) which may be arenaviral fusion peptides. Although we do not rule out that peptide IV might be involved in the fusion process, peptide V seems to be the most probable fusion peptide even though it is not located at the N-terminus of the protein. However, the internal protein regions of some proteins likewise exhibit fusion activity [4]. The hydrophobic peptide V in the GP2 polypeptide is followed by a highly charged region of the amino acid sequence; here the situation is similar to melittin, the bee venom peptide capable of lysing membranes. Melittin charges are assumed to interact with charged heads of phospholipids, stabilizing thereby the contact of the protein with the membrane [19].

It is clear that the analysis of the protein sequences and extrapolation of the properties of known fusion peptides to analogues in other viruses need further experimental verification. For this purpose peptide V was synthesized and tested in model experiments with lipid bilayers and liposomes. Results demonstrate an ability of this peptide not only to fuse liposomes (in assay of resonance energy transfer) and to form single ion channels in planar lipid bilayers, but even to inhibit viral replication in cell culture in concentration 10^{-9}–10^{-5} M (manuscript in preparation). Taken together, these data suggest that peptide V quite probably may be a fusion peptide of arenaviruses, and GP2 their fusion protein.

Acknowledgements: The authors wish to thank Drs A.V. Efimov and N.I. Grebenschikov for much helpful advice and discussion, Professor N.V. Kaverin and Professor A.A. Lev for critical reading of the manuscript. Thanks are also due to I.S. Kochubey for his help in editing the English text.

REFERENCES

[1] Marsh, M. (1984) Biochem. J. 218, 1–10.
[2] Hoekstra, D. and Kok, W. (1989) Biosci. Rep. 9, 273–305.
[3] Wiley, D.C. and Skehel, J.J. (1987) Annu. Rev. Biochem. 56, 356–369.
[4] Gallaher, W.R. (1987) Cell 50, 327–328.
[5] Richardson, C.D., Scheid, A. and Choppin, P.W. (1980) Virology 105, 205–222.
[6] Auperin, D.D., Romanovski, V., Galinski, M. and Bishop, D.H.L. (1984) J. Virol. 52, 897–904.
[7] Auperin, D.D., Sasso, D.R. and McCormick, J.B. (1986) Virology 154, 155–167.
[8] Franze-Fernandez, M.T., Zetina, C., Iapalucci, S., Lucero, M.A., Bouissou, C., Lopez, R., Rey, O., Daheli, M., Cohen, G.N. and Zakin, M.M. (1987) Virus Res. 7, 309–324.
[9] Romanowski, V. and Bishop, D.H.L. (1985) Virus Res. 2, 35–51.
[10] Romanowski, V., Matsuura, Y. and Bishop, D.H.L. (1985) Virus Res. 3, 101–114.
[11] Auperin, D.D. and McCormick, J.B. (1989) Virology 168, 421–425.
[12] Chou, P.Y. and Fasman, G.D. (1987) Adv. Enzymol. 47, 45–148.
[13] Rose, G.D. and Roy, S. (1980) Proc. Natl. Acad. Sci. USA 77, 4643–4647.
[14] Klein, P., Kanehisa, M. and DeLisi, C. (1985) Biochim. Biophys. Acta 815, 468–476.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Jones, D.D. (1975) J. Theor. Biol. 50, 167–183.
[17] Rao, J.K.M. and Argos, P. (1986) Biochim. Biophys. Acta 869, 197–214.
[18] Garnier, J., Osquthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[19] Terwilliger, T.C., Weissman, L. and Eisenberg, D. (1982) Biophys. J. 37, 353–361.