SEQUENCE ANALYSIS OF TRANSFER RNA_7^{SER} OF

DROSOPHILA MELANOGASTER

by

DAVID LAMAR CRIBBS
B.A., The University of Colorado, 1975

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in
THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF BIOCHEMISTRY
FACULTY OF MEDICINE

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 1979

SUMMARY (Abstract)

Sequence analysis of $tRNA_7^{Ser}$ from *Drosophila melanogaster* was carried out, primarily by the formamide degradation and post-labeling method of Stanley and Vassilenko. Preliminary analysis was on the terminal nucleoside-5', 3'-bis [5' - ^{32}P] phosphates of electrophoretically separated [5' - ^{32}P] ribooligonucleotides, identifying the [^{32}P]-nucleotides by chromatography on PEI-cellulose plates. Further analysis of possible modified nucleotides was performed by thin layer chromatography of [5' - ^{32}P] nucleoside phosphates derived by nuclease P_1 digestion from [5' - ^{32}P] oligomers. The partial sequence generated in this fashion was supplemented by ladder gel sequence analysis of [5' - ^{32}P]$tRNA^{Ser}$, and to a limited extent by two dimensional homochromatography. In this way, a sequence was obtained that is complete except for part of the aminoacyl stem and the 5'-end of the extra arm. The data are consistent with a sequence for $tRNA_7^{Ser}$ of pGCAG$_m$UUGUGGCac^{4}CGAGCG$_m$GDDAAGGCXUCUGA—
m^{3}CUIGAi^{6}AAψCAGAU$_m$UCCCUm^{3}CUGGGAGm^{5}CGUAGGTψCGm^{1}AAUCCUACCGACUGCNCCA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I am indebted to the people in our laboratory over the past three years for help and discussions, and particularly to Ian Gillam and Gordon Tener for worthwhile advice and encouragement. Thanks also go to workers not in our lab for their practical advice, including Caroline Astell and David Leung; and especially to Otto Hagenbüchle, whose help was extensive and timely.

Special thanks to Jenny, who gave me much help and support.

Chapter 1

INTRODUCTION

PART A

1.A.1 RNA

Ribonucleic acids (RNA) perform a variety of functions in the cell. Stable RNAs are primarily structural components involved in protein synthesis, and include the ribosomal RNAs (rRNA) and transfer RNAs (tRNA). Messenger RNA (mRNA), in eukaryotes apparently derived from high molecular weight nuclear RNA (hnRNA), is more rapidly metabolized than stable RNA, and is informational, directing the incorporation of amino acids into protein.

RNA is derived by transcription from the genomic DNA of a cell, and its sequence reflects that of the template DNA. The functional RNA may differ greatly from the initial transcript. Mechanisms by which specific nucleases act to generate mature RNA species by removal of terminal nucleotide sequences have been studied (for review, 1). It has recently been found that in many eukaryotic genes, the RNA transcript is not colinear with the DNA, due to the presence of intervening sequences in the DNA (2-8). In expression of some yeast tRNA genes, this requires the presence of enzyme activities capable of excising the intervening sequence from a tRNA precursor and ligating the split tRNA (9,10). Yeast precursor tRNA

is processed to tRNA size by enzyme activities present in oocytes of *Xenopus laevis* (11), suggesting that intervening sequences in tRNA genes, and enzymes like those in yeast needed to excise those sequences, may be general in eukaryotes.

In a comprehensive study of a particular gene's expression, it is desirable to know the sequence of the DNA and of the RNA product, and a precise knowledge of both is necessary to identify intervening sequences. Comparison of a large number of sequences of related RNAs may be of value in studying structure-function relationships (12,13). Further, comparative studies on gene structure are now possible in many cases, due to the genes' accessibility through recombinant DNA techniques. Genes of interest may be isolated as inserts into plasmids or bacteriophage by methods like those of Grunstein and Hogness (14) or of Maniatis *et al.* (15), provided a selective system is available to retrieve DNA of interest. Purified RNAs define specific regions of interest in the genome, those expressed during cellular function; and being complementary to their template DNA, the RNAs are useful in selecting for recombinant clones containing the genes to be studied.

1.A.2 tRNA

The tRNA of a cell is a mixture of at least 20 different species, with one or more acceptors for each amino acid required for synthesis of proteins. Transfer RNAs are small, normally 75-90 nucleotides in length, and contain many modified nucleotides (16). A substantial body of literature on tRNA sequences has been compiled, and certain features of

tRNA structure have become clear (17). The nucleotide sequences of tRNAs may be drawn in a standard "cloverleaf" configuration (Fig. 1, 2). In this structure there are double-stranded stems of standard length, a 3'-terminal CCA sequence, and nearly invariant nucleotides at certain positions (17,18).

Transfer RNAs are remarkable in the variety of cell components interacting with them. The tRNA precursors transcribed from DNA are processed by a number of enzymes, including nucleotidyl transferase which repairs the CCA-end. The modified nucleotides present in mature tRNA are introduced post-transcriptionally by specific tRNA-modifying enzymes. Transfer RNAs are also recognized by many components of the protein-synthetic machinery, such as aminoacyl-tRNA synthetases, ribosomes, several proteins involved in initiation, elongation, and termination of peptide chains, and by messenger RNA. Through their involvement in protein synthesis, tRNAs are intimately tied to a wide variety of cellular functions. It is known that certain tRNAs can have regulatory roles. Several classes of mutations in *Salmonella typhimurium* cause derepression of the histidine operon by their effect on histidine tRNA (19). A number of suppressor tRNAs have been characterized which insert an amino acid into polypeptides in response to a termination codon. The most intensively studied of these is the tyrosine suppressor tRNA ($\underline{su}_3^+$) of *E.coli*, a gene for which has been synthesized (20), based on the sequence of the precursor RNA (21), and is being characterized in detail. Along somewhat different lines, tryptophan tRNA acts in attenuation of the <u>trp</u> operon of *E.coli* (22), or as a primer for transcription of Rous sarcoma virus (23).

The tRNAs of *E.coli* have been best studied. Most of the tRNAs have been isolated and sequenced (17), their codon responses examined (24),

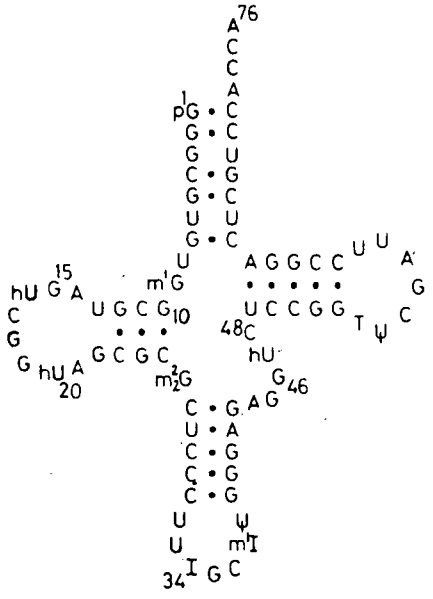


Figure 1. Cloverleaf structure of yeast $tRNA^{Ala}$

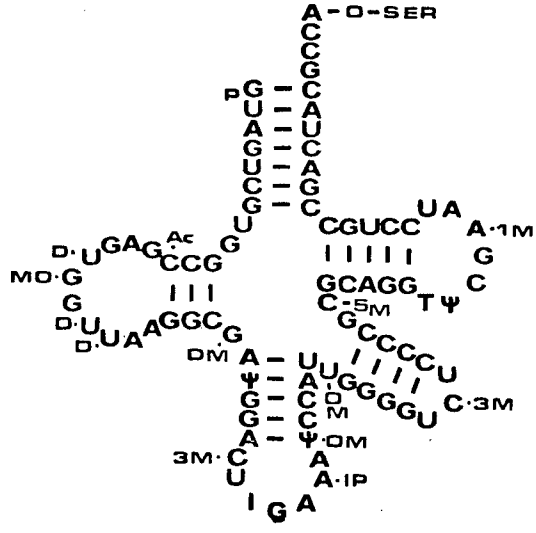


Figure 2. Cloverleaf structure of rat liver $tRNA_1^{Ser}$.

and their genes mapped on the *E.coli* chromosome (25). Mutant strains with altered tRNAs, including informational suppressors (26,28,29,30,31,32) and temperature sensitive lesions (27) have been isolated and characterized.

Some differences are apparent between transfer RNAs of prokaryotes and eukaryotes. The cloverleaf structures of isoaccepting tRNAs may be markedly different (for example, the extra arm of *E.coli* $tRNA^{Tyr}$ contains 8 nucleotides more than that of mammalian $tRNA^{Tyr}$) (17). As organisms become more complex, their tRNAs contain increasingly more modified nucleotides, up to 25% of the total (16). Differences are also seen in the types of modifications present (e.g. s^4U has been found only in prokaryotic tRNAs).

It is not clear how many of the modifications affect tRNA function in protein synthesis. Removal or alteration of a hypermodified base adjacent to the anticodon may result in loss of template binding activity or amino acid transfer function (24). However, $tRNA^{His}$ from his T mutants of *Salmonella typhimurium*, lacking two pseudouridine residues normally present in the anticodon region of wild type $tRNA^{His}$(33), is present in normal amounts, is aminoacylated normally, and appears to function like wild type in protein synthesis, although not in repression control (34). In this light, it is of value to know the sequences of many tRNA species, including modified nucleotides, toward an understanding of tRNA functions and the attributes making them suitable for these functions.

1.A.3 tRNA Genes

Transfer RNA genes are attractive model systems, and many studies on them have been undertaken in recent years. In *E.coli* there are

approximately 60 tRNA genes, mostly located in clusters at several locations on the chromosome (25,35), though genes for several tRNAs are found in the spacer regions between 16S and 23S rRNA genes and are co-transcribed with the rRNAs (36,43). In eukaryotes the tRNA genes are redundant, though the organization of genes for a particular isoacceptor may vary. There are about 320-440 tRNA gene copies in yeast (37). Digestion of yeast DNA with restriction endonucleases and separation give eight distinct fragments hybridizing tyrosine tRNA (38). This is in agreement with genetic studies which identified eight unlinked loci capable of mutation to tyrosine-inserting nonsense suppressors (39), indicating that these tRNA genes may be widely scattered. In *Xenopus laevis* there are about 8000 gene copies, estimated by hybridization kinetics to consist of 43 tRNA sequences each reiterated about 200 times. There is extensive clustering of genes. Most of the tRNA genes are grouped together with spacer DNA over long regions extending up to 10^5 base-pairs or more with an average spacer of about 800 base-pairs (40,41,42). Approximately 100 species of *Drosophila melanogaster* tRNA can be distinguished by RPC-5 column chromatography, some of which are homogeneic, differing only in extent of modification (44). In *D. melanogaster* estimates of gene number have varied, from 750 (45) to a more recent value of 590, with 59 tRNA sequences each 10-fold redundant (46). Studies on the organization of *Drosophila* tRNA genes have been primarily by *in situ* hybridization of radiolabeled purified tRNAs to polytene chromosomes from salivary cells (47,48,49,50). In this manner genes for a number of tRNAs have been located. Work with a recombinant plasmid containing *Drosophila* DNA has shown that there is some clustering of tRNA genes (51,52), in agreement with the results of Dunn *et al.* for $tRNA_{3b}^{Val}$ (49). However,

correlation of gene numbers from DNA-RNA hybridization experiments with the numbers of restriction fragments of *Drosophila* DNA hybridizing purified tRNA species (53) indicates that tandem redundancy like that found for 5S RNA and $tRNA_i^{Met}$ genes of *Xenopus laevis* (54,55,56), and 5S, 18S, and 28S rRNAs of *Drosophila* (57,58) is not general for *Drosophila* tRNA genes.

Drosophila melanogaster is biologically and genetically one of the best-studied eukaryotic organisms. The presence of polytene chromosomes has allowed extensive cytological characterization, with good correlation of genetic and cytological maps. Localization of tRNA genes in *Drosophila* by *in situ* hybridization of purified tRNAs to polytene chromosomes (47,48,49,50) and construction of recombinant bacterial plasmids containing *Drosophila* tRNA genes (51,59) have proceeded rapidly, allowing studies of these genes at two distinct and complementary levels, organization in the genome and fine structure of individual genes, to be carried out. Also, assignments of gene location by *in situ* hybridization have made possible genetic experiments such as construction of duplication and deficiency lesions spanning sites of tRNA genes, to study control of expression in the organism (49).

As part of a study on *Drosophila* tRNA genes carried on recombinant plasmids (59), sequence analyses of the serine isoacceptors $tRNA_4^{Ser}$ and $tRNA_7^{Ser}$ were undertaken. These tRNAs were of interest for several reasons. (a) Mutant serine tRNAs are potential suppressors of amber, ochre, or UGA nonsense mutations, as has been shown in bacteriophage T4 and the yeast *Schizosaccharomyces pombe* (26,60). (b) These *Drosophila* serine tRNAs have been characterized for codon specificity in ribosome binding experiments and shown to have different triplet responses, and were markedly different

on RPC-5 column chromatography (61). However, their modified base contents are very similar (61), and they are indistinguishable by *in situ* hybridization, by hybridization to recombinant plasmids in lysed bacterial clones on filters, and in solution (unpublished results, this lab). It was therefore of interest to know whether tRNAs $^{Ser}_{4,7}$ are from two sets of genes with similar sequences, or from one set of genes and differing only in extent of modification. (c) Further, there is a strong site for *in situ* hybridization on the X chromosome (50), the only such case from about 20 purified tRNAs studied to date. In addition to the site on the X, there are weak sites of hybridization on the left and right arms of the second, and the left arm of the third, chromosome. Comparative studies by DNA sequencing of $tRNA^{Ser}$ genes from the different chromosomes would be of interest. Toward this end, it is necessary to have accurate sequences for both serine tRNAs. In the present work, the sequence analysis of *Drosophila* $tRNA_7^{Ser}$ is presented.

PART B

Sequencing Transfer RNA

The first complete tRNA sequence, for yeast alanine tRNA, was reported by Holley and his co-workers in 1965 (62). They noted that the molecule could be drawn as a "cloverleaf" structure (Fig. 1) in which a number of standard Watson-Crick base pairs were present. This proposal was strengthened when further sequences completed shortly after were found to

fit the same pattern (63,64,65). Other patterns also emerged from this early work, particularly the site-specificity of certain modified nucleotides.

The sequencing strategy developed by Holley and his group for small RNAs involves (a) cleavage of purified RNA with nucleases of differing specificities to generate distinct sets of fragments, which are identified; and (b) partial digestions with the same enzymes to make larger fragments, which can be used to order the limit digest oligonucleotides (66). This strategy was developed for sequencing non-radioactive RNA. It was very good for identification of modified nucleotides, since the large amounts of fragments and nucleotides or nucleosides used could be analyzed by their UV spectra. On the other hand, the amount of material required limited the application of this strategy to those tRNAs that could be isolated in large quantity.

A major advance in sequencing nucleic acids was the introduction by Sanger of the use of the radioisotope phosphorus-32 to label nucleic acids (67) or their degradation products (68,69). This led to development of the microtechniques now commonly used in nucleic acid sequencing. Of particular value was the use of bacteriophage T4-induced polynucleotide kinase, to radiolabel at high specific activity the nucleic acids from organisms which could not be conveniently grown on a radioactive medium (70). As a consequence of such radiolabeling techniques, the amount of nucleic acid sample required for sequence analysis decreased by several hundred fold. Further development of microtechniques for sequence analysis of non-radioactive tRNAs after labeling *in vitro* has allowed the complete

sequence of a tRNA to be established using just several micrograms of purified material (70).

Sequencing strategies that have been used for radioactive RNAs are like that described above for non-radioactive RNA, and normally involve two steps: (a) limit digestion with base-specific nucleases followed by separation and identification of individual fragments; and (b) generation and analysis of larger, partial digestion products to order the limit digest products. Because the tRNAs of most higher organisms are not conveniently radiolabeled *in vivo*, they are commonly labeled *in vitro* using polynucleotide kinase. Only sequencing of *in vitro* labeled RNA will be discussed here.

Using the strategy above, limit-digest oligonucleotides from RNase T_1 or pancreatic RNase digestion are labeled and separated by two-dimensional electrophoresis, on cellulose acetate in the first dimension and DEAE-cellulose paper or thin layer plates in the second (70,71,72,11); or by electrophoresis on cellulose acetate strips, followed by homo-chromatography (71,72,73) or chromatography on PEI-cellulose thin layer plates (74). Identification of fragments is based on several lines of evidence. The first is the position of the fragment in the two-dimensional "fingerprint." Combined with analysis of the 5'-terminal ([^{32}P]-labeled) mononucleotide and the known specificity of the enzyme used in making the fragments, there is sufficient information for identification of many RNase fragments from a tRNA. For fragments five nucleotides or less in length, partial enzymatic digestion of [5'-^{32}P] oligonucleotides and mobility shift analysis of the [5'-^{32}P]-intermediates after one dimensional electrophoresis at acid pH on DEAE-cellulose paper allows identification

of fragments lacking internally modified nucleotides. Those with internal modifications can be identified by mobility shift analysis of successive intermediates in the partial digest after two-dimensional homochromatography, by comparison with standards from tRNAs of known sequence (70, 71,72,75). (Sequencing by two-dimensional homochromatography is discussed below.)

Ordering of limit fragments from tRNAs has been done (a) by sequencing large fragments generated by partial enzymatic or specific chemical cleavage; or (b) by sequencing terminally labeled intact tRNA. Step (a) is done by two-dimensional homochromatography, by two-dimensional polyacrylamide gel elctrophoresis (76), or by ladder gel analysis, that is by analysis on polyacrylamide gels of sets of base-specific partial digests generated either enzymatically (77,78) or chemically (79). Step (b) is done by ladder gel analysis (though two-dimensional homochromatography may be performed to obtain the terminal sequence). Two-dimensional homochromatography affords some discrimination in both dimensions. In the first, electrophoresis at pH 3.5 on cellulose acetate strips, the relative mobilities of oligonucleotide intermediates in a homologous partial digestion series which differ by a single nucleotide will depend both on the pK of the nucleotide by which they differ and on the size and base composition of the sequence common to them (70). In the second dimension, homochromatography on DEAE-cellulose plates, the fragments are displaced by oligomers of greater length in a partial hydrolysate of RNA (70,71,72). Mobility is roughly inversely proportional to fragment size, but purines can be reliably distinguished from pyrimidines in most cases (70,73,75). While the four common nucleotides can be "read" from the fingerprint

reliably in this manner, the differences between modified and unmodified bases are often subtle, resulting from slightly altered pK values and hydrophobic effects (70). Discerning a modified nucleotide within a fragment is done by electrophoretic mobility alone, since discrimination by homochromatography is limited. Thus, a modified nucleotide several residues from the 5'-terminal label of a [5'-^{32}P] oligomer may be identifiable only in the presence of a standard oligonucleotide of known sequence, if then. In the ordering of the identified limit digest oligonucleotides, two-dimensional homochromatography is reliable but the number of nucleotides which may be read from a single experiment is limited by the RNA hydrolysate used (73), and fragments 20 nucleotides or less in length are desirable. In two-dimensional polyacrylamide gel electrophoresis (76), the first dimension is run at pH 3.5 in a gel of sufficiently high porosity that separation of fragments in a limited size range depends on their size and base composition. Electrophoresis in the second dimension is through a denaturing gel of acrylamide concentration sufficiently high to effect a separation on the basis of size alone. The sequence assigned is based primarily on mobility shifts in the first dimension. This method is used to "read" pyrimidines as a complement to the method of Donis-Keller *et al.* for purines (77). It has some advantages over two-dimensional homochromatography; it can allow one to deduce longer sequences from a single experiment, and the size of fragment used for analysis is flexible since acrylamide concentrations are readily manipulated. Overall, however, this method is less effective than other techniques available, and is not used extensively. Analysis of partial digests of terminally radiolabeled tRNA by electrophoresis on denaturing polyacrylamide gels (ladder gel analysis)

is generally the method of choice for ordering limit enzymatic digest oligonucleotides. A purified RNA carrying a terminal radiolabel (normally ^{32}P or ^{125}I) at a unique site, the point of reference, is required for partial digestions, which may be done enzymatically (70,77,78) or chemically (79). Subsequent size separation of digestion products is by electrophoresis in adjacent slots of a denaturing polyacrylamide gel, followed by exposure to X-ray film. All fragments located by autoradiography carry the point of reference; the identity of the opposite terminus is determined by the specificity of the nuclease or reagent used in digestion. Since the sets of fragments are ordered by size following electrophoresis, it is possible to deduce from the distinctive band pattern on the autoradiograph much of the tRNA sequence.

Partial enzymatic digestions of radiolabeled tRNA are performed using ribonucleases RNase T_1 (G-specific), RNase U_2 (A-specific), RNase A (cleaves at pyrimidine residues), and RNase Phy I from the slime mold *Physarum polycephalum* (80; cleaves most sequences except CpN); and with hot formamide or alkali, which cleave randomly at all but ribose-methylated nucleotides, making a reference "ladder" for ordering oligonucleotides from the partial enzymatic digests. Ladder gel sequence analysis using partial enzymatic hydrolyses is not effective for distinguishing most modified nucleotides. The phosphodiester bonds of many modified nucleotides are cleaved slowly if at all by enzymes which specifically cleave 3'-esters of the unmodified counterparts. Analysis is further complicated by the difficulty in getting random cleavage by a nuclease at its potential cleavage sites, due to residual tertiary structure maintained by tRNAs even in 7M urea at elevated temperatures (70,81).

In ladder gel analysis utilizing chemical hydrolyses, independent digestions are carried out using dimethyl sulfate (G-specific), diethyl pyrocarbonate (A-specific), and hydrazine in the presence or absence of sodium chloride (C+U or U specific), to remove the base, followed by treatment with aniline to cleave the sugar-phosphate backbone. Ladder gel analysis of partial chemical hydrolysates as described by Peattie (79) appears to give more reliable sequence information than enzymatic digests, particularly for distinguishing C and U, but is not without shortcomings. Distinguishing the various nucleotides depends on the rates of chemical modification of nucleotides by the reagents used. Nucleotide modifications can greatly affect these rates of modification. Pseudouridine does not give a band on these ladder gels (79). Ribothymidine and 5-methylcytidine would be expected to react much more slowly with hydrazine than uridine or cytidine due to steric hindrance by the methyl group at C-5 (79,82). While the identities of some modified nucleotides could probably be established, such assignments might often be unreliable. Thus, while sufficient for ordering of oligonucleotides, the ladder gel techniques are not sufficient to establish a complete, unambiguous tRNA sequence, and must be used in conjunction with other methods.

An independent method for gel sequencing of tRNA was recently introduced by Stanley and Vassilenko (83) that is more useful as a primary sequencing method. It makes use of the random cleavage by formamide of RNA seen at elevated temperatures. Purified tRNA (5'-phosphorylated) is digested in formamide under conditions such that singly cleaved or uncleaved products predominate. A family of intermediate-sized fragments is generated, with each size class evenly represented. Each internal cleavage of an intact tRNA molecule creates a 5'-fragment bearing a 5'-phosphate and a 3'-cyclic phosphate, and a 3'-fragment that has a 5'-hydroxyl and

3'-terminal CCA. Polynucleotide kinase catalyzes the transfer of a phosphate from [γ-^{32}P] ATP to fragments with a 5'-hydroxyl, that is those with a 3'-terminal CCA. The CCA is thus the point of reference in subsequent analysis of [^{32}P]-oligonucleotides, though not radioactively labeled. The [^{32}P]-oligomers are electrophoresed through denaturing polyacrylamide gels, the bands located by autoradiography, excised, and the RNA eluted. Each fragment is digested to mononucleotides and the 5'-terminal (radiolabeled) nucleotide identified. Because the [^{32}P]-oligonucleotides are ordered by size after polyacrylamide gel electrophoresis and each such fragment has a 3'-CCA, analysis of the 5'-terminal nucleotides of the ordered fragments allows direct determination of the sequence. In this way, much of a tRNA sequence can be established. In principle, it would be possible to obtain a sequence complete except for the 5'- and 3'-terminal nucleotides (which are phosphorylated or present as a nucleoside and not labeled by polynucleotide kinase, respectively) (83), and nucleotides sharing the 3'-phosphate of a ribose-methylated nucleotide (formamide degradation involves formation of a 2', 3'-cyclic phosphate).

The particular strength of this method of sequence analysis is that it allows direct evaluation of modified nucleotides that would not be accessible by standard procedures described above. Its limitations are primarily those of labeling using polynucleotide kinase. The efficiency with which fragments are labeled varies with the 5'-terminal nucleotide (or oligonucleotide), and decreases in the presence of extensive secondary or tertiary structure.

None of the previously developed sequencing methods is generally sufficient to obtain a complete tRNA sequence; combinations of techniques must be employed for a particular tRNA. Thus, it is reasonable to try to

develop new, more general strategies extending those previously used. Sequencing of certain tRNAs remains difficult, due to the high incidence of modified bases, the tight tertiary structure, and the lack of an array of specific enzymes or chemical reaction conditions adequate for many situations encountered. Thus a strategy was devised for tRNA sequencing involving a novel combination of techniques. The basis of the strategy is to use the formamide degradation method as the primary sequencing technique instead of analysis of radiolabeled limit enzymatic digestion products, due to the advantage it offers in identifying modified bases. The formamide degradation sequencing method is limited, as described above. However, its limitations are different from those of the ladder gel sequencing method (above); thus these two methods complement one another in deriving a tRNA sequence. Use of ladder gel sequence analysis in conjunction with analysis of formamide degradation products will in principle yield a complete sequence from a limited amount of purified tRNA. The work presented here describes the sequence analysis of a previously uncharacterized tRNA following this strategy.

Chapter 2

MATERIALS AND METHODS

General:

Unless specifically indicated, all chemicals were reagent grade. Unless noted, reactions and sample preparations were carried out in 1.5 ml polypropylene conical test tubes with caps (referred to as 1.5 ml test tubes in the text). Centrifugations were in a mini-centrifuge (Micro-Scientific Co.) attached to a variable power source. Here, the applied voltages are given rather than rpm. Only glass distilled water was used.

Abbreviations:

TEMED	=	N, N, N', N' - tetramethylethylene diamine
ATP*	=	[γ - ^{32}P] ATP
*	=	$^{32}PO_4$
pN	=	nucleoside-5'-phosphate
pNp	=	nucleoside-5', 3'-bisphosphate
XC	=	xylene cyanol
BPB	=	bromphenol blue
formamide/dye mix	=	deionized formamide, 1% (w/v) XC, and 1% (w/v) BPB, 13:1:1 by volume.

Abbreviations adopted for modified nucleotides were those used in ref. 17 with the exception of X which is used in this text to refer to the unknown modified nucleotide occupying position 26 of *Drosophila* $tRNA_7^{Ser}$, and Y, which in Fig. 8 indicates a pyrimidine nucleotide.

Synthesis of [γ - ^{32}P]ATP by enzyme-catalyzed exchange:

The method is based on that of Glynn and Chappell (108), as modified by Gilbert and Maxam (82).

Preparation of enzymes: Rabbit muscle glyceraldehyde-3-phosphate dehydrogenase (GAPD) was obtained from Worthington Biochemicals, Inc., as a precipitate in 3.0M ammonium sulfate (on suspension, 11 mg/ml and 75 units/mg). The precipitate was suspended with a disposable pipettor tip, 100 μl transferred to a 1.5 ml test tube and centrifuged 10 min. at 55V, 4°. The supernatant liquid was removed and discarded. The protein pellet was suspended in 100 μl of 3.2 M ammonium sulfate (pH 8.0), 50 mM TrisHCl (pH 8.0), 10 mM 2-mercaptoethanol, 1 mM EDTA, and 0.1 mM NAD, stored 3 hr. at 4°, then centrifuged as before. The supernatant liquid was discarded, the pellet suspended in 100 μl of the buffered ammonium sulfate solution, and stored at 4°. 3-Phosphoglycerate kinase (PGK) was obtained from Sigma Chemical Co. as a precipitate in 3.0 M ammonium sulfate (on suspension, 12 mg/ml and 2100 units/mg). The precipitate was suspended with a disposable pipet tip, 10 μl removed and added to 90 μl 3.6 M ammonium sulfate at 0° in a 1.5 ml test tube. The suspension was centrifuged 10 min. at 55 V, 4°, the supernatant liquid removed and discarded. The pellet was suspended

in 100 μl of 3.2 M ammonium sulfate (pH 8.0), 50 mM Tris HCl (pH 8.0), 10 mM 2-mercaptoethanol, 1 mM EDTA, stored 3 hr. at 4°, and centrifuged as before. The supernatant liquid was discarded, the pellet resuspended in 100 μl of the same solution, and stored at 4°. The enzyme suspension had concentrations of 11 mg/ml GAPD and 1.2 mg/ml PGK. There was no apparent loss of either enzymatic activity after 6 months at 4°.

Synthetic reaction: Prior to use, a 1.5 ml test tube was rinsed well with water. Carrier-free $H_3\ ^{32}PO_4$ (3-10 mCi; New England Nuclear) was dried in the tube, under aspirator vacuum in a dessicator containing P_2O_5 and sodium hydroxide pellets. The reaction mix (50 μl) was 50 mM Tris HCl (pH 8.0), 7 mM $MgCl_2$, 1 mM EDTA, 2 mM glutathione, 0.4 mM 3-phosphoglycerate (Sigma Chemical Co., Grade I), and contained 2-4 nmole of ATP (P-L Biochemicals, Inc.). EDTA at 1 mM was needed for consistently high yields apparently because of heavy metal contaminants in some batches of carrier-free $H_3\ ^{32}PO_4$. ATP was previously tested for contaminating ADP (which decreases the extent of reaction) by chromatography on polyethyleneimine-cellulose plates (PEI-cellulose, Brinkmann Instruments, Inc.) Chromatographic development was with 0.25 M $K_2H\ PO_4$ (pH 7.0) at room temperature for 2.5-3 hr. Under conditions where the plate was heavily overloaded with ATP, no ADP was detected under UV light.

GAPD and PGK, added to start the exchange reaction were prepared as follows immediately before addition. The precipitates in ammonium sulfate solution were suspended with a disposable pipet tip, 5 μl from each suspension mixed in a 1.5 ml test tube, and centrifuged 10 min. at 55 V, 4°. The supernatant liquid was discarded, and the pellet dissolved in 50 μl H_2O.

From this solution, 1 μl (containing about .08 unit GAPD and 2.5 units PGK) was added to the reaction mix. The reaction was incubated 20 min. at room temperature, and was terminated by adding 5 μl 0.1 M EDTA and heating 3 min. at 100°. The tube was then chilled on ice, centrifuged 5 min. at 40 V, 4°, and stored at -20°. During the reaction, assay aliquots of about 0.2 μl were taken at 0 (before enzyme addition), 10, and 20 min., and spotted onto a PEI-cellulose plate, then developed as above. The $^{32}PO_4$ and [γ-^{32}P]ATP were located by autoradiography. Marker dyes were not necessary to align the TLC and autoradiograph as some residual radioactivity located the origin for each assay sample. Extent of reaction was determined in the following manner. [γ-^{32}P]ATP and $^{32}PO_4$ were collected from the PEI-cellulose plates: the cellulose was removed from its plastic backing by scraping the appropriate spots and gathered by suction into a 1 ml disposable pipettor tip which was plugged with glass wool and inserted into a vacuum line (similar to the apparatus described in ref. 72). The radiolabeled material was eluted with 3.0 ml 2 M KCl and Cerenkov radiation determined in a liquid scintillation counter. If the sample contained more than about 5 x 10^5 cpm, an aliquot was removed and the Cerenkov radiation determined. Essentially all radiolabel is present as either $^{32}PO_4$ or [γ-^{32}P]ATP. By determining the fraction of radiolabel present in [γ-^{32}P]ATP in an assay aliquot, and knowing the amount of H_3 $^{32}PO_4$ placed in the reaction mix, the yield of [γ-^{32}P]ATP can be calculated. The specific activity is the amount of radiolabel in ATP divided by the molar amount of unlabeled ATP placed in the reaction mix. Conversion of H_3 $^{32}PO_4$ to [γ-^{32}P]ATP was usually 60-70%, giving specific activities of about 1500 Ci/mmole.

Autoradiography:

Autoradiography was as follows. The polyacrylamide gels or thin layer plates were marked with radioactive dye in a distinctive pattern to allow alignment of the autoradiograph with the gel or plate. The dye spots contained enough [^{32}P]-PO_4 to give distinct spots in the exposure time used.

Gels were covered with Saran Wrap, and the dye was applied to AE-cellulose paper (Whatman AE-30) which was taped to the covered gel. The gel was then exposed in a film holder to "no-screen" NS-5T X-ray film (Kodak). For exposure times longer than 4 hr., the exposure was at -20° or -70°. For TLC plates (cellulose, DEAE-cellulose, or PEI-cellulose), the dye was spotted directly onto the plates, which were then covered and exposed to the X-ray film as above. A gel band containing 20,000 [^{32}P]-dpm was clearly visible after a 10 min. exposure to NS-5T film. Somewhat more radiolabel was required in a spot on a thin layer plate, as the area was larger than that of a gel band. Screen film (Kodak-XR-1) was occasionally used for autoradiography. This film was adequate without a calcium tungstate intensifying screen, but required longer exposures (roughly 5-10 fold) than was necessary with "no-screen" film. On occasion, an X-ray intensifying screen was used with "screen" film for autoradiography (84). In that case, the film was activated by flashing with a pulse of light from a photographic flash unit covered with a yellow gelatin filter (Wratten 22, Kodak), placed between the intensifying screen and the wrapped gel or TLC plate in a film holder, and exposed at -70° (84). Sensitivity of the screen film for [^{32}P] was greatly increased by this procedure, and a spot on a thin layer plate containing about 50 Cerenkov cpm was clearly visible after two days' exposure.

The exposed X-ray film was developed by immersing for 7 min. in developer, 2 min. in running water, and 10 min. in fixer. (The film, opaque

after developing, becomes clear after 1-2 min. in the fixer solution, but was left longer to ensure consistent quality of the autoradiographs.)

Purification of $tRNA_7^{Ser}$ from Drosophila melanogaster (Samarkand strain):

The tRNA was purified by established methods (85) and provided for sequencing experiments by Dr. Ian Gillam.

Sequencing $tRNA_7^{Ser}$ by the formamide degradation method (83):

Deionized formamide: the conditions were derived from those of Brownlee and Cartwright (86). Formamide was added to dry Dowex-1 and Dowex-50 (well washed with dH_2O), left for several minutes with stirring, then filtered through a nitrocellulose filter and stored in glass at room temperature. Dowex-1 was equilibrated before use in .01 M NaOH, Dowex-50 in .01 M HCL.

Preparation of RNA fragments: to 5 μg of $tRNA_7^{Ser}$ in 5 μl H_2O in a 1.5 ml test tube was added 12.5 μl (2.5 volumes) deionized formamide. The RNA was hydrolyzed by heating 15 min. at 100°, then cooled on ice. Forty-five μl (2.5 volumes) of 95% ethanol was added, and the sample stored at -70° overnight. The RNA fragments were collected by centrifugation (15 min. at 70V, 4°), the supernatant liquid removed, and the pellet air-dried at room temperature.

Labeling of tRNA fragments using [γ-^{32}P] and polynucleotide kinase: the reaction mix (10 μl), in the same tube, contained 5 μg of partially hydrolyzed tRNA, 85 mM Tris HCl (pH 8.0), 9 mM $MgCl_2$, 2 mM EDTA, 24 mM 2-mercaptoethanol, 40 μM ATP* (specific radioactivity 1300 Ci/mmole), and

2 units of polynucleotide kinase (P-L Biochemicals, Inc.). Immediately before addition of the enzyme, the reaction mix was heated 1 min. at 90° and chilled 1 min. on ice. The reaction was incubated 30 min. at 37°, and was terminated by adding 15 µl formamide/dye mix and heating 1 min. at 100°. Samples were stores at -20° until electrophoresis.

Polyacrylamide gel electrophoresis of [^{32}P] RNA fragments:

polyacrylamide gels were 20% acrylamide (w/v), 1% N, N'-methylene-bis-acrylamide (w/v), 7 M urea, 90 mM Tris-borate, 1 mM EDTA (pH 8.3). This solution was prepared just prior to use, filtered through a nitrocellulose filter (0.45 micron pore size), then chilled on ice under aspirator vacuum for 5-10 min. Ammonium persulfate was added to 0.56% (w/v), TEMED to 0.09% (v/v), and the solution de-gassed briefly. The solution was immediately poured between 20 cm X 40 cm glass plates separated by 0.05 cm Teflon spacers. The slots (made with a Teflon slot former) had dimensions 1 x 1 x 0.05 cm. Following polymerization (complete in 10-15 min.), the slots were washed thoroughly with running buffer (90 mM Tris-borate, 1 mM EDTA, pH 8.3), then filled with the same buffer containing 6.5 M urea. The gel was pre-run 30 min. at 300V, and slots were rinsed with buffer (no urea). Samples were loaded using a drawn-out capillary tube, and electrophoresis carried out at no more than 16 watts of power (constant voltage at about 1200V). Normally, anion exchange paper (Whatman AE-23) was inserted beneath the gel to bind unreacted ATP* and $^{32}PO_4$ from the reaction mix. For long runs, the running buffer in the electrode chambers was changed every 24 hr.

After electrophoresis, the gel was exposed to X-ray film to visualize the [^{32}P]RNA bands (see Methods, Autoradiography). The gel and autoradiograph were aligned by the radioactive marker dye, with the autoradiograph beneath the back glass plate carrying the gel, and the bands excised in order. (Good alignment is critical, and requires special care due to the thickness of the glass plate.)

Analysis of 5'-terminal nucleotides of [5'-^{32}P] riboligonucleotides: each band was excised from the gel and placed in a scintillation vial with 0.50 ml 0.3 M NaCl, 0.1% SDS, and eluted at least 16 hr. at 4° (80-95% recovery of tRNA-size fragments is obtained after 24 hr.). After elution, 0.15 ml of eluate from each band was added to 0.50 ml of 95% ethanol containing 50 μg of *E. coli* tRNA (Schwarz-Mann) as carrier, in a 1.5 ml test tube. The samples were mixed by inverting the capped tubes several times, stored overnight at -70°, and the RNA collected by centrifugation (10 min. at 70V, 4°). The supernatant was removed with a Pasteur pipet and discarded. The pellet (clearly visible) was air-dried for 1-2 hr. at 37°, then dissolved in 5 μl 0.2M NaOH and incubated 16 hr. at 37° (the liquid had usually entirely evaporated into the tube in this time). Each sample was neutralized by adding 5 μl 0.24 M formic acid, and 3 μl of each was spotted in order onto a PEI-cellulose plate at 1 cm. intervals (plates prepared as described by Southern and Mitchell, ref. 87). The plate was washed in distilled water, air-dried, developed with 0.80 M ammonium sulfate at room temperature and exposed to X-ray film. The *pNp's were identified from their mobilities, and much of the tRNA sequence could be deduced from the autoradiograph.

Nucleotides suspected of being modified were analyzed further as *pN's. Of the 0.5 ml of eluate (above) from each gel slice, 0.30 ml was placed with 0.75 ml of 95% ethanol and 50 μg of carrier tRNA in a 1.5 ml test tube, mixing by inverting several times. The sample was chilled at -20° overnight and collected by centrifugation (10 min. at 70V, 4°). The supernatant was removed and discarded after determining that a pellet was present. The sample was air-dried at 37°, then dissolved in 10 μl of 20 mM NH_4 acetate (pH 4.6) containing 0.2 μg nuclease P_1 (obtained from Calbiochem) (70,88). The digestion ran 2 hr. at room tempeature, and was terminated by freezing. Samples were stored at -20° until use.

Analysis of ribose-methylated nucleotides: initial digestion as above with nuclease P_1 of [^{32}P]RNA fragments gave dinucleotides, $^{*}pN_mpN$, as products under conditions where hydrolysis was otherwise complete (88). In this case, digestion was with 6 μg of nuclease P_1, in 20 mM NH_4 acetate (pH 4.6) at 37° for at least 7 hr. (70,89). The increases of 3.5-fold and 30-fold in reaction time and enzyme concentration resulted in nearly quantitative hydrolysis of the dinucleotides to give $^{*}pN_m$'s as product.

Analysis for N^4-acetylcytidine: a [^{32}P] nucleotide sample thought to be pac^{4}C was divided into two fractions. One was the control; the other was dried, dissolved in 4 μl 0.1 M NaOH, incubated 30 min. at 37°, then neutralized with 2 μl 0.24 M formic acid. The half-life of ac^{4}C in 0.1 M NaOH at 37° is 6.8 min. (90), so most ac^{4}C should be converted to C in this time.

Characterization of *pN's by thin layer chromatography:

initial characterization of *pN's was by one-dimensional chromatography. Nucleotides were chromatographed at room temperature in solvent A (66 ml isobutyric acid: 1 ml conc. NH_4OH: 33 ml dH_2O) and solvent B (100 ml 0.10 M Na_2HPO_4, pH 6.8: 60 gm ammonium sulfate: 2 ml 1-propanol) on 20 x 20 cm cellulose TLC plates (E. Merck, purchased through Brinkmann Instruments, Inc.). Non-radioactive standards (1 µl of a mixture containing pA, pG, pU, and pC at 0.2 A_{260} units of each per µl; and 1 µl of appropriate modified nucleotide standards, when available, at about 0.2 A_{260} units per µl) were spotted with the *pN samples, chromatographed, and located under UV light. Spotting samples of no more than 5 µl (preferably 1-3 µl) at 1 cm intervals permitted analysis of up to 17 samples on one plate, leaving 2 cm clear at each side to minimize edge effects. Chromatographic development was continued until the eluting solvent was within 1 cm of the top of the plate (about 8 hr. with solvent A, 6 hr, with B), and the *pN's were located by autoradiography (Methods, above). Most nucleotides could be identified at this point by comparison of their mobilities to those of pA(solvent A) or pU (solvent B) (Table 3 in ref. 70). Any further characterization was by two-dimensional TLC, essentially as described (70). Of particular importance, it was found to be necessary to allow at least 30 hr. at room temperature in a fume hood for solvent A to evaporate from plates after the first dimension. The *pN's were identified by their positions relative to the standards, if possible (Fig. 2, ref. 70).

Nucleoside 3', 5'-bis [5' - ^{32}P] phosphates (*pNp's) were analyzed by one-dimensional chromatography as above, and identified after

autoradiography by comparison to the mobilities of pAp (solvent A) or pUp (solvent B) (Table 2, ref. 70). The standards were *pNp's identified by chromatography on PEI-cellulose (as above) in a formamide degradation sequencing experiment on a tRNA of known sequence (yeast $tRNA_{I}^{Gly}$).

Analysis of [5' - ^{32}P]tRNA by partial digestion with base-specific ribonucleases and gel electrophoresis:

Dephosphorylation of tRNA: The reaction mix (5 µl), in a 1.5 ml test tube, was 20 mM Tris HCl (pH 8.3), 0.2 mM EDTA, and contained 2 µg of $tRNA_{7}^{Ser}$ and 0.026 unit of bacterial alkaline phosphatase (BAPF; Worthington Biochemicals, Inc.). The mix was heated 30 sec. at 90° with the enzyme present, chilled on ice, then incubated 40 min. at 37°. The reaction was stopped by adding 1 µl of 50 mM nitrilotriacetic acid (pH 7.0) (Aldrich Gold Label) and heating to dryness at 90° in an open tube (70).

Labeling (5' - OH)tRNA using ATP* and polynucleotide kinase: the polynucleotide kinase reaction was carried out in the same tube. The reaction mix (10 µl) was 70 mM TrisHCl (pH 8.0), 8.5 mM $MgCl_2$, 24 mM 2-mercaptoethanol, 43 µM ATP* (500 Ci/mmole) and contained 2 units of polynucleotide kinase. Enzyme was added to initiate the reaction immediately after heating 30 sec. at 90° and chilling on ice. The reaction was incubated 15 min. at 37° and terminated by adding 15 µl of formamide/dye mix and heating 1 min. at 90°. This sample was loaded onto a pre-run denaturing polyacrylamide gel (10% acrylamide, 0.5% methylene-bis-acrylamide,

7M urea, 90 mM Tris-borate, 1 mM EDTA, pH 8.3; 38 x 20 x 0.15 cm, slots 1 x 1 x 0.15 cm). The running buffer was 90 mM Tris-borate, 1 mM EDTA, pH 8.3. Electrophoresis continued until the XC dye had nearly left the gel (8 hr. at 500 V, 25 mA). Excess ATP* and $^{32}PO_4$ were trapped on anion exchange paper, as described above. The [5' - ^{32}P]tRNA band was located by autoradiography, excised, and eluted with 0.60 ml 0.3 M NaCl, 0.1 % SDS, overnight at 4°. The eluate was divided into three fractions, 30 μg of carrier tRNA and 2.5 volumes of 95% ethanol added to each, and the samples stored at -70° overnight. The RNA was collected by centrifugation (10 min. at 70V, 4°), the supernatant liquids discarded, the pellets dried briefly under aspirator vacuum, and each sample dissolved in 5 μl dH_2O.

Partial hydrolyses with base-specific ribonucleases:

ribonuclease T_1 and RNase U_2 were from Sankyo, RNase A from Calbiochem, and RNase Phy I from Enzo Biochem. These enzymes specifically cleave 3'-esters of G, A, pyrimidines, or of any nucleotide except C, respectively, under the conditions used (77,78,80,91,92). Enzymatic hydrolyses were carried out in the following manner. Each reaction contained 6 μg RNA (8000 Cerenkov cpm). RNase T_1 and RNase A reaction mixes (10 μl) were 20 mM sodium citrate (pH 5.0), 1 mM EDTA, 6.3 M urea (Schwartz-Mann ultra pure), containing 0.01 unit RNase T_1 and 0.001 unit RNase A activity. Reactions ran 20 min. at 50°; they were stopped by chilling on ice, and XC and BPB dyes were added to 0.05%. RNase U_2 reaction mixes (5 μl) were 50 mM sodium acetate (pH 4.5), 4 mM EDTA, and contained 0.01 unit of enzyme. Reactions ran 20 min. at 0° and were stopped by adding 5 μl

formamide-dye mix and heating 1 min. at 100°. RNase PhyI reaction mixes (5 μl) contained 10 mM sodium acetate (pH 4.5), 1 mM EDTA, and 0.3 unit of enzyme. Reactions ran 20 min. at room temperature and were stopped as described for the RNase U_2 reactions. A random mixture of all possible intermediate-sized hydrolysis products for reference was generated by heating the RNA sample in 5 μl of 80% formamide, 40 min. at 100°. The control reaction (minus enzyme) was carried out as described for RNases T_1 and A.

The samples were loaded onto a denaturing 20% polyacrylamide gel (described above), 0.5 mm thick, and electrophoresed at 1000 V for 5 or 30 hr. to analyze the 5'- or 3'-termini, respectively. The gels were exposed to no-screen X-ray film (Kodak NS-5T) for 3 days at -20%, at which time reference ladder bands and most bands arising from enzymatic hydrolyses were visible in the autoradiograph.

Two dimensional homochromatography (wandering spot analysis):

Two dimensional homochromatography was performed as described by Jay *et al.* (73), except for the partial hydrolyses. Electrophoresis on cellulose acetate strips (Schleicher and Schuell No. 2500, 3 x 55 cm) was for 2 hr. at 2000 V on a Shandon high voltage electrophoresis apparatus, using a pH 3.5 buffer of 5% acetic acid adjusted with pyridine. For the homochromatography dimension, Homomix V (73; provided by Dr. C. Astell) and Polygram Cel 300 DEAE-cellulose plates (Macherey-Nagel and Co.) were used. The [^{32}P]RNA fragments were oligonucleotides from a formamide degradation experiment purified by gel electrophoresis. Partial hydrolyses

were in 100% deionized formamide (5 μl), 2.5 hr. at 100°. Samples of 2 μl (containing 4500 dpm of [^{32}P]-oligonucleotide) were applied to the cellulose acetate strips. These samples are absorbed slowly by the strips relative to aqueous samples of equal volume, but the formamide does not affect electrophoresis. Autoradiography was for one week at -70°, using pre-fogged Kodak XR-1 X-ray film and an intensifying screen (84; Methods, Autoradiography).

RESULTS

The nucleotide sequence of $tRNA_7^{Ser}$ from *Drosophila* was analyzed primarily by methods derived from those of Stanley and Vassilenko (83). [5'-^{32}P]-labeled fragments were generated, isolated from polyacrylamide gels (Fig. 3), and analyzed for their 5' terminal nucleotides. Most such nucleotides could be identified by their chromatographic mobilities as *pNp's on PEI-cellulose (Fig. 4). Preliminary experiments with *pNp's made by labeling nucleoside 3'-monophosphates (93), and *pNp's from [5'-^{32}P] oligomers generated by the formamide degradation method from yeast $tRNA_I^{Gly}$, the sequence of which is known (94), allowed the chromatographic mobilities of pAp, pGp, pCp, pUp, pψp, and pTp to be determined (results not shown). These results have been confirmed and extended in the work presented here by identifying the corresponding pN's by chromatography on cellulose thin layer plates. Chromatographic mobilities of a number of *pNp's on PEI-cellulose are presented in Table 1. A number of *pN's were identified on the basis of their chromatographic mobilities in two solvent systems (Methods), compared with those listed in Table 2 of ref. 70. Though pseudouridine could be identified by its mobility as a *pNp, this was confirmed by one-dimensional chromatography of the *pN on cellulose plates in solvents A and B (Methods). Other nucleotides identified by one-dimensional TLC were pG_m, pU_m, pD, pT, pi^6A, pm^1A, and pm^3C (Fig. 5). Confirmation for pD and pm^3C was by two-dimensional chromatography (Methods). Inosine [5'-^{32}P]

Figure 3. Polyacrylamide gel electrophoresis of [^{32}P]RNA fragments.

[^{32}P]RNA samples prepared by hydrolysis in formamide and post-labeling as described in Methods were electrophoresed through denaturing polyacrylamide gels. The gel running buffer was 50 mM Tris-borate, 1 mM EDTA, pH 8.3, for (a) and (b); and 90 mM Tris-borate, 1 mM EDTA, pH 8.3, for (c). Electrophoresis was at 1000V for (a) 8.6 hr., or (b) 21 hr.; or (c) at 1200V for 40 hr.

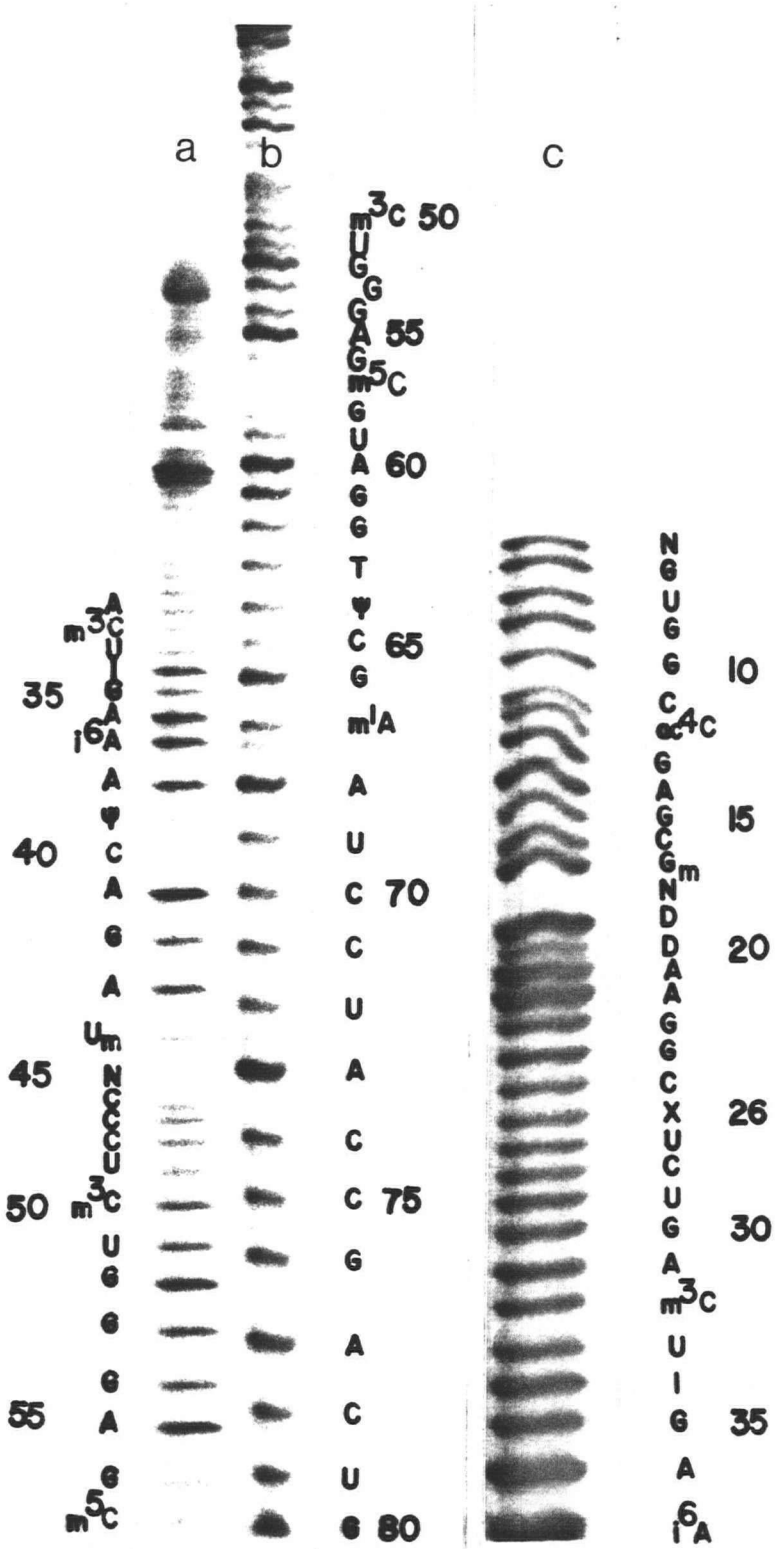
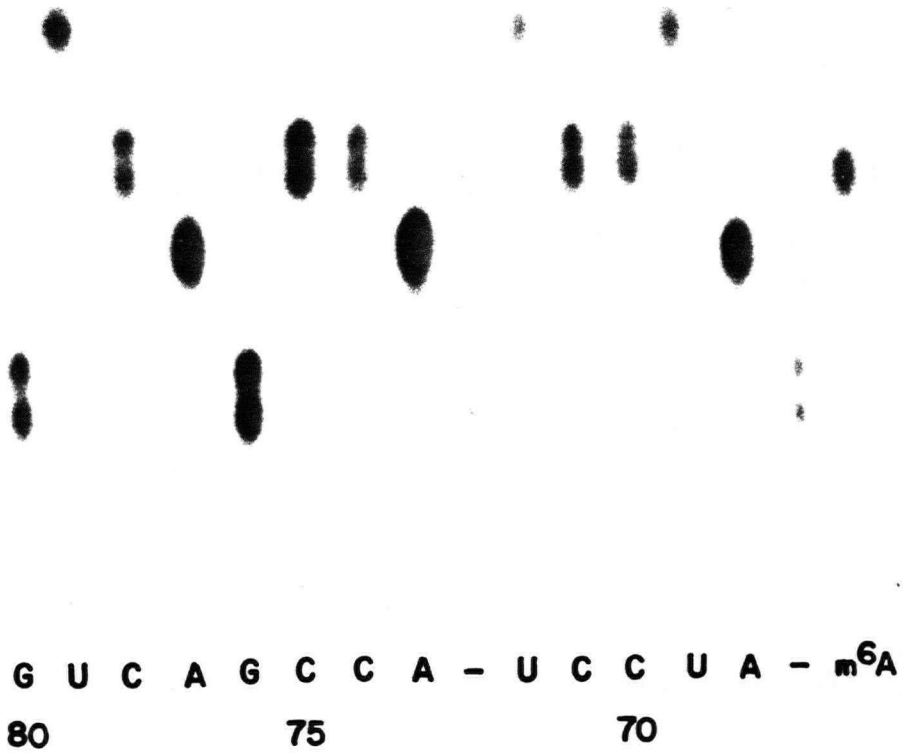

a
b
c
35
40
45
50
55
50
55
60
65
70
75
80
10
15
20
26
30
35

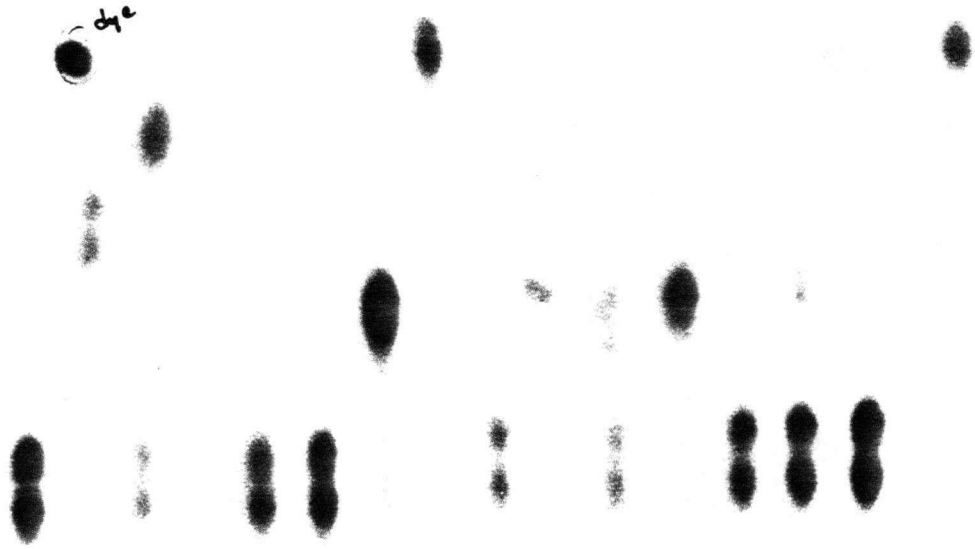
Figure 4. PEI cellulose chromatography of *pNp's.

[^{32}P]RNA fragments separated by polyacrylamide gel electrophoresis as shown in Figure 4 were analyzed for their 5'-terminal nucleotides as described in Methods, and are represented by the standard abbreviations used in Ref. 17.

a

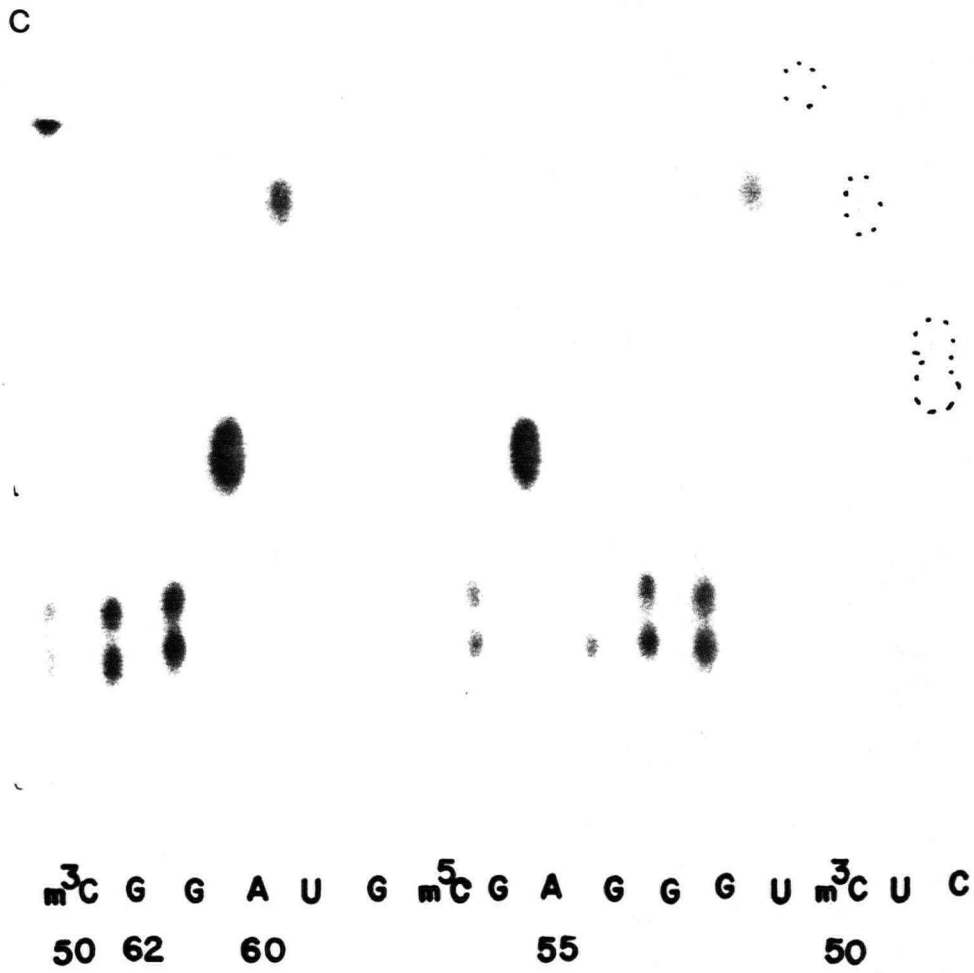


b

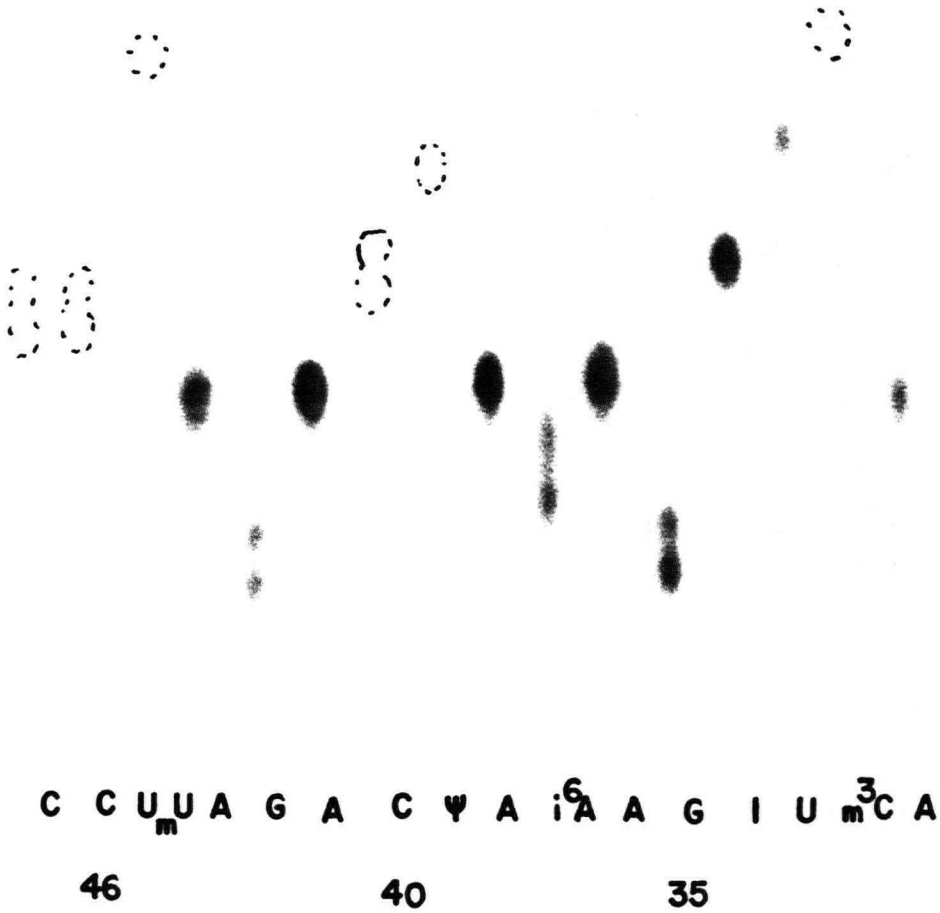


G C Ψ T G G A U G m5C G A G G G U

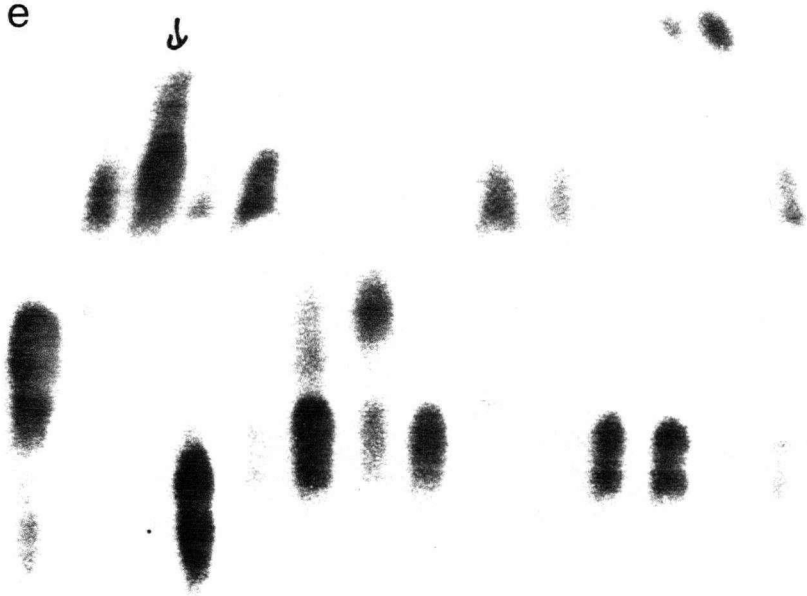
65 60 55


C
m3C G G A U G m5C G A G G G U m3C U C
50 62 60 55 50

d



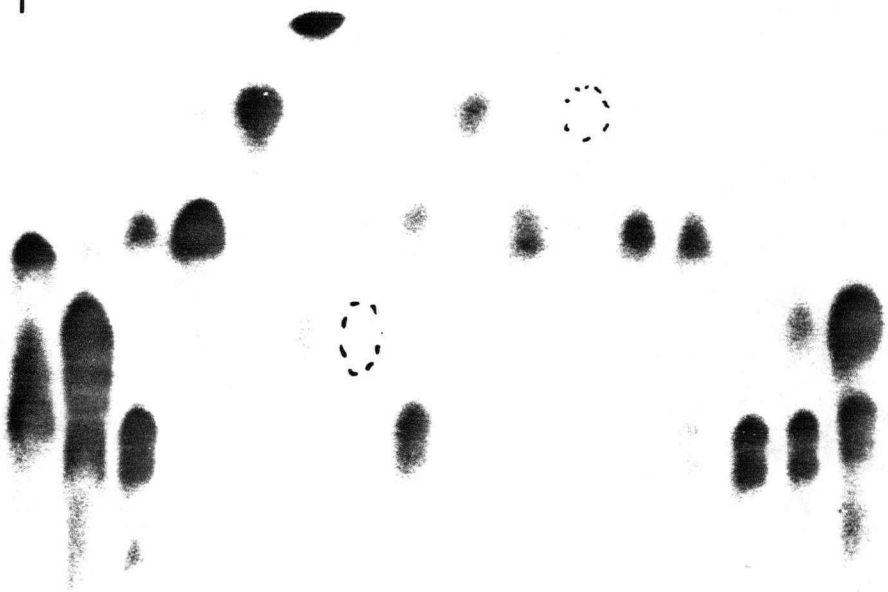
e



A D D G_m G C G A G C C G G U G/C

20 15 10

f



I^6A A G I U m^3C A G U C U X C G G A

35 30 25

Table 1

Mobilities of Nucleoside-5', 3'-Bisphosphates on PEI-Cellulose[a]

pNp	R_{pUp}[b]
pUp	1.00
pTp	1.00
pψp	0.94
pDp	0.85 (1.1)[c]
pAp	0.66
pm^6Ap	0.75
pi^6Ap	0.52, 0.61[d]
pIp	0.84
pCp	0.79, 0.83[d]
pm^3Cp	1.12
pGp	0.40, 0.47[d]

a. Developed with 0.80 M ammonium sulfate (pH 5.3) at room temperature.

b. Mobility relative to pUp.

c. The value in parentheses corresponds to an apparent degradation product of D from alkaline hydrolysis (Methods).

d. Two spots, for the 2' and 3'-isomers, are found.

Figure 5. Identification of modified nucleotides by thin layer chromatography.

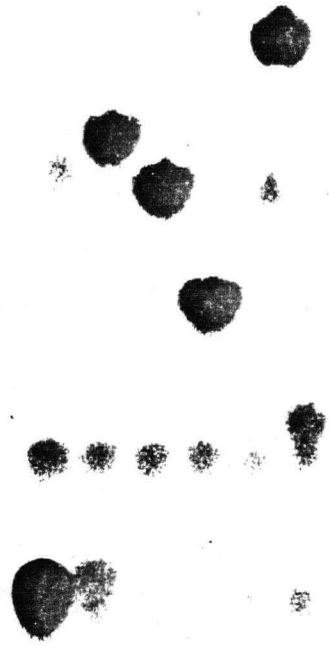
Samples shown in a, c, and e were chromatographed in solvent A, while samples shown in b, d, and f were chromatographed in solvent B, as described in Methods. Identification was based on mobility relative to either pA (solvent A) or pU (solvent B), and the assignments here are based on the mobility values presented in Table 2.

a

A m^1A Ψ T m^3C U_mU

77 67 64 63 50 44

b



A m^1A Ψ T m^3C U_mU

c

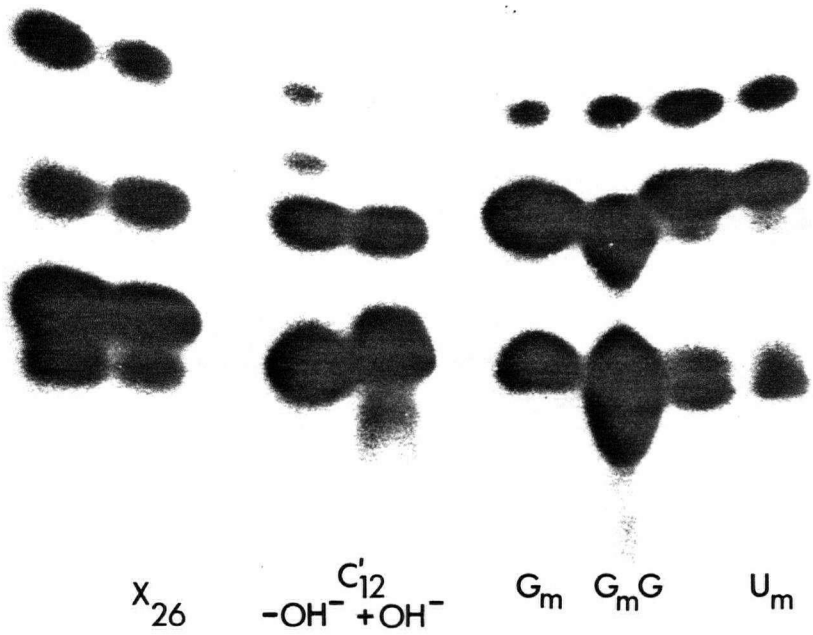
i^6A I m^3C D D G_mG C X

37 34 32 20 19 17 16 12

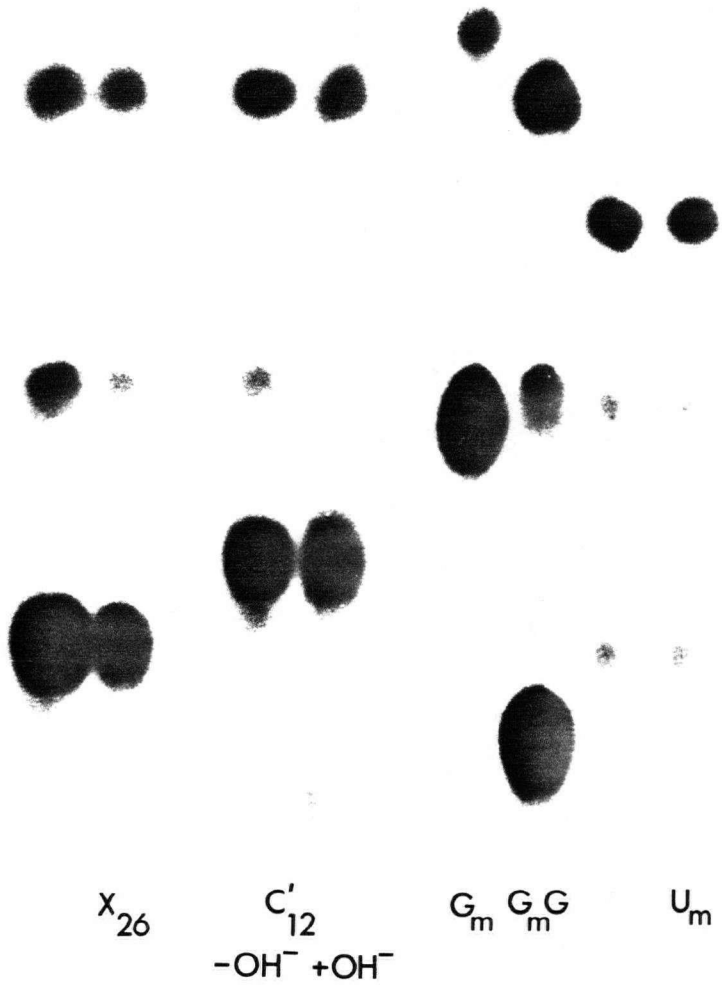
d

i^6A I m^3C D D G_mG C X

e



f



phosphate was identified by comigration with its non-radioactive standard on two-dimensional TLC (Fig. 6). Nucleotide 26 (pX) could not be identified on this basis (Fig. 6). The chromatographic mobilities of these *pN's are presented in Table 2. The positions of these modified nucleotides in the cloverleaf structure are the same as those in rat liver $tRNA_1^{Ser}$ (Fig. 2).

Identification of m^5C in position 57 was based on several lines of evidence. It was considered to be a modified nucleotide, probably a modified C, due to its mobility as a pNp on PEI-cellulose. On cellulose thin layer chromatography in solvent A (Methods), $p(N_{57})p$ ran slightly ahead of pCp. From two-dimensional homochromatography, it is clear that nucleotide 57 contains C or a derivative with a similar pK (Fig. 7). The presence of m^5C was indicated by nucleotide and nucleoside analysis (61). On ladder gel analysis of [5'-^{32}P] 3'-half molecules of yeast $tRNA_I^{Gly}$, a doublet in the RNase A slot was seen at a site where m^5C should be (Fig. 8). Equivalence of the doublet band to m^5C is not certain, since an extra nucleotide was found in this region by ladder gel analysis which was not in the published sequence (94). However, the sequence shown in Fig. 8 (and thus the correlation of the doublet band pattern with m^5C) seems most likely to be the correct one. The same doublet pattern is seen for nucleotide 57 in $tRNA_7^{Ser}$ (not visible in the photo Fig. 9). The nucleotide in the equivalent site in other eukaryotic serine tRNAs sequenced is m^5C. Nucleotide 57 was identified as m^5C on the basis of these various data.

Two modified nucleotides present in $tRNA_7^{Ser}$ were not identified. Nucleotide 12 is probably a derivative of cytidine, pC'. The *pN from

Table 2

Chromatographic Mobilities of Nucleoside-5'-Phosphates on Cellulose Thin Layer Plates

Nucleotide	R_{pA} [a]	R_{pU} [b]	R_{pA} (published) [c]	R_{pU} (published) [c]
pG	0.48	0.65	0.50	0.63
pA	1.00	0.37	1.00	0.34
pC	0.73	1.00	0.83	1.00
pU	0.50	1.00	0.57	1.00
pI	0.53	0.77	--	--
pT	0.61	0.82	0.76	0.81
pψ	0.41	1.01	0.46	1.01
pD	0.53	1.08	0.53	1.07
pm^3C	0.85	1.20	--	--
pm'A	0.89	1.10	0.92	1.07
pi^6A	1.5	0.12	--	--
pU_m	0.87	0.86	0.86	0.93
pG_m	0.84	0.63	0.86	0.58
pX_{26}	0.65	0.36	--	--

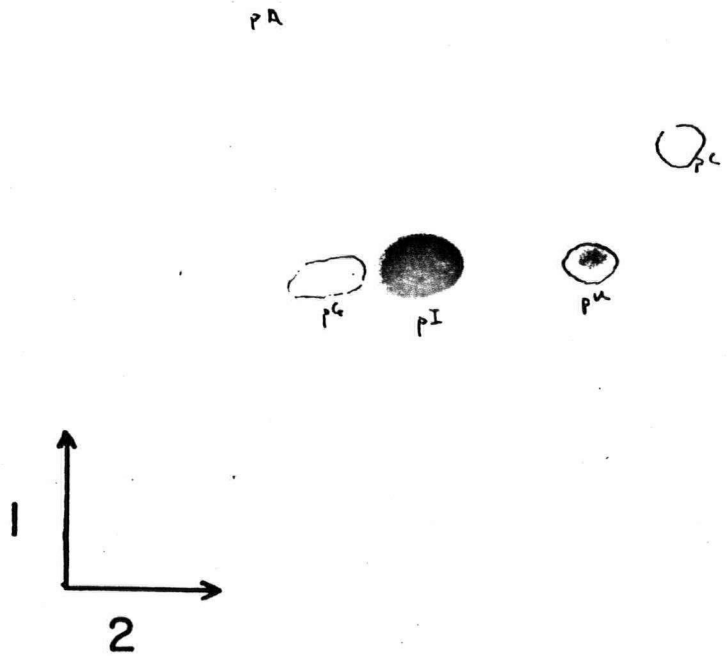
a - Mobility relative to pA, in solvent A (Methods).

b - Mobility relative to pU, in solvent B (Methods).

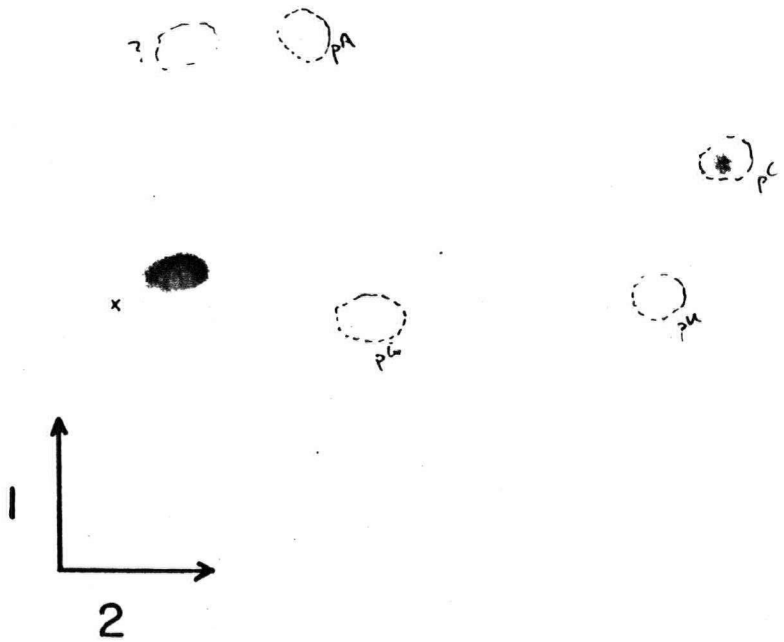
c - Values from Table 3 of reference 70.

Figure 6. Two dimensional thin layer chromatography of *pN's. Chromatography was carried out as described in Methods. A. The main spot co-chromatographed with inosine-5'-phosphate. B. The nucleotide $^{*}pX_{26}$ could not be identified based on its chromatographic mobility: pm^2_2G should run at about the same position as pA.

A.



B.



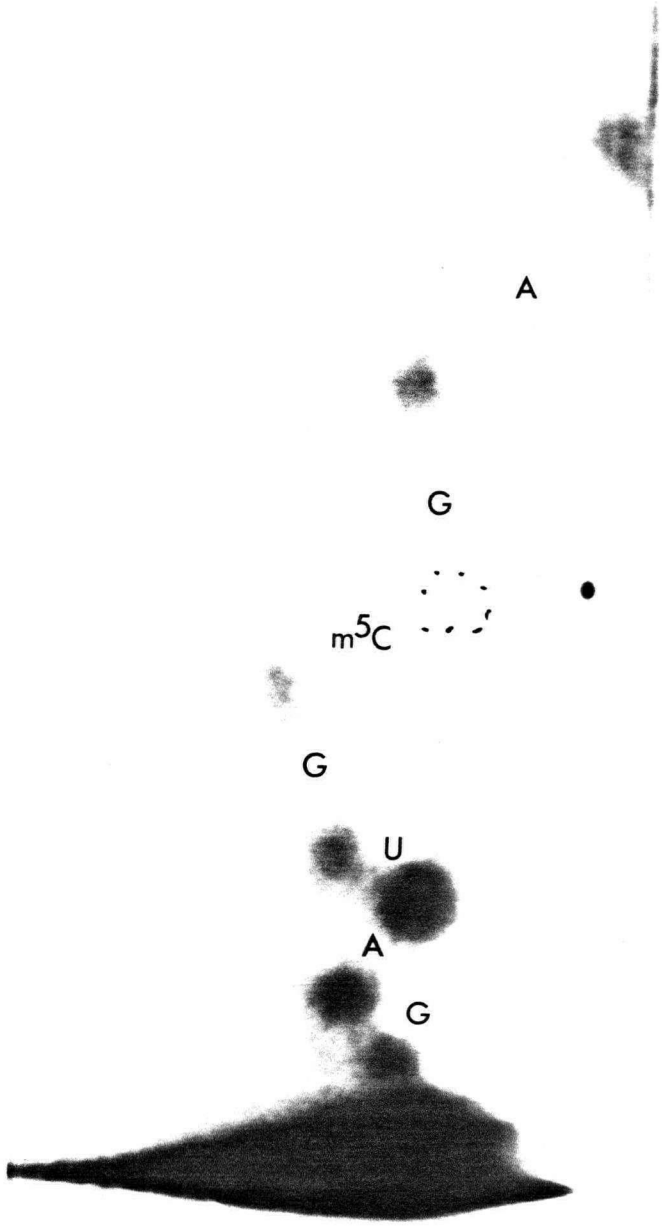


Figure 7. Two-dimensional homochromatography of a [5'-^{32}P]RNA fragment of $tRNA_7^{Ser}$.

Figure 8. Ladder gel analysis of a [5'-^{32}P]RNA fragment of yeast $tRNA_I^{Gly}$.

Analysis was carried out similar to the description in Methods. Each reaction contained 40,000 Cerenkov ^{32}P-cpm of yeast $tRNA_I^{Gly}$ fragment and 4 micrograms of carrier tRNA except for the ladder (L), which contained 80,000 cpm and 8 micrograms carrier tRNA. Reaction conditions were as described in Methods for the RNase T_1 (T) reaction, the ladder, and the control (-E). All others were as described in Methods except for varying enzyme concentrations: the RNase U_2 reactions U-1, U-2, and U-3 contained 0.02, 0.01, and 0.002 unit of enzyme activity; the RNase PhyI reactions Phy-1 and Phy-2 contained 0.08 and 0.013 unit of activity; and the RNase A reactions A-1 and A-2 contained 0.002 and 0.001 unit of activity, respectively. Electrophoresis was for 5 hr. at 1100 V. Autoradiography was for 2.5 days at -20°.

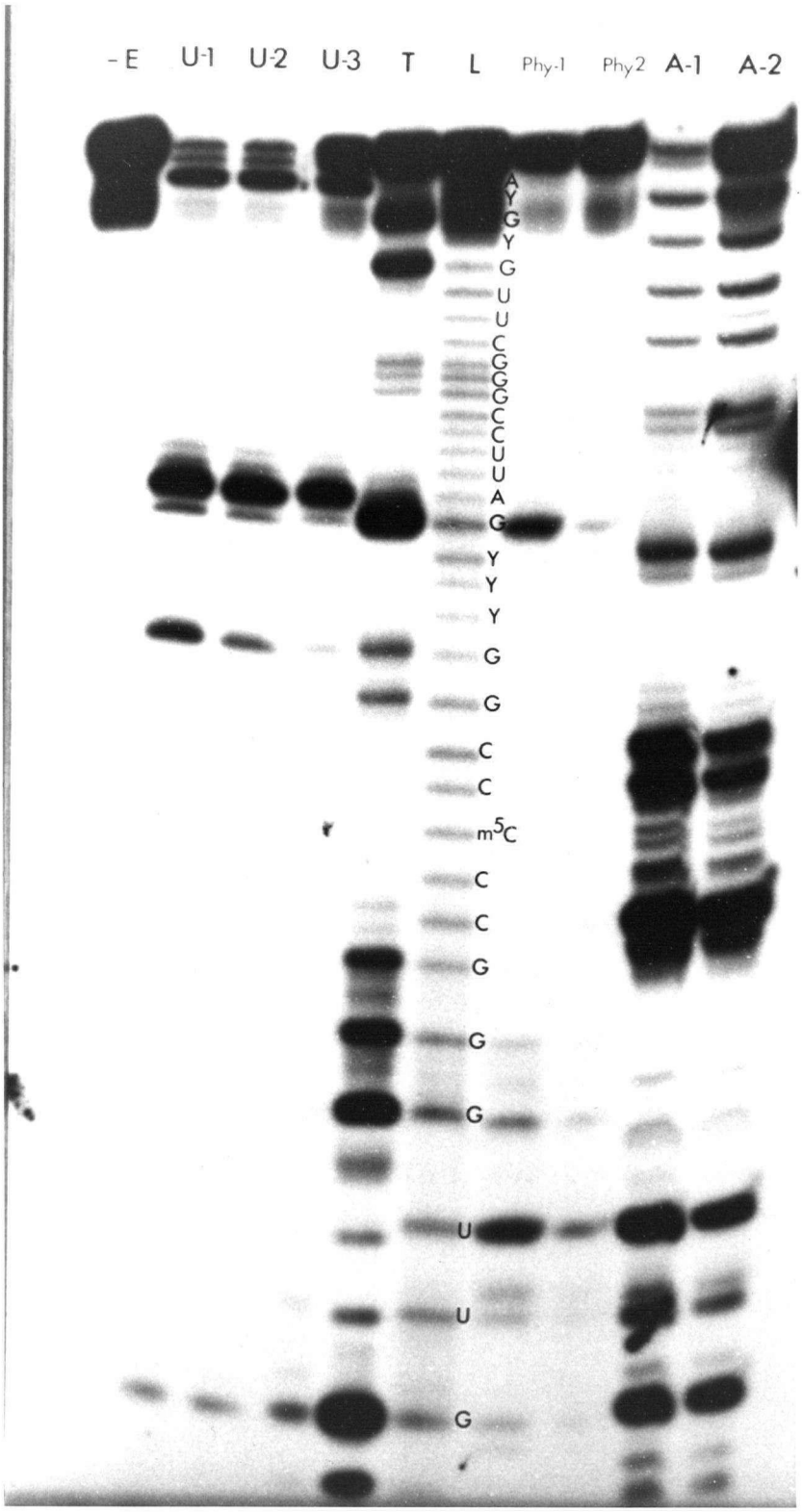
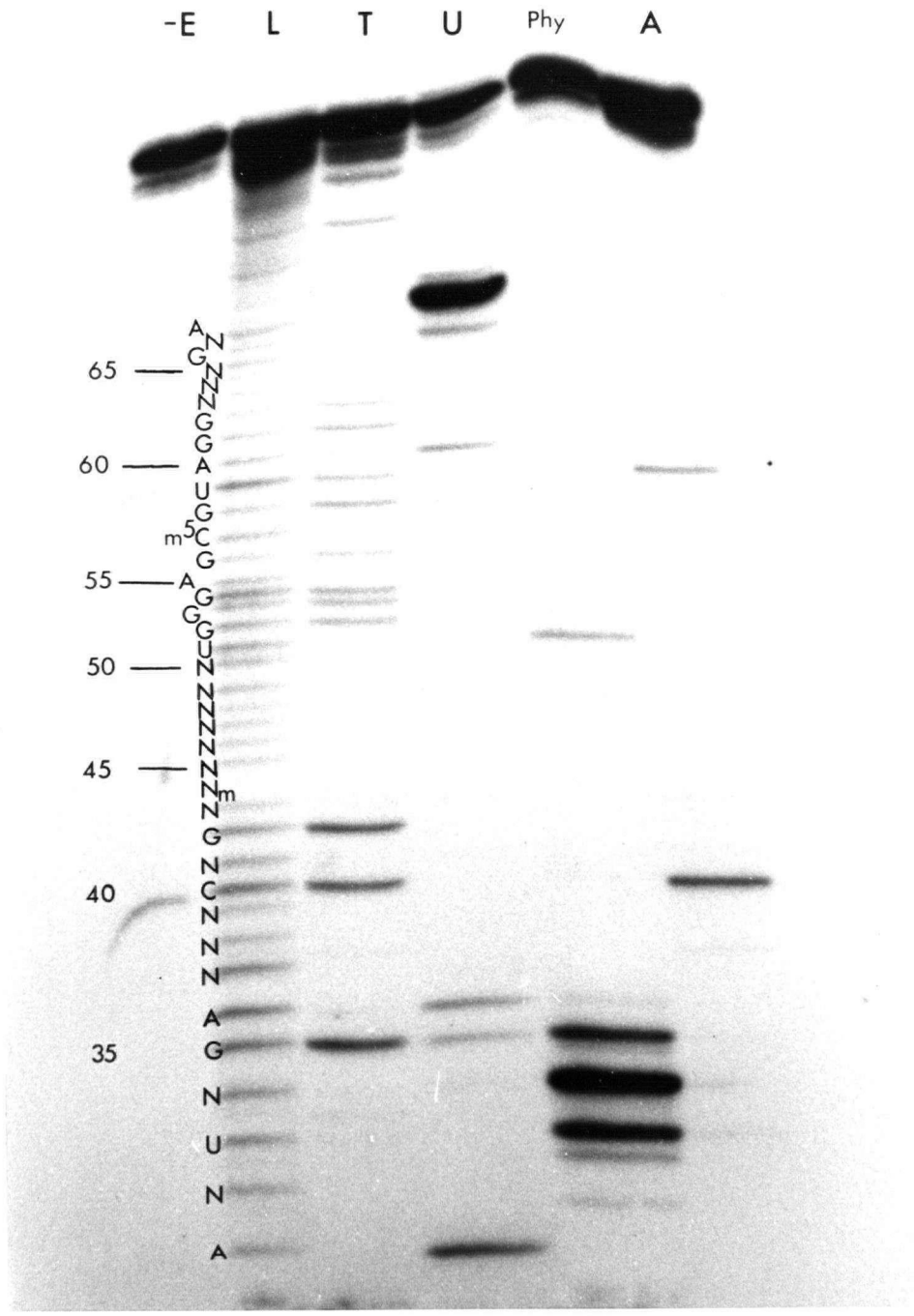

-E
U-1
U-2
U-3
T
L
Phy-1
Phy2
A-1
A-2
A
Y
G
Y
G
U
U
C
G
G
G
C
C
U
U
A
G
Y
Y
Y
G
G
C
C
m5C
C
C
G
G
G
U
U
G

Figure 9. Ladder gel analysis of [5' - ^{32}P] $tRNA_7^{Ser}$, 3'-end.

Analysis was carried out as described in Methods. Electrophoresis was for 24 hr. at 1200 V, followed by exposure to X-ray film for 3 days.


-E
L
T
U
Phy
A
65
60
55
50
45
40
35
A
G
N
N
N
G
G
A
U
G
m5C
G
A
G
G
G
U
N
N
N
N
N
N
N
m
N
G
N
C
N
N
N
A
G
N
U
N
A

position 12 generated by nuclease P_1 digestion of a formamide degradation-derived [5'-^{32}P] oligomer is distinct from pC on cellulose TLC, while the *pNp from alkaline hydrolysis of the same [5'-^{32}P]RNA fragment could not be distinguished from pCp on PEI-cellulose chromatography, indicating that the modificiation is alkali-labile. Nucleotide 12 runs differently on two-dimensional TLC from any nucleotide presented in ref. 70, Fig. 2. The nucleotide found at position 12 in other sequenced eukaryotic serine tRNAs (which are very similar or identical to $tRNA_7^{Ser}$ from U_8 to C_{25}, as in Fig. 11) is N^4-acetylcytidine, and the properties of this modified nucleotide are consistent with the data obtained. Since the *Drosophila* tRNA sequence is very similar to that of $tRNA_1^{Ser}$ from rat liver (Fig. 2, 11), the cloverleaf structure drawn from the *Drosophila* tRNA was that proposed for the rat liver species (95). On this basis, nucleotide 12 is in the D loop stem, and is opposite a G. For correct Watson-Crick base-pairing, C'_{12} should be modified at N-4, C-5, or C-6. The alkali-lability of the modification is consistent with acylation at N-4. No ac^4C was detected in nucleoside or nucleotide analysis (61), but acyl derivatives of cytidine are labile in both acid and alkali (96), and the conditions of analysis may have been sufficiently severe to cause deamination or hydrolysis of the amide linkage (90,96). The data for nucleotide 12 are consistent with ac^4C, and while comparison of $^{*}pC'_{12}$ with or without NaOH treatment (Methods) did not allow unambiguous identification (Fig. 5), pN_{12} is shown as ac^4C in Fig. 12.

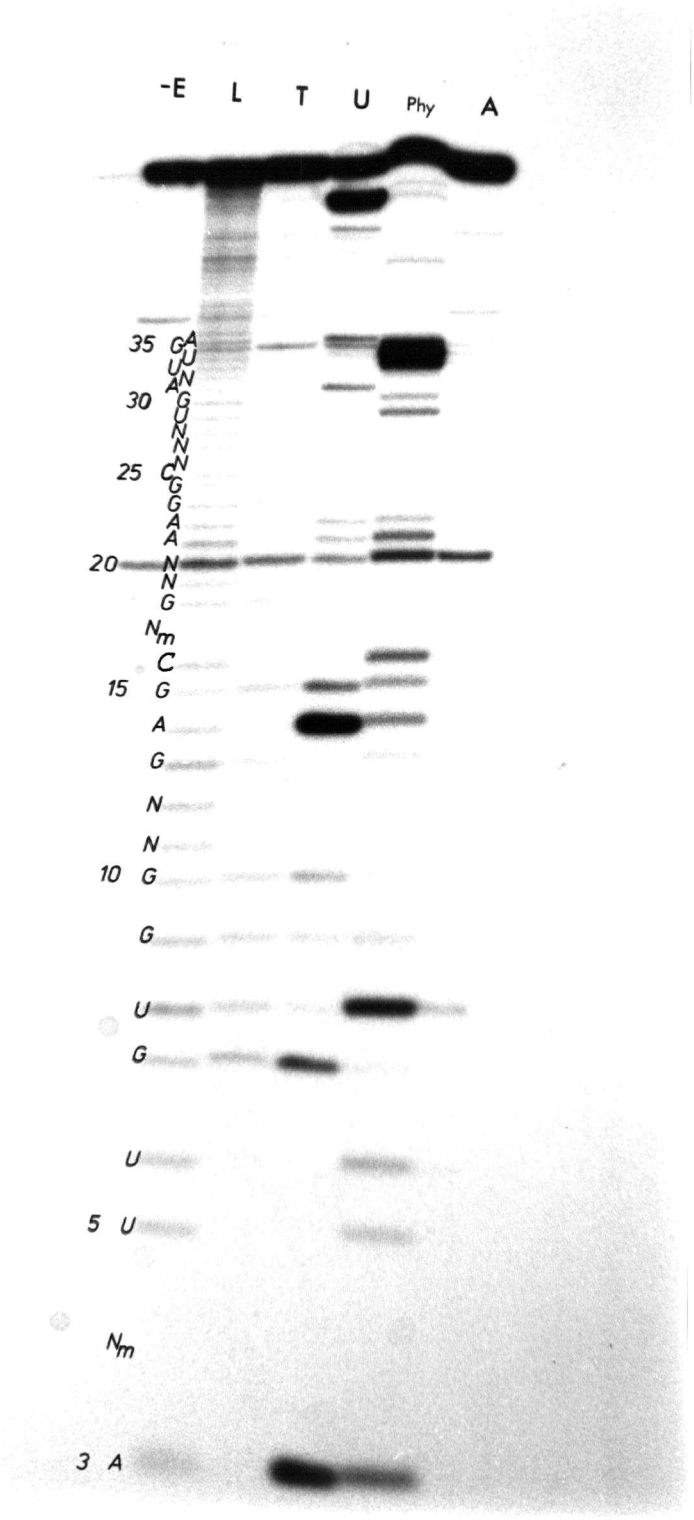
Nucleoside and nucleotide analysis of $tRNA_7^{Ser}$ indicated the presence of N^2, N^2-dimethylguanosine (61). This nucleoside is found only

in the position between the D loop and anticodon stems in other tRNAs (17), equivalent to position 26 of $tRNA_7^{Ser}$. However, on two-dimensional chromatography $^*pX_{26}$ ran much differently than pm_2^2G (Fig. 6; compare with Fig. 2 of ref. 70). Chromatography of $^*pX_{26}$ adjacent to a nuclease P_1 digest of yeast $tRNA_1^{Met}$, which contains pm_2^2G, did not allow unambiguous identification. Though there was a nucleotide in the $tRNA_1^{Met}$ digest which ran with $^*pX_{26}$, the yeast tRNA contains unidentified derivatives of A and G (17), which could have the same R_f as pm_2^2G. No band can be seen for this nucleotide in a ladder gel (Fig. 10), but RNase T_1 cuts at m_2^2G residues slowly (97), and no conclusions can be drawn from this negative result. The possibility that X_{26} is in fact m_2^2G cannot be ruled out based on the data here.

Further sequence data were obtained by partial enzymatic digestion of [5'-^{32}P] tRNA and analysis of the digestion products as described (Methods). The results confirm and extend those obtained by the formamide degradation method. In this way the identities of U_5 and G_{18} on the 3'-side of ribose-methylated nucleotides were established. N_{45}, immediately 3' to $(U_m)_{44}$, could not be identified in this way, but by base-pairing constraints within the cloverleaf structure it should be U (Fig. 12). The single ladder gel covering the 5'-terminus of $tRNA_7^{Ser}$ gave the sequence: 5'—$A_3N_mUUGUGG_{10}$—3'. This sequence is complementary to that obtained by the formamide degradation method for the aminoacyl stem at the 3'-terminus, though it introduces one non-Watson-Crick base-pair (U_6 - G_{76}). N_2, by base-pairing, is probably C since it is opposite G. Similarly, $(N_m)_4$ is opposite C and should be G_m.

Figure 10. Ladder gel analysis of [5'-^{32}P]tRNA$_7^{Ser}$, 5'-end.

Analysis was carried out as described in Methods. Electrophoresis was for 5 hr. at 1100V, and exposure of the gel to X-ray film for 2 days.


-E
L
T
U
Phy
A
35
30
25
20
15
10
5
3

```
            ac4                 m3   i6          m3     m5        m1
pGCAGUUGUGGCCGAGCGGDDAAGGCXUCUGACUIGAAAψCAGAUUCCCUCUGGGAGCGUAGGTψCGAAUCCUACCGACUGCNCCA      D.m. tRNA7Ser
   m            m                            m

        s4          D            o5 ms2i6      C   A
pGGAAGUGUGGCCGAGCGGDGAAGGCACCGGUCUUGAAAACCGGCGACCGAAGGGUUCCAGAGTψCGAAUCUCUGCGCUUCCGCCA      E. coli tRNAISer
                m              m

            ac4          m2          i6                  m5
pGGCAACUUGGCCGAGDGGDDAAGGCGAAAGAψUIGAAAψCUUUUGGGCUCUGCCCGCGCAGGTψCAAAUCCUGCAGUUGUCGCCA      Yeast tRNA2Ser
                m                                m

            ac4          m2   m3   i6          m3     m5        m1
pGUAGUCGUGGCCGAGDGGDDAAGGCGAψGGACUIGAAAψCCAUUGGGGUCUCCCCGCGCAGGTψCGAAUCCUGCCGACUACGCCA      Rat liver tRNA1Ser
                                                 m

            ac4          m2   m3   mt6         m3     m5        m1
pGACGAGGUGGCCGAGDGGDDAAGGCGAψGGACUGCUAAψCCAUUGψGCUCUGCACGCGUGGGTψCGAAUCCCAUCCUCGUCGCCA      Rat liver tRNA3Ser
                m                a:c:      m
```

Common sequence between D.m. $tRNA_7^{Ser}$ and other serine tRNAs

	Stems	Loops	Total
D.m. vs. *E. coli* I	25/48	26/40	51/88
vs. Yeast II	24/48	31/37	55/85
vs. Rat liver I	30/48	34/37	64/85
vs. Rat liver 3	25/48	31/37	56/85

Figure 11. Comparison of the sequence of D.m. $tRNA_7^{Ser}$ with other serine tRNAs

D.m. $tRNA_7^{Ser}$

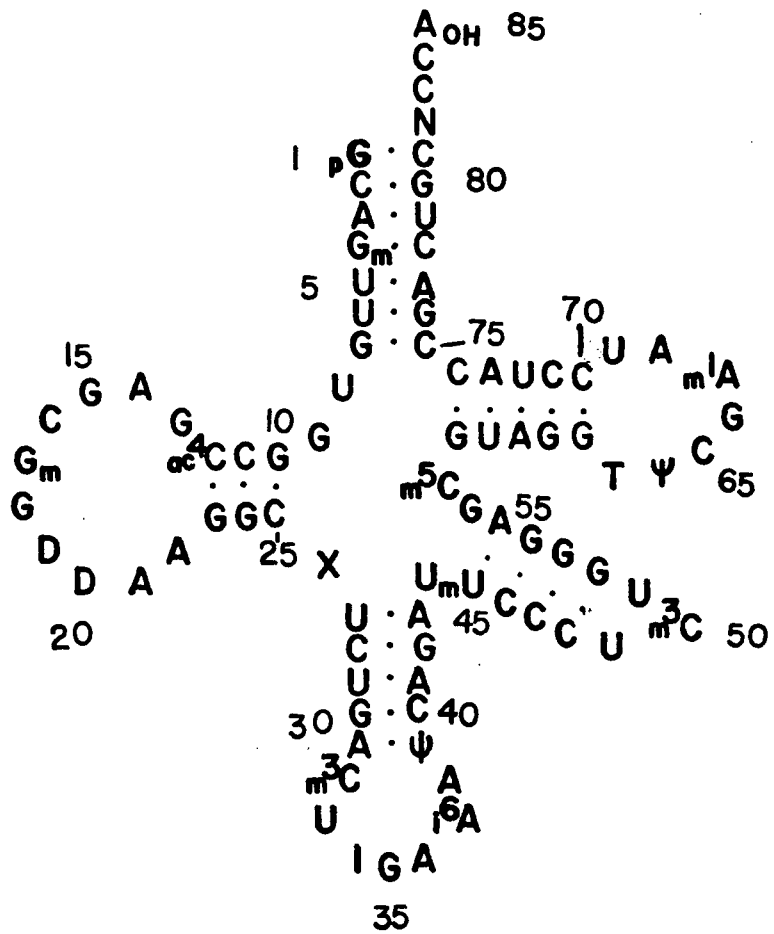


FIGURE 12

The sequence analysis indicates that there are three ribose-methylated nucleotides in $tRNA_7^{Ser}$. Digestion with RNase T_2 should yield the three dinucleotides G_mpUp, G_mpGp and U_mpUp. Consistent with this, the nucleotide analysis of White *et al.* (61) indicates three RNase T_2 stable dinucleotides in the tRNA, two of which are fluorescent in acid and thus contain G or a derivative of it.

The presence of a doublet band corresponding to m^5C in ladder gel analysis of RNase A digests may be diagnostic for that nucleotide. The doublet is the result of enzymatic hydrolysis, since it is not seen in the reference ladder (Figures 8, 9). Normally under conditions of partial enzymatic digestion, fragments with cyclic phosphates at their 3'-termini are generated. At higher ratios of enzyme to substrate, or on prolonged digestion, ring-opening of the cyclic phosphate occurs (72, 78). It appears that the doublet seen in ladder gels reflects particularly efficient ring-opening of the cyclic phosphate on 3'-terminal m^5C by RNase A.

There may be some change in the specificity of RNase T_1 under the conditions for partial hydrolysis used (Methods; 83). RNase T_1 normally cleaves on the 3'-side of inosine residues (91), yet for $tRNA_7^{Ser}$ there was no band in the RNase T_1 channel at position 34 despite that residue's accessibility shown by a strong band in the RNase Phy I channel. This result is contrary to previous reports indicating that the specificity of RNase T_1 is not altered under the conditions used (77,92).

Certain regions of the tRNA molecule are resistant to enzymatic hydrolysis. In Fig. 9, the sequence from N_{43}-N_{50} is highly resistant to cleavage. This is probably due to the tertiary structure of the RNA. Nucleotides in this region were poorly labeled by polynucleotide kinase

in several formamide degradation experiments, but the poor labeling is probably not due to lack of affinity of the enzyme for the terminal nucleotides here (70) and should reflect residual structure of the fragments. Enzymatic hydrolysis of the aminoacyl stem was also weak (Fig. 10), presumably due to double-stranded structure (81).

The clover leaf structure drawn for $tRNA_7^{Ser}$ of *Drosophila* was that constructed for rat liver $tRNA_1^{Ser}$ (based in turn on yeast $tRNA^{Ala}$) (95,62,17), with equivalent numbers of base pairs in each stem. As seen in Fig. 12, the data obtained are consistent with that structure, and no mismatched base-pairing is seen (U-G base pairs like U_6-G_{76} are often found in the aminoacyl stems of tRNAs). A RNase T_2 digest of $tRNA_7^{Ser}$ contains pGp and A (61), consistent with a 5'-terminal G and a 3'-terminal A. Assuming a 3'-terminal-CCA as in all other tRNAs sequenced so far (17), there is direct sequence information on 76 of 81 remaining nucleotides. Four of the remaining five nucleotides are placed within stems of the cloverleaf structure, and are subject to base-pairing constraints. Only for N_{82} is there no information whatever. Identification of X_{26} may be aided by determination of the gene DNA sequence. Further information on the modified nucleotide at position 12 would be desirable. However, most of the sequence information has been verified in separate experiments, and the sequence shown in Fig. 12 should be at least 95% accurate.

Chapter 4

DISCUSSION

There are extensive sequence homologies among the eukaryotic serine transfer RNAs (Fig. 11). This is particularly true for loop regions: drawn in the conventional cloverleaf structure, the two rat liver serine tRNAs are nearly identical for those nucleotides not involved in base pairs. Likewise the nucleotide sequences in the loop regions of yeast $tRNA^{Ser}_{1,2}$ are almost identical to those of the rat serine tRNAs (17). Since the two *Drosophila* tRNAs sequenced ($tRNA^{Lys}_{2}$ and $tRNA^{Met}_{3}$) are very similar to their mammalian counterparts (85,98), the relationship between *Drosophila* $tRNA^{Ser}_{7}$ and rat liver $tRNA^{Ser}_{1}$, which has the same anticodon, was examined. With the exceptions of C_{16}, X_{26}, which has not been shown conclusively to be distinct from m^2_2G, and possibly N_{82}, for which there is no information yet, the *Drosophila* and rat liver tRNAs are identical in non-base paired regions. This high degree of homology is not maintained in stem sequences. The D, Tψ, and aminoacyl stems are very similar, with no more than one base pair varying. Their extra arms are similar in base content (9 of 11 positions) but are inverted relative to each other so that only 3 of 11 nucleotides match. Considering base content as well as primary sequence, only the anticodon stems vary much between these tRNAs (Fig. 11).

The similarities of base sequence of eukaryotic $tRNAs^{Ser}$ are apparently independent of their anticondons. Consistent with this, work in progress indicates roughly 90% sequence homology between *Drosophila* $tRNA_4^{Ser}$ and $tRNA_7^{Ser}$ (which respond to UCG and UCU codons, respectively; ref. 61).

Among the various families of eukaryotic tRNAs recognizing codons beginning with U (Cys, Leu, Phe, Ser, Trp, and Tyr), those tRNAs that have been sequenced show certain similarities. The D loops generally contain AG_D^C—GG— — —A; the anticodon loops are usually $\overset{\psi}{C}_mU$— —AHA, where H is a hypermodified purine nucleotide; and the Tψ loops are usually either TψCGm'AAU or TψCGm'AUC, though there may be variation in as many as three of the seven positions (17). The stems vary more than the loops in these tRNA families, but there are certain nearly invariant residues. In the D stem, $(m^2)G_{10}C_{11}Py_{12}/Pu_{23}G_{24}C_{25}$ is found in most cases (about equal proportions of G and m^2G in position 10); G_{30}-C_{40} and A_{31}-ψ_{39} base pairs in the anticodon stem have been found in 19 of 20 tRNAs examined (m^5C and ψ_m sometimes replace C and ψ); and G_{62}-C_{70} is invariant. These base pairs in the anticodon and Tψ stems are immediately adjacent to the loop structures as drawn in the standard cloverleaf (Fig. 12). These patterns break down outside tRNA families recognizing UNN codons (17).

The yeast tRNAs whose genes have been shown to contain intervening sequences all respond to UNN codons (99,100). As yet, no recognition feature for the enzyme(s) responsible for excision of the intervening sequence, which is included in the precursor tRNA, is apparent, but sequencing of genes from other organisms for the equivalent tRNAs may prove enlightening and is now in progress with recombinant plasmids

containing *Drosophila* $tRNA^{Ser}_{4,7}$ genes. No tRNA genes from any organism except yeast have been found to contain intervening sequences. The gene for yeast $tRNA^{Ser}_{UCG}$ (equivalent to *Drosophila* $tRNA^{Ser}_{4}$) contains an intervening sequence, while that for $tRNA^{Ser}_{UC(A,C,U)}$ (equivalent to $tRNA^{Ser}_{7}$) does not (unpublished results of G. Page). It will be of interest to compare the structures of the equivalent serine tRNA genes from *Drosophila* and yeast.

The sequence of *Drosophila* $tRNA^{Ser}_{7}$ is generally unremarkable; it is similar to other eukaryotic serine tRNAs analyzed. It contained m^3C and likely ac^4C, which, though not found in most tRNAs, are common in the eukaryotic serine tRNAs studied (17). Sixteen of 85 nucleotides are modified, similar to other eukaryotic serine tRNAs (Fig. 11). One unusual feature of $tRNA^{Ser}_{7}$ is the ribose-methylation at position 4 (Fig. 12). The only other known cases of ribose-methylation in the aminoacyl stem are glycine tRNAs of yeast, wheat germ, and *Bombyx mori*, also modified at nucleotide 4 (17).

The codon recognition properties of $tRNA^{Ser}_{7}$ are unusual. The tRNA responds significantly in the ribosome binding assay only to UCU (61), but the 5' nucleotide of the anticodon is inosine. The presence of inosine in the wobble position normally allows a tRNA to translate codons ending in A, C, or U, as is the case for yeast alanine and serine and rat liver serine tRNAs (see Fig. 8, ref. 101). The anticodon loop of $tRNA^{Ser}_{7}$ is identical to that of rat liver $tRNA^{Ser}_{1}$, and differs by only one nucleotide from yeast serine tRNAs 1 and 2 (17). The first two base pairs from the anticodon loop of each stem are also the same, except for the ribose-methylated ψ at position 39 in rat liver $tRNA^{Ser}_{1}$, which does

not affect that tRNA's codon response compared with the equivalent yeast tRNAs lacking that methylation. Thus, it is likely that a structural feature elsewhere is responsible for the codon response of $tRNA_7^{Ser}$. A candidate would certainly be the ribose-methylated nucleotide in position 4, since this is an unprecedented modification for a serine tRNA, but no firm conclusions can be drawn here. The codon response of $tRNA_7^{Ser}$ is not likely to be typical of *Drosophila* tRNAs with inosine in the anticodon. Transfer RNA_4^{Val} contains inosine and responds to codons ending in A, C, and U (102), indicating that inosine is in the first position of the anticodon and consistent with the "wobble" hypothesis (103).

The modified nucleotides found in sequence analysis of *Drosophila* $tRNA_7^{Ser}$ agree with analyses of White *et al.* (61), except for m_2^2G and ac^4C. White *et al.* found m_2^2G in two chromatographic systems, but in the present study the identificationof X_{26} as m_2^2G has not been confirmed. C'_{12} is likely to be ac^4C. Though this nucleotide was not found by White *et al.*, this is undoubtedly due to its instability in both acid and alkali (90,96). Further experiments are required for identification of these two nucleotides.

One purpose of the sequence analysis presented here was to allow identification of the tRNA gene within a sequence of recombinant *Drosophila* DNA; there is sufficient information here for that. Also, the RNA sequence shows that the corresponding gene for $tRNA_7^{Ser}$ should contain recognition sequences for the restriction endonucleases Taq I (TCGA, positions 64-67) and Hae III (GGCC, positions 9-12). This should make sequencing of the tRNA gene by the Gilbert-Maxam method (82) relatively easy, since it

will be possible to sequence in both directions from known starting points within a $tRNA_7^{Ser}$ gene.

The sequencing strategy employed here (Introduction) was reasonably successful. The formamide degradation method made possible extensive characterization of the modified nucleotides, and gave most of the sequence (72 of 85 nucleotides). Gaps in the sequence obtained by this method were filled by ladder gel analysis in most cases. Wandering spot analysis of gel-purified [5'-^{32}P]RNA fragments generated by formamide hydrolysis (Methods) was shown to be effective as a means of verifying regions of difficulty in the sequence (Fig. 7). Ladder gel analysis provided most but not all of the information desired from it. Nucleotide 45, assigned as U_{45} by the criterion of Watson-Crick base pairing, could not be read from the ladder gels (Fig. 9). Thus there is no direct evidence on the identity of this nucleotide, since it is 3' to a ribose-methylated nucleotide and is not accessible by the formamide degradation method. On the other hand, G_{18} was readily identified by ladder gel analysis (Fig. 10). The inability to identify U_{45} by ladder gel analysis may be a consequence of using intact tRNA as substrate: access to many phosphodiester bonds is apparently limited for the nucleases due to tertiary structure (Results). As seen in Fig. 8, this is less a problem with tRNA half-molecules. Thus, generation of half-molecules by limited enzymatic digestion (70,104) may be desirable for these analyses. A limiting factor in ladder gel analysis so far has been obtaining sufficient radiolabeled RNA for substrate, since the recessed 5'-termini of tRNAs make them relatively inaccessible to labeling by polynucleotide kinase (105). Since this should be less of a problem with half-molecules, their use may be advantageous. Another alternative may be to label at the 3'-terminus

using RNA ligase (70,106). Once [^{32}P]RNA is available, ladder gel analysis using either chemical or enzymatic digestions would be compatible with the general strategy employed here. However, the conditions of chemical hydrolysis appear to give a more uniform array of product bands and that method may well be preferable (compare Fig. 1, ref. 79, with Fig. 2, ref. 81). A possible modification of the formamide degradation method is direct transfer of [^{32}P]RNA fragments from gels to anion exchange thin layer plates, followed by enzymatic hydrolysis *in situ* to mononucleotides and chromatography (107). The subjectivity involved in excision of gel bands would then be removed. Placement of poorly labeled modified nucleotides would be easier, since they could be identified, then located precisely within even a high background of contaminating fragments (double "hit" products) by comparison to ladder gels covering the same sequence. A disadvantage of doing such transfers would be that two-dimensional homochromatography could not be performed. Provided the initial transfer from gel to plate maintains resolution and is reasonably efficient such transfer experiments may be sufficient, but the flexibility provided by two-dimensional nomochromatography as described here (Methods) can be quite useful since any region of the tRNA except its 5'-terminus is accessible to analysis in this way.

While some modifications such as those mentioned above may be desired, the general sequencing strategy employed here was reasonably successful. On refinement, it should be at least as fast, accurate, and generally applicable as any other strategy now used for tRNA sequencing.

BIBLIOGRAPHY

1. R.P. Perry (1976). Annu. Rev. Biochem. 45, 605-630.

2. S.M. Tilghman, D.C. Tiemeier, J.G. Seidman, B.M. Peterlin, M. Sullivan, J.V. Maizel, and P. Leder (1978). Proc. Nat. Acad. Sci. USA 75, 725-727.

3. S. Tonegawa, C. Brack, N. Hozumi, and R. Schuller (1977). Proc. Nat. Acad. Sci. USA 74, 3518-3522.

4. S.M. Berget, C. Moore, and P.A. Sharp (1977). Proc. Nat. Acad. Sci. USA 74, 3171-3175.

5. L.T. Chow, R.E. Gelines, T.R. Broker, and R.J. Roberts (1977). Cell 12, 1-8.

6. P.K. Wellauer and I.B. Dawid (1977). Cell 10, 193-212.

7. R.L. White and D.S. Hogness (1977). Cell 10, 177-192.

8. R. Breathnach, J.L. Mandel, and P. Chambon (1977). Nature 270, 314-319.

9. G. Knapp, J.S. Beckmann, P.F. Johnson, S.A. Fuhrman, and J. Abelson (1978). Cell 14, 221-236.

10. P.Z. O'Farrell, B. Cordell, P. Valenzuela, W.J. Rutter, and H.M. Goodman (1978). Nature 274, 438-445.

11. E.M. DeRobertis and M.V. Olson (1979). Nature 278, 137-143.

12. F.E. Baralle and G.G. Brownlee (1978). Nature 274, 84-87.

13. O. Hagenbüchle, M. Santer, J.A. Steitz, and R.J. Mans (1978). Cell 13, 551-563.

14. M. Grunstein and D.S. Hogness (1975). Proc. Nat. Acad. Sci. USA 72, 3961-3965.

15. T. Maniatis, R.C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G.K. Sim, and A. Efstratiadis (1978). Cell 15, 687-701.

16. R. Hall, "The Modified Nucleosides in Nucleic Acids," Columbia University Press, New York and London, 1971.

17. M. Sprinzl, F. Grüter, and D. H. Gauss (1979). Nucleic Acids Res. 6, r1-r9.

18. J.E. Heckman, B. Alzner-Deweerd, and U.L. Raj Bhandary (1979). Proc. Nat. Acad. Sci. USA 76, 717-721.

19. M. Brenner, J.A. Lewis, D.S. Straus, F. DeLorenzo, and B.N. Ames (1972). J. Biol. Chem. 247, 4333-4339.

20. H.G. Khorana, K.L. Agarwal, P. Besmer, H. Büchi, M.H. Caruthers, P.J. Cashion, M. Fridkin, E. Jay, K. Kleppe, R. Kleppe, A. Kumar, P.C. Loewen, R.C. Miller, K. Minamoto, A. Panet, U.L. RajBhandary, B. Ramamoorthy, T. Sekiya, T. Takeya, and J.H. van de Sande (1976). J. Biol. Chem. 251, 565-570.

21. S. Altman and J.D. Smith (1971). Nature New Biol. 233, 35-39.

22. G. Zurawski, D. Elseviers, G.V. Stauffer, and C. Yanofsky (1978). Proc. Natl. Acad. Sci. USA 75, 5988-5992.

23. F. Harada, R.C. Sawyer, and J.E. Dahlberg (1975). J. Biol. Chem. 250, 3487-3497.

24. S. Nishimura, Progress Nucl. Acids Res. Mol. Biol. 12, 49-85.

25. T. Ikemura and H. Ozeki (1977). J. Mol. Biol. 117, 419-446.

26. J.H. Wilson and J.N. Abelson (1972). J. Mol. Biol. 69, 57-73.

27. S.P. Eisenberg, L. Soll, and M. Yarus (1979). J. Biol. Chem. 254, 5562-5566.

28. S. Benzer and S.P. Champe (1962). Proc. Natl. Acad. Sci. USA 48, 1114-1121.

29. S. Brenner and J.R. Beckwith (1965). J. Mol. Biol. 13, 629-637.

30. H.M. Goodman, J. Abelson, A. Landy, S. Brenner, and J.D. Smith (1968). Nature 217, 1019-1024.

31. L. Soll and P. Berg (1969). Proc. Natl. Acad. Sci. USA 63, 392-399.

32. L. Soll and P. Berg (1969). Nature 223, 1340-1342.

33. C. Singer, G. Smith, R. Cortese, and B.N. Ames (1972). Nature New Biol. 238, 72-74.

34. M. Brenner and B.N. Ames (1972). J. Biol. Chem. 247, 1080-1088.

35. J.D. Smith (1976). Prog. Nucleic Acid Res. Mol. Biol. 16, 25-73.

36. E. Lund and J.E. Dahlberg (1977). Cell 11, 247-262.

37. E. Schweizer, C. MacKechnie, and H.O. Halvorson (1967). J. Mol. Biol. 40, 261.

38. M.V. Olson, D.L. Montgomery, A.K. Hopper, G.S. Page, F. Horodyski, and B.D. Hall (1977). Nature 267, 639-641.

39. F. Sherman, S.W. Liebman, J.W. Stewart, and M. Jackson (1973). J. Mol. Biol. 78, 157-168.

40. M.L. Birnstiel, B.H. Sells, and I.F. Purdom (1972). J. Mol. Biol. 63, 21-39.

41. S.G. Clarkson, M.L. Birnstiel, and I.F. Purdom (1973). J. Mol. Biol. 79, 411-429.

42. S.G. Clarkson, M.L. Birnstiel, and V. Serra (1973). J. Mol. Biol. 79, 391-410.

43. E.Lund, J.E. Dahlberg, L. Lindahl, S.R. Jaskunis, P.P. Dennis, and M. Nomura (1976). Cell 7, 165-177.

44. B.N. White, G.M. Tener, J. Holden, and D.T. Suzuki (1973). J. Mol. Biol. 74, 635-651.

45. F.M. Ritossa, R.C. Atwood, and S. Spiegelman (1966). Genetics 54, 663-676.

46. L. Weber and E. Berger (1976). Biochemistry 15, 5511-5519.

47. T.A. Grigliatti, B.N. White, G.M. Tener, T.C. Kaufman, and D.T. Suzuki (1974). Proc. Nat. Acad. Sci. USA 71, 3527-3531.

48. E. Kubli and T. Schmidt (1978). Nucleic Acids Res. 5, 1465-1478.

49. R. Dunn, S. Hayashi, I.C. Gillam, A.D. Delaney, G.M. Tener, T.A. Grigliatti, T.C. Kaufman, and D.T. Suzuki (1979). J. Mol. Biol. 128, 277-287.

50. S. Hayashi, I.C. Gillam, A.D. Delaney, R. Dunn, G.M. Tener, T.A. Grigliatti, and D.T. Suzuki (1979). Chomasoma, in press.

51. P.H. Yen, A. Sodja, M. Cohen, S.E. Conrad, M. Wu, and N. Davidson (1977). Cell 11, 763-777.

52. D. Söll (1979). Abstracts, XI International Congress of Biochemistry.

53. G.M. Tener, S. Hayashi, R. Dunn, A. Delaney, I.C. Gillam, T.A. Grigliatti, T.C. Kaufman, and D.T. Suzuki (1979). Cold Spring Harbor Symposium on Quantitative Biology, in press.

54. D.D. Brown, P.C. Wensink, and E. Jordan (1971). Proc. Nat. Acad. Sci. USA 68, 3175-3179.

55. S.G. Clarkson, M.L. Birnstiel, and V. Serra (1973). J. Mol. Biol. 79, 391-410.

56. S.G. Clarkson and V. Kurer (1976). Cell 8, 183-195.

57. J.D. Procunier and R.J. Dunn (1978). Cell 15, 1087-1093.

58. B.A. Hamkalo, O.L. Miller, and A.H. Bakken (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 915-919.

59. R. Dunn, A.D. Delaney, I.C. Gillam, S. Hayashi, G.M. Tener, T. Griliatti, V. Misra, M.G. Spurr, D.M. Taylor, and R.C. Miller, Jr. (1979). Gene, in press.

60. A. Rafalski, J. Kohli, P. Agris, and D. Söll (1979). Nucleic Acids Res. 6, 2683-2695.

61. B.N. White, R. Dunn, I. Gillam, G.M. Tener, D.J. Armstrong, F. Skoog, C.R. Frihart, and N.J. Leonard (1975). J. Biol. Chem. 250, 515-521.

62. R.W. Holley, J. Apgar, G.A. Everett, J.T. Madison, M. Marquisee, S.H. Merrill, J.R. Penswick, and A. Zamir (1965). Science 147, 1462.

63. J.T. Madison, G.A. Everett, and H. Kung (1966). Science 153, 531.

64. H.G. Zachau, D. Dütting, and H. Feldmann (1966). Z. Physiol. Chem. 347, 212-235.

65. U.L. RajBhandary, S.H. Chang. A. Stewart, R.D. Faulkner, R.M. Hoskinson, and H.G. Khorana (1967). Proc. Nat. Acad. Sci. USA 57, 571.

66. R.W. Holley (1968). Prog. Nucleic Acid Res. Mol. Biol. 8, 37-47.

67. F. Sanger, G.G. Brownlee, and B.G. Barrell (1965). J. Mol. Biol. 13, 373-398.

68. M. Székely and F. Sanger (1969). J. Mol. Biol. 43, 607-617.

69. F. Labrie and F. Sanger (1969). Biochem. J. 114, 29P.

70. M. Silberklang, A.M. Gillum, and U.L. RajBhandary (1979). Meth. in Enzymol. 54, Part G, 58-109.

71. B.G. Barrell, in "Procedures in Nucleic Acid Research" Vol. 2 (G.L. CAntoni and D.R. Davies, eds.) p. 751. Harper and Row, New York, 1971.

72. G.G. Brownlee, "Determination of Sequences in RNA." North-Holland, Amsterdam, 1972.

73. E. Jay, R. Bambara, R. Padmanabhan and R. Wu (1974). Nucleic Acids Research 1, 331-353.

74. B.E. Griffin (1971). FEBS Lett. 15, 165-168.

75. C.-P. D. Tu, E. Jay, C.P. Bahl, and R. Wu (1976). Anal. Biochem. 74, 73-93.

76. R.E. Lockard, B. Alzner-Deweerd, J.E. Heckman, J. MacGee, M.W. Tabor, and U.L. RajBhandary (1978). Nucleic Acids Res. 5, 37-55.

77. H. Donis-Keller, A.M. Maxam, and W. Gilbert (1977). Nucleic Acids Res. 4, 2527.

78. A. Simoncsits, G.G. Brownlee, R.S. Brown, J.R. Rubin, and H. Guilley (1977). Nature 269, 834.

79. D.A. Peattie (1979). Proc. Nat. Acad. Sci. USA 76, 1760-1764.

80. D. Pilly, A. Niemeyer, M. Schmidt, and J.P. Bargetzi (1978). J. Biol. Chem. 253, 437-445.

81. G. Winter and G.G. Brownlee (1978). Nucleic Acids Res. 5, 3129-3139.

82. A.M. Maxam and W. Gilbert (1977). Proc. Nat. Acad. Sci. USA 74, 560-564.

83. J. Stanley and S. Vassilenko (1978). Nature 274, 87-89.

84. R.A. Laskey and A.D. Mills (1977). FEBS Lett. 82, 314-316.

85. S. Silverman, J. Heckman, G.J. Cowling, A.D. Delaney, R.J. Dunn, I.C. Gillam, G.M. Tener, D. Söll, and U.L. RajBhandary (1979). Nucleic Acis Res. 6, 421-433.

86. G.G. Brownlee and E.M. Cartwright (1977). J. Mol. Biol. 114, 93-117.

87. E.M. Southern and A.R. Mitchell (1971). Biochem. J. 123, 613-617.

88. M. Fujimoto, A. Kuninaka, and H. Yoshino (1974). Agr. Biol. Chem. 38, 1555-1561.

89. Y. Yamada and H. Ishikura (1975). Biochim. Biophys. Acta 402, 285.

90. H. Feldmann, D. Dütting, and H.G. Zachau (1966). Z. Physiol. Chem. 347, 236-248.

91. F. Egami, K. Takahashi, and T. Uchida (1964). Prog. Nucleic Acid Res. Mol. Biol. 3, 59-101.

92. R.C. Warrington (1974). Biochim. Biophys. Acta 353, 63-68.

93. M. Silberklang, A. Prochiantz, A.-L. Haenni, and U.L. RajBhandary (1977). Eur. J. Biochem. 72, 465-478.

94. M. Yoshida (1973). Biochem. Biophys. Res. Comm. 50, 779-784.

95. T. Ginsberg, H. Rogg, and M. Staehelin (1971). Eur. J. Biochem. 21, 249-257.

96. D.M. Brown, A. Todd, and S. Varadarajan (1956). J. Chem. Soc. 2384.

97. G.W. Rushizky and H.A. Sober (1968). Prog. Nucleic Acids Res. Mol. Biol. 8, 171-207.

98. S. Silverman, I.C. Gillam, G.M. Tener, and D. Söll (1979). Nucleic Acids Res. 6, 435-442.

99. H.M. Goodman, M.V. Olson, and B.D. Hall (1977). Proc. Nat. Acad. Sci. USA 74, 5453-5457.

100. P. Valenzuela, A. Venegas, F. Weinberg, R. Bishop, and W.J. Rutter (1978). Proc. Nat. Acad. Sci. USA 75, 190-194.

101. C.T. Caskey, M. Beaudet, and M. Nirenberg (1968). J. Mol. Biol. 37, 99-118.

102. R. Dunn, W.R. Addison, I.C. Gillam, and G.M. Tener (1978). Can. J. Biochem. 56, 618-623.

103. F.H.C. Crick (1966). J. Mol. Biol. 19, 548-555.

104. J.R. Penswick and R.W. Holley (1965). Proc. Nat. Acad. Sci. USA 53, 543-546.

105. J.R. Lillehaug and K. Kleppe (1977). Nucleic Acids Res. 4, 373-379.

106. A.G. Bruce and O.C. Uhlenbeck (1978). Nucleic Acids Res. 5, 3665-3677.

107. K. Randerath and R.C. Gupta (1979). Fed. Proc. 38, 499 (abstr. 1420).

108. I.M. Glynn and J.B. Chappell (1964). Biochem. J. 90, 147-149.