# Intrinsic Disorder and Function of the HIV-1 Tat Protein

Shaheen Shojania and Joe D. O'Neil*

*Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada*

**Abstract:** The type 1 Human Immunodeficiency Virus transcriptional regulator Tat is a small RNA-binding protein essential for viral gene expression and replication. The protein binds to a large number of proteins within infected cells and non-infected cells, and has been demonstrated to impact a wide variety of cellular activities. Early circular dichroism studies showed a lack of regular secondary structure in the protein whereas proton NMR studies suggested several different conformations. Multinuclear NMR structure and dynamics analysis indicates that the reduced protein is intrinsically disordered with a predominantly extended conformation at pH 4. Multiple resonances for several atoms suggest the existence of multiple local conformers in rapid equilibrium. An X-ray diffraction structure of equine Tat, in a complex with its cognate RNA and cyclin T1, supports this conclusion. Intrinsic disorder explains the protein's capacity to interact with multiple partners and effect multiple biological functions; the large buried surface in the X-ray diffraction structure illustrates how a disordered protein can have a high affinity and high specificity for its partners and how disordered Tat assembles a protein complex to enhance transcription elongation.

**Keywords:** Circular dichroism; HIV-1; intrinsically disordered proteins; NMR spectroscopy; Tat; X-ray diffraction.

## 1. THE HUMAN IMMUNODEFICIENCY VIRUS

The Type 1 Human Immunodeficiency Virus (HIV-1) is a lentivirus, the latter comprising a genus of retroviruses displaying a characteristically long latency period [1,2]. Retroviruses contain 2 copies of positive-sense single stranded RNA genomes (Fig. (**1a**)) which are reverse-transcribed and then integrated into the DNA of an infected cell [3]. The first cells infected by HIV-1 are CD4+ T-lymphocytes, but macrophages and microglial cells can also become infected [1,4]. The mature HIV-1 virion, depicted in Fig. (**1a**), is characterized by a cone-shaped capsid (CA) core enveloped by a lipid bilayer derived from the host cell. The capsid contains the viral RNA genome along with several viral enzymes necessary for early replication steps [5]. The HIV-1 genome contains nine open reading frames [6] as depicted in Fig. (**1b**). The group-specific antigen gene (*gag*) encodes a polyprotein (Gag) that contains the major structural components of the virus (matrix, MA; capsid, CA; nucleocapsid, NC) and p6. The *pol* gene encodes a polyprotein (Pol) containing reverse transcriptase (RT), integrase (IN) and protease (PR). In the polyproteins, the individual folded proteins are connected by unstructured linker segments which are cleaved by the viral protease [2]. The envelope (*env*) gene encodes the Env proteins glycoprotein 120 (gp120) and glycoprotein 41 (gp41) which make up the surface unit (SU) and transmembrane (TM) complexes [6]. The six additional reading frames contain the genes for the regulatory proteins transactivator of transcription (Tat) and regular expression of virus (Rev), and the accessory proteins viral infectivity factor (Vif), viral protein U (Vpu), viral protein R (Vpr), and negative factor (Nef) [2]. The Tat and Rev proteins are produced by joining discontinuous segments of genomic RNA (Fig. (**1b**)).

*Address correspondence to this author at the Department of Chemistry, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada; Tel. 204-474-6697; Fax. 204-474-7608; E-mail: joneil@cc.umanitoba.ca.

Expression of viral DNA produces three major classes of mRNA including unspliced (9 kb), partially spliced (4 kb) and multiply spliced (1.8 kb) messages [7,8]. The full length, unspliced mRNAs exit the nucleus and can serve for translation or they can assemble at the cell membrane for packaging into a new virion. The most abundant partially spliced 4 kb mRNA encodes the envelope glycoproteins [7]. Multiply-spliced mRNAs are transported to the cytoplasm and translated into the viral regulatory proteins Tat and Rev along with the accessory protein Nef. The viral regulatory protein Tat is essential for virus expression – its main role, described in Section 3, is to enhance the processivity of RNA Polymerase II (RNAP II) leading to the production of full-length transcripts.

## 2. INTRINSICALLY DISORDERED PROTEINS

Over the past 20 years, it has become increasingly clear that disordered proteins can be biologically active and that the disordered segments of folded proteins are often functional. In some cases, protein function is expressed during transitions between different unfolded [9-11], partially folded [9,12,13], molten globule [14], and amyloid states [15-17] of proteins. The extent of protein disorder in the biosphere has been measured with the use of algorithms that detect disordered segments of proteins based on amino acid properties such as charge, hydropathy, secondary structure propensity and backbone flexibility, and their frequency of occurrence throughout a protein [9,18-31]. Genome analysis suggests that disordered segments of 30 or more consecutive residues occur in 2.0% of archaean, 4.2% of eubacterial and 33.0% of eukaryotic proteins [29]. They also indicate that viruses have the highest fraction of proteins containing a disordered segment (37%) compared to bacteria (30%), eukaryotes (21%) and archaea (19%) [32]. Thus, evolution appears to have selected amino acid sequences that can exist in multiple conformational states; one of the advantages of

0929-8665/10 $55.00+.00 

this versatility is the possibility of assigning multiple functions to individual proteins a property sometimes referred to as "*moonlighting*" [33].

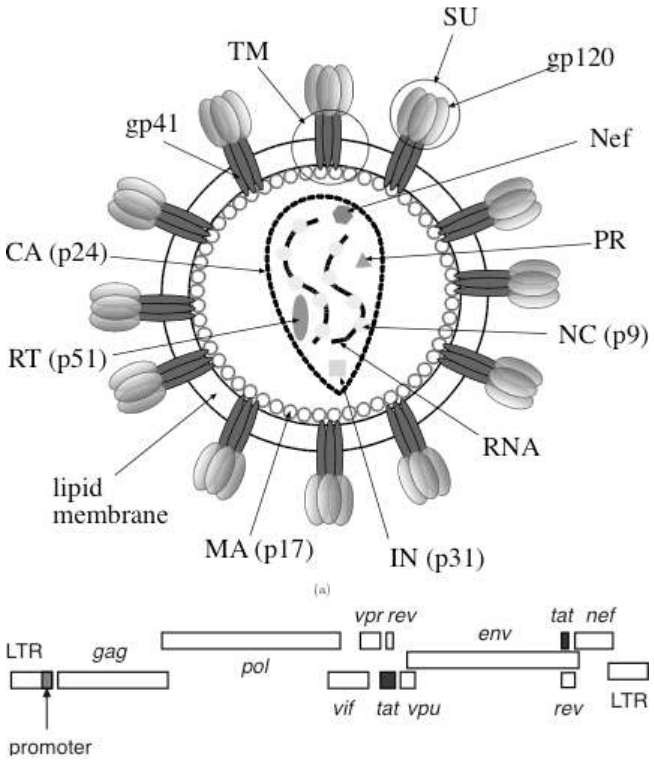



**Figure 1. (a)** Features of the HIV-1 virion showing the conical capsid comprised of ~1500 copies of the capsid protein (CA, p24). The core contains the viral diploid single-strand RNA [5], nucleocapsid protein (NC, p9), protease (PR), integrase (IN), negative factor (Nef), integrase (IN or p31) and reverse transcriptase (RT, p51). The capsid core is enclosed in a protein matrix (MA, p17). The matrix is enveloped by a lipid bilayer derived from the host cell along with some host cell proteins. The surface unit (SU) is comprised of trimers of glycoprotein 120 (gp120) which are anchored to the envelope via the transmembrane (TM) complex consisting of a trimer of glycoprotein 41 (gp41). **(b)** Open reading frames of the HIV-1genome. The HIV-1 long terminal repeat (LTR) has an inducible promoter [159] followed by the genes for: group-specific antigen (*gag*) encoding a polyprotein containing the major structural components of the matrix, capsid and nucleocapsid complexes; polyprotein (*pol*) encoding another polyprotein containing the viral enzymes protease, reverse transcriptase and integrase; accessory proteins viral infectivity factor (*vif*); viral protein R (*vpr*); viral protein U (*vpu*) and negative factor (*nef*); envelope (*env*) encoding the surface and transmembrane glycoproteins; and the regulatory proteins transactivator of transcription (*tat*) and regular expression of virus (*rev*). Both of the regulatory proteins are encoded by two exons.

The functions of intrinsically disordered protein (IDP) have been classified into six broad categories [34-36]: *Effectors*, proteins that activate, or more commonly inhibit, activities such as transcription initiation; *assemblers* of macromolecular complexes such as ribosomes, *chaperones* that provide a protective function to proteins and RNA, *display sites* for post-translational modification, *scavengers* of small ligands such as calcium phosphate by casein, and *entropic chains* that can be dynamic linkers or regulators of access to binding sites, active sites, and pores [10,34-42]. Recent analysis supports earlier observations that intrinsically disordered proteins are particularly abundant in processes involving the binding of DNA and RNA, and in the regulation of cell cycle, transcription, differentiation, and growth [34,43].

Several advantages have been proposed for disordered proteins over folded proteins [44-46]. IDPs have significantly greater surface areas than folded proteins of the same length endowing them with the ability to bind multiple partners and to have multiple functions [17,47] [48]. This binding promiscuity, which can modulate the activity of different targets, has been observed for several IDPs [10,34,49] and can have opposing effects on the same target [34]. However, in all of the categories except for the entropic chains, interactions between a disordered protein or segment and its target typically result in some degree of disorder-to-order transition [34]. The large surface areas of unfolded proteins can contribute to high specificity of binding but the large entropy loss upon folding will, in principle, reduce binding affinity [50]. Surprisingly, a survey of the binding affinities of folded and intrinsically disordered proteins shows that the range of binding affinities is remarkably similar in both classes [51]. High-affinity binding in IDPs may be achieved in part by the counteracting solvent entropy gain that occurs upon burial of a large hydrophobic surface. It has also been suggested that decoupling folding and binding permits faster dissociation of high affinity IDPs than is possible for folded proteins [34,51]. In this regard, recent dynamics analyses of several proteins [52,53] and RNA [54] suggest that many biomolecular interactions entail conformational selection of a target that samples the bound state in its unbound ensemble. However, more work is needed to dissect the multiple enthalpic and entropic contributions that explain IDP folding transitions [50].

### 2.1. NMR Spectroscopy and Disorder

Disorder in proteins can be detected by a variety of methods but perhaps the most frequently used are circular dichroism (CD) spectropolarimetry, nuclear magnetic resonance (NMR) spectroscopy, X-ray diffraction, and protease digestion. With the appropriate sample, NMR spectroscopy is uniquely capable of determining the three-dimensional structures of proteins in solution at atomic resolution [55]. A further advantage of NMR spectroscopy is the ability to obtain high-resolution dynamic information over a range of timescales that spans 12 orders of magnitude from picoseconds to seconds [56]. Advances in NMR spectroscopy over the past decade have made it particularly well suited to studying the dynamics of flexible segments in folded proteins, protein folding intermediates such as the molten globule, and denatured and intrinsically disordered proteins [18,19,57,58]. A $^{1}$H one-dimensional NMR spectrum can indicate disorder through the low dispersion of the resonances [58]; more commonly, disorder is evident in the narrow dispersion of resonances in the $^{1}$H and $^{15}$N dimensions of Heteronuclear Single-Quantum Coherence (HSQC) spectra. Isotope enrichment of proteins in $^{15}$N and $^{13}$C is essential; the comparatively wide dispersion in the $^{15}$N-amide, $^{13}$C-carbonyl, and $^{1}$H-amide resonances in disordered proteins in combination with multi-dimensional experiments

and high magnetic fields permits the unambiguous resonance assignment of disordered proteins that would be nearly impossible in unlabelled proteins [59,60].

A disadvantage of NMR spectroscopy is the fact that it requires sample concentrations in the 0.5-1 mM range, although this can be as low as 100 μM in very high magnetic fields. This can be a serious problem in the study of the disordered state since disordered proteins are prone to aggregation at these concentrations [58]. Disordered proteins are usually not uniformly hydrophilic and their hydrophobic residues lead to low solubility. Thus, preparing and maintaining a monodisperse sample can be challenging [61].

## 3. THE HIV-1 TRANSACTIVATOR OF TRANSCRIPTION

All lentiviruses produce a transactivator of transcription including the viruses found in primates, sheep, horses, cats, and cattle [62]. HIV-1 Tat is expressed during the early stages of viral infection and is essential to viral replication [63]. Shortly after initiation of transcription of the HIV-1 DNA, RNA Polymerase II (RNAP II) halts as a result of the binding of the negative transcription elongation factors N-TEF (negative transcription elongation factor) and DSIF (5, 6-dichloro-1-β-D-ribofuranosylbenzimidazole sensitivity-inducing factor) [64-66]. In the absence of Tat protein, RNA polymerase is hypophosphorylated, exhibits low processivity, and releases only short transcripts [67]. The main effect of Tat is to stimulate hyperphosphorylation of the carboxyl terminus of RNAP II enhancing the processivity of the enzyme and leading to the production of full-length transcripts. A brief outline of the mechanism begins with the binding of Tat to a stable, nuclease-resistant, stem-loop structure referred to as the transactivation response (TAR) element (Fig. (**2**)). The TAR element is located downstream of the long terminal repeat (LTR) promoter and spans nucleotides +1 to +59 of the nascent RNA [68]. Tat stimulates elongation of full-length transcripts by recruiting the positive transcription elongation factor b (P-TEFb), a hetero-dimeric complex of a regulatory cyclin T and CDK9. Upon formation of the P-TEFb/Tat-TAR complex, CDK9 is brought into close proximity of the carboxyl-terminal domain (CTD) of RNAP II. CDK9 then hyperphosphorylates the CTD of RNAP II, the components of N-TEF, and DSIF (Spt5) [69-71]; Tat activates P-TEFb by displacing Hexim1 from its cyclin T1 binding site [72]. Finally, the affinity of the Tat-cyclin T1-CDK9 complex for TAR is decreased through Tat acetylation at Lys-50 by histone acetyl transferases (HAT) in p300/CBP (CREB binding protein) and PCAF (p300-CREB binding protein associated factor) [66,73] leading to dissociation of Tat and P-TEFb from TAR and the enhancement of transcription elongation [66,74].

HIV-1 Tat is a 101-residue protein encoded by two exons (Figs. (**1b**) and (**3**)). Several HIV-1 laboratory strains yield Tat proteins with 86 residues that likely originate from the HXB2 reference strain (subtype B) [75]. The 86-residue variants are not found in natural viral isolates [76] and it has been suggested that they arose during tissue culture passaging of the virus; a single nucleotide correction in the laboratory genomes yields the expected 101 residue protein [77]. HIV-1 Tat exon-1 defines amino acids 1–72 and is shown in Fig. (**3**) where it is compared to the sequence of the equine infectious anemia virus (EIAV) Tat. Based on amino acid composition, several sequence domains have been identified: an acidic and proline-rich N-terminus (1-21), a cysteine-rich region (22–37), a core (38–47), an arginine-rich basic region (48–57), and a Gln-rich segment (58-72) [78]. The second *tat* exon defines residues 73–101 and is highly enriched in polar and charged residues as well as proline. The 72-residue segment of Tat can activate transcription with the same proficiency as the 86-residue protein [65,79-81]. Recently however, HIV-1 expressing the full-length 101-residue protein was shown to have enhanced NFκB activation, transactivation, and replication in T-cells compared to the exon-1 product [82] strongly suggesting that future work on the protein include the segment from exon-2.

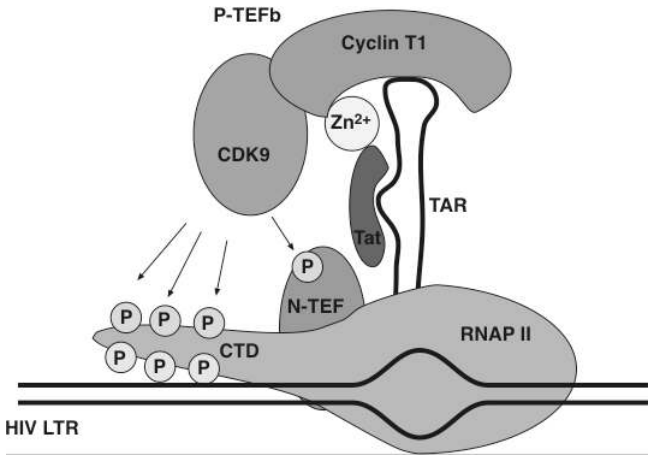



**Figure 2.** The transcription elongation regulatory complex. The regulatory complex formed by recognition of the TAR stem-loop bulge by Tat and the Tat-associated kinases of P-TEFb (Cyclin T1 and CDK9). Tat recognition primarily involves interactions between the Arg-rich region of its basic domain and the phosphates of the UCU bulge in the TAR element of the RNA. The Tat-cyclin T1 interaction may involve cysteine residues in both proteins through coordination with zinc ions. The complex allows the hyperphosphorylation of the CTD of RNAP II by CDK9 leading to transcription elongation.

A recent survey of interactions between human proteins and the proteins of human pathogens indicates that viruses and bacteria mainly manipulate host cell cycle, apoptosis, immune response, and nuclear transport [83]. HIV-1 Tat has been implicated in all but the last of these major cellular processes [84-86]. For example, Tat lengthens the G1 phase of the cell cycle and ten human proteins involved in G1 are known to interact with Tat [83]. Tat has also been implicated in mitochondrial apoptosis of T-cells [87,88] and inducing oxidative stress [89,90]. Tat may also be involved in derepression of heterochromatin, in transcription initiation [91], and in reverse-transcription [92]. More recently, Tat has been shown to regulate capping of HIV-1 mRNA [93], to interact with Dicer and suppress the production of small interfering RNA [94], and to act as a nucleic acid chaperone [95]. Absence of Tat and low levels of Tat-associated host factors cyclin dependent kinase 9 (CDK9) and cyclin T1 in resting CD4+ T-cells are all implicated in the latency of the virus, a major roadblock in antiretroviral therapy [96].

Tat protein also appears to be able to escape infected cells and has been implicated in an extraordinarily wide range of activities including supporting endothelial cell proliferation and thereby contributing to the development of Kaposi's Sarcoma, [97-99], decreasing expression of tight junction proteins [100], and promoting angiogenesis [101], to name a few. A major event in the progression of HIV-1 infection is neuronal damage despite the fact that neurons cannot be infected with the virus [102]. Tat can be released from infected microglia and astrocytes within the central nervous system (CNS) [103] and is able to cross the blood-brain barrier [104] resulting in cell death of neurons [89,105]. NeuroAIDS encompasses a range of disorders arising from damage to the peripheral and central nervous systems including HIV-associated dementia (HAD) and HIV-associated encephalitis (HIVE) [103,106]. Several HIV proteins have been implicated in neural dysfunction including gp120, gp41, Rev, Nef and Tat [105]. However, it is difficult to prove that enough protein is released by infected cells to account for the observed pathology. Furthermore, overactivation of the immune system, and in particular macrophages, may be the main contributor to neuropathy [107]. In general though, the pathological activities of Tat appear to contribute significantly to both immune and non-immune dysfunction resulting in an overall increase in the impact of viral infection.

Some of the biological activities of Tat have been assigned to a sequence domain of the protein. Residues 1-24 form the KIX domain binding site on p300/CBP [108]. Cyclin T1 is thought to interact with the Cys-rich region of Tat [109]. Mutation of any one of the cysteine residues in the protein, except for Cys-31, results in loss of transactivation [65]. Cys-31 is common in most major subtypes of HIV-1, but it is not present in the HIV-1 subtype C strain which accounts for more than 50% of infections worldwide [110]. The end of the Cys-rich region and the core are involved in mitochondrial apoptosis of bystander non-infected cells through their ability to bind tubulin and prevent its depolymerization [111]. The basic region is critical for RNA binding [112] as well as nuclear localization; peptide fragments corresponding to the region of Tat spanning Tyr-47 and Arg-57 have been used to transport a large variety of materials including proteins, DNA, drugs, imaging agents, liposomes, and nanoparticles across cell and nuclear membranes [113]. The Gln-rich region has been implicated in mitochondrial apoptosis of T-cells [87]. The minimum transactivation protein spans residues 1 – 48 [65].

The second exon-encoded polypeptide has far fewer biological activities ascribed to it presumably because it has been studied less intensively [81,92]. Studies have shown that the segment encoded by the second exon is involved in repressing the expression of major histocompatibility complex (MHC) class I molecules whose presence at the cell surface serve as targets for cytotoxic T lymphocytes [114-116]. This repressive function from the exon-2 encoded segment may contribute to HIV-infected cells escaping an immune response [114,115]. The second exon segment also includes an RGD motif that may mediate Tat binding to cell surface integrins [117]. A summary of published interactions between HIV-1 proteins and human proteins counted 830 unique, direct, physical interactions [86] and found that Tat is the most prolific interactor.

The Tat amino acid sequence (Fig. (**3**)) has a high net positive charge, a low overall hydrophobicity, and a low-sequence-complexity, meaning low variability of the 20 amino acids within a segment of the protein and repetition of amino acids. Analyses by several disorder prediction algorithms suggested that it is intrinsically disordered with a potential for order in the cysteine-rich region [118]. Early circular dichroism (CD) spectropolarimetry experiments showed that the protein lacks secondary structure but showed evidence of minor conformational changes upon addition of zinc [108]. The strong negative ellipticity at 198 nm shown in Fig. (**4**) is characteristic of a backbone dominated by the random coil conformation.

There have been several attempts to determine solution conformations of Tat and its segments, both alone and in complexes and here we briefly summarize several. Most of the studies have suffered from the limited resolution attainable in homonuclear $^{1}H$ NMR spectra of unlabelled proteins. At pH 4.5, in the absence of reducing agents a low-resolution, globular conformation with some flexible segments, particularly in the basic region, was deduced for $^{13}C^{\alpha}$-Gly-labelled synthetic $Tat_{1\text{-}86}$ (Bru) (Fig. (**5a**)) [119]. In an NMR model of the 87-residue Tat Mal protein at pH 4.5 under oxidizing conditions, the N-terminal Trp-11 formed a hydrophobic core through interactions with Phe-38 and Tyr-47 (Fig. (**5b**)) [120]. The basic region is in an extended conformation and the Cys-rich region contains β-turns; an α-

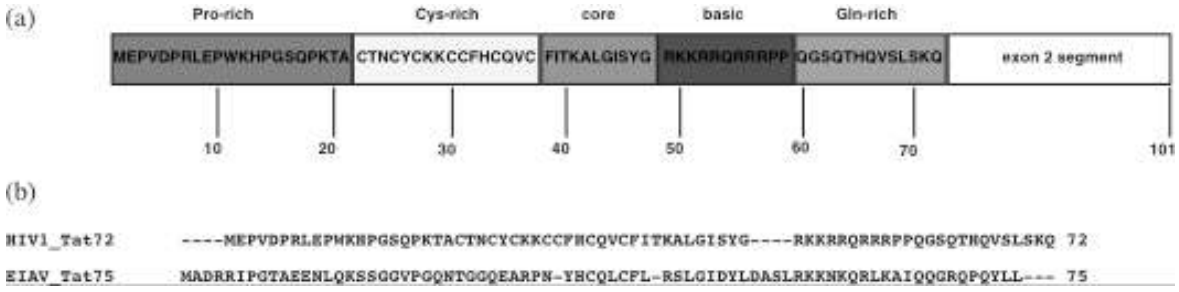



**Figure 3. (a)** The HIV-1 Tat sequence (BH10 isolate) encoded by exon 1. The 72-residue segment encompasses an N-terminal proline-rich region (1-21) containing the only three acidic residues, a cysteine-rich region (22–37), a core (38–47), a basic region (48–57), and a Gln-rich segment (58-72). Residues 73–101 are encoded by exon 2. **(b)** Sequence alignment of the first exon-encoded region of $Tat_{1-72}$ of Tat BH10 isolate and the 75-residue Tat protein from EIAV (Swiss-Prot accession number 20920) [156].

helix is found in the Gln-rich segment. $^{1}$H NMR spectroscopy and molecular dynamics simulations suggested that at pH 6.3 under reducing conditions a Z-variant $Tat_{1-86}$ forms condensed domains encompassing the core and Gln-rich regions, whereas the basic and Cys-rich regions were found to be highly flexible (Fig. (**5c**)) [121]. The lack of agreement among the conformations illustrated in Fig. (**5**) suggests that the protein does not form a stably folded conformation and supports the hypothesis that HIV-1 Tat is an IDP [60]. Since the structures were determined under different conditions it is worth considering that metal ions, oxidized thiols, pH or other solution conditions may have contributed to the heterogeneity of the structures and offers the possibility that binding partners may induce structure in all or part of the protein.

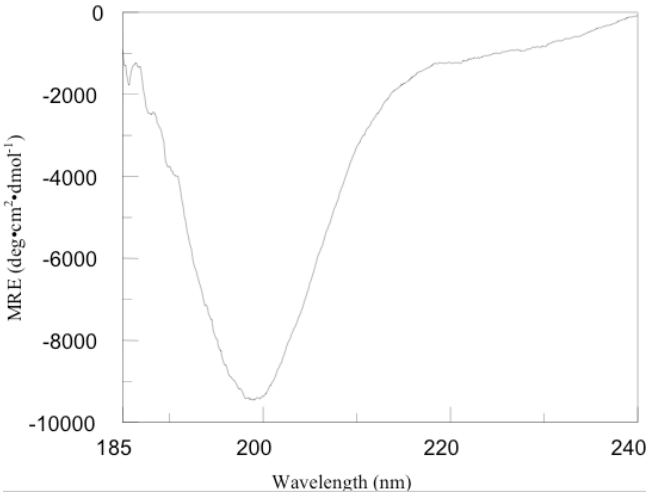



**Figure 4.** A baseline-subtracted far-ultra-violet CD spectrum of HIV-1 $Tat_{1-72}$ at room temperature. Tat protein concentration was 173 μM dissolved in 10 mM acetate buffer at pH 4.17.

There have been several studies of Tat fragments in complex with TAR RNA mainly focusing on the conformation of TAR [122-126]. NMR spectroscopy suggested a conformational change in $Tat_{32-72}$, in the region of Gly-42 and Gly-44, upon binding to TAR [125]. $^{1}$H NMR also showed that $Tat_{46-55}$, acetylated at Lys-50, is bound in an extended conformation to the bromodomain of p300/CBP-associated factor (PCAF) [66]. CD spectra suggested the possibility of a conformational change in $Tat_{1-86}$ upon binding to the KIX domain of p300/CBP [108]. $^{15}$N NMR relaxation measurements showed that $Tat_{47-58}$ becomes slightly more ordered on binding heparin [127], while CD studies of overlapping peptide fragments suggested that the most flexible regions of Tat are those adjacent to the basic region [128]. An oxidized Tendamistat-$Tat_{1-37}$ fusion protein showed multiple conformations with some evidence of helicity in the Cys-rich region (20-33) at pH 3.5 [129]. A fusion protein consisting of the activation domain from the Equine Infectious Anemia Virus and $Tat_{48-57}$ showed high helical content in the basic domain by NMR spectroscopy and CD [130]. One structural study has been published on a 101-residue Tat (Eli) protein at pH 4.5 and at pH 7 [131]. Circular dichroism spectra of the protein exhibit a strong band at 198 nm typical of disordered proteins (see Fig. (**4**)). Structural restraints deduced from $^{1}$H NMR spectra were used to deduce a globular conformation however the chemical shift dispersion of the amide protons suggests that the protein is disordered [131].

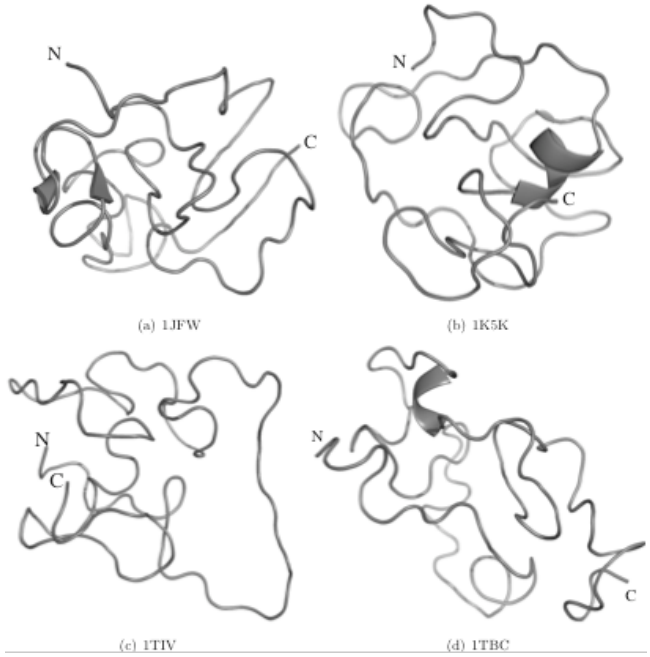



**Figure 5.** NMR derived conformations of HIV-1 Tat found in the RCSB Protein Data bank (PDB): **(a)** Tat Bru (group M subtype B), 86 residues (PDB ID: 1JFW) [119]; **(b)** Tat Mal (group M subtype A) [120], 87 residues (PDB ID:1K5K); **(c)** Tat Z2 (group M subtype D) 86 residues (PDB ID: 1TIV) [121]; and **(d)** Tat Z2 (group M subtype D) 86 residues (PDB ID: 1TBC) (Boehm, M., Sticht, H., Seidel, G., and Roesch, P. unpublished).

Based mainly on binding studies between fragments of Tat and site-specific mutants of TAR, the low resolution features of the interaction between Tat and TAR were delineated as indicated in Fig. (**2**). Tat binds directly to TAR through electrostatic interactions between its basic arginine-rich region and the negatively charged phosphates at a stem-loop UCU-bulge (uridine23-cytidine24-uridine25) of the RNA and the complex has a dissociation constant of $K_d = 12$ nM [132]. The two base pairs ($G_{26}$:$C_{39}$ and $A_{27}$:$U_{38}$) immediately above the TAR bulge are also believed to be critical for Tat recognition and the two base pairs ($A_{22}$:$U_{40}$ and $G_{21}$:$C_{41}$) below the bulge also contribute to the binding affinity [133]. Phosphates at positions 22, 23 and 40 on the RNA are also critical for Tat binding interactions [123]. The basic Arg-rich and Gln-rich regions of Tat govern the binding affinity of Tat to TAR RNA, but it is the core region that seems to control the specificity of Tat for the TAR element [134,135].

## 4. MULTINUCLEAR NMR SPECTROSCOPY AND X-RAY DIFFRACTION STUDIES OF TAT

With its identifiable sequence domains, low-amino acid-sequence-complexity, low overall hydrophobicity, and high net positive charge, Tat has all of the hallmarks of intrinsic disorder making it a good candidate for investigation by NMR spectroscopy. A major stumbling block has been the difficulty in preparing homogenous solutions of the protein

at concentrations amenable to NMR owing to the readily oxidized Cys-rich region of the protein that gives rise to a heterogeneous mixture of multimers. The high net positive charge is also problematic as it causes the protein to bind tightly to anionic surfaces including glass, and to polyanionic species in cell lysates such as DNA and RNA. Preparation of reduced, monomeric exon-1 Tat protein labelled with $^{15}N$ and $^{13}C$ at pH 4.1 was achieved by extraction of *E. coli* with the strong reducing agent tris(carboxyethyl)phosphine and the denaturant guanidine hydrochloride, purification using immobilized metal affinity chromatography, and subsequent removal of the denaturants and reducing agent by dialysis [60]. Experiments for assigning the backbone chemical shifts along with measurements of $^{3}J_{H^N H^\alpha}$ coupling constants, $^{15}N$ relaxation rates, and steady-state nuclear Overhauser effects (NOE) between $^{15}N$ and $^{1}H$ nuclei were done at pH 4.1 in order to maintain the protein in a reduced monomeric state. Tat transactivates transcription in the interior of a cell that is a reduced environment where most proteins do not contain intramolecular disulfide bonds. In the rare cases where intracellular disulfides have been observed they are usually involved in redox cycling, are transiently formed, and not required to maintain the structure of the protein [136]. In agreement with this, evidence suggests that the reduced Tat monomer is the form of the protein that is active in transactivation [137].

Multinuclear labelling of Tat was essential to obtaining correct chemical shift assignments of backbone resonances and for high-resolution dynamics experiments [60]. The assignments of the Cys residues are particularly informative as they confirmed that all of the Cys residues in the NMR sample were reduced (Fig. (**6**)); all of the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts observed in a 3-dimensional HNCACB [138] spectrum, are in the range of the random coil chemical shifts of reduced cysteine (58.6 ppm and 28.3 ppm) [139] differing significantly from those of oxidized cysteine (55.6 ppm and 41.2 ppm) involved in disulfide bonds. The dispersion of the resonances in $^{15}N$-$^{1}H$ HSQC spectra of $Tat_{1-72}$ (Fig. (**7**)) is typical for disordered proteins lacking regular secondary structure and agrees with previous homonuclear NMR studies [119,120,129,140,141]. The resonance line-widths of 15±5 Hz for $^{1}H$ and 6±1 Hz for $^{15}N$ are broad compared to those measured for folded proteins of a comparable size (6-9 Hz and 3 Hz respectively) [142]. The broad linewidths suggest backbone conformational exchange on the µs-ms timescale. Furthermore, multiple minor resonances are detected for many residues suggesting the presence of multiple, likely unfolded, conformers in equilibrium [60].

The use of the chemical shift as an indicator of secondary structure [143] has been successfully applied to the analysis of proteins under denaturing conditions to identify regions of residual structure [144-146] but required adaptation to account for local sequence effects on the reference shifts for the random coil [139,147,148]. In some cases, non-native residual structure has been identified that may indicate transient intermediates along the protein folding path-

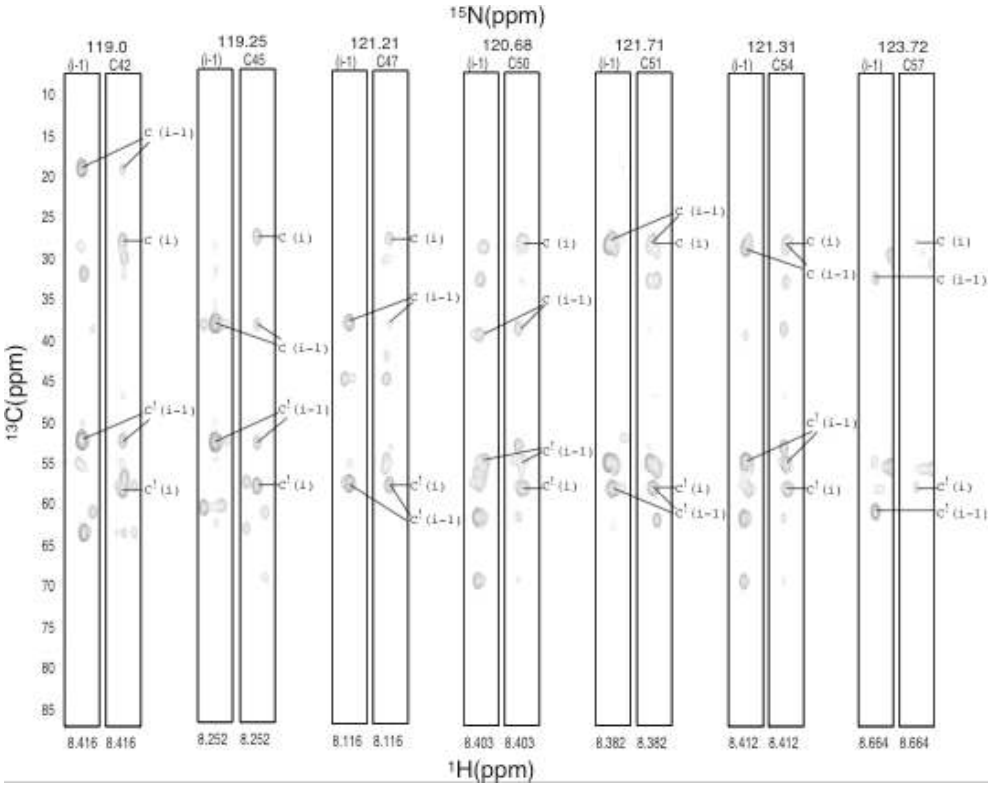



**Figure 6.** Strip plots extracted from three-dimensional, amide-detected heteronuclear NMR experiments for the Cys-rich region of $^{13}C/^{15}N$-labeled His-tagged $Tat_{1-72}$. Inter- and intra-residual correlations are obtained from an HNCACB [138] spectrum correlating $H^N(i)$ and N(i) with $C^{\alpha}(i)$, $C^{\alpha}(i-1)$, $C^{\beta}(i)$, and $C^{\beta}(i-1)$ resonances and a CBCA(CO)NH [160] spectrum correlating $H^N(i)$ and N(i) with $C^{\alpha}(i-1)$ and $C^{\beta}(i-1)$ resonances. Indicated are the chemical shifts of the $C^{\alpha}$ and $C^{\beta}$ resonances for the 7 Cys residues, C54, C45, C47, C50, C51, C42, and C57 from left to right, and the residues preceding the cysteines.

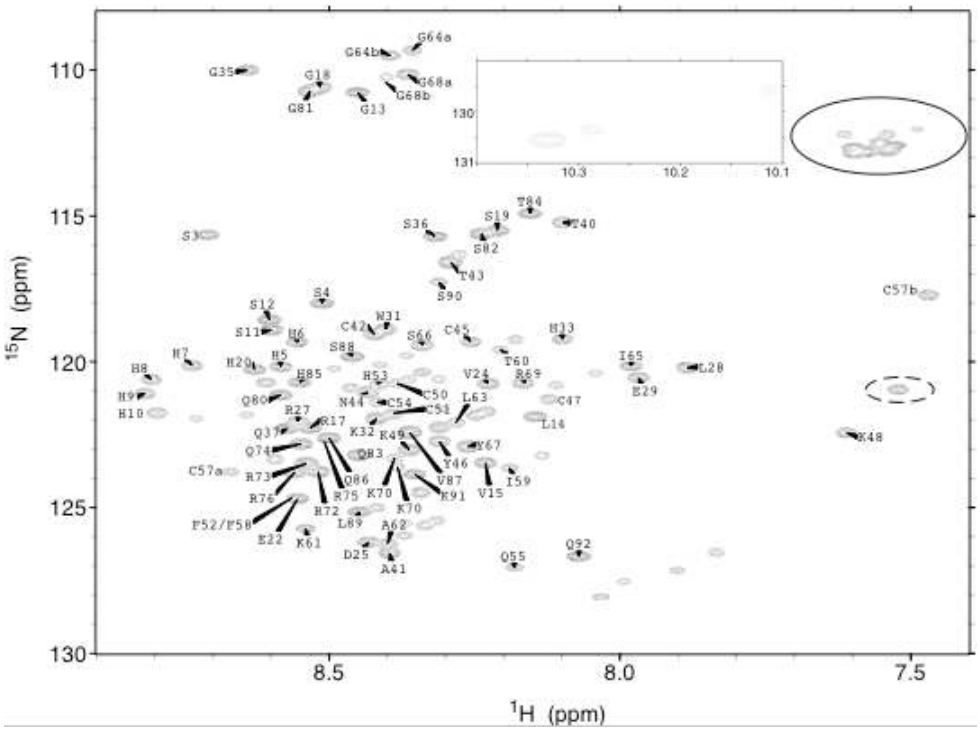



**Figure 7.** Two-dimensional $^{1}$H/$^{15}$N-HSQC spectrum [161] of Tat at pH 4.1 and 293 K. Backbone amide proton resonances are narrowly dispersed over the range of 7.4-8.9 ppm. Backbone resonance assignments of 80 of the 84 non-proline resonances are shown. The inset spectrum shows the three peaks associated with the side chain of the single Trp residue indicating 1 major and 2 minor conformers. Also labelled are several glycine conformers. Side-chain Asn and Gln $NH_2$ resonances are outlined with a solid ellipse and side-chain Arg resonances are outlined in a dashed ellipse.

way [144,145]. In the case of the His-tagged $Tat_{1-72}$ the majority of the backbone chemical shift differences from sequence-corrected random coil shifts were found to be within the bounds of the random coil conformation [60]. Subsequently, Marsh *et al*. described a method for combining the secondary chemical shifts from many nuclei into a single probe score to obtain a secondary structural propensity (SSP) especially suited to detecting transient structure and structural propensity in IDPs [149]. In this procedure, consecutive residues with SSP scores of 1 or -1 indicate a fully formed α- or β-structural element, respectively. Lower scores correspond to the fraction of the ensemble that is in a particular structural state. We have re-analyzed the chemical shifts of Tat [60] and present the SSP scores in Fig. (**8**). They show that most residues lie between the ±20% boundary lines meaning that, among the ensemble of conformations, fewer than 20% of the residues are in any particular regular secondary structure. Particularly noteworthy, are the SSP scores for the region centered on residue 52 that indicate a propensity towards β structure followed by a propensity to α-helix for a region centered on residue 66. The two regions correspond to the Cys-rich region and a region that encompasses the core and basic regions shown in Fig. (**3**).

Despite the relatively poor chemical shift dispersion in the NMR spectra of Tat, relaxation data were obtained for 77% and 72% of the observable resonances (non-proline and non-N-terminal) at 600 MHz and 800 MHz spectrometer frequencies respectively. The steady-state heteronuclear $^{15}$N-$^{1}$H NOEs exhibited a flattened bell-shaped variation with amino acid sequence [60], diagnostic of a disordered protein [150]. The lack of any significant variation in the $^{15}$N-$^{1}$H NOEs, coupled with their negative values, are good indicators of the degree of uniform disorder throughout the protein backbone on the ns-ps timescale. Longitudinal relaxation rates ($R_1$) observed for Tat [60] were consistent with slower relaxation and faster dynamics at the ns-ps timescale than for a folded protein of similar size and both the $^{15}$N-$^{1}$H NOEs and $R_1$ values were very similar to those measured for a guanidine-denatured protein [151]. The rotating frame longitudinal relaxation rates ($R_{1\rho}$) were slightly higher than average at the end of the Cys-rich region (Cys-57 to Ile-59) possibly indicating contributions from slow conformational exchange.

Analysis of the relaxation rates by reduced spectral density mapping [151-153] showed little variation over the length of the sequence and show a range of values similar to those observed for the guanidine-denatured state of drkN SH3 [151], acid-denatured apomyoglobin [145], low pH/urea denatured apomyoglobin [152], and the intrinsically disordered pro-peptide of subtilisin [154]. Interestingly, the acid-urea-unfolded state of apomyoglobin shows minima in the high-frequency plots that appear to be diagnostic of residual structure because they correspond to maxima in buried surface area in the folded protein [145]. The lack of definition in the plots of high-frequency motion for reduced $Tat_{1-72}$ implies a lack of formation of any transient structure at pH 4.1. However, residues Asn-44, Lys-61 and Ala-62 were found to have faster than average motions in the mid-frequency plots. Schwarzinger *et al.* [148] made a similar observation in their studies of acid-urea denatured apomyoglobin and noted that high proportions of glycine and alanine

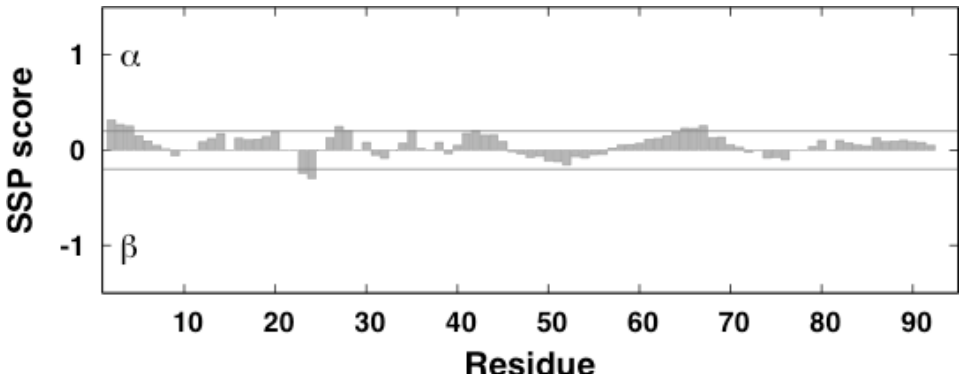



**Figure 8.** Secondary Structure Propensity (SSP) scores [149] based on the $H^N$, N, $C^\alpha$, $C^\beta$, C' and $H^\alpha$ chemical shifts. Upper and lower lines correspond to 20% of the conformers in the ensemble being in either the α- or β-states. Note that the 20-residue affinity tag has been added to the sequence in the figure whereas it is absent from Fig. (**3**).

are present in the most flexible regions of the protein. The region of Tat following the Cys-rich region contains one alanine and two glycines (Ala-62, Gly-64 and Gly-68) which may explain the higher dynamics there. Calculation of Lipari-Szabo order parameters from the relaxation data gave an average value of 0.5 where a value of 1 represents order and 0 represents disorder and where most unfolded proteins show values in the range of 0.4 – 0.6 [151,154].

The cross-peaks in the HSQC spectra of Tat diminish in intensity as the pH is raised from 3.3 to 6.7, whereas the chemical shifts of the resonances move very little. For the most part the loss of resonance intensity could be explained by the pH-dependent increase in hydrogen exchange rates and suggests that the protein does not fold at the elevated pH because buried resonances would be expected to exchange more slowly and remain visible in the spectrum. However, it cannot be ruled out that the protein forms a molten globule conformation at elevated pH in which conformational exchange causes broadening of the resonances. And although the dynamics experiments and chemical shift analysis suggest that the protein is in a random coil conformation at pH 4 it is also conceivable that measurement of $^1$H-NOEs might detect minor condensed conformations in the ensemble.

Although HIV-1 Tat has never been crystallized, a 69-residue EIAV Tat fused to the cyclin box domain of human cyclin T1 formed crystals with electron density observed for residues 8 – 263 of the cyclin T1 [155]. EIAV Tat contains only 2 Cys residues in contrast to the 6 (Type C) or 7 (Type B) present in the HIV-1 proteins, considerably simplifying expression and purification of the protein. No electron density was observed for EIAV Tat indicating that cyclin T1 does not induce a stable conformation in Tat [155]. Crystallization of an EIAV cyclin T1-Tat fusion protein in the presence of TAR RNA resulted the detection of electron density for residues 41 – 69 of Tat [156] (Fig. (**9**)). Residues 41 – 47 in the core region (see Fig. (**3**)) exist in an extended conformation interacting mostly with the cyclin T1. Residues 48 – 59 in the basic segment (see Fig. (**3**)) form a helix that binds to the major groove of the TAR stem-loop structure (Fig. (**9**)). This region is followed by a tight turn and an extended segment that inserts the C-terminal Leu-68 and Leu-69 into a hydrophobic groove on cyclin T1. It is interesting to note that the strongest indication of a propensity to fold based on the chemical shift analysis in Fig. (**8**) is found in the basic region of the Tat protein [81]. The structure of the tripartite complex explains the specificity of Tat proteins for T-type cyclins, the binding specificities of cyclin T1s for Tat-TAR complexes from different species, the structural implication of the acetylation of Lys-50, and likely reveals the binding site for HEXIM-1 on cyclin T1. The cyclin T1-Tat interface encompasses 806 Å$^2$, the Tat-TAR interface covers 582 Å$^2$ and the cyclin T1-TAR complex buries only 193 Å$^2$. Clearly the role of Tat is to bring the two molecules together. Although this structure does not directly reveal the dynamics of Tat and TAR it clearly illustrates the large surface area available to an unfolded protein permitting high affinity and high specificity of binding. The structure also shows the power of molecular structures to illuminate biological activity even in the realm of IDPs and suggests that the crystallization of fusion proteins will be a rich source of information on IDPs.

## CONCLUSIONS

Recent multinuclear NMR [60] and X-ray diffraction studies [155,156] show that Tat proteins are intrinsically disordered but undergo a partial disorder-to-order transition upon formation of the Tat-TAR-Cyclin T1 complex. Intrinsic disorder is likely a key to the protein's capacity to interact with multiple partners and effect multiple biological functions; the large buried surface in the X-ray diffraction structure illustrates how a disordered protein can have a high affinity and a high specificity for its partners and how disordered Tat assembles a protein complex to enhance transcription elongation. An interesting aspect of the binding of Tat is that its target, TAR RNA, is also highly conformationally flexible as revealed by NMR spectroscopy [54] suggesting a significant entropy penalty in the formation of a Tat-TAR complex. Do Tat and TAR passively select a suitable conformation from among the possibilities offered by their partner [18] or do they induce the desired conformation in their partner [19]? Another option is that Tat is presented to TAR in a relatively fixed conformation in the cyclin T1 complex but the current X-ray diffraction structure of the bimolecular complex shows that Tat is disordered [155,156]. Resolution of this problem will require detailed kinetic and thermodynamic analysis of the interactions.

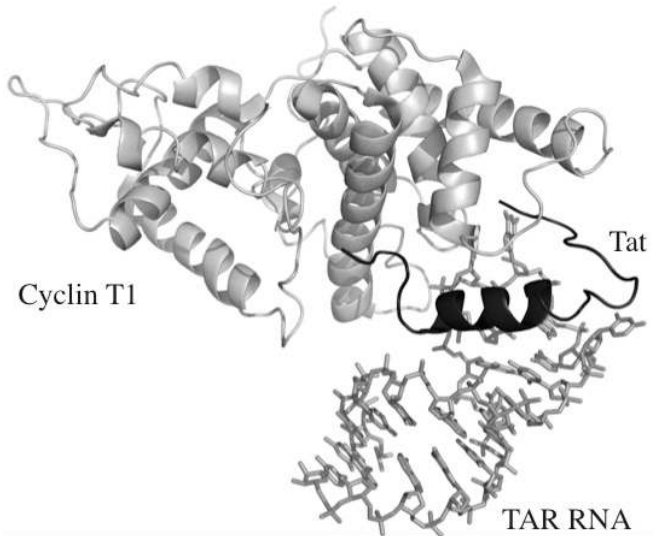



**Figure 9.** X-ray diffraction structure of the Cyclin T1-Tat-TAR complex from EIAV [156] showing the C-terminal region of EIAV Tat (residues 41-69) bound to the cyclin box repeat of cyclin T1 and the major groove and loop region of TAR. The remainder of the EIAV Tat protein is missing from the electron density map

There is substantial evidence that Tat interacts with cyclin T1 by shared coordination of zinc [136,157,158]. Multinuclear NMR data suggested possible transient structure formation in the Cys-rich region of the protein but the X-ray diffraction structure of EIAV Tat does not define the structure of this region of the protein. Furthermore, EIAV Tat contains only two Cys residues and may not be stabilized by zinc binding in a manner similar to the HIV-1 protein. Delineating the role of metal ions in formation of an active Tat transcriptional activation complex requires the development of methods to study the HIV-1 version of Tat. Also important will be multinuclear NMR studies of the structure and dynamics of Tat at physiological pH to determine if lowering the net charge of the protein permits more transient structure formation or induces the formation of a molten globule-like conformation. Future studies of Tat by NMR should also include the full 101-residue protein. Although the 72-residue protein encoded by the first *tat* exon is fully competent in transcription activation, the remaining 29 residues must be of significance to the virus since it is conserved in all natural HIV-1 isolates. With one exception [129], the full-length protein has been ignored in structural studies. In addition to providing epitopes for developing Tat vaccines across many HIV-1subtypes, the presence of residues 73-101 may have an important effect on some of the non-transcriptional activities of Tat that also need to be addressed.

## ABBREVIATIONS

AIDS = Acquired immunodeficiency syndrome

CA = Capsid

CBP = CREB binding protein

CCR5 = CC chemokine receptor 5

CD = Circular dichroism

CD4 = Cluster of differentiation 4

CDK9 = Cyclin dependent kinase 9

CNS = Central nervous system

CTD = Carboxy-terminal domain

drkN SH3 = N-terminal SH3 domain from the adapter protein drk

DSIF, 5 = 6-Dichloro-1-beta-D-ribofuranosylbenzimidazole sensitivity-inducing factor

EIAV = Equine infectious anemia virus

Env = Envelope protein

Gag/gag = Group-specific antigen

gp = Glycoprotein

HAD = HIV-associated dementia

HAT = Histone acetyl transferase

Hexim1 = Hexamethylene bisacetamide-inducible protein 1

HIV-1 = Human immunodeficiency virus type 1

HIVE = HIV-associated encephalitis

HSQC = Heteronuclear single-quantum coherence

IDP = Intrinsically disordered proteins

IN = Integrase

LTR = Long terminal repeat

MA = Matrix

MHC = Major histocompatibility complex

N-TEF = Negative transcription elongation factor

NC = Nucleocapsid

Nef = Negative factor

NFκB = Nuclear factor κ-light-chain-enhancer of activated B cells

NOE = Nuclear Overhauser effect

P-TEFb = Positive transcription elongation factor b

PCAF = p300/CBP-associated factor

Pol/pol = Polyprotein

PR = Protease

PDB = Protein Data Bank

R1 = Longitudinal relaxation rate

R1ρ = Rotating frame longitudinal relaxation rate

R2 = Transverse relaxation rate

Rev/rev = Regular expression of virus

RGD = Arginine-glycine-aspartic acid

RNAP II = RNA Polymerase II

RT = Reverse transcriptase

Spt5 = Suppressor of tyrosine 5

SSP = Secondary structure propensity

SU = Surface unit

TAR = Transactivation response

Tat/tat = Transactivator of transcription

TM = Transmembrane

Vif/vif = Viral infectivity factor

Vpr/vpr = Viral protein r

Vpu/vpu = Viral protein u

## ACKNOWLEDGEMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the University of Manitoba; the Canadian Foundation for Innovation provided funding for the NMR facility at the University of Manitoba. We thank Thach Vo for the CD spectrum of Tat.

## REFERENCES

[1] Emini, E.A. The human immunodeficiency virus: biology, immunology, and therapy. Princeton University Press: Princeton, N.J., Chichester, 2002.

[2] Turner, B.G.; Summers, M.F. Structural Biology of HIV. *J. Mol. Biol*., **1999**, *285*(1), 1-32.

[3] Cullen, B. Regulation of HIV-1 gene expression. *FASEB J*., **1991**, *5*(10), 2361-2368.

[4] Kingsman, S.M.; Kingsman, A.J. The Regulation of Human Immunodeficiency Virus Type-1 Gene Expression. *Eur. J. Biochem*., **1996**, *240*(3), 491-507.

[5] Haseltine, W. Molecular biology of the human immunodeficiency virus type 1. *FASEB J*., **1991**, *5*(10), 2349-2360.

[6] Frankel, A.D.; Young, J.A.T. HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem*., **1998**, *67*(1), 1-25.

[7] Watts, J.M.; Dang, K.K.; Gorelick, R.J.; Leonard, C.W.; Bess Jr, J.W.; Swanstrom, R.; Burch, C.L.; Weeks, K.M. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **2009**, *460*(7256), 711-716.

[8] Freed, E.O. HIV-1 Replication. *Somat. Cell Mol. Genet*., **2001**, *26*(1), 13-33.

[9] Dunker, A.K.; Obradovic, Z. The protein trinity—linking function and disorder. *Nat. Biotech.,* **2001**, *19*(9), 805-806.

[10] Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W. Intrinsically disordered protein. *J. Mol. Graphics Modell*., **2001**, *19*(1), 26-59.

[11] Patriarca, A.; Eliseo, T.; Sinibaldi, F.; Piro, M.C.; Melis, R.; Paci, M.; Cicero, D.O.; Polticelli, F.; Santucci, R.; Fiorucci, L. ATP Acts as a Regulatory Effector in Modulating Structural Transitions of Cytochrome c: Implications for Apoptotic Activity. *Biochemistry,* **2009**, *48* (15), 3279-3287.

[12] Uversky, V.N.; Ptitsyn, O.B. "Partly Folded" State, a New Equilibrium State of Protein Molecules: Four-State Guanidinium Chloride-Induced Unfolding of beta-Lactamase at Low Temperature. *Biochemistry,* **1994**, *33*(10), 2782-2791.

[13] Uversky, V.N.; Ptitsyn, O.B. Further Evidence on the Equilibrium "Pre-molten Globule State": Four-state Guanidinium Chloride-induced Unfolding of Carbonic Anhydrase B at Low Temperature. *J. Mol. Biol.,* **1996**, *255*(1), 215-228.

[14] Ohgushi, M.; Wada, A. 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett*., **1983**, *164*(1), 21-24.

[15] Dobson, C.M. Principles of protein folding, misfolding and aggregation; Protein Misfolding and Human Disease and Developmental Biology of the Retina. *Semin. Cell Dev. Biol*., **2004**, *15*(1), 3-16.

[16] Maji, S.K.; Perrin, M.H.; Sawaya, M.R.; Jessberger, S.; Vadodaria, K.; Rissman, R.A.; Singru, P.S.; Nilsson, K.P.R.; Simon, R.; Schubert, D.; Eisenberg, D.; Rivier, J.; Sawchenko, P.; Vale, W.; Riek, R. Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science*, **2009**, *325*(5938), 328-332.

[17] Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit*., **2005**, *18*(5), 343-384.

[18] Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics,* **2005**, *21*(16), 3433-3434.

[19] Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The Pairwise Energy Content Estimated from Amino Acid Composition Discriminates between Folded and Intrinsically Unstructured Proteins. *J. Mol. Biol.,* **2005**, *347*(4), 827-839.

[20] Li, X.; Romero, P.; Rani, M.; Dunker, A.K.; Obradovic, Z. Predicting protein disorder for N-, C-, and internal regions. *Genome Inform.,* **1999**, *10*, 30-40.

[21] Romero, P.; Obradovic, Z.; Li, X.; Garner, E.; Brown, C.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins*, **2001**, *42*, 38-48.

[22] Romero, P.; Obradovic, Z.; Dunker, A.K. Sequence data analysis for long disordered regions prediction in the calcineurin family. *Genome Inform*., **1997**, *8*, 110-124.

[23] Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K. Predicting intrinsic disorder from amino acid sequence. *Proteins*, **2003**, *53(S6)*, 566-72.

[24] Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K.; Z., O. Abstract Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol*., **2005**, *3*(1), 35-60.

[25] Yang, Z.R.; Thomson, R.; McMeil, P.; Esnouf, R.M. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics*, **2005**, *21*(16), 3369-3376.

[26] Linding, R.; Jensen, L.J.; Diella, F.; Bork, P.; Gibson1, T.J.; Russell, R.B. Protein Disorder Prediction: Implications for Structural Proteomics. *Structure*, **2003**, *11*(11), 1453-1459.

[27] Jones, D.T.; Ward, J.J. Prediction of disordered regions in proteins from position specific score matrices. *Proteins,* **2003**, *53* (S6), 573-578.

[28] Ward, J.J.; McGuffin, L.J.; Bryson, K.; Buxton, B.F.; Jones, D.T. The DISOPRED server for the prediction of protein disorder. *Bioinformatics,* **2004**, *20*(13), 2138-2139.

[29] Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol*., **2004**, *337* (3), 635-645.

[30] Linding, R.; Russell, R.B.; Neduva, V.; Gibson, T.J. GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res*., **2003**, *31*(13), 3701-3708.

[31] Dunker, A.K.; Garner, E.; Guilliot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J.E. Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac. Symp. Biocomput*., **1998**, *3*, 471-482.

[32] Chen, J.W.; Romero, P.; Uversky, V.N.; Dunker, A.K. Conservation of Intrinsic Disorder in Protein Domains and Families: I. A Database of Conserved Predicted Disordered Regions. *J. Proteome Res*., **2006**, *5*(4), 879-887.

[33] Jeffery, C.J. Moonlighting proteins—an update. *Mol. BioSyst*., **2009**, *5*(4), 345-350.

[34] Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci*., **2002**, *27*(10), 527-533.

[35] Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.,* **2005**, *579*(15), 3346-3354.

[36] Tompa, P.; Csermely, P. The role of structural disorder in the function of RNA and protein chaperones. *FASEB J*., **2004**, *18*(11), 1169-1175.

[37] Uversky, V.N. Natively unfolded proteins: a point where biology waits for physics. *Prot. Sci*., **2002**, *11*(4), 739-756.

[38] Hoh, J.H. Functional protein domains from the thermally driven motion of polypeptide chains: A proposal. *Proteins*, **1998**, *32*(2), 223-228.

[39] Rout, M.P.; Aitchison, J.D.; Magnasco, M.O.; Chait, B.T. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol*., **2003**, *13*(12), 622-628.

[40] Denning, D.P.; Patel, S.S.; Uversky, V.; Fink, A.L.; Rexach, M. Disorder in the nuclear pore complex: The FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*(5), 2450-2455.

[41] Brown, H.G.; Hoh, J.H. Entropic Exclusion by Neurofilament Sidearms: A Mechanism for Maintaining Interfilament Spacing. *Biochemistry,* **1997**, *36*(49), 15035-15040.

[42] Mukhopadhyay, R.; Kumar, S.; Hoh, J.H. Molecular mechanisms for organizing the neuronal cytoskeleton. *BioEssays*, **2004**, *26*(9), 1017-1025.

[43] Dunker, A.K.; Oldfield, C.; Meng, J.; Romero, P.; Yang, J.; Chen, J.; Vacic, V.; Obradovic, Z.; Uversky, V. The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics*, **2008**, *9* (Suppl 2), S1.

[44] Tsvetkov, P.; Reuven, N.; Shaul, Y. The nanny model for IDPs. *Nat. Chem. Biol*., **2009**, *5*(11), 778-781.

[45] Schrader, E.K.; Harstad, K.G.; Matouschek, A. Targeting proteins for degradation. *Nat. Chem. Biol.,* **2009**, *5*(11), 815-822.

[46] Radivojac, P.; Vacic, V.; Haynes, C.; Cocklin, R.R.; Mohan, A.; Heyen, J.W.; Goebl, M.G.; Iakoucheva, L.M. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins*, **2010**, *78*(2), 365-380.

[47] Bussell, R., Jr.; Eliezer, D. Residual Structure and Dynamics in Parkinson's Disease-associated Mutants of alpha-Synuclein. *J. Biol. Chem.,* **2001**, *276*(49), 45996-46003.

[48] Hegde, M.L.; Rao, K.S.J. DNA induces folding in alpha-synuclein: Understanding the mechanism using chaperone property of osmolytes. *Arch. Biochem. Biophys*., **2007**, *464*(1), 57-69.

[49] Radivojac, P.; Vucetic, S.; O'Connor, T.R.; Uversky, V.N.; Obradovic, Z.; Dunker, A.K. Calmodulin signaling: Analysis and prediction of a disorder-dependent molecular recognition. *Proteins*, **2006**, *63* (2), 398-410.

[50] Demchenko, A.P. Recognition between flexible protein molecules: induced and assisted folding. *J. Mol. Recognit.,* **2001**, *14*(1), 42-61.

[51] Huang, Y.; Liu, Z. Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding-Binding Process: A Critical Assessment of the "Fly-Casting" Mechanism. *J. Mol. Biol*., **2009**, *393*(5), 1143-1159.

[52] Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature*, **2007**, *450*(7172), 964-972.

[53] Baldwin, A.J.; Kay, L.E. NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol*., **2009**, *5*(11), 808-814.

[54] Zhang, Q.; Stelzer, A.C.; Fisher, C.K.; Al-Hashimi, H.M. Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature*, **2007**, *450*(7173), 1263-1267.

[55] Wüthrich, K. NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture). *Angew. Chem. Int. Ed. Engl*., **2003**, *42*(29), 3340-3363.

[56] Palmer, A.G.; Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chem. Rev*., **2006**, *106*(5), 1700-1719.

[57] Redfield, C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods*, **2004**, *34*(1), 121-132.

[58] Chatterjee, A.; Kumar, A.; Chugh, J.; Srivastava, S.; Bhavesh, N.S.; Hosur, R.V. NMR of unfolded proteins. *J. Chem. Sci.,* **2005**, *117*(1), 3-21.

[59] Dyson, H.J.; Wright, P.E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev*., **2004**, *104*(8), 3607-3622.

[60] Shojania, S.; O'Neil, J.D. HIV-1 Tat is a natively unfolded protein: the solution conformation and dynamics of reduced HIV-1 Tat-(1-72) by NMR spectroscopy. *J. Biol. Chem*., **2006**, *281*(13), 8347-8356.

[61] Shojania, S.; Henry, G.D.; Chen, V.C.; Vo, T.N.; Perreault, H.; O'Neil, J.D. High yield expression and purification of HIV-1 Tat1-72 for structural studies. *J. Virol. Meth*., **2009**, *(in press)*.

[62] Coffin, J.M.; Hughes, S.H.; Varmus, H.E., *Retroviruses*. Cold Spring Harbor Laboratory Press: Woodbury, NY, 1997.

[63] Liang, C.; Wainberg, M.A. The role of Tat in HIV-1 replication: an activator and/or a suppressor? *AIDS Rev*., **2002**, *4*(1), 41-49.

[64] Ping, Y.-H.; Rana, T.M. DSIF and NELF Interact with RNA Polymerase II Elongation Complex and HIV-1 Tat Stimulates P-TEFb-mediated Phosphorylation of RNA Polymerase II and DSIF during Transcription Elongation. *J. Biol. Chem*., **2001**, *276*(16), 12951-12958.

[65] Jeang, K.; Xiao, H.; Rich, E. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *J. Biol. Chem.,* **1999**, *274*, 28837 - 28840.

[66] Mujtaba, S.; He, Y.; Zeng, L.; Farooq, A.; Carlson, J.E.; Ott, M.; Verdin, E.; Zhou, M.M. Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. *Mol. Cell*, **2002**, *9*(3), 575-586.

[67] West, M.J.; Karn, J. Stimulation of Tat-associated kinase-independent transcriptional elongation from the human immunodeficiency virus type-1 long terminal repeat by a cellular enhancer. *EMBO J*., **1999**, *18*(5), 1378-1386.

[68] Berkhout, B.; Silverman, R.H.; Jeang, K.T. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell*, **1989**, *59*(2), 273-82.

[69] Yamaguchi, Y.; Takagi, T.; Wada, T.; Yano, K.; Furuya, A.; Sugimoto, S.; Hasegawa, J.; Handa, H. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell,* **1999**, *97*(1), 41-51.

[70] Bourgeois, C.F.; Kim, Y.K.; Churcher, M.J.; West, M.J.; Karn, J. Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Mol. Cell. Biol.,* **2002**, *22*(4), 1079-1093.

[71] Kim, Y.K.; Bourgeois, C.F.; Isel, C.; Churcher, M.J.; Karn, J. Phosphorylation of the RNA Polymerase II Carboxyl-Terminal Domain by CDK9 Is Directly Responsible for Human Immunodeficiency Virus Type 1 Tat-Activated Transcriptional Elongation. *Mol. Cell. Biol*., **2002**, *22*(13), 4622-4637.

[72] Schulte, A.; Czudnochowski, N.; Barboric, M.; Schonichen, A.; Blazek, D.; Peterlin, B.M.; Geyer, M. Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. *J. Biol. Chem*., **2005**, *280*(26), 24968-24977.

[73] Bannwarth, S.; Gatignol, A. HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr. HIV Res*., **2005**, *3*(1), 61-71.

[74] Wong, K.; Sharma, A.; Awasthi, S.; Matlock, E.F.; Rogers, L.; Van Lint, C.; Skiest, D.J.; Burns, D.K.; Harrod, R. HIV-1 Tat Interactions with p300 and PCAF Transcriptional Coactivators Inhibit Histone Acetylation and Neurotrophin Signaling through CREB. *J. Biol. Chem*., **2005**, *280*(10), 9390-9399.

[75] Opi, S.; Péloponèse, J.-M.; Esquieu, D.; Watkins, J.; Campbell, G.; De Mareuil, J.; Jeang, K.T.; Yirrell, D.L.; Kaleebu, P.; Loret, E.P. Full-length HIV-1 Tat protein necessary for a vaccine. *Vaccine*, **2004**, *22*(23-24), 3105-3111.

[76] Jeang, K.-T., HIV-1 Tat: Structure and Function. In *Human Retroviruses and AIDS 1996: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*, Myers, G.;Korber, B.T.;Foley, B.T.;Jeang, K.-T.;Mellors, J.W.; Wain-Hobson, S., Eds. Los Alamos National Laboratories: Los Alamos, 1996; pp 11-26.

[77] Neuveut, C.; Jeang, K.T. Recombinant human immunodeficiency virus type 1 genomes with tat unconstrained by overlapping reading frames reveal residues in Tat important for replication in tissue culture. *J. Virol*., **1996**, *70*(8), 5572-5581.

[78] Derse, D.; Carvalho, M.; Carroll, R.; Peterlin, B.M. A minimal lentivirus Tat. *J. Virol*., **1991**, *65*(12), 7012-7015.

[79] Kuppuswamy, M.; Subramania, T.; Srinivasan, A.; Chinnadurai, G. Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nucleic Acids Res*., **1989**, *17*, 3551 - 3561.

[80] Garcia, J.A.; Harrich, D.; Pearson, L.; Mitsuyasu, R.; Gaynor, R.B. Functional Domains Required for Tat-Induced Transcriptional Activation of the Hiv-1 Long Terminal Repeat. *EMBO J*., **1988**, *7*(10), 3143-3147.

[81] Smith, S.M.; Pentlicky, S.; Klase, Z.; Singh, M.; Neuveut, C.; Lu, C.Y.; Reitz, M.S.; Yarchoan, R.; Marx, P.A.; Jeang, K.T. An *in vivo* replication-important function in the second coding exon of Tat is constrained against mutation despite cytotoxic T lymphocyte selection. *J. Biol Chem*., **2003**, *278*(45), 44816-44825.

[82] Mahlknecht, U.; Dichamp, I.; Varin, A.; Van Lint, C.; Herbein, G. NF-kappaB-dependent control of HIV-1 transcription by the second coding exon of Tat in T cells. *J. Leuk. Biol*., **2008**, *83*(3), 718-727.

[83] Dyer, M.D.; Murali, T.M.; Sobral, B.W. The Landscape of Human Proteins Interacting with Viruses and Other Pathogens. *PLoS Pathogens*, **2008**, *4*(2), e32.

[84] Gallo, R.C. Tat as one key to HIV-induced immune pathogenesis and Tat (correction of Pat) toxoid as an important component of a vaccine. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*(15), 8324-8326.

[85] Alimonti, J.B.; Ball, T.B.; Fowke, K.R. Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS. *J. Gen. Virol.*, **2003**, *84*(Pt 7), 1649-1661.

[86] Ptak, R.G.; Fu, W.; Sanders-Beer, B.E.; Dickerson, J.E.; Pinney, J.W.; Robertson, D.L.; Rozanov, M.N.; Katz, K.S.; Maglott, D.R.; Pruitt, K.D.; Dieffenbach, C.W. Cataloguing the HIV Type 1 Human Protein Interaction Network. *AIDS Res. Hum. Retroviruses*, **2008**, *24*(12), 1497-1502.

[87] Campbell, G.R.; Pasquier, E.; Watkins, J.; Bourgarel-Rey, V.; Peyrot, V.; Esquieu, D.; Barbier, P.; de Mareuil, J.; Braguer, D.; Kaleebu, P.; Yirrell, D.L.; Loret, E.P. The glutamine-rich region of the HIV-1 Tat protein is involved in T-cell apoptosis. *J. Biol. Chem.,* **2004**, *279*(46), 48197-48204.

[88] Goldstein, G. HIV-1 Tat protein as a potential AIDS vaccine. *Nat. Med*., **1996**, *2*(9), 960-964.

[89] Pocernich, C.B.; Sultana, R.; Mohmmad-Abdul, H.; Nath, A.; Butterfield, D.A. HIV-dementia, Tat-induced oxidative stress, and antioxidant therapeutic considerations. *Brain Res. Rev*., **2005**, *50*(1), 14-26.

[90] Westendorp, M.O.; Shatrov, V.A.; Schulze-Osthoff, K.; Frank, R.; Kraft, M.; Los, M.; Krammer, P.H.; Droge, W.; Lehmann, V. HIV-1 Tat potentiates TNF-induced NF-kappaB activation and cytotoxicity by altering the cellular redox state. *EMBO J*., **1995**, *14*(3), 546-554.

[91] Pumfery, A.; Deng, L.; Maddukuri, A.; de la Fuente, C.; Li, H.; Wade, J.D.; Lambert, P.; Kumar, A.; Kashanchi, F. Chromatin remodeling and modification during HIV-1 Tat-activated transcription. *Curr. HIV Res*., **2003**, *1*(3), 343-362.

[92] Guo, X.; Kameoka, M.; Wei, X.; Roques, B.; Gotte, M.; Liang, C.; Wainberg, M.A. Suppression of an intrinsic strand transfer activity of HIV-1 Tat protein by its second-exon sequences. *Virology*, **2003**, *307*(1), 154-163.

[93] Chiu, Y.-L.; Ho, C.K.; Saha, N.; Schwer, B.; Shuman, S.; Rana, T.M. Tat Stimulates Cotranscriptional Capping of HIV mRNA. *Mol. Cell*, **2002**, *10*(3), 585-597.

[94] Bennasser, Y.; Jeang, K. HIV-1 Tat interaction with Dicer: requirement for RNA. *Retrovirology*, **2006**, *3*, 95 - 101.

[95] Kuciak, M.; Gabus, C.; Ivanyi-Nagy, R.; Semrad, K.; Storchak, R.; Chaloin, O.; Muller, S.; Mely, Y.; Darlix, J.-L. The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities *in vitro*. *Nucleic Acids Res*., **2008**, *36*(10), 3389-3400.

[96] Lassen, K.; Han, Y.; Zhou, Y.; Siliciano, J.; Siliciano, R.F. The multifactorial nature of HIV-1 latency. *Trends Mol. Med*., **2004**, *10*(11), 525-531.

[97] Ensoli, B.; Barillari, G.; Salahuddin, S.Z.; Gallo, R.C.; Wong-Staal, F. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature*, **1990**, *345*(6270), 84-6.

[98] Albini, A.; Benelli, R.; Presta, M.; Rusnati, M.; Ziche, M.; Rubartelli, A.; Paglialunga, G.; Bussolino, F.; Noonan, D. HIV-tat protein is a heparin-binding angiogenic growth factor. *Oncogene*, **1996**, *12*(2), 289-297.

[99] Albini, A.; Soldi, R.; Giunciuglio, D.; Giraudo, E.; Benelli, R.; Primo, L.; Noonan, D.; Salio, M.; Camussi, G.; Rockl, W.; Bussolino, F. The angiogenesis induced by HIV-1 tat protein is mediated by the Flk-1/KDR receptor on vascular endothelial cells. *Nat. Med.,* **1996**, *2*(12), 1371-1375.

[100] András, I.E.; Pu, H.; Deli, M.A.; Nath, A.; Hennig, B.; Toborek, M. HIV-1 Tat protein alters tight junction protein expression and distribution in cultured brain endothelial cells. *J. Neurosci. Res*., **2003**, *74*(2), 255-265.

[101] Wu, R.F.; Gu, Y.; Xu, Y.C.; Mitola, S.; Bussolino, F.; Terada, L.S. Human Immunodeficiency Virus Type 1 Tat Regulates Endothelial Cell Actin Cytoskeletal Dynamics through PAK1 Activation and Oxidant Production. *J. Virol*., **2004**, *78*(2), 779-789.

[102] Kaul, M.; Garden, G.A.; Lipton, S.A. Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature*, **2001**, *410*(6831), 988-994.

[103] King, J.E.; Eugenin, E.A.; Buckner, C.M.; Berman, J.W. HIV tat and neurotoxicity. *Microbes Infect.,* **2006**, *8*(5), 1347-1357.

[104] Banks, W.A.; Robinson, S.M.; Nath, A. Permeability of the blood-brain barrier to HIV-1 Tat. *Exp. Neurol*., **2005**, *193*(1), 218-227.

[105] Nath, A.; Psooy, K.; Martin, C.; Knudsen, B.; Magnuson, D.S.; Haughey, N.; Geiger, J.D. Identification of a human immunodeficiency virus type 1 Tat epitope that is neuroexcitatory and neurotoxic. *J. Virol.,* **1996**, *70*(3), 1475-1480.

[106] Toborek, M.; Lee, Y.W.; Flora, G.; Pu, H.; András, I.E.; Wylegala, E.; Hennig, B.; Nath, A. Mechanisms of the Blood-Brain Barrier Disruption in HIV-1 Infection. *Cell. Mol. Neurobiol*., **2005**, *25*(1), 181-199.

[107] Pardo, C.A.; McArthur, J.C.; Griffin, J.W. HIV neuropathy: Insights in the pathology of HIV peripheral nerve disease. *J. Peripher. Nerv. Syst*., **2001**, *6*(1), 21-27.

[108] Vendel, A.C.; Lumb, K.J. Molecular recognition of the human coactivator CBP by the HIV-1 transcriptional activator Tat. *Biochemistry,* **2003**, *42*(4), 910-916.

[109] Bieniasz, P.D.; Grdina, T.A.; Bogerd, H.P.; Cullen, B.R. Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *EMBO J.,* **1998**, *17*(23), 7056-7065.

[110] Campbell, G.R.; Watkins, J.D.; Singh, K.K.; Loret, E.P.; Spector, S.A. Human Immunodeficiency Virus Type 1 Subtype C Tat Fails To Induce Intracellular Calcium Flux and Induces Reduced Tumor Necrosis Factor Production from Monocytes. *J. Virol.,* **2007**, *81*(11), 5919-5928.

[111] Chen, D.; Wang, M.; Zhou, S.; Zhou, Q. HIV-1 Tat targets microtubules to induce apoptosis, a process promoted by the pro-apoptotic Bcl-2 relative Bim. *EMBO J*., **2002**, *21*(24), 6801-6810.

[112] Weeks, K.M.; Ampe, C.; Schultz, S.C.; Steitz, T.A.; Crothers, D.M. Fragments of the HIV-1 Tat protein specifically bind TAR RNA. *Science*, **1990**, *249*(4974), 1281-1285.

[113] Gupta, B.; Levchenko, T.S.; Torchilin, V.P. Intracellular delivery of large molecules and small particles by cell-penetrating proteins and peptides. *Adv. Drug Deliv. Rev.,* **2005**, *57*(4), 637-651.

[114] Weissman, J.D.; Brown, J.A.; Howcroft, T.K.; Hwang, J.; Chawla, A.; Roche, P.A.; Schiltz, L.; Nakatani, Y.; Singer, D.S. HIV-1 Tat binds TAFII250 and represses TAFII250-dependent transcription of major histocompatibility class I genes. *Proc. Natl. Acad. Sci. USA,* **1998**, *95*(20), 11601-11606.

[115] Carroll, I.R.; Wang, J.; Howcroft, T.K.; Singer, D.S. HIV Tat represses transcription of the beta2-microglobulin promoter. *Mol. Immunol*., **1998**, *35*(18), 1171-1178.

[116] Howcroft, T.; Strebel, K.; Martin, M.; Singer, D. Repression of MHC class I gene promoter activity by two-exon Tat of HIV. *Science*, **1993**, *260*(5112), 1320-1322.

[117] Avraham, H.K.; Jiang, S.; Lee, T.-H.; Prakash, O.; Avraham, S. HIV-1 Tat-Mediated Effects on Focal Adhesion Assembly and Permeability in Brain Microvascular Endothelial Cells. *J. Immunol*., **2004**, *173*(10), 6228-6233.

[118] Goh, G.; Dunker, A.K.; Uversky, V. Protein intrinsic disorder toolbox for comparative analysis of viral proteins. *BMC Genomics*, **2008**, *9*(Suppl 2), S4.

[119] Peloponese, J.-M.; Gregoire, C.; Opi, S.; Esquieu, D.; Sturgis, J.; Lebrun, E.; Meurs, E.; Collette, Y.; Olive, D.; Aubertin, A.-M.; Witvrow, M.; Pannecouque, C.; De Clercq, E.; Bailly, C.; Lebreton, J.; Loret, E.P. 1H-13C nuclear magnetic resonance assignment and structural characterization of HIV-1 Tat protein. *C.R. Accad. Sci., Ser. III*, **2000**, *323*(10), 883-894.

[120] Gregoire, C.; Peloponese, J.; Esquieu, D.; Opi, S.; Campbell, G.; Solomiac, M.; Lebrun, E.; Lebreton, J.; Loret, E. Homonuclear 1H-NMR assignment and structural characterization of human immunodeficiency virus type 1 Tat Mal protein. *Biopolymers*, **2001**, *62*, 324 - 335.

[121] Bayer, P.; Kraft, M.; Ejchart, A.; Westendorp, M.; Frank, R.; Rosch, P. Structural studies of HIV-1 Tat protein. *J. Mol. Biol*., **1995**, *247*(4), 529-535.

[122] Long, K.S.; Crothers, D.M. Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA. *Biochemistry*, **1999**, *38*(31), 10059-10069.

[123] Aboul-ela, F.; Karn, J.; Varani, G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol*., **1995**, *253*(2), 313-332.

[124] Seewald, M.J.; Metzger, A.U.; Willbold, D.; Rosch, P.; Sticht, H. Structural model of the HIV-1 Tat(46-58)-TAR complex. *J. Biomol. Struct. Dyn*., **1998**, *16*(3), 683-692.

[125] Metzger, A.U.; Bayer, P.; Willbold, D.; Hoffmann, S.; Frank, R.W.; Goody, R.S.; Rosch, P. The interaction of HIV-1 Tat(32-72)

with its target RNA: a fluorescence and nuclear magnetic resonance study. *Biochem. Biophys. Res. Commun.*, **1997**, *241* (1), 31-36.

[126] Greenbaum, N.L. How Tat targets TAR: structure of the BIV peptide-RNA complex. *Structure*, **1996**, *4*(1), 5-9.

[127] Hakansson, S.; Caffrey, M. Structural and dynamic properties of the HIV-1 tat transduction domain in the free and heparin-bound states. *Biochemistry*, **2003**, *42*(30), 8999-9006.

[128] Gregoire, C.J.; Loret, E.P. Conformational heterogeneity in two regions of TAT results in structural variations of this protein as a function of HIV-1 isolates. *J. Biol. Chem.*, **1996**, *271*(37), 22641-22646.

[129] Freund, J.; Vertesy, L.; Koller, K.P.; Wolber, V.; Heintz, D.; Kalbitzer, H.R. Complete 1H nuclear magnetic resonance assignments and structural characterization of a fusion protein of the alpha-amylase inhibitor tendamistat with the activation domain of the human immunodeficiency virus type 1 Tat protein. *J. Mol. Biol.*, **1995**, *250*(5), 672-688.

[130] Puglisi, J.D.; Tan, R.; Calnan, B.J.; Frankel, A.D.; Williamson, J.R. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, **1992**, *257*(5066), 76-80.

[131] Watkins, J.; Campbell, G.; Halimi, H.; Loret, E. Homonuclear 1H NMR and circular dichroism study of the HIV-1 Tat Eli variant. *Retrovirology*, **2008**, *5*(1), 83.

[132] Dingwall, C.; Ernberg, I.; Gait, M.J.; Green, S.M.; Heaphy, S.; Karn, J.; Lowe, A.D.; Singh, M.; Skinner, M.A. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J.*, **1990**, *9*(12), 4145-4153.

[133] Pritchard, C.E.; Grasby, J.A.; Hamy, F.; Zacharek, A.M.; Singh, M.; Karn, J.; Gait, M.J. Methylphosphonate mapping of phosphate contacts critical for RNA recognition by the human immunodeficiency virus tat and rev proteins. *Nucleic Acids Res.*, **1994**, *22*(13), 2592-2600.

[134] Churcher, M.J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S.M.; Lowe, A.D.; Butler, J.G.; Gait, M.J.; Karn, J. High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.,* **1993**, *230*(1), 90-110.

[135] Rana, T.M.; Jeang, K.T. Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.*, **1999**, *365*(2), 175-85.

[136] Mallick, P.; Boutz, D.R.; Eisenberg, D.; Yeates, T.O. Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*(15), 9679-9684.

[137] Siddappa, N.; Venkatramanan, M.; Venkatesh, P.; Janki, M.; Jayasuryan, N.; Desai, A.; Ravi, V.; Ranga, U. Transactivation and signaling functions of Tat are not correlated: biological and immunological characterization of HIV-1 subtype-C Tat protein. *Retrovirology*, **2006**, *3*(1), 53.

[138] Wittekind, M.; Mueller, L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J. Magn. Reson. Ser B*, **1993**, *101*(2), 201-205.

[139] Schwarzinger, S.; Kroon, G.J.; Foss, T.R.; Wright, P.E.; Dyson, H.J. Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J. Biomol. NMR,* **2000**, *18*(1), 43-48.

[140] Peti, W.; Smith, L.J.; Redfield, C.; Schwalbe, H. Chemical shifts in denatured proteins: resonance assignments for denatured ubiquitin and comparisons with other denatured proteins. *J. Biomol. NMR*, **2001**, *19*(2), 153-165.

[141] Dyson, H.J.; Wright, P.E. Equilibrium NMR studies of unfolded and partially folded proteins. *Nat. Struct. Biol.,* **1998**, *5*, 499-503.

[142] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skelton, N.J. *Protein NMR Spectroscopy: Principles and Practice*. Academic Press: San Diego, 1996.

[143] Wishart, D.S.; Sykes, B.D. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR*, **1994**, *4*(2), 171-180.

[144] Teilum, K.; Kragelund, B.B.; Poulsen, F.M. Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. *J. Mol. Biol.*, **2002**, *324*(2), 349-357.

[145] Yao, J.; Chung, J.; Eliezer, D.; Wright, P.E.; Dyson, H.J. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry*, **2001**, *40*(12), 3561-71.

[146] Garcia, P.; Serrano, L.; Durand, D.; Rico, M.; Bruix, M. NMR and SAXS characterization of the denatured state of the chemotactic protein Che Y: Implications for protein folding initiation. *Prot. Sci.*, **2001**, *10*(6), 1100-1112.

[147] Schwarzinger, S.; Kroon, G.J.; Foss, T.R.; Chung, J.; Wright, P.E.; Dyson, H.J. Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.*, **2001**, *123*(13), 2970-2978.

[148] Schwarzinger, S.; Wright, P.E.; Dyson, H.J. Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry*, **2002**, *41*(42), 12681-12686.

[149] Marsh, J.A.; Singh, V.K.; Jia, Z.; Forman-Kay, J.D. Sensitivity of secondary structural propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Prot. Sci.*, **2006**, *15*, 2795-2804.

[150] Hu, Y.; Macinnis, J.M.; Cherayil, B.J.; Fleming, G.R.; Freed, K.F.; Perico, A. Polypeptide Dynamics - Experimental Tests of an Optimized Rouse-Zimm Type Model. *J. Chem. Phys.*, **1990**, *93*(1), 822-836.

[151] Farrow, N.A.; Zhang, O.; Forman-Kay, J.D.; Kay, L.E. Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. *Biochemistry*, **1997**, *36*(9), 2390-2402.

[152] Peng, J.W.; Wagner, G. Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry,* **1992**, *31*(36), 8571-8586.

[153] Lefevre, J.F.; Dayie, K.T.; Peng, J.W.; Wagner, G. Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. *Biochemistry*, **1996**, *35*(8), 2674-2686.

[154] Buevich, A.V.; Shinde, U.P.; Inouye, M.; Baum, J. Backbone dynamics of the natively unfolded pro-peptide of subtilisin by heteronuclear NMR relaxation studies. *J. Biomol. NMR*, **2001**, *20*(3), 233-249.

[155] Anand, K.; Schulte, A.; Fujinaga, K.; Scheffzek, K.; Geyer, M. Cyclin Box Structure of the P-TEFb Subunit Cyclin T1 Derived from a Fusion Complex with EIAV Tat. *J. Mol. Biol.,* **2007**, *370*(5), 826-836.

[156] Anand, K.; Schulte, A.; Vogel-Bachmayr, K.; Scheffzek, K.; Geyer, M. Structural insights into the Cyclin T1-Tat-TAR RNA transcription activation complex from EIAV. *Nat. Struct. Mol. Biol.*, **2008**, *15*(12), 1287-1292.

[157] Tosi, G.; Meazza, R.; Barbaro, A.D.L.; D'Agostino, A.; Mazza, S.; Corradin, G.; Albini, A.; Noonan, D.M.; Ferrini, S.; Accolla, R.S. Highly stable oligomerization forms of HIV-1 Tat detected by monoclonal antibodies and requirement of monomeric forms for the transactivating function on the HIV-1 LTR. *Eur. J. Immunol.*, **2000**, *30*(4), 1120-1126.

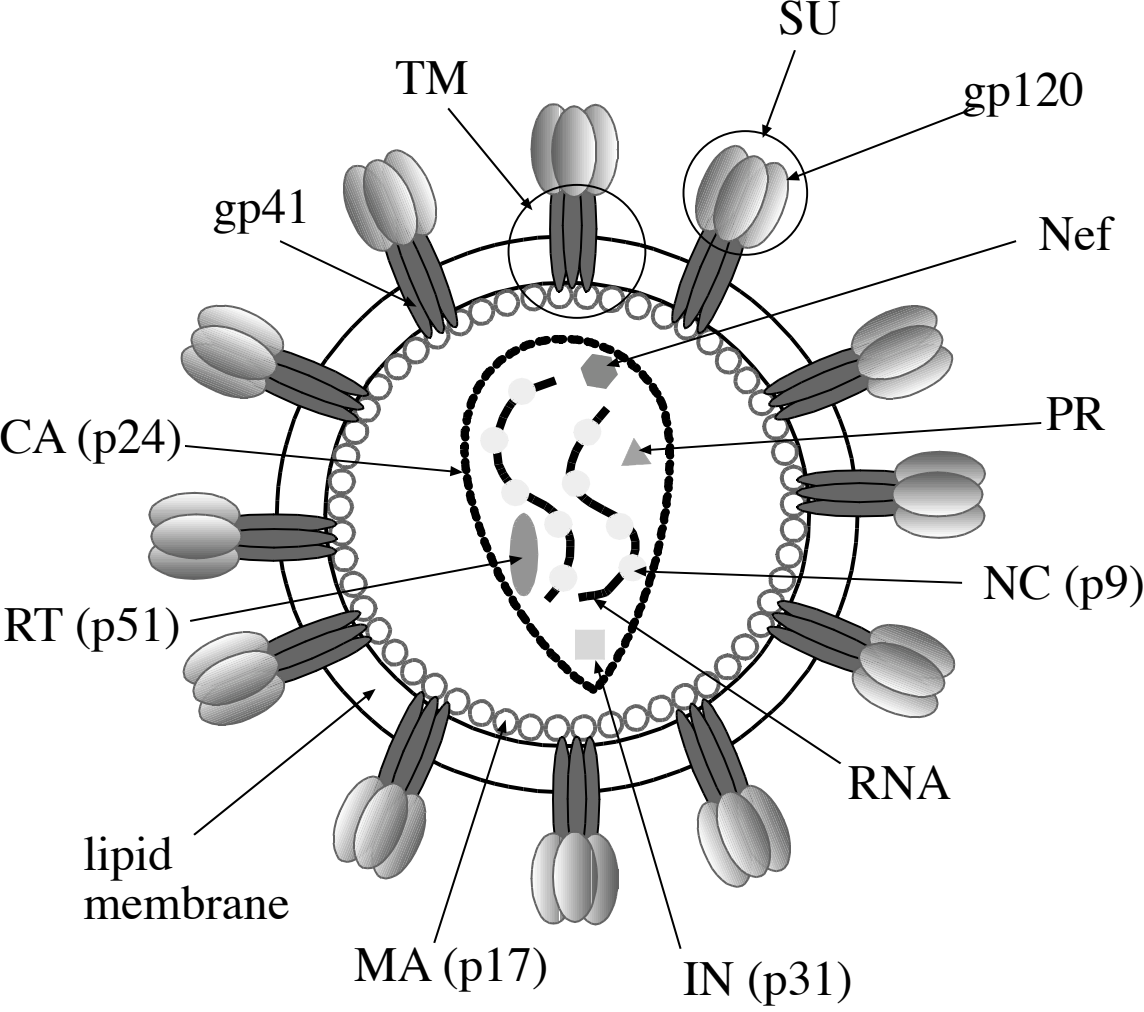
[158] Garber, M.E.; Wei, P.; KewalRamani, V.N.; Mayall, T.P.; Herrmann, C.H.; Rice, A.P.; Littman, D.R.; Jones, K.A. The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes and Dev.*, **1998**, *12*(22), 3512-3527.

[159] Karn, J. Tackling tat. *J. Mol. Biol.*, **1999**, *293*(2), 235-254.

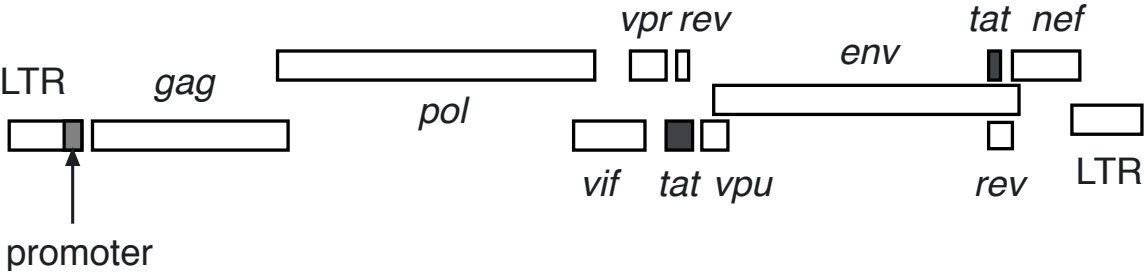
[160] Grzesiek, S.; Bax, A. Correlating Backbone Amide and Side-Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *J. Am. Chem. Soc.*, **1992**, *114*(16), 6291-6293.

[161] Kay, L.E.; Keifer, P.; Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.*, **1992**, *114*(26), 10663-10665.

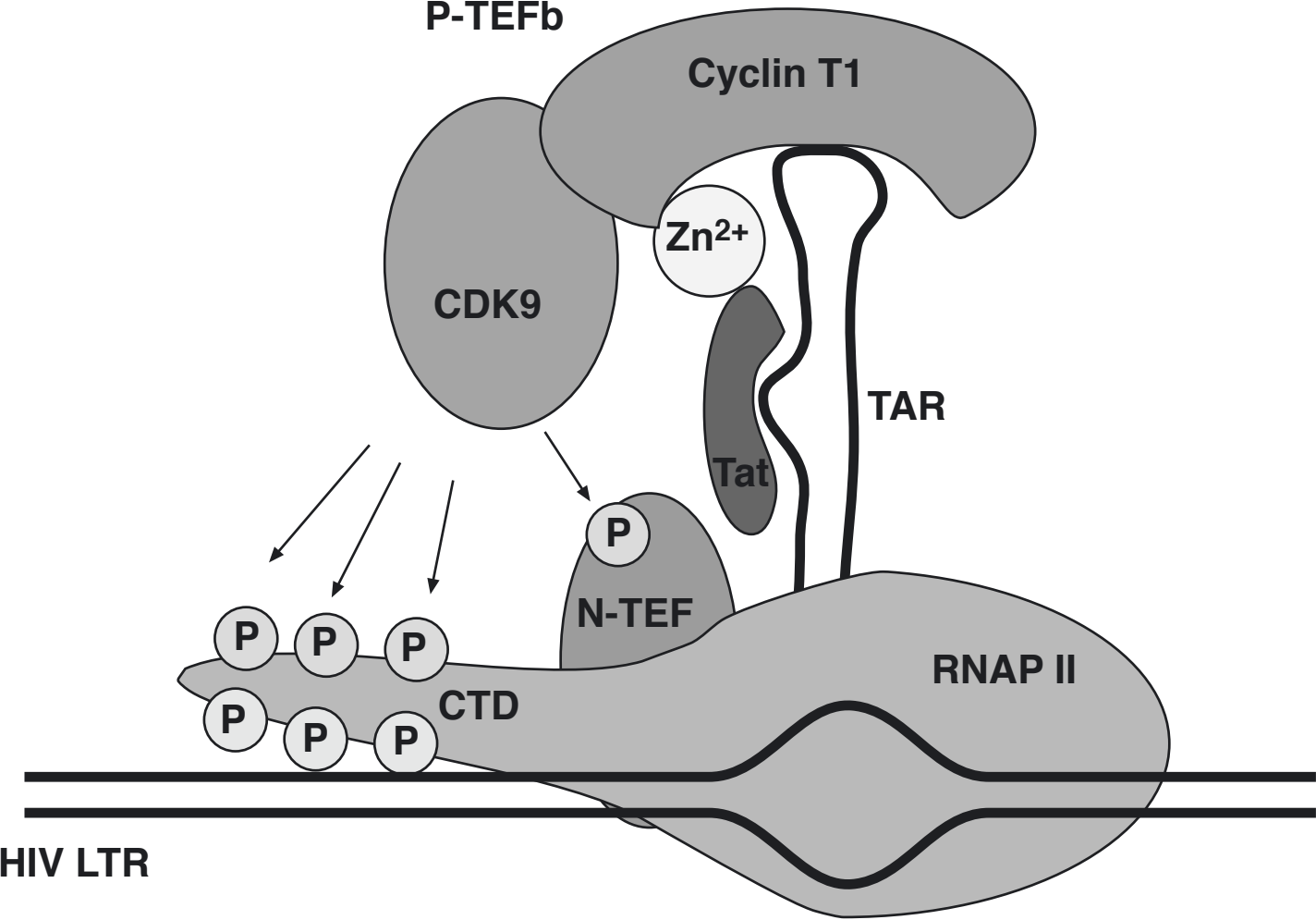
**Received: December 01, 2009 Revised: December 24, 2009 Accepted: December 24, 2009**


SU
TM
gp120
gp41
Nef
PR
CA (p24)
NC (p9)
RT (p51)
RNA
lipid
membrane
MA (p17)
IN (p31)

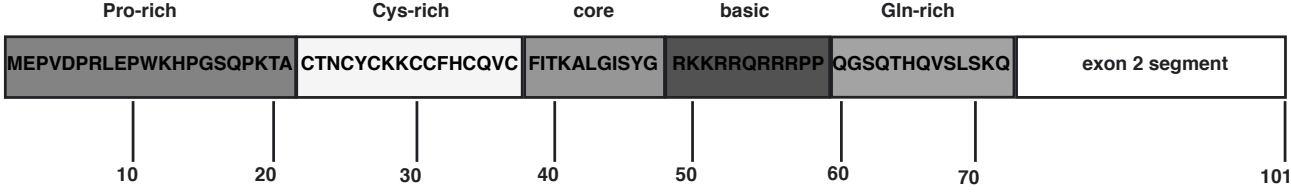
(a)


vpr rev
tat nef
env
LTR
gag
pol
LTR
vif
tat vpu
rev
promoter

(b)

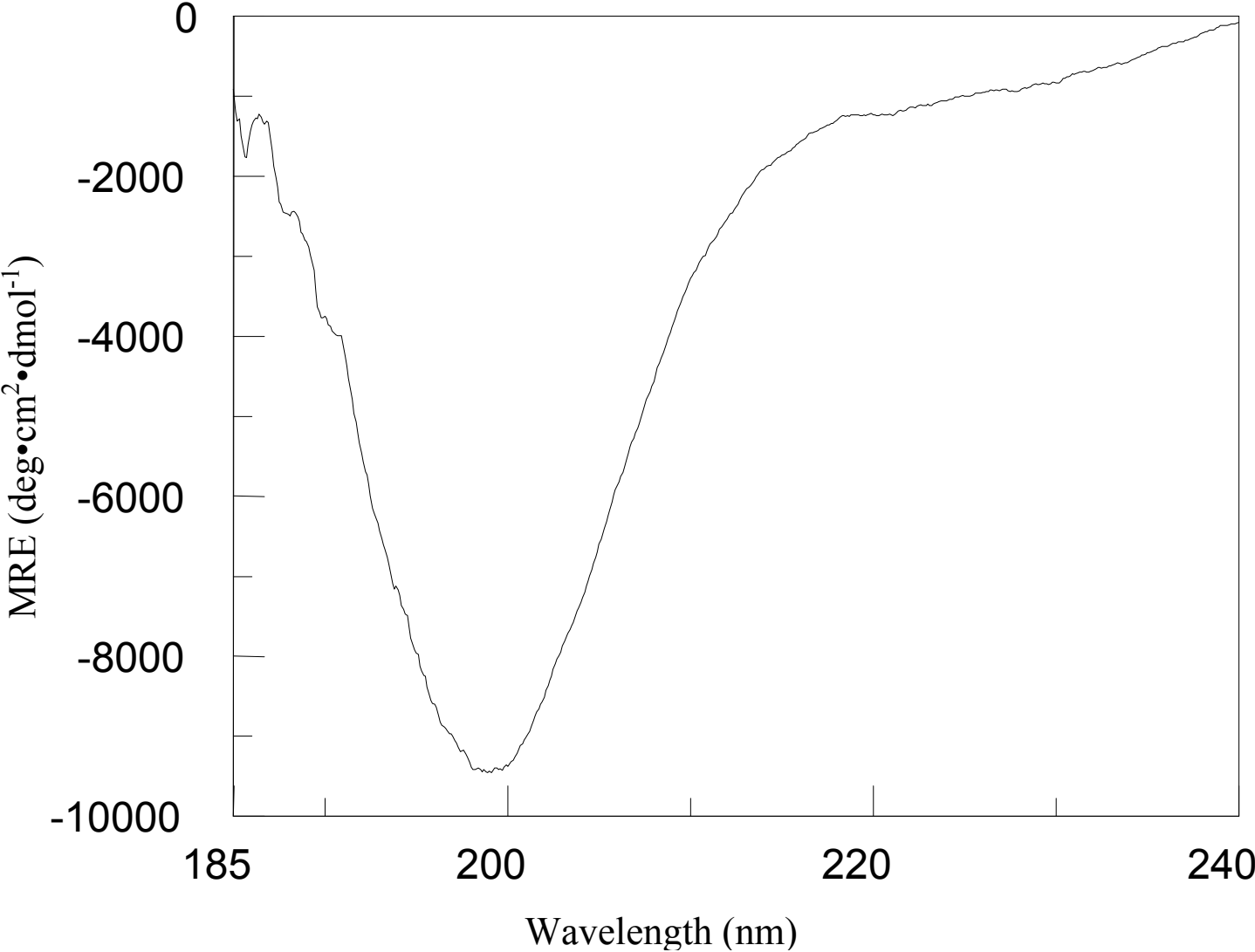

P-TEFb
Cyclin T1
Zn2+
CDK9
TAR
Tat
P
N-TEF
P
P
P
CTD
P
P
P
RNAP II
HIV LTR

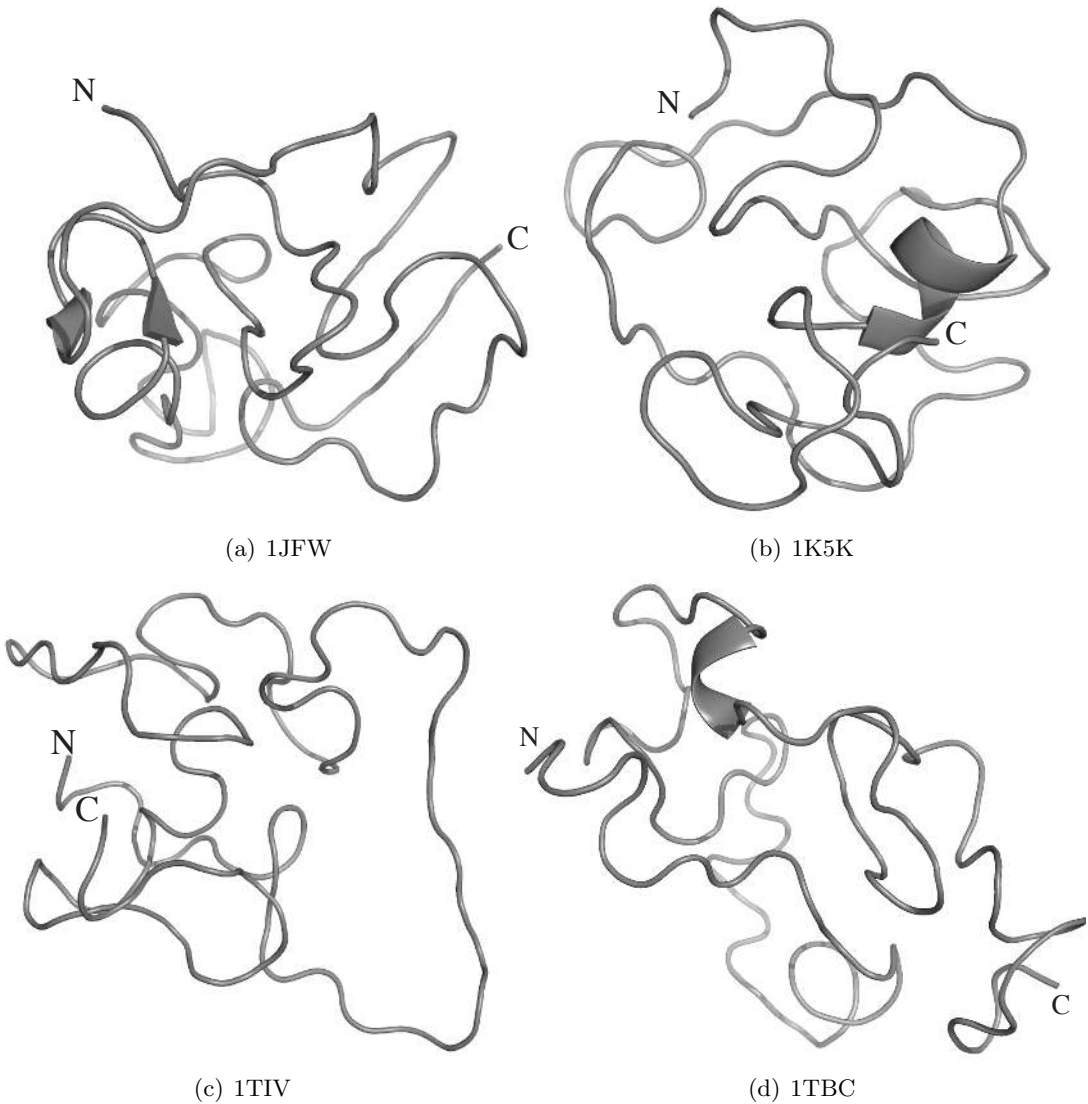
(a)




(b)

```
HIV1_Tat72      ----MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYG----RKKRRQRRRPPQGSQTHQVSLSKQ 72
EIAV_Tat75      MADRRIPGTAEENLQKSSGGVPGQNTGGQEARPN-YHCQLCFL-RSLGIDYLDASLRKKNKQRLKAIQQGRQPQYLL--- 75
```


0
-2000
-4000
-6000
-8000
-10000
MRE (deg•cm2•dmol-1)
185
200
220
240
Wavelength (nm)

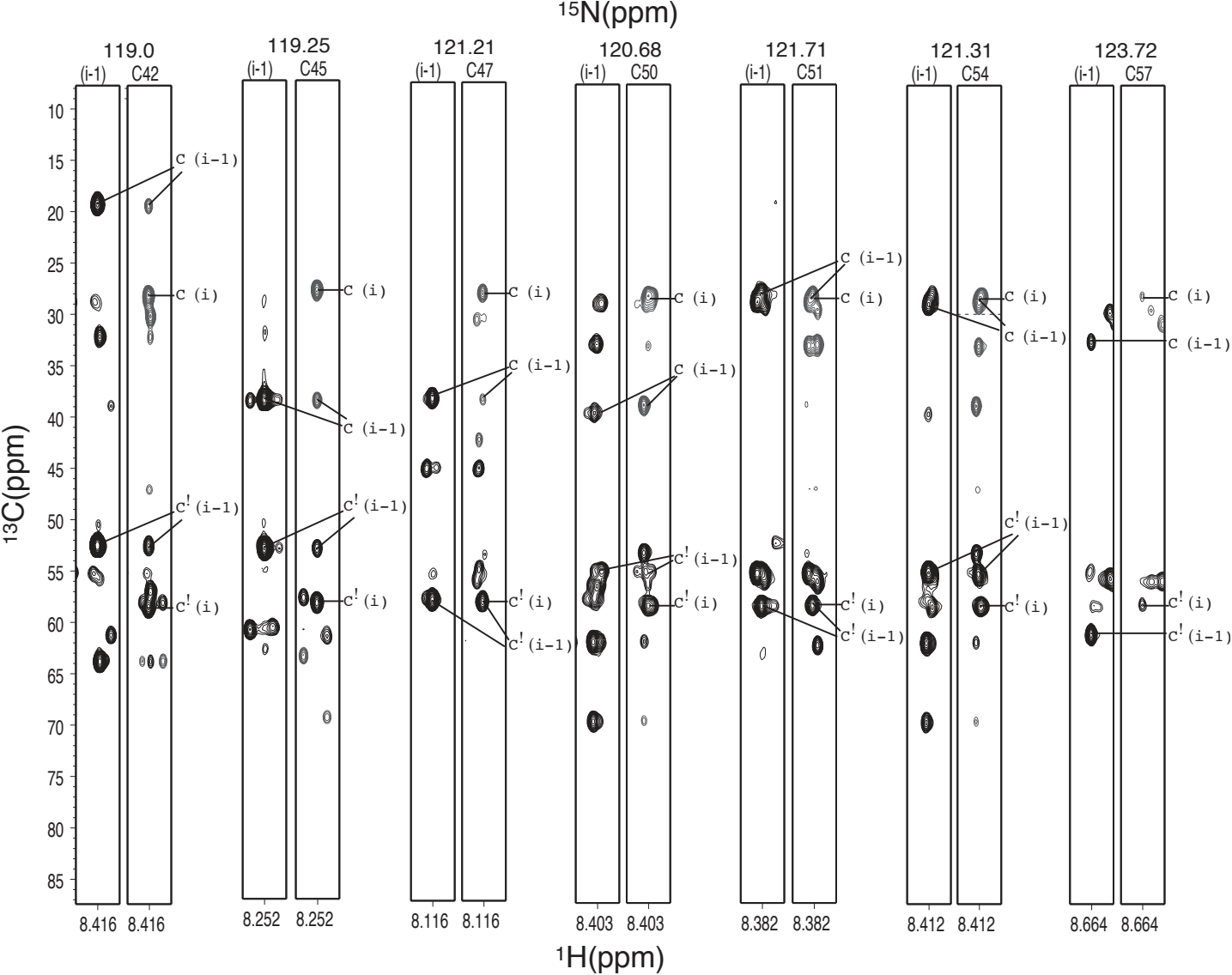


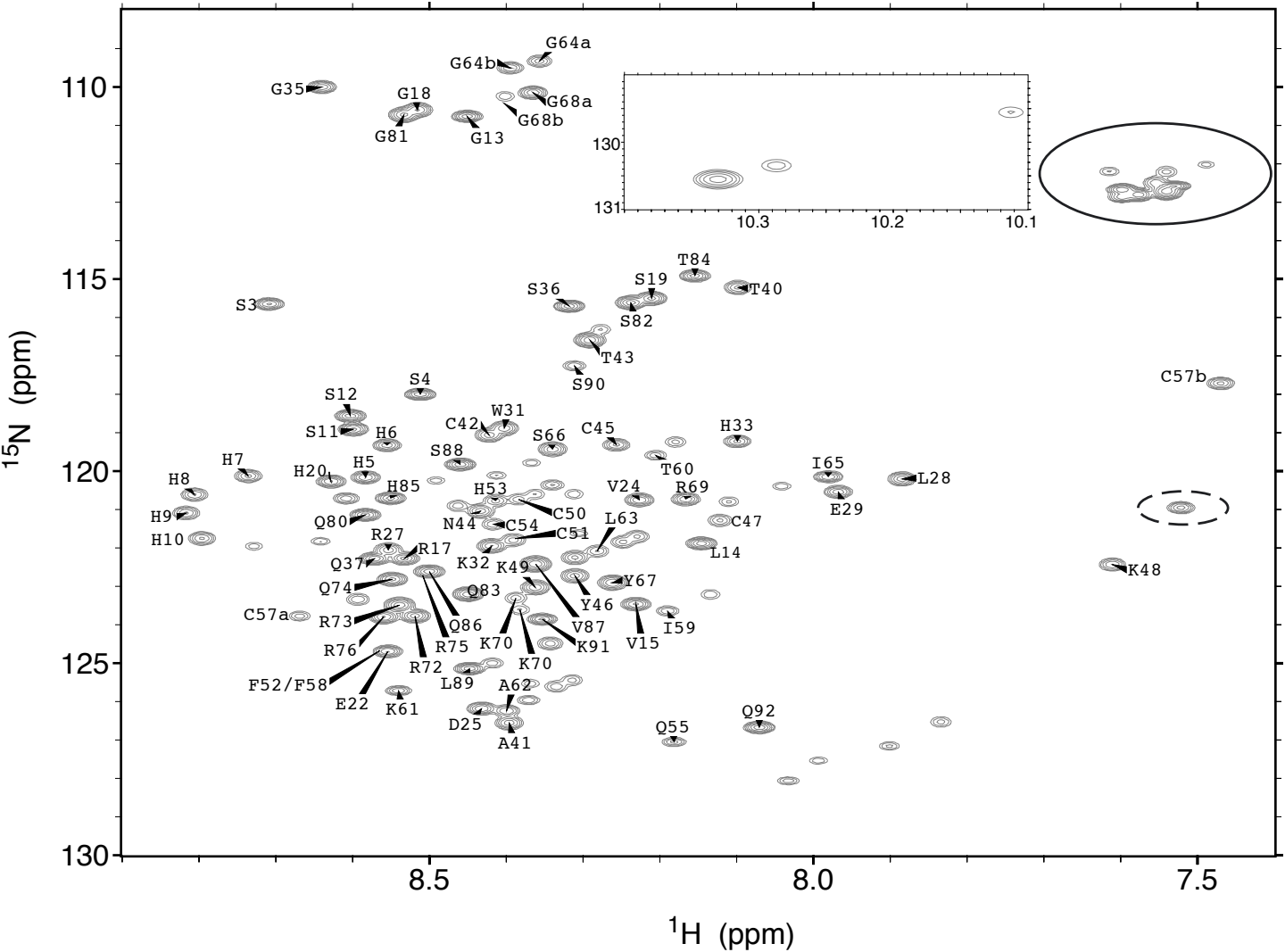

(a) 1JFW

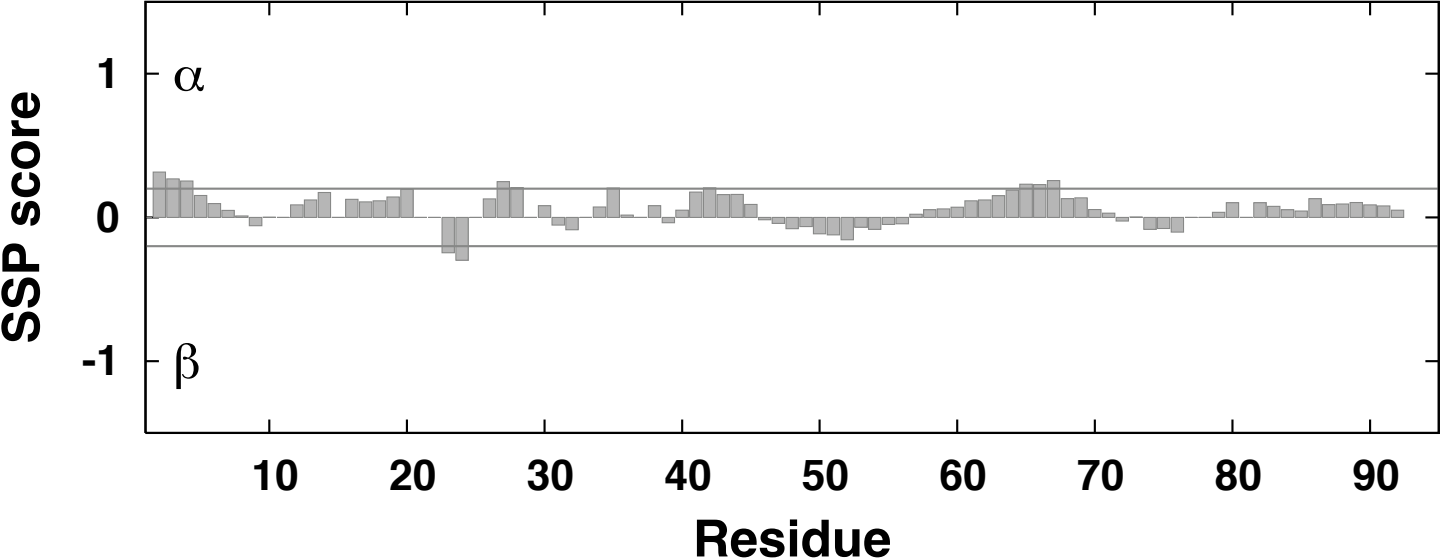
(b) 1K5K

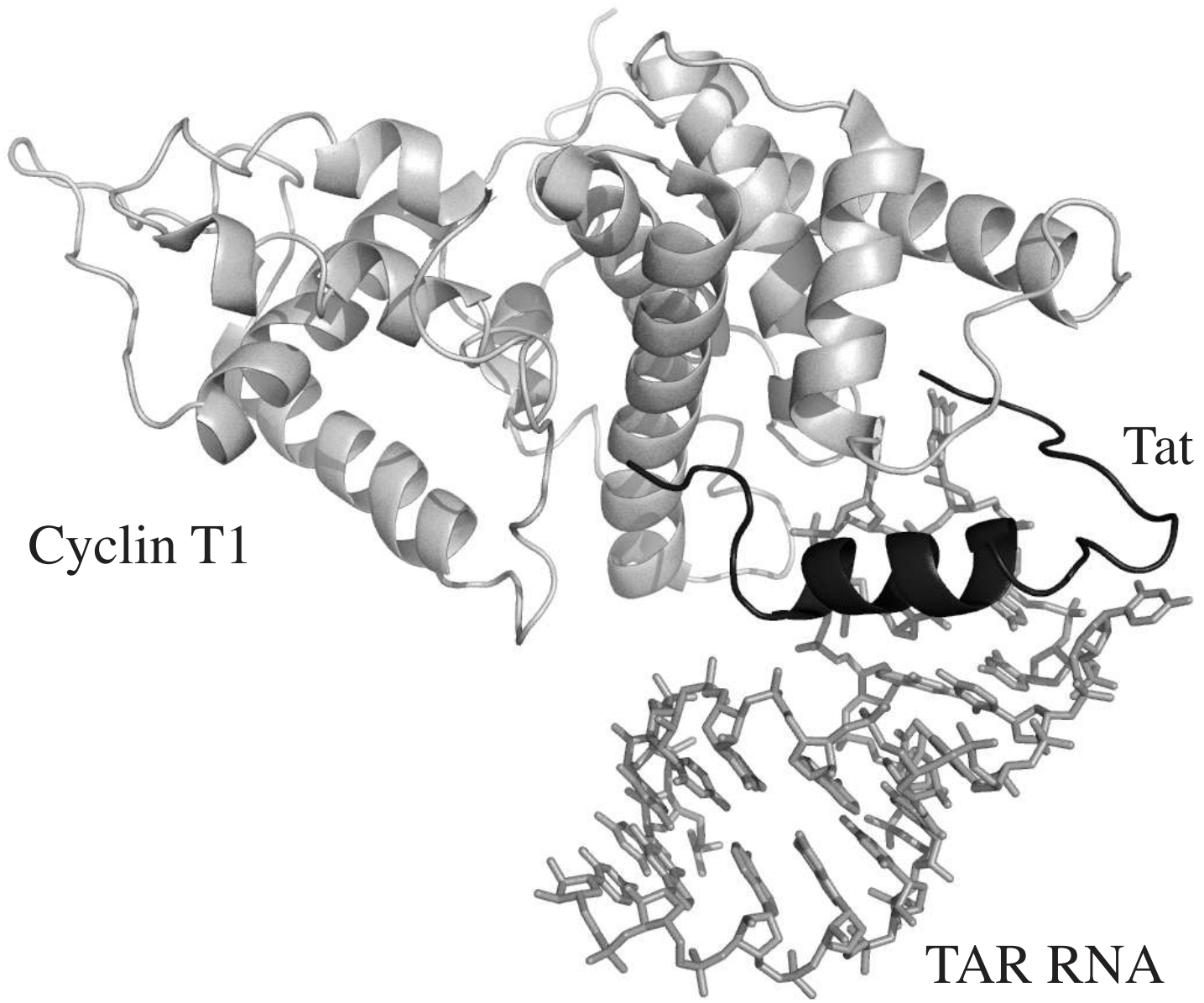
(c) 1TIV

(d) 1TBC


$^{15}$N(ppm)
119.0
(i-1)
C42
119.25
(i-1)
C45
121.21
(i-1)
C47
120.68
(i-1)
C50
121.71
(i-1)
C51
121.31
(i-1)
C54
123.72
(i-1)
C57
$^{13}$C(ppm)
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
c (i-1)
c (i)
c! (i-1)
c! (i)
8.416
8.252
8.116
8.403
8.382
8.412
8.664
$^{1}$H(ppm)


G64a
G64b
G35
G18
G68a
G68b
G81
G13
T84
S19
T40
S36
S3
S82
T43
S90
S4
C57b
S12
W31
C42
S11
H6
C45
H33
S66
S88
T60
I65
H7
H20
H5
H8
H85
H53
V24
R69
L28
E29
H9
Q80
N44
C54
C50
L63
C47
H10
R27
C51
L14
R17
Q37
K32
K49
K48
Q74
Q83
Y67
C57a
R73
Y46
Q86
V87
I59
R76
R75
K70
K91
V15
R72
K70
F52/F58
L89
A62
E22
K61
D25
Q55
Q92
A41
130
131
10.3
10.2
10.1
110
115
120
125
130
8.5
8.0
7.5
$^{15}$N (ppm)
$^{1}$H (ppm)


SSP score
1
0
-1
α
β
10
20
30
40
50
60
70
80
90
Residue


Cyclin T1
Tat
TAR RNA