# MOLECULAR ANALYSIS OF SMALL RNAS OF SACCHAROMYCES CEREVISIAE

A dissertation submitted by John Michael Xavier Hughes in fulfilment of the requirements for the degree of *Philosophiae Doctor* at the University of Leicester, England.

UMI Number: U010104

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U010104 Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author. Microform Edition © ProQuest LLC. All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC 789 East Eisenhower Parkway P.O. Box 1346 Ann Arbor, MI 48106-1346



# **Table of Contents**

Preface Abbreviations List of Figures and Tables Abstract	Page 5 6 8 1 2
Indroduction SCOPE OF THE WORK THE ORGANISM DIVERSITY OF SMALL RNAS 1. Abundant U snRNAs 2. Less abundant U snRNAs 3. Non-capped small RNAs	13 13 14
<ol> <li>4. RNase P</li> <li>5. Small RNAs involved in DNA replication</li> <li>6. Small RNAs in the regulation of transcription</li> <li>7. Small nuclear RNAs of unknown function</li> <li>8. Small RNAs related to <i>Alu</i>-family repetitive DNA sequen</li> <li>9. Signal recognition particle</li> <li>10. Small RNAs and control of translation</li> <li>11. Small RNA required for bacteriophage assembly</li> <li>12. Small RNAs in non- nucleic acid and protein metabolism</li> <li>13. Other small RNAs of unknown function</li> </ol>	ces
RNA DURING THE EARLY COURSE OF EVOLUTION	28
Chapter 1. Preliminary attempts to identify small RNAs and RNPs in Sa	nccharomyces
SUMMARY INTRODUCTION 1. Choice of cell disruption method 2. Yeast nuclei 3. SRP	32 33
4. hnRNP/snRNP RESULTS	36
<ol> <li>SDS/prenor permeabilization</li> <li>Salt precipitation of naked RNA</li> <li>Fractionation of RNPs: nuclear/cytoplasmic distribution</li> <li>Fractionation of RNPs: mobility shift compared to naked F gradients</li> <li>Fractionation of RNPs: sucrose gradients and attempts a purification</li> <li>Fractionation of RNPs: from the nucleus</li> </ol>	RNA on sucrose
	57
Chapter 2. Cloning of yeast small RNA genes and their use as probes to	o study sub-cellular
SUMMARY	58
INTRODUCTION Background to the pEMBL plasmid system	59
METHODS FOR CLONING YEAST SMALL RNA GENES 1. Genomic DNA preparation 2. Genomic library in AEMBL4	61
3. Purification of small RNAs for use as probes 4. Screening of the $\lambda$ EMBL4 yeast genomic library	
<ol> <li>Sub-cloning</li> <li>Determining the copy number of small RNA genes in the y</li> </ol>	east genome

7. Isolation of the second *SNR17* gene
 8. Determining the orientation of small RNA genes sub-cloned on plasmids

-

**...**'

<ul> <li>THE USE OF CLONED GENES AS PROBES TO STUDY THE DISTRIBUTION OF SMALL RNAS IN DIFFERENT FRACTIONS <ol> <li>Nuclear/cytoplasmic distribution</li> <li>SnR30, snR17 and scR1 are not polyadenylated</li> <li>SnR30 and snR17 are m<sub>3</sub>G-capped</li> <li>SnR17 is hydrogen bonded to pre-rRNA</li> <li>The cytoplasmic RNA scR1 is not mitochondrial DISCUSSION</li> </ol> </li> </ul>	87
<ol> <li>SnR30 and snR17 are snRNAs</li> <li>Like U3 snRNA, snR17 is associated with pre-rRNA</li> </ol>	
<ul> <li>Chapter 3. Analysis of the primary and secondary structure of snR17 and SUMMARY</li> <li>METHODS USED FOR SEQUENCING AND MAPPING SMALL RNA GENES <ol> <li>Introduction</li> <li>Methods for creating nested deletions</li> <li>Deleting between restriction sites</li> <li>Deleting with exonuclease Bal31</li> <li>Deletions from random DNase I cuts</li> </ol> </li> </ul>	9 0 9 1
<ul> <li>3. Sequencing strategies</li> <li>4. Mapping of the genes</li> <li>ANALYSIS OF PRIMARY AND SECONDARY STRUCTURE</li> <li>1. Structure of the <i>SNR17</i> genes</li> <li>2. The primary structure of snR17 is similar to U3 snRNA</li> <li>3. The secondary structure of snR17 and U3 snRNA</li> <li>DISCUSSION</li> <li>1. Transcriptional signals of yeast snRNA genes</li> <li>2. Structure of U3 snRNA and possible interactions with pressible</li> </ul>	104 114 e-rRNA
<ul> <li>Chapter 4. Disruption of small RNA genes.</li> <li>SUMMARY</li> <li>INTRODUCTION</li> <li>YEAST CELLS REQUIRE AT LEAST ONE SNR17GENE <ol> <li>Plasmid construction</li> <li>Disruption of SNR17A in vivo</li> <li>snr17a-Δ strains have increased levels of snR17B RNA</li> <li>Disruption of SNR17B in vivo</li> <li>Inheritance of the snr17a-Δ snr17b-Δ double deletion is DISCUSSION</li> <li>SnR17 is essential</li> <li>The two SNR17 genes are complementary</li> </ol> </li> </ul>	121 122 123 lethal 129
<ul> <li>Chapter 5. Yeast contains a family of snRNPs antigenically related to U3 SUMMARY INTRODUCTION RESULTS <ol> <li>Immunoprecipitation of yeast snRNPs with human antisera</li> <li>Attempt to define U3 snRNA box C as an RNP protein bindin DISCUSSION <ol> <li>Precipitation of yeast snRNPs by anti-(U3)RNP serum</li> <li>SnR17, but not U3 snRNP, is Sm-precipitable</li> <li>U3 snRNA box C as a protein binding site</li> </ol> </li> </ol></li></ul>	snRNP 143 144 147 a ng site 156

3

~

4

•

<ul> <li>Chapter 6. Materials and Methods</li> <li>Section 1. Yeast methods</li> <li>1.1 Culture of yeast cells</li> <li>1.2 Preparing spheroplasts</li> <li>1.3 Crude nuclei preparation</li> <li>1.4 Yeast genomic DNA maxiprep</li> <li>1.5 Yeast genomic DNA miniprep</li> <li>1.6 Extraction of yeast RNA</li> <li>1.7 Yeast transformation - Lithium acetate method</li> <li>1.8 Mating, sporulation and tetrad analysis</li> </ul>	174 175
<ul> <li>Section 2. Bacteriophage λ methods</li> <li>2.1 Measurement of λ phage titre</li> <li>2.2 Preparation of filters for plaque hybridization</li> <li>2.3 Isolation of single clones</li> <li>2.4 Phage preparation - plate lysate</li> <li>2.5 Phage DNA miniprep</li> <li>2.6 Phage DNA maxiprep</li> </ul>	181
<ul> <li>Section 3. pEMBL plasmid methods</li> <li>3.1 Blue/white colony distinction</li> <li>3.2 Plasmid DNA miniprep - boiling method</li> <li>3.3 Plasmid DNA miniprep - cleared lysate</li> <li>3.4 Plasmid DNA maxiprep - cleared lysate</li> </ul>	183
Section 4. DNA sequencing 4.1 Preparation of ssDNA from pEMBL plasmids 4.2 DNA sequencing by dideoxynucleotide chain termination	186
<ul> <li>Section 5. Labelling and hybridization</li> <li>5.1 Hybrid-enrichment of RNA</li> <li>5.2 3' end labelling of RNA with <sup>32</sup>pCp</li> <li>5.3 5' end labelling of oligonucleotides with [γ-<sup>32</sup>P]ATP</li> <li>5.4 Isolation of DNA fragments</li> <li>5.5 Labelling DNA by random oligonucleotide priming</li> <li>5.6 Assessing incorporation of label by TLC</li> <li>5.7 DNA and RNA blot hybridization</li> <li>5.8 T7 RNA polymerase <i>in vitro</i> transcription</li> </ul>	189
<ul> <li>Section 6. Immunoprecipitation, and materials and methods used in Chapter 5</li> <li>6.1 Antibodies</li> <li>6.2 Immunoprecipitation</li> <li>6.3 Construction of pSRBC plasmids</li> <li>6.4 Injection of T7 transcripts into <i>Xenopus</i> oocytes</li> </ul>	194
Section 7. BUFFERS, MEDIA and REAGENTS	197
Conclusions	200
Appendix 1. The structure of snR30 and its gene.	202
Appendix 2. The structure and possible functions of scR1. The structure of scR1 and its gene. $scr1-\Delta$ strains segregate $\rho^-$ cells at high frequency	205
DISCUSSION Reference list	210 219

#### PREFACE

This work was carried out between 1983 and 1988 at the European Molecular Biology Laboratory, Heidelberg, Germany, in the laboratory of Dr. Giovanni Cesareni, and has been presented for the degree of *Ph.D* at the University of Leicester, England.

The work represents part of an investigation that set out to identify new small RNA species in the yeast *Saccharomyces cerevisiae* and to define their functions. The work focusses on the molecular analysis of a small nuclear RNA, snR17, but also describes the identification of another small nuclear RNA, snR30, and a small cytoplasmic RNA, scR1, the genes for which were isolated and cloned simultaneously with those of snR17. The detailed analysis of the gene for snR30 was later taken over by Dr. Marc Bally, and that of the gene for scR1 by Dr. Franco Felici, who determined large parts of the nucleotide sequences and performed the deletion analyses of these genes. Some subsequent analysis of the structure of snR30 and scR1, and some analysis and discussion of the phenotype of yeast strains carrying the *scr1-\Delta* deletion has been presented in two appendices.

Some parts of the snR17 work were performed in colaboration with other people, where these people could offer material or expertise usefully applicable to the study of snR17. The contributions of others have been clearly indicated in the relevant figure legends and in the text. These were, namely, Cosima Baldari-Telford, who constructed the yeast genomic DNA library, David Tollervey, who studied associations of yeast snRNAs with pre-rRNA, Danielle Konings, who ran the computer programmes to generate nucleotide sequence alignments and secondary structure models for U3 snRNAs, and Iain Mattaj, who performed injections into, and immunoprecipitations from, *Xenopus* oocytes.

My warmest thanks are due to Gianni Cesareni for his infinite patience, expert guidance and enthusiastic encouragement.

I owe special thanks to Bill Brammar for permitting me to register as a student in the Biochemistry Department of the University of Leicester, and for his help and generous support.

I am grateful to many people for helpful discussion of ideas, but in particular to Danielle Konings and to lain Mattaj.

I wish also to express particular gratitude to the staff of the EMBL Szilard Library and of the EMBL photographic department for their excellent and invaluable services.

# Abbreviations

Α	adenosine
Amp	β-lactamase gene endowing ampicillin resistance in <i>E. coli</i>
ARS	veast autonomously replicating sequence
ATP	adenosine 5'-triphosphate
B-EtSH	R-mercantoethanol
hn	base pair(s) of DNA
DCV DCV	base pair(s) of Div
0	
CAMP	adenosine 3 :5 -cyclic monophosphate
CUNA	DNA complementary to RNA
CIP	calf intestinal alkaline phosphatase
Δ	deletion
dATP	5' deoxyadenosine triphosphate
dCTP	5' deoxycytidine triphosphate
ddNTP	dideoxynucleoside triphosphate
DEAE	diethylaminoethyl (cellulose)
DEPC	diethylpyrocarbonate
dGTP	5' deoxyguanosine trinhosnhate
DMSO	dimothylaulphovido
DNIA	
DNA	
UNase	deoxyribonuclease
<b>dNIP</b>	deoxynucleoside triphosphate
ds	double-stranded
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EthBr	ethidium bromide
EIOH .	ethanol
G	quanosine
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
hnRNP	heterogeneous nuclear ribonucleoprotein
HPLC	high-pressure liquid chromatography
IPTG	isonronyl.B.D.galactosido
	internal transoribed esquence (in rDNA)
110 kh	kilohooo poir(a) of DNA
	kiloDaltene
KD :	
LB	Luria broth
LEU2	yeast gene encoding $\beta$ -isopropylmalate dehydrogenase
m <sub>3</sub> G	2,2,7-trinethylguanosine
moi	multiplicity of infection
mRNA	messenger RNA
mtDNA	mitochondrial DNA
N ·	any nucleoside
nt	nucleotide(s)
0 n	over night
	acatata
	optical donsity
	open reading from
	Open reading frame
ori	DNA replication origin
p	plasmia
p Cp	cytidine 3,5-bisphosphate
PEG	polyethylene glycol
pfu	plaque forming units
PIPES	1,4-piperazine-diethanesulphonic acid
PMSF	phenylmethylsulphonyl fluoride
taa	precipitate
pre-RNA	precursor BNA
nei	nounds per square inch
P31	pourios por square mon

4

.

ρrho: genetic designation for yeast mitochondrial genomerDNADNA encoding ribosomal RNARNAribonucleic acidRNaseribonucleaseRNPribonucleoproteinrpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
rDNADNA encoding ribosomal RNARNAribonucleic acidRNAseribonucleaseRNPribonucleoproteinrpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
RNAribonucleic acidRNaseribonucleoproteinRNPribonucleoproteinrpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
RNaseribonucleaseRNPribonucleoproteinrpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
RNPribonucleoproteinrpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
rpmrevolutions per minuterRNAribosomal RNAr.t.room temperature (~20°C)Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
rRNA ribosomal RNA r.t. room temperature (~20°C) S sedimentation coefficient scRNA small cytoplasmic RNA SD Yeast synthetic dextrose selective medium SDS sodium dodecyl sulphate Sm antibodies or antigens of this serotype snRNA small nuclear RNA snRNP small nuclear nibonucleoprotein particle spnt supernatant ss single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
r.t. room temperature (~20°C) S sedimentation coefficient scRNA small cytoplasmic RNA SD Yeast synthetic dextrose selective medium SDS sodium dodecyl sulphate Sm antibodies or antigens of this serotype snRNA small nuclear RNA snRNP small nuclear nibonucleoprotein particle spnt supernatant ss single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
Ssedimentation coefficientscRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
Scenific futurescRNAsmall cytoplasmic RNASDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
SchurtSinal cytoprasmic http://SDYeast synthetic dextrose selective mediumSDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
SDSsodium dodecyl sulphateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
South antibodiesSuprateSmantibodies or antigens of this serotypesnRNAsmall nuclear RNAsnRNPsmall nuclear nibonucleoprotein particlespntsupernatantsssingle-strandedSSC(1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6TthymidineTBE(1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3TE10mM TrisHCl pH 8.0, 1mM EDTA
snRNA small nuclear RNA snRNP small nuclear nibonucleoprotein particle spnt supernatant ss single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
sind a sinal nuclear nix snRNP small nuclear nibonucleoprotein particle spnt supernatant ss single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
spint supernatant ss single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
SS single-stranded SSC (1 x) 0.15M sodium chloride, 0.015M sodium citrate, pH 7.6 T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
T thymidine TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
TBE (1 x) 89mM Tris-borate, 25mM EDTA, pH 8.3 TE 10mM TrisHCl pH 8.0, 1mM EDTA
TE 10mM TrisHCl pH 8.0, 1mM EDTA
inin-layer chromatography
I ris I ris(hydroxymethyl)aminomethane
IRNA Transfer RNA
ts temperature sensitive (mutation)
TIP thymidine tripnosphate
U uridine or units
UAS upstream activation sequence
UHA3 yeast gene coding for orotidine-5' phosphate decarboxylase
UIP uridine 5 triphosphate
UV ultraviolet light
v/v percentage expressed as volume/volume
vol volume
w/v percentage expressed as weight/volume
w/w percentage expressed as w/w
wt wild type
Xgal 5-bromo-4-chloro-3-indolyl-β-D-galactoside
Y pyrimidine
YPD yeast rich dextrose medium
YPG yeast rich glycerol medium
minute(s) or prime
second(s)
% percentage, expressed as w/v when applied to chemical reagents unless otherwise indicated
unless otherwise indicated.

7

-

# List of Figures and Tables

Figure	9	Page
Introd	luction	
i.1	Base-pairing interactions of yeast U1 snRNA with the 5' splice site of yeast actin pre-mRNA, and of U2 snRNA with the branch point of the intron.	17
i.2	U small nuclear RNPs involved in pre-mRNA splicing.	19
i.3	Base-pairing interaction between U7 snRNA and the conserved sequence of histone pre-mRNA required for 3' end maturation.	20
i.4	Endonucleolytic cleavage site of RNase P generating a mature tRNA 5' end.	2 1
Chapt	er 1	
1.1	Differential extraction of RNA by SDS/phenol permeabilization of yeast cells.	38
1.2	Fractionation of total yeast RNA by precipitation in 2M NaCl.	40
1.3	Comparison of RNA extracted from crude nuclear and cytoplasmic fractions prepared in high or low [salt] from pressure ruptured cells or lysed spheroplasts.	42
1.4	Comparison of RNA extracted from crude nuclear and cytoplasmic fractions prepared in high or low [salt] from pressure ruptured cells or lysed spheroplasts.	44
1.5	Mobility of cytoplasmic RNPs compared to naked RNA on sucrose gradients.	47
1.6	Mobility of cytoplasmic RNPs on sucrose gradients.	49
1.7	Native polyacrylamide gel electrophoresis of cytoplasmic sucrose gradient fractions.	- 51
1.8	RNPs released from nuclei by sonication.	53
1.9	Fractionation of yeast nuclei.	56

.

•

2.1	Purified RNA bands used as probes to screen a genomic DNA library.	63
2.2	Distinguishing small RNA gene clones from rDNA clones.	67
2.3	Strategy for sub-cloning small RNA genes.	69
2.4	SnR30 is encoded by a unique gene.	71
2.5	ScR1 is encoded by a unique gene.	72
2.6	SnR17 is encoded by two genes.	73
2.7	Determining the orientation of small RNA genes.	75
2.8	Distribution of small RNAs in nuclear, cytoplasmic, polyadenylated and non-polyadenylated fractions.	78
2.9	Immunoprecipitation of small RNAs with anti-m <sub>3</sub> G cap antibodies.	80
2.10	Several yeast small nuclear RNAs, including snR17, are hydrogen bonded to pre-rRNA.	82
2.11	The cytoplasmic RNA scR1 does not copurify with mitochondria.	86
Chapt	er 3	
3.1	SNR17A plasmid and phage clones.	94
3.2	SNR17B plamsid clones.	95
3.5	Restriction maps and sequencing strategies for SNR17A and SNR17B.	97
3.7	Mapping of snR17A by S1 nuclease-protection.	101
3.8	Mapping the 5' end of snR17A by primer extension.	103
3.11	Nucleotide sequences of SNR17A and SNR17B.	105
3.12	Alignment of U3 snRNA sequences with snR17A and snR17B.	107
3.13	Secondary structure models for U3 snRNA and snR17.	111
3.19	Comparison of the 5' flanking sequences of SNR genes.	118
3.20	Precursor rRNA processing pathway.	119
3.21	Hypothetical base-pairing interactions between U3 snRNA and pre-rRNA.	120

....

....

.

.

-

Chapter 4

4.1	Disruption of SNR17A.	134
4.2	The lack of snR17A.	135
4.3	The absence of snR17A in <i>snr17a::URA3</i> strains confirmed by hybridization.	137
4.4	Levels of snR17B are increased in snr17a- $\Delta$ strains.	138
4.5	Disruption of SNR17B.	139
4.6	Construction of a doubly heterozygous, <i>snr17a/SNR17B snr17b/SNR17B</i> , diploid strain.	140
<b>4.7</b>	Possible segregation patterns of <i>snr17a</i> and <i>snr17b</i> alleles among tetrads from a doubly heterozygous diploid strain.	127
4.8	Tetrad analysis of <i>snr17a</i> and <i>snr17b</i> deletion strains.	141
4.9	Probing for snR17A and snR17B in <i>snr17a</i> and <i>snr17b</i> strains.	142
Table	List of yeast strains.	131
Table	II Gene deletion analysis of <i>Saccharomyces cerevisiae</i> snRNAs.	132
Chapt	er 5	
5.1	RNA immunoprecipitated from yeast nuclear extracts, end-labelled.	161
5.1 5.2	RNA immunoprecipitated from yeast nuclear extracts, end-labelled. RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.	161 163
5.1 5.2 5.3	RNA immunoprecipitated from yeast nuclear extracts, end-labelled. RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization. Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.	161 163 165
5.1 5.2 5.3 5.4	<ul> <li>RNA immunoprecipitated from yeast nuclear extracts, end-labelled.</li> <li>RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.</li> <li>Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.</li> <li>Hypothetical RNA protein binding sites.</li> </ul>	161 163 165 166
5.1 5.2 5.3 5.4 5.5	<ul> <li>RNA immunoprecipitated from yeast nuclear extracts, end-labelled.</li> <li>RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.</li> <li>Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.</li> <li>Hypothetical RNA protein binding sites.</li> <li>Sequences of T7 transcripts injected into <i>Xenopus</i> oocytes.</li> </ul>	161 163 165 166 167
5.1 5.2 5.3 5.4 5.5 5.6	<ul> <li>RNA immunoprecipitated from yeast nuclear extracts, end-labelled.</li> <li>RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.</li> <li>Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.</li> <li>Hypothetical RNA protein binding sites.</li> <li>Sequences of T7 transcripts injected into <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation of T7 transcripts from <i>Xenopus</i> oocytes.</li> </ul>	161 163 165 166 167 169
5.1 5.2 5.3 5.4 5.5 5.6 5.7	<ul> <li>RNA immunoprecipitated from yeast nuclear extracts, end-labelled.</li> <li>RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.</li> <li>Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.</li> <li>Hypothetical RNA protein binding sites.</li> <li>Sequences of T7 transcripts injected into <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation of T7 transcripts from <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation of T7 transcripts from <i>Xenopus</i> oocytes.</li> </ul>	161 163 165 166 167 169 171
<ol> <li>5.1</li> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>5.5</li> <li>5.6</li> <li>5.7</li> <li>5.8</li> </ol>	<ul> <li>RNA immunoprecipitated from yeast nuclear extracts, end-labelled.</li> <li>RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization.</li> <li>Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP antiserum.</li> <li>Hypothetical RNA protein binding sites.</li> <li>Sequences of T7 transcripts injected into <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation of T7 transcripts from <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation of T7 transcripts from <i>Xenopus</i> oocytes.</li> <li>Immunoprecipitation from yeast and <i>Xenopus</i> whole cell extracts of native U3 snRNP.</li> </ul>	161 163 165 166 167 169 171

~

# Appendix 1

.

A1.1 Nucleotide sequence of the *SNR30* gene.

A1.2	Secondary structure conformation around sequences in snR30 matching the Sm-binding site consensus and the conserved box C of U3 snRNAs.	204
Appen	dix 2	
A2.1	Nucleotide sequence of the SCR1 gene.	214
A2.2	RNA polymerase III internal promoter elements in scR1 and 7SL RNA genes.	216
A2.3	Secondary structure model for scR1 compared to established 7SL RNA structures.	218

#### Abstract

RNA has many diverse functions in living organisms, from serving as genome for many viruses, to regulating DNA replication, transcription, translation and other metabolic processes. Some RNA, like protein, has been shown to have catalytic activity. The great proportion of the mass of RNA in living cells, in the form of ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA), constitutes the machinery of protein synthesis, the remainder (approximately 2%) consists of many heterogeneous RNA species of relatively small size, loosely termed "small RNAs", the functions of many of which are completely unknown.

In an attempt to understand some of these functions, three hitherto undescribed small RNAs of the budding yeast *Saccharomyces cerevisiae* were identified and their genes were cloned. These three small RNAs, which lack polyadenylation at their 3' ends, appear to represent the three most abundant RNA species in this organism after rRNA and tRNA. The most abundant of the three was found to be mainly cytoplasmic and was therefore called "small cytoplasmic RNA 1" (scR1). The other two RNAs, named snR17 and snR30, were found to be enriched in nuclear fractions and to possess trimethyl guanosine cap structures at their 5' ends, identifying them as belonging to the ubiquitous class of "U" small nuclear RNAs (U snRNAs), of which several are required for the endonucleolytic cleavage and splicing reactions in the maturation pathways of nuclear precursor mRNAs (pre-mRNA).

Whereas scR1 and snR30 are both encoded by single genes, snR17 is the only yeast small RNA found so far to be encoded by two genes. SnR17 was found to be essential: haploid yeast strains lacking intact copies of one or other of the genes appeared to grow normally, but strains lacking both genes were inviable.

The nucleotide sequences of the snR17 genes were determined, and the primary and predicted secondary structures of the RNA, 328 nucleotides in length, were found to show significant similarities to those of U3 snRNA, an abundant U snRNA, the function of which is not known. SnR17 belongs to a family of *S. cerevisiae* snRNAs which, unlike those involved in pre-mRNA splicing, are located in the nucleolus hydrogen-bonded to pre-rRNA, and are associated with antigenic protein that is recognised by human antibodies specific for a 36 kD polypeptide of the U3 small nuclear ribonucleoprotein (U3 snRNP) in mammals. U3 snRNA is also nucleolar and associated with pre-rRNA. Given their structural similarities, snR17 and U3 snRNA are presumably homologous. Yeast snRNAs associated with the anti-(U3)RNP antigen share with U3 snRNAs a conserved nucleotide sequence element. This sequence element alone, however, when injected into *Xenopus* oocytes, was not sufficient to direct binding of the antigen. The association of snRNAs with pre-rRNA suggests that they function in ribosomal biogenesis.

### Introduction

#### SCOPE OF THE WORK

Ribonucleic acid (RNA) constitutes between 4% and 20% of the dry mass of living cells. Many distinct classes of RNA exist, all of diverse size and abundance. The major classes, those that together make up 98% of the total mass of cellular RNA, consisting of ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA), are almost exclusively involved in mediating the expression of genetic information through protein synthesis. Other classes, accounting for the remaining 2% (Hellung-Larsen 1977), have been collectively termed "small RNAs" because, at first, very few of them were found to be more than 300 nucleotides in length. Small RNAs have diverse functions, some of which have been well characterized, while many remain completely unknown.

This report describes an investigation that set out to define the nature and function of small RNAs in the bakers' yeast Saccharomyces cerevisiae. An initial foray into the world of yeast small RNAs, in which the size, abundance, intracellular distribution and possible methods of purification for different species were investigated, is described in Chapter 1. Chapter 2 describes the cloning of the genes for what appear to be the three most abundant yeast small RNAs, and how these genes were used as probes to study the intra-cellular distribution of the RNAs by hybridization. The detailed structures of the genes for one of the RNAs, the abundant small nuclear RNA snR17, and its predicted secondary structure, are presented in Chapter 3. Chapter 4 describes how, as a first step towards studying the function of snR17, each of its genes was specifically deleted, and how these deletions affected the phenotypes of haploid yeast strains inheriting them. Finally, Chapter 5 shows how some yeast small RNAs, includung snR17, form complexes with evolutionarily conserved proteins which are recognised by specific antibodies, and describes some attempts to determine the specific RNA structural requirements for protein binding.

#### THE ORGANISM

Saccharomyces cerevisiae, a species of the fungal division ascomycota, is a eukaryotic, unicellular organism that undergoes a budding form of cell division. It exhibits haploid and diploid phases of vegetative growth: two haploid cells of opposite mating type, **a** and  $\alpha$ , will conjugate to form a diploid, and a diploid cell,

on induction, will undergo meiosis forming four haploid ascospores contained within a sac or ascus. In nature, the diploid phase is dominant because haploid strains of a single mating type undergo spontaneous mating type switching, giving rise to populations of mixed mating type, which then mate and form diploids. Most laboratory haploid strains are stabilized as either **a** or  $\alpha$  by a mutation that abolishes spontaneous mating type switching; such strains are said to be "heterothallic". The ease with which genetic manipulations could be performed with yeast in the laboratory led to the widespead use of S. cerevisiae in the study of Mendelian-inheritance, and consequently the genetics of this organism is relatively well understood. Interest in yeast as an experimental organism first arose from its role in the production of alcohol during the fermenting of wine and beer: the discovery that alcoholic fermentation could be carried out by a cellfree extract of yeast originally led to the discovery of enzymes and coenzymes, and the mapping of the glycolytic pathway (reviewed by Roman 1981). The study of mitochondrial genetics was initiated in S. cerevisiae owing to the fact that this yeast, by preference, gains its energy by alcoholic fermentation and can tolerate mutations affecting mitochondrial function (Dujon 1981). In the first half of this century, S. cerevisiae was the organism from which RNA was first extracted and its chemical composition determined (RNA constitutes more than 98% of yeast nucleic acid, Cryer et al. 1975); RNA was first known as "yeast pentose nucleic acid", as opposed to DNA which was known as "thymus pentose nucleic acid, because it was first characterized from the thymus (reviewed by Hellung-Larsen 1977, Letham and Stewart 1977). S. cerevisiae was one of the first eukaryotic organisms in which techniques were developed for transformation, that is, the process by which naked DNA can be introduced into a cell resulting in a heritable change (see Chapter 4); yeast thus became unique as a biochemically and genetically very well characterized eukaryotic organism in which the effects of controlled genetic modifications carried out in vitro could be studied in vivo.

#### **DIVERSITY OF SMALL RNAS**

Small RNAs can be classified variously according to function, mode of action, intracellular distribution, metabolic stability, mode of transcription, distinguishing structural features, or the classes of specific antigenic proteins with which they associate. Small RNAs of eukaryotic, prokaryotic and viral origin have been identified; many occur naturally in the form of RNA combined with protein, as ribonucleoprotein particles (RNPs); some appear to be ubiquitous, while some appear to be organism- or tissue-specific. The known functions of small RNAs or RNPs include roles in nucleic acid metabolism, such

as control of DNA replication, control of transcription, and maturation of RNA; roles in the control of protein synthesis; and diverse functions including bacteriophage packaging, polysaccharide metabolism and the translocation of protein into membranes. The extent of knowledge of the RNAs involved in these functions, and of those RNAs of unknown function, is outlined in the following pages. [The field of small RNAs and RNPs has been reviewed for prokaryotes by Inouye and Delihas (1988) and for eukaryotes by Busch *et al.* (1982), Steitz *et al.* (1982), Reddy and Busch (1983, 1988), Dreyfuss *et al.* (1988a).]

#### 1. Abundant U snRNAs

Historically, the first small RNAs to be identified were the most abundant species of mammalian cells; these could be classified as being cytoplasmic, nuclear or nucleolar (Peacock and Dingman 1967, Hodnett and Busch 1968, Nakamura et al. 1968, Weinberg and Penman 1968, Zieve and Penman 1976). The abundant small nuclear RNAs were found to be rich in uridine and consequently became known as U snRNAs. Six U snRNAs were readily identified (numbered U1 to U6) and characterized as follows (reviewed by Busch et al. 1982): all were metabolically stable, with half-lives in the order of one cellcycle (Weinberg and Penman 1968); each species was represented by between 10<sup>5</sup> and 10<sup>6</sup> molecules per cell. U1 and U2 being the most abundant; all were subject to posttranscriptional modifications in the form of hypermethylation of nucleotides, especially in the 5' portions of the RNAs, and the addition of a 5' terminal cap, which was found to consist of a 2,2,7-trimethyl guanosine pyrophosphate structure (Saponara and Enger 1969), since found to be unique to U snRNAs (U6 is an exception in that it has a non-nucleotide cap of unknown structure); all were located in the nucleoplasmic fraction associated with heterogeneous nuclear ribonucleoprotein (hnRNP), which contains precursor mRNA (pre-mRNA), except U3, which was found in the nucleolus associated with precursor rRNA (pre-rRNA) (Prestayko et al. 1970).

The observed association with precursor RNAs gave rise to the hypothesis that U snRNAs were involved in RNA processing. All RNA species, including U snRNAs, undergo a series of processing events whereby initial transcripts are converted to mature, functional forms. These reactions include endonucleolytic cleavage events resulting in the removal of terminal sequences and excision of intervening sequences (IVS), ligation of cleaved products (splicing), polyadenylation, exonuclease trimming of 5' and 3' ends, and base modification (reviewed for mRNA by Dreyfuss *et al.* 1988b, for rRNA by Hadjiolov 1985, Gerbi 1985, for tRNA in Söll *et al.* 1980, and for U snRNA by Dahlberg and Lund 1988). For eukaryotic, nuclear-encoded transcripts, the site of processing

appears to be in an intracellular compartment separate from that in which the mature RNA performs its function: rRNA, tRNA and mRNA, destined for active roles in the cytoplasm, are thus processed in the nucleus, while U snRNAs, destined for active roles in the nucleus, shortly after transcription appear transiently in the cytoplasm (Zieve and Penman 1976, Zieve *et al.* 1988) where their 3' ends are matured, they are assembled into ribonucleoprotein particles (specific complexes of protein and RNA) and acquire 5' trimethyl guanosine (m<sub>3</sub>G) cap structures, before being imported back into the nucleus (De Robertis *et al.* 1982, Mattaj and De Robertis 1985, Mattaj 1986, reviewed by Mattaj 1988).

Patients with various forms of rheumatic autoimmune diseases have circulating antibodies directed against a variety of nuclear components (reviewed by Pinnas et al. 1973, Tan 1982). Lerner and Steitz (1979) discovered that certain antibodies in sera of patients with systemic lupus erythematosus reacted specifically with ribonucleoprotein particles containing U snRNAs; such ribonucleoprotein particles became known as U snRNPs, and the class of antibodies reacting with them were called "Sm" antibodies after the initials of the patient from whom they were first characterized. SnRNPs containing U1, U2, U4, U5 and U6 snRNAs were Sm-precipitable, but not the nucleolar U3 snRNP. U1, U2 and U5 were precipitable as individual snRNPs, while U4 and U6 were observed to be part of the same particle (Bringmann et al. 1984, Hashimoto and Steitz 1984). Various types of antibodies with specificities for different subsets of U snRNPs have since been instrumental in determining the functions, the conservation, and the structures of the abundant U snRNPs (reviewed by Steitz et al. 1988, Lührmann 1988) as well as providing means by which new, related, but less abundant snRNPs could be isolated (Lerner et al. 1981, Reddy et al. 1983, Reddy et al. 1985, Krämer 1987, see also Chapter 5, Introduction). The results of such studies have shown that the abundant U snRNPs are ubiquitous in eukaryotic organisms, and that their structures and functions are conserved (Guthrie and Patterson 1988).

Owing to its association with pre-rRNA, U3 snRNP was the first snRNP for which a role in RNA processing was proposed (Prestayko *et al.* 1970). However, despite confirmation of this initial observation (Calvet and Pederson 1981, Epstein *et al.* 1984) and several models having been proposed suggesting a function in mediating endonucleolytic cleavage of pre-rRNA through complementary base-pairing interactions (Bachellerie *et al.* 1983, Crouch *et al.* 1983, Tague and Gerbi 1984, Parker and Steitz 1987, see also Chapter 3, Discussion), the actual function of U3 snRNP still remains to be demonstrated. In the meantime, the five abundant nucleoplasmic U snRNPs have been shown

convincingly to be components required in the splicing reaction of nuclear premRNA. Towards the end of the 1970s it became clear that the genes for many eukaryotic proteins were split by intervening sequences, or introns; regions of protein-encoding sequence divided by introns were called exons. Introns were found to be transcribed into pre-mRNA, but then excised in a reaction which resulted in exon sequences being joined together to form one uninterrupted open reading frame (reviewed by Sharp 1985, Padgett et al. 1986, Green 1986, Maniatis and Reed 1987, Sharp 1987, Sharp et al. 1987). Once nucleotide sequences became available, it was observed that the 5' end of U1 snRNA was complementary to a conserved sequence at the 5' ends of introns, at exon-intron junctions (5' splice site). This observation formed the basis for a model in which U1 aligned the upstream (5') exon for correct ligation with the downstream (3') exon (Lerner et al. 1980a, Rogers and Wall 1980). Although the exact details of this model were not completely correct, it was subsequently shown that U1 snRNP bound to the 5' exon-intron boundary (Mount et al. 1983), that the 5' end of U1 was required for splicing (Krämer et al. 1984), and that maintenance of at least some complementary base-pairing between the 5' end of U1 snRNA and the exon-intron boundary was important, but not the sole requirement, for determining the efficiency of use of the 5' splice site (Zhuang and Weiner 1986, Zhuang et al. 1987, Siliciano and Guthrie 1988). The interaction between yeast U1 (snR19) and the 5' splice site is shown in Figure i.1.





A similar base-pairing interaction has been proposed between U2 snRNA and a portion of the intron approximately 30 nucleotides upstream of the 3' splice site. A conserved adenosine nucleotide at this position initiates the splicing reaction by mediating a transesterification at the 5' splice site, resulting in cleavage of the 3'-5' phosphodiester linkage at the 5' exon-intron boundary, and formation of a new 2'-5' phosphodiester bond between the conserved adenosine and the 5' end of the intron, thus forming a loop structure (known as a "lariat") in which the adenosine forms the branch point. The splicing reaction is then completed when the conserved guanosine at the 3' end of the cleaved 5' exon

mediates a second transesterification at the 3' intron-exon boundary, resulting in ligation of the two exon sequences and release of the branched intron (reviewed by Padgett *et al.* 1986, Green 1986, Maniatis and Reed 1987, Sharp 1987, Sharp *et al.* 1987). The sequence at the branch site of yeast nuclear pre-mRNA introns is well conserved, and a sequence in the yeast U2 snRNA, named LSR1 or snR20 (Ares 1986, Riedel *et al* 1987), is complementary to it (see Figure i.1); when point mutations, which abolished splicing, in the branch point sequence of the yeast actin pre-mRNA intron were complemented by mutations in the U2 snRNA sequence predicted to restore base-pairing, splicing of the gene was restored (Parker *et al.* 1987). Mammalian pre-mRNA intron branch point sequences and U2 snRNA do not show such good base-pairing complementarity, but nuclease protection studies indicate that U2 snRNP does bind to this region (Black *et al.* 1985).

The ability to reconstitute the pre-mRNA splicing reaction in vitro, selectively inactivate specific components, and then restore splicing activity by adding back the components in a purified form, has provided evidence that U1, U2 and U4/6 snRNPs are all required for splicing. These snRNPs, along with U5 snRNP, associate with pre-mRNP to form a splicing complex, called a "spliceosome", with a sedimentation coefficient of 60S in mammalian cells (Frendewey and Keller 1985, Grabowski et al. 1985), and 40S in yeast (Brody and Abelson 1985). The identification in yeast of splicing complexes containing snRNAs (Pikielny and Rosbash 1986, Pikielny et al. 1986, Cheng and Abelson 1987) and the subsequent demonstration that these snRNAs were all precipitable as snRNPs by human Sm antibodies (Tollervey and Mattaj 1987, Riedel et al. 1987), led to the identification of the yeast U snRNA homologues snR19 = U1, snR20 or LSR1 = U2, snR14 = U4, snR7 = U5, and snR6 = U6. Nuclear premRNA splicing in yeast was shown to proceed by the same mechanisms as mammalian pre-mRNA splicing (Brody and Abelson 1985, Pikielny and Rosbash 1985, Lin et al. 1985, Parker and Guthrie 1985), and the yeast U snRNAs were found to be required for splicing in vitro, and were shown to be essential for cell viability (Kretzner et al. 1987, Siliciano et al. 1987b, Ares 1986, Siliciano et al. 1987a, Patterson and Guthrie 1987, Brow and Guthrie 1988).

The pattern of formation of the splicing complex, resolved by studies using differential centrifugation, native gel electrophoresis and chromatographic techniques on *in vitro*-formed complexes (Zillmann *et al.* 1987, Bindereif and Green 1987, and reviewed by Maniatis and Reed 1987, Sharp 1987, and Sharp *et al.* 1987, Dreyfuss *et al.* 1988a), is summarized in Figure i.2. The premRNA is initially associated with various hnRNP proteins, whose possible roles in the splicing reaction are not understood. Association of U1 snRNP with the 5'

splice site, an interaction that appears to be weak under some conditions, and of U2 with the branch site occurs early in the order of spliceosome assembly. This and the subsequent stage of assembly requires ATP, although the splicing reaction itself does not. At least two specific proteins are involved in spliceosome assembly: the "U2 snRNP assembly factor" (Ruskin et al. 1988) and the "intron-binding protein" (Tazi et al. 1986, Gerke and Steitz 1986). ATP is possibly required for U4/6 and U5 snRNPs to form a compound particle, which then associates with the pre-mRNA to form a mature spliceosome through an interaction of U5 snRNP with the 3' splice site. The splicing reaction proceeds in two stages: first, the cleavage at the 5' splice site results in two resolvable intermediates, the 5' exon and the branched intron, the latter still covalently joined to the 3' exon; second, the two exons become ligated to form mature mRNA and the intron is released as a branched lariat, which remains associated with the U2, U5 and U6 snRNPs. U4 and probably U1 snRNPs appear to be liberated during the reaction, possibly already before the cleavage event at the 5' slice site.





#### 2. Less abundant U snRNAs

The availability of anti-snRNP antibodies, in particular, those reacting specifically with the  $m_3G$  cap structure, has enabled the isolation of a number of less abundant U snRNPs (the cap structure, rather than the nucleotide composition, is now considered to define a U snRNA, Reddy and Busch 1988).

These have been named U7 to U11 in order of their discovery. They are less than one tenth as abundant as the major U snRNPs, and they can all be precipitated with Sm antibodies, including U8, which like U3 is nucleolar (Reddy *et al.* 1985). The functions of U8, U9 and U10 are unknown.

U7 snRNP was identified by its ability to restore normal 3' end-processing to sea urchin histone pre-mRNA which had been injected into *Xenopus* oocytes (Galli *et al.* 1983, Strub *et al.* 1984, Strub and Birnstiel 1986, reviewed by Birnstiel and Schaufele 1988). Histone pre-mRNA, unlike other pre-mRNAs, do not become polyadenylated at their 3' ends, but are matured by an endonucleolytic cleavage reaction, which requires the presence of a conserved nucleotide sequence downstream of the cleavage site. The reaction has been shown to depend on the conservation of base-pairing between this site and a complementary sequence at the 5' end of U7 snRNA (Schaufele *et al.* 1986). It has been suggested that the availability of a thermolabile factor, which is required for 3' end cleavage in conjunction with U7 snRNP, could regulate the levels of mature histone mRNA throughout the cell cycle (Birnstiel and Schaufele 1988).





U11 snRNP might play an analogous role to U7 in the 3' end cleavage and polyadenylation of non-histone pre-mRNA: cell extracts that support non-histone pre-mRNA cleavage and polyadenylation *in vitro* are sensitive to RNase treatment (Hashimoto and Steitz 1986) and U11 snRNP has been isolated from pre-mRNP-polyadenylation complexes (Krämer 1987).

Many m<sub>3</sub>G-capped snRNAs have been detected in fungi, for which no homologues have been found yet in vertebrates (Riedel *et al.* 1986, Tollervey and Mattaj 1987, Zagorski *et al.* 1988). The total number of snRNAs identified in *S. cerevisiae* to date, including those described in this work, is at least 27 (Riedel *et al.* 1986, Brow and Guthrie 1988, Zagorski *et al.* 1988). Several of these have been found in nucleolar fractions associated with pre-rRNA (Tollervey 1987a, Zagorski *et al.* 1988, see Chapter 2). All the yeast snRNAs, including those involved in pre-mRNA splicing, are much less abundant than the mammalian species; most are represented by between 50 and 500 molecules per cell (Riedel *et al.* 1986, 1987).

#### 3. Non-capped small RNAs

All the U snRNAs, except U6, possess m<sub>3</sub>G cap structures and, like mRNAs, are transcribed by RNA polymerase II (reviewed by Dahlberg and Lund 1988). Other eukaryotic small RNAs (including U6 snRNA, Kunkel *et al.* 1986) are transcribed by RNA polymerase III, and possess 5' triphosphate ends. All RNA polymerase III precursor transcripts (5S rRNA and tRNA included), can be precipitated in the form of RNP by "La" autoimmune antibodies (Lerner *et al.* 1981, Hendrick et al. 1981, Rinke and Steitz 1985). The La antigen binds to the characteristic poly(U) ends of RNA polymerase III transcripts and possibly acts as a transcription termination factor (Gottlieb and Steitz 1987).

#### 4. RNase P

RNase P is an endoribonuclease that cleaves pre-tRNA generating mature tRNA 5' ends (see Figure i.4). The enzyme is an RNP particle, characterized in eubacteria to consist of a ~400 nt RNA and a protein component (reviewed by Pace *et al.* 1987, Altman *et al.* 1988). The RNA, the secondary structure of which is conserved (James *et al.* 1988), is the catalytically active component



**Figure i.4** Endonucleolytic cleavage site of RNase P generating a mature tRNA 5' end.

of the enzyme: it has been shown to perform the cleavage in the absence of protein *in vitro* (Guerrier-Takada *et al.* 1983). The protein appears to function as a screen against anionic repulsion between the RNA-enzyme and substrate (Reich *et al.* 1988). RNase P appears to be ubiquitous, occurring in the nuclei of eukaryotic organisms, and in mitochondria, [RNase P in *S. cerevisiae* mitochondria consists of a nuclear-encoded protein and a 9S (~450 nt), mitochondrial-encoded RNA (Hollingsworth and Martin 1986)].

#### 5. Small RNPs involved in DNA replication

Three distinct RNPs have recently been found to play roles in DNA replication in eukaryotes, two of which are required for priming mitochondrial DNA (mtDNA) replication, while the third is required for nuclear telomere replication. "Mitochondrial RNA processing" (MRP) RNase cleaves RNA primers of mtDNA replication at the site of transition from RNA to DNA synthesis, and requires for its activity a 135 nt RNA species (Chang and Clayton 1987). The DNA primase responsible for initiating mtDNA replication also requires RNA for its function; this RNA appears to be the cytoplasmic 5.8S rRNA (Wong and Clayton 1986). Both the RNA components of these RNPs are nuclear-encoded, and are the first examples of RNAs found to be transported into mitochondria. A telomere terminal transferase ("telomerase") enzyme, which contains small RNAs essential for its function, has been identified in Tetrahymena (Greider and Blackburn 1987). The telomerase adds specific nucleotides to the 3' ends of telomeres (the ends of linear chromosomes), in a template-independent manner, in order to replace telomeric sequences that may have been lost after a cycle of DNA replication.

Some species of tRNA have functions other than being adaptors for mRNA translation. In *E. coli*, a minor tRNA<sup>Arg</sup> was found to be the product of the *dna*Y gene, which is essential for DNA replication (Garcia *et al.* 1986). In chick cells, tRNA<sup>Trp</sup> acts as a cofactor for avian RNA tumour virus reverse transcriptase by priming DNA synthesis on the viral RNA-genome template. The reverse transcriptase, and a region of the RNA template, have specific affinity for the tRNA<sup>Trp</sup> (Panet *et al.* 1975).

In *E. coli*, the replication of ColE1 plasmids is regulated by a 108 nt plasmid-encoded RNA called "RNA I". Plasmid DNA replication is initiated from a 550 nt RNA primer ("RNA II") which is first transcribed as a precursor with an extended 3' end. RNA I is encoded by the complementary DNA strand at the 5' end of the primer coding sequence, and is transcribed in the opposite direction. Maturation of the primer by RNase H, and consequently DNA replication, is inhibited by hybridization of RNA I to the primer, an interaction that is accelerated by a plasmid-encoded protein called "rop" (reviewed by Cesareni and Banner 1985).

#### 6. Small RNAs in the regulation of transcription

The only examples of small RNAs regulating transcription have been found in bacteria. The nucleotide cyclic AMP (cAMP) regulates many metabolic processes in both eukaryotes and prokaryotes. In *E. coli*, cAMP complexed with its receptor protein "CRP" binds to DNA, acting on promoters to regulate the expression of operons. CRP is encoded by the *crp* gene, the transcription of which is inhibited by the cAMP/CRP complex by a mechanism involving a small RNA: binding of cAMP/CRP to a region close to the *crp* promoter stimulates transcription of a small RNA, which is initiated three base-pairs upstream, and in the opposite direction from the *crp* mRNA; a base-pairing interaction between the 5' portions of the two RNAs leads to the formation of a structure that causes termination of further *crp* mRNA transcription (Okamoto and Freundlich 1986).

Both charged (aminoacyl) tRNAs and uncharged tRNAs inhibit transcription by competing with DNA for binding to *E. coli* RNA polymerase. fMet-tRNA<sup>Met</sup>, but not Met-tRNA<sup>Met</sup>, stimulates the affinity of RNA polymerase for specific gene promoters (Nomura *et al.* 1986).

#### 7. Small nuclear RNAs of unknown function

The eukaryotic 7SK RNA was one of the early, relatively abundant small RNAs to be identified (Zieve and Penman 1976), but it has not been assigned a function; it is predominantly nuclear and transcribed by RNA polymerase III (Murphy *et al.* 1984, Murphy *et al.* 1987). Several small RNAs in the 7-8S size range (~300 nt) have been described previously (Reddy *et al.* 1981): of three nucleolar species, 7-1 and 8S were found to be pre-5.8S rRNAs, but 7-2 RNA was found to be a distinct transcript with a 5'-pppG terminus. 7-2 RNA and cytoplasmic 8-2 RNA, also of unknown function, both occur as antigenically related RNPs which are recognised by "To" antibodies from scleroderma patients' sera (Reddy *et al.* 1983).

A new class of small RNAs, apparently conserved in primary structure in all eukaryotes from human to yeast, has been discovered recently (Trinh-Rohlik and Maxwell 1988). 4.5S hybRNAs are 90 to 130 nt in length and contain a highly conserved oligonucleotide sequence complementary to a conserved sequence on all 18S (small ribosomal subunit) rRNAs, to which they appear to hybridize *in vivo*. Over 90% of the RNA has been estimated to be nuclear.

#### 8. Small RNAs related to Alu-family repetitive DNA sequences

Alu-repetitive DNA sequences (the repeat units are cut by Alul restriction endonuclease) make up 4%-5% of the human genome and are interspersed throughout the genome, often within the introns of genes encoding proteins, and as such, can be detected as unstable RNA polymerase II transcripts. Alusequences contain RNA polymerase III promoters within their repeated units, and a number of stable RNA polymerase III transcripts originating from these repetitive sequences have been found. 7SL RNA, of the signal recognition particle (see below), is a polymerase III transcript derived from one class of *Alu*-DNA (Ullu *et al.* 1982b, Li *et al.* 1982); retroposition of truncated 7SL RNA sequences is thought to have given rise to other classes of *Alu*-DNA (Ullu and Tschudi 1984). In rodents, three classes of stable small RNAs, known as 6S (183 nt), 4.5S and 4.5SI (90-100 nt) (Reddy 1985), are transcribed by RNA polymerase III from *Alu*-related sequences. A 4.5S RNA species with similarity to 4.5SI has been found also in the broad bean (Kiss *et al.* 1987). These RNAs were observed to be associated with polyadenylated mRNA in both nuclear and cytoplasmic cell fractions (Reddy and Busch 1983). A 7S RNA transcribed from interspersed repetitive sequences has been detected specifically in chick embryonic muscle (Khandekar *et al.* 1984), and an RNA of less than 160 nt is specifically transcribed in rat neural tissue (Sutcliffe *et al.* 1984). The roles of these RNAs are not understood, but they have been implicated in tissue- or developmental-specific gene expression.

N

#### 9. Signal recognition particle

The signal recognition particle (SRP) is an 11S RNP composed of six polypeptides complexed with one RNA molecule, 7SL RNA, which is stable and abundant. SRP is cytoplasmic and its function has been defined in vitro, using mammalian cell extracts, as mediating the translocation of membrane and secretory proteins into the endoplasmic reticulum ("ER") (reviewed by Walter and Lingappa 1986). SRP has two separable activities that can be distinguished in vitro: translation retardation and promotion of translocation. Translocation is promoted through binding of SRP to the newly synthesised signal-sequence of a nascent polypeptide chain as it emerges from the ribosome; translation is then thought to be retarded or arrested until the signal-sequence comes into contact with the signal-sequence receptor in the ER membrane, from which point translocation proceeds cotranslationally. SRP thus acts as an adaptor between ribosome and ER, via an interaction with its own receptor, the "docking protein". also in the ER. More recent work has suggested that cotranslational translocation is not obligatory, and that, in vivo, in some circumstances, posttranslational translocation does occur (see Chapter 4, Discussion). SRP consists of two distinct domains, one containing that portion of the 7SL RNA homologous to Alurepetitive DNA, the "Alu-domain", and the other containing that portion of 7SL RNA which is absent from most Alu sequences, the "S domain". The S domain is responsible for the signal recognition and translocation functions, while the Aludomain causes translational retardation/arrest (Siegel and Walter 1985, Siegel and Walter 1986, Andrews et al. 1987, Siegel and Walter 1988). The role of SRP in protein translocation has been demonstrated only for mammals, although

7SL-like RNAs have been detected in animal (Ullu *et al.* 1982a, Li *et al.* 1982, Balmain *et al.* 1982, Ullu and Tschudi 1984, Gundelfinger *et al.* 1984), fungal (Ribes *et al.* 1988, Brennwald *et al.* 1988, Poritz *et al.* 1988), plant (H. L. Sänger, in preparation) and even prokaryotic cells (Moritz and Goebel 1985). The secondary structure is well conserved among these RNAs, but the nucleotide sequence is not. Despite deliberate efforts, no 7SL-like RNA nor SRP particle has been detected in *S. cerevisiae* (Poritz *et al.* 1988).

No supportive evidence has been obtained for the hypothesis that an 11S RNP in *E. coli*, containing a 6S RNA (184 nt), has an analogous function to eukaryotic SRP. The RNA is stable and represented by 1000 molecules per cell. Gene disruption showed that it was neither required for normal cell growth nor for protein secretion (Lee *et al.* 1985); its function remains unknown.

### 10. Small RNAs and control of translation

A stable, 4.5S (114 nt) RNA is essential for the viability of Escherichia coli. Without it, E. coli ribosomes cease to be competent to initiate protein synthesis but remain able to elongate polypeptide chains. Ribosomal biosynthesis was found not to be affected in cells partially depleted of the 4.5S RNA, the ratio of ribosomal subunits to assembled ribosomes was normal, and the ribosomes and subunits retained their normal sedimentation values. Ribosomes from depleted cells did not function in a "wild type" cell-free translation system, unless they were first salt-washed, in which case their ability to initiate translation was restored (Bourgaize and Fournier 1987). This RNA appears to be conserved at least among eubacteria: a 271 nt RNA with significant primary and secondary structural similarities has been found in Bacillus subtilis (Struck et al. 1988a). The most interesting structural feature of both these RNAs is the single, long hairpin that contains conserved nucleotides: a hairpin containing most of the same conserved nucleotides, in the same conformation of loops and bulges, is reproduced in all the 7SL-like RNAs described (Struck et al. 1988b); this portion of the 7SL-like RNAs represents the most conserved structural feature of the molecule and corresponds to the lower branch in the "T" of the "S domain" (Chapter 3, Figure 3.18). The significance of this conserved structure is not clear.

Several examples of translational inhibition of specific mRNAs by "antisense" RNA, that is, RNA encoded by the complementary strand of a gene sequence and transcribed in the opposite direction, have been found in prokaryotes. Antisense RNA blocks translation of the mRNA by complementary hybridization, often with the region containing the ribosomal binding site sequence (Shine and Dalgarno 1975a & b) and the AUG initiation codon. Proteins whose expression is regulated in this way have been found to be involved in plasmid replication control, conjugal transfer of plasmids, transposition, and in regulating bacteriophage repressor activity (reviewed by Inouye and Delihas 1988). Antisense RNAs vary in size from 70 to 180 nt. Translation of the mRNA from the *E. coli Omp*F gene, encoding an outer membrane protein, is inhibited by complementary hybridization with a 4.5S RNA (93 nt) resulting in shielding of the Shine-Dalgarno sequence and AUG codon; in this case, the 4.5S RNA is encoded by the *mic*F gene at a different location on the chromosome (Andersen *et al.* 1987).

A number of eukaryotic small cytoplasmic RNAs (scRNAs) appear to regulate translation of cytoplasmic mRNA pools. Those that have been characterized are heterogeneous, originating from a variety of different organisms and tissues, but in all cases they can be isolated from cytoplasmic fractions as scRNPs, and both the scRNP and the naked scRNA inhibit initiation of translation in cell-free translation systems (Sarkar et al. 1981, Winkler et al. 1983, McCarthy et al. 1983, Piot et al. 1984, Lorberboum et al. 1986). ScRNPs that inhibit translation of specific mRNAs, referred to as "translational control" RNAs (tcRNA), have been distinguished from scRNPs that appear to inhibit translation non-specifically, referred to as "inhibitory" RNAs (iRNA). Several scRNPs have been found associated with cytoplasmic pools of mRNP, which have been found to be untranslatable in cell-free systems, whereas the same mRNA in its naked, protein-free state could be translated (McCarthy et al. 1983, Piot et al. 1984, Lorberboum et al. 1986). Such an iRNP, containing an RNA only 35 nt long, could be salt-washed from human placental cytoplasmic mRNP, which could then be translated. This iRNP exerted its inhibitory effect through an interaction with a factor required for translation initiation, and not by direct blocking of the mRNA (Lorberbourn et al. 1986). A 6S iRNP, containing an RNA of ~85 nt, was found to be associated in a  $Mg^{2+}$ -dependent manner with mRNP of the shrimp Artemia salina, preventing translation (Piot et al. 1984). McCarthy et al. (1983) discovered a tcRNP, containing a 102 nt RNA specifically associated with myosin heavy chain mRNP of chick skeletal muscle, specifically inhibiting translation of this mRNA but not of other endogenous mRNAs. A free cytoplasmic 10S iRNP from chick embryonic muscle, containing a 4 S RNA (70-90 nt), appeared to inhibit translation by preventing initiationcomplex formation (Sarkar et al. 1981, Winkler et al. 1983).

- 26

#### 11. Small RNA required for bacteriophage assembly

A small (~120 nt) RNA, encoded by the dsDNA genome of the *Bacillus* subtilis  $\phi$ 29 bacteriophage, has been shown *in vitro* to be required for packaging bacteriophage particles; the RNA does not become packaged in the virion (Guo *et al.* 1987).

## 12, Small RNAs in non- nucleic acid and protein metabolism

Branch formation is an integral part of glycogen and starch synthesis, resulting in a high concentration of chain-terminal substrate molecules which can be broken down rapidly in the event of an energy requirement. In rabbit skeletal muscle, the branch-forming enzyme, amylase isomerase  $(1,4-\alpha-D-glucan: 1,4-\alpha-D-glucan-6-\alpha-[1,4-\alpha-glucano]$  transferase), is a small RNP (Petrova and Belykh 1982) containing a 2.5S (31 nt) RNA that appears to be the catalytic component (Shvedova *et al.* 1987, Petrova and Belykh 1986). This small RNA contains ten modified nucleotides, consists of 40% G-residues, and probably folds into a single hairpin structure. When assayed *in vitro*, the purified, naked RNA was able to recognise the  $\alpha$ -1,4-glucan substrate and transfer it to the C<sub>6</sub> position of a glucose residue on the polysaccharide chain, forming a (branch)  $\alpha$ -1,6-linkage.

It has long been recognised that tRNAs play multiple, complex regulatory roles in the cell as well as acting as adaptors for mRNA translation (Panet *et al.* 1975, Letham and Wettenhall 1977, Garcia *et al.* 1986, Nomura *et al.* 1986, Yanofsky 1988). But in bacteria tRNAs also mediate in the transfer of amino acids for the synthesis of aminoacyl phospholipids and of peptidoglycans (reviewed by Soffer 1974).

#### 13. Other small RNPs of unknown function

"Ro" small cytoplasmic RNPs, so named because they are recognised by "Ro" antibodies from patients with autoimmune disease (reviewed by Keene *et al.* 1987), appear to be associated with the cytoskeleton of mammalian cells (Bachmann *et al.* 1986). They contain "Y" RNAs, which are RNA polymerase III transcripts, and of which five species have been detected in human cells ("hY1hY5", Wolin and Steitz 1983), two in mouse cells ("mY1 and mY2", Hendrick *et al.* 1981) and three in rat ("rY1a, rY1b and rY2", Reddy *et al.* 1983). In humans, they occur in varying abundance in different tissues (Wolin and Steitz 1983). Two classes of RNPs have been identified by their distinct morphology viewed by electron microscopy. One class, isolated from rat liver microsomal fractions, represents the largest RNP particles yet discovered; the particle size (35 nm x 65 nm) is greater than that of ribosomes (Kedersha and Rome 1986). The appearance of these particles put the authors in mind of vaulted gothic arches, therefore they have been called "vaults". A single species of RNA, of ~140 nt, has been found within vaults, but up to nine RNA molecules could be contained within each particle.

The second class is a family of 19S RNP particles that are distinguished by their hollow, cylinder-like morphology (10-14 nm in diameter x 16-19 nm in length) and consequently known as 19S cylinder particles (reviewed by Kloetzel 1987). They appear to be ubiquitous among eukaryotes, they all appear to contain a similar complement of proteins and in most cases, but not all, have been shown to contain a 70-90 nt RNA. There are reports of cylinder particles in both nuclear and cytoplasmic fractions, and in association with a number of cellular activities, such as translational inhibition in association with polyribosomes (Martins de Sa *et al.* 1986), RNase P activity (Castaño *et al.* 1986).

#### RNA DURING THE EARLY COURSE OF EVOLUTION

Several lines of evidence indicate that RNA preceded DNA and protein in the course of evolution, and that the first living entities were probably RNA-like macromolecules with the ability to catalyse their own reproduction (reviewed by Orgel 1987, Cavalier-Smith 1987). RNA molecules combine the properties of being able to fold into complex tertiary structures, allowing them to recognise specific substrates for catalysis in an enzyme-like manner, with the potential to act as template for their own replication by a complementary base-pairing process. The evolution of more complex self-replicating entities would have followed the development of a varied repertoire of catalytic functions, eventually of sufficient complexity to support the evolution of protein synthesis, and of such complexity that it became advantageous to distinguish RNA with catalytic function from RNA with genomic function (Weiner and Maizels 1987), the former with a requirement for specific tertiary structure and the latter with a requirement for stable primary structure, which was eventually better provided by the more chemically stable properties of DNA (Darnell and Doolittle 1986, Lazcano et al. 1988). According to this view, the vestiges of this early stage in evolution when RNA performed all the catalytic functions required for pre-cellular life are represented in present-day organisms by the diversity of functions in which RNA appears to perform a role.

The strongest evidence in support of this view of the early course of evolution comes from the discovery that RNA does indeed have catalytic activity (Cech et al. 1981, Guerrier-Takada et al. 1983, Van der Veen et al. 1987, Hutchins et al. 1986, Shvedova et al. 1987). Thus, in the complete absence of protein in vitro, pre-mRNA and pre-rRNA transcripts containing either group I or group II intervening sequences (Michel and Dujon 1983) undergo an autocatalytic splicing reaction in which structural features of the intervening sequences play catalytic roles (Cech and Bass 1986, Michel and Jacquier 1987); the RNA component of RNase P can cleave pre-tRNA (Pace et al. 1987); plant pathogenic RNAs undergo a distinct type of self-cleavage reaction (Forster et al. 1987); and the 31 nt RNA component of amylase isomerase can catalyse glycosidic branching (Shvedova et al. 1987). An additional variety of reactions can be shown to be catalysed by artificially modified RNAs (Cech and Bass 1986, Sampson et al. 1987), extending the list of structure-specific reactions performed by RNA to include RNA endonuclease, nucleotidyl transferase, phosphotransferase, phosphatase, kinase, ligase and glucanotransferase activities. Moreover, it is likely that the peptidyl transferase activity of ribosomes is a function of rRNA (Noller 1984, Nomura 1987). Given the variety of known reactions performed by RNA, it is not inconceivable that RNA could also function as a template-dependent RNA polymerase (Cech 1986).

The competition for possibly rare nucleotide substrates among early selfreplicating RNAs would have selected for the ability to synthesise substrates from simpler compounds, leading to the evolution of primitive metabolism; at this stage, a variety of catalytic abilities would have been advantageous. The problems of concentrating and storing substrates and products might have eventually provided the selection pressure leading to cellularization (Cavalier-Smith 1987), but initially, concentration and retention of substrate molecules could have been achieved by covalent binding; aminoacyl tRNA synthetases, which catalyse the linking of amino acids to the 3' ends of tRNAs, might have evolved initially for this purpose, and primitive ancestors of ribosomes could have provided the catalytic sites for an ordered sequence of biosynthetic events, involving accurate and even sequential positioning of substrates by means of their tRNA-like adaptors (T. J. Gibson and A. I. Lamond, in preparation).

tRNAs might have originated as molecular "tags" at the 3' ends of genomic RNA, serving as a structural feature signalling the site at which templatedependent RNA synthesis should begin (Weiner and Maizels 1987). Several ssRNA viral genomes have tRNA-like structures at their 3' ends which are recognised by modern enzymes of tRNA metabolism such as tRNA terminal transferase, which adds CCA to the 3' ends of tRNAs, and RNase P (Haenni et al. 1982). Weiner and Maizels (1987) have pointed out the interesting similarity between the activities of tRNA terminal transferase, adding CCA to tRNAs, and of telomerase, which adds telomere sequences to the 3' ends of eukaryotic chromosomes [yeast telomeres consist of repeated units of C1-3A (Walmsley 1987), and telomerase, furthermore, is a small RNP (Greider and Blackburn 1987)]. An RNase P-like enzyme, in removing tRNA-like structures, might have liberated non-genomic RNA strands for various catalytic functions, possibly after a series of further RNA-processing steps. Weiner and Maizels (1987) further point out the similarity between the RNA replicase of bacteriophage QB and ribosomes: both recognise tRNAs or tRNA-like structures, both the processes of protein synthesis and QB replication are initiated with the binding of a guanosine nucleotide, and the protein subunits of QB replicase consist of two host-encoded protein elongation factors and one ribosomal protein; these similarities possibly reflect the early common ancestor of both these enzymes, namely a primitive RNA replicase.

If ribosome-like complexes evolved from an original RNA replicase in response to the demand for ordered, possibly templated, metabolic pathways using tRNA-like molecules as adaptors between substrate and template, then the step to protein synthesis seems less improbable. The selection pressure favouring the synthesis of, at first, simple polypeptide chains could have come from the advantage gained by being attached to, or penetrating lipid membranes with hydrophobic polypeptides, a function that RNA alone, being hydrophilic, could not have performed (Cavalier-Smith 1987, T. J. Gibson and A. I. Lamond, in preparation). Thus one of the earliest functions of proteins could have been as membrane anchors or pores, suggesting that the modern mechanism of cotranslational membrane translocation involving SRP (a small RNP) has a very early origin.

Once template-dependent protein synthesis had evolved, the greater potential variety of conformations of polar and non-polar groups would have made proteins more adaptable molecules for performing very specific catalytic functions; proteins would have gradually taken over many of these roles from RNA, first in the form of RNP, and later dispensing with RNA altogether. However, the ubiquitous role of ribonucleotides in all aspects of modern cellular metabolism, as nucleoside triphosphates, as carriers of electrons and reducing potential, as cytokinins, as cAMP and enzyme co-factors (Cavalier-Smith

1987), suggests that many of the essential mechanisms of catalysis still resemble those of the original processes performed by RNA.

Given the presumed advantages of protein- over RNA-catalysis, the question of why RNA has been retained in modern RNPs remains unanswered: for reactions involving other nucleic acids where specificity can be defined partially by complementary base-pairing interactions, RNA would appear to be admirably suitable, but RNA possibly has additional, more generally advantageous traits for catalysis that remain obscure. It is possible therefore that the essential catalytic roles of all RNPs are performed by RNA components, and that RNP proteins have merely "supportive" functions.

ŝ

#### Chapter 1

# Preliminary attempts to identify small RNAs and RNPs in *Saccharomyces cerevisiae*

#### SUMMARY

Methods for the preparation of RNA from yeast cells were tried. Total RNA could be extracted by permeabilization of whole yeast cells with TE/SDS/phenol at 65°C. When total yeast RNA was fractionated on a denaturing polyacrylamide gel, many bands could be seen by ethidium bromide staining in the region between 5.8S and 18S ribosomal RNA (rRNA). Some of these RNAs could be fractionated by differential permeabilization of yeast cells with TE/SDS/phenol at lower temperatures, or by precipitation of total yeast RNA in the presence of 2M sodium chloride.

Crude nuclear and cytoplasmic fractions could be prepared from yeast cells ruptured by passage through a French press, or by hypotonic lysis of spheroplasts after prior enzymatic removal of the yeast cell wall. Of two relatively abundant RNAs seen in the 5.8S to 18S rRNA size range, one of them appeared to be enriched in cytoplasmic fractions and was named "small cytoplasmic RNA 1" (scR1), and the other appeared to be enriched in nuclear fractions and was named "small nuclear RNA 17" (snR17). Both were observed to have sedimentation velocities in sucrose gradients consistent with being small ribonucleoprotein particles (RNPs). There was some suggestion that snR17 exists in two forms, the first a small RNP either present in the cytoplasm or readily released from nuclei on cell disruption, and the second, tightly bound to the nuclear fraction even after sonication. Like the 7SL RNA component of the signal recognition particle (SRP) in HeLa cells, scR1 appears to be the most abundant small cytoplasmic RNA in *Saccharomyces cerevisiae*.

#### INTRODUCTION

At the beginning of this investigation, no small RNAs from *S. cerevisiae* had been described. Attempts to detect U snRNAs or 7SL RNA by cross-hybridization with specific probes, or by immunoprecipitation with antibodies to human snRNPs had failed (Lerner *et al.* 1980b). It was therefore necessary to make an experimental foray into the world of yeast RNA to find out, first, what the complement of yeast small RNAs looked like, and secondly, what their characteristics were in terms of RNP particle behaviour during intracellular fractionation, and how they could be isolated.

Two methods for the fractionation of RNA were tested: that of SDS/phenol permeabilization of whole cells (Rubin 1975), and that of differential salt precipitation of naked RNA.

Various approaches to intracellular fractionation were tried in order to identify small RNAs associated with proteins in the form of ribonucleoprotein particles (RNPs) and to study their intra-cellular distribution. In chosing the methods, consideration was given to how nuclei and cytoplasm could be separated, and how RNPs have been purified from mammalian cells, especially U small nuclear RNPs (snRNPs) and signal recognition particle (SRP). Continuous sucrose density gradient centrifugation was found to be a powerful method for resolving different RNP components, and attempts were made to purify fractions further from these by DEAE ion exchange chromatography and native gel electrophoresis.

1. Choice of cell disruption method.

Yeast cells possess a cell wall which can be removed enzymatically by digestion with a mixture of  $\beta$ -1,3-glucanases (to form "spheroplasts"), or which can be ruptured mechanically either on release from a pressure cell, or by vortexing with very fine glass beads. Both lysis of spheroplasts in hypotonic buffer, and passage through a French pressure cell have yielded preparations of yeast nuclei that were transcriptionally active and morphologically intact (Wintersberger *et. al.* 1973). Spheroplasting has the advantage of being the gentler method, possibly giving a greater yield of more intact nuclei, whereas French pressing may allow faster processing of more material, avoiding prolonged incubation at 30°C (see Duffus 1975 for a discussion of the relative merits).

#### 2. Yeast nuclei.

Yeast nuclei can be prepared after hypotonic lysis of spheroplasts in slightly acidic buffers (pH6.5) containing a high molecular weight polymer, such as polyvinylpyrrolidone or Ficoll, and divalent cations, all of which stabilize nuclei. Ficoll has been preferred for transcriptional studies of nuclei because polyvinylpyrrolidone inhibits RNA polymerases (Bhargava and Halvorson 1971, Duffus 1975, Schulz 1978). After low speed centrifugation to remove unlysed cells, crude nuclei can be pelletted at 30 000 x g for 25<sup>-</sup>. Mitochondria, which are not abundant in glucose-grown yeast owing to catabolite repression (Linnane and Lukins 1975), vacuoles and other membrane components float in the Ficoll supernatant (Wiemken 1975). Ribosomes remain mainly with the cytoplasmic components in the supernatant, but electron micrographs show that many of them copurify with the nuclei in association with the nuclear envelope (Duffus 1975).

#### 3. SRP

Signal recognition particle acts as an adaptor between ribosomes and endoplasmic reticulum, allowing simultaneous translation and translocation of proteins. It recognizes signal sequences of proteins targetted for translocation (Walter and Lingappa 1986). Six polypeptides are bound up with a single RNA approximately 300 nt long; the free RNA has a sedimentation coefficient of 7S, and is referred to as 7SL RNA, and the particle sediments at 11S. SRP can be washed from rough microsomes (endoplasmic reticulum which binds ribosomes) with buffers of high salt concentration (0.5M KOAc). Rough microsomes stripped of SRP are unable to translocate proteins in cell-free translation systems, and the activity of SRP is defined *in vitro* as stimulating and being essential for translocation when added back to such extracts. Mg<sup>2+</sup> ions are required to maintain structural integrity of SRP, and its activity is stabilized by very low concentrations of non-ionic detergents (eg: 0.002% Triton X-100) and by reducing agents (Walter and Blobel 1983, 1980).

#### 4. hnRNP/snRNP.

SnRNPs have been purified from hnRNP fractions of mammalian cell nuclei (Brunel *et al.* 1981, Liautard *et al.* 1982). Heterogeneous nuclear RNA is precursor messenger RNA which is being cleaved, trimmed, spliced, polyadenylated and otherwise modified (Dreyfuss 1986, Dreyfuss *et al.* 1988b) before being passed into the cytoplasm in a mature form to be translated on ribosomes. These processing steps are mediated by proteins and snRNPs which,
complexed with the RNA, make hnRNP. HnRNP fractions of nuclei can be separated from chromatin and nucleoli. Purification of hnRNP has been achieved by sonicating purified nuclei in low salt buffer containing  $Mg^{2+}$ , spinning the sonicate into a cushion of 30% sucrose at low speed giving a pellet enriched in nucleoli, and loading the material that does not enter the cushion onto a three-step sucrose gradient, which retains the hnRNP fraction in a band of 45% sucrose after spinning at 130 000 x g for 90' (Kish and Pederson 1975, Blanchard *et al.* 1977).

# 1. SDS/phenol permeabilization of whole cells.

The SDS/phenol permeabilization method of Rubin (1975) [Chapter 6, Section 1.6] was tried as a means of purifying total cellular RNA, and of differentially extracting small RNA species. The extent of RNA extraction depends on the harshness of the treatment, and is largely temperature-dependent; smaller RNAs, tRNA and 5S rRNA being readily extractable at room temperature, 5.8S rRNA at intermediate temperatures, and practically all species at 65°C. Figures 1.1A, odd numbered lanes, and 1.1B, lanes 1 and 2, show the RNAs extracted from cells under different conditions, and Figures 1.1A, even numbered lanes, and 1.1B, lanes 3, show the residual RNA remaining in the ruptured cells. [In Figure 1.1B, each of fractions 1, 2 and 3 was prepared in duplicate, and two different quantities of RNA from each sample were loaded on the gel.] Residual RNA was extracted from the ruptured cell pellet, which sediments to the bottom of the tube on separating the aqueous and phenolic phases by centrifugation, by resuspending in TE/0.5% SDS/phenol and heating at 65°C for 30' with intermittent vortexing. To see whether any RNA still remained in the twice extracted cells, a portion of the re-extracted cell pellet was redispersed in TE/SDS with gel loading buffer, heated at 95°C for 5' and loaded on an agarose gel with ethidium bromide. Only a minimal amount of RNA was seen to remain (data not shown). The efficiency of RNA extraction at 65°C was higher when cells grown to mid-exponential growth phase were used, rather than stationary phase cells, whose cell walls are less easy to permeabilize. When cells already treated with a cell wall-digesting enzyme (containing  $\beta$ -1,3-glucanase) were extracted at 65°C, even after only partial digestion, the yield of RNA was 100%. SDS/phenol permeabilization was routinely used for RNA extraction from whole cells and from intracellular fractions.

On the gel in Figure 1.1, one RNA species, intermediate in size between  $5.8S^1$  and 18S rRNA, is more readily extractable than other species of approximately the same size. It is one of the three abundant small RNAs the cloning of whose genes is described in chapter 2, and has been called small cytoplasmic RNA 1 (scR1). RNA extracted from whole cells at low temperatures is enriched in scR1, as shown in Figures 1.1A, odd numbered

<sup>&</sup>lt;sup>1</sup>5.8S rRNA has heterogeneous 5' ends in *Saccharomyces cerevisiae* (Rubin 1974) and appears on gels as a double band.

# Figure 1.1 Differential extraction of RNA by SDS/phenol permeabilization of yeast cells.

Yeast cells (wild type strain ATCC 25657) harvested in exponential growth phase in YPD were pelleted, washed in water and resuspended in one volume of "NaOAc": 0.1M NaOAc pH 5.0, 1% SDS; "NaOAc/EDTA": 0.1M NaOAc pH 5.0, 20mM EDTA, 1% SDS; or "TE": 10mM Tris pH 8.0, 10mM EDTA, 0.5% SDS. One volume of phenol (water or TE-equilibrated for NaOAc or TE, respectively) was added, and the suspensions were incubated for the indicated times and temperatures with intermittent vortexing, then placed on ice for 5'. The aqueous phase was saved after centrifugation and extracted successively with 1 vol. phenol, 1 vol. phenol/chloroform (1:1), 1 vol. chloroform, and then ethanol-precipitated in the presence of 0.3M NaOAc. 70% ethanol-washed pellets were dried, and resuspended in water. RNA concentration was estimated according to 45 µg/ml = 1.0 O.D.260. 10 µg RNA, heated at 95°C, 2' in 80% formamide, was loaded in each lane of a 6% polyacrylamide (30:1 mono/bis), 8M urea gel in TBE (20 cm x 20 cm x 1 mm) run at up to 500 V. A: RNA extracted at various temperatures (odd nos.) and re-extracted at 65°C from the remaining pellets (even nos.). Samples processed in duplicate. B: RNA extracted in "TE" at 20°C (1) and at 65°C (2), and at 65°C after prior 20°C extraction (3) (5 µg and 10 µg of each duplicate pair is shown). The total yield of RNA in each fraction was as follows:

# Yield of RNA from $3 \times 10^9$ cells

(mean of two samples)

1.	20°C, 15′	0.2 mg
2.	65°C, 30′	1.4 mg
3.	65°C, 30' after 20	)°C, 15′ 1.0 mg

C: 10  $\mu$ g yeast RNA (Y) from Figure 1.1B, lane 1, compared with 10  $\mu$ g HeLa cell cytoplasmic RNA (H), a gift from M.-L. Melli.



lanes ' to 11, 1.1B, lanes 1, and 1.1C, lane Y, whereas snR17, another small RNA, is only readily extractable at 65°C (see Figures 1.1A, even numbered lanes 2 to 12, lanes 13 and 15, and 1.1B, lanes 2 and 3).

The RNA extracted from yeast cells in TE/SDS/phenol at room temperature (Figure 1.1C, lane Y) shows a similar pattern of relative band intensities on polyacrylamide gels to HeLa cell cytoplasmic RNA (Figure 1.1C, lane H): both samples have the same relative proportions of the various rRNA species and tRNA. The most abundant HeLa cell cytoplasmic RNA intermediate in size between 18S and 5.8S rRNA is the 7SL RNA of the SRP: interestingly, the only yeast RNA of similar abundance in this size range appears to be scR1.

The yield of RNA extracted in TE/SDS/phenol at room temperature (Figure 1.1B, lanes 1) represents only 16% of the total RNA extractable at 65°C (lanes 2, see Figure 1.1 legend). Therefore, despite the enrichment of scR1 against its background in r.t. extracts (lanes 1), the still strong presence of the band in the pellet RNA (lanes 3) indicates that most scR1 still remains in the cells.

#### 2. Salt precipitation of naked RNA.

When total yeast RNA, thoroughly phenol/chloroform extracted, ethanolprecipitated and redissolved in water, was incubated on ice in the presence of sodium chloride, separation of pellet and supernatant samples enabled fractionation of the RNA roughly according to size; large RNA tended to be precipitated and smaller species remained in solution. The efficiency of precipitation was to some extent dependent on RNA concentration; a minimum of 100-200 µg/ml gave good yields in the precipitates. A concentration of 1.0M NaCi left over 80% of the RNA in the supernatant, 1.5M NaCl split the total yields equally between pellet and supernatant, and 2M NaCl precipitated the greater proportion of RNA including most of the large rRNA and 5.8S rRNA, leaving most of the 5S rRNA and tRNA in solution. Increasing the NaCl concentration above 2M caused no further change in respective yields or relative size distribution of RNA in pellet and supernatant fractions. Figure 1.2 illustrates the method applied to total yeast RNA. Samples were processed in duplicate in order to show the consistency of the band distribution. The soluble fractions (Figure 1.2, Spnt.) contain several small RNA bands that appear to have been enriched. Among these are two species whose genes were cloned, named snR30 and snR17. It is noticeable that scR1, which was readily released by whole cell permeabilization, is selectively precipitated by 2M NaCI, as it is only visible in the pellet and total RNA fractions.

# Figure 1.2 Fractionation of total yeast RNA by precipitation in 2M NaCl.

Yeast RNA, 0.1 mg/ml, extracted from whole yeast cells by treatment at  $65^{\circ}$ C for 30' in TE/0.5% SDS/one volume phenol ("Total RNA") was incubated on ice in 2M NaCl. After centrifugation, pellet ("2M NaCl Pellet") and supernatant ("2M NaCl Spnt.") fractions were precipitated in 2.5 volumes of ethanol, redissolved and loaded (10 µg/lane) on a 6% polyacrylamide (30:1 monomer/bis), 8M urea gel in Tris-borate/EDTA buffer and stained with ethidium bromide. Duplicate samples were processed. SnR30 and snR17 were enriched in the supernatant fraction.



3. Fractionation of RNPs: nuclear/cytoplasmic distribution.

An experiment was performed to test the relative efficiencies of spheroplast lysis compared to French pressure cell disruption as means of disrupting yeast cells for studying RNPs in crude nuclear and cytoplasmic fractions. The experimental scheme is illustrated in Figure 1.3 and the different RNA fractions derived are shown in Figure 1.4. Either spheroplasts were prepared as described in Chapter 6, Section 1.2, washed and vigourously resuspended with a Dounce homogenizer, or a cell suspension was passed through a French press at 20 000 psi. All treatments, except the spheroplast digestion, were at 4°C. After cell disruption, any remaining whole cells were removed by a low speed spin (fraction A in Figures 1.3 and 1.4), the supernatant was spun at 35000 x g for 30' to pellet nuclei (fraction B) leaving a supernatant containing cytoplasmic components (fraction C). Cells were broken in two slightly different types of buffer: 1, based on the buffer containing Mg(OAc)2 used to fractionate SRP (Walter and Blobel 1980), and 3, a buffer used for preparing hnRNP fractions from HeLa cells (Brunel et al. 1981), henceforth known as "hnbuffer". Furthermore, to test whether any SRP-like particles could be selectively salt-washed from the the pellets, cells were also disrupted in buffer 2, consisting of buffer 1 with 0.5M KOAc, and in buffer 4, consisting of buffer 3 with 0.5M KOAc. RNA was extracted from all samples by the 65°C TE/SDS/phenol method (Chapter 6, Section 1.6) and 10 µg from each was analysed by electrophoresis on a denaturing polyacrylamide gel.

The results of this experiment, displayed in Figures 1.3 and 1.4, can be summarized as follows:

- 1. Both cell disruption methods yielded RNA of the same quality, in terms of degradation, from the various fractions.
- 2. The relative yields from fractions A, B, and C (table in Figure 1.3) did not differ significantly between the two cell disruption methods at low salt concentrations (1 and 3). High salt preparations (2 and 4) from spheroplasts gave much higher amounts of whole cells in fraction A (reflected by the RNA yields) due to inefficient lysis, because breakage of the spheroplasts is much more dependent on osmotic shock. This was also reflected by the greater proportion of tRNA in the nuclear pellets (B) owing to contamination by whole cells (Figure 1.4)
- 3. No significant differences in the distribution of any of the "small RNAs" could be seen between samples in high and low salt buffers.
- 4. The enrichment of tRNA in supernatant C and the depletion of tRNA in pellet B indicates that a crude fractionation of nuclei and cytoplasm has occurred.

# Figure 1.3

Comparison of RNA extracted from crude nuclear and cytoplasmic fractions prepared in high or low [salt] from pressure-ruptured cells or lysed spheroplasts (see Figure 1.4).



Supernatant C

(cytoplasm)

# Yield of RNA (%) in each of fractions A, B, and C

Pellet B

(nuclei)

		French Press					Spheroplasts			
		Soln.	1	2	3	4	1	2	3	4
		[Salt]	Lo	Hi	Lo	Hi	Lo	Hi	Lo	Hi
Pellet	Α	7		18	8	15	7	67	29	72
Pellet	В	.15		13	14	15	19	13	11	14
Spnt.	С	78		69	78	71	74	20	60	14

Mean total yield of RNA (sum of fractions A, B & C) for the eight samples, representing the yield from approx. 10  $^{10}\,$  cells, was 4.2mg.

Figure 1.4

Comparison of RNA extracted from crude nuclear and cytoplasmic fractions prepared in high and low [salt] from pressure-ruptured yeast cells or lysed spheroplasts.

The experimental scheme and the total yields of RNA in the different fractions are summarized in Figure 1.3. Yeast cells were harvested in exponential growth phase in YPD, washed in water and treated as described for the preparation of spheroplasts (Chapter 6, Section 1.2) or suspended in each of solutions 1-4 (see Figure 1.3: 1 and 3 were low [salt], 2 and 4 contained 0.5M KOAc) and passed at 4°C through a French pressure cell at 20 000 psi. A, B and C represent samples from the low speed (whole cell) pellet, nuclear pellet and cytoplasm, respectively. T = total yeast RNA, prepared by 65°C SDS/phenol permeabilization. 10  $\mu$ g RNA was loaded in each lane, as described for the previous figures.

43



5. The abundant small RNA scR1 has been enriched in the cytoplasmic fractions C and depleted from the nuclear fractions B: snR17 has been enriched in the nuclear fractions and depleted from the cytoplasmic fractions. SnR30, the third small RNA whose gene was later cloned, is less easy to identify unambiguously.

4. Fractionation of RNPs: mobility shift compared to naked RNA on sucrose gradients.

An experiment was performed in order to find out whether any yeast small RNAs obviously behaved as RNP complexes. Cell extracts and RNA from which protein had been removed ("naked RNA") were centrifuged through continuous sucrose density gradients and the mobilities of individual RNA species were compared.

Yeast cells harvested in mid-exponential growth phase were washed, resuspended in hn-buffer at 4°C and passed twice through a French pressure cell at 20 000 psi. After removing cell debris, 4 ml of supernatant from a 35000 x g spin (corresponding to an extract from approximately  $10^{10}$  cells, or 1.5 g wet pellet) were loaded onto a 35 ml 15-30% sucrose gradient in hnbuffer. Approximately 2 mg of phenol/chloroform extracted RNA from an equivalent cell fraction were loaded onto an identical gradient, and the samples were centrifuged in a Beckmann SW27 rotor at 27 000 rpm for 20 hrs. Gradient fractions were phenol/chloroform extracted, ethanol precipitated, redissolved and analysed on denaturing polyacrylamide gels.

The cell extract gradient (Figure 1.5A) shows tRNA to migrate as a broad peak in the top half of the gradient with its peak one quarter of the way down (Figure 1.5A, lane 4). A small 5S rRNA peak occurs slightly below this, otherwise all the rRNA is in the pellet (Figure 1.5A, lanes P). Two bands, which on other gels were seen to have the same electrophoretic mobility as scR1 and snR17, form peaks just over half way down the gradient (Figure 1.5A, lanes 9 and 10). This migration pattern can be compared to that of naked RNA (Figure 1.5B) in which tRNA, 5S and 5.8S rRNA form a peak one sixth of the way down the gradient (Figure 1.5B, lane 3), RNAs of a size similar to scR1 and snR17 one third of the way down (lane 6), and 18S rRNA half way down (lane 10). A second 5.8S rRNA peak comigrated with 25S rRNA; it is known that 5.8S rRNA associates with the large eukaryotic rRNA by hydrogen bonding (Hadjiolov 1985).

## Figure 1.5

Mobility of cytoplasmic RNPs compared to naked RNA on sucrose gradients.

20 hr 15%-30% sucrose gradients in hn-buffer (10mM Tris pH 7.5, 10mM NaCl, 1.5 mM MgCl<sub>2</sub>) centrifuged in a Beckman SW27 rotor at 27 000 rpm, 4°C.

A: Cytoplasmic supernatant fraction (Sorvall SS34, 17 000 rpm, 30' supernatant, equivalent to fraction "C" of Figure 1.3) In up to 5 ml was loaded onto 35 ml gradients. This was equivalent to the material from approximately 1 g wet cell pellet after two passages through a French press (20 000 psi). The gradient was divided into 16 fractions and an equal proportion of each fraction was loaded onto the gel.

**B**: An identical gradient from the same centrifugation loaded with approximately 2 mg purified RNA from the cytoplasmic supernatant fraction. The gradient was divided into 18 fractions.

ł



These results show that when cell extracts are loaded onto continuous sucrose gradients, most RNA species migrate significantly faster than when naked RNA is loaded onto the same gradients. Bands of the size of scR1 and snR17 on gradients of cell extracts migrate in about the same region as naked 18S rRNA on identical gradients.

# 5. Fractionation of RNPs: sucrose gradients and attempts at further purification.

Figure 1.5A shows peaks of what appear to be scR1 and snR17 half way down 20 hr sucrose gradients of cell extracts (lanes 9 and 10). However, bands of the same size but fainter can still be seen deeper in the gradient (lanes 12 to 16). In order to find out whether peaks of these or other small RNA bands migrated to greater densities, 15-30% sucrose gradients were loaded with 35 000 x g supernatant material as in the previous experiment, but centrifuged only for 6 hrs. The gradient fractions are shown in Figure 1.6. Two bands, clearly comigrating on the gel with the prominent bands labelled scR1 and snR17 in the total RNA marker sample (Figure 1.6, lanes T), form prominent peaks in fractions 3 and 4 (lanes 3 and 4) of the 6 hr gradient, but do not appear to peak again lower down. The most prominent feature of the gradient is the ribosomal peak (80S) in fractions 9-11, distinguished by the heavy presence of large, 5.8S and 5S rRNAs.

It was not clear from the gradient shown in Figure 1.6 whether the large amount of RNA in the ribosomal peak was obscuring denser peaks of scR1 and snR17. For this reason, an attempt was made to fractionate further the gradient fractions containing snR17, scR1 and the ribosomes by electrophoresis on non-denaturing ("native") gels.

Fractions 2, 3, 4, 5 and 10 from a 6 hr sucrose gradient, identical to the one shown in Figure 1.6, were loaded directly onto native gels of 1 mm thick 3% polyacrylamide/bis 20:1 at 4°C and electrophoresed at 7V/cm. Two different gel running buffers were tried: 1. containing 10mM Mg(OAc)<sub>2</sub> at pH7.6 (Figure 1.7A and B), and 2. without Mg<sup>2+</sup> at pH8.4 (Figure 1.7C and D). After the runs, the two gels were both stained, first with ethidium bromide to visualize RNA under UV light (Figure 1.7A and C), and then with Coomassie blue to visualize protein (Figure 1.7B and D). It was hoped that RNP could be identified as bands stainable both by ethidium bromide and Coomassie blue. Ribosomes and ribosomal subunits could be visualized in this way: in the buffer containing Mg<sup>2+</sup>, whole ribosomes migrated as a single, RNA/protein band

# Figure 1.6 Mobility of cytoplasmic RNPs on a sucrose gradient.

6 hr 15%-30% sucrose gradient loaded with a yeast cell cytoplasmic fraction, as described for Figure 1.5.





# Figure 1.7 Native polyacrylamide gel electrophoresis of cytoplasmic sucrose gradient fractions.

Fractions 2, 3, 4, 5 and 10 from a cytoplasmic sucrose gradient, like that shown in Figure 1.6, loaded onto 3% polyacrylamide gels (20:1 mono/bis, 1 mm x 20 cm x 20 cm, 7 V/cm,  $4^{\circ}$ C) run in 25mM Tris, 60mM KOAc, 10mM Mg(OAc)<sub>2</sub>, pH 7.6 (A and B) or 40mM Tris, 20mM NaOAc, pH 8.4, (C and D) and stained with ethidium bromide (A and C) or Coomassie blue (B and D). R = ribosomes, L = large ribosomal subunit, S = small ribosomal subunit, 5S = 5S rRNA.



(Figure 1.7A and B, "R" in lanes 10), while without  $Mg^{2+}$ , dissociated large and small subunits were observed (Figure 1.7C and D, "L" and "S" in lanes 10) and free 5S rRNA was released (1.7C, "5S" in lane 10). The resolution of the gels, however, was insufficient to allow identification of less abundant small RNPs in the fractions containing scR1 and snR17 (Figure 1.7, lanes 2-5).

DEAE ion exchange chromatography was tried as an alternative method to native gel electrophoresis for further purifying small RNPs from sucrose gradient fractions; DEAE columns having the advantage of allowing a large amount of material to be loaded. Gradient fractions were loaded onto DEAE Sepharose equilibrated in hn-buffer, and then eluted with increasing 0.1M steps of KOAc in hn-buffer, up to 1.0M. When RNA extracted from the eluted samples was analysed, no enrichment of small RNPs could be easily detected in any of the fractions.

6. Fractionation of RNPs: from the nucleus.

It seemed worthwhile to have a closer look at yeast cell fractions enriched in nuclei to find out whether scR1 and snR17 could be liberated from these in the form of RNPs with the same sedimentation velocities as those detected in the cytoplasmic (35 000 x g supernatant) fractions.

A 35 000 x g pellet enriched in nuclei was prepared after passage once through a French press in hn-buffer as described for Figure 1.4. The pellet, resuspended in 10 ml cold hn-buffer, was sonicated on ice for 30" using a microprobe attached to a Branson Cell Disrupter B15 (output setting 4). The sonicate was loaded onto a 20 ml cushion of 30% sucrose in hn-buffer and centrifuged at 3000 rpm for 20' as described for the preparation of hnRNP from HeLa cells (Kish and Pederson 1975, Blanchard *et al.* 1977). Material not entering the cushion was loaded onto a 15-30% sucrose gradient as in the previous experiments, and centrifuged for 12 hrs.

The 12 hr gradient, shown in Figure 1.8, resolved a similar range of sedimentation velocities as the 20 hr cytoplasmic gradient of Figure 1.5; the ribosomes having all but pelletted. A small RNA band with a peak between one third and one half the way down the tube (Figure 1.8, lanes 5-7) is probably the scR1 RNP, which has migrated slightly less far into this gradient than into the gradient of Figure 1.5. SnR17 is hardly visible in the nuclear gradient of Figure 1.8: this is surprising because snR17 was shown to be enriched in the nuclear pellet fraction (Figure 1.4).

# Figure 1.8 RNPs released from nuclei by sonication.

12 hr 15%-30% sucrose gradient in hn-buffer (10mM Tris pH 7.5, 10mM NaCl, 1.5 mM MgCl<sub>2</sub>) centrifuged in a Beckman SW27 rotor at 26 000 rpm, 4°C, loaded with material released from crude nuclei by sonication. A nuclear pellet, equivalent to fraction B of Figure 1.3, was sonicated on ice in 10 ml hn-buffer and centrifuged into a cushion of 30% sucrose (3000 rpm, 20'). Material not entering the cushion was loaded onto the gradient.



In order to find out which fractions of the nucleus contained snR17, other fractions of the nuclear preparation were analysed as shown in the scheme in Figure 1.9. Nuclei were prepared following the protocol of Schulz (1978) (Chapter 6, Section 1.3) and then fractionated according to a protocol described for the preparation of HeLa cell hnRNP (Kish and Pederson 1975, Blanchard *et al.* 1977). The nuclear pellet, washed in Ficoll solution, was suspended in 10 ml hn-buffer, sonicated and spun into a 30% sucrose cushion, as in the previous experiment. Material not entering the cushion was loaded onto a three-step sucrose gradient and centrifuged in a SW27 rotor at 26 000 rpm for 90°. The step gradient was fractionated and the RNA analysed as shown in Figure 1.9. HnRNP was expected to be enriched in the 45% sucrose step (fraction 5 in Figure 1.9), while the pellet at the bottom of the 30% sucrose cushion (Figure 1.9, fraction 3) was expected to be enriched in nucleoli (Kish and Pederson 1975, Blanchard *et al.* 1977).

ScR1 features prominently in fraction 4 of the step gradient; very little of snR17 is visible in this fraction. Neither scR1 nor snR17 are visible in the hnRNP fraction (Figure 1.9, fraction 5). This is consistent with the previous experiment in which only scR1, but not snR17, appeared to be released from nuclear fractions after sonication. Nearly all the nuclear snR17 appears to have been retained in the nucleolar fraction. [The pellet from the bottom of the step gradient was not examined (Figure 1.9, fraction 6). This contained less total RNA than the other gradient fractions.]

These results are consistent with there being two distinct populations of snR17 in yeast cells, one, the greater part, being tightly associated with pellet fractions enriched in nucleoli, and the other, the lesser part, being either readily released from nuclei as a small RNP on cell disruption, or cytoplasmic. There was no suggestion of a similar partitioning of scR1, which appeared to be predominantly cytoplasmic, migrating in sucrose gradients as a small RNP. The fraction of scR1 cosedimenting with nuclei could represent cytoplasmic contamination.

54

# Figure 1.9 Fractionation of yeast nuclei.

Spheroplasts were prepared from cells harvested in exponential growth phase (Chapter 6, Section 1.2). Nuclei were prepared by the protocol of Schulz (1978) (Chapter 6, Section 1.3) and fractionated as shown in the figure according to a protocol for the preparation hnRNP from HeLa cells (Kish and Pederson 1975, Blanchard *et al.* 1977). 10, 15 and 20  $\mu$ g of each of fractions 1, 2 and 3, approximately 10 and 15  $\mu$ g of fraction 4, and less than 10  $\mu$ g fraction 5 were loaded on the gel. Fraction 6 was not loaded.



#### DISCUSSION

Electrophoretic analysis of the RNA extracted from whole yeast cells revealed many bands of intermediate size between 18S and 5.8S rRNA, many of them larger than the traditional definition of "small" RNAs, and all of them in much lower relative abundance than rRNA and tRNA. Many of them were consistently present in different yeast RNA preparations implying that they might be discrete small RNAs and not merely non-specific breakdown products. It was possible to enrich for some of these either by SDS/phenol permeabilization of whole cells, or by removing a 2M NaCl cut of high molecular weight RNA from total RNA.

Two of the most abundant small RNA species visible in whole cell RNA extracts, named scR1 and snR17, could be identified by their size with some confidence in intracellular fractions. Crude nuclear and cytoplasmic fractions appeared to be enriched in snR17 and scR1, respectively, but the nuclear pellet was seen to contain significant amounts of scR1, and the cytoplasmic supernatant was seen to contain some snR17. Both scR1 and snR17 from cytoplasmic fractions migrated in sucrose gradients with sedimentation velocities similar to naked (protein-free) 18S rRNA, suggesting that they exist as RNPs with approximately this sedimentation coefficient. The snR17 in the cytoplasmic fractions could have leaked from nuclei during the cell fractionation procedure, or could represent a cytoplasmic component of this RNA. SnR17 remaining in the nuclei after fractionation was not readily released, even after sonication, but was found concentrated in pellet fractions after centrifuging sonicated nuclei into 30% sucrose cushions. Such pellet fractions both from HeLa cells (Kish and Pederson 1975) and from yeast (Sillevis Smitt et al. 1973) have been shown to be enriched in nucleoli.

Just as 7SL RNA is the most abundant small cytoplasmic RNA of HeLa cells, scR1 appears to be the most abundant small cytoplasmic RNA of yeast. ScR1 was the small RNA most readily released from SDS/phenol permeabilized cells, and it was enriched in cytoplasmic fractions as RNP. ScR1 could also be released from nuclear pellets after sonication in the form of RNP which was indistinguishable from the cytoplasmic RNP on sucrose gradients.

It was impossible to make firm conclusions about the intracellular distribution of yeast small RNPs without means of positively identifying the RNAs. In order to obtain small RNA-specific probes, the strategy of the investigation was changed to one of gene cloning.

# Chapter 2

Cloning of yeast small RNA genes and their use as probes to study sub-cellular RNA distribution.

## SUMMARY

Genes for yeast small RNAs were isolated with the aims of: 1. creating specific hybridization probes with which to study the distribution of small RNAs in various sub-cellular fractions; 2. determining the importance and functions of small RNAs by introducing deletion mutations into their coding sequences.

Gel-purified small RNAs were used to screen a yeast genomic DNA library in the bacteriophage vector  $\lambda$ EMBL4. Genes for three small RNAs, named snR30, scR1 and snR17, were isolated and sub-cloned into pEMBL plasmid vectors. When the genes were used as hybridization probes on blots of yeast genomic DNA, it was found that the *SNR30* and *SCR1* probes hybridized to single DNA fragments, indicating that snR30 and scR1 are encoded by unique genes. However, *SNR17* hybridized to two fragments representing two different genes: the second gene was isolated from a separate 20 kb clone in the genomic library using the first gene as a probe. The two genes were called *SNR17A* and *SNR17B*.

The DNA clones were used to probe different RNA fractions. SnR17 was confirmed to be strongly enriched in a purified nuclear fraction, and scR1 was confirmed to be cytoplasmic. SnR30 was also found to be nuclear, but less strongly so than snR17. None of the three RNAs were polyadenylated. Both snR30 and snR17 could be immunoprecipitated by rabbit antiserum against 2,2,7-trimethylguanosine (m<sub>3</sub>G) cap, but not scR1. Nuclear location, lack of polyadenylation, and possession of m<sub>3</sub>G caps are features that distinguish U snRNAs from other transcripts; snR30 and snR17, therefore, can be considered to be members of this class of molecules. Like mammalian U3 snRNA, snR17 was observed to be hydrogen bonded to 37S precursor rRNA (pre-rRNA). Neither the cytoplasmic RNA, scR1, nor the two nuclear RNAs copurified with yeast mitochondria.

#### INTRODUCTION

Saccharomyces cerevisiae is a eukaryote which lends itself to analysis of its biological processes by both biochemical and genetic means. The ability physically to isolate and define functional entities by sub-cellular fractionation combined with the possibility of isolating genes, modifying them in a controlled way *in vitro*, and reintroducing them into yeast cells, offers a powerful approach to the elucidation of cellular functions. As a means of providing handles to get to grips with the the problem of deducing small RNA function, the genes for small RNAs were isolated and cloned.

The cloning strategy employed involved purifying small RNAs of interest for use as hybridization probes to screen a yeast genomic DNA library. Construction of the library required the preparation of large, unsheared fragments of yeast DNA, which then could be partially digested with the restriction endonuclease *Sau*3A, size fractionated to lengths of approximately 20 kb, and cloned into a  $\lambda$  replacement vector. Discussion of the principles and experimental methods, and protocols for constructing and screening  $\lambda$  DNA libraries may be found in Brammar (1982), Murray (1983), Arber *et al.* (1983), Kaiser and Murray (1985) and Frischauf (1987).

Small RNA genes borne on  $\lambda$  clones, identified by hybridization to RNA probes, were isolated on smaller fragments of DNA in plasmid vectors. The isolated genes were then used as probes to study the distribution of specific small RNAs in different sub-cellular fractions.

Background to the pEMBL plasmid system.

The Escherichia coli filamentous bacteriophages fd, f1, M13 constitute a family of rod-shaped nucleoproteins consisting of a circular single-stranded DNA genome ensheathed in about 2400 helically arranged capsid proteins, and capped by three or four "A" proteins at one end. The fact that they can accommodate insertions into the genome of up to 7 x the normal genomic size simply by becoming longer rods, and the fact that they enable the easy preparation of single-stranded DNA for use as templates for DNA sequencing and making single-strand-specific probes have rendered them useful cloning vectors.

Messing *et al.* (1977) first showed that foreign DNA could be inserted into the intragenic region of M13. They were able to insert a 660 bp *Hin*dIII fragment containing the *E. coli lac* promoter and operator, and a portion of the *lac* Z cistron encoding the first 145 amino acid residues of the  $\beta$ -galactosidase enzyme. This C-terminal-truncated  $\beta$ -galactosidase polypeptide was able to complement an operator-proximal deletion, M15 (Beckwith 1964), in the host *lac* Z gene, and so offered a means of selection for the recombinant phage.

The bacterial host for the M13 and plasmid vectors throughout most of this investigation is *E. coli* K12 71-18:  $\Delta$ [*lac, pro*], F' *lac* I<sup>Q</sup>Z  $\Delta$ M15 *pro*<sup>+</sup> (Messing *et al.* 1977, Yanisch-Perron *et al.* 1985), which carries the M15 deletion on the F episome. (The F episome, which is relatively stable in this strain, can be selected on minimal medium lacking proline.)

By introducing a series of unique restriction sites into the *lac* Z gene fragment, without disrupting its intracistronic complementary function, several useful M13 cloning vectors were constructed (Messing 1983, Norrander *et al.* 1983). Insertion of extra DNA into the cloning sites reduced or, if accompanied by a reading frame shift, abolished the ability of the peptide to complement  $\beta$ -galactosidase activity.  $\beta$ -galactosidase activity in a colony on an agar plate could be observed by the blue colour formation due to the cleavage by the enzyme of the chromogenic substrate X-gal (5-bromo-4-chloro-indolyl- $\beta$ -D-galactoside) in the presence of IPTG (isopropyl-thiogalactoside), an inducer of the *lac* operon. The possibility of having blue/white colony selection for insertions into plasmids was successfully realized with pUC8 and pUC9, simply by transferring the truncated *lac* Z fragment with its multiple cloning sites from M13 vectors (Vieira and Messing 1982).

The pEMBL vectors combine the convenience of being able to work with plasmids with the possibility of being able to prepare single-stranded DNA. pEMBL8 and pEMBL9 (Dente *et al.* 1983, 1985) were constructed essentially by introducing into pUC8 and pUC9 a 1300 bp *Eco*RI fragment consisting of the intragenic region of bacteriophage f1, containing all the features necessary in *cis* for phage DNA replication and packaging (Dotto *et al.* 1981). The new plasmids could replicate either as plasmids, using the colE1 replication origin (reviewed by Cesareni and Banner 1985), or as phage using the f1 replication origin, given that helper phage was present to provide all the factors necessary in *trans* (Cesareni and Murray 1987).

### 1. Genomic DNA preparation.

The construction of the yeast genomic library required the preparation of genomic DNA of sufficiently large fragment length to allow generation of partially cut *Sau*3A fragments of about 20 kb. A modification of the method of Bhargava *et al.* (1972), described in Chapter 6, Section 1.4, was found to yield DNA of sufficiently high molecular weight. The method has the following features: first, the prior preparation of nuclei and the retention of DNA in a dialysis bag to which successive additions of RNase and protease can be made, allowing digestion products to diffuse away, thus minimising the number of pipetting steps and thus the occasions for shearing the DNA; secondly, a final purification step of CsCl density gradient centrifugation, which if performed carefully, was as convenient and less damaging to the size of the DNA than phenol extraction.

# 2. Genomic library in $\lambda$ EMBL4.

The yeast genomic library was prepared by Dr. Cosima Baldari in the  $\lambda$  bacteriophage replacement vector EMBL4 (Frischauf *et al.* 1983) according to the protocols and guidelines presented by Kaiser and Murray (1985). High molecular weight genomic DNA was partially digested with the restriction endonuclease *Sau*3A to give fragments of 20 kb average size (this corresponds to a peak of EthBr staining intensity at around 40 kb on an agarose gel), and further size fractionated on 5-25% (w/v) NaCl gradients in 3mM EDTA pH8 (Beckman SW41 rotor, 37000 rpm, 5 hrs.) to remove fragments less than 15 kb which might ligate randomly to each other and become inserted into the vector.  $\lambda$ EMBL4 DNA was cut with restriction endonuclease *Bam*HI, which cleaves the lateral arms of the vector from the replaceable central fragment, and then with *Sal*I to cleave off the small lengths of polylinker sequence remaining at the ends of the *Bam*HI-cleaved ends of the central fragment, thus reducing the likelihood of the createral fragment being religated to the arms of the vector; the small polylinker fragment was rarified by precipitating the DNA with isopropanol.

Recombinant phage was selected by its ability to grow in a bacterial P2 lysogenic host, its so called spi<sup>-</sup> phenotype ("spi" = <u>s</u>ensitive to <u>P2</u> interference). Wild type  $\lambda$  does not form plaques on bacterial strains harbouring the P2 prophage (Zissler 1971), but  $\lambda$  replacement vectors such as EMBL4 lose the *red* and *gam* genes when the central fragment is replaced by foreign DNA, and

because functional *red* and *gam* genes endow sensitivity to P2 interference, recombinant phage is able to grow on P2 lysogens. Without the *red* gene, however, maturation of packageable  $\lambda$  DNA is dependent on the host being recA<sup>+</sup> (Brammar 1982).

After estimating their respective yields, size-fractionated yeast DNA and vector arm were mixed at a molar ratio of approx. 1:1, ligated and then packaged *in vitro* into infective phage heads as described by Karn *et al.* (1980). The titre of the phage suspension diluted in Tris-phage buffer was determined, and about  $10^5$  pfu were plated out in lawns of the *E. coli* P2 lysogenic strain Q359 (Karn *et al.* 1980) in L-top agar onto L-plates at less than confluent plaque density. The library stock, eluted from the plates in L-broth, was stored at 4°C with a few drops of chloroform.

3. Purification of small RNAs for use as probes.

Yeast small RNAs were enriched as described in Chapter 1 either by partial TE/SDS/phenol permeabilization of whole cells or by removing a 2M NaCI precipitate from total, extracted RNA. Additional purification could also be achieved by separating the enriched samples on a sucrose gradient similar to that of Figure 1.5B. Small RNAs were purified by loading up to 0.4 mg RNA onto preparative denaturing 6% polyacrylamide gels (30:1 mono/bis, 8M urea, in TBE, 20 x 20 x 0.2 cm, run at 400 V), and cutting out ethidium bromide-stained bands from the 5.8S to 18S rRNA size range. RNA was eluted from intact polyacrylamide slices by gentle agitation for several hours at 37°C successively in several slice volumes of 5mM Tris pH8, 0.5mM EDTA, 0.5% SDS, 0.1 mg/ml proteinase K, and twice in several slice volumes of water. The combined eluates were lyophilized to less than 0.5 ml, brought to 0.3M NaOAc, phenol/chloroform extracted, ethanol-precipitated and redissolved in water. Glycogen, pretreated with DEPC, autoclaved, and added to a final concentration of 20 µg/ml, could be added as carrier for the ethanol-precipitation. Figure 2.1 shows a sample set of purified RNA bands analysed on a denaturing polyacrylamide gel, including the RNAs snR30, scR1 and snR17.

RNAs were <sup>32</sup>P-labelled using  $[\gamma$ -<sup>32</sup>P]ATP and T4 polynucleotide kinase. (Chapter 6, Section 5.3) at 5'-OH ends created by limited alkaline hydrolysis (pH9.5, 100°C for 1'). RNA probes thus labelled were separated from unincorporated nucleotides by gel filtration through autoclaved Sephadex G-50 columns equilibrated in TE in Pasteur pipettes. Figure 2.1 Purified RNA bands used as probes to screen a genomic DNA library.

RNAs were purified for use as probes by eluting them from slices cut out of preparative denaturing polyacrylamide gels. Lanes 1 to 8 show a selection of purified RNAs; lanes 3, 4 and 6 contain snR30, scR1 and snR17, respectively. Lane 9 contains total yeast RNA as a marker.



# 4. Screening of the $\lambda$ EMBL4 yeast genomic library.

Detailed practical considerations for screening genomic libraries in  $\lambda$  bacteriophage vectors are discussed by Kaiser and Murray (1985); other practical methods appear in Arber *et al.* (1983). <sup>32</sup>P-labelled RNA probes were used to screen for small RNA genes by plaque hybridization (Benton and Davis 1977). A total of about 20 000 plaques were screened. The size of the haploid *S. cerevisiae* genome contained on 16 chromosomes is approximately 16 000 kb. Assuming the average size of yeast DNA contained in a  $\lambda$ EMBL4 phage is 20 kb, the probability *p* of a single clone containing a unique gene is 20/16000 or 0.00125; the probability that none of a number *n* clones contain the unique gene is  $(1-p)^n$  or  $(0.99875)^n$ , thus the probability of a given gene not being represented in 20000 recombinant phages is  $(0.99875)^{20000} = 1.4 \times 10^{-11}$ . At least one of 2400 plaques would contain a unique gene with 95% certainty, and 5500 plaques with 99% certainty.

Phage stock of the amplified genomic library was plated out in lawns of *E. coli* strain 71-18 in 0.7% L-top layer agar on BBL plates at a plaque density of approx. 2000 pfu/plate. As many duplicate nitrocellulose filters were lifted from each plate (Chapter 6, Section 2.2) as the number of different RNA probes and each probe was hybridized to one filter from each plate. As a control, a probe of mixed 25S, 18S 5.8S and 5S rRNA was also used to screen for rRNA genes.

Nearly all of the positive plaques on small RNA-probed filters appeared to be clones of rRNA genes because the same plaques hybridized even more strongly to the rRNA probe. This pointed to one of the following conclusions: 1. the "small RNA" bands were fragments of rRNA; 2. specificity had been lost owing to very small labelled RNA fragments present in the probe generated by the alkaline hydrolysis, or insufficiently stringent hybridization and washing conditions; or 3. the small RNA probes were contaminated by rRNA degradation products coeluted from the gel. Assuming the latter to be likely, small RNA-positive plaques, which showed little or no signal with the rRNA probe, were picked and subjected to two or three further rounds of replating at low plaque density and rehybridization (Chapter 6, Section 2.3).  $\lambda$  DNA was prepared from plate lysates of those clones that gave consistently good signals with small RNA probes and weak signals with the rRNA probe (Chapter 6, Sections 2.4 and 2.5).

Such clones were analysed in detail as follows. The *Eco*RI digestion patterns of the clones were compared on agarose gels with *Eco*RI digested total genomic DNA. Yeast rRNA genes (rDNA) consist of about 100-150 tandem repeats of 9.1 kb units on chromosome XII (Petes 1979, Zamb and Petes 1982),

and the three largest *Eco*RI fragments of these repeated units are quite easy to recognize as prominent bands in total *Eco*RI digested yeast DNA. Any  $\lambda$  clones with *Eco*RI fragments comigrating with two or three of these prominent bands in *Eco*RI-digested genomic DNA were assumed to be clones of rDNA. DNA from the gels was transferred to nitrocellulose filters and probed with labelled small RNAs or rRNA. Clones were analysed until one non-rDNA clone was found for each of the abundant small RNA probes snR30, scR1 and snR17 (Figure 2.2); these clones were named, respectively,  $\lambda$ SR8,  $\lambda$ SR6 and  $\lambda$ SR3A.

To be sure that the non-rDNA clones contained sequences complementary to the RNA bands originally isolated, 100  $\mu$ g DNA from large scale  $\lambda$  DNA preparations (Appendix 2.6) from each of the non-rDNA clones, and from one of the rDNA clones as a control, were denatured and fixed to Millipore filters (Appendix 5.1). The filters were hybridized to total yeast RNA, washed, and the remaining RNA eluted and 3' end labelled with <sup>32</sup>pCp and T4 RNA ligase (England and Uhlenbeck 1978, Appendix 5.2). The labelled products, analysed on a denaturing gel, all contained RNA bands of the expected size for the original probes with which the DNA was isolated. Additional bands were also present in each lane, presumably the transcripts of other genes, but only one of these, from the *SCR1* clone was more abundant than the small RNA bands; this band migrated with the size of tRNA. [A gene for tRNA<sup>IIe</sup> was later found downstream of the *SCR1* gene, see Chapter 3.]

# 5. Sub-cloning

Small RNA genes from the  $\lambda$ EMBL4 clones  $\lambda$ SR8,  $\lambda$ SR6 and  $\lambda$ SR3A were initially sub-cloned into plasmid vectors as illustrated in Figure 2.3. The pEMBL plasmids (Dente *et al.* 1983, 1985) were chosen because they offer a convenient means of selection for insertions into the vector polylinker according to the white or blue colony phenotype of transformants grown on the chromogenic substrate X-gal, and because they can be prepared either as double stranded DNA (dsDNA) when propagated as plasmids, or as single-stranded DNA (ssDNA) when replicated in the presence of IR-1 helper phage (Cesareni and Murray 1987).

The normal procedure for sub-cloning  $\lambda$  DNA fragments into plasmids was as follows: 2.5 µg plasmid DNA (eg: pEMBL9) and 0.5 µg  $\lambda$  bacteriophage DNA (prepared as described in Appendix 3.4) were digested with the restriction endonuclease of choice (eg: *Eco*RI). The DNA was phenol/chloroform extracted, ethanol-precipitated, redissolved in TE and mixed in a final volume of 40 µl with ligation buffer (20mM Tris/HCI pH7.6, 10mM MgCl<sub>2</sub>, 10mM DTT, 0.5mM

# Figure 2.2 Distinguishing smail RNA gene clones from rDNA clones.

When excised small RNA bands were used as probes to identify small RNA gene clones by plaque hybridization, a high background of signals from rDNA clones was observed, probably due to contamination of the probes with rRNA fragments. The rDNA clones could be identified by their characteristic restriction endonuclease digestion pattern.

Upper panel: *Eco*RI digestion pattern of the yeast rDNA repeat unit. The sizes of the fragments were determined from a compilation of the complete nucleotide sequence.<sup>1</sup>

Lower panel: an ethidium bromide-stained agarose gel loaded with various *Eco*RIdigested DNA samples: yeast genomic DNA (lane 5); six different  $\lambda$ EMBL4 clones of rDNA (lanes 1, 2, 3, 4, 6 and 10), each of which contain at least one of the large *Eco*RI fragments A, B or C (confirmed by hybridization to a rRNA probe); and three  $\lambda$ EMBL4 small RNA clones,  $\lambda$ SR8 containing *SNR30* (lane 7),  $\lambda$ SR6 containing *SCR1* (lane 8), and  $\lambda$ SR3A containing *SNR17A* (lane 9). The small RNA genes were mapped to the *Eco*RI fragments indicated by flecks in

lanes 7, 8 and 9, after hybridization to small RNA probes.

<sup>&</sup>lt;sup>1</sup>The complete nucleotide sequence of the *S. cerevisiae* rDNA repeat unit has been assembled by A. A. Hadjiolov (personal communication) from data of Bayev *et al.* (1980), Rubtsov *et al.* (1980), Bayev *et al.* (1981), Georgiev *et al.* (1981), Skryabin *et al.* (1984) and unpublished data of O. I. Georgiev, N. Nikolaev and A. A. Hadjiolov.



ATP final concentration<sup>1</sup>) and T4 DNA ligase, and incubated at room temperature for 4 hrs. As a control for ligation efficiency and for any spontaneous background of white colonies after the subsequent transformation, cut plasmid DNA alone was treated in the same way. 2 μl of the ligation mixture were analysed on 0.7% agarose gels after 0 and 4 hrs. Ligation was deemed to have been successful when linear plasmid DNA had been seen to be converted into circular forms. 20 μl of ligation mixture was incubated on ice with 200 μl competent 71-18 cells for 45', heat shocked for 2' at 42°C, and plated onto several L-plates containing 0.2 mg/ml ampicillin, X-gal and IPTG (Chapter 6, Section 3.1). After incubation overnight and storage for one or two days to allow strong development of the blue colour, single white colonies were picked onto filters laid on fresh ampicillin plates, and treated as described in Chapter 6, Section 3.2 for colony hybridization using the small RNA probes described previously. Colonies giving rise to a positive signal with probe were propagated in L-broth with ampicillin, and plasmid DNA was isolated and analysed in greater detail.

The cleavage sites of various restriction endonucleases were mapped by digesting the DNA of each different clone with single enzymes alone, and two enzymes successively (or together, depending on their salt requirements), each enzyme with each other. Digested fragments were separated by electrophoresis on agarose gels, transferred to nitrocellulose filters and probed again with the small RNA probes. Analysis of the restriction fragment patterns on the gels enabled restriction sites to be mapped relative to each other, and by probing the fragments on filters, those DNA fragments bearing the coding sequences for the RNAs could be identified. Using this technique genes were mapped to within a few hundred bp.

The strategies and the plasmids used for the initial sub-cloning of fragments from  $\lambda$  bacteriophage clones bearing small RNA genes are shown in Figure 2.3. Small RNA genes were isolated on yet smaller yeast DNA fragments by deleting DNA between restriction sites present in both the yeast gene flanking regions and the vector polylinker region. Thus, for example, pSR3.1, containing the *SNR17A* gene on a 1.4 kb fragment (Figure 3.1), was derived from pSR3b (Figure 2.3) by digesting with *Sal*I, which cuts approximately 150 nt upstream of the gene initiation site and also in the polylinker sequence of pEMBL9, diluting to a DNA concentration of less than 5 µg/ml and treating with T4 DNA ligase. At this relatively low DNA concentration nearly all the plasmid

68

<sup>&</sup>lt;sup>1</sup>Ferretti and Sgøramella (1981) have studied the effects of ATP concentration on blunt and sticky end ligation reactons.





DNA was recircularized with no reinsertion of the Sall fragment, which was subsequently lost when the ligation mixture was used to transform competent cells.

6. Determining the copy number of small RNA genes in the yeast genome.

This chapter has described the isolation of three different genes, each of which appears to encode a small RNA. As a first step towards studying the functions of these genes, it was intended that disruptive mutations be made in order to determine their importance to the cell. It was therefore necessary to find out how many copies of each gene were present in the haploid yeast genome, and so how many disruptions would have to be performed to inactivate them all.

Total yeast genomic DNA was digested with several different restriction endonucleases, electrophoresed on agarose gels and blotted to nitrocellulose filters. The DNA was then probed with nick-translated plasmid DNA clones of the small RNA genes. Each of the lanes in Figure 2.4 represents genomic DNA digested with a different restriction endonuclease and probed with the nick-translated plasmid pSR8.1 containing the cloned *SNR30* gene on a 2.3 kb fragment (Figure 3.3). In each lane a single band only is visible, showing that the *SNR30* gene is unique.<sup>1</sup>

Figure 2.5 shows that the *SCR1* gene is also unique. The plasmid probe pSR6.1 (containing all but the first 38 nt of the *SCR1* gene on a 960 bp fragment, Figure 3.4) hybridizes to a single band only in each of the different lanes.

When the plasmid pSR3.4 (containing the SNR17A gene on a 785 bp fragment, Figure 3.1) was used as a probe, two bands were visible in each lane (Figure 2.6A). Yeast DNA was digested with a total of eleven different restriction endonucleases, and the same result was obtained in each case. To eliminate the possibility that the second band might have been due to hybridization to regions in the probe other than the SNR17A coding sequence, a 237 bp Sau3A fragment taken from within the gene coding sequence (Figure 3.5) was purified from a polyacrylamide gel and used as a probe: two bands were still to be seen, as illustrated in Figure 2.6B. The aene already isolated

<sup>&</sup>lt;sup>1</sup>The pSR8.1 plasmid probe hybridizes to an *Eco*RI fragment of 6.6 kb in Figure 2.4. The *Eco*RI fragment bearing the *SNR30* gene cloned from the genomic library, however, was only 4.3 kb (Figure 2.3). This suggests that one of the *Eco*RI sites of the 4.3 kb fragment was derived from the  $\lambda$ EMBL4 polylinker sequence and does not naturally occur in the yeast genome.
# Figure 2.4 SnR30 is encoded by a unique gene.

Yeast genomic DNA was digested by restriction endonucleases EcoRI(E), HindIII(H), Bg/II(B), Sa/I(S), or PvuII(P), transferred to filters after electrophoresis on a 0.6% agarose gel with a size marker, and probed with a labelled DNA fragment comprising the complete length of the 605 bp transcribed sequence of the SNR30 gene. A single band is visible in each lane.



Figure 2.5 ScR1 is encoded by a unique gene.

Yeast genomic DNA was digested by restriction endonucleases Sall (S), Pstl (Ps), Hpal (H), Pvull (Pv), or Bg/II (B), transferred as for Figure 2.4, and probed with a DNA fragment comprising the complete SCR1 transcribed sequence. A single band is visible in each lane.



## Figure 2.6 SnR17 is encoded by two genes.

Yeast genomic DNA, digested with the indicated restriction endonucleases and transferred to filters after electrophoresis, was probed with **A**: plasmid pSR3.4, containing the *SNR17A* gene on a 785 bp yeast DNA fragment, and **B**: an isolated 237 bp *Sau*3A fragment entirely from within the *SNR17A* transcribed sequence. Two bands are visible in each lane; the stronger of the two corresponds to the *SNR17A* gene. The sizes of some of the fragments are indicated in kilobase pairs.



was referred to as *SNR17A*; the signals due to hybridization to this gene were the 4.3 kb *Eco*RI band and the 1.4 kb *Hin*dIII band. The cross-hybridizing fragment was assumed to represent a second gene; the signals due to this gene were the 1.2 kb *Eco*R1 band and the 4.6 kb HindIII band.

7. Isolation of the second SNR17 gene.

Approximately 10 000 plaques from the  $\lambda$ EMBL4 partial *Sau*3A yeast genomic library were transferred to filters and screened by hybridization to the nick-translated plasmid pSR3.4 (Figure 3.1). Twelve hybridizing clones were isolated and  $\lambda$  DNA, prepared by the plate lysate technique, was digested with either *Eco*RI or *Hind*III, run on agarose gels, transferred to filters, and again probed with pSR3.4. Six of the twelve clones showed digestion patterns similar to the original  $\lambda$ SR3A clone bearing the *SNR17A* gene, and the probe hybridized to a 4.3 kb *Eco*RI fragment and a 1.4 kb *Hind*III fragment. The other six clones exhibited a different band pattern, and the probe hybridized to a 1.2 kb *Eco*RI fragment and a 4.6 kb HindIII fragment. *Eco*RI fragments from one of the latter clones, named  $\lambda$ SR3B, were sub-cloned into pEMBL8, and plasmid clones bearing the second gene were identified by colony hybridization. The second gene, isolated on plasmid pSR3B as shown in Figure 2.3, was called *SNR17B*.

8. Determining the orientation of small RNA genes sub-cloned on plasmids.

pEMBL plasmids can be isolated as either double stranded (ds) or singlestranded (ss) DNA forms because they possess both ColE1-type plasmid and filamentous bacteriophage origins of replication. This facility was exploited to determine the orientation of small RNA genes. SsDNA was prepared (Chapter 6, Section 4.1) from clones of genes which, restriction mapping had shown, were present in both orientations with respect to the vector DNA. The ssDNA was spotted onto filters and hybridized to small RNA probes. The RNA probes annealed only to those clones in which the ssDNA represented the anti-sense sequence of the genes (Figure 2.7). From these results the direction of transcription of the genes was deduced. 74



# Figure 2.7 Determining the orientation of small RNA genes.

Arrows represent the polarity of transcription of the *SNR* gene or of replication of plasmid DNA from the phage F1 origin (5' to 3'). For pSR3a, the ssDNA constitutes the anti-sense sequence of the RNA: a labelled RNA probe, therefore, hybridizes to this. For pSR3b, the ssDNA constitutes the sense strand: consequently, an RNA probe cannot hybridize to it.

;

# THE USE OF CLONED GENES AS PROBES TO STUDY THE DISTRIBUTION OF SMALL. RNAS IN DIFFERENT FRACTIONS.

#### 1. Nuclear/cytoplasmic distribution.

In Chapter 1, the two prominent small RNAs scR1 and snR17 were identified by their size in various sub-cellular fractions on ethidium bromidestained RNA gels. To confirm the cytoplasmic and nuclear locations of these RNAs, and to study the intracellular distribution of snR30, the third RNA the cloning of whose gene has been described in this chapter, purified nuclear and cytoplasmic RNA was prepared and probed on filters using labelled DNA clones of the isolated genes.

Nuclei were purified by centrifugation through Percoll gradients after hypotonic lysis of spheroplasts in 18% Ficoll solution (Ide and Saunders 1981, Schulz 1978; see Figure 2.8 legend). Figures 2.8A, B and C show equal amounts of nuclear (N) and cytoplasmic (C) RNA probed, respectively, with *SNR30*, *SNR17* and *SCR1*. Both snR30 and snR17 can be seen to be enriched, per  $\mu$ g of RNA, in the nuclear fraction, whereas scR1 can be seen to be enriched, per  $\mu$ g RNA, in the cytoplasmic fraction. By probing different quantities of RNA from each fraction on the filters and comparing the different strengths of signal (not shown), it was estimated that snR30 was 10-20 fold more enriched (per  $\mu$ g), and snR17 approximately 100 fold more enriched (per  $\mu$ g) in the nuclear RNA than in the cytoplasmic RNA.

2. SnR30, snR17 and scR1 are not polyadenylated.

Figures 2.8D-G show equal amounts of yeast polyadenylated (+) and nonpolyadenylated (-) RNA on filters probed with various genes. Polyadenylated RNA, which comprises less than 2% of total yeast RNA, was separated from nonpolyadenylated RNA by oligo (dT)-cellulose chromatography (Maniatis *et al.* 1982). Messenger RNA of the *TRP1* gene was found exclusively in the polyadenylated RNA fraction (Figure 2.8E), whereas snR30, snR17 and scR1 were found only in the non-polyadenylated RNA fraction (Figures 2.8D, F and G).

3. SnR30 and snR17 are m<sub>3</sub>G capped.

U snRNAs (except U6) have 2,2,7-trimethylguanosine cap structures at their 5' ends. Rabbit anti-sera raised against this cap structure has been shown to be highly specific, and has no detectable cross-reactivity with 7-monomethyl guanosine cap structures characteristic of mRNAs (Lührmann *et al.* 1982).

#### Figure 2.8

# Distribution of small RNAs in nuclear, cytoplasmic, polyadenylated and non-polyadenylated RNA fractions.

RNA was fractionated on denaturing polyacrylamide gels, electrotransferred to nylon membranes (Genescreen, Dupont, or Hybond-N, Amersham, see Chapter 6, Section 5.7) and screened with <sup>32</sup>P-labelled plasmid hybridization probes containing the small RNA genes *SNR30* (pSR8.1), *SNR17A* (pSR3.10), *SNR17B* (pSR3B), *SCR1* (pSR6.1) and the protein coding sequence of the *TRP1* gene (Tschumper and Carbon 1980).

**A**, **B** and **C**: Equal quantities of nuclear (N) and cytoplasmic (**C**) RNA. Yeast nuclei were prepared by hypotonic lysis of spheroplasts in 18% Ficoll (Schulz 1978, Chapter 6, Section 1.3), the washed nuclear pellet was resuspended in 1M sorbitol, 0.5mM CaCl<sub>2</sub>, 10mM K-phosphate pH 6.5 and further purified by centrifugation in an SW27 rotor at 3500 rpm for 20' through preformed 33% Percoll (Pharmacia) gradients at 4°C (Ide and Saunders 1981). The gradients were preformed by centrifuging 33% Percoll, 1M sorbitol, 0.5mM CaCl<sub>2</sub>, 10mM K-phosphate pH6.5 in an SW27 rotor at 13000 rpm (27000 x g) for 1 hr. The nuclear band was removed from the gradient with a pasteur pipette and RNA was extracted directly from all but an aliquot of the sample, which was diluted with 2 volumes of the 1M sorbitol solution, pelleted, resuspended in 0.15M NaCl, 0.1M EDTA and thereafter treated as described in Chapter 6, Section 1.5 for the preparation of genomic DNA; the presence of DNA in the sample, visualized on an agarose gel, confirmed that this fraction contained nuclei. Cytoplasmic RNA was extracted from the 30 000 x g, 30' supernatant of the initial cell lysate.

**D**, **E**, **F** and **G**: Equal quantities of polyadenylated (+) and non-polyadenylated (-) RNA prepared by repeated passage of total RNA through an oligo(dT)-cellulose column (Maniatis *et al.* 1982, pp197-198).



Rabbit anti-m<sub>3</sub>G serum was bound to protein A-sepharose beads and used to immunoprecipitate  $m_3$ G-capped RNAs from yeast. RNA was extracted from a crude nuclear extract preparation, incubated with the bead-bound antibodies, then re-extracted from the beads as described in Chapter 6, Section 6. In Figure 2.9, approximately equal quantities of nuclear extract RNA (NE) and anti-m<sub>3</sub>G precipitated RNA (CAP) have been probed with the *SNR30*, *SNR17* and *SCR1* genes. As a control, RNA extracted from the same amount of beads without antibody was also probed (CON). Both snR30 and snR17 were strongly enriched in the anti-m<sub>3</sub>G precipitated samples, indicating that they possess  $m_3G$  caps. The cytoplasmic RNA scR1, however, was not immunoprecipitated, indicating that this RNA does not possess this cap structure. A significant amount of scR1 was present in the nuclear extract (see Chapter 1). The possibility that scR1 might have been capped in the cytoplasm but not in the nucleus was discounted when RNAs were anti-cap precipitated from whole cell extracts, and the same result was obtained.

4. SnR17 is hydrogen bonded to pre-rRNA.

U3 snRNAs of vertebrates have been observed to be hydrogen bonded to prerRNA in nucleoli (Prestayko *et al.* 1970, Calvet and Pederson 1981, Epstein *et al.* 1984). In the light of the similarity observed between the nucleotide sequences of snR17 and U3 snRNAs (Chapter 3, Figures 3.12, 3.13), an experiment was performed to find out whether snR17 could be detected hydrogen bonded to yeast pre-rRNA.

The following fractionation was performed by David Tollervey (Tollervey 1987a). Yeast extracts were deproteinized under conditions expected to maintain hydrogen bonding by digestion with proteinase K, centrifuged through a continuous sucrose density gradient, and the RNA fractions from the gradient analysed on filters with gene probes for various yeast snRNAs, including the plasmid pSR3.4 (Figure 3.1) for snR17.

The upper panel of Figure 2.10 shows the ethidium bromide-stained RNA of the sucrose gradient fractions. The lower panels show the same gradient fractions probed with various snRNA probes, and probes for pre-tRNA<sup>Leu</sup> and tRNA<sup>Leu</sup> as controls. Lane C is the material loaded onto the gradient, lane 1 is the top and lane 15 is the bottom of the gradient. Analysis of the high molecular weight RNA fractions revealed that the peak of mature 18S rRNA was in lane 5, and that of 25S rRNA was in lane 7. 5.8S rRNA, which is hydrogen bonded to 25S rRNA, also shows a peak in lane 7. The peaks of the major pre-rRNA

Figure 2.9 Immunoprecipitation of small nuclear RNAs with anti-m<sub>3</sub>G cap antibodies.

Nuclear extract RNA before (NE) and after immunoprecipitation with protein Asepharose beads bound to rabbit anti-trimethylguanosine cap antiserum (Lührmann *et al.* 1982) (CAP), or beads without antiserum as control (CON), probed on filters with clones of *SNR30*, *SCR1*, *SNR17* and *SNR3* (Tollervey *et al.* 1983) genes. Immunoprecipitations were performed in duplicate as described in Chapter 6, Section 6.



#### Figure 2.10 Several yeast small nuclear RNAs, including snR17, are hydrogen bonded to pre-rRNA.

RNA was prepared from yeast cell extracts under conditions designed to maintain hydrogen-bonding: Spheroplasts were prepared from cells grown in minimal medium, further incubated in minimal medium after spheroplasting to allow recovery of normal metabolic activity, and then lysed in 100mM NaOAc, 50mM Tris-OAc pH7.9, 5mM Mg(OAc)<sub>2</sub>, 10mM DTT, 10mM vanadyl ribonucleoside complex at 4°C. 2% SDS and 2 mg/ml proteinase K were added, the mixture was incubated at 14°C for 15', clarified by brief centrifugation, then loaded onto 10-30% sucrose gradients in lysis buffer and centrifuged in an SW41 rotor at 22000 rpm for 16 hrs at 2°C. After denaturing electrophoresis, RNA from the gradient fractions was ethBr-stained (upper panel) and probed on filters with various snRNA genes (lower panels). 37S pre-rRNA was shown to peak in fractions 9/10, and 20S pre-rRNA in fractions 5/6 (see Tollervey 1987a). C = material loaded onto the gradient. Lane 1 represents the top of the gradient, and lane 15 represents the bottom.

This experiment was performed by David Tollervey.



species in the gradient were detected with a DNA probe from the first internal transcribed spacer (ITS1) of the yeast rDNA gene (see Figure 3.20), thus 37S, 27S and 20S pre-rRNAs were found to peak in lanes 9/10, 7/8 and 5/6, respectively (not shown in Figure 2.10, but shown by Tollervey 1987a).

The snRNAs snR17, snR10, snR4 and snR3 all form peaks in lanes 9/10 with 37S pre-rRNA, and also at the top of the gradient. SnR8 shows a peak in lanes 5/6 with 20S pre-rRNA, snR9 is smeared from lanes 5 to 10. SnR7 is only visible at the top of the gradient. SnR7 is the *S. cerevisiae* U5 snRNA, involved in pre-mRNA splicing (Patterson and Guthrie 1987).

When the material from the gradient fractions containing 20S and 37S pre-rRNA shown in lanes 9 and 5 were run over a second, identical sucrose gradient, the snRNAs associated with these fractions remained quantitatively bound. When the fractions were first heated to 65°C and then loaded onto a second sucrose gradient, however, the snRNAs were released and remained at the top of the gradient.

Yeast 37S pre-rRNA is the longest detectable initial transcript from *S. cerevisiae* rDNA genes (Bayev *et al.* 1981). It consists of the complete mature 18S, 5.8S and 25S rRNA sequences separated by spacers (ITS1 and ITS2) and flanked by external transcribed sequences (ETS; see Figure 3.20). It is likely that snR17, snR10, snR4 and snR3 were hydrogen bonded to the 37S pre-rRNA transcript in these gradients, and not to any other nucleic acid species: few yeast pre-mRNAs are as large as this (6630 nt), and it is unlikely that DNA would give such a clean fractionation pattern in these gradients.

5. The cytoplasmic RNA scR1 is not mitochondrial.

Cell fractionation studies described in this and the previous chapter have indicated that most scR1 is cytoplasmic. In order to check the possibility that scR1 could be associated with mitochondria, RNA was extracted from purified yeast mitochondria and probed with *SCR1* DNA. At least one small RNA, the RNA component of the mitochondrial MRP RNP, has been shown to be synthesised in the nucleus and transported into mitochondria (Chang and Clayton 1987).

Mitochondria were prepared from spheroplasts of yeast grown in lactate medium by lysis in 0.6M sorbitol (Daum *et al.* 1982). After removing unbroken cells by low speed centrifugation, the mitochondria were sedimented from the lysate supernatant and washed several times (see Figure 2.11 legend). Figure 2.11A shows ethidium bromide-stained RNA from the purified mitochondria

(lane 1), from the post-mitochondrial supernatant (lane 2), from the low speed pellet containing mainly unbroken cells (lane 3) and from freshly spheroplasted cells (lane 4).

The substantial purification of the mitochondrial fraction from other cytoplasmic components is illustrated by the lack of cytoplasmic 5S and 5.8S rRNA in Figure 2.11A, lane 1. EthBr-stained bands corresponding to scR1 and snR17 are visible in the whole spheroplast and 3000 rpm pellet fractions (lanes 4 and 3), but not in the mitochondrial pellet fraction. In Figure 2.11B, the same RNA fractions were probed with DNA clones of the *SNR30, SCR1* and *SNR17* genes, confirming the identity of the RNA bands, and showing that none of these RNAs copurified with mitochondria. [Interestingly, one band of unknown origin, within the 300 to 500 nt size range, is enriched in the mitochondrial fractions (Figure 2.11A, lane 1).]

ł

### Figure 2.11 The cytoplasmic RNA scR1 does not copurify with mitochondria.

Mitochondria were purified from yeast grown aerobically in lactate medium as described by Daum et al. (1982), except that the lysis buffer contained sorbitol and not mannitol. For the preparation of well developed, functional mitochondria, it is important to use a non-fermentable carbon source because fermentable sugars, such as glucose, exert catabolite repression. Spheroplasts, prepared in 1.2M sorbitol solution, (lane 4) were lysed with a Dounce homogenizer in 0.6M sorbitol and centrifuged at 3000 rpm (Sorvall, SS34 rotor) for 5'. The pellet was resuspended and centrifuged again; RNA from the second pellet is shown in lane 3. Mitochondria were pelleted from the combined lysate supernatants by centrifugation at 9000 rpm for 10'. The postmitochondrial supernatant RNA is shown in lane 2 (this fraction contains a high concentration of vacuolar nucleases). The mitochondrial pellet was resuspended, homogenized and centrifuged three more times to give a final pellet, reddish brown in colour owing to the enrichment of cytochromes (lane 1). 10  $\mu$ g RNA were loaded in lane 1, and 20  $\mu$ g into each of lanes 2, 3 and 4 of a denaturing polyacrylamide gel. RNA in the gel was stained with ethidium bromide (A), or electro-transferred to filters and probed with DNA consisting of the SNR30, SCR1 and SNR17 genes (B).



#### DISCUSSION

#### 1. SnR30 and snR17 are U snRNAs.

As a class of molecules, U small nuclear RNAs (U snRNAs) can be defined as ubiquitous, metabolically stable RNA molecules, present in the nuclei of eukaryotic cells complexed with specific proteins, and (except U6) possessing at their 5' ends a unique cap structure, 2,2,7-trimethylguanosine ( $m_3G$ ) (Busch *et al.* 1982). Since the availability of specific anti-bodies against this cap, reacting with the RNA both in the presence and absence of associated proteins (Lührmann *et al.* 1982, Bringmann *et al.* 1983), the cap structure above all has become a criterion on which new snRNAs can be identified and subsequently purified.

The genes for three yeast small RNAs have been isolated. By using them as hybridization probes to analyse different RNA fractions, it could be shown that none of the RNAs they encode are polyadenylated, but that snR30 and snR17 are nuclear RNAs and possess m<sub>3</sub>G cap structures.

The genes for snR30 and the cytoplasmic RNA scR1 were each found to be unique in the yeast genome: snR17, however, was found to hybridize to two distinct loci; these were always found on separate 20 kb cloned DNA fragments so they are probably not closely linked. Christine Guthrie and coworkers have identified more than 24 snRNAs in *S. cerevisiae* on the basis of their 5' m<sub>3</sub>G cap structures (Riedel *et al.* 1986), and all those analysed were found to be encoded by unique genes. SnR17 is unusual in being encoded by two genes.

Comparison by hybridization of clones exchanged with C. Guthrie's laboratory confirmed the identity of the smallest of the three RNAs with an RNA designated snR17 by C. Guthrie's laboratory (Riedel *et al.* 1986). For the sake of consistency and to avoid confusion, this name was adopted.

2. Like U3 snRNA, snR17 is associated with pre-rRNA.

In Chapter 1 it was shown how relatively little snR17 was released from nuclei in low ionic strength buffer containing  $Mg^{2+}$  even after sonication (Figures 1.8 and 1.9); most snR17 appeared to copurify with pellet material that was expected to be enriched in nucleoli. Some snR17 was found in cytoplasmic fractions: this might represent a proportion of less tightly bound nuclear snR17 that "leaked" during the fractionation, or it might represent a proportion of snR17 normally present in the cytoplasm. De Robertis *et al.* (1982) and Mattaj and De Robertis (1985) demonstrated in *Xenopus* that U2 snRNA is exported into the cytoplasm shortly after having been synthesised, where it is recognized and bound by RNP proteins,  $m_3G$  capped and re-imported into the nucleus, presumably as mature RNP. [Mattaj (1986) found that  $m_3G$  capping activity was cytoplasmic and dependent on Sm antigen binding.] Zieve *et al.* (1988) have confirmed that all six major U snRNAs, including U3, appear transiently in the cytoplasm of mouse cells shortly after synthesis, and are there assembled into RNPs that antigenically resemble the mature particles.

It has long been known that U3 snRNA is concentrated in the nucleolus (Nakamura et al. 1968, Weinberg and Penman 1968), and the observation that U3 was apparently hydrogen bonded to pre-rRNA in nucleolar fractions first gave rise to the suggestion that snRNAs might be involved in pre-RNA processing (Prestayko et al. 1970). Prestayko found that when nucleolar RNA from rat cells was extracted with SDS/phenol at 65°C and fractionated on a sucrose gradient, all the 5.8S rRNA (referred to as "7S" by Prestayko et al. 1970), its "8S" precursor and U3 migrated in the 4-7S region of the gradient. When the extraction was performed at 25°C, however, 5.8S rRNA, its precursor and U3 migrated in a 28S peak; a significant quantity of U3 was also still observed at 4-7S. The behaviour of U1 and U2 snRNAs were studied as controls and were always found in the 4-7S range irrespective of the extraction temperature. The observation that U3, 5.8S rRNA and its precursor were associated with a large nucleolar RNA of about 28S, but that U3 was not associated with cytoplasmic 28S rRNA, gave rise to the hypothesis that U3 might play a role in the cleavage of 32S pre-rRNA into mature 5.8S and 28S forms (the pre.rRNA processing pathway is outlined in Figure 3.20).

An interaction of U3 with large nucleolar RNA was confirmed in HeLa cells by Calvet *et al.* (1981) by psoralen cross-linking under conditions which also enabled them to show that U1 could be cross-linked to pre-mRNA in whole nuclei.

Subsequently, Epstein *et al.* (1984) studied the role of proteins in binding U3 in the nucleolus and found that when isolated rat nucleoli were extracted in low ionic strength buffer, 65% of U3 was associated with pre-ribosomal RNP, sedimenting at >60S, while 15% sedimented at less than 20S. Of the 65% in pre-rRNP, relatively little could be released by heating at 55°C, whereas most of the 5.8S rRNA could; but most of the U3 could be released by proteinase K or 22°C SDS/phenol treatment, whereas 5.8S could not, leaving only 10% of the total nucleolar U3 still hydrogen bonded to pre-rRNA, and over 90% of 5.8S; the hydrogen bonding of 5.8S, moreover, was more stable to denaturation than that of U3.

These results indicate that the hydrogen bonding between U3 snRNA and pre-rRNA is much weaker and possibly more dependent on proteins for stability than that between 5.8S rRNA and 28S rRNA.

U3 is among the most abundant snRNAs in vertebrates: snR17 is one of the most abundant snRNAs in *S. cerevisiae.* Like U3, snR17 is associated with pre-rRNA and appears to copurify with nucleoli. However, snR17 peaks in sucrose gradients with 37S pre-rRNA, whereas mammalian U3 was observed to peak with 5.8S rRNA at 28S. The 37S pre-rRNA is the longest pre-rRNA transcript in yeast, containing 18S, 5.8S and 25S sequences in contiguous array (see Figure 3.20): in vertebrates, the equivalent pre-rRNA is 47S. If snR17 is yeast U3, this discrepancy could be explained by assuming that the mammalian 47S pre-rRNA was degraded during the fractionation procedure. More recent observations suggest that human U3 snRNA indeed associates with larger pre-rRNA species (I. L. Stroke, and R. Maser, personal communications).

## Chapter 3

Analysis of the primary and secondary structure of snR17 and its genes.

#### SUMMARY

The nucleotide sequences of the *Saccharomyces cerevisiae SNR17A* and *SNR17B* genes and flanking regions were determined by the dideoxynucleotide chain-termination method. The 5' and 3' ends of snR17A were mapped by a combination of S1 nuclease protection studies and primer elongation; the length of snR17A was determined to be 328 nt. The 5' and 3' ends of snR17B were not mapped precisely, but snR17A and snR17B appear to be the same length, and their ends presumably map to homologous sites. The nucleotide sequences in the transcribed regions of the two *SNR17* genes are 96% matched, the first 100 bp of the 3' flanking sequences are 61% matched, but very little similarity was observed between the 5' flanking sequences. As in other *S. cerevisiae* snRNA genes, "TATA" sequences were found 80 to 100 bp upstream of the *SNR17A* and *SNR17B* transcription initiation sites.

Significant similarities were found between the primary structures of snR17 and U3 snRNAs, mainly in four distinct regions; most notable was an almost perfectly conserved pentadecamer nucleotide sequence close to the 5' ends. An energetically stable secondary structure model could be derived for snR17 which strongly resembled models for U3 snRNAs, the main difference in the yeast model being the presence of an extra hairpin stem. Two of the conserved regions of U3 snRNAs and snR17, boxes B and C, form single-stranded loops in the secondary structure models, and these have been proposed as sites available for base-pairing interactions between U3 snRNA and pre-rRNA: the hypothetical interaction of box B with the 3' external transcribed sequence of pre-rRNA appears to be phylogenetically conserved, and, in yeast, corresponds to a cleavage site beyond the 3' end of the mature 25S rRNA; the sequence in the second internal transcribed spacer of pre-rRNA proposed to be complementary to box C is conserved in vertebrates but not in yeast.

90

### METHODS USED FOR SEQUENCING AND MAPPING SMALL RNA GENES.

#### 1. Introduction

The nucleotide sequences of the *SNR17A* and *SNR17B* genes and their flanking regions were determined by the dideoxyribonucleotide chain termination method of Sanger *et al.* (1977). Protocols for DNA sequencing using this method have been published by Perbel (1984) and Eperon (1987). The RNAs were previously mapped approximately by probing restriction fragments with labelled RNAs (Chapter 2). Nested deletions spanning parts of the genes were generated with the aim of bringing every region of DNA targeted for sequencing ideally within about 200 bp of a site at which the sequencing primer could anneal. Priming was from the "universal" M13 primer (Messing 1983) on single stranded plasmid or M13 bacteriophage DNA templates, or from specially synthesised oligodeoxynucleotides.

The construction of the plasmid and bacteriophage clones used for sequencing the *SNR17A* and *SNR17B* genes is summarized in Figures 3.1 and 3.2. The strategies for sequencing the *SNR17A* and *SNR17B* genes are illustrated in Figure 3.5. The nucleotide sequences of the genes and flanking regions are presented in Figure 3.11.

The mapping and sequencing of snR30 and scR1 were performed by Marc Bally and Franco Felici, respectively. The sequences of the genes for these RNAs are presented in Appendices 1 and 2.

2. Methods for creating nested deletions.

Deletion clones were generated in three ways: 1. excising DNA fragments between suitably placed restriction sites; 2. controlled exonuclease Bal31 digestion from a suitable unique restriction site, cleavage at a second site to release a truncated, exonuclease-treated fragment, and recloning of the fragment into another vector; and 3. addition of *Eco*RI linkers at a single, random cut made by DNase I in the presence of  $Mn^{2+}$ , cutting in the vector polylinker sequence with *Eco*RI, thus deleting a fragment, and then recircularizing.

Conveniently placed restriction sites sometimes allowed a DNA fragment to be deleted from a plasmid, bringing regions targeted for sequencing into close proximity to the sequencing primer annealing site. Plasmid pSR3.1 was thus derived from pSR3b (Figure 3.1) by deleting a DNA fragment between two *Sall* sites, one in the vector polylinker sequence and the other in the *SNR17A* 5' flanking sequence, and recircularizing the plasmid by ligating the two *Sall* sticky ends. The likelihood of the fragment being reinserted was reduced by diluting the cleaved plasmid to a DNA concentration of less than 5 µg/ml before ligating.

This method was also applicable when multiple sites for a restriction enzyme were present: thus, plasmids pSR3B.1 and pSR3B.2 were derived from pSR3B (Figure 3.2) by partially digesting with *Sau*3A, for which there are several cleavage sites within the plasmid, digesting with *Bam*HI, which cuts only once within the vector polylinker sequence, size-fractionating to select those plasmids that had been cut only once by *Sau*3A at the desired site, and recircularizing the plasmid by ligating the complementary *Sau*3A and *Bam*HI sticky ends. This was achieved as follows:

First, a series of pilot digestions were carried out to determine the conditions in which the majority of plasmid was linearized with *Sau*3A. 50  $\mu$ l of pSR3B miniprep DNA (approx. 5  $\mu$ g, Chapter 6, Section 3.4) in 60  $\mu$ l total volume was then linearized with *Bam*Hl in 50mM salt buffer; 0.7 U (determined by trial) *Sau*3A were added, and the mixture was incubated for a further 15' at 37°C; agarose gel loading buffer was added, the sample was heated at 65°C for 5', loaded onto a 0.7% low melting agarose gel (BRL Ultrapure, with ethBr in Tris-OAc buffer), and electrophoresed over-night at 3 V/cm; gel slices (<200  $\mu$ l) from the size range appropriate for the desired deletions were mixed with 0.5 ml water, melted at 70°C for 10', and then placed at 37°C; a 200  $\mu$ l aliquot was mixed with 22  $\mu$ l 10 x ligase buffer, several units of T4 DNA ligase were added, and the mixture to 0.1M CaCl<sub>2</sub>, aliquots could be mixed directly with competent 71-18 cells for transformation.

Deleting with exonuclease Bal31.

The M13 bacteriophage clones QSR10 to QSR19 (Figure 3.1) were constructed as follows: A pilot experiment determined the quantity of Bal31 and the digestion time appropriate for the deletions required. Volume and quantity were then scaled up to the following reaction mixture in 250  $\mu$ l: 50  $\mu$ l (50  $\mu$ g) CsCI-purified pSR3.3 DNA in TE linearized with *Sal*I; 25  $\mu$ l 10 x Bal31 buffer

(2M NaCl, 0.12M CaCl<sub>2</sub>, 0.12M MgCl<sub>2</sub>, 0.1M Tris pH 8.0), 173 µl water and 2 µl Bal31 (BRL, 0.5 U/µl). The mixture was incubated at 30°C, 25 µl aliquots were removed at 2' intervals, immediately phenol/chloroform extracted and ethanol-precipitated. The DNA, redissolved in TE, was digested to completion with *Eco*Rl, and again phenol/chloroform extracted and ethanol-precipitated. Approximately 1 µg of redissolved *Eco*Rl-cut DNA from each aliquot was mixed with 1 µg *Eco*Rl/*Sma*l-cut M13mp18 RF DNA (Norrander *et al.* 1983) in 10µl ligase buffer (20mM Tris pH 7.6, 10mM MgCl<sub>2</sub>, 10mM DTT, 0.5mM ATP) and incubated with 1 U T4 DNA ligase at room temperature for several hrs. Competent 71-18 cells were transformed with the ligation mixture and plated onto X-gal/IPTG plates. White M13 "plaques" were isolated, those which contained bacteria resistant to ampicillin were discarded, and miniprep RF DNA was analysed for deletion size.

Deletions from random DNase I cuts.

DNase I in the presence of  $Mn^{2+}$  makes dsDNA cuts rather than nicks in one DNA strand only. Frischauf *et al.* (1980) exploited this for the construction of random deletion clones. The method as used to generate pSR3.2-pSR3.8 (Figure 3.1) is described below.

A pilot digestion determined the conditions required to linearize 60-70% supercoiled plasmid DNA in  $Mn^{2+}$  buffer. The 10 x scaled up digestion in 200 µl was as follows: 40 µg supercoiled pSR3.1 DNA, 2.5 ng/ml DNase I, 20mM Tris pH7.6, 1.5mM MnCl<sub>2</sub> (stored as 2 x buffer), incubated at 22°C for 20', then phenol/chloroform extracted and ethanol-precipitated. T4 DNA polymerase was used for end-polishing: DNA was redissolved in 76 µl TE, to which was added 10 µl 10 x T4 pol. buffer (0.33M Tris-OAc pH 7.9, 0.66M KOAc, 0.1M Mg(OAc)<sub>2</sub>, 5mM DTT, 1 mg/ml BSA), 10 µl all four dNTPs, each at 2.5mM, and 16 U T4 DNA polymerase (BRL). The mixture was incubated at 37°C for 1 hr, and phenol/chloroform extracted and ethanol-precipitated. 100 fold molar excess of (partially  $[\gamma^{-32}P]$  ATP) kinased *Eco*RI linkers was ligated to one half (20 µl) of the redissolved DNA, the over-night, room temperature reaction was heat inactivated, completely digested with EcoRI (as judged by the disappearance of the "linker ladder" on a polyacrylamide gel), and loaded onto a 0.7% low melting agarose gel. After electrophoresis, gel slices containing the appropriate sized deletions were cut out, plasmid DNA was recircularized in situ, and used to transform 71-18 as described in the previous section.





Various yeast DNA fragments cloned into polylinker sites of pEMBL9 (Dente *et al.* 1983) (pSR3 series) or M13mp18 (Norrander *et al.* 1983) (QSR series). pSR3.10 and pSR3b were obtained by sub-cloning from  $\lambda$ SR3 (a  $\lambda$ EMBL4 library clone, see Figures 2.2 and 2.3). Lengths of the yeast DNA insertions appear in parentheses after the name of each clone. Nucleotides representing deletion boundaries are numbered, with respect to the initiation site of transcription, in bold face. Methods for the construction of deletions are described in the text.





E = EcoRl, S = Sall, S3 = Sau3A, D = Dral, B = Ball, Bm = BamHI

Figure 3.2 SNR17B plasmid clones pSR3B is a sub-clone of  $\lambda$ SR3B in pEMBL8 (see Figure 2.3). Lengths of the yeast DNA insertions appear in parentheses after the name of each clone. Nucleotides representing deletion boundaries are numbered, with respect to the initiation site of transcription, in bold face. 3. Sequencing strategies.

The sequences of *SNR17A* and *SNR17B* were derived using the protocol in Chapter 6, Section 4 with the M13 universal primer and ssDNA templates of the clones shown in Figures 3.1 and 3.2. Additional nucleotide sequence was also obtained for *SNR17A* using an oligonucleotide 5'-TAGATTCAATTTCGGTTT-3' ("oligo 49") complementary to nucleotides +116 to +133 of the sequence in Figure 3.11 and for *SNR17B* using oligonucleotides 5'-TAAATTCGATTTCATTTC-3' ("oligo 50") complementary to nucleotides +116 to +133, and 5'-CATTCATGATTGTAATG-3' ("oligo 51"), complementary to nucleotides +428 to +444. Figure 3.5 illustrates the length of sequence derived from each deletion clone/oligonucleotide. Apart from a small section of *SNR17A*, the nucleotide sequences of the genes were determined from only one strand of the DNA: nucleotide sequences are normally only considered to be absolutely reliable when the sequence of both strands has been determined.



Restriction maps and sequencing strategies for SNR17A and SNR17B. Figure 3.5

Arrows indicate the length (5' to 3') of nucleotide sequence derived from each deletion clone or oligonucleotide. The deletion clones are illustrated in Figures 3.1 and 3.2. The oligonucleotides used were oligo 49 and oligo 50, complementary to nucleotides +116 to +113 of snR17A and snR17B, respectively, and oligo 51, complementary to nucleotides 428 to 444 in the 3' flanking region of SNR17B, (numbering in Figure 3.11). 4. Mapping of the genes.

At this stage in the investigation it was not clear whether both the *SNR17* genes were transcribed. This question was resolved by synthesising the two oligonucleotides mentioned above, oligo 49 and oligo 50, complementary to snR17A and snR17B, respectively, in a region where there are five out of eighteen nucleotide mismatches between the genes. When labelled and used as probes on filters of total yeast RNA, both oligonucleotides hybridized to RNA species of exactly the same size at stringencies which did not allow non-specific annealing. The fact that both genes were transcribed was confirmed later when one or other of the genes were deleted and it could be shown that the transcript from the deleted gene was absent whereas the one from the intact gene was still present. These results are presented in Chapter 4.

SnR17A was mapped by a combination of nuclease S1 protection experiments and primer-extension using reverse transcriptase: the 5' end of the gene was first assigned to a position by S1 nuclease protection mapping; the 3' end was determined relative to this by measuring the length of the complete transcript; the position of the 5' end was subsequently modified on the basis of primer-extension data.

Using the Klenow fragment of DNA polymerase I to extend the M13 universal primer with  $\left[\alpha-3^{32}P\right]$ dNTPs, a series of single-strand labelled probes were generated, complementary to the snR17A RNA, from single-stranded template DNA derived from plasmids pSR3.1 to pSR3.7 (Figure 3.1), which represent a set of nested deletions spanning the region to which the SNR17A gene had been roughly mapped. The nucleotide sequence of this region and the exact boundary of each deletion had already been determined. By annealing the probes to the RNA and analysing the pattern of S1 nuclease-protected fragments, those deletions occurring within the transcribed region of the gene could be identified, and the transcription initiation site could be mapped to a position the same distance from the known deletion boundaries as the lengths of the protected fragments. Probes derived from clones containing the full length of the transcribed sequence gave rise to protected fragments corresponding to the full length of the RNA: the position on the gene corresponding to the 3' end of the RNA was mapped relative to the 5' end by measuring the length of the full-length protected fragments.

Figure 3.7 shows full length S1 nuclease-protected fragments (lanes 1 to 5) and a partial length protected fragment (lane 6) with the nucleotide sequence of pEMBL9 as size marker. Single-strand labelled probes complementary to the

RNA were generated from pSR3.1, pSR3.2, pSR3.3, pSR3.4, pSR3.5, (full length probes, lanes 1 to 5), pSR3.6 and pSR3.7 (partial length probes, lanes 6 and 7). The length of the pSR3.6 protected fragment, 254 nucleotides, would indicate the 5' end of the RNA to be this distance from the boundary of the pSR3.6 deletion, that is, at the G at position 17 in Figure 3.11. The full length protected fragments (lanes 4 and 5) correspond to a size of 312 nt, fixing the 3' end (relative to position 17) as indicated in Figure 3.11.

Primer-extension with reverse transcriptase was performed by using 5'  $[^{32}P]$  end-labelled oligo 49 (complementary to nucleotides +116 to +133 of the sequence in Figure 3.11) to prime a reverse transcript on snR17A as template. The length of the reverse transcript indicated that the 5' end was 16 nt further upstream from the position determined by the S1 nuclease-protection studies. The 5' end of the gene shown in Figure 3.11 is that determined by reverse transcription. The originally designated position of the 3' end was not adjusted because, when, in later experiments, the total length of the end labelled snR17 RNA was measured, it was found to be approximately 330 nt long, consistent with the length of 328 nt between the designated 5' and 3' ends, indicating that the full length S1 nuclease protected fragments (312 nt), as well as the partial length (pSR3.6) fragment, were truncated at their 5' ends.

Figure 3.8 shows the length of the reverse transcript primed from endlabelled oligo 49 on snR17A template, relative to a dideoxy-sequence ladder primed with the same oligonucleotide on a single-stranded DNA template. Two bands are visible: a lower, stronger band and an upper weaker band. The transcription start site has been assigned to the upper, fainter band, which comigrates with the final T of the sequence ...TGTGTI, thus indicating the first A of the complementary sequence <u>AACACA</u>... to be the first transcribed nucleotide.

An attempt was made to map *SNR17B* by nuclease S1 protection, using the same method described for *SNR17A*. The length of a protected, truncated probe complementary to the 5' portion of the RNA was measured, as well as the length of a probe complementary to the whole length of the RNA. Interestingly, the position of the 5' end of snR17B, indicated by the strongest protected fragment, corresponded to the G at position 17, equivalent to that originally determined by S1 protection to be the 5' end of snR17A. The 5' end of snR17B was not confirmed by reverse transcription. Since the two RNAs were observed on gels to be the same length, the 5' and 3' ends were assumed to be at equivalent positions.

### Figure 3.7 Mapping of snR17A by S1 nuclease-protection.

Single strand-labelled probes were generated from plasmids pSR3.1-7 (see Figure 3.1), annealed to RNA and treated with S1 nuclease<sup>1</sup> as follows: on ssDNA templates, chain elongation reactions were performed with  $[\alpha$ -<sup>32</sup>PldATP and dATPmix (but no ddATP), as described in Chapter 6, Section 4.2. The probes were purified over sepharose G-50, and one third of each was mixed with 100 µg total yeast RNA, precipitated with ethanol and redissolved in 30 µl S1 hybridization buffer (40mM PIPES pH 6.4, 1mM EDTA, 0.4mM NaCl, 80% formamide), denatured at 85°C for 15', then transferred immediately to 52°C and incubated overnight; the mixture was transferred to 37°C and, simultaneously, 0.3 ml ice-cold S1 buffer (0.29M NaCl, 50mM NaOAc pH 4.6, 4.5mM ZnSO<sub>4</sub>) containing 290 U S1 nuclease (BRL) was added. After incubation for 30' at 37°C, 100 µl 2M NH<sub>4</sub>OAc, 50mM EDTA and 400 µl isopropanol were added and the nucleic acid precipitated at -20°C. The samples were redissolved in 80% formamide dyes and resolved on a sequencing gel. Fragments derived from plasmids pSR3.1 to pSR3.7 are shown in lanes 1 to 7, respectively, parallel to a sequence ladder (G, A, T, C) primed on pEMBL9 with the M13 universal primer. Probes spanning the whole length of the transcribed sequence gave protected fragments corresponding in size to a C fragment (312 nt) in the sequence lane (underlined), marked with the upper asterisk. pSR3.6 (lane 6) gave a shorter fragment (indicating that the boundary of the deletion of this clone was within the RNA coding sequence) corresponding in size to a T fragment (254 nt) in the sequence lane, marked with the lower asterisk.

<sup>&</sup>lt;sup>1</sup>Methods for mapping with S1 nuclease have been described by Casey and Davidson (1977), Berk and Sharp (1977), Maniatis *et al.* (1982) pp207-209, Calzone *et al.* (1987), Reyes and Wallace (1987).



### Figure 3.8 Mapping the 5' end of snR17A by primer-extension.

Oligonucleotide 49 (complementary to nucleotides +116 to +113 of the RNA in Figure 3.11) was used to prime a reverse-transcript<sup>1</sup> on snR17A template as follows:

100 µg total yeast RNA was denatured in 9 volumes of DMSO at 45°C for 20'. ethanol-precipitated after bringing to 0.3M NaOAc, redissolved in 5 µl 5 x reverse transcriptase buffer (5 vol. 1M Tris pH 8, 5 vol. 0.1M MgCl<sub>2</sub>, 5 vol. 1M KCl, 1 vol. 0.1M DTT) to which 0.1 pmol 5' end-labelled oligonucleotide, and water up to 25 µl final volume was added. The mixture was heated at 80°C for 5', allowed to cool in a beaker of water to 42°C, and 2.5 µl dNTP mix (each dNTP at 2.5mM), 1 µl RNasin (Promega Biotec, 40 U/µl) and 8 U reverse transcriptase were added. The reaction was allowed to proceed for 2 hrs at 42°C, RNA was then hydrolysed by adding 2 µl 5N NaOH and incubating at 56°C for 10', neutralized by addition of 2  $\mu$ l of 1M Tris pH8, 9.5  $\mu$ l 1N HCl, then extracted with phenol/chloroform and precipitated with ethanol. Washed and dried pellets were redissolved in formamide dyes and analysed on sequence gels. The figure shows the reverse transcript run in parallel with a dideoxy-sequence ladder primed with the same oligonucleotide on a single-stranded SNR17A template. Of the two bands visible in the reverse transcript lane, the upper, fainter band has been used to designate the 5' end of the gene in Figure 3.11.



#### ANALYSIS OF PRIMARY AND SECONDARY STRUCTURE.

#### 1. Structure of the SNR17 genes.

The nucleotide sequences of the *SNR17A* and *SNR17B* genes are shown in Figure 3.11. The two sequences were compared and aligned using the UWGCG "Gap" programme (Devereux *et al.* 1984). The coding sequence of *SNR17A* is aligned to an almost identical sequence in *SNR17B* with 96% matching. The 5' flanking sequences of *SNR17A* and *SNR17B* show little similarity either with each other, or with any other snRNA genes: however, both contain long stretches of (dT)s and (dA)s, from -43 to -25 in *SNR17A* and from -147 to -131 in *SNR17B*. [Poly(dA-dT) sequences are a feature of constitutively expressed yeast genes transcribed by RNA polymerase II (Struhl 1985).] "TATA" sequences have been marked at -91 (*SNR17A*) and -85 (*SNR17B*). Downstream, similarity between the two sequences (60% matching with two gaps) persists for a further 100 bp beyond the 3' ends.

#### 2. The primary structure of snR17 is similar to U3 snRNA.

A search was made in the EMBL and Genbank nucleotide sequence data libraries for sequences with similarity to snR17A using the UWGCG programme "Wordsearch" (Devereux *et al.* 1984). This programme assesses similarity between sequences by matching up oligonucleotide "words" of definable length. When the search was carried out using a relatively long word length, a 15 nt perfect match between a region close to the 5' end of snR17 and a region close to the 5' end of mammalian U3 snRNAs was identified; this prompted a closer look at the relationship between these molecules.

Figure 3.12 shows an alignment of the two snR17 sequences with the U3 snRNA sequences of human (Suh *et al.* 1986), rat (Reddy 1985), frog *Xenopus laevis* (Jeppesen *et al.* 1988), and slime mould *Dictyostelium discoideum* (Wise and Weiner 1980). The sequences were aligned using the multiple sequence alignment method "Trials" (Hogeweg and Hesper 1984, Konings *et al.* 1987).<sup>1</sup> The method makes the assumption that any attempt to align more than two molecular sequences implies an attempt to define hypothetical ancestral relationships, and that no attempt to do the first is meaningful without consideration of the second. The procedure of alignment is as follows:

1. Each sequence is aligned independently as a pair with each other using an algorithm common to several alignment methods that allows penalization of gap formation and of gap length.

<sup>1</sup>The snR17 and U3 snRNA sequences were aligned with "Trials" by Danielle Konings.

	-320	-300	-280	-260	-240	
SNR17B	B ACTACTACGATATGTATTTCGAGCATTTTTCTTAGGGCTGAATTTTTCATTAGTTTTTCCATTCTCGATGAGGTGAATGAGTAAAGTTAAAGAAACATAT					
	-220	-200	-180	-160	-140	
SNR17B	AAATAGTGAGCTCTTTCC	CAAACAACTTGATAACTTCAC	TCATTCATCCTTTTCTTAG	ITTTATICTCITCITGITCIT	TTTACTAAAAAAATTGTCGACG	
	-120	-100	-80	-60	-40	
SNR17A	ATTITACCCACAAAAAAACCAGGATTTGAAAAAC <u>TATA</u> GCACCTAAAAGTCTTAGGTACTAGAGTTTTCATTTCGGAGCAGGTTTTTTGAAAAAATTAAATTC					
SNR17B	TACTTCAGTATGTTTTATACCATATACTTTATTAGGAA <u>TATAA</u> CAAAGCATACCCAATAATTAGGCAATGCGATTGTCGTATTCAACAACCATCTTCTAT					
	-20	+1	20	40	60	
snR17A SNR17A	AACCATTGCAGCAGCTT	3mGpppAACACAUUCUACAGU FTGACTAACACATTCTACAGI 	JAGGAUCAUUUCUAUAGGAA AGGATCATTTCTATAGGAA 	JCGUCACUCUUUGACUCUUCA ICGICACTCTTTGACTCTTCA	AAAGAGCCACUGAAUCCAACUU AAAGAGCCACTGAATCCAACTT 1	
SNR17B	TTCACCAGCTTCAGGTT.	ITGACTAACACATTCAACAGI	AGGATCATTTCTATAGGAA	ICGICACTCTITGACTCTTCA	AAAGAGCCACTGAATCCAACTT	
snR17B	:	3mGpppAACACAUUCAACAGU	JAGGAUCAUUUCUAUAGGAAU	UCGUCACUCUUUGACUCUUCA	AAAGAGCCACUGAAUCCAACUU	
	80	100	120	140	160	
SNRI /A	GGUUGAUGAGUCCCAUA		SAAACCGAAAUUGAAUCUAA	AUUAGCUUGGUCCGCAAUCCU	UAGEGGUUEGGECAUEUAUAAU	
SNR1 /A	GGTIGAIGAGICCCATA	ACCITIGIACCCCAGAGIGAG	AAACCGAAATIGAATCTAA	ATTAGCTTGGTCCGCAATCCT	TAGCGGTTCGGCCATCTATAAT	
SNR17B	GGTTGATGAGCCCCATA	ACCITIGIACCCCAGAG <u>IG</u> AC	GAAATGAAATCGAATTTAA	ATTAGCTTGGTCCGCAATCCT	TAGCGGTTCGGCCATCTATAAT	
snR17B	GGUUGAUGAGCCCCAUAACCUUUGUACCCCAGAGUGAGGAAAUGAAAUCGAAUUUAAAUUAGCUUGGUCCGCAAUCCUUAGCGGUUCGGCCAUCUAUAAU					
			•			
	180	200	220	240	260	
an D173						
SIRI /A			JUUUUCCUUUGGAAGUAAUU		AUCUUGACCCAUCCUAUGUACU	
SNRI /A	TTTGAATAAAAATTTTCCTTTGCCGTTGCATTTGTAGTTTTTTCCTTTGGAAGTAATTACAATATTTTATGGCGCGGATGATCTTGACCCATCCTATGTACT					
			1			
SNR17B	TITIGAATAAAATIFIGC.	TTTGAATAAAATTTTGCTTTGCCGTTGCAATTGTGATAACTTCCTTTGGAATTTGGTGCGATATTTAATGGCGCGATGATCTTGACCCATCCTATGTGCC				
snR17B	17/B UUUGAAUAAAAUUUUGCUUUGCCGUUGCAAUUGUGAUAACUUCCUUUGGAAUUUGUUGCGAUAUUUAAUGGCGCGAUGAUCUUGACCCAUCCUAUGUGCU					
	280	300	320	340	360	
	• •	• •	• •	• •		
snR17A	UCUUUUUUGAAGGGAUA	GGCUCUAUGGGUGGGUACAA	AUGGCAGUCUGAC			
SNR17A	TCTTTTTTGAAGGGATAC	GGCTCTATGGGTGGGTACAA	ATGGCAGICTGACAAGTTA	ACCACTTTTTTCCTTTTC	TAAATTGTTTAAAACCAAAGGT	
					8191 INA IN INTINI	
SNR17B	TCCATTTTGGAAGGATAGGGATTCATGGGTGGGTACAAATGGCAGTCTGACAAGTCAAATCATTTATTT					
snR17B	UCCAUJUJUGGAAGGAUA	GGAUUCAUGGGUGGGUACAA	AUGGCAGUCUGAC			
	380	400	420	440	460	
	• •	· ·		• •	• •	
SNR17A	TTGGTTTTCAGTTAAGAAATTGGATTAGTTGGTGTATAAGTATAATTAAATGTAGTCTTGGTTAAGCTTAATGTATAGGTTCTTGATAGACTCCA					
SNR17B	3 TIGATTCAGTATTTTGGAGCCGCATTATACTTAGTATTTAATTAA					
	480	500	520	540	560	
				<b>-</b>		
SNR17A	CCTAAAATTCCTCAACAC	CGACGGATGCGAAAACAAA	AAATCGCTTTGCATCCTTA		·	
~~~~~ //2	CONTRACT COLOMON		**************************************			

#### STRUCTURE OF SNR17A AND SNR17B

SNR178 GTTTTCAAGAGATCGAGACACCTTCTTTAAGATATATAAATGGTTTGCTACTAGCATAGAGGGATCTTCACTAGGCAGCGATGACCACGTTACATAACACG

# Figure 3.11 Nucleotide sequences of *SNR17A* and *SNR17B*.

The SNR17A and SNR17B sequences are 96% matched in their transcribed regions. TATA sequences between 80 and 90 nucleotides upstream have been underlined. These nucleotide sequences were determined from one DNA strand only: nucleotide sequences are normally only considered to be absolutely reliable after sequenceing both strands.

#### Figure 3.12 Alignment of U3 snRNA sequences with snR17A and snR17B.

Possible alignments of human (Suh *et al.* 1986), rat (Reddy 1985), frog *Xenopus laevis* (Jeppesen *et al.* 1988), and slime mould *Dictyostelium discoideum* (Wise and Weiner 1980) U3 snRNA and snR17 sequences were investigated using the multiple sequence alignment method "Trials" (Hogeweg and Hesper 1984, Konings *et al.* 1987). The penalty for gap length was set at zero in order to accommodate reasonably the 50% greater length of the yeast sequences, but the penalty for gap formation was varied. In the alignment shown, the gap formation penalty was 3. Boxes A, B, C and D are conserved regions of nucleotide sequence. Features of the secondary structure models in Figure 3.13 have been marked: hairpins 1, 2, 3 and 4 are indicated by lines above and below the sequences, and the two complementary sides of the **central stem** have been underlined. [N.B: The sequences were aligned without considering possible secondary structure conformations.]

The computational runs to generate the alignments were performed by Danielle Konings.
#### **U3 snRNA SEQUENCE ALIGNMENT**

	10 <b>BOX A</b> 30	50	70	90
	H <u>AIRPIN 1</u>		<u> </u>	·
	AAGACUAUACUUUCAG GGAUCAUUUCUAUAG UGUGUUACUAG	GAGAAGUUUCUCUGAACGUG	UAGAGCACCGAAAACCACGAGG	AAGAGAGG
Rat 113A	AAGACUAUACUUUCAG GGAUCAUUUCUAUAG U UCGUUACUAC	SAGAAGUUUCUCUGACUGUG	UAGAGCACCCGAAACCACG. AG	TI TI SAGG AGACG
Rat U3B	AAGACUAUACUUUCAG GGAUCAUUUCUAUAG UUCGUUACUAC	SAGAAGUUUCUCUGACUGUG	UAGAGCACCCGAAACCACG.AGC	GACG.AGACA
X. laevis	AAGACUAUACUUUCAG GGAUCAUUUCUAUAG GUUGUACCUAG	JUGAAAUGUGCUCGAAAGUG	UCUGAACUCACAAACCACGAGGA	AAGAGCGA
D. discoideum D2	AUGACCAAACUCU.UA GGAUCAUUUCUAGAG UA.UCGUCUAUU	AAAAUUAUUCAUCAAUAA		GAUGAUGAUAC
Yeast snR17A	AACACAUUCUACAGUA GGAUCAUUUCUAUAG GAAUCGUCACUCU	JUUGACUCUUCAAAA.	GAGCCACUGAAUCCAACUUG	GUUGAUGAGUCCCAU
Yeast snR17B	AACA <u>CAUUCAACAGUA GGAUCAUUUCUAUAG</u> GAAUCGUCACUCU	UUGACUCUUCAAAA	GAGCCACUGAAUCCAACUUG	GUUGAUGAGCCCCAU
	HAIRPIN	11		
	CENTRAL 110 BOX B 130	150	170	190
	STEM HAIR	PIN 2	170	
Human	UAGCGUUUUCUCCU, GAGCGUGAAGCCG GCUUUCUGGCGUUGCU			
I ACARCILLE I				
Rat U3A	UAGCGUUCUCCU. GAGCGUGAAGCCG GCUCUUAGUGUUGCUU	JC		
		Hereit States - States		
Rat U3B	UAGCGUCCCCCCUCCU. GAGCGUGAAGCCG GCUCUAGGUGCUGCUU	JC		• • • • • • • • • • • • • • • • • •
X.laevis	CA <u>GUGUUUUCUCC</u> U. GAGCGUGAAGUGA GCUCACAGUGCU		•••••••••••••••••••••••••••••••••••••••	
D discolderm D?	AUACUCACUAUACGA AACCGUGAAACCG UUAUUAUCGAAUGAU	IC		
D. disconcella De				
Yeast snR17A	AACCUUUGUACCCCA GAGUGAGAAACCG AAAUUGAAUCUAAAUU	JAGCUUGGUCCGCAAUCCUU	AGCGGUUCGGCCAUCUAUAAUUU	JUGAAUAAAAUUUU
			THEFT HERE THE THE PARTY OF THE	INTERNETICE.
Yeast snR17B	AACCUUUGUACCCCA GAGUGAGGAAAUG AAAUCGAAUUUAAAUU	JAGCUUGGUCCGCAAUCCUU	AGCGGUUCGGCCAUCUAUAAUUU	<u>UGAAUAAAAUUU</u> UG
		HAIRPIN	14	
	210 220	250 801	<b>c</b> 270	200
	210 230	250 <b>BOX</b>	C 270	290
	210 230 			290 IN 3 .
	210 230 HAI <u>RPIN 2</u> UGGCUGCAACUG	250 BOX	C 270 	290 IN 3
Rat U3A	210 230 HAI <u>RPIN 2</u> UGCCUGCAACU         .CGUGCAACU	250 BOX	C 270 	290 IN 3
Rat U3A	210 230 HAI <u>RPIN 2</u> UGGCUGCAACU IIIIII 	250 BOX	C 270 	290 IN 3 .
Rat U3A Rat U3B	210 230 HAIRPIN 2 UGGCUGCAACU	250 BOX	C 270 HAIRP	290 IN 3
Rat U3A Rat U3B	210 230 HAI <u>RPIN 2</u> 	250 BOX	C 270 . HAIRP	290 IN 3 .
Rat U3A Rat U3B X.laevis	210 230 HAI <u>RPIN 2</u> UGGCUGCAACU IIIIII 	250 BOX	C 270 . HAIRP ICCU UCUUCUCCUCUAUU III III I III ICCU UCU.CGGUCCUCUCU III III III III ICCU UCUCUCUCU.CU.UC III III III III ICCU UCUCCUCCUCU.UAUUAU	290 IN 3 .
Rat U3A Rat U3B X.laevis D. discoideum D2	210 230 HAI <u>RPIN 2</u>	250 BOX	C 270 . HAIRP ICCU UCUUCUCCGUAUU III III I III ICCU UCUUCUCCUCUCU III III I IIII II ICCU UCUUCUCCUCU.UC III III III III ICCU UCUUCUCCCCUUUAUUAU III II ICCU UCUGCUCCCCUUUAUUAU III II ICCU UCUGCUCCCCUUUAUUAU	290 IN 3 .
Rat U3A Rat U3B X.laevis D. discoideum D2	210 230 HAI <u>RPIN 2</u> 	250 BOX	C 270 	290 IN 3 .
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A	210 230 HAIRPIN 2 UGGCUGCAACU IIIIII CGUGCAGCU IIIIII UGUGCAGCU IIIII UGUGCAGCU IIIII UGUGCAGCU IIIII CGUUCAUUGUGGU III CUUUGCCGUUGCAUUUGUAGUUUUUUCCUUUGGAAGUAUUACAA IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 . HAIRP ICCU UCUUCUCCGUAUU III III IIII ICCU UCU.CGGUCCUCUCU III III IIIIII ICCU UCUCUCUCU.U.C III III IIIIIII III IIIIIIIIII CCU UCUCUCUCUCU.U.C III IIIIIIIIIIIIII CUU GACCCAUCCUAUGUACUU GACCAUCCUAUGUACUU	290 IN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAI <u>RPIN 2</u> UGGCUGCAACUA IIIIII CGUGCAACUA IIIIIIIIII CGUGCAACUA IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	250 BOX	C 270 	290 IN 3 UU
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAI <u>RPIN 2</u> 	250 BOX	C 270 	290 IN 3 U CUUUUUUGAAGGGAU I IIIII IIII CAUUUUGGAAGGAU RPIN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis <i>D. discoideum</i> D2 Yeast snR17A Yeast snR17B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Bat U3B	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B	210 230 HAIRPIN 2 GUGCUGCAACUA HAIRPIN 2 GUGCUGCAACUA HIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	250 BOX	С 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis D. discoideum D2	210 230 HAIRPIN 2 	250 BOX	С 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis D. discoideum D2	210 230 HAIRPIN 2 	250 BOX	С 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A	210 230 HAIRPIN 2 	250 BOX	C 270 	290 IN 3
Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A Yeast snR17B Human Rat U3A Rat U3B X.laevis D. discoideum D2 Yeast snR17A	210 230 HAIRPIN 2 	250 BOX	С 270 	290 IN 3

- 2. The alignment of each pair is assigned a match value that is high when a minimum number of mutations must be assumed to account for the divergence of the pair.
- 3. On the basis of the match values, a phyletic tree is constructed in which the most similar sequences are linked by branches high up in the tree (twigs) and the most dissimilar by branches low down (boughs), thus the distance between branch points represents the number of mutations assumed to have had to have occurred to explain the divergence of the various sequences.
- 4. The sequences are realigned following the branching pattern of the tree: the most similar sequences are aligned first, and the resulting consensus or "internode" sequence is aligned to the next most similar sequence; gap formation and length is penalized as before, gaps are treated as fifth characters, and a modified version of the tree is generated on the basis of new match values.
- 5. The procedure iterates until the tree and the alignment stabilize.

The method is restricted in that, once the first tree is constructed, the branching pattern remains the same, although distances between branch points may vary.

The alignments highlighted four regions of sequence conservation: boxes A, B, C and D. Whereas boxes A, B and C have already been noted as being conserved between rat and slime mould (Wise and Weiner 1980), box D, with the consensus GCAGUCUGA, at the 3' ends has not previously been emphasized. The overall similarity is greatest within the first 120 nt: in this region, given the gapping strategy shown, snR17 is 46% matched, and slime mould is 54% matched to human U3.

3. Secondary structure of snR17 and U3 snRNA.

Given the considerable similarity in nucleotide sequence between snR17 and U3 snRNAs, a comparison was made between the predicted secondary structures of these molecules. A secondary structure model for snR17, derived from a study of minimal free energy conformations, was found to have significant features in common with a consensus minimal free energy model derived for U3 snRNAs.<sup>1</sup>

Potential secondary structures were studied using the dynamic programming algorithm described by Williams and Tinoco (1986). Current

<sup>&</sup>lt;sup>1</sup>The minimal free energy models for snR17 and U3 snRNAs were derived by Danielle Konings at EMBL and as a guest of Arthur Williams (Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, CA 94720, USA).

understanding of the thermodynamic rules governing RNA folding is insufficient to predict real biological structures reliably: reliability is dependent on the accuracy of thermodynamic data that describe the free energies of structural features, the folding rules used by an algorithm to find a lowest free energy structure, and the stabilizing effects of environmental conditions on different structures of equal or higher free energy. It is important, therefore, to be able to compare a range of optional structures. Dynamic programming algorithms that predict RNA secondary structure are usually limited in being able to predict only one optimal structure; several runs applying different constraints being necessary to generate other structures of similar or lesser stability. The method of Williams and Tinoco is able to predict a definable number of optional structures from a single computational run: these can then be evaluated according to other available data such as consensus folding, phylogenic or biochemical studies.

The structures shown in Figure 3.13 were derived as follows: Various optimal structures generated for human, rat, *Xenopus* and *Dictyostelium* U3 snRNAs were compared, and a uniform structure was derived, consistent for all four. This consensus structure, represented by rat U3B RNA in Figure 3.13, was used to judge the various optimal yeast structures consistent with both snR17A and snR17B; the model for snR17 in Figure 3.13 was chosen because it showed the most similar conformation of the four conserved sequence boxes.

Secondary structure models for U3 snRNA have already been proposed by studying the patterns of compensatory base changes preserving stem structures between different species (Bernstein *et al.* 1983, Stroke and Weiner 1985). The model for rat in Figure 3.13 is quite similar to these: the very stable "central stem", the "hairpin 1" and the "hairpin 3" are the same as the "stem E", "stem A" and the "stem C", respectively, of the structure of Bernstein *et al.* (1983). The "hairpin 2" is essentially the same as that of Stroke and Weiner (1985) except that it has been extended basally, thus disrupting their "stem D" and opening out a loop containing the conserved box B; the structure in Figure 3.13 at this point has a lower free energy, and furthermore, the suggested stems D for *Dictyostelium* and rat (and bean, for the partial sequence available, Kiss *et al.* 1985,) are not formed by homologous nucleotides. The rat model in Figure 3.13 is the same as a model derived independently for human U3, based on an analysis of sites susceptible to chemical modification and RNase cleavage (Parker and Steitz 1987).

#### Figure 3.13 Secondary structure models for U3 snRNA and snR17.

The rat U3B model represents the consensus for all other U3 snRNAs. The yeast snR17A model is shown; differences in snR17B are indicated by smaller characters and arrows. Boxes A, B, C and D are the conserved regions illustrated in Figure 3.12. The central stem is marked with a line. The sites of possible, additional, complementary base-pairing interactions (variable between species) between nucleotides of box B and box C are indicated with dots. The structures were analysed by the method of Williams and Tinoco (1986) and those illustrated are among the most stable using both the energy values of Freier *et al.* (1986) and those of Cech *et al.* (1983).

The secondary structures were generated and evaluated by Danielle Konings.



The rat and yeast models in Figure 3.13 are quite similar: the striking difference between them is the extra hairpin (hairpin 4) of the yeast structure between the conserved box B and hairpin 2. It is significant that stable secondary structure models for U3 snRNAs and snR17 can be derived in which blocks of conserved nucleotide sequence can take up the same conformations. In both models the highly conserved box A forms part of hairpin 1; boxes B and C are loops (although additional base pairing interactions between these regions are theoretically possible, as indicated by dots in Figure 3.13, but these potential interactions are between different bases in different species); and Box D forms the 3' end. Box D is partially base paired to nucleotides 76 to 95 in the rat U3B model: there is experimental evidence that the 3' end of rat U3B is hydrogen bonded to this region in so far as reverse transcription can be primed from the 3' terminus with the complementary side of the stem acting as template; it has been suggested that this process occurring naturally in vivo could explain the existence of 3'-truncated pseudogenes (Bernstein et al. 1983). Box D of snR17A could form a similar interaction by the pairing of the terminal GAC (nucleotides 326 to 328) with CUG (nucleotides 89 to 87), but the U to C base change at 88 would disrupt this in snR17B, therefore the yeast box D has been left open. Most of the extra nucleotides of snR17 are taken up in the additional hairpin 4: an alternative structure for this, of equal stability, is also shown in Figure 3.13.

The representation of boxes B and C in the models as non-base paired loops suggests they might be available for base pairing with other RNA molecules or for binding to proteins. The fact that box D and the highly conserved box A apparently form parts of stems, the complementary sides of which are not well conserved, indicates that the reason they are conserved is not simply for the purpose of maintaining the stems; the rate of change in nucleotide sequence in structurally important stems is mainly constrained by the rate at which compensatory base changes can occur (Noller *et al.* 1981, Woese *et al.* 1983). Moreover, although the hairpins 1 of yeast and rat are quite stable, that of *Dictyostelium* is much less so. Parker and Steitz (1987) and Jeppesen *et al.* (1988) have observed that nucleotides in this domain are more accessible to chemical modification than in other stems, suggesting that hairpin 1 may not be base-paired as represented in Figure 3.13. It would seem, therefore, that more complex factors determine the structure of hairpin 1 than can be predicted by minimal free energy folding.

The abundant mammalian snRNAs are all found as RNP (ribonucleoprotein) particles consisting of a core set of proteins that they share in common, along with some extra proteins specific to individual particles (reviewed by Lührmann

112

1988). Patients suffering from the autoimmune disease lupus erythematosus have been found to have circulating antibodies directed against the core snRNP proteins. Such antibodies are known as "Sm" antibodies, and the proteins with which they react are known as "Sm" antigens. The binding site for Sm proteins on snRNAs has been defined as a short, single-stranded oligonucleotide sequence with the consensus  $RA(U)_{4-6}GR$  (R = purine) (reviewed by Mattaj 1988). A sequence conforming to this consensus, AAUUUUGA, is present in both snR17A and snR17B (nucleotides 175-182). U3 snRNAs neither contain Sm binding site sequences nor have been observed to be precipitable with Sm antisera (Lerner and Steitz 1979; Lerner *et al.* 1980a & b). In the snR17 model in Figure 3.13, this sequence occurs on hairpin 4, the hairpin which is absent from U3 snRNAs. Characterized Sm binding sites of other snRNAs occur in apparently single-stranded regions (Branlant *et al.* 1982, Liautard *et al.* 1982, Jacob *et al.* 1984, Guthrie and Patterson 1988, Mattaj 1988): the putative snR17 Sm binding site, however, is mainly base-paired as part of a stem.

#### DISCUSSION

This chapter has described the sequencing and mapping of the small nuclear RNA snR17 and its two genes. SnR17 was determined to be 328 nt in length. Analysis of the nucleotide sequences revealed that the transcribed regions of *SNR17A* and *SNR17B* were almost identical, with 96% matching, and significant similarities were found between these sequences and the sequences of U3 snRNAs of other organisms. The fact that stable secondary structure models resembling U3 snRNAs could be derived for both the snR17 species supports the hypothesis that snR17 could be functionally equivalent to U3 snRNA.

#### 1. Transcriptional signals of yeast snRNA genes

SnRNAs that possess 5' trimethylguanosine (m<sub>3</sub>G) cap structures are transcribed by RNA polymerase II: they acquire a monomethyl cap in the nucleus, like mRNA, but then become trimethylated after export into the cytoplasm. before returning again to the nucleus in their mature form (Zieve et al. 1988, reviewed by Dahlberg and Lund 1988, Mattaj 1988). Vertebrate snRNA gene transcription, however, does not proceed by exactly the same mechanism as mRNA gene transcription: vertebrate snRNA genes do not posses canonical "TATA" boxes (although a region in the 5' flanking sequence serves the same function, that is, defining accurately the point of transcription initiation), nor are snRNA and mRNA gene promoters interchangeable (Hernandez and Weiner 1986, Neuman de Vegvar et al. 1986, Dahlberg and Schenborn 1988). It has been suggested that the differences between snRNA and mRNA polymerase II promoters could be due to requirements for very high rates of transcription from snRNA genes, for independent transcriptional control mechanisms between snRNAs and mRNAs, or for a requirement for coupling of the transcription termination mechanism of snRNAs with the mechanism for export of pre-snRNAs into the cytoplasm (Dahlberg and Lund 1988). Contrary to the case in vertebrates, S. cerevisiae snRNA genes appear to have similar promoters to S. cerevisiae mRNA genes in that they possess TATA elements up to 100 bp upstream of their transcription initiation sites. Yeast TATA elements occur at variable distances (-40 to -120) upstream of mRNA initiation sites, and accurate initiation is probably determined by nucleotides within the first 15 upstream bp (Chen and Struhl 1985, Struhl 1987).

Figure 3.19 compares the upstream sequences of other published *S*. *cerevisiae SNR* genes, including *SNR30* (Bally *et al.* 1988), the cloning of which was described in Chapter 2 (the *SNR30* sequence is presented in Appendix 1), and shows that *SNR17* and *SNR30* share features in common with these. TATA

elements occur between -100 and -80, most of which finish with ...AAAG, either directly following the TATA, or separated by a few nucleotides. Transcription initiation is at "A", immediately preceded by a pyrimidine. The non-spliceosome species *SNR30*, *SNR3*, *SNR10*, *SNR17A* and *SNR17B*, more specifically, initiate with "AAC..", and initiation sites of *SNR30*, *SNR3* and *SNR10* are all preceded by (A)<sub>n</sub>GT. Initiation sites of spliceosome *SNR* genes are preceded by mainly "C"s up to -4 or -5, mainly "A"s up to -8 or -9, and "T"-rich tracts from -10 and beyond. SnR17, snR10 and snR3 are associated with pre-rRNA (Figure 2.10, Tollervey 1987a). The sequence specificities around the transcription initiation sites of yeast snRNA genes suggest two categories of genes that delineate the two different areas of proposed functions of the snRNAs: pre-mRNA and pre-rRNA processing; snR30 appears to be related to the latter group.

#### 2. Structure of U3 snRNA and possible interactions with pre-rRNA

On aligning the nucleotide sequence of snR17 with other U3 sequences, four regions of strong conservation were observed, referred to as Boxes A to D. Box A near the 5' end, with the exception of one mismatched nucleotide in the Dictyostelium sequence, is perfectly conserved over 15 nt in all species. This pentadecamer is unique to U3 and neither it nor its complement is to be found in any other published sequence in either the EMBL or the Genbank data libraries. In the secondary structure models of U3, Box A always forms one side of the stem of Hairpin 1: because no complementary, conserved sequence is observed on the other side of the stem, its importance probably lies, not in the features of secondary structure it could maintain, but in its primary structure. The conserved boxes B and C are represented as single-stranded regions in the secondary structure models. They might be conserved because of their particular tertiary folding properties, allowing the formation of tertiary base-base hydrogen bonds as in the case of tRNAs (Kim 1979, Sundaralingam 1979); they could be protein binding sites, or regions available for complementary base pairing with other RNAs. The position of Box D at the 3' end suggests a possible role as a transcription termination or 3' end processing signal.

U3 is known to be hydrogen-bonded in mammalian cell extracts to high molecular weight, nucleolar RNA fractions containing rRNA precursors (Prestayko *et al.* 1970, Calvet and Pederson 1981, Epstein *et al.* 1984; see Discussion in Chapter 2). Figure 2.10 shows that snR17 remains hydrogenbonded to what appears to be the yeast 37S initial pre-rRNA transcript. The general pattern for eukaryotic pre-rRNA processing is represented by that of *S. cerevisiae* in Figure 3.20. The order of cleavage varies somewhat in different organisms, notably the sequence of cuts 4, 5 and 7, and some organisms perform 115

additional processing steps involving excision of intervening sequences ("expansion segments") from the large rRNA sequence and splicing (Gerbi 1985, Hadjiolov 1985). Mature rRNA sequences are evolutionarily conserved and eukaryotic 5.8S rRNA is homologous to the 5' end of prokaryotic 23S (large) rRNA; the internal transcribed spacer 2 (ITS 2) therefore has been likened to an intron of pre-mRNA except that 5.8S rRNA does not become covalently linked to the large rRNA, but remains hydrogen-bonded to it. By analogy to the role proposed for U1 snRNA in the excision of introns (Lerner *et al.* 1980a, Rogers and Wall 1980), it was suggested that U3 snRNA could have a similar function in the excision of ITS 2, involving a hydrogen-bonding interaction with the "exon-intron" boundary (Bachellerie *et al.* 1983, Crouch *et al.* 1983, Tague and Gerbi 1984).

Figure 3.21 summarizes some of the proposed base-pairing interactions between U3 snRNAs and pre-rRNAs that are, at least partially, evolutionarily conserved. The interactions shown are those involving the conserved box B and box C sequences of U3 snRNAs (see Figure 3.12). The hypothetical interaction of the box C with ITS 2 at the 3' boundary of 5.8S rRNA (Bachellerie et al. 1983) is conserved in all mammals, and less stably, but consistently in Xenopus. However, whereas the U3 box C sequence is conserved in snR17, the complementary sequence is neither conserved in S. cerevisiae nor in other invertebrates (Tague and Gerbi 1984). Parker and Steitz (1987) have shown that box B is more accessible to modifying reagents than box C in the U3 snRNP, and therefore probably more readily available for base-pairing interactions: they have pointed out a potential interaction, conserved between vertebrates and yeast, between box B and the external transcribed spacer 2 (ETS 2), downstream of the 3' end of mature large rRNA, and close to processing sites in this region (Figure 3.21). A processing site has been detected +15 nt beyond the mature 25S rRNA 3' end of S. cerevisiae pre-rRNA<sup>1</sup> (Kempers-Veenstra et al. 1986), that is, within the region proposed to interact with the U3 box B, and a processing site at +30 in the mouse ETS 2 occurs in a region similar to rat ETS 2 complementary to box B of U3 snRNA; the proposition that these interactions are close to processing sites at equivalent positions in human and Xenopus, however, remains purely speculative.

Computer searches revealed several potential base-pairing interactions between snR17 and different regions of the yeast initial pre-rRNA, many of them theoretically more stable than, for example, the interaction of the 5' end of U1 116

<sup>&</sup>lt;sup>1</sup>S. cerevisiae 37S pre-rRNA is normally cleaved +10 nt beyond the mature 25S 3' end by the endonuclease product of the RNA82 gene, which also processes pre-5S rRNA. Processing at +15 nt was observed in an rna82 strain (Kempers-Veenstra *et al.* 1986).

with the 5' splice site of pre-mRNA (Mount *et al.* 1983), but none of these was supported by phylogenic comparisons. The only experimental data indicating which regions of U3 snRNA and pre-rRNA might interact appears to be from psoralen cross-linking: preliminary results suggest that U3 binds to the ETS 1, upstream of the mature 5' end of 18S rRNA (I. Stroke and A. Weiner, and R. Maser and J. Calvet, unpublished data). The role of U3 snRNA and the nature of its interaction with pre-rRNA therefore is still very much open to speculation.

Comparison of upstream sequences of SNR genes of S. cerevisiae.

	-11	LO	-100	-90	-80		-30	-20	-10	+	
SNR30		TCTCT(	CCCTATATAAA	CACTTTGAA	. ACTCTCATCGTCGT.	50nt	. TTCTC	LTTGCTTATTT	GTAAAAAAA	AGT AJ	ACCATAGTC
SNR3		GAGGT(	CCGAGCGAAAAA	AAAGTATAA	<u>A</u> GAAGCACAGGCAA	50nt	.TATTT	<b>TGAAGTTTCC</b>	TTCCGAAAAAA	AGT AJ	ACTTTGTCC
SNR10		TTTCTJ	ragc <u>tataaa</u> g	GCAGATCAA	TGCAGCATTCAATT	50nt	. TTCTT?	ATTTAGTGTT	ATTTTATAACC7	AGT AJ	ACGCAAATT
SNR17A (I	J3A)	AAAACO	CAGGATTTGAAA	ACTATAGCA	CCTAAAAGTCTTAG	50nt	. TAAAT	<b>CCAACCATTGC</b>	AGCAGCTTTTG	ACT AJ	ACACATTCT
SNR17B (I	U3B)	TTAT	ACCATATACTTT	'ATTAGGAA <u>T</u>	ATAACAAAGCATAC	50nt	CTTCT2	ATTCACCAGC	TTCAGGTTTTG?	ACT AJ	ACACATTCA
SNR19 (U:	1)	ATTGGC	CAAGACG <u>TATAA</u>	ATAGAGAGA	AGAAGTTCCATTTA	50nt	AGTTT	<b>LTCTTGAATT</b>	TTTTAAAATCO	CTC A	<b>FACTTACCT</b>
<i>SNR20</i> (U:	2)	TCGGT(	SAAAACA <u>TATAA</u>	AAAGGGCTC	TACATTCATTTTTT	50nt	TGTTT(	TACTTGTTTT	TTTTTAAATCO	CCC A(	CGAATCTCT
SNR14 (U	4)	TTTTC	CTGAA <u>TATAAA</u> A	GGAACAGAA	TATTAGTTAGCTTC	50nt	CTTTTC	CTTCACTTCC	TTCCTTAATACT	rcc A	<b>FCCTTATGC</b>
SNR7 (U5)	~	TAGAAA	ACGGAGTGCTCA	G <u>TATAAAA</u>	GCGCATAGTAAGAC	50nt	.ACTTT	гстттсттттт	GTTTAAAACCI	rgc ai	AGCAGCTTT

Figure 3.19 Comparison of the 5' flanking sequences of *SNR* genes. Putative "TATA" boxes are underlined. In addition to those shown, SNR19 has the sequence TATA at -33, and SNR7 has TATATA at -45. Sequence sources: SNR3 Tollervey et al. (1983); SNR10 Tollervey and Guthrie (1985); SNR19 Kretzner et al. (1987); SNR20 Ares (1986); SNR14 Siliciano et al. (1987a); SNR7 Patterson and Guthrie (1987).



#### Figure 3.20 Precursor rRNA processing pathway

The pre-rRNA processing pathway of S. cerevisiae is shown, but the main features are common to other eukaryotes (see reviews by Hadjiolov 1985, and Gerbi 1985). The sizes of the various precursor and mature rRNA species are indicated<sup>1</sup> (those of the corresponding mammalian species are in parentheses). ETS 1 & 2: 5' and 3' external transcribed sequences; ITS 1 & 2: first and second internal transcribed spacers. Pre-rRNA processing takes place within the nucleolus and in association with ribosomal proteins. Association of ribosomal proteins, however, is not an absolute requirement: the 37S yeast pre-rRNA is correctly processed in the absence of ribosomal proteins (Waltscheva et al. 1983). The cleavage steps are numbered, and their essential features are as follows: 1. Primary cleavage of the mammalian 47S pre-rRNA (Craig et al. 1987, Kass et al. 1987), occurs very rapidly, but has not been observed in S. cerevisiae, in which the triphosphorylated 5' end of the predominant, 37S initial transcript (pppApUp..) represents the site of transcription initiation; such cleavage may occur in yeast, but may be less easy to detect. 2. Transcription by RNA polymerase I towards its terminator (possibly +210 bp beyond the mature 3' end of 25S rRNA in yeast; Kempers-Veenstra et al. 1986) allows cleavage within 50 nt downstream of the mature 25S 3' end. 3 and 4 result in separation of the 18S sequence from the 5.8S and 25S sequences; these cleavages may proceed in a manner similar to the cleavage of 16S rRNA from E. coli pre-rRNA by RNase III, involving a staggered cut in a base-paired stem formed between the 5' ETS and the ITS 1 (Veldman et al. 1981). Cleavage at 5 allows the 5.8S rRNA sequence to associate with the 5' end of the 25S sequence by hydrogen-bonding. The mature 3' ends of the 18S and 25S sequences are formed (7 may occur in the small ribosomal sub-unit in the cytoplasm) and the ITS 2 is removed from between 25S and 5.8S rRNA (8 to 10).

<sup>&</sup>lt;sup>1</sup>The sedimentation values assigned to the various precursor and mature yeast rRNA species vary. The nomenclature adopted here is that of Tollervey (1987a) and of Philippsen *et al.* (1978).



# Figure 3.21 Hypothetical base-pairing interactions between U3 snRNA and pre-rRNA.

Potential interactions between the conserved U3 snRNA boxes B and C and sites at which pre-rRNA cleavages might, or are known to occur are shown. The interaction of box C with the 5.8S/ITS 2 boundary, proposed by Bachellerie *et al.* (1983) and Tague and Gerbi (1984), is conserved in vertebrates only. The interactions of box B with ETS 2, proposed by Parker and Steitz (1987), coincides with a cleavage site at +15 in the yeast ETS 2 (Kempers-Veenstra *et al.* 1986), and with rat ETS nts +85 to +92, which match mouse ETS nts +28 to +35, where a cleavage site has been mapped to nt +30 (Grummt *et al.* 1985). No equivalent cleavage sites have yet been shown to exist in the sequences of the other organisms. Nucleotides in ETS 2 are numbered from the 3' end of the mature large rRNA.

#### Chapter 4

#### Disruption of small RNA genes

#### SUMMARY

As a step towards elucidating the function of snR17, mutant yeast strains were created which carried deletions of the *SNR17* genes, and the resulting phenotypes were analysed.

DNA fragments constructed *in vitro*, consisting of genes from which large parts of the coding sequences had been deleted, and containing marker genes for selection *in vivo*, were introduced into diploid yeast cells. Clones were then selected in which homologous recombination had resulted in one of the wild type chromosomal loci being replaced by the introduced fragment; genotypes of the clones were confirmed by Southern blot hybridization. These heterozygous strains were sporulated, and the effects of the mutations in the resulting haploids were analysed.

Haploid strains carrying either the  $snr17a-\Delta$  or the  $snr17b-\Delta$  deletions appeared to grow normally. Higher steady-state levels of snR17B were observed in  $snr17a-\Delta$  strains; however, this increase did not match the wild type levels of snR17A, which appeared to be several fold greater than those of snR17B. Haploid strains inheriting the double mutation  $snr17a-\Delta$   $snr17b-\Delta$  were not viable.

#### INTRODUCTION

Saccharomyces cerevisiae was one of the first eukaryotic organisms for which techniques were developed for transformation, that is, the process by which naked DNA can be introduced into a cell resulting in a heritable change. Plasmid DNA used to transform yeast cells was found either to be integrated into the yeast genome, through recombination between homologous portions of DNA (Hinnen *et al.* 1978), or to replicate extrachromosomally, if the plasmid contained its own origin of DNA replication (Beggs 1978). DNA elements allowing a plasmid to replicate extrachromosomally took the form of the endogenous yeast  $2\mu$  plasmid origin of replication (Beggs 1978), or ARS elements (<u>a</u>utonomous replicating <u>s</u>equences) of chromosomal origin (Struhl *et al.* 1979).

When plasmid DNA was linearized within a sequence homologous to chromosomal DNA and used to transform yeast, it was found that the ends of the linearized plasmid were highly recombinogenic, and integration of the plasmid at the site homologous to the ends of the plasmid took place several orders of magnitude more frequently than at sites homologous to other regions of the plasmid (Orr-Weaver et al. 1981). This phenomenon can be exploited to disrupt wild-type yeast genes in vivo (Rothstein 1983): first, a significant length of gene coding sequence on a plasmid in vitro is replaced by a selectable marker gene for yeast, for example, one capable of complementing an auxotrophic mutation; then, the whole construction, excised as a linear DNA fragment, is used to transform a diploid yeast strain. After homologous recombination between sequences shared by the fragment and one of the wild type loci, one chromosomal allele is replaced by the deleted form containing the selectable marker. Heterozygotes for the deletion are then isolated, which, on sporulation, produce tetrads in which two of the haploid spores carry the mutant allele and two the wild type. The use of a diploid strain for transformation not only enables lethal mutations to be identified by scoring haploid spore viability, but also avoids the selection of suppressors of non-lethal mutations.

The strategy used to analyse the *SNR17* genes was more complex than that outlined above and involved deleting each of *SNR17A* and *SNR17B* independently using two different selectable markers, *URA3* and *LEU2*. Having established that haploid strains carrying either the *snr17a-* $\Delta$  or the *snr17b-* $\Delta$  deletions grew normally, a cross was made to produce a diploid heterozygous for deletions at both the loci; on sporulation, it was found that inheritance of the double mutation by a haploid was lethal.

#### YEAST CELLS REQUIRE AT LEAST ONE SNR17GENE.

#### 1. Plasmid construction

Plasmid pSR3 $\Delta$ *URA3* was designed to disrupt the *SNR17A* gene. It contains an *SNR17A* gene from which 309 bp from nucleotides +74 to +383 downstream of the transcription initiation site have been deleted and replaced by the *URA3* selectable marker gene. This *snr17a::URA3* construction could be excised as a 1.7 kb *Sall/Bal* fragment in which 73 bp of *SNR17A* coding sequence and 150 bp of upstream flanking sequence remained intact at the 5' end, and 170 bp of downstream flanking sequence remained intact at the 3' end.

pSR3AURA3 was constructed as follows: The 520 bp Sall/EcoRI fragment from phage QSR20 RF DNA consisting of the 3' flanking region of the SNR17A gene (Figure 3.1) and a 1330 bp Xhol/Sall fragment containing the URA3 selectable marker gene (Rose and Botstein 1983, Rose et al. 1984) from plasmid pLGD312 (Guarente and Mason 1983, from which the 1.5 kb Smal fragment had been deleted), were isolated from a polyacrylamide gel and purified over DEAE cellulose (Chapter 6, Section 5.4). Plasmid pSR3.7, containing only the 5' 73 bp of the SNR17A coding sequence (Figure 3.1), was linearized with EcoRI, and 5  $\mu$ g were treated with 1.4 U alkaline phosphatase (Böhringer) in 50mM Tris pH8, 37°C for 30', then with 100 µg/ml proteinase K for 30' at 37°C, phenol/chloroform extracted and ethanol-precipitated. 2.5 μg phosphatased plasmid DNA was mixed with approximately equal quantities of the isolated fragments in 50 µl and ligated overnight at 4°C; the mixture was then used to transform E. coli 71-18 cells. A recombinant plasmid clone with the structure illustrated in Figure 4.1 was isolated after analysing the restriction patterns of miniprep DNA.

Plasmid pSR3B $\Delta$ LEU2 was designed to disrupt SNR17B. It contains an SNR17B gene from which 466 bp from nucleotides +22 to +487 (between two Sau3A sites) have been deleted and replaced by the LEU2 selectable marker gene. This snr17b::LEU2 construction could be excised as a 2.5 kb Sall/Ball fragment, in which 21 bp of SNR17B coding sequence and 130 bp of upstream flanking sequence remained intact at the 5' end, and 120 bp of downstream flanking sequence remained intact at the 3' end.

pSR3B $\Delta$ LEU2 was constructed as follows: The DNA between the two Sau3A sites at nucleotides +18 and +487 (numbering as on the sequence in Figure 3.11; see also restriction map in Figure 3.5) was deleted from plasmid pSR3B (Figure 3.2) by partially digesting with Sau3A (one or two cuts per molecule:

10  $\mu$ g DNA in 250  $\mu$ l 50mM NaCl buffer with 8 U *Sau*3A for 15' at 37°C), filling-in, size-fractionating in LM agarose, recircularizing and transforming 71-18 to give plasmid pSR3B.67. During the process of filling-in and religating the *Sau*3A sticky ends, a new *Cla*l site was created (GATCGATC) but the site was methylation-protected after propagation in the dam<sup>+</sup> *E. coll* 71-18 strain and could not, therefore, be cut by the enzyme. The *Cla*l site in the pEMBL vector DNA of pSR3B.67 was destroyed by filling-in and religating, and the plasmid was used to transform the *dam<sup>-</sup>* strain GC344. The new, now unique and unprotected *Cla*l site was cleaved, *Xho*I linkers were added after filling-in, and an isolated 2230 bp *Sall/Xho*I fragment bearing the *LEU2* selectable marker gene from plasmid YEp13 (Broach *et al.* 1979) was inserted. pSR3B $\Delta$ LEU2 was isolated by colony hybridization using the *LEU2* fragment as a probe.

#### 2. Disruption of SNR17A in vivo.

The diploid yeast strain GY45, *ura3/ura3* (see strains list, Table I), was transformed with pSR3 $\Delta$ URA3 plasmid DNA cleaved at the unique Sal and Bal sites to release the 1.7 kb gene replacement fragment. Transformation was by the LiOAc method (Altherr *et al.* 1983). Ura<sup>+</sup> transformants were selected on SD minimal medium lacking uracil and screened for integration events at the SNR17A locus by Southern blot hybridization. [Methods for yeast transformation, selection and DNA preparation are described in Chapter 6, Section 1.]

Figure 4.1 illustrates the scheme used to disrupt the *SNR17A* gene and the expected restriction fragment sizes of the wild type and disrupted loci. The panel shows genomic DNA digested with *Hin*dIII and hybridized to a probe consisting of the *Sall/Hin*dIII fragment spanning the wild type *SNR17A* gene. Lane 1 shows the non-transformed diploid strain GY45 with one 1.4 kb band corresponding to the wild type *SNR17A Hin*dIII fragment, and one 4.6 kb band corresponding to the cross-hybridizing *SNR17B* fragment. Lane 2 shows a Ura<sup>+</sup> diploid transformant, GY68, with the band pattern predicted for a diploid heterozygous for the disruption of *SNR17A*. Lane 3 shows a Ura<sup>+</sup> haploid strain, GY69, grown from an ascospore of the diploid in Lane 2, exhibiting the band pattern of the *snr17a::URA3 SNR17B* genotype.

Southern blot analysis revealed that several Ura<sup>+</sup> transformants had acquired integrated copies of the complete pSR3 $\Delta$ URA3 plasmid. The band patterns were consistent with the linear plasmid having integrated at the *Ball* site. This was probably due to the incomplete digestion of the plasmid DNA. When *Ball*-linearized pSR3 $\Delta$ URA3 was used to transform GY45 as a control, it was

found that integration of the whole plasmid occurred approximately 10 x more frequently than integration of the Sall/Ball fragment. In order to reduce the background of Ura<sup>+</sup> colonies caused by whole plasmid integration, in later experiments, fragments targeted to replace genes were first isolated and purified.

3.  $snr17a-\Delta$  strains have increased levels of snR17B RNA.

The *snr17a::URA3/SNR17A* (= a/A) heterozygous strain GY68 was induced to sporulate and haploid daughter strains were isolated by dissecting asci with a micromanipulator (Chapter 6, Section 1.8). Spore viability was poor, but three Ura<sup>+</sup> strains, GY69, GY70 and GY71 with the genotype *snr17a::URA3 SNR17B* (= aB), confirmed by Southern blot hybridization (GY69 is shown in the panel in Figure 4.1, iane 3), were identified from different asci.

The snr17a- $\Delta$  strains appeared to grow normally at 30°C on YPD. Total cellular RNA was isolated from exponentially growing cultures, electrophoresed on 6% polyacrylamide/urea gels and visualized by ethidium bromide-staining under U.V. light. The total RNA from two *aB* strains and two wild type *AB* sister strains is shown in Figure 4.2: a band at the expected position of snR17A is noticeably lacking from the *aB* strains.

In order to confirm that snR17A RNA was absent, total cellular RNA was transferred to filters and probed with the labelled oligonucleotides 49 and 50, which specifically hybridize to snR17A and snR17B, respectively (see Chapter 2). Figure 4.3 shows RNA from five different strains probed, in the left panel, with oligo 49 (snR17A-specific) and, in the right panel, with oligo 50 (snR17B-specific). SnR17A is absent from the haploid *aB* strain GY69, whereas snR17B is present (lanes 5). Both the haploid *aB* GY69 (lane 5) and its diploid *a/A B/B* parent GY68 (lane 4) show elevated steady-state levels of snR17B, compared to the *A* or *A/A* control strains (lanes 1, 2 and 3). The control samples are RNA from: 1. an *AB* sister strain of GY69; 2. a diploid strain GY67 heterozygous for integration of the complete pSR3 $\Delta$ URA3 plasmid at the *Bal*I site downstream of *SNR17A* (without disrupting the gene); and 3. GY72, a haploid daughter strain of GY67, also carrying the integrated plasmid.

Figure 4.4A shows that the steady-state levels of snR17B are elevated in all three of the haploid *a* strains, GY69, GY70 and GY71 (lanes 6, 7 and 8) as well as in the *a*/*A* heterozygote GY68 (lane 3), compared to the *A* or *A*/*A* controls (lanes 1, 2, 4 and 5). In Figure 4.4B the same samples have been probed with a rDNA clone hybridizing to 5.8S and large rRNA, and the levels of 5.8S rRNA in each lane are the same.

The increased amount of snR17B in  $snr17a-\Delta$  strains was reproducibly observed when the  $snr17a-\Delta$  lesion was introduced into a different strain. It was found, however, that levels of snR17A were not affected by a deletion in the *SNR17B* gene (see Figure 4.9). Figure 4.2 indicates that snR17A is normally in greater abundance than snR17B, and that the increase in snR17B in  $snr17a-\Delta$ strains does not restore the total snR17 to normal levels.

#### 4. Disruption of SNR17B in vivo

SNR17B was disrupted in strain GY87 (*ura3/ura3 leu2/leu2*, see Table I) in the hope of obtaining better spore viability than was observed for GY68. The SNR17A gene was also disrupted in this strain in order that isogenic strains for snr17a- $\Delta$  and snr17b- $\Delta$  would be available.

The scheme for the disruption of *SNR17B* is shown in Figure 4.5. GY87 (Leu<sup>-</sup>) was transformed with the isolated 2.5 kb *Sall/Bal*I replacement fragment containing the *LEU2* gene from plasmid pSR3B $\Delta$ *LEU2*. Leu<sup>+</sup> transformants were selected on minimal medium lacking leucine, and were screened by Southern blot hybridization. The DNA band patterns of non-transformed GY87 and of an *snr17b::LEU2/SNR17B* transformant, GY94, are shown in the panel of Figure 4.5; the probe used was the *Eco*RI fragment spanning the wild type *SNR17B* gene.

SNR17A was likewise disrupted in strain GY87 by transforming with the isolated Sall/Ball fragment of pSR3 $\Delta$ URA3 and selecting for Ura<sup>+</sup> transformants, to give GY93 (snr17a::URA3/SNR17A).

GY93 (a/A B/B) and GY94 (A/A b/B) were sporulated and asci were dissected with a micromanipulator. Both GY93 and GY94 yielded tetrads in which all four spores were viable, and in which Ura<sup>+</sup> and Ura<sup>-</sup> or Leu<sup>+</sup> and Leu<sup>-</sup> phenotypes segregated 2:2. No differences between the growth characteristics of the aB, Ab and AB strains were observed on either YPD (rich) or minimal media.

#### 5. Inheritance of the snr17a- $\Delta$ snr17b- $\Delta$ double deletion is lethal.

GY114, a Ura<sup>+</sup> Leu<sup>-</sup> (*aB*) haploid daughter strain of GY93, was crossed with GY119, a Ura<sup>-</sup> Leu<sup>+</sup> (*Ab*) haploid daughter strain of GY94, and a doubly heterozygous diploid strain GY132 (a/A B/b) was selected on minimal medium lacking uracil and leucine. The genotype of GY132 was confirmed by Southern blot hybridization, as shown in Figure 4.6.

The four diploid strains GY87 (A/A B/B), GY93 (a/A B/B), GY94 (A/A B/b) and GY132 (a/A B/b) were induced to sporulate and full complements of four haploid spores from individual asci were dissected out onto YPD plates. The

products were analysed by scoring the phenotypes, and by DNA and RNA blot hybridization.

The expected segregation patterns of the *snr17a::URA3* and *snr17b::LEU2* genotypes among spores from GY132 are summarized in Figure 4.7. Parental ditype (PD), non-parental ditype (NPD) and tetratype (T) tetrads would be expected to occur at a ratio of 1:1:4 if one or both genes are neither centromere-linked nor linked to each other (linkage of two genes on the same chromosome is reflected by a PD/NPD ratio significantly greater than 1; if two genes are centromere linked on different chromosomes, the frequency of T asci is reduced, ie: PD:NPD:T is 1:1:<4; Mortimer and Hawthorne 1969).

Meiotic segregation of snr17A and snr17B



PD = Parental Ditype NPD = Non-Parental Ditype T = Tetratype

# Figure 4.7 Possible segregation patterns of *snr17a* and *snr17b* alleles among tetrads from a doubly heterozygous diploid strain.

Figure 4.8 shows the phenotypes of tetrad colonies from the four diploid strains. Ura<sup>+</sup> Leu<sup>-</sup> or Ura<sup>-</sup> Leu<sup>+</sup> colonies are represented by diagonal bars. Ura<sup>+</sup> Leu<sup>+</sup> colonies were not observed. The observed PD:NPD:T ratio for GY132 (a/A B/b) tetrads, 1:3:11, is consistent with the expected ratio, 1:1:4, for the

random segregation of two unlinked genes. Spores which did not germinate have not been marked. However, where microscopic examination revealed that spores had germinated but failed to form colonies after going through three or four cycles of cell division, a filled circle has been drawn. Such abortive "lethal" colonies were only observed where a spore could have inherited the the *snr17a::URA3 snr17b::LEU2* double deletion, as predicted from the segregation pattern of marker phenotypes among other spores of the same tetrad. It was concluded, therefore, that inheritance of the double deletion by a haploid was lethal.

The coincidence of the Ura<sup>+</sup> and Leu<sup>+</sup> phenotypes with the absence of snR17A and snR17B, respectively, was confirmed by RNA blot hybridization. Figure 4.9 shows RNA from strains of representative tetrads probed with the snR17A-specific (upper panel) and snR17B-specific (lower panel) oligonucleotides 49 and 50. Columns 1 and 2 illustrate tetrads from GY93 (*a/A B/B*) and GY94 (*A/A B/b*), and columns 3, 4 and 5 represent, respectively, parental ditype, non-parental ditype and tetratype tetrads of GY132 (*a/A B/b*). It is apparent that the levels of snR17B were higher in the strains lacking snR17A: the levels of snR17A, however, did not change detectably when snR17B was absent.

#### 1. SnR17 is essential.

In this chapter it has been shown that snR17 is essential for yeast viability. This observation is particularly interesting in the light of the similarities between snR17 and U3 snRNA, implying that U3 snRNA has an essential function in other eukaryotes. All the S. cerevisiae snRNAs homologous to the abundant vertebrate U snRNAs involved in pre-mRNA splicing have been shown also to be essential (Kretzner et al. 1987, Siliciano et al. 1987b, Ares 1986, Siliciano et al. 1987a, Patterson and Guthrie 1987, Brow and Guthrie 1988). Several yeast snRNAs, including snR17, exhibit properties consistent with being hydrogen-bonded to pre-rRNA (see Chapter 2), namely snR3, snR4, snR5, snR8, snR9, snR10, snR128 and snR190 (Tollervey 1987a, Zagorski et al. 1988). Of these, only snR128 is essential (Zagorski et al. 1988). Strains carrying the deletion mutation snr10 grow slowly and process 37S pre-rRNA in an aberrant manner (Tollervey and Guthrie 1985, Tollervey 1987a), while the genes for the remaining snRNAs are dispensable: even the haploid quintuple deletion mutant snr3 snr4 snr5 snr8 snr9 exhibits no apparent deleterious phenotype, and the sextuple snr3 snr4 snr5 snr8 snr9 snr10 mutant is indistinguishable from the single snr10 mutant (Parker et al. 1988). SnR30, the cloning of which has been described in Chapter 2 (see also Chapter 6, Section 1) has been found to be essential (Bally et al. 1988).

The results of deletion analyses of *S. cerevisiae* snRNA genes published to date are summarized in Table II. SnR17 and snR30 bring the total number of essential snRNAs of *S. cerevisiae* to eight.

#### 2. The two SNR17 genes are complementary.

SnR17 is the only yeast snRNA found to be encoded by two genes. Both *SNR17A* and *SNR17B* are transcribed, and it has been shown that the gene product of one can complement the lack of the other. The strong similarity between the nucleotide sequences (96%), both of which potentially can form the same U3 snRNA-like secondary structure, suggest that the functions of the two transcripts are identical. [Hybridization with the snR17A- and snR17B-specific oligonucleotide probes has shown that both transcripts can be found hydrogenbonded to pre-rRNA (David Tollervey, personal communication).]

The steady-state levels of snR17A were several fold higher than those of snR17B in strains with wild type SNR17 genes. The levels of snR17B were increased in both haploid  $snr17a-\Delta$  strains and in diploid  $snr17a-\Delta/SNR17A$ 

heterozygous strains, but not increased to the level of the wild type snR17A concentration. No increase in the level of snR17A was observed in  $snr17b-\Delta$ strains. Possibly the most likely and simplest explanation of this phenomenon is that snR17A and snR17B are transcribed constitutively at different rates (snR17A several fold higher than snR17B), that the two gene products are equally stable, but that the turnover rate is dependent on the total concentration of snR17, such that, in snr17a- $\Delta$  strains, snR17B can accumulate to a higher concentration than when snR17A is being transcribed. The same would be true for snR17A in snr17b- $\Delta$  strains, but the difference between the steady-state levels of snR17A would be proportionately less, and consequently less readily detectable. An intriguing possibility consistent with this hypothesis is that snR17 could become metabolically more stable when associated with pre-rRNA, the increase in snR17B in snr17a- $\Delta$  strains would then represent the greater proportion of snR17B bound to pre-rRNA in the absence of snR17A. The hypothesis that the promoters of the two genes operate with different efficiencies is supported by the fact that the sequences upstream of the two genes are divergent.

Another explanation could be a feed-back mechanism regulating the rate of transcription according to the cellular concentration of total snR17: here again the simplest model would assume that *SNR17A* is more efficiently transcribed than *SNR17B*, but that both genes bind a repressor with equal efficiency, which only exerts repression when the total snR17 concentration is above a certain concentration; since snR17A represents the greater proportion of total snR17 in wild type cells, the derepression of *SNR17B* in *snr17a-* $\Delta$  strains would be proportionally greater than the derepression of *SNR17A* in *snr17b-* $\Delta$  strains. It is less easy to conceive why a mechanism of feed-back regulation acting specifically on *SNR17B* and not on *SNR17A* should have evolved, unless snR17B performs a specific, distinct function, and this seems unlikely considering the close structural similarity of the molecules.

## Strain

Genotype

GY1 (VB2-20A):	$\alpha$ ura3 leu2 ade2 trp1
GY2 (CHL2881-1C):	a ura3 leu2 lys2 trp1
GY9 (F102-2):	a leu2-3,112 his4-519 can1 [p°]
GY45:	2n GY1 x GY2
GY67:	as GY45, <i>SNR17A</i> (pSR3∆ <i>URA3</i> )/SNR17A
GY68:	as GY45, snr17a::ÜRA3/SNR17A
GY69 (from GY68):	α snr17a::URA3 leu2 lys2 trp1
GY70 (from GY68):	α snr17a::URA3 leu2 lys2 trp1
GY71 (from GY68):	a snr17a::URA3 leu2 lys2 ade2 trp1
GY72 (from GY67):	a SNR17A(pSR3AURA3) leu2 lys2 trp1
GY85 (JR2619B):	α ura3 leu2 lys1-1 his3 ade2-1 can1-100
GY86 (JU4-7):	a ura3 leu2 his4 ade2 trp1 can1 flo
GY87:	2n GY85 x GY86
GY93:	as GY87, snr17a::URA3/SNR17A
GY94:	as GY87, snr17b::LEU2/SNR17B
GY113 (from GY93):	α snr17a::URA3 leu2 trp1 his3 or 4 ade2 lys1 flo
GY114 (from GY93):	a snr17a::URA3 leu2 trp1 his3 or 4 ade2
GY115 (from GY93):	a snr17a::URA3 leu2 lys1 his3 or 4 ade2
GY116 (from GY93):	a snr17a::URA3 leu2 trp1 lys1 his3 or 4 ade2 flo
GY117 (from GY93):	a snr17a::URA3 leu2 trp1 his3 or 4 ade2 flo
GY118 (from GY93):	a snr17a::URA3 leu2 trp1 his3 or 4 ade2
GY119 (from GY94):	α snr17b::LEU2 ura3 lys1 trp1 ade2
GY120 (from GY94):	α snr17b::LEU2 ura3 trp1 ade2 his3 or 4
GY121 (from GY94):	α snr17b::LEU2 ura3 ade2 his3 or 4 flo
GY122 (from GY94):	α snr17b::LEU2 ura3 ade2
GY123 (from GY94):	α snr17b::LEU2 ura3 ade2 flo
GY124 (from GY94):	α snr17b::LEU2 ura3 trp1 ade2 his3 or 4
GY125 (from GY94):	α snr17b::LEU2 ura3 lys1 ade2
GY132:	2n GY114 x GY119
GY133:	2n GY115 x GY123
GY134:	2n GY116 x GY122
GY135:	2n as GY87, <i>scr1-∆</i> ∆( <i>Asu</i> ll-Xhol)URA3/SCR1
GY136 (from GY135):	$\alpha$ scr1- $\Delta$ $\Delta$ (Asull-Xhol)::URA3 his3 his4 ade2 trp1 leu2
GY137 (from GY135):	$\alpha$ scr1- $\Delta$ $\Delta$ (Asull-Xhol)::URA3 his3 ade2 trp1 lys1 leu2
GY138 (from GY135):	a SCR1 ura3 his4 ade2 trp1 leu2
GY139 (from GY135):	a SCR1 ura3 ade2 lys1 leu2
GY140 (from GY135):	a SCR1 ura3 his3 his4 ade2 lys1 leu2 flo
GY141 (from GY135):	a SCR1 ura3 his4 ade2 trp1 lys1 leu2
GY142 (from GY135):	$\alpha$ scr1- $\Delta$ $\Delta$ (Asull-Xhol)::URA3 his3 ade2 trp1 leu2 flo
GY143 (from GY135):	$\alpha$ scr1- $\Delta$ $\Delta$ (Asull-Xhol)::URA3 his3 ade2 trp1 leu2 flo
GY150 (A4772-15A):	<b>a</b> ade2-1 his4-15 [ρ°]

-

æ

# Table IIGene deletion analysis of Saccharomyces cerevisiae snRNAs

The table summarizes the results of published gene deletion analyses of yeast snRNAs.

<u>S. cerevisiae</u> snRNA	<u>Vertebrate</u> equivalent	<u>Gene copy</u> number	Phenotype of null-mutation	<u>Reference</u>
Spliceosomal RN/	As			
snR19	U1	1	Lethal	1,2
snR20/LSR1	U2	1	Lethal	3
snR14	U4	1	Lethal	4
snR7	U5	1	Lethal	5
snR6	U6	1	Lethal	6
Pre-rRNA RNAs				
snR3	?	1	none	7
snR4	?	1	none	8
snR5	?	1	none	8
snR8	?	1	none	8
snR9	?	1	none	8
snR10	?	1	Slow growth,	
			cold sensitivity,	
			defective pre-RNA	
			processing	9,10
snR190	?	1	none	11
snR128	?	1	Lethal	11
snR17	U3	2	Lethal	12
snR30 <sup>*</sup>	?	1	Lethal	13

#### References

- 1. Kretzner et al. (1987)
- 2. Siliciano et al. (1987)b
- 3. Ares (1986)
- 4. Siliciano et al. (1987)a
- 5. Patterson and Guthrie (1987)
- 6. Brow and Guthrie (1988)
- 7. Tollervey et al. (1983)
- 8. Parker et al. (1988)
- 9. Tollervey and Guthrie (1985)
- 10. Tollervey (1987)
- 11. Zagorski et al. (1988)
- 12. This work, Hughes et al. (1987)
- 13. Bally et al. (1988)

\* Association of snR30 with pre-rRNA has not been tested.

#### Figure 4.1 Disruption of SNR17A

The upper diagram represents the structure of the *SNR17A* locus before and after replacement, through homologous recombination, by the *Sall-Ball* fragment of plasmid pSR3 $\Delta$ *URA3*, containing an insertion of the *URA3* marker gene in place of a 309 bp deletion. The lower panel shows the confirmation of the genomic structures by Southern blot hybridization. *Hin*dIII-digested genomic DNA from various yeast strains was probed with a *Sall-Hin*dIII DNA fragment containing the wild type *SNR17A* gene: lane 1, the non-transformed diploid strain GY45; lane 2, the diploid strain GY68, a transformant of GY45 heterozygous for integration of the *Sall-Ball* fragment at the *SNR17A* locus; and lane 3, GY69, a haploid daughter strain of GY68 carrying the *snr17a::URA3* lesion. The probe cross-hybridizes with the 4.6 kb *SNR17B* HindIII fragment (see Figure 4.5).





### Figure 4.2

The lack of snR17A.

Total RNA extracted from four haploid strains, resolved by electrophoresis on a 6% polyacrylamide/urea gel (20  $\mu$ g RNA per lane) stained with ethidium bromide. **a**<sup>-</sup> = *snr17a::URA3*, **A**<sup>+</sup> = *SNR17A*. The prominent snR17 band in the strains carrying the wild type gene is indicated.

# Figure 4.3 The absence of snR17A in *snr17a::URA3* strains confirmed by hybridization.

Total RNA was extracted from various yeast strains, equal quantities were electrophoresed on 6% polyacrylamide/urea gels, transferred to filters and probed with the snR17A-specific oligo 49 (left panel) and snR17B-specific oligo 50 (right panel). [Oligos 49 and 50 are complementary to regions of snR17A and snR17B where there are 5 out of 18 nucleotide mismatches, see Chapter 3.]

Lanes 1: Sister strain of GY69, n SNR17A SNR17B.

Lanes 2: GY67, 2n SNR17A (pSR3AURA3)/SNR17A SNR17B/SNR17B.

Lanes 3: GY72, n SNR17A (pSR3AURA3) SNR17B

Lanes 4: GY68, 2n snr17a::URA3/SNR17A SNR17B/SNR17B.

Lanes 5: GY69, n snr17a::URA3 SNR17B.

Hybridization proceeded as follows: RNA was elecrotransferred to Genescreen (Dupont) filters as recommended by the manufacturers; the filters were prehybridized in 25 ml 6 x SSC, 10 x Denhardt's, 0.1% SDS, 10  $\mu$ g/ml sheared denatured salmon sperm DNA (Sigma) at room temperature for 1-2 hrs; 10 pmol oligo 49 or 50 were 5' end labelled with [ $\gamma$ -<sup>32</sup>P]ATP, purified by DEAE chromatography, added to 10 ml hybridization mix (as above, but without salmon sperm DNA) and incubated with the filters for 1-2 hrs; the filters were washed several times in 6 x SSC at room temperature, three times for 5' at 42°, and then exposed.





Figure 4.5

# Disruption of SNR17B

gene in place of a 466 bp deletion. The diagram represents the structure of the genomic locus before and after replacement. The panel shows the Southern blot analysis of the non-transformed diploid strain GY87 (lane 1) and a Leu<sup>+</sup> transformant GY94 (lane 2), heterozygous for the replacement. Genomic DNA was digested with *Eco*RI and probed with the wild type *SNR17B Eco*RI and *SNR17A Hind*III fragments. The 4.3 kb *Eco*RI Disruption of the SNR17B locus by replacement with the Sall-Ball fragment of plasmid pSR3BALEU2, containing an insertion of the LEU2 marker fragment contains the SNR17A gene (see Figure 4.1).



Figure 4.6 Construction of a doubly heterozygous, *snr17a/SNR17A snr17b/SNR17B*, diploid strain.

Southern blot analysis of A/A B/B GY87 (lane 1), a/A B/B GY93 (lane 2), A/A b/B GY94 (lane 3), and a/A b/B GY132 and GY 133 (lanes 4 and 5) diploid strains. Total genomic DNA was digested with *Hind*III and probed with plasmids pSR3.3 (see Figure 3.1) and pSR3B (see Figure 3.2), containing the wild type *SNR17A* and *SNR17B* genes. The origins of the various DNA fragments are illustrated in Figures 4.1 and 4.5.

GY93 and GY94 were obtained by transforming GY87 (Ura<sup>-</sup> Leu<sup>-</sup>) with the Sall-Ball fragments from plasmids pSR3 $\Delta$ URA3 and pSR3B $\Delta$ LEU2, respectively; GY132 and GY133 were subsequently obtained by crossing snr17a with snr17b haploid daughter strains of GY93 and GY94 (see strains list in Table I).



Tet	rad	analy	sis		
Control A B A B	a B A B	A B A b	a B A b	}	PARENT STRAIN
0000	0000	0000	0000		PARENTAL DITYPE
0000	0000 0000 0000				NON- PARENTAL DITYPE
0000					TETRATYPE
○ = ( ○ = ( ○ = ( ● = (	ura-/leu- Jra+/leu- ura-/Leu+ LETHAL		000 00		

Figure 4.8 Tetrad analysis of snr17a and snr17b deletion strains

Germinated spores are marked as disks, arranged in rows of four corresponding to tetrads. Spores that did not germinate are not marked: only those spores which germinated but failed to go through more than three or four cell divisions are marked LETHAL.

The parent diploid strains were: GY87 (A/A B/B); GY93 (a/A B/B); GY94 (A/A b/B); and GY132 (a/A b/B).

#### Figure 4.9

Probing for snR17A and snR17B in snr17a and snr17b strains.

RNA segregation patterns in representative tetrads from 1. GY93 (a/A B/B); 2. GY94 (A/A B/b), 3. GY132 (a/A B/b), parental ditype; 4. GY132 (a/A B/b), non-parental ditype (only two out of four spores were viable, and only one of the identical, viable spores is shown); 5. GY132 (a/A B/b), tetratype (only three out of four spores were viable). The samples are blots of RNA (equal quantities in each lane), extracted from cultures grown to the same optical density, probed with snR17A-specific oligo 49 (upper panel) or snR17Bspecific oligo 50 (lower panel).


Yeast contains a family of snRNPs antigenically related to U3 snRNP

#### SUMMARY

When human anti-(U3)RNP antibodies, specific for a 36 kD polypeptide of the human U3 snRNP, were incubated with yeast cell extracts, the antibodies recognised a specific subset of yeast RNP containing several distinct species of  $m_3G$ -capped snRNAs. SnR17, which has significant structural similarities to U3 snRNA, was among the most abundant species immunoprecipitated; snR30 appeared to be weakly precipitated; but none of the snRNA species found in the yeast spliceosome reacted with the anti-(U3)RNP antibodies. Yeast RNP precipitated by Sm antibodies not only contained all the snRNAs shown to be involved in pre-mRNA splicing, but also contained snR17, although as a relatively minor component. Unlike vertebrate U3 snRNA, snR17 contains a consensus Sm protein-binding site sequence (AAUUUUGA), which could explain its association with Sm proteins.

Of the snRNAs found in anti-(U3)RNP precipitates, three for which nucleotide sequences are available, snR17, snR3 and snR30, shared a common sequence motif conserved in all U3 snRNAs, referred to as "box C". In order to test the hypothesis that the box C sequence could be a U3 snRNP protein-binding site, monomethyl guanosine-capped RNA transcripts with sequences unrelated to snRNAs, but containing box C motifs, were synthesised *in vitro* and injected into *Xenopus* oocytes; the ability of *Xenopus* U3 snRNP proteins to bind to the transcripts was then assayed by immunoprecipitation of oocyte extracts. The results indicated that the box C sequence alone was probably not sufficient to allow binding of the 36 kD U3 snRNP antigen.

#### INTRODUCTION

Sera of patients with a variety of systemic rheumatic diseases contain autoantibodies directed against various nuclear antigens. Certain types of antibodies tend to be characteristic of certain classes of disease, and "Sm" antibodies (so named after the patient from whom they were first characterized), which give a particular speckled immunohistochemical nuclear staining pattern, are diagnostic of systemic lupus erythematosus (reviewed by Tan 1982). Lerner and Steitz (1979) first discovered that Sm antibodies recognised specifically protein components of RNP particles containing the U small nuclear RNAs (U snRNAs). U snRNAs had been described more than a decade previously (Weinberg and Penman 1968, reviewed by Busch et al. 1982), but their functions were unknown, although a role for the nucleolar U3 snRNA had been proposed in pre-rRNA processing, owing to its association with pre-rRNA (Prestayko et al. 1970). The availability of specific antisera proved very useful for the purification and functional analysis of snRNPs: anti-snRNP antibodies were shown to inhibit pre-mRNA splicing by depleting snRNPs from cell fractions in vitro (Yang et al. 1981, Padgett et al. 1983, Krämer et al. 1984); and various patient antisera reacting with specific subsets of snRNPs, along with antibodies directed specifically against the snRNA 5' trimethylguanosine cap structure (Lührmann et al. 1982), have been used to purify snRNP particles by immune affinity chromatography and in the detailed characterization of their constituent proteins (Bringmann et al. 1983, Pettersson et al. 1984, Bringmann and Lührmann 1986, reviewed by Lührmann 1988).

The Sm proteins, that is, those proteins containing one or more epitopes of Sm specificity, constitute a set of seven core polypeptides making up each of the snRNPs that contain the abundant, vertebrate, non-nucleolar snRNAs U1, U2, U4, U5 and U6 (reviewed by Lührmann 1988). The core polypeptides are named B' (29 kD), B (28 kD), D (16 kD), D' (15.5 kD), E (12 kD), F (11 kD) and G (9 kD). Most Sm antisera contain antibodies directed against a common epitope shared by B', B and D, while antibodies in some sera react with other epitopes on these and the remaining polypeptides. The U1, U2 and U5 snRNPs, in addition, contain one or more polypeptides unique to the individual particles: most systemic lupus erythematosus patients, and patients with mixed connective tissue disease, have U1 snRNP-specific antibodies (Lerner and Steitz 1979) which usually react with the two largest of the three U1 snRNP polypeptides, 70 kD, A (34 kD) and C (22 kD); anti-(U2)RNP patient antiserum has been found to react with the larger of the two U2 snRNP-specific polypeptides, A' (33 kD) and B'' (28.5 kD) (Mimori *et al.* 1984); and a third class of patient antibodies

reacts with both the U1 snRNP A, and the U2 snRNP B" polypeptides (Habets *et al.* 1985). No U5 snRNP-specific patient antibodies have been found, although U5 snRNP has been observed to contain a specific 25 kD polypeptide (LeLay-Taha *et al.* 1986). The so-called "intron-binding protein", which associates with the poly-pyrimidine tract close to the 3' splice sites of pre-mRNAs, copurifies with U5 snRNP at low Mg<sup>2+</sup> concentration (Tazi *et al.* 1986), Gerke and Steitz 1986); this protein has at least one Sm epitope (Gerke and Steitz 1986). U4 and U6 snRNAs are base-paired to each other in the same RNP particle (Bringmann *et al.* 1984, Hashimoto and Steitz 1984), but no proteins specific to this RNP have been identified yet. U6 snRNA is distinct from other snRNAs in that it is transcribed by RNA polymerase III and does not have a characteristic 5' m<sub>3</sub>G cap structure; it does not bind Sm proteins directly, but is Sm-precipitable by virtue of its association with U4 (Guthrie and Patterson 1988).

*S. cerevisiae* homologues for each of the abundant, non-nucleolar vertebrate snRNAs have been found, U1/snR19 (Kretzner *et al.* 1987, Siliciano *et al.* 1987b), U2/snR20 (Ares 1986), U4/snR14 (Siliciano *et al.* 1987a), U5/snR7 (Patterson and Guthrie 1987) and U6/snR6 (Brow and Guthrie 1988), and their associated proteins are recognised by Sm antibodies (Tollervey and Mattaj 1987, Riedel *et al.* 1987, Siliciano *et al.* 1987b). The genes *RNA2, RNA3, RNA4, RNA5, RNA7, RNA8* and *RNA11* have all been shown to be required for pre-mRNA splicing (Lustig *et al.* 1986), and possibly encode snRNP proteins.

The Sm protein binding site of snRNAs has been defined as a short oligonucleotide with the consensus sequence  $RA(U)_{4-6}GR$  (R = purine) occurring in single-stranded regions of all those snRNA examined (Branlant *et al.* 1982, Liautard *et al.* 1982, reviewed by Mattaj 1988). Substitution of this sequence in human U1 (Hamm *et al.* 1987) or *Xenopus* U2 (Mattaj and De Robertis 1985) with random sequences abolishes the ability to be immunoprecipitated with Sm antiserum, whereas deletion of other parts of the RNA does not; moreover, the presence of a consensus Sm binding site on an artificial T7 transcript can direct the assembly of Sm proteins after injection into *Xenopus* oocytes (Mattaj 1986).

Patients with the autoimmune disease scleroderma have high titres of circulating antibodies reacting with nucleolar antigens (Pinnas *et al.* 1973, Reddy *et al.* 1983, reviewed by Tan 1982), some of which are specific to the U3 snRNP (Lischwe *et al.* 1985, Reimer *et al.* 1987). Parker and Steitz (1987) have shown that human U3 snRNP immunoprecipitated with anti-(U3)RNP serum contains six polypeptides, two of which, 74 kD and 59 kD, are

phosphorylated, and four of which, 36 kD, 30 kD, 13 kD and 12.5 kD, are not phosphorylated; the anti-(U3)RNP serum reacted specifically with the 36 kD polypeptide (Lischwe *et al.* 1985, Reimer *et al.* 1987). No reaction of Sm antisera with U3 snRNP has ever been reported.

This chapter describes how human anti-(U3)RNP serum was used to immunoprecipitate extracts of yeast cells in an attempt to determine whether the U3 and snR17 RNPs shared conserved antigenic determinants. It was found that this antiserum could precipitate RNP containing not only snR17, but also a number of other yeast snRNAs, at least one of which, snR3, has also been shown to be associated with pre-rRNA (see Figure 2.10, and Tollervey 1987a).

SnR3 (Tollervey *et al.* 1983) contains a short sequence resembling the box C sequence conserved between snR17 and U3 snRNAs contained in a singlestranded loop of the predicted secondary structure; a similar sequence was also found in snR30 (Chapter 3, Figure 3.15), and snR30 could be detected weakly in anti-(U3)RNP precipitated material. This prompted an investigation to test whether the conserved box C was a binding site for the anti-(U3)RNP serum antigen: the results of introducing representative box C sequences into artificial T7 transcripts indicated that the box C sequence was probably not sufficient to direct the association of the antigen with the RNA in *Xenopus* oocytes.

#### RESULTS

#### 1. Immunoprecipitation of yeast snRNPs with human antisera

Human antiserum from a patient suffering from the autoimmune disease scleroderma has been shown to react specifically with the 36 kD protein of the U3 snRNP (Lischwe et al. 1985, Parker and Steitz 1987).<sup>1</sup> SnR17, along with a number of other snRNAs, was enriched from yeast nuclear extracts by immunoprecipitation with this antiserum.<sup>2</sup> Figure 5.1 shows the RNA species 3' end-labelled after immunoprecipitation from yeast nuclear extracts using different antisera: Lupus erythematosus Sm and anti-(U1)RNP, rabbit antim<sub>3</sub>G cap [note that the anti-cap serum reacts with the RNA irrespective of the presence of protein], and the scleroderma anti-(U3)RNP. Lane "N" shows 3' end-labelled total nuclear extract RNA. The precipitations were performed by incubating the nuclear extract with antisera bound to protein A-sepharose beads, and as controls, beads without antisera were used. As a further control to show that anti-(U3)RNP serum was reacting specifically with protein components in the nuclear extract, antisera were incubated with deproteinized (naked) RNA. The pattern of snRNAs precipitated by anti-cap serum was essentially the same whether from nuclear extract or naked RNA, and corresponds well to the pattern presented by Riedel et al. (1986), by careful comparison with which the major species have been tentatively identified.

At least eleven prominent RNA bands comigrating with anti-capprecipitated RNAs appeared to be enriched specifically by anti-(U3)RNP serum: snR17, 328 nt long, at "3" in Figure 5.1, which was positively identified by hybridization with a DNA probe; a cluster of bands at "5" between 180 and 190 nt (two dimensional gel electrophoresis, performed as described by Wise *et al.* 1983, revealed that these included snR3, snR4, snR5 and snR28, David Tollervey, unpublished results); a band of approximately 169 nt at "7" comigrating with snR15, a band at "9" of approximately 124 nt comigrating with snR13 (also supported by two dimensional electrophoretic analysis), and three bands migrating faster than 5S rRNA (not labelled in Figure 5.1) with approximate sizes of 105, 103 and 97 nt. In addition to these prominent bands, six less intense bands comigrating with  $m_3G$  capped RNAs were specifically

<sup>&</sup>lt;sup>1</sup>The U3 snRNP protein described as 34 kD by Lischwe *et al.* (1985) is referred to in this work and by Parker and Steitz (1987) as 36 kD.

<sup>&</sup>lt;sup>2</sup>Scleroderma anti-(U3)RNP antiserum was obtained from the patient Linda Skinner and was kindly given, along with the anti-(U1) serum, by Kathy Parker of the Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

enriched in the anti-(U3)RNP precipitates, with the following sizes: ~500 nt (migrating below snR19 at "2"); ~393 nt comigrating with snR18a (Riedel *et al.* 1986); ~324 nt (migrating just below snR17 at "3"); ~277 nt; ~257 nt comigrating with snR11 (Riedel *et al.* 1986); and ~200 nt comigrating with snR16 (Riedel *et al.* 1986). A very faint band corresponding to snR30 (605 nt) was also visible; snR30 is the prominent band migrating just above "2" (U1/snR19) in the anti-cap lanes.

In the Sm precipitated samples, three very prominent, specifically enriched bands were observed, at "4", "6" and "8", which correspond in mobility to the "large" (213 nt) and "small" (178 nt) forms of snR7 (S. cerevisiae U5, Patterson and Guthrie 1987), and snR14 (160 nt) (S. cerevisiae U4, Siliciano et al. 1987a), respectively; these species have previously been shown to be Smprecipitable from yeast (Tollervey and Mattaj 1987, Riedel et al. 1987). Several other bands of lesser intensity were visible in the Sm precipitates: among these LSR1 or snR20 (1175 nt) (S. cerevisiae U2, Ares 1986) at "1" and snR19 (569 nt) (S. cerevisiae U1, Kretzner et al. 1987, Siliciano et al. 1987b) at 2 were tentatively identified [these have since been shows to be Sm precipitable from yeast, and to be enriched in yeast spliceosomes (Siliciano et al. 1987b)], and snR17 (328 nt) at "3" was positively identified. SnR17 was not the only RNA present in both the anti-(U3)RNP and the Sm precipitate samples: the fainter bands of ~500 nt, ~324 nt (forming a band of equal intensity to snR17, and migrating just below it in the Sm samples), and ~277 nt in the anti-(U3)RNP lanes are also visible in the Sm lanes. These bands comigrated with m<sub>3</sub>G-capped RNAs of unknown identity.

No strong, specific bands appeared in the anti-(U1)RNP precipitates, and in particular, no band corresponding to the *S. cerevisiae* U1, snR19 (at "2"), was visible.

Figure 5.2 illustrates the same RNA samples as those shown in Figure 5.1 transferred to a filter and probed with snR30-, scR1-, snR17- and snR3-specific probes.<sup>1</sup> The presence of snR17 in the anti-(U3)RNP and the Sm lanes was confirmed, as was the presence of a faint snR30 band in the anti-(U3)RNP lanes. ScR1, visible in the total nuclear extract RNA (Figure 5.2, lanes 1-4) was not found to be precipitated by any of the antisera used, and served as an internal control.

<sup>&</sup>lt;sup>1</sup>A plasmid probe for snR3 (Tollervey *et al.* 1983) was kindly provided by David Tollervey.

The fact that human anti-(U3)RNP serum can precipitate yeast RNP containing snR17 gives much weight to the hypothesis that snR17 is the yeast homologue of U3 snRNA. The observation that several other yeast snRNAs were also precipitated by this serum was interesting since, in human cell extracts, this serum appears to be U3 snRNP-specific (Lischwe *et al.* 1985). A surprising observation was that snR17 was Sm-precipitable: vertebrate U3 snRNP does not appear to contain any Sm antigens (Lerner and Steitz 1979, Lerner *et al.* 1980a & b, Lührmann 1988).

Figure 5.3 shows the results of two further experiments that confirm these observations. In Figure 5.3A and B, yeast nuclear extract RNA ("NE"), anti-(U3)RNP-precipitated RNA ("U3"), and anti-m<sub>3</sub>G cap-precipitated RNA ("CAP") and control samples ("CON") have been hybridized with probes specific to the RNAs indicated; the presence of snR30 and snR3 in the anti-(U3)RNP lanes was confirmed. Figure 5.3C represents an independent experiment which included also anti-(U1)RNP and Sm precipitates, and shows that snR17 is clearly visible in the Sm samples, whereas snR3, which is clearly visible in the anti-(U3)RNP lane, is completely absent from the Sm lane. Riedel *et al.* (1987) have confirmed that snR17 can be weakly, but not quantitatively, precipitated by Sm antiserum.

The various *S. cerevisiae* snRNAs immunoprecipitated by Sm and anti-(U3)RNP antisera are summarized in Table I.

The precipitability of yeast snRNAs by Sm antisera correlates well with the presence of Sm binding site sequences. The yeast spliceosomal RNA homologues all possess Sm binding site sequences (except U6/snR6, which, like its vertebrate counterpart, is combined with U4/snR14 in the same particle, Siliciano et al. 1987, Brow and Guthrie 1988) in regions predicted to be single-stranded (Guthrie and Patterson 1988) and are all precipitated by Sm antisera. SnR17 is consistently, although not quantitatively, precipitated by Sm antiserum and has a consensus Sm binding site sequence in a partially base-paired region. The main difference between the snR17 and U3 snRNA secondary structure models was the presence of an extra hairpin on snR17 (hairpin 4 in Figure 3.13); this is exactly the structure containing the Sm binding site sequence, and therefore conveniently accounts for the fact that snR17 is Sm-precipitable whereas vertebrate U3 snRNPs are not. The fact that the snR17 Sm binding site appears to be partially base-paired might account for the relative inefficiency with which snR17 is Sm-precipitated. An Sm binding site sequence with one mis-match to the consensus, AGUUUUGA, was identified in snR30, and this sequence also occurred in a partially base-paired region (see Appendix 1, Figure A1.2). A

very weak snR30 signal is discernable in the Sm lane of Figure 5.3C, suggesting that snR30 might associate very weakly with Sm protein. In contrast, no signal for snR3 could be seen in the Sm precipitates.

#### 2. Attempt to define U3 snRNA box C as an RNP protein binding site

U3 snRNAs share four regions of conserved nucleotide sequence, referred to as boxes A to D (see Chapter 3, Figure 3.12). Two of these, boxes B and C, are probably single-stranded (Figure 3.13) and therefore potentially available to interact with other molecules, either proteins or RNAs. Parker and Steitz (1987) have shown that fragments of RNA containing the box C region, as well as fragments extending from nucleotides 71 to 92 and 190 to 217 of human U3 (equivalent to approximately the same nucleotide numbers of rat U3 in Figure 3.13), are protected after digestion of the RNP with RNase, and subsequently can be precipitated with anti-(U3)RNP antiserum. The box C region is therefore a good candidate for a conserved U3 snRNP protein binding site.

Precipitability of yeast snRNPs by Sm antiserum appeared to correlate with the presence of a good Sm binding site in a single-stranded region of the molecule; in snR17, where an Sm binding site was predicted to be partially base paired, precipitation was less efficient. [The efficiency of precipitation can be judged to some extent by comparing the relative band intensities of each RNA in the Sm lanes compared to the anti-cap lanes of Figure 5.1.] In a similar way, at least for the three yeast snRNAs for which sequences are available, the efficiency of precipitation by anti-(U3)RNP antiserum appeared to correlate with the presence of a single-stranded U3 box C sequence: snR3 contains a sequence which resembles box C in an apparently single-stranded loop between the second and third stems of a clover leaf structure (Tollervey et al. 1983), and is efficiently precipitated by anti-(U3)RNP serum (Figure 5.3C); snR30 contains a close match to the box C sequence, but in a region which is partially base-paired, as illustrated in Figure A1.2, Appendix 2, and is relatively poorly precipitated by anti-(U3)RNP serum. Interestingly, the only other vertebrate U snRNA identified in the nucleolus, U8, also contains a box C-like sequence, but also in a partially base-paired region (Reddy et al. 1985). The precipitability of U8 snRNP with anti-(U3)RNP has not been tested. Figure 5.4 shows a comparison of all these box C-like sequences.

In order to test the hypothesis that a box C sequence might be sufficient to allow binding of the U3 snRNP antigen to an RNA, representative box C sequences were introduced into T7 transcripts, which, labelled, were then injected into *Xenopus* oocytes; their ability to bind antigen was assessed by subsequent immunoprecipitation.<sup>1</sup> This approach was used by Mattaj (1986) to show that the Sm binding site directed the association of Sm proteins with RNA: T7 transcripts containing Sm binding sites, after injection into Xenopus oocytes, could be precipitated with Sm antiserum, and furthermore, once these transcripts had bound Sm proteins, they were subsequently trimethylguanosine-capped and transported into the nucleus.

The representative box C sequences chosen for the experiment were those of *Xenopus laevis* U3 snRNA and *Saccharomyces cerevisiae* snR17. Two pairs of complementary oligodeoxynucleotides were synthesised, consisting of the two box C sequences flanked by *Eco*RI restriction site "sticky-ends". The annealed oligos were inserted into the *Eco*RI site of the polylinker sequence of a "Bluescript KS" plasmid cloning vector, which has a T7 RNA polymerase promoter in the polylinker flanking sequence enabling transcripts to be synthesised *in vitro*. The sequences of the oligonucleotides are shown below, the box C regions are underlined (*q.v.* Figures 3.12 and 5.4).

Oligo 180 Oligo 181	5′	$\begin{array}{c} \text{AATTCGC} \underline{\text{GATGATCTT}} \text{GACC}^{3'} \\ \underline{3}'  \text{GCGCTACTAGAACTGGTTAA}_{5'} \end{array}$	Yeast
Oligo 182 Oligo 183	5´	AATTATT <u>GATGAACGT</u> TCTG <sup>3'</sup> 3' TAACTACTTGCAAGACTTAA <sub>5'</sub>	Frog

Plasmid clones containing insertions of the oligonucleotides were isolated as described in Chapter 6, Section 6.3 and their structures were confirmed by DNA sequencing. Two of the clones, pSRBC5 and pSRBC11, contained single insertions of the yeast sequence in the forward orientation with respect to the T7 promoter (as shown above for oligos 180/181); of these, owing to an imperfectly synthesised batch of oligonucleotides, pSRBC11 contained a single base deletion of the first "A" in the box C sequence. A third clone, pSRBC27, contained four insertions of the *Xenopus* sequence in the forward orientation, and a fourth clone, pSRBC1, for use as a negative control, contained three insertions of the yeast sequence in the reverse orientation. The structures of the transcripts generated from these clones are shown in Figure 5.5 and in the legend to Figure 5.6.

 $^{32}$ P-labelled T7 polymerase transcripts were generated *in vitro* as described in Chapter 6, Section 5.8. The transcripts were synthesised with a 5' monomethyl guanosine cap structure by priming the transcription with a m<sub>3</sub>GpppG dinucleotide (Contreras *et al.* 1982, Konarska *et al.* 1984, Chen *et al.* 

151

<sup>&</sup>lt;sup>1</sup>Microinjection into, and immunoprecipitation from *Xenopus* oocytes was performed by Iain Mattaj.

1987). The transcripts were generated on plasmid DNA templates linearized at the Xhol site at the end of the polylinker distal to the promoter. The quality and sizes of the transcripts were checked by gel electrophoresis and autoradiography. Transcripts were then micro-injected into Xenopus oocytes, and a selection of different antisera were used to immunoprecipitate cell extracts prepared from the oocytes after incubation for up to 24 hrs. In vitro-synthesised U1 snRNA was included among the transcripts as a positive control for immunoprecipitation by Sm antiserum; as negative controls, a transcript either from pSRBC1, containing three insertions of the box C sequence in the reverse orientation, or from the Bluescript plasmid without insertion were used (Figure 5.5 and legend to Figure 5.6). Total RNA was extracted from oocytes immediately after injection and after over-night incubation in order to assess the stability of the transcripts in vivo. pSRBC27 always gave two transcripts of apparently different lengths; these have been labelled "27" and "27" in Figures 5.6 and 5.7. The shorter transcript, presumably a result of premature termination, was longer than 100 nt and therefore probably contained at least two insertions of the Xenopus box C sequence.

The results of the first injection experiment are shown in Figure 5.6. No precipitation of any of the transcripts by scleroderma anti-(U3)RNP serum was seen. However, the experiment did indicate that transcripts containing box C sequences were recognised by proteins in a specific manner, because these transcripts, along with U1 snRNA, but not the negative control, were all precipitated by anti-m<sub>3</sub>G cap antiserum. *In vivo*, snRNAs are initially synthesised with a monomethyl-cap structure; trimethylation is cytoplasmic and is dependent on Sm protein binding (demonstrated for U2 snRNA, Mattaj 1986). It is likely therefore that specific protein binding is a prerequisite for trimethyl-capping of U3 snRNAs, which do not normally bind Sm proteins.

Lanes 1 and 2 of Figure 5.6 show RNA extracted from oocytes at 0 and 16 hrs, respectively, after injection; the injected transcripts are visible in both these lanes and appear to be stable. Panels A, B, C and D represent separate batches of oocytes injected with different sets of transcripts; individual transcripts are labelled in the figure according to the number of the clones from which they were derived. No transcripts were visible in the anti-(U3)RNP precipitates (lanes 4). The anti-m<sub>3</sub>G cap serum (lanes 6) specifically precipitated the transcripts containing the box C sequences, that is, those from pSRBC5 ("5", in A), pSRBC11 ("11", in B), pSRBC27 (27 and 27', in C), and also U1 snRNA (present in A, B and C), but not the negative control from pSRBC1 ("1", in D). [Note that transcript "1" shows the same gel mobility as U1 snRNA.] Sm antiserum precipitates U1 snRNA alone (lanes 3, A, B and C). A fourth

antibody, "anti-54k" (lanes 5), appears to precipitate all the transcripts in a non-specific manner.<sup>1</sup>

In conclusion, the U3 snRNA box C sequences of both *Xenopus* and yeast, and also a yeast box C sequence with a deleted, conserved A, appear to be sufficient to direct trimethylation of monomethyl-capped RNAs in *Xenopus* oocytes. No indication that these sequences were sufficient for the binding of anti-(U3)RNP antigen was observed.

In order to confirm these results, and confirm that the anti-(U3)RNP antiserum was active, [the previous experiment consumed the last of the original anti-(U3)RNP serum, and its activity may have been reduced], a variety of different anti-(U3)RNP antibodies were obtained, and the abilities of these to precipitate T7 transcripts from oocytes, as well as to precipitate native U3 RNP from *Xenopus* and yeast cell extracts were tested.

The results of this experiment were inconclusive. Weak signals corresponding to the box C-containing transcripts were observed in samples precipitated with three anti-(U3)RNP antibodies, which were shown to precipitate native U3 RNP, and in the sample precipitated with the anti-m<sub>3</sub>G cap serum. It could not be shown confidently, however, that these signals were significantly above background. A fourth anti-(U3)RNP antibody precipitated box C-containing transcripts as well as U1 snRNA, but this antibody could not be shown to precipitate native U3 RNP.

Lanes 1 and 2 of Figure 5.7 show RNA extracted from *Xenopus* oocytes 0 and 24 hrs after injection, respectively. All the transcripts appeared to be stable except the negative control ("C"), which in this experiment was derived from the Bluescript plasmid without insertion. A, B, C and D correspond to different batches of oocytes injected with the sets of transcripts indicated; the lower panels show longer exposures of the gels in the upper panels. The most prominent bands visible correspond to U1 snRNA and several U1 break-down products (best identified in Figure 5.7D). The three antibodies shown to precipitate native U3 snRNP are represented in lanes 5, 6 and 8: faint bands corresponding to the box

<sup>&</sup>lt;sup>1</sup>The anti-54k antibody is a mouse monoclonal against the 54kD polypeptide of canine signal recognition particle, obtained from Bernard Dopperstein. It was first used in yeast as part of an experiment to test whether scR1 could be immunoprecipitated using anti-SRP antibodies (which was completely negative). Remarkably, this antibody did precipitate both snR17 and snR30 (revealed by northern blotting), along with a range of other RNAs, including 5.8S rRNA (revealed by <sup>32</sup>pCp labelling) but not scR1. The significance of this is not understood. The fact that this antibody precipitates the negative control shows that, for the purposes of the experiment in Figure 5.6, it can be considered to be "non-specific".

C-containing transcripts "5", "11" and "27" are visible in the longer exposures of these lanes (Figure 5.7A, B and C, lower panels). These bands are also visible in the anti-m<sub>3</sub>G cap precipitates (lanes 4). Very faint signals corresponding to these transcripts, however, also appear in the Sm lanes (Figure 5.7, lanes 3, lower panels), suggesting that the signals in the anti-(U3)RNP and anti-m3G cap lanes are not significantly above background. The negative control (C) was neither visible in these anti-(U3)RNP- (lanes 5, 6 and 8) nor in the anti-m<sub>3</sub>G cap-precipitated samples (lanes 4), but this transcript appeared to be less stable in the oocytes than the other transcripts after 24 hrs incubation (Figure 5.7C, lanes 1 and 2, upper panel); its absence therefore cannot indicate conclusively a specific reaction of antigen with box C-containing transcripts. U1 snRNA was precipitated by Sm antiserum (lanes 3) and anti-m<sub>3</sub>G cap serum (lanes 4) as expected, but also by the anti-(U3)RNP patient serum "knoblauch" (lane 5), and the autoimmune mouse anti-(U3)RNP serum "712" (lane 7). Anti-(U1)RNP antibodies have been previously observed in "knoblauch" serum (lain Mattaj, unpublished results). Faint U1 signals are visible for the anti-(U3)RNP serum "S1" (lanes 6) and the human anti-(U3)RNP monoclonal antibody "72B9"; the fact that the latter has been shown to react specifically with the U3 snRNP 36 kD polypeptide (Reimer et al. 1987) suggests that these signals are not specific. The reaction of the mouse antiserum 712 (lanes 7) with box C-containing transcripts (Figure 5.7A, B and C) was strong and may indicate a specific reaction, but these lanes also contain a weaker band corresponding to the negative control, and this serum could not be shown to precipitate native U3 snRNP particles.

In order to test whether the various antisera were able to precipitate native U3 or snR17 snRNP, RNA extracted from immunoprecipitates of *Xenopus* and yeast cell extracts was transferred to filters and probed by hybridization. The northern blots in Figure 5.8 show that the anti-(U3)RNP sera "knoblauch" and S1, and the human anti-(U3)RNP monoclonal antibody 72B9 (used in Figure 5.7, lanes 5, 6 and 8) precipitate native snR17 RNP from yeast (Figure 5.8A) and native U3 RNP from *Xenopus* (Figure 5.8B), whereas the mouse antiserum 712 (used in Figure 5.7, lanes 7) and various control sera, including Sm, do not.

The lack of precipitation of snR17 by Sm antiserum (Figure 5.8A, lane 2; compared to lanes 8 and 9 as controls) in this experiment is not consistent with the previous results in which snR17 RNP was observed to be Sm-precipitable (Figures 5.1, 5.2 and 5.3C). The differences between this and the previous experiments was that the Figure 5.8 samples were precipitated from whole cell extracts (presumably reflected by more RNA degradation in Figure 5.8 due to cytoplasmic vacuolar nucleases) whereas, previously, precipitation was from

154

crude nuclear extracts, and that, in Figure 5.8, heparin was not included in the immunoprecipitation mixture (the effects of heparin in these experiments were not studied). Figure 5.8A confirms that snR30 is co-precipitated with snR17 using anti-(U3)RNP antibodies "knoblauch" (lane 5) and 72B9 (lane 7).

In conclusion, it seems probable that the U3 snRNA box C sequence is not sufficient on its own to direct binding of proteins normally antigenic in the native U3 snRNP. This question cannot be resolved firmly, however, without a careful study of the complex antigen-specificities of the antibodies used. The results of the experiment shown in Figure 5.6 suggested that the box C sequence might be recognised specifically by proteins required for trimethylguanosine cap formation, but these results were not supported by the subsequent experiment hown in Figure 5.7.

#### DISCUSSION

This chapter has described how human antibodies reacting specifically with the 36 kD polypeptide of U3 snRNP can immunoprecipitate RNP from *Saccharomyces cerevisiae* containing a variety of m<sub>3</sub>G-capped snRNAs, including snR17, which has been shown to have significant structural similarities to U3 snRNA. Unlike mammalian U3 snRNP, snR17 was also precipitated by Sm antiserum; this correlates with the existence in snR17 of a potential Sm protein binding site which is absent from mammalian U3 snRNA. There appeared to be a correlation between the efficiency of precipitation of yeast snRNPs by anti-(U3)RNP serum and the presence of a conserved snRNA sequence, box C, in an apparently single-stranded region. When *in vitro* RNA transcripts were injected into *Xenopus* oocytes, the box C sequence alone was not sufficient to direct the binding of the 36 kD polypeptide.

#### 1. Precipitation of yeast snRNPs by anti-(U3)RNP serum

The result that human anti-(U3)RNP antibodies can immunoprecipitate yeast snRNPs, including snR17, was interesting in several ways: first, it shows that the U3 snRNP antigen is highly conserved; secondly, it provides strong support for the hypothesis that snR17 is functionally equivalent to U3 snRNA; and thirdly, it shows that *S. cerevisiae* contains a whole new class of snRNAs that share at least one common antigen, and that are distinct from spliceosome-associated snRNAs in this respect, and in that several of them are clearly located in the nucleolus in close association with pre-rRNA. These snRNAs presumably have functions concerned with pre-ribosomal assembly, transport or pre-rRNA processing.

Anti-(U3)RNP antiserum specific for the 36 kD polypeptide has been observed to precipitate only U3 snRNP from human cell extracts (Lischwe *et al.* 1985), but this might reflect the differences in relative abundance of snRNPs in mammalian and yeast cells: U1 and U2 snRNAs are represented by  $10^6$  and 5 x  $10^5$  copies, respectively, per mammalian cell, and U3 to U6 by up to 2 x  $10^5$ (Weinberg and Penman 1968, Reddy and Busch 1988); the copy number per cell of many yeast snRNAs has been estimated as being not more than 500 (Wise *et al.* 1983), although snR17 and snR30, the two most abundant yeast snRNAs, are probably represented by up to 1000 copies. SnRNAs of low abundance associated with the 36 kD U3 snRNP antigen in mammalian cells may yet be found.

The yeast immunoprecipitation experiments involved no attempts to resolve individual snRNP particles, so it is an open question as to whether each precipitated snRNA band represented an individual RNP containing the antigen, or 156

whether several snRNAs were coprecipitated as part of a complex association of RNPs, not all of which were antigenic. No attempt was made to find out whether snR17 could be quantitatively precipitated with anti-(U3)RNP serum: in Chapter 1 it was observed that a proportion of snR17 appeared to be tightly associated with nucleolar fractions whereas a smaller part was either cytoplasmic or readily released from nuclei; whether both these portions were antigenic is not clear. Parker and Steitz (1987) observed that only 10% of human U3 snRNP was immunoprecipitable; this might represent either a distinct population, or the proportion of 36 kD protein remaining stably bound after sonication.

#### 2. SnR17, but not U3 snRNP, is Sm-precipitable

SnR17 was observed to be weakly but consistently precipitated with Sm antiserum from yeast nuclear extracts (Riedel *et al.*, 1987, have confirmed this). Vertebrate U3 snRNP has not been observed to be Sm-precipitable. This difference, however, can be conveniently explained by the extra hairpin structure of snR17 (see Figure 3.13) which contains a consensus Sm protein binding site. SnR17 would not be the first nucleolar snRNP observed to be Sm-precipitable: U8 snRNP (Reddy *et al.* 1985), the only m<sub>3</sub>G-capped vertebrate nucleolar snRNP found to date, is also Sm-precipitable. SnR17 is possibly not the only yeast snRNP that shares both anti-(U3)RNP and Sm antigens, since three fainter RNA bands, observed in the anti-m<sub>3</sub>G cap lanes, were seen in both the anti-(U3)RNP and Sm lanes of Figure 5.1. The fact that the putative Sm binding site of snR17 appears to be partially base-paired in a stem structure could explain the relative inefficiency with which snR17 was Sm precipitated.

The significance of a possible association of Sm proteins with snR17 *in vivo*, in an inefficient manner with a sub-optimal binding site, is not clear. The fact that snR17 was Sm-precipitable from nuclear extracts in the presence of heparin, but was not obviously Sm-precipitable from whole cell extracts when heparin was not used, suggests some artifactual variability that must be explained before it may be concluded that snR17 binds Sm proteins *in vivo*. Sm proteins have been considered to be characteristic of nucleoplasmic snRNPs involved in pre-mRNA processing (Riedel *et al.* 1987). Sm protein binding takes place in the cytoplasm, and is required for efficient trimethylguanosine capping activity, and for transport of snRNPs back into the nucleus (Mattaj and De Robertis 1985, Mattaj 1986). SnRNAs acquire many modified nucleotides (Reddy and Busch 1988), and Sm proteins, by stabilizing snRNAs in specific conformations, could specify the sites of these reactions.

#### 3. U3 snRNA box C as a protein binding site

Of the several yeast snRNAs precipitated as RNP by anti-(U3)RNP serum, snR17, snR3 and snR30 shared a sequence in common, conserved in U3 snRNAs, called box C. SnR17 and snR3, which were among the prominent bands precipitated, contain this sequence in an apparently single-stranded loop between two stems. SnR30, which was only weakly precipitated, contains a box C-like sequence in a partially base-paired region of a stem (see Figure A1.2, Appendix 1), just as snR17, which was weakly Sm-precipitated, contains a partially base-paired Sm binding site.

When the box C sequence was injected into *Xenopus* oocytes, as part of a larger transcript of essentially "random" sequence, and RNP subsequently immunoprecipitated from cell extracts was analysed, no convincing precipitation of box C-containing transcripts was observed with antibodies that could be shown to precipitate native U3 snRNP. The conclusion of these results is that the box C sequence alone is not sufficient for the assembly of the 36 kD protein, nor other proteins normally antigenic in the native U3 snRNP. Box C was not the only RNA fragment found by Parker and Steitz (1987) to be immunoprecipitated by anti-36 kD serum after treatment of the RNP with RNase, suggesting that the 36 kD protein could associate with other proteins which bind to different parts of the RNA, more than one of which are required for 36 kD assembly.

Table I. Antigens associated with Saccharomyces cerevisiae snRNAs.

<u>snRNA</u>	<u>Size</u> ( <u>nt</u> )	Associated structures	<u>U3 snRNP</u> antigen	<u>Sm</u> antigen	Phenotype of deletion mutant
U2/snR20 <sup>1</sup> U1/snR19 <sup>1</sup> U5/snR7 <sup>1</sup> U4/snR14 <sup>1</sup> U6/snR6 <sup>1</sup>	1175 569 213 & 178 160 115	Spliceosome Spliceosome Spliceosome Spliceosome U4/U6 RNP		+ + + +	Lethal Lethal Lethal Lethal Lethal
U3/snR17 2 snR3 3 snR4 3 snR5 3 snR8 3 snR9 snR10 snR10 snR13 snR15 snR28 3 snR190 snR128	328 194 ~192 205 189 188 245 ~124 ~169 190 128	Pre-rRNA Pre-rRNA Pre-rRNA Pre-rRNA Pre-rRNA Pre-rRNA Pre-rRNA Pre-rRNA	+ + + + + + + +	(+)    	Lethal Viable Viable Viable Viable Sick Viable Lethal
x x x	~105 ~104 ~97		+ + +		
x snR18a? x snR11? snR16?	~500 ~393 ~324 ~277 ~257 ~200		+ + + + +	+  + 	
snR30 <sup>2</sup>	605		(+)		Lethal

<sup>1</sup> Presence in Sm-precipitates confirmed (Siliciano et al. 1987b) [U6 as part of U4/6 RNP (Brow and Guthrie 1988)].

<sup>2</sup> Presence in immunoprecipitates confirmed by hybridization (this work).

<sup>3</sup> Presence in immunoprecipitates confirmed by two dimensional gel electrophoretic analysis (D. Tollervey, unpublished).

Other immunoprecipitated snRNAs have been identified tentatively by their migration patterns on one-dimensional gels.

x = RNA band comigrating with unidentified band in anti-cap-precipitated samples of Figure 5.1.

References to yeast snRNAs:

U1/snR19 — Kretzner et al. 1987, Siliciano et al. 1987b; U2/snR20/LSR1 - Ares 1986, Parker et al. 1987; U4/snR14 - Siliciano et al. 1987a; U5/snR7 - Patterson and Guthrie 1987; U6/snR6 - Siliciano et al. 1987a, Brow and Guthrie 1988; U3/snR17 - this work, Hughes et al. 1987; snR30 — this work, Bally et al. 1988; snR3 — Tollervey et al. 1983; snR3, snR4, snR5, snR8 and snR9 — Parker et al. 1988; snR10 — Tollervey 1987a; snR128 and snR190 - Zagorski et al. 1988; other snRNAs - Riedel et al. 1986.

#### Figure 5.1

#### RNA immunoprecipitated from yeast nuclear extracts, end-labelled

• Yeast RNA, fractionated on a 6% polyacrylamide/urea gel, was 3' end-labelled after immunoprecipitation from either sonicated nuclear extract (NE) (see Chapter 6, Section 6) or naked nuclear extract RNA (RNA), with different antisera bound to protein A-sepharose beads. Precipitations (in duplicate) were with CON, control of beads without antiserum; CAP, rabbit anti-m3G cap serum; U3, scleroderma patient anti-(U3)RNP serum; U1, patient anti-(U1)RNP serum; and Sm, lupus erythematosus patient Sm antiserum. N: complete nuclear extract RNA. RNA species were tentatively identified by their gel migration patterns as foliows: 1, LSR1/snR20 = U2 (1175 nt); 2, snR19 = U1 (569 nt); 3, snR17 identified by hybridization = U3 (328 nt); 4, snR7L = U5 (213 nt); 5, five or more snRNAs including snR3, snR4, snR5 and snR28 (180-190 nt); 6, snR7S = U5 (178 nt); 7, snR15 (169 nt); 8, snR14 = U4 (160 nt); and 9, snR13 (124 nt). The quantities of antisera bound to protein A-sepharose beads for each sample

The quantities of antisera bound to protein A-sepharose beads for each sample were:  $20\mu$ l rabbit anti-m<sub>3</sub>G cap serum;  $1 \mu$ l scleroderma anti-(U3)RNP serum;  $5 \mu$ l anti-(U1)RNP serum; and  $10 \mu$ l Sm serum.



#### Figure 5.2

# RNA immunoprecipitated from yeast nuclear extracts, identified by hybridization

The same samples and quantities of RNA shown in Figure 5.1 were fractionated on a denaturing polyacrylamide gel, electro-transferred to a Genescreen filter and probed with labelled plasmids encoding genes for snR30, snR17, snR3 and scR1 as control. Lanes 1, 2, 3 and 4 contain 1.6, 0.8, 0.4 and 0.2  $\mu$ g respectively of nuclear extract RNA. The quantities of RNA loaded in the immunoprecipitate lanes were all less that 0.2  $\mu$ g.

Sm CON 2 3 4 CON CAP U3 U3 CAP U1 RNA snR17 > • snR30> scR1> ~ snR3 • 

#### Figure 5.3

## Both snR30 and snR3, as well as snR17, are precipitated by anti-(U3)RNP serum

Panels A and B show the same filters probed with A: the indicated small RNA probes, and B: a rDNA probe as control. Panel A shows clearly that snR30 and snR3 were enriched in the anti-(U3)RNP precipitates. Panel C shows an independent experiment in which snR17 was clearly Sm precipitated, whereas snR3 was not. NE, nuclear extract RNA; CON, control of beads without antisera; CAP, anti-m<sub>3</sub>G cap serum; U3, anti-(U3)RNP serum; U1, anti-(U1)RNP serum; Sm, Sm antiserum. The same quantities and samples of antisera were bound to protein A-sepharose beads as in Figure 5.1.<sup>1</sup>

<sup>&</sup>lt;sup>1</sup>Efficient precipitation of RNP with anti- $m_3G$  cap serum appeared to require longer incubation times than with anti-(U3)RNP serum. Efficient precipitation with anti-(U3)RNP serum, however, appeared to be sensitive to warming above 4°C (such as by over-sonication). The inter-play of these factors presumably explained the varying relative precipitation efficiencies in different experiments.





## HYPOTHETICAL BINDING SITE FOR:

	anti-(U3)RNP antigen	Sm antigen
Human U3	CAGCCAUU <u>GAUGAUCGU</u> UCUUCUC	
Rat U3A	UGGCCAUU <u>GAUGAUCGU</u> UCUCGGU	
Rat U3B	UUGCCAUU <u>GAUGAUCGU</u> UCUUCUC	
X. laevis	UUGCUAUU <u>GAUGA</u> A <u>CGU</u> UCUGCUC	
X. borealis	UUGCUAUU <u>GAUGA</u> A <u>CGU</u> UCUACUU	
D. discoideum	UUAACAUU <u>GAUGA</u> C <u>CGU</u> CUAAUU	
Bean U3	CAGCAGUA <u>GAUGAUCGU</u> UCGGAGU	
snR17A	UAUGGCGC <u>GAUGAUC</u> U <u>U</u> GACCCAU	<u>AAUUUUGA</u>
snR17B	AAUGGCGC <u>GAUGAUC</u> U <u>U</u> GACCCAU	<u>AAUUUUGA</u>
snR3	UUAUUCGU <u>GAU</u> A <u>A</u> G <u>CGU</u> AAUGUGG	
snR30	UAGGACGC <u>AUGAUC</u> U <u>U</u> GAGCUCU	<u>AGUUUUGA</u>
Rat U8	GAGCUUGC <u>AUGAUC</u> UGCUGAUU	<u>GAUU</u> A <u>U</u> A <u>G</u>

#### Figure 5.4 Hypothetical RNA protein binding sites

The conserved box C regions are shown of U3 snRNAs from human, rat, *Xenopus*, *Dictyostelium*, bean and of *S. cerevisiae* snR17, along with a similar region in snR3, which also occurs as a single-stranded loop, and regions in snR30 and rat U8 small nucleolar RNA which resemble the box C sequence, but form partially base-paired portions of stems. U3 snRNPs, snR17 and snR3 clearly react with anti-(U3)RNP antisera; snR30 reacts less efficiently, and U8 snRNA has not been tested. Also shown are the putative Sm binding sites of snR17, present on an extra hairpin which is absent from other U3 snRNAs, of snR30 and of U8 snRNA. U8 snRNP is Sm-precipitable, and its Sm binding site occurs in a single-stranded region. SnR17 is weakly Sm-precipitable, and snR30 is possibly very weakly Sm-precipitable: the Sm binding sites of these molecules form partially base-paired regions.

#### Bluescript, no insertion

- 1 GGAGCUCCACCGCGGUGGCGGCCGCUCUAGAACUAGUGGAUCCCCCGGGC
- 51 UGCAGGAAUUCGAUAUCAAGCUUAUCGAUACCGUCGACCUCGA (93 nt)

#### pSRBC5, 1 yeast insertion, forward

- 1 GGAGCUCCACCGCGGUGGCGGCCGCUCUAGAACUAGUGGAUCCCCCGGGC
- 51 UGCAGGAAUUCGC<u>GAUGAUCUU</u>GACCAAUUCGAUAUCAAGCUUAUCGAUA
- 101 CCGUCGACCUCGA (113 nt)

#### pSRBC11, 1 yeast insertion, deleted A, forward

- 1 GGAGCUCCACCGCGGUGGCGGCCGCUCUAGAACUAGUGGAUCCCCCGGGC
- 51 UGCAGGAAUUCGC<u>GUGAUCUU</u>GACCAAUUCGAUAUCAAGCUUAUCGAUAC
- 101 CGUCGACCUCGA (112 nt)

#### pSRBC27, 4 frog insertions, forward

- 1 GGAGCUCCACCGCGGUGGCGGCCGCUCUAGAACUAGUGGAUCCCCCGGGC
- 51 UGCAGGAAUUAUU<u>GAUGAACGU</u>UCUGAAUUAUU<u>GAUGAACGU</u>UCUGAAUU
- 101 AUU<u>GAUGAAGCU</u>UCUGAAUUAUU<u>GAUGAACGU</u>UCUGAAUUCGAUAUCAAG
- 151 CUUAUCGAUACCGUCGACCUCGA (173 nt)

#### pSRBC1, 3 yeast insertions, reverse

- 1 GGAGCUCCACCGCGGUGGCGGCCGCUCUAGAACUAGUGGAUCCCCCGGGC
- 51 UGCAGGAAUUGGUCAAGAUCAUCGCGAAUTGGUCAAGAUCAUCGCGAAUU
- 101 GGUCAAGAUCAUCGCGAAUUCGAUAUCAAGCUUAUCGAUACCGUCGACCU
- 151 CGA (153 nt)

#### Figure 5.5 Sequences of T7 transcripts injected into *Xenopus* oocytes.

The primary structures of the T7 transcripts used in the experiments described in Figures 5.6 and 5.7 are shown. The transcripts contain single or multiple insertions of either the *Xenopus* or the *S. cerevisiae* U3 snRNA conserved box C sequences, in forward (underlined) or reverse orientations. The plasmid pSRBC11 contains a yeast sequence in which a single base deletion occurred. The transcript from pSRBC1 was used as a negative control in Figure 5.6, and that of Bluescript without insertion in Figure 5.7.

#### Figure 5.6 Immunoprecipitation of T7 transcripts from *Xenopus* oocytes

Xenopus oocytes were injected with mixtures of the transcripts shown diagrammatically below as indicated in the figure labels opposite. 5, 11 etc.. refer to the plasmid names pSRBC5, pSRBC11 etc.. from which the RNAs were transcribed. Two forms of 27 were always transcribed, the full-length 27 and a shorter 27'.



Total RNA was extracted from oocytes 0 (lanes 1) and 16 hrs (lanes 2) after injection. Oocyte extracts prepared after 16 hrs were precipitated with Sm antiserum (lanes 3), anti-(U3)RNP antiserum (lanes 4), monoclonal antibody against dog SRP 54k polypeptide (lanes 5), and anti- $m_3G$  cap serum (lanes 6). Injections into oocytes, preparation of oocyte extracts and immunoprecipitations from oocyte extracts were performed by lain Mattaj.



#### Figure 5.7 Immunoprecipitation of T7 transcripts from Xenopus oocytes

Mixtures of the same transcripts shown in the legend to Figure 5.6 were injected into oocytes as shown by the labels of the figure. Total RNA extracted from oocytes 0 and 24 hrs after injection is shown in lanes 1 and 2, respectively. Oocyte extracts were precipitated after 24 hrs with (quantities bound to beads in parentheses) Sm antiserum (2  $\mu$ l), lane 3; anti-m<sub>3</sub>G cap serum (50  $\mu$ l), lane 4; "Knoblauch" anti-(U3)RNP serum (10  $\mu$ l), lane 5; "S1" anti-(U3)RNP serum (50  $\mu$ l), lane 6; "712" mouse autoimmune anti-(U3)RNP serum (50  $\mu$ l), lane 7; and "72B9" human monoclonal anti-(U3)RNP antibody (1 ml cell supernatant), lane 8.

Injections into oocytes, preparation of oocyte extracts and immunoprecipitations from oocyte extracts were performed by Jain Mattaj.

 $\langle$ 



Figure 5.8 Immunoprecipitation from yeast and *Xenopus* whole cell extracts of native U3 snRNP

Whole yeast spheroplasts were sonicated in IP buffer and clarified extracts were used for immunoprecipitation as described in Chapter 6, Section 6. RNA extracted from immunoprecipitates was blotted to filters and probed with labelled plasmids encoding yeast small RNAs (A). A *Xenopus* cell line was treated in a similar manner after harvesting the cells and washing in PBS. After immunoprecipitation, the RNA was probed with a 5' end-labelled oligonucleotide complementary to the 5' end of vertebrate U3 snRNAs (B). Lanes 1 contain total cell extract RNA. Precipitations were with Sm antiserum (10  $\mu$ l), lanes 2; anti-m<sub>3</sub>G cap serum (70  $\mu$ l), lanes 3; "712" mouse autoimmune anti-(U3)RNP serum (100  $\mu$ l), lanes 4; "Knoblauch" anti-(U3)RNP serum (25  $\mu$ l), lanes 5; "S1" anti-(U3)RNP serum (200  $\mu$ l), lanes 6; "72B9" human monoclonal anti-(U3)RNP antibody (10 ml cell supernatant), lanes 7; "H125" human anti-SRP serum (100  $\mu$ l) lanes 8; and human non-immune serum as control, lanes 9.

5



## Chapter 6

### Materials and Methods

Sections 1 to 6 of this chapter contain a selection of mainly routine experimental techniques; less routine methods have been described in the text or in figure legends. Section 7 contains a list of buffers, media and reagents. Other general practical techniques are presented by Berger and Kimmel (1987), Boulnois (1987), Glover (1985), and Maniatis *et al.* (1982).

All percentage figures are expressed as weight/volume unless specifically indicated as volume/volume (v/v) or weight/weight (w/w).

Section 1

Yeast methods

#### **1.1 CULTURE OF YEAST CELLS**

YPD rich glucose medium:

1% Bacto Yeast Extract
 2% Bacto Peptone
 2% Dextrose (glucose)
 For plates: 2% Bacto Agar

[YPG is the same as this with 3% (v/v) glycerol instead of glucose.] Population doubling time in YPD at 30°C is approximately 75' in midexponential growth phase. Inoculating  $50\mu$ I of saturated culture into 800 ml YPD gives a mid-exponential phase (O.D.<sub>600</sub> = 1.00) culture after 15 hrs.

Growth curve for wild type Saccharomyces cerevisiae strain D-273-10B

<u>O.D.<sub>600</sub></u>	<u>Cells/ml</u>
0.20	5 x 10 <u>6</u>
0.30	1 x 10 <u>7</u>
0.50	2 x 10 <sup>7</sup>
1.00	6 x 10 <sup>7</sup>
1.30	1 x 10 <sup>8</sup>
saturated	3 x 10 <sup>8</sup>

SD selective dextrose medium:

0.67% Bacto-yeast nitrogen base w/o amino acids 2% glucose (dextrose) amino acid supplements as required (see below) For plates: 2% agar

Yeast grows more rapidly if all amino acids listed are used, omitting only the selected amino acid. All stock solutions can be autoclaved, except those indicated.

<u>Constituent</u>	<u>Final</u> mg/l	<u>Stock solution,</u> per_200_ml	<u>ml stock to add</u> to 600ml
adenine sulphate	20	240 mg*	10
uracil	20	480 mg*	5
L-tryptophan	20	480 mg	5
L-histidine-HCI	20	480 mg	5
L-arginine-HCI	20	480 mg	5
L-methionine	20	480 mg	5
L-tyrosine	30	180 mg*	20
L-leucine	30	720 mg*	5
L-isoleucine	30	720 mg	5
L-lysineHCl	30	720 mg	5
L-phenylalanine	50	600 mg*	<b>10</b> 7
L-glutamic acid	100	1.2 g*	10
L-aspartic acid	100	800 mg*+	15
L-valine	150	3.6 g	5
L-threonine	200	4.8 g+	5
L-serine	375	9 g	5

\*Store at room temperature

+Add after autoclaving the media

J.

#### 1.2 PREPARATION OF SPHEROPLASTS

Sherman, F., Fink, G.R. and Hicks, J.B. (1983) Methods in Yeast Genetics, Cold Spring Harbor Laboratory, New York.

Solutions and Reagents

β -(1,3)-D-glucanase:	e.g."Zymolyase," Kirin Brewery Co. Ltd., Tokyo, or "Cytohelicase", Reactifs IBF,
	France. Dissolve in water at least 3 hrs.
	before use (0.22 g in 1 ml) Stable for several weeks at 4°C.
TE:	10mM Tris pH 8.0, 1mM EDTA
Spheroplasting solution:	1M Sorbitol, 0.1M Na-Citrate pH 7.4, 10mM EDTA
β-mercaptoethanol:	2% (v/v) solution in water (0.27M)

Method

- 1. Grow the cells to mid-exponential phase at  $30^{\circ}$ C In 800 ml YPD (up to  $O.D_{.600} = 1.30$ , otherwise the cell walls become more difficult to digest. The cell wall is easier to remove after growth in galactose, which can be substituted for glucose in YPD, but growth is slower.)
- 2. Spin the cells at 1000 rpm in a Sorvall centrifuge for 5'.
- 3. Resuspend the pellet in 20 ml TE. Spin at 1000 rpm for 5'.
- Resuspend the pellet in 20 ml of water containing 2% (v/v) βmercaptoethanol. Incubate for 15' at room temperature. Spin at 1000 rpm for 5'.
- 5. Resuspend the pellet in 20 ml of spheroplasting solution. Add 0.5 ml (0.1 mg) of cytohelicase and incubate for 30'-60' at 30°C. Check for the formation of spheroplasts during the digestion by adding 2 drops of the cell suspension with a Pasteur pipette to 1 ml of 1% SDS. Lysis is indicated by clearing of the suspension and increased viscosity due to the release of DNA, thus showing the digestion to be complete.
- 6. Spin down the spheroplasts at 1000 rpm for 4' in a Beckmann bench-top centrifuge at 4°C.
- 7. Gently resuspend the pellet in sph. soln. (4°C) and spin twice more.

#### **1.3 PREPARATION OF CRUDE NUCLEI**

Method after Schulz (1978).

Solutions

NucA:	18% Ficol (w/w) Sigma Type 400 dissolved in
	20 mM K-phosphate pH 6.5, 0.5 mM CaCl <sub>2</sub> , 1.7
	mM PMSF - Add 0.6 ml 5% (0.28 M) of stock
	solution in ethanol to 100 ml immediately before
	use.
NucB:	7% (w/w) Ficol, 20% (w/w) Glycerol dissolved
	in 20mM K-phosphate pH 6.5, 0.5 mM CaCl <sub>2</sub> ,
	1.7 mM PMSF added immediately before use.
Nuc Storage Solution:	0.6 M Sucrose, 50 mM Tris pH 7.4, 10 mM
-	NaCl, 1 mM MgCl <sub>2</sub> , 1 mM β-mercaptoethanol.

Method

All steps to be performed at 4°C.

1. Thoroughly resuspend a pellet of yeast spheroplasts, prepared from 800 ml of culture, in 40 ml NucA. Allow the spheroplasts to lyse osmotically

while standing on ice for 5'. Lysis should be complete if the spheroplasting was efficient.

- 2. Spin out any remaining whole cells (ideally none) at 2000 rpm for 4' and resuspend them again in NucA to induce lysis.
- Spin the lysate (clarified of whole cells) at 30 000 x g for 25' (Sorvall SS34, 16000 rpm) Discard the supernatant.
  If highly purified nuclei are required, at this point the pellet may be resuspended and loaded onto Percoll gradients, as described by Ide and Saunders (1981).
- 4. Resuspend the pellet in 2 ml NucB until it is a homogeneous paste, then make up the volume with NucB to 20 ml. (The pellet is sticky and harder to resuspend in the larger volume.)
- 5. Spin at 20 000 x g for 25' (Sorvall SS34, 13 000 rpm.) Discard the supernatant material.
- 6. Repeat steps 4 and 5.
- 7. The pellet of crude nuclei may be stored at -70°C after first resuspending it in a small volume (0.5 ml) of Nuc Storage Soln.

#### 1.4 YEAST GENOMIC DNA MAXIPREP

This method is designed to minimize the number of pipetting steps, and thus reduce shearing of the DNA, giving a good yield of high molecular weight DNA suitable for constructing genomic libraries. Modified from Bhargava *et al.* (1972).

#### Solutions

10 x TE:	100 mM Tris pH 8.0, 10 mM EDTA
NaCI/EDTA:	1M NaCl, 10 mM EDTA, pH 8.0
10% SDS	•
1x SSC:	0.15M NaCl, 15 mM Na-citrate, pH 7.0
Protease:	20 mg/ml Sigma Type 111 No. P-4880
RNase:	20 mg/ml DNase-free

#### Method

- 1. Prepare a pellet of crude yeast nuclei, (from 800 ml of culture)
- 2. Suspend the nuclei in 4 ml of 10 x TE.
- 3. Transfer the suspension to a dialysis bag immersed in 1 litre of NaCl/EDTA at 55°C, leaving the top end of the bag open, hanging over the lip of the beaker.
- 4. Add 1 ml of 10% SDS to the contents of the bag and incubate the bag at 55°C for 20'.
- 5. Transfer the bag to 1 litre of 1x SSC at 37°C.
- 4. Add to the bag 0.5 ml RNase, incubate at 37°C for several hours.
- 5. Add 1 ml protease to the bag, incubate at 37°C overnight.
- 6. Carefully transfer the contents of the dialysis bag to a calibrated, screwcapped tube and make up the volume to 10 ml with 1 x TE. Gently mix.
- 7. Dissolve 50 g of CsCl in 30.5 ml of TE, then gently mix the CsCl solution with the 10 ml of DNA suspension.
- 8. Divide between two Quickseal tubes. Spin in a Beckman VTi 50 rotor at 45000 rpm overnight.
- 9. Puncture the bases of the tubes and collect about 20 fractions, drop-wise from each. (The DNA-containing fractions are more viscous.)
- 10. Analyse about 2 μl of each fraction on a 0.6% agarose gel. Pool the DNAcontaining fractions and dialyse at 4°C against several changes of TE buffer.

#### 1.5 YEAST GENOMIC DNA MINIPREP

Method described by Baldari and Cesareni (1985), adapted from Schulz (1978). Useful for preparing DNA to be analysed by Southern blotting.

Solutions	
SED:	1M Sorbitol; 25mM EDTA pH 8.0; 50mM β- mercaptoethanol (0.4% v/v).
SCE:	1M Sorbitol; 0.1M Na-citrate pH 5.8; 10mM EDTA.
Cytohelicase:	100 mg/ml in water.
NucA:	18% (w/w) Ficoll in 20mM K-phosphate pH 6.5, 0.5mM CaClo
Saline EDTA:	0.15M NaCI, 0.1M EDTA pH 8.
TE/SDS:	2% SDS, 0.2M Tris pH 8.8, 25mM EDTA.
K-OAc:	8M KOAc.
TE:	10mM Tris pH 8.0. 1mM EDTA.

#### Method

- 1. Set up 10 ml yeast cultures in YPD and incubate overnight at 30°C.
- 2. Spin down the cells at 2000 rpm for 4'. Discard the supernatant material.
- 3. Resuspend each pellet in 10 ml of water. Spin down again.
- 4. Resuspend each pellet in 1 ml of SED and transfer to a 1.5 ml Eppendorf tube. Incubate for 10' at 30°C.
- 5. Spin down in an Eppendorf centrifuge for 2". Discard the supernatant.
- 6. Resuspend the cells in 1 ml SCE and add 0.1 ml Cytohelicase. Digest for 45' at 30°C, then spin down the spheroplasts for 2''. Discard the supernatant.
- 7. Lyse the spheroplasts by resuspending them in 1 ml of NucA, mixing well. Allow to stand on ice for 5'.
- 8. Spin for 5' and discard the supernatant. The pellet should contain intact nuclei which are stable in the Ficoll soln.
- 9. Resuspend the pellet in 0.25 ml saline EDTA then add 0.25 ml TE/SDS. Mix well. Add 75  $\mu$ l of K-OAc and allow to stand on ice for 5'.
- 10. Spin for 5'. Retain the supernatant and add it to 0.3 ml of isopropanol. Allow to precipitate on ice for 5'.
- 11. Spin for 10'. Wash the pellet with 1 ml of 70% (v/v) ethanol, dry under vacuum, and resuspend in 0.1 ml of TE. 5  $\mu$ l of this preparation should contain sufficient DNA to detect single copy genes on Southern blots.

#### **1.6 EXTRACTION OF YEAST RNA**

Yeast RNA can be effectively extracted from whole cells (or subcellular fractions) be heating at 65°C in SDS/phenol. See Rubin (1975).

- 1. Grow a yeast culture to mid-exponetial phase  $(O.D._{600} = 1.0)$
- 2. Wash once in water.
- 3. Resuspend the cells in 10mM Tris pH 7.5, 10mM EDTA, 0.5% SDS (at least 5 ml per gram wet weight of cells) and add an equal volume of water-saturated phenol.
- 4. Vortex vigorously and incubate at 65°C for 30' (vortex occasionally during this time).
- 5. Place on ice for 5'.
- 6. Spin at 3000 rpm for 5' at 4°C.
- 7. Re-extract the supernatant material successively with cold phenol, 1:1 phenol/chloroform, and chloroform.
- 8. Bring the supernatant to 0.3M NaOAc, and add 2.5 vols of cold ethanol in order to precipitate the RNA.
- 9. Wash the pellet with 70% ethanol, dry under vacuum, and resuspend in water.
- 10. Determine the concentration by assuming  $O.D._{260} 1.0 = 45 \mu g/ml$ . The yield from 800 ml of culture should be at least 25 mg.
### 1.7 YEAST TRANSFORMATION - LITHIUM ACETATE METHOD

After Altherr et al. (1983).

Method

- 1. Grow 50 ml of yeast culture in YPD at 30°C with good aeration up to O.D.<sub>600</sub> 1.5-2.0
- 2. Harvest the cells by centrifugation at 2000 rpm in a bench-top centrifuge and wash the cell pellet twice with 10 ml of TE buffer (10mM Tris-HCI pH 7.5, 0.1 mM EDTA).
- 3. Resuspend the pellet in 5 ml of 0.1M lithium acetate in TE and incubate for 1 hr at 30°C with agitation.
- 4. Centrifuge, then resuspended in an equal volume of 0.1M lithium acetate in TE containing 15% glycerol. Dispense into aliquots and store at -70°C. Aliquots can be thawed out as required. Alternatively, use fresh cells which give a higher transformation efficiency.
- 5. Add 0.3 ml of competent cells to a sterile Eppendorf tube containing 10 μl (0.1-1.0 μg) of plasmid DNA.
- 6. Add 0.7 ml of 50% polyethylene glycol (PEG 4000), invert the tube three times, then incubate at 30°C for 1 hour.
- Provide a heat-shock at 42°C for 5<sup>-</sup>. Spin the cells in a microfuge for 4", resuspend in 100 μl of sterile water and plate directly onto selective media.

### 1.8 MATING, SPORULATION and TETRAD ANALYSIS

#### Mating

- 1. Using a wire loop, mix two colonies of opposite mating type on a small patch of a YPD plate. Incubate overnight at 30°C.
- 2. Streak out onto selective plates to obtain single colonies of diploids.

### Sporulation

Sporulation occurs when aerobically growing  $\alpha/a$  diploid cells are starved of nitrogen. For laboratory cultures, this is best achieved by growing cells in rich medium, with good aeration, and then transferring to nitrogen-starvation medium, containing a non-fermentable C-source, such as acetate, which does not repress respiration. Sporulation requires that the pH of the medium be able to rise to 9; the medium is therefore best adjusted to pH 7 but not buffered (for a review, see Haber and Halvorson 1975). Many methods for sporulation are available, some may suit certain strains better than others. The following has been found to be simple and very efficient:

- 1. Grow the diploid strain in YPD with good aeration at 30°C until the cuture is freshly saturated (ie: until  $\geq$ 2 hours after O.D.<sub>600</sub> = 2.0, or overnight).
- 2. Wash the cells once in water.
- 3. Dilute by a factor of 1:10 in 2% KOAc (pH7/acetic acid). The dilution factor is important to allow an approprite rise in pH.
- 4. Incubate at 30°C (in the shaker, with good aeration) for 4-6 days, after which asci should be visible on microscopic examination.

# Tetrad analysis

- 1. Mix a suspension of asci with 1 vol of 2% (w/v in water) Zymolyase and incubate for approx. 5' at room temperature; too short an incubation leaves the ascus wall too tough to rupture, too long an incubation reduces spore viability.
- 2. Streak out the suspension with a wire inoculating loop in a line along the side of a fresh, thin YPD plate, and mount the plate, inverted, on the stage of the micro-manipulator microscope stage.
- 3. Dissect the asci with a good needle and plant the spores on a sterile area of the plate. (This takes infinite patience and practice.)

The plates must be fresh because cells are much more easily moved around on moist, rather than dry, wrinkled agar, and they must be especially thin in order to accomodate the focal length of the microscope lens.



# Section 2

## Bacteriophage $\lambda$ methods

2.1 MEASUREMENT OF  $\lambda$  PHAGE TITRE

- 1. Prepare ten-fold dilutions of phage stock in Tris phage buffer (10mM Tris pH7.5, 10mM MgCl<sub>2</sub>).
- 2. In a test tube, mix 0.1 ml of freshly grown, saturated bacterial culture with 0.1 ml of phage dilution.
- 3. Add 3 ml of molten BBL top-layer agar, pre-incubated at 45°C, and pour immediately onto a pre-warmed L-agar plate.
- 4. Allow to set. Incubate at 37°C overnight. Count the plaques.

### 2.2 PREPARATION OF FILTERS FOR PLAQUE HYBRIDIZATION

- 1. Spin down a freshly saturated bacterial culture and resuspend the cell pellet in the same volume of 10mM MgSO<sub> $\Delta$ </sub>.
- 2. Mix 0.1 ml of cells with an appropriate dilution of phage stock (in Tris phage buffer) to allow separated plaques to form (up to 2000 plaques per 85 mm diameter plate.) Incubate for 15' at room temperature.
- 3. Add 3 ml of LB top-layer (0.7%) agarose, pre-incubated at 50°C, and pour immediately onto a warm BBL agar plate. Allow to set. Incubate overnight at 37°C.
- 4. Cool the plates at 4°C for 2 hrs.
- 5. With a pencil, mark a Schleicher and Schüll nitrocellulose filter with the plate number, and lay the filter onto the surface of the agar plate; avoid trapping air bubbles between the filter and the agar. With a needle, mark the exact position of the filter on the agar. Allow the filter to lie on the plate for at least 30<sup>°°</sup>. Several duplicate filters can be prepared successively from a single plate.
- 6. Carefully peel off the filter and lay it, DNA-side up, for 5' on a sheet of blotting paper soaked in denaturing solution (0.2M NaOH, 0.5M NaCl).
- 7. Transfer successively to two sheets of blotting paper soaked in neutralizing solution (IM NH<sub>3</sub>-acetate, 0.02M NaOH) for 5' each.
- 8. Blot on dry blotting paper and bake for 2 hrs. at 80°C under vacuum.

# 2.3 ISOLATION OF SINGLE CLONES

Having identified positive plaques by hybridization, a further round of purification ensures the isolation of single clones:

- 1. Pick a single positive plaque with a sterile Pasteur pipette, expel it into 1 ml of Tris phage buffer, then vortex.
- Dilute by 10<sup>-2</sup>-10<sup>-3</sup>, and mix 0.1 ml of dilution with 0.1 ml of freshly saturated bacterial culture in 10mM MgSO<sub>4</sub>. Incubate at room temperature for 15', and plate out in top-layer agarose as described above.

# 2.4 PHAGE PREPARATION - PLATE LYSATE

- Inoculate a single isolated plaque (~10<sup>5</sup> pfu) into 0.2 ml of freshly saturated bacterial culture in 10mM MgSO<sub>4</sub>. Incubate at room temperature for 15<sup>-</sup>.
- 2. Add 5 ml of top-layer agar and pour onto a fresh, wet L-agar plate. Incubate at 37°C until lysis is just confluent (about 8 hrs).
- 3. Overlay the plate with 5 ml of LB. Leave at 4°C overnight.
- 4. Decant the liquid into tubes, add 2 drops of chloroform, and store at 4°C. Phage titre:  $1-5 \times 10^{10}$  pfu/ml.

## 2.5 PHAGE DNA MINIPREP

- 1. Infect 0.1 ml of feshly saturated bacterial culture in 10mM MgSO<sub>4</sub> with 3  $\times$  10<sup>6</sup> pfu phage. Incubate for 15' at room temperature.
- 2. Add this to 10 ml of LB in a 100 ml flask and incubate at 37°C, shaking, until lysis occurs and the culture is seen to clear (about 9 hrs.)
- 3. Add 0.2 ml of chloroform, shake at 37°C for a further 10'.
- 4. Spin in a Sorvali centrifuge at 8000 x g (SS34 rotor at 7000 rpm) for 10<sup>°</sup> at 4<sup>°</sup>C. Discard the pellet of bacterial debris.
- 5. Transfer the supernatant material to a fresh tube, add DNase I (10  $\mu$ g/ml) and RNase A (50  $\mu$ g/ml). Incubate at 37°C for 30'.
- Add 2 ml of 0.5M Tris pH 9, 0.25M EDTA, 2.5% SDS, incubate at 70°C for 30'.
- 7 Add 2.5 ml of 8M KOAc, incubate on ice for 15' (in order to precipitate the SDS and proteins). Centrifuge at 12 000 x g (Sorvall SS34, 10 000 rpm) and discard the pellet.
- Transfer the supernatant to a fresh tube and add 2 vols. of ethanol. Spin at 10 000 rpm for 20', discard the supernatant and resuspend the pellet in 0.3 ml TE. Transfer to an Eppendorf tube.
- 9. Extract successively with TE-equilibrated phenol, phenol/chloroform 1:1, and chloroform.
- Add 30 μl of 3M NaOAc and 0.7 ml ethanol to precipitate the DNA, spin for 10' in an Eppendorf bench-top centrifuge, wash the pellet in 70% ethanol, dry, and resuspend in 100 μl TE. Analyse 5μl on an agarose gel.

### 2.6 PHAGE DNA MAXIPREP

- 1. Grow up a 5 ml culture of bacterial cells in LB containing 0.2% maltose, at 37°C overnight.
- Inoculate 4 ml of cells into 0.4 I LB containing 10mM MgSO<sub>4</sub> in a 21 conical flask. Incubate, shaking, at 37°C.
- After about 1hr., when O.D.<sub>600</sub> = 0.05-0.1, inoculate the culture with 3 x 10<sup>9</sup> pfu phage (m.o.i. = 0.1). Continue to incubate at 37°C until lysis occurs (5-10 hrs.)
- 4. Add 1 ml of chloroform, shake for 5'. Centrifuge at 7000 rpm for 15', transfer the supernate to a fresh bottle, discard the pellet.
- 5. Add DNase I (2 μg/ml) and RNase A (10 μg/ml), incubate at 37°C for 30'.
- 6. Add 12 g NaCl (to give a final concentration of 0.5M) and 40 g PEG (final concn. 40%), allow to dissolve, then incubate at 4°C for several hours or overnight.
- 7. Centrifuge at 7000 rpm for 20' and discard the supernate. Drain the pellets very well, and remove all residual liquid.
- 8. Resuspend the pellets in 2.5 ml (cold) Tris phage buffer by rolling the tube gently at 4°C. Rinse with a further 1 ml and transfer to a 10 ml tube.
- 9. Spin at 3000 rpm for 10', retain the supernate. Wash the pellet by resuspending it in 1 ml TPB, spin once more and combine the supernates.
- Prepare a CsCl step gradient in TPB in a Beckman SW40 ultracentrifuge tube as follows: 1ml 1.7 g/ml; 1.5 ml 1.5 g/ml; 1.5 ml 1.3 g/ml and layer the phage suspension (up to 7ml) onto the top. Spin at 35 000 rpm for 90' at 15°C in an SW40 rotor.
- 11. A broad band of debris forms in the 1.3 g/ml step; the sharp white phage band forms just below the 1.3-1.5 g/ml boundary.
- 12. Either proceed directly with step 13, or perform a further purification step of equilibrium gradient centrifugation by diluting the phage band with 1.5 g/ml CsCl, and centrifuging overnight at 15°C in a VTi65 rotor (45 000 rpm).
- 13. Dialyse the phage sample at  $4^{\circ}$ C against TPB ( $\geq$ 2hrs).
- 14. Extract succesively, gently, with phenol, phenol/chloroform and chloroform, and dialyse against TE.

Section 3

# pEMBL plasmid methods

3.1 BLUE/WHITE COLONY DISTINCTION

Materials

- Xgal: 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside [SIGMA, Cat. No. B 4252] 20 mg/ml in DMF (N,N-dimethyl formamide) Store dark in aliguots at -20°C.
- IPTG: Isopropyl β-D-thio-galactopyranoside [SIGMA, Cat. No. I5502] 20 mg/ml in water store in aliquots at -20°C. Filter-sterilize through 0.22 μm Millipore filters.

Amp plates: 100 or 200 µg/ml ampicillin in L-agar.

*E.coli*: Competent cells of a strain containing the M15 deletion of the  $\beta$ -galactosidase gene which can be complemented by the *lacZ* gene of pEMBL plasmids, eg: 71-18, JM101, etc..

Xgal is cleaved by active  $\beta$ -galacosidase to produce deep blue dibromodichloroindigo, causing a blue colony phenotype. IPTG induces  $\beta$ galactosidase expression.  $\beta$ -galactosidase is no longer sythesized when a DNA fragment is inserted into the polylinker of the pEMBL plasmid, disrupting the *lacZ* gene, which complements the host cell deletion: thus, Xgal is not cleaved, and a white colony phenotype is observed.

### Method

- 1. Transform competent cells: mix 200  $\mu$ l of competent cells with 20  $\mu$ l of plasmid DNA and incubate for 45'-60' on ice.
- Meanwhile prepare Xgal/IPTG amp plates by spreading 100 μl of the following mixture on each plate:
  - 2 vols. Xgal 1 vol. IPTG
  - 2 vols. water
- 3. Apply a heat shock to the cells by incubating at 42°C for 2'.
- 4. Spread appropriate volumes of transformed cells onto each plate, eg: 30, 60 and 100  $\mu$ l.
- 5. Incubate the plates at 37°C overnight. A mixture of blue and white colonies will grow. Blue colour development usually becomes more intense after the plates are left on the bench for a few days.

# 3.2 PREPARING FILTERS FOR COLONY HYBRIDZATION

- 1. Lay a filter onto an agar plate, avoiding air bubbles; transfer individual bacterial colonies to the filter and allow to grow overnight (but ideally for not more than 7 hrs).
- 2. Lift the filter from the plate; denature by laying (colony-side up) onto blotting paper soaked in 1.5M NaCl, 0.1M NaOH for 5'.
- 3. Neutralize by transferring to blotting paper soaked in 0.2M Tris pH 7.5, 2 x SSC for 15'.
- 4. Repeat step 3.
- 5. Lay onto blotting paper soaked in chloroform for 2'.
- 6. Squash-dry the filter by sandwiching it between sheets of blotting paper and rolling it flat with a glass bottle.
- 7. Bake for 2 hrs. at 85°C.

# 3.3 PLASMID DNA MINIPREP - BOILING METHOD

A fast method for screening large numbers of clones.

Solutions

STETTE8% Sucrose10mM Tris pH 8.05% Triton X-1001mM EDTA50mM Tris pH 8.050mM EDTA

### Method

- 1. Innoculate 5-10 ml LB (containing 100 μg/ml ampicillin) with a single, transformed colony. Grow at 37°C overnight or until saturated.
- 2. Centrifuge the cells at 3000 rpm for 5'. Discard supernatant material.
- 3. Resuspend the pellet in the residual liquid which accumulates at the bottom of the tube (usually about 0.2 ml) and transfer this to an Eppendorf tube.
- 4. Add 0.4 ml STET, vortex and place imediately in boiling water for 2'.
- 5. Spin in an Eppendorf centrifuge for 10'.
- 6. Remove the gelatinous pellet with a sterile toothpick and add 0.4 ml of isopropanol to the supernate. Vortex.
- 7. Centrifuge for 10'.
- Discard the supernatant material, wash the pellet with 70% (v/v) ethanol, vacuum dry, and resuspend in 100 μl TE. Analyse 5μl on an agarose gel.

### 3.4 PLASMID DNA MINIPREP - CLEARED LYSATE

This method produces a better and cleaner yield than above, but takes longer. The DNA is sufficiently clean for direct use in further cloning steps, oligo-labelling, primer elongation, etc..

### Solutions

<u>Sucrose/Tris</u>	Lysozyme	<u>Triton Lytic Mix</u>	<u>RNase A</u>
25% Sucrose	20 mg/ml in	0.1% Triton X-100	2mg/ml in
50mM Tris,	250mM Tris	50mM EDTA	in 0.25mM
pH 8.0	pH 8.0	50mM Tris pH 8.0	EDTA
•	•	(Freshly dissolved)	

### Method

- 1. Inoculate 5-10ml of LB/amp with a single colony, grow overnight at 37°C.
- 2. Centrifuge the cells at 3000 rpm for 5' in a bench-top centrifuge.
- 3. Resuspend pellet in 0.4 ml Sucrose/Tris and transfer to an Eppendorf tube.
- 4. Add 20 μl of lysozyme, mix, incubate at r.t. for 5'.
- 5. Add 10 µl of RNase/EDTA, incubate at r.t. for 5'.
- 6. Add 0.6 ml of triton lytic mix, incubate at r.t. for 1'.
- 7. Spin in an Eppendorf centrifuge for 10', then carefully remove and discard the gelatinous pellet with a wooden toothpick.
- 9. Extract the remaining supernate successively with 1 vol. phenol, 1 vol. phenol/chloroform 1:1, and 1 vol. chloroform.
- 10. Add 0.1 vol. of 3M NaOAc and 1 vol. of isopropanol, spin for 10'.
- 11. Wash the pellet with 70% ethanol, dry under vacuum and resuspend in 0.1 ml of TE. Analyse 2  $\mu$ l on an agarose gel.

### 3.5 PLASMID DNA MAXIPREP - CLEARED LYSATE

### Solutions

SUCROSE/TRIS	LYSOZYME	EDTA	TRITON LYTIC MIX
25% Sucrose	20mg/ml in	0.5M EDTA	0.1% Triton X-100
50mM Tris pH8	250mM Tris pH 8		50mM EDTA
-	(Freshly dissolved)		50mM Tris pH 8.0

This method is suitable for the preparation of both high copy-number plasmid DNA and M13 replicative form DNA. For the latter, set up 200 ml bacterial cultures, infect with phage at  $O.D_{600} = 0.05$ , and grow for  $\leq 5$  hrs, then proceed as described.

#### Method

- 1. From single colonies of each bacterial clone, grow 200 ml cultures overnight at 37°C.
- 2. Harvest the cells by centrifugation at 8000 rpm for 10'.
- 3. Resuspend the cell pellet in 2.5 ml sucrose/Tris and transfer to a 15 ml Corex tube.
- 4. Add 0.5 ml of lysozyme and mix by inverting the tube. Incubate for 5' at room temperature.
- 5. Add 0.5 ml of EDTA and mix. Incubate for 5' at room temperature.
- 6. Add 4 ml of Triton Lytic Mix and mix. Incubate for 20' at room temperature or until the suspension appears to clear.
- 7. Spin in a Sorvall centrifuge SS34 rotor for 1 hr at 12 000 rpm.
- 8. Carefully decant and retain the supernate.
- 9. Make up the volume of the supernate to 8.5 ml with TE and dissolve in it 8.1 g CsCl. When dissolved, add 0.5 ml 2 mg/ml ethidium bromide.
- Centrifuge in two small Quick-Seal tubes in a Beckman VTi65 rotor at 55000 rpm, overnight, or in one medium Quick-Seal tube in a Beckman Ti75 rotor, at 55000 rpm, for ≥20 hrs.
- 11. Visualize the DNA bands with U.V. light if necessary, and remove the large, plasmid band by drawing it out through the wall of the tube with a syringe.
- 12. Dilute the sample with two volumes of TE and extract 4 times with an equal volume of water-equilibrated isobutanol to remove the ethidium bromide.
- Add 2 volumes of ethanol, place at -20°C for 1 hr, spin down the precipitate (Sorvall SS34 rotor, 10 000 rpm, 20'), wash the pellet with 70% ethanol, dry under vacuum, and resuspend in 0.5 ml TE.
- 14. Measure the O.D.<sub>260</sub> and determine the DNA concentration, assuming O.D.<sub>260</sub> 1.0 to be equivalent to 50  $\mu$ g/ml DNA.

## Section 4

### **DNA** sequencing

4.1 PREPARATION OF ssDNA FROM pEMBL PLASMIDS

- 1. Grow overnight cultures of clones in 100 µg/ml ampicillin in LB.
- 2. Dilute by 1:50 in 10 ml LB, without ampicillin, in 100 ml conical flasks. Incubate at 37°C, shaking well to ensure good aeration.
- 3. At O.D.<sub>600</sub>  $\leq$ 0.20 (2 x 10<sup>8</sup> cells/ml for *E. coli* 71-18), or after one hour, super-infect the culture with interference resistant strain of f1 phage, IR1, at an a m.o.i. of  $\geq$ 20.
- 4. Incubate for a further 5 hrs (not longer).
- 5. Pellet the cells by centrifugation (10 000 rpm, Sorvall SS34 rotor, 20') and carefully remove 8 ml of <u>cell-free</u> supernatant material, which can be stored at 4°C until required.
- 6. Spin 2.5 ml supernate for 10' in a large (2.5ml) Eppendorf tube.
- 7. Avoiding any cell pellet, carefully transfer 2 ml spnt to a fresh tube containing 0.5 ml 2.5M NaCl, 10% PEG in a large Eppendorf tube. Mix and leave at room temp. for 15'.
- 8. Spin for 10<sup> $\prime$ </sup>. Discard the spnt with a flick of the wrist, and respin for 1<sup> $\prime$ </sup>. With a drawn out pasteur pipette remove the remaining  $\mu$ ls of spnt.
- 9. Resuspend the pellet in 0.2 ml of TE and add 0.2 ml of TE-equilibrated phenol. Vortex and place on ice for 5'.
- 10. Spin for 2'. Retain the aqueous phase and extract successively with equal volumes of phenol, phenol/chloroform 1:1 and chloroform.
- 11. Add 20  $\mu$ l of 5M LiCl and 0.5 ml of ethanol. Spin immediately for 10'.
- 12. Wash the pellet in 1 ml of 70% ethanol. Dry under vacuum for 5'.
- 13. Resuspend the dried pellet in 30  $\mu$ l water. Analyse 5  $\mu$ l on an agarose gel. Store at 4°C.

Preparation of bacteriophage F1 stock

- 1. Start a fresh culture of *E. coli* 71-18 by diluting an o.n. culture by 1/50 in 20 ml LB. Grow to O.D.<sub>600</sub> = 0.1, shaking with good aeration, at 37°C.
- 2. Inoculate with a single plaque of the phage IR1, picked with a pasteur pipette from a bacterial lawn on an agar plate. (Single plaques can be obtained initially by streaking out a phage suspension with a wire loop on a freshly made lawn of 71-18 and incubating overnight at 37°C.) Incubate the culture at 37°C, shaking, for 5-7 hours. (Overnight also works, but higher titres seem to obtain after a shorter time.)
- 3. Spin out cells at 8000 rpm for 10<sup>'</sup>. Store supernate at 4°C. Expected titre: ≥10<sup>11</sup> pfu/mi - can be upto 10<sup>13</sup> if everything is optimal. The supernate can be heated at 65°C for 30<sup>'</sup> to kill bacteria. Phage viability decreases with a half-life of approximately 6 weeks. Phage stocks should always be stored as cell-free supernatants to inhibit selection of faster growing phage deletion mutants.

Measuring F1 phage titre

- 1. Mix 0.1 ml of fresh overnight 71-18 culture with 0.1 ml of a phage dilution in LB in a sterile test tube. Incubate at 37°C for 15'.
- 2. Add 3 ml of molten BBL top-layer agar at 45°C, and immediately pour this over the surface of a warm L-agar plate; swirl it slightly to obtain a level top layer. (L-top layer agar may be used but gives smaller plaques.)
- 3. When the agar has set, incubate overnight at 37°C. Count the plaques.
- N.B: Infection of bacteria by F1 and related phage is dependent on the extrusion of sex pili. This does not occur below 34°C, therefore always grow cultures at 37°C.

#### 4.2 DNA SEQUENCING BY DIDEOXYNUCLEOTIDE CHAIN-TERMINATION

Stock solutions

10 x TBE:	100 g Tris/base,	55 g boric acid, 9	.3 g Na <sub>2</sub> -EDTA, d	dissolve, and
	make up the volu	ume to 1 I with wa	ater (check pH 8.3	3).
6% 30:1 acry	lamide, 8M urea	(500 ml): Dissolve	e 30 g acrylamid	e, 1 g
	methyl-bis-acryla	mide and 250 g	urea in 240 ml w	vater (do not
	heat to >40°C); a	dd about 20 g ion	exchange resin (E	3IO-RÁD
	AG501-X8(D) an	alytical grade mix	ked bed resin), s	tir until no
	further colour ch	ange occurs; filter	; add 50 ml 10 x	TBE; de-gas.
10 x Hin buff	er: 0.1M Tris pH	7.4, 0.1M MgCl <sub>2</sub> ,	0.5M NaCl.	-
Dideoxy-NTPs	s (ddNTPs): 10mM	1 stocks in 5mM T	ris pH 8.0, 0.1m	VI EDTA, and
•	working solns. (c	concn. determined	by trial) 0.69mM	ddGTP,
	0.09mM ddATP,	1.8mM ddTTP, 0.4	5mM ddCTP.	
dNTPs:	10mM stocks, as	above, and 0.5m	M working solutio	ns.
dNTP mixes:		·	-	
	<u>dGTPmix</u>	dATPmix	<u>dTTPmix</u>	<u>dCTPmix</u>
0.5mM dGTP	1 μl	20 µl	20 μl	20 μl
0.5mM dTTP	20 µl	20 µl	1 μİ	20 µl
0.5mM dCTP	20 µl	20 µl	20 µl	1 μİ
TE	5 µl	5 μİ	5 μΪ	5 μΙ
dNTP chase:	2.5mM each dN1	P. (Store all nucl	eotides at -20°C.'	)
Formamide d	ves: 100 ml forma	mide, deionized v	vith ion exchange	resin, filtered.
	add 30 ma xvlene	e cvanol. 30 mg bro	omophenol blue ar	nd EDTA
	(20mM). Store a	t room temperatu	ire.	
10-32PIDATP	Deoxyadenosine	5'-[a-32P]triphos	nhato triethvam	monium salt

- PJdATP: Deoxyadenosine 5'-[α-<sup>32</sup>P]triphosphate, triethyammonium salt, aqueous solution with 5mM β-mercaptoethanol, 3000 Ci/mmol, 10 μCi/μl. (Amersham.)
- Klenow DNA polymerase, 4.5 U/µl, Böhringer, Mannheim.

#### Gels

Glass gel plates (20 x 40 cm) scrubbed, rinsed, dried: one "sticky" plate treated with 5 ml ethanol containing 150  $\mu$ l 10% (v/v) acetic acid, 15  $\mu$ l "silane" ( $\gamma$ -methacryloxypropyltrimethylsilane), allowed to dry for 30'; one (non-sticky) plate treated with 5 ml 1% (v/v) dimethyldichlorosilane in carbon tetrachloride, allowed to dry for 5'. Gels are assembled with two lateral, 0.2 mm thick, Teflon spacers (the bottom of the gel is left open). 15  $\mu$ l TEMED and 200  $\mu$ l 10% ammonium persulphate is added to 20 ml acrylamide/urea solution at room temperature, which is then mixed and poured into the gel mould, held at a slight angle from horizontal, from the top. A 0.2 mm thick comb is inserted, and the gel allowed to polymerize in a horizontal position.

#### Annealing

For each clone to be sequenced, mix 5  $\mu$ l ssDNA template (see 4.1 above), 2  $\mu$ l 10 x *Hin* buffer, 1  $\mu$ l oligodeoxynucleotide primer (molar ratio of primer:template about 5:1), 12  $\mu$ l water; place at 70°C for 2', then at r.t. for 15'. Add 4  $\mu$ l annealled template/primer to each of four tubes labelled G, A, T, C.

### Chain extension

For *n* different clones to be sequenced, place  $n \ge 0.75 \ \mu l \ [\alpha^{-32}P] dATP$  into each of four Eppendorf tubes, and dry under vacuum. To one tube containing  $[\alpha^{-32}P] dATP$ , add  $n \ge (2 \ \mu l \ dGTP + 2 \ \mu l \ dGTPmix)$ , to the next tube add  $n \ge (2 \ \mu l \ dATP + dATPmix)$ , and so on for T and C. Mix  $n \ge 0.25 \ \mu l$  Klenow into the  $[\alpha^{-32}P] dATP/G$ -mix tube, and transfer 4  $\mu l$  into each primer/template "G" tube. Repeat for the A, T and C reactions. Incubate each for 15' at r.t. Add 2  $\mu l \ dNTP$  chase to each reaction, in the same order, and incubate for a further 15' at r.t. Place the reactions on ice, add 10  $\mu l$  of formamide dyes, boil for 2', quickly chill in ice water, then load 1-2  $\mu l$  onto the gel with a micro-capillary. Run the gel with 1 x TBE running buffer at up to 2000 V. Prise open the plates after the run (the gel sticks to the "sticky" plate), fix the gel by laying the sticky plate, gel side down, in 5% acetic acid for 10'; dry the gel on the plate (vacuum oven 80°C for 30', or microwave oven, 10') and expose by direct contact to X-ray film.

# 189

# Section 5

# Labelling and hybridization

## 5.1 HYBRID-ENRICHMENT OF RNA

Method of Ullu and Melli (1982) for the enrichment of RNA by hybridization to DNA which has been denatured and immobilized on Millipore filters. Fixing DNA to filters:

- 1. To 100  $\mu$ g plasmid DNA in 300  $\mu$ l TE, add 6  $\mu$ l of 5M NaOH. Incubate for 20' at r.t.
- 2. Pre-wash the filters, assembled in a suction apparatus attatched to a vacuum pump, with 2 x 5 ml 2M NaCl.
- 3. Add 5 ml of 2M NaCl to the alkaline-denatured DNA (thus raising the pH to 11.5), then apply the DNA to the filter, still assembled in the suction apparatus.
- 4. Wash filter with 3 x 5 ml of 2M NaCl
- 5. Wash filter with 3 x 10 ml of SSC
- 6. Dry the filter, then bake it at 80°C for 2 hrs.
- Annealing the RNA:
- 1. Pre-soak the filter in 5 ml of 2 x SSC, 100  $\mu$ g/ml proteinase K at 37°C for 15'.
- 2. Prehybridize the filter in 5 ml 2 x SSC, 10 x Denhardt's, 0.2% SDS, 20  $\mu$ g/ml denatured, sheared, non-homologous DNA, 100  $\mu$ g/ml proteinase K, for 2-3 hrs. at 65°C.
- 3. Hybridize by incubating the filter in the mixture of step 2., without nonhomologous DNA, but with 200 μg RNA (eg: total yeast RNA), o.n. at 65°C.
- 4. Wash filters: 30' at 65°C in 2 x SSC, 0.2% SDS; 30' at 65°C in 1 x SSC, 0.2% SDS; 30' at 65°C in 0.1 x SSC, 0.2% SDS; 10' at r.t. in water.
- 5. Elute the RNA from the filter with 2 ml of water at 95°C for 5'. Then precipitate with ethanol, using 20 μg/ml glycogen as carrier.

# 5.2 3' END-LABELLING OF RNA WITH <sup>32</sup>pCp

After England and Uhlenbeck (1978).

Materials:

<sup>32</sup> pCp	Cytidine 3,5'-[5'- <sup>32</sup> p]bisphosphate, trimethylammonium
	salt, aq. soln., 3000 Ci/mmol, 10 mCi/ml, Amersham.
RNA ligase	T4-infected E. coli RNA ligase, RNase and DNase-free,
-	Pharmacia.
10 x buffer	0.5M HEPES pH 7.5, 0.15M MgCl <sub>2</sub> , 33mM DTT.

Labelling mixture:	$\frac{Vol}{2} = \frac{10\mu l}{10}$
Incubate overnight	2.5 $\mu$ RNA (up to 0.5 $\mu$ g) 1 $\mu$ l 10 x buffer 1 $\mu$ l $32$ p Cp
al 4°C.	1 μl DMSO
	4.5 μl water
	0.2 U T4 RNA ligase

# 5.3 5' END-LABELLING OF OLIGONICLEOTIDES WITH [232P]ATP

Oligonucleotides were synthesized by phosphoramidite chemistry on an Applied Biosystems 380B DNA synthesizer and purified by HPLC. As standard procedure, 5' phosphorylation was performed as follows:

Standard mixture:	$Vol = 10 \mu l$
	2 µl (10 pmol) oligonucleotide
Incubate for 30' at 37°C.	1 µl 10 x PNK buffer
	$2 \mu [\gamma^{-32}P]ATP$ , ~5000 Ci/mmol,
	2pmol/µl, Amersham.
	5 µl water
	5 Ü polynucleotide kinase

Stop the reaction with 40  $\mu$ I TE (4°C) + 50  $\mu$ I TE-equilibrated phenol. After separating the phenolic phase, purify the labelled oligos for use as probes by loading directly onto a DEAE column (0.2 ml bed vol.) in a Pasteur pipette, pre-equilibrated with TE; wash out the [ $\gamma$ -<sup>32</sup>P]ATP with a few mls of 0.2M NaCI in TE, then elute the oligo with 1M NaCI in TE and use directly as a probe.

# 5.4 ISOLATION OF DNA FRAGMENTS

A clean and reliable method for isolating DNA fragments from polyacrylamide gels. 50% recovery is possible for fragments of 2 kb from 4% gels.

- 1. Prepare a 4% polyacrylamide (40:1) gel, 20 x 20 x 0.2 cm, in Trisacetate buffer, load several µg of restricted DNA, and fractionate by electrophoresis at 150 V. [For oligonucleotides, use a 20% gel.]
- 2. Stain the gel with ethidium bromide (0.05  $\mu$ g/ml, 10'), visualize the bands under UV, cut out the desired band.
- Crush the polyacrylamide band in a plastic tube, add ≥1 ml of Maxam and Gilbert (1980) elution buffer (0.5M NH<sub>4</sub>OAc, 10mM Mg<sub>2</sub>OAc, 1mM EDTA, 0.1% SDS), then incubate for several hrs. or o.n. at 37°C.
- 4. Pellet the polyacrylamide by centrifugation, remove and retain the spnt., add 1 volume of water to the pellet, and incubate at 37°C.
- 5. Repeat step 4.
- Combine all eluates (ie: 1 vol. buffer + 2 vols. water), and load onto ~0.2 ml bed vol. DEAE cellulose (Whatman DE52), equilibrated with water, in a siliconized, sterile Pasteur pipiette.
- 7. Elute the DNA from the column with 0.4 ml of 1.5M NH<sub>4</sub>OAc pH 10/NH<sub>3</sub>.
- 8. Add 0.8 ml of ethanol,  $\pm 2 \mu l$  10 mg/ml glycogen as carrier. Cool at -20°C, centrifuge, wash the pellet with 70% ethanol, dry under vacuum, and resuspend in 10  $\mu l$  TE.

N.B: This same method can be used to purify oligonucleotides, but to avoid losses on ethanol precipitation, it is desirable to use 1M triethylammonium bicarbonate (TEAB: 1M triethylamine,  $\sim$ pH 7.6/CO<sub>2</sub>) to elute the column, the sample can then be lyophilized.

### 5.5 LABELLING DNA BY RANDOM OLIGONUCLEOTIDE PRIMING

A method for preparing high specific activity DNA probes by random oligodeoxynucleotide priming, after Feinberg and Vogelstein (1983).

Solutions

- 250mM Tris pH 8, 25mM MgCl<sub>2</sub> TM:
- OL: 90 O.D.u/ml random deoxynucleotidehexamers, PL Biochemicals, No.27-2166, in 50mM ß-mercaptoethanol,1mM Tris pH 8, 1mM EDTA.
- LS: 1M HEPES pH 6.6/TM/OL mixed 25:25:7

### Method

- Take (max.) 50 ng DNA to be labelled, either super-coiled plasmid or 1. isolated fragment, in 5 µl water in an Eppendorf tube. Boil for 10'.
- 2. Spin for 2" to collect condensation, then boil again for 1', and quickly chill on ice.
- 3. Immediately add 9 µl of LS.
- Add 0.5 µl of 2 mg/ml BSA 4.
  - 1.0 µl of 0.5mM dGTP
  - 1.0 µl of 0.5mM dTTP

  - 2.0  $\mu$ l of [ $\alpha$ -<sup>32</sup>P]dATP 2.0  $\mu$ l of [ $\alpha$ -<sup>32</sup>P]dCTP
  - 1.0 μl of Klenow (4.5 U/μl)
  - Incubate at room temperature for several hrs. or overnight.
- 5. Add 30 µl of TE and extract with phenol. Separate the unincorporated nucleotides by passage of the material through a 1.5 ml Sepharose G-50 column in TE in a Pasteur pipette.

### 5.6 ASSESSING INCORPORATION OF LABEL INTO DNA PROBES BY TLC

A convenvenient method of assessing incorporation of <sup>32</sup>P into nick-translated, oligo-labelled, or kinased DNA probes; the DNA remains at the origin, while unincorporated nucleosides move with the buffer front.

### **Materials**

0.1 mm polyethylene-impregnated cellulose cards [POLYGRAM pre-coated cards for thin-layer chromatography, Macherey-Nagel & Co., 516 Düren, Werkstr. 6-8, Postfach 307, BRD].

Solvent: 0.75M KH<sub>2</sub>PO<sub>4</sub>, pH3.4/orthophosphoric acid Method

Apply ~0.1 µl of labelling reaction as a spot to the card; develop in buffer for ~20', then expose the card to X-ray film.

### 5.7 DNA AND RNA BLOT HYBRIDIZATION

# After Church and Gilbert (1984).

DNA and RNA was routinely transferred to nylon membranes for hybridization: either Hybond-N [Amersham], GeneScreen [NEN, Dupont], or Zeta-Probe [Bio-Rad], used according to manufacturers' instructons. RNA was electrotransferred from 1 mm thick 6% polyacrylamide/urea gels [for this, Hybond-N was most effective in retaining the smallest RNA species] as follows:

- 1. After electrophoresis, the gel was stained with ethidium bromide (0.1  $\mu$ g/ml ethBr in electro-transfer buffer, 10') and photographed.
- 2. The membrane was pre-soaked in electro-transfer buffer (25mM sodium phosphate, pH 6.5) for several minutes, then assembled into a sandwich with the gel, between pre-soaked blotting paper, and inserted into the electro-transfer tank, containing buffer.
- 3. Transfer proceeded at 4°C at 0.25A for 1 hr, then at 1.0A for 1 hr.
- 4. After transfer, the membrane was briefly rinsed in buffer, exposed to a pre-callibrated UV cross-linking lamp, baked at 80°C for 2 hrs, then stored until required for hybridization.

Hybridization of RNA and DNA blots was conveniently carried out in a 100 ml polythene cylinder, plugged with a silcon-rubber bung, rotated about its axis in a 65°C oven:

- 1. Pre-wet the membrane in water, then roll it around a 10 ml glass pipette (in Swiss roll fashion). Insert the pipette with the membrane into a 100 ml polythene measuring cylinder filled with 0.5M sodium phosphate, pH 7.2; the membrane uncoils from the pipette and adjusts its shape to the inner surface of the cylinder. Remove the pipette.
- 2. Decant the sodium phosphate buffer and pour in 20 ml of hybridization solution (0.5M sodium phosphate, pH 7.2, 1mM EDTA, 7% SDS), prewarmed to 65°C.
- 3. Seal the cylinder with a well-fitting bung, through the axis of which a syringe needle has been inserted in order to allow for adiabatic presure change. Place the cylinder in a horizontal position on a device that will rotate it about its axis in a 65°C oven. Incubate for 2-4 hrs.
- 4. Prepare a high specific activity probe (not more than 50 ng of DNA, purified by gel filtration through a Sephadex G-50 column in a Pasteur pipette, or by ion-exchange chromatography on DEAE cellolose), denature it by boiling, then immediately add it to 10 ml of hybridization solution, pre-warmed at 65°C. Substitute the pre-hybridization fluid in the cylinder for the hybridization mixture. Allow to hybridize, rotating in the 65°C oven, overnight.
- Remove the hybridization mixture, rinse the filter in the cylinder several times with wash-solution (40mM sodium phosphate, pH 7.2, 1mM EDTA, 1% SDS) at r.t., then remove the filter from the cylinder and wash it at 65°C for 4 x 20' with ~250 ml wash-solution. Expose the filter to X-ray film.

### 5.8 T7 RNA POLYMERASE IN VITRO TRANSCRIPTION

Method successfully used to generate full length transcripts of 250 nt.

### Reagents

T7 RNA polymerase: Riboprobe<sup>TM</sup>, Promega Biotec, Wisconsin, USA, 15 U/μl. RNasin: Promega Biotec, 40 U/μl.

Cap dinucleotide:  $m^7G(5')ppp(5')G$ , M<sub>r</sub> 815.3, Böhringer Mannheim GmbH, Nr. 904988, 5 A<sub>250</sub> units resuspended in 18 µl water.

 $[\alpha^{-32}P]UTP$ : Amersham, PB.10203, aqueous solution, 3000 Ci/mmol, 10  $\mu$ Ci/ $\mu$ I.

5 x Buffer: 200mM Tris pH 8, 40mM MgCl<sub>2</sub>, 250mM NaCl, 10mM spermidine. DTT: 300mM dithiothreitol.

ACG mix: Three ribonucleotide triphosphates mixed each at 4mM.

UTP (cold) 4mM.

DNA linearized plasmid template, phenol/chloroform treated, ethanol precipitated and resuspended in an appropriate volume of TE.

Reaction mixture	$Vol = 10  \mu l$
DNA 1 μg	2µ1
5 x buffer	2µ1
ACG mix	1µ1
Cap dinucleotide	1µ1
0.3M DTT	1µ1
RNasin	1µ1
[α- <sup>32</sup> Ρ]UTΡ	2µ1
T7 RNA pol. 15 U/μl	0.5µl
- Incubate for 5' at 37°C	•
Add 4mM UTP	1μl
- Incubate at 37°C for 10'-2	20

Add 40  $\mu$ l of (cold) TE and 50  $\mu$ l of phenol/chloroform, vortex, spin, then load the aqueous phase onto a Sephadex G-50/TE column in a Pasteur pipette; collect the first 20 drops of labelled peak (0.4-0.45 ml), pool the labelled fractions and then preciptate the RNA by adding 50  $\mu$ l of 5M LiCl, 1  $\mu$ l of 10mg/ml glycogen and 1 ml of ethanol.

### Section 6

Immunoprecipitation and materials and methods used in Chapter 5

### 1. Antibodies

Scleroderma patient anti-(U3)RNP serum and patient anti-(U1)RNP serum used in the experiments shown in Figures 5.1, 5.2, 5.3 and 5.6, were generously provided by Kathy Parker (see Parker and Steitz 1987); patient anti-(U3)RNP sera "knoblauch" and "S1" (from patient Dielz), mouse induced autoimmune anti-(U3)RNP serum "712", used in the experiments shown in Figures 5.7 and 5.8, and rabbit anti-2,2,7-trimethylguanosine cap serum (Lührmann et al. 1982) were provided by Reinhardt Lührmann; human anti-(U3)RNP monoclonal antibody 72B9 (Reimer et al. 1987) was provided by Eng Tan. All the anti-(U3)RNP antibodies were either reported (Lischwe et al. 1985, Reimer et al. 1987) or reputed to react with the 36 kD U3 snRNP peptide. Lupus erythematosus Sm antiserum from patient Doris Küng, used throughout, was provided by lain Mattaj (Mattaj and De Robertis 1985). The anti-54k monoclonal antibody with specificity for canine SRP 54k polypeptide, used in Figure 6, was given by Bernard Dopperstein. Human autoimmune serum H125, used in Figure 5.8, containing antibodies against the 68/72 kD protein components of SRP, was provided by W. van Venrooij.

### 2. Immunoprecipitation

The precipitation method was based on that of Matter *et al.* (1982). Binding of beads to antisera:

Protein A-sepharose CL-4B beads (Pharmacia) were pre-swollen in 10mM Tris pH 8.0, 150mM NaCl, 0.1% (v/v) Nonidet NP40 (or Triton X-100), 0.1% (w/v) sodium azide, in which state they could be stored for several months at 4°C. [I ml swollen bead volume contained 2 mg ligand which could bind 20 mg human lgG.] 2 mg (dry wt. equivalent) of beads were used for each precipitate sample; this corresponded to 30  $\mu$ l suspension, which was mixed with antiserum (in amounts described in the figure legends), and made up to 200  $\mu$ l in an Eppendorf tube with IP buffer: 50mM Tris pH 7.4, 150mM NaCl, 0.05% (v/v) Triton X-100. All steps were carried out at 4°C. Antisera were incubated with beads, gently rocking, for 2-4 hrs. [In the case of monoclonal antibodies provided as dilute cell supernatants, the supernatant was brought to 150mM NaCl and pH 7.4 with Tris, and incubated as a large volume (up to 10 ml) with beads over-night.] The beads were centrifuged for 2<sup>---</sup> and, after discarding the supernatant, were washed 4 x with 1 ml IP buffer.

Preparation of yeast nuclear extracts:

Nuclear extracts were derived from crude nuclear pellets prepared by hypotonic lysis of spheroplasts in 18% Ficoll solution (Schulz 1978) as described in Chapter 1 and this chapter, Section 1.3. Washed nuclear pellets were resuspended in 5 ml IP solution containing 10 mg/ml heparin, sonicated on ice for 4 x 10 seconds (with pauses) using a microprobe attached to a Branson Cell Disruptor B15, output setting 4, aliquotted into Eppendorf tubes and clarified by 5' centrifugation. As standard, 0.5 ml nuclear extract, corresponding to the yield from approximately  $10^9$  yeast cells, was added to each antibody/bead sample; the mixtures were incubated for 4 hrs, with rocking, then the beads were centrifuged and washed 4 x with IP buffer. [Precipitation of naked RNA by anticap antiserum was performed in the same manner, but 100 µg total yeast RNA was added instead of extract, and the volume made up to 0.5 ml with IP/heparin.] Removal of RNA from antibody/beads:

After washing, the beads were suspended in 50mM Tris/HCI pH 7.4, 5mM EDTA, 0.3M NaCl, 1.5% SDS, 2 mg/ml proteinase K and incubated at 37°C for 30'. The suspension was placed on ice for 5', the beads (and precipitated SDS) were removed by centrifugation for 2' at 4°C, the supernatant was retained and extracted successively with water-saturated phenol, phenol/chloroform and chloroform, then ethanol-precipitated after adding 2  $\mu$ l of 10 mg/ml (DEPC-treated and autoclaved) glycogen as carrier. Washed and dried RNA pellets were resuspended in 20  $\mu$ l water and analysed by blot hybridization or <sup>32</sup>pCp 3' end-labelling (England and Uhlenbeck 1978); 2.5  $\mu$ l RNA (one eighth of each sample) was labelled in a 10  $\mu$ l reaction volume (this chapter, Section 5.2), or blotted after electrophoresis.

#### 3. Construction of pSRBC plasmids

Oligonucleotides 180, 181, 182 and 183 (see page 192) were synthesised by phosphoramidite chemistry on an Applied Biosystems 380B DNA synthesiser and purified by HPLC by Philippe Neuner. The size of the oligos and their ability to anneal to each other was confirmed by 5' end-labelling with  $[\gamma^{-32}P]$ ATP and electrophoresis in 20% polyacrylamide gels either with or without urea; annealed oligos showed an appropriate band retardation on non-denaturing gels. Equimolar quantities of kinased, complementary oligos were heated at 70'C for 5', incubated at room temperature for 15' in water, then mixed in two different concentrations with *Eco*RI-cut Bluescript (Strategene, San Diego, California, USA) plasmid DNA for ligation: 2 pmol or 0.2 pmol annealed oligos, 0.05 pmol cut plasmid DNA, 1 x ligase buffer and 5U ligase, in 10 µl, and incubated overnight at 4°C. Any of *E. coli* strains 71/18, JM109 or XL1-Blue (Strategene) could be transformed with the ligation mixtures; white colonies were picked after growth on X-gal/IPTG and ampicillin, and these were further screened by hybridization to the labelled oligonucleotide probes. The orientation and number of oligonucleotide insertions in the positive clones were determined by a series of primer-elongation studies on linearized plasmid miniprep DNA. The constructions shown in Figure 5.5 were confirmed by DNA sequencing (using Sequenase [United States Biochemical Corporation]).

# 4. Injection of T7 transcripts into Xenopus oocytes<sup>1</sup>

T7 transcripts were generated as described in this chapter, Section 5.8 (see Contreras *et al.* 1982, Konarska *et al.* 1984, Chen *et al.* 1987, Krieg and Melton 1987), purified by gel filtration over Sephadex G-50, ethanolprecipitated, resuspended in 5 µl water, and mixed as described in the figures. 20-30 nl RNA was injected into the cytoplasm of each *Xenopus laevis* oocyte. After injection, the oocytes were incubated at 19°C over-night in Modified Barth's Solution (see Marcus-Sekura *et al.* 1987). Seven oocytes per immunoprecipitation sample were harvested in 1 ml 10mM Tris/HCl pH 8.0, 150mM NaCl at 4°C, homogenized, and the cell extract, having been clarified by centrifugation and brought to 0.1% Nonidet NP40, was added to antibodies bound to protein A-sepharose beads, and immunoprecipitated as described above.

<sup>&</sup>lt;sup>1</sup>The preparation, handling and injection into *Xenopus* oocytes described in Chapter 5 was performed by lain Mattaj. Descriptions of the techniques involved appear in Hitchcock *et al.* (1987), Marcus-Sekura *et al.* (1987) and Melton (1987).

# Section 7

### **BUFFERS, MEDIA AND REAGENTS**

(In alphabetic order)

BBL-agar. pH 7.2 10 g Trypticase 5 g NaCl 10 g Agar Add 1 litre bidist. water

Baltimore Biological Laboratories MERCK 6404 DIFCO 0140-01

BBL-top laver agar. pH 7.210 g Trypticase5 g NaCl6.5 g AgarAdd 1 litre bidist water

Baltimore Biological Laboratories MERCK 6404 DIFCO 0140-01

BSA - Bovine Serum Albumin - SERVA  $M_r \approx 67000$ 

Coomassie blue stain Coomassie blue stock soln.	
(1% in methanol, filtered)	50 ml
methanol	400 ml
acetic acid	100 ml
water	450 ml

Denhardt's solution. 50 x

BSA 10 g PVP 10 g Ficoll 10 g Water make up to 1 litre Filter-sterilize.

DEPC - Diethyl pyrocarbonate SIGMA use at 0.1% to inhibit RNase [not for Tris buffers!]

Destain (for Coomassie blue)

methanol	2 vol
acetic acid	1 vol
water	7 vol

EDTA: Titriplex III, MERCK

Ficoll (Type 400) - SIGMA

Formamide dyes		
deionized formamide	80%	v/v
TBE	1 x	
xylene cyanol	0.1%	w/v
bromophenol blue	0,1%	w/v

#### <u>Glycoaen</u>

as carrier for RNA precipitation: 10 mg/ml in water, add 0.1% DEPC, autoclave and store as aliquots at -20°C. For precipitation, use at 20  $\mu$ g/ml.

Hammersmith stabs. pH 7.2 Difco Nutrient Broth 9 g Difco Agar 7.5 g NaCl 5 g Thymine 0.1 g Water make up to1 litre Dispense 1-2 ml into stab bottles while hot. Hybridization buffer (DNA and RNA blots). 0.5M sodium phosphate pH 7.2, 1mM EDTA. 7% SDS Wash solution: 40mM sodium phosphate pH 7.2, 1mM EDTA. 1% SDS. Ligation buffer, 10 x 20mM Tris-HCI, pH 7.6 10mM MgCl<sub>2</sub>, 10mM DTT 0.5mM ATP Luria agar. pH 7.2 10 g Bactotryptone DIFCO 0123-01 5 g Yeast extract DIFCO 0127-01 10 g NaCl **MERCK 6404** 15 g Agar DIFCO 0140-01 Add 1 litre bidist. water Luria broth. pH 7.2 10 g Bactotryptone DIFCO 0123-01 5 g Yeast Extract DIFCO 0127-01 5 g NaCl **MERCK 6404** Add 1 litre bidist. water Maxam and Gilbert gel elution buffer 0.5M NH<sub>4</sub>OAc, 10mM Mg(OAc)2, 1mM EDTA, 0.1% SDS M9 salts solution. 10 x 70 g Na<sub>2</sub>HPO<sub>4</sub> **MERCK 6586** 30 g KH<sub>2</sub>PO<sub>4</sub> anhydrous **MERCK 4873** 5 g NaCĪ **MERCK 6404** 10 g NH<sub>4</sub>CI **MERCK 1145** make up to1 litre with bidist. water Phage Buffer 3 g KH<sub>2</sub>PO<sub>4</sub> anhydrous **MERCK 4873** 7 g Na<sub>2</sub>HPO<sub>4</sub> anhydrous **MERCK 6586** 5 g NaČl **MERCK 6404** 1 ml 1 molar MgSO<sub>4</sub> 0.1 ml 1 molar CaCl<sub>2</sub> 1 ml 1% Gelatine solution MERCK 4070 make up to1 litre with bidist water

198

Phenol. phenol/chloroform and chloroform Phenol, containing 0.1% 8-hydroxyquinoline as antioxidant, equilibrated with TE (pH 8.0) for DNA extaction, with water for RNA extraction. Chloroform: 24 vol chloroform + 1 vol isoamyl alcohol. Phenol/chloroform: 1:1 mixture of above, equilibrated with TE (for DNA) or water (for RNA). PNK buffer, 10 x 500mM Tris-HCL pH 7.6, 100mM MgCl<sub>2</sub>, 50mM DTT, 1mM spermidine, 1mM EDTA PVP - Polyvinyl Pirrolidone - SIGMA  $M_r \approx 40000$ Restriction enzyme, 1 x. low-, medium-, high-salt buffers 10mM Tris, pH 7.6 10mM (LS), 50mM (MS), or 100mM (HS) NaCl, 6mM MgCl<sub>2</sub>, 2mM β-mercaptoethanol Silane solution  $\gamma$ -methacryloxypropyltrimethylsilane 15µl [Wacker-Chemie, GmbH, München] acetic acid 150 µl ethanol 5 ml Siliconizing solution dimethyldichlorosilane, 2% in 1,1,1-trichloroethane [BDH Ltd., Poole, England] SSC. 20 x Sodium citrate dihydrate 441 g Sodium chloride 877 g make up to5 litres water adjust to pH 7.0/NaOH TBE Buffer, 10 x, pH 8.3 Tris-base 500 g Boric acid 275 g Na<sub>2</sub>EDTA 46.5 g Water make up to 5 litres Tris-acetate buffer. 20 x. pH 7.8 Tris-base 485 g 113 ml Acetic acid (glacial) Na<sub>2</sub>EDTA 140 g Water make up to 5 litres

Tris-base: Trizma, SIGMA

# Conclusions

This work describes a project which set out to identify new small RNA species in *Saccharomyces cerevisiae* and define their functions, with the eventual aim of better understanding the various mechanisms by which RNA performs possibly essential regulatory roles in living cells.

At the outset, the three most abundant yeast RNA species within the size range of 5.8S to 18S rRNA were identified and their genes were cloned. Two of these, snR17 and snR30, were identified as trimethyl guanosine-capped small nuclear RNAs. Many such small nuclear RNAs have been identified in *S. cerevisiae:* including those in this study, the total number described in the literature is at present 27 (Riedel *et al.* 1986, Brow and Guthrie 1988, Zagorski *et al.* 1988). This complexity of snRNAs is almost certainly common to all eukaryotes (Tollervey 1987b). In *S. cerevisiae*, at least two distinct families of snRNAs can be identified conveniently: those of the nucleoplasm, found associated with Sm antigens, including the snRNAs required for pre-mRNA splicing, and those found to be hydrogen-bonded to pre-rRNA in nucleolar fractions, of which at least some appear to be associated with an anti-(U3)RNP antigen.

SnR17, the main subject of this work, appears to be the most abundant snRNA in *S. cerevisiae*. SnR17 shows a strong resemblance to U3 snRNA of other eukaryotes in that it is found in nucleolar fractions hydrogen-bonded to prerRNA, it associates with anti-(U3)RNP antigen, and its primary and secondary structure is similar to U3. By these criteria, it seems reasonable to suppose that U3 and snR17 are homologous, and perform similar functions.

SnR17 is the only *S. cerevisiae* snRNA shown to be encoded by two genes; all other yeast snRNAs are encoded by unique genes. Yeast cannot survive without snR17, simultaneous null-mutations in both genes being lethal, but the two genes can complement each other, indicating that their transcripts perform the same essential function. The genes are not linked, the nucleotide sequences of the two transcripts are 96% matched, and they are probably transcribed at different rates, since the steady-state levels of the transcripts differ. Interestingly, it was discovered recently that U3 snRNA in *S. pombe* is also encoded by two genes (D. Tollervey, personal communication), whereas other *S. pombe* snRNAs are encoded by unique genes, suggesting either that there is some evolutionary

advantage to be gained by having two U3 snRNA genes instead of one, or that a duplication took place early in the course of evolution.

SnR17 and snR30 both perform essential functions. The yeast snRNAs required for pre-mRNA splicing are also essential (Guthrie and Patterson 1988), snR128 is essential (Zagorski *et al.* 1988), but several other snRNAs associated with pre-rRNA appear to be completely dispensable (Parker *et al.* 1988). This mixed population of essentially required and completely dispensable snRNAs associated with pre-rRNA in the nucleolus presents a complex picture. The potential significance of these snRNAs in ribosomal biogenesis remains to be determined.

# Appendix 1

# The structure of snR30 and its gene.

The cloning of the *S. cerevisiae* gene for the small nuclear RNA snR30 was decribed in Chapter 2. This RNA was mapped and its gene was sequenced by Marc Bally<sup>1</sup> (Bally *et al.* 1988). SnR30 was determined to be 605 nt in length, and haploid yeast strains inheriting a deletion of the unique *SNR30* gene were not viable. The sequence of the *SNR30* gene and flanking regions is presented in Figure A1.1. No significant similarities were found between *SNR30* and any other sequences in the nucleotide sequence data bases. However, a sequence very closely matched to the box C region of snR17 (see Chapter 3, Figure 3.12): CGCAUGAUCUUGA (nucleotides 403 to 415 in Figure A1.1), and a sequence resembling an Sm antigen binding site: AGUUUUGA (nucleotides 245 to 252) were observed. A "TATA" sequence is marked at nucleotide -99.

No attempt was made to define the secondary structure of snR30 because, with no clues from similarity with other molecules to suggest what conformation it might have, there were no criteria with which to evaluate the various structures predicted for such a long molecule. However, the structure was investigated to see whether the putative Sm binding site and the region similar to snR17 box C could occur in stable conformations as single-stranded loops or as base-paired stems. In all the structures analysed, using various sets of free-energy values (those of either Zucker and Sankoff 1984, Cech *et al.* 1983, or Freier *et al.* 1986), the putative Sm binding site consistently formed part of a stem, as it does in snR17, albeit with a bulged loop, as shown in Figure A1.2. The box C-like sequence in snR30 also consistently formed part of a stem (Figure A1.2) whereas the box C in the U3 snRNA models occurs in a loop (Figure 3.13). Chapter 5 describes an investigation that set out to determine whether the box C sequence functions as a specific protein binding site.

<sup>&</sup>lt;sup>1</sup>Present address: C.N.R.S., Laboratoire de Chimie Bactérienne, BP 71-31, Chemin J.-Aiguier, 13277 Marseilles, Cedex 9, France.

-420 -400 - 390 -360 -340 CTGTAAATTGCTCTTACTACAAAACCTGCAATGCCCATATCATCATATTCAAAGTAAAAATTTCATTTTCCGTAATTATTCTTATCTTTGGGTGCTTTGTG -320 -300 -280 -260 -240 GCAGCTACTGGAGTAATTGTGTATAATAGATGCCACTAGTGCAGTTAAATGAAATATCCTACATTGAGAGTTTAAAAAATTGCCTCCTTGTATGTTTCTT -220 -200 -180 -160 -140 GTAAATAATGATTTTATGTCCTTCTAGCGTAGTAACCCTAAACCATCAGCGGCTATTATTAACATTGTTACCCTGACAGTGCTTTTTTATTTCAAGAAAA -120 -100 -80 -60 -40 ATTTTCCTTTAGATTCTCTCGCCTATATAAACACTTTGAAACTCTCATCGTCGTCGTGCATTTAACTTCTATGGTAAACTTCTTTTCCCATTCAAGTTTCTCT -20 20 60 40 80 100 120 140 160 AUUGCUGCUUAGCUUCUCUAAAACACUGGGCUACGUUUUUCAACGCUCGAGAGGCGCGAGAGUCUCAAGGAGCCUCCAAUGGGCCUCACGUAUUCAUCUAGAU ATTGCTGCTTAGCTTCTCTAAAACACTGGGCTACGTTTTTCAACGCTCGAGAGGCAGAGTCTCAAGGAGCCTCCAATGGGCCTCACGTATTCATCTAGAT 220 -200 240 180 260 GGCGCUUCGGACAACGGCAUCACAUAAGAGAUGCAGCUCCUGACUUCUCCUCUGAUCUUCGUGAUCAG<u>AGUUUUGA</u>GUCGUCAGACUACGAGCA<u>GUUUCU</u> GGCGCTTCGGACAACGGCATCACATAAGAGATGCAGCTCCTGGACTTCTCCTCGTGATCAGAGTTTTGAGTCGTCAGACTACGAGCAGTTTCT 280 300 320 340 380 400 420 440 ACGAGUUUAACUUAGAUUAAGUAGGACGC<u>AUGAUCUUGA</u>GCUCUUUUCCUAUACUUUGUCCUAUGGCCAGCUUUCUCCUUAUUACGAAGAGAUUGCGGGA ACGAGTTTAACTTAGATTAAGTAGGACGCATGATCTTGAGCTCTTTTCCTATACTTAGTCCTATGGCCAGCTTTCTCCCTTATTACGAAGAGATTGCGGGGA 500 520 480 540 UGUGGGUGCAGAGUGGGAAAAUCUGAGUUCGGUCAUCUUUGUUGUUGUUCGUCCUACCGCAGUAUAUUCCUAAACACUAUGAAAUGACCCUAGUUGGUCCAUG TGTGGGTGCAGAGTGGGAAAATCTGAGTTCGGTCATCTTTGTTGTTCGTCCTACCGCAGTATATTCCTAAACACTATGAAATGACCCTAGTTGGTCCATG 580 600 620 640 660 AUCAUUUGGGUAAAACCAUACUGCAGACAU ATCATTTGGGTAAAACCATACTGCAGACATCTCTATAATTCACATATTCACATCTTGTCCTCTCGCAGAAGAGAAGGCCTTTGACGTAGAAAATTAAGT 680 700 720 740 760 AAATCGTATACGTTTACGTATATACACCTGTATTCAGCTATATATGTTGTTGTTCTCGTTCTCATAAAATGTACCCTACTTAGTGGTGGAATAAACGAAT 780 800 820 840 860 TGGGTTCTCCACGTAGACGATATCAAATGAATCCGCCGCCTTGGAAGACAACTTTGTCCATGACAGACCCCCGAAGGAGCCGTGATGAACTTTACACCAGAC 880 900 920 940 960 1000 1020 1040 1060 980  ${\tt tgaaaggttcagggacttcttcaaatttgctttcgtaatccaccttttccgggccttctattctttgccctgtaacaaaagatcaatcgacttcgacttt}$ 1080 1100 GTCCGCTCTCTTAATTCCCTTGGCCCACTTCCATG

# Figure A1.1 Nucleotide sequence of the SNR30 gene.

The RNA sequence has been inferred from the DNA sequence and the presence of the 5' trimethylguanosine cap structure has been shown by immunoprecipitation (see Figure 2.9). Underlined regions: putative "TATA" box at -99; Sm binding site consensus at 245; region similar to U3 snRNA conserved box C at 406.

# NUCLEOTIDE SEQUENCE OF SNR30 GENE



Figure A1.2

Secondary structure conformation around sequences in snR30 matching (A) the Sm-binding site consensus and (B) the conserved box C of U3 snRNAs.

# Appendix 2

# Structure and possible functions of scR1

The identification of a cytoplasmic RNA, scR1, in *Saccharomyces cerevisiae* and the cloning of its unique gene was described in Chapters 1 and 2. The detailed characterization of this gene was subsequently performed by Franco Felici<sup>1</sup> (F. Felici, G. Cesareni and J. M. X. Hughes, in preparation). Felici determined the structure of the gene and found scR1 to be 519 nt in length. He then went on to construct a diploid yeast strain heterozygous for a deletion of the *SCR1* gene. On sporulation, haploid strains inheriting the deletion were viable, but grew very slowly, frequently segregated cells that were unable to use glycerol as a sole carbon source, and showed abnormalities at cell division. This Appendix presents some further analysis of the structure of scR1 and its gene and of the phenotype of haploid yeast strains carrying the *scr1-A* mutaion, and discusses the possible functions of scR1.

The structure of scR1 and its gene.

The nucleotide sequence of the *SCR1* gene and its flanking regions, determined largely by Franco Felici, is illustrated in Figure A2.1; the scR1 RNA sequence is aligned to the coding sequence. No sequences with significant similarity to scR1 were found by screening the EMBL nucleotide sequence data base, but downstream from *SCR1*, on the opposite DNA strand, the coding sequence for a tRNA<sup>IIe</sup> (nucleotides 1531 to 1458) was discovered by its identity with other yeast tRNA genes (Felici and Cesareni 1987).<sup>2</sup>

The scR1 coding sequence is 519 nt long. The 3' end coincides with a run of (dT)s which is typical of an RNA polymerase III termination signal. RNA polymerase III genes possess a dual-element internal promoter, each element consisting of about 10 nt, separated from each other by 30 to 45 nt (reviewed by Ciliberto 1983). A search for matches to the internal promoter block A element with the consensus derived for eukaryotic tRNAs and mammalian *Alu*-family repetitive DNA, RGYNNRR-GG (where "R"= purine, "Y"= pyrimidine, "N"= any nucleotide, and "-"= variable presence of 0 to 2 nucleotides; Ciliberto 1983),

<sup>&</sup>lt;sup>1</sup>Present address: Dipartimento di Biologia, II Universitá di Roma, Torvergata, Via O. Raimondo, 00173 Roma, Italy.

<sup>&</sup>lt;sup>2</sup>The tRNA<sup>lle</sup> gene, like other yeast tRNA genes (Nelböck *et al.* 1985), is flanked by a complex pattern of Ty transposable element fragments; see Felici and Cesareni (1987) and Figure 4.11.

revealed two regions close to the 5' end: nucleotides +3 to +12 (numbering on the RNA sequence in Figure A2.1), GCTGTAATGG (with a mismatched C at the second position), and nucleotides +19 to +28, GGTGGGATGG. Using the block B element consensus, GWTCRANNC (where "W"= A or T; Ciliberto 1983), a region from nucleotides +52 to +60, GAACAAATC (with a mismatched A in the third position) was found. These elements have been marked on the sequence in Figure A2.1 and on the secondary structure model in Figure A2.3. Figure A2.2 compares these putative promoter elements with those of 7SL RNAs (see below).

Upstream of the transcription initiation site, from nucleotides -31 to -25, the sequence TATAAAA was found. "TATA" boxes have been considered to be typical of RNA polymerase II promoters, but recently, it was shown that human 7SK RNA and vertebrate U6 snRNA, both of which appear to be polymerase III transcripts, can be transcribed irrespectively of the presence of internal block A and block B elements, purely from their upstream promoters, which also contain TATA boxes at about -30 (Murphy *et al.* 1987, Das *et al.* 1988). Yeast U6 snRNA also has a TATA sequence at -30 (Brow and Guthrie 1988), whereas other yeast snRNAs consistently have TATA sequences between -80 and -100 nucleotides upstream of the transcription initiation site (see Figure 3.19).

In the light of the observations that scR1 is the most abundant small cytoplasmic RNA (other than rRNA and tRNA) of *S. cerevisiae* just as 7SL RNA of SRP is the most abundant small cytoplasmic RNA in mammalian cells (Ullu *et al.* 1982b, and see Chapter 1, Figure 1.1C), that no equivalent of 7SL RNA has yet been found in *S. cerevisiae* and yet 7SL-like molecules have been characterized from mammals (Ullu *et al.* 1982b, Li *et al.* 1982, Balmain *et al.* 1982) amphibia (Ullu and Tschudi 1984), insects (Gundelfinger *et al.* 1984), plants (H. L. Sänger, in preparation), other fungi (Ribes *et al.* 1988, Brennwald *et al.* 1988, Poritz *et al.* 1988) and even bacteria (Moritz and Goebel 1985), and that 7SL RNAs are polymerase III transcripts just as scR1 appears to be a polymerase III transcript (Zieve et al. 1977), it seemed worthwhile to look in detail at the possible structural similarities between 7SL RNAs and scR1. [The term "7SL" is used to denote putative RNA components of signal recognition particles (SRP) and not the sedimentation coefficient.]

Significant nucleotide sequence similarity can be seen between vertebrate 7SL RNAs (Ullu and Tschudi 1984) and to a lesser extent between vertebrate and insect (Gundelfinger *et al.* 1984), but primary structure similarity over a wider evolutionary spectrum is virtually non-existent. Secondary structure, however, is very well conserved among 7SL RNAs (Gundelfinger *et al.* 1984, Zwieb 1985, Poritz *et al.* 1988).

Attempts to align the nucleotide sequences of scR1 and 7SL RNAs revealed no overall sequence similarities. A remarkably conserved block of up to 13 nt, however, was found at the 5' ends of the 7SL RNAs of the yeasts *S. pombe* and *Y. lipolitica*, and scR1. In scR1, this block corresponds to nucleotides 3 to 12 that correspond to the consensus pol III block A sequence. The relative positions and sequences of regions similar to the pol III blocks A and B consensi in various 7SL RNAs and scR1 are shown in Figure A2.2. All the 7SL RNAs have block A sequences close to their 5' ends and those of the two yeast species match perfectly the first block A sequence of *S. cerevisiae* scR1.

Models for the secondary structure of 7SL RNAs have been determined by studying RNase digestion patterns, and patterns of compensatory base changes preserving stem structures between different species (Gundelfinger *et al.* 1984, Zwieb 1985). The secondary structure of 7SL RNAs is remarkably well conserved between species, despite lack of extensive nucleotide sequence similarity. It appears in the form of a "T" with a long stem extending from the ends of the molecule to a junction at which two shorter stems branch off. The 5' end usually forms two small loops at the beginning of the main stem (see Figure A2.3, human 7SL).

Comparison of the secondary structure models for human, *Xenopus*, *Drosophila*, *S. pombe* and *Y. lipolitica* 7SL-like RNAs has revealed a small number of conserved nucleotides in apparently single-stranded loops; the potential significance of these is highlighted by the fact that they occur in regions that broadly correspond to parts of the RNA protected from digestion by the binding of the 68/72 kd and 19 kd proteins (Poritz *et al.* 1988). Canine SRP proteins can associate with 7SL RNAs of other species to yield active SRP *in vitro* (Walter and Blobel 1983, Siegel and Walter 1985), and Poritz *et al.* (1988) have shown that canine proteins even bind to yeast (*S. pombe* and *Y. lipolitica*) 7SL RNAs.

Figure A2.3 shows a secondary structure model for scR1 derived by studying a selection of minimal free energy structures (see Figure A2.3 legend). ScR1 is 519 nt long, much longer than 7SL RNAs, which are approximately 300 nt or less (*S. pombe* RNA is 252 nt and *Y. lipolitica* RNA is 270 nt.) Other *S. cerevisiae* small nuclear RNAs, however, are larger compared to their vertebrate counterparts (Ares 1986, Kretzner *et al.* 1987, Siliciano *et al.* 1987b). The model in Figure A2.3 possesses some of the conserved features of 7SL RNAs. The human and *S. pombe* 7SL structures (after Poritz *et al.* 1988) are illustrated for comparison.

The model for scR1 shows a single 5' end loop, like that of *S. pombe*. There follows a long stem structure that would be equivalent to the stem of the "T" of 7SL RNAs, forming the backbone of the SRP (Andrews *et al.* 1987). Two "additional" hairpin structures branch from near the mid-point of the stem, which ends in a multi-stem loop structure, equivalent to the branch point of the "T". Three hairpin structures radiate from the branch point: an upper, further "additional" stem, and two others that exhibit individual, conserved nucleotides, indicated by circles in Figure A2.3. These two hairpins would correspond to the two branches of the "T", the ends of which have been shown by RNase footprinting to bind the 19 kd polypeptide component of SRP (Siegel and Walter 1988).

Figure A2.3 also shows that the sequences corresponding to the RNA polymerase block A and B elements occur in equivalent parts of the structures of both 7SL RNAs and scR1: block A within the small 5' end loop, and the block B on the main stem. Interestingly, just as the block A sequences of 7SL RNAs and scR1 form parts of loops, so the block A sequences of tRNAs form the first loop of the tRNA clover leaf (Ciliberto 1983). This suggests that the block A and B elements might have some structural importance for RNA polymerase III transcripts in addition to their importance as transcriptional signals; this hypothesis is supported by the observation that the block A and B elements are dispensable for the transcription of 7SK RNA and U6 snRNA, and yet these blocks are conserved in these molecules (Murphy *et al.* 1987, Das *et al.* 1988).

scr1 strains segregate  $\rho^{-}$  cells at high frequency.

F. Felici constructed yeast strains carrying a 235 nt deletion (between the *Clal* and *Aval* restriction sites, shown in Figure A2.1) within the 519 nt scR1 gene coding sequence. Haploid  $scr1-\Delta$  yeast colonies grew slowly, and microscopic observation revealed that approximately 5% to 10% of the cells had multiple buds, each of which could simultaneously contain nuclear material.

The growth of *SCR1* strains and *scr1-* $\Delta$  mutant strains was compared in different conditions and it was observed that cultures derived from *scr1-* $\Delta$  spores contained a high proportion of cells that were unable to grow on media containing glycerol as sole carbon source, indicating a lack of the ability of these cells to respire.

S. cerevisiae is a facultative anaerobe, that is, it can grow indefinitely on a fermentable carbon source without functional mitochondria. Owing to the repression of various respiratory enzymes during growth on certain sugars, notably on glucose (> 0.1%), S. cerevisiae preferentially gains most of its energy through alcoholic fermentation rather than respiration, even in the

presence of oxygen (Lagunas 1986). Respiratory competence can be assessed by the ability to grow on non-fermentable carbon sources such as glycerol, ethanol, lactate or acetate (Linnane and Lukins 1975). Respiratory deficiency can be caused by mutations either in the nuclear or mitochondrial genomes: gross lesions in the mitochondrial genome, designated  $\rho^{-,3}$  give rise to a *petite* phenotype; strains of this type ("cytoplasmic petites") are defined by their small colony size and irreversible loss of the ability to respire due to disruption of mitochondrial protein synthesis.

It was possible to isolate respiratory competent *scr1-* $\Delta$  clones if colonies were streaked onto glycerol plates shortly after germination. However, after growth in the absence of selection (approximately 10 generations on glucose) a high percentage of cells (10 to 95 % depending on the clone) lost the ability to respire. Prolonged culture in glycerol medium often stabilized the *scr1-* $\Delta$  clone with respect to its ability to respire, probably by selecting extragenic suppressors.

It was shown that respiratory deficient  $scr1-\Delta$  strains were  $\rho^-$  by performing a series of crosses with known  $SCR1 \rho^{\circ}$  strains: the respiratory deficiency of an  $scr1-\Delta$  strain could not be complemented by chromosomal loci, and respiratory deficient  $scr1-\Delta$  strains were unable to endow diploids with functional mitochondria. Thus, diploids selected on minimal medium lacking histidine after crossing respiratory deficient  $scr1-\Delta$  his3 HIS4 GY137, GY142 and GY143 with  $SCR1 \rho^{\circ}$  HIS3 his4 GY9 were unable to grow on glycerol as a sole carbon source (Table I). As a control it was shown that diploids from  $SCR1 \rho^{\circ}$ GY150 x  $SCR1 \rho^+$  GY139, and  $SCR1 \rho^+$  GY86 x respiratory deficient scr1GY137,142,143 crosses were able to respire (Table I; see also strains list, Chapter 4, Table I).

These results show that  $scr1-\Delta$  strains segregate  $\rho^-$  cells at high frequency. The small colony phenotype, however, was not simply a reflection of this, since both respiratory competent and  $\rho^- scr1-\Delta$  strains grew slowly.

The possibility that scR1 could be imported into mitochondria was excluded by probing mitochondrial RNA with a labelled *SCR1* probe: as shown in the Northern in Figure 2.11, scR1 was not enriched in RNA isolated from purified mitochondria.

 $<sup>{}^{3}\</sup>rho^{-}$  cells have large deletions and duplications of the mitochondrial genome and are relatively unstable, reverting to a  $\rho^{\circ}$  state, in which mitochondrial DNA has been completely lost (Wilkie 1975, Dujon 1981). During the characterization of *scr1* mutants, no attempt was made to distinguish between  $\rho^{-}$  and  $\rho^{\circ}$ .

#### DISCUSSION

ScR1 appears to be the most abundant single species of RNA in *Saccharomyces cerevisiae*, after rRNAs and tRNAs, and it is the first small cytoplasmic RNA to be described in this organism. Unlike the m<sub>3</sub>G capped snRNAs which are polymerase II transcripts, the *SCR1* gene contains sequence elements typical for RNA polymerase III transcription; this characteristic it shares with other cellular RNAs, namely 5S rRNA, tRNA, U6 snRNA (Das *et al.* 1988), 7SL RNA (Ullu *et al.* 1982b), 4.5S RNA (Leinwand *et al.* 1982), 7SK RNA (Murphy *et al.* 1987), "Y" RNAs (Keene *et al.* 1987) and viral RNAs VAI and VAII of adenovirus (Kitajewski *et al.* 1986) and EBER I and EBER II of Epstein-Barr virus (Lerner *et al.* 1981).

A secondary structure model for scR1 could be derived that showed some of the features conserved in 7SL RNAs of SRP. 7SL RNA can be divided into two functionally separable domains: a portion of the main stem with the 5' loops has been referred to as the "Alu domain", because this represents the part of the molecule homologous to Alu repetitive DNA (Ullu et al. 1982b). [Alu repeats are believed to be truncated pseudogenes of 7SL RNA (Ullu and Tschudi 1984).] The Alu domain binds 9 and 14 kd polypeptides in the SRP, and this portion of the RNP is responsible for the translation arrest phenomenon that characterizes SRP activity in vitro. SRP still mediates protein translocation even when the Alu domain is physically removed. Translocation is dependent on the binding of four other polypeptides: 68 and 72 kd polypeptides, which bind to the RNA close to the "T" junction and allow interaction of SRP with its receptor ("docking protein") in the endoplasmic reticulum; a 19 kd polypeptide, which binds to the stem loops at the ends of the two branches; and a 54 kd polypeptide, responsible for signal recognition, the binding of which is dependent on the 19 kd protein (Siegel and Walter 1985, Siegel and Walter 1986). This latter portion of the RNA has been called the "S domain".

The scR1 model (Figure A2.3) exhibits a similar pattern of stems and loops to the models determined for 7SL, with the addition of three extra hairpin structures. The model also exhibits some of the conserved, bulged nucleotides found in the two branches of the S domain of 7SL. ScR1 is significantly larger than 7SL RNA, but there is a precedent for unusually large "small" RNAs in *S. cerevisiae* (U1 and U2; Kretzner *et al.* 1987, Ares 1986), significant parts of which can be deleted without impediment to their function (Igel and Ares 1988) The tentative similarity between these molecules is interesting in the knowledge that attempts to identify a *S. cerevisiae* 7SL RNA homologue have so far failed (see Discussion in Chapter 4).

As a first step towards determining its function, a mutation in the unique *SCR1* gene was introduced by deleting a large part of the transcribed sequence. Haploid yeast strains carrying the *scr1-* $\Delta$  lesion were found to grow slowly, show signs of abnormal cell division and to give rise to cytoplasmic petites ( $\rho^{-}$ ) at high frequency.

The mechanisms causing  $\rho^{-}$  mutability, that is, the failure by yeast cells to maintain stably the mitochondrial genome, resulting in deletions of mitochondrial DNA (mtDNA) and consequent loss of mitochondrial function, are not understood. Ephrussi et al. (1949a & b) first showed that treatment of yeast cells with acroflavine dyes gave rise to mutants simultaneously defective in the function of several respiratory enzymes showing a non-Mendelian but cytoplasmic pattern of inheritance; these cytoplasmic petites were found to have extensive deletions and repetitions of the mitochondrial genome. Ethidium bromide-treatment was found to produce  $\rho^-$  cells very efficiently (100%) without apparently affecting the nuclear genome (Slonimski et al. 1968). Chemical mutagens, ultraviolet light, unsaturated fatty acid depletion, intramitochondrial ATP depletion, high or low temperatures, heavy metals and protein synthesis inhibitors, such as erythromycin, which specifically inhibits mitochondrial protein synthesis, have all been listed as causes of high frequency petite formation (Dujon 1981). Some nuclear pet mutations, temperaturesensitive for growth on glycerol, at non-permissive temperatures fail to synthesise certain mitochondrial encoded proteins and generate  $\rho^{-}$  cells on glucose (Marmiroli et al. 1980a). The temperature-sensitive cell division cycle mutations cdc8 and cdc21, whose wild type gene products are essential for both mitochondrial and nuclear DNA synthesis, are lethal at non-permissive temperatures, but both segregate  $\rho^{-}$  cells at high frequency at permissive temperatures; these effects were neither due to reduced rates of mtDNA replication nor to increased mtDNA point mutation rates (Newlon et al. 1979). Some rad and rec mutants, deficient in X-ray damage repair and mitotic intragenic recombination, have high  $\rho^-$  mutability (Dujon 1981). Certain nuclear mutations, of the type first described by Ephrussi and Hottinguer (1951), have no apparent phenotype other than high  $\rho^{-}$  mutability. Such genes may function directly in maintaining the stability of the mitochondrial genome during growth (Lusena and James 1976, Marmiroli et al. 1980b).

A direct involvement of scR1 in mitochondrial maintenance cannot be excluded, but in view of the many possible indirect causes of  $\rho^-$  mutability, and

since the scr1- $\Delta$  mutation appears to be pleiotropic, the segregation of  $\rho^-$  cells by scr1- $\Delta$  strains is probably due to an indirect mechanism.

Both  $scr1-\Delta$  strains and clathrin-deficient yeast strains grow much more slowly than wild type strains. Clathrin is a protein that forms a lattice coat on the surface of certain membrane vesicles ("coated vesicles") involved in intracellular membrane traffic, particularly in receptor-mediated endocytosis, where receptor-bound ligands accumulate in clathrin-coated membrane pits, which invaginate forming vesicles; the clathrin coat is subsequently stripped before fusion of the vesicles with their target membranes. The function of clathrin is not known, although roles have been suggested not only in endocytosis but in mediating membrane recycling and in vesicle transport from the Golgi to lysosomes and to secretory vesicles. Deletion of the clathrin heavy chain gene CHC1 in S. cerevisiae causes a delay in invertase transport from the Golgi to the cell surface and some defect in the processing of the mating pheromone precursor, prepro- $\alpha$ -factor, prior to secretion (Payne and Schekman 1985, Payne et al. 1987); the detrimental effect of the chc1- $\Delta$  lesion on secretion might account for the slow growth and aberrant cell morphology of these strains. In scr1- $\Delta$  cultures, a relatively high frequency of doubly budded cells ("Mickey Mouse phenotype"), in which both the buds appeared to contain nuclei, was observed. Aberrations in cell division have also been observed in  $chc1-\Delta$  strains (Lemmon and Jones 1987), although not of the multiple-bud type.

The possible similarity between  $scr1-\Delta$  and  $chc1-\Delta$  mutant phenotypes is interesting given the possibility that scR1 might be homologous to 7SL RNA (see secondary structure comparison in Figure 3.18). The signal-recognition particle (SRP), which contains 7SL RNA, has been shown *in vitro* to mediate the targeting of nascent secretory proteins to the endoplasmic reticulum (ER) membrane; any defect in SRP function therefore would be expected to have a significant effect on the secretory process.

Protein secretion is essential for yeast, as illustrated by the many conditionally lethal *sec* mutants causing defects at various stages in the secretory pathway (Schekman and Novick, 1982). The fact that the gene for 7SL RNA in *Schizosaccharomyces pombe* is required for the viability of this organism gives support to the hypothesis that SRP is absolutely required for secretion *in vivo* (Ribes *et al.* 1988, Brennwald *et al.* 1988, Poritz *et al.* 1988). If the latter hypothesis were correct, partially active SRP containing a truncated form of scR1 could not account for the survival of *scr1-* $\Delta$  *S. cerevisiae* strains because the *scr1-* $\Delta$  lesion consists of a 253 bp deletion in the *SCR1* gene removing all those nucleotides, according to the model presented in Figure A2.3, equivalent to

the conserved "S domain" of the SRP RNA, the portion of the molecule essential for signal-recognition (Siegel and Walter 1988).

The efficiency with which  $scr1-\Delta$  mutants secrete protein was not directly studied. It is clear, however, that these strains are not completely blocked in  $\alpha$ -factor secretion because  $\alpha scr1-\Delta$  strains could be mated with **a** *SCR1* strains, and  $\alpha$ -factor secretion is prerequisite for zygote-formation. Studies with *in vitro* translation/translocation systems have shown that translocation of prepro- $\alpha$ -factor across the ER can occur post-translationally, and in a way that is independent of canine SRP added back to the system (Rothblatt and Meyer 1986, Hansen *et al.* 1986). The possible role of endogenous SRP in these systems is not quite clear: a soluble factor in yeast cytoplasm that stimulates post-translational translocation of prepro- $\alpha$ -factor was found to be insensitive to RNase-digestion and therefore not likely to be RNP (Waters *et al.* 1986, Fecycz and Blobel 1987).

It has been shown that a human membrane-spanning protein can be translocated into the ER membrane post-translationally, although in a manner dependent on SRP (Muekler and Lodish 1986), and that the honeybee secretory polypeptide prepromelittin can be translocated post-translationally without SRP (Zimmermann and Mollay 1986). This suggests that the mechanisms of membrane translocation may be less rigidly defined than has been supposed, and that, *in vivo*, although most secretory proteins are presumably translocated co-translationally, with the aid of SRP as an adaptor between ribosome and ER membrane, this may not be an obligatory requirement. It is conceivable, therefore, that at least some organisms could survive without SRP, although in a rather sick state. The function of SRP could be conserved in *S. cerevisiae*, but this yeast may be more flexible in its ability to bypass SRP-mediated protein targeting, having evolved a more active alternative mechanism facilitating direct interaction of polypeptide signal sequences with signal sequence receptors in the ER membrane.

In conclusion, the abundant small cytoplasmic RNA scR1 plays some important role in *S. cerevisiae*, required to maintain normal growth, cell division and mitochondrial stability, and represents another addition to the increasing number of small RNAs with diverse functions. The hypothesis that the 7SL RNA component of SRP and scR1 share similar functions should be amenable to experimental testing.

Crossing respiratory deficient scr1 strains with known  $\rho$  or  $\rho^+$  strains.

-

Strain	Relevant type x	Strain	Relevant type	Selection	Growth on glycerol
GY9	a HIS3 his4 p°	GY137, 142, 143	α scr1 his3 HIS4 Respiration <sup>-</sup>	-His	I
GY150	α HIS3 his4 ade2 Red ρ <sup>o</sup>	GY139	a <i>SCR1 HIS3</i> <i>HIS4</i> ade2 White	Grande red on -His	+
GY86	<b>a</b> HIS3 his4 p <sup>+</sup>	GY137, 142, 143	α <i>scr1 his3</i> <i>HIS4</i> Respiration <sup>-</sup>	-His	, +

Table I Crosses performed to show that respiratory deficient scrt yeast strains are  $ho^-$ .
## SCR1 gene and flanking sequence

BclI

1	TGATCAACTTAGCCAGGACATCCATAGCTTGTGACAATCTGACTATATCATTTTCAGTCTTATTTGATGATCACACATTCCTTGTTATAACAACCCTCG	; 100
-	ACTAGTIGAATCOGTCCIGTAGGTATCGAACACTGTIAGACTGATATAGTAAAAGTCAGAATAAACTACTAAGIGTGTAAGGAACAATATIGTIGGGAGC	
101	GTATTCTTTTAACAAACTTACGATGACTATTATTGTATTAGTGTCAACCTCATCGCCAGGAACTTTTCATTGAAAATCTTCTGACAAATTTTCTGTGTAAA	200
	CATAAGAAAATTGITTGAATGCTACTGATATAACATAATCACAGTTGGAGTAGCGGTCCTTGAAAAGTAACTTTTAGAAGACTGITTAAAAGACACATTT	
	+1. SCR1 Eccri +50. AGGCUGUAAUGGCUUUCUGGUGGAUGGAUACGUUGAGAAUUCUGGCCAGGAA GGAAAGAAAATGGAG <u>TATAAAA</u> TGGAAAGTTTATTCCAATTGTGCTA <u>GGCTTGTAATGGCTTTCTGGGAAG</u> GATACGTTGAGAATTCTGGCCA <u>GAA</u>	L L
201		300
301	+100. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150. +150	G G
	GTTTAGGAAGGAGGCCGGATCTGTGCCTAACGTGCGGGAAACCGTTGCCTATCAAGAGATAAGGCGTGGCACGGGACAACACCGTTGGCAGAAAGGAGG	400
401	+200. Clai +250. UCGUAAAUUUGUCCUGGSCAGAGCUGUCUGCCCGGAGGCGGGAGAGUCCGUUCUGAAGUGUCCGGGCUAUAAUAAUCGUCUUGCGGGCAGCCCGUUG TCGTAAATTTGTCCTGGGCAGAGCTGTCTGCCCGGAGGCGGGAGGCCGGTAGAGTGGTCCGGCTATAATAAATCGATCTTTGCGGGCAGGCCGGTGG	500
	AGCATTTAAACAGGACCCGTCTCGACAGACGGGCCTCCGCCCTCTCAGGGCAAGACTTCACAGGGCCGATATTATTTAGCTAGAAACGCCCGTCGGGCAAC	
501	+300. +350. GCNGGAGGGUGAGGANUCCGUGUCUGUCUGUCGGGGGGGGGG	600
601	BstEI +400. +450. CGUCAGAAGUGACCCGUGAUGAGGCGGCGGGGAUAAGUGGAUAAUUGUGCAAGUUGAUCGCUUCGGCGGUUUAAUUUGGCGGUGCA CGTCAGAAGGTGACCCGTGATGGAGGCGGCCGGGATAGCACATATCGAGTCGGATAATTGTGCAAGTGGACGCCGTTTAATTTGCGGTGCCA GCAGTCTTCCACTGGGCACTACCTCCGCCGCCCTATCGTGTATAGCTCACCTATTAACACGTTCAACTAGCGAAGCCGCCCAAATTAAACCGCCACGGT	700
701	Avai +500, +519. UCAGGAUUJACUCGCACAUUGUGGCCCUUCGGGGAUGGAGUGUGUGCUGAACCAUAUUUUU TCAGGATTTACTCGCACATTGTGGCCGTTCCCTCGGGGATGGAGTGTGTGCTGAACCATATTTTTTTAACATTTTTTTGAACTAATTAACCGTCCCAATGA 	800
801	AsuII GGGATGITGGAAGAGTAAGTTGATTTAATGTOCTATTCTTGTTCTTTCGAATATAAATGACGATTGGAAGGTGCATGTCAATGTTTATCATCATGTTCT 	900
901	TGTTGTAGGTAGTCCTTACCTTGGTTGTAGAGTGAGTAGCGACAGCGGCAGGGAATATAAGAAGGATAAAATGAGAATACTTAGAACGTGTTGAAAGAA	1000
	ACAACATCCATCAGGAATGGAACCAACATCTCACTCATCGCTGTCCGCCGTCCCTTATATTCTTCCTATTTTACTCTTATGAATCTTGCACAACTTTCTT	1000
.001	GAATAAACGATCATCCATCGACTAGTACTGCTGCTACTAGTGTGTGT	; + 1100 ;
101	AAAATAGTCATCAAAATCTAATGGAAGCTGAAATTCAAGGGTTGATAATGTAATAAAATAAAGAATGACAAAGTATAAAAGGAAGAAGAAGATAGAGTAACATT	r + 1200
1101	TTTTATCAGIAGTTTAGATTACCTTCGACTTTAAGTTCCCAACTATTACATTATTTTATTTCTTACTGTTTCCTACTACTTCTTCCTACTCCATCTACTACTTCTACTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT	4

## Figure A2.1 Nucleotide sequence of the SCR1 gene

The transcribed sequences for scR1 and tRNA<sup>lle</sup> (second page) are marked. Elements resembling RNA polymerase III internal promoter sequences in both the genes have been underlined, aswell as a TATA sequence upstream of the SCR1 transcription initiation site. Restriction sites relevant to the work are indicated.

	TAATACATCTCGATAGTTAAGCGAAAAACATCTAAGGATAAAAAAGAGCTCCTCTTAACGTCGTATAAGCAAGGTGCATTATAAOGCOGGAAATTGTTC
1 201	ATGTAATTOCAACAATTATCTTAAAAATTTGTCTATATTTTTAAAAAAGTATGAACAATTACCAACATGTGGTACATATTGGAGGCCCGCAAAGCTGAG
1301	TACATTAAGGTYGTTAATAGAATTTTTAAACAGATATAAAAAATTTTTTCATACTTGTTAATGGTYGTACACCATGTAAACCTCCGGGGGTTTCGACTC
1401	
1501	ACGGTGCCTTAACCAACTGGGCCAAGAGACCAAATTTTAAATGGAAATACTTTTACTGGAAAGAGGTAACAAACCATGGAAGATGCTTTTCTACCAC
	TGCCACGGAATI <u>GGTTGACCCGG</u> TTCTCTGGTTTAAAATTTACCTTTATGAAAATGACCTTTCTTCCATTGTTTGGTACCTTCTACGAAAAGATGGTG UGCCACGGAAUUGGUUGACCGGUUCUCUGG Gene for tRNA <sup>le</sup>
1601	ATGAAAGTCTGAGATTTTAATGATGTTGGAATAAAAATCAACTATCATCTTGAGAAATGGGTGAATAATTAGATAAGTGTGTGGGATTCCATTG
	TACTITCAGACTCTAAAAATAATTACTACAACCTTATTTTAGTTGATAGTAGAACTCTTTACCCACTTATTAATCTATTCAACAACCCTAAGGTAAC
1 7 6 1	GATAAAGGCTATAATATTAGGTATACATGTTGGAATAGAAATCAACTATCATCTACTAGTATTTACATTACTAGTATATTATCATATACGGTGT
1701	CTATTTCCGATATTATAATCCATATGTACAACCTTATCTTTAGTTGATAGTAGATGATCATAAATGTAATGATCATATAATAGTATATGCCACA
1801	GAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATAAACATATAA 
	ХћоІ
1 <b>9</b> 01	CGGAATGAGGAATAATCGTAATATTAGTATGTAGGAATATAGGAGTCCATTTGAGGATCCTATATCCTGAGGAGGAGAACTTCTGGTATATCGTGTAT +
	TAATATTATAGCTTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCATTTCTCATGGTAGCGCCTGTGCTTCGGTTACTTCTAAGGAA
2001	ATTATAATATCGGAAATAGTTGTTACCTTAGGGTTGTTAATAGAGTTGTAAGTGGGTAAAGAGTACCATCGCGGACACGAAGCCAATGAAGATTCCTT
2101	CCACACAAAATCAAGATCCGTTAGACGTTTCCAGCTTCCAAAAATTCAAGAATATGATAAGGCTTCCACTAAGGCTAACTCTCCAACAGACAACAACACCCTG
	GGTGTGTTTAGTTCTAGGCAATCTGCAAAGTCGAAGGTTTTAAGTTCTTATACTATTCCGAAGGTGATTCCGATTGAGAGTTGTCTGTTGTTGTGTGGGC
2201	TCATCAGCTGTTCCAGAGAAACCCCCATCATGCCTCTCCTCAAACTGCTCAGTCACATTCACCACAGAATGGGCCGTACCCACAGGAGTGCATGATGAC
	AGTAGTCGACAAGGTCTCTTGGGGGTAGTACGGAGGAGGAGTTTGACGAGTCAGTGTGAGTGTCTTACCCGGCATGGGTGTCGTCACGTACTACTG
<b>230</b> 1	AAAACCAAGCCAATCCATCTGATTGGTCATTTTAGGGACGCCATCTATGATTCCGTATACACeTTATCAAATGTCGCTATGTACTTTCCACTGGgCcA
	TTTTGGTTCGGTTAGGTAGACTAACCAGTAAAATGCCTGCGGTAGATACTAAGGCATATGTGGAATAGTTTACAGCGATACATGAAAGGTGACCcGgT

-

RNA polymerase	III internal promoter elements in scR1 and 7SL RNA genes.
Human Frog Fly Tomato S. <i>pombe</i> <i>Y. lipolytica</i> S. <i>cerevisiae</i> scF	5' +1Pol IIIPol III5' +1Block ABlock B6CCG6GCGGGTGGCGCGTGTCTGTGTGTCCGGGAGGCTGGGG6CCG6GCGCGGTGGCGCGTGTTGTGTGTGTCCGGGAGGCTTGGGC6CCG6GCGCGTGGCGTGTGTATCCAGCTGTTGGAGGCTTGGGC6CCG6GCGCGTGGCGTGTGTATCCAGCTGTATCGGGGGGGGGGGGGGCTGGGC6GTGACTAGTAACGTGGCGTTGTAATCCAGCTTCGGGGGGGGGGGGGGCTGGGC6GTGAGCTTAGTAACGTGGCTTGTTAATCCAGGTGGAGGACATCAAAGTGG6GTGAGCTAGTAATCCAGGTGGAGGACATCAAAGTGGCGCGGGGGG6GACTAGCTGTAATCCCTTGTTAGTACCCCGAGAGCATCAAAGTGG6GTAATGGCTGTAATGGCTTTTGTCGCCCGCGAGGAGGACATCAAAGTGGCTTGGTGGGTG
tRNA <sup>Ile</sup>	GGTCTCTT <u>GGCCCAGTTGG</u> TTAAGGCACCGTGCTAATAACGCGGGGATCAGCG <u>GTTCGATCC</u> CGCTAGAGACCA
Figure A2.2 RNA polymerase II	ll internal promoter elements in scR1 and 7SL RNA genes.
Matches to consensi "R"= purine, "Y"= pyr A or T), allowing upto 1982b), <i>X. laevis</i> (UI (Moritz and Geebel 1 gene for tRNA <sup>IIe</sup> (fron	for eukaryotic tRNA genes and mammalian <i>Alu</i> -family repetitive DNA (Ciliberto 1983) block A, RGYNNRR-GG (where imidine, "N"= any nucleotide, and "-"= variable presence of 0 to 2 nucleotides), and block B, GWTCRANNC (where "W"= o one mismatch, have been underlined. Matches within the first 100 nt of the 7SL RNA genes from human (Ullu <i>et al.</i> lu and Tschudi 1984), <i>D. melanogaster</i> (Gundelfinger <i>et al.</i> 1984), tomato (H. L. Sänger, in preparation), <i>H. halobium</i> 985), <i>S. pombe</i> (Ribes <i>et al.</i> 1988), <i>Y. lipolytica</i> (Poritz <i>et al.</i> 1988), the <i>S. cerevisiae</i> scR1 gene and the <i>S. cerevisae</i> n Figure A2.1) are shown.

.

.

•

,

## Figure A2.3 Secondary structure model for scR1 compared to established 7SL RNA structures.

The scR1 structure is compared with *Schizosaccharomyces pombe* and human 7SL RNA structures (after Poritz *et al.* 1988). Conserved nucleotides occurring in loops or bulges of human, *Xenopus laevis*, *Drosophila melanogaster*, *S. pombe*, *Yarrowia lipolitica*, tomato and *Halobacterium halobium* 7SL-like RNA secondary structures are circled. Nucleotides that are conserved but are lacking, or appear to form parts of stems, in more than one of these species have not been marked. Of the 14 circled nucleotides that fulfil this criterion, *Saccharomyces cerevisiae* scR1 possesses 11 of them at comparable positions.

The method of Williams and Tinoco (1986) was used to generate a selection of lowest free energy secondary structures for scR1 that could be compared with the established structures of 7SL RNAs (Gundelfinger et al., 1984; Zwieb, 1985). None of the scR1 structures showed the simple "T" form of 7SL RNAs, they all showed a more complex pattern of stems and loops. One structure, however, (from among the most stable, with  $\Delta G = -161$  kcal), had a similar 5' end conformation to that represented for S. pombe 7SL by Poritz et al. (1988), that is, with a single small loop. [Ribes et al. 1988, and Brennwald et al. 1988 have drawn their S. pombe 5' structures without loops, other 7SL RNAs have a double loop at this position.] The scR1 structure shown differs from the most stable scR1 structure in the conformation of nucleotides 156 to 373. Optional folding patterns for this latter portion were investigated by folding nucleotides 149 to 380 independently: the conformation of this region shown was chosen because it best represented the conserved looped and bulged nucleotides indicated by Poritz et al. (1988). The free energy of this portion of the structure was  $\Delta G$ = -74.5, compared to  $\Delta G$  = -83.6 for the most stable conformation. The distal portions of the hairpin structures between nucleotides 156 and 373 of the scR1 model shown are the same as in the most stable structure; the nucleotide interactions linking the bases of these stems being somewhat different.

Danielle Konings performed the computation to derive the various scR1 structures.



## **Reference** list

.

Altherr, M.R., Quinn, L.A., Kado, C.I. and Rodriguez, R.L. (1983). Transformation and storage of competent yeast cells. *NATO Adv. Sci. Inst., Series A*, 61; 33-36.

Altman,S., Gold,H.A. and Bartkiewicz,M. (1988). **Ribonuclease P as a snRNP**. In Birnstiel,M.L. (ed.) *Structure and function of major and minor small nuclear ribonucleoprotein particles*. Springer-Verlag, Heidelberg.

Andersen, J., Delihas, N., Ikenaka, K., Green, P.J., Pines, O., Ilercil, O. and Inouye, M. (1987). The isolation and characterization of RNA coded by the micF gene in E. coli. Nucl. Acids Res., 15; 2089-2101.

Andrews, D.W., Walter, P. and Ottensmeyer, F.P. (1987). Evidence for an extended 7SL RNA structure in the signal recognition particle. *EMBO* J., 6; 3471-3477.

Arber, W., Enquist, L., Hohn, B., Murray, N.E. and Murray, K. (1983). **Experimental methods for use with lambda**. In Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (eds.) *Lambda II*, Cold Spring Harbor Laboratory, New York. pp433-466.

Ares Jr.,M. (1986). U2 RNA from yeast is unexpectedly large and contains homology to vertebrate U4, U5 and U6 small nuclear RNAs. *Cell*, 47; 49-59

Bachellerie, J.-P., Michot, B. and Raynal, F. (1983). Recognition signals for mouse pre-rRNA processing. *Mol. Biol. Rep.*, 9; 79-86.

Bachmann,M., Mayet,W.J., Schröder,H.C., Pfeifer,K., Meyer zum Buschenfelde,K.-H. and Müller,W.E.G. (1986). Association of La and Ro antigens with intracellular structures in HEp-2 carcinoma cells. *Proc. Natl. Acad. Sci. USA*, 83; 7770-7774.

Baldari,C. and Cesareni,G. (1985) Plasmids pEMBLY: new single stranded shuttle vectors for the recovery and analysis of yeast DNA sequences. *Gene*, 35; 27-32.

Bally,M., Hughes,J. and Cesareni,G. (1988). SnR30: a new, essential small nuclear RNA from Saccharomyces cerevisiae. Nucl. Acids Res., 16; 5291-5303.

Balmain, A., Krumlauf, R., Vass, J.K. and Birnie, J.D. (1982). Cloning and characterization of the abundant cytoplasmic 7S RNA from mouse cells. *Nucl. Acids Res.*, 10; 4259-4277.

Bayev,A.A., Georgiev,O.I., Hadjiolov,A.A., Kermekchiev,M.B., Nikolaev,N., Skryabin,K.G. and Zakharyev,V.M. (1980). The structure of the yeast ribosomal RNA genes. 2. The nucleotide sequence of the initiation site for ribosomal RNA transcription. *Nucl. Acids Res.*, 8; 4919-4926.

Bayev,A.A., Georgiev,O.I., Hadjiolov,A.A., Nikolaev,N., Skryabin,K.G. and Zakharyev,V.M. (1981). The structure of the yeast ribosomal RNA genes. 3. Precise mapping of the 18S and 25S rRNA genes and structure of the adjacent regions. *Nucl. Acids Res.*, 9; 789-799.

Beckwith, J.R. (1964). A deletion analysis of the *lac* operator region in *Eschericia coli. J. Mol. Biol.*, 8; 427-430.

Beggs, J.D. (1978). Transformation of yeast by a replicating hybrid plasmid. *Nature*, 275; 104-109.

Benton,W.D., and Davis,R.W. (1977). Screening of  $\lambda$ gt recombinant clones by hybridization to single plaques in situ. Science, 196; 180-182.

Berger, S.L. and Kimmel, A.R. (eds.) (1987). Guide to molecular cloning techniques. *Methods in Enzymology*, vol. 152. Academic Press.

Berk,A.J. and Sharp,P.A. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell*, 12; 721-732.

Bernstein,L.B., Mount,S.M. and Weiner,A.M. (1983). Pseudogenes for human small nuclear RNA U3 arise by integration of self-primed reverse transcripts of the RNA into new chromosomal sites. *Cell*, 32; 461-472.

Bhargava, M.M. and Halvorson, H.H. (1971). Isolation of nuclei from yeast. J. Cell Biol., 49; 423-429.

Bhargava, M.M., Cramer, J.H. and Halvorson, H.O. (1972). Isolation of high molecular weight DNA from yeast nuclei. Anal. Biochem., 49; 276-284.

Bindereif, A. and Green, M.R. (1987). An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. *EMBO J.*, 6; 2415-2424.

Birnstiel,M.L. and Schaufele,F.J. (1988). Structure and function of the minor snRNPs. In Birnstiel,M.L. (ed.) Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Heidelberg.

Black,D.L., Chabot,B. and Steitz,J.A. (1985). U2 aswell as U1 small nuclear ribonucleoproteins are involved in pre-mRNA splicing. *Cell*, 42; 737-750.

Blanchard, J.-M., Brunel, C. and Jeanteur, P. (1977). Characterisation of an endogenous protein kinase activity in ribonucleoprotein structures containing heterogeneous nuclear RNA in HeLa cell nuclei. *Eur. J. Biochem.*, 79; 117-131.

Botstein,D., Falco,S.C., Stewart,S.E., Brennan,M., Scherer,S., Stinchcomb,D.T., Struhl,K. and Davis,R.W. (1979). Sterile host yeasts (SHY): a eukaryotic system of biological containment for recombinant DNA experiments. *Gene*, 8; 17-24.

Boulnois, G.J. (1987) Gene cloning and analysis: a laboratory guide. Blackwell Scientific Publications, Oxford.

Bourgaize, D.B. and Fournier, M.J. (1987). Initiation of translation is impaired in *E. coli* cells deficient in 4.5S RNA. *Nature*, 325; 281-284.

Brammar, W.J. (1982). Vectors based on bacteriophage lambda. In Williamson, R. (ed.) *Genetic Engineering 3*, Academic Press, London.

Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Bernard, H. and Jacob, M. (1982). U2 RNA shares a structural domain with U1, U4 and U5 snRNAs. *EMBO* J., 1; 1259-1265.

Brennwald, P., Liao, X., Holm, K., Porter, G. and Wise, J.A. (1988). Identification of an essential *Schizosaccharomyces pombe* RNA homologous to the 7SL component of signal recognition particle. *Mol. Cell Biol.*, 8; 1580-1590.

Bringmann, P. and Lührmann, R. (1986). Purification of the Individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. *EMBO J.*, 5; 3509-3516.

Bringmann,P., Rinke,J., Appel,B., Reuter,R. and Lührmann,R. (1983). Purification of snRNPs U1, U2, U4, U5 and U6 with 2,2,7-trimethylguanosine-specific antibody and identification of their constituent proteins reacting with anti-Sm and anti-(U1)RNP antisera. *EMBO J.*, 2; 1129-1135.

Bringmann,P., Appel,B., Rinke,J., Reuter,R., Theissen,H. and Lührmann,R. (1984). Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base-pairing. *EMBO J.*, 3; 1357-1363.

Broach, J.R., Strathern, J.N. and Hicks, J.B. (1979). Transformation in yeast: development of a hybrid cloning vector and isolation of the *CAN1* gene. *Gene*, 8; 121-133.

Brody, E. and Abelson, J. (1985). The "spliceosome": yeast premessenger RNA associates with a 40S complex in a splicingdependent reaction. *Science*, 228; 963-967.

Brow, D.A. and Guthrie, C. (1988). Yeast U6 snRNA is remarkably conserved. *Nature*, 334; 213-218.

Brunel,C., Sri-Widada,J., Lelay,M.-N., Jeanteur,P. and Liautard,J.-P. (1981). Purification and characterization of a simple ribonucleoprotein particle containing small nucleoplasmic RNAs (snRNP) as a subset of RNP containing heterogeneous nuclear RNA (hnRNP) from HeLa celis. Nucleic Acids Research, 9; 815-830.

Busch,H., Reddy,R., Rothblum,L. and Choi,Y.C. (1982). SnRNAs, snRNPs and RNA processing. Ann. Rev. Biochem., 51; 617-654.

Calvet, J.P. and Pederson T. (1981). Base-pairing interactions between small nucear RNAs and nuclear RNA precursors as revealed by psoralen cross-linking *in vivo*. *Cell*, 26; 363-370.

Calzone,F.J, Britten,R.J. and Davidson,E.H. (1987). Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. *Methods in Enzymology*, 152; 538-556.

Casey, J. and Davidson, N. (1977). Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentration of formamide. *Nucl. Acids Res.*, 4; 1539-1552.

Castaño, J.G., Ornberg, R., Koster, J.G., Tobian, J.A. and Zasloff, M. (1986). Eukaryotic pre-tRNA processing nuclease: copurification with a complex cylindrical particle. *Cell*, 46; 377-387. Cavalier-Smith,T. (1987). The origin of cells: a symbiosis between genes, catalysts and membranes. Cold Spring Harbor Symp. Quant. Biol., LII; 805-824.

Cech,T.R. (1986). A model for the RNA-catalysed replication of RNA. *Proc. Natl. Acad. Sci. USA*, 83; 4360-4363.

Cech,T.R. and Bass,B.L. (1986). Biological catalysis by RNA. Ann. Rev. Biochem., 55; 599-629.

Cech,T.R., Zaug,A.J. and Grabowski,P.J. (1981). In vitro splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell*, 27; 487-496.

Cech,T.R., Tanner,N.K., Tinoco,I., Weir,B.R., Zucker,M. and Perlman,P.S. (1983). Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences. *Proc. Natl. Acad. Sci. USA*, 80; 3903-3907.

Cesareni,G. and Banner,D.W. (1985). Regulation of plasmid copy number by complementary RNAs. *TIBS*, 10; 303-306.

Cesareni,G. and Murray,J.A.H. (1987). Plasmid vectors carrying the replication origin of filamentous single-stranded phages. In Setlow,J.K. (ed.) *Genetic Engineering*, Vol. 9; pp135-153, Plenum Publishing Corporation.

Chang, D.D. and Clayton, D.A. (1987). A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. *Science*, 235; 1178-1184.

Chen,W. and Struhl,K. (1985). Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *EMBO J.*, 4; 3273-3280.

Chen,W., Tabor,S. and Struhl,K. (1987). Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell*, 50; 1047-1955

Cheng, S.-C. and Abelson, J. (1987). Yeast spliceosome assembly in yeast. *Genes and Development*, 1; 1014-1027.

Church,G.M. and Gilbert,W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA*, 81; 1991-1995.

Ciliberto, G. (1983). Transcription by RNA polymerase III. In Current Topics in Developmental Biology, Vol. 18; 59-88. Academic Press.

Ciliberto, G., Dathan, N., Frank, R., Philipson, L. and Mattaj, I.W. (1986). Formation of 3' end on U snRNAs requires at least three sequence elements. *EMBO J.*, 5; 2931-2937.

Contreras, R., Cheroutre, H., Degrave, W. and Fiers, W. (1982). Simple, efficient *in vitro* synthesis of capped RNA useful for direct expression of cloned eukaryotic genes. *Nucl. Acids Res.*, 10; 6353-6362.

Craig, N., Kass, S. and Sollner-Webb, B. (1987). Nucleotide sequence determining the first cleavage site in the processing of mouse precursor rRNA. *Proc. Natl. Acad. Sci. USA*, 84; 629-633.

Crouch,R.J., Kanaya,S. and Earl,P.L. (1983). A model for the involvement of the small nucleolar RNA (U3) in processing eukaryotic ribosomal RNA. *Molec. Biol. Rep.*, 9; 75-78.

Cryer, D.R., Eccleshall, R. and Marmur, J. (1975). Isolation of yeast DNA. *Methods in Cell Biology*, XII; 39-44.

Dahlberg, J.E. and Lund, E. (1988). SnRNA genes and their transcription. In Birnstiel, M.L. (ed.) *Structure and function of major and minor small nuclear ribonucleoprotein particles.* Springer-Verlag, Heidelberg.

Dahlberg, J.E. and Schenborn, E.T. (1988). The human U1 snRNA promoter and enhancer do not direct synthesis of messenger RNA. Nucl. Acids Res., 16; 5828-5940.

Darnell, J.E. and Doolittle, W.F. (1986). Speculations on the early course of evolution. *Proc. Natl. Acad. Sci. USA*, 83; 1271-1275.

Das,G., Henning,D., Wright,D. and Reddy,R. (1988). Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *EMBO J.*, 7; 503-512.

Daum,G., Böhni,P.C. and Schatz,G. (1982). Import of proteins into mitochondria. J. Biol. Chem., 257; 13028-13033.

Dente, L., Cesareni, G. and Cortese, R. (1983). pEMBL: a new family of single-stranded plasmids. Nucl. Acid Res., 11; 1645-1655.

Dente,L., Sollazzo,M., Baldari,C., Cesareni,G. and Cortese,R. (1985) The pEMBL family of single-stranded vectors. In Glover,D.M. (ed.) DNA Cloning: a practical approach, Vol.I, pp101-107, IRL Press, Oxford.

De Robertis, E.M., Lienhard, S. and Parisot, R.F. (1982). Intracellular transport of microinjected 5S and smail nuclear RNAs. *Nature*, 295; 572-577.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res., 12; 387-395.

Dotto,G.P., Enea,V. and Zinder,H.D. (1981). Functional analysis of bacteriophage f1 intragenic region. *Virology*, 114; 463-473.

Dreyfuss,G. (1986). Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Ann. Rev. Cell Biol., 2; 459-498.

Dreyfuss, G., Philipson, L. and Mattaj, I.W. (1988) a. Ribonucleoprotein particles in cellular processes. J. Cell Biol., 106; 1419-1425.

Dreyfuss, G., Swanson, M.S. and Piñol-Roma, S. (1988) b. Heterogeneous nuclear ribonucleoprotein particles and the pathway of messenger RNA formation. *TIBS*, 13; 86-91.

Duffus, J.H. (1975). The isolation of yeast nuclei and methods to study their properties. *Methods in Cell Biology* XII; 77-97.

Dujon, B. (1981). Mitochondrial genetics and functions. In Strathern, J.N., Jones, E.W. and Broach, J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, pp505-635.

England, T.E. and Uhlenbeck, O.C. (1978) 3'-terminal labelling of RNA with T4 RNA ligase. *Nature*, 275; 560-561.

Eperon,I. (1987). M13 cloning and dideoxysequencing. In Boulnois,G.J. *Gene cloning and analysis: a laboratory guide*, Blackwell Scientific Publications, Oxford. pp109-122.

Ephrussi, B. and Hottinguer, H. (1951). On an unstable cell state in yeast. Cold Spring Harbor Symp. Quant. Biol., 16; 75-85.

Ephrussi, B., Hottinguer, H. and Chimènes, Y. (1949)a. Action de l'acriflavine sur les levures. I. La mutation "petite colonie". Ann. Inst. Pasteur, 76; 351-367.

Ephrussi, B., Hottinguer, H. and Tavlitzki, J. (1949)b. Action de l'acriflavine sur les levures. II. Étude génétique du mutant "petite coionie". Ann. Inst. Pasteur, 76; 419-450.

Epstein, P., Reddy, R. and Busch, H. (1984). Multiple states of U3 RNA in Novikoff hepatoma nucleoli. *Biochemistry*, 23; 5421-5425.

Fecycz, I.T. and Blobel, G. (1987). Soluble factors stimulating secretory protein translocation in bacteria and yeast can substitute for each other. *Proc. Natl. Acad. Sci. USA*, 84; 3723-3727.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Ananlytical Biochem.*, 132; 6-13.

Felici, F. and Cesareni, G. (1987). Structure of the Saccharomyces cerevisiae gene encoding tRNA<sup>IIe</sup> (IAU). Nucl. Acids Res., 15; 364.

Ferretti,L. and Sgaramella,V. (1981) Specific and reversible inhibition of the blunt end joining activity of the T4 DNA ligase. *Nucl. Acids Res.*, 9; 3695-3705.

Forster,A.C., Jeffries,A.C., Sheldon,C.C. and Symons,R.H. (1987). Structural and ionic requirements for self-cleavage of virusoid RNAs and *trans* self-cleavage of viroid RNA. *Cold Spring Harbor Symp. Quant. Biol.*, LII; 249-259.

Frendewey, D. and Keller, W. (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell*, 42; 355-367.

Frischauf, A.-M., Garoff, H. and Lehrach, H. (1980). A subcloning strategy for DNA sequence analysis. *Nucl. Acid. Res.*, 8; 5541-5549.

Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol., 170; 827-842.

Frischauf, A.-M. (1987). Construction and characterization of a genomic library in  $\lambda$ . Methods in Enzymology, 152; 190-199.

Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Neilson,T and Turner,D.H. (1986). Improved free energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA*, 83; 9373-9377.

Galli,G., Hofstetter,H., Stunnenberg,H.G. and Birnstiel,M.L. (1983). Biochemical complementation with RNA in the *Xenopus* oocyte: a small RNA is required for the generation of 3' histone mRNA termini. *Cell*, 34; 823-828.

Garcia, G., Mar, P.K., Mullin, D.A., Walker, J.R. and Prather, N.E. (1986). The E. coli dnaY gene encodes an argenine transfer RNA. Cell, 45; 453-459.

Georgiev, O.I., Nikolaev, N. and Hadjiolov, A.A. (1981). The structure of the yeast ribosomal RNA genes. 4. Complete sequence of the 25S rRNA gene from Saccharomyces cerevisiae. Nucl. Acids Res., 9; 6953-6958.

Gerbi,S.A. (1985). **Evolution of ribosomal DNA**. Chapter 7 in MacIntyre,R.J. (ed.), *Molecular Evolutionary Genetics*. Plenum Publishing Corporation.

Gerke, V. and Steitz, J.A. (1986). A protein associated with small nuclear ribonucleoprotein particles recognizes the 3' splice site of premessnger RNA. *Cell*, 47; 973-984.

Ghosh, P.K., Reddy, V.B., Swinscoe, J., Lebowitz, P. and Weissman, S.M. (1978). Heterogeneity and 5' terminal structures of the late RNAs of simian virus 40. J. Mol. Biol., 126, 813-846.

Glover, D.M. (1985). DNA Cloning: a practical approach, Vols.I & II, IRL Press, Oxford.

Gottlieb, E. and Steitz, J. (1987). The mammalian La protein is an RNA polymerase III transcription termination factor. *Mol. Biol. Rep.*, 12; 243.

Grabowski, P.J., Seiler, S.R. and Sharp, P.A. (1985). A multi-component complex is involved in the splicing of messenger RNA precursors. *Cell*, 42; 345-353

Green, M.R., (1986). Pre-mRNA splicing. Ann. Rev. Genet., 20; 671-708.

Greider, C.W. and Blackburn, E.H. (1987). The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, 51; 887-898.

Grummt, I., Sorbas, H., Hofmann, A. and Roth, E. (1985). Spacer sequences downstream of the 28S rRNA coding region are part of the mouse rDNA transcription unit. *Nucl. Acids Res.*, 13; 2293-2304.

Guarente, L. and Mason, T. (1983). Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. *Cell*, 32; 1279-1286.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35; 849-857.

Gundelfinger, E.D., Di Carlo, M., Zopf, D. and Melli, M. (1984). Structure and evolution of the 7SL RNA component of the signal recognition particle. *EMBO J.*, 3; 2325-2332.

Guo, P., Erickson, S. and Anderson, D. (1987). A small viral RNA is required for *in vitro* packaging of bacteriophage  $\phi$ 29 DNA. Science, 236; 690-694.

Guthrie, C. and Patterson, B. (1988). Spliceosomal snRNAs. Ann. Rev. Genet., in press.

Haber, J.E. and Halvorson, H.O. (1975) Methods in sporulation and germination of yeasts. Methods in Cell Biology, XI; 45-69.

Habets, W., Hoet, M., Bringmann, P., Lührmann, R. and van Venrooij, W.J. (1985). Autoantibodies to ribonucleoprotein particles containing U2 small nuclear RNA. EMBO J., 4; 1545-1550.

Hadjiolov, A.A. (1985) The Nucleolus and Ribosome Biogenesis, Springer-Verlag, Wien.

Haenni,A.-L., Joshi,S. and Chapeville,F. (1982). tRNA-like structures in the genomes of RNA viruses. *Prog. in Nucl. Acid Res. and Mol. Biol.*, 27; 85-104.

Hamm, J., Kazmaier, M. and Mattaj, I.W. (1987). *In vitro* assembly of U1 snRNPs. *EMBO J.*, 6; 3479-3485.

Hansen,W., Garcia,P.D. and Walter,P. (1986). *In vitro* protein translocation across the yeast endoplasmic reticulum: ATP-dependent post-translational translocation of the prepro-a-factor. *Cell*, 45; 397-406.

Hashimoto, C. and Steitz, J.A. (1984). U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. Nucl. Acids Res., 12; 3283-3293.

Hashimoto, C. and Steitz, J.A. (1986). A small nuclear ribonucleoprotein associates with the AAUAAA polyadenylation signal *in vitro*. *Cell*, 45; 581-591.

Hellung-Larsen, P. (1977). Low molecular weight RNA components in eukaryotic cells. Properties of the LMW RNA components apart from tRNA, 5S rRNA and 5.5S rRNA. FADLs Forlag, København, Denmark.

Hendrick, J.P., Wolin, S.L., Rinke, J., Lerner, M.R. and Steitz, J.A. (1981). Ro cytoplasmic ribonucleoproteins are a sub-class of La ribonucleoproteins: further characterization of the Ro and La small ribonucleoproteins from uninfected mammalian cells. *Mol. Cell. Biol.*, 1; 1138-1149.

Hernandez, N. and Weiner, A.M. (1986). Formation of 3' end of U1 snRNA requires compatible snRNA promoter elements. *Cell*, 47; 249-258.

Hinnen, A., Hicks, J.B. and Finck, G.R. (1978). Transformation of yeast. *Proc. Natl. Acad. Sci. USA*, 75; 1929-1933.

Hitchcock, M.J.M., Ginns, E.I. and Markus-Sekura, C.J. (1987). Microinjection into Xenopus oocytes: equipment. Methods in Enzymology, 152; 276-284.

Hodnett, J.L. and Busch, H. (1968). Isolation and characterization of uridylic acid-rich 7S ribonucleic acid of rat liver nuclei. J. Biol. Chem., 243; 6334-6342.

Hogeweg, P. and Hesper, B. (1984). The alignment of sets of sequences and the construction of phyletic trees: an integrated method. J. Mol. Evol,. 20; 175-186.

Hollingsworth, M. and Martin, N. (1986). RNase P: activity in the mitochondria of Saccharomyces cerevisiae depends on both mitochondrial and nuclear-encoded components. *Mol. Cell. Biol.*, 6; 1058-1064.

Hughes, J.M.X., Konings, D.A.M. and Cesareni, G. (1987). The yeast homologue of U3 snRNA. EMBO J., 6; 2145-2155.

Hutchins, C.J., Rathjen, P.D., Forster, A.C. and Symons, R.H. (1986). Selfcleavage of plus and minus RNA transcripts of avocado sunblotch viroid. Nucl. Acids. Res., 14; 3627-3640.

Ide,G.J. and Saunders,C.A. (1981) Rapid isolation of yeast nuclei. Current Genetics, 4; 85-90.

Igel,A.H. and Ares, M. (1988). Internal sequences that distinguish yeast from metazoan U2 snRNA are unnecessary for pre-messenger RNA splicing. *Nature*, 334; 450-453.

Inouye, M. and Delihas, N. (1988). Small RNAs in the prokaryotes: A growing list of diverse roles. *Cell*, 53; 5-7.

Jacob, M., Lazar, E., Haendler, B., Gallinaro, H., Krol, A. and Branlant, C. (1984). A family of small nucleoplasmic RNAs with common structural features. *Biol. Cell*, 51; 1-10.

James,B.D., Olsen,G.J., Lui,J.S. and Pace,N.R. (1988). The secondary structure of ribonuclease P RNA: the catalytic element of a ribonucleoprotein enzyme. *Cell*, 52; 19-26.

Jeppesen, C., Stebbins-Boaz, B. and Gerbi, S. (1988). Nucleotide sequence determination and secondary structure of Xenopus U3 snRNA. Nucl. Acids Res., 16; 2127-2148.

Johnson, J.R. and Reader, H.P. (1983). Genetic control of flocculation. In Spencer, J.F.T., Spencer, D.M. and Smith, A.R.W. (eds), *Yeast genetics:* fundamental and applied aspects, Springer Verlag. pp205-224.

Kaiser,K. and Murray,N. (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In Glover,D.M. (ed.) DNA Cloning: a practical approach, Vol.I, pp1-47, IRL Press, Oxford.

Karn, J., Brenner, S., Barnett, L. and Cesareni, G. (1980) Novel bacteriophage  $\lambda$  cloning vector. *Proc. Natl. Acad. Sci. USA*, 77; 5172-5176.

Kass,S., Craig,N. and Sollner-Webb,B. (1987). Primary processing of mammalian rRNA involves two adjacent cleavages and is not species specific. *Mol. Cell. Biol.*, 7; 2891-2898.

Kedersha,N.L. and Rome,L.H. (1986). Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. J. Cell Biol., 103; 699-709.

Keene, J.D., Deutscher, S.L., Kenan, D. and Kelekar, A. (1987). Nature of the La and Ro RNPs. Mol. Biol. Rep., 12; 235-238.

Kempers-Veenstra, A.E., Oliemans, J., Offenberg, H., Dekker, A.F., Piper, P.W., Planta, R.J. and Klootwijk, J. (1986). **3'-end formation of transcripts** from the yeast rRNA operon. *EMBO J.*, 5; 2703-2710.

Khandekar, P., Saidapet, C., Krauskopf, M., Zarraga, A., Lin, W.-L., Mendola, C. and Siddiqui, M.A.Q. (1984). Co-ordinate control of gene expression: muscle-specific 7S RNA contains sequences homologous to 3'-untranslated regions of myosin genes and repetitive DNA. J. Mol. Biol., 180; 417-435.

Kim,S.-H. (1979). Crystal structure of yeast tRNA<sup>Phe</sup> and general structural features of other tRNAs. In Schimmel,P.R., Söll,D. and Abelson,J. (eds.) *Transfer RNA: Structure, Properties and Recognition*, Cold Spring Harbor Laboratory, New York, pp. 83-100.

Kish,V.M. and Pederson,T. (1975). Ribonucleoprotein organization of polyadenylate sequences in HeLa cell heterogeneous nuclear RNA. J. Mol. Biol., 95; 227-238.

Kiss,T., Tóth,M. and Solymosy,F. (1985). Plant small nuclear RNAs. Nucleolar U3 snRNA is present in plants: partial characterization. *Eur. J. Biochem.*, 152; 259-266.

Kiss,T., Antal,M. and Solymosy,F. (1987). Plant small nuclear RNAs. II. U6 RNA and a 4.5SI-like RNA are present in plant nuclei. *Nucl. Acids Res.*, 15; 543-560.

Kitajewski, J., Schneider, R.J., Safer, B., Munemitsu, S.M., Samuel, C.E., Thimmappaya, B. and Shenk, T. (1986). Adenovirus VAI RNA antagonizes the antiviral action of interferon by preventing activation of the interferon-induced eIF-2 $\alpha$  kinase. *Cell*, 45; 195-200.

Kloetzel, P.-M. (1987). 19S cylinder particles ubiquitous from plants to man: their morphology, molecular composition and potential functions. *Mol. Biol. Rep.*, 12; 223-227.

Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984). Recognition of cap structure in splicing *in vitro* of mRNA precursors. *Cell*, 38; 731-736.

Konings, D.A.M., Hesper, B. and Hogeweg, P. (1987). Evolution of the primary and secondary structures of the E1a mRNAs of the adenovirus. *Mol. Biol. Evol.*, 4; 300-314.

Krämer, A. (1987). Fractionation of HeLa cell nuclear extracts reveals minor small ribonucleoprotein particles. *Proc. Natl. Acad. Sci. USA*, 84; 8408-8412.

Krämer, A., Keller, W., Appel, B. and Lührmann, R. (1984). The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell*, 38; 299-307.

Kretzner, L., Rymond, B.C. and Rosbash, M. (1987). *S. cerevisiae* U1 RNA is large and has limited primary sequence homology to metazoan U1 snRNA. *Cell*, 50; 593-602.

Krieg, P.A. and Melton, D.A. (1987). *In vitro* RNA synthesis with SP6 RNA polymerase. *Methods in Enzymology*, 155; 397-415.

Kunkel,G.R., Maser,R., Calvet,J. and Pederson,T. (1986). U6 small nucler RNA is transcribed by RNA polymerase III. Proc. Natl. Acad. Sci. USA, 83; 8575-8579.

Lagunas, R. (1986). Misconceptions about the energy metabolism of *Saccharomyces cerevisiae.* Yeast, 2; 221-228.

Lazcano, A., Guerrero, R., Margulis, L. and Oró, J. (1988). The evolutionary transition from RNA to DNA in early cells. J. Mol. Evol., 27; 283-290.

Lee, C.A., Fournier, M.J. and Beckwith, J. (1985). *Escherichia coli* 6S RNA is not essential for growth or protein secretion. *J. Bacteriol.*, 161; 1156-1161.

Leinwand,L.A., Wydro,R.M. and Nada-Ginard,B. (1982). Small RNA molecules related to the *Alu* family of repetitive DNA sequences. *Mol. Cell. Biol.*, 2; 1320-1330.

LeLay-Taha,M.-N., Reveillaud,I., Sri-Wada,J., Brunel,C. and Jeanteur,P. (1986). RNA-protein organization of U1, U5 and U4/U6 small nuclear ribonucleoproteins in HeLa cells. J. Mol. Biol., 189; 519-532.

Lemmon, S.K. and Jones, E.W. (1987). Clathrin requirements for normal growth of yeast. *Science*, 238; 504-509.

Lerner, M.R. and Steitz, J.A. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA*, 76; 5495-5499.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J.A. (1980)a. Are snRNPs involved in splicing? *Nature*, 283; 220-224.

Lerner, M.R., Boyle, J.A., Mount, S.M., Weliky, J.L., Wolin, S.L. and Steitz, J.A. (1980)b. Structure and conservation of small ribonucleoprotein complexes in eukaryotic cells. In Osawa, S., Ozeki, H., Uchida, H. and Yura, T. (eds.) Genetics and evolution of RNA polymerase, tRNA and ribosomes, University of Tokyo Press/Elsevier, Amsterdam. pp299-311.

Lerner, M.R., Andrews, N.C., Miller, G. and Steitz, J.A. (1981). Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA*, 78; 805-809.

Letham, D.S. and Stewart, P.R. (1977). RNA in retrospect. In Stewart, P.R. and Letham, D.S. (eds.), *The Ribonucleic Acids*, Springer-Verlag. pp1-8.

Letham,D.S. and Wettenhall,R.E.H. (1977). Transfer RNA and cytokinins: Possible roles of tRNAs in the regulation of gene expression. In Stewart,P.R. and Letham,D.S. (eds.), *The Ribonucleic Acids*, Springer-Verlag. pp173-193.

Li,W.Y., Reddy,R., Henning,D., Epstein,P., Busch,H. (1982). Nucleotide sequence of 7S RNA: homology to Alu RNA and 4.5S RNA. J. Biol. Chem., 257; 5136-5142.

Liautard, J.-P., Sri-Widada, J., Brunel, C. and Jeanteur, P. (1982). Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. J. Mol. Biol., 162; 623-643.

Lin, R.-J., Newman, A.J., Cheng, S.-C. and Abelson, J. (1985). Yeast mRNA splicing in vitro. J. Biol. Chem., 260; 14780-14792.

Linnane, A.W. and Lukins, H.B. (1975). Isolation of mitichondria and techniqures for studying mitochondrial biogenesis in yeast. *Methods in Cell Biology*, XII; 285-309.

Lischwe,M.A., Ochs,R.L., Reddy,R., Cook,R.G., Yeoman,L.C., Tan,E.N., Reichlin,M. and Busch,H. (1985). Purification and partial characterization of a nuclear scleroderma antigen (Mr=34,000; pl, 8.5) rich in NG,NG-dimethylargenine. J. Biol. Chem., 260; 14304-14310.

Lorberboum,H., Digweed,M., Erdmann,V.A., Servadio,Y., Weinstein,D., De Groot,N. and Hochberg,A.A. (1986). Small cytoplasmic RNAs from human placental free mRNPs: structure and their effect on *in vitro* protein synthesis. *Eur. J. Biochem.*, 155; 279-287.

Lusena,C.V. and James,A.P. (1976). Alterations in mitochondrial DNA of yeast which accompany genetically and environmentally controlled changes in  $\rho^-$  mutability. *Mol. Gen. Genet.*, 144; 119-125.

Lustig, A.J., Lin, R.-J. and Abelson, J. (1986). The yeast RNA gene products are essential for mRNA splicing in vitro. Cell, 47; 953-963.

Lührmann, R. (1988). SnRNP proteins. In Birnstiel, M.L. (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Heidelberg. pp 71-99.

Lührmann,R., Appel,B., Bringmann,P., Rinke,R. and Rothe,S. (1982). Isolation and characterization of rabbit anti-m<sub>3</sub><sup>2,2,7</sup>G antibodies. *Nucl. Acid Res.*, 10; 7103-7113.

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

Maniatis,T. and Reed,R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature*, 325; 673-678.

Marcus-Sekura, C.J. and Hitchcock, M.J.M. (1987). Preparation of oocytes for microinjection of RNA and DNA. *Methods in Enzymology*, 152; 284-288.

Marmiroli,N., Restivo,F.M., Donnini,C., Bianchi, L. and Puglisi,P.P. (1980)a. Analysis of *rho* mutability in *Saccharomyces cerevisiae:* i. Effects of *mmc* and *pet*-ts alleles. *Molec. Gen. Genet.*, 177; 581-588.

Marmiroli,N., Donnini,C., Restivo,F.M., Tassi,F. and Puglisi,P.P. (1980)b. Analysis of *rho* mutability in *Saccharomyces cerevisiae:* II. Role of the mitochondrial protein synthesis. *Molec. Gen. Genet.*, 177; 589-595.

Martins de Sa,C., Grossi de Sa,M.-F., Akhayet,O., Broders,F. and Scherrer,K. (1986). **Prosomes: ubiquity and inter-species structural variation**. *J. Mol. Biol.*, 187; 479-493.

Mattaj,I.W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell, 46; 905-911.

Mattaj,I.W. (1988). U snRNP assembly and transport. In Birnstiel,M.L. (ed.) *Structure and function of major and minor small nuclear ribonucleoprotein particles.* Springer-Verlag, Heidelberg.

Mattaj,I.W. and De Robertis,E.M. (1985). Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell*, 40; 111-118.

Matter, L., Schopfer, K., Wilhelm, J.A., Nyffenegger, T., Parisot, R.F. and De Robertis, E.M. (1982). Molecular characterization of ribonucleoprotein antigens bound by antinuclear antibodies; a diagnostic evaluation. Arth. Rheum., 25; 1278-1283.

Maxam, A.M. and Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods in Enzymology*, 65; 499-560.

McCarthy, T.L., Siegel, E., Mroczkowski, B. and Heywood, S.M. (1983). Characterization of translational control ribonucleic acid isolated from embryonic chick muscle. Biochemistry, 22; 935-941.

Melton, D.A. (1987). Translation of messenger RNA in injected frog oocytes. Methods in Enzymology, 152; 288-296.

Messing, J. (1983). New M13 vectors for cloning. *Methods in Enzymology*, 101; 20-78.

Messing, J., Gronenborn, B., Müller-Hill, B. and Hofschneider, P. (1977). Filamentous coliphage M13 as a cloning vehicle: insertion of a Hindlil fragment of the *lac* regulatory region in M13 replicative form *in vitro*. *Proc. Natl. Acad. Sci. USA*, 74; 3642-3646.

Michel, F. and Dujon, B. (1983). Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast and nuclear-encoded members. *EMBO J.*, 2; 33-38.

Michel, F. and Jacquier, A. (1987). Long-range intron-exon and intronintron pairings involved in self-splicing of class II catalytic introns. Cold Spring Harbor Symp. Quant. Biol., LII; 201-221.

Mimori,T., Hinterberger,M., Pettersson,I. and Steitz,J.A. (1984). Autoantibodies to the U2 small nuclear ribonucleoprotein in a patient with sclerodermapolymyositis overlap syndrome. J. Biol. Chem., 259; 560-565.

Moritz, A. and Goebel, W. (1985). Characterization of the 7S RNA and its gene from halobacteria. Nucl. Acids Res., 13; 6969-6979.

Mortimer, R.K. and Hawthorne, D.C. (1969). Yeast genetics. In Rose, A.H. and Harrison, J.S., (eds.) *The Yeasts*, Vol. 1, Academic Press, London, pp. 385-460.

Mount,S.M., Pettersson,I., Hinterberger,M., Karmas,A. and Steitz,J.A. (1983). The U1 small nuclear RNA-protein complex selectively binds a 5' splice site *in vitro*. *Cell*, 33; 509-518.

Muekler, M. and Lodish, H.F. (1986). The human glucose transporter can insert posttranslationally into microsomes. *Cell*, 44; 629-637.

Murphy,S., Altruda,F., Ullu,E., Tripodi,M., Silengo,L. and Melli,M. (1984). The DNA sequences complementary to human 7SK RNA show structural similarities to the short mobile elements of the mammalian genome. *J. Mol. Biol.*, 177; 575-590.

Murphy,S. Di Liegro,C. and Melli,M. (1987).The *in vitro* transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell*, 51; 81-87.

Murray, M.G. (1986). Use of sodium trichloroacetate and mung bean nuclease to increase sesitivity and prescision during transcript mapping. *Analyt. Biochem.* 158; 165-170.

Murray, N.E. (1983). Phage lambda and molecular cloning. In Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (eds.) Lambda II, Cold Spring Harbor Laboratory, New York. pp395-432.

Nakamura, T., Prestayko, A.W. and Busch, H. (1968). Studies on nucleolar 4 to 6S ribonucleic acid of Novikoff hepatoma cells. J. Biol. Chem., 243; 1368-1375.

Nelböck, P., Stucka, R. and Feldmann, H. (1985). Different patterns of transposable elements in the vicinity of tRNA genes in yeast: a possible clue to transcriptional modulation. *Hoppe-Seglers Z. Physiol. Chem.*, 366; 1041-1051.

Neuman de Vegvar, H.E., Lund, E. and Dahlberg, J.E. (1986). 3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. *Cell*, 47; 259-266.

Newlon,C.S., Ludescher,R.D. and Walter,S.K. (1979). Production of petites by cell cycle mutants of *Saccharomyces cerevisiae* defective in DNA synthesis. *Molec. Gen. Genet.*, 169; 189-194.

Noller,H.F., Kop,J., Wheaton,V., Brosius,J., Gutell,R.R., Kopylov,A., Dohme,F. and Herr,W. (1981). A secondary structure model for 23S ribosomal RNA. Nucl. Acids Res., 9; 6167-6189.

Noller, H.F. (1984). The structure of ribosomal RNA. Ann. Rev. Biochem., 53; 119-162.

Nomura, M. (1987). The role of RNA and protein in ribosome function: a review of early reconstitution studies and prospects for future studies. Cold Spring Harbor Symp. Quant. Biol., LII; 653-663.

Nomura,T., Fujita,N. and Ishihama,A. (1986). Promoter selectivity of *Escherichia coli* RNA polymerase: alteration by fMet-tRNA<sup>Met</sup>. Nucl. Acids Res., 14; 6857-6870.

Norrander, J., Kempe, T. and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene*, 26; 101-106.

Okamoto,K. and Freundlich,M. (1986). Mechanism for the autogenous control of the *crp* operon transcriptional inhibition by a divergent RNA transcript. *Proc. Natl. Acad. Sci. USA*, 83; 5000-5004.

Orgel,L.E. (1987). Evolution of the genetic apparatus: a review. Cold Spring Harbor Symp. Quant. Biol., LII; 9-16.

Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. (1981). Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci. USA*, 78; 6354-6358.

Pace,N.R., Reich,C., James,B.D., Olsen,G.J., Pace,B. and Waugh,D.S. (1987). Structure and catalytic function in ribonuclease P. Cold Spring Harbor Symp. Quant. Biol., LII; 239-248.

Padgett,R.A., Mount,S.M., Steitz,J.A. and Sharp,P.A. (1983). Splicing of messenger RNA precursors is inhibited by antisera to small nuclear ribonucleoprotein. *Cell*, 35; 101-107.

Padgett, R.A., Grabowski, P.G., Konarska, M.M., Seiler, S.R. and Sharp, P.A. (1986). **Splicing of messenger RNA precursors**. *Ann. Rev. Biochem.*, 55; 1119-1150.

Panet,A., Haseltine,W.A., Baltimore,D., Peters,G., Harada,F. and Dahlberg,J.E. (1975). Specific binding of tryptophan transfer RNA to avian myeloblastosis virus RNA-dependent DNA polymerase (reverse transcriptase). *Proc. Natl. Acad. Sci. USA*, 72; 2535-2539.

Parker,K.A. and Steitz,J.A. (1987). Structural analyses of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. *Mol. Cell. Biol.*, 7; 2899-2913.

Parker, R. and Guthrie, C. (1985). A point mutation in the conserved hexanucleotide at a yeast 5' splice junction uncouples recognition, cleavage and ligation. *Cell*, 41; 107-118.

Parker, R., Siliciano, P.G. and Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base-pairing to the U2-like snRNA. *Cell*, 49; 229-239.

Parker, R., Simmons, T., Shuster, E.O., Siliciano, P.G. and Guthrie, C. (1988). Genetic analysis of small nuclear RNAs in *Saccharomyces cerevisiae*: viable sextuple mutant. *Mol. Cell. Biol.*, 8; 3150-3159.

Patterson, B. and Guthrie, C. (1987). An essential yeast snRNA with a U5like domain is required for splicing in vivo. *Cell*, 49; 613-624.

Payne,G.S. and Scheckman,R. (1985). A test of clathrin function in protein secretion and cell growth. *Science*, 230; 1009-1014.

Payne,G.S., Hasson,T., Hasson,M. and Schekman,R. (1987). Genetic and biochemocal characterization of clathrin-deficient *Saccharomyces* cerevisiae. *Mol. Cell. Biol.*, 7; 3888-3898.

Peacock,A.C. and Dingman,C.W. (1967). **Resolution of multiple ribonucleic acid species by polyacrylamide gel electrophoresis**. *Biochemistry*, 6; 1818-1827.

Perbel, B. (1984). A practical guide to molecular cloning. John Wiley & Sons, New York. pp 441-486.

Petes, T.D. (1979). Yeast ribosomal DNA genes are located on chromosome XII. Proc. Natl. Acad. Sci. USA, 76; 410-414.

Petrova, A.N. and Belykh, G.F. (1982). Formation of an intermediate polysaccharide-RNA complex during glycogen biosynthesis. *Dokl. Akad. Nauk SSSR*, 268; 27-30 (English).

Petrova, A.N. and Rudik, O.A. (1986). Enzymatic properties of the 2.5S RNA from the branching enzyme. Dokl. Akad. Nauk SSSR, 289; 281-283 (English).

Pettersson, I., Hinterberger, M., Mimori, T., Gottlieb, E. and Steitz, J.A. (1984). The structure of mammalian small nuclear ribonucleoproteins. Identification of multiple protein components reactive with anti-(U1)ribonucleoprotein and anti-Sm autoantibodies. J. Biol. Chem., 259; 5907-5914.

Philippsen, P., Thomas, M., Kramer, R.A. and Davis, R.W. (1978). Unique arrangement of coding sequences for 5S, 5.8S, 18S and 25S ribosomal RNA in *Saccharomyces cerevisiae* as determined by R-loop and hybridization analysis. J. Mol. Biol., 123; 387-404.

Pikielny, C.W. and Rosbash, M. (1985). mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell, 41; 119-126.

Pikielny, C.W. and Rosbash, M. (1986). Specific small nuclear RNAs are associated with yeast spiiceosomes. Cell, 45; 869-877.

Pikielny, C.W., Rymond, B.C. and Rosbash, M. (1986). Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature*, 324; 341-345.

Pinnas, J.L., Northway, J.D. and Tan, E.M. (1973). Antinucleolar antibodies in human sera. J. Immnuol., 111; 996-1004.

Piot, E., Backhovens, H. and Slegers, H. (1984). The inhibitor ribonucleoprotein of poly(A)-containing non-polysomal messenger ribonicleoprotein of *A. salina* cryptobiotic embryos. *FEBS* Letters, 175; 16-20.

Poritz, M.A., Siegel, V., Hansen, W. and Walter, P. (1988). Small ribonucleoproteins in Schizosaccharomyces pombe and Yarrowia *lipolytica* homologous to signal recognition particle. Proc. Natl. Acad. Sci. USA, 85; 4315-4319.

Prestayko, A.W., Tonato, M. and Busch, H. (1970). Low molecular weight RNA associated with 28S nucleolar RNA. J. Mol. Biol., 47; 505-515.

Reddy,R. (1985). Compilation of small RNA sequences. Nucl. Acids Res. 13; Suppl. r155-163.

Reddy,R. and Busch,H. (1983). Small nuclear RNAs and RNA processing. *Prog. in Nucl. Acid Res. and Mol. Biol.*, 30; 127-162.

Reddy,R. and Busch,H. (1988). Small nuclear RNAs: RNA sequences, structure and modification. In Birnstiel,M.L. (ed.) Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Heidelberg.

Reddy,R., Li,W.-Y., Henning,D., Choi,Y.C., Nohga,K. and Busch,H. (1981). Characterization and subcellular localization of 7-8 S RNAs of Novikoff hepatoma. J. Biol. Chem., 256; 8452-8457.

Reddy,R., Tan,E.M., Henning,D., Nohga,K. and Busch,H. (1983). Detection of nucleolar 7-2 ribonucleoprotein and a cytoplasmic 8-2 ribonucleoprotein with autoantibodies from patients with scleroderma. J. Biol. Chem., 258; 1383-1386.

Reddy,R., Henning,D. amd Busch,H. (1985) Primary and secondary structure of U8 small nuclear RNA. J. Biol. Chem., 260; 10930-10935.

Reddy,V.B., Thimmappaya,B., Dhar,R., Subramanian,K.N., Zain,B.S., Pan,J., Ghosh,P.K., Celma,M.L., and Weissman,S.M. (1978). The genome of simian virus 40. *Science*, 200; 494 -502.

Reich,C., Olsen,G.J., Pace,B. and Pace,N. (1988). Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science*, 239; 178-181.

Reimer,G., Pollard,K.M., Penning,C.A., Ochs,R.L., Lischwe,M.A., Busch,H. and Tan,E.M. (1987). Monoclonal autoantibody from a (New Zealand black x New Zealand white) $F_1$  mouse and some human scleroderma sera target an  $M_r$  34000 nucleolar protein of the U3 RNP particle. Arthritis and Rheumatism, 30; 793-800.

Reyes, A.A. and Wallace, R.B. (1987). Mapping RNA using S1 nuclease and synthetic oligonucleotides. *Methods in Enzymology*, 154; 87-94.

Ribes,V., Dehoux,P. and Tollervey,D. (1988). **7SL RNA from** Schizoscchyromyces pombe is encoded by a single copy essential gene. EMBO J., 7; 231-237.

Riedel,N., Wise,J.A., Swerdlow,H., Mak,A. and Guthrie,C. (1986). Small nuclear RNAs from Saccharomyces cerevisiae: unexpected diversity in abundance, size and molecular complexity. Proc. Natl. Acad. Sci. USA, 83; 8097-8101.

Riedel,N., Wolin,S. and Guthrie,C. (1987). A subset of yeast snRNAs contains functional binding sites for the highly conserved Sm antigen. *Science*, 235; 328-331.

Rinke, J. and Steitz, J.A. (1985). Association of the lupus antigen La with a subset of U6 snRNA molecules. Nucl. Acids Res., 13; 2617-2629.

Rogers, J. and Wall, R. (1980). A mechanism for RNA splicing. Proc. Natl. Acad. Sci. USA, 77; 1877-1879.

Roman,H. (1956). Studies of gene mutaions in Saccharomyces. Cold Spring Harbor Symp. Quant. Biol., XXI; 175-185.

Roman,H. (1981). Development of yeast as an experimental organism. In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, pp1-9.

Rose,M. and Botstein,D. (1983). Structure and function of the yeast URA3 gene: Differentially regulated expression of hybrid  $\beta$ -galactosidase from overlapping coding sequences in yeast. J. Mol. Biol., 170; 883-904.

Rose, M., Grisafi, P. and Botstein, D. (1984). Structure and function of the yeast URA3 gene: expression in Escherichia coll. Gene, 29; 113-124.

Rothblatt, J.A. and Meyer, D.I. (1986). Secretion in yeast: reconstitution of the translocation and gylcosylation of a-factor and invertase in a homologous cell free system. *Cell*, 44; 619-28.

Rothstein, R.J. (1983) **One-step gene disruption in yeast.** Methods Enzymol., 101; 202-211.

Rubin,G.M. (1974). Three forms of the 5.8S ribosomal RNA species in *Saccharomyces cerevisiae*. *Eur. J. Biochem.*, 41; 197-202.

Rubin,G.M. (1975) Preparation of RNA and ribosomes from yeast. *Methods in Cell Biology*, XII; 45-64.

Rubtsov, P.M., Musakhanov, M.M., Zakharyev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980). The structure of the yeast ribosomal RNA genes. 1. The complete nucleotide sequence of the 18S ribosomal RNA gene from Saccharomyces cerevisiae. Nucl. Acids Res., 8; 5779-5794.

Ruskin,B., Zamore,P.D. and Green,M.R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, 52; 207-219.

Sampson, J.R., Sullivan, F.X., Behlen, L.S., DiRenzo, A.B. and Uhlenbeck, O.C. (1987). Characterization of two RNA-catalysed RNA cleavage reactions. *Cold Spring Harbor Symp. Quant. Biol.*, LII; 267-275.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chainterminating inhibitors. *Proc. Natl. Acad. Sci. USA*, 74; 5463-5467.

Saponara, A.G. and Enger, M.D. (1969). Occurrence of 2,2,7-trimethylguanosine in minor RNA species of a mammalian cell line. *Nature*, 223; 1365-1366.

Sarkar,S., Mukherjee,A.K. and Guha,C. (1981). A ribonuclease-resistant cytoplasmic 10S ribonucleoprotein of chick embryonic muscle: a potent inhibitor of cell-free protein synthesis. J. Biol. Chem., 256; 5077-5086.

Schaufele, F., Gilmartin, G.M., Bannwarth, W. and Birnstiel, M.L. (1986). Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of H3 messenger RNA. *Nature*, 323; 777-781.

Schekman, R. and Novick, P. (1982). The secretory process and yeast cell-surface assembly. In Strathern, J.N., Jones, E.W. and Broach, J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*, Cold Spring Harbor Laboratory, New York. pp361-398.

Schulz,L.D. (1978) Transcriptional role of yeast deoxyribonucleic acid dependent ribonucleic acid polymerase III. *Biochemistry*, 17; 750-758.

Sharp, P.A. (1985). On the origin of RNA splicing and introns. Cell, 42; 397-400.

Sharp, P.A. (1987). Splicing of messenger RNA precursors. Science, 235; 766-771.

Sharp, P.A., Konarska, M.M., Grabowski, P.J., Lamond, A.I., Marciniak, R. and Seiler, S.R. (1987). Splicing of messenger RNA precursors. Cold Spring Harbor Symp. Quant. Biol., LII; 277-285.

Sherman, F., Fink, G.R. and Hicks, J.B. (1983) Methods in Yeast Genetics, Cold Spring Harbor Laboratory, New York.

Shine, J. and Dalgarno, L. (1975)a. Determinant of cistron specificity in bacterial ribosomes. *Nature*, 254; 34-38.

Shine, J. and Dalgarno, L. (1975)b. Terminal sequence analysis of bacterial ribosomal RNA. Eur. J. Biochem., 57; 221-230.

Shvedova,T.A., Korneeva,G.A., Otroshchenko,V.A. and Venkstern,T.V. (1987). Catalytic activity of the nucleic acid component of the 1,4- $\alpha$ -glucan branching enzyme from rabbit muscles. *Nucl. Acids Res.*, 15; 1745-1752.

Siegel,V. and Walter,P. (1985). Elongation arrest is not a prerequisite for secretory protein translocation across the microsomal membrane. J. Cell Biol., 100; 1913-1921.

Siegel, V. and Walter, P. (1986). Removal of the *Alu* structural domain of from signal recognition particle leaves its protein translocation activity intact. *Nature*, 320; 81-84.

Siegel, V. and Walter, P. (1988). Binding sites of the 19-kDa and 68/72-kDa signal recognition particle (SRP) proteins on SRP RNA as determined by protein-RNA "footprinting". *Proc. Natl. Acad. Sci. USA*, 85; 1801-1805.

Siliciano, P.G. and Guthrie, C. (1988). 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes and Development, in press.

Siliciano, P.G., Brow, D.A., Roiha, H. and Guthrie, C. (1987)a. An essential snRNA from S. cerevisiae has properties predicted for U4, including interaction with a U6-like snRNA. Cell, 50; 586-592.

Siliciano, P.G., Jones, M.H. and Guthrie, C. (1987)b. *Saccharomyces cervisiae* has a U1-like small nuclear RNA with unexpected properties. *Science*, 237; 1484-1487.

Sillevis Smitt, W.W., Vlak, J.M., Molenaar, I. and Rozijn, Th.H. (1973). Nucleolar function of the dense crescent in the yeast nucleus. *Exptl. Cell Res.*, 80; 313-321.

Skryabin,K.G., Eldarov,M.A., Larionov,V.L., Bayev,A.A., Klootwijk,J., de Regt,V.C.H.F., Veldman,G.M., Planta,R.J., Georgiev,O.I. and Hadjiolov,A.A. (1984). Structure and function of the non-transcribed spacer regions of yeast rDNA. Nucl. Acids Res., 9; 2955-2968.

Slonimski, P.P., Perrodin, G. and Croft, J.H. (1968). Ethidium bromideinduced mutation of yeast mitochondria: complete transformation of cells into respirtory-deficient non-chromosomal "petites". *Biochem. Biophys. Res. Commun.*, 30; 232.

Soffer, R.L. (1974). Aminoacyl-tRNA transferases. Adv. Enzymol., 40; 91-139.

Söll,D., Abelson,J.N. and Schimmel,P.R. (eds.) (1980). Transfer RNA: bioogical aspects. Cold Spring Harbor Laboratory, New York.

Steitz, J.A., Berg,C., Gottlieb,E., Hardin,J.A., Hashimoto,C., Hendrick,J.P., Hinterberger,M., Krikeles,M., Lerner,M.R., Mount,S.M., Pettersson,I., Rinke,J., Rosa,M.D. and Wolin,S.L. (1982). Structure and function of small ribonucleoproteins from eukaryotic cells. In Miwa,M. et al., (eds.) *Primary and Tertiary Structure of Nucleic Acids and Cancer Res.*, Japan Sci. Soc. Press, Tokyo, pp. 101-107.

Steitz, J.A., Black, D.L., Gerke, V. and Parker, K.A. (1988). Functions of the abundant U-snRNPs. In Birnstiel, M.L. (ed.), *Structure and function of major and minor small nuclear ribonucleoprotein particles*. Springer-Verlag, Heidelberg.

Stroke, I.L. and Weiner, A.M. (1985). Genes and pseudogenes for rat U3A and U3B small nuclear RNA. J. Mol. Biol., 184; 183-193.

Strub, C. and Birnstiel, M. (1986). Genetic complementation in the *Xenopus* oocyte: co-expression of sea urchin histone and U7 RNAs restores 3' processing of H3 pre-mRNA in the oocyte. *EMBO J.*, 5; 1675-1682.

Strub,C., Galli,G., Busslinger,M. and Birnstiel,M.L. (1984). The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. *EMBO J.*, 3; 2801-2807.

Struck, J.C.R., Vogel, D.W., Ulbrich, N. and Erdmann, V.A. (1988)a. The Bacillus subtilis scRNA is related to the 4.5S RNA from Escherichia coli. Nucl. Acids Res., 16; 2719.

Struck, J.C.R., Toschka, H.Y.. Specht, T. and Erdmann, V.A. (1988)b. Common structural features between eukaryotic 7SL RNAs, eubacterial 4.5S RNA and scRNA and archebacterial 7S RNA. Nucl. Acids Res., 16; 7740.

Struhl,K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promotor elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA, 82; 8419-8423.

Struhl,K. (1987). Promoters, activator proteins, and the mechanism of transcription initiation in yeast. Cell, 49; 295-297.

Struhl,K., Stinchcomb,D.T., Scherer,S. and Davis,R.W. (1979). High frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl. Acad. Sci. USA*, 76; 1035-1039.

Suh,D., Busch,H. and Reddy,R. (1986). Isolation and characterization of a human U3 small nucleolar RNA gene. *Biochem. Biophys. Res. Commun.*,137; 1133-1140.

Sundaralingam,M. (1979). The structure, conformation and interaction of tRNA. In Schimmel,P.R., Söll,D. and Abelson,J. (eds.) *Transfer RNA: Structure, Properties and Recognition.* Cold Spring Harbor Laboratory, New York, pp. 115-132.

Sutcliffe,G.J., Milner,R.J., Gottesfeld,J.M. and Lerner,R.A. (1984). Identifier sequences are transcribed specifically in brain. *Nature*, 308; 237-241.

Tague, B.W. and Gerbi, S.A. (1984). **Processing of the large rRNA precursor: two proposed categories of RNA-RNA interactions in eukaryotes**. J. Mol. Evol., 20; 362-367.

Tan,E.M. (1982). Autoantibodies to nuclear antigens (ANA): their immunology and medecine. Adv. Immunol., 33; 167-240.

Tanaka,K., li,K. and Ichihara,A. (1986). A high molecular weight protease in the cytosol of rat liver. J. Biol. Chem., 261; 15197-15203.

Tazi, J., Alibert, C., Temsamani, J., Reveillaud, I., Cathala, G., Brunel, C. and Jeanteur, P. (1986). A protein that specifically recognizes the 3' splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. *Cell*, 47; 755-766.

Thomas, J.H. and Botstein, D. (1986). A gene required for the separation of chromosomes on the spindle apparatus in yeast. *Cell*, 44; 65-76.

Tollervey, D. (1987)a. A yeast small nuclear RNA is required for normal processing of pre-rRNA. *EMBO J.*, 6; 4169-4175.

Tollervey, D. (1987)b. High level of complexity of small nuclear RNAs in fungi and plants. J. Mol. Biol., 196; 355-361.

Tollervey, D. and Guthrie, C. (1985). Deletion of a yeast small nuclear RNA gene impairs growth. EMBO J., 4; 3873-3878.

Tollervey, D. and Mattaj, I.W. (1987). Fungal small nuclear ribonucleoproteins share properties with plant and vertebrate U-snRNPs. *EMBO J.*, 6; 469-476.

Tollervey, D., Wise, J.A. and Guthrie, C. (1983). A U4-like small nuclear RNA is dispensable in yeast. *Cell*, 35; 753-762.

Trinh-Rohlik, Q. and Maxwell, E.S. (1988). Homologous genes for mouse 4.5S hybRNA are found in all eukaryotes and their low molecular weight RNA transcripts intermolecularly hybridize with eukaryotic 18S ribosomal RNAs. *Nucl. Acids Res.*, 16; 6041-6056.

Tschumper, G. and Carbon, J. (1980). Sequence of a yeast DNA fragment containing a chromosomal replicator and the *TRP1* gene. *Gene*, 10; 157-166.

Ullu, E. and Tschudi, C. (1984). *Alu* sequences are processed 7SL RNA genes. *Nature*, 312; 171-172.

Ullu, E., Esposito, V. and Melli, M. (1982)a. Evolutionary conservation of the human 7S RNA sequences. J. Mol. Biol., 161; 195-201.

Ullu,E., Murphy,S. and Melli,M. (1982)b. Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an Alu sequence. Cell, 29; 195-202.

Van der Veen, R., Arnberg, A.C. and Grivell, L.A. (1987). Self-splicing of a group II intron in yeast mitochondria: dependence on 5' exon sequences. *EMBO J.*, 6; 1079-1084.

Veldman,G.M., Klootwijk,J., van Heerikhuizen,H. and Planta,R. (1981). The nucleotide sequence of the intergenic region between the 5.8S and 26S rRNA of the yeast ribosomal RNA operon. Possible implications for the interaction between 5.8S and 26S rRNA and the processing of the primary transcript. *Nucl. Acids Res.*, 9; 4847-4862.

Vieira, J. and Messing, J. (1982). The pUC plasmids, and M13mp7derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, 19; 259-268.

Walmsley, R.W. (1987). Yeast telomeres: the end of the chromosome story. Yeast, 3; 139-148.

Walter, P. and Blobel, G. (1980). Purification of a membrane associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA*, 77; 7112-7115.

Walter, P. and Blobel, G. (1983). Signal recognition particle: a ribonucleoprotein required for cotranslational translocation of protiens, isolation and properties. *Methods in Enzymology*, 96; 682-691.

Walter, P. and Lingappa, R. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. Ann. Rev. Cell Biol., 2; 499-516.

Waltscheva,L., Georgiev,O. and Venkov,P. (1983). Relaxed mutant of *Saccharomyces cerevisiae:* proper maturation of ribosomal RNA in absence of protein synthesis. *Cell*, 33; 221-230.

Waters, M.J., Chirico, W.J. and Blobel, G. (1986). Protein translocation across the yeast microsomal membrane is stimulated by a soluble factor. *J. Cell Biol.*, 103; 2629-2636.

Weinberg, R. and Penman, S. (1968). Small molecular weight monodisperse nuclear RNA. J. Mol. Biol., 38; 289-304.

Weiner, A.M. and Maizels, N. (1987). tRNA-like structures tag the 3' ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc. Natl. Acad. Sci. USA*, 84; 7383-7387.

Wiemken, A. (1975). Isolation of vacuoles from yeast. *Methods in Cell Biology*, XII; 99-109.

Wilkie, D. (1975). Cytoplasmic inheritance and mitochondrial genetics in yeasts. *Methods in Cell Biology*, XII; 353-372.

Williams, A.L. and Tinoco, I. (1986). A dynamic programming algorithm for finding alternative RNA secondary structures. *Nucl. Acid Res.*, 14; 299-315.

Winkler, M.M., Lashbrook, C., Hershey, J.W.B., Mukherjee, A.K. and Sarkar, S. (1983). The cytoplasmic 4S translation inhibitory RNA species of chick embryonic muscle. J. Biol. Chem., 258; 15141-15145.

Wintersberger, U., Smith, P. and Letnansky, K. (1973). Yeast chromatin. Preparation from isolated nuclei, histone composition and transcription capacity. *Eur. J. Biochem.*, 33; 123-130.

Wise, J.A. and Weiner, A.M. (1980). *Dictyostelium* small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. *Cell*, 22; 109-118.

Wise, J.A., Tollervey, D., Maloney, D., Swerdlow, H., Dunn, E.J. and Guthrie, C. (1983). Yeast contains small nuclear RNAs encoded by single copy genes. *Cell*, 35; 743-751.

Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983). Detailed analysis of the higher order structure of 16S-like ribosomal ribonucleic acids. *Microbiol. Rev.*, 47; 621-669.

Wolin,S. and Steitz,J.A. (1983). Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell*, 32; 735-744.

Wong, T.W. and Clayton, D.A. (1986). DNA primase of human mitochondria is associated with structural RNA that is essential for enzymatic activity. *Cell*, 45; 817-825.

Yang,V.W., Lerner,M.R., Steitz,J.A. and Flint,S.J. (1981). A small nuclear ribonucleoprotein is required for splicing of adenoviral early RNA sequences. *Proc. Natl. Acad. Sci. USA*, 78; 1371-1375.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 33; 103-119.

Yanofsky, C. (1988) Trancription attenuation. J. Biol. Chem., 263; 609-612.

Zagorski, J., Tollervey, D. and Fournier, M.J. (1988). Characterization of an SNR gene locus in Saccharomyces cerevisiae that specifies both dispensible and essential small nuclear RNAs. Mol. Cell. Biol., 8; 3282-3290.

Zamb,T.J. and Petes,T.D. (1982). Analysis of the junction between ribosomal RNA genes and single-copy chromosomal sequences in the yeast *Saccharomyces cerevisiae*. *Cell*, 28; 355-364.

Zhuang,Y. and Weiner,A.M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell*, 46; 827-835.

Zhuang,Y., Leung,H. and Weiner,A.M. (1987). The natural 5' splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. *Mol. Cell. Biol.*, 7; 3018-3020.

Zieve, G. and Penman, S. (1976). Small RNA species of the HeLa cell: metabolism and subcellular localization. *Cell*, 8; 19-31.

Zieve, G., Benecke, B.-J. and Penman, S. (1977). Synthesis of two classes of small RNA species in vivo and in vitro. Biochemistry, 16; 4520-4525.

Zieve,G.W., Sauterer,R.A., and Freeney,R.J. (1988). Newly synthesized small nuclear RNAs appear transiently in the cytoplasm. J. Mol. Biol., 199; 259-267.

Zillmann,M. Rose,S.D. and Berget,S.M. (1987). U1 small nuclear ribonuclear proteins are required early during spliceosome assembly. *Mol. Cell. Biol.*, 7; 2877-2883.

Zimmermann, R. and Mollay, C. (1986). Import of honeybee prepromelittin into the endoplasmic reticulum. J. Biol. Chem., 261; 12889-12895. Zissler, J., Signer, E.R. and Schaefer, F. (1971). The role of recombination in the growth of bacteriophage lambda: II. inhibition of growth by prophage P2. In Hershey, A.D. (ed.) *The bacteriophage lambda*, Cold Spring Harbor Laboratory, New York. p469-475.

Zoller, M.J. and Smith, M. (1984). Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a singlestranded DNA template. DNA, 3; 479-488.

Zucker, M. and Sankoff, D. (1984). RNA secondary structures and their prediction. Bull. Math. Biol., 46; 591-621.

Zwieb, C. (1985). The secondary structure of the 7SL RNA in the signal recognition particle: functional implications. *Nucl. Acids Res.*, 13; 6105-6124.