HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2010 December 10.

Published in final edited form as:

Nature. 2010 June 10; 465(7299): 747–751. doi:10.1038/nature09131.

Crystal structure of HIV-1 Tat complexed with human P-TEFb

Tahir H. Tahirov[1], **Nigar D. Babayeva**[1], **Katayoun Varzavand**[2], **Jeffrey J. Cooper**[2], **Stanley C. Sedore**[2], and **David H. Price**[2]

[1]Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, NE 68198-7696, USA.

[2]Biochemistry Department, University of Iowa, Iowa City, IA 52242

Abstract

Regulation of the expression of the human immunodeficiency virus (HIV) genome is accomplished in large part by controlling transcription elongation. The viral protein Tat hijacks the host cell's RNA polymerase II elongation control machinery through interaction with the positive transcription elongation factor, P-TEFb, and directs the factor to promote productive elongation of HIV mRNA. Here we describe the crystal structure of the Tat•P-TEFb complex containing HIV-1 Tat, human Cdk9, and human Cyclin T1. Tat adopts a structure complementary to the surface of P-TEFb and makes extensive contacts, mainly with the Cyclin T1 subunit of P-TEFb, but also with the T-loop of the Cdk9 subunit. The structure provides a plausible explanation for the tolerance of Tat to sequence variations at certain sites. Importantly, Tat induces significant conformational changes in P-TEFb. This finding lays a foundation for the design of compounds that would specifically inhibit the Tat•P-TEFb complex and block HIV replication.

According to the most recent UNAIDS AIDS Epidemic Update, as of the end of 2008, over 33 million people were living with HIV. Although the number of people infected has increased about 30% over the last 10 years, part of the increase is due to improved antiretroviral therapies and an increase in the number of people treated. Unfortunately, strains of HIV that are resistant to current anti-HIV drugs arise rapidly due to the error-prone reverse transcription step in the viral life cycle, and to a very high frequency of recombination that leads to a generation of mutant viruses that are unresponsive to the drugs. 1 Because of this and the lack of effective vaccines2, the search for new and better anti-HIV

Correspondence and requests for materials should be addressed to T.H.T. (ttahirov@unmc.edu)..

Author Contributions THT managed the crystallization and structure determination part of the project, solved the crystal structures and prepared the manuscript. NDB obtained the crystals. Diffraction data collection was performed by NDB and THT. Protein cloning, expression, purification and writing the corresponding methods sections were performed by SCS, KV, and JJC, respectively. DHP managed the protein production part of the project, and helped generate and edit the manuscript.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Atomic coordinates and structure factors for Tat•P-TEFb and Tat•P-TEFb•ATP structures have been deposited in the Protein Data Bank with accession numbers 3mi9 and 3mia, respectively. Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing financial interests.

therapies continues. Most current HIV drugs inhibit viral enzymes, but host proteins that interact with viral proteins are also potential targets. One such host protein is the positive transcription elongation factor, P-TEFb, which interacts with the viral transactivator of transcription, Tat.3-5

P-TEFb plays a key role in regulating the elongation phase of transcription by RNA polymerase II.6 It is a cyclin-dependent kinase comprised of Cdk9 paired with one of several cyclin subunits, and in its active form is responsible for relief of the block to promoter proximally paused polymerases.7-9 P-TEFb is generally required for the generation of mRNAs10 and is recruited to genes by a variety of gene specific proteins including NFκB, CIITA, and nuclear receptors, as well as more generally by the bromodomain containing protein Brd4.6, 11 The activity of P-TEFb is globally regulated by reversible association with the 7SK snRNP in which it is held in an inhibited state by an RNA-binding protein HEXIM1 or HEXIM2.6, 11, 12

HIV has evolved to manipulate both the Cdk9•CyclinT1 form of P-TEFb and the cellular process controlling P-TEFb. Expression of HIV Tat in otherwise uninfected cells leads to the release of P-TEFb from the 7SK snRNP and the formation of a Tat•P-TEFb complex.13 During an active HIV infection, the Tat•P-TEFb complex is recruited to the transactivation response element, TAR, which is manifested in the first 59 nucleotides of the nascent HIV transcript.5, 14-16 In addition to providing a means to recruit P-TEFb to promoter proximally stalled polymerases engaged in the HIV LTR, Tat stimulates the kinase activity and changes the substrate specificity of P-TEFb.17

Here we report the high-resolution crystal structures of the Tat•P-TEFb complex and the Tat•P-TEFb complex with bound ATP analog AMPPNP and magnesium ion (further referred to as Tat•P-TEFb•ATP). The structures reveal that Tat interacts with both the Cyclin T1 and Cdk9 subunits of P-TEFb, and provide evidence for a Tat-induced conformational change in P-TEFb. Our findings also provide a structural basis for the design of compounds that would specifically inhibit HIV transcription.

Crystal structure of Tat•P-TEFb complex

The structure of the Tat•P-TEFb complex is refined at 2.1 Å resolution to an R of 21.9% and R_{free} of 26.4%, and the structure of the Tat•P-TEFb•ATP is refined at 3.0 Å resolution to an R of 21.8% and R_{free} of 28.6%. The Cdk9 subunit is phosphorylated at Thr186 and is in a conformation characteristic for active cyclin-dependent kinases.18

The overall structure of the Tat•P-TEFb•ATP complex depicted in Figure 1 reveals that Tat acquires an extended conformation mainly through interactions with Cyclin T1 (see Supplementary Figure 1 for the definition of secondary structure elements). As expected from earlier biochemical studies,19 Tat forms a Zn-mediated bridge with Cys261 of Cyclin T1. Tat binding to P-TEFb buries 3499 $Å^2$ of surface area, of which 88% is on Cyclin T1 and 12% is on Cdk9. Clearly, Tat has evolved to bind tightly to P-TEFb as 37% of its folded portion (amino acids 1-49) surface is complementary to the kinase. The amount of surface buried between P-TEFb and Tat is double the average value for stable protein•protein interactions.20 The CDK•Cyclin interaction surface is smaller in P-TEFb than in other such

complexes.21 However, Tat augments it by inserting in a groove at the heterodimer interface and leading to a more stable and more active P-TEFb complex.17

Tat is encoded by two exons. Amino acids 1-72 encoded by the first exon are highly conserved and comprise five regions (Figure 2a) that are sufficient for transactivation of the viral LTR promoter. The first three regions constitute a minimal activation domain capable of binding to Cyclin T1. Indeed, the residues 1-49 are ordered in the crystal structure of Tat•P-TEFb, while the density for the rest of Tat (residues 50-86) is not defined.

The acidic/proline-rich region is in an extended loop conformation and folded as a random coil with two type II β turns and one type II′ β turn (Figures 2b and 2c). The mutations of the acidic or proline residues as well as deletion mutants within this region impair the activity of Tat.22 A cysteine-rich region and core regions form a more compact structure with a random coil, two helices, a tail, and two bound Zn ions (Figures 2b and 2c). The short 3_{10}-helix is located between the Zn ions (Zn1 and Zn2) and interacts with both zincs. The α-helix interacts only with Zn1 and extends toward the C-terminus of the activation domain. Zn1 is coordinated by Cys22, His33, Cys34 and Cys37 of Tat (Figures 2d and 2e), and Zn2, located 10.7 Å away, is coordinated by Cys25, Cys27 and Cys30 of Tat and Cys261 of Cyclin T1 (Figures 2f and 2g). The observed coordination of Zn ions explains earlier mutational studies pointing to the importance of His33 and all cysteines, except Cys31, in the Zn-dependent function of Tat.23, 24 The Cys31 is replaced by a serine in several Tat variants and this replacement does not diminish the protein function.25 The Zinc Fingers (ZnFs) in Tat should be considered unique because protein fold searches did not uncover similar ZnF motifs in any other protein.

In addition to the main-chain hydrogen bonds involved in secondary structure elements, only six intramolecular hydrogen bonds were found in Tat (Supplementary Figure 2). Three of these hydrogen bonds were between the side chain of Lys41 and main-chain oxygens of Thr23, Cys25, and Cys30. The latter intramolecular interactions appear to be crucial for the structural integrity of the cysteine-rich/core module because mutation of Lys41 impairs the activity of Tat.25, 26

Recently the crystal structure of the fusion of equine Cyclin T1 and equine infectious anemia virus (EIAV) Tat in complex with EIAV TAR RNA was reported.27 The basic RNA-recognition motif of EIAV Tat adopts a helical structure whose flanking regions interact with an equine Cyclin T1 and both proteins coordinate the stem-loop structure of EIAV TAR.27 However, the key residues in EIAV TAT and equine Cyclin T1 involved in ternary complex formation with TAR RNA, are different in HIV-1 Tat and human Cyclin T1.27-29 The TAR RNAs from HIV-1 and EIAV have different structures as well.27, 30 Moreover, the photocross-linking and protein footprinting studies31 show that HIV-1 TAR RNA loop interacts with Cyclin T1 residues 252–260, which is far from the EIAV TAR RNA docking site. These differences make it difficult modeling HIV-1 TAR binding to Tat•P-TEFb complex based on the structure of the EIAV complex.27 However, it is appropriate to apply the structure of HIV-1 Tat•P-TEFb to model the portion of the activation domain for the EIAV Tat sharing high sequence identity with HIV-1 Tat in ZnF and core regions (Supplementary Figure 3a). The proposed model of EIAV Tat ZnF and

core regions in complex with Cyclin T1 is shown in Supplementary Figure 3b and displays an uninterrupted connection with the basic region. The high sequence conservation of the central ZnF-core motif in Tat proteins, compared with poorly conserved flanking regions, indicates that the binding of ZnF-core motif to a Cyclin T1 surface is the most conserved structural feature shared among Tat transcriptional regulators.

Tat•P-TEFb interactions

Interactions between Tat and P-TEFb can be divided into four areas (Figures 3a and 3b). The first area involves a U-shaped, acidic/proline-rich region of Tat and a wide depression between the two cyclin repeats of Cyclin T1. The second area involves a β-turn in the acidic/proline-rich region of Tat and the T-loop of Cdk9. In the third area, the cysteine-rich region and a core region of Tat interact with a depression between the cyclin repeats of Cyclin T1 and a point close to the helix H5′. The fourth area involves the second ZnF of Tat and the Cys261 of Cyclin T1 (Figure 2f and 2g).

Tat possesses an extended hydrophobic patch comprised of the N-terminal methionine and amino acid residues from a cysteine-rich region and the core region which are packed against the hydrophobic surface of Cyclin T1 (Supplementary Figure 4a). Moreover, several intermolecular van der Waals contacts can be seen at discrete locations (Supplementary Figures 4b and 4c).

Tat forms numerous direct and water-mediated hydrogen bonds with both Cdk9 and Cyclin T1 (Supplementary Table 1 and Supplementary Figures 5 and 6). The fewer Tat-stabilizing intramolecular hydrogen bonds compared with intermolecular Tat•P-TEFb hydrogen bonds indicates that overall the conformation of Tat is determined predominantly by its interactions with P-TEFb.

Conformational changes of Tat and P-TEFb

Several solution structures of Tat variants have been reported.32 In the absence of interacting partners, Tat does not have prominent secondary structure elements, indicating the flexible nature of free Tat proteins.32 However, upon interaction with P-TEFb and two Zn ions, the activation domain of Tat acquires a well ordered structure (Figure 1).

To elucidate Tat-induced conformational changes in P-TEFb, we compared the structure of the Tat•P-TEFb•ATP complex with the structure of P-TEFb•ATP (3blq) solved in the Johnson laboratory.21 Although there are differences in P-TEFb subunit constructs and crystal packing, four notable changes were observed. First, Tat binding resulted in the unfolding and disordering of the Cyclin T1 α-helix HC (residues 253-256) and a 4.6 Å displacement of the C-terminal portion of the preceding α-helix H5′ (Figure 4a). These changes expose buried surface in Cyclin T1 that is used for the binding of the ZnF and core region sequences of Tat. Second, when the Cyclin T1 subunits of P-TEFb•ATP and Tat•P-TEFb•ATP were aligned, a significant, 8.5° rotation was detected for the Cdk9 subunit (Figure 4b). Third, the shift of phospho-Thr186 by 2.6 Å resulted in the formation of two hydrogen bonds with Arg65 and one additional hydrogen bond with Arg148 (Figure 4c). Fourth, the concerted allosteric changes result in the switch of the conformations in the $\beta_1\beta_2$-

loop and $\beta_3\alpha$C-loop which are in close proximity to the ATP binding site (Figure 4d). To understand the possible sequence of events leading to the latter conformational changes upon Tat binding, we superimposed the Cdk9 subunit of Tat•P-TEFb•ATP with the Cdk9 subunit of P-TEFb•ATP. Figure 4d demonstrates a significant shift of the first cyclin repeat of Cyclin T1 toward Cdk9 in the structure of Tat•P-TEFb•ATP relative to structure of P-TEFb•ATP. This shift reaches up to 4.8 Å and leads to steric hindrance between the α-helix H5 of Cyclin T1 and the $\beta_3\alpha$C-loop of Cdk9, in particular between Val134 and Ser138 of Cyclin T1 and Gly58 and Met52 of Cdk9, respectively (Figure 4e). To release this steric hindrance, the Gly58 of Cdk9 is shifted 2.4 Å away from Val134 of Cyclin T1 and the part of the Cdk9 $\beta_3\alpha$C-loop comprised of residues 51 to 55 changed their conformation. As result of this conformational change the Leu51 clashes with Phe30, pushing it away and forcing the residues 25 to 32 of $\beta_1\beta_2$-loop to acquire a conformation which is different than in P-TEFb•ATP (Figure 4e). The Tat-induced conformational changes in Cdk9 modify the substrate-binding surface of P-TEFb•ATP. This provides a structure-based explanation for the earlier findings showing that Tat-bound P-TEFb phosphorylates not only the Ser2 but also the Ser5 in the heptad repeats of the RNA polymerase II large subunit C-terminal domain.17

Structural constraints for Tat mutations

HIV-1 genes are prone to rapid mutation leading to high genetic variability. HIV-1 has been classified into three genetic groups: major (M), outlier (O) and new (N) that are non-M and non-O.33 Most HIV-1 infections are caused by group M viruses. The group M is divided into nine subtypes: A–D, F–H, J and K.33 Genetic variation within a subtype is 15–20%, and is 25–35% between subtypes.33 In line with this observation, Tat tolerates up to 40% sequence variation without loss of transcriptional activity.32 Our structure provides insight into how mutations in Tat are tolerated. To facilitate the description of sequence variations, we produced sequence logos34 of the Tat activation domain using the alignment of 1,118 Tat sequences35 (Figure 5a). The alignment indicates that half of the residues have a higher degree of conservation, with 14 invariant residues and 10 residues with randomly occurring polymorphs (Figure 5b). The apparent reasons for the conservation are the following: involvement in Zn-coordination; interaction with the T-loop of Cdk9; stabilization of Tat structure by intramolecular hydrogen bonds; interaction with a hydrophobic surface of Cyclin T1; formation of hydrogen bonds with Cyclin T1; providing rigidity to the loop connecting the acidic/proline-rich region and a cysteine-rich region. There is no alternative for glycine at position 15, because any mutation would create a backlash with Asn53 side chain of Cyclin T1 and disrupt its hydrogen bond with Gln17 of Tat (Supplementary Figure 7). The analysis of Tat•P-TEFb structure alone does not provide an immediate explanation for the functional importance of Lys28. However, it explains an earlier finding indicating that acetylated Lys28 interacts with Asn257 of Cyclin T1 and facilitates the formation of Tat•TAR•P-TEFb complex.36 Indeed, Lys28 is located at the surface of the second ZnF (Figure 2e); in close proximity to the missing Cyclin T1 residues (252-260) linking Cys261 to the Cyclin domain and harboring Asn257. Acetylated Lys28 could reach the Asn257 side chain and form an intermolecular hydrogen bond.

Residues with a high mutation rate are divided into two groups: those with predominantly functionally equivalent substitutions (Figure 5c) and those with a variety of substitutions (Figure 5d). All these residues, except Gln35 and Ile39, are fully or partially exposed. None of the observed mutations would disrupt interactions of Tat with P-TEFb. Gln35 penetrates deeply into a depression on Cyclin T1 and forms a hydrogen bond with Asn180 (Supplementary Figure 7a). All Gln35 mutations (Figure 5a), except tyrosine, could be accommodated with minor adjustments in the positions of surrounding side chains. However, modeling of tyrosine shows that it may shift the N-terminal methionine without disrupting the Tat•P-TEFb (Supplementary Figure 7a). Another frequently mutated residue buried at the Tat•P-TEFb interface is Ile39 (Figure 5a). The Ile39 side chain faces a groove which is occupied by a water molecule (Supplementary Figure 7c), and hence all Ile39 mutations could be easily accommodated without disrupting the Tat•P-TEFb.

Implications for P-TEFb activation

The Tat•P-TEFb structure provides insight into how Tat functions in recruiting P-TEFb to the HIV LTR. The first step of this process involves a Tat-induced release of P-TEFb from the 7SK snRNP-containing complex.13, 37 Most of P-TEFb within the cell is held in an inactive form by HEXIM1, which is bound to 7SK snRNA along with other proteins.38, 39 Several experiments argue that inhibition of P-TEFb by HEXIM1 resembles the inhibition of Cdk2•CyclinA by $p27^{kip1}$.40 In the Cdk2•CyclinA•$p27^{kip1}$ crystal structure the side chains of Phe87 and Tyr88 of $p27^{kip1}$ enter the active site cleft and prevent the binding of ATP.40 Mutational studies of invariant tyrosine and phenylalanine residues of HEXIM1 revealed that Tyr271 and Phe208 mutations relieved the inhibitory activity of P-TEFb without dissociation of HEXIM1.41 These results suggest that HEXIM1 blocks the binding of ATP to Cdk9 in a manner similar to $p27^{kip1}$. In such a case the flipping of the $\beta_1\beta_2$-loop conformation would lead to dissociation from P-TEFb (Supplementary Figure 8). In support of this idea, P-TEFb with a shorter Cyclin T1 (1-254) and a missing α-helix HC was not capable of binding 7SK snRNA,42 indicating the important contribution of this helix to P-TEFb•HEXIM1•7SK snRNA complex formation. Because Tat unfolds the HC of Cyclin T1 this also may facilitate the release of HEXIM1 (Supplementary Figure 8).

Conclusions

The emergence of multidrug resistant HIV-1 strains requires the development of new drugs capable of stopping viral replication.43 Tat plays a major role in the viral replication and remains a viable target for crystallographic studies and structure-based drug design. The structure of Tat•P-TEFb described here revealed a large interaction interface buried between these two factors and showed how local conformational changes caused by Tat mutations could be tolerated without disruption of the complex. The extended and flexible interaction-dependent conformation of Tat suggests that designing compounds that would block the interaction of Tat with P-TEFb would be a challenging task. The disadvantage of such inhibitors might be also the interference with the function of cellular factors (for example HEXIM1, Brd4, etc.) which bind to and are essential for the regulation of P-TEFb. On the other hand, the Tat binding changes the overall shape and surface of P-TEFb. This provides an opportunity for the design of inhibitors which would not bind to or inhibit the normal

cellular function of P-TEFb, but would be specific only to the form of P-TEFb utilized by the virus.

METHODS SUMMARY

C-terminally His tagged human Cdk9 (1-345), human Cyclin T1 (1-266), and HIV-1 Tat (1-86) were co-expressed in insect cells and purified by Ni-NTA and Mono S columns. The purified Tat•P-TEFb complex was crystallized using the sitting-drop vapor-diffusion method. The Tat•P-TEFb•ATP crystals were obtained by soaking AMPPNP and magnesium chloride into the Tat•P-TEFb crystals. Transformation of crystals to a low-humidity form dramatically improved the diffraction enabling the complete data collection at high resolution using synchrotron radiation. The Tat•P-TEFb structure was solved using molecular replacement method and refined at 2.1 Å resolution. The positions of AMPPNP and magnesium in Tat•P-TEFb•ATP structure were determined by the difference Fourier map and the structure was refined at 3 Å resolution. Both structures exhibited electron density maps of high quality and provided the models with excellent geometry.

METHODS

Expression of Tat•P-TEFb complex

Sf9 insect cells were co-infected with three baculoviruses that individually directed the expression of C-terminally His-tagged human Cdk9 (1-345), human Cyclin T1 (1-266), and HIV-1 Tat (1-86). The three coding sequences were individually cloned into the pENTR/SD/D-TOPO vector (Invitrogen K2420) after PCR amplification.

The baculoviruses were generated by LR recombination reactions using BaculoDirect C-Term Expression kit (Invitrogen, 12562-013), and these viruses were transfected into Sf9 cells following the manufacturer's instructions. The viruses were amplified several times to obtain viral stocks with 10^8 pfu/ml. Stocks were titered by infecting individual wells of a 6-well plate containing 1×10^6 Sf9 cells with 0.001, 0.01, 0.1, 1, 10 or 100 μl of the stocks. After 48 hours the cells were fixed with 3.7% formaldehyde for 15 minutes and stained using first a gp64 antibody (Novagen, 70814-3), then by an alkaline phosphatase-coupled goat anti-mouse IgG (Sigma-Aldrich, A 3562); color development with BCIP/ NBT (Sigma-Aldrich, B 6404) followed. The presence of the gene of interest was determined by conducting PCR on DNA isolated by phenol extraction from the stocks.

Large scale production of the Tat•P-TEFb complex was carried out using baffled 2 liter glass Erlenmeyer flasks containing 1 liter of Sf-900 serum-free medium (GIBCO 10902-088) with 1.5×10^6 Sf9 cells/ml. The flasks were rotated at 70 rpm at 27°C. Typically, 6 liters of cells were infected at once by adding 20 ml of each of the three viral stocks to each liter of culture. The flasks were then incubated for an additional 72 hours.

Purification of the Tat•P-TEFb complex

To avoid protein degradation, all solutions and equipment were kept cold and the purification was performed swiftly. Cells were harvested from six 1 liter cultures by spinning for 10 minutes at 4200 rpm in a Beckman JS-4.2 rotor. Each of the 6 pellets was

resuspended in 20 ml of PBS supplemented with 1 μg/ml E-64 protease inhibitor (Calbiochem 324890), 1 × Roche Complete Protease Inhibitor (Roche 11873580001), and 0.1% of a saturated PMSF solution in isopropanol then transferred to 40 ml tubes. After centrifuging at 3000 rpm for 5 minutes in a Beckman JS 13.1 rotor, the supernatants were decanted and pellets were resuspended in 12 ml of Lysis Buffer (5 mM imidazole, 150 mM NaCl, 10 mM Tris (pH 7.8), 1% Triton, 2 mM $MgCl_2$, 1 μg/ml E-64, 1 × Roche Complete Protease Inhibitor, and 0.1% PMSF solution). Lysates were combined into three 50 ml plastic beakers on ice, and each was sonicated 5 times for 15 seconds with a 5-minute pause between bursts. Cell lysates were combined and brought to 500 mM NaCl by adding 5 M NaCl and mixing vigorously then centrifuged at 45000 rpm for 45 minutes in a Beckman 55.2Ti rotor in Oak Ridge tubes. The high-speed supernatant was incubated with 4.5 ml of Qiagen Ni-NTA Agarose that had been equilibrated in Lysis Buffer containing 500 mM NaCl. After 60 minutes the resin was pelleted at 1200 rpm for 3 minutes in a Beckman JS-13.1 rotor. The resin was resuspended in Wash Buffer (5 mM imidazole, 500 mM NaCl, 10 mM Tris [pH 7.8], 1% Triton, and 0.1% PMSF solution) and transferred equally into three 20 ml BioRad disposable columns. Each column was washed successively with 10 ml of Wash Buffer, 15 ml of HSW (25 mM imidazole, 1 M NaCl, 10 mM Tris [pH 7.8], 1% Triton, and 0.1% PMSF solution), and 15 ml of 70 mM HGKE Wash (25 mM imidazole, 70 mM KCl, 25 mM HEPES [pH 7.6], 15% glycerol, 0.1 mM EDTA, 1% Triton, and 0.1% PMSF solution). Each column was then eluted with 6 ml of HGKE Wash supplemented with 300 mM imidazole. The eluted material was pooled and centrifuged at 13000 rpm for 15 minutes in a Beckman JS-4.2 rotor and then loaded onto a 1 ml Mono S column that had been equilibrated in 70 mM HK Buffer (70 mM KCl, 10 mM HEPES [pH 7.6], 1 mM DTT, and 0.1% PMSF solution). After a 5 ml wash the column was eluted with a 17.5 ml gradient from 70 to 620 mM KCl in HK Buffer (Supplementary Figure 9a). Free P-TEFb eluted at 220 mM KCl and the Tat•P-TEFb complex at 420 mM KCl. The peak fractions were relatively free of contaminating proteins (Supplementary Figure 9b). The purification protocol was performed more than 20 times with minor modifications and usually resulted in 5-10 mg of the Tat•P-TEFb complex per 6 liters of cells. The obtained samples were kept in small aliquots at −80 °C.

Crystallization

The frozen Tat•P-TEFb samples were thawed on ice, dialyzed against 400 mM KCl, 10 mM HEPES buffer (pH 7.5), and 2 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), then concentrated to 5-6 mg/ml using Millipore Ultrafree centrifugal devices. Only few fractions reached these concentrations because sudden precipitation occurred on most fractions even at lower concentrations. The state of sample aggregation was monitored by Dynamic Light Scattering. Application of the solubility screening protocol described by Jancarik *et al.*44 did not improve the sample solubility and aggregation. Initial crystallization screening was performed at a temperature of 295 K in 96-well plates using 50% diluted Crystal Screen and Crystal Screen 2 solutions (Hampton Research) by the sitting-drop vapor-diffusion method by mixing 1 μl of complex solution with 1 μl of reservoir solution. 1 mM TCEP was added to each condition and finally 400 mM KCl was added to each reservoir. Crystals appeared in 37^{th} condition of the Crystal Screen 2 containing 50 mM HEPES buffer (pH 7.5), 5% PEG 8,000, 2 mM TCEP, and 2% ethylene

glycol in the crystallization drop plus 400 mM KCl in the reservoir only. Variations in concentration of the components, molecular weight of PEGs, buffers and pH, and inclusion of additives from Hampton Research Additive Screen revealed the following conditions producing crystals of better shape: 50 mM HEPES buffer (pH 7.5), 4.25-5% PEG 20,000, 1 mM TCEP, and 20 mM glycyl-glycyl-glycine mixed with protein solution. The reservoir solution contained also 400 mM KCl. Most drops produced small needle-like crystals in 1-2 days; however, in some of remaining clear drops larger crystals of hexagonal prism shape (Supplementary Figure 10a) appeared in 1-3 weeks and sometimes after 1-3 months. The growth of crystals normally continued for 2-4 weeks, reaching the maximum dimensions of $0.25 \times 0.25 \times 1.5$ mm^3. The polarization of light by these crystals was barely detectable. The SDS-PAGE of dissolved crystals revealed all three subunits of Tat•P-TEFb in stoichiometric (1:1:1) proportions (Supplementary Figure 10b). The Tat•P-TEFb•ATP crystals were obtained by soaking the Tat•P-TEFb crystals in crystallization solution with 1mM ATP analog AMPPNP and 5 mM magnesium chloride for 40 min.

X-ray diffraction data collection and processing

For the preliminary studies the crystals were soaked in cryoprotectant for a few seconds, scooped with a nylon-fiber loop and flash cooled in a dry nitrogen stream at 100 K. The best cryoprotectant was obtained by adding 25% (v/v) glycerol or ethylene glycol to the reservoir solution. Preliminary diffraction data were collected at cryogenic temperature on a Rigaku R-AXIS IV imaging plate using Osmic VariMax™ HR mirror-focused CuK_α radiation from a Rigaku FR-E rotating anode operated at 45 kV and 45 mA. Despite a large volume of crystals, the diffraction was poor for most crystals, reaching a resolution of only 4.2-4.5 Å. Occasionally some crystals diffracted better, up to a resolution of 3.5-3.8 Å. Inspection of unit cell parameters revealed that, in general, better-diffracting crystals had slightly shorter (1-1.5 Å) parameters *a* and *b*. Because the crystals with such properties are often capable of transformation to a low-humidity form, we started such experiments with Tat•P-TEFb crystals with modification of earlier described experimental setup.45 Instead of mounting the crystal directly in capillary, we mounted the cryoprotectant-soaked crystal (with the content of glycerol or ethylene glycol reduced to 12-16% v/v) in a MiTeGen MicroMeshes™, wiped the excess but not all solution around the crystal with the fine-cut filter paper over the top of the well with cryoprotectant solution, immediately placed it in a capillary having a cryoprotectant solution on the wider end and sealed its bottom touching the goniometer base (without using the hot wax). Cryocpotectant in capillary was replaced with solutions each having the PEG 20,000 concentration increased by 3% every 5-10 minutes until initial signs of transformation such as the reduction of the volume of the remaining solution around the crystal. Then the capillary was removed and crystal was flash cooled under the stream of nitrogen gas. It was important not to reach the point of transformation when an apparent “sweating” of crystal occurs because further reduction of the relative humidity results in dramatic shrinkage of crystals, especially along the axis *c*, and complete loss of diffraction. This method of transformation to a low-humidity form was effective only for crystals having cross section of over 0.15 mm, probably due to fast dehydration of smaller crystals. The most successful transformation of Tat•P-TEFb crystals was accompanied with the reduction of unit cell parameters *a* and *c* by 4.5 Å and 1.5 Å, respectively, and extension of the X-ray diffraction resolution limits from 4.5-3.5 Å to 2.1

Å. The high-resolution diffraction data were collected using synchrotron radiation at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines BL24ID-C and BL24ID-E. To minimize the radiation damage, complete data sets were obtained from one crystal of Tat•P-TEFb and one crystal of Tat•P-TEFb•ATP, each exposed at six positions. The data set for the calculation of the Zn anomalous diffraction signal was collected by exposing only one position of the crystal with a beam intensity reduced to 20%. The intensity data were indexed, integrated, and scaled using the HKL2000 program package46 (Supplementary Table S2).

Structure determination and refinement

The structure was determined by the molecular replacement method starting with the coordinates of P-TEFb (PDB entry 3blh).21 All amino acid residues with temperature factors above 80 $Å^2$ were excluded from the positioned model before the calculation of phases. Density modification with CNS software47 reduced the model bias and improved the electron density maps. The building of the Tat structure and missing parts of the Cdk9 and Cyclin T1 subunits, as well as adjustment of the initial model, was performed manually in 5 rounds with TURBO-FRODO software. The positions of Zn atoms were confirmed using anomalous difference Fourier map. The refinement of the completed Tat•P-TEFb structure was performed using conjugate gradient minimization, simulated annealing, and temperature factor refinement CNS protocols.47 Finally, 341 water molecules were added with the water-pick protocol of CNS. Application of zonal scaling48 and bulk solvent correction improved the geometry of the model and the quality of the maps (Supplementary Figure 11). Density was not observed for the amino acid residues 1-7, 89-96, 345 and the C-terminal His-tag of Cdk9, for the residues 1-6, 253-260 and 262-266 of Cyclin T1, and for the residues 50-86 of Tat. The broadened density and high temperature factor for Cys261 indicates to its high mobility or disorder. The partial occupancy of the Cys261 position by a histidine from a disordered portion of Tat or from a disordered His-tag of neighboring molecule can not be excluded. The refined structure of Tat•P-TEFb was used to determine the structure of Tat•P-TEFb•ATP. The difference Fourier map revealed the site of AMPPNP- and magnesium binding (Supplementary Figure 12). The refinement statistics for both structures are provided in Supplementary Table S2. The electron density for the Val190 and Arg284, only residues located in the disallowed region of Ramachandran plot, is well defined. The unfavorable main-chain conformations of Val190 and Arg284 are stabilized by intramolecular hydrogen bonds. All figures displaying the protein structures were prepared with PyMol software from Delano Scientific.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J. Lovelace and G.E. Borgstahl for maintenance and management of the Eppley Institute's X-ray Crystallography facility; D.G. Vassylyev for the zonal scaling instruction files. This work is supported by the NIH grants GM35500 and AI074392 to DHP, by Nebraska Department of Health and Human Services grant LB506 and in part by NIH grant GM082923 to THT. This work is also based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from

the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. The Eppley Institute's X-ray Crystallography facility is supported by Cancer Center Support Grant P30CA036727.

References

1. Onafuwa-Nuga A, Telesnitsky A. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. Microbiol Mol Biol Rev. 2009; 73:451–80. Table of Contents. [PubMed: 19721086]
2. Letvin NL. Virology. Moving forward in HIV vaccine development. Science. 2009; 326:1196–8. [PubMed: 19965456]
3. Zhu Y, et al. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev. 1997; 11:2622–32. [PubMed: 9334325]
4. Mancebo HS, et al. P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev. 1997; 11:2633–44. [PubMed: 9334326]
5. Garber ME, Wei P, Jones KA. HIV-1 Tat interacts with cyclin T1 to direct the P-TEFb CTD kinase complex to TAR RNA. Cold Spring Harb Symp Quant Biol. 1998; 63:371–80. [PubMed: 10384302]
6. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol Cell. 2006; 23:297–305. [PubMed: 16885020]
7. Peng J, Zhu Y, Milton JT, Price DH. Identification of multiple cyclin subunits of human P-TEFb. Genes Dev. 1998; 12:755–62. [PubMed: 9499409]
8. Marshall NF, Peng J, Xie Z, Price DH. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem. 1996; 271:27176–83. [PubMed: 8900211]
9. Fu TJ, Peng J, Lee G, Price DH, Flores O. Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. J Biol Chem. 1999; 274:34527–30. [PubMed: 10574912]
10. Chao SH, Price DH. Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. J Biol Chem. 2001; 276:31793–9. [PubMed: 11431468]
11. Zhou Q, Yik JH. The Yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation. Microbiol Mol Biol Rev. 2006; 70:646–59. [PubMed: 16959964]
12. Michels AA, Bensaude O. RNA-driven cyclin-dependent kinase regulation: When CDK9/cyclin T subunits of P-TEFb meet their ribonucleoprotein partners. Biotechnol J. 2008; 3:1022–1032. [PubMed: 18655042]
13. Sedore SC, et al. Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR. Nucleic Acids Res. 2007; 35:4347–58. [PubMed: 17576689]
14. Wimmer J, et al. Interactions between Tat and TAR and human immunodeficiency virus replication are facilitated by human cyclin T1 but not cyclins T2a or T2b. Virology. 1999; 255:182–9. [PubMed: 10049833]
15. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR. Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. Embo J. 1998; 17:7056–65. [PubMed: 9843510]
16. Fujinaga K, et al. The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. J Virol. 1998; 72:7154–9. [PubMed: 9696809]
17. Zhou M, et al. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Mol Cell Biol. 2000; 20:5077–86. [PubMed: 10866664]
18. Pavletich NP. Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. J Mol Biol. 1999; 287:821–8. [PubMed: 10222191]

19. Garber ME, et al. 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 Dev. 1998; 12:3512–27. [PubMed: 9832504]
20. Janin J. Specific versus non-specific contacts in protein crystals. Nat Struct Biol. 1997; 4:973–4. [PubMed: 9406542]
21. Baumli S, et al. The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. Embo J. 2008; 27:1907–18. [PubMed: 18566585]
22. Kuppuswamy M, Subramanian T, Srinivasan A, Chinnadurai G. Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. Nucleic Acids Res. 1989; 17:3551–61. [PubMed: 2542902]
23. Sadaie MR, Mukhopadhyaya R, Benaissa ZN, Pavlakis GN, Wong-Staal F. Conservative mutations in the putative metal-binding region of human immunodeficiency virus tat disrupt virus replication. AIDS Res Hum Retroviruses. 1990; 6:1257–63. [PubMed: 2078407]
24. Ruben S, et al. Structural and functional characterization of human immunodeficiency virus tat protein. J Virol. 1989; 63:1–8. [PubMed: 2535718]
25. Jeang KT, Xiao H, Rich EA. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. J Biol Chem. 1999; 274:28837–40. [PubMed: 10506122]
26. Rice AP, Carlotti F. Mutational analysis of the conserved cysteine-rich region of the human immunodeficiency virus type 1 Tat protein. J Virol. 1990; 64:1864–8. [PubMed: 2181156]
27. 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; 15:1287–92. [PubMed: 19029897]
28. Taube R, et al. Interactions between equine cyclin T1, Tat, and TAR are disrupted by a leucine-to-valine substitution found in human cyclin T1. J Virol. 2000; 74:892–8. [PubMed: 10623752]
29. Derse D, Newbold SH. Mutagenesis of EIAV TAT reveals structural features essential for transcriptional activation and TAR element recognition. Virology. 1993; 194:530–6. [PubMed: 8389074]
30. Ippolito JA, Steitz TA. A 1.3-A resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. Proc Natl Acad Sci U S A. 1998; 95:9819–24. [PubMed: 9707559]
31. Richter S, Ping YH, Rana TM. TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication. Proc Natl Acad Sci U S A. 2002; 99:7928–33. [PubMed: 12048247]
32. Campbell GR, Loret EP. What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine? Retrovirology. 2009; 6:50. [PubMed: 19467159]
33. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. Aids. 2006; 20:W13–23. [PubMed: 17053344]
34. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004; 14:1188–90. [PubMed: 15173120]
35. Kuiken C, et al. HIV Sequence Compendium 2009. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR 09-03280. 2009
36. D'Orso I, Frankel AD. Tat acetylation modulates assembly of a viral-host RNA-protein transcription complex. Proc Natl Acad Sci U S A. 2009; 106:3101–6. [PubMed: 19223581]
37. Barboric M, et al. Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. Nucleic Acids Res. 2007; 35:2003–12. [PubMed: 17341462]
38. Michels AA, et al. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. Mol Cell Biol. 2003; 23:4859–69. [PubMed: 12832472]
39. Yik JH, et al. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. Mol Cell. 2003; 12:971–82. [PubMed: 14580347]
40. Russo AA, Jeffrey PD, Patten AK, Massague J, Pavletich NP. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. Nature. 1996; 382:325–31. [PubMed: 8684460]

41. Li Q, et al. Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. J Biol Chem. 2005; 280:28819–26. [PubMed: 15965233]

42. Chen R, Yang Z, Zhou Q. Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. J Biol Chem. 2004; 279:4153–60. [PubMed: 14627702]

43. Wang Y, Liu XY, De Clercq E. Role of the HIV-1 positive elongation factor P-TEFb and inhibitors thereof. Mini Rev Med Chem. 2009; 9:379–85. [PubMed: 19275730]

44. Jancarik J, Pufan R, Hong C, Kim SH, Kim R. Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. Acta Crystallogr D Biol Crystallogr. 2004; 60:1670–3. [PubMed: 15333951]

45. Tahirov TH, et al. High-resolution crystals of methionine aminopeptidase from Pyrococcus furiosus obtained by water-mediated transformation. J Struct Biol. 1998; 121:68–72. [PubMed: 9573622]

46. Otwinowski Z, Minor W. Carter JCW, Sweet RM. Processing of X-ray Diffraction Data Collected in Oscillation Mode. 1997

47. Brünger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54:905–21. [PubMed: 9757107]

48. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I. Structural basis for transcription elongation by bacterial RNA polymerase. Nature. 2007; 448:157–62. [PubMed: 17581590]

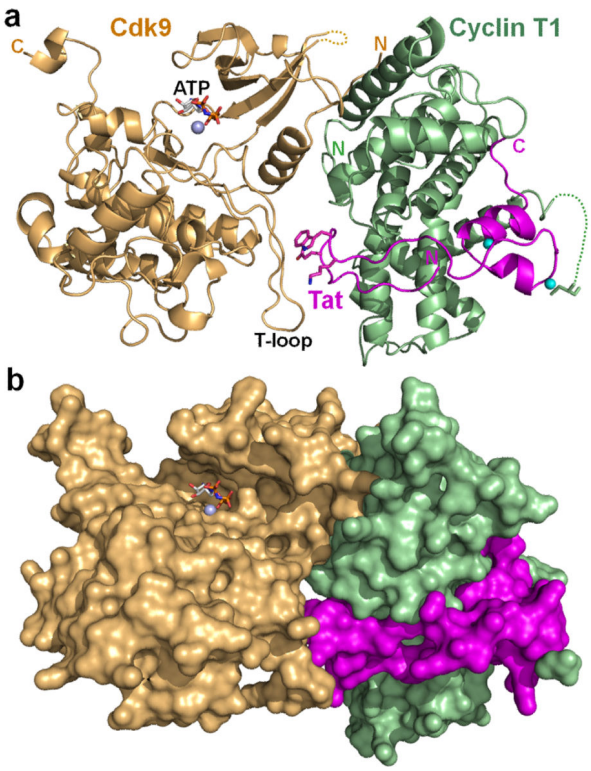

Figure 1. Overall structure of the Tat•P-TEFb•ATP

(**a**) Ribbon representation and (**b**) surface representation of the Tat•P-TEFb•ATP structure. Cdk9 is light orange, Cyclin T1 is pale green and Tat is magenta. The side chains of the Cdk9-interacting residues of Tat, Cys261 of Cyclin T1, and ATP analog are drawn as sticks, and the zinc and magnesium atoms are drawn as cyan and light blue spheres, respectively. The dashed lines represent the missing link between Lys88 and Gly of Cdk9, and between Leu252 and Cys261 of Cyclin T1.

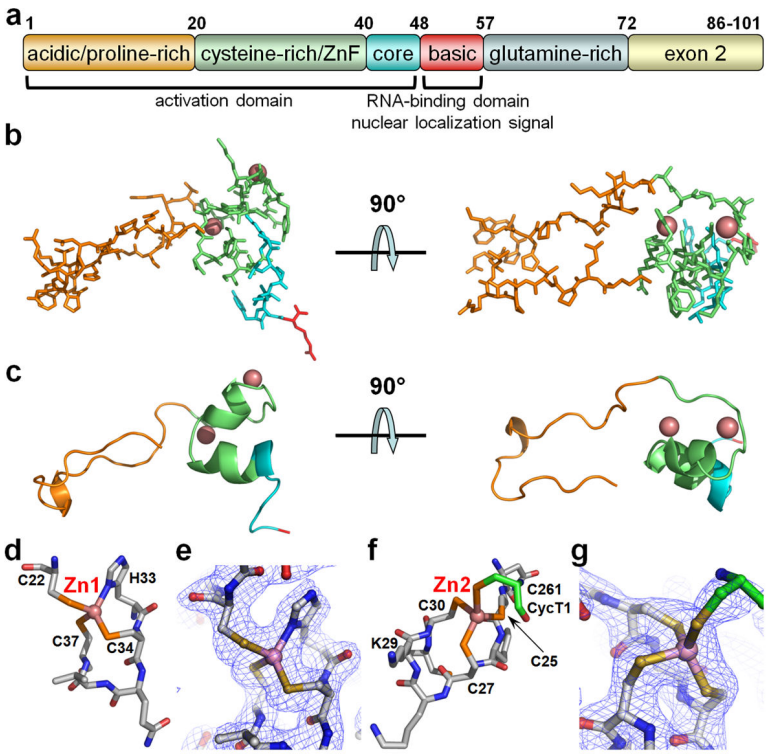

Figure 2. Structure of P-TEFb-bound Tat

(**a**) The regions and functional domains of HIV-1 Tat proteins. (**b**) Stick and (**c**) cartoon representations of the Tat structure in two orientations. The regions of Tat are colored as in panel **a**. A close-up view of Zn1 and Zn2 coordination and a portion of $2F_o$-F_c Fourier maps (mesh) contoured at 1σ are shown (**d-g**).

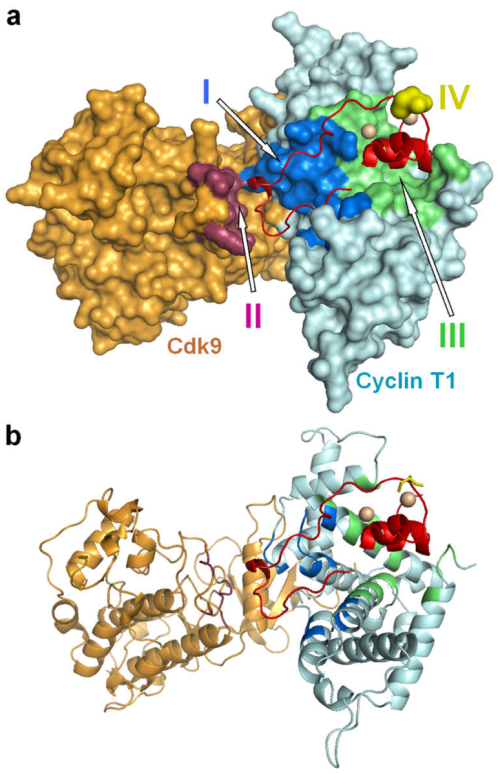

Figure 3. The Tat-interacting areas of P-TEFb

Mapping of the Tat-interacting residues of P-TEFb (**a**) to the surface of P-TEFb and (**b**) to the ribbon diagram of P-TEFb. The subunits and interacting areas are colored according to definitions shown in panel **a**.

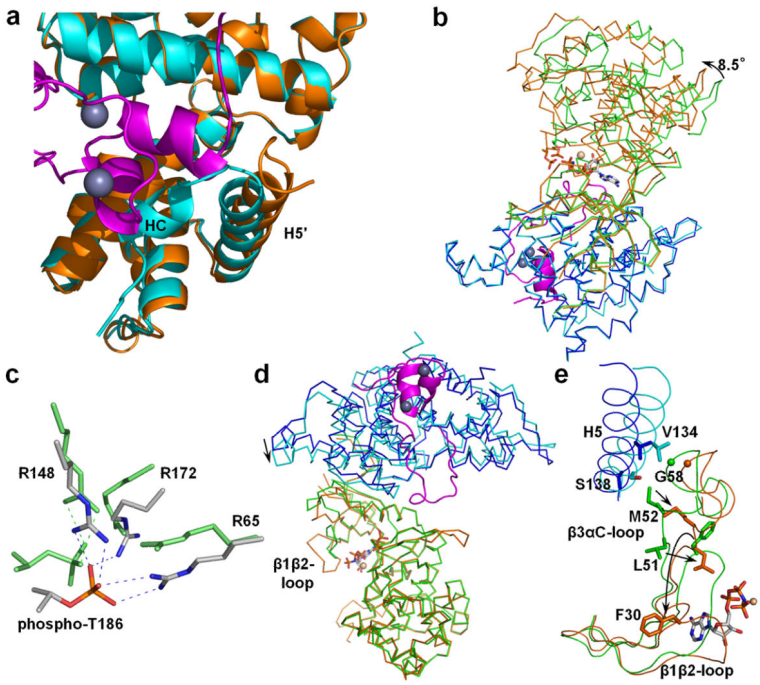

Figure 4. Comparison of the crystal structures of P-TEFb•ATP and Tat•P-TEFb•ATP

(**a**) Shift of H5′ and disordering of HC in Cyclin T1 of Tat•P-TEFb•ATP (orange). (**b**) Superimposition of Cyclin T1 showing rotation of Cdk9. P-TEFb•ATP: Cdk9 (green), Cyclin T1 (blue); Tat•P-TEFb•ATP: Cdk9 (salmon), Cyclin T1 (cyan). (**c**) Interactions of phospho-Thr186 in panel **b**. P-TEFb•ATP (green); Tat•P-TEFb•ATP (colored by atom type). (**d**) Superimposition of Cdk9 showing the shift of Cyclin T1. (**e**) Concerted movement of the Cyclin T1 H5 and the $\beta_3\alpha$C-loop of Cdk9 causing the conformational switch of the Cdk9 $\beta_1\beta_2$-loop upon Tat binding to P-TEFb. The residues causing sterical hindrance are drawn as sticks. Colors in **d** and **e** are as in **b**.

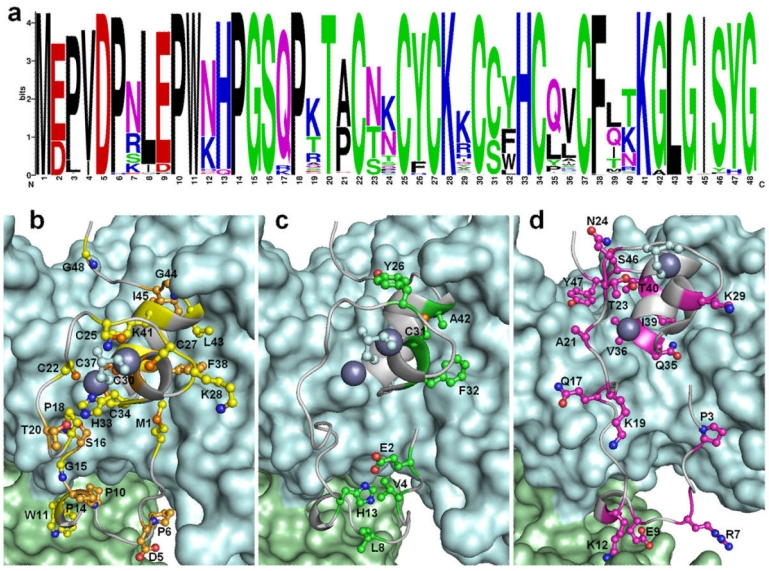

Figure 5. A conservation-dependent location of Tat's amino acid residues

(**a**) Sequence logos of the Tat activation domain. The location of Tat (**b**) invariant residues (yellow) and residues with randomly occurring polymorphs (deep salmon), (**c**) residues with predominantly functionally equivalent substitutions (green), and (**d**) residues with a variety of substitutions (magenta). In panels **b**-**d** Tat is drawn as grey ribbon, and Cyclin T1 (pale cyan) and Cdk9 (pale green) are shown in surface representation. The highlighted Tat residues and Cys261 of Cyclin T1 are represented as ball-and-stick diagrams. The zinc atoms are shown as large grey spheres.