

Cloning, High-Yield Expression in *Escherichia coli*, and Purification of Biologically Active HIV-1 Tat Protein

Thomas Kirsch,* Markus Boehm,* Oliver Schuckert,† Armin U. Metzger,*
Dieter Willbold,* Rainer W. Frank,† and Paul Rösch*

*Lehrstuhl für Biopolymere, Universität Bayreuth, D-95440 Bayreuth, Germany; and †Zentrum für Molekularbiologie Heidelberg, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany

Received December 12, 1995, and in revised form March 27, 1996

We have established an expression system for full-length HIV-1 transactivator (Tat) protein in *Escherichia coli*. By constructing a synthetic gene for high level expression in enteric bacteria, the recombinant protein can be obtained in high yield. Fusion of the Tat sequence to an N-terminal histidine tag allows the rapid purification of the fusion protein through a single chromatographic step. After cleavage of the fusion protein with CNBr, pure Tat can be obtained through the use of a MonoS column. Reduction of the protein with Tris(2-carboxyethyl)phosphine-HCl and subsequent stepwise refolding yields biologically active Tat. Sample purity and the identity of the protein mass with the mass expected from the amino acid sequence was demonstrated by mass spectrometry. Nuclear magnetic resonance spectroscopy showed the identity of bacterially expressed and chemically synthesized Tat protein (P. Bayer *et al.*, 1995, *J. Mol. Biol.* 247, 529–535). The expression of Tat in *E. coli* enables isotope labeling as a prerequisite for multidimensional NMR experiments toward the elucidation of the structure of the Tat–*trans*-activation response element complex.

© 1996 Academic Press, Inc.

Human immunodeficiency virus type 1 (HIV-1) transactivator (Tat) protein (1,2) upregulates viral production by three orders of magnitude (3). The *tat* gene consists of two exons, the first coding for 72 residues, the second coding for an additional 10 to 29 residues. Tat can be divided into five distinct sequence domains: the N-terminal acidic, cysteine-rich, core, basic, and C-terminal glutamine-rich sequences. The structure of the Tat protein from equine infectious anemia virus (EIAV) has been determined by two-dimensional NMR spectroscopy and shows, except for a rigid core region, a highly flexible

protein, whose basic region is weakly helical (4). The NMR structure of HIV-1 Tat exhibits a well defined hydrophobic core and a similarly well defined glutamine rich domain which sandwich the N-terminal residues from Val4 to Pro14. The cysteine rich and the basic region are highly flexible, and the basic region shows no obvious helical tendency. The Arg–Gly–Asp (RGD) loop of the C-terminal region is structured as a rigid solvent exposed hairpin (5). Tat is unique among transcriptional activators as its *cis*-regulatory target is an RNA element (*trans*-activation response element, TAR). TAR is located just downstream of the long terminal repeat (LTR) promoter (6) and forms a hairpin structure that spans nucleotides +1 to +59 of all transcripts. The minimal functional TAR sequence (+19 to +42) includes the upper stem, a uracil-rich trinucleotide bulge, and a structured hexanucleotide loop. TAR is bound by Tat principally through electrostatic interactions between the basic domain of Tat and negatively charged phosphates surrounding the bulge. Binding is further stabilized by hydrogen-bonding interactions in the major groove of the RNA (7). Several cellular proteins are involved in the function of the Tat–TAR complex, for example a loop binding protein and a Tat binding coactivator (8). Whereas in the absence of Tat ~90% of the transcripts terminate around position +59, in the presence of Tat more than 99% of the transcripts are of full length. The Tat protein thus functions as a transcriptional anti-terminator (6).

It has been shown that mice transgene for *tat* develop dermal lesions resembling Kaposi's sarcoma (9) which in an aggressive form is frequently associated with HIV-1 infection. Further studies confirmed that Tat is released by infected human cells and in synergy with basic fibroblast growth factor induces Kaposi's sarcoma (10).

Tat proteins are considered a major target for antiviral therapies. Thus, determining the three-dimensional structure of this protein and its various protein and

RNA complexes is of utmost importance. So far, spectroscopic studies on Tat proteins employed either non-HIV proteins such as the equine infectious anemia virus Tat protein (4,11), chemically synthesized HIV-1 Tat protein (5), or peptide analogs (12–14). For more extensive studies demanding high amounts of protein, and eventually isotopic labeling of the protein for nuclear magnetic resonance spectroscopy, it is highly desirable to employ bacterially produced protein from a high yield expression and purification system.

The preparation of Tat protein has been hampered by its high cysteine content and its susceptibility to oxidation leading to the formation of stable multimers (15). Some reports on the preparation of recombinant Tat restrict on first exon Tat (16,17), lacking the RGD sequence that is coded by the second exon. A model for the cellular adhesion of Tat employs the binding of Tat to integrins via this sequence (18), resulting in cellular uptake of the protein and transformation of endothelial cells to Kaposi's sarcoma cells (19). Other purification protocols give only relatively low yields (20). We report high-yield purification of biologically active full-length Tat protein from a synthetic gene. Additionally, we show that the prepared protein is suitable for further structural investigations by two- and multidimensional NMR spectroscopy which should improve our knowledge of the atomic details of Tat action.

MATERIALS AND METHODS

Chemicals

Unless otherwise indicated, chemicals were obtained from Merck (Darmstadt).

Construction of the Tat Expression System

The Tat amino acid sequence was HIV-1 Zaire 2 (H1Z2) except that an extra leucine residue was introduced after Met1, as Met1 is removed during preparation. A synthetic gene was designed based on the following considerations: First, codon usage was adapted to that for highly expressed genes of enteric bacteria. Second, it was attempted to minimize mRNA secondary structure formation. The gene was divided into four building blocks (Fig. 1). The commercially obtained oligonucleotides (MWG, München) differed in length from 57 to 74 nucleotides with overlaps between 8 and 12 nucleotides to enhance ligation efficiency.

Portions (1 nmol) of the oligonucleotides (except Hi1 and Hi4) were phosphorylated with 10 nmol ATP and 10 units of T4 polynucleotide kinase (USB, Cleveland, OH) in 19 μ l of kinase buffer for 45 min at 37°C. The reaction was stopped by heating to 65°C for 45 min. For hybridization, 200 pmol of complementary oligonucleotides in 20 μ l ligase buffer were heated to 90°C for 5

min and cooled to room temperature in 5 h. The parallel annealing reactions were pooled and adjusted to 100 μ l with 10 mM ATP and 10 units of T4 ligase (USB). Ligation was performed for 16 h at 14°C. The *tat* gene was purified by 1.5% agarose gel electrophoresis, excised and subsequently purified with a QIAquick kit (QUIAGEN, Hilden).

Cloning of the Synthetic Gene

The *tat* gene was ligated with agarose gel-purified *Nde*I/*Esp*I-cut pET16b (Novagen, Madison, WI). This ligation mixture was used to transform 200- μ l portions of *E. coli* strain DH5 α cells, made competent by the standard CaCl₂ method. Ampicillin-resistant clones were picked and grown overnight in LB medium/ampicillin. Plasmid DNA was prepared and restricted with *Sac*II and *Mlu*I to screen for recombinant plasmids. The recombinant plasmid was termed pTK1 and sequenced by the alkali-denatured plasmid chain-termination method using T7 promoter and terminator primers and T7 DNA polymerase (Pharmacia, Freiburg). pTK1 was transformed into *Escherichia coli* cells of strain BL21(DE3) plysS for expression.

Expression

Fermentation of *E. coli* BL21(DE3) plysS pTK1 was performed at 37°C in a 7-liter working volume of LB medium containing ampicillin at 50 mg/liter and chloramphenicol at 34 mg/liter. The medium was inoculated with 150 ml overnight culture. The culture was aerated with 3 liters/min and stirred with 300 rpm until the OD₆₀₀ reached 0.3. Above OD₆₀₀ 0.3 the aeration was increased to 7 liters/min and the stirring to 700 rpm. When the OD₆₀₀ reached 0.8 to 1.0 Tat expression was induced with isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 3 mM. Following induction, cells were grown for 4 h, harvested by centrifugation at 6000 rpm for 15 min at 4°C and stored at -20°C until needed.

To test whether Tat was expressed in minimal medium with 20 mM glucose, 100 ml medium supplemented with 50 mg/liter ampicillin and 34 mg/liter chloramphenicol were inoculated with 4 ml overnight culture of *E. coli* BL21 (DE3) plysS pTK1 and incubated at 37°C. At an OD₆₀₀ of 1.0 a sample of noninduced cells was taken and the expression of Tat was induced by adding IPTG to a final concentration of 1 mM. Samples were taken after 30, 90, and 150 min. Cells were pelleted by brief centrifugation, 50 μ l Laemmli buffer was added, and the cell suspension was heated to 95°C for 10 min. Aliquots of 10 μ l were analyzed by denaturing gel electrophoresis.

Protein Purification

All purification steps were carried out at room temperature, unless otherwise specified. Frozen cells were

a

```

NdeI                                                                                      .49
<-----Hi1----->
T ATG CTC GAC CCG GTC GAC CCG AAC ATC GAA CCG TGG AAC CAC CCC GGG
  AC GAG CTG GGC CAG CTG GGC TTG TAG CTT GGC ACC TTG GTG GGG CCC
  <-----Hic1----->
Met Leu Asp Pro Val Asp Pro Asn Ile Glu Pro Trp Asn His Pro Gly
  1                               10

                                                                                      .100
-----><-----
TCT CAA CCG AAA ACT GCT TGC AAC CGT TGC CAC TGC AAG AAA TGC TGC TAC
AGA GTT GGC TTT TGA CGA ACG TTG GCA ACG GTG ACG TTC TTT ACG ACG ATG
-----><-----Hic2-----
Ser Gln Pro Lys Thr Ala Cys Asn Arg Cys His Cys Lys Lys Cys Cys Tyr
  20                               30

                                                                                      .151
-----Hi2-----><-----
CAC TGC CAA GTT TGC TTT ATC ACT AAA GGT CTG GGA ATT TCT TAC GGT CGT
GTG ACG GTT CAA ACG AAA TAG TGA TTT CCA GAC CCT TAA AGA ATG CCG GCA
-----><-----
His Cys Gln Val Cys Phe Ile Thr Lys Gly Leu Gly Ile Ser Tyr Gly Arg
  40

                                                                                      .202
-----Hi3-----><-----
AAG AAA CGT CGT CAG CGT CGT CGT CCG TCT CAG GGT GGT CAA ACT CAC CAA
TTC TTT GCA GCA GTC GCA GCA GCA GGC AGA GTC CCA CCA GTT TGA GTG GTT
-----Hic3-----
Lys Lys Arg Arg Gln Arg Arg Arg Pro Ser Gln Gly Gly Gln Thr His Gln
  50                               60

                                                                                      .253
-----><-----SacII-----><-----Hi4-----><-----
GAC CCG ATC CCG AAA CAA CCG TCT TCT CAA CCC CGC GGT GAC CCG ACT GGT
CTG GGC TAG GGC TTT GTT GGC AGA AGA GTT GGG GCG CCA CTG GGC TGA CCA
--> <-----Hic4-----><-----
Asp Pro Ile Pro Lys Gln Pro Ser Ser Gln Pro Arg Gly Asp Pro Thr Gly
  70                               80

.260          EspI
----->
CCG AAA GAA TAA TAA GC
GGC TTT CTT ATT ATT CGACT
----->
Pro Lys Glu End End
  86

```

FIG. 1a. Nucleotide sequence of the synthetic *tat* gene. The single oligonucleotides used to build the gene, Hi1 to Hi4 and Hic1 to Hic4, are shown by broken lines. Nucleotide numbers (underlined) are indicated on the right-hand side above the double strand. The corresponding protein sequence is numbered below and starting at the amino-terminal leucine after CNBr cleavage. The *NdeI* and *EspI* sites are indicated above the double strand. The *SacII* site is also indicated.

suspended in 4 ml/g wet cell weight buffer A (6 M urea, 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5 mM imidazole) and stirred for 48 h at 4°C. Insoluble cell debris

was removed by two centrifugation steps: first at 16,500 rpm, 4°C, for 30 min, and second at 40,000 rpm, 4°C, for 30 min. The supernatant was sterile filtered

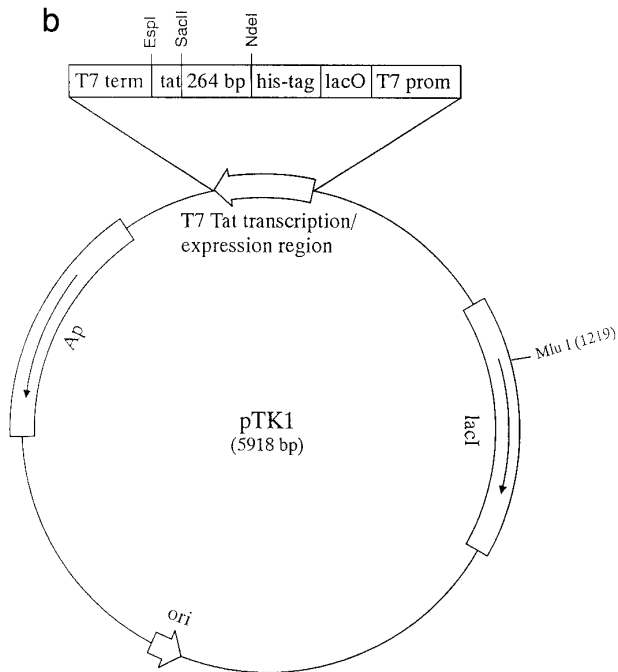


FIG. 1b. The Tat expression system based on the vector pTK16b (Novagen, Madison, WI). The synthetic *tat* gene was cloned into the *NdeI* and *EspI* sites of pTK16b and is under the control of the T7 promoter and *lacO*-Operator. Tat expression is induced by the addition of IPTG.

(diameter 0.2 μm) and applied to a 25 ml chelating Sepharose column (Pharmacia) loaded with 100 ml 50 mM NiSO_4 and preequilibrated with buffer A. The column was washed with four column volumes of buffer A and eight volumes of buffer A with 150 mM imidazole. The fusion protein was eluted with two column volumes of buffer A with 1 M imidazole. Tat containing fractions were pooled and the pH was adjusted to 7.0 by adding 1 M HCl. Then the solution was dialyzed against 5 liters

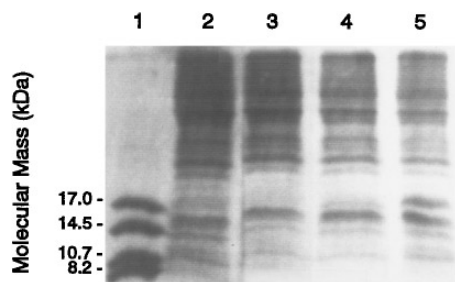


FIG. 2. Expression of the his-tag fusion Tat protein in minimal medium containing 20 mM glucose. Lane 1, molecular weight standards; lane 2, total cell lysate without IPTG; lanes 3, 4, 5, total cell lysate 30, 90, and 150 min after IPTG addition, respectively.

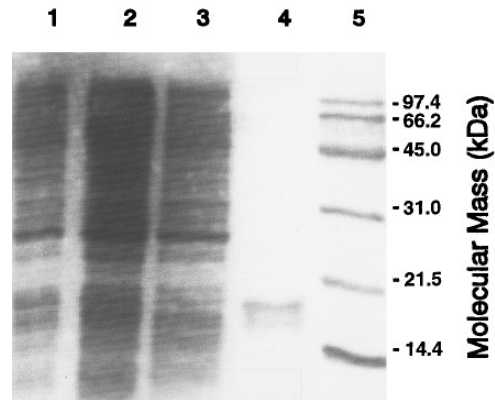


FIG. 3. Purification of the his-tag fusion Tat protein on metal chelate column. The gel contained 20% polyacrylamide-SDS. Lane 1, flow through; lane 2, wash-step with 5 mM imidazole; lane 3, wash step with 150 mM imidazole; lane 4, elution of fusion protein; lane 5, molecular weight standards.

of 0.1 M HCl overnight and lyophilized. The protein lyophilisate was dissolved in 5 ml 1 mM HCl and further desalted on a FastDesalting column (Pharmacia) fitted to an HPLC system. HCl (1 mM) was used as an eluent, and fractions showing absorption at 280 nm were collected, pooled, and lyophilized.

The lyophilized protein was dissolved in 70% (v/v) formic acid, and solid CNBr was added to a final concentration of 20% (w/v). The cleavage reaction was allowed to proceed for 24 h in the dark under vigorous shaking. The mixture was lyophilized to remove remaining CNBr and volatile cleavage products, redissolved in 100 mM NaOAc, pH 5.0, 200 mM NaCl (buffer B), and loaded on a MonoS HR 5/5 column (Pharmacia) preequilibrated in buffer B. A gradient from 200 mM to 1.5 M NaCl in the same buffer was applied. Recombinant Tat protein eluted in fractions containing 1.045 M NaCl. Elution fractions were pooled and lyophilized.

Reduction of Disulfide Bonds and Refolding

The lyophilisate was dissolved in 6 M urea, 1 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$, pH 6.0, 2 mM NaCl, and solid Tris(2-carboxyethyl)phosphine-HCl (TCEP-HCl) was added to get a final molar ratio of TCEP-HCl to Tat of 25 to 1. Reduction of disulfide bonds was carried out at room temperature in the dark for 8 h.

Then, the mixture was dialyzed for 12 h each against:

- (1) 4 M urea, 500 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$, pH 6.3, 1 mM NaCl
- (2) 2M urea, 300 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$, pH 6.3, 600 mM NaCl
- (3) 100 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$, pH 6.3, 200 mM NaCl

TABLE 1
Summary of Purification of Tat Protein

Fraction	Volume (ml)	Total protein ^a (mg)	Tat (μ mol)	Yield (%)
1. Crude extract	92	1760	—	—
2. Ni ²⁺ -column	40	115	9.11 (His-Tat)	100
3. MonoS	8	23	2.26	24.8
4. Refolding	9	18.4	1.82	20.0

^a Measured by the method of Smith *et al.* (26).

(4) for an additional 6 h against 100 mM KH₂PO₄/K₂HPO₄, pH 6.3, 200 mM NaCl.

Thereafter, the dialysate was lyophilized and stored at -20°C until further use.

Characterization of Recombinant HIV-1 Tat

Gel electrophoresis. To monitor the stepwise purification and final preparation grade of the Tat protein electrophoresis of 10- μ l aliquots on discontinuous 20% polyacrylamide-SDS gels was performed with water cooling to avoid temperature gradients and a constant current of 20 mA according to Laemmli (18) with a Hoefer, Inc., MightySmall vertical gel system. The samples were heated with 10 μ l sample buffer (40% (w/v) sucrose, 3% (w/v) SDS, 3% (v/v) 2-mercaptoethanol, 0.075% (w/v) bromphenol blue) at 95°C for 5 min. After 1 h electrophoresis the gels were stained with staining solution (0.05% (w/v) Coomassie brilliant blue G-250, 45% (v/v) methanol, 9.2% (v/v) acetic acid) and destained with destaining solution (26% (v/v) methanol, 1% (v/v) acetic acid) twice for 30 min. Gels were documented with the Cybertech CS-1 Image Documentation System (Cybertech, Berlin).

MALDI-MS. A trace amount of protein was dissolved in H₂O and subjected to matrix-assisted laser desorption-time of flight (MALDI-TOF) mass spectrometric analysis on a Shimadzu Kompact MALDI II mass spectrometer with sinapinic acid as matrix.

N-terminal sequencing. Lyophilized Tat-fusion protein (500 pmol) was dissolved in 10 μ l 0.1% (v/v) trifluoroacetic acid and spotted on a Polybrene-coated glass filter disc for sequence analysis. Sequencing of the first 30 amino acid residues was done on a Model 476A sequencer (Applied Biosystems, Foster City, CA).

¹H-NMR spectroscopy. Lyophilized protein was dissolved in 90% H₂O/10% (v/v) D₂O to yield a final concentration of 1.2 mM. Concentrations of KH₂PO₄/K₂HPO₄ and NaCl were 10 and 20 mM, respectively. The pH was 6.3. Spectra were recorded on a Bruker AMX600 spectrometer at 298 K.

2D-NMR spectroscopy. Lyophilized protein was dissolved in 90% H₂O/10% (v/v) D₂O to yield a final concentration of 1.6 mM in buffer containing 10 mM KH₂PO₄/K₂HPO₄, pH 6.3, 20 mM NaCl. A total correlation spectroscopy experiment with suppression of NOESY-type cross peaks (Clean-TOCSY, 23), and a mixing time of 80 ms was performed on a Bruker AMX600 spectrometer at 298 K.

Activity assays. A reporter construct containing the *P. pyralis* luciferase gene under the control of the HIV-1 NL4-3 LTR was kindly provided by Dr. Georg Sczakiel (Heidelberg). The plasmid was transfected into HeLa cells grown in Dulbecco's modified Eagle's medium (DMEM) with 10% FCS in 60-mm dishes by the DEAE-dextran method (24). Twenty-four hours after transfection cells were washed with phosphate-buffered saline (PBS), and fresh medium containing 200 nM of purified Tat protein in a final volume of 3 ml was added. Chloroquine (100 μ M) was included to enhance transactivation. Luciferase activity was measured 24 h later using the Promega Luciferase Assay System (Serva, Heidelberg) and a Berthold Lumat (Berthold, Bad Wildbad) according to the manufacturers instructions.

RESULTS AND DISCUSSION

Fermentation

The fermentation of the expression strain *E. coli* BL21(DE3) plysS pTK1 in 7 liters LB medium supplemented with 50 mg/liter ampicillin and 34 mg/liter chloramphenicol and inoculated with 150 ml overnight culture resulted in rapid growth to an OD₆₀₀ 0.8 within 2.5–3 h. After induction with IPTG, the culture grew to an OD₆₀₀ 2.8–3.5 within 3 additional h. Cells were harvested after 3 h because there was no further increase in cell growth and Tat expression reached a plateau (data not shown). We obtained reproducibly about 21–24 g of whole cells with our 7-liter fermentor design.

The test for expression of Tat in minimal medium with glucose as sole carbon source was positive (Fig.

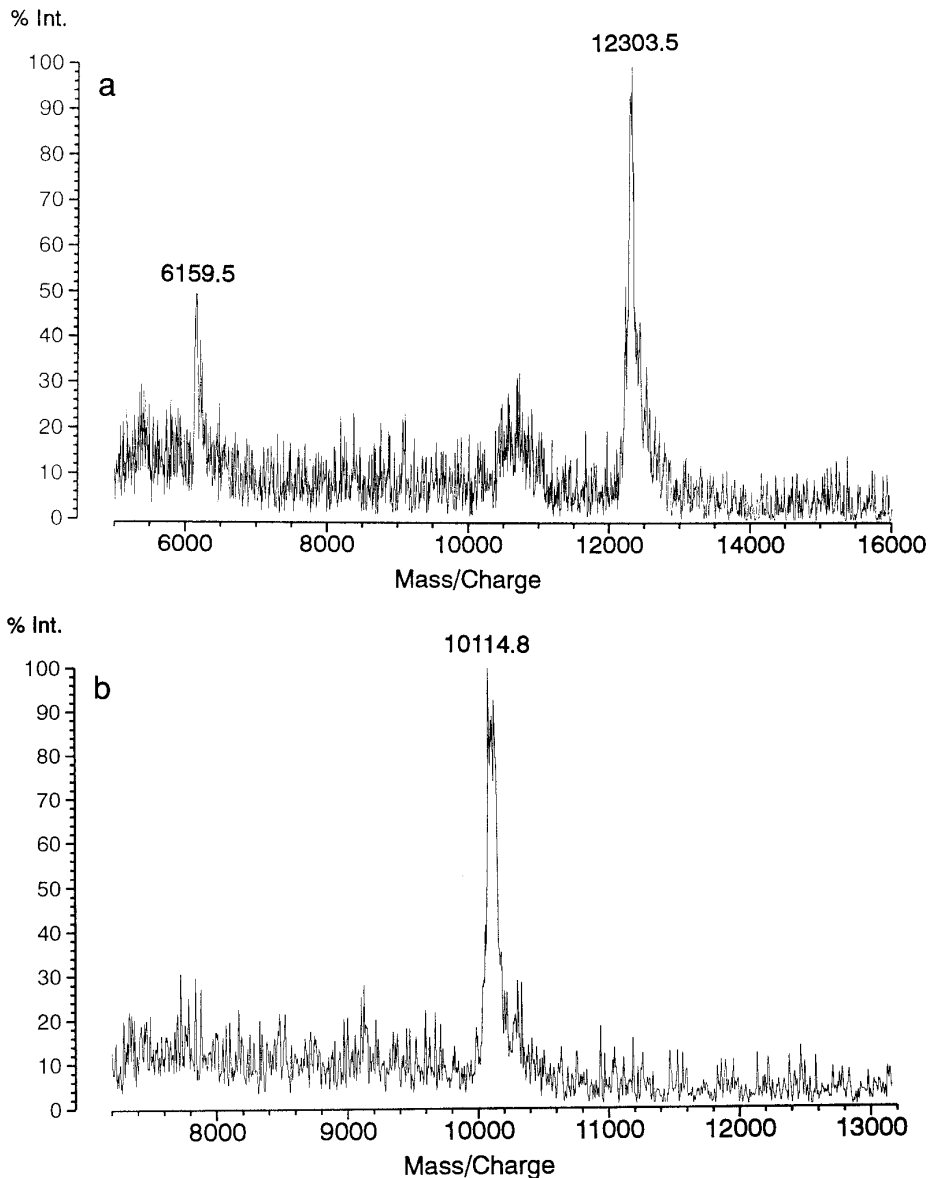


FIG. 4. MALDI-MS spectra of fusion protein (a) and Tat (b). The matrix was sinapinic acid and spectra were recorded on a Shimadzu Kompact MALDI II mass spectrometer. The peak with half mass/charge ratio represents the double charged protein.

2). After 90 min a band on the SDS-PAGE gel representing the fusion protein was visible and expression reached a maximum 180 min after induction of the culture. The high yield expression in minimal medium enables us to prepare uniformly ^{13}C - and ^{15}N -labeled protein samples for heteronuclear NMR experiments which will help to resolve ambiguities in the spectrum due to frequency degeneracy. Especially the assignments for the cysteine-rich and arginine-rich basic region might be improved by using three-dimensional NMR techniques. In addition, la-

beled Tat protein will facilitate the interpretation of Tat-TAR spectra.

Purification of Tat Protein

The fusion of a tag of 10 histidine residues to the Tat sequence allows the rapid purification of the fusion protein with immobilized metal ion affinity chromatography. Cellular proteins are not or only weakly bound by the Ni^{2+} equilibrated chelating Sepharose column. They can easily be removed by

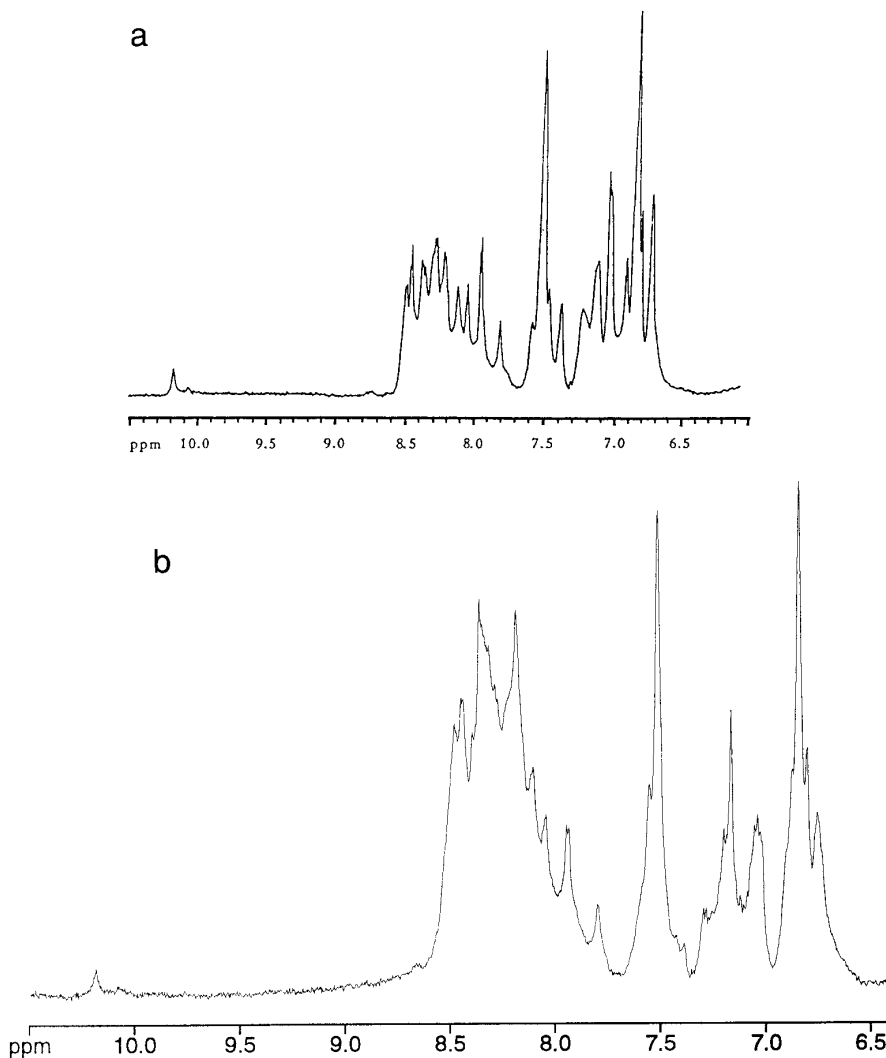


FIG. 5. Amide proton region of the ^1H -NMR spectra of synthetic (a) and recombinant (b) Tat protein. Spectra were recorded on a Bruker AMX600 spectrometer at 293 K in 10% (v/v) D_2O .

washing with buffer containing 5 or 100 mM imidazole (Fig. 3). The fusion protein is strongly bound and elutes at 415 mM imidazole (data not shown). About 5 mg fusion protein is obtained from 1 g cells. Because the Tat protein contains no internal methionine residues, the histidine tag can be removed by cleavage with CNBr. The histidine tag and uncleaved fusion protein cannot be separated from the Tat protein on a Ni^{2+} containing chelating sepharose column, as it is possible for EIAV Tat (4). This is probably due to the seven cysteine residues anchoring the Tat protein to the column. In fact, a purification scheme for recombinant Tat protein without histidine tag has been introduced by Slice *et al.* (25), in which a metal chelate column charged with Zn^{2+} is

applied. This, however, does not result in the preparation of pure Tat protein in a single step, and the authors describe the use of two additional columns (MonoS and MonoQ) to obtain Tat free from other protein contaminants. Unfortunately, no overall yield is given in the cited reference, and thus a comparison of the efficiency of the two purification strategies is not possible.

The MonoS column as strong cation exchanger allows the separation of histidine tag, fusion protein, and Tat protein based on their different charge at pH 5.0. We obtained about 1 mg Tat protein from 1 g cells. The purity of full-length monomeric Tat is estimated to be greater than 98% as judged by densitometry of gel bands with the Cybertech CS-1 system.

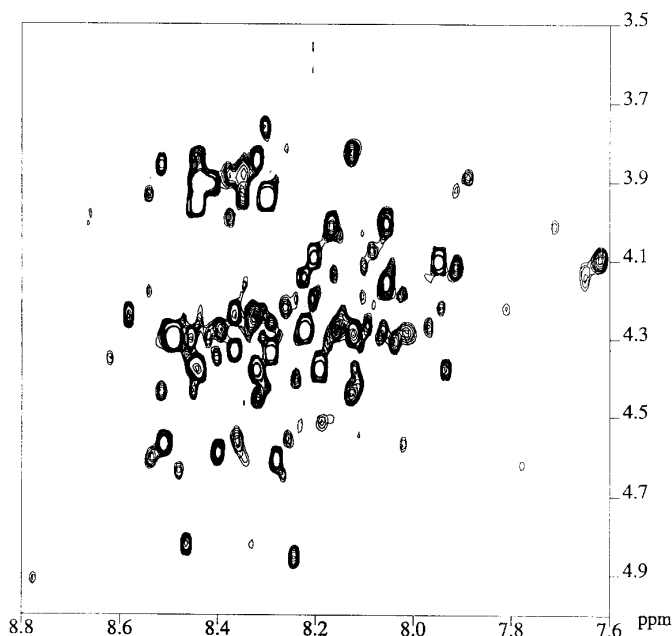


FIG. 6. Fingerprint region of the 80-ms Clean-TOCSY spectrum of a 1.6 mM Tat solution in 10% (v/v) D_2O and 10 mM potassium phosphate buffer, pH 6.3, 20 mM NaCl recorded on a 600 MHz Bruker AMX600 spectrometer. The spectrum shows the excellent resolution of $NH-C\alpha H$ -crosspeaks which enabled sequence-specific assignments of spin systems.

Reduction of Disulfide Bonds and Refolding

Before refolding, the seven cysteine residues of Tat are totally reduced with TCEP-HCl which is a stronger reducing agent than dithiothreitol (DTT). TCEP-HCl offers the additional advantage that it cannot form covalent adducts with Tat, whereas covalent adducts with DTT were observed by MALDI-MS (data not shown).

Refolding the diluted protein by stepwise reduction of the urea concentration avoids the formation of aggregates. The fact that no reducing agent is added during the refolding steps results in the formation of two intramolecular disulfide bonds, as judged by Ellman's assay. No structural data on HIV-1 Tat are available concerning the nature of the cysteine residues involved. A comparison with EIAV Tat, however, indicates that one of the disulfide bonds in HIV-1 Tat could be formed between the two cysteine residues in the core domain, i.e., Cys 34 and Cys 37. Furthermore, transfection assays show that complete reduction of partially oxidized protein results in a decrease in the transactivating ability of Tat (27). Whether Tat protein that is secreted by infected lymphocytes *in vivo* contains disulfide bonds, however, remains to be answered. The total

recovery of Tat after refolding is 18.4 mg with 115 mg fusion protein as starting material. The results of a typical purification are summarized in Table 1.

Characterization of Recombinant HIV-1 Tat

Before and after cleavage with CNBr, the molecular mass of the fusion protein and the Tat protein, respectively, was determined by MALDI-MS. The molecular mass of the fusion protein is measured to be 12,303 Da compared to 12,626 Da as expected from the amino acid sequence, and the molecular mass of Tat is 10,114 Da which is in excellent agreement with the theoretical value of 10,108 Da (Fig. 4). MALDI-MS thus confirms that the protein

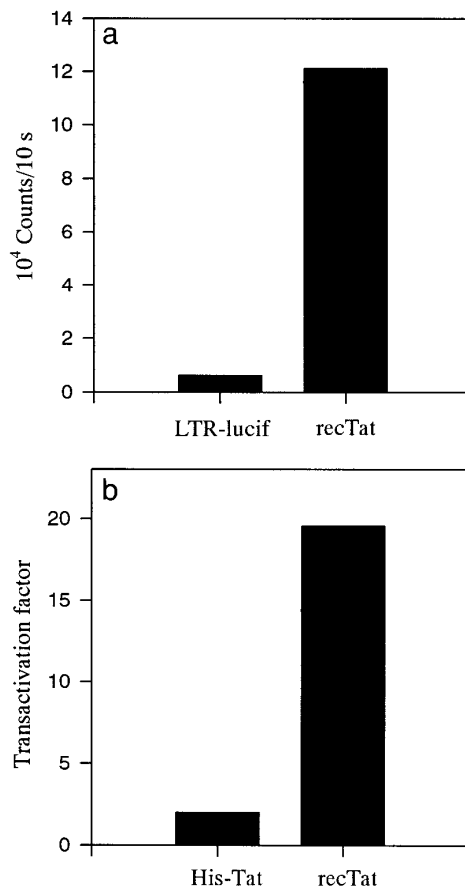


FIG. 7. Biological activity of the recombinant Tat protein. The transactivating function of the purified protein was assessed by the luciferase activity assay. (a) Absolute counts/10 s for LTR-luciferase reporter plasmid alone and 24 h after addition of Tat to the growth medium of LTR-luciferase transfected HeLa cells. (b) Transactivation factor as ratio of counts/10 s with protein to counts/10 s with LTR-luciferase alone for his-tag fusion and Tat protein.

is identical with Tat. The mass spectra show that the proteins are very pure and uniformly monomeric. Sequencing of the N-terminal 30-amino-acid residues of the Tat fusion protein further assures the identity of the purified protein with Tat.

The ¹H-NMR spectrum of reduced and refolded Tat protein (Fig. 5) is largely identical with that of reduced synthetic Tat protein used for initial structural studies (5) providing additional evidence that the prepared recombinant protein is in fact Tat. Differences in the amide proton and aromatic regions of the 1D-spectrum could be due to structural deviations in the two proteins and minor differences in pH values. These might for example be caused by the presence of two disulfide bonds in the recombinant protein. Synthetic HIV-1 Tat, on the other hand, was completely reduced after synthesis, and was measured under reducing conditions.

The Clean-TOCSY NMR spectrum shows an excellent resolution of NH-C_αH crosspeaks (Fig. 6) in the fingerprint region. The high quality of the 2D-NMR spectra enabled us to perform sequence specific assignments of the spin systems as a prerequisite for the calculation a more detailed 3D-structure of HIV-1 Tat protein (data not shown).

Biological activity of purified Tat protein was measured in a cell-based assay in which Tat induces the expression of luciferase placed under control of the HIV-1 LTR in HeLa cells. In the absence of Tat, HeLa cells expressed very low basal luciferase activity. The *in vitro* assay shows 20-fold transactivation 24 h after addition of Tat to the medium (Fig. 7). Gentz *et al.* (16) obtained a comparable 12- to 35-fold elevation with LTR chloramphenicol acetyltransferase (CAT) as reporter gene.

REFERENCES

- Arya, S. K., Guo, C., Josephs, S. F., and Wong-Staal, F. (1985) Trans-activator gene of human T-lymphotropic virus type III. *Science* 229, 69–73.
- Sodroski, J., Patarca, R., and Rosen, C. (1985) Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. *Science* 229, 74–77.
- Feinberg, M. B., Baltimore, D., and Frankel, A. D. (1991) The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proc. Natl. Acad. Sci. USA* 88, 4045–4049.
- Willbold, D., Rosin-Arbesfeld, R., Sticht, H., Frank, R., and Rösch, P. (1994) Structure of the equine infectious anemia virus tat protein. *Science* 264, 1584–1587.
- Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., and Rösch, P. (1995) Structural studies of HIV-1 Tat protein. *J. Mol. Biol.* 247, 529–535.
- Kao, S.-Y., Calman, A. F., Luciw, P. A., and Peterlin, B. M. (1987) Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330, 489–493.
- Gait, M. J., and Karn, J. (1993) RNA recognition by the human immunodeficiency virus Tat and Rev proteins. *Trends Biochem. Sci.* 18, 255–259.
- Jones, K. A., and Peterlin, B. M. (1994) Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* 63, 717–743.
- Vogel, J., Hinrichs, S. H., Reynolds, R. K., Luciw, P. A., and Jay, G. (1988) The HIV tat gene induces dermal lesions resembling Kaposi's sarcoma in transgenic mice. *Nature* 335, 606–611.
- Ensoli, B., Gendelman, R., Markham, P., Fiorelli, V., Colombini, S., Raffeld, M., Cafaro, A., Chang, H.-K., Brady, J. N., and Gallo, R. C. (1994) Synergy between basic fibroblast growth factor and HIV-1 Tat protein in induction of Kaposi's sarcoma. *Nature* 371, 674–680.
- Sticht, H., Willbold, D., Ejchart, A., Rosin-Arbesfeld, R., Yaniv, A., Gazit, A., and Rösch, P. (1994) Trifluoroethanol stabilizes a helix–turn–helix motif in equine infectious-anemia-virus trans-activator protein. *Eur. J. Biochem.* 225, 855–861.
- Loret, E. P., Vives, E., Ho, P. S., Rochat, H., van Rietschoten, J., and Johnson, W. C. (1991) Activating region of HIV-1 Tat protein: Vacuum UV circular dichroism and energy minimization. *Biochemistry* 30, 6013–6023.
- Loret, E. P., Georgel, P., Johnson, W. C., and Ho, P. S. (1992) Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator. *Proc. Natl. Acad. Sci. USA* 89, 9734–9738.
- Mujeeb, A., Bishop, K., Peterlin, B. M., Turck, C., Parslow, T. G., and James, T. L. (1994) NMR structure of a biologically active peptide containing the RNA-binding domain of human immunodeficiency virus type 1 Tat. *Proc. Natl. Acad. Sci. USA* 91, 8248–8252.
- Frankel, A. D., Bredt, D. S., and Pabo, C. O. (1988) Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science* 240, 70–73.
- Gentz, R., Chen, C.-H., and Rosen, C. A. (1989) Bioassay for trans-activation using purified human immunodeficiency virus tat-encoded protein: Trans-activation requires mRNA synthesis. *Proc. Natl. Acad. Sci. USA* 86, 821–824.
- McKenna, M.-C., Murchardt, C., Gaynor, R., and Eisenberg, D. (1994) Preparative scale culture of *Escherichia coli* cells expressing the human immunodeficiency virus type 1 Tat protein. *Protein Expression Purif.* 5, 105–111.
- Brake, D. A., Debouck, C., and Biesecker, G. (1990) Identification of an Arg–Gly–Asp (RGD) cell adhesion site in human immunodeficiency virus type 1 transactivation protein, tat. *J. Cell Biol.* 111, 1275–1281.
- Ensoli, B., Barillari, G., Salahuddin, S. Z., Gallo, R. C., and Wong-Staal, F. (1990) Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* 345, 84–86.
- Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A., and Valerio, R. (1989) Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA *in vitro*. *Proc. Natl. Acad. Sci. USA* 86, 6925–6929.
- Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, P. J. G., Gait, M. J., and Karn, J. (1993) 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.* 230, 90–110.

22. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
23. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) Clean TOCSY for ¹H spin system identification in macromolecules. *J. Am. Chem. Soc.* 110, 7870–7872.
24. Gorman, C. (1985) High efficiency gene transfer into mammalian cells, in "DNA cloning" (Glover, D. M., Ed.), Vol. 2, pp. 143–190, IRL Press, Washington, DC.
25. Slice, L. W., Codner, E., Antelman, D., Holly, M., Wegrzynski, B., Wang, J., Toome, V., Hsu, M.-C., and Nalin, C. M. (1992) Characterization of recombinant HIV-1 Tat and its interaction with TAR RNA. *Biochemistry* 31, 12062–12068.
26. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150, 76–85.
27. Koken, S. E. C., Greijer, A. E., Verhoef, K., van Wamel, J., Bukrinskaya, A. G., and Berkhout, B. (1994) Intracellular analysis of in vitro modified HIV Tat protein. *J. Biol. Chem.* 269, 8366–8375.