N. Veljkovic
D.R. Branch
R. Metlas
J. Prljic
K. Vlahovicek
S. Pongor
V. Veljkovic

Design of peptide mimetics of HIV-1 gp120 for prevention and therapy of HIV disease

Authors' affiliations:
N. Veljkovic, J. Prljic and V. Veljkovic, Center for Multidisciplinary Research, Institute of Nuclear Sciences VINCA, P.O. Box 522, 11001 Belgrade, Yugoslavia.

D.R. Branch, Canadian Blood Services and University of Toronto, 67 College Street, Toronto, Ontario M5G 2M1, Canada.

R. Metlas, Diapharm, Quay House, South Explanade, St. Peterport, Guernsey, Channel Island GY1 4EJ, UK.

K. Vlahovicek and S. Pongor, Protein Structure and Function Group, International Center for Genetic Engineering and Biotechnology, Padriciano 99, 34012 Trieste, Italy.

Correspondence to:
Donald R. Branch
Research & Development, Canadian Blood Services, 67 College Street, Toronto, Ontario M5G 2M1, Canada
Tel: (416)-313-4458
Fax: (416)-974-9757
E-mail: don.branch@utoronto.ca

Dates:
Received 13 March 2003
Revised 22 May 2003
Accepted 11 June 2003

To cite this article:
Veljkovic, N., Branch, D.R., Metlas, R., Prljic, J., Vlahovicek, K., Pongor, S. & Veljkovic, V. Design of peptide mimetics of HIV-1 gp120 for prevention and therapy of HIV disease.
J. Peptide Res., 2003, **62,** 158–166.

ISSN 1397–002X

Key words: drug design; gp120; HIV-1; peptide mimetic

Abstract: It has been reported that the C-terminus of the second conserved region (C2) of the envelope glycoprotein gp120, encompassing peptide RSANFTDNAKTIIVQLNESVEIN (NTM), is important for infectivity and neutralization of the human immunodeficiency virus type 1 (HIV-1). It was also demonstrated that human natural anti-vasoactive intestinal peptide (VIP) antibodies reactive with this gp120 region play an important role in control of HIV disease progression. The bioinformatic analysis based on the time-frequency signal processing revealed non-obvious similarities between NTM and VIP. When tested against a battery of sera from 46 AIDS patients, these peptides, in spite of a significant difference in their primary structures, showed a similar reactivity profiles ($r = 0.83$). Presented results point out that similarity in the periodical pattern of some physicochemical properties in primary structures of peptides plays a significant role in determination of their immunological crossreactivity. Based on these findings, we propose this bioinformatic criterion be used for design of VIP/NTM peptide mimetics for prevention and treatment of HIV disease.

Introduction

Human immunodeficiency virus (HIV) disease progression varies greatly between individuals and it appears that host factors play an important role in determining the clinical outcome in HIV infection. In order to define these host factors, Neurath and co-workers have investigated antibody profiles in two groups of HIV-infected patients: those who remained healthy for at least 10 years and those who developed AIDS within 5 years of the onset of infection (1). They demonstrated that antibodies recognizing the peptide

RSANFTDNAKTIIVQLNESVEINCTRP (amino acids 280–306 within the C2 region of the envelope glycoprotein gp120 from the BH-10 isolate of HIV-1) are significantly more prevalent in asymptomatic carriers than in patients who progressed to AIDS (6/9 in asymptomatic vs. 0/9 in AIDS patients). Based on these results, it appears that the absence or disappearance of these antibodies my represent a possible factor contributing to disease progression. For this reason, it has been proposed that maintenance of a high level of these antibodies by immunotherapy, based on active immunization with antigens containing this peptide and/or administration of the corresponding antibodies, should be considered as a modality for therapy of HIV-1 infection. This assumption was strongly confirmed by recently reported results of therapy performed by passive immunization with human HIV-negative plasma enriched with antibodies reactive with the C2-derived peptide encompassing amino acids 280–302 (2).

Despite the presence of the strongest T-cell epitope of gp120, which is active *in vitro* (3–7) and an exposed B-cell epitope (8,9), the C-terminus of the C2 region encompassing amino acids 280–306 is not immunogenic in humans (3,10–12). Absence of the active B-cell epitope within this peptide indicates that antibodies in sera of HIV patients recognizing this region of HIV-1 gp120 represent autoreactive antibodies elicited by some human antigen. Vasoactive intestinal peptide (VIP) was identified as the human antigen likely inducing these natural antibodies, which are cross-reactive with peptide RSANFTDNAKTIIVQLNQSVEIN (denoted as peptide NTM) derived from the C2 region of HIV-1 gp120 (13,14).

One might conclude from these previous data that this conserved area of HIV-1 gp120 cannot be used as a vaccine component because the human immune system is unresponsive, or tolerant, to epitope(s) within this part of the molecule. In order to overcome this problem in the development of a possible vaccine, it would be necessary to design antigens which are crossreactive with the C-terminus of the C2 region and, which can escape the immune tolerance.

It has been suggested that the so-called electron–ion interaction potential (EIIP) plot of peptide sequences, and its Fourier transform (informational spectrum method, ISM) may be a good general predictor of peptide–ligand interactions (15). The present work was undertaken in order to test the hypothesis that peptide analogs that have a predetermined immunological profile, but bearing little or no sequence similarity with the native peptide antigen, can be designed by ISM. We demonstrated that two peptides with different primary sequences but closely related information spectra may possess similar immune reactivity, while peptides with very similar primary sequences but different information spectra can have quite different immune profiles. Based on these results, we defined the bioinformatics criterion for design of gp120/VIP peptide mimetics for prevention and therapy of HIV disease.

Materials and Methods

Bioinformatic analysis of peptides

Distribution of periodical patterns and tandem repeats of residues in protein and DNA sequences determine structural and functional characteristics of the molecules. The ISM is a virtual spectroscopy technique, which allows investigation of the periodicity of structural motifs with defined physicochemical characteristics, which determine biological properties of protein and DNA sequences. Physical and mathematical basis of ISM is described in detail elsewhere (15–18); here we will only briefly present this bioinformatic method. A sequence of N residues is represented as a linear array of N terms, with each term given a weight. The weight assigned to a residue is EIIP (19,20), determining electronic properties of amino acids and nucleotides, which are responsible for their intermolecular interactions (21,22). In this way, the alphabetic code is transformed into a sequence of numbers. The obtained numerical sequence, representing the primary structure of protein, is then subjected to a discrete Fourier transformation, which is defined as follows:

$$X(n) = \sum x(m)\mathrm{e}^{-\mathrm{j}(2/N)nm}, \qquad n = 1, 2, \ldots, N/2 \qquad (1)$$

where $x(m)$ is the mth member of a given numerical series, N is the total number of points in this series, and $X(n)$ are discrete Fourier transformation coefficients. These coefficients describe the amplitude, phase and frequency of sinusoids, which comprised the original signal. The absolute value of complex discrete Fourier transformation defines the amplitude spectrum and the phase spectrum. The complete information about the original sequence is contained in both spectral functions. However, in the case of protein analysis, relevant information is presented in an energy density spectrum (15,16), which is defined as follows:

$$S(n) = X(n)X*(n) = |X(n)|^2, \qquad n = 1, 2, \ldots, N/2 \qquad (2)$$

In this way, sequences are analyzed as discrete signals. It is assumed that their points are equidistant with the distance

$d = 1$. The maximal frequency in a spectrum defined in this way is $F = 1/2d = 0.5$. The frequency range is independent of the total number of points in the sequence. The total number of points in a sequence influences only resolution of the spectrum. The resolution of the N -point sequence is $1/n$. The nth point in the spectral function corresponds to a frequency $f(n) = nf = n/N$. Thus, the initial information defined by the sequence of amino acids can now be presented in the form of the informational spectrum (IS), representing the series of frequencies and their amplitudes.

The IS frequencies correspond to distribution of structural motifs with defined physicochemical properties determining a biological function of a protein. When comparing proteins, which share the same biological or biochemical function, the ISM technique allows detection of code/frequency pairs which are specific for their common biological properties, or which correlate with their specific interaction. This common informational characteristic of sequences is determined by cross-spectrum or consensus informational spectrum (CIS). A CIS of N spectra is obtained by the following equation:

$$C(j) = \prod S(i,j) \qquad (3)$$

where (i,j) is the jth element of the ith power spectrum and $C(j)$ is the jth element of CIS. Thus, CIS is the Fourier transform of the correlation function for the spectrum. In this way, any spectral component (frequency) not present in all compared informational spectra is eliminated. Peak frequencies in CIS are common frequency components for the analyzed sequences. A measure of similarity for each peak is a signal-to-noise ratio (S/N), which represents a ratio between signal intensity at one particular IS frequency and the main value of the whole spectrum. If one calculates a CIS for a group of proteins, which have different primary structures, and finds strictly defined peak frequencies, it indicates that the analyzed proteins participate in mutual interaction or have a common biological function.

The ISM has been successfully applied in structure-function analyses of different protein sequences, as well as in de novo design of biologically active peptides (23–36).

Multiple alignment and dendograms

The multiple alignment of the sequences was carried out with the CLUSTALW program (37). The dendrograms of the immunological profiles and information spectra were based on a simple Eucledian distance of the form $D = \sqrt{\sum_{i=1}^{n} |(y_i - x_i)|^2}$ where x_i and y_i are the immunoreactivity data against serum i (or the spectrum amplitude at frequency i) for peptide x and y, respectively. The dendrograms were drawn by the program PHYLIP using the nearest neighbor method (38).

Human subjects and their sera

Serum samples were collected from 46 HIV-positive subjects (CD_4 count, 50–600/μL). Samples from 10 healthy HIV-negative subjects were randomly selected from individuals referred to the blood donor service.

Peptides

Peptides NTM (RSANFTDNAKTIIVQLNESVEIN), NTMM (RSANFTDNAKTHVQLNESVEIN), $_{p}32$ (RSAHFTDNAKTPESVEIP) and V3 (KKGIAIGPGRTLY) were synthesized by solid-phase technology by Sigma Chemicals (St. Louis, MO). VIP (1-28) (HSDAVFTDNYTRLRKQMAVKKYLNSILN) was purchased from Sigma Chemical. All peptides except V3 were coupled to bovine serum albumin (BSA) with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC), while V3 peptide was crosslinked to BSA with glutaraldehyde.

ELISA assay

Polystyrene microtiter plates (Greiner, Germany) were incubated overnight at 4 °C with 100 μL of BSA-coupled peptides (1 μg/well) diluted in carbonate buffer, pH 9.6. Plates were washed with phosphate-buffered saline (PBS)–0.05% Tween and non-specific sites were blocked with 200 μL PBS containing 1% BSA for 1 h at room temperature. After further washing, serum specimens were then added to the wells (100 μL/well). Sera were diluted in 0.1% BSA in PBS. Plates were incubated for 1 h at room temperature. After three washings with PBS–0.05% Tween, 100 μL of goat anti-human IgG peroxidase-conjugated anti-

Figure 1. Peptides used in this study. The multiple alignments were produced with the CLUSTALW program (37) using gap open and gap extension penalties of 10 and 0.05, respectively. Black shading denotes >50% identity while gray shading denotes >50% similarity.

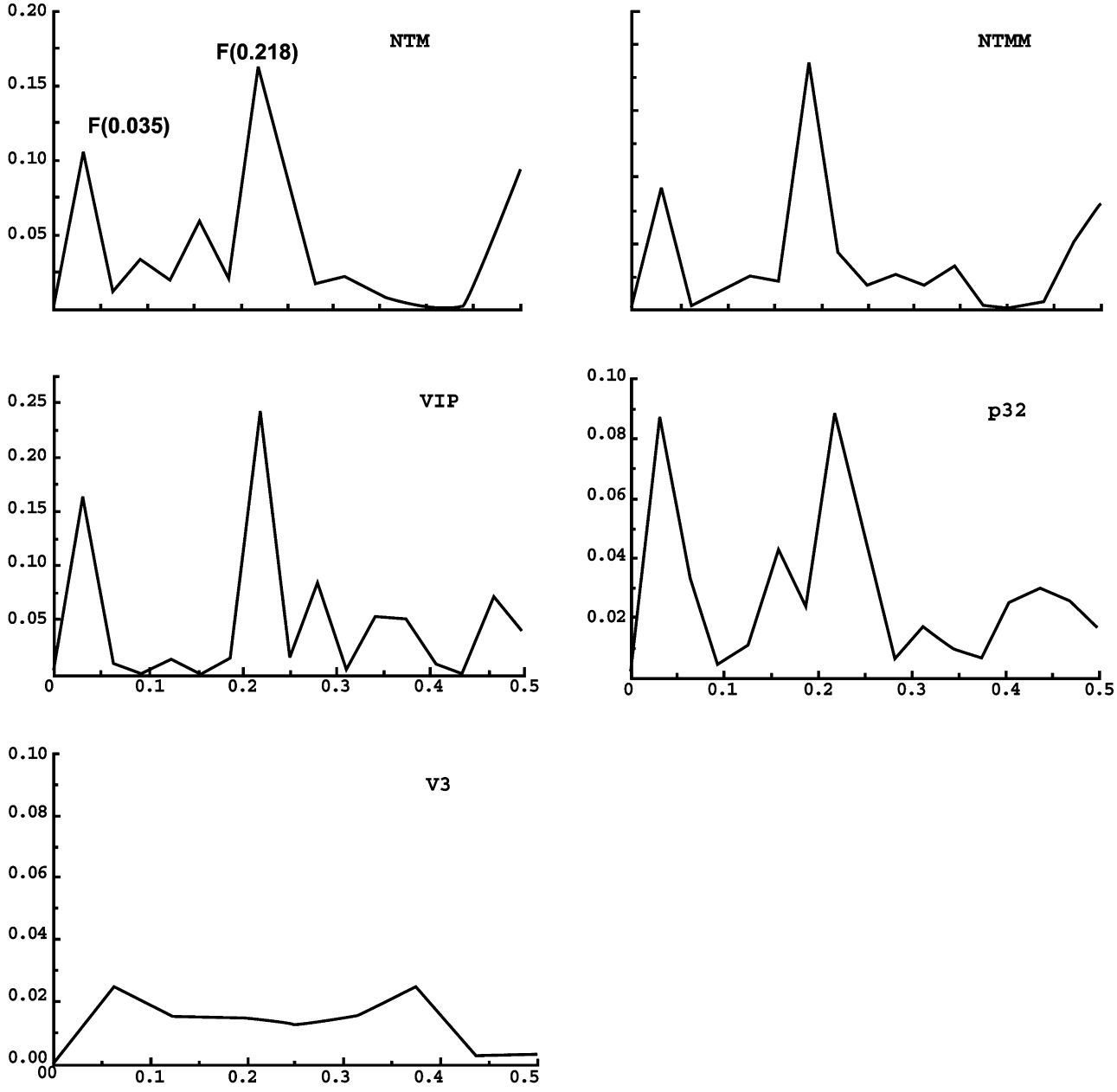

Figure 2. Informational spectra for peptides NTM, NTMM, VIP, $_{p}32$ and V3. For each spectrum, the abscissa represents the frequencies from the Fourier transform of the sequence of electron–ion interaction potential corresponding to the amino acid sequence of the peptide. The lowest frequency is 0.0 and the highest is 0.5. The ordinate represents amplitudes, in arbitrary units, corresponding to each frequency component in the spectrum.

bodies (Sigma), diluted 1 : 5000 was added and the plates were incubated for 1 h. After five washings, the 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonate) (ABTS) substrate was added and the absorbance measured at 405 nm.

Affinity purification of antibodies

One milliliter of HIV^{+} serum from a single donor (diluted with TBS, 2.5 mL TBS/mL gel) was chromatographed on an affinity column consisting of VIP-BSA coupled to CNBr-activated Sepharose 4B (2 mg peptide/5 mL of resin). After the samples had been loaded, the gel was washed with five column volumes of PBS. Washing was continued with 10 column volumes of 20 mM Tris–HCl, pH 7.4 containing 0.5 M NaCl and 0.2% Triton X-100. Washing was repeated with TBS until the base line was stable. Bound antibody was eluted with 0.2 M glycine-HCl pH 2.5 containing 0.15 M NaCl and neutralized immediately with 1 M Tris–HCl pH 9.00.

Table 1. **Immunological reactivities (O.D) of the peptides as determined against 46 HIV-positive sera**

Serum no.	NTM	NTMM	p32	VIP	V3
1	0.361	0.093	0.346	0.433	0.087
2	0.653	0.200	0.582	0.816	0.190
3	0.595	0.432	0.350	0.621	0.308
4	0.897	0.222	1.000	1.010	0.266
5	0.752	0.228	0.586	0.902	0.245
6	0.800	0.321	0.602	0.895	0.195
7	0.398	0.189	0.258	0.298	0.297
8	0.559	0.195	0.547	0.671	0.451
9	0.408	0.174	0.104	0.450	0.291
10	0.373	0.294	0.200	0.450	0.291
11	0.464	0.206	0.378	0.557	0.503
12	0.388	0.135	0.309	0.469	0.417
13	0.470	0.265	0.369	0.560	0.543
14	0.427	0.245	0.421	0.620	0.551
15	0.480	0.256	0.643	0.624	0.591
16	0.420	0.295	0.369	0.546	0.690
17	0.436	0.348	0.334	0.532	0.701
18	0.551	0.255	0.572	0.690	0.563
19	0.477	0.221	0.375	0.578	0.945
20	0.491	0.259	0.329	0.705	0.368
21	0.433	0.335	0.256	0.371	0.258
22	0.509	0.193	0.579	0.631	0.431
23	0.329	0.278	0.258	0.385	0.600
24	0.377	0.111	0.127	0.565	0.418
25	0.497	0.198	0.575	0.606	0.474
26	0.433	0.137	0.377	0.572	0.454
27	0.379	0.225	0.307	0.473	0.519
28	0.363	0.269	0.282	0.360	0.558
29	0.430	0.217	0.177	0.495	0.624
30	0.407	0.319	0.669	0.700	0.763
31	0.442	0.351	0.353	0.516	0.332
32	0.409	0.124	0.315	0.469	0.869
33	0.492	0.233	0.270	0.300	0.358
34	0.432	0.206	0.361	0.532	0.337
35	0.336	0.227	0.277	0.402	0.413
36	0.331	0.192	0.284	0.397	0.465
37	0.503	0.351	0.385	0.593	0.299
38	0.395	0.177	0.170	0.330	0.515
39	0.630	0.330	0.868	0.650	0.387
40	0.499	0.205	0.582	0.596	0.340
41	0.408	0.184	0.302	0.372	0.779
42	0.398	0.224	0.120	0.489	0.347
43	0.304	0.274	0.302	0.372	0.779
44	0.413	0.257	0.122	0.495	0.461
45	0.287	0.282	0.222	0.291	0.503
46	0.513	0.265	0.638	0.640	0.235

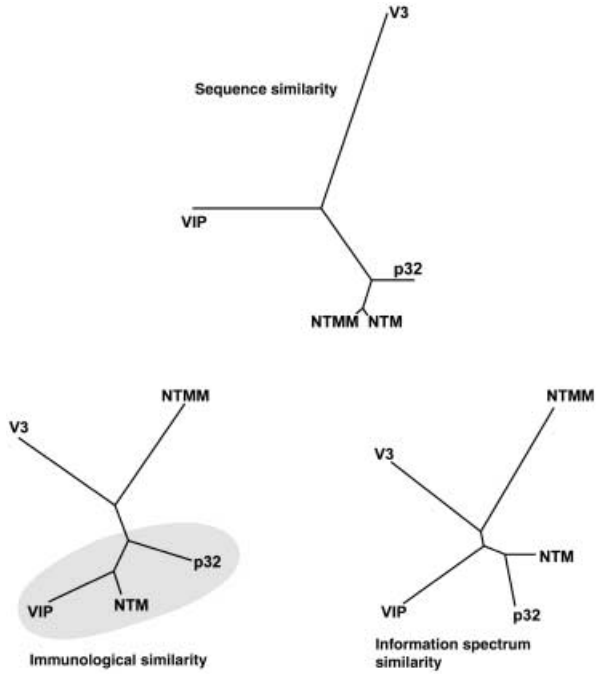

Figure 3. Comparison of the sequence-, immunological and information-spectrum similarities. Unrooted trees were produced with the PHYLIP program (38) as described in the text. The shaded area indicates that peptides NTM, VIP, and $_{p}32$ are closely related in terms of both immunological properties and information spectra.

Statistical analysis of results

Concordance between NTM reactivity in ELISA with HIV^{+} sera and reactivity of other peptides: VIP, NTMM, $_{p}32$ and V3 were measured by Pearson's correlation test (STATISTICA). The character of correlation is described by the correlation coefficient (r) in the following way: 1.00 > $|r|$ > 0.90, very strong correlation; 0.90 > $|r|$ > 0.70, strong; 0.70 > $|r|$ > 0.50, weak; 0.50 > $|r|$ > 0.00, very weak.

Results

The peptide sequences and their IS are shown in Figs 1 and 2, respectively. We used the peptide NTM, derived from the C2 region of the HIV-1 gp120 (13) as a native model antigen. Peptide NTMM was designed as an analog having 91% sequence identity with NTM but significantly different informational spectrum. It is important to note that peptides NTM and NTMM share a common sequence FTDN representing an exposed B-cell epitope (9). Previously we showed that VIP and NTM, despite significantly different primary structures, have very similar IS (13). It has been also shown that NTM and VIP are immunologically cross-reactive (14). For this reason VIP was used as a natural

Table 2. **Functional role of the HIV-1 gp120 C2 region encompassing the peptide NTM**

Function	References
The amino acids critical for the interaction of gp120 with CD4 have been localized to peptide ANFTDNAKTI within the C2 region. Four amino acids from this domain, Asp279, Asn280, Ala281 and Thr283, represent the residues of gp120 in direct contact with the CD4 receptor	Veljkovic, V. & Metlas, R. *Cancer Biochem. Biophys.* 1998; **10**: 191; Pollard, R.S. *et al. Proc. Natl. Acad. Sci. USA* 1991; **88**: 11320; Moore, J.P. *et al. J. Virol.* 1994; **68**: 469; Kwong, P.D. *et al. Nature* 1998; **393**: 648; Wyatt, R. *et al. Nature* 1998; **393**: 705
The region of gp120 extending from 278 to 283 (TDNAK) plays an important role in the early event of viral infection at the level of post-CD4 binding	Turner, S. *et al. Proc. Natl Acad. Sci. USA* 1992; **89**: 1335; Stamatos, L., Cheng-Mayer, C.J. *Virology* 1993; **67**: 5635; Corbeau, P. *et al. Mol. Immunol.* 1993; **31**: 569; Branch, D. R. *et al. AIDS* 2002; **16**: 309
Mutations in C-terminal part of C2 affect the biological phenotype of mutant viruses	Willey, R.L. *et al. J. Virol.* 1988; **62**: 139; Willey, R.L. *et al. J. Virol.* 1993; **67**: 3639; Sabino, E. *et al. AIDS* 1994; **8**: 901
The C2 region of gp120 including its C-terminus adjacent to V3 loop participates in interaction with chemokine receptors	Schols, D. *et al. J. Virol.* 1998; **72**: 4032; Dumanceaux, J. *et al. J. Gen. Virol.* 1999; **80**: 1975; Shimizu, N. & Gojobori, T. *Gene* 2000; **559**: 109; Hamond, A.I. *et al. J. Virol.* 2001; **75**: 5593; Cormier, E.G. *et al. J. Virol.* 2001; **75**: 5541; Nehete, P.N. *et al. Antivir. Res.* 2002: **56**: 233
The sequence EEVVIRSANFTDNAKT is involved in intermolecular interaction within the oligomeric env complex	Lemasson, I. *et al. AIDS Res. Hum. Retrovir.* 1995; **11**: 1177; Seddiki, N. *et al. Biochem. Biophys. Acta* 1997; **1340**: 277
Shielding of the highly conserved region IRSANFTDNAKTIIVQLNQS, participating in gp120/CD4 interaction, with variable V3 loop provides an important viral defense against immune survivalence	D'Costa, S. *et al. Aids Res. Hum. Retrovir.* 2001; **17**: 1205
Absence of N-linked glycosylation site within sequence VIRSANFTDNAKT is crucial for formation of the noncytopathic HIV-1 variant which is unable to induce single-cell killing	Stevenson, M. *et al. J. Virol.* 1990; **64**: 3792
The peptide derived from the jacalin alpha-chain which is homologous with gp120-derived peptide VVIRSANFTDNAKT effectively blocks HIV-1 infectivity *in vitro* and *in vivo* without impairing the lymphocyte function	Favero, J. *et al. Eur. J. Immunol.* 1993; **23**: 179; Lafont, V. *et al. Immunol. Lett.* 1993; **37**: 249; Carbeau, P. *et al. Mol. Immunol.* 1994; **31**: 569; Tamma, S.M.L. *et al. Clin. Immunol. Pathol.* 1996; **80**: 290; Lafont, V. *et al. J. Leukocyte Biol.* 1996; **59**: 691
The peptide QFTDNAKTIIVQLNQS is identified as a CD36-related thrombospondin 1 (TSP-1) binding domain in the HIV-1 p120. Binding of TSP-1 to gp120 significantly retard HIV-1 infectivity	Crombie, R. *et al. J. Exp. Med.* 1998; **187**: 25
The peptide derived from C-terminus of the C2 region blocks gp120/CCR5 chemokine receptor-mediated chemotaxis	Redvine, L.S. *et al. Clin. Immunol.* 1999; **93**: 124
The C2 region of gp120 encompassing NTM peptide is responsible for HIV/co-receptor interaction and the viral tropism	Hammond, A.L. *et al. J. Virol.* 2001: **75**: 5593; Dumonceaux, J. *et al. J. Virol.* 2001; **75**: 5425
Mutations within the sequence ENFTD of gp120 contribute to induction of the AIDS dementia complex (ADC)	Smith, M.K. *et al. J. Neuro Virol.* 2000; **6**: 164
Sequence SVEINC derived from the C-terminus of peptide NTM induces Fas/FasL mediated apoptosis	Zaguri, J.F. *et al. Biomed Pharmacother.* 1993; **47**: 331; Szawlowski, P.W. *et al. AIDS* 1993; **7**: 1018
The region of the gp120 gene encoding peptide VQLN contains Chi promoter of homologous recombination which is active *in vivo* and contributes to HIV variability	Metlas, R. *et al. Biochem. Biophys. Res. Commun.* 1991; **179**: 1056; Veljkovic, V. & Metlas, R. *Immunol. Lett.* 1991; **31**: 11; Prljic, J. *et al. Vaccine* 1999; 17: 1462
Expression of VPAC-1 receptor in CHO cells that recognizes DAVFTDNYT results in HIV-1 infection and integration without a requirement for CD4 or chemokine receptor	Branch, D.R. *et al.* (results submitted for publication)

Table 3. **Immunological properties of peptide NTM**

Immunological properties	References
mAB 110c directed against peptide SANFTD (fine mapped to FTD) and MO20.13.2 directed against the same peptide bound as well to native gp 120 as to denaturated gp 120	Moore, J.P. *et al. J. Virol.* 1994; **68**: 469; Lemasson, I. *et al. AIDS Res. Hum. Retrovir* 1995; **11**: 1177
Resistance of HIV-1 to neutralization by natural antisera occurs through single aminoacid substitution (A281V) within region of C2 extended from 271 to 288 (VIRSVNFTDNAKTIIVQL)	Watkins, B.A. *et al. J. Virol.* 1993: **67**: 7493; Watkins, BA. *et al. J. Virol.* 1996; **70**: 8431
Direct screening of human HIV-1-positive serum by the random fragment expression library revealed B epitope located within peptide VQLNQSVEINCTRPNNNTRKSI	Kusk, P. *et al. AIDS* 1992; **6**: 1451
An unusual set of local homologies between peptide NTM and human proteins indicates that this region of gp 120 contains one of the epitopes around which the human immune network is organized	Veljkovic, V. *et al. Cancer J.* 1995; **8**: 308
Peptide SLAEEEVVIRSANFTDNAKTIIVQ contains T-cell epitope recognized by HIV-1-seropositive and by low risk HIV-1-seronegative individuals	Wahren, B. *et al. J. Acquir. Immune. Defic. Syndr.* 1989; **2**: 448; Mutch, D. *et al. J. Acquir. Immune. Defic. Syndr.* 1994; **7**: 879; Geretti, A.M. *et al. Scand. J. Immunol.* 1994; **39**: 355; Sitz, K.V. *et al. J. Infect. Dis.* 1999; **179**: 817; Mathiesen, T. *et al. Immunology* 1989; **67**: 453; Veljkovic, V. & Methlas, R. *Immunol. Lett.* 1990; **26**: 193; Bradac, J.A. & Mathieson, B.J. An epitope map of immunity to HIV-1: a roadmap for vaccine Development. NIAI, NIH, Bethesda (1991)
Antibodies recognizing peptide RSANFTDNAKTIIVQLNESVENCTRP are significantly more prevalent in sera of asymptomatic carriers than in AIDS patients	Neurath, A.R. *et al. AIDS Res. Human Retrovir.* 1990; **6**: 1183
Antibodies affinity purified on VIP from HIV-positive sera strongly react with peptide NTM	Veljkovic, V. *et al. Biochem. Biophys. Res. Commun.* 1993; **196**: 1019
It has been demonstrated strong therapeutic potential of HIV-negative plasma enriched by NTM-reactive natural antibodies	Vejkovic, V. *et al. Chest* 2001; **120**: 662
Signal-inhibiting antibody directed to the receptor for the ligand sequence DAVFTDDNYT that is shared by NTM inhibits HIV infection	Branch, D.R. *et al. AIDS* **16**: 309

mimetic of NTM. Peptide $_{p}32$ was designed as an analog of NTM with low homology but a similar informational spectrum. The peptide V3 derived from the third hypervariable region of HIV-1 gp120, was used as an unrelated control.

It is conspicuous from Fig. 2 that NTM and VIP, i.e. two peptides of quite different primary structure show two common peaks corresponding to the frequency components $F(0.218)$ and $F(0.035)$, respectively. On the contrary, NTMM, despite a high degree of homology with NTM demonstrates only one of these peaks $F(0.035)$ while the second dominant frequency component $F(0.218)$ is replaced with $F(0.185)$. In other words, substitution of 12 Ile and 13 Ile by one His in NTM has abolished the spectrum characteristics common to NTM and VIP. Peptide $_{p}32$, a rationally designed analog has an information spectrum similar to that of NTM, in spite of the low sequence similarity. Finally, peptide V3 is completely different from NTM both in sequence and IS.

The immunoreactivities of the peptides were determined with a battery of 46 sera obtained from HIV positive patients (Table 1). Linear correlation coefficients were calculated between the data for each peptide (Table 1). A strong immunological crossreactivity was found between NTM and VIP ($r = 0.87$), which is in good agreement with our previous results on HIV-negative sera (2). On the other hand, crossreactivity between NTM and NTMM ($r = 0.19$), as well as the control peptide V3 ($r = -0.38$) is very weak. Peptide $_{p}32$ is highly crossreactive with NTM ($r = 0.75$), even though it has much less sequence similarity to it than does NTMM (56 vs. 91%).

The comparison of these peptides has been graphically summarized in the form of unrooted trees (Fig. 3). In this

representation, the length of the branches connecting two peptides is proportional to their similarity. It is conspicuous that according to the immunological properties and the information spectra: NTM, $_{p}32$ and VIP form one cluster while NTMM and V3 are distant, both from each other and from the cluster. On the other hand, the sequence similarity dendrogram is in contradiction with the immunoreactivity data, as it shows a high similarity between NTM and NTMM.

Discussion

Analysis of the periodical patterns of defined physicochemical property (the charge, the hydrophobicity, the bulkiness, EIIP, etc.) in the primary structure of protein sequences can provide insight into function of biological macromolecules and can also lead to knowledge regarding their biologically active sites. While analysis of protein sequences is often performed directly on the symbolic representation of the amino acid sequence, patterns in the sequence are often too weak to be detected using only pattern of symbols. For this reason, we used ISM in order to understand the role which information determined by periodic distribution of EIIP plays in determination of the immunological properties of peptides. We have used peptides derived from the C-terminus of the C2 region of the HIV-1 gp120 as a model system. This region of the HIV-1 envelope protein is crucial for several important functional and immunological properties of HIV-1 (see Tables 2 and 3). It has been previously demonstrated that antibodies in sera of HIV-infected patients reacting with the C-terminus of the C2 region strongly correlate with disease progression (1). These results indicate that the C2-derived peptides may be promising antigens for the development of a preventive and therapeutic AIDS vaccine, as well as potential AIDS therapeutics (2). We have proposed, based on the ISM analysis of this gp120 region, that antibodies from HIV-positive sera that bind peptide NTM, may, in fact, represent natural anti-VIP autoantibodies (13,14). Recently, we also showed a significant therapeutic potential of these antibodies (2).

Here we demonstrate that information encoded by primary structures of VIP, NTM and their mimetics, and represented by spectral components $F(0.035)$ and $F(0.218)$ in IS of these peptides, seem to be correlated with their immunological crossreactivity. The results presented strongly suggest that informational similarity between investigated peptides is more important for their immunological crossreactivity than their sequence similarity. This conclusion is in accordance with recently reported results demonstrating an important role of the IS in design of artificial peptide antigens which are immunologically crossreactive with the HIV-1 envelope glycoprotein (36).

A corollary of this study is that the information spectrum approach can be used as a simple prediction model for designing VIP/NTM peptide mimetics, which could be used for the prevention and control of HIV disease. Peptides encoding information which is represented by IS frequency components $F(0.035)$ and $F(0.218)$ could be applied as: (i) components of preventive and therapeutic HIV vaccine; (ii) for purification of therapeutic antibodies for the passive immunization of HIV-infected patients; and (iii) as a component of ELISA-based prognostic tests for monitoring the level of protective VIP/NTM-reactive antibodies in HIV-infected patients.

Acknowledgments: This work was supported by the Ministry of Science of Serbia (Yugoslavia) (contract No. 1993) and Diapharm, Ltd.

References

1. Neurath, A.R., Strick, N., Tajlor, P., Rubinstain, P. & Stevans, C.E. (1990) Search for epitope-specific antibody responses to the human immunodeficiency virus (HIV-1) envelope glycoproteins signifying resistance to disease development. *AIDS Res. Hum. Retroviruses* **6**, 1183–1192.
2. Veljkovic, V., Metlas, R., Jevtovic, D.J. & Stringer, W. (2001) The role of passive immunization in HIV-positive patients: a case report. *Chest* **120**, 662–666.
3. Mathiesen, T., Broliden, P.A., Rosen, J. & Wahren, B. (1989) Mapping of IgG subclass and T-cell epitopes on HIV proteins by synthetic peptides. *Immunology* **67**, 453–459.
4. Wahren, B., Rosen, J., Sandstrom, E., Mathiesen, T., Modrow, S. & Wigzell, H. (1989) HIV-1 peptides induce a proliferative response in lymphocytes from infected persons. *J. Acquir. Immune Defic. Syndr.* **2**, 448–456.
5. Geretti, A.M., Van Baalen, C.A., Borleffs, J.C., Van Els, C.A. & Osterhaus, A.D. (1994) Kinetics and specificities of the T helper-cell response to gp120 in the asymptomatic stage of HIV-1 infection. *Scand. J. Immunol.* **39**, 355–362.
6. Mutch, D., Underwood, J., Geysen, M. & Rodda, S. (1994) Comprehensive T-cell epitope mapping of HIV-1env antigens reveals many areas recognized by HIV-1-seropositive and by low-risk HIV-1-seronegative individuals. *J. Acquir. Immune Defic. Syndr.* **7**, 879–890.

7. Sitz, K.V., Ratto-Kim, S., Hodgkins, A.S., Robb, M.L. & Birx, D.L. (1999) Proliferative responses to human immunodeficiency virus type 1 (HIV-1) gp120 peptides in HIV-1-infected individuals immunized with HIV-1 rgp120 or rgp160 compared with nonimmunized and uninfected controls. *J. Infect. Dis.* **179**, 817–824.
8. Kusk, P., Holmback, K., Lindhardt, B.O., Hilgaard, E.F. & Bugge, T.H. (1992) Mapping of two new human B-cell epitopes on HIV-1 gp120. *AIDS* **6**, 1451–1456.
9. Lemasson, I., Housset, V., Calas, B. & Devaux, C. (1995) Antigenic analysis of HIV type1 external envelope (Env) glycoprotein C2 region: implication for the structure of env. *AIDS Res. Hum. Retrovir.* **11**, 1177–1182.
10. Bradac, J.A. & Mathieson, B.J. (1991) *An Epitope Map of Immunity to HIV*-1: a Roadmap for Vaccine Development. Division of AIDS, NIAID, NIH, Bethesda.
11. Sastry, K.J. & Arlinghaus, R.B. (1991) Identification of T-cell epitopes without B-cell activity in the first and second conserved regions of the HIV env protein. *AIDS* **5**, 699–707.
12. Veljkovic, V. & Metlas, R. (1990) Sequence similarity between HIV-1 env protein gp120 and human proteins: a new hypothesis on protective antibody production. *Immunol. Lett.* **26**, 193–196.
13. Veljkovic, V., Metlas, R., Raspopovic, J. & Pongor, S. (1992) Spectral and sequencesimilarity between VIP and the second conserved region of HIV envelope glycoprotein gp120: possible consequences on prevention and therapy of AIDS. *Biochem. Biophys. Res. Commun.* **189**, 705–710.
14. Veljkovic, V., Metlas, R., Vojvodic, D., Cavor, L.J., Pejinovic, N., Dujic, A., Zakhariev, S., Guarnaccia, C. & Pongor, S. (1993) Natural autoantibodies cross-react with a peptide derived from the second conserved region of HIV-1 envelope glycoprotein gp120. *Biochem. Biophys. Res. Commun.* **196**, 1019–1024.
15. Veljkovic, V., Cosic, I., Dimitrijevic, B. & Lalovic, D. (1985) Is it possible to analyze DNA and protein sequence by the method of digital signal processing? *IEEE Trans. BME.* **32**, 337–341.
16. Veljkovic, V. & Cosic, I. (1987) A novel method of protein analysis for prediction of biological function. *Cancer Biochem. Biophys.* **9**, 139–148.
17. Lazovic, J. (1996) Selection of amino acid parameters for Fourier transform-based analysis of proteins. *Comp. Appl. Bio. Sci.* **16**, 553–558.
18. Cosic, I. (1997) *The Resonant Recognition Model of Macromolecular Bioreactivity: Theory & Application*. Birkhauser Verlag, Berlin.
19. Veljkovic, V. & Slavic, I. (1972) Simple general-model pseudopotential. *Phys. Rev. Lett.* **29**, 105–107.
20. Veljkovic, V. (1973) The dependence of the Fermi energy on the atomic number. *Phys. Lett.* **45A**, 41–42.
21. Veljkovic, V. (1980) *A Theoretical Approach to Preselection of Cancerogens and Chemical Carcinogenesis*. Gordon & Breach, New York.
22. Lalovic, D. & Veljkovic, V. (1990) The global average DNA base composition of coding regions may be determined by the electron-ion interaction potential. *BioSystems* **23**, 311–316.
23. Cosic, I., Nesic, D., Pavlovic, M. & Williams, R. (1986) Enhancer binding proteins predicted by informational spectrum method. *Biochem. Biophys. Res. Commun.* **141**, 831–838.
24. Skerl, V. & Pavlovic, M. (1988) Thymopoetins and long postsynaptic neurotoxins share common information in their primary structures. *FEBS Lett.* **239**, 1411–1416.
25. Veljkovic, V. & Metlas, R. (1988) Identification of nanopeptide from HTLV3, LAV and ARV-2 envelope gp120 determining binding to T4 cell surface protein. *Cancer Biochem. Biophys.* **10**, 191–196.
26. Cosic, I. & Nesic, D. (1988) Prediction of 'hot spots' in SV40 enhancer and relation with experimental data. *Eur. J. Biochem.* **170**, 247–252.
27. Cosic, I., Pavlovic, M. & Vojisavljevic, V. (1989) Prediction of 'hot spots' in interleukin-2 based on informational spectrum characteristics of growth regulating factor. *Biochemie* **71**, 333–339.
28. Cosic, I., Vojisavljevic, V. & Pavlovic, M. (1989) The relationship of the resonant recognition model to effects of low-intensity light on cell growth. *Int. J. Radiat. Biol.* **56**, 179–184.
29. Cosic, I. (1990) *Resonant Recognition Model of Protein–Protein and Protein–DNA Interaction in Bioinstrumentation and Biosensors*. Marcel Dekker, New York.
30. Cosic, I., Hodder, A., Aguilar, M. & Hearn, M. (1991) Resonant recognition model and protein topography: model studies with myoglobin, hemoglobin and lysozyme. *Eur. J. Biochem.* **198**, 113–119.
31. Cosic, I. & Hearn, M. (1991) 'Hot spot' amino acid distribution in Ha-ras oncogene product p21: relationship to guanine binding site. *J. Mol. Rec.* **4**, 57–62.
32. Cosic, I. (1994) Macromolecular bioactivity:is it resonant interaction between macromolecules? Theory and applications. *Trans. Biomed. Eng.* **12**, 1101–1108.
33. Cosic, I. & Drummond, A.E., Underwood, J.R. & Hearn, M. (1994) In vitro inhibition of the actions of basic FGF by a novel 16 amino acid peptide. *Mol. Cell. Biochem.* **130**, 1–7.
34. Veljkovic, V., Johnson, E. & Metlas, R. (1995) Analogy of HIV-1 to oncogenic viruses: possible implications for the pathogenesis of AIDS. *Cancer J.* **8**, 308–312.
35. Cosic, I. (1995) Virtual spectroscopy for fun and profit. *Biotechnology* **3**, 236–238.
36. Krsmanovic, B., Cosic, I., Biquard, J. & Hearn, M. (1998) Investigations into the crossreactivity of rabbit antibodies raised against nonhomologous pairs of synthetic peptides derived HIV-1 gp120 proteins. *J. Peptide Res.* **52**, 410–418.
37. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res.* **22**, 4673–4681.
38. Felsenstein J. (1993) PHYLIP (Phylogeny Inference Package), version 3.5c.