tRNA like sequences in the human genome

A Thesis Presented for the

Degree of

Doctor of Philosophy

by

Narinder Singh Bansal

Department of Biochemistry

University of Glasgow.

December, 1986.

Summary

A human foetal DNA library cloned in bacteriophage lambda Charon 4A was screened for human tRNA gene sequences using a *X. laevis* tDNA as a probe (3.18 Kb *X. laevis* DNA fragment containing 8 different tRNA genes, cloned in pBR 322). From the screeing procedure one lambda clone was isolated that showed positive hybridization. This clone, λht3, was further characterized by hybridization of [^{32}P] labelled human 4S RNA probe and the individual subclones of X. laevis tDNA. Of the 5 different subclones of *X. laevis* tDNA, only the tDNA^{Tyr} subclone showed positive hybridization. In addition the hybridization pattern observed was similar to that observed with a [^{32}P] labelled human 4S RNA probe. A restriction cleavage map of λht3 was constructed, and the HindIII restriction fragments of λht3 DNA showing positive hybridization were subsequently subcloned in plasmid pAT153 as pNB1 and pNB4. These subclones were tested in *Xenopus* oocyte nuclei and HeLa cell extract for their transcriptional activity. A fine structure restriction map of the subclones was constructed. The smallest restriction fragments of pNB1 and pNB4 showing positive hybridization to [^{32}P] labelled human 4S RNA probe were identified as 0.6 Kb SmaI-HindIII restriction fragment of pNB1 and 0.9Kb PstI restriction fragment of pNB4. The DNA nucleotide sequence of the 0.6Kb SmaI - HindIII fragment was determined. This nucleotide sequence was examined for tRNA gene sequences and found to contain no full length tRNAs. However when the 0.6Kb sequence was compared with the 3.18Kb DNA sequence of *X. laevis*, one region of homology was identified. This was a stretch of 17 nucleotides long that showed 82% homology to 3' end of *X. laevis* tRNA^{Tyr} gene. The 0.6Kb DNA sequence also contains an incomplete human Alu sequence and the stretch

of 17 nucleotides homologous to *Xenopus* $tRNA^{Tyr}$ gene is contained within the 31bp human insert present in the Alu sequence.

Acknowledgements

I would like to express my sincere thanks to the following people.

Professor R.M.S. Smellie, for making the facilities of the Department of Biochemistry available for this research.

Professor R.H. Burdon for his guidance, encouragement and criticism throughout the course of this work.

Dr. J.P. Goddard for his invaluable advice, helpful comments and criticism.

Dr. A. Cato for his help.

Mrs. Serena Hale for typing this thesis.

My colleagues and friends who saw me through both good and bad days.

My parents for their unfailing support and encouragement.

Mohinder, for her loyal support and forbearance.

TABLE OF CONTENTS

List of Tables Page

List of Figures

Abbreviations

The abbreviations used in this thesis are those suggested in the Instruction to Authors of the Biochemical Journal, with the following additions:-

tRNA	-	transfer RNA
tDNA	-	genes for transfer RNA
bp.	-	base pairs
Kb	-	kilobases
dNTP	-	deoxynucleoside - 5' - triphosphate
pCp	-	cytidine 3' - 5' - bisphosphate
SDS	-	Sodium Dodecyl Sulphate
DMSO	-	Dimethyl Sulphoxide
DTT	-	Dithiothreitol
HEPES	-	N-2-Hydroxythy piperazine-N-2ethane sulphonic acid
TEMED	-	N, N, N', N' - Tetramethylene Diamine

CHAPTER 1 INTRODUCTION

1. General Introduction

Transfer RNAs are a major constituent of every living cell and play a central role in the expression of genetic information. They are relatively small molecules, uniquely designed to serve as adaptors for the translation of the information encoded in the nucleotide sequence of the messenger RNA (mRNA) into specific amino acid sequences. Each transfer RNA (tRNA) molecule is capable of "reading" three adjacent mRNA bases (a codon) and placing the corresponding amino acid at a site on the ribosome at which a peptide bond is formed with an adjacent amino acid. The interaction of tRNAs with the codons in mRNA occurs via base pairing: each tRNA contains a trinucleotide sequence, the anticodon, which is complementary and can thus anneal to the trinucleotide codon.

Transfer RNAs were first discovered in 1957 (Hoagland *et al.*, 1958) following an earlier prediction of Crick of a small adaptor molecule required for molecular transmission of genetic information. According to his "Adaptor Hypothesis": (as cited by Hoagland, 1960) "each amino acid would combine chemically, at a special enzyme, with a small molecule which, having a specific hydrogen-bonding surface, would combine specifically with nucleic acid template". Hoagland and his coworkers (1958) were able to identify these low molecular RNAs and amino acid-activating enzymes from a soluble cytoplasmic fraction of rat liver cells, which had essentially the properties postulated by Crick. This RNA was first known as soluble RNA (sRNA), and later as transfer RNA (tRNA).

The nucleotide sequence of the first tRNA molecule was determined by Holley *et al.*, (1965). This pioneering work was accomplished in seven years and used a full gram of alanine tRNA purified from more than 300 pounds of yeast cells. With the advances in rapid nucleotide sequence techniques, more than 300 different tRNAs for a variety of biological sources have been isolated,

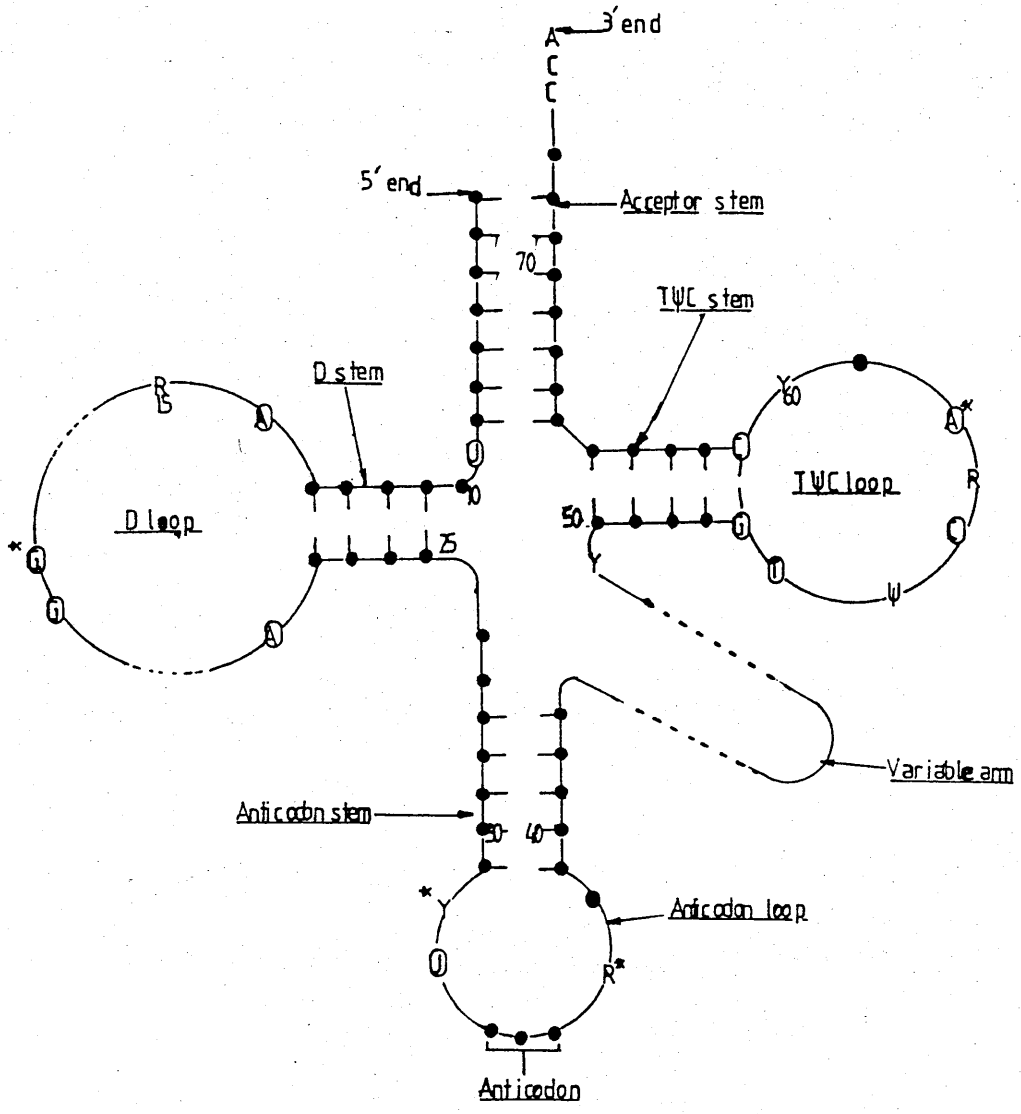
purified and fully sequenced (Sprinzl *et al.*, 1985). This excellent progress in sequence determination has given some general rules about primary and secondary structure of tRNAs. Transfer RNAs are the smallest cellular RNA molecules sedimenting at about 4S, ranging in molecular weight from 23,000 to 25,000. They are distinguished from other small RNA species of the cell by their characteristics high content of modified nucleosides; amongst the more common of these are dihydrouridine (D), pseudo-uridine (ψ), 4-thiouridine, methyl or dimethyl guanosine and methyl adenosine (Zachau, 1972). As many as 15% of nucleotides of different tRNA may take this form.

The recent developments of molecular cloning and rapid DNA sequence analysis have facilitated the study of structure and function of genes which code for transfer RNA. This has also given insight into the organisation of tRNA genes in the genomes of several organisms and allowed the elucidation of DNA sequence of many tRNA genes. The detailed structural analysis of tRNA genes is of central importance to the understanding of their gene expression and regulation. It is now possible to isolate and produce purified segments of single tRNA genes in sufficient quantity, which can be used to study tRNA gene expression and the correlation between gene sequences and the function of its transcription product.

2. Transfer RNA structure

In the two decades that have passed since Holley and his co-workers made their determination, tRNA molecules from different microorganisms, plants and animals have been sequenced. The compilation of different tRNA sequences shows that tRNA molecules as a class have many features in common, although each specific tRNA differs from the others in its nucleotide sequence. The primary structure of all tRNA molecules allows extensive folding and intrastrand complementarity to generate a secondary structure,

Figure 1.1 Generalized clover leaf structure of tRNA. Positions are numbered from 5' to 3' according to the yeast $tRNA^{Phe}$ structure which contains 76 residues. Starred nucleotides are positions where modification of that nucleotide can possibly occur: Invariant nucleosides are circled; Y; represents pyrimidine; R; represents purine; ⟷ indicate bases involved in helical stems containing Watson-Crick base pairs. The dotted stretches are regions of variable lengths. Adapted from Gauss *et al.*, 1979.


3'end
A
C
C
5'end
Acceptor stem
70
TψC stem
D stem
D loop
TψC loop
10
25
50
60
Variable arm
Anticodon stem
30
40
Anticodon loop
Anticodon

which can appear like a clover leaf, as originally proposed by Holley *et al.*, (1965). Study of many of these molecules by physical methods has shown in each case that this is the correct configuration. By comparison to tRNA sequences, certain features have been found to be common to almost all the molecules. This had led to the idea of a "standard" tRNA molecule consisting of 76 nucleotide arranged in a clover leaf configuration as illustrated in Figure 1.1. By convention, the nucleotides are numbered 1 through 76 starting from the 5'- terminus. The general clover leaf form is built up from a stem and four fold-back arms. The 5' and the 3' end of the molecule forms the amino acid acceptor stem, held together by seven Watson-Crick base pairs. The 5'-terminal nucleotide is almost always a guanine residue (except for initiator tRNAs). The 3'-terminal four nucleotides are unpaired, the nucleotide sequence of the last three is CCA in all tRNAs. The presence of the CCA terminus is an absolute requirement for the function of tRNA. The adenine of this terminal sequence serves as the attachment site for the specific amino acid. The first arm, moving in a clockwise direction around the molecule, contains five base pairs and terminal loop with seven bases. This arm is called the TψC arm because the sequence TψC occurs in its loop in almost all tRNAs.

Next in clockwise sequence is the variable arm or extra arm, which may contain between four to twenty-one nucleotide. Below this is the anticodon arm formed, like the TψC arm, from a five base pair stem and a loop containing seven unpaired nucleotides. The three nucleotide anticodon, which pairs with the triplet codon on mRNA during protein synthesis is carried at the tip of this loop. The final arm, called the D arm because it contains the modified base dihydrouridine (D) in most tRNAs, comprises three or four base pairs stem and a loop of variable size. The overall range of tRNA is from 74 to 95

nucleotides. The variation in length is due to differences only in D loop and extra arm. As far as position of some individual nucleotides within the clover leaf structure is concerned, a high degree of homology can be found between different tRNA species, although no long homologous sequences are present in any group of tRNAs. The bases found at some positions are conserved or invariant. Most tRNAs have the same base at that position (circled in Figure 1.1.) Other bases are semi-invariant and limited to either a purine (R) or pyrimidine (Y) at fixed points. The total number of invariant and semi-variant residues is 21. Several exceptions to this generalized tRNA structure have been reported, and are reviewed in detail by Dirheimer *et al.*, (1979) and Clark (1978).

During the early 1970's Rich and Klug and their colleagues (Suddath *et al.*, 1974; Robertus *et al.*, 1974) used X-ray diffraction to deduce the three dimensional crystal structure of purified tRNAs. Their investigations confirm that tRNA molecules contain the clover leaf arrangement, originally proposed by Holley (1965), which is folded further in three dimensions into an L-shaped structure. In addition to the base pairing that are represented in clover leaf structure, further hydrogen bonds form in the tertiary structure. Many of the invariant and semi-invariant bases are involved in the formation of tertiary hydrogen bonds, which both explains their conservation and supports the notion that the general form of the tertiary structure may be common to all tRNAs. The greatest concentration of these tertiary interactions occurs between the D loop and the TψC loop in such a way to create two double helices at right angles to each other while maintaining the base paired double helical stems of the secondary structure. The acceptor stem and the TψC stem form one continuous double helix and the D stem and anticodon stem from another double helix, so that the CCA acceptor and the anticodon are exposed at the opposite ends of the

'L'. The entire structure is about 2.5 nm thick and 9 nm long (Kim 1979, Rich and RajBandary, 1976).

This is only a brief summary of tRNA structure but serves to introduce the subject of tRNA genes and their expression (for a more detailed review see Singhal and Fallis 1979; Clark 1978). The remainder of this introduction summarizes the organisation and structure of tRNA genes with particular emphasis placed on eukaryotic nuclear tRNA genes. For general review of prokaryotic tRNA genes see Fukada *et al.*, (1980), and for mitochondrial tRNA genes see Attardi (1981).

3. Organisation of eukaryotic transfer RNA genes

The organisation of eukaryotic nuclear tRNA genes has been examined in various species. The total number of tRNA genes in any species always exceeds by far the number of different tRNA sequences. (Schweizer *et al.*, 1969; Weber and Berger, 1976; Hattlen and Attardi, 1971; Clarkson *et al.*, 1973; Feldman 1976; Sulston and Brenner 1974). That is, the tRNA genes of a single acceptor (chemically different species of tRNA) are generally represented by several loci, and each tRNA gene is reiterated between an average of 10 times in yeast to about 200 times in *Xenopus*. This poses such intriguing questions regarding the organisation of eukaryotic tRNA genes as where tRNA genes occur on the chromosome, and if there are any functional logic to their arrangement. The answers to these questions are emerging from studies of tRNA genes in yeast, *Drosophila, Xenopus* and mammals. All these studies reveal that there seems to be no general role governing the way in which tRNA genes are organised in eukaryotes.

3.a Organisation of yeast tRNA genes

There are about 360 tRNA genes per haploid genome in yeast (Schweizer, *et al.*, 1969). On average each yeast tRNA gene is reiterated 6 to 10 times

(Feldman *et al.*, 1976). Information on the arrangment of tRNA genes of yeast has come from several sources. Evidence from gene cloning and genetic mapping studies suggests that the tRNA genes are scattered throughout the genome. The first evidence came from the genetic mapping of non-sensense suppressors (Hawthorne *et al.*, 1976), where eight tyrosine inserting ochre suppressors were mapped on eight different genetic loci, on six different chromosomes. These results were later confirmed by genetic mapping of eight different restrictions fragments of yeast DNA generated by restriction endonuclease EcoRI that hybridize to a certain type of labelled $tRNA^{Tyr}$ (Olson *et al.*, 1979). In a similar manner $tRNA^{Ser}_{2}$ has been shown to hybridize to eleven different restriction fragments (Page and Hall, 1981); purified $tRNA^{Leu}$ hybridized to 14 different 3 to 20Kb size restriction fragments (Venegas *et al.*, 1979) and $tRNA^{Phe}$ hybridized to 10 different restriction fragments (Valenzuela *et al.*, 1978). In addition, Beckman *et al.*, (1977) screened a library of yeast DNA fragments generated by restriction endonuclease Hind III, and found 175 clones which hybridize to total yeast tRNA (4S). On screening with 18 different species of purified tRNAs, 97 clones hybridized to at least one tRNA. Other groups have sequenced relatively large DNA fragments cloned from the yeast genome but have shown the presence of single tRNA genes, in each fragment: for example $tRNA^{Met}$ (Olah and Feldman, 1980); $tRNA^{Glu}$, $tRNA^{Val}$ and $tRNA^{Arg}$ (Baker *et al.*, 1982). These and earlier findings led to the tentative conclusion that most yeast tRNA genes are widely scattered throughout the genome and not clustered.

However, in some cases clustering of certain yeast tRNA genes has been observed: In four separate recombinant clones, $tRNA^{Arg}$ gene is closely linked to a $tRNA^{Asp}$ gene. DNA sequence analysis have revealed that in one case the two genes are separated by only nine base pairs (Schmidt *et al.*, 1980). A similar case has been reported for $tRNA^{Ser}$ and $tRNA^{Met}$ gene, the two genes

Figure 1.2 The arrangement of the 18 tRNA genes located in the cytological region 42A of the *Drosophila melanogaster* salivery gland chromosomes. No tRNA genes have been found in the regions from 30 to 55 and from -20 to -40. The distances are given in Kilo bases. The tRNA genes are oriented in both directions. Arrows indicate the polarity of the genes (Adapted from Kubli 1981).

-45
-20
Asn
Asn
-15
-10
Arg
-5
0
Lys
Arg
Ile
5
Lys
Lys
Lys
3Asn
10
Arg
15
Asn
Asn
Asn
20
25
Lys
Arg
30
55

are separated by a spacer DNA of seven base pairs (Mao *et al.*, 1980). Interestingly, in both these cases the tRNA genes are co transcribed *in vitro* as a dimeric precursor.

3.b Organisation of *Drosophila* tRNA genes

The localization of tRNA genes in *Drosophila* is greatly simplified by the powerful technique of *in situ* hybridization to polytene chromosomes (Gall and Pardue, 1969; John *et al.*, 1969). This technique is carried out by treating polytene chromosome fixed to a microscope slide under conditions that denature some of the DNA without disrupting its cytological integrity (structure). Radioactive tRNAs (unfractionated or purified tRNA species) are then hybridized to the chromosome and autoradiography reveals the location of tRNA genes. The results from hybridization of unfractionated tRNA (total tRNA) indicate that there are 600-750 tRNA genes per haploid genome, randomly distributed at about 54-60 different loci over most chromosomes except the fourth (Steffenson and Wimber, 1971; Elder *et al.*, 1980). On the other hand, hybridization of purified tRNA's suggests that the tRNA genes are locally clustered (Kubli and Schmidt, 1978; Hovemann *et al.*, 1980; Yen and Davidson 1980; Robinson and Davidson, 1981). The tRNA genes contained within such clusters may code for the same tRNA (homocluster) or different tRNA species (heteroclusters).

Molecular cloning and DNA sequence analysis of fragments containing such clusters reveals a clearer but somewhat puzzling arrangement of tRNA genes. The most prominent site of hybridization of total tRNA is DNA from region 42A of the chromosome 2R (Schmidt *et al.*, 1978; Hovemann *et al.*, 1980; Yet *et al.*, 1977; Yen and Davidson 1980). Yen and Davidson have cloned a 94 Kb fragment from this region: all the tRNA genes are encoded in the central 42 Kb segment. It contains three widely spaced homoclusters of

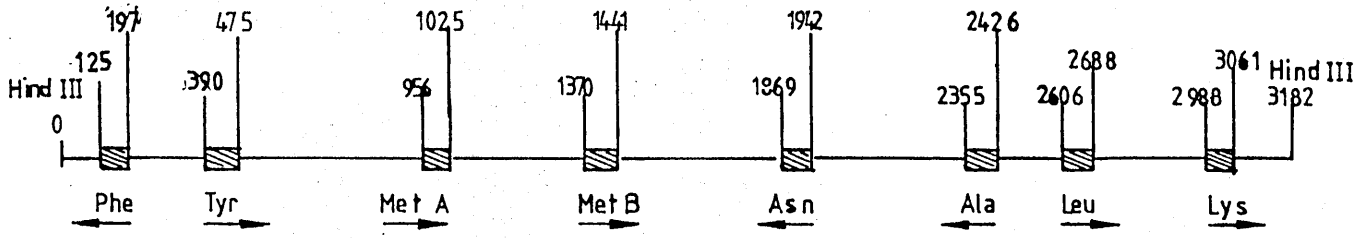
$tRNA^{Asn}$ genes; four widely spaced $tRNA^{Arg}$ genes; five randomly spaced $tRNA^{Lys}$ genes. These tRNA genes are oriented with opposite polarity and coded by both DNA strands (Figure 1.2.) They are separated by DNA regions of varying lengths, which seem to be transcriptionally "silent" (Schmidt *et al.*, 1978); Yen *et al.*, (1977) have reported the occurrence of repetitive elements in these spacer regions. Analysis of another chromosomal region 50 AB (Robinson and Davidson, 1981) reveals that tRNA genes are subclustered in a similar manner to that observed in chromosomal region 42A.

In contrast to clustering of tRNA genes, Sharp *et al.*, (1981) have reported that initiator tRNA genes are widely scattered, not intermingled with other tRNA genes. They occur as single gene copies, within a 415 bp repeated DNA segment, which is separated from other initiator tRNA genes by an average distance of 17 Kb (Sharp *et al.*, 1981). In addition, at chromosomal region 56F, single copies of $tRNA^{Gly}$ gene are found within two direct repeats of 1.1 Kb (Hershey and Davidson, 1980). To add to the diversity of tRNA gene organisation in *Drosophila*, another pattern is observed at chromosomal location 56 EF, where six tRNA genes are located adjacent to 3' end of the 5S RNA gene cluster (Indik and Tartof, 1982). Three of these genes code for $tRNA^{Glu}$, arranged as tandem repeats and are all transcribed in the same direction as 5S RNA, towards the centromere of chromosome II.

3.c Organisation of Xenopus tRNA genes

Xenopus laevis tRNA genes form extensive clusters that behave like a group of satellites, banding at higher buoyant densities in caesium chloride gradients than does bulk *X. laevis* DNA (Clarkson and Kurer, 1976: Clarkson *et al.*, 1973). The arrangement of tRNA genes in *X. laevis* was elucidated by the DNA sequencing of a 3.18 Kb DNA fragment that is tandemly repeated some 150 times in a cluster (Clarkson *et al.*, 1978; Muller and Clarkson, 1980; Clarkson personal communication). This repeat unit contains eight tRNA

Figure 1.3 Schematic diagram of a repeated cluster of tRNA genes of *Xenopus laevis*, and their direction of transcription. (Bryan *et al.*, 1981). Numbers indicate the nucleotide positions based on sequence analysis (S.G. Clarkson personal communication).


Hind III
0
125
197
390
475
956
1025
1370
1441
1869
1942
2355
2426
2606
2688
2988
3061
Hind III
3182
Phe
Tyr
Met A
Met B
Asn
Ala
Leu
Lys

genes: two coding for $tRNA^{Met}$, and single genes for $tRNA^{Phe}$; $tRNA^{Tyr}$; $tRNA^{Asn}$, $tRNA^{Ala}$ $tRNA^{Leu}$ and $tRNA^{Lys}$ (Figure 1.3). These genes are irregularly spaced and coded by both DNA strands. Another feature of this array is its 100 fold tandem repetition in the genome.

3.d Organisation of mammalian tRNA genes

In mammalian cells, the number of reiterated genes for each tRNA species varies from about 10 to 20 copies in the haploid HeLa cell genome (Hattlen and Attardi, 1971) to about 100 in mouse myeloma cells (Marzloff *et al.*, 1975). Information on the arrangement of mammalian tRNA genes is very scanty. In general, there seems to be lack of regularity within mammalian tRNA gene families, and several organisation patterns have been observed. While some single acceptor species are widely dispersed throughout the genome, others are clustered over comparatively small chromosomal domains and small repeating units containing tRNA genes have also been reported. Sekiya *et al.*, (1982) have reported a 11.9 Kb rat DNA fragment containing three tRNA genes for $tRNA^{Pro}$ and three for $tRNA^{Lys}$. Another cluster of human tRNA genes for $tRNA^{Lys}$, $tRNA^{Gln}$ and $tRNA^{Leu}$ is contained with a 1.65 Kb DNA fragment (Roy *et al.*, 1982). These genes are irregularly spaced and some genes have opposite polarity with respect to others. Shibuya and co-workers have reported a 3.3 Kb rat DNA fragment that contains genes for $tRNA^{Asp}$; $tRNA^{Leu}$; $tRNA^{Gly}$ and $tRNA^{Glu}$. This fragment occurs about ten times per haploid rat genome. DNA sequence analyses of six independent clones of these repeating units revealed considerable sequence heterogeneity within the coding region of tRNA genes; while the genes for $tRNA^{Asp}$ and $tRNA^{Leu}$ are highly conserved, the coding regions of $tRNA^{Gly}$ and $tRNA^{Glu}$ lack several nucleotides. In fact, they report one repeating unit that totally lacks the genes for $tRNA^{Glu}$ (Sekiya *et al.*, 1983; Shibuya *et al.*, 1982). In contrast Lasserweiss and coworkers (1981) have isolated and analysed several different

clones from a rat genomic library. Their results indicate that rat tRNA genes are mainly distributed singly or in small clusters spanning 1 to 2 Kb, irregularly spaced and interspaced with large DNA fragments (minimum 8 to 20 Kb) of non-tRNA coding regions. Santos and Zasloff (1982) have reported that human $tRNA_i^{Met}$ genes are scattered throughout the gnome. In addition other groups have shown from DNA sequence analysis of small 1-2 Kb fragments from large genomic clones (12-17 Kb), the presence of a single tRNA gene (Goddard *et al.*, 1983; Han and Harding 1983; Buckland *et al.*, 1983).

This brief survey on the organisation of eukaryotic tRNA genes illustrates the various organisational patterns that exist within the tRNA gene family. The most striking feature of the organisation of tRNA genes is the low degree of regularity. In eukaryotes, tRNA genes are predominantly organised as single transcriptional units. Nonethless in *Drosophila, Xenopus* and mammals sets of tRNA genes are found to be clustered, but in no obvious regular manner. The tRNA genes contained within such clusters may code for the same or different tRNA species, which are intermingled, irregularly spaced, and may be encoded on both strands. Nevertheless, the tRNA genes contained within these clusters are not co-transcribed as a large multimeric precursor as is the case of *E.coli* tRNA genes. Neither are any of the eukaryotic tRNA genes located in the spacer regions between ribosomal RNA genes. So far, little information is available on the structure of the different chromosomal regions in which eukaryotic tRNA genes are located. Furthermore it is not known whether a distinct arrangement of tRNA genes is of functional significance.

4. Biosynthesis of eukaryotic tRNA

The availability of cloned tRNA genes (tDNA) as templates for transcription *in vitro* has given clear insight into the biosynthesis of tRNA. Biosynthesis of tRNAs involves both transcriptional and post-transcriptional

events which occur in a fixed temporal order (Abelson, 1979). The transcriptional product of eukaryotic tRNA genes is not tRNA itself but a longer precursor molecule with extra nucleotides at the 5' and at the 3' extremities. This precursor RNA molecule must be processed subsequently to yield a mature gene product. Some eukaryotic tRNA genes are noncolinear with their gene products, may contain an intron or intervening sequence, that are transcribed to form part of the primary transcript. The primary transcript serves as a substrate for several processing enzymes in a number of sequential steps. These steps include 5' and 3' maturation, nucleotide modification at specific residues, enzymatic addition of CCA terminus and in some cases enzymatic excision (splicing) of intervening sequence. The availability of tRNA genes of known sequence has also allowed the characterization of initiation and termination sites of RNA polymerase III, the enzyme responsible for tRNA gene transcription in eukaryotes. With several tRNA genes it has been established that information essential for the promotion of transcription is located within the coding region (De Franco *et al.*, 1979; Kressmann *et al.*, 1979; Hofstetter *et al.*, 1981; Sharp *et al.*, 1981a).

4.a Transcription of tRNA genes

To understand the basis of tRNA transcription it is essential to identify the primary transcriptional product of a particular tRNA gene. Although precursor tRNAs were first discovered in mammalian cells (Burdon *et al.*, 1967), characterization of these precursors was difficult because of the inherent difficulty in obtaining pure tRNA precursors *in vivo*, from a mixed population of short lived molecules. In recent years, this problem is overcome by using cloned tRNA genes. By micro-injection of tDNA into the nucleus of *Xenopus* oocytes or by incubation in a variety of *in vitro* transcription cell-free systems, (Brown and Gurdon, 1978; Kressman *et al.*, 1979; Cortese *et al.*, 1978; De Robertis and Olson, 1979; Melton and Cortese 1979; Telford *et al.*, 1979;

Organism	Gene	5' Flanking sequence	Reference
		-20 -10 -1	
Yeast	$tRNA^{Ser}$	T T A C A A T G A G T G T C A G A T A A	Mao et al., 1980
	$tRNA^{Tyr}$	G A T G G T T A T C A G T T A A T T G A	DeRobertis and Olson 1979
	SuP4-o $tRNA^{Tyr}$	C T T T C T T C A A C A A T T A A A T A	Koski et al., 1980
Bombyx	$tRNA^{Ala}$	T T T T G C A A G C T T T T T C C G T T	Garber and Gage 1979
Drosophila	$tRNA^{Arg}$	G T T A C A C T C G C A C G T C A A G C	Sharp et al.,1981a
	$tRNA^{Lys}$	T A T C G G C G A T C T T C A C A A G T	DeFranco et al., 1980
	$tRNA^{Lys}$	A G T T T T T G G C T C A T C A A G T G	DeFranco et al., 1980
	$tRNA^{Met}$	A T G A C G C A A C T T T G T A T G C A	Sharp et al., 1981b
	$tRNA^{Met}$	A T G A C G C A A C T T T A T A T G C A	Sharp et al., 1981b
Xenopus	$tRNA^{Asn}$	G G T G G C C G G T G G C A A G C G T C	Clarkson 1983
	$tRNA^{Leu}$	C A C A A C G G C T G G C T T T C A C A	Galli et al., 1981
	$tRNA^{Lys}$	G G A G G A A C G A G C G A C G C A G T	Clarkson 1983
	$tRNA^{Met}$	T C A C C C A A A G G A C G G C A A T C	Koski and Clarkson 1982
	$tRNA^{Met}_{variant}$	T G C G C G T G C C A G C T G C G A T T	Clarkson 1983
Mammals	$tRNA^{Met}$	G A C C G T G T G C T T G G C A G A A C	Zasloff et al., 1983
	$tRNA^{Met}_{variant}$	G T A G A A G C G T G T T T T C C G T T	Zasloff et al., 1982

Table 1.1 Transcription initiation sites of Eukaryotic tRNA genes:- The non-coding DNA strand sequence is numbered negatively from the 5' nucleotide of the mature tRNA in each case. The arrows indicate the initiating nucleotide (compiled from Clarkson 1983).

Melton *et al.*, 1980; Sprague *et al.*, 1980; Weil *et al.*, 1979; Manley *et al.*, 1980; Clarkson *et al.*, 1981; Dingermann *et al.*, 1981; Orden *et al.*, 1979; Mattoccia *et al.*, 1979; Silvermann *et al.*, 1979; Zasloff *et al.*, 1982), it is possible to achieve the synthesis of sufficient amounts of tRNA precursor molecules for fingerprint analysis, or direct RNA sequencing. By comparing the nucleotide sequence of the primary transcript with that of the respective gene and its flanking regions, it is possible to identify precisely the point at which the transcript is initiated, terminated and processed. For several eukaryotic tRNA genes, the sequences of primary transcripts are now known (De Robertis and Olson 1979; Garber and Gage, 1979; Koski and Clarkson, 1982; Zasloff *et al.*, 1982, Table 1.1). All the known tRNA primary transcripts share the following common features.

(i) Their transcription is mediated by RNA polymerase III.

(ii) Most transcripts contain an uncapped purine nucleoside triphosphate as the first base.

(iii) In most cases, the primary transcript contains only a single tRNA sequence. Although some yeast tRNA genes are co-transcribed *in vitro* as a dimeric precursor (Mao *et al.*, 1980; Schmidt *et al.*, 1980).

(iv) All transcripts terminate in oligouridylate sequences.

4.a.i Initiation of transcription

The transcription of eukaryotic tRNA genes invariably starts in this 5' flanking seuqences. In all known cases initiation of eukaryotic tRNA genes occurs at a purine residue located between position -20 and -3 upstream from the 5' end of the tRNA coding sequence (Clarkson 1983, Table 1.1). There is no obvious sequence homology surrounding the initiating residue except that is in most cases preceded by a pyrimidine residue. Most tRNA genes possess a single initiation site except in case of initiator tRNA genes of *D. melanogaster* and *X. laevis* (Table 1.1), where transcription starts at two or even three points

(Sharp *et al.*, 1981b; Koski and Clarkson, 1982; Clarkson 1983). However the underlying features of sequences determining the specific start-site for transcription remain obscure.

4.a.ii Transfer RNA gene promoter

Transcription of purified eukaryotic tRNA genes and a series of insertion and deletion mutants has established that essential promoter elements are located within the transcribed sequence of tRNA genes (Kressmann *et al.*, 1979; DeFranco *et al.*, 1981; Telford *et al.*, 1979; Garber and Gage 1979). More precise studies with *X. laevis* $tRNA^{Met}$ and $tRNA^{Leu}$ genes (Hofstetter *et al.*, 1981; Galli *et al.*, 1981); *C. elegans* $tRNA^{Pro}$ gene (Ciliberto *et al.*, 1982a) and *Drosophila* $tRNA^{Arg}$ gene (Sharp *et al.*, 1981a) revealed that their intragenic promoter is non-contiguous and split into two sequence blocks that are set far apart. They are referred to as 'A' block (or 'D' control region) and 'B' block (or 'T' control region); and have the approximate positions: 8-19 and 52-62 respectively, (according to the standard system of numbering tRNA genes). An intriguing feature of these sequence blocks is that they coincide with the highly conserved nucleotides found in D and TψC stem and loop regions amongst eukaryotic and many prokaryotic tRNAs (Galli *et al.*, 1981). Conservation of these sequences has previously been interpreted solely in terms of their importance to tRNA structure (stabilizing the tertiary structure of RNA molecule). From a comparison of the homologous region of all eukaryotic tRNAs, Traboni *et al.*, (1982) have derived a consensus sequence for the internal split promoter, which is RRYNNARY-GG for block 'A' and G T/A TCCRANNC for block 'B' (where R represents a purine, Y represents a pyrimidine and N represents any nucleotide).

4.a.iii How does the intragenic split promoter function?

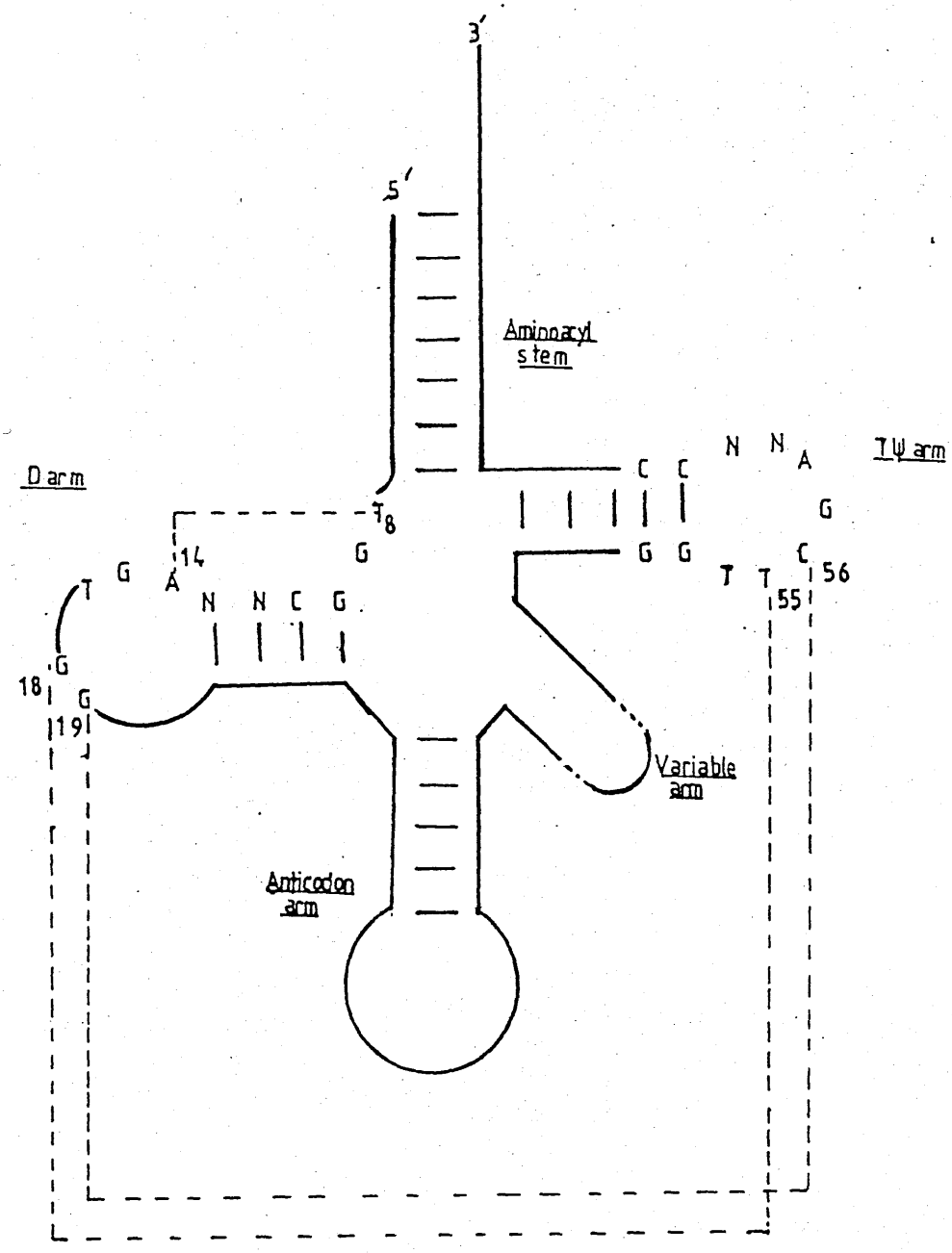
Several mechanistic hypotheses are proposed to explain the role of 'A' and

'B' blocks in transcription of tRNA genes. Purified RNA polymerase III is capable of accurate transcription only if supplemented with additional cofactors. Although tRNA gene transcription factor (s) have not yet been isolated in pure form, partial purification of cell extracts indicate that at least two transcriptional factors are required for tRNA gene expression (Segall *et al.*, 1980; Mattoccia *et al.*, 1979). This has led to the general assumption that 'A' and 'B' blocks interact with transcriptional cofactor(s) rather than the polymerase directly, where a transcriptional factor binds to the 'B' block sequence and the 'A' block is mandated to positioning of polymerase III at the initiating purine nucleotide (Hall *et al.*, 1982). No mechanistic proposal has yet been formulated to explain this interaction. However, transcription competition assay experiments (Kressman *et al..*, 1979; Sharp *et al.*, 1983) support the above assumption. Sharp *et al.*, (1983) used a series of 5' and 3' deletion mutants of *Drosophila* $tRNA^{Arg}$ gene as inhibitors of transcription of a reference tRNA gene and revealed the regions within *Drosophila* $tRNA^{Arg}$ that are required for recognition and binding of transcriptional factor(s). Their results indicate that for efficient transcription a protein factor (T-control factor) binds to the 'B' block sequence and somehow the sequence of the 'A' block effects this interaction: Whilst residues 50 to 58 of the 'B' block are prerequisite for the formation of a stable complex with T-control factor (Schaack et al., 1983), deletion of residues 7 to 12 of the 'A' block leads to a reduced ability of T-control factor binding to the 'B' block. The precise mechanism of this reduced ability is not clear. Sharp *et al.*, (1983) propose that a second factor, D-control factor, binds to residues 7 to 12 of 'A' block and optimal transcription of tRNA gene is accomplished by a cooperative mechanism, where each factor stabilizes the binding of the other (Sharp *et al.*, 1983; Schack *et al.*, 1983).

The clover leaf form in which the secondary structure of tRNA molecule

Figure 1.4 Tertiary interaction model of the eukaryotic tRNA gene promoter, as proposed by Hall *et al.*, (1982).

The non-coding DNA strand is shown in clover leaf form. The essential A block sequences extend from T8 to G19 and the B block sequences from G52 to C62. Dotted lines indicate the tertiary interactions.


3'
5'
Aminoacyl stem
D arm
TΨ arm
Anticodon arm
Variable arm
T8
14
18
19
55
56

is written out has led to the idea that analogous hairpin structures exist at the DNA level, and this secondary structure of tDNA, rather than the linear nucleotide structure plays an important role in the promotion of transcription. Hall *et al.*, (1982) proposed the "tertiary interaction model", based on the fact that 'A' and 'B' blocks encode single stranded loops (D and TψC), which are also engaged in strong mutual tertiary interactions in every tRNA molecule. According to this model the 'B' block DNA sequences exist in an intrastrand stem and loop conformation when a transcriptional cofactor binds to tDNA. This binding promotes tertiary interactions between sequences coding for the D and TψC loops, with the formation of G18 - T55 and G19 - C56 base pairs as illustrated in Figure 1.4 analogous to the well established tertiary interactions between the corresponding structures at the level of tRNA.

According to this model, 'B' block sequences as region of dyad symmetry are an essential feature of the eukaryotic tRNA promoter, while relatively few nucleotides of 'A' and 'B' block are required for transcription factor recognition. This model gains support from some point mutation analysis. For instance, substitution of G or T at position C_{56}, which weakens the proposed G_{19} - C_{56} tertiary interaction , severely reduces *in vitro* transcription of yeast $tRNA^{Tyr}$ gene (Koski, *et al.*, 1980), whereas, G_{57} to T_{57} substitution has no adverse effect on transcription of human $tRNA_i^{Met}$ gene (Zasloff *et al.*, 1982), since this nucleotide is not involved in tertiary pairing.

According to this model a tRNA gene unable to form a TψC stem should be poorly transcribed, even if it contains all the requisite conserved nucleotides of the 'B' block. However, published data on point mutation, deletions and insertion mutations within this region contradict this hypothesis (Folk *et al.*,

1982; Traboni*et al.*, 1982; Mattoccia *et al.*, 1983; Hofstetter *et al.*, 1981; Sharp *et al.*, 1981; Ciliberto *et al.*, 1982a). For example, a *C. elegans* mutant carrying deletion of all $tRNA^{Pro}$ sequences downstream from nucleotide 61, eliminates the possibility of formation of the TψC stem and yet is active in transcription (Ciliberto *et al.*, 1982a). In addition, *E. coli* $tRNA^{Tyr}$ gene with all the normal stem and loops and invariant nucleotides capable of the tertiary interactions between D and TψC loops as required for "tertiary interaction model", is transcriptionally inactive in *X. laevis* oocytes. However, it can be made active if the 'A' block region of *E. coli* $tRNA^{Tyr}$ gene is replaced by the corresponding *C. elegans* sequence (Dente, *et al.*, 1982; Folk *et al.*, 1982). This implies that improper block 'A' sequence are responsible for the inactivity and contradicts the "tertiary interaction model". Moreover, Folk and Hofstetter (1983) have shown that single base pair changes in coding sequence, which do not affect the potential secondary structure and tertiary interactions of the tRNA, do influence dramatically the extent of transcription (Folk and Hofstetter, 1983; Crampi *et al.*, 1982).

The nucleotide sequences located between the two intragenic control regions of the eukaryotic tRNA genes is generally considered to serve as a spacer region,to separate and maintain the two control regions at an optimal distance (Hofstetter *et al.*, 1981; Ciliberto *et al.*, 1982b). The distance between the 'A' and 'B' blocks can vary from 31 to 74 bp., this variability is due to the length of extra arm and the presence or absence of an intervening sequence (Ciliberto *et al.*, 1982b). While the DNA between 'A' and 'B' blocks show little sequence specificity for transcription, its length appears to be important, the optimal length being 30-40 bp., for efficient transcription (Ciliberto *et al.*, 1982b). The intervening sequence is not important for transcription *per se*, but may influence the efficiency of transcription by expanding the distance between 'A' and 'B' blocks. Folk and Hofstetter (1983) have proposed that

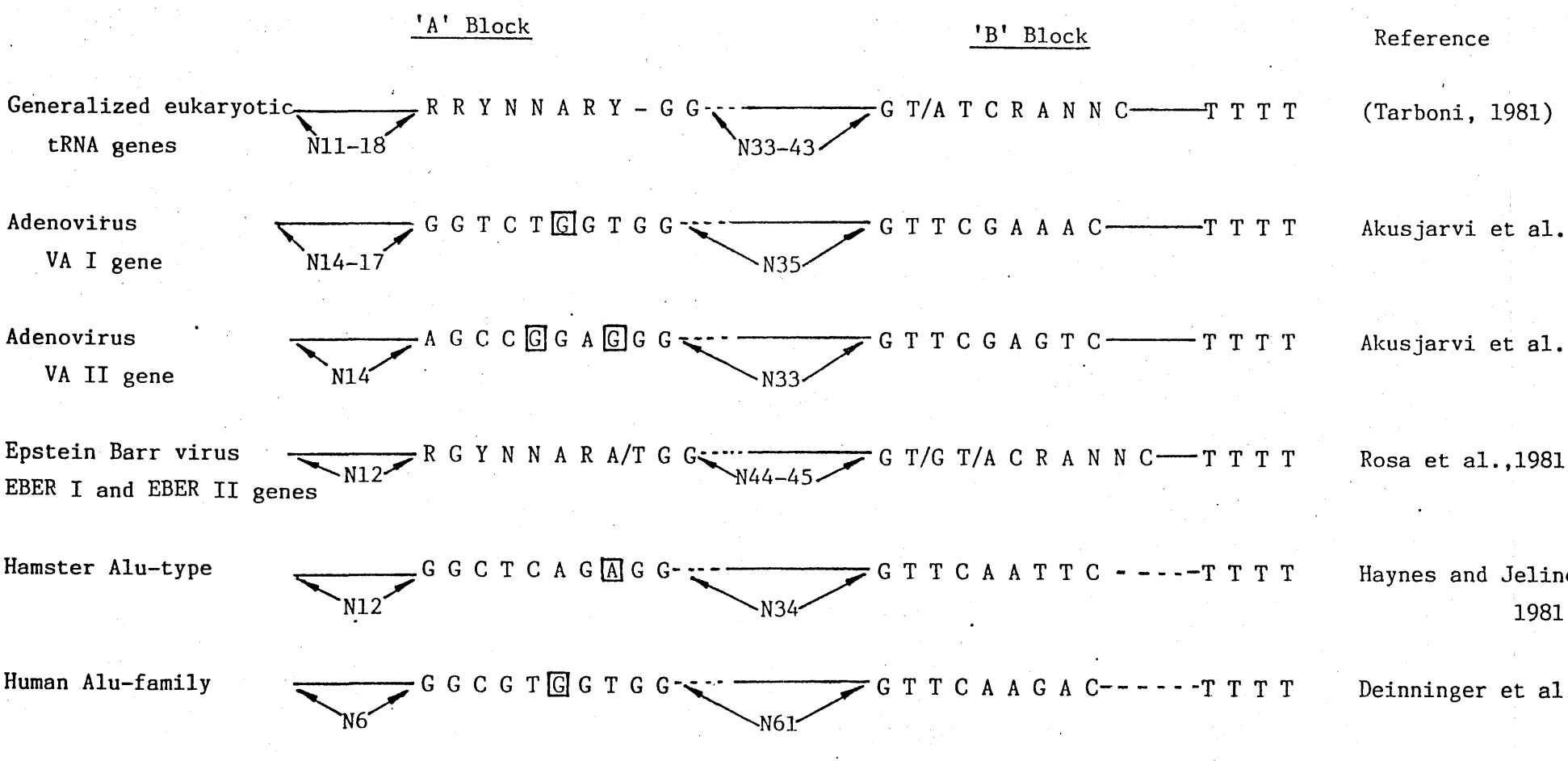
the central region of tDNA may be involved in the formation of a secondary structure analogous to the anticodon arm of a tRNA molecule, thereby establishing a fixed spatial orientation between the promoter 'A' and 'B' blocks. This notion is supported by the finding that mutant tRNA genes unable to form intrastrand stem-loop structure, are transcriptionally inactive (Folk and Hofstetter, 1983), whereas other artificially manipulated tRNA genes that have sequences deleted or substituted between 'A' and 'B' blocks are transcriptionally active only when the potential exists to form intrastrand hairpin structure (Folk and Hofstetter, 1983.) Interestingly, all known natural intervening sequences do not disturb this potential to form such structures.

4.a.iv Diversity of promoters for polymerase III

Eukaryotic RNA polymerase III transcribes a diverse set of genes which do not code for proteins. These include tRNA genes, 5S ribosomal RNA genes, virus-associated RNA encoded by adenovirus 2 (VAI and VAII genes) and Epstein-Barr virus (EBER I and EBER II genes) and certain members of the interspersed middle repetitive DNA genomic sequences (e.g. human Alu family DNA sequence). A common feature of all the genes transcribed by RNA polymerase III is the presence of intragenic promoter (Hofstetter *et al.*, 1981; Galli *et al.*, 1981; Ciliberto *et al.*, 1982b; Sakonju *et al.*, 1981; Sakonju *et al.*, 1980; Bogenhagen *et al.*, 1980; Fowlkes and Shenks 1980; Guilfoyle and Weinmann 1981; Rosa *et al.*, 1981; Paolella *et al.*, 1983. These genes have been classified into two classes (Ciliberto *et al.*, 1983).

Class I genes include; tRNAs, VAI and VAII, EBERI and EBERII, members of human Alu and mouse BI family. A characteristic feature of class I genes is that their intragenic promoter is split into two blocks separated by 30-60 nucleotides. Comparison of sequences of Class I genes (Galli *et al.*, 1981; Ciliberto *et al.*, 1983) revealed that many of the conserved nucleotides of 'A' and 'B' blocks of tRNA genes are also found within most class I genes (as

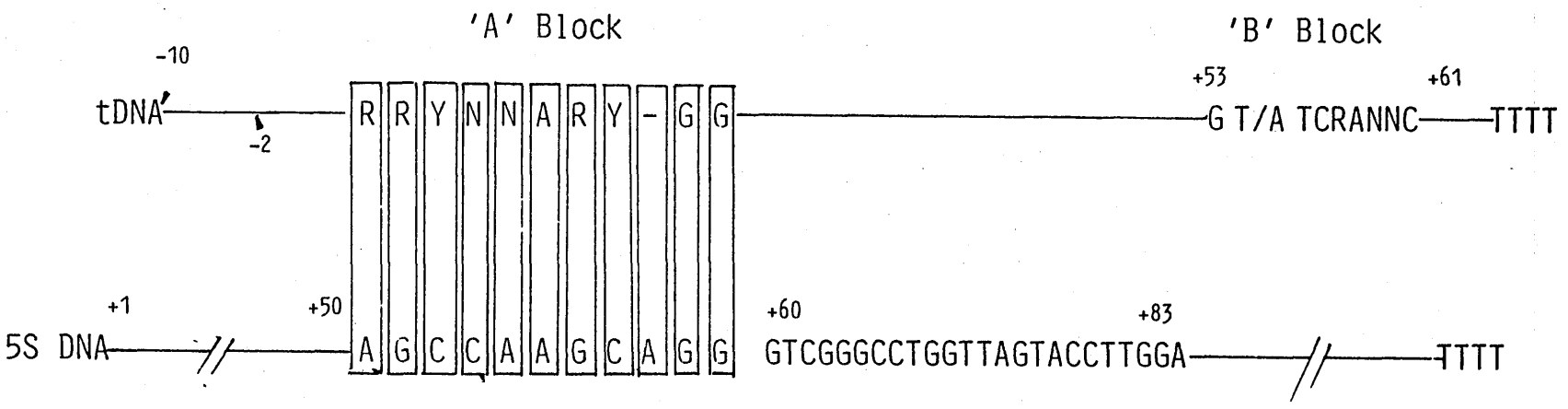
Figure 1.5 Comparison of generalized eukaryotic tDNA consensus sequence with class I genes transcribed by RNA polymerase III. Arrows indicate the site of initiation of transcription. N followed by a number indicates the number of intervening nucleotides between the indicated regions. In the sequences constituting Block A and B, N means that any nucleotide can be found at that position; R and Y represent always a purine or a pyrimidine, respectively. Boxed nucleotides (□) indicate mismatches with the tDNA consensus sequence.


'A' Block
'B' Block
Reference
Generalized eukaryotic tRNA genes
R R Y N N A R Y - G G
N11-18
N33-43
G T/A T C R A N N C
T T T T
(Tarboni, 1981)
Adenovirus VA I gene
G G T C T G G T G G
N14-17
N35
G T T C G A A A C
T T T T
Akusjarvi et al.
Adenovirus VA II gene
A G C C G G A G G G
N14
N33
G T T C G A G T C
T T T T
Akusjarvi et al.
Epstein Barr virus EBER I and EBER II genes
R G Y N N A R A/T G G
N12
N44-45
G T/G T/A C R A N N C
T T T T
Rosa et al.,1981
Hamster Alu-type
G G C T C A G A G G
N12
N34
G T T C A A T T C - - - - -
T T T T
Haynes and Jelin
1981
Human Alu-family
G G C G T G G T G G
N6
N61
G T T C A A G A C- - - - - -
T T T T
Deinninger et al.

illustrated in Figure 1.5. The homology in the 'B' block is particularly striking, whereas, there is lack of homology in the 'A' block, which is also common to some tRNA genes (Galli *et al.*, 1981). The conserved nucleotides are shown by deletion analyses to be essential for transcription of adenovirus VAI genes and human Alu family DNA sequences (Fowlkes and Shenk 1980; Guilfoyle and Weinmann 1981; Rosa *et al.*, 1981; Paolella *et al.*, 1983). Moreover, tRNA genes are shown to be efficient competitors of adenovirus VAI genes in transcription experiments *in vitro*, suggesting that faithful transcription of these genes depends on similar protein cofactors (Segall *et al.*, 1980; Guilfoyle and Weinmann 1981). Initiation of transcription of all class I genes occurs between 10 and 20 nucleotides 5' to the 'A' block boundary (Ciliberto *et al.*, 1982b).

5S RNA gene are the only genes that represent class II genes transcribed by RNA polymerase III. They share some properties with class I genes but show significant differences. The minimum sequence required to direct initiation of transcription of 5S DNA is near the centre of the transcribed region (Sakonju *et al.*, 1980; Bogenhagen *et al.*, 1980). Deletion of nucleotides from each end of 5S DNA has shown that internal control region (ICR) of 5S RNA genes is located between nucleotide 53 and 83 (Figure 1.5). In this case RNA polymerase initiates transcription at approximately 50 nucleotides from the 5' end of ICR. In addition to RNA polymerase III, at least two other protein factors are required. One of these has been purified and referred to as TF111A (Engelke *et al.*, 1980). Purified TF111A directs accurate transcription of 5S RNA genes by binding to central region of 5S DNA between nucleotide 45 to 96 (Engelke *et al.*, 1980; Sakonju *et al.*, 1981), but has no effect on transcription of tRNA genes or other class I genes. TF111A appears to regulate 5S RNA synthesis by binding to products of transcription and resulting in feed back inhibition of transcription (Pelham and Brown 1980; Honda and Roeder 1980).

Figure 1.6 Homology between tDNA Block 'A' consensus sequence and 5S DNA internal control region (ICR). Arrows indicate the site of initiation of transcription. In the sequences consituting Blocks 'A' and 'B', N means any nucleotide, R and Y represent a purine and pyrimidine respectively. Boxed nucleotides indicate homology between the two sequences. (Adapted from Ciliberto *et al.*, 1983).


'A' Block
'B' Block
-10
-2
tDNA
R R Y N N A R Y - G G
+53
+61
G T/A TCRANNC
TTTT
+1
+50
5S DNA
A G C C A A G C A G G
+60
+83
GTCGGGCCTGGTTAGTACCTTGGA
TTTT

No such transcription regulation factor(s) has yet been isolated for tRNA genes. Recent studies of Ciliberto *et al.*, (1983) on the structure of ICR of 5S RNA have shown that it can be internally dissected into two components (illustrated in Figure 1.6). The first half of ICR is homologous and functionally equivalent to 'A' block of tRNA genes. The first eleven bases of the 5S RNA ICR perfectly fit the 'A' block consensus sequence, the remaining sequence shares no homology. Ciliberto *et al.*, (1983) suggest that 'A' block could be responsible for the interaction with common components of polymerase III transcriptional machinery, whereas the second half of the 5S DNA promoter could be responsible for interaction with 5S RNA specific cofactor, for example TF111A (Ciliberto *et al.*, 1983).

4.a.v Termination of transcription

Eukaryotic tRNA genes contain a cluster of four or more thymidine residues (T) in the 3' flanking region of their non-coding DNA strand sequence. This is generally accepted to be the termination site of transcription, because all the tRNA precursor RNAs terminate with corresponding stretch of Uridine (U) residues (Garber and Gage, 1979; Koski and Clarkson, 1982; Sharp *et al.*, 1979; Hagenbuchel *et al.*, 1979). This cluster of T residues normally occurs within a few nucleotides of the 3' end of tRNA gene, but can occasionally occur much further down-steam. For instance, a TTTT sequence occurs 75 bp downstream of *X. laevis* $tRNA^{Lys}$ (Clarkson, 1983) and 18 bp downstream of *B. mori* $tRNA^{Ala}$ gene (Garber and Gage, 1979). Transcription terminates at these two sites yielding primary transcripts of 157 and 98 nucleotides respectively. Termination can also occur at an oligothymidylate stretch of chain length four or higher when removed from its natural position (adjacent to the coding sequence) and placed by in vitro recombination at a considerable distance downstream (Ciliberto *et al.*, 1982a), or when placed in the middle of a tRNA gene. For example, mutations which change the sequence TTTAATTTAT to

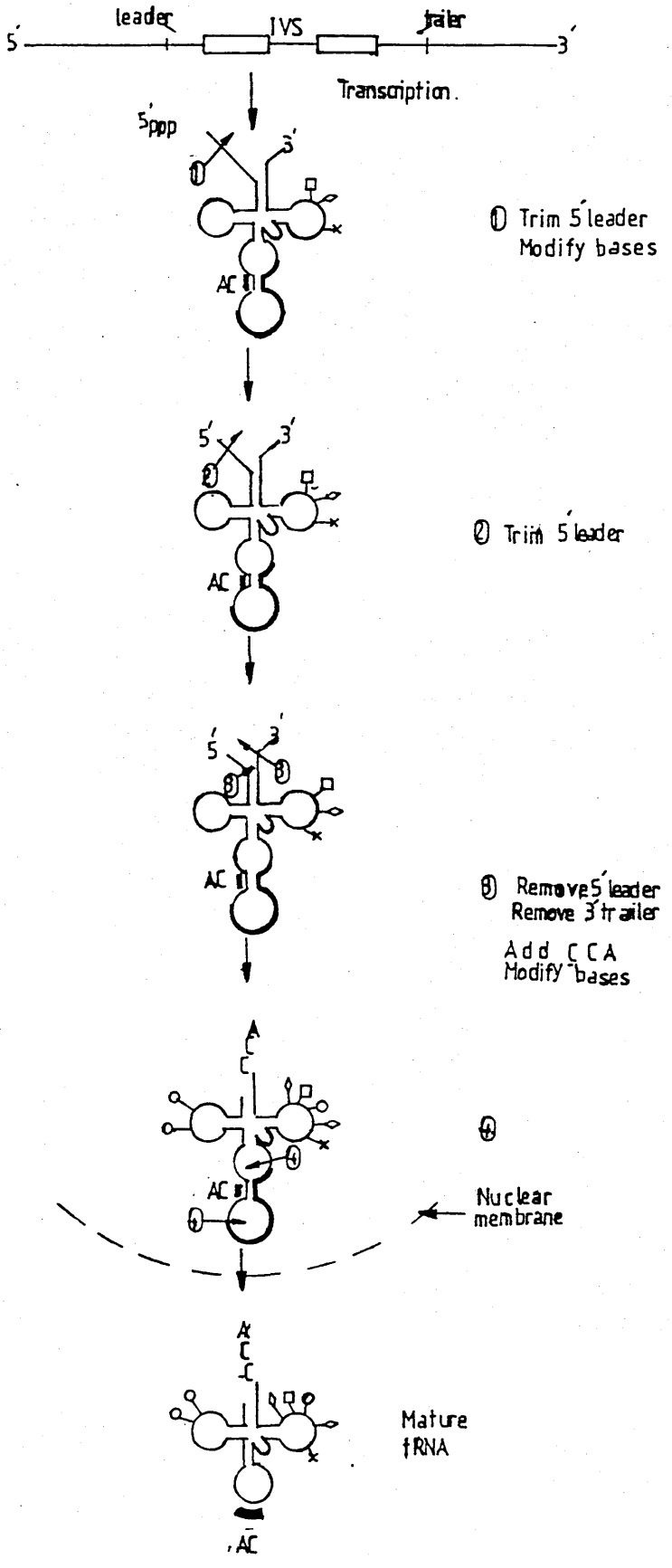
TTTAATTTTT in the non-coding strand of yeast $tRNA^{Tyr}$ gene results in premature termination at the new T stretch with 70-80% efficiency (Koski *et al.*, 1980). Some naturally occurring tRNA genes contain T tracts within their coding structure, for instance, *X. laevis* $tRNA^{Lys}$ gene contains an internal T_4 tract (corresponding to the anti-codon loop). In this case, more than 70% of the transcripts terminate prematurely (Clarkson 1983).

It appears that efficient termination of transcription of eukaryotic tRNA genes is solely dependent on a cluster of T tract and does not require GC rich regions or regions capable of forming hairpin loops as is the case in 5S RNA genes and prokaryotic tRNA gene termination (Korn and Brown 1978; Korn 1982; Rosenberg and Court, 1979).

4.b Post-transcriptional processing

As already mentioned, the maturation of tRNA precursor is a multi step ordered process involving nucleolytic size-reducing reactions, enzymatic addition of CCA to the 3' terminus, nucleotide modification to form dihydrouridine, pseudouridine and various methylated derivatives of all four nucleotides and in some cases excision of intervening sequence. Post-transcriptional processing steps have been investigated in some detail for yeast $tRNA^{Tyr}$ gene, after microinjection into *Xenopus* oocyte nuclei (DeRobertis and Olson, 1979; Melton *et al.*, 1980; Nishikura and DeRobertis 1981; DeRobertis *et al.*, 1982). In *Xenopus* oocyte nuclei yeast $tRNA^{Tyr}$ processing follows a strict temporal order as illustrated in Figure 1.7. The first step that occurs after transcription is the removal of four terminal nucleotides from the 5' leader and five specific bases are modified. Then another seven nucleotides are removed from the 5' leader sequence. In the third stage the remainder of the 5' leader sequence is removed and the trailer is replaced by CCA. At the same time additional bases are modified in the precursor molecule, but it still contains the intervening sequence. The last step before the

Figure 1.7 The strict temporal, ordered, post-transcriptional processing of yeast $tRNA^{Tyr}$ precursors into mature tRNA in injected *Xenopus* oocytes. The yeast $tRNA^{Tyr}$ primary transcript is 108 nucleotides long, contains a 19 nucleotide leader sequence, a trailer of 2 or 3 uridine residues and an intervening sequence (IVS) of 14 nucleotides. The mature yeast $tRNA^{Tyr}$ is 78 nucleotides long. The anticodon (AC) is protected by base pairing with IVS before splicing takes place. (ΨΤΥΨ represent modified bases). (Adapted from Melton *et al.*, 1980 and Kubli 1981).


leader
IVS
trailer
5'
3'
Transcription.
5'ppp
3'
① Trim 5' leader
Modify bases
AC
5'
3'
② Trim 5' leader
AC
5'
3'
③ Remove 5' leader
Remove 3' trailer
Add C C A
Modify bases
AC
④
Nuclear membrane
Mature tRNA
AC

mature tRNA appears in the cytoplasm is the excision of the intervening sequence and a few more bases are modified. There seems to be compartmentalization of the various enzymatic activities involved in processing: splicing and 5' end maturation enzymes are exclusively associated with the cell nucleus (Melton *et al.*, 1980; DeRobertis *et al.*, 1981; Castagnoli *et al.*, 1982), while 3' maturation enzymes including CCA nucleotidyltransferase are present both in the nucleus and the cytoplasm (Melton *et al.*, 1980). All the modification enzymes are present in the nucleus, with exception of enzyme(s) responsible for modification of G at position 26 which occurs in the cytoplasm (Melton *et al.*, 1980).

The reasons for transcription of 5' leader sequences are not known. It has been suggested that precursor tRNA with extra nucleotides at its 5' end is a substrate for some base modification enzymes, while the precursor without it is a substrate for others. Thus, the possible role of 5' leader sequences may be the sequential addition of modified bases or preventing the transport of immature transcripts from the nucleus into the cytoplasm. It seems improbable that the size of the precursor prevents its transport across the nuclear membrane. Perhaps the precursor binds to a processing enzyme, producing a bulky complex that cannot be transported through the nuclear pore (DeRobertis and Nishikura, 1981; Melton and Cortese, 1979). However Castagonli *et al.*, (1982) have illustrated that tRNA processing is dependent on secondary structure of nucleotides uniquely present in the precursor molecule and not affected by altering 5' and 3' flanking sequences. In addition to this Nishikura *et al.*, (1982) have examined 16 different point mutations within yeast $tRNA^{Tyr}$ gene, some of which affect processing steps. They conclude that the three dimensional tRNA like conformation of precursor molecule is essential for its interaction with maturation enzymes (5' and 3' trimming enzyme), whereas the base modification enzymes recognise certain regions of the precursor RNA

molecule.

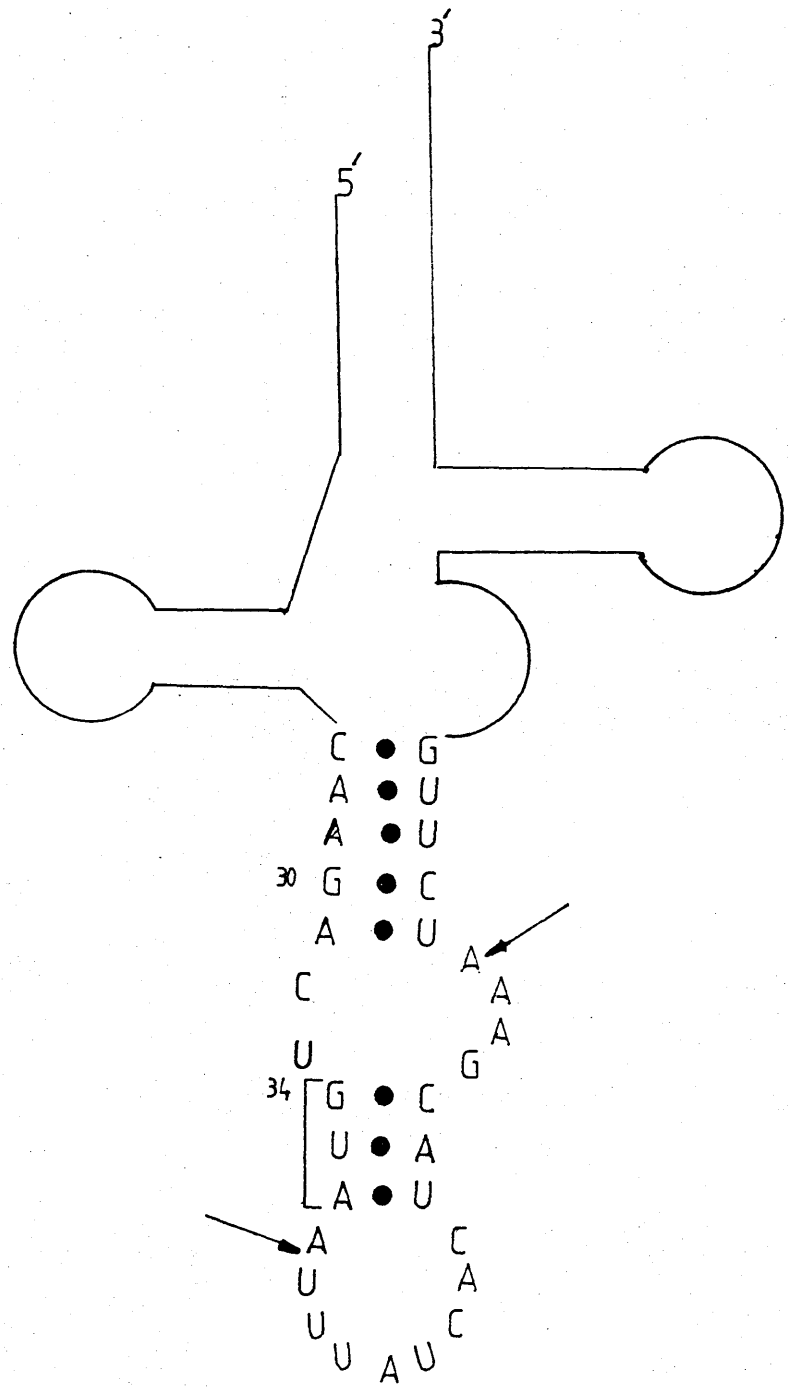
4.b.i Intervening sequences

Genes of some, but not all, eukaryotic tRNA species contain intervening sequences that are transcribed to form part of the primary transcript. These extra sequences are removed from the precursor molecule by an enzymatic process involving specific excision and ligation of the resulting tRNA half molecules (splicing) (Abelson 1979). These are two separate enzymatic reactions. The half-molecule intermediates have unusual 3' phosphate and 5' hydroxyl termini (Knapp *et al.*, 1979; Peebles *et al.*, 1979). Ligation of tRNA half molecules has been described in extracts of yeast, *Xenopus*, HeLa cells and also in *Xenopus* oocytes *in vivo* (Peebles *et al.*, 1979; DeRobertis, *et al.*, 1981; Filipowicz and Shatkin 1983). Interesting, in HeLa cells extract, tRNA half-molecules are ligated by a normal 3' 5' - phosphodiester bond derived from the 3' - terminal phosphate of the 5' half molecule (Filipowicz and Shatkin 1983).

Intervening sequences within tRNA genes were first detected in yeast. Goodman *et al.*, (1977) sequenced several yeast $tRNA^{Tyr}$ genes and found 14 nucleotides which are not represented in the mature tRNA. Similarly Valenzuela *et al.*, (1978) discovered intervening sequences, 18 or 19 nucleotides long in yeast $tRNA^{Phe}$ genes. Several other yeast tRNA genes are also shown to contain intervening sequences: For example the genes coding for $tRNA^{Trp}$ $tRNA^{Leu}$ and $tRNA^{Ser}_{UGG}$ (Ogden *et al.*, 1979; Etcheverry *et al.*, 1979). The presence of intervening sequences is also observed in a wide range of other eukaryotes: *Dictyostelium discoideium* $tRNA^{Trp}$ gene (Paffley and Sogin 1981); *Neurospora crassa* $tRNA^{Phe}$ gene (Selker and Yanofsky 1980) and *Drosophila* $tRNA^{Leu}$ gene (Robinson and Davidson, 1980). In higher eukaryotes intervening sequences are detected in $tRNA^{Tyr}$ gene of *X. laevis* (Muller and Clarkson, 1980) and within two chicken $tRNA^{Lys}$ genes (Wittig

Figure 1.8 The clover leaf model of yeast $tRNA^{Tyr}$ precursor.

Only the nucleotides in the anticodon arm and the 14 bp intervening sequence are shown (in unmodified form). The anticodon is indicated by brackets. Arrows indicate sites for the excision of the intervening sequence.


3'
5'
C G
A U
A U
30 G C
A U
A
C
A
A
U
G
34 G C
U A
A U
A
C
U
A
U
U
U A U
C

and Wittig, 1979). Interestingly intervening sequences have not yet been reported in any of the several different mammalian tRNA genes sequences so far, although mammalian cells contain the requisite "splicing" enzymes (Starding *et al.*, 1981; Laski *et al.*, 1982).

Location of intervening sequence in all known cases of tRNA genes is at 3' side of the base immediately adjacent to the anti-codon. Consequently, in the unspliced precursor RNA molecule, the acceptor stem, the dihydro-U and the TψC stems and loops retain many of the characteristic features of the standard clover lead model of tRNA. Whereas the intact anticodon stem is augmented by a second helical region of variable length. In some cases this helix is separated from the anticodon stem by a few un paired bases that could form bulge or interior loops (Ogden *et al.*, 1980). In most unspliced tRNA precursors the anticodon base pairs with the intervening sequence so that it is not available for codon interaction. However this is not true of every tRNA gene containing intervening sequence. For example the *Dictyostelium* $tRNA^{Trp}$ intervening sequence does not contain a sequence complementary to the anticodon (Paffley and Sogin, 1981) and the intervening sequence of *Xeonopus* $tRNA^{Tyr}$ is capable of base pairing with only one nucleotide of the anticodon (Muller and Clarkson, 1980; Clarkson, 1983).

The intervening sequence of tRNA genes shows great variation in length (from 13 bp. to 60 p.) and divergence in sequence. For example, $tRNA^{Phe}$ genes in yeast and *N. Crassa* contain intervening sequence, located exactly at the same position. While *N. Crassa* $tRNA^{Phe}$ intervening sequence is two or three nucleotides shorter than that of yeast, only 50% of the nucleotides in the two sequences are homologous whereas 91% of their bases are conserved in the mature $tRNA^{Phe}$ coding region. (Selker and Yanofsky, 1980). In contrast, the intervening sequence of yeast and *Drosophila* $tRNA^{Leu}$ genes contain

regions of strong homology. An eight base region (AAAAUCUU) in the *Drosophila* $tRNA^{Leu}$ intervening sequence is repeated exactly at the same location in yeast $tRNA^{Leu}$ (Robinson and Davidson, 1980). Broach *et al.*, (1981) and Olson *et al.*, (1981) have studied yeast $tRNA^{Ser}_{UCG}$ and $tRNA^{Ser}_{UCA}$ genes. They found that an intervening sequence is present in the single $tRNA^{Ser}_{UCG}$ gene but absent from the three $tRNA^{Ser}_{UCA}$ genes. This structual variability within a population of yeast tRNA genes suggest that a few, if any structural constraints are imposed by the function of intervening sequence.

The function of intervening sequences in the pathway of production of biological active tRNA is not fully understood. Intervening sequences within yeast tRNA genes have been extended or completely deleted without affecting transcription (Johnson *et al.*, 1980; Carrara *et al.*, 1981), suggesting that whatever the function of intervening sequence it is not intrinsically required for the expression of eukaryotic tRNA genes. Whether intervening sequences have a role in speeding up the evolution of tRNA genes, as proposed for the intervening sequences in cellular genes is very doubtful (Gilbert, 1978). On the other hand, the presence of intervening sequences within certain genes may hinder illegitimate recombination between closely related tRNA genes. For example, *Xenopus* $tRNA^{Tyr}$ and $tRNA^{Phe}$ (which share 70% sequence homology) are arranged on opposite DNA strands. Only the $tRNA^{Tyr}$ gene contains an intervening sequence, thereby possibly minimising the chance of recombination (Muller and Clarkson, 1980; Clarkson, 1983).

4.c Role of 5' flanking regions in transcription modulation

The control of transcription of eukaryotic tRNA genes depends on primarily intragenic events, but there is evidence that 5' flanking sequences immediately preceding the tRNA gene can influence both the choice of initiation and the efficiency of transcription of some tRNA genes, though not

essential for transcription of all eukaryotic tRNA genes. In some instances 5' flanking sequences are responsible for exerting a negative influence on eukaryotic tRNA gene expression. For example, in the 5' flanking region of *Drosophila* $tRNA^{Lys}$ genes, the consensus sequence GGCAGTTTTTG is responsible for poor transcriptional activity of this and other tRNA genes (DeFranco *et al.*, 1981). The positioning of this consensus sequence is an important factor. When located from position -13 to -23 it inhibits transcription, but not when moved closer or further away from the $tRNA^{Lys}$ gene. By placing this undecanucleotide in the flanking region of a transcriptionally "active" $tRNA^{Arg}$ gene DeFranco and coworkers (1981) have conclusively demonstrated that this consensus sequence is responsible for transcription repression. The mechanism of this external modulation is not clear yet, but an intriguing feature of this inhibitory tract is the stretch of thymidylate residues resembling RNA polymerase III termination signal. A similar tract, GGATTTTT, which precedes one of four *D. melanogaster* $tRNA^{Arg}$ genes, has been implicated in the poor transcription of this gene in a homologous cell-free extract (Dingermann *et al.*, 1982). Another example of an inhibitory effect of 5' flanking sequences is found in front of the variant $tRNA_i^{Met}$ gene of *X. laevis* (Clarkson *et al.*, 1981; Hipskind and Clarkson, 1983). *Xenopus laevis* contains two $tRNA_i^{Met}$ genes (Clarkson *et al.*, 1978. (Figure 1.3), one of which Met- A gene, codes for the initiator methionine tRNA found *in vivo*. The second Met- B gene, differs by a C→T transition at position 65 of the coding sequence and at many positions in its 5' flanking region. Met- B transcripts are not detected *in vivo* and are poorly expressed *in vitro*. Clarkson *et al.*, (1981) have reported that its transcription is stimulated, if the 5' flanking region is replaced by that of Met A gene or by plasmid DNA. Recently, Hipskind and Clarkson (1983) reported the presence of two inhibitory regions in the 5' flanking DNA. The first inhibitory region, contains a 9 bp

stretch of alternating purine and pyrimidine residues in the noncoding strand sequence (TGCGCGTGC) located between position -20 and -12. The second inhibitory region also contains a tract of alternating purine and pyrimidine residues, CATGCACAGCGCA), located between position -43 and -32. This second region has only a weak inhibitory effect when in its normal location, but can reduce transcription severely when brought closer to the gene (Hipskind and Clarkson, 1983). Nothing is yet known of the way in which these purine-pyrimidine tracts inhibit transcription. These tracts have the potential to form short stretches of left-handed Z-form DNA, which may be important in inhibition (Clarkson, 1983). Clarkson (1983) has reported that synthetic oligonucleotides that are capable of forming a Z-DNA status also inhibit expression of Met- B gene when placed in its 5' flanking region. The mechanism of transcription repression in the above examples seems to vary but their identical location emphasizes the importance of 5' flanking sequences.

Another view is that normal 5' flanking sequences of eukaryotic tRNA genes might have a role in their natural context, but are inactive, only in a particular transcriptional system. Evidence supporting this notion comes from transcriptional studies of the $tRNA^{Ala}$ gene of *B. mori* (Sprague *et al.*, 1980), which require its 5' flanking DNA for transcription in an homologous but not in an heterologous, cell-free extract. Raymond and Johnson (1983) have reported an A-rich sequence of fifteen nucleotides, perfectly conserved in the 5' flanking regions of four $tRNA^{Leu}$ genes of *S. cerevisae*, that is absolutely essential for efficient transcription in homologous but not in heterologous cell-free extracts. It is not clear whether this inactivity results from a loss of recognition of the gene by some component of the transcription apparatus or from a loss of the initiation site or the yeast and *B. mori* transcription apparatus are sensitive, at least in some cases, to DNA sequences outside the coding region.

5. Occurrence of repetitive elements in the 5' non-coding region of tRNA genes

Reiterated sequences have been observed in the flanking regions of several eukaryotic tRNA genes (Brodeur *et al.*, 1983; del Ray *et al.*, 1982; Gafner *et al.*, 1983; Eigel and Feldman, 1982; Han and Harding 1983; Rodi *et al.*, 1983). These repeated elements may play a role in the coordinate control of genes which are dispersed in the genome, for example by serving as targets for regulatory molecules (Britten and Davidson 1979). Reiterated sequences may play such a role by regulating the transcription of tRNA genes, RNA processing of tRNA precursor, or may be involved in the process of their dispersion over the genome.

The most overwhelming example of repetitive elements in the 5' non-coding region of tRNA genes is that of *Sigma* (σ). This is a DNA element of about 340 bp, that is reiterated many times in the yeast genome. Brodeur *et al.*, (1983) have estimated that there are approximately 30 copies of *Sigma* present in the haploid genome of *S. cerevisiae*, represented at 25 distinct loci. In each case the *Sigma* element is closely associated with a tRNA gene, though not all tRNA genes in the yeast genome have a *Sigma* element in their proximity. In seven instances, DNA sequence analysis has revealed that the *Sigma* element is 16-18 bp. upstream from the 5' end of a tRNA gene (del Ray *et al.*, 1982; Sandmeyer and Olson 1982) and in six of the seven analysed cases the elements are in the same orientation relative to the tRNA gene. *Sigma* elements bear a number of characteristics features of many transposable elements of both prokaryotes and eukaryotes. *Sigma* has 7 bp., inverted repeats at its end and is flanked by 56 bp., direct repeats. In addition, closely related strains appear to have *Sigma* located at different positions in the genome and its DNA sequence is highly conserved. Despite this, there is no direct evidence for

transposition of *Sigma* and no regions of extensive homology exist between *Sigma* and other known transposable elements.

The consistent association between *Sigma* and tRNA genes suggests that specificity of *Sigma* insertion depends on recognition of an adjacent tRNA encoding region. The proximity of *Sigma* to the site of initiation of tRNA transcription suggests the obvious possibility that *Sigma* confers some special transcriptional properties on the adjacent tRNA gene or it may reflect *Sigma's* preference for some structural feature of tRNA gene during the process of transcription (Brodeur *et al.*, 1983). However, *Sigma* elements show no preference for particular tRNA genes and it cannot be required for the transcription of all tRNA genes because the majority of tRNA genes do not contain *Sigma* element. The exact role of *Sigma* in tRNA gene expression has not yet been elucidated.

Another repetitive element called Ty 1 has been reported in the 5' non-coding region of several yeast tRNA genes (Eigel and Feldman 1982; Gafner *et al.*, 1983). The distance between the Ty 1 elements and the 5' end of the tRNA coding region ranges from 80 to 405 bp. The close association of Ty 1 element to tRNA genes raises the question whether these elements confer some special regulatory effect on the tRNA gene. As has been reported in the case of some structural genes, where wild type activity is completely lost or reduced by Ty insertion (Roeder *et al.*, 1980, Chaleff and Fink 1980). In another case regulated gene expression is turned into an over producing or constitutive type by Ty insertion (Montgomery *et al.*, 1982, Evrede *et al.*, 1980) So far, the available data suggests that the presence of this repetitive element does not affect the in vitro transcription of tRNA genes in yeast (Eigel and Feldman, 1982; Gafner *et al.*, 1983).

Information on the presence of repetitive sequences in the 5' non-coding region of higher eukaryotic tRNA genes is very limited. However Rodi *et al.*,

(1983) have observed at the 5' end of rat $tRNA^{Leu}$ gene a sequence homologous to the split promoter consensus sequence of RNA polymerase III. This sequence is bracketed by a 7 bp direct repeat enclosing a string of adenosine residues typical features of "Alu" type sequence. Han and Harding (1983) have reported a mouse $tRNA^{Lys}$ gene that has short direct repeats in its flanking regions which are characteristic of transposable elements.

Aims

Past studies concerned with the structure, organisation and experession of human tRNA genes have dealt with total populations of tRNAs. These have indicated that there are approximately 1000 tRNA genes contained within the human haploid chromosome complement (Hattlen and Attardi, 1971), distributed on different chromosomes, representing about 60 different tRNA species. Therefore, there are multiple copies of individual tRNAs in the human genome. However at present, relatively little is known about the organisation of human tRNA genes and further progress towards the study of structure and sequences governing the expression of human tRNA genes requires the analysis of specific tRNA genes. One way of isolating these is to screen a large collection of cloned human DNA fragments with a radioactive tRNA sequence probe.

The initial aim of this project was to screen a human genomic library and isolate clone(s) containing sequences corresponding to human tRNA gene(s). Hopefully after the initial isolation and characterization, the recombinant DNA clone(s) containing tRNA gene sequences would be mapped by the procedure of restriction endonuclease digestion and hybridization. This then would permit the subcloning of restriction fragments containing tRNA gene sequences, for use in determining human tRNA gene structure and arrangement and in transcriptional studies.

In practice this general approach revealed the presence of a truncated tRNA gene sequence in a recombinant clone. This recombinant also contains an Alu sequence and examination of this forms the substance of this thesis.

CHAPTER 3 MATERIALS AND METHODS

1. Materials

Unless otherwise specified all chemicals were Analar grade supplied by BDH Chemicals Ltd. or Fisons Scientific Apparatus. Where chemicals or equipment were obtained from other sources this is indicated in the text and a list of the name and addresses of the suppliers is given below:

(a) List of suppliers

Amersham International plc., Amersham, Bucks, England.

Aldrich Chemicals Co., Gillingham, Dorset, England.

A & J Beveridge Ltd., Edinburgh, Scotland.

BDH Chemicals Ltd., Poole, Dorset, England.

Beckman Instruments Inc., High Wycombe, Bucks., England.

Bethesda Research Laboratories (UK) Ltd., Cambridge, England.

The Boehringer Corporation (London) Ltd., Lewes, E. Sussex, England.

James Burrough Ltd., Fine Alcohol Division, London, England.

Calbiochem-Behring Corp., (UK) Bishops Stortford, Herts, England.

Conex-Lighting, Du-point (UK) Huntingdon, Cambs, England.

Collaborative Research Inc., Universal Scientific Ltd., (UK distributor) London, England.

Difco Laboratories, West Molesey, Surrey, England.

Fisons Scientific Apparatus, Loughborough, Leics., England.

Fluka, A.G. Fluorochem Ltd., Colnbrook, Bucks, England.

Kodak Ltd., Kirby, Liverpool, England.

LKB Instruments Ltd., LKB House, South Croydon, Surrey, England.

New England Biolabs, CP Labs. Ltd., (UK distributors), Bishops Stortford, Herts, England.

PL Biochemicals Inc., Northampton, England.

Pharmacia Ltd., Milton Keynes, England.

Schleicher and Schuell, Andermann and Co., (UK distributor)

E. Molesey, Surrey, England.

Sigma London Chemicals Co. Ltd., Poole, Dorset, England.

Whatman Lab. Sales Ltd., Maidstone, Kent, England.

Worthington, Flow Labs. Ltd., Irvine, Scotland.

U.V. Products, Winchester, Herts., England.

Universal Scientific Ltd., London, England.

2. Methods

(a) General Procedures

Experiements involving the cloning or propagation of bacteriophage lambda (λ) and plasmids carrying eukaryotic DNA were performed in accordance with GMAG guidelines for recombinant DNA research. During the course of this work, a number of procedures were frequently used. The following section describes these general procedures.

(b) Description of bacterial strains

Charon 4A: Lac 5 bio 256 KH 54, BW1 nin 5 QSR 80 (Blattner et al., 1977)

E. coli HBH 2600: K12, r_k^-, m_k^+, Sup E^+, Sup F^+, met^- (Hohn, 1979).

The plasmids used in this study are listed and described in Table 3.1. The plasmids were stored both as naked DNA in 10mM Tris HCl. pH 8, 0.1mM EDTA and in cultures of transformed *E. coli* HB 101.

E. coli HB 101: K12, F^-, pro^-, Leu^-, thi^-, Lac Y^-, hsd R^-, end A^-, rec A^-, rps L20, ara-14, gal K2, Xyl-5, mH-1. Sup E44 (Bolivar and Backman, 1979).

Table.3.1 Summary of the plasmids used in this study.

Name	Vector	Antibiotic resistance or other marker	Origin/Reference
PXt 210	pBR 322	ampicillin resistance	A 3.18 Kb repeating unit of X. laevis, containing 8 different tRNA genes. see Figure 1.3 and Bryan et al., 1981) cloned into the HindIII site of pBR322 (A gift from S.G. Clarkson)
pt 38	pBR 322	tetracycline resistance	$tRNA^{Asn}$ gene (from pXt210) as an AluI fragment (residues 1834-2044 inclusive, see Figure 1.3) dC-tailed with synthetic oligonucleotides and then cloned into dG-tailed PstI site of pBR 322 (A gift from S.G. Clarkson)
pt 63	pBR 322	tetracycline resistance	$tRNA^{Tyr}$ gene (from pXt 210) as a HhaI fragment (residues 325-585 inclusive, see Figure 1.3) dC-tailed with synthetic oligonucleotides and then cloned into dG-tailed PstI site of pBR 322 (A gift from S.G. Clarkson. Muller and Clarkson, 1980).
pt 75	pBR 322	tetracycline resistance	$tRNA^{Phe}$ gene (from pXt 210) as a HhaI fragment (residues 163-325 inclusive, see Figure 1.3) dC-tailed with synthetic oligonucleotides and then cloned into dG-tailed PstI site of pBR 322 (A gift from S.G. Clarkson; Muller and Clarkson, 1980).
pt 51	pBR 322	ampicillin resistance	$tRNA^{Lys}$ gene (from pXt 210) as an EcoRI - HindIII fragment (residues 2736-3182 inclusive, see Figure 1.3), cloned into EcoRI and HindIII sites of pBR 322 (A gift from S.G. Clarkson).
pt 145	pBR 322	ampicillin resistance	$tRNA^{Met}$ A gene (from pXt 210) as an EcoRI fragment (residues 931-1112 inclusive, see Figure 1.3), cloned into the EcoRI site of pBR 322 via linkers. (A gift from S.G. Clarkson).
pNB 1 pNB 4 pNB 21	pAT 153	ampicillin resistance	Subclones of λht 3 prepared in this study

(c) Preparation of radioactivity labelled probes

(i) 3' end labelling of tRNAs with T_4 RNA ligase

The method used to label 3' end of tRNAs with T_4 RNA ligase was based on the method of Bruce and Uhlenbeck (1978). RNA ligase from T_4 infected *E. coli* catalyzes the addition of a nucleoside 3', 5' - bisphosphate onto the 3' terminus of tRNA resulting in tRNA molecule one nucleotide longer with a 3' terminal phosphate. The human 4S RNA was end-labelled as follows: 50μCi of 3',5'-[5-^{32}P] cytidine bisphosphate (pCp, specific activity 2000-3000 Ci mmole^{-1}) was lyophilised in a siliconised Eppendorf tube (1.5 ml). To this, lμg 4S RNA, 1ul 10XT_4 ligase buffer (500mM HEPES pH 7.5, 200mM $MgCl_2$, 30mM DTT), 1μl T_4 RNA ligase (P.L. Biochemicals Inc., 1000 units/mg), 1μl (10% v/v) dimethyl sulphoxide (DMSO) and water was added. This reaction mixture (10 μl)was incubated for 12 hours at 4^{o}C. The reaction was terminated by freezing in dry ice-acetone and subsequently lyophilizing the sample. The dried sample was resuspended in 99% deionized formamide, 0.05% Xylene cyanolFF and applied to a pre-electrophoresed 10% denaturing polyacrylamide gel. The labelled tRNA was visualized by autoradiography. The tRNA- containing band was then eluted from the gel as described in Materials and Methods section 2.e.

(ii) 5' - end labelling of tRNAs with T_4 polynucleotide kinase

The method used was based on the method of Chang *et al.*, (1976). Bacteriophage T_4 induced polynucleotide kinase catalyses the transfer of the γ-phoshate from ATP to the 5' - hydroxyl terminus of the polynucleotide.

Because of the requirement of a 5' - hydroxyl terminus, the substrate must be dephosphorylated with alkaline phosphatase prior to the labelling step, and the phosphomonoesterase activity must then be eliminated to avoid degradation of ATP and loss of label from the polynucleotide substrate. The most effective method of inactivation of alkaline phosphatase is repeated extraction with phenol followed by dialysis or ethanol precipitation. However, this method is tedious and results in some losses of polynucleotide. An alternative method of inactivation is boiling in the presence of NTA (Nitrilotriacetic acid) to chelate the zinc from the alkaline phosphatase methalloenzyme (Simsek *et al.*, 1973).

Phosphatase reaction

About one microgramme of human 4S RNA was incubated for 30 min. at 55°C with 5µl of 50mM Tris-HCl pH 8.0 and 2µl of bacterial alkaline phosphate (BAP, 10 mg/ml, Worthington, Flow Labs, Ltd., Irvine, Scotland). (The BAP was purified to a form free of any detectable endo - or - exo - DNAase or RNAase activity by chromatography on Sephadex G.75). The bacterial alkaline phosphatase was inactivated by addition of 40mM NTA pH 8.0 and allowed to stand at room temperature for 20 minutes. The reaction mixture was boiled for 1 minute followed by a brief spin to bring all liquid to the bottom of the reaction tube.

Polynucleotide kinase reaction:

The phosphorylation reaction was carried out in a volume of 20µl with final concentration of 50mM Tris HCl pH 8.0, 10mM $MgCl_2$, 5mM DTT. To this 100 µCi of lyophilised adenosine 5' - [γ-^{32}P] triphosphate (specific activity

5000 Ci/mmol,) and 5 units of T_4 polynucleotide kinase (500 units ml^{-1}, P.L. Biochemical Inc.). The reaction mixture was incubated for 30 minutes at 37°C. The reaction was terminated as in the 3' - end labelling method.

(iii) DNA nick-translation

50 uCi (α-^{32}P) dATP (specific activity 3000 Ci $mmol^{-1}$) was dried in a siliconised 1.5ml Eppendorf tube. To this was added, 1µg DNA, 3µl 10X nick translation buffer (500mM Tris-HCl pH 7.9, 50mM $MgCl_2$) 1µl of 100mM 2-mercaptoethanol (freshly diluted), unlabelled deoxyribonucleotides: dCTP, dGTP, dTTP each to a final concentration of 100µM and enough distilled H_20 to bring the total volume to 30µl. Then 1µl of DNA polymerase I (Boeringer Corp.) (this commercial preparation had contaminating DNA endo nuclease activity) was added and incubated at 14°C for 1 hour. The reaction was stopped by addition of 1µl of 0.5M EDTA and boiling at 65°C for 10 mins. A blue Eppendorf tip was filled with sephadex G50 and equilibrated with 10mM Tris HCl pH 8.0, 10mM NaCl, 1mM EDTA. This tip was placed in a 5ml screw-top ampoule with a small Eppendorf tube at the bottom. The entire column was spun at 3000 rpm (5000 g) for 15 min. until the G50 appeared dry. Then the sample was added and the column was spun again at 3000 rpm for 15 min. The nick-translated material was collected in the Eppendrof tube. The labelled DNA was precipitated from 0.3M sodium acetate by additon of 2.5 volumes of cold 95% ethanol. The pellet was washed with 1ml ethanol, dried and resuspended in distilled H_20.

(d) Polyacrylamide denaturing slab gels

These gels were used to isolate the labelled tRNA species. The 200 x

400 mm. glass plates were washed thoroughly, wiped with ethanol, siliconised ('Repelcote') and then rinsed with distilled H_20. The dry plates were assembled using the slot former (comb), the plasticard side spacers (1mm thick) between the two plates and carefully sealing the sides and bottom with waterproof tape (Univeral Scientific Ltd., London). The acrylamide stock solution was 40% acrylamide 1.34% bisacrylamide, and was deionized by stirring in 2% Aberlite monobed resin MB-1 for 1 hour. To pour a 10% gel, 25 ml of the acrylamide mix was added to 42 gm. of Urea, 10 ml 10X Tris-borate, EDTA, buffer (T.B.E.: 0.89M Tris-base, 0.89M Boric acid, 0.025M NA_2 EDTA, pH 8.3). The volume was adjusted to 100ml with distilled H_20. 0.6 ml of freshly made 10% (w/v) ammonium persulphate was then added and the mixture was filtered and degassed. 32 µl of NNN'N' tetramethyl ethylene diamine (TEMED) was added, the gel was poured and allowed to polymerise. The polymerised gel was pre-electrophoresed for 2 hrs (with TBE). The labelled sample was heated at 90°C for 2 min, then chilled and loaded on the gel. Electrophoresis was carried out at 30mA (1200V) until the dye marker, bromophenol blue, had migrated the required distance, (or at 200 V overnight). After electrophoresis, one of the glass plates was removed, the gel was covered with cling film (Saranwrap) and autoradiographed with Kodak X-0mat HI film. After the required exposure (1/2 hour - 1 hour) the film was developed for 4 minutes in Kodak DX 80, rinsed and fixed in Kodak FX40 for twice the clearing time.

(e) Elution of Nucleic Acids from polyacrylamide gels

The method used was based on the established method of Maxam and Gilbert (1980). The elution buffer composition was 0.5M ammonium acetate,

10mM magnesium acetate, 1mM Na_2 EDTA, 0.1% SDS. The small end of a 100µl blue Eppendorf tip was melted and sealed over a low flame. It was then plugged with a small piece of siliconised spun glass wool. The required band was cut from the gel, transferred into the sealed Eppendorf tip, and crushed into fine bits with a glass rod. The glass rod was rinsed with 0.6 ml elution buffer. The Eppendorf tip was covered and incubated for 16 hrs at 37°C. Then a corner of the Eppendorf tip was cut to allow the eluate to drain through the glass wool plug into a siliconised corex tube. The Eppendorf tip was rinsed thoroughly twice with 0.2 ml elution buffer. The nucleic acids were precipitated by addition of 2 ml cold 95% ethanol and chilled at -70°C for 1 hr. The nucleic acids were recovered by centrifugation at 12000g for 15 min at 4°C. The supernatant was removed, and the nucleic acids were reprecipitated from 0.3M sodium acetate in 2.5 times the volume of cold 95% ethanol. The pellet was washed with 1 ml cold ethanol, dried and resuspended in distilled H_20 or 10mM Tris HCl pH 7.8 0.1mM Na_2 EDTA.

(f) Preparation of *E. coli* HBH2600 host cells

A single colony of *E. coli* HBH2600 was innoculated to 10 ml of LB broth (LB (Luria-Bertani) broth: 1% Bacto-tryptone (Difco, West Molesey, Surrey), 0.5% Bacto-yeast extract (Difco), 0.5% NaCl, adjust to pH 7.5 with NaOH, autoclaved at 15 psi for 20 min.) and incubated overnight at 37°C. 0.4ml of this overnight culture was diluted into 40 ml of fresh LB broth and incubated at 37°C until absorbance reached 0.5 (absorbance measured at 650 nm). The cells were then pelleted by centrifugation at approximately 5000 g for 10 min at 4°C. The supernatant was poured off, the cells were resuspended in 4 ml of 10mM $MgSO_4$ and aerated by shaking at 37°C for 1 hr. These cells

were stored at 4°C until further use.

(g) Plating of λ Phage particles

LAM plates: 1% Bacto-tryptone, 0.5% Bacto-yeast extract, 1% NaCl, 1% Bacto-agar (Difco), 0.288g $MgSO_4$ per 100 ml of medium, autoclaved at 15 psi for 20 mins. After sterilization 1 ml of sterile 20% Maltose and 0.5 ml 0.5M $CaCl_2$ solution were added to every 100 ml of medium. LAM 'top' agar: same as above, except instead of 1% Bacto agar, 0.7% Agarose was used.

3×10^4 recombinant phage particles were absorbed onto 0.3 ml of freshly prepared bacterial competent cells (HBH 2600) in a sterile conical Sterilin screw cap plastic centrifuge tube and incubated at 37°C for 10 mi. Then approximately 10 ml of LAM top agar (stored at 42°C) was added and the mixture was poured onto 15 cm LAM petri dishes. The plates were incubated at 37°C overnight (14-16 hrs), after which time the plaques were near confluence. The plates then were refrigerated for an hour or longer to harden the top agar.

(h) Transfer of λ phage DNA to Nitrocellulose filters

The method used was based on the method of Benton and Davis (1977). Nitrocellulose filter discs, to fit the 15 cm plate were cut from rolls of millipore filter paper (HAMP 00010, type HA, pore size 0.45μm) or Schleicher and Schull filter paper (BA 85, 0.45μm). The dry nitrocellulose filter discs were marked with a soft pencil or ballpoint pen. Phage on the LAM plates were absorbed (transferrred) to the nitrocellulose filter disc by placing the dry filter with marked side down on the surface of agar (avoiding air bubbles). During

this absorption the filters and underlying agar were marked for orientation in three asymmetric locations by stabbing through with an 18- gauge needle. After absorption (2-3 min) the filters were carefully lifted from the agar plate. The DNA was denatured, neutralised and bound to the filters as follows: The filters were first dipped in a solution of 0.1M NaOH/1.5M NaCl for 1 min. then dipped in a solution of 0.2M Tris-HCl pH 7.5 and finally rinsed in a solution of 2XSSC for 1 min. (1XSSC = 0.15M NaCl, 0.015M sodium citrate pH 7.0). The filters were then blot dried and baked in a vacuum over for 2 hrs. at 80^{o}C.

(i) Hybridization of DNA bound to Nitrocellulose filters

The method used was based on the methods of Wahl *et al.*, (1979) and Thomas (1980). The filters were first wetted in 4XSSC at room temperature for 30 min. Then the filters were washed in pre-hybridization buffer for 3-10 hrs. at 42^{o}C (10 ml per filter). The pre-hybridization buffer contained 5XSSC; 50% formamide; 50mM sodium phosphate buffer pH 6.5; 10X Denhardt's solution (1X Denhardt's solution = 0.02% bovine serum albumin, 0.02% polyvinylphyrolidone, 0.2% Ficoll 400); 10µg/ml poly A; 10 µg/ml poly C; 50 µg/ml sonicated, heat denatured salmon sperm DNA and 0.1% SDS. After pre- hybridization, the filters were hybridized with gentle, continuous agitation at 42^{o}C for 10-16 hrs. in fresh pre-hybridization buffer containing ^{32}P-labelled DNA or RNA probe (specific activity 0.5 - 5 x 10^{7} cpm/µg). Prior to hybridization the probe was denatured by boiling at 100^{o}C for 10 min. The pre-hybridization and subsequent hybridization were carried out in thermally sealed plastic bags.

After hybridization, the filters were washed twice (with continuous agitation) in 300 ml of 5XSSC, 0.1% SDS solution at room temperature. This removed the majority of unbound counts. Thereafter the filters were washed

with 4 changes of 0.5XSSC, 0.1% SDS for 15 min. at 50°C. If the background on the filters, detected with a monitor was unacceptably high, more stringent washing conditions (0.3 - 0.1XSSC, 0.1% SDS) were used as a final step. The filters were blot dried, mounted on cardboard, covered with saran wrap and exposed to Kodak X-Omat HI X-ray film with intensifying screen (Cronex lighting - plus, Du Pont, U.K. Huntingdon, Camb). at -70°C for 2-4 days.

(j) Preparation of Recombinant λ phage suspension

After exposure for the appropriate time, the autoradiographs were examined for positive signals corresponding to the plaques on the respective LAM agar plates. The keyed markings on the agar plates and nitrocellulose filters facilitated the locations of these 'positives'. The plaque from the region of a plate corresponding to a positive on the autoradiograph were picked using a sterile short stem Pasteur pipette. The plug of agar containing the phage particles was then resuspended by gentle agitation in 1 ml of sterile SMC (100mM NaCl, 10mM $MgSO_4$, 50mM Tris-HCl pH 7.5, 0.01% gelatine). A drop of chloroform was added to the suspension to aid lysis and prevent further bacterial growth.

(k) Preparation of λ Phage DNA

The phage (10^5 to 10^6 plaque forming units) was mixed with 0.3ml of fresh HBH2600 competent cells (stationary culture) and then incubated at 37°C for 15 min. The mixture was diluted into a litre of warmed NZY broth and shaken overnight at 37°C. The NZY broth contains (per litre) 10 g of NZ amine (a casein hydrolyzate), 5 g of yeast extract, 5 g of NaCl and 2 g of $MgCl_2$

-6 H_2O. The overnight culture was harvested when lysis had occurred. Chloroform was added (20 ml/l) and shaken for 2 min. at 37°C. This immediately killed all surviving bacteria without affecting the phage. The chloroform was allowed to settle on ice. The bacteria were pelleted by centrifugation at 6,000 g for 10 min. at 4°C. The clear supernatant was made up to 0.5M with respect to NaCl and 120 g of polytheylene glycol 6000 (PEG) was added. The PEG was dissolved by gentle shaking and the solution was allowed to stand for more than 2 hrs at 4°C. The PEG precipitated phage were pelleted after centrifugation at 12000 g for 30 mins at 4°C and subsequently resuspended gently in 10mM $MgSO_4$, 50mM Tris HCl pH 7.5. Any resuspended PEG was extracted with an equal volume of chloroform. The supernatant was recovered, caesium chloride was added such that the final refractive index of solution was 1.3813. This was then centrifuged for 18 hrs in a Beckman Ti 50 Rotor at 42K (126,000 g) at 12°C. The opaque phage band was collected through the side of the centrifuge tube with a syringe and dialysed against 25mM NaCl, 1mM $MgSO_4$, 10mM Tris-HCl pH8. The dialysate was ethanol precipitated. The pellet was resuspended in 10mM Tris-HCl pH 7.5, 0.1mM EDTA.

(l) Analysis of Recombinant DNA

Digestion of Recombinant DNA with restriction endonuclease

Restriction enzymes were obtained from New England Biolabs. (U.K. distributed: CP Laboratories, Bishops Stortford, Herts), Bethesda Research Laboratories (Cambridge, U.K.) or Boehringer Mannheim (Boehringer Corp., London, Lewes, East Sussex). For all restriction enzymes, the digestion buffer used was similar to that recommended by the suppliers. All reactions were performed in 1.5ml Eppendorf microfuge tubes. A typical

reaction mix would include: DNA, the appropriate amount of 10X buffer to make the final concentration 1X, the amount of restriction enzyme and reaction time indicated by the specific activity of the enzyme. Volumes, quantities of enzyme and time of digestion also depend on the requirement of the particular experiment. For instance for an analytical experiment enough enzyme was used to digest the DNA in one hour, but for large scale preparations 1/10 as much enzyme (proportional to DNA) was used and digested overnight.

(m) Electrophersis of DNA

In general agarose gels were used to separate fragments over 1000 bp and acrylamide gels were used to separate fragments less than 1000 bp.

(i) Agarose gels

Vertical Slab gels:

Clean, dry 200 x 200mm glass plates were assembled using side spacers (3mm) between the two plates and sealing the sides and bottom with waterproof tape (Universal Scientific Ltd., London). To pour a 1% agarose gel, 1.5g of agarose (Type II Sigma, Poole, Dorset) was dissolved in 150ml of 1X electrophoresis buffer (40mM Tris-HCl, 20mM Sodium acetate 2mM EDTA, pH 8.0) by heating in boiling water (or on a hot plate), cooled to 65^{o}C and poured between the sealed glass plates. The slot former was positioned and the gel was allowed to set. After the gel was completely set (30-40 minutes at room temperature), the slot former was carefully removed and the gel was mounted in the electrophoresis tank. The samples were mixed with loading buffer (5-10% glycerol, 7% Sucrose, 2.5% Ficoll, 0.025% bromophenol blue), loaded into the slots and electrophoresed at 100-150V (approximately 30mA). The electrophoresis time was varied according to the requirement of the particular experiment. In this system bromophenol blue dye marker migrates

with DNA molecules of 600-800 bp.

After electrophoresis, the agarose gel was carefully immersed in 500 ml of H_2O containing 10 μg/ml ethidium bromide. The DNA bands were visualised by placing the gel on U.V. transilluminator (Model C-62, U.V. products, Winchester, Hants). (U.V. safety goggles were worn). The gel was photographed using a Polaroid Cu-5 land camera fitted with a yellow filter and Polaroid 665 (positive/negative) film. The negative was immersed in 1M sodium sulphite solution to remove all the developer and then thoroughly washed in cold water.

Horizontal slab gel

The horizontal slab gels were prepared in the horizontal gel tank apparatus Model H1 or H5 (obtained from Bethesda Research Laboratories, Cambridge, U.K.). The H1 apparatus was set up with 2% agarose wicks. The main gel was prepared in the electrophoresis buffer as described for vertical slab gels and poured to a thickness of approximately 3mm. The samples were prepared and applied as in vertical slab gels. Electropheresis was carried out at 150-200V until the dye marker had travelled three-quarters of the gel. After electropheresis the gel was stained and viewed as described for agarose vertical slab gels.

(ii) Acrylamide gels

Acrylamide gels were prepared in the vertical gel apparatus Model V.16 (obtained from Bethesda Research Laboratories, Cambridge, U.K.) Clean, dry 16X16 cm. glass plates were assembled using 1.5mm side spacers and a spacer across the bottom. For a 4% acrylamide gel, gel mixture was prepared containing 10mls acrylamide stock (19% acrylamide, 1% N, N' -methylene

bisacrylamide) 5ml glycerol, 5ml 10X TBE, 30ml H_2O, 0.4 ml of freshly prepared 10% (w/v) ammonium persulphate was added, the mixture was filtered and degassed. 5 ml of the gel mixture was removed and 40 µl TEMED was added. This was used to seal around the side spacers and form a small plug at the bottom of the gel. When the plug was set (approximately 5 min) 30µl of TEMED was added to the rest of the mixture, stirred and poured between the sealed glass plates. The slot former was positioned and the gel was allowed to polymerise. The polymerised gel was pre-electrophoresed at 150V for 30-60 min. Electrophoresis was carried out at 200V for 3-4 hrs. In this system bromophenol blue migrates with DNA molecules of 50-100 bp. and Xylene cyanol with DNA molecules 400 bp. After electrophoresis one of the glass plates was removed, leaving the gel attached to the back plate. The gel was stained and visualised as described for vertical agarose slab gels.

(n) Transfer of DNA from agarose gels to Nitrocellulose filters

The method used was based on the method of Southern (1975). After electrophoresis was completed the agarose gel was stained and photographed as described above. The unused area of the gel was trimmed with a razor blade. The trimmed gel was soaked in a solution (500 mls) of 0.2M NaOH, 0.6M NaCl at room temperature for 30 min. to denature the DNA. The gel was then neutralised by soaking in a solution of 1M Tris-HCl (pH 7.5), 0.6M NaCl for 30 min. at room temperature. A glass plate (20x20 cm) was covered with Whatman filter paper (no.17), which was placed in turn over a tray (or reservoirs) containing transfer buffer (20xSSC) such that the filter paper formed a continuous wick. The neutralised gel was placed on the damp filter and two perspex spacers were placed 0.5 cm. to either side of the gel. A piece of

nitrocellulose filter (Schleicher and Schull, BA 85) was cut about 1-2mm larger than the gel in both dimensions. The nitrocellulose filter was floated in a solution of 2XSSC until completely wet. The damp nitrocellulose filter was placed very carefully on top of the gel so that no air bubbles were trapped between the filter and gel. Two pieces of 3MM Whatman filter paper were cut slightly larger than the gel, soaked in 2XSSC and layered on top of the nitrocellulose filter (again avoiding air bubbles) with overlaps sitting on the perspex spacers. The filters were then covered with 5 pieces of 3MM paper and a stack of paper towels (cut slightly smaller than 3MM paper). The towels were weighed-down and the transfer was left for 16 hrs. at room temperature. After the transfer, the nitrocellulose filter was removed, rinsed in 2XSSC for 5 mins. air-dried and baked at 80^{o}C under vacuum for 2 hrs.

(o) Extraction of DNA from agarase gels

(i) Electroelution of DNA

After electrophoresis, the agarose gel was stained and the DNA fragment band of interest was located and cut out of the gel using a sharp scalpel. The slice of gel was transferred to a dialysis bag filled with 0.5X electrophoresis buffer (TBE). The gel slice was allowed to sink to the bottom and the dialysis bag was tied just above the gel slice (avoiding any air bubbles), leaving enough fluid to keep the gel slice in constant contact with electrophoresis buffer. The tied bag was immersed in an electrophoresis tank. The DNA was electroeluted by electrophoresis at 100 V (20mA) for 2-3 hrs. During this time DNA electroeluted out of the gel and onto the inner wall of the dialysis bag. The polarity of the current was reversed for 1-2 minutes to release the DNA from the wall of the dialysis bag. Then the dialysis bag was carefully opened, all the liquid surrounding the gel slice was collected and extracted twice with phenol and once with chloroform. The DNA was recovered by ethanol precipitation.

(ii) Agarase extraction method

Another method of extracting DNA from agarose gel is to use the enzyme agarase (Calbiochem. Cat. No. 121811). This method was based on the method of Finkelstein and Rownd (1978). In this case the excised gel slice (approximately 0.1 g) was transferred to a 1-ml plastic syringe containing 100µl of 0.1M Tris-HCl (pH 5.95) and crushed by forcing through a 21 gauge hypodermic needle into a pyrex centrifuge tube. Agarase (50µg/0,1 g gel slice), dissolved in (1mg ml^{-1}) 0.1M Tris-HCl (pH 5.95), was added to the homogenate. The centrifuge tube containing the gel slurry was incubated at 37°C for 2 hrs. The agarose was removed by centrifugation at 4°C for 30 min. at 48,000g. The DNA supernatant was filtered by passing through a 0.45µm Swinnex filter (Millipore). The filtrate was phenol extracted and DNA was recovered by ethanol precipitation.

(iii) Low melting point Agarose Method

Another method of extracting DNA from agarose gel was to use low melting point agarose (Bethesda Research Laboratories, Cambridge, U.K. Cat.No. 5517). In this case the excised gel slice was heated to 65°C until the agarose was melted. Then two volumes of 50mM Tris HCl (pH 8) 0.5mM EDTA were added to the melted gel and the mixture was transferred to a 37°C water bath. The mixture was extracted twice with equal volume of phenol, saturated with 0.1M Tris HCl (pH 8) and once with chloroform. The DNA was recovered by ethanol precipitation.

(p) Construction of Subclones

(i) Ligation

2μg of plasmid pAT 153 DNA cleaved with restriction endonuclease HindIII was incubated with 1.4 units of bacterial alkaline phosphatase in 50μl of 60mM NaCl, 7mM $MgCl_2$, 60mM Tris HCl pH 8 for 60 min. at 60°C. The mixture then was extracted twice with phenol, once with chloroform and three times with diethyl ether. 30μg of λht3 DNA was cleaved with restriction endonuclease HindIII and the products were fractionated on a 1% low melting point agarose gel. The restriction fragment(s) to be subcloned (2.8 Kb double band) was excised from the gel and DNA was extracted. The HindIII cleaved pAT 153 DNA was mixed with 5-10 times molar excess of 'insert' DNA and ethanol precipitated. The precipitate was collected by centrifugation and resuspended in 50μl of ligation buffer containing 50mM Tris HCl pH 7.6, 10mM $MgCl_2$, 10mM DTT, 1mM ATP, 10 units of polynucleotide ligase (Bethesda Research Laboratories, Cambridge, U.K.) was added to mixture and incubated at 4°C for 2 hrs.

(ii) Transformation

The method used was based on the method of Dagert and Ehrlich (1979). An overnight starter culture of *E. coli* HB 101 (5ml) was diluted into 50mls of fresh LB broth and incubated at 37°C until $O.D._{650}$ reached 0.5. The cells were harvested by centrifugation at 2,500g for 10 mins. at 2°C. The pelleted cells were resuspended in 20ml ice cold 0.1M $CaCl_2$ and maintained at 0.4°C

for 30 min. The cells were pelleted again by centrifugation and resuspended in 0.5ml of ice cold 0.1M $CaCl_2$ and incubated at 4°C overnight. The cells were now ready for transformation. The treated cells (100µl) were mixed with 10µl of ligation mixture and incubated on ice for 10 min. The temperature was raised to 37°C for 5 min. The bacteria then were diluted into 2ml LB broth and grown for 1 hr at 37°C. The transformed bacteria (200µl) were plated on LB agar plates containing either 50µg/ml ampicillin or 20µg/ml tetracycline and scored for drug resistance. A small scale preparation of plasmid DNA was made from several transformants and checked for size and restriction pattern.

(q) Preparation of Plasmid DNA

A rapid preparation of plasmid DNA:

This small scale method based on the method of Holmes and Quigley (1979) was used to prepare plasmid DNA rapidly from a single colony on LB agar plate. Using a sterile toothpick, a transformant colony was transferred into a 1.5ml Eppendrof microfuge tube. The cells were suspended in 25ul STET buffer (8% Sucrose, 5% Trition X-100, 50mM EDTA, 50mM Tris HCl pH 8) by vortex mixing. To this suspension, 2ul of freshly prepared lysozyme (10mg/ml) was added. The mixture was swirled on ice until it became viscous, and then the tube containing reaction mixture was placed in boiling water for 1 min. The suspension was then centrifuged at 12,000g for 10 min. at room temperature. The supernatant was recovered. DNA was precipitated at -18°C for 10 min. by addition of an equal volume of isopropanol and centrifugation at 12,000g for 5 min. The nucleic acid concentration from such preparation was approximately 0.3µg plasmid DNA, sufficient for one or two restriction digests.

Large Scale preparation of plasmid DNA

The method used was based on the method of Clarke and Carbon (1976). A single colony of transformed bacteria was innoculated into 10ml of LB broth (containing appropriate antibiotic) and grown overnight at 37°C. The overnight culture was diluted into 1000ml of fresh LB broth and incubated at 37°C until

$O.D_{650}$ reached 0.6, at which time 1.6g of chloramphenicol (160µg/ml) was added to the culture which was shaken overnight at 37°C. The cells were harvested by centrifugation at approximately 5000g for 10 min at 4°C. The pelleted cells were resuspended in 250 ml of cold T.E, buffer (10mM Tris HCl pH8, 1mM EDTA) and resedimented by centrifugation. The pellet was resuspended in 25ml of ice cold sucrose solution (25% sucrose in 50mM Tris HCl pH 8.0) and 5ml of freshly prepared lysozyme (10 mg/ml in 250mM Tris HCl pH 8.0). After addition of lysozyme it was essential to keep the mixture on ice to minimise contaminating DNAase activity). To this mixture 20ml of cold detergent solution (0.1% Triton X-100, 0.05M Tris HCl pH 8, 0.05M EDTA) was added to promote lysis. The mixture was swirled on ice until it became very viscous (10 min) and then centrifuged at 30,000 rpm for 30 min at 4°C in a Beckman 60 Ti rotor (90,720g). The supernatant was collected in a graduated tube, 0.95g of solid caesium chloride and 0.1ml of 10mg/ml ethidium bromide was added per ml of supernatant. The refractive index of the final solution was between 1.340 and 1.396. The solution was transferred to quick-seal polyallomer tubes (Beckman No. 342414, 1" x 3.5") and centrifuged for 16 hrs. in a Beckman 50 VTi rotor at 50,000 rpm (241,200g) at 20°C. After centrifugation the nucleic acid bands were visualised by U.V. illumination (long wavelength). The lower band which contained the covalently closed circular plasmid DNA was collected by puncturing the side of the tube with a 21 gauge syringe needle. The closed circular plasmid DNA was further purified by a

second centrifugation in a Beckman 50 Ti rotor at 33,000 rpm (98,550g) for 24-48 hrs and again the lower band was collected. Ethidium bromide was removed from the DNA by three extractions with isoamyl alcohol (3-methyl 1-butanol, Koch-Light) and caesium chloride was removed by extensive dialysis against 10mM Tris-HCl pH 7.5, 1mM EDTA, 50mM NaCl. The dialysate was ethanol precipitated. The pelleted DNA was resuspended in 10mM Tris HCl pH 8, 0.1mM EDTA.

(r) *In vitro* transcription of DNA

(i) HeLa cell extract

A cell-free extract (S100) was prepared from HeLa cells by the method of Weil et al., (1979). The cells were grown in Eagle's minimal essential medium (Gibco) supplemented with 10% calf serum at 37°C. It is important that the cells are healthy and fast growing. Cells from a 6 litre culture at density of $5\text{-}6\text{x}10^5$ cells per ml. were harvested and washed in Dulbecco's phosphate-buffered saline (140mM NaCl, 3mM KCl, 8mM Na_2HP0_4, 0.5mM $CaCl_2$ and 0.5mM $MgCl_2$). The packed cell volume was measured after centrifugation at 800g for 5 min at 4°C. The cells were then washed with hypotonic buffer containing 10mM HEPES-KOH pH 7.9, 10mM KCl, 1.5mM $MgCl_2$ and 0.5mM DTT. The cells were resuspended in 2 packed cell volumes of hypotonic buffer containing phenylmethyl sulphonyl fluoride (PMSF) freshly prepared, 6mg/ml in ethanol). The cells were allowed to swell on ice for 10 min. and then disrupted by homogenisation in a Kontes all glass Dounce homogenizer (B pestel) with 10 strokes. The extent of cell lysis was generally 90-95% as monitored by light microscope after staining with toluidene blue. To this homogenate, 1/10 volume of buffer containing 0.3M HEPES-KOH pH 7.9, 1.4M KCl and 0.03M $MgCl_2$ was added and the suspension was

centrifuged at 100,000g for 1 hr to yield S100 extract. The S100 cell free extract was stored in small aliquots at -70°C.

ii) Transcription assay.

The protocol used for transcription assay was based on the method of Standring et al., (1981). A typical transcription reaction mixture (50µl) contained 12mM HEPES-KOH (pH 7.9), 60mM KCl, 0.12mM EDTA, 0.3mM DTT, 3mM $MgCl_2$, 0.6mM unlabelled ATP, CTP and UTP, 10µCi of [α^{32}P] GTP (specific activity 500 Ci $mmol^{-1}$), 25µM GTP, 4µg of supercoiled plasmid DNA, 2.2µg α-amanitin per ml, 5mM creatin phosphate, 12% (Vol/Vol) glycerol and 30µl of S100 HeLa cell free extract. The mixture was incubated at 30°C for 1 hr. and the reaction was quenched by the addition of 50µl of a solution containing 0.8% SDS, 0.5mg/ml yeast tRNA, 100µg/ml proteinase K, 0.2M NaCl and 20mM EDTA. The reaction mixture was maintained at 22°C for 30 min. and then diluted with 300µl of 0.2M NaCl; 0.02M EDTA pH 8. The mixture was extracted once with phenol and once with chloroform. The aqueous phase was recovered and 1ml of 95% ethanol was added. The nucleic acids were precipitated at -70°C for 15 min. followed by centrifugation at 12,000g for 15min. The pellet was dried and resuspended in 10µl of loading buffer containing 99% deionised formamide, 0.05% Xylene cyanol FF, 0.05% bromophenol blue. The samples were heated at 90°C for 2 min. chilled, loaded on a 10% acrylamide-7M urea gel (1mm thick) and electrophoresed at 300V for 12 hrs. After electrophoresis the gel was autoradiographed on a sheet of X-ray film.

(s) Purification of tRNA from DNA filters

The method used was based on the method of Bovre and Syzbalski (1971). The overall procedure involves the immobilisation of denatured DNA on nitrocellulose filters, the hybridization of the specific probe to this DNA, washing non-specifically bound probe RNA from the filter, and the elution of the hybridisable material. Between 10 and 50μg of DNA was put in a glass tube and H_2O was added to make a total volume of 2.5ml. This was made 0.25M with respect to NaOH and allowed to stand at room temperature for 10 minutes, at which time, 7.4ml H_20, 10ml 10XSSC and 0.1ml 4N HCl (0.02M) was added. This mixture was then slowly filtered through Schleicher-Schull nitrocellulose filter discs (HA 0.45μm), which had been pre-soaked in 5XSSC. This was followed by a slow wash of 5XSSC, and 3 quick washes (20 ml) of 5XSSC. The filters were then air-dried and baked for 2 hrs. at 80°C under vacuum.

To hybridise the tRNA probe to the DNA bound on the filter, the filter was placed in clean screw up scintillation vial and covered with 2XSSC solution. The radioactively labelled probe was added, and the vial was shaken gently for 16 hrs. at 67°C. Then, the liquid was removed and the filter given 4Xl hourly washes in 2XSSC. The filter was then air-dried, returned to a clean vial and the hybridized Cerenkov counts determined. To elute the hybridized material, the filter was covered with 1-2ml of 0.01XSSC and boiled for 15 min. The liquid was transferred to a centrifuge tube with 0.2ml of 2M sodium acetate pH 5, 0.5 mg/ml yeast carrier tRNA and 3.5ml cold 95% ethanol. The RNA was precipitated at -80°C for 1 hr, followed by centrifugation at 33,000 rpm in Beckman Ti 50 rotor (98,500g) for 1 hr. The precipitate was dried under vacuum, resuspended in loading buffer, and subjected to gel electrophoresis

(10% acrylamide, 7M Urea gel). The nucleic acid were detected by autoradiography.

(t) DNA sequencing by the Maxam and Gilbert chemical method

(i) 5' end labelling of DNA fragments

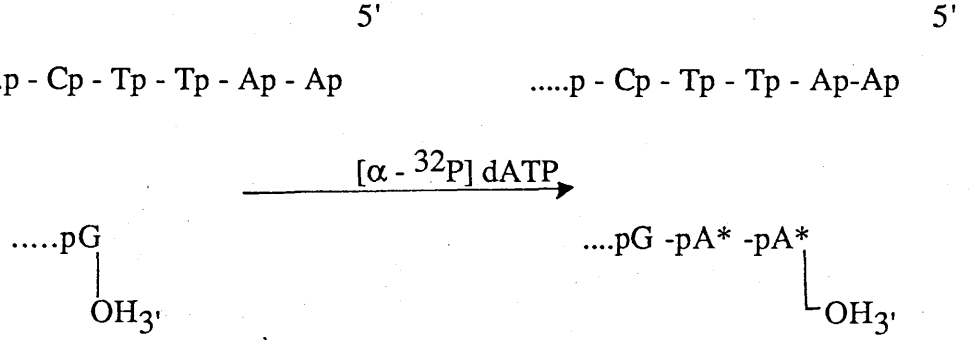
The DNA restriction fragment was first subjected to phosphatase action to remove the 5' phosphate group and then end- labelled at the 5' end using [γ-^{32}P] ATP and T_4 polynucleotide kinase. To remove the 5' phosphate group from DNA, 1-2µg of the restirction fragment was incubated with 10µg (10mg/ml) of bacterial alkaline phosphatase in 100µl 10mM Tris-HCl (pH 8.0), 0.1mM EDTA, at 37°C for 1-1 1/2 hr. Then the reaction mixture was extracted with an equal volume of phenol, the phenol was removed by repeated extractions with an equal volume of diethyl ether and the DNA was precipitated from 0.3M sodium acetate in 2.5 volumes of cold 95% ethanol. For the kinase reaction, the dephosphorylated DNA was dissolved in 10µl of kinase buffer(50mM Tris HCl pH 8.0, 10mM $MgCl_2$, 5mM DTT). To this 60µCi [γ -^{32}P] ATP (Specific activity 5,000 Ci/mmol, 10mCi/ml, in aqueous solution), followed the 1ul polynucleotide kinase (P.L. Biochemicals Inc., U.K.) was added and incubated at 37°C for 30 min. After this incubation, 40µl 2.5M ammonium acetate and 160µl ethanol were added. The mixture was chilled at -70°C for 15 min. and centrifuged (12,000g) for 5 min. The supernatant was removed to radioactive waste, the pellet was dissolved in 100µl 0.3M sodium acetate and 300µl of ethanol was added. The mixture was again

chilled at -70°C and centrifuged. The supernatant was removed and the pellet was dried under vacuum.

This protocol was used for DNA fragments which have 5' 'protruding' ends, for instance after cleavage with EcoRI, whereas fragments that have flush ends e.g. Sma or recessed ends e.g. HhaI, the kinase reaction was modified. In these cases, Tris-HCl pH 9.5 (instead of pH 8.0) and 1mM spermidine was included and the reaction mixture was heated to 90°C for 2 min, then chilled on ice before polynucleotide kinase was added.

(ii) 3' end- labelling of DNA fragments

The large fragment of *E. coli* polymerase I (Klenow fragment) can be used to fill in the 3' recessed end of DNA molecule with ^{32}P labelled deoxyribonucleotides. The [α - ^{32}P] dNTP added to the reaction depends on the sequence of the protruding 5' termini at the end of DNA. For example ends created by cleavage of DNA with EcoRI can be labelled with [α - ^{32}P] dATP.



A typical 3' end labelling reaction (filling-reaction) mixture contained 50mM Tris-HC1 pH 7.2, 10mM $MgSO_4$, 0.1mM DTT unlabelled deoxyribonucleotides dCTP, dGTP, dTTP each to final concentration of 100uM, 50 uCi of [α - ^{32}P] dATP (Specific activity ~3000 Ci $mmol^{-1}$).To this

mixture, lμg of DNA and 1 unit of Klenow fragment of DNA polymerase I (Bethesda Research Laboratories, Cambridge, U.K.) was added. The mixture was incubated at room temperature for 30 min. The reaction was stopped by addition of 1μl 0.5M EDTA. Then the reaction mixture was phenol extracted and ethanol precipitated as described in the 5' end labelling procedure.

(iii) Separation of labelled ends

To produce DNA fragments labelled at one end only for DNA sequence analysis, the DNA fragment was either cleaved with a restriction enzyme (that cleaved the DNA fragment asymetrically, so that the resulting fragments can be separated by gel electrophoresis) or the two strands were separated by strand separation procedure.

The principle of strand separation is that two complementary DNA strands may form differnt secondary structures when allowed to renature. This secondary structure confers different electrophoretic mobilities upon the complementary strands. In a typical strand separation experiment, the labelled DNA fragment was dissolved in 10ul of 0.3M NAOH, 10% glycerol and 0.03% bromophenol blue. The mixture was allowed to stand at room temperature for 10 min. then chilled on ice and immediately loaded onto a non-denaturing gel (8% (w/v) acrylamide, 0.24% (w/v) N', N'-methylene bisacrylamide in 0.5XTBE) and electrophoresed at 300v (8v/cm) for 12 hrs. The separated strands were located by autoradiography and eluted from the gel.

(iv) DNA sequence analysis

The method used was based on the established method of Maxam and Gilbert (1980). The ^{32}P- end labelled DNA (preferably at least 100,000

Cerenkov cpm) was dissolved in 10µl H_2O and divided into 4 aliquots of 2.5µl each, one aliquot for each track. The reactions used for DNA sequence analysis were specific for guanine (G), guanine + adenine (A+G), cytosine + thymine (C+T) and cytosine (C). These reactions were done in 1.5ml siliconised Eppendorf tubes with identification scratched on the respective tube. The reaction times were arrived at empirically to give cleavage equally distributed over the length of DNA. All 4 reactions were in 2 steps, the first reaction serving to weaken the chain and the second resulting in complete cleavage. All reactions were timed such that the precipitation steps were all done simulataneously.

For the A+G reaction, 2.5µl ^{32}P labelled DNA, 1µl carrier DNA (1mg/ml sonicated calf thymus DNA), 11µl H_2O and 2.5µl pyridinium formate (4% formic acid adjusted to pH 2.0 with pyridine) were incubated at 30°C for 70 min. After this incubation, the reaction mixture was frozen and then lyophilised. The resulting residue was redissolved in 10µl H_2O, frozen and relyophilised to dryness. This sample was then ready for the piperidine (strand cleavage) reaction.

Meanwhile for the G reaction, 2.5µl of ^{32}P labelled DNA, 1µl carrier DNA, 90µl DMS Buffer (50mM sodium cacodylate, 10mM Magnesium chloride, 0.1mM EDTA pH 8.0) and 0.5µl dimethylsulphate ('Gold label', Aldrich Chemical Co. Gillingham, Dorset) were mixed and incubated at 20°C for 5 min. Then 24µl of DMS stop (1.5M sodium acetate, 1M β-mercaptoethanol, 100µg/ml yeast tRNA) and 400µl ethanol (Analar grade)

were added. The mixture was chilled at -70°C for 15 min.

The C and C+T reactions were done simultaneously, and were identical but for the fact that C reaction was done in presence of high salt. For the C+T reaction, 2.5µl of ^{32}P labelled DNA, 1µl carrier DNA and 6µl H_20 were added to the Eppendorf tube. For the C reaction 2.5µl of ^{32}P labelled DNA, 1µl carrier DNA and 8µl saturated NaCl were added to the Eppendorf tube. To each of these tubes 15µl hydrazine (Kodak Ltd., Kirkby, Liverpool) was added and incubated at 20°C for 8 minutes (C+T) and 10 minutes (C only). Then, 60ul of 'Hydrazine Stop' (0.3M sodium acetate, 0.1mM EDTA, 50µg/ml yeast tRNA) and 250µl of ethanol were added. The mixture was chilled at -70°C for 15 min. It was important not to allow the ethanol mix to solidify as hydrazine may precipitate into a separate phase, carrying the labelled DNA with it.

The G, C+T and C tubes were centrifuged at 12,000g for 5 min. The supernatant was removed with a drawn-out Pasteur pipette. The pellet was redissolved in 60µl 0.3M sodium acetate, the 250µl of ethanol was added. The mixture was chilled at -70°C for 15 min. and centrifuged at 12,000g. The pellet was twice washed with 70% ethanol, chilled for 5 min and centrifuged for 5 min. after each wash. These samples were ready for the piperidine reaction.

Strand scission with piperidine

The dry DNA pellet from all the four modification reactions was dissolved in 100µl of 10% (v/v) piperidine (Koch - Light Laboratories, Colnbrook, Bucks) and heated to 90°C for 30 min. Weights were placed on top of the reaction tubes to prevent the caps from popping open. After this incubation, the

samples were frozen and lyophilised to dryness. The residue was redissolved in 20µl H_2O and lyophilised again. This step helped to remove the residual piperidine and ensured that the sample was concentrated at the bottom of the tube. The dry samples were dissolved in 99% deionized formamide, 0.05% Xylene cyanol FF.

(v) Sequencing gels

For most of the DNA sequence analysis, 8% gels containing 7.6% (w/v) acrylamide, 0.4% (w/v) N.N-bisacrylamide, 50% (w/v) Urea (8.3M), 100mM Tris-borate pH 8.3, 2mM EDTA, 0.07% (w/v) ammonium persulphate and TEMED catalyst, were used. The above gel mixture was injected into a 0.3X200X400mm mould and a gel slab was prepared as described earlier. The polymerised gel was pre-electrophoresed for 2 hrs (in TBE buffer) limiting the current to 25mA. The initial voltage was usually between 1.0 - 1.5KV and increased later to 2KV. After pre-electrophoresis, the samples were heated at 90°C for 2 min. then chilled, 1-1.5µl of each sample was loaded onto the gel (approximately 10,000 cpm per track) and electrophoresed at 25mA, 1.5-2KV such that the gel heats up to about 45°C. The length of run (electrophoresis) depended on the particular requirement (size of fragment). Usually three successive loadings of the sample were applied to the gel in order to vary the time of electrophoretic separation of the DNA fragments. After electrophoresis, one of the glass plates were removed, the gel was covered with saran wrap and autoradiographed at -70°C.

CHAPTER 4 RESULTS

In this study a human foetal liver DNA library cloned in bacteriophage lambda (λ) Charon 4A, was screened for human tRNA gene sequence. The library was a generous gift from Dr. T. Maniatis. The construction of this library was as described by Lawn *et al.*, (1978).

1. Probes used for screening the library

The first choice of probe to screen the library was either HeLa cell 4S RNA (kindly provided by Professor R.H. Burdon) or human placental 4S RNA (kindly provided by Dr. J.P. Goddard). The 4S RNA was labelled either at the 5' end by treatment with polynucleotide kinase and 5' - [γ - ^{32}P] ATP or at the 3' end with cytidine 3' - 5' - [5-^{32}P] bisphosphate (pCp) using RNA ligase. The ^{32}P-labelled 4S RNA was further fractionated by polyacrylamide gel electrophoresis and the band corresponding to full length tRNAs was eluted from the gel (see Figure 4.1) and used as a sequence probe for subsequent screening experiments. The details of procedures for labelling 4S RNA and preparing the radioactive probe are described in Materials and Methods section 2c. The specific activity of this electrophoretically purified '4S RNA' probe was 1-6x10^6 cpm/μg. Initial attempts to screen the human genomic library using this '4S RNA' probe (labelled either at 5' end or 3' end) were unsuccessful. One of the possible reasons was the low specific radioactivity of the labelled '4S RNA' probe.

As an alternative, it was decided to use a DNA probe. DNA probes have the advantage that they can be labelled to higher specific activities than RNA probes, for example by the *in vitro* nick-translation procedure (see Materials and Methods section 2.c.(iii). For this purpose, a cloned DNA fragment

Figure 4.1 Gel electrophoresis of 3' end labelled human 4S RNA.

Human 4S RNA was labelled at the 3' end with 3', 5'-[5'-^{32}P] cytidine bisphosphate using RNA ligase. The labelled RNA was electrophoresed on 16% acrylamide-7M area gel. Full length tRNA were located by electrophoresis of unlabelled yeast tRNA in an adjacent track, and staining the strip of gel with "stains-all". The radioactive band comigrating with tRNA was excised from the gel and used as a probe for tRNA gene sequences ('4S RNA').

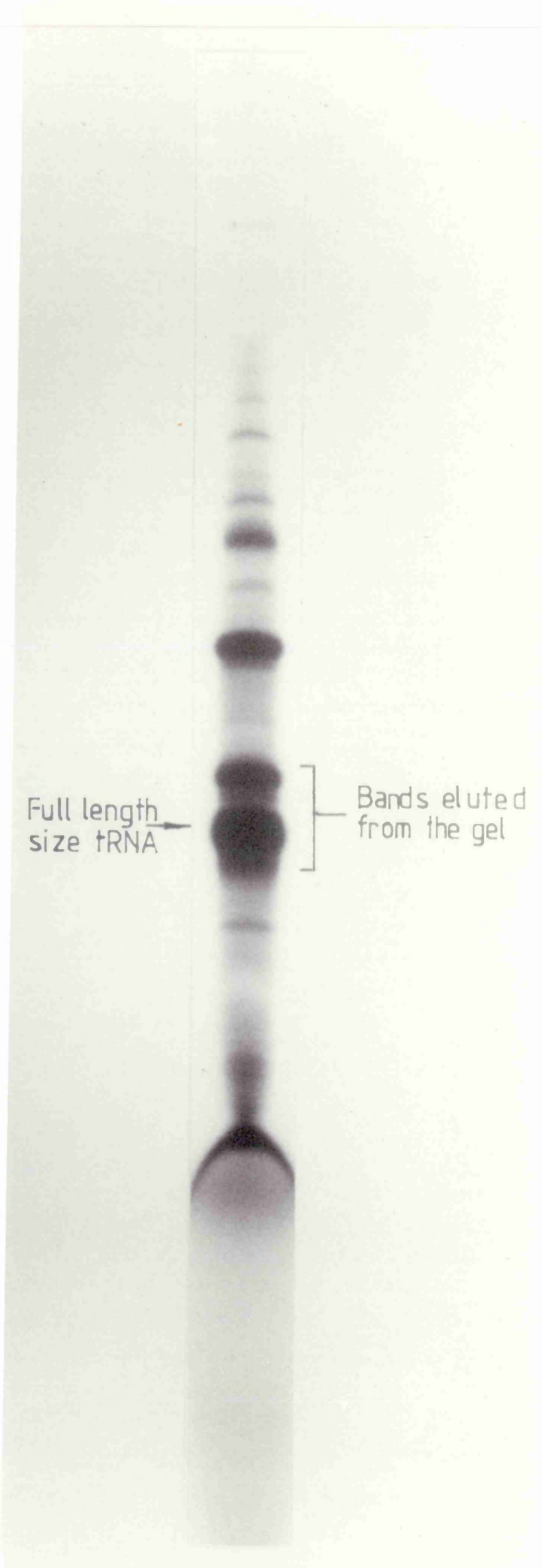

Full length
size tRNA
Bands eluted
from the gel

Figure 4.2 Hybridization of [^{32}P] pCp labelled human tRNA probe ('4S RNA') to pXt 210 DNA.

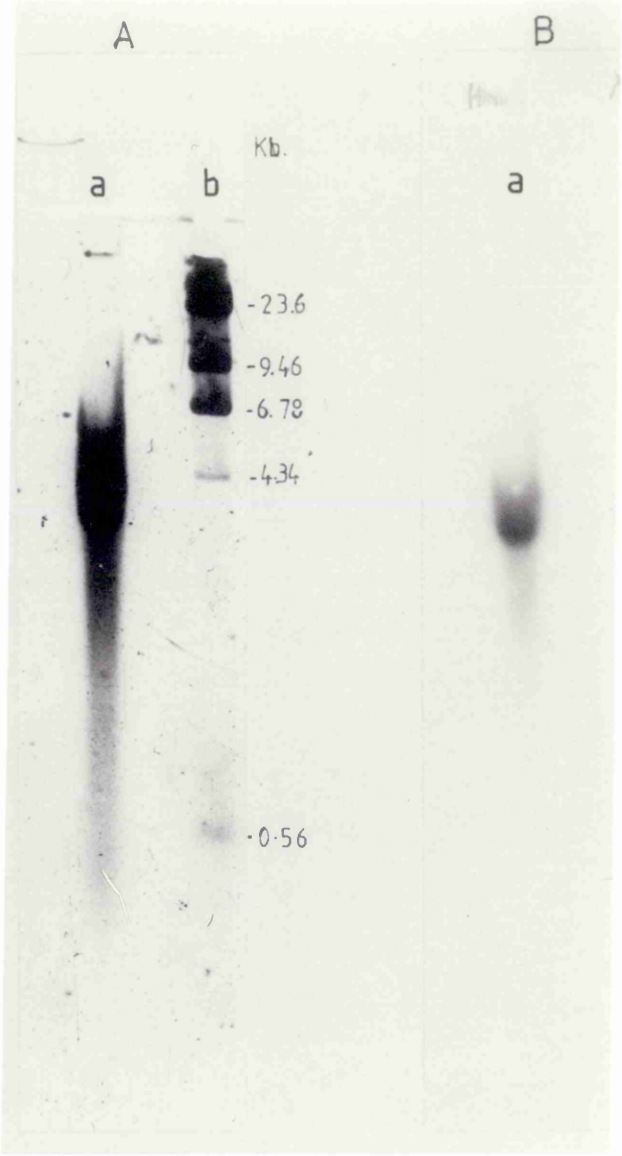
The plasmid pXt210 DNA was digested with restriction endonuclease HindIII to release the 3.18 Kb *Xenopus* insert. The digestion products were fractionated on an 0.8% agarose gel, then transferred to nitrocellulose filter by the method of Southern, (1975) and hybridized to [^{32}P] labelled human '4S RNA'.

A. Electrophoretic analysis of HindIII cleaved pXt210 DNA on 0.8% agarose gel.

lane a - pXt210 DNA cleaved with restriction endonuclease HindIII

lane b - size marker - λDNA, cleaved with HindIII.

B. Autoradiograph of Southern Blot prepared from the gel shown in A.


A
B
Kb.
a
b
a
-23.6
-9.46
-6.78
-4.34
-0.56

containing *X. laevis* tRNA genes was selected as a possible probe. In *X. laevis* eight tRNA genes lie on a 3.18 Kb fragment of DNA, that is repeated tandemly several hundred times. A plasmid with this 3.18 Kb segment cloned in the HindIII site of pBR322 (referred to as pXt210) was kindly gifted by Dr. S.G. Clarkson. While this study was in progress two other groups reported on the isolation of human tRNA genes: Santos and Zasloff (1982), cloned human chromosomal fragments containing $tRNA^{Met}$ sequences, and Buckland *et al.*, (1983) isolated three DNA clones containing human tRNA genes. These groups used pXt210 DNA as a hybridization probe for their screening experiments. Another reason for using this heterologous probe was the high degree of conservation between the coding regions of the tRNAs in the two organisms. A comparison of individual mammalian tRNA sequences with those of *X. laevis* genes present in pXt210, indicates that some of the tRNAs represented on the recombinant plasmid pXt210 share more than 90% homology with the respective mammalian tRNA sequence. For example, the published mammalian tRNA sequence for phenylalanine, tyrosine and asparagine are identical to the respective *X. laevis* tRNA sequence, the tRNA sequence for methionine and leucine show some differences and the mammalian tRNA sequence for alanine is not yet known (Gauss and Sprinzl, 1983; Johnson *et al.*, 1985). To confirm this homology between *Xenopus* and human tRNA genes, the recombinant plasmid pXt210 DNA was cleaved with restriction endonuclease HindIII, the digestion products were then fractionated on an 0.8% agarose gel, transferred to nitrocellulose filter by the method of Southern (1975) and hybridized to [^{32}P] pCp 3'-labelled '4S RNA'. As shown in Figure 4.2, the labelled human '4S RNA' strongly hybridizes to the 3.18 Kb fragment of plasmid pXt210. The conclusion from this preliminary analysis is that *Xenopus* tDNA and human '4S RNA' cross-hybridize under the conditions of high stringency used (for hybridization conditions see Materials and Methods section 2i). As a result of this experiment, nick-translated 3.18 Kb fragment of

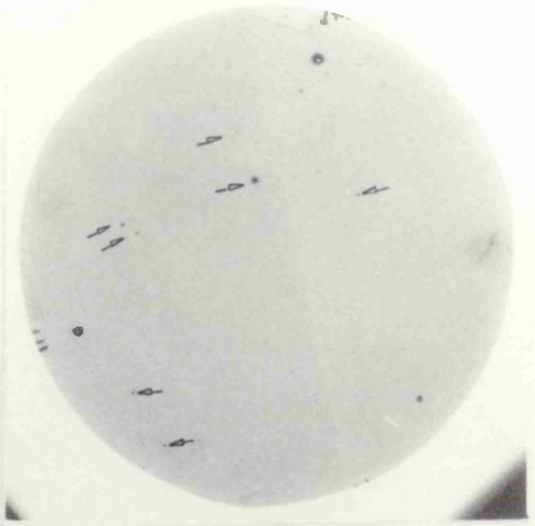
Figure 4.3 Screening of the human library for tRNA gene sequences, using nick-translated, [^{32}P] labelled 3.18 Kb. DNA (*Xenopus* tDNA) as a hybridization probe (specific activity >2x10^7 cpm/μg).

A - Preliminary round of screening. Autoradiograph of a nitrocellulose filter from a 15 cm. agar plate containing approximately 1.2x10^4 recombinant plaques. Details of screening procedure and hybridization conditions are described in Materials and Methods section 2f;g;h and i. Arrows indicate the location of some "positive" signals on the autoradiograph. ( ) and side marking were used for keying the filters.

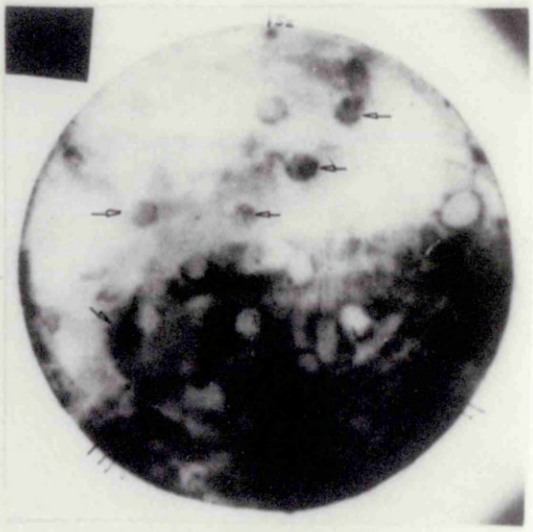
B - Second round of screening. Autoradiograph of a filter prepared from rescreening the "positives" isolated in A. In this case 5 μl of phage suspension of the individual 'positive' (isolated in A) was spotted onto a gridded lawn of HBH 2600 cells. Arrows indicate some of the 'positives'.

C - Third round of screening. Autoradiograph of a filter prepared from a plate containing approximately 100 plaques of the individual 'positive' isolated in B.

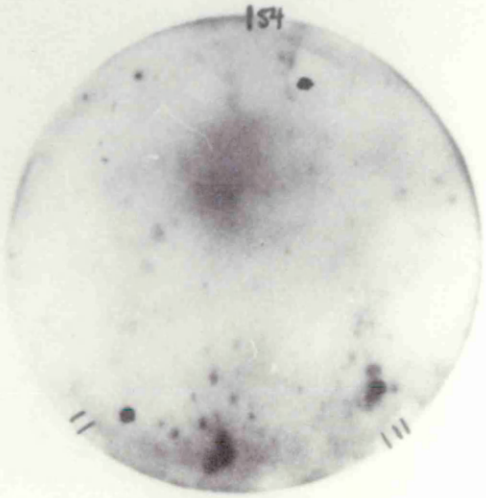
A



B



C



pXt210 was employed for subsequent screening experiments.

2. Isolation of lambda clones

The human foetal liver library was screened by the i*n situ* plaque hybridization procedure of Benton and Davis (1977) with modification described by Maniatis *et al.*, (1978). Approximately $1.2X10^5$ pfu (plaque forming units) were plated on *E. coli* HBH 2600 cells on agar plates. The plaque DNA was lifted by application of nitrocellulose filters. The DNA bound to nitrocellulose filters was denatured, neutralized and fixed to the filters. The filters were then hybridized to ^{32}P-labelled probe in thermally sealed plastic bags. For details of screening procedures and conditions for prehybridization, hybridization and post hybridization see Materials and Methods section 2i. These filters were then autoradiographed. The autoradiographs were examined for 'positive signals' corresponding to plaques on the agar plates. Keyed markings on both the agar plates and nitrocellulose filters facilitated the location of 'positive' plaques (see Figure 4.3A). The plaques from the region of the agar plate corresponding to a 'positive signal' on the autoradiogram were picked up aseptically and phage suspension prepared as described in Materials and Methods section 2j. From this preliminary cycle of screening, 71 putative positives (or areas of positive signal) were selected. Due to the high incidence of 'false' positives and artefacts during screening experiments, the positives selected in the first round of screening were subsequently rescreened by growing 5µl of phage suspension of the individual 'positive' on a lawn of *E.coli* HBH 2600 (gridded plates) (Figure 4.3.B). This spotting of phage suspension was done on duplicate plates to avoid false positives. From this second round of screening nine 'positives' were selected. These were then rescreened at a

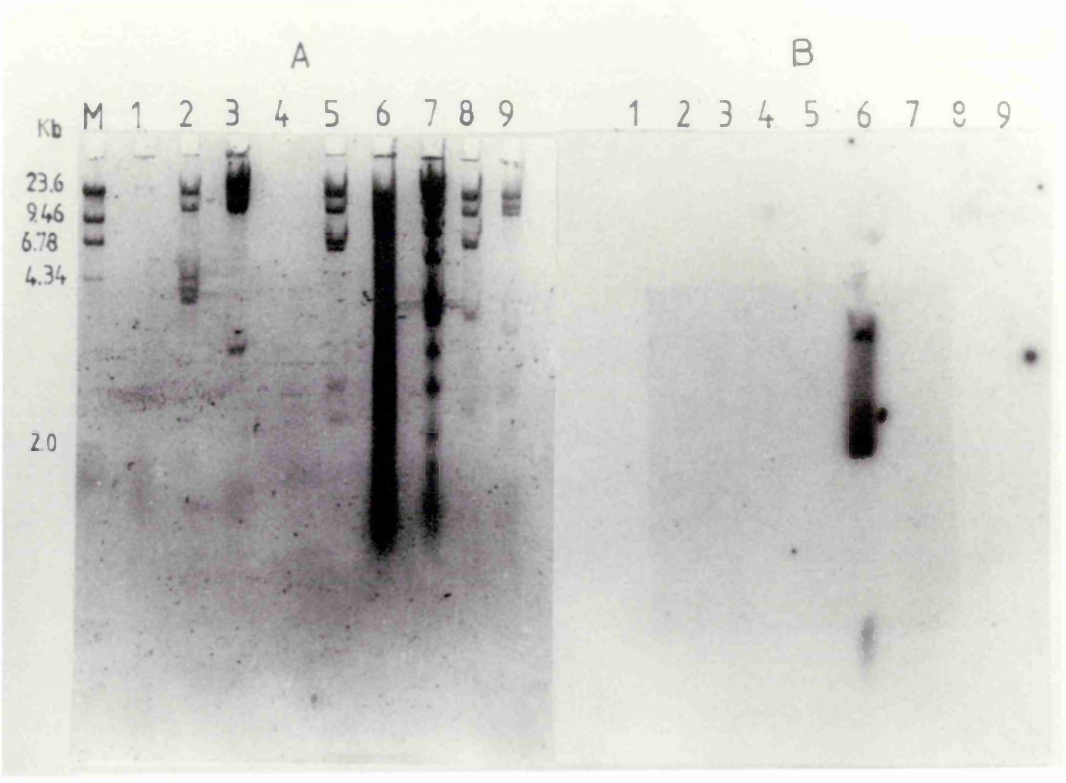
Figure 4.4 EcoRI cleavage pattern of the DNAs from the recombinant clones selected in the screening procedure. DNAs from the respective clones was cleaved with restriction endonuclease EcoRI. The digestion products were fractionated on an 0.8% agarose gel, transferred to a nitrocellulose filter and hybridized to [^{32}P] labelled *Xenopus* tDNA (3.18 Kb insert of pXt210).

A - Ethidium bromide stained gel.

lane M: - size marker, λDNA cleaved with HindIII

lanes 1-9: the different recombinant clones selected in screening procedure.

B - Autoradiograph of the nitrocellulose filter prepared from the gel shown in panel A.


A
B
Kb
M 1 2 3 4 5 6 7 8 9
1 2 3 4 5 6 7 8 9
23.6
9.46
6.78
4.34
2.0

much lower plaque density, to allow the isolation of a single 'positive' plaque (plaque purification) (see Figure 4.3.c). These 'positives' were then amplified in HBH 2600 by the plate lysate procedure (Maniatis *et al.*, 1978), and phage DNA was purified from each of the respective 'positives'. The presence of tRNA sequences in the 'positive' clones selected by screening procedure was confirmed by digesting the phage DNA from each of the respective clone in a separate reaction with restriction endonuclease EcoRI and hybridizing the digestion products to the ^{32}P labelled 3.18 Kb DNA (insert of pXt210). Since the human library was constructed by ligation of human DNA fragments with EcoRI fragments of λCharon 4A (Lawn *et al.*, 1979), cleaving the DNA of recombinant clones with EcoRI should release the insert in several fragments whenever the cloned sequence contained additional EcoRI sites. Figure 4.4A shows the EcoRI cleavage pattern of the respective recombinant clones, and in parallel is shown the resulting autoradiograph of the Southern blot. This analysis showed that only one of the nine clones selected in the screening procedure hybridizes to the labelled *Xenopus* tDNA probe. This recombinant clone was designated λht3 and pursued for further studies.

3. Characterization of λht3

Before the recombinant clone λht3 could be used for further studies it was necessary to demonstrate the λht3 DNA contains sequences that hybridize to human tRNA. It was also useful to know the identity of tRNA species present in λht3. Thus, the approach adopted was first, to hybridize λht3 DNA to [^{32}P] pCp labelled human tRNA ('4S RNA') and second, to hybridize λht3 DNA to the pXt210 subclones, containing each of the individual *X. laevis* t RNA genes.

Figure 4.5 Characterization of λht3 DNA

λht3 DNA was cleaved with the following restriction endonucleases: EcoRI; HindIII; EcoRI + HindIII (double digest). The digestion products were fractionated on an 1% agarose gel, transferred to nitrocellulose filter and hybridized to [^{32}P] pCp labelled human tRNA probe ('4S RNA').

A - Ethidium bromide stained gel.

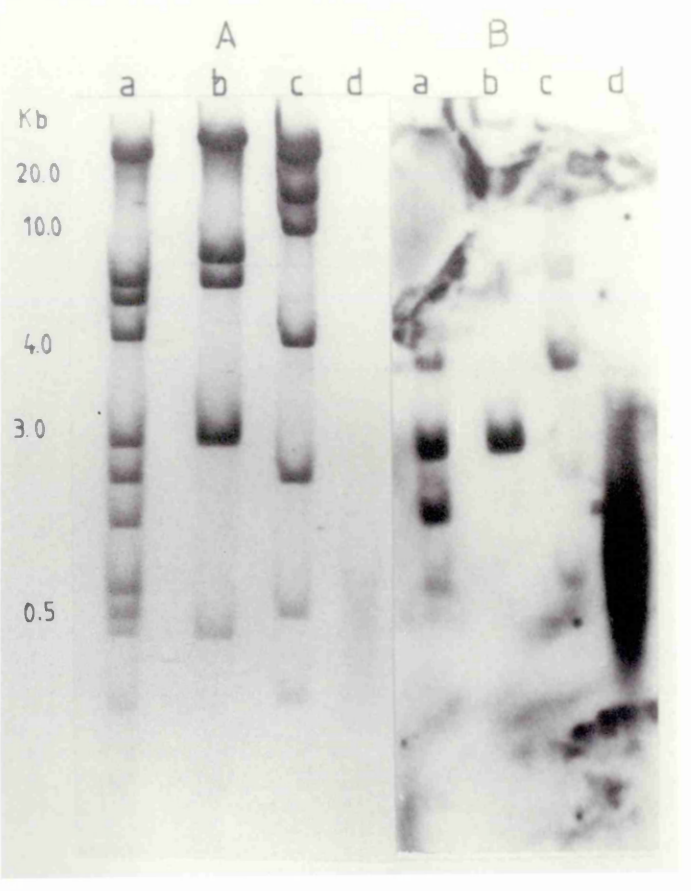
lane a - EcoRI + HindIII double digestion products

lane b - HindIII digestion products

lane c - EcoRI digestion products

lane d - denatured 3.18 Kb *Xenopus* tDNA

B - Autoradiograph of the Southern blot prepared from the gel shown in A.


A
B
a
b
c
d
a
b
c
d
Kb
20.0
10.0
4.0
3.0
0.5

3.a Hybridization of λht3 DNA to pCp labelled human '4S RNA'

To confirm that λht3 contains sequences that hybridize to human tRNA, λht3 DNA was cleaved in separate reactions with restriction endonuclease EcoRI, HindIII, and EcoRI plus HindIII (double digestion). The products were fractionated on a 1% agarose gel, transferred to nitrocellulose filter and hybridized to pCp labelled human tRNA ('4S RNA'). The results of this analysis are shown in Figure 4.5. The pattern emerging from EcoRI cleavage indicates that pCp labelled human tRNA hybridizes to fragments 8.3 and 4 Kb in size. The pattern resulting from HindIII cleavage indicates the possibility of a double band approximately 2.8 Kb in size contains sequences homologous to human tRNA. Whereas EcoRI plus HindIII double digestion products show hybridization to a 2.8 Kb and a 1.9 Kb fragment. Some weak hybridization observed in Figure 4.5B is due to partial digestion products. The result of this double digestion suggests that one of double bands hybridizing in HindIII digestion is retained while the other band is cleaved into approximately 2 Kb and 0.8 Kb fragments. These hybridization results indicate that λht3 DNA contains human tRNA like sequences.

3.b Hybridization of λht3 DNA to the *X. laevis* individual tRNA subclones

In order to determine the identity of the tRNA gene sequence present in λht3, 6μg of λht3 DNA was cleaved with restriction endonucleases

Figure 4.6 Hybridization of [^{32}P] labelled tDNA^{Tyr} subclone of pXt210 to λht3 DNA.

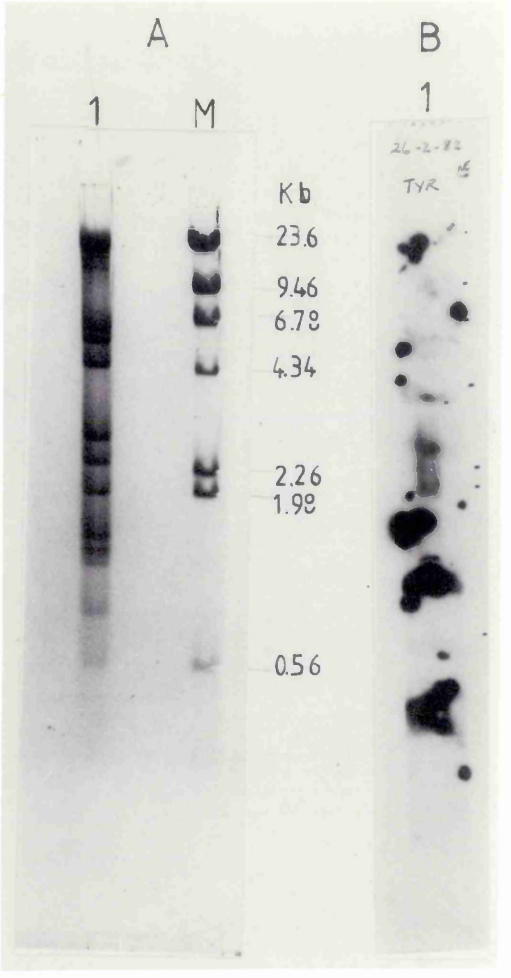
The DNA from λht3 was cleaved with EcoRI + HindIII (double digest). The digestion products were fractionated on 1.5% agarose gel, then transferred to nitrocellulose filter and hybridized to nick-translated pt63 DNA (tDNA^{Tyr} subclone of pXt210).

A - Ethidium bromide stained gel

lane M - size marker, λDNA cleaved with HindIII

lane 1 - EcoRI + HindIII digest of λht3.

B - Autoradiograph of Southern blot prepared from A.


A
B
1
M
1
Kb
23.6
9.46
6.78
4.34
2.26
1.98
0.56
TYR

EcoRI plus HindIII (double digestion), the products fractionated in six different slots on a 1.5% agarose gel, and transferred to a nitrocellulose filter. The resultant nitrocellulose filter was cut into six respective strips. Five of these strips were hybridized to an individual nick-translated subclone derived from pXt210. These individual subclones were a generous gift from Dr. S.G. Clarkson and contain the following tRNA genes: methionine (pt 145); tyrosine (pt 63); lysine (pt 51); asparagine (pt 38) and phenylanine (pt 75) (see Materials and Methods section 2b). The sixth nitrocellulose strip was hybridized to nick-translated pBR322 DNA, in order to rule out that the hybridization observed is not due to pBR322 sequences. The result of this analysis indicated that only the tyrosine containing subclone (pt 63) hybridized to λht3 DNA. Figure 4.6 shows that the two fragments showing positive hybridization to ^{32}P labelled, nick-translated pt 63 DNA are approximately 2.8 Kb and 2 Kb in size. In fact the hybridization pattern is similar to that observed with [^{32}P] pCp labelled human tRNA (Figure 4.5 lane a).

4. Restriction cleavage map of λht3 DNA

A physical restriction map of λht3 was constructed on the basis of single and double digests with restriction endonucleases: EcoRI, HindIII, EcoRI plus HindIII, BamHI, BamHI plus EcoRI, BamHI plus HindIII, XbaI, XbaI plus EcoRI, and XbaI plus HindIII. The resultant cleavage products were fractionated by gel electrophoresis on 1% or 1.5% agarose gels (see Figure 4.7A). The DNA fragments were visualized by staining the gels with ethidium bromide and the sizes of the fragments were determined by comparison with the mobility of standards of known size. A computer program devised by

Table 4.1 Sizes of restriction digest fragments of λht3

	EcoRI	Hind III	EcoRI + HindIII	Bam HI	BamHI + EcoRI	BamHI + HindIII	XbaI	XbaI + EcoRI	XbaI + HindIII
1.	19.8	27.8	19.8	15.5	14.3	15.9	24.5	19.8	23.0
2.	10.9	6.6	5.7	7.5	7.1	7.4	22.4	10.9	6.6
3.	8.3	5.6	5.2	5.6	5.6	5.6	0.7	8.2	5.6
4.	4.2	2.8	4.2	5.5	5.6	5.6		4.2	4.1
5.	2.2	2.7	2.7	4.8	3.9	3.9		2.3	2.8
6.	1.4	1.4	2.3	3.9	3.2	2.8		0.8	2.7
7.	0.8	0.5	1.9	3.2	1.7	2.5		0.8	1.4
8.			1.5	1.5	1.5	1.5		0.6	0.8
9.			1.4	0.5	1.4	1.3			0.5
10.			1.3		0.8	0.7			
11.			0.8		0.6	0.6			
12.			0.5		0.5				
	47.6	47.4	47.3	48.0	46.2	47.8	47.6	47.6	47.5

All sizes are in Kilobases.

The average total size of different restriction products = 47.4

λCharon 4A arms = 10.9 + 19.8 = 30.7 Kb

Therefore, Insert size of λht = 47.4 - 30.7 = 16.7 Kb

Figure 4.7 Cross-hybridization of ^{32}P labelled EcoRI fragments of λht3 DNA to the various other restriction products of λht3. The digestion products of λht3 DNA were fractionated on 1% agarose gel, then transferred to nitrocellulose filter and hybridized to ^{32}P labelled 4.2 Kb EcoRI fragment of λht3.

A - Ethidium bromide stained gel showing the various restriction digestion products of λht3.

lane 1 - size marker, λDNA cleaved with HindIII.

lane 2 - EcoRI, lane 3 - HindIII; lane 4 - EcoRI + HindIII;

lane 5 - Bam HI, lane 6 - Bam HI + EcoRI; lane 7 - Bam HI + HindIII;

lane 8 - XbaI; lane 9 - XbaI + EcoRI; land 10 - XbaI + HindIII; lane 11 - KpnI; digestion products of λht3 DNA.

B - Autoradiograph of Southern blot obtained by hybridization of ^{32}P labelled 4.2 Kb EcoRI fragment of λht3 to A.

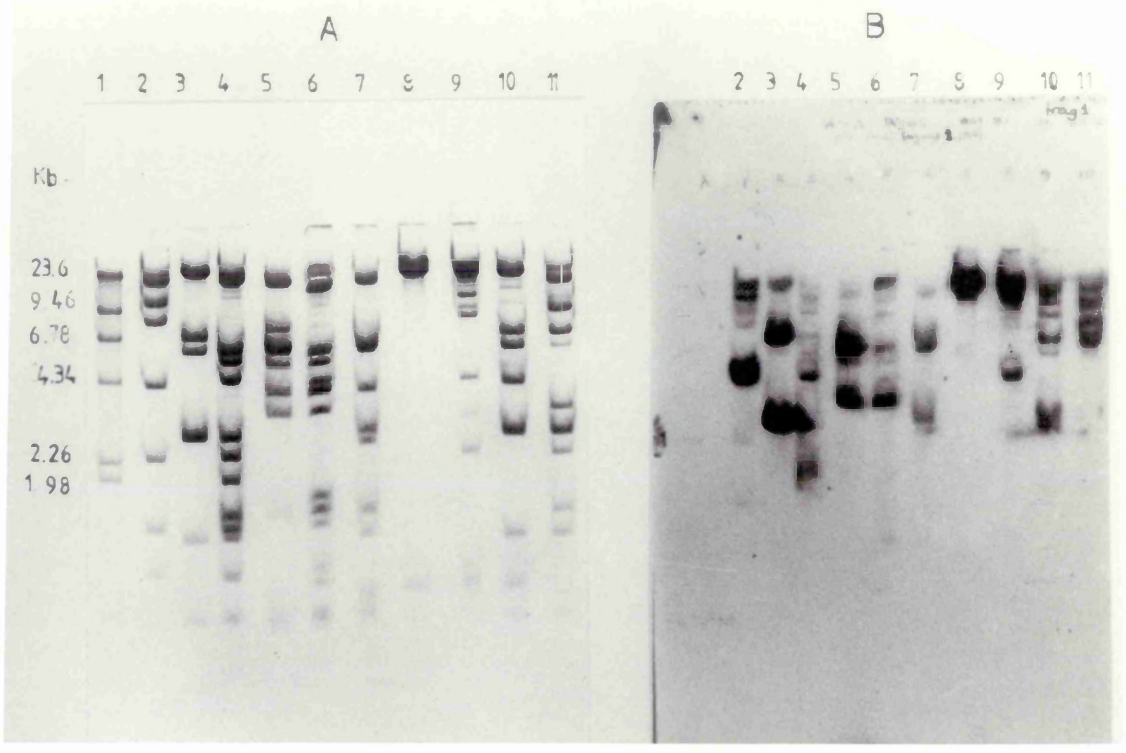

A
B
1 2 3 4 5 6 7 8 9 10 11
2 3 4 5 6 7 8 9 10 11
Kb
23.6
9.46
6.78
4.34
2.26
1.98

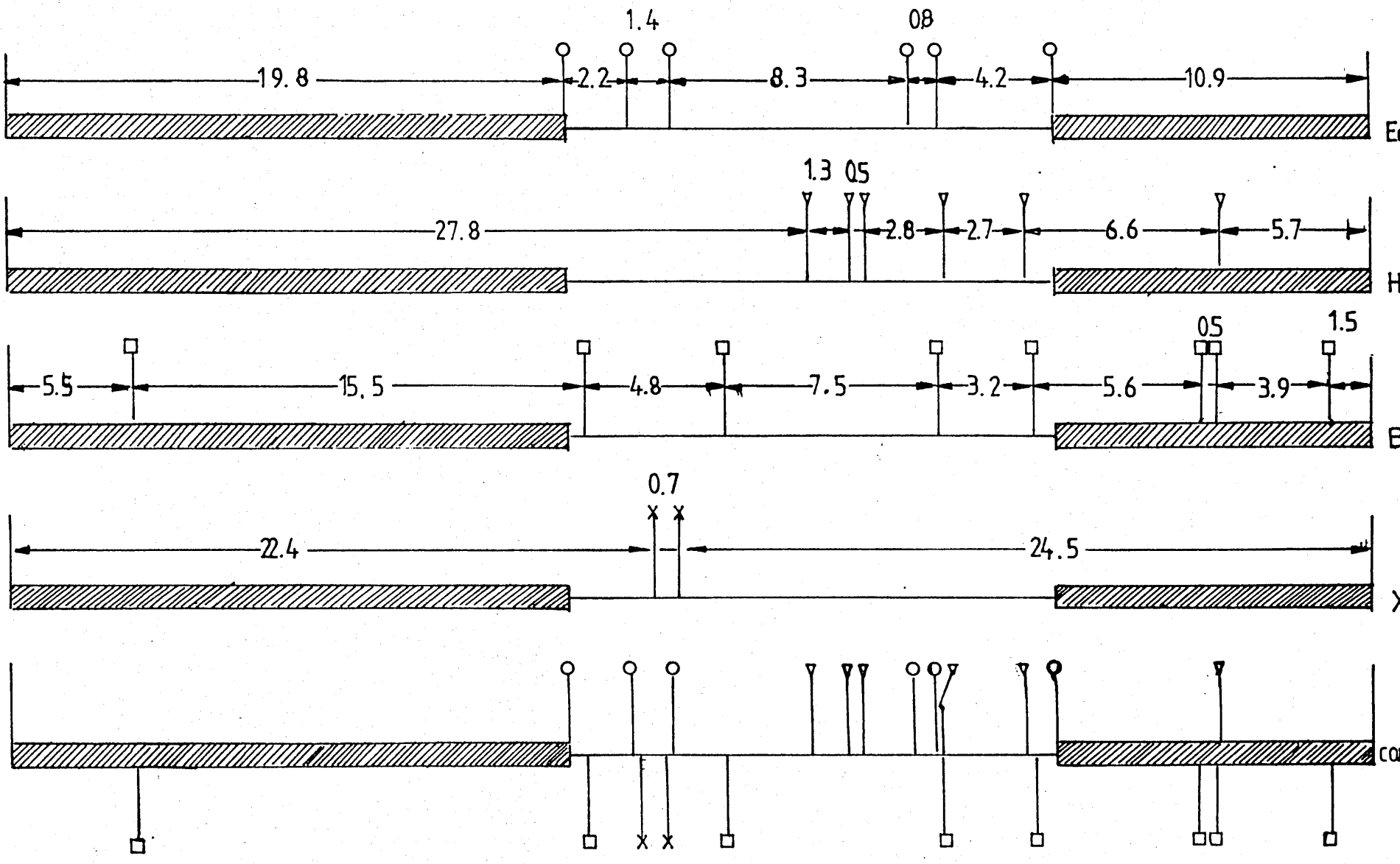
Table 4.2: Cross hybridization of labelled EcoRI restriction fragment of λht3 (4.2; 2.25; 1.4; 0.8 Kb) to other restriction digestion products of λht 3.

Restriction digest	^{32}P-labelled EcoRI restriction fragment of λht3			
	4.2	2.25	1.4	0.8
HindIII	6.6; 2.7	27.8	27.8	2.8
EcoRI + HindIII	4.2; 2.7	2.25	1.4	0.8
Bam HI	5.6; 3.2	15.5; 4.8	4.8	7.5; 2.8
Bam HI + EcoRI	3.2	1.7	1.4	0.8
Bam HI + HindIII	5.6; 2.7	15.8; 4.8	4.8	2.8
Xba I	24.5	22.45	0.76	24.5
XbaI + EcoRI	4.2	2.25	0.76	0.8
XbaI + HindIII	2.7; 6.6	23.0	2.8; 0.76	2.8

Numbers represent fragment size in Kilobases.

Figure 4.8 Restriction map of λht3.

Restriction sites were mapped for several restriction endonucleases by cross-hybridization of several EcoRI fragments (4.1, 2.2, 1.4 and 0.8 Kb) of the insert of λht3 and by reference to known restriction sites in the λCharon 4A arms (Blattner *et al.*, 1977).


1.4
0.8
19.8
2.2
8.3
4.2
10.9
1.3
0.5
27.8
2.8
2.7
6.6
5.7
0.5
1.5
5.5
15.5
4.8
7.5
3.2
5.6
3.9
0.7
22.4
24.5

Duggleby et al., (1981) which fits a parabola to a set of standards was used to calculate the sizes of restriction cleavage products from their mobility. The results of this analysis are summarized in Table 4.i.

The size of the inserted human DNA in λht3 was calculated from the sum of restriction cleavage fragments after subtracting the known size of the λ Charon 4A arms (Table 4.i). Such calculations showed that λht3 has an insert of human DNA of approximately 16.9 Kb. In general, λ Charon 4A has a cloning capacity of 12-20 Kb (Maniatis *et al.*, 1978).

Construction of the restriction map of λht3 and alignment of restriction fragments was facilitated by the fact that fragments adjacent to the λCharon 4A arms could often be identified by virtue of their vector DNA components (Blattner *et al.*, 1977) (Figure 4.8). Other restriction fragments of λht3 were located by cross-hybridization experiments. DNA from λht3 was cleaved with EcoRI, and the resultant restriction fragments (4.1; 2.3; 1.4 and 0.8 Kb in size) were isolated from the agarose gel by the procedure of Finkelstein and Rownd, (1978). Each of these isolated fragments was nick-translated and hybridized to the various restriction cleavage products of λht3 transferred onto nitrocellulose filters, as illustrated in Figure 4.7. The results of this analysis are summarized in Table 4.2. Figure 4.8 shows the construction of the physical restriction map of λht3.

5. Subcloning of tRNA sequence containing fragments of λht3

The principal objective of subcloning was reduction in the physical

and genetic complexity of λht3, by elimination of all genetic information that does not pertain to the biological property under investigation. The strategy adopted for subcloning was to isolate the tRNA sequence - containing HindIII fragments of λht3 (that is, the 2.7 and 2.8 Kb fragments, co-migrating as a doublet (see Figure 4.5 lane b)) and covalently ligate them to the unique HindIII restriction site of the vector, plasmid pAT153 (Twigg and Sherrat, 1980). The HindIII restriction site in pAT153 lies in the promoter involved in expression of tetracycline resistance and subcloning at this site inactivates the expression of tetracycline resistance.

Consequently, 30 μg of λht3 DNA was cleaved with restriction endonuclease HindIII, the products fractionated on 1% agarose gel and the 2.8 Kb double band was excised from the gel. The recovered DNA was purified as described in Materials and Methods section 2.o.(ii). Similarly pAT153 DNA was restricted with endonuclease HindIII and treated with bacterial alkaline phosphate to minimise recirculization during the ligation reaction. The two fragments were ligated *in vitro* using molar excess of '2.8 Kb' DNA (as described in Materials and Methods section 2p). The ligation mixture was used to transform *E. coli* HB 101 bacterial cells. The transformed cells were plated on agar plates containing ampicillin or tetracycline. Most of the transformed colonies appeared on ampicillin containing plates. The yield of ampicillin resistant transformants was approximately twenty fold higher than that of tetracycline resistant transformants, thus suggesting that the inserted DNA had inactivated expression of tetracycline resistance. To confirm this result thirty-five ampicillin resistant colonies were picked at random and innoculated on ampicillin - or tetracycline - containing gridded plates. As expected, most colonies grew on the ampicillin - containing plates. Five

Figure 4.9 Restriction cleavage analysis of pNB subclones.

DNA prepared from single colonies of the respective subclones was cleaved with restriction endonuclease EcoRI plus HindIII. The digestion products were fractionated on 1.5% agarose gel and visualized by ethidium bromide staining.

lane 1 - size marker, λDNA cleaved with HindIII

lane 2 - pAT 153 DNA cleaved with HindIII
(shown at two different DNA concentrations)

lane 3 - pNB23 cleaved with EcoRI + HindIII

lane 4 - pNB21 cleaved with EcoRI + HindIII

lane 5 - pNB4 cleaved with EcoRI + HindIII

lane 6 - pNB1 cleaved with EcoRI + HindIII

lane 7 - pNB51 cleaved with EcoRI + HindIII

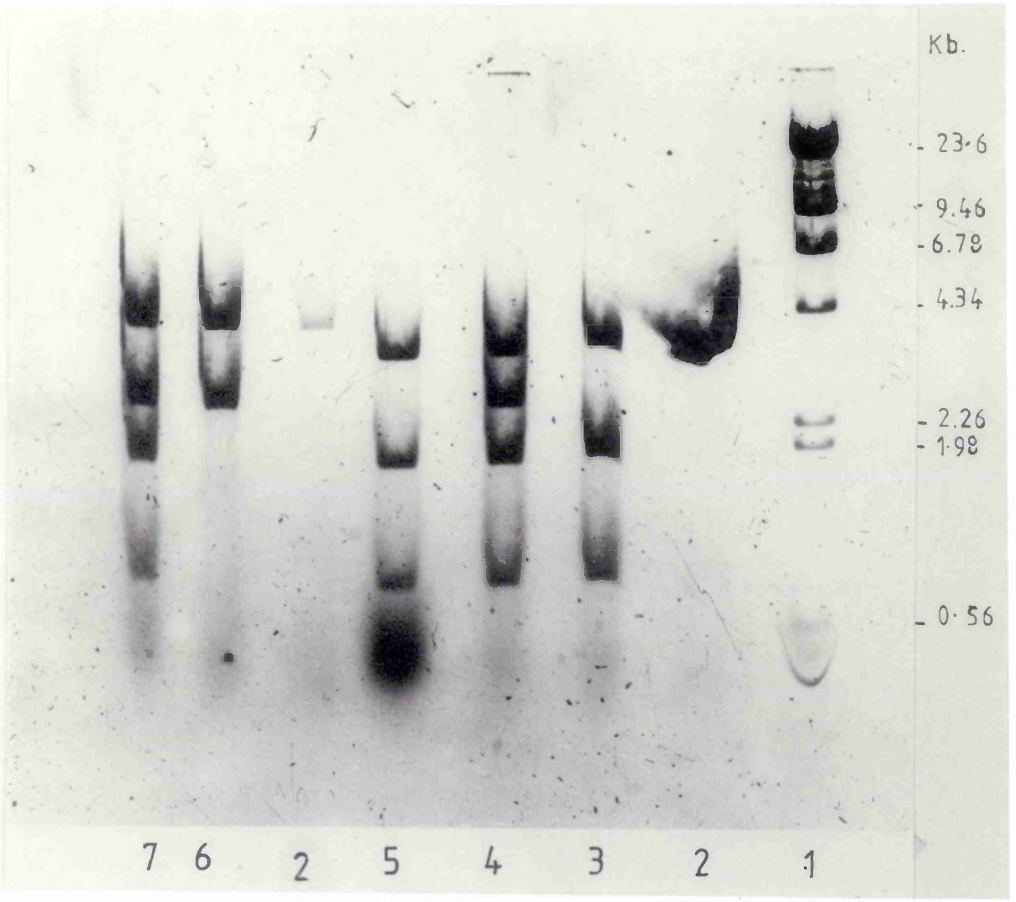
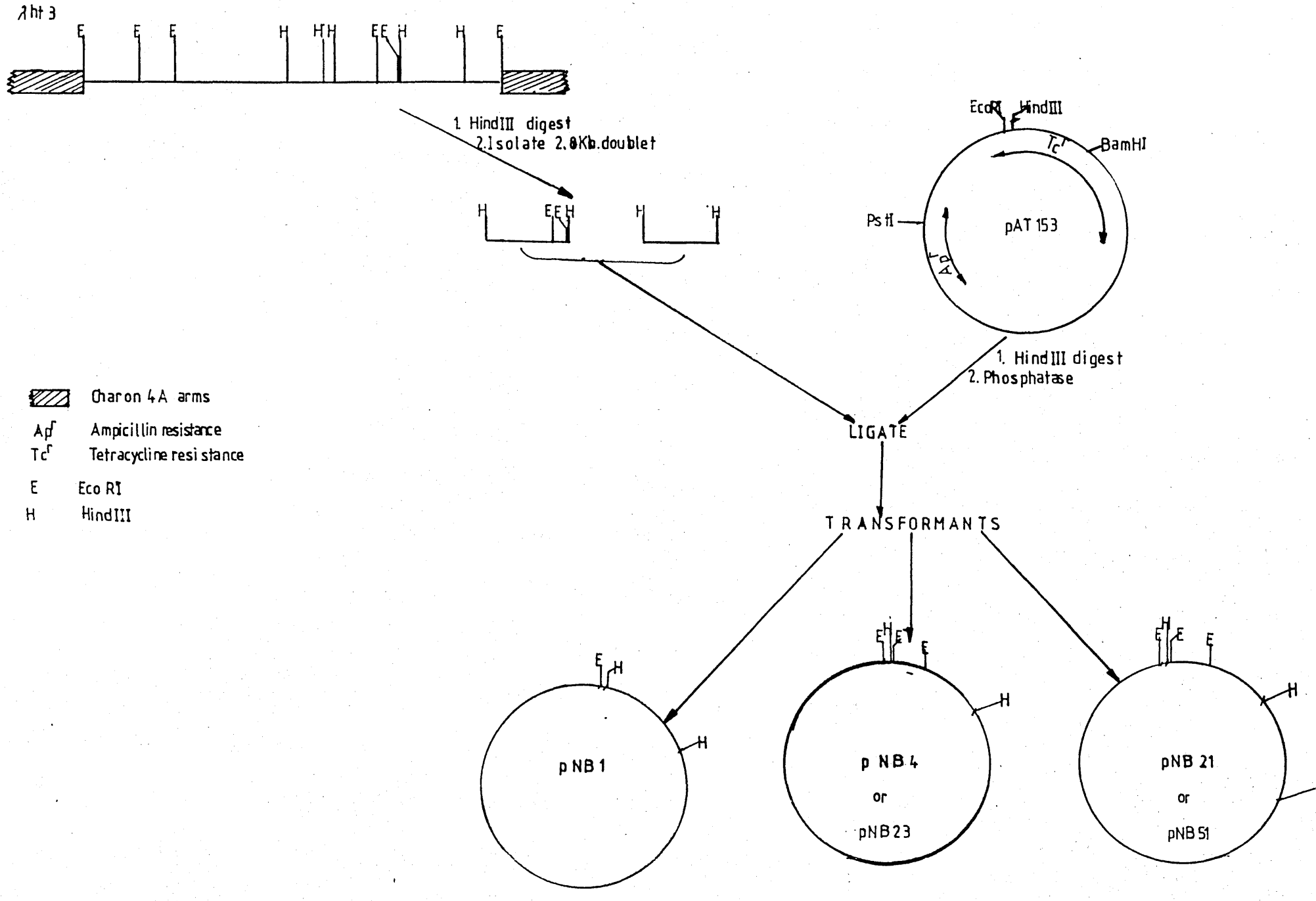

Kb.
23·6
9.46
6.78
4.34
2.26
1·98
0·56
7
6
2
5
4
3
2
1

Figure 4.10 Summarized construction of the subclones.

5μg of the 2.8 Kb HindIII fragments ("doublet") of λht3 was isolated and ligated with 0.5μg of HindIII cleaved pAT153 DNA. Three different types of ampicillin resistant, tetracycline sensitive subclones were identified: 1) pNB1 - contains a 2.7 Kb insert of λht3; 2) pNB4 - contains a 2.8 Kb insert of λht3 and is cleaved by EcoRI; 3) pNB21 - contains head to tail 2.7 Kb and 2.8 Kb insert of λht3.


λht 3
1. HindIII digest
2. Isolate 2.0Kb. doublet
EcoRI
HindIII
BamHI
PstI
pAT 153
1. HindIII digest
2. Phosphatase
LIGATE
TRANSFORMANTS
pNB 1
pNB 4
or
pNB 23
pNB 21
or
pNB 51
Charon 4A arms
Ampicillin resistance
Tetracycline resistance
Eco RI
HindIII

such colonies, which were tetracycline sensitive and ampicillin resistant, were selected and plasmid DNA was prepared from single colonies by the rapid method of Holmes and Quigley (1981). These were designated pNB1, pNB4, pNB21, pNB23 and pNB51. The DNA from each preparation was cleaved with restriction endonucleases EcoRI plus HindIII, and products fractionated on 1.5% Agarose gel. The results of this analysis are shown in Figure 4.9. The sizes of the restriction fragments were computed from the mobility of the marker DNA fragments. These results showed that all the subclones contain vector DNA migrating at 4.14 Kb. The plasmids pNB4 and pNB23 showed a similar restriction pattern (Figure 4.9 lane 3 and lane 5) both contain a 1.9 Kb and 0.9 Kb fragments, whereas pNB1 contains a 2.7 Kb fragment. The plasmids pNB21 and pNB51 also showed a similar restriction pattern (Figure 4 lane 4 and lane 8). Both contain the 2.7 Kb, 1.9 Kb and 0.9 Kb restriction fragments. These results are explained diagramatically in Figure 4.10. The plasmids pNB1, pNB4 and pNB21 were studied further.

6. Restriction cleavage patterns of pNB1 and pNB4

Initial information on the approximate number of different restriction cleavage sites present in the human DNA insert of recombinant plasmids pNB1 and pNB4 determined by subjecting these DNAs to various restriction endonucleases (AvaI; AvaII; AluI; HaeIII; HhaI; HinfI; HpaI; HpaII; PstI; TaqI; SmaI; SalI and XhoI), commonly used in this laboratory.

The 2.8 Kb and 2.7 Kb human DNA inserts of plasmids pNB4 and pNB1, were excised from the respective recombinant parent plasmid with restriction endonuclease HindIII, and recovered electrophoretically as described

Figure 4.11 Restriction cleavage pattern of pNB1 and pNB4.

Approximately 1μg of the human DNA insert of pNB1 and pNB4 DNA was digested with 1 unit of restriction endonuclease. The digestion products were electrophoresed on 1.5% agarose gel. Different marker DNAs were also included.

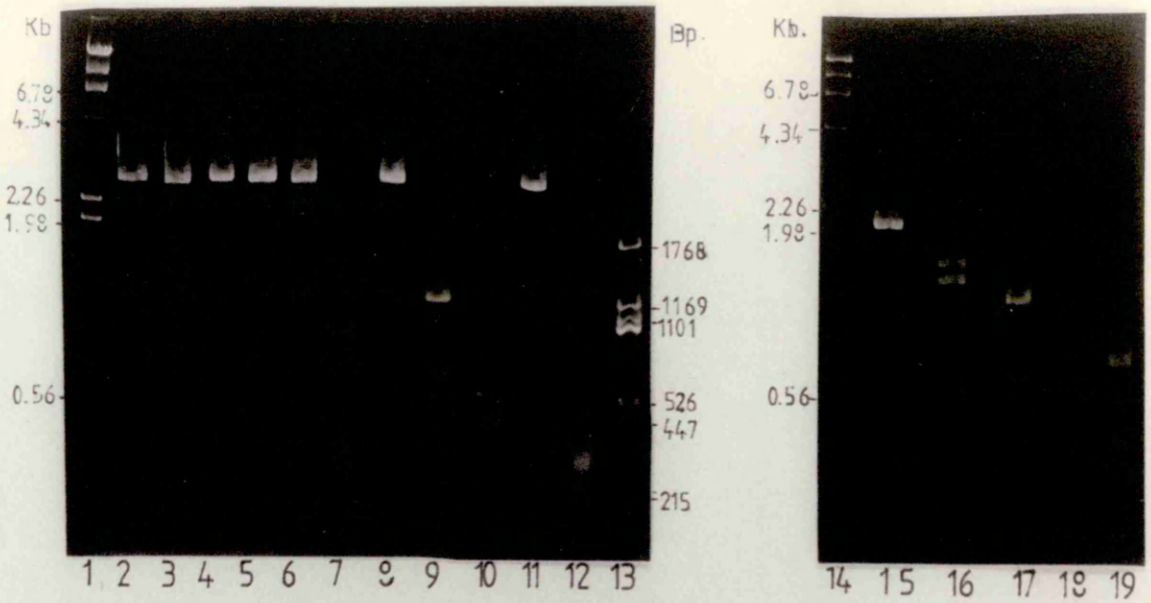
A - Digestion products of pNB1 insert

lane 1 and 14, marker λDNA HindIII cleaved; lane 1 TaqI;
lane 3 - SmaI; lane 4 - XhoI; lane 5 - SalI; lane 6 - PstI;
lane 7 - AluI; lane 8 - HpaII; lane 9 - HhaI; lane 10 - HinfI;
lane 11 - HpaI; lane 12 - Hae III; lane 13 - marker SV40 DNA
cleaved with HindIII; lane 15 - SmaI; lane 16 - PstI; lane 17 - HhaI;
lane 18 - AvaI, lane 19 - AvaII.

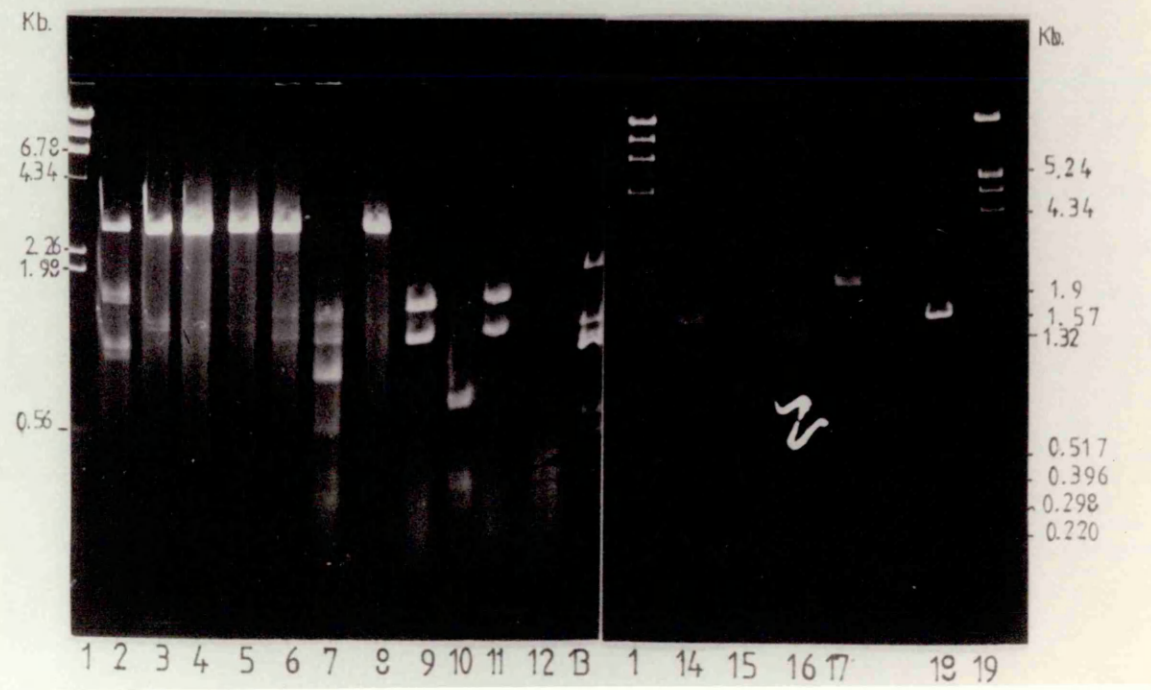
B - Digestion products of pNB4 insert

lane 1 - marker λDNA HindIII cleaved; lane 2 - Taq I;
lane 3 - Sma I; lane 4 - XhoI; lane 5 - SalI; lane 6 - PstI;
lane 7 - AluI; lane 8 - HpaII; lane 9 - HhaI; lane 10 - HinfI;
lane 11 - HhaI; lane 12 - HaeIII; lane 13 - marker SV40 DNA cleaved
with HindIII; lane 14 - PstI; lane 15 - SmaI; lane 16 - BglI; lane 17 -
AvaI; lane 18 - marker pAT153 DNA cleaved with HinfI, lane 19 -
marker λDNA cleaved with EcoRI + HindIII.

A


Kb
6.78
4.34
2.26
1.98
0.56
Bp
1768
1169
1101
526
447
215
1 2 3 4 5 6 7 8 9 10 11 12 13
Kb.
6.78
4.34
2.26
1.98
0.56
14 15 16 17 18 19

B


Kb.
6.78
4.34
2.26
1.98
0.56
1 2 3 4 5 6 7 8 9 10 11 12 13
Kb.
5.24
4.34
1.9
1.57
1.32
0.517
0.396
0.298
0.220
1 14 15 16 17 18 19

in Materials and Methods section 2.o.(i). The purified DNA was digested with different restriction endonucleases and the products fractionated on 1.5% agarose gel. The results of this restriction endonuclease cleavage are shown in Figure 4.11. The human DNA inserts appear to exhibit multiple restriction cleavage sites for HaeIII, HinfI and AluI, a few SmaI; PstI; HhaI; AvaI and AvaII sites and no cleavage sites for XhoI; SalI and HpaII.

6.a Fine structure restriction map of pNB1 and pNB4

A fine structure restriction map of subclones pNB1 and pNB4 was determined by the method of Smith and Birnstiel (1976). This method is based on DNA fragments labelled at 5' termini with a [^{32}P]-phosphoryl group. The labelled fragment is then cleaved asymmetrically with a suitable restriction enzyme into two fragments that are separable by gel electrophoresis. Each DNA segment, labelled at only one end, is recovered and digested with a specific (chosen) restriction enzyme so as to produce a spectrum of partial digestion production all with a common labelled terminus, representing a cut at each of the restriction site in the molecule. Gel electrophoresis followed by autoradiography reveals the length of each labelled fragment and hence the restriction site can be mapped.

The purified human DNA inserts of plasmids pNB1 and pNB4 were treated with bacterial alkaline phosphatase, labelled at 5' end with [γ - ^{32}P] ATP using T_4 polynucleotide kinase and cleaved into two assymmetric fragments using a restriction endonuclease that cleaves the respective fragment only once. The 2.8 Kb insert of pNB4 was cleaved with EcoRI into 1.9 Kb and 0.9 Kb DNA fragments, and the 2.7 Kb insert of pNB1 was cleaved with SmaI into 2.1 Kb and 0.6 Kb DNA fragments. These fragments were isolated and subjected to further analysis except for the 0.9 Kb EcoRI fragment of

Figure 4.12 Partial restriction products of the 2.1 Kb HindIII - SmaI fragment of pNB 1 using restriction endonuclease PstI and AvaII.

The restriction fragment labelled at only one end with 5' - [γ-^{32}P] ATP was cleaved as follows: The reaction mixture (60µl) containing 1000 cpm of the labelled fragment, 6µl of the appropriate 10X restriction endonuclease buffer and 1 unit of restriction endonuclease was incubated at 37°C. Aliquots of 10µl were removed at 0.5; 1; 4; 10 and 20 minutes into 2.5ul of stopper solution, heated at 65°C for 5 minutes prior to loading onto a 1.5% agarose gel, along with labelled standard marker DNA (λDNA cleaved with EcoRI + HindIII) and electrophoresed at 120 V. The gels were dried and autoradiographed.

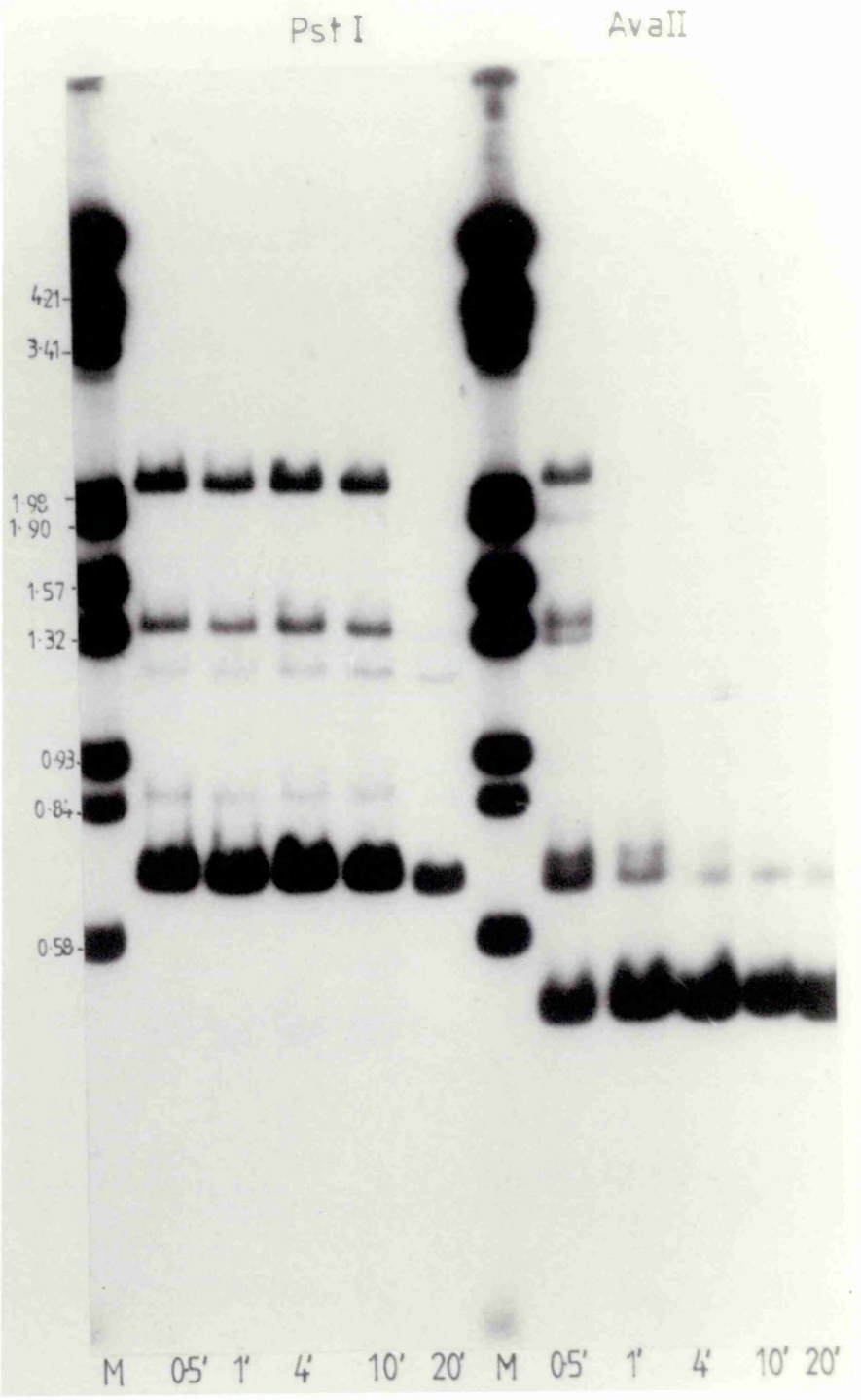

Pst I
AvaII
4·21
3·41
1·98
1·90
1·57
1·32
0·93
0·84
0·58
M
0·5'
1'
4'
10'
20'
M
0·5'
1'
4'
10'
20'

Figure 4.13 Master compilation gel of the partial restriction digestion products of the 1.9 Kb HindIII - EcoRI fragment of pNB4.

lane 1 - marker pAT 153 DNA cleaved with HinfI;

lane 2 - marker, λ DNA cleaved with EcoRI + HindIII;

lane 3 - PstI; lane 4 - HhaI; lane 5 - Bg1I;

lane 6 - AvaI; lane 7 - SmaI; lane 8 - Alu I;

lane 9 - Hinf I; Lane 10 - Hae III

lane 11 - Ava II; lane 12 - undigested 1.9 Kb DNA

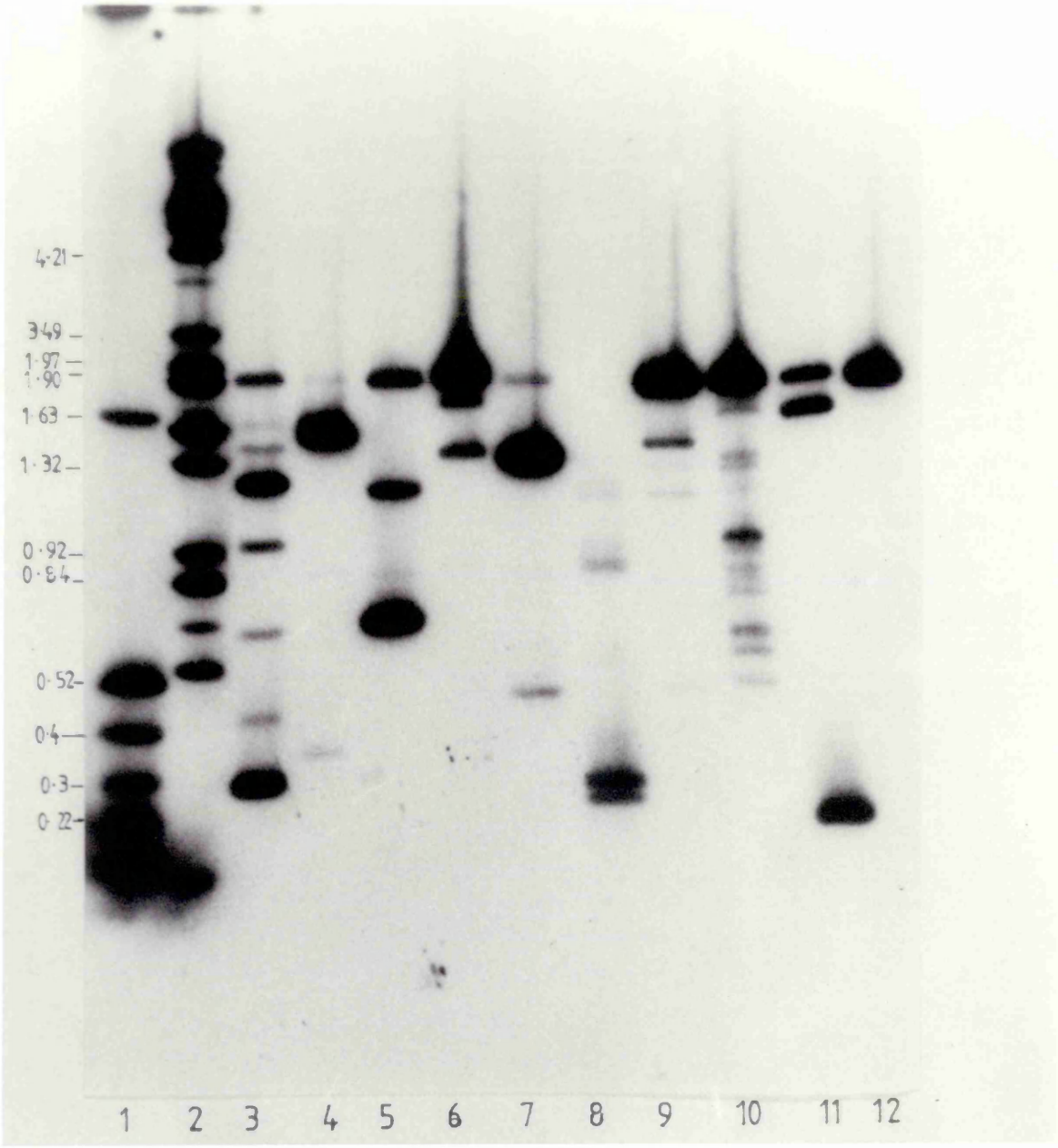

4·21
3·49
1·97
1·90
1·63
1·32
0·92
0·84
0·52
0·4
0·3
0·22
1 2 3 4 5 6 7 8 9 10 11 12

Figure 4.14 Complete restriction cleavage products of pNB1 and pNB4 DNA. The isolated 1.9 Kb EcoRI - HindIII fragment of pBN4, the 2.1 Kb and the 0.6 Kb SmaI - HindIII fragments of pNB1 were cleaved with various restriction endonucleases in separate reactions. The digestion products were fractionated on 1.75% Agarose gel or 6% polyacrylamide gels and visualized by ethidium bromide staining.

A - Digestion products of 1.9 Kb EcoRI - Hind III fragment of pNB4 (1.75% agarose gel).

lane 1 - marker, pAT153 DNA cleaved with HaeIII;

lane 2 - undigested 1.9 Kb DNA; land 3 - HhaI;

lane 4 - PstI; lane 5 - SmaI; lane 6 - AvaI;

lane 7 - Ava II; lane 8 - marker pAT 153 DNA cleaved with HinfI.

B - Digestion of 2.1 Kb SmaI - HindIII fragment of pNB 1 (1.75% agarose gel).

lane 1 - marker, λDNA cleaved with EcoRI + HindIII;

lane 2 - undigested 2.1 Kb DNA; lane 3 - HhaI; lane 4 - PstI;

lane 5 - Ava II; lane 6 - Ava I; lane 7 - marker, pAT 153 DNA cleaved with Hae III.

C - Digestion products of 0.6 Kb SmaI - HindIII fragment of pNB1 (6% polyacrylamide gel).

lane 1 - marker pAT 153 DNA cleaved with Hae III; lane 2 - 0.6 Kb DNA; lane 3 - Pst I; lane 4 - Hae III; lane 5 - Ava II; lane 6 - Ava I, lane 7 - marker, pAT 153 DNA cleaved with Hinf I.

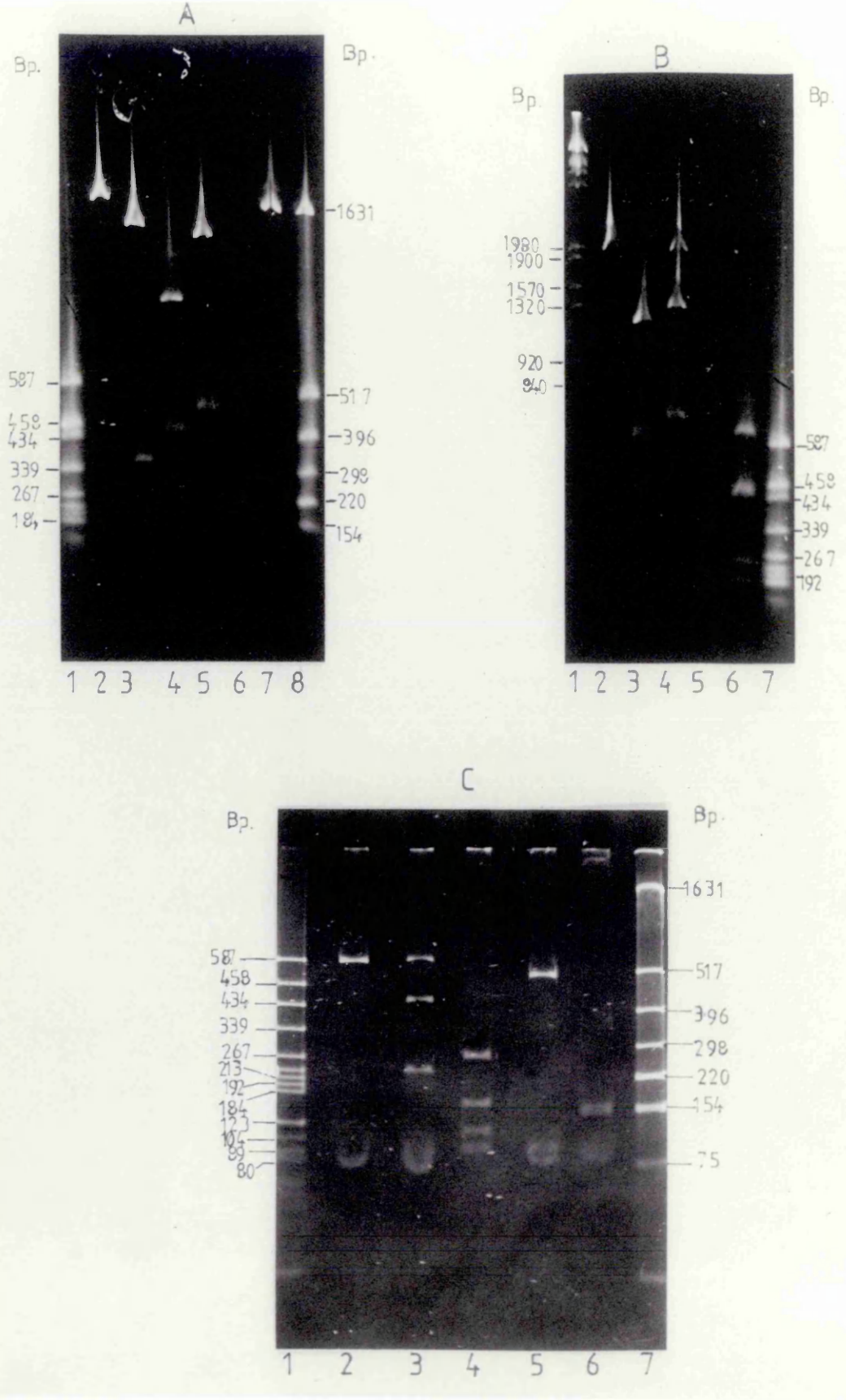
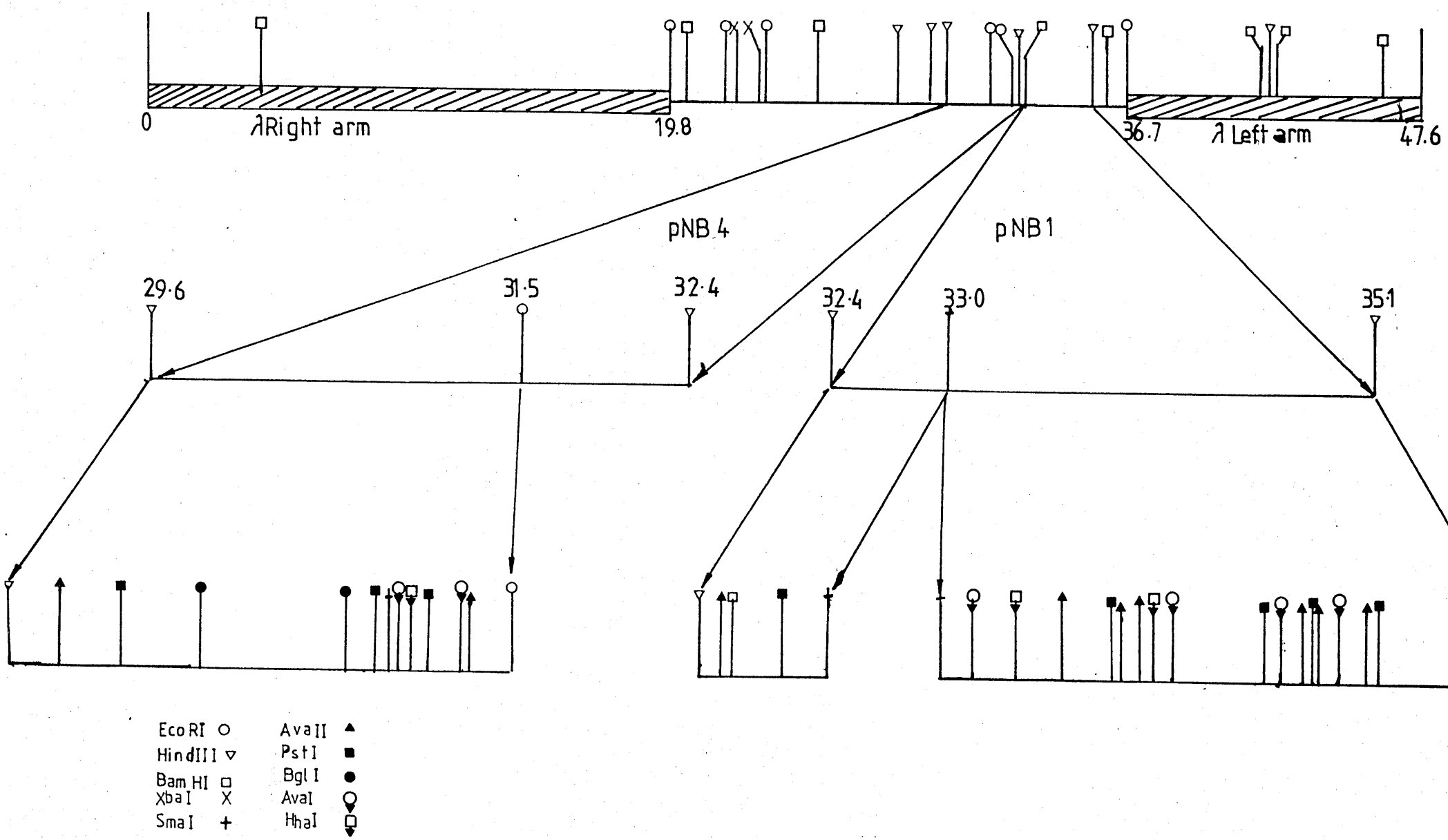

A
Bp.
Bp.
1631
587
517
458
434
396
339
298
267
220
192
154
1 2 3 4 5 6 7 8
B
Bp.
Bp.
1980
1900
1570
1320
920
840
587
458
434
339
267
192
1 2 3 4 5 6 7
C
Bp.
Bp.
1631
587
517
458
434
396
339
298
267
213
220
192
184
154
104
99
80
75
1 2 3 4 5 6 7

Figure 4.15 A fine structure restriction map of pNB 1 and pNB 4.


0
λRight arm
19.8
36.7
λ Left arm
47.6
pNB 4
pNB 1
29·6
31·5
32·4
32·4
33·0
35·1
EcoRI
HindIII
Bam HI
XbaI
SmaI
AvaII
PstI
BglI
AvaI
HhaI

pNB4, since earlier hybridization results showed that it does not hybridize to human tRNA or the *X. laevis* 3.18 Kb probe (see Results section 3).

For restriction site mapping analysis the fragments labelled at only one of its termini were cleaved with the following restriction enzymes: PstI; HhaI; AvaI; AvaII; SmaI; HinfI and HaeIII, samples were removed for electrophoretic analysis at 0.5; 1; 2; 4; 10; 20; and 40 minutes. The results of such a time course are shown in Figure 4.12. In most cases the 0.5 and 1minute samples contained a large proportion of undigested DNA, whereas at 20 and 40 minutes, the reaction was approaching completion. The optimal time point for the various restriction endonucleases were simultaneously compared on a 'master compilation gel' as illustrated in Figure 4.13. The size of each partial fragment was determined by computing the relative mobility of the partial products and comparing them with markers of known size.

The results of the partial restriction mapping were confirmed by complete digestion analysis of the respective fragment (Figure 4.14). The cleavage sites for the differnt restriction endonucleases present in the human DNA insert of recombinant plasmids pNB1 and pNB4 are summarized in Figure 4.15.

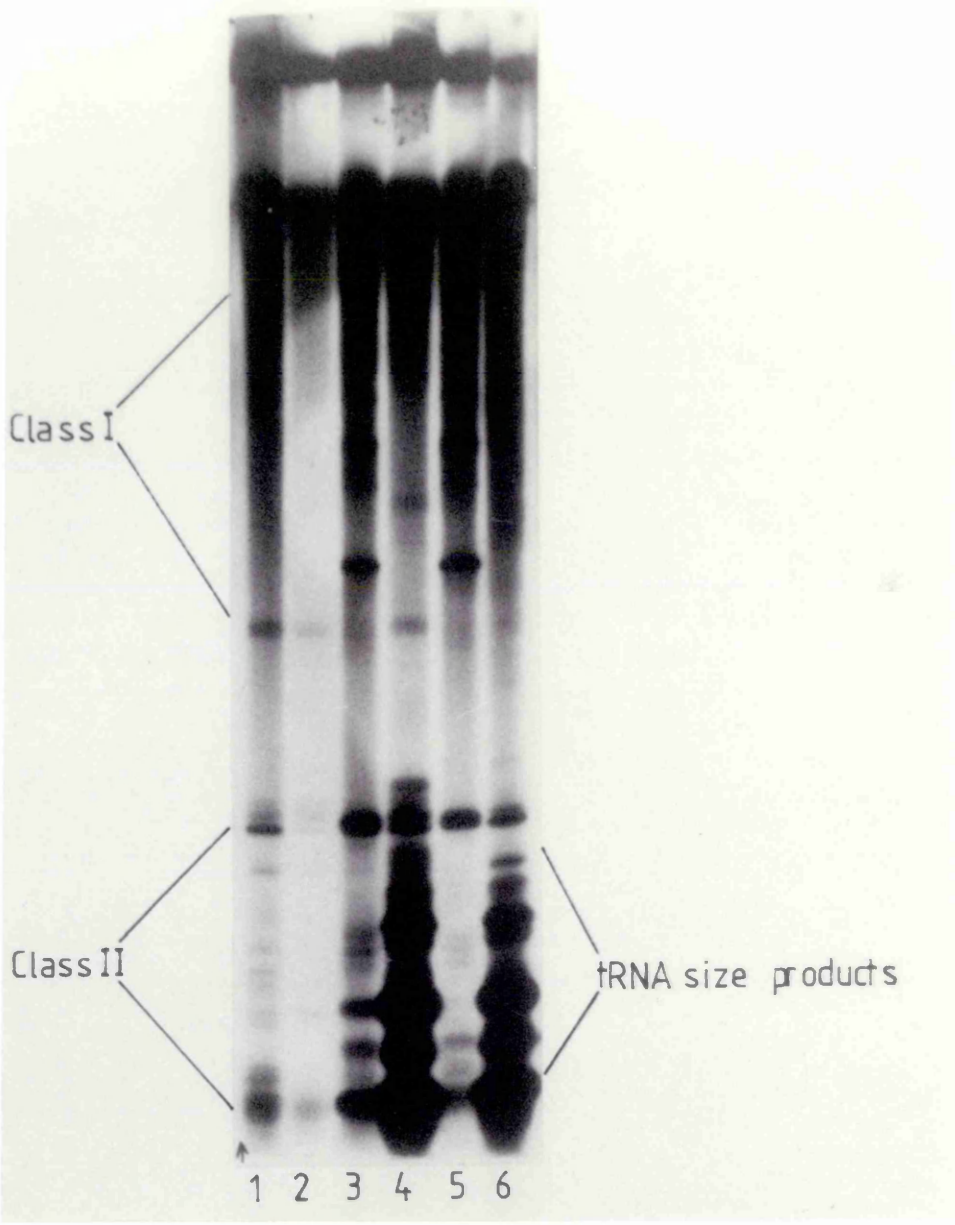
7. Transcription studies of pNB1, pNB4 and pNB21

In order to study the transcription of the subclones, purified DNAs of pNB1, pNB4 and pNB21 were microinjected into the nuclei of *Xenopus* oocytes (This experiment was conducted at M.R.C. laboratory of Molecular Biology, Cambridge, in collaboration with Dr. J.P. Goddard). The RNAs produced in the *Xenopus* oocytes following nuclear injection were resolved by polyacrylamide gel electrophoresis as shown in Figure 4.16. The results of this

Figure 4.16 Electrophoretic analysis of RNAs produced by *Xenopus* oocytes after microinjection of recombinant DNA into nuclei. Purified DNAs of pNB1, pNB4, pNB21 and pXt 210 (50 nanoliters at 100µg/ml) was mixed with [α–^{32}P] GTP and microinjected into stage V and VI Xenopus oocytes. After incubation for 5 hr. at 2^{o}C the RNA was extracted and analysed on 8% polyacrylamide gel under denaturing conditions.

lane 1 - pNB4; lane 2 - pNB1; lane 3 and 5 pNB21

lane 4 and 6 - pXt 210.


Class I
Class II
tRNA size products
1 2 3 4 5 6

analysis indicate that two classes of RNA species are synthesized in *Xenopus* oocytes. These RNAs are indicated in Fig. 4.16 as Class I and Class II. Class I represents RNAs larger than tRNAs. Class II represents tRNA size products. In the case of subclones pNB21 (lanes 3 and 5) and pNB4 (lane 1) tRNA size products along with other class I size RNAs, are synthesized in Xenopus oocytes, whereas in the case of pNB1 (lane 2) there is little evidence of synthesis of any RNA products. It is very difficult to comment on the size of class I RNAs, since the sizes of the RNA products was estimated by the RNAs synthesized by pXt210 DNA (a subclone containing *X. laevis* tRNA genes).

At a later stage in this study, pNB1, pNB4 and pNB21 DNAs were tested in HeLa cell extracts, containing RNA polymerase III activity (S100 extract, prepared according to Weil *et al.*, 1979); and an *in vitro* transcription extract prepared according to the protocol of Manley *et al.*, 1980 (a gift from Dr. S. Gilmour), for the ability to direct synthesis of tRNA size products. However, these analysis revealed that none of the subclones were active in the *in vitro* transcription systems, although pXt210 DNA used as control showed synthesis of tRNA size products.

8. Analysis of RNA complementary to pNB21 DNA

The results presented in this section were obtained in collaboration with Dr. J.P. Goddard.

In order to analyse the RNA forming hybrids with the subclone pNB21, purified pNB21 DNA was bound to nitrocellulose filter discs (see Materials and Methods section 2.s.), and hybridized to labelled RNA probes. The RNAs forming hybrids with the recombinant DNA were released by denaturation of these hybrids and analysed by polyacrylamide gel

Figure 4.17 Schematic diagram showing the principle of periodate-treatment of crude 4S RNA.

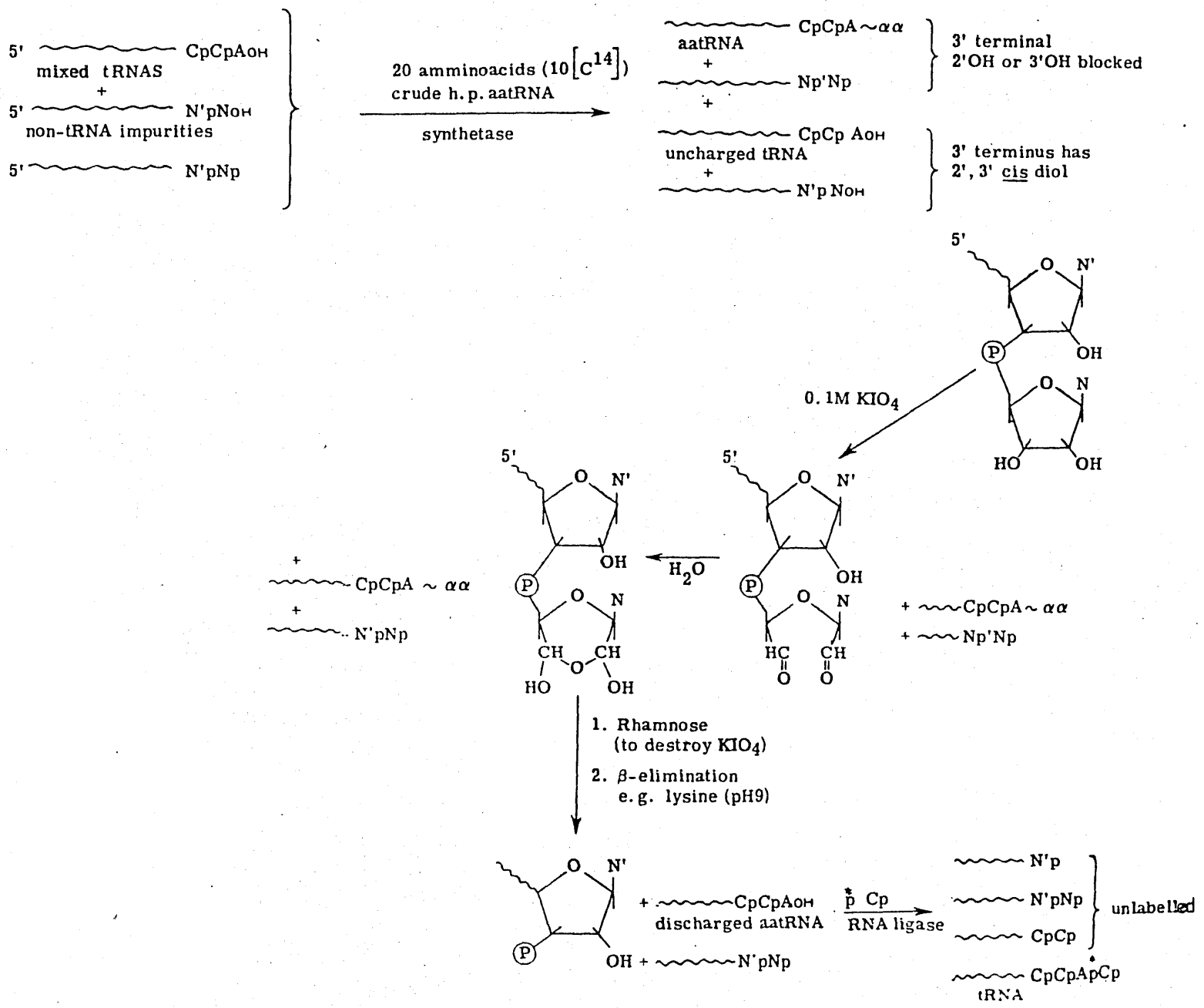

5' CpCpAOH
mixed tRNAS
+
5' N'pNOH
non-tRNA impurities
5' N'pNp
20 amminoacids (10[C14])
crude h.p. aatRNA
synthetase
CpCpA~αα
aatRNA
+
Np'Np
+
CpCp AOH
uncharged tRNA
+
N'p NOH
3' terminal
2'OH or 3'OH blocked
3' terminus has
2', 3' cis diol
0.1M KIO4
H2O
+ CpCpA ~ αα
+ N'pNp
+ CpCpA~αα
+ Np'Np
1. Rhamnose
(to destroy KIO4)
2. β-elimination
e.g. lysine (pH9)
+ CpCpAOH
discharged aatRNA
OH + N'pNp
p Cp
RNA ligase
N'p
N'pNp
CpCp
CpCpApCp
tRNA
unlabelled

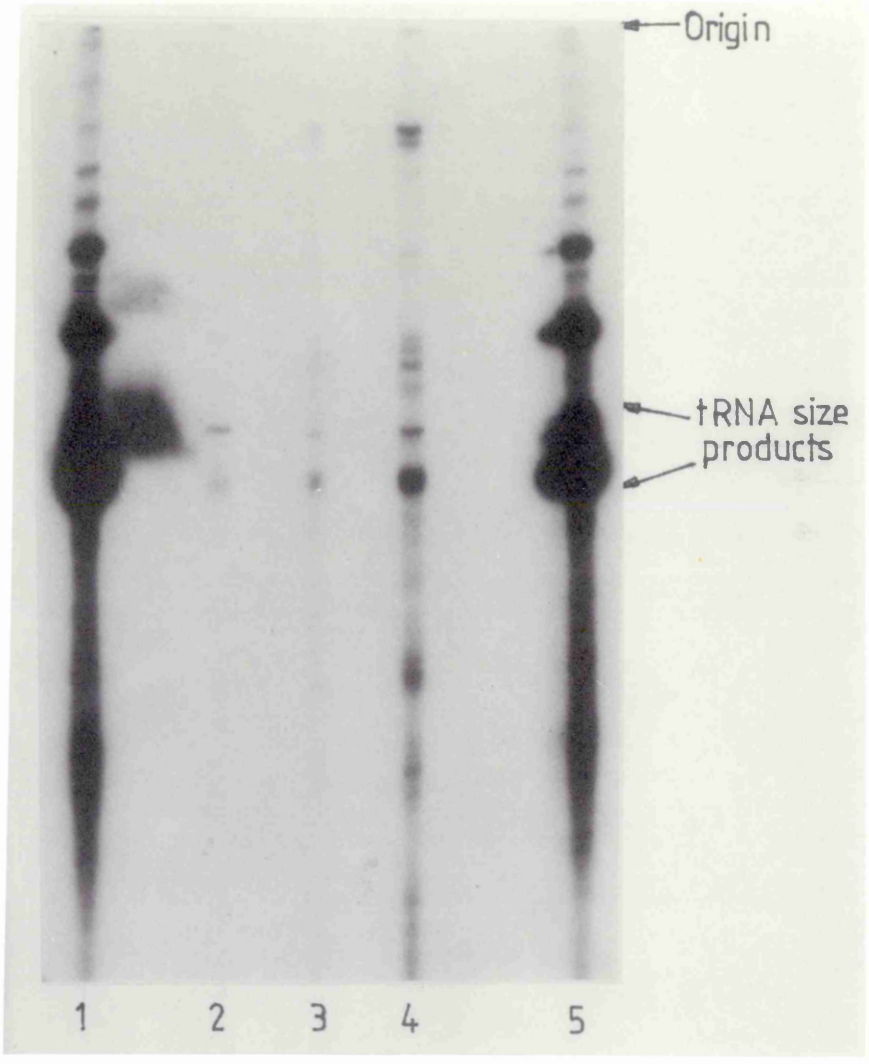
Figure 4.18 Gel electrophoresis of RNA forming hybrids with pNB21 DNA.

Purified pNB21 DNA was bound to nitrocellulose filters and hybridized to [^{32}P] pCp labelled crude human placental 4S RNA. The labelled RNA forming hybrids with pNB21 DNA was released (isolated) and analysed on 10% polyacrylamide gel under denaturing conditions.

lane 1 and 5 - [^{32}P] pCp labelled crude 4S RNA

lane 2 - control (no DNA bound to filter)

lane 3 and 4 - hybrid selected RNA.


Origin
tRNA size
products
1
2
3
4
5

electrophoresis. Two different types of [^{32}P] pCp labelled 4S RNA probes were used for this analysis: a) crude 4S placental RNA, labelled at the 3' - terminus (not purified electrophoretically) and b) crude 4S placental RNA, periodate-treated prior to pCp labelling. Periodate treatment (or periodate protection) is a means of specifically labelling tRNAs. The principle of this procedure (summarized in Figure 4.17) is as follows: Crude 4S RNA (from human placenta) was incubated in the presence of a crude aminoacyl- tRNA synthetase preparation containing most aminoacyl- tRNA synthetases and twenty amino acids (For details of this procedure see Goddard *et al.*, 1983). This achieves the purpose of charging the desired tRNA species, thereby protecting their 3' OH -termini. Periodate oxidation, followed by β-elimination destroys any free, uncharged 3' OH. After deacylation of the protected aminoacyl tRNAs, the only free 3' OH groups will be those of the desired tRNA species, which can be labelled with [^{32}P] pCp using RNA ligase.

The results of the hybrid-selected RNA analysis are shown in Figure 4.18. These results indicate that in the case of non-periodate treated 4S RNA (untreated "crude") pNB21 DNA hybridizes to RNAs of tRNA size along with other RNAs of somewhat lesser electrophoretic mobility. However, in the case of periodate- treated 4S RNA, no labelled RNA was found to hybridize to the recombinant DNA.

At first sight this result appears to suggest that contrary to expectations pNB21 DNA does not actually contain sequences homologous to the abundant human tRNA species. This may be an over simplification. The hybridization that is observed using non-periodate treated RNA may be due to the presence of RNAs other than genuine tRNAs. The preparation of crude human 4S RNA used has not been subjected to prior electrophoretic fractionation (Figure 4.18, lanes 1 and 5) and contains other low molecular

weight RNAs. Such additional RNAs, since they are not removed after the periodate treatment may compete with [^{32}P] labelled tRNAs for complementary sites on the plasmid due to sequence similarities.

To clarify this ambiguous situation, it was necessary to identify the restriction fragments of the respective subclones that hybridize to electrophoretically purified [^{32}P] labelled human '4S RNA' and determine the nucleotide sequence of some of these regions.

9. Hybridization of [^{32}P] pCp labelled human '4S RNA'to the restriction fragments of pNB1 and pNB4

In order to identify restriction fragments of pNB1 and pNB4 that hybridize to the [^{32}P] pCp labelled human '4S RNA', the DNAs from plasmids pNB1 and pNB4 were digested with restriction endonuclease that cleave the insert of these subclones infrequently. The digestion products were fractionated by gel electrophoresis and transferred to nitrocellulose filters. The sequences homologous to human tRNAs were located on the filter by DNA-RNA hybridization followed by autoradiography. Intact plasmid DNA was used for SmaI and PstI digestion, since the vector pAT153 has a unique PstI site and no known SmaI site (Sutcliffe, 1978). This avoids isolation of the human DNA insert. On the other hand, isolated human DNA inserts of the respective plasmids were used for HhaI and AvaI digestion.

Figure 4.19B shows the autoradiograph obtained from this analysis. The bands showing positive hybridization to [^{32}P] pCp labelled human '4S RNA' are as follows: for pNB4, the 1.9 Kb EcoRI plug HindIII band (land 2), 1.2 Kb SmaI - HindIII band (land 3); 0.9Kb PstI band (land 4)

Figure 4.19 Hybridization of [^{32}P] pCp labelled human '4S RNA' to the restriction fragments of pNB1 and pNB4 DNAs.

A - Ethidium bromide stained gel

lane 1 - marker, λDNA cleaved with EcoRI + HindIII

lane 2 - EcoRI + HindIII digest of pNB 4 DNA

lane 3 - SmaI + HindIII digest of pNB 4 DNA

lane 4 - PstI digest of pNB4 DNA

lane 5 - EcoRI digest of the 2.8 Kb insert of pNB4

lane 6 - HhaI digest of the 2.8 Kb insert of pNB4

lane 7 - AvaI digest of the 2.8 Kb insert of pNB4

lane 8 - HindIII digest of pNB1 DNA

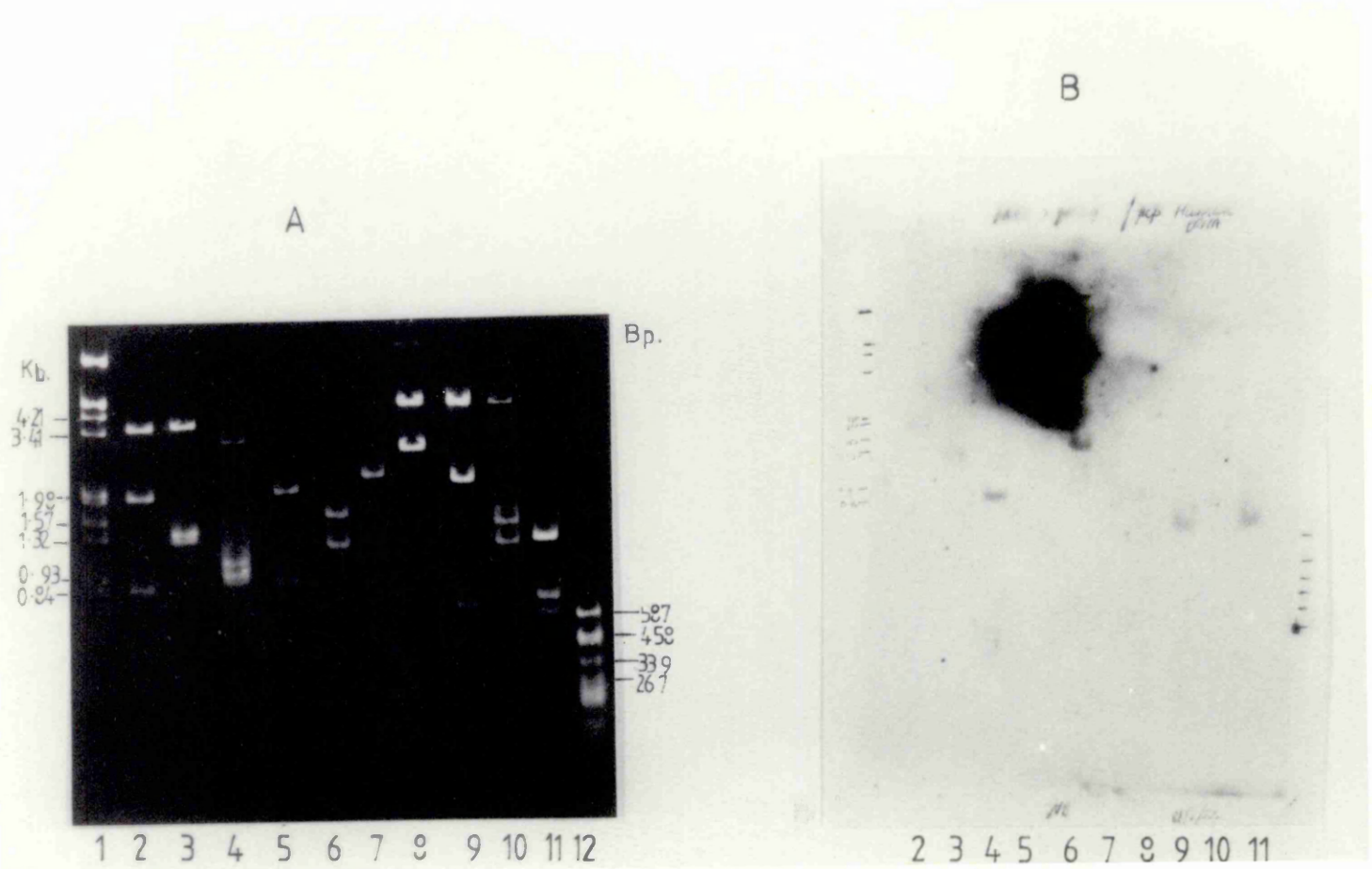
lane 9 - HindIII + SmaI digest of pNB1 DNA

lane 10 - PstI digest of pNB1 DNA

lane 11 - HhaI digest of the 2.7 Kb insert of pNB1

lane 12 - marker, pAT153 DNA cleaved with Hae III

B - Autoradiograph of the Southern blot prepared from gel shown in A.


A
Kb.
4·27
3·43
1·90
1·57
1·32
0·93
0·84
Bp.
527
458
339
267
1 2 3 4 5 6 7 8 9 10 11 12
B
2 3 4 5 6 7 8 9 10 11

and 1.5 Kb HhaI (land 6). For pNB1: 0.6 Kb SmaI - HindIII band (land 9) and 0.72 Kb HhaI - HindIII band (lane 11). Thus the smallest fragments that could be used for nucleotide sequence determination analysis are the 0.9 Kb PstI restriction fragment of pNB4 and the 0.6 Kb SmaI - HindIII fragment of pNB1.

10. Nucleotide sequence determination of the 640bp HindIII - SmaI fragment

The 640 bp HindIII - SmaI restriction fragment of the recombinant plasmid pNB1 was sequenced exclusively by the chemical sequencing method of Maxam and Gilbert (1980). This method requires a unique DNA fragment labelled at its terminus with high specific activity ^{32}P phosphate. Such DNA fragments were created by digestion with suitable restriction endonucleases, followed by *in vitro* labelling of both termini. The two labelled ends are then separated by either a secondary restriction endonuclease cleavage or by strand-separation procedure. These procedures resulted in fragments that were labelled only at one end and could then be directly sequenced. The procedure for sequencing end-labelled DNA fragments is discussed in Materials and Methods section 2.t.

The complete sequence of the 640 bp HindIII - SmaI fragment was derived from a large number of individual determination of overlapping restriction fragments on both strands. As the nucleotide sequence was determined the exact positions of further restriction enzymes became known (For example DdeI; Sau 96A; MspI etc.). In the following section preparation of fragments for sequencing is discussed.

10.a Secondary restriction preparations

The majority of fragments for nucleotide sequence analysis were

Figure 4.20 DNA sequencing strategy of the 640 bp SmaI - HindIII fragment of pNB 1.

The duplex DNA is represented by a single line. The arrows represent the direction and extent of sequence from the labelled end of the different restriction site. Arrows above the line represent the sequence of upper strand of DNA and arrows below the line represent sequence of the lower strand.

1. - Restriction endonuclease cleavage map of the 640 bp SmaI - HindIII fragment.

2,3 and 4 - fragments prepared for secondary restriction.

5 and 6 - fragments prepared for strand separation.

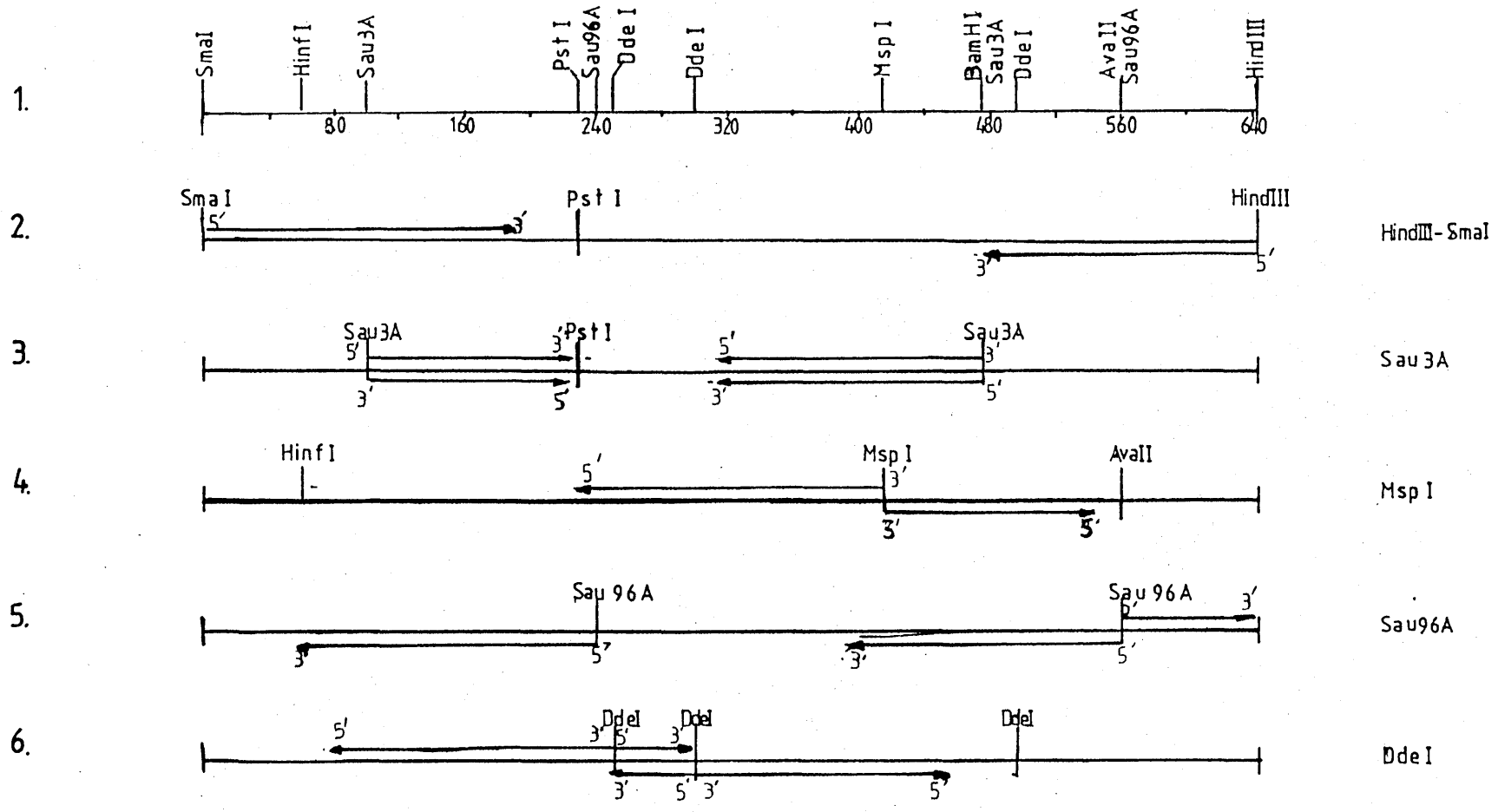
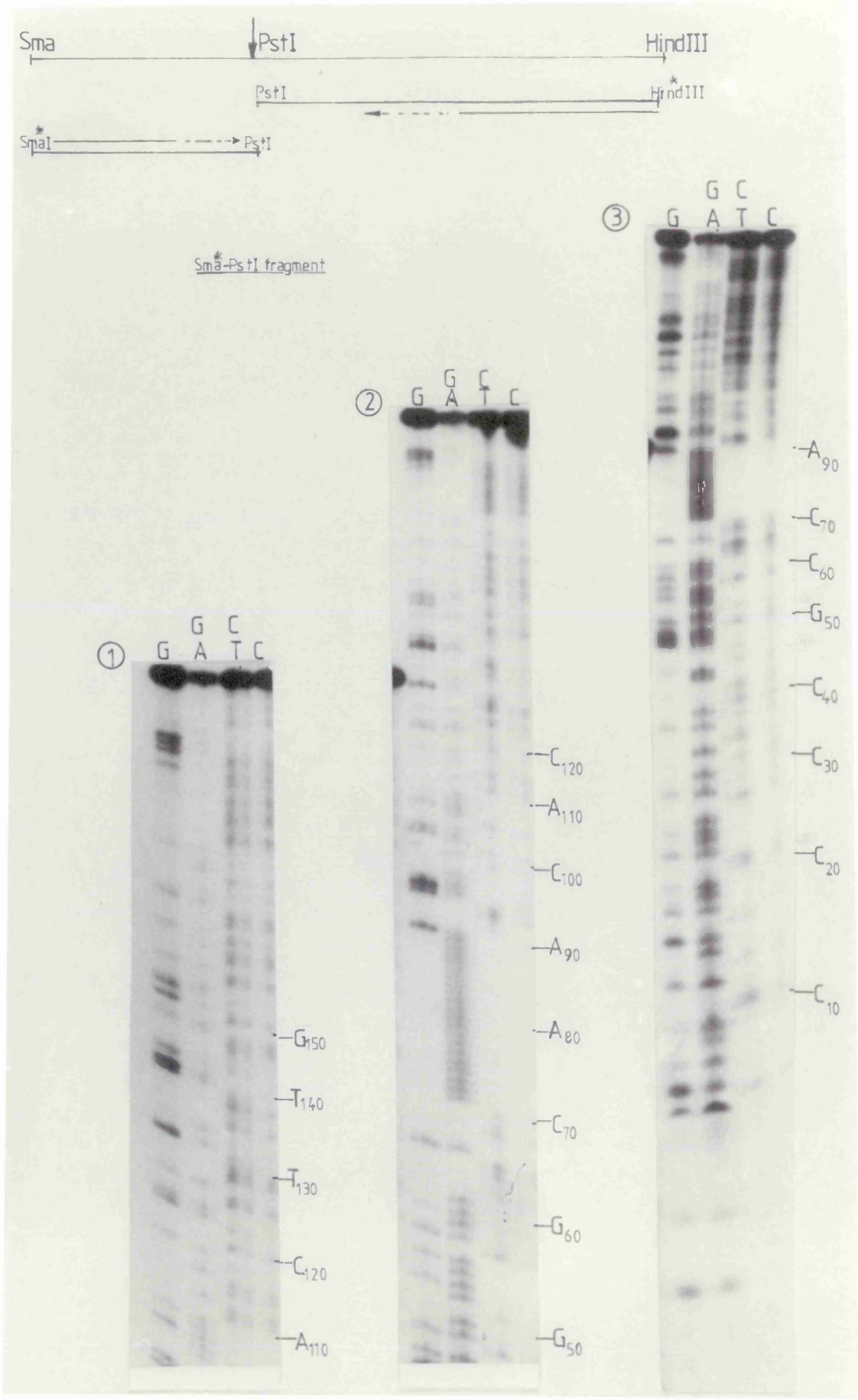

1.
SmaI
HinfI
Sau3A
PstI
Sau96A
DdeI
DdeI
MspI
BamHI
Sau3A
DdeI
AvaII
Sau96A
HindIII
80
160
240
320
400
480
560
640
2.
SmaI
PstI
HindIII
5'
3'
3'
5'
HindIII-SmaI
3.
Sau3A
PstI
Sau3A
5'
3'
3'
5'
5'
3'
3'
5'
Sau3A
4.
HinfI
MspI
AvaII
5'
3'
3'
5'
Msp I
5.
Sau 96A
Sau 96A
3'
5'
5'
3'
Sau96A
6.
DdeI
DdeI
DdeI
5'
3'
5'
3'
3'
5'
3'
5'
Dde I

Figure 4.21 Autoradiographs of sequencing gels from SmaI* - PstI and HindIII* - PstI fragments of the 640 bp SmaI - HindIII fragment. Approximately 5µg of the purified 640 bp, SmaI - HindIII fragment was 5' end labelled (*) using [γ-^{32}P]-ATP and cleaved asymmetrically with restriction endonuclease PstI. The resulting fragments: SmaI* - PstI and HindIII* - PstI were subjected to sequence analysis by the method of Maxam and Gilbert (1980). In each case samples from the four reactions were loaded on a sequencing gel (40x20 cm) and electrophoresed until the Xylene cyanol marker dye had travelled 20cm(1). A second loading was made in the four adjacent wells and the marker dye was allowed to travel 17 cm (2). A third and final loading was made and electrophoresis continued until the marker dye had travelled for 13 cm (3). Thus in (1) the marker dye has travelled for a total of 50 cm, and in (2) for 30 cm. The shortest molecules, representing the 5' end of the fragment are resolved in (3) while in (1) the shortest fragments have moved off the end of the gel but the longer molecules are more clearly resolved. The DNA sequence which can be determined is shown for each loading. The numbers given correspond to the distance from the SmaI restriction site and are derived from the complete sequence (Figure 4.25). The overlap between successive loading is illustrated by the numbers.


Sma
PstI
HindIII
PstI
HindIII
SmaI
PstI
Sma-PstI fragment
①
G A T C
G150
T140
T130
C120
A110
②
G A T C
C120
A110
C100
A90
A80
C70
G60
G50
③
G A T C
A90
C70
C60
G50
C40
C30
C20
C10

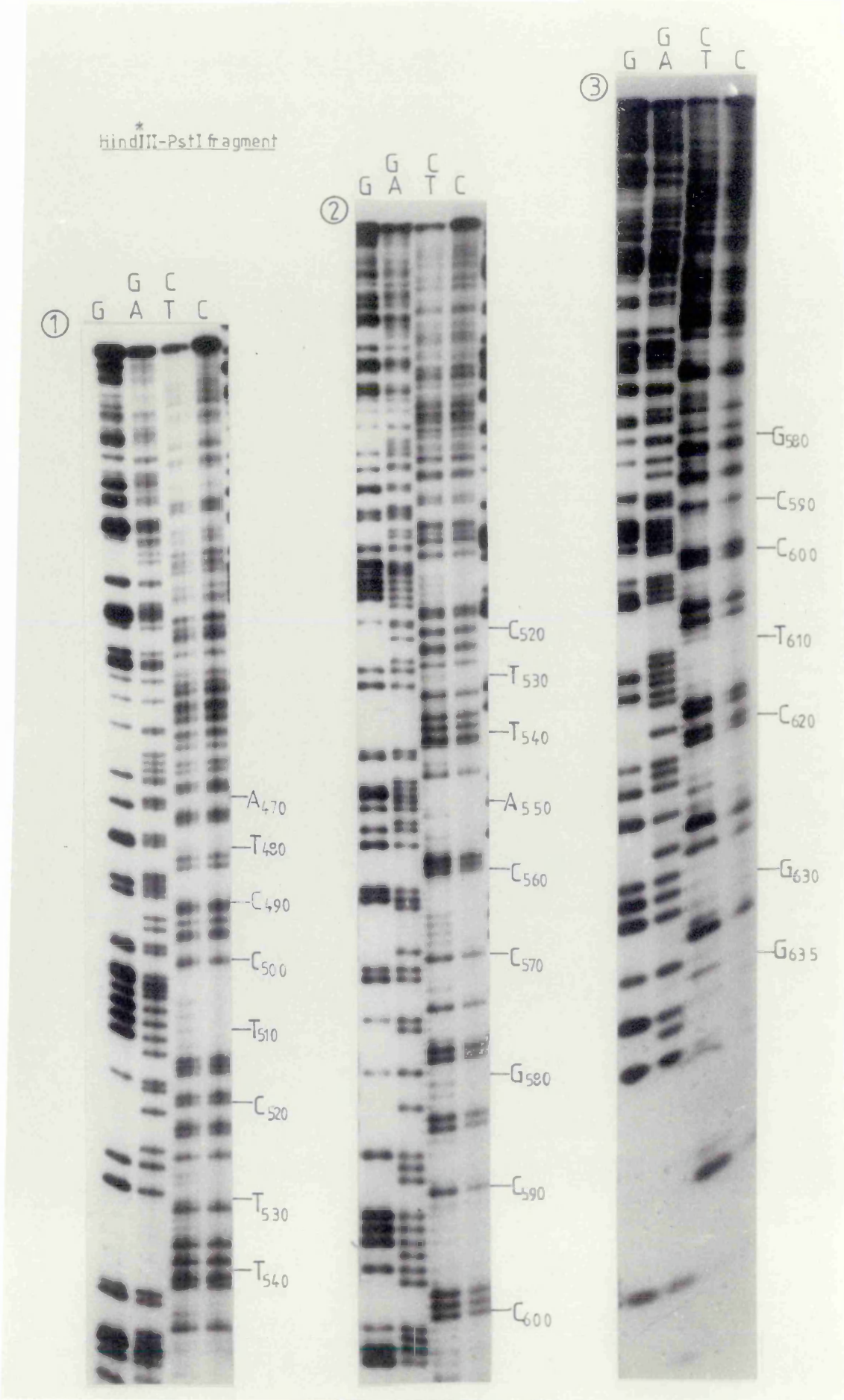
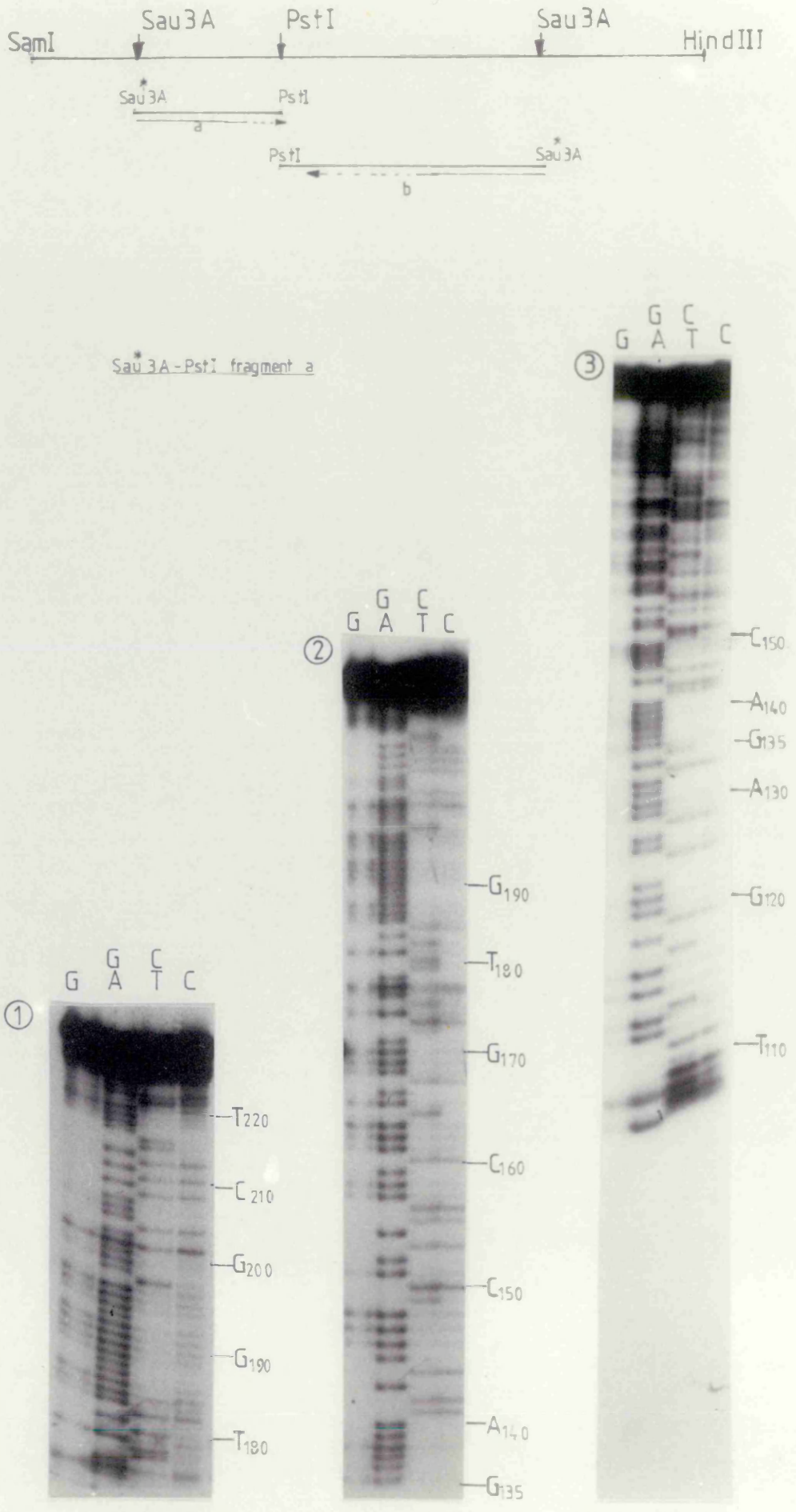
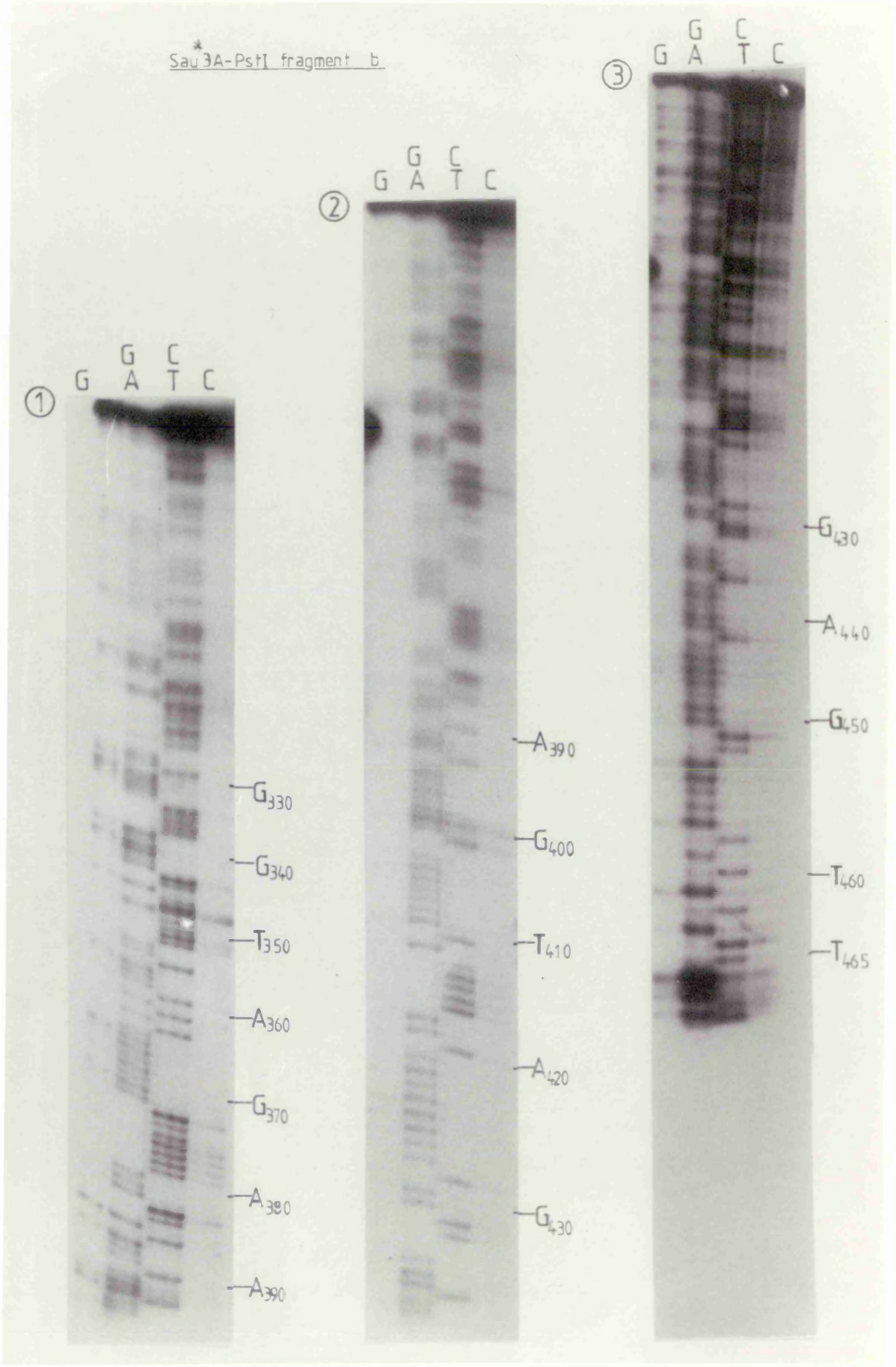

*HindIII-PstI fragment
G C
G A T C
①
A470
T480
C490
C500
T510
C520
T530
T540
②
C520
T530
T540
A550
C560
C570
G580
C590
C600
③
G580
C590
C600
T610
C620
G630
G635

Figure 4.22 Autoradiographs of DNA sequencing gels from Sau 3A* - PstI fragments of the 640 bp. SmaI - HindIII fragment. Approximately 5ug of the purified 640 bp. SmaI - HindIII fragment was cleaved with restriction endonuclease Sau 3A. The 400 bp Sau 3A - Sau 3A fragment was isolated, labelled at 3' end with α dGTP and subjected to secondary restriction digest with restriction endonuclease PstI. The resulting Sau 3A* - PstI fragments were subjected to sequence analysis. The number correspond to the distance from the SmaI restriction site and are derived from the complete sequence of the 640 SmaI - HindIII fragment.


SamI
Sau3A
PstI
Sau3A
HindIII
Sau3A
PstI
a
PstI
Sau3A
b
Sau3A-PstI fragment a
G A T C
①
T220
C210
G200
G190
T180
G A T C
②
G190
T180
G170
C160
C150
A140
G135
G A T C
③
C150
A140
G135
A130
G120
T110


Sau3A-PstI fragment b
G A T C
①
G330
G340
T350
A360
G370
A380
A390
②
A390
G400
T410
A420
G430
③
G430
A440
G450
T460
T465

prepared by digesting a pure end-labelled DNA, with a restriction endonuclease which cleaved the fragment asymmetrically. After this secondary restriction, the fragments can again be separated and purified by gel electrophoresis (6% polyacrylamide), and the band located by ethidium bromide staining. Figure 4.20 shows fragments which were prepared in this way and indicates the approximate length of sequence determined.

In the initial experiment, the 640 bp HindIII - SmaI fragment was isolated, end - labelled with ^{32}P at the 5' end using γ - ^{32}P-ATP and T4 polynducleotide kinase (Maxam and Gilbert 1980). This was then cleaved asymmetrically with restriction endonuclease PstI. The resulting two fragments were separated and isolated on a 6% polyacrylamide gel. These fragments were then sequenced. Two of the sequencing gels obtained from this determination are shown in Figure 4.21. Similarly, the 640 bp fragment was separately subjected to further restriction endonucleases cleavage with Sau 3A, BamHI and MspI. The resulting restriction fragments were end-labelled at their 5' end as before or at 3' end using Klenow fragment of DNA polymerase I and appropriate [α - ^{32}P] deoxyribonucleoside triphosphate. The specific fragments to be sequenced were purified by electrophoresis and recleaved with another restriction endonuclease. The sequencing gels obtained from some of the these determinations and shown in Figure 4.22.

10.b Strand separation preparations

In the first of these preparations the whole of 640 bp SmaI - HindIII fragment was cleaved with restriction endonuclease DdeI and the products end-labelled, denatured and separated on a non-denaturing strand-separation gel (Materials and Methods 2.t iii). As shown in Figure 4.23A a complex pattern of bands were produced, the complementary strand of many fragments of DNA migrated through the gel with different electrophoretic

Figure 4.23 Preparation of restriction fragments for strand separation procedure.

Approximately 2µg of the 640 bp. SmaI - HindIII fragment was cleaved with restriction endonuclease DdeI. The restriction cleavage products were end-labelled, denatured by treatment with 0.2M NaOH and subjected to strand separation on an 8% non-denaturing gel.

lane a - Strand - separated products

land b - DdeI restriction cleavage products (Untreated)

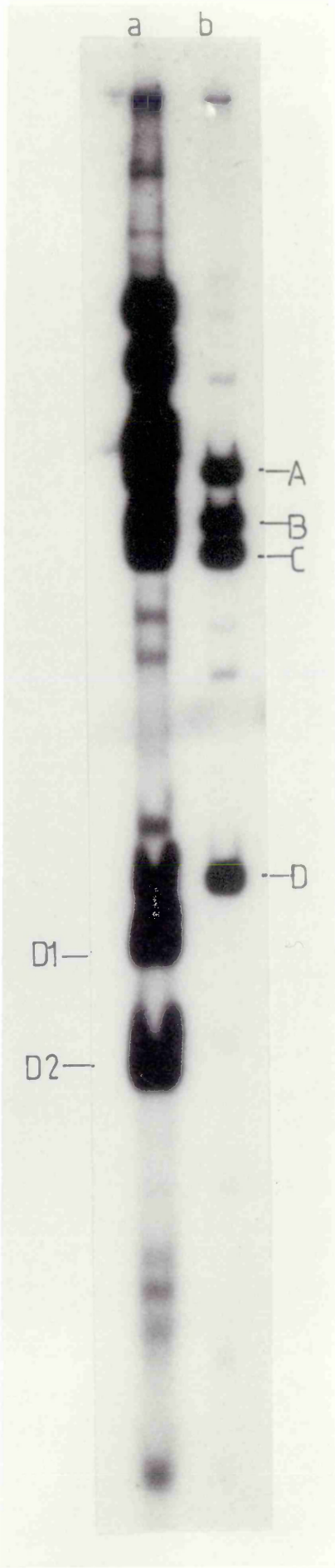

a
b
A
B
C
D
D1
D2

Figure 4.24 Autoradiograph of DNA sequencing gels from the strand separated DdeI - DdeI fragment of the 640 bp SmaI - HindIII fragment. Sequencing gels (15 cm loading only) from bands D_1 and D_2 (Figure 4.23) are shown. It can be seen that the two sequences are complementary, being derived from the opposite strands of a single restriction fragment. Numbering is according to the complete sequence (Figure 4.25).

The following symbols were used for the identity of some uncertain nucleotide bands in the sequence determination analysis (adapted from Stadan, 1979)

R = A or G

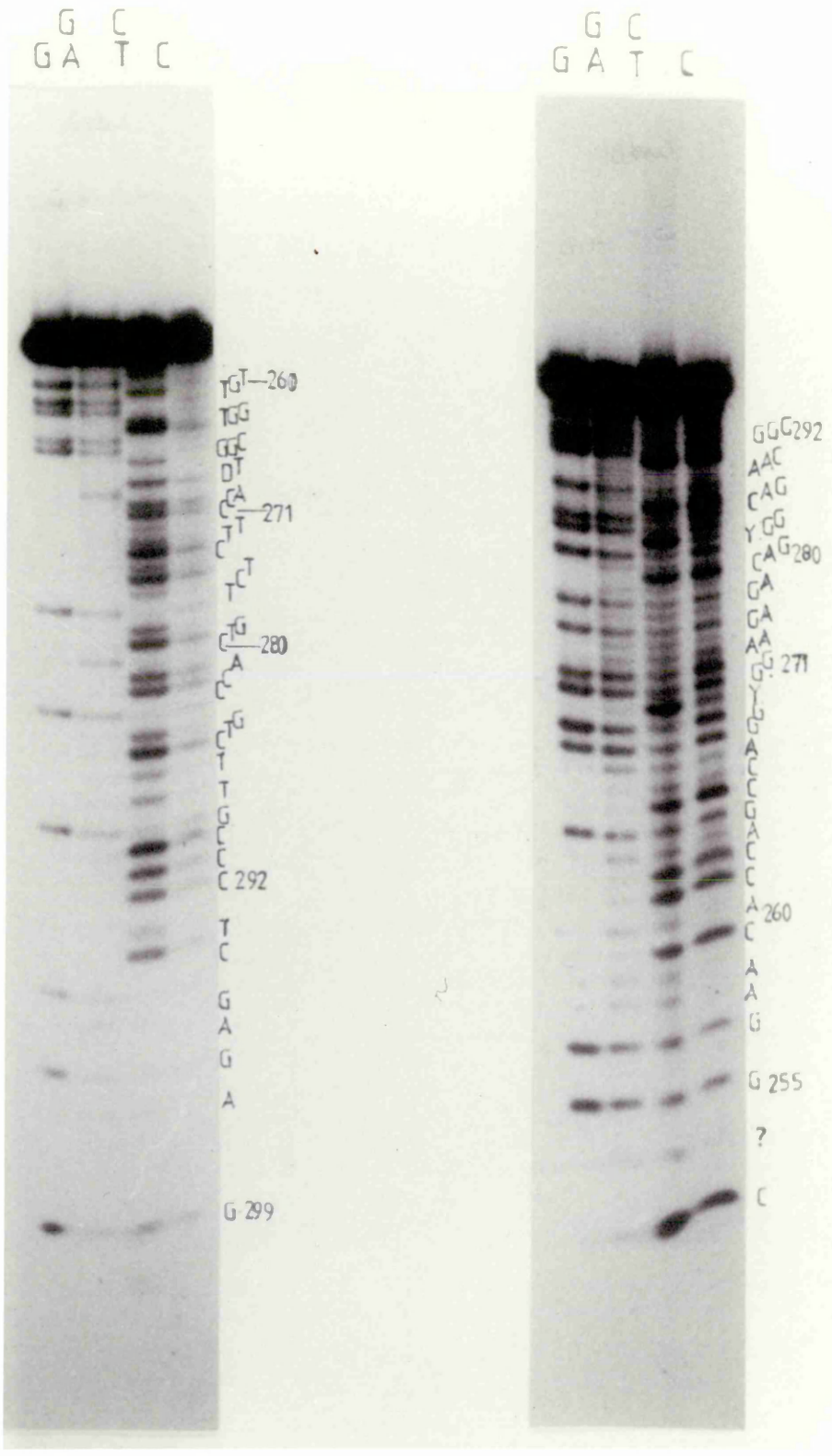
Y = C or T

D = C, possibly 2 or more

H = G possibly 2 or more

X = any nucleotide

? = presence uncertain


G C
G A T C
260
271
280
292
299
G C
G A T C
292
280
271
260
255
?

mobilities (probably due to secondary structures, as a consequence on intra-strand base pairing). However, some strands fail to separate (Figure 4.23). The bands from the separated strands were eluted and sequenced, the length of nucleotide sequence determined from each band is shown in Figure 4.20. Sequencing gels obtained from the two strands of the smallest DdeI fragment are shown in Figure 4.24A.

10.c Interpretation of sequencing gels

A number of sequencing gels are shown in Figure 4.21, 4.22 and 4.24. In most cases the gels can be read unambiguously up to the point where the bands are no longer resolved. The length of sequence which can be read from each gel depends on the resolution of individual gel and on the extent of attack of each chemcial modification (as well as length of the fragment). However, ambiguities were seen in some ladder patterns particularly in the relative intensities of Cs and Ts. Throughout this sequencing study, repeated sequencing of fragments was carried out and confirmed by sequencing the complementary strand in opposite direction, until such ambiguities were eliminated and definitive results obtained. As shown in Figure 4.20, 85% of the nucleotide assignments of the 640 bp SmaI - HindIII fragment were determined by at least two independent sequencing schemes. The complete sequence which was deduced from the combined determination is shown in Figure 4.25.

11. Search for tRNA gene sequences

The complete nucleotide sequence of the 640 bp SmaI - HindIII fragment was searched for tRNA gene sequences with the aid of a tRNA computer program devised by Staden (1980). This program reads through a

Figure 4.25 The complete DNA sequence of the 640 bp SmaI - HindIII fragment of pNB1, deduced from a combined determination of overlapping fragments. The topstrand (5' to 3', left to right) is termed the non-coding strand since it contains part of the non-coding (RNA-like) strand of the human Alu consensus sequence derived by Schmid and Jelinek (1982) (see also Figure 4.28 and Figure 5.2). Some key restriction sites used in the nucleotide sequence determination are indicated by overlying lines.

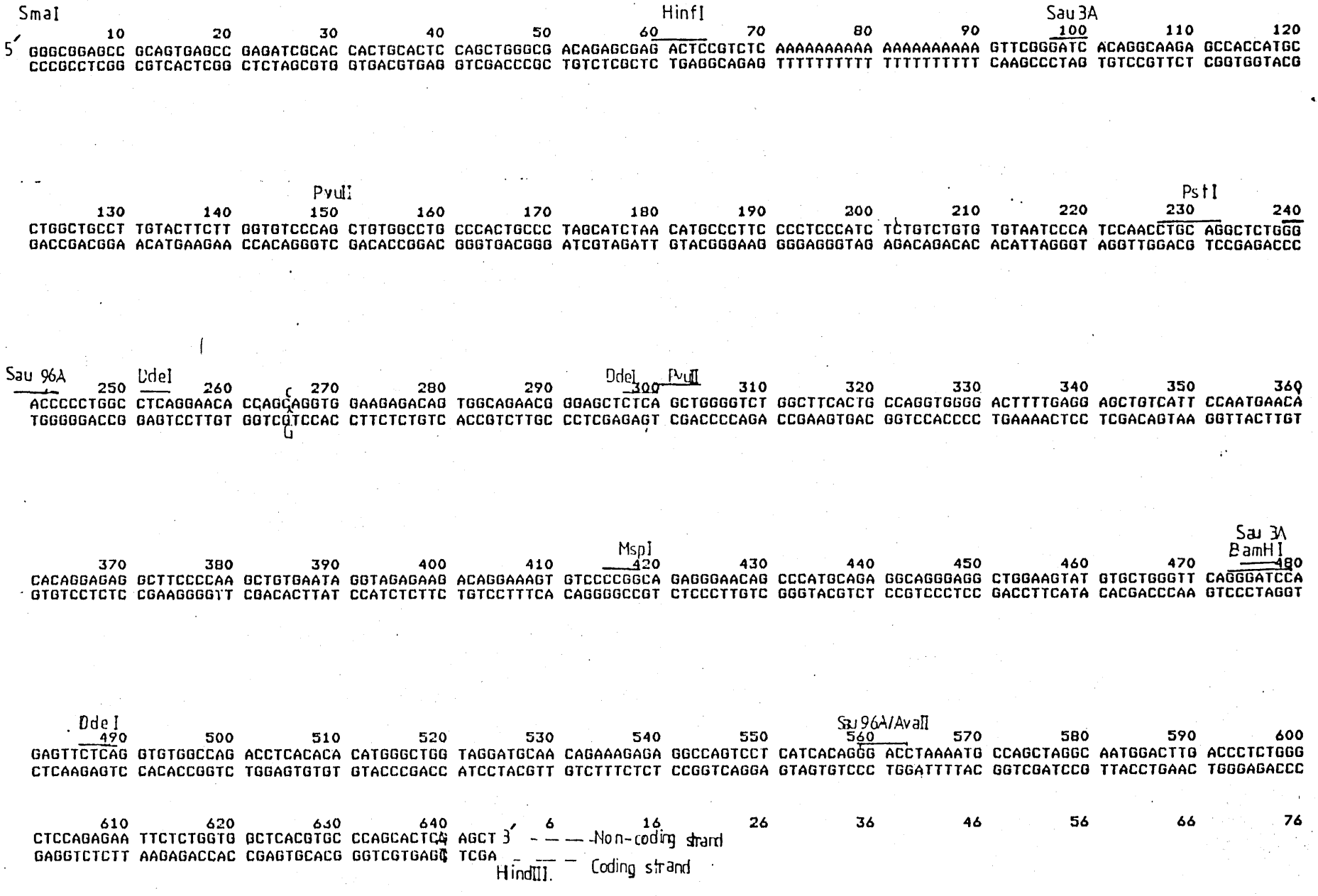

SmaI
HinfI
Sau3A
PvuII
PstI
Sau 96A
DdeI
DdeI
PvuII
MspI
Sau 3A
BamHI
DdeI
Sau96A/AvaII
5′
3′
Non-coding strand
Coding strand
HindIII
FIRST SEQ NO =-1

Figure 4.26 Some regions of the 640 bp SmaI - HindIII sequence that could fold into a clover leaf structure as determined by using the computer program of Staden (1980). These structures were examined for the 18 conserved bases normally present in a typical tRNA. The position of these conserved bases, according to the standard method of numbering tRNAs (Gauss *et al.*, 1979) are as follows:-

T-8; G-10; C or T-11; A-14; A or G-15; A-21; C or T-32; T-33; A-37; T or C-48; G-53; T-54; T-55; C-56; A or G-57; A-58; C or T-60; C-61.

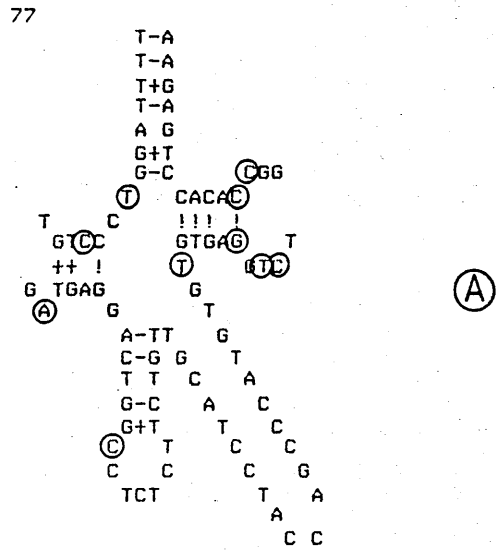
A circled base indicates that the conserved base is present. The nucleotide sequence of these regions is that of the lower strand of the sequence. Consequently the residue numbering shown here is in reverse order to that defining the upper strand of Figure 4.25.

A - residues 77-162 are complementary to residues 482-568 of the top strand of Figure 4.25.

B - residues 350-431 are complementary to residues 213-295 of the top strand of Figure 4.25.

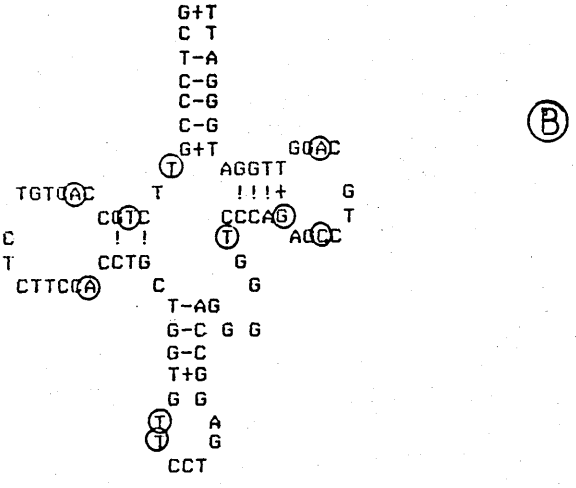
C - residues 403-502 are complementary to residues 142-242 of the top strand of Figure 4.25.

77
TTTTAGGTCCCTGTGATGAGGACTGGCCTCTCTTTCTGTTGCATCCTACCAGCCCATGTGTGTGAGGTCTGGCCACACCTGAGAAC
******0T 0!!!AR !!!! ----0YT A ----- ++++G C C+++1******



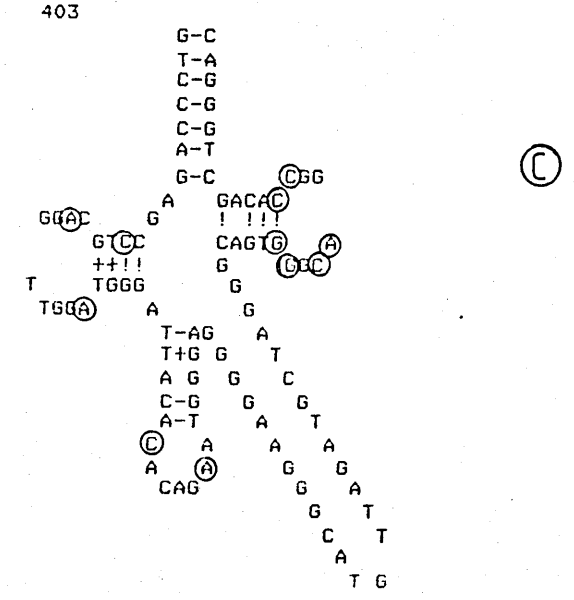
A

350
GCTCCCGTTCTGCCACTGTCTCTTCCACCTGCTGGTGTTCCTGAGGCCAGGGGGTCCCAGAGCCTGCAGGTTGGATGGGATTA
******0T 0!!!AR !!!! ----0YT A ----- +++++ G C C+++0*******



B

403
GTCCCAGAGCCTGCAGGTTGGATGGGATTACACACAGACAGAGATGGGAGGGGAAGGGCATGTTAGATGCTAGGGCAGTGGGCAGGCCACAGCTGGGAC
******4T 2!!!AR !!!! ----0YT A ----- ++++G C C+++1*******



C

Figure 4.27 Comparison of the homologous regions of the 640 bp SmaI - HindIII sequence and the 3.18 Kb *X. laevis* tDNA sequence. The top line represents residues 54 to 5of the lower strand of the sequence shown in Figure 4.25. The bottom line represents part of the RNA-like strand of the *X. laevis* $tRNA^{Tyr}$ gene (Muller and Clarkson, 1980). An asterisk (*) at any position indicates the same nucleotide in the two sequences.

5′ TCTGTCGCCC AGCTGGAGTG CAGTGGTGCG ATCTCGGCTC ACTGCGGCTC C 3′ - 640 bp. SmaI - HindIII sequence

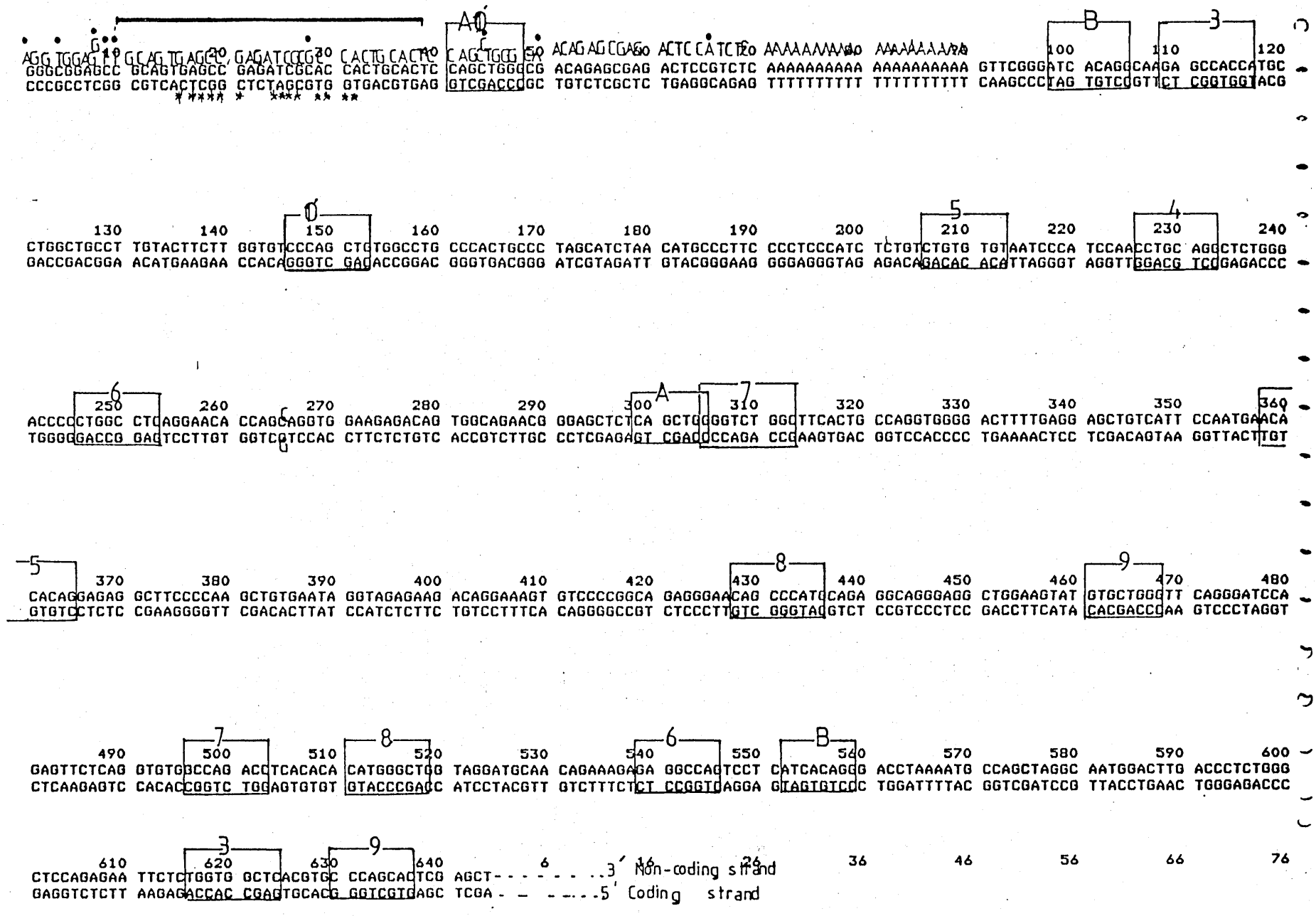
* * * ** * **** ** ** ****** *

5′ TGATCGAGCA ATCCTTAGGT CGCTGGTTCG ATTCCGGCTC GAAGGACGCT T 3′ - 3.18 Kb. X.Laevis sequence

given sequence and searches for sections that could fold into a clover leaf. In doing so it also searched for the 18 conserved bases, that almost always appear at certain positions in the clover leaf. Using this computer program several sections of the 640 bp. SmaI - HindIII fragment were identified that could fold into a clover leaf structure. The sequence of these possible tRNA regions was checked for the tRNA conserved bases and secondary structures. This revealed that in all these possible tRNA regions have much smaller loops and few of the conserved bases (see Figure 4.26). In most cases only 9 or 10 of the 18 conserved bases were present, suggesting that none of these regions conform to the typical tRNA gene structure.

Since the λht3 clone was originally isolated using *X. laevis* pXt210 DNA, it was decided to compare the nucleotide sequence of the 640 bp SmaI - HindIII fragment with that of the 3.18 Kb sequence of pXt210 (the sequence data of 3.18 Kb, a gift from S.G. Clarkson). In order to find homologies between the two sequences a computer program 'SEQFIT' (Staden, 1977) was employed. The only region of homology was found at nucleotide 16 to 32 of the SmaI - HindIII fragment and nucleotides 453-470 of the 3.18 Kb sequence. These regions share 82% homology since 14 of the 17 residues are identical. Figure 4.27 shows the regions of homology between these two sequences. Residues 453-470 of the 3.18 Kb sequence are within the 3' end (TψCloop) of the $tRNA^{Tyr}$ gene, corresponding to nucleotides 51 to 68 numbered according to the standard method for mature tRNA (Gauss *et al.*, 1979), (Muller and Clarkson 1980).

Figure 4.28 DNA nucleotide sequence of the 640 bp SmaI - HindIII fragment showing repeat sequences and regions of homology to human Alu sequence. Boxed areas with numbers on top represent inverted repeats, boxed areas with letters on top represent direct repeats. Overlying bold sequence on top represent human Alu consensus sequence. Overlying dot (.) at any position indicates a different nucleotide as to that present in the human Alu consensus sequence. Overlying bold line indicates the 31 bp insert of human Alu sequence. Asterisk (*) indicates homology to *Xenopus* $tRNA^{Tyr}$ gene seqeunce.



FIRST SEQ NO =-1
TT10 -- STOP

12. Other features of the 640 bp SmaI - HindIII sequence

The nucleotide sequence of the 640 bp SmaI - HindIII fragment was examined for open reading frames. It contains three initiator codons at positions 116, 353 and 524. In the same reading frames there are termination sites at positions 171, 392 and 563 respectively. Therefore the nucleotide sequence appears to contain no long open reading frames. However there are several direct or indirect repeat sequences of eight or more nucleotides within the 640 bp sequences, as shown in Figure 4.28. Another interesting feature of this sequence is the stretch of 20 oligo d(A) residues (position 71-90), such oligo d(A) tracts are a characteristic feature of Alu family members (repetitive DNA element) and some pseudogenes of non-messenger RNA. An examination of the neighbouring sequences to the oligo d(A) residues revealed that 1-90 of the 640 bp SmaI - HindIII sequence share 90% homology with human Alu consensus sequence derived by Schmid and Jelinek (1982). (see Figure 4.28). This region also contains the 31 bp human insert normally present in the right (distal) monomer of the human Alu sequence. Furthermore, the 17 residue region sharing 82% homology with *X. laevis* $tRNA^{Tyr}$ gene is also located within this 31 bp insert, as shown in Figure 4.28.

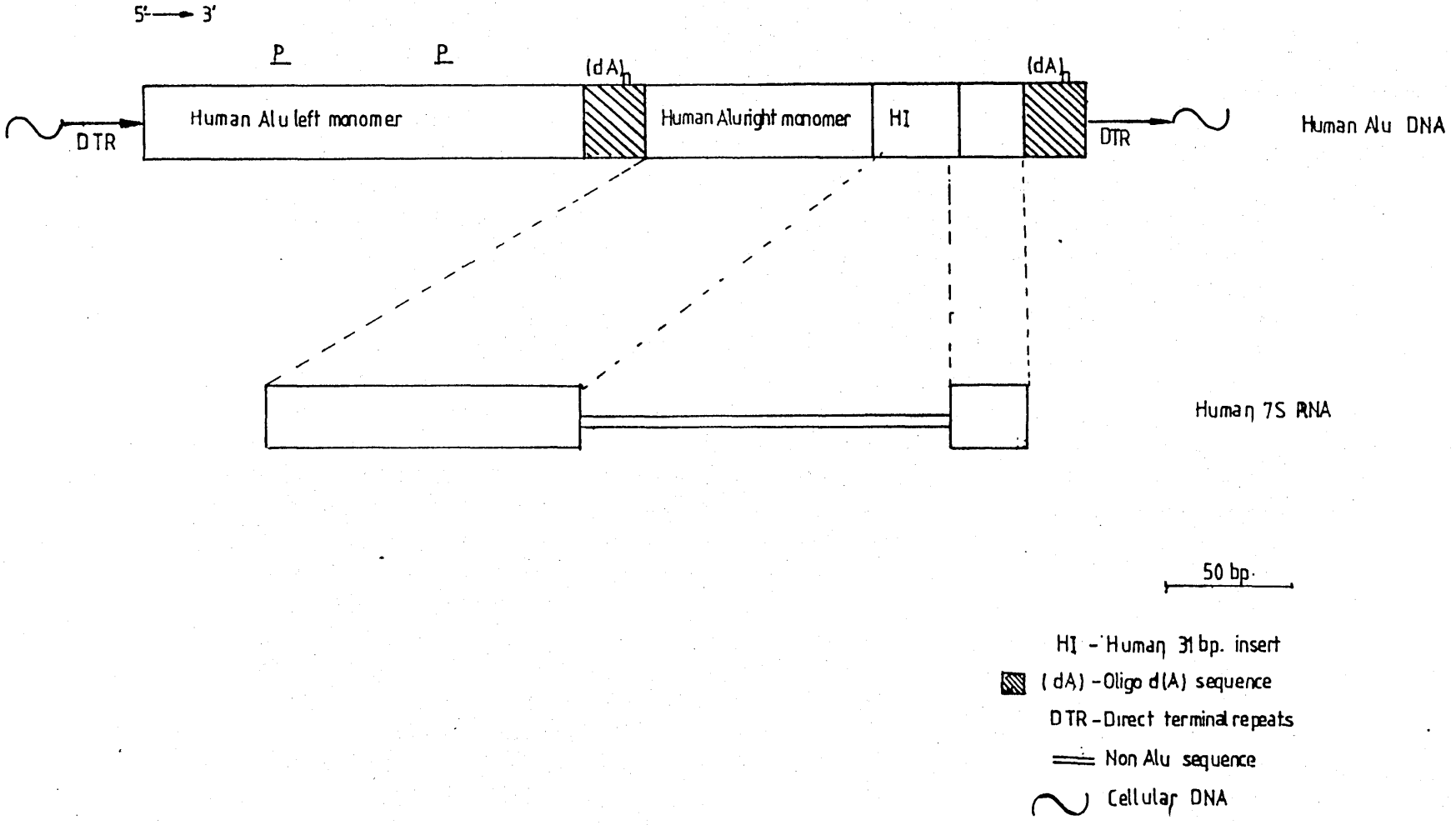
CHAPTER 5 DISCUSSION

The aim of this study was to screen a human genomic library and isolate DNA clone(s) homologues to human tRNA(s) for a study of tRNA gene structure and arrangment. The clone λht3, isolated in this study, was of potential interest.

It might be argued that the homology observed between *X. laevis* $tRNA^{Tyr}$ and the 17 nucleotide region of the 640 bp fragment is too small to be the explanation of the original isolation of λht3. There could have been non-specific hybridization, possibly due to high G+C content of λht3 DNA. Another problem could be that the ^{32}P-labelled human '4S RNA' used as a probe could have been contaminated with 7S RNA molecules that might hybridize to the Alu sequence detected. The human 7S RNA (Weiner, 1980) is an abundant cytoplasmic RNA of 300 nucleotides that shows regions of homology with Alu DNA. Approximately 100 nucleotides at the 5' end and 45 nucleotides at the 3' end of the 7S RNA are homologous with human Alu right monomer consensus sequence (see Figure 5.1) (Ullu *et al.*, 1982; Ullu and Tschudi, 1984). If the hybridization was due to Alu sequence DNA, then more than one positive clone should have been isolated in the screening experiments (Results section 4.2), since the Alu family is present in the human genome at a level of 500,000 members, representing about 3 to 6 per cent of the genome (Houck *et al.*, 1979); Rubin *et al.*, 1980; Schmid and Jelinek, 1982).

There is some evidence to suggest that the hybridization of electrophoretically purified ^{32}P-labelled human '4S RNA', is not due to contaminating 7S RNA molecules. The restriction endonuclease SmaI cleaves pNB1 DNA into a 640 bp and a 2.1 Kb fragment. Only the 640 bp fragment, however, hybridizes to ^{32}P-labelled human '4S RNA' probe (as shown in Figure 4.19). The DNA sequence analysis of the 640 bp SmaI - HindIII

Figure 5.1 Structure of human Alu sequence and its relationship to human 7S RNA.


5'⟶3'
P
P
(dA)n
(dA)n
Human Alu left monomer
Human Alu right monomer
HI
DTR
DTR
Human Alu DNA
Human 7S RNA
50 bp.
HI - Human 31 bp. insert
(dA) - Oligo d(A) sequence
DTR - Direct terminal repeats
Non Alu sequence
Cellular DNA

fragment reveals that the first ninety nucleotides starting from SmaI site are homologous to the 3' end of the right monomer of human Alu DNA sequence (Figure 4.28). The remainder of the human Alu DNA sequence is presumed to be contained in the 2.1 Kb fragment of pNB1. Therefore, if the observed hybridization is due to 7S RNA molecules, then positive hybridization should also be observed to the 2.1 Kb fragment. The hybridization of ^{32}P-labelled human '4S RNA' to HhaI digestion products of pNB1 DNA also support the view that 7S RNA contamination is not a likely explanation (Figure 4.10 lane 11).

As to possible lack of specificity of hybridization DNA sequence analysis of the 640 bp. SmaI - HindIII fragment reveals that the G+C content of this sequence is only 56.6% and there are no long runs of Gs and Cs. This argues against the possiblitiy of G+C rich non-specific hybridization. Moreover, all hybridization were in fact performed under very stringent conditions (5XSSC, 50% formamide, 42^{o}C, see Materials and Methods section 2i) that require at least 80% homology between complementary sequences for duplex formation (Maniatis *et al.*, 1982). Furthermore, the above argument does not explain the hybridization of pXt 210 and pt 63 DNA to the restriction fragment of λht3 DNA (Figure 4.4 and Figure 4.6). In addition, although *X. laevis* tRNA^{Phe} and tRNA^{Tyr} are themselves 70% homologous (Muller and Clarkson, 1980), only the tDNA^{Tyr} subclone of *X. laevis* (pt 63) hybridizes to restriction fragments of λht3 DNA. These findings strongly suggest that the results of this study are not based on artefactual hybridization but are due to sequences homologous to part of tDNA^{Tyr}.

One possibility is that the short DNA region homologous to a segment of tRNA^{Tyr} might represent a truncated tRNA gene or a variety of pseudogenes. A pseudogene is described as being similar to a functional gene,

but incorporating sufficient changes that it is not biologically active, and in most cases not transcribed into RNA. Pseudogenes have been found for an increasing number of structural genes and several cases of low molecular weight RNA genes have also been reported. There are numerous examples of tRNA genes that differ from the expected DNA sequence by one or more nucleotides (Hosbach *et al.*, 1980; Addison *et al.*, 1982; Sharp *et al.*, 1981; Ng and Abelson 1980; Santos and Zasloff, 1981). In addition to minor sequence variation, some tRNA genes are extensively disrupted, so that only small fragment of the coding sequence remains intact. Such incomplete genes have been described in *Drosophila* (Sharp *et al.*, 1981) and in rat (Shibuya *et al.*, 1983; Rosen *et al.*, 1984). Sharp and co-workers (1981) have isolated and subcloned a 1150 bp *Drosophila* DNA fragment that hybridizes strongly to initiator tRNA and also displays properties of moderately repetitive DNA. This fragment contains four regions of homology with initiator tRNA, the longest of which is only 34 bp and the others are 14 bp, 14 bp and 13 bp long respectively. This fragment also hybridizes to about 30 dispersed chromosomal sites (Sharp *et al.*, 1981). These findings have led Sharp *et al.*, (1981) to suggest that these pseudogenes may have arisen by the repeated insertion and excision of a transposable element into an intact tRNA gene. In another case, a 22 Kb DNA fragment was isolated by hybridization of $tRNA^{Phe}$ and found to contain a sequence of 35 nucleotides corresponding to the 3' end of $tRNA^{Phe}$ gene (Rosen *et al.*, 1984). Reilly *et al.*, (1982) have reported a murine DNA clone containing 38 contiguous nucleotides to the 3' end of $tRNA^{Phe}$. This clone also encodes the 3' terminal CCA. However, all available evidence suggests that CCAs are added post-transcriptionally in eukaryotes. In addition to this Shibuya *et al.*, (1983) have isolated a 3.3 Kb DNA sequence of rat genome that reoccurs about 10 times and contains a cluster of four different

tRNA genes: $tRNA^{Asp}$, $tRNA^{Leu}$, $tRNA^{Glu}$ and $tRNA^{Gly}$. Nucleotide sequence analysis of six of these repeat units revealed that four clones carry incomplete genes (Shibuya *et al.*, 1984). In crude cell free transcription systems, none of these genes with nucleotide deletions produce any transcriptional products (Rosen *et al.*, 1984; Shibuya *et al.*, 1982).

The suggestion that tRNA pseudogenes have arisen by repeated insertion and excision of transposable element into an intact tRNA gene (Sharp *et al.*, 1981) is further substantiated by the consistent occurrance of repetitive elements in the 5' non-coding regions of several eukaryotic tRNA genes (Eigel and Feldman, 1982; Gafner *et al.*, 1983; Rodi *et al.*, 1983; Han and Harding, 1983; Brodeur *et al.*, 1982; del Ray *et al.*, 1982). In some cases these repetitive sequences have been identified as mobile transposable elements (Brodeur *et al.*, 1982; Gafner *et al.*, 1983; Eigel and Feldman, 1982).

The above examples of transfer RNA pseudogenes which share degenerate or incomplete homologies with the mature coding sequence indicates that their occurrence is a common phenomenon in eukaryotes. Although no human tRNA pseudogene has yet been reported, it is possible that the stretch of 17 nucleotides homologous to the 3' end of $tRNA^{Tyr}$ reported in this study is a candidate for human tRNA pseudogene. However, if the sequence homologous to $tRNA^{Tyr}$ represents a tRNA pseudogene (or truncated gene) then its relationship to the human Alu sequence warrants an explanation.

In the human genome, the predominant class of short interspersed, repetitive DNA is approximately 300 nucleotides long. The majority of these 300 bp repeated sequences have a common cleavage site for restriction endonuclease AluI and so are commonly referred to as Alu- like elements or Alu family. (Houck *et al.*, 1979; Rubin *et al.*, 1980; Schmid and Jelinek, 1982).

The Alu family is both abundant and widely dispersed throughout the human genome. It has been mapped in the neighbourhood of most structural genes isolated so far (Bell *et al.*, 1980; Barralle *et al.*, 1981; Duncan *et al.*, 1981; Denison *et al.*, 1981). Individual members of this family are readily recognised by their base sequence which resembles a single consensus sequences (Deininger *et al.*, 1981; Schmid and Jelinek, 1982). Each member is on the average about 13% divergent from this common consensus sequence (Deininger *et al.*, 1981). Structurally, the human Alu family members actually represent an imprecise head to tail dimer of two approximately 130 bp monomers (Figure 5.1). The two halves of the dimer contain about 70% homology to each other, with an additional 31 bp segment (insert) located within the right monomer unit. (Deininger *et al.*, 1981; Schmid and Jelinek, 1982; Ullu 1982). At the 3' end of each monomer is a stretch of 10-20 deoxyadenylic acid residues (Oligo dA tract). The human Alu family members are generally flanked on either side by direct repeats of 7 to 20 bp (Figure 5.1). The sequence of these flanking repeats is not conserved among different members, but is unique to each member. Sakaki *et al.*, (1983) have demonstrated that Alu family DNA sequences can form a mobile composite with another longer repetitive sequence element known as *KpnI*.

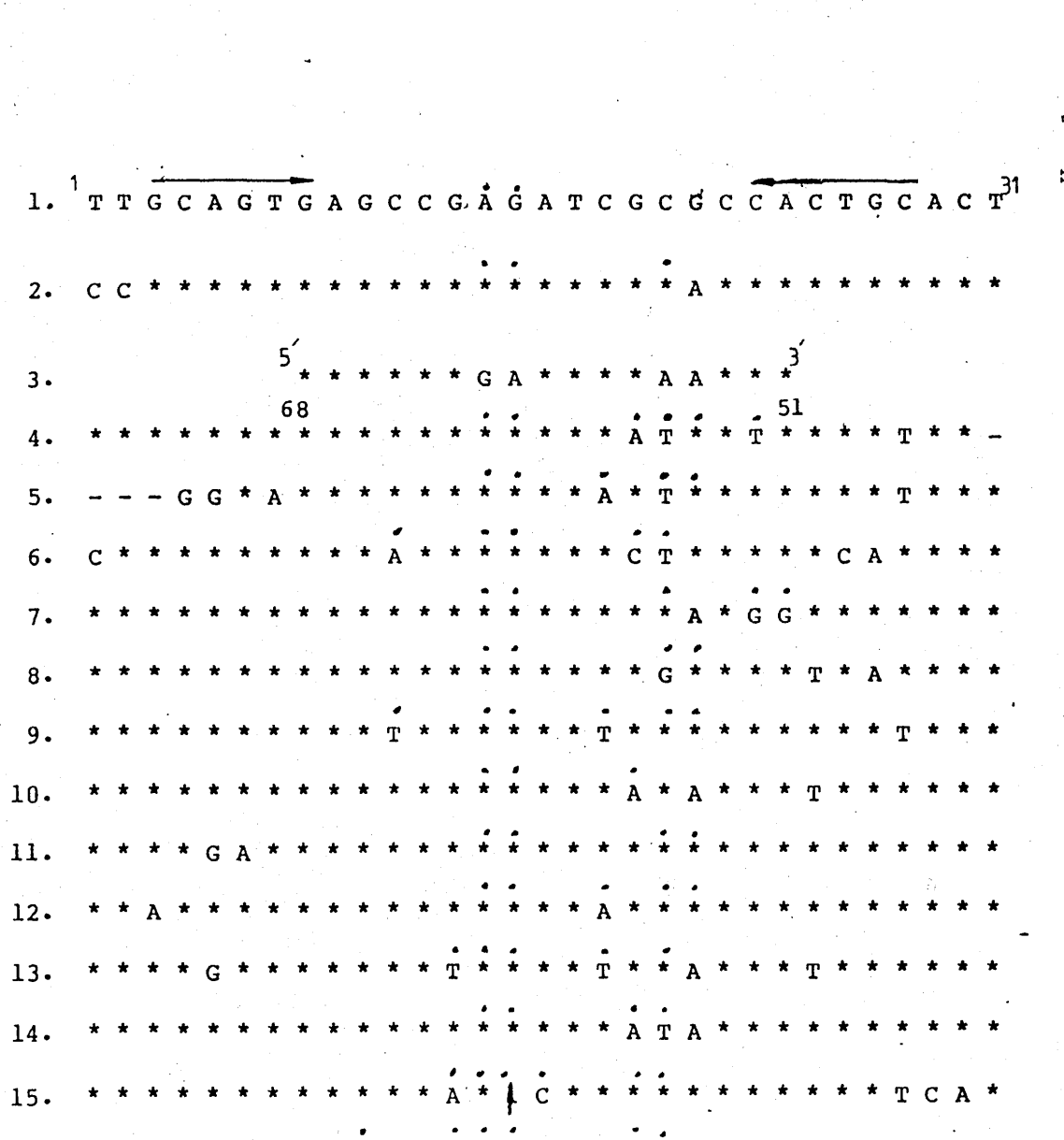
Alu equivalent families have been described in primates (Grimalidi and Singer, 1982) rodents (Haynes and Jelinek, 1981) and insects (Dover 1981). Ullu (1982) has reported Alu like sequences in the genome of slime moulds, echinoderms, amphibians and birds. A comparison of human Alu sequence with rodent and monkey Alu equivalents demonstrates a considerable sequence homology (Schmid and Jelinek, 1982). The dimeric structure appears to be conserved in primates, while the rodent Alu equivalent sequence is

organized as 130 bp (129 bp in mouse and 134 bp in Chinese hamster) monomeric sequence (Schmid and Jelinek, 1982). The rodent Alu seuqence also contains a 32 bp insert, which is different in nucleotide sequence from the human insert and is located nine nucleotides downstream from the insertion point of the human insert. The remainder of the rodent Alu sequence is approximately 80% homologous to human right monomer. This has led to the tenetative suggestion that the dimeric Alu sequence has originated by tandem duplication of an ancestral Alu monomer (Ulla 1982).

Some cloned human Alu sequences are efficiently transcribed *in vitro* by RNA polymerase III, producing discretely sized RNAs (Duncan *et al.*, 1981; Fritsch *et al.*, 1981; Di Segni *et al.*, 1981; Paolella *et al.*, 1983). Recently, Paolella and co-workers (1983) have demonstrated that the promoter of Alu sequences are organised in a bipartite structure, similar to eukaryotic tRNA split promoter (see Figure 1.5 section 1.4a iv). The RNA polymerase III promoter is present in the left half of the dimer, transcription starts close to the first nucleotide of the Alu sequence, proceeds through the entire sequence and terminates beyond the dAMP rich region. Although there is no direct evidence of specific *in vivo* transcription of Alu sequences, transcripts with homologies to Alu DNA are found in several categories of RNA. Alu sequences are well represented in polyadenylated heterogenous nuclear RNA (hnRNA) isolated from human cells (Elder *et al.*, 1981).

The location of Alu sequences down or upstream from functional genes, within introns (Della Fovira *et al.*, 1981), as part of messenger RNA and as a major constituent of 7S RNA, suggests that the Alu sequences fulfil a variety of functions in the cell. Their precise biological function is unclear, although it has been suggested that they might play a regulatory role in the

Figure 5.2 Comparison of the nucleotide sequence of the 31 bp insert of the human Alu consensus with inserts of other published human Alu sequences, the 17 bp of the X. Laevis $tRNA^{Tyr}$ sequence (coding strand) and the insert of African green monkey Alu-equivalent sequence. The top line is the consensus of human Alu repeats determined by Jelinek and Schmidt (1982). The second line represents the 31 bp insert present in the 640 bp SmaI-HindIII fragment of pNB1. The third line represents the 17 bp sequence of *X. laevis* $tRNA^{Tyr}$ (coding-strand), numbered according to the convention of numbering mature tRNAs (Gauss *et al.*, 1979). The lines 4 to 12 represent the BLUR series of human Alu family DNA clones isolated by Deininger et al., (1981). The line 13 represents the Alu family sequence present 5' to the human γ globin (foetal) gene (A36) (Duncan *et al.*, 1981). The line 14 represents Alu family sequence present 5' to human δ globin gene (R 3.1) (Duncan *et al.*, 1981). The line 15 represents the human DNA clone pJP53 reported by Pan *et al.*, (1981). The line 16 represents African green monkey Alu family sequence (Dhruva *et al.*, 1980). An asterisk (*) at any position indicates the same nucleotide as in the human Alu consensus sequence at the corresponding position. A dot (.) overlying indicates the difference with *X. laevis* $tRNA^{Tyr}$ sequence. A dash (-) indicates nucleotide not present. An arrow (↑) indicates an extra nucleotide. The straight overlying arrows (⟶) indicates the inverted repeats.



	Homology to Alu consensus	Homology to X. laevis tRNA^{Tyr}	
1.	100%	76%	Human Alu consensus sequen (non-coding str
2.	90%	82%	pNB 1 (640 bp fragment, nc coding strand)
3.			tRNA^{Tyr} (Coding-strand)
4.	87%	64%	BLUR 2 DNA
5.	80%	70%	BLUR 6 DNA
6.	80%	70%	BLUR 7 DNA
7.	90%	70%	BLUR 8 DNA
8.	90%	76%	BLUR 10 DNA
9.	90%	64%	BLUR 11 DNA
10.	90%	82%	BLUR 13 DNA
11.	93.5%	76%	BLUR 14 DNA
12.	93.5%	70%	BLUR 19 DNA
13.	83%	70%	A36 - DNA
14.	90%	76%	R 3.1 DNA
15.	83%	64%	pJP 53
16.	77%	64%	African green monkey
	87.7%	70%	

process of gene expression (Davidson and Britten, 1979). Jelinek and co-workers (1980) propose that Alu elements represent centres of chromosomal replication. This suggestion is based on their high homology with an inverted repeat sequence found at or near the origin of replication of several papovaviruses and Hepatitis B virus (Jelinek *et al.*, 1980). Alu sequences have been cited as examples of selfish DNA (parasitic DNA). This is a sequence whose only purpose within the genome is to ensure its own survival without playing any role in the host functions of the genome (Doolittle and Saprenza 1980; Orgel and Crick, 1980). Many other functional roles have been ascribed to Alu type elements, including modulators of chromatin structure (Duncan *et al.*, 1981), hot spots for recombination (Jagadeeswaren *et al.*, 1982) or catalysts of chromosome rearrangement (Calaberetta *et al.*, 1982). However, there is little evidence to substantiate any of these hypotheses.

In all the studies to date, there have been no reports on the role of the 31 bp insert found in the right monomer of human Alu DNA sequence. In Figure 5.2 the nucleotide sequence of 12 different published human Alu sequence inserts in compared. This analysis indicates that although there is high degree of conservation between the nucleotide sequence of different inserts (on average share 88% homology), only pNB 1 (line 2) and BLUR 13 (line 10) show homology greater than 80% to $tRNA^{Tyr}$ sequence (hybridization conditions used in this study require at least 80% homology between complementary sequences for duplex formation). This possibly explains the reason why only a single $tRNA^{Tyr}$ related recombinant (λht3) was obtained in screening approximately 20-25% of the human genome in this study. However, the presence of sequences homologous to the coding strand of

$tRNA^{Tyr}$ in the 31 bp insert of the 640 bp fragments of pNB1 is of interest and warrants an explanation. A close examination of this insert indicates the presence of repeated sequences (inverted repeats) (overlying arrows in Figure 5.2) located on both sides of the 17 bp region homologous to $tRNA^{Tyr}$. Inverted repeats are characteristics features of insertion elements. Thus, suggesting that the truncated $tRNA^{Tyr}$ sequences may have arisen by insertion and excision of transposable element into an intact gene, as has been proposed by Sharp *et al.,* (1981) in case of Drosophila $tRNA^{Met}$ pseudogene. Another possibility is that information of the mature tRNA may be "reverse transcribed" and integrated into the genome. Reilly *et al.*, (1982) have suggested a similar phenomenon for mouse $tRNA^{Phe}$ pseudogene. Alternatively the truncated $tRNA^{Tyr}$ sequences may be the remnants of the tRNA molecule from which the human Alu sequence is derived.

Recently, several mammalian repetitive sequence elements from diverse species have been reported that share significant homology with tRNA molecules. The following five groups of repetitive elements have been reported that contain tRNA-like sequences: i) the C family of artiodactyls (cows and goats) has a consensus sequence that is 65% homologous with mammalian glycine tRNA (Lawrence *et al.*, 1985; Sakamoto and Okada, 1985); 2) the *Galago* monomer consensus sequence (the Alu- like family of the bushbaby) shows 60% homology to human methionine (initiator) tRNA (Daniels and Deininger, 1985); 3) the rat R der 1 (brain identifier repetitive sequence element) sequence shows 64% homology with human phenylalanine tRNA and 73% homology to alanine tRNA isolated from *B. mori* (mammalian $tRNA^{Ala}$ has not yet been sequenced); (Lawrence *et al.*, 1985; Sakamoto and Okada, 1985, Daniels and Deininger, 1985, Roger 1985b); 4) the members of rodent

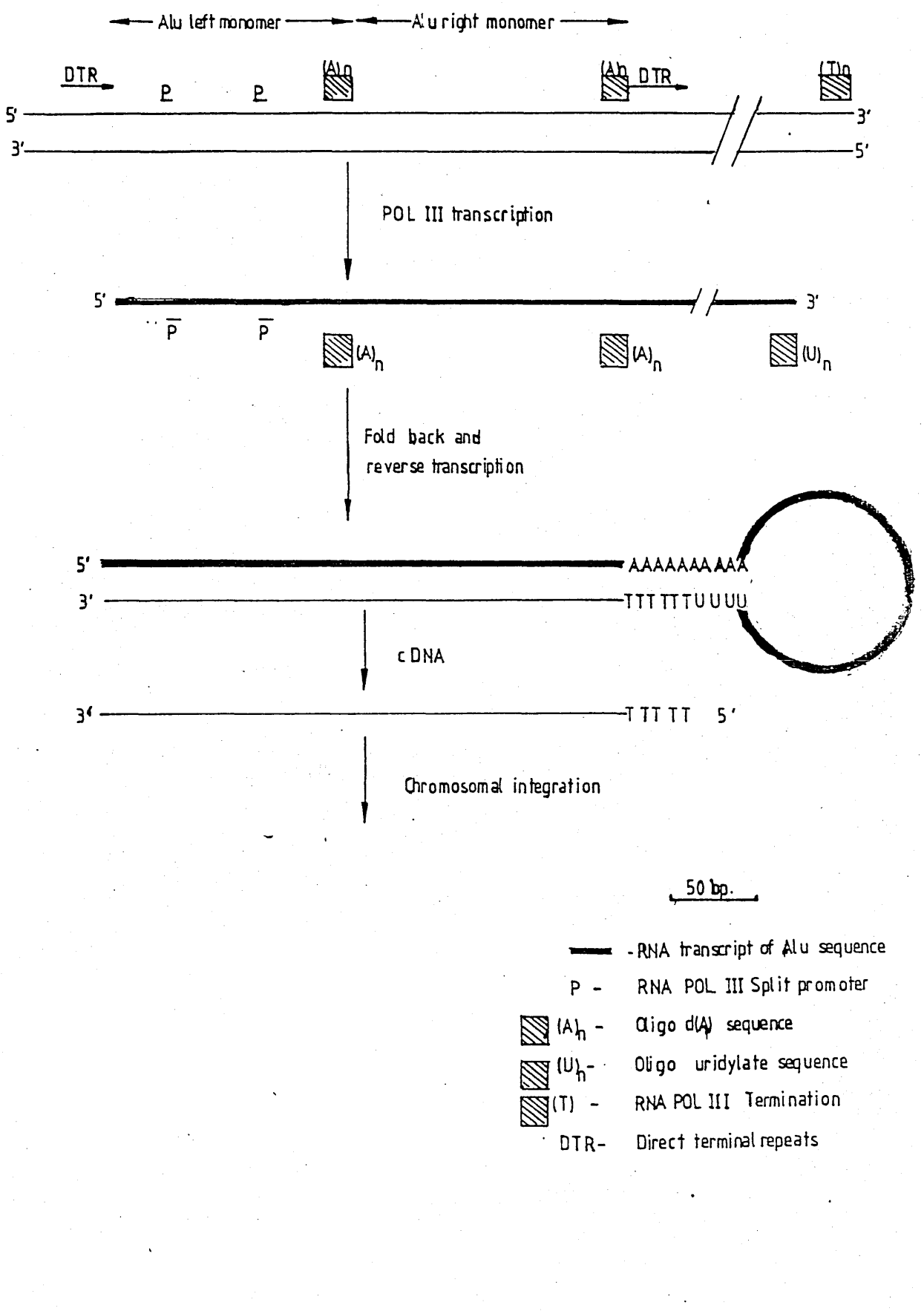
B2 family contains a tRNA like portion that is 60% homologous to rabbit lysine tRNA (Daniels and Deininger, 1985; Sakamotoand Okada, 1985, Lawrence *et al.*, 1985); 5) the rabbit C family (no relation to C family of artiodactyls) shows approximately 60% homology to human glycine tRNA (Sakamoto and Okada, 1985). In all these examples the regions of homology are not confined to the promoter regions, as one would expect to find conserved residues within the internal promoter regions of sequences transcribed by RNA polymerase III (Galli *et al.*, 1981, see section 1.4.a.iv). In addition to the sequence homology, the tRNA-like sequence within the repetitive elements can be folded into the clover leaf secondary structure model characteristic of tRNA molecule. For example, the consensus sequence of the C family of artiodactyls can in theory be folded up into a secondary structure with all the invariant residues, the four stems and loops in the normal position of the tRNA except for the amino acyl stem (Rogers 1985, Lawrence *et al.*, 1985). The anticodon of the folded structure corresponds to cysteine (mammalian tRNA for cysteine has not yet been sequenced). Similarly, the *Galago* monomer consensus sequence can be folded into clover leaf structure that coincides perfectly with human methionine tRNA except for aminoacyl stem (Daniels and Deininger 1985). However, the R. der. 1., B2 and Rabbit C family repetitive sequences contain tRNA like sequences which can be folded into secondary structures with less perfect anticodon and aminoacyl stems (Rogers 1985; Daniels and Deininger, 1985; Sakamoto and Okada, 1985; Lawrence *et al.*, 1985). The remarkable sequence homology and close structural resemblance of repetitive sequence to some specific tRNAs has led to the suggestion that tRNAs or their genes may be the evolutionary ancestors of repetitive sequences (Rogers 1985b; Lawrence *et al.*, 1985; Sakamoto and Okada, 1985; Daniels and Deininger 1985). The

observation that tRNA-derived sequences in the repetitive sequence elements lack the aminoacyl stem structure has led Sakamoto and Okada (1985) to suggest that these repeats may be generated by nonhomologous recombination between a tRNA gene and a tRNA-unrelated sequence block. The above findings suggest that the tRNA genes are progenitors of repetitive DNA families, and further substantiate the results of this study that the 17 bp region homologous to *X. laevis* $tRNA^{Tyr}$ contained in the human Alu family consensus sequence might be derived from authentic tRNA sequences.

Another possibility is that the tRNA-like sequences contained in the repetitive sequence elements may play a functional role in the transposition of repeat sequences. Although the exact molecular mechanism which has led to the abundance and interspersion of repetitive DNA sequences through the genome is yet unlear, some evidence suggest that these sequences are dispersed by means of RNA intermediates and reverse transcription (Van Arsdell *et al.*, 1981, Jagadeeswaran *et al.*, 1981, Hollis *et al.*, 1982, Sharp 1983). Such dispersed sequences created by reverse flow of genetic information from cellular RNA back into chromosomal DNA have been called retroposons by Rogers (1985a). These include several processed genes (intronless genes) (pseudogenes) that correspond to an integrated genomic copy of a spliced mRNA (Hollis *et al.*, 1982; Sharp 1983), some pseudogenes of non-messenger RNA, such as U series small nuclear RNA (U Sn RNA) (Van Arsdell *et al.*, 1981; Van Arsdell and Weiner, 1983; Bernstein *et al.*, 1983) and Alu-like elements (for a detailed review on retropsons see Rogers, 1985a).

As to the transposition of Alu sequences, Van Arsdell *et al.*, (1981) and Jagadeeswaran *et al.*, (1981) have proposed a model that involves the reverse transcription of an RNA polymerase III dependent RNA transcript of the

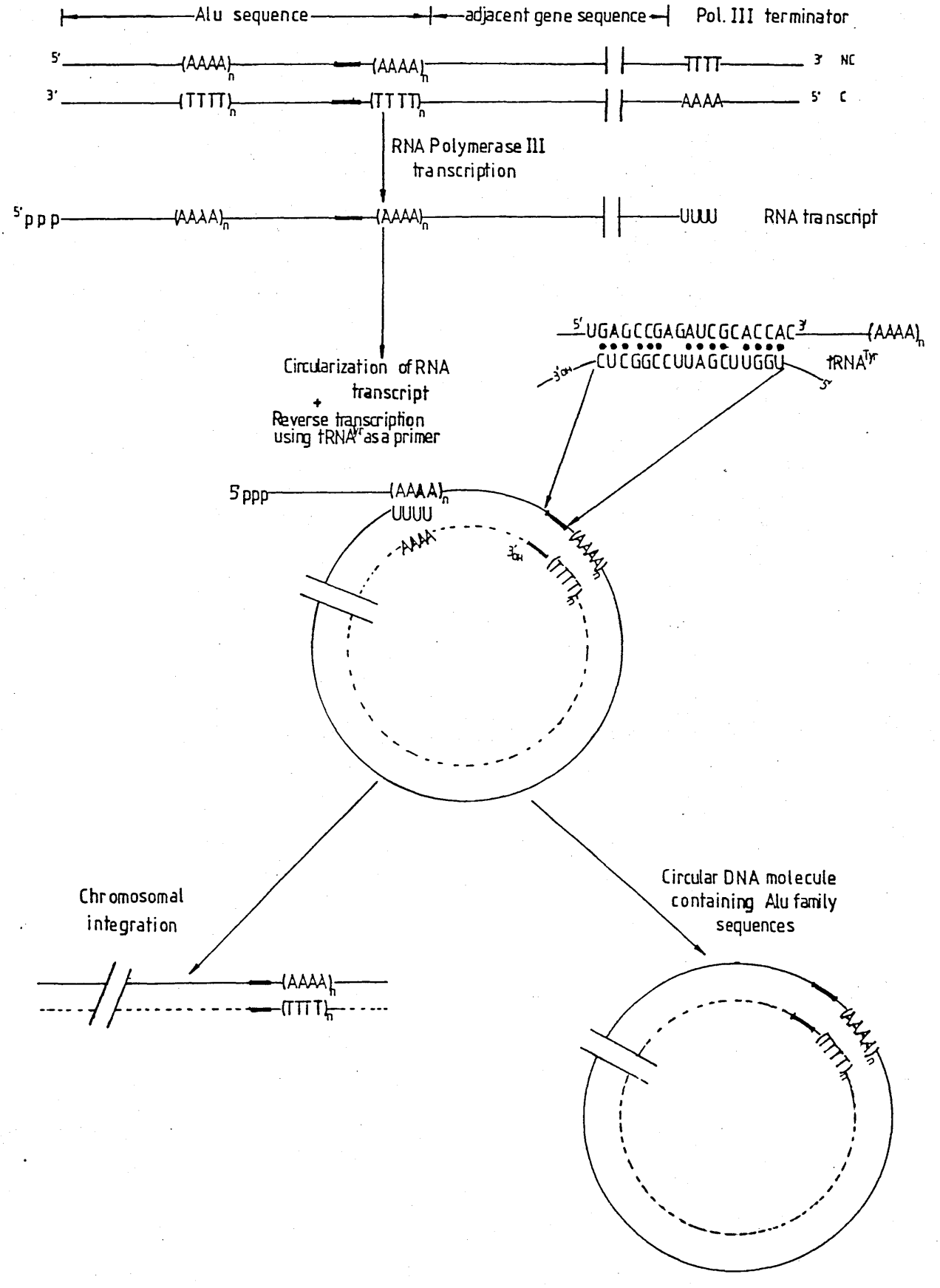
Figure 5.3 Model for the transposition of Alu sequence as proposed by Van Arsdell *et al.*, (1981) and Jagadeeswaran *et al.*, (1981).


Alu left monomer
Alu right monomer
DTR
P
P
(A)n
(A)n
DTR
(T)n
5'
3'
3'
5'
POL III transcription
5'
3'
P
P
(A)n
(A)n
(U)n
Fold back and
reverse transcription
5'
AAAAAAAAAA
3'
TTTTTTUUUU
c DNA
3'
TTTTT 5'
Chromosomal integration
50 bp.
- RNA transcript of Alu sequence
P - RNA POL III Split promoter
(A)n - Oligo d(A) sequence
(U)n - Oligo uridylate sequence
(T) - RNA POL III Termination
DTR - Direct terminal repeats

Alu sequence and the subsequent insertion of the complementary molecule in the genome. According to this model, the cDNA synthesis is primed by the base pairing of the terminal 3' d UMP sequences (termination region of RNA transcript of Alu sequence) with the internal 3' dAMP rich reigon (as shown in Figure 5.3). Thus Alu family members can probably undergo multiple rounds of transcription and reinsertion. This speculative model is based on studies of U3 Sn RNA pseudogenes, where a U3 Sn RNA molecule can fold into a secondary structure such that the 3' end of the molecule folds onto itself and provides the necessary priming site for reverse transcription (Bernstein *et al.*, 1983). However, it is possible that the tRNA like sequences contained in the Alu sequence might serve as a primer binding site for DNA synthesis, analogous to replication of retroviruses (Varmus 1983). The mechanism of replication of retroviruses is well established. It involves a tRNA primer and reverse transcription, the jumping of a short primed oligonucleotide from one end of the RNA genome to the other and subsequent initiation of the second strand at a particular site (For details of replication of retroviruses see Varmus, 1983). Various tRNA primers have been identified for the initiation of retrovirus DNA synthesis by reverse transcription, these include tRNA: Trp, Pro, and Lys. (Harada et al., 1979; Peter and Glover 1980). In all these cases the nucleotide sequence of the primer binding site is complementary to the 18 nucleotides present in the 3' end of the respective tRNA. These similarities suggest that the 17 bp region homologous to $tRNA^{Tyr}$ in the human insert of Alu sequence might also serve as a primer binding site for the initiation of DNA synthesis in the transposition of some Alu family members.

The model proposed in this study involves the circularization of the RNA polymerase III transcript of Alu family member and the subsequence

Figure 5.4 Model for transposition of Alu and adjoining sequences. The RNA polymerase III transcript of Alu sequence terminates in an oligouridylate region downstream from the Alu sequence. The oligouridylate residues base pair with the internal oligo (dA) region, resulting in a circular RNA transcript and the subsequent cDNA synthesis (dotted line) by reverse transcriptase is initiated by $tRNA^{Tyr}$ primer (bold line).


Alu sequence
adjacent gene sequence
Pol. III terminator
5'
(AAAA)n
(AAAA)n
TTTT
3' NC
3'
(TTTT)n
(TTTT)n
AAAA
5' C
RNA Polymerase III
transcription
5'ppp
(AAAA)n
(AAAA)n
UUUU
RNA transcript
5'UGAGCCGAGAUCGCACCAC3'
(AAAA)n
CUCGGCCUUAGCUUGGU
tRNATyr
3'OH
5'
Circularization of RNA
transcript
+
Reverse transcription
using tRNATyr as a primer
5'ppp
(AAAA)n
UUUU
AAAA
3'OH
(TTTT)n
(AAAA)n
Chromosomal
integration
(AAAA)n
(TTTT)n
Circular DNA molecule
containing Alu family
sequences
(AAAA)n
(TTTT)n

cDNA synthesis by reverse transcription initiated by $tRNA^{Tyr}$ (Figure 5.4). The termination of RNA polymerase III transcript of human Alu family members occurs at different positions and depends on the presence of a T-rich region downstream from the Alu sequence. Therefore the RNA transcript from different Alu family members may vary in length, containing a variable length of flanking non-repetitive DNA and end in an oligouridylate region. Most human Alu family DNA transcripts terminate within a hundred nucleotides of the repeated element (Haynes and Jelinek, 1981; Deiniger *et al.*, 1981, Duncan *et al.*, 1979, Pan *et al.*, 1981). Several longer RNA polymerase III transcripts ranging from 600 bp to 800 bp have been reported (Haynes and Jelinek, 1981; Duncan *et al.*, 1979). However some Alu family transcripts could be even longer and may contain adjacent functional gene(s) sequences. For instance, Manley and Colozoo (1982) have described a 2.3 Kb long RNA transcript from the Alu sequence in the neighbourhood of a human β-globin gene. According to the model in Figure 5.4, the terminal U-rich region of these long transcripts of Alu family members might base pair with the internal oligo (dA) region as shown in Figure 5.4. This could result in a circularized molecule and the cDNA synthesis initiated by a $tRNA^{Tyr}$ primer, where the 3' end is hydrogen-bonded to the 17 base sequence in the human insert of the Alu DNA sequence (Figure 5.4). The resultant DNA transcript could therefore comprise an Alu sequence fused with other DNA sequences corresponding to a hypothetical gene sequence. This model suggests an alternative means whereby Alu sequences could promote the transposition and rearrangement of other sequences in eukaryotic genome, as has been implied in many previous reports (Calaberetta *et al.*, 1982; Sharp 1983; Jagadeeswaran *et al.*, 1982; Schmid

and Jelinek 1982). Orkin and Michelson (1980) have reported one instance of such Alu sequence involvement in DNA sequence rearrangement in vivo.

In this case a portion of a human α globin gene is deleted and the remainder is fused to a portion of an Alu family member. The proposed model might also account for the existence of mammalian nuclear extrachromosomal circular DNA containing Alu family sequences (Calabretta *et al.*, 1982). This model presents several interesting possibilities with respect to movement of Alu family DNA, which require further examination.

REFERENCES

Abelson, J. (1979) Ann. Rev. Biochem. 48, 1035-1069.

Attardi, G. (1981) Trends in Biochemical Sciences 6, 86-89.

Attardi, G., Cantatore, P., Chomyn, A., Crews, S., Gelfand, R., Merkel, C., Montoya, J. and Ojala, D. (1982) "In Mitochondrial Genes". pp.51-71. Sloninmski, P., Brost, P., and Attardi, G. Eds. Cold Spring Harbor Laboratory, CSH New York.

Baker, R.E., Elgel, A., Voegtel, D. and Feldmann, H. (1982) The EMBO Journal 1, 291-295.

Baralle, F.E., Shoulder, C.C., Goodboum, S., Jeffreys, A. and Proudfoot, N.J. (1981) Nucl. Acids Res. 8, 4393-4404.

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

Bell, G.I., Picter, R. and Rutter, W.J. (1980) Nucl. Acids Res. 8, 4091-4109.

Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.

Bernstein, L.B., Mount, S.M. and Weiner, A.M. (1983) Cell 32, 461-472.

Blattner, F.R., Williams, B.G., Blechl, A.E., Thomson, K.D., Faber, H.E., Furlong, L.A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) Science 196, 161-169.

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

Bolivar, F., and Backman, K. (1979) In Methods in Enzymology 68, 245-267, Wu, R., Ed. (Academic Press, New York).

Bovre, K. and Syzbalski, N. (1971) In Methods in Enzymology 12, 350-383, Grossman, L. and Moldave, K., Eds. (Academic Press, New York).

Britten, R.J. and Davidson, E.H. (1979) Science 204, 1052-1059.

Broach, J.R., Friedman, L. and Sherman, F. (1981) J. Mol. Biol. 150, 375-387.

Brodeur, G.M., Sandymeyer, S.B., Olson, M.V. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 3292-3296.

Brown, D.D. and Gurdon, J.B. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 2849-2853.

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

Buckland, R.A., Cooke, H.J., Roy, K.L., Dahlberg, J.E. and Lund, E. (1983) Gene 22, 211-217.

Burdon, R.H., Martin, B.T., and Lal, B.M. (1967) J. Mol. Biol. 28, 357-371.

Byran, P.N., Hofstetter, M., and Birnstiel, M.L. (1981) Cell 27, 459-466.

Calabretta, B., Robberson, D.L., Berrera-Saldana, H.A., Lambrou, T.P. and Saunders, G.F. (1982) Nature 296, 219-225.

Carrara, G., Di Segni, G., Otsuka, A. and Tocchim-Valentini, G.P. (1981) Cell 27, 371-379.

Castagnoli, L., Ciliberto, G. and Cortese, R. (1982) Nucl. Acids Res. 4135-4145.

Chaleff, D. and Fink, G.R. (1980) Cell 21, 237-238.

Chang, S.H., Brun, G.K., Silberklang, M., RajBhandary, U.L., Hecker, L.I. and Barnett, W.E. (1976) Cell 9, 717-723.

Ciliberto, G., Castagnoli, L., Melton, D.A., and Cortese, R. (1982a) Proc. Natl. Acad. Sci. U.S.A. 79, 1195-1199.

Ciliberto, G., Rangei, G., Costanzo, F., Dente, L. and Cortese, R. (1983) Cell 725-733.

Ciliberto, G., Traboni, C., and Cortese, R. (1982b) Proc. Natl. Acad. Sci. U.S.A. 79, 1921-1925.

Clark, B.F.C. (1978) in Transfer RNA. Edited by Sidney Altman. MIT Press Cambridge, Massachusettes. pp. 14-47.

Clarke, L. and Carbon, J. (1976) Cell 9, 91-99.

Clarkson, S.G. (1983) , "Eukrayotic Genes: their structure, activity and regulation pp. 239-261. Eds. Maclean, N., Gregory, S.P. and Flavell, R.A. (Butterworths Press London).

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

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

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

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

Clarkson, S.G., Koski, R.A., Corlet, J. and Hipskind, R.A., (1981) In "Developmental Biology using Purified Gene", ICN-UCCLA Symposia on Molecular and Cellular Biology, Vol. XXIII (Eds. D.D. Brown and C.F. Fox), pp. 463-472; Academic Press, N.W.

Cortese, R., Melton, D., Tranquilla, T. and Smith, J.D. (1978) Nucl. Acids. Res. 5, 4593-4611.

Crampi, M.S., Melton, D.A. and Coretese, R. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 1195-1199.

Crick, F.H.C., Griffith, J. and Orgel, L. (1957) Proc. Natl. Acad. Sci. 43, 416-427.

Dagert, M. and Ehrlich, S.D. (1979) Gene 6, 23-28.

Daniels, G.R. and Deininger, P.L. (1985) Nature 317, 819-824.

Daniels, G.R., and Deininger, P.L. (1983) Nucl. Acids Res. 11, 7595-7610.

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

De Franco, D., Sharp, S. and Söll, D. (1981) J. Biol. Chem. 256, 12424-12429.

De Robertis, E.M., Black, P. and Nishikura, K. (1981) Cell 23, 89-93.

De Robertis, E.M. and Olson, M.V. (1979) Nature 278, 137-143.

Del Rey, F.J., Dohanhue, T.F., and Fink, G.R. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 4138-4142.

Della Favira, R.D., Gelmann, E.P., Gallo, R.C. and Wong-Staal (1981) Nature 292, 31 - 32

Denninger, P.L., Jolly, D.J., Rubin, C.M.. Friedmann, T. and Schmid, C.W. (1981) J. Mol. Biol. 151, 17-33.

Denison, R.A., Van Arsdell, S.W., Berstein, L.B. and Weiner, A.M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 810-814.

Dente, L., Fasano, O., Costanzo, F., Traboni, G., Aliberto, G. and Cortese, R. (1982) EMBO J. 1, 817-820.

Dhruva, B.R., Shenk, T., Subramanian, K.N. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 4514-4518.

Dirheimer, G., Keith, G., Sibler, A.P. and Martin, R.P. (1979) In Transfer RNA: Structure, Properties and Recognition. pp. 19-41 ed. Schimmel. Söll, D. and Abelson, J.N. Eds. Cold Spring Harbour Laboratory, CHS, New York.

Di Segni, G., Carrara, G., Tocchini-Valentini, G., Shoulders, C.C. and Baralle, F.E. (1981) Nucl. Acids Res. 9, 6709-6722.

DiGiovanni, L., Haynes, R.S., Misra, R. and Jelinek, W.R. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 6533-6537.

Dingerman, T., Sharp, S., Appel, B., De Franco, D., Mount, S., Heiermann, R.R., Pongs, O., and Söll, D. (1981) Nucl. Acid Res. 2, 3907-3918.

Dingerman, T., Burker, D.J., Sharp, S., Schaack, J. and Söll, D. (1982) J. Biol. Chem. 257, 14738-14744.

Doolittle, W.F. and Saprenza, C. (1980) Nature 284, 601-603.

Duggleby, R.G., Kinns, H. and Rood, J.I. (1981) Analytical Biochemistry 110, 49-55.

Duncan, C., Biro, P.A., Choudary, P.V., Elder, J.T., Wang, R.R.C., Forget, B.G., De Riel, J.K. and Weissmann, S.M. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 5095-5099.

Duncan, C.M., Jagadeeswaran, P., Wang, R.R.C. and Weissman, S.M. (1981) Gene 13, 185-196.

Eigel, A. and Feldmann, H. (1982) EMBO J. 1, 1245-1250.

Dover, G.A. (1981) Nature 285, 618-620.

Elder, J.T., Pan, J., Duncan, C.H. and Weissmann, S.M. (1981) Nucl. Acids. Res. 9, 1171-1189.

Elder, R., St. John, T.P., Stincomb, D.T. and Davis, R.W. (1980) Cold Spring Harbor Symp. Quant. Biol. 45, 581-591.

Elder, R., Szabo, and Uhlenbeck, O. (1980) In Transfer RNA: Biological Aspects: pp.317-323, Söll, D., Abelson, J.N., and Schimmel, P.R., Eds. Cold Spring Harbour Laboratory, CHS, New York.

Engelke, D.R., Ng, S-Y., Shastry, B.S. and Roeder, R.G. (1980) Cell 19, 717-728.

Errede, B., Cardillo, T.S., Sherman, F., Duvois, E., Deschamp, J. and Wiame, J.H. (1980) Cell 25, 427-436.

Etcheveryy, T., Colby, D. and Guthrie, C. (1979) Cell 18, 11-26.

Farabaugh, P.J. and Fink, G.R. (1980) Nature 286, 352-356.

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

Filipowicz, W. and Shatkin, A.J. (1983) Cell Vol. 32, 547-557.

Finkelstein, M. and Rownd, R.H. (1978) Plasmid 1, 557-562.

Folk, R.W., Hofstetter, H. (1983) Cell 33, 585-593.

Folk, W.R., Hofstetter, H., and Birnstiel, M.L. (1982) Nucl. Acids Res. 10, 7153-7162.

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

Fritsch, E.F., Lawn, R.W. and Maniatis, T. (1980) Cell 19, 959-972.

Fukada, K., Gossen, L., and Abelson, J. (1980) J. Mol. Biol. 137, 213-234.

Gafner, J. and Philippsen, P. (1980) Nature 286, 414-418.

Gafner, J., De Robertis, M.E. and Phillippsen, P. (1983) EMBO J. 2, 583-591.

Gall, J.G., and Pardue, M.L. (1969 Proc. Natl. Acad. Sci. U.S.A. 63, 378-382.

Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) Nature 294, 626-631.

Garber, R.L. and Gage, L.P. (1979) Cell 18, 817-828.

Gauss, D.H. and Sprinzl, M. (1982) Nucl. Acids Res. 10, 71-23.

Gilbert, W. (1978) Nature 271, 501.

Goddard, J.P. (personal communication)

Goddard, J.P., Squire, M., Bienz, M. and Smith, J.D. (1983) Nucl. Acid. Res. 11, 2551-2562.

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

Gauss, D.H. and Sprinzl, M. (1983) Nucl. Acids Res. 11, r1-r103.

Gauss, D.H., Grutter, F., and Sprinzl, M. (1979) Nucl Acids. Res. 6, r1-r19.

Grimaldi, G., Singer, M.F. Proc. Natl. Acad. Sci. U.S.A. (1982) 79, 1497-1500.

Guilfoyle, R. and Weinmann, R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 3378-3382.

Hagenbüchle, O., Larson, D., Hall, G.I. and Sprague, K.U. (1979) Cell 18, 1217-1229.

Hall, B.D., Clarkson, S.G. and Tocchini-Valentini, G. (1982) Cell 29, 3-5.

Hall, J.H. and Harding, J.D. (1982) Nucl. Acid. Res. 10, 4891-4900.

Hans, J.H. and Harding, J.D. (1983) Nucl. Acids. Res. 11, 2053-2059.

Harada, F., Peters, G.G. and Dehlberg, J.E. (1979) J. Biol. Chem. 254, 10979-10985.

Hattlen, L. and Attardi, G. (1971) J. Mol. Biol. 56, 535-553.

Hawthorne, D.C. (1976) Biochimie 58, 179-182.

Haynes, S.R. and Jelinek, W.R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 6130-6134.

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

Hoagland, M.B. (1960) in Nucleic Acids Vol. 3. E. Chargaff and J.N. Davidson Eds. (New York, Academic Press) pp. 349-408.

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

Hoagland, M.B., Zamecnik, P.C., and Stephenson, M.C. (1957) Biochem. Biophys. Acta 24, 215.

Hofstetter, H., Kressmann, A. and Birnstiel, M.L. (1981) Cell 24, 573-583.

Hohn, B. (1979) In Methods in Enzymology 68, 299-309, Wu, R., Ed. (Academic Press, New York).

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

Hollis, G.F., Heiter, P.A., McBride, O.W., Swan, D., and Leder, P. (1982) Nature 296, 321-325.

Honda, B.M. and Roeder, R.G. (1980) Cell 22, 119-126.

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

Houck, M.C., Rinehart, F.P. and Schmid, C.W. (1979) J. Mol. Biol. 132, 289-306.

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

Hu, J.C., Cote, B.D., Lund, E. and Dahlberg, J.E. (1983) Nucl. Acid Res. 11, 4809-4821.

Indik, and Tartof (1982) Nucl. Acid. Res. 10, 4159-4172.

Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) Cell 26, 141-142.

Jagadeeswaran, P., Tuan, D., Forget, B.G. and Weissman, S.M. (1982) Nature 296, 469-470.

Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.M., Biro, P.A., Choudary, R.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Delininger, P.L., and Schmid, C. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 1398-1402.

John, M.A., Birnsteil, M.L. and Jones, K.W. (1969) Nature 223, 582-584.

Johnson, G.D., Pirtle, I.L. and Pirtle, R.M. (1985) Archives of Biochemistry and Biophysics 236, 448-453.

Johnson, J.D., Ogden, R., Johnson, P., Abelson, T., Dembeck, P. and Hakura, K. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2564-2568.

Kim, S.H. (1979) in Transfer RNA: Structure, Properties and Recognition pp. 83-100, Schimmel, P.R., Soll, D., and Abelson, J.N., Eds. Cold Spring Harbour Laboratory, CHS New York.

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

Korn, L.J. (1982) Nature 295, 101-105.

Korn, L.J. and Brown, D.D. (1978) Cell 15, 1145-1156.

Koski, R.A. and Clarkson, S.G. (1982) J. Biol. Chem. 257, 4514-4521.

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

Kressman, A., Hofstetter, H., Di Capua, E., Grosschedl, R. and Birnstel, M.L. (1979) Nucl. Acids Res. 7, 1749-1763.

Kubli, E. (1981) Experientia 37, 1-9.

Kubli, E. and Schmidt, T. (1978) Nucl. Acids Res. 5, 1465-1470.

Laski, F.A. Alzner-DeWeerd, B., RajBhandary, U.L. and Sharp, P.A. (1982) Nucl. Acids. Res. 10, 4609-4626.

Lasser-Weiss, M., Bawnik, N., Rosen, A., Sand, S., and Daniel, V. (1981) Nucl. Acid. Res. 9, 5965-5978.

Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) Cell 15, 1157-1174.

Lawrence, C.B., McDonnell, D.P. and Ramsey, W.J. (1985) Nucl. Acids Res. 13, 4239-4249.

Lemischka, I. and Sharp, P.A. (1982) Nature 300, 330-335.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory (CSH, New York).

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

Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Gefter, M.L. (1980) Proc. Natl. Acad. Sci. U.S.A. &2, 3855-3860.

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

Marzloff, W.F., White, E.L., Benjamin, R. nd Huang, R.C.C. (1975) Biochemistry 14, 3715-3724.

Mattoccia, E., Baldi, M.I., Carrara, G., Fruscolinin, P., Benedetti, P., and Tochini-Valentini, G.P. (1979) Cell 18, 643-648.

Mattoccia, E., Baldi, M.I., Pande, G., Ogden, R. and Tocchini-Valentini, G.P. (1983) Cell 32, 67-76.

Maxam, A.M. and Gilber, W. (1980) In Methods in Enzymolqy, 65, 498-560.

Melton, D.A. and Cortese, R. (1979) Cell 18, 1165-1172.

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

Montgomery, D.L., Boss, J.M., McAndrew, S.J., Marr, L., Walthall, D.A. and Zitomer, R.S. (1982) J. Biol. Chem. 257, 7756-7761.

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

Ng, R. and Abelson, J.N. (1980) Proc. Natl. acad. Sci. U.S.A. 77, 3912-3916.

Nishikura, K. De Robertis, E.M. (1981) In International Cell Biolqy pp.60-65, Schweiger, H.G. Ed. Springer-Verlaq.

Nishikura, K., Kurjan, J., Hall, B.D. and De Robertis, E.M. (1982) The EMBO Journal 1, 263-268.

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

Ogden, R.C. Knapp, G., Peebels, C.L., Kang, H.S. Beckmann, J.S. Johnson, P.F., Fuhraman, S.H. and Abelson, J.N. (1980) In Transfer RNA: Biological Aspects, pp.173-190 Ed. Söll, D. Abelson, J.N. and Schimmel, P.R. Eds. Cold Spring Harbor, CHS, New York.

Olah, J. and Feldman, H. (1980) Nucl. Acids. Res. 8, 1975-86.

Olson, M.V., Loughney, K., Hall, B.D. (1979) J. Mol. Biol. 132, 387-410.

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.

Ono, M. and Ohisha, H. (1983) Nucl. Acids. Res. 11, 7169-7179.

Orgel, L.E. and Crick, F.H.C. (1980) Nature 284, 604-607.

Orkin, S.H. and Michelson, A., (1980) Nature 286, 538-540.

Page, G.S. and Hall, B.D. (1981) Nucl. Acids Res. 9, 921-934.

Pan, J., Elder, J.T., Duncan, C.H. and Weissmann, S.M. (1981) Nucl. Acids. Res. 9, 1151-1170.

Paolella, G., Lucero, M.A., Murphy, M.H., Barelle, F.E. (1983) EMBO. J. 2, 691-695.

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

Peffley, D.M. and Sogin, M.L. (1981) Biochemistry 20, 4015-4021.

Pelham, H.R.B. and Brown, D.D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 4170-4174.

Peters, G. and Glover, C. (1980) J. Virology 35, 31-40.

Reilly, J.G., Ogden, R., Rossi, J.J. (1982) Nature 300, 287-289.

Rich, A. and RajBhandary, U.L. (1976) Ann. Rev. Biochem. 45, 805.

Robbins, J., Rosteck, P., Haynes, J.R., Freyer, G., Cleary, M.L., Katler, H.D., Smith, K., and Lingrel, J.B. (1979) J. Biol. Chem 254, 6187-6195.

Robertus, J.D., Ladner, J.E., Finch., J.T., Rodes, D., Brown, R.S., Clark, B.F.C. and Klug, A., (1974). Nature 250, 546-551.

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

Rodi, D.J., Hass, R. and Grunberger, D. (1983) "The Genomic Organization of a Rat tRNA Gene Cluster". In International tRNA Workshop, 1983 Hakone, Japan.

Roeder, G.S. and Fink. G.R. (1980) 21, 239-249.

Rogers, J.H. (1985a) International Reviews of cytology 93, 231-279.

Rogers, J.H. (1985b) Nature 317, 765-766.

Rosa, M.D., Gottlieb, E., Lerber, M.R. and Steitz, J.A. (1981) Mol. Cell Biol 1, 785-796.

Rosen, A., Bawnik, N., Sarid, S., Beckmann, J.S. and Daniel, V. (1983) "Structure and Organization of Rat and Petunia tRNA Genes. In International tRNA Workshop, 1983 Hakone, Japan.

Rosen, A., Sarid, S. and Daniels, V. (1984) Nucl. Acid. Res. 12, 4893-40906.

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

Roy, K.L., Cooke, H., Buckland (1982) Nucl. Acid. Res. 10, 7313-7321.

Rubin, C.M., Houck, M.C., Deininger, P.L., Friedmann, T., and Schmid, C.W. (1980) Nature 284, 372-375.

Sakamoto, K. and Okada, N. (1985) J. Mol. Evol. 22, (in press)

Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980) Cell 19, 13-25.

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

Sandmeyer, S.B. and Olsen, M.V. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7674-7678.

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

Schaack, J., Sharp, S., Dingermann, T. and Söll, D. (1983) J. Biol. Chem. 258, 2447-2458.

Schmid, C.W. and Jelinek, W.R. (1982) Science 216, 1065-1070.

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

Schmidt, O., Mao, J., Silverman, S., Hovemann, B. and Söll, D. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 4819-4823.

Schweizer, E., MacKennie, C., Halvorsen, H.O. (1969) J.M. Biol. 40, 261-77.

Segall, J. Matsui, T. nd Roeder, R.G. (1980) J. Biol. Chem. 255, 11986-11991.

Sekiya, T., Nishizawa, R., Matsuda, K., Taya, Y., and Nishimura, (1982) Nucl. Acids. Res. 10, 6411-6419.

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

Sharp, P.A. (1983) Nature 301, 471-472.

Sharp, S., De Fanco, D., Silberlang, M., Hosbach, H.A., Schmidt, T., Kubli, E., Gergen, J.P., Wensink, P.C., Söll, D. (1981) Nucl. Acids Res. 9, 5867-5882.

Sharp, S., De Franco, D., Silberklang, M., Hosbach, H.A., Schmidt, T., Kubli, E., Gergen, J.P., Wensink, P.C. and Söll, D. (1981b) Nucl. Acids Res. 2, 5867-5882.

Sharp, S., De Franco, D., Dingermann, T., Farrel, P. and Söll, D. (1981a) Proc. Natl. Acad. Sci. U.S.A. 78, 6657-6661.

Sharp, S., Dingermann, T, Schaack, J., DeFranco, D. and Söll, D. (1983) J. Biol. Chem. 258, 2440-2446.

Sharp, S., Dingermann, T., and Söll, (1982) Nucl. Acids. Res. 10, 5393-5406.

Shibuya, K., Noguchi, S., Nishimura, S., and Sekiya (1982) Nucl. Acids Res. 10, 4441-4449.

Shibuya, K., Sekiya, T., Noguchi, S., Yamaki, M. and Nishimura, S. (1983) "Transfer RNA pseudogenes in the Rat Chromosome". In International tRNA Workshop, 1983 Hakone, Japan.

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

Simsek, M., Ziegenmeyer, J., Heckman, J.,m and Rajbhandary, U.L. (1973) Proc. Nat. Acad. Sci. U.S.A. 70, 1041-1043.

Singhal, R.P. and Fallis, P.A.M. (1979) In Progress in Nucleic Acid Research and Molecular Biology 23, 227-289, Davidson, J.N. and Cohn, W.E. Eds., (Academic Press, New York).

Smith, O.H. and Birnstiel, M.L. (1976) Nucl. Acids. Res. 9, 2387-2399.

Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.

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

Sprinzl, M., Vorderwulbecke, T., and Hartman, T. (1985) Nucl. Acids. Res. 13, r51-r104.

Staden, R. (1977) Nucl. Acid. Res. 4, 4037-4051.

Staden, R. (1979) Nucl. Acids. Res. 6, 2601-2610.

Staden, R. (1980) Nucl. Acids. Res. 8, 817-825.

Staden. R. (1982) Nucl. Acids. Res. 10, 2951-2961.

Standring, D.N., Venegas, A. and Rutter, W.J. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 5963-5967.

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

Suddath, F.L., Quiglely, J., McPherson, A., Sneden, D., Kim, J., Kim, S.H. and Rich, A. (1974) Nature 248, 20-24.

Sulston, J.E. and Brenner, S. (1974) Genetics, 77, 95-104.

Sutcliffe, J.G. (1978) Nucl. Acids. Res. 5, 2721-2728.

Telford, J.L., Kressmann, A., Koski, R.A., Grosschedl, R., Muller, F., Clarkson, S.G. and Birnstiel, M.L. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 2590-2594.

Thomas, P.S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5201-5205.

Traboni, C., Ciliberto, G. and Cortese, R. (1982) EMBO J. 1, 415-420.

Twigg, A.J. and Sherratt, D. (1980) Nature 283, 216-218.

Ullu, E. (1982) TIBS June 1982.

Ullu, E. and Tschudi, C. (1984) Nature 312, 171-172.

Ullv, E., Murphy, S., Melli, M. (1982) Cell 29, 195.

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

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

Van Arsdell, S.W. and Weiner, A.M. (1984) Nucl. Acids. Res. 12, 1463-1469.

Van Arsdell, S.W., Denison, R.A., Bernstein, L.B. and Weiner, A.W. (1981) Cell 26 (part 1) 11-17.

Van Arsdell, S.W., Denison, R.A., Bernstein, L.B., Weiner, A.M., Manser, T. and Gesteland, R.F. (1981) Cell 26, 11-17.

Varmus, H.E. (1982) Science 216, 812-820.

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

Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 3683-3687.

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

Weil, P.A., Segall, J., Harris, B., Ng, S.Y. and Roeder (1979) J. Biol. Chem. 254, 6163-6173.

Weil, P.A., Segall, J., Harris, B., Ng, S-Y. and Roeder, R.G. (1979) J. Biol. Chem. 254, 6163-6173.

Weiner, A.M. (1980) Cell 22, 209-218.

Wittig, B. and Wittig, S. (1979) Cell 18, 1173-1183.

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

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

Zachau, H.G. (1972) In the mechanism of protein synthesis and its regulation, pp 173-218, Bosch, L. Ed., North Holland, Amsterdam.

Zasloff, M., Santos, T., Romeo, P. and Rosenberg, M. (1982) J. Biol. Chem 257, 7857-7863.