The nucleotide sequence of a major glycine transfer RNA from the posterior silk gland of Bombyx mori L.

Martha C. Zúñiga† and Joan A. Steitz*

†Biology Department and *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510, USA

Received 11 October 1977

ABSTRACT

The nucleotide sequence of tRNA$_{\text{Gly}}$ isolated from the posterior silk gland of Bombyx mori has been determined. This transfer RNA is present in high amounts in the posterior silk gland during the fifth larval instar. It has a GCC anticodon, capable of decoding a major glycine codon in the fibroin messenger RNA, GGU. Structural features of Bombyx tRNA$_{\text{Gly}}$ and its homology to other eukaryotic glycine tRNAs are discussed.

INTRODUCTION

The posterior silk gland of the fifth instar larva of Bombyx mori is an attractive system for studies on the regulation of protein synthesis because its major product is a single polypeptide, silk fibroin$^{1,2}$. Not only is the fibroin protein well studied, but its messenger RNA has been isolated and chemically characterized$^3$. The fibroin message exhibits selective codon utilization for the major amino acids which occur in the fibroin protein (glycine, 46%; alanine, 29%; serine, 12.9%; and tyrosine, 6%)$^4$. The triplets GGU and GGA appear in 78.5% of the glycine positions, in a ratio of 1.4 GGU to 1.0 GGA$^3$.

The unusual amino acid composition of the fibroin protein$^5$ prompted an analysis of the transfer RNAs from the posterior silk gland of the fifth instar larva. Previous aminoacylation studies showed that there is an increased amino acid acceptance activity for glycine, alanine, and serine during the maturation of the posterior silk gland$^5,6$. These observations raised several questions. First, is the increased amino acid acceptance activity in the fifth instar posterior silk gland due to the synthesis of novel isoacceptors or to the accumulation of tRNA species present in other larval tissues? Second, does the anticodon population of the posterior silk gland tRNAs reflect the pattern of codon utilization in the fibroin mRNA? Third, do the transfer RNAs which accumulate for fibroin biosynthesis possess any unusual structural features which may enhance their biological activity?
To answer these questions, $^{32}$P-labeled transfer RNAs from the posterior silk gland (and several other tissues) of Bombyx larvae were fractionated on polyacrylamide gels, identified by their aminocacylation properties, and analyzed by RNA sequencing methods. The sequences of the two major alanine isoacceptors have been reported elsewhere. The elucidation of the primary structure of one of the major glycine acceptors, tRNA$_{\text{Gly}}$, is described here.

**EXPERIMENTAL PROCEDURES**

**Sources of enzymes:** T$_1$ ribonuclease (B grade) and T$_2$ ribonuclease (A grade) were purchased from Calbiochem. Ribonuclease A (beef pancreas), beef spleen phosphodiesterase, snake venom phosphodiesterase, and micrococcal nuclease were obtained from Worthington.

U$_2$ ribonuclease was a gift of Sankyo, Ltd. P$_1$ nuclease was a gift of Alan Weiner (Yale University); S$_1$ nuclease was a gift of Roger Wiegand (Yale University); and the beef spleen phosphodiesterase used for partial digestions of primary oligonucleotides was a gift from E. G. Niles (State University of New York at Buffalo).

**Preparation of Bombyx mori RNA:** Bombyx mori were reared as described by Suzuki and Brown and Sprague.

$^{32}$P-labeled RNA was prepared from fifth larval instar posterior silk glands, middle silk glands, or midgut tissue as described previously, except that each larva was administered 2.0 to 2.5 mCi $^{32}$P$_{\text{O}}$ (carrier-free, New England Nuclear) on the fifth and sixth days after ecdysis, for a total of 4 to 5 mCi $^{32}$P$_{\text{O}}$ per larva. Animals were dissected on the eighth day of the fifth larval instar. Specific activities obtained were 2 to 6 x $10^4$ cpm/ug RNA.

Low molecular weight RNA was prepared in one of two ways: by sedimentation through sucrose gradients or by DEAE-cellulose column chromatography. When the latter method was used, total cellular RNA or low molecular weight RNA obtained by precipitation of high molecular weight RNA with 1M NaCl was fractionated by passage through DE-52 (Whatman) as described by Landy, et al., except that the buffer used throughout was 0.05M Na acetate, pH 5.0, 0.01M MgCl$_2$.

Unlabeled RNA was extracted from posterior silk glands of fifth instar larvae on the 8th day after ecdysis. Approximately 0.26 mg of total low molecular weight RNA was obtained from each pair of posterior silk glands.

**Fractionation of low molecular weight RNAs:** Bombyx mori low molecular weight RNA was fractionated in one-dimensional and two-dimensional gels using slight modifications of the systems reported by Ikemura and Dahlberg.
The one-dimensional gel (and the first dimension of the two-dimensional gel system) contained 9% acrylamide, 0.3% N, N'-bisacrylamide, and 7M urea, in the 0.09M Tris-borate, pH 8.3 buffer previously described\textsuperscript{13} except that EDTA was omitted. RNA samples were dissolved in 7M urea and then mixed with a small amount of dye solution (0.2% xylene cyanol FF in 40% sucrose) for a total volume of 20 to 40 µl per 1 cm-wide gel slot. In some cases the samples were heated to 60° C immediately prior to loading the gel; however no differences were found between electrophoretic patterns of preheated and unheated samples. Nine percent gels (14 x 40 x 0.15 cm) were run on a vertical slab apparatus\textsuperscript{14} with 0.045M Tris-borate, pH 8.3, in the upper reservoir and 0.09M Tris-borate, pH 8.3, in the lower reservoir. Electrophoresis was carried out at 4°C and 500 volts until the dye was 3 cm from the bottom of the gel.

For two-dimensional gel electrophoresis, RNAs fractionated on a 9% polyacrylamide gel were detected by autoradiography or by staining with methylene blue (0.2% in 0.2M Na acetate, pH 4.7, for 20 minutes). A gel slice containing the 4S RNAs was excised and positioned on the short side of a 14 x 16 cm gel plate. The second dimension gel (20% acrylamide, 0.45% N, N'-bisacrylamide, 0.09M Tris-borate, pH 8.3) was then formed as described by Ikemura and Dahlberg\textsuperscript{12}. Prior to electrophoresis, Trypan red (0.2% in 40% sucrose) was layered onto the top of the gel. Electrophoresis was carried out at 4°C in 0.09M Tris-borate, pH 8.3, at 450 volts until the major red-blue doublet of the Trypan red dye migrated 2/5 of the way down the gel.

Quantitation of $^{32}$P-RNAs in gel pieces was achieved by directly measuring Cerenkov radiation using a liquid scintillation counter. Calculations assume that all RNAs are equally labeled, a conclusion based on specific activity measurements on three RNA species eluted from gels (M. Zúñiga, unpublished observations).

RNA was routinely eluted from gel pieces by electrophoresis\textsuperscript{15} into dialysis casing containing 0.036M Tris-borate, pH 8.3, immersed in the same buffer made 1M in NaCl.

**Aminoacylation studies:** Cell extracts containing tRNA-aminoacyl ligase activity were standardly prepared from three pairs of posterior silk glands taken from fifth instar larvae seven days after ecdysis. The protocol used was that of White et al.\textsuperscript{16}, except that 0.20M PTU (phenylthiourea, Sigma) was included in all buffers (including the 0.15M NaCl, 0.015M Na citrate, pH 7.0, used for dissection) and the DEAE-cellulose chromatography step utilized Cellex D (Biorad) as the resin. Activity was assayed across the elution profile, using the assay reported elsewhere\textsuperscript{7}. Each fraction was made 50% in glycerol and stored up to two months at -20°C.
Amino acid acceptance activity of larval RNAs was assayed using $^{3}$H- or $^{14}$C-amino acids as described previously $^{7,17}$. Background values for each amino acid were obtained from reactions omitting either tRNA or the tRNA-aminoacyl-ligase preparation. In early experiments, tRNA was deacylated by incubation at high pH $^{7}$ prior to aminoacylation. However, values obtained with gel-fractionated tRNAs subjected to this treatment did not differ from those obtained with untreated RNAs. For calculations of percentage aminoacylation an absorbance of 1.0 OD per ml at 260 nm was assumed to represent 40 μg of tRNA.

**Sequencing procedures:** The sequence of purified tRNA$_{\text{Gly}}$ was analyzed using the methods developed by Sanger and his colleagues $^{18}$. Unless otherwise stated, digested products were quantitated by counting in toluene scintillation fluid (containing 0.42% PPO, New England Nuclear).

Products of complete T$_{1}$ RNase or RNase A digestion were fractionated by electrophoresis on cellulose acetate (Kalex) at pH 3.5 in the first dimension followed either by electrophoresis on DEAE-paper (Whatman) in 7% formic acid $^{18}$ or by homochromatography on PEI-cellulose thin layers (20 x 20 cm, Brinkman) $^{19}$. Where necessary for further analysis, oligonucleotides obtained from homochromatograms were purified of carrier RNA by chromatography in pyridine-formate, pH 3.4 and 7M urea $^{7}$ on DEAE-cellulose thin layer plates (Machery-Nagel CEL300). For short oligonucleotides (five nucleotides or less in length), 1.5M pyridine-formate was used; for long oligonucleotides (six residues or longer), 2.0M pyridine-formate.

**Analysis of primary digestion products:** Minor nucleotides (and pGp) were characterized by thin layer chromatography in two dimensions $^{20}$, 3MM paper electrophoresis $^{18}$, and descending paper chromatography $^{18,21}$ after digestion of T$_{1}$ ribonuclease oligonucleotides with T$_{2}$ ribonuclease $^{18}$, spleen phosphodiesterase $^{7}$, snake venom phosphodiesterase $^{18}$, P$_{1}$ nuclease $^{22}$, micrococcal nuclease $^{23}$, or alkaline hydrolysis $^{18,24}$. Identification of modified nucleotides was based both on the mobilities of experimental samples relative to unmodified marker nucleotides and on published mobility values $^{18,25,20}$.

Micrococcal nuclease was employed in the analysis of the dinucleotide X$^{\text{f}}$(UmC) from T$_{1}$ ribonuclease oligonucleotide 11. As described for dTpCp by Mikulski et al. $^{23}$, the sample was dissolved in 5 ul of 0.05M TrisHCl, pH 9.0, 0.01M CaCl$_{2}$ containing 5 μg micrococcal nuclease (enzyme to substrate ratio of 5:1) and incubated at 37°C for six hours. The reaction product and marker nucleotides were then spotted onto two thin layer chromatography plates (Avicel), and one dimensional chromatography was run using the two solvent systems of Nishimura $^{20}$.
P₁ nuclease was also used to analyze the dinucleotide X₁ (UmC). Since X₁ proved resistant to digestion by spleen phosphodiesterase and venom phosphodiesterase under conditions where other dinucleotides are hydrolyzed (reference 7 and this report), harsh P₁ nuclease digestion conditions were utilized. The sample (eluted from paper and estimated to contain about 1 to 2 μg RNA) was incubated with 2.5 μg enzyme in 0.04M NH₄ acetate, pH 6.0 for one hour at 50°C. The reaction product was applied to thin layer chromatography plates along with micrococal nuclease digestion products and analyzed as described above. P₁ nuclease contains a 3'-phosphomonoesterase activity, thus the major product from the dinucleotide UmC was pC.

The isomers CUCG and CCUG (T₁ ribonuclease oligonucleotide 14) were separated by one-dimensional ascending chromatography at room temperature on DEAE-cellulose thin layer plates (Machery-Nagel CEL300) in 7.5% triethylamine carbonate, pH 10.27.

Partial digestion with spleen phosphodiesterase was carried out on carrier-free oligonucleotides using 0.2 to 1.0 mg/ml enzyme in 10 μl of 0.1M NH₄-acetate, pH 5.7, 0.002M EDTA, 0.05% TWEEN 80 at 37°C for 1, 2 and 3 hours. The samples, along with an untreated control and marker oligonucleotides, were fractionated in one dimension either by chromatography on PEI-cellulose in 3% homomixture C₁9 or by electrophoresis on DEAE-paper in 7% formic acid₁8. The composition of each product was then analyzed by complete digestion with T₂ RNase or spleen phosphodiesterase. The length of the 3'-terminal oligonucleotide of tRNA₁ Gly was determined by a modification of a procedure reported by Keith and Gilham for the stepwise degradation of a polynucleotide chain. Oxidation of the sample in 15 μl H₂O at 0°C was initiated by adding 2 μl of freshly made 0.2M Na periodate. After one hour, 2 μl of 0.4M glycerol was added and the reaction incubated an additional 30 minutes at 0°C. Then 6 μl of 2.0M cyclohexamineHCl, pH 8.0, was added and the reaction incubated at 45°C for 30 minutes before analysis.

Partial digestion with T₁ ribonuclease: Limited digestion of tRNA₁ Gly with T₁ ribonuclease was carried out at 0°C for 10 minutes in 5 to 10 μl of either 0.01M TrisHCl, pH 7.5, 0.01M MgCl₂ or 0.01M TrisHCl, pH 7.5, 0.001M EDTA. In two experiments, the enzyme to substrate ratio was 1:1000 in the magnesium-containing buffer. In a third experiment, one half of the sample in the magnesium-containing buffer was digested at an enzyme to substrate ratio of 1:750, whereas the other half in the EDTA-containing buffer was treated at an enzyme to substrate ratio of 1:1500.

Partial digestion products were fractionated by pH 3.5 ionophoresis on cellogel (Kalex) in the first dimension followed by homochromatography on
Nucleic Acids Research

DEAE-cellulose thin layer plates (Analtech). Fragments ranging from 15 to 50 nucleotides were fractionated with 3% homomixture B, while more dilute homomixture B (about 1 to 1.5%) fractionated fragments ranging from 4 to 20 nucleotides in length.

Each partial digestion product was identified by complete digestion with T1 ribonuclease and ribonuclease A. The secondary digestion products, fractionated by homochromatography in 3% homomixture C in one dimension, were then analyzed by digestion with T2 ribonuclease, T1 ribonuclease (in the case of ribonuclease A digestion products) or ribonuclease A (in the case of T1 ribonuclease digestion products). This method allowed the ordering of most of the T1 ribonuclease partial digestion fragments. However, to order the C(U,C)G isomers, secondary T1 RNase products suspected of containing these oligonucleotides were also analyzed by incubation with carbodiimide reagent followed by digestion with ribonuclease A.

Digestion with S1 nuclease: Specific cleavage of tRNA$_{\text{Gly}}$ at the anticodon loop was achieved by incubating carrier-free $^{32}$P-tRNA (4 to 6 ug, 2 to 5 x $10^5$ cpm, eluted from a two dimensional polyacrylamide gel) with S1 nuclease in 10 uL of 0.09M NaCl, 0.03M Na acetate, pH 4.5, 0.001M ZnCl$_2$, 0.01M MgCl$_2$, 5% glycerol at 20°C for eight to twelve hours. An enzyme to substrate ratio of 1 unit nuclease to 1 ug tRNA was used. In some experiments, the tRNA was renatured by heating to 60°C for three minutes and then cooled on ice for five minutes before the addition of the nuclease. The inclusion of magnesium in the buffer minimized the cleavage by S1 nuclease at tRNA loop regions other than the anticodon, as determined by T1 ribonuclease fingerprint analysis of cleaved tRNA (M. Zúñiga, unpublished observations).

RESULTS

Purification and identification of Bombyx mori tRNA$_{\text{Gly}}$: Bombyx mori low molecular weight RNAs can be fractionated into approximately eight bands by electrophoresis in one dimension on a 9% polyacrylamide gel in 7M urea (Figure 1, upper panel). Fractionation in a second dimension (20% acrylamide, no urea) generates about 20 spots as detected by autoradiography of $^{32}$P-labeled RNA (Figure 1, lower panel). Note that gel bands 1, 2 and 3 of the first dimension each give rise to a single major spot in the second dimension. Fingerprint analysis of the RNAs eluted from the two dimensional gel indicated that spots 1, 2 and 3 indeed contain single RNA species which are greater than 80% pure (Figure 2 and reference 7).

Bombyx mori low molecular weight RNAs fractionated on the first dimension 9% gel were examined for their aminoacylation properties. Data for bands 1
Figure 1. Two dimensional polyacrylamide gel electrophoresis of Bombyx mori posterior silk gland tRNAs. $^{32}$P-labeled low molecular weight RNA ($4 \times 10^6$ cpm) was prepared from fifth instar posterior silk glands by DEAE-cellulose chromatography as described in Experimental Procedures. Fractionation of RNAs in the first dimension (left to right, upper panel) was by electrophoresis on a 9% polyacrylamide gel containing 7M urea and in the second dimension (top to bottom) by electrophoresis on a 20% polyacrylamide gel without urea. Xcff marks the position of the xylene cyanol FF dye in the first dimension. The positions of the major "red" and "blue" components of Trypan red in the second dimension are indicated. Aminoacylation data obtained on the first dimension bands labeled 1, 2, 3 and 4 are shown in Table I.

through 4 (which comprise the majority of the material in the 4S region, Figure 1, upper panel) are given in Table I. Bands 1 and 2 aminoacylate with alanine and have been described previously. Band 3 aminoacylates primarily with glycine (77% of the theoretical value) and is called tRNA$^\text{gly}$. The small response to alanine (4.8%) is probably due to contamination of band 3 by band 2. Band 4 also responds primarily to glycine (about 76% of the theoretical value). Together bands 3 and 4 account for about 70% of the total glycine acceptance activity in the posterior silk gland of the fifth instar larva.

The occurrence of pure RNA species in gel spots 1, 2 and 3 of the two dimensional fractionation system (Figure 1, lower panel) permitted an examination of the relative level of these RNAs in several larval tissues. Table II
Figure 2. T₁ ribonuclease fingerprint analysis of $^{32}$P-labeled tRNA$^{\text{Gly}}_3$. $^{32}$P-labeled spot 3 tRNA (about $2.5 \times 10^5$ cpm), eluted from a two dimensional gel, was digested with T₁ ribonuclease. Products were fractionated in the first dimension by electrophoresis on cellulose acetate at pH 3.5 and in the second dimension by electrophoresis on DEAE-paper (DE81) in 7% formic acid. O, B, P, and Y represent the origin and the blue, pink, and yellow dyes, respectively. Oligonucleotides are labeled as in Table III.

shows that the level of tRNA$^{\text{Gly}}_3$ (spot 3) relative to tRNA$^{\text{Ala}}_2$ (spot 2) is greater in late fifth instar posterior silk glands than it is in late fifth instar middle silk glands or midgut tissue. These data are consistent with the hypothesis that tRNA$^{\text{Gly}}_3$, like tRNA$^{\text{Ala}}_2$, is selectively synthesized for fibroin biosynthesis. Although Table II also shows an increase in the relative level of tRNA$^{\text{Gly}}_3$ during the maturation of the posterior silk gland
Aminoacylation of tRNA isolated from posterior silk glands of *Bombyx mori* fifth instar larvae

Table I. Low molecular weight RNA ("P-labeled in three preparations, unlabeled in remaining preparations) was purified by DEAE-cellulose chromatography (sucrose gradients in one experiment), fractionated on a one dimensional 9\% polyacrylamide gel and aminoacylated as described in Experimental Procedures. At least two different tRNA preparations were examined for each amino acid. Quantitation of the tRNA bands was achieved as described in Experimental Procedures; data are expressed relative to the amount of band 2 = 1. The observed percentage of aminoacylation was calculated by dividing the observed cpm incorporated into TCA-precipitable counts above background by the expected TCA-precipitable counts if 100\% aminoacylation occurred. In all experiments the specific activity of \(^3\)H- or \(^14\)C-amino acid was set so that at least 10,000 cpm were incorporated when 0.04 nM tRNA was used as acceptor. Background values were about 100 cpm.

(fourth versus fifth instar), such results are not consistently obtained (see Table II in reference 7).

Sequence analysis of tRNA\(^{\text{Gly}}\): products of complete digestion with T\(_1\) ribonuclease: Figure 2 shows a two dimensional fractionation of the digestion products generated by complete hydrolysis of spot 3 tRNA with T\(_1\) ribonuclease. All of the eighteen T\(_1\) RNase digestion products are numbered in Figure 2, with the exception of 19b which does not occur in all tRNA\(^{\text{Gly}}\) molecules. Table III
Table II. 32P-labeled molecular weight RNAs were prepared as described in Experimental Procedures and fractionated by two dimensional polyacrylamide gel electrophoresis as in Figure 1. The age of each larva was measured in days since last molt (column two). The ratios shown for tRNA1a (spot 1), tRNA2a (spot 2), and tRNA1 Gly (spot 3) are the mean values obtained from the number of independent measurements given in parenthesis in column one. Minimum Cerenkov radiation assayed was 10^6 cpm per gel spot.

<table>
<thead>
<tr>
<th>tissue</th>
<th>stage</th>
<th>tRNA1a/tRNA2a</th>
<th>tRNA1 Gly/tRNA2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>posterior silk gland (12)</td>
<td>V instar (day 7)</td>
<td>1.3</td>
<td>3.2</td>
</tr>
<tr>
<td>middle silk gland (8)</td>
<td>V instar (day 7)</td>
<td>0.2</td>
<td>2.9</td>
</tr>
<tr>
<td>midgut (12)</td>
<td>V instar (day 7)</td>
<td>0</td>
<td>1.9</td>
</tr>
<tr>
<td>posterior silk gland (6)</td>
<td>IV instar (day 5)</td>
<td>0.5</td>
<td>2.4</td>
</tr>
</tbody>
</table>

The sequences of most of the oligonucleotides from the T1 ribonuclease fingerprint were readily established by quantitation of the products of ribonuclease A and T2 digestions, supplemented by additional analyses using ribonuclease U2 and/or RNase A after chemical modification with carbodiimide reagent. These data are presented in Table III. In some cases the sequences were confirmed by complete digestion with snake venom phosphodiesterase or by partial hydrolysis with spleen phosphodiesterase (Table III). Oligonucleotides which gave special problems are discussed below.

T1 ribonuclease digestion product 5: The U*p in oligonucleotide 5 (m1-AU*UCCCG) has mobilities identical to those of Up in thin layer chromatography in two solvent systems20 and on 3 MM paper upon electrophoresis at
Table III. Products of T2 RNase digestion were fractionated by electrophoresis on 3MM paper at pH 3.5.

Other treatments include the following:
(a) U2 ribonuclease digestion
(b) ribonuclease A digestion after chemical modification with carbodiimide reagent
(c) digestion with snake venom phosphodiesterase
(d) total digestion with beef spleen phosphodiesterase
(e) partial digestion with beef spleen phosphodiesterase
(f) alkaline hydrolysis
(g) total digestion with P1 nuclease
(h) digestion with micrococcal nuclease
(i) periodate oxidation and β-elimination followed by U2 ribonuclease digestion
(j) periodate oxidation and β-elimination followed by ribonuclease A digestion
(k) treatment (b) on isolated isomers (see text for details)

Products of treatments (c), (d), (g), and (h) were fractionated by thin layer chromatography in the two Nishimura solvent systems.

† These digestions were performed on isolated dinucleotides.

<table>
<thead>
<tr>
<th>Digestion product number</th>
<th>Enzyme a</th>
<th>T2RNase</th>
<th>Other treatments</th>
<th>Sequence deduced</th>
<th>Molar yield</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.80-1.3C</td>
<td>1.20-1.5C</td>
<td>(b) 3'AU*DCG, CG, U(C)</td>
<td>m3AU*C</td>
<td>0.9 1</td>
</tr>
<tr>
<td>4</td>
<td>0.90-1.4C</td>
<td>0.96-1.5C</td>
<td>(b) 3'AU*DCG, CG, U(C)</td>
<td>m3AU*C</td>
<td>0.9 1</td>
</tr>
<tr>
<td>5</td>
<td>0.20-1.0C</td>
<td>0.98-1.1C</td>
<td>(b) 3'AU*DCG, CG, U(C)</td>
<td>m3AU*C</td>
<td>0.9 1</td>
</tr>
</tbody>
</table>

pH 3.5. However, we believe U*p to be a modified nucleotide for the following reasons:

1. The 3'-5' phosphodiester linkage of m1AU* is not cleaved by T2 RNase
under conditions which hydrolyze m^1AU isolated from B. mori tRNA^Ala^.

2. m^1AU* has a different mobility than m^1AU upon electrophoresis on DEAE-paper at pH 3.5.

3. The 3'-5' phosphodiester linkage of m^1AU* is hydrolyzed by alkali, ruling out the possibility of a 2'-O-methyl group on the adenosine residue. Moreover, the adenosine residue displays the characteristics of m^1A, including rearrangement to N^6-mA in the presence of alkali^24 and its mobility in two different systems. Thus, it is unlikely that the failure of T_2 RNase to hydrolyze m^1AU* is due to further modification of the A residue.

4. Treatment of oligonucleotide 5 with RNase A (0.3 mg/ml, 37° C, one hour) produces m^1AU*U in addition to m^1AU*, U, C, and G (Table III). Although this slow hydrolysis of the U*U bond could be due to other causes, it is nonetheless consistent with possible modification.

**T_1 ribonuclease digestion product 11:** In oligonucleotide 11 (CAX^G), the assignment of UmC to the alkali-resistant dinucleotide X^f is supported by the fact that X^f has the same mobility as the published value for CmU in four different systems: electrophoresis on 3MM paper at pH 3.5^18, descending paper chromatography in the isopropanol:HCl:H_2O system^18-21, electrophoresis on DEAE-paper in 7Z formic acid^25, and two dimensional thin layer chromatography^25. Digestion of X^f with P_1 nuclease gave pC (Table III), confirming that C is the 3' nucleotide in X^f. Micrococcal nuclease digestion of X^f gave products whose mobilities in the isobutyric acid:NH_3:H_2O and isopropanol:HCl: H_2O systems^20 are consistent with the assignment UmC. However, the latter system did not give an unambiguous result, since UmP and P^1 have the same mobility. Finally, the assignment UmC is consistent with the observation that a small percentage of the tRNA^Gly molecules are undermodified at position 4 (Um) and give free U (and increase the molar amount of C) upon RNase A or T_2 RNase digestion of oligonucleotide 11 (Table III).

**T_1 ribonuclease digestion product 14:** There are two moles of this tetranucleotide in tRNA^Gly (Table III). Analysis of the products obtained by ribonuclease A digestion with and without carbodiimide modification and of the products of snake venom phosphodiesterase suggested that this oligonucleotide is an equimolar mixture of two isomers, CUCG and CCUG (Table III). Chromatography of oligonucleotide 14 on a DEAE-cellulose thin layer plate in 7.5% triethylamine carbonate, pH 10, yielded two components. RNase A digestion after carbodiimide modification of these fractionated isomers yielded unique U- and G-containing products (Table III). The sequence assignments CUCG and

---

(4186)
CUGC were confirmed by partial digestions of oligonucleotide 14 with spleen phosphodiesterase (Table III).

T₁ ribonuclease digestion products 19 and 19b: heterogeneity at positions 8 and 9 in tRNA Gly. Most preparations of tRNA Gly contained only T₁ ribonuclease product 19, U₃G (Table III). However several preparations contained, in addition to oligonucleotide 19, another product 19b in low yields (0.4 moles, *#*40).

![Figure 3. Two dimensional fractionation of the ribonuclease A digestion products of tRNA Gly.](image)

Figure 3. Two dimensional fractionation of the ribonuclease A digestion products of tRNA Gly. ³²P-tRNA Gly (spot 3 tRNA, 2.5 x 10⁶ cpm) was obtained and fingerprinted as in Figure 2, except that RNase A rather than T₁ RNase was used. O, B, P, and Y show the positions of the origin and the blue, pink and yellow marker dyes, respectively. Oligonucleotides are numbered as in Table IV.
Table III). The occurrence of product 19b did not appear to be correlated with the developmental stage of the larva. Partial digestions of tRNA\textsubscript{Gly} with T\textsubscript{1} ribonuclease showed that 19 and 19b occurred in the same location (positions 8 and 9) in the tRNA molecule.

Oligonucleotide 19b appears to be U**G*, where U** and G* are both modified nucleotides (Table III). U** is not \( \psi \), T or S"U as determined by its mobility on thin layer chromatography\textsuperscript{20}. G* is not m\textsuperscript{2}G or m\textsuperscript{3}G on the basis of its chromatographic mobility\textsuperscript{20} or its electrophoretic mobility on 3MM paper at pH 3.5\textsuperscript{18}. It may be that G* is a m\textsuperscript{1}G residue which is further modified. A second methylation (for example at the N\textsuperscript{2} position) could occur, however no tRNA sequenced to date has m\textsuperscript{2}G. Alternatively, heterogeneity at position 9 in tRNA\textsubscript{Gly} could be due to degradation or undermodification of this G residue (see for example the yields of CGU in RNase A digestions of this RNA, Table IV). Heterogeneity at position 9 has also been reported for tRNA\textsubscript{Lys} isolated from haploid yeast\textsuperscript{30-31}; in this case m\textsuperscript{2}G occurs with 30\% frequency and G occurs with 70\% frequency in all preparations of tRNA.

T\textsubscript{1} ribonuclease digestion product 21: This oligonucleotide contains no G and therefore must be the 3' terminus of the tRNA\textsubscript{Gly} molecule. RNase A digestion consistently gave 3.9 C to 1.0 mole AC (Table III), and U\textsubscript{2} RNase digestion gave only CA and another product which yielded C upon T\textsubscript{2} RNase digestion. This suggested the sequence CACCCA\textsubscript{OH}. However, equal yields of pA and pC from snake venom phosphodiesterase (svpde) digestion of oligonucleotide 21 (Table III) were compatible with the sequence CACCA\textsubscript{OH} rather than CACCCA\textsubscript{OH}. The conflicting RNase A and svpde data were reconciled by the following experiment. Oligonucleotide 21 was subjected to periodate oxidation and B-elimination and the product chromatographed on PEI-cellulose thin layers in 3\% homomixture C, along with CCG, m\textsuperscript{5}Cm\textsuperscript{5}CG, AAUG, CCAm\textsuperscript{5}CG, and UUCAG (as size markers). The resulting oligonucleotide comigrated with m\textsuperscript{5}Cm\textsuperscript{5}CG and AAUG, suggesting it was a tetranucleotide. Subsequent RNase A and RNase U\textsubscript{2} digestion yielded products which gave the order CACC to the periodate oxidation-B-elimination product (Table III). The high yields of C obtained in RNase A (and T\textsubscript{2} RNase) digests of oligonucleotide 21 may be due to a higher specific activity of C residues in the CCA\textsubscript{OH} sequence, which is usually added to the tRNA posttranscriptionally\textsuperscript{32}. 

Products of complete digestion with ribonuclease A: Figure 3 is a ribonuclease A fingerprint of tRNA\textsubscript{Gly}. The molar yields of the digestion products are listed in Table IV. All of the RNase A oligonucleotides were sequenced by secondary digestions with ribonucleases T\textsubscript{1} and T\textsubscript{2} (Table IV), with the
exception of ribonuclease A digestion product 3, m\(^1\)GGU, which was solved by partial spleen phosphodiesterase digestions (Table IV).

Partial digestion with \(T_1\) RNase: derivation of an unambiguous sequence for \(tRNA_{Gly}^{1}\): The sequence of \(tRNA_{Gly}^{1}\) was established by overlapping the products generated by limited digestion with \(T_1\) RNase. The data obtained in three experiments are summarized in Figure 4. Each partial digestion fragment (except the two discussed below) was readily identified by analysis of its constituent products after complete RNase \(T_1\) and A digestion.

\(T_1\) RNase partial digestion product 1: The structure of this product must be \(\text{GGCG}^5\text{m}^5\text{C}^5\text{m}^5\text{CG}\) for the following reasons. Overlapping of the products of total \(T_1\) RNase (\(\text{m}^5\text{C}^5\text{m}^5\text{CG}\)) and RNase A (\(\text{GGm}^5\text{C}\)) digestion gave the partial sequence \(\text{GGm}^5\text{C}^5\text{m}^5\text{CG}\). Since \(\text{GGm}^5\text{C}\) (and not \(\text{Gm}^5\text{C}\)) was observed among the
Figure 4. Derivation of the sequence of \textit{B. mori} tRNA\textsubscript{Gly}. The brackets above and below the linear sequence indicate the T\textsubscript{1} ribonuclease partial digestion products used to derive the sequence. These were obtained from three different experiments. All products shown were obtained more than once except for CCGAUG.

**Figure 5a.** Two dimensional fractionation of a T\textsubscript{1} ribonuclease digest of the 3'-half of tRNA\textsubscript{Gly}. T\textsubscript{1} partial digestion product 3 (Figure 4) was purified of other surrounding partial digestion products by two dimensional polyacrylamide gel electrophoresis according to De Wachter and Fiers\textsuperscript{33}. The eluted sample was then digested to completion with T\textsubscript{1} RNase, and the products separated by electrophoresis on cellulose acetate (Kalex) at pH 3.5 in the first dimension followed by homochromatography in the second dimension. O, B, P, and Y refer to the origin and the blue, pink, and yellow dye markers, respectively. Note the breakdown of oligonucleotide 5 (\textsuperscript{m1}AU*UCCCG), which is accelerated by the high temperatures used in homochromatography.

**Figure 5b.** Two dimensional fractionation of a T\textsubscript{1} ribonuclease digestion of tRNA\textsubscript{Gly}. Intact tRNA\textsubscript{Gly} was digested with T\textsubscript{1} and the products fractionated as described in Figure 5a. Oligonucleotides were analysed by digestion with RNase A and T\textsubscript{2} RNase.
RNase A digestion products of T₁ partial digestion product 1, GGM₅C cannot occur at the 5' terminus and must be preceded by a pyrimidine in the fragment. This placed the T₁ RNase-generated dinucleotide, CG, 5' to Gm₅Cm₅CG to give GCGm₅Cm₅CGm₅CG. Furthermore, RNase A digestion of T₁ RNase partial digestion product 1 yielded the dinucleotide GC. Thus, the sequence of this RNA fragment is GCGm₅Cm₅CGm₅CG.

T₁ RNase partial digestion product 2: T₁ RNase partial digestion product 2 must be a mixture of two partial digestion products: C(U,C)G,CCAm₅CG and CCAm₅CGCG..., where the exact length of the latter product is unknown. The

<table>
<thead>
<tr>
<th>T₁ ribonuclease</th>
<th>molar yield tRNA₅Gly&lt;sup&gt;C&lt;/sup&gt;</th>
<th>molar yield tRNA₅Gly&lt;sup&gt;C&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>digestion product</td>
<td>control</td>
<td>S₁ nuclease-treated</td>
</tr>
<tr>
<td>e₁AU₅UCCCG</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>T₇CG</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>AUCG</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>CGCG</td>
<td>2.1</td>
<td>1.6</td>
</tr>
<tr>
<td>UGC</td>
<td>1.6</td>
<td>1.9</td>
</tr>
<tr>
<td>pG</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>U₅G</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>CCAm₅CG</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td>CACCA₅U₅C</td>
<td>1.7</td>
<td>--</td>
</tr>
<tr>
<td>e₁Ur₅U₅C₅C₅C</td>
<td>1.4</td>
<td>1.2</td>
</tr>
<tr>
<td>G</td>
<td>2.6</td>
<td>1.9</td>
</tr>
<tr>
<td>G</td>
<td>8.9</td>
<td>11.2</td>
</tr>
</tbody>
</table>

Table V. ³²P-tRNA₅Gly<sup>C</sup> (2 to 5 x 10⁴ cpm), purified by two-dimensional polyacrylamide gel electrophoresis (Figure 1), was digested with S₁ nuclease and the products fractionated from undigested material on a 9% polyacrylamide gel. Half molecules (which comigrated in this gel system) were eluted and analyzed by T₁ ribonuclease fingerprint analysis using homochromatography in the second dimension. T₁ RNase digestion products were quantitated by direct Cerenkov counting. The molar yields are average values obtained from three experiments (expressed relative to DAG = 1). The yields of Um₅G varied considerably from preparation to preparation, presumably because this oligonucleotide migrates near the front on PEI-cellulose and is sometimes lost. The identities of the T₁ RNase digestion products were confirmed by secondary analysis with RNase A.
reasons for this assignment are as follows. We obtain two moles of \( \text{CCAm}^5\text{CG} \) for each mole of \( \text{C(U,C)G} \) in complete Ti RNase digests of partial digestion product 2. Since there is only one mole of \( \text{CCAm}^5\text{CG} \) and two moles of \( \text{C(U,C)G} \) in intact tRNA\(_1^{\text{Gly}}\), product 2 must be a mixture of digestion products both of which contain \( \text{CCAm}^5\text{CG} \). In the complete RNase A digests of the mixture, the products Am\(^5\text{C} \) and GC occur in equal amounts, from which we conclude that there must also be two moles of each in partial digestion product 2. However, there is only one mole of CG present in the complete Ti RNase digests. Thus, the two RNA fragments present in partial digestion product 2 must be \( \text{C(U,C)G,CCAm}^5\text{CG} \) and \( \text{CCAm}^5\text{CG,CG} \), ..., in which the constituent Ti RNase products are not ordered. The orders given in Figure 4 are based on the observation that Ti partial product 3, which is the 3'-half of the molecule (Figure 5), contains two moles of CG (as determined by visual inspection of the autoradiograph) and one mole of \( \text{CCAm}^5\text{CG} \) but no \( \text{C(U,C)G} \). Thus, \( \text{C(U,C)G} \) must be 5' to \( \text{CCAm}^5\text{CG} \), and by process of elimination, CG must be 3' to \( \text{CCAm}^5\text{CG} \). The penta-nucleotide \( \text{CCAm}^5\text{CG} \) is placed in the anticodon loop on the basis of data obtained in experiments using S\(_1\) nuclease, which cleaves intact tRNA\(_1^{\text{Gly}}\) specifically within this oligonucleotide (Table V).

**DISCUSSION**

\( \text{tRNA}_1^{\text{Gly}} \) is a major decoding species in the mature posterior silk gland: Bombyx mori tRNA\(_1^{\text{Gly}}\), as isolated from our two dimensional gel system, was identified as a pure species which represents 10% to 20% of the total tRNA and about 40% of the glycine acceptance activity in the mature posterior silk gland.

Although this tRNA is not unique to the posterior silk gland, we conclude that tRNA\(_1^{\text{Gly}}\) accumulates to meet the needs of rapid fibroin synthesis for the following reasons:

(a) The relative amount of tRNA\(_1^{\text{Gly}}\) is greater in the mature posterior silk gland than in other larval tissues.

(b) The derivation of the complete sequence of tRNA\(_1^{\text{Gly}}\) shows that the molecule has a GCC anticodon, which is theoretically capable of decoding the major glycine codon in the fibroin mRNA, GGU. Moreover, comparison of our data with codon-response studies on glycine-accepting tRNAs fractionated by counter-current distribution\(^3\) suggests that tRNA\(_1^{\text{Gly}}\) is the major GGU-decoding species in the posterior silk gland of B. mori.

The occurrence of GCC rather than an ICC anticodon (which could also pair with GGU) in glycine tRNAs was predicted by Ninio\(^3\) on the basis of kinetic and fidelity arguments. ACC would not be expected to appear since A is never
Analysis of the frequency of codon usage in MS2 bacteriophage RNA has led to the suggestion (H. Grosjean, R.J. Cedergren, D. Sankoff, W. Min-Jou, W. Fiers, personal communication) that there is an optimal half-life for codon-anticodon interaction. The occurrence of a GCC anticodon in a major posterior silk gland glycine tRNA and of a GGU codon as the major glycine codon in the fibroin mRNA would appear to support this hypothesis. Also consistent is the fact that the glycine codon GCC occurs in only 10.4% of the possible glycine positions in the fibroin mRNA.

Another interesting feature of the tRNA\textsubscript{Gly} sequence is the appearance of a U residue 5' and an unmodified A residue 3' to the anticodon triplet. While these nucleotides are not unexpected in these positions, their potential contribution to tRNA-mRNA interaction can now be evaluated since the crystalline portion (greater than 60%) of the silk fibroin protein is known to have the repeating sequence:

\text{Gly-Ala-Gly-Ala-Gly-Ala-Gly-Gly-Tyr.}

Thus, tRNA\textsubscript{Gly} could potentially form five base pairs with the fibroin mRNA when glycine resides between two alanines, for which the major codon is GCU; four base pairs could be formed when glycine occurs between an alanine and a serine residue, for which the major codon is UCA. These observations are compatible with ideas concerning the current nature of the protein synthetic apparatus.

Structural homology of B. mori tRNA\textsubscript{Gly} to other glycine tRNAs: The sequence of eight prokaryotic and three eukaryotic (yeast and wheat germ) are now known. On the basis of these sequences one can generalize that glycine tRNAs tend to be small (73 to 76 nucleotides in length), to have small variable loops (three to five nucleotides) and to have few modified residues. Otherwise, there are no clear sequence homologies among all the various prokaryotic and eukaryotic glycine tRNAs.

Among the three eukaryotic glycine tRNAs (all of which have a GCC anticodon) distinct similarities do exist. Excluding posttranscriptional modifications (modified bases and the CCA\textsubscript{OH} end), B. mori tRNA\textsubscript{Gly} has 66% homology with yeast tRNA\textsubscript{Gly} and 78% homology with wheat germ tRNA\textsubscript{Gly}, suggestive of an evolutionary relationship. Interestingly, the homologies among these three tRNAs are localized in certain areas (Figure 6): loop I (the D loop), loop IV (T\textsuperscript{4}C stem) and the T\textsuperscript{4}C stem. The acceptor stems also exhibit some common sequences, but the major feature of interest here is a posttranscriptional
Figure 6. Cloverleaf structure of \( B. \) mori tRNA\(_{\text{Gly}}^{\text{Gly}} \). Regions of homology among the three eukaryotic glycine tRNAs sequenced to date are indicated by boxed areas. Arrows denote differences in modification where there is \( \text{Gly} \) sequence homology; in each case only one of three is different. Yeast tRNA\(_{\text{Gly}}^{\text{Gly}} \) (73 nucleotides in length) has a three-membered extra loop (loop III), whereas \( B. \) mori tRNA\(_{\text{Gly}}^{\text{Gly}} \) and wheat germ tRNA\(_{\text{Gly}}^{\text{Gly}} \) (74 nucleotides in length) each have a four-membered extra loop.

Modification at position 4. All three molecules have a pyrimidine with a 2'-O-methyl moiety on the ribose group at this position. A modification of this type in the acceptor stem is thus far unique to these tRNAs. It is tempting to suggest that the above mentioned homologous features may be important to recognition by the glycyl tRNA aminoacyl-ligase.

In addition, \( B. \) mori tRNA\(_{\text{Gly}}^{\text{Gly}} \) shares two further sequence homologies with wheat germ tRNA\(_{\text{Gly}}^{\text{Gly}} \): identical anticodon loops with \( \text{m}^5\text{C} \) at position 37 and three additional adjacent 5-methyl cytidine residues in positions 46 to 48 (Figure 6) of their sequences (corresponding to positions 48 to 50 in the standard tRNA structure\(^51\,52\)). Five-methyl-cytidine has been found only in eukaryotic tRNAs. In 29 of 35 known sequences it occurs at position 48 (where it presumably participates in the Levitt base pair\(^51\)) or at position 49 (in the WC stem) of the standard tRNA structure. Of the six eukaryotic tRNAs lacking \( \text{m}^5\text{C} \), five are from the yeasts (Saccharomyces cerevisiae and Torulopsis utilis). Thus \( \text{m}^5\text{C} \) at positions 48 and/or 49 or the standard tRNA structure appears to be a general feature of tRNAs in "higher" eukaryotes, though it has yet to be assigned a function.
Although the occurrence of one or two m^5C residues is common in eukaryotic tRNAs, the four m^5C residues in tRNA^Gly\textsubscript{Bombyx mori} and wheat germ are novel. It has been suggested previously that the high content of m^5C in wheat germ tRNA^Gly, coupled with its lack of ribothymidine at U_52, allows it to function more efficiently in protein synthesis. The fact that B. mori tRNA^Gly possesses both ribothymidine and a high m^5C content (and yet is presumably highly active on the ribosome) suggests that the functional relationship of these two methylated residues may be more complex. Studies of the protein synthetic capabilities of B. mori tRNA^Gly are needed to evaluate the role of these modified residues in tRNA function.

Recently, Garel and Keith, using a different approach, deduced a comparable sequence for B. mori tRNA^Gly.

ACKNOWLEDGEMENTS

We thank N. Farber, P.M.M. Rae, D. Söll and K.U. Sprague for advice, and M. Krikeles and C. Stocking for assistance. This work was supported by National Science Foundation Grant PCM74-01136 and forms part of a dissertation submitted by M.C.Z. to the Graduate School of Yale University in partial fulfillment of the requirements for the PhD degree. M.C.Z. was a recipient of the Ford Foundation Fellowship for Mexican Americans.

REFERENCES

1 Sprague, K.U. (1975) Biochem. 14, 925-31
13 Peacock, A.C. and Dingman, C.W. (1967) Biochem. 6, 1818-1827
17 Zuñiga, M.C. (1977) PhD Thesis, Department of Biology, Yale University, New Haven
20 Nishimura, S. (1972) Prog. Nucleic Acid Res. Mol. Biol. 12, 49-85
50 Marcu, K.B., Mignery, R.E. and Dudock, B.S. (1977) Biochem. 16, 797-806