NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.
THE ORGANIZATION AND EXPRESSION OF 5S RIBOSOMAL RNA GENES FROM THE NEMATODES MELOIDOGYNE ARENARIA AND CAENORHABDITIS ELEGANS

by

Haleh Vahidi
B.Sc.(Honors), Simon Fraser University, 1986

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

in the Department of Biological Sciences

- Haleh Vahidi, 1991

SIMON FRASER UNIVERSITY

December 1991

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L’auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L’auteur conserve la propriété du droit d’auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

APPROVAL

Name: HALEH VAHIDI
Degree: Doctor of Philosophy

Title of Thesis:
THE ORGANIZATION AND EXPRESSION OF 5S RIBOSOMAL RNA GENES
FROM THE NEMATODES MELOIDOGYNE ARENARIA AND
CAENORHABDITIS ELEGANS

Examining Committee:
Chairman: Dr. B.A. McKeown, Professor

Dr. B.M. Honda, Associate Professor,
Senior Supervisor,
Dept. Biological Sciences, SFU

Dr. D.L. Baillie, Professor,
Dept. Biological Sciences, SFU

Dr. M.J. Smith, Professor,
Dept. Biological Sciences, SFU

Dr. T. Borgford, Associate Professor,
Dept. of Chemistry, SFU

Dr. J.V. Price, Assistant Professor,
Dept. Biological Sciences, SFU
Public Examiner

Dr. R.H. Reeder, Section Chief,
Hutchinson Cancer Research Centre, Seattle
External Examiner

Date Approved December 9, 1991
I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

THE ORGANIZATION AND EXPRESSION OF 5S RIBOSOMAL RNA GENES FROM THE NEMATODES MELOIDOGYNE ARENARIA AND CAENORHABDITIS ELEGANS

Author: ___________________________

(signature)

HALEH VAHIDI

(name)

Dec. 9, 1991

(date)
ABSTRACT

The objective of this thesis was to study 5S ribosomal RNA gene transcription in the nematode *Caenorhabditis elegans*. Two approaches were taken: (i) a search for evolutionary conserved sequences, which might be functionally important, in the more distantly related nematode *Meloidogyne arenaria*; and (ii) functional studies of the DNA sequence and factor requirements for transcription in *C. elegans* extracts *in vitro*.

Surprisingly, *M. arenaria* 5S rRNA genes were found to be linked, in the opposite orientation, to the larger rRNA coding units. This is unusual because except for some lower eukaryotes, such as fungi and protozoa, and only one higher eukaryote studied to date, 5S rRNA genes occur as repeating units separate from the larger rRNA coding sequences.

The *M. arenaria* rDNA repeats occur in two major size classes of 9 kb and 5 kb, which are interspersed in the genome with other size classes. The 5 kb rDNA repeat and the other size classes may have been generated from 9 kb-like repeats by deletion/recombination events in the intergenic spacer, which contains tandem 129 bp subrepeats. Deletion events involving these subrepeats did occur when subcloned rDNA was propagated in a recA*+* host.

Further characterization of the *M. arenaria* rDNA repeats indicates that 5S DNA from the 9 kb repeat could be functional. The 5S DNA from the 5 kb repeat is probably a pseudogene because it would not be able to form the proper secondary structure and it does not have a proper transcription termination sequence. The 5 kb rDNA repeat also contains other deletions in the 3' end of the 28S rRNA genes.

A comparison of the *M. arenaria* and *C. elegans* 5S DNA sequences shows sequence conservation in the internal control region, but not upstream of
the 5S rRNA coding region. The *M. arenaria* 5S DNA, like that of *Xenopus*, cannot be transcribed efficiently in the *C. elegans* transcription extract. This may either be because they lack the upstream regulatory sequences or that the *C. elegans* factors cannot interact properly in the presence of heterologous internal control sequences. Fractionation of the *C. elegans* transcription extract on phosphocellulose columns results in a flowthrough fraction which is required for 5S rRNA but not tRNA transcription. This suggests the presence of a 5S-specific transcription factor.
QUOTATION AND DEDICATION

"With them the seed of wisdom did I sow,
And with my own hand labored it to grow:
And this was all the harvest that I reaped-
I came like water, and like wind I go"

From: The Rubaiyat of Omar Khayyam
translated by Edward FitzGerald

To my parents, Susan and Jamshid Vahidi who gave up everything they had, in search of comfortable life and good education for their children.

Also, to my grandmothers, Ghamar Ghafoori and Ehteram Janfaza, who have been separated from their children and grandchildren, for the same reasons.

Finally, to my brother Homayoun Vahidi, the other members of my family and all the hardworking true Iranians in the world, who have come a long way to get where they are.
I take this opportunity to express my gratitude to the members of my supervisory committee, my senior supervisor Dr. Barry Honda as well as Dr. David Baillie and Dr. Michael Smith, for their continued support and encouragement throughout both my undergraduate and graduate years at Simon Fraser University. I would also like to thank Dr. Thor Borgford for letting me experiment with biochemistry in his lab. Special thanks go to Don Nelson whose early guidance and encouragement has been of great value to me.

I would like to thank my colleagues Rob Linning and Les Willis, who have been like brothers to me the last few years.

I would like to thank all my wonderful friends, especially Anahita Modarress, Troy Riskie and Armin Samiei who have been there for me, through thick and thin. I also thank my new but dear friend R.S. who kept me company in the last stretch.

I would also like to thank people whom without their help throughout the years this work would not have been possible: Dr. A.T. Beckenbach, K. Beckenbach, K. Thomas, G. Wilson, D. Banfield, J. Boom, G. Kalmar, I. Scott, M. Khosla, R. Devlin, K. Sechley, F. Ouellette, J. Brown, V. Bourne, M. Rogalski, R. Rosenbluth, J. Price and many others.
TABLE OF CONTENTS

APPROVAL............................................................................................................... ii
ABSTRACT .................................................................................................................. iii
QUOTATION AND DEDICATION........................................................................... v
ACKNOWLEDGEMENTS......................................................................................... vi
TABLE OF CONTENTS............................................................................................... vii
TABLE OF FIGURES.................................................................................................. x

GENERAL INTRODUCTION....................................................................................... 1
  GENERAL .................................................................................................................. 2
  RNA POLYMERASE............................................................................................... 3
  TRANSCRIPTIONAL REGULATION IN PROKARYOTES........................................ 4
  TRANSCRIPTIONAL REGULATION IN EUKARYOTES.......................................... 4
    RNA Polymerase II ............................................................................................. 5
    RNA Polymerase I ............................................................................................. 7
    RNA polymerase III .......................................................................................... 9

THE REGULATION OF 5S RIBOSOMAL RNA TRANSCRIPTION IN
  CAENORHABDITIS ELEGANS.............................................................................. 11
REFERENCES ........................................................................................................... 13

CHAPTER ONE: Unusual sequences, homologous to 5S rRNA, in ribosomal DNA
  repeats of the nematode Meloidogyne arenaria.................................................. 20
  INTRODUCTION ..................................................................................................... 21
  MATERIALS AND METHODS.............................................................................. 22
    Source of nematodes .......................................................................................... 22
Enzymes and radionucleotides......................................................... 22
Genomic DNA isolation..................................................................... 22
Southern blots and hybridizations............................................... 22
Source of probes........................................................................... 23
DNA subcloning and characterization....................................... 23
DNA sequence analysis............................................................... 23
RESULTS.............................................................................................. 24

Linkage of M. arenaria 5S DNA to rDNA........................................ 24
Potentially functional 5S DNA and 5S rRNA pseudogenes in the M.
arenaria rDNA repeats................................................................ 27
DISCUSSION ...................................................................................... 34
REFERENCES...................................................................................... 37

CHAPTER TWO: Repeats and subrepeats in the intergenic spacer of rDNA from
the nematode *Meloidogyne arenaria* .............................................. 40

INTRODUCTION ................................................................................ 41
MATERIALS AND METHODS.............................................................. 42
Enzymes and radionucleotides.................................................. 42
Subcloning the deletion area..................................................... 42
DNA sequence analysis............................................................ 42
Smith-Birnstiel mapping............................................................. 42
RESULTS .............................................................................................. 43
Deletions within the M. arenaria 9 kb rDNA repeat......................... 43
Characterization of the rDNA deletion area.................................. 43
DISCUSSION ...................................................................................... 47
REFERENCES...................................................................................... 52
# CHAPTER THREE: The DNA sequence and factor requirements for in vitro transcription of 5S rRNA genes in the *Caenorhabditis elegans* transcription extracts

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>55</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>56</td>
</tr>
<tr>
<td>Preparation of the <em>C. elegans</em> pol III transcription extracts</td>
<td>56</td>
</tr>
<tr>
<td>Ion exchange chromatography of extracts on phosphocellulose columns</td>
<td>56</td>
</tr>
<tr>
<td>Preparation of exogenous <em>C. elegans</em> pol III</td>
<td>57</td>
</tr>
<tr>
<td>RNA polymerase assays</td>
<td>57</td>
</tr>
<tr>
<td>Transcription assays</td>
<td>58</td>
</tr>
<tr>
<td>RESULTS</td>
<td>59</td>
</tr>
<tr>
<td>Sequence requirements for 5S rRNA gene transcription in the <em>C. elegans</em> extracts</td>
<td>59</td>
</tr>
<tr>
<td>Transcription factor requirements for 5S rRNA gene transcription in the <em>C. elegans</em> extracts</td>
<td>64</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>66</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>71</td>
</tr>
</tbody>
</table>

# GENERAL DISCUSSION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFERENCES</td>
<td>80</td>
</tr>
</tbody>
</table>

# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFERENCES</td>
<td>82</td>
</tr>
</tbody>
</table>
TABLE OF FIGURES

CHAPTER ONE
1-1. Genomic blots of *M. arenaria* probed with rDNA and 5S DNA ..................25
1-2. Restriction map of *M. arenaria* rDNA repeats ........................................28
1-3. Ribosomal DNA sequences associated with *M. arenaria* 5S DNA in potentially functional repeats .................................................................30
1-4. DNA sequence analysis of potentially functional 5S DNA and 5S rRNA pseudogenes in the *M. arenaria* rDNA repeats ...............................................33

CHAPTER TWO
2-1. Deletions in the subcloned *M. arenaria* 9 kb rDNA repeat .......................44
2-2. Nucleotide sequences of the *Marenaria* rDNA deletion area ..................46
2-3. Deletion of subrepeats in cloned rDNA spacer .........................................48

CHAPTER THREE
3-1. The efficiency of transcription of 5S DNA from other nematodes and *Xenopus* in the *C. elegans* cell-free extracts .........................................................60
3-2. Comparison of nematode 5S DNA sequences in the region corresponding to the *Xenopus* internal control .........................................................62
3-3. Comparison of nematode 5S DNA sequences upstream of the coding region ........................................................................................................65
3-4. A P-11 breakthrough fraction of *C. elegans* extract is required in addition to the step fraction for 5S rRNA transcription ............................................67
GENERAL INTRODUCTION
GENERAL

Transcription may at first seem simple. It can occur in a test tube, even on a synthetic polynucleotide template DNA, when RNA polymerase is incubated at the appropriate temperature, pH and salt conditions with a supply of ribonucleotides (reviewed in Roeder 1976). However, many levels of complexity can be added to this picture, even in prokaryotes, in order for transcription to initiate only on specific genes, at specific sites and at specific times. This becomes obvious when transcription is studied in whole or reconstituted transcription systems in vitro or where possible in vivo. The regulation of transcription initiation has been a subject of interest for many investigators and that is what this thesis is about.

RNA polymerase, whether in eukaryotes or prokaryotes, must somehow recognize sequences just upstream of a gene to bind, to form an open complex and to initiate transcription at an precise nucleotide (usually a purine). The transcriptional regulatory sequences required for accurate initiation of transcription are called promoters. The efficiency of these promoters is determined by how well they correspond to a consensus sequence. The polymerase usually needs to interact with other protein factor(s) to recognize the promoter. Furthermore, promoters may be subject to positive or negative regulation through binding of still other protein factors to the same and/or different sequences. This provides another level of regulation.

Prokaryotes and unicellular eukaryotes, such as yeast, are the simplest systems to study the regulation of gene expression, under different environmental conditions. However, in metazoa, there are additional levels of control to allow expression of only subsets of the entire genome, at specific rates, in different types of cells and at different times in development. Many of the classical studies
were done with mammalian and *Xenopus* (reviewed in Wolfe 1991a, Polyanovsky and Stepchenko 1990, Muller et al. 1988, Wasylyk 1988b) cell-free transcription systems. However, simple genetic systems such as that of *C. elegans*, provide many advantages over other systems for the study of gene expression in eukaryotes.

**RNA POLYMERASE**

In bacteria, one RNA polymerase is responsible for the transcription of all RNA and protein coding genes. Eukaryotes have three different classes of polymerase; I, II and III. Pol I is only responsible for transcription of larger ribosomal RNA genes; 18S, 5.8S and 28S. Pol III is responsible for transcription of 5S rRNA, tRNA and other small RNAs. All the protein coding genes are transcribed by pol II, which also transcribes some RNA components of the splicing machinery (Rowland and Glass 1990).

Except for the monomeric RNA polymerases of some bacteriophages and mitochondria, all RNA polymerases are composed of many subunits (Rowland and Glass 1990, Lazcano et al. 1988). The *Eschericia coli* enzyme is composed of a core polymerase with a subunit composition of $\alpha_2\beta\beta^\prime\omega$. Different $\sigma$-factors associate with the polymerase under vegetative, heatshock or nitrogen fixation conditions (Travers 1987). The eukaryotic polymerases are composed of more than ten subunits; two to three large subunits and a number of small subunits. The $\beta^\prime$ subunit of the eubacterial RNA polymerase is homologous to the largest subunit of eukaryotic polymerases. The two largest subunits of the eukaryotic polymerases are highly conserved in different organisms. Furthermore, some of the smaller subunits are shared between the three polymerases. This is all as a result of similarities in the basic function of the RNA polymerases.
TRANSCRIPTIONAL REGULATION IN PROKARYOTES

In bacteria, the most commonly used promoter elements are the -10 TATAAT and the -35 TTGACA which are recognized when a predominant sigma factor (σ70 in E. coli) associates with the core polymerase (Gralla 1991, Travers 1987). The holoenzyme can directly bind to the -35 region, extend to the -10 region and initiate strand separation at those AT-rich sequences (reviewed in Watson et al. 1987). However, variations from the consensus lead to weaker promoters. Such promoters are often also regulated by activator proteins under appropriate conditions. There are also alternate sigma factors which recognize other specific promoters (reviewed in Helmann and Chamberlin 1988). These promoters are sometimes regulated by enhancer-like elements located at further distances. All the promoters discussed above may also be regulated by repressor proteins which often interfere with RNA polymerase binding.

TRANSCRIPTIONAL REGULATION IN EUKARYOTES

The basic mechanism of transcription initiation in eukaryotes bears many similarities to the prokaryotic picture, especially with respect to pol II. A minimal promoter is required for accurate initiation by the polymerase. There are then additional levels of regulation with sequence elements modulating the level of expression of different genes in different tissues and at different times in development.

In eukaryotes, the higher order structure of the DNA into chromatin also has a major impact on the availability of genes for transcription in vivo. Each 200 bp of DNA forms a nucleosome by wrapping around an octameric histone core of proteins consisting of two copies of H2A, H2B, H3 and H4 (reviewed in Van
A fifth histone, H1 pulls adjacent nucleosomes together into a repeating array. Furthermore, H1 molecules tend to cluster together in groups of eight and play a major role in condensation of DNA, serving as a general repressor mechanism for transcription (Zlatanova 1990). There is also further condensation of chromatin into coiled and looped domains leading to metaphase chromosomes. Different regions of chromosome may also vary on how tightly they are condensed, forming heterochromatin and euchromatin (Watson et al. 1987).

The point at which the control regions of genes can be subjected to transcriptional regulation, as it is observed in vitro with naked DNA, is not clear. During replication and cell division, DNA has to be repackaged into nucleosomes. Therefore, at this stage, the availability and the relative affinity of the regulatory proteins for DNA play a critical role in establishing a pattern of gene expression in a particular cell line (Zlatanova 1990). Histones can also undergo modifications such as acetylation and phosphorylation, leading to alterations in their tightness of binding to DNA and their allowing the access of regulatory proteins (Alberts et al. 1983).

There are still many other factors such as DNA torsion (Hirose and Suzuki 1988) and DNA methylation (Dynan 1989, Cedar 1988) that can affect the competence of a gene for transcription (Van Holde 1989). These in vivo effects are now being studied on genes whose transcription has been well studied in vitro, such as the 5S rRNA gene from Xenopus (eg. Tremethick et al. 1990, Morse 1989, Shimamura et al. 1988, Schlissel and Brown 1984).

**RNA Polymerase II**

The promoters for pol II transcription are upstream of the transcription start site (reviewed in Polyanovsky and Stepchenko 1990, Watson et al. 1987).
are general transcription factors which must bind to the basal promoter, before
the polymerase can bind and initiate accurately. One promoter element which
appears to be upstream of most pol II genes is the TATA box, positioned at
around -25 (reminiscent of the bacterial -10 element).

There is a temporal order of general transcription factors associating with
the TATA element, for transcription initiation to take place (reviewed in Greenblatt
transcription factor which directly interacts with the TATA element is TFIID (BTF1,
TBP). This interaction is further stabilized by binding of TFIIA (STF) and/or TFIIB
(BTF3). This preinitiation complex can then interact with pol II alone or in
association with the transcription initiation factors TFIIF (RAP30/74) and TFIIE
(BTF2). However, so far only the gene for TFIID has been cloned from various
organisms (eg. Schmidt et al. 1989) and a recombinant protein is available (eg.
Simmen et al. 1991). This leaves the exact nature of the interaction of TFIID with
the other factors ambiguous, to some extent, since many of the above studies
have been carried out with impure fractions.

In addition to the TATA box, there is often a requirement for other
regulatory elements for transcription to occur. One such sequence element is the
-80 CCAAT box. This element can be bound by a family of protein factors; NFI,
Another ubiquitous regulatory element upstream of many 'housekeeping' genes
is a GC-rich sequence which binds the Sp1 transcription factor. There are still
other general elements such as the 'octamer' site (reviewed in Kingsman and
Kingsman 1988).

There are also promoter elements and enhancers which are regulated; i.e.
they are activated only under certain conditions or in certain tissues (Muller et al.
1988, Wasylyk 1988b). Such promoters include the heatshock elements, the
steroid receptor binding elements and the AP1 binding sites. The best studied tissue specific enhancers are from immunoglobulin genes. Usually in these cases, for quick response to take place, a regulatory protein exists in the cell but requires a post-translational modification for DNA binding. These regulatory proteins may affect transcription in a positive or a negative fashion by direct interaction with the transcriptional control elements and/or by protein-protein interaction with proteins bound to ubiquitous promoter elements (reviewed in Herbomel 1990, Latchman 1990, Levine and Manley 1989, Ptashne 1988).

RNA Polymerase I

The promoters for pol I transcription, like those of pol II, are upstream of transcription start site (reviewed in Sollner-Webb and Mougey 1991, Reeder 1990, Sollner-Webb and Tower 1986, Mandal 1984). Although no homology can be discerned in the upstream sequences from rDNA of different species (Gerbi 1985), a ‘core promoter domain’ of about 35 bp around the initiation site as well as an 'upstream promoter domain' around residue -150 has been found to be required for transcription. The requirement for the upstream element was only evident through more stringent assay conditions in vitro (Henderson and Sollner-Webb 1990). Furthermore, there is a stereo-specific spacing requirement between the upstream and the core promoter domains, in order for transcription initiation to occur on the appropriate side of the DNA helix (Pape et al. 1990).

In most eukaryotes, rDNA occurs in tandem repeating units of 18S, 5.8S and 28S rRNA coding regions separated by a spacer. These spacers are often variable in length and consist of repeated DNA. Promoter duplications (Grimaldi et al. 1990, De Winter and Moss 1986) and enhancer elements (Pikaard et al. 1990, Dunaway and Droge 1989, Pape et al. 1989), as well as terminators
(Labhart and Reeder 1986), have been found within this repeated DNA (Reeder 1990, 1984).

In vitro studies, from mainly Xenopus and mouse transcription systems, have lead to the identification of two factors for transcription initiation by pol I. One factor is UBF (SF1 in rat, also TFIS in X. laevis) which interacts with both the core and upstream promoter domains (Pikaard et al. 1989, Dunaway 1989). The gene for the human UBF has been cloned and shown to have homology to the DNA binding domains of the chromosomal proteins HMG1 and HMG2 (Jantzen et al. 1990). Another factor, SL1 (factor D, TFID, TIF-IB, Rib-1 in X. laevis), which is responsible for species-specificity of rDNA transcription (Bell et al. 1989, Safrany et al. 1989), interacts with the upstream half of both promoter domains. These factors form stable complexes which remain through many rounds of transcription (McStay et al. 1991, Smith et al. 1990, Bell et al. 1988, Learned et al. 1986). In Acanthamoeba, however, a single factor TIF-I is sufficient for in vitro transcription (Paule 1990, Bateman et al. 1985).

There is also another protein factor (TTF I or Rib-2 in X. laevis) which has been shown to bind to terminator sequences within the rDNA spacer (Kuhn et al. 1990, McStay and Reeder 1990, Bartsch et al. 1988, Grummt et al. 1986). The binding of the termination factor to the termination sites at the end of one repeat seems to stimulate transcription initiation at the next repeat (Firek et al. 1989, Grummt et al. 1986, Henderson and Sollner-Webb 1986, McStay and Reeder 1986).

The regulation of rDNA transcription is linked to changes in cell growth. This type of regulation can be achieved rapidly through the modulation of the activity of the polymerase itself, since it is involved in the transcription of only rDNA (Nogi et al. 1991). Two proteins may be involved in this process. One, is TIFIC (TIF-IA, factor C), found associated with the polymerase which is required
for initiation, but not elongation (Schnapp et al. 1990, Mahajan and Thompson 1990). Another factor involved is protein kinase NII. Coordinate regulation of rDNA transcription with other growth-regulated pol II protein coding genes may be possible through shared sequence elements and transcription factors between the two polymerases (Walker and Reeder 1988).

**RNA polymerase III**

The promoters for genes transcribed by pol III were originally found to be internal to the transcription start site, unlike all other genes studied (reviewed in Wolfe AP 1991b, Palmer and Folk 1990, Geiduschek and Tocchini-Valentini 1988). The promoter of tRNA genes is bipartite and consists of domains called the A box and the B box (Hall et al. 1982). The A box is located at about nucleotides 8-19 and the B box is located at nucleotides 52-62, corresponding to the coding region of the D arm and the T arm of the tRNA product, respectively. Depending on the length of the variable arm, the two promoter domains can be separated by over 70 nucleotides. The promoter of the 5S rRNA genes were found to be located at about nucleotides 50-83. This promoter consists of a tRNA-like A box, and intermediate region and a 5S-specific C box (Geiduschek and Tocchini-Valentini 1988). Although 5'-flanking sequences were found to have both positive and negative modulating effect on tRNA transcription, originally, they were found to be dispensable for 5S rRNA transcription (Hall et al. 1982).

The protein factors required for transcription of pol III genes were identified as the general transcription factors TFIIIB and TFIIIC, as well as a 5S-specific transcription factor, TFIIIA (Taylor and Segall 1985, Shastry et al. 1982 and Segall et al. 1980). These factors were shown to form a complex on promoter regions and remain stable through many rounds of transcription (Lassar et al. 1983).
Many aspects of the above picture are now changing as in vitro transcription systems are being developed in more organisms and through technical advances in studying protein-nucleic acid interactions (see Palmer and Folk 1990, Murphy et al. 1989, Sollner-Webb 1988). In terms of the number of the general pol III transcription factors, the picture might not be so simple. First, with respect to TFIIIC, in yeast it is believed to be composed of two covalently linked DNA binding domains (Schultz et al. 1989), while the human TFIIIC can be chromatographically split into two components (Dean and Berk 1987). Second, fractionation of the Bombyx mori pol III transcription system indicates additional general factors TFIIID (Otonello et al. 1987) and TFIIIR (Young et al. 1991a), the latter of which is composed of RNA.

Another recent finding which is of interest is that in the yeast pol III system, it was shown that TFIIIB is the only true transcription initiation factor (Kassavetis et al. 1990). TFIIIC and TFIIIA are just required for the assembly of TFIIIB, upstream of the 5S rRNA gene where it interacts with pol III. Once TFIIIB is positioned and transcription is initiated, it can maintain multiple rounds of transcription, in absence of the other factors. Furthermore, the yeast TFIIIB has also been found to be chromatographically separable into two fractions (Kassavetis et al. 1991).

Further analysis of certain other small RNAs transcribed by pol III indicate promoters completely upstream of the coding region (Sollner-Webb 1988). The coding regions for U6 snRNA (Das et al. 1988) and 7SK RNA (Murphy et al. 1987) do have an A box, but in vitro transcription can solely be dependent on the upstream promoters. This is also the case for the pol III transcribed mitochondrial RNA processing RNA, MRP (Yan and Reddy 1991). These upstream promoters consist of pol II-like TATA boxes. In fact, the transcription factors required for the in vitro expression of mammalian U6 snRNA have been identified as TFIIIB and the pol II TATA box binding factor, TFIID (Waldschmidt et al. 1991). Most
surprisingly TFIIIC, which has been found to be required for the transcription of all pol III genes studied, is not required for the mammalian U6 transcription. However, the most bizarre example of the pol III upstream regulatory sequences is the RNase P RNA gene of *S. cerevisiae*. The coding region for this gene has a transcribed 5' leader sequence with tRNA-like A and B boxes (Lee et al. 1991).

Transcriptional regulatory elements have also been found downstream of pol III genes. Notable among those identified include the silk moth tRNA gene (Young et al. 1991b) and the yeast U6 snRNA gene (Brow and Guthrie 1990).

The picture is also now changing for 5S rRNA gene promoters. Studies from the last few years have indicated the additional requirement of regulatory sequences upstream of the coding region. Upstream regulatory sequences, usually consisting of a pol II-like TATA box, have been found upstream of 5S rRNA genes from *Drosophila* (Sharp and Garcia 1988), *Bombyx mori* (Morton and Sprague 1984), *Neurospora crassa* (Tyler 1987), *Acanthamoeba castellanii* (Zwick et al. 1991), human (Sorensen and Frederiksen 1991) and the loach *Misgurnus fossilis* (Felgenhauer et al. 1990). Furthermore, a transcription stimulatory factor was found to bind upstream of *Xenopus* 5S rRNA and tRNA genes (Oei and Pieier 1990).

THE REGULATION OF 5S RIBOSOMAL RNA TRANSCRIPTION IN *CAENORHABDITIS ELEGANS*

The objective of this thesis was to study the regulation of 5S rRNA gene transcription in the pol III *in vitro* transcription system of the nematode *Caenorhabditis elegans* (Honda et al. 1986). This was achieved through two approaches. The first approach involved a search for evolutionary conserved sequences, which might be functionally important. Previously, a similar approach
led to the identification of conserved sequences upstream of 5S rRNA genes from 1 kb repeats in the two related nematode species, *C. elegans* and *C. briggsae* (Nelson and Honda 1989, 1985). The first two chapters of this thesis are a study of the 5S rRNA genes and their unusual organization in the plant parasitic nematode, *Meloidogyne arenaria*. In the third and final chapter, the regulatory regions for 5S rRNA transcription from *M. arenaria* are compared to *C. elegans*.

The second approach involved functional analysis of DNA sequence and factor requirements for 5S rRNA gene transcription in the *C. elegans* extracts. Again, previously this approach was used to show that the conserved sequences upstream of *C. elegans* and *C. briggsae* 5S DNA repeats are required for *in vitro* transcription, in addition to the internal control regions (Nelson et al. 1991). In the third chapter of this thesis, sequence requirements for 5S rRNA gene transcription are further addressed by transcription of heterologous templates, from more distantly related nematodes. A preliminary analysis of factor requirements is also carried out by fractionation and reconstitution of 5S rRNA gene transcription *in vitro*. 
REFERENCES


De Winter RFJ, Moss T (1986) Spacer promoters are essential for efficient enhancement of X. laevis ribosomal transcription. Cell 44:313-318


Kassavetis GA, Braun BR, Nguyen LH, Geiduschek EP (1990) S. cerevisiae TFIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIBA and TFIIBC are assembly factors. Cell 60:235-245


Learned RM, Learned TK, Haltiner MM, Tjian RT (1986) Human rRNA transcription is modulated by the coordinate binding of two factors to an upstream control element. Cell 45:847-857


McStay B, Hu CH, Pikaard CS, Reeder RH (1991) xUBF and Rib1 are both required for formation of a stable polymerase I promoter complex in *X. laevis*. EMBO J 10:2297-2303

Morse RH (1989) Nucleosomes inhibit both transcriptional initiation and elongation by RNA polymerase III *in vitro*. EMBO J 8:2343-2351


Ottonello S, Rivier DH, Doolittle GM, Young LS, Sprague KU (1987) The properties of a new polymerase III transcription factor reveal that transcription complexes can assemble by more than one pathway. EMBO J 6:1921-1927

Palmer JM, Folk WR (1990) Unraveling the complexities of transcription by RNA polymerase III. Trends Biochem 15:300-304


17


Shastry BS, Ng SY, Roeder RG (1982) Multiple factors involved in the transcription of class III genes in Xenopus laevis. J Biol Chem 257:12,979-12,986


Yan Y, Reddy R (1991) 5' Flanking sequences of human MRP/7-2 RNA gene are required and sufficient for the transcription by RNA polymerase III. Biochim Biophys Acta 1089:33-39


CHAPTER ONE

Unusual sequences, homologous to 5S rRNA, in ribosomal DNA repeats of the nematode *Meloidogyne arenaria*
INTRODUCTION

One of the components of the large subunit of prokaryotic and eukaryotic ribosomes is the 5S rRNA. This RNA is small (about 120 nucleotides), it has a high degree of structural constraint and it is required in equimolar amounts with the other ribosomal components. Thus the gene encoding 5S rRNA has served as an attractive model for studies involving gene organization, gene expression and evolution (reviewed in Gerbi 1985).

In prokaryotes, the 5S rRNA coding region is linked to the larger rRNA genes (16S and 23S) and they are transcribed together in one large transcript which is then processed. In eukaryotes, the 5S rRNA gene is transcribed by pol III while the larger rRNA genes (18S, 5.8S and 28S) are co-transcribed by pol I and processed. In some lower eukaryotes, similar to the prokaryotic gene organization, the 5S DNA is linked to the larger rRNA genes (although mostly in the opposite orientation). However in all higher eukaryotes studied to date, with the exception of a marine copepod (Drouin et al. 1987), 5S rRNA genes occur in separate repeating units from the rDNA (Gerbi 1985).

The objective of this chapter was to isolate and characterize 5S DNA from the plant parasitic nematode *M. arenaria*. This heterologous 5S DNA was to be used later to identify sequences important for transcription in the *C. elegans* cell-free extracts (Honda et al. 1986), based on conservation of functionally important sequences with the *C. elegans* 5S DNA and the ability of the heterologous DNA to interact with the *C. elegans* transcription machinery.

Surprisingly, the *M. arenaria* 5S DNA was found to be linked to the larger rDNA. Furthermore, the rDNA repeats appear to have a complex organization, where functional and defective rDNA repeats of various size classes can be found interspersed in the genome. The characterization of the major 9 kb and 5 kb
M. arenaria rDNA repeats is presented in this chapter (Vahidi et al. 1988, Vahidi et al. 1991).

MATERIALS AND METHODS

Source of nematodes

M. arenaria race B was provided by M. McClure (University of Arizona).

Enzymes and radionucleotides

Enzymes were purchased and used as recommended by suppliers. 32P-dNTPs (800 or 3000 Ci/mmol) were supplied by Amersham.

Genomic DNA isolation

Genomic DNA was prepared as previously described (Curran et al. 1986).

Southern blots and hybridizations

Genomic DNA was digested with the appropriate restriction enzymes and 1 ug of DNA was loaded in each lane of a 0.7% agarose gel. Following electrophoresis, the DNA was blotted onto nitrocellulose bidirectionally (Smith and Summers 1980). The DNA probes were labeled by nick translation (Rigby et al. 1977). Hybridization of probes to duplicate filters were done at 62° C, in 5xSSPE and washed at for 1 hr at the same temperature in 2xSSPE (Maniatis et al. 1982).
Source of probes

The 5S DNA probe was a 5’ exo III, S1 deletion of the C. elegans 5S DNA (Nelson and Honda 1985), consisting of the entire coding region plus 6 bp of 5’ sequence. The 18S, 5.8S and 28S rDNA probes were obtained from appropriate restriction fragments derived from the total 7 kb rDNA repeat from C. elegans (Ellis et al. 1986).

DNA subcloning and characterization

C. elegans 5S DNA and rDNA probes were used to identify and isolate M. arenaria 5S DNA containing repeats, either from plasmid libraries of PstI digested genomic DNA cloned into pUC13 (Vahidi et al. 1988) or from lambda phage libraries of partial Sau3A digested genomic DNA cloned into EMBL3 (Frischauf et al. 1983). The various coding regions and their orientations were identified by hybridization with the appropriate restriction fragments from C. elegans 5S DNA and rDNA.

DNA sequence analysis

All DNA sequencing was done according to Hattori and Sakaki (1986). For sequencing the M. arenaria 5S DNA from the 5 kb PstI rDNA repeat, the central 800 bp Hind III fragment containing 5S rRNA homology (see Fig. 1-2) was subcloned into pUC19 (Vieira and Messing, 1982) and overlapping deletions were constructed using exo III and S1 (Henikoff 1984). For sequencing the 5S DNA from the 9 kb PstI rDNA repeat, the central 3.1 kb Hind III fragment (Fig. 1-2) was subcloned into pUC13 and deletions were constructed as before. The 5S DNA
from the phage rDNA repeats were sequenced similarly, using the appropriate restriction fragments (Vahidi et al. 1991).

The 5.8S rDNA sequences from the 9 kb and the 5 kb rDNA repeats were obtained by sequencing the ends of the *PstI* subclones. Sequence from the 5' end of the 18S rDNA from the 9 kb repeat was obtained from the end of the central 3.1 kb Hind III subclone (Fig. 1-2). Sequence from the 3' end was obtained from sequencing the ends of subcloned central 4.5 kb EcoRI fragment. The 3' end of 28S rDNA from the 9 kb repeat was obtained from sequencing the opposite end of the same 4.5 kb EcoRI fragment.

RESULTS

**Linkage of *M. arenaria* 5S DNA to rDNA**

When 5S DNA and rDNA probes from *C. elegans* were hybridized to bidirectional southern blots of *M. arenaria* genomic DNA digested with various enzymes, surprisingly they hybridized to identical bands (Fig. 1-1). Similar results were obtained with *M. hapla* genomic DNA (data not shown). Also, when *M. arenaria* EMBL3 genomic libraries were screened with the same probes, every plaque that hybridized to 5S DNA also hybridized to rDNA (data not shown).

Two major bands of 9 kb and 5 kb, in size, hybridized with *C. elegans* probes on *M. arenaria* blots of genomic DNA digested with *PstI* or *SmaI*, along with some other dispersed bands (Fig. 1-1). Single *PstI* digested 9 kb and 5 kb rDNA repeats were subcloned into pUC13 (D.W. Nelson). Contiguous rDNA repeats were obtained by cloning partial Sau3A digested genomic DNA into EMBL3 lambda phage.
FIGURE 1-1

Genomic blots of *M. arenaria* probed with rDNA and 5S DNA

*M. arenaria* genomic DNA was digested with *PstI* (P) or *SmaI* (S), electrophoresed, bidirectionally blotted and probed with *C. elegans* cloned rDNA (A) or 5S RNA gene (B) as described in Materials and Methods. The major 9 kb and 5 kb rDNA repeats are indicated.
The appropriate *M. arenaria* rDNA subclones were isolated and further characterized. Figure 1-2 shows the restriction maps of 9 kb (M9) and 5 kb (M5) rDNA repeats from the plasmid subclones, in the middle. Above and below are similar 9 kb-like (A2) and 5 kb-like (B1) repeats isolated from the phage (characterized by A. Purac). Various regions of the *C. elegans* 5S DNA and rDNA were used to identify and orient the different coding regions, as described in Materials and Methods.

The *M. arenaria* rDNA repeats from different size classes have very similar organization. Moreover, they all have 5S rRNA homologous sequences located in the opposite orientation to the 18S, 5.8S and 28S rRNA coding sequences. The organization of the phage rDNA repeats (Fig. 1-2) indicate that functional and nonfunctional repeats of variable length may occur adjacent in the genome (A. Purac and J. Leblanc). Some repeats contain major deletions in the 5' end of 18S DNA (A2, repeat III), both the 5' (A2, repeat II) and 3' (M5 and B1) ends of 28S DNA and/or 3' end of 5S DNA (M5 and B1). These deletions have been verified by DNA sequencing (data not shown).

The rDNA sequences from potentially functional repeats (M9 and A2, repeat I) have been verified by DNA sequencing of certain regions (see Materials and Methods). Figure 1-3 shows comparison of the *M. arenaria* rDNA to *C. elegans* rDNA sequences.

**Potentially functional 5S DNA and 5S rRNA pseudogenes in the *M. arenaria* rDNA repeats**

The 5S rRNA homologous sequences in *M. arenaria* rDNA repeats were subjected to sequence analysis to see if they are likely to be functional. Comparison of these sequences to *C. elegans* 5S DNA is shown in Figure 1-4.
FIGURE 1-2

Restriction map of *M. arenaria* rDNA repeats

The major 9 kb (M9) and 5 kb (M5) *PstI* rDNA repeats from plasmid subclones and 9 kb-like (A2) and 5 kb-like (B1) rDNA repeats from phage (characterized by A. Purac) were restriction mapped and the individual rRNA coding regions were roughly localized as described in Materials and Methods. The phage rDNA is shown split at *PstI* sites and the Roman numerals refer to the number of contiguous rDNA repeats. The *EcoRI* and *HindIII* sites are not shown for phage B1. Open arrows indicate presumed defective sequences. Enzyme sites: P = *PstI*, E = *EcoRI*, S = *SmaI*, X = *XbaI* and H = *HindIII*. 
FIGURE 1-3
Ribosomal DNA sequences associated with *M. arenaria* 5S DNA in potentially functional repeats

A) The 5' and 3' ends of *M. arenaria* 18S rDNA from the 9 kb repeat (M9) and the phage 9 kb-like repeat number I (A2) were compared with the same region of *C. elegans* rDNA (Ce; 5': nts 3-148; 3': nts 1423-1520). Sequence identity is indicated by "=" and deletions are indicated by ".".

B) The *M. arenaria* 5.8S rDNA from the same repeats as in A is compared to *C. elegans* 5.8S rDNA sequences

C) The 3' end of *M. arenaria* 28S rDNA from the same repeats as in A is compared to the same region of *C. elegans* rDNA (nts 3316-3380).

A

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 18-5'</td>
<td>-----CC==A.==..............G==T==TT==C----------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ce 18-5'</td>
<td>CCTGATTGATTCTGTCAGCGCGATATGCTCAAGTAATAAAGATTAAAGCATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2 18-5'</td>
<td>-----CC==A.==..............G==T==TT==C----------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 18-5'</td>
<td>-----TA==A==A==GTTT==CGA==CC----------------------------C=A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ce 18-5'</td>
<td>CATGCTTTAGATTCAATCA...AT...GAAATGCTACGGCTATGAAGA.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2 18-5'</td>
<td>-----TA==A==A==GTTG==CGA==CC----------------------------A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 18-5'</td>
<td>TG=C==T==T==.==.==.==T==.======T==T==.==.==.==</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ce 18-5'</td>
<td>.G.CA.GATATCACCTTTATCCGGGATTCTCAATGGAATGA.A.CTGCCTGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
C

<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 28-3’</td>
<td>=.==CAA====C====ATC======TT.C.TTCT.ACT==AA=T.=GTA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ce 28-3’</td>
<td>GCTTAT..GCCCTAGTTA.TGGCCTTGCGTGTAAGTAATTCTGCGACG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2 28-3’</td>
<td>=.==CA==C====ATC====G====TT.C.TTCT.ACT==AA=T.=GTA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

60

<table>
<thead>
<tr>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 28-3’</td>
</tr>
<tr>
<td>Ce 28-3’</td>
</tr>
<tr>
<td>A2 28-3’</td>
</tr>
</tbody>
</table>
DNA sequence analysis of potentially functional 5S DNA and 5S rRNA pseudogenes in the *M. arenaria* rDNA repeats

The 5S RNA-like sequences within the 9 kb (M9), 9 kb-like (A2, repeat I), 5 kb (M5) and 5 kb-like (B1, all repeats) *M. arenaria* rDNA repeats (isolated from both plasmid and phage subclones) were compared to 5S DNA sequences from *C. elegans* (Ce). See Fig. 1-3 for the description of sequence identity and deletion symbols. The A at position 110 of M9/A2 is not present in M9.

<table>
<thead>
<tr>
<th>Position</th>
<th>M9/A2</th>
<th>Ce</th>
<th>M5/B1</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-50</td>
<td>=A====C=G=A==</td>
<td>GCTTACGACCATATCACGTGGAATGCACGCATCCCGTGCCGATCTGGCAA</td>
<td>=A====C=G=A==</td>
</tr>
<tr>
<td>60-100</td>
<td>======C=G=CT==.=====A=====T==</td>
<td>GTTAAGCAACGTGAG..TCCAGTTAGTACTTGGAATCGGAGACGCCTGG</td>
<td>======C=G=CT==.=====A=====T==</td>
</tr>
<tr>
<td>110-120</td>
<td>=====G.==AG==C.=.==T====T</td>
<td>GAATCCTGG.ATGTGGTAAGCTTTTTG</td>
<td>=====G.==C==A==.==GGGG</td>
</tr>
</tbody>
</table>
Similar to the rDNA sequences, the 5S DNA from the 5 kb repeat (M5) and the 5 kb-like phage repeat (B1, all repeats) are not likely to be functional and they are probably pseudogenes. These 5S DNA sequences have no pol III transcription terminator (run of Ts, Platt 1986) and they cannot form appropriate secondary structure (Erdmann and Wolter 1986) at the 3’ end.

The 5S DNA from the 9 kb repeat (M9) and the 9 kb-like phage repeat (A2, repeat I) are, however, likely to be functional. The 5S DNA from these two repeats are identical, except for a single base insertion in A2 (position 110). Furthermore, Northern blots probed with an oligo complementary to the 3’ end of the phage 9 kb-like 5S DNA suggest that these genes could be expressed in vivo (data not shown). There is also a duplication of the 3’ end of the 5S DNA, a few nucleotides downstream of the full-length coding sequence. However, the 5S DNA sequences from this duplication diverge from the 5S DNA in the 9 kb repeat, at the same nucleotide as the 5S DNA sequences from the 5 kb repeat (the nucleotide sequence for this duplication is shown with the rest of the rDNA intergenic sequences in the next chapter; Fig. 2-2).

DISCUSSION

The results presented in this chapter raise some questions as to the evolutionary significance of M. arenaria 5S DNA organization. There have been conflicting views on whether linkage of 5S DNA to rDNA represents a transition state between prokaryotes and eukaryotes (Maizels 1976) or whether it has occurred independently in different evolutionary lineages (Appels and Honeycutt 1986, Clark 1987). The latter view is supported by the findings in the most primitive forms of life, archaebacteria (Culham and Nazar 1988) and mycoplasma (Taschke et al. 1986), where 5S rRNA genes are not found to be linked to rDNA.
Similarly, the latter view would be supported by the finding of 5S rRNA sequences within the rDNA of a higher eukaryote, as observed in *M. arenaria* (Vahidi et al. 1988; this chapter) and in the genus *Calanus* (Drouin et al. 1987). Movement and fixation of 5S rRNA genes into variant forms of organization have been suggested by the tandem 5S DNA pseudogene cluster found in the rDNA of a thermophilic fungus (Wong et al. 1984, Nazar and Wong 1985), the dispersed 5S DNA of *Neurospora crassa* (Selker et al. 1981) and the tandemly arranged variant 5S DNA of *Saccharomyces cerevisiae* (McMahon et al. 1984).

Movement of 5S DNA into the rDNA during the evolution of *Meloidogyne* could have occurred after the separation of the phylum *Nematoda*, as 5S DNA is not found to be linked to rDNA in *C. elegans* (Nelson and Honda 1985, Files and Hirsh 1981), *C. briggsae* (Nelson and Honda 1989) and *A. lumbricoides* (Nilsen et al. 1989, Back et al. 1984). This arrangement would then be fixed by mechanisms of "molecular drive" (Dover 1982). These mechanisms include gene conversion and unequal crossing over, which would be possible through the *M. arenaria* rDNA spacer subrepeats (discussed in Chapter Two, Vahidi and Honda 1991). The *Meloidogyne* reproductive cycle (meiotic and mitotic parthenogenesis) and the variable ploidy levels in different species (Triantaphyllou 1985) may also have contributed to the unusual rDNA evolution.

The functional significance of the linkage of the *M. arenaria* 5S DNA to rDNA is not clear. The most obvious reason for this linkage would be coordinate expression of these RNA components of ribosomes. In eukaryotes, 5S DNA and rDNA are transcribed by different polymerases. If the level of transcription of these genes is kept comparable, coordinate regulation may be achieved by co-regulation of gene copy number. Alternatively, *In vitro* studies of *Neurospora* 5S DNA, rDNA and ribosomal protein gene (pol II) transcription suggest that a common sequence element may solve the problem of coordinate expression and
the genes do not have to be linked (Tyler 1990). Binding of a common transcription factor to conserved sequence elements in genes coding for ribosomal components may serve as a signal to the different polymerases. Such a conserved element has also been found upstream of C. elegans 5S DNA and rDNA (Nelson et al. 1991).
REFERENCES


Curran J, McClure MA, Webster JM (1986) Genotypic differentiation of Meloidogyne populations by detection of restriction fragment length difference in total DNA. J Nematol 18:83-86


Tyler BM (1990) Two complex regions, including a TATA sequence, are required for transcription by RNA polymerase I in Neurospora crassa. Nucl Acids Res 18:1805-1811


CHAPTER TWO

Repeats and subrepeats in the intergenic spacer of rDNA from the nematode *Meloidogyne arenaria*
INTRODUCTION

In eukaryotes, rDNA is usually composed of tandem repeating units coding for 18S, 5.8S and 28S rRNA, separated by an intergenic spacer (IGS) which is often repetitious in sequence and variable in length (reviewed in Gerbi 1985). The function of the IGS and the extent to which it is transcribed is not clear. It has been suggested that the function of the IGS is to promote unequal crossing over (Dover 1982), and maintain homogeneity between rDNA repeats of the same organism (Gerbi 1985). More recently, promoters, enhancers and transcription termination signals, as well as origins of replication have been identified in the IGS of different organisms (see Gerbi 1985). On the other hand, the IGS from the rDNA of the nematode Ascaris lumbricoides, like that of C. elegans, is short and does not appear to have any repeated elements and promoter duplications (Muller et al. 1990).

Unlike transcription by pol II and pol III, sequences and transcription factors required for pol I transcription have proven to be highly species-specific, and there is no sequence conservation between the IGS sequences of distant species (Gerbi 1985).

Ribosomal DNA repeats of the plant-parasitic nematode, Meloidogyne arenaria, are heterogeneous in size and appear to contain 5S rRNA gene sequences within the IGS (Vahidi et al. 1988). Moreover, in a recA+ host, plasmid clones of a 9 kb rDNA repeat show deletion events within the 2 kb IGS, between 28S and 5S DNA sequences. These deletions appear to result from a reduction in the number of tandem 129 bp repeats. The loss of such repeats might explain how rDNA length heterogeneity, observed in the Meloidogyne genome, could have arisen.
Each 129 bp IGS subrepeat, of the *M. arenaria* rDNA, also contains three copies of an 8 bp subrepeat, which has sequence similarity to an element found in the IGS repeats of some plant rDNAs.

**MATERIALS AND METHODS**

**Enzymes and radionucleotides**

Restriction and other enzymes were purchased from and used as recommended by the suppliers. $^{32}$P-dNTPs (800 or 3000 Ci/mM) were supplied by Amersham.

**Subcloning the deletion area**

The central 3.1 kb *Hind* III fragment (see Fig. 1-2), from the 9 kb *M. arenaria* rDNA repeat, encompassing the deletion area (pM205H), as well as three deleted versions (pM205Ha, pM205Hb and pM205Hc) were subcloned into pUC13 (Vieira and Messling 1982).

**DNA sequence analysis**

All DNA sequencing was done according to Hattori and Sakaki (1986). Overlapping deletions were constructed on pM205H (see section on subcloning), using exo III and S1 (Henikoff 1984).

**Smith-Birnstiel mapping**

Plasmid DNA of pM205H, pM205Hc, pM205Hb and pM205Hc was prepared by the alkaline lysis method and supercoiled, nick-free plasmid DNA was obtained by gel-purification (Maniatis et al. 1982). Plasmid DNA was linearized by digestion with *Sal*I and end labelled, using klenow fragment of DNA
polymerase I (Maniatis et al. 1982). The DNA was digested with Rsal and the appropriate Sam-Rsal end labelled fragments were gel-purified. About 10,000 cpm of Sam-Rsal end labelled fragments were partially digested with HinfI in the presence of unlabelled carrier plasmid DNA. The digests were electrophoresed on a 0.7% agarose gel. The gel was dried and subjected to autoradiography (see Fig. 2-3).

RESULTS

Deletions within the *M. arenaria* 9 kb rDNA repeat

When the subcloned *M. arenaria* 9 kb rDNA repeat was propagated in a recA+ host (*Escherichia coli* JM83), and resultant plasmid DNA was digested with restriction enzymes, a ladder of smaller bands were seen below fragments originating from the central region of the 9 kb repeat (Fig. 2-1). These deletions can be localized to a 2 kb region of the IGS, between the 3' ends of the 28S and the 5S DNA homologous sequences.

Characterization of the rDNA deletion area

The *M. arenaria* rDNA IGS and three deleted versions (pM205H, pM205Ha, pM205Hb and pM205Hc) were subcloned, by gel-purifying (Maniatis et al. 1982) smaller HindIII fragments derived from the original 9 kb repeat (Fig. 2-1 B). Sequence analysis of pM205H revealed that the spacer is composed of fourteen complete copies and a defective copy of a 129 bp repeat (Fig. 2-2). The repeats start about 60 bp away from the 3' end of the 28S gene and the defective copy ends 20 bp from the end of 5S rRNA homologous sequences. Sequence analysis did not indicate the presence of these 129 bp repeats in the *M. arenaria* 5 kb rDNA IGS (data not shown).
FIGURE 2-1

Deletions in the subcloned *M. arenaria* 9 kb rDNA repeat

(A) Electrophoresis of a *Hind*III (H) digest of the cloned 9 kb rDNA on a 0.7% agarose gel, showing deletions from a 3.1 kb fragment.

(B) A *Hind*III map of the 9 kb rDNA, showing also the position of the various coding regions (derived from Vahidi et al. 1988). The subclone pM205H can be seen on the right, showing the area in which the deletions occurred. The internal *Rsa*I (Rs) sites are also shown on this subclone. Three deleted versions of the IGS were also subcloned (pM205Ha, pM205Hb and pM205Hc) by gel-purifying (Maniatis et al. 1982) smaller *Hind*III fragments, derived from the largest 3.1 kb fragment.
The image shows a gel electrophoresis and a diagram of a DNA molecule. The gel contains bands atkb 4.7, 3.1, 2.0, and 1.9, which are likely fragments of DNA. The diagram on the right side of the page illustrates a DNA molecule labeled as pM205H, with regions marked as pUC13, 5.8S, 18S, 5S, and 26S. The area of deletions is indicated by brackets.
FIGURE 2-2

Nucleotide sequences of the *M. arenaria* rDNA deletion area

The plasmid pM205H was sequenced from the 28S end up to the 5S rRNA coding regions (see Fig. 2-1 B). The sequences enclosed by the bracket are tandemly repeated fourteen times. The asterisk above the T in the middle of the repeat indicates the single base that is A instead, in repeat number 4. A 9 bp core sequence repeated three times just downstream of the 28S DNA and also once within each repeat is double underlined. The extended sequence identity between some of these repeats is indicated by single line underneath. The boxed nucleotides are the 8 bp subrepeats CAGGCA(A/T)T within each repeat. The 3' ends of the 28S and the 5S DNA sequences are also indicated with double underline and their orientation is indicated by an arrow.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>28S</td>
<td></td>
</tr>
<tr>
<td>AAGCTTAGCAGCTCGTGCATGTGAACGTGGGGCGCAAAATCATTTGTA</td>
<td>50</td>
</tr>
<tr>
<td>TACGACTTGAATACTGGTCAGGGGTTGTAATCATAGACGCTCTTCGC</td>
<td>100</td>
</tr>
<tr>
<td>TGCAATCTTGGAGACTTTAGCCTTGTACCGATGATTGTCTTTTTATAT</td>
<td>150</td>
</tr>
<tr>
<td>ACTATAGACCTGTTTATCCCAAGTGTGCTAGTATCCCAAGTGTCGCAATTTGACCAAGTGCTGTCGATGGGT</td>
<td>200</td>
</tr>
<tr>
<td>[GCCCTAAGCAGGCAATTTGACCAAGTGCTGACCAAGTGGGTT</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>294</td>
</tr>
<tr>
<td>TTGCTAGGTCGACGTGATCCGATTTAAATTTT</td>
<td>341</td>
</tr>
<tr>
<td>GTAGACCCAAAATTAAAAAAACCTTAAATTATTATTATTTTTTGGATGAGAGAC</td>
<td>391</td>
</tr>
<tr>
<td>TCATTCCAGGACGCTCCCTGCCGTTCCCAAGTAC</td>
<td>421</td>
</tr>
</tbody>
</table>
We wanted next to determine if the deletions in the *M. arenaria* cloned rDNA involved reduction in the number of 129 bp repeats, as would be expected by unequal crossing over. To do this, the single *Hinf*I site in each repeat was used to compare the deleted subclones to pM205H, the largest clone from which they derived. This was accomplished by labelling the 28S end and doing a partial *Hinf*I digest (Smith and Birnstiel 1976). It can be seen from Figure 2-3 that the deletion subclones do indeed have missing bands. This experiment demonstrates that there are two repeats missing from pM205Ha, four missing from pM205Hb and five from pM205Hc. Sequence analysis indicates that all the deletion subclones have the complete 28S coding region (data not shown).

**DISCUSSION**

Spacer heterogeneity is not uncommon in rRNA gene clusters (reviewed by Gerbi, 1985). In terms of a possible mechanism for generating repeat heterogeneity in *M. arenaria*, we have presented evidence of potential recombinational events and deletions associated with an individual *M. arenaria* 9 kb rDNA repeat. These recombinational/deletion events could be mimicking events in the evolution of the shorter rDNA size classes. This has also been previously observed with mouse, *Xenopus* and *Vicia faba* cloned rDNA (Arnheim and Kuehn 1979; Morgan and McMahon 1986; Rogers and Bendich 1988). Recombinational “hot spots” have also been observed in yeast rDNA (Keil and Roeder 1984).

There are several interesting features in the *M. arenara* rDNA spacer repeat sequences. One is that they are identical except for a single base change in repeat number 4. Most IGS analyzed to date are more complex; either they are not composed of just one type of repeat, as in the case of *Xenopus* (Boseley et al. 1979), or there is some sequence variability between the individual repeats,
(A) Strategy for comparing the number of 129 bp repeats in subclones from the deleted area (see also Materials and Methods).

(i) The plasmid DNA was linearized by digestion with SalI (S) and endlabeled, using klenow (Maniatis et al. 1982). The labelled ends are shown with an asterisk.

(ii) The DNA was digested with Rsa I (Rs) and the appropriate SalI-RsaI endlabeled fragments were gel-purified.

(iii) About 10,000 cpm of SalI-Rsal endlabeled fragments were partially digested with Hinfl (Hf) in the presence of unlabeled carrier plasmid DNA.

(B) Autoradiogram of the partial Hinfl digests from (A) electrophoresed on a 0.7% agarose gel. Lane 1-4, partial Hinfl digests of pM205H, pM205Ha, pM205Hb and pM205Hc; respectively.
(i) pUC13

(ii) *Rsa* I digest

(iii) *Hin* fl partial digest

B 1 2 3 4
as in the case of *Drosophila* (Hayward and Glover 1989). The structure of the *Meloidogyne* rDNA IGS is similar to that of *Vicia faba* where the spacer repeats are completely homogeneous (Yakura et al. 1984). This complete homogeneity may be as a result of a recent amplification or maintenance by mechanisms of molecular drive (Dover 1982) and fixation.

Another interesting feature of the spacer repeats is a sequence, repeated three times downstream of the 28S coding region, as well as once within each 129 bp repeat. The core sequence is CCAAGTGTC and it is shown double underlined in Figure 2-2. There is also some extended sequence identity between some of the repeats, which is indicated by a single line underneath.

Another, repeated sequence in the *M. arenaria* rDNA IGS is a partial duplication of the 3' end of the 5S DNA, about 35 bp downstream of the original 3' end (Fig. 2-1 B).

Finally, the sequence CAGGCAAT is repeated three times (the third subrepeate is CAGGCATT), within the *M. arenaria* IGS 129 bp repeats. These subrepeats, spaced at a distance of 10 bp from each other, are boxed in Figure 2-2. This element does not seem to bear any similarity to sequences found in animal rDNA IGS repeats, such as those of *Xenopus* (Boseley et al. 1979), *Drosophila* (Hayward and Glover 1989) or mammals (Yavachev et al. 1986; La Volpe et al. 1985). However, it is interesting that this element has some sequence identity with an element found in the rDNA IGS repeats of the pea, *Pisum sativum* L. cv. Alaska. The sequence of the pea element is CATGCM and it is present once within each of nine 180 bp IGS repeats, as well as 19 bp upstream of the transcription initiation site and twice more between the end of the repeats and the site of initiation (Piller et al. 1990). The authors suggest that this pea sequence represents a core promoter/enhancer element. The nematode transcription initiation site is, however, not immediately downstream of the 129 bp subrepeats.
and is preceded by 5S DNA and more IGS sequence (Fig. 2-1). A sequence identical to the pea 8 bp element can also be found twice in the rDNA IGS repeats of *Vicia faba* (Yakura et al. 1984) and once in the IGS of wheat rDNA, with a single base change to CATGGAAA (Barker et al. 1988).

Should this plant rDNA sequence element prove to be important functionally, the presence of a related sequence in a plant parasitic nematode could have interesting implications in terms of host-parasite evolutionary relationships. Identification of closely related sequences in such diverged organisms, provides strong evidence for horizontal (non-sexual) gene transfer. Other examples of nonviral horizontal gene transfer exist with the glyceraldehyde-3-phosphate dehydrogenase gene of *E. coli* (Doolittle et al. 1990) and transposable elements of nematodes (Abad et al. 1991) and *Drosophila* (Houck et al. 1991).

In the case of *D. melanogaster*, P elements may have been acquired from another *Drosophila* species through a semi-parasitic mite which could have acted as a vector for the gene transfer (Houck et al. 1991). The parasitic interaction of *Meloidogyne* with its plant hosts may also allow such horizontal gene transfer. Since there is no sexual reproduction in *M. arenaria*, horizontal gene transfer may serve as a means of exchanging genetic material. Alternatively, *M. arenaria* may serve as a vector for transfer of genetic material between plants (see General Discussion for further discussion of where in *Meloidogyne* life cycle such gene transfer could occur).
REFERENCES


Muller E, Neuhaus H, Tobler H, Muller F (1990) Unusual transcription termination of the ribosomal RNA genes in Ascaris lumbricoides. EMBO J 9:2849-2856


CHAPTER THREE

The DNA sequence and factor requirements for \textit{in vitro}
transcription of 5S rRNA genes in the \textit{Caenorhabditis elegans}
transcription extracts
INTRODUCTION

The genes coding for 5S ribosomal RNA are transcribed by RNA polymerase III in eukaryotes. These genes were originally found to be unusual because of the sole requirement of an internal promoter for transcription (reviewed in Geiduschek and Tocchini-Valentini 1988, Murphy et al. 1989, Palmer and Folk 1990). Cell-free transcription extracts, developed in yeast (Taylor and Segall 1985), Xenopus (Shastry et al. 1982) and mammalian (Segall et al. 1980) systems, contain general pol III transcription factors TFIiIB and TFIIIC and the 5S-specific transcription factor TFIIIA. These factors bind to the 5S DNA internal promoters and form stable complexes, lasting through many rounds of transcription (Lassar et al. 1983).

Later, upstream regulatory sequences were found 5' to several 5S rRNA genes, either consisting of a pol II-like TATA box (Tyler 1987, Morton and Sprague 1984) and/or sequence elements conserved only in closely related species (Sharp and Garcia 1988, Niefeld et al. 1988). The interaction of this region with factors is not clear. One objective of this chapter was to see if a conserved element, found upstream of 5S rRNA genes from 1 kb repeats of C. elegans and C. briggsae (Nelson and Honda 1989), also existed in more distantly related nematodes.

Although there is some conservation when the sequences corresponding to the internal control region (ICR) are compared between M. arenaria and C. elegans, there is no sequence similarity upstream of the 5S rRNA coding sequences. Furthermore, in a functional assay, the heterologous template failed to be transcribed efficiently in the C. elegans extracts. This suggests that the nucleotide sequence differences in the M. arenaria 5S DNA ICR and/or the 5' -
flanking sequences are critical for interaction with the *C. elegans* transcription machinery.

Another objective of this chapter was to fractionate the *C. elegans* transcription extract to see if there is a 5S-specific transcription factor. Fractionation of the extract on phosphocellulose columns results in a step fraction sufficient for tRNA gene transcription, but requires the breakthrough fraction for 5S rRNA gene transcription.

**MATERIALS AND METHODS**

**Preparation of the *C. elegans* pol III transcription extracts**

All transcription extracts were prepared from the *ncl-1(e1865)* mutant strain of *C. elegans* var. (Bristol) N2, grown on *E. coli*-chicken egg plates (Kennedy and Kenichi, personal communication). *C. elegans* eggs were prepared and stored as previously described (Honda et al. 1986). The preparation of the extracts was as previously described except, instead of breaking the *C. elegans* eggs using a French press cell, they were ground with a mortar and a pestle, while kept frozen in liquid nitrogen.

**Ion exchange chromatography of extracts on phosphocellulose columns**

Phosphocellulose (Whatman P-11) was prepared and used according to manufacturer's instructions. All chromatography steps were performed at 4°C in 200 ul columns, prepared in 1 ml syringes. *C. elegans* extracts were diluted in PC buffer (40 mM Tris pH 7.9, 1 mM EDTA, 20% glycerol, 1mM DTT, 1mM PMSF), containing 0.1 M KCl and 200 ul (10 mg protein/ml packed volume) was applied to the P-11 columns pre-equilibrated in 10 volumes of TGED buffer (50 mM Tris
pH 7.9, 25% glycerol, 0.1 mM EDTA, 1 mM DTT) containing 0.1 M KCl. The column was washed in 3 volumes of the same buffer and the bound proteins were eluted with 3 volumes of TGED containing 1.0 M KCl. Fraction size was 50% of bed volume and the peak fractions were pooled based on absorbance at 280 nm. The 1.0 M KCl step fraction was dialyzed on dialysis membranes (Millipore; 0.05 um) with TGED buffer to 0.1 M KCl. Fractions were stored at -80°C.

Preparation of exogenous *C. elegans* pol III

Partially purified *C. elegans* pol III was a gift from M. Golomb. This pol III containing fraction was subjected to further chromatography on a DEAE-Sephadex column to separate pol I activity (Sanford et al. 1985).

RNA polymerase assays

The activity of the exogenous pol III, the pol III in the *C. elegans* crude transcription extract and the pol III in fractions obtained from P-11 chromatography of the transcription extract were measured on poly[d(A-T)] templates at 0.08 M ammonium sulfate, as described previously (Sanford et al. 1983). The changes made in the assay include the use of 9 mM [³²P]UTP (400 Ci/mmol; supplied by Amersham), instead of [³H]UTP. Due to problems encountered in reproducible inhibition of pol II and pol III activity with α-amanitin, Tagetin (supplied by Epicentre Technologies) was used instead, as a specific inhibitor of pol III (Steinberg et al. 1990). With the specific batch of Tagetin that was used, at least 30 fold higher concentration (60 units per 25 ul assay) was found to be required for inhibition of *C. elegans* pol III, compared to the amount reported for the inhibition of pol III from other organisms. It is not clear, at this
point, whether this is as a result of the *Caenorhabditis elegans* pol III being different or a problem with the specific batch of the Tagetin. One unit of pol III activity represents the incorporation of 1 pmol of UMP in 20 min under standard assay conditions.

**Transcription assays**

Transcription assays were carried out in a 25 ul volume for 60 min. at 25°C, stopped by addition of SDS to 1%, treated with phenol:chloroform:isoamyl-alcohol (25:24:1) and electrophoresed on 10% non-denaturing polyacrylamide gels, as described previously (Honda et al. 1986). For transcription of tRNA genes, a template concentration of 10 nM and 2.5 ul of extract was used. For transcription of 5S rRNA genes, a template concentration of 2 nM and 5.0 ul of extract was used. The templates used in this study were as follows: 1 kb *Bam*H1 subclone of *Caenorhabditis elegans* 5S DNA (Nelson and Honda 1985) and a 3.1 kb *Eco*RI-*Hind*III subclone of SUP7 tRNA<sub>TRP</sub> (Kondo et al. 1988); *C. briggsae* 5S DNA *Hind*III subclones from 1.0 and 0.7 kb repeats (Nelson and Honda 1989); *Ascaris lumbricoides* 1 kb *Cla*I subclone of 5S DNA (Nilsen et al. 1989); *M. arenaria* 5S DNA from the 9 kb repeat (a 1 kb deletion subclone of the central 3 kb *Hind*III fragment generated for sequencing in the previous chapter; this fragment is missing most of the 3' rDNA spacer sequences); *X. borealis* somatic 800 bp *Eco*RI 5S DNA subclone (Peterson et al. 1980).
RESULTS

Sequence requirements for 5S rRNA gene transcription in the C. elegans extracts

Previously, it was shown that 5S DNA from the closely related nematode species C. briggsae, can be transcribed in the C. elegans extracts (Nelson and Honda 1989). However, the 5S DNA of the more distantly related nematode M. arenaria, like that of Xenopus, cannot be transcribed efficiently in the C. elegans extracts (Fig. 3-1). On the other hand, 5S DNA from another distantly related nematode, A. lumbricoides, is transcribed efficiently.

Based on earlier competition studies with the C. elegans 5S DNA template (Nelson et al. 1991), the ICR is required for sequestering limiting components of the extract. In Figure 3-2 A, a 50 bp region, corresponding to the A and C boxes and an intermediate (M) region (contact sites for TFIIIA in the ICR of Xenopus 5S DNA; Vrana et al. 1988), from the 5S DNA of C. elegans is compared to the same region in the templates used in Figure 3-1. The purpose of this comparison is to see to what extent these sequences have diverged and if it is possible to explain the differences in transcription of these templates in the C. elegans extract. The summary of the nucleotide differences is tabulated in Figure 3-2 B. Looking at the total number of differences, Ascaris 5S DNA ICR is most similar to that of C. elegans, followed by Meloidogyne and then Xenopus. The significance of the differences in each region will be discussed later.

Deletion studies with 5S DNA from both C. elegans and C. briggsae indicate that for efficient transcription in the C. elegans extracts, the 5S DNA 5'-flanking sequences are required in addition to the ICR (Nelson et al. 1991).
FIGURE 3-1

The efficiency of transcription of 5S DNA from other nematodes and *Xenopus* in the *C. elegans* cell-free extracts

Autoradiograph, showing the relative transcription of 5S DNA from *C. elegans*, *A. lumbricoides*, *M. arenaria* and *X. borealis* (Ce, Asc, M9 and Xbs1, respectively)
FIGURE 3-2
Comparison of nematode 5S DNA sequences in the region corresponding to the Xenopus internal control

A) The region corresponding to the ICR of the 5S rRNA gene from C. elegans (Ce) is compared to 5S DNA from M. arenaria (M9; from 9 kb repeat), A. lumbricoides (Asc) and X. borealis (Xbs1). Sequence identity is shown by "=" signs and deletions are indicated by a ".". The regions corresponding to the A and C boxes and the intermediate region (M), known to be contact sites for TFIIIA in Xenopus, are underlined (Vrana et al. 1988).

B) The number of nucleotide differences, compared to C. elegans, in each region of the ICR (as shown in A) is summarized in a table.
A

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9</td>
<td>==</td>
<td>==</td>
<td>G=</td>
<td>C==</td>
<td>T==</td>
</tr>
<tr>
<td>Ce</td>
<td>AAGTTAAGCAACGTTGAG...TCCAGTTAGTACTTGAGATCGGAGACGACGCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asc</td>
<td>==</td>
<td>==</td>
<td>G==</td>
<td>C=</td>
<td>T==</td>
</tr>
<tr>
<td>Xbsl</td>
<td>==</td>
<td>==</td>
<td>C==</td>
<td>G==</td>
<td>T==</td>
</tr>
</tbody>
</table>

NUCLEOTIDE DIFFERENCES

<table>
<thead>
<tr>
<th></th>
<th>Asc</th>
<th>M9</th>
<th>Xbsl</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>NUCLEOTIDE DIFFERENCES</th>
<th>Asc</th>
<th>M9</th>
<th>Xbsl</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(DELETION)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>OTHER</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(INSERTION FOR 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>
However, comparison of 5S DNA 5'-flanking regions is less straight forward and sequence similarities are less obvious, even between closely related species, making the precise delineation of possible regulatory sequences much more difficult. In Figure 3-3, the 5S DNA 5' flanking sequences from *C. elegans* and *C. briggsae* (where there are two 5S DNA repeats of 1 kb and 0.7 kb) are shown along with the same region from the templates used in Figure 3-1. A 13 bp box conserved upstream of *C. elegans* and *C. briggsae* 5S DNA from 1 kb repeats (although at different positions relative to the start; Nelson et al. 1991) is underlined. No similarity to this sequence is observed in any of the other templates, including the 5S DNA from the *C. briggsae* 0.7 kb repeat, which is transcribed efficiently in the *C. elegans* extracts. This 5S DNA instead has TATA-like sequences, which were shown by *in vitro* mutagenesis to be required for transcription (Nelson et al. 1991). These sequences, along with other TATA-like sequences in the *Ascaris* and *Meloidogyne* templates are also shown underlined in Figure 3-3. The *Xenopus* template has no similar sequences.

Transcription factor requirements for 5S rRNA gene transcription in the *C. elegans* extracts

Similar to other systems studied, the *C. elegans* 5S DNA ICR sequences seem to be important for interaction with the *C. elegans* transcription machinery. The *C. elegans* transcription extract was fractionated on phosphocellulose columns to see if there is a 5S-specific transcription factor, analogous to TFIIIa.

Figure 3-4 shows that the *C. elegans* extracts can be fractionated into a P-11 step fraction that is sufficient for tRNA gene transcription (lane 3) but the breakthrough fraction is also required for 5S rRNA gene transcription (lane 8, 10).
Comparison of nematode 5S DNA sequences upstream of the coding region

The region corresponding to the 5'-flanking sequences of 5S rRNA genes from *C. elegans* (Ce) is compared to the same regions in *C. briggsae* (Cb1, Cb0.7), *M. arenaria* (M9; from 9 kb repeat), *A. lumbricoides* (Asc) and *X. borealis* (Xbs1). The *Caenorhabditis* 13 bp boxes are underlined, as well as TATA-like sequences.

<table>
<thead>
<tr>
<th></th>
<th>-50</th>
<th>-40</th>
<th>-30</th>
<th>-20</th>
<th>-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xbs1</td>
<td>GC</td>
<td>GC</td>
<td>GC</td>
<td>GC</td>
<td>GC</td>
</tr>
<tr>
<td>M9</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
</tr>
<tr>
<td>Asc</td>
<td>AT</td>
<td>GT</td>
<td>GT</td>
<td>GT</td>
<td>GT</td>
</tr>
<tr>
<td>Ce</td>
<td>CG</td>
<td>GT</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
</tr>
<tr>
<td>Cb1</td>
<td>CC</td>
<td>CT</td>
<td>AG</td>
<td>TG</td>
<td>GA</td>
</tr>
<tr>
<td>Cb0.7</td>
<td>CA</td>
<td>GC</td>
<td>AG</td>
<td>GA</td>
<td>AA</td>
</tr>
</tbody>
</table>
There is also an increase in tRNA transcription when the breakthrough fraction is added to the step fraction (lane 5). This was shown not to be as a result of limiting amounts of pol III in the step fraction, by addition of exogenous pol III (lane 4).

DISCUSSION

The results presented in this chapter further support the requirement for 5'-flanking sequences for 5S rRNA gene transcription in the C. elegans extracts. Previously, this requirement was suggested based on transcription of deletion templates, lacking a 13 bp box conserved between 1 kb 5S DNA repeats of C. elegans and C. briggsae (Nelson et al. 1991). Similar results were obtained when the TATA box upstream of a second 0.7 kb 5S DNA repeat in C. briggsae was altered by in vitro mutagenesis.

In this chapter, the requirement for 5S DNA 5'-flanking sequences for transcription in the C. elegans extracts is further supported by the lack of transcription of templates with similar ICR sequences but no similarity in the 5'-flanking sequences.

While the Xenopus tRNA genes can be transcribed in the C. elegans extracts, the 5S DNA cannot be transcribed (Honda et al. 1986; see also Fig. 3-1). This suggests that there might be some problems associated with the interaction of the heterologous template with the C. elegans transcription machinery and that this is specific to 5S DNA transcription. The lack of interaction of the Xenopus 5S DNA with the C. elegans transcription machinery could be explained by the many differences observed in the entire ICR of 5S DNAs from C. elegans and Xenopus (Fig. 3-2).
FIGURE 3-4

A P-11 breakthrough fraction of *C. elegans* extract is required in addition to the step fraction for 5S rRNA transcription.

P-11 breakthrough and step fractions (see materials and methods) of the *C. elegans* transcription extract were tested in various combinations for *C. elegans* 5S rRNA (lanes 6-10) or tRNA (lanes 1-5) gene transcription. The volume of each fraction per reaction is indicated in ul. The units of pol III activity in each fraction was assayed and the total activity in each reaction is indicated. Lanes 4 and 9 include 7 units of exogenous pol III.
The *Ascaris* 5S DNA has the exact same differences as the *Xenopus* 5S DNA in the M region, but it is different from *C. elegans* 5S DNA by only one nucleotide in the A and C regions. Since the *Ascaris* 5S DNA is transcribed in the *C. elegans* extracts, those nucleotide differences are not likely to be critical.

The *M. arenaria* 5S DNA is also similar to that of *C. elegans*, with only three nucleotide differences in the more critical A and C regions. However, these differences involve deletions and insertions. Furthermore, the *M. arenaria* 5S DNA template cannot be transcribed efficiently in the *C. elegans* extracts. It is possible that the ICR differences no longer allow this template to interact with the *C. elegans* transcription machinery. Alternatively, this lack of interaction may be as a result of differences in the 5'-flanking sequences.

Based on what is known with the *Caenorhabditis* templates, it is difficult to see what allows the *Ascaris* 5S DNA to be transcribed in the *C. elegans* extracts. Although a TATTA sequence can be found upstream of *Ascaris* 5S DNA, it is not clear if it plays a role in transcription in the *C. elegans* extract, because the same element is present further upstream in the *Meloidogyne* template which is not transcribed efficiently. It may just be that other TATA-like sequences are involved in this process and should be delineated precisely.

Preliminary results from this chapter also indicate that the *C. elegans* transcription extract can be fractionated on phosphocellulose columns to give a step fraction sufficient for tRNA gene transcription but requiring reconstitution with the breakthrough fraction for 5S rRNA gene transcription. However, the breakthrough fraction is very crude and we were not able to show specific binding of any protein to the *C. elegans* 5S DNA sequences, using gel mobility-shift assays (data not shown).

Addition of the *C. elegans* extract phosphocellulose breakthrough fraction to the step also results in an increase in tRNA transcription. This was shown not
to be as a result of limiting pol III in the step fraction. This increase in tRNA transcription might be as a result of contamination of the breakthrough fraction with the general pol III transcription factors or the presence of a factor which cannot be split properly. Another possibility is the presence of an activity similar to the transcription stimulatory factor found to bind upstream of *Xenopus* 5S rRNA and tRNA genes, in an equivalent phosphocellulose fraction (Oei and Pieler 1990).
REFERENCES


Palmer JM, Folk WR (1990) Unraveling the complexities of transcription by RNA polymerase III. Trends Biochem 15:300-304


Shastry BS, Ng SY, Roeder RG (1982) Multiple factors involved in the transcription of class III genes in Xenopus laevis. J Biol Chem 257:12,979-12,986


GENERAL DISCUSSION
In the first two chapters of this thesis, 5S rRNA genes were characterized from the plant-parasitic nematode *Meloidogyne arenaria*. The purpose was to see if a conserved 13 bp block, required for *in vitro* transcription, upstream of 5S DNA from 1 kb repeats of the closely related nematodes *C. elegans* and *C. briggsae* (Nelson and Honda 1989), also existed in more distantly related nematodes.

Characterization of 5S DNA from *M. arenaria* demonstrated an unusual organization. The *M. arenaria* 5S DNA was found to be linked to the larger rRNA coding region, similar to the organization found in prokaryotes and some lower eukaryotes (reviewed in Gerbi 1985). Furthermore, functional and non-functional rDNA repeats of various size classes occur adjacent in the genome. In all higher eukaryotes studied to date, except for a marine copepod (Drouin et al. 1987), 5S rRNA coding regions occur in repeated units, unlinked to the larger rDNA.

The 5S DNA repeats of the nematodes *C. elegans* (Nelson and Honda 1985) and *A. lumbricoides* (Nilsen et al. 1989) also occur independent of the rDNA coding sequences. However, these nematode 5S DNA repeats do have another unusual feature. They also code for a short pol II-transcribed leader RNA which is trans-spliced to many protein coding genes (Krause and Hirsh 1987, Nilsen et al. 1989). The significance of this linkage, if any, is not known. There is also another case of a pol II gene linked to 5S DNA in the genome of the brine shrimp *Artemia*. Here, 5S DNA is linked to the histone multigene family (Andrews et al. 1987).

A comparison of the *Meloidogyne* and the *Caenorhabditis* life cycles may provide a clue as to how such different 5S rRNA gene organization may have evolved in the phylum *Nematoda*. The nematode genera to which the *Meloidogyne* and the *Caenorhabditis* species belong to are in the class *Secernentea* (Poinar 1983). However, the *Caenorhabditis* species belong to the
order *Rhabditida* which are mostly terrestrial microbotrophs, while the *Meloidogyne* species belong to the order *Tylenchida* which are stylet-bearing parasites of plants and invertebrates.

The *Meloidogyne* and *Caenorhabditis* species have very different life cycles and mode of reproduction. *C. elegans* is a free-living soil nematode, feeding mainly on bacteria (Wood 1988). It can easily be reared in the laboratory on petri plates. The *C. elegans* life cycle is only three days from egg to adult and the main mode of reproduction is by self-fertilizing hermaphrodites which can lay up to 300 eggs. These hermaphrodites can also be outcrossed to males.

The genome of *C. elegans* is also simple. It is composed of $8 \times 10^7$ base pairs, $80\%$ of which is single copy sequences. The haploid set of chromosomes is composed of 5 autosomes and one sex chromosome. These characteristics have led to extensive studies of genome organization and development.

The life cycle of *Meloidogyne*, however, is much more complex and it is much more difficult to obtain enough material for molecular or biochemical studies. *Meloidogyne* belongs to a genus of "root-knot" nematodes. They produce characteristic galls on the roots of many agriculturally important crops in both temperate and tropical climates (Dropkin 1989). In order to study these nematodes, a greenhouse of the appropriate plant host is required. It is also possible to grow these nematodes on tissue culture of plant roots.

The *Meloidogyne* life cycle starts with freshly hatched female juvenile nematodes which use their stylet to enter cells, usually in the region of root elongation. The nematodes migrate around in between cells and eventually become sedentary in an area of a developing side root. This is followed by transformation of the cells around the nematode head to specialized polyploid "giant cells", consisting of hundreds of nuclei. The nutrients normally sent to growing leaves, flowers, and fruits of the infected plants are now diverted to these
giant cells which are the source of nutrition for the nematodes throughout development. The biochemical interactions that go on in this process are not well understood. This might be the most likely stage where plant DNA might be taken up by the parasitic nematode, leading to horizontal gene transfer (see Discussion in Chapter two).

Within the *Meloidogyne* infected plant roots, the adult females eventually become sacs of eggs. The mode of reproduction is by meiotic or mitotic parthenogenesis. Intersex males may arise through unfavorable environmental conditions, but their role is not understood especially in the obligatory parthenogenetic species.

The cytogenetics available from *Meloidogyne* species, indicate variable ploidy levels throughout the genus and even within a species (Triantaphyllou 1985). One of these complex species is *Meloidogyne arenaria*. Different populations of this nematode may have chromosome numbers varying from $2n = 30$ to $2n = 56$. However the most common form is a triploid with greater than 50 chromosomes.

Therefore, the mode of reproduction and the variable ploidy levels, observed in *Meloidogyne* species, may have contributed to the evolution and fixation of their unusual rDNA repeats.

The functional significance of the linkage of 5S DNA to rDNA in a higher eukaryote is not clear. With the prokaryotic transcription operons, linkage of 5S DNA, rDNA and some ribosomal proteins allows equimolar expression of at least subsets of ribosomal components. However, in eukaryotes only the larger rRNA genes are transcribed in one long transcript which is then processed into 18S, 5.8S and 28S rRNA. The other ribosomal components, 5S rRNA and ribosomal proteins, are even transcribed by different polymerases; pol III and pol II, respectively.
In cases where 5S DNA is linked to the larger rRNA coding regions, in lower eukaryotes, it is not known if this aids in the mechanisms involved for coordinate regulation (reviewed in Gerbi 1985). In vivo studies with yeast have demonstrated that transcription of 5S DNA is at least independent of the rDNA enhancer region (Neigeborn and Warner 1990).

However, it is still possible for other cis-acting elements to direct the coordination of transcription of ribosomal components. This is certainly evident in studies done with N. crassa, where 5S DNA and rDNA are not linked. In vitro transcription of 5S rRNA genes in this fungus requires a -24 TATA box and an additional internal element (D box; +18 to +34) upstream of the A and C boxes (Tyler 1987). The transcription of larger rRNA genes in N. crassa requires a TATA box containing element in the immediate upstream regulatory region as well as another region further upstream (Tyler 1990). There is an 18 bp element within the latter region with homology to the N. crassa 5S DNA internal D box and an element found in triplicate upstream of two ribosomal protein genes (Tyler and Harrison 1990). This element has been termed the 'ribo box' and may play a role in coordination of transcription of ribosomal components.

Interestingly, there is an unrelated 13 bp element conserved upstream (although at different positions) of 5S DNA from 1 kb repeats in C. elegans and C. briggsae (Nelson et al. 1991). This sequence is also present upstream of the C. elegans larger rRNA genes (Ellis et al. 1986). In vitro transcription of C. elegans 5S DNA indicates that upstream sequences including this element are required for transcription. It is possible that the 13 bp element plays a similar role to the N. crassa ribo box, in coordination of transcription of ribosomal components. However this remains to be investigated by cloning of C. elegans ribosomal protein genes and development of in vitro transcription systems capable of transcribing pol I and pol II genes.
One purpose of this thesis was to see if the *C. elegans* 5S DNA upstream regulatory sequences were found in distantly related nematodes, such as *M. arenaria*. However, analysis of sequences upstream of possible functional 5S DNA from *M. arenaria* does not indicate any sequence similarity and this template is not transcribed efficiently in the *C. elegans* extracts (see Chapter Three). This incompatibility with the *C. elegans* extracts may also be as a result of differences in the ICR. A clearer picture can be obtained by construction of chimeric templates, containing the 5'-flanking sequences from *C. elegans* 5S DNA.

The 5S DNA from another distantly related nematode species, *A. lumbricoides*, is transcribed in the *C. elegans* extracts. There is no apparent sequence similarity between the 5'-flanking sequences from these nematodes. Therefore, this heterologous template could also be useful in elucidation of the type of 5'-flanking sequences required for 5S DNA transcription in the *C. elegans* extracts.

Finally, *C. elegans* extracts were fractionated on phosphocellulose columns to study the factors required for 5S rRNA transcription. Preliminary results indicate that a step fraction is sufficient for transcription of a tRNA gene, but requires reconstitution with the breakthrough fraction for 5S DNA transcription. However, since the breakthrough fraction also increases tRNA transcription, it is not clear at the moment how well the factors are separated, or if there are additional factors required for transcription of both genes in the breakthrough fraction. Such factors may be similar to the stimulatory activity found in an equivalent fraction of *Xenopus* extracts (Oei and Pieler 1990) or the pol III factor composed of RNA found in *Bombyx mori* (Young et al. 1991).

It would be of great interest to identify the regulatory proteins involved in the *C. elegans* 5S DNA transcription and define their interaction with the promoter regions. Attempts to study such interactions in the crude extract or
phophocellulose fractions failed to show specific DNA binding to internal or upstream sequences of the C. elegans 5S DNA, using gel mobility-shift assays (data not shown). Such studies should be possible with further purified fractions.
REFERENCES


Tyler BM (1990) Two complex regions, including a TATA sequence, are required for transcription by RNA polymerase I in *Neurospora crassa*. Nucl Acids Res 18:1805-1811


ABBREVIATIONS

A.: *Ascaris*

ASC: *Ascaris lumbricoides*

bp: basepairs

C.: *Caenorhabditis*

CB0.7: *Caenorhabditis briggsae* 0.7 kb 5S DNA

CB1: *Caenorhabditis briggsae* 1 kb 5S DNA

CE: *Caenorhabditis elegans*

D.: *Drosophila*

E.: *Escherichia*

ICR: 5S DNA internal control region

IGS: rDNA intergenic spacer

M.: *Meloidogyne*

Ma: *Meloidogyne arenaria*

nt: nucleotide

kb: kilobase

P-11: phosphocellulose ion exchange resin

pol: RNA polymerase

rDNA: genes encoding for larger (18S, 5.8S and 28S) ribosomal RNA

5S DNA: genes encoding 5S ribosomal RNA

X.: *Xenopus*

XBS1: *Xenopus borealis* somatic 5S DNA