STUDIES ON PROTEIN-NUCLEIC ACID INTERACTIONS IN *Xenopus laevis* oocyte 5S RIBOSOMAL RNA GENE EXPRESSION

by

QIMIN YOU

BACHELOR OF MEDICINE, Jiamusi Medical College, 1982

A Dissertation Submitted in Partial Fulfullment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in the Department
of
Biochemistry and Microbiology

ACCEPTED
FACULTY OF GRADUATE STUDIES

DEAN

DATE May 3/93

We accept this thesis as conforming
to the required standard

Dr. Paul J. Romaniuk, Supervisor (Dept. of Biochemistry & Microbiology)

Dr. Alastair T. Matheson, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Santosh Misra, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Arthur R. Fontaine, Outside Member (Dept. of Biology)

Dr. Michael J. Ashwood-Smith, Outside Member (Dept. of Biology)

Dr. L. Brakier-Gingras, External Examiner (University of Montreal)

Supervisor: Dr. Paul J. Romaniuk

ABSTRACT

The experiments were focussed on three protein-nucleic acid interactions in the *Xenopus* oocyte: TFIIIA-5S RNA, TFIIIA-5S DNA and ribosomal protein L5-5S RNA. The binding affinities and contact sites of the proteins to the nucleic acids were studied.

For studying the TFIIIA-5S RNA interaction, block mutations were constructed in helical stems II, III, IV and V of *Xenopus laevis* oocyte 5S RNA. The affinities of these mutants for binding to transcription factor IIIA were determined using a nitrocellulose filter binding assay. Mutations in stems III and IV had little or no effect on the binding affinity of TFIIIA for 5S RNA. However, single mutants in stems II and V (positions 16-21, 57-62, 71-72, and 103-104) which disrupt the double helix, reduce the binding of TFIIIA by a factor of two- to three- fold. In contrast, double mutants (16-21/57-62, 71-72/103-104) which restore the helical structure of these stems, but with altered sequences, fully restore the TFIIIA binding affinity. The experiments reported here indicate that the double helical structures of stems II and V, but not the sequences, are required for optimal TFIIIA binding.

The effects on TFIIIA binding affinity of a series of substitution mutations in the *Xenopus laevis* oocyte 5S RNA gene were quantified. These data indicate that TFIIIA binds specifically to 5S DNA by forming sequence-specific contacts with three discrete sites located within the classical A and C boxes and the intermediate element of the internal control region. Substitution of the nucleotide sequence at any of the three sites significantly reduces TFIIIA binding affinity, with a 100-fold reduction observed for substitutions in the box C subregion. These results are consistent with a direct interaction of TFIIIA with specific base pairs within the major groove of the DNA. In contrast, the TFIIIA binding data for the same mutations expressed in 5S RNA indicates that the protein does not make

any strong sequence-specific contacts with the RNA. Although the protein footprinting sites on the 5S DNA and 5S RNA are coincident, nucleotide substitutions in 5S RNA which moderately reduce TFIIIA binding affinity do not correspond at all to the three specific TFIIIA interaction sites within the gene.

For investigating the L5-5S RNA interaction, a cDNA encoding ribosomal protein L5 of *Xenopus laevis* was subcloned into a T7 expression vector and expressed in *Escherichia coli*. The resulting soluble fusion protein with a histidine tag at the N-terminus was purified by affinity chromatography to 95% homogeneity. The equilibrium binding of recombinant L5 to *Xenopus* 5S ribosomal RNA was characterized, and the affinity of the protein for a set of 5S RNA mutants was quantitatively measured using a nitrocellulose filter binding assay. L5 binds to 5S RNA with properties similar to those of the TFIIIA-5S RNA interaction. However, unlike TFIIIA, L5 was insensitive to changes in either the sequence or the secondary structure of the 5S RNA.

The results from these studies indicate that the specific protein-nucleic acid interactions in the biological pathway of 5S RNA use distinct mechanisms.

Examiners:

Dr. Paul J. Romaniuk, Supervisor (Dept. of Biochemistry & Microbiology)

Dr. Alastair T. Matheson, Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Santosh Misra Departmental Member (Dept. of Biochemistry & Microbiology)

Dr. Arthur R. Fontaine, Outside Member (Dept. of Biology)

Dr. Michael J. Ashwood-Smith, Outside Member (Dept. of Biology)

Dr. L. Brakier-Gingras, External Examiner, (University of Montreal)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A: Adenine

ATP: Adenosine triphosphate

BSA: Bovine serum albumin

BPB: Bromophenol blue

C: Cytosine

cDNA: Complementary deoxyribonucleic acid

cpm: Counts per minute

CTP: Cytidine triphosphate

dATP: Deoxyadenosine triphosphate

dCTP: Deoxycytidine triphosphate

dGTP: Deoxyguanosine triphosphate

dTTP: Deoxythymidine triphosphate

ddATP: Dideoxyadenosine triphosphate

ddCTP: Dideoxycytidine triphosphate

ddGTP: Dideoxyguanosine triphosphate

ddTTP: Dideoxythymidine triphosphate

DNA: Deoxyribonucleic acid

DTT: Dithiothreitol

E. coli: *Escherichia coli*

EDTA: Ethylenediamine-tetraacetic acid

G: Guanine

GTP: Guanosine triphosphate

HPLC: High performance liquid chromatography

ICR: Internal control region

IE: Intermediate element

L5: Ribosomal protein L5

L18: Ribosomal protein L18

L25: Ribosomal protein L25

LB: Luria-Benton broth

mRNA: Messenger ribonucleic acid

MW: Molecular weight

N: Nucleotide

NMR: Nuclear magnetic resonance

nt.: Nucleotide

PAGE: Polyacrylamide gel electrophoresis

PEG: Polyethylene glycol

Pol. III: RNA polymerase III

PMSF: Phenylmethylsulfonyl fluoride

RNA: Ribonucleic acid

RNP: Ribonucleoprotein particle

rRNA: Ribosomal ribonucleic acid

RNase: Ribonuclease

RNasin: Ribonuclease inhibitor

S: Svedberg unit

SAAB: Selected amplification and binding

SDS: Sodium dodecyl sulphate

TBE: Tris; Borate; EDTA

TFIIIA: Transcription Factor IIIA

TFIIIB: Transcription Factor IIIB

TFIIIC: Transcription Factor IIIC

tRNA: Transfer ribonucleic acid

Tris-HCl: Tris (hydroxymethyl) aminomethane-HCl

U: Uracil

UTP: Uridine triphosphate

W.G.: Wheat germ

W.T.: Wild-type

XC: Xylene cyanol

Xlo: *Xenopus laevis* oocyte

Xls: *Xenopus laevis* somatic

Xlt: *Xenopus laevis* oocyte trace

ACKNOWLEDGEMENTS

I am grateful to Dr. Alastair T. Matheson, Dr. Santosh Misra, Dr. Arthur R. Fontaine and Dr. Michael J. Ashwood-Smith, the members of my supervisory committee, for their advice, expertise and positive input over this period of study.

I am indebted to my colleagues, Jon Faris, Nik Veldhoen, Colleen Nelson, Wei-Qing Zang, Judith Wise and Isabel Stevenson. They are my best friends I can never forget. I thank them for their help in routine work, special instructions, and the English language.

I acknowledge the enjoyment of my studies at the University of Victoria in the Department of Biochemistry and Microbiology. I thank the departmental staff (secretaries and technicians) for their asistance.

I am financially supported by a grant from NSERC.

The most special thanks of all goes to Dr. Paul J. Romaniuk, my supervisor and friend. Every academic progress I achieved during the last six years has been related to his guidance, enthusiasm, patience and (especially) understanding. Paul, I am grateful.

CHAPTER 1
GENERAL INTRODUCTION

The oocyte of the South African toad *Xenopus laevis* has been an ideal subject for developmental biology and molecular biology research. The young female frog carries a large number of oocytes in her ovary, and each oocyte is large. The large size of the oocytes makes them convenient for *in vivo* experimental operations, such as microinjection. Fertilized *Xenopus* oocytes (eggs) can develop in cell culture, providing an ideal system for embryology studies. Furthermore, immature oocytes contain high copy numbers of 5S ribosomal RNA genes (5S DNA), and store millions of transcription factor IIIA (TFIIIA) protein molecules, 5S ribosomal RNA (5S RNA) molecules, and later in oogenesis, a large pool of ribosomal protein L5 (L5). These advantages facilitate research on the developmental control of gene expression and ribosome formation.

In early studies, the oocytes were used to isolate the abundant macromolecules (5S DNA, 5S RNA, TFIIIA and L5). More recently, all these biological materials have been available from other sources: 5S DNA can be synthesized *in vitro* and amplified in bacteria, eg. *Escherichia coli* (*E. coli*); 5S RNA can be obtained by *in vitro* transcription; and the genes of TFIIIA and L5 proteins have been cloned and expressed in *E. coli*. It is no longer necessary to isolate these macromolecules from the frogs for *in vitro* experiments.

1.1. 5S RNA - FROM THE GENE TO THE RIBOSOME

Since the discovery of catalytic RNA (Zaug *et al.*, 1986), there has been a theory that postulates there was an RNA world. In this early stage of the evolution of life, RNA could fulfill its functions and re-produce itself without the presence of any protein or DNA.

However, in the present world, protein-nucleic acid interactions are essential for almost all DNA/RNA biological functions. For instance, protein-nucleic acid interactions occur in every stage of the biological pathway of 5S ribosomal RNA.

The initiation of 5S RNA gene transcription begins with the binding of transcription factor IIIA to a 50 bp region within the coding sequence of the gene, known as the internal control region (ICR), followed by two other protein factors: TFIIIB and TFIIIC. The transcription initiation complex is then joined by RNA polymerase III (pol III). The details of *Xenopus* 5S RNA gene transcription initiation will be discussed in a later section.

The termination of 5S DNA transcription involves the recognition of a sequence signal at the end of the gene by RNA polymerase III (Bogenhagen & Brown, 1981, Cozzarelli *et al.*, 1983). Although pol III alone is capable of recognizing the terminal signal, it is clear that a 50 kD protein, La antigen, is involved in facilitating the termination process (Rinke & Steitz, 1982; Stefano, 1984). La-protein is structurally distinct to TFIIIA, and belongs to a family of RNA binding proteins sharing a domain of 80 amino acid residues, which is known as the RNA recognition motif (RRM) (Query *et al.*, 1989). It seems that the proteins recognize the termination signal on the nascent 5S RNA rather than the DNA template. It has been demonstrated that La protein binds to the U-rich 3' terminus of the transcript, and leaves the 5S DNA template with the transcript as a ribonucleoprotein complex (5S*RNP, Steitz *et al.*, 1988).

The La protein-5S RNA complex is a transient form of RNP. La-protein is replaced by TFIIIA soon after transcription. The newly formed complex, namely the 7S RNP, consists of one molecule each of 5S RNA and TFIIIA (Picard & Wegnez, 1979). 7S RNPs are exported from the nucleus to the cytoplasm. 7S RNPs are very stable and serve as storage particles for 5S RNA, accumulating in the cytoplasm of the oocyte until 5S RNA is needed by the cell. There is evidence that the association of 5S RNA and TFIIIA is

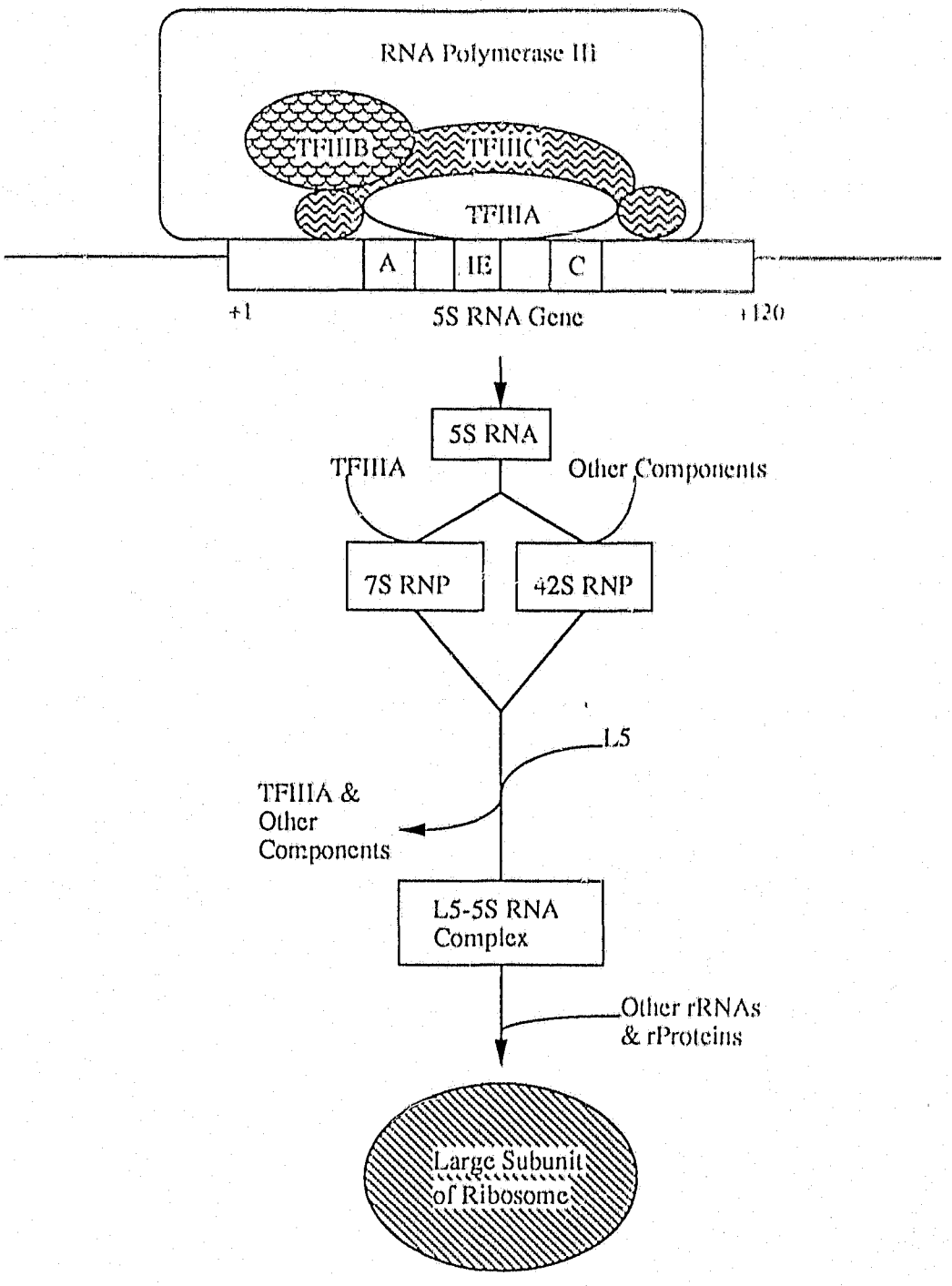


Figure 1: The biological pathway of 5S RNA

required for 5S RNA export from the nucleus (Allison *et al.*, 1991). Another 5S RNA binding protein, ribosomal protein L5, is also believed to be involved in the migration of 5S RNA from the nucleus into the cytoplasm (Allison *et al.*, 1991). Exactly how these two types of RNPs, TFIIIA-5S RNA and L5-5S RNA, migrate out of the nucleus is still under investigation.

In addition to the 7S RNP, roughly half of 5S RNA molecules are stored in 42S RNPs, which are much larger ribonucleoprotein particles consisting of one 5S RNA, two proteins and some tRNA molecules. 5S RNA can be stored as 7S and 42S RNPs in immature *Xenopus* oocytes for months before the massive assembly of ribosomes in the nucleus begins after fertilization.

When it is needed, 5S RNA migrates back into the nucleus with L5 protein. L5-5S RNA is the only form of 5S RNA complex detected in the nuclei during later oogenesis-stages, suggesting that L5 might be the only 5S RNA binding protein responsible for the import of 5S RNA into the nucleus (Allison *et al.*, 1991). The L5-5S RNA complex has been identified as a precursor to ribosome assembly (Steitz *et al.*, 1988).

5S RNA is incorporated into the large subunit of ribosomes in the nucleolus. 5S RNA is then once again transported into the cytoplasm compartment, within the ribosome subunits, to fulfil the great demand for protein synthesis in the later stage of oogenesis. Protein-5S RNA interactions play a crucial role in all the movements of the RNA between nucleus and cytoplasm compartments. These events are summarized in Figure 1.

My Ph.D. project has focussed on three of the protein-nucleic acid interactions: TFIIIA-5S RNA, TFIIIA-5S DNA and L5-5S RNA. Details of those studies are presented in Chapters 2, 3 and 4, respectively.

1.2. THE COORDINATION OF RIBOSOMAL COMPONENT SYNTHESIS

Ribosomes consist of two subunits. In *E. coli*, the intact ribosome has a sedimentation coefficient of 70S. When disassociated it produces two subunits, 30S and 50S in size. In eucaryotes, the ribosomes are larger: the 80S intact ribosome consists of 60S and 40S subunits. The eucaryote ribosome contains an 18S rRNA in its small subunit, and the large subunit contains three rRNA species, 5S, 5.8S and 28S rRNA. The two subunits together contain over 70 ribosomal proteins.

The formation of ribosomes requires all three types of RNA polymerase: RNA polymerase I is involved in the transcription of 28S, 18S and 5.8S rRNAs; RNA polymerase III transcribes 5S rRNA (and tRNA); the ribosomal proteins are translated from mRNAs that have been synthesized by RNA polymerase II. The coordination of ribosomal component synthesis is a very complex but interesting puzzle - a puzzle that has not been solved yet.

The synthesis of 5S RNA is not coupled with that of the other rRNA species. The RNA polymerase III transcripts (5S RNA and tRNA) are stored in RNPs and accumulate in early *Xenopus* oocytes before the other rRNA synthesis begins. When the other rRNA levels increase in the later stages of *Xenopus* development, 5S RNA transcription has been repressed to a low level, and the stored 5S RNA matches the requirement of balanced rRNA molecules for ribosome assembly.

The expression of other rRNA (28S, 18S and 5.8S rRNA) genes by RNA polymerase I shows a stoichiometric pattern and is regulated largely at the level of transcriptional initiation. These rRNA genes are organized in a single transcription unit, ensuring equimolar transcription. Upstream enhancer elements (60/81 bp elements) have a strong positive effect on transcription of these rDNA. Numerous investigations have demonstrated that: 1. These elements may be involved in the stable binding of RNA polymerase I transcription factors (Culotta & Sollner-Webb, 1985; Dunaway & Reeder,

1985); 2. These enhancer elements are polymerase-specific, for they do not stimulate synthesis from a juxtaposed polymerase II or polymerase III promoter (Sollner-Webb *et al.*, 1985); 3. The rDNA 60/81 elements have a cumulative effect, the rDNA genes having more such elements being preferentially transcribed (Busby & Reeder, 1983; Reeder *et al.*, 1983). The transcription of rRNA by polymerase I is also dependent upon the levels of the enzyme and some protein factors, for example, factor C, which may be a dissociatable part of polymerase I.

In summary, synthesis of RNA polymerase I-related rRNA is uncoupled with the synthesis of 5S RNA. They are transcribed by different polymerases, at different stages of *Xenopus* development and regulated by different mechanisms. The balanced levels of rRNA molecules for the ribosome assembly appear to be achieved by the pre-storage of 5S RNA and the increase in transcription of other rRNAs that occurs later during oogenesis.

Similarly, the synthesis of ribosomal protein L5 is not coordinated with the synthesis of other ribosomal proteins. The expression of the L5 gene precedes the maximal synthesis of other ribosomal proteins. An unusually large non-subunit-associated pool of L5-5S RNA is present before the formation of ribosomes (Wormington, 1989). The complexes of these two independently regulated ribosomal components (L5 and 5S RNA) serve as the precursor for the 60S subunit assembly.

The expression of other ribosomal proteins is clearly controlled in a way that equimolar synthesis can be achieved. In *E. coli*, ribosomal protein genes are organized in several operons. One of the ribosomal proteins encoded in each operon is a translational repressor to that operon. This is a direct feedback regulation system that controls the expression of all ribosomal proteins in the unit (reviewed by Nomura *et al.*, 1984; Mager, 1988). In contrast, the control of ribosomal protein expression in eucaryotes is more complicated. There is no operon-like gene arrangement in eucaryotes; instead, ribosomal protein genes are dispersed throughout the genome. The eucaryote regulatory mechanism

is based upon the same feedback principle but uses different molecular strategies. These strategies may involve controls at the levels of transcription, mRNA splicing, mRNA stability and translation. It was observed that over-produced ribosomal protein can be quickly degraded in order to maintain the balance of ribosomal components (Kongsuwan *et al.*, 1985).

The co-ordination between ribosomal RNA and ribosomal protein synthesis is still poorly understood. However, some experimental observations have shown clues for the existence of such a coupling mechanism. For example, experiments showed that when produced in excess over ribosomal RNA, the ribosomal proteins can inhibit the splicing of their own messengers.

1.3. THE 5S RNA GENE FAMILIES

The *Xenopus* genome contains at least three kinds of 5S RNA genes. Oocytic 5S RNA genes (Xlo 5S gene) are the major 5S genes actively transcribed in the immature oocytes, but is repressed in mature oocytes and somatic cells. This is the largest 5S RNA gene family. There are 20,000 to 24,000 copies of Xlo 5S RNA gene per haploid genome (Brown *et al.*, 1977), comprising 0.7% of total genomic DNA (Brown *et al.*, 1971). The Xlo 5S genes are located at the telomere of the long arms in most of the chromosomes (Pardue *et al.*, 1973). These genes are organized as tandemly repeated units, which probably arose through duplication of a primordial gene (Brown *et al.*, 1971).

The spacer between the repeated 5S genes varies in length from about 360 to 570 or more nucleotides. There are AT-rich regions and GC-rich regions. The AT-rich sequence is internally repetitive. The GC-regions are immediately adjacent to both 5' and 3' termini of the Xlo 5S gene. They are much less repetitive. Each 5S gene repeat unit also contains a pseudogene following the 3' GC-rich spacer region (Fedoroff & Brown, 1978; Miller *et al.*, 1978). Pseudogenes are a duplication of the first 101 bp of the full length gene with

90% homology and are transcriptionally silent. The entire repeat unit (5S gene, spacer and pseudogene) is heterologous in length (600-1000 bp). Analysis of the repeat sequence indicates that duplication of the oocyte specific gene was a recent event in the evolutionary development of *Xenopus* (Federoff & Brown, 1978). The spacers are sites of most recent duplications and therefore consist mainly of unselected sequences with little function in gene regulation (Federoff & Brown, 1978). For example, the promoter of the gene is within the gene coding sequence, not the 5'-flanking spacer sequence. Likewise, the termination of 5S gene transcription does not require the 3' flanking spacer (Bogenhagen & Brown, 1981).

In contrast, the somatic 5S RNA genes (Xls 5S gene) have only 400 copies per haploid genome (Peterson *et al.*, 1980), and do not contain a pseudogene. The Xls 5S gene family is located on more than one chromosome (maybe 3-4), as indicated by genetic crossing experiments (Peterson *et al.*, 1980). This gene family has a GC-rich spacer and the repeat units are homogeneous in length. The Xls 5S gene differs from the Xlo 5S gene by six nucleotides, three of them located within the intragenic promoter region. Xls 5S genes are expressed constitutively in all stages of *Xenopus* development, comprising 8% to 10% of 5S RNA in oocytes and over 95% in somatic cells. The transcript of the Xls 5S gene is not processed, as it is expressed as the mature 120 nt. 5S RNA.

The third 5S gene family is the *Xenopus* trace oocyte 5S gene (Xlt 5S gene). The Xlt 5S gene family contains 1200 to 2000 copies per haploid genome, comprising 0.03% of total genomic DNA (Brown *et al.*, 1977). These 5S genes are present on at least 4 chromosomes (Peterson *et al.*, 1980). The Xlt 5S gene clusters have AT-rich spacers and are homogeneous in repeat unit length (ca. 350 bp). Xlt genes are expressed with a similar pattern to the Xlo 5S genes. This minor group of 5S genes is not studied in great detail, and therefore is not the major subject of this thesis.

All three classes of 5S genes are organized in simple tandem repeat units. This arrangement of genes is thought to be involved in a control mechanism that would shut off many genes at once (Peterson *et al.*, 1980). Such a mechanism could be chromatin packaging, as will be discussed in detail in a later section of this chapter.

1.4. STRUCTURE AND FUNCTIONS OF 5S RNA

1.4.1. Structure of 5S RNA.

5S RNA is a small 120 nucleotide molecule. This simple molecule generally does not contain modified nucleotides (Erdmann, 1981). The generalized secondary structure of 5S RNA was established mainly by two approaches: a computer search for the most thermodynamically favourable structure, and a comparison of 5S RNA species from different organisms covering a large evolutionary range (Nussinov *et al.*, 1982). The established 5S RNA structure model was further investigated by various ribonuclease and chemical probes (Noller & Garrett, 1979; Toots *et al.*, 1981; Trout *et al.*, 1982; Silberklang *et al.*, 1983; Kjems *et al.*, 1985; Digweed *et al.*, 1986; Sneath *et al.*, 1986 & Romaniuk *et al.*, 1988). 5S RNA has been a popular subject for research because of its small size, simple structure and ubiquity. Hundreds of 5S RNA sequences of different organisms, ranging from bacteria, archaebacteria, fungi, protozoa, and plants to vertebrates, have been determined (Wolters & Erdmann, 1988). The generalized 5S RNA secondary structure is shown in Figure 2 and the general eucaryotic 5S RNA model is shown in Figure 3.

RNA is capable of forming double helical stems by folding the same chain back on itself in a anti-parallel orientation so that complementary bases may hydrogen bond. Such helices of RNA usually are structures similar to the A form of DNA double helix. Since RNA molecules do not possess the regular interstrand hydrogen-bonded structure

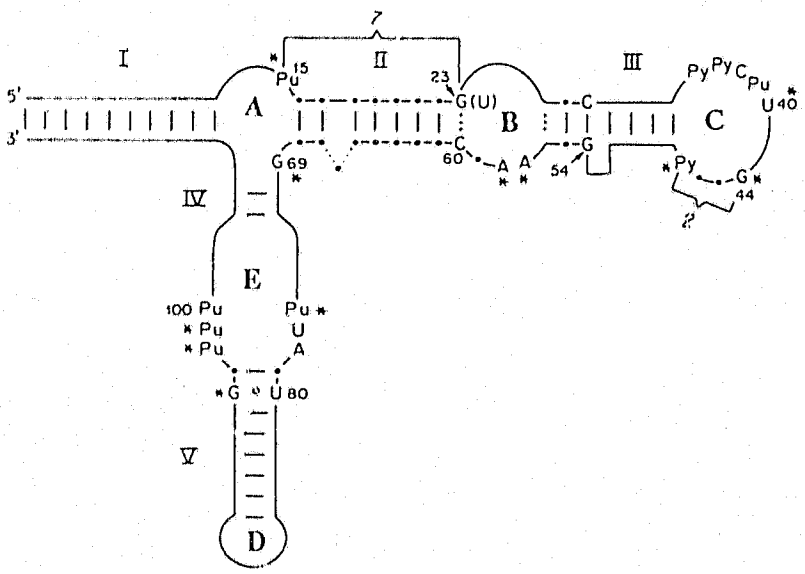


Figure 2: Generalized 5S RNA secondary structure.
The numbering system corresponds that of the *E. coli* 5S RNA. The five helices are labelled I-V, the loops A-E. Nucleotide positions shown are universal positions. Positions marked with asterisks designate total invariance. Dots represent constant chain lengths between universal positions found in most compared 5S RNAs; the chain lengths between Pu15 and G23, G44 and Py47 are always 7 and 2 nucleotides, respectively. Pu represents purine; Py pyrimidine (after Delihas & Andersen, 1982).

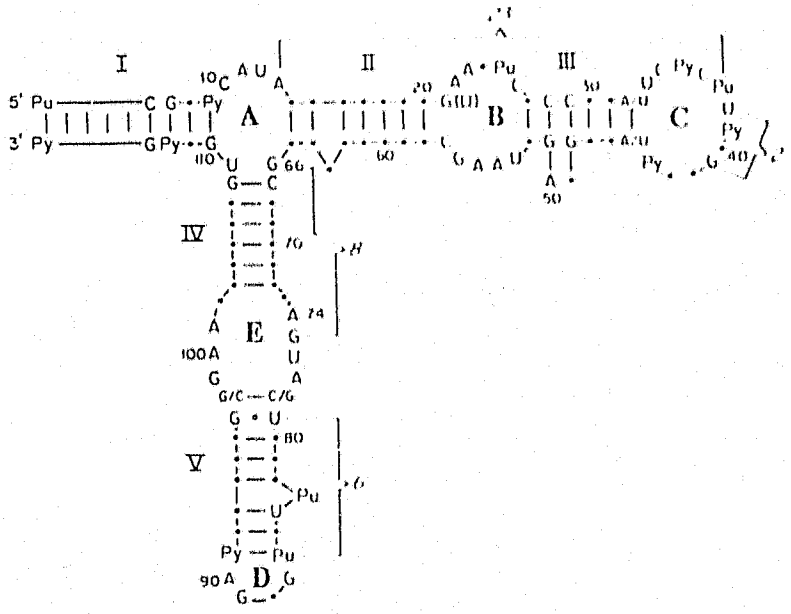


Figure 3: Eucaryotic 5S RNA secondary structure showing common positions found in more than 90% of compared eucaryotic 5S RNAs. The numbers between these positions represent common chain lengths found in eucaryotic 5S RNAs (after Delihas & Andersen, 1982).

characteristic of DNA and do not have long stretches of complementary sequences, the helical stems are usually short, with the uncomplementary nucleotides "looped out" of the structure. These stem-loops are the most common structural features of RNA molecules.

The secondary structure of 5S RNA consists of five stems (I to V), five loops (A to E), and several "bulged" nucleotides which are individual nucleotides that are excluded from the double helical structure (Fig.2). These basic structural features have been well conserved during evolution. Also well conserved are the nucleotides at some particular positions throughout the RNA molecule. These positions are marked in Figure 2 and Figure 3. The conservation of specific secondary structural features and universal nucleotides implies their importance to the molecule's function, for which they have constantly been chosen by evolutionary selection.

Unlike DNA double helical structure, non-Watson Crick base pairing is common in RNA structures. Some examples of these non-canonical base interactions are listed in Table 1. Non-canonical base pairing confers complexity to the understanding of RNA structure, because different structural models can be built based on different theoretical base pairing of the non-canonical interactions. For example, loop E of 5S RNA could be turned into a double helical structure by non-Watson Crick base pairing, making a long stem that includes stem IV and V as well as the loop E region. Evidence for this extended base pairing comes from cobra venom ribonuclease V_1 digestion (which is specific for double stranded regions of RNA) (Anderson *et al.*, 1984) and from chemical reactivity data (Romaniuk *et al.*, 1988). Similar structure was found in *Sulfolobus acidocaldarius* 5S RNA, in which the loop E region is completely base-paired (Stahl *et al.*, 1981). Even more structural alternatives are seen in the tertiary structure models of 5S RNA.

Less is known about the tertiary structure of 5S RNA. It seems that conventional Watson Crick base pairing as well as non-Watson Crick base pairing are both important in

Table 1. Non-canonical interactions in nucleic acids.*

Interactio	Hydrogen Bond	Observed in
A-U	(N7, N6)-(N3, O2)	$tRNA^{phe}$, $tRNA^{asp}$: A_{14}-U_8, A_{58}-T_{54}[1,2,3],
A-U	(N1, N6)-(N3, O2)	$tRNA^{asp}$: A_{15}-U_{48}[3]
A-G	(N1, N6)-(N1, O6)	$tRNA^{phe}$, $tRNA^{asp}$: A_{44}-G_{26}[1,2,3]
A-G	(N6)-(N7)	$tRNA^{asp}$: A_{46}-U_{13}[3]
A-A	(N6, N7)-(N7, N6)	$tRNA^{phe}$, $tRNA^{asp}$: A_9-A_{23}-U_{12}[1,2,3]
G-G	(N3, N2)-(N2, N3)	β-dodecamer[4]
G-G	(N6)-(N7, O6)	$tRNA^{asp}$: G_{45}-G_{10}-U_{25}[3]
G-G	(N6)-(O6)	$tRNA^{phe}$: G_{45}-G_{10}-U_{25}[1,2]
G-G	(N1, N2)-(N7, O6)	$tRNA^{phe}$: m^7G_{46}-G_{22}-C_{13}[1,2]
U-U	(N3, O4)-(O2, O3)	$tRNA^{asp}$: U_{35}-U_{35}[3] (anticodon-anticodon interaction)
C-C	(N4, N3)-(N3, O2)	$tRNA^{gly}$: C_{35}-C_{35}[5] (anticodon-anticodon interaction)
C-C	(N4, N3, O2)-(O3, N3, N4)	PolyC[6]

1 Quigley *et al.* (1975)
2 Jack *et al.* (1976)
3 Westhof *et al.* (1985)
4 Wing *et al.* (1980)
5 Romby *et al.* (1986)
6 Cantor and Schimmel (1980)
* (after Ehresmann *et al.*, 1987)

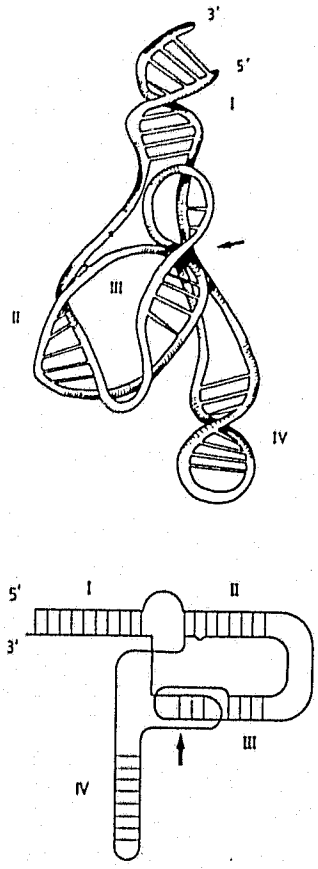


Figure 4: Schematic illustration of the *E. coli* pseudoknotted 5S RNA structure.
A: Spatial illustration of the 5S RNA structure; B: Schematic arrangement of the 5S RNA structure. Helices I to V are indicated by roman numberals. Arrows point to the regions where pseudoknots are formed (after Göringer & Wagner, 1986).

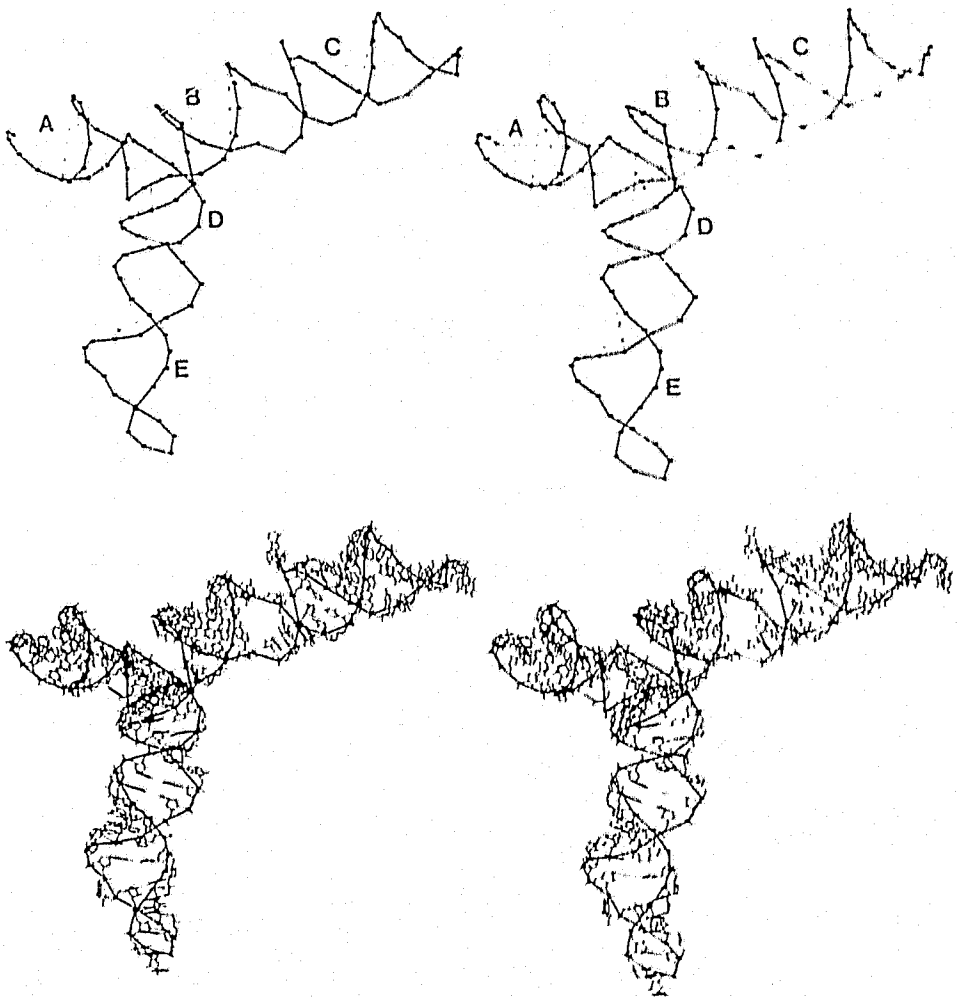


Figure 5: Three-dimensional stereo model of *Xenopus laevis* oocyte 5S RNA
(Top) Phosphate backbone with the Watson-Crick base pairs joined by lines and the non-canonical or tertiary base pairs joined by dotted lines. (Bottom) Atomic view in the same orientation with the phosphate backbone shown in heavy lines. Stems marked with A-E equal to I-V in other figures. All stereo views were drawn with the program PLUTO (after Westhof *et al.*, 1989).

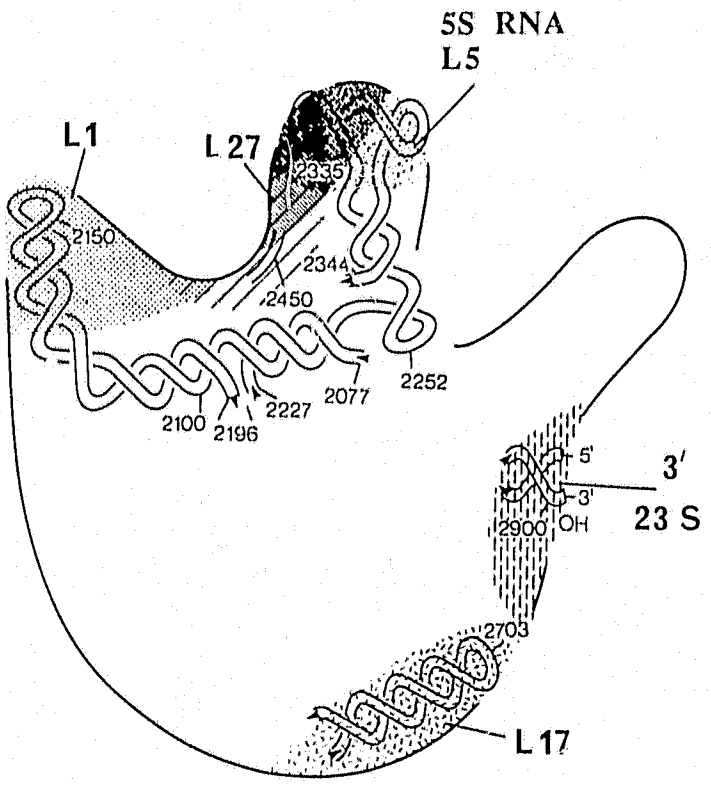


Figure 6: Location of 5S RNA and L5 ribosomal protein in the 50S subunit of *E. coli* ribosome (adapted from Noller & Lake, 1984).

determining the 5S RNA conformation. Base pairing between loops of the same molecule, forming a structure called a pseudo-knot, are an important part of the tertiary structures. Such a structure proposed for *E. coli* 5S RNA is shown in Figure 4. A three-dimensional model of *Xenopus* 5S RNA has been built by Westhof *et al.* (1989). This model was constructed by computer graphic modelling using stereochemical constraints and experimental data on the accessibility of bases and phosphates to structure-specific probes (Fig.5). This model shows that the molecule is Y-shaped; that loop A serves as a hinge that controls the coaxial stacking of the helical domains; that no tertiary interactions occur between loop structures; and that internal loop E contains several non-canonical of A-A, U-U and A-G type base pairs. This model follows closely the generalized secondary structure of 5S RNA, and is well supported by previous observations on the 5S RNA structure. However, no tertiary model of 5S RNA has been thought to be universal and to accurately reflect the true conformation of the molecule in solution. New models will be constructed with the accumulation of more detailed data from future experiments.

1.4.2. The Functions of 5S RNA.

5S RNA is a universal component of the ribosomal large subunits of various organisms or organelles, including procaryote, eucaryote cytoplasmic, mitochondria and plant chloroplasmic ribosomes. The universal existence of the small ribosomal RNA as well as the conservation of its structure and nucleotides make it likely that they are important for the ribosome functions. But whether 5S RNA is important to the ribosome simply because of its structural role in holding the ribosomal proteins, or because of its involvement in the protein synthesis activities of the ribosome, has long been in dispute.

Immune electron microscopy has located the L5-5S RNA complex in the central protuberance of the *E. coli* large ribosomal subunit, (Shatsky *et al.*, 1980; Clark & Lake, 1983; Fig.6). 5S RNA also contacts a number of other ribosomal proteins in the 60S subunit. It has been shown that 5S RNA of rat liver ribosomes forms a complex with L5,

L6 and L18, and that ribosomal proteins L7, L8 and L35 bind to this complex, but more loosely (Spirin, 1986). This eucaryotic ribosomal complex is capable of tRNA binding.

Just before I started to write this thesis, an exciting discovery was published in *Science* (June 5, 1992) by Noller *et al.* The authors demonstrated that it is indeed the RNA components of the ribosome, not the protein components, that possess the peptidyl transferase activity. This discovery means that it is not the protein, but the RNA that is making protein. This new evidence strongly supports the "RNA world" hypothesis. The experiment originated from a simplified protein synthesis system, which includes only the large ribosomal subunits, appropriate ionic strength, and 33% methanol or ethanol, in addition to a f-met-oligonucleotide (a fragment of f-met-tRNA) and puromycin substrates. The reaction transfers the f-met amino acid residue from the tRNA fragment to the puromycin substrate, forming f-met-puromycin without the presence of small ribosomal subunits, mRNA or GTP. Noller *et al.* removed 95% of the protein components of the ribosomal large subunit by SDS treatment, protease K digestion, and three successive phenol extractions. It was a surprise that after large ribosomal subunits from the thermophilic eubacterium *Thermus aquaticus* were treated in this fashion, 80% of the ribosomal peptidyl transferase activity remained in the almost protein-free system. In contrast, when the ribosomal subunit was treated with RNase T_1, no detectable transferase activity was found. Even though the authors pointed out that about 5% of the original ribosomal protein was still present in the reaction system - and a firmly established conclusion requires a protein-free system - the results from Noller *et al.* are already convincing: the vigorous phenol extractions should have destroyed any activity of the remaining protein, and the peptidyl transferase activity did not correspondingly decrease with the decrease of the remaining protein levels after each extraction, indicating that the enzymatic activity is not dependent on the amount of protein present in the system. People used to ask: "What is RNA doing in the ribosome?", and since the proteins are not required

for linking amino acids together, people must now ask: "What are the proteins doing?" (Waldrop, 1992). The authors have evidence to suggest that the 23S ribosomal RNA in the large subunit is responsible for this peptide bond linkage, but since the other rRNA molecules are present in the system, their possible role in the activity can not yet be ruled out.

In fact, it was suggested long ago that 5S RNA might play a role in the peptidyl transferase activity (Raackee, 1971). The hypothesis was based on a space filling model of the ribosome. The position of 5S RNA in the large subunit of the ribosome places the 3'-OH of 5S RNA in a stereochemically perfect position for executing a nucleophilic attack on the carbonyl of the peptidyl-tRNA, thereby effecting the peptide transfer to 5S RNA. Furthermore, when the peptide is on the 5S RNA, the amino group of the acceptor aminoacyl-tRNA is in a perfect position for a nucleophilic attack on the carbonyl of the peptide, causing its transfer to the aminoacyl-tRNA, and forming a new peptide bond.

Related to this possible activity of 5S RNA, the *in vitro* reconstitution of the 50S ribosomal subunit (Nomura & Erdmann, 1970) showed that particles lacking 5S RNA exhibited greatly reduced enzymatic binding of amino acyl-tRNA to the ribosomal A-site (Erdmann *et al.*, 1971; Nierhaus & Dohme, 1974; Dohme & Nierhaus, 1976). This is the second suggested 5S RNA function in ribosome: tRNA binding. It has been suggested that the G_{44}-A-A-C_{47} conserved 5S RNA sequence may base pair with the T-Ψ-C-G of tRNA, which is also strongly conserved in loop IV of all transfer RNAs (Brownlee *et al.*, 1968; Ofengand & Henes, 1969). However, the proposal that the T-Ψ-C-G sequence in tRNA is involved in binding to the A-site of the ribosome was not supported by a experiment conducted by Pace *et al.* (1982). It was demonstrated in this experiment that the conserved 5S RNA sequence G-A-A-C is not essential for the mechanics of protein synthesis.

In another experiment, when two adenines, A73 and A99 in *E. coli* 5S RNA, were modified with monoperphthalic acid, the 5S RNA was still capable of being incorporated into the ribosomal large subunit, but the protein synthesis activity of the ribosomal particle thus formed was reduced by 50% (Erdmann *et al.*, 1973; Silberklang *et al.*, 1983). The authors suggested that these two adenines may be located on the ribosomes surface and be involved in the binding of tRNA.

Other proposed 5S RNA functions include GTPase activity and involvment in ribosomal subunit association. However, more experimental data is required before these hypotheses can be confirmed. In contrast, strong evidence has emerged to suggest that the 23S and 16S ribosomal RNAs are involved in most of the above mentioned ribosomal activities (reviewed by Noller, 1991). While the role of 5S RNA in these ribosomal functions is uncertain, it is commonly accepted that the 5S RNA-L5 complex is the core structure for ribosome assembly. Steitz *et al.* (1988) have shown that this 5S RNA complex is the precursor for ribosome formation in *Hela* cells.

The ribosome is a large piece of molecular machinery for protein synthesis. This complicated organelle contains four ribosomal RNA molecules and over seventy ribosomal proteins. It would be hard to believe that any single component, whether an rRNA or a ribosomal protein, could fulfil a major activity of the ribosome by itself. More likely, coordination between the various components is essential. The discover of RNA catalytic activity demonstrates the importance of the rRNA in the ribosome, not only structurally, but also functionally. A more clearer picture of the 5S RNA role in these activities will emerge from future investigations.

1.5. STRUCTURE AND FUNCTIONS OF TFIIIA

1.5.1. Structure of TFIIIA

5S RNA interacts with a number of protein factors including the termination factor La antigen, the positive transcription factor TFIIIA, storage particle component p43 and ribosomal protein L5. These 5S RNA binding proteins are not structurally related, except for TFIIIA and the p43 protein, which both contain a now well studied structural motif called the "zinc finger."

Xenopus TFIIIA was the first protein found to contain a sequence motif of the form X_3-Cys-X_{2-4}-Cys-X_{12}-His-X_{3-4}-His-X_4 (where X is any amino acid) (Miller *et al.*, 1985). This motif is known as a "C_2H_2" type zinc finger, because it contains two cysteine and two histidine invariable residues in each unit which coordinate to a zinc ion. Since the discovery of this motif, hundreds of similar protein motifs have been reported. Examples of zinc fingers include proteins involved in regulation of differentiation and growth (EGR1, Sukhatme *et al.*, 1988; Christy *et al.*, 1988), (EGR2, Chavrier *et al*, 1988), DNA binding proteins with regulate proto-oncogene (GL1, Kinzler *et al*, 1988), Wilm's tumor gene (Call *et al.*, 1992), *Drosophila* segmentation genes (Hunchback *et al.*, 1987; Kruppel *et al.*, 1986) and steroid hormone receptors (Härd *et al.*, 1990). The cysteine and histidine composition can vary. For example, the motif can be "C_3H" or "C_4" types, although the "C_2H_2" type is the major group of zinc finger proteins.

TFIIIA is a basic protein, 38.5 kD in size, comprising 344 amino acids (Ginsberg *et al.*, 1984, Miller *et al.*, 1985, Bieker & Roeder, 1984). Sequence alignment of this protein revealed 9 tandemly and imperfectly repeated units through the N-terminal two thirds of the protein. Each unit contains 30 amino acid residues, including two cysteines and two histidines highly conserved at the same positions of each unit (Fig.7). Since it was known that TFIIIA requires zinc for its activity (Hanas *et al.*, 1983), and cysteine and

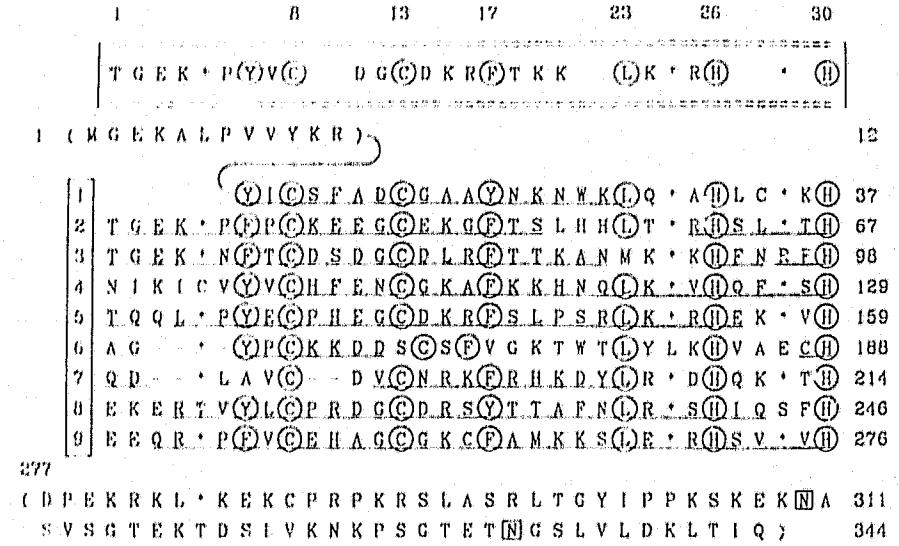


Figure 7: Amino acid sequence of TFIIIA from *Xenopus laevis* oocytes, aligned to show the repeated units. The molecule contains: an amino end region (residues 1-12); a lysine-rich zone (277-309) near the carboxyl end. The repeat units are numbered 1-9 on the left side of the diagram. The box-in consensus sequence at the top shows the characteristic features of a typical repeat unit, numbered as for a length of 30 residues. The end-point of each unit has been chosen arbitrarily after His-30. The best-conserved residues are ringed. Asterisks (*) mark positions where an insertion somtimes occurs in the normal pattern, and dots (.) mark variable positions in the sequence. In the main body of the repeats a dash (-) indicates an alignment gap. The underlined regions are those which show clear evidence of a relationship with at least one other unit (after Miller *et al.*, 1985).

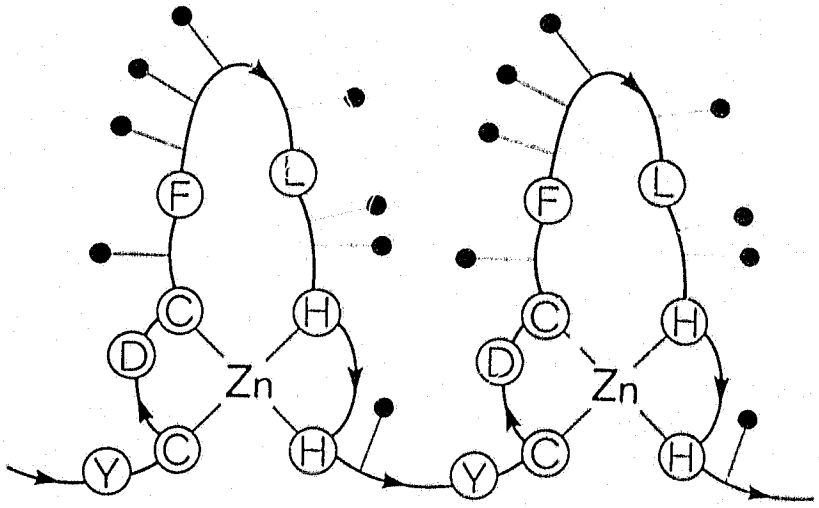


Figure 8: Folding scheme for a linear arrangement of repeated units. Ringed residues are the conserved amino acids. Black circles mark the most probable DNA binding side chains (after Miller *et al.*, 1985).

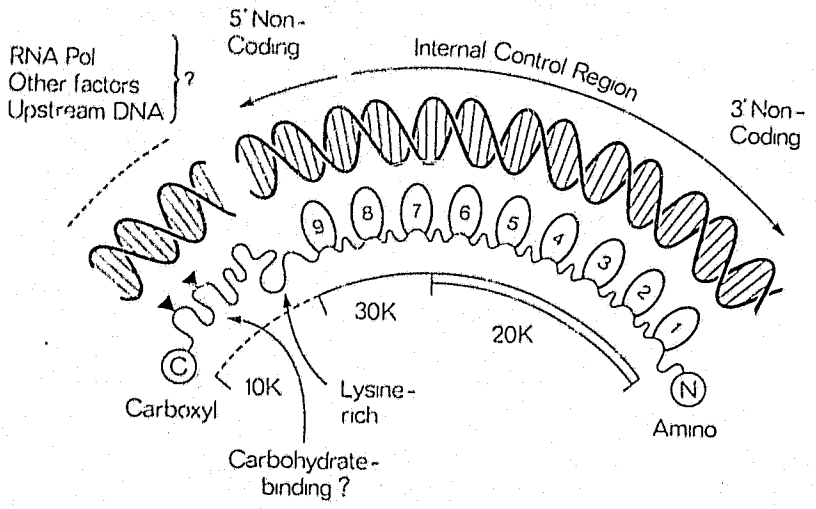


Figure 9: An interpretation of the structural features of the protein TFIIIA and its interaction with DNA (after Miller *et al.*, 1985).

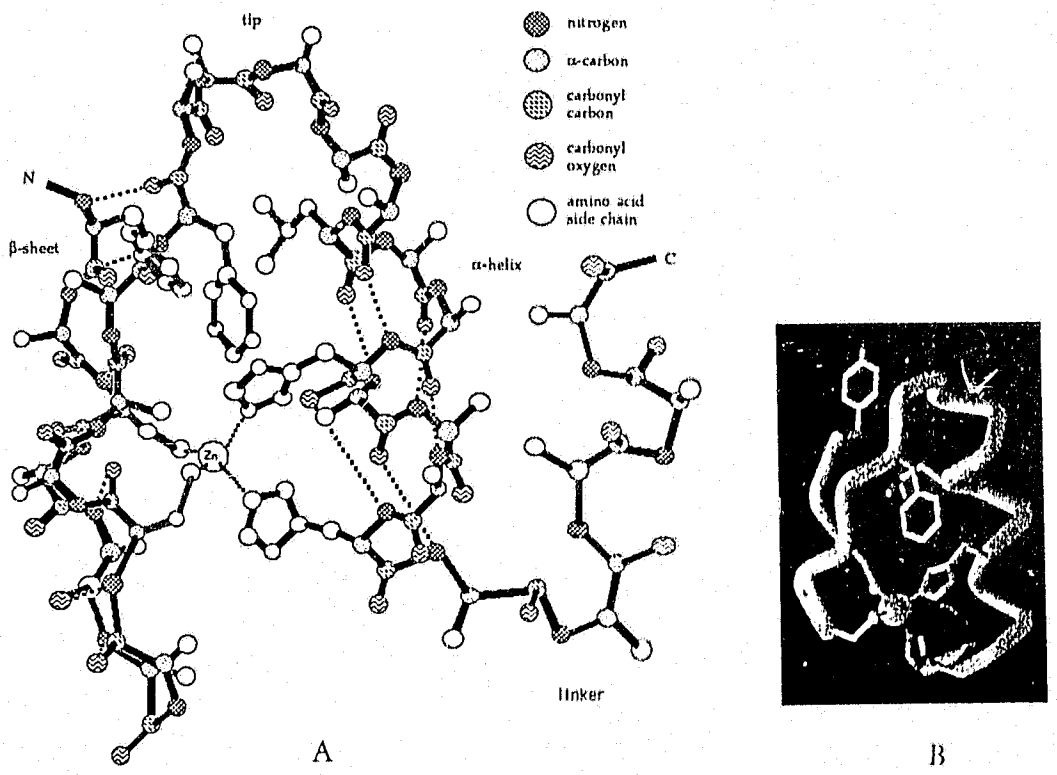


Figure 10: Three-dimensional structure of zinc finger.
A: a proposed model of a TFIIIA zinc figure. (Modified from Berg, 1988).
B: model of a single zinc finger. (Illustration by Lee *et al.*, 1989).

histidine are the most common ligands in enzymes (Fersht, 1977), Miller *et al.* suggested that each repeat unit may fold into a compact motif that is formed around a central zinc ion to confer stability. Each zinc would be coordinated tetrahedrally to the two invariant pairs of Cys and His residues in each unit. This motif is called a "zinc finger" (Fig.8).

TFIIIA contains nine such finger subdomains in the N-terminal two-thirds of the protein. The C-terminal one-third is not repetitive but is Lys rich. Proteolysis produces three fragments of TFIIIA, namely 30K, 20K and 10K fragments (Smith *et al.*, 1984). DNA footprinting showed that these fragments protect different parts of the internal control region of 5S DNA. The N-terminal 20K fragment makes contacts over smaller regions of the ICR than the 30K fragment does. The C-terminal 10K region of TFIIIA does not directly bind to the DNA, but enables the protein to enhance transcription, presumably by contacts with RNA polymerase III or other transcription factors. These observations indicate that only the finger region is responsible for the specific DNA binding, and since the 20K fragment lacking three of the nine fingers is still fully capable of DNA binding, the fingers may be functionally independent. Each finger may make individual contacts with the DNA and occupy half a turn of the DNA double helix. The orientation of TFIIIA zinc fingers on the binding of 5S DNA was hypothesized by Miller *et al.* and is shown in Figure 9.

In the first paper published on the finger protein, Miller *et al.* predicted that the number of fingers might vary in other proteins containing zinc fingers, which they expected to be identified in later studies. This prediction has been proven true. The number of fingers among the known proteins of this class varies from two to thirty seven!

Later investigations of the finger structure revealed more details. The nuclear magnetic resonance (NMR) studies have shown that each TFIIIA-like zinc finger contains an antiparallel ß-sheet (ribbon) and an α-helix. The four invariant residues coordinate the central zinc ion (Párraga *et al.*, 1988; Lee *et al.*, 1989) (Fig.10). The folding of the finger

motif is further stabilized by the side chains of aromatic amino acids at different positions within the motif. These side chains stack in the center of the finger, forming a hydrophobic core which provides additional stability to the structure in solution (Kochoyan *et al.*, 1991).

1.5.2. Functions of TFIIIA.

As a positive transcription factor, the first identified function of TFIIIA was the activation of 5S RNA gene transcription. As discussed in previous sections, the binding of TFIIIA is the beginning of the transcriptional initiation process. The nine fingers of TFIIIA are arranged linearly on the DNA internal promoter, occupying a long stretch of 50 base pair (bp) within the ICR. This arrangement may allow the transcription complex to remain bound after multiple rounds of RNA polymerase III passages. Transcription is carried out continually without forming a new initiation complex for each round of activity. Since the formation of the initiation complex is time consuming (approximately 45 minutes), the linearly arranged nine finger structure of TFIIIA and the way the protein is bound to 5S DNA is a good example of efficiency.

In addition to 5S DNA, TFIIIA also specifically binds to 5S RNA, the very product of the 5S gene. TFIIIA and 5S RNA constitute the well studied storage particle, 7S RNP. Cytoplasmic 5S RNA molecules are stably stored in these particles for months before being transported back to the nucleus for ribosome formation. The dual binding activity of TFIIIA to both 5S RNA and its gene implies a possible feed-back regulation mechanism of 5S RNA gene expression: the binding of TFIIIA to 5S RNA may reduce the transcription of its own gene by depleting the free transcription factors. However, the extreme stability of the transcription complex (particularly the somatic complex) may argue against such a regulatory mechanism, although *in vitro* experiments did show that the addition of exogenous 5S RNA to the transcription system inhibits the synthesis of the RNA (Pelham & Brown, 1980).

The third known function of TFIIIA is its role in the regulation of the differential expression of oocyte and somatic 5S genes. TFIIIA is one of the key factors in the competition between the oocyte and somatic genes for transcription complex formation during *Xenopus* development (see below).

Finally, TFIIIA is required for 5S RNA transportation in the cell. The protein has been shown to be important in the export of nuclear 5S RNA to the cytoplasm in the form of ribonucleoprotein particles (Allison *et al.*, 1991).

The various functions of TFIIIA listed above are all related to 5S RNA. It seems that, unlike the other transcription factors, the activity of TFIIIA is commited to 5S RNA only. This commitment is demonstrated not only by the fact that TFIIIA is a specific factor involved only in 5S RNA transcription, but also by the fact that 7S RNP consists of only TFIIIA and 5S RNA. When 5S RNA is stored with other RNA molecules (tRNA) in 42S particles, TFIIIA is not present. Instead, a similar but distinct protein, p43, takes the place of TFIIIA. It remains unclear whether there is a biological significance to such a restrictive commitment.

1.6. DEVELOPMENTAL REGULATION OF SOMATIC AND OOCYTE 5S RNA GENE EXPRESSION

1.6.1 The Developmental Pattern of Xlo and Xls Gene Expression

The two 5S RNA genes are accurately and efficiently transcribed by RNA polymerase III when they are microinjected into oocyte nuclei or incubated in extracts of these same nuclei. However, in immature oocytes of *Xenopus*, over 90% of the 5S RNA population is oocytic type. Contrastingly, in somatic cells, somatic genes represent only 2% of all 5S RNA genes but express over 95% of the 5S RNA population. This is equal to a 1,000-fold higher transcriptional activity of Xls genes over the Xlo genes. How can this happen? There must be a mechanism that activates the Xls genes and inactivates the Xlo

genes. The phenomenon is an example of what may be a common developmental mechanism, where two or more gene families have similar (but not identical) cis-acting regulatory elements that are recognized by the same factor, but are nonetheless differently expressed.

1.6.2 The Repression of Xlo 5S RNA Genes and Activation of Xls 5S RNA Genes.

The mechanism of activation of Xls 5S RNA genes and repression of Xlo genes in *Xenopus* somatic cells was explained by Wolffe and co-workers (reviewed by Wolffe & Brown, 1988). According to this model, the regulation of differential 5S RNA gene expression is dependent on three major conditions: the stability of transcription complexes, availability of TFIIIA, and chromatin assembly. Differences in these conditions in oocyte and somatic cells determine whether a 5S RNA gene will be transcribed or repressed.

Pol III transcription starts with the formation of the initiation complex. In *Xenopus*, TFIIIA is the specific transcription factor that recognizes only 5S RNA gene and is the first factor that binds to the gene's intragenic promoter. The binding of TFIIIA alone is not very stable. The stability of the complex is enhanced by interaction with a second transcription factor, TFIIIC. TFIIIC is believed to contact both the DNA and the bound TFIIIA protein (Hayes *et al.*, 1989; Majowski *et al.*, 1987). The initiation complex is then completed by the additional interaction of the last activity, TFIIIB. TFIIIB may not be a DNA binding protein, but acts on the TFIIIA/TFIIIC/5S DNA complex. The binding of TFIIIB to the complex is the "rate-limiting" step, accounting for a lag period before synthesis of 5S RNA reaches maximal rates (Setzer & Brown, 1985; Bieker *et al.*, 1985). Pol III is not a part of this transcription complex, but recognizes the complex for transcription (Lassar *et al.*, 1983; Bieker *et al.*, 1985; Setzer & Brown, 1985). TFIIIC

and TFIIIB are not 5S RNA gene specific since they are also involved in transcription of other pol III-transcribed genes.

The transcription complex on the somatic 5S gene has greater stability as compared to the one formed on the oocyte specific gene (Wolffe & Brown, 1988). The stability of the transcription complex may depend on several mechanisms, including a conformational change in TFIIIA structure on binding, a cooperative interaction between TFIIIA and TFIIIC, or a covalent modification of TFIIIA, for example, a dephosphorylation or phosphorylation catalyzed by TFIIIC. But these general features do not explain the different stabilities between Xlo and Xls transcription complexes. It is commonly accepted that the difference in stability is due to the six nucleotides that differ between Xlo and Xls, especially the three nucleotides that lie in the 5' part of the ICR (Wolffe & Brown, 1988). These somatic specific nucleotides have only minimal effects on the binding of TFIIIA to the Xls 5S gene, but produce different DNase I footprint patterns (Xing & Worcel, 1989). Nucleotide changes at positions +53, +55 and +56 (from oocytic to somatic) alter the transcription complex pattern from oocytic to somatic type and markedly enhance the level of transcription of the mutant 5S RNA gene above that of Xlo gene (Xing & Worcel, 1989). It is presumed that the nucleotide differences between oocytic and somatic 5S genes may affect the TFIIIA/TFIIIC interaction and therefore affect the stability of the complex (Wolffe & Brown, 1988).

In early stage oocytes, when TFIIIA is abundant, both Xlo and Xls genes are transcribed. Because Xlo 5S genes are present in higher number, the Xlo 5S RNA become the overwhelming population. TFIIIA levels decrease during embryogenesis, and drop to only 0.25 TFIIIA molecule per gene in the nuclei of somatic cells (Table 2). This low level of TFIIIA alone indicates that only a fraction of the genes can be transcribed in the somatic cell. Since Xls genes have a four-fold higher binding affinity for TFIIIA over Xlo genes, it is likely that the transcription complex will be preferentially formed on the Xls genes, and

thus they will be activated preferentially over the Xlo genes. But this fact alone is not sufficient to explain the 1,000-fold transcription strength difference between the two genes. The competition for forming a transcription complex accounts for some of the differential expression, but is not the major reason for the phenomenon. The major difference arises from process of chromatin assembly. The Xlo gene, with a less stable transcription complex, will be packaged into chromatin in the late stages of *Xenopus* development, and lose its transcriptional activity, while the more stable Xls complex is resistant to the assembly of chromatin, and therefore remains in the transcriptionally active state.

1.6.3. The Effect of Chromatin Assembly on Gene Expression.

Oocytic 5S RNA genes isolated in the chromatin of somatic cells are transcriptionally inactive (Bogenhagen *et al.*, 1982; Schlissel & Brown, 1984). Even the addition of all factors (TFIIIA and fractions containing TFIIIB and TFIIIC) plus RNA polymerase III to the chromatin template does not activate the oocyte 5S RNA genes. In order to transcribe the oocytic 5S RNA genes in somatic chromatin, the histone H1 in the nucleosome must be removed (for example, by a high salt wash; Korn & Gurdon, 1981; Bogenhagen *et al.*, 1982), and then all the transcription factors plus RNA polymerase III added (Wormington *et al.*, 1983; Schissel & Brown, 1984). Neither pol III alone nor a nuclear extract lacking TFIIIA can transcribe the oocyte 5S RNA genes in somatic chromatin, even if they have been washed with high salt buffer to remove histone H1. Furthermore, the readdition of histone H1 to chromatin that has been previously depleted of the protein will restore the repressed state of oocyte 5S RNA genes (Schissel & Brown, 1984). That is, the genes already activated by removal of histone H1 will become transcriptionally silent again by the readdition of histone H1. These observations lead to at least two conclusions: 1. The oocyte 5S RNA genes in somatic chromatin are not

Table 2: Levels of TFIIIA during *Xenopus* embryogenesis.

Stage	Molecules/Cell	Molecules/Gene
Immature oocyte	3×10^{11}	5×10^{7}
Mature oocyte	3×10^{10}	5×10^{6}
Unfertilized egg	1.5×10^{9}	4×10^{5}
Blastula embryo	4×10^{5}	10
Gastrula embryo	1×10^{5}	2
Swimming tadpole	1.7×10^{4}	0.4
Cultured cell	1×10^{4}	0.25

(after Wormington *et al.*, 1983)

complexed with the transcription factors. 2. Histone H1 plays an important role in the repression of oocyte 5S RNA genes, presumably by packaging the genes into chromatin. Therefore the genes become inaccessible to the transcription factors to form an initiation complex (Wolffe, 1988; Brown, 1984). The assembly of oocyte 5S RNA genes into chromatin makes them "invisible" to the RNA polymerase III as well as to the transcription factors.

In contrast, somatic 5S RNA genes isolated from somatic cells are transcriptionally active. The *in vitro* transcription of the Xls gene in somatic chromatin requires only the addition of RNA polymerase III; the addition of transcription factors is not needed (Bogenhagen *et al.*, 1982; Wormington *et al.*, 1983). The interpretation of this observation is that the somatic 5S RNA genes in somatic cell chromatin are associated with stable, active transcription complexes. The complexes on the Xls genes are so stable that they stay bound to the genes for up to 40 rounds of pol III passage during transcription, resist the assembly of chromatin, and are able to survive a gentle chromatin isolation procedure. The isolated Xls 5S genes pre-complexed with the transcription factors therefore need only pol III, which is not a part of the complex, for its transcription.

In summary, the activation and repression of 5S RNA genes are determined by the environment around the genes. Both Xlo and Xls genes are transcribed in the oocytes, where the histone H1 is not present, the DNA is not packaged in a higher order structure and TFIIIA is at a saturating level. There is no competition for transcription factors and no repression factors. In the somatic cells, however, histone H1 is present, and the TFIIIA concentration decreases to an extremely low level. Under these conditions, there is competition not only for the limited transcription factors, but also for the formation of either a transcription complex or a repressive nucleosome on the 5S DNA. The four-fold weaker affinity of oocyte 5S DNA versus somatic results in a discrimination against forming an active transcription complex on the Xlo gene, and this discrimination becomes a repression

when the nucleosome dominantly forms on the unoccupied DNA. Once nucleosomes have formed, the genes become progressively repressed with the maturation of chromatin. The XIs genes remain active due to the stability of the transcription complex and its resistance to the chromatin packaging. This picture provides a simple explanation for the differential gene expression in the oocytic and somatic stages of *Xenopus* development. The simplicity of the picture becomes more complicated with the question: when the cell divides (DNA replicates) during embryogenesis, how can the differential gene expression pattern pass over to the daughter cells? Or to be more specific: What is the effect of DNA replication on this gene expression pattern?

1.6.4. The Effect of DNA Replication on Gene Expression

Cell memory is an extremely important biological feature. One example of cell memory is cells retaining their biological phenotype for long periods of time or after many rounds of cell divisions. The differential 5S RNA gene expression pattern is carried over in *Xenopus* embryogenesis during which a fertilized egg keeps dividing to form a multiple cell embryo. For simplicity the term "cell division" will be used interchangeably with "DNA replication". DNA is replicated at the S phase before each round of cell division. How are the active and repressed states of 5S genes maintained by the daughter DNA after replication? One explanation is that a transcription complex could play a role in such a memory if one or more of its components remained in place long enough after passage of a replication fork to nucleate the formation of a new transcription complex (Lin & Riggs, 1975; Wolffe, 1988). This possibility was tested and ruled out by a direct *in vitro* DNA replication experiment (Wolfe & Brown, 1986).

By using an *in vitro* DNA replication system that mimics *in vivo* DNA replication (Li & Kelly, 1984, 1985; Stillman & Gluzman, 1985; Wobbe *et al.*, 1985; Yamaguchi & De-Pamphilis, 1986), Wolffe and Brown demonstrated that the transcription complex is

disrupted by passage of the replication fork, specific transcription factors are displaced, and the daughter 5S RNA genes are inactivated. There is no evidence to show that any "memory" of the pre-existing transcription complex is transmitted to the daughter DNA duplexes following replication. The authors' conclusion is that the same transcription complexes that can withstand passage of RNA polymerase are erased by the replication fork at each round of cell division.

If the replication fork erases the pre-existing transcription complex and leaves no memory of the gene expression pattern, then the decision for the gene expression pattern must be made after each round of DNA replication. The transcription activity of the Xlo and Xls 5S RNA genes will thus be dependent on the relative abundance of the genes, the binding affinities of transcription factors, the kinetics of complex formation, the concentrations of transcription factors and histone proteins, and the timing of the interaction.

The replication of DNA in the S phase provides a brief interval in which both Xlo and Xls 5S RNA genes must face the competition of forming either an active transcription complex or a repressed nucleosome structure. Once the chromatin on the newly replicated DNA matures, the active or repressed states of 5S RNA genes will once again be established and maintained until the next round of DNA replication.

In somatic cells, a somatic 5S RNA gene forms a stable complex rapidly even at low levels of transcription factors. This stable complex not only activates the somatic 5S RNA gene, but its presence inhibits the competitive process of nucleosome assembly. In contrast, the relatively unstable oocyte 5S RNA gene complex requires continued high levels of factors to maintain activity. When the concentration of the key factor, TFIIIA, drops sharply before and during embryogenesis (Table 2), this gene is unable to form a transcription complex but instead is packaged into chromatin. Once unprogrammed oocyte 5S RNA genes become incorporated into chromatin, even high levels of factors cannot

reverse the repression (Andrews & Brown, 1987). The outcome of these competitive processes is that the oocyte 5S RNA genes that have been repressed before DNA replication will be repressed again, and the somatic 5S RNA genes that have been active will remain active after DNA replication. The differential gene expression pattern remains unchanged after each round of cell division, not by cell memory, but by the same decision being made following each round of DNA replication (Wolffe, 1988).

1.6.5. The Effect of DNA Replication Timing on Gene Expression

The timing of DNA replication plays another very important role in differential gene expression. Provided transcription factors are limited, the first replicated genes have the first opportunity to form active complexes. If the complexes formed are stable enough to resist the chromatin assembly, such as that formed on Xls 5S RNA genes, then the continued activity of the genes is assured until the next round of DNA replication. On the other hand, if the genes are replicated after the limited transcription factors have already been depleted by the genes replicated earlier, they will have less or no chance to form the active complexes. The timing of DNA replication alone can sometimes determine the active or inactive state of genes even when the genes are identical, such as in the case with yeast mating type determination (Nasmyth, 1987; Sternberg *et al.*, 1987).

In the S phase of the *Xenopus* cell cycle during embryogenesis, the somatic 5S RNA genes replicate earlier than the oocyte genes (Gilbert, 1986). The instability of transcription complexes on the oocyte 5S RNA genes may be enhanced by the decreasing free TFIIIA pool that is depleted by somatic genes replicated earlier (Gottesfeld & Bloomer, 1982). The temporal order of 5S RNA gene replication in the dividing *Xenopus* cells plays another crucial role in differential gene expression.

1.6.6. Maintenance of the Gene Expression Pattern.

The repression of oocyte 5S RNA genes in embryo cells is progressive with the maturation of chromatin. The stability of the repressed state of the genes *in vivo* is likely to vary depending upon the extent of compaction of the chromatin and the availability of histone H1 for exchange with the chromatin (McGhee & Felsenfeld, 1986; Thomas & Rees, 1983) (Fig.11). Chromatin maturation is a relatively slow, step-by-step process. Before histone H1 joins the nucleosome octamer and packages the chromatin into higher order structures, the repression is likely reversible in the presence of high levels of transcription factors. Once higher order structure has been established by histone H1, such as in the non-dividing, terminally differentiated somatic cells, the highly compact genes are "closed" to any transcription factor indefinitely.

On the other hand, the active complexes on the somatic genes are stable for a long period of time. Even in cells that no longer synthesize 5S RNA, such as the nucleated *Xenopus* erthyrocyte, a cell that has no detectable RNA polymerase III and no detectable free TFIIIA, somatic 5S RNA genes remain in stable transcription complexes for days and probably for weeks (Wolffe, 1988).

The differential gene expression pattern should reach its maximum stability in the non-dividing terminal somatic cells, because at this stage of development, the oocyte 5S RNA genes are deeply repressed in the fully matured chromatin, and the somatic genes remain active with no competition for the decreased level of TFIIIA (3,000 TFIIIA per cell). The low level of TFIIIA is not sufficient for displacing histone H1 from the well compacted oocyte 5S genes, but seems to be sufficient to keep the somatic 5S genes active. Furthermore, there will be no DNA replication to disrupt the established active and inactive states of these genes.

The above model for differential gene expression in *Xenopus* development is summarized in Figure 12.

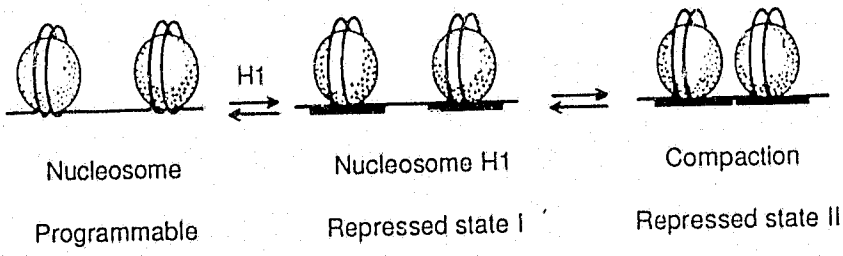


Figure 11: The repressed state.
A nucleosome consisting of DNA wrapped around an octamer of histone proteins does not inhibit the binding of transcription factors to the promoter (ICR) of a 5S RNA gene. However, when histone H1 (the solid bar) binds to nucleosomes and to part of the linker DNA between nucleosomes, the underlying DNA in this structure cannot be programmed into a transcription complex. The nucleosomes containing histone H1 interact, compacting the chromatin. This state of chromatin can occur when there are more than six nucleosomes with H1 in a row and is considered to be a more inaccessible or repressed structure, because histone H1 is much less likely to exchange. Maturation of chromatin after DNA replication proceeds from left to right (after Wolffe & Brown, 1988).

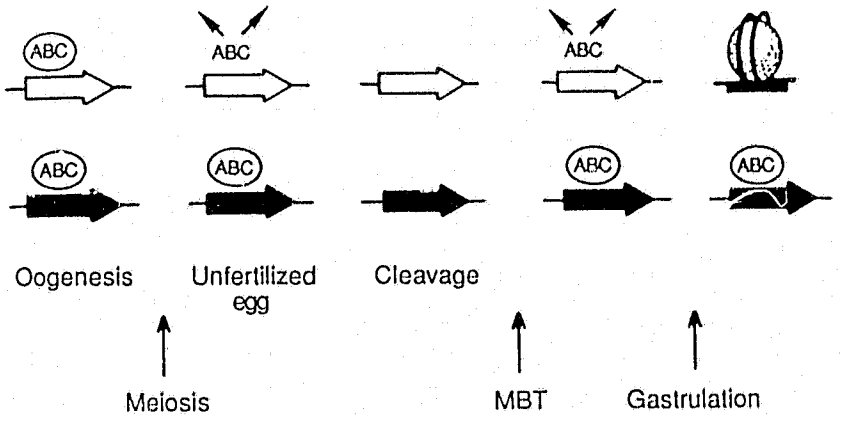


Figure 12: Model for developmental control of 5S RNA gene transcription.
This diagram summarizes the occupancy of oocyte (upper) and somatic (lower) 5S RNA gene with transcription complexes or chromatin during oogenesis and embryogenesis. A stable complex is represented by the factor A, B, and C encircled. The loss of the circle around a complex along with the arrows indicates unstable transcription complexes. These genes are still accessible to high levels of factors. The end result by late gastrulation is stable transcription complexes assembled on somatic 5S RNA genes and a repressed chromatin structure on oocyte 5S DNA. MBT: midblastula transition (after Wolffe & Brown, 1988).

1.6.7. Can TFIIIA Be a Repressor?

The 5S RNA gene expression regulation model discussed so far has been mainly established by Wolffe and coworkers (reviewed by Wolffe, 1988; Brown, 1984). Theirs is the best established and well supported model, but other possibilities still exist. For example, a second form of TFIIIA was identified and proposed to play a role in the regulation of 5S gene expression. This form of TFIIIA is 42 kD in size, in contrast to the well known 39 kD TFIIIA. The 42 kD is present in late stage embryos and in somatic cells (Pelham *et al.*, 1981), and may also exist in oocytes (Blanco *et al.*, 1989). The expression of the 42 kD TFIIIA is at a low level in immature oocytes (2-4% of the 39 kD TFIIIA level) and increases dramatically during oogenesis (Blanco *et al.*, 1989).

The relationship between the two types of TFIIIA molecules is unclear. They are structurally related, similar in charge and antigenicity, but differ in size and trypsin and V8 protease peptide mapping. The two TFIIIA species may be products of two different genes, or the result of alternative splicing of the TFIIIA primary mRNA, or a result of post-translational modification. However, the attempt to detect common post-synthetic modifications on the 42 kD TFIIIA has failed (Blanco *et al.*, 1989). Blanco *et al.* reported that the 42 kD TFIIIA binds to both oocyte and somatic 5S RNA genes with comparable affinities to those of the 39 kD TFIIIA, and that the DNase I footprints of the two proteins on 5S RNA genes are similar. However, the difference is that, unlike the 39 kD TFIIIA, the 42 kD protein does not support the transcription of oocyte-type 5S genes in a fractionated transcription system derived from mature oocytes. The observation that 42kD TFIIIA binds to both Xlo and Xls 5S RNA genes but activates transcription on only one of them led to an alternate model for the differential gene expression pattern. In this model the second form of TFIIIA may serve as an activator of somatic 5S RNA gene transcription and as a repressor of oocyte 5S RNA gene transcription during early embryogenesis (Blanco *et al.*, 1989).

A possible mechanism for the oocyte-specific repression by 42 kD TFIIIA was proposed by Blanco *et al.*. They suggested that the 42 kD TFIIIA disallows the subsequent productive binding of the other components of the transcription complex to the oocyte 5S gene. However, the same protein, as well as the 39 kD TFIIIA, allows the active complex to form on the somatic 5S gene. If this is true, the competition between the active 39 kD TFIIIA and the repressive 42 kD TFIIIA for binding the oocyte 5S RNA genes could play an important role in the initial inactivation of those genes, an additional regulation mechanism to the model established by Wolffe and coworkers.

CHAPTER 2
TFIIIA-5S RNA INTERACTION

2.1. INTRODUCTION

2.1.1. Characterization of TFIIIA-5S RNA Interaction

The binding properties of TFIIIA to 5S RNA were characterized by Romaniuk (1985). The protein binds *Xenopus* oocyte 5S RNA with an association constant of 1.4 x 10^9 M^{-1} in 0.1M salt, pH 7.5 at 20 °C, determined by a nitrocellulose filter binding assay. In contrast, the binding strength between *Xenopus* TFIIIA and the yeast $tRNA^{phe}$ is over a 100-fold weaker than the specific interaction of TFIIIA-5S RNA. On the other hand, *Xenopus* TFIIIA has a two-fold higher affinity for wheat germ 5S RNA and a four-fold lower affinity for *E. coli* 5S RNA.

A thermodynamic analysis of the interaction showed a $\Delta H° = -8.3$ Kcal/$mole^{-1}$ and $\Delta S° = +13.1$ cal•mol^{-1}•deg^{-1}. Approximately five cationic charges are released when TFIIIA binds to 5S RNA. The high affinity and relatively low ionic strength-dependence of the interaction signify that a relatively high proportion, approximately 68%, of the free energy of complex formation in 0.1 M KCl is contributed by non-electrostatic interactions between the protein and 5S RNA.

The binding of TFIIIA to 5S RNA appears to be unaffected by pH changes from 6.0 to 8.0, and the complex is stable at temperatures below 22 °C. Scatchard analysis indicated that the complex consists of one molecule of TFIIIA and one of 5S RNA. These TFIIIA-5S RNA interaction properties, as well as the Ka value, are similar to those of TFIIIA-5S DNA binding.

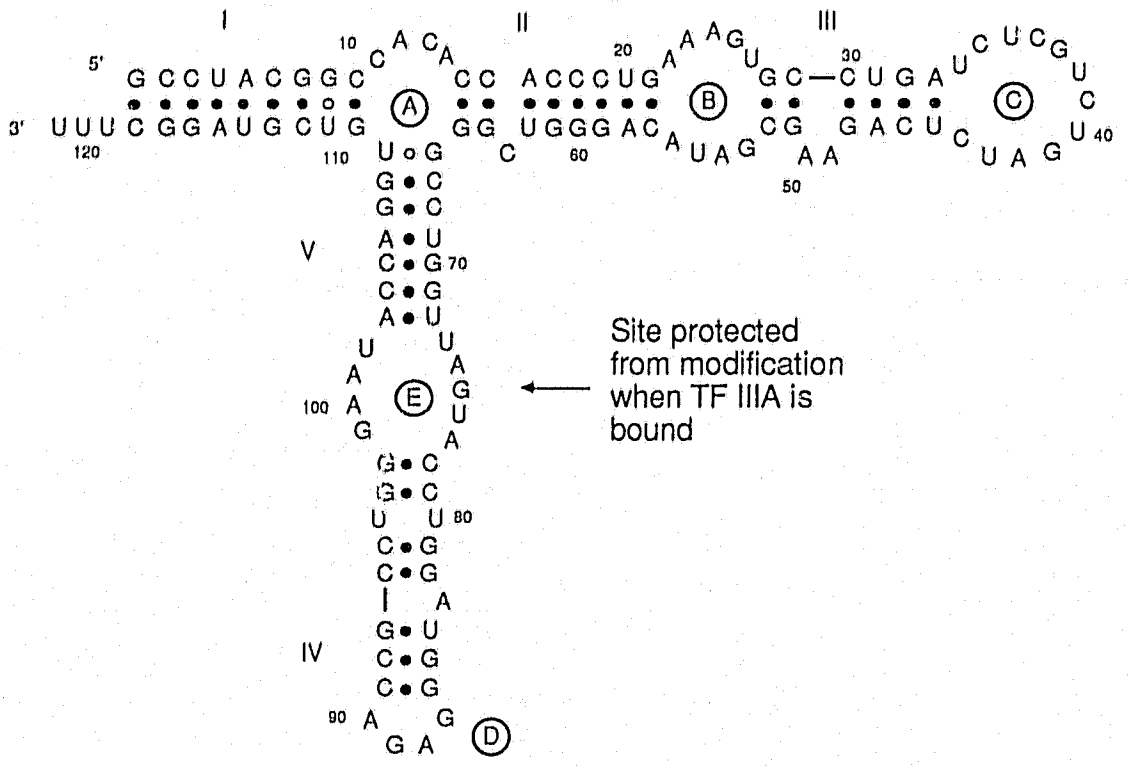
The measurement of TFIIIA-5S RNA complex dissociation kinetics showed that different populations of the complexes can be distinguished: a small population (10%) underwent extremely rapid dissociation, another 40% underwent dissociation with a half-life of approximately 45 minutes, and 50% of the complexes were extremely resistant to dissociation. In contrast, the TFIIIA-5S DNA complex population undergoes a steady dissociation with a half-life of only 6-15 minutes (see introduction in chapter 3). The formation of complexes with slow dissociation properties is consistent with the role of the 7S RNP as a storage particle for 5S RNA molecules in immature oocytes.

2.1.2. Probing TFIIIA Binding Site(s) on 5S RNA.

2.1.2.1. The general TFIIIA protection site of 5S RNA.

Extensive efforts have been made to investigate the TFIIIA-5S RNA interaction. RNase and chemical footprinting experiments provide useful information on the TFIIIA binding site. These methods are based on the protection that bound TFIIIA gives to the 5S RNA binding site. The RNA site(s) bound by the protein are inaccessible to the enzymes or chemical reagents, and therefore are not cleaved or modified. The data from those experiments delineated a general TFIIIA protection region, which covers the nucleotides from 53 to 110, including the helix II/loop B and helix IV/loop E/helix V structural domains of the 5S RNA (Romaniuk, 1985; Pieler & Erdmann, 1983; Anderson *et al.*, 1984; Huber & Wool, 1986; Christensen *et al.*, 1987) (Fig. 13).

In a more recent study (Darsillo & Huber, 1991), Hydroxyl radical footprinting and missing nucleoside techniques were applied to investigate the *Xenopus* TFIIIA-5S RNA interaction. Based on the data from these experiments, Darsillo and Huber conclude that (1), the secondary structure, rather than the sequence, is the principal determinant for



Figgure 13: The general TFIIIA protection site on 5S RNA.

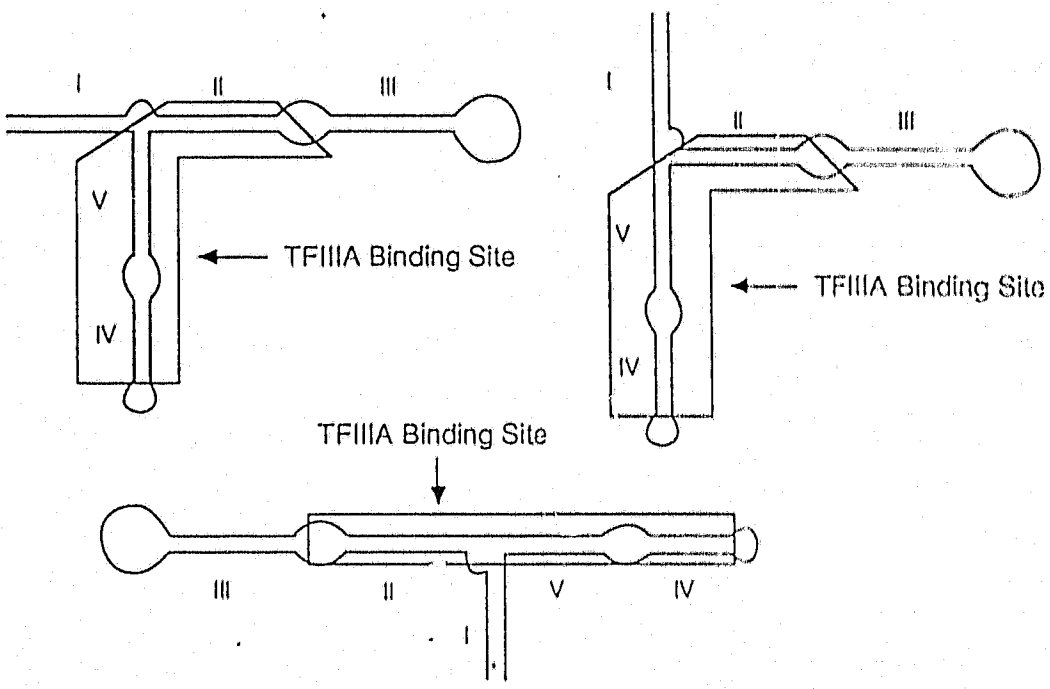


Figure 14: Possible coaxial stacking arrangements of *Xenopus* 5S RNA, indicating the position of the TFIIIA binding site on each (after Romaniuk, 1989).

TFIIIA binding; (2), the protein protects a substantial amount of the 5S RNA molecule; and (3), the stem IV-loop E-stem V region provides the most important contact sites for the factor.

It is interesting to notice that this TFIIIA protection site of 5S RNA is very similar to the ICR region of 5S DNA (bp +45 to +97). Recall the proposal that the ICR region of 5S DNA may have an A form helical structure that is similar to the RNA helix, one may ask whether TFIIIA recognizes the same sequences in the 5S DNA and the 5S RNA. Indeed, structural analysis of 5S RNA indicates an alternative secondary 5S RNA structure which occurs at a "hinge" region in loop A, turning the TFIIIA binding site into a successive linear region (Fig. 14). Other studies (Westhof *et al.*, 1989; Stevenson *et al*, 1991) demonstrated that loop E in this region is actually closer to helical structure than to loop. Three dimensional mapping techniques have been applied to the analysis of the chemical footprinting data and the results suggest that the contacts made with TFIIIA are clustered in three specific areas of the 5S RNA, and are similar in distribution to the three elements of the ICR on the gene (Christiansen *et al.*, 1987). Thus, these results support the view that TFIIIA may bind to 5S RNA and its gene in a similar fashion.

In order to ascertain exactly which elements of 5S RNA structure and sequence are required for TFIIIA binding, various mutant 5S RNA molecules have been created and their affinities for TFIIIA measured with a filter binding assay developed in our laboratory. By using 3' and 5' truncated 5S RNA mutants, it was found that nucleotides 11 to 108 of intact 5S RNA provide the necessary sequence and conformational information required for TFIIIA binding (Romaniuk *et al.*, 1987a). TFIIIA binding is more sensitive to the deletion of nucleotides from the 5' terminus of 5S RNA as opposed to the 3' terminus. This is not surprising because it is known that the somatic 5S RNA has a three-fold higher affinity for TFIIIA over the oocyte 5S RNA, and the increased somatic affinity is conferred by nucleotide substitutions in the 5' half of the molecule (Romaniuk *et al.*, 1987a).

2.1.2.2. Nucleotide substitutions in the loops.

Nucleotide substitutions within the single stranded loop regions were shown to have little or no effect on TFIIIA binding with the exception of nucleotide substitutions in loop A, which lies outside the TFIIIA footprint region (Romaniuk, 1989). The results indicate that these sequence elements in the loops are not important for TFIIIA recognition. The obvious negative effect observed in loop A substitution mutants was explained by the conformational change of 5S RNA caused by the substitution rather than by the change of sequence itself. As pointed out by Christiasen *et al.*(1987), it has long been thought that loop A forms a hinge that controls the coaxial stacking of the three helical domains of the 5S RNA via the formation of non-canonical hydrogen bonding between one nucleotide in loop A and the highly conserved G at position 66. One of the possible co-axial stacking arrangements would provide a co-linear arrangement of the TFIIIA binding site on 5S RNA by stacking the helix IV-V domain on the helix II-III domain (Fig. 14). Perhaps the substitution of the nucleotides in loop A reduces the stability of this coaxial stacking arrangement, therefore reducing the affinity for TFIIIA (Romaniuk, 1989).

The fact that most loop mutations did not affect TFIIIA binding was surprising because the majority of nucleotides that are highly conserved among all eucaryotic 5S RNAs are found in the single stranded loops, and would therefore be excellent candidates for the formation of sequence specific protein-RNA contacts that would explain the similar affinity that TFIIIA has for a variety of eucaryotic 5S RNAs.

2.1.2.3. Deletions of the bulged nucleotides.

The bulged structure has been considered to a highly conserved structural feature among all eucaryotic 5S RNAs. Therefore those bulged nucleotides, the nucleotides that stack out of a double helix, are expected to offer unique opportunities for the formation of bonding contacts with the amino acid side chains of proteins. In fact, the bulged

nucleotides have been shown to be essential in *E. coli* ribosomal protein L18-5S RNA interaction (Peattie *et al.*, 1981) and other protein-RNA interactions (Romaniuk *et al.*, 1987; Mougel *et al.*, 1987). The importance of bulged nucleotides in the *Xenopus laevis* TFIIIA-5S RNA interaction was tested by Baudin and Romaniuk (1989). Mutant 5S RNA genes were created to delete the bulged nucleotides at positions A_{49-50} (ΔA49-50), C_{63}(ΔC63) and A_{83}(ΔA83) of the 5S RNA. Mutants ΔC63 and ΔA83 are located in the 5S RNA site that is protected by TFIIIA, and ΔA49-50 is close to the 5' boundary of the TFIIIA protection site. TFIIIA binding affinities for both the DNA and RNA mutants were determined by filter binding assays. It was found that the deletions of bulged nucleotides have no effect on the TFIIIA binding to 5S RNA. Only one DNA deletion mutant, ΔA83, which is located in box C of the ICR of 5S DNA, showed a four-fold decrease in binding affinity compared to the wild-type 5S DNA. Thus, the bulged nucleotides are not important in the TFIIIA-5S RNA interaction. The fact that the ΔA83 mutation has no effect on the TFIIIA binding to 5S RNA but greatly reduces the binding of TFIIIA to 5S DNA does not support the proposal that the protein may bind to both nucleic acids in the same fashion.

2.1.2.4. Disrupting 5S RNA helical structures

The nucleotides on loops and bulges of 5S RNA did not show any significant effects on TFIIIA binding (except loop A), and since several studies have suggested that the overlapping region of TFIIIA binding on 5S RNA could co-axially stack into a continuous helix similar to the ICR of the gene, the stem domains of *Xenopus laevis* oocyte 5S RNA became the last structural target for investigation. TFIIIA binds to 5S DNA by making sequence-specific contacts with the base pairs. Is the binding of the protein to 5S RNA also dependent on specific base pairs in the RNA double helical region or, alternately, are the double helical structures of 5S RNA even required for TFIIIA

binding? To investigate these questions, twenty-five different mutations with altered sequence and base pairing properties of the double helical stems II, III, IV and V of 5S RNA were constructed, and the effects of these mutations on the TFIIIA binding affinity for the RNA determined. Stem region I, which is located outside of the TFIIIA protection site and has been studied with the truncation mutants (see above), was not included in this mutagenesis analysis.

By testing the TFIIIA affinity for these 5S RNA mutants using the filter binding assay, the effects of nucleotide substitutions and base pairing disruptions can be distinguished (see Result and Discussion sections of this chapter). In the case of stems II and V, results showed that full TFIIIA binding affinity requires the presence of the double helical conformation, but is sequence independent. In other words, specific base pairs do not play an important role in the TFIIIA-5S RNA interaction, as they do in the TFIIIA-5S DNA interaction.

2.1.2.5. Substitution mutations at the somatic specific sites.

Somatic 5S RNA has a three-fold higher binding affinity for TFIIIA than the oocyte type 5S RNA (Romaniuk *et al.*, 1989). There are six nucleotides that differ between these two 5S RNA species. Three of them are located in the stem regions of 5S RNA, and the other three are clustered in loop B (position 53, 55 and 56). We created a set of nine oocytic mutants by substituting the nucleotides at somatic specific positions clustered in loop B, three mutants for each position (eg. C_{53}, A_{53} and U_{53} were created at G_{53}). The binding of TFIIIA to these mutants was determined. None of these substitution mutants showed any significant reduction in TFIIIA binding. The picture which emerged from all the above-mentioned data showed that unlike the TFIIIA-5S DNA interaction, TFIIIA binds to 5S RNA via many weak contacts, and requires an intact 5S RNA conformation rather than a particular sequence (See following sections).

2.2. MATERIALS AND METHODS

2.2.1. Materials

Xenopus laevis : The young female frogs (six month or younger) were purchased from Ann Arbor, MI, USA.

Enzymes: T4 DNA ligase, T4 DNA kinase, restriction enzymes and RNasin were purchased from Pharmacia; T4 RNA ligase was purchased from New England BioLabs; T7 RNA polymerase was purified from *E. coli* strain BL21/pAR1219 by a published procedure (Davanloo *et al.*, 1984).

Chemicals: Commonly used chemicals were purchased from BDH inc., Fisher Scientific Company or Sigma; Urea, acrylamide and bis-acrylamide were purchased from SERVA; Tris base and $tRNA^{phe}$ were purchased from Boehringer Mannheim.

Cell culture: Yeast extract, tryptone and agar were purchased from BBL or BDH.

Nucleotides and plasmid: ATP, GTP, CTP, UTP (NTPs) were purchased from Pharmacia; dNTPs and ddNTPs were also from the same company; Oligonucleotides were synthesized by Dr.Romaniuk in a Biosearch 8600 DNA synthesizer using the phosphoramidite method (Sinha *et al.*, 1984).

Others: Bio-Rex 70 and DE-52 columns for TFIIIA purification were purchased from BIORAD Laboratories; Filters for the binding assays were purchased from Millipore corporation; XAR-5 film was from Kodak; Agarose and low melting temperature agarose were purchased from FMC Bioproduct; [α-^{32}P] GTP (600 Ci/mmol) and [5'-^{32}P] pCp (3000 Ci/mmol) isotopes for RNA labelling were purchased from New England Nuclear (Du Pont).

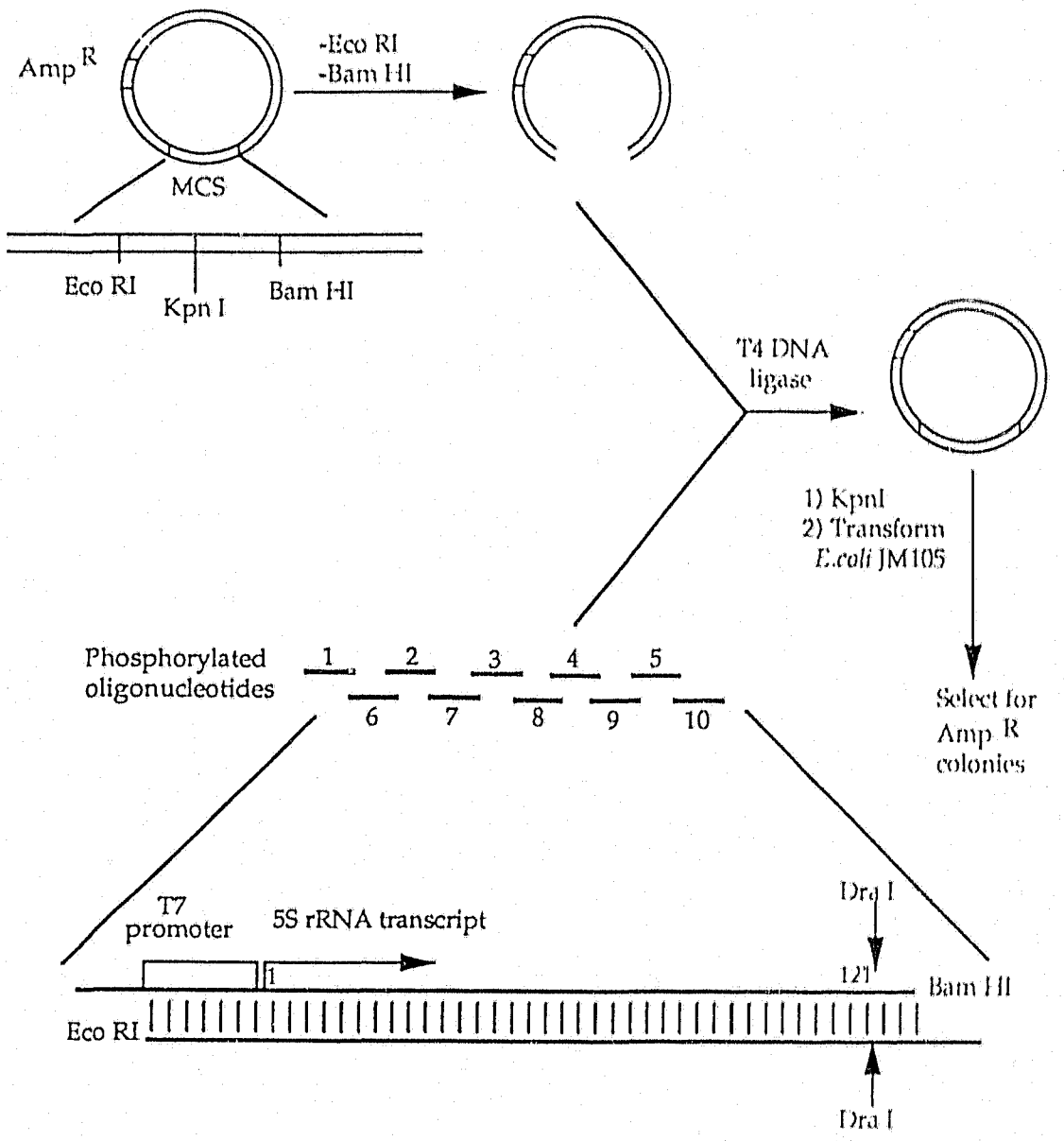


Figure 15: Construction of mutant 5S RNA genes using microscale shotgun ligation method. The oligonucleotides contain sequences that generated both Eco R1 and Bam H1 restriction sites and the T7 promoter. MCS: Multiple Cloning Site; Amp R: Ampicillin Resistance (adapted from Romaniuk *et al.*, 1987a).

2.2.2. Methods

2.2.2.1. Construction of mutant genes

Overall strategy The construction of mutant 5S RNA genes is shown in Figure 15. The method is a modification of oligonucleotide-directed mutagenesis by micro-scale ligation (Gründstrom *et al*, 1985). The cloned genes were synthesized as a series of oligonucleotides. For each mutant the appropriate oligonucleotides for both strands were synthesized and substituted for the corresponding wild type oligonucleotides in a shotgun ligation reaction. These oligonucleotides were subsequently annealed and ligated into vector pUC18. The mutant 5S RNA genes were placed under the control of the T7 promoter. An Eco R1 linker at the 5'-end of the T7 promoter and a Bam H1 linker at the 3'-end of the gene were introduced after ligation, the reaction mixture was digested with Kpn 1 to linearize any re-ligated pUC18 vector with no insertion of the gene.

Deprotection and purification of oligonucleotides Synthesized oligonucleotides were deprotected by incubation in 1.0 ml of concentrated ammonium hydroxide solution, first at room temperature for one hour and then at 50 °C overnight. After this incubation, the ammonium hydroxide solution was evaporated on a RH40-11 Speed Vac concentrator. The pellet was dissolved in 100 µl dH_2O and the oligonucleotide concentration determined by absorbance at 260 nm. Ten O.D. units of oligonucleotides were purified by polyacrylamide (20%) gel electrophoresis and reverse phase chromatography using a C18 Sep-Pak column (Atkinson & Smith, 1984).

Five µg of oligonucleotides were phosphorylated by incubating at 37 °C for 60 minutes in 60 µl of a buffer containing 50 mM HEPES pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 50 µg/ml BSA, 0.1 mM ATP and 3 units of T4 polynucleotide kinase. The reaction was then diluted with 1.0 ml TE buffer (10 mM Tris:HCl pH 8.0, 1.0 mM EDTA). The oligonucleotides were purified by reverse phase-chromatography and stored as 10 µM solution at -20 °C.

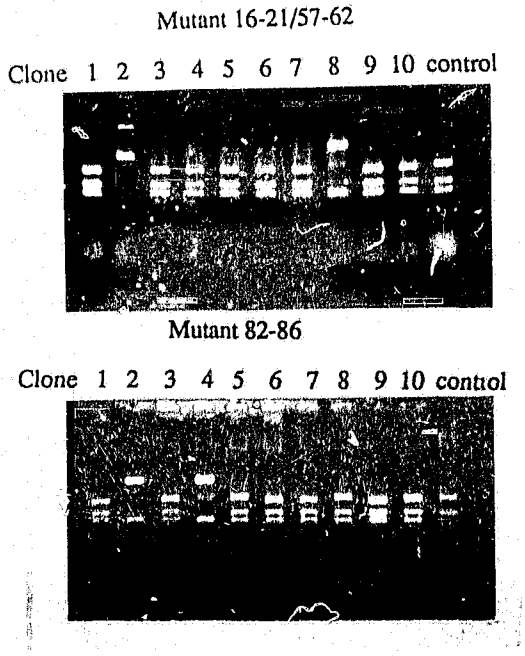


Figure 16: Screening of cloned 5S RNA genes
Plasmid mini-preparations from transformed clones were digested with restriction enzyme Dra I. Plasmids with 5S RNA gene insertions contain an extra Dra I site and show three bands on the gel. The control plasmid contains w.t. 5S RNA gene which has been confirmed by DNA sequencing.

Ligation reaction The appropriate phosphorylated and purified oligonucleotides (0.5 pmol) were annealed by incubating in 8 µl of 50 mM Tris:HCl pH 7.5, 10 mM $MgCl_2$, 20 mM DTT, 0.1 mM spermidine and 1.0 mM ATP, at 37°C for 60 minutes. The pUC vector, which had been previously digested with Eco R1 and Bam H1, was added to the annealing mixture together with 200 units of T4 DNA ligase. The ligation reaction was complete after overnight incubation at room temperature. The reaction was heated at 70 °C for 10 minutes to inactivate T4 DNA ligase and then 20 µl of Kpn 1 buffer (10 mM Tris:HCl pH 7.5, 10 mM $MgCl_2$, 1.0 mM DTT, 0.01% Triton X-100) was added together with 10 units of the restriction nuclease Kpn 1 to linearize vectors which had re-ligated to the multiple cloning site rather than the desired insert.

Transformation of *E. coli* cells and screening for cloned genes The ligation reaction was used to transform *E. coli* strain JM105, made competent by $CaCl_2$ treatment (Maniatis *et al.*, 1982). Cells that contained the plasmid DNA were selected on the basis of ampicillin resistance. For screening purposes, mini-preparations of plasmid DNA were made from 20 colonies according to Maniatis *et al.* (1982). The DNA from these preparations was analyzed both for the presence of a new Dra 1 restriction site (Fig. 16), and the ability to produce 5S RNA transcripts. After this analysis, the correct sequence of each clone was verified by dideoxynucleotide sequencing of plasmid DNA using the M13 reverse sequencing primer and reverse transcriptase (Zaug *et al.*, 1984). In general, 40-70% of the plasmids from mini-preparations contained the correct sequence.

Construction of somatic specific mutants The somatic-specific substitution mutant genes were constructed by a different procedure, which is based upon a modified single primer method developed by Eckstein's group (Taylor *et al.*, 1985a; Taylor *et al.*, 1985b; Nakamaye *et al.*, 1986). The WT 5S RNA gene was inserted into phage M13mp18. Synthetic oligonucleotide primers containing nucleotide substitutions at the somatic specific positions (+53, +55 and +56) were used for the site-directed mutagenesis (according to a

protocol from Amersham Corporation). For each mutant, phage were grown from twenty plaques randomly chosen from the resulting transformation, and then screened by dot blot hybridization with the mutant oligonucleotide under stringent conditions. Each mutant was verified by dideoxynucleotide sequencing of the double stranded phage DNA.

2.2.2.2. Large-scale preparation of plasmid DNA

The large-scale DNA preparation protocol was a modification of a Promega Biotec procedure. *E. coli* cells were grown in 250 ml of LB-Ampicillin medium (0.01g/ml tryptone, 0.005g/ml yeast extract, 0.01g/ml NaCl, pH to 7.0, + Amp. at 100 μg/ml) for over night at 37 °C, with shaking at 300 rpm. The culture was then centrifuged at 3,800 x g in a JA-14 rotor (Beckman). The pellet was resuspended in 6 ml of freshly prepared lysis buffer: 25 mM Tris:HCl pH 7.5, 10 mM EDTA, 15% sucrose, 2 mg/ml lysozyme. After 20 minute incubation on ice, 12 ml of 0.2 M NaOH/1% SDS was added. Incubation continued for 10 minutes on ice and then 7.5 ml of 3 M NaoAc pH 4.6 was added. Incubation on ice continued for 20 minutes. The lysis solution was centrifuged in a JA-20 rotor at 39,000 x g for 25 minutes at 4 °C. The supernatant was saved and 50 μl of 1.0 mg/ml RNase A was added to digest the RNA at 37 °C for 20 minutes. The solution was then extracted twice with phenol: chloroform, followed by a chloroform extraction. DNA was pelleted by ethanol precipitation and resuspended in 1.6 ml dH_2O, plus 0.4 ml 4 M NaCl. 2 ml of 13% PEG 8000 was added, and the solution was incubated on ice for 60 minutes. After the incubation, the solution was centrifuged at 39,000 x g for 20 minutes to pellet the plasmid DNA. The pellet was washed twice with 70% ethanol and was resuspended in 300 μl of dH_2O or TE buffer. The typical yield was 200-300 μg of pure plasmid.

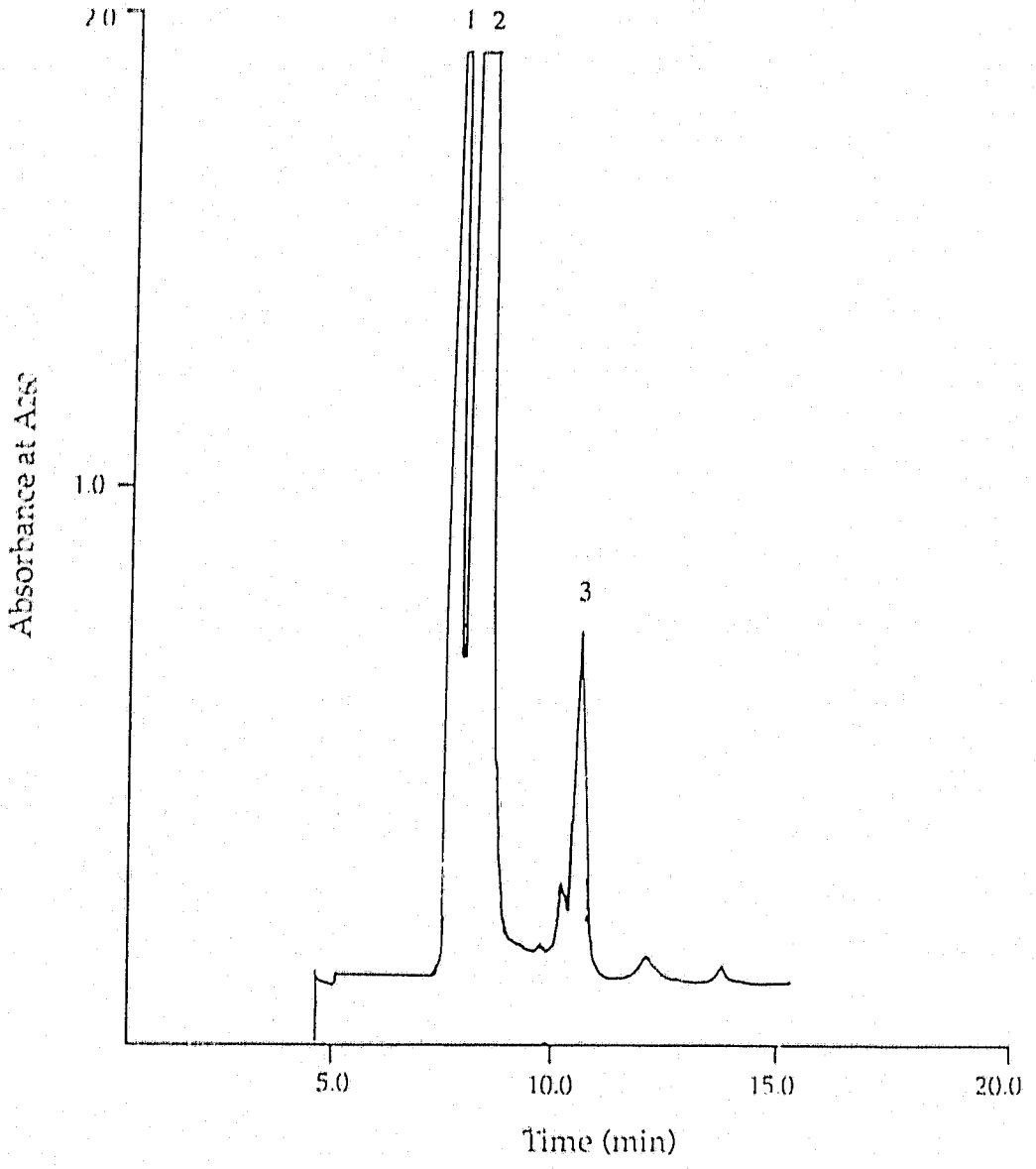


Figure 17: Elution profile of mutant Xlo 67-70/105-108 by gel permeation HPLC. Legend: (1) plasmid DNA; (2) Xlo 67-7-/105-108 5S RNA; (3) unincorporated nucleotide triphosphates.

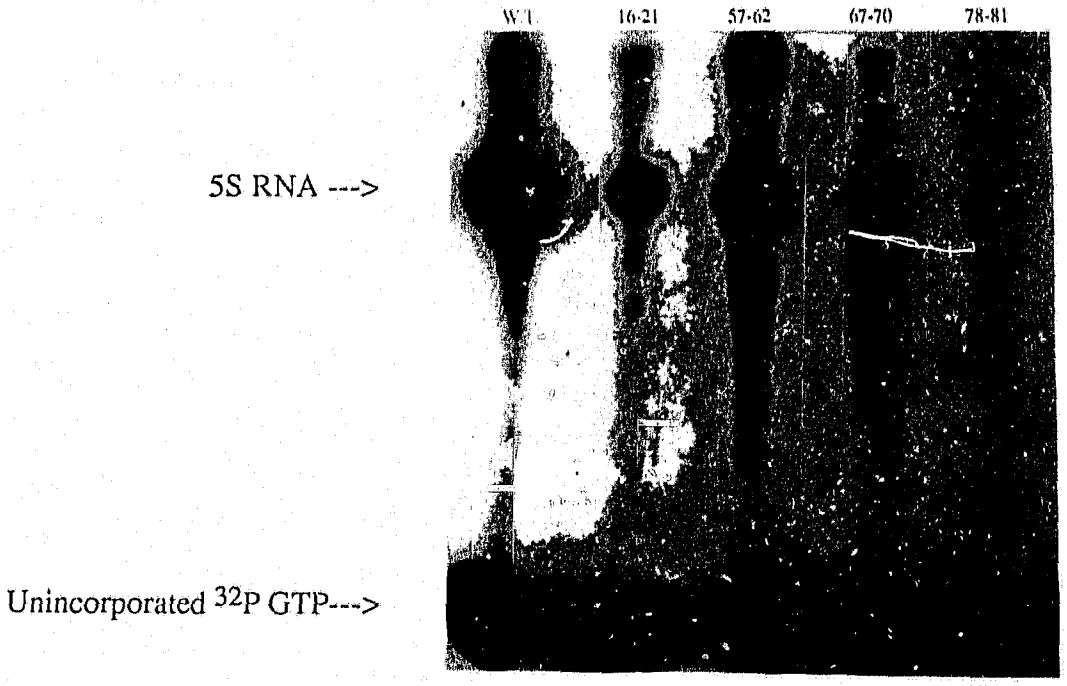


Figure 18: Purification of internally labelled 5S RNA on acrylamide gel.

2.2.2.3. Synthesis and labelling of mutant 5S RNA

In vitro transcription was carried out using T7 RNA polymerase purified by a published procedure (Davanloo *et al.*, 1984) from *E. coli* strain BL21/pAR1219, kindly provided by Dr. F.W. Studier. Prior to transcription, each plasmid containing a 5S RNA gene was digested with the restriction enzyme Dra I, which defines the 3' terminus of the transcripts as nucleotide +121 of the gene. Unlabeled RNA was synthesized in an assay mixture containing: 40 mM Tris:HCl pH 8, 30 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine, 100 µg/ml BSA, 8% PEG8000, 0.01% Triton X-100, 5 mM each NTP, 20 µg of linearized template DNA and 10 µg T7 RNA polymerase in a final volume of 200 µl. After incubation for 4 h at 37 °C, the reaction was extracted with 200 µl of phenol:chloroform, followed by extraction with 200 µl chloroform. The crude RNA was recovered by ethanol precipitation, and purified by gel permeation HPLC (Romaniuk *et al.*, 1987) (Fig. 17). Yields of pure 5S RNA were typically 100 µg.

For the direct binding assay, internally labeled RNA was synthesized in a mixture containing 40 mM Tris:HCl pH 8, 15 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine, 100 µg/ml BSA, 1000 U/ml RNasin, 0.5 mM each ATP, CTP, UTP, 0.0125 mM GTP, 50 µCi [α-^{32}P] GTP (600 Ci/mmol), 1 µg of linearized template DNA, and 0.6 µg T7 RNA polymerase in a final volume of 10 µl. After incubation for 2.5 h at 37 °C, 10 µl of urea-dye sample buffer was added, and the 5S RNA was purified on a 8 M urea, 12% polyacrylamide gel (Fig. 18).

The 3' end-labelling was carried out in a 25 µl reaction containing 50 mM Tris:HCl pH 7.8, 6 mM $MgCl_2$, 20 mM DTT, 5 µM ATP, 6 µl [5'-^{32}P] pCp, 3 µg RNA and 2000 U/ml T4 RNA ligase. The reaction was incubated at 4 °C overnight. Labelled RNA was precipitated with ethanol using $tRNA^{phe}$ as carrier. The RNA was then purified as described above.

2.2.2.4. Preparation of TFIIIA

Homogenization of the ovaries The 7S RNP particle was isolated from ovaries of immature *Xenopus laevis* (*Xenopus* I, Ann Arbor, MI) by a procedure developed by Pelham and Brown (1980), Hanas *et al.*(1983), Pieler and Erdmann (1983) and Romaniuk (1985). Buffers used for the purification were as follows: Homogenization buffer, 50 mM Tris:HCl pH 7.5, 25 mM KCl, 5 mM $MgCl_2$, 0.25 mM DTT; DEAE-50 buffer, 20 mM Tris:HCl pH 7.5, 50 mM KCl, 1.5 mM $MgCl_2$, 1.0 mM DTT; DEAE-160 and DEAE-320 buffers were the same as DEAE-50, except that the KCl concentrations were 160 mM and 320 mM, respectively.

Ovaries were removed from young *Xenopus* (ca. 5 cm long) and rinsed with ice cold homogenization buffer. The ovaries were then homogenized in 5 g amounts with the addition of 1.0 ml of homogenization buffer. After homogenization, the suspension was centrifuged for 15 minutes at 15,000 x g at 4 °C. The supernatant was pooled, taking care to avoid the fat layer on top. The supernatant was then layered onto 10-30% glycerol gradients made with homogenization buffer. The gradients were centrifuged for 20 hours at 40,000 rpm in a SW41 rotor at 4 °C.

Fractionation of the gradients Gradients were pumped out in the cold room, and 0.5 ml fractions were collected. The position of the 7S RNP in the first gradient was determined by non-denaturing polyacrylamide gel electrophoresis. 5 µl aliquots of every second fraction were loaded onto a mini-gel composed of 10% acrylamide in 0.3 x TBE buffer. 1.0 µg of 5S RNA in a glycerol-dye sample buffer was loaded as a marker. The gel was run at 300 V at 4 °C until the Bromophenol blue dye was at the bottom of the gel. The gel was stained by soaking for 20 minutes in 1µg/ml Ethidium bromide in 0.3 x TBE. The position of RNA bands was detected by UV light (Fig. 19). Using the gel as a guide, the 7S RNP peak was pooled from the appropriate fractions from each gradient. EDTA was added to a final concentration of 3.5 mM.

Isolation of 7S RNP by DE-52 column The pooled 7S RNP was loaded onto a 1.0 ml DE-52 column that had been equilibrated with DEAE-50 buffer. The column was washed with 10 ml of DEAE-50 buffer, followed by 10 ml of DEAE-160 buffer and 10 ml of DEAE-320 buffer. 1.0 ml fractions were collected. The 7S RNP was normally eluted in the first 3-4 ml of DEAE-320 buffer. The concentration of 7S RNP was determined by O.D. readings at 260 nm. The purity was checked by a SDS mini-gel and a 10% non-denaturing acrylamide gel (Fig. 20). The 7S RNP could be stored in DEAE-320 buffer for approximately one month at 0 °C.

Separation of TFIIIA from 7S RNP Two mini Bio-Rex 70 ion exchange columns (100 µl volume) were pre-eqilibrated with buffer A-0.1 M KCl (50 mM HEPES:KOH pH 7.5, 5 mM $MgCl_2$, 1.0 mM DTT, 10 µM $ZnCl_2$, 0.1 M KCl and 20% glycerol). The 7S RNA solution (ca. 10 O.D unit) was diluted with two volumes of buffer A-0 KCl. The sample was first loaded onto the Bio-Rex 70 column, to remove any inactive free TFIIIA. The 7S RNP solution that had passed through the column was then digested with 2 µl of 10 mg/ml RNase A for 1 hour at 20 °C. The digested solution was loaded onto a second Bio-Rex 70 column. The column was washed with buffer A containing 0.1 M, and then 0.5 M KCl. TFIIIA was eluted with buffer A-1.0 M KCl. TFIIIA was collected in 50 µl fractions. The binding activity of TFIIIA in buffer A remained approximately one week at 0 °C, without significant decrease.

Protein concentrations were determined by the Bradford method (1976) and the purity was checked by SDS-PAGE (Fig. 21). Fractional 5S RNA binding activity of each protein preparation was determined by comparing the apparent association constant (K_a) measured for the preparation with the value of 1.3×10^9 M^{-1} determined by Scatchard analysis to be the apparent K_a obtained with 100% active TFIIIA and *Xenopus* oocyte 5S RNA (Romaniuk, 1985). Only those preparations which were >90% active were used to study the binding of mutant 5S RNAs.

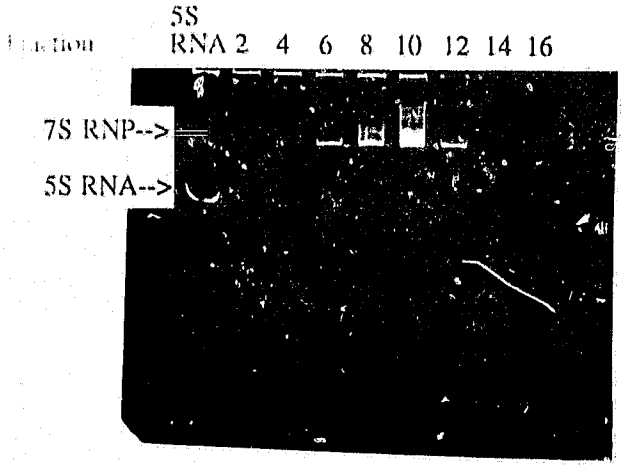


Figure 19: Fractions of 7S RNP after gradient centrifugation. Fraction #6 to #12 were pooled for further procedure.

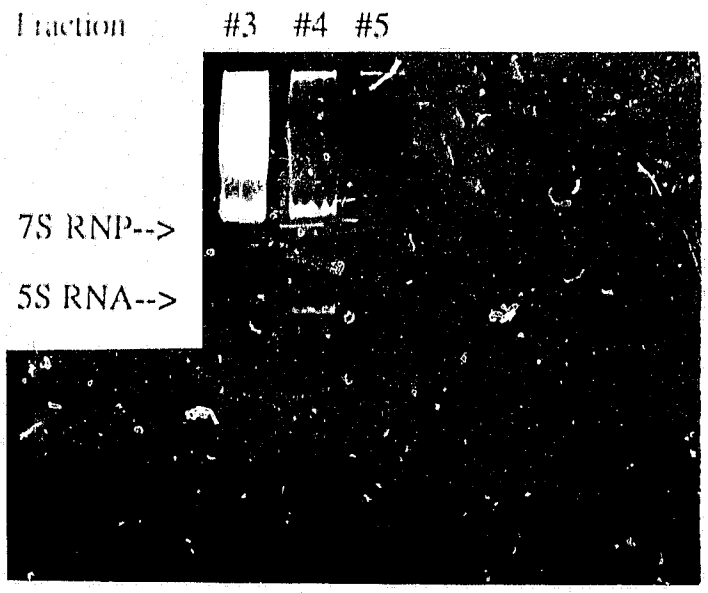


Figure 20: Analysis of purified 7S RNP on non-denaturing acrylamide gel.
7S RNP was purified using a DE-52 column. Fractions #3 and #4 were used for TFIIIA purification.

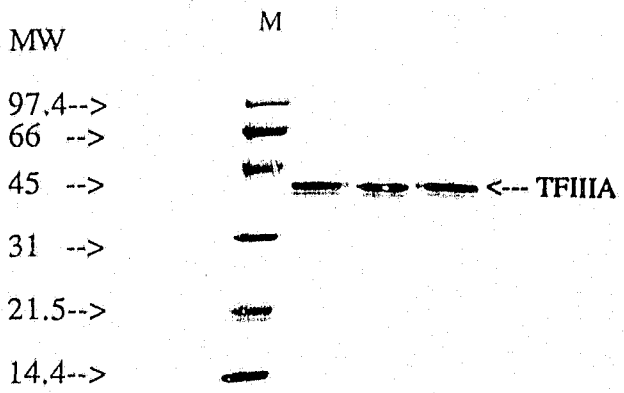


Figure 21: Analysis of purified TFIIIA on SDS PAGE.
Xenopus TFIIIA after Bio-Rex 70 column, showing TFIIIA collected from fractions #2,#3 and #4.

2.2.2.5. Filter binding assays

The standard TMK buffer for the filter binding assays contained 20 mM Tris:HCl, 5 mM $MgCl_2$, 100 mM KCl, 100 µg/ml BSA, 1.0 mM DTT, adjusted to pH 7.5 at the incubation temperature. TFIIIA was serially diluted in 180 µl of TMK buffer to give final concentrations ranging from 0.12 nM to 30 nM and equilibrated for 10 minutes at room temperature. The assay was started by the addition of 20 µl of end-labelled RNA (3-5 nCi, ca. 0.1-0.5 nM final RNA concentration) and allowed to equilibrate for 15 minutes. A 180 µl aliquot was then removed and filtered through a nitrocellulose filter pre-soaked in TMK which was then dried and counted in a toluene based scintillant. Retention of free RNA on the filter was typically 5-10% of input and this value was used to correct measurements of complex formation.

For competition assays, labelled 5S RNA (ca. 0.5 nM) was mixed with unlabelled competitor RNAs at the indicated concentrations. The assay was started by the addition of TFIIIA to a final concentration of 3 nM and after 15 minute incubation at room temperature, aliquots were withdrawn and filtered. The filter was then dried and the radioactivity counted as described.

2.3. RESULTS

2.3.1. Selection of Mutation Sites

Experiments were designed to investigate the contribution of the double stranded stems of 5S RNA to the free energy of TFIIIA binding. Figure 22 shows the secondary structure of 5S RNA, and summarizes the results of footprinting experiments (Pieler & erdmann, 1983; Andersen *et al.*, 1984; Romaniuk, 1985; Huber & Wool, 1986; Christiansen *et al.*, 1987) and the effect that site-specific mutation of single stranded nucleotides had on TFIIIA binding affinity (Romaniuk, 1989; Baudin & Romaniuk, 1989). Substitution of the nucleotides in stems II, III, IV, and V was accomplished by the creation

of a set of 25 mutant 5S RNAs. Stem I, which is located outside of the TFIIIA binding site, has been shown by deletion mutagenesis to be unimportant for TFIIIA binding (Romaniuk *et al.*, 1987) and was therefore not included in this analysis.

So called "single mutants" were created by substituting a contiguous stretch of two to four nucleotides on one strand of a double helix. Corresponding mutants were also created which altered the sequence of the opposite strand of the double helical stem. The resulting mismatch in each mutant results in the formation of single stranded regions within the target helix. "Double mutants" were obtained by combining two single mutants, substituting complementary nucleotides on both strands of the helix to restore the helical stem structure, but introduce a new sequence of base pairs (Fig. 23). Three 5S RNAs with multiple mutations were created by combining the 16-21, 67-70, 95-98, 57-62 and 78-81 substitutions in several ways (Fig. 24) to test the additive effects on TFIIIA binding affinity.

Two loops were also selected for mutagenesis. Mutant 22-25 changed the nucleotides in the 5' half of loop B so they would Watson-Crick base pair with the nucleotides on the opposite side of the loop. This mutation creates a long stem region that includes helix II, loop B, and helix III. Nucleotides 33-36 were substituted to be complementary to nucleotides 41-44 by replacing bases U and C at positions 33 and 34, respectively. The resulting mutant (Xlo 33-34) extended stem III into loop C (Fig. 23). In summary, all of the single mutations studied altered both sequence and conformation, while the double mutations resulted only in changes in the sequence of the 5S RNA.

2.3.2. Determination of the TFIIIA Binding Affinities of Mutant 5S RNAs

The affinity of each mutant 5S RNA for TFIIIA was measured using a nitrocellulose filter binding assay, in which a constant concentration of labeled 5S RNA

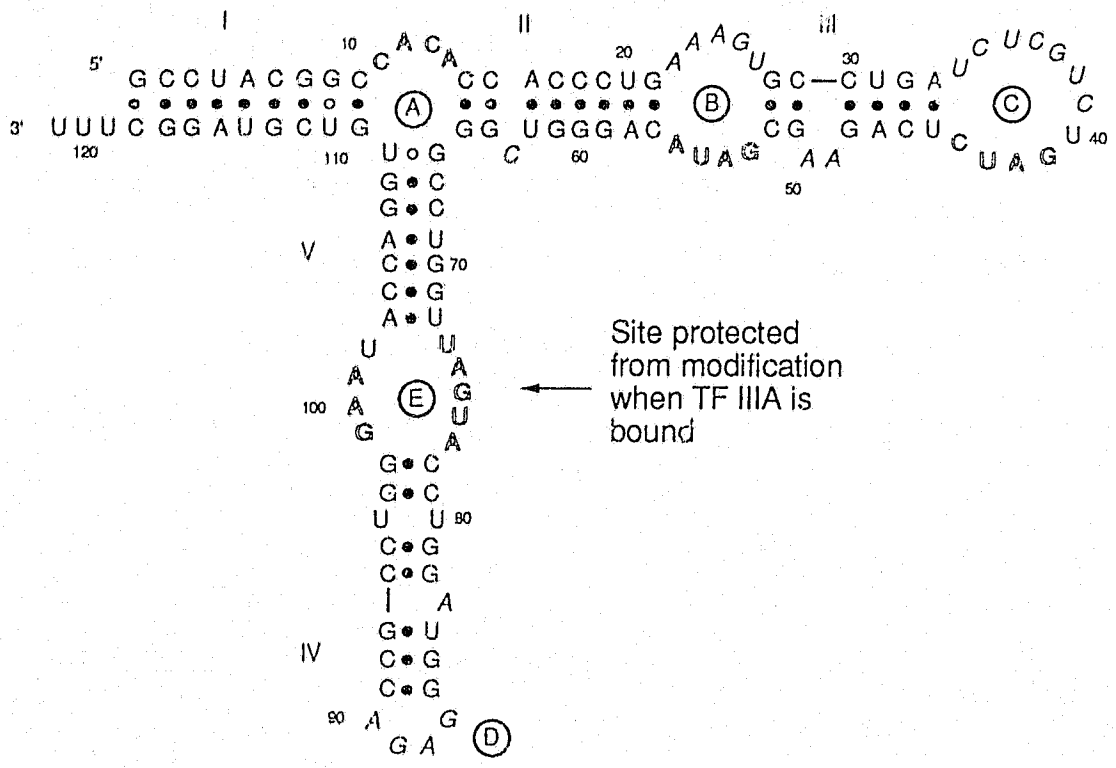


Figure 22: Secondary structure of *Xenopus laevis* oocyte 5S RNA. Shaded area represents the region protected from modification when TFIIIA is bound. Outlined nucleotides represent areas where site-specific mutations have reduced TFIIIA binding affinity by at least two-fold. Italicized nucleotides represent areas where site-specific mutations have little or no effect on TFIIIA binding affinity.

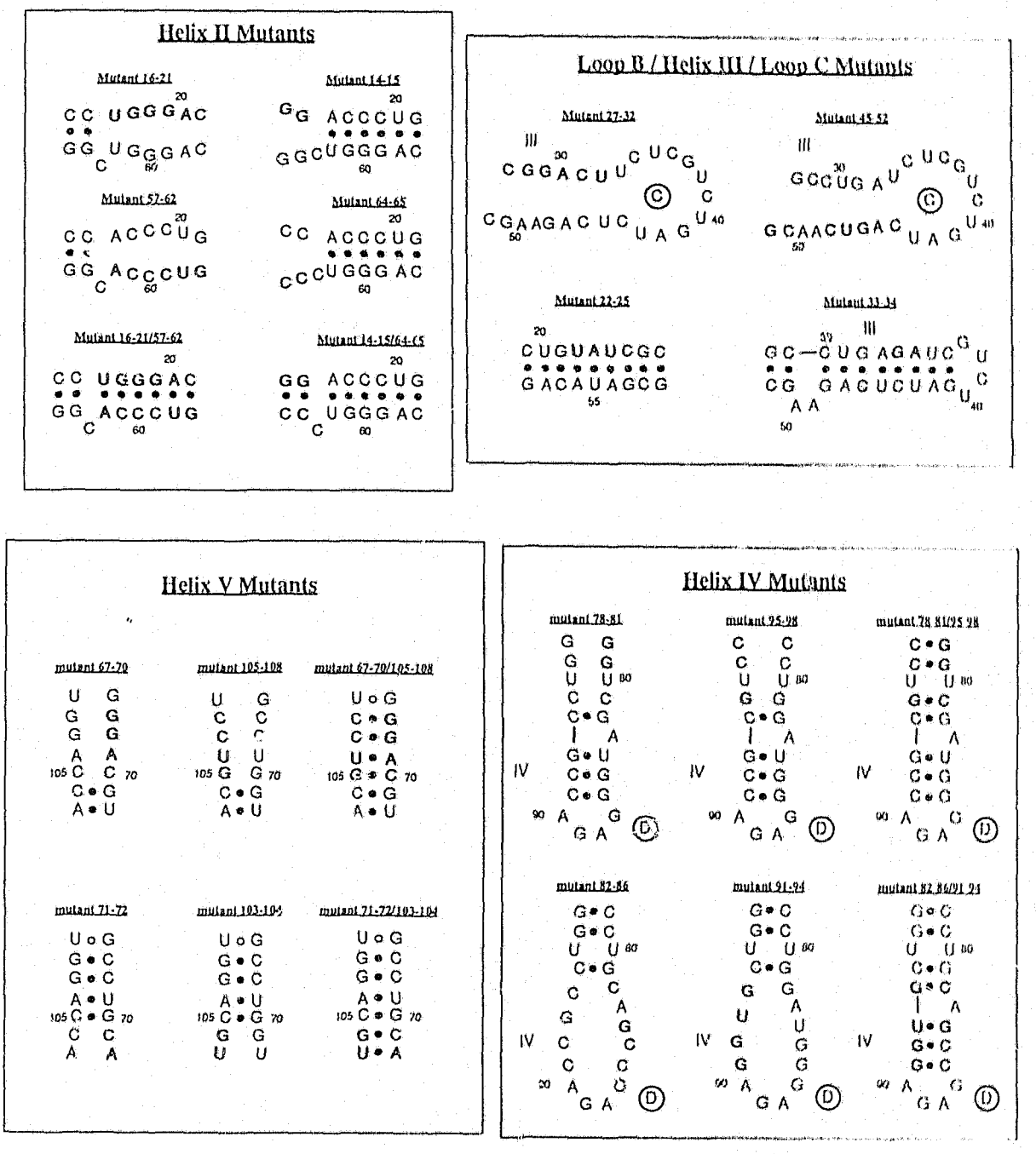


Figure 23: Stem and loop mutants used in this study. Nucleotide substitutions are indicated with outlined characters. All mutations were introduced into the intact 5S RNA molecule; only the affected region is shown for clarity.

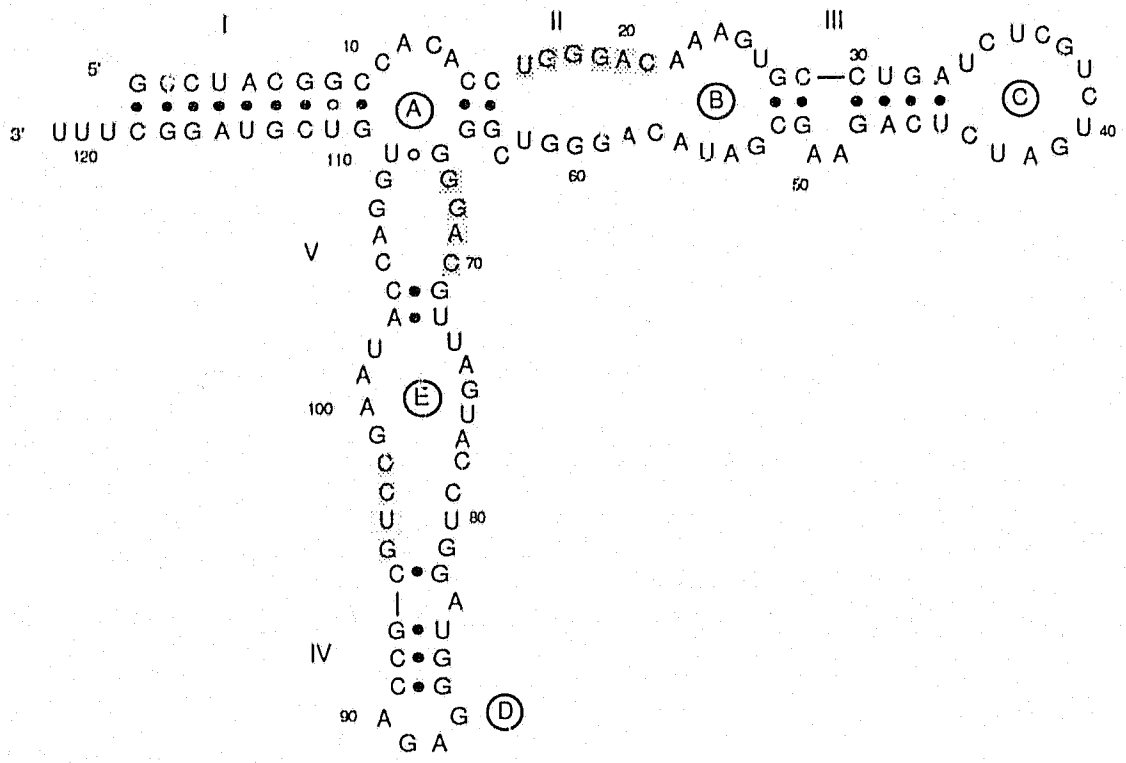


Figure 24: The disruption of secondary structure of the combination mutant 16-21/67-70/95-98. Shaded nucleotides were the substitutions.

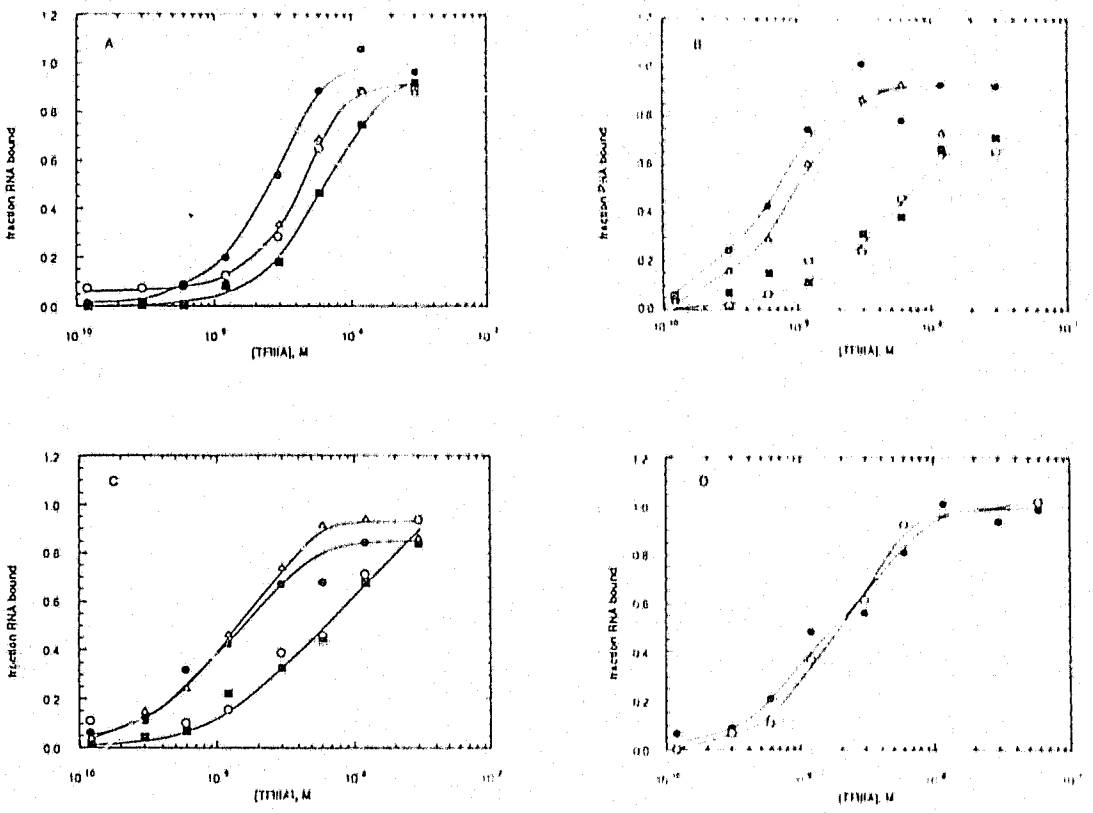


Figure 25: Nitrocellulose filter binding experiments with mutant 5S RNAs. A. (●) wild type 5S RNA, (○) 67-70 mutant, (■) 105-108 mutant, (Δ) 67-70/105-108 mutant. B. (●) wild type 5S RNA, (○) 16-21/95-98 mutant, (■) 16-21/67-70/78-81 mutant, (Δ) 16-21/57-62/78-81/95-98 mutant. C. (●) wild type 5S RNA, (○) 71-72 mutant, (■) 103-104 mutant, (Δ) 71-72/103-104 mutant. D. (●) wild type 5S RNA, (○) 33-34 mutant.

Table 3. Relative Binding Data for Mutant *Xenopus* oocyte 5S rRNA Molecules

Mutant 5S RNA[a]	Relative K_a Value[b]
Stem II	
14-15	0.85±0.22
64-65	0.74±0.24
14-15/64-65	1.11±0.32
16-21	0.32±0.15
57-62	0.40±0.15
16-21/57-62	1.09±0.48
Stem III	
27-32	0.75±0.10
45-52	0.76±0.12
Stem IV	
78-81	0.88±0.01
95-98	0.78±0.02
78-81/95-98	0.86±0.01
82-86	0.81±0.30
91-94	0.96±0.18
82-86/91-94	1.21±0.35
Stem V	
67-70	0.75±0.12
105-108	0.39±0.06
67-70/105-108	0.71±0.01
71-72	0.35±0.21
103-104	0.50±0.23
71-72/103-104	1.18±0.32
loop B	
22-25	0.84±0.12
loop C	
33-34	1.09±0.21
Combination mutants	
16-21/95-98	0.15±0.03
16-21/67-70/95-98	0.16±0.05
16-21/57-62/78-81/95-98	0.63±0.08

[a] Numbers refer to those nucleotides in the wild type 5S rRNA which have been substituted

[b] Determined experimentally as the ratio: K_a(mutant)/K_a(wild type). For each mutant, the value shown is the mean of at least three independent determinations.

Table 4. Relative TFIIIA Binding Affinities for 5S RNA Substitution Mutants at the Somatic specific sites.

mutant	Binding[a]	Competition[a]
53A	1.12 ± 0.06	0.75 ± 0.25
53C	1.08 ± 0.14	1.97 ± 1.09
53T	1.23 ± 0.04	0.83 ± 0.46
55A	1.35 ± 0.30	0.61 ± 0.19
55C	1.50 ± 0.33	1.28 ± 0.54
55G	1.24 ± 0.41	1.13 ± 0.49
56C	1.17 ± 0.30	0.82 ± 0.28
56G	1.13 ± 0.33	1.31 ± 0.41
56T	1.03 ± 0.33	1.19 ± 0.28

[a]Determined as the ratio of the apparent association constant for the mutant nucleic acid to the apparent association constant for the wild type nucleic acid. Average of two or more independent determinations.

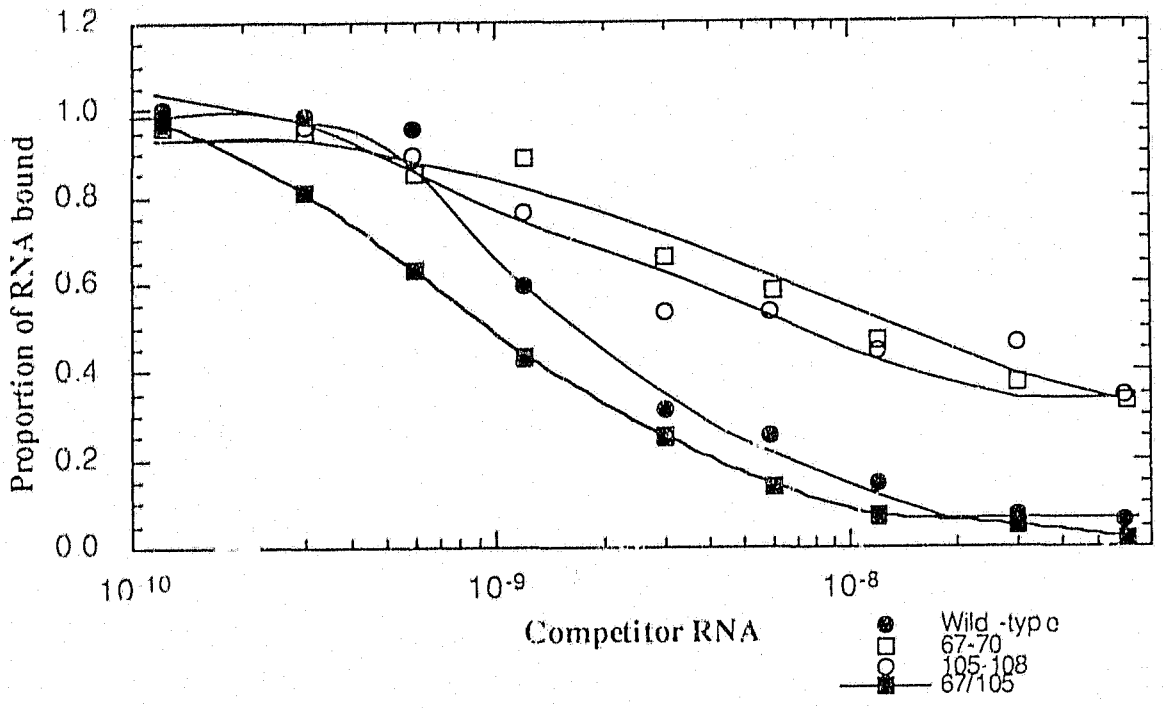


Figure 26: Competition of wild-type 5S RNA and mutants 67-70, 105-108. 67-70/105-108. Unlabelled Xlo wild-type and mutants compete with labelled Xlo wild-type 5S RNA for TFIIIA binding. [TFIIIA] was kept at 2.0 nM.

was titrated with increasing concentrations of highly purified TFIIIA. The results of several experiments are shown in Figure 25, and the complete data are presented in Table 3. A number of single mutations in the stems resulted in a decreased TFIIIA binding affinity, while double mutations generally restored TFIIIA binding to wild type levels. As the data in Table 3 show, mutations which did reduce TFIIIA binding affinity had only a modest effect on the association constant. The two single mutations that extend the helix II-helix III stem either within loop B or in loop C, had virtually no effect on TFIIIA binding affinity. The somatic specific substitutions did not show any significant decrease in TFIIIA binding affinity (Table 4).

2.4. DISCUSSION

Since TFIIIA binds to the 5S RNA gene and the transcript of this gene, it could have been logical to consider whether these two nucleic acids share some common structural features which are required for the binding of TFIIIA. There are essentially two possibilities: the TFIIIA binding region of 5S DNA adopts a 5S RNA-like conformation, or the TFIIIA binding sites of 5S RNA stack into a DNA-like conformation. Although there is evidence to support both models, neither has been established unequivocally because of contradictory evidence. For example, although TFIIIA interacts primarily with the noncoding strand of 5S DNA, which is identical in sequence to 5S RNA (Sakonju & Brown, 1982), the ICR of the gene does not appear to adopt a cruciform conformation (Reynolds & Gottesfeld, 1985). Evidence on the potential formation of an A-type helical structure in the ICR of the gene, and its functional significance, is similarly contradictory (Gottesfeld *et al.*, 1987; Aboul-ela *et al.*, 1988; Be[illegible] & Wang, 1989; Fairall *et al.*, 1989; McCall *et al.*, 1986). The second model is supported primarily by the observation that there is a high degree of overlap in the TFIIIA binding sites on DNA and RNA, the similarity being enhanced by the observation that the loop E region of the 5S RNA is helical in nature

(Romaniuk *et al.*, 1988; Westhof *et al.*, 1989; Varani *et al.*, 1989). The similarity of the DNA and RNA binding sites would be strengthened further if helix II of 5S RNA stacks on helix V (Christiaser. *et al.*, 1987; Romaniuk, 1989) but there is no direct evidence to support that particular conformation of the helical arms. Both models predict that TFIIIA binds DNA and RNA in a similar fashion and that the double stranded stems of 5S RNA therefore may be a primary target for TFIIIA binding.

The present analysis was basically designed to determine whether disruptions within the stems of the 5S RNA will affect the affinity of the RNA for TFIIIA. Nucleotide substitutions within the stems were selected to change both the sequence and the conformation of the RNA, both elements potentially being required for TFIIIA binding. These two different effects were then distinguished by the creation of compensating double mutations which change the primary sequence of the 5S RNA while maintaining the secondary structure conformation of the molecule.

From the data in Table 3, it is clear that disruption of stems II or V has negative effects on TFIIIA binding. This result is in agreement with the observation that these two stems are located within the TFIIIA binding site defined by footprinting techniques, and is also in agreement with results obtained from the study of TFIIIA binding to truncated RNA molecules (Romaniuk *et al.*, 1987). In a previous study with linker-scanning mutations of *X. borealis* somatic 5S RNA, disruption of base pairing in stem II and stem V/loop E by the substitution of nucleotides 57-67, 67-78 and 92-105 was observed by a gel shift assay to reduce the binding of TFIIIA significantly (Sands & Bogenhagen, 1987).

In the present study, the effects of a comprehensive set of smaller mutations, including double mutations that restore base pairing in the stems, on the dissociation constant provides a clear picture of the role of stem nucleotides in the binding of 5S RNA to TFIIIA. Substitution of nucleotides 16-21 or 57-62 within stem II reduces the TFIIIA binding affinity by a factor of 2.5-3 (Table 3). The double mutant 16-21 / 57-62, which restores

base pairing in stem II, also fully restores TFIIIA binding. The single mutations 14-15 and 64-65 located in the same stem, along with the related double mutation, have very little effect on the binding reaction (Table 3). In contrast, replacement of the nucleotides in loop A (nucleotides 10-13) results in a three-fold decrease in TFIIIA binding affinity, and has an even larger effect on the competition strength of the mutant 5S RNA compared to the wild type RNA (Romaniuk, 1989). Evidently replacement of the nucleotides in loop A has a more disruptive effect on TFIIIA binding than disruption of the two base pairs in stem II which close the loop. Substitution of nucleotides 16-21 or 57-62 which disrupt 4 base pairs has a more dramatic effect on the conformation of stem II, and results in a decreased affinity for TFIIIA. The fact that the double mutant 16-21 / 57-62 fully restores TFIIIA binding indicates that the helical structure of stem II, but not its native sequence, is required for full binding affinity.

Similar results were obtained for stem V, which is also located in the TFIIIA protected area. Substitutions (71-72 and 103-104) within this stem that disrupt base pairing reduce TFIIIA binding, while a double mutation 71-72 / 103-104 which yields a stem structure with an altered sequence of base pairs restores full TFIIIA binding activity. The substitution of nucleotides 67-70 and the related double mutation 67-70 / 105-108 result in a similar reduction in TFIIIA binding affinity (Table 3), suggesting that there may be a minor sequence specific interaction between TFIIIA and these base pairs. Substitutions of highly conserved nucleotides in the neighbouring loop E region also modestly reduce TFIIIA binding affinity (Romaniuk, 1989). Studies on the solution structure of *Xenopus laevis* oocyte 5S RNA have indicated that this region adopts a sequence-specific, helix-like conformation consisting of several non-canonical base pairs (Romaniuk *et al.*, 1988; Westhof *et al.*, 1989). Nucleotide substitutions in this region disrupt this quasi-helical structure (Stevenson *et al.*, 1990) and it is therefore unclear whether TFIIIA makes sequence specific or conformation specific contacts in the loop E region of the stem V-IV

domain. However, the combined data from the stem V and loop E mutants indicate that this extended stem structure is required for full TFIIIA binding affinity, with the possible formation of several weak sequence-specific contacts between the protein and the 5S RNA.

The disruptions of stem III or IV has very little effect on the TFIIIA-5S RNA interaction (Table 3). In addition, two mutants which extend Watson-Crick base pairing in this region of the 5S RNA by converting the structures of loop B or loop C to stems (mutants 22-25 and 33-34) have no effect on TFIIIA binding. These data agree with previous observations that stem III is not protected by TFIIIA from chemical modification or RNase digestion (Pieler & Erdmann, 1983; Romaniuk, 1985; Christiasen *et al.*, 1987; Huber & Wool, 1986), and can be disrupted by large linker-scanning substitutions without significantly reducing TFIIIA binding (Sands & Bogenhagen, 1987). The six mutations made within stem IV show little or no effect on TFIIIA binding affinity (Table 3), even though this stem is considered to be within the TFIIIA protected region. Huber *et al.* proposed that tandem CCUGG box regions within helices IV and V were critical for the binding of TFIIIA to 5S RNA (Huber & Wool, 1986). The results of the present experiments do not support this suggestion. As shown in Table 3, double mutations of the CCUGG base pairs in each stem (mutants 67-70/105-108 and 78-81/95-98) which alter the sequence of these base pairs have very little effect on the binding of TFIIIA.

In previous studies, Romaniuk *et al.* have demonstrated that substitution of certain conserved nucleotides in loops B and C can reduce or enhance TFIIIA binding to the 5S RNA (Romaniuk *et al.*, 1987; Romaniuk, 1989). In the current study, the importance of the conformation of these loops for TFIIIA was tested by introducing nucleotide substitutions which would extend Watson-Crick base pairing from the neighbouring stem through the loop. Somatic 5S RNA has a three-fold higher affinity for TFIIIA than oocyte 5S RNA, primarily as a result of the three somatic-specific substitutions in loop B at positions 53, 55 and 56 (Romaniuk *et al.*, 1987). This result suggested that nucleotides 53

to 56 in loop B of the 5S RNA constitute one of the elements required for optimal TFIIIA binding. Mutant 22-25 was constructed to form a double helix in loop B that links stems II and III producing an extended base paired structure. This mutation has no effect on TFIIIA binding, indicating that the conformation of loop B may not be essential for the formation of an interaction between nucleotides 53-56 and TFIIIA.

The importance of positions 53, 55 and 56 was further investigated by individual nucleotide substitutions at each position. Table 4 shows that none of these mutants has any effect on TFIIIA binding. Even though the six somatic specific nucleotides make a three-fold difference between Xlo and Xls affinities for TFIIIA, single substitutions of these nucleotides are not sufficient to make a detectable change in the TFIIIA binding.

A potential interaction between TFIIIA and nucleotides 41-44 of loop C has been suggested by patterns of protection from chemical modification (Christiansen *et al.*, 1987), and nucleotide substitution mutagenesis (Romaniuk, 1989). In the present study, mutant 33-34 was constructed within loop C to extend the base pairing of stem III and consequently reduce the size of the single stranded loop considerably. Binding assays indicated that this mutation has no effect on the interaction of TFIIIA with the 5S RNA (Table 3), although it has been shown previously that substitution of nucleotides 41-44 reduces the binding affinity for TFIIIA by a factor of two (Romaniuk, 1989). In the mutation studied here, substitution of nucleotides 33 and 34 allow the four nucleotides at positions 33-36 to form base pairs with nucleotides at positions 41-44 (Fig. 23), an interaction that has been confirmed by solution structural studies (Brunel *et al.*, 1990). However, this conformational change does not affect TFIIIA binding, implying that it is not the conformation, but the sequence information at positions 41-44 that is involved in the interaction between TFIIIA and loop C.

It has been suggested that loop C of the 5S RNA may form a long range tertiary interaction with either loop D (Toots *et al.*, 1982; McDougall & Nazar 1983) or loop E

(Goringer *et al.*, 1986; Hancock & Wagner, 1982; Pieler & Erdmann, 1982). If loop C did make contact with another part of the molecule, the formation of base pairs between nucleotides 33-36 and 41-44 would break this tertiary interaction causing a major change in 5S RNA conformation. Such an alteration in conformation would almost certainly adversely affect TFIIIA binding. However, the binding affinity measured for the interaction of TFIIIA with mutant 33-34 is identical to that measured for wild type oocyte 5S RNA. This result is in agreement with conformational studies which do not support the formation of long range contacts between loops C and D in *Xenopus* oocyte 5S RNA (Stevenson *et al.*, 1991; Brunel *et al.*, 1990).

Larger reductions in TFIIIA binding affinity were observed with combination mutations which disrupt more of the secondary structure of the 5S RNA. The mutants 16-21/95-98 and 16-21/67-70/95-98 both demonstrate a greater reduction in TFIIIA binding than was observed for any of the parent single mutants (Table 3). It is interesting to note that the binding affinity of these combination mutants is roughly equivalent to the sum of the effects of each parent mutation. This result suggests that although the effects observed for each single mutant are relatively small, each nucleotide region that was substituted contributes directly to the overall free energy of TFIIIA binding. In comparison, the combination of the 16-21/57-62 and 78-81/95-98 mutations, which alters the sequence but not the base pairing of stems II and IV, results in a mutant 5S RNA that binds TFIIIA with only a slightly reduced affinity compared to the wild type 5S RNA (Table 3). This result is consistent with a view that structure, but not nucleotide sequence, is the main feature of helical stems essential for TFIIIA binding.

Exactly how do the helical stems of *Xenopus* 5S RNA participate in the binding of TFIIIA? The deep and narrow major groove of an A-type RNA double helix is generally inaccessible for interaction with functional groups on protein secondary structural domains. A more accessible region of A-type helical stems is the minor groove. The array of base

pair functional groups oriented towards the minor groove provide for very little discrimination of base pair sequence (Seeman *et al.*, 1976). Although sequence-specific protein-RNA contacts formed in the minor groove of an RNA stem have been observed in the crystal structure of a tRNA synthetase-tRNA complex (Rould *et al.*, 1989), the apparent lack of sequence specificity in the interaction of 5S RNA stems with TFIIIA is consistent with the structural constraints of the RNA double helix. It is possible that TFIIIA contacts the sugar-phosphate backbone of the RNA stems, which would be sensitive to their conformational context far more than their sequence context. Many such potential contacts have been identified in the glutaminyl-tRNA synthetase:tRNA complex (Rould *et al.*, 1989). The contribution of these types of contacts to the overall high degree of specificity with which TFIIIA binds to 5S RNA would seem at first glance to be unlikely. However, the results of this and previous studies indicate that TFIIIA contacts two separate arms of the 5S RNA along the central base paired stems and neighbouring loops as indicated in Figure 27 (Romaniuk *et al.*, 1987; Romaniuk, 1989; Sands & Bogenhagen, 1987; Baudin *et al.*, 1989). Although a detailed three dimensional structure of the 5S RNA is not available, a graphic model constructed on the basis of probing the solution structure of *Xenopus* oocyte 5S RNA suggests that stems II and V are co-axially stacked, but are not colinear (Westhof *et al.*, 1989) (Fig. 5). Therefore the specific contacts on the 5S RNA for TFIIIA may be presented in a three dimensional array that is not duplicated by other RNA molecules, nor by the 5S DNA.

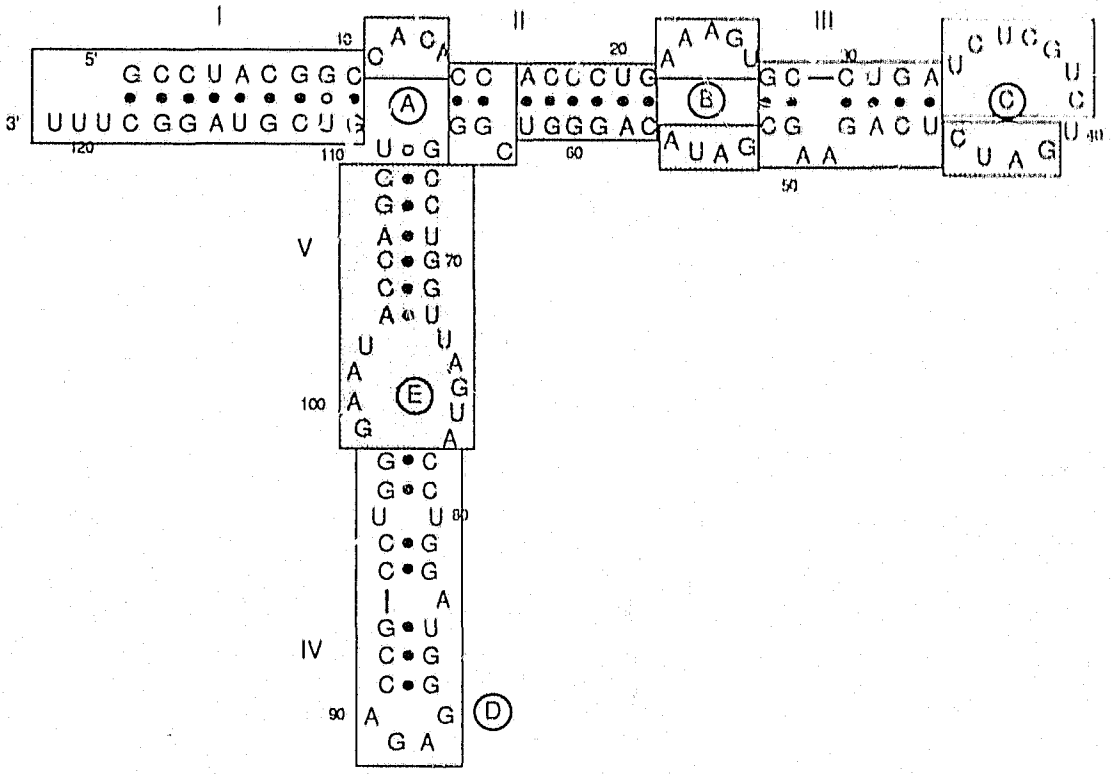


Figure 27: TFIIIA contacts on *Xenopus* oocyte 5S RNA. Shaded boxes indicate regions in which nucleotide substitution results in at least a two-fold reduction in binding affinity. Open boxes indicate regions where TFIIIA binding is essentially insensitive to nucleotide substitution.

CHAPTER 3
TFIIIA-5S DNA Interaction

3.1. INTRODUCTION

3.1.1. Characterization of TFIIIA-5S DNA Interaction.

The interaction of TFIIIA with 5S DNA has been extensively studied. Methods used for the investigation include DNA footprinting, chemical modifications, substitution/deletion mutagenesis of DNA, gel retardation, filter binding assays, immunoprecipitation and numerous other available techniques. However, the general binding properties of TFIIIA-5S DNA were not characterized until 1990 (Romaniuk, 1990).

The characteristics of the TFIIIA-5S DNA interaction were quantitatively measured using a nitrocellulose filter binding assay. TFIIIA has a non-specific binding affinity for DNA, but the non-specific binding can be eliminated by the presence of as little as 5 ng/ml of poly d(I-C), and experiments have shown that the specific affinity of TFIIIA for 5S DNA is independent of poly d(I-C) concentrations. Therefore, poly d(I-C) was present in all the filter binding assays for measuring the specific TFIIIA affinity for 5S DNA.

TFIIIA binds to the 5S DNA with an apparent association constant (Ka) of $1.9 \pm 0.44 \times 10^9\ M^{-1}$, while the BstN I fragment from pUC18 (non-specific control) has a Ka of less than $3.3 \times 10^7\ M^{-1}$. Thus, TFIIIA binds to 5S DNA with over a 100-fold greater strength than to the control DNA fragment. In other experiments, the selectivity ratio was even higher (Baudin & Romaniuk, 1989). These results quantitatively confirmed the previous observations that the TFIIIA-5S DNA interaction is highly specific.

Scatchard analysis of the TFIIIA-5S DNA complex formation indicated that the complex is comprised of one molecule of TFIIIA per 5S DNA. The affinity constant determined by this method was 4.5×10^9 M^{-1}, a value that is reasonably close to that determined by direct binding assay.

Although TFIIIA binding to 5S DNA is highly sequence specific, the complex is metastable. The dissociation rate constant determined by the filter binding assay was $7.53 \pm 0.9 \times 10^{-4}$ S^{-1}, with a half-life of 15.6 ± 1.9 min. Other laboratories have estimated the half-life of the complex as being even shorter, about 5-6 min (Bogenhagen *et al.*, 1982; Hanas *et al.*, 1984). An extremely stable complex can be formed only by the additional interactions of TFIIIC and TFIIIB.

The optimal conditions for TFIIIA binding to 5S DNA require a pH range of 6 to 8, 5 mM $MgCl_2$, 0.1M KCl, and temperatures below 22 °C. This information is a valuable guide for further investigation of the protein-DNA interaction. The experiments described in this chapter were performed basically under these conditions.

3.1.2. The 5S DNA Promoter Is Located Within the Gene.

RNA polymerase III synthesizes 5S RNA, tRNA and some other small nuclear and cytoplasmic RNAs. In general, the transcription of 5S RNA genes is under the control of intragenic DNA sequences (reviewed by Ciliberto *et al.*, 1983; Geiduschek & Tocchini-Valentini, 1988). The intragenic promoter region is commonly called the internal control region (ICR). The ICR was first identified by Sakonju *et al.* (1980) in *Xenopus borealis* somatic 5S RNA genes. A detailed analysis of the effects of 5' and 3' gene deletions revealed the boundaries of the ICR. It was found that the minimal sequence required for transcription initiation is from +50 to +83 (where +1 denotes the site of transcriptional initiation). Efficient transcription of these mutants was correlated to TFIIIA binding ability. Similar studies using a series of point mutants extended this internal control region from

+45 to +97 (Pieler, 1985a). Functionally important areas within the ICR were delineated by studying the transcriptional activity of a number of linker-scanning and substitution mutations, thus identifying the three sub-domains (Box A, the intermediate element, and Box C) necessary for the promoter activity of the ICR (Bogenhagen, 1985; Majowski *et al.*, 1987; Pieler *et al.*, 1987; Pieler *et al.*, 1985). A combination of experimental approaches including chemical modification/ selection, DNase footprinting, and template-commitment assays yielded data that indicated TFIIIA interacts most strongly with the non-coding strand of the 5S DNA within Box C, and forms weaker interactions with Box A and the intermediate element (Bogenhagen, 1985; Fairall *et al.*, 1986; Majowski *et al.*, 1987; Pieler *et al.*, 1987; Pieler *et al.*, 1985; Sakonju & Brown, 1982). These studies also indicated that TFIIIC interacts with the DNA at sites located within the ICR (Majowski *et al.*, 1987). Precise localization of TFIIIA binding interactions within the ICR resulted from hydroxyl radical footprinting of 5S DNA with sequential N-terminally and C-terminally deleted TFIIIA (Vrana *et al.*, 1988). Spacer regions between the promoter elements are quite tolerant to changes in sequence. The box A element (bp 50-64) has a relatively low affinity for TFIIIA and is believed to be involved in TFIIIC binding. The Box C (bp 80-90) and Intermediate Element (bp 67-72) are the main determinants of affinity for TFIIIA. Box C has been proven to be the most crucial sequence. Point mutations within the 3' box C region (+80 to +90) severely affect assembly of a functional transcription complex, while changes in box A or intermediate element sequence result in moderate effects on transcription (Pieler *et al.*, 1985a; 1985b). The importance of Box C for TFIIIA binding was also demonstrated by my scanning substitution mutation results (see results section of this chapter). Recently, experiments using a bacterially expressed polypeptide that includes the first three fingers of TFIIIA showed that the majority of the free energy of binding is due to the interaction of these N-terminal fingers to base pairs from +80 to +90 of the internal control region (Liao *et al.*, 1992). Thus, current experimental evidence strongly

supports a model in which the primary strong binding contacts between TFIIIA and the 5S gene reside within the Box C promoter element. The Box A of the 5S gene is homologous to the Box A or D-control region of tRNA genes, while the other elements are 5S gene specific. It is not surprising that the 5S gene shares some common sequence elements with tRNA genes, because they are transcribed by the same polymerase using some common transcription factors.

TFIIIA interacts with an extended DNA binding site covering about 50 bp of DNA. It is oriented in such a way that the C-terminal is toward the 5' end of the gene and the N-terminal is toward the 3' end. This orientation is consistent with the proposal that the N-terminal part of TFIIIA is mainly for specific binding and that the C-terminal part may contact with polymerase III or the other transcription factors. The overall interaction of TFIIIA-5S DNA can be summarized as follows:

1. There are nine zones of interaction along the ICR, located approximately 5 bp apart. It is plausible that each of these zones may contact with an individual zinc finger of TFIIIA;

2. The interactions with DNA in these nine zones are not equally strong;

3. The strongest binding sites are located near the 3' end of the ICR, in the position +80 to +90 of the Box C segment. But the 5' end of the ICR must also make a contribution to the selective binding of TFIIIA, since changes at bp +53 and +55, which include some of the nucleotides that distinguish the somatic and oocytic 5S gene, are important for the developmental regulation of 5S gene expression (see chapter 1).

4. In at least some of these nine zones the bound protein contacts both DNA strands.

3.1.3. The TFIIIA-5S DNA Interaction Models.

As discussed above, TFIIIA binds to 5S DNA in a colinear fashion. The data from the DNA footprints of TFIIIA using DNase I (Smith *et al.*, 1985; Vrana *et al.*, 1988) or hydroxyl radical (Vrana *et al.*, 1988), and from model building (Berg, 1988; Gibson *et al.*, 1988) suggest two possible models:

1. The "wrapping around" model, where TFIIIA follows the helical path of the major groove, successive zinc fingers making contacts with opposite sides of the DNA double helix without crossing over the minor groove;
2. The "alternating" model, where alternate fingers bind on one face of the DNA helix, rather than following the screw, so that successive fingers must cross over the minor grooves.

Protection experiments that use DNase I and DNase II, micro coccal nuclease, and diethyl sulfate suggest that TFIIIA binds in the major groove and mainly to one face of the DNA helix, and therefore support the "alternating" model (Churchill *et al.*, 1990) (Fig. 28). However, data from model building (Berg, 1988) and the NMR structure determination of the TFIIIA-5S DNA complex (Parraga *et al.*, 1988; Lee *et al.*, 1989) favour the "wrapping around" model.

A detailed structural analysis of a zinc finger protein-DNA complex was reported by Pavletich and Pabo (1991). A complex containing a three zinc finger peptide (Zif 268) and a consensus DNA-binding site was crystallized, and the structure of this complex was solved by X-ray scattering at 2.1Å resolution. The overall structure of this complex shows that the three zinc fingers are arranged in a semicircular ("C"-shaped) structure that fits directly into the major groove of B-DNA (Fig. 29). The arrangement of the protein-DNA complex is antiparallel, that is, the N-terminal of the peptide is toward the 3' end of the DNA strand, and the C-terminal toward the 5' end of the DNA strand (Fig. 29 & 30).

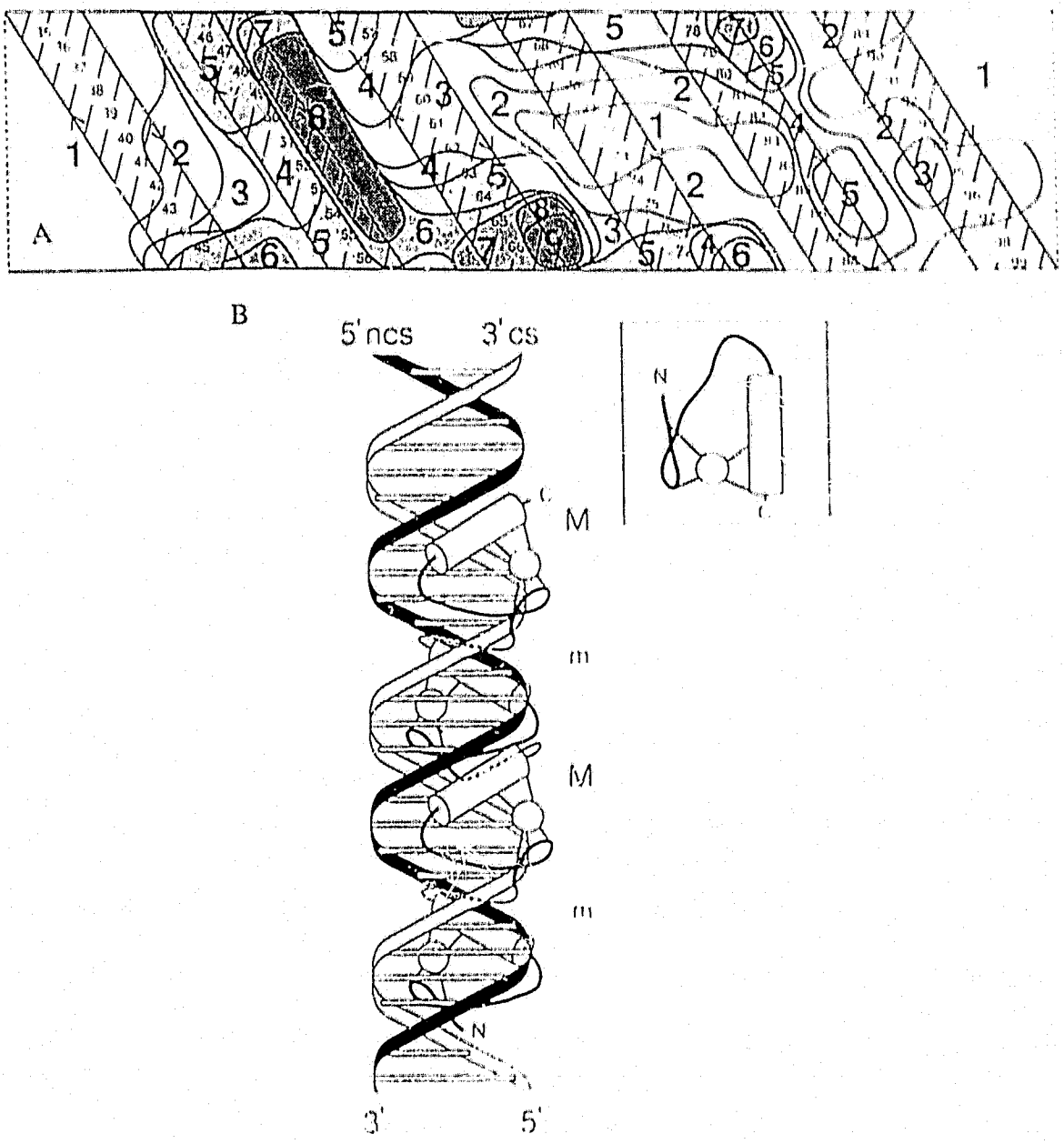


Figure 28: TFIIIA-5S DNA interaction based on hydroxyl radical data.
A: Map of protection and exposure to hydroxyl radical for the TFIIIA-5S gene complex. Contours were drawn and levels labeled 1-9 in an increasing scale of protection and shaded accordingly so that the most exposed regions appear the lightest. The small numbers shown on the minor groove of the DNA denote nucleotide positions. This cylindrical projection has been produced by slitting the cylinder along the helix axis opposite to the exposed face of the DNA.
B: Schematic diagram of zinc fingers placed on a DNA helix to show the mode of interaction favored by hydroxyl radical footprinting data. ncs: noncoding strand; cs: coding strand; M: major groove; m: minor groove (after Churchill *et al.*, 1990).

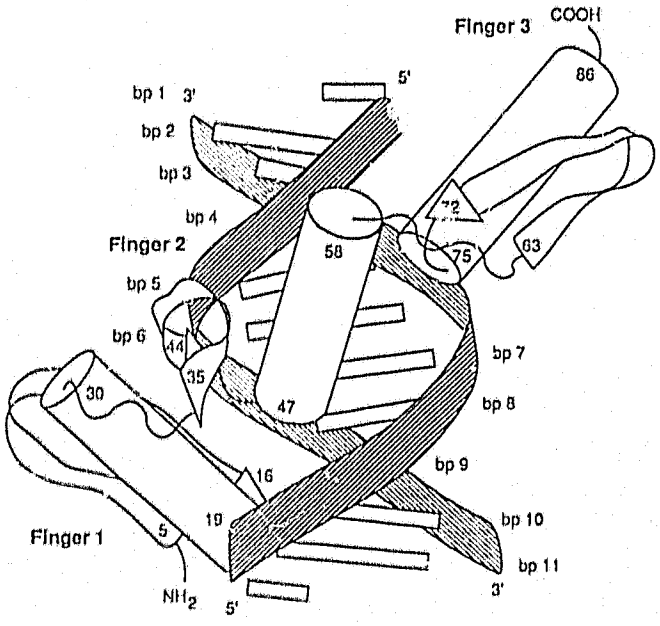


Figure 29: Zinc finger-DNA complex structure at 2.1 Å resolution.
The three Zif268 zinc fingers are arranged in a semicircular, C-shaped structure that fits into the major groove of DNA. Sketch showing the relation of the zinc fingers with respect to each other and with respect to the DNA. The starting and ending residues of each α helix and β sheet are indicated, together with the base pair numbers (after Pavletich & Pabo, 1991).

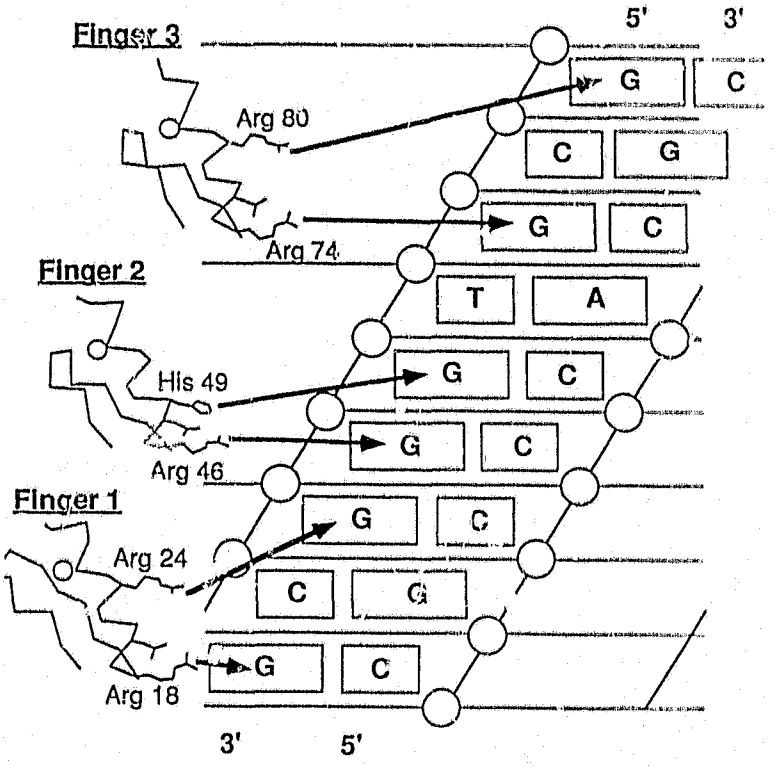


Figure 30: Zinc finger contacts on DNA.
Sketch summarizing all the base contacts made by the Zif268 peptide. The DNA is represented as a cylindrical projection (after Pavletich & Pabo, 1991).

Each zinc finger contacts a three base pair subsite of the DNA. Most contacts are made between guanine residues on the DNA and arginine on the protein, although Zif268 also uses a non-zinc ligand histidine in the second finger for a contact. In general, the a-helix of the finger contacts the N7 and O6 groups of the guanine bases in the center of the DNA helix, while the ß-sheet forms hydrogen bonds with the phosphodiester oxygen of the sugar-phosphate backbone. The arginine-guanine contacts are stabilized by nearby aspartic acid residues. These details confirmed the previous suggestions that in the binding of 5S DNA the zinc fingers of TFIIIA are linearly arranged and that each finger makes individual contacts with the consensus DNA site.

3.1.4. 5S DNA Conformations.

The ICR of 5S DNA was considered to be an A form helix, a form commonly seen in RNA double helical structures. A determination of the structure of bp +81 to +90 (Box C) of the *Xenopus laevis* 5S genes to 3Å showed that the DNA has a geometry closer to A than to B helix (McCall *et al.*, 1986). Since there are undetermined contributions of lattice forces to the relative stabilities of alternate structures in crystals, this result does not require the entire ICR to be in the A form in solution. Indeed a 22 bp DNA fragment from the 3' end of the same gene in solution has been measured by a topoisomer bandshift method and was found to be close to the B form (Peck & Wang, 1981). The more recent crystallographic study of zif268-DNA complex at 2.1Å (discussed above) also found a B form DNA helix in the complex (Paletich & Pabo, 1991). There is also a report that the ICR region may have a conformation lying between A and B forms (Fairall *et al.*, 1989). Since the A form DNA is structurally related to the RNA double helix, the A DNA-protein interaction becomes an important consideration in the RNA-protein research (See chapter 2).

3.1.5. The Other Specific Sequence of 5S DNA.

In *Drosophila* $tRNA^{arg}$ and $tRNA^{val}$ genes, either positive or negative effects on transcription have been found to relate to the 5' flanking sequence (Schaack *et al.*, 1984; Sajjadi *et al.*, 1987). These influential sequences are located within about 70 bp of the initiation site. The position of these sequences is important: removal of a few base pairs changes the transcription efficiency dramatically (Hipskind & Clarkson, 1983; Arnold & Gross, 1987; reviewed by Geidnschek, 1988). In *Xenopus*, replacement of the entire 5' flanking sequence of *Xenopus borealis* 5S gene has no effect on the transcription by oocyte nuclear extracts, but the 5' flanking sequences of *B. mori* and *N. crassa* 5S RNA genes are essential for transcription by homologous cell-free extracts (Larson *et al.*, 1983; Selker *et al.*, 1986). In these cases, the 5' flanking sequences may have contacts with the polymerase and/or with transcription factors other than TFIIIA.

The 5S DNA sequence between the start site and the 5' end of the ICR also influences transcription of the gene. These sequences regulate the efficiency of binding by TFIIIC. A collaboratory work (Keller *et al.*, 1990) using our substitution mutants 10-13, 22-26, 27-32 and 33-39 showed reduced levels of transcription in S-150 extract which support transcription of the wild type oocyte gene. Mutant 33-39 is most severe and shows a loss of stable TFIIIC binding when assayed by competition with a tRNA gene for limiting amounts of TFIIIC. These sequence are located within the 5S DNA coding region but are outside the ICR. The reduced transcription levels in those mutants were considered to be a result of decreased affinity for TFIIIC.

Surprisingly, mutant 16-21 showed a five to seven-fold stimulation of transcription in the same S-150 extract. The substitution of nucleotides 16-21 did not affect the binding of TFIIIA (See result of this chapter), but indeed it has a four-fold higher affinity for TFIIIC than wild type oocyte 5S DNA, as indicated by a TFIIIA stabilization assay, active

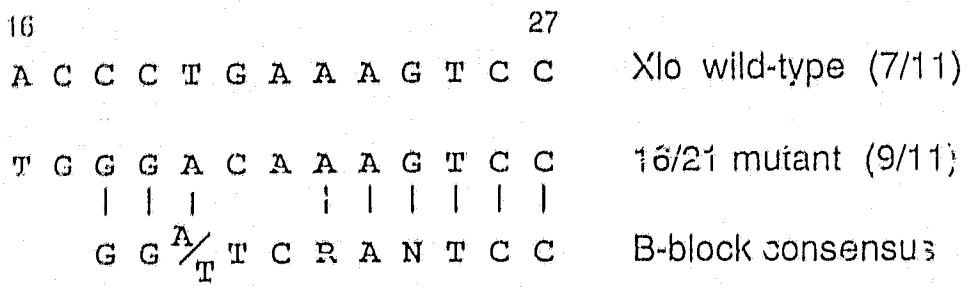


Figure 31: Comparason of tRNA B-block sequence to the corresponding regions of wild-type and mutant 16-21 RNA genes. Substitution mutant 16-21 generates a B-block within the 5S gene. The number of nucleotides out of 11 with similarity to the B-block concensus sequence is shown in parentheses.

transcription under TFIIIC limiting conditions and competition with tDNA for limiting TFIIIC (Keller *et al.*, 1990). This increased affinity for TFIIIC is likely due to the fortuitous creation of a B-block homology, which is a consensus sequence in tRNA genes directly involved in TFIIIC binding. Nine nucleotides in mutant 16-21 match with the eleven tDNA B-block consensus sequence. A weaker B-block homology is found in wild type 5S gene (Fig.31). No B-block homology is found within the ICR. The result from this study demonstrates the importance of TFIIIC and the involvement of sequence other than the ICR in 5S DNA transcription activity.

* * * * *

Although the protein-5S DNA interactions have been studied extensively, a quantitative measurement of the contribution that specific sequences in the 5S DNA make towards the overall free energy of TFIIIA binding was not obtained from any of the studies mentioned above. In this chapter, I will present the quantitative results of TFIIIA-5S DNA interaction studies using a nitrocellulose filter binding assay, which was developed in our laboratory to accurately measure the bimolecular equilibrium leading to the formation of the binary TFIIIA-5S DNA complex from free protein and DNA (Romaniuk, 1990). Using this assay, and an extensive collection of substitution mutations in *Xenopus* oocyte 5S DNA, we have now measured the relative contribution that specific sequences in the DNA make to the binding affinity for TFIIIA. From the results reported here, it is apparent that the specific sub-regions within the ICR previously reported to be target sites for TFIIIA binding do indeed contain specific nucleotide sequences that contribute directly to the free energy of binary complex formation. In particular, nucleotides 57-62 of Box A, and nucleotides 67-70 (the intermediate element) make roughly equivalent contributions to TFIIIA binding as indicated by the observation that substitution of these sequences reduced

the binding constant by a factor of 3 to 4. In addition, the data indicate that nucleotides 78-86 of Box C represent the major TFIIIA recognition feature since substitution of these nucleotides reduces TFIIIA binding affinity by two orders of magnitude.

We also investigated the contribution of individual bases within box C to the overall interaction of TFIIIA with the *Xenopus borealis* somatic-type 5S gene. The effects of point mutations within the intermediate and box C elements on the association constant for TFIIIA were quantified using the same nitrocellulose filter binding assay. In addition, a selected amplification and binding (SAAB) assay (Blackwell & Wentraub, 1990) was performed to identify box C sequences, from position +78 to +86, that bind with high affinity to TFIIIA. Specific guanine residues on the non-coding strand of 5S DNA within box C contribute significantly to TFIIIA binding. Contacts between the transcription factor and 5S DNA may include hydrogen bond formation to the O6 as well as the N7 of individual guanine residues. Sequences between the guanine clusters of 5S DNA, particularly conserved thymine residues, may also provide the proper DNA conformation for TFIIIA site recognition. These results provide insight into the contacts of TFIIIA to the individual nucleotides within the genes.

3.2. MATERIALS AND METHODS

3.2.1. Materials

Commonly used materials have been described in chapter 2. Taq DNA polymerase for the PCR reactions was purchased from Promega; Gelatin was purchased from Sigma; Poly d(I-C) for the binding assays was purchased from Boehringer Mannhein.

3.2.2. Methods

3.2.2.1. Construction of 5S DNA mutants

The construction of 5S DNA block scanning mutants has been described in chapter 2. The phagemid containing point mutants of *Xenopus borealis* 5S somatic DNA used in this study were provided by Dr. Setzer and is described in a previous publication (Setzer *et al.*, 1990). DNA vectors were maintained in *E. coli* K-12 strain NM522, and were prepared in large-scale as described in chapter 2.

3.2.2.2. Recombinant TFIIIA purification from *E. coli* containing the cloned gene

TFIIIA used for most of the experiments in this chapter was purified from *E. coli* that contains the cloned TFIIIA recombinant gene. Expression and purification of recombinant TFIIIA was performed as described by Del Rio and Setzer (1991) with modifications.

A 100 ml volume of LB media containing 50 mg/ml ampicillin was inoculated with 1.0 ml of an overnight culture of BL21(DE3)/pTF3, which is the *E. coli* strain containing the cloned TFIIIA gene in the expression vector. The culture was grown at 37 °C with shaking at 300 rpm to an $O.D._{600}$ of 0.4 to 0.6, and then supplemented with 50 μM $ZnSO_4$ followed by induction with 1.0 mM IPTG for 3-4 hours. Cells were harvested by centrifugation at 3,800 x g for 10 minutes in a Beckman JA-14 rotor and washed once with 5 ml TAB buffer (20 mM HEPES pH 7.4, 5 mM $MgCl_2$, 50 μM $ZnSO_4$, 5 mM DTT, 10% glycerol, 250 mM NaCl and 1.0 mM PMSF). The cell sample was pelleted once more at 3,800 x g for 10 minutes in a Beckman JA-20 rotor and resuspended in 4 ml TABP buffer (TAB + 1.0 mM PMSF). Sonication was performed for 6 x 20 second intervals,

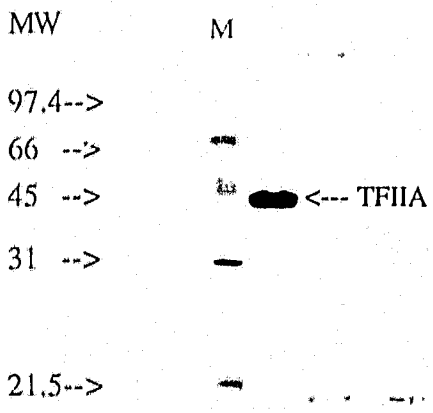


Figure 32: Purification of *Xenopus* TFIIIA expressed in *E. coli*.

with 1 minute cooling between each pulse, using a microtip sonicator at setting 3.5, 50% duty cycle. Samples were immersed in an ethanol-ice bath during sonication. The sonicate was centrifuged for 10 minutes at 12,000 x g in a JA-20 rotor. The pellet was resuspended in 1.0 ml of TABUP buffer (TABP + 5 M urea) and mixed by inversion overnight (16 hours) at 4 °C.

The extract was centrifuged in a microcentrifuge at 14,000 x g for 20 minutes and the supernatant was removed to a new tube (approx. 1.0 ml). 666 µl of TABAS buffer (TABU + 100% saturation of $[NH_4]_2SO_4$) at 4 °C was added to bring the sample to 40% saturation with the salt and mixed by inversion at 4 °C for 60 minutes. The 40% cut was pelleted by centrifuging for 20 minutes at 12,000 x g in a Beckman JA-20 rotor at 4 °C. Additional TABAS (3.33 ml) was added to the 40% supernatant (1.666 ml sample) to achieve 80% saturation and mixed by inversion for 60 minutes at 4 °C. After the sample was centrifuged for 20 minutes at 12,000 x g, the 80% saturation supernatant was removed (5 ml) and the pellet redissolved in 10 ml TABU-250 buffer (TABU, 250 mM NaCl). The resuspended pellet was applied to an 800 µl BioRex 70 column pre-equilibrated in TABU-250. The loaded column was washed with 1.0 ml of TABU-250 buffer, followed by 1.0 ml of TABU-400 buffer (TABU, 400 mM NaCl). TFIIIA was eluted with 800 µl of TABU-600 (TABU, 600 mM NaCl). An additional wash with 800 µl of TABU with 800 mM NaCl sometimes yielded the purest protein. Protein concentration and purity were determined by the method of Bradford and by SDS-PAGE analysis (Fig. 32), respectively. The protein was stored in the elution buffer at 4 °C and used for the binding assay directly. Activity gradually decreased over one week and was optimal within four days of purification. The sample can be dialysed against TAB-500 to remove urea. This purification procedure usually provides TFIIIA at a concentration of 5 to 10 µM.

3.2.2.3. PCR-based labeling of 5S DNA

The 5S rRNA gene was internally labeled in a 10 µl PCR reaction containing; 1 x PCR buffer (10 mM Tris:HCl pH 8.3 at 20 °C, 1.5 mM $MgCl_2$, 25 mM KCl), 0.05 mg/ml gelatin, 1.0 ng template DNA, 150 pmol (dTTP, dGTP, dCTP), 5 pmol dATP, 6.6 pmol α^{32}P-dATP, 0.5 pmol primers 550 and 632, and 0.25U Taq DNA polymerase. The PCR machine (Techne) was programmed as following:

Initial Cycle	24 Cycles	Final Cycle
94 °C 7 min.	45 °C 1.5 min.	45 °C 1.5 min
45 °C 1.5 min.	72 °C 2 min.	72 °C 10 min.
72 °C 2 min.	94 °C 1.5 min.	
94 °C 1.5 min.		

The coding template strand of Xbs 5S DNA was primed by oligomer 550 ($5'$CCCCCAGAAGGCAGCACAAG$3'$) and its 5' end corresponds to position -45 upstream of the transcription start site. Oligomer 632 ($5'$AAGCCTACGACACCTGGT$3'$) anneals to the non-coding template strand of 5S DNA and its 5' end corresponds to position +120. The labeled DNA products were purified on a 10 percent non-denaturing polyacrylamide gel (29:1, acrylamide:bis), eluted for 16 hours in 250 µl of elution buffer (0.6 M NH_4Ac, 1.0 mM EDTA, 0.1% SDS), and ethanol precipitated with 15 µg of $tRNA^{phe}$ as carrier.

3.2.2.4. Selected amplification and binding (SAAB) assay

An 87 nt. oligomer was synthesized (Applied Biosystems 391 DNA synthesizer) that contains the 5S RNA gene sequence from position +55 to +95 as well as unique restriction sites for cloning flanked by M13 universal priming sites at either end. The

sequence from +78 to +86 was randomized for the SAAB assay. This oligomer was made into a labeled double stranded product using Klenow and α^{32}P-dATP in a reaction described below.

Three initial rounds of selection were performed using the nitrocellulose filter binding assay described above with TFIIIA at a concentration of 100 nM. Following each selection by TFIIIA, the nitrocellulose filter was vortexed with 500 µl of dH_2O and 500 µl of phenol:$CHCl_3$ (1:1) to recover bound DNA. The aqueous layer was saved and the DNA ethanol precipitated with 20 µg of glycogen added as a carrier. This selected sample was resuspended in 30 µl of dH_2O and amplified in a 100 µl reaction containing; 1 x PCR buffer, 0.05 mg/ml gelatin, 10 µl selected DNA, 50 pmol M13 forward and reverse universal primers, 5 nmol (dATP, dTTP, dCTP, dGTP), 2.5U Taq DNA polymerase. The thermocycle program was the same as described above except that only 11 intermediary cycles were carried out to minimize heteroduplex formation. The amplified product was phenol:$CHCl_3$ extracted, ethanol precipitated with 20 µg of glycogen, and resuspended in 15 µl dH_2O.

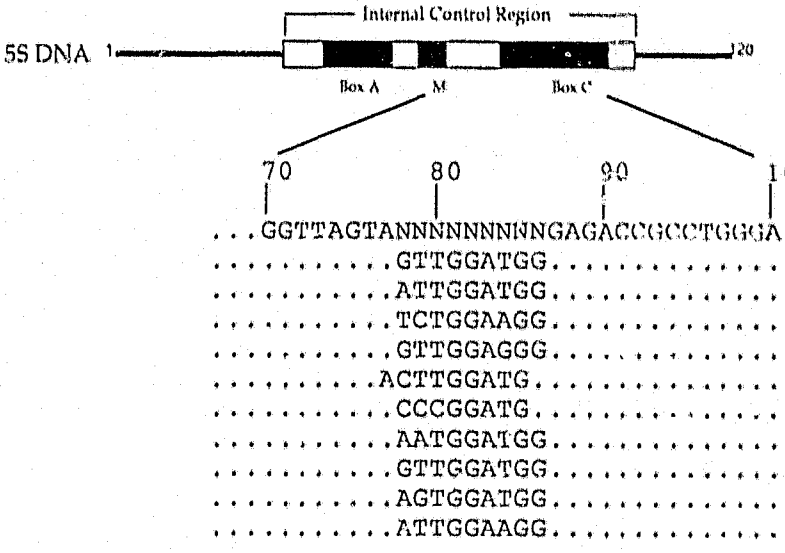
Amplified DNA was subsequently labeled by Klenow DNA polymerase in a 40 µl reaction mixture containing; 50 mM Tris-HCl pH 7.2 at 20 °C, 10 mM $MgSO_4$, 0.1 mM DTT, 50 µg/ml BSA, 5 µl to 10 µl amplified DNA (4 pmol to 8 pmol), 40 pmol M13 forward and reverse universal primers, 8 nmol (dTTP, dCTP, dGTP), 13 pmol α^{32}P-dATP (3000 mCi/mmol), 10U Klenow. Template DNA and primers were initially incubated at 94 °C and cooled slowly to 15 °C followed by the addition of other reaction components. The fill-in labeling reaction was incubated at room temperature for 45 minutes and chased with 8 nmol dATP for 5 minutes. The labeled DNA was purified on a 10% non-denaturing polyacrylamide gel, eluted, ethanol precipitated and resuspended in 20 µl of dH_2O. Approximately 0.5 pmol (10^5 cpm) was used in the next round of selection.

Once TFIIIA-dependent selection was detected by the indication of increased counts over background, three subsequent rounds of selection using a gel retention assay were performed with TFIIIA at 30 nM, 5 nM, and 2 nM final concentration. Amplified DNA was labeled with Klenow DNA polymerase as described and incubated with TFIIIA in a 30 µl reaction mixture containing; 20 mM HEPES pH 7.5, 70 mM NH_4Cl, 7 mM $MgCl_2$., 0.1% NP40, 10 µM $ZnCl_2$, 2.5 mM DTT, 100 µg/ml BSA, 6% glycerol, 5 x 10^4 cpm labeled DNA, 1.0 µg poly d[I-C]. Binding reactions were incubated at room temperature for 20 minutes and loaded onto a 5% non-denaturing polyacrylamide gel (29:1, acrylamide:bis) prepared in 0.3 X TBE buffer (30 mM Tris-borate, 0.7 mM EDTA). The gel was pre-run at 300V for 20 minutes at 4 °C. Electrophoresis was for 2 hours at 4 °C followed by autoradiography of the wet gel for one hour. Bound DNA was eluted and purified from gel slices, ethanol precipitated and resuspended in 20 µl dH_2O. Half of this sample was used for PCR amplification and further selection.

The DNA from the final round of selection was amplified and sequenced using the method described by Blackwell and Weintraub (1991). In addition, individual DNA sequences selected by TFIIIA were identified by digestion of the amplified DNA with the appropriate restriction enzymes, ligation into pUC19, and sequencing individual clones. Ten sequences of such individual samples are shown in Figure 33.

3.2.2.5. Determination of the equilibrium binding of TFIIIA to 5S DNA

The binding assays performed in this study were under the same conditions as described in chapter 2, except that 5 µg/ml poly d(I-C) was added into the reactions.



	Sequence Position								
	78	79	80	81	82	83	84	85	86
Base Occurance/ 10 samples A	4	1	0	0	0	10	2	0	0
C	2	2	1	0	0	0	0	0	0
G	3	1	0	10	10	0	1	10	10
T	1	6	9	0	0	0	7	0	0
SAAB consensus[a]	A/G	T/C	T	G	G	A	T	G	G
Native consensus[b]	G/C	C/T	N	N	G	A/G	T	G	G

[a]Determined by SAAB.

[b]Consensus sequence from Erdmann, *et al.* 1985.

Figure 33: The sequences of ten 5S DNA box C element selected by TFIIIA.

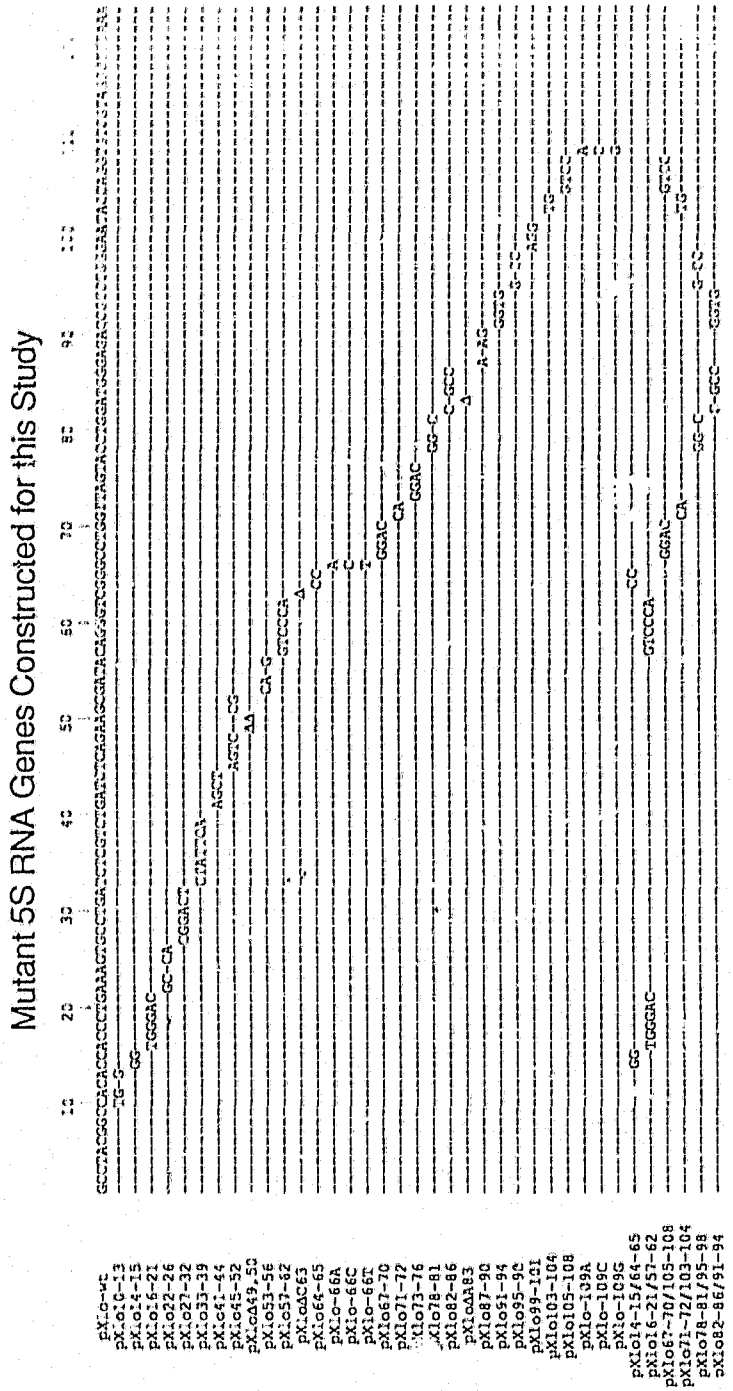


Figure 34: Sequence of the 5S RNA genes (non-coding strand) used in this study. For each mutant, a dash indicates that the nucleotide at that position is identical to that of pXlo-wt (wild type *Xenopus laevis* oocyte 5S RNA gene). A Δ symbol indicates that the nucleotide has been deleted in the mutant.

3.3. RESULTS

3.3.1. Effect of Block Scanning Mutagenesis of *X. laevis* Oocyte 5S DNA on TFIIIA Binding

The block scanning mutagenesis of 5S DNA covers from nucleotide +10 to +108, including the ICR of the gene (Fig. 34). The binding affinity of each mutant to TFIIIA was determined using the filter binding assay under conditions that successfully discriminate between the specific and nonspecific DNA binding activities of TFIIIA (see introduction of this chapter). Examples of the binding curves obtained are shown in Figure 35 and the binding affinity data are presented in Table 5. Our results agree well with the previous observations that TFIIIA primarily contacts three intragenic elements within the ICR. Figure 36 shows clearly that the sequences important for the binding of TFIIIA to the 5S DNA do correspond to specific subregions within the promoter elements of the ICR, while substitutions of nucleotides outside these elements did not affect TFIIIA binding. Nucleotides 57-62 of box A and nucleotides 67-70 of the intermediate element make similar, moderate contributions to the TFIIIA binding affinity. Nucleotides 78-86 of box C clearly are critical for TFIIIA binding, since substitutions of these nucleotides reduce the apparent K_a by as much as two orders of magnitude. TFIIIA has been shown to protect guanine N7 atoms throughout the ICR from methylation, although the protection is stronger near the 3' end (Fairall *et al.*, 1986). Nucleotides 80-86 have been shown by point mutagenesis to be extremely important promoter elements in a TFIIIA dependent transcription assay (Pieler *et al.*, 1987). These results agree well with the quantitative data presented here, and taken together provide a clear picture of TFIIIA interaction sites within the promoter, and their relative contribution to the overall free energy of binding.

The results of hydroxyl radical footprinting studies of amino and carboxyl terminal deletions of TFIIIA on 5S DNA identified three interaction sites on the DNA at which

Table 5. Relative TFIIIA Binding Affinities for *X. laevis* 5S RNA Genes.

mutant	Relative Binding Affinity* of 5S DNA
wild type	1.00
10-13	1.18±0.12
14-15	0.71±0.07
16-21	0.80±0.07
22-26	1.24±0.15
27-32	1.02±0.15
33-39	0.89±0.09
41-44	0.89±0.11
45-52	0.93±0.07
Δ49,50	1.00±0.02
53-56	1.03±0.07
57-62	0.26±0.13
Δ63	1.00±0.02
64-65	0.64±0.12
66A	1.00±0.08
66C	1.00±0.01
66T	1.00±0.05
67-70	0.33±0.0
71-72	1.31±0.17
73-76	1.07±0.16
78-81	0.06±0.02
82-86	0.01±0.005
Δ83	0.25±0.05
87-90	0.64±0.18
91-94	0.61±0.24
95-98	1.00±0.21
99-101	1.81±0.08
103-104	1.40±0.40
105-108	0.97±0.27
109A	1.00±0.09
109C	1.00±0.05
109G	1.00±0.05
14-15/64-65	0.70±0.05
16-21/57-62	0.41±0.10
67-70/105-108	0.32±0.06
71-72/103-104	1.78±0.58
78-81/95-98	0.015±0.005
82-86/91-94	0.01±0.001

*Determined as the ratio of the apparent association constant for the mutant nucleic acid to the apparent association constant for the wild type nucleic acid. Average of two or more independent determinations.

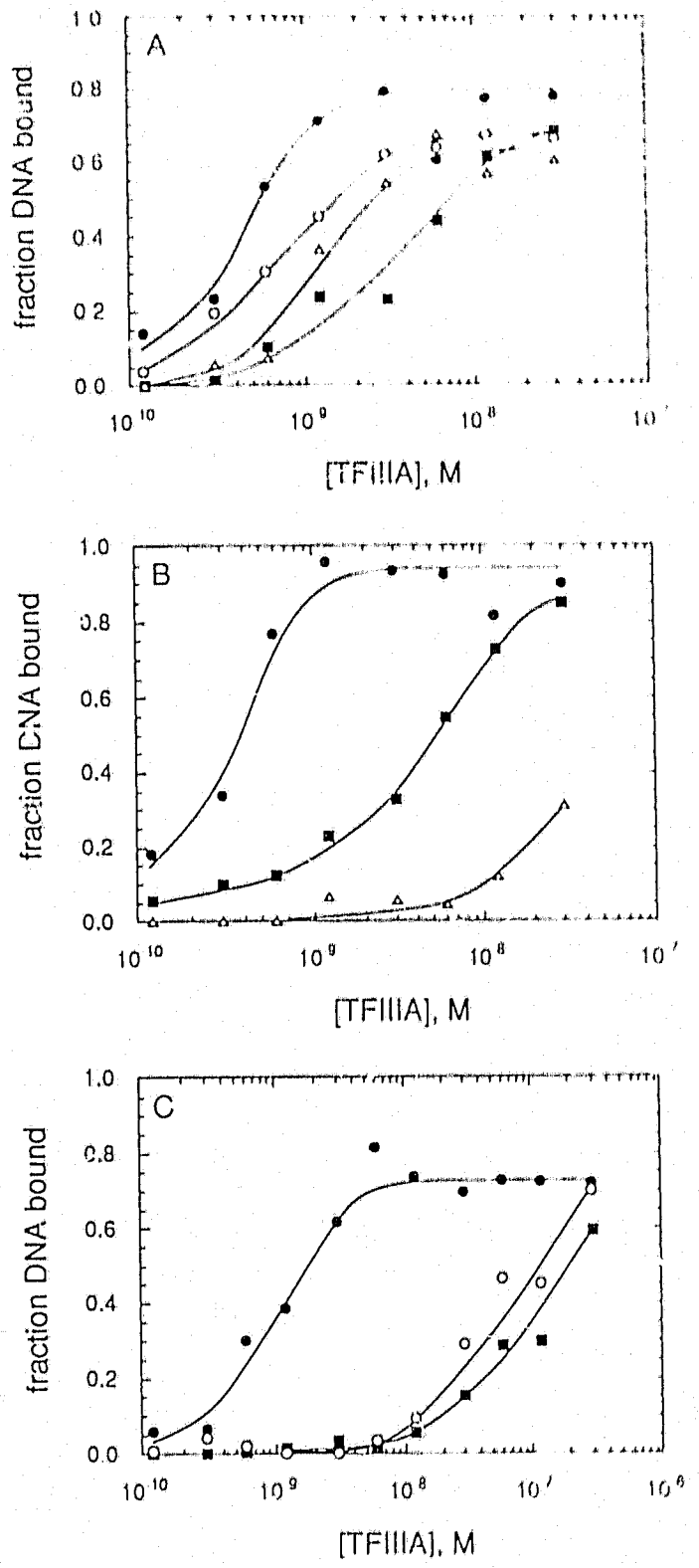


Figure 35: Nitrocellulose filter binding experiments with mutant 5S RNA genes. A. wild type (●), pXlo16-21 (○), pXlo57-62 (■), pXlo91-94 (△). B. wild type (●), pXlo78-81 (■), pXlo78-81/95-98 (△). C. wild type (●), pXlo82-86 (○), pXlo91-95 (■).

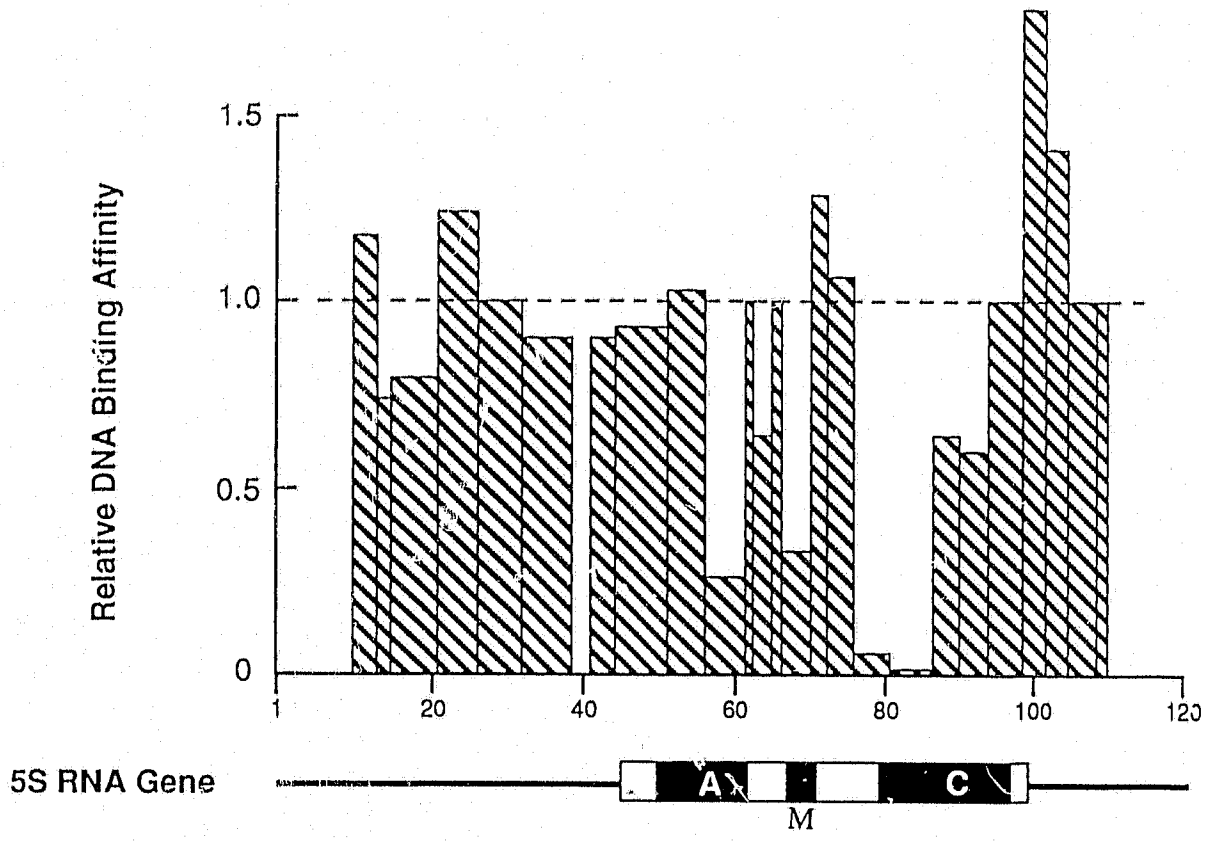


Figure 36: Binding strength of TFIIIA to mutant 5S RNA genes with alignment to the ICR.

TFIIIA fingers cluster upon binding (Vrana *et al.*, 1988). The data in this study indicate that nucleotide substitutions within these three sites decrease the binding of TFIIIA to the 5S DNA. Thus there is a convincing correspondence in the information on the TFIIIA interaction sites obtained by mutagenesis of the protein and the 5S DNA.

3.3.2. Contribution of Individual Nucleotides of *X. borealis* somatic 5S Gene to TFIIIA Binding

To further investigate the sequence specific interaction between TFIIIA and 5S DNA, we tested the importance of individual nucleotides within Box C and the intermediate element in TFIIIA binding. By measuring the TFIIIA binding affinities for a set of *Xenopus borealis* 5S DNA point mutants, we found that the contact strength of individual nucleotides with TFIIIA are non-equivalent and their mutation reduces promoter binding anywhere from two to ten-fold, depending on the base substituted. Most notably, mutation of positions G81, G85, or G89 reduces TFIIIA association five to ten-fold. Base substitution at positions G82, G86, or G87 within the Box C subregion produces moderate reductions of two to four-fold (Fig. 37 & Table 6).

The box C element of the ICR is guanine-rich (+81-GGATGGGAGA-+90). Strong interactions with TFIIIA involve a number of specific guanine bases located within this region. As mentioned above, all dramatic effects of the nucleotide substitutions occurred at guanine positions (G81, G82, G85, G86, G87 and G89). For guanine residues bound by TFIIIA, adenine base substitutions are generally tolerated over those of thymine and cytosine (A>>T>C). This is exemplified by substitution of the guanine at position +86 (Table 6). This trend suggests that the N7s of important guanine residues are involved in the contacts of 5S DNA to TFIIIA and that these contacts are maintained in mutants containing adenine substitutions. Evidence for the involvement of the guanine N7 position in TFIIIA binding has been observed previously using methylation interference

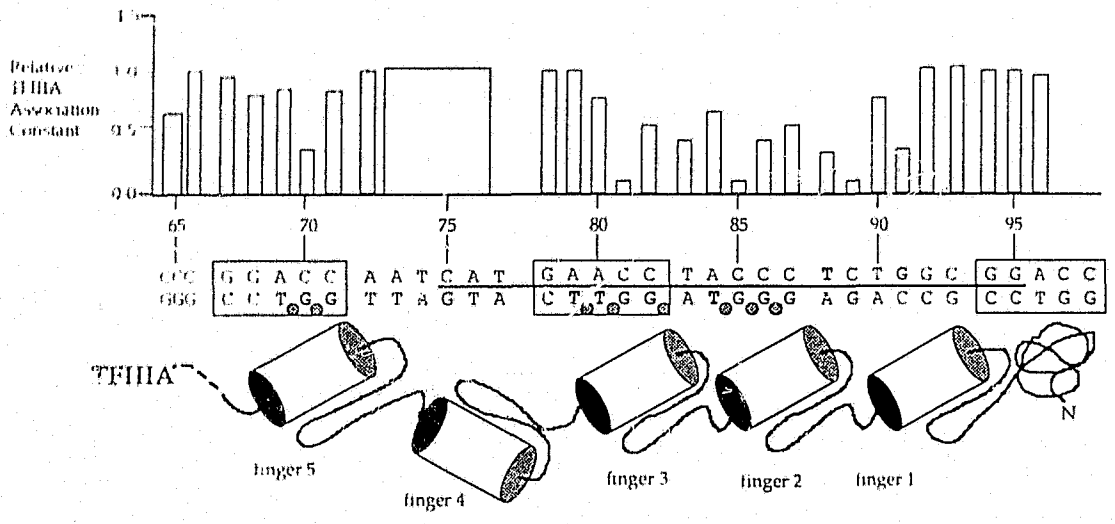


Figure 37: Proposed alignment of the N-terminal TFIIIA zinc fingers on the intermediate and box C elements of the 5S RNA gene internal control region. Xbs 5S DNA sequence of the coding and noncoding strands from residues 65 to 97 are shown on the top and bottom, respectively. The effects of point and block mutations on TFIIIA binding are indicated above the DNA sequence. Guanine bases that interfere with TFIIIA binding when methylated are indicated by boldface letters and DNA backbone phosphates that contact TFIIIA are marked by grey circles (Sakonju & Brown, 1982). Sequence protected by an N-terminal peptide of TFIIIA containing the first three fingers are underlined (Liao *et al.*, 1992). Thymine residues that closely approach TFIIIA are shown in outlined letters (Lee *et al.*, 1991). A-DNA-like sequence within this promoter region are boxed (Huber *et al.*, 1991).

Table 6. Relative TFIIIA Binding Affinities[a] for *Xenopus borealis* 5S DNA Point Mutants[b]

Base Position	Nucleotide Substitution A	C	G	T
		Intermediate Element		
C67	0.89 ± 0.07	•	1.08 ± 0.20	1.09 ± 0.18
C68	1.04 ± 0.06	•	1.05 ± 0.14	1.07 ± 0.03
T69	0.92 ± 0.03	0.98 ± 0.10	0.33 ± 0.09	•
G70	0.31 ± 0.09	0.25 ± 0.04	•	0.91 ± 0.14
G71	0.64 ± 0.11	0.74 ± 0.09	•	0.82 ± 0.11
		Box C Element		
C78	1.11 ± 0.25	•	1.22 ± 0.10	1.00 ± 0.20
T79	1.12 ± 0.13	1.12 ± 0.13	1.09 ± 0.21	•
T80	0.44 ± 0.17	0.98 ± 0.02	0.84 ± 0.04	•
G81	0.30 ± 0.06	0.13 ± 0.03	•	0.22 ± 0.05
G82	0.40 ± 0.01	0.54 ± 0.08	•	0.90 ± 0.10
A83	•	0.70 ± 0.07	0.58 ± 0.04	0.20 ± 0.09
T84	0.73 ± 0.11	0.39 ± 0.10	0.82 ± 0.14	•
G85	0.23 ± 0.07	0.23 ± 0.15	•	0.19 ± 0.07
G86	0.75 ± 0.18	0.35 ± 0.14	•	0.43 ± 0.22
G87	0.49 ± 0.08	0.61 ± 0.19	•	0.56 ± 0.07
A88	•	0.53 ± 0.13	0.70 ± 0.06	0.39 ± 0.07
G89	ND	0.19 ± 0.06	•	0.25 ± 0.08
A90	•	0.90 ± 0.10	ND	0.50 ± 0.06
C91	0.48 ± 0.06	•	0.24 ± 0.05	0.39 ± 0.13
C92	0.99 ± 0.22	•	1.27 ± 0.19	1.07 ± 0.04
G93	1.34 ± 0.14	1.14 ± 0.16	•	1.06 ± 0.10
C94	1.02 ± 0.19	•	1.13 ± 0.09	1.20 ± 0.25
C95	1.14 ± 0.04	•	1.06 ± 0.05	1.07 ± 0.05
T96	0.90 ± 0.14	0.92 ± 0.08	0.94 ± 0.10	•

[a]Determined as the ratio of apparent association constant for the mutant nucleic acid to the apparent association constant for the wild type nucleic acid. Average of three or more independent determinations.
[b]Some of the data in this table were produced by Nik Veldhoen (unpublished observations).

studies (Sakonju & Brown, 1982). In contrast, all substitution mutations of guanine residues at positions 81 and 85 reduce TFIIIA binding to the same extent suggesting a role for the guanine O6 in the protein-DNA interaction at these base positions.

Interestingly, mutations at positions A83, A88, and A90 on the non-coding strand that substitute a thymine residue for the native adenine base reduce TFIIIA binding approximately two-fold. This may be due to steric effects of the methyl group of thymine that interferes with TFIIIA binding to adjacent guanine residues.

Point mutations within the intermediate element of the 5S gene promoter were also analyzed for their TFIIIA binding ability. The transcription factor contacts position G70 and mutation of this guanine reduces binding three- to four-fold.

3.3.3. Selected Amplification and Binding of Box C Subregion

Since earlier experiments using 5S gene scanning substitution mutants identified a region extending from position +78 to +86 as being the primary contact region for TFIIIA (this chapter), this subregion of the box C promoter element was subsequently targeted for SAAB analysis. The native 5S DNA sequence from positions +80 to +83 and +85 to +86 is predominant in the ICR population selectively bound by TFIIIA from the randomized pool, indicating the importance of these positions for TFIIIA binding. In contrast, bases at positions +78, +79, and +84 remain relatively unselected and are not involved in high affinity protein-DNA interactions (Fig. 33). Selection of the guanine-rich native box C sequence further suggests the involvement of guanine residues on the non-coding stand in TFIIIA binding.

3.4. DISCUSSION

3.4.1. Mutagenesis Analysis of TFIIIA-5S DNA Contact

Our results from both block scanning and point mutations confirmed the existence of three sequence elements within the ICR that contact with TFIIIA. The results also demonstrate that the contributions of these elements to the highly specific protein-DNA interaction are unequal. Box C is the most important site for TFIIIA binding. Detailed analysis of the Box C point mutations and SAAB results revealed the contributions of individual nucleotides to the interaction. Our results, combined with previous studies, provide a clear picture that the TFIIIA-5S DNA interaction is dependent upon specific sequence recognition.

The hydroxyl radical footprinting studies conducted with N and C terminal deletions of TFIIIA identified regions within the ICR in intimate contact with specific fingers of TFIIIA (Vrana *et al.*, 1988). Apparently fingers 8 and 9 and a short stretch of the C-terminal domain of TFIIIA shield base pairs +42 to +60 of the DNA from modification by hydroxyl radicals. The results of our study suggest that the sequence-specific interactions between this region of TFIIIA and the 5S DNA occur within the much smaller site of base pairs +57 to +62. It was suggested, based on the footprinting studies, that fingers 5 and 6 protect the minor groove of the ICR at the intermediate element and make critical contacts with the DNA in the major groove at adjoining nucleotides including the two G residues at base pairs +70 and +71 (Vrana *et al.*, 1988). Our block scanning mutation data indicate that nucleotide substitutions from base pairs +67 to +70 of the intermediate element significantly reduce TFIIIA binding affinity, in agreement with the hydroxyl radical footprinting results. We further investigated the intermediate element with point mutations. Single nucleotide substitutions from +67 to +70 constituted a set of mutants in this region. The binding affinities of these mutants for TFIIIA showed that only mutations at position +70 gave a significant decrease in binding activity (Table 6). The

importance of G70 was indicated by the relative binding strengths of A70 (0.31) and C70 (0.25). The single nucleotide substitutions at this position reduced the binding of TFIIIA to levels equal to that of the block substitution 67-70 (0.33). Considering the fact that point mutants at positions +67, +68 and +69 respectively did not show any significant effect on the TFIIIA binding, we concluded that G70 is the only nucleotide in the sequence 67-70 that is important for the TFIIIA contact. Furthermore, the down-stream nucleotide G71 was tested by substitutions and did not show a significant reduction in the TFIIIA binding (Table 6). The results we obtained showed clearly that not all nucleotides in the intermediate element are important for the TFIIIA contact, and that the guanine residues may be involved in the direct contacts with TFIIIA. The result strongly agreed with the hydroxyl radical footprinting observations of Vrana *et al.*.

Results obtained with the N-terminal deletion mutants of TFIIIA indicated that the largest contribution to binding strength was obtained from the interaction of fingers 1 and 2 with about 17 base pairs (+80 to +97) within Box C of the ICR (Vrana *et al.*, 1988). This is supported by functional studies, using 5S DNA point mutants, that identify the 3' boundary of the ICR as position +97 (Pieler *et al.*, 1985a). However, another collection of point mutants centered around residue +90 suggests the 3' border of the ICR extends only to C91 (McConkey & Bogenhagen, 1987). Point mutants at residues +92, +93, +94, and +95 do not detectably affect TFIIIA binding, as determined by DNase I protection, or transcription efficiency (McConkey & Bogenhagen, 1987). The extent of TFIIIA contacts within box C has also been defined by methylation protection and ethylation interference and show base contacts to position +91 (Sakonju & Brown, 1982). These studies suggest that the 3' end of the ICR involved in energetically favourable contacts with TFIIIA extends to residue +91. Residues +92 to +97 may be involved in subtle contacts with TFIIIA that further define the functional complex but are not energetically important or contact other components of the polymerase III transcription complex (possibly TFIIIC).

The hydroxyl radical footprinting data indicated that the protein appeared to be in close contact at the major groove with almost every base pair through at least one turn of the DNA double helix within the Box C region. This area of the ICR includes six G residues demonstrated to be important for TFIIIA binding by a methylation interference assay (Sakonju & Brown, 1982). The data in Table 5 indicate that a strong sequence-specific interaction occurs between nucleotides 78-86 of the 5S DNA and the N-terminal fingers of TFIIIA. Thus there is excellent agreement with regards to the relative contributions of protein and DNA regions to the overall free energy of the TFIIIA-5S DNA interaction obtained from these independent binding studies conducted with TFIIIA mutants and 5S DNA mutants. In addition, our data clearly identifies which sequences within the hydroxyl radical footprint regions are likely to be in direct contact with TFIIIA.

More detailed results came from our point mutations in this region (+78 to +90) (Table 6). The data showed that the most severe effects came from the substitutions at G81, G85 and G89. Mutations at G82, G86 and G87 showed moderate decreases in the binding affinities. The importance of guanine residues in the Box C region was further emphasized by the SAAB result: among the randomized nucleotide sequence +78 to +86, all four guanine residues were selected either 100% (G82 and G86) or 90% (G81 and G85) (Fig. 33). These independent experiments agreed well with each other in the determining the TFIIIA contact sites in the Box C region.

The guanine contacts are non-equivalent in strength which suggests that the N7 and O6 atoms are involved in hydrogen bond formation to different degrees, depending on the position of the guanine base. Interestingly, the inactive *Drosophila* 5S gene differs from its active counterpart by a G to A substitution at position +86 (Sharp *et al.*, 1984). The corresponding point mutation in *Xenopus* 5S DNA does not result in a significant reduction in TFIIIA binding (relative Ka 0.78) (Table 6), although transcription efficiency is reduced ten-fold (Pieler *et al.*, 1985b). In contrast, a G to C substitution at position +86 results in a

larger reduction in TFIIIA binding (relative Ka 0.35). Thus, the nature as well as the position of promoter mutations may direct subtle changes in protein-DNA complex stability and biological activity. This is clearly evident in the 5S gene family, where base differences at positions +53, +55, and +56 between somatic and oocyte forms lead to significant differences in TFIIIA-5S DNA complex stability and developmental regulation of transcription (see chapter 1).

T80 and A83 were also selected in a SAAB experiment at high levels (90%) (Fig. 33). Nucleotide substitutions at these two positions had moderate effects on the TFIIIA-5S DNA interaction, except for the five-fold decreased affinity of mutant T83, in which the native adenine base was replaced by a thymine base. We propose that this may be due to steric effects by the methyl group of thymine that interferes with TFIIIA binding to adjacent guanine residues. Similarly, the most obvious effect at the highly selected T80 position was observed with mutant A80, a T to A substitution. It seems likely that the native T80 and A83 were selected in the SAAB experiment not because of their direct contacts with TFIIIA, but because of the indirect steric effects that may affect the TFIIIA binding.

In summary, TFIIIA interacts strongly with the box C element of the 5S gene promoter through multiple contacts to guanine residues over one complete turn of the DNA helix (position +81 to +91). The N-terminal zinc fingers of TFIIIA provide the majority of the free energy of association with this region of 5S DNA. The other nucleotides in this region may also be important for the contact because of their steric effects. It is likely that some TFIIIA fingers also contact with the Box A and intermediate elements.

3.4.2. Comparison of TFIIIA-5S RNA and 5S DNA Interactions

The relative TFIIIA binding affinities for the 5S DNA and 5S RNA mutants are compared in Table 7 and Figure 38. It is striking that these mutations have a much larger

Table 7. Comparison of Relative TFIIIA Binding Affinities for 5S DNA and 5S RNA mutants.

mutant	Relative Binding Affinity[a]	
	RNA[b]	DNA
wild type	1.00	1.00
10-13	0.30±0.01	1.18±0.12
14-15	0.85±0.22	0.74±0.07
16-21	0.32±0.15	0.80±0.07
22-26	1.00±0.02	1.24±0.15
27-32	0.75±0.10	1.02±0.15
33-39	1.00±0.02	0.89±0.09
41-44	0.40±0.10	0.89±0.11
45-52	0.76±0.12	0.93±0.07
Δ49,50	1.00±0.02	1.00±0.02
53-56	1.62±0.12	1.03±0.07
57-62	0.40±0.15	0.26±0.13
Δ63	1.00±0.02	1.00±0.02
64-65	0.74±0.24	0.64±0.12
66A	0.14±0.01	1.00±0.08
66C	0.12±0.01	1.00±0.01
66T	0.27±0.03	1.00±0.05
67-70	0.75±0.12	0.33±0.05
71-72	0.35±0.21	1.31±0.17
73-76	0.57±0.02	1.07±0.16
78-81	0.88±0.01	0.06±0.02
82-86	0.81±0.30	0.01±0.005
Δ83	1.00±0.02	0.25±0.05
87-90	0.71±0.10	0.64±0.18
91-94	0.96±0.18	0.61±0.24
95-98	0.78±0.02	1.00±0.21
99-101	0.61±0.08	1.81±0.08
103-104	0.50±0.23	1.40±0.40
105-108	0.39±0.06	0.97±0.27
109A	0.29±0.01	1.00±0.09
109C	0.50±0.03	1.00±0.05
109G	0.17±0.05	1.00±0.05
14-15/64-65	1.11±0.32	0.70±0.05
16-21/57-62	1.09±0.48	0.41±0.10
67-70/105-108	0.71±0.01	0.32±0.06
71-72/103-104	1.18±0.32	1.78±0.58
78-81/95-98	0.86±0.01	0.015±0.005
82-86/91-94	1.21±0.35	0.01±0.001

[a]Determined as the ratio of the apparent association constant for the mutant nucleic acid to the apparent association constant for the wild type nucleic acid. Average of two or more independent determinations.
[b]Data taken from Romaniuk *et al.* (1987) ; Romaniuk (1989); Baudin and Romaniuk (1989); You and Romaniuk (1990); Baudin *et al.* (1990).

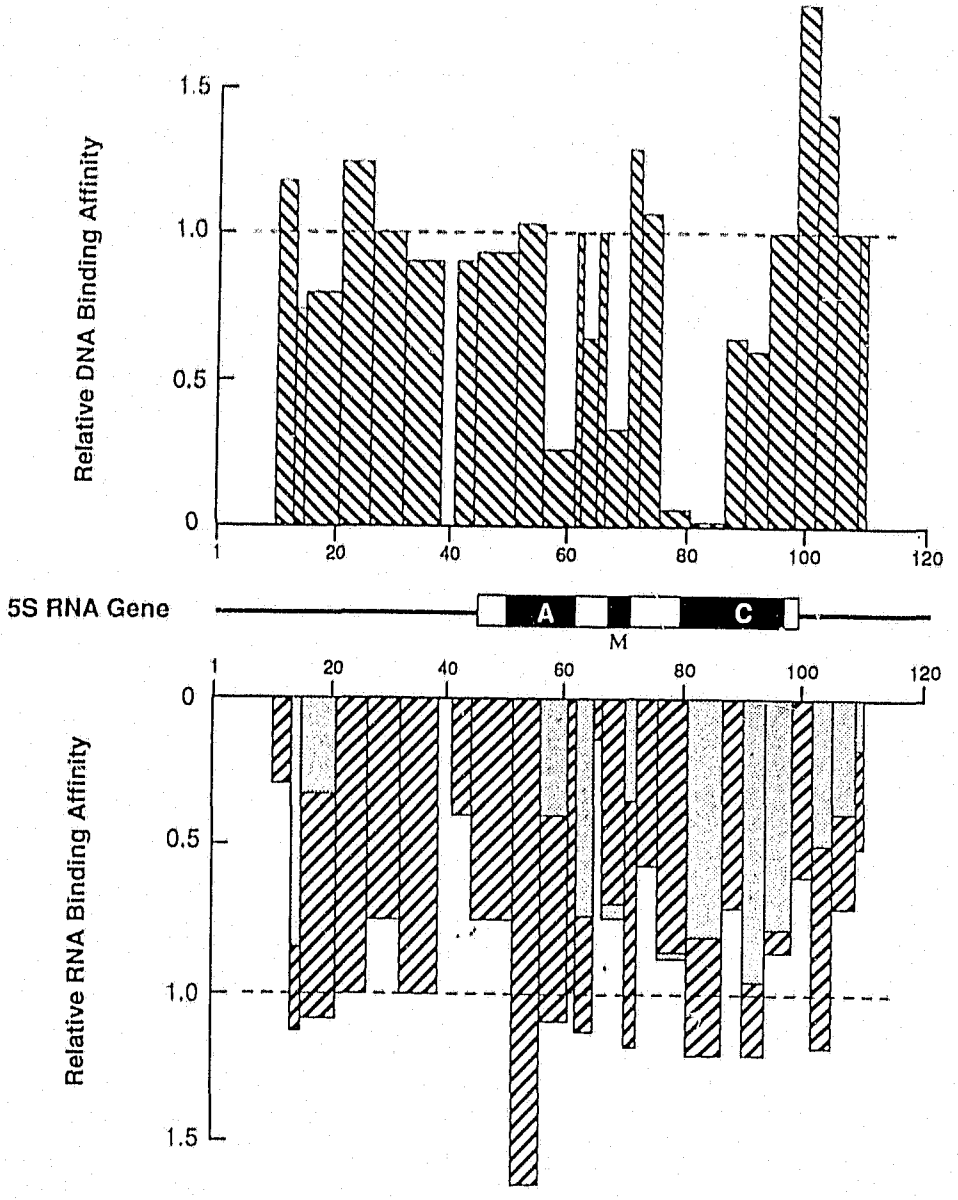


Figure 38: Comparison of the TFIIIA binding affinities of mutant 5S RNA genes and their corresponding 5S RNA transcripts relative to the wild type nucleic acids. For mutations in the 5S RNA that span helical stems, the stippled bars indicate the relative TFIIIA binding affinities of single mutations which disrupt the stem structure, while the hatched bars indicate the relative TFIIIA binding affinities of compensating double mutations which maintain a stem structure with an altered base pair sequence.

effect on the affinity of 5S DNA for TFIIIA than they have on the affinity of 5S RNA for the protein. The largest effect observed on TFIIIA binding to the 5S RNA is a decrease of eight-fold in the apparent K_a value. In comparison, certain nucleotide substitutions in the 5S RNA gene result in a 100-fold reduction in the binding affinity of the DNA for TFIIIA. Clearly, TFIIIA binding to DNA is much more sensitive to nucleotide sequence than is the binding of the protein to 5S RNA. This effect is even more striking if one considers that a number of the nucleotide substitutions in the 5S RNA disrupt a base paired stem (Figures 23 & 24), while proper Watson-Crick base pairing is maintained in all of the 5S DNA mutants. Compensating double mutations in the 5S RNA that maintain the base pairing within a target stem, while changing the sequence of base pairs, have little effect on TFIIIA binding affinity. Detailed analyses of the solution structures of a number of these mutant 5S RNAs have indicated that the decrease in TFIIIA binding affinity can in each case be explained by alterations in the secondary or tertiary structure of the 5S RNA (Baudin *et al.*, 1990; Brunel *et al.*, 1990; Romaniuk *et al.*, 1988; Stevenson *et al.*, 1991; Westhof *et al.*, 1989 & chapter 2).

Studies with linker-scanning mutations in the *X. borealis* somatic 5S RNA gene indicated that mutations which disrupt 5S RNA structure and decrease TFIIIA binding do not fall within those regions of the 5S DNA sequence important for promoter function (Sands & Bogenhagen, 1987). Figure 38 compares the effects of substitution mutations in *X. laevis* oocyte 5S DNA and 5S RNA mapped against the three promoter elements within the ICR of the gene. It is evident from this Figure that there is little correspondence between the sequence elements required for the optimal binding of TFIIIA to the two nucleic acids.

It has been suggested, based upon the results of several studies, that regions of the 5S DNA adopt a conformation that is closer to an A-type double helix rather than the more classical B-type double helix normally associated with DNA (Diakun *et al.*, 1986; Fairall *et*

al., 1989; Rhodes & Klug, 1986). However, other studies have demonstrated that the ICR region of the 5S DNA has a B type conformation in both the absence and presence of TFIIIA (Aboul-ela *et al.*, 1988; Becker & Wang, 1989; Gottesfeld *et al.*, 1987). A carefully conducted study of the conformation of a number of model oligonucleotides by circular dichroism has demonstrated that the ICR of 5S DNA has a uniform conformation that is partway between the classical A and B type double helices (Fairall *et al.*, 1989).

A quantitative investigation of the hydroxyl radical footprint of TFIIIA on 5S DNA has lead to a proposal of how the fingers of TFIIIA interact with the 5S DNA (Churchill *et al.*, 1990). In this model, alternate fingers of TFIIIA bind on one face of the DNA double helix in an equivalent manner, and with the same polarity, to the major groove, so that successive minor grooves are crossed by the linker peptides between fingers. The model is a general one, and does not currently explain the precise effects of local irregularities in protein or DNA structures. The effects of nucleotide substitutions in 5S DNA on TFIIIA binding affinity reported here are consistent with the loss of major groove hydrogen bonding interactions between TFIIIA and 5S DNA.

The higher-order structure of DNA is relatively simple compared with that of RNA. Proteins differentially bind linear DNA molecules mainly depending upon the specific sequence. Sequence recognition has been considered as a general characteristics of protein-DNA interactions. The best example of the specificity of this kind of recognition is the restriction enzyme-DNA interaction. In comparison, protein-RNA interactions can be dependent on the recognition of sequence as well as specific conformation. Because RNA has more conformational possibilities for evolutionary selection to choose, protein-RNA interactions are generally less dependent on sequence recognition than protein-DNA interaction.

However, DNA does undergo conformational changes. The most common conformations are the B form and A form, or a structure between these two forms. The

potential role of a specific DNA structure at any of the three interaction sites within the ICR cannot be addressed by the block-mutation results, since in each case several neighboring nucleotides were substituted simultaneously, making it unlikely that a sequence-specific DNA conformation, if it existed in the wild type 5S DNA, would be maintained in the mutant. However, the results from point mutation analysis indicates that the specific interaction of TFIIIA and 5S DNA is essentially dependent on the sequence. The placement of the three interaction sites within the 5S DNA further suggests that a special local DNA conformation is unlikely to be a feature of the TFIIIA-5S DNA interaction. A 5S RNA gene consists of short regions of highly conserved base pair sequences which encode the corresponding conserved sequences found in single stranded loops of the 5S RNA, flanked by more highly variable sequences which encode the non-conserved sequences of the base paired stems of the 5S RNA molecule. It is striking that all of the sequence-specific regions identified from our study to be important for TFIIIA binding (base pairs 57-62, 67-70 and 78-86) map to those base pairs in the gene which encode the variable sequences of base paired stems of the 5S RNA molecule. If the TFIIIA-5S DNA interaction pattern identified from our results is a universal feature of the expression of 5S RNA genes within the eucaryotic kingdom, it is clear that the interaction sites for TFIIIA molecules would be highly variable in sequence. This observation supports the view that a special, sequence-specific conformation in the DNA is not a general requirement for TFIIIA binding to 5S RNA genes. However, it does not rule out the possibility that a special DNA conformation is required specifically for the *Xenopus* TFIIIA-5S DNA interaction.

The sequence of RNA may affect the protein binding in two ways: nucleotides may be directly involved in the contacts or they may play a role in the contact by maintaining a proper RNA conformation required for the interaction. The TFIIIA footprint on 5S RNA extends from nucleotides +53 to +90, and formation of the correct conformation in this binding site requires the complementary sequence to this footprint (Andersen & Delihas,

1986; Andersen *et al.*, 1984; Christiansen *et al.*, 1987; Huber & Wool, 1986; Pieler & Erdmann, 1983; Romaniuk, 1985; Romaniuk *et al.*, 1987). Thus, there is a striking correspondence in the location of the TFIIIA binding regions on 5S RNA and the 5S RNA gene. A number of techniques have been employed to determine which features of the RNA are important to the TFIIIA-5S RNA interaction, including gel shift assays, chemical cross-linking, chemical modifications, and filter binding assays. Results obtained from the measurement of TFIIIA binding affinities of site-specific mutants of 5S RNA indicate that the major determinants for protein binding are primarily structural (Baudin & Romaniuk, 1989; Baudin *et al.*, 1990; Romaniuk, 1989; Romaniuk *et al.*, 1987; You & Romaniuk, 1990 & chapter 2). Virtually all of the mutations which significantly reduce TFIIIA binding affinity disrupt the structure of the 5S RNA around the junction of the three helical domains (Fig. 27). Chemical crosslinks between TFIIIA and 5S RNA are also clustered in this region of the RNA molecule (Baudin *et al.*, 1989). It is therefore apparent that the secondary and tertiary structures of the 5S RNA form the essential recognition features for TFIIIA, and there is currently no evidence to indicate that TFIIIA forms any strong sequence-specific contacts with nucleotides on the RNA.

Several models that have been proposed in attempts to explain how TFIIIA could interact specifically with both 5S DNA and 5S RNA suggest that TFIIIA makes identical contacts within similar conformational contexts on both nucleic acids. The data presented in this study allows an examination of this concept, and indicates that such models are unnecessary. It is apparent that TFIIIA binds specifically to 5S DNA by forming sequence-specific contacts within three discrete regions of the coding part of the gene. Substitution of the sequence at any of these sites significantly reduces TFIIIA binding affinity, by as much as 100-fold in the case of the Box C subregion. Although these experiments cannot address the question of whether a special conformation within the 5S DNA is one of the requirements for optimal TFIIIA binding, the placement of the target

sequences within highly variable sequences of 5S RNA genes does suggest that a special conformation is not an evolutionarily conserved feature of TFIIIA-5S DNA interactions in general.

In contrast, TFIIIA does not make any strong sequence-specific contacts with the 5S RNA. Although the protein binding sites on the 5S DNA and 5S RNA are coincident, those regions of the 5S RNA that have been identified as being moderately involved in TFIIIA binding do not correspond at all to the three specific TFIIIA interaction regions within the 5S DNA. It is evident that substitution of nucleotides 57-62 and 67-70 in the 5S RNA lowers TFIIIA binding affinity, but the effect of these mutations results from the disruption of base paired stems. Incorporation of compensating double mutations (e.g. 16-21/57-62) that restore the local base pairing in the 5S RNA results in the recovery of full TFIIIA binding affinity. In contrast, the decreased protein binding activity of the 57-62 mutation in the 5S DNA cannot be compensated for by the substitution of base pairs 16-21, thus illustrating another fundamental difference in the role of sequence vs. structure in the interaction of DNA and RNA with TFIIIA. Finally, substitution of nucleotides 78-86 in the 5S RNA leads to a disruption of a helical stem, but has virtually no effect on TFIIIA binding. Substitution of these nucleotides in the 5S DNA results in a very large reduction in the affinity for TFIIIA.

Our results suggest that the DNA and RNA binding activities of TFIIIA are distinct, even though there is apparently a single nucleic acid binding site on the protein. It is possible that the two activities have evolved separately from an ancestor TFIIIA molecule which exhibited general nucleic acid binding properties. This concept predicts that fingers (or regions within any one finger) were optimized for binding either to 5S DNA or 5S RNA. Directed mutagenesis studies of the TFIIIA protein, combined with detailed binding studies of the mutants to DNA and RNA, should provide useful information on this point.

The work presented in this chapter was a collaborated effort of Qimin You (Mutagenesis, Filter binding assays) and Nik Veldhoen (SAAB, Filter binding assays). Point mutant 5S DNA of *Xenopus borealis* was provided by Dr. Setzer of Case Western University, USA. A part of the mutants were from the laboratory's collection created by Dr. Paul J. Romaniuk, Qimin You, Isabel Stevenson and Florence Baudin.

CHAPTER 4
Ribosomal Protein L5-5S RNA Interaction

4.1. INTRODUCTION

Eucaryotic 5S RNA interacts with three structurally distinct protein groups. They are represented by La-antigen, TFIIIA and ribosomal proteins, respectively (see Chapter 1). In *E. coli*, three ribosomal proteins (L5, L18 & L25) bind to 5S RNA. In eucaryotes, L5 is the only ribosomal protein associated with 5S RNA. However, 5S RNA may make additional contacts with other ribosomal proteins once complexed within eucaryotic ribosomes (Ulbrich & Wool, 1978; Metspalu *et al.*, 1978; Kärgel *et al.*, 1987). No special structural motif has been identified in L5. In a computer analysis of the sequence of rat L5, Chan *et al.* (1987) did not find any repeat sequence reminiscent of zinc finger domains and there is no evidence that L5 requires zinc for its activity. Comparison of L5-related ribosomal protein sequences among *E.coli.*, *H. cutirubrum* and eucaryotes revealed that there is a conserved region rich in basic amino acids located near the N-termini of these proteins (Kenmochi *et al.*, 1991), but the functional significance of this region, if any, has not been elucidated. It is not clear how the three types of structurally unrelated proteins make specific contacts with the same 5S RNA molecule (for a rewiew, see Bandziulis *et al.*, 1989).

On the other hand, 5S RNAs from different organisms bind to the same ribosomal proteins. Reconstitution experiments with *Bacillus stlearothermophilus* 50S subunits and 5S RNAs from different organisms showed that 5S RNAs other than that of *B. stearothermophilus* itself can be incorporated into the *Bacillus* ribosome and yield biologically active particles. These eubacteria 5S RNAs were from organisms including *E.*

coli, Pr. vulgaris, B. subtilis, S. avreus, M. lysodeicticus, Ps. fluorescens, A. vinelandii (Wrede & Erdmann, 1973) and *B. licheniformis* (Raué *et al.*, 1981). The list was later extended to include spinach chloroplast 5S RNA (Vogel *et al.*, 1984), caulobacter 5S RNA (Erdmann *et al.*, 1986), archaebacteria *Halobacterium* and *Thermoplasma acidophilum* 5S RNAs and eucaryote 5S RNAs of *Saccharomyces carlsbergensis* and *Equisetum arvense*, but the 50S ribosomal subunits reconstituted with eucaryotic 5S RNAs showed significantly reduced biological activity (Erdmann *et al.*, 1986). The fact that so many heterologous 5S RNAs with wide sequence diversity can be incorporated into the eubacterium ribosomal 50S subunit indicates great flexibility in the 5S RNA sequence that ribosomal protein(s) contact. This view was supported by experiments in which heterologous 5S RNA-ribosomal protein complexes were constructed (Horne & Erdmann, 1972; Erdmann, 1976; Erdmann *et al.*, 1982). For instance, *E. coli* L5, L18 and L25 will bind to 5S RNAs from archaebacteria, and *B. stearothermophilus* 5S RNA binding proteins also recognize *E. coli* 5S RNA.

The interaction of 5S RNA and L5 is crucial for the formation of the ribosome. The L5-5S RNA complex is the presumptive precursor to ribosome assembly in mammalian somatic cells (Steitz *et al.*, 1988), in yeast (Brow & Geiduschek, 1987), and in *Xenopus laevis* (Wormington, 1989, Allison *et al.*, 1991). Our understanding of the nature of this protein-RNA interaction is very limited. Efforts have been made to determine the ribosomal protein binding site(s) on the 5S RNAs of *E. coli*. By using the cytotoxic nuclease α-sarcin, the *E.coli* L5 protein protection region was found to be near both the 5' and 3' ends of the 5S RNA stem I region (Huber & Wool, 1984; Huber & Wool, 1986). The L18 binding site was determined to lie in the loop A/stem II/loop B region, and L25 protects a site including stem IV, V and loop D, E from nuclease attack and chemical modifications (Garrett *et al.*, 1981; Egebjerg *et al.*, 1989). There is evidence that stem III and loop C may also be involved in L18 binding (Speek & Lind, 1982; Huber & Wool,

1984; Christiansen & Garrett, 1986; Egebjerg *et al.*, 1989). A protein-RNA cross-linking experiment (Oβwald *et al.*, 1990) indicated that *E. coli* L5 cross-linxs to 5S RNA at position 34-41 in loop C. Thus the combination of the three *E. coli* ribosomal protein binding sites covers almost the entire 5S RNA molecule (Fig.39). Similar results were obtained from archaebacteria *H.cutirubrum* (HL13 and HL19) (Schnier & Faist, 1985), and rat L5 (Huber & Wool, 1986) interactions with 5S RNA: The rat L5 and *H. cutirubrum* H13, H19 ribosomal proteins protect 5S RNA sites that are similar to each other and also close to the sum of the protected sites observed for *E. coli* ribosomal proteins L5, L18 and L25 on *E. coli* 5S RNA.

It is interesting that there are three 5S RNA binding ribosomal proteins in bacteria (L5, L18 & L25 in *E. coli*), two in archaebacteria (HL13 & HL19 in *H. thermoplasma*) and only one in eucaryotes (YL3 in yeast, L5 in rat & *Xenopus*), and they all bind to the same RNA at similar sites. It has been noted that YL3 has some amino acid sequence homology to the L5 and L18 of *E. coli*. Nazar *et al.* (1979) suggested that the yeast YL3 (or L5 in other eucaryotes) might be a fusion of the three procaryotic ribosomal proteins, developed in the course of evolution. This proposal explains the similarity of various ribosomal protein binding sites on 5S RNA, and implies that the eucaryote L5 may have domains that independently contact 5S RNA subregions.

The parameters for the interaction of *E. coli* L5, L18 and L25 with 5S RNA were estimated by Spierer *et al.* (1978) using a filter binding assay. The strongest interaction was observed between L18 and 5S RNA with a Ka value of 2.3×10^8 M^{-1}. The Ka values for L25 and L5 were 1.5×10^7 M^{-1} and 2.3×10^6 M^{-1}, respectively. The affinities of L18 and L5 for 5S RNA differed by 100-fold. It was found that interactions involving L5 and L18 possesses a strongly cooperative character (Feunteun *et al.*, 1975; Spierer &

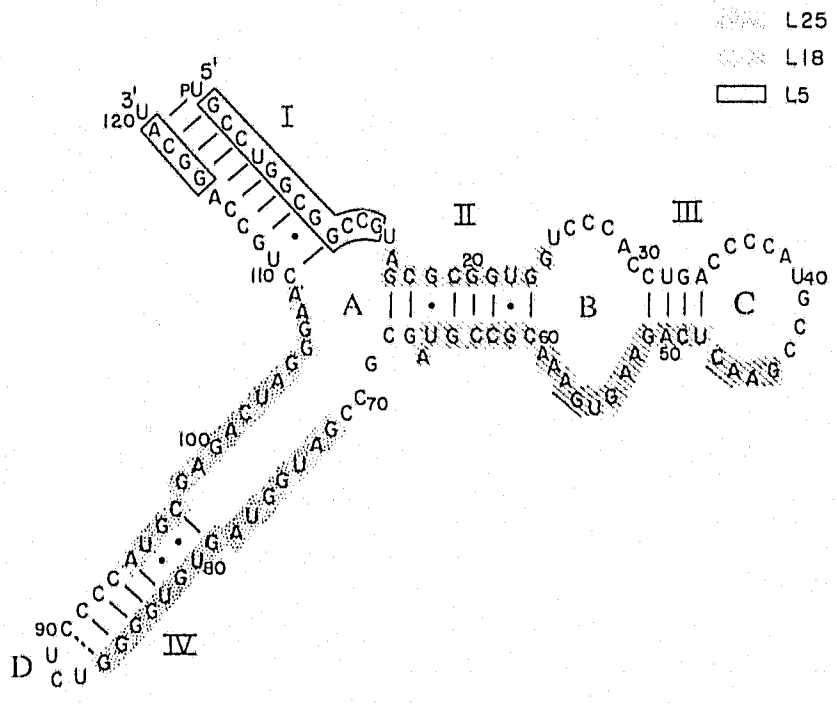


Figure 39: The binding sites for *E. coli* ribosomal proteins L5, L18, and L25 on 5S RNA determined with α-sarcin (after Huber & Wool, 1984).

Zimmermann, 1978; Spierer *et al.*, 1978). Circular dichroism (CD) studies showed that the L18-5S RNA complex has a 20-25% higher amplitude at wavelength 268 mM than the free 5S RNA, indicating that a protein-induced shift in secondary and/or tertiary structure occurs in the 5S RNA when it is complexed with L18. No similar structural shift was observed in the interactions of 5S RNA with L5 or L25. The binding affinity of L5 to 5S RNA was found to increase roughly an order of magnitude as the result of L18 cooperative stimulation.

In summary, studies on *E. coli* ribosomal protein-5S RNA interactions revealed the ribosomal proteins bind to extensive 5S RNA sites, and the secondary/tertiary structure appears to be more important than the sequence.

Much less is known about L5-5S RNA interactions in eucaryotes. Allison *et al.* (1991) demonstrated nucleotides 11-108 of *Xenopus* 5S RNA are sufficient for the binding of L5 protein *in vivo*. In another *in vivo* experiment, Guddat *et al.* (1990) observed that L5 binding is notably insensitive to a number of mutations of *Xenopus* 5S RNA, including large block deletions, substitutions or insertions of nucleotides. In contrast, some of these same mutations dramatically affected the binding of TFIIIA to 5S RNA.

A detailed study of the *Xenopus* L5-5S RNA interaction has been difficult to achieve, partly because of the poor solubility of the protein. We have developed an expression-purification system that yields highly purified, soluble L5 protein in sufficient quantity for *in vitro* study of the L5-5S RNA interaction. The properties of the equilibrium binding of *Xenopus* L5 to 5S RNA were characterized using a nitrocellulose filter binding assay. By using a set of 5S RNA substitution/deletion mutants, we also studied the sequence and structural features of 5S RNA required for the specific binding of L5. The data we obtained shows that the interaction of *Xenopus* L5-5S RNA is unusually insensitive to significant sequence and structure disruptions throughout the entire 5S RNA molecule.

4.2. MATERIALS AND METHODS

4.2.1. Construction of Expression Plasmid

Plasmid spL5AT was kindly provided by Dr. Wormington (see Wormington, 1989). In order to create a Nde I restriction site for subcloning the gene into the expression plasmid pET-16b, which contains a histidin tag for affinity column purification of the fusion protein, PCR mutagenesis was performed. The 5' and 3' primers for PCR were synthesized using a Biosearch DNA synthesizer. The 5' primer matches the 21 nucleotides around the ATG codon of the L5 gene, except for the -1 A, which was substituted with a T to create a CATATG Nde I restriction site. The 3' primer was complementary to a sequence downstream of a Kpn I site within the gene. The sequences of the primers are as following:

Oligo 91-1, L5-5'

5' AAATAGAGC<u>CATATG</u>GGGTTC 3'
A
Nde I

Oligo 91-2, L5-3'

5' TGGATCACCAGTCGCTTG 3'

The PCR reaction was carried out in 10 mM Tris:HCl, pH 8.3 at 20 °C, 1.5 mM $MgCl_2$, 25 mM KCl, 50 µg/ml gelatin, 50 µM dNTPs, 0.1 µg/ml DNA template, 0.2 µM each forward and reverse primers and 25 U/ml Taq DNA polymerase.

After 27 cycles of reaction, the PCR product was flush ended with the Klenow fragment of DNA polymerase I and then blunt-end ligated with DNA ligase. The self-ligated DNA fragments were successfully digested with Nde I and Kpn I, producing the 5' portion of the L5 gene with appropriate "sticky" ends. The 3' portion of the L5 gene was obtained by the digestion of plasmid spL5AT with Kpn I and Bam HI. The 5' and 3'

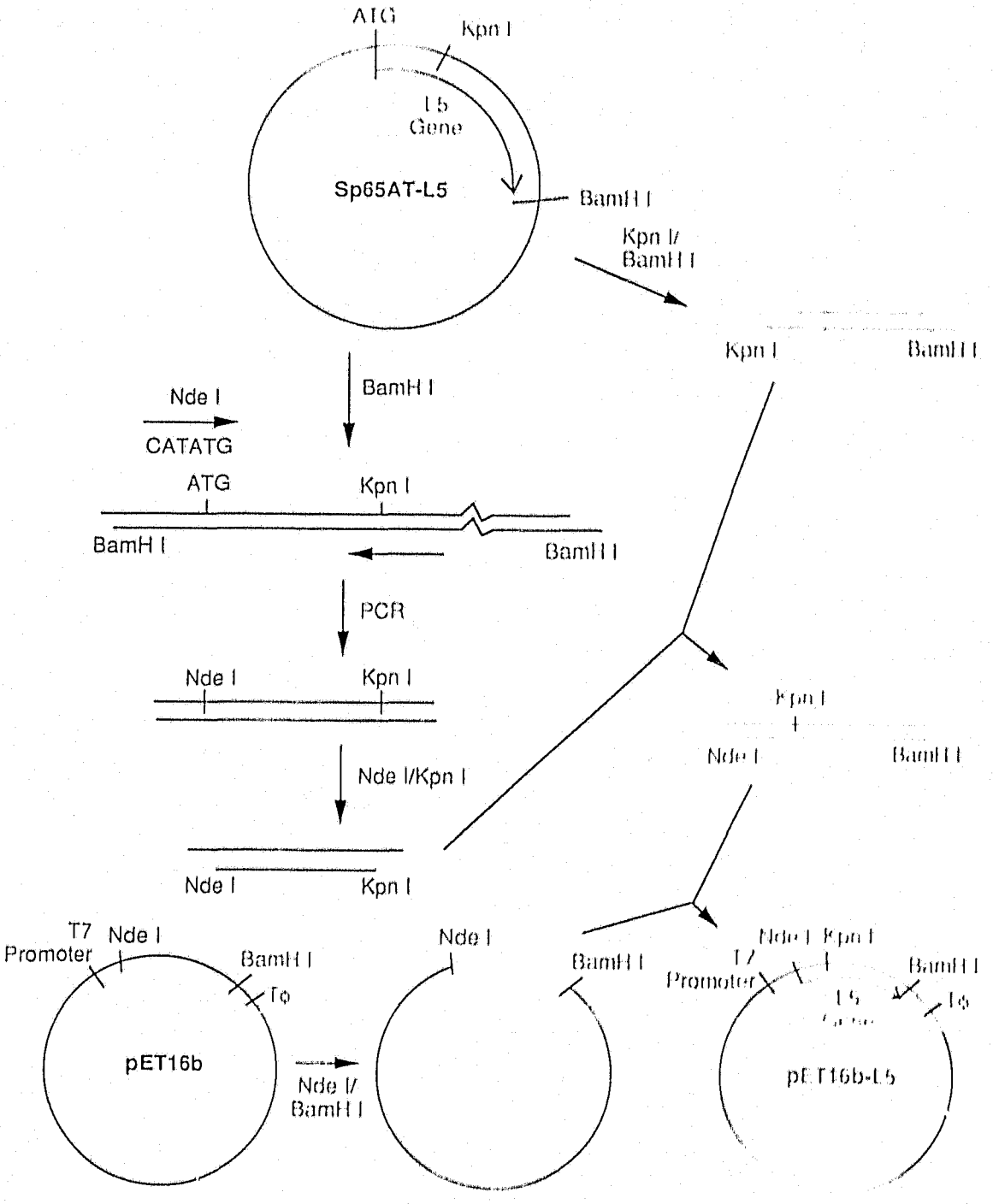


Figure 40: Construction of pET-L5 plasmid.

fragments of the L5 gene were then ligated into pET-16b that had been digested with Nde I and Bam III. The construction of the pET-L5 plasmid is summarized in Figure 40.

Construction of the desired clone was confirmed by restriction site analysis and DNA sequencing.

4.2.2. Construction of 5S RNA Mutant Genes

The construction of various 5S RNA mutant genes in the plasmid pUC18, 5S RNA *in vitro* synthesis and ^{32}P labelling are described elsewhere (Baudin & Romaniuk, 1989; Romaniuk, 1989; Romaniuk *et al.*, 1987; You & Romaniuk, 1990, also see Chapter 2). Detailed description of these mutations is presented in Results.

4.2.3. Expression and Purification of L5 From *E.coli* strain BL21(DE3)

Plasmid pET-L5 was used to transform *E.coli* strain BL21(DE3), a lysogen carrying the gene for T7 RNA polymerase under control of the lac operon. Bacteria were incubated at 37 °C in LB medium supplemented with ampicillin, until an A_{600}=0.6 was reached. IPTG was added to a final concentration of 1.0 mM. Cells were grown for 3 more hours, and then harvested by centrifugation at 3,000 x g in a Beckman JA-20 rotor for 10 minutes. The following procedures were all performed at 4 °C. Cells were washed with 10 ml of IMAC-5 buffer (20 mM Tris:HCl pH 7.9 at 4 °C, 0.5 M NaCl, 10% glycerol, 1.0 mM PMSF, 5 mM Imidazole), and resuspended in 10 ml of the same buffer. Cells were lysed by sonication, and then cell debris was pelleted by centrifugation at 40,000 x g in JA-20 rotor for 30 minutes. The supernatant was filtered using a Whatman filter paper before loading onto the His-Bind Resin affinity column (Novagen, 69670-1, #1). The principle of this protein purification procedure is based on the immobilized metal affinity chromatography (IMAC). The His-Tag-L5 fusion protein expressed from pET 16b-L5 plasmid binds strongly to the immobilized nicle atam on the column, and can only be eluted by buffers containing high concentrations of imidazol or EDTA.

The column was prepared as follows: The resin was packed under gravity flow. 1.0 ml settled resin was sufficient for our purpose. The column was then washed with 3 volumes of 0.1 M EDTA, followed by 3 volume dH_2O. 5 volumes of 50 mM $Ni(II)SO_4$ was loaded on the column, and the column was then equilibrated with 3 volumes of IMAC-5 buffer. The filtered sample was loaded onto the column, followed by successive washes with 4 volumes of IMAC-5, 4 volumes of IMAC-70 (IMAC with 70 mM Imidazole), 4 volumes of IMAC-100 and 4 volumes of IMAC-150. The L5 protein was eluted by IMAC-EDTA (IMAC with 0.1 M EDTA, no imidazole). Protein was collected in 500 µl fractions. The concentration of purified L5 was determined by a Bradford method assay (Bradford, 1971). The purity of the samples were checked by SDS-PAGE (Fig. 41). L5 retains full 5S RNA binding activity for a week when stored at 4 °C in the column elution buffer.

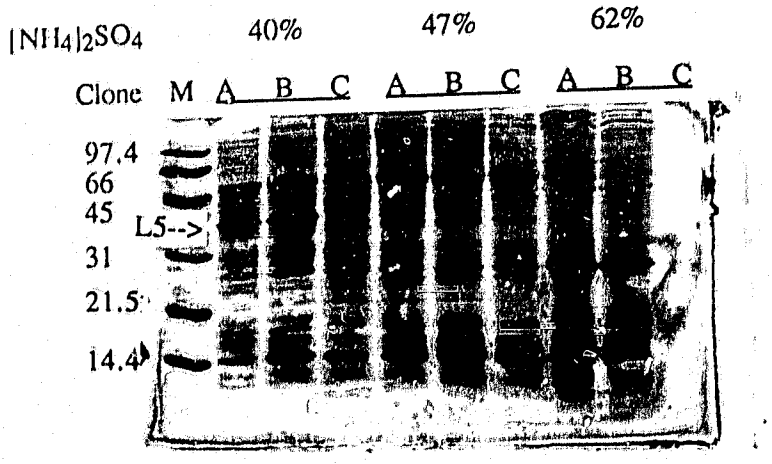
4.2.4. The Determination of the Equilibrium Binding of L5 to 5S RNA

The equilibrium constants for the binding of radioactively labelled wild type 5S RNA and 5S RNA mutants were determined using a nitrocellulose filter binding assay as described by Romaniuk for the TFIIIA-5S RNA interaction (Romaniuk,1985 & Chapter 2). The competition assays were also described in the same paper. The TFIIIA-5S RNA interaction data have been published previously (Baudin & Romaniuk, 1989; Baudin *et al.*, 1990; Romaniuk, 1989; Romaniuk *et al.*, 1987; You & Romaniuk, 1990 & Chapter 2).

4.3. RESULTS

4.3.1. Expression and Characterization of Recombinant L5 Protein

The L5 cDNA cloned into the pET-L5 vector was sequenced, to ensure that the gene was inserted correctly. Figure 41 shows that L5 protein of the expected molecular



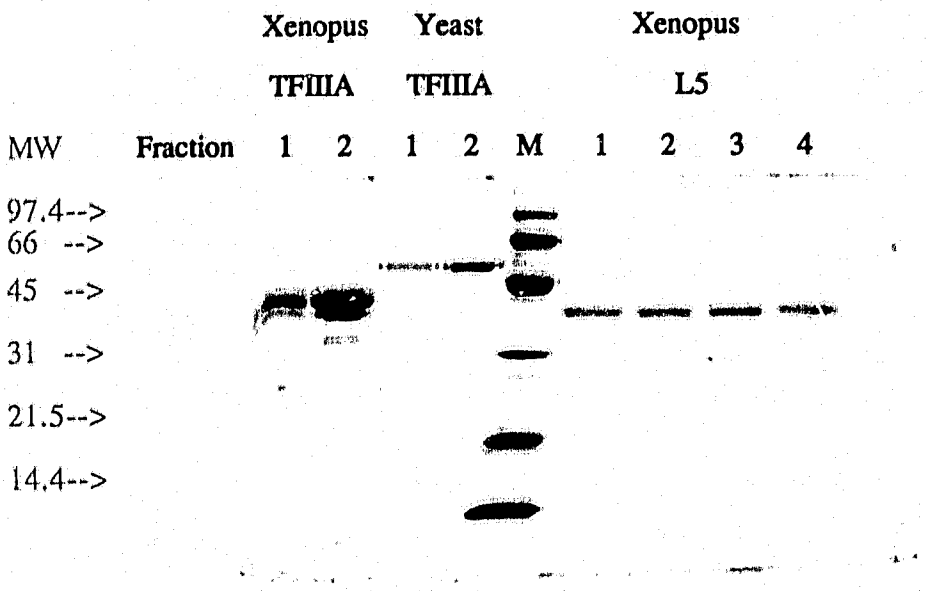


Figure 41: Purification of *Xenopus* ribosomal protein L5 expressed in *E. coli*
Top. induction of cloned L5 expression. Clone A and B contain L5 gene, C is a control cell strain transformed by the plasmid without the gene. Ammonium sulfate precipitations performed at 40%, 47% and 63%. L5 was presented in the 40% Amino sulfate precipitant. Bottom. purified L5 protein after the affinity column. The molecular weight of L5 is compared with those of *Xenopus* TFIIIA (Kd=39.5) and *Yeast* TFIIIA (Kd=50).

weight was produced only in the induced culture. To verify the identity of recombinant L5, amino acid sequence data were obtained. The recombinant protein could not be sequenced directly from the N-terminus presumably because of the presence of a blocking group. The protein was treated with cyanogen bromide, producing three major peptide bands on a SDS-polyacrylamide gel. Using one of these peptides, we obtained a sequence of 16 amino acids (MRLLIEEDEDAYKKQF) that align with amino acids 208-223 of the published L5 sequence (Wormingon, 1989), with one difference. There is a single mismatch: an isoleucine (underlined) in the peptide sequence was predicted to be a methionine from the published cDNA sequence. This apparent mismatch may have resulted from a point mutation at the wobble position of the codon, or may represent an error in the original cDNA sequence. Alignment of known eucaryotic ribosomal L5 sequences indicates that this methionine is not conserved.

The specific binding activity to 5S RNA of the expressed L5 was tested by nitrocellulose filter binding assays (Fig. 42) and by gel retardation (Fig. 43). An RNA excess binding assay was designed to further investigate the activity of this protein in 5S RNA binding. Figure 44 shows that the purified L5 is 100% active and that one L5 molecule interacts with one 5S RNA. Factor Xa (Novagen) was used to remove the histidine tag from the fusion protein, and the 5S RNA binding affinity of cleaved and uncleaved L5 were tested. No significant difference in the Ka values were found (Ka for L5 with His tag: $0.40 \times 10^9\ M^{-1}$, and for L5 without His tag: $0.47 \times 10^9\ M^{-1}$). These experiments allow us to conclude that the L5 protein is correctly expressed in the *E.coli* and that the protein retains complete 5S RNA binding activity throughout the purification process.

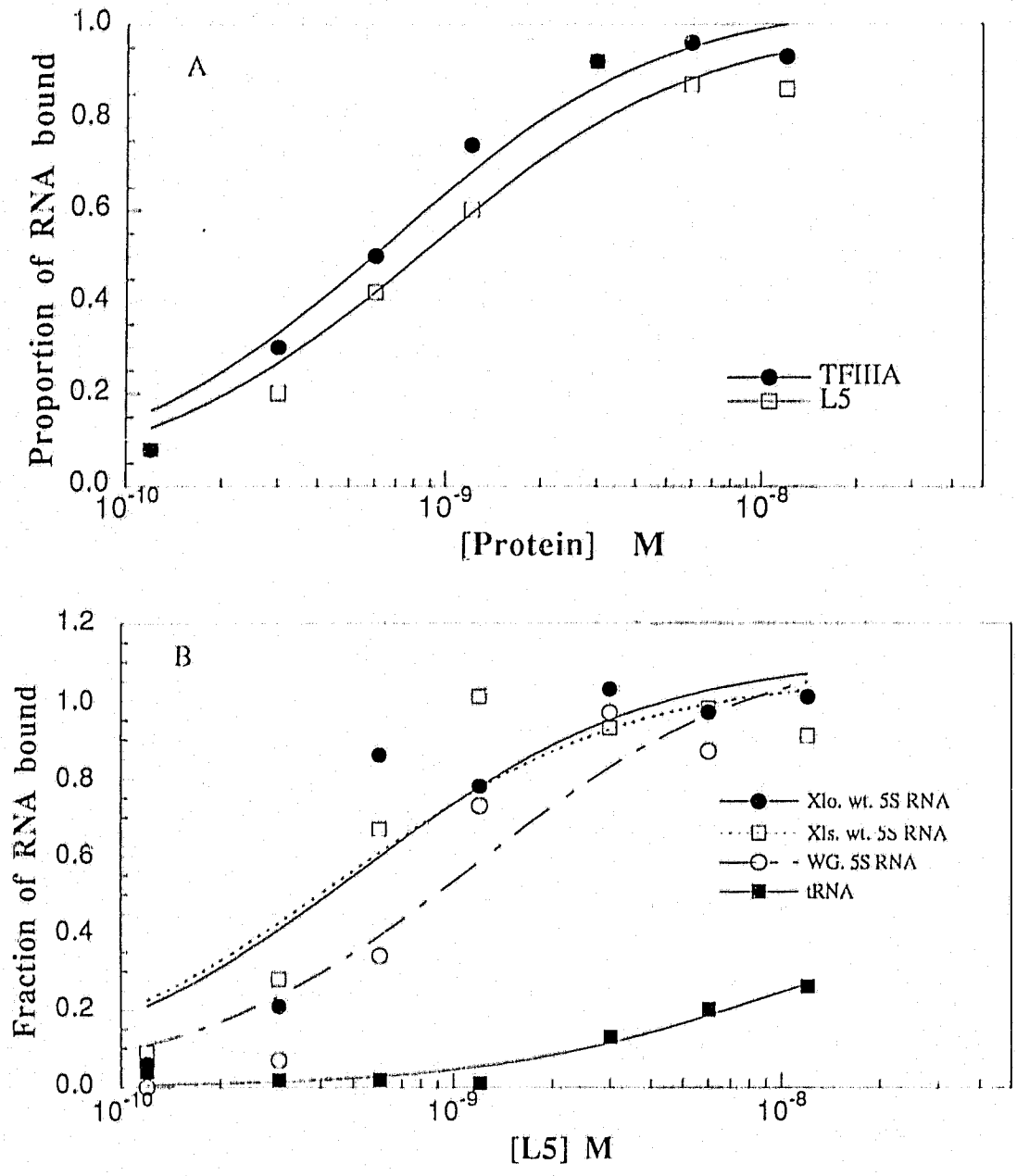


Figure 42: The specific binding of L5 to 5S RNA.
A: Titration of 5S RNA with L5 and TFIIIA under equilibrium binding conditions.
B: Binding of L5 to various RNA species.

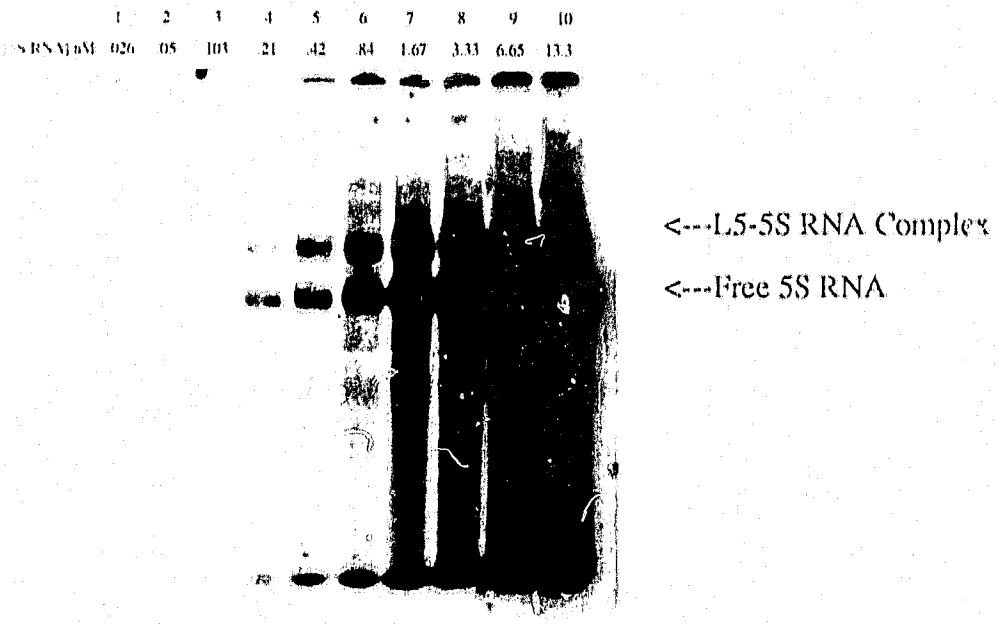


Figure 43: Gel retardation of L5-5S RNA complex. [L5] was kept at 2.0 nM. Complexes pre-formed with increasing concentrations of ^{32}P labelled 5S RNA were loaded on 10% non-denaturing acrylamide gel. Electrophoresis was performed at 4 °C.

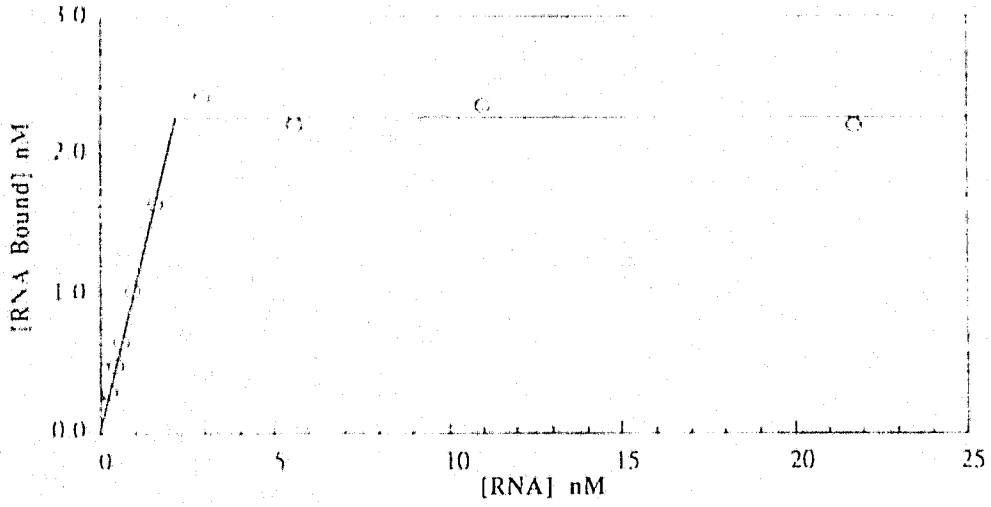


Figure 44: RNA Excess Binding Assay. L5 protein concentration was kept at 2.0 nM and the concentration of labelled 5S RNA was kept at 0.25 nM in all fractions. Unlabelled 5S RNA was added to the concentrations as indicated. The binding mixes were filtered and the radioactivity was counted as described previously. The amount of bound RNA was calculated as: [Bound RNA]=[Bound labelled RNA] x Folds of dilution with unlabelled RNA.

Table 8. Comparison of L5 and TFIIIA binding affinities to various RNA species.

	Binding affinities relative to Xlo. wt. 5S RNA	
	TFIIIA[a]	L5
Xlo. wt. 5S RNA	1.00	1.00
Xls. wt. 5S RNA	1.67	1.56
Wheat germ 5S RNA	2.7	0.58
Yeast $tRNA^{phe}$	0.01	0.018
Xlo. wt. 5S DNA	2.0	0.034
Drosophila wt. 5S RNA	ND	1.08
Drosophila 121 CG	ND	1.07

[a]Data taken from Romaniuk (1985).

4.3.2. Characterization of L5-5S RNA Interaction

4.3.2.1. Equilibrium constants and the binding specificity

Association constants (Ka) of the protein-RNA interaction in the filter binding assays were determined by measuring the L5 concentration at which 50% of saturation is achieved, provided that 100% of the protein is active. As indicated by Figure 44, the purified L5 meets this requirement. The L5-5S RNA association constant was estimated to be $2.0 \pm 0.5 \times 10^9\ M^{-1}$ under standard binding conditions. This value is close to the Ka of $1.0 \times 10^9\ M^{-1}$ for the TFIIIA-5S RNA interaction (Romaniuk, 1985).

Using filter binding assays to determine association constants involves several assumptions that can be tested: that complexes do not dissociate during filtration and that free protein bound to filters will not trap free RNA during filtration (Giacomoni, 1981a, b; Carey *et al.*, 1983; Romaniuk, 1985). We tested the possibility that once bound, L5-RNA complexes may dissociate during the completion of the filtration process. When measuring dissociation rates of L5-RNA complexes, one set of reactions was washed after filtration with 400 µl TMK (lacking BSA), while the other set of reactions was filtered without further washing. The dissociation rates measured from these two experiments were virtually identical, indicating that once bound to nitrocellulose filters, L5-5S RNA complexes are stable to dissociation during the filtration process. The other concern was that the free protein on the filter may trap free 5S RNA. This possibility was tested by first filtering free L5 solution and then followed by a free 5S RNA solution. At L5 concentration of less than 60 nM, less than 15% of free 5S RNA was trapped by protein pre-bound on the filter.

The specificity of the RNA binding activity of L5 was investigated by measuring the binding affinities of L5 for *Xenopus laevis* oocytic wild type 5S RNA (Xlo WT), *Xenopus* somatic 5S RNA (Xls WT), wheat germ 5S RNA (WG 5S RNA) and yeast

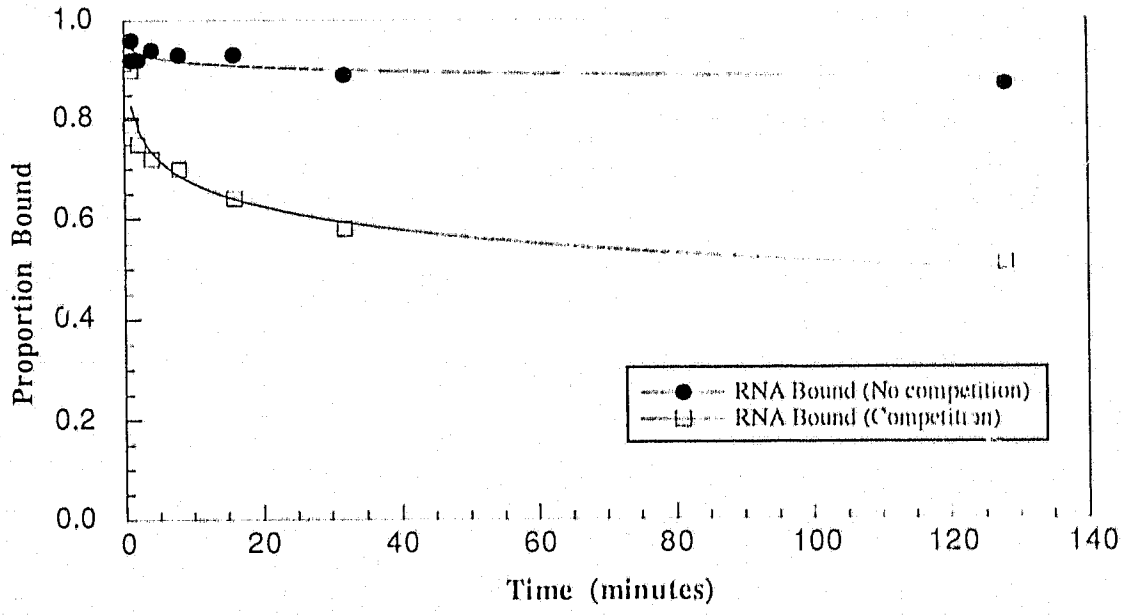


Figure 45: Dissociation of L5-5S RNA complex formed between 0.5 nM 5S RNA and 20 nM L5 protein. 0.4 μM unlabelled 5S RNA was added in the competition reaction.

$tRNA^{phe}$ (Fig. 42). The relative Ka values of these RNAs are compared with those reported previously for TFIIIA (Table 8). L5 has an RNA binding specificity similar to TFIIIA, with the exception of WG 5S RNA. Apparently, both proteins have no specific affinity for $tRNA^{phe}$ in the binding assays. In a competition assay, $tRNA^{phe}$ showed very weak competition strength against Xlo WT for L5 binding (Fig. 51). In separate experiments, the binding of L5 to Xlo 5S DNA and both precursor and mature *Drosophila* 5S RNA were tested (Table 8). The L5-5S DNA interaction has a Ka value that is 30-fold lower than the Ka of L5-Xlo 5S RNA, resulting from a non-specific nucleic acid binding activity of L5. In contrast, both mature and precursor *Drosophila* 5S RNA bind to L5 as strongly as the Xlo WT 5S RNA.

4.3.2.2. Dissociation of the L5-5S RNA complex

The dissociation of the homologous complex was measured by pre-forming complexes between ca. 0.5 nM ^{32}P-labelled *Xenopus* 5S RNA and 20 nM L5 protein at 20 ^{o}C in standard TMK buffer, supplemented with the addition of 200 U/ml RNasin. After a 15 minute incubation, an aliquot was removed, filtered and the filter washed with 400 µl TMK lacking BSA. Unlabelled *Xenopus* 5S RNA was added to a final concentration of 0.4 µM and aliquots were removed at the indicated times and filtered with washing. In the control experiment, buffer was added at time zero rather than unlabelled 5S RNA. Figure 45 shows the results. A rapid dissociation of ca. 19% of the complexes occurred during the first few minutes and the remainder of the complexes appear to be quite stable. The initial dissociation rate for the slow dissociation population of *Xenopus* L5-5S RNA complex was calculated to be 2.97×10^{-5} s^{-1}, Compared with the estimated rate constant of 4.5×10^{-4} s^{-1} for the dissociation of the TFIIIA-5S RNA complex (Romaniuk, 1985), the L5-5S RNA complex appears to be much more stable. After an extended 24 hour incubation at room temperature, 72% of the complexes in the control solution remained

bound and in the competition reaction, 24% of labelled 5S RNA was still bound to L5, despite the presence of a twenty-fold excess of unlabelled 5S RNA (data not shown).

The dissociation of the complex was also measured by simply diluting the complexes 25 fold rather than adding unlabelled RNA as competitor. The kinetic pattern measured in this way is indistinguishable from that measured by the first method (data not shown). This result ruled out the possibility that the incomplete dissociation observed was result of an inability of the unlabelled RNA to bind to free protein.

4.3.2.3. The optimal binding condition

The experimental conditions were explored for the optimal pH, temperature and ion strength for the L5-5S RNA binding interaction. All of these experiments were performed using standard filter binding assays, except for the necessary changes made for each purpose. In the pH dependence experiment, the standard TMK pH (7.5) was substituted by a buffer with appropriate pKa. Binding assays under different pH values were performed. It was found that the L5-5S RNA interaction has a broad pH range of 6.0 to 8.0, and shows a sharp decrease in the binding affinity at pH values higher than 8.0 (Fig. 46), which is similar to the TFIIIA-5S RNA interaction (Romaniuk, 1985).

To determine the optimal temperature for the binding, the pH of binding buffers was carefully adjusted for individual differences in the temperature dependence of pKa, and reactions were carried out at 4 °C, 15 °C, 22 °C, 30 °C and 37 °C, respectively. The filtrations were also performed at the incubation temperature and the filters were washed with 400µl of TMK buffer (lacking BSA) equilibrated at the incubation temperature. The results show there is no significant difference in the Ka values at temperatures below 22 °C. However, an obvious decrease in the Ka value is observed when the temperature reaches 30 °C and above (Fig. 47).

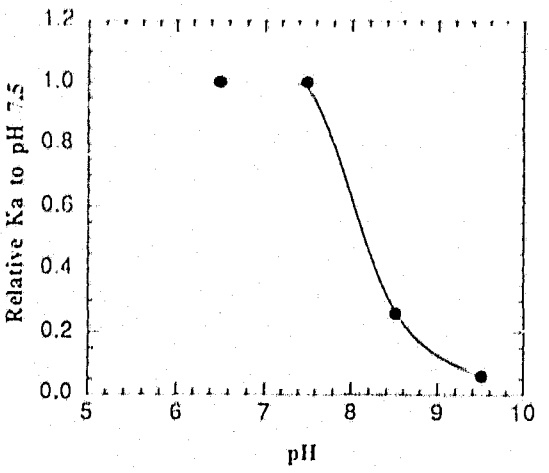


Figure 46: pH dependence of Ka.

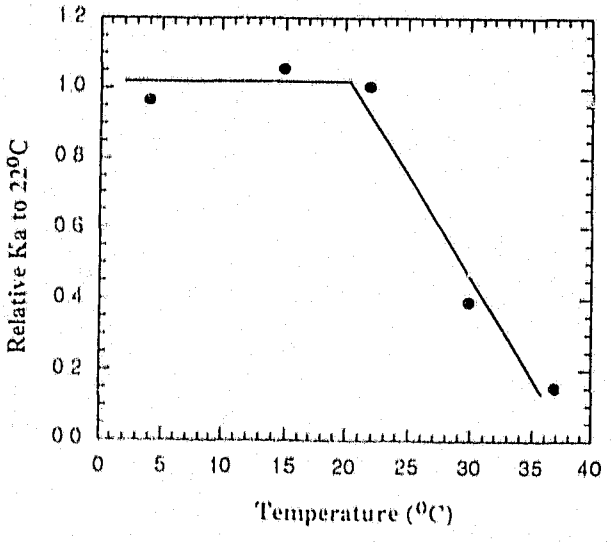


Figure 47: Temperature dependence of Ka.

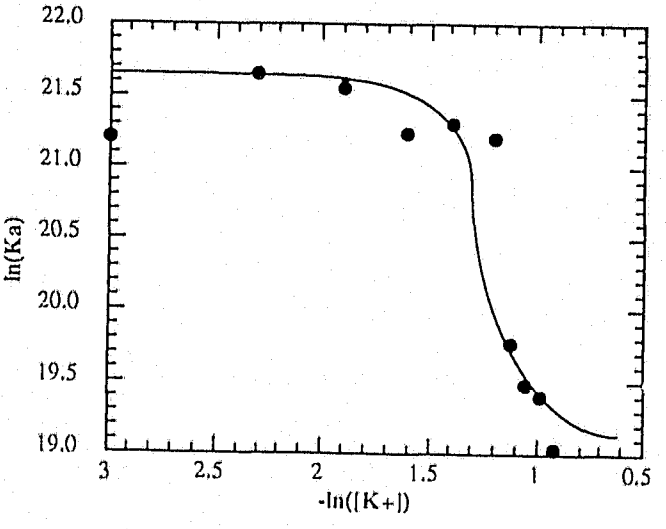


Figure 48: Ionic strength dependence of Ka.

In the experiments to test the ionic strength dependence of Ka, the concentration of monovalent salt in the TMK buffer was varied between 0 and 0.4 M. Figure 48 shows that in the 0 to 0.25 M range, L5 binds well to 5S RNA. At salt concentrations above 0.25 M, the Ka value decreased significantly, and virtually no binding activity was observed when the monovalent salt concentration reached 0.4 M.

From these investigations, we determined that optimal 5S RNA binding conditions for L5 are: TMK buffer containing 0.1 M KCl, at room temperature and pH 7.5.

4.3.2.4. The mutagenesis analysis

A large number of mutants have been created from the *Xenopus* oocytic 5S RNA. These substitution/deletion mutations span almost the entire 5S RNA molecule. As figure 13 shows, the secondary structure of *Xenopus* oocyte 5S RNA consists of double helical stems, single stranded loops and several bulged nucleotides. Any of these special structures may play a role in the specific interaction with L5 and therefore are subject to the mutagenesis analysis.

Besides the stem mutants described in Chapter 2, we also tested interactions of L5 with the other mutants created on the 5S RNA. Loops of 5S RNA secondary structure were also target for mutagenesis. In addition to simple nucleotide replacements, where the loop structure was retained, some mutants were designed to alter the native structure of the target loop. Mutant 22-25 changed the nucleotides on one half of loop B, so that they complement the nucleotides on the other half of the loop. This mutant created a long stem region including helix II, loop B and helix III. In loop C, nucleotides 33-36 were made complementary to nucleotides 41-44 on the opposite position of the loop by replacing positions 33 and 34 with appropriate nucleotides. The resulting mutant, named 33-34, extends stem III into loop C (Fig.49). Finally, loop E was eliminated in mutant 96-101 and both loop E and B were "closed" in the combination mutant 22-25/96-101.

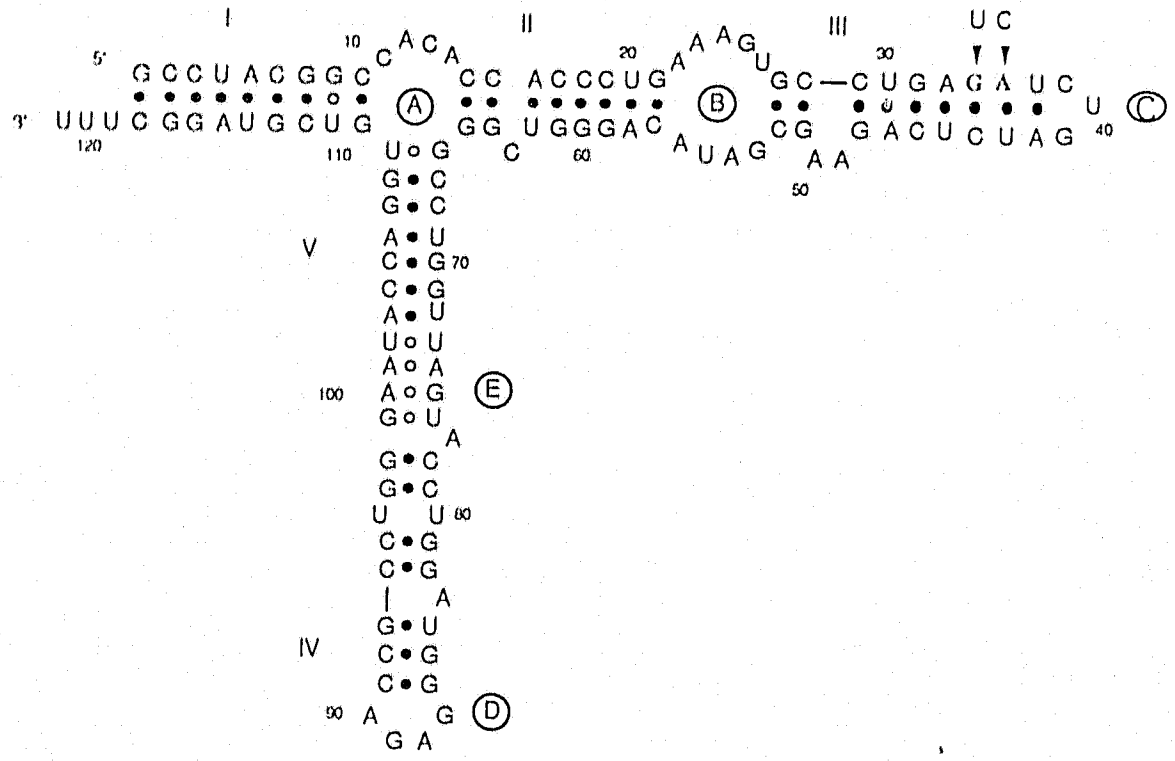


Figure 49: An example of loop to stem mutations. Loop C is "closed" by substitutions at nucleotides 33-34. Loop E is also shown as the proposed stem structure.

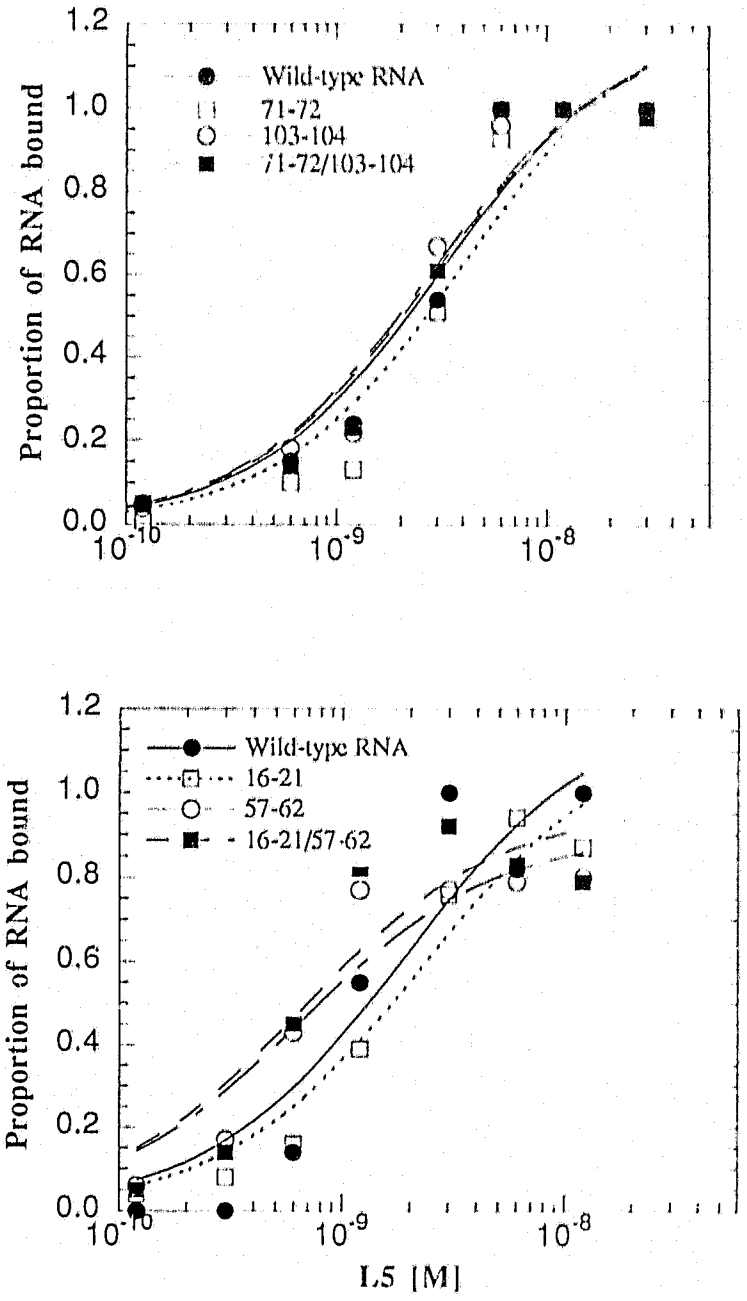


Figure 50: Standard filter binding assays for L5-5S RNA mutant interactions.

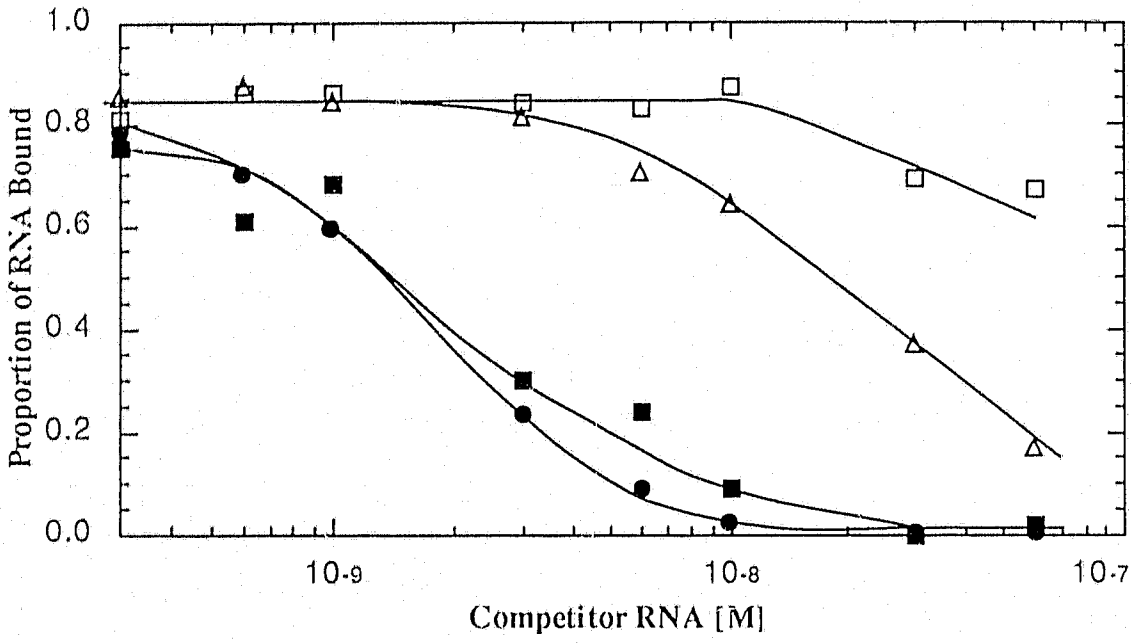


Figure 51: Competition binding assay. Unlabelled Xlo. 5S RNA was used as competitor. The concentration of L5 was kept at 2 nM. Legend:

- ● wt. 5S RNA
- □ tRNA
- ■ 1-68
- △ 50/51-120

Table 9. Relative L5 and TFIIIA Binding Affinities for 5S RNA Mutants

Mutant	relative binding affinity L5	TFIIIA
wt. 5S RNA	1.00	1.00
Stem II		
14-15	1.11±0.05	0.85±0.22
64-65	1.02±0.13	0.74±0.24
14-15/64-65	1.00±0.28	1.11±0.32
16-21	0.82±0.05	0.32±0.15
57-62	0.99±0.03	0.40±0.15
16-21/57-62	0.94±0.13	1.09±0.48
Stem III		
27-32	0.52±0.15	0.75±0.10
45-52	0.54±0.12	0.76±0.12
Stem IV		
78-81	1.11±0.13	0.88±0.01
95-98	0.99±0.10	0.78±0.02
78-81/95-98	1.00±0.26	0.86±0.01
82-86	1.28±0.19	0.81±0.30
91-94	1.06±0.09	0.96±0.18
82-86/91-94	1.42±0.26	1.21±0.35
Stem V		
67-70	1.18±0.11	0.75±0.12
105-108	1.11±0.01	0.39±0.06
67-70/105-108	1.03±0.07	0.71±0.01
71-72	0.88±0.17	0.35±0.21
103-104	1.08±0.08	0.50±0.23
71-72/105-108	1.08±0.10	1.18±0.32
Loop A		
10-13	1.19±0.23	0.30±0.01
Loop B		
22-25	0.72±0.19	0.84±0.12
22-26	0.86±0.05	1.00±0.02
53-56	0.88±0.24	1.62±0.12
Loop C		
33-39	0.92±0.41	1.00±0.02
33-34	0.73±0.29	1.09±0.21
41-44	1.03±0.22	0.40±0.10
Loop D		
87-90	1.20±0.02	0.71±0.10
Loop E		
73-76	0.77±0.25	0.57±0.02
96-101	0.63±0.04	0.59
99-101	0.77±0.29	0.61±0.08
Deletion Mutants		
Δ49-50	0.91±0.02	1.00±0.02
Δ63	1.05±0.18	1.00±0.02
Δ83	1.22±0.25	1.00±0.02
nt. 1-68	0.58±0.08	0.48±0.21
nt.51-120	0.76±0.16	
Combination Mutants		
16-21/95-98	0.59±0.12	0.15±0.03
16-21/67-70 /95-98	0.62±0.02	0.16±0.05
16-21/57-62 /78-81/95-98	0.81±0.19	0.63±0.08
22-25/96-101	0.64±0.06	

Bulged nucleotides were deleted in mutants ΔA49-50, ΔC63 and ΔA83, respectively. Large block deletions were created in truncated 5S RNA mutants: nt. 1-68, nt. 50/51-120 and nt. 66-108. The first two mutants deleted the 3' or 5' end regions of the 5S RNA molecule, which were not included in the substitution mutagenesis.

All of the 41 5S RNA mutants were subjected to filter binding assay with L5 protein. In each experiment, 3 related mutants (ie. two single and one double mutants at the same site) were tested along with wild type 5S RNA. Relative Ka values were obtained under identical experiment conditions. The results of these experiments are summarized in Figure 50 and Table 9. The binding affinities of some the mutants were investigated further by competition assays (Fig. 51).

Surprisingly, even though modest effects on the association constants were observed for some mutants (for example, stem III substitution mutants, the truncated mutants and the multiple mutations), the L5-5S RNA interaction generally is unaffected by most alterations either in the sequence or the secondary structure of the 5S RNA. No particularly important binding site has been identified for L5. In comparison, TFIIIA-5S RNA seems to be more sensitive to the disruption of 5S RNA secondary structures (Chapter 2 & You & Romaniuk. 1990).

4.4.DISCUSSION

4.4.1. L5 Expression and Purification

The poor solubility of L5 has been a problem in isolation of this protein. In an initial attempt, L5 was synthesized and labelled with ^{35}S *in vitro*, using L5 cDNA in SpL5AT plasmid (Wormington, 1989). The protein obtained from this method was small in quantity, and mixed with a relatively large amount of other proteins from the translation system. The obtained L5 might have aggregated due to its insolubility and was not very active. The binding result was unsatisfactory. A further effort was made by subcloning

the L5 gene into pET 3a and pET 11b vectors, and expressed the protein in *E.coli* strain BL21(DE3). Purification was carried out by ammonium sulfate precipitation and Bio-Rex 70 ion exchange column. L5 protein isolated by this way was active in 5S RNA binding, but the purity was still poor (60 to 80%). In this process, we found that the protein was overexpressed in large quantity, but most of the protein was in the form of aggregates, and was found in the cell debris pellet after centrifugation. The yield of active L5 protein was estimated to be less than 5% of the total. Use of 5M urea in the buffers increased the yield, but not the purity.

By expressing L5 with the pET 16b vector and purifying it using a His-tag affinity column, the purified protein was at least 95% homogeneous. pET 16b has 10 histidine codons located between the initiation codon and the Nde I cloning site (Fig. 40). The expressed L5 therefore is a fusion protein with a histidine tag at its N-terminus. Histidine has a strong affinity for nickel atoms on the affinity column. The fusion protein binds to the column so strongly that the protein could only be eluted by a buffer containing 100 mM EDTA, which frees the protein by stripping off the nickel atoms from the column. Histidine tags can be removed by digestion with the restriction protease Factor Xa. There is no significant difference in the Ka values for L5 proteins with or without the histidine tag. The yield of pure L5 obtained from this procedure was still relatively low (typically 300-400 μg from a 250 ml culture), compared with the total overexpressed protein. Guanidine can be applied in this procedure to solublize the protein and increase the yield. But since this amount of L5 is sufficient for our study, we chose to purify the protein in the absence of denaturant.

4.4.2. Comparison of L5-5S RNA and TFIIIA-5S RNA Interactions

The specific binding of L5 to 5S RNA has been demonstrated in previous studies (Lastick & McConkey, 1976; Phillips & McConkey, 1976; Blobel, 1971; Lebleu, *et al.*,

1971; Marion & Reboud, 1981; Petermann *et al.*, 1972; Ulbrick & Wool, 1978; Wormington, 1989), and efforts have been made to investigate the 5S RNA binding site(s) for L5 in some eucaryote species (Huber & Wool, 1986; Huber & Wool, 1984; Allison *et al.*, 1991; Guddat *et al.*, 1990). However, a direct measurement of the interaction kinetics has not yet been reported. The success of purifying L5 protein expressed in *E.coli* and the extensive collection of 5S RNA mutants allowed us to investigate the interaction directly and quantitatively. Extensive filter binding assays were applied to study the general binding features and the mutagenic effects on L5-5S RNA complex. These results were compared with TFIIIA-5S RNA data we obtained from previous experiments.

The equilibrium constant of L5-5S RNA was estimated to be 2.0 x 10^9 M^{-1}, slightly stronger than that of TFIIIA-5S RNA (Romaniuk, 1985). The relative values for Xls 5S RNA and WG 5S RNA were determined to be 1.52 and 0.58, respectively. In contrast, TFIIIA binds to WG 5S RNA 4 fold better than to Xlo 5S RNA. The binding strength of L5 to $tRNA^{phe}$ was shown to be 54-fold weaker than the L5-Xlo 5S RNA interaction, compared with that of 100 fold decrease in TFIIIA-$tRNA^{phe}$. Despite these differences, the L5-5S RNA equilibrium constants are essentially similar to those of TFIIIA-5S RNA bindings. Both proteins bind to 5S RNA specifically and strongly. The close Ka values of L5-5S RNA and TFIIIA-5S RNA interactions are consistent with the fact that the two types of complexes can exist simultaneously at the same intracellular location and have exchangeable protein components (see Introduction).

The Ka of 2.3 x $10^8 M^{-1}$ was reported for *E. coli* L18-5S RNA interaction, and the Ka values for L25 and L5 were one and two orders of magnitude lower than that of L18, respectively. Our data indicates that *Xenopus* L5 binding strength to 5S RNA is 10 fold greater than the *E. coli* L18-5S RNA, and 1,000 fold greater than that of *E. coli* L5. If the eucaryotic L5 is a fusion protein of the three procaryotic 5S RNA-binding ribosomal proteins (Nazar *et al.*, 1979), and considering the cooperation of the three proteins in 5S

RNA binding (Spierer *et al.*, 1978), one may suggest that the strong binding affinity of eucaryote L5 is a result of evolutionary development in which three separate 5S RNA-binding proteins were incorporated into a more efficient single protein.

The measurement of the dissociation kinetics for the reconstituted *Xenopus* L5-5S RNA complex revealed two populations: a rapid dissociating species and a extremely stable species, with a dissociation constant of 2.97×10^{-5} s^{-1}. The rapid dissociation of reconstituted complexes had been observed in the TFIIIA-5S RNA interaction (Romaniuk, 1985). In this study, approximately 19% of the complexes dissociated within the first few minutes of the experiments, while 29% underwent rapid dissociation in the TFIIIA-5S RNA case. The rapid dissociation was considered to be related to the preparation of the protein, or due to the non-specific binding of the protein to the RNA. The later seems to be unlikely because L5-tRNA^{phe} binding experiments showed that the non-specific interaction was extremely weak. The nature of rapid dissociation is not clear. 72% of the L5-5S RNA complexes remained bound after 24 hour incubation at 25 °C. The extreme stability of L5-5S RNA complexes is consistent with its function as a core structure for the ribosome assembly.

In the search for optimal binding conditions for L5-5S RNA, we found that the binding requirements (pH, temperature and ion strength) for L5-5S RNA are similar to those for TFIIIA-5S RNA. It appears that the L5-5S RNA interaction could tolerate higher ionic strength than that of TFIIIA-5S RNA. Otherwise, the general binding properties of the two proteins to 5S RNA are rather similar, consistent with the fact that the two types of RNPs co-exist in an identical environment.

4.4.3. Mutagenesis Analysis

The 5S RNA mutations were designed according to the secondary structural pattern of the molecule. Structural sites, like stems, loops, bulged nucleotides and potential

pseudonots were selected for mutagenesis and the binding affinities of mutants for L5 were carefully compared to that of wild type 5S RNA.

L5 is less sensitive to disruptions of helical structures than TFIIIA in contacting 5S RNA. Disruptions either at stem II or V reduced the TFIIIA affinity to 5S RNA by a factor of 2- to 3-fold (You & Romaniuk, 1990), but have no effect on L5-5S RNA binding (Table 9). Mutations of stem IV had no effect on binding of either protein. The only modest decrease in L5 Ka (rel. Ka 0.52) was observed in stem III (mutants 27-32 & 45-52). This region is out of the TFIIIA protection area on 5S RNA and appears to be unimportant in the TFIIIA-5S RNA interaction. Loop B/Stem III was reported to be within the *E. coli* L18 protection site (Huber & Wool, 1984). A much larger deletion in this region of *Xenopus* 5S RNA (Δ11-41) also showed a moderate decrease in L5 binding (Guddat *et al.*, 1990). Unfortunately, for some reason we could not clone the double mutant gene at this site. Therefore we could not conclude whether this modest decrease in Ka values is due to the destruction of the helix or due to the substitution of the nucleotides at this location. Since the mutations on each strand resulted in an almost identical decrease in the Ka values, it is likely that the structural factor, rather than the sequences, contributed to the poorer binding. There is evidence that the double stranded stem region of 5S RNA may adopt an A form helix, similar to that of the internal control region (ICR) of 5S DNA (Diakun *et al.*, 1986; Fairall *et al.*, 1989; Rhodes & Klug, 1986). It was proposed that TFIIIA may recognize the double helical structure of 5S RNA for its specific binding. However, L5 is not structurally related with TFIIIA and it is not considered as a DNA binding protein. Our data with the stem mutants showed that disruption of helical structures of 5S RNA has only moderate effects on the L5-5S RNA interaction. These observations are further supported by the binding results from the stem multiple mutants 16-21/95-98 and 16-21/67-70/95-98. These two combination mutants disrupted two or three stems of 5S RNA, respectively. While TFIIIA binding was reduced to 15% of that of the wild type, L5 kept 60% of its

binding strength. However, the mutant 16-21/57-62/78-81/95-98, which restores the stem II and stem V secondary structure, restores 80% of the binding strength, again suggesting a conformational effect in the determining of L5-5S RNA interaction.

The substitutions of nucleotides in the loops did not have any significant effect on the L5-5S RNA binding. In contrast, mutant 10-13 at loop A dramatically reduced the TFIIIA binding strength to 5S RNA by a factor of 3-fold and 20-fold in competition strength (Romaniuk *et al.*, 1989). It was suggested that the conformational change, not the sequence substitution, was responsible for the reduced binding affinity of TFIIIA. Loop A has been suggested to serve as a hinge that controls co-axial stacking of the helical domains of the 5S RNA (Christiansen *et al.*, 1987). Obviously the conformational change caused by nucleotide 10-13 substitution did not affect the L5 contacts to 5S RNA, as indicated by the relative Ka of 1.19 of direct binding and 2.1 in the competition assay (Table 9). Major changes in other loops did not show obvious effects on the binding, either. Loops B, C and E were turned into helixes in mutants 22-25, 33-34 and 96-101, respectively. Simple substitution mutants (loops were retained) in these three loops were also created (22-26 at loop B; 33-39 & 41-44 at loop C; 73-76 & 99-101 at loop E). None of these mutants appear to significantly interfere with the binding of L5 to 5S RNA (Table 9). The single stranded sequence GAGA in a loop-hairpin motif has been suggested to be important for some protein-RNA interactions (reviewed by Frankel *et al.*, 1991). A similar structure was found in loop D and stem IV. In mutant 87->0, the GAGA at loop D was replaced by AAAG. Again, the mutant did not respond to the sequence alteration, with a relative Ka value of 0.87.

Bulged nucleotides are crucial for some protein-RNA complex formation (Romaniuk *et al.*, 1987; Peattie *et al.*, 1981). Bulged nucleotides may provide sites for direct protein-RNA bonding, or provide a conformation necessary for the interaction. It was reported that the substitution of *E.coli* 5S RNA bulged nucleotide A66 did not affect

the binding of L18 to the mutant (Meier *et al.*, 1986), but the deletion of this bulged A66 resulted in a 7-fold decrease in binding affinity (Christiansen *et al.*, 1985). In the case of the R17 coat protein-RNA interaction, nucleotide substitution of the bulged A reduced the binding affinity by as much as 1,000-fold (Romaniuk *et al.*, 1987). The *Xenopus* 5S RNA bulged nucleotides were deleted in mutants ΔA49-50, ΔC63 and ΔA83. The results in Table 9 indicated that these highly conserved nucleotides are clearly not required for L5-5S RNA or TFIIIA-5S RNA interactions.

Finally, truncated 5S RNA molecules were constructed to test the effect of losing large 5S RNA fragments on L5 binding. RNA fragment nt. 1-68 represent the 5' half of the molecule, nt.50/51-120 represent the 3' half of the molecule. In a standard binding assay, nt. 1-68 had a moderately decreased relative Ka of 0.58, and the relative Ka of nt.50/51-120 was 0.76 (Table 9). Both fragments kept relatively strong binding affinities despite deletion of almost half of the 5S RNA molecule. However, a more significant difference was revealed in the competition assays. While the 5' fragment nt. 1-68 competes almost as well as the wild type (relative strength = 86%), the 3' fragment nt.50/51-120 has only 9% of the competition strength (Fig. 51). The other 5S RNA fragment, nt. 66-108, consisting of stem IV, V and loops D and E, also shows very poor competition strength in the same experiment, suggesting that the 5' half of the 5S RNA is likely more important than the 3' half for the protein to contact. Although these data indicate the different binding strength of the RNA fragments, further investigation is needed before we can make precise conclusions.

It is surprising that no particularly important 5S RNA site for L5 binding was found after the extensive mutagenesis analysis. In previous microinjection experiments, Allison *et al.* (1991) showed that 5S RNA nucleotides 11-108 provide sufficient sequence information for binding by L5. From our study, the size of 5S RNA required for L5 binding is even shorter, for example, nt. 1-68 was bound well by L5. The result from

Allison *et al.* indicates that both strands of stem I of 5S RNA are not required for binding, in agreement with our binding data of nt. 1-68 and 50/51-120 that either strand of stem I is not important. The stemII/Loop B/Stem III region was determined as the *E. coli* L18 binding site (Egebjerg *et al.*, 1989; Garrett *et al.*, 1981; Huber & Wool, 1984). Our data show that disrupting stem III resulted in the largest but still moderate decrease of 50% in binding affinity, and that 5S RNA fragment 1-68 competed strongly with the intact wild type 5S RNA, indicating that this region is probably also a *Xenopus* L5 binding site. The other fragments, either completely (66-108) or partly (51-120) losing the stem II/Loop B/StemIII region, were still bound by L5 but competed poorly with the wild type 5S RNA (Fig. 51), suggesting these fragments (equivalent to the *E. coli* L25 binding site) may be a weaker binding site for the eucaryote L5 ribosomal protein. Recall that *E. coli* L18-Stem II/Loop B/Stem III interaction was 10 fold stronger than the L25-Stem IV/Loop E/Stem V contact (Spierer *et al.*, 1978), we speculate that the proposed fusion eucaryote L5 may have retained these procaryote features during the course of evolution.

It seems that L5 protein is less sensitive to either sequence or higher order structural changes of 5S RNA than TFIIIA. Guddat *et al.* (1990) reported similar observations to those reported here. *Xenopus* L5 binds well to a number of large-block deletion/substitution mutants which together covered 50% of the 5S RNA sequence and disrupted major secondary structures. The only moderate effect of those mutations on L5 binding was mutant (Δ11-41), in which stem II/Loop B/Stem III region was completely deleted (Guddat *et al.*, 1990). The same set of mutants that bound well to L5 could not form complexes with TFIIIA, indicating that L5 protein could tolerate more changes in 5S RNA molecule. Combined with our data, it is apparent that *Xenopus* L5 binding sites are dispersed over the entire 5S RNA molecule, while the stem II/loopB/stem III region seems to be the most important. Interestingly, they also showed that, like L5, the La protein binds very well to these mutants, including the deletion mutant (Δ11-41). Our unpublished

data on p43 protein, a component of the 42S RNP, also demonstrate the unusual insensitivity of this protein to 5S RNA mutations. Do these three structurally distinct 5S RNA binding proteins have a similar binding pattern that is different from the TFIIIA-5S RNA pattern?

At present, we do not have sufficient data to define a *Xenopus* L5-5S RNA interaction site and the nature of the interaction. However, the combination of observations on *E. coli* L5, L18 and L25-5S RNA binding features and the data from *Xenopus* L5-5S RNA interactions provide us these facts: (1) The binding sites of the three procaryote ribosomal proteins are dispersed widely along almost the entire 5S RNA molecule, and their binding strengths are unequal; (2) The eucaryote L5 appears to possess binding features similar to the combination of the three procaryote ribosomal proteins, including the dispersed binding sites and uneven binding strength; (3) The *Xenopus* L5-5S RNA interaction is insensitive to alterations in both sequence and secondary structures.

In contrast to RNA-protein interactions, the protein-DNA interactions are essentially dependent on specific consensus sequence recognition, due to the relatively simple DNA structural features. TFIIIA contacts a few nucleotides within three short elements (Box A, C, & IE) of the 5S DNA ICR by direct amino acid side-chain and base contacts (Veldhoen *et al.*, in preparation). No such elements or nucleotides have been identified in the 5S RNA molecule for the protein interaction. Protein-RNA interactions have more options. Proteins may contact RNA via its bases or phosphate backbone, recognize special secondary structural motifs as well as tertiary features, or any combination of them.

In the case of the *Xenopus* TFIIIA-5S RNA interaction, we have shown that the secondary structures of 5S RNA, but not the sequence, play an important role in the specific interaction (Chapter 2 & You & Romaniuk, 1990). TFIIIA may make electrostatic contacts to 5S RNA sugar-phosphate backbone of the stems, which would be sensitive to their conformational context far more than their sequence context. We suggest that the

same proposal may also apply to the *Xenopus* L5-5S RNA interaction with modifications: L5 may recognize the local conformations of the binding sites; the binding may be dependent upon the overall contribution of multiple contacts rather than any particular site; these contacts could be widely dispersed along most of the RNA, and each contact is weak. Therefore, even though the interaction is specific, it can still be able to tolerate drastic alterations in sequences and secondary structures.

CONCLUSIONS

The protein-nucleic acid interactions of TFIIIA-5S DNA, TFIIIA-5S RNA and L5-5S RNA were studied in this work, and the features of these interactions compared. The data presented in this thesis clearly demonstrate that although different mechanisms are involved in each protein-nucleic acid interaction, the protein-RNA interactions are more related to each other than to the protein-DNA interactions.

1. TFIIIA-5S DNA interaction. TFIIIA binds specifically to 5S DNA by forming sequence-specific contacts with three discrete sites located within the classical A and C boxes and the intermediate element of the internal control region. Substitution of the nucleotide sequence at any of the three sites significantly reduces TFIIIA binding affinity with a 100-fold reduction observed for substitutions in the box C subregion. These results are consistent with those of a selected amplification and binding (SAAB) experiment, in which the native sequence of box C subregion was highly selected from a randomized pool. The importance of individual nucleotides for TFIIIA contact was investigated with an extensive set of point mutants in the ICR region. This study reveals that the most important individual nucleotides for TFIIIA contact are located on the non-coding strand of the box C element, and they are all guanine residues. A single base substitution of G81, G85, or G89 reduces TFIIIA association five to ten-fold.

These observations show clearly that the TFIIIA-5S DNA interaction is dependent on base recognition, a typical protein-DNA contact mechanism. The general characteristics of this type of interaction are high sequence-specificity; depending on a few individual nucleotides and forming a few but crucial bonds.

2. TFIIIA-5S RNA interaction. The interaction of TFIIIA-5S RNA appears not to depend directly on the sequence. Previous results have shown that nucleotide substitutions

in loops and bulged nucleotides have no obvious effect on TFIIIA binding, except for the mutation in loop A, which was explained as a result of conformational change rather than the alteration of sequence. Sequence substitutions in stems were performed in this study. Mutations in stems III and IV had little or no effect on the binding affinity of TFIIIA for 5S RNA. However, single mutants in stem II and V which disrupt the double helix reduce the binding of TFIIIA by two- to three-fold. In contrast, double mutants at the same site, which restore the helical structure of these stems, but alter the sequences, fully restore the TFIIIA binding affinity.

The experiment shows that the TFIIIA-5S RNA interaction is basically limited to the center of the 5S RNA molecule (stem II, loop B/stem V loop E), and is likely dependent on the local conformation, not the sequence. In contrast to TFIIIA-5S DNA interaction, the specificity of TFIIIA-5S RNA interaction appears to be achieved by recognition of local 5S RNA conformation rather than individual nucleotides, and TFIIIA may contact the sugar-phosphate backbone rather than the bases. In summary, it is apparent that TFIIIA interacts with 5S DNA and 5S RNA by fundamentally different mechanisms. This concept predicts that fingers (or regions within any one finger) were optimized for binding either to 5S DNA or 5S RNA.

3. L5-5S RNA interactions. Ribosomal protein L5 is structurally distinct from TFIIIA: the protein does not have zinc finger motif. However, both proteins specifically bind to the same 5S RNA, with similar binding characteristics. There is no sequence-specific contact observed in either case. A significant difference in the binding of TFIIIA and L5 to 5S RNA is that L5 is insensitive not only to alterations in sequence but also to conformational changes.

It is premature to make a firm conclusion on the nature of L5-5S RNA interaction. However, existing data suggests that the eucaryotic L5 protein may contact with 5S RNA on the sites that are scattered along most of the molecule. These contacts may be weak,

depending on hydrogen bonds or electrostatic contacts to the sugar-phosphate backbone rather than the bases. These features of L5-5S RNA interaction are similar to those of TFIIIA-5S RNA, but the L5-5S RNA may make more extensive contacts and therefore be less dependent on any individual interaction.

5S RNA is a small (120 nt.) molecule and it interacts with many structurally distinct proteins during various stages of its biological pathway. It is not surprising that these interactions may use different strategies. It is possible that the activities based on the interaction of TFIIIA with 5S DNA and 5S RNA have evolved separately from an ancestor molecule, and developed distinct mechanisms. On the other hand, the TFIIIA and L5 binding to 5S RNA may have evolved from two distinct ancestors, but resulted in a similar yet still distinguishable mechanism.

LITERATURE CITED

Aboul-ela, F., Varani, G., Walker, G. T., & Tinoco, I., Jr. (1988) *Nucleic Acids Res.* **16**, 3559-3572.

Allison, L. A., Romaniuk, P. J. & Bakloen, A. H. (1991) *Dev. Biol.* **144**, 129-144.

Altman, S. (1990) *J. Biol. Chem.* **265**, 20053-20060.

Andersen, J. & Delihas, N. (1986) *J. Biol. Chem.* **261**, 2912-2917.

Andersen, J., Delihas, N., Hanas, J. S. & Wu, C.-W. (1984) *Biochemistry* **23**, 5759-5766.

Andrews, M. T. & Brown D. D. (1987) *Cell* **51**, 445-453.

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

Atkinson, T. & Smith, M. (1984) In: *Oligonucleotide Synthesis: A Practical Approach*, Gait, M. J. ed. pp35-82, IRL Press, Oxford.

Bandziulis *et al.* (1989) *Genes Dev.* **3**, 431.

Baudin, F., Romaniuk, P. J., Romby, P., Brunel, C., Westhof, E., Ehresmann, B. & Ehresmann, C. (1991) *J. Mol. Biol.* **218**, 69-81.

Baudin, F. & Romaniuk, P. J. (1989) *Nucleic Acids Res.* **7**, 2043-2056.

Baudin, F., Romby, P., Romaniuk, P. J., Ehresmann, B. & Ehresmann, C. (1989) *Nucleic Acids Res.* **17**, 10035-10046.

Becker, M. M. & Wang, Z. (1989) *J. Biol. Chem.* **264**, 4163-4167.

Berg, J. M. (1988) *Proc. Natl. Acad. Sci. USA.* **85**, 99-102.

Berg, J. M. (1986) *Science* **232**, 485-487.

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

Bieker, J. J. & Roeder R. G. (1984) *J. Biol. Chem.* **259**, 6158-6164.

Blackwell, T. K. & Weintraub, H. (1990) *Science* **250**, 1104-1110.

Blanco, J., Millstein, L., Razik, M. A., Dilworth, S., Cote, C. & Gottesfeld, J. (1989) *Gene & Dev.* **3**, 1602-1612.

Blobel, G. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1881-1885.

Bogenhagen, D. F. (1985) *J. Biol. Chem.* **260,** 6466-6471.

Bogenhagen, D. F. & Brown, D. D. (1981) *Cell* **24**, 261-270.

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

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

Borer, R. A., Lehner, C. F., Eppenberger, H. M. & Nigg, E. A. (1989) *Cell* **56**, 379-390.

Bradford, M. M. (1976) *Analytical Biochem.* **72**, 248-254.

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

Brown, D. D., Carroll, D. & Brown, R. D. (1977) *Cell* **12**, 1045-1056.

Brown, D. D., Wensink, P. C. & Jordan, E. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 3175-3179.

Brown, D. A. & Geiduschek, E. P. (1987). *J. Biol. Chem.* **262**, 13953-13958.

Brown, R. S., Sander, C. & Argos, P. (1985) *FEBS Lett.* **186,** 271-274.

Brownlee, G. G., Sanger, F. & Barrel, B. G. (1968) *J. Mol. Biol.* **34**, 374-386.

Brunel, C., Romby, P., Ehresmann, C., Romaniuk, P. J., Ehresmann, B. & Westhof, E. (1990) *J. Mol. Biol.* **215**, 103-111.

Burma, D. P., Srivastava, S., Srivastava, A. K., Mahanti, S. & Dash. D. (1986) *Conformational Change of 50S Ribosomes during Protein Synthesis.* In: *Structure, Function, and Genetics of Ribosomes,* eds. Hardesty, B. & Kramer, G., Springer-Verlag, New York, pp438-453.

Busby, S. J. & Reeder, R. H. (1983) *Cell* **34**, 989-996.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C. & Housman, D. E. (1990) *Cell* **60**, 509-520.

Cantor, C. R. & Schimmer, P. R. (1980) In: *Biophysical Chemistry*, Part III. W. F. Freeman & Co., San Francisco, pp. 1146-1160.

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

Chan, Y. L., Lin, A., MacNally, J. & Wool, I. G. (1987). *J. Biol. Chem.* **262**, 12879-12886.

Chavrier, P., Lemaire, P., Relevant, O., Bravo, R. & Charnay, P. (1988) *Mol. Cell. Biol.* **8**, 1319-1326.

Chen-Schmeisser, U. & Garrett, R. A. (1977) *FEBS Lett.* **74**, 287-291.

Christiansen, J., Brown, R. S., Sproat, B. S. & Garrett, R. A. (1987) *EMBO J.* **6**, 453-460.

Christiansen, J., Douthwaite, S. R., Christensen, A. & Garrett, R. A. (1985) *EMBO J.* **4**, 1019-1024.

Christiansen, J. & Garrett, R. A. (1986) In: *Structure, Function, and Genetics of Ribosomes*, eds. Hardesty, B. & Kramer, G., Spinger-Verlag, New York, pp254-269.

Christy, B. A., Lau, L. F. & Nathans, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7857-7861.

Churchill, M. E. A., Tullius, T. D. & Klug, A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5528-5532.

Ciliberto, G., Raugei, G., Constanzo, F., Dente, L. & Cortese, R. (1983) *Cell* **32**, 725-733.

Clark, W. & Lake J. A. (1984) *J. Bact.* **157**, 971-974.

Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D. & Bogenhagen, D. F. (1983) *Cell* **34**, 829-835.

Culotta, V. A. & Sollner-Webb, B. (1985) *J. Cell. Biochem.* **B9**, 156.

Cutruzzolá, F., Loreni, F. & Bozzoni, I. (1986) *Gene* **49**, 371-376.

Darby, M. K., Andrews, M. T.& Brown, D. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5516-5520.

Darsillo, P & Huber, W (1991) *J. Biol. Chem.* **266**, 21075-21082.

Davanloo, P., Rosenberg, A. H., Dunn, J. J. & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035-2039.

Delihas, N. & Andersen, J. (1982) *Nucleic Acids Res.* **10**, 7323-7344.

Del Rio, S. & Setzer, D. R. (1991) *Nucleic Acids Res.* **19**, 6197-6203.

Diakun, G. P., Fairall, L. & Klug, A. (1986). *Nature* **324**, 698-699.

Digweed, M., Pierler, T., Kluwe, D., Schuster, L., Walker, R. & Erdmann, V. A. (1986) *Eur. J. Biochem.* **154**, 31-39.

Dohme, F. & Nierhaus, K. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2221-2225.

Dunaway, M. & Reeder, R. H. (1985) *Mol. Cell. Biol.* **5**, 313-319.

Egebjery J., Christiansen, J., Brown, R. S., Larsen, N. & Garrett, R. A. (1989) *J. Mol. Biol.* **206**, 651-668.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P. & Ehresmann, B. (1987) *Nucleic Acids Res.* **15**, 9109-9128.

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

Erdmann, V. A. (1976) *Prog. Nucleic Acid Res. Mol, Biol.* **18**, 45-90.

Erdmann, V. A. (1981) *Nucleic Acids Res.* **9**, 225-242.

Erdmann, V. A., Appel, B., Digweed, M., Kluwe, D., Lorenz, S., Lüke, A., Schreibe, A. & Schuster, L. (1980) *Structure and Function of 5S and 5.8S ribosomal RNAs*, In: *Genetics and Evolution of RNA polymerase, tRNA and Ribosomes*, eds. Osawa, S. *et al.*, University of Tokyo Press, Elsevier/North Holland Biochemical Press, Amsterdam, pp553.

Erdmann, V. A., Doberer, H. G. & Sprinzl, M. (1971) *Mol. Gen. Genet.* **114**, 89-94.

Erdmann, V. A., Pieler, T., Wolters, J., Digweed, M., Vogel, D. & Hartmann, R. (1986) *Comparative Structural and Functional Studies on Small Ribosomal RNAs*, In: *Structure, Function and Genetics of Ribosomes*, eds. Hardesty, B. & Kramer, G., Springer-Verlag, New York, pp. 165-183.

Erdmann, V. A., Sprinzl, M. & Pongs, O. (1973) *Biochem. Biophys. Res. Comm.* **54**, 942-948.

Erdmann, V. A., Wolters, J., Huysmans, D. & DeWachter, R. (1985) *Nucleic Acids Res.* **13**, 103-153.

Evans, R. M. & Hollenberg, S. M. (1988) *Cell* **52**, 1-3.

Fairall, L., Martin, S. & Rhodes, D. (1989) *EMBO J.* **8**, 1809-1817.

Fairall, L., Rhodes, D. & Klug, A. (1986) *J. Mol. Biol.* **192**, 577-591.

Fersht (1977) *Enzyme Structure and Mechanism*, Published by W. H. Freeman & Co. San Francisco.

Feunteun, J., Monier, R., Garrett, R. A., Le Bret, M. & Le Pecq (1975) *J. Mol. Biol.* **6**, 637-648.

Frankel, A. D., Berg, J. M. & Pabo, C. O. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4841-4845.

Frankel, A. D., Mattaj, I. W. & Rio, D. C. (1991) *Cell* **67**, 1041-1046.

Garey, J., Cameron, V., de Haseth, P. L. & Uhlenbeck, O. C. (1983) *Biochemstry* **22**, 2601-2610.

Garrett, R. A., Douthwaite, S. & Noller, H. F. (1981) *Trends Biochem. Sci.* **5**, 137-139.

Geidnschek, E. P. & Tocchini-Valentini, G. P. (1988) *Ann. Rev. Biochem.* **57**, 873-914.

Giacomoni, P. U. (1981) *Biochem. Int.* **2**, 389-397.

Giacomoni, P. U. (1981) *Biochem. Int.* **2**, 399-410.

Gibson, T. J., Postma, J. P. M., Brown, R. S. & Argos, P. (1988) *Protein Engin.* **2**, 209-218.

Gilbert, D. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2924-2928.

Ginsberg, A. M., King, B. O. & Roeder, R. G. (1984) *Cell* **39**, 479-489.

Göringer, H. U., Bertram, S. & Wagner, R. (1986) *Nucleic Acids Res.* **14**, 7473-7485.

Gottesfeld, J. M., Blanco, J. & Tennant, L. L. (1987) *Nature* **329**, 460-462.

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

Gottlieb, E. & Steitz, J. A. (1989) *EMBO J.* **8**, 851-861.

Gründstrom, T., Zenke, W. M., Wintzerith, M., Matthes, H. W. D., Staub, A. & Chambon, P. (1985) *Nucleic Acids Res.* **13**, 3305-3316.

Guddat, U., Bakken, A. H. & Pieler, T. (1990) *Cell* **60**, 619-628.

Hanas, J. S., Bogenhagen, D. F. & Wu, C.-W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2142-2145.

Hanas, J. S., Bogenhagen, D. F. & Wu, C.-W. (1984) *Nucleic Acids Res.* **12**, 2745-2758.

Hanas, J. S., Hazuda, D. J., Bogenhagen, D. F., Wu, F.Y.-H. & Wu, C.-W. (1983) *J. Biol. Chem.* **258**, 14120-14125.

Hancock, J. & Wagner, R. (1982) *Nucleic Acids Res.* **10**, 1257-1269.

Härd, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Guslafsson, J.-Å. & Kaptein, R. (1990) *Science* **249**, 157-160.

Hayes, J., Tullius T. D. & Wolffe, A. P. (1989) *J. Biol. Chem.* **264**, 6009-6012.

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

Horne, J. R. & Erdmann, V. A. (1972) *Mol. Gen. Genet.* **119**, 337-344.

Huber, P. W. & Wool, I. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 322-326.

Huber, P. W. & Wool, I. G. (1986) *J. Biol. Chem.* **261**. 3002-3005.

Huber, P. W. & Wool, I. G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1593-1597.

Huber, P. W., Morii, T., Mei, H-Y. & Barton, J. K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10801-10805.

Jack, A., Ladner, J. E. & Klug, A. (1976) *J. Mol. Biol.* **108**, 619-649.

Joho, K. E. & Darby, M. K. (1990) *Cell* **61**, 293-300.

Kargel, H-J., Gross J. S. B., Knespel, S., Bielka, H. & Saarma M. (1987) *FEBS Lett.* **220**, 126-128.

Keller, H. J., You, Q., Romaniuk. P. J. & Gottesfeld, J. M. (1990) *Mol. Cell. Biol.* **10**, 5166-5176.

Kenmochi, N., Maeda, N. & Tanaka, T. (1991) *Biochim. Biophys. Acta.* **1088**. 445-447.

Kinzler, K. W., Ruppett, J. M., Bigner, S. H. & Vogelstein, B. (1988) *Nature* **332**, 371-374.

Kjems, J., Olesen, S. E. & Garrett, R. A. (1985) *Biochemistry* **24**, 241-250.

Kochoyan, M., Keutmann, H. T. & Weiss, M. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8455-8459.

Kongsuwan, K., Yn, Q., Vincent, A., Frisardi, M. C., Rosbash, M., Lengyel, J. A. & Merriam, J. (1985) *Nature* **317**, 555-558.

Korn, L. K. & Gurdon, J. B. (1981) *Nature* **289**, 461-467.

Larson, D., Brandford-Wilcox, J., Young, L. S. & Sprague, K. U. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3416-3420.

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

Lastick, S. M. & McConkey, E. H. (1976) *J. Biol. Chem.* **251**, 2867-2875.

Lebleu, B., Marbaix, G., Huez, G., Temmerman, J., Burny, A. & Chantrenne, H. (1971) *Eur. J. Biochem.* **19**, 264-269.

Lee, D. K., Evans, R. K., Blanco, J., Gottesfeld, J. M. & Johnson, J. D. (1991) *J. Biol. Chem.* **266**, 16478-16484.

Lee, M. S., Gippert, G. P., Soman, K. V., Case D. A. & Wright, P. E. (1989) *Science* **245**, 635-637.

Liao, X., Clemens, K. R., Tennant, L., Wright, P. E. & Gottesfeld, J. M. (1992) *J. Mol. Biol.* **223**, 857-871.

Lin, J. J. & Keller, T. J. Jr. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6973-6977.

Lin, J. J. & Keller, T. J. Jr. (1985) *Mol. Cell. Biol.* **5**, 1238-1246.

Lin, S-Y. & Riggs, A. D. (1975) *ibid* **4**, 107-111.

Loreni, F., Ruberti, I., Pierandrei-Amaldi & Amaldi, F. (1985) *EMBO J.* **4**, 3483-3488.

Macharias, M. & Wagner, R. (1986) *FEBS Lett.* **204**, 89-95.

Mager, W. H. (1988) *Biochem. Biophys. Acta* **949**, 1-15.

Majowski, K., Mentzel, H. & Pieler, T. (1987) *EMBO J.* **6**, 3057-3063.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Mannual.* Cold Spring Harbor, pp250-251.

Marion, M. J. & Reboud, J. P. (1981) *Biochim. Biophys. Acta.* **652**, 193-203.

McConkey, G. A. & Bogenhagen, D. F. (1987) *Mol. Cell. Biol.* **7**, 486-494.

McGhee, J. & Felsenfeld, G. (1986) *Cell* **44**, 375-377.

McCall, M., Brown, T., Hunter, W. N. & Kennard, O. (1986) *Nature* **322**, 661-664.

McDougall, J. & Nazar, R. N. (1983) *J. Biol. Chem.* **258**, 5256-5259.

Metspalu, A., Saarma, M., Villems, R., Ustav, M. & Lind, A. (1978) *Eur. J. Biochem.* **91**, 73-81.

Miller, J., McLachlan, A. D. & Klug, A. (1985) *EMBO J.* **4**, 1609-1614.

Miller, J. R., Cartwright, E. M., Brownlee, G. G., Federoff, N. V. & Brown, D. D. (1978) *Cell* **13**, 717-725.

Mougel, M., Eyermann, F., Westhof, E., Romby, P., Expert-Bezancon, A., Ebel, J.-P., Ehresmann, B. & Ehresmann, C. (1987) *J. Mol. Biol.* **198**, 91-107.

Nakamaye, K. L. & Eckstein, F. (1986) *Nucleic Acids Res.* **14**, 9679-9698.

Nasmyth, K. (1987) *EMBO J.* **6**, 243-248.

Nazar, R. N., Yaguchi, M., Willick, G. E., Rollin, C. F. & Roy, C. (1979) *Eur. J. Biochem.* **102**, 573-582.

Nierhaus, K. M. & Dohme, F. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4713-4717.

Noller, H. F. (1991) *Ann. Rev. Biochem.* **60**, 191-227.

Noller, H. F. & Garrett, R. A. (1979) *J. Mol Biol.* **132**, 621-636.

Noller, H. F., Hoffarth, V. & Zimniak, L. (1992) *Science* **256**, 1416-1419.

Nomura, M. & Erdmann, V. A. (1970) *Nature* (London) **228**, 744-748.

Nomura, M., Gourse, R. & Baughman, G. (1984) *Ann. Rev. Biochem.* **53**, 75-117.

Nussinov, R., Tinoco, I. Jr. & Jacoson, A. B. (1982) *Nucleic Acids Res.* **10**, 351-363.

Oβwald, M., Greuer, B. & Brimacombe, R. (1990) *Nucleic Acids Res.* **18**, 6755-6760.

Ofengand, J. & Henes, C. (1969) *J. Biol. Chem.* **244**, 6241-6253.

Pace, B., Matthews, E. B., Johnson, K. D., Cantor, C. R. & Pace, N. R. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 36-40.

Pardue, M. L., Brown, D. D. & Birnstiel, M. L. (1973) *Chromosoma* **42**, 191-203.

Párraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T. & Klevit, R. E. (1988) *Science* **241**, 1489-1492.

Pavletich, N. P. & Pabo, C. O. (1991) *Science* **252**, 809-817.

Peattie, D. A., Douthwaite, S., Garret, R. A. & Noller, H. P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7331-7335.

Peck, L. & Wang, J. C. (1981) *Nature* **292**, 375-378.

Pelham, H. R. B. & Brown, D. D. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4170-4174.

Petermann, M. L., Hamilton, M. G. & Pavlove, A. (1972) *Bochemistry* **2**, 2323-2326.

Peterson, R. C., Doering, J. L. & Brown, D. D. (1980) *Cell* **20**, 131-141.

Phillips, W.F. & MacConkey, (1976) *J. Biol. Chem.* **251**, 2876-2881.

Picard, B. & Wegnez, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 241-245.

Picard, B., le Maire, M., Wegnez, M. & Denis, H. (1980) *Eur. J. Biochem.* **109**, 359-368

Pieler, T. & Erdmann, V. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4599-4603.

Pieler, T. & Erdman, V. A. (1983). *FEBS Lett.* **157**, 283-287.

Pieler, T., Erdmann, V. A. & Appel, B. (1984) *Nucleic Acids Res.* **12**, 8393-8406.

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

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

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

Query, C. C., Bentley, R. C. & Keane, I. D. (1989) *Cell* **57**, 89-101.

Quigley, G. J., Wang, A., Seeman, N. C., Suddath, F. L., Rich, A., Sussman, J. L. & Kim, S. H. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4866-4870.

Raacke, I. D. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2357-2360.

Ramé, H. A., Lorenz, S., Erdmann, V. A. & Planta, R. J. (1981) *Nucleic Acids Res.* **9**, 1263-1269.

Reeder, R. H., Roan, J. & Dunaway, M. (1983) *Cell* **35**, 449-456.

Reynolds, W. F. & Gottesfeld, J. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4018-4022.

Rhodes, D. & Klug, A. (1986) *Cell* **46**, 123-132.

Rinke, J. & Steitz, J. A. (1982) *Cell* **29**, 149-159.

Romby, P., Westhof, E., Moras, D., Giegé, R., Houssier, C. & Grosjéan, H. (1986) *J. Biomol. Str. Dyn.* **4**, 193-203.

Romaniuk, P. J. (1985) *Nucleic Acids Res.* **13**, 5369-5387.

Romaniuk, P. J. (1988) *Biochemistry* **28**, 1388-1395.

Romaniuk, P. J., Leal de Stevenson, I. & Wong, H.-H. A. (1987a) *Nucleic Acids Res.* **15**, 2737-2755.

Romaniuk, P. J., Lowary, P., Wu, H. N., Stormo, G. & Uhlenkeck, O. C. (1987b) *Biochemistry* **26**, 1563-1568

Romaniuk, P. J., Stevenson, I. L. & You, Q. (1988) In "*The Molecular Biology of RNA*", T. R. Cech ed. Alan R. Liss Inc. pp 123-132

Romaniuk, P. J. (1989) *Biochemistry* **28**, 1388-1395.

Romaniuk, P. J. (1990) *J. Biol. Chem.* **265**, 17593-17600.

Romaniuk, P. J., Leal de Stevenson, I., Ehresmann, C., Romby, P. & Ehresmann, B. (1988) *Nucleic Acids Res.* **16**, 2295-2312.

Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989) *Science* **246**, 1135-1142.

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

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

Sakonju, S. & Brown, D. D. (1981) *Cell* **23**, 665-669.

Sakonju, S. & Brown, D. D. (1982) *Cell* **31**, 395-405.

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

Sands, M. S. & Bogenhagen, D. F. (1987) *Mol. Cell. Biol.* **7**, 3985-3993.

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

Shang, Z., Windsor, W. T., Liao, Y.-D. & Wu, C.-W. (1988) *Anal. Biochem.* **168**, 156-163.

Shang, Z., Liao, Y.-D., Wu, F. Y.-H. & Wu, C.-W. (1989) *Biochemistry* **28**, 9790-9795.

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

Schnien, J. & Faist, G. (1985) *Mol. Gen. Genet.* **200**, 476-481.

Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 804-808.

Selker, E. U., Morzycka-Wroblewska, E., Stevens, J. N. & Metzenberg, R. L. (1985) *Mol. Gen. Genet.* **205**, 189-192.

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

Setzer, D. R., Hmiel, R. M. & Liao, S. (1990) *Nucleic Acids Res.* **18**, 4175-4178.

Sharp, S., Garcia, A., Cooley, L. & Soll, D. (1984) *Nucleic Acids Res.* **12**, 7617-7623.

Shatsky, I. N., Evstafievea, A. G., Bystrova, A. A., Bogdanov, A. A. & Vassiliev, V. D. (1980) *FEBS Lett.* **121**, 97-100.

Silberklang, M., RajBhandary, U. L., Lüke, A. & Erdmann, V. A. (1983) *Nucleic Acids Res.* **11**, 605-617.

Sinha, N. D., Biernat, J., McManus, J. & Koster, H. (1984) *Nucleic Acids Res.* **12**, 4539-4557.

Smith, D. R., Jackson, I. J. & Brown, D. D. (1984) *Cell* **37**, 645-652.

Smith, N., Matheson, A. T., Yaguchi, M., Willick, G. E. & Mzar, R. N. (1978) *Eur. J. Biochem.* **89**, 501-509.

Sneath, B., Vary, C., Pavlakis, G. & Vournakis, J. (1986) *Nucleic Acids Res.* **14**, 1365-1378.

Sollner-Webb, B., Miller, K. G., Tower, J., Culotta, V. C. & Wildle, J. (1985) *J. Cell. Biochem.* **B9**, 174.

Speek, M. & Lind, A. (1982) *Nucleic Acids Res.* **10**, 947-965.

Spierer, P., Bogdanov, A. A. & Zimmermann, R. A. (1978) *Biochemistry* **17**, 5394-5398.

Spierer, P. & Zimmermann, R. A. (1978) *Biochemistry* **17**, 2474-2479.

Spirin, A. S. (1986) In: *Ribosome Structure and Protein Biosynthesis*, Menlo Park CA: Benjamin/Cummings, pp. 260-261.

Sprinzl, M., Wagner, T., Lorenz, S. & Erdmann, V. A. (1976) *Biochemistry* **15**, 3031-3039.

Stahl, D., Luehrsen, K. R., Woese, C. R. & Pace, N. R. (1981) *Nucleic Acids Res.* **9**, 6129-6137.

Stefano, J. (1984) *Cell* **36**, 145-154.

Steitz, J. A., Berg, C., Hendrick, J. P., Branche-Chabot, H. L., Metspalu, A., Rinke, J., & Yaro, T. (1988) *J. Cell Biol.* **106**, 545-556.

Sternberg, P. W., Stern, M. J., Clark, I. & Herskowitz, I. (1987) *Cell* **48**, 567 577.

Stevenson, I. L., Baudin, F., Brunel, C., Romby, P., Ehresmann, C., Ehresmann, B. & Romaniuk, P. J. (1991) *J. Mol. Biol.*, **219**, 243-255.

Stillman, B. W. & Gluzman, Y. (1985) *Mol. Cell. Biol.* **5**, 2051-2060.

Studier, F. W., Rosenberg, A. H., Dunn, J. J & Dubendorff, J. W. (1990) *Methods Enzymology* **185**, 60-88.

Sukhatme, V. P., Cao, X., Chang, L. C., Tsai-Morris, C.-H., Stamenkovich, O., Ferreira, P. C. P., Cohen, D. R., Edwards, S. A., Shows, T. B., Curran, T., Le Beau, M. M. & Adamson, E. D. (1988) *Cell* **53**, 37-43.

Taylor, J. W., Schmide, W., Cosstick, R., Okruszed, A. & Echstein, F. (1985a) *Nucleic Acids Res.* **13**, 8749-8764.

Taylor, J. W., Ott, J. & Eckstein, F. (1985b) *Nucleic Acids Res.* **13**, 8765-8784.

Terao, K., Takahashi, Y. & Ogata, K. (1975) *Bochim. Biophys. Acta.* **402**, 230-237.

Thomas, J. D. & Rees, C. (1983) *Eur. J. Biochem.* **134**, 109-115.

Toots, I., Metspalu, A. & Saarma, M. (1981) *Nucleic Acids Res.* **9**, 5331-5343.

Toots, I., Misselwitz, R., Bohm, S., Welfle, H., Villems, R. & Saarma, M. (1982) *Nucleic Acids Res.* **10**, 3381-3389.

Trout, A., Savin, T., Curtis, W. C., Celentano, J. & Vournakis, J. (1982) *Nucleic Acids Res.* **10**, 653-663.

Ulbrich, N. & Wool, I. G. (1978) *J. Biol. Chem.* **253**, 9049-9052.

Varani, G., Wimberly, B. & Tinoco, I. J. (1989) *Biochemistry* **28**, 7760-7772.

Veldhoen, N., You, Q., Setzer, D. R. & Romaniuk, P. J. (1992) In preparation.

Vincent, A. (1986) *Nucleic Acids Res.* **14**, 4385-4381.

Vrana, K. E., Churchill, M. E. A., Tullius, T. D. & Brown, D. D. (1988) *Mol. Cell. Biol.* **8**, 1684-1696.

Vogel, D. W., Hartmann, R. K., Bartsch, M., Subramanian, A. R., Kleinow, W., O'Brien, T. ... Pieler, T. & Erdmann, V. A. (1984) *FEBS Lett.* **169**, 67-72.

Waldrop, M. M. (1992) *Science* **256**, 1396-1397.

Westhof, E., Dumas, P. & Moras, D. (1985) *J. Mol. Biol.* **184**, 119-145.

Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J.-P., Ehresmann, C. & Ehresmann, B. (1989) *J. Mol. Biol.* **207**, 417-431.

Wing, R., Drew, H., Takano, T., Broca, C., Tanaka, S., Itakura, K. & Dickerson, R. E. (1980) *Nature* **287**, 755-758.

Wobbe, C. R., Dean, F., Weissbach, L. & Hurwitz, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5710-5714.

Wolffe, A. P. (1988) *EMBO J.* **7**, 1071-1079.

Wolffe, A. P. & Brown, D. D. (1986) *ibid* **47**, 217-227.

Wolffe, A. P. & Brown, D. D. (1988) *Science* **241**, 1626-1631.

Wolffe, A. P., Jordan, E. & Brown, D. D. (1986) *Cell* **44**, 381-389.

Wolters, J. & Erdmann, V. A. (1988) *Nucleic. Acids Res.* **16** Supl., r1-r70.

Wormington, W. M. (1989) *Mol. & Cell. Biol.* **9**, 5281-5288.

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

Wormington, W. M., Schlissel, M. & Brown, D. D. (1983) *Cold Spring Barbor Symp. Quant. Biol.* **47**, 879-884.

Wormington, W. M. & Brown, D. D. (1983) *Dev. Biol.* **99**, 248-257.

Wrede, P. & Erdmann, V. A. (1973) *FEBS Lett.* **33**, 315-31'

Xing, Y. Y. & Worcel, A. (1989) *Mol. Cell. Biol.* **9**, 499-514.

Yamaguchi, M. & De-Pamphilis, M. L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1646-1650.

You, Q. & Romaniuk, P. J. (1990) *Nucleic Acids Res.***18**, 5055-5062.

You, Q., Veldhoen, N., Baudin, F. & Romaniuk (1991) *Biochemistry* **30**, 2495-2500.

Zaug, A. J., Been, M. D. & Cech, T. R. (1986) *Nature* **324**, 429-433.

Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988) *Biochemistry* **27**, 8924-8931.

Zaug, A. T., Kent, J. R. & Cech, T. R. (1984) *Science* **224**, 574-578.

Zengel, J. M. & Lindahl, L. (1986) *J. Bact.* **151**, 1261-1268.