AD_____

CONTRACT NO: DAMD17-93-C-3122

TITLE: ISOLATION AND PRELIMINARY CHARACTERIZATION OF A RECOMBINANT TAT PROTEIN FROM HUMAN IMMUNOEFFICIENCY VIRUS

SUBTITLE: Isolation of Recombinant HIV-TAT and Mutant TAT Protein for Structural and Functional Studies

PRINCIPAL INVESTIGATOR: Lichun H. Walls, Ph.D.

CONTRACTING ORGANIZATION: Morgan State University (HBC)
Cold Spring Lane & Hillen Road
Baltimore, Maryland 21239

REPORT DATE: June 28, 1994

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research, Development, Acquisition and Logistics Command (Provisional), Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 28, 1994	Annual Report (5/24/93-5/23/94)

4. TITLE AND SUBTITLE
Isolation and Preliminary Characterization of a Recombinant Tat Protein From Human Immunodeficiency Virus

5. FUNDING NUMBERS
Contract No.
DAMD17-93-C-3122

6. AUTHOR(S)
Lichun H. Walls, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Morgan State University (HBC)
Cold Spring Lane & Hillen Road
Baltimore, Maryland 21239

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research, Development, Acquisition and Logistics Command (Provisional), Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Subtitle: Isolation of Recombinant HIV-TAT and Mutant TAT Protein for Structural and Functional Studies

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Human Immunodeficiency rirus type-1 (HIV-1) Tat protein is a transactivator for viral gene expression. In order to carry out the structural and functional studies of the protein, it is necessary to isolate and purify the protein in large quantity. To produce this large quantity, the fusion protein Rop-Tat was over-expressed. The Rop-Tat is isolated from insoluble aggregates of inclusion bodies. Cyanogen bromide cleavage of Rop-Tat yields numerous protein fractions from S-Sepharose and Q-Sepharose chromatography. Efforts are under way to determine which fractions are related to Tat.

14. SUBJECT TERMS
Human Immunodeficiency Virus (HIV), Transactivator protein Tat, Isolation, Recombinant DNA

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

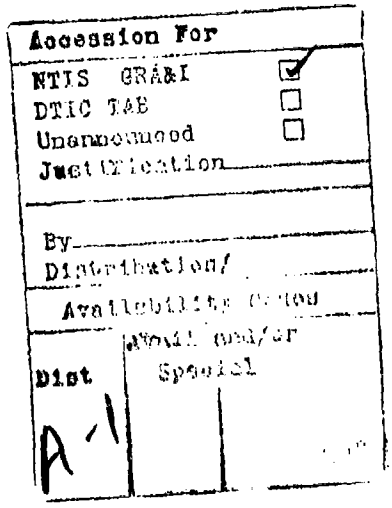
____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

JW ✓ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



Lihun H Wallace 6/28/94

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Human immunodeficiency virus type - 1 (HIV-1) trans-activator (Tat) is a transactivator of viral gene expression. The Tat protein is made of 86 amino acids, including 7 cysteine residues that may participate in metal binding (*i.e.*, Zn^{++}), followed by a string of basic amino acids (1,2,3).

A highly efficient prokaryotic expression system has been developed by Dr. Ru Chih C. Huang's laboratory (co-principal investigator) that produced the fusion protein, Rop-Tat at high levels (4). Rop is a 7.2 kDa, 63 amino acid protein. Only 52 amino acids of Rop are fused with the 86 animo acid Tat resulting in a 15.8 kDa fusion protein, Rop-Tat (4). The fusion protein, Rop-Tat is accumulated as insoluble aggregates of Inclusion Bodies (IB). The over-expressed bacterial cells were supplied by Dr. Huang's laboratory. In our laboratory we first prepared the IB, and then used the cyanogen bromide cleavage at methionine residues. The Rop-Tat fusion protein has two internal methionine residues. It should yield protein (or peptide) 2-11 (no. 1 is a methionine residue) 12-53, 54-138. The 54-138 is the 86 amino acid residue of Tat (see Appendeix).

This report has four parts:

I. Inclusion Bodies (IB) preparation

II. Cyanogen Bromide Cleavage of the Fusion Protein, Rop-Tat

III. Purification of the CNBR Cleavage Protein Mixtures

IV. Discussion

BODY OF THE REPORT

I. Inclusion Bodies (IB) Preparation

The Rop-Tat protein in the *E. coli* microorganism (over-expressed) was supplied by Dr. R.C. Huang of Johns Hopkins University, co-principal investigator. The initial procedure used a cell homogenizer to prepare the IB, which yielded Rop-Tat. Upon electrophoresis (see Figure 1, Lanes 4 & 5) the Rop-Tat was not a major band. The initial procedure was modified whereby a French Press was substituted for the cell homogenizer to disrupt the *E. coli* cells. The result is shown in Figure 1, Lane 2. Rop-Tat appears to be a major band. The use of the French Press effectively prepared the IB, but this process generated a significant amount of heat. A new cell disruption system, the BioNeb Disruption System, disrupts the cell under helium gas and generated no heat, therefore, is better for our IB preparation. There seems to be fewer contaminant bands (see Figure 1, Lane 1).

II. Cyanogen Bromide Digest of Rop-Tat

In order to obtain Tat protein, it was necessary to incubate the inclusion bodies with cyanogen bromide in the presence of 6 M Guanidine hydrochloride (GudnCl) and hydrochloric acid (HCl) because of the solubility problem of IB. Various concentrations of HCl were tried (0.1 M, 0.2 M, 0.5 M, 1 M, *etc.*). The problem of not using too weak a concentration of HCl was that of incomplete breakage of Rop-Tat, resulting in a mixture of Rop-Tat and various sizes of Tat. The danger of using too high a concentration of HCl was that it could break too many protein bonds. The final condition of HCl concentration chosen to be used was 0.75 M (See Figure 3).

III. Isolation of Tat Protein by Ion-Exchange Column Chromatography

A. S-Sepharose Cation Exchange Column Chromatography

The protein mixture after cyanogen bromide digest was applied onto a cation S-Sepharose column for chromatography. Three elution fractions (0.1 M; 0.33 M; 0.5 M NaCl) were obtained from the column. Figure 3 shows the scheme for purification of the Tat protein(s). Figure 2 shows the column profile. Although the runthrough fraction appears to be the major fraction, it is mostly low molecular weight peptide from gel electrophoresis. However, if the S-Sepharose column is overloaded, then larger molecular weight protein also begins to appear in this runthrough fraction. At present efforts are concentrated on purifying the three elution fractions, 0.1 M, 0.33 M and 0.5 M from the S-Sepharose column. (The Sodium Dodecyl Detergent (SDS) gel electrophoresis of S-Sepharose column is shown on Figure 5.) Each of these fractions was subjected to size exclusion, using Sephacryl S-200 column chromatography. Sephacryl S-200 columns are used to determine molecular weights. Results are shown in Figure 4. Previously, because of the solubility problem, the size column had to be run in the presence of 5 M urea, and this rendered tests of protein size inclusive. We then discovered that the Tat protein can be dissolved in alkali pH 8 -> 10 in the absence of the denaturant urea. In this pH range, the 0.1 M elute dissolves in pH 8 tris buffer, whereas the 0.5 M elute requires pH 10 buffer. For uniformity, all gel filtration columns were run on 0.2 M tris, 0.15 M NaCl, pH 10. The 0.15 M NaCl is added to prevent interaction between the gel matrix and the protein. The molecular weights of the elution fractions are

0.1 M ≈ 9,900 Daltons
0.33 M ≈ 19,000 Daltons
0.5 M ≈ 77,000 Daltons.

However, when the 0.5 M elute is subjected to both reducing and non-reducing conditions of the SDS gel electrophoresis, two bands occur, both of which are below the 20,000 D molecular weight standard in the non-reduced state (see Figure 5, Lanes 4 & 5). Since the detergent, SDS, is unable to break covalent bonds, this means that the 77,000 D molecular weight is probably a non-covalent aggregate.

B. Q-Sepharose Anion Exchange Column Chromatography

Each elution fraction (0.1 M; 0.33 M; 0.5 M) from S-Sepharose were dialyzed against 10 mm HCl and lyophilyzed. The 0.1 M fraction was then dissolved in a 1 M urea, 0.05 M tris, pH 8.2 buffer, and sodium hydroxide was added to adjust to pH 8.2 as necessary. This fraction was then applied onto a Q-Sepharose column equilibrated with the same buffer. The column was eluted with 0.1 M NaCl and 0.5 M NaCl added to the column buffer. Runthrough and two more fractions were collected as noted on Figure 3. SDS electrophoresis of the runthrough (Figure 5, Lane 6) and 0.1 M elutes (electrophoresis not shown) showed a smear pattern, indicating low molecular impurities. SDS electrophoresis of the 0.5 M elute appears to show a pure protein (see Figure 5, Lane 5; also Figure 7, Lane 4). This fraction was subjected to amino acid sequence analysis.

The 0.33 M fraction from S-Sepharose, after being dialyzed and lyophilyzed, was dissolved in 2 M urea, 0.05 M tris, pH 8.2. It was then dialyzed against the same 2 M urea buffer for several days. The fraction was then applied onto a Q-Sepharose column (see Figure 3). Runthrough and three fractions were collected and subjected to SDS electrophoresis (results shown on Figure 6, Lanes 2 to 5).

The 0.5 M fraction from the S-Sepharose column, after being dialyzed and lyophilyzed, was dissolved and dialyzed in 4 M urea, 0.05 M tris, pH 8.2. The

fraction was then applied onto a Q-Sepharose column (see Figure 3). Runthrough and three fractions were collected and subjected to SDS electrophoresis. The results are shown on Figure 6, Lanes 6 to 9.

Alternately, the 0.5 M fraction from the S-Sepharose column, after being dialyzed against 20 mM sodium phosphate, 5 M urea, 20% acetonitrile pH 7.8, and reapplied onto the S-Sepharose column can further dissociate into a 9,900 D molecular weight moiety (see Figure 7, Lanes 1 & 2).

DISCUSSION & CONCLUSIONS

Using the 15% gel electrophoresis, all the fractions from the 0.1 M, 0.33 M, 0.5 M of S-Sepharose, have bands in line with the 14,000 D molecular weight standard (see Figure 6, Lanes 5 to 10). Since the S-Sephacryl column of the 0.1 M elute gave a molecular weight of 9,900 D, we might assume that all the bands are the same protein in a reduced state. The reason they might elute with different sodium chloride concentrations is because some of them might be monomers, dimers, *etc.* However, when some of these fractions were subjected to low molecular weight SDS gel electrophoresis (using tricine instead of glycine buffer), the 0 1 M elute of the 9,900 D bands appeared under the 8,100 D standard (see Figure 7, Lane 4). When this fraction was subjected to amino acid sequence analysis, it proved to be the peptide 12-53 (from the Rop part of Rop-Tat). This peptide should have a molecular weight of 4,876 D. The low molecular weight gel electrophoresis (see Figure 7. Lane 4) shows a band below the 8,160 D standard, but above the 7,200 D Rop protein (Figure 7, Lane 3). Since there is a cysteine residue in this peptide, it may have formed a dimer, an unexpected finding. The Tat protein is known to self-associate to form dimers and even heterodimers (6). The question is whether the band which is very close to the above Rop dimer molecular weight is also a

peptide from the Rop part of the Rop-Tat (compare Lanes 1 and 4 of Figure 7). The bands in Figure 7, Lanes 1 and 2 are from the 0.5 M elute of the S-Sepharose and the band from Lane 4 (the 12-53 peptide) is from 0.1 M elute. They should not be the same protein. Since Tat is a basic protein it should absorb more strongly than the Rop protein onto the cation S-Sepharose (see Figure 6, Lanes 5 & 10). These bands from the 0.33 M elute (Figure 6. Lanes 5 & 10) also look to be more or less the same size as those from the 0.1 M and 0.5 M elutes. The 15% SDS gel electrophoresis cannot separate proteins very well at molecular weights below 10,000 D. Even the low molecular weight gel electrophoresis cannot well differentiate between proteins of 8.000 - 10,000 D. Therefore, the most pure fractions from each of the sodium chloride elutes should be subjected to amino acid sequence analysis to identify which are the Tat protein and which are from the Rop part of the fusion protein.

Future Plan:

The most pure fraction from the Q-Sepharose is subjected to amino acid sequence analysis. Once the Tat protein is identified effort can then be made to purify that fraction to homogeneity.

REFERENCES

1. Arya S.K., Guo C., Josephs S. and Wong-Staal F. Transactivator gene of the human T-Lymphotropic virus type (HTLV-III). *Science*, 229:69-73, 1985.

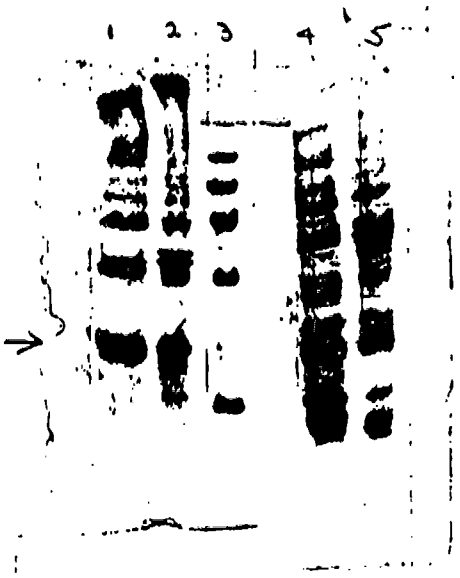
2. Jeang K.T., Chang K., Berkhout B., Hammarskjold M.L. and Rekosh D. Regulation of HIV expression: mechanism of action. *AIDS*, 5(suppl 2):S3-14. 1991.

3. Bohan C.A., Knaschanchi F., Ensoli B., Buonaguro L., Boris-Lawrie K.A. and Brady J.N. Analysis of Tat transactivation of human immunodeficiency virus transcription in vitro. *Gene Expression*, 2:391-407, 1992.

4. Giza P.E. and Huang R.C. A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. *Gene*, 78:73-84, 1989.

5. Schagger H. and von Jagow G. Tricine-sodium Dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166:368-379, 1987.

6. Frankel A.D., Ling C., Cotter R.J. and Pabo C.O. Dimerization of the Tat protein from human immunodeficiency virus: A cysteine-rich peptide mimics the normal metal-linked dimer interface. *Proc. Natl. Acad. Sci USA*, 85:6297-6300, 1988.



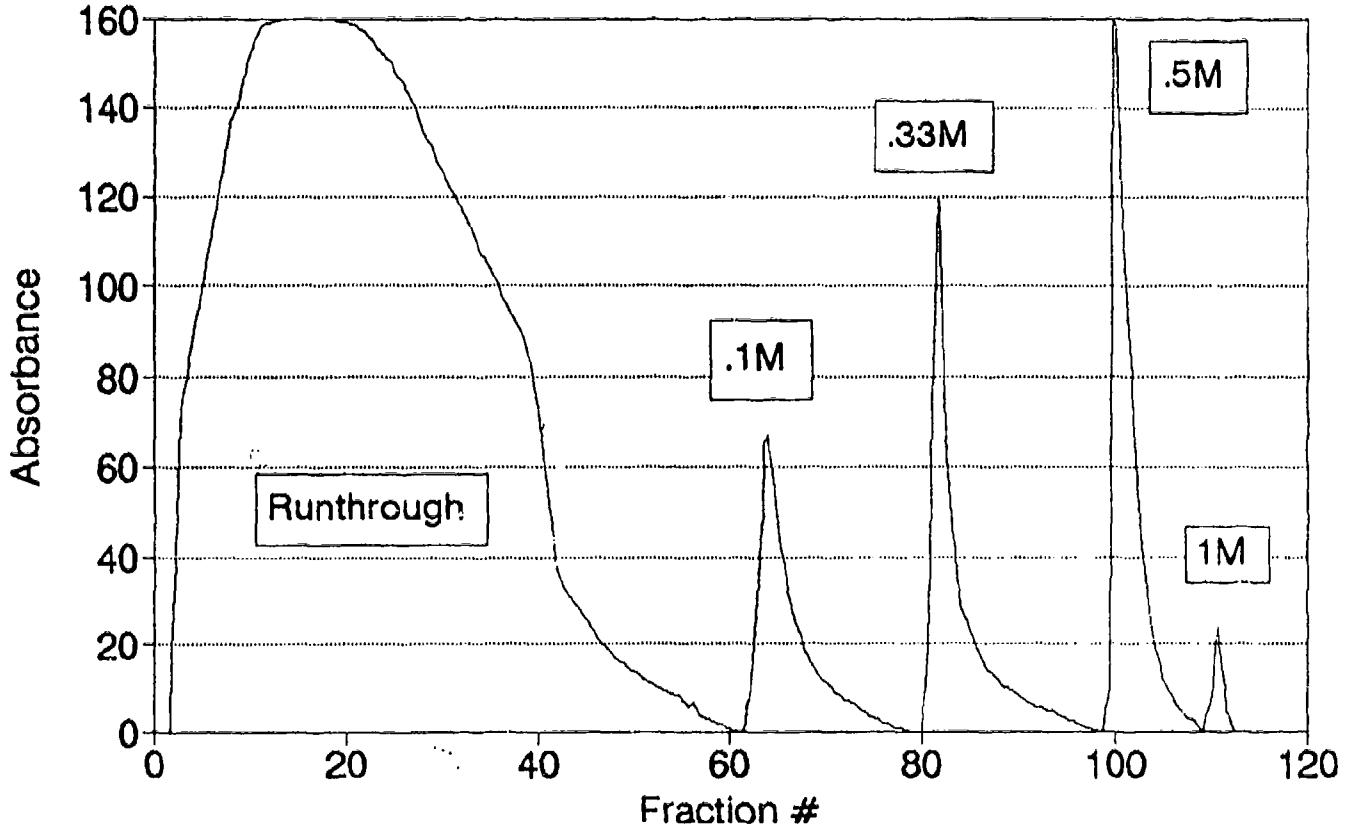


Figure 1 (upper). Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. Lane 1 IB prepared by BioNeb; Lane 2, IB by French Press; Lane 3, molecular weight standards; Lanes 4 & 5, IB by cell homogenizer. The arrow indicates the position of Rop-Tat.

Figure 2 (lower). S-Sepharose chromatography of cyanogen bromide-Rop-Tat. The sample was loaded onto a column (1.5 x 4 cm) preequilibrated with 5 M urea containing 0.05 M sodium acetate pH 5. The column was eluted stepwise with 0.1 M, 0.2 M, 0.33 M, 0.5 M sodium chloride in the column buffer.

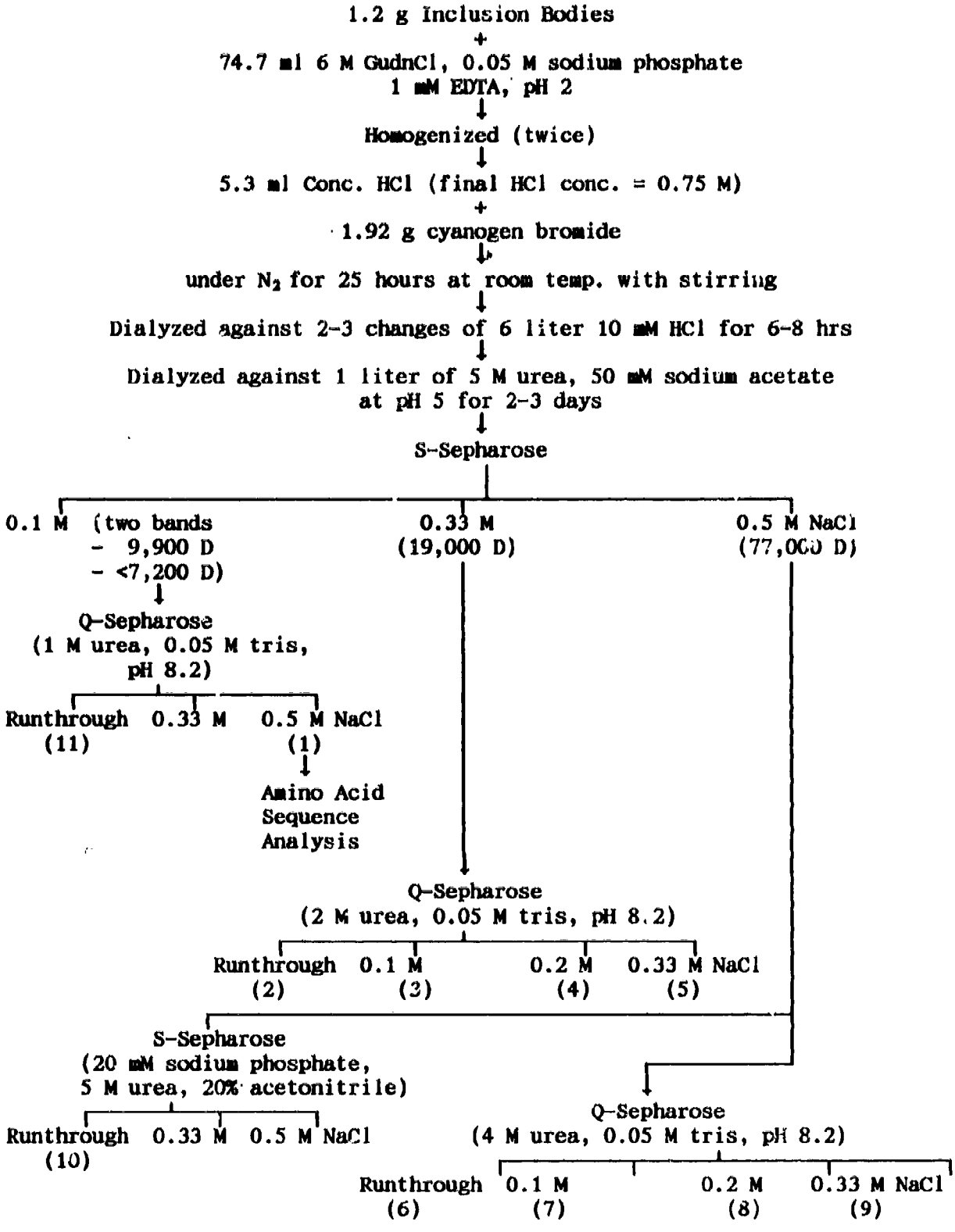


Figure 3. Scheme of purification of Tat protein.

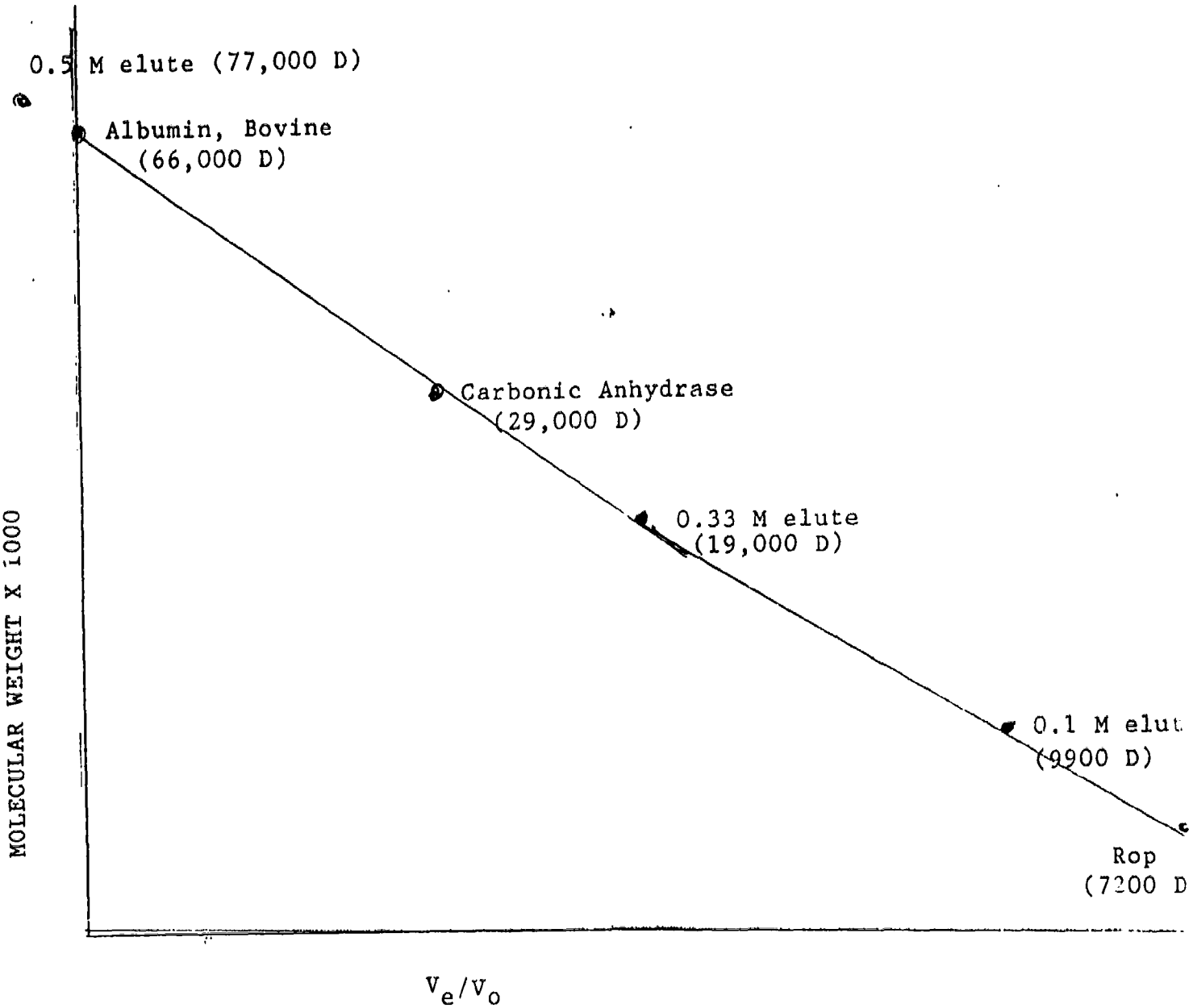


Figure 4 Sephacryl S-200 gel filtration (column size 1 x 48 cm) of fractions from S-Sepharose column (0.1 M , 0.33 M, 0.5 M NaCl elutes). The fractions were compared with standards as indicated above.

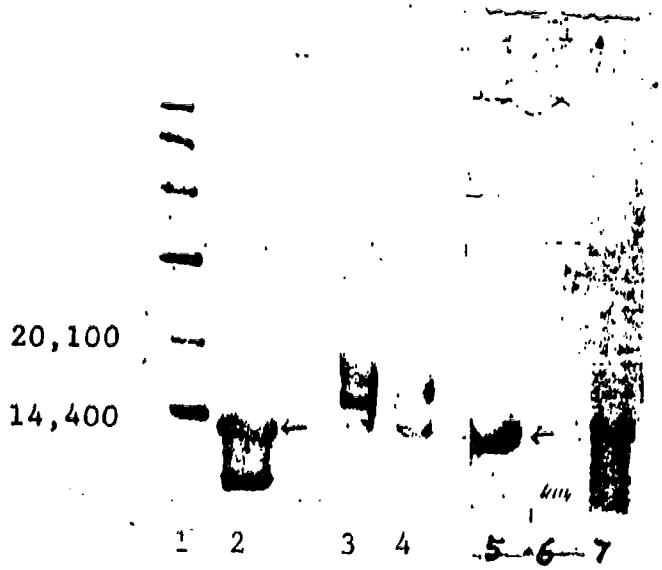


Figure 5 SDS-gel electrophoresis in 15% acrylamide gel.

All samples in reduced state except sample on Lane 3, where β-mercaptoethanol is omitted from the sample.

Lane 1. Molecular weight standard.
Lane 2. 0.1 M elute from the S-Sepharose column.
Lane 3. Non-reduced 0.5 M elute from the first S-Sepharose column.
Lane 4. Reduced state of 0.5 M NaCl elute of sample from Lane 3.
Lane 5 Sample no. 1, Figure 3.
Lane 6. Sample no. 11, Figure 3.
Lane 7. 0.33 M elute of S-Sepharose column

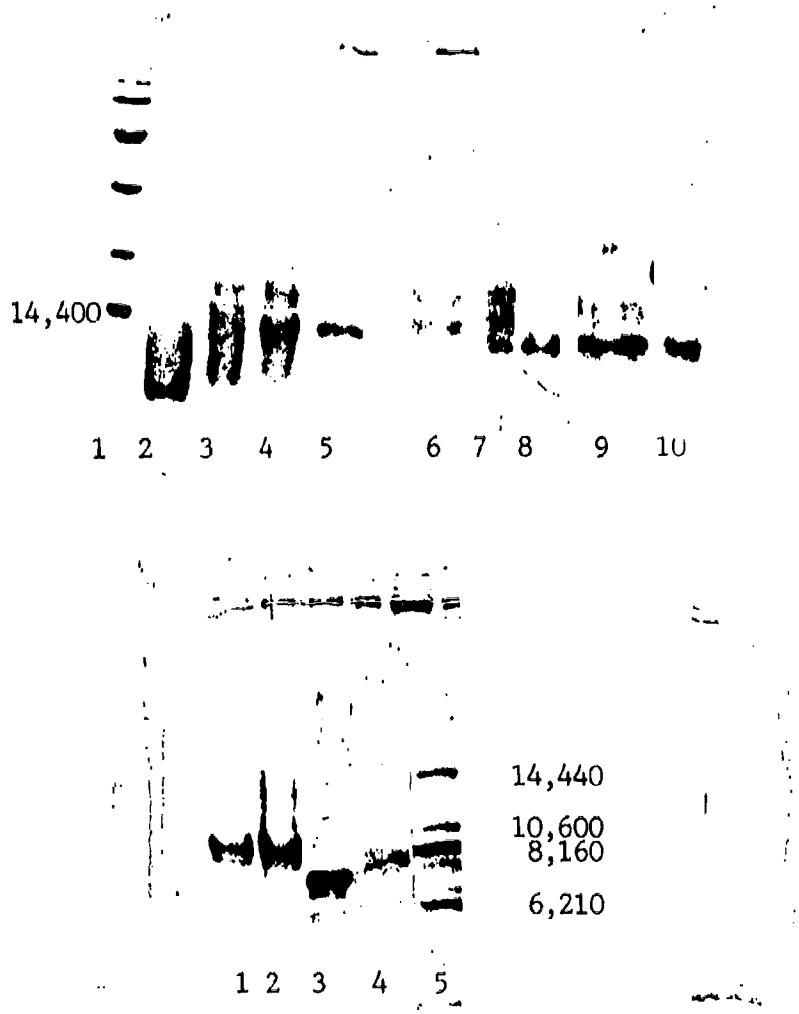


Figure 6 (upper gel) 15% SDS Gel Electrophoresis. All samples are in reduced state.
Lane 1. Molecular weight standard.
Lane 2. Sample 2, Figure 3.
Lane 3. Sample 3, "
Lane 4. Sample 4, "
Lane 5 and 10. Sample no. 5 , Figure 3.
Lane 6. Sample no. 6, Figure 3.
Lane 7. Sample no. 7, "
Lane 8. 8, "
Lane 9. 9, "

Figure 7 (Lower gel) Low Molecular Weight Tricine Gel Electrophoresis (5).
Lane 1 and 2 . Sample no. 10, Figure 3.
Lane 3. Rop protein , molecular weight 7,200
Lane 4. Sample no. 1, Figure 3.
Lane 5. Low Molecular Weight Standards

REFERENCES

1. Arya S.K., Guo C., Josephs S. and Wong-Staal F. Transactivator gene of the human T-Lymphotropic virus type (HTLV-III). *Science*, 229:69-73, 1985.

2. Jeang K.T., Chang K., Berkhout B., Hammarskjold M.L. and RekoshD. Regulation of HIV expression : mechanism of action. *AIDS*, 5(suppl 2):S3-14. 1991.

3. Bohan C.A., Knaschanchi F., Ensoli B., Buonaguro L., Boris-Lawrie K.A. and Brady J.N. Analysis of Tat transactivation of human immunodeficiency virus transcription in vitro. *Gene Expression*, 2:391-407, 1992.

4. Giza P.E. and Huang R.C. A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. *Gene*, 78:73-84, 1989.

5. Schagger H. and von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166:368-379, 1987.

6. Frankel A.D., Ling C., Cotter R.J. and Pabo C.O. Dimerization of the Tat protein from human immunodeficiency virus: A cysteine-rich peptide mimics the normal metal-linked dimer interface. *Proc. Natl. Acad. Sci USA*, 85:6297-6300, 1988.

APPENDIX A

Amino acid sequence of Rop-tat fusion

met thr lys gln glu lys thr ala leu asn met (11) ala arg phe ile
arg ser gln thr leu thr leu leu glu lys leu asn glu leu asp
ala asp glu gln ala asp ile cys glu ser leu his asp his ala
asp glu leu tyr arg ser cys met (53) glu pro val asp pro arg leu
glu pro trp lys his pro gly ser gln pro lys thr ala cys thr
asn cys tyr cys lys lys cys cys phe his cys gln val cys phe
ile thr lys ala leu gly ile ser tyr gly arg lys lys arg arg
gln arg arg arg pro pro gln gly ser gln thr his gln val ser
leu ser lys gln pro thr ser gln ser arg gly asp pro thr gly
pro lys glu (138)

Molecular weight of Rop protein: 7.2 kD

Molecular weight of Rop-tat fusion: 15.8 kD

Molecular weight of residue 12-53 = 4,876

Molecular weight of residue 54-138 = 9673, Tat Protein

Residues 1-53 is Rop protein , molecular weight = 6,153

APPENDIX B

Annual Report
(05/24/93-05/23/94)

Subcontractor: Ru Chih C.Huang
Department of Biology
Johns Hopkins University

Introduction

The tat gene plays an important role in transactivation of HIV-1 replication and in HIV virus production. The protein encoded by the tat gene consists of 86 amino acids and is characterized by a cysteine-rich region (amino acid 22 to 37) and a high lysine/arginine region (amino acid 49 to 57). The two regions are separated from each other by a core of eleven amino acids containing a lysine residue at position 41 and glycine at position 42 (strain SF-2)(1).

We have reported earlier on the analysis of seven Tat mutants and their ability to enhance HIV-LTR directed expression in human embryonal carcinoma NTERA-2 cells (2). These constructs consisted of wild-type and mutated tat genes (strain SF-2) driven by the SV40 early promoter (1). The plasmids were co-transfected into cells with an LTR-CAT reporter plasmid to assess the effect of the mutation (2). Of the seven mutants assayed only tat42, a glycine to alanine change, was shown to be transactivation positive. The other mutants ($Tat_{26/28}$, $Tat_{28/31}$, $Tat_{29/30}$, Tat_{27}, Tat_{28}, and Tat_{41}) showed little enhancement of HIV-LTR directed enhancement.

In order to obtain Tat protein (and other HIV proteins) in sufficient amounts for further study, they must be chemically synthesized or produced in eukaryotic or prokaryotic expression systems, since they cannot be isolated from their natural source.

We have achieved high level synthesis of Tat proteins by subcloning mutant tat genes into a novel prokaryotic expression system that is especially useful for the production of small proteins and polypeptides. Since we first described the system (3) during the first granting period (05/24/93-05/23/94), we have gathered insight into how it functions, and have refined conditions for the synthesis and large scale purification of Tat proteins. The essential findings are summarized as follows:

The Body of This Annual Report

a. Description of the Rop-expression system

The expression system consists of two plasmids; an ampicillin (Ap) resistant expression vector, pPGtrpRopAp, that carries the rop gene (4) as the bacterial fusion mojety; and a tetracycline (Tc) resistant helper plasmid that allows for propagation of rop gene fusion plasmids by suppressing their lethal runaway plasmid replication phenotype. Each plasmid carries a single copy of the trp promoter-operator upstream of the rop gene and promoter, and exhibits normal plasmid copy numbers when grown at 30°C. Gene fusions with the rop sequence in pPGtrpRopAp cause lethal runaway plasmid replication because of the loss of copy number control

mediated by the 63 amino acid Rop protein. Runaway replication can be suppressed by supplying Rop activity in-trans by co-transforming cells with fusion plasmid and helper plasmid and maintaining selective pressure with both Ap and Tc.

Induction growth is accomplished by culturing co-transformed cells in medium containing only Ap. The Ap resistant cells lose the helper plasmid, the expression plasmid begins runaway replication in early stationary phase, and Rop-fusion proteins accumulate as the trp promoter is derepressed via repressor titration. No chemical inducers or temperature shifts are required for the induction of the trp promoter. Since cellular growth and promoter induction can be carried out in rich growth medium, high culture densities ($4x10^9$ cells/ml) combine with high cellular fusion protein accumulation (30-40%) to produce fusion protein yields as high as 1 gram per liter of culture. The efficient loss of the Tc resistant helper plasmid is crucial to the functioning of this system. At the end of induction growth, more than 99.5% of the cells have lost the helper plasmid which results in plasmid DNA amplification as high as 35 μg/ml of culture.

b. Hypersynthesis of Two Mutant Rop-Tat Proteins

We have earlier shown that Rop-Tat_{41} (lysine → glu at position 41) protein has no transactivation ability while Rop-Tat_{42} (gly → ala at position 42) is a transactivation-positive mutant. A large production of these two proteins was accomplished.

We use the constructs pPGtrpRopTat41 and pPGtrpRopTat42 to illustrate the method of protein production. All constructs produced Tat fusion proteins at comparable levels. HB101 was transformed with the helper plasmid pPGtrpRopTc and then co-transformed with pPGtrpRopTat41 and pPGtrpRopTat42. Cultures were maintained at 30°C in medium supplemented with Ap and Tc. After 48 hours of induction growth RopTat fusion proteins accumulated to 42% of the bacterial proteins. We have found that growth in TB medium (5) not only increases levels of cellular fusion protein synthesis, but also enables cultures to reach higher densities. The OD_{660} of induced $RopTat_{41}$ and $RopTat_{42}$ cultures were 10.60 and 7.76, respectively.

c. Purification and the Yields of Rot-Tat_{41} and Rop-Tat_{42}

Initial purification of the fusion proteins was accomplished by cellular lysis and centrifugation from inclusion bodies. The recovered fusion protein pellets were washed to remove residual soluble protein contaminants, and finally dissolved in 6M guanidine hydrochloride (GuHCl). The partially purified yields of RopTat41 and RopTat42 were 1.25 g/liter and 0.9 g/liter, respectively. The fusion proteins recovered in GuHCl were determined to be 80-85% pure. Further large scale purification was achieved by taking advantage of the high binding capacity of hydroxyapatite (HAP). We found that RopTat proteins could be adsorbed to HAP in 6M urea at pH 8.0, while non-RopTat contaminating proteins exhibited little

binding affinity under these conditions. Approximately 270 mg of partially purified RopTat41 was adsorbed to 10 g (dry weight) HAP. After removal of unbound proteins by washing of the HAP bed, 207 mg of purified RopTat41 was recovered upon elution with phosphate buffer. Any RopTat protein remaining in the unbound fraction because of HAP saturation could be recovered by repeating the binding procedure with fresh HAP. HAP adsorption of the same amount of RopTat42 GuHCl extract resulted in nearly identical purification and recovery.

Analysis of RopTat fusion proteins by reducing SDS-PAGE reveals a main band at approximately 16 kilodaltons.

d. Biological Function of the Rop-Tat Proteins

The proteins produced by the expression system were first tested for activity using microinjection. To assess whether recombinant Tat proteins derived from our Rop expression system can participate directly in transactivation of HIV-LTR in vivo, HIV template plasmid DNA containing either full-length HIV-LTR or a NF-KB (B) and Sp1 (III,II) depleted HIV-LTR (HIVΔLTR) were used. Test proteins Rop, RopTat, and cyanogen bromide (CnBr) cleaved products of RopTat and RopTat41 (designated Tat and Tat41, respectively) were injected either with or without templates (HIV-LTR-CAT or HIVΔLTR-CAT) into either HeLa cells or NIH 3T3 cells. The CAT protein synthesized as the result of template expression was monitored in situ by autoradiographic analysis of the immuno-stained cells. Data have shown that both RopTat protein and the CNBr cleaved Tat product from RopTat are excellent transactivators, while Rop and Tat41 proteins have no activities in vivo. This in vivo transactivation was shown to require the proximal *cis* elements of HIV-LTR, the deletion template HIVΔLTR-CAT has no basal activity and is not responsive to Tat activation.

The transacivation of Rop-Tat proteins has been also analyzed in a transcriptional assay system in vitro. For these studies, the Rop-Tat proteins were tested for transcription in HeLa whole cell extract using the same HIV-LTR-CAT as template (6,7). A positive "control" mutant, RopTat42 (gly→ala) was included for comparison. High concentration of test protein stock (10-20 mg/ml) was diluted 10 to 20 fold by stepwise addition of Tat storage buffer and immediately added to the incubation mixture in the presence of HIV template DNA (7,8). Production of run-off transcripts of HIV-LTR-CAT (325 b) was analyzed by acrylamide gel electrophoresis. It was found that at concentrations of 5-10 μm, RopTat42 enhances HIV-LTR-CAT transcription 32 to 37 fold, slightly better than that of full length RopTat (22 fold). In contrast, RopTat41 only enhanced HIV transcription 3-5 fold. A manuscript which described the entire study as indicated in this report is submitted with the annual report (see Appendix).

Conclusions and Work for Year 2 of this Grant (05/24/94--5/24/95)

With the availability of gram level amounts of Tat and mutant Tat proteins and an *in vitro* system to assay them, procedures are currently being developed to isolate the properly folded molecules. The determination of the cysteines that may be involved in intramolecular disulfide bonds could give important structural information. These bonds may be responsible for the *in vitro* stability of the active, properly folded molecule. If the correct disulfide bonds were formed *in vivo*, maintaining them may facilitate the proper refolding of the molecules upon removal of the denaturants. In the coming year, in collaboration with Dr. B. C. Wang, these crystallographic studies will be made. In addition, we will begin to study the specific molecular interactions between Tat, TAR and other transcriptional factors that may be responsible for the transactivation process.

References

1. B. Matija Peterlin, P. A. Luciw, P. J. Barr, M. D. Walker, Proc. Natl. Acad. Sci. U.S.A. **83**, 9734 (1986).
2. R. C. Huang *et al.*, *in* Gene Regulation and AIDS, Adv. in Applied Biotech. Series, T. S. Papas, Ed. (1989), Vol. 7, 147-160.
3. P. E. Giza and R. C. Huang, Gene **78**, 73 (1989).
4. G. Cesareni, M. A. Muesing, B. Polisky, Proc. Natl. Acad. Sci. U.S.A. **79**, 6313 (1982).
5. K. D. Tartof and C. A. Hobbs, Bethesda Res. Lab Focus **9**, 12 (1987).
6. R. A. Marciniak, and P. A. Sharp, EMBO J. **10**, 4189 (1991).
7. C. A. Bohan *et al.*, Gene Exp. **2**, 391 (1992).
8. F. Kashanchi *et al.*, Nature **367**, 295 (1994).

Hypersynthesis of the HIV Tat Protein in *Escherichia coli* by the Rop Expression System

Paul E. Giza, Mark R. Smith*1, Hsiang-fu Kung*2,
John N. Brady**, and Ru Chih C. Huang

Department of Biology, Johns Hopkins University,
Baltimore, Maryland 21218
*1Biological Carcinogenesis and Development Program,
PRI/DynCorp, Biological Response Modifiers Program,
National Cancer Institute, Frederick Cancer Research and
Development Center, Frederick, MD 21702
*2Laboratory of Biochemical Physiology, BRMP, DCT,
National Cancer Institute, Frederick, Maryland 21702
**Laboratory of Molecular Virology,
National Cancer Institute, Bethesda, Maryland 20892

Correspondence: Ru Chih C. Huang, Department of Biology, Johns Hopkins University, Baltimore, Maryland 21218. Tel. 410-516-5181
Fax 410-516-5213

ABSTRACT

High level synthesis of the HIV Tat protein and mutated Tat proteins has been accomplished by fusion of the *tat* gene sequences to the *E. coli rop* gene sequence. Gram level amounts of the RopTat fusion proteins have been produced in bacteria, and the ability of the purified proteins to transactivate the HIV promoter has been analyzed in eukaryotic cells and by an *in vitro* transcription system. RopTat fusion proteins have been found to be as active as Tat (non-fused) protein, whereas several mutated Tat proteins exhibit little or no activity. A lysine to glutamic acid mutation at amino acid 41 has been compared to wild-type Tat, and has shown little activity *in vivo* or *in vitro*. The availability of large amounts of active and inactive Tat proteins, that can be systematically analyzed by *in vitro* methods should aid in the structural determination of the Tat protein.

The tat gene plays an important role in transactivation of HIV-1 replication and in HIV virus production. The protein encoded by the tat gene consists of 86 amino acids and is characterized by a cysteine-rich region (amino acid 22 to 37) and a high lysine/arginine region (amino acid 49 to 57). The two regions are separated from each other by a core of eleven amino acids containing a lysine residue at position 41 and glycine at position 42 (strain SF-2)(1). Cotransfection studies using deletion mutant tat constructs with a plasmid DNA containing HIV LTR linked reporter gene have revealed that only the first 58 amino acids of Tat are essential for enhancing the promoter activity of the long terminal repeat (LTR) of HIV in mammalian cells (2). The cysteine residues at positions 27 and 30, and lysine residues at pcsitions 29 and 41 have been shown to be important in transactivation by many site-specific mutagenesis studies (3,4,5,6). In addition, the involvement of a *cis*-acting RNA, a 58nt short transcript (transactive responsive RNA, TAR) located immediately proximal to the viral mRNA cap site, has been shown (7,8,9,10).

Recent *in vitro* transcription studies using highly purified recombinant Tat proteins have given some direct insights in regard to the mechanism of Tat transactivation at the level of initiation and elongation (11,12,13). Recombinant Tat protein, when added either to crude nuclear extract (12) or to reconstituted HeLa cell nuclear extracts (13), yielded a 10-60 fold increase in basal HIV LTR promoter activity. Kashanchi et

al. (13) recently reported that the Tat protein affected HIV LTR transcription in part by interactions with TATA-binding protein (TBP) and associated factors (TAFs) of the transcriptional factor TFIID complex. The Tat core sequence (amino acids 36-50) was found to be essential for this TFIID interaction. Mutant Tat41 protein, although it binds TAR RNA, does not interact with TFIID and is inactive in transactivation (13). Three-dimensional structural differences between wild type and mutant Tat proteins would be extremely important in elucidating the mechanism of Tat regulation. Such studies would require sufficient amounts of purified proteins for x-ray crystallographic studies, and for binding analysis involving components of a functional HIV transactivation complex.

We have reported earlier on the analysis of seven Tat mutants and their ability to enhance HIV-LTR directed expression in human embryonal carcinoma NTERA-2 cells (6). These constructs consisted of wild-type and mutated tat genes (strain SF-2) driven by the SV40 early promoter (1). The plasmids were co-transfected into cells with an LTR-CAT reporter plasmid to assess the effect of the mutation (6). Of the seven mutants assayed only tat42, a glycine to alanine change, was shown to be transactivation positive. The other mutants (Fig. 1) showed little enhancement of HIV-LTR directed enhancement.

In order to obtain Tat protein (and other HIV proteins) in sufficient amounts for further study, they must be chemically synthesized or produced in eukaryotic or prokaryotic expression

systems, since they cannot be isolated from their natural source. Isolation, purification, and identification of the properly folded, active protein is problematical, because the expression system may not produce large amounts of the active form of the protein, or the purification procedure itself may introduce spurious changes. Preparation of Tat protein crystals therefore requires sufficient amounts of starting material.

We have achieved high level synthesis of Tat proteins by subcloning mutant tat genes (Fig. 1) into a novel prokaryotic expression system that is especially useful for the production of small proteins and polypeptides. Since we first described the system (15), we have gathered insight into how it functions, and have refined conditions for the synthesis and large scale purification of Tat proteins.

The expression system consists of two plasmids; an ampicillin (Ap) resistant expression vector, pPGtrpRopAp, that carries the rop gene (16) as the bacterial fusion moiety; and a tetracycline (Tc) resistant helper plasmid that allows for propagation of rop gene fusion plasmids by suppressing their lethal runaway plasmid replication phenotype. Each plasmid carries a single copy of the trp promoter-operator upstream of the rop gene and promoter, and exhibits normal plasmid copy numbers when grown at 30°C. Gene fusions with the rop sequence in pPGtrpRopAp cause lethal runaway plasmid replication because of the loss of copy number control mediated by the 63 amino acid Rop protein. Runaway replication can be suppressed by supplying

Rop activity *in-trans* by co-transforming cells with fusion plasmid and helper plasmid and maintaining selective pressure with both Ap and Tc.

Induction growth is accomplished by culturing co-transformed cells in medium containing only Ap. The Ap resistant cells lose the helper plasmid, the expression plasmid begins runaway replication in early stationary phase, and Rop-fusion proteins accumulate as the *trp* promoter is derepressed via repressor titration. No chemical inducers or temperature shifts are required for the induction of the *trp* promoter. Since cellular growth and promoter induction can be carried out in rich growth medium, high culture densities ($4x10^9$ cells/ml) combine with high cellular fusion protein accumulation (30-40%) to produce fusion protein yields as high as 1 gram per liter of culture. The efficient loss of the Tc resistant helper plasmid is crucial to the functioning of this system. At the end of induction growth, more than 99.5% of the cells have lost the helper plasmid which results in plasmid DNA amplification as high as 35 μg/ml of culture. We have found that this helper plasmid expulsion could not be reproduced using other repressible bacterial promoters. The construction of an equivalent system by the replacement of the *trp* promoters on both plasmids with *lacZ* promoters did not result in comparable loss of the *lacZ* helper (data not shown). A possible explanation is that excessive copies of the *trp* PO cause sufficient titration of the *trp* repressor pool that the genomic *trp* operon is induced. Bacterial cultures grown under conditions

of trp promoter derepression tend to grow poorly compared to repressed cultures. Thus, loss of the helper plasmid with its compliment of trp PO favors the growth of cells harboring only the Ap resistant expression vector. Since the expression plasmid DNA synthesis and fusion protein accumulation begin in early stationary phase, problems typically associated with trp based expression systems; such as plasmid instability and poor culture growth, are overcome (see Ref. 17 for an excellent review of trp based expression systems).

We use the constructs pPGtrpRopTat41 and pPGtrpRopTat42 to illustrate the method of protein production. All constructs (Fig. 1) produced Tat fusion proteins at comparable levels. HB101 was transformed with the helper plasmid pPGtrpRopTc and then co-transformed with pPGtrpRopTat41 and pPGtrpRopTat42. Cultures were maintained at 30°C in medium supplemented with Ap and Tc (Fig. 2A, lanes 1 and 3) (18). After 48 hours of induction growth RopTat fusion proteins accumulated to 42% of the bacterial proteins (Fig. 2A, lanes 2 and 4). We have found that growth in TB medium (19) not only increases levels of cellular fusion protein synthesis, but also enables cultures to reach higher densities. The OD_{660} of induced $RopTat_{41}$ and $RopTat_{42}$ cultures were 10.60 and 7.76, respectively,

Initial purification of the insoluble fusion proteins was accomplished by cellular lysis and centrifugation (18). The recovered fusion protein pellets were washed to remove residual soluble protein contaminants, and finally dissolved in 6M

guanidine hydrochloride (GuHCl). The partially purified yields of RopTat41 and RopTat42 were 1.25 g/liter and 0.9 g/liter, respectively. The fusion proteins recovered in GuHCl (Fig. 2B, lane G) were determined to be 80-85% pure. Further large scale purification was achieved by taking advantage of the high binding capacity of hydroxyapatite (HAP). We found that RopTat proteins could be adsorbed to HAP in 6M urea at pH 8.0, while non-RopTat contaminating proteins exhibited little binding affinity under these conditions.

The purification of RopTat41 is shown in Fig. 2. Approximately 270 mg of partially purified RopTat41 (Fig. 2B, lane G) was adsorbed to 10 g (dry weight) HAP (20). After removal of unbound proteins (Fig. 2, lane U) by washing of the HAP bed, 207 mg of purified RopTat41 was recovered upon elution with phosphate buffer (Fig. 2B, lanes E1, E2, E3). Any RopTat protein remaining in the unbound fraction because of HAP saturation could be recovered by repeating the binding procedure with fresh HAP (20). HAP adsorption of the same amount of RopTat42 GuHCl extract resulted in nearly identical purification and recovery.

Analysis of RopTat fusion proteins by reducing SDS-PAGE (Fig. 2B) reveals not only a main band at approximately 16 kilodaltons, but also several higher molecular weight species. The bands that occur at approximately 16 kilodalton intervals are likely multimeric molecules linked by disulfide bridges. The bands migrating slightly behind the main band may be due to a

small percentage of the molecules that cannot be completely reduced by the loading buffer and therefore bind less SDS.

To determine if the other bands were due to oxidation of the 9 cysteines in the RopTat fusion (2 cysteines from Rop) during electrophoresis or to incomplete reduction prior to electrophoresis, we treated the protein samples with 50 mM iodoacetic acid (IAA) after complete denaturation and reduction in 8 M urea and 10 mM dithiothreitol (DTT)(21). IAA reacts with sulfhydryls to form stable carboxymethyl derivatives which block the reformation of disulfide bonds. Protein samples analyzed on reducing and non-reducing SDS gels following IAA treatment show a single band (data not shown).

Treatment of protein samples with IAA without prior reduction can also be used to evaluate the number of disulfide bonds and free sulfhydryls in a protein. Marks et al. (21) used this technique to compare the disulfide status of native, properly folded bovine trypsin inhibitor with the E. coli produced protein. When cells that had synthesized RopTat proteins were pelleted and rapidly dissolved in 8 M urea, 2% SDS, 50 mM IAA, we found that most (perhaps 7), but not all, of the cysteine residues of RopTat were reactive to IAA and were therefore in the free sulfhydryl form. The remaining unreactive cysteines are either involved in disulfide bonding, or are inaccessible for reaction with IAA even in the presence of 8 M urea. When the GuHCl extracts or HAP purified RopTat proteins were treated with IAA in 8 M urea without prior reduction, only

25-30% of the molecules corresponded to the sulfhydryl status of the non-extracted bacterial fusion protein. These results indicate that unless precautions are taken to avoid oxidation, the isolation of bacterially produced Tat yields a mixture of molecules with different intra- and intermolecular disulfide bond arrangements that were formed by the oxidation of the sulfhydryls during the extraction procedure.

The proteins produced by the expression system were first tested for activity using microinjection. Functional examination of bacterially expressed recombinant proteins *in vivo* by microinjection were developed shortly after several of these proteins became available in the early eighties. The technique involves skillful hand injection of the test protein in small volume (10^{-15} liters) and high concentration (0.1-1.0 mg/ml) to allow a delivery of approximately 10^6 to 10^7 molecules to individual cells and to analyze its effect *in situ*. By using this method, cellular effects of a given mutant protein, either naturally occurring or created by design, can be compared with its wild type counterpart directly. Functions of several proteins have been analyzed in this manner. For instance, V-Ha-ras p21 protein, which differs from cellular p21 in one amino acid was shown to be able to transform NIH 3T3 cells following microinjection (22). Similar observations have also been reported for oncogenic raf protein (23).

We have reported earlier that for Tat transactivation, in addition to the TAR, the *cis*-regulatory elements NF-KB and SP1

were also important (6). HIV-LTR contains two copies of NF-KB (A,B) and three repeats of SP1 (III, II, I) in similar sequences. A deletion mutant of the HIV-LTR containing only one copy of each type of *cis*-regulatory elements (NF-KB A, and SP1, I) is not a functional promoter and is not responsive for Tat transactivation in NTERA-2 cells (6). To assess whether recombinant Tat proteins derived from our Rop expression system can participate directly in transactivation of HIV-LTR in vivo, HIV template plasmid DNA containing either full-length HIV-LTR or a NF-KB (B) and Sp1 (III,II) depleted HIV-LTR (HIVΔLTR) were used. Test proteins Rop, RopTat, and cyanogen bromide (CnBr) cleaved products of RopTat (24) and RopTat41 (designated Tat and Tat41, respectively) were injected either with or without templates (HIV-LTR-CAT or HIVΔLTR-CAT) into either HeLa cells or NIH 3T3 cells. The CAT protein synthesized as the result of template expression was monitored in situ by autoradiographic analysis of the immuno-stained cells (Table I and Fig. 3). For each experiment, approximately 100 cells were injected. Nearly 80% of the cells survived the injection. A score of "++" (Table I) indicated that at least 80% of the surviving cells were CAT protein positive as shown in Fig. 3B while a score of "-" indicated that there was no CAT protein detected (Fig. 3A). From these studies, we have concluded that both RopTat protein and the CNBr cleaved Tat product from RopTat are excellent transactivators, while Rop and Tat41 proteins have no activities in vivo. This in vivo transactivation was shown to require the proximal *cis* elements of

HIV-LTR, the deletion template HIVΔLTR-CAT has no basal activity and is not responsive to Tat activation (Table 1).

In vivo cotransfection and co-injection experiments have shown that RopTat and Tat proteins produced by the Rop expression system are functional in transactivation of HIV-LTR while mutant Tat41 (lys→glu) is inactive. To examine whether such a functional difference can also be demonstrated in vitro, these proteins were tested for transcription in HeLa whole cell extract using the same HIV-LTR-CAT as template (11,12). A positive "control" mutant, RopTat42 (gly→ala) was included for comparison. High concentration of test protein stock (10-20 mg/ml) was diluted 10 to 20 fold by stepwise addition of Tat storage buffer and immediately added to the incubation mixture in the presence of HIV template DNA (12,13). Production of run-off transcripts of HIV-LTR-CAT (325 b) was analyzed by acrylamide gel electrophoresis (Fig. 4). It was found that at concentrations of 5-10 μm, RopTat42 enhances HIV-LTR-CAT transcription 32 to 37 fold, slightly better than that of full length RopTat (22 fold). In contrast, RopTat41 only enhanced HIV transcription 3-5 fold (Fig. 4). Two short nonspecific transcription bands are also seen in transcription assays using HeLa nuclear extracts with or without addition of template DNA.

In addition to the Rop expression system, several other prokaryotic expression systems have been used to produce the Tat protein. The vectors pOTS (25) and pLCBC (26), have been used to produce non-fused Tat at 2-5% of the bacterial protein. Two

systems that produce truncated proteins make use of the pGEX vector, a glutathione S-transferase gene fusion (27) and the pLE vector, a trpE gene fusion (28). Tat was produced eukaryotically using the baculovirus vector pAc 610 (29) at yields of 1-5 mg/liter of infected SF9 cells. The Tat proteins produced by these systems have been shown to be functionally active as has been demonstrated by a variety of elegant experiments (12,13,25,26). Our results with the Rop expression system demonstrate that regardless of the system used to produce Tat, isolation and purification of the protein from the highly reducing environment of the expressing cells produces a mixture of molecules with different disulfide bond arrangements. Since our purified Tat protein is active, a portion of the molecules in the purified protein sample are correctly folded and stably formed despite extensive in vitro oxidation and perhaps, disulfide interchange. Whether the correct intramolecular disulfide bonds, if there are any, were formed in vivo or upon oxidation in vitro remains to be determined.

Our observations are in general agreement with the recent report of Koken et al. (27) on the analysis of the biological activity of in vitro modified, bacterially produced Tat protein. Using a variety of reagents to assess the status of the cysteines in Tat, they concluded that a monomer is the active form of the molecule, and that two or more cysteines are involved in intramolecular disulfide bonding. They also observed that Tat rapidly oxidized to polymeric forms when special precautions were

not taken and that the transcriptionally active form of the molecule is probably a minor fraction of the Tat protein isolate.

The precise mechanism underlying Tat transactivation of HIV is currently unresolved. Recent in vitro transactivation studies have suggested that Tat enhances HIV initiation via binding to transcriptional factor TFIID (13). Other types of in vivo and in vitro experiments have also pointed out that correct Tat/TAR interaction is essential for optimal HIV-LTR expression. Deletion of four nucleotides CGAG, +35 to +38 from TAR abolishes such interaction and the Tat-dependent transactivation. However, Tat and TAR may not interact directly and such binding may also involve the participation of other cell-specific TAR binding proteins. One of these TAR binding proteins, TARP, has now been cloned and it has been shown to cooperate synergistically for TAR binding and for HIV-LTR activation in vivo (30). It is likely that Tat action depends on proper interaction with proteins specific for both TAR RNA element and proximal DNA elements such as NF-KB, Sp1, and TATA box.

Tat functions to enhance basal HIV transcription and to enhance RNA elongation. In the absence of Tat, HIV LTR template-dependent short transcripts were found to accumulate. With the expression of Tat, the long HIV transcript was observed in HIV infected human cells. A working model has been proposed recently to explain some of the results concerning HIV transcription. The model considers that there may be two overlapping promoter elements within HIV-LTR, one nonprocessive and another processive

(31,32). In the absence of Tat, the nonprocessive transcripts are made and are terminated randomly. The sizes of these transcripts are mostly short. Occasionally, however, the nonprocessive RNA synthesis may proceed to allow the formation of long transcript and the expression of Tat. The Tat protein, in turn, drives the initiation and elongation of long viral mRNA using the processive promoter of the HIV-LTR.

With the availability of gram level amounts of Tat and mutant Tat proteins and an *in vitro* system to assay them, procedures are currently being developed to isolate the properly folded molecules. The determination of the cysteines that may be involved in intramolecular disulfide bonds could give important structural information. These bonds may be responsible for the *in vitro* stability of the active, properly folded molecule. If the correct disulfide bonds were formed *in vivo*, maintaining them may facilitate the proper refolding of the molecules upon removal of the denaturants. In addition to these crystallographic studies, we have begun to study the specific molecular interactions between Tat, TAR and other transcriptional factors that may be responsible for the transactivation process.

References and Notes

1. B. Matija Peterlin, P. A. Luciw, P. J. Barr, M. D. Walker, Proc. Natl. Acad. Sci. U.S.A. **83**, 9734 (1986).
2. L. J. Seigel et al., Virology **148**, 226 (1986).
3. J. A. Garcia, D. Harrich, L. Pearson, R. Mitsuyasu, R. B. Gaynor, EMBO J. 7, 3143 (1988).
4. M. E. Sadie et al., Proc. Natl. Acad. Sci. U.S.A. **85**, 9224 (1988).
5. S. Ruben et al., J. Virol. **63**, 1 (1989).
6. R. C. Huang et al., *in* Gene Regulation and AIDS, Adv. in Applied Biotech. Series, T. S. Papas, Ed. (1989), Vol. 7, 147-160.
7. A. I. Dayton, J. G. Sodroski, C. A. Rosen, W. C. Goh, W. A. Haseltine, Cell **44**, 941 (1986).
8. A. G. Fisher et al., Nature **320**, 367 (1986).
9. J. Leonard et al., J. Virol. **63**, 4917 (1989).
10. D. Harrich, J. Garcia, R. Mitsuyasu, R. Gaynor, EMBO J. **9**, 4417 (1990).
11. R. A. Marciniak, and P. A. Sharp, EMBO J. **10**, 4189 (1991).
12. C. A. Bohan et al., Gene Exp. **2**, 391 (1992).
13. F. Kashanchi et al., Nature **367**, 295 (1994).
14. S. K. Arya, C. Guo, S. F. Josephs, F. Wong-Staal, Science **229**, 69 (1985).
15. P. E. Giza and R. C. Huang, Gene **78**, 73 (1989).
16. G. Cesareni, M. A. Muesing, B. Polisky, Proc. Natl. Acad. Sci. U.S.A. **79**, 6313 (1982).

17. R. L. Somerville, Biotechnology and Genetic Engineering Reviews Vol. 6, 1 (1988).

18. Induction growth was initiated by inoculating two liters (six liter flask) of TB medium (400 μg/ml Ap) with 0.1 ml of starter culture (15) prepared by growing co-transformed cells in LB medium supplemented with 10 μg/ml Tc and 100 μg/ml Ap. Growth in both antibiotics mantains normal copy number and trp promoter repression. Growth in Ap alone leads to loss of the trans helper plasmid, runaway plasmid replication, induction of the trp promoter, and accumulation of RopTat fusion proteins. Cells were prepared for electrophoresis by centrifugation and boiling in loading buffer (35). The cell densities at the end of induction growth were 3.2×10^9/ml (RopTat41) and 2.3×10^9/ml (RopTat42). The entire two liter cultures were harvested by centrifugation, and the insoluble inclusion bodies were isolated (15,36) and dissolved in 40 ml 6 M GuHCl, 50 mM Tris (pH 7.8), 200 mM NaCl, 0.1 mM EDTA, 5% glycerol by homogenization (40 ml Dounce). The solution was centrifuged (one hour at 35,000 RPM in Beckman 75Ti rotor), and the pellets were extracted with an additional 10 ml of GuHCl solution and centrifuged as before. The two extracts were pooled, recentrifuged, and the final GuHCl stocks were stored at a concentration of 50-80 mg/ml at -70°C.

19. K. D. Tartof and C. A. Hobbs, Bethesda Res. Lab Focus 9, 12 (1987).

20. A portion of the RopTat41 GuHCl stock was diluted by dropwise addition of a stock solution of 6 M urea, 50 mM Tris HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, (adsorption buffer) to a final protein concentration of 10 mg/ml. This solution was dialyzed at 4°C against two changes of adsorption buffer in order to eliminate the GuHCl. The dialyzed solution was mixed with HAP (22 cc hydrated bed volume, 10 g dry weight) by resuspending the HAP bed with a stream from a Pasteur pipette. After the HAP settled by gravity, it was resuspended by inverting the centrifuge tube. This was done several times over a 1 h period. The suspension was centrifuged for 10 min (1500 RPM), and the supernatant was decanted, mixed with a second bed volume of HAP (5 g dry weight), and the binding was repeated. Both HAP beds were washed 2 times with 2 bed volumes of adsorption buffer. Following the second wash and centrifugation, the first HAP bed was suspended in 50 ml (2.5 bed volumes) of adsorption buffer supplemented with 400 mM sodium phosphate (pH 8.0) to elute bound proteins (lane E1). The elution was repeated 2 more times with 11 ml (lanes E2 and E3), the three fractions were analyzed for protein content, and 4 μl of each eluant was applied to gel. Proteins from the second HAP bed were eluted twice with single bed volumes (11 ml each). The unbound protein (lane U) consists of pooled aliquots of the initial HAP binding supernatant and all HAP bed washes.

21. C. B. Marks, M. Vasser, P. Ng, W. Henzel, S. Anderson, J. Biol. Chem. **261**, 7115 (1986).

22. D. W. Stacey, and H. F. Kung, Nature **310**, 508 (1984).

23. M. R. Smith, G. Heidecker, U. R. Rapp, H. F. Kung, Mol. Cell Biol. **10**, 3828 (1990).

24. Rop protein was isolated by periplasmic extraction and further purified by HAP chromatography (15). RopTat consists of a rop gene fusion at amino acid 51 with the entire tat gene (86 amino acids) isolated from the plasmid pCV-1 (14). Cyanogen bromide cleavage reactions have been described (6,15).

25. A. Aldovini et al., Proc. Natl. Acad. Sci. U.S.A. **83**, 6672 (1986).

26. C. H. Hung, L. Tang, B. Perilli-Palmer, D. E. Heliand, E. A. Beltz, *in* Gene Regulation and AIDS, Adv. in Applied Biotech. Series, T. S. Papas, Ed., (1989) vol. 7, pp. 183-191.

27. S. E. C. Koken et al., J. Biol. Chem. **269**, 8366 (1994).

28. M. A. Muesing, D. H. Smith, D. J. Capon, Cell **48**, 691 (1987).

29. K.-T. Jeang, P. R. Shank, A. B. Rabson, A. Kumar, J. Virol. **62**, 3874 (1988).

30. A. Gatignol, A. Buckler-White, B. Berkhout, K.-T. Jeang, Science **251**, 1597 (1991).

31. M. Sheldon, R. Ratnasabapathy, N. Hernandes, Mol. Cell Biol. **13**, 1251 (1993).

32. B. R. Cullen, Cell 73, 417 (1993).

33. HeLa cells and NIH 3T3 cells were grown in DMEM (Gibco) supplemented with 10% fetal calf serum (Gibco) or 10% calf serum (NIH 3T3 cells) L-glutamine, and sodium pyruvate. Cells were plated on etched glass coverslips at a concentration of $3x10^4$/ml and grown to confluence. Cells around the etched area were microinjected with various samples of Rop, RopTat, mutant RopTat, HIV-CAT cDNA expression plasmid, and mutant HIV-CAT cDNA expression plasmid.

34. J. Vieira and J. Messing, Gene 6, 259 (1979).

35. U. K. Laemmli, Nature 227, 680 (1970).

36. H. B. Kaplan and E. P. Greenberg, Proc. Natl. Acad. Sci. U.S.A. 84, 6639 (1987).

37. This work was supported by National Institutes of Health 5R01AI32301 and by U.S. Army Medical Research DAMD17-93-C-3122.

Table I. Enhancement of HIV-LTR-CAT expression by RopTat and Tat proteins in mammalian cells. Plasmid DNA (0.5-1 mg/ml in 10 mM Tris-HCl, pH 8.0) and protein (2 mg/ml in 0.5 M Tris-HCl, pH 8.0) were mixed together and microinjected (10^{-15} L) into HeLa and NIH 3T3 cells (33). Cultures were maintained at 37°C, and the presence of CAT protein was determined 12-20 h after injection. The score "++" indicates that at least 80% of the surviving cells were CAT protein positive. The score "-" indicates that no CAT protein was detected.

Sample	CAT protein
pHIV-LTR-CAT + RopTat	++
pHIV-LTR-CAT + Tat	++
pHIV-LTR-CAT + Tat:41	-
pHIV-LTR-CAT + Rop	-
pHIV-LTR-CAT	-
pHIVΔLTR-CAT + RopTat	-
pHIVΔLTR-CAT + Tat	-
pHIVΔLTR-CAT	-
RopTat	-
Tat	-
Rop	-

Fig. 1. Construction of RopTat mutant expression plasmids. The plasmid pSVtat41 (6) was restricted with Sau3A + SstI and the *tat* gene fragment (162 bp) was cloned between the BamHI-SstI sites of pUC12 (34). The fragment was excised from pUC12 by digestion with BamHI + EcoRI (171 bp) and a 5' adapter (encoding the first four amino acids of *tat*) was ligated at the Sau3A site (15). The modified fragment (189 bp) was ligated in-frame to expression vector pPGtrpRopAp (15) to produce pPGtrpRopTat41. This plasmid was restricted with XbaI + SstI and the *tat41* gene fragment was replaced with 6 different pSVtat XbaI-SstI mutant gene fragments. Restriction of these plasmids with SstI and ligation of the oligomer 3'-TCGAATCATC to the 5' recessed ends produced double stop (TAG) codons upon treatment of the ligated adapter with Klenow frament and dNTPs. The linear plasmids were recircularized with T4 ligase, cloned, and induced cultures were screened by SDS-polyacrylamide gels (PAGE) for those that produced the properly terminated fusion proteins.

Fig. 2. Protein isolation and purification. **(A)** Host HB101[pPGtrpRopTc] was co-transformed with pPGtrpRopTat41 and pPGtrpRopTat42, separately, and cultured under conditions of trp promoter repression, lanes 1 (pPRtrpRopTat41) and 3 (pPGtrpRopTat42), and trp promoter induction, lanes 2 and 4 (same order). Total SDS soluble proteins from 6 μl of culture ($2x10^7$ cells) were analyzed by a 15.5% SDS PAGE (18)(35). **(B)** The bacteria were pelleted and lysed, and the RopTat41 fusion proteins were isolated by centrifugation (36). The partially purified protein pellets were dissolved in 6 M Guanidine HCl (GuHCl) and 12 μg protein (18) was applied to gel (lane G). Further purification of the GuHCl solubilized proteins was accomplished by batch hydroxyapatite (HAP) chromatography (20). RopTat41 was adsorbed by HAP, non-adsorbed proteins were removed by washing of the HAP bed, and RopTat fusion proteins were desorbed by three sequential elutions with 400 mM sodium phosphate (pH 8.0). The unbound (lane U) and eluted proteins (lanes E1, E2, and E3) were analyzed on gel after removal of salts by dialysis. The unbound fraction is concentrated 10 fold over the eluted fractions. Arrows denote migration of RopTat fusion proteins. Molecular sizes (M) in kilodaltons are indicated.

Fig. 3. Induction of CAT gene expression following microinjection of purified proteins. HeLa cells were co-injected with pHIV-LTR-CAT DNA + Rop protein (A), or pHIV-LTR-CAT DNA + Tat protein (B). The injected cells were maintained in growth media for 16 hrs, fixed in cold methanol, immunoabsorbed with murine anti-CAT monoclonal antibody, and rabbit anti-mouse polyclonal antibodies (Cappel). Injected cells expressing CAT protein were visualized with ^{125}I-protein A (1 μCi/ml, Amersham) after autoradiography in nuclear track emulsion (Kodak) for three days. The cells were stained with Geisma and photographed (magnification 100X).

Fig. 4. *In vitro* Tat transactivation of HIV-LTR promoter. The preparation of DNA templates for the *in vitro* transcription reactions, and the preparation of the HeLa whole cell extracts have been described (12,13). The final concentrations of reaction ingredients were: DNA templates to 8.3 μg/ml (125 ng/reaction), HeLa cell extracts to 2.4 mg/ml (40 μg/reaction), purified Tat protein (in Tat storage buffer: PBS without Ca^{++} and Mg^{++}, 0.1 percent BSA [RNAse and DNAse free], 0.1 mM DTT) to 2.5-10 μM, and nucleoside triphosphates in water to 500 mM, using 20 μCi (2 μl) of [α-^{32}P]UTP (400 Ci/mmole). The *in vitro* transcription buffer contained 10 mM HEPES (pH 7.9), 50 mM KCl, 0.5 mM EDTA, 1.5 mM DTT, 6.25 mM $MgCl_2$, and 8.5% glycerol. Reactions were incubated for 1 h at 30°C and were terminated by the addition of 20 mM Tris-HCl (pH 7.8), 150 mM NaCl, and 0.2% SDS. RNA was purified and analyzed on a denaturing 4% polyacrylamide-urea gel (12). Lane B reflects basal promoter activity. RopTat41, RopTat42, and RopTat were added to the *in vitro* transcription reaction at the amounts indicated. The size of the HIV-LTR run-off transcript (arrow) is 325 bases.

FIG. 1

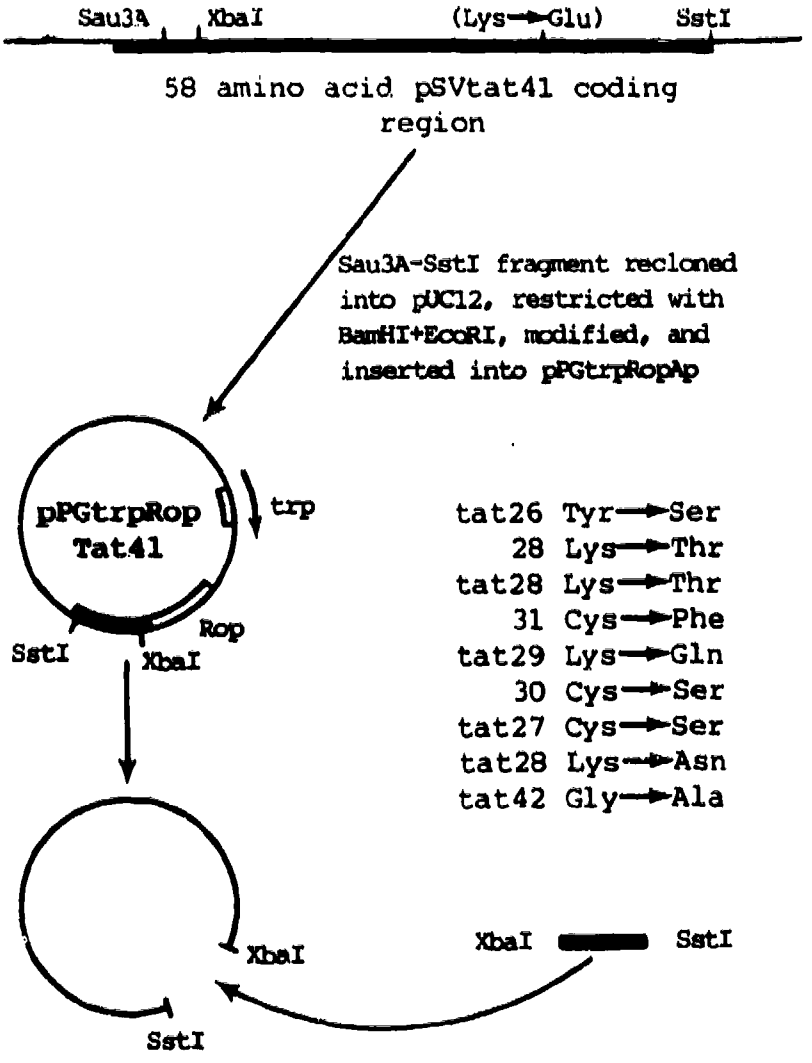
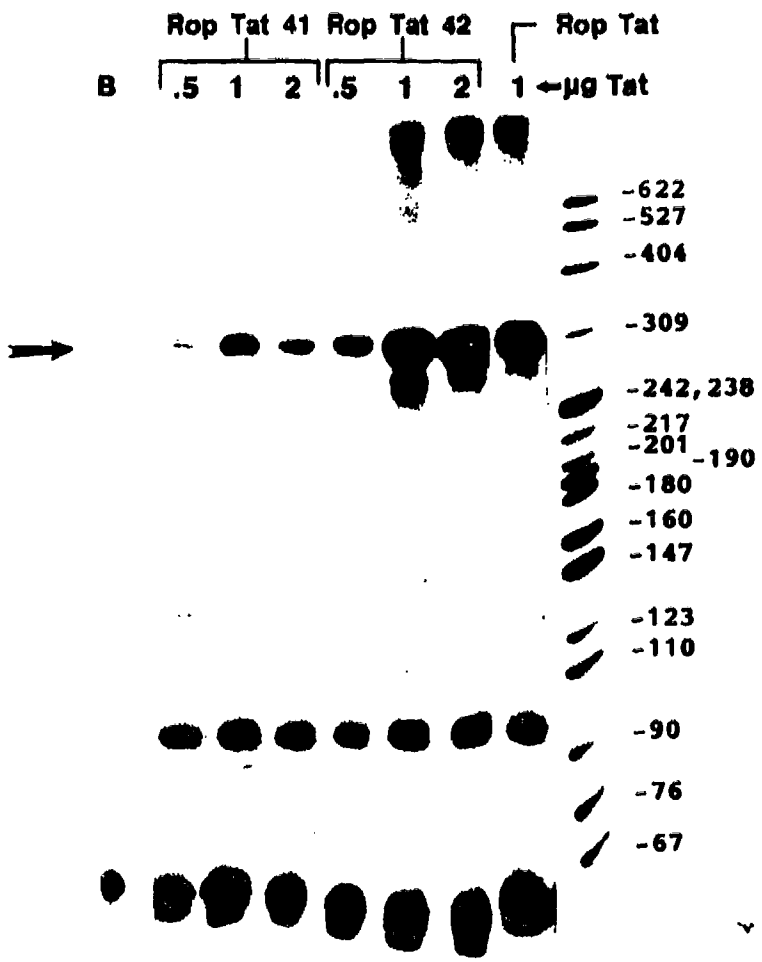

Sau3A
XbaI
(Lys→Glu)
SstI
58 amino acid pSVtat41 coding region
Sau3A-SstI fragment recloned into pUC12, restricted with BamHI+EcoRI, modified, and inserted into pPGtrpRopAp
pPGtrpRop Tat41
trp
Rop
SstI
XbaI
tat26 Tyr→Ser
28 Lys→Thr
tat28 Lys→Thr
31 Cys→Phe
tat29 Lys→Gln
30 Cys→Ser
tat27 Cys→Ser
tat28 Lys→Asn
tat42 Gly→Ala
XbaI
SstI
XbaI
SstI

FIG. 2

FIG. 3

FIG. 4


Rop Tat 41
Rop Tat 42
Rop Tat
B
.5 1 2 .5 1 2 1 ←µg Tat
-622
-527
-404
-309
-242, 238
-217
-201
-190
-180
-160
-147
-123
-110
-90
-76
-67

APPENDIX C

Annual Report (5/24/93 - 5/23/94)

Subcontract of Contract No. DAMD17-93-C-3122
Bi-Cheng Wang, Ph.D.
University of Pittsburgh

Introduction

In the past year, we have been focused on the solubilization problem of a RopTat fusion protein. RopTat contains 51 amino acids of Rop followed by the first 58 amino acids of Tat with a Gly mutated to Ala at position 42. The solubilization problem has long been associated with bacterially expressed Tat protein (Frankel et al., 1988; Harper and Logsdon, 1991). This problem is the biggest challenge in growing crystals of Tat for X-ray diffraction studies. Many attempts have been made to overcome this problem, resulting in improved methods in terms of obtaining functional and soluble protein fractions (Harper and Logsdon, 1991; Rhim et al., 1993; Waszak et al., 1992; Slice et al., 1992). However, none of these approaches has yielded pure Tat protein in high concentration for crystallization trials.

Our goal is to obtain single crystals of the RopTat fusion protein for structural study by X-ray crystallography. The first step is to develop a refolding procedure which will renature the protein from high content of denaturant, 6 M urea. The protein sample of RopTat has been overexpressed in *E. coli* as an inclusion body, dissolved and purified in the presence of 6 M urea (Giza and Huang, 1989). So we have faced a similar situation as in the solubilization of Tat. We decided to first examine the effectiveness of renaturing methods similar to those that have been used by other labs in the refolding of Tat. The details of the approaches we have tested to date and the results are described in the following section.

Body

Solubilization of RopTat Fusion Proteins

We have used the following procedure to renature RopTat from a sample containing 6 M urea. All manipulations were done at 4°C.

1. The original samples were diluted 10 times with buffer A containing 6 M urea..

2. Step-wise dialysis against:
   - 1° 5 M urea in buffer A
   - 2° 4 M " "
   - 3° 3 M " "
   - 4° 2 M " "
   - 5° 1 M urea in buffer B
   - 6° No urea in buffer B

3. Sample collected and centrifuged at 10,000 rpm for 30 min to remove the insoluble part.

Buffers:

A	B
50 mM Tris (pH 8.0)	50 mM Tris (pH 8.0)
50 mM Na_2HPO_4	50 mM Na_2HPO_4
20 mM NaCl	20 mM NaCl
1 mM EDTA	1 mM EDTA
1% Triton X-100	1% Triton X-100
200 mM βME	50 mM βME

Results:

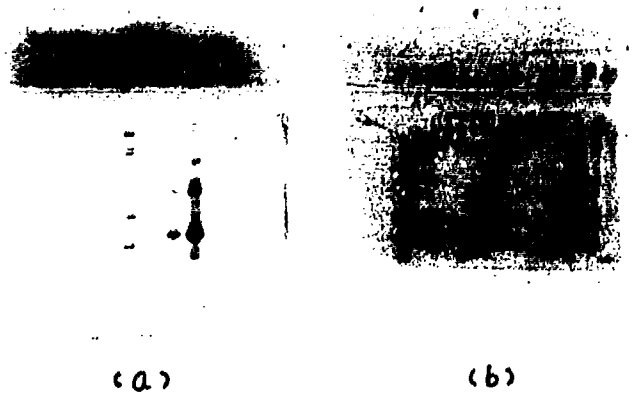
1. SDS-PAGE shows two major bands. One is apparently the RopTat. The other has a M.W. similar to that of the dimer of RopTat. Other minor bands may represent trimer, tetramer or higher oligomer as indicated by the regularity of spacing between them.

2. Recovery of RopTat by the above procedure is about 4%, based on protein assay by BCA method (Pierce).

3. The sample thus prepared remained stable in solution for at least three weeks.

Attempts to Concentrate Solubilized RopTat

RopTat solubilized using the above procedure was concentrated in Centriprep (Amicon) to about 5.8 mg/ml (BSA protein assay, Pierce). The sample then became very viscous (glycerol like). Using higher content of βME (200 mM), did not prevent the formation of high viscosity during concentration. Methyl-Hg Cl (Cystein reactive) did not help to overcome this problem. SDS-PAGE (20% gel) showed a severe trailing pattern for the concentrated sample (see below).



(a) (b)

(a)
1 Low molecular weight markers
2 Blank
3 Renatured RopTat by dialysis, ~0.15 mg/ml
4 RopTat dissolved in 6 M urea

(b)
1 RopTat dissolved in 6 M urea
2 Low molecular weight markers
3 Renatured RopTat by dialysis, ~0.1 mg/ml, sample not heated
4 Renatured RopTat by dialysis, ~1.0 mg/ml, sample not heated
5 Renatured RopTat by dialysis, ~5.8 mg/ml, sample not heated
6 Low molecular weight markers
7 Same sample as in lane 3, heated in boiling water bath
8 Same sample as in lane 4, heated in boiling water bath
9 Same sample as in lane 5, heated in boiling water bath
10 Same sample as in lane 6, heated in boiling water bath

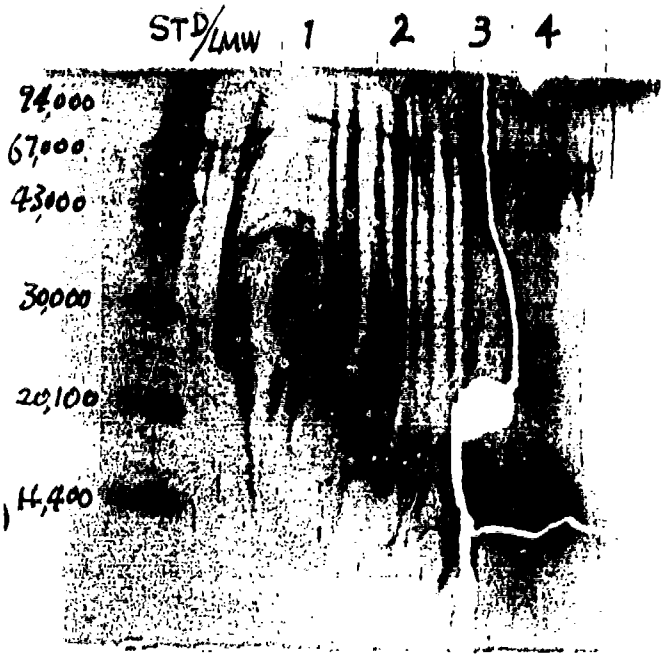
Gel-filtration of Concentrated RopTat

The concentrated sample (at ~ 5.8 mg/ml) was then run through a gel-filtration column, Superdex 75 on FPLC. One major peak appeared at elution volume corresponding to MW 65,000 Da. The peak was broad. SDS-PAGE of this fraction showed two bands, one at 67,000 Da, the other at 50,000 Da. This peak is called **p60k** in the report. There were very high molecular weight components eluted with void volume. There was a long tail following p60k all the way to the small molecule positions (see elution profiles on next page).

Problems in Concentrating Fractionated RopTat

Fractions corresponding to p60k were pooled, concentrated to ~ 0.5 mg/ml and re-run on the Superdex column (2nd); profile similar to the first Superdex column (1st) appeared. There was again a peak in the void volume and a long tail following. The relative height of p60k significantly dropped (see profiles next page).

SDS-PAGE also showed a trailing pattern for the concentrated p60k sample (see below).



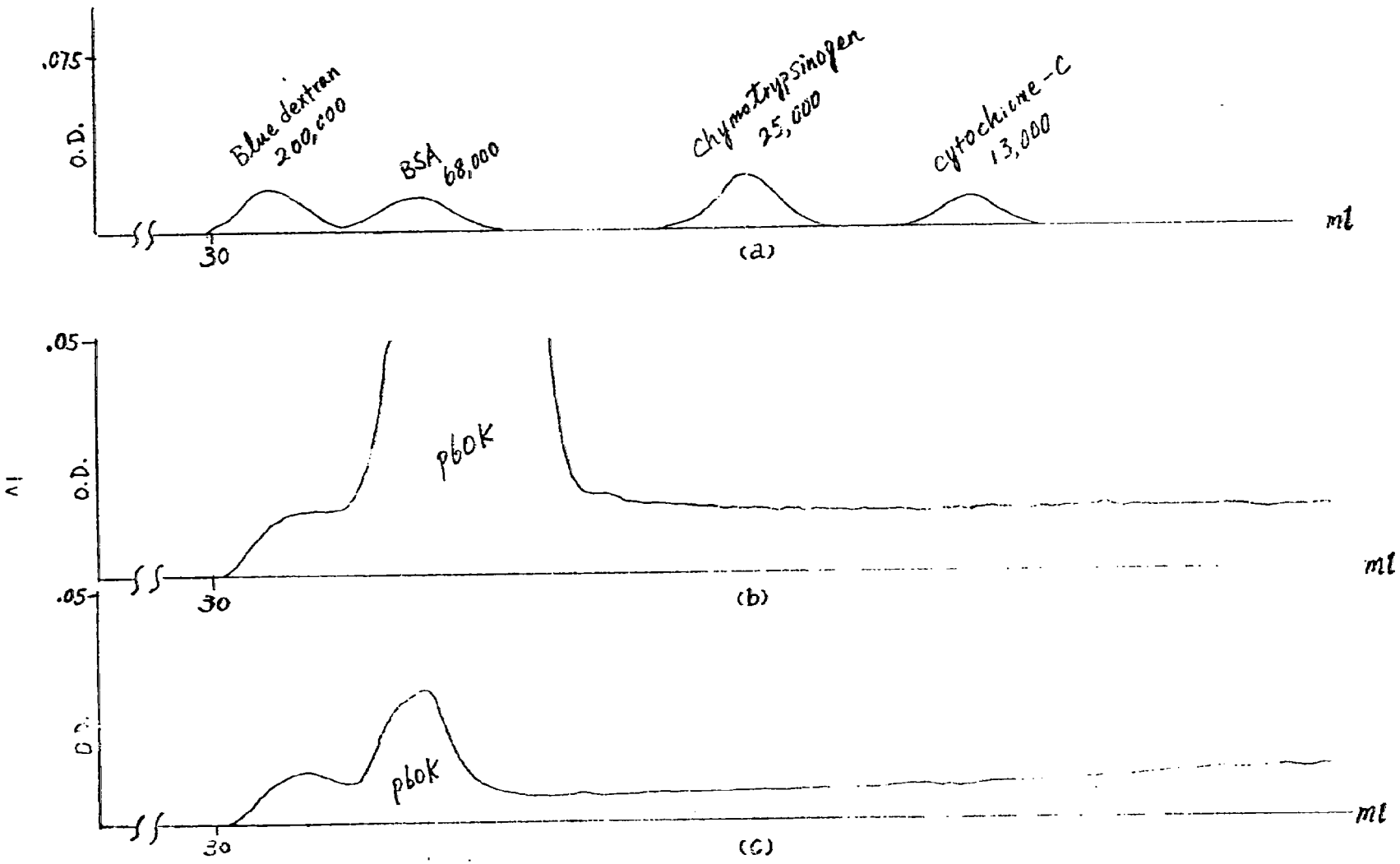
SDS-PAGE, gel percentage 15%

Lane 1 RopTat dialyzed to remove urea and concentrated to 5.8 mg/ml

2 Sample corresponding to a major peak in the gel-filtration (1st) of the protein in lane 1, ~0.5 mg/ml

3 Sample corresponding to the same peak (2nd) as in the 1st gel-filtration, unconcentrated

4 Same protein in lane 3 concentrated to ~0.1 mg/ml



Gel-filtration on Superdex 75 (Pharmacia)

(a) Standard molecular weight size markers

(b) First run, protein concentration 5.8 mg/ml

(c) Second run; fractions of the major peak in the first run were collected and concentrated to 0.5 mg/ml, run through the same gel-filtration column. Note the peak hight decreased significantly while absorptions in the void volume appeared so did that in the long tail following the major peak

Conclusion

So far we have tested commonly used renaturation conditions for recovering RopTat fusion protein from denaturing states under 6 M urea. The effectiveness of these conditions is very good in terms of gettting soluble RopTat protein samples at low concentrations. But, the protein samples produced display strong self-association upon concentrating to moderate levels, imposing a serious obstacle to crystallization attempt.

The difficulty in concentrating renatured RopTat may or may not due to theses conditions or procedures employed. The severe self-association at high concentration may be inherent in the Tat protein, assuming the so renatured RopTat adopts the same conformation as in vivo. Therefore, we should consider different conditions in which the RopTat protein displays different surface properties; for example, a ligand-bound state. Ligand peptides has been successfully used in supporting the refolding of class I histocompatibility antigen (HLA-A2) from 8 M urea (Garboczi et al., 1992). The presence of ligands helps refolding, probably by either pulling the protein conformation toward the correct one or by forming a stable complex with the correctly folded species.

We plan to conduct renaturation experiments using in vitro made RNAs carrying the TAR sequence to which Tat binds. This approach may produce a conformationally more stable protein-RNA complex with surface characteristics, more suitable for crystallization.

References

Frankel, A. D., Bret, D. S. and Pabo, C. O. (1988), Tat protein from human immunodeficiency virus forms a metal linked dimer. *Science*, **240**, 70-73.

Giza, P. E. and Huang, R. C. C. (1989), A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. *Gene*, **78**, 73-84.

Garboczi, D. N., Hung, D. T. and Wiley, D. C. (1992), HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *PNAS*, **89**, 3429-3433.

Harper, J. W. and Logsdon, N. J. (1991), Refolded HIV-1 tat protein protects both bulge and loop nucleotides in TAR RNA from ribonucleolytic cleavage. *Biochem.*, **30**, 8060-8066.

Rhim, H., Chan, F., Echetebu, C. O. and Rice, A. P. (1993), Rapid purification of monomer HIV-2 tat protein expressed in Escherichia coli. *Protein Expression and Purification*, **4**, 24-31.

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. *Biochem.*, **31**, 12062-12068.

Waszak, G. A., Hasler, J. M., McQuade, T. J., Tarpley, W. G. and Deible, Jr., M. R. (1992), Purification of an active monomeric recombinant HIV-1 trans-activator. *Protein Expression and Purification*, **3**, 126-133.