Structural Characterization of the HIV-1 Vpr N Terminus

EVIDENCE OF cis/trans-PROLINE ISOMERISM*[S]

Karsten Bruns§‡, Torgils Fossen‡, Victor Wray‡, Peter Henklein‡, Uwe Tessmer§, and Ulrich Schubert§§§§§§§

From the §Department of Structural Biology, Gesellschaft für Biotechnologische Forschung, D-38124 Braunschweig, Germany, §Heinrich-Heine-Institute, University of Hamburg, D-20251 Hamburg, Germany, *Department of Chemistry, University of Bergen, 5007 Bergen, Norway, ‡Institute of Biochemistry, Humboldt University, D-10115 Berlin, Germany, the **Laboratory of Viral Diseases, NIAID, National Institutes of Health, Bethesda, Maryland 20892, and the ¶¶Institute for Clinical and Molecular Virology, University of Erlangen-Nürnberg, D-91054 Erlangen, Germany

The 96-residue human immunodeficiency virus (HIV) accessory protein Vpr serves manifold functions in the retroviral life cycle including augmentation of viral replication in non-dividing host cells, induction of G2 cell cycle arrest, and modulation of HIV-induced apoptosis. Using a combination of dynamic light scattering, circular dichroism, and NMR spectroscopy the N terminus of Vpr is shown to be a unique domain of the molecule that behaves differently from the C-terminal domain in terms of self-association and secondary structure folding. Interestingly, the four highly conserved proline residues in the N terminus are predicted to have a high propensity for cis/trans isomerism. Thus the high resolution structure and folding of a synthetic N-terminal peptide (Vpr1–40) and smaller fragments thereof have been investigated. 1H NMR data indicate Vpr1–40 possesses a higher helical structure between residues 17–32, and for the first time, this helix, which is bound by proline residues, was observed even in aqueous solution devoid of any detergent supplements. In addition, NMR data revealed that all of the proline residues undergo a cis/trans isomerism to such an extent that ~40% of all Vpr molecules possess at least one proline in a cis conformation. This phenomenon of cis/trans isomerism, which is unprecedented for HIV-1 Vpr, not only provides an explanation for the molecular heterogeneity observed in the full-length molecule but also indicates that in vivo the folding and function of Vpr should depend on a cis/trans-proline isomerase activity, particularly as two of the proline residues in positions 14 and 35 show considerable amounts of cis isomers. This prediction correlates well with our recent observation (Zander, K., Sherman, M. P., Tessmer, U., Bruns, K., Wray, V., Prechtel, A. T., Schubert, E., Henklein, P., Luban, J., Neidleman, J., Greene, W. C., and Schubert, U. 2003) J. Biol. Chem. 278, 43170–43181) of a functional interaction between the major cellular isomerase cyclophilin A and Vpr, both of which are incorporated into HIV-1 virions.

Viral protein R (Vpr) is a small 96-amino acid virion-associated (1) nucleocytoplasmic shuttling (2) regulatory protein that is encoded by (and conserved among) primate lentiviruses, the human immunodeficiency viruses, type 1 and type 2 (HIV-1/2)1 and simian immunodeficiency viruses (SIV). In addition to the canonical retroviral Gag, Pol, and Env proteins HIV-1 encodes further small proteins with either regulatory (Tat and Rev) or accessory (Vpu, Vif, Nef, and Vpr) functions. Although dispensable for growth of HIV-1 in dividing cultured T-cells Vpr appears to play an important role for virus replication in vivo, because deletion of vpr and the related vpx genes in SIV severely compromises the pathogenic properties in experimentally infected rhesus macaques (3, 4). Furthermore, Vpr is highly conserved in HIV-1 and other primate lentiviruses, HIV-2 and SIV, which also encode an additional Vpr-related protein termed Vpx, which is believed to function synergistically with Vpr. Vpr of HIV-1 is reported to exhibit numerous biological activities including nuclear localization based on the presence of at least two nuclear localization signals (2, 5–9), ion channel formation (10), transcriptional activation of HIV and heterologous promoters (11–14), co-activation of the glucocorticoid receptor (15, 59), regulation of cell differentiation (16), induction of apoptosis (17, 18), cell cycle arrest (16, 19), and transduction through cell membranes (20). Although significant amounts of Vpr (~0.15-fold molar ratio to viral core proteins (21)) are packaged into budding HIV-1 particles in a process dependent on interaction of Vpr with the C-terminal p6Gag domain of the Gag polyprotein precursor Pr55 (1), the biological role(s) of virion-associated Vpr remains to be fully elucidated. In particular, the mechanism that regulates the bidirectional transport of de novo synthesized Vpr between nucleus and budding virions is unknown.

Although Vpr is non-essential for virus replication in T-cells this accessory protein is known to augment virus replication in cultures of terminally differentiated monocytes/macrophages, a function that was directly related to its karyophilic properties in non-dividing target cells (5). Vpr is thought to participate in the import of the viral pre-integration complex (PIC) across the nuclear membrane by causing dynamic disruption in the nu-

1 The abbreviations used are: HIV-1/2, human immunodeficiency virus-1/2; CypA, cyclophilin A; DLS, dynamic light scattering; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; MS, mass spectrometry; PIC, pre-integration complex; PPIase, peptidyl-prolyl cis/trans isomerase; TFE, trifluoroethanol; TOCSY, total correlation spectroscopy; SIV, simian immunodeficiency virus; vpr, synthetic full-length Vpr; Fmoc, N-(9-fluorenylmethoxycarbonyl) hydroxymethylglycine; r.m.s.d., root mean square deviation(s).
clear envelope architecture (22), thereby sanctioning integration of HIV proviral DNA into the host genome. In contrast to prototype retroviruses, lentiviruses can efficiently replicate in non-dividing cells. Import of the PIC, which contains the viral RNA/DNA and the machinery that assists import into the nucleus and subsequent integration of the viral genome into chromosomes, is dependent on the cooperation among matrix protein p17MA, integrase, and Vpr, as well as a “DNA flap” produced during reverse transcription (reviewed in Refs. 23–25).

A second well established biological function of Vpr is its ability to cause proliferating CD4+ T-cells to undergo an arrest or delay at the G2 cell cycle checkpoint (19, 26; reviewed in Ref. 23). Numerous cellular binding partners of Vpr have been identified, and for some a role in G2 arrest has been proposed (27–31). Nevertheless, the molecular mechanism underlying Vpr-induced G2 arrest remains unclear. A potential explanation for the obvious paradigm that Vpr prevents proliferation of infected T-cells by arresting them in the G2 phase was provided recently (32) by the observation that viral gene expression is optimal in G2, and that Vpr can increase virus production by delaying cells at this point in the cell cycle. Interestingly, there are sufficient amounts of Vpr in incoming virus particles to induce G2 cell cycle arrest even prior to the initiation of de novo synthesis of viral proteins (33, 34).

The biological importance of Vpr for lentivirus replication in vivo suggests that selective modification of Vpr function with small molecule inhibitors might yield a new class of antiviral agents, particular those that affect the specific interaction of cellular proteins with Vpr. However, the design of effective Vpr antagonists requires more detailed knowledge of its dynamic behavior. As a consequence, the molecular heterogeneity observed for sVpr may arise, at least in part, from cis/trans isomerism of the four highly conserved Pro residues located in position 5, 10, 14, and 35 (Fig. 1). Thus, we explored the high resolution structure and folding of the N terminus of Vpr under various solution conditions using a combination of CD, DLS, and NMR spectroscopic techniques. We found that Vpr1-40 possesses helical structure between residues 17–32, and only one pronounced turn is found in the N terminus. In contrast to previous work by others (36), we are now able to detect stable structures for the N terminus of Vpr even in the absence of organic solvents. Surprisingly, even under the most stabilizing solution conditions, in the presence of organic solvents such as TFE, all proline residues in Vpr (Fig. 1) undergo cis/trans isomerism to such an extent that ~40% of Vpr molecules possess at least one proline in a cis conformation.

As a consequence, the molecular heterogeneity observed for sVpr may arise, at least in part, from cis/trans isomerism of proline residues located in the N-terminal region of the molecule at positions 5, 10, 14, and 35 (Fig. 1). We propose that this unusual cis/trans phenomenon contributes to the high flexibility of Vpr and that this so far unreported phenomenon indicates a requirement for a cellular cis/trans peptidyl-prolyl isomerase (PPIase) activity to regulate folding of Vpr in vivo. Indeed in an accompanying paper (38) we report that the N terminus of Vpr interacts with a major host cell PPIase cyclophilin A (CypA) that, like Vpr, is specifically incorporated into HIV-1 virions and that this CypA-Vpr interaction regulates expression and biological function of Vpr.

EXPERIMENTAL PROCEDURES

Peptide Synthesis, Purification, and Characterization—Synthesis was performed on an Applied Biosystems automated peptide synthesizer 433A on a 0.09 mmol scale with TentaGel R-RAM Fmoc resin (capacity 0.21 mmol/g)1 using the Fmoc/tert-butyl strategy. The following side chain protecting groups were used: 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Arg), t-butoxycarbonyl (Trp, Lys), t-butylen ether (Thr, Ser, Tyr), t-butylen ester (Asp, Glu), and trityl (Asn, Gln, and His). Coupling was performed with 1 mmol Fmoc amino acids us-
ing. N-N'-benzetrizoxol-1-yl(dimethylamino)-methylene-N-methyl-
ethanaminium hexafluorophosphate-N-oxide in N-methylpyrrolidone
as coupling agent with a cycle time of 45 min for single coupling and 75
min for double couplings. The crude protein was purified by reverse
phase HPLC on a VYDAC C18 column (40 × 300 mm, 150 µm, 300 Å)
with a linear gradient of 25% A to 78% B in 50 min. (A, 1000 ml of water,
2 ml of trifluoroacetic acid; B, 500 ml of acetonitrile, 100 ml of water, 1
ml of trifluoroacetic acid) at a flow rate of 100 ml/min. 1H NMR spec-
photometric monitoring at A = 214 nm. The fractions were analyzed by
reverse phase HPLC (Shimadzu L10) on a VYDAC C 18 column (4.6 ×
250 mm, 5 µm, 300 Å) with a linear gradient of 10 to 100% B in 45 min.
Vpr1–40 and Vpr1–46, as well as their Asn mutants, were synthesized
and purified analogously. All peptides showed the expected molecular
masses.

**Peptide Sequencing and Mass Spectrometry**—The complete sequence
of Vpr1–46 was confirmed on an Applied Biosystems 473A pulsed-liquid
phase sequencer according to a standard protocol. Positive ion electro-
spray ionization mass spectra were recorded on a micromass Q-ToF-2TM
mass spectrometer. Protein samples were dissolved in 70% aqueous
methanol and infused into the electrospray chamber with an electro-
spray needle voltage of 0.8 kV. Matrix-assisted laser desorption/ioniza-
tion time-of-flight (MALDI-TOF) mass spectra were recorded on a
Bruker Reflex II MALDI/TOF mass spectrometer using an N2 laser (337
nm). Both spectra did not show any appreciable amounts of contamina-
tion or byproducts.

**SDS-PAGE**—Discontinuous polyacrylamide SDS-PAGE using Tris-
Tricine cathode buffer was carried out as described by Schägger and
von Jagow (39). 0.5, 1, 2.5, and 5 µg of Vpr1–46 dissolved and boiled for 5 min
in sample buffer (5% SDS, 2.5% mercaptoethanol, 20% glycerol, 50 m M
Tris, 0.02% Serva blue G, pH 6.8) were loaded into the sample wells of
the gel. After electrophoresis the peptides were fixed, stained (0.025%
Serva blue G), and destained according to a standard protocol.

**DLS Measurements**—DLS was performed on a DynaPro-801 mole-
cular sizing instrument and evaluated as described previously (20).

**CD Spectroscopy**—CD spectra were recorded at room temperature in
0.5-mM cuvettes on a Jasco J-600 spectropolarimeter in a wavelength
range from 280 to 180 nm under various solution conditions (TFE
content and pH, 0.8 M Tris base) as described above. The secondary
structure content was quantified with the program VARSELEC (40).

**1H NMR Spectroscopy**—1H NMR spectra were recorded without spin-
ning on a Bruker Avance DNX 600 NMR spectrometer using a triple-
resonance probe head with a gradient unit. Vpr1–46 (7.35 mg/ml; 1.5
mm) was dissolved without pH adjustment either in pure water con-
taining 10% D2O by volume or in 50% aqueous TFE-water (7.35 mg/ml;
1.5 mm) for 10% D2O and titrated to pH 7.35. In some cases, 50% TFE
was dissolved and boiled for 5 min. CD spectra were recorded at room
temperature in 50% TFE-water (7.35 mg/ml) (1.5 mm) for 10% D2O
and N-terminal residues could be assigned. Closer inspection of
the final superimposition analysis. Structure fitting criteria were ob-
jectively derived using a consecutive segment method described previ-
ously by us (44). Final structures were displayed and manipulated on a
Silicon Graphics INDIGO 2 work station using the program PROCHECK
(47).

**RESULTS**

**Characterization of Full Length Vpr by 1H NMR Spectroscopy**—Our earlier studies (20) indicated that Vpr already pos-
sessed secondary structure in pure aqueous solution and that this
mainly α-helical structure was dependent on the hydro-
phobicity of the solvent. Further, and most strikingly, a pH-de-
pendent folding switch was observed for Vpr, indicating its
secondary structure and tertiary fold is dependent on the pres-
ence of specific binding factors (e.g. nucleic acids, proteins, or
membrane components) or the environment of the cytosol, nu-
cleus, mitochondrion, cellular membranes, and the extracellu-
lar space where Vpr is present in vivo and exerts biological
function (25). However, the molecular basis for this consider-
able structural variability of Vpr has not yet been evaluated.

To obtain further details of the structure of the full-length
molecule we analyzed the previously characterized synthetic
peptide sVpr (20) that comprises the sequence of a Vpr protein
derived from the isolate HIV-1_DL4_3 (48). 1H NMR spectra of
sVpr were recorded in both water alone and in 50% TFE as this
latter solution tends to stabilize secondary structure and alle-
viates to some extent problems associated with intermolecular
interactions at concentrations necessary for NMR (20). Secondary
structure determinations (49). However, even under the most favorable conditions
(1.3 mM sVpr in 1:1 CF3CD2OH/H2O at pH 3.1) the one-
and two-dimensional 1H NMR spectra (Fig. 2A) showed broadening of
the 1H signals corresponding to the central region of the
molecule. Under these conditions only a limited number of C-
and N-terminal residues could be assigned. Closer inspection of
the low field region (10–9.3 ppm) of the TOCSY spectrum shows
three signals corresponding to the Hδ2/Hε1 of three tryptophan residues. It is evident from the expansion of these
cross-peaks (Fig. 2B) that the sample appears to be heteroge-
eous as each main signal shows at least two smaller signals.
A similar situation was observed for the expansion of the
TOCSY spectrum corresponding to histidine (Fig. 2C). Accord-
ing to the cross-peak intensities these smaller signals account
for at least 20% of the major signal. Given the fact that the
NMR samples of sVpr were apparently homogenous by means of
SDS-PAGE, protein sequencing, and mass spectrometry (20),
the only plausible explanation for the unusual degree of mole-
cular heterogeneity observed would be an extraordinary exis-
tence of a cis/trans isomerism of the four conserved prolines that
are present in the N-terminal region of sVpr in positions 5, 10,
and 35 (Fig. 1). Hence, the next logical step was to charac-
terize the structure and folding of the N-terminal domain.

We have used synthetic Vpr1–46 and various related frag-
ments and mutants thereof to establish the folding behavior of
this domain, potential for oligomerization, and consequences of
the cis/trans isomerization of the four proline residues. Each
peptide was synthesized using a previously established solid
phase synthesis protocol and purified to homogeneity. This
is demonstrated here by HPLC (Fig. 3A) and SDS-PAGE (Fig. 3D)
for Vpr1–46. Even at the highest load of 5 μg of peptide per lane
no significant amount of byproducts that could originate from
incomplete synthesis or proteolytic degradation was detectable
in the gel. In contrast to our previous findings (20) for sVpr,
where dimers of sVpr were found under non-reducing conditions, Vpr\textsubscript{1-40} showed no signs of dimerization. This observation is in good agreement with the fact that the putative dimerization domain of sVpr is only present in the C-terminal portion of the molecule as a single cysteine residue in position 76 capable of forming Vpr homodimers via a disulfide bond. The
identity of purified peptide was further confirmed by N-terminal sequencing and positive ion electrospray ionization MS. The experimental data showed a well defined multiply charged spectrum (Fig. 3B) that was deconvoluted to give an intense envelope for the molecular ion cluster at a molecular mass of 4902.1 Da (Fig. 3C), corresponding to the molecular mass calculated for Vpr1–40. In all cases the MS and sequence analyses indicated that the fragments were homogenous and showed no detectable evidence of byproducts.

Analysis of Vpr1–40 in Solution by CD and DLS—A first insight into the folding of Vpr1–40 was achieved by analysis of the peptide at ambient temperature under various solutions conditions by CD spectroscopy (Fig. 4). In pure water at pH 3.9 the CD spectrum exhibited negative ellipticities at 203 and a strong positive band at 190 nm indicative of the presence of significant amounts of helical structure (Fig. 4A). Deconvolution of the data set using the variable selection method (40) afforded ~20% helical content (Fig. 4D). Increasing the hydrophobicity of the solution by addition of TFE, an organic solvent that is known to favor intramolecular interactions and thus stabilizes secondary structures, increased the α-helical content resulting in more intensive curves and shifted negative extrema. The helical content increased on addition of TFE and had reached a maximum at 20% TFE as there was no further change on going to 50% TFE (Fig. 4A). Deconvolution of the spectra indicates that at the maximum there was ~45% α-helix present that is equivalent to the involvement of 18 residues in the limiting structure (Fig. 4D). To investigate Vpr1–40 under more physiological conditions the same CD analyses was performed in Pi buffer at pH 7.2 (Fig. 4B and C). In contrast to previous results where it was shown that sVpr and the C-terminal domain, where there is no evidence of helix at pH 7.2 in buffer or buffer containing 20% TFE (20), the N-terminal domain maintains some helical structure at pH 7.2 (Fig. 4D). Together, these results demonstrate that similar to sVpr (20) and C-terminal fragments (37) the N terminus possesses secondary structure

![Figure 4](image-url)
in pure water. However, in contrast to the pH-mediated folding switch observed for sVpr or the C-terminal fragment Vpr47−96, the N-terminal Vpr1−40 maintained some α-helical structure even at the critical neutral pH where sVpr and Vpr47−96 were completely unstructured (20). Furthermore, the structure stabilizing effect of TFE shown previously for sVpr and Vpr47−96 appeared to be essentially pH-independent for folding of the N-terminal peptide Vpr1−40.

Previous cross-linking, DLS, and SDS-PAGE analyses demonstrated a tendency of Vpr to form high order complexes in water that correspond in molecular weight to those of decamers (20, 50). Their presence can be drastically reduced by the introduction of detergents such as SDS or TFE, which favor the formation of low order oligomers with S-S-linked dimers being the most pronounced. Hence, we have investigated the self-association of the N terminus by conducting DLS on Vpr1−40. Similar to the domain-dependent folding of sVpr (Fig. 4), the DLS data for Vpr1−40 also show considerable differences to those established previously for sVpr and Vpr47−96 (20, 50) indicating that the N-terminal region of Vpr does not contribute to the high tendency of Vpr for homo-oligomerization (Table I). In non-buffered aqueous solution at low pH the major species of Vpr1−40 present is the monomer. This contrasts markedly with DLS results obtained with both sVpr and Vpr47−96 where substantial oligomerization occurs even at lower protein concentrations indicating that the N terminus is not involved in the oligomerization of Vpr.

In summary, our combined data provide strong evidence that Vpr is organized into different domains; in contrast to the C terminus the N-terminal domain does not oligomerize and does not undergo a pH-mediated folding switch. However, both full-length sVpr and the N-terminal domain have the same tendency to adopt secondary structure even in pure aqueous solution.

Identification of a Proline cis/Trans Isomerism in the N Terminus of Vpr—Inspection of the primary sequence of Vpr shows that the N terminus contains all the proline residues in the molecule, and of particular importance, these highly conserved residues form a CyPA binding motif of biological relevance (see accompanying paper (38)). From a consideration of previous investigations of the effect of adjacent residues on the cis/trans conformations of proline in model systems (51–53) one would expect relatively high percentages of cis isomers for prolines at residues 14 and 35 that are adjacent to aromatic residues. At least two of the tryptophans of Vpr and two histidines are near to these sites (Fig. 1). To investigate this potential cis/trans-proline phenomenon independent of the context of the full-length Vpr we analyzed two N-terminal adjacent fragments, Vpr1−20 and Vpr21−40, in 50% TFE (Fig. 5). For the latter fragment only one proline is present at position 35, and this manifests itself by a doubling of a considerable number of signals for residues up to at least five residues away from the proline. This is readily seen in the TOCSY spectrum for the Vpr1−20 as the third isomer signal, which coincides with the major cis isomer signal, at 300 K showing the correlations between protons Hα/Hε1 of His-33 and His-40 (A) of Vpr1−20 at 276 K showing the correlations between the protons Hα/Hε1, of Trp-18 (B). See text for lettering assignments.

A similar phenomenon is evident for the peptide Vpr1−20 as the low field cross-peak (Hα/Hε1) of the single Trp residue at position 18 exhibits a number of resolvable cross-peaks (Fig. 5B). The most intense peak of these must correspond to the all-trans peptide (Fig. 5B, a), whereas the two more intense of the remaining signals correspond to peptides containing single cis isomers (Fig. 5B, b and c), probably corresponding to residues Pro-14 and Pro-10, respectively. A broadening of the most intense signal, apparent at 300 K (data not shown) but most readily seen at low temperature (Fig. 5B), suggests a third cis isomer coincides with this signal (presumably that from Pro-5, which is furthest away in the linear peptide from Trp-18, Fig. 5B, d). Clearly, the least intense signal (Fig. 5B, e) belongs to the cross-peak Hε1/Hε2 of the all-trans peptide. We assume that the third cis isomer signal, which coincides with the major trans signal, has the same intensity as the less intense of the two separate single cis peptide signals. According to the integration of the signals the total content of cis peptides is 25%, with the major single cis isomer of ~15% (assigned to Pro-14 as
it is adjacent to an aromatic residue), and both minor ones of ~5%.

On the reasonable assumption that the cis peptide content for each proline is independent of the fragments of Vpr studied then these findings suggest that ~40% of the full-length Vpr molecules contain at least one cis-proline residue, and a further small percentage (~5%) contain two or more cis-prolines. From the observation that the tryptophan signals are influenced by prolines at least eight residues away it is logical to conclude that the molecular heterogeneity of the proline isomers also leads to a broadening of the NMR signals and subsequent loss of signal intensity observed for sVpr (Fig. 2).

To further define the structural role of cis/trans-proline isomerism in our system we have investigated the N-terminal fragment Vpr

\[ \text{Vpr}^{1-40} \]

the two equal length related fragments Vpr

\[ \text{Vpr}^{1-20} \]

and Vpr

\[ \text{Vpr}^{21-40} \]

used above, and also mutants that carry proline to asparagine exchanges. Deciphering the high resolution structure of the N terminus of Vpr was particular important as Wecker and Roques (36) have reported the structure of the N terminus of a closely related N-terminal Vpr fragment, as well as that of the full-length molecule (35), derived from a HIV-1 isolate different from HIV-1NL4-3 used in our experiments. In these studies no attempt was made to assess either the effects of a potential cis/trans isomerism or the specific role of proline residues in the molecule, nor was the effect of solution conditions on folding and self-association of the N terminus of Vpr investigated.

Secondary Structure of Vpr

\[ \text{Vpr}^{1-40} \]

from \( ^{1}H \) TOCSY and NOESY NMR spectra afforded complete assignments of the all-trans-proline isomers of Vpr

\[ \text{Vpr}^{1-40} \]

in 50% TFE at 310 K and in water alone and of its two related fragments, Vpr

\[ \text{Vpr}^{1-20} \]

and Vpr

\[ \text{Vpr}^{21-40} \]

(all measure-
ments at 300 K and pH ~3; see Supplemental Material). In such molecules qualitative information about the nature and position of secondary structure is deducible from the α-proton chemical shifts (54). This provides a convenient method of comparing and following changes in such structures while varying solution conditions or introducing point mutations. In particular, upfield shifts of the α-proton in four adjacent residues relative to the random coil values are indicative of local helical structure, whereas the downfield shift of three adjacent residues is indicative of β-sheets. In the present case the plots of the 1H chemical shift differences for 50% TFE solutions are shown in Fig. 6. For Vpr1-40 helical secondary structure is clearly present between residues Trp-18 and Arg-32, and this extends back in a less well defined form to residue Tyr-15 (Fig.

**Fig. 7.** Chemical shift differences of the α-protons between the experimental values and those for residues in a random coil for trans-proline Vpr1-40 in 50% TFE and water (A) and chemical shift differences of the α-protons for the all-trans-proline isomer of Vpr1-40 in 50% TFE and in water alone (B).
In addition, a number of unusual low field shifts were observed for the \( \alpha \)-protons of N-terminal residues adjacent to the all-trans-proline residues that can not be readily interpreted in terms of the secondary structure. To investigate these shifts signals more thoroughly and to test how stringent the structural changes as little has been reported on the effects of proline substitution calculated above belonging to the adjacent preceding residue in the sequence (+0.28 ± 0.1 ppm) and similar but smaller shifts (+0.08 ± 0.03 ppm) for those that are two residues toward the N terminus that is independent of the presence of secondary structure.

In contrast, the last eight residues of the mutant of Vpr\(^{21-40}P35N\) (Fig. 6E) clearly show a tendency for helical structure, which is seen as an extension of the structure already apparent in the wild type sequence of peptides Vpr\(^{21-40}\) (Fig. 6C) and Vpr\(^{1-40}\) (Fig. 6A). Taking into consideration the inherent shifts caused by proline substitution calculated above the apparent unusual shift of Phe-34 in the wild type peptide would be much less pronounced and nearer zero, a value implying loss of helical structure in this region of the molecule.

**The Hydrophobicity of the Solvent Affects Secondary Structure of Vpr\(^{21-40}\)**—As it has been reported recently (57) that the N terminus of Vpr interacts with biological membranes it was important to investigate the impact of TFE on the structure of Vpr\(^{1-40}\). Although the structure of this region derived from a different HIV-1 Vpr protein was reported recently (36) for a 30% TFE solution the influence of the organic solvent could not be quantitatively assessed in this previous report. Using different strategies in peptide synthesis and purification we have been able to assign the NMR data of Vpr\(^{1-40}\) in pure water and are now able to probe the environmentally related structural changes of the molecule.

The patterns and magnitude of the \( \alpha \)-shift differences obtained for Vpr\(^{1-40}\) (Fig. 7A) are very similar between residues 1 to 23 for both solutions, with and without TFE, implying that the helical structure in this part of the molecule is maintained upon removal of the organic solvent. In water the \( \alpha \)-shifts are consistently to lower field between residues 24 and 40 indicating a loss in the stability of structures in this area (Fig. 7B). Clearly, in water alone the N-terminal side of the helix (residues 18 to 23) is still present whereas, in contrast, the C-terminal section of the molecule has become more flexible. The implication for these foremost structural analyses of the N terminus of Vpr in pure aqueous solutions is that residues 18 to 23 adopts a particularly stable structure, the presence of which is independent on the hydrophobicity of the solvent.

We have shown here that there is evidence for a considerable percentage of cis-proline isomers present in all four proline residues of Vpr. The complexity, the weak intensity, and the multiplicity of signals caused by this phenomenon prevented an unambiguous assignment of the H\(a\)-signals in the cis iso-
mers in most of the peptides studied, except for Vpr\textsuperscript{21–40}, where only one proline residue is present, allowing an assignment of both cis and trans isomers in both water and 50% TFE (see Supplemental Material). Although there may be some local loss in secondary structure in the vicinity of Pro-35 the data obtained with Vpr\textsuperscript{1–40} in both solutions indicate there is no struc-

**Fig. 8.** Mean r.m.s.d. for the backbone atoms in each residue, calculated using a consecutive segment method that averages the differences for segments two to five residues in length, plotted against the residue number (A). The figure represents supposition of the 20 best final restrained structures of Vpr\textsuperscript{1–40} in water and in 50% TFE after alignment of the backbone atoms of residues 14 to 39 (B and D, respectively) and 17 to 32 (C and E, respectively).
of Vpr1 in pure water and aqueous TFE indicate that the helical structure to 30.

In summary, the qualitative chemical shift data obtained in pure water and aqueous TFE indicate that the helical structure of Vpr1–40, which is also evident from the CD data, is situated between residues 18 and 32, and this continues at least partially into the C-terminal region of the molecule including Pro-35. On removal of the organic solvent there is an increase in the flexibility of the C-terminal section of the helix indicating the helix encompassing residues 18 to 23 is the most stable structured moiety in the molecule. Independent of the presence of TFE, the initial 14 residues show no tendency for extended secondary structure formation, a property that is not a consequence of TFE, the initial 14 residues show no tendency for extended secondary structure formation, a property that is not a consequence of the three proline residues in this region and their cis/trans isomerization as this phenomenon is also observed for the all proline-deficient mutant Vpr1–20P5,10,14N.

Calculated Solution Structures of Vpr1–40—Finally we have determined the quality of the secondary structure elements present in the high resolution structure of Vpr1–40 and the dependence of these on the solution conditions. Quantitative NOE data for both pure water and 50% TFE were used as distance constraints in molecular dynamic/energy minimization calculations using a standard protocol (43) (details of the distribution of the quantitative NOEs used in these calculations are available in the Supplemental Material). A total of 461 and 370 distance restraints (from 246/207 intraresidue, 148/131 sequential, and 67/32 medium range NOEs) for Vpr1–40 in 50% TFE and water, respectively, were used to generate 100 conformations for each regime. In each case 20 conformations with the lowest NOE (E_{NOE} 377 kJ/mol for Vpr1–40 in 50% TFE and 585.9 kJ/mol for Vpr1–40 in water) and total energies (E_{total} 1183.6 kJ/mol for Vpr1–40 in 50% TFE and 1342.7 kJ/mol for Vpr1–40 in water) and showing no (50% TFE) or not more than two (water) constraint violations greater than 0.2 Å were used for the final fitting analysis (Table II).

We have shown previously (44) that the heterogeneity within a final set of molecular conformations can be visualized using the consecutive segment approach in which the r.m.s.d. of the backbone atoms for short segments, two to five residues in length, are systematically compared pairwise for all selected final structures. In principle, this function provides a relative measure of how well the backbone atom positions in each amino acid in all final structures are defined. Consequently, such a comparison provides an objective method for the recognition of stable structural elements in the ensemble of final structures as the lower the mean r.m.s.d. the more similar the conformations in the final structures. Fig. 8A shows the results of the analysis of the 20 final structures of Vpr1–40 obtained in water with or without 50% TFE. The best defined region of the molecule, with the lowest r.m.s.d. (<0.15) is between and includes residues 17 to 32, which correspond to a well defined α-helix (Fig. 8, C and E, without and with TFE, respectively). The least defined region of the molecule, incorporating residues 1 to 14 (r.m.s.d. >0.6 Å), has no apparent extended regular structure and affords no observable medium range NOEs. The clustering of structures in the C-terminal region (Fig. 8, B and D) in both solutions implies that some regular structure is present and that this structure is oriented with respect to the central well defined helix. This feature in the structure begins in the region of residues 33 and 34, and clearly before Pro-35.

Generally, the consecutive segment approach as it was applied for the calculation of Vpr1–40 tends to underestimate the stability of fragments smaller than five residues in length. Thus, the method provides a suitable means for detecting α-helical structure but is probably less successful for identification of turns in which only three or four residues are held in a particular conformation. To test whether short structural elements exist in our 20 final structures, we have plotted the change in the summed r.m.s.d. according to the individual segment lengths evaluated (Fig. 9). Surprisingly, using this method a short structured element is evident in TFE solution within the 14-residue N terminus, characterized above as mainly unstructured. This two- to four-residue-long segment is centered on residues 5/6 and appears to exhibit more stable structure than the surrounding residues of the otherwise highly flexible N terminus. The existence of this short structure in the region of Pro-5 at the very N terminus is evident in all the individual analyses but most easily observed in the two to four-amino acid interval analyses.

In summary, the analysis of the ensemble of backbone conformations derived from the quantitative NOE data confirms...
the presence of a stable α-helix between Glu-17 and Arg-32 in 50% TFE. Most significantly, the same major helical domain has also been defined for the first time in water albeit shorter and less structured at its C-terminal end. In TFE a pronounced kink in the structure at residues 33/34 leads into a less well defined frayed helical C-terminus. The flexible N-terminal region incorporating three proline residues has a short relatively stable turn-like structure centered at residues 5 to 6.

**DISCUSSION**

Recent studies have shown that sVpr used in all our studies so far is biologically active when added at nanomolar concentration to the cell culture medium (20, 58); sVpr supports virus replication of vpr-deficient HIV-1 in cultured human macrophages, it translocates into cells where it modulates apoptosis, and it causes cell cycle arrest following transport into the nucleus. The nature of the mechanism(s) how extracellular and virus-free Vpr that was also found in the peripheral blood of AIDS patients (59) passes through the cell and mitochondrial membranes (60), and enters the nucleus is currently unknown (25). In its passage through the nuclear membrane Vpr displays nucleocytoplasmic shuttling properties reflecting the presence of an exportin-1-dependent nuclear export signal (61) and it causes alterations in the nuclear lamina of cells that lead to dynamic herniations and rupture of the nuclear envelope (22). For the transport of virus-associated Vpr released after uncoating of the incoming virus particle, Vpr is believed to traverse as part of the pre-integration complex the cytoplasm toward the nucleus through the microtubule network after interaction with cytoplasmic dynein (62).

To understand the transport of Vpr through membranes and its functional process in different intra- and extracellular compartments it is particularly important to elucidate the structural behavior of Vpr in various environments, and to determine its molecular interactions with other molecules. Structural studies of full-length Vpr by NMR techniques continue to be difficult because of protein oligomerization and a strong dependence of the structure on the solution condition (35), as well as an apparent heterogeneity in the composition of the molecule (36). In addition, previous reports (63, 64) demonstrated that this unusual behavior of Vpr results in difficulties during the synthesis of full-length Vpr derived from different HIV-1 isolates. Our previously established and now further optimized solid phase peptide synthesis protocol (20) permits the synthesis of sufficient sVpr necessary for structural analysis of either the full-length protein or fragments thereof. Noteworthy, all peptides are stable and devoid of any sign of protein precipitation even at concentrations suitable for long term NMR and DLS studies.

It is now reported that the oligomerization and folding of Vpr occurs domain-dependent and that the molecular heterogeneity observed for sVpr and its N-terminal fragments under all solution conditions arises through a so far unrecognized isomerization phenomenon that is associated with the conserved four N-terminal prolines. Two of these proline residues are present in cis conformations at an unusually high level implying a PPIase activity may be required for efficient folding of Vpr in vivo. The identification of a cis/trans phenomenon in Vpr, together with sequence similarity in the proline-rich N terminus of Vpr with a CypA binding domain located in HIV-1 capsid (for review see Ref. 65), led us to explore the as yet unidentified interactions of Vpr with CypA and its biological relevance, which are reported in detail in the accompanying paper (38).

Early studies (50) indicated that Vpr exists as an oligomer with an apparent molecular mass over 100 kDa and that residues 36 to 42 are critical for this phenomenon. Several follow-up analyses have tried to identify the oligomerization domain of Vpr (36, 66, 67). Among these it has been hypothesized that an N-terminal helix-turn-helix motif is the driving force in self-association of Vpr (36). The present DLS data analyzing self-association of sVpr and its fragments in its dynamic state in solution without artificial cross-linking reveal the existence of sVpr as a dimer and its C-terminal fragment as a hexamer whereas the N-terminal domain is monomeric. Thus, our data unambiguously show that the region(s) responsible for self-oligomerization lay in part or in total outside the 1–40 region and argue strongly against a previous assumption that considers the N terminus as driving force in the self-association of Vpr (36). However, complete clarification would require the DLS analysis of sVpr molecules that carry scrambled side chain positions within the C-terminal helix region. Nonetheless, it can be recapitulated that there are pronounced differences in the folding behavior and oligomerization of Vpr1-40 compared with the full-length protein and its C-terminal fragment, which again supports the idea that Vpr is organized into at least two different structural modules.

It is of interest to analyze the structure of the N-terminal domain of Vpr as mutational analyses had suggested that this domain of Vpr encompassing the first 40 residues is required for nuclear localization, packaging into virions, and binding of transcription factor (TFIIIB, Sp1) (68, 69), and viral (p65avag (67)) and cellular proteins (RIP1, UNG, karyopherins (70–72)). Since cis/trans isomerism at the N-terminal proline residues has the potential of influencing folding of the full-length molecule, it was imperative to investigate the high resolution structure of the N-terminal region and its dependence on the solution conditions. The major structure in Vpr1-40 is a well defined helix from residues 17–32 followed by a kink starting at residues 33/34 that leads into a less well defined fraying structure, which is compatible with the previously elucidated α-helix-turn-α-helix motif (36). Clearly, structures calculated using the quantitative NOE data for the major isomer identified in TFE resembles those determined previously (36) for the slightly longer N-terminal Vpr fragment derived from a different HIV-1 isolate. In our calculation, the helix is slightly longer than previously proposed and locates the turn prior to the proline residue at position 35. As it is common with all small flexible peptides, the calculation, employed cannot provide information on the stability of the helical conformations or information on the occurrence of a dynamic equilibrium between helical and non-helical structures. The r-6 dependence of the NOE tends to favor compact structures in an ensemble of structures, but this does not exclude other structures being present. Inspection of the relative intensities of the sequential d-N NOEs compared with d-N and the medium-range d-N (i, i+3) NOEs compared with d-N (i, i+3) are stronger than one would expect for an ideal stable helix and imply the occurrence of a dynamic equilibrium between helical and non-helical structures. The helical population for the helical region found above between residues 17 to 32 can be estimated from the ratios of the integrated intensities of the sequential d-N and d-N following a procedure enumerated previously (73). This affords ~89% (± 5%) on average over the pairs of residues 17/18, 20/21, 26/27, 31/32, and 32/33 observed for the helix in Vpr1-40.

Our structural analyses using the quantitative NOE data do not take into account the presence of signals from the cis isomers that can be distinguished from their trans partners in the vicinity of the proline residues (see above). Hence, some of the distance restraints in the vicinity of proline will be underestimated, particularly as distances are too large relative to those in the majority of the peptide where trans and cis signals are coincident. Consequently, the relative flexibility of these regions of the peptide must be considered as overemphasized in
our calculation. For the case in point, a critical area would be located between residues 31 and 38. Nevertheless, even a 15% loss in intensity of the peaks in this area consequent to the presence of the cis isomer would not account for the absence of a considerable number of (i, i-3) NOEs in the region between His-33 and His-40, and the presence of these NOEs in the region up to His-33. Consequently the presence of the cis isomer does not affect the conclusions drawn regarding the relative positions or stabilities of structured regions in the all-trans molecule.

Our data differ from previous NMR studies on Vpr (36) with regard to the existence of stable structures in the first 14 residues. Both our qualitative (1H chemical shifts and qualitative NOEs) and quantitative NOE data show little evidence of any extended secondary structure in this area. Furthermore, our r.m.s.d. analysis of the final structures in 50% TFE identified no unambiguous evidence of the three short structural elements, defined previously (36) as various types of turns that involve the first three proline residues. In the published analysis it was clear that those turns involving Pro-10 and Pro-14 had considerably higher r.m.s.d. (0.85) than the r.m.s.d. found for the long helix (0.20). To check the validity of such structures we have modified our consecutive segmental analysis to allocate stable structures that are even as short as 2 to 4 residues. This analysis locates only one relatively stable structure that incorporates Pro-5 and is centered at residue 6. Although Pro-14 appears to be in an area associated with the fraying at the N-terminal end of the well defined α-helix, we did not find any evidence for short turn structures involving other proline residues, particularly at position 10.

We have been able to monitor the dependence of the helical structure of the N-terminal half of Vpr on various solution conditions. There is a considerable differential in the shift changes on going from TFE solution to water under acidic conditions. Particularly, there is some loss in stability of the helical structure in the region after residue 23 whereas residues 18 to 23 prove to be the most stable part of the helix, which is present independently of stabilizing organic solvents. This is noteworthy as the leucine-rich region 20–26 is highly conserved among different HIV-1 isolates, and the N-terminal face of the helix incorporates part of the LLEEL motif that contributes to the Vpr-mediated co-activation of the glucocorticoid receptor (61). In keeping with our observations, a short peptide, Vpr19–36, corresponding to the N-terminal helix, showed high helical content in aqueous solution at pH 7 (74). In another study N-terminal peptides of Vpr caused ion channel formation and cell death in neuronal cells (57). It would be plausible that the strong helical fold in this region might provide the structural constraints for membrane interaction and formation of an ion conductive pore shown previously for full-length Vpr in lipid bilayers (10). The folding of the first 14-residue N-terminal proline-rich domain might be different in the lipid environment of bilayers and thus different from the data recorded in aqueous solution, in particular the turn near Pro-5 might gain in importance near or in a membrane. It would be interesting to verify the solution structures obtained in organic solvent using micelle solutions and solid state NMR techniques.

The heterogeneity of the signals in the NMR spectra of Vpr 1–14 provides the first evidence of cis/trans isomerization phenomena related to the prolines of the N terminus. These four prolines that are conserved among all known HIV-1 isolates (75) show significant proportions of cis isomer such that at least ~40% of all Vpr molecules contain one proline residue in a cis configuration. Evidently, Pro-35 and most probably Pro-14 are the residues possessing the highest cis-content. According to studies on different proteins, this might arise from interactions with adjacent aromatic residues (76). Two aromatic residues are located distal and proximal to the N-terminal helix. Most strikingly, Pro-35 resides in a relatively flexible but still structured region that is part of a turn leading into a second helical region (36). Clearly, the isomerization of this residue will have a profound influence on the relative orientation of the two most stable N-terminal helices, residues 17–32 and 35–46 (36). This is further substantiated by the observation that Pro-35 regulates the functional interaction between Vpr and CypA in vivo (38). Isomerization might be less important for Pro-5 and Pro-10 where the percentages of cis isomer are much lower. Both residues are located in considerably more flexible regions, and only a short section of stable structure centered on residue 6.

The discovery of a cis/trans isomerism at the proline residues of Vpr is relevant for the biological function of Vpr (see accompanying paper (38)). First, the major cellular cis/trans-proline isomerase CypA binds to Vpr in a fashion that is dependent on the N-terminal proline residues. Second, the position of the prolines of Vpr share striking sequence similarities with the CypA binding motif in HIV-1 CA. Third, CypA inhibitors, mutation of proline residues in Vpr, or genetic inactivation of CypA expression block de novo synthesis of Vpr and interfere with cell cycle arrest mediated by Vpr in HIV-1-infected cells. Fourth, both Vpr and CypA are present in significant amounts in HIV-1 particles (for review see Ref. 65).

Vpr is the only regulatory HIV-1 protein that is specifically encapsidated into virions at significant amounts by a mechanism that is mediated by interaction between Vpr and the C-terminal p60Gag domain of the Gag polyprotein Pr55 (77). However, it is assumed that during virus maturation this firm Vpr-p60Gag interaction disappears as Vpr and p60Gag localizes to different structures inside maturing viral particles (21, 78). Thus, changes in folding of binding partners Vpr and Gag, which might even exist in a tertiary complex together with CypA in budding viruses, are likely to occur during virus maturation, a process that could be influenced by the isomerase activity of CypA. Studies using CypA inhibitors, mutants of CA, or gene silencing of CypA (78) did not reveal significant impact of CypA on morphology or maturation of HIV-1 virions although it is well established that CypA increases the infectivity of HIV-1. In addition to the supposed effect on Gag in the process of virus uncoating, CypA might also regulate functional folding of Vpr after it is released from incoming virions necessary to support nuclear import of the PIC. Alternatively, CypA might also regulate the late function of Vpr in the virus replication cycle. Indeed, data from our ongoing work demonstrate that expression of Vpr requires co-translational interaction with CypA and that CypA activity is required for the Vpr-mediated G2 cell cycle arrest in HIV-1-infected T cells (38). As the interaction of CypA with HIV-1 proteins Gag and Vpr can be blocked by CypA inhibitors such as cyclosporin A, our observation might open new avenues for anti-viral therapies, particularly as it was shown recently (79) that highly active antiretroviral therapy can benefit from combining with cyclosporin A treatment.

Acknowledgments—We thank S. Weißlag for DLS, M. Nimtz for MS, and F. Kastner for peptide synthesis. We are indebted to J. W. Yewdell and J. R. Bennink for continuous support. We thank K. Zander and D. Ott for helpful comments to the manuscript.

REFERENCES
1. Paxton, W., Connor, R. I., and Landau, N. R. (1993) J. Virol. 67, 7229–7237
2. Sherman, M. P., de Noronha, C. M., Heusch, M. I., Greene, S., and Greene, W. C. (2001) J. Virol. 75, 1522–1532
3. Gibbs, J. S., Lackner, A. A., Lang, S. M., Simon, M. A., Sehgal, P. K., Daniel, M. D., and Desrosiers, R. C. (1995) J. Virol. 69, 2378–2383
4. Lang, S. M., Weeger, M., Stahl-Hennig, C., Coulbaly, C., Hunsmann, G.
