Spectroscopic investigations of HIV-1 trans-activator and related peptides in aqueous solutions

Dagmar Klostermeier a,b, Peter Bayer a,b,*, Margot Kraft c, Rainer W. Frank c, Paul Rösch a

a Lehrstuhl für Biopolymere, Universität Bayreuth, D-95440 Bayreuth, Germany
b Max-Planck-Institut für Molekulare Physiologie, Rheinlanddamm 201, Postfach 102664, D-44026 Dortmund, Germany.
c Zentrum für Molekularbiologie Heidelberg (ZMBH), D-69120 Heidelberg, Germany

Received 4 March 1996; revised 28 September 1996; accepted 30 September 1996

Abstract

The 86 amino acid trans-activator (Tat) protein of human immunodeficiency virus type 1 (HIV-1) is an RNA-binding transcriptional regulator. HIV-1 Tat proteins (wild type and Thr40Lys mutant) and the HIV-1 Tat peptide fragments Tat(32–48) and Tat(32–72) were chemically synthesized. One- and two-dimensional nuclear magnetic resonance spectroscopy experiments were performed to elucidate the structural features of these proteins. In fluorescence quenching studies of the full-length Tat protein (Thr40Lys), Trp11 was found to be only partially protected against solvent accessibility. Circular dichroism melting studies monitored a slight cooperative change in the conformation of the Tat with increasing temperature. Backbone NH protons of amino acids located in the main core element of the protein are partially protected against exchange.

Keywords: Tat protein; Solution structure; Protein RNA interactions; HIV; EIAV; Lentiviruses

1. Introduction

Trans-activator (Tat) proteins are essential regulatory elements in the replication cycle of all known lentiviruses (see [1,2] and references cited therein). Tat proteins of immunodeficiency viruses form a complex with a viral RNA sequence (Tat-responsive element; TAR) and enhance the expression of viral genes in the host cell DNA in the presence of cellular co-factors. Human immunodeficiency virus (HIV) is the causative agent of acquired human immunodeficiency syndrome (AIDS). HIV type 1 (HIV-1) Tat protein was suggested to form a com-
plex with transcription factor TFIID [3], but details of the mechanism of action of Tat proteins in complex formation and gene activation are largely unknown. In addition to its role in transcriptional regulation, HIV-1 Tat protein may be involved in the generation of Kaposi’s sarcoma in AIDS patients [4] and in T-cell apoptosis [5,6].

Conventionally, the HIV-1 Tat protein sequence is subdivided into regions on the basis of local homologies [7]; specifically, in an NH₂-terminal region (Met1–Ala21), a cysteine-rich region (Cys22–Cys31), a core region (Phe32–Tyr47), a basic region (Gly48–Arg57), a glutamine-rich region (Ala58–Gln72) and a COOH-terminal region (Pro73–Glu86).

To date, structural studies of HIV Tat proteins could only be carried out on a very limited scale, because of the shortage of transcriptionally active protein, although the development of an efficient expression, purification and refolding procedure was reported recently [8]. Investigations have been limited to experiments using circular dichroism (CD) spectroscopy [9] and nuclear magnetic resonance (NMR) spectroscopy of synthetic peptides with the sequence of the RNA binding domain and, in some cases, adjacent residues [10]. Nevertheless, the structures of the HIV-1 TAR RNA [11] and a peptide–TAR complex of bovine immunodeficiency virus (BIV) [12,5] have recently been determined.

The structure of the Tat protein of equine infectious anemia virus (EIAV), which is a close relative of HIV-1, is known from two- and three-dimensional NMR spectroscopy and molecular dynamics calculations [13]. EIAV Tat protein consists mainly of a hydrophobic core, and NH₂- and COOH-terminal loop regions. It was inferred to be a highly flexible protein with a tendency to form very weak helical elements in the NH₂-terminal, core and basic regions [13]. Under mildly hydrophobic conditions, however, the helical secondary structure elements are stabilized, at the expense of the stability of the hydrophobic core [14].

The structure of a chemically synthesized HIV-1 Tat (Thr40Lys) protein was recently determined in solution by NMR spectroscopy [15,16]. In general, the structure was similar to the EIAV Tat protein structure, but did not show any tendency towards helix formation. Here, we present details of spectroscopic studies to gain some insight into the time-dependent behaviour of this chemically synthesized protein.

2. Material and methods

2.1. Protein synthesis and refolding

Wild-type and Thr40Lys mutant HIV-1 Tat protein (SwissProt data bank sequence entry HV1Z2) were chemically synthesized, purified and refolded as follows [15,17]. 10 ml of a 1 mg ml⁻¹ solution of Tat protein was unfolded in 1 l of buffer solution that contained 20 mM potassium phosphate (pH 6.5), 100 mM ZnCl₂, 50 mM mannitol, 10 mM ascorbic acid, 0.5 mM phenylmethylsulphonyl fluoride, 0.5 mM dithiothreitol (DTT) and 6 M urea. Refolding was performed by stepwise dialyses against 4 M, 2 M and 0 M urea in the same buffer, and two final dialyses against 10 mM ammonium acetate (procedure A). For comparison, the same protocol was followed, except that 100 mM ZnCl₂ was present in all the solutions (procedure B). Helium-degassed buffers were used throughout, and the protein was lyophilized twice to remove ammonium acetate and then stored as a lyophilized powder.

2.2. NMR sample preparation

100 ml H₂O, 30 mM NaCl and 0.02 vol.% NaN₃ was degassed for 2 min at a pressure of 10 N. The buffer was then bubbled with N₂ to saturation and kept under an N₂ atmosphere. Na₂S,O₃ was added to give an amount of 50 μM, such that the final pH was approximately 6.0. The lyophilized protein was dissolved rapidly in 900 μl of buffer solution, still under N₂ atmosphere, and the volume was reduced to approximately 450 μl in a Speedvac vacuum concentrator. An NMR sample tube was flushed with N₂ for approximately 10 min and the pH value was determined to be approximately 6.3. N₂ was subsequently passed over the sample in the NMR tube and the sample tube was then sealed. Protein concentrations of up to 2.8 mM could be obtained without precipitation with this procedure.

During dialyses in the presence of DTT in excess of 5 mM, HIV-1 Tat protein precipitated, and an increase in molecular mass of the order of an integer
multiple of the DTT molecular masses was detected by matrix-assisted laser desorption ionization–time of flight mass spectrometry. Observations that HIV-1 Tat protein is inactivated by DTT [18] are in agreement with these results. Thus, Na₂S₂O₄ was used as a reducing agent.

2.3. NMR spectroscopy

For HIV-1 Tat (wild type and Thr40Lys), the experimental conditions and data evaluation procedures were as described before [16]. Spectra at pH 8 were recorded with a 2 mM oxygen-free protein solution (Thr40Lys) in 20 mM sodium borate buffer–HCl (pH 8) with 200 μM Na₂S₂O₄ and 0.04% NaN₃ at 298 K. For the peptides, double quantum-filtered correlated spectroscopy (DQF-COSY) and CLEAN total correlation spectroscopy (TOCSY) were used with a mixing time of 90 ms, along with nuclear Overhauser enhancement spectroscopy (NOESY) with a mixing time of 200 ms at a temperature of 298 K. The peptide concentration was 2 mM, with 50 mM NaCl and 0.04 vol.% NaN₃ in potassium phosphate buffer (pH 6.3), as well as 40 vol.% trifluoroethanol (TFE).

In the HD-exchange experiment, an oxygen-free protein solution (1 mM Thr40Lys) in 10 mM phosphate buffer (pH 6.0–6.3) with 100 μM Na₂S₂O₄ and 0.02 vol.% NaN₃ was lyophilized and stored at −20 °C for 10 h. Prior to the experiment, the protein powder was diluted in 500 μl of a preincubated (298 K) D₂O–flushed H₂O solution and injected into a 5 mm NMR tube. After a short shimming procedure, one-dimensional NMR spectra at 298 K, using a 90° pulse of 10 μs and 16–32 scans per experiment. The dead time before starting acquisition (dilution and shimming) was about 2 min. The data were evaluated on X-window workstations using the NDee program package (Software Symbiose GmbH, Bayreuth).

2.4. Molecular dynamics calculations

Structures were calculated using the X-PLOR program package [19], as described elsewhere [16].

2.5. CD spectroscopy

CD spectra were recorded from 250 to 190 nm using a Jasco J 600 CD spectropolarimeter at 10 and 60 °C. The sample used contained 20 μM HIV1Z2 Tat protein (Thr40Lys) in 10 mM potassium phosphate buffer (pH 6.3) with 10 mM NaCl. The spectra were accumulated fourfold with a scan speed of 20 nm min⁻¹ and corrected for buffer effects. For measurements of the temperature dependence of [θ]MRE222nm, the protein solution was equilibrated at 10, 20, 30, 35, 40, 45, 50 and 60 °C for 10 min. The [θ]MRE222nm values for each temperature were averaged (2 min) and corrected for buffer effects.

2.6. Fluorescence spectroscopy

Fluorescence emission spectra were recorded using a Hitachi F-4500 spectrofluorimeter equipped with a stirrer. The fluorescence was excited at 280 nm (bandwidth, 3 nm), and the emission was recorded with a bandwidth of 5 nm. The protein concentration (Thr40Lys) was 2 μM in 10 mM potassium phosphate buffer (pH 6.3) that contained 10 mM NaCl, or in 10 mM potassium phosphate buffer (pH 6.3) that contained 10 mM NaCl and 6 M GdmCl. All the spectra were corrected for buffer fluorescence.

2.7. Fluorescence quenching

Aliquots of a 6 M acrylamide solution in 10 mM potassium phosphate buffer (pH 6.3) with 10 mM NaCl or 100 mM tris (pH 7.0), with 150 mM NaCl were added to a 5 μM protein solution (Thr40Lys) in identical buffer solutions up to an acrylamide concentration of 0.4 M. The fluorescence emission at 355 nm was recorded after each addition (bandwidth, 10 nm), following selective excitation of tryptophan fluorescence at 295 nm (bandwidth, 1 nm). The fluorescence intensities were multiplied by 10^A², where A is the UV absorption at 295 nm in 1 cm cuvettes, to correct for the inner filter effect caused by the absorption of acrylamide at this wavelength; the results were also corrected for volume increases.

3. Results and discussion

The biological activity of Tat protein obtained as described above was shown by the transactivation of HIV-1 long terminal repeat (LTR) in HeLa cells transfected with an HIV-1 LTR–chloramphenicol...
The structure of the free HIV-1 Tat protein (Thr40Lys) in solution under oxygen-free conditions exhibited a hydrophobic core built partly from the core- and glutamine-rich regions, and partly from the NH$_2$-terminal region, including the amino acid Trp11 [16]. Only a few long-range nuclear Overhauser effects (NOEs) from side-chain protons characterize this main structural element. Because NOEs measure only a time-averaged distance, they may be caused by two protons being in close proximity for a short period of time only. In this case, the high interior dynamics of the HIV-1 Tat protein NH$_2$-terminus could lead to a more solvent-exposed Trp11 than would expected from the rigid structural model, where this residue exhibits NOEs on core amino acids.

To investigate this possibility, fluorescence measurements of the protein (Thr40Lys) in the phosphate buffer solution were performed. The emission spectra exhibit a maximum at 355 nm in the presence as well as in the absence of 6 M guanidinium chloride (Fig. 1), which corresponds to the emission maximum of tryptophan in aqueous solutions [22]. Thus, there is no indication of a change in the polarity of the environment of the Trp11 on unfolding. Therefore, the solvent accessibility of Trp11 was investigated with acrylamide as a quenching agent (Fig. 2). The protection factor for the tryptophan in HIV-1 Tat protein is around 4 ($K_Q = 3 \text{ M}^{-1}$), taking N-acetyl-tryptophan-amide as a fully solvent-exposed standard ($K_Q = 11 \text{ M}^{-1}$). Thus, it is suggested that Trp11 is only partially protected and that it is not entirely buried in a rigid structure. This is in good agreement with the proposed high internal dynamics of the NH$_2$-terminal chain.

Melting studies using CD at various temperatures from 10 to 60°C were carried out to monitor possible cooperativity in the temperature behaviour of Tat (Fig. 3(a)). The spectral changes observed with increasing temperature are reversible. Similar spectra have been recorded in studies on hydrophobic peptides of platelet factor 4 [23]. A small increase in B-sheet could be concluded from the CD difference spectrum, with its typical maximum and minimum (Fig. 3(b)). B-sheet structures could also be detected in Tat proteins binding to their corresponding TAR-RNA, as in the case of BIV-Tat [12,24]. Moreover, an extended B-type structure on HIV-1 Tat–TAR complexation was suggested from CD and Fourier transform IR studies [25].
Fig. 3. CD spectra of HIV-1 Tat (Thr40Lys, 20 μM) in potassium phosphate buffer (10 mM, pH 6.3, with 10 mM NaCl) at various temperatures: (a) spectra at 10 and 60°C; (b) CD difference spectrum (CD_{60°C} - CD_{10°C}). Reversible refolding of the protein was achieved by cooling down to 10°C in about 1 h, followed by an equilibration step of duration 30 min at 10°C.
CD and NMR spectroscopy of HIV Tat (Thr40Lys) did not indicate the presence of extended secondary structural elements at room temperature [16]. To verify this, HD-exchange experiments were performed with full-length HIV-1 Tat protein (Thr40Lys). A sodium buffer solution (10 mM, pH 6.0–6.3) with 100 μM Na_{2}S_{2}O_{4} and 0.02 vol.% NaN_{3} of the protein was lyophilized and resuspended in D_{2}O (see Section 2). NH protons involved in hydrogen bridging, such as in regular secondary structures with some exceptions (first four residues of an α-helix, first three residues of a 3_{10}-helix and every second residue of a peripheral β-strand; [26]), should (for a certain timescale) be protected against exchange, so should be visible in ^{1}H-NMR spectra. All unprotected NH protons, however, undergo a much faster exchange process and should not be observable.

In one-dimensional NMR spectra at pH 6, we found that all the NH protons disappeared within the deadtime of the experiment. From these data, one would conclude that the protein exhibits no secondary structural elements stabilized by hydrogen bridging under the given conditions. Because Tat seems to be a very flexible molecule, internal dynamical processes that take place on a millisecond or second timescale might lead to solvent accessibility of NH protons buried in a hydrophobic core. Therefore, an experimental deadtime of the order of minutes would allow all the amide protons to exchange.

To make a distinction between NH protons that are protected and those that are not, we recorded

Fig. 4. Temperature dependence (10–60°C) of the mean residue ellipticity at 222 nm for Tat (Thr40Lys, 20 μM) in potassium phosphate buffer (10 mM, pH 6.3, with 10 mM NaCl).

Following the ellipticity at 222 nm, we could observe a slight cooperative change in the HIV-1 Tat conformation (Fig. 4). This behaviour may be explained by strengthened hydrophobic interactions in the main core structure of HIV-1 Tat protein at increased temperatures. The formation of a β-sheet structure might be involved in this process. A red-shift of the Trp11 absorption with increasing temperature confirms the change in the aromatic side-chain to a more hydrophobic environment (data not shown).
one-dimensional NMR spectra at pH 8 (2 mM protein (Thr4OLys), 20 mM borate buffer, 200 μM Na₂S₂O₇ and 0.04 vol.% NaN₃) and a temperature of 298 K in H₂O. The intrinsic exchange rate constants of unprotected backbone amide protons at pH 8 (298 K) were in the range 10⁻¹⁻¹⁰⁻⁴ min⁻¹ [27]. Thus, the observation of labile backbone amide protons with ¹H-NMR spectroscopy is only possible if solvent exchange is slow, i.e. much less than 10⁻¹ min⁻¹. The spectra showed non-exchanged backbone amide protons belonging to Val4 or Val36 (ambiguous), Leu43, Gln66, Ile69, Lys71, Glu86, and Ile8 or Ile45 (ambiguous; Fig. 5).

The backbone amide protons of the eight amino acids are probably less solvent accessible, so are better protected against exchange, but are not involved in rigid hydrogen bridging, according to the HD-exchange experiments. Although a structure determination of Tat at pH 8 has not been possible up to now, there is an obvious correlation between the amide protons involved in NOE contacts ([16]: Fig. 6) at pH 6.0–6.3 and the 'slow' exchanging NH protons observed at pH 8. All these amino acids—except for Glu86—are involved in the main core element of the protein.

Statistical evaluation of 90 calculated structures of HIV-1 Tat (Fig. 7) using about 600 NOE constraints [16] and an X-PLOR molecular dynamics protocol [28] demonstrates that these amino acids are located in structural elements that indicate partially turn characteristics. Only in a few structures could helical elements be detected. An additional D₂O exchange
Fig. 8. Chemical shift differences of Cα protons between random coil values and Tat (Thr40Lys) at 298 K. The random coil values (pH 5, 298 K) of the Cα-shifts were chosen from [29]. Correction was made for shifts of residues prior to a proline in the sequence. 4,4-dimethyl 4-silapentane sodium sulfonate was used as an internal standard. Interpretation of the Cα-shifts of amino acids that have pK values between 5 and 7 (such as histidine) has to be carried out cautiously, because no pH correction of the sample could be achieved.

Experiment with Tat(32–72) in aqueous solution indicated amino acids that contained amide backbone protons that exchange with a half-life of approximately 1 h at pH 6.0. Again, these protons belong to the core- and glutamine-rich domain, but not to the basic domain (Fig. 5). This result is in good agreement with the NMR structure of the full-length HIV-1 Tat protein.

From analysis of chemical shift values of Cα protons [29] in aqueous solution at room temperature, a tendency to form helical secondary structures rather than β-sheet structures may be inferred for the HIV-1 Tat protein (Fig. 8). Most of the medium-range NOEs are observed between resonances of side-chain protons (Fig. 6), corroborating the lack of rigid secondary structure under the present conditions. Only a small number of very weak non-sequential NOESY cross-peaks close to the NH2-terminus at the beginning of the core sequence region and in the glutamine-rich region could be interpreted as indicating the transient presence of helical elements or turns. A lack of regular secondary structure elements in uncomplexed HIV-1 Tat protein is also indicated by studies of peptide analogues to the basic region [30,31]. In contrast, an α-helical structure was reported for the basic sequence region in a fusion peptide spanning the EIAV Tat protein core region and a modified HIV-1 Tat protein basic region [32].

In previous studies on EIAV Tat protein, elements of regular secondary structure could only be detected as limiting structures or as structures in hydrophobic solutions [13,14]. Because the full-length HIV-1 Tat was not soluble in TFE–water solutions in concentrations suitable for NMR spectroscopy, we chemically synthesized a Tat peptide that contained the 15 amino acid core domain only. This peptide could be studied with the same conditions under which the EIAV Tat protein (TFE/buffer, 40/60 (v/v); pH 6.3) was found to be helical [14]. For Tat(32–48) in TFE–buffer, αN(i,i+3), αβ(i,i+3) and αN(i,i+4) NOEs indicate the presence of an α-helix from amino acids Gln35 to Lys41, whereas the

![Diagram](attachment:image.png)

**Fig. 9.** NOE connectivities of Tat(32–48) in (a) TFE–buffer (40/60 (v/v)) and (b) DMSO (black bars). Positions where NOEs could not be assigned unambiguously are marked by shaded bars. The NOESY mixing time was 200 ms.
peptide shows random coil behaviour in dimethylsulphoxide (DMSO) solution (Fig. 9).

4. Conclusions

The situation with the HIV-1 Tat protein is similar to the situation with the EIAV Tat protein [13]. The number of \(|i-j| > 5\) NOESY cross-peaks that define the tertiary structure is very low in both proteins (EIAV Tat, 34; HIV-1 Tat, 25). The main structural element of both Tat proteins is a hydrophobic core (EIAV Tat, Tyr35–Tyr49; HIV-I Tat, Phe/Tyr32–Tyr47). NH protons of amino acids that belong to this region are partially protected against exchange with solvent protons. However, fluorescence experiments indicate that the side-chain of Trp11 is not entirely buried in the hydrophobic core. This may be explained by high flexibility of the whole N-terminal domain.

Acknowledgements

Drs. T. Parslow and T. James (UCSF) made unpublished data available. We thank Drs. M. Beissinger, F.X. Schmid, S. Walter, C. Frech and T. Schindler for help with the CD and fluorescence measurements. This work was supported by fellowships from the Fonds der Chemischen Industrie, the Deutsche Forschungsgemeinschaft (ROE 17/4-1 and FR881/2-1) and the BIOMED I-AIDS RESEARCH program of the European Union (BMH1-CT93-1175). The NDee program is available from the authors on request.

References