Interaction of HIV Tat model peptides with tRNA and 5S rRNA

Małgorzata Giel-Pietraszuk, Miroslawa Z. Barciszewska, Piotr Mucha, Piotr Rekowski, Gotryf Kupryszewski and Jan Barciszewski

1Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704 Poznań, Poland;
2Faculty of Chemistry, University of Gdańsk, J. Sobieskiego 18, 80-952 Gdańsk, Poland

Key words: Tat-peptides, TAR RNA, ribosomal 5S RNA, tRNA\(^{\text{Phe}}\), protein-nucleic acid interactions

New data are presented on the interaction of model synthetic peptides containing an arginine-rich region of human immunodeficiency virus (HIV-Tat), with native RNA molecules: tRNA\(^{\text{Phe}}\) of Saccharomyces cerevisiae and 5S rRNA from Lupinus luteus. Both RNA species form complexes with the Tat1 (GRKKRRQRRRA) and Tat2 (GRKKRRQRRRAFQDSQTHQASLSKQPA) peptides, as shown by electrophoretic gel shift and RNase footprint assays, and CD measurements. The nucleotide sequence UCGG located in the dihydrouridine loop of tRNA\(^{\text{Phe}}\) as well as in the loop D of 5S rRNA is specifically protected against RNases. Our data indicate direct interactions of guanine of RNA moieties with arginine residues. These interactions seem similar to those observed in DNA-protein complexes, but different from those previously observed in the TAR RNA-Tat complexes.

Over recent years the mechanism of protein-nucleic acid interactions in solutions and crystal state was the subject of numerous studies. On the basis of crystallographic data several protein structural motifs responsible for specific recognition of DNA have been found [1]. A recurring feature in all these domains is a recognition \(\alpha\)-helix, which binds to, and perfectly fits into, the major groove of B-DNA where amino-acid side chains form specific hydrogen bonds and non-specific contacts with bases and backbone phosphates, respectively. In some cases, high specificity of binding of DNA to protein through bound water molecules has been achieved [2]. Detailed analysis of crystal structures of various complexes has allowed several groups to formulate a “code” of interaction of protein transcription factors with DNA [3–5]. Generally, it predicts that the guanine and adenine residues form specific contacts with arginine and glutamine moiety, respectively. However, the data collected up to now do not confirm the existence of a similar coding system for RNA-protein complexes [6].

Recent advances in understanding of RNA-protein interactions come mostly from studies of three dimensional structure of the complexes of glutaminyl-, aspartyl- and seryl-tRNA synthetases with their cognate tRNAs [7–9], R17 coat protein with a 18 nucleotide long RNA hairpin [10] and U1A ribonucleoprotein domain with a 20 nucleotide long RNA hairpin [11]. A general conclusion drawn from those data indicates that it is rather the tertiary structure than nucleotide sequence of RNA which is the most important feature for its specific recognition by a cog-

*This study was supported by a grant from the State Committee for Scientific Research (KBN) No. PB 218/T09/95/00.

Abbreviations: ARM, arginine-rich motif; BIV, bovine immunodeficiency virus; FZ-CE, free zone capillary electrophoresis; TAR RNA, transactivation response RNA.
nate protein. On the protein part, an RNA-protein binding motif is responsible for specific interaction with ribonucleic acid. One of such domains consists of a short string of basic amino acids which contains mainly arginine residues (arginine-rich motif, ARM) [6, 12]. Similar amino-acid stretches have been identified also in other proteins, e.g., as bacterial antiterminators, ribosomal proteins, coat proteins of RNA viruses, human immunodeficiency virus Tat (HIV-Tat) and Rev (HIV-Rev) proteins [13], and a bovine immunodeficiency virus Tat (HIV-Tat) protein [14]. A Tat protein activates expression of the HIV-1 genome by its specific binding to the transactivation response RNA (TAR RNA) sequence, located at the 5' end of the untranslated leader region of HIV mRNA [15]. The latter acquires a specific secondary structure involving a helical stem interrupted by a trinucleotide bulge ended with a six-residue loop [15]. Some time ago the complex structure of TAR RNA complexed with arginine amide was solved by NMR spectroscopy [16]. Hydrogen bonds between arginine and guanine residues in position 26 and two phosphates (positions 22 and 23) have been identified. The binding of arginine induces structural changes in RNA with formation of a U23-A27-U38 base triplet [16, 17]. These data have not been confirmed in other studies of the HIV RNA-Tat, where the base triplet has not been observed [18]. The computer modeling studies of TAR RNA-Tat complex have demonstrated that it is possible to accommodate the α-helix of peptide into the RNA major groove,widened significantly by the bulge [19]. In that location of the α-helix of Tat-peptide, the guanidine group of arginine fits to the binding pocket, and side chains of amino acid interact with A22 phosphoryl group and O6 oxygen of G26, but have hydrogen bonds with the N7 atom of A22 and the O4 atom of U23. The proposed interactions in this model differ slightly from the model proposed by Puglisi et al. [16] where guanidine group of arginine forms hydrogen bonds with N7 and O6 atoms of G26, and electrostatic interactions with U23 and A22 phosphoryl group.

In order to determine which RNA domain binds to arginine-rich peptides, we used yeast tRNA$^{Phe}$ and 5S rRNA from yellow lupin as model RNA molecules. These RNA molecules were chosen because they have a 5'-UGGG sequence in some of their loops identical to that in TAR RNA. Moreover, the three-dimensional structure of yeast tRNA$^{Phe}$ is well known for over 20 years [20] and the tertiary structure of 5S rRNA has also been proposed [21]. As model proteins we used two short synthetic peptides with amino-acid sequence corresponding to the arginine-rich RNA-binding domain of HIV-1 Tat. In this paper we have shown by electrophoretic, CD and RNases protection assays that RNA molecules form complexes with these peptides, the Tat-peptides bind to the dihydouridine loop of tRNA$^{Phe}$ and to the D-loop of 5S rRNA. These interactions resemble very much the guanine-arginine recognition mode observed previously in numerous DNA-protein complexes. The data presented here are at variance with those proposed earlier for TAR-Tat, in which protein was suggested to recognize a distorted helical RNA fragment [15].

**EXPERIMENTAL**

**Synthesis of polypeptides**

Tat1 and Tat2 peptides, having the amino-acid sequences shown in Table 1, were synthesized manually on a cross linked polystyrene resin (capacity 0.68 mmol/g) by solid-phase method, using Boc chemistry. The following side-chain protection groups were employed: benzyl (Ser, Thr), 2-chlorobenzyloxy-carbonyl (1-Lys), cyclohexyl (Asp), tosyl (Arg) and Np-benzylhexymethyl (His) [22]. The peptides were cleaved from the resin by the HF procedure and lyophilized. The crude products were desalted on a Sephadex G-25 column (2.8 cm × 105 cm) in 30% CH$_3$COOH and chromatographed twice by preparative RP-HPLC on a Vyad C-18 column (32 mm × 240 mm, 15–20 mm particle size). Separation was done with the use of a linear gradient of 0–15% acetonitrile in 0.1% trifluoroacetic acid. Relative molecular ions mass of the peptides were determined by fast atom bombardment mass spectrometry method (FAB
MS) on AMD-604 mass spectrometer. Amino-
acid analysis was performed on a Beckman
model 121M analyzer. The peptide was hy-
drolyzed with hydrochloric acid containing
1% of phenol at 110°C for 24 h.

Table 1. Sequences of Tat model peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-1 Tat*</td>
<td>...39RLKLGISYGRKKRRQRRRQPDSQTHQASLSKQPA37...</td>
</tr>
<tr>
<td>Tat1 peptide</td>
<td>1GRKKRRQRRRA11</td>
</tr>
<tr>
<td>Tat2 peptide</td>
<td>1GRKKRRQRRRQPDSQTHQASLSKQPA27</td>
</tr>
</tbody>
</table>

*Data from [27].

Isolation of RNA

Ribosomal 5S RNA was isolated from yel-
low lupin (Lupinus luteus) seeds by phenol
extraction, purified on Sephadex G-75 and
repurified on 15% polyacrylamide gel con-
taining 7 M urea, 50 mM Tris/borate buffer,
ph 8.3 and 1 mM EDTA (TBE), as described
previously [23]. tRNA\textsuperscript{Phe} extracted from
yeast was additionally purified on 15% polyacrylamide gel containing 7 M urea in 50
mM TBE buffer. 5S tRNA and tRNA were
3' end labelled with \[^{32}P\]P\textsubscript{Cp} and RNA ligase
[24]. \[^{32}P\]-Labelled RNAs were purified by
10% polyacrylamide gel electrophoresis
(PAGE) in the presence of 7 M urea, eluted
from the gel and renatured [25].

RNA binding reactions and electrophoretic mobility gel shift assays

The RNA-Tat peptide complex formation assay was performed at 22°C for 40 min in a
buffer containing 50 mM Tris/HCl, pH 7.5,
70 mM NaCl, 1 mM EDTA, 0.1% Nonidet
P-40, 1 mM 5S tRNA (or tRNA\textsuperscript{Phe}), 2 mg
crude tRNA, 3 mM Tat peptide in a total
volume of 10 ml. Electrophoretic analysis of
the complexes was carried out on 0.7%
agarose gel in 50 mM Tris/borate/EDTA
(TBE) buffer.

RNase footprint assay

For a footprint reaction of the RNA-peptide
complex the following amounts of RNases
were used T1 (2 × 10\textsuperscript{-4} U), V1 (6 × 10\textsuperscript{-2} U),
S1 (3 U) (Pharmacia). For localization of Tat1
binding site, 4 mg of cold 5S tRNA or tRNA
and 40000 c.p.m. of labelled 5S tRNA or
tRNA\textsuperscript{Phe}, respectively, were digested with
0.04 U of T1 RNase in a buffer containing: 20
mM CH\textsubscript{3}COONa, pH 4.5, 7 M urea, 1 mM
EDTA and 0.05% xylene cyanol, and ana-
lyzed on 10% polyacrylamide gel in the pres-
ence of 7 M urea in 0.09 M TBE buffer.

Circular dichroism measurements

CD spectra of Tat1 and Tat2 peptides as
well as RNAs and their complexes were re-
corded at room temperature on a Jasco J-20
spectropolarimeter over an 190–310 nm in-
terval using a 1.0 mm pathlength cell. RNA-
Tat peptide 1:1 complexes were prepared in
100 mM Tris, pH 7.0, using concentrations of
5.2 × 10\textsuperscript{-5} M and 3.3 × 10\textsuperscript{-6} M for tRNA and
5S tRNA, respectively. Concentration of the
Tat peptides used for CD spectra was 2 × 10\textsuperscript{-5} M. The molar ellipticity of Tat-RNA
complexes and free peptides was calculated
per nucleotide of RNA.

Capillary electrophoresis analysis

Free zone capillary electrophoresis (FZ-CE)
(Beckman P/ACE System 2100) was used for
analysis (measured at 214 nm) of peptides
purity and RNA-Tat complexes formation.
The samples of peptides were separated in
0.1 M sodium phosphate buffer, pH 2.5, using
an uncoated fused silica capillary (75 mm ×
57 cm, 5 cm to the detector). Runs were done
at 18 kV (140 µA) and 25°C. Samples were
injected by pressure for 5 s. The tRNA\textsuperscript{Phe}
Tat1 complex formation was analyzed in 25
mM sodium phosphate buffer, pH 7.0, at
constant voltage of 30 kV (68 µA) and 25°C.
All solutions and samples were filtered
through 0.22 µm membrane filters. Some
properties of Tat fragments are summarized in Table 2.
Table 2. Physicochemical properties of Tat peptides.

<table>
<thead>
<tr>
<th>Compound</th>
<th>FAB-MAS Relative molecular mass M+H⁺</th>
<th>FZ-CE Tm [min]</th>
<th>RP-HPLC Rₑ [min]</th>
<th>Amino-acid analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>calc.</td>
<td>found</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tat1</td>
<td>C₅₉H₁₁₄N₃₂O₁₃+H⁺ 1465.8</td>
<td>8.54</td>
<td>8.8</td>
<td>Glu1.10(1), Gly1.05(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1468.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tat2</td>
<td>C₁₂₀H₂₂₇N₅₅O₇₅+H⁺ 3173.6</td>
<td>13.28</td>
<td>9.67</td>
<td>Glu5.21(5), Ser3.11(3), Lys3.00(3),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3173.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSION

Many recent studies have been focused on determination of the mechanism of protein-nucleic acids recognition. The NMR data of the HIV TAR-Tat complex indicated that an arginine residue recognizes an UCU bulge [17]. This mode of interactions is totally different from those observed in some DNA-protein complexes, where direct binding of guanine and arginine residues was observed [21]. Why the interactions within RNA-protein and DNA-protein complexes are different and why arginine residue does not interact with guanine-rich loop? To solve this question we turned to native RNA molecules for which there is a lot of structural data including their three dimensional structure, i.e. yeast tRNA^{Phe} and lupin 5S rRNA. tRNA^{Phe} from yeast was found to contain a (5')-UGGG sequence in the dihydrouridine loop. The tetranucleotide string resembles very well a part of the TAR RNA hairpin tip (Fig. 1). Two guanosine residues (G18G19), are conserved in all cytoplasmic tRNAs.

To check the specificity of RNA binding with the arginine rich peptide, we prepared two synthetic peptides, Tat1 and Tat2, 11 and 27 amino acid long, respectively (Table 1). As one can see, the amino-acid sequence of these peptides corresponds exactly with the RNA binding motif of the HIV-1 Tat protein.

![Figure 1. The secondary structure of: A, TAR RNA of HIV-1 [17], B, tRNA^{Phe} of yeast [20], C, 5S rRNA of Lupinus luteus [23]. UGGG sequence is highlighted black box.](image-url)
The first step in our studies was to check whether, or not, these peptides form complexes with tRNA\textsuperscript{Phe}. Results of the agarose gel shift assay show that both Tat1 and Tat2 bind to yeast tRNA\textsuperscript{Phe} (Fig. 2). In the presence of a three fold excess of the peptides in the reaction mixture, 100% of RNA was bound. The protein-RNA complex formation was also confirmed by capillary electrophoresis (Fig. 3). The data presented in Fig. 3 for Tat 1 indicate that the peaks corresponding to the Tat1 peptide and tRNA\textsuperscript{Phe} with migration time (M\textsubscript{r}) of 3.57 min and 17.10 min, respectively, disappear after complex formation and a new peak corresponding to the Tat 1-tRNA\textsuperscript{Phe} complex is the only one to be seen.

In order to determine the peptide binding sites on the RNA studied, a limited RNase hydrolysis with RNase T1 (G-specific), S1 (single-stranded RNA specific), V1 (double-stranded RNA specific) was carried out. For these studies we used the Tat1 peptide at two different concentrations (Figs. 4 and 5). The footprints of the RNA-Tat2 complex were so smeared as to make their interpretation practically impossible. Therefore, further footprint analysis in this paper concerns the Tat1-RNA complexes. T1 RNase hydrolysis of tRNA\textsuperscript{Phe} gave two weak bands corresponding to cuts at nucleotides G18 and G19 (Fig. 4). These cuts are absent in the digest of the complex. At the same time, there are no differences in the hydrolysis by RNase V1 of tRNA alone and in the complex. Interestingly, RNase S1, in contrast to T1 RNase, hydrolyses tRNA\textsuperscript{Phe} complexed with Tat1 at each of the three nucleotides G17G18G19 (Fig. 6A) but not the free tRNA. Results of

![Figure 2. Electrophoretic agarose gel-shift analysis of binding of Tat peptides to [3'-\textsuperscript{32}P]5S rRNA (lanes 1-3) and [3'-\textsuperscript{32}P]tRNA\textsuperscript{Phe} (lanes 4-6) on 0.7% agarose gel (lanes 2, 5 — Tat1, 3, 6 — Tat2). Arrow points the RNA-Tat peptide complexes.](image1)

![Figure 3. Free zone capillary electrophoresis of Tat1, tRNA\textsuperscript{Phe} and Tat1-tRNA\textsuperscript{Phe} complex. Analysis were carried out in uncoated fused-silica capillary (75 μm x 57 cm, 50 cm to the detector, constant voltage of 30 kV (68 μA), temp. 25°C.](image2)
hydrolysis of the complex with T1, S1 and V1 RNases clearly suggest that the Tat1 peptide interacts with dihydrouridine loop of tRNA\textsuperscript{Phe} molecule.

From the crystallographic structure of yeast tRNA\textsuperscript{Phe} it is known that G18 and G19 of the loop D and C56 and Y55 in ribothymidine loop (T) form hydrogen bonds, respectively [20]. In the case of a single-stranded RNA specific enzyme (RNase S1), lack of cleavages at G18 and G19 in the absence of Tat1 and strong hydrolysis in its presence suggest that binding of the peptide induces in tRNA\textsuperscript{Phe} some conformational changes so that the sugar-phosphate backbone in the complex becomes accessible. The guanosine residues in dihydrouridine loop are protected against T1 RNase by the peptide.

Circular dichroism spectra of free tRNA\textsuperscript{Phe} (Fig. 7A) show a Cotton effect at 266 nm which disappears after addition of the Tat1 and Tat2 peptides; this means that the conformation of tRNA\textsuperscript{Phe} becomes changed upon peptide binding. These changes may be interpreted as a disruption of the tertiary interaction between the D and T loops induced by these peptides [28, 29]. From other studies it is known that if the A-form of RNA is perturbed, the intensity of 265 nm band decreases and that of the other band at 240 nm increases [29]. These observations led us to the conclusion that the tertiary hydrogen bonds of tRNA, G19–C56 and G18–Y55, are disrupted in the complex of tRNA\textsuperscript{Phe} with Tat peptides. Therefore, we postulate that the Tat1 and Tat2 peptides should form hy-

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4}
\caption{ Autoradiogram of 10% polyacrylamide gel containing 7 M urea showing the products of [3'-\textsuperscript{32}P]tRNA\textsuperscript{Phe},Tat1 complex hydrolysis by RNase T1 (lanes 4–6), RNase VI (lanes 7–9) and RNase S1 (lanes 10–12).}
\label{fig:autoradiogram}
\end{figure}

Lanes: 1, control: 5S rRNA incubated in the reaction buffer (10 min/32°C); 2, ladder; 4, 7, 10, tRNA\textsuperscript{Phe}; 5, 8, 11, tRNA\textsuperscript{Phe} + 1.45 nM Tat1 peptide; 6, 9, 12, tRNA\textsuperscript{Phe} + 3 nM Tat1 peptide. Differences in hydrolysis pattern are indicated by arrow. Numbers on the left side correspond to nucleotides. Lane 3, T1 nuclease digestions for guanosine ladder.
Figure 5. Autoradiogram of 10% polyacrylamide gel containing 7 M urea of the products of [3'–32P]5S rRNA-Tat1 hydrolysis by RNase T1 (lanes 4–6), RNase V1 (lanes 7–9) and RNase S1 (lanes 10–12).

Lanes: 1, control: 5S rRNA incubated in reaction buffer (10 min/32°C); 2, ladder; 3, T1 nuclease digestions for guanosine ladder; 4, 7, 10, 5S rRNA; 5, 8, 11, 5S rRNA + 1.45 nM Tat1-peptide; 6, 9, 12, 5S rRNA + 3 nM Tat1-peptide. Differences in hydrolysis pattern are indicated by arrow. Numbers on the left side correspond to nucleotides.

The agarose gel shift assay indicated that 5S rRNA also forms complexes both with peptide Tat1 and Tat2 (Fig. 2). The patterns obtained on hydrolysis of 5S rRNA-Tat1 complex with RNases T1 and V1 (Fig. 5) show that nucleotides G89, G56 and 93–95 were not affected at all. However, strong bands due to digestion of the backbone at G85–G87 nucleotides by RNase S1 were observed (Figs. 5, 6A). This suggests that the Tat1 peptide interacts with D-loop and a part of stem IV of 5S rRNA molecule. According to the three-dimensional model of plant 5S rRNA, the nucleotides G85–G87 of the loop D are involved in interaction with C34–C36.
Figure 6. The secondary structure of tRNA\textsuperscript{Phe} from yeast (A) and 5S rRNA from L. lactis (B) with marked results of footprint experiments on the complex with Tat1-peptide.

Open arrows indicate the nucleotides protected from hydrolysis by RNase T1 and black ones the sites of enhanced digestion by RNase S1. Nucleotides protected from RNase V1 are in box.

of the loop C [21, 30]. The results of hydrolysis of the 5S rRNA-Tat complex (summarized in Fig. 6B) with RNases are similar to those observed with the tRNA\textsuperscript{Phe}-Tat1 complex. Inspection of CD spectra (Fig. 7B) shows that maximum of Cotton effect of 5S rRNA at 265 nm decreases in the presence of Tat peptides, as it has been observed in the case of tRNA\textsuperscript{Phe} (Fig. 7A). Interestingly, CD spectra of both RNA's complexed with Tat2 suggest that this peptide has a smaller influence on the conformation of either of them than Tat1. This observation is similar to the data obtained on interaction of poly dGm\textsuperscript{d}C with two synthetic peptides containing the KPKKKKEK sequence, 22 and 8 amino acids long, derived from lupin glutaminyl-tRNA synthetase [31]. It was observed by CD that the presence of the shorter peptide brought about changes in B-DNA conformation of poly dGm\textsuperscript{d}C while the longer peptide did not produce such changes. These data suggest that shorter peptide has a higher ability to induce conformational changes in RNA.

Figure 7. CD-spectra of complexes of Tat1 and Tat2 with tRNA\textsuperscript{Phe} (A) and 5S rRNA (B).

Full line, tRNA\textsuperscript{Phe}, 5S rRNA; dashed line, Tat2-tRNA\textsuperscript{Phe}, Tat1-5S rRNA; short dashed line, Tat1-tRNA\textsuperscript{Phe}, Tat2-5S rRNA.
which can be explained by its flexible conformation. Indeed, NMR analysis of 25 amino acids long Tat-peptide, comprising the TAR binding domain of HIV-1, showed that in aqueous solution at pH 7.0 it has α-helical structure [19]. This suggest that Tat2, due to its more stable structure, could not recognize RNA as precisely as the shorter Tat1. This conclusion is also supported by NMR analysis of arginine amide-DNA-hairpin complex. The 24-mer DNA with sequence: GATCGA-

AACGTAGCGCCCTCGATC undergoes an adaptive conformational transition after binding of amino acid and finally arginine is encapsulated by the loop folded down toward the stem [31].

In summary, the data on Tat-RNA complexes presented in this paper lead to conclusion that the interactions between nucleotides C34C35C36 of loop C and G85G86G87 of loop D in the model of plant 5S rRNA structure [30] are very similar to those observed between the loops D and T in tRNA\textsuperscript{Phe} from yeast. Another conclusion is that the UGGG sequence in the loop of RNA is a target for binding of arginine rich proteins. Finally, we postulate that the binding site for Tat protein is the guanosine rich loop and not the UCU bulge.

REFERENCES


