STUDIES ON THE VALINE TRANSFER RNAS AND THEIR GENES IN *DROSOPHILA MELANOGASTER*

by

WILLIAM ROBERT ADDISON

B.Sc., University of Regina, 1976

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF BIOCHEMISTRY

FACULTY OF MEDICINE

UNIVERSITY OF BRITISH COLUMBIA

We accept this thesis as conforming

to the required standard.

THE UNIVERSITY OF BRITISH COLUMBIA

February, 1982

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Biochemistry

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date March 5, 1982

Abstract

The coding properties of the 3 major valine tRNA isoacceptors of Drosophila melanogaster, the nucleotide sequences of $tRNA_{3b}^{Val}$ and $tRNA_{4}^{Val}$ and the nucleotide sequences of genes for these two tRNAs have been determined. Valyl-$tRNA_{3a}^{Val}$ binds strongly to ribosomes in response to the trinucleotide GUA and to a lesser extent with GUU and GUG. Valyl-$tRNA_{3b}^{Val}$ binds strongly in the presence of GUG and very weakly with the other 3 triplets whereas valyl-$tRNA_{4}^{Val}$ binds strongly in the presence of GUU, GUC, and GUA and weakly with GUG.

The nucleotide sequences of $tRNA_{3b}^{Val}$ and $tRNA_{4}^{Val}$ were determined by a combination of techniques. For both tRNAs most of the sequence was determined by the method of Stanley and Vassilenko. The sequences at the 5' and 3'-ends of the molecules were determined by wandering-spot analysis. Regions of the molecules that could not be sequenced by these two techniques were determined by the gel read-off method. The use of tRNA modified with chloroacetaldehyde to overcome problems in sequencing RNA by the gel read-off method caused by secondary structure in the RNA is described. The nucleotide sequence of $tRNA_{4}^{Val}$ is: GUUU$_m$CCGUm^{1}GGUGΨAGCGGDU (acp^{3}U)AUCACAΨCUGCC$_m$UIACAm^{5}CGCAGAAGm^{7}GCCCCCGGTΨC Gm^{1}AUCCCGGGCGGAAACACCA. About 50% of the U residues at position 20 are modified to acp^{3}U. One of the C residues at position 48 or 49 is probably modified to m^{5}C. The nucleotide sequence of $tRNA_{3b}^{Val}$ is: GUUUCCGUAGUGΨAGCGGDacp^{3}UAUCACGΨGUGCUUC ACACGCACAAGm^7-GDCCCCGGTΨCGm^{1}AACCC GGGCGGGAACACCA. The C residue at position 48 is probably modified to m^{5}C. The observed codon responses of the two tRNAs are discussed in relation to the anticodons found.

The two $tRNA_{4}^{Val}$ genes of the recombinant plasmid pDt55 were sequenced by the Maxam and Gilbert method. This plasmid hybridizes to the

70BC site on the polytene chromosomes, a major site of $tRNA_4^{Val}$ hybridization. The two genes are of opposite polarity and are separated by 525 bp of DNA. The genes have identical sequences, which correspond to that expected from the sequence of $tRNA_4^{Val}$.

The nucleotide sequence of the $tRNA_{3b}^{Val}$ gene of recombinant plasmid pDt78R was also determined. This plasmid hybridizes to the 84D site, a major site of $tRNA_{3b}^{Val}$ hybridization. The sequence of the gene corresponds to that expected from the sequence of $tRNA_{3b}^{Val}$.

Comparison of the valine tRNA genes sequenced in this study and those determined by other workers shows that tRNA genes from major sites of $tRNA_{3b}^{Val}$ or $tRNA_4^{Val}$ hybridization to polytene chromosomes correspond exactly to the $tRNA^{Val}$ sequences while $tRNA^{Val}$ genes from minor sites of $tRNA^{Val}$ hybridization differ at 4 positions from the sequences expected on the basis of the tRNA sequences. The possible significance of this finding is discussed.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Page

Acknowledgments

I wish to thank Gordon Tener and Ian Gillam for much encouragement, guidance and good advice during the course of this work. I would also like to thank Allen Delaney, Shizu Hayashi, Dave Cribbs, Caroline Astell and Rob Dunn for many helpful discussions.

DEDICATION

TO

MY PARENTS

Abbreviations Used

A_{260}	- absorbance at 260 nm
A_{260} unit	- the amount of material giving an absorbance of 1.0 in 1.0 ml of solution in a 1 cm light path at 260 nm at neutral pH
acp^3U	- 3-(3-amino-3-carboxypropyl)uridine
bp	- base-pair
Cm	- 2'-O-methylcytidine
D	- dihydrouridine
i^6A	- $\underline{N}^6$-isopentenyladenosine
kb	- kilo base-pair
m^5C	- 5-methylcytidine
m^1G	- 1-methylguanosine
m^7G	- 7-methylguanosine
MOPS	- morpholinopropane sulfonic acid
OAc	- acetate
PEI-cellulose	- polyethyleneimine impregnated cellulose
Pu	- a purine nucleoside
Py	- a pyrimidine nucleoside
Q	- 7-(4,5-cis-dihydroxy-1-cyclopenten-3-ylaminomethyl)-7-deazaguanosine
QAE-Sephadex	- quaternary aminoethyl Sephadex
RNase	- ribonuclease
RPC-5	- reverse phase chromatography system 5
rT	- ribosylthymine
t^6A	- N-[(9-β-D-ribofuranosylpurin-6-yl)carbamoyl]threonine
TLC	- thin layer chromatography
Um	- 2'-O-methyluridine
Val 3a	- $tRNA_{3b}^{Val}$

Val 3b - $tRNA_{3b}^{Val}$

Val 4 - $tRNA_{4}^{Val}$

Y - α-(carboxyamino)-4,9-dihydro-4,6-dimethyl-9-oxo-1H-imidazo[1,2-a]-purine-7-butyric acid dimethyl ester

Introduction

In 1955 Crick (Judson, 1979) postulated the existence of a class of "adaptor" molecules capable of chemically combining with an amino acid and of hydrogen bonding to specific sequences in a template nucleic acid. Adaptors with such properties were proposed as intermediates required for the transfer of genetic information from the nucleotide sequences of nucleic acids to the amino acid sequences of proteins. Three years later Zamecnik and his coworkers, using a cell-free protein synthesizing system, (Hoagland *et al.*, 1958) detected small RNA molecules with some of the properties predicted for the adaptors. The study of the structure and function of transfer RNA, as the adaptors came to be called, continues to be an active field of research. In the sections that follow the structure and function of tRNA and tRNA genes are briefly described.

I. Structure of the tRNA Molecule

A. Primary Structure - RNA Sequencing

A prerequisite for understanding how tRNA functions at the molecular level is knowledge of its primary structure. In 1965 Holley published the first nucleotide sequence of any nucleic acid, yeast $tRNA^{Ala}$ (Holley *et al.*, 1965). Ultraviolet spectrophotometry was used to identify each mononucleotide produced during sequence analysis. The sensitivity of this method of detection is low, and large amounts of purified tRNA (95 mg) were required to determine the nucleotide sequence. High resolution tRNA purification procedures and the use of tRNA labelled *in vivo* or *in vitro* with the radioisotope phosphorus-32 have made feasible the rapid sequencing of very small amounts of tRNA.

A commonly used method of tRNA sequence analysis evolved from Holley's original strategy (reviewed by Silberklang *et al.*, 1979). First, the

nucleoside composition of the tRNA is determined. Knowledge of the number and nature of the modified nucleotides present in the tRNA is provided by this analysis. These data are required for all methods of tRNA sequence determination. Next, the tRNA is digested to completion with RNase T_1 (cleaves at G residues) or RNase A (cleaves at C and U residues) and the 5'-hydroxyl groups of the resulting oligonucleotides are phosphorylated with [^{32}P]phosphate. The radiolabelled oligonucleotides are separated by electrophoresis on cellulose acetate in the first dimension followed by chromatography on ion-exchange paper or TLC plates in the second dimension. Each purified oligonucleotide is then partially digested with snake venom phosphodiesterase and subjected to "wandering spot" analysis. In this analysis the phosphodiesterase digest undergoes electrophoresis at pH 3.5 on cellulose acetate strips which separates the RNA fragments according to their charges. At this pH the charge of an oligonucleotide reflects its base composition. The RNA fragments are transferred from the cellulose acetate strip to DEAE-cellulose TLC plates and fractionated according to size by homochromatography, a process of displacement analysis by an unlabelled mixture of oligonucleotides. After autoradiography of the chromatogram, the nucleotide sequence of the original oligonucleotide can be deduced from the pattern of mobility shifts displayed by its labelled degradation products. Removal of an A, G, C or U residue from the 3'-end of an oligonucleotide results in base-specific shifts in the direction and distance the remaining oligonucleotide migrates during "wandering-spot" analysis. Similar sequencing of large oligonucleotides produced by partial RNase T_1 or RNase A cleavage of the tRNA allows ordering of all the sequenced fragments into a unique tRNA sequence.

Recently, a new method of RNA sequencing has been developed that is similar, in principle, to the Maxam and Gilbert DNA sequencing procedure

(Donis-Keller et al., 1977; Simoncsits et al., 1977). Transfer RNA is labelled at the 5' or 3'-end with ^{32}P. In separate reactions the tRNA is partially digested with base-specific ribonucleases such as RNase T_1 for cleavage at G residues, RNase U_2 for cleavage at A and G residues, RNase A for cleavage at C and U and RNase Phy I for cleavage at all bases except C. Partial hydrolysis with alkali or hot, aqueous formamide produces indiscriminate cleavage of the RNA. The products of each reaction are applied to a series of slots in a denaturing polyacrylamide gel and fractionated according to size by electrophoresis. Autoradiography of the gel reveals the position of each G, A, U and C residue in the RNA. In practice, tRNA sequences derived by the "gel read-off" method are often ambiguous. The most serious difficulty is poor discrimination between U and C residues by the nucleases. More specific ribonucleases from chicken liver (Boguski et al., 1980), Staphylococcus aureus and Neurospora (Krupp and Gross, 1979) have been demonstrated to overcome this problem in their proponent's hands but are not yet widely used. Because the gel read-off method does not identify the positions of modified bases its utility in tRNA sequence analysis is limited. Finally, the strong secondary structure characteristic of tRNAs results in poor enzymatic cleavage in parts of the molecule. This makes interpretation of the gels difficult.

Conditions for base-specific chemical cleavage of RNA have been developed by Peattie (1979). The discrimination between U and C residues is reported to be excellent and secondary structure does not affect the reactivity of the RNA with the sequencing reagents. In spite of these advantages the chemical sequencing method is not widely used at present. This may reflect difficulty in reproducing the cleavage conditions used by Peattie in the original paper. Although the gel read-off method cannot, by itself, be used to determine tRNA sequences it can provide data supporting

sequences obtained by other methods.

An RNA sequencing procedure developed by Stanley and Vassilenko (1979) has proved particularly useful in sequencing tRNAs. This method requires purified RNA that is homogeneous in length and that has a phosphate group esterified to its 5'-hydroxyl group. The RNA is subjected to very limited, random hydrolysis. The conditions of hydrolysis ensure that the small number of molecules cleaved only once greatly exceeds the number that undergo multiple cleavages. Cleavage of a polynucleotide chain at a single site results in two fragments, only one of which has a free 5'-hydroxyl group. These hydroxyl groups are labelled with [^{32}P]phosphate and the labelled fragments are fractionated by size on a denaturing polyacrylamide gel. The autoradiograph of the gel reveals a "ladder" of bands. Each band is excised from the gel, its RNA is eluted from the gel fragment and then hydrolysed to mononucleotides. Only the 5' nucleotide of each fragment is radioactively labelled. This nucleotide is identified by thin layer chromatography in several solvent systems. By identifying the nucleotide at the 5'-end of all the fragments the sequence of the RNA can be determined. Each nucleotide of the sequence can be treated individually, thus modified nucleotides can be chromatographed in the solvent systems best suited for their identification. This feature is particularly useful in the sequencing of tRNA.

Gupta and Randerath (1979) and Tanaka *et al.* (1980) have developed modifications of the Stanley and Vassilenko procedure. The ladder of RNA fragments on the polyacrylamide gel is printed directly onto ion-exchange TLC plates. The RNA fragments are digested with ribonuclease *in situ* and the plates are then developed in an appropriate solvent system or the nucleotides are separated by electrophoresis. The TLC plates are autoradiographed and the nucleotide sequence of the RNA can be determined. The modi-

fied procedures have two advantages over the original method. First, the manipulations required in excising and eluting the RNA fragments are eliminated. This results in a great saving in time and effort. Second, less RNA is required to determine a nucleotide sequence. Much of the flexibility inherent in the Stanley and Vassilenko method is, however, lost in the more rapid procedures derived from it.

The advances made in RNA sequencing methods in the last few years mean that a complete tRNA sequence can now be obtained using very small amounts of purified tRNA (10-20 μg). As a result sequencing tRNAs from a wider variety of organisms becomes practicable. Comparison of the primary structure of tRNAs from many sources will undoubtedly deepen our understanding of tRNA structure, function and evolution.

B. Secondary Structure

In their paper presenting the primary structure of yeast $tRNA^{Ala}$ Holley and his coworkers (1965) described three possible secondary structures for the RNA. Since 1965 about 175 tRNAs have been sequenced (Gauss and Sprinzl, 1981). All these tRNAs can be folded into a form similar to one of those proposed by Holley, the "cloverleaf" form shown in Figure 1 (Rich and RajBhandary, 1976).

The most prominent features of the structure are the four base-paired stem regions, three of which are closed by nonbase-paired loops. The acceptor stem contains the 5' and 3'-ends of the tRNA. It consists of a 7 base-pair stem and 4 unpaired nucleotides. The 5'-hydroxyl group is phosphorylated while the 3'-ends of all tRNAs have the same -CCA sequence. The dihyrouridine arm ("D-arm") has a 3 or 4 base-pair stem and a loop of variable size, ranging from 7 to 11 nucleotides. Two regions of variable length, α and β, flank the pair of G residues always found in the D-loop. α and β contain from 1 to 3 nucleotides each. These are usual-

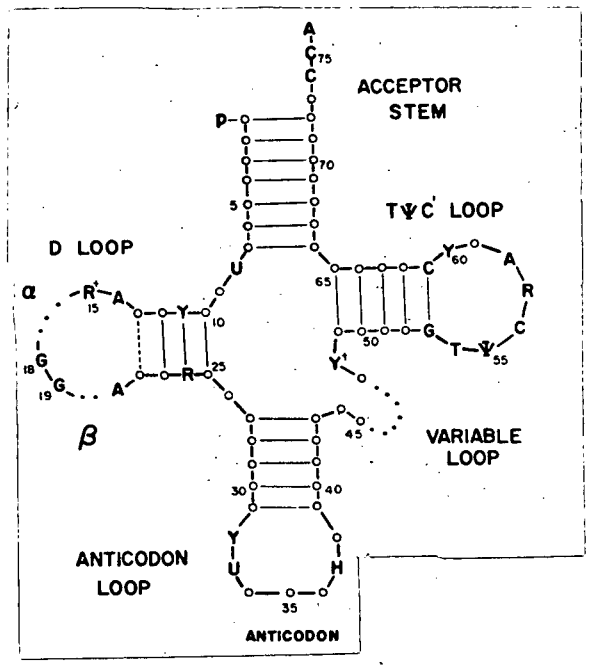


Figure 1. The Cloverleaf Structure of tRNA

ly pyrimidines, a high proportion of which is dihydrouridine. The anticodon arm is made up of a 5 base-pair stem and a 7 nucleotide loop. The third region of varying length in tRNAs is the variable loop. Two classes of tRNAs can be distinguished by the size of their variable loops. Most tRNAs fall into Class 1, tRNAs with 4 or 5 unpaired nucleotides in this region. Leucine and serine accepting tRNAs and prokaryotic $tRNAs^{Tyr}$ form Class 2. They have a long variable arm of 13-21 nucleotides which can form a stem and loop structure. The TψC-arm ("T-arm") consists of a 5 base-pair stem and a 7 nucleotide loop.

When the cloverleaf structures of sequenced tRNAs are compared, it becomes apparent that the nucleotides at certain positions in the structure are strongly conserved. Some positions are almost invariably occupied by the same nucleotide. These are: U8, A14, G18, G19, A21, U33, T54, ψ55, C56, A58, C61, C74, C75 and A76 (Figure 1). Other positions are almost always occupied by a pyrimidine (Y) and other sites by purines (R). These are: Y11, R15, R24, Y32, R37, Y48, R57 and Y60 (Figure 1). Purine 37,

adjacent to the 3' end of the anticodon (H in Figure 1) is often hypermodified and probably plays a role in the codon-anticodon interaction. Purine 15 and pyrimidine 48 (R^+, Y^+ in Figure 1), though widely separated in the cloverleaf structure, are usually complementary. This suggested that they might be close enough to form a base-pair in the tertiary structure of tRNA. X-ray diffraction studies of yeast $tRNA^{Phe}$ have confirmed this prediction and shown that 20 of the 23 strongly conserved nucleotides in this tRNA are involved in maintaining the molecule's tertiary structure (Rich and RajBhandary, 1976).

The patterns of primary and secondary structure described above are generalizations drawn from many tRNA sequences, and any particular tRNA species may differ in some respects from this pattern. One class of tRNAs, the initiator methionine tRNAs, are distinctly different from other tRNAs. Prokaryotic $tRNA_f^{Met}$ lacks a base-pair at the 5'-end of the acceptor stem and has an A11•U24 base-pair in the D-stem rather than the usual Y11•R24 pair. In eukaryotic $tRNA_i^{Met}$ T54Ψ55 is replaced by A54U55 and Y60 is replaced by A. In $tRNA_i^{Met}$ of higher eukaryotes the normally invariant U33 adjacent to the anticodon is replaced by a C residue.

C. Tertiary Structure

In 1968 tRNA was crystallized for the first time. This development meant the powerful technique of X-ray diffraction analysis could be used to determine the three-dimensional structure of tRNA. It was several years before the highly ordered crystals required for high resolution X-ray crystallography could be prepared. In 1975 two research groups published the crystal structure of yeast $tRNA^{Phe}$ to 2.5 Å resolution (reviewed by Kim, 1978; Rich and RajBhandary, 1976).

Yeast $tRNA^{Phe}$ is a flat, L-shaped molecule (Figure 2)(Rich and RajBhandary, 1976). It is 20-25 Å thick and each arm of the "L" is about

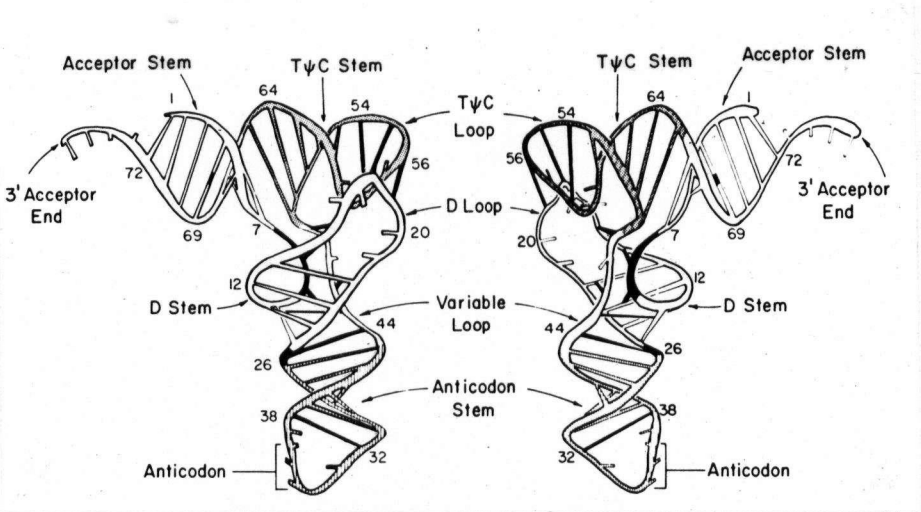


Figure 2. Two Views of the Tertiary Structure of Yeast $tRNA^{Phe}$

70 Å long. The 4 helical stem regions predicted in the cloverleaf structure of tRNA are indeed present in the tertiary structure. The acceptor stem is stacked on the T-stem to form one limb of the "L". The other limb is made up of the D-stem stacked onto the anticodon stem. In the latter helix there is a 25° deviation from linearity between the two component stem regions. The double-stranded regions of $tRNA^{Phe}$ approximate an RNA A helix. In this helix the base-pairs are tilted with respect to the helix axis and do not intersect with the axis. This results in a 6 Å hole running through the center of the helix. The A helix has a very deep major groove and a very shallow minor groove. There are 11 base-pairs per helical turn.

The T, D and variable loops are all clustered together where the two arms of the "L" intersect. There, a complex array of hydrogen bonds plays an important role in maintaining the molecule's tertiary structure. The anticodon loop is found at one end of the molecule. The conformation of this loop ensures that the anticodon is readily accessible for hydrogen

bonding to the codon on the ribosome. At the other end of the molecule the -CCA sequence, the site of aminoacylation, extends out into the solvent in a continuation of the acceptor stem helix.

What forces hold the intricate $tRNA^{Phe}$ structure together? Of great importance are tertiary hydrogen-bonds formed between the bases. These bonds are shown in Figure 3 (Kim, 1978).

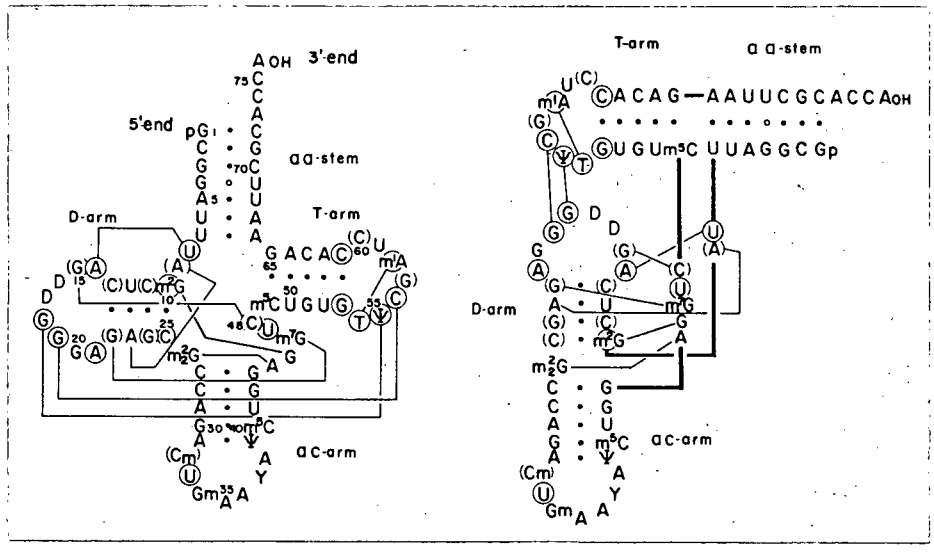


Figure 3. Tertiary Hydrogen-Bonds of Yeast $tRNA^{Phe}$. At left the clover-leaf form of $tRNA^{Phe}$, at right a drawing that more closely resembles the tertiary structure of $tRNA^{Phe}$.

Strongly conserved nucleotides are involved in almost all of these bonds. This suggests that interactions similar to those seen in yeast $tRNA^{Phe}$ are present in other tRNAs. Hydrogen bonds form a network holding the two arms of the tRNA in the correct orientation to one another. All the base-pairs in the major groove of the D-stem are involved in tertiary hydrogen bonds with the variable loop. The conserved GG doublet in the D-loop is bonded to the ΨC sequence in the T-loop. Uridine 8 and A9, located between the acceptor stem and the D stem, are hydrogen bonded to A14 and A23 respectively. These bonds may help stabilize the sharp bend in the RNA chain in the

U8-U12 region. With the exception of the G19-C56 bond none of the tertiary hydrogen bonds are of the Watson-Crick type. A wide variety of bonds are observed including reverse Watson-Crick bonds (G15-C48) and reverse Hoogsteen bonds (U8-A14; T54-m^1A58).

Other important interactions involve hydrogen bonds between bases and the 2'-hydroxyl group of ribose or the oxygen of phosphate residues. These bonds probably play an important part in maintaining tRNA structure but are difficult to detect with certainty by X-ray crystallography. Sharp bends in the polynucleotide chain appear to be stabilized by these types of hydrogen bonds. For example, A21 is bonded to the 2'-hydroxyl of U8 at the bend in the chain between the acceptor and D-stems. Similarly, the sharp bends in the T-loop and anticodon loop are probably stabilized by bonds between Ψ55 and phosphate 58 in the former and between U33 and phosphate 36 in the latter.

Five tightly bound divalent metal ions have been located in $tRNA^{Phe}$ (Rich et al., 1980). All these ions are octahedrally coordinated by water molecules and phosphate groups. One of the Mg^{++} ions binds at the 5'-end of the tRNA and the others are found at sharp bends in the RNA where segments of the negatively charged polynucleotide backbone come into close contact with one another. Highly ordered crystals of $tRNA^{Phe}$ can only be formed if spermine is present. Two molecules of spermine are associated with each tRNA in these crystals. One is found near the variable loop and likely stabilizes the bend in the U8-U12 region of the molecule. The other binds to the major groove of the RNA helix at the top of the anticodon stem.

A major contribution to the stability of tRNA structure is made by the extensive base-stacking present in the molecule. All but 5 bases in $tRNA^{Phe}$ are involved in base-stacking (Holbrook et al., 1978). One column of stacking interactions involves the acceptor stem, T-stem and T-loop. A

second column, roughly perpendicular to the first, involves the D-loop and stem and the anticodon arm.

Since the structure of yeast $tRNA^{Phe}$ was published, the tertiary structures of several other tRNAs have been determined at varying degrees of resolution (E. coli $tRNA_f^{Met}$, Woo et al., 1980; yeast $tRNA^{Asp}$, Moras et al., 1980; yeast $tRNA_i^{Met}$, Shevitz et al., 1980). The structure of yeast $tRNA^{Phe}$ appears to be typical of at least those tRNAs with a small variable loop. This is not surprising since the major determinants of tRNA structure, the cloverleaf pattern of base-pairing and the strongly conserved nucleotides, are common to almost all tRNAs.

The structure of tRNA in solution is very similar to its structure in a crystal lattice. A large body of data gathered by a wide variety of techniques including oligonucleotide binding, tritium exchange, base-specific chemical modification and NMR spectroscopy supports this conclusion (reviewed by Kim, 1978).

D. Mitochondrial tRNAs

Mitochondrial tRNAs (mt tRNA) are exceptions to the general pattern of secondary structure seen in other tRNAs. The mitochondrial genome contains genes for a set of mt tRNAs required for protein synthesis within the organelle. In lower eukaryotes the mitochondrial genome is relatively large (about 80 kb in yeast). Most of the tRNAs produced in these mitochondria are, judging from the sequence of their genes, similar in structure to cytoplasmic tRNAs (yeast, Bonitz and Tzagoloff, 1980; Aspergillus nidulans, Kochel et al., 1981). All these tRNAs can be folded into a standard cloverleaf structure but some lack a few of the invariant nucleotides found in cytoplasmic tRNAs.

The mitochondrial genomes of higher eukaryotes are about one-fifth the size of that of yeast (Borst and Grivell, 1981). Genes found in these small

genomes code for tRNAs that differ greatly from their cytoplasmic counterparts (human, Anderson et al., 1981; mouse, Van Etten et al., 1980). Most of the differences lie in the D and T-loops. The D-loop ranges from 3 to 10 nucleotides in length and the T-loop can be from 3 to 9 nulceotides long. Some or all of the normally invariant nucleotides found in these loops are missing. The most extreme example of the differences between mitochondrial and cytoplasmic tRNAs is mt $tRNA^{Ser}_{GCU}$ (de Bruijn et al., 1980; Arcari and Brownlee, 1980). This tRNA completely lacks a D-arm. In contrast, the mammalian mt $tRNA^{Leu}$ gene codes for a tRNA with all the structural features common to cytoplasmic tRNAs (Van Etten et al., 1981). Homologous mitochondrial tRNAs from different mammalian species show great sequence divergence. For example, human and bovine mt $tRNA^{Thr}$ are only 74% homologous. Homology between equivalent cytoplasmic tRNAs from different mammalian species is usually complete. Evidently, mammalian mitochondrial tRNA genes are evolving much more rapidly than their cytoplasmic counterparts (Borst and Grivell, 1981).

E. Modified Nucleotides Found in tRNAs

The large number and wide variety of unusual nucleotides found in tRNAs are a characteristic feature of these nucleic acids (reviewed by Nishimura, 1978). Over 50 different modified bases have been isolated from tRNA. Most of these result from methylation of the base or ribose moieties of nucleotides or from replacement of oxygen atoms in the bases by sulfur. More extensive alterations in nucleotide structure produce the hypermodified bases. Most of the modified nucleotides occur only at one or a few characteristic positions in the tRNA structure. Often the same modified nucleotide or its derivatives occupy the same site in homologous tRNAs from a wide variety of organisms. This suggests the modified bases of tRNA play important roles in tRNA structure and function.

Modified nucleotides in the first, or "wobble", position of the anticodon (when written in the conventional 5' to 3' direction) are directly involved in the codon-anticodon interaction. Unmodified A or U residues are almost never found in the "wobble" position, adenosine commonly being modified to inosine (I). In ribosome-binding experiments I in this position is capable of pairing ("wobbling") with A, C or U in the third position (3'-end) of the codon. Uridine in the "wobble" position is often modified to 2-thiouridine or its derivatives. These nucleotides will pair only with A in the third position of the codon. In Escherichia coli tRNA a U in the "wobble" position is sometimes modified to uridin-5-oxyacetic acid allowing it to base-pair with A, G or U in the codon. The hypermodified base Q (derived from G) or its glycosylated derivatives are found in the first position of the anticodon of some tRNAs. This base will pair with either U or C but has greater affinity for U.

The third position (3'-end) of the anticodon base-pairs with the first position (5'-end) of the codon during translation. This interaction must be very specific if errors in protein synthesis are to be avoided. If a tRNA has an A residue in the third position of the anticodon the A is almost invariably flanked on the 3' side by a hydrophobic base such as N^6-isopentenyladenine, Y base, or their derivatives. If a tRNA has a U residue in the third position of the anticodon the hydrophilic nucleoside t^6A or its derivatives are found immediately 3' to the anticodon. The function of these hypermodified bases may be to stabilize the A-U base-pair between the first position of the codon and the third position of the anticodon. G or C residues at the third position of the anticodon are flanked by simple methylated purines or by unmethylated A.

Other modified bases are found at specific sites in the tRNA molecule. Some of these are nearly universal such as rT and Ψ in the T-loop and

dihydrouridine in the D-loop of most tRNAs. Some, such as m^7G, are found only in tRNAs specific for certain amino acids. Yet others are found only in the tRNAs of some organisms. For example, 4-thiouridine is found only in prokaryotic tRNA while 5-methylcytosine is found only in that of eukaryotes. The function of modified nucleotides such as these is poorly understood.

II. Functions of Transfer RNA

A. Protein Synthesis

The most fundamental function of tRNA is its role in protein synthesis. Aminoacyl-tRNA is a crucial intermediate in the transfer of information from the nucleotide sequence of a mRNA to the amino acid sequence of a protein. The formation of aminoacyl-tRNA is catalyzed by the aminoacyl-tRNA synthetases (reviewed by Igloi and Cramer, 1978; Cramer et al., 1980). These enzymes (ENZ) catalyze a two step reaction. In the first step the amino acid (AA) is activated by forming a mixed anhydride with AMP:

$$\text{ENZ} + \text{AA} + \text{ATP} \rightleftharpoons \text{ENZ}\cdot\text{AA-AMP} + \text{PPi}$$

In the second step the amino acid is esterified to the free 2' or 3'-hydroxyl groups of the cognate tRNA:

$$\text{ENZ}\cdot\text{AA-AMP} + \text{tRNA} \longrightarrow \text{AA-tRNA} + \text{AMP} + \text{ENZ}$$

To ensure that protein synthesis is accurate an aminoacyl-tRNA synthetase must display high specificity for its substrates. Because structural differences among some amino acids are so slight substrate binding alone cannot explain the accuracy displayed by the synthetases. "Proofreading" of some kind is needed to prevent misacylation of tRNA. A number of proof- reading mechanisms have been proposed. In one scheme, proffered by Fersht (1980), all amino acids larger than the correct one are simply excluded from the synthetase's aminoacylation site because of their size. Amino acids smaller

or the same size as the correct substrate (e.g. valine instead of isoleucine) may be tranfered to the tRNA. These misacylated tRNAs are substrates for a second, hydrolytic, site. The correct aminoacyl-tRNA is excluded from the hydrolytic site because of its size or chemical properties and is released from the enzyme. Other mechanisms have been proposed which differ in detail from that of Fersht but share the same two-step discrimination process (Igloi and Cramer, 1978).

The features of tRNA recognized by the aminoacyl-tRNA synthetases are poorly known. Chemical modification studies indicate the acceptor stem, D-stem, anticodon loop and variable loop are all involved in tRNA recognition by the synthetases. In the crystal structure of $tRNA^{Phe}$ most of these regions lie on the side of the tRNA enclosed by the two arms of the L-shaped molecule.

The complex process of protein synthesis in eukaryotes is briefly outlined below (reviewed by Weissbach and Ochoa, 1976; Revel, 1977; Benne and Hershey, 1978). Initiation of protein synthesis requires the formation of a ternary complex between the charged initiator tRNA, methionyl-$tRNA_i^{Met}$, GTP and an initiation factor eIF-2. This complex binds to the 40S ribosomal subunit with the participation of factors eIF-3 and eIF-4C. Other initiation factors (eIF4A, eIF4B, eIF1) promote the binding of mRNA to the 40S subunit. ATP hydrolysis occurs at this step. The factor eIF-5 is required for the joining of the 60S ribosomal subunit to the 40S subunit. Concomitantly GTP is hydrolyzed and the initiation factors are released from the initiation complex. In this initiation complex Met- $tRNA_i^{Met}$ is in the ribosome's peptidyl (P) site (Revel, 1977). During elongation a ternary complex of aminoacyl-tRNA, an elongation factor (EF-1) and GTP binds to the ribosome. Hydrolysis of GTP releases an EF-1•GDP complex from the ribosome and leaves the aminoacyl-tRNA specified by the codon being

translated at the ribosome's aminoacyl (A) site. The peptidyl transferase activity of the ribosome catalyzes the transfer of the peptidyl (or methionyl) group from the tRNA at the P site to the free amino group of aminoacyl tRNA at the A site. In a concerted series of reactions deacylated tRNA is removed from the P site, the newly formed peptidyl-tRNA is transferred from the A to the P site and the ribosome moves to the next codon. Translocation is accompanied by GTP hydrolysis and requires the elongation factor EF-2. At this point the next cycle of elongation can begin. When a termination codon occupies the A site a release factor (RF) binds to the ribosome in a GTP dependent reaction. Peptidyl transferase then hydrolyses the peptidyl-tRNA bond, releasing the protein. GTP hydrolysis occurs and the RF is released from the ribosome. Dissociation of the ribosome into its component subunits can then occur, possibly catalyzed by the initiation factor eIF-3.

B. Regulation of Transcription

Aminoacyl-tRNA plays a role in the regulation of several bacterial operons. A large body of evidence supports a model for the regulation of these operons proposed by Lee and Yanofsky (1977). As an example, the regulation of the trp operon of *E. coli* will be briefly described in terms of this model (Yanofsky, 1981; Platt, 1981). The trp operon consists of a regulatory sequence and 5 genes coding for enzymes in the pathway from chorismate to tryptophan. Transcription of the operon is controlled by an operator-repressor system and by attenuation. Only the latter control mechanism will be described here. The leader sequence of trp operon mRNA contains a ribosome binding site, coding sequence for a short peptide containing two adjacent tryptophan residues and a transcription termination signal (the attenuator). The terminator consists of a G-C rich region of dyad symmetry followed immediately by a tract of uridine residues. Termina-

tion at this site is prevented if the G-C rich region cannot form a hair-pin. The leader region of the transcript can form a number of mutually exclusive secondary structures (Figure 4)(Platt, 1981). Region 1 contains

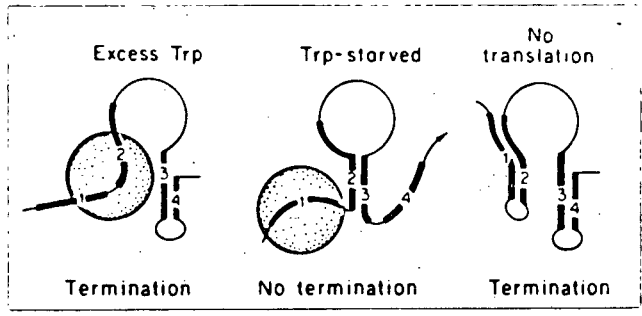


Figure 4. Model for Attenuation in the E. coli trp Operon

the end of the peptide coding sequence, including the two tryptophan codons. Region 2 is complementary to region 1 but part of it can base-pair with region 3. Regions 3 and 4 can form the essential stem and loop structure of the termination signal. In Yanofsky's model translation of the leader peptide is tightly coupled to transcription of the operon. If the tryptophanyl-tRNA concentration in the cell is low, the ribosome stalls when it reaches the tandem tryptophan codons. The ribosome covers region 1 of the transcript. Region 2, therefore, base-pairs to region 3 as soon as the latter is transcribed. This prevents formation of the hair-pin required for termination and transcription of the structural genes of the operon can occur. If, however, tryptophanyl-tRNA is abundant in the cell the ribosome reads through the tryptophan codons and stalls at the peptide's termination codon. At this site the ribosome covers regions 1 and 2 leaving region 3 and 4 free to form a hair-pin when they are transcribed. Termination of transcription occurs at the attenuator and the structural genes of the operon cannot be expressed.

C. Aminoacyl Group Transfer

In addition to its universal participation in protein synthesis

aminoacyl-tRNA serves as a donor of the aminoacyl group in a few biosynthetic reactions.

Lysylphosphatidylglycerol is one of the major lipids found in Staphylococcus aureus. This lipid is synthesized by transfer of lysine from lysyl-$tRNA^{Lys}$ to the 3'-hydroxyl group of phosphatidylglycerol. Alanylphosphatidylglycerol, a lipid found in Clostridium welchii, is made in an analogous manner (Soffer, 1974).

All eukaryotes assayed to date contain an enzyme that can transfer arginine from arginyl-$tRNA^{Arg}$ to the terminal amino group of acceptor proteins (Soffer, 1980). To serve as an acceptor a protein must have an N-terminal aspartate, glutamate or cysteine residue. A similar enzyme activity in Gram-negative bacteria transfers leucine, phenylalanine or methionine from their tRNAs to proteins with N-terminal arginyl, lysyl or histidyl residues. The physiological function of these enzymes is not understood.

The rigid cell walls of bacteria are made of murein. In this complex substance strands of peptidoglycan are crosslinked by short peptides. The length and sequence of these linker peptides varies from species to species but they are usually made by the sequential transfer of amino acids from aminoacyl-tRNA to the N-terminus of the growing linker (Soffer et al., 1974). In Staphylococcus epidermidis the linker peptide contains serine and glycine residues. This organism contains 1 $tRNA^{Ser}$ isoacceptor and 2 $tRNA^{Gly}$ isoacceptors that cannot function in protein synthesis. These tRNAs do, however, participate in cell wall synthesis. Roberts (1974) has determined the nucleotide sequence of the glycine isoacceptors. In the cloverleaf form these tRNAs have a 6 base-pair anticodon stem while the anticodon loop is reduced to 5 nucleotides. The GTΨC sequence found in almost all tRNAs is replaced by GUGC and the GG sequence usually present in

the D-loop is replaced by UU. A C residue substitutes for the purine adjacent to the anticodon in other tRNAs. These changes may produce tRNAs with tertiary structures optimized for their function in cell wall synthesis and that preclude their participation in protein synthesis.

D. tRNAs and tRNA-Like Structures Associated with Viruses

Transfer RNA is intimately associated with animal retrovirus replication. During retrovirus infection a viral reverse transcriptase produces a DNA copy of the RNA viral genome. Like other DNA polymerases, reverse transcriptase requires a primer. In these viruses a specific tRNA, of cellular origin, serves this function (reviewed by Dahlberg, 1980). In the avian retroviruses studied, the primer is a $tRNA^{Trp}$ species while in the murine viruses $tRNA^{Pro}$, or in some cases $tRNA_3^{Lys}$, is used. Only the 3' terminal 16-18 nucleotides of the tRNA hybridize to the viral RNA at the priming site near the 5'-end of the viral genome. Each retrovirus virion contains two copies of the RNA genome, to which the priming tRNAs are already hybridized, 80-100 molecules of reverse transcriptase, and 80-100 additional tRNA molecules. These tRNAs are a subset of total host cell tRNA. They are probably selected for inclusion in the virion by association with reverse transcriptase molecules.

The 3'-ends of many plant viruses contain remarkable structures that can be aminoacylated *in vitro* (reviewed by Haenni and Chapeville, 1980). Viruses classified in the same group usually accept the same amino acid. For example, turnip yellow mosaic virus and other tymoviruses accept valine while bromoviruses, typified by brome mosaic virus (BMV), accept tyrosine. In addition to aminoacyl-tRNA synthetases the viruses are substrates for other enzymes, such as RNase P, tRNA nucleotidyl transferase and some nucleotide modifying enzymes, that normally act upon tRNA substrates. The nucleotide sequences at the 3'-ends of several of these viruses are known. They

show little structural resemblance to the isoaccepting tRNA of the host cell. For example, BMV contains regions homologous to the acceptor stem and anticodon arm of host cell $tRNA^{Tyr}$ but in the viral RNA these homologies are separated by only 5 nucleotides. Further studies of the 3'-ends of plant viruses may yield new insights into features of tRNA structure recognized by aminoacyl-tRNA synthetases and other enzymes.

III. Drosophila melanogaster tRNA

The well developed state of Drosophila genetics makes this organism attractive as a system in which to study tRNA structure and function. The nucleotide sequences of 5 Drosophila tRNAs have been published, $tRNA_4^{Glu}$, $tRNA_1^{His}$, $tRNA_2^{Lys}$, $tRNA_i^{Met}$ and $tRNA_2^{Phe}$ (Gauss and Sprinzl, 1981). Though the number of known sequences is small a few conclusions can be drawn at this preliminary stage. As expected Drosophila tRNAs can be drawn in the standard cloverleaf form and contain the strongly conserved nucleotides present in most tRNAs. Drosophila $tRNA_2^{Phe}$, like that of Bombyx mori, is unusual among eukaryotic $tRNAs^{Phe}$. It has m^1G adjacent to the anticodon rather than the hypermodified Y base. Homology between Drosophila tRNAs and the corresponding tRNAs of vertebrates is great. This indicates that genes for cytoplasmic tRNAs have evolved very slowly in the higher eukaryotes.

Drosophila goes through several distinct stages in its life cycle: embryo (egg), first, second and third instar larva, pupa and adult. The tRNA species present in first and third instar larvae and in adult flies were resolved by RPC-5 chromatography. (White et al., 1973a). A total of 99 tRNA species were detected; of these 63 were major isoacceptors. The pattern of major isoacceptors for cysteine, glutamine, methionine, serine and threonine was found to change during the course of development. The

isoacceptor patterns for asparagine, aspartic acid, histidine and tyrosine show changes in tRNA modification during development (White et al., 1973b). The "δ" forms of these tRNAs contain a derivative of base Q at the first position of the anticodon. The "δ" forms contain G at this position. During larval growth the amount of Q-containing isoacceptors in a larva decreases. In an adult fly, however, the proportion of each tRNA species in the δ form is greatly increased.

The patterns of tRNA isoacceptors present in young and old flies were investigated by Hosbach and Kubli (1979 a,b). Transfer RNA isolated from old males (35 days) could not be aminoacylated to the same extent as the tRNA of young males (5 days). Most of the amino acids tested showed a drop in charging of 10-25% but leucine acceptance was reduced by 50%. The activity of the aminoacyl-tRNA synthetases for some amino acids did not change with age while the synthetases for alanine, leucine, serine and arginine were reduced by 50% in the old flies. The isoacceptor patterns for a number of amino acids were determined in old and young flies. Only tRNAs containing Q base show a change in pattern with age. Those isoacceptors containing Q (δ forms) increase with age while the δ forms decrease.

IV. Transfer RNA Genes

A. E. coli

E. coli contains about 60 tRNA genes (Brenner et al., 1970). Genetic and biochemical evidence suggested that most tRNA genes were tightly clustered at several sites on the chromosome (Smith, 1976; Ikemura and Ozeki, 1977). These findings have been confirmed by DNA sequence analysis of tRNA operons.

The 7 rRNA operons of E. coli all contain tRNA genes (Lund et al., 1976; Morgan et al., 1980). Within these operons rRNA genes are arranged in

the order 16S rRNA-spacer-23S rRNA-spacer-5S rRNA (Young et al., 1979a). Transfer RNA genes are found in two regions of the operon, in the spacer separating the 16S and 23S rRNA genes and distal to the 5S gene. Three of the operons have identical 16S-23S spacers each containing a $tRNA^{Ile}$ and a $tRNA^{Ala}_{1B}$ gene. The spacer in the other operons is shorter and contains a $tRNA^{Glu}_{2}$ gene. The two types of spacer share extensive regions of sequence homology which may be important for accurate processing of the operons' primary transcripts. Four rRNA operons contain tRNA genes distal to the 5S gene. Two have a single $tRNA^{Asp}_{1}$ gene in this region, in another the $tRNA^{Asp}_{1}$ is tightly linked to a $tRNA^{Trp}$ gene while in the fourth a $tRNA^{Thr}_{1}$ gene is found in the spacer between two 5S genes (Young, 1979b; Sekiya et al., 1980; Duester and Holmes, 1980).

Those tRNA genes not associated with rRNA operons are found in clusters, each cluster forming a single transcription unit. Each cluster may contain multiple copies of a single tRNA species (e.g. Duester et al., 1981) or genes for several different tRNAs (e.g. Nakajima et al., 1981). Within a cluster the spacers between genes vary greatly in length (9-200 bp) but are typically 20-40 bp long. With few exceptions (Schedl et al., 1974) E. coli tRNA genes code for the terminal -CCA sequence common to all tRNAs.

The sequences needed for transcription of tRNA operons are the same as those found in other E. coli operons. About 35 nucleotides upstream from the transcription start site is a sequence related to TTGACA. This element of the promoter is thought to interact with the σ subunit of RNA polymerase. The "Pribnow box", TATPuATPu, is found 10 nucleotides upstream from the start site and is thought to be recognized by RNA polymerase core (containing only $\alpha_2\beta\beta'$ subunits) (reviewed by Pribnow, 1979; Rosenberg and Court, 1979). In tRNA operons about 40 nucleotides separate the initiation site from the first tRNA gene. At the other end of the operons are

the transcription termination sites. These can be of two types. Both contain a G·C-rich region of dyad symmetry. The transcript of this region can presumably form a hair-pin structure. In rho-independent terminators the inverted repeat is followed by a series of adjacent T residues in the non-transcribed strand. Termination occurs in the corresponding series of U residues in the transcript. Rho-dependent terminators lack the T-rich sequence and require the presence of rho factor for efficient termination (Pribnow, 1979; Duester et al., 1981).

When the concentration of any aminoacyl-tRNA becomes the rate-limiting factor in protein synthesis, E. coli cells undergo a complex change in metabolism known as the stringent response. A prominent feature of the stringent response is a 10-20 fold reduction in the synthesis of rRNA and tRNA (Gallant, 1979). Travers (1980) has noted that promoters sensitive to the stringent response have a G·C-rich sequence between the Pribnow box and the transcription start site. Sequenced tRNA operons follow this pattern (Rossi et al., 1980; Duester et al., 1981; Nakajima et al., 1981). How, or even if, these sequences are involved in reducing stable RNA synthesis during the stringent response is unknown.

B. Organelles

Mitochondria and chloroplasts both contain small, circular DNA molecules. These genomes code for a complement of tRNAs required for the translation of mitochondrial and chloroplast mRNAs. Human mitochondrial DNA (mt DNA) is 16.5 kb long and codes for 2 rRNAs, 13 proteins, and 22 tRNAs of highly unusual structure (Anderson et al., 1981). The most notable feature of this genome is its extreme economy of organization. Both strands of the DNA are completely transcribed but the heavy (H) strand transcript is the precursor of the mitochondrial rRNAs, most of the mRNAs and most of the tRNAs. Usually, there are no spacer nucleotides between the gene tran-

scripts; instead tRNA sequences are interspersed among the rRNA and mRNA sequences (Montoya, 1981; Ojala, 1981). Attardi and his coworkers have proposed a "tRNA punctuation" model for the processing of the H strand primary transcript (Ojala, 1981). In this model mRNAs and rRNAs are generated by nucleases that precisely cut the transcript at the 5' and 3'-ends of the tRNA sequences.

Yeast mt DNA is about 5 times larger than that of higher eukaryotes (Borst and Grivell, 1981). The extreme compactness of organization characteristic of smaller mitochondrial genomes is absent in yeast. Most yeast mt tRNA genes are clustered together in one segment of the mitochondrial DNA.

The genes for spinach chloroplast tRNAs are located at many positions in the chloroplast DNA. Clusters of tRNA genes are found at some sites (Steinmetz et al., 1980). Recently, genes for $tRNA^{Ile}$ and $tRNA^{Ala}$ were discovered in the spacer between the 16S and 23S rRNA genes of maize chloroplasts (Koch et al., 1981). This arrangement is reminiscent of the tRNA genes found in some E. coli rRNA operons. The chloroplast isoleucine and alanine tRNA genes show strong homology to the corresponding prokaryotic tRNAs but contain enormous intervening sequences of 949 and 806 bp respectively.

C. Yeast

Saccharomyces cerevisiae contains about 360 tRNA genes per haploid genome coding for between 40 and 60 different tRNA species (Feldmann, 1976). By studying the sensitivity of yeast tRNA genes to U.V. light Feldmann (1977) determined that, unlike those of E. coli, most yeast tRNA genes have monocistronic transcription units. Subsequently, sequence analysis of tRNA genes from a wide variety of eukaryotes indicates that the great majority of eukaryotic tRNA genes are transcribed monocistronically. Exceptions to this general rule are, however, known. In bakers' yeast a

$tRNA^{Asp}$ gene is found only 10 bp downstream from a $tRNA^{Arg}$ gene (Schmidt et al., 1980). Similarly, in Schizosaccharomyces pombe 7 bp separates a $tRNA^{Ser}$ gene from one coding for $tRNA_i^{Met}$ (Mao et al., 1980). Transcription of either gene pair in a Xenopus oocyte extract produces a dimeric tRNA precursor.

The arrangement of nuclear yeast tRNA genes on the chromosomes is not known in detail. Genetic studies have shown that none of the 8 yeast $tRNA^{Tyr}$ genes are tightly linked and that they are found on 6 different chromosomes. EcoR I fragments bearing 7 of these genes have been cloned. Only one of the fragments contains more than a single tRNA gene (Olson et al., 1979). Similar cloning experiments by Beckmann et al. (1977) support the hypothesis that in yeast tRNA genes are not clustered but are widely distributed on the chromosomes.

Within some yeast tRNA genes are sequences not found in the mature tRNA. Originally discovered in the $tRNA^{Tyr}$ genes (Goodman et al., 1977), intervening sequences ranging in size from 14 to 34 bp have been found in yeast phenylalanine, serine, leucine and tryptophan tRNA genes (Valenzuela et al., 1978; Olson et al., 1981; Venegas et al., 1979; Ogden et al., 1979). The site of the intervening sequence in the gene is always between the first and second nucleotides after the anticodon. It is intriguing that the codons translated by the products of these genes all begin with U. In a family of isoaccepting tRNAs, genes for some isoacceptors may contain an intervening sequence while other, even closely related, isoacceptors may not (Olson et al., 1981). The transcripts of genes with intervening sequences can all be folded into a secondary structure with certain common features. In the tRNA precursors the anticodon may base-pair to a complementary region in the intervening sequence. There are two loops in the proposed secondary structure, the beginning of the first loop and the end of the second are the

sites at which the intervening sequence is excised during processing of the precursor (Figure 5)(Selker and Yanofsky, 1980).

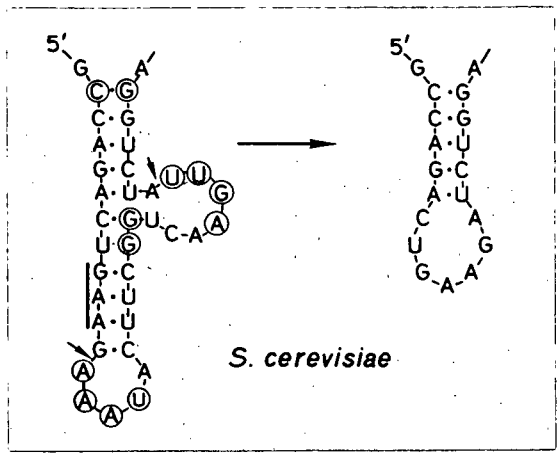


Figure 5. Removal of the Intervening Sequence from Yeast $tRNA^{Phe}$ Precursor

Intervening sequences are not unique to yeast tRNA genes. Similar intervening sequences have been found in tRNA genes of Drosophila, Xenopus, and Neurospora (Robinson and Davidson, 1981; Muller and Clarkson, 1980; Selker and Yanofsky, 1980). The function of intervening sequences in tRNA genes is unknown.

D. Xenopus

In contrast to the dispersed arrangement of tRNA genes in yeast, the tRNA genes of Xenopus laevis are organized in tandemly repeated arrays. Hybridization data indicate there are 8000 tRNA genes in Xenopus coding for about 43 types of tRNA (Clarkson et al., 1973a). The number of genes for different tRNA species is quite variable. For example, there are 310 genes for $tRNA_i^{Met}$ but only 170 genes for $tRNA_2^{Met}$ (Clarkson et al., 1973a). Clarkson et al. (1973b) sheared Xenopus DNA and centrifuged it to equilibrium on a CsCl density gradient. Even when the DNA was of high average molecular weight the genes for $tRNA_i^{Met}$, $tRNA_2^{Met}$ and $tRNA^{Val}$ were segregated from one another on DNA fragments of different buoyant

density ("cryptic" satellite DNAs). DNA containing $tRNA_i^{Met}$ genes was partially purified by repeated density gradient centrifugation. Restriction analysis of this DNA showed that the $tRNA_i^{Met}$ genes were located on a tandemly repeated 3.1 kb fragment of DNA (Clarkson and Kurer, 1976). The repeat unit was cloned and contained 2 $tRNA_i^{Met}$ genes and single copies of 6 other tRNA genes (Clarkson et al., 1978; Hofstetter et al., 1981). In this repeat unit the genes are irregularly spaced and are transcribed from both strands of DNA. It is likely the other "satellite" bands contain tandemly repeated families of tRNA genes. The arrangement of tRNA genes in Xenopus parallels that of the highly redundant 5S RNA, rRNA, and histone genes in this organism. The arrangement of these genes may be adapted to the requirements of oogenesis when large amounts of stable RNAs are accumulated.

V. Drosophila melanogaster tRNA Genes

A. Gene number

One of the most elementary questions that can be asked about Drosophila tRNA genes is: how many tRNA genes are there in the Drosophila genome? Ritossa et al. (1966) attempted to answer this question by quantitating the hybridization of a tRNA probe (total mixed tRNAs) to Drosophila DNA. They found that 0.015% of the total Drosophila DNA coded for tRNA, equivalent to 750 tRNA genes per haploid genome. In a more recent study Weber and Berger (1976) followed the kinetics of hybridization of Drosophila 4S RNA to Drosophila DNA. By measuring the rate at which the probe hybridized to the DNA the kinetic complexity of the tRNA genes could be determined (Britten et al., 1974). The observed complexity indicated there are 59 families of tRNA genes in Drosophila. The total number of tRNA genes was estimated to be about 590 per haploid genome. This figure is probably more accurate than

the earlier estimate because the 4S probe used in these experiments was of higher purity. The 56 gene families detected by hybridization are significantly fewer than the approximately 90 species of tRNA found in Drosophila by White et al. (1973a). The hybridization technique would probably not detect slight heterogeneity among tRNA genes, this would lead to an underestimate of sequence complexity. The difficulties in determining tRNA gene number by hybridization were demonstrated by the studies of Delaney (Tener et al., 1980). Purified tRNA isoacceptors were used as probes to determine the number of genes for individual tRNA species. The "saturation" level of hybridization was found to vary greatly with the RNA concentration, though in all cases there was an excess of RNA present. For example, the number of $tRNA_2^{Lys}$ genes ranged from 4 to 18 depending on the RNA:DNA ratio. The hybrids that did form were of two classes, one about 10 times more stable than the other. Seven of 18 $tRNA_2^{Lys}$ genes formed labile hybrids with the probe, the remainder were much more stable.

The number of genes for a particular tRNA can be estimated by digesting Drosophila DNA to completion with restriction enzymes and determining the number of DNA fragments (separable by gel electrophoresis) that contain genes for the tRNA of interest. This approach was used to estimate the number of genes for two valine isoacceptors: 17-19 $tRNA_4^{Val}$ genes (Dudler, 1981) and 6-7 $tRNA_{3b}^{Val}$ genes were detected (Tener et al., 1980).

B. Gene Location

The first information about the location of tRNA genes in the Drosophila genome came from the early hybridization studies of Ritossa et al.(1966). They were able to show that few, if any, tRNA genes were associated with rRNA genes in Drosophila. Since then the technique of in situ hybridization of tRNA to polytene chromosomes has greatly increased our

knowledge of tRNA gene location in Drosophila.

Polytene chromosomes are found in a number of Drosophila tissues, notably the larval salivary gland. They are the result of repeated rounds of DNA replication in the absence of cell division. All the chromatids resulting from the replication of each chromosomal arm are exactly aligned with one another (Beermann, 1972; Rudkin, 1972). Polytene chromosomes are much thicker and longer than mitotic chromosomes. Their most striking feature is a pattern of bands seen on all the chromosomal arms. The pattern of bands is constant for a given stock of flies and they serve as useful chromosomal markers. A system of numbers and letters is used to describe the location of sites on the polytene chromosomes (Lefevre, 1976).

In situ hybridization (Gall and Pardue, 1969) is widely used to locate genes on Drosophila polytene chromosomes. In this technique a ^{3}H or ^{125}I labelled probe is hybridized to its genes in a squashed preparation of the polytene chromosomes from larval salivary glands. Subsequent autoradiography reveals the sites of hybridization as silver grains clustered above specific loci on the chromosomes.

In situ hybridization of total 4S RNA to polytene chromosomes reveals 54 sites of hybrid formation, 26 strong sites and 28 weak sites (Steffensen and Wimber, 1971; Elder et al., 1980). The sites are randomly scattered over the chromosomal arms with the exception of the X chromosome, which has proportionally fewer sites than the others. No tRNA genes were detected on the small fourth chromosome. Any tRNA genes in the under-replicated and highly condensed heterochromatin may have escaped detection. The number of tRNA sites identified on the chromosomes is much less than the 600-750 genes detected by in vitro hybridization studies. This implies that many chromosomal sites contain several tRNA genes.

Atwood has hypothesized that the class of about 55 dominant mutations

known as Minutes is caused by deletion of tRNA genes (Ritossa et al., 1966). The Minute phenotype (delayed development, small bristles, and recessive lethality) was attributed to reduced protein synthesis due to suboptimal levels of tRNA. The location of Minute mutations does not coincide, on more than a chance basis, with the location of tRNA genes (Elder et al., 1980). Genetic evidence gathered by Huang and Baker (1976) supports the conclusion that the Minute phenotype is not due to deficiencies for tRNA genes.

Genes coding for 25 purified tRNA isoacceptors have been located by in situ hybridization (reviewed by Hayashi et al., 1981a). In general, genes for a particular tRNA species are found at more than one site. These sites may or may not be on the same chromosomal arm and a single site can contain genes for several different tRNAs. The chromosomal sites for the 3 major valine isoacceptors are presented in Table I (Hayashi et al., 1980, 1981a).

Table I. Sites of Drosophila melanogaster $tRNA^{Val}$ Genes on the Polytene Chromosomes

Isoacceptor	Major Site	Location	Minor Site[a]	Location
Val 3a	$64D_{1-2}$	3L	-	-
Val 3b	$84D_{3-4}$	3R	90BC	3R
	$92B_{1-9}$	3R		
Val 4	$56D_{3-7}$	2R	90BC	3R
	70BC	3L	89B	3R

a - About 1/5 the silver grains are found over a minor site as are found over a major site of hybridization.

C. Organization of Cloned Drosophila tRNA Genes

The technique of in situ hybridization of tRNAs to polytene chromosomes

provides much information about the large scale organization of tRNA genes. The development of techniques such as gene cloning, restriction mapping and DNA sequence analysis has made study of the fine structure of tRNA genes possible. No uniform pattern of tRNA gene organization has emerged from these studies. Rather, Drosophila tRNA genes are found to be organized in a variety of ways.

Davidson and his colleagues cloned overlapping fragments of Drosophila DNA that span 94 kb from the 42A region of chromosome 2 (Hovemann et al., 1980; Yen and Davidson, 1980). A total of 18 tRNA genes: 8 $tRNA^{Asn}$, 4 $tRNA_2^{Arg}$, 5 $tRNA_2^{Lys}$ and 1 $tRNA^{Ile}$, have been located in this large segment of DNA. The organization of these genes is illustrated in Figure 6 (Yen and Davidson, 1980).

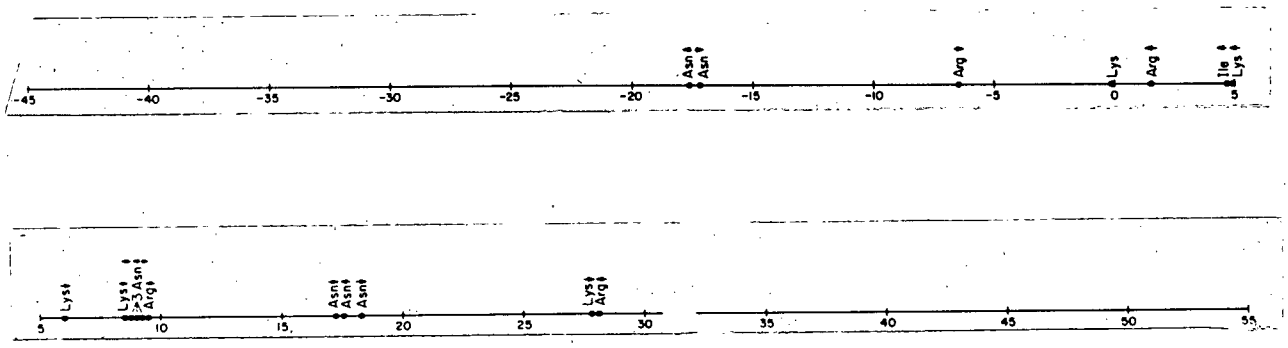


Figure 6. tRNA Genes at the 42A Region of the Drosophila Chromosome. Arrows denote polarity of the genes, arrowheads are at the 3'-ends of the genes.

They are widely dispersed and irregularly spaced within the central 46 kb of the DNA segment. Several clusters of genes occur. These clusters may contain identical genes or genes for several tRNAs. It is noteworthy that wherever identical genes are clustered two of the genes are of opposite polarity. The arrangement of some $tRNA_2^{Lys}$ genes as closely linked

inverted repeats may explain the puzzling results of hybridization of $tRNA_2^{Lys}$ to Drosophila DNA (Tener et al., 1980) described in a previous section (V.A). Under hybridization conditions two such genes could form hair-pin structures with each other rather than an RNA:DNA hybrid.

Eleven of the 18 genes have been sequenced. None contains an intervening sequence or codes for the -CCA end of the tRNA. All the genes for a given tRNA species are identical. All the sequenced genes are followed by runs of T residues in the non-coding strand. These are thought to be termination signals for RNA polymerase III (Valenzuela et al., 1977; Silverman et al., 1979). Some genes coding for the same tRNA have regions of homology in their 5'-flanking sequences. All but one of the cloned $tRNA_2^{Lys}$ genes in the 42A region have a sequence closely related to GGCAGTTTTTA about 25 bp upstream from the gene. A similar sequence is found 5' to a $tRNA_2^{Lys}$ gene from a different, but not precisely known, chromosomal location (De Franco et al., 1980). Similarly, all the sequenced $tRNA^{Arg}$ genes from the 42A region have a conserved nucleotide sequence (TCTTNACA or TGTTACA) about 20 nucleotides 5' to the genes (Yen and Davidson, 1980). A cloned fragment of DNA that hybridizes to the 50AB site on chromosome 2 codes for a cluster of 5 $tRNA^{Ile}$ genes and 2 $tRNA^{Leu}$ genes (Robinson and Davidson, 1981). All the $tRNA^{Ile}$ genes are preceded by a region strongly homologous to GCNTTTTG. At a similar position, about 25 bp before the coding sequence, GANTTTGG precedes the $tRNA^{Leu}$ genes. The significance of these conserved sequences is not yet known.

A number of other plasmids containing Drosophila tRNA genes have been isolated and characterized. Some of these display features of gene organization not found at the 42A site. The genes for $tRNA^{Leu}$ mentioned above are the first in Drosophila found to contain intervening sequences (Robinson and Davidson, 1981). The two intervening sequences differ from one another

in size but display regions of homology with the intervening sequence found in the yeast $tRNA_3^{Leu}$ genes.

Hershey and Davidson (1980) cloned a group of overlapping fragments of Drosophila DNA spanning 22 kb of sequence. Within this DNA segment they found a 1.1-2 kb fragment to be duplicated. Each repeat unit contains a single $tRNA^{Gly}$ gene. In situ hybridization studies showed that this repeat unit occurs only at the 56F site. The arrangement of these tRNA genes parallels the tandem arrangement of repeat units containing 5S RNA genes, also found at 56F (Prensky et al., 1973). It will be interesting to see how other tRNA genes, known to be present at 56EF (Hayashi, et al., 1981b), are arranged.

Heterogeneity in closely related tRNA gene sequences was found in a cloned cluster of $tRNA^{Glu}$ genes (Hosbach et al., 1980). Four of 5 $tRNA^{Glu}$ genes are identical but one contains a C to T transition at position 4 of the non-transcribed strand. Homologies among the flanking sequences of these genes indicate that the present gene cluster arose from a single ancestral gene by a combination of gene duplication and unequal crossing over.

Although few of the hundreds of tRNA genes in the Drosophila genome have been thoroughly studied some generalizations about the organization of these genes can be made. Many tRNA genes occur in clusters of a few genes. The clusters often contain genes for more than one species of tRNA. Within a cluster tRNA genes are irregularly spaced and may be transcribed from both strands of DNA. Frequently, closely spaced identical tRNA genes are arranged as inverted repeats. Often genes for a particular tRNA species are preceded by a short sequence common to other genes for the same tRNA. A single chromosomal tRNA hybridization site may contain several gene clusters as well as isolated tRNA genes.

VI. Transcription of Eukaryotic tRNA Genes

Eukaryotes contain 3 types of RNA polymerase. Each type transcribes a unique class of genes. RNA polymerase I transcribes rRNA genes, RNA polymerase II synthesizes the precursors to mRNAs and RNA polymerase III transcribes genes for a number of small RNAs including tRNA and 5S RNA. All three polymerases are of high molecular weight and are composed of from 10 to 15 subunits (Paule, 1981). The nucleotide sequences of many eukaryotic genes transcribed by RNA polymerase III have been determined. By analogy with bacterial promoters it might be expected that inspection of the sequences adjacent to the 5'-ends of these genes would reveal a conserved sequence, the RNA polymerase III promoter. This expectation has not been fulfilled. The 5'-flanking sequences of 5S and tRNA genes show no strongly conserved sequences.

A number of approaches have been used to determine the features of a tRNA gene essential for efficient transcription. Kurjan *et al.* (1980) selected spontaneous mutants of the yeast SUP4 $tRNA^{Tyr}$ locus that could no longer produce functional suppressor tRNA molecules. Sixty-nine mutants in the tRNA gene were mapped to 10 tightly linked sites. Several mutant genes from each of these clusters were cloned and the nucleotide sequence of each mutant gene was determined. None of the mutants had changes in the sequences flanking the gene. This suggests the gene's promoter does not lie in these flanking regions. Transcription of 29 of the sequenced mutant genes was tested in a transcription system derived from *Xenopus* kidney cells (Koski *et al.*, 1980). Most of the mutant genes were transcribed although changes in the length or quantity of the transcript were noted in some cases. Two of the mutant genes were not transcribed in this system. Both are altered at the same site within the gene. In one mutant, the C residue corresponding to the invariant C56 of the TΨC sequence in the tRNA is

changed to a G. In the other, it is changed to a U. Thus, it is likely that the C-G base-pair at this site in the gene is an important part of the RNA polymerase III promoter.

The elements of the eukaryotic tRNA gene promoter have been more precisely defined by the studies of Birnstiel and his coworkers (Kressmann et al., 1979; Hofstetter et al., 1981). From a point in the 5'-flanking sequence of a cloned Xenopus $tRNA_i^{Met}$ gene a series of deletion mutants were created, each with progressively more of the 5'-flanking sequence or gene sequence removed. The ability of each mutant gene to support specific transcription in vitro or after injection into Xenopus oocytes was tested. All of the 5'-flanking sequence and up to 9 bp of coding sequence could be deleted without impairment of transcription. These changes in the 5'-flanking sequence did alter the exact site of transcript initiation. Transcription did not occur if more than the first 10 nucleotides of the gene were deleted. Studies of a similar series of deletions from the 3'-end of the gene showed that deletion of the 3'-flanking sequence did not prevent transcription of the gene. A mutant gene lacking the 3'-end of the coding sequence (from nucleotide 55 on) as well as the 3'-flanking sequence could not be transcribed. Two deletion mutants were created by removing sequences of DNA from within the gene. These mutants lacked DNA between positions 12 and 28 or between positions 33 and 46. Neither gene could be transcribed in vitro. Transcription of the latter mutant gene, but not the former, could be restored by replacing the deleted section with DNA of unrelated sequence. This indicates that the particular nucleotide sequence between position 12 and 28 is required for transcription while the sequence of the DNA between positions 33 and 46 does not matter so long as it is at least 13 base-pairs in length. Insertion mutants were made by inserting short pieces of DNA at several sites in the $tRNA_i^{Met}$ gene. All the insertion mutants

could be transcribed. The transcripts of these mutants had the same initiation point as the transcript of the wild-type gene but were longer by the number of nucleotides introduced into the tRNA gene when the mutants were constructed.

These experiments indicate that the $tRNA_i^{Met}$ gene contains an internal promoter sequence. A model for this promoter has been proposed by Hofstetter et al. (1981). The promoter has two essential elements, one located between nucleotides 8 and 30 and the other between positions 51 and 72. These regions coincide with two highly conserved sequences in eukaryotic tRNA genes, GTPuGCGPyAGTNGG (acceptor stem, D-arm) and GGTTCGA(A/T)-PyCC(T-arm) respectively. The middle portion of the gene maintains the two elements of the promoter at a critical distance from one another. This distance may be increased but any significant decrease destroys promoter activity. Other genes transcribed by RNA polymerase III, such as the 5S RNA genes of Xenopus and the adenovirus VAI gene, contain internal promoters similar to that found in the $tRNA_i^{Met}$ gene (Sakonju et al., 1980; Bogenhagen et al., 1980; Fowlkes and Shenk, 1980).

In their study of the Xenopus $tRNA_i^{Met}$ gene Hofstetter et al. (1981) found that the gene's 5'-flanking sequence had only a minor effect on transcription, merely altering slightly the exact point of initiation. Several other studies indicate that the 5'-flanking sequence can play a much more significant role in the control of tRNA gene transcription. De Franco et al. (1980) transcribed two cloned Drosophila $tRNA_2^{Lys}$ genes in the Xenopus germinal vesicle system. One gene was transcribed about 10 times more efficiently than the other. The 5'-flanking sequence of the poorly transcribed gene was shown by these workers to be responsible for its transcriptional inactivity. Both genes are downstream from a conserved sequence found near many Drosophila $tRNA_2^{Lys}$ genes (Section V.C).

Sprague et al. (1980) transcribed a cloned Bombyx mori $tRNA_2^{Ala}$ gene in two systems, a Xenopus germinal vesicle extract and a system derived from Bombyx tissues. The intact gene was well transcribed in both systems. If all but 11 bp of the gene's 5'-flanking sequence was removed the gene was still transcribed in the heterologous Xenopus system but not in the homologous Bombyx system. In contrast to one of the Drosophila $tRNA_2^{Lys}$ genes described above, the 5'-flanking sequence of the gene stimulates rather than inhibits transcription. Thus, it is concluded that sequences in the 5'-flanking region seem to play a role in modulating the activity of adjacent tRNA genes.

VII. Processing of tRNA Transcripts

The primary transcripts of tRNA genes, in both prokaryotes and eukaryotes, contain nucleotide sequences not present in the mature tRNAs. The processing of these transcripts to produce functional tRNA molecules is briefly described below (reviewed by Mazzara and McClain, 1980).

The events, and particularly the enzymology, of tRNA processing are best understood in E. coli. In prokaryotes the primary transcript usually contains several tRNA sequences separated by spacers. Enzymatic cuts are made within the spacers liberating monomeric tRNA precursors that are then trimmed to mature size. RNase P_2 is an endonuclease that can cut tRNA precursors in the spacer region. Another endonuclease, RNase P, cuts monomeric pre-tRNAs, precisely at the 5'-end of the tRNA sequence. While endonucleolytic cleavage by RNase P generates the mature 5'-end of the tRNA, the 3'-end is produced by the action of an exonuclease. A great deal of confusion exists as to which nuclease is responsible for this trimming in vivo. RNase D removes nucleotides from the 3'-end of a precursor but stops when it reaches the terminal -CCA sequence (Ghosh and Deutscher, 1980). RNase D may

be the same enzyme as the previously described RNases P3 and Q. Other nucleases that have been implicated in E. coli tRNA processing are BN exonuclease and RNase III (Mazzara and McClain, 1980).

The primary transcripts of eukaryotic tRNA genes are usually monomeric. An RNase P-like activity and an exonuclease capable of processing the 3'-ends of tRNA precursors have been detected in a wide variety of eukaryotic cells. These enzymes have been partially purified from Bombyx mori (Garber and Altman, 1979). They are capable of accurately processing both eukaryotic and prokaryotic tRNA precursors. Unlike those of E. coli, eukaryotic tRNA precursors do not contain the 3'-terminal -CCA sequence. This sequence is added post-transcriptionally by the enzyme tRNA nucleotidyl transferase. Some eukaryotic tRNA precursors contain intervening sequences. Abelson and his coworkers have used a soluble extract from yeast that excises these intervening sequences to investigate the splicing reaction (Peebles et al., 1979; Knapp et al., 1979). Splicing occurs in 2 steps. First an endonuclease activity excises the intervening sequence to produce two half-molecules. The 3'-end generated by the excision is phosphorylated while the new 5'-end bears a free hydroxyl group. The second step is the ATP dependent ligation of the two half-molecules.

Transfer RNA maturation is a complex process. The intracellular location of processing and the order in which the processing steps occur was investigated by Melton et al. (1980). Plasmids carrying a yeast $tRNA^{Tyr}$ gene were injected into Xenopus oocytes. At several times after injection the nucleus and cytoplasm of some of the oocytes were separated, the RNA of each fraction was isolated and the processing intermediates were fractionated by polyacrylamide gel electrophoresis. The primary transcript was 108 nucleotides long and was confined to the nucleus. Trimming of nucleotides from the 5'-end of this transcript generated discrete processing inter-

mediates 104 and 97 nucleotides long. Base modifications at specific sites in the tRNA were already present in the larger intermediate. Next all the extra nucleotides were removed from the 5' and 3'-ends of the tRNA, the -CCA sequence was added, and base modification occurred at other sites in the tRNA. The product was an intermediate 92 nucleotides long. The final step was the excision of the intervening sequence. All these processing steps occurred within the nucleus.

VIII. The Present Investigation

The valine tRNAs of D. melanogaster are an attractive system in which to study the structure of tRNA isoacceptors and the genes that code for them. The investigation of Drosophila valine tRNAs reported in this thesis can be divided into two parts. In the first part, the relationship between the coding properties and the nucleotide sequences of Drosophila valine tRNA isoacceptors was investigated. These studies were prompted by reports of the peculiar coding behavior of rabbit liver $tRNA^{Val}_{(IAC)}$ as revealed by the ribosome-binding assay (Jank et al., 1977a). This tRNA, with an inosine residue in the first position of the anticodon, binds to ribosomes in the presence of any valine codon (GUA, GUC, GUU or GUG) but most strongly in the presence of GUG. The wobble hypothesis predicts that $tRNA^{Val}_{(IAC)}$ should bind to ribosomes in the presence of any valine codon except GUG (Crick, 1966). The nucleotide sequence of rabbit liver $tRNA^{Val}_{(IAC)}$ has several unusual features which may be responsible for its anomalous coding properties (Jank et al., 1977). Seven Drosophila $tRNA^{Val}$ isoacceptors have been resolved by RPC-5 chromatography, 3 major species $tRNA^{Val}_{3a}$, $tRNA^{Val}_{3b}$ and $tRNA^{Val}_{4}$, and the minor species $tRNA^{Val}_{1}$, $tRNA^{Val}_{2}$, $tRNA^{Val}_{5}$ and $tRNA^{Val}_{6}$ (Figure 7) (Dunn et al., 1978). In this study the coding properties of the 3 major valine isoacceptors were

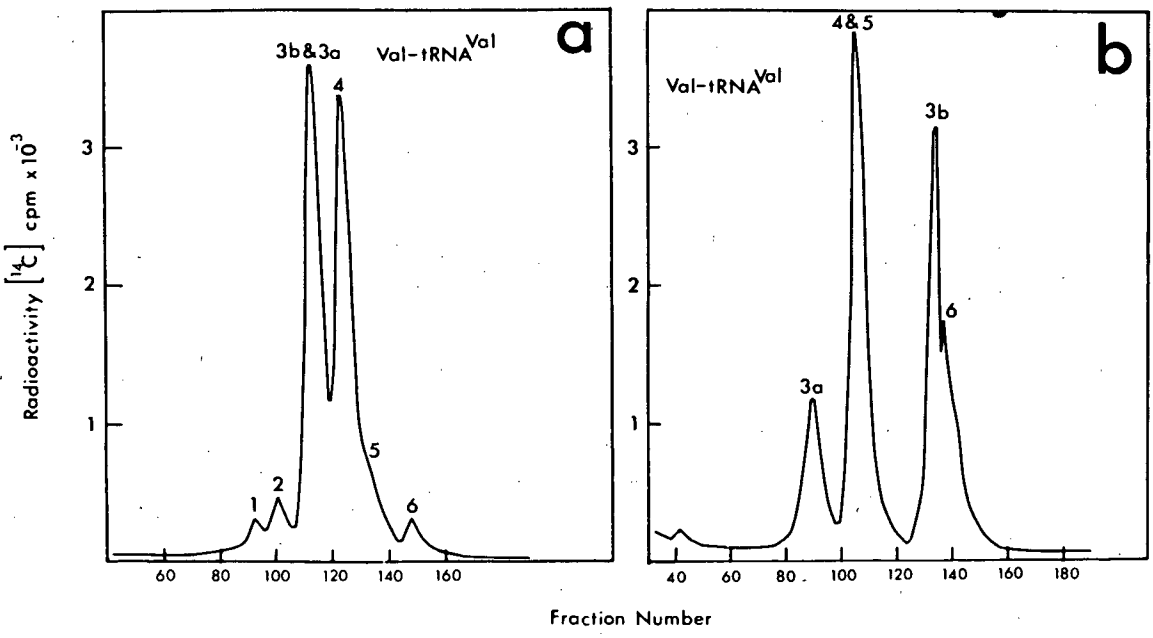


Figure 7. (a) Separation of [^{14}C]valyl-tRNAs^{Val} of crude Drosophila tRNA on an RPC-5column (0.9 x 21 cm) using elution system A, a 100 ml linear gradient of sodium chloride from 0.5 to 0.7 M in 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 10 mM sodium acetate pH 4.0. Flow rate 0.25 ml/min; fraction size, 0.5 ml. (b) Separation of [^{14}C]valyl-tRNA^{Val} as in (a) but with elution system C, a linear gradient of sodium chloride from 0.5 to 0.7 M, 1 mM EDTA, 1 mM 2-mercaptoethanol, and 10 mM sodium formate pH 3.8.

Table II. Nucleoside Analysis of tRNAs^{Val} from Drosophila

Nucleoside	Number of residues per 75-nucleotide tRNA molecule					
	tRNA$^{Val}_{3a}$		tRNA$^{Val}_{3b}$		tRNA$^{Val}_{4}$	
	I[a]	II[b]	I	II	I	II
A	14.6	15.1	14.6	15.2	13.7	14.7
U	8.7	8.7	8.8	7.8	8.1	6.3
G	16.2	17.8	18.7	23.9	16.9	21.0
C	20.5	18.3	22.8	18.0	22.3	19.0
ψ	5.7	4.6	3.7	4.0	4.0	4.0
hU	2.2	Nt[c]	2.3	Nt	1.6	Nt
rT	1.0	1.1	1.0	1.2	1.0	1.0
m^{1}A	1.0	0.80	0.83	0.73	1.0	0.90
m^{7}G	0.84	1.1	0.80	1.1	0.72	0.92
m^{5}C	0.80	0.73	0.88	0.96	2.0	1.7
m^{1}G	—	—	—	—	0.81	1.0
I	—	—	—	—	—	1.3
Cm	Nt	0.75[d]	Nt	—	Nt	1.3
Um	Nt	0.75[d]	Nt	—	Nt	0.9
Am	Nt	0.41	Nt	—	Nt	—
X	1.7	—	—	—	—	—
A*		0.80[e]	—	—	—	—
N		0.57[f]	—	—	—	—

[a]Columns labelled I are values obtained by the tritium-labelling technique (Materials and Methods).
[b]Columns labelled II are values obtained from the ultraviolet spectra of separated nucleosides (Materials and Methods).
[c]Nt, not tested.
[d]Estimated from two poorly resolved nucleoside spots.
[e]Assuming a molar extinction of 14 600 at 260 nm.
[f]Assuming a molar extinction of 9900 at 260 nm.

determined by the ribosome-binding assay of Nirenberg and Leder (1964). The nucleotide sequences of $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ were determined by a combination of modern RNA sequencing methods. The base composition of the Drosophila valine tRNAs (Table II), previously determined by Dunn (Dunn, et al., 1978), provided information essential to the sequence determinations.

In the second part of this investigation the nucleotide sequences of Drosophila valine tRNA genes were determined. Dunn et al. (1979) isolated and characterized recombinant plasmids carrying $tRNA^{Val}$ genes (Table III). In the present investigation the two $tRNA_4^{Val}$ genes of plasmid pDt55 and the single $tRNA_{3b}^{Val}$ gene of plasmid pDt78R were sequenced. Other investigators in this laboratory have sequenced several of the other genes coding for these two tRNAs. As a result different genes for the same tRNA can be compared to each other and to the nucleotide sequence of the tRNA. Such comparisons may help answer some important questions about Drosophila tRNA gene structure and expression. Are regulatory sequences present in the DNA flanking Drosophila tRNA genes? Are the gene sequences the same as the sequences of the tRNAs that hybridize to them? Are genes for the same tRNA from different chromosomal locations identical? How is the sequence homogeneity of widely scattered copies of a tRNA gene maintained?

Table III. Recombinant Plasmids Containing Drosophila $tRNA^{Val}$ Genes

Each of the recombinant plasmids listed here was cleaved by *Hin*dIII, electrophoresed in agarose gels and hybridized with complementary ^{125}I-labeled tRNA. All of the inserted *Hin*dIII fragments from the original isolates are listed; the 4.4 kb pBR322 parental fragment is omitted. In order to distinguish different plasmids containing similar sized *Hin*dIII inserts, the plasmids were cleaved with other endonucleases and analyzed by the procedure of Southern (1975). Frequently all of the fragments are listed in this Table; these include fragments liberated from the parental pBR322 DNA as well as those from the *Drosophila* insert DNA. Sometimes only the fragments which anneal with the [^{125}I]tRNA are listed.

tRNA	Plasmid No.	Prep. No.	Fragments *Hin*dIII	Other nucleases
Val 3b	pDt21	2	5.2*,0.6	3.4[ac]*
	pDt41R	3	2.0*,5.2	
	pDt48	3	2.4*	
	pDt78R	4	5.2*, 2.2	3.4[ac]*
	pDt85	4	5.2*,4.2	3.4[ac]*
Val 4	pDt14	2	12.0*	0.4[b]*
	pDt23	2	12.0*	0.4[b]*
	pDt55	4	8.0*	4.7,3.3[ac];0.8*,0.6[bc]*
	pDt62	4	8.0*,3.1	4.7,3.3[ac]
	pDt70	4	8.0*,3.2,1.1	4.7,3.3[ac]
	pDt92	4	1.7,0.5*	0.3[b]*;1.3[e]*
	pDt109	5	8.0*	4.7*, 3.3[ac]
	pDt110	5	2.0*,5.4	2.0*,2.1,2.7[ac]
	pDt112	5	8.0*	9.0*,3.3[a]
	pDt113	5	8.0*	9.0*,3.3[a]
	pDt114	5	8.0*,2.5,1.9	4.7*,3.7,3.3,2.6,1.7,1.4[ac]
	pDt115	5	8.0*	9.0*,3.3[a]
	pDt117	5	8.0*	9.0*,3.3[a];4.7*,3.3[ac]
	pDt118	5	8.0*	9.0*,3.3[a]
	pDt119	5	8.0*	9.0*,3.3[a];0.8*,0.6*[bc]
	pDt120	5	2.0*,5.4	3.2[a]*;1.9[e]*

[a] Fragment generated by digestion with *Eco*RI.
[b] Fragment generated by digestion with *Hae*III.
[c] Fragment generated by digestion with *Hin*dIII.
[d] Fragment generated by digestion with *Pst*I.
[e] Fragment generated by digestion with *Hha*I.
* Insert to which [^{125}I]tRNA hybridizes as determined by the procedures of Southern (1975).
R signifies that the parent plasmid has been recloned to give a new plasmid containing only the *Hin*dIII fragment bearing the tRNA gene.

Materials and Methods

I. General

A. Thin-layer Chromatography Solvents

Solvent A: 0.8 M $(NH_4)_2SO_4$, 10 mM Na_2EDTA

Solvent B: isobutyric acid: 30% NH_4OH: 1 mm EDTA 66:1:33

Solvent C: 0.1 M sodium phosphate buffer pH 6.8: $(NH_4)_2SO_4$: 1-propanol 100:60:2 v/w/v

Solvent D: 2-propanol: 12 N HCl: H_2O 70:15:15

B. Scintillation Counting

Aqueous samples containing ^{3}H, ^{14}C or ^{125}I were prepared for counting by mixing them with 7 volumes of a scintillation cocktail containing 66 g Omnifluor, 8.1 l xylene and 3.1 l Triton N-101. Radioactive precipitates deposited on paper, glass fibre or nitrocellulose filters were counted in a cocktail containing 12 g PPO, 0.3 g dimethyl POPOP and 8 l toluene. The radioactivity of samples containing ^{32}P was measured in water. The radioactivity of samples was measured in a Nuclear Chicago Isocap 300 scintillation counter.

C. Polyacrylamide Gel Electrophoresis

Acrylamide stock solution (45%) was deionized by stirring it successively with Dowex-1 (OH^-) and Dowex-50 (H^+) resin. Polyacrylamide gels of the required porosity were prepared by polymerization of a mixture containing the appropriate concentration of acrylamide, N,N-methylene-bisacrylamide (1/20 concentration of acrylamide, Eastman), 45 mM Tris-borate buffer pH 8.3, 1 mM EDTA, 0.08% thiourea and 0.15% v/v hydrogen peroxide (30%). Denaturing gels contained 7M urea. Analytical and sequencing gels were usually 0.5 mm thick while preparative gels were 1.5 mm thick.

D. Agarose Gel Electrophoresis

Agarose gels (0.5-2% agarose) were prepared by gelling an agarose

(Biorad) solution containing 45 mM Tris-borate pH 8.3, 1 mM EDTA, 1 µg/ml ethidium bromide in a Studier-type gel apparatus (McDonell et al., 1977). The sample wells were filled with electrophoresis buffer and the DNA samples (30 µl), containing 10% sucrose and 0.1% bromphenol blue, were carefully added. The wells were sealed by covering them with warm agarose solution. Electrophoresis was done at 100-250 V on a Savant flat-bed electrophoresis apparatus cooled with running tap water. Gels were exposed to ultraviolet light and photographed through an orange filter using Polaroid Type 57 or 667 film.

E. Autoradiography

TCL plates or polyacrylamide gels were autoradiographed by first wrapping them in Saran Wrap then pressing them into intimate contact with X-ray film (Kodak X-OMAT R, Kodak NS-5T, Agfa-Gevaert Curix RP-1). The films were exposed at -20°C and developed according to the manufacturers' instructions. When extra sensitivity was required, the film was preflashed (Laskey and Mills, 1977), sandwiched between the gel or TLC plate and an intensifying screen (Dupont Cronex Lightning-Plus), and exposed at -70°C.

F. Purification of $tRNA_{3b}^{Val}$ and $tRNA_{4}^{Val}$

$tRNA_{3b}^{Val}$ and $tRNA_{4}^{Val}$ were purified by Dr. I.C. Gillam as previously described (Dunn et al., 1978).

G. Synthesis of [5'-^{32}P]pCp

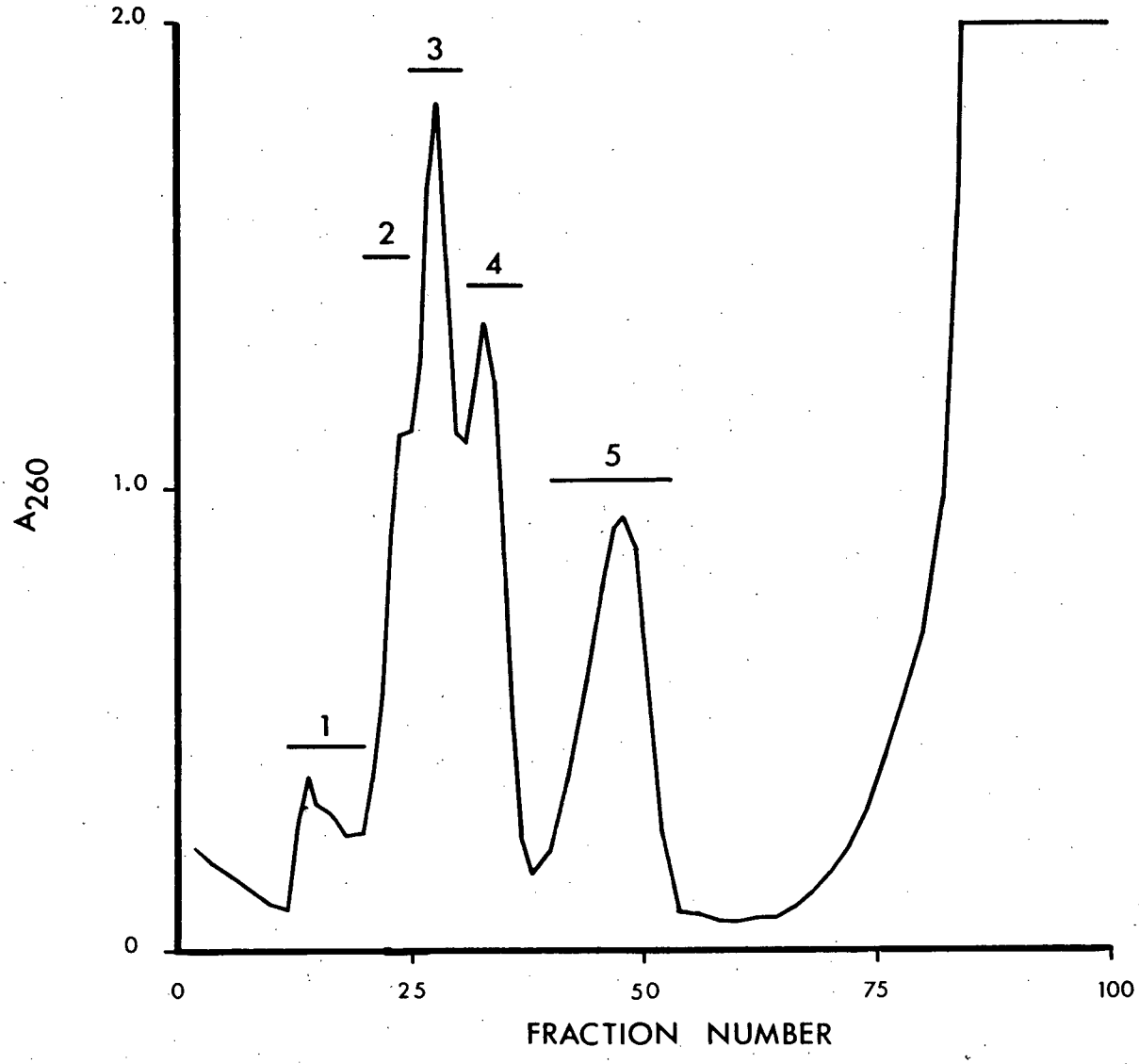
[5'-^{32}P]pCp was prepared by incubating a 10 µl reaction mixture containing 1.2 mM Cp(2' and 3'), 20 mM Tris HCl pH 8.3, 10 mM $MgCl_2$, 10 mM dithiothreitol, 35 µM [γ-^{32}P]ATP (>300 Ci/mmol, Amersham) and 2 units of T4 polynucleotide kinase (P-L Biochemicals) for 60 min at 37°C. The reaction was stopped by heating the mixture to 100°C for 1 min. The [5'-^{32}P]pCp was used without further purification.

H. Isolation and Characterization of acp^3U

3-(3-Amino-3-carboxypropyl) uridine-5'-phosphate was isolated from E. coli tRNA by the following method. Crude E. coli B tRNA (0.5 g) (Schwarz-Mann) was dissolved in 5 ml 50 mM NH_4OAc buffer pH 5.3 containing 2.5 mM EDTA, heated to 100°C for 2 min, cooled and incubated with 0.4 mg of nuclease P_1 (Calbiochem) for 16 h at 37°C. At the end of this time a further 0.2 mg of nuclease P_1 was added and incubation was continued for 7 more hours. 82% of the RNA was rendered acid soluble by the nuclease treatment. The enzyme digest was diluted with 25 ml of 50 mM Na_2CO_3 pH 10.5, 50 mM NaCl and applied to a QAE-Sephadex column (1.2 x 39 cm) previously equilibrated with the dilution buffer. Nucleoside-5'-phosphates were eluted from the column with a 600 ml linear gradient of NaCl (0.05 M to 0.8 M) in 50 mM Na_2CO_3 pH 10.5. The ultraviolet absorption of the eluted fractions was monitored at 260 nm and the major nucleotide in each peak of absorbance was identified by its ultraviolet absorption spectrum. The fractions making up the third peak, containing mostly uridine-5'-phosphate, were pooled and acidified with 12N HCl to pH 3. The nucleotides were desalted by adsorption on charcoal as described by Thomson (1960). The desalted nucleotides (500 A_{260} units) were dissolved in 10 ml distilled H_2O and applied to a 0.6 x 8 cm column of BioRad AG1-X2 resin (200-400 mesh, formate form). The nucleotides were eluted with a 100 ml linear gradient of formic acid (0 to 3 M). The elution profile of the column is presented in Figure 8. Peaks 4 and 5 contained nucleotides with ultraviolet absorption spectra (at pH 5 and 12) similar to those of acp^3U (Ohashi et al., 1974). Aliquots (0.2-0.5 A_{260} units) from each peak were applied to a cellulose TLC plate (Eastman) and developed in Solvent B. The plate was thoroughly dried, the spots of UV absorbing material were located and the chromatogram was sprayed with 0.1% ninhydrin in water-saturated 1-butanol.

Figure 8. Chromatography of the uridine-5'-phosphate fraction recovered from the QAE-Sephadex column on BioRad AG1-X2.

Fractions from the QAE-Sephadex column containing predominantly uridine-5'-phosphate (503 A_{260} units) were pooled, desalted, dissolved in distilled water and applied to a column (0.6 x 8 cm) of BioRad AG1-X2 (200-400 mesh, formate form) that had been thoroughly washed with distilled water. Adsorbed nucleotides were eluted with a 100 ml linear gradient of formic acid (0 to 3 M) at a flow-rate of 20 ml/hr. Eluted nucleotides were detected by their UV absorbance at 260 nm. Fractions corresponding to peaks of UV absorbance (1-5) (marked by solid lines in the Figure) were pooled and reduced to dryness over KOH at reduced pressure (10 mm Hg). The residue from each pool was dissolved in 1 ml distilled water and characterized as described in the text. Peak 5 was found to contain acp^3U-5'-phosphate (7.8 A_{260} units).


2.0
1.0
0
A_{260}
1
2
3
4
5
0
25
50
75
100
FRACTION NUMBER

The plate was heated to 80°C for 5 min. Only material from peak 5 gave the positive ninhydrin reaction expected of acp^3U (Ohashi et al., 1974). The position on the TLC plate of this ninhydrin positive material coincided with the position of the UV absorbing material in peak 5. Some of the nucleotide in peak 5 was dephosphorylated by sealing 8 µl of 0.1 M HCl containing 2 A_{260} units of the nucleotide in a capillary tube and heating the tube in the autoclave to 126°C for 75 min. The contents of the tube were applied to a cellulose TCL plate and developed in solvent D. Two spots of UV absorbing material were seen on the chromatogram possibly because dephosphorylation was not complete. Both spots were ninhydrin positive and one had an R_f very similar to that reported for acp^3U (Ohashi et al., 1974).

II. Trinucleotide-stimulated Binding of Valyl-tRNA to Ribosomes

A. Synthesis of Valine Codons

The four valine codons, GUA, GUG, GUC and GUU were made by the polynucleotide phosphorylase catalysed addition of the appropriate nucleotide to a GpU primer (Thach and Doty, 1965). Each 0.4 ml reaction mixture typically contained 0.2 M glycine pH 9.3, 0.4 M NaCl, 10 mM $Mg(OAc)_2$, 0.1 mM $CuSO_4$, 0.8 mM NDP, 5 mM GpU (Serva Feinbiochemica) and 1 unit primer-dependent polynucleotide phosphorylase (Micrococcus lysodeikticus, P-L Biochemicals). The reaction mixture was heated to 70°C for 5 min then cooled before adding the enzyme. After incubation at 37°C for 12-36 h the reaction was stopped by heating the mixture to 100°C for 2 min. Unreacted NDP was dephosphorylated by incubating the mixture with 0.2 mg bacterial alkaline phosphatase (Worthington) for 1 h at 37°C. The reaction mixture was diluted to 8 ml with distilled H_2O and applied to a column of DEAE-cellulose (HCO_3^- form, 0.7 x 50 cm). The nucleotides were eluted with a linear gradient (300 ml) of $(NH_4)HCO_3$ from 0 to 0.275 M. The trinucleotide diphosphate

product was detected as a peak of absorbance at 260 nm that emerged from the column after the major peak of unreacted dinucleoside monophosphate starting material. The trinucleoside diphosphate solution was desalted by repeated evaporation with ethanol/water under reduced pressure. The yields of trinucleoside diphosphate ranged from 48% for GpUpA to 10% for GpUpG.

The trinucleoside diphosphates were characterized by hydrolysing a small amount (0.5 A_{260} unit) of each in 10% piperidine at 100°C for 2 h or with RNase A (4 units in 10 µl 10 mM Tris-HCl pH 7.5) at room temperature for 2 h. The hydrolysates were applied to cellulose TLC plates along with the appropriate nucleotide and nucleoside standards and the plate was developed in Solvent B. All the trinucleoside diphosphates gave the expected hydrolysis products.

B. Isolation of [^{3}H]Valyl-tRNA^{Val} Isoacceptors

Crude Drosophila tRNA was isolated from adult Oregon R flies by the method of Roe (1975). Crude tRNA (1.3 mg) was aminoacylated with 0.1 mCi L-[2,3,4 -^{3}H]valine (10 Ci/mmol, New England Nuclear) using the conditions of White and Tener (1973c) for the reaction. After aminoacylation the reaction mixture (1.25 ml) was diluted with 3.5 ml of 50 mM NaOAc pH 4.5, 0.3 M NaCl, 10 mM $Mg(OAc)_2$ and applied to a small DEAE-cellulose column (1.1 x 5 cm) equilibrated with the same buffer. The column was washed with 60 ml of the above dilution buffer then the tRNA was eluted by increasing the NaCl concentration to 1.1 M. 1 ml fractions were collected manually and those fractions containing [^{3}H]valyl-tRNA were pooled. The nucleic acid was precipitated with 2 volumes of ethanol at -20°C and the precipitate collected on a Millipore filter (Type HA), washed with ethanol and dried.

The 3 major [^{3}H]valyl-tRNA^{Val} isoacceptors were separated by RPC-5 chromatography using the buffer systems of Dunn et al. (1978). The precipitated tRNA (above) was dissolved in 3 ml of 0.45 M NaCl in Buffer A (10 mM

NaOAc pH 4.0, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol) and applied to an RPC-5 column (0.9 x 60 cm) equilibrated with the same buffer. The tRNA was eluted from the column with a linear gradient (600 ml) of Buffer A containing 0.5 to 0.65 M NaCl. The column temperature was 37°C and the flow rate was 15 ml/h. Fractions containing the two major peaks of radioactivity, corresponding to [^{3}H]valyl-$tRNA^{Val}_{3a+3b}$ and [^{3}H]valy-$tRNA^{Val}_4$ (Fig. 10A), were each pooled and precipitated with ethanol as described above. The two valine tRNA fractions were further purified by RPC-5 chromatography in a second buffer system. Each sample was dissolved in 3 ml of Buffer C (10 mM sodium formate pH 3.8, 1 mM Na_2 EDTA, 1 mM 2-mercaptoethanol) containing 0.45 M NaCl and applied to an RPC-5 column (0.9 x 60 cm) that had been equilibrated with the second buffer. The column was eluted with a linear gradient (600 ml) of Buffer B containing 0.55 to 0.75 M NaCl, the column temperature and flow rate were the same as for the previous column. The 3 purified [^{3}H]valyl-tRNA isoacceptors (Fig. 10B, Fig. 10C) were each precipitated with ethanol and the precipitates were collected on Millipore filters for storage at -70°C.

C. Determination of Codon Triplet- Stimulated Binding of Valyl-$tRNA^{Val}$ to Ribosomes

The codons recognized by each of the valine tRNA isoacceptors were determined using the ribosome binding assay of Nirenberg and Leder (1964). Each ribosome-binding reaction (50 µl) contained 0.1 M Tris-OAc pH 7.2, 20 mM $Mg(OAc)_2$, 50 mM KCl, 2 A_{260} units of ribosomes, 4 pmol [^{3}H]valyl-$tRNA^{Val}$ isoacceptor and from 0 to 5 nmoles of one of the valine codon triplets. The reactions were incubated at 24°C for 30 min then stopped by the addition of 5 volumes of ice-cold wash buffer (0.1 M Tris-OAc pH 7.2, 20 mM $Mg(OAc)_2$, 50 mM KCl). The ribosomes were collected on Millipore filters (Type HA) and washed with 120 ml of cold wash buffer. The filters were dried and the

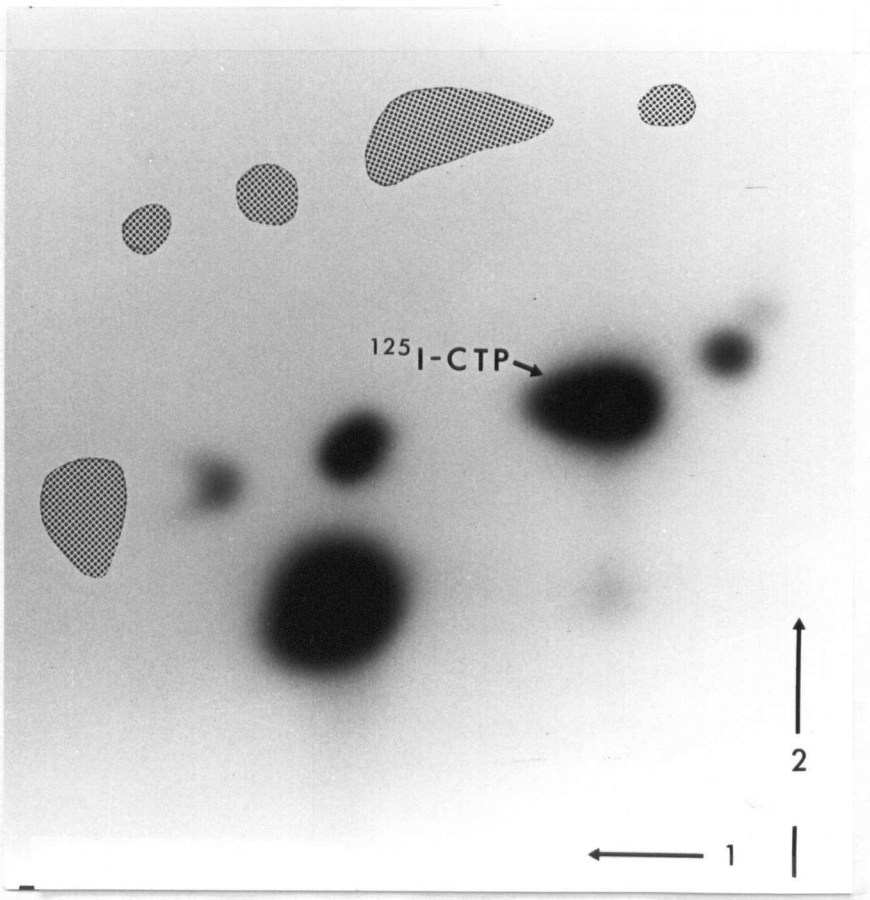
bound radioactivity measured by scintillation counting.

III. Synthesis of ^{125}I-CTP

^{125}I-CTP(^{125}I-5-iodocytidine-5'-triphosphate) of high specific activity was made by a modification of the procedure of Scherberg and Refetoff (1974). Two µl of carrier-free Na^{125}I (Amersham, 16 mCi/µg of iodine, approximately 550 mCi/ml) was added to 18 µl of cold (0°C) 44 mM CTP, 0.17 M NaOAc pH 4.0, 1.1 mM $TlCl_3$ in a capped 1.5 ml conical polypropylene tube and heated to 70°C for 20 min. The mixture was chilled on ice and 2 µl of 2-mercaptoethanol was added. To purify the ^{125}I-CTP, the reaction mixture was applied as a 1 cm streak near one corner of a 20 x 20 cm PEI-cellulose TLC plate (HCO_3^- form)(Polygram Cel 300 PEI, Machery-Nagel). The plate was developed with 0.2 M NH_4HCO_3, 10 mM EDTA in the first dimension, dried thoroughly, and developed with 0.1 M NaH_2PO_4, pH 6.8 (NaOH)/$(NH_4)_2SO_4$/0.25 M EDTA/1-propanol (100 ml:60 g:4 ml:2 ml) in the second dimension. This procedure separates ^{125}I-CTP from unreacted CTP and from iodinated side products (^{125}I-CDP, ^{125}I-CMP, etc.). The ^{125}I-CTP, which was located on the TLC plate by autoradiography (Figure 9), was scraped from the plate and eluted from the scrapings with 6 x 0.5 ml portions of 2.5 M NaCl. The ^{125}I-CTP was desalted by applying it to a column (0.7 x 1.0 cm) of acid-washed charcoal. The column was washed with 15 ml distilled water and the ^{125}I-CTP was eluted with 2 ml of ethanol/water/30% ammonium hydroxide (20:20:5). The eluate was reduced to dryness over P_2O_5 in a vacuum dessicator at 10 mm Hg and the residue taken up in 50 µl ethanol:water (1:1) for storage at -20°C. Twenty-five percent of the input ^{125}I could be incorporated into ^{125}I-CTP by this method, giving a product containing essentially undiluted ^{125}I-CTP.

Figure 9. Autoradiogram of a two-dimensional chromatogram used to separate ^{125}I-CTP from unlabelled CTP.

Shaded areas represent UV-absorbing materials. The PEI-cellulose TLC plate was developed in 0.2 M $(NH_4)HCO_3$, 10 mM EDTA in the first dimension (1) and in 0.1 M NaH_2PO_4, pH 6.8 (NaOH)/$(NH_4)_2SO_4$/0.25 M EDTA/1-propanol (100 ml:60 g:4 ml:2 ml) in the second dimension (2).


^{125}I-CTP
2
1

IV. Isolation of tRNA Nucleotidyl Transferase from Yeast

A. Enzyme Isolation

tRNA nucleotidyl transferase was isolated from bakers' yeast by modifications to the procedure of Morris and Herbert (1970). Compressed bakers' yeast (125 g, Fleishmann) was crumbled into fine pieces and added to a Dewar flask containing 300 ml of toluene at -70°C. The temperature was maintained at -70°C for 3 h by the addition of dry ice as needed. The frozen yeast was collected on a Buchner funnel (no paper), transferred to a 250 ml beaker and 40 ml of 1.25 M MOPS buffer pH 8 (NaOH) containing 10 mM EDTA and 10 mM 2-mercaptoethanol were added. The rest of the enzyme purification was carried out at 4°C. The cells were thawed on ice for 8 h, excess toluene was removed and the cells were allowed to autolyse for 24 h. During autolysis it was important to prevent the pH of the autolysate from falling below pH 7.5. The pH was adjusted by addition of 30% NH_4OH as needed. Cellular debris was removed from the autolysate by centrifugation at 12,000 x g for 15 min. The supernatant was brought to 58% saturation with $(NH_4)_2SO_4$ by addition of saturated $(NH_4)_2SO_4$ (saturated at 4°C) in 50 mM Tris-HCl pH 7.5, 25 mM EDTA and stirred for 2 h. The precipitated proteins were collected by centrifugation at 13,200 x g for 20 min and suspended in 40% saturated $(NH_4)_2SO_4$, 20 mM Tris-HCl pH 7.5, 10 mM EDTA, 15 mM $MgCl_2$, 10 mM 2-mercaptoethanol. After stirring for 2 h the solution was clarified by centrifugation at 16,000 x g for 20 min. One volume of saturated $(NH_4)_2SO_4$ in 25 mM Tris-HCl pH 7.5, 10 mM 2-mercaptoethanol was added to the supernatant and the mixture was stirred for 30 min at 4°C. The precipitate was collected by centrifugation at 16,000 x g for 20 min, drained well and dissolved in 10 ml of 20% glycerol, 5 mM Tris-HCl pH 7.5, 10 mM 2-mercaptoethanol (glycerol-buffer). The enzyme solution was desalted on a column (2 x 47 cm) of Sephadex G-25 (coarse) equilibrated with

glycerol-buffer. The peak of material absorbing at 280 nm from the Sephadex column was diluted to < 1 A_{280} unit/ml with glycerol-buffer and applied to a wide, short column (5.5 x 2 cm) of DEAE-cellulose. The material not adsorbed on the DEAE-cellulose column was applied to a column (1.2 x 10 cm) of phosphocellulose (Whatman P 11) that had been equilibrated with 20% glycerol, 25 mM Tris-HCl pH 7.5, 1 mM EDTA and 10 mM 2-mercaptoethanol. The enzyme was eluted from the column with a linear gradient (100 ml total volume) of 0 to 0.8 M NaCl in column buffer. The column fractions (1 ml) were assayed for tRNA nucleotidyl transferase activity and the enzyme was seen to emerge as a broad peak centered about 0.5 M NaCl. The enzyme was concentrated by diluting the pooled active fractions from the phosphocellulose column with 3 volumes of 40% glycerol, 0.25 M Tris-HCl pH 7.5, 1 mM EDTA, 10 mM 2-mercaptoethanol and applying the solution to a small column of phosphocellulose (1.2 x 2 cm) equilibrated with the dilution buffer. The enzyme was eluted with dilution buffer containing 0.6 M NaCl. The active fractions from the column were pooled, divided into 0.5 ml aliquots and stored at -70°C. A total of 4700 units (defined below) of activity were recovered and the enzyme concentration was 0.65 units/µl.

B. tRNA Nucleotidyl Transferase Assay

Substrate for the tRNA nucleotidyl transferase assay was prepared by periodate oxidation of the 3'-terminal ribose residue of tRNA followed by the lysine catalysed β-elimination of the oxidized nucleoside. The 3'-phosphate of the shortened RNA was removed with alkaline phosphatase. The 3'-terminal nucleotide was removed from 50 mg of crude yeast tRNA by the method of Khym and Uziel (1968). The product tRNA ("tRNA-CC") was obtained as its cetyltrimethylammonium salt. This precipitate was converted to the ammonium salt by extracting it 3 times with 0.1 M NH_4OAc in 70% ethanol.

The enzyme assay mixture (0.125 ml) contained 50 mM glycine pH 9.5

(NaOH), 10 mM $MgSO_4$, 0.15 mM CTP, 6 A_{260} units tRNA-CC, 2 mM [^{3}H]ATP (4 mCi/mmol) and 10 µl of sample solution. The mixture was incubated for 40 min at 37°C then 2 ml cold (0°C) 5% TCA was added. The precipitated RNA was collected on a glass fibre filter (Reeve Angel) and the filters were washed successively with two 2 ml portions of 5% TCA, 5 ml 95% ethanol and 5 ml ether, dried and counted in a scintillation counter. One unit of tRNA nucleotidyl transferase is the amount required to incorporate 1 nmol of ATP into substrate tRNA in 1 min at 37°C under the above assay conditions.

V. End-Labelling of tRNA

A. 3' End-Labelling

Radiolabelled nucleotides were incorporated into the 3'-end of tRNA molecules by either tRNA nucleotidyl transferase or T4 RNA ligase. In preparation for labelling with nucleotidyl transferase, the 3'-CCA end of the purified tRNA was removed by limited digestion of the tRNA with snake venom phosphodiesterase (Sprinzl et al., 1972). A 10 µl reaction mixture containing 10 µg tRNA, 0.6 units (Razzell, 1963) of snake venom phosphodiesterase (Worthington), 50 mM Tris, pH 8.0 and 10 mM $MgCl_2$ was incubated at 25°C for 30 min. The reaction was stopped by heating the mixture to 100°C for 1 min. An [α-^{32}P]AMP residue could be incorporated into the 3'-end of tRNA by incubating 10 µg of snake venom phosphodiesterase treated tRNA in 10 µl of 100 mM Tris-HCl buffer pH 9.0, containing 100 mM KCl, 10 mM magnesium acetate, 2 mM dithiothreitol, 1 mM EDTA, 0.15 mM CTP, 60 µCi [α-^{32}P]ATP (12 Ci/mmol, Amersham) and 1.2 units of tRNA nucleotidyl transferase for 20 min at 32°C. The reaction was stopped by adding 10 µl 7 M urea containing 0.5% xylene cyanol and bromphenol blue. The mixure was applied to the sample slot in a 16 x 17 x 0.15 cm 20% polyacrylamide gel containing 7 M urea. After electrophoresis the terminally labelled tRNA was

located in the gel by autoradiography and eluted from the gel slice by soaking in 0.5 ml 0.1 M sodium acetate, 0.1% SDS for 1-2 days at 4°C. The labelled tRNA and 10 µg of added carrier tRNA was precipitated from the eluate with 3 volumes of ethanol and was used for sequence analysis.

tRNA nucleotidyl transferase was also used to incorporate ^{125}I-CMP residues into the 3'-ends of tRNA molecules. Labelling was carried out in a 10 µl volume containing 2 µg phosphodiesterase-treated tRNA, 6-9 x 10^7 dpm ^{125}I-CTP, 50 mM glycine buffer, pH 9.5, 10 mM $MgCl_2$, 0.1 mM ATP, 5 mM 2-mercaptoethanol and 1.2 units of tRNA nucleotidyl transferase. After incubation at 32°C for 45 min 10 µl of a saturated solution of urea containing 0.5% xylene cyanol and 0.5% bromphenol blue was added to stop the reaction. The reaction mixture was applied to the sample slot in a 16 x 17 x 0.15 cm 20% polyacrylamide gel containing 7 M urea. After electrophoresis the terminally labelled tRNA was located on the gel by autoradiography, eluted from the gel slice by soaking in 0.4 ml 0.1 M NaOAc, pH 6.0, 0.1% SDS, 1 mM EDTA for 1 to 2 days at 4°C, and then precipitated from the eluate with ethanol in the presence of 20 µg carrier E. coli tRNA. The precipitate was rinsed with cold (-20°C) 70% ethanol, dried in vacuo and dissolved in water. 60 to 80% of the input ^{125}I-CTP was recovered in the tRNA.

tRNA was also labelled at the 3'-end with [5'-^{32}P]pCp using phage T4 RNA ligase (P-L Biochemicals) as described by England and Uhlenbeck (1978).

B. 5' End-Labelling

Labelling the 5'-end of the valine tRNAs using [γ-^{32}P]ATP and polynucleotide kinase was inefficient unless the tRNA was first cut in the anticodon loop by limited digestion with RNase U_2 to give half molecules. The 10 µl reaction mixture containing 2.5 µg $tRNA^{Val}$, 10 mM ammonium acetate buffer, pH 4.5 and 0.01 unit of RNase U_2 was incubated at 4°C for 1 h, heated to 100°C for 1 min, then chilled on ice. 1 µl of 1 M Tris-HCl

pH 7.5 and 1 µl 100 mM $MgCl_2$ were added to the reaction mixture and the tRNA fragments were dephosphorylated by adding 1 unit of calf intestine alkaline phosphatase (Boehringer Mannheim) and incubating the mixture at 45°C for 30 min. The reaction was stopped by adding 2 µl 50 mM nitrilotriacetic acid, 1 µl 2-mercaptoethanol, 1 µl 1% SDS and heating it to 100°C for 2 min. The RNA was precipitated with 3 volumes of ethanol at -70°C, the precipitate was rinsed with 95% ethanol and dried in vacuo. The RNA fragments were end-labelled with [^{32}P]phosphate by incubating them in a 10 µl reaction mixture containing 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 20 mM KCl, 10 mM $MgCl_2$, 2 mM spermine hydrochloride, 2 mM dithiothreitol, 20-30 µCi [γ-^{32}P]ATP (2000-3000 Ci/mmol, Amersham) and 1.5 unit phage T4 polynucleotide kinase (New England Biolabs) at 37°C for 30 min. The kinase reaction was stopped by adding 10 µl of 7 M urea containing 0.5% xylene cyanol and bromphenol blue and was applied to the sample slot in a 35 x 15 x 0.05 cm denaturing 20% polyacrylamide gel. After electrophoresis labelled tRNA fragments were located on the gel by autoradiography. Fragments approximately 35 nucleotides long were eluted from gel slices and used for sequence analysis.

VI. RNA Sequence Analysis

A. Stanley and Vassilenko Method

The sequencing of $tRNA_{3b}^{Val}$ and $tRNA_4^{Val}$ was done by a combination of methods. Most of the nucleotide sequence was obtained using the method of Stanley and Vassilenko (1978). Purified tRNA was prepared for sequencing by incubating it in a 10 µl reaction mixture containing 45 µg tRNA, 50 mM glycine pH 9.5 (NaOH), 10 mM $MgSO_4$, 2 mM ATP, 0.15 mM CTP, 2 mM dithiothreitol, 1.2 units tRNA nucleotidyl transferase and 2 units T4 polynucleotide kinase (New England Biolabs) for 40 min at 37°C. 20 µl of

saturated urea solution containing 0.1% xylene cyanol and bromphenol blue was added to the reaction mixture before it was applied to 3 wide (2.5 cm) sample slots in a denaturing 20% polyacrylamide gel (35 x 15 x 0.15 cm). Electrophoresis was carried out until the xylene cyanol marker dye had reached the bottom of the gel. The gel was stained with ethidium bromide (1 µg/ml), the topmost fluorescent band was excised and the RNA eluted by soaking the gel fragment overnight in 1.5 ml of 0.1 M NaOAc pH 6.0, 0.1% SDS, 1 mM EDTA at room temperature. The eluted RNA was precipitated with 3 volumes of ethanol and the precipitate was dried in vacuo.

The Stanley and Vassilenko (1978) sequencing procedure was modified in the following ways. Gel-purified tRNA (2 µg) was sealed in a glass capillary tube with 10 µl of either 5 mM MOPS buffer pH 7.2, 0.1 mM EDTA or 10 mM NH_4OAc pH 4.5, 1 mM EDTA and heated to 100°C for 3 min. The acetate buffer produced limited, random hydrolysis of the tRNA more consistently than did the MOPS buffer. Alkaline hydrolysis of the labelled RNA fragments was replaced by digestion with 0.05 unit RNase T_2 in 10 µl of 10 mM NH_4OAc buffer pH 4.6 containing 2 mM EDTA for 16 h at 37°C. The labelled nucleoside-5', 3'-bisphosphates (pNp's) were identified by comparing their mobilities on PEI-cellulose TLC plates (Polygram Cel 300 PEI, Machery-Nagel) in Solvent A and on cellulose TLC plates (E. Merck or Eastman Kodak) in Solvent B with published values (Silberklang et al., 1979; Cribbs, 1979). Modified pNp's were incubated with 1 µg nuclease P_1 (Calbiochem) for 16 h at 25°C in 5 µl 10 mM NH_4OAc buffer pH 4.6, 2 mM EDTA and the resulting nucleoside-5'-phosphates were identified by chromatography on cellulose TLC plates in Solvent B and Solvent C or D. In some cases the labelled RNA fragments were hydrolysed directly to nucleoside-5-phosphates with nuclease P_1.

B. Wandering-spot Analysis

The sequences near the 5' and 3'-ends of the tRNAs were determined by wandering spot analysis. The procedure followed was a modification of that of Silberklang et al.(1979). End-labelled RNA and 5 µg of carrier RNA were dissolved in 4 µl of deionized 98% formamide and partially hydrolysed by heating the solution for 1 h at 100°C. 1 µl of a dye mix containing 0.33% of xylene cyanol FF, orange G and acid fuchsin was added to the hydrolysate and the mixture was applied to a thoroughly blotted strip of cellulose acetate membrane (3 x 55 cm, Schleicher and Schuell No. 2500) that had been soaked in a buffer containing 5% HOAc, 7 M urea and 5 mM EDTA (pH 3.5). While the drop of RNA hydrolysate soaked into the cellulose acetate strip the sample was covered with an inverted 1 ml beaker to retard evaporation. The rest of the strip, except for about 1.5 cm around the origin, was covered with a layer of Saran Wrap. After the sample had soaked in, the strip was covered with single layer of Saran Wrap and connected by 3 MM paper wicks to the electrode compartments of a Shandon high voltage electrophoresis apparatus. The electrophoresis buffer contained 5% HOAc, pH 3.5 (pyridine) and 5 mM EDTA. Electrophoresis was done at 500 V till the dyes had separated, then the voltage was increased to 3000 V. Electrophoresis was stopped when the xylene cyanol dye had migrated 6-8 cm. The separated oligonucleotides were transferred from the cellulose acetate membrane to a 20 x 40 cm DEAE-cellulose TLC plate (Polygram Cel 300 DEAE, Machery-Nagel) by the method of Southern (1974). The TLC plate was washed in distilled H_2O, dried, and taped to a glass plate (20 x 38 cm). The TLC plate was developed with 1 mM EDTA at 65°C till the solvent front was 10 cm beyond the origin, then the plate was transferred to a second tank where development was continued with Homo-mix V (Jay et al., 1974) at 65°C till the solvent front reached the top of the plate. The plate was dried and autoradio-

graphed.

C. Gel Read-Off Method

RNA sequencing by partial, base-specific,enzymatic cleavage of end-labelled RNA (gel read-off method) was used to fill in gaps in the sequence and confirm results obtained by other methods. The procedures described are modifications of those of Donis-Keller et al. (1977), Simoncsits et al. (1977) and Gupta and Randerath (1977). Each RNase A (Worthington) or RNase T_1 (Calbiochem) reaction was done in 10 µl of a buffer containing 6 mM sodium citrate pH 5.0, 6 mM EDTA, 4.2 M urea and 0.5 µg of tRNA (end-labelled tRNA plus carrier tRNA). The 1.9 ml plastic centrifuge tubes (Evergreen) containing the reactions were heated to 100°C for 20 s then chilled on ice before the enzymes were added. Usually several reactions, each containing a different amount of enzyme, were done to ensure that a good distribution of partial digestion products was obtained. From 0.01-0.005 unit of RNase A and 0.1-0.01 unit of RNase T_1 were used. The reactions were incubated for 15 min at 50°C. RNase U_2 (Calbiochem, 0.01-0.001 unit) and RNase Phy I (Enzo Biochem, 1-0.1 unit) were incubated for 15 min at room temperature in 10 µl reactions containing 10 mM NH_4OAc pH 4.5, 1 mM EDTA and 0.5 µg RNA. All the enzymatic reactions were stopped by adding 1 µl of 0.5% SDS, 50% 2-mercaptoethanol to each reaction and heating them to 100°C for 30 s. The samples were reduced to dryness in a vacuum desiccator. Those samples containing urea were dissolved in 5 µl distilled H_2O containing 0.1% xylene cyanol and bromphenol blue, the remaining samples were dissolved in 7 M urea containing the same dyes. The samples were heated to 100°C for 30 s and applied to the slots of a polyacrylamide sequencing gel.

A "ladder" of fragments derived from breaks at every phosphodiester bond in the tRNA was produced by partial hydrolysis of the RNA in aqueous

formamide. Formamide was deionized by stirring it consecutively with Dowex-1 (OH^-) and Dowex-50 (H^+) resin and was stored at 4°C. End-labelled RNA (1 µg) and 12 µl of 66% formamide were sealed in a glass capillary tube and heated at 100°C for 20 min. 1 µl of 0.1 M NH_4OAc pH 4.5, 10 mM EDTA was added to the contents of the tube before the hydrolysate was applied to a slot in a sequencing gel.

D. Modification of tRNA with Chloroacetaldehyde

The occurrence of five consecutive Cp residues (C_{47}-C_{51}) in the variable loop and T-stem of $tRNA_4^{Val}$ could not be satisfactorily demonstrated by either the Stanley and Vassilenko or the gel read-off methods, presumably because of strong secondary structure in this region. Modification of the RNA with chloroacetaldehyde allowed the sequence to be resolved. A 50% solution (v/v) of chloroacetaldehyde was prepared by heating 25 µl of 97% chloroacetaldehyde dimethyl acetal (Aldrich) with an equal volume of 0.2 M HCl in a sealed glass tube at 100°C for 30 min. End-labelled tRNA plus carrier tRNA (2-5 µg) was dissolved in 15 µl of 1 M sodium acetate buffer pH 4.0, 1 mM EDTA containing 2 µl of the chloroacetaldehyde solution. The solution was sealed in a glass capillary tube and heated to 100°C for 1 min, then chilled on ice. The contents of the tube were expelled into 100 µl of 0.1 M sodium acetate buffer pH 4.0 containing 0.1 mM EDTA and the solution was incubated at 80°C for 20 min. The RNA was precipitated with ethanol and its sequence was determined by the gel read-off method.

VII. Recombinant Plasmid DNA Isolation

Bacteria containing recombinant plasmids were grown in M9S medium (Champe and Benzer, 1962) containing 0.4% glucose, 0.25% casamino acids, 0.025% $MgSO_4$, 0.02% uridine and 0.001% thiamine. Cultures were incubated

at 30°C with vigorous agitation (300 rpm) in an air shaker bath (New Brunswick Scientific) until the A_{650} of the culture was about 0.6. Chloramphenicol (80 mg/ml in 95% ethanol) was then added to a final concentration of 200 µg/ml and incubation was continued for 12-16 h. The cells in a 1 liter culture were collected by centrifugation and suspended in 10 ml of cold buffer (0°C) containing 12.5% sucrose, 25 mM Tris-HCl pH 8 and 0.25 M EDTA (pH 8). Egg white lysozyme (2.5 ml, Sigma 3x crystallized, 5 mg/ml) was added and the mixture was incubated for 15 min at 0°C. The cells were lysed by rapidly injecting 7 ml of 2% Triton X-100 into the cell suspension with a syringe (50 ml, 18 gauge needle). The lysate was centrifuged at 25,000 rpm for 1 h in a Beckman 30 rotor. The supernatant was extracted with 20 ml phenol:$CHCl_3$ (1:1, equilibrated with 50 mM Tris-HCl pH 8). The upper phase was reextracted as above, then twice with 20 ml portions of $CHCl_3$:isoamyl alcohol (24:1). The deproteinized cleared lysate was incubated with 1 ml of RNase A (1 mg/ml, Sigma, heated 100°C/5 min) for 10 min at 37°C then it was extracted once with phenol:$CHCl_3$ and once with $CHCl_3$:isoamyl alcohol as above. Nucleic acid was precipitated with 3 volumes of ethanol at -20°C for 16 h. The precipitate was dissolved in 10 ml 10 mM Tris-HCl pH 8, 10 mM NaCl, 1 mM EDTA then 3.75 mg of ethidium bromide and 10 g of CsCl (Koweki Berylco) were dissolved in the DNA solution. The solution was centrifuged to equilibrium at 35,000 rpm in a Beckman Ti50 rotor. Plasmid DNA was collected from the CsCl gradient. The DNA solution was extracted 3 times with water-saturated 1-butanol and twice with ether then dialyzed against 1 liter of 50 mM NaCl, 10 mM Tris-HCl pH 8, 1 mM EDTA. The buffer was changed once during the course of dialysis. The plasmid DNA was precipitated with 3 volumes of ethanol and dissolved in 1 ml of 5 mM Tris-HCl pH 8, 5 mM NaCl, 0.5 mM EDTA. In some cases the DNA was purified further on a second CsCl gradient, in others CsCl gradients were

replaced by chromatography on a Bio-Gel A-5m column.

VIII. Restriction Mapping of Recombinant Plasmids

A. Restriction Endonuclease Cleavage of DNA

Restriction buffers A, B and C each contained 6 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$ and 6 mM 2-mercaptoethanol. The NaCl concentration of the three buffers differed, being 6 mM, 50 mM and 150 mM respectively. Buffer D contained 100 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$ and 6 mM 2-mercaptoethanol. Decreased cleavage site specificity (EcoR I*, Dde I* activities) was displayed if DNA was cut with EcoR I or Dde I in a low salt buffer containing 20 mM Tris HCl pH 8.5, 2 mM $MgCl_2$, 5 mM 2-mercaptoethanol (Polisky et al., 1975). For complete digestion of plasmid DNA the reaction mixture contained 0.5 unit of restriction enzyme per µg of DNA and the incubation time was 2 h. A unit of enzyme is defined as the amount required to completely digest 1 µg of phage λ DNA in 15 min at 37°C. Restriction endonuclease Xma I was prepared by the method of Endow and Roberts (1977) or was the gift of Dr. C. Astell. Restriction enzymes Fnu4H I and FnuE I were the gift of Dr. M. Smith.

B. Restriction Mapping

Restriction maps of plasmids pDt55 and pDt78R were constructed by the multiple enzyme digest method described by Danna (1980). Hind III cut phage λ DNA provided size standards (23.1, 9.9, 6.6, 4.4, 2.46, 2.15 and 0.49 kb) useful in determining the size of restriction fragments. The method of Southern (1979) was used to estimate the length of DNA fragments from their electrophoretic mobilities.

IX. DNA Sequence Analysis

The nucleotide sequence of plasmid DNA restriction fragments was determined using the Maxam and Gilbert (1980) procedure as modified by Dr. A. Delaney.

Restriction fragments were labelled for sequencing by incubating the DNA (5-20 µg) in 20 µl of restriction buffer B containing 30-80 µCi of the appropriate [α-^{32}P]deoxynucleoside triphosphate (2000-3000 Ci/mmol, Amersham) and 1 unit of the Klenow fragment of E. coli DNA polymerase I (Boehringer Mannheim) for 15 min at room temperature. The solution was mixed with an equal volume of 15% sucrose containing 0.1% xylene cyanol and bromphenol blue and applied to the sample slot of a thin, non-denaturing polyacrylamide gel (35 x 15 x 0.05 cm) for fractionation of the labelled DNA fragments by electrophoresis.

If the labelled fragments were to be strand-separated, solid urea (to a concentration of 50% w/v) and the dyes were added to the labelling reaction mixture. The solution was then heated to 100°C for 3-5 min and quickly applied to the sample slot of a non-denaturing polyacrylamide gel. The sample was subjected to electrophoresis at 400 V for 10 min then at 100-250 V till the desired degree of separation had been achieved. Some fragments that could not be strand-separated at room temperature separated cleanly if the electrophoresis was performed at 4°C.

Sequence data were stored, edited and analyzed using the SEQNCE computer program of Dr. A. Delaney.

Results and Discussion

I. The Coding Properties of the Drosophila Valine tRNAs

The ribosome-binding assay developed by Nirenberg and Leder (1964) is used for determining the coding properties of tRNA isoacceptors. In the assay E. coli ribosomes are "programmed" with a trinucleoside diphosphate ("triplet") of known sequence. These ribosomes will bind aminoacyl-tRNAs with anticodons complementary to the triplet. Ribosome-bound aminoacyl-tRNA can be separated from unbound material by filtering the reaction mixture through a nitrocellulose membrane. Only ribosome-bound aminoacyl-tRNA is retained on the filter.

In determining the codon recognition properties of individual tRNA isoacceptors it is important that the isoacceptors are completely separated from one another. In this study the valine tRNA isoacceptors in a crude mixture of Drosophila tRNAs were aminoacylated with L-[^{3}H]valine and separated by chromatography on two successive RPC-5 columns. $tRNA_3^{Val}$ and $tRNA_4^{Val}$ were resolved by chromatography in a buffer containing Mg^{++} ions (Figure 10A). [^{3}H]valyl-$tRNA^{Val}$ from each of these peaks was chromatographed on a second column in a buffer containing EDTA (no free Mg^{++}). This chromatographic system separated $tRNA_3^{Val}$ into two subspecies, $tRNA_{3a}^{Val}$ and $tRNA_{3b}^{Val}$ (Figure 10B) and freed $tRNA_4^{Val}$ of contaminating $tRNA_{3a}^{Val}$ and $tRNA_{3b}^{Val}$ (Figure 10C).

The response of the 3 major valine tRNAs to each of the four valine codon triplets (GUG, GUA, GUU and GUC) is shown in Figure 11. $tRNA_{3a}^{Val}$ responded strongly to GUA and weakly to GUU and GUG. $tRNA_{3b}^{Val}$ responded only to GUG. $tRNA_4^{Val}$ responded strongly to GUU, GUC and GUA and weakly to GUG.

On the basis of ribosome-binding experiments, tRNA nucleotide sequences,

Figure 10. Purification of Drosophila valyl-tRNA^{Val} isoacceptors by RPC-5 chromatography

A. Crude Drosophila tRNA (54 A_{260} units) was aminoacylated with [^{3}H]valine and chromatographed on an RPC-5 column in buffer A (10 mM NaOAc, pH 4.0; 10 mM $Mg(OAc)_2$; and 1 mM 2-mercaptoethanol) as described in Materials and Methods. The numbers denote the elution positions of the 6 valine tRNA isoacceptors. Those fractions containing valyl-tRNA$_3^{Val}$ or valyl-tRNA$_4^{Val}$ that were pooled for further purification are indicated by the solid black lines.

B. Valyl-tRNA$_3^{Val}$ was fractionated into two subspecies, valyl-tRNA$_{3b}^{Val}$ (3a) and valyl-tRNA$_{3b}^{Val}$ (3b), by RPC-5 chromatography in buffer B (50 mM sodium formate, pH 3.8; 1 mM EDTA; 1 mM 2-mercaptoethanol) as described in Materials and Methods. The valyl-tRNA^{Val} of fractions marked by solid black lines was used for the ribosome-binding experiments.

C. Valyl-tRNA$_4^{Val}$ was purified by RPC-5 chromatography in buffer B as described in Materials and Methods. Valyl-tRNA^{Val} of those fractions marked by the solid black line was used for the ribosome-binding experiments.

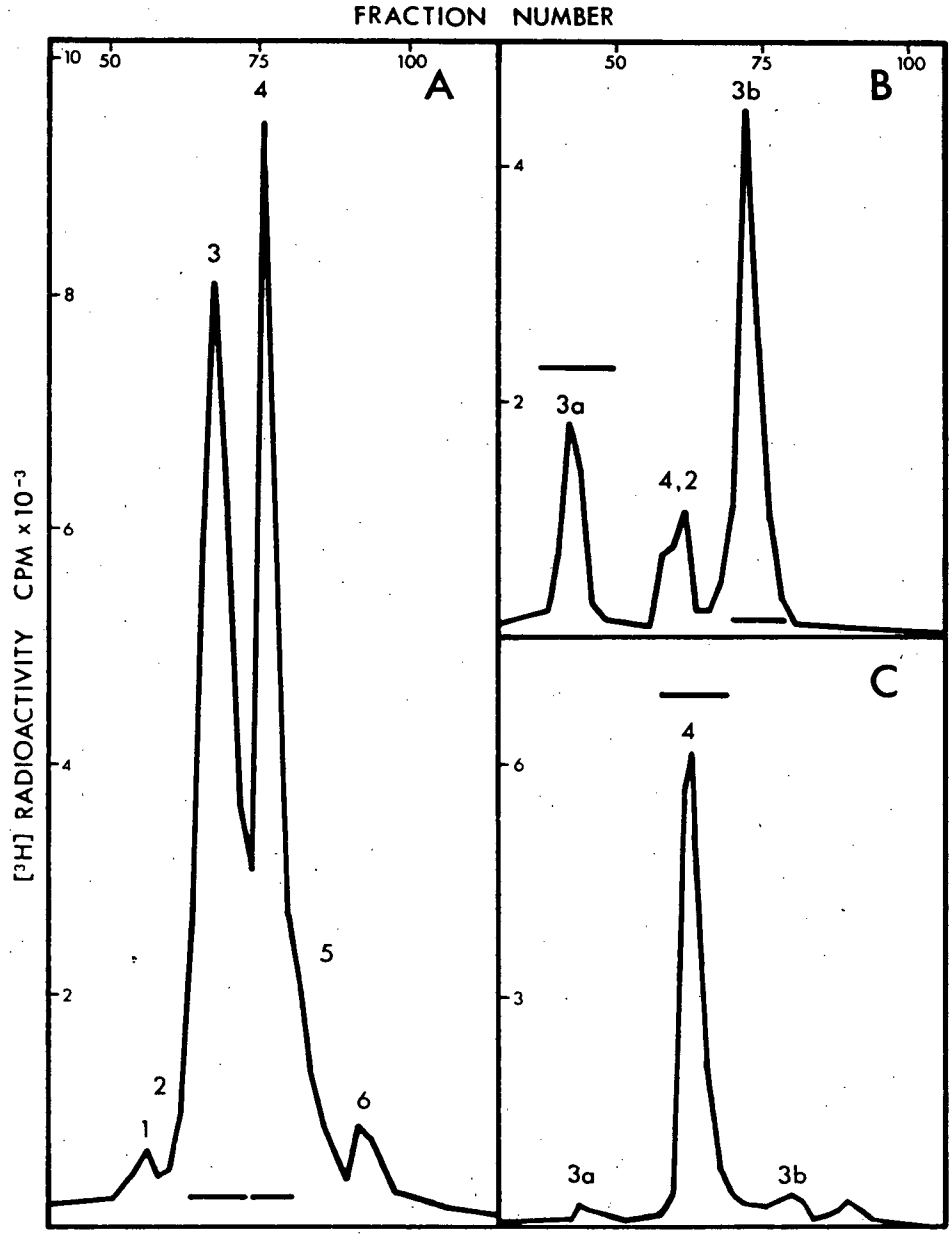
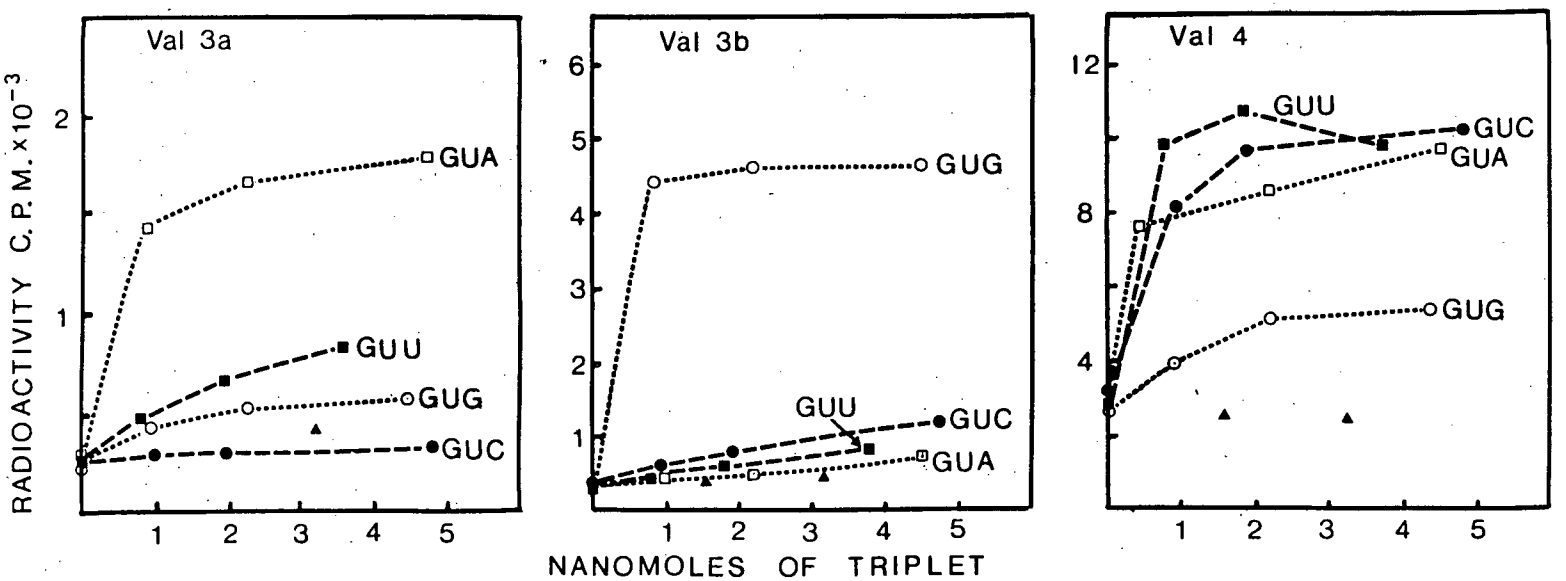

FRACTION NUMBER
[³H] RADIOACTIVITY CPM x 10⁻³
A
B
C
1
2
3
4
5
6
3a
4,2
3b

Figure 11. The coding properties of the Drosophila valine tRNAs.

Binding of Drosophila [^{3}H]-tRNA to E. coli ribosomes in the presence of triplets: GUU (■), GUC (●), GUA (□), GUG (o), and UCG (▲). Each 50 µl reaction mixture contained 20 mM magnesium acetate, 50 mM KCl, 0.1 M Tris-acetate pH 7.2, 2 A_{260} unit salt-washed ribosomes (1 A_{260} unit in the case of Val 3a) and [^{3}H]valyl-tRNA, i.e., Val 3a, 0.77 pmol; Val 3b, 2.2 pmol; and Val 4, 1.4 pmol. After 20 min at 25°C, the ribosome complexes were collected on Millipore filters, washed, dried, and the radioactivity determined in a liquid scintillation counter.


Val 3a
Val 3b
Val 4
GUA
GUU
GUG
GUC
RADIOACTIVITY C.P.M. x10⁻³
NANOMOLES OF TRIPLET

the organization of the genetic code, and structural considerations, Crick (1966) proposed the "wobble" hypothesis. Crick postulated that there is a certain amount of "play" or wobble possible in the base-pairing between the first ("wobble") nucleotide of the anticodon and the third nucleotide of the codon. This wobble allows some nucleotides in the first position of the anticodon to form unorthodox base-pairs with the codon. The wobble rules, as formulated by Crick, allow the following base-pairs:

1st position of anticodon		3rd position of codon
A	pairs with	U
C	" "	G
G	" "	C, U
U	" "	A, G
I	" "	A, C, U

The wobble rules allow predictions to be made about the anticodon sequences of the Drosophila valine tRNAs. Thus $tRNA_{3b}^{Val}$ and $tRNA_{4}^{Val}$ would be predicted to have the anticodons CAC and IAC respectively. The nucleotide sequences of these tRNAs, presented later in this study (Figure 12, Figure 20) confirm these predictions. The marked preference of $tRNA_{3a}^{Val}$ for the GUA triplet over the GUG triplet suggests that the wobble position of this tRNA contains a 2-thiouridine residue (or derivative)(Nishimura, 1979). However, $tRNA_{3a}^{Val}$ binds weakly to ribosomes in response to the GUU triplet, behavior not consistent with the presence of a thiouridine residue in the first position of the anticodon (Nishimura, 1979). The unidentified nucleoside N found in the nucleoside analysis of $tRNA_{3a}^{Val}$ (Table II) may be responsible for its somewhat unusual coding properties.

The coding properties and nucleotide sequences of $tRNA_{(IAC)}^{Val}$ from yeast, rabbit liver, rat liver and Drosophila are now known. Yeast $tRNA_{(IAC)}^{Val}$ has the coding properties predicted by the wobble hypothe-

sis (Mirzabekov et al., 1968). In contrast, rabbit liver $tRNA^{Val}_{(IAC)}$ binds most strongly to ribosomes in the presence of GUG, somewhat less strongly in the presence of GUU and weakly in the presence of the other two valine triplets (Jank et al., 1977a). Such base-pairing between an IAC anticodon and a GUG codon violates the wobble base-pairing rules. It could be argued that some unique structural feature of rabbit liver $tRNA^{Val}_{(IAC)}$ is responsible for its unusual coding properties. However, $tRNA^{Val}_{(IAC)}$ of rat ascites hepatoma cells, with a nucleotide sequence identical to $tRNA^{Val}_{(IAC)}$ of rabbit liver, binds to ribosomes solely in response to GUU (Shindo-Okada et al., 1981). The ribosome-binding assay conditions were very similar for the two mammalian $tRNA^{Val}_{(IAC)}$ species. Clearly further studies of these two tRNAs are required to reconcile these contradictory findings. Drosophila $tRNA^{Val}_{4}$ is quite similar to yeast $tRNA^{Val}_{(IAC)}$ in its coding properties but does show some ribosome-binding in the presence of the GUG triplets.

How closely do ribosome-binding studies of the codon-anticodon interaction reflect the interactions that occur in vivo? Mitra et al. (1977) attempted to answer this question by translating phage MS2 RNA in an E. coli protein synthesizing system in which protein synthesis was completely dependent on the addition of purified valyl-$tRNA^{Val}$ isoacceptors. The capacity of each isoacceptor to translate the valine codons of MS2 coat protein mRNA was determined. In this system all valine tRNAs (E. coli $tRNA^{Val}_{(UAC)}$, $tRNA^{Val}_{(GAC)}$ yeast $tRNA^{Val}_{(IAC)}$, Torulopsis utilis $tRNA^{Val}_{(IAC)}$) translated all valine codons. The authors concluded that, in their system, the codon-anticodon interaction was much less restrictive than ribosome-binding studies suggested. Holmes et al. (1978) showed that the error rate in a protein synthesizing system similar to that described above was very high with misreading occurring in 15-30% of the protein synthesized.

Further experiments under more physiological conditions indicated that codon anticodon base-pairing was less restrictive than predicted by the wobble hypothesis but much less indiscriminate than originally suggested by Mitra et al. (Mitra et al., 1979; Goldman et al., 1979: Lustig et al., 1981).

Genetic evidence suggests that, in vivo, codon-anticodon interactions are at least as restrictive as those predicted by the wobble hypothesis. E. coli contains 4 species of glycine tRNA. Ribosome-binding experiments showed that only $tRNA^{Gly}_{(UCC)}$ translates the GGA glycine codon. There is only 1 copy of the $tRNA^{Gly}_{(UCC)}$ gene in the E. coli genome. Therefore, elimination of this gene should be lethal to the cell. A mutation in which $tRNA^{Gly}_{(UCC)}$ has been changed into an AGA suppressor tRNA is indeed lethal in the haploid state (Murgola and Pagel, 1980). In Schizosaccharomyces pombe there are two genes for a $tRNA^{Ser}_{(UGA)}$ that translates the UCA codon. Munz et al. (1981) inactivated one of these genes in one strain of S. pombe and the other gene in another strain. When the two strains were crossed spores containing both of the inactivated genes were not viable. The cells did contain a $tRNA^{Ser}_{(IGA)}$ isoacceptor that should have been able, according to the wobble hypothesis, to translate the UCA codons in the absence of the $tRNA^{Ser}_{(UGA)}$ species. This suggests that inosine may not wobble base-pair with adenosine in vivo. Perhaps the recognition of NNA codons by INN anticodons, commonly observed in the ribosome-binding assay, is artifactual. The Mg^{++} concentration (20 mM) in the ribosome-binding assay mixture is much higher than the physiological concentration (4 mM in bacteria, Lusk et al., 1968; 1 mM in rat cells, Veloso et al., 1973). High Mg^{++} concentration is known to reduce the specificity of some anticodon-triplet interactions in the ribosome-binding assay (Rudloff and Hilse, 1975). The suggestion that inosine cannot wobble base-pair with adenosine in vivo is attractive. Drosophila $tRNA^{Val}_{4}$

would then translate only GUU and GUC codons, GUA codons would be translated solely by $tRNA_{3a}^{Val}$ and $tRNA_{3b}^{Val}$ would translate GUG codons. The amounts of the three major $tRNA^{Val}$ species in Drosophila (Figure 7) would also roughly parallel the frequency of valine codons in the mRNAs of animals (30 GUC + GUU:5 GUA:33 GUG, Grantham et al., 1981).

II. The Nucleotide Sequence of D. melanogaster $tRNA_4^{Val}$

The nucleotide sequence of $tRNA_4^{Val}$, determined in this study is shown in cloverleaf form in Figure 12. A combination of 3 different RNA sequencing techniques was used to determine the sequence. The results obtained by each of these techniques are described and discussed below.

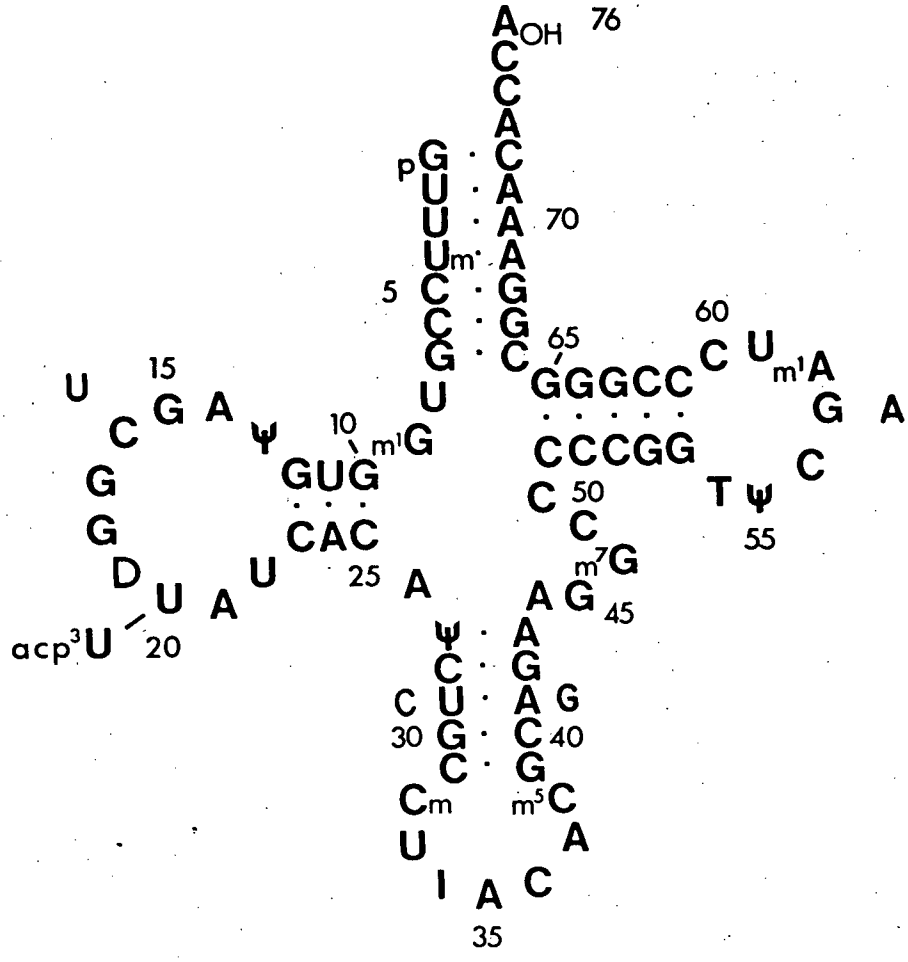
A. The Stanley and Vassilenko Method

Most of the tRNA sequence was obtained by the Stanley and Vassilenko method (nucleotides 1, 2, 8-32, 34-46 and 51-71). The [5'-^{32}P]nucleoside-5', 3'-bisphosphates (pNp's) produced by this technique were separated by TLC on PEI-cellulose ion-exchange plates and cellulose plates in Solvents A and B respectively. The pNp's were identified by comparing their chromatographic mobilities to published values (Silberklang et al., 1979; Cribbs, 1979). The sequence data obtained by the Stanley and Vassilenko method was, for almost all positions in the tRNA, unambiguous (Figure 13). Experience has shown that the most important factor in obtaining such results is high tRNA purity. Heterogeneity in the length of the tRNA, presumably due to hydrolysis during storage, manifests itself as a "messy" ladder of 5' end-labelled fragments. In such a ladder there are a great many bands, the bands are fuzzy and they vary greatly in intensity. To avoid these problems tRNA to be sequenced by the Stanley and Vassilenko method was routinely purified by gel electrophoresis just before use (see Materials and Methods).

Figure 12. The nucleotide sequence of D. melanogaster $tRNA_4^{Val}$ arranged as a cloverleaf.

The uridine residue at position 20 is partially modified to acp^3U. One of the cytidine residues between positions 47 and 50 is probably modified to m^5C. Transcription of the $tRNA_4^{Val}$-like genes of plasmids pDt92R, pDt120R and pDt14 would produce a tRNA identical to $tRNA_4^{Val}$ except for the replacement of C16, U29, A41 and G57 by U,C,G and A respectively.

D.m. $tRNA_4^{Val}$


A OH 76
C
C
A
C
pG
U
U
Um
5 C
C
G
C
70
65
60
C Um1A
G A
C
T Ψ
55
GGGCC
CCCGG
C 50
C
m7G
45
A G
A
Ψ
C
C U
30 G
C
Cm
m5C
A
U
I A C
35
15
U
C G A Ψ
G
G
D U A
acp3U 20
10 m1G
GUG
CAC
25 A
U

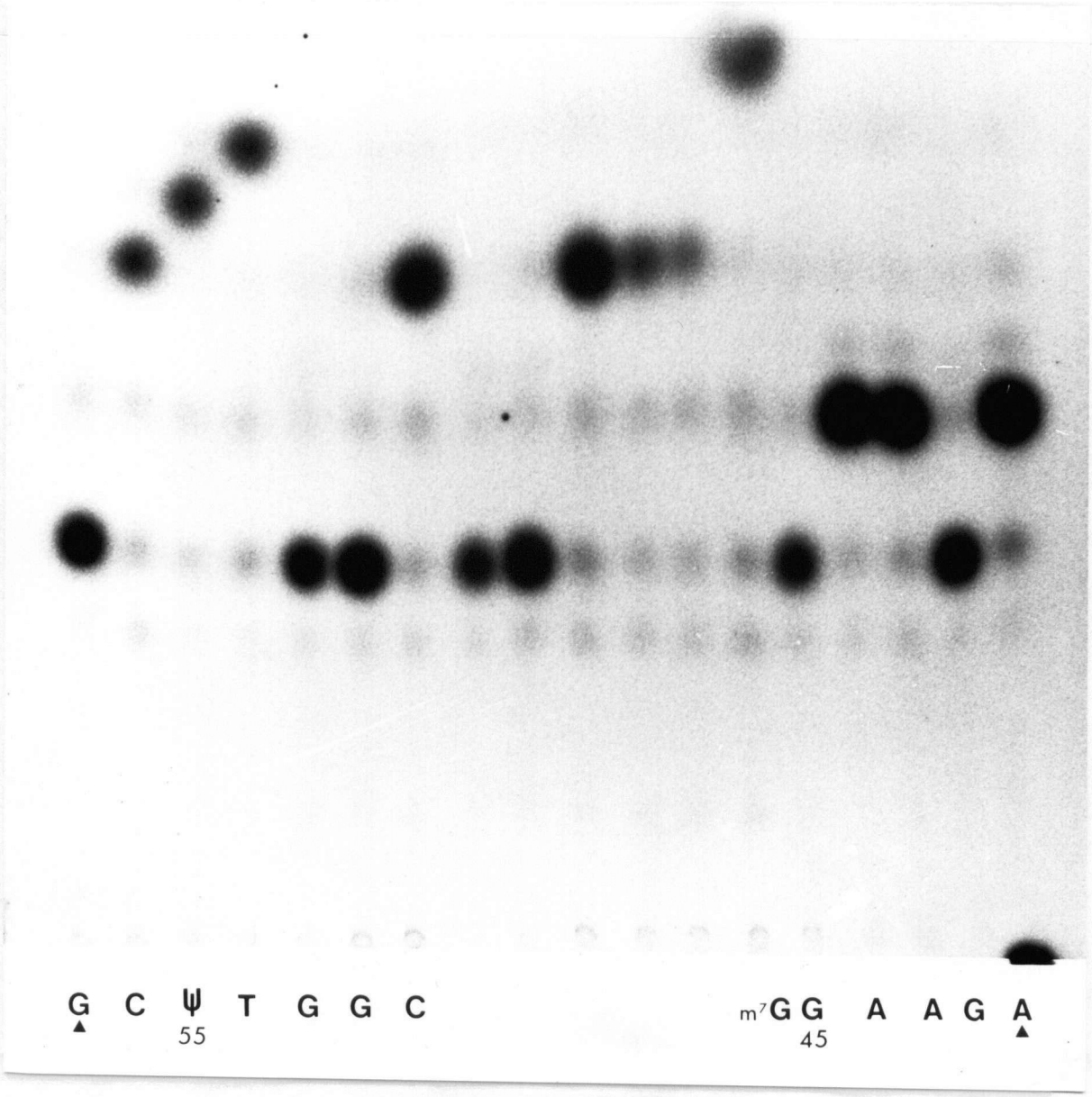
Stanley and Vassilenko sequencing is uniquely suited to determining the identity and position of modified nucleotides in tRNAs. Those pNp's that, because of their chromatographic behavior or position in the putative tRNA sequence, were suspected of being modified were characterized more fully by first converting them to [5'-^{32}P]nucleoside-5'-phosphates (pN's) using the 3'-phosphatase activity of nuclease P_1 then identifying them by their mobility on cellulose TLC plates developed in solvents B (Figure 14) and C.

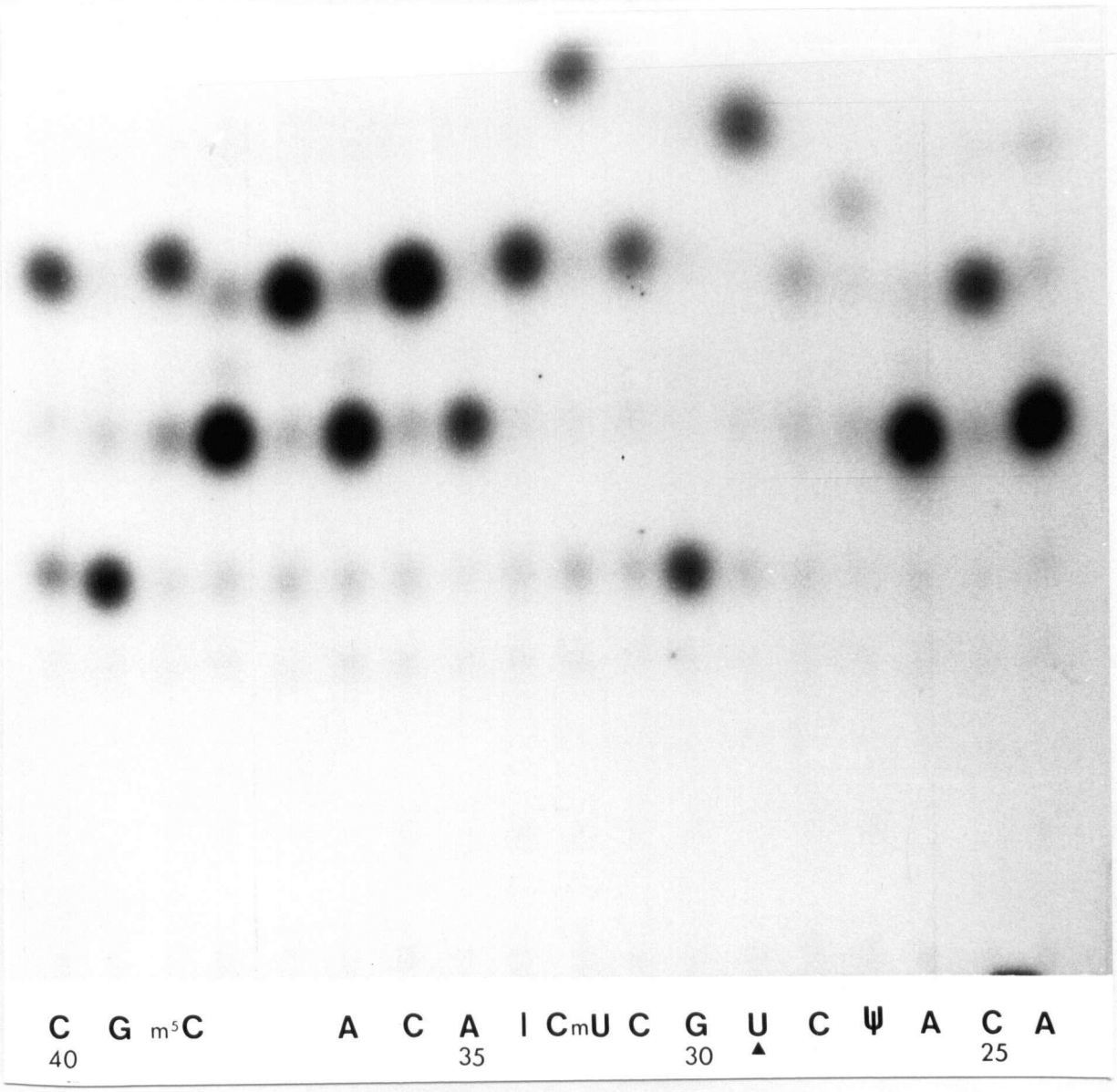
The identification of several modified nucleotides deserves special comment. The chromatographic mobility of the pNp from position 9 of $tRNA_4^{Val}$ (Solvent A, R_{pUp}= 0.83; Solvent B, R_{pAp}= 0.71) did not match any of the published mobility values for modified pNps. The sequence of [5'-^{32}P]endlabelled $tRNA_4^{Val}$ obtained by the gel read-off method showed that the RNA was slowly cleaved at nucleotide 9 by RNases T_1 and U_2. This suggested that nucleotide 9 was a modified G residue. On the basis of the nucleoside analysis (Table II) 1-methylguanosine (m^1G) was assigned to position 9. This assignment is consistent with the observation that m^1G is found only at position 9 in previously sequenced eukaryotic tRNAs (Dirheimer et al., 1979).

About 50% of the uridine residues at position 20 (Figure 13) are modified to 3-(3-amino-3-carboxypropyl) uridine (acp^3U). Authentic acp^3U-5'-phosphate, isolated from a nuclease P_1 digest of crude E. coli tRNA and characterized by its UV absorption spectrum, positive ninhydrin reaction and the chromatographic properties of the nucleoside (Ohashi et al., 1974), was used as a standard for identification of the modified nucleotide. The presence of acp^3U in $tRNA_4^{Val}$ is consistent with the work of White (1980) who has shown that Drosophila $tRNA^{Val}$ reacts with cyanogen bromide and the N-hydroxysuccinimide ester of naphthoxyacetic acid, reagents thought to react with the amino group of acp^3U in tRNAs. An acp^3U

Figure 13. Stanley and Vassilenko sequencing of $tRNA_4^{Val}$: PEI-cellulose chromatography of [5'-^{32}P]pNp's.

$tRNA_4^{Val}$ was sequenced by the Stanley and Vassilenko method as described in Materials and Methods. The resulting [5'-^{32}P]pNp's were identified by chromatography on PEI-cellulose TLC plates in Solvent A. The nucleotides at positions 57-51, 46-34 and 31-8 of $tRNA_4^{Val}$ were identified by this method. Positions at which the sequence of $tRNA_4^{Val}$ differs from a hypothetical transcript of the $tRNA_4^{Val}$-like genes of plasmids pDt92R, pDt120R and pDt14 are marked (▲). The A and C residues seen between m^5C38 and A37 are a repetition of A37 and C36. The order of G and C residues seen between C51 and m^7G46 is not known because of band compression in the ladder of 5' end-labelled fragments generated during sequencing by the Stanley and Vassilenko method.


G C Ψ T G G C
55
m7G G A A G A
45


C G m5C A C A I CmU C G U C Ψ A C A
40 35 30 25

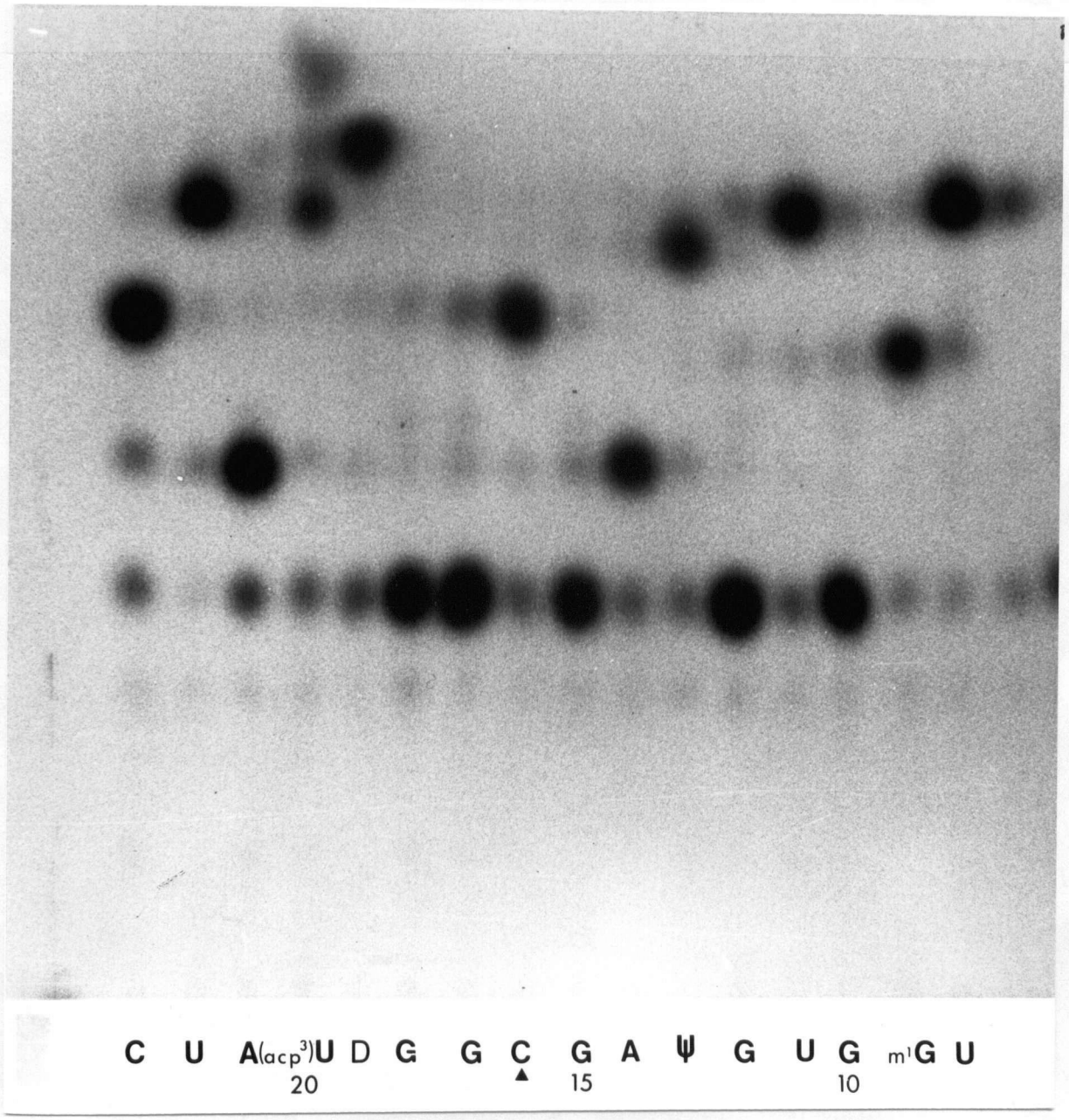
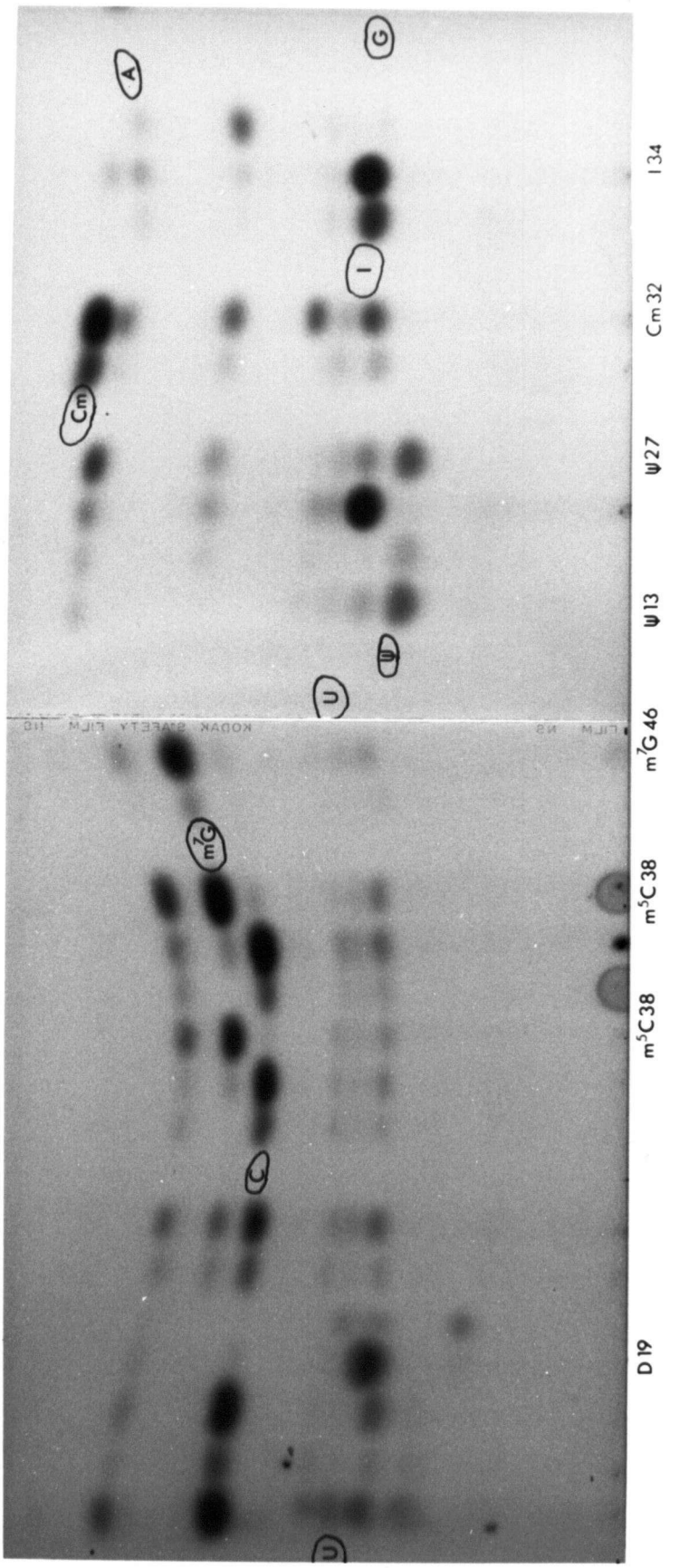

C U A(acp³)U D G G C G A Ψ G U G m¹G U
20
15
10

Figure 14. Stanely and Vassilenko sequencing of $tRNA_4^{Val}$: Identification of modified nucleotides by thin-layer chromatography.

The [5'-^{32}P]3',5'-bisphosphates of modified nucleotides produced by the Stanley and Vassilenko method were dephosphorylated to the corresponding [^{32}P]nucleoside-5'-phosphate by digestion with nuclease P_1 as described in Materials and Methods. These nucleotides were identified by chromatography on cellulose TLC plates developed in Solvent C. If commercially available, the 5'-phosphates of the appropriate modified nucleosides were chromatographed on the same TLC plates as standards. The positions of these standards after chromatography were detected by their UV absorbance and are circled in the autoradiograms shown in the Figure.


D 19
m5C 38
m5C 38
m5C 38
m7G 46
Ψ13
Ψ27
Cm 32
134
A
G
Cm
I
Ψ
U
m7G
C
U

residue, also designated nucleoside X, was detected in the nucleoside analysis of $tRNA_{3a}^{Val}$ but not in those of the other two major valine isoacceptors (Table II). The reason why acp^3U escaped detection is not apparent. It may have simply run off the end of the TLC plate used to separate the nucleosides during the analysis. Alternatively, the $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ used for nucleoside analysis may have contained significantly less acp^3U than the sequenced tRNAs.

An acp^3U residue is found at position 20 in the D-loop of both $tRNA_4^{Val}$ and rat liver $tRNA_{(QUU)}^{Asp}$, the only other sequenced eukaryotic tRNA that contains this nucleotide (Chen and Roe, 1978). tRNAs for the same amino acids contain acp^3U in both Drosophila and rat (Asp, Ile, Thr, Tyr and Val tRNAs; White, 1980). This distribution is quite different from that found among E. coli tRNAs (Lys, Phe, Val, Ile, Arg and Met tRNAs; Friedman, 1973). The position of acp^3U within E. coli tRNAs is also different from its position in the tRNAs of higher eukaryotes. In prokaryotes and plant chloroplasts acp^3U is found exclusively 3' to the 7-methylguanosine (m^7G) residue in the variable arm of those tRNAs that contain m^7G (Gauss and Sprinzl, 1981). acp^3U has not been detected in yeast tRNA (White, 1980).

In RNA the phosphodiester bond between a 2'-O-methylated nucleotide and its 3' neighbouring nucleotide is resistant to hydrolysis by alkali and most ribonucleases. This had two consequences for the sequencing of $tRNA_4^{Val}$ by the Stanley and Vassilenko method. First, since the bond between Cm32 and U33 was not cleaved during partial hydrolysis of the tRNA the nucleotide at position 33 was not labelled and could not, therefore, be identified by this sequencing method. Second, RNase T_2 digestion of that labelled fragment with Cm33 as its 5'-terminal nucleotide did not produce a pNp but a dinucleoside triphosphate, [5'-^{32}P]pCmpUp. Digestion of this nucleotide

with nuclease P_1, however, liberated [5'-^{32}P]pCm which could be easily identified (Figure 14).

Stanley and Vassilenko sequencing showed that Drosophila $tRNA_4^{Val}$ has a typical T-loop with rT at position 54 and a C residue at position 60. In contrast, the T-loops of mammalian $tRNA^{Val}_{(IAC)}$ are unique among sequenced cytoplasmic tRNAs in having a U residue at position 54 and an A at position 60 (Piper and Clark, 1974; Piper, 1975; Chen and Roe, 1977; Jank et al., 1977a; Shindo-Okada et al., 1981). These two nucleotides could base-pair to produce a tRNA with a 6 base-pair T-stem and a 5 nucleotide T-loop. Gross and his coworkers have produced some evidence favouring such a structure (Jank et al., 1977b). As previously described in this Discussion (Section I) mammalian $tRNA^{Val}_{(IAC)}$ has unusual coding properties while Drosophila $tRNA_4^{Val}$ has coding properties consistent with the wobble hypothesis. Comparison of the two tRNA structures suggests that the anomalous coding properties of the mammalian $tRNA^{Val}_{(IAC)}$ may be related to its unusual T-loop structure.

B. Wandering-Spot Analysis

Wandering-spot analysis of end-labelled $tRNA_4^{Val}$ was used to determine the sequences at the 5' and 3'-ends of the molecule (nucleotides 1-16, and 62-76). The wandering-spot analysis of the 5'-end of $tRNA_4^{Val}$ is presented in Figure 15. The 5'-terminal nucleotide was shown to be pG by the Stanley and Vassilenko method. The residue at position 4 is resistant to the formamide-induced hydrolysis used in the wandering-spot procedure. This would be expected if nucleotide 4 was 2'-O-methylated. Of the two ribose-methylated nucleosides detected in the nucleoside analysis (Table II) 2'-O-methylcytidine had been detected at position 32 by the Stanley and Vassilenko method. Therefore, the 2'-O-methyluridine (Um) residue was assigned to position 4. A Um residue at this position could hydrogen bond

Figure 15. Wandering-spot analysis of the 5'-terminal nucleotides of $tRNA_4^{Val}$.

$tRNA_4^{Val}$ was cleaved in the anticodon loop with RNase U_2 and 5' end-labelled with [^{32}P]phosphate as described in Materials and Methods. The labelled RNA fragment containing the 5'-end of $tRNA_4^{Val}$ (2.5 x 10^5 cpm Cerenkov) was dissolved in 5 µl of 98% formamide and heated to 100°C for 50 min. The partially hydrolysed RNA was subjected to wandering-spot analysis as described in Materials and Methods. In the first dimension (1), electrophoresis at pH 3.5, the xylene cyanol marker dye was allowed to migrate 6 cm. A "10 mM KOH" homomix prepared according to the procedure of Silberklang et al. (1979) was used for homochromatography in the second dimension (2). R and Y are the positions of the red and yellow marker dyes. The identification of the 2'-O-methyluridine residue at position 4 is discussed in the text.

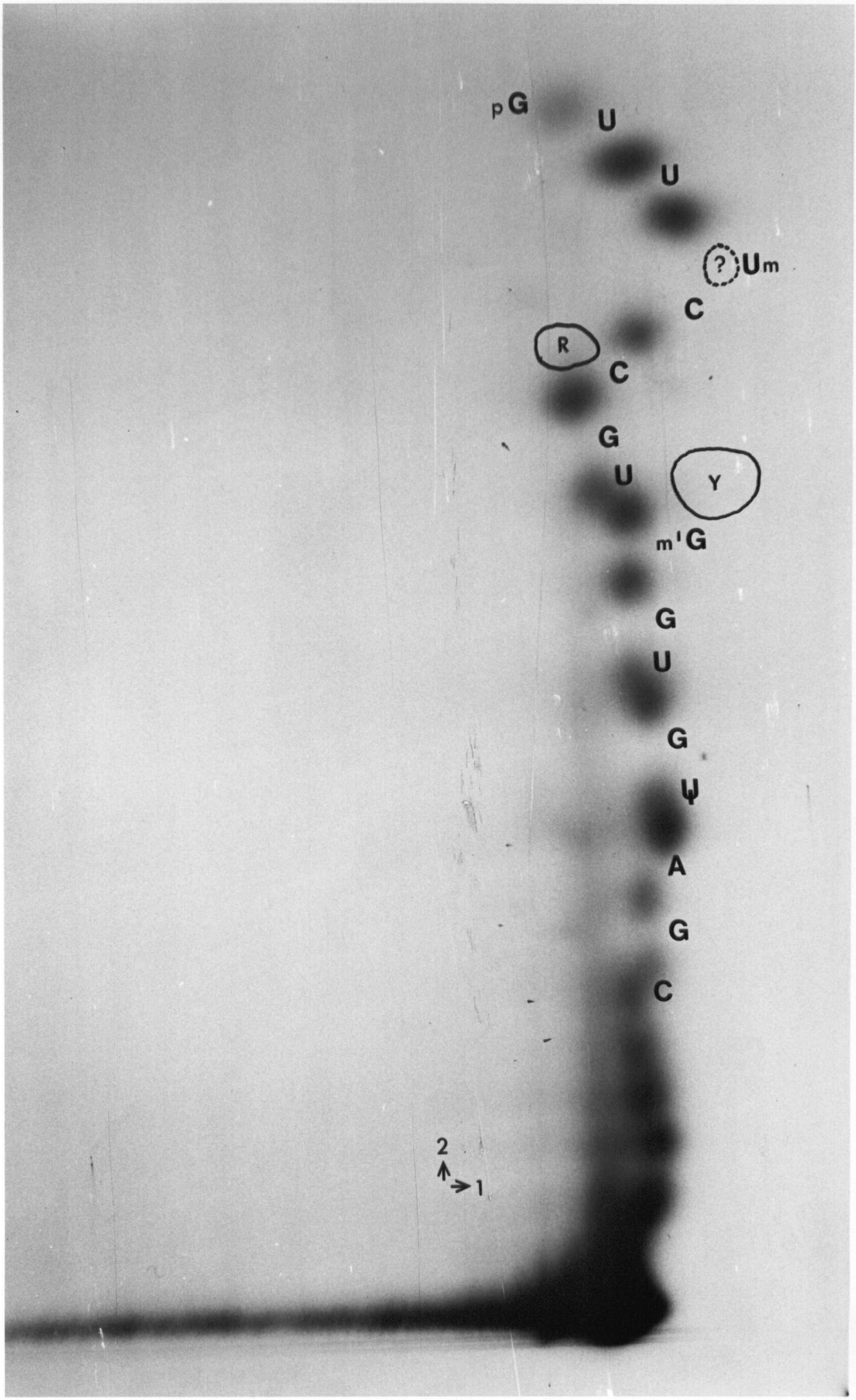
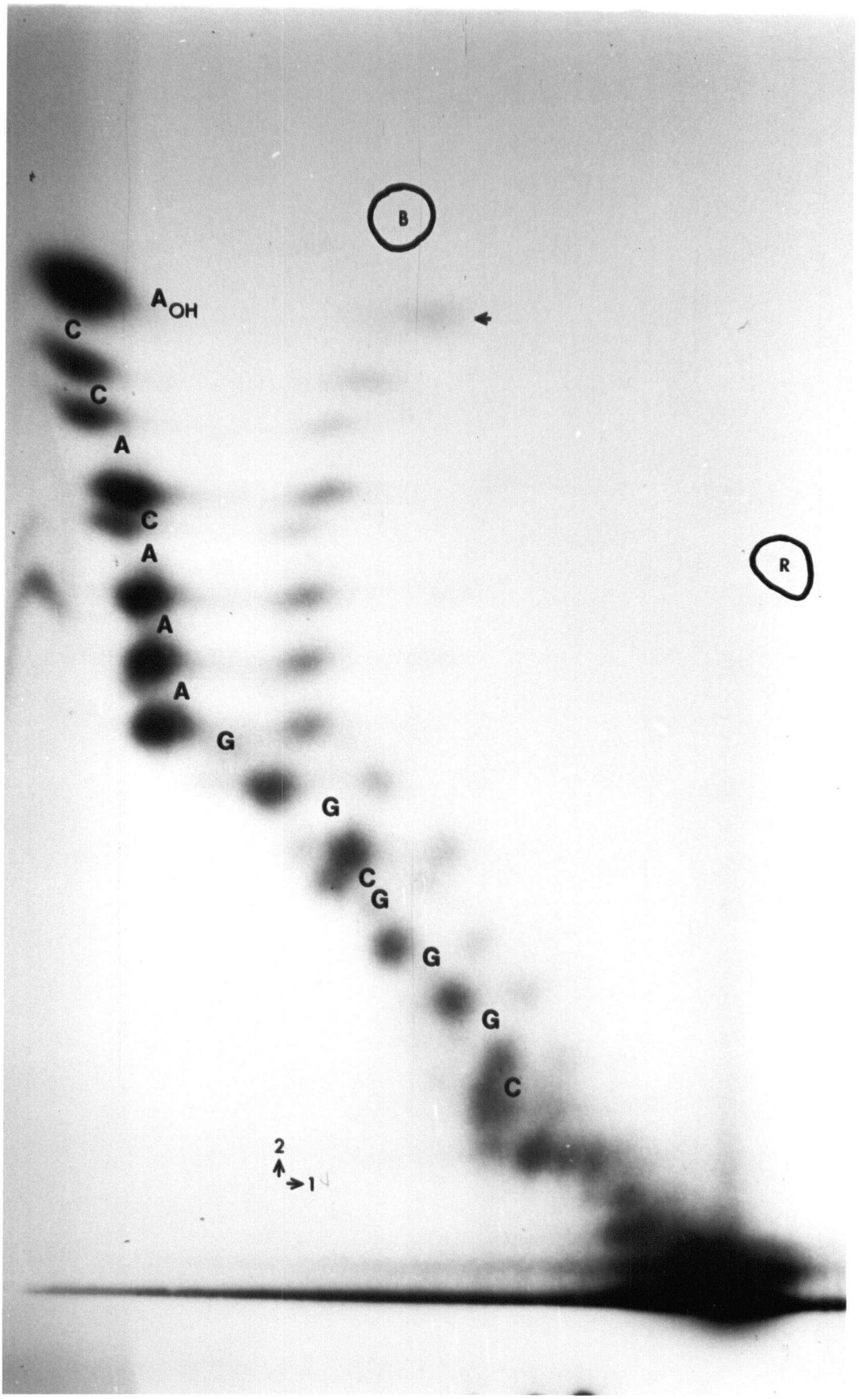

pG
U
U
?Um
C
R
C
G
U
Y
m'G
G
U
G
Ψ
A
G
C
2
1

Figure 16. Wandering-spot analysis of the 3'-terminal nucleotides of $tRNA_4^{Val}$.

$tRNA_4^{Val}$ was labelled at the 3'-end with [5'-^{32}P]pCp as described in Materials and Methods. The end-labelled RNA (4.4 x 10^4cpm Cerenkov) was dissolved in 5 µl 98% formamide and partially hydrolysed by heating the solution to 100°C for 1 h. The partial hydrolysate was subjected to wandering-spot analysis as described in Materials and Methods. The xylene cyanol marker dye (B) was allowed to migrate 8 cm during electrophoresis in the first dimension (1). "10 mM KOH" homomix (Silberklang et al., 1979) was used for homochromatography in the second dimension (2). R shows the position of the red marker dye.


B
R
A OH
C
C
A
C
A
A
A
G
G
C G
G
G
C
2
1

to the complementary A residue at position 69. Ribose-methylated nucleotides are present at position 4 of all sequenced glycine and proline tRNAs (Gauss and Sprinzl, 1981) and $tRNA^{Ser}_{4,7}$ of D. melanogaster (Cribbs, 1979). With these exceptions nucleotides with methylated base or sugar residues are very rare in the aminoacyl stem of tRNAs (Gauss and Sprinzl, 1981).

Figure 16 shows the wandering-spot analysis of the 3'-end of $tRNA^{Val}_{4}$. The interpretation of the mobility shifts was straightforward except for the presence of a faint pattern of spots "paralleling" the major one (indicated by the arrow in Figure 16). The nucleotide sequence inferred from either pattern is the same. The weaker spots display much greater mobility in the first dimension (electrophoresis at pH 3.5) than the corresponding major spots but exactly the same mobility in the second dimension (homochromatography). For wandering spot analysis the tRNA was end-labelled with [5'-^{32}P]pCp using phage T4 RNA ligase. The series of minor spots are probably the result of a small amount of the RNA being labelled with a [5'-^{3}P]pUp contaminant in the [5'-^{32}P]pCp used to label the tRNA. Replacement of 3'-terminal pCp by pUp would increase the negative charge at pH 3.5 of otherwise identical oligonucleotides and hence increase their electrophoretic mobility. For example, at pH 3.5 ApCp is calculated to have a charge of -0.52 while ApUp would have a charge of -1.36 (Brown, 1979).

C. Sequencing by the Gel Read-Off Method: Chloroacetaldehyde Modification as an Aid to RNA Sequencing

The gel read-off method was used to determine the nucleotide sequence of regions of the tRNA that could not be sequenced by the other two methods (nucleotides 33, 46-52). Chloroacetaldehyde modification of the tRNA facilitated sequencing by this method.

At pH 4.0 chloroacetaldehyde reacts with nonbase-paired cytidine and adenosine residues in nucleic acids to form fluorescent etheno derivatives. These derivatives cannot form Watson-Crick base-pairs with other nucleotides (Kochetkov et al., 1971; Barrio et al., 1972; Kimura et al., 1977; Biernat et al., 1978). The currently accepted mechanism for the reaction of chloracetaldehyde with cytidine residues is presented in Figure 17 (Krzyzosiak et al., 1979).

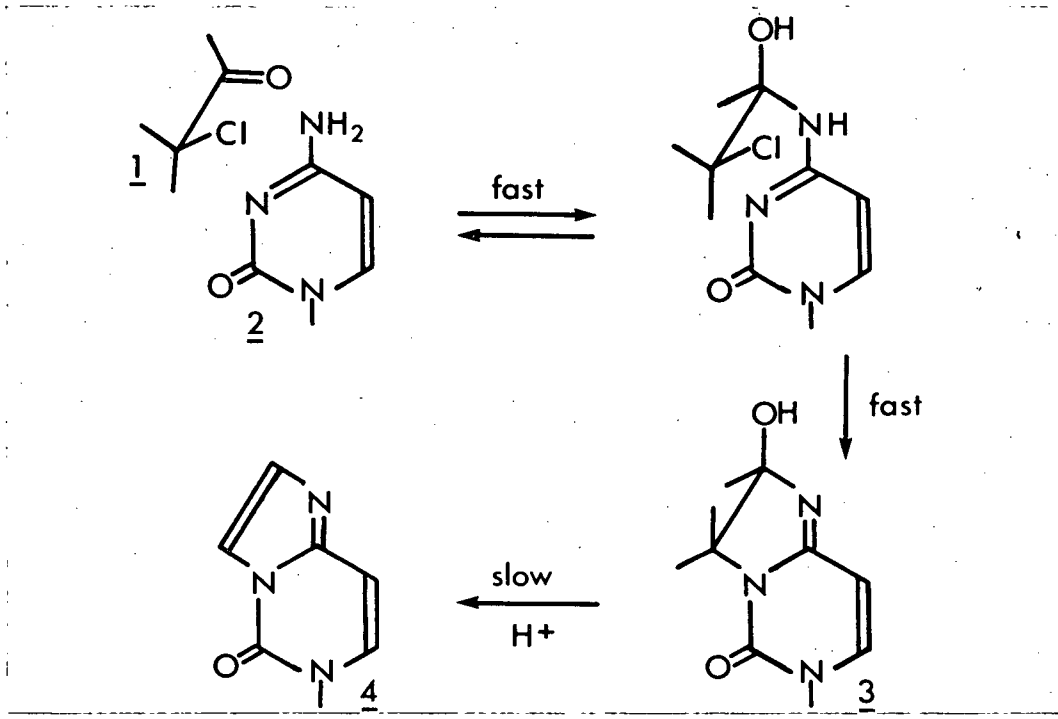


Figure 17. Reaction of Chloroacetaldehyde with Cytidine - Compounds 1, 2, 3 and 4 are respectively chloroacetaldehyde, cytidine, hydroxyethnocytidine and ethenocytidine. The rate-limiting dehydration step is acid catalysed.

Adenosine reacts with chloroacetaldehyde in an analogous manner. Thus if a tRNA were denatured and then modified with chloroacetaldehyde it should lose much of its secondary structure. In the technique developed in this study end-labelled tRNA is heated to 100°C at pH 4.0 with a 6.6% chloroacetaldehyde solution (v/v) for 1 min. The reaction mixture is greatly diluted with buffer (pH 4.0) and incubated at 80°C for 20 min. The modified tRNA is then sequenced by the gel read-off method. If the 80°C post-treatment is omitted

very fuzzy bands are seen on autoradiographs of the sequencing gels. Krzyzosiak et al. (1981) have recently shown that such a "maturation" step is required for the dehydration of the stable hydroxyetheno reaction intermediate to form the etheno derivative.

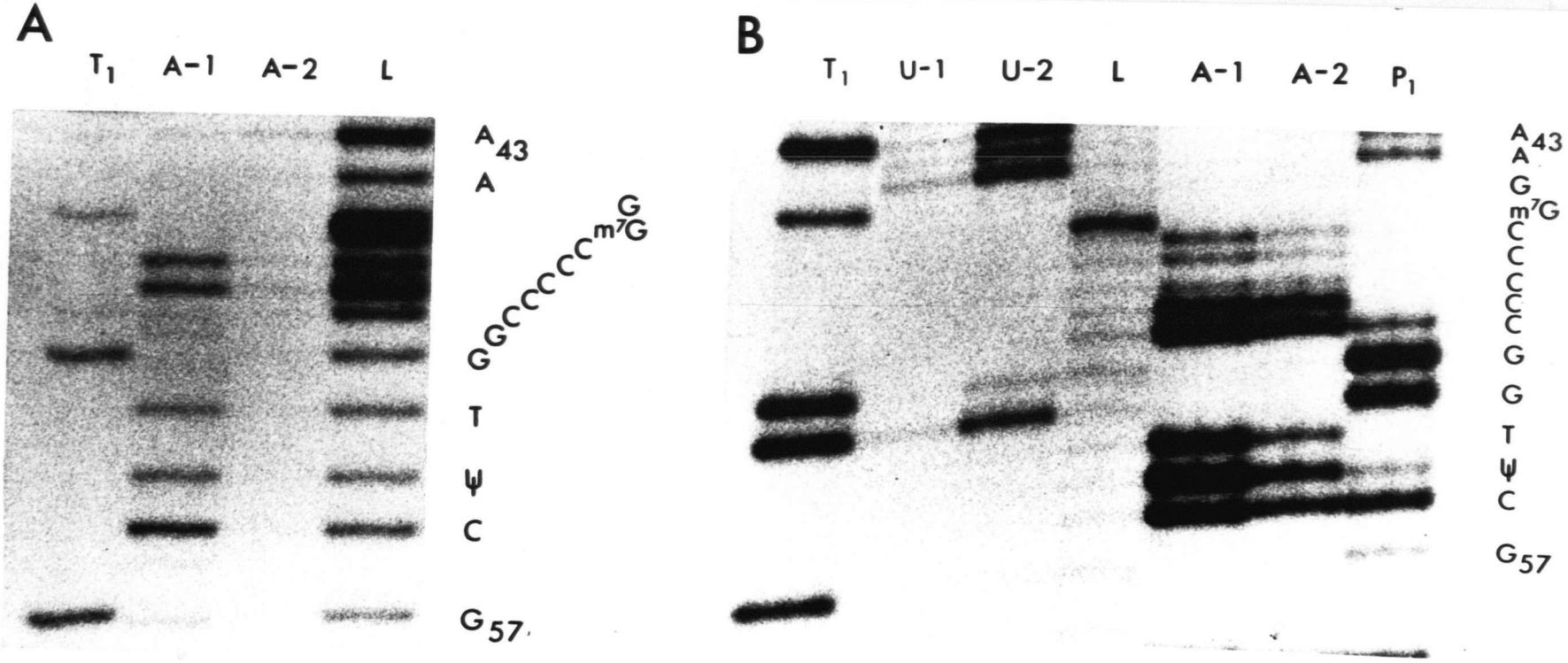
The series of five cytidine residues at positions 47-51 (variable loop, T-stem regions) could not be determined by either the gel read-off or Stanley and Vassilenko methods. Electrophoresis of partial enzymatic or formamide digests of $tRNA_4^{Val}$ on denaturing polyacrylamide gels exhibited both strong band compression and incomplete enzymatic cleavage in the variable loop and T-stem regions. In yeast $tRNA^{Phe}$ these regions are involved in strong secondary and tertiary interactions with other parts of the molecule (Rich and RajBhandary, 1976). Similar interactions in $tRNA_4^{Val}$ may be particularly strong because of the high G-C content of these regions. The observed band compression indicates that the products of such RNase cleavage as does occur fold back on themselves and migrate anomalously during polyacrylamide gel electrophoresis.

Figure 18 compares the results obtained when unmodified $tRNA_4^{Val}$ and chloroacetaldehyde modified $tRNA_4^{Val}$ are sequenced by the gel read-off method. The severe band compression and inadequate enzymatic cleavage observed when intact $tRNA_4^{Val}$ was the substrate (Figure 18A) was relieved when the RNA was modified in this way before sequencing (Figure 18B). The ribonucleases used for sequencing display the same substrate specificity towards modified RNA as to the unmodified form. This is surprising considering the strict requirement of RNase A for an unsubstituted nitrogen at position 3 of pyrimidines in its substrate (Richards and Wyckoff, 1971). Thus, the ethenocytidine residues produced by modification with chloroacetaldehyde would not be expected to be sites for RNase A cleavage. Tolman et al. (1974) modified dinucleoside monophosphates containing cytidine with chloroacetalde-

Figure 18. Chloroacetaldehyde modification relieves band compression on sequencing gels of the A43-G57 region of $tRNA_4^{Val}$.

A. $tRNA_4^{Val}$ was labelled at the 3'-end with [α-^{32}P]ATP and tRNA nucleotidyl transferase (Materials and Methods). The end-labelled tRNA was sequenced by the gel read-off method (Materials and Methods). 10 µl reaction volumes contained: T_1 - 0.1 unit RNase T_1; A-1 - 0.01 unit RNase A; A-2 - 0.005 unit RNase A; L - 66% formamide, 100°C for 60 min. The sequence between positions 45-52 could not be read because of band compression.

B. $tRNA_4^{Val}$ was labelled at the 3'-end as above. The end-labelled tRNA was modified with chloroacetaldehyde (Materials and Methods) and sequenced by the gel read-off method. 10 µl reaction volumes contained: T_1 - 0.01 unit RNase T_1, U-1 - 0.01 unit RNase U_2; U-2 - 0.001 unit RNase U_2; L - 66% formamide, 100°C for 45 min; A-1 - 0.05 unit RNase A; A-2 - 0.01 unit RNase A; P_1 - 1 unit RNase Phy I. The nucleotide sequence between positions 45 and 52 could be read when chloroacetaldehyde-modified $tRNA_4^{Val}$ was used as the substrate for the sequencing reactions.


A
T1
A-1
A-2
L
A43
A
G
m7G
CCCCC
G
T
Ψ
C
G57
B
T1
U-1
U-2
L
A-1
A-2
P1
A43
A
G
m7G
CCCCCC
G
G
T
Ψ
C
G57

hyde and examined the sensitivity of the etheno derivatives to hydrolysis by RNase A. They found ethenocytidylyl uridine (εCpU) and εCpεA to be completely resistant to hydrolysis while εCpG and εCpεC showed slight hydrolysis after prolonged RNase A treatment. It is possible, therefore, that RNase A recognizes the modified cytidine residues but it is more likely that modification was not complete and that unmodified cytidine residues were the sites of RNase A cleavage.

The nucleoside analysis (Table II) suggests that $tRNA_4^{Val}$ contains two 5-methylcytidine (m^5C) residues per molecule. One of these was located at position 38; the other is assigned to a probable position at residue 48 or 49 since m^5C has been found in this region of other eukaryotic valine tRNAs (Gauss and Sprinzl, 1981). The bands corresponding to C_{49} in Figure 18B move anomalously fast compared to the surrounding cytidine bands. This suggests modification of C_{49}.

Ribonuclease Phy I is often assumed to hydrolyse all phosphodiester bonds in RNA except those 3' to a C residue. Figure 18B shows that RNase Phy I cleaves the C-G bond between residues 51 and 52 and residues 56 and 57. The work of Pilly et al. (1978) showed that the rate at which RNase Phy I cleaves a phosphodiester bond depends on the nucleotides both 5' and 3' to the bond. For the 5'-nucleotide the rate of cleavage decreases in the order U > G = A > C. For the 3'-nucleotide the order of preference is A > G = C > U. Thus UpA should be most susceptible to cleavage by RNase Phy I and CpU most resistant. Other phosphodiester bonds should be of intermediate sensitivity. Of the four CpN phosphodiester bonds, only CpG was cleaved under the partial digestion conditions used for sequencing by the gel read-off method. This made distinguishing UpG from CpG difficult. Bonds 3' to modified uridine residues such as rT and Ψ were poorly cleaved by RNase Phy I and these residues could, therefore, be mistaken for C residues.

Nucleotide 33 of $tRNA_4^{Val}$ (Figure 12) is flanked on the 5'-side by a 2'-O-methylcytidine residue. The phosphodiester bond between these two nucleotides is resistant to hydrolysis, hence nucleotide 33 could not be identified by the Stanley and Vassilenko method. $tRNA_4^{Val}$ was modified with chloroacetaldehyde, labelled at the 5'-end with [^{32}P]phosphate and sequenced by the gel read-off method. Figure 19 shows that position 33 is occupied by a U residue. With the exception of the initiator $tRNA_i^{Met}$ of higher eukaryotes, all sequenced tRNAs have a U at this position (Gauss and Sprinzl, 1981). The tRNA was modified with chloroacetaldehyde not because this facilitated interpretation of the sequencing gel but because modification made efficient labelling of the 5-end of the molecule possible.

D. Sequencing tRNA End-Labelled with ^{125}I-CMP

$tRNA_4^{Val}$, labelled at the 3'-end with ^{125}I-CMP using the enzyme tRNA nucleotidyl transferase (see Materials and Methods), was sequenced by the gel read-off method. The results were inferior to those obtained with tRNA end-labelled in a similar manner with [α-^{32}P]AMP residues because of the wide, fuzzy bands produced on autoradiograms of sequencing gels when an ^{125}I label was used.

E. Homologies Between $tRNA_4^{Val}$ and Other Valine tRNAs

The homology between $tRNA_4^{Val}$ and 18 published $tRNA^{Val}$ or $tRNA^{Val}$ gene sequences is presented in Table IV. It should be remembered that, because of the invariant and highly conserved nucleotides present in all tRNAs, unrelated tRNAs display about 40% homology to each other. $tRNA_4^{Val}$ displays the greatest homology (85%) to the mammalian valine tRNAs. In this $tRNA_4^{Val}$ is unexceptional, homologies between Drosophila tRNAs and the corresponding tRNAs of mammals are high, ranging from 100% for $tRNA_2^{Lys}$ of Drosophila and rabbit liver to 92% for $tRNA_i^{Met}$ of mammals (Gauss and Sprinzl, 1981). $tRNA_4^{Val}$ shows an interesting pattern of homology with

Figure 19. Identification of the nucleotide present at position 33 of $tRNA_4^{Val}$.

$tRNA_4^{Val}$ (7 µg) was modified with chloroacetaldehyde as described in Materials and Methods. The modified tRNA was dephosphorylated then 5' end-labelled with [^{32}P]phosphate by the procedures of Silberklang et al. (1979). Chloroacetaldehyde-modified tRNA was labelled about 5 times more efficiently then unmodified tRNA by this procedure. The 5' end-labelled $tRNA_4^{Val}$ was sequenced by the gel read-off method as described in Materials and Methods. 10 µl reaction volumes contained: U-1 - 0.005 unit RNase U_2; U-2 - 0.001 unit RNase U_2; T-1 - 0.01 unit RNase T_1; T-2 - 0.001 unit RNase T_1; L - 66% formamide, 100°C for 45 min; A-1 - 0.01 unit RNase A; A-2 - 0.001 unit RNase A; P-1 - 1 unit RNase Phy I; P-2 - 0.1 unit RNase Phy I. Nucleotide 33 is a U residue. The cause of the band present in the A-1 slot between positions 34 and 35 is not known.

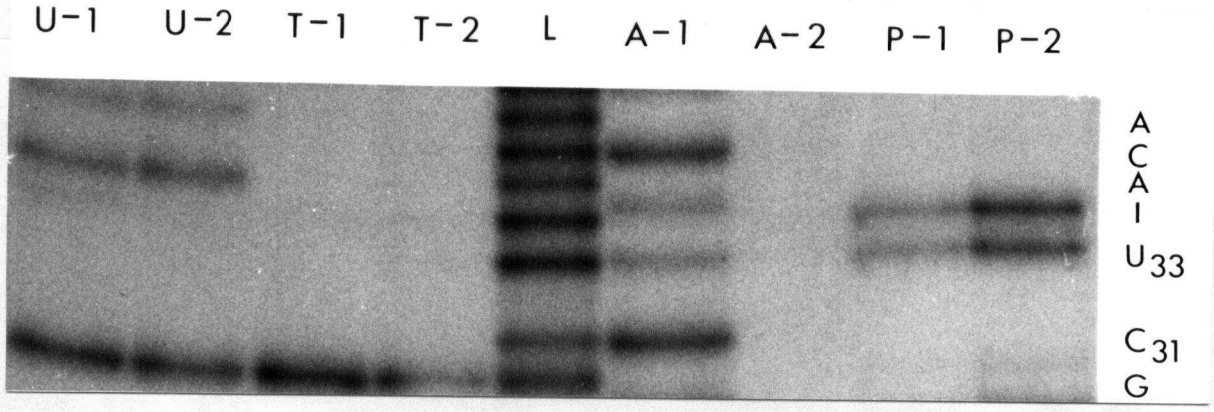

U-1
U-2
T-1
T-2
L
A-1
A-2
P-1
P-2
A
C
A
I
U33
C31
G

Table IV. Homology Between $tRNA_4^{Val}$ and Other Sequenced Valine tRNAs and Valine tRNA Genes

Source	Valine Anticodon	Percent Homology	Reference
D. melanogaster	IAC	100	-
mammalian	IAC	85	1
human placenta	CAC	85	1
yeast	IAC	81	1
yeast	CAC	70	1
yeast	UAC	61	1
T. utilis	IAC	79	1
B. stearothermophilus	GAC	60	1
B. subtilis	UAC	57	1
E. coli	UAC	57	1
E. coli 2a*	GAC	56	1
E. coli 2b	GAC	49	1
spinich chloroplast	UAC	48	2
maize chloroplast	UAC	41	3
N. crassa mitochondria	UAC	44	1
yeast mitochondria	UAC	44	4
mouse mitochondria	UAC	42	5
A. nidulans mitochondria	UAC	36	6
human mitochondria	UAC	37	7

* *E. coli* contains two valine tRNAs, $tRNA_{2a}^{Val}$ and $tRNA_{2b}^{Val}$, with the same GAC anticodon.

References

1. Gauss and Sprinzl, 1981
2. Sprouse et al., 1981
3. Schwarz et al., 1981
4. Li and Tzagoloff, 1979
5. Van Etten et al., 1980
6. Kochel et al., 1981
7. Anderson et al., 1981

the three yeast $tRNA^{Val}$ isoacceptors. There is strong homology (81%) to the yeast isoacceptor with the same IAC anticodon as $tRNA_4^{Val}$. Homology is significantly weaker between $tRNA_4^{Val}$ and yeast $tRNA^{Val}_{(CAC)}$ (70%). Yeast $tRNA^{Val}_{(UAC)}$ displays only 61% homology to $tRNA_4^{Val}$, about the same level of homology as is found between $tRNA_4^{Val}$ and most bacterial $tRNA^{Val}$ species. $tRNA_4^{Val}$ displays the least homology to $tRNA^{Val}$ of mitochondria and chloroplasts (36-48%). Examination of the sites of homology among the sequenced valine tRNAs shows a nonrandom distribution of these sites. There is considerably more homology among the loop regions of these tRNAs than in the stem regions.

Eighteen valine tRNA species or their genes have been sequenced to date. Inspection of these sequences reveals a number of sites occupied by the same nucleotides in all valine tRNAs or among the members of a particular subset of these sequences (in addition, of course, to the invariant nucleotides found in all tRNAs). Some of these similarities have been previously noted by Sprouse et al. (1981). All sequenced valine tRNAs contain an A73 residue and all, with the exception of maize chloroplast $tRNA^{Val}$, contain a G10-C25 base-pair. All the non-organelle $tRNA^{Val}$ contain G1-C72 and C31-G39 base-pairs. The conserved Y11-R24 base-pair (Figure 1) is occupied by a U-A pair in all sequenced eukaryotic $tRNAs^{Val}$ and by a C-G pair in the prokaryotic and organelle tRNAs. Similarly, all eukaryotic valine tRNAs have a G12-C23 base-pair while the prokaryotic and organelle tRNAs have a U-A base-pair at these positions. The last nucleotide in the anticodon loop is C38 in all eukaryotic, and with the exception of maize chloroplast $tRNA^{Val}$, all organelle tRNAs. No sequenced prokaryotic $tRNA^{Val}$ has a C residue at this position. All sequenced eukaryotic $tRNA^{Val}$ contain a U3-A70 and Ψ27-A43 base-pairs and an A residue at position 44. Eukaryotic $tRNAs^{Val}$ also contain C49-G65 and C 50-G64 base pairs. In bacteria, with the exception of B. stearothermophilus

$tRNA^{Val}$, these positions are occupied by G49-C65 and G50-C64 pairs. The conserved nucleotides described above probably play important roles in the structure, function and recognition of valine tRNAs. The Drosophila valine tRNAs can be efficiently aminoacylated by E. coli valyl-tRNA synthetase (Dr. I.C. Gillam, personal communication). Therefore, those conserved nucleotides found exclusively among eukaryotic valine tRNAs are probably not involved in the recognition of $tRNA^{Val}$ by its synthetase.

III. The Nucleotide Sequence of D. melanogaster $tRNA_{3b}^{Val}$

A. Nucleotide Sequence Determination

The nucleotide sequence of Drosophila $tRNA_{3b}^{Val}$ is presented in Figure 20. The strategy used to sequence $tRNA_4^{Val}$ was applied successfully to the sequence determination of $tRNA_{3b}^{Val}$. Again the Stanley and Vassilenko method provided most of the sequence information (nucleotides 9-45, 51-69). No single Stanley and Vassilenko experiment allowed the nucleotides at all these sites to be identified. The results obtained by this method and displayed in Figure 21 are, therefore, drawn from several different experiments. In sequencing $tRNA_{3b}^{Val}$ a slight modification was made in the Stanley and Vassilenko procedure. Solvent D (2-propanol:12 N HCl:H_2O, 70:15:15) was used for the separation of labelled nucleoside-5'-phosphates on cellulose TLC plates. This solvent proved particularly useful for the identification of modified pyrimidine nucleotides. Table V lists the relative mobility value of several nucleotides in this solvent system.

The 5' and 3'-ends of $tRNA_{3b}^{Val}$ were sequenced by wandering spot analysis. The results of these experiments are shown in Figures 22 and 23. The 3'-terminal nucleotide of $tRNA_{3b}^{Val}$ was not determined. However, it is undoubtedly an A residue since, without exception, all sequenced tRNAs have a 3'-terminal -CCA sequence (Gauss and Sprinzl, 1981).

Figure 20. The nucleotide sequence of D. melanogaster $tRNA_{3b}^{Val1}$ arranged as a cloverleaf.

The C residue at position 48 is probably modified to m^5C. Transcription of the $tRNA_{3b}^{Val1}$-like genes of plasmids pDt48 and pDt41R would produce a tRNA identical to $tRNA_{3b}^{Val1}$ except for the replacement of C5, C16, G68 and G69 by U, U, A and A respectively.

D.m. $tRNA^{Val}_{3b}$

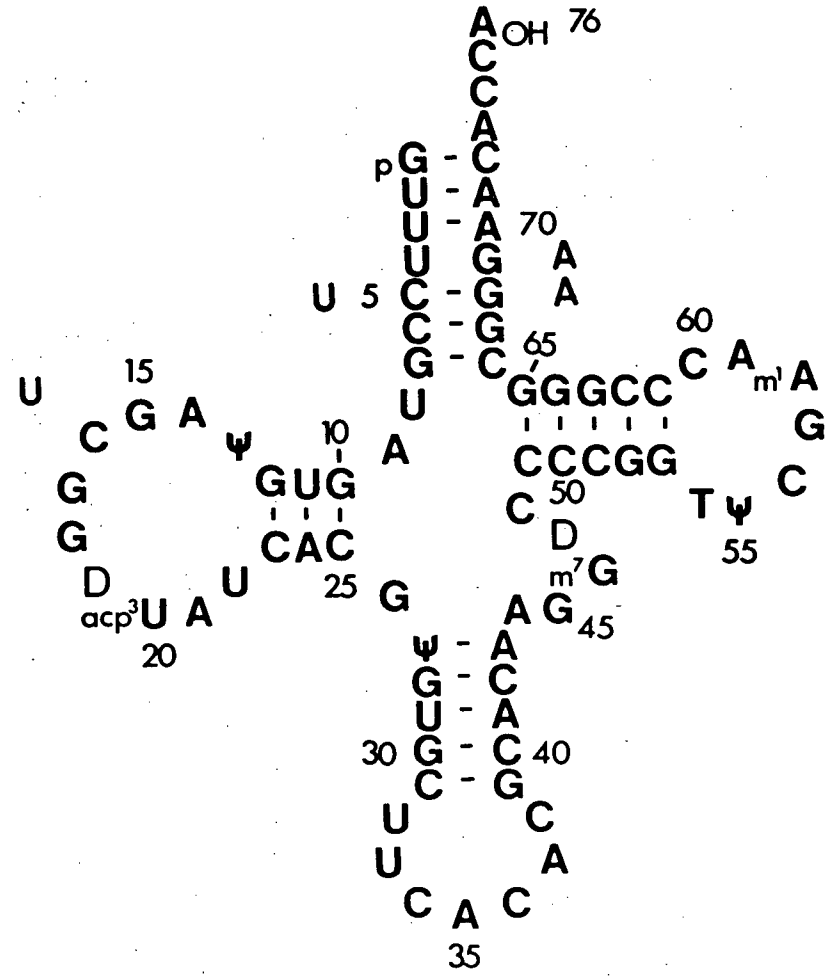


Figure 21. Stanley and Vassilenko sequencing of $tRNA_{3b}^{Val}$.

$tRNA_{3b}^{Val}$ (2 µg) was dissolved in 10 µl 10 mM NH_4OAc pH 4.5, 1 mM EDTA and the solution was heated to 100°C for 3 min. The RNA fragments produced by this limited hydrolysis were 5' end-labelled and fractionated according to size by electrophoresis on a 20% polyacrylamide gel as described in Materials and Methods. After autoradiography of the gel, bands containing labelled RNA fragments were excised from the gel and the RNA was separately eluted from each band and precipitated with ethanol (Materials and Methods). The eluted RNAs were each dissolved in 5 µl 10 mM NH_4OAc pH 4.5 containing 1 µg of nuclease P_1. After incubation at 37°C for 14 h the resulting [^{32}P]nucleoside-5'-phosphates (pN's) were identified by thin layer chromatography. In the Figure, positions at which the sequence of $tRNA_{3b}^{Val}$ differs from a hypothetical transcript of plasmids pDt48 and pDt41R are marked (▲). The origins of the chromatograms are also marked (O).

A. Chromatography of pN's from positions 35-37 and 58-69 on cellulose TLC plates developed in Solvent B (Materials and Methods). The positions of pG, pA and pU standards on the thin layer plate were detected by their UV absorbance and are circled on the autoradiogram.

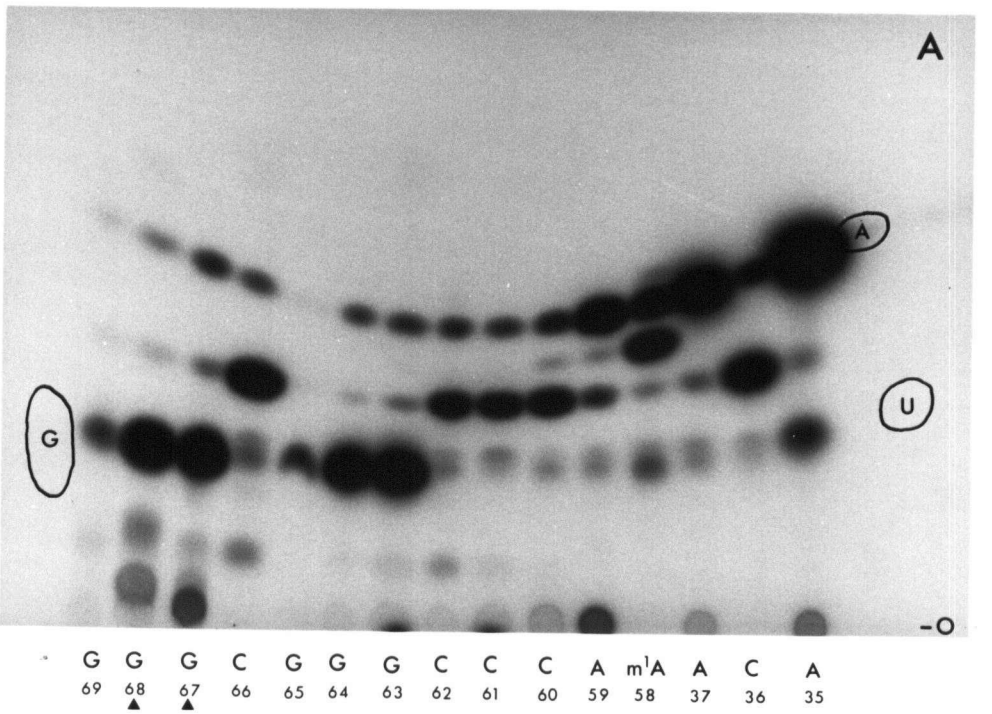
B. Chromatography of pN's from positions 13, 18, 19, 27 and 51-56 on cellulose TLC plates developed in Solvent D. The position of nucleotide standards on the TLC plate are circled on the plates's autoradiogram.

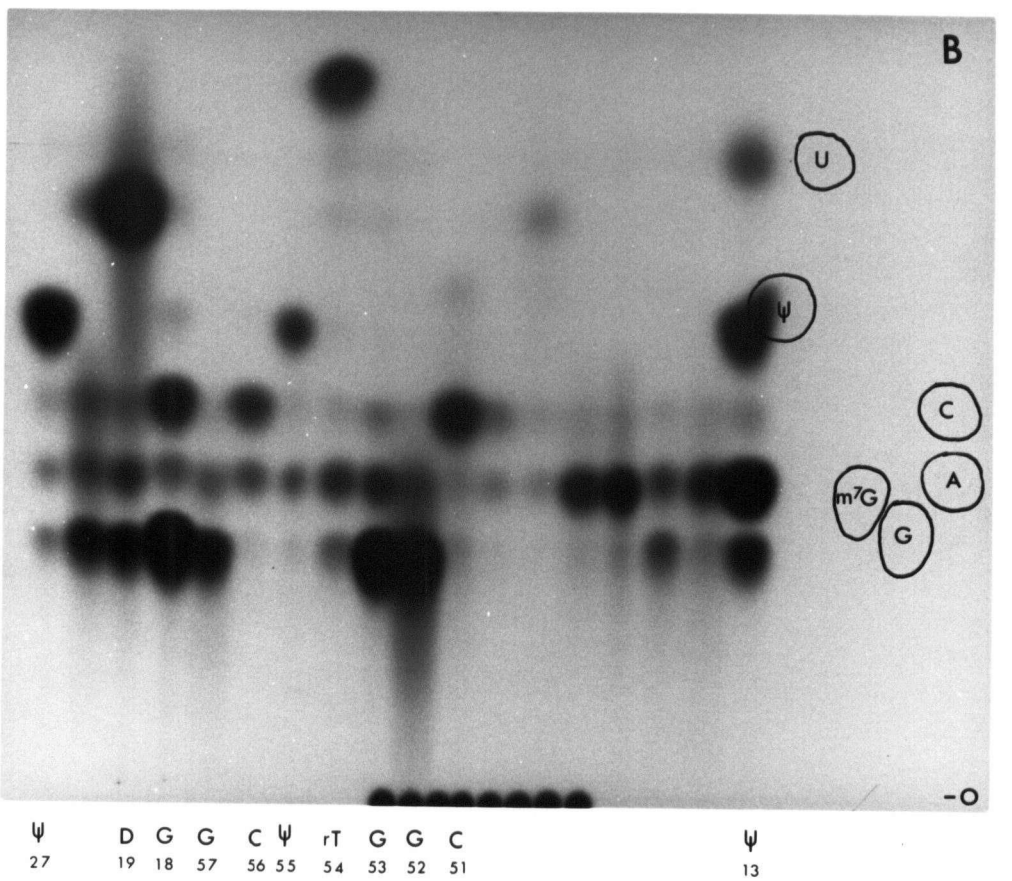
C. The results presented in panel C are from a different experiment from those shown in panels A and B. In this experiment the 5', 3'-bisphosphates (pNp's) of nucleotides 37-45 and 47 were identified by chromatography on PEIcellulose TLC plates developed in Solvent A.

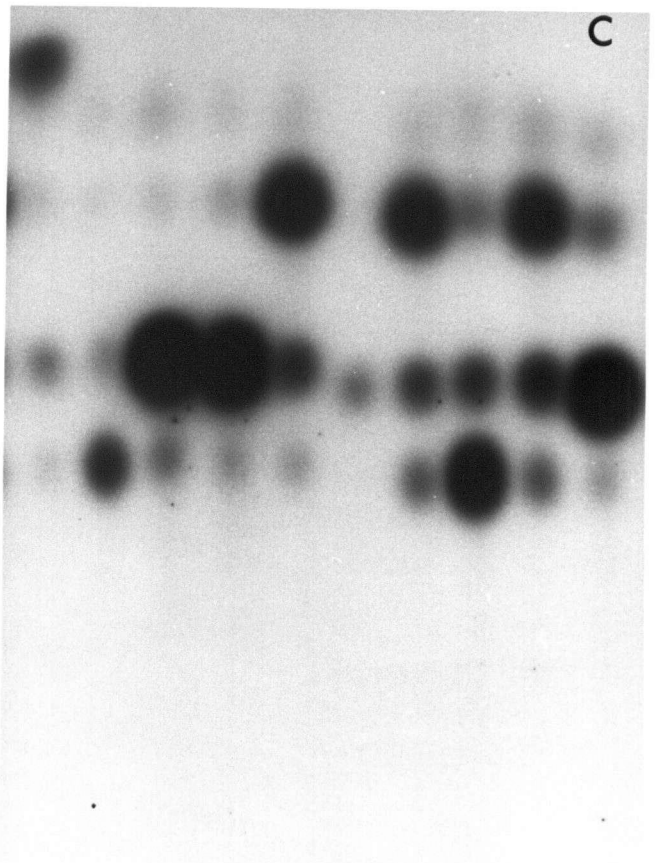
D. Chromatography of the pN's from positions 16, 17, 21-26 and 29-34 of $tRNA_{3b}^{Val}$ on cellulose TLC plates developed in Solvent B.

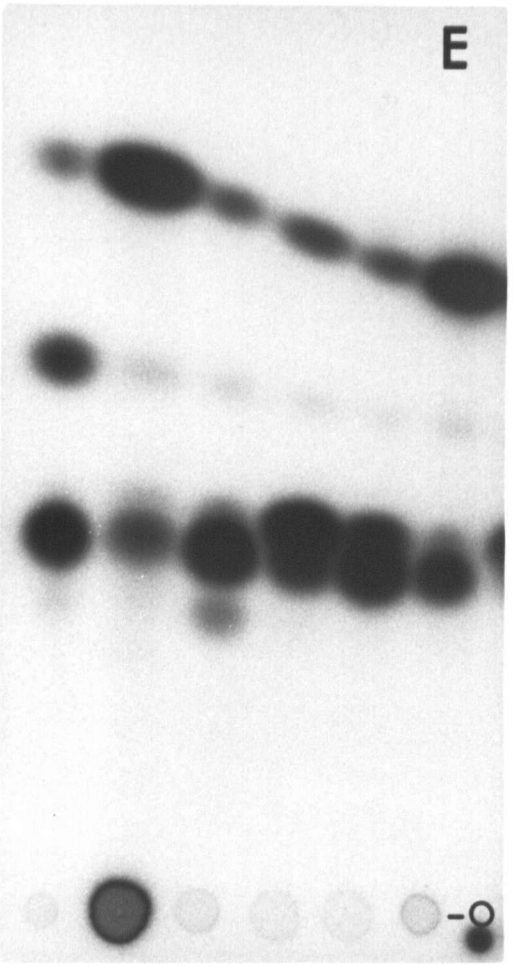
E. Chromatography of pN's from positions 9-12 and 14, 15 of $tRNA_{3b}^{Val}$ on cellulose TLC plates developed in Solvent B.

F. Identification of the acp^3U residue at position 20 of $tRNA_{3b}^{Val}$. pNp's resulting from the Stanley and Vassilenko sequencing of the D-loop of $tRNA_{3b}^{Val}$ were chromatographed on PEI-cellulose TLC plates in Solvent A.

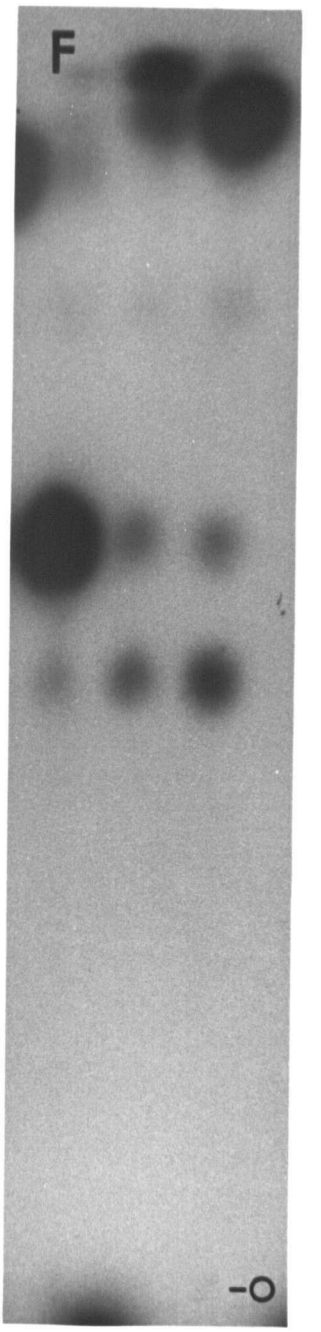

A
A
U
G
-O
G G G C G G G C C C A m¹A A C A
69 68 67 66 65 64 63 62 61 60 59 58 37 36 35


B
U
Ψ
C
A
m⁷G
G
-O
Ψ D G G C Ψ rT G G C Ψ
27 19 18 57 56 55 54 53 52 51 13


C
D G A A C A C G C A
47 45 44 43 42 41 40 39 38 37
D
U
C
Ψ
C U U C G U G G C A C U A G C
34 33 32 31 30 29 28 26 25 24 23 22 21 17 16



G A G U G A
15 14 12 11 10 9



A acp^{3}U D
21 20 19

Figure 22. Wandering-spot analysis of the 5'-end of $tRNA_{3b}^{Val}$.

In an attempt to cleave $tRNA_{3b}^{Val}$ at its m^7G residue (Silberklang et al., 1979), $tRNA_{3b}^{Val}$ (2 µg) was incubated at room temperature with 50 mM NaOH, 0.25 mM EDTA (40 µl). After 15 min, 3 µl of 1 M acetic acid and 40 µl of aniline-HCl pH 4.5 were added to the reaction mixture and incubation was continued at 37°C for 4 h. The cleaved RNA was concentrated by precipitation with ethanol, dephosphorylated and then 5' end-labelled as described in Materials and Methods. The labelled RNA fragments were separated from each other by polyacrylamide gel electrophoresis. Autoradiography of the gel revealed a large number of RNA fragments. Several of the most prominent of these bands were excised from the gel. RNA was eluted from each of the bands. The 5'-terminal nucleotide of each fragment was identified by incubating an aliquot of RNA from each band in 5 µl 10 mM NH_4OAc pH 4.5, 1 mM EDTA containing 0.1 unit of RNase T_2 for 3 h at 37°C. The [5'-^{32}P]pNp's produced by the RNase T_2 digestion were identified by TLC on PEI-cellulose plates developed in Solvent A. One of the fragments had a 5'-terminal G residue. The sequence of this fragment was determined by the wandering-spot procedure (Materials and Methods). The xylene cyanol marker dye was allowed to migrate 6 cm during electrophoresis in the first dimension (1). Homo-mix V of Jay et al. (1974) was used for homochromatography in the second dimension (2). R indicates the position of the red marker dye. The mark (▲) at C5 indicates that a transcript of the $tRNA_{3b}^{Val}$-like genes of plasmids pDt48 and pDt41R would contain a U residue at that position.

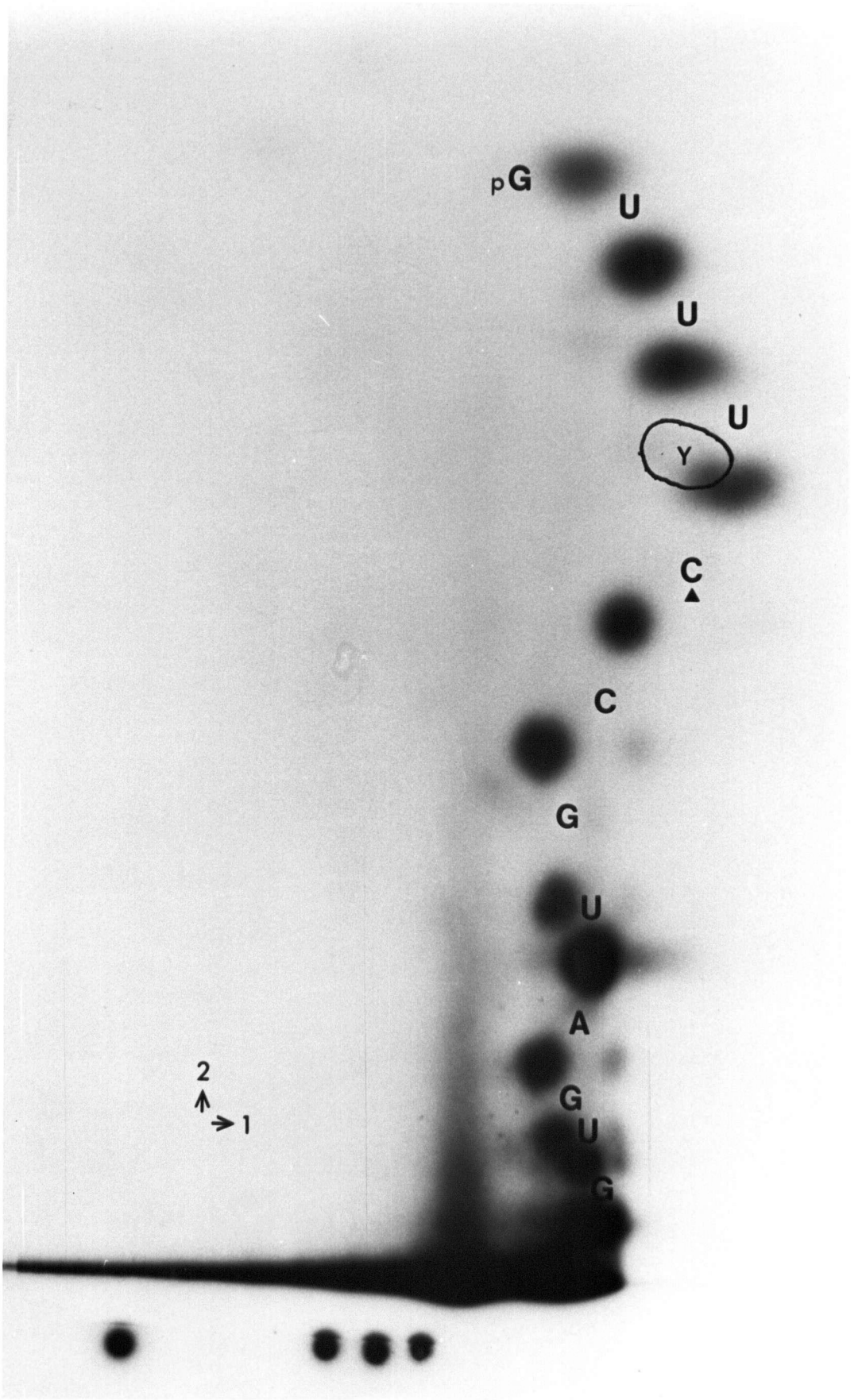

pG
U
U
U
Y
C
C
G
U
A
G
U
G
2
1

Figure 23. Wandering-spot analysis of the 3'-end of $tRNA_{3b}^{Val1}$.

$tRNA_{3b}^{Val1}$ was labelled at the 3'-end with [5'-^{32}P]pCp as described in Materials and Methods. The end-labelled RNA was heated to 100°C for 40 min in 4 µl of 98% formamide. The partially hydrolysed RNA was subjected to wandering-spot analysis as described in Materials and Methods. The xylene cyanol marker dye was allowed to migrate 8 cm during electrophoresis in the first dimension (1). Homo-mix V of Jay *et al.* (1974) was used for homochromatography in the second dimension (2).

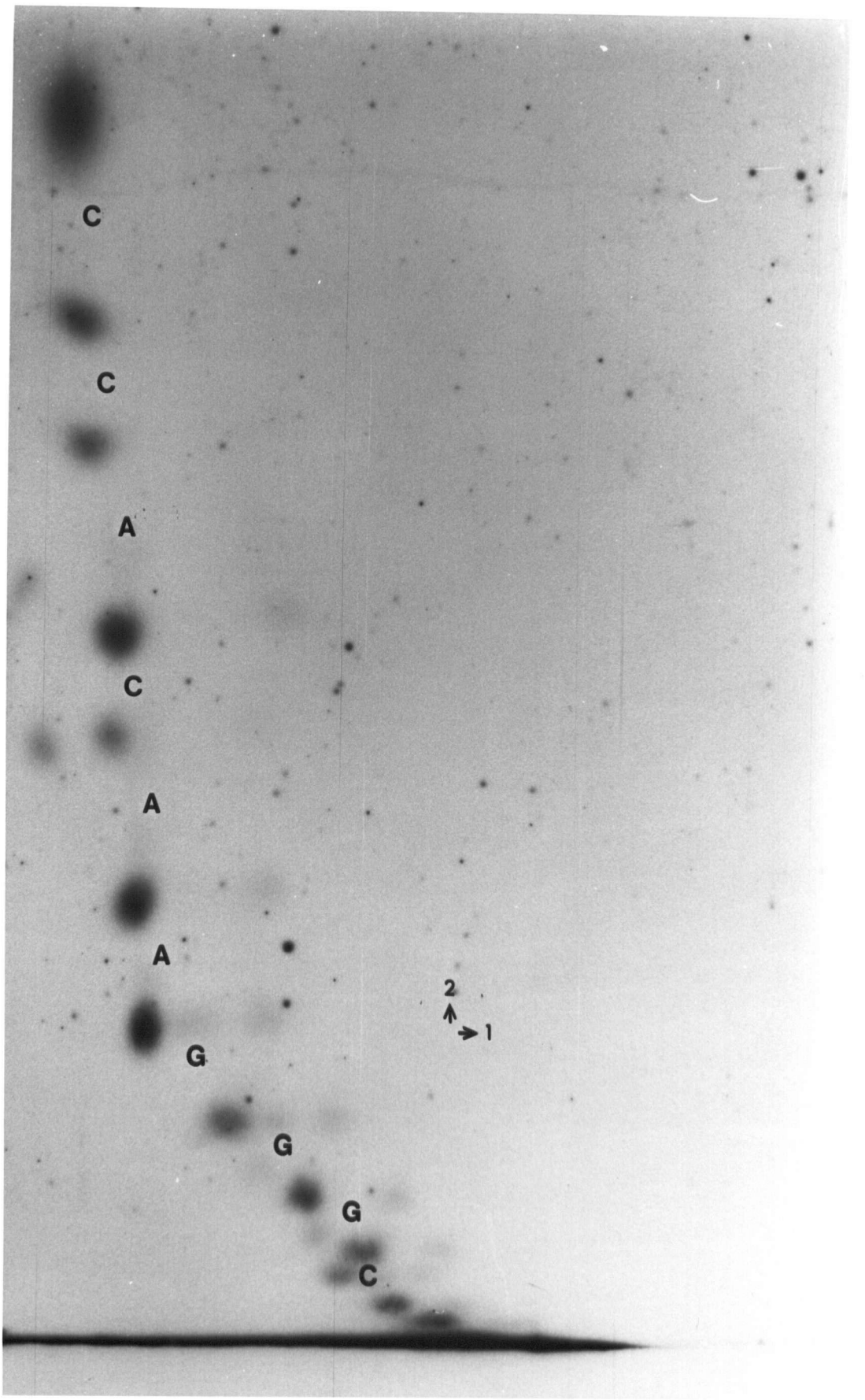

C
C
A
C
A
A
G
G
G
C
2
1

Table V. Chromatographic Mobilities of Nucleoside-5'-Phosphates on Cellulose TLC Plates[a] in Solvent D

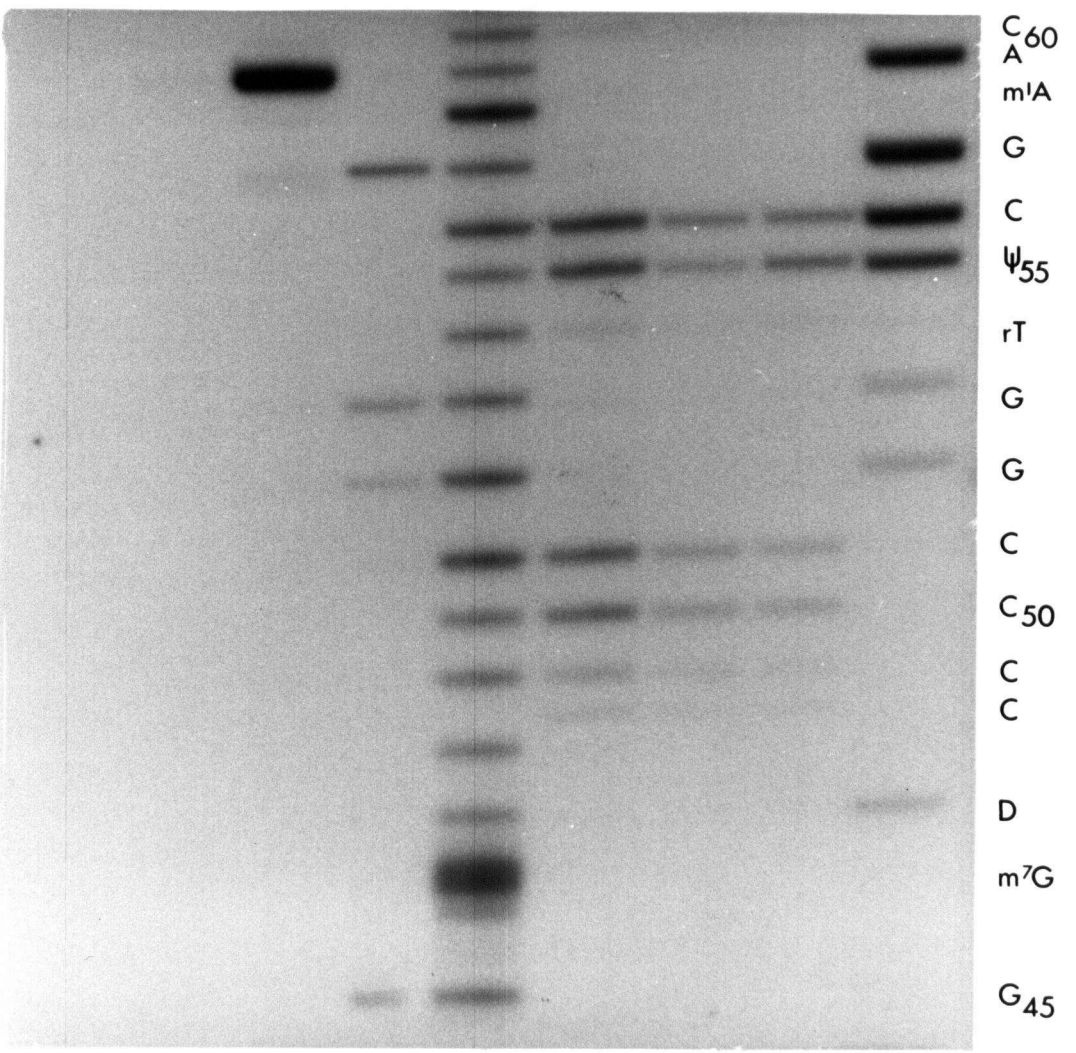
Nucleotide	R^{b}_{pU}
pU	1.00
pC	0.59
pA	0.40
pG	0.36
$pm^{7}G$	0.47
pΨ	0.72
prT	1.10
pD	0.90

a - cellulose on glass TLC plates (E. Merck)
b - mobility relative to pU

Like $tRNA_4^{Val}$, the variable loop region of $tRNA_{3b}^{Val}$ could not be sequenced by the Stanley and Vassilenko method presumably because of strong secondary structure of RNA fragments produced by cleavage in this region. Instead $tRNA_{3b}^{Val}$ was cleaved in the anticodon by limited digestion with RNase U_2 to produce half-molecules. The half-molecules could be efficiently 5' end-labelled and were then sequenced by the gel read-off method. Figure 24 shows an autoradiogram of part of such a sequencing gel. The nucleotide sequence of the variable loop region of $tRNA_{3b}^{Val}$ can be read from this autoradiogram. The strong band compression evident in the sequencing gels of the corresponding region in 3'-end labelled $tRNA_4^{Val}$ (Figure 18A) is absent in this autoradiogram. When present in the same molecule, the pyrimidine-rich variable loop of $tRNA_{3b}^{Val}$ or $tRNA_4^{Val}$ can presumably hydrogen bond to purine-rich regions near the 3'-ends of these tRNAs. Such intramolecular interactions may result in the anomalous electrophoretic behavior observed when 3' end labelled tRNA is sequenced by the gel read-off method. These interactions would also explain the poor 5' end-labelling of tRNA fragments produced by cleavage of $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ in the variable loop during

Figure 24. Gel read-off sequencing of $tRNA_{3b}^{Val}$: The nucleotide sequence of the variable arm of $tRNA_{3b}^{Val}$.

$tRNA_{3b}^{Val}$ was partially hydrolysed with RNase U_2 and the resulting fragments were 5' end-labelled as described in Materials and Methods. One of these fragments was sequenced by the gel read-off method (Materials and Methods). 10 µl reaction volumes contained: NE - no enzyme; U-1 - 0.01 unit RNase U_2; U-2 - 0.001 unit RNase U_2; T - 0.02 unit RNase T_1; L - 66% formamide, 100°C for 20 min; A-1 - 0.02 unit RNase A; A-2 - 0.01 unit RNase A; A-3 - 0.02 unit RNase A, 2 µl aliquots removed and the digestion stopped after 1, 2, 4, 8 and 16 min of incubation at 50°C; P - 0.1 unit of RNase Phy I. In the autoradiogram of the sequencing gel a very faint band is visible in the A-1 slot at position 47, the probable site of a D residue. C48 may be modified to m^5C.


NE
U-1
U-2
T
L
A-1
A-2
A-3
P
C
A60
m1A
G
C
Ψ55
rT
G
G
C
C50
C
C
D
m7G
G45

Stanley and Vassilenko sequencing of these tRNAs. When 5' end-labelled tRNA is sequenced by the gel read-off method cleavages in the variable loop by base-specific ribonucleases separate the two interacting regions of the tRNA. The end-labelled products of such cleavages migrate normally on the sequencing gel.

The nucleoside analysis of $tRNA_{3b}^{Val}$ indicates that each tRNA molecule contains a single N^7-methylguanosine (m^7G) residue (Table II). This nucleotide is found at only one position in sequenced tRNAs, in the middle of the variable loop (nucleotide 46) (Dirheimer et al., 1979). Figure 24 shows that none of the sequencing ribonucleases cleave $tRNA_{3b}^{Val}$ at position 46 but that hydrolysis in formamide cleaves particularly efficiently there giving a band in the formamide ladder at this position more intense than the other bands. This is characteristic of an m^7G residue (for example see Figure 18B).

Several lines of evidence suggest that there is a dihydrouridine (D) residue at position 47 of $tRNA_{3b}^{Val}$. The tRNA is poorly cleaved by RNase A but appreciably cleaved by RNase Phy I at position 47 (Figure 24). This suggests a modified uridine residue may be present at this site. Stanley and Vassilenko sequencing of $tRNA_{3b}^{Val}$ indicated that a D residue was present in the variable loop region of the tRNA (Figure 21) although the position of the nucleotide could not be established by this method. On the basis of these data one of the two D residues detected in the nucleoside analysis of $tRNA_{3b}^{Val}$ (Table II) was assigned to position 47. This assignment is consistent with the partial sequence of $tRNA_{3b}^{Val}$ published by Altwegg (1980).

$tRNA_{3b}^{Val}$ is cleaved at position 48 by RNase A but not by RNase Phy I. This indicated that a C residue is present at that position. Examination of Figure 24 shows that the labelled RNA fragment produced by RNase A

cleavage at position 48 migrates more slowly during eletrophoresis than the corresponding fragment produced by hydrolysis in hot, aqueous formamide. The reason for this anomalous behavior is not clear. It is tempting to speculate that nucleotide 48 is the 5-methylcytidine (m^5C) residue detected in the nucleoside analysis of $tRNA_{3b}^{Val}$ (Table II). Cribbs (1979) has previously noted that cleavage at an m^5C residue in a tRNA by RNase A produced a double band on a sequencing gel. He attributed this behavior to particularly efficient RNase A-catalysed ring-opening of cyclic phosphates on 3'-terminal m^5C residues. It should be mentioned that, under the condition used for gel read-off sequencing, the products of formamide hydrolysis and partial ribonuclease digestion are thought to bear 2',3'-cyclic phosphate moieties at their 3'-termini (Simoncsits et al., 1977). Thus formamide hydrolysis at position 48 of $tRNA_{3b}^{Val}$ would yield a product with a cyclic phosphate at the newly generated 3'-end. This fragment would migrate normally on the sequencing gel. Assuming that nucleotide 48 is an m^5C residue, cleavage at this position with RNase A would produce, according to Cribbs' proposal, substantial amounts of RNA with a non-cyclic 3'-phosphate residue. This material may migrate anomalously during electrophoresis. m^5C residues are found at position 48 of all sequenced valine tRNAs of higher eukaryotes (Gauss and Sprinzl, 1981).

B. Features of the $tRNA_{3b}^{Val}$ Sequence

Although the nucleotide sequence of $tRNA_{3b}^{Val}$ is very similar to that of Drosophila $tRNA_4^{Val}$, it does contain a number of unique features. Some of these features are briefly discussed below. $tRNA_{3b}^{Val}$ has a mismatched U4-G69 base-pair in its acceptor stem (Figure 20). Such mismatches are quite common in the stem regions of tRNAs. Clark (1978) has hypothesized that the irregularity introduced into double helical regions of tRNAs by these mismatched base-pairs are important sites of enzyme recogni-

tion. It is interesting that in $tRNA_4^{Val}$ the same base-pair is of the Watson-Crick type but that the U4 residue is modified to 2'-O-methyluridine. Examination of a space-filling model of yeast $tRNA^{Phe}$, presumed to be similar in structure to the structure of $tRNA_4^{Val}$, shows that a 2'-O-methyl group on nucleotide 4 would lie in the shallow groove of the helical stem. Some alteration in the helix backbone would probably be required to accomodate the bulky methyl group. Thus in both $tRNA_{3b}^{Val}$ and $tRNA_4^{Val}$ the base-pair between nucleotides 4 and 69 is distinctive.

Unlike the case in $tRNA_4^{Val}$, the uridine residue at position 20 of $tRNA_{3b}^{Val}$ is completely modified to acp^3U (Figure 20). As with $tRNA_4^{Val}$, acp^3U (nucleoside X) was not detected in the nucleoside analysis of $tRNA_{3b}^{Val}$ (Table II).

$tRNA_{3b}^{Val}$ has a CAC anticodon. According to the wobble hypothesis (see Results and Discussion, Section I) this tRNA should, therefore, decode only the GUG codon. The tRNA's behaviour during ribosome-binding studies was consistent with this prediction (Figure 11). The anticodon loop of $tRNA_{3b}^{Val}$ ends with an unmodified C residue at position 38. In $tRNA_4^{Val}$ this position is occupied by an m^5C residue. This difference in methylation of C38 among isoacceptors is unusual. Human $tRNA_{(IAC)}^{Val}$ and $tRNA_{(CAC)}^{Val}$ and the 3 yeast $tRNA^{Val}$ isoacceptors all have unmethylated C at position 38. In $tRNA_{(IAC)}^{Val}$ of mouse myeloma cells and rabbit liver an m^5C residue is found at this position.

The variable loop of $tRNA_{3b}^{Val}$ contains a dihydrouridine residue at position 47 (Figure 20). In this respect $tRNA_{3b}^{Val}$ resembles all sequenced eukaryotic valine tRNAs except Drosophila $tRNA_4^{Val}$ which contains a C residue at this site.

C. Homologies Between $tRNA_{3b}^{Val}$ and Other Valine tRNAs

Drosophila $tRNA_{3b}^{Val}$ contains all the invariant and strongly con-

served nucleotides present in all cytoplasmic tRNAs. In addition, it contains all the nucleotides common to eukaryotic valine tRNAs noted in Section II.D of the Results and Discussion. One difference between $tRNA_{3b}^{Val}$ and all other sequenced valine tRNAs is the G28-C42 base-pair; in the other tRNAs a pyrimidine 28-purine 42 base-pair is found at this site.

$tRNA_{3b}^{Val}$ displays 88% homology to Drosophila $tRNA_4^{Val}$. This degree of homology is much stronger than the 61% homology found between the corresponding $tRNA_{(CAC)}^{Val}$ and $tRNA_{(IAC)}^{Val}$ isoacceptors of yeast. There are several possible explanations for this. The two Drosophila tRNAs may have diverged from a common ancestral sequence more recently than the two species from yeast. Alternatively, stronger selective pressures may have prevented the two Drosophila species from diverging as rapidly a those of yeast. It is also possible that the two Drosophila tRNAs are the products of convergent evolution from ancestral sequences that were less similar to one another. Of course, any combination of these processes is also possible. Deciding which, if any, of these alternatives is correct is beyond the scope of this thesis. A recent review by Cedergren et al. (1981) describes the difficulties associated with the determination of phylogenetic relationships among tRNA sequences. While $tRNA_{3b}^{Val}$ and $tRNA_4^{Val}$ of Drosophila are more similar than the corresponding tRNAs of yeast they are much less similar than the sequenced $tRNA_{(CAC)}^{Val}$ and $tRNA_{(IAC)}^{Val}$ of human placenta. These two tRNAs differ by a single nucleotide change in the first position of the anticodon (Chen and Roe, 1977). It remains to be seen if the sequenced human $tRNA_{(CAC)}^{Val}$ is the only valine tRNA in human placenta with the CAC anticodon.

The degree of homology between Drosophila $tRNA_{3b}^{Val}$ and the other sequenced valine tRNAs is presented in Table VI. $tRNA_{3b}^{Val}$ shows no more homology to Drosophila $tRNA_4^{Val}$ than it does to the mammalian

Table VI. Homology Between $tRNA_{3b}^{Val}$ and Other Sequenced Valine tRNAs and Valine tRNA Genes

Source	Valine Anticodon	Percent Homology	Reference
D. melanogaster	CAC	100	-
D. melanogaster	IAC	88	-
mammalian	IAC	88	1
human placenta	CAC	89	1
yeast	IAC	74	1
yeast	CAC	77	1
yeast	UAC	66	1
T. utilis	IAC	70	1
B. stearothermophilus	GAC	63	1
B. subtilis	UAC	53	1
E. coli	UAC	51	1
E. coli 2a*	GAC	55	1
E. coli 2b	GAC	49	1
spinich chloroplast	UAC	48	2
maize chloroplast	UAC	42	3
N. crassa mitochondria	UAC	45	1
yeast mitochondria	UAC	45	4
mouse mitochondria	UAC	40	5
A. nidulans mitochondria	UAC	42	6
human mitochondria	UAC	34	7

* *E. coli* contains two valine tRNAs, $tRNA_{2a}^{Val}$ and $tRNA_{2b}^{Val}$, with the same GAC anticodon.

References

1. Gauss and Sprinzl, 1981
2. Sprouse et al., 1981
3. Schwarz et al., 1981
4. Li and Tzagoloff, 1979
5. Van Etten et al., 1980
6. Kochel et al., 1981
7. Anderson et al., 1981

valine tRNAs. This suggests that the two Drosophila tRNAs diverged from a common ancestral sequence about the time the lines leading to arthropods and chordates diverged from each other. tRNA sequence data from a wider variety of organisms and more sophisticated analysis of the data are required to prove or disprove this hypothesis. Drosophila $tRNA_{3b}^{Val}$ displays about equal homology to yeast $tRNA_{(IAC)}^{Val}$ and $tRNA_{(CAC)}^{Val}$ (about 75%) but distinctly lower homology to yeast $tRNA_{(UAC)}^{Val}$ (66%). Bacillus stearothermophilus $tRNA_{(GAC)}^{Val}$ shows considerably greater homology to $tRNA_{3b}^{Val}$ (63%) (and to a lesser extent $tRNA_{4}^{Val}$) than do other bacterial valine tRNAs (about 52%). Most of the increased homology can be attributed to the great similarity in the acceptor stems of the two tRNAs. $tRNA_{3b}^{Val}$, like $tRNA_{4}^{Val}$, displays the least homology to the valine tRNAs of subcellular organelles (34-48%).

IV. The Nucleotide Sequence of $tRNA_4^{Val}$ Genes

The recombinant plasmid pDt55 consists of an 8 kb segment of Drosophila DNA cloned into the Hind III restriction site of the plasmid vector pBR322 (Table III). The plasmid can form an RNA:DNA hybrid with $tRNA_{4}^{Val}$ and was, therefore, presumed to contain a gene for this tRNA (Dunn et al., 1979). In situ hybridization of pDt55 to Drosophila polytene chromosomes showed that the Drosophila DNA of the plasmid came from the 70BC region, one of the major sites of $tRNA_{4}^{Val}$ hybridization (Table I). In the following sections the nucleotide sequence of the two $tRNA_{4}^{Val}$ genes of pDt55 will be presented and their sequences will be compared to those of other Drosophila $tRNA_{4}^{Val}$ genes.

A. Strategy Used to Sequence the $tRNA_4^{Val}$ Genes of pDt55

Plasmid pDt55 carries a large segment of Drosophila DNA. The first objective in sequencing the $tRNA_{4}^{Val}$ genes of this plasmid was to iden-

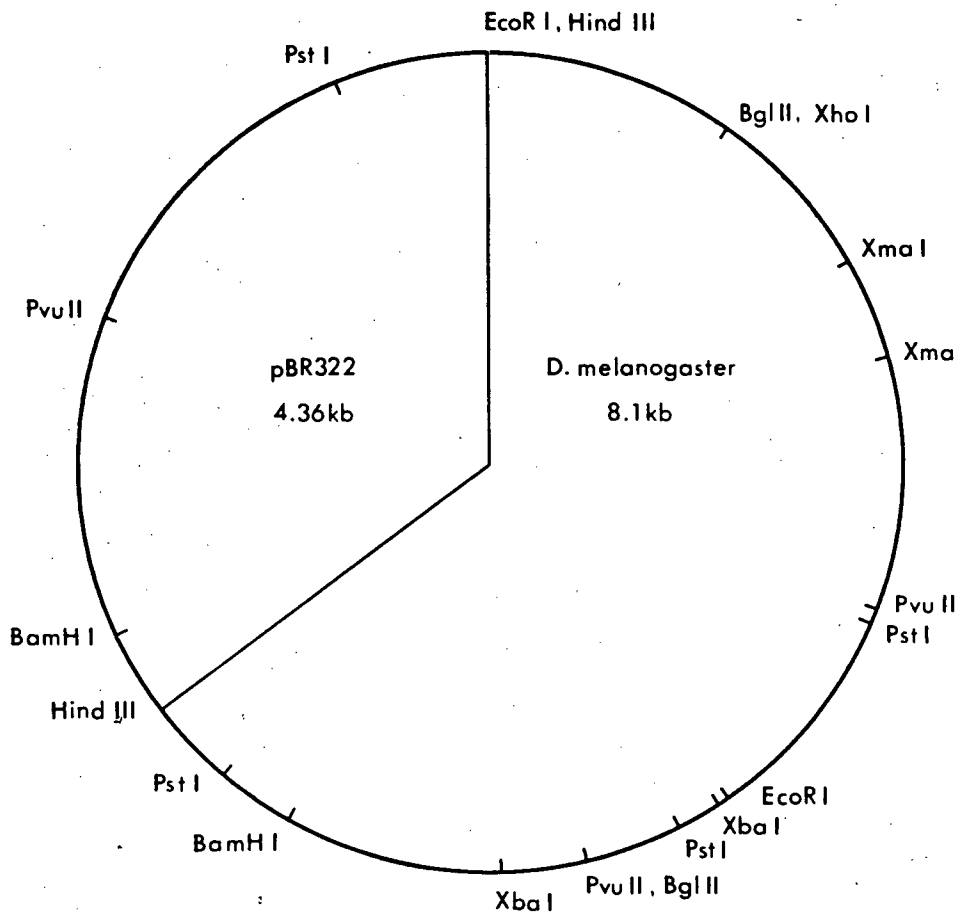
tify smaller fragments of DNA, amenable to sequence determination by the Maxam and Gilbert method, that contained the genes of interest. To this end a restriction endonuclease cleavage map of the plasmid was constructed. This map is shown in Figure 25. The two sites of cleavage by restriction endonuclease Xma I were of special significance. $tRNA_4^{Val}$ contains the sequence CCCGGG (nucleotides 60-65, Figure 12), the recognition sequence of Xma I. Thus the genes for this tRNA should be cleaved by this restriction enzyme.

Restriction fragments of pDt55 containing the $tRNA_4^{Val}$ genes were identified by their sensitivity to cleavage with Xma I. pDt55 DNA was cleaved with a restriction nuclease and the resulting fragments were end-labelled. An aliquot of this mixture of labelled restriction fragments was digested with Xma I. The original restriction enzyme digest and the DNA that had been cleaved with both the original enzyme and Xma I were applied to adjacent slots in a polyacrylamide gel and the restriction fragments were separated by electrophoresis. Autoradiography revealed the position of labelled DNA fragments on the gel. Any fragments present in DNA cut only with the first restriction enzyme but absent in the DNA cut with both the first enzyme and Xma I must contain an Xma I restriction site and hence perhaps a $tRNA_4^{Val}$ gene. Figure 26 shows the results of such an experiment. Digestion of pDt55 DNA with restriction enzyme Hinf I produces a single large fragment (about 1 kb long) sensitive to cleavage by Xma I. This fragment was used as the source of DNA for most of the sequencing experiments done on plasmid pDt55.

In a typical sequencing experiment plasmid DNA (20 µg) was cleaved with Hinf I and the fragments were end-labelled to a low specific activity with ^{32}P (see Materials and Methods). The fragments were separated by electrophoresis on a 5% polyacrylamide gel and the labelled fragments were

Figure 25. The restriction map of plasmid pDt55.

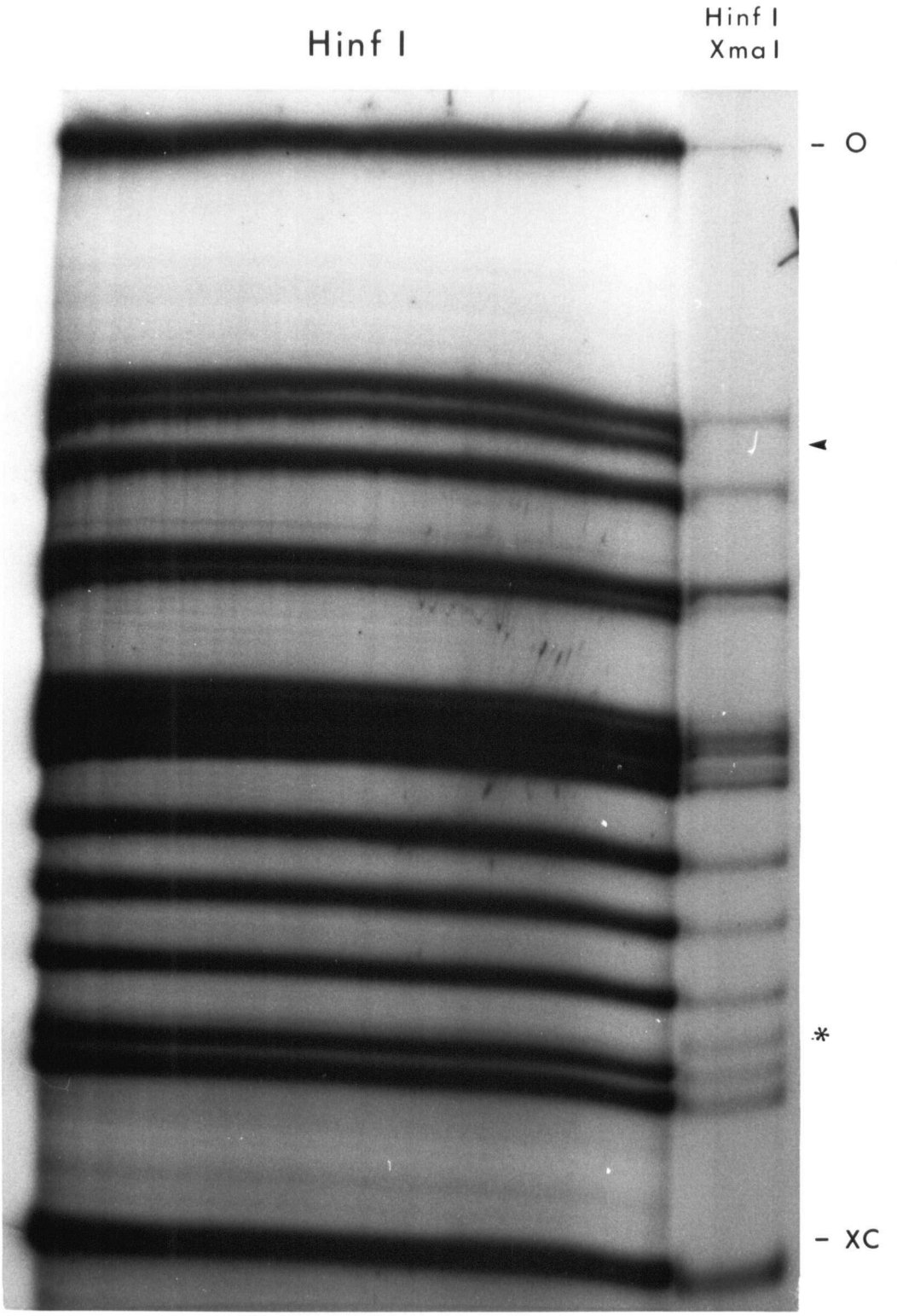
Recombinant plasmid pDt55 consists of a 8.1 kb Hind III fragment of Drosophila DNA cloned into the single Hind III site of the plasmid vector pBR322. A restriction endonuclease cleavage map of pDt55 was constructed by the multiple enzyme digest method described by Danna (1980). The two Xma I sites of pDt55 mark the sites of the plasmid's two $tRNA_4^{Val}$ genes.


EcoR I, Hind III
Pst I
Bgl II, Xho I
Xma I
Xma I
Pvu II
Pst I
EcoR I
Xba I
Pst I
Pvu II, Bgl II
Xba I
Pst I
BamH I
Hind III
BamH I
Pvu II
pBR322
4.36kb
D. melanogaster
8.1kb

pDt55

Figure 26. Identification of the Hinf I restriction fragment of pDt55 that contains the $tRNA_4^{Val}$ genes.

pDt55 DNA (20 µg) was dissolved in 400 µl of restriction enzyme buffer B (Materials and Methods) containing 22 units of restriction endonuclease Hinf I (New England Biolabs) and incubated for 3.5 h at 37°C. The cleaved DNA was precipitated with ethanol and redissolved in 50 µl of buffer B containing 80 µCi of [α-^{32}P]dATP (Amersham, 2000-3000 Ci/mmol) and 1 unit of Klenow DNA polymerase I (Boehringer-Mannheim). An aliquot (10 µl) of the reaction mixture was withdrawn and its DNA was precipitated with ethanol in preparation for cleavage with Xma I. EDTA (5 µl, 0.25 M) and sucrose (15 µl, 40% w/v) were added to the remainder of the reaction mixture which was then chilled on ice until it was loaded onto the gel. The aliquot of Hinf Icleaved DNA was dissolved in 10 µl of restriction buffer A (Materials and Methods) that contained 2 units of Xma I. After incubation at 37°C for 2 h, 3 µl of 40% sucrose containing 0.1% xylene cyanol was added and it, together with the Hinf I-cleaved DNA, was loaded onto a non-denaturing 5% polyacrylamide gel (15 x 40 x 0.05 cm) and run for 16 h at 100 V (2.8 V/cm). The autoradiogram of the gel showed that a large Hinf I fragment of pDt55 (marked by the arrow) contained at least one Xma I cleavage site. Cleavage of this Hinf I fragment with Xma I generated a new, much smaller fragment (*). "O" is the origin of the gel, "XC" marks the position of xylene cyanol.


Hinf I
Hinf I
Xma I
- O
◄
*
- XC

located by autoradiography. The second largest band (Figure 26) on the gel was excised and the DNA was eluted from it. This DNA was cleaved with any one of several restriction enzymes and the fragments were end-labelled to high specific activity. These labelled fragments were strand separated as described in Materials and Methods. Figure 27 shows the autoradiograph of such a strand separation. Bands containing single-stranded fragments of end-labelled DNA were excised from the gel and the DNA was eluted from the gel slices. The nucleotide sequence of the eluted fragments was determined by the method of Maxam and Gilbert (1980). The autoradiogram of one sequencing experiment is shown in Figure 28. The strategy used to sequence the two tRNA genes of pDt55 is summarized in Figure 29. For 50% of the Hinf I fragment depicted in Figure 29 both strands of the DNA were sequenced. The rest of the DNA, except for 70 bp that was sequenced only once, was sequenced at least twice in the same direction. The DNA between the two tRNA genes is very A-T rich (73% A-T) and contains few restriction enzyme recognition sequences. The lack of suitable restriction enzyme sites made this region difficult to sequence. At pH 8.5 under conditions of low ionic strength restriction enzyme EcoR I is reported to recognize and cleave the sequence AATT in addition to its normal recognition sequence, GAATTC (Polisky et al., 1975). The less specific recognition properties of EcoR I under these conditions are termed the EcoR I* activity. The use of EcoR I* permitted the region between the two $tRNA_4^{Val}$ genes to be sequenced. The major sites of cleavage by EcoR I* were: AAATTC, GAATTT and GAGTTC. All these sites are related to the normal recognition sequence of EcoR I by a single base change and all retain at least 1 G-C base-pair in the recognition sequence.

B. The Nucleotide Sequence of the tRNA Genes of pDt55

The nucleotide sequence of the $tRNA_4^{Val}$ genes of plasmid pDt55 is

Figure 27. Strand separation of Dde I fragments of pDt55 DNA.

pDt55 DNA (20 µg) was cleaved with Hinf I and end-labelled by incorporation of [α-^{32}P]dATP into the restriction enzyme cleavage sites as described in Fig. 26 except that only 2 µCi of [α-^{32}P]dATP (2000-3000 Ci/mmol) was used for the labelling. The Hinf I fragments were separated by polyacrylamide gel electrophoresis. The DNA fragments were located on the gel by autoradiography and the second largest Hinf I fragment was eluted from the gel. This DNA was precipitated with ethanol and redissolved in 25 µl of buffer containing 100 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 1.25 unit of Dde I. After incubation at 37°C for 2 h, 30 µCi of [α-^{32}P]TTP (Amersham, 2000-3000 Ci/mmol) and 1 unit of Klenow DNA polymerase I (Boehringer-Mannheim) was added and incubation was continued at room temperature for 15 min. TTP (1 µl, 0.5 mM was added to the mixture and 30 s later EDTA (2 µl, 0.25 M) was added to stop the polymerase reaction. 3 µl of the reaction mixture was withdrawn and mixed with 3 µl of a solution containing 20% sucrose w/v, 0.1% xylene cyanol and 0.1% bromphenol blue. This sample was used as a control for the strand separation. Ultrapure urea (25 mg) was added to the remainder of the reaction mixture. The solution was heated to 100°C for 3 min, then quickly applied to a wide sample slot (8 cm) in a 15 x 15 x 0.05 cm non-denaturing 5% polyacrylamide gel. The control sample was applied to an adjacent narrow sample slot (C). The samples were subjected to electrophoresis at 500 V for 45 min. Electrophoresis was continued in the cold room (4°C) at 500 V (14.2 V/cm) for 4 h. Those bands (numbered 1-7) in the strand separation slot that had electrophoretic mobilities different from bands in the control slot were presumed to contain single-stranded DNA. DNA from these bands was eluted from the gel and subjected to Maxam and Gilbert sequence analysis. "XC" and "BPB" mark the positions of xylene cyanol and bromphenol blue marker dyes.

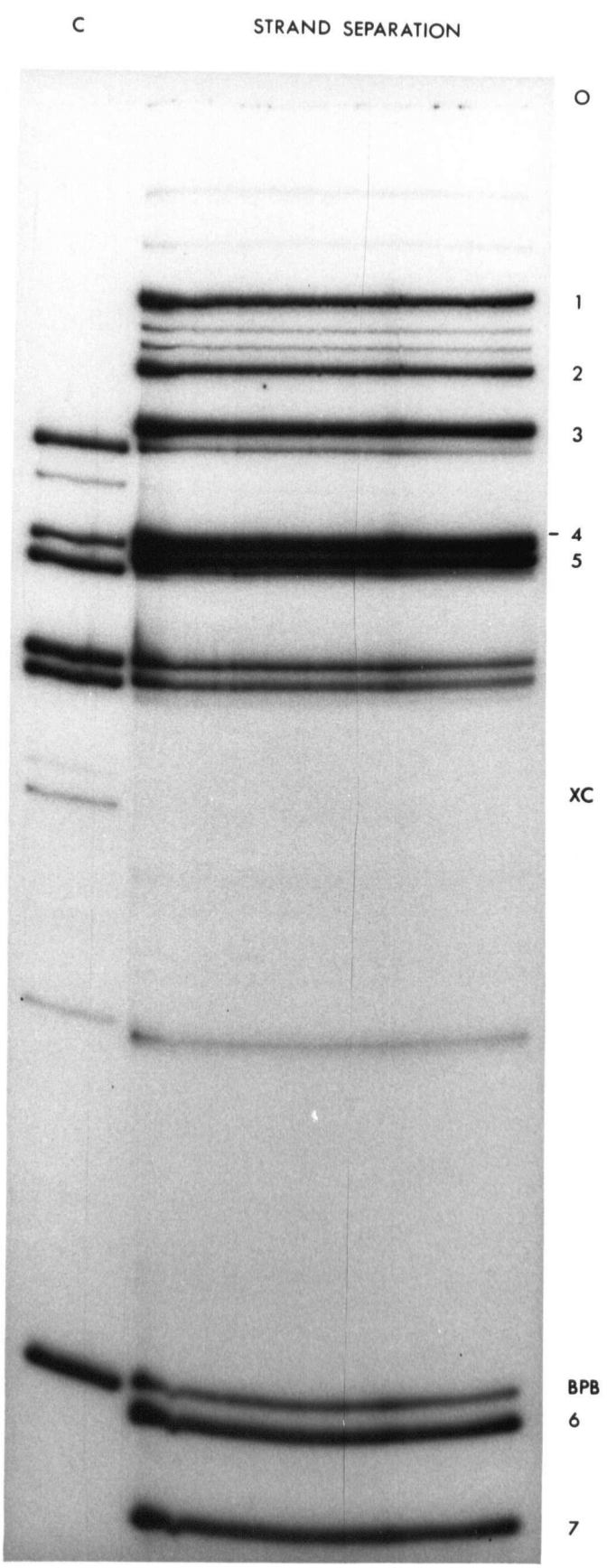

C
STRAND SEPARATION
O
1
2
3
4
5
XC
BPB
6
7

Figure 28. Maxam and Gilbert sequencing: The first $tRNA_4^{Val}$ gene of pDt55.

The sequence of DNA fragment 4 of Figure 27 was determined by the Maxam and Gilbert method as described in Materials and Methods. Samples of the sequencing reactions were applied to a denaturing 12% polyacrylamide gel and subjected to electrophoresis for 5.5 h (4 lanes at left) or 3 h (4 lanes at right) at 1800 V (51 V/cm). Fragment 4 contains the DNA sequence of the untranscribed strand of the first $tRNA_4^{Val}$ gene of pDt55 (Fig. 30). The 5' and 3'-ends of the $tRNA_4^{Val}$ gene are indicated in the Figure.

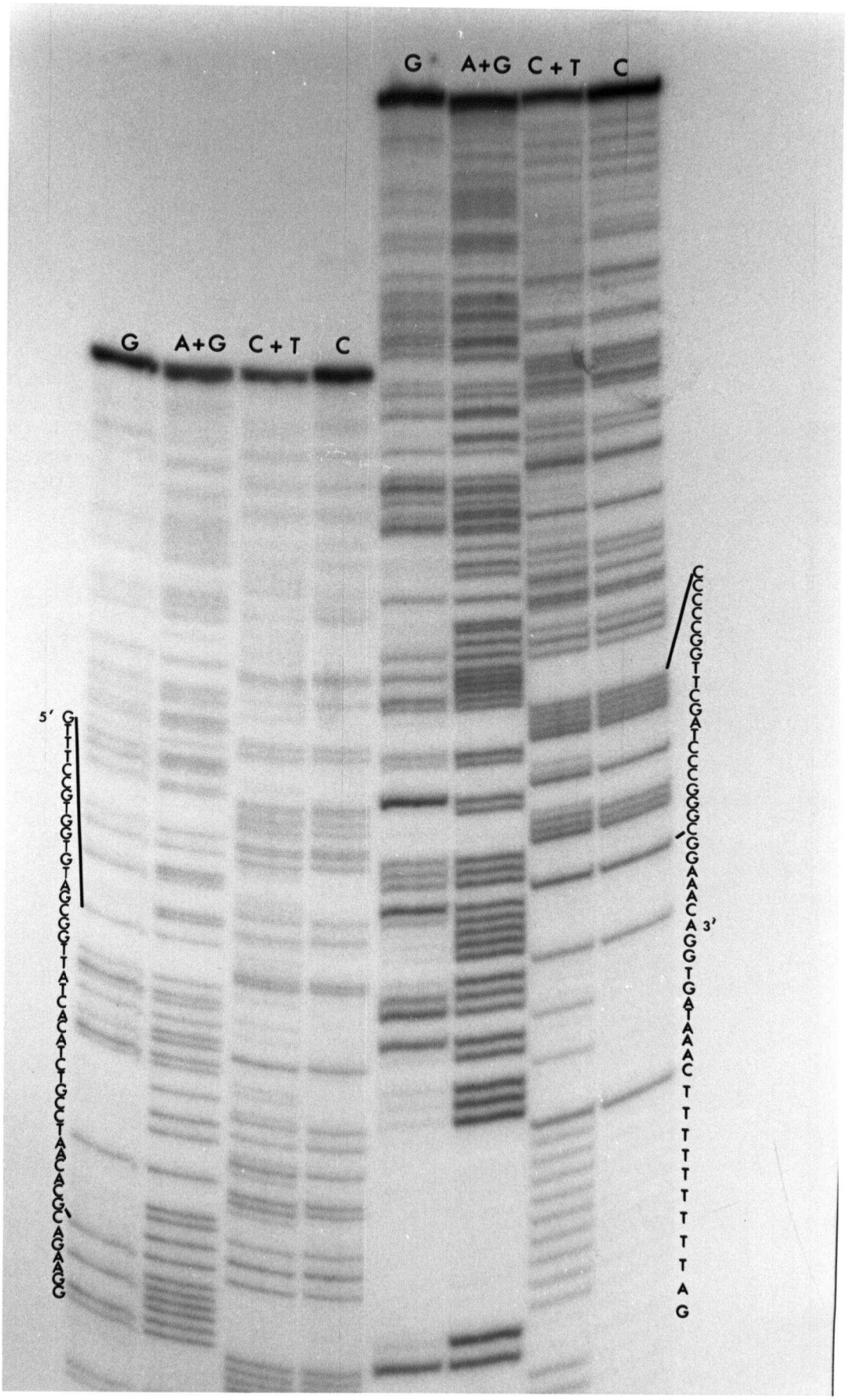
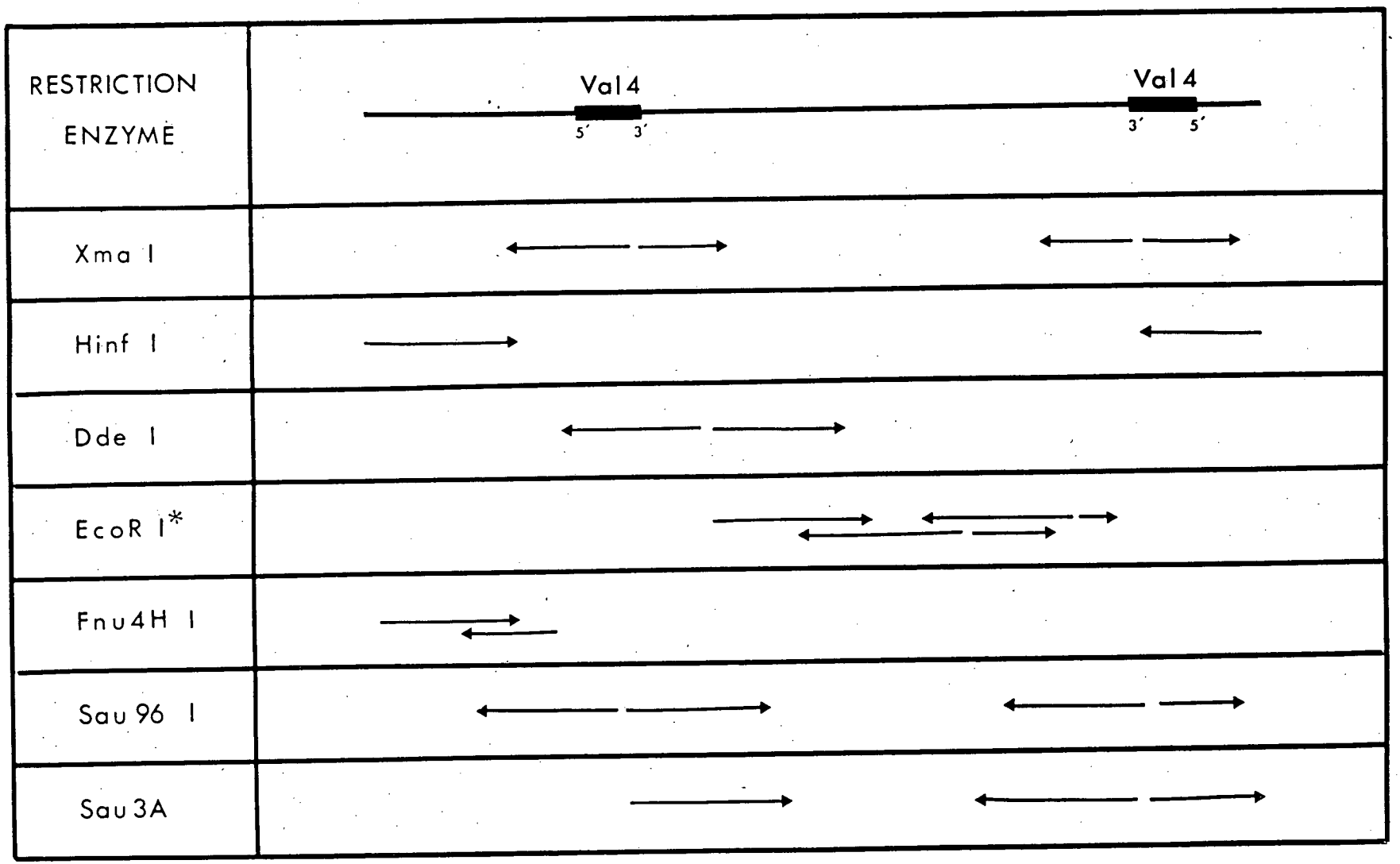

G
A+G
C+T
C
G
A+G
C+T
C
5'
3'

Figure 29. The strategy used to sequence the $tRNA_4^{Val}$ genes of pDt55.

Each arrow represents a nucleotide sequence determined by the method of Maxam and Gilbert (1.5 cm = 100 nucleotides). For the sequencing, pDt55 DNA was end-labelled at cleavage sites for the restriction endonucleases listed in the Figure. Arrows pointing to the left indicate that the sequence of the strand shown in the top panel of the Figure was determined. Arrows pointing to the right indicate the sequence of the complementary strand was determined.


RESTRICTION ENZYME
Val 4
5′ 3′
Val 4
3′ 5′
Xma I
Hinf I
Dde I
EcoR I*
Fnu4H I
Sau 96 I
Sau 3A

presented in Figure 30. The sequenced segment of Drosophila DNA contains 2 identical $tRNA_4^{Val}$ genes of opposite polarity 525 bp apart. With respect to the pDt55 restriction map (Figure 25) the first gene shown in Figure 30 is nearest the pBR322-Drosophila DNA boundary. The sequence of both genes corresponds to the sequence of $tRNA_4^{Val}$ (Figure 12). The sequence homology between the two genes extends 7 bp beyond the 5'-ends and 5 bp beyond the 3'-ends of the structural genes. In fact, with loop-outs the homology extends for 24 bp beyond the 3'-end. These regions of homology may be important for the functioning of the genes or they may be the result of a duplication by which a single ancestral gene gave rise to the present gene pair. Similar homologies are noted in the flanking sequences of a group of 5 Drosophila $tRNA^{Glu}$ genes thought to have arisen by a combination of gene duplication and unequal crossing over (Hosbach et al., 1980). A short distance beyond the 3'-ends of the $tRNA_4^{Val}$ genes of pDt55 are clusters of T residues in the non-transcribed strand (T_9 for the first gene, T_6 for the second). These sequences are probably termination signals for eukaryotic RNA polymerase III (Valenzuela et al., 1977; Silverman et al., 1979).

The 5'-flanking sequences of the two $tRNA_4^{Val}$ genes contain a number of interesting features. For both genes there is the possibility of forming hair-pin structures in the 5'-flanking sequences. These hair-pins, starting 15 bp upstream from the first gene and 14 bp upstream from the second are shown in Figure 31. If loop-outs and G-T base-pairs are allowed, the hair-pin adjacent to the first gene would have a 13 bp stem and a 7 nucleotide loop while the hair-pin upstream from the second gene would have a 17 bp stem and a 3 nucleotide loop.

Figure 30. The nucleotide sequence of a segment of the Drosophila DNA insert of plasmid pDt55.

The strand shown is the non-transcribed strand with respect to the first $tRNA_4^{Val}$ gene and the transcribed strand of the second $tRNA_4^{Val}$ gene. The two $tRNA_4^{Val}$ genes are underlined.

The Nucleotide Sequence of pDt55

```
          10         20         30         40         50         60         70         80         90        100
CTCAGCAGCC ACCTTAAAAT AATTCTATTA TCAGTTGTGC TCTTTCCCCT TCACTGAGCT GAATACCATT AACAAAGACA AACTGCCCAA TCATTGGGTC
         110        120        130        140        150        160        170        180        190        200
TCCTTGAAAC ATTTCCCATA AAAATCACTC AAATAGATAC AATATACGAT TTTATTCAAG CAACCAGTTT TATTTTTGAC CCTTGGCAGT TGAGGTCGCT
         210        220        230        240        250        260        270        280        290        300
GAAGTTGACC TCTCTGCCGC TTAAGTTTCA ACTGTTTCCG TGGTGTAGCG GTTATCACAT CTGCCTAACA CGCAGAAGGC CCCCGGTTCG ATCCCGGGCG
         310        320        330        340        350        360        370        380        390        400
GAAACAGGTG ATAAACTTTT TTTTTAGTTT TTATACAATT CGTATTTTAA GAAACCACCA GACTAAATGG CTGAGTTCTC CTCTAACGAT ATTTAGGTAT
         410        420        430        440        450        460        470        480        490        500
AAAGTATTTG AGTATTAATT GAAATTTATA GATATGTGCA TAAATATTTC ACTTTTTTTT GCTAGTTCTT GATTGCTCGC TTTATGTGTA ACTTAAACGT
         510        520        530        540        550        560        570        580        590        600
TTTAAGCCAT TAGTATATAC TTGCAATAAA GTATTTAAGC ATTAATTAAA TTATCTATCA ATCCTAGCTT GTTATTTAGT GTACGCATCA TGAGTTACTT
         610        620        630        640        650        660        670        680        690        700
TGAACCTTAC AATTATTGTA TTAAAAAAAC GTACATTACA TTTTCCTATT GGAATTTATC ATAGAAAATA TATGACCAAA TGAAACCCTT TTACTCAAAT
         710        720        730        740        750        760        770        780        790        800
ATGTCAACTA AATACATCTA CTCCATACAG GCTGACTTAA AATATCGCAT AGCAACTACA GTTTCATTGA TAAAAATTCA AACATCTTTT ACATCAATCC
         810        820        830        840        850        860        870        880        890        900
GAATAACAAA AAATTAAAAA ATTTTTTCAC CTGTTTCCGC CCGGGATCGA ACCGGGGGCC TTCTGCGTGT TAGGCAGATG TGATAACCGC TACACCACGG
         910        920        930        940        950        960        970
AAACAGTTGA ATATAGTCCG TCCGAAGAGC CTCTTGTGAG CCACAGGAGG CTTTCGGTTG GGCAAAGTGC CAGTA
```

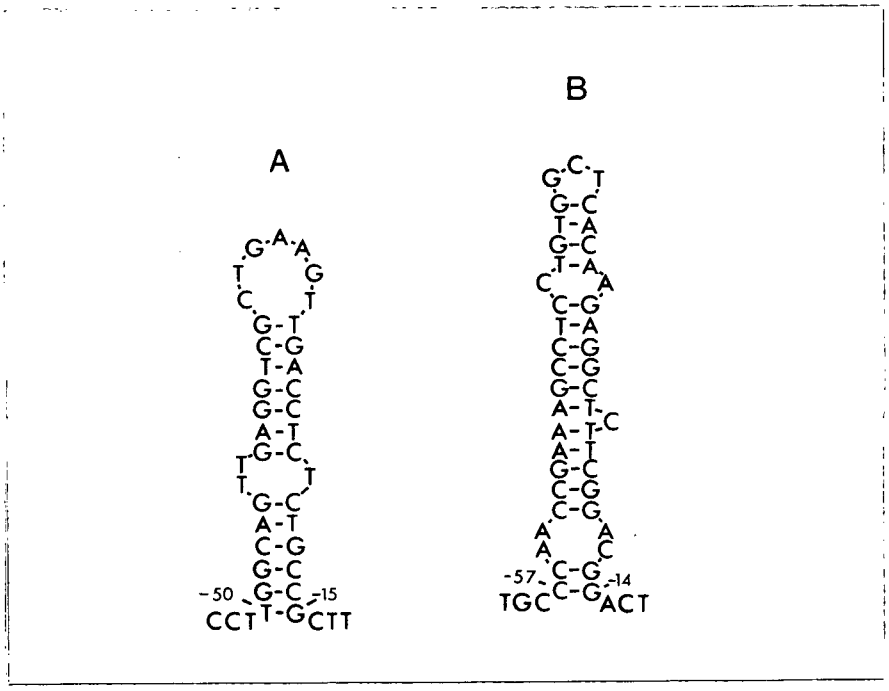


Figure 31. Possible Hair-pins in the 5'-Flanking Sequences of the $tRNA_4^{Val}$ Genes of pDt55 - A. A potential hair-pin structure 15 bp upstream from the first $tRNA_4^{Val}$ gene of pDt55. The strand shown is the non-transcribed strand.

B. A potential hair-pin structure 14 bp upstream from the second $tRNA_4^{Val}$ gene of pDt55. The strand shown is the non-transcribed strand of this gene; in the complementary strand the stem would be 3 bp longer because G-T base-pairs are considerably more stable than A-C pairs.

In the "Introduction" (Section V.C) it was noted that many sequenced *Drosophila* $tRNA_2^{Lys}$ genes have a sequence closely related to GGCAGTTTTTA about 25 bp upstream from the genes. Examination of the 5'-flanking sequences of the two $tRNA_4^{Val}$ genes shows that the first gene has a similar sequence, GGCAGTT, 49 bp upstream from the gene. Sixty-seven base-pairs upstream from the second gene (on the non-transcribed strand) is a related GGCACTT sequence. At present, the function of these sequences, if any, is unknown.

In their study of human initiator methionine tRNA genes Santos and Zasloff (1981) noted that short segments of the $tRNA_i^{Met}$ gene sequence

were repeated in the 5'-flanking sequences of these genes. They reviewed published tRNA gene sequence data and found other examples of duplicated sequences similar to those associated with human $tRNA_i^{Met}$ genes. In several genes for which the transcription start sites are known they fall within a 6 or 7 bp sequence that is also found between positions 15 and 40 of the structural gene. This suggests that the duplicated sequence may play a role in determining the transcription initiation site. The first $tRNA_4^{Val}$ gene of pDt55 contains the sequence TCTGCC between positions 27 and 32 (nucleotides 260-265, Figure 30). This sequence is also present 21 nucleotides upstream from the gene (nucleotides 213-218, Figure 30). It will be interesting to see if the transcript of this gene is initiated within the repeated sequence. Segments of the second $tRNA_4^{Val}$ gene of pDt55 are not repeated in the sequenced portion of the gene's 5'-flanking sequence.

C. Other $tRNA_4^{Val}$ Genes of Drosophila: Comparison to the $tRNA_4^{Val}$ Genes of pDt55

Recombinant plasmids containing $tRNA_4^{Val}$ genes were identified by the capacity of these plasmids to hybridize with $tRNA_4^{Val}$ (Dunn et al., 1979). Four different size Hind III fragments of Drosophila DNA containing $tRNA_4^{Val}$ genes were isolated in these plasmids (12 kb, 8 kb, 2 kb and 0.5 kb; Table III). The sequences of the $tRNA_4^{Val}$ genes of the 8 kb Hind III fragment of Drosophila DNA cloned in pDt55 were presented above. Other workers have sequenced the $tRNA^{Val}$ genes present in recombinant plasmids containing each of the three other Hind III fragments. These sequences are shown in Figure 32. Plasmids pDt92R and pDt120R (sequenced by Dr. C. Astell and B. Rajput), representative of plasmids containing the 0.5 kb and 2 kb Hind III fragments respectively, both hybridize to the 90C site on the Drosophila polytene chromosomes, a minor site of $tRNA_4^{Val}$ hybri-

Figure 32. The nucleotide sequence of segments of the Drosophila DNA inserts of plasmids pDt92R, pDt120R and pDt14.

Only the non-transcribed strands are shown. tRNA genes are underlined. At sites where the nucleotide sequences of pDt120R and pDt92 differ the nucleotides found in pDt92R are shown directly under the corresponding nucleotides in pDt120R. Dashes beneath the pDt120R sequence indicate regions which have not been sequenced in pDt92R. The G in parentheses at position 515 of pDt92R is postulated to occur in the chromosome to account for a Hind III recognition sequence found at this site. The sites at which the $tRNA_4^{Val}$-like genes of these plasmids differ from the $tRNA_4^{Val}$ genes of pDt55 are overlined in the Figure. The first gene in pDt14 codes for $tRNA_2^{Phe}$.

The Nucleotide Sequence of pDt120R with the differences in pDt92R under the sequence

```
          10         20         30         40         50         60         70         80         90        100
AAGCTTCGAG GTAGGTATGT AGCTTCACGG CTTGCTGCTT AAGTTGTTAC AATACCATTG GGAGGAGAGT GGG TAAAGG CAAGCCACTA TATAAGCGT
---------                                             G                        A  G      A
         110        120        130        140        150        160        170        180        190        200
CACTTTTTAA ATTATTTCCT ATAATTACAT TTTATAATTA CTTTGTGCTT TTATTATAAC AGATATATTT GCTAACTTAT CTTAAATTGT CTATGAGGAA
                                                                                          G
         210        220        230        240        250        260        270        280        290        300
AACGTTCGTC ATCCGAGTTT CCGTGGTGTA GTGGTTATCA CATCCGCCTA ACACGCGGAA GGCCCCCGGT TCAATCCCGG GCGGAAACAG TTGGAATTTA

         310        320        330        340        350        360        370        380        390        400
TTTTTTGCTA AATATTTATT TATCATAATG TTCAGTTGTA AAACACACAT AGCTAATAGT ATTTATAGCT GCATCATGGC CTTAAACTTA TCACGTTGCT
                                                       A    T                                              C
         410        420        430        440        450        460        470        480        490        500
TTTGCTTCAA GGCCTCGTGC TTCTTACGAT CCACATTTTT AAGCAGAAAT TCTTGAAATT TCCTACGCAT ATCTAACGAT AGACCTGTAT TTCGAAGGTC

         510        520        530        540        550
CAACCTCTCA AGAACCTTGT TGCAGCTAAT TATCTTCATC AAATGCTTGC CAAAGTC
            (G)---- ---------- ---------- ---------- -------
```

The Nucleotide Sequence of pDt14

```
          10         20         30         40         50         60         70         80         90        100
TGTATGCTCA TAATGCGCTT TAAAAAAAAG GATGTACTGA AAAATAATGT CTGTACTTTA AACAGGGAAG GTTCAAATAT TTTGGCAAGT ACGGCAAATT
         110        120        130        140        150        160        170        180        190        200
CACAACTTTG ATGTGAAGAA TCCCGTGGCC GAAATAGCTC AGTTGGGAGA GCGTTAGACT GAAGATCTAA AGGTCCCCGG TTCAATCCCG GGTTTCGGCA
         210        220        230        240        250        260        270        280        290        300
ATAATAATTT TTGCACAAAT TAGGCAGAAC TCGCATAAAA AAATAATATA AATTTTGGAA TAATTTTAAG GCATAATACA ACACATACCG TAAACACTGA
         310        320        330        340        350        360        370        380        390        400
ACACTTTTTA ATATTTGAAC GATCTGATAG TTCAATAAGA CCGTTATCCA AGTCTTATTT AAAATTATTT AATCACCCTC AAAACCGAAA AGCTGTGAAT
         410        420        430        440        450        460        470        480        490        500
CCACCCCATC ACGTGTTTCC GTGGTGTAGT GGTTATCACA TCCGCCTAAC ACGCGGAAGG CCCCCGGTTC AATCCCGGGC GGAAACATTG GAAATATTTA
         510        520        530        540        550        560        570        580        590        600
TTTTAATGCA TTTCCCAAAT TATTTTGCCT GTATAACTTA AATATATATA TTTTGTAATG TGATTTATGT GTCACTTTTG TCGGTCACCT GAAGTGCTTA
         610        620        630        640        650        660        670        680        690        700
AATTGTATAG TAAGTTTGAG GTCTCCACTG GCAAACCTCC CCTAGATCAG CGGCATGCCA GAATCTTCGT CCTGGACTCG CACCTACCTC AACTGGAAGC
         710        720        730        740        750        760
GGGTGCTGTT CCTCTCGCTG AAGTGCGTAT ACTTAATGAT GGTGGTGAAG TTCAGCCAGA GGT
```

dization (Table I). These two plasmids differ at only 8 sites in the 506 bp between the Hind III site at the left end of each Drosophila DNA insert and a Hind III site at the right end of the pDt92R insert. pDt120R lacks this second Hind III site and its insert is about 1500 bp longer. The differences in pDt92R are noted in Figure 32 by letters under the sequence of pDt120R. These two plasmids may have been generated because the DNA used for cloning the tRNA genes was isolated from a non-isogenic stock of D. melanogaster (Oregon R). The variations in nucleotide sequence outside the structural genes for the tRNAs may represent either differences in the DNA of homologous chromosomes, making the regions flanking the tRNA genes "allelic", or differences between repeated sequences in the DNA of the same chromosome. At present, we do not know which alternative is correct.

Plasmid pDt14 is representative of the $tRNA_4^{Val}$ plasmids containing the large 12 kb fragment of Drosophila DNA. This plasmid hybridizes to the 89B site on the polytene chromosomes, the other minor site of $tRNA_4^{Val}$ hybridization. A 763 bp segment of this plasmid was sequenced by Dr. A Delaney. The sequence is shown in Figure 32. This region contains two tRNA genes of the same polarity. The second is a $tRNA^{Val}$ gene located 214 bp downstream from the first, a $tRNA_2^{Phe}$ gene.

The $tRNA^{Val}$ genes of plasmids pDt92R, pDt120R and pDt14 are identical and differ from the $tRNA_4^{Val}$ genes of pDt55 at 4 sites. Transcription of the $tRNA_4^{Val}$-like genes would produce a tRNA with a U instead of a C at position 16 (in the D-loop), a C-G base-pair instead of a U29-A41 base-pair in the anticodon stem and an A57 instead of the G57 residue present in the T-loop of $tRNA_4^{Val}$ (Figure 12). Are these $tRNA_4^{Val}$-like genes expressed in vivo? At present a definite answer to this question cannot be given. Sequence analysis of $tRNA_{3a}^{Val}$ and $tRNA_{3b}^{Val}$ shows that the $tRNA_4^{Val}$-like genes do not code for these tRNAs. There is no hetero-

geneity in the sequence of $tRNA_4^{Val}$ that could be attributed to the presence of transcripts of the $tRNA_4^{Val}$-like genes in the sample of $tRNA_4^{Val}$ used for sequence determination (Figure 13). It is possible that the $tRNA_4^{Val}$-like genes code for one of the 4 minor valine tRNAs of Drosophila (Figure 7). None of these tRNAs have been characterized because of the paucity of material available for such investigations.

A number of facts hint that the $tRNA_4^{Val}$-like genes may be expressed and are not inactive pseudo-genes. First, the $tRNA_4^{Val}$-like genes of pDt14 and pDt92R/120R are well separated on the chromosomes yet have exactly the same sequence. This suggests that the sequence homogeneity of these genes has been actively maintained, presumably to preserve their ability to function in the cell. It could, however, be argued that whatever mechanism maintains the sequence homology of dispersed copies of functional tRNA genes also acts on pseudo-genes. Second, two of the differences between the structure of $tRNA_4^{Val}$ and the $tRNA_4^{Val}$-like genes are the substitution of a C-G base-pair for a U-A base-pair in the anticodon stem of the potential transcript of the $tRNA_4^{Val}$-like genes (Figure 12). Such coupled substitutions would be very unlikely if the $tRNA_4^{Val}$-like genes were the product of mutations in a non-functional gene. Third, all the nucleotides characteristic of eukaryotic valine tRNAs (Results and Discussion, Section II.D) would be present in any transcript of the $tRNA_4^{Val}$-like gene. Such a transcript would, however, be the first valine tRNA with an A57 rather than a G57 residue (Gauss and Sprinzl, 1981). While a G at position 57 is most common, many tRNAs have been sequenced that contain an A at that position.

If the $tRNA_4^{Val}$-like genes were expressed the tRNA produced would have, like $tRNA_4^{Val}$, an IAC anticodon. There are several examples in the literature of organisms that contain several slightly different tRNAs,

each with the same anticodon ("isocoding tRNA"). For example, yeast $tRNA_1^{Ser}$ and $tRNA_2^{Ser}$ both have an IAG anticodon but differ from each other at 3 other sites, one tRNA has a G at position 57 the other having an A residue at that position (Zachau et al., 1966). Bovine $tRNA^{Trp}$ is a mixture of up to 3 subspecies differing in the extent of some base modifications and in the nucleotides present at 3 positions (Fournier et al., 1978). Two of the sites of variability (C/U 16, G/A 57) are the same as those between $tRNA_4^{Val}$ and the hypothetical isocoding $tRNA^{Val}$. Keith and Dirheimer (1980) reported that $tRNA^{Phe}$ of Bombyx mori posterior silk gland is a mixture of two species; 20% of the tRNA has an A at position 57 the rest contains a G residue at that position.

In several organisms new isocoding species of tRNA are produced during cell differentiation. In bovine lens the epithelial cells contain a single $tRNA_2^{Phe}$ isoacceptor with a sequence identical to that of beef liver $tRNA^{Phe}$. Lens epithelial cells differentiate to produce lens fibre cells. The differentiated cells contain $tRNA_1^{Phe}$ as well as $tRNA_2^{Phe}$. The two tRNAs differ by a single nucleotide. $tRNA_2^{Phe}$ contains G57 while $tRNA_1^{Phe}$ contains an A residue at this position (Lin et al., 1980). It is noteworthy how frequently isocoding tRNAs differ in the nucleotide present at position 57. During development of the Bombyx posterior silk gland a new species of $tRNA^{Ala}$ is produced ($tRNA_1^{Ala}$) in addition to the $tRNA_2^{Ala}$ found in all Bombyx tissues. The two species differ by a single nucleotide. The C residue present at position 40 of $tRNA_2^{Ala}$ is replaced by a U or Ψ residue in $tRNA_2^{Ala}$. This introduces a mismatched base-pair into the anticodon stem of the latter species (Sprague et al., 1977). What is the function of the isocoding tRNAs? Bovine lens fibre cells and Bombyx posterior silk gland cells make large amounts of specific proteins, lens crystallins and silk fibroin respectively. Lens crystallins

are relatively rich in phenylalanine while fibroin contains 29% alanine. Both Lin et al. (1980) and Sprague et al. (1977) speculate that the new tRNA isoacceptors produced during the differentiation of lens fibre and posterior silk gland cells are required for the most efficient synthesis of the major proteins produced by these specialized cells. Perhaps the mRNAs for these proteins contain codons that require the new isocoding tRNAs for their efficient translation. Evidence for such context effects has been found in E. coli (Bossi and Roth, 1980). It is possible that the $tRNA_4^{Val}$-like genes of Drosophila are expressed only in certain tissues or only at specific periods during development. The pattern of valine isoacceptors in Drosophila revealed by RPC-5 chromatography is the same in first and third instar larvae and in adult flies (White et al., 1973a). However, the pattern of isoacceptors present in the embryonic (egg) and pupal stages remains unknown.

In summary, $tRNA_4^{Val}$ or $tRNA_4^{Val}$-like genes have been isolated and sequenced from 3 of the 4 sites of $tRNA_4^{Val}$ hybridization to Drosophila polytene chromosomes. The genes of plasmid pDt55 originate from a major site of $tRNA_4^{Val}$ hybridization (70BC). The two $tRNA_4^{Val}$ genes cloned in this plasmid correspond to the sequence of $tRNA_4^{Val}$. No $tRNA_4^{Val}$ genes have been cloned from the other major site of $tRNA_4^{Val}$ hybridization, 56D. The $tRNA_4^{Val}$-like genes of plasmids pDt92R, pDt120R and pDt14 originate at the minor sites of $tRNA_4^{Val}$ hybridization to polytene chromosomes (90BC and 89BC). Transcription of these genes would produce a tRNA that differs at 4 sites from the nucleotide sequence of $tRNA_4^{Val}$.

V. The Nucleotide Sequence of $tRNA_{3b}^{Val}$ Genes

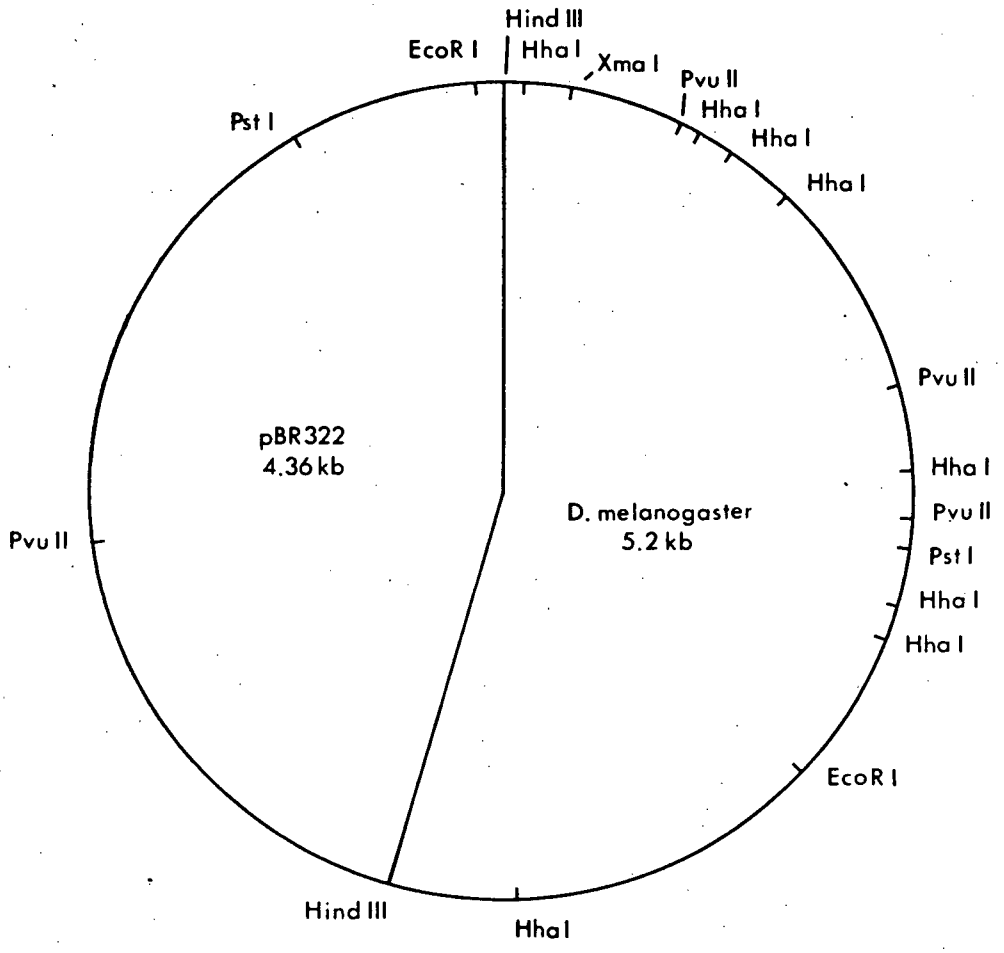
A. The Nucleotide Sequence of the $tRNA_{3b}^{Val}$ Gene of pDt78R

The recombinant plasmid pDt78R contains a 5.2 kb fragment of Drosophila DNA that can hybridize to $tRNA_{3b}^{Val}$ (Table III). The Drosophila DNA of this plasmid was shown by in situ hybridization to have originated from the 84D region of the polytene chromosomes, one of the two major sites of $tRNA_{3b}^{Val}$ hybridization (Table I). Thus pDt78R was predicted to contain a gene for $tRNA_{3b}^{Val}$. The restriction map of pDt78R is presented in Figure 33 (the sites of cleavage by Hha I were determined by D. St.Louis). Particularly noteworthy is the presence of a single Xma I site in the plasmid. The nucleotide sequence of $tRNA_{3b}^{Val}$ indicates that $tRNA_{3b}^{Val}$ genes should contain the restriction site for this enzyme. The strategy developed to sequence the $tRNA_4^{Val}$ genes of pDt55 was successfully applied to the sequence analysis of the $tRNA_{3b}^{Val}$ gene of pDt78R. The sequencing of pDt78R is summarized in Figure 34. Cleavage of pDt78R with the restriction endonuclease Hha I produces a 640 bp fragment of DNA (the third largest fragment present in the digest) that contains the $tRNA_{3b}^{Val}$ gene. This fragment provided a convenient source of DNA for many of the sequencing experiments.

The nucleotide sequence of the $tRNA_{3b}^{Val}$ gene of pDt78R is presented in Figure 35. The sequence of the gene corresponds exactly to that of $tRNA_{3b}^{Val}$. The 3'-flanking sequence of the gene contains a series of 5 consecutive T residues (in the non-transcribed strand). Like similar sequences in other eukaryotic tRNA genes, this sequence probably serves as a transcription termination signal. The five nucleotides immediately adjacent to the 5'-end of the gene together with the first nucleotide of the gene form the sequence CACAAG. This sequence recurs within the gene itself beginning at position 40 (nucleotide 98 in Figure 35). As previously

Figure 33. The restriction map of plasmid pDt78R.

Recombinant plasmid pDt78R consists of a 5.2 kb fragment of Drosophila DNA cloned into the single Hind III site of the plasmid vector pBR322. A restriction endonuclease cleavage map of pDt78R was constructed by the multiple enzyme digest method described by Danna (1980). The Xma I site marks the location of the plasmid's single $tRNA_{3b}^{Val}$ gene. The Hha I restriction sites in the pBR322 DNA of the plasmid are not shown.


Hind III
EcoR I
Hha I
Xma I
Pvu II
Hha I
Hha I
Hha I
Pst I
Pvu II
pBR322
4.36 kb
Hha I
D. melanogaster
5.2 kb
Pvu II
Pvu II
Pst I
Hha I
Hha I
EcoR I
Hind III
Hha I

pDt78R

Figure 34. The strategy used to sequence the $tRNA_{3b}^{Val}$ gene of pDt78R.

Each arrow represents a nucleotide sequence determined by the method of Maxam and Gilbert (1 cm = 20 nucleotides). For the sequencing, pDt78R DNA was end-labelled at cleavage sites for the restriction endonucleases listed in the Figure. Arrows pointing to the left indicate that the sequence of the non-transcribed strand (shown in the top panel of the Figure) was determined. Arrows pointing to the right indicate the sequence of the complementary strand was determined.

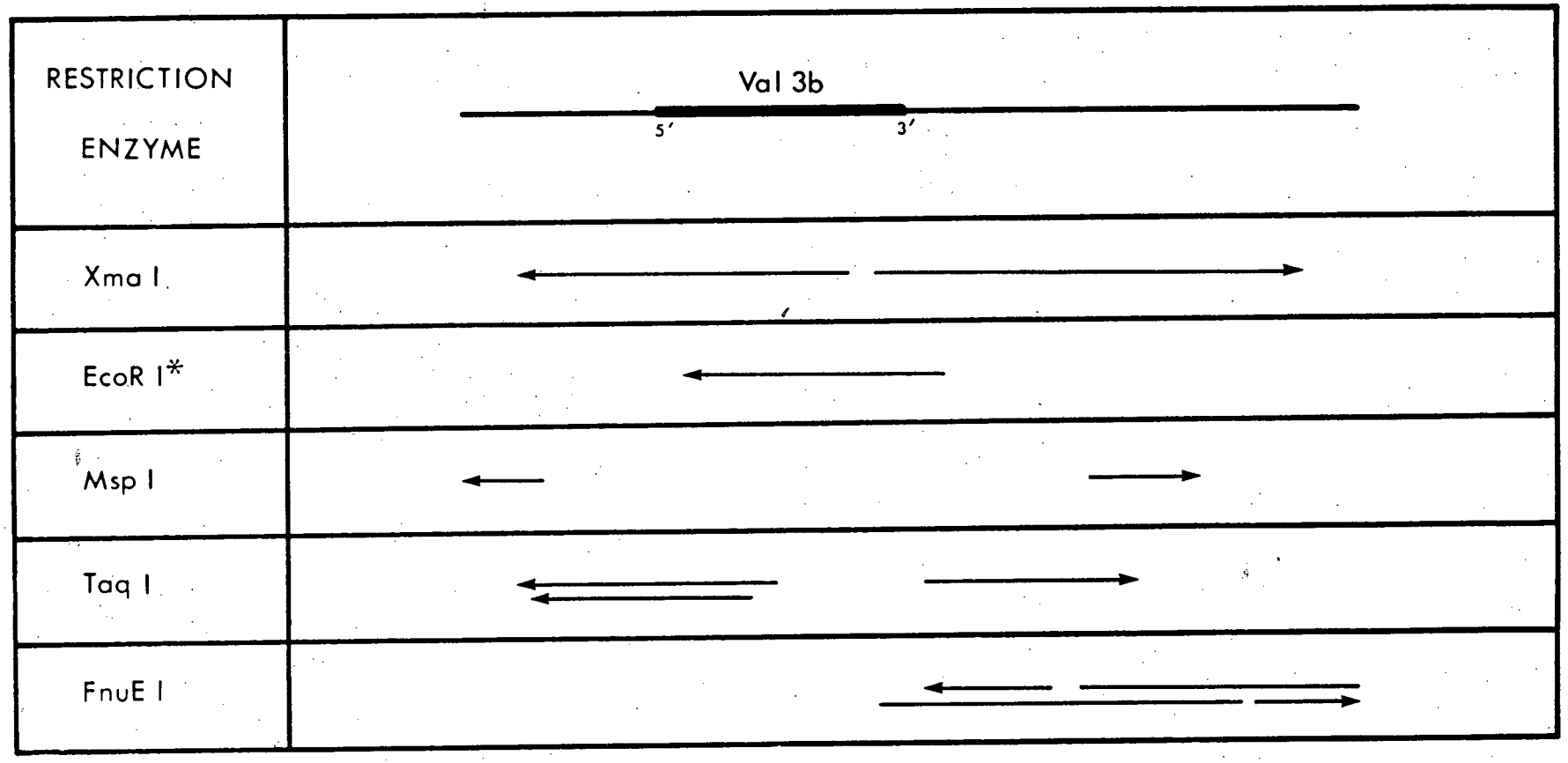

RESTRICTION ENZYME
Val 3b
5′
3′
Xma I
EcoR I*
Msp I
Taq I
FnuE I

Figure 35. The nucleotide sequences of two segments of Drosophila DNA containing $tRNA_{3b}^{Val}$ genes.

The nucleotide sequences of the non-transcribed strands of two $tRNA_{3b}^{Val}$ genes, one found in pDt78R the other sequenced by Silberklang, are shown. The $tRNA_{3b}^{Val}$ genes are underlined.

The Nucleotide Sequence of pDt78R

```
          10         20         30         40         50         60         70         80         90         100
TCCGTGAATT TATACTAGAC TTTATAATAT AGGTCTTGTG ATGTCAGCAC CGCCACAAGT TTCCGTAGTG TAGCGGTTAT CACGTGTGCT TCACACGCAC

          110        120        130        140        150        160        170        180        190        200
AAGGTCCCCG GTTCGAACCC GGGCGGGAAC ATGCGATCCT TTTTGAATTA ATTTATCAAT AATTATTTTG TATTATTTTA CGTTTTTAGT ATGTGGGAAA

          210        220        230        240        250        260
AAGGACTGTT CAGCCAGAAA CGAAGTTTTT CCGTTCAGAT GTGTTGTTGG GATGCATATG TGAAGGGA
```

Val

THE NUCLEOTIDE SEQUENCE OF THE tRNA3b GENE DETERMINED BY SILBERKLANG

```
          10         20         30         40         50         60         70         80         90         100
CGGCCGTCTC TTAAGAGTTT CCGTAGTGTA GCGGTTATCA CGTGTGCTTC ACACGCACAA GGTCCCCGGT TCGAACCCGG GCGGGAACAG TCGAAATAGT

          110        120        130        140
TTTTGTCTTT TTTTATTTCA TTACTCACTA GTTATTTTGC GATAATATT
```

described (Section IV.B, Results and Discussion), duplications of tRNA gene sequences in the 5'-flanking sequences of the genes are found for one of the $tRNA_4^{Val}$ genes of pDt55 and have been noted in the tRNA genes of a number of organisms by Santos and Zasloff (1981). These sequences may influence the site at which transcription of these tRNAs genes is initiated. An inverted complement of the CACAAG hexanucleotide occurs 24 bp upstream from the $tRNA_{3b}^{Val}$ gene of pDt78R (nucleotides 35-40, Figure 35). Hybridization of this complementary sequence to either of the two CACAAG sequences would produce hair-pin structures in the DNA. With loop-outs, two longer inverted complementary regions, one containing the intragenic CACAAG sequence and the other its inverted complement, can be recognized. The region within the gene extends from nucleotide 86 to nucleotide 103 of Figure 35. An imperfect inverted complementary sequence to this region is found between nucleotides 35 and 50 of Figure 35. The significance, if any, of these regions is unknown.

The CACAAG sequence is also found in the 5'-flanking sequence of another Drosophila tRNA gene. The sequence occurs 34 bp upstream from the 5'-end (in the non-transcribed strand) of the second $tRNA_4^{Val}$ gene of plasmid pDt55 (nucleotide 933-938, Figure 30). The association of the CACAAG sequence with two different $tRNA^{Val}$ genes from widely separated chromosomal locations suggests that the sequence plays some role in tRNA gene expression. However, the sequence is not found in the 5'-flanking sequences of the other Drosophila tRNA genes discussed in this thesis.

B. Other $tRNA_{3b}^{Val}$ Genes: Comparisons with the $tRNA_{3b}^{Val}$ Gene of pDt78R

In addition to the $tRNA_{3b}^{Val}$ gene of pDt78R three other $tRNA_{3b}^{Val}$ genes have been sequenced. One of these genes is located at the 92B site on the polytene chromosome, a major site of $tRNA_{3b}^{Val}$

hybridization. This gene was sequenced by Dr. M. Silberklang (personal communication). The other two genes were cloned and sequenced in this laboratory. One gene was cloned in the recombinant plasmid pDt48, the other in plasmid pDt41R (Table III). Both of these genes originate from the minor site of $tRNA_{3b}^{Val}$ hybridization, 90BC (Table I). Thus $tRNA_{3b}^{Val}$ genes from each of this tRNA's chromosomal loci have been sequenced.

The nucleotide sequence of a $tRNA_{3b}^{Val}$ gene from the 92B site was determined by Dr. M. Silberklang and is shown in Figure 35. The transcript of this gene would have the same sequence as $tRNA_{3b}^{Val}$. Thus the sequence of this gene is identical to that of the $tRNA_{3b}^{Val}$ gene of pDt78R. The flanking sequences of the genes from the two plasmids, however, show no significant homology other than the 5 adjacent T residues that are probably termination signals for RNA polymerase III.

Recombinant plasmids pDt48 and pDt41R contain segments of Drosophila DNA (2.4 kb and 2.0 kb long respectively) that can hybridize with $tRNA_{3b}^{Val}$. The $tRNA_{3b}^{Val}$ gene of pDt48 was sequenced by Dr. A. Delaney, that of pDt41R was sequenced by J. Leung. Their results are shown in Figure 36. The sequenced portion of pDt48 contains two tRNA genes, one for a tRNA similar to $tRNA_{3b}^{Val}$, the other for a $tRNA_{(UGG)}^{Pro}$. A transcript of the $tRNA_{3b}^{Val}$-like gene of pDt48 would differ from $tRNA_{3b}^{Val}$ at 4 sites. Nucleotides C5, C16, G68 and G69 of $tRNA_{3b}^{Val}$ would be replaced by U5, U16, A68 and A69 in a transcript of the $tRNA_{3b}^{Val}$-like gene (Figure 20). Except for the C16 to U16 transition these differences are not the same as those between $tRNA_{4}^{Val}$ and hypothetical transcripts of the $tRNA_{4}^{Val}$-like genes of plasmids pDt120R, pDt92R or pDt14. It is perhaps significant that, unlike $tRNA_{3b}^{Val}$, a transcript of the $tRNA_{3b}^{Val}$-like gene would have a perfectly base-paired aminoacyl stem. With the possible exception of C16, the sequence of

Figure 36. The nucleotide sequences of segments of the Drosophila DNA inserts of plasmids pDt48 and pDt41R.

The non-transcribed strand of the $tRNA_{3b}^{Val}$-like genes is shown. The second gene, coding for $tRNA^{Pro}$, is of opposite polarity to the first. Homologies between the sequences from pDt48 and pDt41R are marked by asterisks. Dashes were inserted into the sequences wherever necessary to ensure maximum homology. A 14 bp sequence from position 325 to 338 of pDt48 is absent in pDt41R. A large segment of pDt41R, between positions 409 and 483, has not yet been sequenced. tRNA genes are overlined. Differences between the $tRNA_{3b}^{Val}$-like genes of pDt48 and the $tRNA_{3b}^{Val}$ gene of pDt78R are underlined.

The Nucleotide Sequences of pDt48 and pDt41R - Homologies between the two sequences are marked by asterisks.

```
                10        20        30        40        50        60        70        80        90       100
48  AAAATAAATCTAAGTATGCAACTTTGGCAAGATCAGAAGAATAAGTTAAACGGCCATTGAAAATGTGTTTCTCCAATTGTTCTAAAAAAAATGTAATAAA

               110       120       130       140       150       160       170       180       190       200
48  ATTTTAAAATAAGCAAATAGTTCCACAGGAAACTAGAGTCATGCAGGTTAGTCCTTTTTGTGTTGTGTGAACACAATACGCTATACTGTTAGTTTAAACT
                                                                ********************************* ***** ********
41R                                                             TGTGTTGTGTGAACACAATACGCTATAGTGTTATTTTAAACT

               210       220       230       240       250       260       270       280       290       300
48  AAATGTTTCATGATGTTTTCGTAGTGTAGTGGTTATCACGTGTGCTTCACACGCACAAGGTCCCCGGTTCGAACCCGGGCGAAAACAGATTGATTTTTTT
    ***** ******* ************************************************************************** ** ** ******
41R AAATGATTCATGACGTTTTCGTAGTGTAGTGGTTATCACGTGTGCTTCACACGCACAAGGTCCCCGGTTCGAACCCGGGCGAAAACAAATCGACTTTTTT

               310       320       330       340       350       360       370       380       390       400
48  TTAATTTCTTTTTACATTTTCGATGAATCTTAGGGTTGAAAACGGTAACACAAATAAAATATTTTAATACCCTTAAGGAATAATTGAAAAAAGAC--CGA
    * ****************  *                ************************** **** *********  * ********* *   ***
41R TAAATTTCTTTTTACATTTTTTAC---Deletion---AAAACGGTAACACAAATAAAATATTTGAATATCCTTAAGGATAATTTGAAAAAATAAAACGA

               410       420       430       440       450       460       470       480       490       500
48  TGCTATAGAAACGTACCAATATTTGAATAAGCCAATGGGGTTGAAATCCATACATATTGTTTACGGGTCAAACCATTTACTTTCTATAGTTTAAATATTT
    *********                                                                          ******************
41R TGCTATAGA--------------------------Not Yet Sequenced--------------------------CTATAGTTTAAATATTT

               510       520       530       540       550       560       570       580       590       600
48  TCTTTAATTTTCAGAAAAATTAGCAAAGAAAAAATTTGTACGTGCGGTTGA--GTTGAGCAATAAAAACAGTACAGCTGGGCTCAACCGGGATTTGAACC
    ************** ****** ********************* ****** *****  ***** *** ****  *  ****************************
41R TCTTTAATTTTCAGGAAAATTTGCAAAGAAAAAATTTGAACGTGCAGTTGACCGTTGAACAACAAAA-TAACACAGCTGGGCTCAACCGGGATTTGAACC

               610       620       630       640       650       660       670       680       690       700
48  CGGGACCTCTCGCACCCAAAGCGAGAATCATACCCCTAGACCATTGAGCCTCATAACGAGTATGCAGCTGTTCGCCAGTTTCCAACTAGGCAGAGCAAGG
    *************************
41R CGGGACCTCTCGCACCCAAAGCGAG

               710       720       730       740       750       760       770       780       790       800
48  CATTTTCTTGAGTGAGCCTCTAAAGACAAACAAAAATCATTTCTTCGATGTAACAATCTAAAATATATATTTAGTAAACATAGAACAATATCTCTCAGCG

48  TCGCTAAGTG
```

$tRNA_{3b}^{Val}$ shows no heterogeneity at the sites at which it differs from the $tRNA_{3b}^{Val}$-like genes (Figure 21, 22). It is concluded, therefore, that $tRNA_{3b}^{Val}$ does not contain transcripts of the $tRNA_{3b}^{Val}$-like genes. Preliminary sequencing data indicate these genes do not code for $tRNA_{3a}^{Val}$. The transcript of the $tRNA_{3b}^{Val}$-like genes would contain all the nucleotides characteristic of eukaryotic valine tRNAs (Results and Discussion, Section II.E) and these genes may well be expressed as one of the minor $tRNA^{Val}$ species (Figure 7). There is no significant homology between the DNA sequences flanking the $tRNA^{Val}$ genes of pDt78R and pDt48.

The proline tRNA gene of pDt48 is located 287 bp downstream from the $tRNA_{3b}^{Val}$-like gene and is of opposite polarity to the latter. Neither of the two major $tRNA^{Pro}$ species of Drosophila (White et al., 1973a) has been sequenced so a comparison between the gene sequence and the proline tRNA sequence cannnot yet be made. However, the $tRNA_{(UGG)}^{Pro}$ gene (non-transcribed strand) has a sequence very similar to that of the corresponding tRNA of mouse and chicken (95% homology)(Gauss and Sprinzl, 1981).

The sequenced portion of pDt41R is very similar, but not identical, to corresponding sequences in pDt48 (Figure 36). The $tRNA_{3b}^{Val}$-like genes and most of the $tRNA^{Pro}$ genes (the sequencing of the $tRNA^{Pro}$ gene of pDt41R is not yet completed) of the two plasmids are identical. Outside the genes, the two plasmids are very similar (about 86% homology) but are more divergent than the homologous regions around the $tRNA_{4}^{Val}$-like genes of plasmids pDt92R and pDt120R. Most of the differences between the sequenced portions of pDt48 and pDt41R are due to scattered base changes or the insertion or deletion of one or two base-pair segments of DNA. There is, however, a 14 bp sequence present between nucleotides 325 and 339 of pDt48 (Figure 36) that is not found in pDt41R. It seems likely that both sequences are present in the Drosophila genome. The Drosophila DNA of pDt41R is a Hind

III fragment 2.0 kb long. A fragment of this size was detected by hybridizing ^{125}I-labelled $tRNA_{3b}^{Val}$ to a Southern blot of Drosophila DNA fragments generated by cleavage of genomic DNA with Hind III (Tener et al., 1980). pDt48 contains a 2.4 kb Hind III fragment of Drosophila DNA. No fragment this size carrying $tRNA_{3b}^{Val}$ genes was detected in Hind III-cleaved Drosophila DNA (Tener et al., 1980). The 2.0 kb fragment of DNA cloned in pDt41R is, therefore, likely an "allele" of the fragment cloned in pDt48. This "allele" may not have been detected in Hind III-cleaved Drosophila DNA because it is present in low frequency in the population of flies from which the DNA was isolated. If the 2.0 and 2.4 kb Hind III fragments were derived from different repeat units of a DNA segment duplicated in the Drosophila genome, they should always be present in a 1:1 ratio in the Drosophila DNA. In this case both the 2.4 and 2.0 kb fragments should have been detected in the Hind III cleaved genomic DNA.

In summary, the $tRNA_{3b}^{Val}$ genes present at the major sites of $tRNA_{3b}^{Val}$ hybridization to Drosophila polytene chromosomes (pDt78R from 84D, Silberklang's gene sequence from 92B) correspond to the sequence of the $tRNA_{3b}^{Val}$. The genes present at the minor hybridization site (pDt48 and pDt41R from 90BC) are similar, but not identical, to the $tRNA_{3b}^{Val}$ sequence. These results parallel those obtained for the cloned $tRNA_4^{Val}$ genes. Perhaps the genes for $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ present at the major hybridization sites are isolated from the closely related $tRNA_4^{Val}$-like and $tRNA_{3b}^{Val}$-like genes present at the minor sites. The weaker hybridization of $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ to the minor sites for these tRNAs would then be due to the decreased homology between the genes at these sites and the tRNA probes used to locate them. If this hypothesis is true the genes for $tRNA_4^{Val}$ and $tRNA_{3b}^{Val}$ would not be so widely dispersed on the chromosome as previously thought.

Bibliography

Altwegg, M. (1980) Inaugural Dissertation, Philosophischen Fakultat II Der Universitat Zurich.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457-465.

Arcari, P. and Brownlee, G.G. (1980) Nucl. Acids Res. 8, 5207-5212.

Barrio, J.R., Secrist III, J.A. and Leonard N.J. (1972) Biochem. Biophys. Res. Commun. 46, 597-604.

Beckmann, J.S., Johnson, P.F. and Abelson, J. (1977) Science 196, 205-208.

Beermann, W. (1972) in Developmental Studies on Giant Chromosomes (Beermann, W., ed.), Results and Problems in Cell Differentiation Vol. 4, pp 1-31, Springer-Verlag, New York.

Benne, R. and Hershey, J.W.B. (1978) J. Biol. Chem. 253, 3078-3087.

Biernat, J., Ciesiolka,J., Gornicki, P., Adamiak, R.W., Krzyzosiak, W.J. and Wiewiorowski, M. (1978) Nucl. Acids Res. 5, 789-804.

Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) Cell 19, 27-35.

Boguski, M.S., Hieter, P.A. and Levy, C.C. (1980) J. Biol. Chem. 255, 2160-2163.

Bonitz, S.G. and Tzagoloff, A. (1980) J. Biol. Chem. 255, 9075-9081.

Borst, P. and Grivell, L.A. (1981) Nature 290, 443-444.

Bossi, L. and Roth, J.J. (1980) Nature 286, 123-127.

Brenner, D.J., Fournier, M.J. and Doctor, B.P. (1970) Nature 227, 448-451.

Britten, R.J., Graham, D.E., and Neufeld, B.R. (1974) Methods in Enzymology 29, 363-418.

Brown, N.L. (1979) in Companion to Biochemistry (Bull, A.T., Legnado, J.R., Thomas, J.O. and Tipton, K.F., eds.) Vol. 2, pp 1-48, Longman, London and New York.

Cedergren, R.J., Sankoff, D., LaRue, B. and Grosjean, H. (1981) Crit. Rev. Biochem. 11, 35-104.

Champe, S.P. and Benzer, S. (1962) Proc. Natl. Acad. Sci. USA 48, 532-546.

Chen, E.Y. and Roe, B.A. (1977) Biochem. Biophys. Res. Commun. 78, 631-640.

Chen, E.Y. and Roe, B.A. (1978) Biochem. Biophys. Res. Commun. 82, 235-245.

Clark, B.F.C. (1978) in Transfer RNA (Altman, S., ed.) pp 14-47, MIT Press Cambridge, Mass. and London.

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

Clarkson, S.G., Birnsteil, M.L. and Prudom, I.F. (1973b) J. Mol. Biol. 79, 411-429.

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

Clarkson, S.G., Kurer, V. and Smith, H.O. (1978) Cell 14, 713-724.

Cramer, F., von der Haar, F. and Igloi, G.L. (1980) in Transfer RNA: Structure, Properties and Recognition (Schimmel, P.R., Soll, D., and Abelson, J.N., eds.) pp 267-279, Cold Spring Harbor Laboratory.

Cribbs, D.L. (1979) M.Sc. Thesis, University of British Columbia.

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

Dahlberg, J.E. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 507-516, Cold Spring Harbor Laboratory.

Danna, K.J. (1980) Methods in Enzymology 65, 449-467.

de Bruijn, M.H.L., Schreier, P.H., Eperon, I.C., Barrell, B.G., Chen, E.Y., Armstrong, P.W., Wong, J.F.H. and Roe, B.A. (1980) Nucl. Acids Res. 8, 5213-5222.

DeFranco, D., Schmidt, O., and Söll, D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 3365-3368.

Dirheimer, G., Keith, G., Silber, A.-P. and Martin, R.P. (1979) Transfer RNA: Structure, Properties and Recognition (Schimmel, P., Söll, D. and Abelson, J.N., eds.) pp 19-41, Cold Spring Harbor Laboratory.

Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucl. Acids Res. 4, 2527-2538.

Dudler, R., Egg, A.H., Kubli, E., Artavanis-Tsakonas, S., Gehring, W.J., Steward, R. and Schedl, P. (1980) Nucl. Acids Res. 13, 2921-2937.

Dudler, R.K. (1981) Inaugural Dissertation, Zoologisches Institut der Universitat Zurich.

Duester, G.L. and Holmes, W.M. (1980) Nucl. Acids Res. 8, 3793-3807.

Duester, G., Campen, R.K. and Holmes, W.M. (1981) Nucl. Acids Res. 9, 2121-2139.

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

Dunn, R., Delaney, A.D., Gillam, I.C., Hayashi, S., Tener, G.M., Grigliatti, T., Misra, V., Spurr, M.G., Taylor, D.M. and Miller, R.C. Jr. (1979) Gene 7, 197-215.

Elder, R.T., Uhlenbeck, O.C. and Szabo, P. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abeleson, J.N. and Schimmel, P.R., eds.) pp 317-323, Cold Spring Harbor Laboratory.

Endow, S.A. and Roberts R.J. (1977) J. Mol. Biol. 112, 521-529.

England, T.E. and Uhlenbeck, O.C. (1978) Nature 275, 560-561.

Feldmann, H. (1976) Nucl. Acids Res. 3, 2379-2386.

Feldmann, H. (1977) Nucl. Acids Res. 4, 2831-2841.

Fersht, A.R. (1980) in Transfer RNA: Structure, Properties and Recognition (Schimmel, P.R., Söll, D. and Abelson, J.N., eds.) pp 247-254.

Fournier, M., Labouesse, J., Dirheimer, G., Fix, C. and Keith, G. (1978) Biochim. Biophys. Acta 521, 198-208.

Fowlkes, D.M. and Shenk, T. (1980) Cell 22, 405-413.

Friedman, S. (1973) Nature New Biol. 244, 18-20.

Gall, J.G. and Pardue, M.L. (1969) Proc. Natl. Acad. Sci. USA 63, 378-383.

Gallant, J.A. (1979) Ann. Rev. Genet. 13, 393-415.

Garber, R.L. and Altman, S. (1979) Cell 17, 389-397.

Gauss, D.H. and Sprinzl, M. (1981) Nucl. Acids Res. 9, r1-r23.

Ghosh, R.K. and Deutscher, M.P. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 59-69, Cold Spring Harbor Laboratory.

Goldman, E., Holmes, W.M. and Hatfield, G.M. (1979) J. Mol. Biol. 129, 567-585.

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

Grantham, R., Gautier, C., Gouy, M. Jacobzone, M. and Mercier, R. (1981) Nucl. Acids Res. 9, r43-r74.

Gupta, R.C. and Randerath, K. (1977) Nucl. Acids Res. 4, 3441-3454.

Gupta, R.C. and Randerath, K. (1979) Nucl. Acids Res. 6, 3443-3458.

Haenni, A.-L. and Chapeville, F. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 539-556, Cold Spring Harbor Laboratory.

Hayashi, S., Gillam, I.C., Delaney, A.D., Dunn, R.D., Tener, G.M., Grigliatti, T.A. and Suzuki, D.T. (1980) Chromosoma 76, 65-84.

Hayashi, S., Gillam, I.C. and Tener, G.M. (1981a), Manuscript in preparation.

Hayashi, S., Addison, W.R., Gillam, I.C., Grigliatti, T.A. and Tener, G.M. (1981b) Chromosoma 82, 385-397.

Hershey, N.D. and Davidson, N. (1980) Nucl. Acids Res. 8, 4899-4910.

Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L. and Zamecnik, P.C. (1958) J. Biol. Chem. 231, 241-257.

Hofstetter, H., Kressmann, A. and Birnstiel, M.L. (1981) Cell 24, 573-585.

Holbrook, S.R., Sussman, J.L., Warrant, R.W. and Kim, S.-H. (1978) J. Mol. Biol. 123, 631-660.

Holley, R.W., Apgar, J. Everett, G.A., Madison, J.T., Marquisee, M., Merrill, S.H., Penswick, J.R. and Zamir, A. (1965) Science 147, 1462-1465.

Holmes, W.M., Hatfield, G.W. and Goldman, E. (1978) J. Biol. Chem. 253, 3482-3486.

Hosbach, H.A. and Kubli, E. (1979a) Mech. Ageing Dev. 10, 131-140.

Hosbach, H.A. and Kubli, E. (1979b) Mech. Ageing Dev. 10, 141-149.

Hosbach, H.A., Silberklang, M. and McCarthy, B.J. (1980) Cell 21, 169-178.

Hovemann, B., Sharp, S., Yamada, H. and Söll, D. (1980) Cell 19, 889-895.

Huang, S.L. and Baker, B.S. (1976) Mut. Res. 34, 407-414.

Igloi, G.L. and Cramer, F. (1978) in Transfer RNA (Altman, S., ed.) pp 294-349 MIT Press, Cambridge, Mass. and London.

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

Jank, P., Shindo-Okada, N., Nishimura, S. and Gross, H.J. (1977a) Nucl. Acids Res. 4, 1999-2008.

Jank, P., Riesner, D. and Gross, H.J. (1977b) Nucl. Acids Res. 4, 2009-2020.

Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) Nucl. Acids Res. 1, 331-353.

Judson, H.F. (1979) The Eighth Day of Creation, pp 285-293, Simon and Schuster, New York.

Keith, G. and Dirheimer, G. (1980) Biochem. Biophys. Res. Commun. 92, 109-115.

Khym, J.X. and Uziel, M. (1968) Biochemistry 7, 422-426.

Kim, S.-H. (1978) in Transfer RNA (Altman, S., ed.) pp 248-293 MIT Press, Cambridge, Mass. and London.

Kimura, K., Nakanishi, M., Yamamoto, T. and Tsuboi, M. (1977) J. Biochem. 81, 1699-1703.

Knapp, G., Ogden, R.C., Peebles, C.L. and Abelson, J. (1979) Cell 18, 37-45.

Koch, W., Edwards, K. and Kossel, H. (1981) Cell 25, 203-213.

Kochel, H.G., Lazarus, C.M., Basak, N. and Kuntzel, H. (1981) Cell 23, 625-633.

Kochetkov, N.K., Shibaev, V.N. and Kost, A.A. (1971) Tetrahedron Lett. 22, 1993-1996.

Koski, R.A., Clarkson, S.G., Kurjan, J., Hall, B.D. and Smith, M. (1980) Cell 22, 415-425.

Kressmann, A., Hofstetter, H., DiCapua, E., Grosschedl, R. and Birnstiel, M.L. (1979) Nucl. Acids Res. 7, 1749-1763.

Krupp, G. and Gross, H.J. (1979) Nucl. Acids Res. 6, 3481-3490.

Krzyzosiak, W.J., Biernat, J., Ciesiolka, J., Gornicki, P. and Wiewiorowski, M. (1979) Polish J. Chem. 53, 243-252.

Krzyzosiak, W.J., Biernat, J., Ciesiolka, J., Gulewicz, K. and Wiewiorowski, M. (1981) Nucl. Acids Res. 9, 2841-2851.

Kurjan, J., Hall, B.D., Gillam, S. and Smith, M. (1980) Cell 20, 701-709.

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

Lee, F. and Yanofsky, C. (1977) Proc. Natl. Acad. Sci. USA 74, 4365-4369.

Lefevre, G. Jr. (1976) in The Genetics and Biology of Drosophila (Ashburner, M. and Novitski, E., eds.) Vol. 1a, pp 31-66, Academic Press, London, New York, San Francisco.

Li, M. and Tzagoloff, A. (1979) Cell 18, 47-53.

Lin, F.-K., Furr, T.D., Chang, S.H., Horwitz, J., Agris, P.F. and Ortwerth, B.J. (1980) J. Biol. Chem. 255, 6020-6023.

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

Lusk, J.E., Williams, R.J.P. and Kennedy, E.P. (1968) J. Biol. Chem. 243, 2618-2624.

Lustig, F., Elias, P., Axberg, T., Samuelsson, T., Tittawella, I. and Lagerkvist, U. (1981) J. Biol. Chem. 256, 2635-2643.

Mao, J., Schmidt, O. and Söll, D. (1980) Cell 21, 509-516.

Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology 65, 499-560.

Mazzara, G.P. and McClain, W.H. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 3-27, Cold Spring Harbor Laboratory.

McDonell, M.W., Simon, M.N. and Studier, F.W. (1977) J. Mol. Biol. 110, 119-146.

Melton, D.A., De Robertis, E.M. and Cortese, R. (1980) Nature 284, 143-148.

Mirzabekov, A.D., Grunberger, D., Krutilina, A.I., Holy, A., Bayev, A.A. and Sorm, F. (1968) Biochim. Biophys. Acta 166, 75-81.

Mitra, S.K., Lustig, F., Akesson, B., Lagerkvist, U. and Strid, L. (1977) J. Biol. Chem. 252, 471-478.

Mitra, S.K., Lustig, F., Akesson, B., Axberg, T., Elias, P. and Lagerkvist, U. (1979) J. Biol. Chem. 254, 6397-6401.

Montoya, J., Ojala, D. and Attardi, G. (1981) Nature 290, 465-470.

Moras, D., Comarmond, M.B., Fischer, J., Weiss, R., Thierry, J.C., Ebel, J.P. and Giege, R. (1980) Nature 288, 669-674.

Morgan, E.A., Ikemura, T., Post, L.E. and Nomura, M. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J., and Schimmel, P.R., eds.) pp 259-266, Cold Spring, Harbor Laboratory.

Morris, R.W. and Herbert, E. (1970) Biochem. 9, 4819-4827.

Muller, F. and Clarkson, S.G. (1980) Cell 19, 345-353.

Munz, P., Leupold, U., Agris, P. and Kohli, J. (1981) Nature 294, 187-188.

Murgola, E.J. and Pagel, F.T. (1980) J. Mol. Biol. 138, 833-844.

Nakajima, N., Ozeki, H. and Shimura, Y. (1981) Cell 23, 239-249.

Nirenberg, M. and Leder, P. (1964) Science 145, 1399-1407.

Nishimura, S. (1978) in Transfer RNA (Altman, S., ed.) pp 168-195, MIT Press, Cambridge, Mass. and London.

Nishimura, S. (1979) in Transfer RNA: Structure, Properties and Recognition (Schimmel, P.R., Söll, D. and Abelson, J.N., eds.) pp 59-79, Cold Spring Harbor Laboratory.

Ogden, R.C., Beckmann, J.S., Abelson, J., Kang, H.S., Söll, D. and Schmidt, O. (1979) Cell 17, 399-406.

Ohashi, Z., Maeda, M., McCloskey, J.A. and Nishimura, S. (1974) Biochem. 13, 2620-2625.

Ojala, D., Montoya, J. and Attardi, G. (1981) Nature 290, 470-474.

Olson, M.V., Hall, B.D., Cameron, J.R. and Davis R.W. (1979) J. Mol. Biol. 127, 285-295.

Olson, M.V., Page, G.S., Sentenac, A., Piper, P.W., Worthington, M., Weiss, R.B., and Hall, B.D. (1981) Nature 291, 464-469.

Paule, M.R. (1981) Trends Biochem. Sci. 6, 128-131.

Peattie, D. (1979) Proc. Natl. Acad. Sci. USA 76, 1760-1764.

Peebles, C.L., Ogden, R.C., Knapp, G. and Abelson, J. (1979) Cell 18, 27-35.

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

Piper, P.W. (1975) Eur. J. Biochem. 51, 295-304.

Piper, P.W. and Clark, B.F.C. (1974) FEBS Lett. 47, 56-59.

Platt, T. (1981) Cell 24, 10-23

Polisky, B., Greene, P. Garfin, D.E., McCarthy, B.J., Goodman, H.M. and Boyer, H.W. (1975) Proc. Natl. Acad. Sci. USA 72, 3310-3314.

Prensky, W., Steffensen, D.M. and Hughes, W.L. (1973) Proc. Nat. Acad. Sci. USA 70, 1860-1864.

Pribnow, D. (1979) in Biological Regulation and Development (Goldberger, R.F., ed.) Vol. I, pp 219-277, Plenum Press, New York and London.

Razzell, W.E. (1963) Methods in Enzymology 6, 237-238.

Revel, M. (1977) in Molecular Mechanisms of Protein Biosynthesis (Weissbach, H. and Pestka, S., eds.) pp 245-321, Academic Press, New York, San Francisco and London.

Rich, A. and RajBhandary, U.L. (1976) Annu. Rev. Biochem. 45, 805-860.

Rich, A., Quigley, G.J., Teeter, M.M., Ducruix, A. and Woo, N. (1980) in Transfer RNA: Structure, Properties and Recognition (Schimmel, P.R., Söll, D. and Abelson, J.N., eds.) pp 101-113, Cold Spring Harbor Laboratory.

Richards, F.M. and Wyckoff, H.W. (1971) in The Enzymes 3rd ed. (Boyer, P.D., ed.) Vol. IV, pp 647-806, Academic Press, New York and London.

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

Roberts, R.J. (1974) J. Biol. Chem. 249, 4787-4796.

Robinson, R.R. and Davidson, N. (1981) Cell 23, 251-259.

Roe, B.A. (1975) Nucl. Acids Res. 2, 21-42.

Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. 13, 319-353.

Rossi, J., Egan, J., Berman, M.L. and Landy, A. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 221-244, Cold Spring Harbor Laboratory.

Rudkin, G.T. (1972) in Developmental Studies on Giant Chromosomes (Beermann, W., ed.) Results and Problems in Cell Differentiation Vol. 4, pp 58-80, Springer-Verlag, New York.

Rudloff, E. and Hilse, K. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1359-1367.

Sakonju, S., Brown, D.D., Engelke, D., Ng, S.-Y., Shastry, B.S. and Roeder, R.G. (1981) Cell 23, 665-669.

Santos, T. and Zasloff, M. (1981) Cell 23, 699-709.

Schedl, P., Primakoff, P. and Roberts, J. (1974) Brookhaven Symp. Biol. 26, 53-76.

Scherberg, N.H. and Refetoff, S. (1974) Biochim. Biophys. Acta 340, 446-451.

Schevitz, R.W., Podjarny, A.D., Krishnamachari, N., Hughes, J.J. and Sigler, P.B. (1980) in Transfer RNA: Structure, Properties, and Recognition (Schimmel, P.R., Söll, D. and Abelson, J., eds.) pp 133-143.

Schmidt, O., Mao, J., Ogden, R., Beckmann, J., Sakano, H., Abelson, J. and Söll, D. (1980) Nature 287, 750-752.

Schwarz, Z., Kossel, H., Schwarz, E. and Bogorad, L. (1981) Proc. Natl. Acad. Sci. USA 78, 4748-4752.

Sekiya, T., Mori, M., Takahashi, N. and Nishimura, S. (1980) Nucl. Acids Res. 8, 3809-3827.

Selker, E. and Yanofsky, C. (1980) Nucl. Acids Res. 8, 1033-1042.

Shindo-Okada, N., Kuchino, Y., Harada, F., Okada, N., and Nishimura, S. (1981) J. Biochem. 90, 535-544.

Silberklang, M., Gillum, A.M. and RajBhandary, U.L. (1979) Methods in Enzymology 59, 58-109.

Silverman, S., Schmidt, O., Söll, D. and Hovemann, B. (1979) J. Biol. Chem. 254, 10290-10294.

Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977) Nature 269, 833-836.

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

Soffer, R.L. (1974) Adv. Enzymol. 40, 91-139.

Soffer, R.L. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 493-505, Cold Spring Harbor Laboratory.

Southern, E.M. (1974) Anal. Biochem. 62, 317-318.

Southern, E.M. (1979) Anal. Biochem. 100, 319-323.

Sprague, K.U., Hagenbuchle, O. and Zuniga, M.C. (1977) Cell 11, 561-570.

Sprague, K.U., Larson, D. and Morton, D. (1980) Cell 22, 171-178.

Sprinzl, M., von der Haar, F. and Cramer, F. (1972) Europ. J. Biochem. 25, 262-266.

Sprouse, H.M., Kashdan, M., Otis, L. and Dudock, B. (1981) Nucl. Acids Res. 9, 2543-2547.

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

Steffensen, D.M. and Wimber, D.E. (1971) Genetics 69, 163-178.

Steinmetz, A., Mubumbila, M., Keller, M., Burkard, G., Weil, J.H., Driesel, A.J., Crouse, E.J., Gordon, K., Bohnert, H.-J. and Herrmann, R.G. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R. eds.) pp 281-286, Cold Spring Harbor Laboratory.

Tanaka, Y., Dyer, T.A. and Brownlee, G.G. (1980) Nucl. Acids Res. 8, 1259-1272.

Tener, G.M., Hayashi, S., Dunn, R., Delaney, A., Gillam, I.C., Grigliatti, T.A., Kaufman, T.C. and Suzuki, D.T. (1980) in Transfer RNA: Biological Aspects (Söll, D., Abelson, J.N. and Schimmel, P.R., eds.) pp 295-307, Cold Spring Harbor Laboratory.

Thach, R.E. and Doty, P. (1965) Science 148, 632-634.

Thomson, R.Y. (1960) in Chromatographic and Electrophoretic Techniques (Smith, I., ed.) Vol. I, pp 242-243, Wm. Heinemann Publishers, London.

Tolman, G.L., Barrio, J.R. and Leonard, N.J. (1974) Biochem. 13, 4869-4878.

Travers, A.A. (1980) J. Bacteriol. 141, 973-976.

Valenzuela, P., Bell, G.I., Venegas, A., Sewell, E.T., Masiarz, F.R., DeGennaro, L.J., Weinberg, F. and Rutter, W.J. (1977) J. Biol. Chem. 252, 8126-8135.

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

Van Etten R.A., Walberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170.

Veloso, D., Guynn, R.W., Oskarsson, M., and Veech, R.L. (1973) J. Biol. Chem. 248, 4811-4819.

Venegas, A., Quiroga, M. Zaldivar, J., Rutter, W.J. and Valenzuela, P. (1979) J. Biol. Chem. 254, 12306-12309.

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

Weissbach, H. and Ochoa, S. (1976) Ann. Rev. Biochem. 45, 191-216.

White, B.N., Tener, G.M., Holden, J. and Suzuki, D.T. (1973a) Develop. Biol. 33, 185-195.

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

White, B.N. and Tener, G.M. (1973c) Can. J. Biochem. 51, 896-902.

White, B.N. (1980) Comp. Biochem. Physiol. 67B, 103-107.

Woo, N.H., Roe, B.A. and Rich, A. (1980) Nature 286, 346-351.

Yanofsky, C. (1981) Nature 289, 751-758.

Yen, P.H. and Davidson, N. (1980) Cell 22, 137-148.

Young, R.A., Macklis, R., and Steitz, J.A. (1979a) J. Biol. Chem. 254, 3264-3271.

Young, R.A. (1979b) J. Biol. Chem. 254, 12725-12731.

Zachau, H.G., Dutting, D. and Feldmann, H. (1966) Hoppe-Seyler's Z. Physiol. Chem. 347, 212-235.