Volume 1 number 10 October 1974 Nucleic Acids Research

Nucleotide sequence determination of bacteriophage T4 glycine transfer ribonucleic acid

Stephen Stahl*, Gary V. Paddock, and John Abelson

Department of Chemistry, University of California, San Diego, La Jolla, California 92037, USA

Received 31 July 1974

ABSTRACT

The nucleotide sequence of a T4 tRNA with an anticodon for glycine has been determined using ^{32}P-labeled material from T4-infected cultures of Escherichia coli. The sequence is: pGCGGAUAUCGUAUAAUGmGDAUUACCUCAGACUUCCA-AψCUGAUGAUGUGAGTψCGAUUCUCAUUAUCCGCUCCA-OH. The 74 nucleotide sequence can be arranged in the classic cloverleaf pattern for tRNAs. The anticodon of T4 $tRNA^{Gly}$ is UCC with a possible modification of the U. The tRNA molecule would thus be expected to recognize the glycine codons GGG and GGA. Comparative analysis of $tRNAs^{Gly}$ from T2 and T6 indicate that their sequences are identical with that from T4.

INTRODUCTION

When bacteriophage T4 infects Escherichia coli, several new stable RNA species which include eight new tRNAs are transcribed from the T4 genome.[1-5] The RNA species produced are clustered in a small region of the T4 genome between genes e and 57.[5] Under conditions of T4 infection, there is no host synthesis and the T4 RNA species can thus be selectively labeled.

Polyacrylamide gel electrophoresis of ^{32}P-labeled RNA extracted from T4-infected cells results in the characteristic band pattern seen in Fig. 1.[4,6-8] It is known that one of the T4 tRNAs accepts glycine,[3] and we have previously reported the sequence of this $tRNA^{Gly}$ extracted from 10% polyacrylamide gel band 6.[9] We report here the details of the sequence determination of T4 $tRNA^{Gly}$ and the comparative analysis of $tRNA^{Gly}$ coded for by bacteriophages T2 and T6.*

MATERIALS AND METHODS

The materials and methods used here are the same as described previously,[7] with the following additions and modifications.

*Subsequent to our preliminary report,[9] the following article was published by Barrell et al. reporting on an independently determined nucleotide sequence of T4 $tRNA^{Gly}$ in agreement with the sequence determination reported here: Barrell, B. G., Coulson, A. R. and McClain, W. H. (1973) FEBS Lett. 37, 64-69.

Enzymes. Deoxyribonuclease was obtained from Worthington Biochemical Corp.

^{32}P-labeled T4 tRNA. One-tenth volume of cold 20X standard saline citrate (3.0 M NaCl and 0.3 M sodium citrate) was added to the infected culture before the cells were harvested by centrifugation. After centrifugation the cells were resuspended in a total volume of 4.0 ml of 0.01 M Tris pH 7.4, 0.1 M NaCl, 0.002 M mercaptoethanol. The cells were then treated with 10 µl of 2 mg/ml DNase in 0.025 M sodium acetate pH 4.0, 0.002 M $MgCl_2$, 50% glycerol. After incubating 30 min at 25°C, the cells were extracted with phenol as described previously.

The unfractionated T4 [^{32}P]tRNA was purified only through the DEAE-cellulose step and was then electrophoresed on a 10% polyacrylamide gel. T4 tRNA^{Gly} was eluted directly from the gel and prepared for sequence analysis.

It was desirable to have especially pure T4 [^{32}P]tRNA^{Gly} for quantitation experiments. Such purification was achieved by subjecting tRNA^{Gly} eluted from gels to RPC-2 column chromatography as described previously.[7]

Enzymic digestions. Digestions of oligonucleotides with silkworm nuclease were carried out in 5 µl of the enzyme at 125 µg/ml in 0.5 M Na_2CO_3-$NaHCO_3$ pH 10.5, 0.1 M NaCl, 0.005 M MgAc. Incubation was at 37°C for 180 min. The digestion products were separated by two-dimensional electrophoresis on cellulose acetate and DEAE paper at pH 3.5.

RESULTS

Purification of T4 tRNA^{Gly}

[^{32}P]tRNA was isolated from T4-infected cells as described in Materials and Methods. The [^{32}P]tRNA was then fractionated by electrophoresis on a 10% polyacrylamide gel. Fig. 1 shows the results obtained when [^{32}P]tRNA is subjected to polyacrylamide gel electrophoresis. The autoradiograph of the gel reveals eight sharp bands. Seven of these bands (A, B, 1, 2, 3, 4, 6) are pure species of RNA. Band 5, which has the same mobility as bulk *E. coli* tRNA, is a mixture of five species of tRNA.[4] Band 6 is T4 tRNA^{Gly}. All of these RNA species hybridize specifically to T4 DNA.[5]

Sequence Analysis of T4 tRNA^{Gly}

Pancreatic ribonuclease digestion products. T4 [^{32}P]tRNA^{Gly} was digested to completion with pancreatic ribonuclease, and the products separated using the standard two-dimensional system developed by Sanger et al.[10]

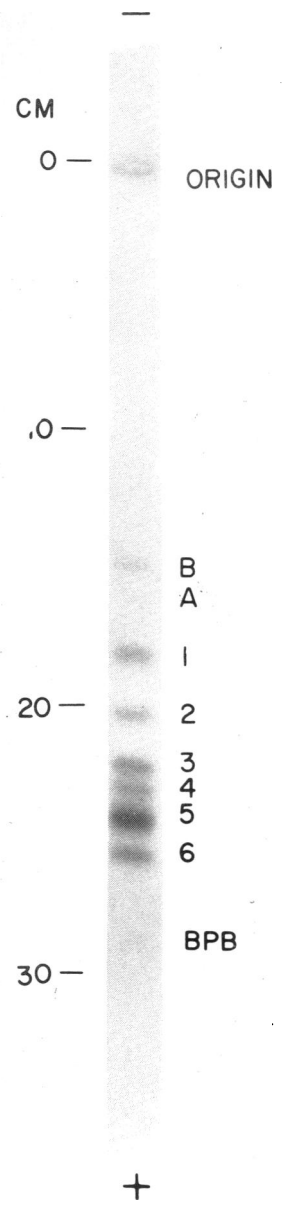

Fig. 1. Polyacrylamide gel electrophoresis of T4 [^{32}P]tRNA. Unfractionated T4 [^{32}P]tRNA purified only through the DEAE-cellulose step was electrophoresed on a 10% polyacrylamide gel as described in Materials and Methods. Distance migrated is shown in centimeters on the left. Bands of radioactivity are numbered A, B and 1-6. BPB indicates the position of migration of the bromphenol blue dye marker.

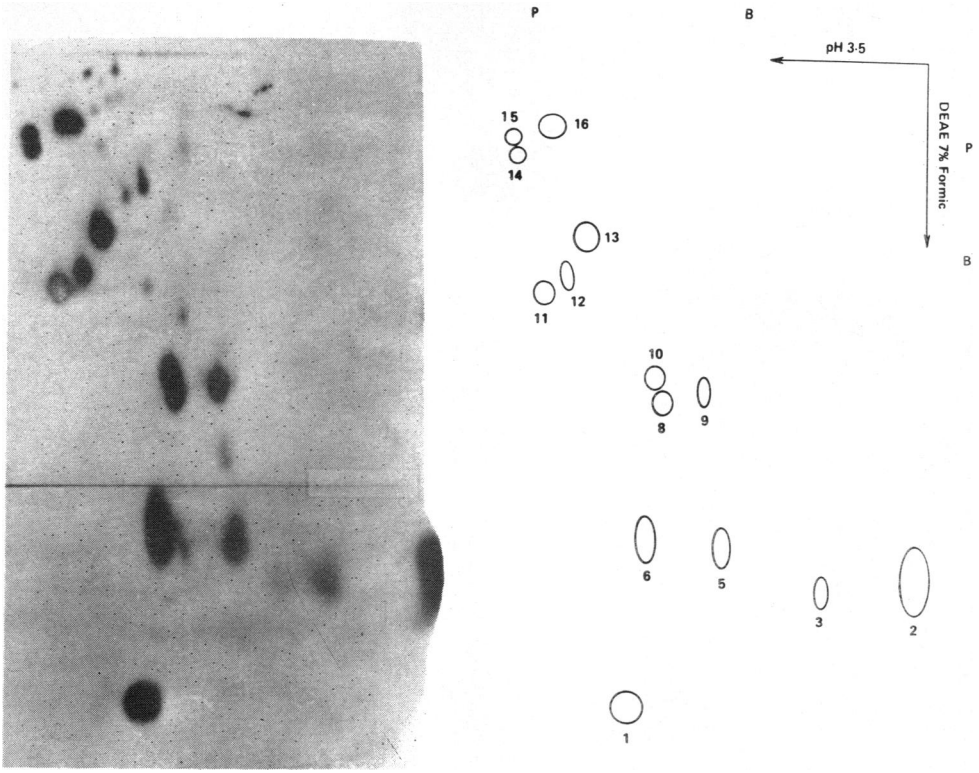

Fig. 2. Standard two-dimensional fingerprint of pancreatic ribonuclease digest of T4 [^{32}P]tRNA^{Gly}. Material used was from band 6 of Fig. 1. A, Autoradiograph of fingerprint on DEAE paper. B, Diagram identifying pancreatic nucleotide products. Products p1, "U", include ψ, as well as U. Letter P indicates the position of pink dye marker (acid fuchsin) and B shows the position of blue dye (xylene cyanol) in the first and second dimensions.

Fig. 2A shows the resulting fingerprint, while Fig. 2B identifies by number the oligonucleotides corresponding to each spot.

Analysis of the pancreatic digestion products and their sequence is shown in Table I. All but one nucleotide, p16, were sequenced by the combined data from alkaline and T1 ribonuclease digestions. Nucleotide p16 is in fact a mixture of two unresolved nucleotides, each containing 1 Ap, 1 Up or Tp, and 2 Gp residues. Since one of the T1 ribonuclease digestion products is the nucleotide ApUp, the sequence of one of the nucleotides comprising p16 must be GGAU. The second nucleotide of p16, p16B, contains a Gp, an ApGp and a Tp residue. Thus the sequence of this nucleotide could be either AGGT or GAGT. The sequence of this nucleotide was determined by examination of the products of partial T1 digestion of the entire tRNA molecule (see below). The sequence of T1 partial product 852-9 (see Table VII) is AGTψCG. Since there is only one Tp residue in the entire tRNA molecule, the sequence of p16B is GAGT.

Experimentally determined molar yields of the pancreatic nucleotides of Fig. 2 are shown in Table II. The figures are based on data from a pancreatic ribonuclease digest of T4 $tRNA^{Gly}$ that had been further purified by chromatography on an RPC-2 column as described in Materials and Methods. Nucleotides p14 and p15 contain various modified forms of the same sequence. Nucleotide p14 is completely modified with regard to 2,O-methyl guanosine and contains 2/3 molar yield of D and 1/3 of U. Nucleotide p15 is unmodified with respect to 2,O-methyl guanosine and contains 1/3 molar yield of D and 2/3 of U. The combined molar yield of p14 and p15 is 1.1.

T1 ribonuclease digestion products. A standard two-dimensional fingerprint of the T1 ribonuclease digestion products of T4 $tRNA^{Gly}$ is shown in Figs. 3A and 3B. A standard fingerprint of a combined T1-bacterial alkaline phosphatase digest is shown in Figs. 3C and 3D. Nucleotides t12 and t13 were not always separated on these fingerprints and required separation by homochromatography on DEAE-cellulose thin layer plates. Nucleotides t8 through t13 could also be separated by homochromatography of the T1 digest products in one dimension.

Table III shows the composition and sequence of the T1 ribonuclease digestion products, as far as they could be determined by initial analysis of their alkaline and pancreatic ribonuclease digestion products. The quantitative data from alkaline and pancreatic digestion above were sufficient to establish the sequences of nucleotides t1 through t6. Tables IV and V

TABLE I. Analysis of pancreatic ribonuclease digestion products[a]

Nucleotide	Alkaline digestion products[b]	T1 ribonuclease digestion products[c]	Deduced sequence or composition
p1	Up[d]	Up[d]	Up[d]
p2	Cp	Cp	Cp
p3	1.1 Ap + 0.9 Cp	A-Cp	A-Cp
p5	0.9 Gp + 1.1 Cp	1.0 Gp + 1.0 Cp	G-Cp
p6	1.0 Ap + 0.9 Up	A-Up	A-Up
p8	2.0 Ap + 1.0 Up	A-A-Up	A-A-Up
p9	2.1 Ap + 0.9 Gp + 1.0 Cp	1.0 A-Gp + 0.6 A-Cp	A-G-A-Cp
p10	2.1 Ap + 0.9 ψp	A-A-ψp	A-A-ψp
p11	1.0 Gp + 1.0 Up	1.0 Gp + 1.0 Up	G-Up
p12	pGp + Cp	1.0 pGp + 0.8 Cp	pG-Cp
p13	1.1 Ap + 0.9 Gp + 1.0 Up	1.0 Gp + 1.6 A-Up	G-A-Up
p14[e]	1.0 Gm-Gp + 1.0 Dp	1.0 Gm-Gp + 1.0 Dp	Gm-G-Dp
p15[e]	2.0 Gp + 1.2 Up	2.0 Gp + 0.5 Up	G-G-Up
p16A[f] } p16B	1.0 Ap + 1.8 Gp + 1.2 (Up + Tp)	1.0 A-Gp + 2.8 Gp + 1.1 A-Up + 0.8 Tp	G-G-A-Up (Gp,A-Gp)Tp

Pancreatic RNase digestion products from fingerprints as shown in Fig. 2 were analyzed by standard alkaline and enzymatic digestions. The heading "Nucleotide" of the first column refers to any pancreatic digestion product, either mononucleotide or oligonucleotide.

(a) Dashes between bases indicate phosphodiester bonds. Relative yields were determined by measuring the radioactivity of the paper containing the nucleotide products in a Nuclear Chicago scintillation counter. (b) Products were identified by their electrophoretic mobilities at pH 3.5 on Whatman 540 paper. Identifications of modified bases Tp, ψp, Dp and GmpGp were verified in various chromatographic systems (see Table VIII). (c) All products other than mononucleotides were analyzed by alkaline hydrolysis. Nucleotides p3, p6, p8 and p10 were predictably unaffected by T1 ribonuclease digestion and had the expected electrophoretic mobilities. (d) Nucleotide p1, more than one spot in most fingerprints, includes ψp. (e) Nucleotides p14 and p15 are two versions of one variably modified pancreatic product. See also Table II, and text. In p14, D is predominant, but there is also 1/3 unmodified U. In p15, U is predominant, but 1/3 is modified to D. These ratios were determined through analysis of the T1 digestion products electrophoresed on DEAE paper at pH 3.5. (f) Nucleotides p16A and p16B are two unresolved pancreatic ribonuclease digestion products. See text.

TABLE II. Pancreatic ribonuclease digestion products

Nucleotide	Sequence	Molar Yields	
		Experimentally determined[a]	Expected from final structure
p1	Up (Np, ψp)	9.5	10
p2	Cp	9.0	13
p3	A-Cp	0.9	1
p5	G-Cp	0.9	1
p6	A-Up	4.8	5
p8	A-A-Up	1.1	1
p9	A-G-A-Cp	0.9	1
p10	A-A-ψp	1.1	1
p11	G-Up	2.1	2
p12	pG-Cp	0.8	1
p13	G-A-Up	3.0	3
p14	Gm-G-Dp	0.7	1
p15[b]	G-G-Up	0.4	-
p16A	G-G-A-Up	} 2.0	1
p16B	G-A-G-Tp		1

Relative yields of products obtained in fingerprints such as in Fig. 2.

(a) Yields were determined by measuring the radioactivity of the paper containing the nucleotide in a Nuclear Chicago scintillation counter.
(b) Nucleotide p15 is the unmodified form of p14 as explained in the text and Table I.

show the results of additional experiments needed to establish the sequences of the remaining longer T1 nucleotides.

Sequence analysis of t7. Digestion with silkworm nuclease and bacterial alkaline phosphatase of 3'-dephosphorylated t7 obtained from fingerprints such as that shown in Fig. 3C were carried out as described in Materials and Methods. Silkworm nuclease cleaves 3'-phosphodiester bonds leaving 3'-OH and 5'-p products. When bacterial alkaline phosphatase is included in the reaction, the 5' phosphates are also removed from the digestion products. This digestion of t7 yielded the product CpG as identified by subsequent alkaline and snake venom phosphodiesterase (SVP) digestion. This establishes the sequence of t7 as (T,ψ)CG. Analysis of the pancreatic ribonuclease digestion products from T4 $tRNA^{Gly}$ indicates that there is only one Tp residue in the tRNA molecule, and it is present in nucleotide p16B, GAGT. Thus t7 must have the Tp residue at its 5' end, and the sequence of t7 is established as TψCG.

Sequence analysis of t8. Silkworm-phosphatase digestions of t8 were carried out and two of the products were ApUpCpG and UpCpG. The silkworm phosphatase product, UpCpG, could be sequenced by alkaline and SVP digestions. This permitted the sequencing of the next longer product, ApUpCpG. Thus the complete sequence of t8 is determined to be AUAUCG.

This sequence was verified by analysis of digestion products generated by U2 ribonuclease, which cleaves after purines. The U2 digestion products were A, UA, and (UC)G, as determined by alkaline hydrolysis.

Sequence analysis of t9. Alkaline and SVP digestion of 3' dephosphorylated t9 were carried out in order to determine the residue at the 5' end of the nucleotide. Alkaline digestion yielded 3.0 Ap + 3.1 Up, while SVP digestion yielded 2.8 pA + 1.0 pG + 2.2 pU. This indicates that the sequence of t9 is U(AU,AAU)G.

The products ApUpG and ApApUpG were obtained from silkworm-phosphatase digestions of t9 and establish its sequence as UAUAAUG.

Sequence analysis of t10. A partial digestion with SVP of 3'-dephosphorylated t10 obtained from one-dimensional DEAE thin-layer homochromatography was carried out using an enzyme concentration of 0.05 mg/ml and an incubation time of 30 min at 37°C. These relatively harsh reaction conditions were necessary since nucleotides separated by homochromatography have large amounts of unlabeled RNA mixed with them. Separation and analysis of the partial digestion products were carried out as described by Barrell.[11]

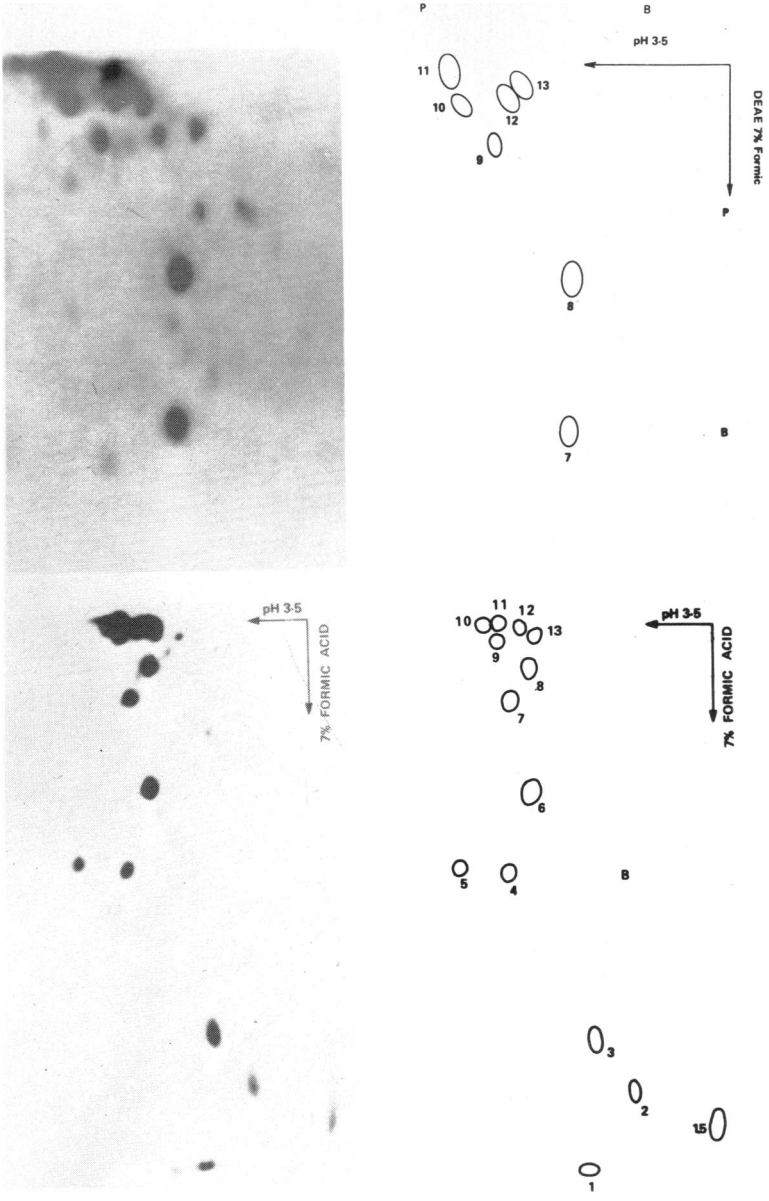

Fig. 3. Two-dimensional fractionations of T1 ribonuclease digests of T4 [^{32}P]tRNA^{Gly}. A, Standard fingerprint on DEAE-paper, of T4 material from Band 6 of Fig. 1. B, Diagram identifying nucleotides in A. Letter B indicates the position of the blue dye marker (xylene cyanol). C, Standard fingerprint of combined T1-bacterial alkaline phosphatase digest of T4 material from Band 6. The digestion products are 3'-OH rather than 3'-p as they are in A and B. D, Diagram identifying nucleotides in C. Letters P and B indicate the position of the dye markers in the first and second dimensions.

TABLE III. Initial analysis of T1 ribonuclease digestion products[a]

Nucleotide	Alkaline digestion products[b]	Pancreatic ribonuclease digestion products[c]	Deduced sequence or composition
t1	G	G	G
t2	1.0 C + 1.1 G	0.7 C + 1.0 G	CG
t3	1.2 A + 1.0 G	AG	AG
t4	1.1 U + 1.0 G	1.0 U + 1.0 G	UG
t5	pG	pG	pG
t6	1.1 A + 1.0 U + 0.9 G	1.4 AU + 1.0 G	AUG
t7	1.1 C + 2.0 (T + ψ) + 1.0 G	1.0 C + 1.8 (T + ψ) + 1.0 G	(T,ψ,C)G
t8	2.2 A + 1.0 C + 1.9 U + 0.9 G	2.4 AU + 0.7 C + 1.0 G	(2AU,C)G
t9[d]	3.2 A + 3.0 U + 0.9 G	0.9 AAU + 1.3 AU + 1.0 U + 0.6 G	(AAU,AU,U)G
t10	3.0 A + 3.2 U + 0.9 GmG	1.0 AAU + 1.1 AU + 1.0 U + 1.0 GmG	(AAU,AU,U)GmG
t11	3.3 A + 4.0 C + 5.9 U + 0.9 G	3.5 AU + 3.8 C + 2.7 U + 1.0 G	(3AU,4C,3U)G
t12	3.0 A + 3.0 C + 4.1 "U"[e] + 0.9 G	1.2 AU + 0.9 AC + 2.1 C + 2.8 "U"[e] + 1.0 AG	(AU,AC,2C,3"U"[e])AG
t13	3.0 A + 3.9 C + 3.8 "U"[f] + 1.3 G	1.1 AAψ + 1.0 AC + 3.1 C + 2.8 U + 1.0 G	(AAψ,AC,3C,3U)G
t15	3.0 C + 0.8 U	3.0 C + 0.8 U	(3C,U)

T1 RNase digestion products from fingerprints such as shown in Fig. 3A were analyzed by standard alkaline and enzymatic digestions. The term "Nucleotide" of column 1 refers to any T1 digestion product, either mononucleotide or oligonucleotide. (a) From considerations of space, all representations of phosphates are omitted except for the 5'-p of nucleotide t5. All bases shown are 3'-p, except as noted below. Relative yields were determined by measuring the radioactivity of the paper containing the nucleotide products in a Nuclear Chicago scintillation counter. Final analysis of the longer nucleotides required further experiments, shown in Tables IV and V. (b) All mononucleotides shown as alkaline digestion products are 3'-p. In t7 the notation (T + ψ) indicates one spot containing both Tp and ψp. In t10, GmG is the dinucleotide GmpGp. (c) All bases shown are 3'-p. All products other than mononucleotides were analyzed by alkaline digestion. (d) As explained in the text, t9 results as a T1 ribonuclease digestion product of unmodified t10. (e) The notation "U" refers to unresolved Up and Dp. (f) The notation "U" refers to unresolved Up and Dp. As explained in the text, there is evidence that another of the U residues in t13 is also modified.

TABLE IV. Analysis of snake venom phosphodiesterase partial digestion products from T1 nucleotides

T1 Nucleotide		Alkaline digestion products[a]	Snake venom digestion products[b]	Deduced sequence of partial snake venom digestion product[c]	M[d]
t10	(1)[e]	1.5 Ap + 1.0 Up	0.6 pA + 1.0 pU	UpApU	
	(2)	1.0 Ap + 1.7 Up	1.7 pA + 1.0 pU	UpApUpA	1.7
	(3)	2.0 Ap + 2.0 Up	[2.7 pA + 2.0 pU][f]	UpApUpApA	2.5
	(4)	3.0 Ap + 2.0 Up	[2.1 pA + 2.0 pU][f]	UpApUpApApU	1.5
t11	(1)	1.0 Ap + 1.0 Up	n(pU)	ApUpU	
	(2)	1.0 Ap + 1.8 Up	1.0 pC + 2.2 pU	ApUpUpC	1.0
	(3)	1.1 Cp + 1.0 Ap + 1.9 Up	1.0 pC + 3.7 pU	ApUpUpCpU	1.9
	(4)	1.1 Cp + 1.0 Ap + 2.9 Up	2.0 pC + 3.5 pU	ApUpUpCpUpC	1.0
	(5)	1.8 Cp + 1.0 Ap + 2.4 Up	[2.0 pC + 0.6 pA + 4.9 pU][f]	ApUpUpCpUpCpA	1.9
	(6)	1.9 Cp + 2.1 Ap + 3.0 Up	[1.6 pC + 1.0 pA + 5.1 pU][f]	ApUpUpCpUpCpApU	1.1
t12	(1)	1.1 Ap + 1.0 Dp	1.0 pA + 1.2 pU	DpApU	
	(2)	1.0 Ap + 0.9 Up	1.0 pA + 1.2 pU	UpApU	
	(3)	1.0 Ap + 1.8 "Up"[g]	1.0 pA + 2.0 pU	DpApUpU	1.9
	(4)	1.0 Ap + 1.7 Up	1.0 pA + 2.0 pU	UpApUpU	2.1
	(5)	1.0 Ap + 2.5 "Up"[g]	2.0 pA + 2.4 pU	$^{D}_{U}$pApUpUpA	2.6
	(6)	[3.1 Ap + 3.0 "Up"][f,g]	1.0 pC + 1.9 pA + 2.2 pU	$^{D}_{U}$pApUpUpApC	0.9
t13	No partial products obtained				
t15	(1)	1.0 Cp + 1.0 Up	0.7 pC + 1.0 pU	CpUpC	
	(2)	2.7 Cp + 1.0 Up	2.0 pC + 1.0 pU	CpUpCpC	0.7
	(3)	3.0 Cp + 1.0 Up	2.0 pC + 1.0 pA + 1.2 pU	CpUpCpCpA	2.1

3'-dephosphorylated nucleotides t10, t11, t12 and t13 were partially digested with snake venom phosphodiesterase and products separated as in Materials and Methods. Nucleotide t15 is the naturally occurring 3'-OH product from the 3' end of the tRNA molecule.

(a) Products identified and yields determined as in Table I. Alkaline digestion yields 3'-mononucleotides for all except the 3'-terminal nucleoside of the partial product. (b) Products identified by electrophoretic mobilities at pH 3.5 on Whatman 540 paper and yields determined as in Table I. Snake venom phosphodiesterase digestion yields 5'-mononucleotides for all except the 5'-terminal nucleoside of the partial product. (c) Partial products are 3'-OH. (d) M = x/y where y = distance from origin to the digestion product and x = the distance from the digestion product to the next smaller degradation product as discussed in Brownlee and Sanger.[15] M does not fall in the ranges for terminal nucleotides as determined by Brownlee and Sanger (C 0.6 to 1.2, A 2.1 to 2.9, U 1.7 to 1.9, and G 2.6 to 4.4)[15] for all of the products listed. However, it is obvious from the data presented that the sequence determinations are correct. Others have observed that the M values for the terminal nucleotides do not always fall within the expected ranges.[7,16] (e) Partial products are listed in order of decreasing mobility for each T1 nucleotide. Numbers in parentheses are arbitrary. (f) Data shown in brackets do not support the deduced sequence shown. (g) The notation "Up" refers to unresolved Up and Dp.

TABLE V. Silkworm nuclease digestion products of T1 nucleotides

Nucleotide		SW[a]	SWP[b]
t11		ApUpU	ApUpU
		ApUpUpC	
		pUpCpU	
		pUpCpUpC	
		pCpUpC	
		pCpA	
		pApUpU	ApUpU
		pApUpUpA	
		pUpUpApU	
		pApUpCpCpG	
		pUpCpCpG	UpCpCpG
		pCpCpG	
		pCpG	CpG
	Deduced sequence:	A-U-U-C-U-C-A-U-U-A-U-C-C-G	
t12		D/U pApUpU	
		pUpA	
		pUpApCpC	
		pApCpC	ApCpC
		pApCpCpU	ApCpCpU
		pCpCpU	
		pCpUpC	
			CpApG
		pApG	
	Deduced sequence:	D/U-A-U-U-A-C-C-U-C-A-G	
t13		ApCpU	ApCpU
		ApCpUpU	
		pUpUpC	UpUpC
		pUpUpCpC	
		pUpCpC	UpCpC
		pUpCpCpA	
		pCpCpA	CpCpA
		pCpApA	
		pApApψ	
		pApApψpC	
		pApψpC	
		pψpCpUpG	
		pCpUpG	CpUpG
	Deduced sequence:	A-C-U-U-C-C-A-A-ψ-C-U-G	

3'-dephosphorylated T1 nucleotides separated by homochromatography were digested with silkworm nuclease, and combined silkworm nuclease-bacterial alkaline phosphatase, and products were separated as in Materials and Methods. Products were sequenced by analysis of alkaline and snake venom phosphodiesterase digestion. (a) "SW" indicates products of digestion with silkworm nuclease alone. Products are listed in order from 5'-end to 3'-end of T1 nucleotide, rather than by mobility. (b) "SWP" indicates products of combined silkworm-phosphatase digestion.

The products are listed in Table IV. Combined with the overall composition seen in Table III, this series of products establishes the sequence of t10, since only the terminal GmG remains to be added to the sequence of partial product number (4). The complete sequence of t10 must be UAUAAUGmG.

Sequence analysis of t11. A partial SVP digestion of t11 purified by homochromatography was carried out using an enzyme concentration of 0.1 mg/ml and an incubation time of 45 min at 37°C. The results (Table IV) establish the sequence of the first eight bases in t11.

The complete sequence of t11 was established by analysis of the products obtained from silkworm nuclease and combined silkworm-phosphatase digestions.

Two basic classes of products result from silkworm nuclease digestion of a T1-phosphatase nucleotide. The enzyme cleaves 3'-phosphodiester bonds leaving 3'-OH and 5'-p products. Possible kinds of products are: Type A, which include the 5'-end of the nucleotide, and Type B, which do not. Type A products from t11, shown in Table V-A, are ApUpU and ApUpUpC. Type B products can be either internal products like pUpCpU, or 3'-end products like pUpCpCpG.

Type A (5'-end) products are both 5'-OH and 3'-OH, and can be analyzed by alkaline and snake venom digestion in the same way as the partial snake venom products of Table IV.

Both kinds of Type B products (internal and 3'-end) are 5'-p and 3'-OH. The 5'-base of a Type B product appears as a nucleoside diphosphate pXp after alkaline digestion. The pXp's have characteristic mobilities and identification of a given pXp is verified by the appearance of the 5'-mononucleoside pX after snake venom digestion. An example is product pUpCpU in Table V-A. Alkaline digestion yields pUp plus Cp. Snake venom digestion yields 2 pU and 1 pC which verifies the identification of pUp.

Silkworm digestion products from t11 in column "SW" of Table V-A are shown in order, from the 5'-end to the 3'-end, rather than by mobility. Silkworm-phosphatase products are in column "SWP". These results establish the sequence of t11 as AUUCUCAUUAUCCG. Nucleotide t11 was also digested with U2 ribonuclease. The digestion products found, (UUCUC)A, UUA, and (UCC)G, as determined by alkaline hydrolysis, confirm the given sequence.

Sequence analysis of t12. Partial SVP digestions of t12 isolated from a T1-phosphatase fingerprint such as shown in Fig. 3C using enzyme concentrations of 0.01 mg/ml and incubation times of 15 and 30 min at 37°C, and 0.05 mg/ml enzyme for 15 min at 37°C. All three reaction mixtures were

combined and the products separated as described in Materials and Methods. These results (Table IV) establish the sequence of the first six bases in t12. These results also show that the U residue at the 5'-end of t12 was incompletely modified to D. In our preparations of $tRNA^{Gly}$ approximately 50% of the U in this position is modified to D.

Table V-B shows the results of silkworm and silkworm-phosphatase digestions of t12. These results establish the complete sequence of t12, which is DAUUACCUCAG. The U2 digestion products DA, UUA, and (CCUC)A verify this sequence.

Sequence analysis of t13. Silkworm and silkworm-phosphatase digestions of t13, shown in Table V-C, generated a series of products which unambiguously give the sequence of this nucleotide as ACUUCCAAψCUG. This sequence was confirmed by the results of a U2 digestion of t13 and subsequent alkaline hydrolysis of its U2 products. As predicted by the sequence above, the U2 products (CUUCC)A and ψCU(G) were found.

As seen in Table IV, no partial SVP digestion products were obtained from t13 in sufficient quantity to permit analysis, although the nucleotide was susceptible to complete SVP digestion. This resistance to SVP has also been found in the case of nucleotide u15C of T4 $tRNA^{Leu}$,[7] and in both cases these nucleotides contain the anticodon of their respective tRNA molecules. It is possible that these nucleotides have a special tertiary structure, even when separated from the rest of the tRNA molecule, that renders them resistant to digestion with SVP. It is also possible that the ψ modification near the 3'-end of each nucleotide interferes with the partial venom digestions.

There is some evidence that there is a modified U in t13 in addition to ψ. If the Up residues from a pancreatic ribonuclease digest of t13 are eluted and chromatographed in the systems described in Table VIII, a nucleotide with different mobilities than U also appears. Alkaline digestions of t13 do not yield this nucleotide, so this modification is apparently alkaline labile. We shall refer to this nucleotide as U^{+}. We have no evidence for the actual position of U^{+} in t13, although it is likely that it is in the wobble position of the anticodon.

Sequence analysis of t15. A partial SVP digestion was carried out on t15 isolated from a T1 fingerprint as shown in Fig. 3A using an enzyme concentration of 0.01 mg/ml and an incubation time of 15 min at 37°C. The results are listed in Table IV. Nucleotide t15 is the naturally occurring

3'-OH product from the 3'-end of the tRNA molecule, the complete sequence of which is CUCCA-OH. This sequence is verified by silkworm-phosphatase products CpUpC and CpCpA.

Experimentally determined molar yields of the T1 ribonuclease digestion products are shown in Table VI. The figures are based on data from a T1-ribonuclease digest of T4 $tRNA^{Gly}$ that had been further purified by chromatography on an RPC-2 column as described in Materials and Methods. Nucleotide t9 is the unmodified form of t10. When the G in the seventh position of t10 is not modified, T1 ribonuclease cleaves in this position and yields t9. The combined molar ratio of t9 and t10 is 0.9. The experimental molar yields agree very well with those expected on the basis of the finally determined structure.

Partial Enzymatic Digestion Products and Overall Sequence of T4 $tRNA^{Gly}$

Partial digestion of T4 [^{32}P]$tRNA^{Gly}$, using T1 and modified ε-carboxymethyllysine-41-pancreatic RNases, were carried out and the products fractionated as in Materials and Methods. Table VII shows a series of such partial digestion products. These partial products are more than sufficient to establish the overall sequence of the molecule.

Each partial digestion product was analyzed by subsequent total digestion with T1 and pancreatic RNases. Mobilities of the total digestion products in 7% formic acid were compared with known mobilities of total digestion products of the entire molecule which were run as markers on the same paper. In addition the total digestion products of one enzyme were analyzed by further digestion with the other enzyme, or with alkali, or with both.

The combined data usually, but not always, led to a definitive sequence for the partial product. Where a partial product could not be independently sequenced, however, the results from other partials could be applied to order it. For instance, partial product T1-852-14 is AUGUGAGTψCG. It consists of t6, t4, t3, and t7. Total T1 digests yield these four products, which are clearly identifiable. Inspection will show, however, that the pancreatic digestion products of order t6-t4-t3-t7 are identical in composition with those of t6-t3-t7-t4. However, several partial products in Table VII show definitively that t3 is immediately preceded by t4. Furthermore, product T1-852-19 shows that the 3'-end of t7 is linked to t11. Therefore the order of T1-852-14 must be as shown.

TABLE VI. T1 ribonuclease digestion products

Nucleotide	Sequence	Molar Yields	
		Experimentally determined[a]	Expected from final structure
t1	Gp	1.2	1
t2	C-Gp	0.9	1
t3	A-Gp	1.0	1
t4	U-Gp	1.0	1
t5	pGp	0.8	1
t6	A-U-Gp	2.0	2
t7	T-ψ-C-Gp	1.2	1
t8	A-U-A-U-C-Gp	1.0	1
t9[b]	U-A-U-A-A-U-Gp	0.3	-
t10	U-A-U-A-A-U-Gm-Gp	0.6	1
t11	A-U-U-C-U-C-A-U-U-A-U-C-C-Gp	1.0	1
t12	D/U-A-U-U-A-C-C-U-C-A-Gp	1.1	1
t13	A-C-U-U-C-C-A-A-ψ-C-U-Gp	0.8	1
t15	C-U-C-C-A-OH	0.8	1

Relative yields of products obtained in fingerprints such as in Fig. 3A.

(a) Yields were determined by measuring the radioactivity of the paper containing the nucleotide in a Nuclear Chicago scintillation counter.
(b) As explained in the text, t9 results as a T1 ribonuclease digestion product of unmodified t10.

An example of a partial product that can be individually sequenced is T1-852-21. Total digestion of this partial product with T1 RNase yields t6, t12 and t13. Complete pancreatic RNase digestion of this partial yields products including AGAG and GAU, and G. This unambiguously establishes the sequence of T1-852-21 as t12-t13-t6.

The final sequence of T4 $tRNA^{Gly}$ is shown both in Table VII and in the familiar cloverleaf configuration in Fig. 4.

Identification and Characterization of Modified Nucleotides

Modified nucleotides were identified and further characterized as described in Pinkerton et al.[7] The mobilities of the modified and unmodified nucleotides in the various systems employed are summarized in Table VIII.

All modified nucleotides in the molecule with the exception of U^{+} are identified and their mobilities in the various systems agree with published values.[7,11]

Sequence Comparison with $tRNA^{Gly}$ from Bacteriophages T2 and T6

Bacteriophages T2 and T6 also code for a tRNA with an electrophoretic mobility on 10% polyacrylamide gels which is the same as that for the T4 $tRNA^{Gly}$ (band 6). Pancreatic and T1 digestion products from these tRNAs give fingerprints that are identical to the pancreatic and T1 fingerprints obtained for T4 $tRNA^{Gly}$. Sequence analysis of the pancreatic and T1 ribonuclease digestion products of these tRNAs from T2, T4, and T6, in which each product was digested with the other enzyme and the electrophoretic mobilities of these digestion products compared on DEAE paper at pH 3.5, indicates that the T2 and T6 $tRNAs^{Gly}$ sequences are identical to that of T4 $tRNA^{Gly}$.

DISCUSSION

From the data presented above, a nucleotide sequence has been unambiguously determined for T4 $tRNA^{Gly}$. The evidence presented indicates incomplete nucleotide modification at some sites in the tRNA. Such incomplete modification of nucleotides has been seen before in tRNA isolated from phage-infected cells.[7] The data concerning pancreatic nucleotides p14, GmGD, and p15, GGU, shows partial modification of U to D in both cases. This indicates that the modificat on of U to D and the 2'-O methylation of G occur independently of each other.

The anticodon of T4 $tRNA^{Gly}$ is UCC which would be expected to recognize glycine codons GGA and GGG. Our evidence suggests the possibility

TABLE VII. Partial enzymatic digestion products of T4 [^{32}P]tRNA^{Gly}

	pGCGGAUAUCGUAUAAUGmGDAUUACCUCAGACUUCCAAψCUGAUGAUGUGAGTψCGAUUCUCAUUAUCCGCUCCA-OH
T1-852-4	pGCG
T1-852-8	pGCGG
MP-851-23	pGCGGAU
MP-851-32	pGCGGAUAU
MP-909-19	pGCGGAUAUC
T1-852-16	pGCGGAUAUCG
T1-852-17	AUAUCGUAUAAUGmG
MP-909-6	AUAAUGmGD
MP-909-3	AAUGmGD
MP-909-27	AAUGmGDAUUACCUC
T1-852-21	DAUUACCUCAGACUUCCAAψCUGAUG
MP-851-39	ACCUCAGACUUCC
MP-851-16	AGACU
MP-851-25	AGACUUCC
T1-852-18	ACUUCCAAψCUGAUG
MP-909-26	AAψCUGAUGAUGU
MP-909-35	AAψCUGAUGAUGUGAGT
MP-851-15	GAUGU
T1-852-14	AUGUGAGTψCG
MP-851-37	GUGAGTψCGAU
MP-851-40	GUGAGTψCGAUU
MP-851-15	GAGTψ
MP-851-19	GAGTψC
MP-851-28	GAGTψCGAU
MP-851-32	GAGTψCGAUU
T1-852-9	AGTψCG
T1-852-19	TψCGAUUCUCAUUAUCCG
T1-852-22	TψCGAUUCUCAUUAUCCGCUCCA-OH
T1-852-20	AUUCUCAUUAUCCGCUCCA-OH
MP-851-26	AUCCGCUCC
MP-851-27	AUCCGCUCCA-OH

Partial digestions of T4 [^{32}P]tRNA^{Gly} were carried out using T1 and modified ε-carboxymethyllysine-41-pancreatic RNases, and the products fractionated as in Materials and Methods. In product nomenclature MP or T1 indicate the enzyme used. The second group of three numbers refers to a given experiment and the last number refers to a specific partial product for a given experiment. For T1 RNase, an enzyme-to-substrate ratio of 1:1000 was used for experiment 852. For modified pancreatic RNase, the ratios were 1:20 for experiment 851, and 1:100 for 909.

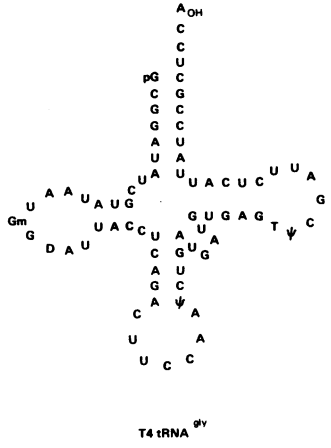

Fig. 4. Cloverleaf arrangement of T4 tRNA^{Gly}.

TABLE VIII. Nucleotide characterization

Nucleotide[f]	Electrophorestic systems[a]		Chromatographic systems[b]		
	DEAE pH 3.5 Ru	Whatman 540 pH 3.5 Ru	System A[c] Ru	System B[d] Ru	System C[e] Ru
Cp	0.81	0.18	1.30	0.65	-
Ap	-	0.38	1.44	0.53	-
Gp	0.66	0.77	0.90	0.44	-
GmGp	0.20	0.94	-	0.16	-
Dp	1.05	1.00	0.90	1.07	0.86
Tp	0.98	0.98	1.17	1.11	-
ψp	0.98	0.98	0.78	0.75	-
$G^{m}_{p}G$	0.59	0.46	-	-	-
pG^{m}	-	0.70	-	-	-
pG	-	0.68	-	-	-
$U^{+}p$	1.00	-	0.91	1.21	-

Mobilities are shown for nucleotides contained in $tRNA^{Gly}$ relative to the mobility for Up.

(a) Mobilities were obtained for products derived from alkaline, pancreatic, or T1 digestions of the parent T1 or pancreatic oligonucleotides, except that $G^{m}_{p}G$ was derived from alkaline digestion of dephosphorylated t10, and pG^{m} and pG were derived from total snake venom digestions of dephosphorylated t10. (b) Mobilities were obtained for nucleotides which were first isolated by electrophoresis on Whatman 540 paper after alkaline digestion of the parent T1 or pancreatic oligonucleotide. An exception to this rule is that the mobility of $U^{+}p$ was determined by first isolating the Up residues after a pancreatic ribonuclease digestion of t13 by electrophoresis on DEAE paper at pH 3.5. See text. (c) System A is ascending thin-layer cellulose chromatography with isobutyric acid and 0.5 M NH_4OH (5:3, v/v) as the solvent. R_f for Up = 0.53. (d) System B is ascending thin-layer cellulose chromatography with isopropanol, conc. HCl, and water (70:15:15, v/v/v) as the solvent. R_f for Up = 0.85. (e) System C is ascending thin-layer cellulose chromatography with isopropanol, water, and conc. NH_4OH (70:30:1, v/v/v) as the solvent. R_f for Up = 0.15. (f) All nucleotides are the nucleoside 3' phosphates except GmGp is the dinucleoside diphosphate, $Gp^{m}G$ is the dinucleoside monophosphate, and pG^{m} and pG are the nucleoside 5' phosphates. Identical nucleotides from different sources (i.e., the parent T1 or pancreatic oligonucleotide) had the same mobilities when run in the various characterization systems.

that the U in the anticodon is modified but the nature and extent of the modification remains undetermined.

The results of our experiments indicate that the $tRNA^{Gly}$ coded for by the T2 and T6 bacteriophages probably have a sequence which is the same as that of T4 $tRNA^{Gly}$. Since partial digestions were not performed on the T2 and T6 $tRNA^{Gly}$ to determine the order of the T1 and pancreatic digestion products, the sequence identities are, of course, not rigidly proved. However, the small size of these tRNAs which allows for a minimum of sequence permutations that would still obey the known rules for tRNA sequences and the several unique overlaps provided by matching the T1 and pancreatic oligonucleotide sequences suggests that the T2, T4, and T6 $tRNA^{Gly}$ nucleotide sequences are indeed identical.

Sequence comparisons among the $tRNAs^{Gly}$ coded for by T4, E. coli, and Staphlococcus epidermis have been previously reported.[9,12] These comparisons implicate the 3'-5' stem and the anticodon as potential sites for amino acyl synthetase recognition of $tRNA^{Gly}$. It has already been shown that the anticodon is important in the synthetase recognition of $tRNA^{Gly}$.[13,14] Unlike the E. coli and T4 leucyl-tRNAs,[7] the length of the $tRNA^{Gly}$ molecules appears to be inessential to synthetase recognition since T4 $tRNA^{Gly}$ has 74 nucleotides and the three E. coli tRNAs have 76 nucleotides.[13]

Acknowledgments. We thank Peter Johnson for his help, and Drs. John Carbon and Richard Roberts for communicating their unpublished results. We are indebted to June Miller for preparing the manuscript. This work was supported by a grant from the National Cancer Institute (CA-10984). J.A. is a Faculty Research Associate of the American Cancer Society (FRA-80). G.V.P. was supported by a USPHS Training Grant (GM-01045).

* Department of Biochemistry, University of California, Berkeley, California 94720, USA

/ Department of Biology, University of California, Los Angeles, California 90024, USA

REFERENCES

1 Hsu, W. T., Foft, J. W. and Weiss, S. B. (1967) Proc. Nat. Acad. Sci. USA 58, 2028-2035
2 Daniel, V., Sarid, S. and Littauer, U. Z. (1970) Science 167, 1682-1688
3 Scherberg, N. H. and Weiss, S. B. (1970) Proc. Nat. Acad. Sci. USA 67, 1164-1171
4 McClain, W. H., Guthrie, C. and Barrell, B. G. (1972) Proc. Nat. Acad. Sci. USA 69, 3703-3707
5 Wilson, J. H., Kim, J. S. and Abelson, J. N. (1972) J. Mol. Biol. 71, 547-556
6 Pinkerton, T. C., Paddock, G. and Abelson, J. (1972) Nature New Biol. 240, 88-90
7 Pinkerton, T. C., Paddock, G. and Abelson, J. (1973) J. Biol. Chem. 248, 6348-6365
8 Adesnik, M. and Levinthal, C. (1970) J. Mol. Biol. 48, 187-208
9 Stahl, S., Paddock, G. and Abelson, J. (1973) Biochem. Biophys. Res. Commun. 54, 567-569
10 Sanger, F., Brownlee, G. G. and Barrell, B. G. (1965) J. Mol. Biol. 13, 373-398
11 Barrell, B. G. (1971) Proc. Nucl. Acid Res. 2, 751-779
12 Hill, C. W., Combriato, G., Sternhart, W., Riddle, D. L. and Carbon, J. (1973) J. Biol. Chem. 248, 4252-4262
13 Squires, C. and Carbon, J. (1971) Nature New Biol. 233, 274-277
14 Carbon, J. and Curry, J. B. (1968) J. Mol. Biol. 38, 201-216
15 Brownlee, G. G. and Sanger, F. (1967) J. Mol. Biol. 23, 337-353
16 Goodman, H. M., Abelson, J. N., Landy, A., Zadrazil, S. and Smith, J. D. (1970) Eur. J. Biochem. 13, 461-483