Journal of Peptide Science
*J. Peptide Sci.* **4**: 426–435 (1998)

# Solution Structure of Peptides from HIV-1 Vpr Protein that Cause Membrane Permeabilization and Growth Arrest

**SHENGGEN YAO, ALLAN M. TORRES[1], AHMED A. AZAD, IAN G. MACREADIE and RAYMOND S. NORTON***

Biomolecular Research Institute, Parkville, Victoria, Australia

*Received 27 January 1998*
*Accepted 21 March 1998*

Abstract: Vpr, one of the accessory gene products encoded by HIV-1, is a 96-residue protein with a number of functions, including targeting of the viral pre-integration complex to the nucleus and inducing growth arrest of dividing cells. We have characterized by 2D NMR the solution conformations of bioactive synthetic peptide fragments of Vpr encompassing a pair of H(F/S)RIG sequence motifs (residues 71–75 and 78–82 of HIV-1 Vpr) that cause cell membrane permeabilization and death in yeast and mammalian cells. Due to limited solubility of the peptides in water, their structures were studied in aqueous trifluoroethanol. Peptide $Vpr^{59-86}$ (residues 59–86 of Vpr) formed an $\alpha$-helix encompassing residues 60–77, with a kink in the vicinity of residue 62. The first of the repeated sequence motifs (HFRIG) participated in the well-defined $\alpha$-helical domain whereas the second (HSRIG) lay outside the helical domain and formed a reverse turn followed by a less ordered region. On the other hand, peptides $Vpr^{71-82}$ and $Vpr^{71-96}$, in which the sequence motifs were located at the N-terminus, were largely unstructured under similar conditions, as judged by their $C^{z}H$ chemical shifts. Thus, the HFRIG and HSRIG motifs adopt $\alpha$-helical and turn structures, respectively, when preceded by a helical structure, but are largely unstructured in isolation. The implications of these findings for interpretation of the structure–function relationships of synthetic peptides containing these motifs are discussed. © 1998 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: HIV-1; viral protein; solution structure; sequence motifs; helices

## INTRODUCTION

One of the accessory genes of HIV-1, *vpr*, encodes a 96-residue protein Vpr (viral protein R or virion-associated protein R) with a molecular mass of 11.4 kDa. Vpr protein is not essential for HIV-1 replication in established cell lines [1–4], but augments viral replication in primary macrophages [5,6] and human mononuclear phagocytes [7]. Because of its association with the virion, Vpr may have an early role in HIV-1 infection, possibly in penetration or uncoating of the virus [8–10] or in nuclear targeting of the pre-integration complex [11]. Several distinct domains have been recognized in the amino acid sequence of Vpr, some of which correlate with specific biological activities of the protein [12–15]. Of these domains, a region near the C-terminus of Vpr has been shown to cause growth arrest in


Figure 1 Amino acid sequence of HIV-1 Vpr protein with the repeated H(F/S)RIG sequence motifs highlighted. The region of Vpr for which a solution structure was determined is boxed.

Abbreviations: AIDS, acquired immunodeficiency syndrome; HIV-1, human immunodeficiency virus type 1; Vpr, viral protein R; DQF-COSY, double-quantum filtered correlation spectroscopy; RMS, root mean square; TOCSY, total correlation spectroscopy.

* Correspondence to: Biomolecular Research Institute, 343 Royal Parade, Parkville, Victoria 3052, Australia.
[1] Current address: Department of Biochemistry, University of Sydney, Sydney 2006, Australia.


CCC 1075–2617/98/070426-10$17.50

human cells [12,16] and structural defects in yeast [17]. It is suggested that dividing cells are arrested in the $G_2$ phase of the cell cycle, where they are capable of supporting greater virus production than cells progressing normally through the cell cycle [18].

In particular, a pair of H(F/S)RIG amino acid sequence motifs at position 71–82 in HIV-1 Vpr has been shown to cause cell membrane permeabilization and death when added externally to various species of yeast [19]. In $CD4^+$ T lymphocytes, synthetic peptides containing these motifs also caused permeabilization, mitochondrial dysfunction and death by apoptosis [20], and the C-terminal region of Vpr containing these motifs was implicated in mitochondrial dysfunction in yeast expressing Vpr [21]. It has been proposed that these sequence motifs may be responsible for the profound effects of full-length Vpr on cell proliferation [16,22] and thus on the ability of the C-terminal part of Vpr to inhibit chronic infection of dividing cells [16,23]. In addition, as virion-free Vpr protein has been found in the serum of HIV-1 infected individuals [24], it is possible that extracellular Vpr, or peptide fragments thereof, including these sequence motifs, could contribute to the death of bystander cells in AIDS. In view of the possible importance of the C-terminal region of Vpr in HIV-1 infection, we have undertaken a 2D $^1$H NMR study of biologically active synthetic peptide fragments of HIV-1 Vpr from this region (Figure 1) with the aim of defining their structures in solution.

## MATERIALS AND METHODS

### Sample Preparation and NMR Spectroscopy

Peptides were synthesized on an Applied Biosystems 430A peptide synthesizer using the FastMoc solid-phase technique. The crude peptides were purified by reversed-phase HPLC and finally verified by sequencing and/or mass spectrometry. NMR samples were prepared by dissolving 3.5–4.0 mg of each peptide in 600 μl of $H_2O$ containing 50% TFE-$^2H_3$ by volume. Approximately 0.7 mg ($\sim$ 10 mM) DTT was added to each sample in order to minimize possible sample oxidation due to the free cysteine at position 76. The pH was 3.3, as measured at 295 K without correction for isotope or solvent effects.

Spectra were recorded at 298 K on Bruker AMX-500, AMX-600 and DRX-600 spectrometers. Conventional 2D phase-sensitive TOCSY, NOESY and DQF-COSY spectra were obtained for each sample using 2048 complex data points in the directly detected dimension (F2) and 512 increments in the F1 dimension, with 64 scans for each increment. Spectra were processed using UXNMR or XWINNMR (Bruker AG, Karlsruhe) with 60° or 90° phase shifted sine-squared window functions manipulated in both dimensions. Analysis of the spectra was undertaken using XEASY [25], with all spectra being referenced internally to the residual methylene protons of TFE-$^2H_3$, at 3.90 ppm from 2,2-dimethyl-2-silapentane-5-sulfonate. $^3J_{NHC^{\alpha}H}$ coupling constants were estimated from DQF-COSY spectra as described previously [26].

### Structural Constraints and Structure Calculation

The volumes of cross-peaks measured from a NOESY spectrum with a mixing time of 300 ms were used to derive upper bound interproton distance restraints. Conversion from NOE volumes to distance bounds was accomplished using the program CALIBA [27], with correction for pseudoatoms [28], and distances were calibrated using the $\beta$-methylene protons of His78. Five predefined classes were employed in CALIBA: (i) intraresidue cross-peaks, except those between backbone or $\beta$-protons, (ii) intraresidue and sequential cross-peaks between backbone protons or between backbone and $\beta$-protons, (iii) medium range (less than five residues apart) cross-peaks between backbone protons or between backbone and $\beta$-protons, (iv) other (i.e. long-range) cross-peaks between backbone protons and (v) all others. The NOESY cross-peak volumes were considered to be proportional to $r^{-6}$ for classes 2, 3 and 4 and $r^{-4}$ for classes 1 and 5. After these conversions, a further 0.5 or 1.0 Å was added to distance constraints involving only backbone protons or at least one side-chain proton, respectively, to allow for conformational averaging and possible errors in volume integration. Backbone dihedral angle ($\phi$) constraints were inferred from $^3J_{NHC^{\alpha}H}$ coupling constants as follows: $^3J_{NHC^{\alpha}H} < 5$ Hz, $\phi = -60° \pm 30°$; $5 \text{ Hz} < {}^3J_{NHC^{\alpha}H} < 6$ Hz, $\phi = -60° \pm 40°$.

Procedures described previously [26,29] were followed for the structure determination of $Vpr^{59-86}$. Initial structures were generated with the distance geometry program DIANA, version 2.8 [27], using upper bound distance restraints and backbone dihedral angle restraints. The resulting structures were then used to resolve ambiguous NOESY cross-peak assignments. Once the final set of restraints

Table 1 $^{1}$H NMR Chemical Shifts of Vpr$^{59-86}$ at pH 3.3 and 298 K in 50% TFE-$^{2}H_3$/50% $H_2O$[a]

| Residue | NH | C$^{\alpha}$H | C$^{\beta}$H | Others |
|---|---|---|---|---|
| Ala59 | — | 4.15 | 1.55 | |
| Ile60 | 8.40 | 4.20 | 1.94 | C$^{\gamma}H_2$ 1.50, 1.25; C$^{\gamma}H_3$ 0.96; C$^{\delta}H_3$ 0.89 |
| Ile61 | 7.73 | 3.96 | 1.88 | C$^{\gamma}H_2$ 1.51, 1.28; C$^{\gamma}H_3$ 0.95; C$^{\delta}H_3$ 0.88 |
| Arg62 | 7.73 | 4.19 | 1.88, 1.69 | C$^{\gamma}H_2$ 1.68; C$^{\delta}H_2$ 3.21; N$^{\varepsilon}$H 7.17 |
| Ile63 | 7.33 | 3.90 | 2.00 | C$^{\gamma}H_2$ 1.56, 1.19; C$^{\gamma}H_3$ 0.93; C$^{\delta}H_3$ 0.86 |
| Leu64 | 7.67 | 4.06 | 1.72, 1.52 | C$^{\gamma}$H 1.67; C$^{\delta}H_3$ 0.92, 0.84 |
| Gln65 | 8.27 | 3.89 | 2.19, 2.04 | C$^{\gamma}H_2$ 2.48; N$^{\delta}H_2$ 6.75, 6.37 |
| Gln66 | 7.68 | 4.13 | 2.31, 2.25 | C$^{\gamma}H_2$ 2.52; N$^{\delta}H_2$ 6.86, 6.49 |
| Leu67 | 8.34 | 4.15 | 2.02, 1.52 | C$^{\gamma}$H 1.87; C$^{\delta}H_3$ 0.91, 0.87 |
| Leu68 | 8.61 | 4.06 | 1.93, 1.55 | C$^{\gamma}$H 1.78; C$^{\delta}H_3$ 0.86, 0.81 |
| Phe69 | 8.25 | 4.35 | 3.35 | C(2,6)H 7.23; C(3,5)H 7.28 |
| Ile70 | 8.56 | 3.60 | 2.02 | C$^{\gamma}H_2$ 1.96, 1.23; C$^{\gamma}H_3$ 0.92; C$^{\delta}H_3$ 0.89 |
| His71 | 8.33 | 4.11 | 3.31 | C(2)H 8.35; C(4)H 6.84 |
| Phe72 | 8.58 | 4.35 | 3.21 | C(2,6)H 7.19; C(3,5)H 7.25 |
| Arg73 | 8.35 | 3.87 | 1.73, 1.47 | C$^{\gamma}H_2$ 1.44; C$^{\delta}H_2$ 2.95; N$^{\varepsilon}$H 7.00 |
| Ile74 | 8.16 | 3.83 | 1.87 | C$^{\gamma}H_2$ 1.60, 1.12; C$^{\gamma}H_3$ 0.82; C$^{\delta}H_3$ 0.74 |
| Gly75 | 8.10 | 3.82, 3.82 | | |
| Cys76 | 7.89 | 4.26 | 2.85, 2.73 | |
| Arg77 | 7.84 | 4.15 | 1.86, 1.79 | C$^{\gamma}H_2$ 1.63; C$^{\delta}H_2$ 3.15; N$^{\varepsilon}$H 7.16 |
| His78 | 8.22 | 4.62 | 3.33, 3.15 | C(2)H 8.47; C(4)H 7.26 |
| Ser79 | 8.03 | 4.43 | 3.90, 3.86 | |
| Arg80 | 8.17 | 4.34 | 1.89, 1.80 | C$^{\gamma}H_2$ 1.63; C$^{\delta}H_2$ 3.16; N$^{\varepsilon}$H 7.14 |
| Ile81 | 7.85 | 4.12 | 1.81 | C$^{\gamma}H_2$ 1.48, 1.15; C$^{\gamma}H_3$ 0.89; C$^{\delta}H_3$ 0.83 |
| Gly82 | 8.15 | 3.98, 3.89 | | |
| Val83 | 7.73 | 4.16 | 2.12 | C$^{\gamma}H_3$ 0.91, 0.89 |
| Thr84 | 7.94 | 4.38 | 4.23 | C$^{\gamma}H_3$ 1.18 |
| Arg85 | 8.08 | 4.36 | 1.88, 1.75 | C$^{\gamma}H_2$ 1.62; C$^{\delta}H_2$ 3.16; N$^{\varepsilon}$H 7.14 |
| Gln86 | 8.00 | 4.25 | 2.14,1.95 | C$^{\gamma}H_2$ 2.23; N$^{\delta}H_2$ 7.38, 6.57 |

[a] A single entry for methylene protons indicates that it was not established whether the two resonances were degenerate. Spectra were referenced internally using the residual methylene protons of TFE-$^{2}H_3$ at 3.90 ppm.

was obtained, a new family of structures was generated using DIANA, and the 50 with the lowest penalty function were then refined by simulated annealing and energy minimization in X-PLOR, version 3.1 [30]. Conversion of the data format from DIANA to X-PLOR was accomplished using the program COFIMA [27]. The best 20 structures based on their stereochemical energies (i.e. the sum of all contributions to the calculated energy excluding the electrostatic term) and NOE energies were chosen for structural analysis. These structures and the NMR restraints on which they were based have been deposited in the Brookhaven Protein Data Bank (accession no. 1dsk). Structural analyses were carried out using the program MOLMOL [31].

## RESULTS

Sequence-specific assignments were made using standard procedures. Table 1 lists the assignments of Vpr$^{59-86}$ at pH 3.3 and 298 K in 50% TFE-$^{2}H_3$ and 50% $H_2O$. Figure 2 shows the NH–C$^{\alpha}$H and NH–NH regions of a NOESY spectrum of peptide Vpr$^{59-86}$ and Figure 3 summarizes the sequential and medium-range NOE connectivities, including the $^{3}J_{NHC^{\alpha}H}$ coupling constants. Due to spectral overlap, $^{3}J_{NHC^{\alpha}H}$ coupling constants for Arg62 and Val83 were not measured. The numerous $d_{\alpha N}(i, i+3)$, $d_{\alpha N}(i, i+4)$ and $d_{\alpha\beta}(i, i+3)$ NOE connectivities and small $^{3}J_{NHC^{\alpha}H}$ coupling constants ($< 6$ Hz) indicate that an $\alpha$-helical structure is present. The

structure of $Vpr^{59-86}$ was calculated using a total of 267 upper bound distance restraints (excluding those redundant with the covalent geometry, which were eliminated by the program DIANA) and 12 backbone dihedral angle constraints. These 267 upper bound distance restraints consisted of 128 intraresidue, 92 sequential and 47 medium range ($1 < |i-j| < 5$) as summarized in Figure 4a. No lower bound restraints and no sidechain dihedral angles restraints were employed. A summary of geometric and energetic parameters for the best 20 structures of $Vpr^{59-86}$ is given in Table 2.

Backbone RMS deviations over the well-defined region were quite low, as shown in Figure 4b and Table 2. Analysis of the backbone angular order parameters (S) [32,33] of these final 20 structures indicated that residues 60–80 were well defined ($S > 0.9$ for both $\phi$ and $\psi$), as shown in Figure 4c,d. A stereo view of the overall conformation of $Vpr^{59-86}$ is shown in Figure 5, where the backbone heavy atoms of the 20 best structures are superimposed over the well-defined region. On the basis of $\phi$ and $\psi$ angles for the angular average structures and backbone hydrogen bonds (Table 3), residues 60–77 were $\alpha$-helical. Large numbers of $i$ to $i-4$ backbone hydrogen bonds were present in the final structures (Table 3), even though hydrogen bond restraints were not included in the structure calculations. A helical wheel diagram of the $\alpha$-helical residues 60–77 is shown in Figure 6.

Table 2 Structural Statistics for the 20 Final Energy-Minimized Structures of $Vpr^{59-86}$ in TFE/Water at 298 K

| | |
|---|---|
| RMS deviations from experimental distance restraints (Å) (267) | $0.025 \pm 0.002$ |
| RMS deviations from experimental dihedral restraints (deg) (12) | $0.012 \pm 0.024$ |
| RMS deviations from idealized geometry | |
| bonds (Å) | $0.0096 \pm 0.0004$ |
| angles (deg) | $2.30 \pm 0.04$ |
| impropers (deg) | $0.25 \pm 0.03$ |
| Energies (kcal/mol) | |
| $E_{NOE}$ | $8.6 \pm 1.1$ |
| $E_{cdih}$ | $0.0007 \pm 0.0014$ |
| $E_{L\text{-}J}$ | $-109.3 \pm 3.9$ |
| $E_{bond}+E_{angle}+E_{improper}$ | $71.6 \pm 2.7$ |
| $E_{elec}$ | $-356.1 \pm 7.1$ |
| Mean pairwise RMSD (Å) | Backbone heavy atoms/All heavy atoms |
| Residues 59–86 | $2.90 \pm 0.84/4.13 \pm 0.90$ |
| Residues 60–80 (S$\phi$, $\psi$ > 0.9) | $0.92 \pm 0.29/1.94 \pm 0.37$ |

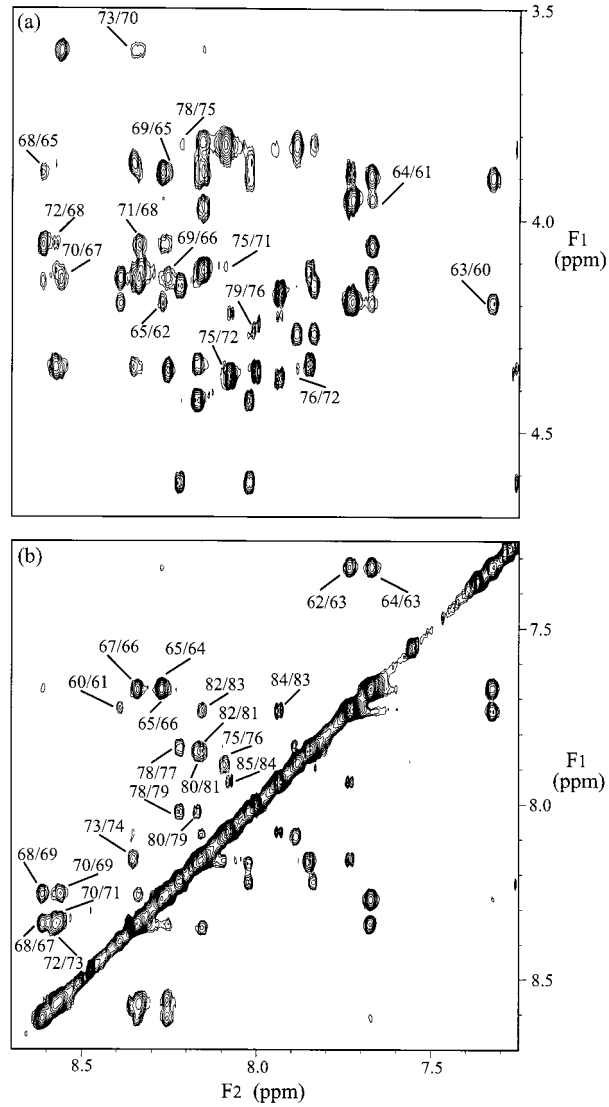


Figure 2 NH–C$^{\alpha}$H (a) and NH–NH (b) regions of a 300 ms mixing time NOESY spectrum of $Vpr^{59-86}$ in aqueous TFE at pH 3.3 and 298 K. In the NH–C$^{\alpha}$H region the $d_{\alpha N}(i, i+3)$ and $d_{\alpha N}(i, i+4)$ connectivities, which are indicative of $\alpha$-helical structure, are labelled, while in the NH–NH region sequential $d_{NN}$ connectivities are indicated.

Sequential assignments were also made for peptides $Vpr^{71-82}$ and $Vpr^{71-96}$, in which the H(F/S)RIG repeat motifs were not preceded by the helical domain. One of these, $Vpr^{71-82}$, consisted of just the two motifs plus the two intervening residues in the native Vpr sequence, while the second, $Vpr^{71-96}$,

incorporated this region plus the basic C-terminus of Vpr (Figure 1). Plots of the deviations of $C^z H$ chemical shifts from random coil values [34] for these two peptides are shown in Figure 7.

## DISCUSSION

In the solution structure of peptide $Vpr^{59-86}$ the helix is slightly kinked around residue 62 and shows slight curvature (Figure 5, Table 3). Curvature towards the less polar face of a helix is expected in amphipathic helices such as those from Vpr (Figure 6), but, as has been noted previously, the direction and degree of curvature may also be influenced by the NMR restraint set [35]. The structure of peptide $Vpr^{59-86}$ shows clearly that the first of the repeated sequence motifs (HFRIG, corresponding to residues 71–75) is part of the well-defined $\alpha$-helical domain (residues 60–77). The second (HSRIG, corresponding to residues 78–82) lay outside the helical domain, and the last two residues were not even particularly well defined in our structures. The failure of the second motif to participate in the helix may reflect the lower preference of Ser for a helical conformation compared with Phe [36]. Inspection of the assembly of structures also suggests that there is a chain reversal around residues 78 and 79, and the $\phi$ and $\psi$ angles (Table 3) are consistent with a distorted Type I turn

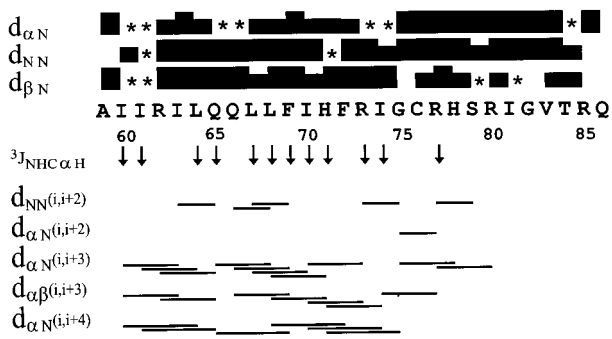


Figure 3 Summary of NMR data for $Vpr^{59-86}$ in aqueous TFE at pH 3.3 and 298 K. Filled bars indicate sequential NOE connectivities, with the heights of the bars reflecting their relative strength (strong, medium or weak); a uniform height is used for the medium-range NOE connectivities. An asterisk (*) indicates that the presence of an NOE could not be confirmed unambiguously due to overlap. Residues with values of $^3J_{NHC^zH} < 6$ Hz are indicated with ↓. Residues with values of $^3J_{NHC^zH}$ in the range of 6–8 Hz and residues for which $^3J_{NHC^zH}$ could not be measured are left blank; no residues with $^3J_{NHC^zH} > 8$ Hz were observed.

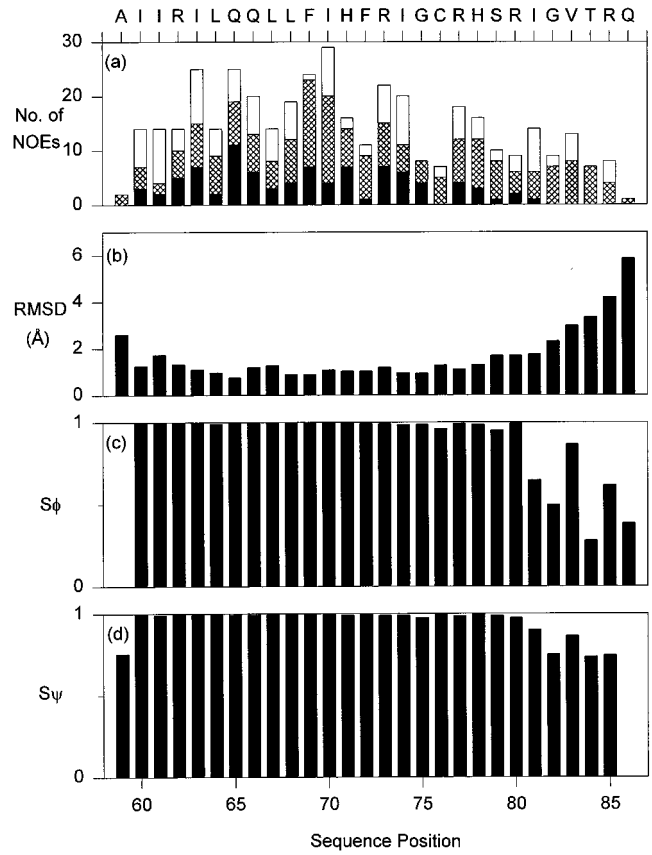


Figure 4 Summary of parameters characterising the final structures of $Vpr^{59-86}$. (a) Distribution of upper bound restraints employed in the final round of structure calculations, with black, cross-hatched and open bars representing medium-range, sequential and intraresidue restraints, respectively. (b) RMS deviations from the mean structure for backbone heavy atoms (N, $C^z$ and C) over the final 20 structures superimposed across the entire length of the peptide. (c, d) Angular order parameters for the backbone dihedral angles $\phi$ and $\psi$, respectively.

[37]. In six of the 20 structures there is an $i$ to $i-3$ hydrogen bond from Arg80 to Arg77 (Table 3). Thus, it may be concluded that in full-length Vpr the first motif would be helical and the second would form a reverse turn followed by a less ordered region. This region may include $3_{10}$-helix, as residues 81–84 formed a turn of $3_{10}$-helix in nearly half of the 20 structures (with average $\phi/\psi$ angles of $-54/-13$, $-52/-19$, $-60/-23$ and $-117/-8$, respectively, in the structures identified by MOLMOL as having a $3_{10}$-helix in this region), although this is not apparent in Figure 5 because of the poorer definition of the C-terminal residues of $Vpr^{59-86}$ relative to the rest of the peptide.

For the peptides $Vpr^{71-82}$ and $Vpr^{71-96}$, in which the H(F/S)RIG repeat motifs are not preceded by the helical domain, there are some similarities between the patterns for residues 71–82 in the two peptides (Figure 7b, c), but most of the $C^{\alpha}H$ chemical shifts are close to random coil values, leading to the conclusion that in isolation these peptide motifs do not adopt a stable secondary structure. This marked contrast with the chemical shifts for the corresponding residues when preceded by a helical region, particularly for the first of the repeated motifs, indicates that the ability of these motifs to participate in a helix and reverse turn is context-dependent. It is likely that the helical structure of the first motif will be present in peptides that have one or two turns of helix preceding the start of the motif. By contrast, extension of these peptides in the C-terminal direction had less effect on the chemical shifts, and by inference the solution structure, of the motifs. Furthermore, the C-terminal region itself appears to be mainly unstructured, although the slightly negative $C^{\alpha}H$ chemical shift deviations for residues 81–92 (Figure 7c) suggest that there is a weak tendency for helix formation in this region.

Table 3 Average Backbone $\phi$, $\psi$ Angles Over the 20 Final Energy-Minimized Structures, and the Number of C=O···HN Intramolecular H-bonds ($i$ to $i+4$ and $i$ to $i+3$) in the 20 Final Structures[a]

| Residue | $\phi$ (deg) | $\psi$ (deg) | C=O···HN H-bonds | |
|---|---|---|---|---|
| | | | $i$ to $i+4$ | $i$ to $i+3$ |
| Ala59 | – | $53 \pm 49$ | 20 | 18 |
| Ile60 | $-49 \pm 4$ | $-43 \pm 7$ | 20 | |
| Ile61 | $-70 \pm 4$ | $13 \pm 9$ | | |
| Arg62 | $-135 \pm 7$ | $-43 \pm 5$ | | |
| Ile63 | $-60 \pm 5$ | $-51 \pm 4$ | | 12 |
| Leu64 | $-68 \pm 8$ | $-30 \pm 6$ | 19 | 19 |
| Gln65 | $-59 \pm 5$ | $-21 \pm 3$ | 20 | |
| Gln66 | $-71 \pm 4$ | $-30 \pm 3$ | 6 | |
| Leu67 | $-81 \pm 6$ | $-46 \pm 3$ | 12 | 4 |
| Leu68 | $-62 \pm 5$ | $-33 \pm 5$ | 20 | |
| Phe69 | $-51 \pm 4$ | $-47 \pm 5$ | 20 | |
| Ile70 | $-65 \pm 6$ | $-47 \pm 6$ | 20 | |
| His71 | $-67 \pm 5$ | $-26 \pm 9$ | 10 | |
| Phe72 | $-90 \pm 7$ | $-37 \pm 7$ | 15 | |
| Arg73 | $-68 \pm 8$ | $-51 \pm 9$ | | 4 |
| Ile74 | $-66 \pm 11$ | $-41 \pm 8$ | | 13 |
| Gly75 | $-58 \pm 10$ | $-49 \pm 14$ | | 12 |
| Cys76 | $-70 \pm 17$ | $0 \pm 7$ | 3 | 1 |
| Arg77 | $-82 \pm 8$ | $-26 \pm 11$ | | 6 |
| His78 | $-48 \pm 10$ | $-36 \pm 7$ | | |
| Ser79 | $-168 \pm 19$ | $34 \pm 9$ | 1 | 5 |
| Arg80 | $38 \pm 7$ | $50 \pm 14$ | 1 | 5 |
| Ile81 | $-28 \pm 53$ | $0 \pm 27$ | 1 | 10 |
| Gly82 | $-55 \pm 76$ | $-12 \pm 45$ | | 5 |
| Val83 | $-55 \pm 16$ | $-27 \pm 17$ | | |
| Thr84 | $-130 \pm 93$ | $-25 \pm 51$ | | |
| Arg85 | $-53 \pm 57$ | $82 \pm 45$ | | |
| Gln86 | $-86 \pm 87$ | – | | |

[a] Data shown in the table were generated using the program MOLMOL [31].

It has been proposed previously that a leucine/isoleucine-rich (LR) domain from residues 60 to 81 is capable of forming a leucine-zipper-like structure which, when fused to the basic domain of the cyclic AMP-responsive element binding protein CREB, supports specific DNA binding by the CREB basic domain [38]. The LR domain is also important in the interaction of Vpr with RIP, a host protein of unknown function [39]. Our results also confirm that the region of Vpr from residues 60 to 81 is largely helical, as required by the proposal that this region behaves as a leucine zipper [38]. Under the conditions examined here, however, the helix did not extend beyond residue 78 and thus presents only three Ile/Leu residues (60, 67 and 74) at a single position on the hydrophobic face of the helix (Figure 6), one less than suggested by Wang *et al.* [38], who included Ile81. Moreover, the structure was well defined as far as residue 80 and was clearly non-helical, so some energetic penalty would be needed to force this region into a helical structure in a complex with other proteins or nucleic acid. In model leucine zippers, leucine residues must occupy at least two of the four *d* positions that make up the core of the dimer interface [40]. In the HIV-1 Vpr sequence investigated here, the 60–81 region would contain only one leucine and three isoleucine residues even if the helix did extend to residue 81. Thus, our data imply that, at least structurally, this region is not a strong candidate for a leucine zipper.

Recently, the activities of Vpr peptides containing the HFRIGCRHSRIG sequence have been assessed in a yeast $LD_{50}$ assay that measures the concentration of peptides required to cause 50% cell killing in 1 h [41]. In these assays the HFRIGCRHSRIG sequence alone ($Vpr^{71-82}$) had an $LD_{50}$ of 750 nM, while additional surrounding sequences appeared to cause an increase in toxicity. For example, the $Vpr^{59-86}$ peptide investigated by NMR had an $LD_{50}$ of 200 nM. While the second motif fused to the downstream Vpr sequence, $Vpr^{76-96}$, has also been shown to be cytotoxic [19,20], there have been no previous reports on the effects of the first motif fused to upstream sequences. The non-acetylated

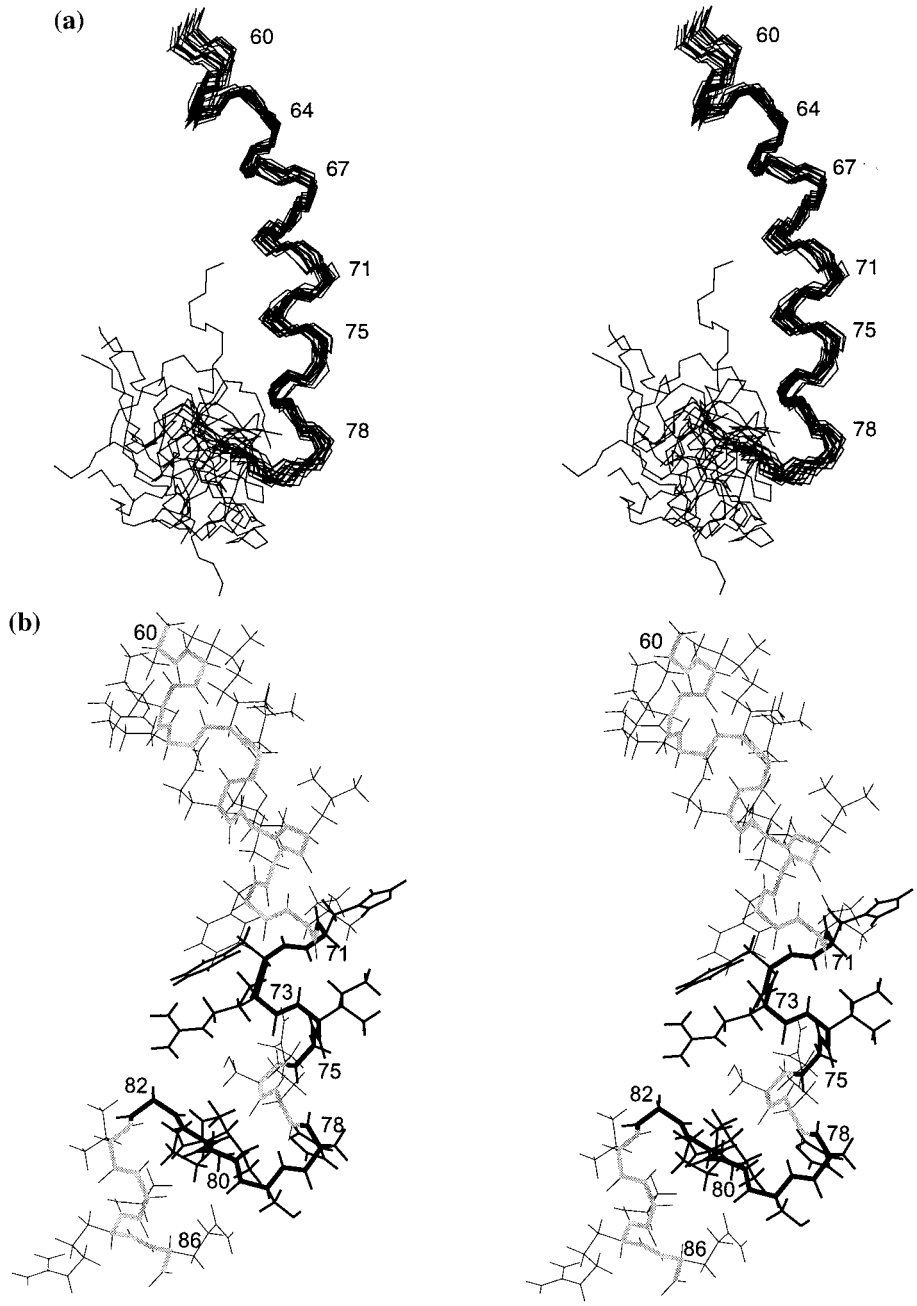


Figure 5 Stereo views of (a) the backbone atoms of 20 final structures of $Vpr^{59-86}$ superimposed over the backbone heavy atoms of residues 60–80 and (b) the structure closest to the average conformation of the 20 structures, with the backbone highlighted with darker shading and the repeated H(F/S)RIG sequence motifs shown in bold. Residues are labelled at backbone $C^{\alpha}$ positions.

form of $Vpr^{50-75}$ had an $LD_{50}$ of 1000 nM, similar to the toxicity of the pair of motifs in isolation ($Vpr^{71-82}$) and slightly weaker than that of peptides containing the pair of motifs flanked by additional residues on both sides (e.g. $Vpr^{59-86}$). The most cytotoxic peptide in yeast was $Vpr^{63-90}$, with an $LD_{50}$ of 10 nM [41]. It appears that maximum expression of the cytotoxic activity of the pair of sequence motifs required the presence of an N-terminal extension, which would have the effect of stabilising the helical structure of the first motif. The addition of residues 83–86 (VTRQ) also in-

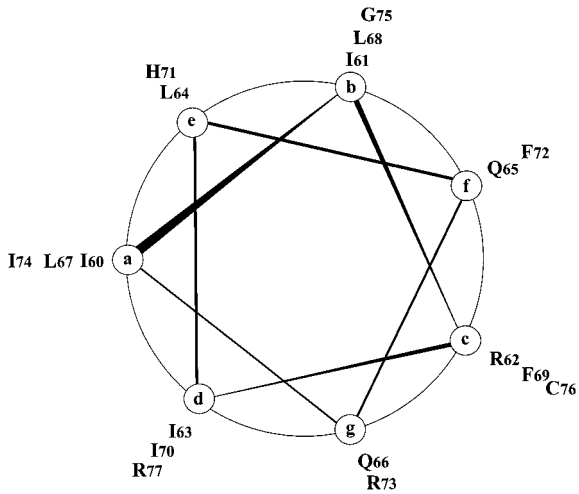


Figure 6 The $\alpha$-helical region of $Vpr^{59-86}$ viewed as an axial projection in a helical wheel presentation.

creased toxicity, possibly by stabilising the reverse turn structure of the second motif, but Macreadie *et al.* [41] observed variable effects on cytotoxicity of further extension to include the basic residues 87–90 (RRAR), depending on which peptide was assayed.

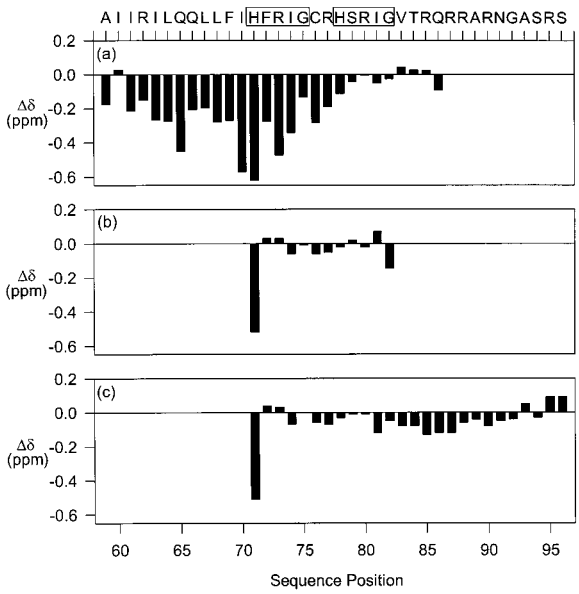


Figure 7 Deviations from random coil values of the $C^{\alpha}H$ chemical shifts of (a) $Vpr^{59-86}$, (b) $Vpr^{71-82}$ and (c) $Vpr^{71-96}$, all in 50% TFE-$^2H_3$/50% $H_2O$ at pH 3.3 and 298 K. Note that the $C^{\alpha}H$ chemical shift of Ile60 in $Vpr^{59-86}$ deviates from the pattern expected for $\alpha$-helical residues, but this probably reflects local effects, as all other criteria [31] are consistent with inclusion of this residue in the helix.

In conclusion, interpretation of the structure–function relationships of HIV-1 Vpr has been complicated by the multiplicity of functions ascribed to it, and uncertainty as to which of these functions are the most important in the establishment and progression of HIV-1 infection in man. In this study we have focused on the C-terminal half of the protein, which includes regions implicated in cell membrane permeabilization, growth arrest and apoptosis. Our solution structures show that the HFRIG and HSRIG sequence motifs adopt helical and turn structures, respectively, when preceded by a helical structure, as in full-length Vpr, but are largely unstructured in isolation. In elucidating the structure–function relationships of these motifs in relation to their effects on membrane integrity and cell proliferation, therefore, it will be necessary to take account of the effect of flanking residues on the structure. In particular, 7–8 residues, sufficient to support at least 1–2 turns of helix, should be included at the N-terminus of synthetic peptide analogues of Vpr to ensure that they are able to adopt the same structure as in the full-length protein.

## Acknowledgements

We thank Mark Hinds, Kevin Barnham, Paul Pallaghy, Jane Tudor and Steve Monks for their valuable assistance and advice, Alan Kirkpatrick for peptide synthesis, Phil Strike for peptide purification, and Jeff Gorman for mass spectral analysis.

## REFERENCES

1. D. Dedera, W. Hu, N. Vander Heyden and L. Ratner (1989). Viral protein R of human immunodeficiency virus type 1 and 2 is dispensible for replication and cytopathogenicity in lymphoid cells. *J. Virol. 63*, 3205–3208.
2. E.A. Cohen, E.F. Terwilliger, Y. Jalinoos, J. Proulx, J.G. Sodroski and W.A. Haseltine (1990). Identification of HIV-1 *vpr* product and function. *J. Acquir. Immune Defic. Syndr. 3*, 11–18.
3. Y. Nishino, M. Kishi, M. Sumiya, K. Ogawa, A. Adachi, K. Maotani-Imai, S. Kato, K. Hirai and K. Ikuta (1991). Human immunodeficiency virus 1 *vif: vpr:* and *vpu* mutants can produce persistently infected cells. *Arch. Virol. 120*, 181–192.
4. Y.-L. Lu, P. Spearman and L. Ratner (1993). Human immunodeficiency virus type 1 viral protein R localization in infected cells and virions. *J. Virol. 67*, 6542–6550.

5. C. Balotta, P. Lusso, R. Crowley, R.C. Gallo and G. Franchini (1993). Antisense phosphothioate oligonucleotides targeted to the *vpr* gene inhibit human immunodeficiency virus type 1 replication in primary macrophages. *J. Virol. 67*, 4409–4414.
6. Z. Matsuda, X. Yu, Q.C. Yu, T.-H. Lee and M. Essex (1993). A virion-specific inhibitory molecule with therapeutic potential for human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA 90*, 3544–3548.
7. R.I. Connor, B.K. Chen, S. Choe and N.R. Landau (1995). Vpr is required for efficient replication of human immunodeficiency virus type 1 in mononuclear phagocytes. *Virology 206*, 935–944.
8. E.A. Cohen, G. Dehni, J.G. Sodroski and W.A. Haseltine (1990). Human immunodeficiency virus *vpr* product is a virion-associated regulatory protein. *J. Virol. 64*, 3097–3099.
9. X. Yuan, Z. Matsuda, M. Matsuda, M. Essex and T.-H. Lee (1990). Human immunodeficiency virus *vpr* gene encodes a virion-associated protein. *AIDS Res. Hum. Retroviruses 6*, 1265–1271.
10. X.-F. Yu, M. Matsuda, M. Essex and T.-H. Lee (1990). Open reading frame *vpr* of simian immunodeficiency virus encodes a virion-associated protein. *J. Virol. 64*, 5688–5693.
11. N.K. Heinzinger, M.I. Bukrinsky, S.A. Haggerty, A.M. Ragland, V. Kewalramani, M.A. Lee, H.E. Gendelman, L. Ratner, M. Stevenson and M. Emerman (1994). The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. USA 91*, 7311–7315.
12. P. Di Marzio, S. Choe, M. Ebright, R. Knoblauch and N.R. Landau (1995). Mutational analysis of cell cycle arrest, nuclear localization, and virion packaging of human immunodeficiency virus type 1 Vpr. *J. Virol. 69*, 7909–7916.
13. S. Mahalingam, S.A. Khan, M.A. Jabbar, C.E. Monken, R.G. Collman and A. Srinivasan (1995). Identification of residues in the N-terminal acidic domain of HIV-1 Vpr essential for virion incorporation. *Virology 207*, 297–302.
14. S. Mahalingam, R.G. Collman, M. Patel, C.E. Monken and A. Srinivasan (1995). Functional analysis of HIV-1 Vpr: Identification of determinants essential for subcellular localization. *Virology 212*, 331–339.
15. S.C. Piller, G.D. Ewart, A. Premkumar, G.B. Cox and P.W. Gage (1996). Vpr protein of human immunodeficiency virus 1 forms cation-selective channels in planar lipid bilayers. *Proc. Natl. Acad. Sci. USA 93*, 111–115.
16. M.E. Rogel, L.I. Wu and M. Emerman (1995). The human immunodeficiency virus type 1 *vpr* gene prevents cell proliferation during chronic infection. *J. Virol. 69*, 882–888.
17. I.G. Macreadie, L.A. Castelli, D.R. Hewish, A. Kirkpatrick, A. Ward and A.A. Azad (1995). A domain of human immunodeficiency virus type 1 Vpr containing repeated H(S/F)RIG amino acid motifs causes cell growth arrest and structural defects. *Proc. Natl. Acad. Sci. USA 92*, 2770–2774.
18. R.H. Miller and N. Sarver (1997). HIV accessory proteins as therapeutic targets. *Nature Med. 3*, 389–394.
19. I.G. Macreadie, C.K. Arunagiri, D.R. Hewish, J.F. White and A.A. Azad (1996). Extracellular addition of a domain of HIV-1 Vpr containing the amino acid sequence motif H(S/F)RIG causes cell membrane permeabilization and death. *Mol. Microbiol. 19*, 1185–1192.
20. C. Arunagiri, I.G. Macreadie, D. Hewish and A.A. Azad (1997). A C-terminal domain of HIV-1 accessory protein Vpr is involved in penetration, mitochondrial dysfunction and apoptosis of human $CD4^+$ lymphocytes. *Apoptosis 2*, 69–76.
21. I.G. Macreadie, D.R. Thorburn, D.M. Kirby, L.A. Castelli, N.L. de Rozario and A.A. Azad (1997). HIV-1 protein Vpr causes gross mitochondrial dysfunction in the yeast *Saccharomyces cerevisiae*. *FEBS Letts. 410*, 145–149.
22. D.N. Levy, L.S. Fernandes, W.V. Williams and D.B. Weiner (1993). Induction of cell differentiation by human immunodeficiency virus 1 *vpr*. *Cell 72*, 541–550.
23. T. Nakaya, K. Fujinaga, M. Kishi, S. Oka, T. Kurata, I.M. Jones and K. Ikuta (1994). Nonsense mutations in the *vpr* gene of HIV-1 during in vitro passage and in HIV-1 carrier derived peripheral blood mononuclear cells. *FEBS Letts. 354*, 17–22.
24. D.N. Levy, Y. Refaeli, R.R. MacGregor and D.B. Weiner (1994). Serum Vpr regulates productive infection and latency of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA 91*, 10873–10877.
25. C. Bartels, T.-H. Xia, M. Billeter, P. Güntert and K. Wüthrich (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR 6*, 1–10.
26. K.J. Barnham, S.A. Monks, M.G. Hinds, A.A. Azad and R.S. Norton (1997). Solution structure of a polypeptide from the N terminus of the HIV protein Nef. *Biochemistry 36*, 5970–5980.
27. P. Güntert, W. Braun and K. Wüthrich (1991). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol. 217*, 517–530.
28. K. Wüthrich, M. Billeter and W. Braun (1983). Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol. 169*, 949–961.
29. S.A. Monks, P.K. Pallaghy, M.J. Scanlon and R.S. Norton (1995). Solution structure of the cardiostimulant polypeptide anthopleurin-B and comparison with anthopleurin-A. *Structure 3*, 791–803.

 

30. A.T. Brünger: *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University, New Haven, CT 1992.
31. R. Koradi, M. Billeter and K. Wüthrich (1996). MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, *51–55*.
32. S.G. Hyberts, M.S. Goldberg, T.F. Havel and G. Wagner (1992). The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci. 1*, 736–751.
33. P.K. Pallaghy, B.M. Duggan, M.W. Pennington and R.S. Norton (1993). Three-dimensional structure in solution of the calcium channel blocker $\omega$-conotoxin. *J. Mol Biol. 234*, 405–420.
34. D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges and B.D. Sykes (1995). $^{1}H$, $^{13}C$ and $^{15}N$ random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR 5*, 67–81.
35. T.D. Mulhern, G.J. Howlett, G.E. Reid, R.J. Simpson, D.J. McColl, R.F. Anders and R.S. Norton (1995). Solution structure of a polypeptide containing four heptad repeat units from a merozoite surface antigen of *Plasmodium falciparum*. *Biochemistry 34*, 3479–3491.
36. P. Prevelige, Jr. and G.D. Fasman in: *Prediction of Protein Structure and the Principles of Protein Conformation*, G.D. Fasman Ed., p. 391–416, Plenum, New York 1989.
37. J.A. Smith and L.G. Pease (1980). Reverse turns in peptides and proteins. *CRC Crit. Rev. Biochem. 8*, 315–399.
38. L. Wang, S. Mukherjee, O. Narayan and L.-J. Zhao (1996). Characterization of a leucine-zipper-like domain in Vpr protein of human immunodeficiency virus type 1. *Gene 178*, 7–13.
39. L.-J. Zhao, S. Mukherjee and O. Narayan (1994). Biochemical mechanism of HIV-1 Vpr function. *J. Biol. Chem. 269*, 15577–15582.
40. J.C. Hu, E.K. O'Shea, P.S. Kim and R.T. Sauer (1990). Sequence requirements for coiled-coils: Analysis with lambda repressor-GCN4 leucine zipper fusions. *Science 250*, 1400–1403.
41. I.G. Macreadie, A. Kirkpatrick, P.M. Strike and A.A. Azad (1997). Cytocidal activities of HIV-1 Vpr and Sac1p peptides bioassayed in yeast. *Protein Peptide Letts. 4*, 181–186.