Organization and *in vitro* expression of human $tRNA^{Glu}$ genes.

A Thesis Presented for the

Degree of

Doctor of Philosophy

by

Efstathios S. Gonos

Department of Biochemistry

University of Glasgow

... ποτε θ' ακουσουμε τους χτυπους μιας καρδιας

ποτε το βλεμμα απεναντι θα σκιαζει το δεν πρεπει

και θα πλανιεται αγερωχο δακρυζοντας το φοβο

ηδονικα παντοτινα ζηταμε

ματαια παντα ...

To those who dare to look

beyond their existence.

Summary

Prior to the start of this work several recombinants containing tRNA genes were selected from a bank of 15-20 kb fragments of a human foetal DNA, cloned in λ Charon 4A, by screening with a mixed placentae tRNA[^{32}P]pCp probe (Goddard *et al.*, 1983). Two of these recombinants, λht137 and λht190, generated restriction fragment profiles which were very similar but not identical. Two common 2.4 kb HindIII fragments which hybridized the probe used, were subcloned from λht137 and λht190 into the vector pAT153 to generate recombinants pLB4 and pTC51 respectively. The characterized recombinant pLB4 was found to contain a single tRNA^{Glu} gene, notable for its high transcriptional activity and the potential of its 5'-flanking sequence for forming a tRNA-like structure (Goddard *et al.*, 1983).

The sequence of 940 base pairs of human DNA in the pTC51 was determined. This sequence showed 97.3% homology to the sequence already known from plasmid pLB4 and the presence of a second human tRNA^{Glu} gene copy. This copy differs from the first gene copy in two nucleotides which are not invariant or semi-invariant , or part of the obligatory intragenic control regions. The second characterized tRNA^{Glu} gene, also has a potential tRNA-like structure in its 5'-flanking sequence and is transcribed similarly to the first gene copy after using HeLa S3 cell free extracts. Restriction site mapping of recombinants λht137 and λht190 showed that the two characterized tRNA^{Glu} genes belong to highly homologous DNA fragments of at least 10.7 kb, in the human genome.

Hybridization studies using human placental genomic DNA suggest that the copy number for the tRNA^{Glu} gene family is thirteen. Nine of these copies appear to belong to highly conserved 2.4 kb HindIII fragment repeats in the human genome. Additional restriction mapping analysis in 8-10 kb genomic DNA regions, containing tRNA^{Glu} genes, show that the organization of the copies in the human genome varies in a way which may allow them to be divided into three inter-related subgroups, of 3-4 copies each.

The role of the tRNA-like structure in the 5'-flanking sequence on transcription of a human tRNA^{Glu} gene was also investigated. Several deletion mutants of

recombinant MtGlu6 (a M13mp9 recombinant containing the $tRNA^{Glu}$ gene from λht137 on a 481 bp fragment), which lacked part or all of the tRNA-like structure, were characterized. Optimal conditions for transcription were established using MtGlu6 and these conditions were used to compare the transcriptional efficiency of MtGlu6 and its deletion mutants. The transcriptional efficiency of four mutants was the same as the "wild type" MtGlu6, two mutants had decreased transcriptional efficiency, one was more efficient and one lacking part of the 5' intragenic control region, was inactive. Correlation of the transcriptional efficiencies with the position and the size of the 5'-flanking sequence deleted, indicated that the tRNA-like structure may be deleted without loss of transcriptional efficiency. Current models for the modulation of tRNA gene transcription by the 5'-flanking sequence were assessed in the light of the results obtained and a potential model presented.

Acknowledgements

I would like to thank Prof. M.D. Houslay and the late Prof. R.M.S. Smellie for providing the facilities for this research.

I am extremely grateful to Dr. J.P. Goddard for his friendly supervision, persistent encouragement and constructive criticism during the course of this work.

I am thankful to Dr. R.L.P. Adams and Dr. G.I. Jenkins for their helpful discussion and advice.

I would also like to thank all members of Lab C30, and in particular Tom Carr, for their invaluable assistance, friendship and for introducing me to the "English" language.

I express my gratitude to the University of Glasgow for the financial support.

Most of all I am deeply indebted to my parents for their endless support and encouragement throughout all these years.

Finally I express my deepest emotions to Letizia Pascalino for the moral support, patience and love she willingly offered me in good and bad days.

Contents

List of tables

List of figures

Abbreviations

The abbreviations used in this thesis are those suggested in the Instruction to Authors of the Biochemical Journal (1989), with the following additions :

tRNA	transfer RNA
tDNA	genes for transfer RNA
mRNA	messenger RNA
VA RNA	virus-associated RNA
SV 40 DNA	simian virus 40 DNA
dNTP	deoxynucleoside-5'-triphosphate
ddNTP	dideoxynucleoside-5'-triphosphate
ICR	internal control region
TF	transcription factor
pCp	cytidine 3'-5'- biphosphate
D	dihydrouridine
ψ	pseudouridine
nt	nucleotide
pfu	plaque forming unit
rpm	revolution per minute
u.v.	ultra-violet
IPTG	isopropyl-β-D-thio-galactopyranoside
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside
PBS	phosphate buffer saline
SDS	sodium dodecyl sulphate
DMSO	dimethyl sulphoxide
DTT	dithiothreitol
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid
TEMED	N,N,N',N',-tetramethylene diamine

CHAPTER 1

INTRODUCTION

1.1. General introduction.

Transfer RNAs (tRNAs) are small molecules (23,000-28,000 daltons) which account for some 10-15% of total cellular RNA. The major function of tRNA is to serve as the adaptor molecule during translation of mRNA into protein, although tRNA is also involved in other cell functions (e.g. primers for synthesis of retrovirus DNA by reverse transcriptase, or donors of their charged amino acids in the formation of some cell walls). They were first discovered by Hoagland *et al.*, (1958), following a pioneering prediction by Crick in 1955 who had proposed : "prior to incorporation into proteins, amino acids are first attached to specific adaptor molecules, which in turn possess unique surfaces that can bind specifically to bases on RNA templates". However the nucleotide sequence of the first tRNA molecule (a yeast alanine tRNA) was not determined until 1965 by Holley *et al.* This work was recognized by the award of the Nobel Prize in 1968.

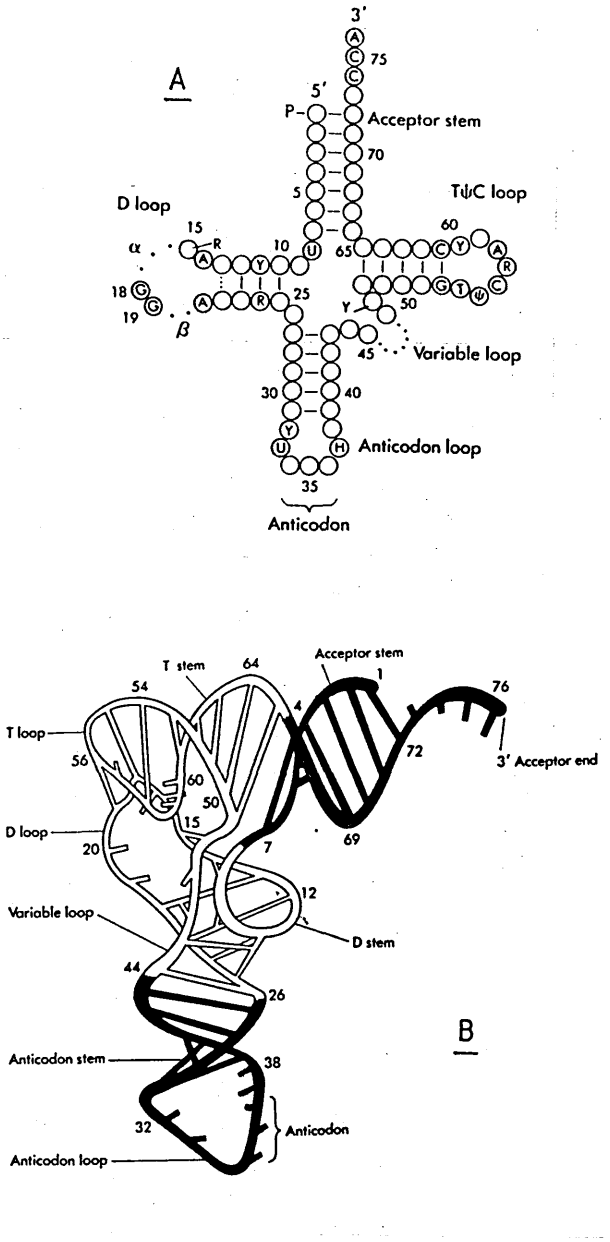
Transfer RNA provided a stimulus model during mid-1960s and 1970s for studying protein nucleic acids recognition. In those years the nucleotide sequences of several hundred tRNAs were determined and general rules concerning their secondary structure have emerged since then (see section 1.2.). In addition other studies identified the tertiary structure of tRNA molecules, opening the way to understanding the formation of tRNA aminoacyl-tRNA synthetase interactions (see section 1.2.). The astonishing developments in recombinant DNA technology over the last twelve years have substantially facilitated the study of tRNA genes. From

such studies, assessments about the gene structure emerged (see section 1.3.), as well as about the organization and evolution of tRNA genes (see section 1.4.). In addition the availability of transcription systems in conjunction with other techniques, made feasible the direct comparison between tRNA gene primary transcripts and tRNAs themselves (see section 1.5.). Nowadays, as over 400 tRNAs and 660 tRNA genes have been characterized from various species (Sprinzl *et al.*, 1987), it is possible to understand the function of tRNAs and the features of their identity.

1.2. tRNA structure and identity.

The sequences of several hundred tRNAs from a wide variety of organisms have been determined and they nearly all conform to a general "cloverleaf" secondary structure, originally proposed by Holley *et al.*, (1965). This structure, which is maintained by base pairing between complementary short regions, is illustrated in figure 1.1. Although the overall range of tRNA lengths varies from 74 to 95 nucleotides (see below), the numbering of tRNA nucleotides, in 5' to 3' direction, is according to a yeast $tRNA^{Phe}$ which has 76 residues. This was adopted at the Cold Spring Harbour Symposium on tRNA in 1978 (Schimmel *et al.*, 1979) largely to allow relation of all tRNAs to the $tRNA^{Phe}$ sequence. Each cloverleaf contains 4 hydrogen-bonded stems and a number of non hydrogen-bonded loops. The 5' and the 3' end of the molecule forms the amino acid acceptor stem, held together by seven Watson-Crick base pairs. The 3' terminus consists of 4 unpaired nucleotides, the last 3 of which are always CCA. The amino acid becomes attached to the 3'-terminal A. Arm I ("D arm", see figure 1.1.) almost always contains the modified base dihydrouridine (D). This arm consists of a 3-4 bp stem and a variable size loop. Arm II (anticodon arm) consists of a 5 base pair stem and a 7 nucleotide loop. The middle three of these unpaired nucleotides form the tRNA anticodon, which pairs with the triplet codon on mRNA during protein synthesis. Arm III (extra arm) has a variable size (4-21 nts) and arm IV ("TψC arm") consists of a 5 base pair stem and a 7 nucleotide loop which always contains the sequence 5'-TψC-3'. All tRNAs contain some modified nucleotides (Sprinzl *et al.*, 1987) as well as invariant or semi-invariant nucleotide positions (constant

Figure 1.1.
Secondary and tertiary structure of tRNA.



A. The general cloverleaf secondary structure for tRNAs is shown. The positions of invariant and semi-invariant bases are shown. The numbering system is that of yeast $tRNA^{Phe}$. Y stands for pyrimidine, R for purine, Ψ for pseudouridine and H for hypermodified purine. R_{15} and Y_{48} are usually complementary. The dotted regions α and β in the D loop and in the variable loop contain different numbers of nucleotides in various tRNA sequences.
Diagram B illustrates the tertiary structure of a yeast $tRNA^{Phe}$. The ribose-phosphate backbone is drawn as a continuous ribbon and internal hydrogen bonding is indicated by crossbars. Position of single bases are indicated by rods that are intentionally shortened. The anticodon arm and the acceptor stem are shaded.
(Adapted from Watson *et al.*, 1987).

purine (R) or constant pyrimidine (Y)). The total number of invariant and semi-invariant nucleotides is 23 (see figure 1.1.; for review see Rich and RajBhandary, 1976).

During the 1970s several laboratories used X-ray diffraction with a view to deduce the three dimensional structure of yeast $tRNA^{Phe}$ crystals (Suddath *et al.*, 1974; Robertus *et al.*, 1974). The major characteristics of this structure (see figure 1.1.) which is sometimes called L-shaped, are the following : the acceptor stem with the TψC stem form an extended double helix and the D stem with the anticodon stem form another extended double helix. The region between the double helices, contains the TψC loop and the D loop. The folded molecule is stabilized by H- bonding and additional base stacking. The amino acid acceptor CCA is located at one end of the L, and the anticodon loop forms the other end. In addition, it was also observed that many tertiary hydrogen-bonding interactions involve invariant or semi-invariant nucleotides, strongly supporting the belief that all tRNAs have basically the same tertiary structure. The entire structure is about 2.5 nm thick and 9 nm long (Kim, 1979). Subsequent studies in solution confirm the tertiary tRNA structure, which is probably common to all tRNAs.

Recently new statements have emerged about tRNA identity. Each amino acid is represented by a group of tRNAs, called isoacceptors. Each group of isoaccepting tRNAs must be charged only by its single aminoacyl-tRNA synthetase, specific for the corresponding amino acid. Thus each group of isoaccepting tRNAs must share some common features enabling recognition only by the cognate aminoacyl-tRNA synthetases. Experiments performed by Normanly *et al.*, (1986), show that a $tRNA^{Leu}$ could loose its identity becoming a $tRNA^{Ser}$, by changing 12 nucleotides not involving the anticodon. More recently McClain and Foss (1988), proposed an identity set of 10 nucleotides for *E. coli* $tRNA^{Phe}$, whereas Hou and Schimmel (1988) showed that the G_3-U_{70} pair present in an *E. coli* $tRNA^{Ala}$, predominantly controls the aminoacylation of the $tRNA^{Ala}$ with alanine *in vivo* and *in vitro* . These and other studies demonstrate that for half of the tRNAs tested their identity survived substitution of the anticodon, but it is very dependent on the nucleotides present in the acceptor stem (for review see Yarus, 1988; Schulman and Abelson, 1988). De Duve (1988) called the latter region a paracodon, coining the term " the second genetic code"! However further analysis

is needed in order to establish general rules, if any, which govern tRNA identity.

The remainder of this study is concentrated on eukaryotic tRNA genes, their structure, organization and expression, with particular emphasis on mammalian tRNA genes.

1.3. Eukaryotic tRNA gene structure.

In eukaryotes each tRNA isoacceptor is nearly always encoded by more than one tRNA gene copy, the number of which varies from isoacceptor to isoacceptor, as well as from organism to organism (see section 1.4.). Nucleotide sequences of these copies are similar but not necessarily identical. As a consequence the sequence of the isoacceptor tRNA, might differ from some of the corresponding gene copies.

Furthermore as a result of post-transcriptional modifications and processing, the tRNA products may differ in several respects from the genes and their primary transcripts. The primary transcripts are processed by removal of 5' and 3' extremities, by modification of several nucleotides (e.g. formation of pseudouridine at position 54; see also figure 1.1.) and in some cases by excision of intervening sequence (see also section 1.5.6.). In addition all eukaryotic tRNA genes lack the 3' terminal CCA sequence in the acceptor arm which is added post-transcriptionally by RNA nucleotidyltransferase (For review see Deutscher, 1982).

An example available of the primary sequence change is that of rat $tRNA^{Asp}$. This tRNA is encoded by at least 10 $tRNA^{Asp}$ gene copies in the rat haploid genome. Eight of these copies have been sequenced and none of them has the same sequence of that found for rat $tRNA^{Asp}$. The sequence CT was found adjacent to the 5'-end of the anticodon in all characterized gene copies (residues 32-33; Shibuya *et al.*, 1982; Makowski *et al.*, 1983; Rosen *et al.*, 1984), but the sequence UC was found in the same position in the $tRNA^{Asp}$ (Kuchino *et al.*, 1981). Although at least two gene copies remain to be analyzed and may reveal the sequence apparently displayed by $tRNA^{Asp}$, it seems probable that some novel enzyme activity is responsible for post-transcriptional modification of the sequence CT to UC.

About 10 to 20 % of eukaryotic genes appear to contain intervening sequences

(Sharp *et al.*, 1985). Such genes have been found in *Drosophila melanogaster* (Robinson and Davidson, 1981); in yeast (Goodman *et al.*, 1977; Baldi *et al.*, 1983; Gamulin *et al.*, 1983; Swerdlow and Guthrie, 1984; Willis *et al.*, 1984); in *Xenopus Laevis* (Muller and Clarkson, 1980; Laski *et al.*, 1982a); in chicken (Wittig and Wittig, 1979); in human (McPherson and Roy, 1986; Van Tol and Beier, 1988) and in other eukaryotes. The length of eukaryotic tRNA gene introns varies from 8 to 113 nucleotides. Families of isoacceptor tRNA genes generally appear to have homologous intervening sequences. A tRNA gene family which has been studied on the subject, is the human $tRNA^{Tyr}$ gene family. All the 12 copies of the gene for $tRNA^{Tyr}$ contain introns (Van Tol and Beier, 1988) whose size, in four characterized copies, is 21, 21, 20 and 20 nucleotides (McPherson and Roy, 1986; Van Tol *et al.*, 1987). In *Drosophila melanogaster* 5 out of 6 characterized $tRNA^{Tyr}$ gene copies contain introns of 20 or 21 nucleotides; however the sixth characterized copy contains a much larger intron, of 113 nucleotides (Suter and Kubli, 1988).

The location of all intervening sequences in sequenced tRNA genes is one nucleotide to the 3'-end of the tRNA anticodon (i.e. after position 37). In addition, for some precursor tRNAs, transcribed from such genes, 3 contiguous nucleotides of the intervening sequence can base pair with the anticodon, while the rest of the intervening sequence usually forms a stem and a loop. The function of intervening sequences in the expression of such tRNA genes will be discussed in section 1.5.6.

1.3.1. tRNA pseudogenes.

One of the most intriguing observations, after sequencing cloned fragments of genomic DNA which hybridized tRNA probes, is the finding of sequences which resemble tRNA genes. These sequences, called "tRNA pseudogenes", consist either of a partial sequence of a corresponding tRNA sequence, or of DNA sequences displaying considerable sequence difference from the corresponding tRNA. Some of the tRNA pseudogenes cannot be folded into a cloverleaf configuration, and/or do not contain all the invariant and semi-invariant nucleotides present in tRNA genes.

TransferRNA pseudogenes displaying nucleotide heterogeneity have been found in several eukaryotes. A human opal suppressor phosphoserine tRNA pseudogene

differs in 7 nucleotides from the corresponding "true" gene (O' Neil *et al.*, 1985), whereas a human $tRNA^{Gly}$ pseudogene differs in just 2 nucleotides (including the invariant C at position 56) from the corresponding $tRNA^{Gly}$ gene (Pirtle *et al.*, 1986). The latter pseudogene has 8 nucleotides of 5'-flanking sequence identical to those found in the $tRNA^{Gly}$ gene, as well as a transcription termination site.

Several "incomplete tRNA genes" (pseudogenes) have also been found in eukaryotic genomes. A fragment which contains four regions of approximately 50% homology with the initiator tRNA, the longest of which is only 34 bp, is dispersed in about 30 sites in the *Drosophila melanogaster* genome (Sharp *et al.*, 1981). These fragments also have a hybridization pattern characteristic of a mobile DNA element, suggesting that the pseudogenes may have arisen by the repeated insertion and excision of a transposable element into an intact tRNA gene. In *Drosophila melanogaster* a $tRNA^{His}$ pseudogene is clustered with a $tRNA^{His}$ gene. The pseudogene contains eight consecutive base pairs different from the region of the gene which codes for the 3' portion of the anticodon stem of $tRNA^{His}$. Interestingly this pseudogene was found to be transcriptionally active *in vitro* when its 5'-flanking sequence was replaced with the 5'-flanking sequence of an active $tRNA^{His}$ gene (Cooley *et al.*, 1984).

In rat a 3.3 kb DNA fragment which contains single copies for the genes of $tRNA^{Leu}$, $tRNA^{Asp}$, $tRNA^{Gly}$ and $tRNA^{Glu}$ is reiterated, with a copy number of 10 per haploid genome. Nucleotide sequence analysis of several of these DNA fragment repeats revealed the presence of 9 pseudogenes. Three of the DNA fragment repeats were each found to contain incomplete $tRNA^{Gly}$ and $tRNA^{Glu}$ sequences instead of the corresponding genes; two DNA fragment repeats were found to contain incomplete $tRNA^{Gly}$ sequences and an additional repeat was found to contain a $tRNA^{Leu}$ pseudogene instead of the $tRNA^{Leu}$ gene. The $tRNA^{Leu}$ pseudogene differs in 15 residues from the $tRNA^{Leu}$ gene (Rosen *et al.*, 1984). Two of the $tRNA^{Glu}$ pseudogenes are lacking the last 11 nucleotides of the $tRNA^{Glu}$ gene (Shibuya *et al.*, 1982), while the third $tRNA^{Glu}$ pseudogene is lacking the last 17 nucleotides of the same gene (Makowski *et al.*, 1983). Finally all five $tRNA^{Gly}$ pseudogenes are lacking 7 nucleotides of the $tRNA^{Gly}$ gene (residues 20 to 26); in addition three of them differ in 1, 2 and 3 nucleotides from the true gene (Shibuya *et al.*, 1982; Makowski *et al.*, 1983).

However the most interesting case reported so far, concerns a mouse $tRNA^{Phe}$ pseudogene (Reilly *et al.*, 1982). This pseudogene consists of 38 contiguous nucleotides of the $tRNA^{Phe}$ 3'-end region (residues 39 to 76), including the terminal CCA. Since CCA is added post-transcriptionally to all known eukaryotic tRNAs for which the genes have been identified, the authors suggested that this pseudogene represents an example of a possible reverse flow of genetic information. In addition in human, a 17 nucleotides long sequence has been identified, which shows 82% homology to the 3'-end of a *Xenopus Laevis* $tRNA^{Tyr}$ gene (Bansal,1986). Notably the latter human tRNA-like sequence, as well as the human opal suppressor tRNA pseudogene reported earlier (O'Neil *et al.*, 1985), are both located close to Alu sequences; (a class of short interspersed repetitive DNAs, approximately 300 nucleotides long, found exclusively in the human genome; for review on Alu family see Kariya *et al.*, 1987).

No tRNA pseudogenes have been observed in yeast genomes so far. Since many yeast tRNA genes have now been characterized, it appears that tRNA pseudogenes are absent from this organism.

All the eukaryotic tRNA pseudogenes so far investigated are inactive *in vitro*. (The only potential exception concerning a *Drosophila melanogaster* $tRNA^{His}$ pseudogene was discussed earlier; Cooley *et al.*, 1984). However it is still to be answered whether the pseudogenes are functional in terms of binding transcription factors, or merely "evolutionary remnants" (Sharp *et al.*, 1985).

1.4. Organization of eukaryotic tRNA genes.

The organization of eukaryotic tRNA genes has been examined in various species. Up to now (end of 1988), approximately 200 tRNA genes have been identified and characterized via nucleotide sequence analysis. These genes appear either as solitary copies or as clusters. The tRNA genes contained within such clusters which may code for the same or different tRNAs, are irregularly spaced, occur on either DNA strand and usually are not cotranscribed as large multimeric precursors, as is the case with *E. coli* tRNA genes. The copies of tRNA gene families are not always embedded within homologous genomic environments, although they are usually located in the same chromosomal locus. In addition the

gene copy number for the same isoacceptor tRNA, varies from organism to organism, although it is usually proportional to the genome size. As a result no general rules concerning the organization of eukaryotic tRNA genes have emerged.

TransferRNA genes have been sequenced from 17 different eukaryotic species (Sprinzl *et al.*, 1987). However the remainder of this section focuses on studies concerning yeast, *Drosophila melanogaster*, *Xenopus Laevis*, rat, mouse and human tRNA genes; in these species several genes have been sequenced and assessments about the organization of tRNA genes have emerged.

1.4.1. Organization of yeast tRNA genes.

The yeast haploid genome (10^7 bp) contains approximately 360 tRNA genes (Schweizer *et al.*, 1969). Each gene is reiterated on average 6 to 10 times (Feldman, 1976). Several studies on yeast tRNA gene families show that gene copies are solitary throughout the genome. The best studied yeast tRNA gene family is that of the gene for $tRNA^{Glu}$ (Hauber *et al.*, 1988), where eight cloned gene copies have been characterized. These copies are identical in their gene sequence, contain no introns and are flanked by transposable elements. They are embedded in entirely different genomic environments and have been assigned to different chromosomal loci.

However in some cases clusters of yeast tRNA genes have also been observed. A $tRNA^{Ser}$-$tRNA^{Met}$ gene pair has been found in three different loci in *S. pombe* . The genes are separated by only 7 bp and are transcribed as a dimeric precursor. The total length of the sequence homology among the three pairs is about 200 bp; however the flanking sequences outside that region are totally different (Mao *et al.*, 1980; Amstutz *et al.*, 1985). Similarly four separate clones have been identified, in which genes for $tRNA^{Arg}$ and $tRNA^{Asp}$ are separated by only 9 bp and are transcribed *in vitro* as a dimeric precursor (Schmidt *et al.*, 1980).

The majority of yeast tRNA genes contains introns (see also section 1.3.), although a few intronless genes have also been reported (Baker *et al.*, 1982; Hauber *et al.*, 1988). Finally transposable elements such as *sigma*, *delta*, Ty1 and *tau* are frequently associated with yeast tRNA genes (Hauber *et al.*, 1988 and references). The most common of these transposons is the *sigma* element. This element has been found to occur 16 to 18 bp upstream from the 5'-end of several

yeast tRNA genes (Sandmeyer and Olson, 1982; Del Rey *et al.*, 1982; Brodeur *et al.*, 1983; Sandmeyer *et al.*, 1988). *Sigma* is about 340 bp long, consists of 8 bp inverted repeats at its ends and is flanked by 5 bp direct repeats. *Sigma* elements are so far exclusively associated with tRNA genes and their role in evolution, organization as well as transcription initiation of tRNA genes is still under discussion. (For review see Sharp *et al.*, 1985).

1.4.2. Organization of *Drosophila melanogaster* tRNA genes.

The *Drosophila melanogaster* haploid genome contains on average 600-750 tRNA genes (Weber and Berger, 1976). The localization of these genes has been substantially simplified by *in situ* hybridization to polytene chromosomes (Gall and Pardue, 1969) and shows that tRNA genes are distributed at about 54-60 different loci over most chromosomes,as solitary genes or as clusters of genes which code for either different tRNAs (heteroclusters), or the same tRNA (homoclusters; Steffenson and Wimber, 1971; Elder *et al.*, 1980).

The best studied cluster is that located in region 42A of the chromosome 2R (Hovemann *et al.*, 1980; Yen and Davidson, 1980). Yen and Davidson's group isolated a region of 94 kb of which a central 46 kb region contains 3 widely spaced homoclusters of $tRNA^{Asn}$, $tRNA^{Arg}$ and $tRNA^{Lys}$ genes. These genes were found in both DNA strands. Another homocluster which consists of five nearly identical copies of $tRNA^{Glu}_{CTC}$ has also been characterized. The genes are organized as follows: a gene triplet spanning approximately 0.55 kb, followed by a 0.45 kb gene doublet 3.0 kb downstream. Examination of the immediate flanking sequences of each gene revealed a striking pattern of sequence homologies among certain of the copies. As a result, the authors proposed that two ancestral $tRNA^{Glu}_{CTC}$ genes each gave rise to gene doublets by duplication, while one of these gene pairs then gave rise, in turn, to a trio of genes as a result of unequal crossover (Hosbach *et al.*, 1980).

Nine $tRNA^{Ser}$ genes (for $tRNA_4^{Ser}$ and $tRNA_7^{Ser}$) have also been characterized, eight of which compose a major $tRNA_{4,7}^{Ser}$ homocluster at

chromosomal locus 12DE. The $tRNA_4^{Ser}$ and $tRNA_7^{Ser}$ genes are 96% homologous and all the gene copies show several blocks of homology in their flanking sequences (Cribbs *et al.*, 1987). In addition 6 $tRNA^{Tyr}$ genes have been characterized; three compose a homocluster, two a second homocluster, while the sixth gene copy is a solitary gene. The genes contain different size introns (see section 1.3.) although their exons are identical. No homologies were observed in their flanking sequences (Suter and Kubli,1988).

However several solitary tRNA genes have also been reported. Sharp *et al.*, (1981), identified single copies for the gene of $tRNA_i^{Met}$ at chromosomal locus 61D, within a 415 bp repeated DNA fragment, which is separated from the next $tRNA_i^{Met}$ gene by a distance of approximately 17 kb. Finally Hershey and Davidson (1980) characterized two direct duplicate fragments from chromosomal locus 56F, both 1.1 kb long, and each containing single copies of $tRNA^{Gly}$ genes.

1.4.3. Organization of *Xenopus Laevis* tRNA genes.

The *Xenopus Laevis* genome contains a large population of tRNA genes (6500-7800 genes per haploid complement; see Land and Dawid, 1980). The gene copies, scattered or clustered, are usually located in 150 to 300-fold repeated DNA fragments, which can well be separated from the rest of the genomic DNA by buoyant density centrifugation (Clarkson and Kurer, 1976). One of these DNA fragments, 3.18 kb long, contains 8 tRNA genes and is tandemly repeated approximately 150 times per *Xenopus Laevis* haploid genome (Clarkson *et al.*, 1978; Muller and Clarkson, 1980; Fostel *et al.*, 1984; Muller *et al.*, 1987). These genes (two coding for $tRNA^{Met}$ and single copies for $tRNA^{Phe}$, $tRNA^{Tyr}$, $tRNA^{Asn}$, $tRNA^{Ala}$, $tRNA^{Leu}$ and $tRNA^{Lys}$) are irregularly spaced and coded in both DNA strands. In addition a $tRNA^{Val}$ gene was found to be located in two 250-fold repeated fragments, 892 bp and 4100 bp long (Peterson, 1987). However a dispersed $tRNA^{Tyr}$ gene has been characterized from a 9.4 kb DNA fragment, which is not reiterated with high frequency in the *Xenopus Laevis* genome. This gene differs from the clustered $tRNA^{Tyr}$ gene by a single purine transition within the coding region, has a different intron, as well as different 5'

and 3'-flanking sequences (Gouilloud and Clarkson, 1986).

1.4.4. Organization of mammalian tRNA genes.

The average copy number of genes for mammalian tRNA species varies from 10 to 20 copies per haploid genome, for man (Hattlen and Attardi, 1971) and rat (Lasser-Weis *et al.*, 1981) and up to 100 for mouse (Marzluff *et al*, 1975). Although there are no general rules concerning the organization of mammalian tRNA genes, several similarities have been observed between genes, which code the same isoacceptor, from different mammalian species. In the first part of this section the organization of rat and mouse tRNA genes will be examined, while the last section focuses on the organization of human tRNA genes.

1.4.4.1. Organization of the rat and mouse tRNA genes.

In rat and mouse genomes, tRNA genes have been found in small clusters, although solitary genes have also been observed. The best studied cluster in rat contains 4 tRNA genes (single copies for $tRNA^{Leu}$, $tRNA^{Asp}$, $tRNA^{Gly}$ and $tRNA^{Glu}$). The genes are found in both transcription directions and they are separated by 450-600 bp DNA "spacers". The whole unit, 3.3 kb long, is reiterated about 10 times per haploid genome (Sekiya *et al.*, 1981; Shibuya *et al.*,1982 and 1985). Nucleotide sequence analysis of 8 of these repeats revealed the following : five of the repeats are highly homologous. The 5'- and 3'-flanking sequences of the gene copies for $tRNA^{Asp}$, $tRNA^{Gly}$ and $tRNA^{Glu}$ in each repeat are almost identical. (The $tRNA^{Leu}$ genes were not sequenced in these studies). Surprisingly, four of these repeats contain incomplete "$tRNA^{Gly}$ genes", of which two also contain incomplete "$tRNA^{Glu}$ genes" (Shibuya *et al.*, 1982; see also section 1.3.1.). A sixth cluster also contains $tRNA^{Asp}$ and $tRNA^{Leu}$ genes as well as $tRNA^{Gly}$ and $tRNA^{Glu}$ pseudogenes (Makowski *et al.*, 1983). Rosen *et al.*, (1984) characterized the two remaining repeats, one of which contains a $tRNA^{Leu}$ pseudogene instead of a $tRNA^{Leu}$ gene. The same authors also compared the nucleotide sequence data from all characterized clusters. Seven clusters appear to

have homologous sequences, although deletions, insertions or base substitutions at several positions were also observed. However one repeat (the one containing the $tRNA^{Leu}$ pseudogene) shows minimal homology to the rest of the repeats. All these data led Rosen *et al.*, (1984) to suggest that the 10 copies of the cluster were derived by duplication, followed by mutations and deletions of non essential sequences (see also figure 1.2.). In addition Looney and Harding (1983) characterized a mouse tRNA gene cluster which is homologous to one of the above mentioned rat sequences (Sekiya *et al.*, 1981). The rat and the mouse $tRNA^{Asp}$ and $tRNA^{Gly}$ genes are identical, while the mouse $tRNA^{Glu}$ gene differs from the rat $tRNA^{Glu}$ gene in two residues. The flanking sequences of the genes of the two clusters, show a degree of homology. The same group also characterized a solitary mouse $tRNA^{Asp}$ gene which is identical to the clustered $tRNA^{Asp}$ gene.

Several other tRNA gene clusters have been found in rat and mouse. One cluster in rat contains 3 $tRNA^{Lys}$ genes and 3 $tRNA^{Pro}$ genes. The genes occur in both strands and show no homology in their 5'-flanking sequences (Sekiya *et al.*, 1982). In mouse a cluster containing at least 7 tRNA genes, two of which (for $tRNA^{Pro}$ and $tRNA^{Gly}$) have been sequenced (Hu *et al.*, 1983), and a cluster containing single copies of genes for $tRNA^{Ala}$, $tRNA^{Ile}$, $tRNA^{Pro}$ and $tRNA^{Lys}$ has been reported (Russo *et al.*, 1987).

Scattered tRNA genes have been characterized for : a mouse $tRNA^{His}$ gene, which differs in only one residue from a sheep $tRNA^{His}$ gene (Han and Harding, 1982); a mouse $tRNA_i^{Met}$ gene, whose 5'-flanking region is homologous to a human $tRNA_i^{Met}$ gene (Han *et al.*, 1984); and finally a rat $tRNA^{Phe}$ gene which is identical but for having one extra nucleotide in the variable arm, from a rabbit $tRNA^{Phe}$ gene (Rosen and Daniel, 1988).

1.4.4.2. Organization of human tRNA genes.

The human haploid genome contains approximately 1200 tRNA genes (Hatlen and Attardi, 1971), representing about 60 different genes of 10 to 20 copies each. These genes are dispersed either as solitary genes or as clusters. Although at least 30 gene copies have been characterized so far, very little is known about the organization of tRNA gene families in the human genome.

About 12 $tRNA_i^{Met}$ genes have been found at scattered locations in the human genome. Four DNA fragments of sizes from 11 to 18 kb, each containing a single $tRNA_i^{Met}$ gene, were shown to represent two different loci with homologies limited to several dispersed repetitive sequences within each chromosomal neighbourhood. Analysis of two $tRNA_i^{Met}$ gene regions show that the genes are identical, their flanking sequences have several blocks of homology and are both located on human chromosome 6 (Zasloff and Santos, 1980; Santos and Zasloff, 1981; Naylor *et al.*, 1983).

At least 12 independent gene loci for $tRNA^{Tyr}$ have also been detected in the human genome. Nucleotide sequence analysis of four of these copies (McPherson and Roy, 1986; Van Tol *et al.*, 1987) and further analysis performed by Van Tol and Beier (1988) show that all $tRNA^{Tyr}$ genes carry an intron (see also section 1.3.). In addition all four characterized copies have nearly identical exons, while 2 of them, characterized by McPherson and Roy (1986), belong to homologous 200 bp DNA fragment repeats in the human genome.

Studies on the human $tRNA^{Val}$ gene family show that 3 of the copies occur only once in the haploid genome. Two of the genes encode the major and minor $tRNA^{Val}$ isoacceptor while the third copy does not correspond to any known $tRNA^{Val}$. Comparison of the flanking sequences of the copies did not reveal significant homologies. In addition an Alu-type repeat was found within several hundred base pair distance from two of the $tRNA^{Val}$ genes (Arnold *et al.*, 1986).

Three clones containing 15-17 kb DNA fragments were isolated from a human DNA library and found to contain single copies of a $tRNA^{Gly}_{GCC}$ gene (Pirtle *et al.*, 1986; Doran *et al.*, 1988). The three genes are identical to human $tRNA^{Gly}_{GCC}$ (Gupta *et al.*, 1979), but their flanking sequences are different and one of them is 758 bp away from a $tRNA^{Gly}_{GCC}$ pseudogene (Pirtle *et al.*,1986). Finally the nucleotide sequence of a $tRNA^{Gly}_{CCC}$ has also been determined and found to be of a striking homology (96%) with the human $tRNA^{Gly}_{GCC}$ (Shortridge *et al.*, 1985).

Several other solitary human tRNA genes have also been reported: one $tRNA^{Glu}$ gene (Goddard *et al.*, 1983; see also chapter 2); two $tRNA^{Asn}$ genes identical in all but in 3 nucleotides which belong to 90% homologous 500 bp DNA

fragments (Ma *et al.*, 1984); 4 $tRNA^{Ser}$ genes (Yoo, 1984; Hong *et al.*,1987; Hoe *et al.*, 1987; Krupp *et al.*, 1988) and finally one opal suppressor phosphoserine tRNA gene which is located in chromosome 19 (O'Neil *et al.*, 1985; McBride *et al.*, 1987).

A 12.5 kb clone containing tRNA genes, was also isolated from a human DNA library. A 1.65 kb fragment of this clone was sequenced and found to contain three tRNA genes; one for $tRNA^{Lys}$, one for $tRNA^{Gln}$ and one for $tRNA^{Leu}$. The three genes are separated by about 500 bp intergenic regions (Roy *et al.*, 1982; Buckland *et al.*, 1983). Characterization of another clone from the same DNA library , revealed the presence of a nearly identical tRNA gene cluster (McLaren and Goddard, 1986a) which is separated from a single copy $tRNA^{Gly}$ gene by a 4.85 kb DNA spacer (Bourn and Goddard, unpublished results). This second cluster, which has been assigned to chromosome 17 (Morrison *et al.*, manuscript in preparation), shares common restriction sites, extending over at least 4.5 kb, with the first cluster but differs at regions 3 kb upstream from this common region (Bourn and Goddard, unpublished results).

Finally two additional tRNA gene clusters have also been characterized. One of them contains 4 tRNA genes arranged as two tandem pairs, separated by 3 kb. A $tRNA^{Pro}$ gene is separated from a $tRNA^{Leu}$ gene by a 724 bp intergenic region in the first pair and a second $tRNA^{Pro}$ gene is 316 bp away from a $tRNA^{Thr}$ gene in the second pair. In addition a putative Alu-like element was found to occur within a 2.0 kb DNA fragment, at least 700 bp away from the tRNA gene cluster. The $tRNA^{Pro}$ genes are identical although no apparent homology occursbetween their 5'-flanking regions (Chang *et al.*, 1986). The other cluster also contains 4 tRNA genes. Two of the genes code for $tRNA^{Lys}$ and the other two for $tRNA^{Phe}$. The $tRNA^{Lys}$ genes are identical, while the $tRNA^{Phe}$ genes differ in just 1 nucleotide (at position 57). The $tRNA^{Lys}$ and $tRNA^{Phe}$ genes are organized in alternating order and are irregularly spaced, by intergenic regions of approximately 1.0, 2.6 and 5.0 kb. Finally no homology was observed in the flanking sequences of the same isoacceptor gene copies (Doran *et al.*, 1987).

1.5. Transcription of eukaryotic tRNA genes.

The availability of cloned tRNA genes in combination with the use of cell-free transcription systems or their micro-injection into *Xenopus* oocyte nuclei, have substantially facilitated the study of tRNA gene transcription. By using these techniques it is possible to achieve the synthesis of sufficient amounts of tRNA precursor molecules for fingerprint analysis, or direct RNA sequencing. In addition by comparing the nucleotide sequence of the primary transcript with that of the respective gene and its flanking regions, it is feasible to identify how the transcript is processed. Such studies revealed that the transcriptional product of eukaryotic tRNA genes is not tRNA itself but a longer precursor molecule. This precursor molecule must be processed post-transcriptionally by several enzymes in a number of fixed temporal sequential steps in order to yield a mature tRNA. Transcriptional studies on tRNA genes have also allowed the characterization of initiation and termination sites for RNA polymerase III, the enzyme responsible for tRNA gene transcription. Additional analysis has shown that promotion of transcription occurs within the coding region of the gene and also involves the binding of specific transcription factors.

1.5.1. Genes transcribed by RNA polymerase III.

Eukaryotic RNA polymerase III consists of 9 to 15 subunits having a total molecular weight of about 650,000 daltons (for review see Ciliberto *et al.*, 1983a and Geiduschek and Tocchini-Valentini, 1988). The enzyme transcribes a diverse set of genes, such as : tRNA genes, 5S ribosomal RNA genes, virus-associated RNA encoded by adenovirus 2 (VA I and VA II genes), Epstein-Barr virus genes (EBER I and EBER II), as well as numerous RNAs encoded by repetitive elements and a variety of other types of small nuclear and cytoplasmic RNAs. In general the transcription of RNA polymerase III genes is under the control of intragenic DNA sequences (Geiduschek and Tocchini-Valentini, 1988 and references therein). Two kinds of internal control regions (ICR) have been identified, one for 5S RNA genes (type 1) and another for tRNA genes (type 2). The ICRs of the genes encoding for VAI, VAII, EBERI, EBERII,

7SL and mouse 4.5S RNAs as well as for human Alu and mouse BI families, can also be classified to type 2. However the U6 and 7SK genes appear to belong to a separate type of RNA polymerase III genes, since their promoter functions are determined by 5'-flanking sequences (for review see Folk,1988 and Sollner-Webb, 1988).

After a detailed analysis the sequence required to direct initiation of transcription of type 1 genes (5S RNA) has been characterized near the centre of the transcribed region. Deletion of nucleotides from each end of 5S DNA and oligonucleotide-directed mutagenesis have shown that the ICR of 5S RNA genes is located between nucleotides 50 and 97 of the gene, but that it is not a continuous structure (Sakonju *et al.*, 1980; Bogenhagen *et al.*, 1980; Pieler *et al.*, 1985a, 1985b and 1987). In these genes RNA polymerase III initiates transcription at approximately 50 nucleotides from the 5' end of ICR. Further analysis revealed that the ICR consists of a box A (bp 50-64), an intermediate element (bp 67-72) and a box C (bp 80-97) where at least three transcription factors (TFIIIA, TFIIIB and TFIIIC; Segall *et al.*, 1980; Lassar *et al.*, 1983; Bieker *et al.*, 1985; Setzer and Brown, 1985) are required to bind prior to the addition of RNA polymerase III. TFIIIA, a Zn containing protein, binds to the whole ICR of 5S genes (although it has no effect on transcription of type 2 genes; Gottesfeld and Bloomer, 1982) and helps recruit TFIIIB and TFIIIC to a stable transcription complex (Engelke *et al.*, 1980; Sakonju *et al.*, 1981; Bogenhagen *et al.*, 1982; Brown, 1984; Schlissel and Brown, 1984).

A characteristic feature of all type 2 genes is that their intragenic promoter is split into two blocks (A and B), separated by 30-60 nucleotides. The sequences of both blocks A and B are highly conserved among type 2 gene promoters (see table 1.1.); in addition block A sequence is structurally and functionally homologous with the first component (box A) of the promoter of type 1 genes (Ciliberto *et al.*, 1983b; Geiduschek and Tocchini-Valentini, 1988). Moreover tRNA genes are shown to be efficient competitors of adenovirus VA I genes as well as of type 1 genes (5S RNA), in transcription experiments *in vitro* , suggesting that faithful transcription of these genes depends on similar transcription factors (Segall *et al.*, 1980; Burke *et al.*, 1980; Wormington *et al.*, 1981; Guilfoyle and Wienmann, 1981). Initiation of transcription of all type 2 genes occurs between 10 and 20 nucleotides 5' to the A block boundary (Ciliberto *et al.*, 1982b). The

Table 1.1.
Comparison of ICRs of type 2 genes.

	5' ICR	3' ICR
tDNA	TGGCnnAGTGG	GTTCGA
Alu(CHO)	TGGCtcgagGt	GTTCaA
Alu(human)	TGGCtcAGgct	GTTCaA
EBER I	TGcCctAGaGG	GTaCaA
EBER II	TGcCctAGTGG	GgTCaA
Human 7SL	TGGCgcgtgcc	GTTCtg
Mouse 4.5S	TGGCgcAcgcc	GTTCGA
VAI	TGGtctgGTGG	GTTCGA
VAII	TaGCcggagGG	GTTCGA
consensus sequence →	TGGCnnRR$^{T}_{G}$GG	GTTCRA

This table compares the relative common ICR sequences of several type 2 genes. Capital letters indicate conserved residues for all type 2 genes compared. As a result a general consensus sequence for ICRs of type 2 genes can be derived. (Adapted from Sharp *et al.*, 1985).

mechanism of initiation, promotion and termination of tRNA gene transcription will be discussed in the following sections.

1.5.2. Initiation of transcription.

The transcription of eukaryotic tRNA genes initiates at a purine nucleotide usually located between position -20 and -3 upstream from the tRNA coding sequence (Clarkson, 1983). There is no obvious sequence homology surrounding the initiating residue except that in most cases the purine is preceded by a pyrimidine nucleotide. All known tRNA genes possess a single initiation site except two reported initiator tRNA genes of *Drosophila melanogaster* and *Xenopus Laevis*, where transcription starts at two or three different residues (Sharp *et al.*, 1981b; Koski and Clarkson, 1982b).

1.5.3. The tRNA gene promoter.

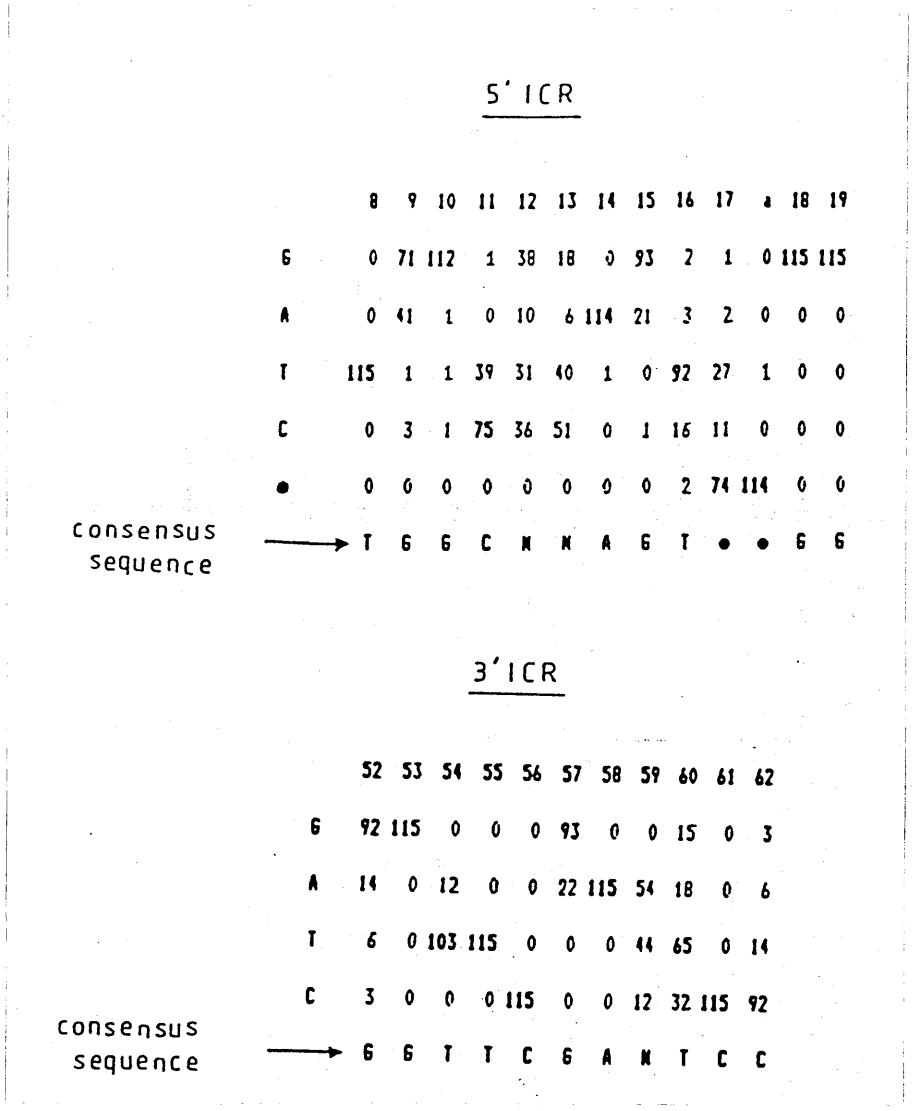
Transcription studies on eukaryotic tRNA genes have shown that these genes contain intragenic control regions (ICR) that direct transcription. Detailed analysis of a *Xenopus Laevis* $tRNA_i^{Met}$ gene (Kressmann *et al.*, 1979; Hofstetter *et al.*, 1981; Folk and Hofstetter, 1983), a $tRNA^{Leu}$ gene (Galli *et al.*, 1981), a *Drosophila melanogaster* $tRNA^{Lys}$ gene (De Franco *et al.*, 1980), a $tRNA^{Arg}$ gene (Sharp *et al.*, 1981a, 1982, 1983a and 1983b), a yeast suppressor $tRNA^{Tyr}$ gene (Koski *et al.*, 1980 and 1982; Allison *et al.*, 1983), and a *C. elegans* $tRNA^{Pro}$ gene (Ciliberto *et al.*, 1982; Traboni *et al.*, 1984), have revealed that the intragenic promoter is non-contiguous and split into two blocks that are set apart. These blocks are called the A block (or "D" control region or 5'- ICR) and the B block (or "T" control region or 3'-ICR), having the approximate positions 8-19 and 52-62 respectively. (According to the standard system of numbering tRNA and tRNA genes). The two ICRs coincide with the highly conserved nucleotides found in the D and TψC arm regions among eukaryotic tRNA genes. Comparative study of the ICRs of all sequenced tRNA genes has shown that the 3' ICR sequence shows little variation, while the 5' ICR sequence shows some variation with regard to sequence as well as to the number of nucleotides present

(Sharp *et al.*, 1985; Geiduschek and Tocchini-Valentini, 1988; see also table 1.2.). The outcome of this comparison has revealed the following consensus sequence for the internal split promoter : TRGCNNAGY-GG for the A block and GGTTCGANTCC for the B block (where R represents a purine, Y a pyrimidine and N any nucleotide; Geiduschek and Tocchini-Valentini, 1988).

1.5.4. The function of tRNA gene promoter; interaction with transcription factors.

Initial *in vitro* experiments performed on VAII RNA and tRNA genes, using cell free extracts, have shown that multiple factors are required for the accurate transcription of these genes (Segall *et al.*, 1980). Other studies revealed that mutations in the ICRs of tRNA genes drastically reduce the rate and processing of their transcription (Newmann *et al.*, 1983; Traboni *et al.*, 1984; Sullivan and Folk, 1987). tRNA genes appear to bind rapidly and stably a limiting transcription factor when added to cell-free extracts, as shown by the time-dependent inhibition of transcription of a reference tRNA gene added to the extract after the addition of a competitor gene. In these experiments the 3'- half of the gene was shown to be the most important region in the transcription competition assays (Kressman *et al.*, 1979). In other transcription assays, mutants having regions deleted from either the 3' or the 5' side of a *Drosophila melanogaster* $tRNA^{Arg}$ gene were tested for ability to compete for limiting transcription factors (Sharp *et al.*, 1983a), or in pre-incubation-competition assays designed to test for stable complex formation (Schaack *et al.*, 1983). These assays also demonstrated that the B block is required for the gene to form stable complexes. While the B block was shown to be the most important region for stable complex formation, other regions of the DNA, and in particular the presence of the A block, are also required for maximal rate and strength of stable complex formation. In addition it was also found, by the use of footprinting methods, that transcription factors bind to both ICRs of tRNA genes (Klemenz *et al.*, 1982; Stillman and Geiduschek, 1984b; Fuhrman *et al.*, 1984; Stillman *et al.*, 1985; Camier *et al.*, 1985), whereas parallel studies showed that TFIIIC interacts predominantly with B block but also with A block (Stillman and

Table 1.2.
Comparison of ICR sequences of tRNA genes.



The nucleotide occurence of each position within the 5' and 3' ICR (residues 8-19 and 52-62 respectively) was computed from the collection of 115 tRNA gene sequences shown in Sharp *et al.*,(1985; table 1). Bold numbers correspond to the base predominantly found in each position numbered according to yeast $tRNA^{Phe}$ (figure 1.1.). Dark circles indicate the absence of base at the corresponding position. The consensus sequence derived as a result of this comparison is shown. Further statistical analysis performed by Geiduschek and Tocchini-Valentini (1988) revealed that residue 9 is generally purine and residue 16 is generally pyrimidine.
(Adapted from Sharp *et al.*, 1985).

Geiduschek, 1984b; Baker and Hall, 1984; Baker *et al.*, 1986; Huibregtse *et al.*, 1987). Finally other authors have revealed that transcription of tRNA genes requires TFIIIB (Burke and Soll, 1985; Carey *et al.*, 1986).

The experiments earlier described, demonstrate that there are a minimum of two transcription factors (TFIIIB and TFIIIC) in addition to RNA polymerase III which are required for accurate tRNA gene transcription. TFIIIC interacts predominantly with the B block and it remains bound for several rounds of transcription initiation. The stable complex formation of TFIIIC-ICR was found to be dependent on the right ionic environment (Ruet *et al.*, 1984; Stillman *et al.*, 1984a; Gabrielsen and Oyen, 1987). The tRNA gene also requires TFIIIB for the formation of a stable complex, but TFIIIB does not remain stably bound and appears to be rapidly recycled. However TFIIIB has no recognizable ability to bind specifically on its own (Carey *et al.*, 1986). Whether TFIIIB directly contacts DNA while also anchored to TFIIIC, or whether it indirectly affects the binding of the conjugate domain of the TFIIIC to A block, remains to be decided (Geiduschek and Tocchini-Valentini, 1988).

The nucleotide sequence located between the two ICR of the 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). Several experiments have been performed where mutant tRNA genes, contained from 12 to 1530 nucleotides spacers, were tested in terms of their ability for factor binding and transcriptional efficiency. (Ciliberto *et al.*, 1982b; Dingermann *et al.*, 1983; Baker and Hall, 1984). These studies have shown that a distance of 34-53 nucleotides separating the A and B blocks allows the highest TFIIIC-tDNA stability.

During the last two years several groups have resolved TFIIIC into two functional components (TFIIIC1 and TFIIIC2; Yoshinaga *et al.*, 1987; Van Dyke and Roeder, 1987; Dean and Berk, 1987). Both TFIIIC1 and TFIIIC2 fractions are required for *in vitro* transcription of type 2 genes, tested as VAI genes, whereas footprinting analysis has shown that TFIIIC2 fraction binds to the B block. In addition Ottenelo *et al.* (1987) have reported the resolution of an unidentified factor (TFIIID) which is required for *in vitro* transcription of RNA polymerase III templates (tested as tRNA and 5S RNA genes). However the function of TFIIIC1 and TFIIIC2 fractions as well as TFIIID on tRNA gene transcription is still under

discussion.

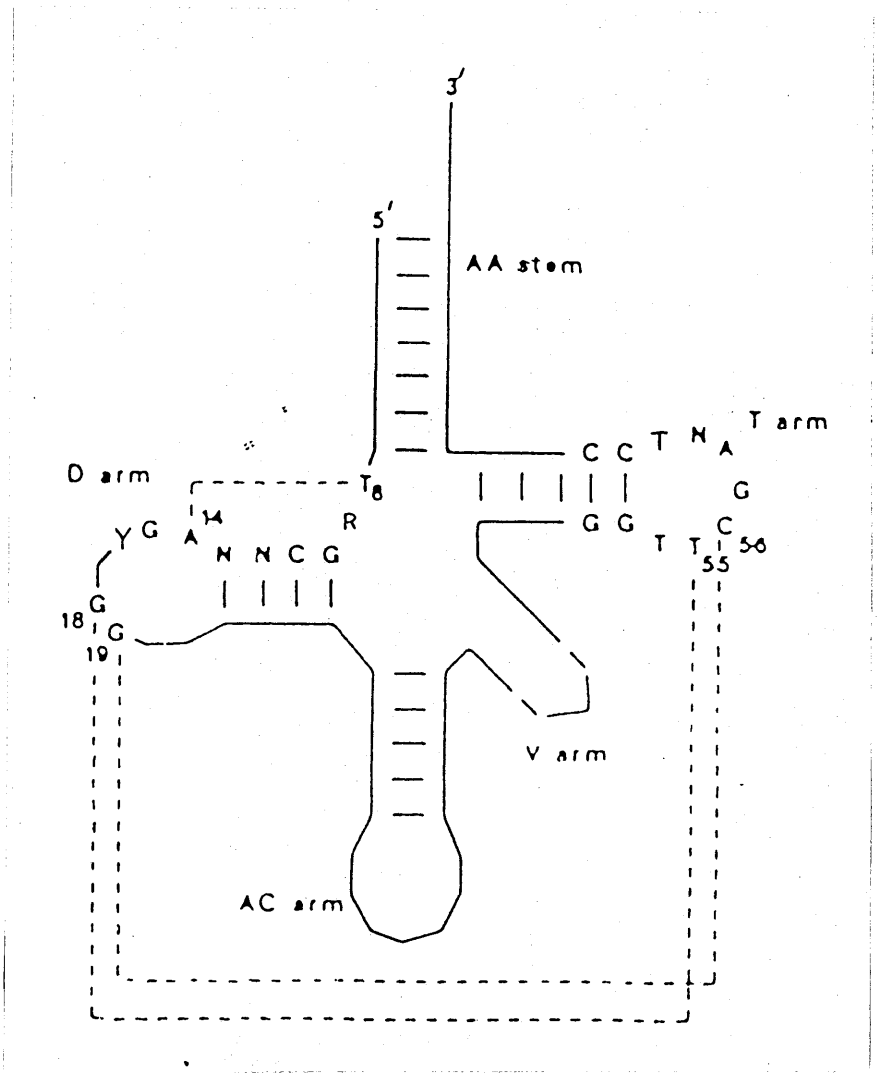
An alternative model which explains the function of tRNA gene promoter has been proposed by Hall *et al.*, (1982). This "tertiary interaction model" is based on the fact that A and B blocks encode single stranded loops (D and TψC), which are also engaged in strong mutual tertiary interactions in every tRNA molecule . According to this model the B block sequence exists in an intrastrand stem and loop conformation, when a transcription factor binds to tDNA. This binding promotes tertiary interactions which occur between sequences coding for the D and TψC loops, with the formation of G_{18}-T_{55} and G_{19}-C_{56} bp as illustrated in figure 1.3., and which are analogous to the established tertiary interactions occurring in the tRNA structure. However data reported by Newman *et al.*, (1983) do not support the importance of a "tertiary interaction model" in tRNA gene transcription; in addition McLaren and Goddard (1986b) showed that the above higher-order structure is not present in supercoiled DNA containing a human $tRNA^{Leu}$ gene.

Finally there are reports concerning transcription of eukaryotic tRNA dimeric precursors (Mao *et al.*, 1980; Schmidt *et al.*, 1980; Reyes *et al.*, 1986; Straby, 1988), or other studies which demonstrated that the natural arrangement of a tRNA pair (Hottinger *et al.*, 1985) or of a tRNA cluster (Russo *et al.*, 1987) is required for the expression of these genes. These cases cannot be explained with the transcriptional models now available. In contrast to these reports Ciliberto *et al.* (1984), constructed a dimeric and a trimeric arrangement of a $tRNA^{Pro}$ gene from *C. elegans* with no sequence in between, showing that each coding region functions as an internal promoter directing the synthesis of independent transcriptional products.

1.5.5. Termination of transcription.

Eukaryotic tRNA genes contain a cluster of four or more thymidine residues (T) in the 3'-flanking sequence which is the termination site of transcription (Garber and Gage, 1979; Koski *et al.*, 1982a; Koski and Clarkson, 1982). This cluster of T residues normally occurs within a few nucleotides of the 3'- end of tRNA genes, although exceptions have also been reported (Garber and Gage, 1979; Clarkson, 1983). Deletion of a T cluster results in read-through transcription to

Figure 1.3.
Tertiary interaction model of the eukaryotic tRNA gene promoter.



The non-coding DNA strand is shown in clover leaf form. The A block sequence extends from T_8 to G_{19} and the B block sequence from G_{52} to C_{62}. Dotted lines indicate the tertiary interactions as proposed by Hall *et al.*, (1982). (Adapted from the same paper).

the next available T cluster (Adeniyi-Jones *et al.*,1984; Allison and Hall, 1985). Shortening a T_7 cluster to T_5 downstream of a yeast $tRNA^{Tyr}$ gene results in significant read-through both *in vitro* and *in vivo* ; and shortening to T_4 abolishes termination completely (Allison and Hall, 1985). In human a $tRNA_i^{Met}$ gene produces multiple transcripts *in vitro* as a result of a series of inefficient termination sites (Tn, n<4) present downstream of the gene (Vnencak-Jones *et al.*, 1987). Finally a *Xenopus Laevis* $tRNA^{Lys}$ gene has two T_4 clusters, one located within the coding sequence and the other at some distance downstream from the gene. *In vitro* both these T_4 clusters behave as equally efficient terminators (Mazabraud *et al.*, 1987).

1.5.6. Maturation of tRNA transcripts.

As mentioned earlier, the maturation of a tRNA precursor involves excision of 5'- leader and 3'- termini sequences, addition of CCA, nucleotide modification to form dihydrouridine, pseudouridine and other derivates of all four nucleotides and in some cases splicing of intervening sequence.

There are several reports which cover the processing of tRNA precursors (for review see Deutscher, 1984). A complete study on the subject concerns the maturation of a yeast $tRNA^{Tyr}$ primary transcript (De Robertis and Olson, 1979; Melton *et al.*, 1980; De Robertis *et al.*, 1981). Injection of precursor $tRNA^{Tyr}$ into *Xenopus* oocyte nuclei or into the oocyte cytoplasm of the nucleus, revealed that the processing of the 5' and 3' ends and of the intervening sequence only occurs in the nucleus. The primary $tRNA^{Tyr}$ precursor is approximately 108 nucleotides long, containing a 5'- leader and a 3'- terminus and an intervening sequence. In *Xenopus* oocytes the 5'- leader is removed in three stages, the last of which appears to be accompanied by excision of the 3'- terminus and the addition of the 3' CCA. At the same time bases are modified in the precursor molecule which still contains the intervening sequence. The last step, which takes place in the nucleus, is the excision of the intervening sequence.

The intervening sequences of tRNA genes when present (see section 1.3.), are also present in the primary transcript. These sequences are removed from the

precursor molecule by a two step enzymatic process, involving specific excision and ligation of the resulting tRNA half molecules (Abelson, 1979). Such a process has been described in several cases (O'Farrell *et al.*, 1978; Knapp *et al.*, 1978; Standring *et al.*, 1981; De Robertis *et al.*, 1981; Willis *et al.*, 1984; Baldi *et al.*, 1986). The function of intervening sequences on tRNA gene transcription is not fully understood. Intervening sequences within yeast tRNA genes have been extended or completely deleted without affecting transcription (Johnson *et al.*, 1980; Wallace *et al.*, 1980; Carrara *et al.*, 1981). However Johnson and Abelson (1983) found that the intron of a yeast $tRNA^{Tyr}$ gene is essential for correct modification of its product. While transcription in general is not affected by the presence of an intron, the structure of the resultant intervening sequence in the transcript can affect the rate of intervening sequence removal (Laski *et al.*, 1983). In addition although complementarity between the intervening sequence and the anticodon stem and loop may not be essential, its occurence may have a kinetic effect on excision of the intervening sequence from the precursor tRNA (Swerdlow and Guthrie, 1984).

It is still not clear why tRNA goes through such a complicated maturation process. In addition different precursor tRNAs have been shown to be processed along different pathways (Rooney and Harding, 1986; Van Tol *et al*, 1987). It is possible to understand the 3'- end maturation. A functional adaptor tRNA should not contain the stretch of Ts which would interfere with the correct structure of the amino acid stem and consequently affect its aminoacylation. The reasons for transcription of the 5'- leader sequences are still unknown. It has been suggested that precursor tRNAs with extra nucleotides at their 5'- end are preventing the transport of immature transcripts from the nucleus into the cytoplasm, perhaps due to binding of a processing enzyme that produces a bulky complex (De Robertis *et al.*, 1981; Melton and Cortese, 1979). However in at least one case, the primary transcript of a phosphoserine tRNA gene lacks a 5'- leader sequence (Lee *et al.*, 1987). Finally other authors have illustrated that tRNA processing is dependent on the secondary and tertiary structure of the molecule and not affected by altering 5' and 3' flanking sequences (Castagnoli *et al.*, 1982; Mattocia *et al.*, 1983).

1.5.7. The role of flanking sequences in transcription modulation.

Although the transcription of eukaryotic tRNA genes is directed by control regions within the mature coding sequences, flanking sequences serve to modulate transcription. With only few exceptions (e.g. Shaw and Olson, 1984; Schaack and Soll, 1985; Raymond *et al.*, 1985; Marschalek and Dingermann, 1988) the role of flanking sequences has been investigated *in vitro* only. The predominant observation of such studies is that deletions into 5' flanking regions diminish and in some instances almost abolish transcription. However in a few cases the opposite effect has been found, suggesting that positive as well as negative modulator sequences are present in the 5' flanking sequences of some tRNA genes.

Positive modulators have been identified in the 5' flanking sequences of several tRNA genes (Sprague *et al.*, 1980; Larson *et al.*, 1983; Shaw and Olson, 1984; Johnson and Raymond, 1984; Raymond *et al.*, 1985; Morry and Harding, 1986; Sajjadi and Spiegelman, 1987a; Arnold and Gross, 1987; Lofquist *et al.*, 1988; Horvath and Spiegelman, 1988; Rooney and Harding, 1988). These sequences have lengths of 5 to 35 nucleotides and have been located in the first 53 nucleotides upstream of various tRNA genes. Comparison of the sequences of these positive modulators has revealed no common features. However some of these have been shown to be involved in transcription factor binding (Raymond and Johnson, 1987; Rooney and Harding, 1988; Marschalek and Dingermann, 1988). In addition to these reports, there are other tRNA genes where faithful transcription also requires 3'-flanking sequences. For the *B. mori* $tRNA_2^{Ala}$ gene, sequences which are important for transcription extend 48 nucleotides downstream of the coding sequence of the gene (Wilson *et al.*, 1985), while 35 nucleotides into the 3'-flanking sequence of a *Drosophila* $tRNA^{Arg}$ gene are required for maximal stable complex formation (Schaack *et al.*, 1983). Moreover *in vitro* studies of *Drosophila melanogaster* $tRNA_4^{Val}$ and $tRNA^{Arg}$ genes have identified both inhibitory and stimulatory sequences in their 5'-flanking regions (Schaack *et al.*, 1984; Sajjadi *et al.*, 1987b).

Potential Z-DNA sequences, present in the 5'-flanking regions of several tRNA genes appear to inhibit their transcription (Hipskind and Clarkson, 1983;

Nordheim and Rich, 1983; Santoro *et al.*, 1984; Glew *et al.*, 1986). In addition, in the 5'-flanking region of a *Drosophila* $tRNA_2^{Lys}$ gene an undecanucleotide sequence is responsible for the poor transcriptional activity of the gene (De Franco *et al.*, 1981). This sequence is well conserved within the 5'-flanking region of all *Drosophila* $tRNA_2^{Lys}$ genes, although the position relative to the coding sequence varies (Hovemann *et al.*, 1980). After mutational analysis De Franco *et al.*, (1981) observed that the position of this oligonucleotide, relative to a downstream element, influences the extent of transcriptional repression of the *Drosophila* $tRNA^{Lys}$ genes.

More striking are the results of studies where transcription of isoacceptor tRNA genes with different 5'-flanking sequences was tested. The general outcome of these experiments is that isoacceptor tRNA genes are transcribed differently due to different 5'-flanking sequences (Lofquist and Sharp, 1986; Gouilloud and Clarkson, 1986; Arnold *et al.*, 1986 and1988; Doran *et al.*, 1987 and 1988; MacKay *et al.*, 1988). Young *et al.*, (1986), reported that 5'-flanking sequence elements are responsible for tissue-specific expression of a *B. mori* $tRNA_2^{Ala}$ gene, whereas Waterson and his collegues came to the same conclusion after studying five *C. elegans* suppressor $tRNA^{Trp}$ genes (Hodgkin *et al.*, 1987; Kondo *et al.*, 1988). Other authors have demonstrated that 5'-flanking sequences modulate transcription of tRNA genes differently in different transcription systems (Raymond and Johnson, 1983; Johnson *et al.*, 1984; Schaack and Soll, 1985). In addition Drabkin and RajBhandary (1985) have shown that a human $tRNA^{Met}$ gene is not expressed in yeast cell-free extracts although a yeast $tRNA^{Met}$ gene with the same ICR is expressed efficiently. In conclusion all the studies so far have not revealed any common feature of flanking sequences by which the transcription of tRNA genes is modulated. The requirements for faithful transcription might differ from species to species or even from tissue to tissue in the same eukaryote.

CHAPTER 2

AIMS AND BACKGROUND OF THE PROJECT

2.1. Organization of $tRNA^{Glu}$ genes in the human genome.

Several recombinants containing tRNA genes have been selected from a bank of 15-20 kb fragments of a human foetal liver DNA, cloned in λ Charon 4A (Lawn *et al.*, 1978) by screening with a mixed placental tRNA[^{32}P]pCp probe (Goddard *et al.*, 1983). Two of these recombinants λht137 and λht190, generate restriction fragment profiles which are very similar, but not identical. The most notable similarity is a common 2.4 kb HindIII fragment which hybridized to the tRNA[^{32}P]pCp probe (Goddard, unpublished results). These 2.4 kb HindIII fragments were subcloned from λht137 and λht190 into the vector pAT153 to generate the recombinants pLB4 and pTC51 respectively (see section 3.4.1.). The recombinant pLB4 has been already characterized and found to contain a single $tRNA^{Glu}$ gene, notable for its high transcriptional activity in *Xenopus* oocyte nuclei and the potential of its 5'-flanking sequence for forming a tRNA-like structure (Goddard *et al.*, 1983).

After the preliminary experiments on the restriction fragment profiles of λht137 and λht190, it was expected that the 2.4 kb HindIII fragment from λht190 also contained a sequence very much resembling the one already found in the corresponding fragment from λht137. However it was not yet clear whether the fragments represented distinct portions of the human genome which have arisen by duplication, or whether they were derived from the same portion of the genome by overlapping portions of the genomic DNA cloned in λht137 and λht190. Nevertheless it was interesting to characterise the 2.4 kb HindIII fragment from

λht190, since it probably contained a second tRNA^{Glu} gene copy, and later on to construct and compare the restriction maps of both λht137 and λht190.

Another objective of the project was to estimate the copy number of the human tRNA^{Glu} gene family, as well as to compare the copies and their flanking sequences. One preliminary experiment on the subject suggested that most of the tRNA^{Glu} gene copies belong to long (several kb) DNA repeats in the human genome.

2.2. Transcription studies on human tRNA^{Glu} genes.

As has been mentioned earlier, the tRNA^{Glu} gene characterized from λht137 was found to be of a remarkable high transcriptional activity, when micro-injected into *Xenopus* oocyte nuclei, producing more tRNA than a cluster of seven active *Xenopus* tRNA genes (Goddard *et al.*, 1983). The same gene was subcloned into M13mp9 vector and the recombinant derived (MtGlu6; see section 3.4.3.) was suitable for further studies in an *in vitro* homologous transcription system, using HeLa S3 extracts. If recombinant pTC51 (see section 2.1.) contains another tRNA^{Glu} gene copy, its transcriptional efficiency will be compared with the one of the first gene copy, using an *in vitro* transcription system.

The 5'-flanking sequence preceding the characterized tRNA^{Glu} gene has a potential for forming a tRNA-like structure at residues 292-347 (Goddard *et al.*, 1983). The fact that the gene appears to be of a remarkable high transcriptional efficiency, in conjuction with the possibility that the tRNA-like structure could well allow recognition and binding of RNA polymerase III and transcription factors, led to the obvious question whether this structure affects the transcriptional efficiency of the gene or not. Several deletion mutants of recombinant MtGlu6, which lacked part or all of the tRNA-like structure, were therefore constructed and partly characterized (Goddard, unpublished results; see also section 3.4.3.). The transcriptional efficiency of these "MtGlu6 deletion mutants" has to be compared in a homologous transcription system with the "wild type" recombinant MtGlu6. Preliminary experiments using a heterologous

transcription system in which two deletion mutants were tested in parallel with MtGlu6, showed that part of the 5'-flanking sequence is required for the maximum $tRNA^{Glu}$ gene transcription (Goddard *et al.*, 1985).

CHAPTER 3

MATERIALS AND METHODS

3.1. Materials, list of suppliers.

Unless otherwise specified all chemicals were Analar grade, supplied by BDH Chemicals, Poole, Dorset, or Fisons Scientific, Loughborough, Leics. All radiochemicals were provided by Amersham International plc. Where chemicals or equipment were obtained from other sources, this is indicated in the text. A list of suppliers of special reagents is given below:

Aldrich Chemicals Co., Gillingham, Dorset

Amersham International plc., Amersham, Buckinghamshire

Anglian Biotechnology Ltd., Colchester

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

Gibco/BRL (Bethesda Research Laboratories) Ltd., Paisley

Kodak Ltd., Kirby, Liverpool

Pharmacia Ltd., Milton Keynes

Sigma Chemicals Company, Poole, Dorset

3.2. Media and Buffers.

All media and solutions used in the handling of nucleic acids were sterilised wherever possible by autoclaving for 20 min, at 15 p.s.i., otherwise they were filter sterilized.

3.2.1. Bacterial growth media.

L-broth (per litre) : 10 g bactotryptone (Difco)
5 g yeast extract (Difco)
10 g NaCl
Adjusted to pH 7.2 with NaOH.

2 x YT (per litre) : 16 g bactotryptone
10 g yeast extract
5 g NaCl

LAM plates (per litre) : 960 ml L-broth
15 g agar
10 ml 20% maltose
10 ml 1 M $MgSO_4$

LAM top (per litre) : 960 ml L-broth
10 g agar
10 ml 20% maltose
10 ml 1 M $MgSO_4$

LBT plates (per litre) : 980 ml L-broth
15 g agar
7.5 mg tetracycline

minimal glucose plates : 5.8 g Na_2HPO_4
(per litre) 3.0 g KH_2PO_4
1.0 g NH_4Cl
0.5 g NaCl
10 ml 1M $MgSO_4$
1 ml 2mg/ml vitamin B1
15 g agar
10 ml glucose 20%

H-plates (per litre)	:	10 g	bactotryptone
		8 g	NaCl
		12 g	agar
H-top (per litre)	:	10 g	bactotryptone
		8 g	NaCl
		8 g	agar

3.2.2. Buffers.

λ diluent	:	50 mM	Tris-HCl, pH 7.5
		100 mM	NaCl
		10 mM	$MgSO_4$
		0.1%	gelatin
PBS mix	:	8/1/1	mix of solutions A, B, C.
Solution A	:	170 mM	NaCl
		0.5 mM	KCl
		10 mM	Na_2HPO_4
		2 mM	KH_2PO_4
		Adjusted to pH 7.2	
Solution B	:	7 mM	$CaCl_2.2H_2O$
Solution C	:	5 mM	$MgCl_2.6H_2O$
10 x TBE	:	0.89 M	Tris
		0.89 M	boric acid
		0.025 M	EDTA
		It should give pH 8.3	
20 x SSC	:	3 M	NaCl
		0.3 M	sodium citrate

TE : 10 mM Tris-HCl
1 mM EDTA
Adjusted to pH 7.6 with HCl

3.3. Storage of bacteria; list of bacterial strains.

Most bacteria could be kept for at least one month at 4° C on tightly sealed L-agar or H-plates.

The following *E.coli* strains were used in this study:

A. Strains used for bacteriophage M13 recombinants.

JM105: thi, rps L, end A sbc B15, hsd R4, Δ(lac-pro AB), [F', tra D36, pro AB, lac IqZΔM15] (Yannish-Perron *et al.*, 1985)

JM109: rec A1, end A1, gyr A96, thi, hsd R17, Sup E44, rel A1, λ-, Δ(lac-pro AB), [F', tra D36, pro AB, laclqZΔM15] (Yannish-Perron *et al.*, 1985)

B. Strains used for bacteriophage λ recombinants.

Q358: hsd R_k-, hsd M_k-, Sup F, ø80^r, rec A$^+$ (Karn *et al.*, 1980)

LE392: F$^-$, hsd R514(r_k-, m_k-), sup E44, sup F58, lac Y1 or Δ(lac IZY)6, gal K2, gal T22, met B1, trp R55, λ$^-$ (Murray, 1977)

C. The strain used as the host for the plasmids pTC51 and pLB4 was HB101.

3.4. DNA vectors and recombinants used in this study.

DNA vectors and recombinants were stored in TE (see section 3.2.2.) in tight fitting Eppendorf tubes, at -20° C. DNA stored in this way remains stable for several years and could be used to retransform bacterial host cells should the need arise.

3.4.1. pLB4 and pTC51.

Plasmid recombinant pLB4 was derived by subcloning a 2.4 kb HindIII fragment from λht137 (see figure 4.8.) into pAT153 vector (Twigg and Sherratt, 1980). The plasmid had already been characterised and found to contain a human $tRNA^{Glu}$ gene (Goddard *et al.*, 1983). pLB4 is a gift from L. Brown.

Plasmid recombinant pTC51 was derived by subcloning a 2.4 kb HindIII fragment from λht190 (see figure 4.8.) into pAT153 vector (Twigg and Sherratt, 1980). The insert of the plasmid, was found to hybridize a periodate treated tRNA[^{32}P]pCp probe (Goddard, unpublished results). pTC51 is a gift from T. Carr. The construction of both pLB4 and pTC51 is shown in figure 3.1.

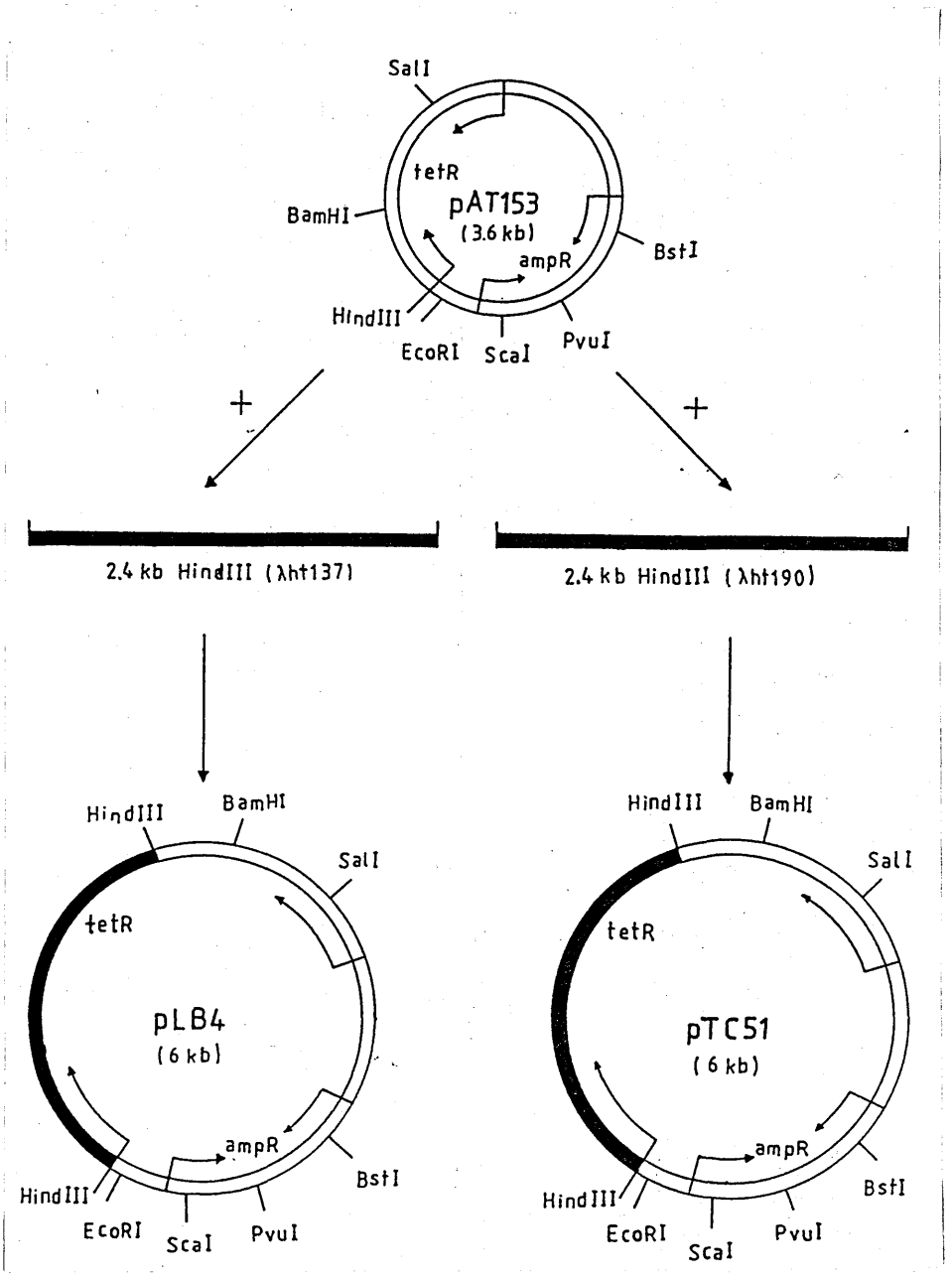
3.4.2. M13mp18 and M13mp19 bacteriophage DNA vectors.

The bacteriophage vectors M13mp18 and M13mp19 (Yannish-Perron *et al.*, 1985) were used for subcloning DNA fragments from pTC51 into the multiple cloning sites of both vectors' replicative forms (see table 4.1.). These fragments were then sequenced by the use of the "chain termination" method (Sanger *et al.*, 1980). M13mp18 and M13mp19 vectors are shown in figure 3.2. Other M13 vectors used in this work (M13mp9, M13 mp12) differ only in the smaller number of restriction sites in the polylinker multiple cloning site.

3.4.3. MtGlu6 and "MtGlu6 deletion mutants".

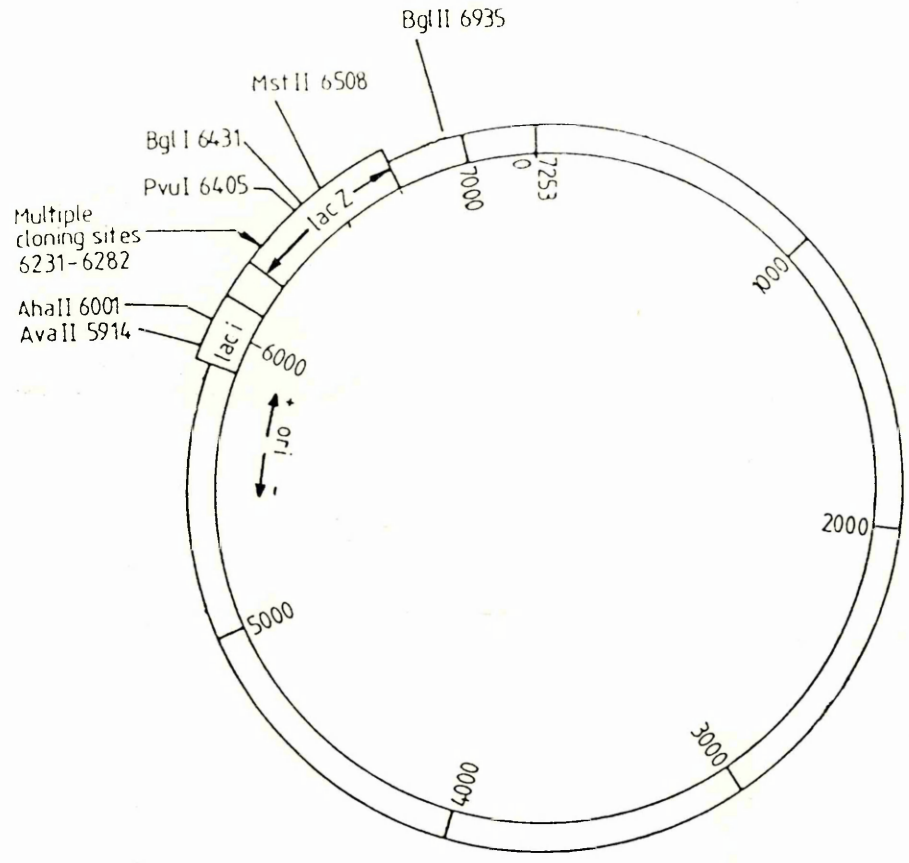
As has been mentioned in section 2.1., recombinant λht137 contains a single $tRNA^{Glu}$ gene (Goddard *et al.*, 1983). This $tRNA^{Glu}$ gene, cloned as a 937 bp fragment in M13mp9 for sequence analysis, was subcloned further as a 481 bp EcoRI-NaeI fragment into the same vector. The latter recombinant, called MtGlu6, contains the $tRNA^{Glu}$ gene as well as 355 nucleotides and 54 nucleotides of its 5' and 3' flanking sequences respectively (see figure 3.3.). MtGlu6 was a gift from Dr. J.P. Goddard.

Figure 3.1.
Construction of plasmids pLB4 and pTC51.



Plasmids pLB4 and pTC51 derived by subcloning 2.4 kb HindIII fragments from λht137 and λht190 respectively into pAT153 vector. The size of both recombinants is 6 kb . The restriction sites, as well as the Tetracycline resistance and the Ampicillin resistance present in pAT153 are shown.

Figure 3.2.
Bacteriophage DNA vectors M13mp18 and M13mp19.



Multiple cloning sites in M13 mp18:

6230 6289

AccI
HincII
EcoRI SstI KpnI BamHI XbaI SalI PstI SphI HindIII

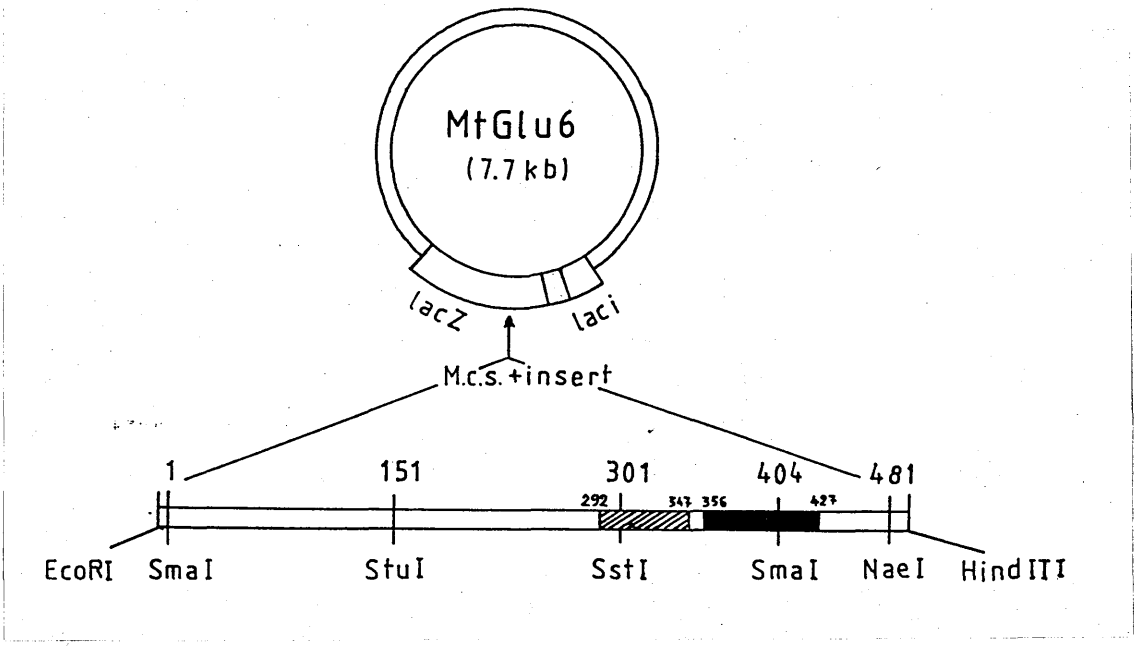
GAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGC

SmaI
XmaI
TaqI

The bacteriophage DNA vectors M13mp18 and M13mp19 were used for the nucleotide sequence analysis by the Sanger chain termination method. These are single stranded circular molecules (single stranded in viral DNA, double stranded in replicative form)7253 bases in length and differ only in the orientation of the 54 bp polylinker that they carry. The polylinker includes 10 recognition sites for 13 different restriction enzymes. The multiple cloning sites in the orientation present in M13mp18 are shown below the map. In M13mp19 the same sites are present in the opposite orientation. The map shows a number of restriction sites, the position of the *lac* gene fragment and the origins of plus and minus strand replication.

Figure 3.3.

Restriction map of MtGlu6. List of MtGlu6 deletion mutants.



Mutant	Deleted sequence of MtGlu6	Sequence at the position of deletion
MtGluD5	152 - 304	...$AGG_{151-305}GGA$...
MtGluD7	152 - 311	...$AGG_{151-312}GTC$...
MtGluD20	152 - 343	...$AGG_{151-344}CCC$...
MtGluD33	296 - 303	...$CCC_{295-304}GGG$...
MtGluD44B	273 - 330	...$GAC_{272-331}GGG$...
MtGluD47	152 - 318	...$AGG_{151-319}CTG$...
MtGluD92	294 - 366	...$CTC_{293-367}CTA$...
MtGluD104C	175 - 342	...$GTG_{174-343}TCC$...

Recombinant MtGlu6 contains a 481 bp SmaI-NaeI fragment from pLB4 subcloned into the vector M13mp9. The size of the recombinant is 7.7 kb. The restriction sites in the multiple cloning region of the M13mp9 used in this study, the sites present at the 481 bp insert as well as the position of the $tRNA^{Glu}$ gene (residues 356-427) and the tRNA-like structure (residues 292-347) are shown. A list of the "MtGlu6 deletion mutants" together with the sequence which has been deleted in every case, are shown as well.

The 5'-flanking sequence preceding the above mentioned $tRNA^{Glu}$ gene, has a potential for forming a tRNA-like structure at residues 292-347 (Goddard *et al.*, 1983). Several mutants of recombinant MtGlu6 have been constructed, taking advantage of a unique SstI restriction site (residues 297-302) which allowed access to the tRNA-like structure region, for sequential removal of nucleotides by partial digestion with Bal31 exonuclease. These partially characterized "MtGlu6 deletion mutants", where part or all of the tRNA-like structure is missing, were gifts from Dr. J.P. Goddard. A list of "MtGlu6 deletion mutants" is given in figure 3.3.

3.4.4. MtGlu(282/419) and MtGlu(301/1).

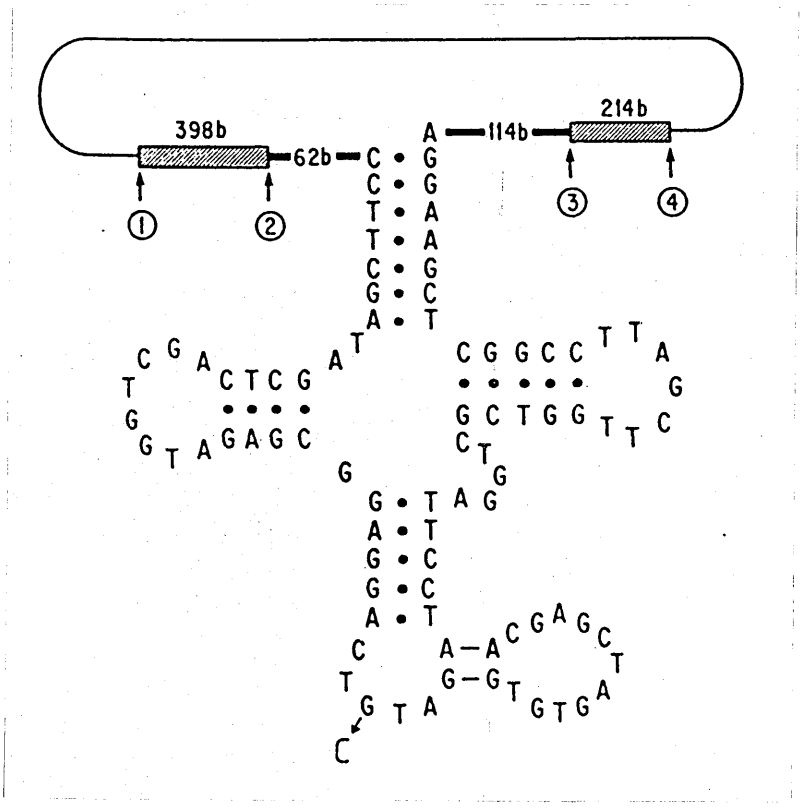
MtGlu(282/419) and MtGlu(419/282) were derived by subcloning a 137 bp HaeIII fragment from MtGlu6 by blunt end ligation in both orientations into SmaI digested M13mp18. The 137 bp insert (residues 282-419 of MtGlu6) represents 63 nucleotides of the $tRNA^{Glu}$ gene found in λht137, as well as 74 nucleotides of its 5' flanking sequence (Goddard *et al.*, 1983). MtGlu(282/419) was mainly used for preparation of "$tRNA^{Glu}$ gene" probes (see section 3.5.9.2.); it was a gift from A. McLaren.

MtGlu(301/1) were derived by subcloning a 301 bp SmaI-SstI fragment from pTC51 (see section 3.4.1.) into M13mp19. The 301 bp insert represents part of the 5'-flanking sequence of the $tRNA^{Glu}$ gene found in λht190 (see sections 4.1.1. and 4.1.2.). The recombinant was used for preparation of probe"3" (see section 3.5.9.2.).

3.4.5. M13tT-(su^+).

M13tT-(su^+) contains a 263 bp HaeII-HhaI fragment of *X. laevis* DNA, derived from recombinant pSV-tT (Laski *et al.*, 1982a), flanked by 398 bp (to its 5'-end) and 214 bp (to its 3'-end) SV40 DNA sequences, and cloned into the HindIII site of M13mp5. The 263 bp HaeII-HhaI fragment contains an amber suppressor $tRNA^{Tyr}_{GTA}$ gene, as well as 62 and 114 nucleotides of its 5' and 3' flanking sequences respectively (see figure 3.4.). This suppressor

Figure 3.4.
Structure of M13tT-(su$^+$).



M13tT-(su$^+$) contains a *X.Laevis* 263 bp HaeII-HhaI DNA fragment derived from recombinant pSV-tT (Laski *et al.*, 1982a), flanked by 398 bp and 214 bp SV40 DNA sequences and cloned into the HindIII site of M13mp5. Description of recombination sites are as follows : 1) the HindIII site of SV40 (m.u. 0.65) joined with the HindIII of the M13mp5 polylinker; 2) the HhaI site of *X. Laevis* DNA joined with the Hha I site of SV40 (m.u. 0.72); 3) the HaeII site of *X.Laevis* DNA joined with the same site of SV40 (m.u. 0.82); 4) the HindIII site of SV40 (m.u. 0.86) joined with the Hind III site of M13mp5 polylinker. The sequence of the coding region of the tRNA^{Tyr} gene, its intervening sequence and the site of suppression are shown.
(Adapted from Laski *et al.*, 1982b).

(su$^+$)tRNA gene, which has a thirteen nucleotide intervening sequence in the anticodon loop, starting after position 37 (Laski *et al.*, 1982a and 1982b). M13tT-(su$^+$) was used as a DNA competitor in the *in vitro* transcription assays, which are presented in section 4.3.4.5.; M13tT-(su$^+$) was a gift from Dr. A. Fire.

3.5. Methods.

3.5.1. Large scale preparation of plasmid DNA.

The method used was based on the method of Birnboim and Doly (1979). An overnight culture was prepared from a single colony of transformed bacteria in L-broth (see section 3.2.1.) supplemented with the appropriate antibiotic (in the case of pLB4 or pTC51, 20µg/ml of ampicillin were used; see figure 3.1.). The main culture was set up by inoculating 2x10 ml of overnight culture into 2x1000 ml of L-broth in 2x2 litre conical flasks. The bacteria were shaken at 37° C until A_{610} reached 0.8. Chloramphenicol (25mg/ml in 50% ethanol) was added to a final concentration of 165 µg/ml and incubation continued for a further 16 to 20 hr.

Bacteria were then harvested by centrifugation at 5,000 rpm for 10 min at 4° C and resuspended in 9.5 ml of a solution containing 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl (pH 8.0) and 0.5 ml lysozyme (Sigma grade I; 40 mg/ml in the same solution). After 30 min on ice, 20 ml of a solution of 0.2 N NaOH, 1% SDS were added, and the mixture left for a further 5 min on ice. Finally, 15 ml of 3 M sodium acetate (pH 4.8) were added and the mixture left standing for 1 hr on ice. The high molecular weight DNA and bacterial debris were removed by centrifugation at 30,000 rpm for 30 min in a Beckman Ti60 rotor at 4° C. Total DNA was precipitated from the resultant supernatant by the addition of 0.6 volume of isopropanol and left standing for 10 min at room temperature. DNA was sedimented by centrifugation at 8,000 rpm for 15 min, at room temperature.

The DNA was dissolved in 30 ml TE (see section 3.2.2.), 28.9 g of CsCl and 1.8 ml of ethidium bromide (10 mg/ml) were added. The solution, clarified by centrifugation in a Beckman "table-top" centrifuge at 1,500 rpm for 30 min,

was then carefully transferred to sealable tubes and centrifuged at 48,000 rpm for 16 hr and 20° C in a VTi50 rotor of a Beckman ultracentrifuge.

The DNA was visualised under u.v. (long wave) illumination, and two bands were usually seen. The upper band consisted of linear bacterial DNA and nicked plasmid DNA, while the lower band consisted of closed circular DNA. The plasmid DNA was collected after piercing the tube with a 18g needle just below the band. A second CsCl centrifugation was usually performed using sealable tubes and centrifugation at 63,000 rpm overnight in a VTi65 rotor. The major band containing plasmid DNA was removed as described above.

Ethidium bromide was removed by repeated extraction with an equal volume of n-butanol; CsCl was removed by dialysing the sample against three changes (1hr each) of 500 ml TE. Boiled pancreatic RNase A (Boehringer, grade I) was added to a final concentration of 10 μg/ml and samples were then left standing at room temperature for 30 min. The DNA solution was then extracted with phenol/chloroform twice (see section 3.5.5.1.), diluted with four volumes of TE and precipitated by the addition of 2.5 volumes of ethanol and 0.1 volumes of 5 M ammonium acetate. DNA was sedimented by centrifugation at 10,000 rpm for 20 min at 4° C, washed with 70% ethanol, briefly dried under vacuum and finally redissolved in 500 μl TE.

The concentration of DNA was determined by measuring the A_{260}, using the assumption that a 50 μg/ml solution of DNA has an A_{260}=1.0, when measured in a spectrophotometer cell with a 1.0 cm light-path. A scanning of A_{220}-A_{290} was also performed to confirm the absence of protein and phenol from DNA. Finally a sample of 0.5 μg was subjected to electrophoresis through a 1% agarose gel to check the quality of the preparation.

3.5.2. Large scale preparation of bacteriophage λ DNA.

A single colony of the bacterial host *E.coli* Q358 (see section 3.3.), was inoculated into an overnight culture of L-broth (see section 3.2.1.), supplemented with 10 mM $MgSO_4$ and 0.2% maltose. A suitable dilution of phage,

giving 10 to 100 plaques, was adsorbed onto 200 µl of the overnight culture of bacterial host, mixed with 3 ml of LAM-top (see section 3.2.1.), poured onto a LAM-plate (see section 3.2.1.) and inoculated overnight.

A single plaque was removed from the plate, added to 200 µl of an overnight culture of bacterial host and left standing at room temperature for 20 min. The culture was then transferred to a fresh flask containing 20 ml L-broth supplemented with 10 mM $MgSO_4$ and 0.2% maltose. The flask was shaken at 37° C until lysis occured (usually between 4 to 6 hr), when 1 ml of chloroform was added. After 10 min of shaking, the upper layer was sedimented by centrifugation in a "table-top" centrifuge for 20 min and the supernatant which could be stored at 0° C was transferred to a fresh tube. The supernatant was titred and was generally found to be approximately 10^{10} pfu/ml.

Two 500 ml batches of L-broth supplemented with 10 mM $MgSO_4$ and 0.2% maltose, were inoculated with 2 x 5.0 ml of overnight bacterial culture. The culture was grown until A_{610} reached 0.3 when a total of 5 x 10^{10} pfu/ml of phage was added. (That gives a ratio, pfu/*E.coli* cells, 1/200.) The infected culture was shaken until lysis occured (usually about 3.5 hr) and then 5.0 ml of chloroform was added to each flask and shaking continued for a further 10-15 min.

The lysed cultures were decanted and bacterial debris was removed by centrifugation at 4,000 rpm for 20 min. Solid NaCl was added to 2%, solid polyethylene glycol (PEG-6000 Serva) was added to 8% and the culture was left standing overnight at 4° C to precipitate the phage particles.

The phage was sedimented by centrifugation at 6,000 rpm for 30 min and carefully resuspended in 20 ml of lambda diluent (see section 3.2.2.). Solid CsCl was added, giving a final concentration of 0.71 g/ml. After a clarifying centrifugation in the "table-top" centrifuge for 30 min, the solution was transferred into sealable tubes and centrifuged at 48,000 rpm and 20° C overnight using a VTi50 rotor of a Beckman ultracentrifuge.

The white phage band was removed (as described in section 3.5.1.) and a second CsCl centrifugation was performed at 63,000 rpm and 20° C overnight. The phage was then dialysed against three changes (1hr each) of 500 ml buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 10 mM $MgSO_4$) to remove CsCl. Boiled

pancreatic RNase A was added to a final concentration of 10 µg/ml and samples were left standing at room temperature for 30 min. DNA was extracted with phenol/chloroform twice, precipitated with ethanol, briefly dried under vacuum, resuspended in 500 µl TE and the concentration determined as described in section 3.5.1.

Note: By using the above method, DNA from several different λ recombinants was routinely prepared without difficulty. However in the case of λht190 it became impossible to purify an appreciable amount of DNA, due to the low titre obtained (10^8 pfu/ml maximum) during the phage amplification step. In order to overcome the "low yield" problem, several variations of the method were attempted: different *E.coli* host cells (LE392, see section 3.3.) were tried or, different ratios during the phage amplification step (pfu/*E.coli* cells from 1/5 up to 1/1000), or even a solid amplification was attempted, by using a series of LAM plates instead of the L-broth liquid amplification (Maniatis *et al.*, 1982). Unfortunately none of these variations increased the λht190 yield significantly.

3.5.3. Large scale preparation of replicative form bacteriophage M13 DNA.

Competent cells were prepared from *E.coli* JM105 or JM109 strains as in section 3.5.11.2. 300 µl of competent cells were transformed with 10 ng of the appropriate M13 DNA and plated out as described in section 3.5.11.2. Transformed cells yielded blue plaques if DNA used for transformation was the M13 vector, or white plaques if DNA used for transformation was an M13 recombinant (see section 3.5.11.2.).

A single plaque was transferred to a conical flask containing 20 ml of 2 x YT (see section 3.2.1.) and 200 µl of an overnight culture of *E. coli* bacterial host cells. The flask was shaken at 37° C for 4 or 5 hr, when extruded M13 phage was separated from the cell debris by centrifugation at 3,000 rpm for 10 min at 4° C. The supernatant was transferred to a fresh tube which could be stored at 0° C. The supernatant was titred and generally found to be approximately 10^{10} pfu/ml.

Two 500 ml batches of 2 x YT were inoculated with 2 x 5.0 ml of overnight bacterial culture. The culture was grown until A_{610} reached 0.3 and then a total

of $5x10^{10}$ pfu of M13 phage added. The infected cultures were shaken for 5 hr, when they were centrifuged at 9,000 rpm for 20 min at 4° C. The pellets, containing infected cells harbouring the replicative form of M13 were washed with TE, recentrifuged and stored overnight at -20° C.

To the cells, resuspended in 9 ml buffer containing 25% sucrose and 50 mM Tris-HCl pH 7.4, were added 0.9 ml of a fresh solution of lysozyme (20 mg/ml in 50 mM Tris-HCl, pH 7.4) and 3.6 ml of 0.25 M EDTA, pH 8.0. After 5 min on ice, 14.5 ml of "lysis mix" (1% Triton x100; 50 mM Tris-HCl, pH 8.0; 60 mM EDTA) were added and the resulting highly viscous solution was centrifuged at 33,000 rpm for 30 min at 8° C, in a Beckman Ti60 rotor.

The supernatant was carefully removed, the volume adjusted to 27.5 ml with TE, 26.13 g CsCl dissolved and 2.75 ml ethidium bromide (10 mg/ml) added. After a clarifying centrifugation in a "table top" centrifuge at 1,500 rpm for 30 min, the solution was transferred into sealable tubes and centrifuged at 48,000 rpm for 16 hr at 20° C in a VTi50 rotor of a Beckman ultracentrifuge.

The M13 RF DNA band was collected as described in section 3.5.1. and recentrifuged at 63,000 rpm and 20° C overnight in a similar CsCl gradient. Ethidium bromide and CsCl were removed (see section 3.5.1.), and the DNA after phenol/chloroform extraction (see section 3.5.5.1.) was precipitated with ethanol (see section 3.5.5.2.). DNA was sedimented at 10,000 rpm for 20 min at 4° C, briefly dried under vacuum and resuspended in 500 µl TE. The concentration of DNA was determined by measuring the A_{260}, as described in section 3.5.1. Finally a sample of 0.5 µg was subjected to electrophoresis through a 1% agarose gel, to check the quality of the preparation.

3.5.4. Large scale preparation of high molecular weight human placentae DNA.

Fresh human placentae, lacking the umbilical cord (kindly provided by Yorkhill Infirmary, Glasgow), were immediately chopped into small pieces; some were frozen on dry ice and stored at -70° C, the remainder washed twice with cold PBS mix (see section 3.2.2.), and ground in two volumes of cold PBS mix ,using a Philips blender, for 2-3 min. The finely ground placentae were homogenised with

8-10 strokes of a loose fitting pestle in a Dounce homogeniser, in the presence of 10 ml of "guanidinium chloride mix" (8 M guanidinium chloride, 20 mM sodium acetate, 50 mM EDTA, 0.7 M 2-mercaptoethanol; pH 7.0). To the homogenate an equal volume of isopropanol was added to precipitate DNA, which was then spooled onto a glass rod. The DNA was washed three times with 70% ethanol, air dried for 2 min, dissolved in 10 ml of buffer A (10 mM NaCl; 0.5% SDS; 10 mM EDTA; 10 mM Tris-HCl, pH 8.0), containing 50 µg/ml of Proteinase K (Boehringer) and incubated at 37° C overnight.

DNA was extracted with phenol/chloroform three times by gentle rolling, and dialysed against four changes (4 hr each) of 200 ml TE, to remove the remnants of phenol.

Boiled pancreatic RNase A (Boehringer, grade I) was added to DNA to 50 µg/ml and the sample incubated for 30 min at 37° C. DNA was then phenol/chloroform extracted and dialysed as above. DNA was left standing for an additional 24 hr at room temperature in the same buffer, to ensure its complete dissolution; finally it was stored at -20° C.

The concentration of DNA was determined by measuring the A_{260}, as described in section 3.5.1. A sample of 0.5 µg was subjected to electrophoresis through a 0.6% agarose gel, in the presence of high molecular weight markers, to check the size of the DNA, which was found to be greater than 50 kb.

3.5.5. Extraction and precipitation of DNA.

3.5.5.1. Phenol/Chloroform extraction.

Phenol was redistilled before use, saturated with TE and stored at room temperature. Extraction with phenol/chloroform was performed as follows: an approximately equal volume of phenol was added to the DNA solution to be extracted, mixed and centrifuged for 2 min; the upper aqueous layer transferred to a fresh tube and the extraction with phenol was repeated; two volumes of chloroform were added to the aqueous phase, mixed and centrifuged for 1 min; the aqueous layer was transferred to a fresh tube and the extraction with chloroform repeated.

3.5.5.2. Ethanol precipitation.

The ethanol was of the absolute alcohol 100 grade and was stored at -20° C. To the solution of DNA, 0.1 volume of 5 M ammonium acetate and 2.5 volumes of cold ethanol were added, mixed and placed at -20° C overnight or in dry ice for 15 min. Samples were then centrifuged at 10,000 rpm for 10 min at 0° C, when the precipitated DNA was washed with 70% ethanol and briefly dried under vacuum. DNA was usually stored in TE at -20° C.

3.5.6. DNA digestion with restriction endonucleases.

Restriction digestions were routinely carried out in 1.5 ml Eppendorf tubes in the presence of the appropriate buffer. One unit of enzyme activity is defined as the amount of enzyme required to digest 1 μg of λ DNA to completion in 1 hr. However a several fold excess of enzyme per digest was usually added to ensure complete digestion.

A typical reaction mixture contained 1 μg of DNA, the appropriate restriction enzyme buffer (usually low, medium or high salt buffers; see Maniatis *et al.*, 1982, unless otherwise recommended by the manufacturer) and 5 units of the desired restriction enzyme in a final volume of 25 μl. The mixture was incubated at 37° C (unless otherwise recommended by the manufacturer) for 2.5 hr , when an additional 5 units of the same restriction enzyme were added and the incubation extended for another 2.5 hr. The completion of the digestion was monitored by electrophoresis of a small aliquot, through a 1% agarose gel (see section 3.5.7.1.). If the restriction enzyme used required Mg^{2+} and/or was sensitive to denaturation at high temperature (>60° C), then the digestion was stopped by the addition of 20 mM EDTA, pH 7.5 and/or a further incubation at 68° C for 10 min; if not the digestion mixture was stored at -20° C.

In order to test for the completion of digestion of genomic DNA, the following variation of the method described was performed: 0.1 volumes of the initial digestion mixture containing 0.1 μg of genomic DNA and 0.5 units of enzyme were transferred to a separate tube, to which 0.5 μg of λ DNA and distilled H_20 to 10 μl were added. The sample containing the λ DNA was incubated in parallel with

the sample containing the genomic DNA (i.e. for 2.5 hr at 37° C). The completion of the digestion was monitored by electrophoresis of the λ DNA digestion mixture through a 0.6% agarose gel, when a restriction fragments profile of λ DNA typical to the enzyme used was observed. A typical result of such procedure is shown in figure 4.9. In addition, as has been mentioned above, to the sample containing the genomic DNA digestion mixture, an additional 5 units/μg of the corresponding restriction enzyme were added and the digestion was extended for another 2.5 hr.

In multiple digestions where different buffers were required, the digestion which required the lower ionic strength was carried out first, then the ionic conditions were adjusted, the second enzyme added and the incubation continued.

3.5.7. Separation of DNA fragments by agarose gel electrophoresis.

3.5.7.1. DNA electrophoresis in agarose gels.

Agarose gels were routinely used to separate DNA fragments of size range between 0.3 kb and 20 kb. They were prepared by dissolving agarose (Sigma, type II) in 1xTBE (see section 3.2.2.) to the desired concentration (0.6-2.0%; Maniatis *et al.*, 1982) by heating. Ethidium bromide was added to 10 μg/ml and the gel poured onto a horizontal tank (usually 14x12 cm, BRL model H5). Wells formed by insertion of a comb into the gelling agarose, were usually 0.6x0.3x0.1 cm. DNA samples were mixed in 2:1 ratio with "loading buffer" (10% glycerol, 25% ficoll 400, 1 mM EDTA and 0.25% bromophenol blue), loaded into the wells and electrophoresed at 100 volts (20-40 mA). The electrophoresis time was varied according to the requirement of the particular experiment.

After electrophoresis the gel was carefully immersed in 200 ml of H_2O for 20 min, to remove excess ethidium bromide. The DNA fragments were visualised by the use of a long wave length u.v. transilluminator (U.V. Products, Inc.) and photographed (if desired) by using a Polaroid camera (Cu-5 hand camera), with type 665 positive/negative film.

The sizes of the restriction fragments were determined by comparison with

DNA marker fragments of known size, subjected to electrophoresis alongside the fragments of unknown size. The distances between the well and the positions where DNA fragments of known sizes had travelled, were measured and plotted on semi-log graph paper, as distance travelled (mm), against log size of DNA (kb). Similarly, the distance travelled by DNA fragments of unknown sizes were then measured and their sizes were determined from the standard curve.

3.5.7.2. Recovery of DNA from low melting agarose.

Low melting agarose (BRL) gels were prepared exactly as agarose gels, but they were electrophoresed in a cold room.

The DNA band of interest was excised and transferred to an Eppendorf tube. Three volumes of TE were added and the sample was placed into a 65° C waterbath for 15 min to melt the agarose. The sample was then phenol/chloroform extracted (see section 3.5.5.1.) to remove agarose and the DNA was ethanol precipitated (see section 3.5.5.2.).

3.5.8. DNA transfer to nylon membrane (Southern blotting).

This method is based on that of Southern (1975). After electrophoresis the agarose gel was immersed in 400 ml of 0.25 N HCl for 15 min at room temperature, with gentle agitation. (This "depurination" step was performed for more efficient transfer of large DNA fragments.) The gel was then incubated in 400 ml of denaturing solution (0.4 N NaOH, 0.6 M NaCl) for 30 min at room temperature with gentle agitation and the gel was finally immersed in 400 ml of neutralising solution (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.5) for 30 min at room temperature with gentle agitation.

The arrangement of the blotting components was as follows: 500 ml of 10xSSC (see section 3.2.2.) were added to the buffer reservoirs; two strips of Whatman 3 MM paper were connected over the solid support into both reservoirs from the wicks; the gel was placed onto the bridge (ideally the same width); a Saranwrap sheet covered the whole apparatus, leaving a window of the same size

and position of the gel, in order to avoid side capillary transfer; a nylon filter (Gene Screen Plus, Du Pont Company, Boston; cut to size) was carefully placed onto the gel after previously wetting in 10xSSC; six sheets of 3 MM papers were further placed on the top; an excess stack of absorbent pads (cut to size) was then placed on the top and finally the whole system was compressed using a glass plate and a 1 kg weight. Transfer of DNA was allowed to proceed for 16-20 hr at room temperature.

After blotting the nylon filter was marked for orientation, washed in an excess of 0.4 N NaOH for 1 min, immersed in an excess of 0.2 M Tris-HCl (pH 7.5), 2 x SSC, placed onto two pieces of cling film and left standing on the top of an u.v. transilluminator (DNA face up) for 3 min, to bind the DNA to the membrane. The membrane was then stored moist at 0° C.

3.5.9. Radiolabelling DNA fragments.

3.5.9.1. Random oligonucleotide-primed synthesis from DNA fragments.

This method (Feinberg and Vogelstein, 1983 and 1984) was used to make probes for the construction of the restriction map of λht137 experiments (see section 4.1.3.2.).

λht137 DNA was cleaved with an appropriate restriction enzyme and the fragments derived were electrophoretically separated in a low melting point agarose gel, as described in section 3.5.7.2.

The labelling reaction was carried out as follows: to 100 ng of DNA were added 10 µl of oligolabelling buffer (OLB buffer, see below), 20 µg of bovine serum albumin, 50 µCi of α-[^{32}P]-dCTP (Amersham No: PB10205, 3000 Ci/mmol), 3 units of large fragment DNA polymerase I (Klenow; BRL) and distilled H_2O to 50 µl. The reaction was carried out at room temperature overnight and stopped by the addition of 50 µl of a solution containing 20 mM NaCl; 20 mM Tris-HCl, pH 7.5; 2 mM EDTA; 0.25% SDS and 1 µM dCTP. Purification of the labelled DNA was not performed.

Oligolabelling buffer was made up from the following components: 2 volumes

of 0.018% 2-mercaptoethanol, 0.5 mM of the remaining unlabelled dNTPs in 1.25 M Tris-HCl (pH 8.0) and 0.125 M $MgCl_2$; 5 volumes of 2 M HEPES (pH 6.6); and 3 volumes of random sequence hexadeoxyribonucleotides (Pharmacia) at 90 OD units/ml in TE. This was stored at -20° C.

The specific activity of the DNA labelled by this method determined after ethanol precipitation and using a scintillation counter (see section 3.5.12.3.) was between 10^8 and 10^9 cpm per µg of DNA.

3.5.9.2. Single stranded DNA probes.

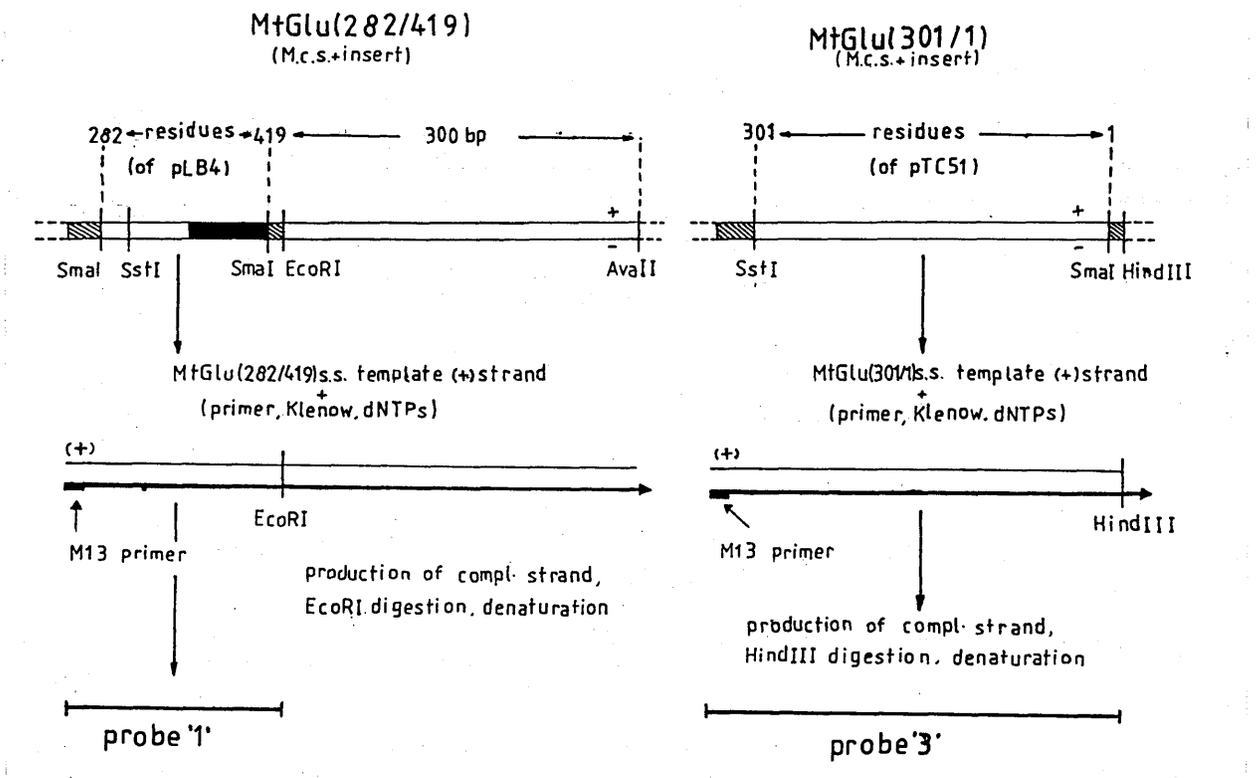
This method, summarized in figure 3.5., is a modification of the method developed by Ricca *et al.* (1982) and by Hu and Messing (1982).

The method requires single stranded M13 DNA templates prepared as in section 3.5.11.3. Approximately 1 µg of a template containing the desired DNA inserts was annealed with the appropriate primer (M13 universal primer or synthetic oligonucleotide; see section 3.5.11.4.) for 90 min at 60° C. To the annealed template dTTP, dCTP and dGTP were added to 250 µM each, as well as unlabelled dATP to 1.5 µM, 50 µCi of α-[^{32}P]-dATP (Amersham No: PB 10204, 3000 Ci/mmol), 6 units of a large fragment DNA polymerase I (Klenow, BRL) and distilled H_2O to 30 µl. The reaction was carried out for 5 min at room temperature when unlabelled dATP was added to 250 µM and the reaction extended for a further 15 min. The reaction was stopped by heating the samples at 68° C for 10 min. The solution was then adjusted to the appropriate ionic conditions, 30 units of the desired restriction enzyme to cut the DNA insert were added and the samples were incubated at 37° C for 90 min.

The samples were then mixed with an equal volume of formamide dye mix (0.03 % xylene cyanol FF, 0.03 % bromophenol blue, 1mM EDTA in deionised formamide), boiled for 5 min, quickly loaded on a 4.2% polyacrylamide sequencing gel (see section 3.5.11.4.) and electrophoresed at 30 mA for 2.5 hr.

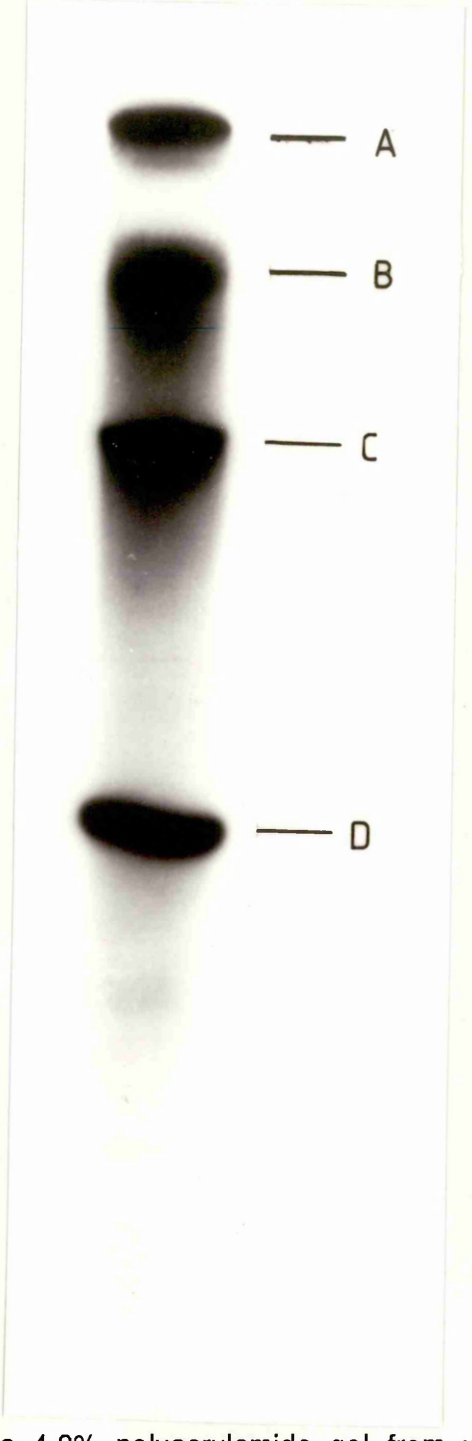
A sheet of X-ray film (Kodak) was then exposed briefly (2-5 min) on top of the gel and the resulting autoradiograph, showed a predominant sharp band derived from the single stranded DNA insert. (An example of such autoradiograph is shown in figure 3.6). This band was then excised from the gel, crushed via a 1 ml

Figure 3.5.
Preparation of probes "1", "2" and "3".



Probes "1" and "3" are single stranded DNA fragments derived from recombinants MtGlu(282-419) and MtGlu(301/1) as summarized above. Probe "2" also derives from recombinant MtGlu(282-419) but its respective template was annealed with the synthetic oligonucleotide "1"(see section 3.5.11.4.) at residues 348-364 of the insert, while the reaction was extended to an AvaII restriction site, shown in the figure. A schematic representation of the method used, the multiple cloning sites (dashed box), the insert of each recombinant and the position of the $tRNA^{Glu}$ gene (dark box) are shown.

Figure 3.6.
Purification of probe "2".



An autoradiograph of a 4.2% polyacrylamide gel from which probe "2" was purified (band D) is shown. Probe "2" was prepared after annealing a single stranded template (derived from MtGlu(282/419)) with synthetic oligonucleotide "1", after production of the complementary strand, digestion with AvaII, denatutarion and electrophoresis on the shown gel. D represents the probe (primer to AvaII site); C DNA copies from vector beyond the AvaII site; B is undigested DNA. The well (A) is also shown.

syringe and transferred to an Eppendorf tube. 1 ml of distilled H_2O was added and the sample was placed in a 68° C waterbath for 40 min, to elute the DNA from the polyacrylamide. The sample was then centrifuged for 5 min and the aqueous phase containing the "probe" was transferred to a fresh tube.

The recovery of DNA was estimated by measurment of Cerenkov cpm in an aliquot of the eluted material and in the remaining acrylamide. It was found that with this procedure 50% to 70% of the DNA generally having an activity of approximately 0.5 to $5x10^6$ cpm, was recovered.

3.5.9.3. End labelling DNA fragments.

The following method was used for preparation of radioactive DNA size markers : to 1 µg of digested, phenol/chloroform extracted and precipitated marker (commonly λ DNA digested with either HindIII or EcoRI + HindIII) were added dTTP, dCTP and dGTP, to 100 µM each, 10 µCi α-[^{32}P]-dATP (Amersham No: PB10204, 3000 Ci/mmol), 3 units of large fragment DNA polymerase I (Klenow, BRL) and distilled H_2O to 50 µl. The reaction was carried out at room temperature for 30 min. An aliquot of the above mix was then used as radioactive size marker.

3.5.10. Hybridization of blotted DNA.

3.5.10.1. Hybridization of DNA.

The following hybridization and working conditions were used for the detection of blotted nucleic acid sequences with DNA probes.

The nylon membrane with blotted DNA was placed in a polythene bag containing 10 ml of hybridization buffer (0.2% polyvinyl-pyrrolidone, 0.2% ficoll, 0.2% bovine serum albumin, in deionised H_2O stored at 4° C; 50 mM Tris-HCl, pH 7.5; 1 M NaCl; 0.1% sodium pyrophosphate; 1% SDS; 10% dextran sulphate; 50 % deionised formamide and 1 mg denatured salmon sperm DNA). The bag was sealed and the membrane pre-hybridized for 6 hr at 42° C with constant agitation.

The appropriate DNA probe was denatured in the presence of 1 mg salmon sperm DNA, by boiling the sample for 10 min and cooling it on ice for 5 min.

After pre-hybridization, the membrane was then incubated overnight at 42° C with the denatured probe in the same hybridization buffer. The hybridized membrane was washed successively in 2x100 ml of 2 x SSC for 5 min at room temperature; 2x200 ml of 2 x SSC and 1% SDS for 30 min at 65° C, followed by 2x100 ml of 0.1 x SSC for 30 min at room temperature. For a high strigency wash, i.e. genomic DNA blots, an additional wash in 2x200 mls of 0.1 x SSC and 1% SDS for 30 min at 68° C was also performed. After washing, the membrane was wrapped with cling film, and exposed to a preflashed X-ray film (Kodak) using two intensifying screens (Dupont, Cronex-lighting plus) for the desired period of time at -70° C.

The DNA fragments which were complementary to the sequence of the DNA probe were identified from the autoradiograph. After autoradiography, the membrane was washed in 200 ml of 0.4 N NaOH at 42° C for 30 min, followed by 200 ml of 0.1 x SSC; 0.1 % SDS; 0.2 M Tris-HCl, pH 7.5, at 42° C for 30 min and then stored moist at 4° C. By using the above procedure at least 95 % of the probe used was removed (i.e. after two days autoradiography no hybridization was detected); as a result the membrane was ready for another hybridization.

3.5.10.2. Densitometric analysis of autoradiographs.

The autoradiographs obtained from most experiments of hybridization of the human genomic DNA were scanned, parallel to the direction of migration of blotted DNA, by the use of a Laser densitometer (2202 Ultrascan, LKB Bromma). From each visible band of the autoradiograph at least four values were obtained and the average value expressed in the area of the band was calculated by the use of a standard deviation equation. An example of scanning and such analysis are given in figure 4.11.

The linear response between the intensity of the autoradiographs and the amount of respective DNA fragments which were hybridized was also checked. Under the same experimental conditions (autoradiography at -70° C, preflashed film using two intensifying screens) radioactive spots of a ^{32}P from 0.1 cpm up to 800 cpm were autoradiographed and analyzed as above. The densitometeric

data obtained from spots containing 1 to 100 cpm showed a linear response to the amount of radioactivity of the respective spot, from one day (i.e. 16 hr) to 14 days exposures.

3.5.11. Nucleotide sequence analysis by the Sanger chain termination method.

The nucleotide sequence analysis was performed as described by Sanger *et al.* (1980). The desired DNA fragments are subcloned into the multiple cloning sites of M13 replicative form DNA vectors (Yanish-Perron *et al.*,1985) followed by a transformation of *E. coli* cells by the recombinant DNA. Plaques containing the recombinant DNA phage (colourless in the presence of X-gal because lacking a functional β-galactosidase) were selected and single-strand (+) phage DNA prepared. An oligonucleotide, the M13 universal primer, was annealed to each of these DNA templates and extended in a 5' to 3' direction, by the use of DNA polymerase I large fragment (Klenow) and the addition of dNTP/ddNTPs mixes. Four separate reaction mixtures were prepared, each having all four dNTPs (one radioactive) but only one of the four ddNTPs.

By carefully controlling the ratio between each dNTP/ddNTP mix, incorporation of the base-specific dideoxynucleotide and hence chain length at termination is random but base-specific. The end result of such reactions is a family of DNA strands of different lengths, each terminated with a specific dideoxynucleotide. Samples were then electrophoresed on a thin polyacrylamide gel, which was then fixed, dried and autoradiographed to show DNA fragments which migrated according to their sizes.

3.5.11.1. Subcloning DNA fragments into M13 vectors.

Insert DNA and M13 vector DNA were digested with the appropriate enzyme(s) and the extent of digestion checked by gel electrophoresis. After digestion, the insert and vector DNA were purified by extraction with phenol/chloroform and precipitated with ethanol. The restricted insert DNA was

resuspended in TE to 40 ng/µl.

If the vector had been linearized by digestion with a single enzyme, 5' phosphates were usually removed by alkaline phosphatase to reduce the background of the blue plaques due to religation of the vector : the restricted vector DNA (2 µg) was redissolved in a final volume of 40 µl in a solution of 10 mM Tris-HCl (pH 9.2), 0.1 mM EDTA and 1 unit of alkaline phosphatase (calf intestinal, Boehringer grade I) was added. After incubation at 45° C for 30 min a further unit of alkaline phosphatase was added and incubation continued at 45° C for 30 min. The vector DNA was then purified by extraction with phenol/chloroform, precipitated with ethanol and redissolved in TE to 40 ng/µl.

For high efficiency cloning, it was essential to have the correct molar ratio of vector and insert. Consequently three different ligations were performed at weight ratio 3 :1, 1 : 1 and 1 : 3. Ligation reaction mixture contained : insert DNA (40-120 ng), vector DNA (40-120ng), 1 mM ATP, 10 mM DTT and 1 unit of T4 DNA ligase (BRL) in a final volume of 10 µl in ligase reaction buffer (50 mM Tris, pH 7.4; 10 mM $MgCl_2$, 2 mM spermidine; 100 µg/ml bovine serum albumin). The ligation mixture was incubated at 15° C for 16 hr.

3.5.11.2. Transformation of *E. coli* by recombinant DNA and plating out.

Competent cells were prepared by the method of Hanahan (1983) : 350 µl of an overnight culture of the bacterial host cells (*E. coli* strains JM105 or JM109) were inoculated into 35 ml SOB medium (see below). It was shaken at 37° C and allowed to grow until an A_{550} of 0.5 was reached. The cells were left standing on ice for 15 min and then were harvested by centrifugation at 2,500 rpm for 10 min at 4° C. The supernatant was removed and the cells gently resuspended in a 0.3 volume of cold TFB buffer (see below). After allowing to stand on ice for 15 min, the cell suspension was recentrifuged at 2,500 rpm for 10 min at 4° C and the cells were gently resuspended in 0.08 volumes of cold TFB buffer. Fresh DMSO (0.035 volumes) was added, the cells left on ice for 10 min, and DTT added to 75 mM. Cells were maintained on ice for an additional 10 min when a further 0.035 volumes of DMSO were added. The resultant competent cells

could be stored at 4° C for several hours, without loss of activity (10 ml of SOB medium was also inoculated with a drop of overnight culture to provide exponentially growing cells for plating).

To 300 µl of competent cells, 5 µl of DNA ligation mix were added, mixed and left on ice for 40 min. The cells were then maintained at 42° C for 3 min and returned to ice. During this time the following were added to 3 ml of molten H-top agar (kept at 42° C, see section 3.2.1.) to the final concentrations shown : 0.03% IPTG (in water), 0.03% X-gal (in dimethyl-formamide), 0.06% fresh JM105 or JM109 cells. Immediately after mixing, the molten agar was added to the transformation mix and directly spread onto a prewarmed minimal-glucose plate (see section 3.2.1.). The plate was allowed to set and incubated at 37° C overnight.

Transformed cells harbouring recombinant M13 phage appeared as opaque colourless plaques on a lawn of uninfected cells, while those containing religated vectors appeared as blue plaques. The efficiency of the method was generally found to be 10^7 plaques (i.e. transformed cells) per µg of DNA.

SOB medium was made up from the following components : 20 g/l bacto-tryptone, 5 g/l yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgS0_4$, 10 mM $MgCl_2$. The pH was adjusted to 7.0 and the medium autoclaved and stored at 4° C. The $MgCl_2$ and $MgSO_4$ were added after autoclaving.

TFB buffer contains 10 mM 2-(N-Morpholino)ethane-sulphonic acid, pH 6.2; 100 mM RbCl; 45 mM $MnCl_2$; 10 mM $CaCl_2$; 3 mM hexamminocobalt trichloride. The pH was adjusted to 6.3; the buffer filter sterilized (Nalgin filter) and stored at 4° C.

3.5.11.3. Preparation of single stranded DNA templates.

A single white plaque was inoculated into 1.5 ml 2 x YT containing 0.01 volume of an overnight culture of *E. coli* JM105 or JM109.

This culture was vigorously shaken for 4.5 hr at 37° C and the cells sedimented by centrifugation while the supernatant containing extruded phage was transferred to a fresh tube and recentrifuged to remove residual cells. The second

supernatant (1 ml) was added to 200 µl of a solution of 20% polyethylene glycol 6,000 and 2.5 M NaCl mixed and left standing at room temperature for 15 min. The precipitated phage was then sedimented by centrifugation and the supernatant was aspirated via a drawn out pasteur pipette to ensure removal of all traces of polyethylene glycol. The phage was redissolved in 100 µl TE, extracted twice with phenol/chloroform, precipitated with ethanol, briefly dried, redissolved in 25 µl TE and stored at -20° C.

3.5.11.4. Sequencing of single stranded DNA templates.

Sequencing was carried out using the protocols supplied by Amersham plc in the form M13 Cloning and Sequencing Handbook.

A. WORKING SOLUTIONS.

All nucleotide stock and working solutions were stored at -20° C.

Deoxy NTP working solutions : 10 mM stocks were prepared and diluted to 0.5 mM working solutions. 0.5 mM dATP was not required when sequencing with α-[^{35}S]-dATP.

Deoxy NTP mixes (A^o, C^o, G^o, T^o) :

	A^o	C^o	G^o	T^o
0.5 mM dCTP	20 µl	1 µl	20 µl	20 µl
0.5 mM dGTP	20 µl	20 µl	1 µl	20 µl
0.5 mM dTTP	20 µl	20 µl	20 µl	1 µl
TE, pH 8.0	20 µl	20 µl	20 µl	20 µl

Dideoxy NTP working solutions : 10 mM stocks were prepared and diluted to 0.1 mM ddATP, 0.1 mM ddCTP, 0.3 mM ddGTP and 0.5 mM ddTTP. These concentrations were altered for the sequencing reaction, if required.

Deoxy NTP/Dideoxy NTP mixes : An equal volume of dNTP mixes (i.e. A^o, C^o, G^o, T^o) was added to the corresponding ddNTP working solution.

B. ANNEALING PRIMER TO TEMPLATE.

The primer mainly used was a 17mer universal primer with the sequence 5' d[GTAAAACGACGGCCAGT] 3' complementary to a sequence adjacent to the multiple cloning sites of M13(+) strand. However the following 'tailor-made' synthetic oligonucleotide primers were also used : oligo "1" (complementary to 348-364 sequenced region of pLB4; Goddard *et al.*, 1983), oligo "9" (complementary to 130-146 sequenced region of pLB4), oligo "11" (complementary to 271-287 sequenced region of pLB4) and oligo "58" (complementary to 635-619 sequenced region of pLB4). These were made in the Department by Dr. Math, using the phosphite-triester method (Atkinson and Smith, 1984).

To anneal the primer to the template, single stranded template DNA (5 µl of the preparation approximately 1 µg) was mixed with 20 ng of the primer in a final volume of 10 µl in Klenow reaction buffer (10 mM Tris HCl, pH 8.5; 5 mM $MgCl_2$), and incubated at 60° C for 90 min.

C. SEQUENCING REACTION.

To the annealed primer/template mixture, 15 µCi α-[^{35}S]-dATP (Amersham, >600 Ci/mmol, SJ304) and 2 units of Klenow fragment (Boehringer) were added and mixed. An aliquot of the mixture (2.5 µl), was placed into each of the four tubes marked T, C, G and A in a microcentrifuge rotor. The relevant dNTP/ddNTP mix (2 µl) was placed inside the rim of each tube and a brief spin mixed the contents. After 20 min, 2 µl of chase mixture (0.5 mM of all four dNTPs) were placed into each tube, mixed and allowed to stand at room temperature for a further 15 min. The chase reaction was stopped by the addition of 4 µl of formamide dye (0.03% xylene cyanol FF, 0.03% bromophenol blue, and 20 mM EDTA in deionised formamide).

D. GEL ELECTROPHORESIS AND AUTORADIOGRAPHY.

Polyacrylamide-urea denaturing gels were prepared as follows : to 50 ml 1xTBE buffer (see section 3.2.2.) containing 5.7% acrylamide/0.3% N'N'-methylene bisacrylamide, 21 g Urea (Ultrapure ^{TM}) was dissolved by warming at 37° C. (The acrylamide was deionised and stored as a 40% stock solution in deionised H_2O). To the mixture 300 µl of a freshly made 10% ammonium persulphate solution, and 50 µl TEMED were added and it was then poured between two electrophoresis plates (40 x 20 x 0.04 cm clean glass plates, using a 0.4x0.3x0.04 cm comb). The gel was left to set for 1 hour and then transferred to the sequencing gel apparatus (61x53 Vertical Slab unit 400; Shandon, Southern), the reservoirs of which were filled with 1 x TBE buffer.

The sequencing reaction samples were boiled for 3 min, then loaded immediately onto the gel, using a drawn-out capillary. Electrophoresis was performed at 30 mA and 40 W usually until the bromophenol blue reached the bottom of the gel (approximately 2 hr).

After electrophoresis and removal of the notched plate the gel was fixed by soaking in a 2 litre bath of 10% v/v acetic acid and 10% v/v methanol for15-20 min to remove the urea. The gel was drained for a few minutes, transferred onto a sheet of Whatman 3 MM paper, covered with cling film and dried under vacuum on a gel drier (model 1125B Bio-Rad) for 30 min at 80° C.

After drying the gel was exposed directly onto Kodak X-Omat H film overnight at room temperature. A longer exposure was sometimes subsequently required.

Following autoradiography it was possible to read between 150 to 200 nucleotides from one loading. Two separate loadings were necessary to maximise the length of the sequence that could be read : the first loading was subjected to electrophoresis for 3 to 4 hr and the second loading for 2 hr. A total of 280 to 330 nucleotides could be read from two loadings. Buffer gradient gels were sometimes used to give up to 280 nucleotides with increased resolution in the lower section of the gel. These were generated as follows : 4 ml of solution A (5.7% acrylamide / 0.3% N'N'-methylenbisacrylamide, 7 M Urea, 1 x TBE, 0.1 % TEMED, 0.06% ammonium persulphate) were drawn up into a 25 ml pipette followed by 6 ml of solution B (5.7 % acrylamide / 0.3 % N'N'-methylenbisacrylamide, 7 M Urea, 5xTBE, 0.03% bromophenol blue, 0.1%

TEMED, 0.06% ammonium persulphate). One air bubble was drawn up through the solution to generate a TBE concentration gradient in the acrylamide solution within the pipette. The solution was immediately poured in the usual manner, followed by the addition of 30 ml of solution A into the electrophoresis plates set.

3.5.12. *In vitro* transcription of DNA.

3.5.12.1 HeLa S3 cell extract.

Three different methods for preparing HeLa cell nuclear extract were attempted (Dignam *et al.*, 1983, Shapiro *et al.*, 1988, Weil *et al.*, 1979). Twelve preparations (2-5 attempts from each method) were performed and the extracts derived were tested in a standard *in vitro* transcription assay (see section 3.5.12.2 and 3.5.12.3). The method of Weil *et al.* (1979) produced the most active extract in the transcription system used.

HeLa S3 cells (Gibco) were grown in suspension in 700 ml stirring flasks (Techne, Cambridge) in an Eagle's minimal essential medium (see Adams, 1980; provided by Gibco), supplemented with 10% calf serum, 1% glutamine, 1% non-essential amino acids and 1 N NaOH to give pH ≈7.5. The cells were daily monitored by microscopic examination, since it is important that cells are healthy and fast growing. Cells from a 3 litre culture (6 flasks) at a density of ≈5 x10^5 cells/ml, were harvested and washed with cold PBS mix (see section 3.2.2.). The packed cell volume was measured after centrifugation at 800 g for 5 min at 4° C. The cells were then washed with hypotonic buffer (10 mM HEPES, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT) and the cell pellet was resuspended in twice the packed cell volume of hypotonic buffer. The cells were allowed to swell for 10 min on ice and then disrupted in a Kontes all glass Dounce homogeniser by 8-10 strokes of the B pestle. The extent of cell lysis was generally 90-95 % as monitored by light microscopy, while nuclear lysis was minimal. One tenth volume of a solution containing 0.3 M HEPES (pH 7.9), 1.4 M KCl, 30 mM $MgCl_2$ was added and the lysate was centrifuged at 100,000 g for 60 min at 4° C. To the supernatant from this centrifugation, glycerol was added to

20% to yield S100 extract. The S100 cell free extract was stored in 400 µl aliquots at -70° C.

3.5.12.2. *In vitro* transcription assay.

The protocol used for the transcription assays was initially based on a consensus of the methods of Laski *et al.*, (1983), Makowski *et al.*, (1983), Murphy and Baralle (1983), Schaak *et al.*,(1984), and Segall *et al.*,(1980). However the ultimate protocol used was derived as a result of a series of optimization experiments, described in section 4.3.1.

A typical transcription reaction mixture (20 µl) contained 14 mM HEPES-KOH (pH 7.9), 80 mM KCl, 20 mM NaCl, 3.5 mM $MgCl_2$, 10% (v/v) glycerol, 0.3 mM DTT (these are final concentrations derived from components added in the extract as well as exogenously, see section 4.3.1.), 600 µM unlabelled ATP, CTP and GTP, 2.5 µCi of α-[^{32}P]-UTP (Amersham, PB: 163, 400 Ci/mmol), 25 µM unlabelled UTP, 10 mM creatine phosphate, 10 µl HeLa S100 extract, the desired amount of supercoiled recombinant DNA and M13mp12 DNA "carrier" to 150 ng of total DNA in the assay. The mixture was incubated at 30° C for 60 min and the reaction was stopped by the addition of 20 µl of a solution containing 1% SDS, 2 mg/ml proteinase K and 0.5 mg/ml *E. coli* crude tRNA. The reaction mixture was maintained at 37° C for 20 min and then an equal volume of 1 M ammonium acetate was added. The mixture was extracted once with phenol/chloroform and the aqueous layer was transferred to a fresh Eppendorf tube. To the phenol/chloroform an equal volume of 0.5 M ammonium acetate was added, mixed, centrifuged and the aqueous layer transferred to the tube containing the reaction mixture. Cold ethanol was added to 70% and the RNA precipitated at -70° C for 15 min, was recovered by centrifugation at 12,000 rpm for 10 min at 4° C. The pellet was briefly dried under vacuum, resuspended in 10 µl TE and mixed with 10 µl formamide dye mix (0.03% bromophenol blue, 20 mM EDTA in deionised formamide).

The samples were then loaded onto a 10% polyacrylamide-4M Urea 20x20x0.1 cm vertical gel (prepared essentially as described in section 3.5.11.4.; gel apparatus is a model V161 provided by BRL) and electrophoresed at

250 volts for 2.5 hr.

After electrophoresis the notched plate was removed, the gel covered with a cling film and exposed to a sheet of X-ray film (Kodak) in the presence of two intensifying screens for a defined time (usually 6-16 hr) at -70° C.

Every transcription assay was usually performed in duplicate. In order to eliminate "pipetting errors" which could lead to a misinterpretation of the results of the transcriptional experiments, all salts and ribonucleotides used were mixed and added together in every transcription assay. Finally controls where no DNA was added were also performed in each experiment.

3.5.12.3 Analysis of transcripts by the use of a scintillation counter.

Soon after the autoradiography, the autoradiographs were aligned to the gels by the help of radioactive markers and the regions of the gel corresponding to the labelled transcripts were excised and transferred to scintillation vials. Regions of the gel where no transcripts were detected were also excised to be used as background. All these "gel pieces" having the same dimensions (usually 2 x 1 cm) were "Cerenkov counted" in a Beckmann LS6800 scintillation counter. To the values obtained from transcripts, an average background value was substracted. Then an average value of the duplicates was estimated, any decay factor calculated and finally the value of each transcript expressed in counts per minute (cpm).

3.5.12.4. Calculation of tRNA products.

The tRNA products derived from every transcription assay were calculated according to the following example :

80 ng of MtGlu6 produced transcripts of 4300 cpm. The efficiency of Cerenkov counts was generally found to be 35-40%. Therefore the transcripts produced are 11466.6 dpm (A).

In every 20 µl assay, 2.5 µCi of α-[^{32}P]-UTP (400 Ci/mmol) were used, which corresponds to 0.3 µM. In addition, 25 µM of unlabelled UTP were also used; therefore the total UTP concentration is 25.3 µM or 506 pmol (B).

The 2.5 μCi of α-[^{32}P]-UTP used correspond to 5.5x10^{6} dpm (C). As a result [(A/C)xB] 1055 *f*mol UTP were incorporated into RNA (D). The tRNA^{Glu} primary transcript contains 20 uridines (E) (see section 4.3.3.3.); therefore (D/E) the transcripts produced by 80 ng MtGlu6 are 52.7 *f*mol tRNA (F).

The same amount of MtGlu6 corresponds to 15.6 *f*mol (G), (MtGlu6 molecular weight is 5.1x10^{6}). The reaction time of the assay was 60 min. Therefore (F/G) 80 ng MtGlu6 produced ≃3 transcripts/gene/hour.

CHAPTER 4

RESULTS

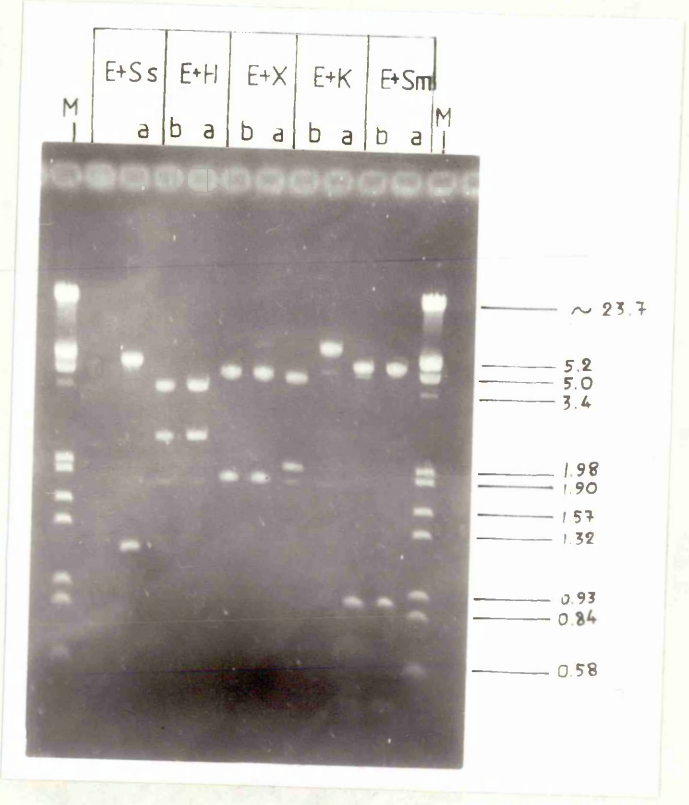
4.1. Characterization of a second copy of the gene for $tRNA^{Glu}$ and comparison with the first gene copy.

4.1.1. Restriction map of plasmid pTC51.

As has been mentioned in section 3.4.1., plasmid pTC51 contains a 2.4 kb HindIII fragment from λht190, cloned in pAT153 (figure 3.1.). The restriction map analysis of the plasmid was based on the estimation of the size of the fragments released after digestion of the DNA with a series of restriction enzymes, (EcoRI, HindIII, KpnI, PvuII, SmaI, SstI, StuI or XhoI), followed by electrophoresis through an agarose gel (see sections 3.5.6. and 3.5.7.). The construction of the map was facilitated by a direct comparison of pTC51 restriction sites with those already found in plasmid pLB4. Usually a second DNA digestion with a combination of two or three of the above mentioned restriction enzymes, was used to confirm data earlier obtained (see figure 4.1.).

The restriction map of pTC51 is shown on figure 4.2. The sites found for the restriction enzymes PvuII, SmaI, SstI, StuI and XhoI are identical to those known in plasmid pLB4. However pTC51 lacks a KpnI restriction site, which is present in pLB4.

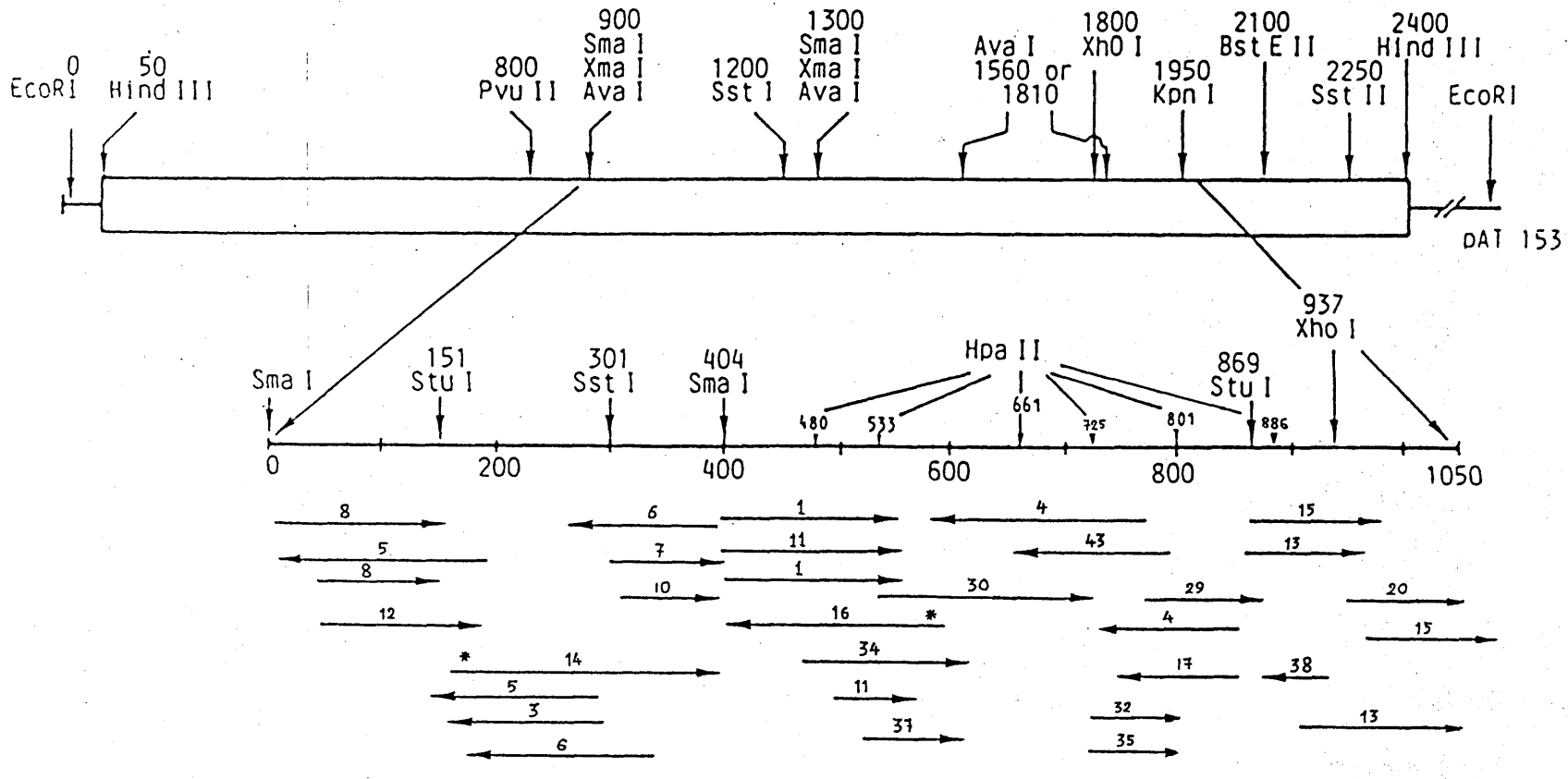
Figure 4.1.
Restriction digestions of pLB4 and pTC51.



The photograph shows an ethidium bromide stained 1.5% agarose gel where the following pTC51 (a) and pLB4 (b) digestions were electrophoresed : EcoRI+SstI (E+Ss), EcoRI+HindIII (E+H), EcoRI+XhoI (E+X), EcoRI+KpnI (E+K) and EcoRI+SmaI (E+Sm). The marker (M= λDNA/EcoRI+HindIII) fragment sizes are indicated.

Figure 4.2.

Restriction map of pLB4 and pTC51. Sequencing strategy of pTC51.



Restriction map of pLB4 containing the 2.4 kb HindIII fragment (box) from λht137 cloned into pAT153. The restriction map and the extent of sequence determination on the 2.4 kb fragment from λht190 (pTC51) is shown below. The restriction sites used to generate fragments for cloning into M13 are shown (see also table 4.1.). Sequence obtained by extension of "tailor-made" primers is shown by an asterisk.

4.1.2. Nucleotide sequence analysis of plasmid pTC51.

The sequence analysis of plasmid pTC51 was performed by the chain termination method (Sanger *et al.*, 1980). The fragments SstI-SmaI 301-404 and 301-0, SmaI-XhoI 404-937 and StuI-HindIII 869- to approximately 1500, which were obtained and isolated from a digestion of pTC51 with the respective restriction enzymes, were subcloned into M13mp19 vector (Yanisch-Perron *et al.*, 1985, figure 3.2.) as described in section 3.5.11.1. The fragments SmaI-SmaI 0-404, SmaI-StuI 404-869 and HpaII-HpaII 480-661, 533-661, 533-725, 533-801, 725-801 and 801-886 (released after a SmaI-XhoI 404-937 fragment was further digested with HpaII; see figure 4.2.) were subcloned into M13mp18 vector. The derived recombinants were used to transform *E. coli* cells, and phages were selected and amplified (see section 3.5.11.2.), for preparation of single stranded DNA templates (see section 3.5.11.3.), which were sequenced as described in section 3.5.11.4. A summary of the subclones used and the fragments sequenced, is shown in table 4.1.

By following the above experimental procedure it became possible to determine a sequence for both strands for 940 nucleotides in all but three small regions. Two of these regions were sequenced by the use of "tailor-made" synthetic oligonucleotide primers, complementary to 130-146 and 635-619 sequenced regions (see figure 4.2. and table 4.1.).

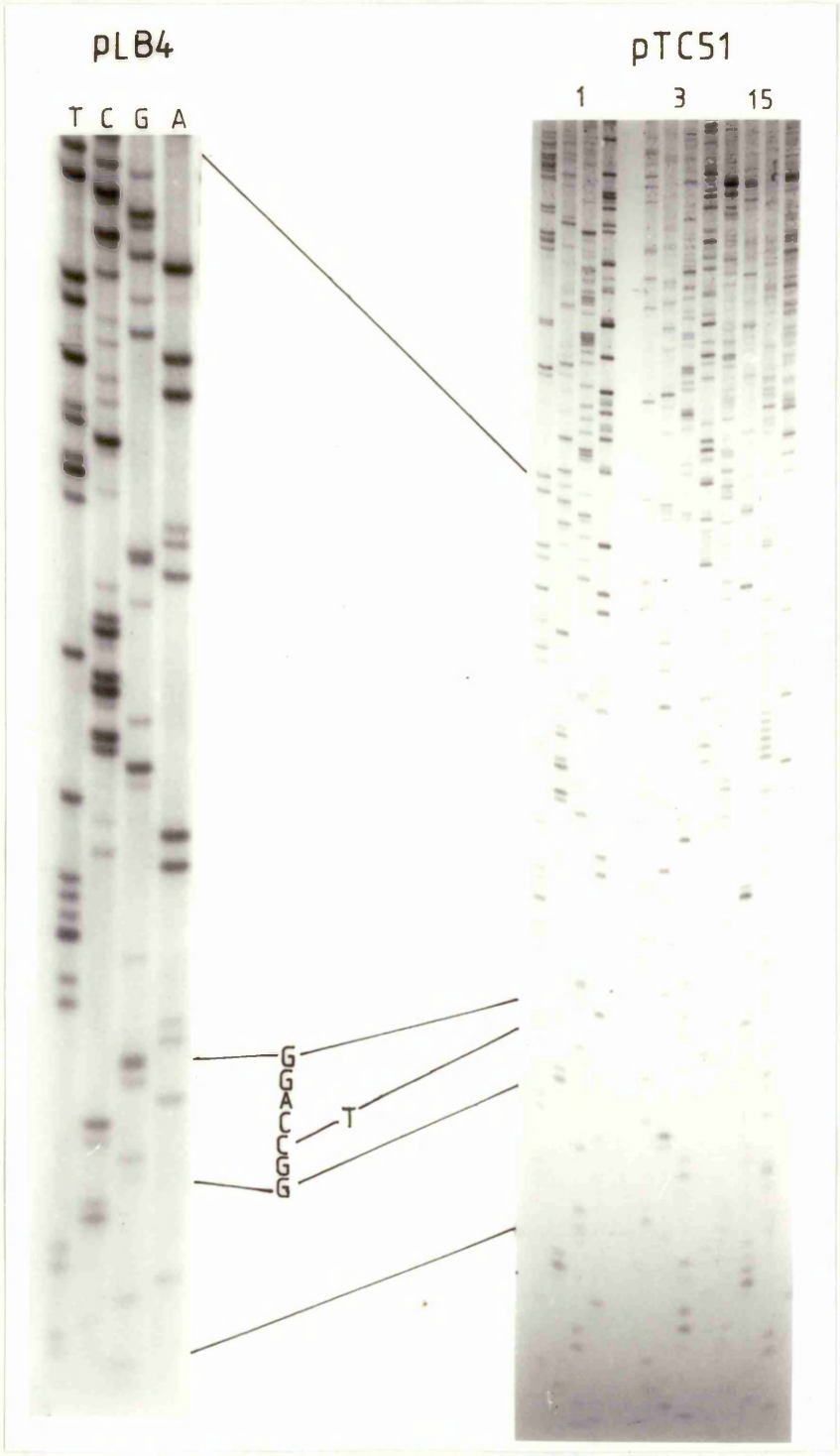
The sequence obtained from plasmid pTC51 revealed a 97.3% homology to the sequence already known from pLB4. Eleven transitions and nine transversions were found, as well as three deletions and two insertions. The sequence data also revealed the presence of a second human $tRNA^{Glu}$ gene, which differs in two nucleotides from the first characterized copy (Goddard *et al.*, 1983). In addition a sequence immediately preceding the $tRNA^{Glu}$ gene has a potential for forming a tRNA-like structure (residues -9 to -64 upstream of the gene). The features of the tRNA-like structure which is identical in all but one nucleotide to the one preceding the first gene copy (Goddard *et al.*, 1983; see also section 2.2.) will be discussed in section 5.1.2. A representative autoradiograph of the nucleotide sequence analysis of pTC51 is shown in figure 4.3. The data obtained from pTC51 aligned to those known from pLB4 and are presented in figure 4.4.

Table 4.1.
Summary of sequence data of pTC51.

Subclone	Insert		Vector	Primer	Sequenced region
8	SmaI-SmaI	1-404	M13mp18	Universal	1-165, 45-16
12	SmaI-SmaI	1-404	M13mp18	Universal	52-193
14	SmaI-SmaI	1-404	M13mp18	Oligo "9"	168-404
6	SmaI-SmaI	404 -1	M13mp18	Universal	404-252, 343-164
5	SstI-SmaI	301-1	M13mp19	Universal	297-148, 192-1
3	SstI-SmaI	301-1	M13mp19	Universal	297-154
7	SstI-SmaI	301-404	M13mp19	Universal	301-404
10	SstI-SmaI	301-404	M13mp19	Universal	316-404
1	SmaI-XhoI	404-937	M13mp19	Universal	404-550, 404-551
11	SmaI-XhoI	404-937	M13mp19	Universal	404-550, 515-581
34	HpaII-HpaII	480-661	M13mp18	Universal	481-636
37	HpaII-HpaII	533-661	M13mp18	Universal	540-636
30	HpaII-HpaII	533-725	M13mp18	Universal	535-725
32	HpaII-HpaII	725-801	M13mp18	Universal	725-801
35	HpaII-HpaII	725-801	M13mp18	Universal	725-801
29	HpaII-HpaII	801-886	M13mp18	Universal	801-886
38	HpaII-HpaII	937-886	M13mp18	Universal	937-886
43	HpaII-HpaII	801-533	M13mp18	Universal	801-650
4	StuI-SmaI	869-404	M13mp18	Universal	869-724, 785-563
17	StuI-SmaI	869-404	M13mp18	Universal	869-737
16	StuI-SmaI	869-404	M13mp18	Oligo "58"	592-404
15	StuI-HindIII	869~1500	M13mp19	Universal	869-985, 961-1125
13	StuI-HindIII	869~1500	M13mp19	Universal	869-964, 921-1051
20	StuI-HindIII	869~1500	M13mp19	Universal	954-1052

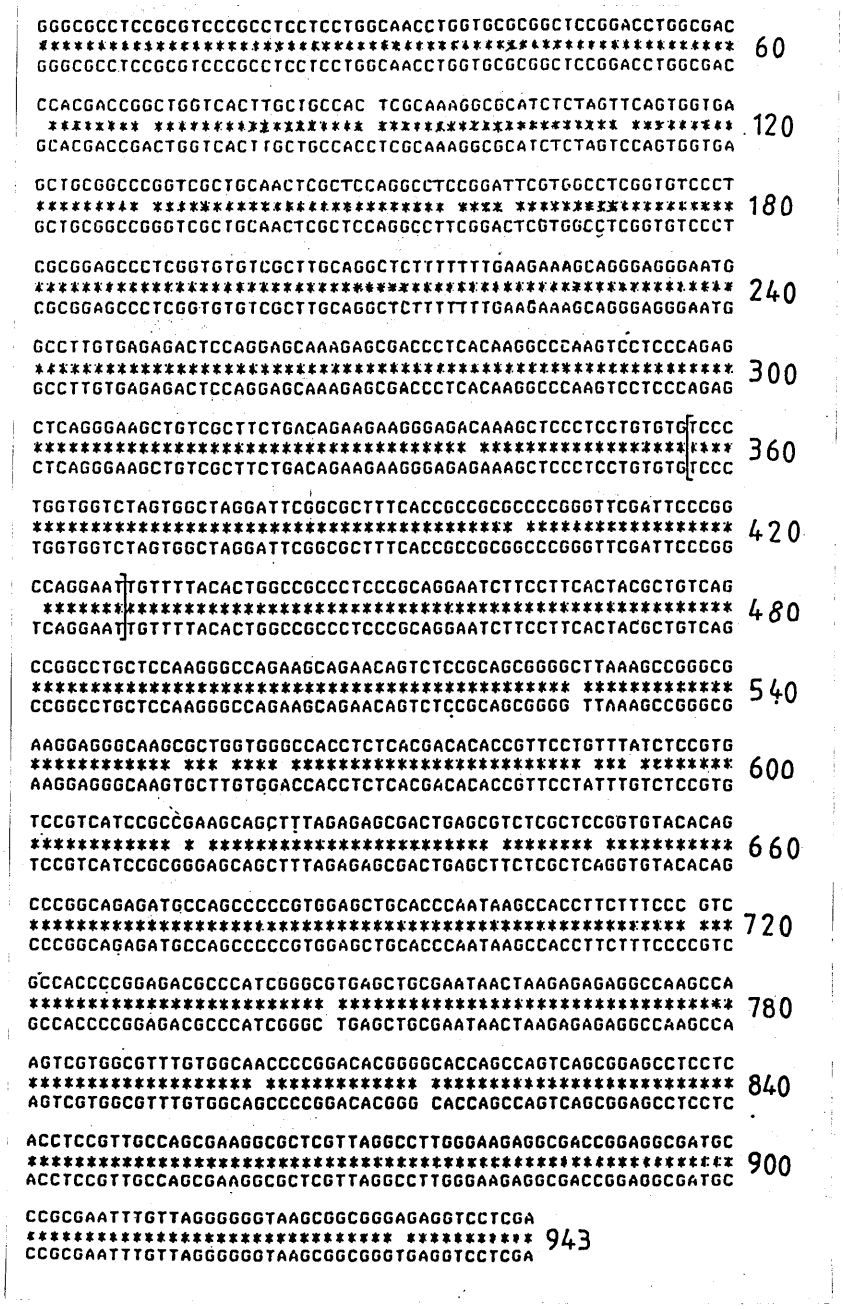
This table summarizes the subclones which were used and the pTC51 fragments sequenced by the chain termination method. The insert and the vector of each subclone, the primer used and the sequenced region are shown.

Figure 4.3.
Autoradiograph of sequencing gel of pLB4 and pTC51.



An autoradiograph of a sequencing gel of subcloned fragments from pTC51 is shown. Subclone 1 contains a SmaI-XhoI 404-937 nts fragment, subclone 3 contains a SstI-SmaI 301-1 nts fragment and subclone 15 contains a StuI-HindIII 869~1500 nts fragment (see table 4.1.). Residues 404-427 of subclone 1 correspond to the 3'-end of the $tRNA^{Glu}$ gene found in pTC51. For comparison, the sequence found in pLB4 and the nucleotide transition found in the gene (C→T) are shown.

Figure 4.4.
Nucleotide sequence of pLB4 and pTC51.



The sequence of the gene and flanking sequence from pLB4 (top line) and pTC51 (bottom line) is shown. Each line contains a sequence of 60 nts. Sequence identity is indicated (*). The $tRNA^{Glu}$ genes are boxed.

4.1.3. Restriction map of λht137 and λht190.

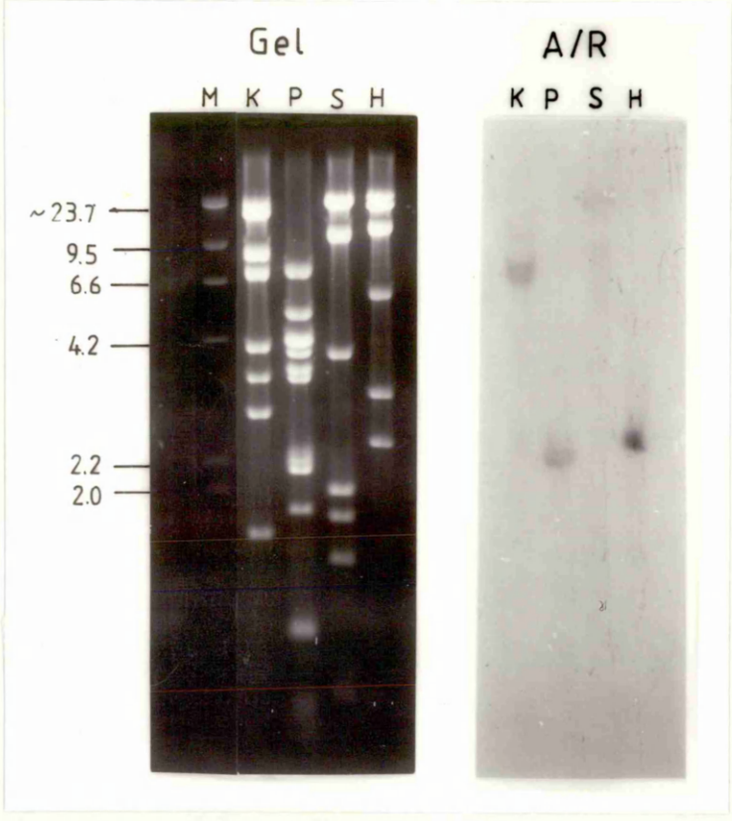
In order to understand better the organization of $tRNA^{Glu}$ genes in the human genome, it was important to construct the restriction map of both λht137 and λht190. Unfortunately throughout the period of this project, it became impossible to purify DNA from λht190 in such amounts as necessary for the completion of its restriction mapping experiments (see section 3.5.2.). Nevertheless the data obtained from λht137, together with the already known restriction fragment profiles from both λht137 and λht190, allowed the construction of an unambiguous map for λht137, as well as for λht190 in all but a small region.

4.1.3.1. Location of the $tRNA^{Glu}$ genes in λht137 and λht190.

In order to locate the position of the $tRNA^{Glu}$ gene on λht137 the following experimental procedure was used : recombinant λht137 DNA was digested with the restriction enzymes HindIII, SstI, PvuII or KpnI and fragments separated by electrophoresis through an agarose gel. The separated DNA fragments were then transferred to a nylon membrane as described in section 3.5.8. and hybridized (see section 3.5.10.1.) versus a radiolabelled single stranded DNA probe. The probe, which contains 63 out of 71 nucleotides of a $tRNA^{Glu}$ gene, as well as 74 nucleotides from its 5' prime flanking sequence, was made as described in section 3.5.9.2. (probe "1", see figure 3.5.). The washed hybridized membrane, was then exposed to a X-ray film (section 3.5.10.1.), and the resulting autoradiograph showed the following DNA fragments to have hybridized the probe which was used : 2.4 kb HindIII, approximately 22.5 kb SstI, 2.15 kb PvuII and 7.3 kb KpnI (see figure 4.5.).

At this point it has to be mentioned that similar experiments have been performed on both λht137 and λht190, by Dr. J.P. Goddard (Goddard, unpublished results). In these experiments λht137 and λht190 DNAs were digested with the restriction enzymes EcoRI, BamHI, HindIII, EcoRI+BamHI, EcoRI+HindIII and

Figure 4.5.
Hybridization of probe "1" to fragments of λht137.



The hybridization of probe"1" (see text) to KpnI (K), PvuII (P), SstI (S) or HindIII (H) fragments of λht137 is shown. The marker (M=λ DNA/HindIII) fragment sizes are indicated. (A/R= autoradiograph). See also fig. 4.12.

BamHI+HindIII, followed by a DNA electrophoresis, Southern transfer and hybridization of the blotted DNA versus a periodate-treated tRNA[^{32}P]pCp probe, as described by Goddard *et al.*,(1983). The probe, being RNA, contained sequences corresponding to $tRNA^{Glu}$, but no flanking sequences. The DNA fragments which hybridized the probe are the following: 7.8 kb EcoRI, 14.9 kb BamHI, 2.4 kb HindIII, 7.8 kb (EcoRI+BamHI), 2.4 kb (EcoRI+HindIII) and 2.4 kb (BamHI+HindIII) for λht190, and 5.2 kb EcoRI, 21.5 kb BamHI, 2.4 HindIII, 5.2kb (EcoRI+BamHI), 2.4 kb (EcoRI+HindIII) and 2.4 kb (BamHI+HindIII) for λht137 (see for example figure 4.6).

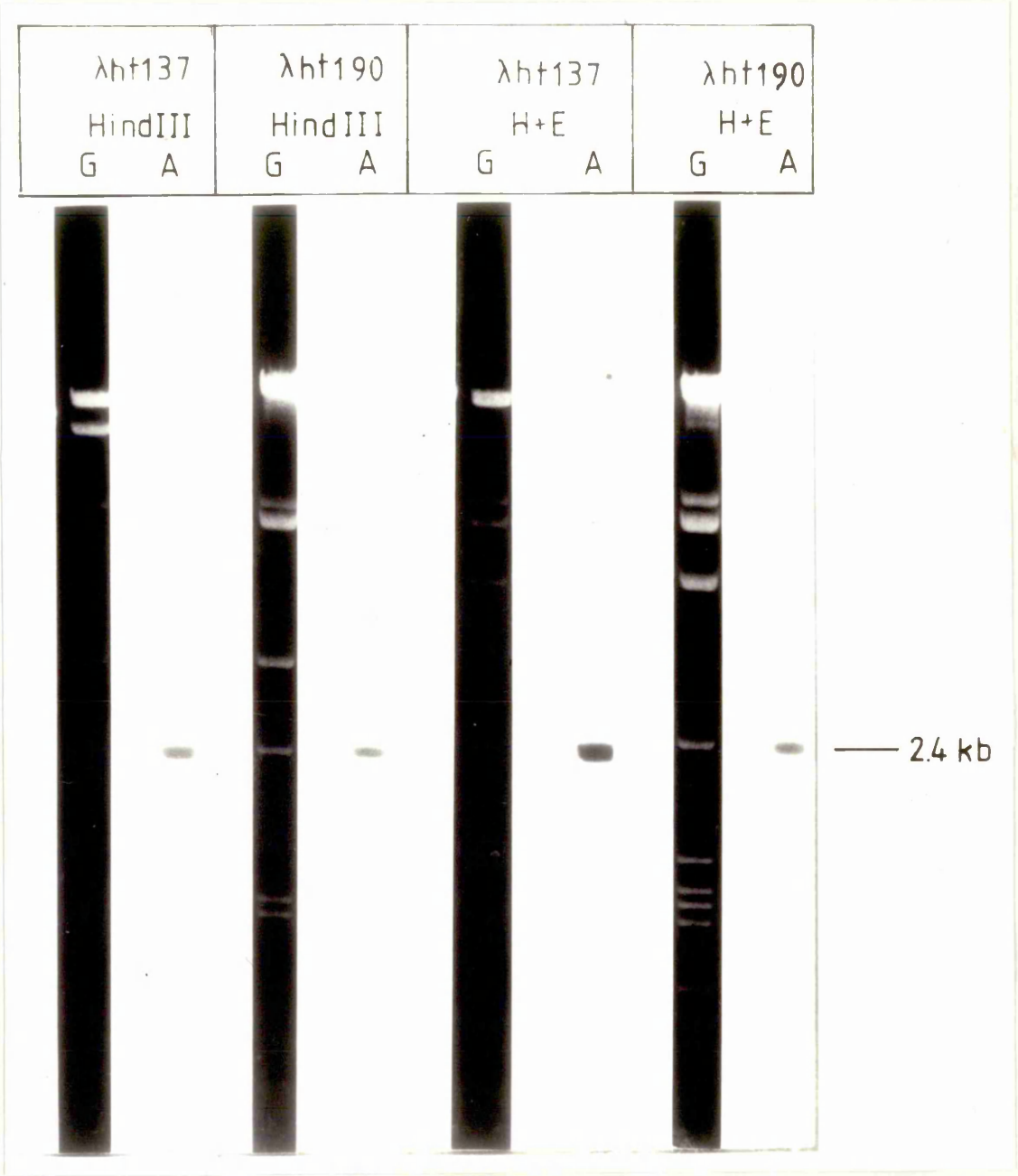
4.1.3.2. Construction of the restriction maps of λht137 and λht190.

The construction of the restriction map of λht137 was based on digestion of its DNA with the restriction enzymes EcoRI, BamHI, HindIII, EcoRI+BamHI, EcoRI+HindIII and BamHI+HindIII, followed by electrophoresis through an agarose gel and transfer to a nylon membrane as described in section 3.5.8. The blotted λht137 digests were hybridized versus different radiolabelled DNA fragments. These fragments were obtained by digestion of λht137 DNA with the desired restriction enzymes and recovery after electrophoresis through a low melting point agarose gel (see section 3.5.7.2.). The fragments were then radiolabelled with the large fragment of DNA polymerase I (Klenow), using random oligonucleotides as primers, as described by Feinberg and Vogelstein (1983 and 1984) and in section 3.5.9.1. An autoradiograph from one of these hybridization experiments, together with a photograph of the corresponding digested and electrophoresed λht137 DNA, is shown in figure 4.7. A summary of the results attained is shown in table 4.2.

The interpretation of the autoradiographs was based on logical explanation of the fragments hybridized to the probe used. An example is given below:

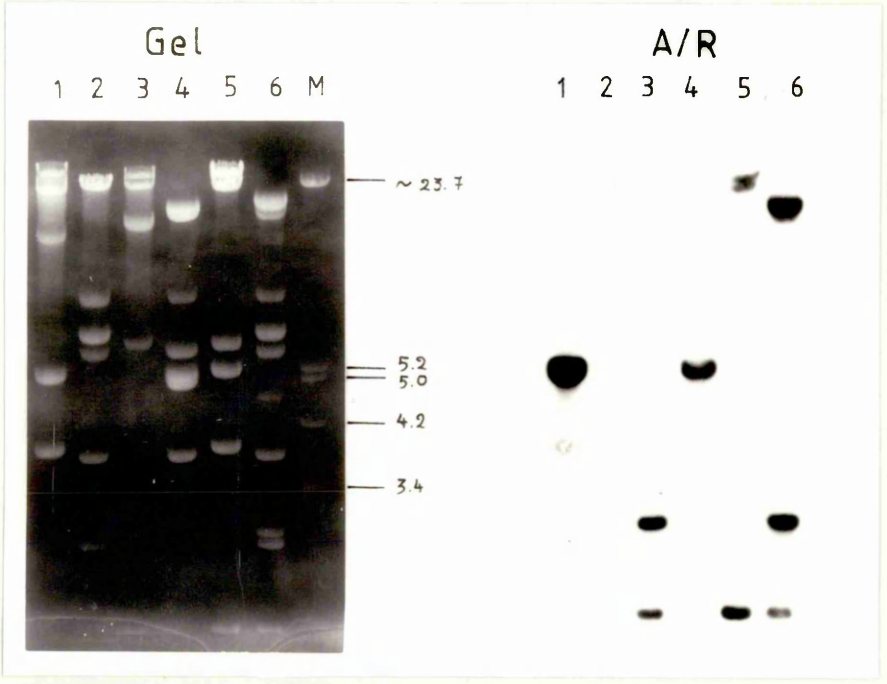
The 2.4 kb HindIII fragment hybridized the following fragments : 5.2 kb (EcoRI digest), 21.5 kb (BamHI digest), 2.4 kb (HindIII digest), 5.2 kb (EcoRI+BamHI digest), 2.4 kb (EcoRI+HindIII digest) and 2.4 kb (BamHI+HindIII digest); (see table 4.2.). As a result the 2.4 kb HindIII fragment

Figure 4.6.
Hybridization of tRNA[^{32}P]pCp probe to fragments of λht137 and λht190.



The hybridization of tRNA[^{32}P]pCp probe to HindIII or HindIII+EcoRI (H+E) fragments of λht137 or λht190 is shown. The location of the 2.4 kb HindIII fragment is shown (G=gel, A=autoradiograph). Kindly provided by Dr. J.P. Goddard.

Figure 4.7.
Hybridization of 5.2 kb EcoRI fragment to fragments of λht137.



The hybridization of 5.2 kb EcoRI fragment to EcoRI (1), BamHI (2), HindIII (3), EcoRI+BamHI (4), EcoRI+HindIII (5) or BamHI+HindIII (6) fragments of λ ht137 is shown. The weak top band of EcoRI+HindIII digest (5) hybridized presumably due to partial digestion of the DNA. Large fragments (>20 kb) transfer was low since this gel was not pretreated with acid to fragment DNA prior to transfer. The markers (M= λ DNA/EcoRI+HindIII) fragment sizes are indicated. (A/R= autoradiograph).

Table 4.2.
Hybridization data of λht137.

Probe used	*Fragments hybridized*	
Probe "1"		
	2.4 kb	(HindIII digest)
	22.5 kb	(SstI digest)
	2.15 kb	(PvuII digest)
	7.3 kb	(KpnI digest)
2.4 kb HindIII fragment		
	5.2 kb	(EcoRI digest)
	21.5 kb	(BamHI digest)
	2.4 kb	(HindIII digest)
	5.2 kb	(EcoRI+BamHI digest)
	2.4 kb	(EcoRI+HindIII digest)
	2.4 kb	(BamHI+HindIII digest)
5.2 kb EcoRI fragment		
	5.2 kb	(EcoRI digest)
	2.4 kb, 3.1 kb	(HindIII digest)
	5.2 kb	(EcoRI+BamHI digest)
	2.4 kb, 1.4 kb, 1.3 kb	(EcoRI+HindIII digest)
	15.7kb, 3.1 kb, 2.4 kb	(BamHI+HindIII digest)

Note: According to the map later constructed, (see figure 4.8.) the fragments 21.5 kb BamHI, and 21.2 kb HindIII should also hybridize the 5.2 kb EcoRI probe. This did not happen presumably due to lack of depurination of the gel used and therefore resulting insufficient blotting of large (20 kb<) DNA fragments (see figure 4.7.)

belongs to the 5.2 kb EcoRI fragment, while the latter belongs to the 21.5 kb BamHI fragment. In addition the approximately 21.5 kb BamHI fragment should contain 14.4 kb from the left arm of λ Charon 4A (see restriction map of λ Charon 4A, Maniatis *et al.*, 1982; no other λht137/BamHI fragment is bigger than 14.4kb; see table 4.3.) and consequently approximately 7.1 kb of the insert of λht137. Therefore there is a BamHI restriction site in λht137, located 7.1 kb away from the left arm EcoRI linker.

The 5.2 kb EcoRI fragment hybridized to the following fragments: 5.2 kb (EcoRI digest), 2.4 kb and 3.1 kb (HindIII digest), 5.2 kb (EcoRI+BamHI digest),2.4 kb, 1.4 kb and 1.3 kb (EcoRI+HindIII digest) and 15.7 kb, 3.1 kb and 2.4 kb (BamHI+HindIII digest) (see table 4.2.). The 15.7 kb (BamHI+HindIII) fragment should contain 14.4 kb from the left arm of λ Charon 4A (see above) and consequently 1.3 kb of the insert of λht137. From the comparison of the restriction fragment profiles of λht137/BamHI digest and λht137/BamHI+HindIII digest it is clear that the 15.7 kb is a BamHI-HindIII fragment. Therefore there is a HindIII site located 1.3 kb away from the left arm EcoRI linker. However the 5.2 kb EcoRI fragment belongs to a 7.1 kb, left arm neighbouring, EcoRI-BamHI fragment and contains the 2.4 kb HindIII fragment (see above). These data in conjunction with the EcoRI+HindIII fragments which hybridized the same probe, reconfirm the location of the previously identified HindIII site and also demonstrate that the next HindIII site is located 2.4 kb away from the first one.

The data obtained from such analysis, led to the construction of the map of λht137 for the region around the $tRNA^{Glu}$ gene. However the close relationship between the restriction fragment profiles which λht137 and λht190 generate, facilitated the construction of the rest of the map of λht137, on the assumption that identical size DNA fragments are located on similar positions on both λht137 and λht190 (see table 4.3.). Additional analysis of the restriction fragment profiles of λht137, KpnI or SstI digests, and further analysis of hybridization data described earlier (see section 4.1.3.1.), led also to the mapping of all KpnI and SstI restriction sites.

Using these methods and analyzing the data from λht190 (see section

Table 4.3.
Restriction fragment profiles of λht137 and λht190.

EcoRI		BamHI		HindIII		EcoRI+BamHI		EcoRI+HindIII		BamHI+HIndIII	
λht137	λht190	λht137	λht190	λht137	λht190	λht137	λht190	λht137	λht190	λht137	λht190
20 *	20	**21.5**			25.7	14.4 *	14.4	20 *	20	15.7	
11 *	11		**14.9**	21.2			**7.8**	5.7 *	5.7		14.8
	7.8		14.8	12.1		5.5 *	5.5	5.3 *	5.3	5.9	
5.2		5.9		5.7 *	5.7	**5.2**		4.0 --	4.0	5.5 *	5.5
4.0 --	4.0	5.5 *	5.5		5.5	5.1 *	5.1	**2.4** --	**2.4**		5.3
2.05		3.9 *	3.9	3.1 --	3.1	3.9 *	3.9	2.05		3.9 *	3.9
1.8 --	1.8	3.05 --	3.05	**2.4**--	**2.4**	1.9 --	1.9	1.75--	1.75	3.1 --	3.1
	1.6	1.50 *	1.5		1.7	1.8 --	1.8		1.70	3.05 --	3.05
0.7		0.95 --	0.95		1.65	1.5 *	1.5		1.65	**2.4** --	**2.4**
		0.90 --	0.90		0.45	1.15 --	1.15		1.60		1.7
		0.80				0.95 --	0.95	1.40 --	1.40		1.65
		0.50 *	0.50			0.90 --	0.90	1.30		1.50 *	1.50
						0.80		0.70		0.95 --	0.95
						0.70			0.45	0.90 --	0.90
						0.50 *	0.50		0.20	0.80	
							0.40				0.45
						0.20--	0.20			0.30 *	0.30
										0.20 *	0.20
										0.20 --	0.20

λht137 total length: 44.5 kb

λht190 total length: 46.0 kb

All sizes are expressed in kb. Fragments indicated by an asterisk belong exclusively to the arms of λCharon 4A. A dashed line indicates homology of the fragments from the genomic DNA insert of λht137 and λht190. The fragment sizes in bold contain the $tRNA^{Glu}$ genes. DNA fragments smaller than 0.2 kb are not mentioned.

4.1.3.1.), it became also possible to construct an unambiguous restriction map for λht190, in all but a small region. The restriction maps of both λht137 and λht190 are shown in figure 4.8.

4.2. Organization of the human tRNA^{Glu} gene family in the human genome.

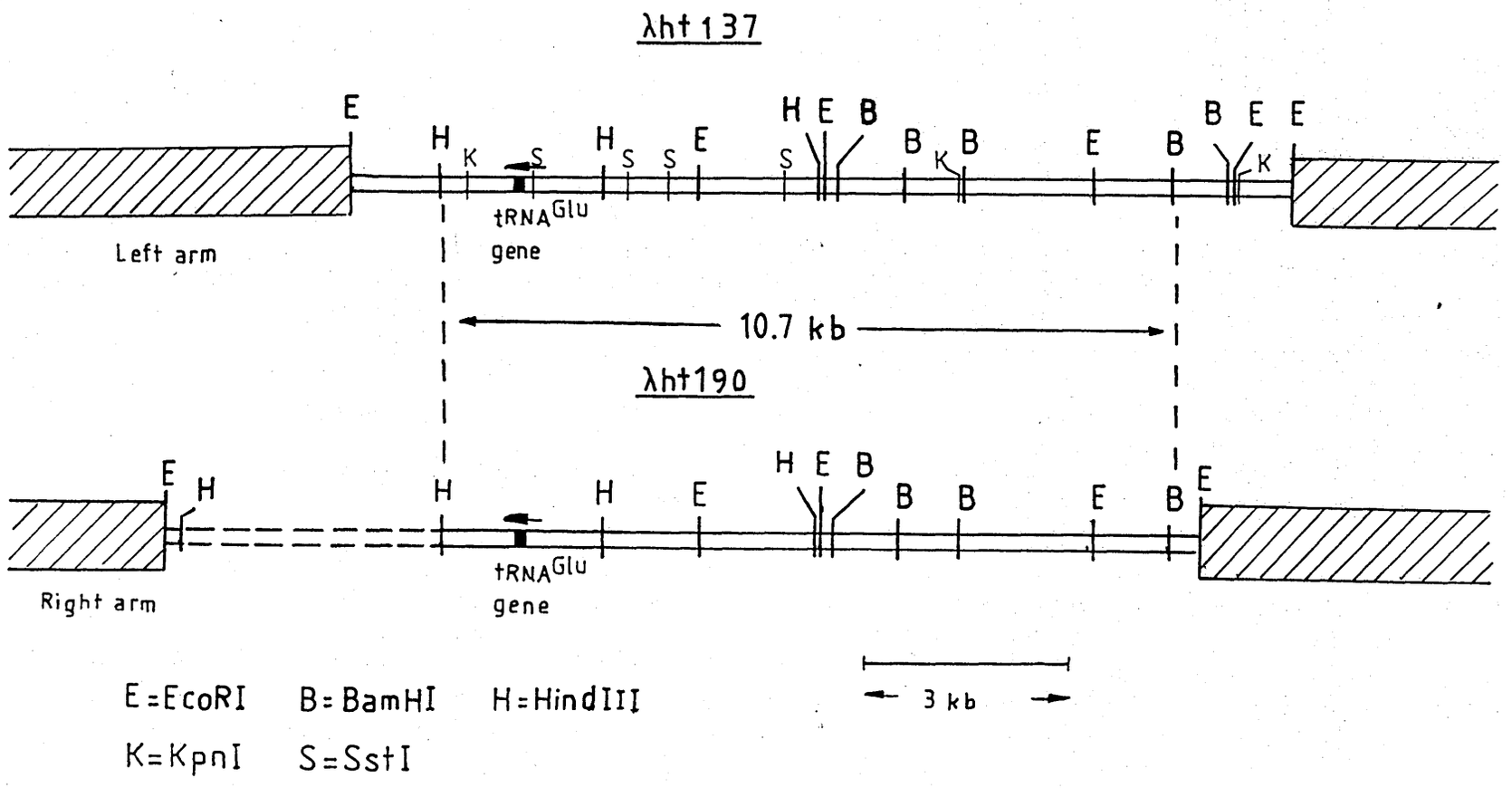
One preliminary experiment (Goddard, unpublished results) suggested that many of the tRNA^{Glu} genes belong to long (several kb) DNA repeats in the human genome. The data obtained from recombinants λht137 and λht190 showed that two tRNA^{Glu} gene copies, belong to at least 10.7 kb highly homologous DNA fragment repeats (see figure 4.8.). As a consequence, the estimation of the tRNA^{Glu} gene copy number and the comparison of the copies and their flanking sequences became one of the first objectives of this project.

4.2.1. Estimation of the tRNA^{Glu} gene copy number.

In order to estimate the copy number for the human tRNA^{Glu} gene family, the following experimental procedure was followed:

Human placental DNA (prepared as described in section 3.5.4.) was digested with the restriction enzyme HindIII. For comparative purposes three different amounts (1.0; 1.3 and 2.0 µgs) of genomic DNA/HindIII digest were electrophoresed through an agarose gel in parallel with pTC51 digests with HindIII (to release a 2.4 kb fragment containing the tRNA^{Glu} gene). The amounts of pTC51/HindIII digests (10 to 50 pgs, equivalent to 0.25 to 1.25 x 10^{-18} mols of gene) were chosen to generate approximately the same number of tRNA^{Glu} genes as 1.0 - 2.0 µg of genomic DNA. The pTC51/HindIII digests were electrophoresed in the presence of undigested λ DNA "carrier". The separated DNA fragments were blotted into a nylon membrane as described in section 3.5.8. and hybridized (as in section 3.5.10.1.) versus a single stranded "tRNA^{Glu} gene probe". The probe ("probe 1", see figure 3.5.) which contains 63 out of 71 nucleotides of a tRNA^{Glu} gene as well as 74 nucleotides of its 5'-flanking sequence, was prepared as described in

Figure 4.8.
Restriction maps of λht137 and λht 190.



The restriction maps of λht137 and λht190 are shown. The extent of homology in λht137 and λht190 and the positions of the $tRNA^{Glu}$ gene copies are shown. The broken region of λht190 has not been mapped; however it contains two HindIII restriction sites, releasing the following fragments : 1.7 kb, 1.65 kb, 0.45 kb.

section 3.5.9.2. and in figure 3.5. The hybridized membrane was then exposed for a defined time to a preflashed X-ray film, at -70° C. The resulting autoradiograph showed a predominant 2.4 kb band as well as three other weak bands from the genomic DNA HindIII digestion, and a series of intensities of 2.4 kb HindIII bands from the pTC51 HindIII digestions (see figure 4.9.). The autoradiograph was then scanned as described in section 3.5.10.2. to determine the extent of hybridization to each fragment and the data obtained were processed as shown in figure 4.10. An example of a densitometeritic analysis is given in figure 4.11.

The results of such analysis show that 9 copies of the gene for $tRNA^{Glu}$ belong to 2.4 kb HindIII fragment repeats per human haploid genome ($3x10^9$ bp), while one copy belongs to a 5.5 kb HindIII fragment, two copies in 8.2 kb HindIII fragments and finally one more copy in a very large (30 kb approximately) HindIII fragment. It was concluded that the copy number for the human $tRNA^{Glu}$ gene family is thirteen.

The membrane was cleaned by removal of the probe previously used (see section 3.5.10.1.) and rehybridized versus another single stranded DNA probe. This probe ("probe 2", see figure 3.5.) contains 63 out of 71 nucleotides of a $tRNA^{Glu}$ gene, but only 6 nucleotides of its 5'-prime flanking sequence, prepared as described in section 3.5.9.2. The membrane was then processed as described above. The results obtained confirmed the data from the previous experiment, that the copy number is thirteen.

4.2.2. Organization of the $tRNA^{Glu}$ gene copies in the human genome.

The experiments on the estimation of the human $tRNA^{Glu}$ gene copy number revealed that the vast majority of the copies (nine out of thirteen), belong to 2.4 kb HindIII fragment repeats in the human genome. In order to further investigate the length of these DNA repeats, as well as to understand more about the organization of the copies, a series of human genomic DNA digestions were performed and hybridized versus a variety of probes.

In the first experiment human placental DNA was digested with nine different restriction enzymes. These restriction enzymes were chosen, either because they release different average size DNA fragments or, because their recognition sites

In gel A the HindIII fragments derived after the simultaneous digestion of 0.1 μg human genomic and 0.5 μg λ DNA with 0.5 units HindIII (b) have been electrophoresed in parallel to the HindIII fragments derived after digestion of 1 μg λ DNA with the same enzyme (a; see section 3.5.6).
In gel B three different amounts of human genomic DNA digested with HindIII (1=1μg, 2=1.3 μg, 3=2 μg) have been electrophoresed in parallel with 8 different amounts of pTC51 digested with HindIII (4=10 pg, 5=15 pg, 6=20 pg, 7=25 pg, 8=30 pg, 9=35 pg, 10=40 pg and 11=50 pg). The pTC51/HindIII digests were electrophoresed in the presence of λ DNA "carrier". The hybridization of probe "1" to the blotted gel is shown. The position of the three weak bands in lanes 1-3 of the autoradiograph (A/R) are marked (•) (see text and also figure 4.12.). The marker (M=λDNA/HindIII) fragment sizes are indicated .

Figure 4.9.

Hybridization of probe "1" to HindIII fragments of human genomic DNA.

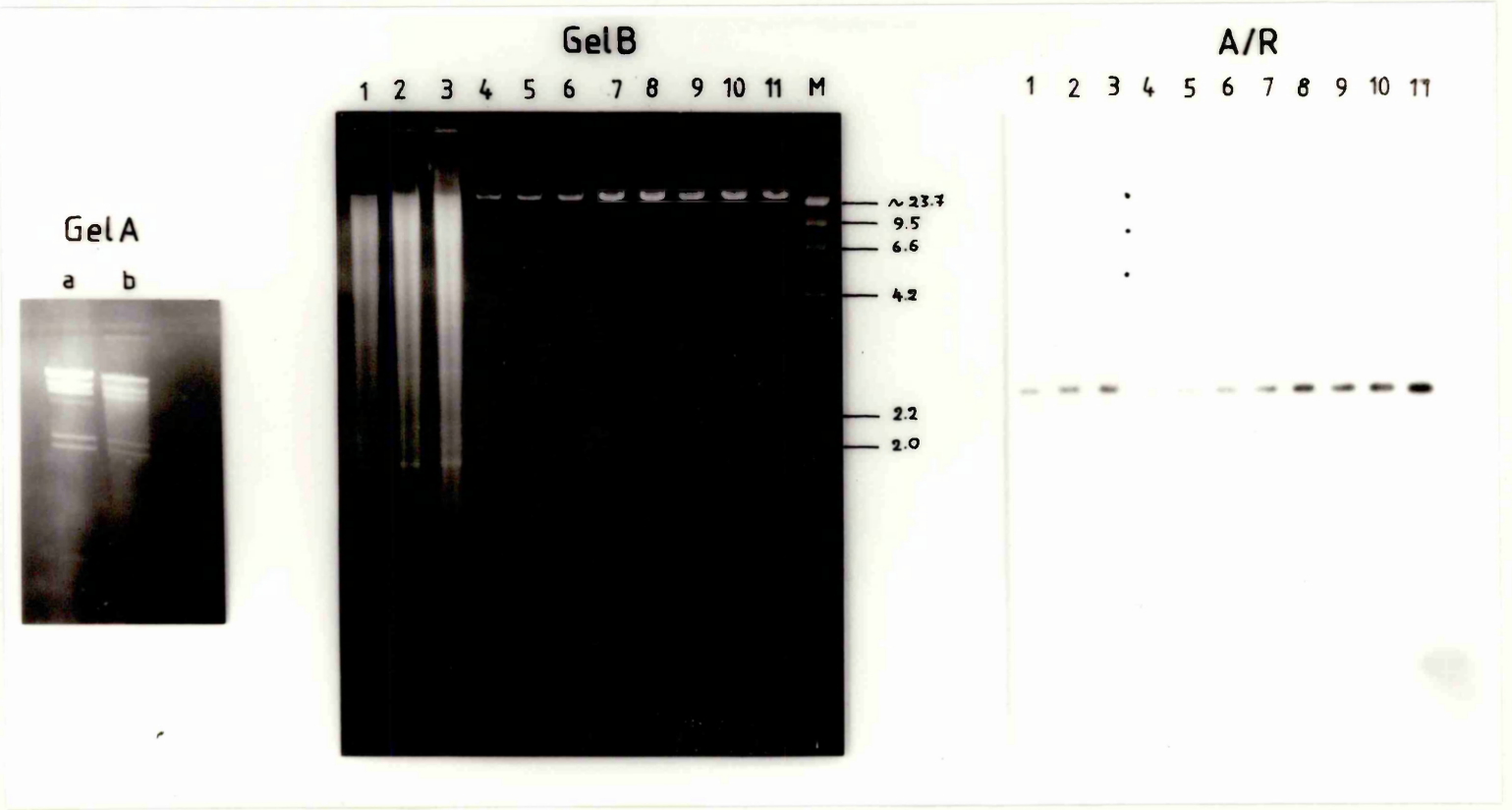
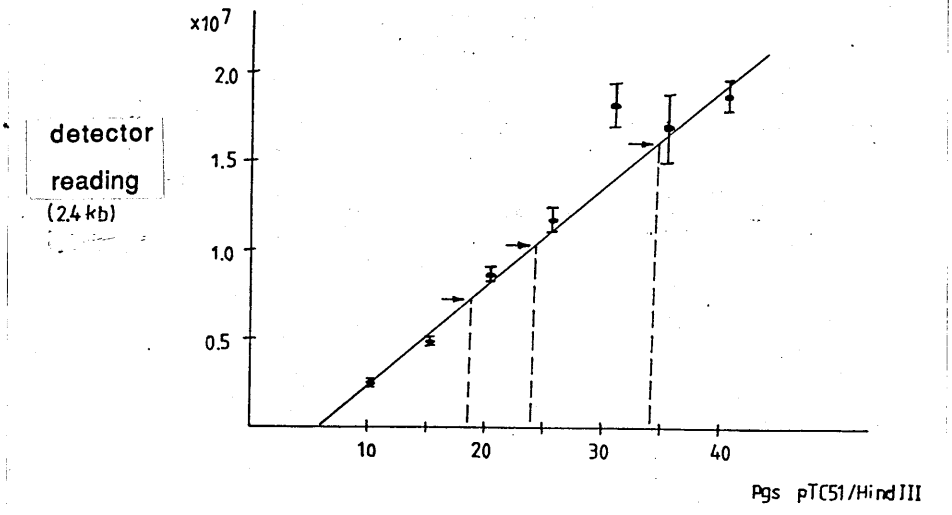


Figure 4.10.
Estimation of human $tRNA^{Glu}$ gene copy number.



EQUATION FOR THE ESTIMATION OF THE COPY NUMBER :

$$\frac{\text{amount of genomic DNA used} \times \text{copy number}}{\text{M.W. human DNA}} = \frac{\text{amount of pTC51/HindIII used}}{\text{M.W. pTC51}}$$

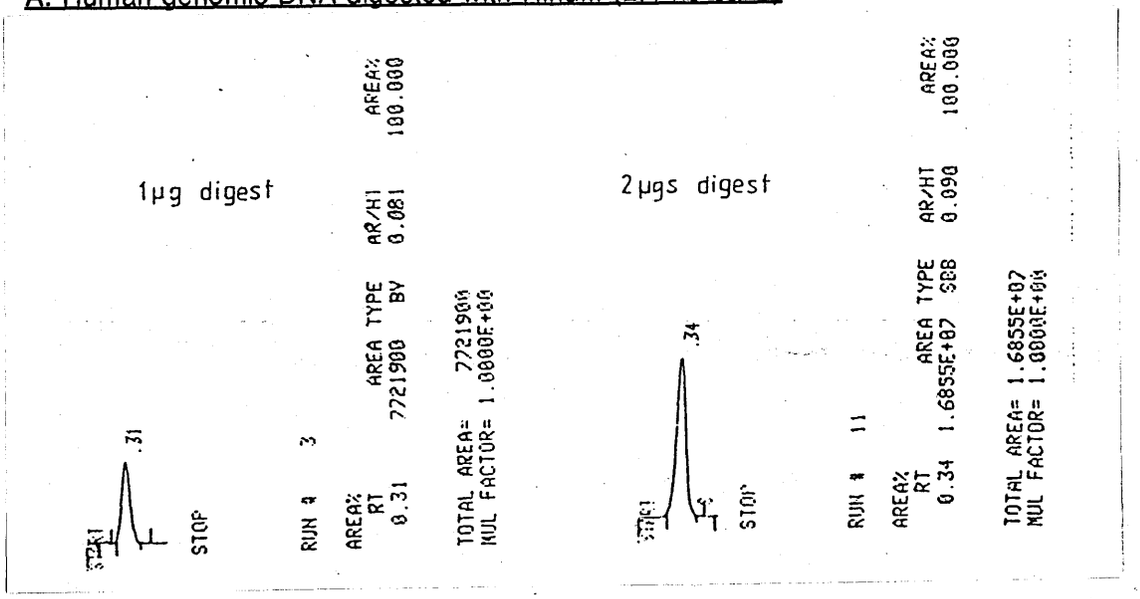
Amount of genomic DNA used	Detector reading* (for 2.4 kb band)	Corresponding amount of pTC51/HindIII digest	Copy number (from equation)
1.0 =	0.76 x10^7 ±0.20 =	18.5 pgs ±0.5	9.25±0.25
1.3 µg =	1.07 x10^7 ±0.05 =	23.9 pgs ± 0.8	9.19±0.31 → 8.96±0.4
2.0 µg =	1.65 x10^7 ±0.02 =	33.8 pgs ± 0.4	8.45±0.10

The standard curve derived from the pTC51/HindIII digestions (see figure 4.9) and the equation for the estimation of the $tRNA^{Glu}$ gene copy number are shown. Error bars indicate deviation of densitometeric readings from each band. They were estimated according to the standard deviation equation. Arrows represent the densitometeric readings obtained from the 2.4 kb bands from 1, 1.3 and 2 µg of genomic DNA digested with HindIII. An example of the densitometeric analysis is given in figure 4.11. The number of copies belonging to the 2.4 kb HindIII band was concluded to be nine. For the three weak bands at 5.5 kb, 8.2 kb and approximately 30 kb (see fig.4.9.&4.12) detector readings were at 0.5/2/1 ratio. Although the 'band' at 5.5 kb appeared rather as an artifact than a real hybridization signal, additional experiments reconfirmed its presence as a single copy. In conclusion the copy number for the gene of $tRNA^{Glu}$ resulted thirteen.

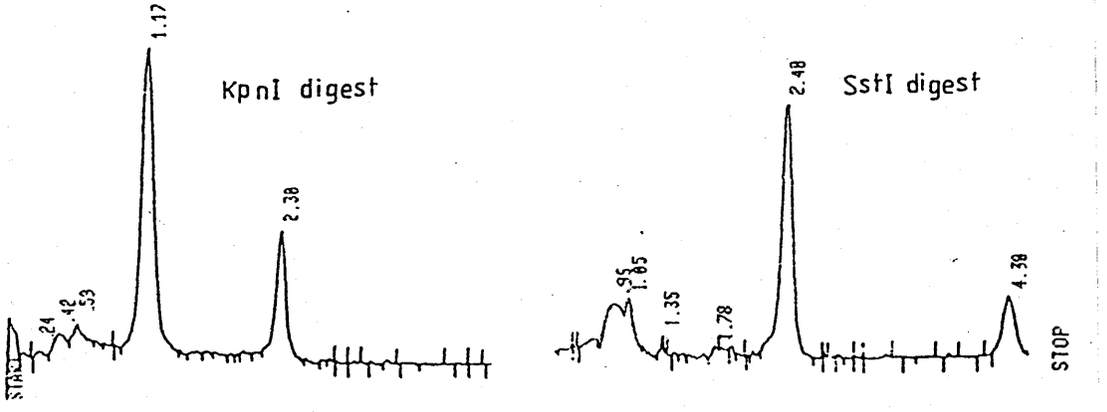
*Standard deviations were derived from four readings.

Figure 4.11.
An example of densitometeric analysis of genomic DNA hybridization autoradiographs.

A. Human genomic DNA digested with HindIII (2.4 kb band)



B. Human genomic DNA digested with Kpn I or Sst I



Part A shows the densitometric readings obtained for the 2.4 kb band in an autoradiograph of a Southern blot of 1 and 2 µg of human genomic DNA digested with HindIII and hybridized to probe "1". The values on the right side of the patterns, expressing the area of the peak, were analyzed as shown in figure 4.10. Part B shows the densitometric readings obtained from an autoradiograph of a Southern blot of human genomic DNA digested with KpnI or SstI and hybridized to probe "1". The respective readings obtained (data not shown) were in the ratio 1/1/6/3 and 2/1/1/5/2 which corresponded to the number of gene copies per haploid genome complement.
Note : The left side peak of the SstI digest (position 0.95-1.05) appeared as a doublet due to background in the autoradiograph (see figure 4.12.). A and B are not in scale.

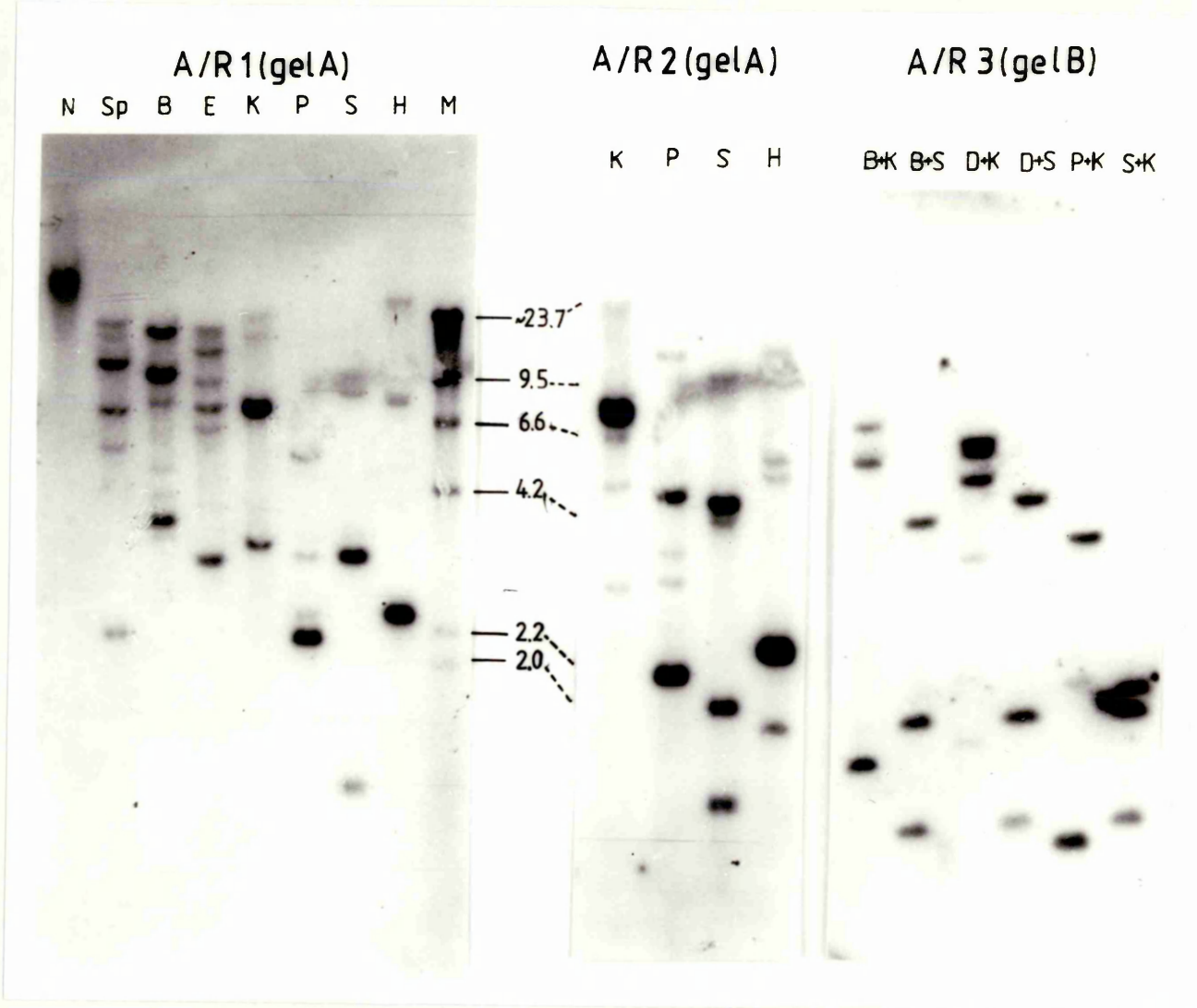
were known to occur in or around the characterized $tRNA^{Glu}$ gene copies within recombinants λht137 and λht190. Thus the restriction enzymes NotI, which cut DNA very rarely , DraI or SphI, which cut DNA every 4 kb on average, as well as EcoRI, BamHI, SstI, PvuII, KpnI or HindIII, all of which occur in λht137 and λht190, were selected. After the appropriate digestions, the DNA fragments were electrophoresed through an agarose gel, Southern blotted (see section 3.5.8.) and hybridized (see section 3.5.10.1.) versus a single stranded "$tRNA^{Glu}$ gene probe" ("probe 1", see figure 3.5.), which was prepared as described in section 3.5.9.2. The hybridized membrane was then exposed to a preflashed X-ray film for a defined time as described in section 3.5.10.1. and the resulting autoradiograph (see figure 4.12.) was scanned as described in section 3.5.10.2.

The densitometric analysis of the autoradiograph (an example of such analysis is given in figure 4.11.) showed again 9 copies to belong to the 2.4 kb HindIII fragment repeats in the human genome, while 5, 5 and 6 copies appear to belong to 2.15 kb PvuII, 3.15 kb SstI and 7.3 kb KpnI fragment repeats respectively. For the rest of DNA digestions, although a few copies appear to belong to identical sized DNA fragments, greater restriction fragments length polymorphism was observed. However in all DNA digestions, single copies were also detected. The total copy number of the $tRNA^{Glu}$ gene family in every case was found to be between 11 and 14. A summary of the results attained is shown in table 4.4.

The membrane was cleaned by removal of the probe previously used (see section 3.5.10.1.) and rehybridized versus another single stranded "$tRNA^{Glu}$ gene probe" ("probe 2", see figure 3.5.) which was prepared as described in section 3.5.9.2. The membrane was then washed as usual and the resulting autoradiograph was scanned as described in section 3.5.10.2. The data obtained from this experiment show only one difference to those found by using "probe 1" (see table 4.4. and accompanying comments).

From these primary experiments it became clear, that most of the $tRNA^{Glu}$ gene copies belong to similar repeats of several kb in the human genome. To test the extent of homology in the 5'-flanking sequence of the copies, a single stranded probe which only contains 301 nucleotides of the 5'-prime flanking sequence of a $tRNA^{Glu}$ gene (probe "3", residues -55 to -355; see figure 3.5) was employed to

Figure 4.12.
Hybridization of probes "1" and "3" to fragments of human genomic DNA.



The hybridization of probe "1" to NotI (N), SphI (Sp), BamHI (B), EcoRI (E), KpnI (K), PvuII (P), SstI (S) and HindIII (H) fragments of human genomic DNA electrophoresed in gel A, is shown. (A/R 1; here shown reduced compared to A/R 2). The hybridization of probe "3" to the same fragments (digestions with KpnI, PvuII, SstI and HindIII) is also shown. (A/R 2). The radiolabelled marker (M= λ DNA/HindIII) fragment sizes are indicated in between autoradiographs 1 and 2. The hybridization of probe "3" to BamHI+KpnI (B+K), BamHI+SstI (B+S), DraI+KpnI (D+K), DraI+SstI (D+S), PvuII+KpnI (P+K) and SstI=KpnI (S+K) fragments of human genomic DNA electrophoresed in gel B is shown on the right hand side of the figure (A/R 3).

In each section of the table, the restriction enzyme and the probes (P=probe) used are indicated. All DNA fragment sizes are expressed in kb. The number of copies to the corresponding DNA fragments are shown in bold type. An asterisk indicates copies which hybridize all three probes. The data from probes "1&2" represent identical results obtained in series of experiments using either probes. However in two cases (4.35 kb SphI and 4.55 kb SstI fragments size) the number of corresponding copies found was inconsistent. In the SstI digestion the results from probes "1" and "2" do not overlap with those from probe "3", presumably due to the presence in most of the copies of a SstI site, present in pLB4 at 297-302 (see figures 4.2. and 5.2.). The probes used contain sequences from different directions of the referred SstI site (see figure 3.5.). The estimation of the copy number which corresponds to every DNA fragment was based on densitometeritic analysis of the autoradiogtraphs.

Table 4.4.
Human genomic DNA hybridization data.
A. Single digestions.

DraI				SphI				BamHI				EcoRI			
P"1&2"			P"3"	P"1&2"			P"3"	P"1&2"			P"3"	P"1&2"			P"3"
7.3	4	*	3 7.3	19.0	1	*	1 19.0	15.4	3	*	3 15.4	16.2	2	*	2 16.2
6.5	4	*	4 6.5	14.8	1	*	1 14.				1 12.0	11.8	2	*	2 11.8
5.3	1	*	1 5.3	10.8	3	*	3 10.8	9.5	4	*	4 9.5	9.2	1	*	2 9.2
4.95	1						1 8.3	7.7	1	*	1 7.7	7.3	2	*	1 7.3
			1 4.40	7.0	2	*	1 7.0				4 5.6	6.5	1		
			1 3.90	5.55	1	*	1 5.55	4.75	1						1 5.9
3.75	1	*	1 3.75	4.35(1)	3	*	3 4.35	4.15	1			4.55	1		
3.50	1			2.30	2			3.65	3			4.20	1	*	1 4.20
			1 2.15								1 2.15	3.05	3	*	3 3.05
1.50	1														2 1.10

Copies detected: 13	Copies detected: 13	Copies detected: 13	Copies detected: 13
Copies hybridized both probes: 9	Copies hybridized both probes: 10	Copies hybridized both probes: 8	Copies hybridized both probes: 10
Additional fragments hybridized probe "3" : 3	Additional fragments hybridized probe "3" : 1	Additional fragments hybridized probe "3" : 6	Additional fragments hybridized probe "3" : 4

KpnI				PvuII				SstI				HindIII			
P"1&2"			P"3"	P"1&2"			P"3"	P"1&2"			P"3"	P"1&2"			P"3"
22.5	1	*	1 22.5				1 15.2				1 10.0	~30	1		
14.8	1	*	1 14.8				1 11.5	8.6	2						1 12.0
7.3	6	*	6 7.3	10.3	1						1 7.4	8.2	2		
			1 6.2	5.2	1			4.55(1)	4	*	4 4.55	5.5	1	*	1 5.5
			1 4.85	4.75	3	*	3 4.75				1 4.10				1 4.9
			1 3.55	3.45	1	*	1 3.45	3.45	1			2.40	9	*	9 2.40
3.30	3	*	1 3.30	3.15	1	*	1 3.15	3.15	5						2 1.80
			1 3.10	2.35	1						4 2.00				
				2.15	5	*	5 2.15				3 1.40				
								1.20	2						

Copies detected: 11	Copies detected: 13	Copies detected: 14	Copies detected: 13
Copies hybridized both probes: 9	Copies hybridized both probes : 10	Copies hybridized both probes : 4	Copies hybridized both probes: 10
Additional fragments hybridized probe "3" : 4	Additional fragments hybridized probe "3" : 2	Additional fragments hybridized probe "3" : 10	Additional fragments hybridized probe "3" : 4

hybridize the membrane used earlier. The experimental procedure followed was the same as the one described above. The data of this experiment (table 4.4.) show, that indeed 9 or 10 copies do share homologous 5' flanking sequences, at least as far as 350 nucleotides upstream the gene, although the remaining 3 or 4 copies might not. However from every DNA digestion, an additional 3 or 4 fragments were hybridized, (which did not hybridize the probes containing a $tRNA^{Glu}$ gene), when "probe 3" was used (see figure 4.12.).

Finally two of the probes described above (probes "1" and "3") were used individually to hybridize another membrane which contained the electrophoresed and Southern blotted human genomic DNA fragments obtained by digestion with : BamHI+KpnI, BamHI+SstI, DraI+KpnI, DraI+SstI, PvuII+KpnI or SstI+KpnI. The experimental procedure in both cases was the same as described above. The data acheived from these experiments (see table 4.5.), showed that although most of the copies belong to highly similar 2-3 kb DNA fragment repeats, three classes of restriction fragments length polymorphism (6-10 kb) were observed, where 3 to 4 copies belong to each of them (see figure 4.12).

4.3. Transcription studies on human $tRNA^{Glu}$ genes.

One of the major objectives of this project was to study the role of the 5'-flanking sequence on a human $tRNA^{Glu}$ gene transcription. Consequently several recombinants lacking different parts of the 5'-flanking sequence of the above mentioned gene had been constructed and partially characterized (see section 3.4.3.), for use in transcriptional studies. The use of a HeLa S100 extract provided an advantage over *Xenopus* oocyte nuclear micro-injection (see section 2.2.) of studying the above recombinants in a homologous transcriptional system. However before such analysis took place, it was important to establish a reproducible and efficient transcription system. Therefore the same HeLa S100 extract had to be used in all experiments, under invariant optimized conditions.

Table 4.5.
Human genomic DNA hybridization data.
B. Double digestions.

BamHI+KpnI Probe "1"- Probe "3"					DraI+KpnI Probe "1"- Probe "3"					PvuII+KpnI Probe "1"- Probe "3"				
8.8	**1**	*	**1**	8.8	6.5	**4**	*	**4**	6.5	20	**1**	*	**1**	20
7.3	**3**	*	**3**	7.3	6.0	**3**	*	**3**	6.0				**1**	14.5
5.5	**3**	*	**3**	5.5	5.15	**3**	*	**3**	5.15	3.7	**4**	*	**4**	3.7
1.95	**1**	*	**1**	1.95	3.75	**1**	*	**1**	3.75	3.4	**1**			
1.75	**4**	*	**4**	1.75	3.30	**2**				2.15	**2**	*	**2**	2.15
								1	3.10				**1**	1.80
								1	2.40	1.20	**6**	*	**6**	1.20
								1	1.85					
Copies detected:				**12**	Copies detected:				**13**	Copies detected:				**14**
Copies hybridized both probes:				**12**	Copies hybridized both probes:				**11**	Copies hybridized both probes:				**13**
Additional fragments hybridized Probe "3":				-	Additional fragments hybridized Probe "3":				**3**	Additional fragments hybridized Probe "3":				**2**

BamHI+SstI Probe "3"		DraI+SstI Probe "3"		SstI+KpnI Probe "3"	
4.10	**4**	4.70	**4**	4.80	**1**
2.60	**1**	3.90	**1**	4.30	**1**
2.20	**1**	2.00	**4**	3.70	**1**
2.00	**4**	1.40	**3**	2.10	**4**
1.40	**3**			2.00	**4**
				1.40	**3**
Fragments hybridized:	**13**	Fragments hybridized:	**12**	Fragments hybridized:	**14**

In each section of the table, the restriction enzymes and the probes used are indicated. All DNA fragment sizes are expressed in kb. The number of copies to the corresponding DNA fragment are shown in bold type. An asterisk indicates copies which hybridize both probes. The hybridization data obtained from the BamHI+SstI, DraI+SstI and SstI+KpnI digestions using probe "1" have not been analyzed, due to cross-hybridization of the probe at both directions of a conserved to most copies SstI site (residues 297-302 of pLB4; see figures 3.5., 4.2., 5.2. and table 4.4.).

4.3.1. Optimization of an *in vitro* transcription system.

From preliminary experiments it was observed that the transcription reaction has a linear response (transcription products versus time of the reaction) at 30° C, between 30 and 120 minutes. For a 60 minute reaction time it was also observed that 100-200 ng of added DNA saturate the reaction, producing the maximum amount of transcription products for the amount of the extract used. (A more detailed analysis of the time course and DNA concentration experiments is given in sections 4.3.3.2. and 4.3.3.1. respectively). Having established these conditions (i.e. 60 minute reaction at 30° C), twelve different preparations of HeLa S100 extracts were tested, using 150 ng MtGlu6 as template (see section 3.4.3 and figure 3.3.). These extracts were prepared by the methods of Weil *et al.*, (1979), Dignam *et al.*, (1983) and Shapiro *et al.*, (1988) (see section 3.5.12.1.). The transcription assays were performed as described in section 3.5.12.2. and samples were then electrophoresed, autoradiographed and finally the transcription products derived were analyzed as described in section 3.5.12.3. The selection of the best extract preparation, was based upon the efficiency of the extract to produce the maximum amount of MtGlu6 transcripts, and give rise to no endogenous transcription. The reproducibility of the extract chosen in efficient transcription was tested in additional experiments. This extract was used in all subsequent experiments, unless otherwise stated.

The extract chosen had been prepared as described in section 3.5.12.1. All 20µl reactions contained 10 µl extract which itself contained 3 mM $MgCl_2$ and 120 mM KCl, so that in the absence of additional Mg^{2+} and K^+ final concentrations were 1.5 mM $MgCl_2$ and 60 mM KCl (see section 3.5.12.1. and 3.5.12.2.). However a further optimization of the salt concentration was attempted. In the first experiment final concentrations of 60-180 mM KCl were tested. Having established the optimum KCl final concentration (80 mM) an optimization of $MgCl_2$ concentration was attempted. In these assays final concentrations of 1.5-13.5 mM $MgCl_2$ were tested in the presence of 80 mM KCl. The optimum final concentration for $MgCl_2$ was found to be 3.5 mM. Finally using 80 mM KCl

and 3.5 mM $MgCl_2$ as final concentrations, an optimization of NaCl concentration was also attempted. In these assays concentrations of 0-140 mM NaCl were tested and the optimum concentration was found at 20 mM. All assays were performed as decribed in sections 3.5.12.2. and 3.5.12.3., having 150 ng of MtGlu6 as template. Figure 4.13 illustrates results of salt optimization experiments.

Finally the effect of the total amount of exogenous DNA upon the amount of transcripts of the DNA itself, was also tested. In these experiments different concentrations of MtGlu6 were used, in the presence or absence of M13mp12 DNA "carrier". The results (see section 4.3.3.1.) indicated that at low MtGlu6 concentrations in the assay, the presence of M13mp12 "carrier" considerably stimulated the transcription of MtGlu6, although it was not itself transcribed . Even 150 ng of M13mp12 did not produce any detectable transcript. Consequently DNA "carrier" was added to give final total DNA of 150 ng/assay.

These experiments established the optimum salt concentrations as well as the optimum DNA concentration for reaction times of 60 min. The method described in section 3.5.12.2. is the outcome of such optimization and was followed in all subsequent experiments, unless otherwise indicated.

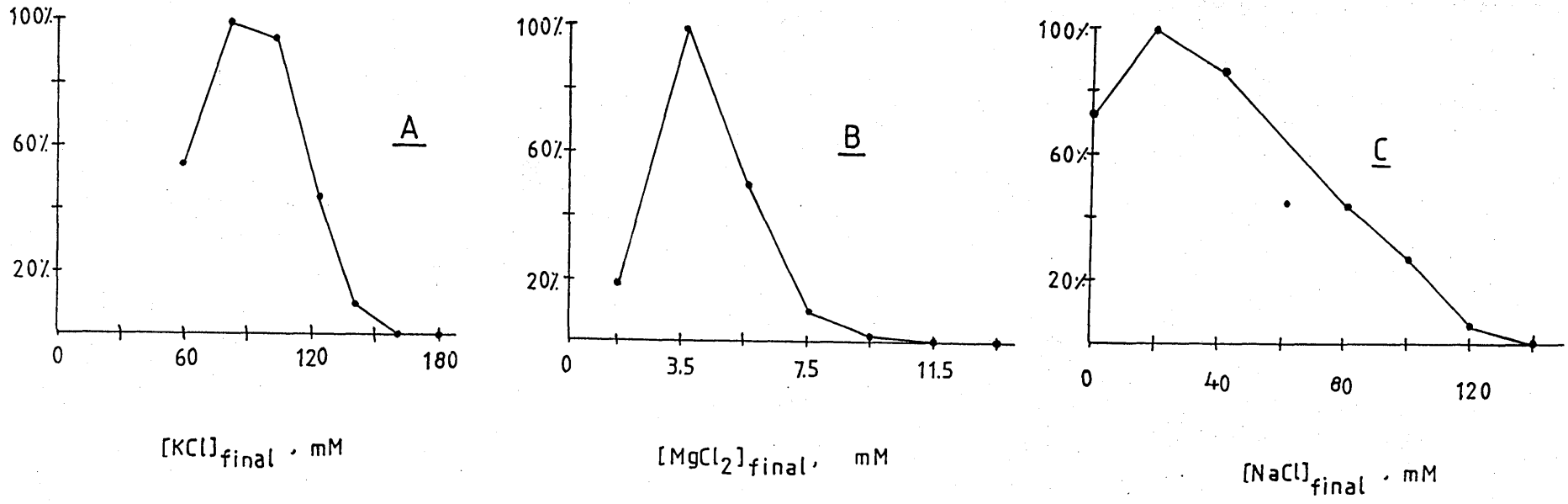
4.3.2. *In vitro* transcriptional efficiency of pLB4 and pTC51.

As has been reported earlier, plasmids pLB4 and pTC51 contain two different human $tRNA^{Glu}$ genes (see section 4.1.2. and figure 4.4.). 150 ng of each plasmid were used as templates in identical *in vitro* transcription assays, as described in section 3.5.12.2. The results of these assays, shown in figure 4.14., reveal that both plasmids were transcribed with very similar efficiency.

4.3.3. *In vitro* transcriptional analysis of MtGlu6.

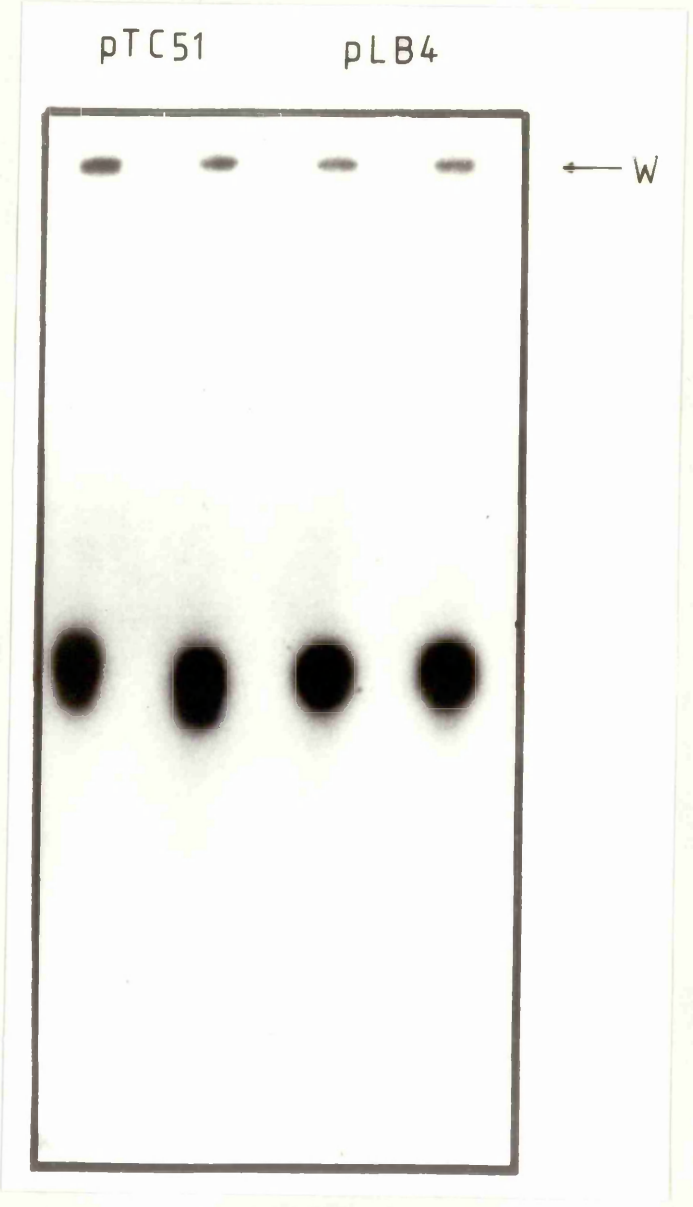
MtGlu6 contains a single $tRNA^{Glu}$ gene, derived from recombinant λht137 (see sections 2.1., 3.4.3. and figure 3.3.). This recombinant was used for studying the transcriptional activity of its $tRNA^{Glu}$ gene, in the *in vitro* transcription system, which has been described earlier.

Figure 4.13.
Salt optimization of the *in vitro* transcription system.



-A. The transcription products derived from 150 ng of MtGlu6 in the presence of 1.5 mM $MgCl_2$ and 60 to 180 mM final concentrations KCl. 100% = transcripts obtained in the presence of 80 mM KCl.
-B. The transcription products derived from 150 ng of MtGlu6 in the presence of 80 mM KCl and 1.5 to 13.5 mM final concentrations $MgCl_2$. 100% = transcripts obtained in the presence of 3.5 mM $MgCl_2$.
-C. The transcription products derived from 150 ng of MtGlu6 in the presence of 3.5 mM $MgCl_2$, 80 mM KCl and 0 to 140 mM NaCl. 100% = transcripts obtained in the presence of 20 mM NaCl.

Figure 4.14.
Autoradiograph of *in vitro* transcription of pTC51 and pLB4.



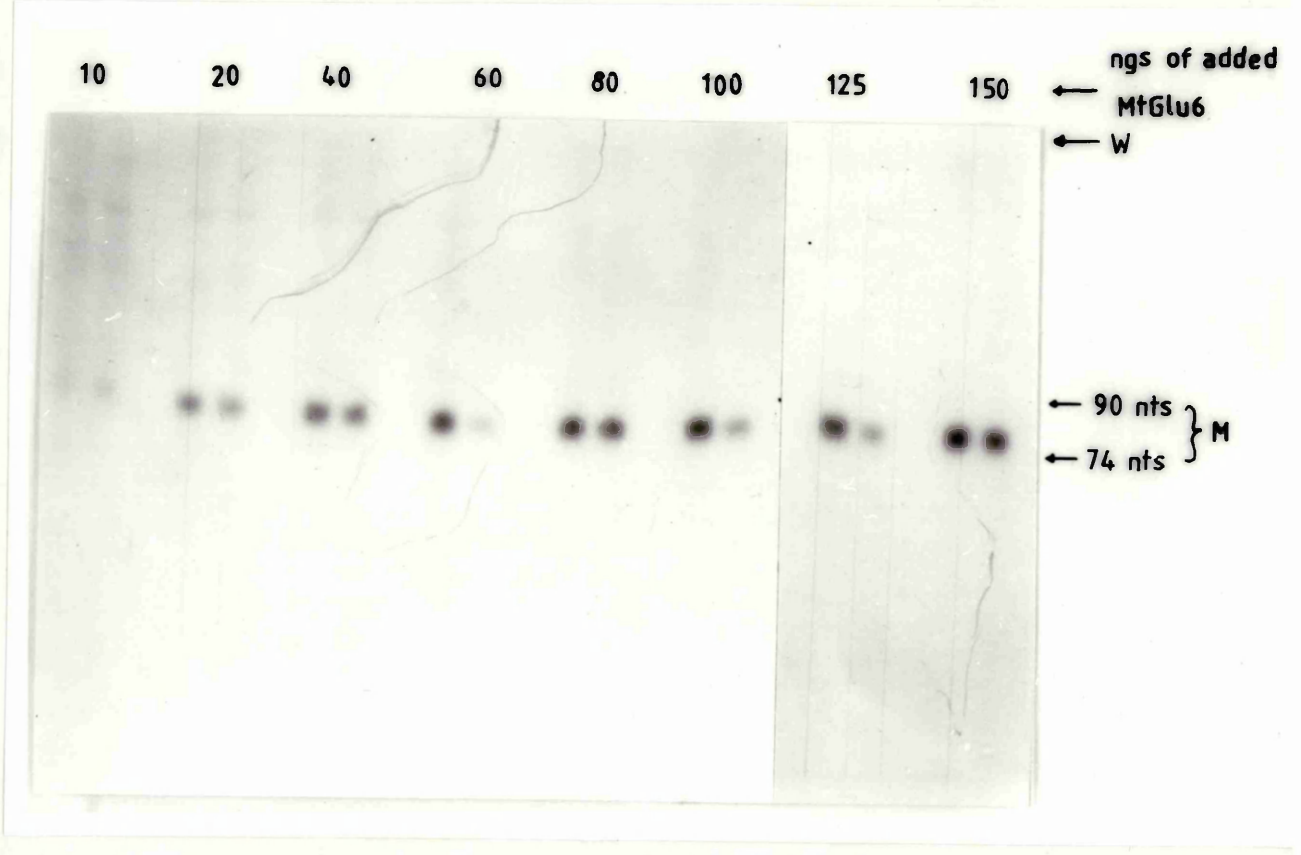
An autoradiograph of products of transcription of 150 ng pTC51 or 150 ng pLB4, using HeLa S100 extract is shown. All assays are shown in duplicates. The position of the wells (W) is also shown.

4.3.3.1. MtGlu6 concentration assays.

The effect of increasing amounts of added $tRNA^{Glu}$ gene upon the extent of transcription was investigated. Using the procedure described earlier (see section 3.5.12.2.) and optimal metal ion concentrations (see section 4.3.1.) samples containing 0, 5, 10, 20, 40, 60, 80, 100, 125, 150, 175 or 200 ng of MtGlu6 plus an amount of M13mp12 DNA "carrier" to give a final concentration of 150 ng/20 µl wherever required. MtGlu6 samples without M13mp12 DNA "carrier" were also processed. After 60 min at 30° C the reaction was terminated and samples were then electrophoresed, autoradiographed (see for example figure 4.15.) and finally the transcription products derived were analyzed as described in sections 3.5.12.3. and 3.5.12.4.

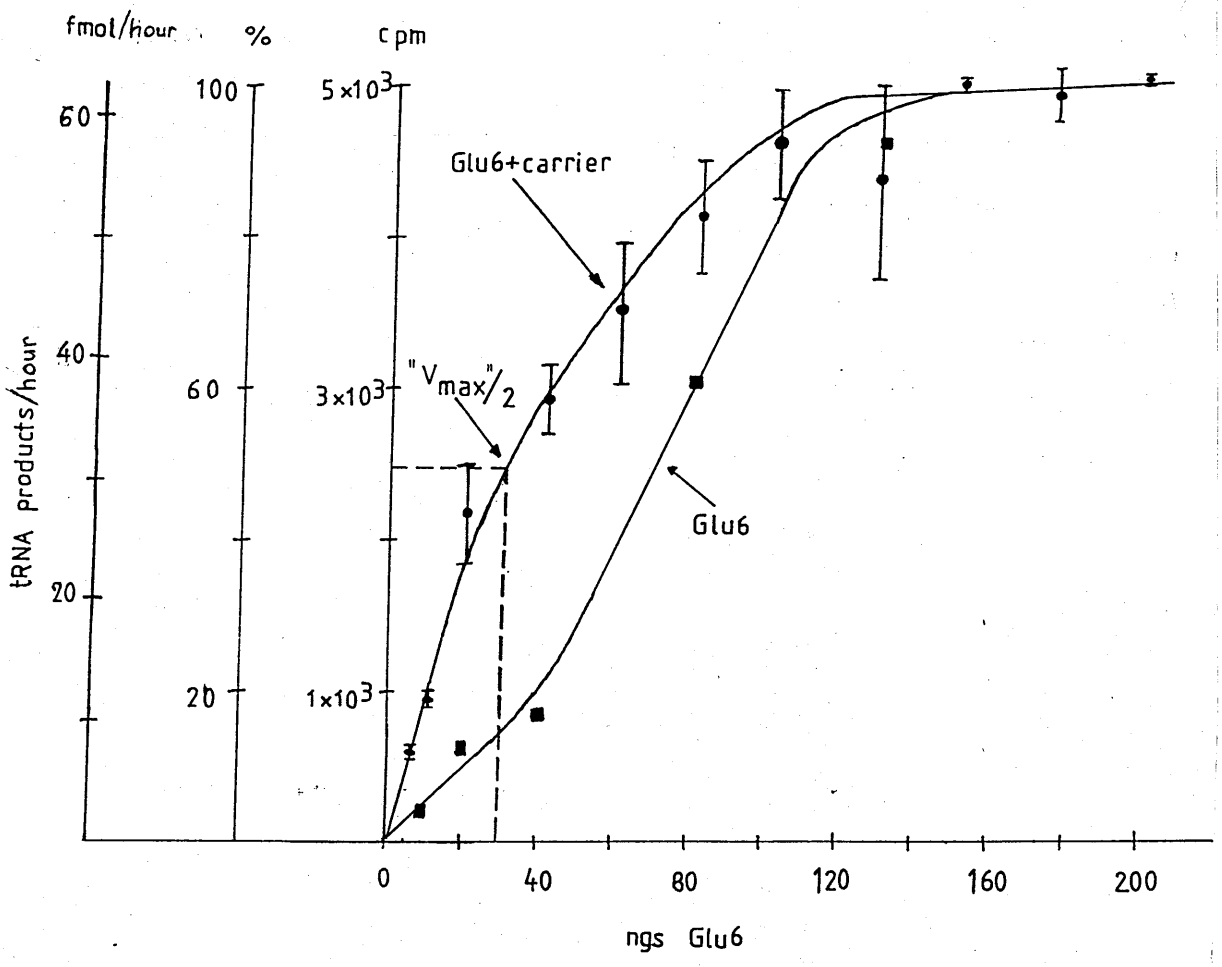
The results of such analysis (see figure 4.16.), show that in the transcription system which was used, MtGlu6 transcription products have an increasing response from 5 to 90 ng of added MtGlu6. However 100 ng of MtGlu6 appear to saturate the reaction and no additional increase in transcription was observed at higher MtGlu6 concentrations (i.e. up to 200 ng/20µl). The transcription products derived from 150 ng added MtGlu6 were found to be 61 *f*mol (see section 3.5.12.4.). Day to day variation was checked by always assaying 40 ng MtGlu6 (in the presence of 110 ng M13mp12) and 150 ng MtGlu6. The results were ±2 *f*mol and ±3 *f*mol respectively, compared to the results shown in figure 4.16. Consequently for comparative purposes the amount of transcripts derived from 150 ng added MtGlu6 was considered as 100% efficiency in subsequent experiments and therefore in all subsequent figures presented in this thesis. 30 ng of MtGlu6 (in the presence of M13mp12 carrier) produced half of the maximum amount of transcripts (i.e. 50% efficiency or "Vmax"/2). The ratio of transcription at this concentration (30 ng/20µl="Vmax"/2) was found to be ~5 transcripts per gene per hour (see section 3.5.12.4.). MtGlu6 samples without added M13mp12 DNA carrier, produced considerably fewer transcripts than the corresponding samples with the M13mp12 carrier present (see figure 4.16). The same experiment performed twice again, under similar experimental conditions, gave similar results (i.e. similar concentration curve and saturation concentration, in the presence of M13mp12 carrier) to those described.

Figure 4.15.
Autoradiograph of transcription of MtGlu6 (DNA concentration assays).



An autoradiograph of products of transcription of 10, 20, 40, 60, 80, 100, 125, 150 ng of MtGlu6 using HeLa S100 extract is shown. The position of *E. coli* crude tRNA marker (M, 74-90 nucleotides) and the position of the wells (W) are indicated. All assays are shown as duplicates.

Figure 4.16.
Analysis of MtGlu6 concentration assays.



The tRNA products in *f*mol, or cpm or percentage (100% = maximum MtGlu6 efficiency) derived from 0-200 ng of MtGlu6 (in the presence of M13mp12 "carrier") are shown (•). The concentration curve derived in the absence of M13mp12 "carrier" is also shown (■). The amount of MtGlu6 which gave half maximal transcription products ("Vmax"/2) is indicated. Error bars indicate deviation between four samples, estimated according to the standard deviation equation.

4.3.3.2. MtGlu6 time course assays.

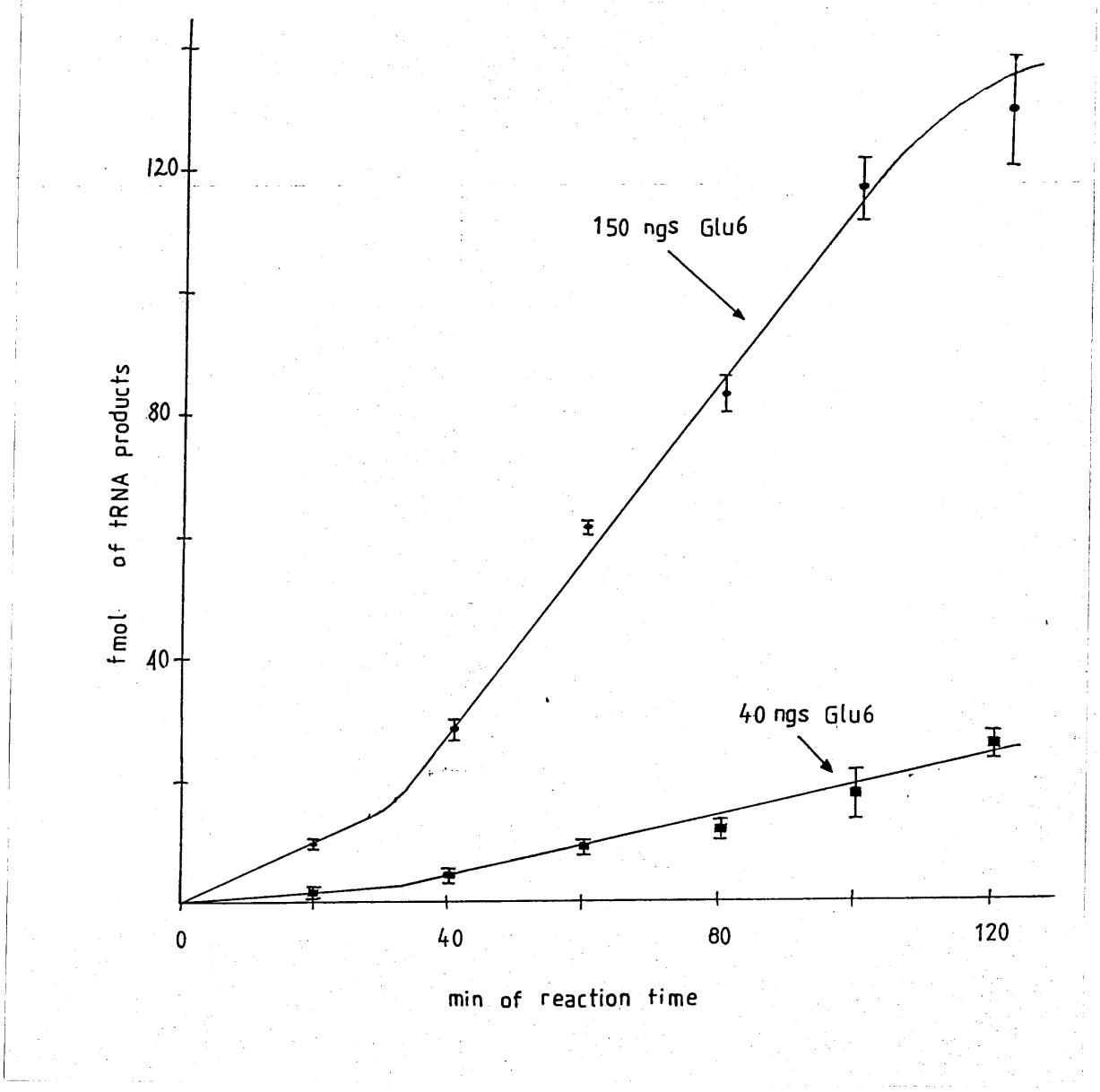
Two different MtGlu6 amounts, 40 ng (in the absence of DNA carrier) as well as 150 ng, were incubated and then processed separately for 0, 20, 40, 60, 80, 100 and 120 minute reaction times at 30° C, as described in section 3.5.12.2. (In these assays a different HeLa S100 extract, prepared and optimized as described in sections 3.5.12.1. and 4.3.1. respectively, was used). Samples were then electrophoresed, autoradiographed and finally the transcripts derived were analyzed as described in sections 3.5.12.3. and 3.5.12.4. The data obtained, presented in figure 4.17., show that the tRNA products from both MtGlu6 amounts, have a linearly increasing response from 35 to 120 minute reaction times.

4.3.3.3. Further investigation of MtGlu6 transcription products.

In order to check whether MtGlu6 transcripts are real RNA polymerase III products or not, a standard transcription assay (see section 3.5.12.2.) of 150 ng MtGlu6, was performed in the presence of 0.5 μg/ml α-amanitin. It is known that α-amanitin inhibits the activity of RNA polymerase II at concentrations as low as 50 ng/ml, whereas RNA polymerase III is affected at much higher concentrations (5 μg/ml; Jacob, 1973). Identical samples, in the absence of α-amanitin, were employed as a control. These assays show that MtGlu6 transcripts are not affected at all by the presence of 0.5 μg/ml α-amanitin (see autoradiograph 1; figure 4.18.).

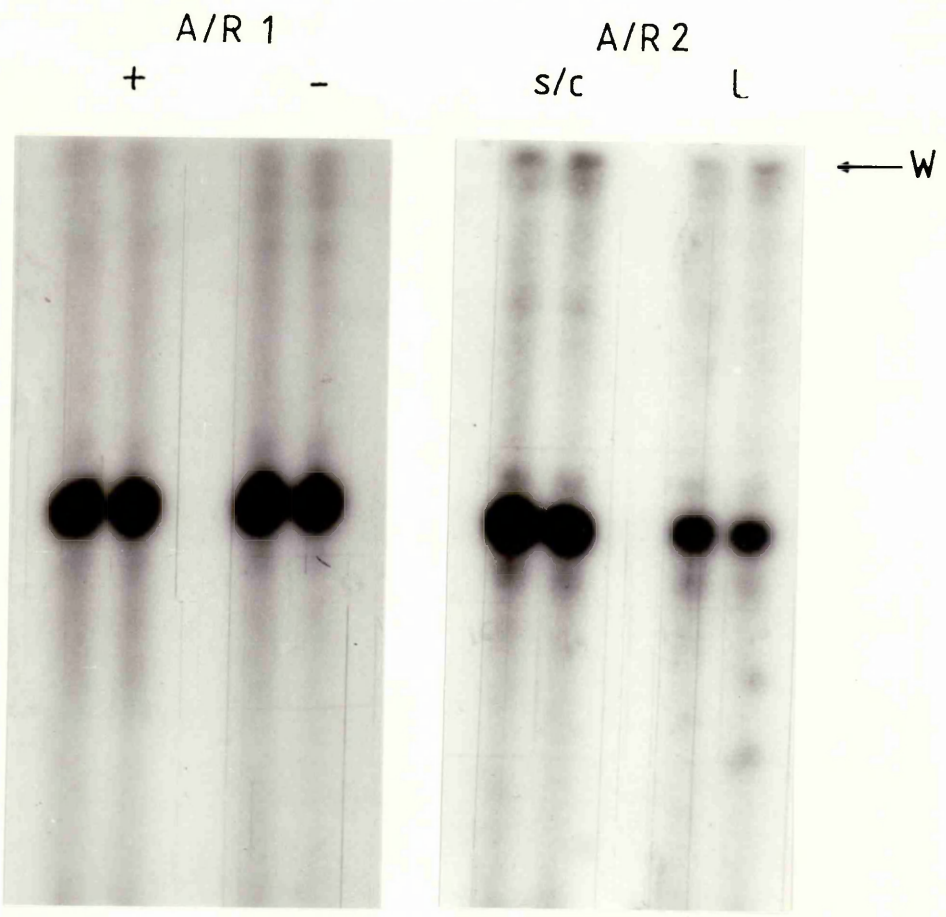
The size of MtGlu6 transcript was also estimated. As described in section 3.5.12.2., the tRNA products derived from all transcription reactions were further processed and electrophoresed in the presence of 250 μg/ml *E. coli* crude tRNA carrier. These tRNAs have sizes from 74 to 95 nucleotides and can be visualized by staining the transcription gel with 10 μg/ml ethidium bromide for 20 min. As a result it is possible to compare directly the position of MtGlu6 transcript detected from the autoradiograph with the positions of *E. coli* tRNAs in the stained gel. Such analysis shows that by using the chosen extract, the MtGlu6 transcript is a 75-80 nucleotides precursor molecule (see also figure 4.15.).

Figure 4.17.
MtGlu6 time course assays.



The tRNA products in *f*mol derived from 40 ng or from 150 ng of MtGlu6 (in absence of M13mp12 DNA "carrier"), from 0, 20, 40, 60, 80, 100 and 120 min reaction times are shown. Error bars indicate deviation between two samples.

Figure 4.18.
Autoradiographs of transcription products of MtGlu6. Transcription in the presence of α-amanitin and transcription of linearized recombinant.



Autoradiograph 1 shows the products of transcription of 150 ng MtGlu6 in the presence (+) or absence (-) of 0.5 μg/ml α-amanitin. Autoradiograph 2 shows the products of transcription of 150 ng of supercoiled MtGlu6 (s/c) or 150 ng of linearized MtGlu6 (l). All assays are shown in duplicates. The position of the wells (W) is indicated.

That was concluded after using the same extract in parallel with other extract preparations which produced both precursor and smaller size MtGlu6 transcripts.

Finally the effect of supercoiling of the MtGlu6 template upon the transcriptional efficiency of its gene, as well as the efficiency of MtGlu6 to transcribe in the presence of another RNA polymerase III template, were tested. These experiments were performed for both MtGlu6 and "MtGlu6 deletion mutants" and they are presented in section 4.3.4.3. and 4.3.4.5. respectively.

4.3.4. The role of the 5'-flanking sequence on the $tRNA^{Glu}$ gene transcription.

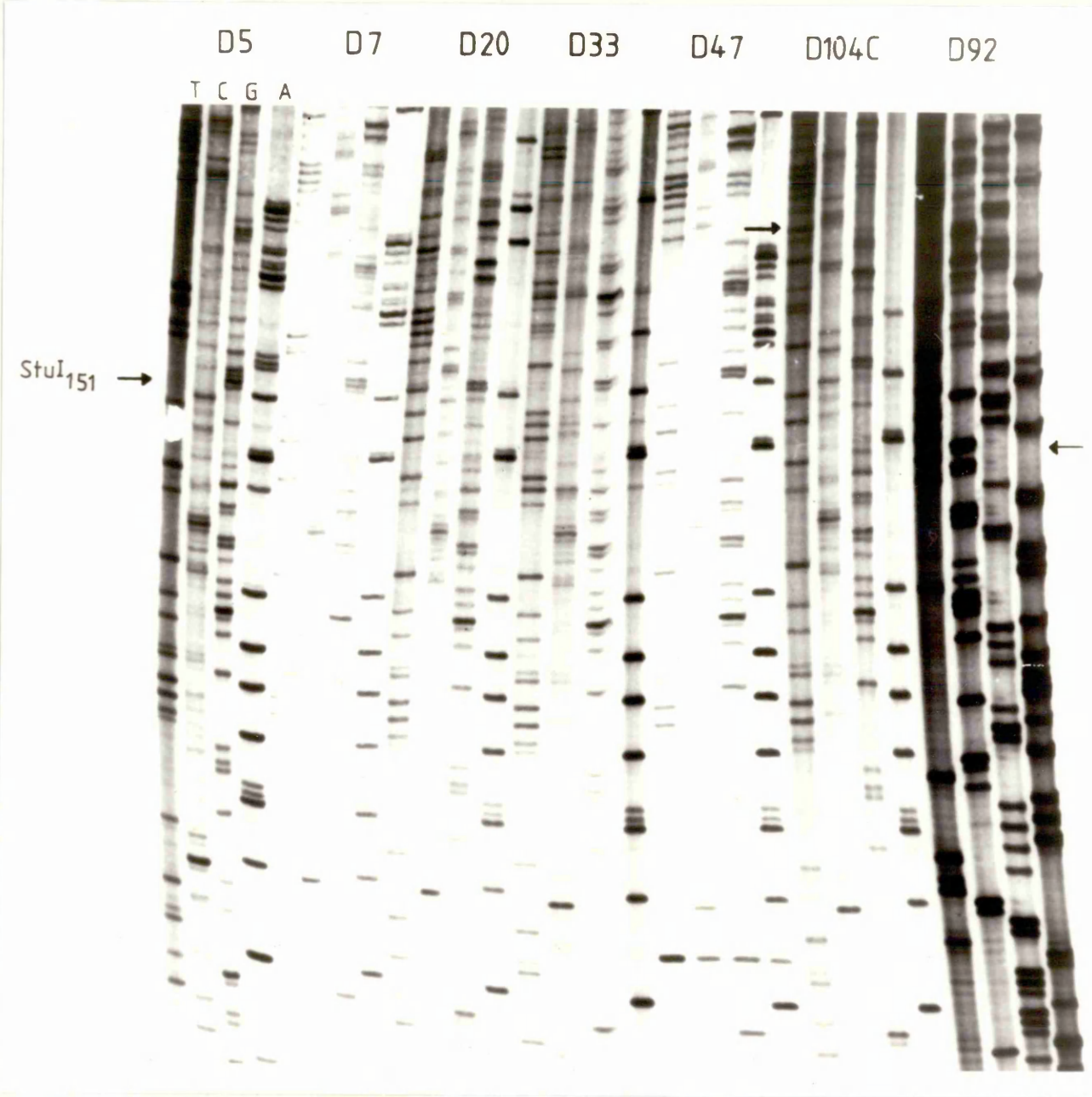
4.3.4.1. Confirmation of nucleotide sequence of "MtGlu6 deletion mutants".

As has been mentioned earlier "MtGlu6 deletion mutants" (see section 3.4.3.) were partly characterized (see section 2.2). In order to confirm the identity of those recombinants two different sets of experiments were performed.

The nucleotide sequence of the mutants was confirmed by the use of chain termination method (Sanger *et al.*, 1980; see section 3.5.11.). These mutants being M13mp9 recombinants were used directly to transform *E. coli* cells. The single stranded DNA templates, which were prepared from mutants MtGluD5, MtGluD7, MtGluD20, MtGluD47 and MtGluD104C, were annealed with the universal M13 primer, while the templates prepared from the mutants MtGluD33, MtGluD44B and MtGluD92 were annealed with the "tailor made" synthetic oligonucleotide "9", which is complementary to 130-146 residues of MtGlu6 (see figure 3.3. and section 3.5.11.4.). All were sequenced as described in section 3.5.11.4. using gradient gels. The resulting autoradiographs (see for example figure 4.19.), confirmed the deleted regions which are presented in figure 3.3. In addition reading of sequencing gels was made against the MtGlu6 sequence and no anomalies detected.

These experiments were performed in order to reconfirm previous data, so no additional nucleotide sequence analysis was attempted. However most of the

Figure 4.19.
Autoradiograph of sequencing gel showing the position of the deletion of mutants MtGluD5, MtGluD7, MtGluD20, MtGluD47, MtGluD104C and MtGluD92.



An autoradiograph of a gradient gel of mutants MtGluD5, MtGluD7, MtGluD20, MtGluD33, MtGluD47, MtGlu104C and MtGluD92 is shown. The position of the deletions is shown by arrows. For mutants MtGluD5, MtGluD7, MtGluD20 and MtGluD47 the point of deletion starts after residue 151 (StuI restriction site) as indicated on the left hand side of the autoradiograph. The point of the deletion for MtGluD33 is not shown. All mutants were annealed with the universal primer, except MtGluD92 which was annealed with synthetic oligonucleotide "9" (see text).

mutants were digested with the restriction enzymes EcoRI+HindIII, the fragments derived were electrophoresed through a 2% agarose gel and their estimated size was found to be as expected. In addition to these confirmation experiments, the DNA concentration of all "MtGlu6 deletion mutants" was accurately estimated. Finally all the recombinants were electrophoresed through agarose gels showing a predominant (>90%) supercoil form of their DNA.

4.3.4.2. "MtGlu6 deletion mutants" concentration assays.

Previous studies on the transcriptional efficiency of genes have almost without exception reported the extent of transcription at only one concentration of template DNA. The preliminary results and those reported below show that the amount of transcript is dependent upon the amount of template DNA in the assay. It was therefore considered wise to compare the amount of transcript produced using different concentrations of $tRNA^{Glu}$ gene or its deletion mutants as templates. Two parameters are used for the sake of this comparison. These are named "Vmax" and "Km" only because the shape of the curve superficially resembles that found in Michaelis-Menten kinetics and do not necessarily imply that the terms are considered in their normal sense or that the assumptions required for Michaelis-Menten kinetics are valid in this case. Thus "Vmax" is considered here to represent the maximum rate of transcript produced for a constant amount of extract (i.e. constant amount of RNA polymerase III and transcription factors) with varying concentrations of template DNA. "Km" is considered to represent the template concentration which results in half of the maximum rate of transcript produced for a constant amount of extract.

In these experiments samples of 0,10, 20, 40, 60, 80, 100, 125 and 150 ng from each mutant, in the presence of M13mp12 DNA "carrier" to give a final concentration of 150 ng/20 µl were assayed as described in section 3.5.12.2. Concurrent with each mutant assayed in this manner, MtGlu6 was also assayed at two concentrations (40 ng in the presence of carrier and 150 ng per 20 µl assay). This allowed the direct comparison of the efficiency of the mutants with MtGlu6. Samples were then electrophoresed, autoradiographed and finally the transcripts derived were analyzed as described in sections 3.5.12.3. and

3.5.12.4. The results of such analysis are presented in figure 4.20.A. and B.

The concentration curves derived from mutants MtGluD20, MtGluD33, MtGluD44B and MtGluD104C as well as their maximum transcriptional efficiencies, reveal no significant difference from the MtGlu6 data which have been reported earlier (see section 4.3.3.1. and figure 4.16.). For mutant MtGluD47, although its maximum transcriptional efficiency was essentially the same as MtGlu6 (105%), a lower "Km" was observed (see table 4.6.). However mutants MtGluD5 and MtGluD7 produced considerably fewer transcripts than MtGlu6 (60% and 74% of MtGlu6 maximum efficiency respectively) although their "Km" was essentially the same as MtGlu6 (see table 4.6.). Finally mutant MtGluD92 was not transcribed at all, presumably due to its lack of residues 1-12 of the $tRNA^{Glu}$ gene, including the first four nucleotides of the "A block" (see figure 3.3.)

In order to ensure that MtGluD47 has lower "Km" than MtGlu6, while MtGluD5 and MtGluD7 have the same "Km" as MtGlu6 but lower "Vmax", double reciprocal plots for those recombinants were attempted. These plots confirm what has been stated above (see figure 4.21.). However it has to be mentioned that since it is not clear whether in the transcription system used there is excess of RNA polymerase III or not, these plots are only to allow comparison between the recombinants used under the described experimental conditions.

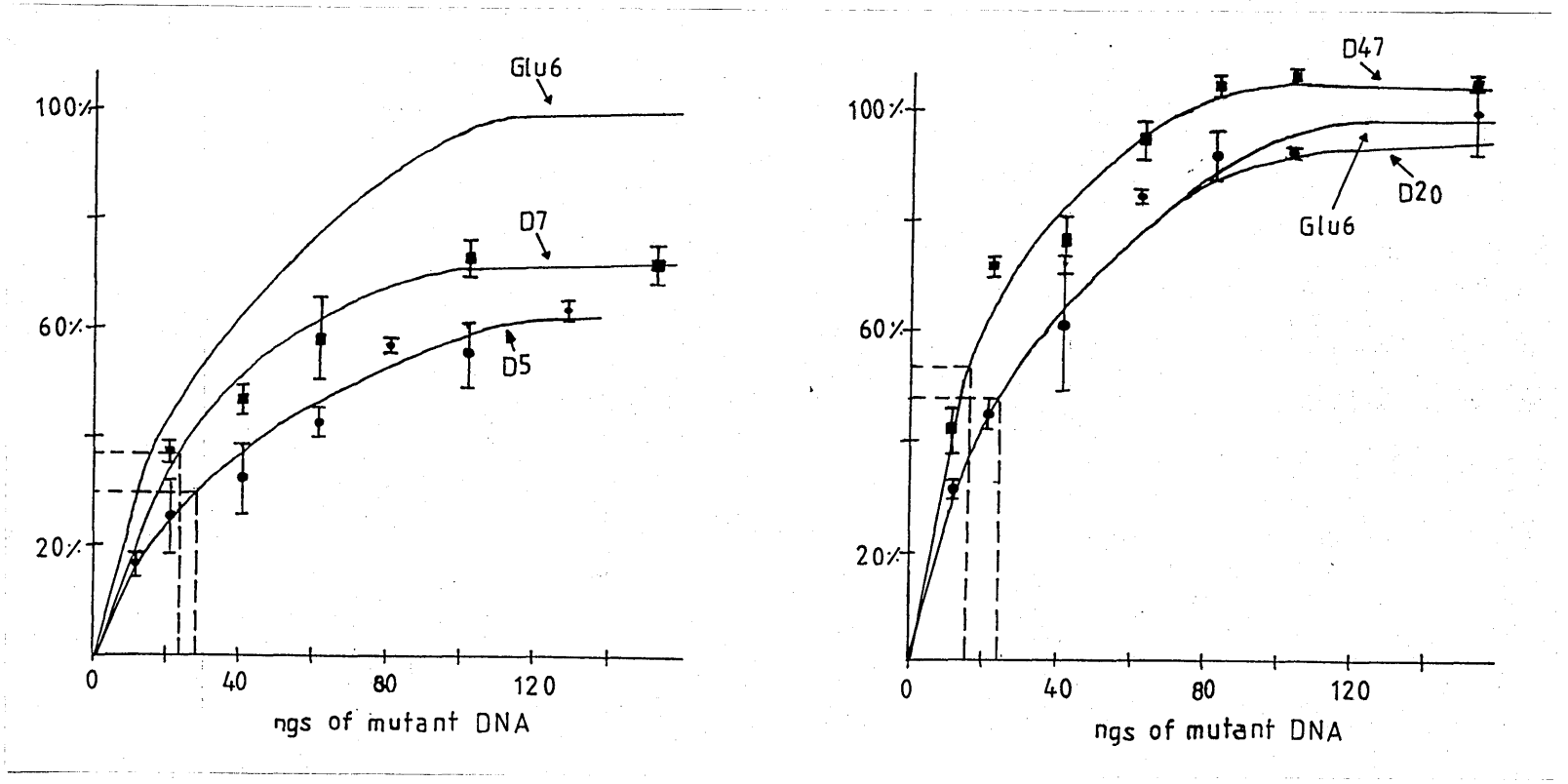
The above experiments were repeated at least once for most deletion mutants, giving similar results to those quoted. A summary of all the concentration assays data is given in table 4.6.

4.3.4.3. The effect of the supercoil form of MtGlu6 and "MtGlu6 deletion mutants", upon their transcriptional efficiency.

The following experimental proœdure was followed to test the effect of the supercoil form of the recombinants, upon their transcriptional efficiency :

250 ng of MtGlu6, or MtGluD5, or MtGluD20, or MtGluD47 were digested with the restriction enzyme BglII which has a unique restriction site in the recombinants, located at residue 6935 of the M13 vector, approximately 700 bp

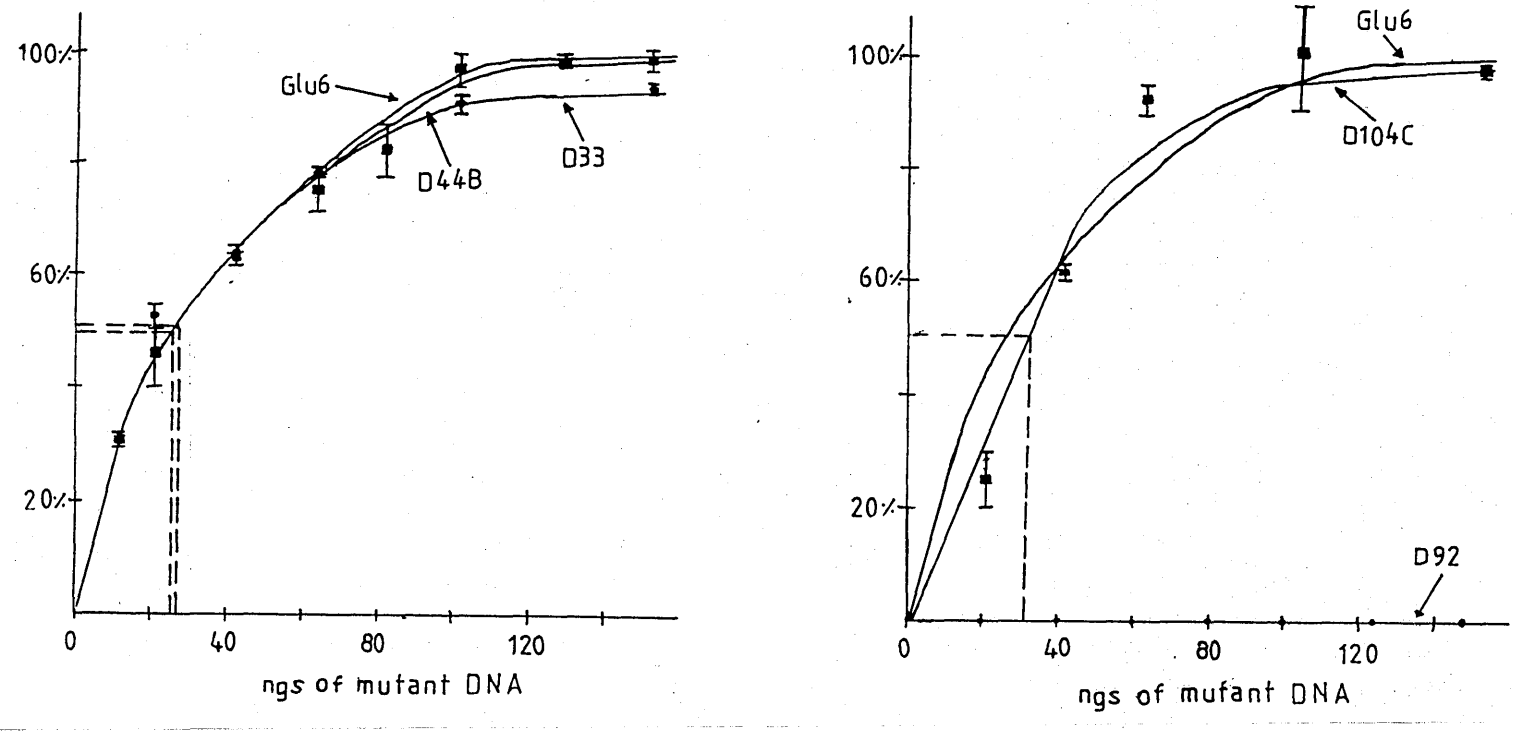
Figure 4.20.A.
MtGluD5, MtGluD7, MtGluD20 and MtGluD47 concentration assays.



The percentage of the tRNA products (100% = maximum transcription of MtGlu6) derived from 0 to 150 ng of MtGluD5 (•), MtGluD7 (■), MtGluD20 (•)or MtGluD47 (■)are shown. The amounts of mutant DNA which give "Vmax"/2 are indicated. For comparative studies the MtGlu6 concentration curve (see section 4.3.3.1. and figure 4.16.) is also shown.
Error bars indicate deviation between two samples.

Figure 4.20.B.

MtGluD33, MtGluD44B, MtGluD92 and MtGluD104C concentration assays.



The percentage of the tRNA products (100% = maximum transcription of MtGlu6) derived from 0 to 150 ng of MtD33 (•), MtD44B (▪), MtD92 (•) or MtD104C (▪)are shown. The amounts of mutant DNA which give "Vmax"/2 are indicated. For comparative studies the MtGlu6 concentration curve (see section 4.3.3.1.and figure 4.16.) is also shown.

Error bars indicate deviation between two samples.

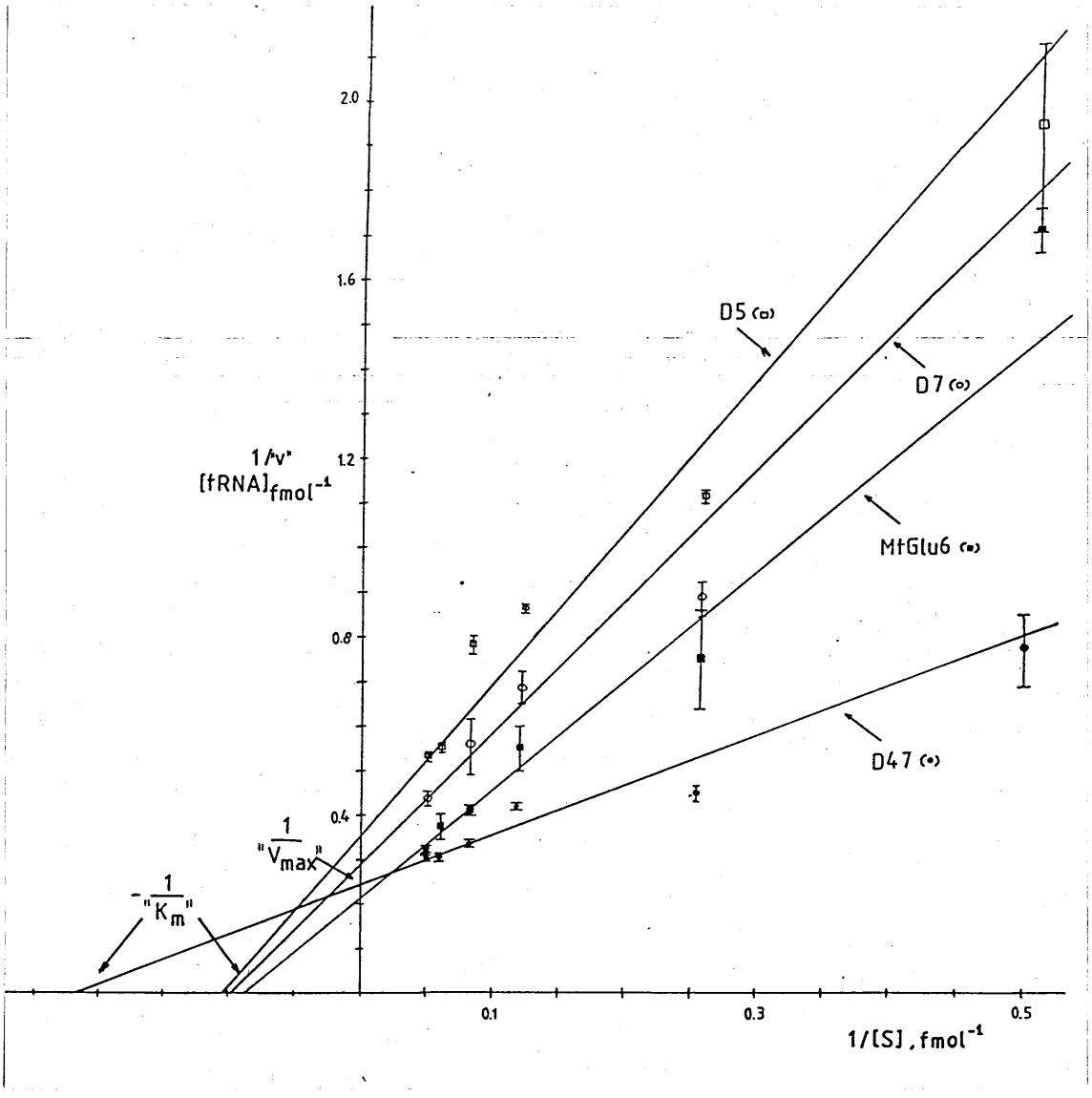
Table 4.6.

A summary of MtGlu6 and "MtGlu6 deletion mutants" concentration assays and competition assays data.

Recombinant	"Vmax" (150 ng DNA)	Relative "Km " (for "Vmax"/2)	Transcription efficiency of 120 ngs M13tT(su^+) in presence of 80 ng recombinant	Transcription efficiency of 80 ng M13tT(su^+) in presence of 120 ng recombinant
MtGlu6	100%	1.00	39%	16%
MtGluD5	60%	0.93	51%	ND
MtGluD7	74%	0.80	47%	ND
MtGluD20	95%	0.83	ND	15%
MtGluD33	93%	0.83	38%	ND
MtGluD44B	99%	0.90	ND	14%
MtGluD47	105%	0.53	ND	14%
MtGluD92	0%	(-)	ND	20%
MtGluD104C	96%	1.03	41%	ND
no recombinant			58%	33%

100% = maximum transcription ("Vmax") of 150 ng MtGlu6.
ND = not determined.
"Km"1.00 = the MtGlu6 concentration which gives "Vmax"/2. The "Km" values for all "MtGlu6 deletion mutants" are expressed relative to MtGlu6 Km.

Figure 4.21.
Double reciprocal plots for MtGlu6, MtGluD5, MtGluD7 and MtGluD47.



The double reciprocal plots for MtGlu6, MtGluD5, MtGluD7 and MtGluD47 are shown. 1/[S] values are expressed in $f mol^{-1}$ of added recombinant in the assay and 1/"V" values are expressed in $f mol^{-1}$ of corresponding tRNA products (see figures 4.16. and 4.20.). The 1/"Vmax" and -1/"Km" values for all recombinants are shown. Error bars indicate deviation between two samples.

from the multiple cloning region (see figure 3.2.). 100 ng from each DNA digestion were electrophoresed through an agarose gel to test for the complete linearization of the DNA. The remaining 150 ng were phenol extracted, ethanol precipitated and used in transcription assays as described in section 3.5.12.2. 150 ng of native supercoil DNA from each of the above recombinants were also processed as a control. Samples were then electrophoresed, autoradiographed and finally the transcripts derived were analyzed as usual. Such analysis revealed that the linearized recombinants MtGlu6, MtGluD5, MtGluD20 and MtGluD47 produced 38%, 32%, 31% and 37% of the transcripts produced by the supercoiled covalently closed circular DNA. A representative autoradiograph is shown in figure 4.18. Some of these experiments were repeated by using another extract preparation, showing a similar decrease in the efficiency of linearized recombinants.

4.3.4.4. Sequential assays.

The DNA concentration assays (see section 4.3.4.2.), show that MtGluD47 has lower "Km" than MtGlu6, while MtGluD5 and MtGluD7 have lower "Vmax" than MtGlu6. In order to investigate more fully these results, sequential assays were performed. These assays were constructed with the view to test the transcriptional efficiency of "MtGlu6 deletion mutants" at DNA concentrations below saturation. Therefore the following experimental procedure was followed :

40 ng of MtGlu6, MtGluD5, MtGluD7, MtGluD20, MtGluD44B or MtGluD47 in the absence of DNA "carrier", were added in a transcription assay which was incubated for 40 minutes. At that time 100 ng of MtGlu6 were added to each sample and the reaction allowed for an additional 20 minutes. Samples were then processed and analyzed as usual. The following efficiencies from these sequential assays were obtained relatively to the products derived from 150 ng of MtGlu6, for a 60 minute reaction time:

40 ng MtGlu6 + 100 ng MtGlu6 = 44 ± 0.5%
40 ng MtGluD5 + 100 ng MtGlu6 = 26 ± 1.5%
40 ng MtGluD7 + 100 ng MtGlu6 = 33 ± 6%
40 ng MtGluD20 + 100 ng MtGlu6 = 47 ± 6%
40 ng MtGluD44B + 100 ng MtGlu6 = 48 ± 3%
40 ng MtGluD47 + 100 ng MtGlu6 = 58 ± 3%

In a very similar experiment to that just described, 40 ng of MtGlu6, MtGluD20, MtGluD44B, or MtGluD47, in the absence of DNA "carrier", were added to the transcription assay which was incubated for 40 minutes. At that time, 100 ng of the same recombinant were added to the corresponding sample and the reaction continued for an additional 20 minutes. Samples were then processed and analyzed as usual. The following efficiencies were obtained, on the base that 100% efficiency corresponds to the products derived from 150 ng of MtGlu6, for a 60 minute reaction time:

40 ng MtGlu6 + 100 ng MtGlu6 = 44 ± 0.5%
40 ng MtGluD20 + 100 ng MtGluD20 = 46 ± 0.5%
40 ng MtGluD44B + 100 ng MtGluD44B = 49 ± 1.5%
40 ng MtGluD47 + 100 ng MtGluD47 = 60 ± 6.0%

In the last sequential assay, a series of samples of 100 ng of MtGlu6, in the absence of DNA "carrier", were added to a transcription assay which was incubated for 40 minutes. At that time, 40 ng of MtGlu6, MtGluD5, MtGluD20, MtGluD33, MtGluD44B, MtGluD47 or MtGluD104C, were added separately to each of the 100 ng MtGlu6 samples and the reaction continued for an additional 20 minutes. Samples were then processed and analyzed as described previously; finally the following efficiencies were obtained:

100 ng MtGlu6 + 40 ng MtGlu6 = 59 ± 2%
100 ng MtGlu6 + 40 ng MtGluD5 = 56 ± 3%
100 ng MtGlu6 + 40 ng MtGluD20 = 57 ± 8%
100 ng MtGlu6 + 40 ng MtGluD33 = 56 ± 6%
100 ng MtGlu6 + 40 ng MtGluD44B = 58 ± 1%
100 ng MtGlu6 + 40 ng MtGluD47 = 58 ± 6%
100 ng MtGlu6 + 40 ng MtGluD104C = 52 ±10%

The data obtained from the first two experiments show that the assays containing MtGluD47 as initial template produced more transcripts than the

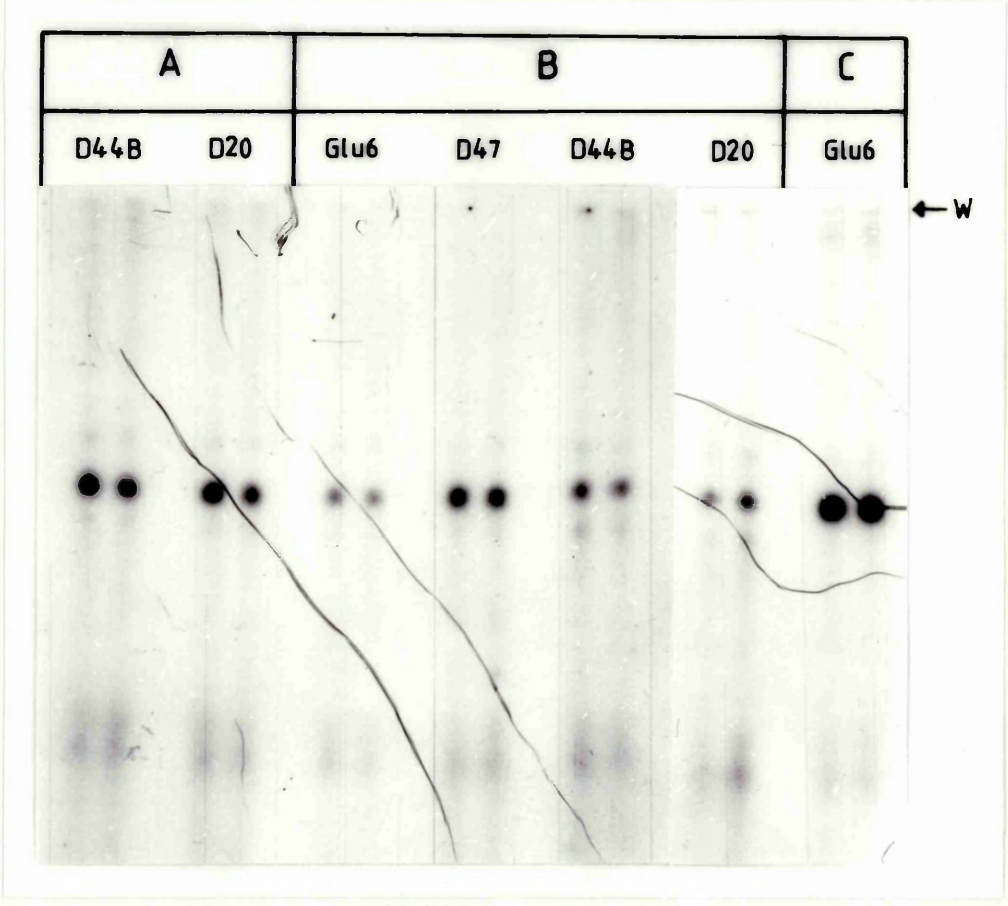
assays containing MtGlu6 as initial template, while the assays containing MtGluD5 and MtGluD7 as initial templates produced fewer transcripts. The same experiments also show that the nature of the DNA added after the initial 40 minute reaction time does not really affect the rate of transcripts produced from samples containing the same initial template. For example the 40 ng MtGluD44B + 100 ng MtGlu6 assay produced the same amount of transcripts as the 40 ng MtGluD44B +100 ng MtGluD44B assay. In addition the last experiment indicates that the nature of the second DNA added in the assays does not affect substantially the total amount of transcripts produced. For example although MtGluD47 is a more efficient template than MtGluD5, the 100ng MtGlu6+40 ng MtGluD5 assay produced the same amount of transcripts as the 100 ng MtGlu6+40 ng MtGluD47 assay. As a result, the different efficiencies obtained from the first experiment may derive from the different efficiency of the initial template added to the assay. The latter implies that MtGluD47 is a more efficient template than MtGlu6, and MtGluD5 and MtGluD7 are less efficient templates at DNA concentration of 40ng/20µl assay. However the transcription products from these assays were not estimated prior to the addition of the second DNA. Such a procedure would have ensured that the different transcriptional efficiencies obtained from some of the assays, were due to the initial 40 minute reaction products. A representative autoradiograph is shown in figure 4.22. A summary of the results obtained from all the sequential assays is presented in table 4.7.

4.3.4.5. Competition assays.

In these transcription assays MtGlu6 or "MtGlu6 deletion mutants" were employed to compete for RNA polymerase III and/or transcription factors with another tRNA gene. The gene used for this purpose was an amber suppressor tRNA gene, derived from a *X. Laevis* $tRNA^{Tyr}_{GTA}$ gene, (Laski *et al.*, 1982a and 1982b). This suppressor (su)tRNA gene (M13tT-(su^+), see section 3.4.5.) has a thirteen nucleotide intervening sequence in the anticodon loop, starting after position 37. The total length of M13tT-(su^+) gene is 86 nucleotides, which helps to distinguish its transcription products from those derived from MtGlu6 or "MtGlu6 deletion mutants".

In the first competition experiment, 0, 20, 40, 80 and 120 ng of MtGlu6,

Figure 4.22.
Autoradiograph of transcription of MtGluD20, MtGluD44B, MtGluD47 and MtGlu6 (DNA sequential assays).



The autoradiograph shows the transcription products of :
A. 100 ng of MtGlu6 for 40 min reaction time, followed by the addition of 40 ng of MtGluD44B or MtGluD20 and extension of the reaction for further 20 min.
B. 40 ng of MtGlu6 or MtGluD47 or MtGluD44B or MtGluD20 for 40 min reaction time, followed by the addition of 100 ng of MtGlu6 and extension of the reaction for further 20 min.
C. 150 ng of MtGlu6 for 60 min reaction time .
All assays are shown as duplicates. The position of the wells (W) is indicated.

Table 4.7.
Summary of MtGlu6 and "MtGlu6 deletion mutants" sequential assays data.

A. 40 ng recombinant, 40 min reaction time; addition of 100 ng MtGlu6, 60 min total reaction time.

1st DNA	2nd DNA	Maximum transcription
MtGlu6	MtGlu6	44 %
MtGluD5	MtGlu6	26 %
MtGluD7	MtGlu6	33 %
MtGluD20	MtGlu6	47 %
MtGluD44B	MtGlu6	48 %
MtGluD47	MtGlu6	58 %

B. 40 ng recombinant, 40 min reaction time; addition of 100 ng of the same recombinant, 60 min total reaction time.

1st DNA	2nd DNA	Maximum transcription
MtGlu6	MtGlu6	44 %
MtGluD20	MtGluD20	46 %
MtGluD44B	MtGluD44B	49 %
MtGluD47	MtGluD47	60 %

C. 100 ng MtGlu6, 40 min reaction time; addition of 40 ng recombinant, 60 min total reaction time.

1st DNA	2nd DNA	Maximum transcription
MtGlu6	MtGlu6	59 %
MtGlu6	MtGluD5	56 %
MtGlu6	MtGluD20	57 %
MtGlu6	MtGluD33	56 %
MtGlu6	MtGluD44B	58 %
MtGlu6	MtGluD47	58 %
MtGlu6	MtGluD104C	52 %

100% = transcription products derived from 150 ng MtGlu6 from 60 min reaction time.

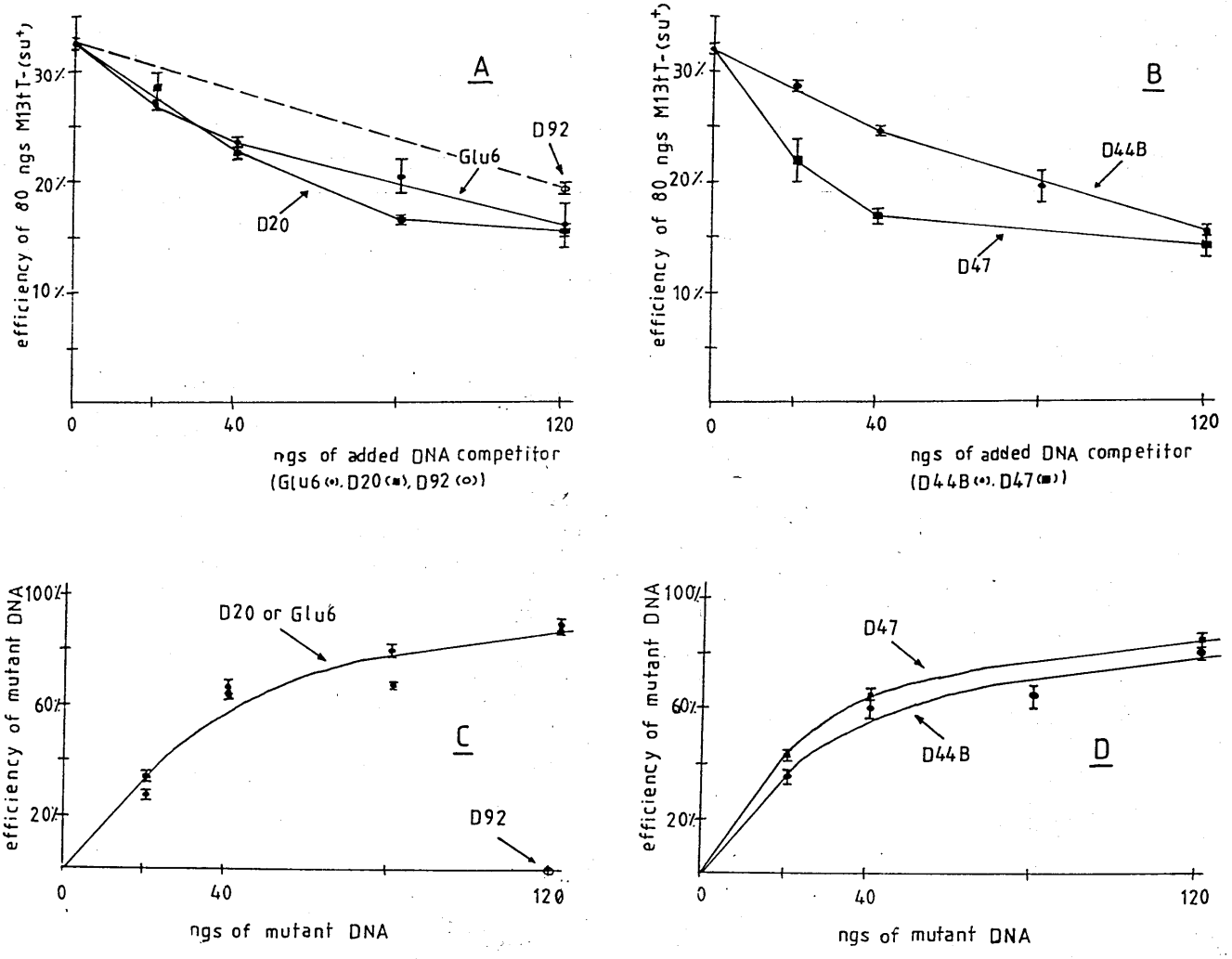
MtGluD20, MtGluD44B, or MtGluD47, or 120 ng of MtGluD92, in the presence of M13mp12 DNA "carrier" to give a final concentration of 120 ng/20 µl, were assayed together with 80 ng of M13tT-(su^+), as described in section 3.5.12.2. 150 ng of MtGlu6 in the absence of M13tT-(su^+) were also processed as a control. Samples were then electrophoresed, autoradiographed and further analyzed as usual.

The data obtained from this experiment, presented in figure 4.23.A., show that the transcription products derived from 80 ng of M13tT-(su^+) in the absence of MtGlu6 or of the "MtGlu6 deletion mutants", were only 33% of the transcription products derived from 150 ng of MtGlu6. However when 80 ng of M13tT-(su^+) were assayed together with 120 ng of MtGlu6, M13tT-(su^+) efficiency was reduced two-fold (16%); the ratio of 120 ng MtGlu6/80 ng M13tT-(su^+) transcription products was 5.7/1.0. When the same amount of M13tT-(su^+) was assayed together with 120 ng of MtGluD20, MtGluD44B, MtGluD47 or MtGluD92 its efficiency was reduced from 33% to 15%, 14%,14% and 20% respectively.

In the last of these series of experiment, 0, 20, 40, 60 and 80 ng of MtGlu6, MtGluD5, MtGluD7, MtGluD33, or MtGluD104C, in the presence of M13mp12 DNA "carrier" to give a final concentration of 80 ng/20 µl, were assayed together with 120 ng of M13tT-(su^+), as described in section 3.5.12.2. 150 ng of MtGlu6 in the absence of M13tT-(su^+) were also processed as a control. Samples were then electrophoresed, autoradiographed and further analyzed as usual.

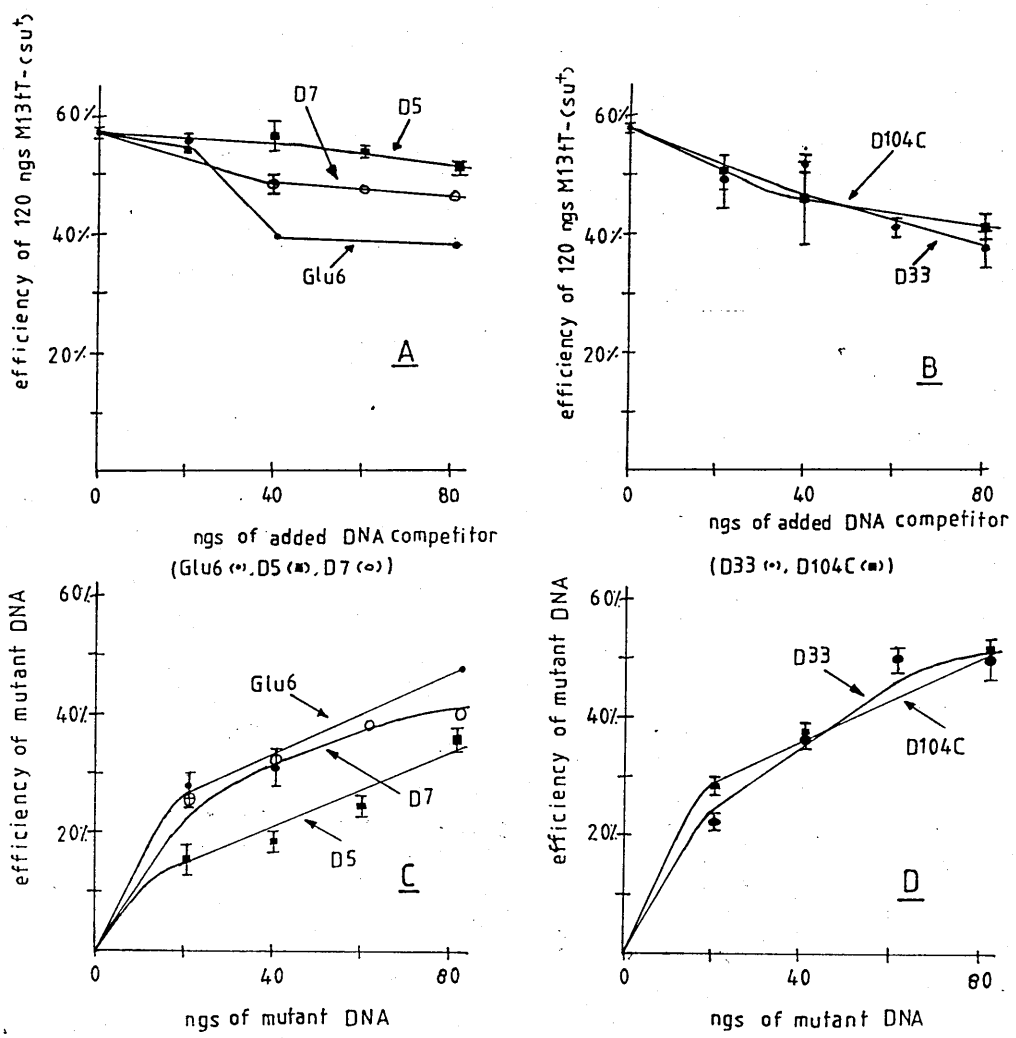
The data obtained, presented in figure 4.23.B., show that the transcription products derived from 120 ng of M13tT-(su^+) in the absence of MtGlu6 or of its "deletion mutants", were 58% of the transcription products derived from 150 ng of MtGlu6. However when 120 ng of M13tT-(su^+) were assayed together with 80 ng of MtGlu6, M13tT-(su^+) efficiency was reduced to 39%; the ratio of 80 ng MtGlu6/120 ng M13tT-(su^+) transcription products was 1.3/1.0. When the same amount of M13tT-(su^+) was assayed together with 80 ng of MtGluD5, MtGluD7, MtGluD33 or MtGluD104C, its efficiency was reduced from 58% to 51%, 47%, 38% and 41% respectively. A summary of the results achieved from both competition assays is given in table 4.6. A representative autoradiograph is shown in figure 4.24.

Figure 4.23.A.
MtGlu6,MtGluD20, MtGluD47, MtGluD92 competition assays.



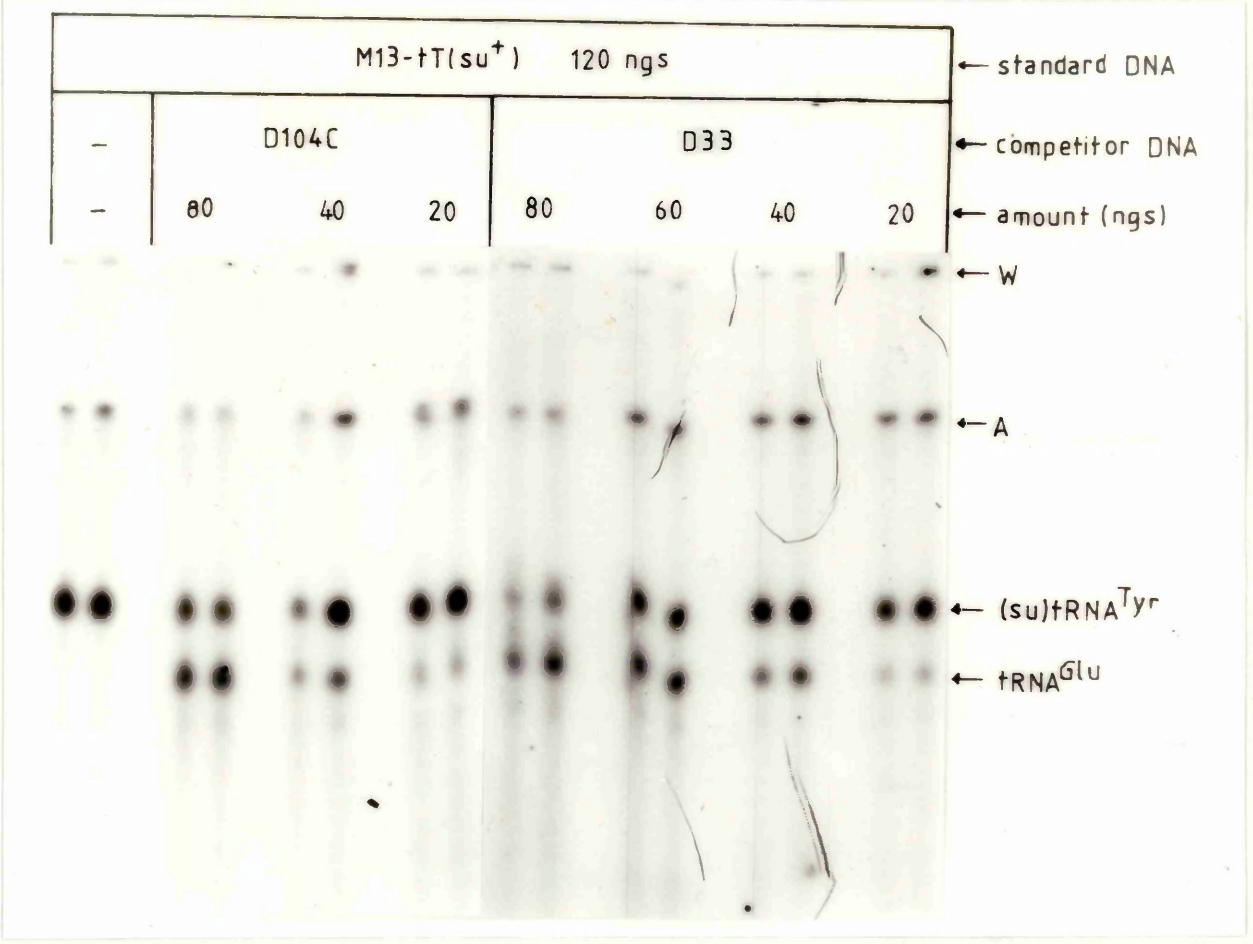
Graphs A and B show the decrease of transcription products of 80 ng M13tT-(su+) in the presence of 0 to 120 ng of MtGlu6, MtGluD20, MtGlu44B, MtGluD47 or MtGluD92 (for MtGluD92 only one concentration, 120 ng/20µl, was tested).
Graphs C and D show the transcription products derived from 0 to 120 ng of the same mutants (MtGlu6, MtGluD20, MtGluD44B, MtGluD47 and MtGluD92) when assayed together with 80 ng M13tT-(su+).
100% = maximum transcription of 150 ng MtGlu6 (see figure 4.16.).
Error bars indicate deviation between two samples.

Figure 4.23.B.
MtGlu6, MtGluD5, MtGluD7, MtGluD33 and MtGluD104C competition assays.



Graphs A and B show the decrease of transcription products of 120 ng M13 tT-(su$^+$) in the presence of 0 to 80 ng of MtGlu6, MtGluD5, MtGluD7, MtGluD33 or MtGluD104C.
Graphs C and D show the transcription products derived from 0 to 80 ng of the same mutants (MtGlu6, MtGluD5, MtGluD7, MtGluD33, MtGluD104C) when assayed together with 120 ng M13tT-(su$^+$).
100% = maximum transcription of 150 ng MtGlu6 (see figure 4.16.)
Error bars indicate deviation between two samples.

Figure 4.24.
Autoradiograph of transcription of MtGluD33, MtGluD104C and M13tT-(su$^+$) (DNA competition assays).



The autoradiograph shows the transcription products of : 120 ng of M13-tT(su$^+$) (standard DNA); 120 ng of M13-tT(su$^+$) assayed together with 80 ng or 40 ng or 20 ng of MtGluD104C; 120 ng of M13-tT(su$^+$) assayed together with 80 ng or 60 ng or 40 ng or 20 ng of MtGluD33. The position of $tRNA^{Glu}$ and $(su)tRNA^{Tyr}$ transcripts are shown. Transcript A is a RNA polymerase III 200-250 nucleotides product and derives from M13-tT(su$^+$). All assays are shown in duplicates. The position of the wells (W) is indicated.

CHAPTER 5

DISCUSSION

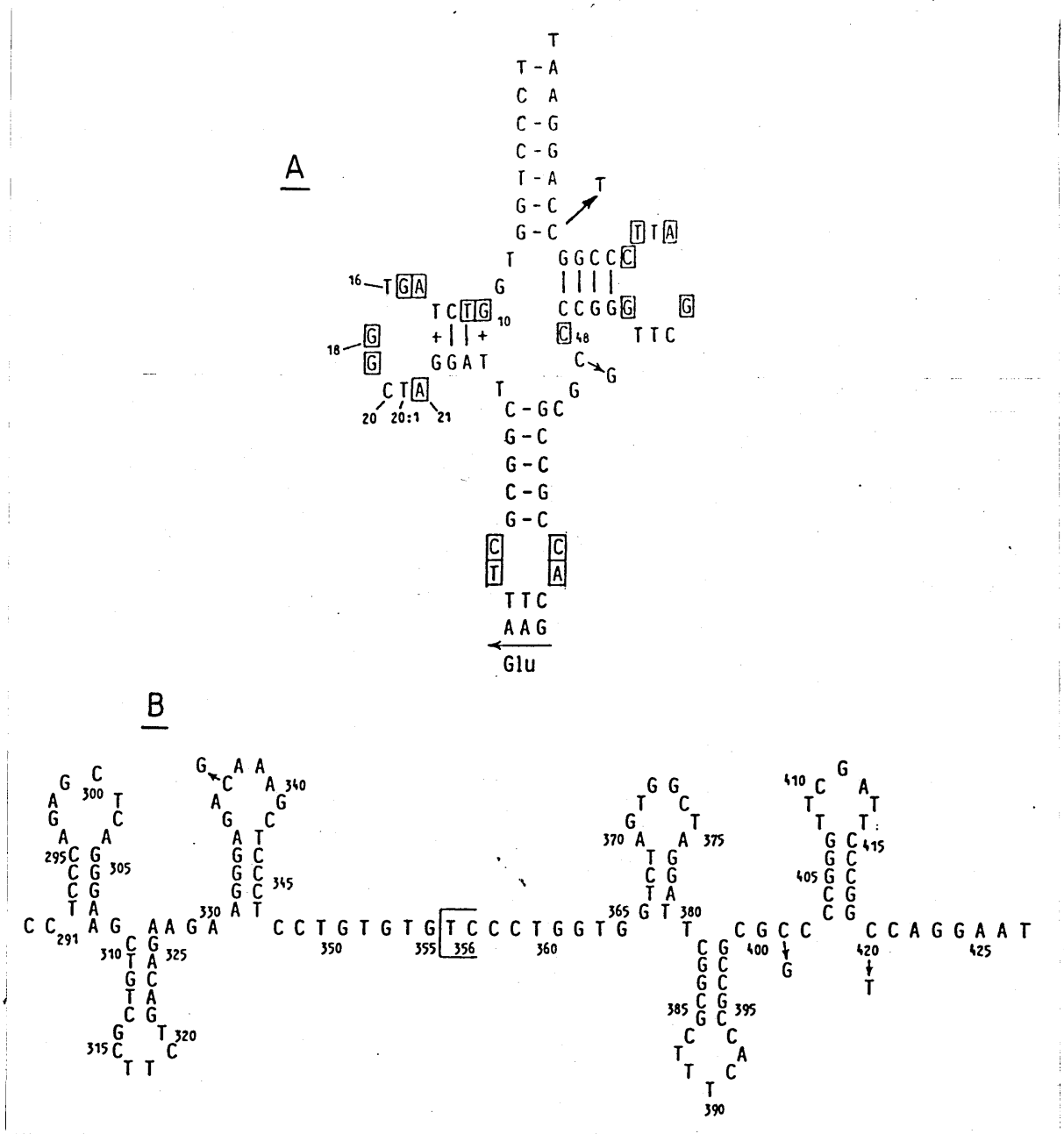
An overall analysis of the results obtained during this study will be presented in this chapter. Section 5.1. examines the structure of a second human $tRNA^{Glu}$ gene copy and compares this gene and its flanking sequences with other mammalian $tRNA^{Glu}$ genes and their genomic environments. Section 5.2. focuses on the organization of the $tRNA^{Glu}$ genes in the human genome. Finally section 5.3. concentrates on the *in vitro* expression of human $tRNA^{Glu}$ genes; section 5.3.1. examines the nature and the properties of the transcription of a $tRNA^{Glu}$ gene and section 5.3.2. concludes on the role of the 5'-flanking sequence on the transcription of the same gene.

5.1. Comparison of the two characterized human $tRNA^{Glu}$ gene copies and their flanking sequences.

5.1.1. Structure of a second human $tRNA^{Glu}$ gene copy; comparison with other mammalian $tRNA^{Glu}$ gene copies.

The nucleotide sequence analysis of 1 kb of sequence within the plasmid pTC51 (see section 4.1.2. and figure 4.4.) revealed the presence of a second human $tRNA^{Glu}$ gene copy. The identification of the gene was facilitated by the use of a computer program to search for tRNA genes (Staden, 1980). This copy differs in two nucleotides from the first characterized $tRNA^{Glu}$ gene copy (Goddard *et al.*, 1983); these two nucleotides (residues 47 and 66 of the RNA-like strand; see figure 5.1.) are not invariant or semi-invariant nor are they part

Figure 5.1.
Sequence of the $tRNA^{Glu}$ genes and tRNA-like structures.



A. The sequence of the $tRNA^{Glu}$ gene from λht137 (tRNA-like strand) arranged in a cloverleaf structure with the yeast $tRNA^{Phe}$ numbering is shown. Invariant and semi-invariant nucleotides are boxed. Differences found in the $tRNA^{Glu}$ gene present in λht190 are shown.

B. A possible secondary structure of the RNA-like strand of the 5'-flanking sequence (residues 292-347) and the $tRNA^{Glu}$ gene (residues 356-427) from λht 137 are shown. Differences found in the corresponding λht190 sequence are shown. Residues are numbered according to Goddard *et al.*, (1983).

of the obligatory intragenic control regions. Both copies belong to the same isoacceptor $tRNA^{Glu}$ gene family (anticodon TTC) and also have an unusual structural feature, a C_2-A_{71} pair in the acceptor stem, as well as a T at position 73 instead of a purine nearly always found in $tRNA^{Glu}$ gene sequences (Sprinzl *et al.*,1987). Figure 5.1. shows the sequence of both $tRNA^{Glu}$ gene copies, as a cloverleaf.

The second human $tRNA^{Glu}$ gene copy differs in 3 nucleotides from a characterized human $tRNA^{Glu}_{CUC}$ (Smardo and Calvet, 1987b); in 4 or 5 nucleotides from 5 characterized copies (3 and 2 copies respectively) of a rat $tRNA^{Glu}_{CTC}$ gene (Shibuya *et al.*, 1982 and Rosen *et al.*, 1984); and in 4 nucleotides from a mouse $tRNA^{Glu}_{CTC}$ gene (Looney and Harding, 1983). Much greater divergence (17-29 nucleotides) was observed when the above human $tRNA^{Glu}$ gene was compared with $tRNA^{Glu}$ genes from lower eukaryotes (Hosbach *et al.*, 1980; Indik and Tartof, 1982; Eigel *et al.*, 1981; Feldmann *et al.*, 1981; Hauber *et al.*, 1988; Stucka *et al.*, 1987; Gamulin *et al.*, 1983; Corlet *et al.*, 1986).

The two unusual features of the $tRNA^{Glu}$ gene sequences which are noted above led Smardo and Calvet (1987a) to suggest that the original sequence (Goddard *et al.*, 1983) may be a pseudogene, probably derived from a $tRNA^{Glu}_{CTC}$ gene. They suggest that this sequence is derived from a functional $tRNA^{Glu}_{CTC}$ gene by acquisition of six base changes, including one which converted the anticodon from CTC to TTC. However there are several reasons to believe that both copies are "true" genes. There are many cases where tRNA genes from one isoacceptor family resemble genes or even the tRNA from another isoacceptor family: a human $tRNA^{Ser}_{AGA}$ gene differs in just 2 nucleotides from a human $tRNA^{Ser}_{TGA}$ gene (Krupp *et al.*, 1988), while a transcriptionally active human $tRNA^{Gly}_{CCC}$ gene differs in 3 nucleotides from the human $tRNA^{Gly}_{GCC}$ (Shortridge *et al.*, 1985). Similar observations have also been reported in other eukaryotes, by Sekiya *et al.*, (1982) and by Cribbs *et al.*, (1987). In addition, most human $tRNA^{Glu}$ gene copies (including the two characterized here) appear to belong to 2-3 kb highly homologous DNA fragment repeats, and have very similar flanking sequences (see sections 4.2. and 5.2.). However the flanking sequences of the two

characterized human $tRNA^{Glu}_{TTC}$ genes show no sequence homology with the flanking sequences of the above mentioned sequenced mammalian $tRNA^{Glu}_{CTC}$ gene copies which are also highly conserved not only between different copies in rat, but also between rat and mouse (see also section 1.4.4.1.). As a result the possibility of duplication of an ancestral sequence in mammals, containing a $tRNA^{Glu}_{CTC}$ gene, from which eventually both the human $tRNA^{Glu}_{TTC}$ and the rat and mouse $tRNA^{Glu}_{CTC}$ gene copies were derived, seems unlikely. Moreover since no human $tRNA^{Glu}_{UUC}$ has been characterized so far it is not possible at present to compare the sequence of the characterized $tRNA^{Glu}_{TTC}$ gene copies with the sequence of the corresponding $tRNA^{Glu}_{UUC}$. Even when this is possible, uncertainties may remain if the major tRNA species differs from the gene sequence. For example, as stated earlier, all characterized copies of the rat $tRNA^{Asp}$ gene (eight of an estimated total of ten) differ from the major $tRNA^{Asp}$ sequence (see also section 1.3.).

The first $tRNA^{Glu}$ gene copy (Goddard *et al.*, 1983) is remarkably active *in vitro* , producing more tRNA than a cluster of seven active *Xenopus* tRNA genes, when micro-injected into *Xenopus* oocyte nuclei (Goddard *et al.*, 1983). The work reported here shows that both human $tRNA^{Glu}$ gene copies were also transcribed with similar high efficiency using HeLa S100 extracts (see section 4.3.2. and figure 4.14.). In all the known literature up to date tDNA sequences have been named pseudogenes because either they display considerable sequence difference to the corresponding tRNA (e.g. are lacking invariant or semi-invariant nucleotides or even a termination site), or they are truncated, or they are inactive *in vitro* (see also section 1.3.1.). According to all these criteria both human $tRNA^{Glu}$ gene copies are "true" genes. However it will be worth testing whether the processed transcripts of these genes can be charged with glutamic acid. Moreover the characterization of the remaining $tRNA^{Glu}$ gene copies will clarify whether the unusual features found in these two $tRNA^{Glu}_{TTC}$ gene copies are conserved in all copies or not. Since nearly all copies appear to belong to highly homologous DNA fragments, the unidentified copies could be amplified by the polymerase-chain reaction method for cloning and sequencing (Engelke *et al*, 1988).

5.1.2. Comparison of the flanking sequences of the two $tRNA^{Glu}$ gene copies.

Comparative analysis of the nucleotide sequence data from plasmids pLB4 (first copy) and pTC51 (second copy) shows a 97.3 % homology from a region which extends at least 350 bp upstream and over 500 bp downstream of the two $tRNA^{Glu}$ gene copies (see figure 4.4.). Additional sequence data to residue 1100 were obtained for one strand only and were therefore not included in figure 4.4. This sequence also showed high homology with the corresponding sequence in pLB4. All nucleotides differences which were found, appeared as single nucleotides rather than as contiguous sequences (see section 4.1.2.). However, two "hot spot" regions were observed (residues 60-160 and 550-650; see figure 4.4.) where 7 and 9 point mutations were found respectively. In addition the construction of the restriction maps of λht137 and λht190 showed that the two characterized $tRNA^{Glu}$ gene copies belong to highly homologous DNA fragment repeats of at least 10.7 kb in the human genome (see section 4.1.3. and figure 4.8.). However it is still to be answered whether the high homology of the DNA fragments represented by the inserts of the two λ recombinants is present in regions much longer than 10.7 kb.

Both copies have potential tRNA-like structures in their 5'-flanking sequences which are identical in all but 1 nucleotide (see figure 5.1.). It would be misleading to call these sequences tRNA pseudogenes (see section 1.3.1.). Both sequences have a TTC present in the region which corresponds to the anticodon loop, as well as similar relative spacing of the arms to those found in tRNA genes. However the "anticodon" and "TψC" loops do not have the expected number of residues (6 and 8 respectively, instead of the normal 7) and the proportion of invariant or semi-invariant nucleotides (no more than 50 %) is such that we can not consider these sequences tRNA pseudogenes.

5.2. Organization of tRNA^{Glu} gene copies in the human genome.

5.2.1. The human tRNA^{Glu} gene copy number.

The experiments described in section 4.2.1. showed that the copy number for the human tRNA^{Glu} gene family is thirteen. The copy number which was found appeared to be almost consistent (12-14) in all experiments which were performed during this study (see sections 4.2.1. and 4.2.2. and tables 4.4. and 4.5.). This number is in between the expected average number of copies (10-20) for human tRNA gene families (Hatlen and Attardi, 1971) and is remarkably similar to those found so far for all other human tRNA genes, with the exception of the single copy gene and pseudogene for the suppressor phosphoserine tRNA (O'Neil *et al.*, 1985; McBride *et al.*, 1987). Santos and Zasloff (1981) have reported about 12 human tRNA$_i^{Met}$ gene copies; Doering *et al.*, (1982) have reported approximately 14 tRNA$_3^{Lys}$ genes; Arnold *et al.*, (1986) detected at least 13 tRNA^{Val} loci and Van Tol and Beier (1988) detected at least 12 gene loci for tRNA^{Tyr} in the human haploid genome. All copies of the gene for tRNA^{Glu} have been assigned to human chromosome 1 (Boyd *et al.*, 1989).

5.2.2. Conclusions on the organization of the human tRNA^{Glu} gene family.

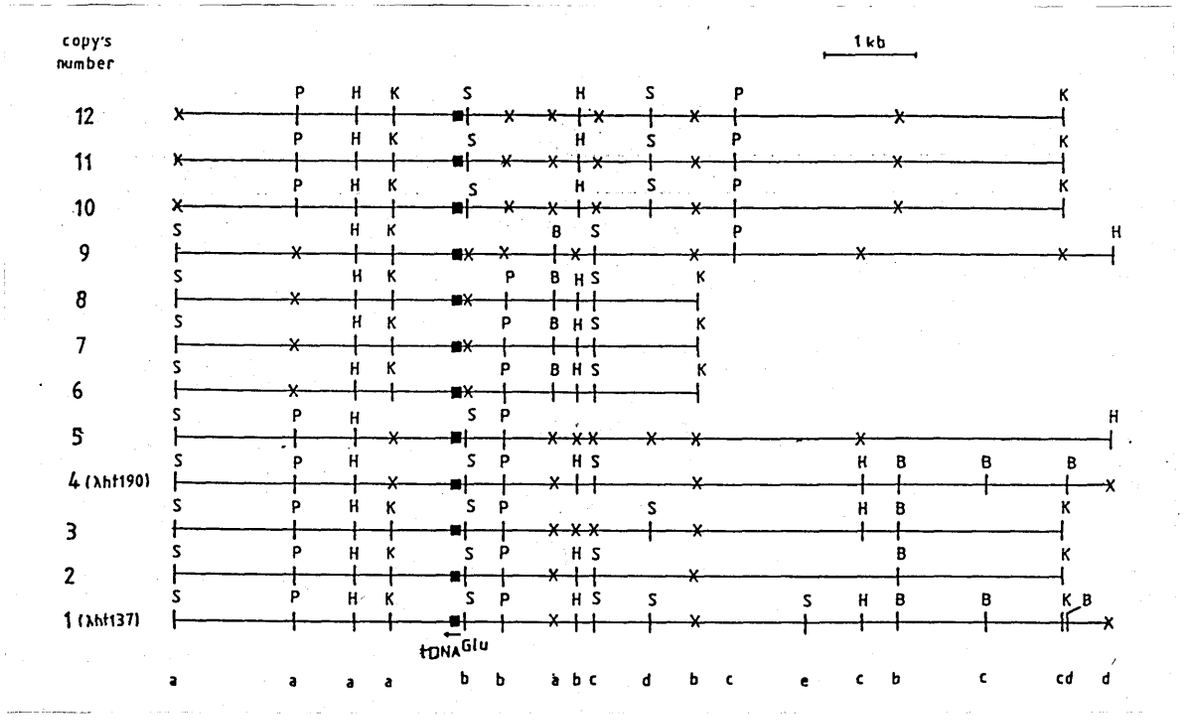
Preliminary analysis of the data from the organization experiments (see section 4.2.2.) shows that 9 copies of the gene for tRNA^{Glu} belong to a highly conserved 2.4 kb HindIII fragment repeat in the human genome. In addition restriction sites, such as KpnI, PvuII or SstI found to occur within the characterized 2.4 kb HindIII fragments from recombinants λht137 and λht190, also appeared to occur in the same position in most of the remaining copies. For example the tRNA^{Glu} gene copy found in λht137 belongs to a 1.2 kb KpnI-PvuII fragment. The genomic DNA hybridization data show that 6 copies belong to 1.2 kb KpnI-PvuII fragment repeats in the human genome (see sections 4.1.3.1. and 4.2.2., figure 4.2. and

tables 4.4. and 4.5.). The data obtained by using probe "3" (which contains 301 nucleotides of the 5'-flanking sequence but not a $tRNA^{Glu}$ gene; see section 3.4.4.) also show that 9-10 copies do share highly homologous 5'-flanking sequences, at least as far as 350 bp upstream of the gene (see tables 4.4. and 4.5.). These preliminary observations encouraged me to attempt to construct the most likely disposition of restriction sites occurring within the remaining unmapped $tRNA^{Glu}$ gene copies. Starting with the λht137 and λht190 restriction maps (copies 1 and 4 of figure 5.2.) it was clear that seven other copies (number 2, 6-8, 10-12 in figure 5.2.) share the HindIII sites (a) and (b) and that the single copy 5.5 kb HindIII fragment would arise by loss of HindIII site (b) (copy 3, figure 5.2.). Such a positively identified "lost" site is represented by X in figure 5.2. Only those sites which may be positively inferred from the data, are shown so that the omission of a site from any of the copy maps (e.g. SstI site (d) in copies 2, 4, 6-9) does not imply its absence unless so denoted by X. Within these limitations, the results shown are consistent with the data obtained and permit some conclusions on the nature of the $tRNA^{Glu}$ gene copies.

First, most of the copies (9 or 10) have a similar flanking sequence at least as far as 350 bp upstream of the gene, as noted earlier by the very similar results obtained for the distribution of the gene copies when probe "3" (containing only the 5'-flanking sequence -55 to -355) was used instead of probe "1" (gene and 5'-flanking sequence -1 to -74) or probe "2" (gene only). Second, inspection of figure 5.2. suggests that the copies share common restriction sites with each other over a much larger region, which is consistent with high sequence homology between copies in that region. If each copy shares 97% sequence homology with any other copy, as found for the 1.1 kb of sequence in and around the gene copies found in λht137 and λht190 (see section 5.1.2.), then one would expect each copy to share an average 4 of 5 common 6 bp recognition sites, or 80% of the copies to share a common restriction site. A situation approximating to this is found for copies 1-12 in figure 5.2. The occurence of SstI (a), PvuII (a), HindIII (a) and KpnI (a) in the 3 kb region downstream of the gene copies, varies from 9/12 to 12/12 and totals 39/48 or 81%. Similarly the occurence of SstI (b), PvuII (b), BamHI (a), HindIII (b) and SstI (c) in the 1.5 kb region upstream of the gene copies was found to be 60%. The remaining 5 kb upstream of the gene region is less complete but again shows many conserved sites between copies. Third, the

Figure 5.2.

Potential restriction maps of human genomic DNA fragments containing $tRNA^{Glu}$ genes.



Potential restriction maps of 12 human genomic DNA fragments containing $tRNA^{Glu}$ genes are shown. The disposition of the restriction sites present (H=HindIII; S= SstI; K= KpnI; P= PvuII, B=BamHI) is a logical explanation of the results of the hybridization experiments described in section 4.2.2. Positively identified "lost" sites are represented by X. The position and orientation of $tRNA^{Glu}$ genes, which have been numbered from 1 to 12 are indicated. Multiple sites identified for the same restriction enzyme are marked by letters.

conservation of sites in this 8-10 kb region of the copies varies in a way which may allow them to be divided into three inter-related subgroups each of 3-4 copies. For example copies 6-8 closely resemble one another but differ from the closely related copies 10-12, since they have PvuII site (b) (found also in copies 1-5) and KpnI site (b) (peculiar to copies 6-8).

However in several instances uncertainties remain, particularly for regions distant from the gene and the probe sequences used. The most notable uncertainties in the *potential* restriction maps of figure 5.2. are : i) the inconsistent data obtained from probes "1" and "3" with regard to the number of copies belonging to 4.55 kb SstI and 3.3 kb KpnI fragments (see table 4.4.); ii) the lack of additional data of HindIII digests (i.e. double digests) which might confirm the HindIII sites (a), (b) and (d); iii) the fact that 3-4 additional DNA fragments hybridized probe "3" but did not hybridize the probes containing $tRNA^{Glu}$ gene sequences. It is still unclear whether these fragments contain $tRNA^{Glu}$ gene sequences or not.

The inter-relatedness of the apparent subgroups earlier mentioned, complicates any interpretation of how these $tRNA^{Glu}$ gene copies evolved. A series of duplications of an ancestor sequence (at least 10 kb long), followed by insertions, deletions, or bases substitutions seems likely to have happened during evolution of the human genome. The latter hypothesis is supported by several other studies, where tRNA gene copies have been found to share homologous flanking sequences. Such examples have been reported : in human (Santos and Zasloff, 1981; Ma *et al.*, 1984; McPherson and Roy, 1986; Goddard *et al.*, unpublished results), in rat (Rosen *et al.*, 1984), as well as in other eukaryotes (Hershey and Davidson, 1980; Hosbach *et al.*, 1980; Sharp *et al.*, 1981; Fostel *et al.*, 1984; Amstutz *et al.*, 1985; Cribbs *et al.*, 1987). However there are other reports where tRNA genes have different flanking sequences and therefore gene duplication is unlikely to have happened (Sekiya *et al.*, 1982; Arnold *et al.*, 1986; Chang *et al.*, 1986; Gouilloud and Clarkson, 1986; Doran *et al.*, 1987 and 1988; Hauber *et al.*, 1988; Suter and Kubli, 1988). Weiner and Denison (1983), in order to explain the evolution of a multigene family encoding human U1 small nuclear RNA, have proposed that gene conversion is the only mechanism of homogenization which would result in conserving the sequence of the gene, while not necessarily conserving the flanking sequences. However if this statement is correct, it is not clear why conversion would not extend into flanking sequences,

especially with regard to 3'-flanking sequences since these contain the functional sequences providing signals for transcription termination. Nevertheless the conservation or not of a particular flanking sequence may arise from mechanisms of transcriptional regulation. Several groups have demonstrated that tRNA genes have tissue specificity due to different flanking sequences (see section 1.5.7.). However in the case of the human $tRNA^{Glu}$ gene family, where several kb sequences upstream and downstream of the gene copies have been conserved, it is still to be answered whether sequences far upstream of the gene modulate the expression of the gene copies or even if other important elements (like transposable elements in yeast; see section 1.4.1.) have also been conserved close to tRNA genes during evolution of the human genome. One approach to further investigation of the organization of human $tRNA^{Glu}$ genes would be to perform additional restriction mapping analysis on genomic DNA using probes containing sequences more distant (up to several kb upstream or downstream) from the $tRNA^{Glu}$ gene region (e.g. radiolabelled fragments of recombinant λht137).

5.3. *In vitro* expression of human $tRNA^{Glu}$ genes.

5.3.1. The properties of *in vitro* transcription of MtGlu6.

The *in vitro* transcription analysis of the two characterized human $tRNA^{Glu}$ gene copies (tested as pLB4 and pTC51), showed that both are transcribed with a similar efficiency (see section 4.3.2.). As a result only one of them (from recombinant λht137) was further used to study the characteristics of its transcription.

Comparative studies of this gene copy (tested as MtGlu6) and an active *X. Laevis* amber suppressor tRNA gene (M13tT-(su^+); see section 4.3.4.5.) showed that MtGlu6 is 1.5-2.0 times more efficient as a template at 120 ng/20μl assay than M13tT(su^+)(see figure 4.23B.). MtGlu6 was also able to compete with M13tT-(su^+) for transcription factors and RNA polymerase III when the two recombinants were assayed together (see section 4.3.4.5.). In DNA saturated

assays, when increasing amounts of MtGlu6 were added, fewer transcripts were produced from the same amount of M13tT-(su^+) DNA. 80 ng of MtGlu6 produced approximately 1.3 times more transcripts than 120 ng of M13tT-(su^+) and 120 ng of MtGlu6 produced approximately 5.7 times more transcripts than 80 ng of M13tT-(su^+). These results which show that MtGlu6 is an efficient RNA polymerase III template, are in accordance with what was found after micro-injection of the same gene (as pLB4) into *Xenopus* oocyte nuclei where the requirement for both the 3' ICR and the 5' ICR was also demonstrated (Goddard *et al.*, 1983; see also section 2.2.).

MtGlu6 transcription products were identified in the presence of α-amanitin (see section 4.3.3.3.). In addition the size of MtGlu6 transcripts was estimated relative to single stranded DNA size markers (data not shown) and to stained crude tRNA on polyacrylamide gels, as 75-80 nucleotides. It was concluded that the extract used in the experiments described in section 4.3., produced only precursor size MtGlu6 transcripts, since with some other HeLa extract preparations, larger precursors as well as smaller size transcripts were observed.

The transcriptional efficiency of MtGlu6 was found to be substantially dependent on the supercoil form of its DNA. The linear recombinant, when tested, had lost two thirds of its transcriptional activity. These results were also found with all the "MtGlu6 deletion mutants" which were tested (see section 4.3.4.3.). This is consistent with the findings of Lamond (1985), that transcription *in vitro* of an *E. coli* $tRNA^{Tyr}$ gene was dependent on the extent of negative supercoiling of its DNA. More recently Shapiro *et al.*, (1988) have shown a similar 2 to 3-fold reduction in the transcription efficiency of covalently closed circular DNA containing the adenovirus major late promoter, when the template is linearized.

Finally the salt optimization experiments show that the transcription efficiency of MtGlu6 is also dependent on the ionic conditions of the *in vitro* transcription system. All salts which were tested (KCl, $MgCl_2$, NaCl) appear to affect the MtGlu6 transcription efficiency. The observed critical dependence on $MgCl_2$ concentration (a 2 mM $MgCl_2$ decrease resulted in a 5-fold decrease of MtGlu6 efficiency) has also been reported by Weil *et al.*, (1979), after testing a similarly prepared KB extract with VAI or 5S RNA genes.

In the light of these initial remarks, the characteristics of the *in vitro*

transcriptional analysis of MtGlu6 can be summarized as follows :

The DNA concentration assays showed that tRNA size transcription products, which were detected even with 5 ng of DNA (0.25 µg/ml), increased with increasing amounts (from 20 ng (1 µg/ml) to 90 ng (4.5 µg/ml)) of added MtGlu6. No further increase was observed with MtGlu6 concentrations from 5µg/ml up to 10µg/ml. Higher concentrations than 10 µg/ml were not tested, although earlier observations had shown a considerable decrease of transcription after using extremely high DNA concentrations (J.P. Goddard, personal communication). At low concentrations of MtGlu6, assays where no DNA "carrier" was added, produced considerably fewer transcripts than those with DNA "carrier" present (see section 4.3.3.1.). The time course assays showed that both MtGlu6 concentrations which were tested (2 µg/ml and 7.5 µg/ml) produced detectable transcripts, during a 20 min reaction time, and that the transcription products increased linearly with time from 35 min to 120 min (see section 4.3.3.2.). These aspects are now discussed more fully :

There are plenty of transcription factors and RNA polymerase III available in the HeLa S100 extract at the lower DNA concentrations used to process transcription. In addition at low DNA concentrations transcription of the $tRNA^{Glu}$ gene was stimulated as much as 5-fold by addition of "carrier" DNA. The reason why the DNA carrier stimulates the transcribed DNA is not clear; however it seems likely that it is due to sequence independent DNA binding proteins which can inhibit transcription by randomly binding to DNA, masking the promoter to prevent factor binding (Fowlkes and Shenk, 1980; Sharp *et al.*, 1983a; Shapiro *et al.*, 1988). Since each assay in this study contained the same amount of extract and therefore the same amount of sequence independent DNA binding proteins, "carrier" DNA was added to give constant total DNA in the assay (see section 4.3.1.).

The amount of transcription products reaches a plateau at higher concentrations of MtGlu6 (5 µg/ml). At this stage presumably, there are no more transcription factors and/or RNA polymerase III available in order to form additional initiation complexes. The rate of tRNA products is 61±3 *f*mol/hr with 150 ng MtGlu6 per 20 µl assay. This corresponds to ~2 transcript/gene/hr (see section 4.3.3.1.). The calculated number of transcripts per gene per hour obtained at lower MtGlu6 concentration (e.g. at "Vmax"/2) is higher

(approximately 5) but still well below the expected *in vivo* rates. However this is not a result of the particular RNA polymerase III template which was used, but due to deficiencies of the HeLa S100 extract. The extract used in the experiments reported here is comparable to those used by other groups. Sharp *et al.*, (1983a) showed that about 6 transcripts per gene per hour were obtained from a *Drosophila* $tRNA^{Arg}$ gene using *Drosophila* Kc or HeLa S100 extracts; similarly Weil *et al.*, (1979) found about 3 transcripts per gene per hour from a VAI RNA gene using a KB cell free extract. Recently Shapiro *et al.*, (1988) have reported a method which results in a 20-fold more efficient HeLa cell nuclear extract. This method was tried twice, but in my hands yielded an extract which was no more efficient than the extract used in the reported experiments.

In the time course assays the lag observed before appreciable amounts of tRNA were synthesized (see figure 4.17.) presumably reflects the time taken for the formation of active initiation complexes. Thereafter the number of full-length transcripts increases linearly with time for at least two hours. These observations might suggest the requirement of a releasing factor in the extract to allow the tRNA polymerase III to be released soon after the completion of one round of transcription. As a result, although each RNA polymerase III performs several rounds of transcription (since the transcripts per gene per hour at DNA concentrations below saturation are greater than 1) the whole "transcription apparatus" works considerably slower than expected.

From the above discussion it is clear that many parameters, apart from the nature of the 5'-flanking sequence, can affect the apparent transcriptional efficiency of the $tRNA^{Glu}$ gene. These parameters include the amount and the supercoiling of the DNA; the time and the temperature of the reaction; the nature of the extract and the ionic conditions. Therefore several precautions were taken in order to eliminate errors which could lead to a misinterpretation of the results of the transcriptional experiments. All the recombinants used were purified by CsCl centrifugations and their DNA was >90% supercoil when electrophoresed through ethidium bromide stained agarose gels; pipetting errors were eliminated by adding constant volumes of DNA, salts and ribonucleotides in every assay (see also section 3.5.12.2.), finally the same extract, time and temperature for the reaction were used in all experiments, except when otherwise stated.

5.3.2. Conclusions on the role of the 5'-flanking sequence of $tRNA^{Glu}$ gene transcription.

The *in vitro* transcriptional analysis of "MtGlu6 deletion mutants" shows, in summary, the following :

The DNA concentration assays (see section 4.3.4.2.) revealed that the "Vmax" of MtGluD20, MtGluD33, MtGluD44B, MtGluD47 and MtGluD104C is essentially the same as MtGlu6. (The ±5% divergence could be considered as experimental error rather than a real deviation.) Mutants MtGluD5 and MtGluD7 appear to have considerably lower "Vmax" (60% and 74% of MtGlu6 maximum efficiency respectively) and mutant MtGluD92 is totally inactive. In addition all active mutants have a similar "Km" to MtGlu6, with the exception of MtGluD47 which has a lower "Km" (see figure 4.21. and table 4.6.).

At this point it is appropriate to discuss how the differences in the number of transcripts produced from varying amounts of MtGlu6 or of its deletion mutants are to be interpreted. As previously stated, use of the terms "Vmax" and "Km" should not be taken to imply the applicability of simple Michaelis-Menten kinetics to the complex mechanism of eukaryotic tRNA gene transcription. The mechanism of this process is still not completely understood (see section 1.5.4.), but involves the sequential and probably cooperative binding of factors to the gene promoter and its surrounding region to permit binding of RNA polymerase III to the active initiation complex and transcription. By analogy with 5S RNA gene transcription, one would expect the transcription to occur without complete disruption of the factors from the DNA, so that several rounds of transcription will occur on each active complex if RNA polymerase III is released on termination of transcription. Thus it is clear that the freely dissociated enzyme-substrate complex required for Michaelis-Menten kinetics is not present. Secondly the rate of "product formation" (transcripts produced) is not constant, since the time course experiments (see figure 4.17.) show an initial lag, probably representing the formation of the initiation complexes.

In these circumstances "Vmax" is simply considered as the extent of transcription which is achieved under defined conditions, when a given DNA template is in excess, so that transcription factor(s) and/or RNA polymerase III

are limiting. As such its value will depend upon the efficiency of the template and on the nature of the extract. Different extracts assayed under the standard conditions of saturating MtGlu6 produced fewer transcripts, presumably because the limiting factors/RNA polymerase III were at low concentration.

The term "Km" in the DNA concentration assays is the MtGlu6 or mutant DNA concentration which produced half the rate of transcription as that produced under the same conditions when DNA was in excess. As such a significant difference in "Km" of two templates would reflect a difference in their transcriptional efficiencies, which would not be discerned under conditions of excess DNA. Such a difference is found between MtGlu6 and MtGluD47, where "Km" of MtGluD47 is about half of MtGlu6 (see table 4.6.) suggesting that the mutant is a more efficient template. Thus in order to draw any conclusions about the efficiency of "MtGlu6 deletion mutants", the amount of transcripts obtained from each of them at DNA concentrations below saturation should also be compared. At DNA concentrations of 40 ng/ 20 µl assay the transcriptional efficiency of MtGlu6, MtGluD5, MtGluD7 and MtGluD47 relative to the efficiency at saturating MtGlu6 were found at : 60%, 35%, 48% and 82% respectively (see figure 4.20.A. and 4.20.B.). This comparison strongly implies that MtGluD47 is a more efficient template than MtGlu6, in accordance with the "Km" values obtained (see table 4.6. and figure 4.21.), and that MtGluD5 and MtGluD7 are less efficient templates than MtGlu6 as indicated by "Vmax". Inspection of figure 4.20.A. and 4.20.B. shows that the order of transcriptional efficiency MtGluD47 > MtGlu6 > MtGluD7 > MtGluD5, inferred above from the data at 40 ng template DNA per assay, is found at all non-saturating DNA concentrations. The amount of transcripts produced by MtGluD20, MtGluD33, MtGluD44B and MtGluD104C is very similar to that of MtGlu6 at all DNA concentrations, indicating that within the error limits of the experiment (see section 4.3.4.2.) the transcriptional efficiency of these mutants is the same as that of MtGlu6.

The changes in transcriptional efficiency of MtGluD5, MtGluD7, MtGluD47 relatively to MtGlu6 presumably reflect the efficiency of active complex formation and transcription initiation. The sequential assays and the competition assays reported in section 4.3.4.4. and 4.3.4.5. respectively were attempts at checking the conclusions made from the DNA concentration assays. Although a complete interpretation of these experiments is difficult, they indicate that

MtGluD47 is a more efficient template and stronger competitor than MtGlu6, while MtGluD5 and MtGluD7 are less efficient templates and weaker competitors. Interestingly though, when 120 ng MtGlu6 were assayed together with 80 ng M13tT-(su^+), it resulted in the same decrease of M13tT-(su^+) transcription products as MtGluD47. MtGluD47 appears to be a stronger competitor at lower DNA concentrations. (For example when 40 ng of MtGluD47 were assayed together with 120 ng M13tT-(su^+), M13tT-(su^+) efficiency was reduced from 33% to 17%, but when the same amount of MtGlu6 was assayed together with 120 ng M13tT-(su^+), M13tT-(su^+) efficiency was reduced to 24%; see figure 4.23.A.).

All these experiments show that the tRNA-like structure is transcriptionally inactive and all the transcriptional activity of recombinant MtGlu6 is derived from the $tRNA^{Glu}$ gene. However regions of the 5'-flanking sequence of the gene could modulate its transcriptional efficiency. These results appear to be consistent within the homologous *in vitro* transcription system which was used.

Prior to this study, preliminary experiments using a heterologous transcription system, in which the mutants MtGluD5 and MtGluD7 were tested in parallel with MtGlu6 (Goddard *et al.*, 1985), produced results which differ from these reported here using HeLa S100 extract. The amount of tRNA products produced by MtGluD5 or MtGluD7, after micro-injection into *Xenopus* oocyte nuclei, (42% or 19% of MtGlu6 products respectively) was considerably lower than that obtained by using HeLa S100 extract (60% or 74% respectively). However these data are limited since only one concentration from each recombinant was micro-injected (10 ng per oocyte). Nevertheless, there are many occasions where contradictory results have been obtained by using two different transcription systems. Raymond and his collegues have reported that parts of the 5'-flanking sequence of a yeast $tRNA_3^{Leu}$ gene modulate differently the transcription of the gene in different transcription systems.With replacement of the 5'-flanking sequence, the gene remains active in *Xenopus* or HeLa extracts but is nearly inert in yeast extracts (Raymond and Johnson, 1983; Johnson *et al.*,1984b). In addition Schaack and Soll (1985) have shown that the transcription kinetics of a *Drosophila* $tRNA^{Arg}$ gene are different in a *Drosophila* Kc extract from a *S. cerevisiae* extract. As a result it has been proposed that 5'-flanking regions of RNA polymerase III genes could be weak promoters for one

transcription system (Ciliberto *et al.*, 1983a); or that the requirement of *in vitro* transcription for flanking sequences is dependent on the cellular source of the extract (Schaack *et al.*, 1984).

The limited data from micro-injection of $tRNA^{Glu}$ genes into *Xenopus* oocyte nuclei supported the simple hypothesis that the tRNA-like structure considerably stimulates the transcriptional efficiency of the gene itself. The more extensive and detailed study reported here is inconsistent with such a simple hypothesis. While MtGluD5 and MtGluD7 (which lack part of the tRNA-like structure) have reduced transcriptional activity in both systems, other deletion mutants where the 5'-flanking sequence is removed, have the same transcriptional activity as MtGlu6. In the HeLa S100 *in vitro* transcription system a 5'-flanking sequence even as close as 12 nucleotides from the $tRNA^{Glu}$ gene, can be deleted without affecting the transcriptional efficiency of the gene itself. As expected when the first 11 nucleotides of the $tRNA^{Glu}$ gene, including 3 from the A block, were removed the corresponding recombinant (MtGluD92) showed no transcriptional activity, although it was able to compete weakly with M13tT-(su^+) (see section 4.3.4.5.). The latter implies that TFIIIC can still bind to the B block of the mutated $tRNA^{Glu}$ gene even in the absence of the first 11 nucleotides of the gene itself.

The 5'-flanking sequence of the gene, does not contain, but for one exception, any sequence which has been reported to modulate positively or negatively transcription of eukaryotic genes (see section 1.5.7.). Sajjadi and Spiegelman (1987a and 1987b) have proposed that a general form of sequence TNNCT (TCGCT especially) is a positive transcription modulator for a class of *Drosophila* tRNA genes. Although a similar study has not been undertaken with higher eukaryotic tRNA genes , the human $tRNA^{Glu}$ gene contains 7 such sequences (of which 4 are TCGCT) in its immediately preceding 250 nucleotides. (The statistical probability for a TCGCT sequence is once in every 1024 nucleotides). However no obvious outcome can arise from the location of these sequences relative to the transcriptional efficiency of the $tRNA^{Glu}$ gene. For instance although deletion mutants, MtGluD47 and MtGluD104C, both lack the same three such sequences (residues 175-179, 198-202 and 312-316), MtGluD104C transcribes similarly to MtGlu6 but MtGluD47 does not.

The effect of the distance of a possible modulator sequence, present at least

200 nucleotides upstream of the gene (i.e. upstream from the StuI site used in the construction of mutants with the largest deletions), upon its transcriptional activity, was also examined (see figure 5.3.A.). However the activities obtained from "MtGlu6 deletion mutants" which contain from 8 up to 192 nucleotides shorter 5'-flanking sequences of MtGlu6 (see figure 3.3.), do not lead to any obvious explanation. For example although MtGluD47 and MtGluD104C are lacking 167 and 168 nucleotides from the 5'-flanking sequence of the gene respectively, they transcribe differently.

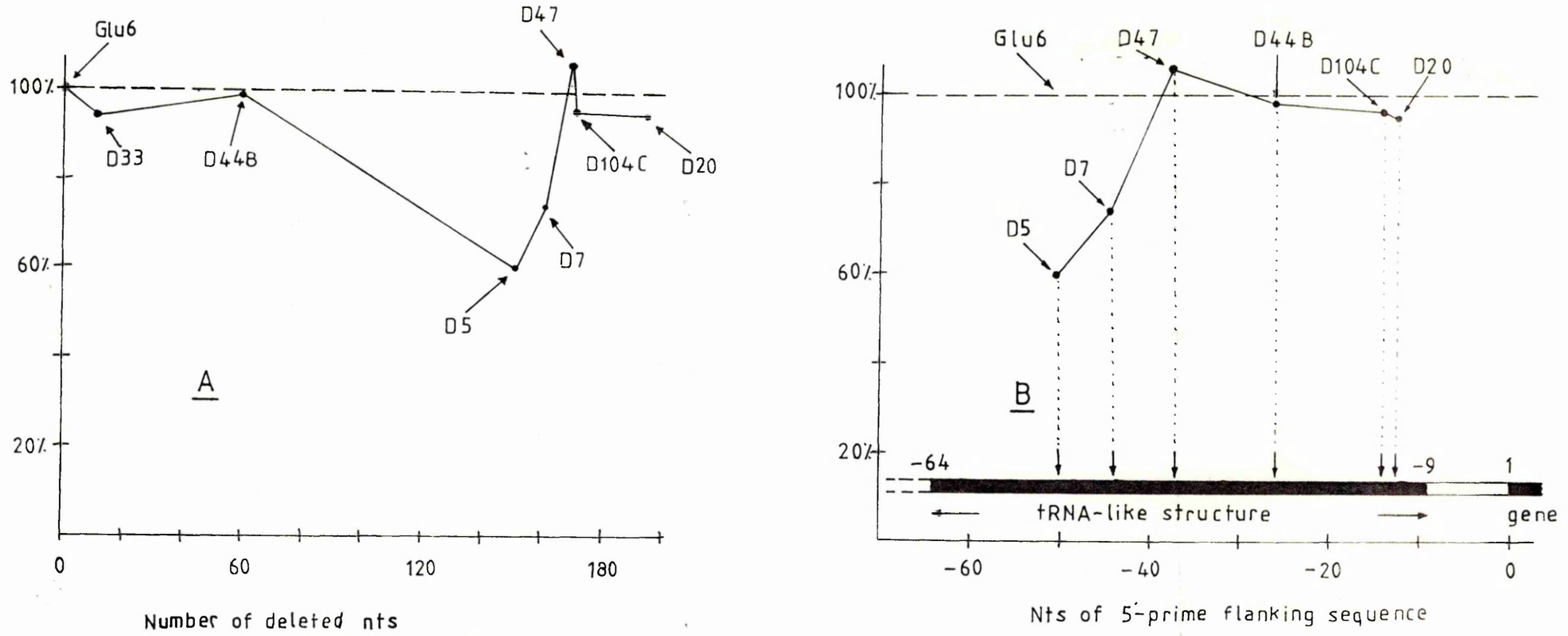
Figure 5.3.B. shows the effect of the presence of different length intact immediate 5'-prime flanking sequences upon the transcriptional efficiency of the $tRNA^{Glu}$ gene. The latter graph indicates that the deletion mutants MtGluD20, MtGluD104C, MtGluD44B and MtGluD47 which contain immediate 5'-flanking sequences of 12,13, 25 and 37 nucleotides respectively, have the same maximum efficiency as MtGlu6. However the deletion mutants MtGluD7 and MtGluD5 which contain immediate 5'-flanking sequences of 44 and 51 nucleotides respectively, have lower maximum efficiency (74% and 60% of MtGlu6 efficiency respectively).

In conclusion a possible interpretation of the role of the 5'-flanking sequence on the $tRNA^{Glu}$ gene transcription, could be the following :

A sequence located between residues 152-304 (of the sequence published in Goddard *et al.*, 1983; see also figure 3.3.) might be a positive modulator of the *in vitro* transcription of the $tRNA^{Glu}$ gene. When this sequence was removed (mutant MtGluD5; deleted residues 152-304 of the same sequence), the maximum transcription efficiency of the gene was reduced to 60%. Mutants MtGluD7 and MtGluD47 (deleted residues 152-311 and 152-318 respectively) have maximum efficiencies 74% and 105% of MtGlu6 respectively. If the prior statement is correct (a positive modulator is located between residues 152-304), then the residues 305-318 should contain a negative modulator. This last hypothesis is supported by the fact that the more of that sequence is removed, the higher the maximum transcription efficiency of the gene is. When both positive and negative modulator are removed (mutant MtGluD47) the new recombinant has substantially lower "Km" than MtGlu6, but similar "Vmax". Mutant MtGluD20 (deleted residues 152-343) transcribes similarly to MtGlu6. This means that residues 319-342 have no effect on the transcription of the gene.

Figure 5.3.

Analysis of "MtGlu6 deletion mutants" transcriptional efficiencies.



A. The transcriptional efficiencies of "MtGlu6 deletion mutants" are plotted versus their total number of deleted nucleotides. The broken line indicates the maximum efficiency of MtGlu6 (100%).

B. The transcriptional effficiencies of "MtGlu6 deletion mutants" are plotted versus the length (i.e. number of nucleotides) of the intact immediate 5'-prime flanking sequence of the $tRNA^{Glu}$ gene they contain. The location of the tRNA-like structure (residues -9 to -64 upstream the gene) is shown. The broken line indicates the maximum efficiency of MtGlu6 (100%).

The data obtained from the remaining deletion mutants (MtGluD33, MtGluD44B, MtGluD104C) do not eliminate the prior hypothesis. However, since these deletion mutants contain a different class of deletions (the deleted sequence does not start from position 152; MtGluD33 : deleted residues 296-303, MtGluD44B : deleted residues 273-330, MtGluD104C : deleted residues 175-342) it is preferable not to make any further statement about the precise location of the potential modulator sequences. The principal conclusions which can be drawn unequivocally are that, firstly, the deletion mutant MtGluD47 contains a sequence which maximizes the transcription efficiency of the $tRNA^{Glu}$ gene and, secondly that the deletion mutants MtGluD5 and MtGluD7 contain sequences which inhibit transcription.

There are two reports about eukaryotic tRNA genes which contain both positive and negative transcription modulators in their 5'-flanking sequences (Schaack *et al.*, 1984; Sajjadi *et al.*, 1987b; see also section 1.5.7.). However in the case of the human $tRNA^{Glu}$ gene, additional recombinants have to be constructed and tested in order to identify the sequences which modulate its transcriptional efficiency. One approach to the subject could be the construction of series of new deletion mutants all lacking 167 nucleotides (as in the case of MtGluD47), but with a variable distance between the deleted region of the 5'-flanking sequence and the gene. The transcriptional analysis of these deletion mutants might facilitate the location of any modulator sequence. In addition other deletion mutants could be constructed and further studied, lacking different regions of either positive (residues 152-304) or negative (residues 305-318) potential modulator sequences. Any outcome of such analysis could be further tested by competition transcription assays where recombinants containing different 5'-flanking sequences, but not the gene itself, could compete with MtGlu6. The latter experiments as well as footprint analysis of MtGlu6 DNA will show whether any transcription factor or inhibitor protein binds to the 5'-flanking sequence of the gene, or not.

REFERENCES

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

Adams, R.L.P. (1980) Cell Culture for Biochemists. Elsevier/N.Holland Biomedical Press.

Adeniyi-Jones, S., Romeo, P.H. and Zasloff, M. (1984) *Nucleic Acids Res.* **12**, 1101-1115.

Allison, D.S., Goh, S.H. and Hall, B.D. (1983) *Cell* **34**, 655-664.

Allison, D.S. and Hall, B.D. (1985) *EMBO J.* **4**, 2657-2664.

Amstutz, H., Munz, P., Heyer, W.D., Leupold, U. and Kohli, J. (1985) *Cell* **40**, 879-886.

Arnold, G.J., Schmutzler, C., Thomann, U., Toll, H.V. and Gross, H.J. (1986) *Gene* **44**, 287-297.

Arnold, G.J. and Gross, H.J. (1987) *Gene* **51**, 237-246.

Arnold, G.J., Schmutzler, C. and Gross, H.J. (1988) *DNA* **7**, 87-97.

Atkinson, T. and Smith, M. (1984) in Oligonucleotide Synthesis, a Practical Approach, pp. 35-81, Gait,M.J.(ed), I.R.L. Press.

Baker, R.E., Eigel, A., Vogtel, D. and Feldmann, H. (1982) *EMBO J.* **1**, 291-295.

Baker, R.E. and Hall, B.D. (1984) *EMBO J.* **3**, 2793-2800.

Baker, R.E., Gabrielsen, O. and Hall, B.D. (1986) *J. Biol. Chem.* **261**, 5275-5282.

Baldi, M.I., Mattocia, E. and Tocchini-Valentini, G.P. (1983) *Cell* **35**, 109-115.

Baldi, M.I., Mattocia, E., Ciafre, S., Attardi, D.G. and Tocchini-Valentini, G.P. (1986) *Cell* **47**, 965-971.

Bansal, N.S. (1986) Ph.D. Thesis, Glasgow University.

Bieker, J.J., Martin, P.L. and Roeder, R.G. (1985) *Cell* **40**, 119-127.

Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513-1523.

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

Bogenhagen, D.F., Wormington, W.M. and Brown, D.D. (1982) *Cell* **28**, 413-421.

Boyd, E., Theriault, A., Goddard, J.P., Kalaitsidaki, M., Spathas, D.H. and Connor, J.M., (1989) *Hum. Genet.* ***81***, 153-156.

Brodeur, G.M., Sandameyer, S.B. and Olson, M.W. (1983) *Proc. Nat. Acad. Sci. U.S.A.* **80**, 3292-3296.

Brown, D.D. (1984) Cell **37**, 359-365.

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

Burke, D.J., Schaak, J., Sharp, S., Soll, D., (1980) *J. Biol. Chem.* **258**, 15224-15231.

Burke, D.J. and Soll, D. (1985) *J. Biol. Chem.* **260**, 816-823.

Camier, S., Gabrielsen, O., Baker, R. and Sentenac, A. (1985) *EMBO J.* **4**, 491-500.

Carey, M.F., Gerrard, S.P. and Cozzarelli, N.R. (1986) *J. Biol. Chem.* **261**, 4309-4317.

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

Castagnoli, L., Ciliberto, G. and Cortese, R. (1982) *Nucleic Acids Res.* **10**, 4135-4145.

Chang, Y.N., Pirtle, I.L. and Pirtle, R.M. (1986) *Gene* **48**, 165-174.

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

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

Ciliberto, G., Castagnoli, L. and Cortese, R. (1983a) *Curr. Topics Dev. Biol.* **18**, 59-87.

Ciliberto, G., Costanzo, R.F., Dente, L. and Cortese, R. (1983b) *Cell* **32**, 725-733.

Ciliberto, G., Traboni, C. and Cortese, R. (1984) *Nucleic Acids Res.* **12**, 1277-1285.

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

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

Clarkson, S.G. (1983) in Eukaryotic genes: their structure, activity and regulation, pp 239-261, Maclean,N., Gregory, S.P. and Flavell R.A. (ed), Butterworths Press, London.

Cooley, L., Schaack, J., Burke, D.J., Thomas, B. and Soll, D. (1984) *Mol. Cell. Biol.* **4**, 2714-2722.

Corlet, J., Clarkson, S.G., Fournier, A. and Guerin, M.A. (1986) *Nucleic Acids Res.* **14**, 1916.

Cribbs, D.L., Leung, J., Newton, C.H., Hayashi, S., Miller, R.C. and Tener, G.M. (1987) *J. Mol. Biol.* **197**, 397-404.

Dean, N. and Berk, A.J. (1987) *Nucleic Acids Res.* **15**, 9895-9907.

De Duve, C. (1988) *Nature* **333**, 117-118.

De Franco, R., Schmidt, O. and Soll, D. (1980) *Proc. Nat. Acad. Sci. U.S.A.* **77**, 3365-3368.

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

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

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

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

Deutscher, M.P. (1982) *The Enzymes* **15**, 183-232.

Deutscher, M.P. (1984) *Crit. Rev. Biochem.* **17**, 45-71.

Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475-1489.

Dingerman, T., Sharp, S., Schaack, J. and Soll, D. (1983) *J. Biol. Chem.* **258**, 10395-10402.

Doering, J.L., Jelachich, M.L. and Hanlin, K.M. (1982) *FEBS Letters* **146**, 47-51.

Doran, J.L., Wei, X. and Roy, K.L. (1987) *Gene* **56**, 231-243.

Doran, J.L., Bingle, W.H. and Roy, K.L. (1988) *Gene* **65**, 329-336.

Drabkin, H.J. and RajBhandary, U.L. (1985a) *J. Biol. Chem.* **260**, 5580-5587.

Drabkin, H.J. and RajBhandary, U.L. (1985b) *J. Biol. Chem.* **260**, 5588-5595.

Drabkin, H.J. and RajBhandary, U.L. (1985c) *J. Biol. Chem.* **260**, 5596-5602.

Eigel, A., Olah, J. and Feldman, H. (1981) *Nucleic Acids Res.* **9**, 2961-2970.

Eigel, A. and Feldman, H. (1982) *EMBO J.* **10**, 1245-1250.

Elder, R., Szabo, A. and Ohlenbeck, O. (1980) in Transfer RNA : Biological Aspects. pp.317-323, Soll, D. Abelson, J.N. and Scimmel 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.

Engelke, D.R., Hoener, P.A. and Collins, F.S. (1988) *Proc. Nat. Acad. Sci. U.S.A.* **85**, 544-548.

Feinberg, A.R. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13.

Feinberg, A.R. and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266-267.

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

Feldmann, H., Olah, J. and Friedenreich, H. (1981) *Nucleic Acids Res.* **9**, 2949-2959.

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

Folk, W.R. (1988) *Genes and Development* **2**, 373-375.

Fostel, J., Narayanswani, S., Hankalo, B., Clarkson, S.G. and Pardlk, M.L. (1984) *Chromosoma* **90**, 254-260.

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

Fuhram, S.A., Engelke, D.R. and Geiduschek, E.P. (1984) *J. Biol. Chem.* **259**, 1934-1943.

Gabrielsen, O.S. and Oyen, T.B. (1987) *Nucleic Acids Res.* **15**, 5699-5713.

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

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

Gamulin, V., Mao, J., Appel, B., Smith, M.S., Yamao, F. and Soll, D. (1983) *Nucleic Acids Res.* **11**, 8537-8546.

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

Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) *Annual Rev. Biochem.* **57**, 873-914.

Glew, L., Lo, R., Reece, T., Nichols, M., Soll, D. and Bell, J. (1986) *Gene* **44**, 307-314.

Goddard, J.P., Squire, M., Bienz, M. and Smith, J.D. (1983) *Nucleic Acids Res.* **11**, 2551-2561.

Goddard, J.P., Mc Laren, A., Carr, T., Bienz, M. and Smith, J.D. (1985) *Biochem. Soc. Transactions* **13**, 754.

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

Gottesfeld, J. and Bloomer, L.S. (1982) *Cell* **28**, 781-791.

Gouilloud, E. and Clarkson, S.G. (1986) *J. Biol. Chem.* **261**, 486-494.

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

Gupta, C., Roe, B.A. and Randerath, K. (1979) *Nucleic Acids Res.* **7**, 959-970.

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

Han, J.H. and Harding, J.D. (1982) *Nucleic Acids Res.* **10**, 4891-4890.

Han, J.H., Rooney, R.J. and Harding, J.D. (1984) *Gene* **28**, 249-255.

Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557-580.

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

Hauber, J., Stucka, R., Krieg, R. and Feldmann, H. (1988) *Nucleic Acids Res.* **16**, 10623-10634.

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

Hipskind, R.A. and Clarkson, S.G. (1983) *Cell* **34**, 881-890.

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

Hodgkin, J., Kondo, K. and Waterson, R.H. (1987) *Trends in Genet.* **3**, 325-329.

Hoe, K.L., Hong, H.J., Yoo, S.H. and Yoo, O.J. (1987) *Nucleic Acids Res.* **15**, 10045.

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

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

Hong, H.J., Yoo, S.H. and Yoo, O.J. (1987) *Nucleic Acids Res.* **15**, 4987.

Horvath, D. and Spiegelman, G.B. (1988) *Nucleic Acids Res.* **16**, 2585-2599.

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

Hottinger-Werlen, A., Schaak, J., Lapointe, J., Mao, J., Nichols, M. and Soll, D. (1985) *Nucleic Acids Res.* **13**, 8739-8747.

Hou, Y.M. and Schimmel, P. (1988) *Nature* **333**, 140-145.

Hovemann, B., Sherp, S., Yamada, H. and Soll, D. (1980) *Cell* **19**, 889-895.

Hu, J.C., Lund, E. and Dahlberg, J.E. (1983) *Nucleic Acids Res.* **11**, 4809-4821.

Hu, N. and Messing, J. (1982) *Gene* **17**, 271-277.

Huibregtse, J.M., Evans, C.F. and Engelke, D.R. (1987) *Mol. Cell. Biol.* **7**, 3212-3220.

Indik, Z.K. and Tartof, K.D. (1982) *Nucleic Acids Res.* **10**, 4159-4172.

Jacob, S.T. (1973) *Prog. Nucleic Acids Res. Mol. Biol.* **13**, 93-99.

Johnson, J.D., Ogden, R., Johnson, P.F., Abelson, J., Dembeck, P. and Itakura, K. (1980) *Proc. Nat. Acad. Sci. U.S.A.* **77**, 2564-2568.

Johnson, J.D. and Raymond, G.J. (1984a) *J. Biol. Chem.* **259**, 5990-5994.

Johnson, J.D., Raymond, G.J. and De Parasis, J. (1984b) *Mol. Gen. Genet.* **197**, 55-61.

Johnson, P.F. and Abelson, J. (1983) *Nature* **302**, 681-687.

Kaiser, K. and Murray, N.E. (1985) in DNA cloning, a practical approach.,Vol. 1. pp1-47, IRL Press.

Kariya, Y., Kato, K., Hayashizaki, Y., Hineno, S., Tarui, S. and Matsubara, K. (1987) *Gene* **53**, 1-10.

Karn, J., Brenner, S., Barnett, L. and Cesareni, G. (1980) *Proc. Nat. Acad. Sci. U.S.A.* **77**, 5172-5176.

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.

Klemenz, R., Stillman, D.J. and Geiduschek, E.P. (1982) *Proc. Nat. Acad. Sci. U.S.A.* **79**, 6191-6195.

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

Kondo, K., Hodgkin, J. and Waterson, R.H. (1988) *Mol. Cell. Biol.* **8**, 3627-3635.

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

Koski, R.A., Allison, D.S., Worthington, M. and Hall, B.D. (1982a) *Nucleic Acids Res.* **10**, 8127-8143.

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

Kressmann, A., Hofstetter, H., Di Capua, E., Grosschedl, R. and Birnstiel, M.L. (1979) *Nucleic Acids Res.* **7**, 1749-1763.

Krupp, J.L., Shu, H.H. and Martin, N.C. (1988) *Nucleic Acids Res.* **16**, 770.

Kuchino, Y., Okada, N.S., Ando, N., Watanabe, S. and Nishimura, S. (1981) *J.Biol. Chem.* **256**, 9059-9062.

Lamond, A.I. (1985) *EMBO J.* **4**, 501-507.

Larson, D., Bradford-Wilcox, J., Young, L.S. and Sprague, K.U. (1983) *Proc. Nat. Acad. Sci. U.S.A.* **80**, 3416-3420.

Laski, F.A., Alzner-De Weerd, B., RajBhandary, U.L. and Sharp, P.A. (1982a) *Nucleic Acids Res.* **10**, 4609-4626.

Laski, F.A., Belagaje, R., RajBhandary, U.L. and Sharp, P.A. (1982b) *Proc. Nat. Acad. Sci., U.S.A.* **79**, 5813-5817.

Laski, F.A., Fire, A.Z., RajBhandary, U.L. and Sharp, P.A. (1983) *J. Biol. Chem.* **258**, 11974-11980.

Lassar, A.B., Martin, P.L. and Roeder, R.G. (1983) *Science* **222**, 740-748.

Lasser-Weis, M., Bawnick, N., Rosen, A., Sarid, S. and Daniel, V. (1981) *Nucleic Acids Res.* **9**, 5965-5978.

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

Lee, B.J., De La Rena, P., Tobian, J.A., Zasloff, M. and Hatfield,D. (1987) *Proc. Nat. Acad. Sci. U.S.A.* **84** , 6384-6388.

Lofquist, A. and Sharp, S. (1986) *J. Biol. Chem.* **261**, 14600-14606.

Lofquist, A., Garcia, A. and Sharp, S.J. (1988) *Mol. Cell. Biol.* **8**, 4441-4449.

Looney, J.E. and Harding, J.D. (1983) *Nucleic Acids Res.* **11**, 8761-8775.

Ma, D.P., Lund, E., Dahlberg, J.E. and Roe, B.A. (1984) *Gene* **28**, 257-262.

Makowski, D.R., Haas, R.A., Dolan, K.P. and Grunberger, D. (1983) *Nucleic Acids Res.* **11**, 8609-8624.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (eds) (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory. C.S.H. New York.

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

Marschalek, R. and Dingermann, T. (1988) *Nucleic Acids Res.* **16**, 6737-6752.

Marzluff, W.F., White, E.L., Benjamin, R. and Huang, R.C.C. (1975) *Biochemistry* **14**, 3715-3724.

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

Mazabraud, A., Scherly, D., Muller, F., Runger, D. and Clarkson, S.G. (1987) *J. Mol. Biol.* **195**, 835-845.

McBride, O.W., Rajagopalan, M. and Hatfield, D. (1987) *J. Biol. Chem.* **262**, 11163-11166.

McClain, W.H. and Foss, K. (1988) *J. Mol. Biol.* **202**, 697-709.

McKay, R.M., Horvath, D., Duncan, L. and Spiegelman, G.B. (1988) *Nucleic Acids Res.* **16**, 4137-4153.

McLaren, A. and Goddard, J.P. (1986a) *Nucleic Acids Res.* **14**, 5938.

McLaren, A. and Goddard, J.P. (1986b) *Biochem. Soc. Transactions* **14**, 1169-1170.

McPherson, J.M. and Roy, K.L. (1986) *Gene* **42**, 101-106.

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

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

Morry, M.J. and Harding, J.D. (1986) *Mol. Cell. Biol.* **6**, 105-115.

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

Muller, F., Clarkson, S.G. and Galas, D.J. (1987) *Nucleic Acids Res.* **15**, 7191.

Murphy, M.H. and Baralle, F.E. (1983) *Nucleic Acids Res.* **11**, 7695-7700.

Murray, N.E., Brammer, W.J. and Murray, K. (1977) *Mol. Gen. Genet.* **150**, 53-68.

Naylor, S.L., Sakagushi, A.Y., Shows, T.B., Grzeschik, K.H., Holmes, M. and

Zasloff, M. (1983) *Proc. Nat. Acad. Sci. U.S.A.* **80**, 5027-5031.

Newman, A.J., Ogden, R.C. and Abelson, J. (1983) *Cell* **35**, 117-125.

Nordheim, A. and Rich, A. (1983) *Nature* **303**, 674-679.

Normanly, J., Ogden, R.C., Horvath, S.J. and Abelson, J. (1986) *Nature* **321**, 213-219.

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

O'Neil, V.A., Eden, F.C., Pratt, K. and Hatfield, D.L. (1985) *J. Biol. Chem.* **260**, 2501-2508.

Ottonelo, S., Rivier, D.H., Doolittle, G.M., Young, L.S. and Sprague, K.U. (1987) *EMBO J.* **6**, 1921-1927.

Peterson, R.C. (1987) *Biochem. Biophys. Acta* **908**, 81-89.

Pieler, T., Appel, B., Oei, S.L., Mentzel, H. and Erdmann, V.A. (1985a) *EMBO J.* **4**, 1847-1853.

Pieler, T., Oei, S.L., Hamm, J., Engelke, U. and Erdmann, V.A. (1985b) *EMBO J.* **4**, 3751-3756.

Pieler, T., Hamm, J. and Roeder, R.G., (1987) *Cell* **48**, 91-100.

Pirtle, I.L., Shortridge, R.D. and Pirtle, R.M. (1986) *Gene* **43**, 155-167.

Raymond, G.J. and Johnson, J.D. (1983) *Nucleic Acids Res.* **11**, 5969-5988.

Raymond, G.J. and Johnson, J.D. (1987) *Nucleic Acids Res.* **15**, 9881-9894.

Raymond, K.C., Raymond, G.J. and Johnson, J.D. (1985) *EMBO J.* **4**, 2649-2656.

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

Reyes, V.M., Newman, A. and Abelson, J. (1986) *Mol. Cell. Biol.* **6**, 2436-2442.

Ricca, G.A., Taylor, J.M. and Kalinyak, J.E. (1982) *Proc. Nat. Acad. Sci. U.S.A.* **79**, 724-728.

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

Robertus, J.D., Ladner, J.E., Finch, J.T., Rhodes, 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.

Rooney, R.J. and Harding, J.D. (1986) *Nucleic Acids Res.* **14**, 4849-4864.

Rooney, R.J. and Harding, J.D. (1988) *Nucleic Acids Res.* **16**, 2509-2521.

Rosen, A., Sarid, S. and Daniel, V. (1984) *Nucleic Acids Res.* **12**, 4893-4906.

Rosen, A. and Daniel, V. (1988) *Gene* **69**, 275-285.

Roy, K.L., Cooke, H. and Buckland, R. (1982) *Nucleic Acids Res.* **10**, 7313-7322.

Ruet, A., Camier, S., Smagowiez, W., Sentenac, A. and Fromageot, P. (1984) *EMBO J.* **3**, 343-350.

Russo, T., Oliva, A., Duilio, A., Ammendola, R., Costanzo, F., Zannini, M. and Cimino, F. (1987a) *Bioch. Bioph. Res. Commun.* **149**, 1118-1124.

Russo, T., Duilio, A., Ammendola, R., Costanzo, F. and Cimino, F. (1987b) *Nucleic Acids Res.* **15**, 8562.

Sajjadi, F.G. and Spiegelman, G.B. (1987a) *Gene* **60**,13-19.

Sajjadi, F.G., Miller, R.C. and Spiegelman, G.B. (1987b) *Mol. Gen. Genet.* **206**, 279-284.

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

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

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

Sandmeyer, S.B., Bilanchone, V.W., Clark, D.J., Marcos, P., Carle, G.F. and Brodeur, G.M. (1988) *Nucleic Acids Res.* **16**, 1499-1515.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) *J.Mol. Biol.* **143**, 161-178.

Santoro, C., Costanzo, F. and Ciliberto, G. (1984) *EMBO J.* **3**, 1553-1559.

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

Schaack, J., Sharp, S., Dingermann, T. and Soll, D. (1983) *J. Biol. Chem.* **258**, 2447-2453.

Schaack, J., Sharp, S., Dingermann, T., Burke, D.J., Cooley, L. and Soll, D. (1984) *J. Biol. Chem.* **259**, 1461-1467.

Schaack, J. and Soll, D. (1985) *Nucleic Acids Res.* **13**, 2803-2814.

Schimmel, P.R., Soll, D. and Abelson, J.N. (eds)(1979) Transfer DNA : Structure, Properties and Recognition, Cold Spring Harbour Laboratory, New York.

Schlissel, M.S. and Brown, D.D. (1984) *Cell* **37**, 903-913.

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

Schulman, L.H. and Abelson, J. (1988) *Science* **240**, 1591-1592.

Schweizer, E., MacKennie, C. and Halvorsen, H.O. (1969) *J. Mol. Biol.* **40**, 261-277.

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

Sekiya, T., Kuchino, Y. and Nishimura, S. (1981) *Nucleic Acids Res.* **9**, 2239-2250.

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

Setzer, D.R. and Brown, D.D. (1985) *J. Biol. Chem.* **260**, 2483-2492.

Shapiro, D.J., Sharp, P.A., Whali, W.W. and Keller, M.J. (1988) *DNA* **7**, 47-55.

Sharp, S., De Franco, D., Dingermann, T., Farell, P. and Soll, D. (1981a) *Proc. Nat. Acad. Sci. U.S.A.* **78**, 6657-6661.

Sharp, S., De Franco, D., Silberklang, M., Hosbach, H.A., Schmidt, T., Kubli, E., Gergen, J.P., Wensink, P.C. and Soll, D. (1981b) *Nucleic Acids Res.* **9**, 5867-5882.

Sharp, S., Dingermann, T. and Soil, D. (1982) *Nucleic Acids Res.* **10**, 5393-5406.

Sharp, S., Dingermann, T., Schaack, J., De Franko, D. and Soll, D. (1983a) *J.Biol. Chem.* **258**, 2440-2446.

Sharp, S., Dingermann, T., Schack, J., Sharp, J.A., Burke, D.J., De Robertis, E.M. and Soll, D. (1983b) *Nucleic Acids Res.* **11**, 8677-8690.

Sharp, S., Schaack, J., Cooley, L., Burke, D.J. and Soll, D. (1985) *Crit. Rev. Biochem.* **19**, 107-144.

Shaw, K.J. and Olson, K.V. (1984) *Mol. Cell. Biol.* **4**, 657-665.

Shibuya, K., Noguchi, S., Nishimura, S. and Sekiya, T. (1982) *Nucleic Acids Res.* **10**, 4441-4448.

Shibuya, K., Nogushi, S., Yamaki, M., Nishimura, S. and Sekiya, T. (1985) *J.Biochem.* **97**, 1719-1725.

Shortridge, R.D., Pirtle, I.L. and Pirtle, R.M. (1985) *Gene* **33**, 269-277.

Smardo, F.L. and Calvet, J.P. (1987a) *Gene* **57**, 213-220.

Smardo, F.L. and Calvet, J.P. (1987b) *Nucleic Acids Res.* **15**, 661-681.

Sollner-Webb, B. (1988) *Cell* **52**, 153-154.

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., Hartmann, T., Meissner, F., Moll, J. and Vorderwulbecke, T. (1987) *Nucleic Acids Res.* **15**, r53-r188.

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

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

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

Stillman, D.J., Sivertsen, A.L., Zetner, P.G. and Geiduschek, E.P. (1984a) *J. Biol. Chem.* **259**, 7955-7962.

Stillman, D.J. and Geiduschek, E.P. (1984b) *EMBO J.* **3**, 847-853.

Stillman, D.J., Better, M. and Geiduschek, E.P. (1985) *J. Mol. Biol.* **185**, 451-455.

Straby, K.B. (1988) *Nucleic Acids Res.* **16**, 2841-2857.

Stucka, R., Hauber, J. and Feldmann, H. (1987) *Curr. Genet.* **12**, 323-328.

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

Sullivan, M.A. and Folk, W.R. (1987) *Nucleic Acids Res.* **15**, 2059-2068.

Suter, B. and Kubli, E. (1988) *Mol. Cel. Biol.* **8**, 3322-3331.

Swerdlow, H. and Guthrie, C. (1984) *J. Biol. Chem.* **259**, 5197-5207.

Traboni, C., Ciliberto, G. and Cortese, R. (1984) *Cell* **36**, 179-187.

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

Van Dyke, M.W. and Roeder, R.G. (1987) *Nucleic Acids Res.* **15**, 5031-5039.

Van Tol, H., Stange, N., Gross, H.J. and Beier, H. (1987) *EMBO J.* **6**, 35-42.

Van Tol, H. and Beier, H. (1988) *Nucleic Acids Res.* **16**, 1951-1966.

Vnencak-Jones, C.L., Wahab, S.Z., Zehener, Z.E. and Holmes, W.M. (1987) *Mol. Cell. Biol.* **7**, 4134-4138.

Wallace, R.B., Johnson, P.F., Tanaka, S., Schold, M., Itakura, K. and Abelson, J. (1980) *Science* **209**, 1396-1400.

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A. and Weiner, A.M. (eds)(1987) Molecular Biology of the Gene, The Benjamin Publishing Company.

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

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

Weiner, A.M. and Denison, R.A. (1983) *Cold Spring Harbour Symp. Quant. Biol.* **47**, 1141-1168.

Willis, J., Hottinger, H., Pearson, D., Chisholm, V., Leupold, U. and Soll, D. (1984) *EMBO J.* **3**, 1573-1580.

Wilson, E.T., Larson, D., Young, L.S. and Sprague, K.U. (1985) *J. Mol. Biol.* **183**, 153-163.

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

Wormington, W.M., Bogenhagen, D.F., Jordan, E. and Brown, D.D. (1981) *Cell* **24**, 809-817.

Yanish-Perron, C., Viera, J. and Messing, J. (1985) *Gene* **33**, 103-119.

Yarus, M. (1988) *Cell* **55**, 739-741.

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

Yoo, S.H. (1984) *Corean J. Genet.* **6**, 101-108.

Yoshinaga, S.K., Boulanger, P.A. and Berk, A.J. (1987) *Proc. Nat. Acad. Sci. U.S.A.* **84**, 3585-3589.

Young, L.S., Takahashi, N. and Sprague, K. U. (1986) *Proc. Nat. Acad. Sci. U.S.A.* **83**, 374-378.

Zasloff, M. and Santos, T. (1980) *Proc. Nat. Acad. Sci. U.S.A.* **77**, 5668-5672.

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