

Secondary structure prediction of the hemagglutinin-neuraminidase from a porcine rubulavirus

R. Zenteno-Cuevas^{1,2}, J. Hernández², B. Espinosa³, J. Reyes², and E. Zenteno¹

¹Laboratorio de Inmunología, Departamento de Bioquímica,
Facultad de Medicina, UNAM, México

²Centro de Investigación Biomédica de Oriente, IMSS, Puebla, México

³Departamento de Bioquímica, Instituto Nacional de Enfermedades Respiratorias,
Tlalpan, México

Accepted August 28, 1997

Summary. The Hemagglutinin-Neuraminidase (HN) from ‘La Piedad, Michoacan’ porcine rubulavirus (LPMV) interacts specifically with NeuAc α 2,3 lactose residues on the target cell. In this work we report the secondary structure of this protein, determined with five different theoretical algorithms. Results indicate that the HN protein is organized in: an intracellular region (from amino acid 1 to 25); in a β -strand transmembrane region (residue 26 to 47), typically hydrophobic, rigid and solvent inaccessible; and extracellular region (residue 48 to 576), which possesses hemagglutinating and neuraminidase activity. The secondary structure in this region is organized in a β -loop- β alternated with few α -helices. Regions with structural and functional implications were determined by pattern search and multiple alignment of the HN from LPM with 12 rubulaviruses and paramyxoviruses HN sequences. The low diversity observed among the HN sequences evaluated indicates that in general the structural organization of the protein, and in particular its sugar binding domain, is closely related among both genera, thus suggesting that the sugar binding domain is well preserved through evolution.

Introduction

The porcine rubulavirus, La Piedad, Michoacan (LPMV), is responsible for the blue eye syndrome in pigs. It induces reproductive alterations in adults and neural alterations with high mortality rates in piglets [1]. This virus has the same structural organization as all rubula- and paramyxoviruses, and is composed of two proteins: the Fusion protein (F), and the Hemagglutinin-Neuraminidase (HN). HN is a tetrameric protein that recognizes and binds specifically saccharidic structures on the surface of the host cell [2, 3]. The agglutinating activity of LPMV

is directed specifically to α 2,3 sialyllactose containing structures, suggesting that the pathogenicity of this virus is due to the capacity of HN to recognize specific receptors on target organs, such as the respiratory tract, the nervous system and the epididimus [4, 5]

The LPMV rubulavirus has recently been classified in the genus *Rubulavirus*, subfamily *Paramyxovirinae*, family *Paramyxoviridae*, order *Mononegavirales* [2] and its HN protein has been characterized at the molecular level [6–7]. The gene is 1906 nucleotides long, and expresses a 576 amino acid sequence, with a molecular weight of 63.3 kDa. This protein seems to possess four potential N-glycosylation sites and a major hydrophobic region near the N-terminal, suggesting that this is a membrane anchor domain. Comparisons of the protein amino acid sequence with that of other paramyxoviruses revealed great similarity with HN from Simian virus 5 and mumps viruses [8].

Many efforts have been made to characterize the structure of paramyxovirus and rubulavirus HN proteins taking advantage of techniques such as theoretical algorithms [9], structural correlation with influenza neuraminidase [10], structural evaluation using monoclonal antibodies [11–14], circular dichroism [15], and protein crystallization for X-ray diffraction studies [16–18]. However, the structure and function of HN proteins from paramyxoviruses in general, and LPMV, in particular, has not been completely elucidated. The aim of this work is to define the structural organization and physicochemical properties of the HN protein from the LPMV rubulavirus, and to find regions with potential sugar binding capacity, in order to understand the specific interaction with target cells.

Materials and methods

Structural characterization

Five different theoretical algorithms were used to predict the secondary structure of the HN protein of LPMV based on the reported amino acid sequence [7]:

1. Chou and Fasman [19] reported that their method based on the calculation of conformational parameters from observed frequencies of amino acids in proteins with a well-known structure, achieved 51% correct prediction.
2. Garnier [20], considered the parameters for the secondary from a subset of 26 proteins with a well known three-dimensional structure.

$$I(k, i) = \ln[P(k/i)/P(k)]$$

where: k = designates the given state (α helix, α sheet or turn coil)

$P(k/i)$ = is the probability of the k state for an amino acid (AA) i

$P(k)$ = the probability of the k state.

$I(k, i)$ = is the information on how an AA(I) behaves on state k . The optimized decision constants used were fixed according to the optimized set determined by the authors. The level of correct prediction was 58–60%.

3. Deleage and Roux [21], developed an algorithm that allows the prediction of the secondary structure in four step: a) prediction of the structural class of a protein from the amino acid composition, b) preliminary estimation of the secondary structure with a simple algorithm; c) comparison between two independent predictions, and d) optimization of parameters and re-prediction of the secondary structure. The parameters for secondary struc-

ture prediction were determined from a set of 60 proteins with well-known three-dimensional structures according to the formula:

$$P(k, i) = P(k/i)/P(k).$$

where: k = designates the state (α helix, β sheet or turn coil)

$P(k/i)$ = is the probability of the k state for AA i .

$P(k)$ = is the frequency of the k state.

$P(k, i)$ = is the parameter of AA i for state k .

The percentage of correct prediction is 61–72%

4. Levin et al. report that their algorithm [22] allows the prediction of the secondary structure using similarities with proteins of known secondary structure. Each stretch of residues is compared with all the stretches on the same longitude on all the proteins of the Kabsch and Sander database [23]; this method leads to 62% correct predictions.

5. The Rost and Sanders, or PHD method, (version 5.94–317) which uses a three-layered neural network approach (25–26), is based upon a jury decision that averages several independently trained networks before making the final assessment. This method groups the eight secondary structure classes defined by DSSP [23] into three output classes: H, helix, (α , 3_{10} , π); E, extended β -strand; and C, turn, bend, coil, loop, β -bridge, and non-periodic structures. This method has an average cross-validated successful prediction above 72% on a set of 250 sequence-unique chains with known structures. This analysis was made by e-mail server [27].

Utilising these prediction approaches a consensus was reached regarding the nature of the secondary structure: 1. For a correct prediction at least three different algorithms should be considered for each residue with the same structure. 2. For any predicted α -helix, it was necessary to identify at least six sequentially consecutive residues assigned to the helical conformation; the minimum length necessary for determining β -strands was four residues, while for loops it was three residues. 3. When doubts arose concerning allocation, the influence of neighbor residues was considered. 4. Discrepancies between different methods on the length of each α -helix segment were corrected by manual inspection. Taking into account amino acid preference for location at the ends of helical segment [28]. 5. A loop structure was considered when three algorithms yielded the same residue with turn or coil structure, or if many uncertainties arose in the allocation. The exact denotation of turn or coil structure was excluded. The justification for using several algorithms to predict secondary structures was to improve prediction by increasing accuracy [20, 29].

Protein folding recognition analysis

To confirm the validity of our predictions, a comparison was made with the TOPITS program. This program which uses a set of new methods for predicting secondary structures, is based on the observation that similar three-dimensional structures are often adopted by proteins, even if no significant sequence homology is detected. The aim of this program is to optimally combine a sequence with one structure, a process referred to 'threading'. In practice, one must use a database of known structures and a computer program with enough potential for threading, sorting and ranking the possible solutions [30]. In our case the threading TOPITS method was used by e-mail server service [31]; and the database was the PDB release # 77 (July 1996).

Physicochemical characterization

Three different theoretical algorithms were used to characterize some physicochemical aspects of the HN protein, 1) flexibility, determined by the method of Karplus and Schulz

[32], based upon scales of known temperature factors of the α -carbons from proteins with a known structure, 2) hydrophathy determined, as described by Kyte and Doolittle [33], by averaging the hydrophobicity index for each individual amino acid; in both cases a window of eleven residues was used, 3) solvent accessibility was calculated as described by Kabsch and Sander [23].

Transmembrane regions analysis

Transmembrane regions were analyzed using three different and specific algorithms as defined by Eisenberg et al. [34], Roa and Argos, [35] and Klein [36]. Only those regions with the adequate physicochemical properties, defined by the three-algorithms, were considered as transmembranal.

Detection of sites and signatures in the sequence (PROSITE-dictionary)

The Prosite dictionary release No21-1995 (37) was used to find signatures or residues with potential biological significance.

Alignment process

Twelve HN sequences were aligned using the Clustal program [38], considering a gap penalty of five in the pairwise similarity comparison and an open gap cost of 10 for the final alignment: Rubulavirus LPMV (access number S77541, from the GenBank Data Base 1994); two mumps viruses MUMPM (P11235), MUMPI (P19762); two Simian viruses SV41 (P25180), SV5LN (P28885); three Newcastle disease viruses NDVH4 (P12559), NDVI (P12556), and NDVU (P12558); one Parainfluenza 2 virus, PI2H (P25465); two Parainfluenza viruses, PI3B (P06167), PI3HU (PI2563), and one Sendai virus SENDZ (P04853) from the paramyxovirus genus. Another alignment was carried out by increasing the database to 25 HN sequences for consideration of the paramyxoviruses and rubulaviruses with the same program and parameters. The new sequences were, one MUMPS virus: MUMPR (P10866); six Newcastle disease viruses: NDVB (P32884), NDVC (P35740), NDVJ (P35742), NDVM (P12557), NDVQ (P13850), and NDVTG (P12553); two Parainfluenza viruses PI2HT (P25466) and PI3HT (P12562); one Sendai virus SENDH (P03425), and three Simian viruses, SV5 (P04850), SV5CM (P28883) and SV5CP (P28884). No manual modification was made to the final alignments; all HN sequences were obtained from the Swiss-Prot database 1994.

Most prediction and sequence analyses were made with the package Antheptot Version 5.5 [39], Windows version V2.3, on a 486 personal computer.

Variability analysis

Variability analysis was performed using the VIR program [40] and the algorithms of Wu and Kabat, [41] and Jores et al. [42]. The calculation of variability indeces was performed in each position of the 25 HN multiple alignment sequences previously obtained, according to these definitions for each algorithm:

i) The Wu-Kabat Index (Vwk):

$$Vwk = k/p_1,$$

where k is the number of different amino acid types that appear at a given position, and p_1 is the frequency of the most common amino acid in that position.

ii) Jores index (1990)(V_J):

$$V_J = j/np_1,$$

were j is the number of distinguishable amino acid pairs that appear at a given position, and p_1 is the frequency of the most common amino acid pair in that position.

*A search of common physicochemical properties using hierarchies
of amino acid categorizations*

This analysis was performed only with the residues from the potentially hemagglutinating domain, considering the average amino acids in each position and its correlation with the Jimenez-Montaño hierarchy of amino acid [43].

Results

Structural characterization

Table 1 shows the secondary structure of HN from LPMV, determined with each algorithm. The consensus structure shows a low content of α helix (11.7%), and a high content of β strand (43.3%) and loop structure (turn and coil) (45%) (Table 1 and Fig. 1). It is important to note that the structural β -loop- β arrangement was interrupted in a few cases by α -helix segments in the middle and at the end of the sequence (Figs. 1 and 2).

Physicochemical properties and transmembrane regions

Hydrophobic and solvent accessible properties coincided with the structure in their alternate tendency, although the flexibility did not show this tendency. The three methods used for identification of the transmembrane region indicated the presence of several regions; however, only a region close to the amino terminal (residues 26–47) was defined by all the methods (Fig. 1). This region also showed high rigidity, hydrophobicity, and solvent inaccessibility values.

Table 1. Comparison of the secondary structure content (%) on the HN of LPMV determined by several methods

Method	Alpha-Helix	Beta-Strand	Loop
Chou-Fasman	19.9	55.2	24.8
Levin	23.0	26.0	28.9
Garnier (1996)	16.3	23.9	59.6
Deleage-Roux	6.2	45.8	47.9
Rost	19.6	21.4	59.0
TOPITS ^a	18.0	25.9	56.1
Consensus ^a	11.7	43.3	45.0

^aThese values were obtained from Fig. 2

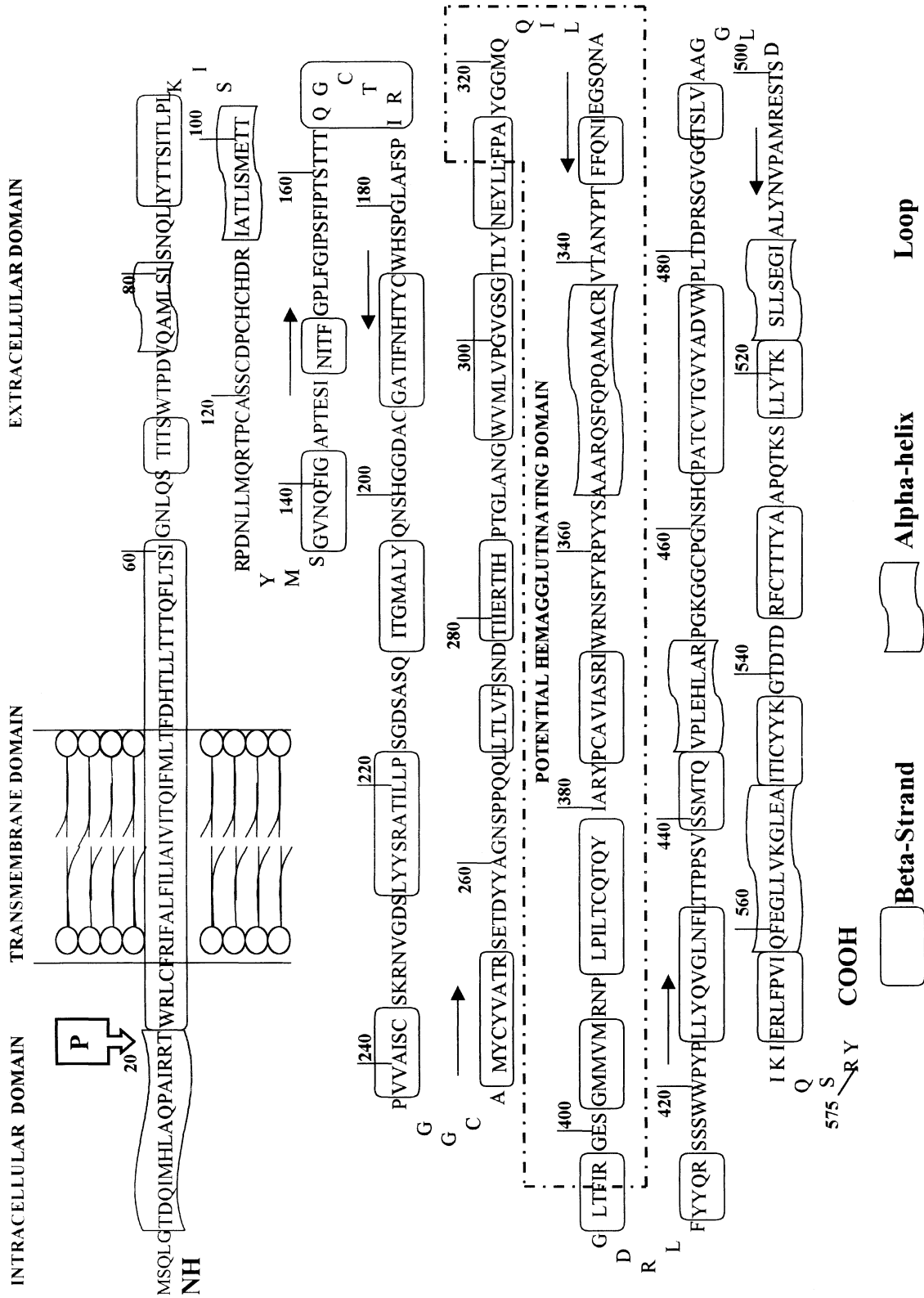


Fig. 1. Secondary structure representation for the HN of LPMV rubulavirus, see Fig. 2 for details

Sites and signature analysis

The HN protein from LPMV showed seven potential phosphorylation residues with a Protein Kinase C activity and eight Casein Kinase C activity, and also twelve N-myristoylation sites, and four potential N-glycosylation sites, which correspond to the glycosylation sites suggested by Sundqvist et al. [8].

Homology score

The comparison of homology scores between HN LPMV and other HNs (Table 2) revealed that the highest homology scores were with SV5LN and MUMPM HN's with 43 and 42% respectively, and that the lowest was with NDVI HN with a 30% homology score. Table 2 also shows a clear differentiation between the genera analyzed. Homology score of 30 to 95% were found in all the HN of the rubulavirus genus, whereas a 48 to 77% score was obtained with the paramyxovirus HN's.

Multiple alignment

The 12 and 25 HN aligned sequences from paramyxo- and rubulavirus, including the LPMV, indicated the presence of common patterns previously reported for the family, such as the NRKSCSI/V/L, FXXYGGV/L/M and GA/SEGRI/V/L (Figs. 3:2, 3:3 and 3:4 respectively, only twelve HN aligned sequence are shown) [44]. The Prosite analysis of the HN from LPMV showed an intracellular TXR pattern, where the threonine 21 could be phosphorylatable by a Protein Kinase C. This pattern seems to be a new genus-specific one, since it is well conserved in all rubulaviruses (Fig. 3:1), and is absent in the paramyxovirus genus. These alignments led us to identify in the HN from rubulaviruses, a new pattern, which seems to be species-specific. This pattern is situated between residues 16 to 25; in the HN from LPMV it contains the sequence AIARRTWRLC, the P12H contains the sequence SIPKRTCRII, the mumps viruses AADKKTFRTC, and the sequence REAKNTWRLV in the Newcastle disease virus. This new pattern contains a phosphorylation site and is delineated at the carboxyl region by two highly conserved residues: phenylalanine and arginine (FR). The HN from LPMV shows these residues at the beginning of the transmembrane region (Fig. 1 and 3:1).

Identification of the potential hemagglutinin domain

The potential hemagglutinin domain has been excluded in previous comparative analysis [10]; however, the multiple alignment of 25 HN sequences performed in this work led us to identify one region, from residues 314 to 404 (Figs. 3:3 and 4, and 7), that seems to contain the hemagglutinin domain. It possesses two highly conserved patterns at their extremities, which could mark the beginning and the end of the hemagglutinating domain. One pattern, from residues 314 to 320 (346–437 alignment numbering), contains the sequence FXXYGGV/L/M (Fig. 3:3 and 7); and is found inside a β -strand as part of a loop in the HN from LPMV (Fig. 2). The other, from residues 398 to 404 (431–437) with the sequence GA/SEGRI/V/L, is found inside a β -strand and as part of a loop in the HN from

Table 2. Homology scores for the HN from LPMV, versus eight HN's from rubulavirus and three HN's from the paramyxovirus genus

	Rubulavirus genus										Paramyxovirus genus	
	LPMVV	SV5LN	MUMPM	SV41	MUMP1	P12H	NDVH4	NDVU	NDVI	P13B	P13HU	SENDZ
LPMVV	–	43	42	42	41	39	31	30	30	28	27	26
SV5LN		–	44	52	43	46	32	32	32	28	24	26
MUMPM			–	42	95	41	32	33	32	26	25	24
SV41				–	42	62	35	35	35	28	28	27
MUMP1					–	40	31	31	31	25	27	25
P12H						–	32	32	32	25	26	26
NDVH4							–	94	91	26	24	26
NDVU								–	95	25	24	26
NDVI									–	28	25	26
P13B										–	77	49
P13HU											–	48
SENDZ												–

LPMV (Fig. 2). The relevance of this fragment is indicated by point mutations in the GA/SEGRI/V/L pattern which lead to little or no mature protein, affect the intracellular processing and produce a mature protein with altered antigenicity [45].

Protein folding recognition analysis TOPITS

Comparison between the TOPITS prediction and the prediction we obtained, concerning the secondary structure of the HN from LPMV, in percentage terms (Table 1) as well as in a more detailed comparison (Fig. 2), show great similarities; however, the main differences were found in the transmembranal region, and other segments as in the 79–100 and the 375–420 (Fig. 2).

Variability analysis

The two-variability indexes, led us to analyze the variability of a multiple alignment from 25 HN sequences from different species. Figure 4 shows the variability of the HN sequences determined according to the Wu-Kabat index [41]. The results obtained with this method indicated a widely spread variability of HNs; however, it was possible to find three regions with high levels of variability: the first with the highest values from residues 80–110, was very close to the potential transmembrane of the HN from LPMV (Fig. 4:1), the second was located between residues 125–160 (Fig. 4:2), and the third lay between residues 300–340; (Fig. 4:3) very close to the potential hemagglutinin domain [46]. Fig. 5 shows the variability determined by the Jores method [42], indicating the three variable regions mentioned above (Fig. 5:1,2 and 3); and two other variable regions, located after the potential hemagglutinating domain on residues 470–510 (Fig. 5:4) and

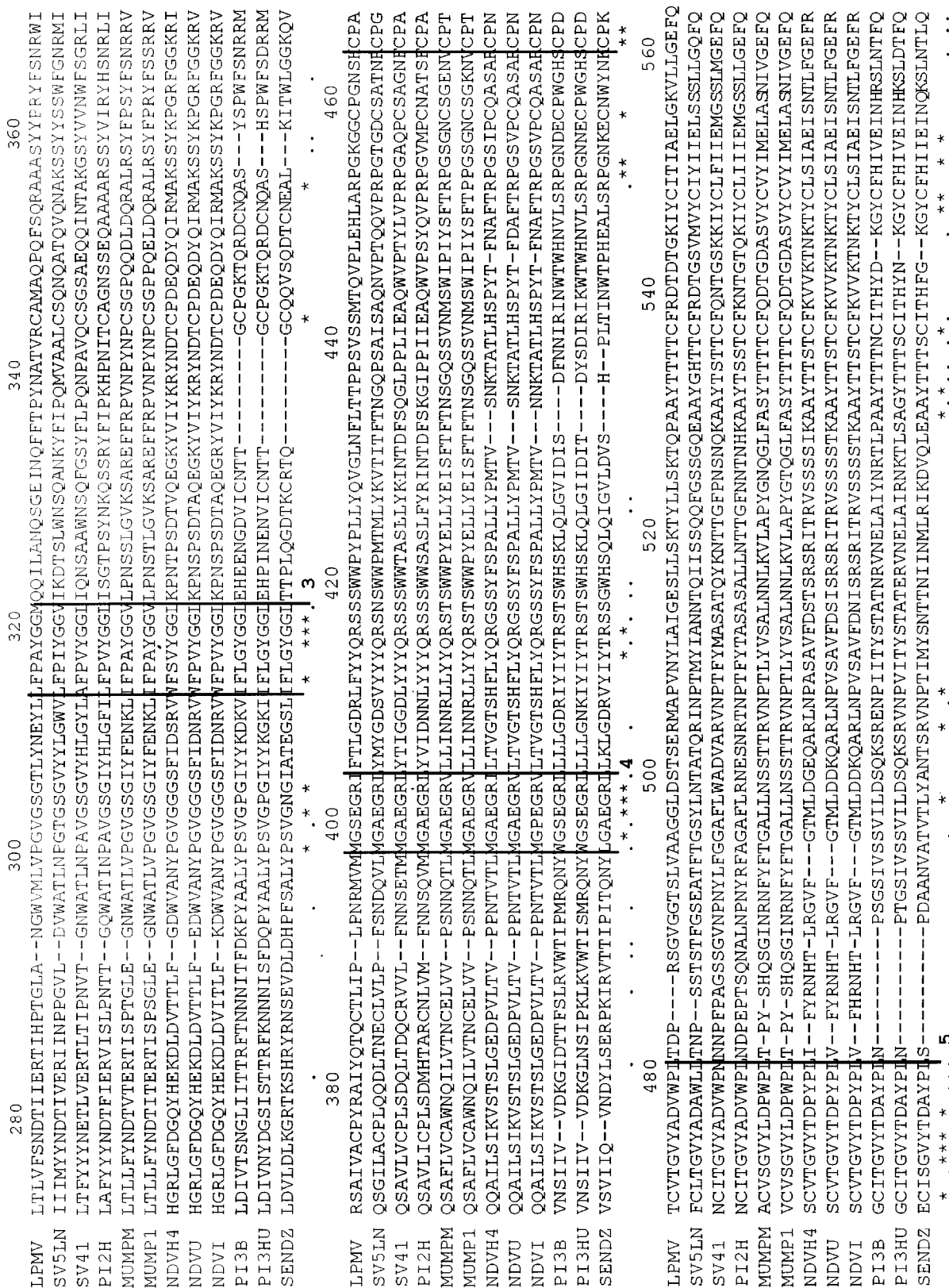


Fig. 3 (see caption on p. 344)

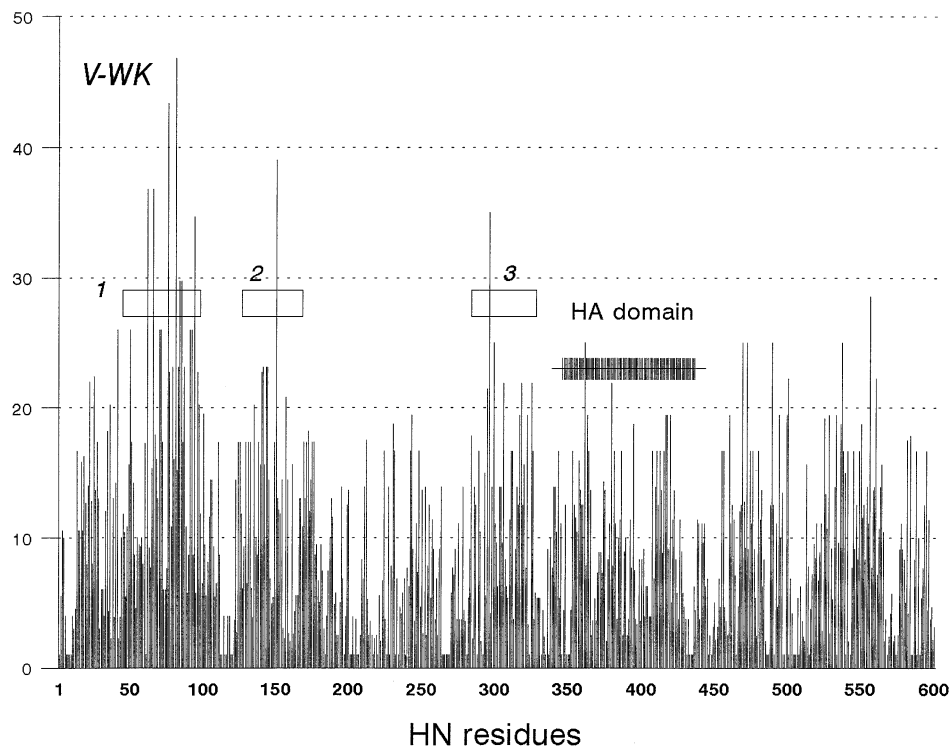


Fig. 4. Variability plot for HN sequences using VK index. Variable regions are boxed and numbered, the HA domain is delineated by a black box

540–550 (Fig. 5:5). It was interesting to note that the profile of the variability, obtained with this index, had a much better resolution than the variability profile obtained with the Wu-Kabat index.

Variability of the potential hemagglutinin domain

The analysis of the variability of the potential hemagglutinating domain with the 25 multiple alignment HN sequence (Fig. 7), performed with the Wu-Kabat index (Fig. 6a), did not lead us to see positions of high variability in the domain; however, the use of the Jores index showed a better resolution for variable and

←
Fig. 3 (p. 342f.) Multiple alignment of 12 paramyxoviruses and rubulavirus HN amino acid sequences, italic numbers indicate the alignment numeration for the 25 alignment sequence and inferior numbers indicate the sequence numeration (Fig. 2), the numbers on the bottom show the highly conserved patterns: 1 the species-specific pattern, (*) on the top indicates the highly conserved threonine; 2 the pattern with the neuraminidase function NRKSCS; 3 the pattern at the beginning of the potential Hemagglutinating domain, FXXGGV/L/M; 4 the pattern at the end of the potential hemagglutinating domain and with activity function, GA/SEGR/V/L and 5 a new very long pattern CPXXCI/V/LS/T/KGVYXD without known function

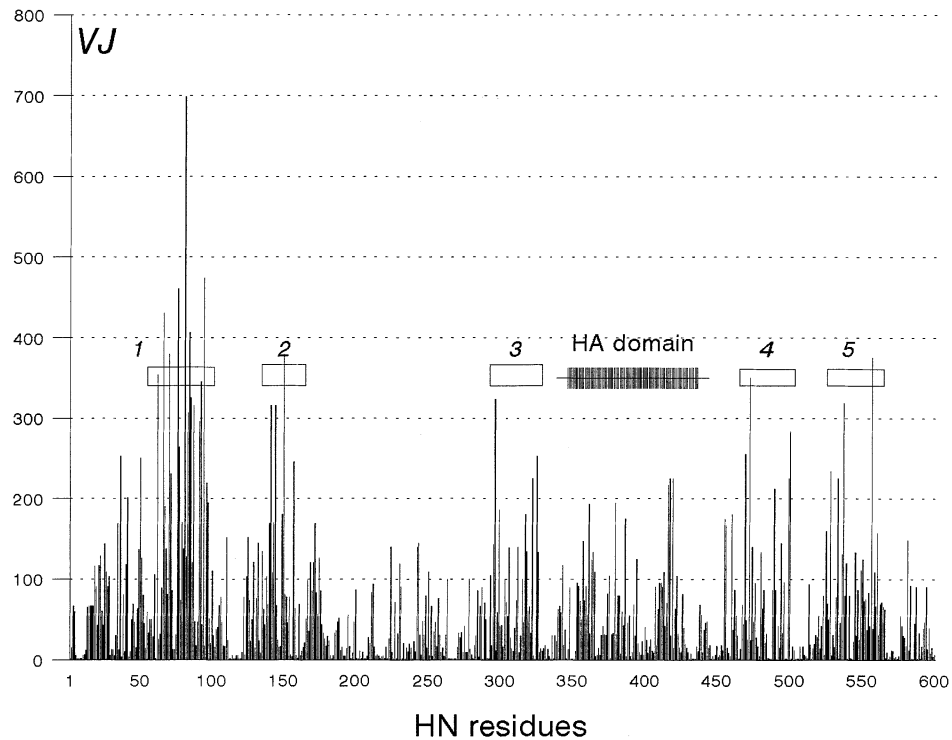


Fig. 5. Variability plot for HN sequences using VJ index; variable regions are boxed and numbered 1, 2, 3, 4, and 5, the HA domain is delineated by a black box

conserved positions; moreover, two stretches of residues with high variability values, from 358 to 369 and from 409 to 424 (Fig. 6b:6 and 6b:7), were defined.

Common physicochemical properties at each position of the potential hemagglutinin domain

From the positions 314 to 404 (346–437) (Fig. 3:3, 4 and 7), considered as the potential hemagglutinin domain, only 12 positions: 353, 361, 363, 374, 379, 386, 388, 412, 413, 416, 419, and 422 show no common physicochemical properties, despite their high diversity and variability values (Figs. 2 and 6b). Only five positions showed high variability values and common physicochemical properties, these were the positions 357, 364, 375, and 394; where 76%, 64%, 76% and 76% of these amino acids were hydrophilic, and position 417, where 96% of the residues were polar amino acids.

We also observed seven highly conserved positions, devoid of common physicochemical properties: position 360, with 36% hydrophobic, and 40% polar residues; position 391, with 52% hydrophobic residues; position 409, with 44% hydrophilic and 40% non polar residues; position 418, with 44% hydrophobic residues and 36% proline, position 421, with 44% hydrophobic residues; position 423, where 52% of the amino acids were hydrophobic and finally, position 428, with 52% of hydrophilic amino acids.

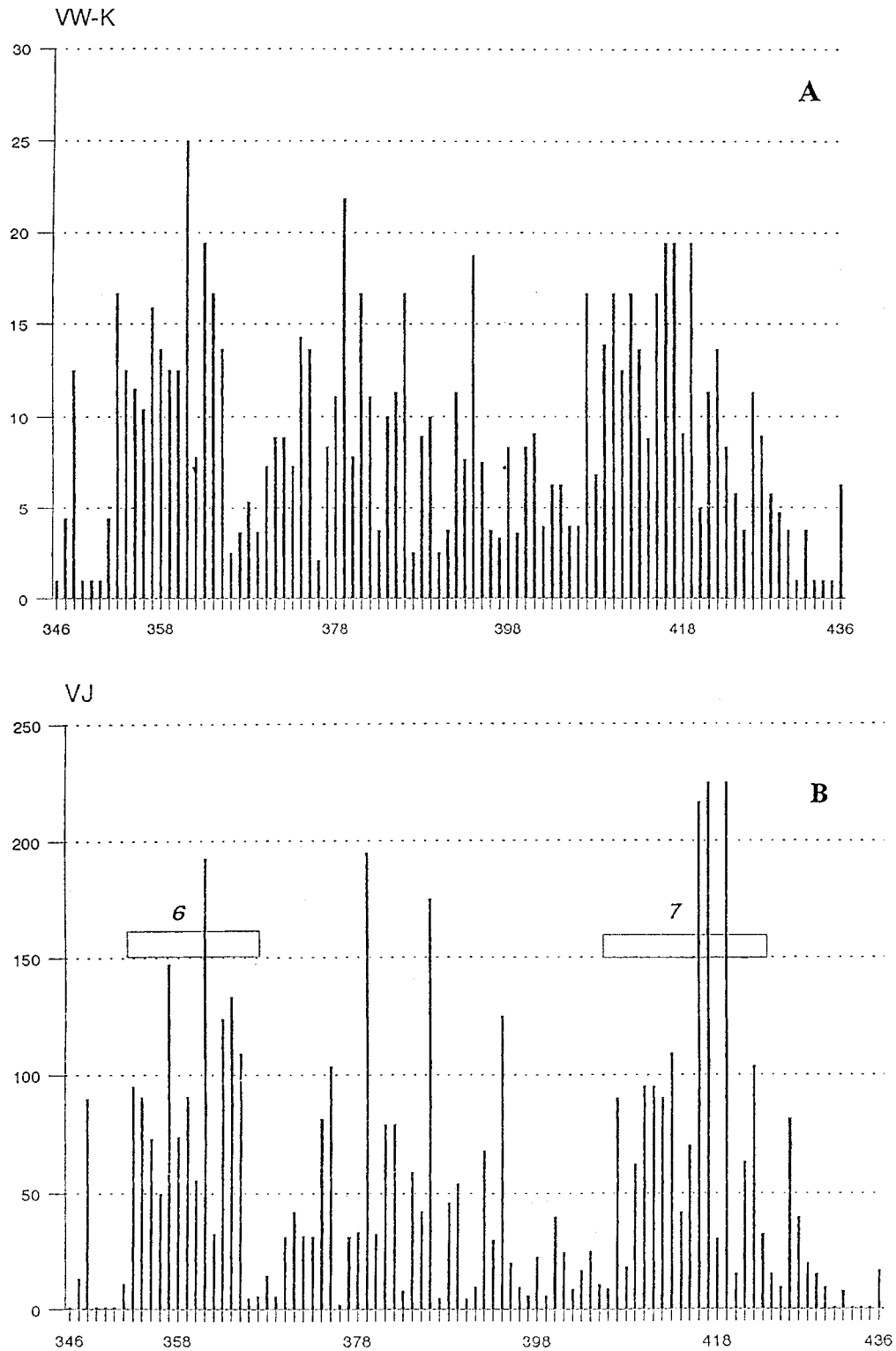


Fig. 6. Variability plots for Hemagglutinin domain, using VK (A) and VJ (B); variable regions are boxed and numbered 6 and 7

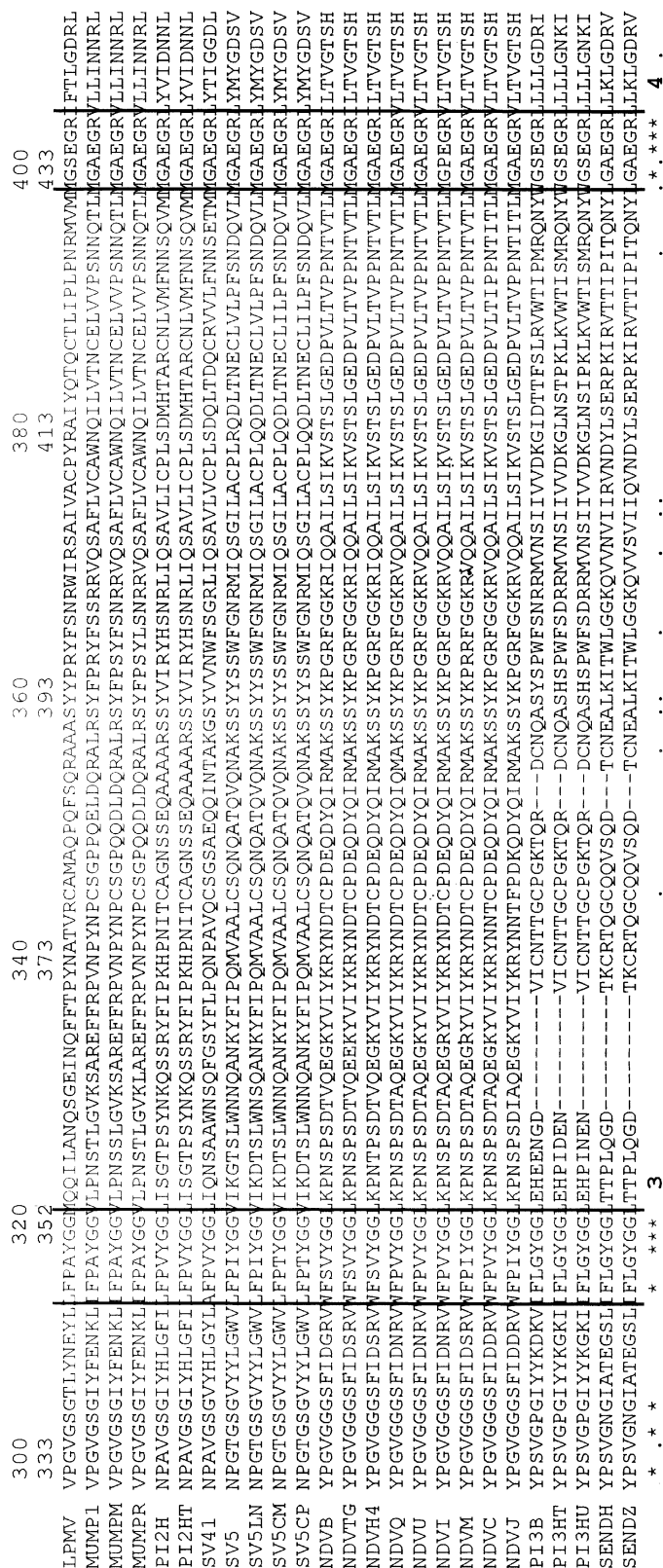


Fig. 7. HN multiple sequence alignment

Discussion

The structural β -loop- β arrangement, as well as the hydrophobic and accessibility alternated design, observed in this study of the hemagglutinin-neuraminidase from the porcine paramyxovirus LPMV, indicates that the HN is organized in an anti-parallel β strand fold. According to our results, the interpretation of the HN structure for LPMV is more complex than in previous structural reports for the HNs from other paramyxoviruses [10]. Discrepancies could be due to the interference of 90 to 100 amino acids, containing the hemagglutinin domain, this segment could induce an alteration in the three-dimensional structural organization. For this reason we consider that the structural organization of these proteins should be explained in terms of their particular dual functions. The results obtained with our prediction, indicate that the transmembrane region of the HN was a β -strand; although this is an uncommon structure for a transmembrane domain, X-ray crystallography of OmpF and PhoE from *E. coli* [47] and of the proaerolysin toxin from *Aeromonas sp* [48], shows a similar transmembrane β -strand domain. Moreover, as the HN from rubulavirus LPMV these proteins show a high tendency to form tetramers and trimers.

The multiple alignment permitted us to see in the HN's from the rubulavirus genus, a new species-specific pattern between residues 16 to 25 (Fig. 3:1), which contains the phosphorylatable residue threonine 21. The exclusivity of the sequences observed in each viral species, could be due to the interaction with a particular host cell element or with other specific viral components such as the fusion protein [49]. Alternatively this species-specific sequence could be explained in terms of the interaction with the Protein Kinase C or other activation proteins on infected cells. The relevance of the Protein Kinase C in the infection process has been suggested by experiments performed with measles virus, which clearly show that Kinase activity increased in the infection; moreover when antibodies directed against the hemagglutinating protein were added, the viral activity and protein Kinase C activity diminished to basal levels [50]. No similar information has been described for other members of the paramyxovirus family.

Our alignments allowed us to identify some patterns reported as highly conserved in HN proteins [44]; which possess structural and enzymatic functions. One of these patterns is the NRKSCSI/V/L segment from residues 233 to 239 (Fig. 3:2), found inside a loop (perhaps a β -turn) and a β -strand segment in the HN from LPM. It has been proved that this pattern generates in the HN, architectural alterations, antigenic variations, syncytium formation, and a drastic inhibition of neuraminidase activity [51]. Another preserved pattern was found on residues 398 to 404 (431–437, Fig. 7) were the amino acids GA/SEGRI/V/L (Fig. 3:4), in the LPMV are situated inside a loop-like structure (perhaps a β -turn) and two β -structure regions. Recent studies by site directed mutations in this region for the HN from Newcastle disease virus demonstrated that this region is relevant for intracellular processing, and correct folding of the protein [45]. We have also identified a new sequence pattern between residues 464 to 475, CPXXCI/V/LS/T/KGVYXD (Fig. 3:5). In the absence of further data, only

mutations in this pattern may help us to understand its conservation or biological function.

The variability analysis of the 25 HN aligned sequences, including LPMV, performed with the Wu-Kabat index allowed us to define three regions with high rates of variability (Figs. 4 and 5); moreover, the use of the Jores' index [42], (Fig. 5:4 and 5:5) clearly defined five regions, confirming that for HN proteins, variability with the Jores algorithm index was more useful than the Wu-Kabat index. In view of the extent of conservation of the sequences analyzed, and the localization in specific stretches of high variability, this behavior could be explained by the conservation of sequences. The region with the highest rate of variability (region one of variability (Figs. 4:1 and 5:1) was found near to the transmembrane domain. This region seems to be implicated in the interaction with homologous proteins in the formation or stabilization of a functional tetramer, thus suggesting that this association could be species-specific. The third variability region (Figs. 4:3 and 5:3), is antigenically active, since immunogenic peptides have been identified in this region in other rubulaviruses, such as Newcastle virus [46].

The variability rate determined with the Jores index, together with the amino acid physicochemical properties, allowed us to identify three groups of residues in the potential hemagglutinin domain (Figs. 3 and 4): 1) A group of residues with a high variability rate and without common properties (12 residues); 2) A group of residues with a high variability rate and with common properties (5 residues); 3) A group with conserved residues without common physicochemical properties (7 residues). Participation of these residue groups, in sugar specificity or structural function should be confirmed by point mutations.

Overall, our results demonstrate that the HN protein from the porcine rubulavirus LPMV possesses a structure organized predominantly in a β -loop- β arrangement. The highly preserved patterns and homology between the amino acid sequences observed in the paramyxoviruses analyzed, suggests that the HN from the rubulavirus and paramyxovirus genera are well preserved proteins throughout evolution; moreover, their common structural organization could explain similarities in the recognition mechanisms of sugar residues on the target cell.

Acknowledgements

This work was supported in part by PAPIIT Program (IN20995) UNAM; and the Consejo Nacional de Ciencia y Tecnologia (2151PM and F643-M9406), Mexico.

References

1. Stephano HA, Gay M, Ramírez TC (1988) Encephalomyelitis, reproductive failure and corneal opacity (blue eye) in pigs, associated with Paramyxovirus infection. *Vet Rec* 122: 6–10
2. Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) (1995) *Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses*. Springer, Wien New York (Arch Virol [Suppl] 10)

3. Kingsbury DW (1990) Paramyxovirus and their replication. In: Fields BN, Knipe DM (eds) *Virology*. Raven Press, New York, pp 945–961
4. Reyes Leyva J, Hernandez-Jauregui P, Montaña LF, Zenteno E (1993) The porcine Paramyxovirus LPMV specifically recognizes sialyl (α 2,3) lactose-containing structure. *Arch Virol* 133: 195–200
5. Reyes-Leyva J, Espinosa B, Hernández J, Zenteno R, Vallejo V, Hernandez-Jauregui P, Zenteno E (1997) NeuAc α 2–3 Gal-glycoconjugate expression determines cell susceptibility to the porcine rubulavirus LPMV. *Comp Biochem Physiol* 118B (in press)
6. Sundqvist A, Berg M, Hernandez-Jauregui P, Linné T, Moreno-López J (1990) The structural proteins of a porcine paramyxovirus (LMPV). *J Gen Virol* 71: 609–613
7. Linné T, Berg M, Bergvall AC, Hjertner B, Moreno-Lopez J (1992) The molecular biology of the porcine paramyxovirus LPMV *Vet Microbiol* 33: 263–273
8. Sundqvist A, Berg M, Moreno-Lopez J, Linné T (1992) The hemagglutinin-neuraminidase glycoprotein of the porcine paramyxovirus LPMV: comparison with other paramyxoviruses revealed the closest relationship to Simian virus 5 and mumps virus. *Arch Virol* 122: 331–340
9. Blumberg B, Giorgi C, Roux L, Raju K, Dowling P, Cholet A, Kolakofsky D (1985) Sequence determination of the virus HN gene and its comparison to influenza glycoproteins. *Cell* 41: 269–278
10. Colman PM, Hoyne PA, Lawrence MC (1993) Sequence and structure alignment of paramyxovirus hemagglutinin-neuraminidase with the influenza virus neuraminidase. *J Virol* 67: 2972–2980
11. Yewdell J, Gerhard W (1982) Delineation of four antigenic sites on a paramyxovirus glycoprotein via which monoclonal antibodies mediate distinct antiviral activities. *J Immunol* 128: 2670–2675
12. Portner A, Scroggs RA, Metzger DW (1987) Distinct functions of antigenic sites of the HN glycoprotein of Sendai virus. *Virology* 158: 61–68
13. Deshpande KL, Portner A (1984) Structural and functional analysis of Sendai virus nucleocapsid protein NP with monoclonal antibodies. *Virology* 139: 32–42
14. Deshpande KL, Portner A (1985) Monoclonal antibodies to the protein HN of Sendai virus define its structure and role in transcription. *Virology* 140: 125–134
15. Baiocchi M, Pescarmona M, Gallina A, Dalocchio F, Tomasi M (1993) Simple purification and secondary structure evaluation of the Sendai virus neuraminidase water soluble fragment. *Biochem Mol Biol Int* 31: 389–398
16. Murty FKG, Takimoto T, Laver WG, Portner A (1993) Crystals of hemagglutinin neuraminidase of parainfluenza virus contain triple-stranded helices. *Proc Natl Acad Sci USA* 90: 1523–1525
17. Takimoto T, Laver WG, Murti KG, Portner A (1992) Crystallization of biologically active hemagglutinin-neuraminidase glycoprotein dimers proteolytically cleaved from human parainfluenza virus type 1. *J Virol* 66: 7597–7600
18. Crenell SJ, Portner A, Takimoto T, Laver WG, Taylor GL (1996) Towards the structure of hemagglutinin-neuraminidase from the Newcastle disease virus. XVII. Congress and General assembly of the International Union of Crystallography, Washington
19. Chou PY, Fasman GD (1974) Prediction of protein conformation. *Biochemistry* 13: 222–244
20. Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120: 97–120
21. Deleage G, Roux B (1989) Use of class prediction to improve protein secondary structure

- prediction. In: Fasman GD (ed) Prediction of protein structure and the principles of protein conformation, vol. 13. Plenum Press, New York, pp 587–597
22. Levin JM, Robson B, Garnier J (1986) An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Lett* 205: 303–308
 23. Kabsch W, Sanders C (1983) Dictionary of protein secondary structural patterns, recognition of hydrogen bonded and geometrical features. *Biopolymer* 22: 2 577–2 637
 24. Rost B, Sanders C (1993) Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 232: 584–99
 25. Rost B, Sanders C (1994) Combining evolutionary information and neural networks to predict protein secondary structures. *Protein* 19: 55–72
 26. Sanders C, Rost B (1993) Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc Natl Acad Sci USA* 90: 7 558–7 562
 27. Rost B, Sanders C, Schneider R (1994) PHD – an automatic mail server for protein secondary structure prediction. *Comput Appl Biosci* 10: 53–60
 28. Richardson JS, Richardson DC (1988). Amino acid preferences for specific locations at the ends of alpha helices. *Science* 240: 1 648–1 652
 29. Chou PY (1989) In: Fasman GD (ed) Prediction of protein structure and the principles of protein conformation, vol 12. Plenum Press, New York, pp 549–586
 30. Flöckner H, Braxenthaler M, Lackner P, Jaritz M, Hortner M, Sippl MJ (1995) Progress in fold recognition. *Proteins* 23: 376–386
 31. Rost B (1995) In: Rawlings C, Clark D, Altman R, Hunter L, Lengauer T, Wodak S (eds) “The third International Conference on Intelligent Systems for Molecular Biology (ISMB)”, Cambridge, UK, July 16–19, AAAI Press, Menlo Park, pp 314–321
 32. Karplus PA, Schuls GE (1985) Prediction of chain flexibility in proteins. *Naturwissenschaften* 72: 212–213
 33. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–132
 34. Eisenberg D, Schwartz E, Komaromy M, Wall R (1984) Analysis of membrane and surface protein sequences with hydrophobic moment plot. *J Mol Biol* 179: 125–142
 35. Roa M, Argos A (1986) A conformational preference parameter to predict helices in integral membrane proteins. *Biochim Biophys Acta* 869: 197–214
 36. Klein P, Kanehisa M, Delisi C (1985) The detection and classification of membrane spanning proteins. *Biochim Biophys Acta* 815: 468–476
 37. Bairoch A (1992) PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res* 20: 2 013–2 018
 38. Higgins DG, Sharp PM (1989) Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* 5: 151–15 332
 39. Deleage G, Clerc FF, Roux B, Gautheron DC (1988) Antheprot: A package for protein sequence analysis using a microcomputer. *Comp Biosci* 4: 351–356
 40. Almagro JC, Vargaz Madrazo E, Zenteno-Cuevas R, Hernandez-Mendiola V, Lara Ochoa F (1995) VIR: A computational tool for analysis of immunoglobulin sequences. *Bio Systems* 35: 25–32
 41. Wu TT, Kabat EA (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med* 132: 211–250
 42. Jores R, Alzari PM, Meo T (1990) Resolution of hypervariable regions in T-cell receptor β -chains by a modified Wu-Kabat index of amino acids diversity. *Proc Natl Acad Sci USA* 87: 9 138–9 142
 43. Jimenez-Montaña MA (1984) On the syntactic structure of protein sequence and the concept of grammer complexity. *Bull Math Biol* 46: 641–659

44. Morrison TG, Portner A (1991) Structure, function, and intracellular processing of the glycoproteins of Paramyxoviridae. In: Kingsbury DW (ed) *The paramyxovirus*. Plenum Press, New York, pp 347–382
45. Sergel T, McGinnes L, Morrison T (1993) Role of a conserved sequence in the maturation and function of the NDV HN glycoprotein. *Virus Res* 30: 281–294
46. Iorio RM, Glickman RL, Sheehan JP (1992) Inhibition of fusion by neutralizing antibodies to the haemagglutinin-neuraminidase glycoprotein of Newcastle disease virus. *J Gen Virol* 73: 1 167–1 176
47. Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN, Rosenbusch JP (1992) Crystal structures explain functional properties of two *E. coli* porins. *Nature* 358: 727–733
48. Parker M, W, Buckley J, Postma PMJ, Tucker A, Leonard K, Pattur F, Tsernoglou D (1994) Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states. *Nature* 367: 292–295
49. Lamb RA (1993) Paramyxovirus fusion: A hypothesis for changes. *Virology* 197: 1–11
50. Segev Y, Rager-Zisman B, Isakov N, Schneider-Chaulies S, ter Meulen V, Udem S, Segal S, Wolfson M (1994) Reversal of the measles virus increase of phosphorylating activity in persistently infected mouse neuroblastoma cells by anti-measles virus antibodies. *J Gen Virol* 75: 819–827
51. Mirza A, Deng R, Iorio M (1994) Site-directed mutagenesis of a conserved hexapeptide in the paramyxovirus hemagglutinin-neuraminidase glycoprotein effects on antigenic structure and function. *J Virol* 68: 5 093–5 099

Authors' address: Dr. E. Zenteno, Dep. Bioquímica, Facultad de Medicina, UNAM, P.O. Box 70159, 04510 México.

Received January 23, 1997