Footprinting RNA – protein complexes following gel retardation assays: application to the R-17-procoat – RNA and tat – TAR interactions

Lori Pearson, C.-h.B.Chen, Richard P.Gaynor⁺ and David S.Sigman*

Department of Biological Chemistry, School of Medicine, Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90024-1570, USA

Received March 30, 1994; Accepted April 25, 1994

ABSTRACT

RNA – protein complexes isolated following a gel retardation assay can be footprinted within the gel matrix using the chemical nuclease activities of 4,7-dimethyl-, 5,6-dimethyl-, and 3,4,7,8-tetramethyl-1,10-phenanthroline – copper. These complexes are more reactive than 1,10-phenanthroline – copper but share its reaction preference for bulges and loops. The interaction of the coat protein of R-17 with its viral RNA target and tat- and tat-derived peptides with HIV TAR RNA have been studied. In both cases, the RNA sequence opposite a 2–3 nucleotide bulge are protected. Tat-derived peptides inhibit cleavage at sites which intact tat does not protect. These results are consistent with transcription studies which have suggested that truncation of tat increases nonspecific binding.

INTRODUCTION

Gel retardation assays provide a rapid and sensitive method for characterizing the stability, stoichiometry and specificity of protein – RNA and protein – DNA interactions (1–4). We have previously demonstrated that the protein – DNA complexes isolated by gel retardation can be footprinted with the chemical nuclease activity of 1,10-phenanthroline – copper ion (OP-Cu) within the gel matrix (5, 6). This coupling of gel retardation and OP-Cu footprinting provides a convenient procedure to determine the sequence specific contacts of an electrophoretically discrete protein – DNA complex.

Interest in the analysis of the binding interactions of both cellular proteins and tat and tat-derived peptides to the TAR region of HIV RNA prompted us to extend this methodology to the study of RNA – protein interactions (7). The chemical nuclease activity of 1,10-phenanthroline – copper has previously been used as a footprinting reagent for RNA – protein interactions in solution (8–10). Its reaction preference for loops or bulges and reactivity with all four nucleotides makes it useful as a footprinting reagent for ligands which bind to these regions (11).

The chemical mechanism of the reaction involves the reversible association of the tetrahedral cuprous complex to RNA followed by the hydrogen peroxide dependent formation of the oxidative species responsible for scission (12, 13) (Eq. 1). The chemical

Equation 1

nuclease will not only detect the binding of ligands which protect the oxidatively sensitive C-1 H of the ribose but also those ligands which destroy the binding site for the tetrahedral complex. For example, the hybridization of a complementary oligonucleotide to a single-stranded loop can be assayed even though the oxidatively sensitive hydrogen atoms are fully exposed to solvent (11).

Although the procedure devised for the study of protein – DNA interactions could have been used without modification, the RNA scission activity of 1,10-phenanthroline – copper proved more sluggish in the gel matrix than the DNA cleavage activity. In an attempt to improve the procedure, the copper complexes of a variety of 1,10-phenanthroline derivatives were screened in order to determine the efficiencies of their RNase activities. Previous studies of the DNA scission activities of the copper complexes of 1,10-phenanthroline have revealed that uncharged substituents at the 5 position (6 position is sterically equivalent) of the 1,10-phenanthroline are tolerated (14) whereas only methyl groups are permitted at the 4 (or 7) position. The scission of DNA is reduced by substitution of methyl groups at the 3 (or 8) positions because they sterically hinder the close approach of the copper ion of the tetrahedral complexes to the floor of the minor groove where the oxidatively sensitive C-1H is located (11).

In this report, we have examined the reactivity of the cuprous complexes of different 1,10-phenanthroline derivatives with HIV and R-17 viral RNA fragments. The cuprous complexes of

*To whom correspondence should be addressed

⁺Present address: Department of Internal Medicine, Southwestern Medical School, 5323 Harry Hines Blvd, Dallas, TX 75235, USA

5,6-dimethyl- and 4,7-dimethyl-, and 3,4,7,8-tetramethyl-1,10-phenanthroline cleave RNA more efficiently than the copper complex of unsubstituted 1,10-phenanthroline within the gel matrix. Indeed, the methylated phenanthroline complexes cleave segments of the RNA more efficiently than the copper complex of the unsubstituted parent ligand. The copper complex of the 4,7-dimethyl-1,10-phenanthroline has been used in the gel retardation/*in situ* footprinting procedure first to examine the interaction of the R-17 coat protein with the replicase translational start site and then to assay the binding of tat and tat-derived peptides to TAR.

EXPERIMENTAL PROCEDURES

Synthesis of RNAs

HIV TAR RNA. The HIV TAR RNA was expressed from a pGEM construct and then 5'-labelled using polynucleotide kinase and γ - 32 P-ATP (7).

R-17 RNA fragment. To study the R-17 coat protein interaction, RNAs were prepared by *in vitro* transcription of synthetic DNAs using T-7 RNA polymerase (15). A Pharmacia gene assembler was used to prepare the following DNAs.

- a) 5'-TAATACGACTCACTATA-3' T7 non-template strand
b) 3'-ATTATGCTGAGTGATAT-CCTGTACTCCTAATGGGTACA-5'

T7 template strand 'U' loop

- c) 3'-ATTATGCTGAGTGATAT-CCTGTACTCCTAGTGGGTACA-5'

T7 template strand 'C' loop

The synthetic template were annealed by adding 250 pm of each oligo in 50 μ l final volume of diethylpyrocarbonate(depc)-treated water, heating to 65°C for 10 min and then placing on ice for 10 min. The DNA fragments were then transcribed with T-7 RNA polymerase in a reaction mixture containing the following components: 2 μ l annealed oligonucleotide template (10 pm); 8 μ l 5 $\times$ transcription buffer; 16 μ l NTPs each at 1 mM concentration each; 0.8 μ l RNasin (Promega); 9.2 μ l water; and 4 μ l T7 RNA polymerase (Promega) for a total volume of 40 μ l. The components of the 5 $\times$ transcription buffer were: 200 mM Tris-HCl pH 8.1 at 37°C; 25 mM DTT; 5 mM spermidine; 0.05% T $\times$ 100; 400 mg/ml PEG 8000; 100 mM MgCl₂; and 250 μ g/ml BSA. The reaction was carried out for 4 hrs at 37°C. The products were loaded onto a 20% denaturing acrylamide gel and later excised after visualization by UV shadowing. They were eluted overnight using an RNA elution buffer composed of 0.5M NH₄OAc, 10 mM MgAc, 1 mM EDTA and 0.1% SDS, ethanol precipitated, and end labelled. The labeled product was then purified again on a 20% denaturing acrylamide gel.

R-17 coat protein purification. BL21(DE3) cells with pTCT 5 (16), a gift of Dr O. Uhlenbeck, were grown to an O.D. of 0.600 in YT (per liter: 10 g bactotryptone, 5 g yeast extract, 5 g NaCl) with ampicillin added to 150 μ g/ml in a 500 ml culture. They were then treated with IPTG (400 μ M) and incubated for an additional 3 hours at 37°C. Cells were pelleted at 5000 rpm for 15 minutes, resuspended in 50 mM NH₄Cl, 20 mM Tris pH 8.0 containing 3 mM DTT, 1 mM EDTA, 5% glycerol, and lysozyme (0.45 mg/ml) and incubated at 4°C for 15 min. Then 14 mM MgCl₂ and 15 units/ml RNase free DNase (Promega) were added and incubation continued for 30 min on ice. The solution was made 200 mM in NaCl, and then purified using

Figure 1. Scission of HIV TAR RNA by the copper complexes of a series of 1,10-phenanthroline derivatives. All cleavage reactions were carried out within the acrylamide gel matrix at the indicated concentrations in a buffer containing 50 mM Tris pH 8.0; 50 mM NaCl, 1 mM MgCl₂; 5.8 mM 3-mercaptopropionic acid; 6.4 mM H₂O₂. The reaction was quenched with 2,9-dimethyl-1,10-phenanthroline.

an ECTEOLA cellulose (Sigma E-9126) column equilibrated with buffer containing 200 mM NaCl, 50 mM MOPS pH 7, 1 mM EDTA, 5 mM mercaptoethanol. Salt concentration was increased with a gradient to 1M NaCl. The fractions containing protein

were dialyzed against 1 mM acetic acid with 3 changes of solution at 4°C. The protein was stored at -70°C in 5 mM mercaptoethanol and 10% glycerol.

Preparation of tat using a glutathione-transferase fusion protein. The procedure for the isolation of glutathione-S-transferase fusion proteins provided by the vendor (Pharmacia) was used as the starting point for the preparation. The pGEX-2T vector containing the tat gene was used to transform DH5 α competent cells (BRL). Two 50 ml overnight cultures were grown from single colonies. The cultures were ten fold diluted to 500 ml and allowed to grow for 1 hr, induced with 0.1 mM IPTG and then grown for an additional 4 more hrs. The cultures were harvested by spinning at 5000 rpm for 15 min at 4°C and resuspended in 5 ml of 1 $\times$ PBS (150 mM NaCl, 16 mM NaHPO₄, 4 mM NaH₂PO₄, (pH 7.3)) + 1 mM DTT.

The cells were then sonicated and centrifuged at 12,000 rpm for 15 min at 4°C. Batch purification of the supernatant involved the following: The resin is prepared by adding a 600 μ l slurry of 50% Glutathione Sepharose 4B (Pharmacia) to each of 5 eppendorf tubes, centrifuging at 1000 rpm for 2 min at 4°C, then washing 5 $\times$ each with 300 ml of 1 $\times$ PBS + DTT. The sample is loaded onto the resin and incubated for 2 hours at 4°C. The tubes are inverted repeatedly using a tilting platform. The resins are washed 5 $\times$ each with 500 μ l 1XPBS followed by 5 $\times$ each with 500 μ l wash buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2.5 mM CaCl₂, 1 mM DTT).

To cleave the fusion protein, 500 μ l of 3.3 μ g/ml solution of thrombin dissolved in wash buffer was added to each sample. After a 40 min incubation with inversion at room temperature, the samples were centrifuged at 1000 rpm for 2 min at 4°C. The supernatants were filtered to remove all the affinity resin and dialyzed against two changes of the following buffer: 20 mM Tris pH 7.9; 0.2 mM EDTA; 100 mM KCl; 20% glycerol; 1 mM DTT; 0.5 mM PMSF. Samples were stored at -70°C.

Synthetic tat-derived peptides. The sequence of tat is:

MEPVDPKLEPWKHPGSPKTA $\overline{\text{CTNCYCKKCFHCQVCFITLALGISYGRKKRRQRRRPQGSQTHQVLSKQPTSQSKGDPTGPKG}}$

The following basic peptides were synthesized by solid phase method and purified to homogeneity by HPLC after labelling the N-terminal residue with ¹⁴C-acetic anhydride.

NH₂ $\overline{\text{CKKKRRQRRRPQGSQTHQVLSKQ}}$ -CO₂-
628
NH₂-CYGR $\overline{\text{KKKKRRQRRRP}}$ -CONH₂
935
NH₂YGR $\overline{\text{KKKKRRQRRRP}}$ C-CONH₂
930

The common sequences of the intact tat and the tat derived peptide have been underlined.

Gel retardation/in situ phenanthroline-copper footprinting

This procedure consists of two distinct components: a) gel retardation; and b) *in situ* footprinting. Gel retardations can be carried out using any detailed procedure appropriate for a given protein-RNA interaction. In the present study, the incubated samples were loaded onto polyacrylamide gels of different compositions as described below. No dyes were included in the samples.

R-17 coat protein binding conditions. The binding of the R-17 coat protein to the C-loop RNA was accomplished by incubating 100,000 cpm of 5' ³²P labeled C-loop RNA with 10 μ g of purified protein in a 50 mM Tris-HCl pH 8.5 buffer containing 40 mM KCl, 5 mM MgAc, 10% glycerol, 0.4 μ g/ μ l yeast tRNA in a final volume of 50 μ l for 20 min at 4°C. Prior to incubation, the C-loop RNA was heated to 90°C for 3 min and then allowed to cool on ice for 10 min to allow proper formation of the stem-loop structure.

The samples were loaded onto a 10% (75:1) polyacrylamide 2% glycerol 1/2 $\times$ TBE gel that had been prerun for 1 hr at 4°C. The gel was run at 250 V for 2 hrs at 4°C.

Tat protein binding conditions. Tat protein was incubated with 200,000 cpm of 5' ³²P labeled TAR RNA (+1 to +80) and up to 200 ng of protein in a buffer containing 10 mM Tris pH 7.9, 50 mM KCl, 0.1 mM EDTA, 10% glycerol, 0.5 mM DTT, 0.25 mM EDTA, 1 mM MgCl₂, 0.05 μ g/ μ l poly I/C in a final volume of 20 μ l for 20 min at room temperature.

The reaction mixture was loaded onto a 6% (29:1) polyacrylamide 2% glycerol 1/2 $\times$ TBE gel that had been prerun 1 hour at 4°C. The gel was run at 250V for 2 hours at 4°C.

Tat derived peptide binding conditions. Peptide-RNA complexes were prepared by incubating the peptide with 100,000 cpm of TAR RNA in 20 μ l of a buffer containing 70 mM NaCl, 0.2 mM EDTA, 10 mM Tris-HCl (pH 7.5), 5% glycerol, 0.1% NP40, 0.5 μ g/ μ l yeast tRNA for 30 min at room temperature. The samples were loaded onto a 10% (75:1) polyacrylamide 1/2 $\times$ TBE gel that had been prerun for 1 hr at 4°C. The gel was run at 300 V for 2 hrs at 4°C.

Detailed methods for the cleaving of protein-DNA complexes by OP-Cu have been published (5,6). The major modification of the cleavage protocol for RNA-protein complexes is the inclusion of hydrogen peroxide in the reaction mixture. Since the chemical nuclease cuts only in bulge and loop regions of the RNA, the cleavage reagent reacts less robustly with RNA than DNA. The cleavage has been carried out as follows. After the protein-RNA complexes were separated from the unbound RNA as visualized by autoradiography, the slice of acrylamide (approx. 100 μ l volume) containing the complex was excised with a razor blade and then immersed for 20 min at room temperature in a solution of the scission reagent which contains the following components: 50 mM Tris pH 8.0; 50 mM NaCl; 1 mM MgCl₂; 5.8 mM 3-mercaptopropionic acid; 6.4 mM H₂O₂; 400 μ M 4,7-dimethyl-OP and 200 μ M cupric ion. The reaction was quenched by adding 2,9-dimethyl-1,10-phenanthroline to a final concentration of 3 mM.

All components were diluted into the buffer from 10 fold more concentrated stock solutions. 2,9-Dimethyl- and 4,7-dimethyl-OP are readily soluble in absolute ethanol. Stock solutions of (4,7-dimethyl-OP)₂-Cu⁺⁺ are stable in ethanol/water at 4°C for at least one month.

After the cleavage reaction is quenched, the products are eluted overnight from the gel with 270 μ l of the following elution buffer: 0.5 M NH₄OAc; 10 mM MgAc; 1 mM EDTA; 0.1% SDS and then ethanol precipitated. The RNA is then redissolved in 10 μ l of 80% formamide containing 10% bromophenol blue and 10% xylene cyanol, heated for 5 min at 90°C, loaded on a denaturing acrylamide sequencing gel, and then run at 55 W for 2 hrs. The gel was exposed wet at -70°C with a screen.

RESULTS

RNase activities of the copper complexes of 1,10-phenanthroline and its derivatives

OP-Cu cuts RNA preferentially in single-stranded regions including bulges and loops (8, 11). The specificity is sufficiently stringent to allow experimental tests of postulated models of secondary structure (11, 17). For example, mismatches generated within double-helical stem structures can be identified as new sites for the chemical nuclease activity (11). Although the reaction of OP-Cu with RNA could be demonstrated within the acrylamide matrix and gave the same single-stranded preference observed in solution, the reaction was not efficient enough to be reliably used for the in gel cleavage reaction. As a result, we undertook a comparison of the reactivity and specificity of cuprous complexes of a variety of 1,10-phenanthroline derivatives.

Since the chemical nuclease activity of copper complexes of 1,10-phenanthroline with nucleic acids involves the reversible formation of a noncovalent intermediate followed by an oxidative step (13, 14), both the stability of the intermediate and the efficiency of the chemical cleavage step will affect the overall reactivity of the complex at any site. Substituents which favor the stability of the intermediate in the nucleolytic reaction do not necessarily increase the rate of the oxidative attack on the ribose.

The OP-Cu catalyzed oxidation of thiols by molecular oxygen provides a measure of the efficiency of this oxidative step (14, 18). Generally, conjugative substituents like phenyl and nitro groups decrease the reactivity more than methyl groups because they stabilize the cuprous complex relative to the cupric complex. Methyl substituents have the reverse effect. For example, the copper complexes 4,7-dimethyl-, 5,6-dimethyl-, 5-methyl-, and 3,4,7,8-tetramethyl-1,10-phenanthroline are more efficient catalysts for the oxidation of thiols by O₂ than the copper complexes of 5-phenyl- or 5-nitro-1,10-phenanthroline (14). The only exception to this generalization is the tetrahedral cuprous complex of 2,9-dimethyl-1,10-phenanthroline which is stabilized in the univalent state because the square-planar cupric (divalent) complex is destabilized by the steric clash of the ortho methyl groups. In fact, 2,9-dimethyl-1,10-phenanthroline has been used as an analytical reagent for copper because the stable cuprous complex has a strong visible absorption maximum at 454 nm and can be extracted readily into butanol and other water immiscible solvents (19).

The secondary structure of the TAR region of HIV RNA has been mapped by base-specific cleavage reagents, mutagenesis, and NMR (11, 20–23). This RNA was used to compare the relative scission efficiencies of the cuprous complexes of substituted 1,10-phenanthrolines. The primary sites of reactivity to 1,10-phenanthroline–copper are in bulge or loop regions. The cuprous complexes of 5,6-dimethyl-, 4,7-dimethyl-, and 3,4,7,8-tetramethyl-1,10-phenanthroline not only react at the same sites as the cuprous complex of unsubstituted 1,10-phenanthroline but also react at additional sites not measurably cleaved by the parent complex. These additional reactive sites include positions 37–39 (CUC) opposite to the bulge at positions 23–25 in a secondary structure map; positions 45–46 (CU) opposite the adenosine (position 17) bulge in the stem; and positions 55–56 (AA) opposite the cytosine bulge (sequence position 5). The cuprous complex of 3,4,7,8-tetramethyl-1,10-phenanthroline cleaves 38–40 efficiently but it is substantially less reactive than the dimethyl derivatives at positions 43–44.

HIV-1 TAR

R-17 coat protein–viral RNA interaction

The initial system used to test the gel retardation *in vitro* footprinting system was the R-17 coat protein–viral RNA interaction (24). This interaction has been extensively studied as the prototype for the sequence specific binding of a protein to an RNA (25). This interaction is central to two viral processes, the inhibition of replicase translation and the initiation of phage assembly late in infection. The phage protein binding site is located within the 21 nucleotide long hairpin shown below (24). To increase the efficiency of the *in vitro* RNA synthesis using T-7 RNA polymerase, the G-G dinucleotide at the 5' terminus has been substituted for 5' A-A dinucleotide of the wild-type. Gel retardation analysis (Figure 2A) supports previous observations which indicate that substitution of the italicized uridine in the loop by a cytosine increases the binding affinity of the wild-type coat protein by at least a factor of ten (26). At 42 μM of coat protein the 'C'-loop RNA is completely bound whereas the 'U'-loop RNA is bound about 10%.

The scission reaction of both the unbound RNA and the coat protein–'C-loop' RNA complex was carried out within the gel matrix using the cuprous complex of 4,7-dimethyl-1,10-phenanthroline as the cleavage reagent. Since the ribooligonucleotide is only 21 bases long, it is too short to provide a definitive profile of the cleavage specificity of the (4,7-dimethyl-OP)₂Cu⁺. However, the strong cut sites at the cytosines 16, 17 are consistent with the previously reported specificity of the

tetrahedral coordination complexes for bulges, loops and sequence positions which border these sites. The strong reactivities of the G's which bracket the A-bulge and are immediately 5' to the loop also fit this pattern. All the nucleotides within the loop are equivalently reactive including the C which increases the binding affinity of the coat protein.

The digestion pattern of the R-17 coat protein-RNA complex is of interest from a number of points of view (Figure 2B). First, there is no protection of the C residue which increases the binding affinity of the coat protein. Since the scission chemistry involves the ribose, this is not inconsistent with the interaction between the base and the coat protein implied by photocrosslinking data (16). However, it does indicate that this interaction does not interfere with the access of the reagent to the oxidatively sensitive bonds of the ribose. The most strongly protected sites are at C-16 and C-17 which oppose the adenosine bulge in the stem structure. Another strongly protected site is U-6 within a canonical A-helix. U-19 which is on the opposing strand in the 3'-direction and part of the initiation codon for the replicase, is cut more efficiently, possibly reflecting a protein induced conformational change within the stem structure.

Tat-TAR interaction

In previous studies of the tat-TAR interaction, the 24 amino acid basic peptide (628) was derivatized at its N-terminus with 5-iodoacetamido-1,10-phenanthroline and used to cleave the TAR target upon the activation of the nucleolytic activity by the addition of copper ion and thiol (27). The preferred sites of cleavage by this target nuclease are in the loop region and also clustered in the A-16 bulge region (27). No scission is observed in the UCU bulge region 23-25 and in the domain opposing it in the secondary structure, sequence positions 38-40.

The interaction of recombinant tat was examined with the TAR region by gel retardation/*in situ* footprinting. As demonstrated in Figure 3A, a discrete retarded protein-RNA complex was identified. The unbound RNA and the retarded band were then cleaved within the gel matrix. The clearest difference in the scission pattern was at positions 38-40, the predominant cleavage site of the cuprous complex of 4,7-dimethyl-OP (Figures 3B & C).

The binding of the 24 amino acid peptide used for the targeted scission as well as the peptides with 13 amino acids derived from the basic region were also characterized by gel retardation/*in situ* footprinting experiments. Distinct retarded bands were obtained

Figure 2. Binding of R-17 coat protein to fragments of R-17 RNA. **A.** gel retardation of R-17 coat protein with 'C'-loop and 'U'-loop forms of R-17 RNA. The indicated proteins concentration were incubated with labelled RNA in a 50 mM Tris-HCl pH 8.5 buffer containing 40 mM KCl, 5 mM MgAc, 10% glycerol, 4 μ g/ μ l yeast RNA at 4°C in a final volume of 50 μ l for 20 min. No dye was included in the incubation mixtures when the sample was loaded. **B.** *in situ* footprinting pattern provided by R-17-'C'-loop complex. A gel slice containing protein-RNA complexes was reacted with 4,7-dimethyl-OP-Cu and then analyzed on a sequencing gel.

Figure 3. Interaction of tat and TAR analyzed by gel retardation/*in situ* footprinting. **A.** (left panel) Binding of tat to HIV TAR RNA analyzed by gel retardation. 5'-³²P labeled RNA (100,000 cpm) and 200 ng of tat protein were incubated in a 20 μ l reaction composed of 10 mM Tris pH 7.9, 50 mM KCl, 0.1 mM EDTA, 10% glycerol, 0.5 mM DTT, 0.25 mM PMSF, 1 mM MgCl₂, and 0.05 μ g/ μ l poly I/C for 20 min at room temperature. Lane 1: HIV TAR RNA and tat protein; Lane 2: HIV TAR RNA. **B.** (right panel) *in situ* footprint of tat-TAR retarded band by 4,7-dimethyl-OP-Cu. The acrylamide slice containing the protein-DNA complex was footprinted and analyzed on a sequencing gel as summarized in the Experimental Section. **C.** Histogram presentation of footprint. The intensity of the bands of the tat-TAR digest were subtracted from the TAR digest.

as illustrated in the case of the 24 amino acid peptide (628) (24-C of Figure 4A). However, the pattern of protection by the peptides was different from that observed with intact tat. Although sequence positions 38–40 were strongly protected, the inhibition of scission was observed at other positions as well. For example, the shorter peptides protected sequence positions 20–25 roughly 50% as effectively as sequences 38–40 (Figure 4C).

DISCUSSION

Footprinting has played a central role in defining protein–nucleic acid interactions. DNase I has been the reagent of choice in studying interactions with DNA because of its robust cleavage and ready availability. Chemical reagents which have played a role in the analysis of ligand-DNA interactions include

Figure 4. Interaction of tat-derived peptides with TAR. **A.** Gel retardation. Gel 5'-³²P labelled HIV TAR RNA (+1 to +80) was incubated with the 24 amino acid tat peptide (628) at the indicated concentrations in 20 μ l reaction mixtures containing 70 mM NaCl, 0.2 mM EDTA, 10 mM Tris-HCl (pH 7.5), 5% glycerol, 0.1% NP40, and 0.5 μ g/ μ l t-RNA for 30 min at room temperature. Electrophoresis was carried out on a 10% (75:1) polyacrylamide gel with 1/2 $\times$ TBE. Lane a: RNA only. Lanes b & c: peptide concentration of 5 μ M and 0.5 μ M, respectively. **B.** *in situ* footprinting. Lane a: Alkaline hydrolysis of HIV TAR RNA; Lane b: T1 RNase digest of HIV TAR RNA; Lane c: HIV TAR RNA; Lanes d-f: scission of HIV TAR RNA by 4,7-dimethyl-OP/Cu; Lane d: RNA alone; Lane e: HIV TAR RNA + 24 C-tat peptide; Lane f: HIV TAR RNA + 13 tat peptide (C-COO⁻)(935); Lane g: HIV TAR RNA + 13 tat peptide (C-amino)(930). **C.** Histogram presentation of footprinting. The band intensities of the 24 amino acid peptide-TAR digest were subtracted directly from the TAR digest. a) 24 amino acid peptide '628'; b) 13C amino acid peptide '935'.

1,10-phenanthroline-copper, ferrous-EDTA, uranyl acetate and methidium propyl-EDTA (13). Chemical nucleases are advantageous because they provide higher resolution footprints owing to their small size. 1,10-Phenanthroline-copper has proven useful as a footprinting reagent because its robust reactivity and ready diffusibility allows it to be used in conjunction with gel retardation assays to obtain footprinting of electrophoretically distinct protein-DNA complexes (5).

The experiments presented here show that this procedure can now be applied to protein-RNA complexes as well. Modification of the methodology involves the use of the more reactive methylated phenanthroline-copper complexes in place of the parent 1,10-phenanthroline-copper complex and the inclusion of hydrogen peroxide in the reaction mixture. In the present case,

4,7-dimethyl-1,10-phenanthroline has been used as the ligand although 5,6-dimethyl- and 3,4,7,8-tetramethyl-phenanthroline could also serve as adequate substitutes.

The chemical nucleases of the copper complexes of substituted phenanthrolines therefore provide a diverse set of cleavage activities to study RNA structure-function interactions. Although all of these complexes cleave single-stranded loops and bulges more efficiently than double-stranded RNA, the copper complexes of 4,7-dimethyl-, 5,6-dimethyl-, and 3,4,7,8-tetramethyl-1,10-phenanthroline are not identical. The difference in the specificity between the different coordination complexes most likely arises from the stability of the essential noncovalent intermediate at a given site and not from any underlying difference in the chemistry of scission. An important

advantage of the phenanthroline-copper complex is that it cleaves at all nucleotides whereas biological nucleases are frequently specific for nucleotides as well as secondary structure. For example, RNase A is specific for pyrimidines and T₁ is specific for guanosine residues.

The results obtained with the R-17 and HIV systems studied here show an interesting similarity. In both cases, the sequences opposing a bulge are predominant sites of protection. In the case of the R-17 interaction, for example, C16 and C17 which are opposite the purine bulge (adenosine in this case) are strongly protected by the coat proteins. In the tat-TAR interaction, sequence positions 38-40, which are strongly cleaved by the copper complexes of the methylated phenanthroline derivatives, are well-protected by the binding of tat and its related peptides.

The heightened reactivity of the nucleotides which oppose the bulge may be attributable to the disruption of the stem structure caused by the bulge. The altered electrophoretic migration of RNA-DNA or RNA duplexes containing bulges relative to the canonical duplex structures provides strong physical evidence of the major distortions that they cause in stem structures (28). Molecular models suggest that the bulges widen the deep major groove of the A-helix and expose the bases more fully to solvent (4). For example, diethyl pyrocarbonate reacts with the major groove residing N-7 nitrogen of an adenine which participates in Watson-Crick base pairing when it is adjacent to a bulge but not if it is in an uninterrupted double-helix. The biochemical consequence of the greater accessibility is that the sequence specific interactions of the RNA with the protein would be facilitated. An additional consequence would be the narrowing of the minor groove of the double-stranded RNA stem which is otherwise flat and shallow and therefore a poor host for the tetrahedral cuprous complexes of the substituted phenanthrolines. Compression of the minor groove so that it approaches the geometry of the minor groove of the B-helix of DNA would increase the binding affinity for the nucleolytic coordination complexes.

These hyperreactive sites opposite the bulges will be lost if the protein disrupts the local structure imposed by the bulge, alters the minor groove, and/or sterically blocks the access of the coordination complex either to its binding site or scission site. Since the disruption of the UCU bulge by arginine binding has been demonstrated by NMR studies, the loss of local structure may be the reason for the loss of the strong cleavage sites (21). In previous studies, we have shown that a deoxyoligoribonucleotide can preferentially block scission of loop regions by the chemical nuclease presumably because it disrupts and/or blocks its binding site (11).

Intact tat only protects the nucleotides opposite the bulge at positions 38-40 (Figure 3B); no other positions are protected by the protein. In contrast, the 24 amino acid and 13-amino acid peptides (e.g. 628, 935 and 930) not only protect nucleotide 38-40 but also the nucleotides within the bulge (20-25) and the loop region (30-33). The footprints obtained in these studies are consistent with the work of Jayasena and Johnston who derivatized the 24 amino acid peptide with 1,10-phenanthroline (27). In these targeted scission experiments, the principal cleavage sites were in the A-16 bulge and the loop but not at the sequences opposing the UCU because the OP-derivatized peptide would not be able to react with its own principal binding site at sequence positions 38-40.

Initially, the greater inhibition of OP-Cu scission by the peptides appears contradictory because the protein should be

capable of more extensive interactions with the TAR binding site than the truncated peptides. These studies however are consistent with recent results of Karn and colleagues who contend that the high affinity binding observed with arginine and the short peptides includes interactions which are not functionally important (22, 29-31). Using an accurate competitive binding assay (22, 32) these workers have demonstrated that short peptides bind to specific as well as nonspecific sites. The gel retardation/*in situ* footprinting methodology therefore reveals the reduced discrimination in the binding of short peptides to TAR.

In summary, both the specificity and efficiency of the chemical nuclease activities of the copper complexes of 1,10-phenanthroline and its derivatives are a function of the structure of the ligand. Given their range of reactivities, one particular phenanthroline-copper complex may be a more useful footprinting reagent than another for a particular RNA. Since all these complexes react within the acrylamide gel matrix, these different nuclease activities can be used in the gel retardation/*in situ* footprinting protocol to determine the sequence specific contacts in those domains where the cleavage by the chemical nuclease is most robust. A clear advantage of this methodology is that it provides footprints of electrophoretically discrete protein/peptide-RNA complexes.

ACKNOWLEDGEMENTS

This research was supported by USPHS GM 39558 and USPHS GM 21199.

REFERENCES

- Garner, M. M., and Revzin, A. (1986) *TIBS* **11**, 395-396.
- Crothers, D. M. (1987) *Nature* **325**, 464-465.
- Leibold, E. A., and Munro, H. N. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2171-2175.
- Weeks, K. M., and Crothers, D. M. (1991) *Cell* **66**, 577-588.
- Kuwabara, M. D., and Sigman, D. S. (1987) *Biochemistry* **26**, 7234-7238.
- Sigman, D. S., Kuwabara, M. D., Chen, C.-h. B., and Bruice, T. W. (1991) *Methods in Enzymology* **208**, 414-433.
- Gaynor, R., Soultanakis, E., Kuwabara, M., Garcia, J., and Sigman, D. S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4858-4863.
- Murakawa, G. J., Chen, C.-h. B., Kuwabara, M. D., Nierlich, D., and Sigman, D. S. (1989) *Nucleic Acids Res.* **17**, 5361-5369.
- Murakawa, G. J., and Nierlich, D. P. (1989) *Biochemistry* **28**, 8067-8072.
- Darsillo, P., and Huber, P. W. (1991) *J. Biol. Chem.* **266**, 21075-21082.
- Mazumder, A., Gaynor, R. B., and Sigman, D. S. (1992) *Biochem. Biophys. Res. Comm.* **187**, 1503-1509.
- Sigman, D. S. (1990) *Biochemistry* **29**, 9097-9105.
- Sigman, D. S., and Chen, C.-h. B. (1990) *Ann. Rev. Biochem.* **59**, 207-236.
- Thederahn, T. B., Kuwabara, M. D., Larsen, T. A., and Sigman, D. S. (1989) *J. Am. Chem. Soc.* **111**, 4941-4946.
- Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 8783-8798.
- Gott, J. M., Willis, M. C., Koch, T. H., and Uhlenbeck, O. C. (1991) *Biochemistry* **30**, 6290-6295.
- Wang, Y.-H., Sczekan, S. R., and Theil, E. C. (1990) *Nucleic Acids Res.* **18**, 4463-4468.
- Graham, D. R., Marshall, L. E., Reich, K. A., and Sigman, D. S. (1980) *J. Am. Chem. Soc.* **102**, 5419-5421.
- Luke, C. L., and Campbell, M. E. (1953) *Anal. Chem.* **25**, 1588-1595.
- Feng, S., and Holland, E. C. (1988) *Nature* **334**, 165-167.
- Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) *Science* **257**, 76-80.
- 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) *J. Mol. Biol.* **230**, 90-110.
- Hamy, F., Asseline, U., Grasby, J., Iwai, S., Pritchard, C., Slim, G., Butler, P. J. G., Karn, J., and Gait, M. J. (1993) *J. Mol. Biol.* **230**, 111-123.

24. Carey, J., Lowary, P. T., and Uhlenbeck, O. C. (1983) *Biochemistry* **22**, 4723–4730.
25. Witherell, G. W., Gott, J. M., and Uhlenbeck, O. C. (1991) *Prog. Nucl. Acid Res. and Mol. Bio.* **40**, 185–220.
26. Lowary, P. T., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 10483–10493.
27. Jayasena, S. D., and Johnston, B. H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3526–30.
28. Bhattacharyya, A., Murchie, A. I. H., and Lilley, D. M. J. (1990) *Nature* **343**, 484–487.
29. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) *EMBO J.* **9**, 4145–4153.
30. Karn, J., and Graeble, M. A. (1992) *TIG* **8**, 365–368.
31. Graeble, M. A., Churcher, M. J., Lowe, A. D., Gait, M. J., and Karn, J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6184–6188..
32. Weeks, K. M., and Crothers, D. M. (1992) *Biochemistry* **31**, 10281–10287.