An Analysis of Lsm Protein Complexes

Martin Reijns

A thesis presented for the degree of Doctor of Philosophy
The University of Edinburgh
December 2006

Declaration

I hereby declare that I alone have composed this thesis and that, except where stated otherwise, the work presented here is my own.

Martin Reijns,
December 2006

Acknowledgements

First of all I'd like to thank Jean (Professor Jean D. Beggs CBE, FRS that is) for giving me the chance to do my PhD project in her lab and for her guidance and encouragement throughout three years and three months of hardship (:-p). Thank you also to everyone else in the Beggs' lab, present and past members (in alphabetical order: Boon, David, Florance, Karen, Liz, Marian, Michael, Olivier, Parastoo, Richard, Ross, Steve, Tatsiana and Vanessa) for their help both in and outside the lab. In particular thank you to Olivier and Steve for ploughing through this thesis (well, some of it anyway) to point out any typos and other mistakes or ambiguities (sorry it wasn't on quilted paper!). Thank you to everyone else in the Michael Swann and Darwin Buildings for their help, advice, materials and/or friendship. I'd like to thank David Leach and lab members for materials and advice that helped me create the *hfq* deletion strains; Tollervey lab members for help, advice and materials; Kevin Hardwick and lab members for help, advice and materials; Adele Marston for advice and materials; Ohkura lab members for friendship and cakes; Martin Wear and others on the third floor for help and advice on, and materials for gel filtration; David Brown and the people in the kitchen for making everything run a little bit more smoothly. Thank you also to the Wellcome Trust for their generous financial support.

Thank you to people on my postgraduate committee for feedback: David Tollervey, Paul McLaughlin, Margarete Heck and Robin Allshire. Thank you to examiners David Tollervey and Reinhard Lührmann for an interesting discussion during my PhD defense.

Thank you to other researcher for providing materials: Poul Valentin-Hansen for *E. coli* strains, Andrew Feig, Christine Guthrie and Roy Parker for plasmids, Steven Rader and Mark Schmitt for yeast strains and plasmids, Sandra Wolin and Javier Jiménez for yeast strains, Olivier Cordin for Ded1 protein and Christian Kambach for human Lsm proteins.

Thank you to friends and family for making me forget about the lab once in a while. Thank you to my parents for financial and moral support throughout my career as a "professional" student (DANKJEWEL!). But most of all, thank you to my wife Karen for putting up with my bad moods because of yet another bad day in the lab and for everything else you've done to support me and keep me sane. THANK YOU!

Acknowledgements

First of all I'd like to thank Jean (Professor Jean D. Beggs CBE, FRS that is) for giving me the chance to do my PhD project in her lab and for her guidance and encouragement throughout three years and three months of hardship (:-p). Thank you also to everyone else in the Beggs' lab, present and past members (in alphabetical order: Boon, David, Florance, Karen, Liz, Marian, Michael, Olivier, Parastoo, Richard, Ross, Steve, Tatsiana and Vanessa) for their help both in and outside the lab. In particular thank you to Olivier and Steve for ploughing through this thesis (well, some of it anyway) to point out any typos and other mistakes or ambiguities (sorry it wasn't on quilted paper!). Thank you to everyone else in the Michael Swann and Darwin Buildings for their help, advice, materials and/or friendship. I'd like to thank David Leach and lab members for materials and advice that helped me create the *hfq* deletion strains; Tollervey lab members for help, advice and materials; Kevin Hardwick and lab members for help, advice and materials; Adele Marston for advice and materials; Ohkura lab members for friendship and cakes; Martin Wear and others on the third floor for help and advice on, and materials for gel filtration; David Brown and the people in the kitchen for making everything run a little bit more smoothly. Thank you also to the Wellcome Trust for their generous financial support.

Thank you to people on my postgraduate committee for feedback: David Tollervey, Paul McLaughlin, Margarete Heck and Robin Allshire. Thank you to examiners David Tollervey and Reinhard Lührmann for an interesting discussion during my PhD defense.

Thank you to other researcher for providing materials: Poul Valentin-Hansen for *E. coli* strains, Andrew Feig, Christine Guthrie and Roy Parker for plasmids, Steven Rader and Mark Schmitt for yeast strains and plasmids, Sandra Wolin and Javier Jiménez for yeast strains, Olivier Cordin for Ded1 protein and Christian Kambach for human Lsm proteins.

Thank you to friends and family for making me forget about the lab once in a while. Thank you to my parents for financial and moral support throughout my career as a "professional" student (DANKJEWEL!). But most of all, thank you to my wife Karen for putting up with my bad moods because of yet another bad day in the lab and for everything else you've done to support me and keep me sane. THANK YOU!

Abstract

Lsm (Sm-Like) protein complexes have been implicated in a large number of RNA processing events in every kingdom of life. In *Saccharomyces cerevisiae* two distinct heteroheptameric complexes exist: Lsm1-7p in the cytoplasm where it is involved in mRNA decapping and Lsm2-8p in the nucleus where it is implicated in a range of processes from pre-mRNA splicing and degradation to pre-snoRNA, pre-tRNA and pre-rRNA processing. Their involvement in these processes and all that is known about these and other Sm-like proteins suggest that they function as RNA chaperones. To investigate this function, recombinant yeast Lsm proteins were purified from *Escherichia coli* and tested for their ability to promote annealing of the U4 and U6 snRNAs and to unwind a DNA/RNA duplex resembling the 3' stem-loop of U6 snRNA. For comparison, Hfq, the Sm-like protein from *E. coli,* and recombinant human Lsm complexes were used in the same *in vitro* assays. The results are consistent with their ability to promote RNA/RNA annealing and to modulate RNA secondary structure, which, *in vivo*, may also allow them to affect RNA/protein interactions. To complement the *in vitro* analyses, studies were carried out to elucidate the contributions made by individual Lsm polypeptides to the function and localisation of these complexes *in vivo*. The function of *LSM4* was studied by overexpression and depletion of wild-type Lsm4p, and by expression of Lsm4p C-terminal deletion mutants. Results suggest that Lsm4p affects cell morphology and that its C-terminus promotes efficient recruitment of Lsm1-7p to P-bodies. *LSM5* was shown to be dispensable for cell viability, and its deletion was shown to affect levels of U6 and U4/U6 RNAs similar to effects in *lsm6Δ* and *lsm7Δ* strains. A dramatic increase in the level of U4 snRNA in the *lsm5Δ* strain suggests Lsm5p may have a task in degradation of this small stable RNA. The involvement of the different domains of Lsm1p and Lsm8p in localisation of these proteins to the cytoplasm (to P-bodies under stress conditions) and nucleus respectively was investigated by creating (deletion) mutants and hybrids of various domains of these proteins. Results suggest that the N-termini of both proteins play a central role in targeting them to their respective cellular locations. The *in vitro* studies reveal that the RNA chaperone activities of Sm-like proteins appear to have been conserved from bacteria to eukaryotes. Presumably, gene duplication and formation of hetero-multimeric complexes in higher organisms has allowed functional diversification. The *in vivo* results begin to distinguish functions specific to individual Lsm polypeptides and protein domains, however, undoubtedly others remain to be found.

Acknowledgements *1*

Abstract *2*

Abbreviations *10*
Amino acid abbreviations 13
The genetic code 13

CHAPTER 1 Introduction **14**

1.1 RNA, a versatile biomolecule *15*
1.1.1 Messenger RNA 15
1.1.2 Non-coding RNA 16
1.1.2.1 Transfer RNA 16
1.1.2.2 Ribosomal RNA 16
1.1.2.3 Small nucle(ol)ar RNA 17
1.1.2.4 Catalytic RNA 17
1.1.2.5 Regulatory RNA 18
1.1.2.6 Other non-coding RNAs 18
1.1.3 RNA processing and ribonucleoprotein particles 19

1.2 Pre-mRNA splicing and the spliceosome *19*
1.2.1 The splicing reaction 20
1.2.2 Spliceosome assembly 21

1.3 The spliceosomal snRNAs *23*
1.3.1 Biogenesis of non-U6 spliceosomal snRNPs 23
1.3.2 U6 snRNA 25
1.3.3 snRNP assembly and the SMN complex 26
1.3.4 U6 snRNA recycling 28

1.4 The U6 snRNP proteins *29*
1.4.1 Lsm proteins form the core of the U6 snRNP 29
1.4.2 Lhp1p is a U6 snRNP chaperone 31
1.4.3 Prp24p promotes U4/U6 annealing 31
1.4.4 Interactions between Prp24p and Lsm2-8p 32

1.5 Additional nuclear functions of the Lsm proteins *33*
1.5.1 Lsm proteins in tRNA, rRNA and snoRNA processing 33
1.5.2 Lsm proteins in nuclear pre-mRNA turnover 34

1.6 Lsm proteins in cytoplasmic mRNA turnover *35*
1.6.1 Two general mRNA degradation pathways 35
1.6.2 mRNA decapping 36
1.6.3 The cytoplasmic Lsm1-7 complex functions in mRNA decapping 36
1.6.4 Lsm1-7p protects shortened mRNA 3' ends 37
1.6.5 Cytoplasmic 5' to 3' mRNA decay occurs in P-bodies 40

1.7 Two (or more) complexes, many questions *41*

1.8 Structural and functional studies of Sm(-like) complexes *42*
1.8.1 The Sm heteroheptameric ring 42
1.8.2 The Lsm heptamers 44
1.8.3 Archaeal homomultimeric rings 46
1.8.4 Hfq, a bacterial Lsm complex with RNA chaperone activity 48

1.8.5 (L)Sm RNA binding activity 50
1.8.6 Alternative (L)Sm complexes 51
1.8.6.1 U7 snRNP 51
1.8.6.2 snRNP-specific Sm complexes in trypanosomes 51
1.8.6.3 Telomerase 52
1.8.6.4 Hexameric Lsm complexes? 52
1.8.6.5 Other *S. cerevisiae* Sm-like proteins 52
1.8.7 Evolutionary considerations 54

1.9 Lsm proteins in disease *56*
1.9.1 Systemic lupus erythematosus 56
1.9.2 Lsm proteins and spinal muscular atrophy 56
1.9.3 Lsm proteins and cancer 57
1.9.4 Other links between Sm-like proteins and disease 58
1.9.4.1 Hfq and bacterial virulence 58
1.9.4.2 Lsm1-7p in viral replication and translation 58

1.10 Aims of this thesis *59*

CHAPTER 2 Materials and Methods 61

2.1 General reagents *62*
2.1.1 Sources of reagents 62
2.1.1.1 Chemicals 62
2.1.1.2 Enzymes 62
2.1.1.3 Growth reagents 62
2.1.1.4 Antibiotics 62
2.1.2 Bacterial and yeast growth media 62
2.1.2.1 General information 62
2.1.2.2 Bacterial media 63
2.1.2.3 Yeast media 63
2.1.2.4 Drop-out powder 64
2.1.2.5 Antibiotics 64
2.1.3 Commonly used buffers 65
2.1.4 Sterilising solutions and making solutions RNase free 66
2.1.5 *Escherichia coli* strains 67
2.1.6 *Saccharomyces cerevisiae* strains 68
2.1.7 Plasmids 72
2.1.8 Oligonucleotides 85
2.1.9 Antisera 100

2.2 Microbiological methods *101*
2.2.1 Growth of strains 101
2.2.1.1 Growth of bacteria 101
2.2.1.2 Growth of yeast 101
2.2.2 Preservation of strains 102
2.2.2.1 Preservation of bacteria 102
2.2.2.2 Preservation of *S. cerevisiae* 102
2.2.3 Transformation of *E. coli* 102
2.2.3.1 Preparation of chemically-competent cells for cloning 102
2.2.3.2 Transformation of chemically-competent cells 103
2.2.3.3 Preparation of electro-competent cells 103
2.2.3.4 Transformation of electro-competent cells 104
2.2.3.5 Preparation of calcium-competent cells for transformation with plasmid DNA 104
2.2.3.6 Transformation of calcium-competent cells 105
2.2.4 Creation of *E. coli* genomic deletion 105

2.2.5 Transformation of *S. cerevisiae* 106
2.2.5.1 Yeast transformation with plasmid DNA 106
2.2.5.2. Yeast transformation for genomic integration of DNA 107
2.2.6 Yeast sporulation and tetrad dissection 108
2.2.8 Mating type determination 109
2.2.9 Growth curves 109
2.2.10 Spot assays 110
2.2.11 Plasmid shuffle on 5-FOA 110

2.3 Nucleic acid methods *110*
2.3.1 General methods 110
2.3.1.1 Spectrophotometric quantification of nucleic acids 110
2.3.1.2 Phenol/chloroform extraction 111
2.3.1.3 Precipitation of nucleic acids 111
2.3.2 DNA methods 111
2.3.2.1 Small scale preparation of plasmid DNA 111
2.3.2.2 Large scale preparation of plasmid DNA 112
2.3.2.3 Plasmid rescue from yeast 112
2.3.2.4 DNA restriction digests 112
2.3.2.5 Removal of phosphates from DNA ends 113
2.3.2.6 Yeast genomic DNA preparation for PCR 113
2.3.2.7 Amplification of DNA by Polymerase Chain Reaction 114
2.3.2.8 Colony PCR 115
2.3.2.9 Site directed mutagenesis 115
2.3.2.10 Primer design 117
2.3.2.11 Purification of PCR products and restriction digests 118
2.3.2.12 Agarose gel electrophoresis 118
2.3.2.13 Purification of DNA from agarose gels 119
2.3.2.14 Ligation of DNA molecules 119
2.3.2.15 De-salting small amounts of DNA 119
2.3.2.15 DNA sequencing 120
2.3.2.16 End-labelling of oligonucleotides 120
2.3.2.17 Making DNA probes for Northern analysis 121
2.3.2.18 Radioactive labelling of DNA size marker 122
2.3.2.19 Purification of labelled DNA from unincorporated nucleotides 122
2.3.3 RNA methods 122
2.3.3.1 Three methods for total RNA isolation from yeast 122
2.3.3.2 *In vitro* transcription and RNA purification 124
2.3.3.3 Denaturing PAGE of RNA 126
2.3.3.4 Native PAGE of RNA 127
2.3.3.5 Electrophoresis of glyoxylated RNA through agarose gels 127
2.3.3.6 Northern blotting 128
2.3.3.7 Northern blot hybridisation 129
2.3.3.7 Denaturing northern blot analysis of snRNAs 130
2.3.3.8 Primer extension 130

2.4 Protein methods *131*
2.4.1 Crude protein preparation from yeast 131
2.4.2 Preparation of yeast whole cell extract 132
2.4.2.1 Preparation of splicing extracts 132
2.4.2.2 Preparation of crude extracts for immunoprecipiation 133
2.4.3 SDS Polyacrylamide gel electrophoresis (SDS-PAGE) 133
2.4.4 Coomassie staining of protein gels 134
2.4.4.1 Standard Coomassie staining protocol 134
2.4.4.2 GelCode Blue staining for mass spectrometry 134
2.4.5 Western blotting 135
2.4.5.1 Transfer of proteins to PVDF membrane 135
2.4.5.2 Western blot hybridisation 135
2.4.5.3 Enhanced Chemiluminescence (ECL) 135

2.4.5.4 Stripping Western blots 136
2.4.6 Immunoprecipitation of proteins 136
2.4.7 Recombinant protein expression in *E. coli* 137
2.4.7.1 Recombinant protein induction with IPTG 137
2.4.7.2 Recombinant protein expression using ONEX 137
2.4.7.3 Harvesting cultures after recombinant protein expression 137
2.4.7.4 Expression of Lsm proteins and Hfq 138
2.4.7.5 Expression of Prp24p 138
2.4.8 His-tagged protein purification 138
2.4.8.1 Cell lysis for protein purification 138
2.4.8.2 His-tagged protein purification under native conditions 139
2.4.8.3 His-tagged protein purification under denaturing conditions 140
2.4.9 Protein quantification 140
2.4.10 *In vitro* assembly of Lsm complexes 141
2.4.10.1 Assembly of Lsm complexes in 8M urea 141
2.4.10.2 Assembly of Lsm complexes in 4M urea 141
2.4.10.3 Optimisation of yeast Lsm complex assembly 142
2.4.11 Size exclusion chromatography of protein complexes 142
2.4.12 Mass spectrometry of proteins 143

2.5 Mixed protein/RNA assays *143*
2.5.1 RNA bandshift assay 143
2.5.2 U4/U6 RNA annealing assay 144
2.5.3 Strand displacement assay 145
2.5.3.1 Strand displacement substrate preparation 145
2.5.3.2 Strand displacement reactions 145

2.6 Microscopy methods *146*
2.6.1 Live cell imaging 146
2.6.2 Imaging of fixed and DAPI stained cells 146

2.7 Bioinformatics *147*

CHAPTER 3 Lsm protein expression, purification and in vitro complex formation 149

3.1 Introduction *150*

3.2 Creating protein expression vectors *150*
3.2.1 Cloning yeast *LSM* genes into *E. coli* expression vectors 150
3.2.2 Confirmation of a change in *LSM8* ORF designation 151

3.3 Optimisation of expression and purification of individual Lsm proteins *154*

3.4 Co-expression of Lsm proteins improves solubility *156*
3.4.1 Co-expressed pairs of Lsm proteins can be co-purified 159
3.4.2 Co-expression and purification of Lsm1-7p and Lsm2-8p 160

3.5 Lsm complex formation in vitro *162*
3.5.1 Re-naturation of a mixture of denatured Lsm proteins allows for complex assembly *in vitro* 162
3.5.2 Soluble complexes can form in the absence of single subunits 164
3.5.3 Lsm complexes aggregate and/or form higher order structures 164
3.5.4 *In vitro* assembly of human LSm complexes 168
3.5.5 Optimisation of *in vitro* Lsm complex assembly 168

3.6 Contamination of recombinant Lsm proteins with E. coli Sm-like protein Hfq *170*
3.6.1 Lsm8p co-purifies *E. coli* Hfq 170
3.6.2 Creation of an *hfq* deletion strain for Lsm protein production 171

3.7 Discussion *174*
3.7.1 Change to the *LSM7* annotation 174
3.7.2 Conclusions 174

CHAPTER 4 RNA binding and annealing activities of Sm-like Proteins 176

4.1 Introduction *177*

4.2 RNA binding by recombinant Sm-like proteins *177*
4.2.1 Individual yeast Lsm proteins and U6 snRNA binding 177
4.2.2 *In vitro* assembled yeast Lsm1-7p and Lsm2-8p bind U4 and U6 snRNA 178
4.2.3 Recombinant Hfq binds U4 and U6 snRNA 181
4.2.4 RNA binding by *in vitro* assembled human LSm complexes 183

4.3 RNA annealing activity of recombinant Sm-like proteins *186*
4.3.1 Recombinant yeast Lsm complexes do not display U4/U6 annealing activity 187
4.3.2 Recombinant Hfq promotes U4/U6 annealing 188
4.3.2.1 Hfq RNA binding is essential for its annealing activity 188
4.3.2.2 Hfq affects the conformation of the U4/U6 duplex 190
4.3.3 Recombinant human LSm complexes promote annealing of yeast U4 and U6 snRNA ... 192
4.3.3.1 hLSm1-7 and hLSm2-8 annealing activity is affected by the method of complex assembly 192
4.3.3.2 hLSm4 may be essential for annealing activity 192

4.4 Discussion *198*

CHAPTER 5 Investigating a possible mechanism for RNA annealing by Sm-like proteins 200

5.1 Introduction *201*

5.2 Effects of mutations that destabilise the U6 3' stem-loop *201*
5.2.1 U6 3' stem-loop destabilising point mutations do not rescue *lsm6Δ* and *lsm7Δ* growth phenotypes 201
5.2.2 Stem-loop destabilising mutations do not increase U4/U6 levels *in vivo* 205
5.2.3 Stem-loop destabilising mutations increase U4/U6 annealing efficiency *in vitro* 208

5.3 Can Lsm complexes change RNA structure? *209*
5.3.1 Designing a strand displacement assay 209
5.3.2 Proof of principle: Ded1p helicase unwinds the U6-like RNA/DNA duplex structure 210

5.4 Yeast Lsm complexes have strand displacement activity *212*

5.5 RNA-binding by recombinant Hfq is required for its strand displacement activity *214*

5.6 hLSm4 may be essential for strand displacement by the human LSm complexes *216*

5.7 Discussion *219*

CHAPTER 6 Interactions between Prp24p and Lsm proteins **222**

6.1 Introduction *223*

6.2 Investigating genetic interactions between PRP24, LSM6 and LSM7 ..... *224*

6.3 Investigating physical interactions between Prp24p and Lsm proteins in vitro *228*

6.4 Discussion *231*

CHAPTER 7 A functional analysis of Lsm4p **237**

7.1 Introduction *238*

7.2 Investigating LSM4 overexpression phenotypes *239*
7.2.1 Overproduced Lsm4p and its C-terminus accumulate in foci 239
7.2.2 Toxic effects of *LSM4* overexpression 240
7.2.3 Other Lsm proteins accumulate in UFO bodies upon Lsm4p overproduction 241
7.2.4 Lsm8p co-localises with Lsm4p, but not Lsm4C aggregates 244
7.2.5 Reduced U6 snRNA levels may be the cause of Lsm4 toxicity 246

7.3 Investigating a role for Lsm4p in Lsm1-7p recruitment to P-bodies *247*
7.3.1 UFO bodies are not P-bodies 247
7.3.2 The Lsm4p carboxy-terminal domain is not essential for localisation to P-bodies 247
7.3.3 Efficiency of Lsm protein recruitment to P-bodies is reduced in the absence of the Lsm4p C-terminus 250
7.3.4 The GFP-tag aggravates the *lsm4ΔC2* defect 252
7.3.5 The Lsm4p C-terminus may contribute to efficient P-body formation 255

7.4 A potential link between Lsm4p and cytokinesis? *255*
7.4.1 Depletion of Lsm4p leads to abnormal cell morphology 255
7.4.2 Increased Lsm4p production rescues a *dbf2-Ts* phenotype 256

7.5 Discussion *257*

CHAPTER 8 A functional analysis of Lsm5p **262**

8.1 Introduction *263*

8.2 lsm5 deletion strains are viable but show severely reduced growth rates *263*

8.3 The Arabidopsis sad1/lsm5 mutation does not show a growth defect in yeast *265*

8.4 Effects of lsm5Δ on spliceosomal snRNA levels, a comparison with lsm1Δ, lsm6Δ and lsm7Δ *267*
8.4.1 *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains shows similar effects on the levels of spliceosomal snRNAs 268
8.4.2 Levels of U6 and U4/U6 snRNA are decreased and U4 snRNA levels increased in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains 270
8.4.3 Changed levels of spliceosomal snRNAs in *lsm* deletion strains are not fully corrected by increased U6 expression 270

8.5 scR1 levels are increased in lsm1Δ *273*

8.6 Discussion *274*

CHAPTER 9 Dissection of Lsm1p and Lsm8p in sub-cellular localisation **279**

9.1 Introduction *280*

9.2 Creation and expression of Lsm1p and Lsm8p hybrids and deletion mutants *281*
9.2.1 Structure prediction of Lsm1p and Lsm8p 281
9.2.2 Hybrids and mutant proteins are stably expressed 284

9.3 The N- and C-terminal domains of Lsm1p and Lsm8p do not suffice as localisation signals *284*

9.4 No single domain of Lsm8p is essential for nuclear localisation *286*
9.4.1 Reduced RNA binding may affect Lsm8p nuclear localisation 287
9.4.2 Any of the Lsm8p domains can be deleted or replaced without complete loss of nuclear accumulation 289
9.4.3 The N-terminal ten amino acids of Lsm8p have a function in nuclear localisation 291

9.5 Lsm1p, Lsm1p mutants and Lsm1/8p hybrids localise to P-bodies *291*
9.5.1 No single domain of Lsm1p is essential for localisation to P-bodies 292
9.5.2 Lsm1p, Lsm1p mutants and Lsm1/8p hybrids co-localise with Dcp2-RFP 292

9.6 The Lsm1p and Lsm8p N-terminal domains are required for distinct cellular localisation *295*

9.7 A correlation between viability and correct localisation *297*

9.8 Levels of mutant and hybrid protein expression affect viability *301*

9.9 Effects of Lsm8 mutant and hybrid proteins on levels of spliceosomal snRNAs *303*

9.10 Discussion *306*

CHAPTER 10 Final discussion **310**

10.1 Sm-like proteins are RNA chaperones *311*

10.2 Lsm protein promiscuity *314*

10.3 Potential new functions of Lsm proteins *314*

10.4 Localisation of Lsm1-7p versus Lsm2-8p *317*

10.5 Final conclusions *318*

Appendices **319**

A.1 Effects of Hfq co-purification on RNA binding, annealing and strand displacement by yeast Lsm proteins *320*
A.1.1 Effects of Hfq co-purification on RNA binding of yeast Lsm proteins 320
A.1.2 Effects of Hfq co-purification on U4/U6 annealing activity of yeast Lsm proteins 324
A.1.3 Effects of Hfq co-purification on strand displacement activity of yeast Lsm proteins 326

A.2 Hfq interaction with Lsm8p *329*

A.3 Effects of hfq expression in yeast *331*

References ***334***

Abbreviations

A_{260}	absorption at 260 nm
aa	amino acid(s)
Amp	ampicillin
ARS	autonomous replication sequence
ATP	adenosine 5'-triphosphate
bp	base-pair
BPB	bromophenol blue
BLAST	basic local alignment search tool
BSA	bovine serum albumin
°C	degrees Celsius
cDNA	complementary DNA
CDS	coding sequence
CEN	centromere element
Ci	Curies
cm	centimetre
Cm	chloramphenicol
cpm	counts per minute
Cs	cold-sensitive
CTP	cytidine 5'-triphosphate
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
dNTP	deoxynucleoside 5'-triphosphate
dTTP	deoxythymidine 5'-triphosphate
DEPC	diethylpyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ds	double-stranded
DTT	dithiothreitol
ECL	enhanced chemi-luminescence
EDTA	ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
EtBr	ethidium bromide
Fig.	figure
5-FOA	5'-fluoroorotic acid
g	gram(s)
Gal	D-galactose
Glu	D-glucose
GTP	guanosine 5'-triphosphate
h	hour(s)
HA	haemaglutinin HA-1 protein epitope
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
His	histidine
Hph	hygromycin phosphotransferase (hygromycin resistance gene)

HRP	horse radish peroxidase
Hyg	hygromycin
IAA	iso-amyl alcohol
Ig	immunoglobulin
IPTG	isopropyl-β-D-thiogalactopyranoside
Kan	kanamycin
kb	kilobase
kDa	kilo Dalton
l	litre(s)
LB	Luria broth
M	molar
mA	milliampere(s)
MBq	megabequerelle
MCS	multiple cloning site
mg	milligram(s)
min	minute(s)
ml	millilitre(s)
mM	millimolar
MOPS	3-[N-morpholino]propanesulphonic acid
mRNA	messenger ribonucleic acid
MW	molecular weight
MWCO	molecular weight cut-off
Nat	nourseothrycin
NCBI	National Centre for Biotechnology Information
ncRNA	non-coding RNA
NES	nuclear exclusion or export signal
ng	nanogram(s)
NLS	nuclear localisation signal
nm	nanometre(s)
NP-40	nonidet P-40 detergent
nt	nucleotide
NTA	nitriloacetic acid
NTC	Prp19 complex
OD_{600}	optical density at 600 nm
ORF	open reading frame
p	plasmid designation
PAGE	polyacrylamide gel electrophoresis
P-bodies	(mRNA) processing bodies
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pI	isoelectric point
pmol	picomole(s)
PNK	T4 polynucleotide kinase
PVDF	polyvinylidine difluoride
RNA	ribonucleic acid
RNase	ribonuclease
RNP	ribonucleoprotein particle

rNTP	ribonucleoside 5'-triphosphate
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT	room temperature
s	seconds
SD-	synthetic defined *or* dropout medium minus
SDS	sodium dodecyl sulphate
SGD	*Saccharomyces* Genome Database
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein particle
sRNA	small (non-coding) RNA
SSC	sodium choride-sodium citrate (buffer)
ssDNA	single stranded DNA *or* salmon sperm DNA
SSPE	sodium chloride-sodium phosphate-EDTA (buffer)
Strep	streptomycin
SV40	Simian virus 40
TAE	Tris-acetate (buffer)
TBE	Tris-Borate-EDTA (buffer)
TBS	Tris-buffered saline
TBST	Tris-buffered saline-Tween-20
TE	Tris-EDTA (buffer)
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Tet	tetracycline
tRNA	transfer ribonucleic acid
Tris	Tris(hydroxymethyl)aminomethane
Ts	temperature-sensitive
µF	microfarad(s)
µg	microgram(s)
µl	microlitre(s)
µm	micrometre(s)
U	unit(s)
UFO bodies	unidentified Lsm4 bodies
Ura	uracyl
UTP	uridine triphosphate
UTR	untranslated region
UV	ultra violet
V	volt(s)
v/v	volume per unit volume
W	Watt(s)
w/o	without
WT or wt	wild type
w/v	weight per unit volume
XCB	xylene cyanol (blue) FF
YMM	yeast minimal medium
YPDA	yeast-peptone-dextrose-adenine (medium)
YPGalA	yeast-peptone-galactose-adenine (medium)

Amino acid abbreviations

Amino acid	Three letter abbreviation	Single letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Iso	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The genetic code

	U	C	A	G	
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

CHAPTER 1

Introduction

1.1 RNA, a versatile biomolecule

Ribonucleic acid (RNA) molecules can be both carriers of information and catalysts. Deoxyribonucleic acid (DNA) is an informational molecule with limited catalytic capabilities and polypeptides can carry out a wide variety of enzymatic functions, but do not usually transfer information. Both functions are needed to allow self-replication. As RNA molecules have the right characteristics, an early "RNA world" may have existed, in which self-replicating RNA molecules were the first life-forms. It is unclear whether such a world really existed and, if so, how it would have evolved into more complex life-forms that are based on DNA, RNA and proteins (Joyce and Orgel, 1999). It is clear however, that RNA is essential for life as we know it and central to many processes in the cell, not the least of which is the transfer of information encoded by DNA into polypeptide chains.

1.1.1 Messenger RNA

Messenger RNAs (mRNAs) contain open reading frames (ORFs) that encode proteins. Pre-mRNAs are transcribed by RNA polymerase II and are processed into mature translatable mRNAs. Processing takes place on both the 5' and 3' ends. The 5' end is capped by a guanosine nucleotide connected to the mRNA via an unusual 5' to 5' triphosphate linkage. Addition of a methyl group to N^7 of the guanine base results in m^7 GpppN or cap 0, which is the cap structure of mRNAs in yeast and plants. The cap regulates nuclear export, promotes 5' proximal intron removal, protects the mRNA from 5' to 3' degradation and is required for translation initiation. In eukaryotes a poly(A) tail is added to the 3' end by PAP1 after endonucleolytic cleavage at the polyadenylation site. This tail is 50-250 nucleotides long (depending on the organism), protects the mRNA from degradation, plays an essential role in translation initiation, and promotes export to the cytoplasm (Wahle and Keller, 1997). In contrast, in prokaryotes poly(A) tails usually target the mRNA for degradation. In eukaryotes, the mRNA coding sequence is often interrupted by non-coding "intronic" sequences that are removed by the spliceosome (see 1.2).

1.1.2 Non-coding RNA

In addition to protein-encoding mRNA a large number of so-called non-coding RNAs (ncRNAs) have a wide variety of functions in the cell.

1.1.2.1 Transfer RNA

Nucleotide triplets (codons) in the mRNA are recognised by complementary "anti-codons" in transfer RNA (tRNA) molecules. When "charged", each tRNA carries a specific amino acid that gets incorporated into a polypeptide chain by the ribosome. The tRNA is transcribed as a precursor sequence by RNA polymerase III with 5' and 3' extensions, called leader and trailer sequences respectively. The leader sequence is removed by endonucleolytic cleavage by the RNase P ribozyme. The trailer sequence is removed by endonucleolytic as well as exonucleolytic cleavages catalysed by a range of ribonucleases (RNases; Nakanishi and Nureki, 2005). In addition, introns need to be removed from tRNA in all three kingdoms of life. In bacteria, introns are self-splicing and are removed by a group I splicing mechanism. In other organisms removal of intronic sequences involves successive actions of a tRNA splicing endonuclease (yeast Sen2p) and ligase (yeast Rlg1p; Trotta and Abelson, 1999).

1.1.2.2 Ribosomal RNA

Ribosomal RNA (rRNA) molecules play both structural and catalytic roles in the ribosome, a multi megadalton ribonucleoprotein (RNP) complex that carries out translation. Each ribosome contains four different rRNA molecules and a large number (>75) of ribosomal proteins. Three of these rRNAs (5.8S, 18S and 25-28S) are transcribed by RNA polymerase I as a large polycystronic precursor. The fourth, 5S rRNA, is transcribed separately by RNA polymerase III. Processing of the pre-rRNA into mature rRNAs involves a large number of endo- and exonucleolytic cleavage steps, including those carried out by the 5' to 3' exonuclease Rat1p and the exosome, a 3' to 5' exonuclease complex. In addition, numerous sites undergo chemical modifications, mostly ribose 2'-O-methylation and pseudouridylation. Most of these processing steps take place in the nucleolus. Proteins that carry out the chemical modifications are guided to the correct sites by complementary small nucleolar RNAs (Venema and Tollervey, 1999; Granneman and Baserga, 2004).

1.1.2.3 Small nucle(ol)ar RNA

A large number of small nucleolar RNAs (snoRNAs) carry out a variety of functions. Some are involved in nucleolytic processing of rRNAs, but the majority play a role in ribose 2'-O-methylation and pseudouridylation, not only of rRNA but also of small nuclear RNAs (snRNA), tRNAs and possibly even mRNAs. There are two classes of snoRNAs each with specific structural features and functions: the box C/D snoRNAs function in 2'-O-methylation and the box H/ACA snoRNAs in pseudouridylation. These so-called guide snoRNAs contain sequences complementary to the target sequence. Although some snoRNAs are individually transcribed by RNA polymerase II or III, in metazoa the majority is encoded within intronic sequences of protein encoding genes, whereas in yeast and plants they are mostly synthesised in the form of polycystronic transcripts. Maturation of snoRNAs involves endonucleolytic and exonucleolytic processing, with roles for Rntlp endonuclease, Ratlp and the exosome (Tollervey and Kiss, 1997; Kiss, 2002; Filipowicz and Pogacic, 2002).

Yeast also has five small nuclear RNAs, U1, U2, U4, U5 and U6 snRNA that play a crucial role in pre-mRNA splicing (see 1.3). Other snRNAs exist, including U11, U12, $U4_{ATAC}$ and $U6_{ATAC}$ in higher eukaryotes that function in the minor U12-dependent spliceosome. These won't be discussed here, but for a recent review see Will and Lührmann (2005).

1.1.2.4 Catalytic RNA

Small RNAs with catalytic activity are called ribozymes. One of the most well-known examples is RNase P, a ribonucleoprotein complex required for the 5' processing of tRNA, in which the RNA component carries out the endonucleolytic cleavage. Other examples of catalytic RNAs include small ribozymes, like hammerhead, hairpin, hepatitis delta virus and *Neurospora* VS ribozymes that naturally function without proteins. Others are larger and can function *in vitro* in the absence of proteins, but function *in vivo* with the help of protein cofactors. These

include RNase P, and group I and II self-splicing introns (Jacquier, 1997; Altman and Kirsebom, 1999; Kazantsev and Pace, 2006).

1.1.2.5 Regulatory RNA

Over the past decade or so a vast array of small regulatory RNAs has been found in eukaryotes as well as prokaryotes. In higher eukaryotes these are micro RNAs (miRNAs) and small interfering RNAs (siRNAs). They can downregulate gene expression by promoting mRNA cleavage or by translational repression. They do this by annealing to complementary sequences often, but not always, in untranslated regions (UTRs) of the mRNA. Whether degradation or translational repression takes place usually depends on the extent of complementarity. miRNAs and siRNAs are derived from larger precursor RNAs by endo- and exonucleolytic processing and eventually get incorporated into an RNP complex called RISC, for RNA-induced silencing complex (Bartel, 2004). In bacteria regulatory ncRNAs are usually called sRNAs (for small RNAs). They can function in translational repression or stimulation, and RNA protection or degradation (Gottesman *et al.*, 2001; Storz *et al.*, 2004; Majdalani *et al.*, 2005). Their interaction with target mRNAs is often modulated by the Sm-like protein Hfq (Valentin-Hansen *et al.*, 2004). *Saccharomyces cerevisiae* does not have a recognised RISC complex and no known sRNA equivalents.

1.1.2.6 Other non-coding RNAs

Examples of other non-coding RNAs that fall outside the types of ncRNA described above include the RNA component of telomerase RNP (yeast *TLC1*) which is used as a template for reverse transcription to extend the end of the chromosomal DNA (Blackburn, 1999; Evans, 2001). Another is the RNA component (yeast scR1) of the signal recognition particle (SRP) RNP which is involved in targeting signal sequence containing proteins to the endoplasmic reticulum (Halic and Beckmann, 2005). Recently RNA was shown to associate with the mitotic spindle and play a direct translation-independent role in spindle assembly in *Xenopus* (Blower *et al.*, 2005). Other ncRNAs exist that are not discussed here, and undoubtedly other functions for ncRNAs remain to be discovered.

1.1.3 RNA processing and ribonucleoprotein particles

It will be clear from the above that RNA has a wide variety of functions. Many different RNAs are synthesised, with the overwhelming majority needing extensive processing before they can properly carry out their function. With very few exceptions, these RNAs are associated with proteins from the moment they are transcribed, throughout their biogenesis and lifetime as a mature functional RNA until the moment they are targeted for degradation. The protein composition of these ribonucleoprotein or RNP particles is often variable throughout their life cycle, but specific RNAs often (stably) associate with specific "core" proteins.

1.2 Pre-mRNA splicing and the spliceosome

Since the discovery in the late seventies that non-coding sequences can interrupt mRNA in eukaryotes (Berget *et al.*, 1977; Chow *et al.*, 1977a; Chow *et al.*, 1977b) much effort has been put into elucidating the mechanism of removal of these so-called introns. The two phosphoryl transfer steps of pre-mRNA splicing (Fig. 1.1) are catalysed by the spliceosome. In yeast, this large RNP complex consists of more than 70 proteins (Stevens *et al.*, 2002) and five small nuclear RNAs (U1, U2, U4/U6 and U5 snRNA). A multitude of dynamic RNA-RNA, RNA-protein and protein-protein interactions navigate the spliceosome through the different steps that ensure the precision that is needed to excise only those sequences that need to be removed to yield properly coding mRNA. It is widely accepted that the chemical steps that remove the intronic sequences and yield a mature mRNA molecule are in effect catalysed by the spliceosomal snRNAs (with a central role for U6 snRNA). The many associated proteins have tasks in assuring specificity and accuracy, in guiding and promoting changing interactions throughout the different stages involved in splicing, and ultimately enable these small RNAs to perform catalysis (Staley and Guthrie, 1998; Will and Lührmann, 2001).

1.2.1 The splicing reaction

In the first step of splicing a nucleophylic attack by the 2'-OH of the branchpoint adenosine on the 5' splice site phosphate cleaves the phosphodiester bond between exon-1 and the intron. This creates a free OH on the 3' end of exon-1, which in the second step attacks the phosphate of the 3' splice site. The products of these two transesterification reactions are the mature mRNA and the excised intron in the form of a lariate structure (Fig. 1.1). The lariate is subsequently debranched and degraded, whereas the mature mRNA is exported to the cytoplasm where it is used for translation by the ribosome.

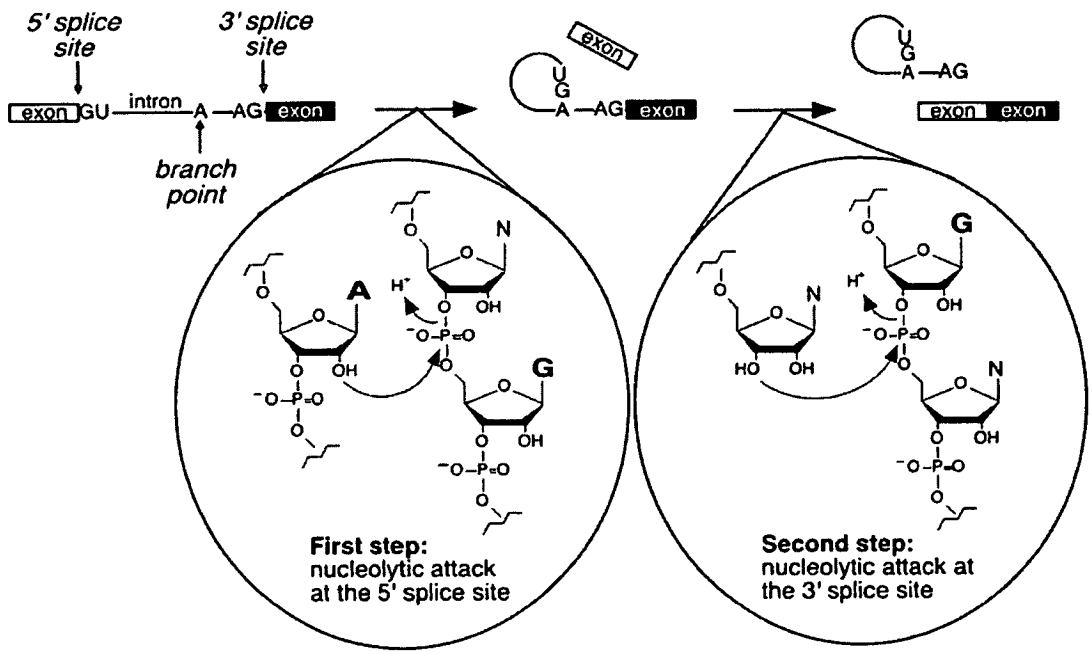


Fig. 1.1 The two chemical steps of splicing. Step 1: nucleophylic attack of the 2'-OH of the branchpoint adenosine on the 5' splice site phosphate. Step 2: the free OH on the 3' end of exon-1 attacks the phosphate of the 3' splice site. The end products are the mature mRNA and the excised intron. Reprinted from Jurica and Moore (2002)

1.2.2 Spliceosome assembly

There is evidence that at least in some cases, splicing occurs co-transcriptionally. Although some splicing factors have been shown to interact with the C-terminal domain (CTD) of RNA polymerase II, it is as yet unclear whether spliceosomal snRNPs interact with RNA Polymerase II or whether they are recruited to the newly synthesised RNA through an alternative mechanism (Maniatis and Reed, 2002; Bentley, 2005). Nonetheless, the moment the 5' splice site appears from the exit groove of the RNA polymerase the U1 snRNP associates with it. Similarly, the branchpoint sequence is recognised by U2AF (yeast Mud2p), which is involved in bringing together both parts of the pre-mRNA in the commitment complex. The U2 snRNP is recruited to the branchpoint and associates with U1. U4, U5 and U6 enter as a pre-assembled U4/U6-U5 tri-snRNP particle to form a complete spliceosome (Fig. 1.2). Soon after this U4 and U1 leave the complex or become less strongly associated, and new RNA-RNA interactions are formed between U2 and U6, between U6 and the 5' splice site, and between U5 and the 5' splice site. These conformational re-arrangements create the right conditions for splicing to take place, with U2 and U6 as the most likely candidates to form the catalytic centre and a role for U5 in aligning the two splice sites (Staley *et al.*, 1998; Collins and Guthrie, 2000; Will *et al.*, 2001; Villa *et al.*, 2002; Valadkhan, 2005).

A spliceosome complex containing all five snRNAs was affinity purified from yeast cell extracts, suggesting that a stable penta-snRNPs exists (Stevens *et al.*, 2002; Nilsen, 2002), but it seems that the normal state of affairs is that there is a number of ordered steps, as described above, in which the spliceosome assembles (see Fig. 1.2)(Brow, 2002). Recent studies examined spliceosome assembly *in vivo* using chromatin immunoprecipitation (ChIP). U1 and U2 snRNP were shown to bind co-transcriptionally to the nascent pre-mRNA, prior to the U4/U6-U5 tri-snRNP (Lacadie and Rosbash, 2005; Görnemann *et al.*, 2005; Tardiff and Rosbash, 2006). This is in support of step-wise recruitment of snRNPs to the pre-mRNA during transcription.

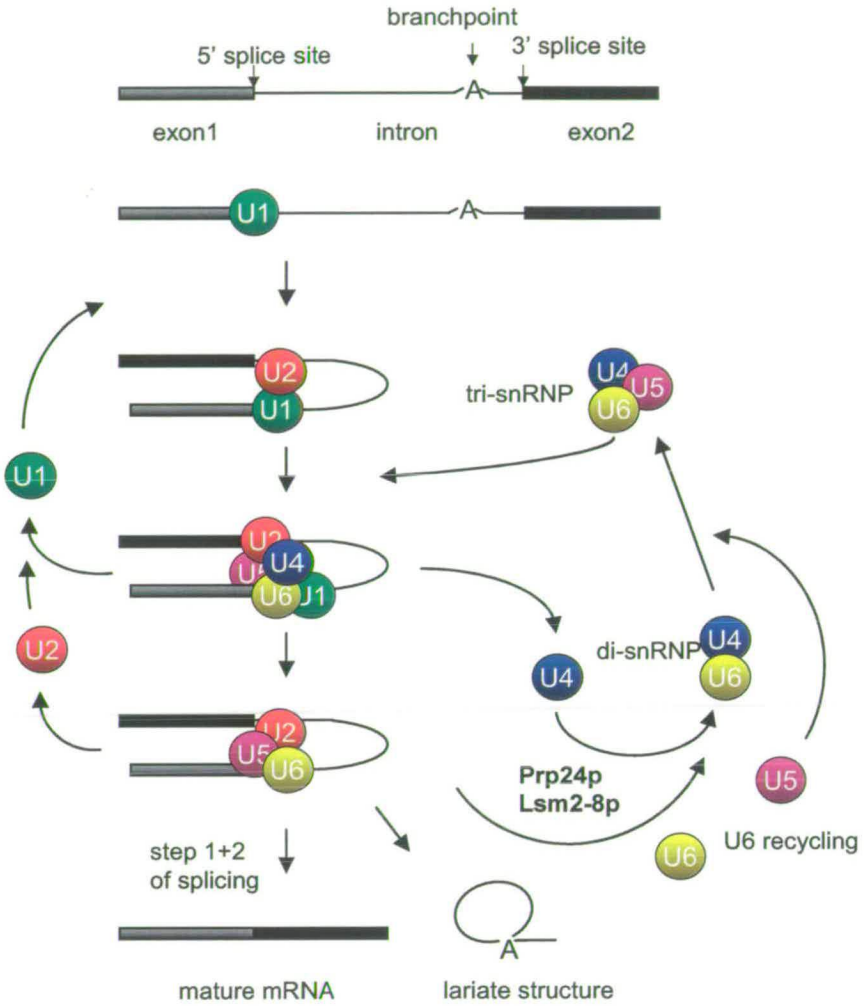


Fig. 1.2 Spliceosome assembly and U6 snRNP recycling. The U1 snRNP binds to the 5' splice site (E complex). The U2 snRNP binds to the branchpoint to form the commitment or A complex. Pre-formed U4/U6-U5 tri-snRNP associates with the A complex, to form the B complex. The spliceosome is activated after dissociation of U1 and U4 forming the catalytically active B* complex. After splicing has occured, the mature mRNA is released and spliceosome components are recycled.

1.3 The spliceosomal snRNAs

1.3.1 Biogenesis of non-U6 spliceosomal snRNPs

The spliceosomal small nuclear RNAs (snRNAs), also termed U (for uridine-rich) snRNAs, have been shown to play a central role in aligning splice sites and are considered the most likely candidates to form the actual catalytic core. Following transcription by RNA polymerase II and export from the nucleus to the cytoplasm U1, U2, U4 and U5 snRNAs stably associate with seven Sm proteins (SmB/B', D1, D2, D3, E, F and G) that form a heteroheptameric core complex on each of these snRNAs. Sm proteins were named after Stephanie Smith, a patient suffering from systemic lupus erythematosus in whom the first autoantibodies against the Sm-antigen, the first known non-histone antigen causing this disease, were found (Tan and Kunkel, 1966; Tan and Kunkel, 2006). These antibodies were later used to identify the Sm antigen as small nuclear ribonucleoprotein particles (snRNPs), consisting of snRNAs and the Sm core proteins (Lerner and Steitz, 1979). Formation of the Sm core RNP is essential for snRNA metabolic stability, hypermethylation of the 5' cap, 3' end processing, import of the particle into the nucleus, and recruitment of at least some particle specific proteins (Will *et al.*, 2001). It has also been suggested that the Sm complex has a function in aiding weak interactions, e.g. those between the U1 snRNP and the 5' splice site, in which the charged C-terminal tails of SmB, D1 and D3 would play a central role (Zhang *et al.*, 2001).

Following transcription by RNA polymerase II the snRNAs acquire a monomethylated m^7-GpppG cap and are exported to the cytoplasm in a Crm1p/RanGTP dependent fashion. U snRNA export requires additional factors including the cap-binding complex (CBC) and PHAX (phosphorylated adaptor for RNA export). In the cytoplasm the Sm proteins assemble on the conserved Sm site (consensus PuAU4-6GPu with two flanking stem-loop structures). Initially the SmD1/D2 and SmE/F/G subcomplexes interact with this sequence and the doughnut-shaped ring is completed by binding of SmB/D3. snRNP assembly and maturation require the SMN complex, which is discussed in more detail in section 1.3.3. The cap is tri-methylated and the 3' end is processed, which in yeast involves RNase III and

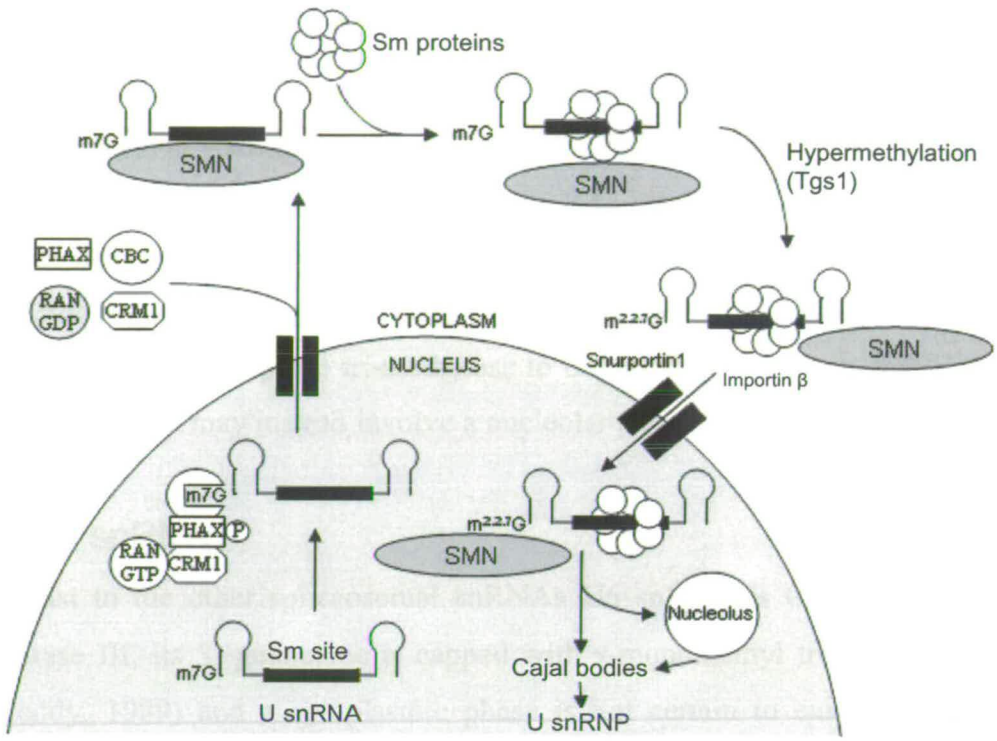


Fig. 1.3 U snRNP assembly in higher eukaryotes. Newly transcribed U snRNA acquires a methyl guanosine (m7G) cap in the nucleus, to which the cap binding complex (CBC) binds. This then associates with phosphorylated PHAX which in turn binds to CRM1/Xpo1 which associates with Ran-GTP. This complex is exported to the cytoplasm where the complex dissociates after PHAX desphosphorylation and GTP hydrolysis by Ran. Assisted by SMN, the Sm proteins assemble around the Sm site of the U snRNA, which leads to the hypermethylation of the m7G cap, forming an $m^{2,2,7}G$ cap. This enables snRNP reimport into the nucleus via Importin-β and Snurportin1. Back in the nucleus the U snRNP associates with snRNP-specific proteins, probably in Cajal bodies, forming a mature U snRNP particle.

U6 is the most highly conserved of the snRNAs (Brow and Guthrie, 1988), and certain conserved residues in U6 are thought to contribute to the catalytic centre of the spliceosome (reviewed by Valadkhan, 2005). Support for a ribozyme type role for U6 snRNA in the splicing reaction comes from experiments showing protein-free catalysis of a reaction resembling the first step of splicing by U2 and U6 fragments (Valadkhan and Manley, 2001; Valadkhan and Manley, 2003). U6 snRNA exists both in a free U6 snRNP and base-paired with U4 in a U4/U6 di-snRNP and U4/U6-U5 tri-snRNP. In each of these snRNPs U6 RNA is associated with Lsm2-8p via its uridine-rich 3' end (Achsel *et al.*, 1999; Mayes *et al.*, 1999; Gottschalk *et al.*, 1999; Stevens and Abelson, 1999), but additional interactions with the flanking stem-loop structure may exist (Vidal *et al.*, 1999). U4 snRNA has been proposed to have a chaperone-like role in masking catalytically important residues in U6 snRNA until they are required for catalysis in the spliceosome (Nilsen, 1994). The U4/U6 di-snRNP must associate with the U5 snRNP to form a U4/U6-U5 tri-snRNP for assembly into spliceosomes, after which U6 forms basepairing interactions with U2 and part of the 5' splice site sequence, and U4 leaves the spliceosome. The U6 3' stem-loop structure plays a central role in regulation of U6 interactions with other RNAs. Involvement of the same nucleotides in the formation of this intramolecular structure and basepairing to U4, and on the other hand the ability of U6 to basepair with the 5' splice site in the presence of the 3' stem-loop indicate that a delicate balance exists between intra- and intermolecular RNA-RNA interactions (Brow and Vidaver, 1995; Ryan and Abelson, 2002). For an overview of some RNA-RNA interactions involving U6 snRNA see Fig. 1.4. Thus the U6 snRNA is involved in many dynamic interactions during the assembly and activation of spliceosomes, and recycling of spliceosome components.

1.3.3 snRNP assembly and the SMN complex

In higher eukaryotes RNP assembly is aided by the survival of motor neuron (SMN) complex (reviewed by Yong *et al.*, 2004). The SMN complex locates both to the cytoplasm and to the nucleus, where it concentrates in nuclear bodies termed gems. It consists of the SMN protein and the Gemin2-7 proteins. Interestingly, Gemin6 and Gemin 7 form a heterodimer with an Sm-like structure, in spite of limited sequence

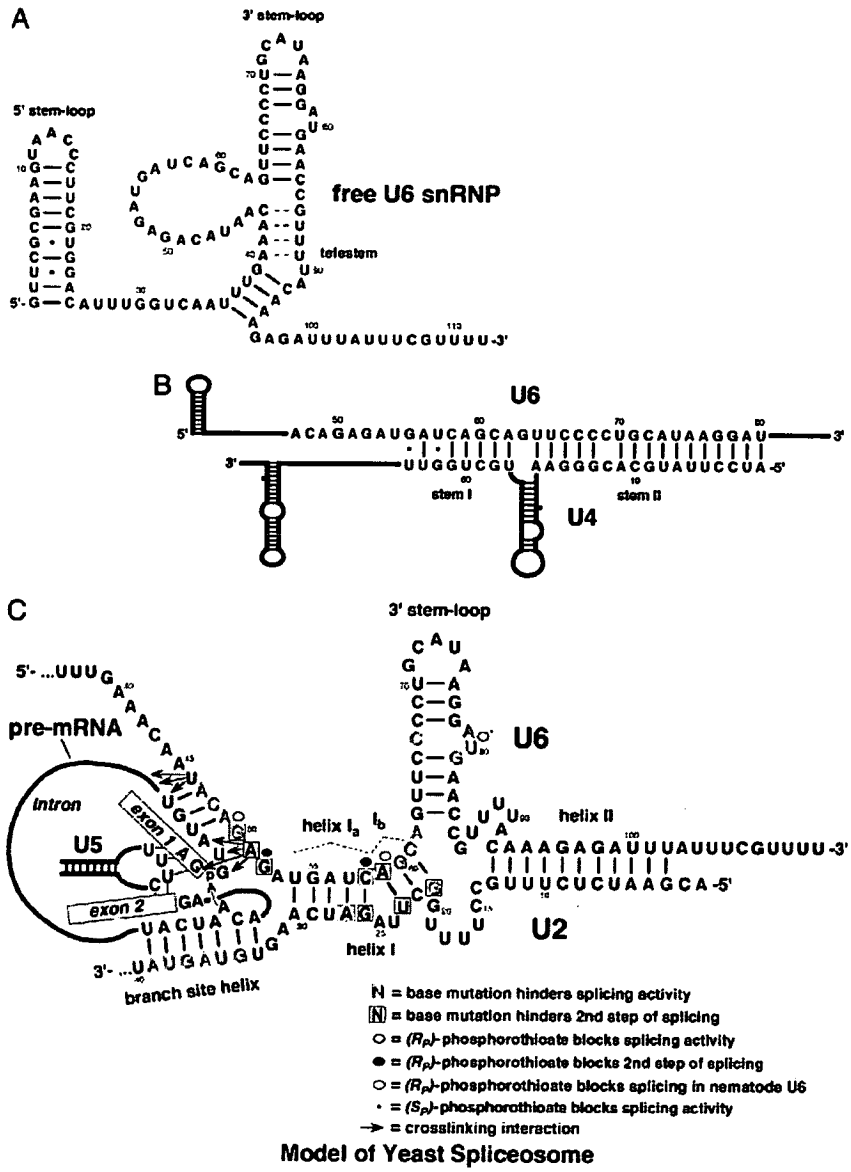


Fig. 1.4 **(A)** Proposed secondary structure of U6 snRNA in *S. cerevisiae.* Recent evidence suggests that nt 40-43 are present in single stranded form (Karaduman *et al.*, 2006). **(B)** Duplex interaction between yeast U4 and U6 snRNAs in U4/U6 snRNP complexes. **(C)** Working model for U2 and U6 snRNA interactions in fully assembled spliceosomes. Mutationally sensitive nucleotides and essential phosphoryl oxygens are highlighted. Sites of previously crosslinked interactions between U6 snRNA and the 5' splice site region are indicated. Reprinted from Ryan and Abelson (2002).

homology, and may bind and organise the (L)Sm proteins[1] in preparation for snRNP assembly (Ma *et al.*, 2005). SMN interacts with the snRNP core proteins SmB, -D1 and -D3, as well as Lsm4 via unique carboxy-terminal arginine- and glycine-rich domains. This interaction is enhanced greatly by a posttranslational, symmetric dimethylarginine modification that occurs at specific arginines in these RG domains, a reaction carried out by the 20S methylosome that contains an arginine methyltransferase called PRMT5 (yeast Hsl7p). The snRNAs are recognised by SMN through specific stem-loop structures. The main function of the SMN complex seems to be to increase efficiency and specificity of snRNP assembly, by preventing incorrect protein-protein and protein-RNA interactions. These would be likely to occur in the absence of a snRNP chaperone, because (L)Sm proteins by themselves can form stable heptameric rings with minimal RNA sequence specificity. Such promiscuous interactions could be dangerous to the cell. This may not be restricted to snRNPs, as SMN has also been implicated in formation of other types of RNPs, including snoRNPs and micro RNPs (Yong *et al.*, 2004).

In light of this, it is surprising that yeast does not seem to have an SMN homologue and is therefore presumed not to have an SMN-type complex. Interestingly however, Gemin2 (SIP1) has a yeast homologue called *BRR1* (Liu *et al.*, 1997). This protein interacts with a number of Sm proteins (Krogan *et al.*, 2006) and its deletion seems to affect snRNP assembly, as newly transcribed snRNAs are destabilised, whereas mature particles are not affected (Noble and Guthrie, 1996a; Noble and Guthrie, 1996b). Brr1p may have a function in snRNP biogenesis in yeast equivalent to that of the SMN-complex in higher eukaryotes. Lhp1p (yeast La protein) has also been implicated in U snRNP assembly as it interacts with and stabilises 3' extended pre-U snRNAs, thus facilitating binding by Sm proteins (Xue *et al.*, 2000).

1.3.4 U6 snRNA recycling

A yeast penta-snRNP particle was isolated which was suggested to go through several rounds of splicing without any of the snRNPs leaving the spliceosome (Stevens *et al.*, 2002). The biological significance of this penta-snRNP particle is

[1] (L)Sm or Sm(-like) is used when referring to both Sm and Lsm proteins

unclear and it is generally assumed that the spliceosome is dismantled after completion of the splicing reaction. The U4/U6-U5 tri-snRNP particle is re-formed and takes part in a new round of splicing. Thus, formation of the U4/U6 di-snRNP and the U4/U6-U5 tri-snRNP are important steps in both initial spliceosome assembly and recycling of spliceosome components. The underlying mechanism for this has not yet been fully clarified. Proteins that have been shown to be involved in U4/U6 duplex formation are Prp24p (Ghetti *et al.*, 1995; Raghunathan and Guthrie, 1998a), its human orthologue p110nrb/SART3 (Bell *et al.*, 2002; Medenbach *et al.*, 2004), and the Lsm2-8 complex (Achsel *et al.*, 1999; Verdone *et al.*, 2004). Verdone *et al.* (2004) proposed specific involvement of Lsm6p and Lsm7p in di-snRNP and tri-snRNP assembly respectively. Prp24p (Ghetti *et al.*, 1995; Vidaver *et al.*, 1999) as well as the U5 snRNP associated RNA unwindase Brr2p (Raghunathan and Guthrie, 1998b; van Nues and Beggs, 2001) have also been implicated in subsequent rearrangements of the U6 snRNA interactions with U4 snRNA, U2 snRNA and the pre-mRNA 5' splice site. In higher eukaryotes U4 and U6 snRNA and p110 are enriched in nuclear foci, called Cajal bodies (Stanek *et al.*, 2003; Stanek and Neugebauer, 2004). This localised increase in concentration contributes to U4/U6 di-snRNP assembly (Klingauf *et al.*, 2006). In humans, U4/U6-U5 tri-snRNP formation also takes place in Cajal bodies and requires Prp31 and Prp6 (Schaffert *et al.*, 2004).

1.4 The U6 snRNP proteins

1.4.1 Lsm proteins form the core of the U6 snRNP

U6 snRNA interacts with seven Lsm (Like Sm) proteins that are structurally related to the highly conserved Sm proteins (reviewed by Beggs, 2005). Eight Lsm proteins were initially identified in *S. cerevisiae*, but only seven of these, Lsm2p to Lsm8p (Lsm2-8p) associate with U6 (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999; Beggs, 2005), whereas a second Lsm complex consisting of the Lsm1 to Lsm7 proteins (Lsm1-7p) plays a role in mRNA turnover in the cytoplasm (He and Parker, 2000; Tharun *et al.*, 2000; Bouveret *et al.*, 2000)(see section 1.6 and Fig. 1.5/1.6). Most of these proteins are essential for viability, but deletion strains of *LSM1, LSM6* and *LSM7* are temperature-sensitive, whereas viability of *lsm5Δ* seems to depend on

the strain background (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999; Pannone *et al.*, 2001). Decreased levels of U6 snRNA in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* and after switch off of P_{GAL}-regulated *LSM2, 3, 4, 5* and *8*, but not *LSM1* showed the importance of the Lsm2 to 8 proteins for U6 stability (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999). Also, overexpression of U6 snRNA makes these proteins dispensable for growth (Mayes *et al.*, 1999). Co-immunoprecipitation experiments showed stable association of the Lsm2-8 proteins with U6 snRNA in free U6 snRNPs, as well as in U4/U6 di-snRNPs and U4/U6.U5 tri-snRNPs (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999). Lsm2-8p leave U6 snRNA after U4/U6 unwinding and Prp19p-complex (NTC) binding to the spliceosome, which coincides with a change in basepairing interactions between U6 and the 5' splice site. It was suggested that NTC-dependent Lsm2-8p dissociation may be required for this change in RNA-RNA interactions (Chan *et al.*, 2003).

The human LSm2-8 complex (hLSm2-8[2]) was purified and shown to be stable in the absence of RNA. Electron micrographs of these complexes revealed a doughnut shaped ring with a diameter of about 80Å and a 20Å central stain, similar to that of the Sm heptamer. The 3'-terminal residues (uuCCAUAUUUUu) of human U6 snRNA were shown to be essential and sufficient for LSm2-8 binding, but dissociation of the RNA oligonucleotide at relatively low salt concentrations suggests that additional contacts may be needed to stabilise the interaction (Achsel *et al.*, 1999). Interactions between the yeast Lsm proteins and U6 snRNA were investigated by photo-crosslinking and immunoprecipitation experiments. Lsm4p, Lsm6p and Lsm7p showed strong crosslinking to U6 RNA precipitated by Lsm4p, Prp24p or Prp8p antibodies indicating that the Lsm proteins are associated with U6 in the single, di- and tri-snRNP particles. Others showed no (Lsm3p and Lsm5p) or only weak (Lsm2p and Lsm8p) crosslinking. In contrast to the results for human U6 snRNA, the 3' end of yeast U6 was shown to be essential for Lsm2-8p binding, but not sufficient by itself (Vidal *et al.*, 1999). A 3' truncated U6 (U6Δ3') still gets incorporated into U4/U6 di-snRNP and U4/U6.U5 tri-snRNP particles (Fabrizio *et al.*, 1989) and *in vitro* splicing still occurs when wildtype U6 is replaced by U6Δ3'

[2] In metazoa LSm is used, as opposed to Lsm for the yeast orthologues

(Fabrizio *et al.*, 1989; Ryan *et al.*, 2002). This suggests that interactions outside the single stranded U-rich track are sufficient for U6 snRNA incorporation into snRNPs *in vitro*, in spite of weakened binding to the Lsm2-8 complex. Binding of Lsm2-8p to the 3' stem-loop would have functional significance as the Lsm proteins are implicated in destabilisation of this structure which favours U4/U6 duplex formation. Indeed, *in vitro* assembly of the U4/U6 di-snRNP was promoted in the presence of human LSm2-8 affinity purified from HeLa cell extracts (Achsel *et al.*, 1999). However, whether U4 and U6 were part of a stable duplex within this particle was not confirmed.

1.4.2 Lhp1p is a U6 snRNP chaperone

Lhp1p is only transiently associated with the 3' terminal uridine residues of newly transcribed U6 snRNA. *LHP1* is non-essential, but becomes essential in strains with mutations in or deletion of *LSM* genes (*lsm8-1*, *lsm8-2*, *lsm5Δ*, *lsm6Δ* and *lsm7Δ*). This requirement disappears upon *SNR6* (the U6 gene) overexpression. In an *lhp1Δ* strain levels of U6-containing snRNPs are almost normal, but U6 RNA is sensitive to 3' trimming. Lhp1p has therefore been proposed to be a chaperone promoting U6 snRNP assembly by stabilising newly transcribed U6 and handing it on to Lsm2-8p (Pannone *et al.*, 1998; Pannone *et al.*, 2001).

1.4.3 Prp24p promotes U4/U6 annealing

The *S. cerevisiae* Prp24 splicing protein was initially identified as a suppressor of a cold-sensitive mutation in U4 snRNA (Shannon and Guthrie, 1991). The authors' suggestion that Prp24p may be stably associated with U6 snRNA and involved in U4/U6 association was subsequently proven by gel shift and *in vitro* U4/U6 annealing assays (Ghetti *et al.*, 1995). Prp24p contains 4 RNA recognition motifs (RRMs; Rader and Guthrie, 2002) and was shown to bind preferentially to a U4/U6 duplex *in vitro*, but under normal conditions only precipitates U6 snRNA *in vivo* (Shannon and Guthrie, 1991). It was therefore proposed that the U4/U6-Prp24p complex is only a transient intermediate in the annealing of U4 and U6 *in vivo* (Jandrositz and Guthrie, 1995). Regions of interaction were narrowed down by hydroxyradical footprinting using recombinant Prp24p bound to U6 RNA alone or

U4/U6. These data suggest that Prp24p preferentially binds an incomplete U4/U6 duplex (Ghetti *et al.*, 1995). The importance of Prp24p U4/U6 annealing activity *in vivo* was confirmed by Prp24p immunodepletion experiments in splicing extracts (Raghunathan *et al.*, 1998a). Prp24p depletion from these extracts lead to reduced U4/U6 di-snRNP formation and splicing efficiency. However, annealing activity of recombinant Prp24p was heavily reduced in deproteinised snRNPs, suggesting that additional proteins are involved in U4/U6 annealing. Attempts to show *in vitro* U4/U6 annealing activity of the human Prp24p homologue, p110, failed altogether (Bell *et al.*, 2002).

1.4.4 Interactions between Prp24p and Lsm2-8p

Lsm2-8p is likely to contribute to more efficient U4/U6 annealing. Purified human LSm2-8 was shown to promote U4/U6 di-snRNP formation (Achsel *et al.*, 1999), and the structure of the yeast native U6 snRNP, containing both Lsm2-8p and Prp24p, was recently shown to be more favourable for U4/U6 annealing than that of a U6-Prp24p complex (Karaduman *et al.*, 2006). Yeast two-hybrid interactions exist between the Lsm proteins and Prp24p (Fromont-Racine *et al.*, 2000; Rader *et al.*, 2002) and a conserved C-terminal motif in Prp24p, which is essential for this interaction, was identified (Rader *et al.*, 2002). Deletion of the 3' uridine-rich track of U6 snRNA not only affects Lsm binding, but also Prp24p association with U6, although no direct interactions between Prp24p and the U6 3' end exist. This suggests that Prp24p binding is enhanced by binding of Lsm2-8p to U6. In contrast, binding of Lsm2-8p does not appear to depend on U6 binding to Prp24p, as U6 mutations that affect Prp24p binding do not affect the U6-Lsm2-8p association (Ryan *et al.*, 2002). Both Rader and Guthrie (2002) and Ryan *et al.* (2002) suggested that Lsm2-8p initially binds U6 snRNA. Prp24p binding then becomes more favourable because of direct Prp24p-Lsm2-8p interactions and/or an Lsm2-8p induced change in U6 RNA structure that promotes the interaction with Prp24p. Prp24p could then recruit U4 snRNA and promote U4/U6 annealing. The proposed function of Prp24p in this mechanism is that of a tethering factor. Biochemical dissection of Prp24p showed that RRM1 and RRM2 are important for high affinity binding to U6 snRNA. RRM3 and RRM4 were suggested to promote U4/U6 association, possibly by

binding to U4 snRNA (Kwan and Brow, 2005). More structural studies of U6 snRNP complexes, like that of Karaduman *et al.* (2006) would be of great help in determining the potential U4/U6 annealing mechanism. A recent paper reported on the structure of the N-terminal RRM1 and RRM2 domains of Prp24p by NMR (Reiter *et al.*, 2006). More may follow. A potential mechanism for U4/U6 annealing involving both Prp24p and Lsm2-8p is discussed in detail in Chapter 6.

1.5 Additional nuclear functions of the Lsm proteins

1.5.1 Lsm proteins in tRNA, rRNA and snoRNA processing

Several of the Lsm proteins are also involved in a number of functions in the nucleus non-related to pre-mRNA splicing (Fig. 1.5). In budding yeast, depletion of Lsm2p to Lsm5p or Lsm8p leads to strong accumulation of unspliced, 5' and 3' unprocessed pre-tRNA species, whereas the absence of Lsm1p, Lsm6p or Lsm7p results in much weaker phenotypes. Co-precipitation of pre-tRNAs with TAP-Lsm3p and reduced association of Lhp1p with pre-tRNAs in Lsm depleted strains suggest a chaperone function for a nuclear Lsm complex facilitating interactions between pre-tRNA and other (processing) proteins. This may also be the case for other RNA polymerase III transcripts as Lhp1p binding to scR1 and pre-P RNA was strongly reduced in Lsm depleted strains (Kufel *et al.*, 2002). In addition, strains depleted of any of the same five essential nuclear Lsm proteins accumulated high levels of (aberrant) rRNA processing intermediates as well as mature rRNAs. This is consistent with a role for the Lsm proteins in promoting 5' and 3' exonuclease activity (Kufel *et al.*, 2003a). Depletion of these same Lsm proteins also affected U3 snoRNA processing, with accumulation of mature and truncated fragments of pre-U3. Binding of Lsm proteins and Lhp1p to pre-U3 was shown to be interdependent (Kufel *et al.*, 2003b). hLSm4 was also shown to interact with human pre-U3 snoRNA and hLSm8 was shown to be essential for U3 accumulation (Watkins *et al.*, 2004), suggesting a conserved role for LSm proteins in U3 snoRNP biogenesis. Involvement of Lsm proteins in these RNA processing pathways is consistent with their proposed role as RNA chaperones. As suggested, an additional function for the nuclear Lsm complex may be in stimulating degradation and/or processing by exonucleases. As polyadenylation of the TRAMP

complex was recently shown to stimulate degradation of highly structured RNAs by the exosome (LaCava *et al.*, 2005; Vanacova *et al.*, 2005; Wyers *et al.*, 2005; Houseley *et al.*, 2006) the Lsm proteins may stimulate the TRAMP complex and/or exosome. Multiple yeast two-hybrid interactions between these complexes exist (Fromont-Racine *et al.*, 2000). Intruigingly, the *Escherichia coli* Sm-like protein Hfq stimulates polyadenylation of mRNA by PAP1 (Hajnsdorf and Régnier, 2000; Le Derout *et al.*, 2003; Folichon *et al.*, 2005) and promotes degradation by PNPase, and can even be found in a complex with these two enzymes (Mohanty *et al.*, 2004).

Nuclear Lsm proteins were also shown to interact with snoRNAs seemingly without being involved in their processing. In *Xenopus*, LSm2 to 8 proteins (excluding LSm5) were shown to bind specifically to U8 snoRNA, which is essential for maturation of the large ribosomal subunit RNAs, 5.8S and 28S. These LSm proteins bind to a conserved element within mature U8 snoRNA and were suggested to be involved in stabilisation of U8 and possibly in promoting its interaction with the rRNA (Tomasevic and Peculis, 2002). More recently it was suggested that an Lsm2-7p complex exists in *S. cerevisiae* as these proteins but not Lsm8p were shown to associate with mature snR5 snoRNA This box H/ACA snoRNA functions in guiding site-specific pseudouridylation of rRNA (Fernandez *et al.*, 2004). It is uncertain what the function of the interaction with snR5 is and whether Lsm2-7p really exist as a separate complex, or whether Lsm8p failed to co-precipitate snR5 because its epitope is masked in the snR5 RNP.

1.5.2 Lsm proteins in nuclear pre-mRNA turnover

Analysis of nuclear mRNA degradation in *lsm* mutants revealed a role for Lsm2-8p in decapping and 5' to 3' degradation of nuclear (pre-)mRNAs. To distinguish between nuclear and cytoplasmic degradation, inducible *MET* and heat shock transcripts were analysed in a *nup145N* temperature-sensitive strain, that fails to export poly(A)+ RNA when incubated at nonpermissive temperature. 5'-Capped and polyadenylated transcripts accumulate in strains lacking any of the Lsm2-8p proteins, but not Lsm1p. Analysis of pre-mRNAs containing intronic snoRNAs show similar

accumulation of capped 5' fragments. In addition, Lsm8p interacts directly with the poly(A) tail, as the two were shown to crosslink. It was concluded that the nuclear Lsm2-8 complex normally targets nuclear RNA substrates for decapping, independent of deadenylation (Kufel *et al.*, 2004)

1.6 Lsm proteins in cytoplasmic mRNA turnover

1.6.1 Two general mRNA degradation pathways

Mature mRNA molecules are transported to the cytoplasm where they are generally used in several rounds of translation. The lifetime of different mRNAs can vary widely and the speed in which they are targeted for degradation is a potential regulatory point to influence resulting levels of protein. Two general decay pathways have been characterised in yeast by which mRNA is degraded in the cytoplasm (reviewed in Tucker and Parker, 2000; Wilusz *et al.*, 2001). Both pathways have the initial step in common: shortening of the poly(A) tail to 10-12 nucleotides by the Pan2p/Pan3p and Ccr4p/Pop2p containing deadenylase complexes. The major pathway consists of 5' to 3' degradation by exonuclease Xrn1p (also known as Kem1p or Ski1p), whereas 3' to 5' degradation by the exosome, a multiprotein complex of 3' to 5' exonucleases, provides an alternative mRNA turnover mechanism. The exosome consists of a large number of protein factors, which in yeast includes Rrp41p, Rrp46p, Mtr3p, Rrp42p, Rrp43p, Rrp45p, Csl4p, Rrp4p, Rrp40p, Rrp44p and Ski7p. Ski7p is only part of the cytoplasmic exosome, whereas Rrp6p and Rrp47p are unique to the nuclear exosome (Houseley *et al.*, 2006). Xrn1p will only act after the 5' cap structure has been removed by the Dcp1p decapping enzyme (Beelman *et al.*, 1996; LaGrandeur and Parker, 1998), a step which in turn is activated by deadenylation. In the alternative pathway, that is independent of decapping, the exonucleolytic activity of the cytoplasmic form of the exosome is required for degradation following the removal of the remaining adenosine residues in a process called terminal deadenylation. Binding of the poly(A) tail by Pab1p (poly(A) binding protein) plays a vital role in controlling deadenylation. It seems to be a negative regulator of deadenylation in metazoa, but at the same time is required for optimal shortening of the poly(A) tail in yeast. Loss of Pab1p binding when the

poly(A) tail reaches a certain minimal length has been suggested to be central to the initiation of decapping. Circularisation of the mRNA transcript through interactions between the 3' (Pab1p) and 5' ends (eIF4E/G translation initiation factors) may prevent access of the decapping complex (Tucker and Parker, 2000; Wilusz *et al.*, 2001).

1.6.2 mRNA decapping

Regulation of decapping is complex and involves proteins with functions in translation initiation and deadenylation (Tucker *et al.*, 2000; Mitchell and Tollervey, 2001; Wilusz *et al.*, 2001). Competition for binding to the cap structure between the decapping complex and other cytoplasmic cap binding factors, including initiation factor eIF4E (Schwartz and Parker, 2000), and deadenylation complexes, plays a pivotal role in this regulation (Fig. 1.6). Decapping by Dcp1p is also influenced by the presence of certain sequences in the RNA. It was found that decapping was stimulated by the presence of AU-rich elements (AREs) and that this was dependent on ARE-binding proteins (Gao *et al.*, 2001). These sequence elements often occur in the 3' UTR and can promote as well as inhibit mRNA decay (reviewed by Wilusz *et al.*, 2001). In addition, a number of decapping cofactors, Dcp2p, Pat1p (Mtr1p), Edc1p, Edc2p, Edc3p (Dcp3p/Lsm16p) and Dhh1p (Hatfield *et al.*, 1996; Dunckley and Parker, 1999; Bonnerot *et al.*, 2000; Coller *et al.*, 2001; Dunckley *et al.*, 2001; Fischer and Weis, 2002), and potentially the phophorylation status of Dcp1p itself (LaGrandeur *et al.*, 1998) regulate its decapping activity. The Lsm1-7 complex was also shown to be required for efficient decapping (Tharun *et al.*, 2000; Bouveret *et al.*, 2000).

1.6.3 The cytoplasmic Lsm1-7 complex functions in mRNA decapping

Involvement of Lsm1p in mRNA decapping was initially discovered because an *spb8-2* (*lsm1*) mutation that allows growth of a *pab1Δ* strain, was shown to accumulate capped, poly(A)-deficient mRNAs (Boeck *et al.*, 1998). Later, a cytoplasmic Lsm heptamer, Lsm1-7p, in which Lsm1p interacts with Lsm2 to 7p, was shown to exist. The full Lsm1-7 complex was shown to be required for mRNA

decapping, as capped mRNAs accumulated in strains with *lsm* mutations. Direct interactions were shown to exist between Lsm1-7p, Pat1p and Xrn1p, whereas an interaction with Dcp1p was RNA dependent (Tharun *et al.*, 2000; Bouveret *et al.*, 2000). Pat1p is not only required for efficient decapping but also for translation initiation by promoting stable association of the 40S ribosomal subunit with the mRNP (Wyers *et al.*, 2000). Whereas mutations in eIF4E can suppress certain mutations in Dcp1p, they do not alleviate decapping defects in *lsm1Δ* or *pat1Δ* strains (Schwartz *et al.*, 2000). These data suggest that Pat1p may make the cap structure available to both the translation machinery and the decapping complex (He *et al.*, 2000). As Lsm1-7p mainly seems to interact with deadenylated mRNAs the Lsm proteins are more likely to act at a step distinct and most likely downstream from the competition between Dcp1p and eIF4E. Binding of Lsm1-7p may only occur after deadenylation, which may in turn stimulate Dcp1p binding and/or activity. Lsm1-7p may mediate this by changing the mRNP structure (Tharun and Parker, 2001).

1.6.4 Lsm1-7p protects shortened mRNA 3' ends

In addition to an important role in mRNA decapping the Lsm1-7p/Pat1p complex seems to protect mRNA 3' ends from degradation. *Pat1* and *lsm1-7* mutant strains were shown to accumulate mRNA transcripts that had been trimmed, i.e. shortened by 10-20 nucleotides at their 3' ends. Trimming was independent of decapping or exosome activity, but trimmed mRNAs did seem to get degraded by the exosome. Loss of mRNA 3' end protection and subsequent degradation by the exosome was proposed to cause temperature-sensitive (Ts) growth of *lms1-7* and *pat1* mutants, as Ts growth was suppressed by mutations in exosome components (He and Parker, 2001). Protection by Lsm1-7p of deadenylated 3' ends from degradation may or may not normally have a biologically significant function. It does however suggest that Lsm1-7p interacts with elements in the 3' UTR after deadenylation. This is consistent with a role for this complex in 5' to 3' degradation at a step in between deadenylation and decapping.

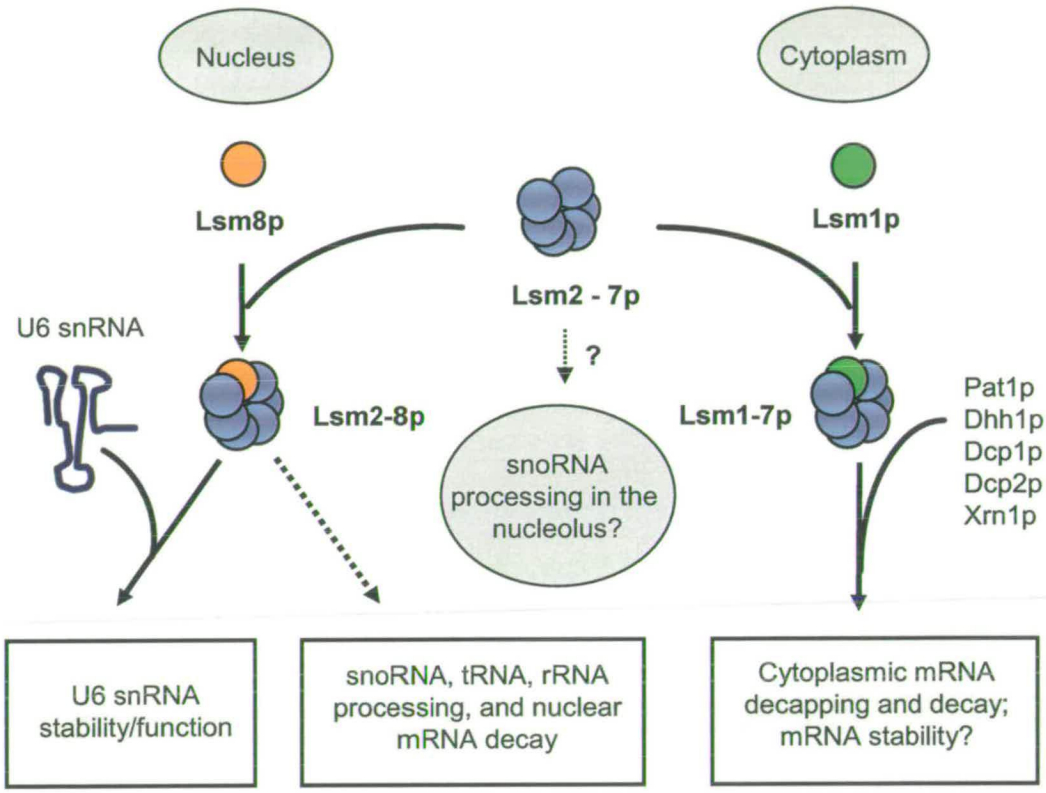


Figure 1.5 Overview of the wide range of functions of the Lsm proteins. The Lsm2 to 7 proteins either associate with Lsm8p to form Lsm2-8p, which functions in the nucleus, or with Lsm1p to form Lsm1-7p, which functions in the cytoplasm. There may be alternative Lsm complexes, e.g. an Lsm2-7p complex may exist with a potential function in snR5 snoRNA processing, function and/or stability.

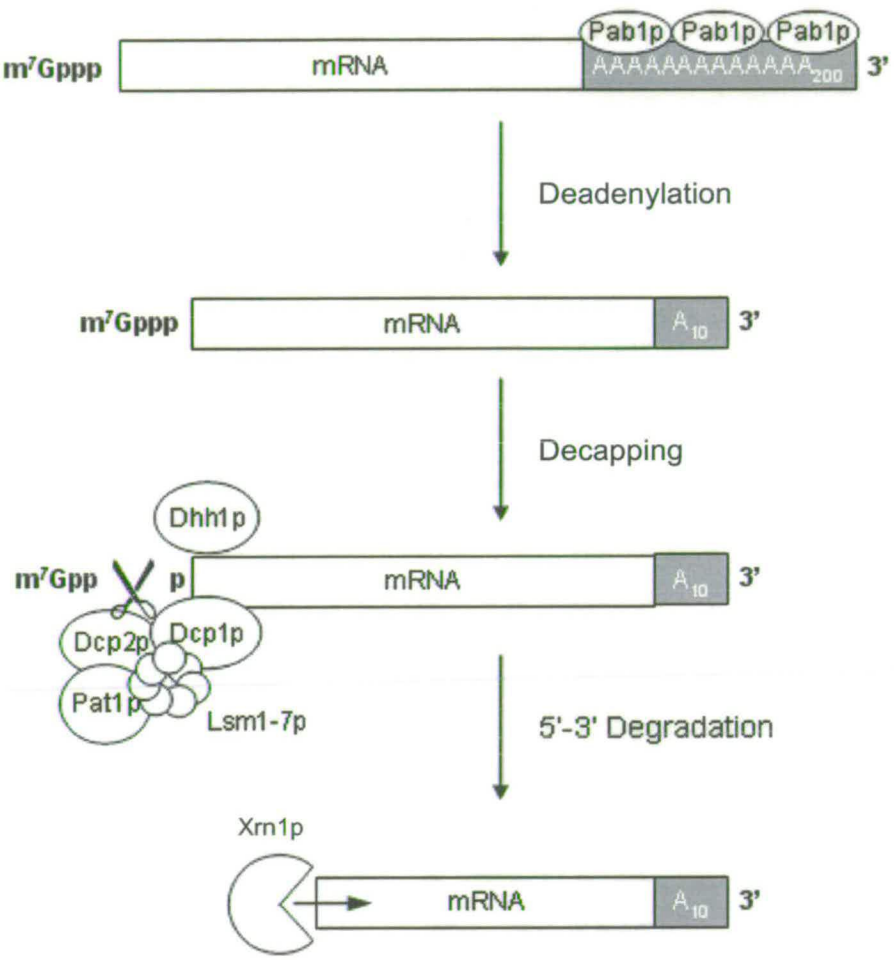


Figure 1.6 Schematic diagram of the deadenylation-dependent 5' to 3' mRNA degradation pathway in yeast. The poly(A) tail at the 3' of the mRNA is first degraded, causing the loss of Pab1p binding. This allows removal of the m^7G cap at the 5' end of the mRNA by Dcp1p, a step which requires a number of protein cofactors, including Dcp2p, Dhh1p, Pat1p and Lsm1-7p. After the cap has been removed, exonuclease Xrn1p degrades the mRNA in a 5' to 3' direction.

1.6.5 Cytoplasmic 5' to 3' mRNA decay occurs in P-bodies

In contrast to the exclusively nuclear Lsm8p, Lsm1p was shown to localise to the cytoplasm (Tharun *et al.*, 2000; Spiller, 2005). This is in line with its involvement in cytoplasmic mRNA decapping and decay. Human LSm1-7 (hLSm1-7), Dcp1/2 and Xrn1 were shown to colocalise in distinct cytoplasmic foci in HeLa cells (Ingelfinger *et al.*, 2002). These were later named processing or P-bodies, and in humans are also called GW bodies after one of its other protein components, GW182 (Eystathioy *et al.*, 2003). Other proteins involved in cytoplasmic mRNA decapping and degradation were found in the same foci, including hCcr4, rck/p54 (hDhh1) and eIF4E (van Dijk *et al.*, 2002; Cougot *et al.*, 2004; Andrei *et al.*, 2005). In yeast Lsm1p, Pat1p, Dcp1p, Dcp2p, Xrn1p and Dhh1p were shown to localise to similar cytoplasmic foci, with weak, possibly transient, focal localisation of Crr4p. Decapping and 5' to 3' mRNA degradation were shown to occur in these P-bodies, as inhibition of mRNA turnover at steps before decapping (*ccr4Δ*) lead to loss of visible foci. In contrast, inhibition of steps at or after decapping (*dcp1Δ* or *xrn1Δ*) lead to increased P-body numbers and size (Sheth and Parker, 2003). LSm complex formation was shown to be essential for localisation of hLSm1-7 to P-bodies as hLSm4 that carries mutations in residues that are proposed to be required for interactions between LSm subunits failed to localise to foci. Deletion of the RG-rich C-terminus of hLSm4 however, did not prevent its focal accumulation (Ingelfinger *et al.*, 2002).

In yeast, accumulation of mRNA in P-bodies was shown to occur in response to a variety of stress conditions, including glucose deprivation, osmotic stress and exposure to ultraviolet light, but not heat or oxidative stress. Additionally, P-bodies grew in number and size during late stages of cell growth. Inhibition of translation initiation promotes formation of P-bodies, whereas inhibition of translation elongation leads to the loss of P-body formation. P-body formation depends on mRNA and transcripts that accumulate in P-bodies are translationally repressed (Teixeira *et al.*, 2005). RNA that accumulates in P-bodies is therefore not necessarily targeted for degradation, but can be stored there in a translationally repressed state.

In summary, the Lsm1-7p/Pat1p/Dhh1p complex is involved in maintaining a proper balance between translating and non-translating pools of mRNA, and can act as a switch to target these mRNAs for degradation (Tharun *et al.*, 2001; Noueiry *et al.*, 2003; Brengues *et al.*, 2005; Teixeira *et al.*, 2005). Dhh1p and Pat1p were in fact shown to be required for a general mechanism of translational repression under stress conditions (Teixeira *et al.*, 2005; Coller and Parker, 2005). Retrograde movement between P-bodies and polysomes shows that this repression is reversible (Brengues *et al.*, 2005). In addition, specific sequences, like AREs, in certain cellular transcripts may affect posttranscriptional regulation by affecting the switch from translation to translationally repressed storage to degradation in response to intracellular or environmental cues. Translational repression and degradation of certain transcripts in response to miRNAs have also been shown to take place in P-bodies and RNA binding protein factors are involved in regulating, and potentially reversing this (Bruno and Wilkinson, 2006; Bhattacharyya *et al.*, 2006; reviewed by Parker and Sheth, 2007)).

1.7 Two (or more) complexes, many questions

It is fascinating that the simple difference of one subunit in otherwise identical Lsm complexes (Lsm1-7p vs. Lsm2-8p) causes them to localise and function differently (Fig. 1.5). Intriguingly, the *Trypanosoma brucei* genome was shown to encode only seven Lsm proteins. Functional characterisation of the Lsm3p and Lsm8p homologues showed that they were essential for U6 stability, *cis*- and *trans*-splicing as well as mRNA decay. This suggests that in trypanosomes only a single Lsm complex may exist (Liu *et al.*, 2004). An interesting question is what factors within the Lsm1 and 8 proteins are responsible for the unique subcellular localisation of the corresponding heteroheptameric complexes: the cytoplasm (P-bodies) or nucleus (Cajal bodies/nucleolus)? Do specific protein- or RNA- interaction domains affect their localisation or their function, or are their distinct functions a direct effect of discrete localisations? Is there competition in terms of formation of one complex or the other? If so, is this regulated simply by changes in expression levels of the Lsm1

and 8 proteins or do post-translational modifications play a role? This is even further complicated by the potential existence of alternative (L)Sm complexes.

1.8 Structural and functional studies of Sm(-like) complexes

1.8.1 The Sm heteroheptameric ring

Seven Sm proteins assemble around the U-rich Sm site present in the U1, U2, U4 and U5 spliceosomal snRNAs. The SmB/B' (SmB' is a splicing variant of SmB specific to humans), D1, D2, D3, E, F and G proteins each have two homologous sequence motifs, Sm1 and Sm2, separated by a short variable linker (Hermann *et al.*, 1995; Cooper *et al.*, 1995; Séraphin, 1995). These Sm motifs were shown to be essential and sufficient for the formation of the SmB/D3 dimer, which suggested the existence of a stable Sm domain that mediates protein-protein interactions (Hermann *et al.*, 1995). Stable subcomplexes SmB/D3, SmD1/D2 and SmE/F/G exist. SmD1/D2 and SmE/F/G can associate together on the snRNAs, with binding of SmB/D3 completing the snRNP core domain (Raker *et al.*, 1996). The crystal structures of the SmD1/D2 and SmB/D3 dimers reveal a common fold containing an N-terminal α-helix followed by a strongly bent five-stranded antiparallel β-sheet made of the Sm1 and Sm2 motifs (Fig. 1.7). Intermolecular interactions are mainly hydrophobic in nature. Pairing of β4 of one protein and β5 of its neighbour is stabilised by hydrophobic clustering and the angle between these strands in each protein determines the geometrical relationship between adjacent subunits. In addition to extensive clustering of hydrophobic residues on helices and β-sheets of binding partners a small number of salt bridges and hydrogen bonds are involved in stabilising multimerisation. Computer modelling, using these crystal structures together with data from yeast two-hybrid and immunoprecipitation experiments predicts the order of subunits in the ring: D1, D2, F, E, G, D3 and B. The resulting doughnut shaped ring (Fig. 1.7A) has an outer diameter of 70Å and a diameter of the central hole of 20Å (Kambach *et al.*, 1999).

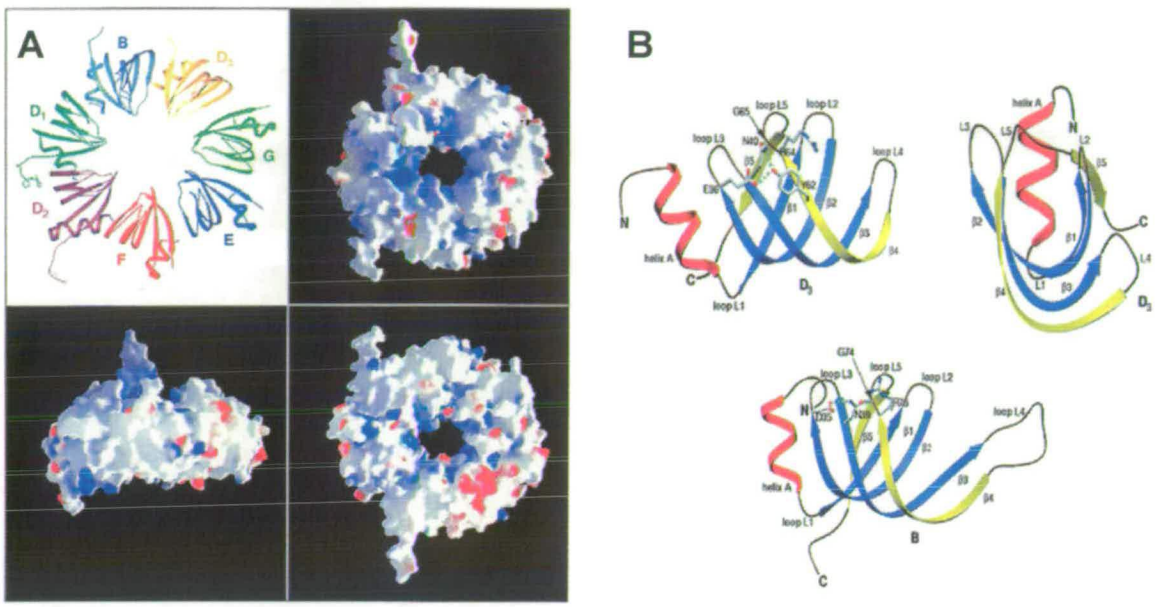


Fig. 1.7 Proposed higher-order structure of the human U snRNP core proteins. **(A)** Top left: Ribbon diagram of the heptamer model. The seven Sm subunits are arranged within the ring structure based on the crystal structures of the SmD1/D2 and SmB/D3 sub-complexes and pairwise interactions deduced from biochemical and genetic experiments. Others, moving in clockwise direction: Surface representation of the heptameric ring model with electrostatic potential (blue, positive; red, negative) viewed as in ribbon diagram, side view. and bottom view. Note that the central hole is highly basic and the surface of the assembly is asymmetric both in shape and charge. **(B)** Ribbon representations of the crystal structure of the SmD3 and B Proteins. Top: SmD3 protein (front view and side view) with the hydrogen bonding network involving Tyr-62 and highly conserved residues Glu-36, Asn-40, Arg-64, and Gly-65. Bottom: B protein (front view) including highly conserved residues Asp-35, Asn-39, Arg-73, and Gly-74. Reprinted from Kambach *et al.* (1999).

Based on these same structural studies it was suggested that the loops of each of the Sm proteins that protrude into the central hole of the ring may interact with the U-rich Sm site sequence through a number of highly conserved residues. The central hole is highly basic and large enough to accommodate single-stranded RNA (Kambach *et al.*, 1999). Electron micrographs of U2 snRNP (Kastner *et al.*, 1991) and U1 snRNP (Stark *et al.*, 2001) show that the 5' and 3' ends of these snRNAs are present on opposite sides of the core domain, supporting the view that RNA passes through the centre of the ring. Cross-linking of loop L3 residues of the Sm1 motif of G and B/B' to the first and third U-residues respectively of an Sm site RNA oligonucleotide confirm the importance of the central hole in RNA binding (Urlaub *et al.*, 2001). The uridine tract in the consensus Sm site sequence ($PuAU_{4-6}GPu$) provides one of the main recognition determinants for the Sm protein interactions, but is usually flanked by stem-loop structures, which may also contribute to Sm protein binding (Raker *et al.*, 1999). Mutagenesis of the U-rich tract leads to severely reduced binding in most cases, but does not have the same effect on all U snRNAs, e.g. yeast U5 snRNA (Jones and Guthrie, 1990) is more tolerant to point mutations in the Sm site than U4 snRNA (Hu *et al.*, 1995). Also, additional non-exchangeable snRNA elements were found to contribute to the assembly of U1 and U5 snRNP particles in *Xenopus* oocytes (Jarmolowski and Mattaj, 1993). Finally, the importance of the Sm site flanking stem-loop structures was confirmed by nucleotide analogue interference mapping in the U1 snRNP. The stem loops are therefore suggested to act as a clamp to hold the Sm ring in place (McConnell *et al.*, 2003).

1.8.2 The Lsm heptamers

Most functional studies that were done on the eukaryotic Lsm1-7 and Lsm2-8 complexes are discussed in sections 1.4 - 1.6. Comparatively little progress has been made in terms of structural studies of these Lsm complexes. Most of what is known about them comes from yeast two-hybrid analyses (Fromont-Racine *et al.*, 2000; Lehner and Sanderson, 2004) and sequence comparison to the canonical Sm proteins (Fig. 1.8). The order of the proteins in the ring has never been confirmed, but is assumed to be Lsm1p/8p, Lsm4p, Lsm7p, Lsm5p, Lsm6p, Lsm3p, Lsm2p (Fig. 3.5),

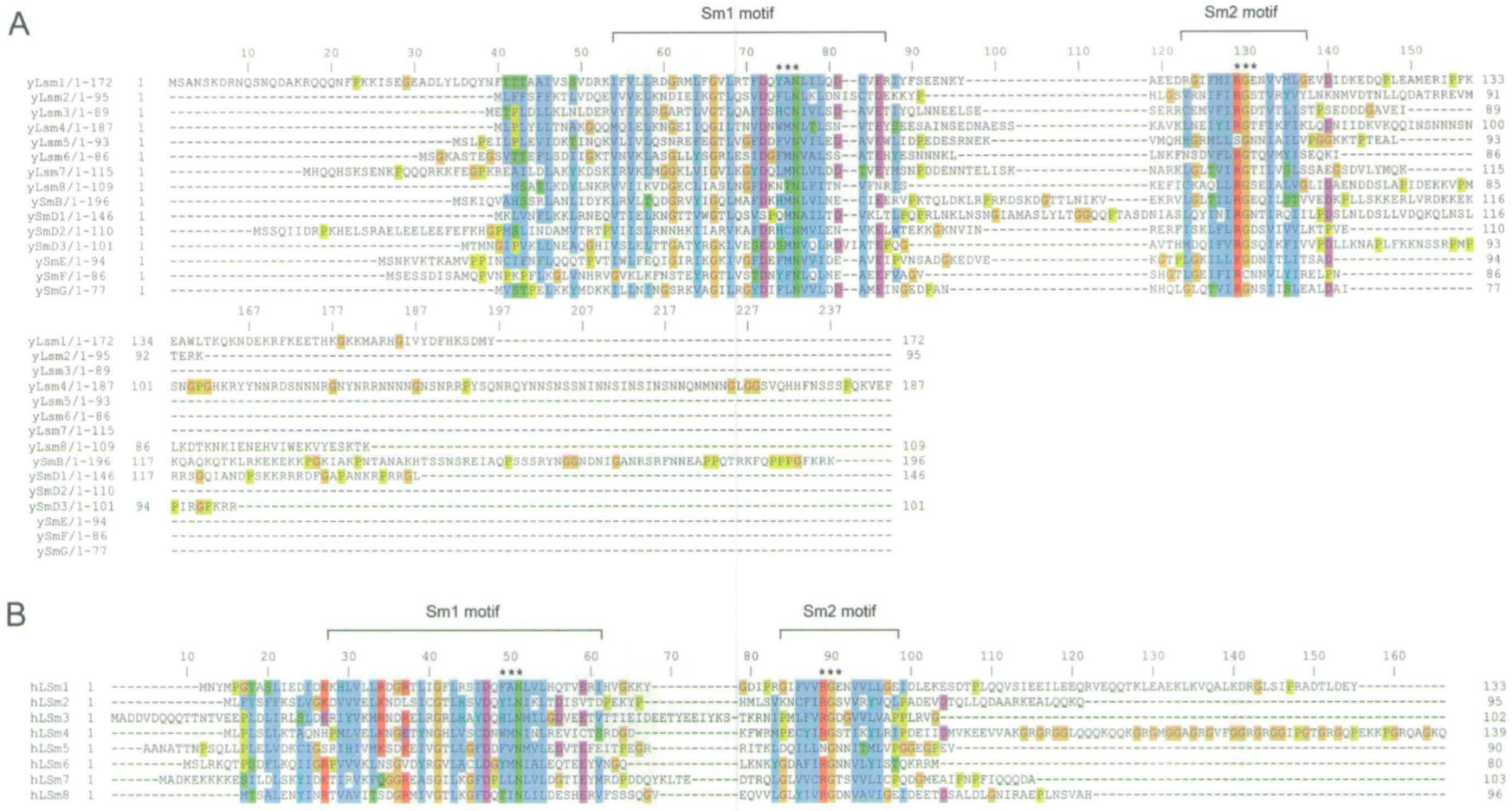


Fig. 1.8 Alignment of amino acid sequences of the yeast and human (L)Sm proteins. **(A)** The canonical yeast Sm and Lsm proteins were aligned using ClustalW and adjusted by hand using Jalview; **(B)** The human LSm proteins were similarly aligned. Conserved Sm1 and Sm2 motifs are indicated as well as putative RNA binding residues (*). The most recently predicted amino acid sequences for *S. cerevisiae* Lsm6p, Lsm7p and Lsm8p were used (see Chapter 3). To adjust alignments by hand, previously published data were used from Cooper *et al.* (1995), Séraphin (1995), Achsel *et al.* (1999), Palfi *et al.* (2000) and Bordonné (2000).

by analogy to the Sm complex. Yeast two-hybrid data on the human LSm proteins by Lehner and Sanderson (2004), who analysed the stringency of all pair-wise interactions, convincingly support this order. Sequence analysis indicates the presence of Sm domains in all Lsm proteins (Fig. 1.8) and electron microscopy on the purified human LSm2-8 complex revealed a doughnut shape (Achsel *et al.*, 1999). Recent attempts to assemble the human LSm1-7 and LSm2-8 complexes *in vitro* using recombinant subcomplexes were successful, showing a similar shape for the LSm1-7 complex. Attempts to crystalise these complexes were however, unsuccessful (Zaric *et al.*, 2005).

1.8.3 Archaeal homomultimeric rings

Since the discovery of the Sm domain and its involvement in oligomerisation of proteins into ring structures with RNA binding capacity, a range of Sm-like proteins from archaea has been identified and studied (Collins *et al.*, 2001; Törö *et al.*, 2001; Achsel *et al.*, 2001; Mura *et al.*, 2001; Thore *et al.*, 2003; Mura *et al.*, 2003a; Mura *et al.*, 2003b). Most archaea only contain one or two different Sm-like proteins and so far these have been shown to form homoheptameric or hexameric ring structures (Fig. 1.9), although interactions between rings and even polymerisation into polar fibres have been shown to occur for some (Mura *et al.*, 2003a; Mura *et al.*, 2003b; Arluison *et al.*, 2006). The biological significance of these multimers of multimers is unclear. Like their eukaryotic counterparts, archaeal Sm-like complexes have nucleic acid binding properties, with a strong preference for oligo(U) RNA (Achsel *et al.*, 2001). They have been implicated in pre-tRNA processing, through binding with RNase P RNA (Törö *et al.*, 2001). Some also bind DNA (Mura *et al.*, 2003a), although this may be non-specific and therefore biologically irrelevant. Many of the crystal structures have been solved and show highly homologous structural features and similar dimensions to the Sm ring. Co-crystalisation of *Archaeoglobus fulgidus* Sm1 (Törö *et al.*, 2001) and *Pyrococcus abyssi* Sm1 (Thore *et al.*, 2003) with uridine oligonucleotides give clues to how these homoheptamers bind single-stranded RNA. Uridine specific binding-pockets are present on the inside of the ring, whereas a secondary RNA binding site on the surface of the ring is involved in additional contacts with the RNA outside the oligo(U) sequence.

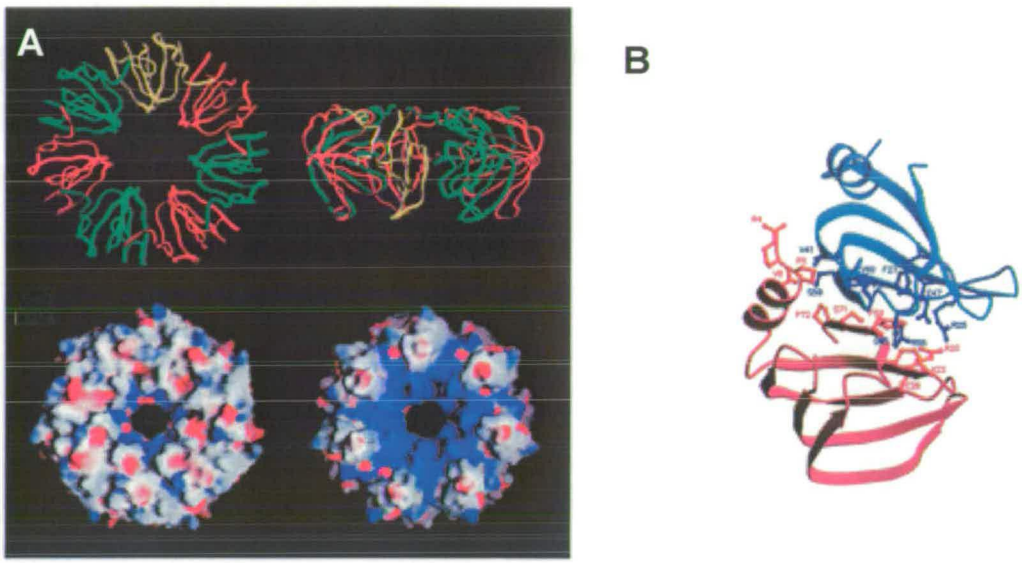


Fig. 1.9 Structure of the *Archaeoglobus fulgidus* Sm1 heptamer. **(A)** Ribbon representation of the AF-Sm1 heptamer (top and side view) and electrostatic surface charge potential of the two face of the ring. Shown on the left is the side binding the RNA and containing the N-terminal helix (top view). On the right is the bottom view showing positively charged grooves (blue) emanating from the centre. **(B)** Subunit contacts in the AF-Sm1 heptamer. The molecules are shown as ribbons of different colours for the pairs of interacting molecules. Side chains involved in contacts, mainly located on β-strands 4 and 5 as well as the N-terminal helix, are represented in ball-and-stick mode. Reprinted from Törö *et al.* (2001).

These findings neither disprove, nor prove the possibility of the RNA threading through the inner hole of the complex. The oligomerisation behaviour of the Sm2 type protein from *Archaeoglobus fulgidus* was shown to depend on the pH and presence of RNA. This Sm-like protein formed a hexamer at acidic pH in the absence of RNA and a heptamer in the presence of U_{10}-RNA over a wide pH range. The stability of both the hexamer and the RNA-bound heptamer increased at lower pH (Kilic *et al.*, 2006).

1.8.4 Hfq, a bacterial Lsm complex with RNA chaperone activity

Important implications for the evolution of Sm-like proteins and their functions came from the discovery of the Hfq Sm-like protein from *E. coli* and homologues in other bacteria. Initially Hfq (also known as HF-I) was identified as an *E. coli* host factor required for initiation of plus-strand synthesis by the replicase of the Qβ RNA bacteriophage (Franze de Fernandez *et al.*, 1968). Later, it was shown to specifically bind the 3' end of the Qβ RNA and to melt out a secondary structure, thus facilitating binding by the replicase protein (Schuppli *et al.*, 1997). Much more recently it was shown to be the bacterial equivalent of the eukaryotic (L)Sm proteins. It associates with various RNAs, binds to U-rich tracks of sRNAs and an Sm-like motif is present in the N-terminal portion of the protein. Electron microscopy studies revealed a doughnut-shaped oligomeric molecule (Møller *et al.*, 2002; Zhang *et al.*, 2002) and the crystal structure of *Staphylococcus aureus* Hfq (Schumacher *et al.*, 2002) and a C-terminally truncated Hfq from *E. coli* (Sauter *et al.*, 2003) confirmed its structural homology to (L)Sm proteins (Fig. 1.10). It is not clear what makes Hfq form hexamers, whereas most other Sm-like proteins form heptamers, but the result is a similarly doughnut shaped ring. *S. aureus* Hfq was also co-crystalised with an AU_5G oligonucleotide giving insight into Hfq RNA binding. Neighbouring Hfq monomers form nucleotide binding pockets with specificity for adenosine and uridine residues located around the central cavity on one face of the ring (Schumacher *et al.*, 2002). In agreement with this Hfq generally recognises a minimal RNA domain consisting of an A/U-rich track and, analogous to the Sm proteins, one or more flanking hairpins. Additional RNA-binding sites have been shown to exist on the opposite face of the ring (Mikulecky *et al.*, 2004). Hfq has a long hydrophilic N-terminal tail,

which may be involved in binding to the hairpins or in binding to the mRNA targets. Alternatively, it could provide a platform for other cellular binding partners like the ribosome with which the majority of Hfq is associated (Kajitani *et al.*, 1994; Sauter *et al.*, 2003).

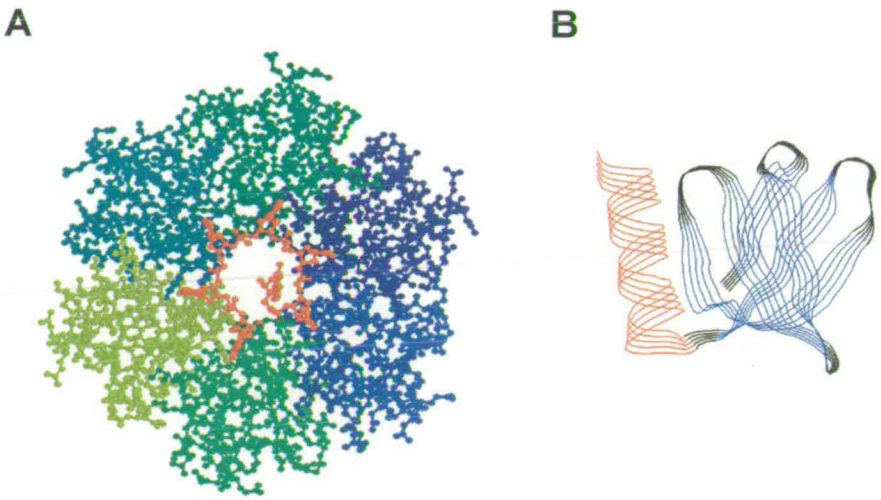


Fig. 1.10 Structure of the *Staphylococcus aureus* Hfq hexamer bound to an RNA oligonucleotide. **(A)** Ball-and-stick representation of the Hfq hexamer (top view). Each Hfq monomer is painted a different colour; the RNA oligonucleotide is shown in orange; **(B)** Ribbon representation of a Hfq monomer. The N-terminal helix is shown in red, β -strands 1 to 5 in blue and loops in black. Coordinates from PDB (1kq2) were used in Rasmol to produce these images. The crystal structure was solved by Schumacher *et al.* (2002).

As has been proposed for other Lsm complexes, Hfq acts as an RNA chaperone, promoting RNA-RNA interactions, often, but not always by modulating secondary structure of its target RNA molecules (reviewed by Storz *et al.*, 2004; Valentin-Hansen *et al.*, 2004). It acts on a range of non-coding sRNAs and their target mRNAs. Some sRNAs are involved in translational control, e.g. *OxyS* sRNA basepairs with and represses translation of *fhlA* mRNA, whereas *DsrA* sRNA promotes translation of *rpoS* mRNA by preventing the formation of an inhibitory secondary structure. A striking example of its involvement in mRNA stability is the regulation of degradation of *sodB* mRNA which encodes an iron superoxide dismutase. Under conditions of limiting iron, *RhyB* sRNA is expressed and promotes the degradation of *sodB* by making it more sensitive to RNase E digestion. Under conditions of high iron the transcriptional regulator protein Fur represses *RhyB* expression, thus downregulating *sodB* degradation and that of at least five other mRNAs encoding proteins involved in the iron metabolism. Storz *et al.* (2004) suggested that sRNAs may also protect certain mRNAs from degradation by binding to them. Binding of Hfq itself to the poly(A) tail of *rpsO* mRNA has already been shown to protect it from RNase degradation *in vitro* (Folichon *et al.*, 2003).

In summary, Hfq has been shown to alter RNA-RNA and RNA-protein interactions by changing RNA secondary structures in some cases (Schuppli *et al.*, 1997; Møller *et al.*, 2002; Zhang *et al.*, 2002; Moll *et al.*, 2003; Geissmann and Touati, 2004), but has been shown to be able to promote RNA-RNA interactions without affecting RNA secondary structures in others (Brescia *et al.*, 2003; Mikulecky *et al.*, 2004). This RNA chaperone function may be conserved between all Sm-like protein complexes. Apart from this, multiple other parallels exist between Hfq and other Sm-like proteins, indicating that lessons about the more intricate eukaryotic Lsm complexes may be learnt from its relatively simple bacterial counterpart (Wilusz and Wilusz, 2005).

1.8.5 (L)Sm RNA binding activity

Comparison of available crystal structures of Sm(-like) complexes with and without RNA, and alignment of known Sm(-like) proteins has provided insight into the way

these complexes bind RNA (Khusial *et al.*, 2005). Bases of the bound nucleotides are held in a pocket created by two three amino acid motifs in loops between strands β2 and β3, and between β4 and β5. Khusial *et al.* (2005) named these motifs Knuckle I and Knuckle II respectively. Differences in residues used to form these nucleotide binding pockets were suggested to be the cause of more promiscuous RNA binding by the Lsm complexes, compared to the Sm complexes (also see Chapter 9).

1.8.6 Alternative (L)Sm complexes

1.8.6.1 U7 snRNP

In metazoa, processing of the non-polyadenylated 3' end of replication-dependent histone mRNA is cell-cycle regulated. An essential factor for endonucleolytic cleavage of the 3' UTR is the U7 snRNP. This RNP consists of U7 snRNA and an unusual (L)Sm core complex. It has a combination of Sm (SmB, D3, E, F and G) and LSm proteins (LSm10 and 11). The latter two effectively replace SmD1 and D2. U7 snRNA has a degenerate Sm binding site and a free 5' end that is complementary to the histone mRNA 3' UTR. Basepairing between U7 snRNA and histone mRNA helps to position the endonucleolytic cleavage activity close to the appropriate site. Interactions between N-terminal extension of LSm11p and other protein factors involved in the processing stabilise the RNA-RNA interaction and provide additional binding specificity (reviewed by Schümperli and Pillai, 2004).

1.8.6.2 snRNP-specific Sm complexes in trypanosomes

The composition of the Sm core complex of the *Trypanosoma brucei* U2 and U4 snRNPs was recently shown to be different than that usually seen for the spliceosomal U snRNPs. In the *T. brucei* U2 snRNP SmB and D3 proteins are replaced with homologous proteins SSm2-1 (Sm16.5K) and SSm2-2 (Sm15K). This exchange of two of the Sm subunits seems to contribute to RNA binding specificity of the U2 snRNA Sm binding site. In addition to specific RNA binding the alternative Sm subunits may contribute to U2 snRNP specific functions, possibly through unique protein-protein interactions (Wang *et al.*, 2006; Tkacz *et al.*, 2006). Surprisingly, in the U4 snRNP only SmD3 is replaced by a snRNP specific Sm protein designated SSm4 (Tkacz *et al.*, 2006).

1.8.6.3 Telomerase

The RNA component of both yeast and human telomerase was shown to associate with Sm proteins. In yeast this occurs via a canonical Sm site, identical to that of U4 snRNA, at the 3' end of *TLC1* RNA. Tagged SmD1p and SmD3p precipitated *TLC1* RNA, but no other (L)Sm proteins were tested (Seto *et al.*, 1999). In humans tagged SmB and SmD3, but none of the other Sm proteins co-immunoprecipitated hTR RNA. This human telomerase RNA does not contain an Sm site, but the interaction was instead shown to require the CAB box sequence, that targets RNA to Cajal bodies. SmB and D3 were also shown to interact with other CAB box containing small Cajal body RNAs (scaRNAs; Fu and Collins, 2006). The precise mode of interaction and composition of the telomerase associated Sm complex, particularly in humans, is unclear and it may consist of yet another combination of (L)Sm proteins.

1.8.6.4 Hexameric Lsm complexes?

Multiple Lsm complexes have been proposed to exist in which one or two Lsm proteins may be missing. LSm5 was not found associated with U8 snoRNA in *Xenopus*; Lsm8p did not co-immunoprecipitate yeast snR5; and defects in tRNA, rRNA and U3 snoRNA processing were strong in yeast strains depleted of any of the essential Lsm proteins, but minimal in the absence of Lsm1p, Lsm6p or Lsm7p (see 1.5). This may suggest existence of alternative Lsm complexes. Importantly, none of these studies conclusively demonstrated the actual composition of any of these complexes. However, existence of functional hexameric complexes or heptamers with alternative subunits can not be ruled out.

1.8.6.5 Other *S. cerevisiae* Sm-like proteins

In addition to seven canonical Sm proteins and eight Sm-like proteins, the *S. cerevisiae* genome encodes a number of other proteins that contain Sm domains. One of these, *LSM9* is also called *MAK31* (maintenance of killer). It is non-essential, but a deletion mutant strain displays elongated telomere length (Askree *et al.*, 2004) and exhibits defects in the structural stability of L-A family of dsRNA-containing viral particles (Wickner and Leibowitz, 1976). In addition, its deletion leads to a synthetic

phenotype with *csm3* and *tof1* (Pan *et al.*, 2006), both of which are required for mitotic sister chromatid cohesion. *LSM9* was also found as a high copy suppressor of a temperature-sensitive mutation affecting the RSC (remodels the structure of chromatin) complex, the yeast SWI/SNF chromatin remodelling complex (Treich *et al.*, 1998). It is most closely related to Lsm7p and SmBp (Fig. 1.11A), but has no direct structural relatives in other organisms (Salgado-Garrido *et al.*, 1999). It has yeast two-hybrid interactions with Lsm2p and Lsm8p (Fromont-Racine *et al.*, 2000), but has a poorly conserved Sm2 motif. A usually highly conserved glycine or cysteine residue in the putative RNA binding motif of Sm2 is absent (Séraphin, 1995) and the motif seems to be disrupted by a proline residue (Fig. 1.11A). It is therefore unclear whether it could be part of an alternative Lsm complex or whether it can bind RNA. Its genetic interactions strongly suggest that it may have a function in chromatin remodelling and/or telomere maintenance. However, lack of L-A viral particle assembly was associated with loss of NatC N-terminal acetyl transferase activity of a Mak31p containing complex (Polevoda and Sherman, 2001). Its effects on chromatin remodelling and telomere maintenance may involve this activity and may be indirect.

Various bioinformatic approaches found a number of other yeast proteins that contain divergent Sm domains (Albrecht and Lengauer, 2004; Anantharaman and Aravind, 2004). Lsm12p is a protein of unknown function that localises to the nucleus and the cytoplasm, interacts with Pbp1p (Poly(A) binding protein binding protein) and Pbp4p (Pbp1p binding protein), and associates with ribosomes (Fleischer *et al.*, 2006). Pbp1p and Pbp4p are involved in regulation of polyadenylation (Mangus *et al.*, 1998; Mangus *et al.*, 2004). Interestingly, Pbp1p itself also contains an Sm domain (Fig. 1.11B). Lsm12p and Pbp1p may therefore interact via their Sm domains and may be part of the same RNA processing machinery. Lsm13p is also called Scd6p (suppressor of clathrin deficiency) and localises to the cytoplasm. Its overproduction suppresses a null mutation in *CHC1*, which encodes the heavy chain of clathrin (Lemmon, unpublished; see SGD). It is a P-body component and orthologues (*Drosophila* trailer hitch, *C. elegans* CAR-1 and human RAP55) have links to various cellular processes including spindle pole

regulation and endoplasmic reticulum dynamics (Barbee *et al.*, 2006; Decker and Parker, 2006 and references therein). *C. elegans* CAR-1 was shown to bind poly(U) RNA (Audhya *et al.*, 2005) and the Lsm domain of Xenopus RAP55 was shown to be important for translational repression and P-body localisation (Tanaka *et al.*, 2006). Another Sm-like gene, *LSM16* (*EDC3/DCP3*) also encodes a cytoplasmic protein that localises to P-bodies. It is an enhancer of deadenylation-independent mRNA decapping by Dcp1p/Dcp2p (Kshirsagar and Parker, 2004; Badis *et al.*, 2004). Interactors found by affinity capture include Dcp1p, Dcp2p, Lsm1p, Lsm3p and Xrn1p (Gavin *et al.*, 2006; Krogan *et al.*, 2006). Others, found by yeast two-hybrid analysis include Dhh1p (Uetz *et al.*, 2000), Edc3p, Lsm2p and Lsm8p (Fromont-Racine *et al.*, 2000).

Lsm12p (LSm12), Lsm13p (LSm14a/RAP55 and b), Lsm16p (Edc3) and Pbp1 (Ataxin-2) all have homologues in human (indicated between brackets). Both Sm motifs for these proteins are weakly conserved, with no significant conservation of the putative RNA binding motifs. This does not necessarily mean that these proteins are not able to bind RNA. *E. coli* Hfq is included in the alignment for comparison (Fig. 1.11B). Proposed RNA binding residues are poorly conserved between Lsm6p and Hfq, but Hfq does bind RNA and Hfq residues (YKH) that line up with Lsm6p (RGT) in Sm motif 2 are required for RNA binding (Mikulecky *et al.*, 2004). Also, the Sm-like domain of CAR-1 was already shown to bind RNA (Audhya *et al.*, 2005). It is unclear whether these proteins are able to form complexes with other (L)Sm proteins. The nature of the Sm-fold makes it likely that they at least form homomultimeric complexes.

1.8.7 Evolutionary considerations

It appears that bacteria have retained a unique RNA chaperone involved in many stages of RNA metabolism. Most archaeal genomes contain one or two Sm-like genes, which may be the result of gene duplication. In eukaryotes further gene and/or genome duplications may have lead to structural and functional divergence of a large number of (L)Sm proteins. Heteromultimerisation of different combinations of these

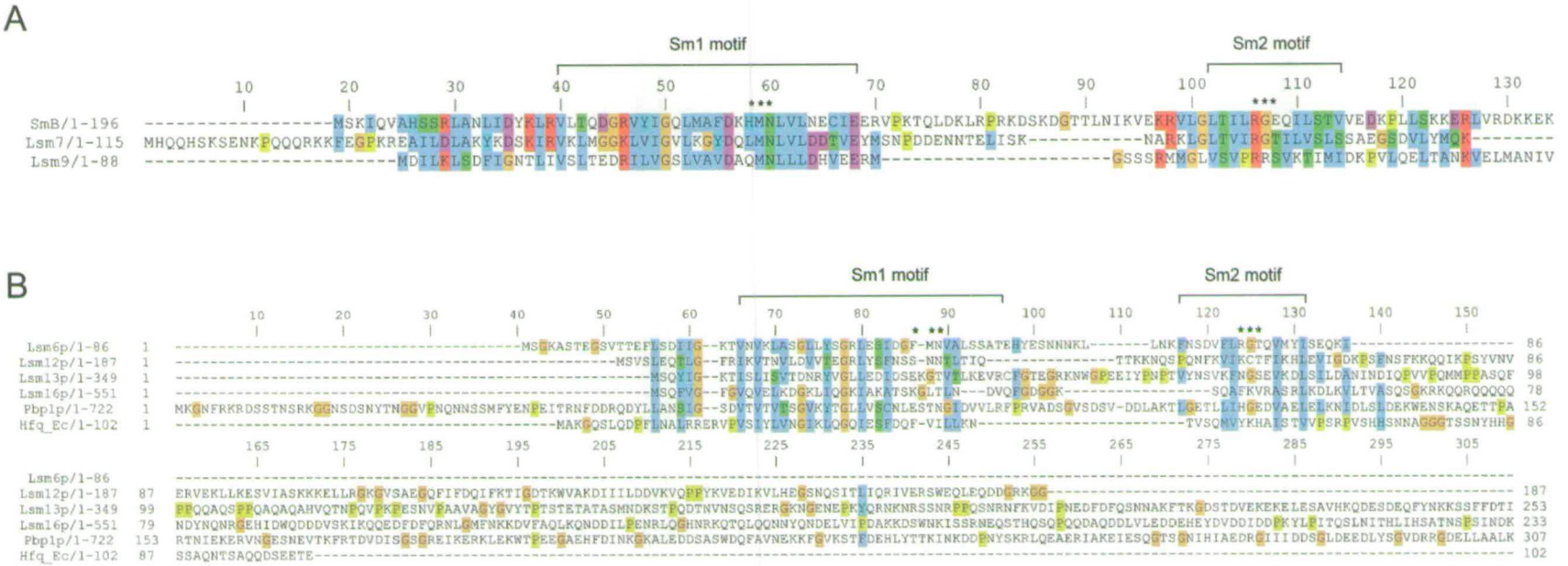


Fig. 1.11 Alignment of amino acid sequences of putative *S. cerevisiae* Lsm proteins. **(A)** The yeast SmB, Lsm7 and Lsm9 proteins were aligned using ClustalW and adjusted by hand using Jalview. Previously published data were used from Séraphin (1995); **(B)** *S. cerevisiae* Lsm6p, Lsm12p, Lsm13p, Lsm16p and Pbp1p were similarly aligned with *E. coli* Hfq. Conserved Sm1 and Sm2 motifs are indicated as well as putative RNA binding residues (*). To adjust alignments by hand previously published data were used from Albrecht and Lengauer (2004) and Anantharaman and Aravind (2004).

polypeptides allows a much higher level of complexity. Thus, heteromers with a variety of specialised functions could evolve (reviewed by Beggs, 2005).

1.9 Lsm proteins in disease

1.9.1 Systemic lupus erythematosus

Patients with the auto-immune disorder systemic lupus erythematosus (SLE) often have autoantibodies to Sm-protein antigens (Lerner *et al.*, 1979). The majority of anti-Sm sera from patients with SLE also recognise hLSm4, while a small percentage shows exclusive reactivity to the hLSm complex in the absence of reactivity to Sm core proteins. Additionally, the finding of anti-hLSm4 antibodies in infectious mononucleosis suggests that infection with Eppstein-Bar Virus may play a role in initiating the production of autoantibodies (Eystathioy *et al.*, 2002).

1.9.2 Lsm proteins and spinal muscular atrophy

The spinal muscular atrophy (SMA) disease gene product, SMN, is crucial for snRNP biogenesis. A complex containing SMN interacts with the snRNP core proteins, both Sm and LSm (Friesen and Dreyfuss, 2000). A direct correlation exists between reduced levels of SMN polypeptide in SMN patient cell extracts and their ability to promote snRNP assembly (Wan *et al.*, 2005). However, recent evidence suggests that SMN has a novel function in motor axon outgrowth that is relevant to SMA but independent of snRNP biosynthesis. Mutations in SMN were identified that failed to rescue motor axon defects, but retained the ability to oligomerise and bind Sm proteins, functions required for snRNP biogenesis (Carrel *et al.*, 2006). Interestingly, a modified U7 snRNP has recently been shown to have potential as a form of gene therapy for the treatment of SMA. Wild-type U7 snRNA associates with an unusual (L)Sm core complex (reviewed by Schümperli and Pillai, 2004; see 1.8.6.1). SMA, a usually fatal, hereditary disease is caused by absence of SMN1 expression and differential splicing of a second gene copy, SMN2. Exon 7 skipping in SMN2 leads to production of a biologically inactive protein isoform. To increase exon 7 inclusion of SMN2, vectors expressing modified U7 snRNAs containing antisense sequences complementary to the 3' splice site of SMN exon 8 were

transiently expressed in HeLa cells. Increased exon 7 inclusion in SMN2 resulted in an increase in the concentration of SMN protein (Madocsai *et al.*, 2005). A similar approach was shown to have potential for the treatment of β-thalassemias (Schümperli and Pillai, 2004).

1.9.3 Lsm proteins and cancer

Human LSm1 is also called CaSm (cancer-associated Sm protein) as its expression is upregulated in cancers of the pancreas (Schweinfest *et al.*, 1997; Gumbs *et al.*, 2002), breast (Streicher *et al.*, 2006) prostate (Fraser *et al.*, 2005) and various cell lines from cancers originating in liver, ovary, lung, and kidney (Schweinfest *et al.*, 1997). LSM1 overproduction has been directly linked to a transformed cell phenotype (Schweinfest *et al.*, 1997; Yan *et al.*, 2005). Delivery of antisense LSM1 RNA using adenovirus inhibits growth of prostate cancer cell lines by altering cell cycle progression (Kelley *et al.*, 2003), and is associated with altered expression of cyclin B1, CDK1, NEK2 and nucleophosmin/B23 proteins. Elevated LSM1 expression in cancers may therefore involve (de)stabilisation of multiple gene transcripts (Fraser *et al.*, 2005). In a mouse model of pancreatic cancer (Yan *et al.*, 2005), tumor growth was shown to be inhibited by intratumoral administration of an adenoviral vector encoding LSM1 antisense RNA and systemic administration lead to decreased metastasis and increased survival time (Kelley *et al.*, 2000a; Kelley *et al.*, 2001; Yan *et al.*, 2006). Surprisingly, others found that LSM1 was downregulated in advanced human prostate cancers (Takahashi *et al.*, 2002), suggesting that not only LSM1 overexpression is associated with cancer, but emphasising the importance of correct levels of LSm1. Interestingly, increased copynumber of other LSM genes has recently been linked to cancer. Increased copy number and expression of LSM3 were shown to be associated with metastatic phenotypes of cervical cancers (Lyng *et al.*, 2006). In addition, LSM7 expression was shown to be increased in malignant thyroid tumours (Rosen *et al.*, 2005). The latter was suggested to involve an interaction between LSM7 and TACC1, which itself is downregulated in various carcinomas. The function of TACC proteins is poorly understood, but orthologues from Xenopus (Maskin) and Drosophila (D-TACC) are involved in translational regulation and dynamics of spindle pole

microtubules during cell division (Conte *et al.*, 2002; Conte *et al.*, 2003). It seems that the LSm proteins, most likely LSm1-7, are involved in cell cycle regulation and changes in their levels can result in carcinogenesis. The underlying mechanism is likely to involve regulated degradation of cell cycle specific transcripts, but may include as yet unknown functions for LSm1-containing Sm-like complexes.

1.9.4 Other links between Sm-like proteins and disease

1.9.4.1 Hfq and bacterial virulence

The bacterial Sm-like protein Hfq has been implicated in the regulation of virulence factors in a wide range of pathogenic bacteria, including *Yersinia enterocolitica*, *Brucella abortus*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Vibrio cholerae* and *Listeria monocytogenes*. Experiments with *hfq* deletion strains have shown a direct effect of the deletion on the production of virulence factors through the actions of small regulatory RNAs (Christiansen *et al.*, 2006 and references therein).

1.9.4.2 Lsm1-7p in viral replication and translation

The Lsm1-7p/Pat1p/Dhh1p deadenylation dependent decapping activator complex was shown to be essential for Brome Mosaic Virus (BMV) replication in yeast. On the other hand, there was no apparent involvement of other decapping factors or deadenylation independent degradation factors (Diéz *et al.*, 2000; Mas *et al.*, 2006). The same complex was shown to be needed for efficient translation of the viral RNA (Noueiry *et al.*, 2003). Polysome profiles showed that, while mutations in the Lsm1-7p/Pat1p complex completely inhibited viral RNA translation, the levels of viral RNA associated with ribosomes were only slightly reduced in mutant yeast. Inhibition of BMV RNA translation was selective, with no effect on general cellular translation. Thus a defective Lsm1-7p/Pat1p complex specifically inhibits BMV RNA translation at the level of translation elongation. A switch between translation and recruitment of BMV RNAs to viral RNA replication seems to be regulated by a pathway that usually transfers mRNAs from translation to degradation (see section 1.6.5). As any (+)strand RNA virus needs to make this switch during its life cycle

Lsm1-7p is likely to be involved in replication of other such viruses (Noueiry *et al.*, 2003; Mas *et al.*, 2006).

1.10 Aims of this thesis

To investigate a role for *S. cerevisiae* Lsm2-8p in U4/U6 annealing, attempts were made to assemble this heteroheptameric complex *in vitro* from its purified recombinant constituents. To determine specificity of RNA binding and U4/U6 annealing activity, the Lsm1-7 complex was similarly produced. For comparison *E. coli* Hfq and *in vitro* assembled human complexes were used. In **Chapter 3**, I describe the development of the methods used to purify, assemble and analyse these complexes. In **Chapter 4**, I analyse RNA-binding and U4/U6 annealing activities for the different complexes. In **Chapter 5**, I investigate in more detail a potential mechanism for the suggested U4/U6 annealing activity. This includes the design of a strand displacement assay to examine the ability of Sm-like proteins to unwind a duplex structure. In summary, Chapters 4 and 5 examine the putative RNA chaperone function of Sm-like protein complexes.

Prp24p and Lsm2-8p are both part of the U6 snRNP and have both been implicated in U4/U6 annealing. In **Chapter 6**, I investigate genetic interactions between *lsm* and *prp24* mutations and attempt to clarify direct Lsm-Prp24p interactions by immunoprecipitation of purified recombinant proteins. Using these data and the available literature I propose a model for the involvement of these proteins in U4/U6 di-snRNP assembly.

Previous members of this lab found specific phenotypes associated with Lsm4p overproduction. In **Chapter 7**, I describe experiments that investigate toxicity associated with Lsm4p overproduction, aggregation of Lsm4p into foci and a potential role for the Lsm4 C-terminal extension in Lsm1-7p localisation to P-bodies. In addition, I investigate a potential role for Lsm4p in cell morphology and cytokinesis.

Deletion of *LSM1*, *LSM6* or *LSM7* leads to temperature-sensitive growth and the effects of these deletions are known in some detail. In **Chapter 8**, I investigate effects of mutation and deletion of *LSM5* on growth and on the levels of spliceosomal snRNAs, and compare these to effects of *lsm1Δ*, *lsm6Δ* and *lsm7Δ*.

Lsm1-7p and Lsm2-8p only differ by one subunit, but localise to different subcellular sites where they carry out distinct functions. In **Chapter 9**, I describe experiments with Lsm1p and Lsm8p mutant proteins as well as hybrids that combine domains of both polypeptides. I investigate their ability to support viability; I analyse *in vivo* localisation of fluorescently tagged mutant and hybrid proteins; and show the importance of specific domains for localisation to nuclei and P-bodies.

In summary, this thesis attempts to elucidate functions of Sm-like proteins, yeast Lsm1-7p and Lsm2-8p in particular, and the underlying mechanisms. In addition, specific functions associated with individual Lsm polypeptides and protein domains are investigated.

CHAPTER 2

Materials and Methods

2.1 General reagents

2.1.1 Sources of reagents

2.1.1.1 Chemicals

Chemicals were purchased from the following sources, except where stated otherwise: Sigma, BDH Laboratory Supplies (AnalaR, VWR), Fisher Chemicals, Melford Labs, Scotlab and Amersham Biosciences (GE Healthcare). Water used to make aqueous solutions was deionised using a Purite Prestige Deioniser to 1-10 μS (with C750 carbon pre-treatment and HP700 deioniser columns).

2.1.1.2 Enzymes

Restriction enzymes, DNA and RNA polymerases and other enzymes were purchased from the following sources unless otherwise stated: Epicentre Technologies, Gibco BRL, New England Biolabs, Pharmacia, Promega and Roche.

2.1.1.3 Growth reagents

Reagents for all growth media were purchased from the following sources: BD Biosciences, Difco Laboratories, Fisher Scientific, Formedium, Gibco BRL, Oxoid and Sigma.

2.1.1.4 Antibiotics

Ampicillin was purchased from Beecham Research and GlaxoSmithKline, kanamycin from Merck, hygromycin from Calbiochem, G418 from Melford, tetracycline from Duchefa and chloramphenicol and streptomycin were purchased from Sigma. Nourseothricin was purchased from HKI, Werner BioAgents (Jena, Germany) under the trade name clonNAT.

2.1.2 Bacterial and yeast growth media

2.1.2.1 General information

All liquid growth media were autoclaved (Thermo Denley, Diamond 36, 10 min at 121°C), stored at room temperature and warmed to the appropriate temperature immediately prior to use. For solid media, 2% (w/v) agar (Oxoid) was added prior to

autoclaving. Plates of 15-25 ml each were poured in 9 cm triple-vented petridishes (Greiner Bio-One), allowed to set at room temperature and stored at 4°C.

2.1.2.2 Bacterial media

Escherichia coli were grown in Luria Broth (LB; also known as Luria-Bertani or Lennox Broth, but originally intended to mean Lysogeny Broth (Bertani, 2004)). Its composition is as follows: 1% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract and 0.5% (w/v) NaCl with the pH adjusted to 7.2 with sodium hydroxide. Rich SOC medium was used after transformation by electroporation (2.2.3.4) and contains 2% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 0.06% (w/v) NaCl, 0.02% (w/v) KCl, 0.1% (w/v) $MgCl_2$, 0.12% (w/v) $MgSO_4$, 0.4% (w/v) glucose.

2.1.2.3 Yeast media

Table 2.1.2.3 Yeast media

Medium	Components
YPDA	1% (w/v) Yeast extract, 2% (w/v) Bacto-peptone, 2% (w/v) Glucose, 0.003% (w/v) Adenine sulphate
YPGalA	As YPDA, but 2% (w/v) galactose instead of glucose
YMM	0.67% (w/v) Yeast nitrogen base without amino acids, 2% (w/v) Glucose or 2% (w/v) galactose
SD- or SDGal-	YMM with 2 g/l of dropout powder (2.1.2.4)
5-FOA (*)	0.67% (w/v) Yeast nitrogen base without amino acids, 2% (w/v) Glucose or 2% (w/v) galactose, 0.2% (w/v) Ura drop-out powder, 0.005% (w/v) Uracil, 0.1% (w/v) 5-Fluoroorotic acid (Toronto Research Chemicals)
Pre-sporulation (minimal)	10% glucose or galactose, 0.2% appropriate dropout powder, 0.67 % nitrogen base without amino acids
Sporulation (minimal)	1% Potassium acetate, 0.005% Zinc acetate, 0.05% dropout powder or simply 2% Potassium acetate

(*) For each 100 ml of 5-FOA medium: 0.67 g of yeast nitrogen base without amino acids, 0.2 g of -Ura drop out powder, 5 mg of uracil, 0.1 g of 5-FOA and 10 ml of 20% glucose (or galactose) were dissolved in a total volume of 50 ml by gently heating (microwave) till about 37°C. This mixture was filter sterilised using a 0.2 μm filter (Minisart Cellulose acetate, Sartorius) and pre-heated to 55°C, mixed with a 55°C solution of autoclaved 2g agar in 50 ml of water. Resulting plates were stored at 4°C protected from the light.

2.1.2.4 Drop-out powder

Drop-out powder was prepared by adding and mixing 2 g of each of the following nutrients; adenine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, tryptophan, uracil and valine. In addition, 4 g of leucine was added to the drop-out powder when required. For a specific drop-out powder, the relevant nutrients were omitted. This mix was then ground with a mortar and pestle to ensure complete mixing and was added to yeast media to a final concentration of 0.2% (w/v) prior to autoclaving. To adjust the pH, one pellet of NaOH was added to each 500 ml of media containing drop-out powder.

2.1.2.5 Antibiotics

Antibiotics were added to liquid media immediately prior to use. Solid media were autoclaved and cooled down to approximately 50°C before adding the appropriate antibiotics. Ampicillin, kanamycin, streptomycin, chloramphenicol, tetracycline and nourseothrycin stock solutions (all at 1,000 times that of the required working concentration) were stored at -20°C, G418 was stored at room temperature as a powder and hygromycin stock solution was stored at 4°C. Those stock solutions made in water were filter sterilised using 0.2 μm filters (Minisart Cellulose acetate, Sartorius) before aliquoting and storing them at -20°C.

Table 2.1.2.5 Antibiotics

Antibiotic	Abbreviation	Solvent	Final concentration (μg/ml)
Ampicillin	Amp	water	100
Chloramphenicol	Cm	ethanol	34
Hygromycin	Hyg	PBS[†]	300
Kanamycin	Kan	water	30
G418 (Geneticin)	G418	N/A	200
Nourseothrycin	Nat	water	100
Streptomycin	Strep	water	50
Tetracycline	Tet	ethanol	5

[†] Dissolved in PBS (at 50 mg/ml) when purchased; 3 ml of this stock was added per 500 ml of medium.

2.1.3 Commonly used buffers

All commonly used buffers were made using autoclaved deionised water. Their compositions are indicated in Table 2.1.3 below.

Table 2.1.3 Commonly used buffers

Buffer	Components
50x TAE	2 M Tris-base, 5.71% (v/v) acetic acid, 50 mM EDTA
10x TE	100 mM Tris-HCl, 10 mM EDTA pH 8.0
10x TBE	0.89 M Tris base, 0.89 M boric acid, 20 mM EDTA
10x TBST	0.5 M Tris-HCl pH 7.5, 1.5 M NaCl, 1% (v/v) Tween-20 (USB Corporation)
20x MOPS SDS-PAGE running buffer	1 M MOPS, 1 M Tris base, 2% (w/v) SDS, 20 mM EDTA

Table 2.1.3 Commonly used buffers (continued)

Buffer	Components
10x Tris-Glycine SDS-PAGE running buffer	250 mM Tris base, 1.9 M Glycine, 1% (w/v) SDS
10x Western transfer buffer	200 mM Tris base, 1.5 M Glycine
10x BPTE	100 mM PIPES, 200 mM Bis-Tris, 10 mM EDTA pH 8.0, DEPC treated or filtered
20x SSC	3 M NaCl, 0.3 M Sodium citrate (pH adjusted to 7.0 with 1 M HCl)
20x SSPE	3.6 M NaCl, 0.2 M NaH_2PO_4, 20 mM EDTA (pH adjusted to 7.4 with NaOH)
Buffer B	0.1 M Potassium phosphate pH 7.4, 1.2 M Sorbitol
SES1	0.5 M Sodium phosphate pH 7.2, 7% (w/v) SDS, 1 mM EDTA
NT and NTN	50 mM Tris-HCl pH7.5, 150 mM NaCl For NTN only: 0.1% (v/v) NP-40 (Roche)
2x IP_{150}	12 mM HEPES-KOH pH 7.9, 300 mM NaCl, 5 mM $MgCl_2$ and 0.1% (v/v) NP-40 (Roche)

2.1.4 Sterilising solutions and making solutions RNase free

To sterilise solutions they were commonly autoclaved (Thermo Denley, Diamond 36, 10 min at 121°C). For those that could not be autoclaved but for which sterility was required solutions were filter-sterilised using disposable filter units with PES filters (Polyethersulphone, Nalgene). For those solutions that needed to be RNase free similar disposable filter units were used with CN filters (Cellulose nitrate, Nalgene). Alternatively aqueous solutions were made RNase free by DEPC treatment. DEPC (Diethylpyrocarbonate) was added to a concentration of 0.1% (v/v), the solution was incubated for a minimum of 1 h at 37°C and autoclaved to hydrolyse the DEPC.

2.1.5 *Escherichia coli* strains

The bacterial strains used throughout this work are listed in Table 2.1.5 below. DH5α and TOP10 cells were used for cloning and transformation of plasmid DNA, and subsequent propagation. All other strains were used for production of recombinant proteins and/or were used in making *hfq* deletion strains (2.2.4).

Table 2.1.5 *Escherichia coli* strains

Strain	Genotype	Source
DH5α	*F⁻, φ80dlacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(rk⁻, mk⁺), phoA, supE44, λ⁻, thi-1, gyrA96, relA1*	GIBCO BRL
TOP10	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR nupG recA1 araD139 Δ(ara-leu)7697 galU galK rpsL(Str^R) endA1 λ⁻*	Invitrogen
BL21(DE3)	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3)*	Novagen
BL21-CodonPlus (DE3)-RIL	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3) endA Hte [argU ileYleuW Cam^R]*	Novagen
C41(DE3)	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3) and an uncharacterised mutation* (Miroux and Walker, 1996)	Avidis
ER2566	*F⁻ λ⁻ fhuA2 [lon] ompT lacZ::T7 gene1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm]*	NEB
MG1655	*F⁻ λ⁻ ilvG- rfb-50 rph-1*	David Leach
MC4100	*F⁻ araD139 Δ(argF-lac)U169 rspL150 relA1 flbB5301 fruA25 deoC1 ptsF25 e14⁻*	n/a
MRE1 (MG1655 *Δhfq*)	*F⁻ λ⁻ ilvG- rfb-50 rph-1 Δhfq*	This thesis
MRE2 (ER2566 *hfq-1*)	*F⁻ λ⁻ fhuA2 [lon] ompT lacZ::T7 gene1 gal sulA11 D(mcrC-mrr)114::IS10 R(mcr-73::miniTn10TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm] hfq-1::Ω (Kan^r, BclI)*	Poul Valentin-Hansen

Table 2.1.5 *Escherichia coli* strains (continued)

Strain	Genotype	Source
MRE3 (C41 *hfq-1*)	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3) hfq-1:: Ω (Kan^r, BclI) and an uncharacterised mutation* (Miroux *et al.*, 1996)	Poul Valentin-Hansen
MRE4 (ER2566 *Δhfq*)	*F⁻ λ⁻ fhuA2 [lon] ompT lacZ::T7 gene1 gal sulA11 D(mcrC-mrr)114::IS10 R(mcr-73::miniTn10TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm] Δhfq*	This thesis
MRE5 (C41 *Δhfq*)	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3) Δhfq and an uncharacterised mutation* (Miroux *et al.*, 1996)	This thesis
MRE6 (GS081)	*F⁻ araD139 D(argF-lac)U169 rspL150 relA1 flbB5301 fruA25 deoC1 ptsF25 e14⁻ hfq-1::Ω (Cm^r, BclI)*	Gisella Storz

2.1.6 *Saccharomyces cerevisiae* strains

The strains used in and created for this study are listed below in Tables 2.1.6.1 and 2.1.6.2. The parent strains of those that were created for this thesis are indicated. A source and/or references are given for all other strains.

Table 2.1.6.1 *Saccharomyces cerevisiae* strains created for this thesis

Strain	Genotype	Derived from
MRY1	*MAT**a** ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pBS1191]*	BSY557
MRY2	*MAT**a** ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm7Δ::KANMX4 [pBS1191]*	BSY557
MRY3	*MAT**a** ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 [pSX6]*	BSY557
MRY4	*MAT**a** ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pSX6]*	MRY1
MRY5	*MAT**a** ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm7Δ::KANMX4 [pSX6]*	MRY2
MRY6	*MAT**a** ade2-1 his3-11,-15 leu2-3, -112 trp1Δ ura3-1 lsm5Δ::TRP1*	AEMY48
MRY7	*MAT**a** ade2-1 his3-11,-15 leu2-3, -112 trp1Δ ura3-1 lsm5Δ::TRP1 [pMR9]*	AEMY48

Table 2.1.6.1 *Saccharomyces cerevisiae* strains created for this thesis (continued)

Strain	Genotype	Derived from
MRY8	*MATa ade2-1 his3-11,-15 leu2-3, -112 trp1Δ ura3-1 lsm5Δ::TRP1 [pMR10]*	AEMY48
MRY9	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pSR53]*	GLS618
MRY10	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pSR70]*	GLS618
MRY11	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pSR86]*	GLS618
MRY12	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pSR93]*	GLS618
MRY13	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pMR60]*	GLS618
MRY14	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pSR151]*	GLS618
MRY15	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm6Δ::HpHMX3 [pPR097]*	GLS618
MRY16	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm7Δ::NATMX6 [pPR097]*	GLS618
MRY17	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ1 lsm6Δ::HpHMX6*	BY4741
MRY18	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ1 lsm7Δ::HpHMX6*	BY4741
MRY21	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 [pMR61]*	BSY557
MRY22	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 [pMR62]*	BSY557
MRY23	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pMR61]*	MRY1
MRY24	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pMR62]*	MRY1
MRY25	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm7Δ::KANMX4 [pMR61]*	MRY2
MRY26	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm7Δ::KANMX4 [pMR62]*	MRY2
MRY27	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ1 lsm5Δ::HpHMX6*	BY4741
MRY28	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm6Δ::HpHMX3 [pMR63]*	MRY15
MRY29	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm6Δ::HpHMX3 [pSR86]*	MRY15
MRY30	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm6Δ::HpHMX3 [pSR151]*	MRY15
MRY31	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm7Δ::NATMX6 [pMR63]*	MRY16
MRY32	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm7Δ::NATMX6 [pSR86]*	MRY16

Table 2.1.6.1 *Saccharomyces cerevisiae* strains created for this thesis (continued)

Strain	Genotype	Derived from
MRY33	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm7Δ::NATMX6 [pSR151]*	MRY16
MRY34	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm6Δ::HpHMX3 [pMR59]*	MRY15
MRY35	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN lsm7Δ::NATMX6 [pMR59]*	MRY16
MRY36	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pMR65]*	GLS618
MRY37	*MATα his3Δ1 leu2Δ0 lys2Δ0 met15Δ0 ura3Δ1 prp24Δ::KAN [pMR66]*	GLS618
MRY61	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 [pMR109]*	BSY557
MRY62	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 [pMR64]*	BSY557
MRY63	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pMR109]*	MRY1
MRY64	*MATa ade2 arg4 leu2-3,-112 trp1-289 ura3-52 SNR6::ADE2 lsm6Δ::KANMX4 [pMR64]*	MRY2
MRY65	*MATa ade1-101 his3Δ1 trp1-289 ura3-52 LEU2-GAL1-LSM4 lsm8Δ::HIS3 GAL$^+$ [pMR63]*	MCY4
MRY68	*MATa ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL$^+$ [pMR157]*	MCY5
MRY71	*MATa ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL$^+$ [pUSS1]*	MRY68
MRY72	*MATa ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL$^+$ [pMC11]*	MRY68
MRY73	*MATa ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL$^+$ [pMC12]*	MRY68

Table 2.1.6.2 Other *Saccharomyces cerevisiae* strains

Strain	Genotype	Origin
BMA38α	*MATα ade2-1 his3Δ200 leu2-3,-112 trp1Δ1 ura3-1 can1-100*	Albers, 1999
BMA38a	*MATa ade2-1 his3Δ200 leu2-3,-112 trp1Δ1 ura3-1 can1-100*	Albers, 1999
AEMY25	*MATa ade2-1 his3-11,-15 leu2-3,-112 trp1Δ1 ura3-1 lsm1Δ::TRP1*	Mayes *et al.*, 1999
AEMY48	*MATa ade2-1 his3-11,-15 leu2-3,-112 trp1Δ1 ura3-1 lsm5Δ::TRP1 [pBM125-GAL1-HA-LSM5]*	Andrew Mayes, this lab
MCY4	*MATa ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 LEU2-GAL1-LSM4 GAL$^+$*	Cooper *et al.*, 1995

Table 2.1.6.2 Other *Saccharomyces cerevisiae* strains (continued)

Strain	Genotype	Origin
MCY5	*MAT* a/α *ADE/ade1-101 HIS3/his3-1 TRP1/trp1-289 ura3-52/ura3-52 leu2-3,112/leu2-3,112 LSM4/lsm4Δ::LEU2, GAL*$^+$	Cooper, 1995
LMA4	*MAT* a/α *ade2-1/ad21 his3-11,-15/his3-11,-15 trp1Δ1/ trp1Δ1 ura3-1/ura3-1 leu2-3,-112/leu2-3,-112 LSM8/ lsm8Δ::TRP1*	Mayes, 1998
BSY557	*MAT*a *ade2 arg4 leu2-3, 112 trp1-289 ura3-52 snr6Δ::ADE2 pBS1191 (CEN-URA3-SNR6)*	Luukkonen and Séraphin, 1998
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ ura3Δ0*	Euroscarf; Brachmann *et al.*, 1998
BY4742	*MAT*α *his3Δ1 leu2Δ0 lys2Δ ura3Δ0*	Euroscarf; Brachmann *et al.*, 1998
BY4743	*MAT* a/α *his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 MET15/met15Δ0 ura3Δ0/ura3Δ0*	Euroscarf; Brachmann *et al.*, 1998
GLS618	*MAT*α *his3Δ leu2Δ0 lys2Δ0 met15Δ ura3Δ01 prp24Δ::KAN [pPR097]*	Rader *et al.*, 2002
GLS619	*MAT*α *his3Δ leu2Δ0 lys2Δ0 met15Δ ura3Δ01 prp24Δ::KAN [pPR113]*	Rader *et al.*, 2002
GLS620	*MAT*α *his3Δ leu2Δ0 lys2Δ0 met15Δ ura3Δ01 prp24Δ::KAN [pSR039]*	Rader *et al.*, 2002
Y01301	*MAT*a *his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 lsm1Δ::kanMX4*	Euroscarf
MPS11	*MAT*a *ade2-1 his3-11,-15 leu2-3,-112 trp1Δ1 ura3-1 can1-100 lsm8::TRP1 [pHIS-GAL1-HA-LSM8] LSM7:13myc-HphMX6*	Michael Spiller, this lab
MPS26	*MAT*a *ade2-1 his3-11,-15 leu2-3,-112 trp1Δ1 ura3-1 can1-100 lsm8::TRP1 [pYX172] LSM7:13myc-HphMX6*	Michael Spiller, this lab
L119-7D	*MAT*α *ura3-52 trp1/2 ade1 dbf2-1*	Parkes and Johnston, 1992
AMy1292	*MAT*a *ura3 leu2 ade2 his3 trp1 dbf2-2*	Adele Marston
AMy1277	*MAT*a *ura3 leu2-3 ade1 trp1-1 omns cdc15-2*	Adele Marston
CY3	*MAT*a *ura3 lys2 ade2 trp1 his3 leu2*	Yoo and Wolin, 1997
BP4	*MAT*a *lsm8-1 LHP1 ura3 lys2 ade2 trp1 his3 leu2*	Pannone *et al.*, 1998
KY117	*MAT*a *his3Δ200 lys2-801 ade2 trp1Δ1 ura3-52 GAL*$^+$	Chen and Struhl, 1985
KY118	*MAT*α *his3Δ200 lys2-801 ade2 trp1Δ1 ura3-52 GAL*$^+$	Chen *et al.*, 1985

2.1.7 Plasmids

Plasmids used throughout this study are listed in Tables 2.1.7.1 – 2.1.7.5 below. Those that were created for this work were constructed using standard cloning techniques and oligonucleotides described in Tables 2.1.8.1a and b.

Table 2.1.7.1 *E. coli* expression vectors

Plasmid	Description	Source
pET19b	AmpR *E. coli* expression vector with 10x EK-cleavable N-terminal His-tag from T7 promoter	Novagen
pET28b	AmpR *E. coli* expression vector with N-terminally 6xHis tagging MCS controlled by T7 promoter	Novagen
pCDFDuet-1	StrepR *E. coli* expression vector with two MCS, one with 6x N-terminal His-tag, second with C-terminal S-tag both from T7 promoter	Novagen
pETDuet-1	AmpR *E. coli* expression vector with two MCS, one with 6x N-terminal His-tag, second with C-terminal S-tag both from T7 promoter	Novagen
pCOLADuet-1	KanR *E. coli* expression vector with two MCS, one with 6x N-terminal His-tag, second with C-terminal S-tag both from T7 promoter	Novagen
pACYCDuet-1	CmR *E. coli* expression vector with two MCS, one with 6x N-terminal His-tag, second with C-terminal S-tag both from T7 promoter	Novagen
pIP24	AmpR *E. coli* expression vector with N-terminally 6xHis tagged PRP24 under the lac promoter	Vidal *et al.*, 1999
pET28b Hfq Y55A	*E. coli hfq* Y55A in frame with C-terminal 6xHis tag in pET28b expression vector, gift from A.Feig	Mikulecky *et al.*, 2004

Table 2.1.7.2 *E. coli* expression vectors created for this thesis

Plasmid	Alias	Description
pMR81	pET19 HA	single HA sequence introduced into *Nde*I site of pET19b, *Nde*I site remains only downstream of HA-tagging sequence
pMR82	pET19 Myc	single c-Myc sequence introduced into *Nde*I site of pET19b, *Nde*I site remains only downstream of HA-tagging sequence
pMR2	pET19b Lsm2	*LSM2* (intronless) cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR3	pET19b Lsm3	*LSM3* cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR4	pET19b Lsm4	*LSM4* cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR5	pET19b Lsm5	*LSM5* cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR6	pET19b Lsm6	*LSM6* cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR7	pET19b Lsm7	*LSM7* (intronless) cloned in frame with N-terminal 10xHis tag of pET19b expression vector; *Nde*I-*Bam*HI
pMR18	pET19 HA Lsm8	*LSM8* (correct, shorter ORF) cloned in frame with N-terminal 10xHis+HA tag of pET19 HA expression vector; *Nde*I-*Bam*HI
pMR19	pET19 HA Prp24	PRP24 cloned in frame with N-terminal 10xHis+HA tag of pET19 HA expression vector; *Nde*I-*Bam*HI
pMR20	pET19 HA Prp24Δ10	Prp24 with 10AA C-terminal deletion cloned in frame with N-terminal 10xHis+HA tag of pET19 HA expression vector; *Nde*I-*Bam*HI
pMR21	pET19 Myc Lsm1	*LSM1* cloned in frame with N-terminal 10xHis+Myc tag of pET19 Myc expression vector; *Nde*I-*Bam*HI
pMR29	pET19 Myc Prp24	PRP24 cloned in frame with N-terminal 10xHis+Myc tag of pET19 Myc expression vector; *Nde*I-*Bam*HI
pMR30	pET19 Myc Prp24Δ10	Prp24 with 10AA C-terminal deletion cloned in frame with N-terminal 10xHis+Myc tag of pET19 Myc expression vector; *Nde*I-*Bam*HI
pMR33	pET19 HA Lsm8 (old)	*LSM8* (incorrect, longer ORF) cloned in frame with N-terminal 10xHis+HA tag of pET19 HA expression vector; *Nde*I-*Bam*HI
pMR39	pCDFDuet Lsm1/Lsm4	*LSM1* cloned in frame with N-terminal 6xHis-tag of pMR35; *Bam*HI-*Pst*I

Table 2.1.7.2 *E. coli* expression vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR40	pETDuet Lsm7(His)	*LSM7* (intronless) cloned in frame with N-terminal 6xHis-tag of pETDuet-1 expression vector; *Bam*HI-PstI
pMR41	pETDuet Lsm7(S)	*LSM7* (intronless) cloned in frame with C-terminal S-tag of pETDuet-1 expression vector; *Nde*I-*Xho*I
pMR46	pCOLADuet Lsm3/Lsm2	*LSM3* cloned in frame with N-terminal 6xHis-tag of pMR45; *Bam*HI-PstI
pMR47	pACYCDuet Lsm5/Lsm6	*LSM5* cloned in frame with N-terminal 6xHis-tag of pMR42; *Bam*HI-PstI
pMR48	pCDFDuet Lsm8/Lsm4 (new)	*LSM8* (correct, shorter ORF) cloned in frame with N-terminal 6xHis-tag of pMR35; *Bam*HI-PstI
pMR52	pET19 HA Hfq	*E. coli hfq* cloned in frame with N-terminal10xHis+HA tag of pET19 HA expression vector; *Nde*I-*Bam*HI
pMR53	pET19 Myc Hfq	*E. coli hfq* cloned in frame with N-terminal 10xHis+Myc tag of pET19 Myc expression vector; *Nde*I-*Bam*HI
pMR107	pET19 Myc HfqΔC83	*E. coli hfq* (up to aa83) cloned in frame with N-terminal 10xHis+Myc tag of pET19 Myc expression vector; *Nde*I-*Bam*HI

Table 2.1.7.3 Yeast shuttle vectors

Plasmid	Description	Source
pBM125	Yeast *E. coli* shuttle vector: *CEN4*, *ARS1*, *URA3*, AmpR, *GAL1* promoter	Johnston and Davis, 1984
pGFP-N-FUS	*CEN6*, *URA3* shuttle vector with MCS for N-terminal GFP tagging under control of *MET25* promoter	Niedenthal *et al.*, 1996
pGFP-C-FUS	*CEN6*, *URA3* shuttle vector with MCS for C-terminal GFP tagging under control of *MET25* promoter	Niedenthal *et al.*, 1996
pRS313	*CEN6*, *HIS3* shuttle vector	Sikorski and Hieter, 1989
pRS314	*CEN6*, *TRP1* shuttle vector	Sikorski and Hieter, 1989
pRS315	*CEN6*, *LEU2* shuttle vector	Sikorski and Hieter, 1989
pRS316	*CEN6*, *URA3* shuttle vector	Sikorski and Hieter, 1989

Table 2.1.7.3 Yeast shuttle vectors (continued)

Plasmid	Description	Source
pRS423-pMet25	381bp of pMET25 (2785-2404) in pRS423 (2 μ, *HIS3*) shuttle vector	R. Grainger, this lab
YEp24	2μ, *URA3* shuttle vector; routinely used in high-copy screens in yeast	Parent *et al.*, 1985
pSX6	pSE358 (*CEN4*, *TRP1*) with *SNR6* and flanking sequences, *Sph*I site and *Xho*I sites were created at 5' and 3' ends respectively of the U6 gene, gift from C. Guthrie	Madhani *et al.*, 1990
pBS1191	Contains *SNR6* with its own regulatory sequences (*CEN*, *URA3*)	Luukkonen *et al.*, 1998
pYX172	2μ, *LEU2* plasmid with *SNR6* and 1kb up and downstream	Hu *et al.*, 1994
pSR53	pSE362 (*CEN4*, *HIS3*) with *prp24-RRM3sub* and its own regulatory sequences, gift from S. Rader	Rader *et al.*, 2002
pSR70	pSE362 (*CEN4*, *HIS3*) with *prp24-RRM4sub* and its own regulatory sequences, gift from S. Rader	Rader *et al.*, 2002
pSR86	pSE362 (*CEN4*, *HIS3*) with *prp24-RRM1sub* and its own regulatory sequences, gift from S. Rader	Rader *et al.*, 2002
pSR93	pSE362 (*CEN4*, *HIS3*) with *prp24-RRM1sub/RRM4sub* and its own regulatory sequences, gift from S. Rader	Rader *et al.*, 2002
pSR151	pSE362 (*CEN4*, *HIS3*) with *prp24Δ10/RRM1sub* and its own regulatory sequences, gift from S. Rader	Rader *et al.*, 2002
pGFP-N-LSM1	*LSM1* in frame with N-terminal GFP tag in pGFP-N-FUS	M. Spiller, this lab
pGFP-N-LSM8	*LSM8* in frame with N-terminal GFP tag in pGFP-N-FUS (pMPS8)	Spiller, 2005
pMPSLsm4	pGFP-N-FUS, with *LSM4* coding region inserted in frame with *GFP* coding region	Spiller, 2005
pMPSLsm4D1	pGFP-N-FUS, with *LSM4* domain I coding region (aa 1-93) inserted in frame with *GFP* coding region (add IDTVDLESCN at C-term.)	Spiller, 2005
pMPSLsm4D2	pGFP-N-FUS, with *LSM4* domain II coding region (aa 92-187) inserted in frame with *GFP* coding region	Spiller, 2005
pMPSLsm2	pGFP-N-FUS, with *LSM2* coding region inserted in frame with *GFP* coding region	Spiller, 2005
pGFP-N-Lsm6	pGFP-N-FUS, with *LSM6* coding region inserted in frame with *GFP* coding region	Michael Spiller, this lab

Table 2.1.7.3 Yeast shuttle vectors (continued)

Plasmid	Description	Source
pGFP-N-Lsm7	pGFP-N-FUS, with *LSM7* coding region inserted in frame with *GFP* coding region	Michael Spiller, this lab
pMPS2	pGFP-N-FUS, with *LHP1* coding region inserted in frame with *GFP* coding region	Spiller, 2005
pUSS1	1.7 kb *Hind*III genomic fragment with *LSM4* ORF, promoter and terminator sequences cloned into pFL39 (*CEN6*, *TRP1*)	Cooper, 1995
pMC11	Modified pFL39 (*CEN6*, *TRP1*) with a 1.49 kb *Hind*III/*Sal*I fragment containing truncated *LSM4* (aa 1-136 + CRLVS) with own promoter	Cooper, 1995
pMC12	Modified pFL39 (*CEN6*, *TRP1*) with a 1.36 kb HindII/SalI fragment containing truncated *LSM4* (aa 1-92 + VD) with its own promoter	Cooper, 1995
pRP1155	*DCP2* with C-terminal RFP tag in *CEN*, *LEU* shuttle vector, gift from R.Parker	Teixeira *et al.*, 2005

Table 2.1.7.4 Yeast shuttle vectors created for this thesis

Plasmid	Alias	Description
pMR9	pRS313 LSM5	*LSM5* with flanking genomic DNA cloned into pRS313 (*CEN6*, *HIS3*)
pMR10	pRS313 LSM5 E10K	*LSM5* with E10K mutation and flanking genomic DNA cloned into pRS313 (*CEN6*, *HI3S*)
pMR55	pSR105 = pSR53+RRM4sub	re-made pSR105 (RRM3/4sub; S. Rader) by making RRM4sub mutation in pSR53
pMR58	pPR097	pSE360 (*CEN4*, *URA3*) with *PRP24* rescued from GLS618 (S. Rader)
pMR59	pPR113	pSE362 (*CEN4*, *HIS3*) with *PRP24* rescued from GLS619 (S. Rader)
pMR60	pMR59 F87A	F87A mutation in *PRP24* in pSE362
pMR63	pSR39	pSE362 (*CEN4*, *HIS3*) with *prp24Δ10* rescued from GLS620 (S. Rader)
pMR65	pMR59 R158S	R158S mutation in *PRP24* in pSE362
pMR66	pMR59 F257I	F257I mutation in *PRP24* in pSE362
pMR61	pSX6 G63A	G63A mutation made in *SNR6* in pSX6 (Guthrie-lab)

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR62	pSX6 C69A	C69A mutation made in *SNR6* in pSX6 (Guthrie-lab)
pMR64	pSX6 G63A/A82G	G63A mutation made (and spontaneous A82G mutation introduced) in *SNR6* in pSX6 (Guthrie-lab)
pMR109	pSX6 G63A/C69A	G63A mutation introduced in *SNR6* in pMR58
pMR69	pGFP-N-Lsm118	Hybrid of *LSM1* (N-term&Sm domains; up to D121) and *LSM8* (C-term; from S74) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *Xba*I/*Xho*I
pMR70	pGFP-N-Lsm881	Hybrid of *LSM8* (N-term&Sm domains; up to D73) and *LSM1* (C-term; from Q122) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *Xba*I/*Xho*I
pMR71	pBM125 Lsm118	Hybrid of *LSM1* (N-term&Sm domains; up to D121) and *LSM8* (C-term; from S74) cloned in pBM125 under GAL Promoter; *Bam*HI
pMR72	pBM125 Lsm881	Hybrid of *LSM8* (N-term&Sm domains; up to D73) and *LSM1* (C-term; from Q122) cloned in pBM125 under GAL Promoter; *Bam*HI
pMR73	pGFP-N-Lsm881 rna1	*rna1* (N28A; D31A) mutation introduced into pMR70
pMR74	pGFP-N-Lsm881 rna2	*rna2* (T34A; N35A) mutation introduced into pMR70
pMR75	pGFP-N-Lsm881 rna3	*rna3* (R57A; G58W; S59A) mutation introduced into pMR70
pMR76	pGFP-N-Lsm8 rna1	*rna1* (N28A; D31A) mutation introduced into pGFP-N-Lsm8
pMR77	pGFP-N-Lsm8 rna2	*rna2* (T34A; N35A) mutation introduced into pGFP-N-Lsm8
pMR78	pGFP-N-Lsm8 rna3	*rna3* (R57A; G58W; S59A) mutation introduced into pGFP-N-Lsm8
pMR79	pGFP-N-Lsm1ΔC	*LSM1* with C-terminal deletion after D121 in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *Xba*I/*Xho*I
pMR80	pGFP-N-Lsm8ΔC	*LSM8* with C-terminal deletion after D73 in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *Xba*I/*Xho*I
pMR83	pGFP-C-Lsm8	*LSM8* (correct, shorter ORF) cloned in frame with C-terminal GFP-tag under *MET25* promoter; *Xba*I/*Cla*I
pMR84	pGFP-C-Lsm881	*lsm881* cloned in frame with C-terminal GFP-tag under *MET25* promoter; *Xba*I/*Cla*I
pMR85	pGFP-C-Lsm1ΔC (ClaI)	*LSM1* up to internal *Cla*I site (up to I117) cloned in frame with C-terminal GFP-tag under *MET25* promoter; *Xba*I/*Cla*I

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR86	pGFP-C-Lsm1	*LSM1* (partial digest) cloned in frame with C-terminal GFP-tag under *MET25* promoter; *XbaI*/*ClaI*
pMR87	pGFP-C-Lsm118	*lsm118* (partial digest) cloned in frame with C-terminal GFP-tag under *MET25* promoter; *XbaI*/*ClaI*
pMR88	pMET25-Lsm1	*LSM1* cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR89	pMET25-Lsm118	*Lsm118* cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR90	pMET25-Lsm8	*LSM8* (correct, shorter ORF) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR91	pMET25-Lsm881	*lsm881* cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR92	pGFP-C-Lsm8 rna1	*lsm8 rna1* (N28A; D31A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR93	pGFP-C-Lsm8 rna2	*lsm8 rna2* (T34A; N35A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR94	pGFP-C-Lsm8 rna3	*lsm8 rna3* (R57A; G58W; S59A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR95	pGFP-C-Lsm881 rna1	*lsm881 rna1* (N28A; D31A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR96	pGFP-C-Lsm881 rna2	*lsm881 rna2* (T34A; N35A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR97	pGFP-C-Lsm881 rna3	*lsm881 rna3* (R57A; G58W; S59A) cloned in frame with C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*ClaI*
pMR98	pMET25-Lsm8 rna1	*lsm8 rna1* (N28A; D31A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR99	pMET25-Lsm8 rna2	*lsm8 rna2* (T34A; N35A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR100	pMET25-Lsm8 rna3	*lsm8 rna3* (R57A; G58W; S59A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR101	pMET25-Lsm881 rna1	*lsm881 rna1* (N28A; D31A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR102	pMET25-Lsm881 rna2	*lsm881 rna2* (T34A; N35A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR103	pMET25-Lsm881 rna3	*lsm881 rna3* (R57A; G58W; S59A) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR104	pGFP-C-Lsm8ΔC	*lsm8ΔC* (up to D73) cloned in frame with C-terminal GFP-tag under *MET25* promoter; *XbaI*/*Cla*I
pMR105	pMET25-Lsm8ΔC	*lsm8ΔC* (up to D73) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *XbaI*/*Bam*HI
pMR110	pGFP-N-Lsm8ΔC NLS	*lsm8ΔC* (up to D73) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter with SV40 NLS fused to C-terminus (PKKKRKV); *XbaI*/*Xho*I
pMR111	pGFP-C-Lsm8ΔC NLS	*lsm8ΔC* (up to D73) cloned in frame with C-terminal GFP-tag under *MET25* promoter with SV40 NLS in between C-terminus (PKKKRKV); *XbaI*/*Cla*I
pMR112	pMET25-Lsm8ΔC NLS	*lsm8ΔC* (up to D73) cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS with SV40 NLS fused to C-terminus (PKKKRKV); *XbaI*/*Bam*HI
pMR113	pMET25-Lsm8 rna3 NLS	*lsm8 rna-3* cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS with SV40 NLS fused to C-terminus (PKKKRKV); *XbaI*/*Bam*HI
pMR114	pGFP-N-Lsm181	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*Xho*I
pMR115	pGFP-N-Lsm811	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains and C-terminal (aa52-172) extension cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*Xho*I
pMR116	pGFP-N-Lsm818	Hybrid of *LSM8* N-terminal extension (aa1-10), *LSM1* Sm domains (aa52-121) and *LSM8* C-terminal extension (aa74-109) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*Xho*I

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR117	pGFP-C-Lsm188	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains and C-terminal extension (aa11-109) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR118	pGFP-C-Lsm81ΔC (ClaI)	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains up to internal *ClaI* site (aa52-118) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR119	pMET25-Lsm181	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172), *LSM8* Sm domains and *LSM1* C-terminal extension cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR120	pMET25-Lsm811	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains and C-terminal (aa52-172) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR121	pMET25-Lsm818	Hybrid of *LSM8* N-terminal extension (aa1-10), *LSM1* Sm domains (aa52-121) and *LSM8* C-terminal extension (aa74-109) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR122	pMET25-Lsm188	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains and C-terminal extension (aa11-109) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR123	pGFP-C-Lsm181	Hybrid of *LSM1* N-terminal extension(aa1-51), *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR124	pGFP-C-Lsm811	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains and C-terminal (aa52-172) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/Bam*HI
pMR125	pGFP-C-Lsm818	Hybrid of *LSM8* N-terminal extension (aa1-10), *LSM1* Sm domains (aa52-121) and *LSM8* C-terminal extension (aa74-109) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/Bam*HI
pMR126	pGFP-N-Lsm188	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains and C-terminal extension (aa11-109) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI/Xho*I

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR129	pGFP-N-lsm81ΔC	Hybrid of *LSM8* N-terminal extension (aa1-10) and *LSM1* Sm domains (aa52-121) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*XhoI*
pMR130	pGFP-C-lsm81ΔC	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI*/*Bam*HI
pMR131	pMET25-lsm81ΔC	Hybrid of *LSM8* N-terminal extension (aa1-10) with *LSM1* Sm-domains loned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI*/*Bam*HI
pMR132	pGFP-C-lsm8N	*LSM8* N-terminal extension (aa1-10) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI*/*ClaI*
pMR133	pGFP-N-lsm1C	*LSM1* N-terminal extension (aa1-51)cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*XhoI*
pMR134	pGFP-N-lsmΔN11	*LSM1* Sm-domains and C-terminal (aa52-172) extension cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*XhoI*
pMR135	pGFP-C-lsmΔN11	*LSM1* Sm-domains and C-terminal (aa52-172) extension cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI*/*Bam*HI
pMR136	pMET25-lsmΔN11	*LSM1* Sm-domains and C-terminal (aa52-172) extension cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI*/*Bam*HI
pMR137	pGFP-N-lsm18ΔC	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains (aa11-73) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*XhoI*
pMR138	pGFP-C-lsm18ΔC	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI*/*ClaI*
pMR139	pMET25-lsm18ΔC	Hybrid of *LSM1* N-terminal extension (aa1-51), *LSM8* Sm domains cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI*/*Bam*HI
pMR140	pGFP-N-lsmΔN88	*LSM8* Sm domains and C-terminal extension (aa11-109) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI*/*XhoI*

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR141	pGFP-C-lsmΔN88	*LSM8* Sm domains and C-terminal extension (aa11-109) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR142	pMET25-lsmΔN88	*LSM8* Sm domains and C-terminal extension (aa11-109) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR143	pGFP-N-lsmΔN18	Hybrid of *LSM1* Sm domains (aa52-D121) and *LSM8* (C-term; from S74) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI/XhoI*
pMR144	pGFP-C-lsm1N	*LSM1* N-terminal extension(aa1-51) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR145	pGFP-C-lsmΔN18	Hybrid of *LSM1* Sm domains (aa52-D121) and *LSM8* (C-term; from S74) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/Bam*HI
pMR146	pMET25-lsmΔN18	Hybrid of *LSM1* Sm domains (aa52-D121) and *LSM8* (C-term; from S74) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR147	pGFP-N-lsmΔN81	Hybrid of *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI/XhoI*
pMR148	pGFP-C-lsmΔN81	Hybrid of *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/ClaI*
pMR149	pMET25-lsmΔN81	Hybrid of *LSM8* Sm domains (aa11-73) and *LSM1* C-terminal extension (aa122-172) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR150	pGFP-N-lsmΔN1ΔC	*LSM1* Sm-domains (aa52-121) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI/XhoI*
pMR151	pGFP-C-lsmΔN1ΔC	*LSM1* Sm-domains (aa52-121) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *XbaI/Bam*HI
pMR152	pMET25-lsmΔN1ΔC	*LSM1* Sm-domains (aa52-121) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *XbaI/Bam*HI
pMR153	pGFP-N-lsmΔN8ΔC	*LSM8* Sm domains (aa11-73) cloned in frame with N-terminal GFP-tag of pGFP-N-FUS under *MET25* Promoter; *XbaI/XhoI*

Table 2.1.7.4 Yeast shuttle vectors created for this thesis (continued)

Plasmid	Alias	Description
pMR154	pGFP-C-lsmΔN8ΔC	*LSM8* Sm domains (aa11-73) cloned in frame with C-terminal GFP-tag of pGFP-C-FUS under *MET25* Promoter; *Xba*I/*Cla*I
pMR155	pMET25-lsmΔN8ΔC	*LSM8* Sm domains (aa11-73) cloned in pGFP-C-FUS without GFP-tag under *MET25* Promoter; *Xba*I/*Bam*HI
pMR156	pGFP-N-lsm8C	*LSM8* C-terminus (aa65-end) cloned with 140bp of 3'UTR in pGFP-N-FUS with N-terminal GFP-tag under *MET25* Promoter; *Xba*I/*Xho*I; made by Michael Spiller, rescued by me
pMR127	pRS423-pMet25-Hfq	*E. coli hfq* cloned behind *MET25* promoter in this 2μ *HIS* plasmid made by Richard Grainger; *Nde*I-*Xho*I (into *Bam*HI-*Xho*I sites)
pMR128	pGFP-N-Hfq	*E. coli hfq* cloned in frame with N-terminal GFP behind *MET25* promoter in pGFP-N-FUS;*Xba*I-*Xho*I
pMR157	pBM125-Lsm4	*LSM4* ORF and terminator sequence cloned into pBM125; *Bam*HI/*Sph*I
pMR158	pBM125-Lsm4D2	*LSM4* C-terminal domain (aa92-187) and terminator sequence cloned into pBM125; *Bam*HI/*Sph*I
pMR159	pRS314 DCP2-RFP	*DCP2-RFP* with promoter and terminator cloned from pRP1155 into pRS314; *Sal*I/*Sac*I
pMR160	pRS313 P_{GAL}-HA-Lsm8	*HA-LSM8* cloned from pAEMY76 into pRS313 by Michael Spiller; *Cla*I/*Xho*I; rescued from MPS11
pMR161	pMET25-Lsm4	*LSM4* ORF cloned without C-terminal GFP-tag under *MET25* promoter into pGFP-C-FUS; *Xba*I/*Bam*HI
pMR163	pRS315 PMet-Lsm8-RFP	P*MET25-LSM8* amplified by PCR from pMR90 cloned into pRP1155 (without *DCP2*+promoter); *Sal*I/*Bam*HI
pMR165	YEp24-SDB23	3.2kb *Bam*HI genomic *SDB23* fragment re-cloned into YEp24; *Bam*HI
pMR168	YEp24-sdb23 lsm4ATG-stop	Site-directed mutagenesis of *LSM4* start ATG to TAG in pMR165
pMR171	pRS313 DCP2-RFP	Promoter-*DCP2-RFP*-PGK1 terminator from pRP1155 cloned into pRS313; *Sal*I/*Sac*I

Table 2.1.7.5 Other vectors

Plasmid	Description	Source
pBR322	one of the most commonly used *E. coli* cloning vectors, here used to make radio labeled marker	Watson, 1988
pTOF24	KanR, CmR plasmid with SacB (gives sucrose sensitive growth) and temperature-sensitive replication (grow at 30°C), gift from D. Leach	D. Leach
pMR54	pTOF24 Δhfq = deletion construct with *hfq* flanking regions cloned into pTOF24; *PstI/SalI*	This thesis
pU4	Plasmid containing the U4 snRNA coding sequence (*SNR14*); used for PCR amplification of T7U4 for *in vitro* transcription	This lab
pT7U6	Modified pUC9 containing the T7 promoter followed by U6 snRNA coding sequence (*SNR6*) with a *DraI* site at the 3' end; used for PCR amplification of T7U6 for *in vitro* transcription	Peter Hodges, this lab
pU2Δ107	pRS316 containing the U2 snRNA coding sequence (*SNR20*; on a *SalI/Eco*RI fragment) with internal 892nt deletion (McPheeters *et al.*, 1989); PCR amplification of T7U2Δ107 for *in vitro* transcription	A. Newman
pIVEX2.5d	*E. coli* HA-tagging expression vector; HA tag from pMR81 is based on this sequence	Roche
pPROTet.E	*E. coli* Myc-tagging expression vector; Myc tag from pMR82 is based on this sequence	Clontech
pAG34 (hphMX3)	AmpR vector with TEF regulated Hygromycin B resistance gene; used as template for PCR products in yeast genomic deletions	Goldstein and McCusker, 1999
pFA6a-kanMX6	AmpR vector with TEF regulated G418 resistance gene; used as template for PCR products in yeast genomic deletions	Longtine *et al.*, 1998
pFA6a-natMX6	AmpR vector with TEF regulated Nourseothricin (clonNAT) resistance gene; used as template for PCR products in yeast genomic deletions	Hentges *et al.*, 2005
pFA6a-hphMX6	AmpR vector with TEF regulated Hygromycin B resistance gene; used as template for PCR products in yeast genomic deletions	Hentges *et al.*, 2005

2.1.8 Oligonucleotides

Oligonucleotides that were used in this study are listed in the tables below (2.1.8.1a – 2.1.8.5b). All oligonucleotides were synthesised by and purchased from MWG Biotech or Sigma-Genosys Ltd., unless otherwise stated. For oligonucleotides that were a gift from from others inside or outside this lab the source is given.

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR

Oligo Name	Alias	Sequence	Description
T7F	mro1	TAATACGACTCACTATAGGG	Forward primer in the T7 promoter region
T7R	mro2	GCTAGTTATTGCTCAGCGG	Reverse primer in the T7 terminator region
NdeI-Lsm2	mro3	AAAAAAACATATGCTTTTCTTCTCC	*LSM2* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm2-BamHI	mro4	AAGGATCCTTATTTTCTTTCAGTCA	*LSM2* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
NdeI-Lsm3	mro5	AAAAAAACATATGGAGACACCTTTG	*LSM3* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm3-BamHI	mro6	AAGGATCCTTATATCTCCACTGCGC	*LSM3* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
NdeI-Lsm4	mro7	AAAAAAACATATGCTACCTTTATAT	*LSM4* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm4-BamHI	mro8	AAGGATCCTTAAAATTCGACCTTTT	*LSM4* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
NdeI-Lsm5	mro9	AAAAAAACATATGAGTCTACCGGAG	*LSM5* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm5-BamHI	mro10	AAGGATCCTTACAACGCCTCCGTAG	*LSM5*forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
NdeI-Lsm6	mro11	AAAAAAACATATGTCCGGAAAAGCT	*LSM6* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm6-BamHI	mro12	AAGGATCCCTATATTTTTGTTCAC	*LSM6* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
NdeI-Lsm7	mro13	AAAAAAACATATGCATCAGCAACAC	*LSM7* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
Lsm7-BamHI	mro14	AAGGATCCCTATTTTTGCATATATA	*LSM7* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
HA-F	mro15	TACCAGCTACCCATACGATGTTCCAGA TTACGCTCA	can be cloned with HA-R directly into *Nde*I site introducing a single HA tag with a 3' *Nde*I site for cloning, used to make pMR81
HA-R	mro16	TATGAGCGTAATCTGGAACATCGTATG GGTAGCTGG	can be cloned with HA-F directly into *Nde*I site introducing a single HA tag with a 3' *Nde*I site for cloning, used to make pMR81
Myc-F	mro17	TATCGGTGAACAGAAACTGATCAGCGA AGAGGATCTGCA	can be cloned with Myc-R directly into *Nde*I site introducing a single c-Myc tag with a 3' *Nde*I site for cloning, used to make pMR82
Myc-R	mro18	TATGCAGATCCTCTTCGCTGATCAGTT TCTGTTCACCGA	can be cloned with Myc-F directly into *Nde*I site introducing a single c-Myc tag with a 3' *Nde*I site for cloning, used to make pMR82
Lsm1-BamHI	mro19	AAGGATCCTTAGTACATGTCAGAT	*LSM1* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
NdeI-LSM1	mro20	AAAAAAACATATGTCTGCAAATAGC	*LSM1* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
LSM8-BamHI	mro21	AAGGATCCTTATTTTGTCTTTGATTC	*LSM8* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
NdeI-LSM8	mro22	AAAAAAACATATGGACCAGCAAGC	*LSM8* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
pIP24delFC	mro25	TGGCGCCATTATTATATTTAG	Forward primer (nt1058-1079 of PRP24) to check for Δ10 deletion in prp24 together with pIPdelRC
pIPdelRC	mro26	CATTTTAGCTTCCTTAGCTCC	Reverse primer in pIP24 plasmid to check for Δ10 deletion in *PRP24* together with pIP24delFC
NdeI-Prp24	mro30	AAAAAAACATATGGAGTATGGACATC	*PRP24* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
Prp24-BamHI	mro31	AAGGATCCCTACTCACCTAGAAAC	*PRP24* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Lsm4-XhoI	mro32	AAAACTCGAGAAATTCGACCTTTTG	*LSM4* reverse primer introducing a 3' *Xho*I site for pETDuet vectors MCS2

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
BamHI-Lsm8	mro33	AAGGATCCGATGGACCAGCAAGC	*LSM8* (incorrect) forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm8-PstI	mro34	AACTGCAGTTATTTTGTCTTTGATTC	*LSM8* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1
ACYCDuetUP1	mro35	GGATCTCGACGCTCTCCCT	sequencing primer for Novagen Duet series MCS1 forward
DuetDOWN1	mro36	GATTATGCGGCCGTGTACAA	sequencing primer Novagen Duet series MCS1 reverse
DuetUP2	mro37	TTGTACACGGCCGCATAATC	sequencing primer Novagen Duet series MCS2 forward
Lsm2-XhoI	mro38	AAAACTCGAGTTTTCTTTCAGTCATTACC	*LSM2* reverse primer introducing a 3' *Xho*I site for pETDuet vectors MCS2
BamHI-Lsm3	mro39	AAGGATCCGATGGAGACACCTTTGG	*LSM3* forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm3-PstI	mro40	AACTGCAGTTATATCTCCACTGCGC	*LSM3* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1
BamHI-Lsm1	mro41	AAGGATCCGATGTCTGCAAATAGC	*LSM1* forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm1-PstI	mro42	AACTGCAGTTAGTACATGTCAGAT	*LSM1* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1
BamHI-Lsm2	mro43	AAGGATCCGATGCTTTTCTTCTCC	*LSM2* forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm2-PstI	mro44	AACTGCAGTTATTTTCTTTCAGTC	*LSM2* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1
Lsm6-XhoI	mro45	AAAACTCGAGTATTTTTTGTTCACTG	*LSM6* reverse primer introducing a 3' *Xho*I site for pETDuet vectors MCS2
Lsm7-XhoI	mro46	AAAACTCGAGTTTTTGCATATATAGTAC	*LSM7* reverse primer introducing a 3' *Xho*I site for pETDuet vectors MCS2
BamHI-Lsm7	mro47	AAGGATCCGATGCATCAGCAACAC	*LSM7* forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm7-PstI	mro48	AACTGCAGTTATTTTTGCATATATAG	*LSM7* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
BamHI-Lsm5	mro49	AAGGATCCGATGAGTCTACCGG	*LSM5* forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
Lsm5-PstI	mro50	AACTGCAGTTACAACGCCTCC	*LSM5* reverse primer introducing a 3' *Pst*I site for pETDuet vectors MCS1
Lsm8Rpe2	mro56	ATTGGTTATGAATAGATTAG	reverse primer anneals to 101-120 of *LSM8* (used for primer extension)
BamHI-Lsm8new	mro58	AAGGATCCGATGTCAGCCACCTTG	*LSM8* (correct) forward primer introducing a 5' *Bam*HI site for pETDuet vectors MCS1
NdeI-Lsm8new	mro59	AAAAAAACATATGTCAGCCACCTTG	*LSM8* (correct) forward primer introducing a 5' *Nde*I site for pET19b
Prp24d10-BamHI	mro62	AAGGATCCCTACTCACCCATCTGC	*prp24Δ10* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
BamHI-327LSM5-F	mro83	AAGGATCCAGAACGTATACGTATATTG	Forward primer introducing *Bam*HI site at 5' end of –327 position of *LSM5* start
XhoI+127LSM5-R	mro84	AAAACTCGAGAGCAAATCACCAGTGTC	Reverse primer introducing *Xho*I site at 3' end of +127 position of *LSM5* stop
U6_check	mro98	CAATATTCACATTCATAAAGC	Reverse complement to downstream sequence of *SNR6*
Ptef-R	mro99	ATGTATGGGCTAAATGTAC	Reverse complement to TEF promoter (*HpH*, *NAT* and *KAN* pFA6 based cassettes)
NdeI-Hfq-F	mro100	AAAAAAACATATGGCTAAGGGGCAATC	*hfq* forward primer introducing a 5' *Nde*I site for pET19b, pMR81 and pMR82
BamHI-Hfq-R	mro101	AAGGATCCTTATTCGGTTTCTTCGCT	*hfq* reverse primer introducing a 3' *Bam*HI site for pET19b, pMR81 and pMR82
Hfq-del-F1	mro102	AAAAACTGCAGTGCATCAGTTCGCCATCG	*Pst*I site at 5' end of nt-449 to -432 of *E. coli* Hfq
Hfq-del-R1	mro103	GAAACCTTATTCATTCTCTCTTTTCCTTATATG	Reverse complement of *E. coli hfq* nt+302 to +313 followed by reverse complement of nt-19 to +2
Hfq-del-F2	mro104	AAAAGAGAGAATGAATAAGGTTTCGGGCTG	nt-10 to +2 of *E. coli hfq* followed by nt+302 to +320
Hfq-del-R2	mro105	AAAAAGTCGACCTGCAACTTACCCTCATG	Reverse complement of SalI site at 3' end of nt+741 to +758 of *E. coli hfq*

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
U6F_check	mro108	CCTGGCATGAACAGTGG	Covers nt-78 to -62 of *SNR6*, used for sequencing and colony PCRs
Lsm1N-8C-F	mro122	GACATCGATAAAGAAGATAGTCTAGCT CCTATAGAC	*LSM1* nt.+346 to +363 followed by *LSM8* nt.+220 to +237
Lsm1N-8C-R	mro123	GTCTATAGGAGCTAGACTATCTTCTTT ATCGATGTC	Reverse complement of *LSM1* nt.+346 to +363 and *LSM8* nt.+220 to +237
Lsm8N-1C-F	mro124	GATGCAGAAAATGATGACCAACCCCTT GAGGCC	*LSM8* nt.+202 to +219 followed by *LSM1* nt.+364 to +378
Lsm8N-1C-R	mro125	GGCCTCAAGGGGTTGGTCATCATTTTC TGCATC	Reverse complement of *LSM8* nt.+202 to +219 and *LSM1* nt.+364 to +378
Lsm8-XhoI	mro126	AAAACTCGAGTTATTTTGTCTTTGATT C	*LSM8* reverse primer introducing a 3' *Xho*I site for pGFP-N-FUS
Lsm1dC-XhoI-R	mro129	AAAACTCGAGTTAATCTTCTTTATCGA TGTC	*LSM1* reverse primer introducing a 3' *Xho*I site to *lsm1ΔC* (behind nt.363) for pGFP-N-FUS
Lsm8dC-XhoI-R	mro136	AAAACTCGAGTTAGTCATCATTTTCTG CATC	*LSM8* reverse primer introducing a 3' *Xho*I site and stop codon to *lsm8ΔC* (behind nt.219) for pGFP-N-FUS
ClaI-Lsm1-R	mro137	AAATCGATGTACATGTCAGATTTATGG	*LSM1* reverse primer introducing a 3' *Cla*I site for pGFP-C-FUS
ClaI-Lsm8-R	mro138	AAATCGATTTTTGTCTTTGATTCGTAC	*LSM8* reverse primer introducing a 3' *Cla*I site for pGFP-C-FUS
PMet25F	mro139	GTCAGATACATAGATAC	Forward primer at 3' end of *MET25* promoter used for sequencing of inserts of pGFP-C-FUS
Lsm8dC-ClaI-R	mro140	AAATCGATGTCATCATTTTCTGCATC	*LSM8* reverse primer introducing a 3' *Cla*I site to *lsm8ΔC* (behind nt.219) for pGFP-C-FUS
Lsm8dC-BamHI-R	mro141	AAGGATCCTTAGTCATCATTTTCTGCA TC	*LSM8* reverse primer introducing a 3' *Bam*HI site to *lsm8ΔC* (behind nt.219) for pGFP-C-FUS
BamHI-HfqdC83	mro143	AAGGATCCTTAGTAGTTACTGCTGGTA C	Reverse primer introducing a *Bam*HI site and stop codon to the 3' end of *E. coli hfq* deleted after aa83
XhoI-NLS-Lsm8dC-R	mro146	AAAACTCGAGTTAGACCTTTCTCTTCT TCTTTGGGTCATCATTTTCTGCATC	Reverse primer introducing *Xho*I site, SV40 NLS and stop codon at 3' end of *lsm8ΔC* (behind nt.219) for pGFP-N-FUS

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
ClaI-NLS-Lsm8dC-R	mro147	AAATCGATGACCTTTCTCTTCTTCTTTGGGTCATCATTTTCTGCATC	Reverse primer introducing *Cla*I site and SV40 NLS at 3' end of *lsm8ΔC* (behind nt.219) for pGFP-C-FUS
BamHI-NLS-Lsm8dC-R	mro148	AAGGATCCTTAGACCTTTCTCTTCTTCTTTGGGTCATCATTTTCTGCATC	Reverse primer introducing *Bam*HI site and SV40 NLS at 3' end of *lsm8ΔC* (behind nt.219) for pGFP-C-FUS
Lsm1NN8CN-F	mro149	CTATTGTAAGCTCAGTAGACAAAAGAGTTGTTATAATCAAAGT	*LSM1* nt.+134 to +153 followed by *LSM8* nt.+31 to +53
Lsm1NN8CN-R	mro150	ACTTTGATTATAACAACTCTTTTGTCTACTGAGCTTACAATAG	Reverse complement of *LSM1* nt.+134 to +153 and *LSM8* nt.+31 to +53
Lsm8NN1CN-F	mro151	CACCTTGAAAGACTACTTAAATCGTAAAATCTTCGTTCTTTTG	*LSM8* nt.+9 to +30 followed by *LSM1* nt.+154 to +174
Lsm8NN1CN-R	mro152	CAAAAGAACGAAGATTTTACGATTTAAGTAGTCTTTCAAGGTG	Reverse complement of *LSM8* nt.+9 to +30 and *LSM1* nt.+154 to +174
BamHI-Lsm1ns-R	mro153	AAGGATCCGTACATGTCAGATTTATGG	Reverse complement introducing *Bam*HI site to 3' end of *LSM1* without stop codon for pGFP-C-FUS
BamHI-Lsm8ns-R	mro154	AAGGATCCTTTTGTCTTTGATTCGTAC	Reverse complement introducing *Bam*HI site to 3' end of *LSM8* without stop codon for pGFP-C-FUS
BamHI-Hfq-F	mro155	AAGGATCCATGGCTAAGGGGCAATC	Introduces *Bam*HI site at 5' end of *E. coli hfq*
XbaI-Hfq-F	mro156	AAAATCTAGAATGGCTAAGGGGCAATC	Introduces *Xba*I site at 5' end of *E. coli hfq* for pGFP-N-FUS
XhoI-Hfq-R	mro157	AAAACTCGAGTTATTCGGTTTCTTCGCT	Reverse primer introducing *Xho*I site at 3' end of *E. coli hfq* for pGFP-N-FUS
ClaI-lsm8N-R	mro158	AAATCGATATTTAAGTAGTCTTTCAAGGTG	Reverse primer introducing *Cla*I site at 3' end of *LSM8* N-terminal extension (behind nt.+30)
XbaI-lsm1dN-F	mro159	AAATCTAGACGTAAAATCTTCGTTCTTTTG	Introduces *Xba*I site at 5' end of *LSM1ΔN* (from nt.+154)
XbaI-lsm1C-F	mro160	AAATCTAGACAACCCCTTGAGGCCATG	Introduces *Xba*I site at 5' end of LSM C-terminal extension (from nt.+364)
XbaI-lsm8dN-F	mro161	AAATCTAGAAAAAGAGTTGTTATAATCAAAG	Introduces *Xba*I site at 5' end of *LSM8ΔN* (from nt.+31)

Table 2.1.8.1a Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo Name	Alias	Sequence	Description
BamHI-Lm1dC-R	mro162	AAGGATCCTTAATCTTCTTTATCGATGTC	Reverse primer introducing *Bam*HI site and stop codon at 3' end of *LSM1ΔC* (behind nt.+363)
BamHI-Lsm1dCns-R	mro163	AAGGATCCATCTTCTTTATCGATGTCTAC	Reverse primer introducing *Bam*HI site at 3' end of *LSM1ΔC* (behind nt.+363)
XbaI-lsm1dNATG-F	mro164	AAATCTAGAATGCGTAAAATCTTCGTTCTTTTG	Introduces *Xba*I site and start codon at 5' end of *LSM1ΔN* (from nt.+154)
XbaI-lsm8dNATG-F	mro165	AAATCTAGAATGATGAAAAGAGTTGTTATAATCAAAG	Introduces *Xba*I site and start codon at 5' end of *LSM8ΔN* (from nt.+31)
ClaI-lsm1N-R	mro166	AAATCGATGTCTACTGAGCTTACAATAG	Reverse primer introducing *Cla*I site at 3' end of *LSM1* N-terminal extension (behind nt.+153)

Table 2.1.8.1b Oligonucleotides used for cloning, sequencing and colony PCR

Oligo name	Sequence	Discription	Source
S8949	CCCGGATCCATGGAGTATGGACATCACGC	Introduces *Bam*HI site at 5' end of PRP24 (used for *PRP24* sequencing)	Vidal, 1998
S8949	CCACTCTAGAGTCTACTCACCTAGAAAC	Reverse primer introducing *Xba*I site at 3' end of *PRP24* (used for *PRP24* sequencing)	Vidal, 1998
JMB11	GGTGTCACCAAGCTTAATC	*PRP24* nt. +966 to +983; used for sequencing (used for *PRP24* sequencing)	Vidal, 1998
N6089	GTATACGGCTTCCCAG	*PRP24* nt. +437 to +451; used for sequencing (used for *PRP24* sequencing)	Vidal, 1998
Oligo-1	GCTTCGTTCGCCTTC	Reverse complement of *PRP24* nt. +83 to +98 (used for *PRP24* sequencing)	Vidal, 1998
Oligo-2	GCGAATCCGCAACATC	Reverse complement of *PRP24* nt. +210 to +226 (used for *PRP24* sequencing)	Vidal, 1998
Oligo-3	TCGCCCCTCTAGTGTG	Reverse complement of *PRP24* nt. +615 to +630 (used for *PRP24* sequencing)	Vidal, 1998

Table 2.1.8.1b Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo name	Sequence	Discription	Source
P5117	GGCGCCATTGAAGTC	Reverse complement of *PRP24* nt. +1050 to +1065 (used for *PRP24* sequencing)	Vidal, 1998
P5118	GAAGGCGAACGAAGC	*PRP24* nt. +83 to +98 (used for *PRP24* sequencing)	Vidal, 1998
M13-F	GTAAAACGACGGCCAG	M13 forward sequencing primer	
M13-R	CAGGAAACAGCTATGAC	M13 reverse sequencing primer	
pQE-F	CGTATCACGAGGCCCTTTCG	Sequencing primer for pQE vectors, approximately 100nt from *Bam*HI site	Markus Bohnsack Tollervey lab
pQE-R	GTTCTGAGGTCATTACTGG	Sequencing primer for pQE vectors, approximately 100nt from *Bam*HI site	Markus Bohnsack Tollervey lab
S680	GGGGACAAGTTTGTACAAAAAAGCAGGCT ctGGCTAATTAACTGGGTTATGTG	Forward primer used to check genomic *lsm6* deletion	Liz Thompson, this lab
S681	GGGGACCACTTTGTACAAGAAAGCTGGGT tCTCCGTCACTTTCTGGCTCT	Reverse primer used to check genomic *lsm6* deletion	Liz Thompson, this lab
S674	GGGGACAAGTTTGTACAAAAAAGCAGGCT CTGTCTTCTTCAACTGCCGTTTC	Forward primer used to check genomic *lsm7* deletion	Liz Thompson, this lab
S677	GGGGACCACTTTGTACAAGAAAGCTGGGT tCAGTTTGTTTCGTTTTTCCTCA	Reverse primer used to check genomic *lsm7* deletion	Liz Thompson, this lab
Lsm8genF (MTO42)	CCCATTCCAGTATCTGCTTGTC	Anneals 443 bp upstream of start of *LSM8* ORF; used to check for *lsm8* deletion	Marian Thomson, this lab
R-lsm8-485	GGGGAAGAGAAGAAGCAT	Anneals 485 bp downstream of *LSM8* ORF; used to check for lsm8 deletion	Kum Loong Boon, this lab
rpTEF	GGATGTATGGGCTAAATGTAC	Reverse primer in the TEF promoter region; used to check for presence of deletion/tagging cassette on the genome	Marian Thomson, this lab
pKO.F	AGGGCAGGGTCGTTAAATAGC	pTOF24 sequencing primer forward	David Leach
pKO.R2	AGGGAAGAAAGCGAAAGGAG	pTOF24 sequencing primer reverse	David Leach
XBALSM1 F	GCTCTAGAATGTCTGCAAATAGCAAGGAC	Introduces *Xba*I site at 5' end of *LSM1*, for cloning into pGFP-N-FUS	Michael Spiller, this lab

Table 2.1.8.1b Oligonucleotides used for cloning, sequencing and colony PCR (continued)

Oligo name	Sequence	Discription	Source
XHOLSM1 R	CCGCTCGAGTTAGTACATGTCAGATTTATGG	Introduces *Xho*I site at 3' end of *LSM1*, for cloning into pGFP-N-FUS	Michael Spiller, this lab
XBALSM8 F	CGTCTAGAATGTCAGCCACCTTGAAAGACTACTTAAATAAA	Introduces *Xba*I site at 5' end of *LSM8*, for cloning into pGFP-N-FUS	Michael Spiller, this lab
XBAGFPF	GCTCTAGAGTGGATCCCCCCGCTGAATTCATGAGT	Introduces *Xba*I site at 5' end of *GFP* ORF from pGFP-N-FUS	Michael Spiller, this lab

Table 2.1.8.2 Oligonucleotides used in site directed mutagenesis

Oligo Name	Alias	Sequence	Description
U6-G63A-F	mro111	CAGAGATGATCAGCAATTCCCCTGCATAAGGATG	U6-G63A primers can be used to introduce the C69A mutation in *SNR6* using the Quik Change method
U6-G63A-R	mro112	CATCCTTATGCAGGGGAATTGCTGATCATCTCTG	U6-G63A primers can be used to introduce the C69A mutation in *SNR6* using the Quik Change method
U6-C69A-F	mro109	GATGATCAGCAGTTCCCATGCATAAGGATGAACCG	U6-C69A primers can be used to introduce the C69A mutation in *SNR6* using the Quik Change method
U6-C69A-R	mro110	CGGTTCATCCTTATGCATGGGAACTGCTGATCATC	U6-C69A primers can be used to introduce the C69A mutation in *SNR6* using the Quik Change method
U6-G63A/C69A-F	mro144	CAGAGATGATCAGCAATTCCCATGCATAAGGATG	U6-G63A/C69A primers can be used to introduce G63A and C69A mutations into *SNR6* using the Quik Change method
U6-G63A/C69A-R	mro145	CATCCTTATGCATGGGAATTGCTGATCATCTCTG	U6-G63A/C69A primers can be used to introduce G63A and C69A mutations into *SNR6* using the Quik Change method
PIP24Fb	mro23	CCTGACAAACAAGAGCAGATGGGTGAGTAGACTC	Used with PIP24rB to introduce *prp24Δ10* in pIP24 by Quik change method

Table 2.1.8.2 Oligonucleotides used in site directed mutagenesis (continued)

Oligo Name	Alias	Sequence	Description
PIP24rB	mro24	AGGTCGCCACTCTAGAGTCTACTCACC CATCTGCTC	Used with PIP24Fb to introduce *prp24Δ10* in pIP24 by Quik change method
RRM4sub-F	mro106	CTTCTCGTTAGCGACGCCAATGCCGCC GCTATTATATTTAGAGATAG	RRM4sub primers can be used to introduce the RRM4sub mutation in *PRP24* using the Quik Change method
RRM4sub-R	mro107	CTATCTCTAAATATAATAGCGGCGGCA TTGGCGTCGCTAACGAGAAG	RRM4sub primers can be used to introduce the RRM4sub mutation in *PRP24* using the Quik Change method
Prp24-F87A-F	mro113	CGTTTTGCACGTATTGAAGCTGCCAGG TATGATGGAGC	Prp24-F87A primers can be used to introduce the F87A mutation in *PRP24* using the Quik Change method
Prp24-F87A-R	mro114	GCTCCATCATACCTGGCAGCTTCAATA CGTGCAAAACG	Prp24-F87A primers can be used to introduce the F87A mutation in *PRP24* using the Quik Change method
Prp24-R158S-F	mro118	CGATTCAATACAAGCAGTAGGTTCGCT TACATCGATG	Prp24-R158S primers can be used to introduce the R158S mutation in *PRP24* using the Quik Change method
Prp24-R158S-R	mro119	CATCGATGTAAGCGAACCTACTGCTTG TATTGAATCG	Prp24-R158S primers can be used to introduce the R158S mutation in *PRP24* using the Quik Change method
Prp24-F257I-F	mro120	GCACAGTTTCAATAACTGTTGTGCAAT TATGGTTTTTGAAAATAAAGATTC	Prp24-F257I primers can be used to introduce the F257I mutation in *PRP24* using the Quik Change method
Prp24-F257I-R	mro121	GAATCTTTATTTTCAAAAACCATAATT GCACAACAGTTATTGAAACTGTGC	Prp24-F257I primers can be used to introduce the F257I mutation in *PRP24* using the Quik Change method
Lsm5-E10K-F	mro85	GTCTACCGGAGATTTTGCCTTTAAAAG TCATAGATAAAACAATTAACCAG	Add E10K mutation and *DraI* site to *LSM5* using the Quik Change method
Lsm5-E10K-R	mro86	CTGGTTAATTGTTTTATCTATGACTTT TAAAGGCAAAATCTCCGGTAGAC	Add E10K mutation and DraI site to *LSM5* using the Quik Change method
Lsm8-N28A/D31A-F	mro130	CCTCATAGCAAGCCTAGCCGGCTTCGC CAAAAATACTAATCTATTC	Lsm8-N28A/D31A primers can be used to introduce N28A and D31A mutations into *LSM8*

Table 2.1.8.2 Oligonucleotides used in site directed mutagenesis (continued)

Oligo Name	Alias	Sequence	Description
Lsm8-N28A/D31A-R	mro131	GAATAGATTAGTATTTTTGGCGAAGCC GGCTAGGCTTGCTATGAGG	Lsm8-N28A/D31A primers can be used to introduce N28A and D31A mutations into *LSM8* using the Quik Change method
Lsm8-T34A/N35A-F	mro132	CTAAACGGCTTCGACAAAAATGCTGCT CTATTCATAACCAATGTTTTC	Lsm8-T34A/N35A primers can be used to introduce T34A and N35A mutations into *LSM8* using the Quik Change method
Lsm8-T34A/N35A-R	mro133	GAAAACATTGGTTATGAATAGAGCAGC ATTTTTGTCGAAGCCGTTTAG	Lsm8-T34A/N35A primers can be used to introduce T34A and N35A mutations into *LSM8* using the Quik Change method
Lsm8-RGS57AWA-F	mro134	CTGCAAGGCACAGTTACTTGCATGGGC CGAGATTGCTCTTGTTGGCCTC	Lsm8-RGS57AWA primers can be used to introduce R57A, G58W and S59A mutations into *LSM8* using the Quik Change method
Lsm8-RGS57AWA-R	mro135	GAGGCCAACAAGAGCAATCTCGGCCCA TGCAAGTAACTGTGCCTTGCAG	Lsm8-RGS57AWA primers can be used to introduce R57A, G58W and S59A mutations into *LSM8* using the Quik Change method

Table 2.1.8.3a Oligonucleotides used for *in vitro* transcription and strand displacement assays

Oligo Name	Alias	Sequence	Description
T7Fnew	mro27	CCTAATACGACTCACTATAGG	for PCR on T7 promoter on pT7U6 (minimal T7 promoter followed by first G of U6)
T7U4Fnew	mro28	CCTAATACGACTCACTATAGGATCCTT ATGCACGGG	for PCR on pU4 with addition of minimal T7 promoter (anneals to first 15 bp of U4 gene)
U6-FokI-R	mro29	AAACGGTTCATCCTTATGCAG	Reverse primer on U6 gene (69-89) to create *in vitro* transcription template with T7Fnew for U6Δ3'
T7U6rc	mro51	CCTAATACGACTCACTATAGGAAAACG AAATAAATCTCTTTG	Introduces a minimal T7 promoter at the 3' end of U6 to produce a reverse complement template with U6F for U6_rc RNA

Table 2.1.8.3a Oligonucleotides used for *in vitro* transcription and strand displacement assays (continued)

Oligo Name	Alias	Sequence	Description
U6F	mro52	GTTCGCGAAGTAACCC	5' end of U6 gene used to produce reverse complement template in combination with T7U6rc for U6_rc RNA
T7U2F	mro54	CCTAATACGACTCACTATAGGATCTCT TTGCCTTTTGG	introduces a minimal T7 promoter to the 5' end of the U2 gene
U2R	mro55	ATAAAAGAGCGAACGGG	reverse primer anneals to 3' end of U2 gene
T7U6(71-86)F	mro60	CCTAATACGACTCACTATAGGCATAAG GATGAACCG	Introduces a minimal T7 promoter 5' of U6 sequence 71-86
U6(55-70)rc	mro61	AGGGGAACTGCTGATC	reverse primer anneals to 55-70 of U6 gene
U6(76-91)rc	mro63	TAAAACGGTTCATCCT	anneals to 76-91 of U6 gene (or U6 RNA)
U6(77-92)rc	mro64	GTAAAACGGTTCATCC	anneals to 77-92 of U6 gene (or U6 RNA)
16nt	mro65	TTGCTTTACGGTGCTA	random 16nt (taken from unwinding substrate in (Cordin *et al.*, 2004))
T7F2	mro66	TAATACGACTCACTATAGG	T7 promoter for small RNA *in vitro* transcriptions
T7U6(77-112)rc	mro67	AAAACGAAATAAATCTCTTTGTAAAAC GGTTCATCCTATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-112) RNA
T7U6(77-108)rc	mro68	CGAAATAAATCTCTTTGTAAAACGGTT CATCCTATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-108) RNA
T7U6(77-103)rc	mro69	TAAATCTCTTTGTAAAACGGTTCATCC TATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-103) RNA
T7U6(77-98)rc	mro70	CTCTTTGTAAAACGGTTCATCCTATAG TGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of 16nt-U6(92-112) RNA
T7-16nt-U6(92-112)rc	mro71	AAAACGAAATAAATCTCTTTGTTGCTT TACGGTGCTACCTATAGTGAGTCGTAT TA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-112) RNA
U6(77-98)rc	mro72	CTCTTTGTAAAACGGTTCATCC	anneals to 77-98 of U6 gene (or U6 RNA)
U6(71-86)rc	mro73	CGGTTCATCCTTATGC	anneals to 71-86 of U6 gene (or U6 RNA)
U6(83-98)rc	mro74	CTCTTTGTAAAACGGT	anneals to 83-98 of U6 gene (or U6 RNA)
U6(26-35)rc	mro75	TTGACCAAATTT	anneals to 26-35 of U6 gene (or U6 RNA)

Table 2.1.8.3a Oligonucleotides used for *in vitro* transcription and strand displacement assays (continued)

Oligo Name	Alias	Sequence	Description
T7U6(36-70)rc	mro76	AGGGGAACTGCTGATCATCTCTGTATT GTTTCAAACCTATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(36-70) RNA
T7U6(40-70)rc	mro77	AGGGGAACTGCTGATCATCTCTGTATT GTTTCCTATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(40-70) RNA
T7-16nt-U6(92-98)rc	mro78	CTCTTTGTTGCTTTACGGTGCTACCTA TAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of 16nt+U6(92-98) RNA
T7U6(77-92)rc	mro79	GTAAAACGGTTCATCCTATAGTGAGTC GTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-92) RNA
T7-16nt-rc	mro80	TTGCTTTACGGTGCTACCTATAGTGAG TCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of 16nt RNA
T7U6(77-92)20nt-rc	mro81	AGCTTGCTTTACGGTGCTAGGTAAAAC GGTTCATCCTATAGTGAGTCGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-92)+20nt RNA
T7rcU6(77-92)	mro82	GGATGAACCGTTTTACCTATAGTGAGT CGTATTA	Template when annealed with T7F2 for *in vitro* transcription of U6(77-92)rc RNA
T7U6(71-112)	mro87	AAAACGAAATAAATCTCTTTGTAAAAC GGTTCATCCTTATGCCTATAGTGAGTC GTATTA	Should be T7U6(71-112)rc; Template when annealed with T7F2 for *in vitro* transcription of U6(71-112) RNA
T7U6(93-112)(77-92)rc	mro88	GTAAAACGGTTCATCCAAAACGAAATA AATCTCTTTCCTATAGTGAGTCGTATT A	Template when annealed with T7F2 for *in vitro* transcription of U6(93-112)+(77-92) RNA
T7U6(112-93)(77-92)rc	mro89	GTAAAACGGTTCATCCTTTCTCTAAAT AAAGCAAAACCTATAGTGAGTCGTATT A	Template when annealed with T7F2 for *in vitro* transcription of U6(112-93)+(77-92) RNA
T7U6F	mro117	CCTAATACGACTCACTATAGGTTCGCG AAGTAACCC	for PCR adding T7 promoter to U6 5'end

Table 2.1.8.3b Oligonucleotides used for *in vitro* transcription and strand displacement assays

Oligo name	Sequence	Discription	Source
T7U4F	CGTAATACGACTCACTATAGGGAGAATCC TTATGCACGGG	Introduces a T7 promoter to the 5' end of *SNR14*	Page, 2002
U4R	AAAGGTATTCCAAAAATTCCC	Reverse primer from 3' end of U4 snRNA coding sequence	Page, 2002
LORY1	CGTAATACGACTCACTATAGGGAGAGTTC GCGAAGTAACCC	Introduces a T7 promoter to the 5' end of *SNR6*	Page, 2002
S5761	AAAACGAAATAAATCTCTTTGTAAAAC	Reverse primer from 3' end of U6 snRNA coding sequence	Page, 2002

Table 2.1.8.4 Oligonucleotides used as probes in Northern hybridisation

Oligo name	Sequence	Discription	Source
G8102	CACGCCTTCCGCGCCGT	U1 snRNA Northern probe	Boon, 2005
G8103	CTACACTTGATCTAAGCCAAAAG	U2 snRNA Northern probe	Boon, 2005
G6568	AGGTATTCCAAAAATTCCC	U4 snRNA Northern probe	Boon, 2005
U5-7wtsmnr	AAGTTCCAAAAAATATGGCAAGC	U5 snRNA Northern probe	Boon, 2005
U6_knockout (V4528)	ATCTCTGTATTGTTTCAAATTGACC AA	U6 snRNA Northern probe	Boon, 2005
U6_knockout_rc (mro53)	TTGGTCAATTTGAAACAATACAGAG AT	reverse complement of U6_knockout primer (29-54), used as a probe for hybridisation to U6 reverse complement RNA	This thesis
TSA1 probe (#1455)	GGAGTATTCGGAGTCAGTGGAGGCG AAAAGAACT	Used as probe in Northern blotting hybridisation of *TSA1* mRNA	Tollervey lab
SCR1 probe (#250)	ATCCCGGCCGCCTCCATCAC	Used as probe in Northern blotting hybridisation of *scR1* RNA	Tollervey lab

Table 2.1.8.5a Oligonucleotides used for yeast genomic deletion and tagging

Oligo Name	Alias	Sequence	Description
Lsm5-HpH-F	mro115	ATAGCGTATACACGTGCACTCACAAGC ATACACAGAAGCGCGCAAACGGATCCC CGGGTTAATTAAG	Forward primer covering 1-46nt of *LSM5* followed by pFA6a forward sequence (F1 deletion primer, Longtine *et al.*, 1998)
Lsm5-HpH-R	mro116	ATTTTTTTTAGATGCACTATACAAATT CGCGTTAATTTTTGCCTTTGAATTCGA GCTCGTTTAAAC	Reverse primer covering the last 46nt of *LSM5* followed by pFA6a reverse sequence (R1 deletion primer, Longtine *et al.*, 1998)
Lsm6-HpH-F	mro90	TTCTATTATTAAAACATTTTACCCGGG GGCAGCTGTTCGTGTACACGGATCCCC GGGTTAATTAAG	Forward primer covering 45nt of *LSM6* followed by pFA6a forward sequence (F1 deletion primer, Longtine *et al.*, 1998)
Lsm6-HpH-R	mro91	ACCAACTTGCTCATTCCTACATATTAA TCCATTAGAGGAGATAAGTGAATTCGA GCTCGTTTAAAC	Reverse primer covering 46nt of *LSM6* followed by pFA6a reverse sequence (R1 deletion primer, Longtine *et al.*, 1998)
Lsm7-HpH-F	mro92	TGAGAGCAGCACTTTGTTTACTACACA GAACATTAACCAAAAAAACCGGATCCC CGGGTTAATTAAG	Forward primer covering 46nt of *LSM7* followed by pFA6a forward sequence (F1 deletion primer, Longtine *et al.*, 1998)
Lsm7-HpH-R	mro93	CGTTTTTAGAAATCAGTTCAGTGTTGT TTTCATCATCAGGATTAGGAATTCGAG CTCGTTTAAAC	Reverse primer covering 45nt of *LSM7* followed by pFA6a reverse sequence (R1 deletion primer, Longtine *et al.*, 1998)

Table 2.1.8.5b Oligonucleotides used for yeast genomic deletion and tagging

Oligo name	Sequence	Discription	Source
S682	TTCTATTATTAAAACATTTTACCCGGG GGCAGCTGTTCGTGTACACGCGCCAGA TCTGTTTAGC	Forward primer for *LSM6* genomic deletion using the KANMX4 cassette	Liz Thompson, this lab
S686	ACCAACTTGCTCATTCCTACATATTAA TCCATTAGAGGAGATAAGTAATTCGAG CTCGTTTAAACTGG	Reverse primer for *LSM6* genomic deletion using the KANMX4 cassette	Liz Thompson, this lab

Table 2.1.8.5b Oligonucleotides used for yeast genomic deletion and tagging (continued)

Oligo name	Sequence	Discription	Source
S678	TGAGAGCAGCACTTTGTTTACTACACA GAACATTAACCAAAAAAACCGCGCCAG ATCTGTTTAGC	Forward primer for *LSM7* genomic deletion using the KANMX4 cassette	Liz Thompson, this lab
S687	CGTTTTTAGAAATCAGTTCAGTGTTGT TTTCATCATCAGGATTAGAATTCGAGC TCGTTTAAACTGG	Reverse primer for *LSM7* genomic deletion using the KANMX4 cassette; deletes Exon1 and Exon2 to end of Sm domain (no disruption of putative YNL147C-A 3'UTR)	Liz Thompson, this lab

2.1.9 Antisera

The antisera used in this study are listed in Table 2.1.9 below. Santa Cruz antibodies were also used for Western blotting in horseradish peroxidase (HRP) conjugated form, making the use of a secondary antibody superfluous. For non-HRP conjugated primary antibodies the appropriate secondary HRP-conjugated antibody was used. The same Santa Cruz antibodies were also used for immunoprecipitation in agarose conjugated form. S-protein HRP conjugate was used according to the manufacturer's instructions.

Table 2.1.9 Antisera

Antibody	Description	Reference
Anti-HA (F-7)	Mouse monoclonal antibody (F-7; IgG_{2a}) raised against the influenza hemaglutinin HA-1 protein epitope; 1:1,000 dilution for Western blotting	Santa Cruz Biotechnology
Anti-His (H-3)	Mouse monoclonal antibody (H-3; IgG1) raised against the polyhistidine domain of expression vectors containing the polyhistidine sequence; 1:1,000 dilution for Western blotting	Santa Cruz Biotechnology

Table 2.1.9 Antisera (continued)

Antibody	Description	Reference
Anti-His (H-15)	Rabbit polyclonal antibodies (H-15) raised against a peptide corresponding to the polyhistidine domains of pET and Xpress polyhistidine expression vectors; 1:1,000 dilution for Western blotting	Santa Cruz Biotechnology
Anti-GFP	Mouse monoclonal antibody (IgG_{2a}) raised against full length *A. Victoria* GFP; 1:1,000 dilution for Western blotting	BD Biosciences
Anti-c-Myc (9E10)	Mouse monoclonal antibody (9E10; IgG_1) raised against the C-terminus of human c-Myc; 1:1,000 dilution for Western blotting	Santa Cruz Biotechnology
Anti-Mouse IgG-HRP	Sheep anti-rabbit IgG covalently linked to horseradish peroxidase; 1:10,000 dilution for Western blotting	Amersham Biosciences
Anti-Rabbit IgG-HRP	Donkey anti-rabbit IgG covalently linked to horseradish peroxidase; 1:10,000 dilution for Western blotting	Amersham Biosciences
S-protein HRP conjugate	Ribonuclease S-protein covalently linked to HRP has high affinity to the 15aa S-tag; 1:5,000 dilution for Western blotting	Novagen

2.2 Microbiological methods

2.2.1 Growth of strains

2.2.1.1 Growth of bacteria

Strains of *E. coli* were grown in/on rich LB medium at 37°C, unless otherwise stated. To select for or maintain the incorporation of a plasmid, the transformed bacteria were grown in/on rich LB medium containing the appropriate antibiotic(s) (2.1.2.5).

2.2.1.2 Growth of yeast

Unless otherwise stated, yeast strains were grown at 30°C in/on the appropriate medium (2.1.2.3). To maintain selection for plasmid DNA and/or for a nutritional reporter gene inserted into the genome, cells were grown in/on YMM (2.1.2.3)

complemented with the appropriate dropout powder (2.1.2.4), i.e. synthetic defined (SD) medium. To maintain selection for an antibiotic resistance gene inserted into the genome, cells were grown in/on the appropriate medium with the relevant antibiotic(s) added at the required concentration (2.1.2.5)

2.2.2 Preservation of strains

2.2.2.1 Preservation of bacteria

For storage of *E. coli* for time periods up to four weeks, bacteria were stored on solid medium at 4°C. In order to preserve strains indefinitely, *E. coli* were grown overnight to stationary phase. 1.5 ml of this culture was added to 150 µl of 87% (w/v) glycerol, resulting in the cells being suspended in 8% (w/v) glycerol, in 2 ml Cryo-s screw capped tubes (Greiner). This was subsequently snap-frozen on dry ice and stored at –70°C.

2.2.2.2 Preservation of *S. cerevisiae*

Yeast strains were preserved on solid medium for up to eight weeks at 4°C. In order to preserve strains indefinitely, yeast cells were grown overnight into late log phase or early stationary phase. 1.5 ml of this culture was added to 300 µl of 87% (w/v) glycerol, resulting in the cells being suspended in 15% (w/v) glycerol, in 2 ml Cryo-s screw capped tubes (Greiner). This was subsequently snap-frozen on dry ice and stored at –70°C.

2.2.3 Transformation of *E. coli*

2.2.3.1 Preparation of chemically-competent cells for cloning

Preparation of chemically-competent *E. coli* was performed according to Inoue *et al.*, 1990 (Inoue *et al.*, 1990). *E. coli* DH5α or TOP10 cells (2.1.5) were grown overnight on LB-agar. The next day, a single colony was taken and 5 ml of rich LB with 20 mM $MgSO_4$ was inoculated and grown overnight to stationary phase. Following this, 250 ml of LB with 20 mM $MgSO_4$ was inoculated with 2 ml of the stationary-phase culture and incubated at 23°C in a shaking incubator at a minimum of 200 rpm until OD_{600} reached 0.4-0.6 (usually 8-10 h). The culture was then cooled on ice for about 15 min and cells were kept on ice for all subsequent steps. Cells were sedimented (10

min, 3,000 rpm, 4°C; Beckman 10,500 rotor) and gently resuspended in 80 ml of ice-cold sterilised TB buffer (10 mM Pipes-HCl pH 6.7, 15 mM $CaCl_2$, 0.25 M KCl, 55 mM $MnCl_2$). Cells were left on ice for 10 minutes, centrifuged for 10 minutes at 3,000 rpm at 4°C, and gently resuspended in 20 ml of ice-cold TB buffer. After adding 1.5 ml of DMSO followed by a final 10 minutes incubation on ice, cells were dispensed into 200 µl aliquots in cold, sterile tubes and snap-frozen in liquid nitrogen. Aliquots were stored at –70°C until required for transformation (2.2.3.2).

2.2.3.2 Transformation of chemically-competent cells

For transformation of chemically-competent *E. coli* (2.2.3.1), aliquots were thawed on ice and either up to 10 µl of plasmid DNA, or 1-15 µl of a ligation reaction (2.3.2.14) was added to 50-200 µl of competent cells and gently mixed. Cells and DNA were incubated on ice for 30 min before a 45 seconds heat-shock in a 42°C waterbath, after which they were allowed to cool on ice for a further 2 min. Cells were pelleted (5 sec, 14,000 rpm at RT), resuspended in 150 µl of LB, pre-warmed at 37°C, before spreading them on LB-agar plates (100 µg/ml ampicillin; pre-warmed at 37°C) using the 'Copacabana' method (Worthington *et al.*, 2001). For antibiotics selection other than ampicillin 1 ml of LB (37°C) was added to the cells after the 2 min incubation on ice and the cells were incubated at 37°C for a further 30-60 min. Cells were then pelleted and resuspended in 150 µl of LB before spreading them on LB-agar plates containing the appropriate antibiotics. These were incubated overnight at 37°C to give cells transformed with the required plasmid time to produce visible colonies.

2.2.3.3 Preparation of electro-competent cells

A single fresh colony of *E. coli* DH5α or TOP10 cells (2.1.5) was used to inoculate 5 ml of LB and grown overnight to stationary phase. This culture was used to inoculate 500 ml of LB, which was grown at 37°C till an OD_{600} of 0.6-0.8 was reached. Cells were then cooled on ice for 15 min and sedimented by centrifugation for 15 min at 4,000 rpm, 4°C in a Beckman 10,500 rotor. From this point onwards care was taken to keep the cells ice-cold at all times. Cells were then washed twice in 500ml of ice-cold sterile water and sedimented as before, each time removing as much of the

supernate as possible. Cells were then washed twice in 50 ml of ice-cold 10% (w/v) glycerol and sedimented as before. Finally, cells were resuspended in 2 ml of ice-cold 10% (w/v) glycerol and 40 µl aliquots were snap-frozen in liquid nitrogen and stored at -70°C.

2.2.3.4 Transformation of electro-competent cells

Electro-competent *E. coli* cells were thawed on ice and 1-15 µl of DNA (plasmid rescued from yeast (2.3.2.3) or ligation mix (2.3.2.14)) was added. This mix was then transferred to a pre-chilled electroporation cuvette (2 mm electrode gap; Equibio or Eurogentec). The cuvette was dried and electroporation was performed using a Bio-Rad Gene Pulser II using the following settings: 200 Ω, 25 µF, 2.5 kV. Immediately after electroporation 1 ml of room temperature SOC or LB (2.1.2.2) was added. Cells were incubated for 30 min to 1 h at 37°C with agitation. Cells were then sedimented by centrifugation (30 s, 14,000 rpm), resuspended in 150 µl of medium and spread using the 'Copacabana' method (Worthington *et al.*, 2001) on LB plates containing the appropriate antibiotics (2.1.1.4).

2.2.3.5 Preparation of calcium-competent cells for transformation with plasmid DNA

For transformation of *E. coli* with purified plasmid DNA (2.3.2.1), e.g. in preparation for expression of recombinant protein (2.4.7), calcium-competent cells were prepared as follows. A single fresh colony was used to inoculate 5 ml of LB, which was incubated overnight in a shaking incubator. The following day 200 µl of this stationary phase culture was used to inoculate 20 ml of LB. This culture was grown for 90 min to 2 h with good aeration until OD_{600} reached 0.4-0.8. The cells were sedimented by centrifugation for 1 min at 10,000 rpm in a Beckman 10,500 rotor, resuspended in 10 ml of ice-cold sterilised 50 mM $CaCl_2$ and incubated on ice for 30 min to 2 h before repeat sedimentation (the presence of a hole in the cell pellet indicates competence). The cells were then resuspended gently in 1 ml of ice-cold 50 mM $CaCl_2$ and incubated on ice for a maximum of 24 h or until required for transformation. For long-term storage a half volume of ice-cold sterilised 30% glycerol was added to the cell suspension, before dispensing into 200 µl aliquots in

cold, sterile tubes. Aliquots were snap-frozen in liquid nitrogen and stored at –70°C until required for transformation.

2.2.3.6 Transformation of calcium-competent cells

For transformation of calcium-competent *E. coli*, aliquots were thawed on ice or freshly prepared competent cells were dispensed and 0.5 µl of plasmid DNA was added to 50-100 µl of competent cells. After gentle mixing the suspension was incubated on ice for 30 min and then 1 min in a 42 °C water bath. The cells were immediately transferred to ice and left there for 2 min to recover. 1 ml of LB (37°C) was then added to the cells followed by incubation at 37°C for a further 30-60 min. Cells were then sedimented, resuspended in 150 µl of LB and spread using the 'Copacabana' method (Worthington *et al.*, 2001) on LB-agar plates containing the appropriate antibiotics (2.1.1.4).

2.2.4 Creation of *E. coli* genomic deletion

The method used to make a genomic deletion of the *hfq* gene was based on that of Merlin *et al.* (Merlin *et al.*, 2002) using the pTOF24 plasmid (2.1.7). The cassette for integration into the bacterial chromosome was cloned using the SalI and PstI sites of pTOF24. The cassette was created by PCR with Hfq-del-F1, Hfq-del-R1, Hfq-del-F2 and Hfq-del-R2 primers (2.1.8) on BL21 (DE3) (2.1.5) genomic DNA (30 µl of water was boiled for 20 min with one BL21 (DE3) colony; use 1 µl for a 50 µl PCR). For design of the cassette see 3.6.2. Selection for the newly created plasmid was on LB-Cm plates (2.1.1.4) at 30°C (pTOF24 replication is temperature sensitive). Colonies were first checked for the absence of Kan resistance to ensure the presence of an insert in place of the Kan resistance gene. The insertion was then sequenced using pKO.F and pKO.R2 (2.1.8) to check for the presence of the correct sequence. The resulting plasmid (pMR54) was transformed using the $CaCl_2$ method into the strain in which the deletion was to be made (2.2.3.6). Transformed cells were selected on LB-Cm plates at 30°C. Single colonies were streaked to LB-Cm plates and incubated overnight at 42°C to induce integration of the plasmid into the genome. Large colonies representing clones that have the integration were re-streaked on LB-Cm plates and grown overnight at 42°C to purify the large colony

away from small ones. At this stage single large colonies were selected and used to inoculate 5 ml LB cultures which were incubated overnight at 200 rpm and 30°C. Serial dilutions of each of these cultures (usually 100 µl of 10^{-5} and 10^{-6} dilutions) were plated on LB + 5% sucrose and incubated overnight at 30°C. pTOF24 carries the SacB gene which makes cells sensitive to sucrose. Growth in the presence of sucrose therefore induces excision by recombination of the integrated cassette and loss of the excised plasmid. Since sucrose is not 100% effective as a selectable marker, 100 colonies (or more if necessary) were replica-plated onto LB with 5% sucrose and LB-Cm with 5% sucrose and grown overnight at 30°C. Colonies that grow on LB-sucrose but not on LB-Cm-sucrose are ones that have lost the genomic insertion and the plasmid. These were then tested by colony PCR using Hfq-del-F1 and Hfq-del-R2 primers (2.1.8) to check for the presence of the genomic deletion.

2.2.5 Transformation of *S. cerevisiae*

Yeast cells were transformed using the method of Gietz *et al.* or slight adaptations thereof (Gietz *et al.*, 1992). Different versions of this method were used for the transformation of plasmid DNA and for the transformation with linear DNA for integration into the genome.

2.2.5.1 Yeast transformation with plasmid DNA

Method 1, standard transformation with plasmid DNA

A single colony was taken and inoculated into 5 ml of appropriate liquid medium (Table 2.1.2). This was incubated overnight at 30°C until the cells had grown into stationary phase. The cells were sedimented in a Mistral 1000 centrifuge (3 min, 3,500 rpm, RT). Cells were then washed two times with 1 ml of deionised water and sedimented in an Eppendorf microfuge (1 min, 3,500 rpm) and finally resuspended in 1 ml of deionised water. 50µl of cell suspension was mixed with 240 µl of 40% PEG_{3350}, 36 µl of 1 M LiAc, 200-800 ng of plasmid DNA (1 µl from a plasmid prep; 2.3.2.1) in 29 µl of deionised water and 50 µg (5 µl of 10 mg/ml) of denatured salmon sperm carrier DNA (Gibco BRL). This was vortexed briefly to mix and incubated at room temperature for 30 min on a rotating wheel. Following this, cells were heat-shocked at 42°C for 15 min, then sedimented (14,000 rpm, 5 s, Eppendorf

microfuge). The pelleted cells were then resuspended in 150 µl of deionised water and plated onto the appropriate solid medium to select for transformants. Plates were incubated at 30°C for 2-3 days, or longer if required for slow growing strains.

Method 2, 'Lazy Bones' plasmid transformation of yeast colonies

This method (Elble, 1992) was used when in hurry or when the appropriate selective liquid medium to grow the strain to be transformed was not available. One fresh colony (2-3 mm in diameter) was picked from a plate with a sterile pipette tip and transferred directly into the transformation mixture containing the plasmid DNA (360 µl total volume as for standard method; see above). The cells were resuspended by vortexing and incubated at room temperature for 5 min. They were then heat shocked for 15 min at 42°C, pelleted and plated as above.

2.2.5.2. Yeast transformation for genomic integration of DNA

Preparing DNA for genomic integration

In order to prepare DNA for transformation and integration into the yeast genome, a linear DNA fragment was produced by PCR (2.3.2.7). The PCR template was chosen to contain the relevant DNA sequence required following the integration. The primers were designed so that the final PCR product (the integration cassette) would be flanked at both ends by approximately 45 bp of sequence identical to the DNA immediately 5' and 3' to the sequence to be replaced or to the site of integration. At least one 50 µl PCR (but up to five depending on the yield) was ethanol precipitated (2.3.1.3) and used to transform the relevant strain (2.1.6).

High efficiency yeast transformation

A high efficiency transformation protocol based on that of Gietz and Woods was used for the transformation with DNA for genomic integration (Gietz and Woods, 2002). A single colony was taken and inoculated into 10 ml of appropriate liquid medium (Table 2.1.2). This was incubated overnight until the cells had entered stationary phase. This culture was used to inoculate 50 ml of fresh medium to an OD_{600} of approximately 0.1. This was then incubated at 30°C until the OD_{600} of the culture was 0.8-1.0. The cells were sedimented in a Mistral 1000 centrifuge (3 min,

3,500 rpm, RT), washed once with 20 ml, once with 10 ml of deionised water and sedimented as before. Cells were finally resuspended in 0.5 ml of deionised water. 50 µl of cell suspension was mixed with ice-cold transformation mixture (240 µl of 40% PEG_{3350}, 36 µl of 1 M LiAc, DNA in 29 µl deionised water (generally 1-5 µg; see below) and 100 µg (10 µl of 10 mg/ml) of denatured salmon sperm carrier DNA (Gibco BRL). This was vortexed briefly to mix and incubated at 42°C for 40 min, then sedimented (14,000 rpm, 5 s, Eppendorf microfuge). The cells were resuspended in 150 µl of 1x TE (Table 2.1.4) and plated onto the appropriate solid medium to select for transformants in case of the use of a nutritional marker. For antibiotics markers (HpH, NAT, KAN) cells were resuspended in YPDA and incubated for 3 h at room temperature before spreading onto selective medium or spread directly on YPDA-agar plates, incubated overnight at 30°C and replica-plated onto plates containing the appropriate antibiotics. Plates were incubated at 30°C (or the appropriate temperature where temperature-sensitive strains were used or being created) for up to two weeks.

Screening colonies for genomic integration

Colonies growing on the appropriate selective medium were re-streaked onto fresh solid medium and the integration was investigated using colony PCR (2.3.2.8) and, where relevant, Western blotting (2.4.1 and 2.4.5). Positive colonies were streaked twice onto selective medium taking a single colony each time to ensure a pure recombinant strain before use in experiments.

2.2.6 Yeast sporulation and tetrad dissection

Fresh colonies of the diploid strain were streaked thickly onto pre-sporulation medium (2.1.2.3) and grown for 1-2 days at 30°C. Cells were scraped off the surface using a streaking loop and streaked thinly onto sporulation medium (2.1.2.3). These were incubated for 3-7 days at 30°C to allow the cells to sporulate. The efficiency of sporulation was investigated microscopically by looking for the presence of tetrads. When a satisfactory number of tetrads was observed (25% or more) one loopful of cells was scraped off the plate and resuspended in a 0.2 mg/ml cytohelicase solution (Biosepra) and incubated overnight on ice to digest tetrad sporewalls. The next day 2

μl of this suspension was spotted onto a selective plate and a sterile toothpick was used to spread the cells and spores from this point in a straight line down the middle of the plate. The Singer MSM series 100 micromanipulator was now used to look for tetrads with properly digested sporewalls. These were picked up with the needle and dissected, positioning each of the four spores on the plate at set coordinates. The plate was subsequently incubated at the appropriate temperature to allow for spores to form single colonies. Single colonies were then tested on appropriate selective media for the presence or absence of markers, e.g. those corresponding to a deletion.

2.2.8 Mating type determination

Following tetrad dissection, haploid progeny (usually Leu-, Lys+) were streaked in lines on the appropriate solid medium (2.1.2.3). KY117 and KY118 cells (Leu+, Lys-) (2.1.6) were simultaneously grown on YPDA medium by streaking in lines. Plates were incubated overnight at 30°C and replica-plated to SD-Leu-Lys with the lines of the strains to be tested perpendicular to those of the KY117 and KY118 strains. Plates were incubated at 30°C. Only those cells which had undergone successful mating (at the intersection point of two lines) were able to grow. The mating type of the haploid progeny could thus be determined since only *MAT***a** cells can mate with KY118 (*MATα*), and only *MATα* cells can mate with KY117 (*MAT***a**).

2.2.9 Growth curves

Growth curve analysis was performed to compare growth rates of different yeast strains and in case of temperature-sensitive strains to test their reaction to a shift in temperature. Cells were grown in liquid media to mid-logarithmic phase under permissive conditions. Cultures were maintained in logarithmic growth by diluting with pre-warmed growth medium to ensure that OD_{600} readings were below 0.8. The growth rate was measured by recording the OD_{600} at regular intervals, typically every 1 or 2 hours. To test temperature-sensitive strains logarithmic cultures were shifted to non-permissive temperature by diluting the culture with pre-heated medium giving the resulting culture the required temperature whilst keeping the culture within the logarithmic phase. They were then incubated at the appropriate temperature and the growth rate was measured as described.

2.2.10 Spot assays

To compare growth of different yeast strains under a variety of conditions spot assays were used. Small cultures (5 – 10 ml) were grown overnight. The OD_{600} was measured for each and serial dilutions were made in the appropriate medium, corresponding to OD_{600} readings of 0.0003, 0.003, 0.03, 0.3 (and 3). These dilutions (200 µl each) were made in sterile 96-well flat bottom plates (Bibby Sterilin, Ltd.) and a metal "hedgehog" was used to spot these on the appropriate solid medium. This could be repeated on various media and plates could be incubated at various temperatures to allow for growth of these virtual replicas to be compared.

2.2.11 Plasmid shuffle on 5-FOA

Plasmids containing a *URA3* marker were replaced with plasmids containing a different nutritional marker by streaking a colony containing both plasmids on 5-FOA containing plates (2.1.2.3). Strains that grew on 5-FOA plates were put through a second round of selection by streaking one colony on 5-FOA again. Resulting colonies were tested for the absence of the *URA3* marker and the presence of the other marker by streaking them on SD-Ura and the appropriate SD droupout plates.

2.3 Nucleic acid methods

Unless otherwise stated, all DNA and RNA manipulations were based on those described by Ausubel *et al.* (Ausubel *et al.*, 1994).

2.3.1 General methods

2.3.1.1 Spectrophotometric quantification of nucleic acids

The concentration of DNA or RNA solutions was determined by measuring the absorption of diluted solutions at 260 nm using a Cecil CE 2040 spectrophotometer, and a quartz cuvette (10 mm light path, Sigma). Dilutions were made in 80 µl of deionised water aiming for a reading of 0.1-1.0. For double-stranded DNA, an A_{260} value of 1.0 represents a DNA concentration of approximately 50 ng/µl. For single-stranded RNA an A_{260} value of 1.0 represents an RNA concentration of

approximately 40 ng/µl. The purity of the nucleic acid sample was determined by measuring the absorption at both 260 and 280 nm. Protein-free preparations of DNA and RNA give an A_{260}:A_{280} ratio between 1.8 and 2.2. The A_{260}:A_{230} ratio should be higher than 1.7 indicating the absence of carbohydrate, lipid and phenol.

2.3.1.2 Phenol/chloroform extraction

Nucleic acids were purified away from proteins in aqueous solutions by extraction with phenol and chloroform. For DNA this was done by adding an equal volume of phenol:chloroform:isoamyl alcohol in a 25:24:1 ratio (equilibrated with 10mM Tris, 1 mM EDTA, pH8.0). For RNA an equal volume of phenol:chloroform in a 5:1 ratio (acid equilibrated, pH 4.7) was added. The mixture was vortexed thoroughly and centrifuged (14,000 rpm, 3 min). The upper, aqueous, phase containing nucleic acids was moved to a fresh container, being careful not to transfer any of the lower phase.

2.3.1.3 Precipitation of nucleic acids

Nucleic acids were precipitated from aqueous solution by adding 0.1 volumes of 3 M sodium acetate (pH 5.2) and 2-2.5 volumes of ethanol for DNA or 2.5-3 volumes of ethanol for RNA. After thorough mixing these were incubated at -70°C for a minimum of 60 min. Precipitated nucleic acids were then sedimented by centrifugation (14,000 rpm, 15-30 min, 4°C) and the resulting pellet was washed with 70% (v/v) ethanol. The pellet was then air-dried and dissolved in an appropriate volume of deionised water or TE.

2.3.2 DNA methods

2.3.2.1 Small scale preparation of plasmid DNA

Plasmid DNA was prepared using the QIAprep Spin Miniprep Kit (Qiagen), following the manufacturer's instructions. DNA was extracted from 3 ml of stationary phase *E. coli* culture, eluted in 50 µl of deionised water or TE and stored indefinitely at –20°C.

2.3.2.2 Large scale preparation of plasmid DNA

Plasmid DNA was prepared using the HiSpeed Plasmid Maxi Kit (Qiagen), following the manufacturer's instructions. DNA was extracted from 150 ml of stationary phase *E. coli* culture for high-copy plasmids or 250 ml for low-copy plasmids. The volumes of buffers P1, P2 and P3 were doubled for low-copy plasmids. Plasmid DNA was finally eluted in 500 µl of deionised water or TE and stored indefinitely at –20°C.

2.3.2.3 Plasmid rescue from yeast

To rescue plasmid DNA from yeast the Zymoprep kit (Zymo Research) was used according to the manufacturer's instructions. Plasmid DNA was extracted from 1ml of full-grown yeast culture. The isopropanol precipitated DNA was finally dissolved in 20 µl TE. Between 1 and 5 µl of this preparation was used to transform *E. coli* (2.2.3.2).

2.3.2.4 DNA restriction digests

DNA was digested by restriction endonucleases (mostly NEB) in the appropriate buffer (supplied and recommended by the manufacturer). To perform restriction analysis on a plasmid the digest was usually performed in 10 µl with 0.5-1 µg of DNA and 2-10 units of the appropriate restriction enzyme(s), ensuring that the total enzyme volume did not exceed 10% of the total reaction volume. The digest was incubated at the appropriate temperature, typically for 1-2 hours. To digest plasmids or PCR products to be used in later cloning steps, the digest was usually performed in 50 µl with 2 – 5 µg of DNA and 1 µl of the appropriate enzyme(s) and incubated overnight at the appropriate temperature. For double digests the manufacturers' guidelines were followed with regards to the optimal buffer conditions for both enzymes (NEB catalogue). Where necessary, sequential digests were performed, cleaning up the DNA in between. After digestion DNA was gel (2.3.2.13) or column (2.3.2.11) purified.

2.3.2.5 Removal of phosphates from DNA ends

Plasmid DNA digested with a single restriction endonuclease to be used for cloning was dephosphorylated to prevent recircularisation during ligation. To this end 5 U of Calf Intestinal Alkaline Phosphatase (CIP, NEB) were added directly after digestion (2.3.2.4). This was incubated for 30 min at 37°C to remove terminal phosphate groups from the DNA. DNA was then gel (2.3.2.13) or column (2.3.2.11) purified.

2.3.2.6 Yeast genomic DNA preparation for PCR

Yeast genomic DNA was prepared using one of the following two methods. The first method is quicker and easier; the second is more likely to give successful results for difficult PCRs (e.g. amplification in the centromeric region).

Method 1

Cells from a 10 ml stationary phase culture were sedimented in 1.5 ml portions in microfuge tubes (3,500 rpm, 1 min, RT). Cells in each tube were resuspended in 100 µl of freshly prepared lysis solution (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% SDS, 1.4 M β–Mercaptoethanol) and vortexed for 15 minutes. 500 µl of TE was added and the mixture was extracted with 500 µl of phenol (pH 8.0). The aqueous phase was transferred to fresh microfuge tubes, 50 µl of 3 M sodium acetate (pH 5.2) was added and DNA was precipated by mixing with 600 µl of isopropanol and 1 hour incubation at -20°C. After centrifugation for 20 min at 14,000 rpm at 4°C the resulting pellet was washed twice with 70% ethanol. The pellets were then air-dried and each pellet was dissolved in 20 µl of TE with 0.2 mg/ml RNase A (DNase free, Roche). Routinely 0.5-1 µl of this solution was used as a template in a 50 µl PCR.

Method 2

Five ml of stationary phase culture was sedimented by centrifugation and washed with one ml of deionised water. The cells were resuspended in 360 µl of spheroplasting solution (0.9 M sorbitol, 0.1 M EDTA pH 8.0) and 15 µl of 1 M DTT was added together with 15 µl of zymolase (15 mg/ml; Seikagaku Corporation, Tokyo, Japan). This was incubated at 37°C for 30 – 60 min to break down the cell walls. The cells were centrifuged and resuspended in 300 µl of lysis solution (50 mM

Tris-HCl, 50 mM EDTA pH 8.0) and mixed with 60 µl of 10% SDS. This was incubated at 65°C for 5 min before adding 100 µl of 5 M potassium acetate. This mixture was incubated on ice for 5 min to allow proteins to precipitate. Proteins were sedimented by centrifugation for 2 min at 14,000 rpm. The supernate was transferred to fresh microfuge tubes and centrifugation was repeated if necessary to lose all precipitated proteins. To precipitate the DNA 300 µl of isopropanol was added. After thorough mixing the DNA was sedimented by centrifugation for 2 min at 14,000 rpm. The pellet was washed with 500 µl of 70% ethanol and left to air dry. The dried DNA pellet was finally resuspended in 50 µl of TE and 0.5 – 1 µl of this was used as a template in a 50 µl PCR.

2.3.2.7 Amplification of DNA by Polymerase Chain Reaction

Specific regions of DNA were amplified using the polymerase chain reaction (PCR). Specifically designed oligonucleotide primers (2.1.8 and 2.3.2.10) were annealed to denatured template DNA (1-10 ng of plasmid DNA or genomic DNA (2.3.2.6)) and extended by a thermostable DNA polymerase. A typical PCR was assembled as follows:

5 µl	10x DNA Polymerase Buffer with $MgCl_2$
5 µl	2.5 mM dNTPs (dATP, dCTP, dGTP, dTTP; Roche)
1 µl	Forward oligonucleotide primer (10-50 µM)
1 µl	Reverse oligonucleotide primer (10-50 µM)
1 µl	Template DNA
1 µl	Thermostable DNA polymerase
36 µl	Sterile deionised water

Depending on the type and purpose of PCR different DNA polymerases were used: *Taq* (Roche, 2 U/µl) and TaKaRa LA *Taq* (Takara Bio Inc., 5 U/µl) for highly efficient amplification (at 72°C) but with relatively high error rates; *Pfu* or *Pfu* Turbo (Promega, 3 U/µl) and Expand Long Template (Roche, 5 U/µl) for amplification (at 68°C) with low error rates. For amplification of DNA with a high GC content (e.g. the NAT cassette) 5% DMSO was added to the PCR. All PCRs were performed in a

Biometra UNO II Thermoblock or an MJ Research PTC-200 Peltier Thermal Cycler programmed according to the length of the desired product and the annealing temperature of the oligonucleotide primers being used[†]. A typical PCR cycle was as follows and repeated 25-35 times:

Denaturation		1min	94°C
25 - 35x	Denaturation	30 s	94°C
	Annealing	30s	40-60°C[†]
	Extension	1 min per kb at 72°C or	
		1 min per 500 bp at 68°C	

[†] For oligonucleotides up to 25 nt in length: T_m (°C) = 4(G+C) + 2(A+T)

2.3.2.8 Colony PCR

To check *E. coli* colonies for plasmids with a specific insert (in a specific orientation) or for the presence of a genomic deletion (2.2.4) PCR was performed directly using the colony as a source of template. Yeast colonies were similarly checked for the presence of genomic insertions or deletions. To this end a 10 µl tip was used to lift (part of) a colony from a plate. Each colony was resuspended in 20 µl PCR mix and run using the standard PCR protocol (2.3.2.7), instead starting with a 6 min heating step at 96°C (to break the cells) and denaturing at 96°C for each cycle thereafter. The PCR mix was made by mixing 15 µl of water, 2 µl of 10x PCR buffer with $MgCl_2$, 2 µl of 2.5 mM dNTP mix, 0.4 µl of 10 µM forward primer, 0.4 µl of 10 µM reverse primer, 0.2 µl of *Taq* DNA Polymerase (Roche, 2 U/µl; or Fast Start *Taq* DNA Polymerase, Roche, 5 U/µl; the latter works particularly well for colony PCR because of the initial heat activation step) or multiples thereof for each colony to be checked.

2.3.2.9 Site directed mutagenesis

Site directed mutagenesis was performed based on the QuikChange method (Stratagene). 50 ng of plasmid (usually 1 µl of a 1 in 10 dilution of plasmid miniprep) was used for a PCR in which the entire plasmid was amplified by primers with imperfect complementarity to the region of the gene to be mutated. These

primers typically contained around 40 bp of sequence complementary to the target gene, as well as nucleotide changes giving the desired mutation (usually located close to the middle of the primers). The forward and reverse primers were perfectly complementary to each other, with a minimum annealing temperature of 78°C (2.3.2.10). Another strategy can be used (particularly useful for creating deletions) in which the forward and reverse primers only overlap partially (15 – 20 nt in the region to be mutated; see (Boon, 2005). The product of this type of PCR is a nicked circle, with the nicks occurring at the end of each synthesised strand.

A typical reaction mix was as follows:

5 µl	10x DNA Polymerase Buffer with $MgCl_2$
5 µl	2.5 mM dNTPs (Roche)
1 µl	Plasmid (50 ng)
1 µl	Forward primer (10 µM)
1 µl	Reverse primer (10 µM)
1 µl	DNA polymerase
36 µl	Water

Pfu polymerase (Promega, 3 U/µl) was used routinely. If this was not successful *Pfu* Turbo (Stratagene, 2.5 U/µl; more efficient amplification, especially for long (>10 kb) products) or Expand Long Template Enzyme Mix (Roche, 5 U/µl; buffer 2 with 2.75 mM $MgCl_2$; more efficient amplification, but higher error rate) was used. A typical PCR cycle would be as follows:

Denaturation		95°C	1 min
18x	Denaturation	95°C	30 sec
	Annealing	55°C	1 min (but sometimes as low as 50°C)
	Extension	68°C	2 min per kb
Extension		68°C	10-20 min

The PCR product was incubated with 1 µl of DpnI (NEB, 20 U/µl) at 37°C for a minimum of 2 h. DpnI digests dam-methylated DNA, so that parental plasmid is digested, but not non-methylated PCR-product. At this stage 10 µl of the DpnI

treated sample could be run on an agarose gel (2.3.2.12) to see if any correctly sized product was present. Depending on the amount visible on the gel, 1-10 µl of the DpnI treated DNA was used to transform *E. coli* (2.2.3.2 or 2.2.3.4). Resulting colonies were screened for presence of the desired mutation by restriction digest (2.3.2.4), colony PCR (2.3.2.8) and/or by sequencing (2.3.2.15).

2.3.2.10 Primer design

Primers to amplify DNA for cloning

Primers were designed by eye, aiming for a T_m of around 50°C (40-60°C) taking care that the melting temperatures of two primers to be used in the same PCR did not differ by more than 5°C. The formula T_m (°C) = 4(G+C) + 2(A+T) was used for primers up to 25 nt in length. Where possible, primers were designed to terminate in at least one G or C base at the 3' end. Restriction enzyme cleavage sites were included at the 5' end of primers with an additional number of random nucleotides (usually A) added to the 5' end to allow for efficient digestion of the PCR product. For this, the NEB catalogue (2002-2003; p.242-243) was used as a reference.

Primers for site-directed mutagenesis

Mutagenic primers for use in site-directed mutagenesis were designed according to the desired mutation. The following considerations were made for their design. First, both mutagenic primers contain the desired mutation and anneal to the same sequence on opposite strands of the plasmid. Second, primers should be between 25 and 45 bases in length, with a melting temperature (T_m) of ≥78°C (see below), although primers longer than 45 bases were also used successfully. Third, the desired mutation (deletion or insertion) should be near the middle of the primer with a minimum of ~10–15 bases of correct sequence on either side. Fourth, the primers optimally should have a minimum GC content of 40% and should terminate in one or more C or G bases. The following formula was used to estimate the T_m of mutagenic primers:

$$T_m = 81.5 + 0.41(\%GC) - 675/N - \%\ \text{mismatch}$$

in which N is the primer length in bases and the values for %GC and % mismatch are whole numbers. To calculate T_m for primers intended to introduce insertions or deletions, a modified version of the above formula was used:

$$T_m = 81.5 + 0.41(\%GC) - 675/N$$

in which N does not include the bases which are being inserted or deleted. Alternatively the QuikChange T_m calculator was used (available online at http://www.stratagene.com/QPCR/tmCalc.aspx).

2.3.2.11 Purification of PCR products and restriction digests

To purify PCR products and endonuclease digested DNA away from salts, dNTPs, oligonucleotides and/or other DNA fragments shorter than 100 bp the QiaQuick PCR Purification kit (Qiagen) was used according to the manufacturer's instructions. DNA was eluted in 30-50 µl of water or TE buffer and stored indefinitely at -20°C. Alternatively PCR reaction mixes or digests were phenol/chloroform extracted (2.3.1.2) and/or ethanol precipitated (2.3.1.3) or separated by size by agarose gel electrophoresis (2.3.2.12) and gel purified (2.3.2.13).

2.3.2.12 Agarose gel electrophoresis

Plasmid DNA and DNA fragments produced by restriction endonuclease digestion (2.3.2.4) or generated by PCR (2.3.2.7) were analysed in 0.8-2% (w/v) agarose gels. Gels were prepared by dissolving agarose (SeaKem LE Agarose, Cambrex) in 100 ml of 1x TAE buffer (2.1.3) by boiling and adding ethidium bromide to a final concentration of 0.5 µg/ml before pouring the mixture into a horizontal gel tray (BRL, Horizontal System for Submerged Gel Electrophoresis, Model H5) and allowing it to set. Ficoll loading buffer (6x Ficoll loading buffer: 15% (w/v) Ficoll 400, 0.025% (w/v) Bromophenol Blue (BPB), 0.025% (w/v) Xylene Cyanole FF (XCB) and 50 mM EDTA pH 8.0) was added to samples to be analysed. The samples were loaded in the gel and 100-150 V were applied across the gel separating the DNA fragments by size. The DNA was subsequently visualised using a ChemiGenius 2 UV transilluminator (Syngene, Cambridge, UK) and Genesnap 6.03

software to process the image. DNA markers (1kb PLUS DNA Ladder (Invitrogen) or 100bp DNA Ladder (NEB)) with fragments of known sizes were separated simultaneously, allowing fragment sizes to be estimated.

2.3.2.13 Purification of DNA from agarose gels

DNA fragments produced by endonuclease digestion (2.3.2.4) or generated by PCR (2.3.2.7) were separated according to size by agarose gel electrophoresis (2.3.2.12). Fragments were visualised on a UV transilluminator (Herolab UVT-28M at low intensity setting) and the desired band was cut from the gel using a clean surgical blade. The DNA was then isolated from the agarose gel using the QiaQuick Gel Extraction kit (Qiagen) according to the manufacturer's instructions. DNA was eluted in 30-50 µl of sterile deionised water or TE and stored indefinitely at -20°C.

2.3.2.14 Ligation of DNA molecules

Ligations were performed using the Fast-Link DNA Ligation Kit (Epicentre Biotechnologies) according to the manufacturer's instructions. Briefly, 100-200 ng of vector DNA, 2 to 5 times this molar amount of insert DNA, 1 mM ATP and 2 U of Fast-Link DNA Ligase were incubated in 1x Ligation Buffer in a final volume of 15 µl (30 min to 16 h, RT). Some or all of this ligation mixture was then used for transformation into *E. coli* (2.2.3.2 or 2.2.3.4). Oligonucleotide linkers (e.g. HA and Myc tag inserts for pMR81 and pMR82) were cloned using ≥100:1 molar ratio of linker to plasmid (with approximately 0.1 pmol of linker) and incubating for 16 h at 16°C. Complementary linker oligonucleotides were mixed at a 1:1 molar ratio (50µM each) in TE buffer, heated to 95°C and slowly cooled to room temperature to allow duplex to form, before using them for ligation.

2.3.2.15 De-salting small amounts of DNA

In preparation for transformation by electroporation ligation mixes (2.3.2.14) were de-salted by floating a Millipore 0.025 µm Type VS membrane on a petridish filled with deionised water (shiny side up) with the ligation mix spotted on top and incubating them for a minimum of 30 min at room temperature.

2.3.2.15 DNA sequencing

Plasmid DNA to be sequenced was purified (2.3.2.1) and quantified by spectrophotometry (2.3.1.1) and/or comparison to a marker by visualisation on an agarose gel (2.3.2.12). Sequencing reactions were performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). A typical reaction mix was as follows:

Template DNA (200-500 ng)	2 µl
BigDye Terminator mix	2 µl
Better Buffer (Web Scientific Ltd.)	3.3µl
Primer (1.6 µM)	1 µl
Water	1.7µl

Reactions were performed in a Biometra UNO II Thermoblock or an MJ Research PTC-200 Peltier Thermal Cycler for a minimum of 25 cycles:

Denaturation		3min	96°C
>25x	Denaturation	15s	96°C
	Annealing	10s	50°C
	Extension	4 min	60°C

Samples were processed and run at the School of Biological Sciences Sequencing Service (SBSSS) on an Applied Biosystems 3730 DNA Analyzer. Sequence data were analysed using Sequencher 4.2 or 4.5 (Gene Codes Corp.), and later using ContigExpress (Invitrogen, Vector NTI Advance 10).

2.3.2.16 End-labelling of oligonucleotides

Oligonucleotides were 5'-end radiolabelled with [γ-^{32}P] ATP (Redivue, 10µCi/µl; ~6,000 Ci/mmol; Amersham Biosciences) using T4 Polynucleotide Kinase (PNK, NEB) and the provided 10x PNK buffer. Variations of the recommended protocol were used depending on the downstream application: Northern analysis (2.3.3.7), primer extension (2.3.3.8) or strand displacement assays (2.5.3). To label an oligonucleotide probe for Northern analysis a 20 µl reaction was performed in 1x T4 PNK buffer with 10 pmol of oligonucleotide, 2 µl of [γ-^{32}P] ATP and 5 U of T4

PNK (NEB). To label primers for primer extension reactions were performed in 10 µl in 1x T4 PNK buffer with 10 pmol of oligonucleotide, 1 µl of 100mM DTT and 2 µl of [γ-^{32}P] ATP. To prepare substrate for strand displacement assays oligonucleotides were labelled in 20 µl reactions: 2 µl of 33 µM oligo, 2 µl of 10x PNK buffer (NEB), 13 µl of water, 1 µl of PNK (10 U/µl, NEB) and 2 µl of [γ-^{32}P] ATP. All labelling reactions were incubated at 37 °C for 45 min to 1 h. For use in Northern probing the labelled oligonucleotide was passed through a 0.45 µm filter in 2 – 3 ml of hybridisation buffer to remove any particulate matter. For use in primer extension and strand displacement assays the labelled oligonucleotides were purified by gel filtration (2.3.2.19) in a total volume of 20 µl. Labelling efficiency was estimated and generally between 30 and 60%.

2.3.2.17 Making DNA probes for Northern analysis

Method 1, DNA oligomer end-labelling

Ten pmol of oligonucleotide was end-labelled in a 20 µl PNK reaction (2.3.2.16) and passed through a 0.45 µm filter (Minisart Cellulose acetate, Sartorius) in 2 – 3 ml of SES1 hybridisation buffer (2.1.3) to remove any particulate matter.

Method 2, Single stranded DNA labelling by PCR

A 50 µl reaction was assembled in 1x *Taq* DNA polymerase buffer with 20 ng of template DNA (purified, double stranded PCR product (2.3.2.7) of the target sequence), 0.2 mM each of dATP, dGTP and dTTP, 1.2 µM dCTP, 10 pmol of primer (reverse complement to the sequence to be probed), 5 µl of [α-^{32}P] dCTP (Amersham) and 2.5 U of *Taq* DNA polymerase (Roche). This was incubated for 35 cycles (30 s 94°C, 30 s 40 - 60°C (2.3.2.10), 1 min/kb 72°C) in a Biometra UNO II Thermoblock. The labelled single stranded DNA probe was purified by gel filtration (2.3.2.19) and heated at 95°C for 2 min before adding it to the Northern blot for hybridisation (2.3.3.7).

2.3.2.18 Radioactive labelling of DNA size marker

Labelled pBR322 DNA- *Msp*I digest (NEB) and 100bp DNA ladder (NEB) were used as size markers. A 10 µl reaction was assembled and incubated for 15 - 30 min at room temperature:

1 µl marker DNA (1 µg /µl)

1 µl 10x Filling buffer (Roche)

1 µl 3 dNTPs mix (2.5 mM each of dATP, dGTP and dTTP; Roche)

3 µl [α-^{32}P] dCTP (Redivue, 10 µCi/µl; ~6,000 Ci/mmol, Amersham)

3.5 µl water

0.5 µl Klenow (Roche, 2 U/µl)

The reaction was stopped by adding 1 µl of 0.5 M EDTA pH 8.0 and 40 µl of TE. The labelled ladder was purified by gel filtration (2.3.2.19). A small volume (2 µl) of a large dilution equivalent to about 50 cps/µl was usually enough to give a good picture after denaturing PAGE (2.3.3.3).

2.3.2.19 Purification of labelled DNA from unincorporated nucleotides

Radioactively labelled DNA molecules were separated from unincorporated nucleotides by gel filtration using Mini Quick Spin Oligo columns (Roche) following the manufacturer's instructions. In brief, the column matrix was resuspended by gentle flicking. The column was packed by centrifugation for 1 min at 1,000 g. The column was transferred to a clean microfuge tube and the labelling reaction (20-50 µl) was applied to the centre of the column. After 4 min centrifugation at 1,000 g the flow-through contained labelled nucleic acids whereas unincorporated nucleotides were stuck on the column. Efficiency of labelling could be estimated by comparing the number of counts in both.

2.3.3 RNA methods

2.3.3.1 Three methods for total RNA isolation from yeast

Method 1, RNA isolation for primer extension

Total yeast RNA was prepared using the method of Schmitt *et al.* (Schmitt *et al.*, 1990). A 20 ml culture was grown in YPDA (2.1.2.3) to OD_{600} 0.5 – 1.0. Cells were sedimented by centrifugation for 3 min at 3,500 rpm (Mistral 1000 centrifuge). The pellet was washed with 400 µl of AE buffer (50 mM Na acetate pH 5.2, 10 mM EDTA), sedimented and resuspended in 400 µl of AE buffer. Then 40 µl of 10% SDS was added together with 500 µl of phenol (equilibrated with AE buffer). This was vortexed two times for 10 s at a time, followed by 4 min incubation at 65°C and snapfreezing on dry ice. After thawing and 2 min centrifugation at 20,800 g in a microfuge at 4°C the aqueous phase was phenol/chloroform extracted (2.3.1.2). Finally one tenth of a volume of 3 M Na acetate (pH 5.2) was added together with 2.5 volumes of 100% ethanol. The RNA was precipitated at -70°C overnight, sedimented by centrifugation, washed with 70% ethanol, air-dried and finally dissolved in 50 µl RNase-free TE.

Method 2, for Northern analysis

A 50 ml yeast culture was grown in the appropriate medium (2.1.2.3) and at the appropriate temperature to an OD_{600} of 0.5. The cells were sedimented by centrifugation for 5 min at 1,000 g at 4°C. The pellet was washed with 20 ml of ice-cold water and centrifuged again. In a 4°C room the following were added ice-cold: 0.5 ml of GTC mix, 0.5 ml of phenol and 1 ml of glass beads (acid-washed, Sigma). This was vortexed for 5 min (2 times 2 min and 1 time 1 min with 2 min cooling on ice in between) after which another 7.5 ml of ice-cold GTC mix and 7.5 ml of phenol were added. The mixture was incubated for 5 min at 65°C and subsequently allowed to cool on ice for 5 min. Then, 4 ml of NaAc mix (0.1 M Na acetate, 1 mM EDTA and 10 mM Tris-HCl pH8.0) and 7.5 ml of chloroform were added, vortexed briefly and centrifuged for 5 min at 4,600 rpm at 4°C (4,570 g; Sorvall Legend RT). Ten ml of the aqueous phase was removed and added to 10 ml of cold PCI (25:24:1 pH 4.7), vortexed and centrifuged for five min at 4,600 rpm at 4°C. This was repeated with 8 ml of the aqueous phase (plus an equal volume of PCI) and finally 6ml of the aqueous phase was mixed with 20 ml of ice-cold 100% ethanol. After mixing by inversion this was incubated for a minimum of 30 min at -70°C. The precipitated RNA was sedimented by centrifugation for 30 min at 4,600 rpm at 4°C, washed with

5 ml ice-cold 70% ethanol, air-dried and finally dissolved in 100 – 200 µl of DEPC-treated water.

GTC mix: 100 g of Guanidine thiocyanate, 100 ml of DEPC-treated water, 10.6 ml of 1 M Tris-HCl pH8.0 and 4.24 ml of 0.5 M EDTA pH8.0 were mixed and left to dissolve overnight with agitation. This mixture was then heated for 10 min at 60 -70°C and finally 21.2 ml of 20% L.Lauroylsarcosine and 2.1 ml of β-Mercaptoethanol were added

Method 3, for Northern analysis of non-denatured RNA

To look at levels of U4/U6 duplex by Northern anaylsis, total RNA was isolated under non-denaturing conditions using the method described by Lygerou *et al.* (1999). Cells were grown in the appropriate medium to OD_{600} 0.5 – 1.0. Approximately 10^8 cells (10 units at OD_{600}) were then collected, washed with ice-cold deionised water and resuspended in ice-cold 250 µl of RNA extraction buffer (100 mM LiCl, 1 mM EDTA, 100 mM Tris-Cl pH 7.5, 0.2% (w/v) SDS). Cells were broken in a cooled Thermomixer Comfort (Eppendorf) for 15 min at 4°C with 250 µl of phenol-chloroform (5:1, pH 4.7) and 100 µl of Zirconia beads (Ambion). The aqueous phase was mixed with an equal volume of 2x native RNA loading buffer or 2x formamide loading buffer for native (2.3.3.4) or denaturing PAGE respectively (2.3.3.3)

2.3.3.2 *In vitro* transcription and RNA purification

Template preparation

Three different kinds of templates were used for *in vitro* transcription. For *in vitro* transcription of *ACT1* pre-mRNA p283 plasmid DNA (2.1.7) was digested (2.3.2.4) and purified (2.3.2.13). For *in vitro* transcription of snRNAs or parts of snRNAs templates were created by PCR (2.3.2.7) or by annealing two oligonucleotides (Table 2.1.8.3a,b). In each case a minimal T7 promoter sequence (TAATACGACTCACTATAGG) was fused upstream of the sequence to be transcribed. By PCR this was accomplished with a forward primer which contained the T7 promoter sequence followed by the first nucleotides of the sequence (giving an annealing temperature between 48 – 56°C) and a reverse primer complementary to

the last nucleotides of the sequence (with the same annealing temperature). The PCR product was purified (2.3.2.11), taking care to keep it RNase free. In case of an oligonucleotides template, a DNA oligomer was designed by taking the reverse complement of the sequence that covered the T7 promoter sequence and the desired sequence (Milligan and Uhlenbeck, 1989). This oligomer (12.5 µl of 100 µM) was mixed with an equal amount of DNA oligomer covering the full T7 promoter sequence (T7F2, Table 2.1.3.7a). The mixture was heated at 90°C for 3 minutes and snap-chilled on ice. The resulting duplex was diluted in 225 µl TE to a final concentration of 5 µM and used directly in the *in vitro* transcription reaction or stored at -20°C until required.

The in vitro transcription reaction

For each 100 µl reaction the following were mixed according to the NEB T7 polymerase product information:

10 µl 10x T7 RNA polymerase reaction buffer (NEB)

10 µl 100mM DTT

10 µl DNA template

10 µl 20 mM NTPs mix (10mM ATP, CTP, GTP and UTP; Amersham)

1 µl BSA (10 mg/ml)

2.5 µl RNasin (Promega, 40 U/µl)

2 µl T7 RNA polymerase (NEB, 50 U/µl)

54.5 µl RNase free water

The reaction was incubated at 37°C for a minimum of 2 h (overnight incubation may increase yield). The reaction was stopped by adding 2 µl of 0.5 M EDTA pH 8.0 and 10 µl of 3 M Na acetate pH 5.2. This was extracted with an equal volume of phenol/chloroform (2.3.1.2) and the RNA was ethanol precipitated (2.3.1.3).

Purification of in vitro *transcribed RNA*

The RNA was pelleted by centrifugation at 14,000 rpm at 4°C for 30 min, washed with 500 µl 70% ethanol and finally air dried. The pellet was dissolved in 20 µl of 1x

formamide loading dye (48% deionised formamide, 10 mM EDTA pH 8.0, some BPB and XCB) and heated at 95°C for 2 minutes before running it on a 1x TBE 8% denaturing polyacrylamide (19:1) gel for 30 – 90 min at 25 W (2.3.3.3). The RNA was visualised by UV shadowing using a hand-held UV lamp (254 nm; UVItec Ltd.) and cut out with a clean RNase-free surgical blade. The RNA was eluted from the gel by electro-elution in a Schleicher and Schuell Biotrap (according to the manufacturer's instructions; using BT1 and BT2 Elutrap-Membranes, Whatman) in 0.25x TBE at 4°C for 30 min - 1 h at 200 V or by diffusion in two times 200 µl of RNA elution buffer (10 mM Tris-HCl pH 7.6, 0.1% SDS, 1 mM EDTA, pH 8.0) for two times 2 h at 30°C. The RNA was again ethanol precipitated (after electro-elution with 0.3 M Na acetate pH5.2; after passive elution with 150 mM NaCl), washed with 70% ethanol and air-dried before dissolving in 20 µl of TE. Yields were variable (0.5 – 20 µg). If desired the size and integrity of the purified RNA fragment was tested by running 1 µl on a 1x TBE 8% denaturing polyacrylamide (19:1) gel for 30 – 90 min at 25 W (2.3.3.3) along side 1 µg of pBM322 *Msp*I digest (NEB). The RNA was then visualised using Stains-All (Sigma). Ten millilitres of a 0.1% (w/v) Stains-All solution in formamide was mixed with 20 ml of isopropanol and 70 ml of deionised water. The gel was incubated in this mixture in the dark for 30 minutes and then destained in deionised water by exposure to light until bands were visible against a clear background.

2.3.3.3 Denaturing PAGE of RNA

For separation of small to medium sized single stranded RNA fragments samples were mixed with an equal volume of 2x formamide loading dye (96% (v/v) formamide, 20mM EDTA pH 8.0, some BPB and XCB) heated for 2 min at 95°C and loaded on a pre-run denaturing polyacrylamide gel in a vertical gel tank (Cambridge Electrophoresis, Model No. EV200). Denaturing polyacrylamide-urea gels (0.2 – 1.0 mm thick) were prepared using Sequagel 6 or 8 (National Diagnostics) for 6% or 8% (19:1) gels according to the manufacturer's instructions. The polymerised gel was pre-run in 1x TBE (2.1.3) for 30 min – 1 h at 25 W with a metal plate to distribute heat evenly. Wells were rinsed thoroughly before loading the samples. The gel was run for the required length of time (usually 45 – 120 min)

depending on the size of the fragments and the desired separation. Bromophenol Blue and Xylene Cyanol FF were used as a reference (BPB corresponds to 19 nt and 26 nt fragments on 8% and 6% gels respectively, XCB corresponds to 75 nt and 106 nt fragments).

2.3.3.4 Native PAGE of RNA

For separation of native RNA, RNA-RNA duplexes (annealing assay; 2.5.2) or RNA protein complexes (bandshift assay; 2.5.1) samples were mixed with native RNA loading buffer (2x native loading buffer: 30% glycerol, 80 mM HEPES-KOH pH7.9, 100 mM KCl, 2 mM Mg acetate, some BPB and XCB) before loading on a pre-run native polyacrylamide gel. Native polyacrylamide gels (6%, 20:1, 1.0 mm thick) were prepared by mixing 25.4 ml of water with 6 ml of 40% (w/v) acrylamide 6 ml of 2% (w/v) bis-acrylamide (Bio-Rad), 2 ml of 10x TBE, 0.4 ml of 10% APS and 40 µl of TEMED. After polymerisation the comb was removed and wells were rinsed thoroughly with 0.5x TBE. The gel was pre-run for 1 h at 5 W, 4°C in a vertical gel tank (Cambridge Electrophoresis, Model No. EV200) in pre-cooled 0.5x TBE running buffer. The gel was run for the required length of time (usually 3 - 4 h) depending on the size of the fragments and the desired separation. Bromophenol Blue and Xylene Cyanol FF were used as a reference (BPB corresponds to 45 nt and 60 nt fragments on 8% and 6% gels respectively, XCB corresponds to 160 nt and 240nt fragments).

2.3.3.5 Electrophoresis of glyoxylated RNA through agarose gels

For separation of medium to large sized single stranded RNA, samples were glyoxylated and run on agarose gels. Five µg of total RNA in 2 µl was mixed with 10 µl of glyoxal reaction mixture (6 ml of DMSO, 2 ml of deionised glyoxal, 1.2 ml of 10x BPTE electrophoresis buffer (2.1.3), 0.6 ml of 80% glycerol and 20 µl of 10 mg/ml ethidium bromide, divide into small aliquots and store at -70°C) and heated for 1 h at 55°C. The samples were then snap-chilled in ice water and collected at the bottom of the tube by centrifugation. The samples were loaded on a horizontal 1.2% agarose gel (24 cm in length) in 1x BPTE. The gel was run overnight at 55 V in a EMBL H135 gel tank. Lanes not containing RNA were loaded with RNase-free

standard DNA loading dye (2.3.2.12) containing Bromophenol Blue and Xylene Cyanol FF. These were used as a reference for migration. BPB was run to near the bottom of the gel.

2.3.3.6 Northern blotting

Method 1, after polyacrylamide gel electrophoresis

RNA separated by size on polyacrylamide gels, native (2.3.3.4) or denaturing (2.3.3.3), was transferred to a membrane to allow for detection of specific RNA species with radioactively labelled probes (2.3.2.17). The gel was stacked between two layers of Whatman 3MM paper with Hybond-N membrane (Amersham Biosciences) directly on top of the gel, all of which were cut to the size of the gel. This was sandwiched between sponge pads all of which were submerged in 0.5x TBE (2.1.3) taking care not to leave any bubbles between any of the layers. The transfer was performed in 0.5x TBE in a water-cooled Bio-Rad Trans-Blot Cell for 1 h at 60 V after which the RNA was crosslinked to the membrane in a UV Stratalinker 1800 (Stratagene) using the 'autocrosslink' function.

Method 2, after electrophoresis of glyoxylated RNAs through agarose gels

RNA separated by size on agarose gels was transferred to a membrane using passive transfer. The gel was washed in deionised water and incubated in 75 mM NaOH for 20 min to break the RNA molecules into smaller pieces to facilitate transfer. The reaction was stopped by incubation in Tris/NaCl (0.5 M Tris-HCl pH 7.0, 1.5 M NaCl) for 15 min, after which the gel was washed a further two times in 6x SSC (2.1.3). The gel was then stacked on top of Whatman 3MM paper on a bridge suspended over a container with 6x SSC with the filter paper hanging into the solution. A piece of Hybond-N^+ membrane (Amersham Biosciences) was cut to the exact size of the gel and after soaking it in 6x SSC was placed on top of the gel taking care not to leave any bubbles in between. Three more layers of soaked 3MM filter paper cut to size, followed by three layers of dry 3MM paper were added. Stacks of hand towels and a weight were placed on top of these to ensure constant pressure. Parafilm was used around the gel to ensure the towels only soaked up liquid that passed through the gel and the membrane. The transfer was left overnight after

which the RNA was crosslinked to the membrane in a UV Stratalinker 1800 (Stratagene) using the 'autocrosslink' function.

2.3.3.7 Northern blot hybridisation

Hybridisation of oligonucleotide probes to membrane immobilised RNA

End-labelled oligonucleotides (2.3.2.16) were passed through a 0.45 μm filter (Minisart Cellulose acetate, Sartorius) in 2 – 3 ml of SES1 buffer to remove particulate matter and hybridised to RNA on Hybond-N membrane (Amersham Biosciences) in 50 ml of SES1 hybridisation buffer (0.5 M Na phosphate pH 7.2, 7% (w/v) SDS, 1 mM EDTA) by incubation with agitation overnight at 37°C. Before addition of the probe, membranes were pre-hybridised in 50 ml of SES1 buffer at 37°C for a minimum of 30 min. After hybridisation, membranes were washed twice for 15 minutes in 6x SSPE buffer (2.1.3) at 37°C, followed by two washes of 30 minutes in 6x SSPE (pre-warmed to 42°C) at 37°C. Following the final wash, membranes were wrapped in Saran Wrap and subjected to autoradiography using Kodak BioMax MS Film and/or a Phosphor Imaging screen to allow for quantification using a STORM 860 scanner and ImageQuant 5.2 software (Molecular Dymamics).

Hybridisation of single stranded DNA probe to membrane immobilised RNA

Northern blot membranes (2.3.3.6) were pre-hybridised in Neverfail Hyb Solution at 42°C with agitation for a minimum of 2 h. Neverfail Hyb Solution was prepared as follows: 5 ml of 10% SDS, 250 ml of deionised formamide, 125 ml of 20x SSC (2.1.3), 12.5 ml of 1 M NaH_2PO_4, 12.5 ml of 1 M Na_2HPO_4, 25 ml of 100x Denhardts (10 g Ficoll 400, 10 g polyvinylpyrrolidone, 10 g BSA (Pentax Fraction V, Miles Laboratories) in 500 ml, filter sterilized, aliquots stored at -20°C), 1 ml of 10 mg/ml salmon sperm DNA and deionised water up to 500 ml. Neverfail Hyb Solution was stored at room temperature and passed through a 0.45 μm filter immediately prior to use. Purified single stranded DNA probe (2.3.2.17) was heated at 95°C for 2 min and added directly to the hybridisation solution. After overnight incubation at 42°C the hybridisation solution containing the labelled probe was stored in a 50 ml tube at -20°C (for re-use this mixture was heated for 15 min at

100°C). The membrane was washed 2 times with 50 ml of 6x SSC while heating the incubator to 55°C, then 2 times for 15 min with 50 ml of 0.2x SSC with 0.1% SDS at 55°C. Following the final wash, membranes were wrapped in Saran Wrap and subjected to autoradiography using Kodak BioMax MS Film and/or a Phosphor Imaging screen to allow for quantification using a STORM 860 scanner and ImageQuant 5.2 software (Molecular Dymamics).

Membrane stripping for reprobing

For successful removal of probes, membranes were prevented from drying out after hybridisation. To strip, a boiling solution of 0.1% (w/v) SDS was poured on the membrane and incubated on a rotating platform for 10 minutes. This was repeated 3 times, or until all counts were gone. The membrane was then either hybridised with another probe, or blotted dry and stored for future use.

2.3.3.7 Denaturing northern blot analysis of snRNAs

Yeast total RNA (2.3.3.1) was separated by denaturing PAGE (2.3.3.3; 25 W for 45 min or until XCB was 2cm from the bottom of the gel). The RNA was then transferred to Hybond-N membrane by electrophoretic transfer (2.3.3.6). All five spliceosomal snRNAs were detected simultaneously by hybridising the membrane with end-labelled RNA specific oligonucleotide probes for U1, U2, U4, U5 and U6 (Table 2.1.8.4)

2.3.3.8 Primer extension

The oligonucleotides to be used in primer extension were 5'-end radiolabelled with [γ-^{32}P] ATP (2.3.2.16) and purified by gel filtration (2.3.2.19). Primer extension reactions were set up as follows. 1 µl of 10 mM dNTPs, 0.5 µl of total RNA (0.7 µg; 2.3.3.1) and 4.5 µl of labelled primer were mixed, incubated for 5 min at 65°C and snap-chilled on ice. To this 4 µl of the following mixture was added: 2 µl 5x RT buffer (Invitrogen), 0.5 µl of 100 mM DTT, 0.125 µl of SUPERase-In (Ambion, 20 U/µl), 0.875 µl of water and 0.5 µl of SuperScript II Reverse Transcriptase (Invitrogen, 200U/µl). The reverse transcription reaction was then incubated for 1 h at 42°C and inactivated for 15 min at 70°C. Ten µl of 2x formamide loading dye

(96% (v/v) deionised formamide, 20mM EDTA, some BPB and XCB) with 10 mM NaOH was added to each reaction. Samples were heated for 5 min at 95°C to degrade the RNA and to denature the DNA products. They were then snap-chilled on ice and 10 µl of each was loaded onto a tall pre-run 6% (19:1; 0.2 mm thick) denaturing polyacrylamide gel (Sequagel-6, National Diagnostics) in a vertical gel tank (Cambridge Electrophoresis, Model EV500) in 1x TBE (2.1.3) using a metal plate to distribute the heat evenly. The samples were separated by electrophoresis for 2 h at 30 W until the BPB had just run off the bottom of the gel. The gel was dried onto Whatman 3MM filter paper on a Hybaid Gel-Vac drier by heating for 1.5 h at 80°C. The gel was exposed overnight to a Phosphor Imaging screen and an image was obtained by scanning the screen on a STORM 860 and processing the image using ImageQuant 5.2 (Molecular Dymamics). The sizes of fragments were estimated by comparing them to a radiolabelled marker (2.3.2.18).

2.4 Protein methods

2.4.1 Crude protein preparation from yeast

To prepare crude protein from yeast a method described by Volland *et al.* was used (Volland *et al.*, 1994). Briefly, 10 ml of the appropriate medium (2.1.2.3) was incubated overnight with a single fresh colony of the appropriate yeast strain (2.1.6). The optical density of the culture was measured at 600nm and approximately 3 units of yeast culture was sedimented by centrifugation. The cells were lysed by resuspending and incubation in 0.5 ml of 0.2 M NaOH on ice for 10 minutes. Proteins were then precipitated by adding 50 µl of 50% TCA and incubation on ice for a minimum of 10 minutes. Proteins were sedimented by centrifugation for 5 minutes at 14,000 rpm at 4°C. The pellet was finally resuspended in 35 µl of dissociation buffer (0.1 M Tris-HCl pH 6.8, 4 mM EDTA, 4% SDS, 20% (v/v) glycerol, 2% (v/v) β-mercaptoethanol, 0.02% (w/v) BPB) and the solution neutralised by the addition of 15 µl of 1 M Tris base. Samples were heated at 95°C for 10 min before SDS-PAGE (2.4.3; 2 – 10 µl was generally loaded per sample).

2.4.2 Preparation of yeast whole cell extract

2.4.2.1 Preparation of splicing extracts

100 - 200 ml of the appropriate medium was inoculated with a single colony of the yeast strain under investigation and incubated overnight at the appropriate temperature. This overnight culture was used to inoculate 2-3 l of pre-warmed medium to OD_{600} of approximately 0.1, and grown to an OD_{600} of approximately 0.8. Cells were then centrifuged in a Beckman JLA 10,500 rotor (5,000 rpm, 5 min, RT) and washed in 50 ml of ice-cold AGK buffer. Cells were sedimented (3,500 rpm, 3 min RT; Mistral 1000 centrifuge), resuspended in 20 ml AGK buffer and centrifuged as above. The resulting yeast pellet was weighed (typically 3 - 5 g) and cells were resuspended in a volume of AGK buffer (ml), corresponding to 0.4x the cell pellet weight (g). DTT was added to a final concentration of 2 mM and the yeast suspension was frozen by dropwise addition to liquid nitrogen and stored at -70°C.

The frozen cell suspension was added to a mortar containing liquid nitrogen and ground to a fine powder keeping everything from thawing by occasional addition of liquid nitrogen. The frozen powder was then transferred to a pre-chilled polycarbonate centrifuge tube, thawed on ice and centrifuged in a JA25.50 rotor (17,000 rpm, 30 min, 4°C). The supernate was transferred to a pre-chilled polycarbonate centrifuge tube, taking care to avoid the lipid-containing top layer of the supernate, and centrifuged in a 70.1 Ti rotor (40,000 rpm, 60 min, 4°C; Beckman). The resulting supernate was transferred to SnakeSkin dialysis membrane (Pierce, MWCO 3.5 kDa) and dialysed against 1 l of pre-chilled dialysis buffer for 3 hours at 4°C, with two changes of dialysis buffer. The dialysed cell extract was cleared of insoluble material one last time by centrifugation in pre-chilled microfuge tubes (14,000 rpm, 10 min, 4°C). The resulting supernate was finally aliquoted into pre-chilled microfuge tubes and frozen at -70°C until use.

AGK buffer:	10 mM HEPES-KOH, pH 7.9
	1.5 mM $MgCl_2$
	200 mM KCl
	10% (v/v) Glycerol

Dialysis buffer:	20 mM HEPES-KOH, pH 7.9
	50 mM KCl
	0.4 mM DTT

2.4.2.2 Preparation of crude extracts for immunoprecipiation

Cultures of 1 litre were grown to OD_{600} of 0.8 – 1.0. Cells were centrifuged, washed with 20ml of ice-cold AGK buffer and pelleted again. The wet pellet was weighed and 0.4x volume of AGK buffer was used to resuspend the cells. DTT was added to a final concentration of 2 mM before freezing the cell suspension in liquid nitrogen in a dropwise fashion. The cells were ground using a mortar and pestle, as described above. Frozen powder was transferred to a pre-chilled polycarbonate centrifuge tube, thawed on ice and centrifuged in a JA25.50 rotor (17,000 rpm, 30 min, 4°C). The supernate was aliquoted and stored at -70°C or used directly for immunoprecipitation (2.4.6).

2.4.3 SDS Polyacrylamide gel electrophoresis (SDS-PAGE)

Samples to be analysed by SDS-PAGE had an equal volume of 2x SDS Loading Buffer (see below) added before being heated at 95°C for 5 - 10 minutes. Samples were then loaded onto a pre-cast 4-12% polyacrylamide Bis-Tris gel (NuPAGE, Invitrogen), 4-20% Tris-Glycine gel (Invitrogen or iGels from Life Gels) or 10-20% Tris-Glycine gel (Invitrogen) alongside molecular weight markers of known size (SeeBlue Plus2, Invitrogen) for reference. Bis-Tris gels were subjected to electrophoresis in an XCell SureLock Mini-Cell gel tank (Invitrogen) with 1x MOPS SDS-PAGE running buffer (2.1.3); Tris-Glycine gels with 1x Tris-Glycine SDS-PAGE running buffer (2.1.3). Electrophoresis was performed at 150 V until the bromophenol blue dye front reached the bottom of the gel or until the pre-stained marker had reached the desired level of separation (usually 90 min). Home-made gels (Tris-Glycine) were also used at the appropriate concentration (most commonly 15% resolving gel with 5% stacking gel), poured and run using the Bio-Rad Mini-PROTEAN 3 System. Ten ml of 15% resolving gel mix for two 1.0 mm gels was made as follows: mix 2.3 ml of water, 5 ml of 30% acrylamide/bis-acrylamide

(Design-A-Gel, 37.5:1, Severn Biotech Ltd.), 2.5 ml of 1.5 M Tris-HCl pH 8.8, 0.1 ml of 10% SDS, 0.1 ml of 10% APS and 10 µl of TEMED. Gels were allowed to set covered with isopropanol and subsequently washed with water. Then 5% stacking gel mix (for 3 ml mix: 1.66 ml of water, 0.5 ml of 30% acrylamide/bis-acrylamide, 0.78 ml of 0.5 M Tris-HCl pH 6.8, 30 µl of 10% SDS, 30 µl of 10% APS and 3 µl of TEMED) was poured and allowed to set on top. For different percentage gels the amount of bis-acrylamide solution and water were adjusted accordingly. Alternatively, for better resolution of proteins of low molecular weight the high-TEMED protocol described by (Lehmeier *et al.*, 1990) was used. For SDS-PAGE of mixtures containing multiple Lsm proteins 4-20% Tris-Glycine gels were most commonly used (occasionally 10-20%).

2x SDS Loading Buffer:	100 mM Tris-HCl, pH 6.8
	1.25 M β-Mercaptoethanol
	4% (w/v) SDS
	0.2% (w/v) bromophenol blue
	20% glycerol

2.4.4 Coomassie staining of protein gels

2.4.4.1 Standard Coomassie staining protocol

After electrophoresis the gel was stained for 30 min – 2 h in Coomassie staining solution (50% methanol, 10% acetic acid and 0.1% (w/v) Coomassie Brilliant Blue R, Sigma B7920) until the gel was completely blue. It was then transferred to destaining solution (10% methanol, 10% acetic acid) and incubated with gentle mixing for several hours or overnight if necessary with a piece of film to soak up the Coomassie until the background was clear compared to the stained protein bands.

2.4.4.2 GelCode Blue staining for mass spectrometry

To allow detection of protein bands after SDS-PAGE for isolation and subsequent Mass spectrometry (2.4.12) GelCode Blue Stain Reagent (Pierce) was used according to the manufacturer's instructions. The gel was handled with clean gloves only and stained in a clean container. After electrophoresis the gel was washed three

times five minutes in 200 ml of deionised water. It was then incubated with gentle mixing in 20 ml of the staining solution for a minimum of 1 h, followed by destaining for 2 – 3 h with several changes of deionised water.

2.4.5 Western blotting

2.4.5.1 Transfer of proteins to PVDF membrane

Proteins were transferred electrophoretically from the SDS-polyacrylamide gel to PVDF membrane (Immobilon-P, Millipore) using the Invitrogen XCell II Blotmodule system according to the manufacturer's instructions. Electrophoretic transfer was performed in 1x Western Transfer Buffer (2.1.3) at 25 - 30 V for 1 - 2 hours. Transfer was confirmed by staining of the membrane with Ponceau S or simply by observing the transfer of the pre-stained marker.

2.4.5.2 Western blot hybridisation

Non-specific protein interactions to the membrane were prevented by blocking the PVDF membrane directly after transfer. To block, the membrane was incubated for a minimum of 30 min at room temperature in 20 ml TBST-M (1x TBST (2.1.3) with 5% (w/v) Marvel dried skimmed milk) with constant shaking. Primary antiserum (2.1.9) was diluted to the appropriate concentration in 2.5 - 10 ml TBST-M, added to the blocked membrane and incubated (2 h, RT or o/n, 4°C) with constant shaking. The membrane was then washed 5 times 5 minutes in 20 ml of 1x TBST. For HRP conjugated primary antibodies the Western blot was ready for development (2.4.5.3). For non-conjugated primaries, HRP-conjugated secondary antiserum was diluted in 10 ml of 1x TBST-M, applied to the membrane and incubated with constant shaking (1 – 2 h, RT). The membrane was then washed as described above and developed as described below.

2.4.5.3 Enhanced Chemiluminescence (ECL)

1.5 ml of ECL solution (Amersham Biosciences) was prepared according to the manufacturer's instructions, added to the protein side of the PVDF membrane and incubated (1 min, RT). Excess liquid was removed, the membrane was placed in Saran wrap and exposed to photographic film (Kodak MXB Film). Exposure times

varied depending on the strength of the signal, but were typically between 10 seconds and 5 min. Film was developed using a Konika SRX-101A Developer. For weak signals ECL Plus (Amersham Biosciences) was used according to the manufacturer's instructions.

2.4.5.4 Stripping Western blots

To remove antibodies from a Western to allow subsequent hybridisation with a different antibody Restore Western Blot Stripping Buffer (Pierce) was used. To this end the blot was incubated with agitation for 30 min in 20 – 30 ml of stripping buffer at 50°C. The blot was then washed in TBS-T, blocked and hybridised as usual (2.4.5.2). The Stripping Buffer was kept at 4°C and reused multiple times.

2.4.6 Immunoprecipitation of proteins

Pre-swollen Protein A Sepharose beads (30 µl; Amersham Biosciences) were incubated with antiserum (2.1.9) overnight at 4°C. Beads were washed 3 times with 1 ml of NTN buffer (2.1.3) and blocked with 100 µl blocking solution (0.1 mg/ml tRNA, 0.1 mg/ml BSA and 0.1 mg/ml glycogen) for 45 minutes at room temperature. Pre-conjugated agarose beads (30 µl; Santa Cruz Biotechnology; 2.1.9) were washed with 1 ml of NTN and blocked the same way. Beads were then washed three times with 1 ml of NTN buffer and once with 1 ml of 1x IP_{150} buffer (2.1.3). For immunoprecipitation of proteins from whole cell extracts 30 µl of extract (2.4.2) was added in a final volume of 200 µl of 1x IP_{150} buffer. For co-immunoprecipitation of recombinant proteins appropriate amounts (0.2 – 2 µg) of each of the purified proteins were added in IP_{150} buffer. The suspension was incubated for two hours at 4°C with gentle mixing. Beads were then washed 2 times with 1 ml of NTN and 2 times with 1 ml of NT buffer. For problematic "sticky" proteins 0.1% Triton X-100 was added to the IP_{150} and NTN buffers. Beads were finally resuspended in 20 µl 1x SDS loading buffer (2.4.3). Supernate and pellet samples (and where appropriate input samples) were heated in 1x SDS loading buffer for 5 – 10 min at 95°C and subjected to SDS-PAGE (2.4.3) and Western blotting (2.4.5).

2.4.7 Recombinant protein expression in *E. coli*

Two methods of recombinant protein expression in *E. coli* were used and many variations of both methods to obtain optimal levels of expression and/or optimal levels of soluble protein. For optimal levels of expression colonies of the appropriate *E. coli* strain (2.1.5) freshly transformed with the expression vectors (Tables 2.1.7.1 and 2.1.7.2) were used to inoculate an appropriate volume of LB in the presence of selective antibiotics (2.1.2.5).

2.4.7.1 Recombinant protein induction with IPTG

For one method of expression a small culture (usually 5 ml) was grown overnight at 37°C. The following morning a 1 in 100 dilution was made in fresh LB (100 ml - 1 L) with antibiotics. This culture was grown at 18°C, 23°C, 25°C, 30°C or 37°C to an OD_{600} of 0.4, at which stage 0.1mM of isopropyl-β-D-thiogalactopyranoside (IPTG) was added. The culture was then grown for another 1h to overnight before harvesting (usually short for expression at 37°C; long for 18°C or 23°C).

2.4.7.2 Recombinant protein expression using ONEX

The second method of expression used the Novagen Overnight Expression (ONEX) system. For this method 2 – 10 fresh colonies were used to inoculate an appropriate volume of LB (1 - 100 ml) with antibiotics and ONEX solutions 1, 2 and 3 (0.01, 0.05 and 0.001 times the culture volume respectively). This culture was grown at the appropriate temperature for the appropriate length of time (14 – 16h for 37°C, 24 - 26 h for 30°C or 30 - 34 h for 23°C) before harvesting.

2.4.7.3 Harvesting cultures after recombinant protein expression

A small volume of each expression culture was taken and mixed with an equal volume of 2x SDS-PAGE loading buffer to allow to check for expression by SDS-PAGE (2.4.3) followed by Coomassie staining (2.4.4) and/or Western analysis (2.4.5). Cultures were harvested by pelleting the cells for 5 min at 6,000 rpm, 4°C in a Beckman JLA10,500 rotor. Cell pellets were stored at -70°C until required for lysis and recombinant protein purification (2.4.8). Even if the purification was carried out the same day, the pellet was frozen at -70°C because this step aids in cell lysis.

2.4.7.4 Expression of Lsm proteins and Hfq

The most successful methods for expression of Lsm proteins and *E. coli* Hfq were overnight expression at 23°C in the presence of 0.1 mM IPTG (500 ml culture volume) and 24 h expression at 30°C using the ONEX system (10 – 50 ml culture volume). When performed as described high levels of expression and satisfactory levels of soluble protein were obtained, although the latter varied considerably between the different Lsm proteins.

2.4.7.5 Expression of Prp24p

The most successful method for expression of Prp24 protein was expression at 37°C in the presence of 0.1 mM IPTG (500 ml culture volume) for 1 – 2 h. When performed as described high levels of expression were obtained. Levels of soluble Prp24p were generally high; the main concern was to limit proteolysis. Degradation of recombinant Prp24p was minimised by adding IPTG no later than OD_{600} 0.4, by not exceeding 2 h of incubation in the presence of IPTG and by cooling the culture quickly after expression.

2.4.8 His-tagged protein purification

Two methods of His-tagged protein purification were used, native and denaturing, both based on Qiagen Ni-NTA Agarose or Superflow and the manufacturer's instructions. Methods were optimised for individual proteins as described below.

2.4.8.1 Cell lysis for protein purification

Two methods were used for cell lysis. The first method employed lysozyme and Triton X-100 detergent to break the cells and exonucleases to degrade DNA and RNA; in the second (preferred) method cells were lysed by sonication while shearing DNA and RNA simultaneously.

Cell lysis for native purification

Both methods were used for cell lysis in preparation for native purification of recombinant proteins. In the first method cells were resuspended in Lysis Buffer (20

mM Tris-HCl pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 10 mM imidazole and 5 mM β-Mercaptoethanol; 30 ml for 500 ml IPTG expression culture; 5 ml for 10 ml ONEX expression culture). Lysozyme was added to a final concentration of 0.2 mg/ml and the suspension was incubated for 30 minutes with mixing at 4°C. Triton X-100 was then added to a final concentration of 0.2% (v/v) together with DNase and RNase A (final 5 μg/ml for both), and the mixture was incubated a further 15 min at 4°C. To lyse cells by sonication a Vibra Cell microtip sonicator (Sonics & Materials Inc., Jencons Scientific Ltd.) was used. It was set to amplitude 35% with 5 sec bursts and 5 sec breaks for 2 min. To make sure the mixture stayed cold during sonication it was kept on watery ice throughout.

Cell lysis for denaturing purification

Harvested cell pellets (2.4.7.3) were resuspended in freshly prepared 8 M urea pH 8 buffer (100 mM NaH_2PO_4, 10 mM Tris, 8 M urea, pH 8.0 set with NaOH) and cells were lysed by sonication as described above.

After lysis cell extracts were transferred to polypropylene centrifuge tubes and centrifuged (17,000 rpm, 30 min, 4°C; Beckman JLA 10,500 rotor). Supernates were transferred to polycarbonate tubes and used for purification immediately (2.4.8.2 and 2.4.8.3) or stored at -70°C until required.

2.4.8.2 His-tagged protein purification under native conditions

Cleared cell lysate (2.4.8.1) was incubated with Ni-NTA agarose (Qiagen; washed with lysis buffer) for 30 min -1 h at 4°C with rotation. The amount of Ni-NTA agarose used depended on the level of soluble recombinant protein present after lysis and was chosen for each case. It is preferable to have excess His-tagged protein over resin, to avoid high levels of non-specific binding of *E. coli* proteins. As a rule of thumb 0.5 to 2 ml of 50% slurry was used for 100 to 500 ml IPTG induced cultures (2.4.7.1) or 10 to 100 ml ONEX induced cultures (2.4.7.2). The agarose beads were sedimented by centrifugation for 1 min at 1,000 rpm in a Mistral 1000 centrifuge. The beads were washed with 10 ml of lysis buffer, centrifuged, resuspened in 5 ml of lysis buffer, applied to a Polyprep Chromatography column (Bio-Rad, 731-1550) and

allowed to drain by gravity. Another 10 ml of lysis buffer was applied to the resin and left to drain. The resin was washed another three times with 10 ml of wash buffer (20 mM Tris-HCl pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 40 mM imidazole and 5 mM β-Mercaptoethanol), fully resuspending the resin during one of these steps. Finally elution buffer (wash buffer with 250 mM imidazole) was applied to the column 1 ml at a time and 0.2 to 0.5 ml fractions were collected and stored at -70°C. Samples were taken at each step, including all elution fractions, mixed with equal volumes of 2x SDS-PAGE loading buffer. All samples were separated by SDS-PAGE (2.4.3) followed by Coomassie staining (2.4.4) and/or Western analysis (2.4.5) to determine the presence and purity of purified recombinant protein. Elution fractions with good amounts and purity of recombinant protein were pooled and dialysed against 3x 330 ml of BC100 buffer (20 mM Tris-HCl pH7.9, 100 mM KCl, 20% (v/v) glycerol) at 4°C in Spectra/Por (Spectrum, MWCO 12 - 14 kDa) or SnakeSkin (Pierce, MWCO 3.5 kDa) dialysis tubing to remove all imidazole. For proteins that tended to precipitate at low salt (Lsm4 in particular) KCl or NaCl in the buffer up to 0.5 M was beneficial. After dialysis insoluble material was removed by centrifugation and supernate was aliquoted and stored at -70°C.

2.4.8.3 His-tagged protein purification under denaturing conditions

Cleared cell lysate (2.4.8.1) was incubated with Ni-TNA agarose (Qiagen; washed with 8M urea lysis buffer) for 30 min at 4°C with rotation. The amount of Ni-NTA agarose used depended on the level of recombinant protein present after lysis and was chosen for each case (see 2.4.8.2). Resin was washed as described under 2.4.8.2, but 8 M urea pH 8 lysis buffer (100 mM NaH_2PO_4, 10 mM Tris, 8 M urea, pH 8.0 set with NaOH) and 8 M urea pH 6.3 wash buffer were used instead. Protein was finally eluted with 8 M urea pH 4.5 buffer (pH adjusted with HCl). All buffers were prepared freshly each time. Elution fractions with good amounts and purity of recombinant protein were pooled and stored at -70°C.

2.4.9 Protein quantification

To determine the concentration of proteins in purified samples or in total protein extracts the Bradford assay was used in the form of the Bio-Rad Protein Assay,

according to the manufacturer's instructions. In brief, a standard curve was made using a BSA concentration range of 0.2, 0.4, 0.6, 0.8 and 1.0 mg/ml. 20 µl of each solution was mixed with 1 ml of dye reagent (made by diluting 1 part of 5x Dye Reagent Concentrate with 4 parts of deionised water). These were mixed thoroughly and their A_{595} was measured between 5 min and 1 h of mixing. Absorbance readings were plotted against protein concentration and a linear fit was used to estimate unknown protein concentrations. To this end A_{595} was measured for a mix of 20 µl of protein sample with 1 ml of 1x Bradford reagent. Where necessary dilutions of protein samples were made to ensure readings were between those of the 0.2 and 1.0 mg/ml BSA standards.

2.4.10 *In vitro* assembly of Lsm complexes

2.4.10.1 Assembly of Lsm complexes in 8M urea

Individual Lsm proteins were expressed in *E. coli* (2.4.7.4) and purified under denaturing conditions using Ni-NTA resin (2.4.8.3). Equal amounts of each of the members of the Lsm1-7p or Lsm2-8p complexes as estimated by Coomassie stained SDS-PAGE gels (2.4.3) were mixed together and dialysed in Spectra/Por dialysis tubing (Spectrum, MWCO 12 - 14 kDa) against BC100 (2.4.8.2) at 4°C (3 times 330 ml, 2 h). Soluble and insoluble proteins were separated by centrifugation for 15 min at 14,000 rpm, and samples of each were separated by SDS-PAGE (2.4.4). Coomassie staining of the gel showed the yield of soluble protein and presence of the different members of the Lsm complexes. Mass spectrometry of overlapping bands (Lsm2 and 5, and Lsm7 and 8; 2.4.12) confirmed the presence of all members of the ring (3.5.3). Co-immunoprecipitation (2.4.6) of all Lsm proteins using the epitope tag on one of them showed their interaction (3.5.1). Size exclusion chromatography (2.4.11) was used to estimate the size of the resulting complexes.

2.4.10.2 Assembly of Lsm complexes in 4M urea

Human LSm complexes were assembled *in vitro* using purified recombinant subcomplexes by a method based on the one described by Zaric *et al.* (Zaric *et al.*, 2005). In brief, equal molar amounts of each of the subunits were mixed and the urea concentration was brought up to 4 M by the addition of the appropriate amount of

urea. The mix of proteins was dialysed for 4 h against 20 mM Tris-HCl pH7.5, 1 M NaCl and 4 M urea at 30°C, for 16 h against 20 mM Tris-HCl pH7.5, 1 M NaCl at 30°C and finally for 2 – 4 h against 20 mM Tris-HCl pH7.5, 0.5 M NaCl at 4°C.

2.4.10.3 Optimisation of yeast Lsm complex assembly

Comparison of the 4M and 8M urea methods used for assembly of the yeast and human Lsm complexes, their yields and RNA binding, annealing and strand displacement activities (described in Chapters 3-5), shows that there is still room for improvement, in particular for *in vitro* assembly of recombinant yeast Lsm complexes. Possibilities for improvements are discussed in detail in 3.5.4.

2.4.11 Size exclusion chromatography of protein complexes

Protein samples to be analysed by size exclusion chromatography were dialysed into gelfiltration buffer (20mM Tris-HCl pH 7.9, 0.1 – 0.5 M KCl, 10 mM β-Mercaptoethanol) and concentrated by centrifugation in Vivascience Vivaspin concentrators (MWCO 30,000; VS0121). Samples were centrifuged directly before injection to loose any insoluble matter. For separation of proteins in the range between 10 and 600 kDa a SUPERDEX 200 HR10/30 size exclusion chromatography column was used connected to an ÄKTApurifier10 FPLC unit (Amersham Biosciences). For better resolution in a higher molecular weight range (30 to 2,000 kDa) a SUPEROSE 6 HR10/30 (Amersham Biosciences) size exclusion column was used. After a pumpwash purifier run, the column was equilibrated with at least one column volume (V_c) of gelfiltration buffer (filtered and degassed at 4°C). Protein was loaded in a volume of 200 µl using a 200 µl loop and injected onto the column using 300 µl of buffer. During a run with 1.2 column volumes of buffer at 0.3 - 0.5 ml/min the A_{260} of flowthrough was measured and collected in 0.5 ml fractions. Data were collected using Unicorn 4.10 software. To calculate protein molecular weights (Mw) from their corresponding elution volumes (V_e) proteins of known size were run across the same columns. Proteins with molecular weights of 15.6, 20.4, 49.1, 67.4, 176, 219, 416 and 699 kDa had previously been used to create a calibration curve for the SUPERDEX 200 column. Albumin (67 kDa), aldolase (158 kDa), catalase (232 kDa), ferritin (444 kDa) and thyroglobulin (669 kDa) from

the Amersham High Molecular Wight Gel Filtration Calibration Kit (17-0441-01) were used to calibrate the SUPEROSE 6 column. Blue dextran 2,000 was used to estimate the column void volumes (V_0). The Kav = $(V_e-V_0)/(V_c-V_0)$ was plotted against the Log(Mw), the linear fit to these data points was used to calculate the molecular weight of detected A_{260} peaks.

2.4.12 Mass spectrometry of proteins

Proteins were visualised after SDS-PAGE by Coomassie staining (2.4.4). Appropriate bands were excised using a clean scalpel. Samples were processed and analysed by Sam Clokie and Andrew Cronshaw in the Mass Spectrometry facility of The Wellcome Trust Centre for Cell Biology. In brief, gel bands were cut and digested with trypsin (Promega). Peptides from the digested bands were then run on a MALDI-TOF MS spectrometer (Matrix-assisted Laser Desorption/Ionisation, Time Of Flight; Voyager DE-STR MALDI-TOF MS, Applied Biosystems) and some on a nanoLC-MS mass spectrometer (Liquid Chromatography, Mass Spectromertry; Famos autosampler, Switchos column switching unit and Ultimate nanoLC, Dionex). Processed spectra were searched against the NCBI non-redundant database using Protein Prospector or MASCOT (Matrix Science).

2.5 Mixed protein/RNA assays

2.5.1 RNA bandshift assay

RNA bandshift reactions were carried out in a volume of 7.5 µl. *In vitro* transcribed RNA (2.3.3.2) and purified proteins (2.4.8) were mixed and incubated at room temperature before separating bound and unbound RNA by native PAGE (2.3.3.4). RNA was denatured at 85°C for 3 min and snap-chilled on ice just before the assay. Each reaction was assembled as follows in 1x RNA annealing buffer (40 mM HEPES-KOH pH 7.9, 50 mM KCl, 10 mM Mg acetate):

0.75 µl	10x RNA annealing buffer
0.75 µl	10 mM DTT
0.1 µl	RNasin (Promega, 40 U/µl)

1 µl	U6 RNA (7.5 ng; 0.2 pmol) or other RNA
0.5 – 5 µl	purified protein (complex)

Bring total volume up to 7.5 µl with deionised water

Reactions were incubated for 10 min at room temperature. They were then transferred to ice before separation by native PAGE (2.3.3.4), followed by Northern blotting (2.3.2.6) and hybridisation with radioactively labelled RNA specific probes (2.3.2.7; Table 2.1.8.4).

2.5.2 U4/U6 RNA annealing assay

U4/U6 annealing reactions were carried out in a volume of 7.5 µl. *In vitro* transcribed U4 and U6 RNA (2.3.3.2) and purified proteins (2.4.8) were mixed and incubated before degrading the protein with Proteinase K (and in some cases removing all protein by phenol/chloroform extraction; 2.3.1.2). RNA was denatured at 85°C for 3 min and snap-chilled on ice just before the assay. Each reaction was assembled as follows in 1x RNA annealing buffer (40 mM HEPES-KOH pH 7.9, 50 mM KCl, 10 mM Mg acetate):

0.75 µl	10x RNA annealing buffer
0.75 µl	10 mM DTT
0.1 µl	RNasin (Promega, 40 U/µl)
1 µl	U6 RNA (7.5 ng; 0.2 pmol)
0.5 – 4 µl	purified protein (complex)

Bring total volume up to 6.5 µl with deionised water

Reactions were incubated for 10 min at room temperature before adding 1 µl U4 RNA (15 ng, 0.27 pmol), followed by incubation at 30°C for 30 min – 1 h. 2.5 µl of native loading buffer (15% glycerol, 40 mM HEPES-KOH pH7.9, 50 mM KCl, 1 mM Mg acetate, some BPB and XCB) with 1 µg/µl Proteinase K was added to each reaction and incubated for a further 5 – 10 min at 30°C. In some cases phenol/chloroform extraction was performed with an equal volume of

phenol/chloroform (2.3.1.2) and the aqueous phase was loaded directly onto a pre-run 6% (20:1) native PAGE gel (2.3.3.4) and run for 3 h at 5 W, 4°C in 0.5x TBE. RNA was transferred to Hybond-N by Northern blotting (2.3.3.6) and hybridised with radioactively labelled probes (2.3.2.16) for U4 or U6 RNA (2.3.2.7; Table 2.1.8.4).

2.5.3 Strand displacement assay

The strand displacement assay was adapted from the helicase assay described by Cordin *et al.* (Cordin *et al.*, 2004).

2.5.3.1 Strand displacement substrate preparation

Substrate for the strand displacement assay consisted of a duplex formed between non-labelled RNA and a labelled complimentary DNA oligonucleotide. The RNA fragment was produced by *in vitro* transcription and gel purified (2.3.3.2). The DNA oligonucleotide was end-labelled (2.3.2.16) and purified by gel filtration (2.3.2.19). Duplex was formed by mixing the following: 5 µl of 5x annealing buffer (100 mM Tris-HCl pH 8.0, 1 M potassium acetate, 0.5 mM EDTA), 1.25 µl of 100 µM unlabeled oligo (final 5 µM), RNA at a final concentration of 5 µM, 2.5 µl of labelled oligo (final 0.33 µM), and water up to 25 µl. This mixture was incubated for 2 min at 90 °C and snap-chilled on ice, and used directly in the strand displacement assay or stored at -20 °C until required.

2.5.3.2 Strand displacement reactions

Strand displacement reactions were carried out in a volume of 10 µl: 0.5 µl of duplex (25 - 250 nM final), 0.5 µl of 100 µM oligo, 2 µl of 5x helicase assay buffer (100 mM Tris-HCl pH 7.5, 250 mM potassium acetate, 25 mM Mg acetate, 0.5 mg/ml BSA), 1 µl of 10 mM DTT, 0.2 µl of RNAsin (Promega, 40 U/µl), 0.5 – 4 µl of protein and deionised water up to 10 µl. Duplex was added last and reactions were mixed thoroughly (for the Ded1p control 1 mM ATP was added last). These were then incubated for various times (usually between 1 min and 1 h) at 37 °C. Reactions were stopped by transferring them to ice and adding 5 µl of helicase loading buffer (40% glycerol, 10 mM EDTA, some BPB and XCB). Samples were loaded on a 15%

native 1x TBE PAGE (20:1) (Bio-Rad, MiniPROTEAN gel-kit; 1.0 mm gels) and run in pre-cooled 1x TBE at 4°C for 30 min at 8 W per gel. Gels were dried onto Whatman 3MM paper on a Hybaid Gel-Vac drier by heating for 1.5 h at 80°C. Dried gels were exposed to film (Kodak BioMax MS Film) and/or to a Phosphor Imaging screen to allow for quantification using a STORM 860 scanner and ImageQuant 5.2 software (Molecular Dymamics).

2.6 Microscopy methods

2.6.1 Live cell imaging

For live cell imaging, cells were grown in the appropriate liquid medium. For simple visualisation, 5 µl of (log-phase) culture was spotted on a poly-lysine coated microscope slide (Polysine, VWR, 631-0107), a coverslip was placed on top, and cells were investigated immediately.

Live cells were observed by bright field microscopy and expression and localisation of fluorescently tagged proteins were detected by fluorescence microscopy using a Leica FW4000 microscope. The microscope contained a narrowband filter wheel with the capability for detecting fluorescence from GFP (gfp; 489 nm excitation, 508 nm emission), RFP (Cy3; 550 nm excitation, 570 nm emission) and DAPI (A4; 359 nm excitation, 461 nm emission). Exposure times for GFP and RFP fluorescence were typically between 200 and 500 milliseconds, while exposure times for DAPI were between 50 and 100 milliseconds. Images were captured with a CoolSnap camera (Roper Scientific) using Leica FW4000 V1.1.1 software. Only one Z-section was taken for each sample, because live cells are constantly in motion.

2.6.2 Imaging of fixed and DAPI stained cells

Cells were grown in the appropriate liquid media, fixed for 10 minutes in 4% formaldehyde and washed 3 times with buffer B (2.1.3). To detect the localisation of nuclei the nuclear marker DAPI, which stains DNA, was used. Fixed and washed cells were resuspended in a suitable volume of buffer B and 2 µl of cell suspension was spread on a poly-lysine coated microscope slide (Polysine). This was left to

airdry for a few minutes, 3 µl of Vectashield (H1200), mounting medium for fluorescence with DAPI, (Vector Laboratories Inc., Burlingame, CA, USA) was added and a coverslip was placed on top. Cells were observed 1 – 24 h later to allow the DAPI to stain the DNA properly.

2.7 Bioinformatics

A range of bioinformatics tools was used, most of which are internet-based. The Saccharomyces Genome Database (**SGD**; http://www.yeastgenome.org/) was used to get general information on *S. cerevisiae* genes of interest. This includes DNA and amino acid sequence information, physical and genetic interaction data and description of mutant phenoypes. **NCBI BLAST** (http://www.ncbi.nlm.nih.gov/blast/) was used to find homologues of proteins (blastp) or to find out the origin of a piece of sequenced DNA (blastn or blastx; Altschul *et al.*, 1990). **Protein Calculator v3.3** (http://www.scripps.edu/~cdputnam/protcalc.html) was used to get information on a protein, such as its molecular weight, pI and molar absorbance. **NEBcutter V2.0** (http://tools.neb.com/NEBcutter2/index.php) was used to get a restriction map of a piece of DNA (Vincze *et al.*, 2003). **Reverse complement** (http://bioinformatics.org/sms/rev_comp.html) was used to get the reverse complement of a DNA sequence (usually to make a reverse primer). **ExPASy translate tool** (http://www.expasy.org/tools/dna.html) was used to get all possible frames of translation of a DNA or RNA sequence. To calculate the annealing temperature of primers for site-directed mutagenesis the **Stratagene QuikChange Primer T_m calculator** was used (http://www.stratagene.com/QPCR/ tmCalc.aspx). To order primers and get information on their annealing temperature and likelihood of secondary structure formation the **Sigma-Genosys** website was used (https://orders.sigma-genosys.eu.com/). To align multiple amino acid sequences by conservation **ClustalW** was used (http://www.ebi.ac.uk/clustalw/) (Thompson *et al.*, 1994) and **Boxshade 3.21** for a clearer graphical output of the results (http://www.ch.embnet.org/software/BOX_form.html). **Jalview version 2.2** (http://www.jalview.org/) was used to make manual adjustements to automatic multiple sequence alignments (Clamp *et al.*, 2004). To check for the presence of nuclear

localisation signals in proteins of interest **PredictNLS** (http://cubic.bioc. columbia.edu/ predictNLS/) was used (Cokol *et al.*, 2000). To check for the presence of putative phophorylation sites in proteins of interest **NetPhos 2.0** (http://www.cbs.dtu. dk/services/NetPhos/) was used (Blom *et al.*, 1999). To predict the 2D and 3D structure of proteins of interest based on homology to proteins with known (crystal) structures **3D-PSSM** (http://www.sbg.bio.ic.ac.uk/~3dpssm/ html/ffrecog_simple. html) was used (Kelley *et al.*, 2000b). For prediction of secondary RNA structures **mfold** version 3.2 (http://www.bioinfo.rpi.edu/ applications/mfold/rna/form1.cgi) was used (Mathews *et al.*, 1999; Zuker, 2003) Finally, to check the sequence and trace of a sequenced piece of DNA and to align it with a known sequence **Sequencher 4.2 or 4.5** (Gene Codes Corp.; Macintosh) or **ContigExpress** (Invitrogen, Vector NTI Advance 10; PC) were used.

CHAPTER 3

Lsm protein expression, purification and in vitro complex formation

3.1 Introduction

To investigate the structure and function of the *S. cerevisiae* Lsm1-7 and Lsm2-8 complexes recombinant Lsm proteins needed to be produced. Initial experiments with Lsm proteins produced by *in vitro* translation in wheat germ extracts (Page, 2002; Silvia Galardi, this lab, unpublished data) were promising, but the potential presence of wheat Lsm proteins and/or other functionally related proteins made interpretation of results problematic. A similar approach using *E. coli* cell-free extracts to produce Lsm proteins by *in vitro* translation proved unsuccessful. Large amounts of most Lsm proteins could be produced, but they were often largely insoluble. No specific U6 RNA binding activity of mixtures of Lsm2-8 proteins produced this way could be detected (data not shown). An alternative approach was therefore taken, in which recombinant Lsm proteins were produced in *E. coli*, purified and used to assemble the appropriate complexes *in vitro*. Complexes thus produced could be used to study their function in isolation of contaminant eukaryotic proteins. In addition, the aim was to elucidate the correct order of the Lsm polypeptides within the ring-shaped complex using different tags on different proteins to allow examination of interactions between individual subunits by co-immunoprecipitation. Efforts to optimise expression, purification and *in vitro* assembly of Lsm complexes are described in this chapter. Functional and structural analyses of these complexes are described in Chapters 4-6.

3.2 Creating protein expression vectors

3.2.1 Cloning yeast *LSM* genes into *E. coli* expression vectors

Coding sequences of *S. cerevisiae LSM1* to *8* were cloned into pET19b (pMR1-8, 32) enabling T7 promoter driven expression of N-terminally His-tagged proteins in BL21(DE3) cells or other T7 expression systems. Predicted coding sequences were taken from the *Saccharomyces* Genome Database (SGD) and cloned, PCR-amplified sequences were inspected to ensure absence of any mutations. Intronless *LSM2* and *LSM7* were amplified by PCR using previously cloned sequences (Mayes, 1998) as template. Two new vectors were created with a single HA (pMR81) or Myc tag

(pMR82) inserted between the pET19b 10xHis-tag and the multiple cloning site, allowing for the production of double tagged proteins. All Lsm coding sequences were cloned into these new vectors (pMR11-18, 21-28, 33, 34). A different, shorter *LSM6* open reading frame (ORF) was used than that proposed by Mayes *et al.* (1999) and indicated on SGD. Bouveret *et al.* (2000) suggested a new start site for *LSM6* because it is consistent with the protein's apparent molecular mass of 9.3 kDa, and is supported by comparison of the yeast Lsm6p sequence to homologues from other organisms (Bouveret *et al.*, 2000). The size of the protein is therefore 86 amino acids (aa) instead of 123 aa. This change was later confirmed by genome-wide mapping of yeast transcription start sites (Zhang and Dietrich, 2005) and amended by SGD.

3.2.2 Confirmation of a change in *LSM8* ORF designation

The *LSM8* ORF proposed by SGD was initially used to create expression vectors for the production of recombinant Lsm8p. When *S. cerevisiae* Lsm8p was aligned with orthologues from other organisms a discrepancy between the N-terminus of yeast Lsm8p (scLsm8p) and other Lsm8 polypeptides became apparent (Fig. 3.1A). Further investigation revealed that the *LSM8* ORF designation had been changed based on the automated comparison of closely related *Saccharomyces* species (Kellis *et al.*, 2003). These authors suggested to move the start site for *LSM8* 57 nucleotides (nt; 19 aa) downstream. This is also the predicted start methionine in the majority of *Saccharomyces* species orthologues (Kellis *et al.*, 2003; Cliften *et al.*, 2003) and significant sequence conservation begins abruptly at this methionine (Fig. 3.1B). To confirm this new start site primer extension was performed to determine the length of the 5' untranslated region (UTR) of *LSM8*. The main product after primer extension with Lsm8Rpe2 was approximately 128 nt long (Fig. 3.2A), suggesting a short 5' UTR for *LSM8* of only about 8 nt consistent with the new suggested start site. The majority of yeast leader sequences are upward of 25 to 30 nt, but 5' UTRs of 8 nt or even shorter do occur (Zhang *et al.*, 2005). Unfortunately the *LSM8* 5' UTR was not mapped in the Zhang study. A recent large-scale cDNA analysis using a vector-capping method to clone the yeast transcriptome reported the *LSM8* 5' UTR to be 15 nt long (Miura *et al.*, 2006). The authors of this study

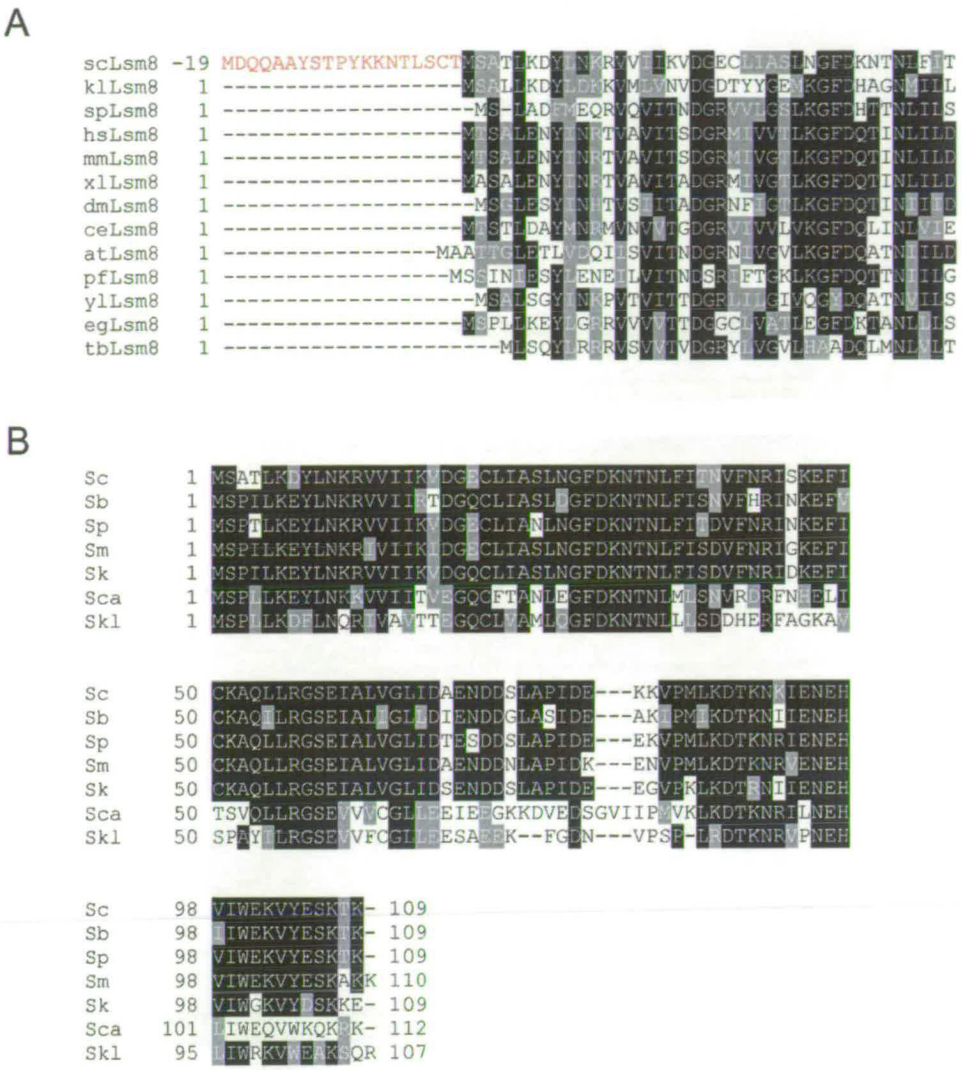


Fig. 3.1 The *S. cerevisiae LSM8* ORF was initially predicted incorrectly as alignments with Lsm8p sequences from different organisms suggest. **(A)** Alignment of Lsm8p sequences from 13 different organisms suggests that the predicted *S. cerevisiae* (sc) sequence may be 19 amino acids too long. kl = *Kluyveromyces lactis*; sp = *Schizosaccharomyces pombe*; hs = *Homo sapiens*; mm = *Mus musculus*; xl = *Xenopus laevis;* dm = *Drosophila melanogaster*; ce = *Caenorhabditis elegans*; at = *Arabidopsis thaliana*; pf = *Plasmodium falciparum*; yl =*Yarrowia lipolytica*; eg = *Eremothecium gossypii*; tb = *Trypanosoma brucei*; **(B)** Alignment of Lsm8p from seven different *Saccharomyces* species shows that these 19 amino acids are not conserved amongst related yeasts either, confirming a new start for *S. cerevisiae LSM8*. Sc = *Saccharomyces cerevisiae*; Sb = *S. bayanus*; Sp = *S. paradoxus*; Sm = *S. mikatae*; Sk = *S. kudriavzevii*; Sca = *S. castellii*; Skl = *S. kluyveri*

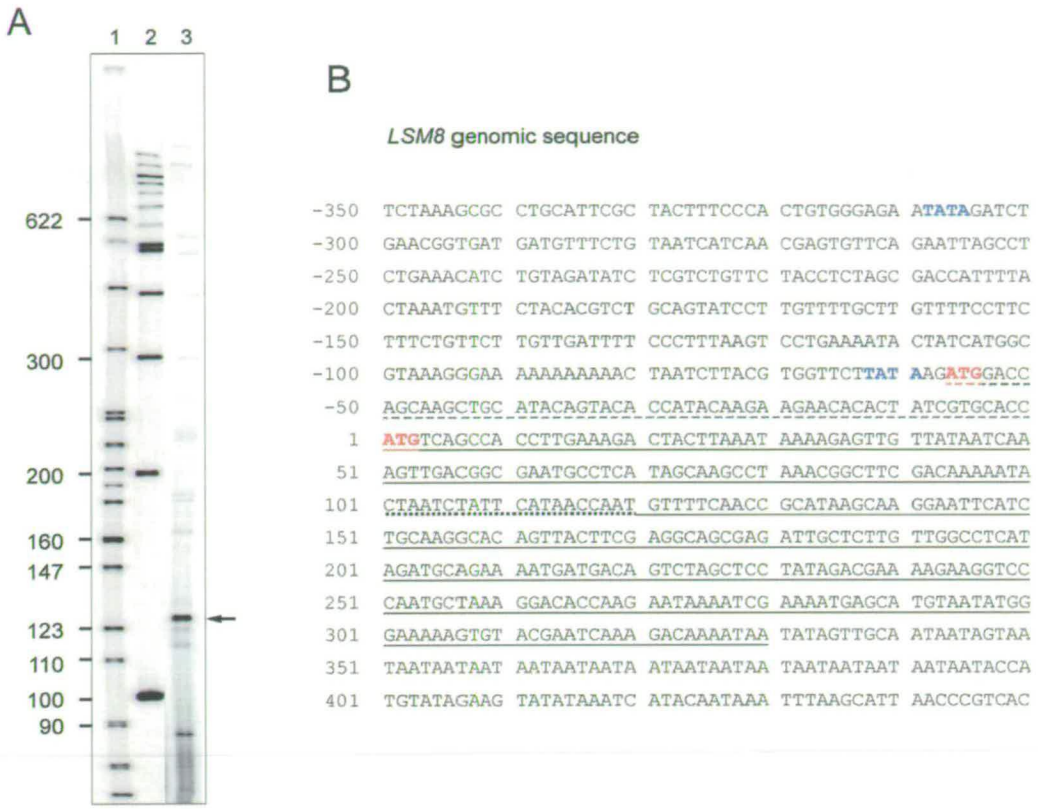


Fig. 3.2 Verification of new *LSM8* start by primer extension. **(A)** Primer extension with primer Lsm8Rpe2 on yeast total RNA (3), with labelled pBR322 *Msp*I digest (1) and NEB 100 bp ladder (2). The main product is estimated to be 128 nt long (see arrow), indicating that the *LSM8* 5' UTR from the new start may only be 8 nt long; **(B)** *LSM8* coding sequence and surrounding genomic sequence. Potential TATA-boxes are indicated in blue, old and new ATG start codons in red; the predicted coding sequence is underlined with the old prediction dashed; the Lsm8Rpe2 annealing site is underlined with a dotted line. Numbering starts with +1 at the new ATG start. The putative TATA box at position -63 is a favourable distance away from the new transcription start site. The new start codon is in a sequence context favourable for translation intiation.

showed variable lengths of 5' UTRs for many transcripts that were cloned multiple times. In fact, the vast majority of yeast genes was shown to have two or more transcription start sites. As the *LSM8* transcript was only cloned once, there may be alternative 5' UTRs for *LSM8* as well. To further confirm the new start site the genomic sequence surrounding the *LSM8* start-ATG was examined (Fig. 3.2B). TATA elements in *S. cerevisiae* are typically located 40 to 120 nt upstream of the transcription initiation site and have a TATAAA consensus. Derivatives of this sequence can also confer TATA function, possibly with diminished activity (Hampsey, 1998). A TATAAG sequence 63 nt upstream of the new start ATG fits these criteria. In addition, the new start codon seems to have a sequence context favourable for translation initiation. Protein synthesis is regulated by the sequence and structure of the 5' UTR of the mRNA transcript. In mammals, the Kozak sequence GCCACCAUGG, which often lies within a short 5' UTR region, directs translation of mRNA (Kozak, 1991). In yeast, the preferred sequence context is $^{A}/_{y}\mathbf{A}^{A}/_{u}$AAUGUCU with a particular preference for A in position -3 and U in position +4 (Yoon and Donahue, 1992).

In summary, the correct *LSM8* ORF is 130 nt long and encodes a 109 aa (12.4 kDa) polypeptide with a pI of 7.18. The *LSM8* transcript has a short 5' UTR of 8-15 nt. This correct *LSM8* coding sequence was used throughout the work presented in this thesis, unless otherwise stated.

3.3 Optimisation of expression and purification of individual Lsm proteins

Because of problems with solubility of recombinant Lsm proteins overexpressed in *E. coli* a range of expression conditions were tested to optimise the level of expression as well as solubility, whilst minimising degradation. Details are given in Materials and Methods (2.4.7). The best conditions for general expression of Lsm proteins were 16 h expression at 23°C in the presence of 0.1 mM IPTG or 24 h expression at 30°C using the ONEX system (Novagen; see Materials and Methods; Fig. 3.3), although further optimisation for individual Lsm proteins may still be

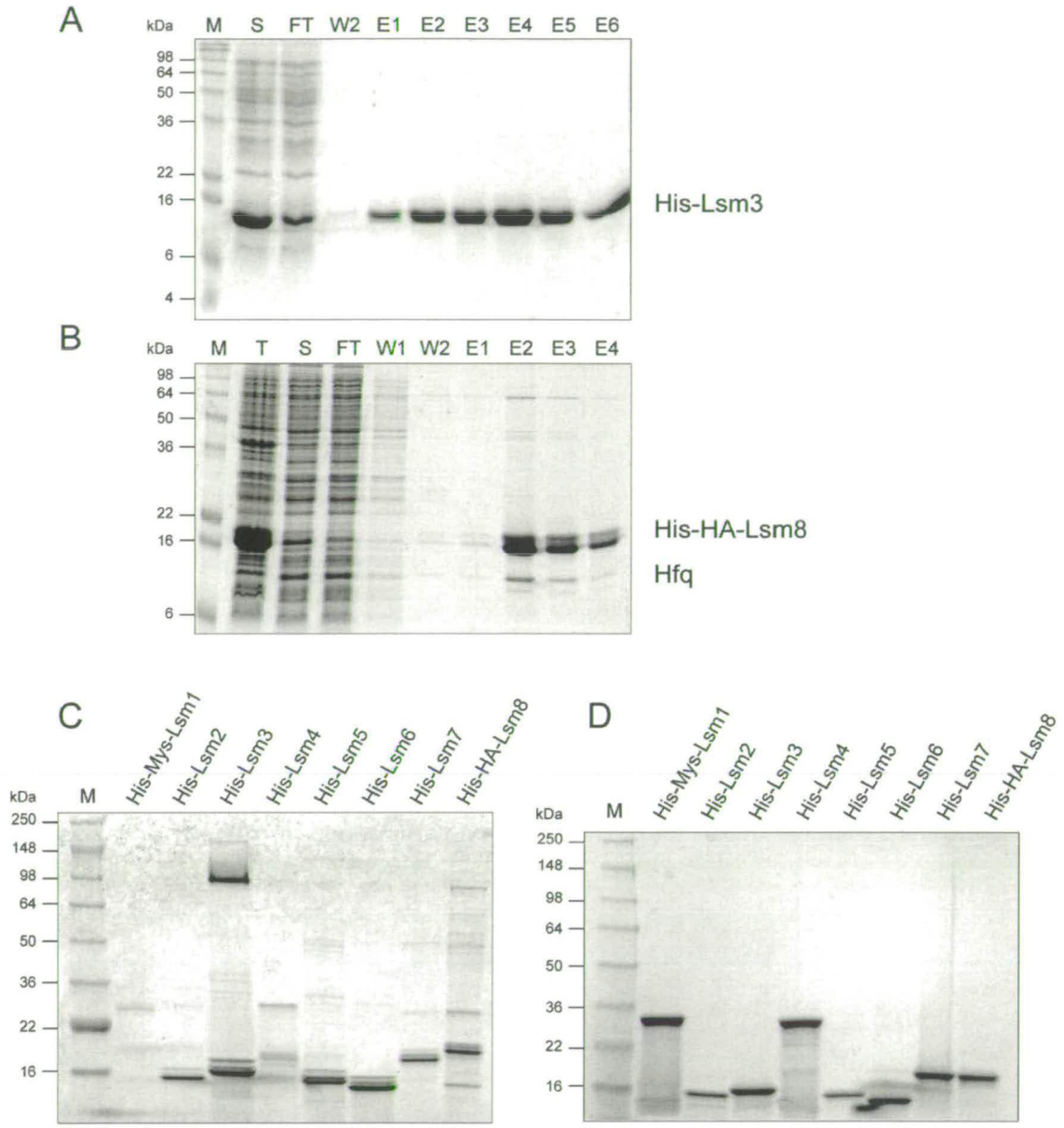


Fig. 3.4 Ni-NTA affinity purification of yeast Lsm proteins expressed in *E. coli*. **(A)** His-Lsm3 purified under native conditions after expression in BL21(DE3) cells; **(B)** His-HA-Lsm8 purified under native conditions after expression in BL21(DE3) cells; **(C)** Recombinant Lsm proteins expressed in BL21(DE3) cells and purified under native conditions; **(D)** Purified recombinant Lsm proteins after expression in MRE4 cells and purification in the presence of 8 M urea. M = SeeBlue Plus2 marker; T = total protein; S = soluble proteins; FT = Ni-NTA agarose unbound fraction; W1 = 20 mM imidazole wash; W2 = 55 mM imidazole wash; E1-E6 = elution fractions.

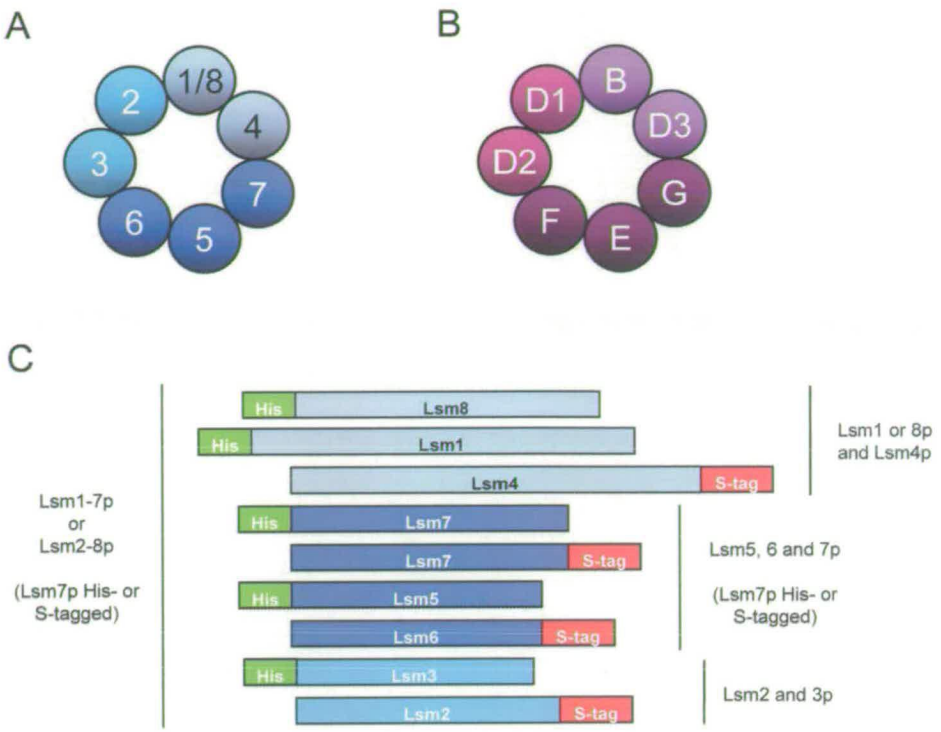


Table 3.1 Levels of expression and solubility of co-expressed recombinant Lsm proteins

Proteins	Expression (ONEX 37°C)	Solubility (ONEX 37°C)	Expression (ONEX 23°C)	Solubility (ONEX 23°C)
Lsm1/Lsm4	+++	+	+++	+++
Lsm3/Lsm2	+++	+++	+++	+++
Lsm5/Lsm6	+++		+++	+++
Lsm8/Lsm4	+++	++		
Lsm5/6, Lsm7-Stag	++		+++	+++
Lsm5/6, His-Lsm7	+		++	+
Lsm1-7, Lsm7-Stag	+		++	+++
Lsm1-7, His-Lsm7	+/-		+	+
Lsm2-8, Lsm7-Stag	+		++	+++
Lsm2-8, His-Lsm7	+/-		+	+/-

Fig. 3.5 Co-expression of Lsm proteins improves solubility. **(A)** Proposed order of Lsm proteins in the ring-shaped complex, based on homology of individual proteins to those in the Sm complex **(B)** and yeast two-hybrid analysis; **(C)** Recombinant Lsm proteins were cloned for co-expression in pairs corresponding to the known Sm subcomplexes: SmB/D3 (Lsm1/4 or Lsm8/4), SmD1/D2 (Lsm2/3) and SmF/E/G (Lsm6/5/7). **Table 3.1** gives scores for levels of expression and solubility for different combinations of co-expression.

Lsm5 and 6 (pMR47). *LSM7* was cloned in pETDuet-1 either with the N-terminal His-tag or the C-terminal S-tag (pMR40 and 41). Pairs of Lsm proteins, as well as Lsm5-6-7p, Lsm1-7p and Lsm2-8p were co-expressed in BL21(DE3) at 23 or 37°C using the ONEX system. Expression levels were generally high (Table 3.1), but not always equal for co-expressed proteins. Expression levels seemed to be better for expression at 23 than 37°C when more than two proteins were co-expressed at one time. The level of solubility was examined by SDS-PAGE of total, soluble and insoluble protein after cell lysis. Western analysis with α-His-HRP and S-protein-HRP allowed for comparison of the levels of recombinant protein in all fractions. Solubility of pairs of proteins was significantly improved (>90% soluble; Table 3.1) compared to that of individually expressed Lsm proteins (data not shown). Surprisingly there was a significant difference between solubility of Lsm5-6-7p, Lsm1-7p and Lsm2-8p including N-terminally His-tagged or C-terminally S-tagged Lsm7p. Reduced solubility with His-Lsm7p suggests that the N-terminal tag may interfere with Lsm7p's interaction with its neighbouring subunits, thus making them more susceptible to aggregation.

3.4.1 Co-expressed pairs of Lsm proteins can be co-purified

Improved solubility is presumed to be the result of formation of hetero-multimeric complexes that are less prone to aggregation than monomers or homo-multimeric complexes. This was also shown to be the case for the human LSm proteins (hLSm) with stable complexes formed for hLSm2/3, hLSm4/8 and hLSm5/6/7 (Zaric *et al.*, 2005). Surprisingly however, in the Zaric study, improved solubility did not depend on co-expression of assumed nearest neighbours as other combinations showed similarly improved solubility. This was suggested to be due to a low level of LSm-LSm interaction specificity. Nonetheless, I decided to test whether co-expressed yeast Lsm proteins could be co-purified by affinity chromatography using the His-tag on one of the proteins. Lsm2p and Lsm3p (pMR46), as well as Lsm5p and Lsm6p (pMR47) were co-expressed in BL21(DE3) cells using the ONEX system at 23°C. Soluble protein after cell lysis was used for Ni-NTA affinity chromatography under native conditions. Washes were performed with 20 mM and 55 mM imidazole and finally proteins were eluted with 250 mM imidazole. Lsm2-Stag was co-purified

with His-Lsm3 (Fig. 3.6A) and Lsm6-Stag was co-purified with His-Lsm5 (Fig. 3.6B), indicating that co-expressed Lsm proteins interact. These interactions may not be specific to assumed nearest neighbours, as was shown by Zaric *et al.* (2005) for the human LSm proteins, but lead to decreased aggregation and improved solubility.

3.4.2 Co-expression and purification of Lsm1-7p and Lsm2-8p

Lsm1-7p and Lsm2-8p were co-expressed using S-tagged Lsm7p in BL21(DE3) cells and the ONEX system at 23°C. Soluble protein after cell lysis was used for Ni-NTA affinity chromatography under native conditions. Washes were performed with 20 mM and 55 mM imidazole and finally proteins were eluted with 250 mM imidazole. Proteins were expressed at similar levels and were mostly soluble. Exceptions were His-Lsm1 and Lsm2-Stag, with His-Lsm1 expressed at significantly higher levels and Lsm2-Stag expressed at extremely low levels. All proteins were purified at similar levels, with the same exceptions of Lsm1p and Lsm2p (Fig. 3.6C and D). Efficient co-purification of S-tagged Lsm proteins with His-tagged Lsm proteins suggest that heteromultimeric complexes are formed during expression in *E. coli*. The absence of high levels of Lsm2p suggests that the majority of the formed complexes may not contain Lsm2p. Alternatively, Lsm2p and/or its C-terminal S-tag may be particularly susceptible to proteolytic cleavage leading to reduced levels of Lsm2p that are detectable using S-protein-HRP. However, the fact that no such problems occurred during co-expression with His-Lsm3p (Fig. 3.6A) argues against this. These results show that it should be possible to co-express seven recombinant Lsm proteins at once and to co-purify them using affinity tags. Additional purification steps will be required to obtain homogenous complexes that contain all seven subunits. Use of combinations of dual expression vectors, as was done here, may not be optimal. Likely problems include differences in expression levels due to variable plasmid copy-numbers and possible interference of the large number of tags with Lsm complex formation and functionality. A better approach would be to create a heptacistronic expression vector with only one protein carrying a cleavable affinity tag. This approach is challenging in terms of devising a successful cloning strategy and because of potential problems with instability of such a construct due to recombination (Christian Kambach, personal communication).

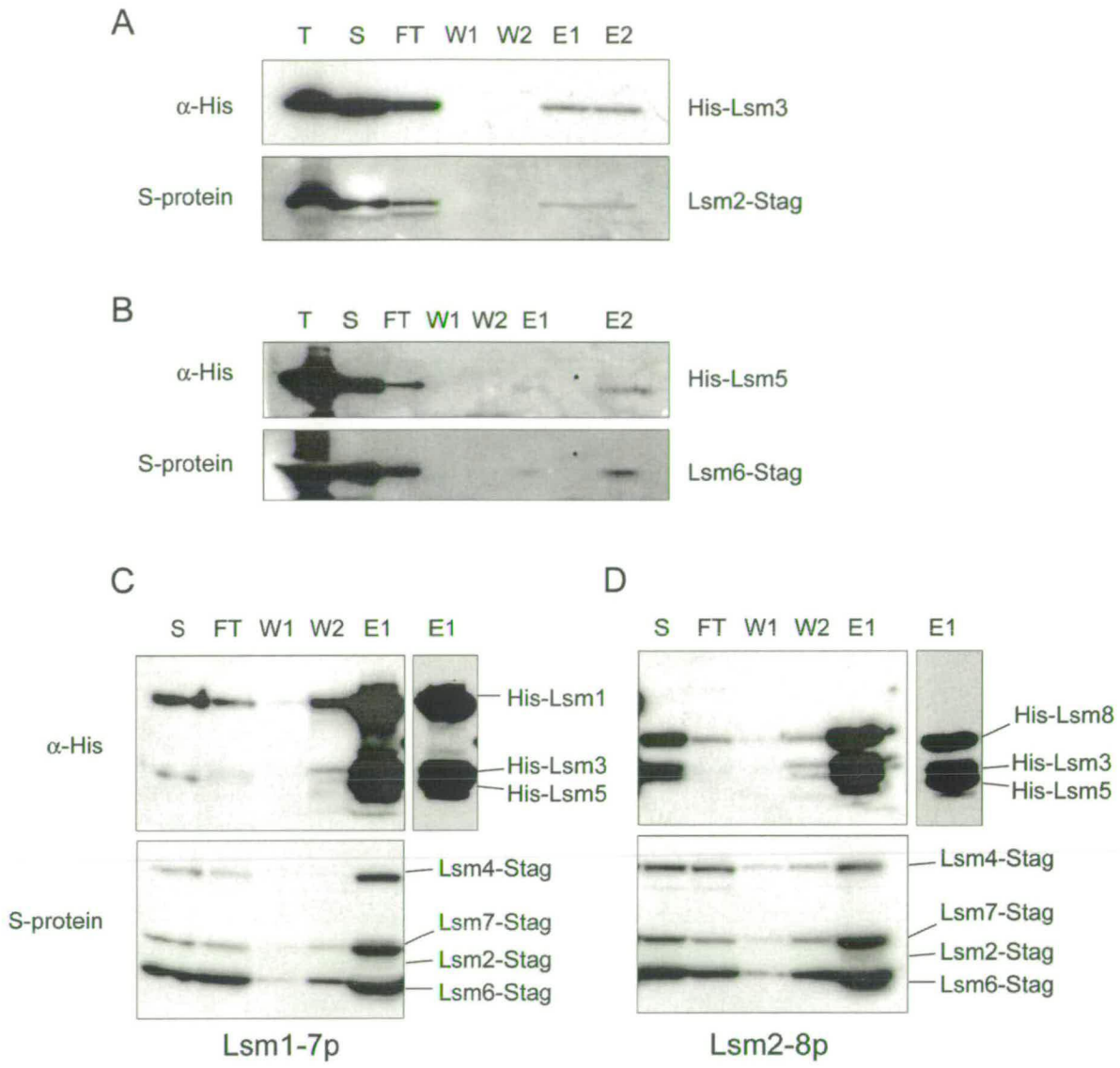


Fig. 3.6 Co-expressed Lsm proteins can be co-purified. Recombinant Lsm proteins were co-expressed using (combinations of) dual expression vectors in BL21(DE3) cells and the ONEX system at 23°C, and purified by Ni-NTA chromatography. Samples from various fractions were separated by SDS-PAGE and analysed using α-His and S-protein-HRP Westerns. **(A)** His-Lsm3 purified with Ni-NTA agarose co-purifies Lsm2-Stag; **(B)** His-Lsm5 similarly co-purifies Lsm6-Stag; **(C)** Co-expression of His-Lsm1, 3 and 5 with Lsm2, 4, 6 and 7-Stag allows purification of all Lsm proteins by Ni-NTA affinity chromatography, although levels of Lsm2-Stag are extremely low; **(D)** Co-expression of Lsm2 to 8 proteins similarly allows their co-purification, although levels of Lsm2-Stag are low here also. T = total protein; S = soluble protein; W1 = 20 mM imidazole wash; W2 = 55 mM imidazole wash; E = elution with 250 mM imidazole

3.5 Lsm complex formation *in vitro*

3.5.1 Re-naturation of a mixture of denatured Lsm proteins allows for complex assembly *in vitro*

An alternative strategy to *in vivo* complex formation by co-expression was *in vitro* assembly of complexes from purified components. Initial attempts to create Lsm1-7p and Lsm2-8p complexes *in vitro* by mixing together individually purified Lsm proteins under native conditions were unsuccessful (data not shown). This is likely to be due to formation of homomultimeric complexes of which Lsm3p is a prime example. To allow for heteromultimeric complex formation these self-interactions need to be disturbed first. His-Lsm2 to -Lsm7p were individually expressed in BL21(DE3) cells and affinity purified under denaturing conditions, i.e. in the presence of 8 M urea (Fig. 3.7A). Approximately equimolar amounts of each of these monomers were then mixed with natively purified His-Myc-Lsm1p or His-HA-Lsm8p, and dialysed against buffer without urea. Soluble and insoluble proteins were separated after dialysis by centrifugation. Samples of each were analysed by SDS-PAGE and Coomassie staining. Yields of soluble proteins varied from experiment to experiment, but generally between 50 and 90% of proteins were soluble after dialysis (Fig. 3.7B). Approximately equimolar amounts of each of the Lsm monomers were recovered, suggesting that correct heteroheptameric complexes may have formed. To show that all proteins were present in the same complex the Myc- and HA-tags of Lsm1p and Lsm8p respectively were used to immunoprecipitate these proteins. Both Coomassie staining after SDS-PAGE and Western analysis with α-His antibody showed that all Lsm proteins were pulled down by Lsm1p and Lsm8p (Fig. 3.7C). These results show that it is possible to assemble heteromultimeric Lsm complexes *in vitro* by re-naturation of a mixture of denatured Lsm monomers. The homogeneity of the resulting complexes is as yet uncertain.

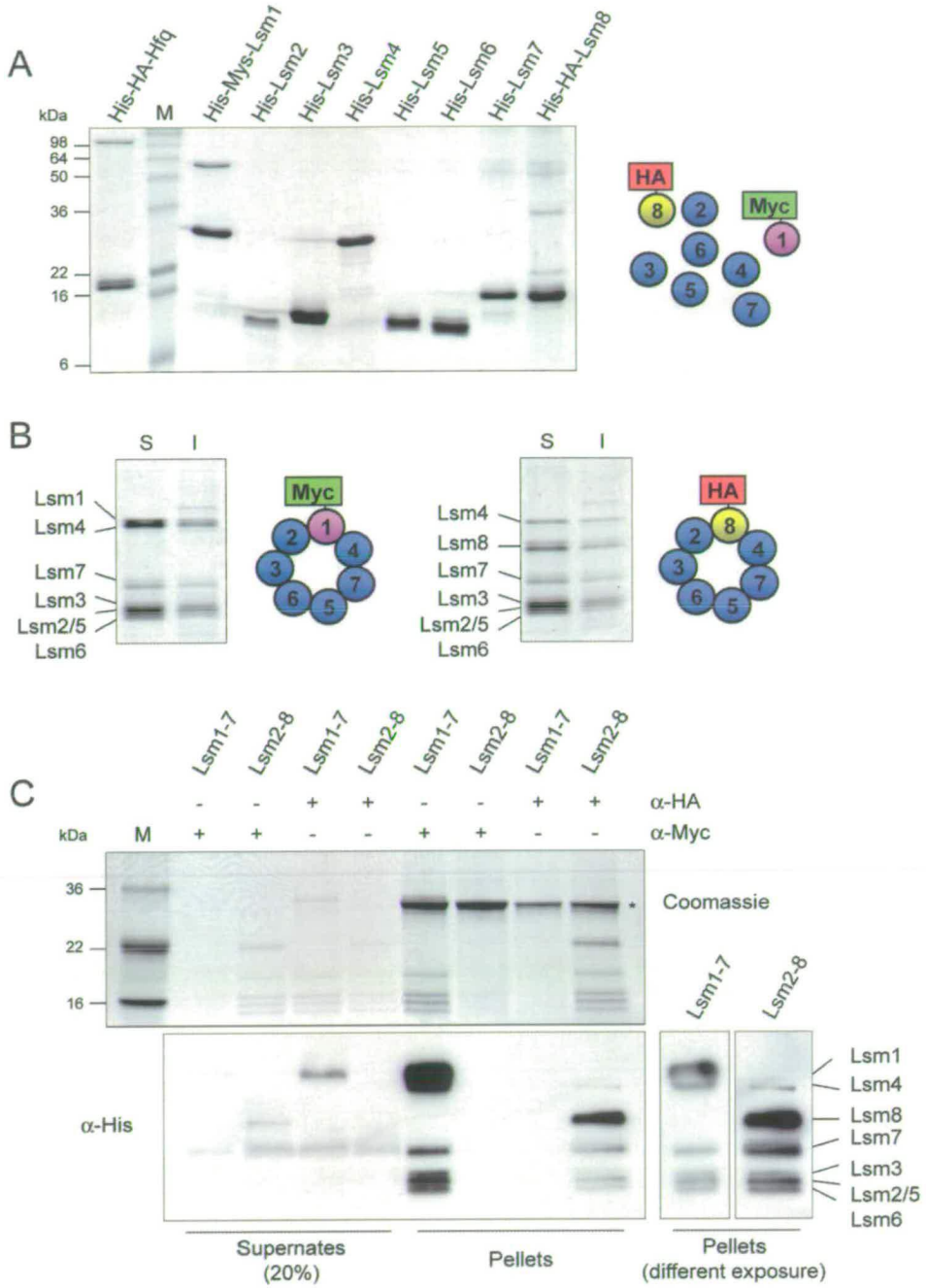


Fig. 3.7 *In vitro* assembly of yeast Lsm complexes. **(A)** His-Myc-Lsm1, His-Lsm2 to 7 and His-HA-Lsm8, affinity purified under denaturing conditions, were separated by SDS-PAGE and Coomassie stained; **(B)** Natively purified His-Myc-Lsm1 or His-HA-Lsm8 were mixed with His-Lsm2 to 7 purified with 8 M urea and re-natured by dialysis. Soluble (S) and insoluble (I) fractions were analysed by SDS-PAGE and Coomassie staining. **(C)** Soluble fractions of Lsm1-7 and Lsm2-8 were used for immunoprecipitation with α-Myc and α-HA antibodies. Pellets and supernates were separated by SDS-PAGE and Coomassie stained or Western blotted and hybridised with α-His-HRP antibody. M = See Blue Plus2 marker (Invitrogen) * Light chain of antibodies used in CoIP.

3.5.2 Soluble complexes can form in the absence of single subunits

It seems that Lsm proteins are promiscuous in their interactions. Formation of the correct heteromultimeric complexes *in vivo* therefore presumably depends on the fact that interactions within the Lsm1-7 and Lsm2-8 complexes are the most energetically favourable. In higher eukaryotes the SMN complex contributes to formation of the correct complexes, but in *S. cerevisiae* no such snRNP assembly chaperone is known to exist. The success of these *in vitro* assembly experiments also depends on formation of these most energetically favourable complexes. To test whether formation of correct complexes can be measured by the yield of soluble protein after dialysis, various combinations of six Lsm proteins were mixed and re-natured. Soluble and insoluble proteins were separated after dialysis by centrifugation and samples of each were separated by SDS-PAGE. Coomassie stained gels showed that yields of six-protein mixtures were not significantly different from those that contained all seven Lsm proteins (Fig. 3.8). This suggests that heterohexameric Lsm complexes can form *in vitro*. Whether this is a reflection of the potential existence of such hexameric complexes *in vivo* is unclear, but it does mean that: a) levels of soluble protein after re-naturation are not necessarily a measure of correct heteroheptameric complexes; b) hexameric complexes may form in mixtures of seven proteins with a heterogeneous mixture of complexes as a result.

3.5.3 Lsm complexes aggregate and/or form higher order structures

To investigate the homogeneity of the *in vitro* assembled Lsm1-7 and Lsm2-8 complexes size exclusion chromatography was performed. In addition to a measure of homogeneity (elution in a single peak), size-exclusion chromatography allows an estimation of the apparent molecular weight of a protein or protein complex which corresponds to the time it takes the complex to leave the column. Lsm1-7 and Lsm2-8 complexes were assembled *in vitro* as described. They were concentrated to approximately 1 mg/ml and subjected to size exclusion chromatography on a SUPERDEX 200 HR10/30 column (data not shown) or a SUPEROSE 6 HR10/30 column (Fig. 3.9). Proteins eluting off the column were detected by measuring the

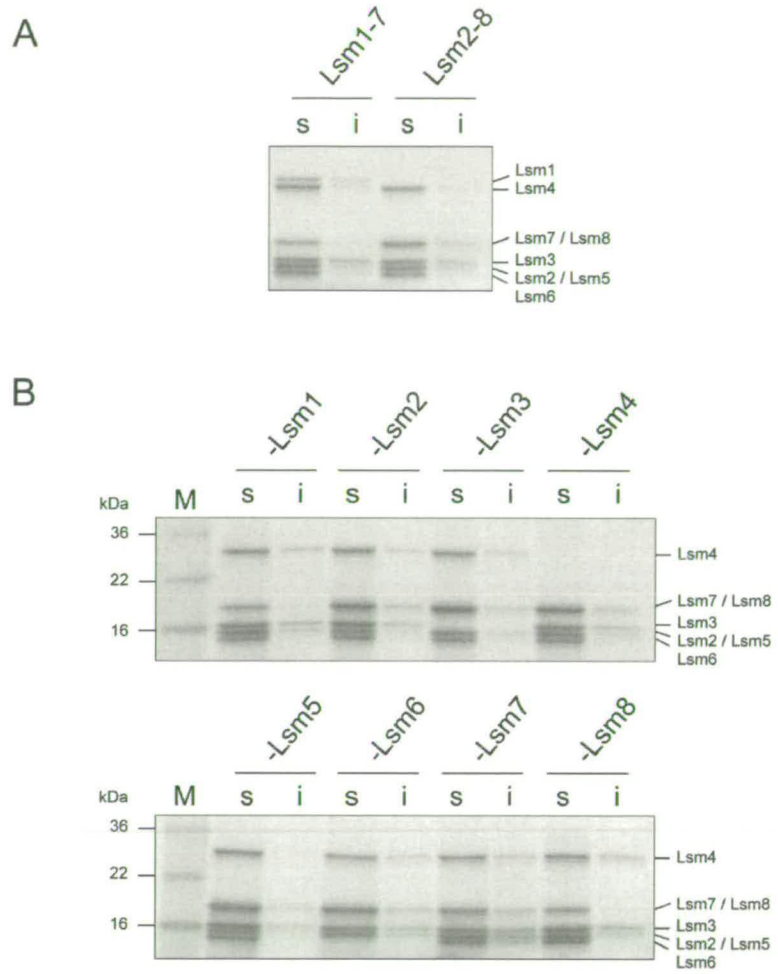


Fig. 3.8 *In vitro* Lsm complex formation with six subunits. To test if soluble complexes can form in the absence of one of the seven Lsm subunits, six of them were mixed under denaturing conditions and re-natured as described. Soluble (s) and insoluble (i) proteins after dialysis were separated by centrifugation and analysed by SDS-PAGE and Coomassie staining. M = SeeBlue Plus2 Protein Marker (Invitrogen) **(A)** Soluble and insoluble fractions of mixtures of Lsm1 to 7p and Lsm2 to 8p after dialysis; **(B)** Soluble and insoluble fractions of mixtures of 6 different Lsm proteins after dialysis

absorbance of the eluate at a wavelength of 280 nm. The extinction coefficients for Lsm1-7p and Lsm2-8p were estimated at 0.55 $(mg/ml)^{-1}cm^{-1}$ and 0.52 $(mg/ml)^{-1}cm^{-1}$ respectively using the online Protein Calculator (see section 2.7). Here, extinction coefficients are estimated based on the number of tryptophans and tyrosines using an absorbance curve of known lyophilized proteins (Gill and von Hippel, 1989). As the proteins are strongly diluted between loading onto and eluting off the column, A_{280} readings are low, reducing the accuracy and making them sensitive to background noise. Nonetheless, both complexes eluted off the column in a single main peak, with a shoulder, indicating a major and a minor peak that are not properly resolved (Fig. 3.9A and C). Additional peaks were detected corresponding to molecular weights of less than 10 kDa. To determine protein content of each of the peaks, the appropriate fractions were collected, TCA precipitated and analysed by SDS-PAGE and Coomassie staining (Fig. 3.9B and D). All Lsm proteins were detected in fractions corresponding to the main peak and shoulder, but no proteins were detected in the <10 kDa peaks, suggesting that the latter may have been due to background noise. Lsm7p and Lsm8p, and Lsm2p and Lsm5p migrate in single bands during SDS-PAGE. To confirm the presence of all Lsm polypeptides these bands were submitted to mass spectrometry (2.4.12). In the Lsm1-7p Lsm2/5 band Lsm2p was covered by 8 peptides (57% aa) and Lsm5p by 3 peptides (31% aa). In the Lsm2-8p Lsm2/5 band Lsm2p was covered by 12 peptides (73% aa) and Lsm5p by 2 peptides (19%). The Lsm7/8 band produced 2 Lsm8p peptides (13% aa) and 1 Lsm7p peptide (8%). In addition most bands produced small numbers of peptides from Lsm6p and some of the other smaller Lsm proteins. In conclusion, the main peaks contain all seven Lsm polypeptides. The apparent molecular weights corresponding to the main peaks and shoulders were calculated. The molecular weight corresponding to the Lsm1-7p main peak was estimated at 530-580 kDa, that of the minor peak at 300-310 kDa, corresponding to approximately 5 and 3 Lsm1-7p heptamers respectively. Calculated molecular weights corresponding to the Lsm2-8p peaks were more variable between experiments and were 560-730 kDa for the main peak and 250-310 for the minor peak, corresponding to 5-7 and 2-3 Lsm2-8p heptamers respectively. Increasing the salt concentration from 0.1 M to 0.5 M KCl did not significantly change the outcome.

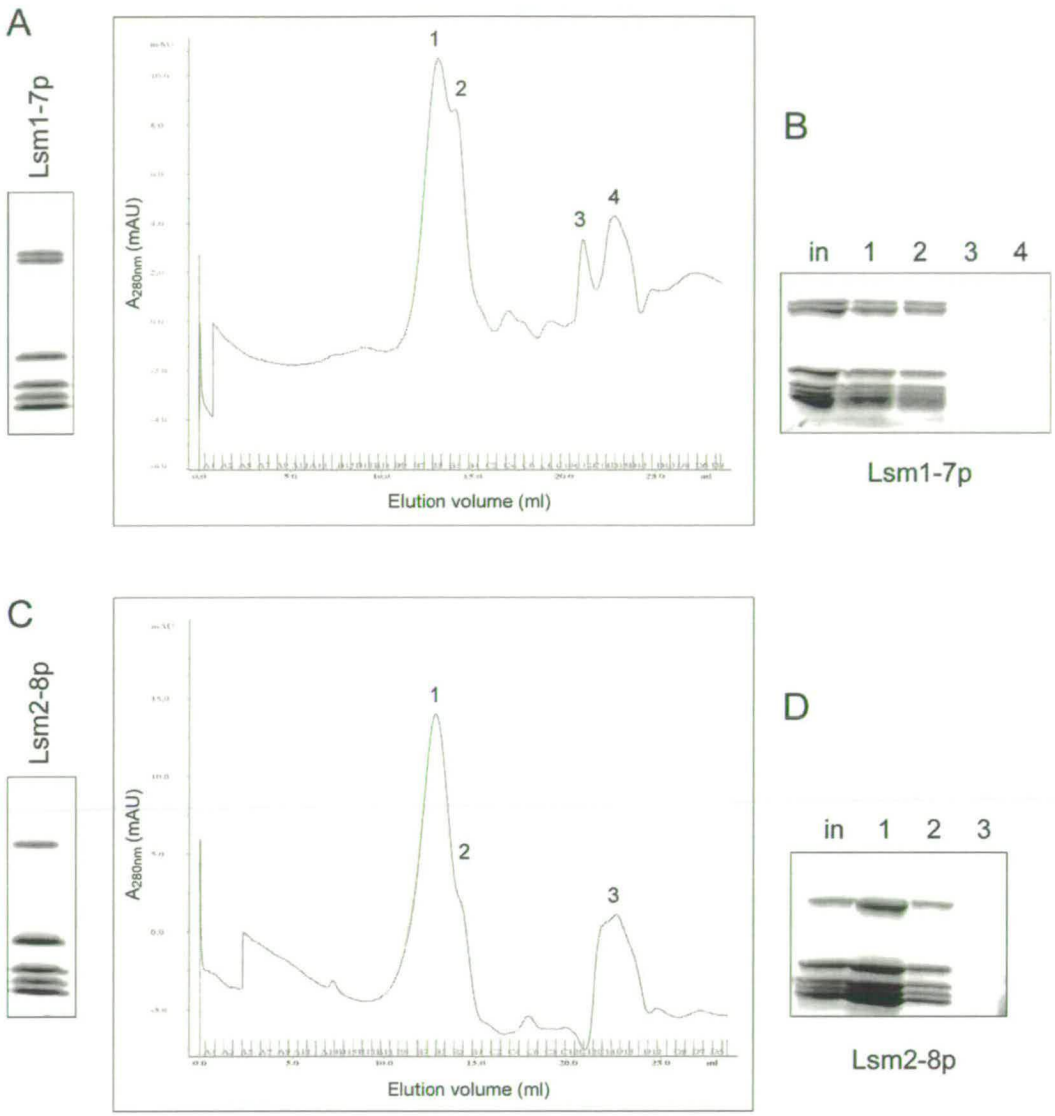


Fig. 3.9 Size exclusion chromatography of *in vitro* assembled yeast Lsm1-7p and Lsm2-8p suggests these complexes aggregate. **(A)** *In vitro* assembled Lsm1-7p was separated on a SUPEROSE 6 size exclusion chormatography column into 4 main peaks; **(B)** Fractions corresponding to peaks 1-4 were TCA precipitated, separated by SDS-PAGE and Coomassie stained; **(C)** *In vitro* assembled Lsm2-8p was separated into 3 main peaks; **(D)** Fractions corresponding to peaks 1-3 were TCA precipitated, separated by SDS-PAGE and Coomassie stained.

In summary, the *in vitro* assembled Lsm complexes do not resolve like heptameric complexes during size exclusion chromatography. This may well reflect multimerisation of (heptameric) complexes, in this case into structures containing anywhere between 2 and 7 Lsm rings. As it has been shown that Sm-like complexes are prone to aggregation, in some cases into well-ordered higher order structures (Arluison *et al.*, 2006), this is not unlikely.

3.5.4 *In vitro* assembly of human LSm complexes

Purified human LSm protein (hLSm) subcomplexes (kindly provided by Christian Kambach) were used to assemble hLSm complexes *in vitro* as described in Materials and Methods (2.4.10.2). hLSm1, hLSm2/3, hLSm4, hLSm5/6/7 and hLSm4/8 (Fig. 3.10A) were mixed at equimolar amounts under semi-denaturing conditions and dialysed to form hLSm1-7 or hLSm2-8. Yields were generally high (>90%) and although not all protein bands could be resolved by SDS-PAGE relative intensities of Coomassie stained bands and published data (Zaric *et al.*, 2005) suggest that all seven hLSm proteins are present in the resulting complexes (Fig. 3.10B). A number of complexes were assembled with less than seven subunits to act as controls in RNA bandshift, annealing and strand displacement experiments (Chapter 4 and 5): hLSm2-7 (hexamer), hLSm23567 (pentamer without hLSm4), hLSm4-8 (pentamer with hLSm4), hLSm123567 (hexamer without hLSm4) and hLSm48 (dimer). Surprisingly, similar yields were obtained for each of these control complexes (Fig. 3.10C shows a Coomassie stained SDS-PAGE gel with approximately equimolar amounts of each of these complexes). RNA bandshift experiments with these protein mixtures (Chapter 4) suggest that alternative complexes are formed that retain some functionality.

3.5.5 Optimisation of *in vitro* Lsm complex assembly

Comparison of the 4M and 8M urea methods used for assembly of the yeast and human Lsm complexes and their functional integrity (described in Chapters 4 and 5) shows that there is still room for improvement, in particular for *in vitro* assembly of recombinant yeast Lsm complexes. Presence of soluble protein after renaturing dialysis is not necessarily a measure of correctly assembled active complexes.

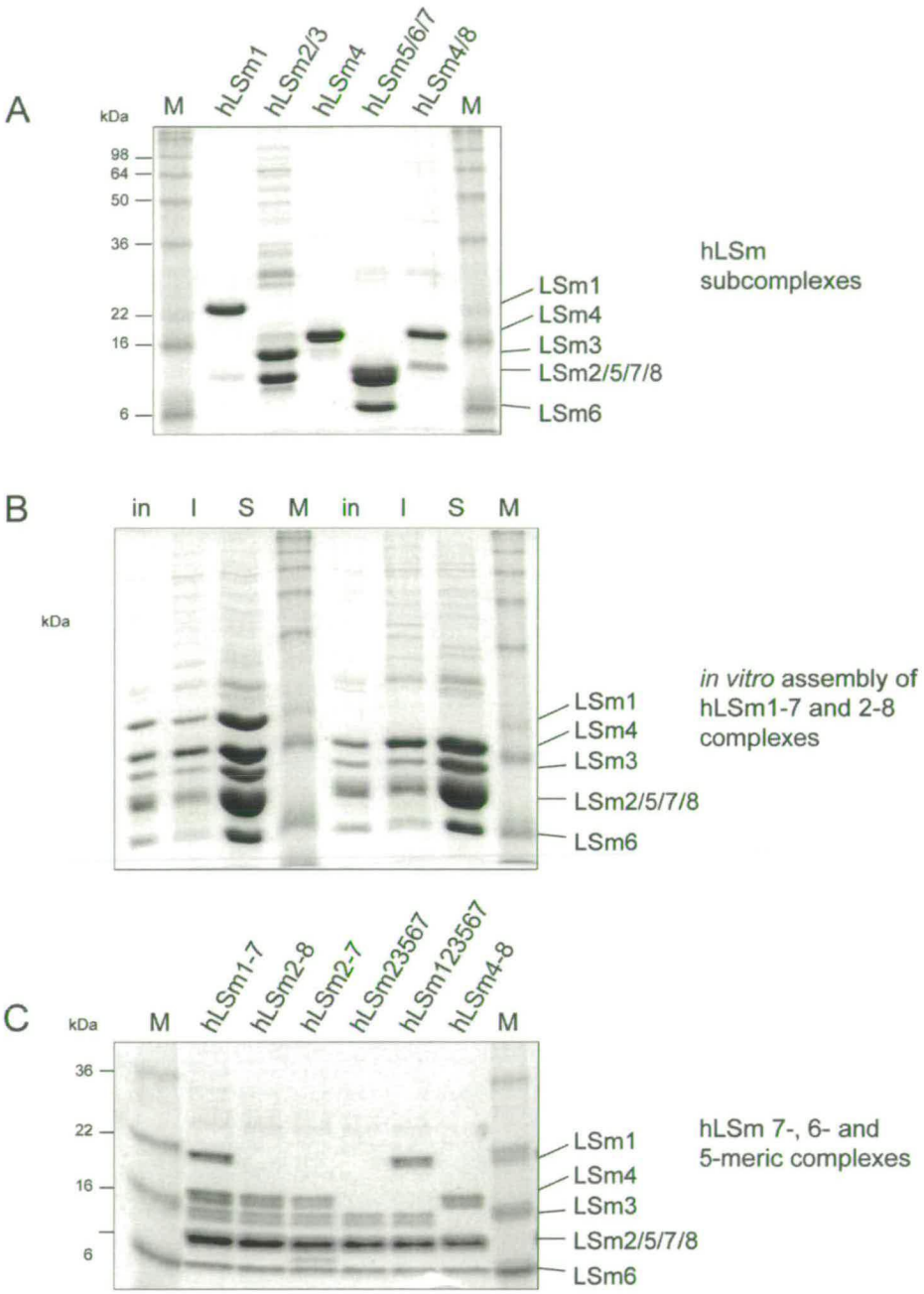


Fig. 3.10 *In vitro* assembly of hLSm complexes. **(A)** Coomassie stained SDS-PAGE gel with purified hLSm proteins and subcomplexes; **(B)** Coomassie stained SDS-PAGE gel with input (in), insoluble (I) and soluble (S) samples before and after dialysis; **(C)** Equimolar amounts of heptameric, hexameric and pentameric hLSm complexes assembled *in vitro*.

Efficiency of U6 RNA binding may be the best and easiest way to measure this (see Chapter 4). Different RNA annealing activities of human LSm complexes assembled from subcomplexes in the presence of 4 M and 8 M urea also shows that 8 M urea may disrupt a structure needed for this activity (section 4.3.3.1). The best way to optimise *in vitro* assembly of yeast Lsm complexes may therefore be to purify subunits under semi-denaturing conditions (e.g. in the presence of 4 M urea), dialysis of purified proteins to buffer of physiological pH in the presence of 4 M urea and assembly into heteromultimeric complexes under conditions similar to those described for the human LSm proteins (2.4.10.2). Strong homomultimerisation by some of the Lsm proteins (yLsm3p in particular) may still prevent successful complex assembly. A combination of co-expression of predicted sub-complexes and their purification under semi-denaturing conditions, analogous to the method used for the human LSm proteins (Zaric *et al.*, 2005), may therefore be necessary. Attempts to purify recombinant yeast Lsm proteins at lower urea concentrations (<8M) to assemble complexes *in vitro* under semi-denaturing conditions were unsuccessful (data not shown).

3.6 Contamination of recombinant Lsm proteins with *E. coli* Sm-like protein Hfq

3.6.1 Lsm8p co-purifies *E. coli* Hfq

His-HA-Lsm8p purified from BL21(DE3) cells under native conditions contained a 10 kDa fragment detectable by Coomassie staining (Fig. 3.4B) but not by α-His Western analysis (data not shown). This was initially thought to be an Lsm8p degradation fragment without the His-tag, but problems with the interpretation of results from functional assays (discussed in Appendix A1) lead me to investigate further. The 10 kDa band was taken from a Coomassie stained SDS-PAGE gel and analysed by mass spectrometry. No Lsm8p peptides, but *E. coli* Hfq peptides were detected, showing that Hfq protein is co-purified with recombinant Lsm8p at significant levels. The nature of the interaction between Lsm8p and Hfq is unclear, but there are some indications that they may form a heteromultimeric complex (Appendix A2). Results from U6 RNA bandshift assays (compare Fig. 4.1A to Fig.

A1.1) suggest that Lsm5p and Lsm6p may also co-purify Hfq, but at lower levels. No additional Coomassie stained bands were observed for these proteins however, as Lsm5p, Lsm6p and Hfq migrate at around the same molecular weight. Hfq is the *E. coli* Sm-like protein, it binds RNA and has been shown to act as an RNA chaperone (Moll *et al.*, 2003). Its presence in functional assays of the yeast Lsm proteins is therefore undesirable.

3.6.2 Creation of an *hfq* deletion strain for Lsm protein production

To prevent contamination by Hfq an *hfq* deletion strain was constructed to be used for the production of Lsm proteins. The method used to make this deletion was based on that of Merlin *et al.* (2002). This method is described in detail under Materials and Methods (2.2.4) and in Fig. 3.11A-D. The standard method was successful for the MG1655 control strain (Fig. 3.11G), but for unknown reasons no deletion strains were obtained from BL21(DE3) after screening 350 colonies by colony PCR. As an alternative, *hfq* was deleted from *hfq-1* mutant strains (kindly provided by Poul Valentin-Hansen) that are based on T7 polymerase expression strains ER2566 and C41(DE3) (MRE2 and 3; Table 2.1.5). These are effectively *hfq* knockout strains and were created by P1 transduction of the *hfq-1 Ω-Kan*^{R} insertion mutation in the *hfq* *Bcl*I site (at nt 117) from strain TX2808 (Tsui *et al.*, 1994). To allow co-expression of Lsm proteins (see section 3.4) using the Kan^R pCOLADuet-1 plasmid and to ensure no fragments of Hfq protein would be produced, the entire Hfq coding sequence with the insertion was removed in both strains (MRE4 and 5; Table 2.1.5). The same procedure for creating a genomic deletion was followed, but resulting colonies were additionally screened for the absence of kanamycin resistance (Fig. 3.11E). Colonies that were kanamycin sensitive were tested by colony PCR and shown to carry the genomic deletion (Fig. 3.11G). In addition *hfq* deletion strains could be detected because of their changed colony morphology, which was described previously (Tsui *et al.*, 1994). Colonies of *hfq* deletion strains generally look more transparent and spread out, an example of which can be seen for MG1655 with and without *hfq* (Fig. 3.11F). The T7-polymerase expressing *hfq* deletion strains (MRE4 and 5; Table 2.1.5) were used for expression and purification of the Lsm proteins that were used in functional assays described in Chapter 4 and 5.

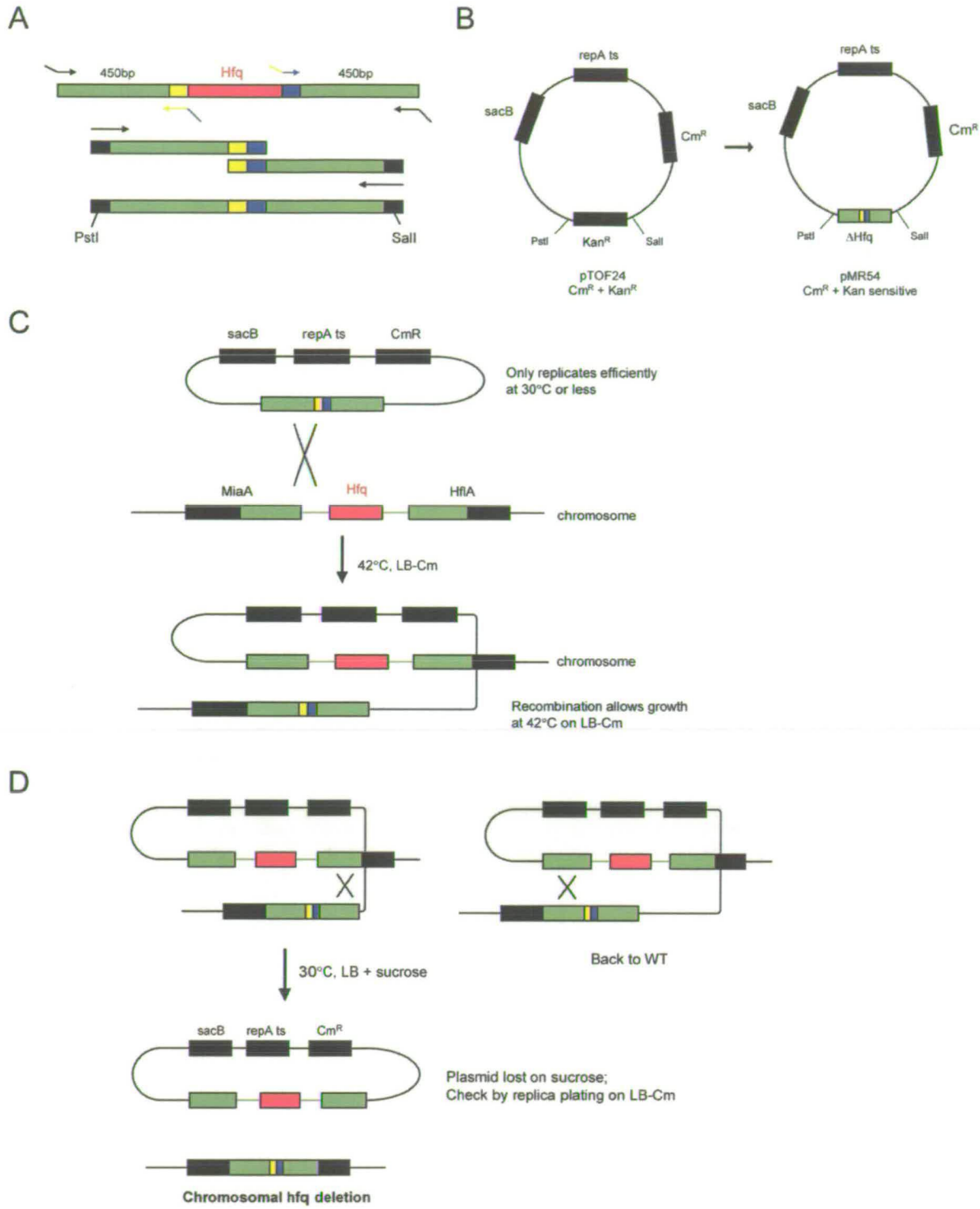


Fig. 3.11 Making a genomic *hfq* deletion in *E. coli*; **(A-D)** The strategy; **(E-G)** Verification of the *hfq* deletion

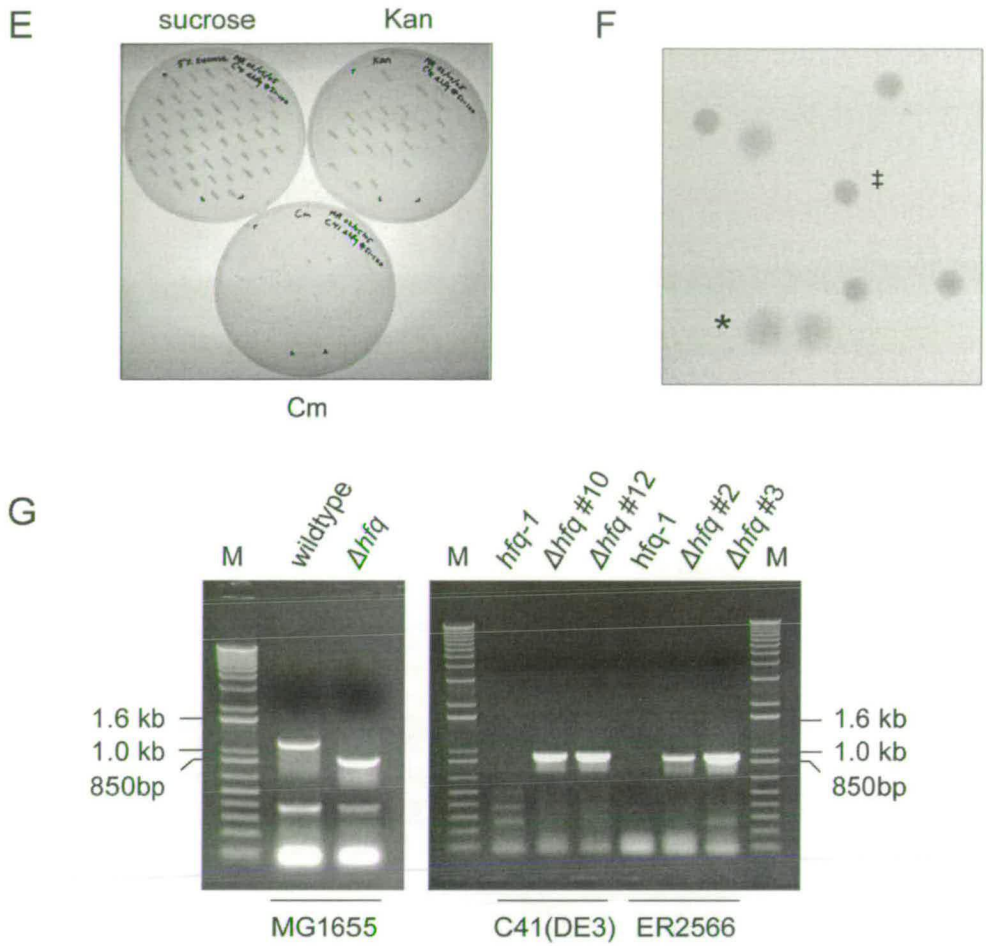


Fig. 3.11 Making a genomic *hfq* deletion in *E. coli*; **(A-D)** The strategy; **(E-G)** Verification of the *hfq* deletion. **(A)** A 900 bp fragment covering 450 bp of *E. coli* genomic sequence on either side of the Hfq coding sequence, with *Pst*I and *Sal*I restriction sites on either end, was created by two-step PCR using primers mro102-105; **(B)** The Kan^R gene from pTOF24 was removed by *Pst*I/*Sal*I digestion and replaced with the 900 bp genomic sequence (pMR54); **(C)** *E. coli* was transformed with pMR54. Subsequent growth on LB-Cm at 42°C promoted recombination of pMR54 onto the genome; **(D)** Growth at 30°C allowed recombination out of the genome with two possible outcomes: revert to wild type or deletion of *hfq*. Growth on LB with 5% sucrose is indicative of plasmid loss (SacB on pTOF24 makes the cells sensitive to sucrose); **(E)** Resulting colonies were tested for growth on sucrose (growth = no plasmid), Cm (no growth = no plasmid) and where appropriate on Kan (no growth = loss of *hfq-1*); **(F)** For MG1655 wild-type (‡)and Δ*hfq* (*) colonies could be distinguished by different colony morphology as previously described by Tsui *et al.* (1994); **(G)** Potentially positive colonies were tested for the *hfq* genomic deletion by colony PCR with primers mro102 and 105.

3.7 Discussion

3.7.1 Change to the *LSM7* annotation

Recent evidence from a large-scale cDNA analysis from the yeast transcriptome (Miura *et al.*, 2006) suggests that the *LSM7* 3' splice site may have been annotated incorrectly (chromosomal coordinate 351055 instead of 351079) leading to insertion of an additional 24 nt in the mRNA and 8 aa (KSENKPQQ) after S6 in the Lsm7 polypeptide. This insertion eliminates a gap in the alignment between Lsm7p and some of its fungal orthologues. In addition, the authors detected many rare splicing variants of known introns that often disrupted the ORFs, including an alternative 3' splice site in *LSM7* and alternative 5' and 3' splice sites in *LSM2*. The biological significance of this is unclear. Mayes (1998) cloned intron-less *LSM7* by PCR using genomic DNA as a template and a forward primer covering all of the first exon and spanning the intron based on the annotated sequence at the time. This explains the discrepancy between the intron-less *LSM7* cloned by Mayes and the cDNA cloned in the Miura study. The *LSM7* sequence as described on SGD and in Mayes *et al.* (1999) was used throughout the work presented in this thesis. The consequences of the absence of 8 aa in the N-terminus of Lsm7p for this work is uncertain. These 8 residues were however shown to be non-essential for growth at 37°C, as intron-less *LSM7* cloned by Mayes (1998) and expressed from the *GAL1* promoter rescued temperature-sensitive growth of *lsm7Δ* (Mayes, 1998). This suggests that the outcome of experiments described here are unlikely to have been affected by the use of the incorrect *LSM7* sequence.

3.7.2 Conclusions

In this chapter I have shown that it is possible to assemble both yeast and human Lsm complexes *in vitro* from purified recombinant Lsm subunits expressed in *E. coli*. These complexes contain all seven Lsm polypeptides, but may aggregate, as was shown by size exclusion chromatography. The homogeneity of the resulting complexes is unclear, as it seems that not all seven subunits are required to form soluble complexes. In addition, Lsm proteins seem to be highly promiscuous in their interactions. They are able to form homo-multimeric complexes and seem to be able to interact with Lsm proteins other than the proposed nearest neighbours in the ring-

shaped complex. Lsm8p even interacts with the *E. coli* Sm-like protein, Hfq, leading to co-purification of this polypeptide in significant amounts during affinity chromatography. It is therefore important that Lsm proteins are purified from *hfq* deletion strains to prevent downstream problems due to contamination of Lsm complexes with this known RNA chaperone. The challenge in forming the correct complex *in vitro* lies in finding conditions that disrupt pre-existing undesirable Lsm-Lsm interactions without affecting potentially important structures within each of the subunits, and subsequently conditions in which correct Lsm-Lsm interactions are most likely to form. No easy failsafe methods exist to test for the formation of correctly assembled Lsm complexes. The homogeneity of complexes can be characterised by biochemical means such as ion-exchange and size-exclusion chromatography, although the propensity to complex aggregation may complicate interpretation of results. Due to the unknown level of homogeneity of these *in vitro* assembled complexes, additional purification steps may be desirable in any case. The RNA binding efficiency of human LSm2-8 increased with additional purification steps after the initial complex formation (Christian Kambach, personal communication). Another way of testing for proper complex formation is to see if they are functionally active. Lsm2-8p can be tested for its ability to bind U6 snRNA and may be tested for its ability to promote U4/U6 di-snRNP formation, which was shown to be an activity of the native human LSm2-8 complex (Achsel *et al.*, 1999). However, no RNA binding target or biological activity is known for Lsm1-7p. One of the aims of this study was to elucidate biological activities and underlying mechanism for the Lsm complexes. This can only be done with functionally active complexes, posing a catch-22 situation. Nonetheless, it seems that the methods described here can form at least partially functionally active Lsm complexes as will be discussed in Chapters 4 and 5.

CHAPTER 4

RNA binding and annealing activities of Sm-like Proteins

4.1 Introduction

The Lsm2-8 complex forms the core of the U6 snRNP and although the Lsm proteins interact with a wide range of other RNAs, U6 snRNA is the only known target with which they stably associate. The ability of *in vitro* assembled Lsm2-8p to bind U6 RNA is therefore a good test for its functional competence, at least in terms of specific RNA binding. No specific binding sequence for Lsm1-7p is known, but it is expected to be a single-stranded uridine-rich sequence similar to the proposed Lsm2-8p binding site in U6 snRNA (Achsel *et al.*, 1999; Vidal *et al.*, 1999), as these complexes have six out of the seven subunits in common. In addition, a stem-loop structure adjacent to the single-stranded sequence may be required for efficient binding (Vidal *et al.*, 1999). *S. cerevisiae* Lsm2-8p has been implicated in U4/U6 di-snRNP formation *in vivo* (Verdone *et al.*, 2004) and purified native human LSm2-8 has been shown to promote U4/U6 di-snRNP formation *in vitro* (Achsel *et al.*, 1999). It is as yet unclear whether Lsm2-8p directly promotes U4/U6 annealing, like Prp24p has been shown to do *in vitro* (Ghetti *et al.*, 1995; Raghunathan *et al.*, 1998a; Rader *et al.*, 2002), or whether it acts simply as a co-factor to Prp24p. The *E. coli* Sm-like protein Hfq has been shown to function both *in vivo* and *in vitro* in promoting basepairing between non-coding small RNAs and their target mRNAs (reviewed by Storz *et al.*, 2004; Valentin-Hansen *et al.*, 2004). This RNA chaperone activity may be conserved in some, if not all, Sm-like protein complexes. This chapter describes RNA binding and U4/U6 annealing activities of individual recombinant yeast Lsm proteins, Lsm1-7p and Lsm2-8p, and compares these to recombinant Hfq and human LSm complexes.

4.2 RNA binding by recombinant Sm-like proteins

4.2.1 Individual yeast Lsm proteins and U6 snRNA binding

During purification of some of the recombinant yeast Lsm proteins *E. coli* Hfq was co-purified (section 3.6.1 and Appendix A.1). Initial RNA binding and U4/U6 annealing experiments were therefore unknowingly performed with Lsm-Hfq mixtures. To put the later experiments that are discussed here in perspective, results

of the earlier experiments with protein mixtures are presented in Appendix A1. To examine RNA binding activities of individual Lsm proteins in the absence of Hfq, recombinant His-tagged yeast Lsm proteins 1 to 8 were expressed in MRE4 or MRE5 (*Δhfq*) cells and purified by Ni-NTA affinity chromatography. These purified proteins were used together with *in vitro* transcribed gel-purified U6 RNA in RNA bandshift assays (Fig. 4.1A). Proteins were added in excess over RNA, but none of the yeast Lsm proteins showed significant U6 RNA binding with a discreet protein-RNA complex as a result, like those observed for the Lsm8p (Hfq+) preparation or recombinant Hfq. However, some of the Lsm proteins did seem to interact with U6 RNA. The exception was His-Myc-Lsm1p which caused a discrete U6 RNA mobility shift that may be due to binding by an Lsm1p homomultimer. Lsm1 protein was added in large excess over RNA however (more than 100x), indicating low affinity. His-Lsm4p and -Lsm7p shifted U6 RNA into the wells, indicative of aggregation of proteins with RNA, and His-Lsm5p and His-HA-Lsm8p showed a small shift that may be caused by RNA binding of Lsm5p and 8p monomers. In contrast, no U6 mobility shift was observed for His-Lsm2p, -Lsm3p or -Lsm6p. A shift observed for Lsm8p (Hfq+) was due to binding by co-purified Hfq as no such shift was observed for Lsm8p (Hfq-) (discussed in detail in Appendix A1).

In summary, none of the recombinant yeast Lsm proteins seems to form discrete homomultimeric complexes with significant U6 RNA binding activity (only weak binding by a possible Lsm1p multimer). Lsm4p and Lsm7p bind U6 RNA but aggregate simultaneously, whereas small shifts in the presence of Lsm5p and Lsm8p suggests that these may interact with U6 RNA as monomers.

4.2.2 *In vitro* assembled yeast Lsm1-7p and Lsm2-8p bind U4 and U6 snRNA

Recombinant yeast Lsm proteins purified from *hfq* deletion strains were used to assemble Lsm1-7p and Lsm2-8p *in vitro* as described in Chapters 2 and 3. The ability of these complexes to bind U6 RNA was investigated in an RNA bandshift assay with *in vitro* transcribed U6 RNA. Both Lsm1-7p and Lsm2-8p bind U6 RNA (Fig. 4.1B and C). The observed RNA binding by these *in vitro* assembled

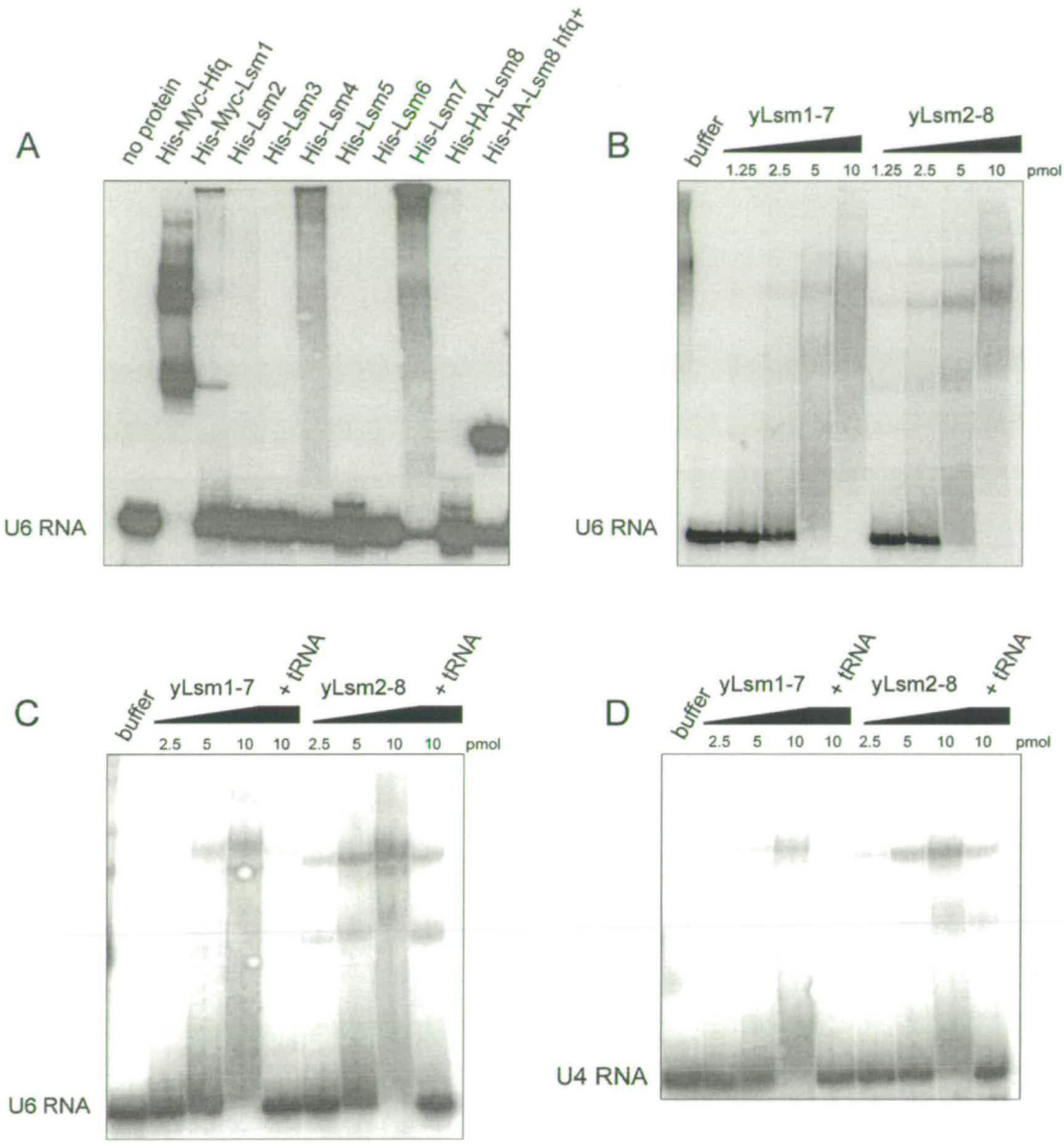


Fig. 4.1 RNA binding by individual recombinant yeast Lsm proteins and *in vitro* assembled yLsm1-7p and yLsm2-8p. **(A)** His-Myc-Hfq, His-Myc-Lsm1, His-HA-Lsm8 and His-Lsm2 to -Lsm7 were purified under native conditions from a *hfq* deletion strain, and His-HA-Lsm8 was also purified from BL21(DE3) (a *hfq*+ strain). Proteins were used at concentrations in excess over *in vitro* transcribed gel-purified U6 RNA: 7.5 ng (0.2 pmol) of U6 RNA, 10 pmol of Hfq, 23 pmol of Lsm1, 2 pmol Lsm2, 26 pmol Lsm3, 0.4 pmol Lsm4, 65 pmol Lsm5, 44 pmol Lsm6, 8 pmol Lsm7, 49 pmol of Hfq-free Lsm8 and 5 pmol of Lsm8 (monomers; Hfq+); **(B)** *In vitro* assembled Hfq-free yLsm1-7p and yLsm2-8p were used at a maximum of 10 pmol and 2-fold serial dilutions thereof; **(C)** As in (B) with 2 µg of tRNA added to the highest protein concentration; **(D)** As in (C), but with 0.2 pmol U4 RNA (10 ng) instead of U6 RNA.

complexes in contrast to the absence of significant specific RNA binding of the individual Lsm proteins suggests that at least some functional (heteroheptameric) complexes may have formed. Smeary shifts for Lsm1-7p in particular and failure to fully resolve into distinct bands suggests that these RNA-protein complexes may not be very stable under the bandshift conditions. The mobility of U6-Lsm1-7p and U6-Lsm2-8p bands compared to those observed with recombinant Hfq suggests that Lsm1-7p binds mainly with two complexes per U6 molecule. Lsm2-8p seems to bind with one, two or three complexes per U6 molecule, depending on the protein concentration and with a preference for two complexes. The difference in mobility between Lsm1-7p and Lsm2-8p-bound U6 RNA suggests that correct heptameric complexes may have formed as Lsm1-7p has a slightly higher molecular weight than Lsm2-8p leading to a less mobile U6-protein complex. Although the homogeneity of the *in vitro* formed complexes is unknown, the level of U6 binding by Lsm1-7p and Lsm2-8p suggests that Lsm2-8p has a slightly higher affinity for U6 RNA. Competition with a large excess of tRNA (270x) reduces U6 binding by both complexes significantly, but also seems to reduce smearing a lot (Fig. 4.1C). The level of protein bound U6 RNA in the presence of competitor RNA suggests that Lsm2-8p may also have a higher specificity for U6 RNA than Lsm1-7p. Using the same proteins in a U4 RNA bandshift assay (Fig. 4.1D) shows that both complexes bind similarly to U4 RNA, with a preference for two rings binding per RNA molecule and slightly higher affinity and specificity for Lsm2-8p than for Lsm1-7p. These conclusions are preliminary, as these experiments need to be repeated with complexes of known homogeneity. Binding to U6 as well as U4 RNA may indicate promiscuous RNA binding by both Lsm complexes, but may be biologically significant for Lsm2-8p if it could bind these RNAs simultaneously. Alternatively two complexes may interact with one binding U4 RNA (possibly the Sm heptamer) and the other binding U6 RNA thus favouring duplex formation by bringing the RNAs into close proximity. Favoured binding of two rings per RNA molecule and the tendency to aggregate may be significant in this respect.

4.2.3 Recombinant Hfq binds U4 and U6 snRNA

As *E. coli* Hfq was co-purified with a number of the yeast Lsm proteins, I wanted to investigate the ability of recombinant Hfq to bind U6 RNA to allow comparison with the Lsm complexes and interpretation of results of experiments with Lsm-Hfq protein mixtures (Appendix A1). His-Myc- and His-HA-tagged Hfq were expressed in *Δhfq* strains and purified under native and denaturing conditions respectively. The latter was followed by dialysis to remove all urea identical to the protocol used for *in vitro* yeast Lsm complex assembly. Denaturing purification of Hfq may produce more pure protein as more protein and RNA contaminants are likely to be removed. The Hfq hexamer has been shown to be extremely stable as many published purification methods of native Hfq include heating steps and/or 8 M urea steps which may only partially disrupt hexamer formation and may not affect the structure of the Hfq monomers. Indeed, recombinant Hfq was shown to still shift some RNA when separated by denaturing PAGE after heating in formamide loading buffer (data not shown). Recombinant Hfq purified under native and denaturing conditions displayed similar levels of U6 RNA binding in bandshift assays (Fig. 4.2A and B), suggesting that both methods result in equally functionally competent Hfq complexes. There was a slight difference in levels of aggregation (compare Fig. 4.2A and B), but this varied between experiments as well as between protein batches (data not shown). Increased co-purification of RNA with natively purified Hfq may affect its observed RNA binding and aggregation *in vitro*. RNA binding seemed specific for U6 as a large excess of tRNA (270x) did not strongly affect Hfq-U6 binding. The specificity may be for a single-stranded U-rich sequence adjacent to a stem-loop structure, as recombinant Hfq displayed similar binding to U4 RNA in a bandshift assay (Fig. 4.2C), although specificity for U4 RNA binding appeared to be slightly lower than for U6 RNA (see lane with added competitor tRNA in Fig. 4.2C).

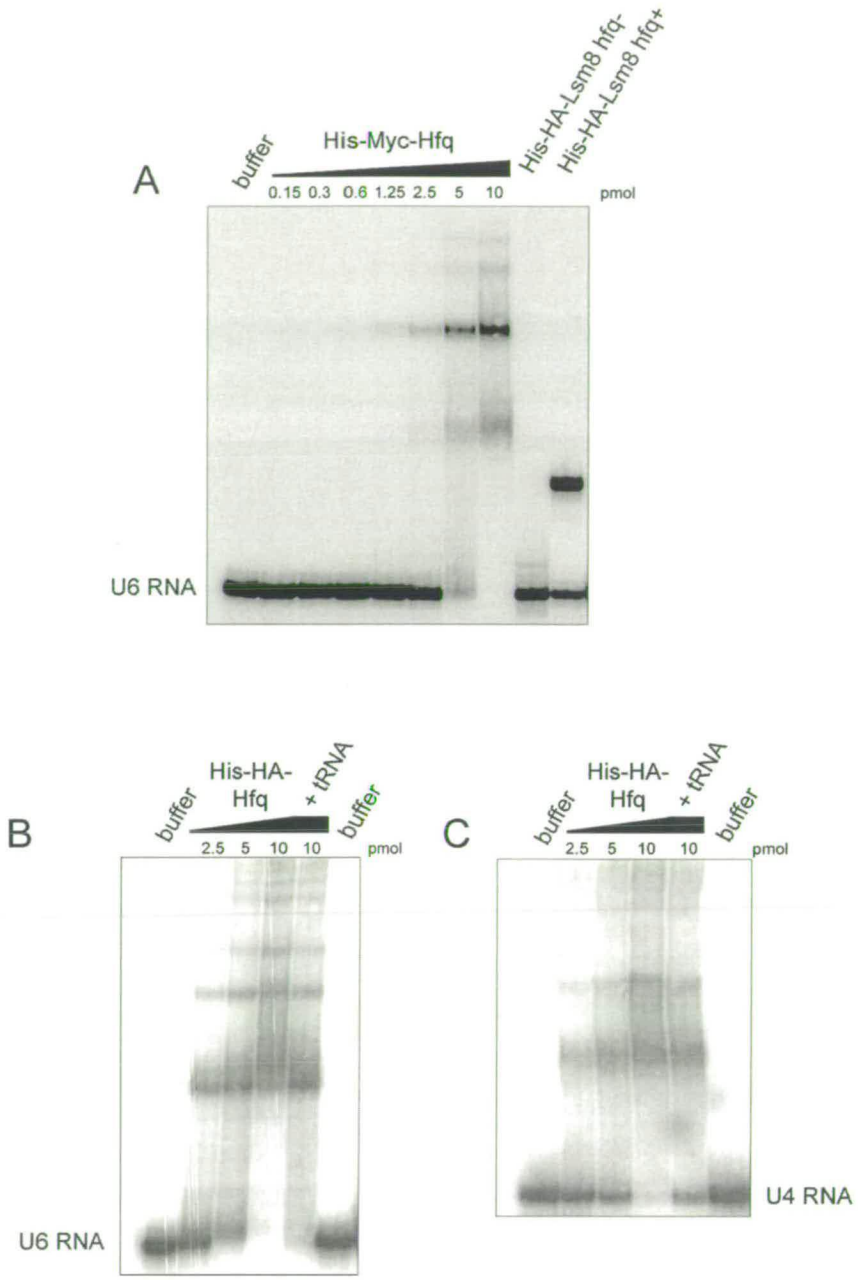


Fig. 4.2 RNA binding by purified recombinant Hfq. **(A)** His-Myc-Hfq was purified under native conditions from a *hfq* deletion strain. 7.5 ng (0.2 pmol) *of in vitro* transcribed gel-purified U6 RNA was incubated with a maximum of 10 pmol of Hfq (monomers) and 2-fold serial dilutions thereof. His-HA-Lsm8 was used as in Fig. 4.1A; **(B)** Dialysed His-HA-Hfq purified under denaturing conditions was used at a maximum of 10 pmol of Hfq and 2-fold dilutions thereof with 2 µg of tRNA added to the highest Hfq concentration; **(C)** As in (B), but with 0.2 pmol U4 RNA (10 ng) instead of U6 RNA.

4.2.4 RNA binding by *in vitro* assembled human LSm complexes

Human LSm1-7 and LSm2-8 were assembled *in vitro* from purified recombinant sub-complexes as described in Chapters 2 and 3. The ability of these complexes to bind U6 RNA was investigated in an RNA bandshift assay with *in vitro* transcribed U6 RNA. Both hLSm1-7p and hLSm2-8p were shown to bind U6 RNA with similar affinities (Fig. 4.3A). Just like the yeast Lsm complexes and Hfq, the human LSm complexes seem to aggregate at higher protein concentrations, but not as severely and at these concentrations only one or two complexes were seen to bind one RNA molecule at one time. A slight difference in mobility between U6-hLSm1-7 and U6-hLSm2-8 suggests that the correct heptameric complexes may have formed as hLSm1-7 has a higher molecular weight than hLSm2-8. I was surprised to find that both hLSm1-7 and hLSm2-8 bind U6 RNA, as Zaric *et al.* (2005) had found that hLSm2-8, but not hLSm1-7 binds U6 snRNA. However, Zaric *et al.* used *Xenopus tropicalis* U6 snRNA for the bandshift assay, whereas the results presented here are for *S. cerevisiae* U6 snRNA. *Xenopus tropicalis* and *Homo sapiens* U6 snRNA only differ by one nucleotide in the 5' stem-loop, but the difference with yeast U6 is much more considerable. Although the secondary structures of these RNAs are very similar and both have uridine-rich 3' ends, differences in their primary sequence could significantly contribute to a difference in RNA binding by the LSm complexes. To test if I could distinguish between RNA binding specificity by the hLSm complexes, I used various competitor RNAs in the U6 RNA bandshift assay. However, up to four times excess of U6(77-98), U6(77-112) and U2Δ107 RNA did not significantly affect U6 RNA binding by either of the human LSm complexes (data not shown). A large excess of competitor RNA was needed to abolish U6 RNA binding (Fig. 4.3B). To lose half of U6 binding an excess of approximately $2.7 \cdot 10^3$ times (w/w) of tRNA was needed, with all binding abolished at $2.7 \cdot 10^4$ times excess. In contrast most U6 RNA binding by hLSm2-8 and a significant amount of binding by hLSm1-7 are abolished by addition of 100 times excess (w/w) of U6(77-112) RNA. This suggests that U6 RNA binding by both hLSm1-7 and hLSm2-8 is highly specific, with the single-stranded uridine-rich U6(77-112) RNA as a moderately effective competitor. To determine what part of U6 RNA is important for binding by the hLSm complexes U6(1-90) and U6(77-112) RNAs were used in bandshift assays (Fig. 4.4A-D). The

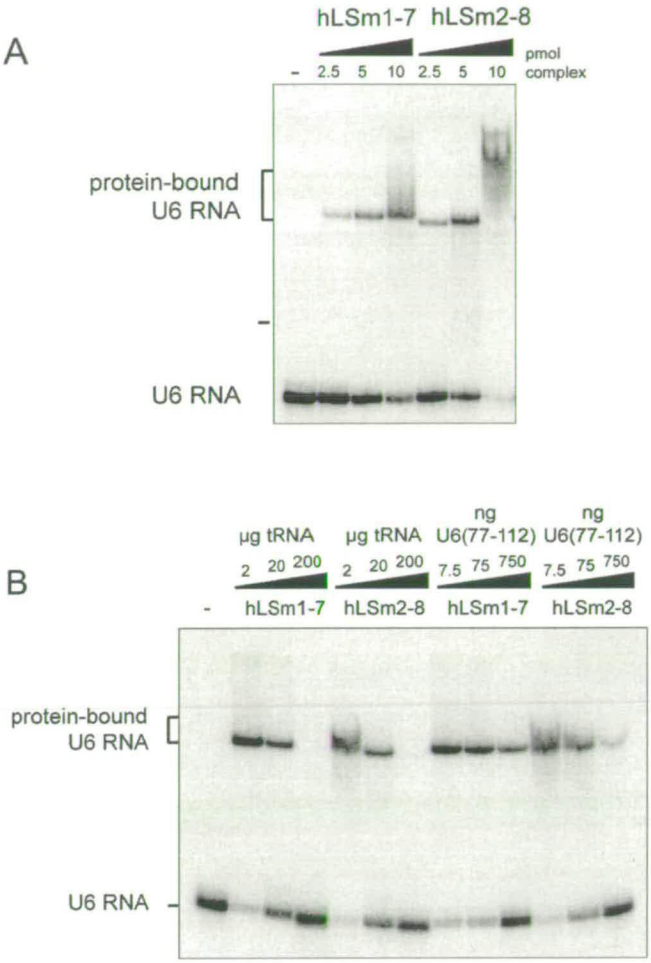


Fig. 4.3 RNA binding by *in vitro* assembled hLSm1-7 and hLSm2-8. **(A)** Human LSm1-7 and 2-8 complexes were assembled from purified recombinant subcomplexes and incubated at a maximum of 10 pmol and 2-fold serial dilutions thereof with 0.2 pmol of U6 RNA; **(B)** 5 pmol of hLSm1-7 and hLSm2-8 was incubated with 0.2 pmol (7.5 ng) of U6 RNA and 2, 20 or 200 µg of tRNA, or 7.5, 75 or 750 ng of U6(77-112) competitor RNA.

affinity of both complexes was higher for U6(77-112) than for the U6Δ3' RNA, which lacks the U-rich single stranded tail. The affinity of hLsm2-8 was slightly higher than that of hLSm1-7, particularly for U6(77-112) RNA. This confirms the data with this RNA as a more effective competitor for hLSm2-8p than for hLsm1-7 in the U6 bandshift assay (Fig. 4.3B). Binding to both of these U6 RNA fragments is specific as a large excess of tRNA was needed to abolish the majority of binding (Fig. 4.4B and D). All binding by hLSm1-7 disappears with $8.3 \cdot 10^3$ (w/w) for U6Δ3' RNA and $2 \cdot 10^4$ (w/w) for U6(77-112) RNA excess of tRNA, whereas hLSm2-8 still binds some of the U6 RNA fragments at these tRNA concentrations.

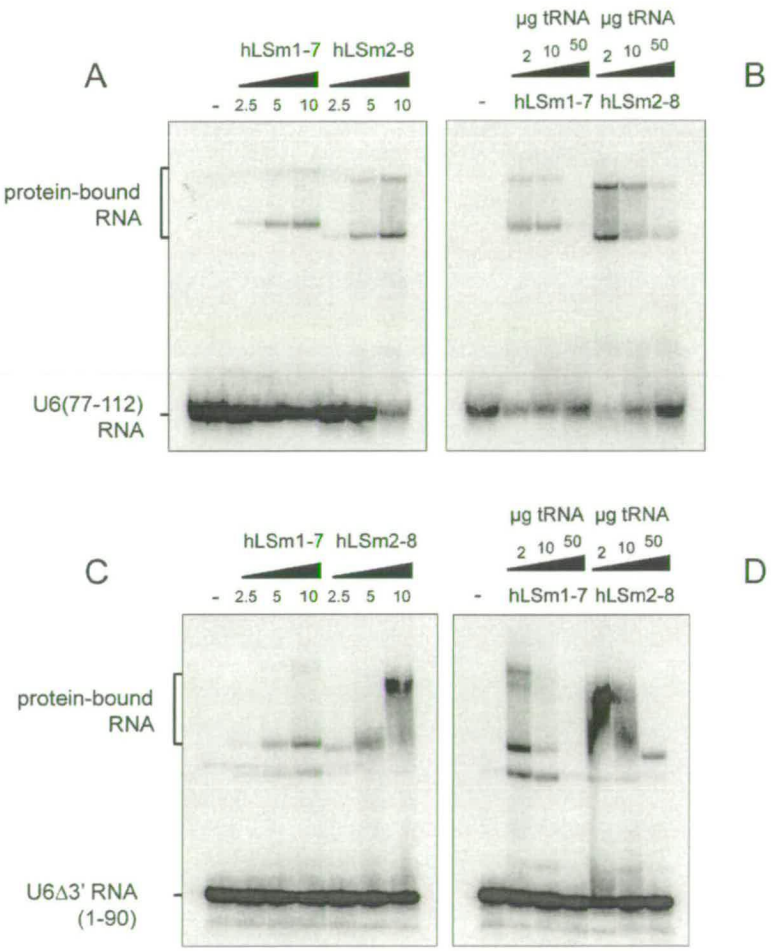


Fig. 4.4 RNA binding requirements of *in vitro* assembled hLSm1-7 and hLSm2-8. **(A)** 10 pmol of hLSm1-7 and hLSm2-8 and 2-fold dilutions thereof were incubated with 0.2 pmol (2.5 ng) of U6(77-112) RNA, or **(B)** 20 pmol of complexes was incubated with 0.2 pmol of RNA and 2, 10 or 50 µg of competitor tRNA; **(C)** As in (A) but with 0.2 pmol (6 ng) of U6Δ3' RNA; **(D)** As in (B) but with U6Δ3' RNA.

The results of these experiments are summarised in Fig. 4.5. The presented data suggest that both hLSm1-7p and hLSm2-8p bind yeast U6 snRNA with similar affinities and high specificity (possibly slightly higher for hLSm2-8p than for hLSm1-7p). Both complexes show preferred binding to full length U6 RNA over U6 fragments lacking either the 3' uridine-rich single stranded tail or the structured 5' half. This strongly suggests that, like yeast Lsm2-8p (Vidal *et al.*, 1999), both hLSm1-7p and hLSm2-8p prefer to bind U-rich single stranded RNA with an adjacent stem-loop structure.

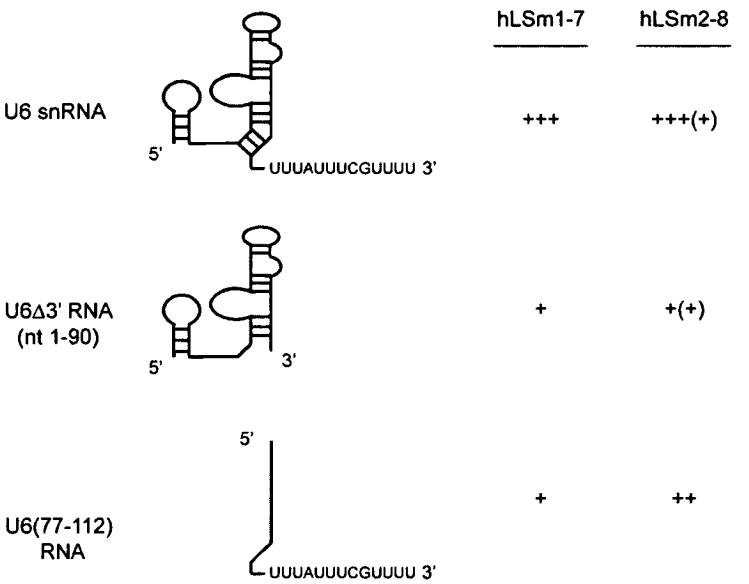


Fig. 4.5 Overview of hLSm1-7 and hLSm2-8 binding to yeast U6 snRNA or fragments thereof

4.3 RNA annealing activity of recombinant Sm-like proteins

The *in vitro* U4/U6 annealing activities of yeast and human Lsm complexes as well *E. coli* Hfq were tested with purified recombinant proteins and *in vitro* transcribed gel-purified U4 and U6 RNA. For these experiments it was important not to use the RNAs at too high a concentration as this was shown to significantly increase the amount of spontaneous U4/U6 duplex formation in the absence of any protein. In addition the RNAs were denatured at 85°C and snap chilled on ice just prior to use.

Failure to do so lead to increased amounts of U6 bands of slow mobility, suggestive of U6 dimer formation. After incubation RNA was deproteinised by Proteinase K digestion and/or phenol-chloroform extraction to allow determination of the level of U4/U6 duplex by native PAGE and Northern analysis.

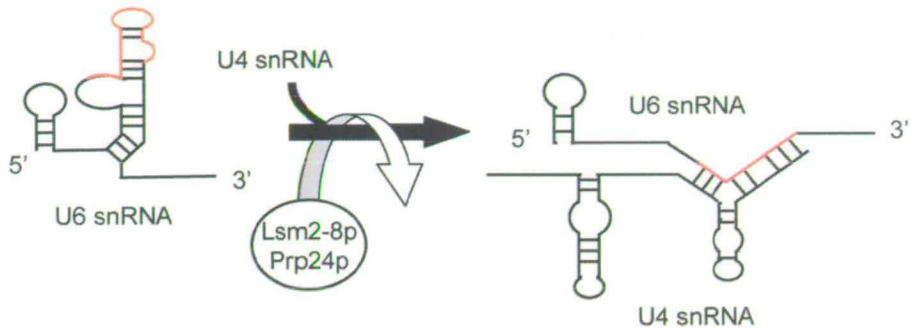


Fig. 4.6 Schematic representation of U4/U6 annealing. The proposed secondary structure of U6 snRNA on the left contains a 3' stem-loop structure part of which (indicated in red) is directly involved in base-pairing interactions with U4 snRNA in the U4/U6 duplex on the right.

4.3.1 Recombinant yeast Lsm complexes do not display U4/U6 annealing activity

The *in vitro* assembled yeast Lsm1-7 and Lsm2-8 complexes failed to promote annealing between U4 and U6 RNA *in vitro* (data not shown), in spite of their U6 (and U4) RNA binding ability (Fig. 4.1). Their failure to do so may mean that yeast Lsm2-8p itself does not have U4/U6 annealing activity, at least *in vitro*. However, it could also indicate that the *in vitro* assembled complexes are not fully functional. Annealing experiments with human LSm complexes assembled under different conditions suggest that denaturation with 8 M urea may affect a structure within the complex that is essential for U4/U6 annealing activity (4.3.3.1). Attempts to assemble yeast Lsm complexes using individual, natively purified Lsm proteins and the method used for the human LSm complexes (with 4 M urea) failed to produce complexes that bind U6 RNA (data not shown), so this hypothesis could not be tested. As discussed in Chapter 3, a method analogous to that used for the human LSm proteins (i.e. co-expression in pairs and assembly under semi-denaturing

conditions) or co-expression of all seven proteins using a heptacistronic construct may be more likely to allow formation of recombinant complexes that are fully functional.

4.3.2 Recombinant Hfq promotes U4/U6 annealing

Hfq has been shown to function as an RNA chaperone both *in vivo* and *in vitro* (reviewed by Storz *et al.*, 2004; Valentin-Hansen *et al.*, 2004; Arluison *et al.*, 2007). This bacterial Sm-like protein promotes basepairing interactions between non-coding small RNAs and their target mRNAs. As recombinant yeast Lsm proteins that co-purified Hfq were shown to be able to promote U4/U6 annealing (Appendix A1) and Hfq-free Lsm complexes did not display this activity (4.3.1) it seems that this Hfq activity is not restricted to its biological targets. To confirm that Hfq can indeed promote U4/U6 duplex formation and to further investigate potential mechanisms for this, recombinant Hfq was purified and tested in U4/U6 annealing assays

4.3.2.1 Hfq RNA binding is essential for its annealing activity

Recombinant purified Hfq, either natively purified or purified under denaturing conditions, displayed U4/U6 annealing activity *in vitro* (Fig. 4.7 and 4.8). To test if Hfq needs to be able to bind RNA to promote U4/U6 annealing Hfq-Y55A mutant protein, which has been shown to have much reduced RNA binding (Mikulecky *et al.*, 2004) was purified. In addition, a C-terminal deletion mutant, HfqΔC (aa 1-83), was purified to test if the C-terminal extension of Hfq plays a role in the RNA-RNA annealing process by participating in RNA-binding as was suggested by Sauter *et al.* (2003). The Y55A mutation was indeed shown to abolish binding to U6 RNA at concentrations at which wild-type Hfq shifts all U6 RNA. In contrast, the C-terminal deletion had only a small effect on U6 RNA binding (Fig. 4.7A). These same proteins were used in a U4/U6 annealing assay (Fig. 4.7B). HfqΔC was shown to have annealing activity similar to that of wild-type Hfq (only slightly reduced again), whereas Hfq-Y55A did not display any annealing activity above background levels. These data indicate that Hfq needs to be able to bind stably to RNA to promote RNA-RNA duplex formation and that the C-terminal extension does not play a

significant role in either RNA binding or RNA-RNA annealing, at least under the conditions tested here.

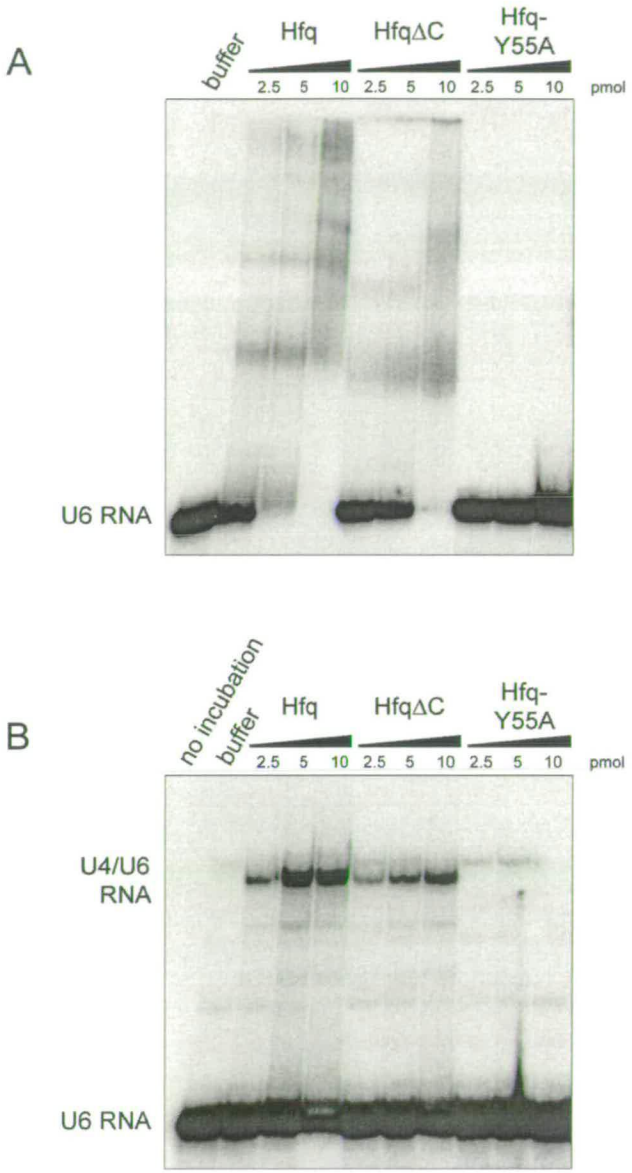


Fig. 4.7 U6 RNA binding and U4/U6 annealing by recombinant Hfq and Hfq mutants. **(A)** Recombinant His-HA-Hfq, His-Myc-HfqΔC and Hfq-Y55A-His were incubated at 10 pmol (of monomers) or 2-fold dilutions thereof with 0.2 pmol of U6 RNA; **(B)** The same proteins and amounts as in (A) were incubated with 0.2 pmol of U6 RNA and 0.3 pmol of U4 RNA and treated with Proteinase K before separation by native PAGE.

4.3.2.2 Hfq affects the conformation of the U4/U6 duplex

Interestingly, U4/U6 duplex formed in the presence of Hfq displayed a different (faster) mobility during native PAGE compared to the duplex formed in the absence of protein or with recombinant Prp24p (Fig. 4.8 and A1.3). Denaturing PAGE and Northern analysis of U4 and U6 RNA after annealing with or without Hfq showed that the faster migrating duplex is not caused by partial degradation of the RNA (data not shown). U4/U6 duplex formed by recombinant Prp24p, which migrates like spontaneously formed duplex, was converted to the more mobile duplex when Hfq (co-purified with Lsm8p) was added (Fig. 4.8 and A1.3C). U4 and U6 RNA were present in both duplexes (Fig. 4.8B). These data suggest that Hfq changes the structure of the U4/U6 duplex to a more compact one. Phenol/chloroform extraction of the RNA-protein mixture before separating the RNA by native PAGE caused the fast-migrating duplex to be converted back into the slow-migrating duplex (Fig. 4.8A). This is not affected by the pH (4.7 or 8.0) of the phenol/choloroform mixture (data not shown). Separation of the two different duplexes on a 1x TBM instead of a 1x TBE gel (1 mM $MgCl_2$ instead of EDTA) made them migrate at the same rate (data not shown), suggesting that the $MgCl_2$ either affects the Hfq-fragment RNA interaction, or that Mg^{2+} favours the structure of the less mobile duplex. The melting temperature of the two duplexes is the same as both start melting at 56°C (Fig. 4.8C), in agreement with the melting temperature of the native U4/U6 duplex of approximately 55°C (Brow *et al.*, 1988; Raghunathan *et al.*, 1998a). Interestingly, at 52 and 54°C some of the "fast" duplex changes to the "slow" duplex suggesting that the protein-duplex interaction may be destabilised at these temperatures before the U4/U6 duplex itself is destabilised. In summary, these data strongly suggests that there is an Hfq fragment that is resistant to degradation by Proteinase K and that is stably bound to RNA leading to a U4/U6 duplex with a more compact structure, but the same duplex stability, that migrates faster in spite of its presumably larger molecular weight. This alternative U4/U6 structure is not maintained in the absence of the Hfq fragment.

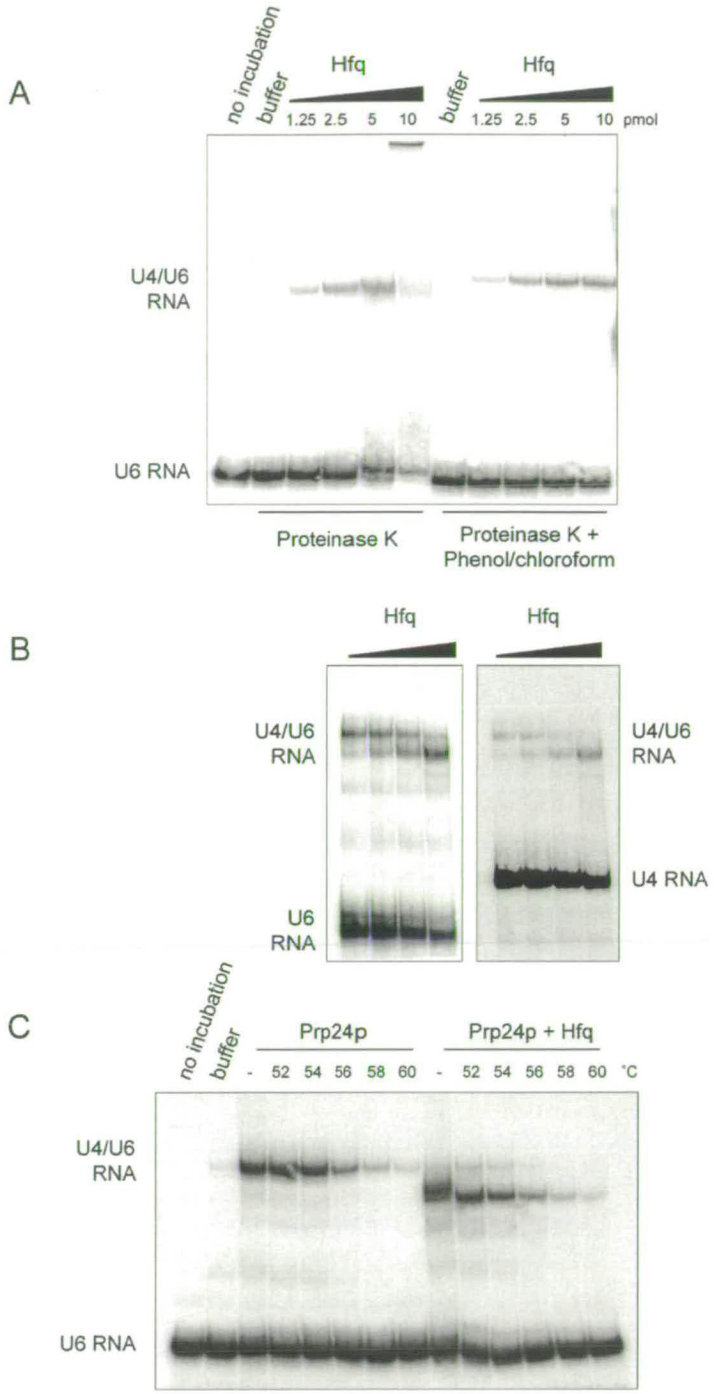


Fig. 4.8 Hfq affects the structure of the U4/U6 duplex **(A)** Recombinant His-Myc-Hfq was incubated at 10 pmol (of monomers) or 2-fold dilutions thereof with 0.2 pmol of U6 RNA and 0.27 pmol of U4 RNA. Samples were treated with Proteinase K and/or phenol-chloroform extracted before separation by native PAGE; **(B)** 1.3 pmol of U6 and 1.8 pmol of U4 RNA were annealed in the presence of increasing amounts of Hfq and Proteinase K treated before native PAGE; **(C)** U6 and U4 RNA were annealed with 7.5 pmol of Prp24p and/or Hfq, Proteinase K treated and heated for 5 min at various temperatures and snap chillled before separation by native PAGE.

4.3.3 Recombinant human LSm complexes promote annealing of yeast U4 and U6 snRNA

4.3.3.1 hLSm1-7 and hLSm2-8 annealing activity is affected by the method of complex assembly

Both the human LSm1-7 and LSm2-8 complexes show U4/U6 annealing activity *in vitro* (Fig. 4.9B), when complexes are assembled using the method described by Zaric *et al.* (2005) or a slight adaptation thereof (2.4.10.2). However, when the human LSm complexes are assembled *in vitro* using the method described in 2.4.10.1, with a denaturing step at 8 M urea followed by one-step dialysis to 0 M urea, almost all U4/U6 annealing activity disappears (Fig. 4.9B). In contrast, complexes assembled with each of these methods show almost identical binding to U6 RNA (Fig. 4.9A), suggesting that both methods produce similar complexes that are competent in terms of RNA binding. Both hLSm1-7 and hLSm2-8 seem to have similar U4/U6 annealing as well as U6 RNA binding activity. These data suggest that, as for Hfq, the level of RNA binding correlates with the level of U4/U6 annealing, but that there may be an additional requirement for RNA-RNA annealing activity, at least in case of the human LSm complexes.

4.3.3.2 hLSm4 may be essential for annealing activity

To confirm the necessity of RNA binding for U4/U6 annealing activity of the LSm complexes and to determine what the additional requirement for this activity may be a number of controls were prepared. As the LSm complexes are heteroheptameric it is difficult to know whether point mutations in putative RNA-binding residues of single subunits would affect RNA binding of the complexes *in vitro*, as is the case for Hfq-Y55A. An alternative approach was therefore taken in which single subunits were left out of the mixture used for *in vitro* complex assembly. Soluble complexes can be formed this way (see Fig. 3.10). These are presumably complexes containing less than seven subunits per ring (5- or 6-mers) or with certain subunits occurring twice, with alternative heteroheptameric rings as a result. hLSm2-7 and hLSm23567 were produced this way and tested for U6 RNA binding and U4/U6 annealing (Fig. 4.10A and B). Both complexes have significantly reduced affinities for U6 RNA (>2x for hLSm2-7 and >4x for hLSm23567), but the nature of the complexes is

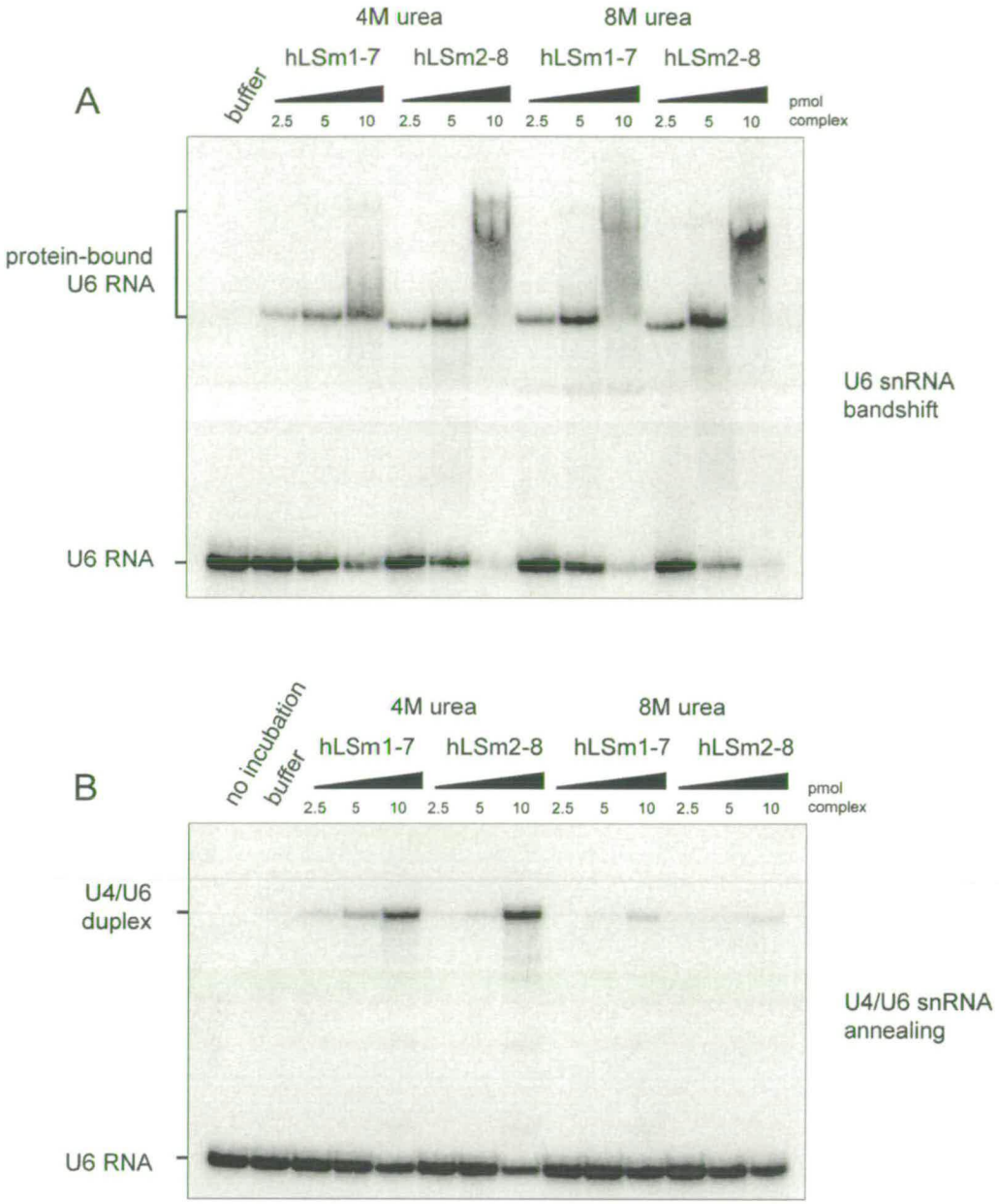


Fig. 4.9 U4/U6 annealing activity of *in vitro* assembled human LSm complexes is affected by 8 M urea. **(A)** Both hLSm1-7 and hLSm2-8 assembled either using 4 M urea or 8 M urea bind U6 RNA with equal affinity; **(B)** hLSm1-7 and hLSm2-8 assembled using 4 M urea have similar U4/U6 annealing activity, whereas the complexes assembled with 8 M urea only display less than one quarter of this activity. U6 RNA was used at 0.2 pmol for both assays, U4 RNA at 0.3 pmol in the annealing assay and proteins at 10 pmol or 2-fold serial dilutions thereof.

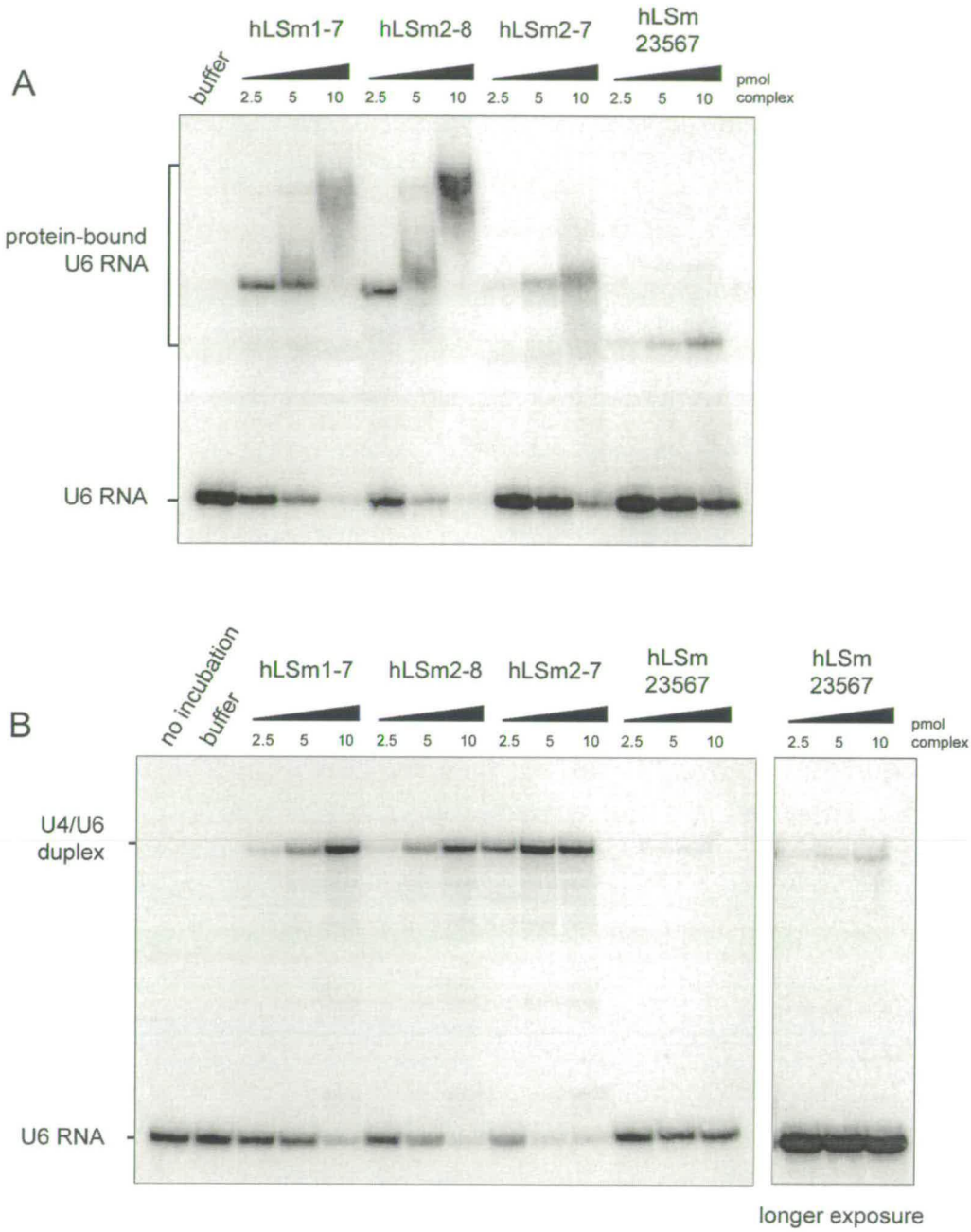


Fig. 4.10 U6 binding and U4/U6 annealing by *in vitro* assembled human LSm complexes with less than seven different subunits. hLSm1-7, 2-8, 2-7 and 23567 were assembled *in vitro* using the 4 M urea protocol. Proteins were used at 10 pmol (assuming 7-, 6- and 5-mers respectively) and 2-fold serial dilutions thereof **(A)** RNA bandshift assay with 0.2 pmol of U6 RNA and the described complexes; **(B)** U4/U6 annealing assay with the same complexes.

uncertain as U6-hLSm2-7 migrates similar to hLSm1-7 and hLSm2-8, whereas U6-hLSm23567 migrates significantly faster and could therefore represent an actual pentameric complex. In contrast the U4/U6 annealing activity of hLSm2-7 is about twice as high as that of hLSm1-7 and hLSm2-8 and the activity of hLSm23567 is relatively very low (>8x lower). This indicates that there is no direct correlation between RNA-binding affinity and U4/U6 annealing activity of these complexes.

As these complexes were a "hexamer" with hLSm4 (active) and a "pentamer" without hLSm4 (inactive) it may be either that a hexameric complex is the minimal requirement for efficient U4/U6 annealing or that hLSm4 is an essential component of the LSm complexes in terms of annealing. To test these hypotheses a "pentameric" complex with hLSm4 (hLSm4-8) and a "hexameric" complex without hLSm4 (hLSm123567) were assembled and tested for U6 RNA binding and U4/U6 annealing (Fig. 4.11A and B). The exact nature of the complexes that were formed was again uncertain as mobility of U6-protein complexes may not fit with formation of pentameric and heptameric complexes respectively. Nonetheless both displayed U6 RNA binding with reduced affinity compared to hLSm2-8 (at least 8 times lower). In contrast, hLSm123567 displayed very little U4/U6 annealing, whereas hLSm4-8 displayed annealing activity that was slightly higher than that of hLSm2-8. This suggests that hLSm4 may be the essential subunit of hLSm complexes in terms of U4/U6 annealing. To test if hLSm4 and the hLSm48 subcomplex are able to promote annealing both were dialysed to the same buffer without urea. Soluble proteins were tested for U6 RNA binding and U4/U6 annealing. Surprisingly, much of the free U6 RNA disappeared in the U6 bandshift assay without showing a shift (Fig. 4.12A). This is unlikely to be due to RNA degradation as no significant degradation of U6 RNA was observed during the U4/U6 annealing assay (Fig. 4.12B) even though proteins and RNA are incubated much longer in the latter assay. As U4/U6 annealing reactions are Proteinase K treated before native PAGE and U6 bandshift reactions are loaded directly onto the gel, hLSm4 and hLSm48 may aggregate while binding RNA and thus prevent the U6-protein complexes from entering the gel. This could be equivalent to what was observed for yeast Lsm4p (Fig. 4.1A) but the U6-hLSm4(8) aggregates may not have transferred to the

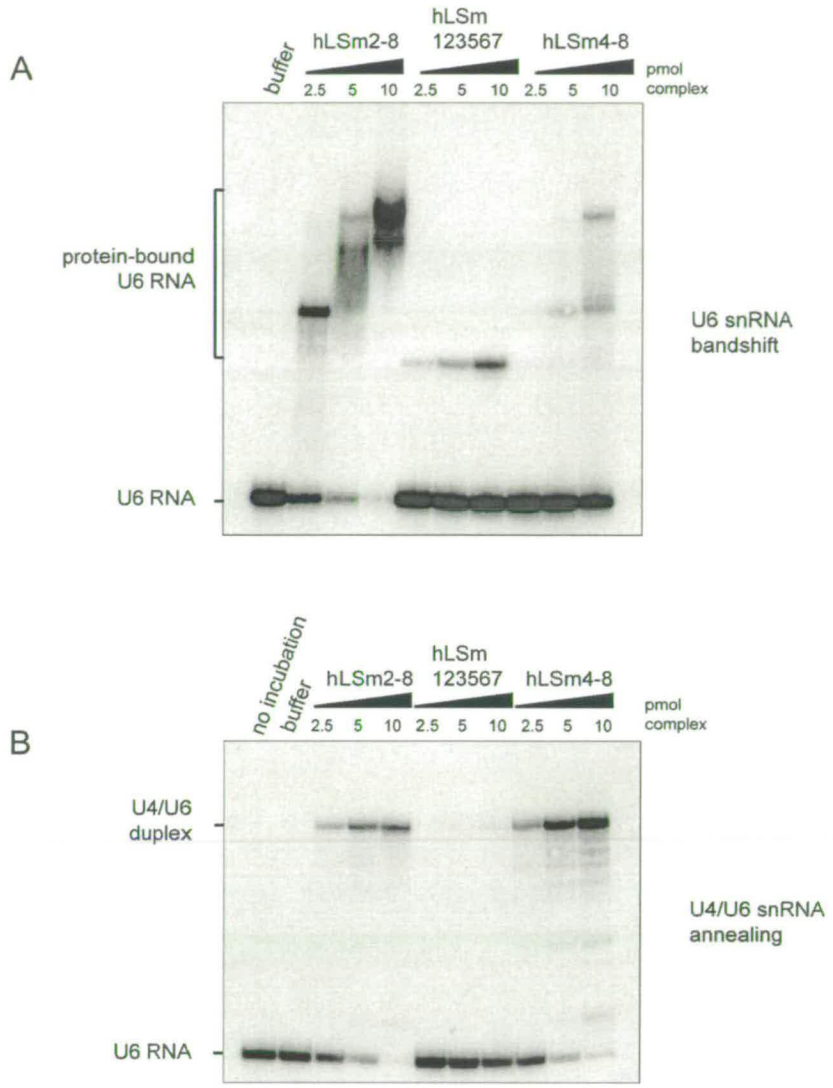


Fig. 4.11 hLSm4 is essential for annealing activity of human LSm complexes. hLSm2-8, 123567 and 4-8 were assembled *in vitro* using the 4 M urea protocol. Proteins were used at 10 pmol (assuming 7-, 6- and 5-mers respectively) or 2-fold dilutions thereof, with 0.2 pmol of U6 RNA and/or 0.3 pmol U4 RNA. **(A)** U6 RNA bandshift assay with the hLSm2-8, 123567 and 4-8 complexes; **(B)** U4/U6 annealing assay with the complexes from (A).

membrane during Northern transfer. U4 and U6 RNA are annealed efficiently by both hLSm4 and hLSm48. As the concentration of hLSm4 (subunits) is the same in these reactions as in the reaction performed with other hLSm4 containing complexes it seems that the annealing activity depends on the number of hLSm4 monomers present. As the annealing activity of hLSm48 is approximately half of that of hLSm4, the presence of hLSm8 may inhibit some of the annealing activity. Similarly, hLSm1-7 and hLSm2-8 both display less annealing activity than hLSm2-7, which may indicate that hLSm1 also inhibits the LSm4-dependent annealing activity. Significantly recombinant yeast Lsm4 also displayed U4/U6 annealing activity (Fig. 4.12C).

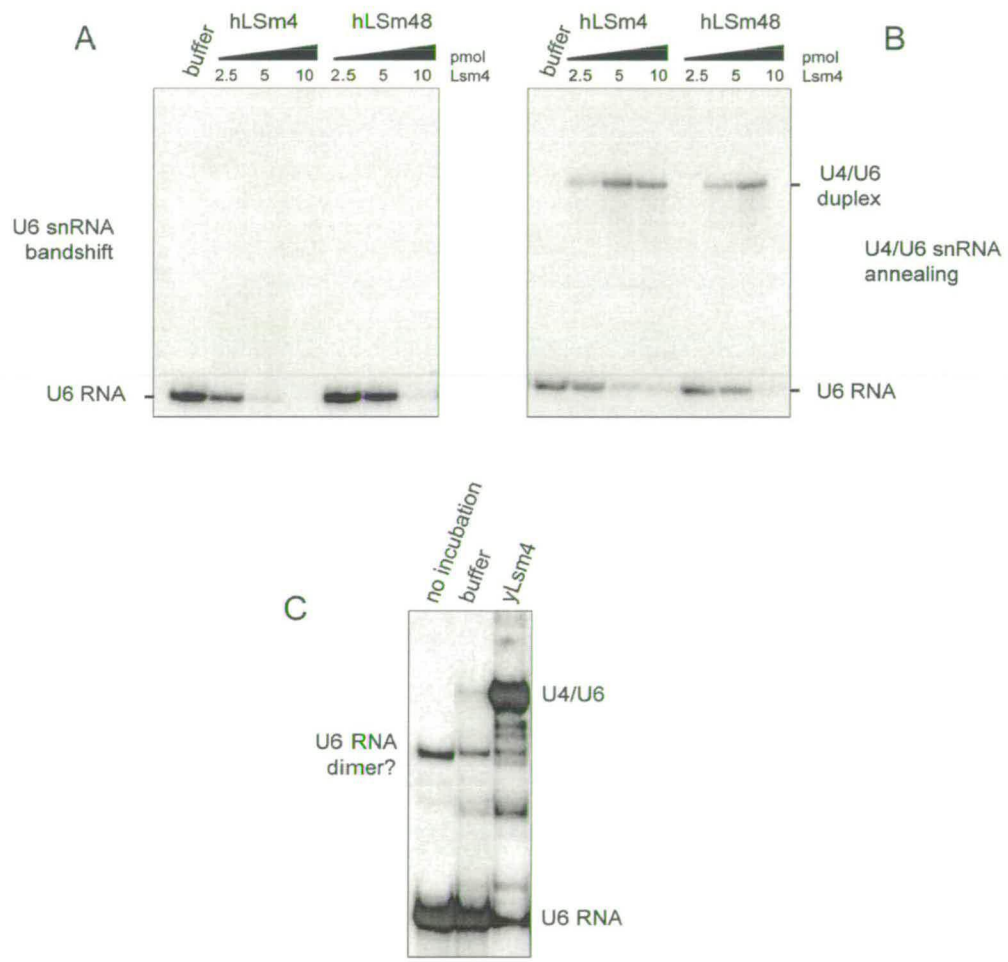


Fig. 4.12 Lsm4 is sufficient for annealing activity.
(A) hLSm4 and 48 were dialysed into 0 M urea buffer. Proteins were used at 10 pmol of hLSm4 monomers or 2-fold dilutions thereof, with 0.2 pmol of U6 RNA for bandshift assays, **(B)** and 0.3 pmol U4 RNA for U4/U6 annealing assays; **(C)** U4/U6 annealing assay with 0.4 pmol of hfq-free recombinant yeast Lsm4p (as in Fig. 4.1A).

4.4 Discussion

In this chapter I have shown that recombinant yeast and human LSm1-7 and 2-8 complexes assembled *in vitro* as well as affinity purified recombinant *E. coli* Hfq are capable of binding U6 and U4 snRNA. The affinity and specificity for binding to U6 RNA may be slightly higher for yLsm2-8p than for yLsm1-7p. The human LSm complexes show a similar but subtle difference in RNA binding specificity, but seem to prefer a uridine-rich single stranded sequence with an adjacent stem-loop structure as was previously proposed for yLsm2-8p (Vidal *et al.*, 1999). The affinity of Hfq for both U4 and U6 RNA seems to be considerably higher, but this difference may reflect the low efficiency of correct Lsm complex assembly *in vitro* more than an actual difference in affinity for RNA. In addition, it shows the promiscuity of RNA binding by Hfq. Hfq has a preference for A/U rich sequences and a nearby stem-loop structure, which may be a feature that is conserved between all Sm-like complexes. *In vitro* assembled yeast Lsm complexes failed to show any annealing activity. As these complexes were assembled from individual Lsm proteins in the presence of 8M urea and the human LSm complexes had much reduced activity when assembled under these same conditions, this may be due to the failure to produce fully functional complexes this way. However, it can not be ruled out that the yeast Lsm complexes do not have U4/U6 annealing activity by themselves, but assist Prp24p (p110 in humans) in this process (see also Chapter 6).

In contrast, hLSm1-7 and hLSm2-8 showed high levels of U4/U6 annealing. Presence of recombinant hLSm4 was shown to be essential and sufficient for this activity. Although it is appealing to conclude that the hLSm4 polypeptide itself determines the annealing activity of the Lsm complexes, co-purification of an *E. coli* protein that has this activity with hLSm4 can not be ruled out at this stage. In that case, reduced annealing for hLSm1-7 and hLSm2-8 compared to hLSm4 alone may indicate inhibition of activity of this factor because of high affinity hLSm-U6 binding. This potential hLSm4-specific contaminant is unlikely to be Hfq for two reasons: firstly, no Hfq-sized band is observed on (Coomassie stained) SDS-PAGE gels; secondly, a denaturation step with 8 M urea decreases the hLSm annealing

activity, whereas increased activity may be expected for Hfq (see Appendix A1 for results with Hfq-contaminated yLsm complexes). If it is hLSm4 that provides annealing activity it poses a number of questions. How does it promote annealing? What happens to hLSm4 in the presence of 8 M urea that leads to loss of most of this activity? In this respect it is interesting to note that yeast Lsm4p was shown to aggregate, both *in vivo* (see Chapter 7) and *in vitro* (Fig. 4.1A), and was shown to promote U4/U6 annealing *in vitro* (Fig. 4.12C). Is it possible that hLSm4 does the same and that its aggregation while bound to RNA promotes U4/U6 annealing? Annealing activity of yeast Lsm4p again emphasises the possibility that the yeast Lsm complexes would also have U4/U6 annealing activity if assembled under the right conditions. Both hLSm1-7 and hLSm2-8 display similar levels of U6 RNA binding as well as U4/U6 annealing activity. This suggests that there may only be subtle differences between the RNA-binding specificity of these two complexes and that a big part of their functional difference *in vivo* may come from their specific subcellular localisations (see also Chapter 9).

As both Hfq and the human LSm complexes were shown to be able to promote annealing between U4 and U6 RNAs, it suggests that RNA chaperone activity may be conserved between Sm-like complexes across the different kingdoms of life. It is as yet uncertain however, what the mechanism is by which Sm-like complexes promote RNA-RNA annealing. Experiments that attempt to elucidate this intriguing question will be discussed in Chapter 5.

CHAPTER 5

Investigating a possible mechanism for RNA annealing by Sm-like proteins

5.1 Introduction

Sm-like proteins have been shown to promote RNA-RNA interactions (see Chapter 4 and references therein). Two different mechanisms for this RNA-RNA annealing activity have previously been proposed, although they are not mutually exclusive. One mechanism involves simultaneous binding of both RNAs, thus bringing them into close proximity, effectively increasing their local concentration and chance of forming an intermolecular duplex. A second mechanism involves destabilisation of inhibitory intramolecular secondary structures (5.1B). For U4/U6 duplex formation this second mechanism is likely to play an important role, as destabilisation of the 3' stem-loop and telestem structures in U6 snRNA is required to allow U6 to form basepairing interactions with U4 snRNA (Fig. 5.1A). The "bridging" or "matchmaker" mechanism may however, be equally important. Although both Prp24p and Lsm2-8p are known to be involved in U4/U6 di-snRNP formation, it is unclear exactly what parts the individual proteins play in U4/U6 duplex formation. This chapter describes experiments that try to address this question. The main focus is on the potential role of Lsm2-8p in destabilisation of the U6 snRNA 3' stem-loop structure. Effects of U6 point mutations predicted to disrupt the 3' stem-loop were investigated both *in vivo* and *in vitro*. A strand displacement assay using an RNA/DNA duplex that resembles the stem-loop was developed and activities of the recombinant yeast and human LSm complexes as well as *E. coli* Hfq were examined.

5.2 Effects of mutations that destabilise the U6 3' stem-loop

5.2.1 U6 3' stem-loop destabilising point mutations do not rescue *lsm6Δ* and *lsm7Δ* growth phenotypes

Strains with a deletion of *LSM6* or *LSM7* are viable, but temperature sensitive, have reduced levels of U6 snRNA and are defective in efficient incorporation of U6 snRNA into U4/U6 di-snRNPs and U4/U6-U5 tri-snRNPs respectively (Mayes *et al.*, 1999; Verdone *et al.*, 2004). If Lsm2-8p promotes U4/U6 annealing by destabilisation of the U6 3' stem-loop, this step is likely to be affected by *lsm* deletions. Point mutations that are predicted to disrupt this stem-loop (Fig. 5.1C) may

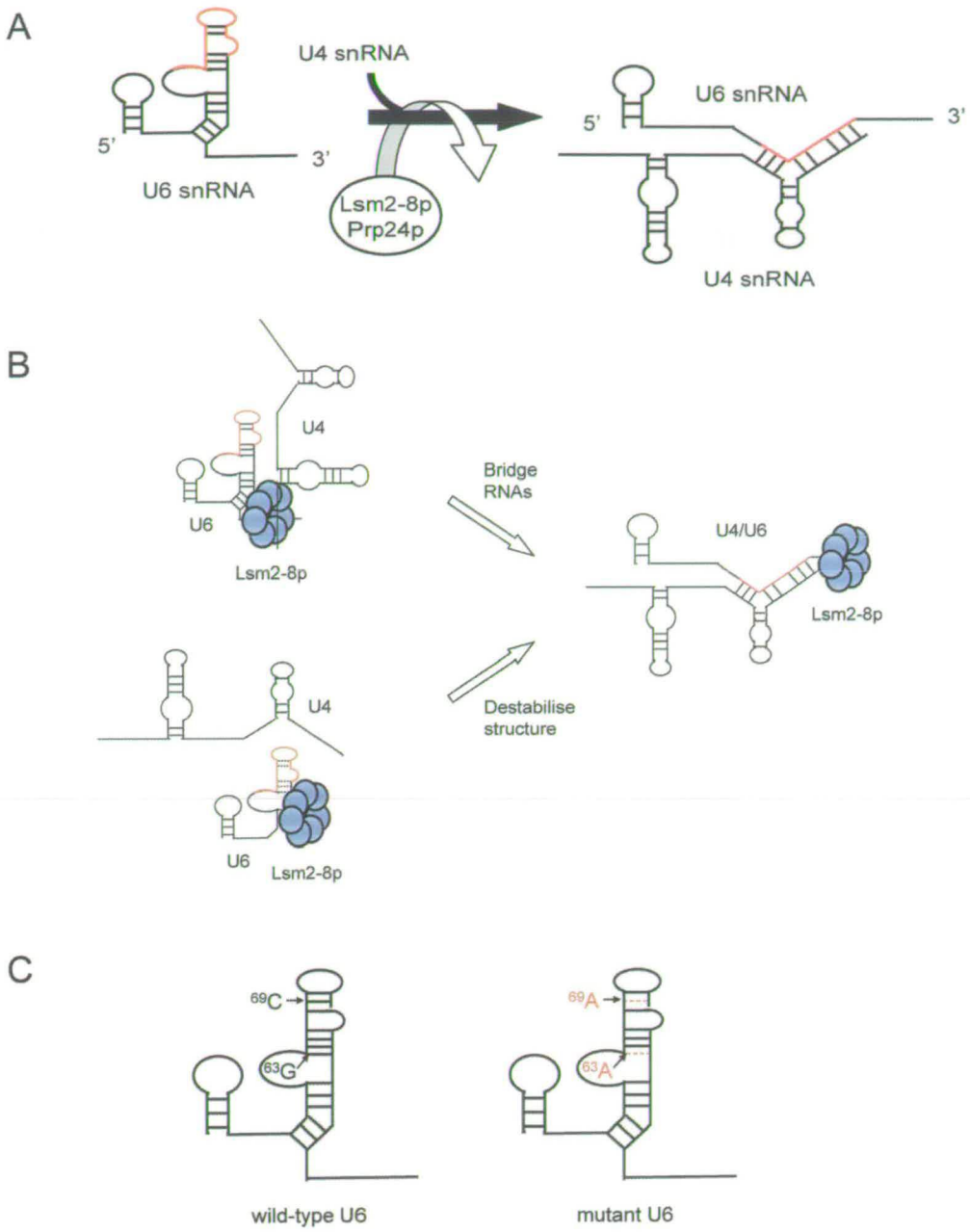


Fig. 5.1 U4/U6 annealing. **(A)** To allow U4 and U6 snRNA to anneal the U6 3' stem-loop (red) needs to be destabilised; **(B)** Two possible mechanisms by which Lsm2-8p could promote U4/U6 annealing; **(C)** G63A and C69A mutations destabilise the 3' stem-loop, but do not affect basepairing in the U4/U6 duplex.

therefore rescue this defect. This was previously investigated by Page (2000) who reported that C69A and G63A/C69A mutations in *SNR6*, but not a G63A mutation, rescued the temperature sensitive growth of an *lsm6Δ* strain, but not of an *lsm7Δ* strain. These mutations were inserted into the genomic copy of *SNR6* using a PCR-based *snr6-HpH* cassette. However, strains created by Page (2000) could not be recovered from glycerol stocks and attempts to recreate these strains either using the same or other strategies to incorporate these mutations into the genome were unsuccessful. Therefore an alternative strategy was used to determine the effect of these *snr6* mutations on the growth phenotype of wild-type, *lsm6Δ* and *lsm7Δ* strains. The G63A, C69A and G63A/C69A mutations were introduced in a plasmid based copy of the *SNR6* gene (pSX6) by PCR-based site-directed mutagenesis. Resulting plasmids were checked for the presence of the correct mutations and absence of any other mutations by sequencing. One other spontaneous double mutation, G63A/A82G was found. Just like the G63A and C69A mutations, the A82G mutation is not involved in basepairing interactions in the U4/U6 duplex, but may destabilise the 3' stem-loop, in this case by disturbing an A-U Watson-Crick interaction. All four mutations were introduced into BSY557 (WT), MRY1 (*lsm6Δ*) and MRY2 (*lsm7Δ*) by plasmid shuffle (Fig. 5.2A), resulting in strains that only express U6 snRNA from these plasmids: MRY3 to 5, MRY21 to 26 and MRY61 to 64 (see Table 2.1.6.1). These strains were tested for growth at 30 and 37°C (Fig. 5.2B). The *snr6* mutations had no noticeable effect on growth of *LSM+*, *lsm6Δ* and *lsm7Δ* strains. To confirm the presence of the *snr6* mutations, plasmids were rescued and re-sequenced. All mutations were present where expected. To confirm that G63A and C69A really did not rescue growth defects of the *lsm6Δ* strain, as had been reported by Page (2000), growth rates of the appropriate strains (MRY4, 23 and 24) were measured at 30°C in liquid YPDA. Growth rates were identical and all strains stopped growing approximately 12 h after shifting the cultures to 37°C (data not shown).

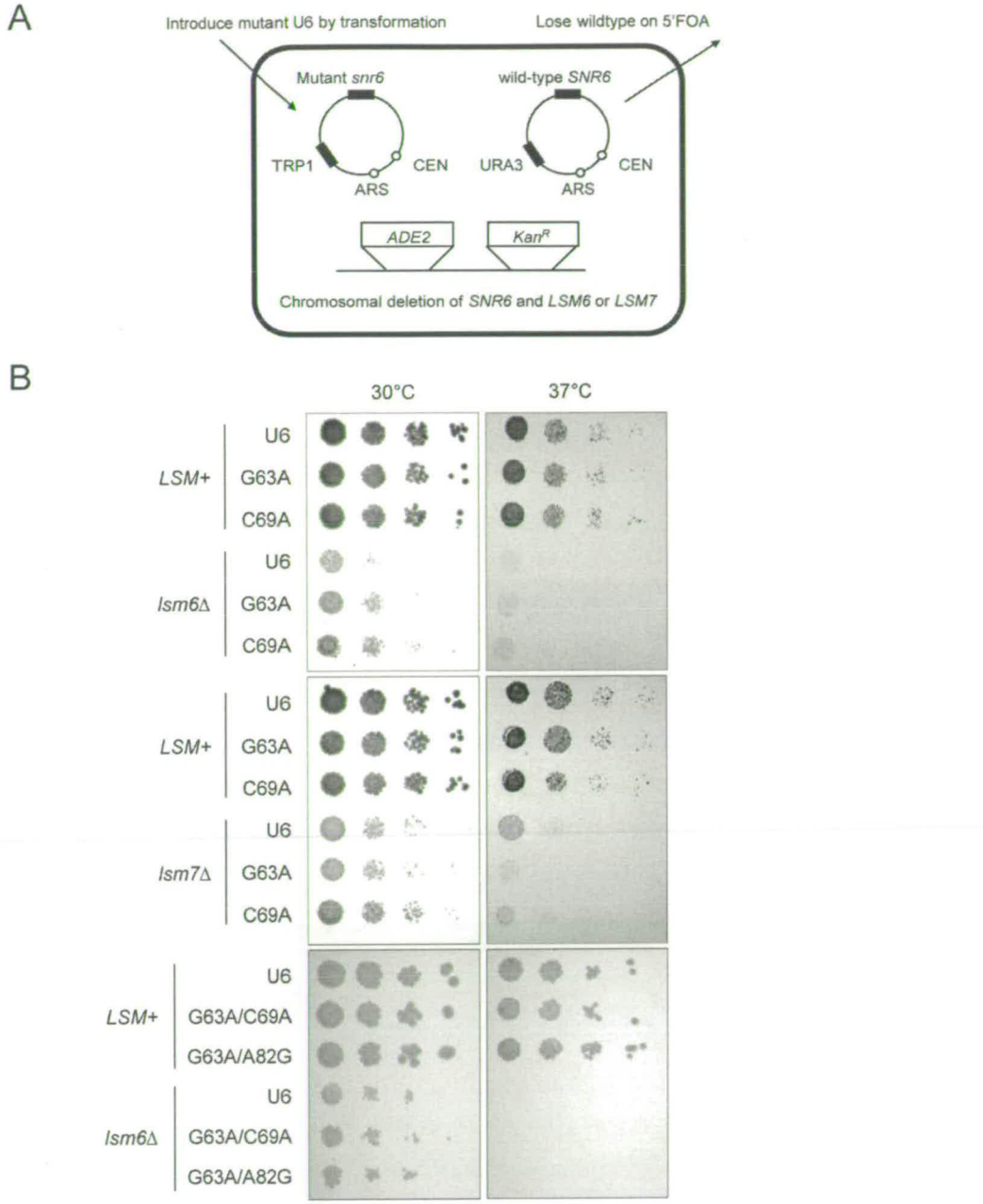


Fig. 5.2 U6 3' stem-loop destabilising mutations do not rescue temperature-sensitive growth of *lsm6Δ* and *lsm7Δ*. **(A)** *snr6* mutations were introduced by plasmid shuffle into BSY557 (WT), MRY1 (*lsm6Δ*) and MRY2 (*lsm7Δ*). **(B)** The resulting strains were spotted on YPDA in 10-fold serial dilutions and incubated at 30 or 37°C. MRY3, MRY21, MRY22, MRY61 and MRY62 (*LSM+* with *SNR6* (U6), *snr6-G63A*, *C69A*, *G63A/C69A* and *G63A/A82G* respectively) did all display similar amounts of growth; MRY4, MRY23, MRY24, MRY63 and MRY64 (*lsm6Δ* with *SNR6* (U6), *snr6-G63A*, *C69A*, *G63A/C69A* and *G63A/A82G* respectively) were all similarly temperature-sensitive, as were MRY5, MRY25 and MRY26 (*lsm7Δ* with *SNR6, snr6-G63A* and *C69A*)

In summary, *snr6-G63A*, *–C69A*, *-G63A/C69A* and *-G63A/A82G* mutations that are predicted to destabilise the U6 snRNA 3' stem-loop structure (Fig. 5.1C) do not rescue slow growth of *lsm6Δ* at permissive temperature or lethality at non-permissive temperature. Nor do the *snr6-G63A* and *–C69A* mutations rescue the growth defects of *lsm7Δ* and, surprisingly, no growth defect was associated with any of the *snr6* mutations in the *LSM+* background (at least not at 30 or 37°C).

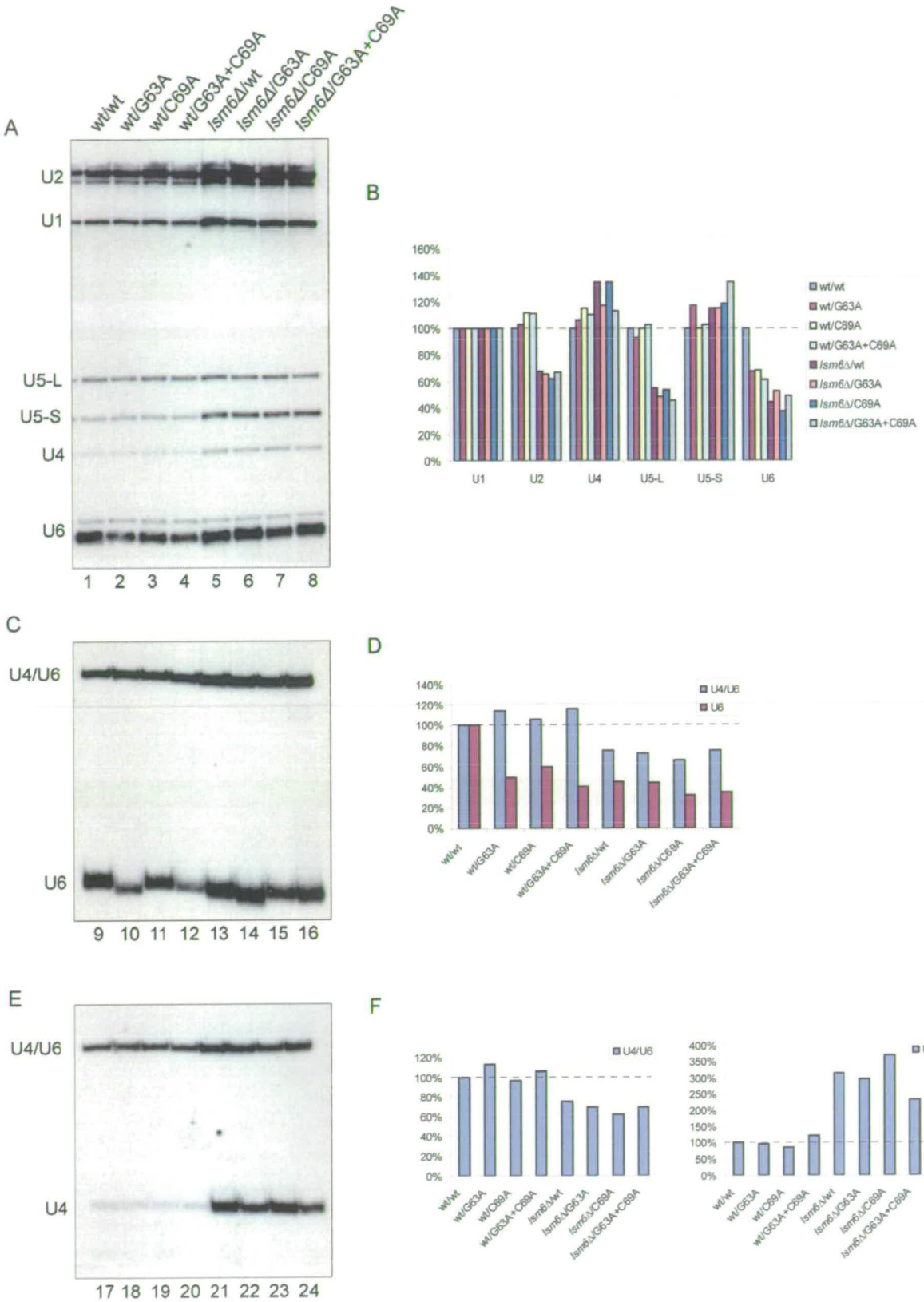
5.2.2 Stem-loop destabilising mutations do not increase U4/U6 levels *in vivo*

Temperature-sensitive phenotypes associated with the *lsm6* and *lsm7* deletions may be caused by reduced protection of shortened mRNA 3' ends by the Lsm1-7 complex as was previously proposed for *lsm1Δ* (He *et al.*, 2001) and not by reduced levels of U6 snRNA, U4/U6 di-snRNPs and/or U4/U6-U5 tri-snRNPs. If this is indeed the case the *snr6* mutations may still increase U4/U6 levels of *lsm6Δ* and/or *lsm7Δ* strains, in spite of their failure to rescue their growth phenotypes. To examine this possibility the levels of the snRNAs and the U4/U6 duplex were determined in *LSM6* and *lsm6Δ* strains with *SNR6*, *snr6-G63A*, *snr6-C69A* or *snr6-G63A/C69A* (Fig. 5.2). Total RNA was isolated from all strains under non-denaturing conditions. Levels of all snRNAs were examined by Northern blotting analysis with snRNA-specific probes, after denaturing PAGE (Fig. 5.3A and B). Levels of free U4 and U6 as well as U4/U6 duplex were determined by Northern blotting analysis after non-denaturing PAGE with U6 (Fig. 5.3C and D) and U4-specific probes (Fig. 5.3E and F). Loading was not entirely equal between lanes and was corrected for in the quantifications by normalising against the level of U1 snRNA in each lane. This was shown to be a valid loading control (see Fig. 8.5)

The *snr6* mutations do not significantly affect the levels of any of the other snRNAs in an *LSM6* strain, but surprisingly cause reduced levels of U6 snRNA (61-68% of *SNR6*; Fig. 5.3A and B). This suggests either that transcription of these mutant RNAs is less efficient or that the RNAs themselves are less stable. As expected the levels of U6 snRNA are reduced in the *lsm6Δ* strain, with slight variations in the exact level between U6 mutants (37-53% of *LSM6/SNR6*). Interestingly, all *lsm6Δ* strains also

show reduced levels of U2 snRNA (62-67% of *LSM6/SNR6*) and U5-L snRNA (46-55% of *LSM6/SNR6*). The reason for this is unclear, but this phenomenon is consistently observed in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains and is discussed in more detail in Chapter 8. Analysis of levels of free U6 snRNA and U4/U6 by native PAGE shows that the U6 mutations in the *LSM6* background only affect levels of free U6 snRNA (41-60% of *SNR6*) and not the levels of U4/U6 (106-116%; Fig. 5.3C and D). This suggests, that in spite of overall reduced levels of U6 RNA the mutant RNAs are efficiently (possibly even more efficiently) incorporated into U4/U6 di-snRNPs. In contrast, the levels of free U6 snRNA and U4/U6 are similarly reduced (to 32-46% and 66-76% of *LSM6/SNR6* respectively) for wild-type U6 and U6 mutants in the *lsm6Δ* background. The U6 mutations clearly do not correct the U4/U6 assembly defect caused by the *lsm6* deletion. This is confirmed by analysis of the U4/U6 levels using a U4 specific probe (Fig. 5.3E and F). Levels of free U4 are significantly increased for all *lsm6Δ* strains (232-369% of *LSM6/SNR6*), but because of the reduced U4/U6 levels the total level of U4 is only slightly increased (113-135% of *LSM6/SNR6*; Fig. 5.3A and B). An increase in free U4 levels is consistently seen for *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains, with a variable, but significant increase of total U4 snRNA levels as well. This seems to be, at least in part, an indirect effect of reduced U6 snRNA levels. A possible mechanism is discussed in Chapter 8.

Fig. 5.3 U6 3' stem-loop destabilising mutations do not rescue U4/U6 levels in *lsm6Δ*. RNA was extracted under non-denaturing conditions from MRY3, MRY21, MRY22, MRY61 (*LSM6* (wt) with *SNR6* (wt), *snr6-G63A*, *C69A* and *G63A+C69A*), MRY4, MRY23, MRY24, MRY63 (*lsm6Δ* with *SNR6* (wt), *snr6-G63A*, *C69A* and *G63A+C69A*). RNA was separated by denaturing and non-denaturing PAGE, Northern blotted and probed for the spliceosomal snRNAs. (A) Northern blot after denaturing PAGE and (B) quantification of all spliceosomal snRNAs. (C) Northern blot after non-denaturing PAGE probed for U6 RNA and (D) quantification of U4/U6 duplex and free U6 snRNA. (E) Northern blot after non-denaturing PAGE probed for U4 RNA and (F) quantification of U4/U6 duplex and free U4 snRNA. Levels were expressed as a percentage of *LSM6/SNR6* (wt/wt) and were corrected for equal loading by comparison to the amount of U1 snRNA present in each lane.


A
wt/wt
wt/G63A
wt/C69A
wt/G63A+C69A
lsm6Δ/wt
lsm6Δ/G63A
lsm6Δ/C69A
lsm6Δ/G63A+C69A
U2
U1
U5-L
U5-S
U4
U6
1 2 3 4 5 6 7 8
B
160%
140%
120%
100%
80%
60%
40%
20%
0%
U1 U2 U4 U5-L U5-S U6
wt/wt
wt/G63A
wt/C89A
wt/G63A+C69A
lsm6Δ/wt
lsm6Δ/G63A
lsm6Δ/C69A
lsm6Δ/G63A+C69A
C
U4/U6
U6
9 10 11 12 13 14 15 16
D
140%
120%
100%
80%
60%
40%
20%
0%
U4/U6
U6
wt/wt
wt/G63A
wt/C69A
wt/G63A+C69A
lsm6Δ/wt
lsm6Δ/G63A
lsm6Δ/C69A
lsm6Δ/G63A+C69A
E
U4/U6
U4
17 18 19 20 21 22 23 24
F
120%
100%
80%
60%
40%
20%
0%
U4/U6
400%
350%
300%
250%
200%
150%
100%
50%
0%
U4
wt/wt
wt/G63A
wt/C69A
wt/G63A+C69A
lsm6Δ/wt
lsm6Δ/G63A
lsm6Δ/C69A
lsm6Δ/G63A+C69A

5.2.3 Stem-loop destabilising mutations increase U4/U6 annealing efficiency *in vitro*

As the *snr6* mutations lead to reduced levels of U6, but not U4/U6 in an *LSM+* background, it suggests that the stem-loop destabilising mutations may increase U4/U6 annealing. Failure to correct U4/U6 levels in the *lsm6Δ* background suggests that these mutations may only do so *in vivo* in the presence of a fully functional Lsm2-8 complex. Alternatively, the U6 mutants are more susceptible to degradation unless they are basepaired to U4. To test if the stem-loop destabilising mutations lead to increased annealing *in vitro*, wild-type U6, U6-G63A and U6-C69A were

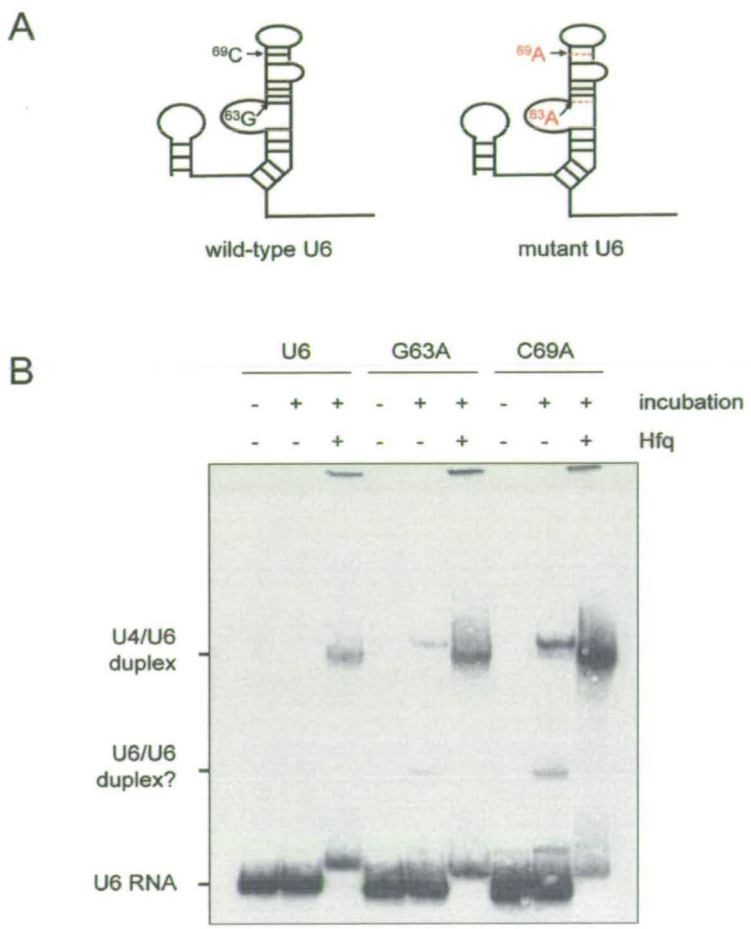


Fig. 5.4 U6 3' stem-loop destabilising mutations make U4/U6 annealing *in vitro* more efficient. **(A)** Schematic representation of U6 snRNA secondary structure with putative stem-loop destabilising mutations, G63A and C69A, indicated; **B)** *In vitro* U4/U6 annealing with *in vitro* transcribed and gel-purified U4 RNA and U6, U6-G63A and U6-C69A RNA in the absence and presence of recombinant Hfq. Annealing reactions were separated by non-denaturing PAGE and the Northern blot was probed for U6 snRNA.

transcribed *in vitro* and gel-purified. U4/U6 annealing reactions were carried out as in Chapters 4, with or without recombinant Hfq present. Both mutations increased the amount of U4/U6 duplex significantly, both with and without Hfq there to promote annealing, but C69A had a stronger effect than G63A (Fig. 5.4B). U4 was shown to be present in the band indicated with U4/U6 duplex, but not in the additional bands of slower mobility (data not shown), that may represent a U6/U6 duplex. These data suggest that the 3' stem-loop structure in U6 RNA inhibits U4/U6 annealing *in vitro* as well as *in vivo*, and that when this secondary structure is destabilised by point mutations and/or protein binding, formation of a U4/U6 duplex is more efficient.

5.3 Can Lsm complexes change RNA structure?

One of the functions of the Lsm2-8 complex in addition to protecting the U6 snRNA from degradation may be to destabilise the U6 3' stem-loop structure, thus promoting U4/U6 annealing. The Lsm1-7 complex may similarly destabilise RNA secondary structures and thereby affect RNA-protein interactions. Hfq has been shown to alter RNA-RNA and RNA-protein interactions by changing RNA secondary structures in some cases (Schuppli *et al.*, 1997; Møller *et al.*, 2002; Zhang *et al.*, 2002; Moll *et al.*, 2003; Geissmann *et al.*, 2004), but has been shown to be able to promote RNA-RNA interactions without affecting RNA secondary structures in others (Brescia *et al.*, 2003; Mikulecky *et al.*, 2004). This RNA chaperone function may be conserved between all Sm-like protein complexes.

5.3.1 Designing a strand displacement assay

To investigate the ability of Lsm2-8p to melt out the U6 3' stem-loop structure an *in vitro* strand displacement assay was designed. Initial attempts to create a U6 RNA structure *in vitro* by mixing *in vitro* transcribed and gel-purified 5' (nt 1-70) and 3' (nt 71-112) U6 fragments were unsuccessful. This was thought to be due to the formation of stable intramolecular structures within both of these U6 fragments, the 5' fragment in particular (predicted by mfold (Mathews *et al.*, 1999; Zuker, 2003); data not shown). Therefore an alternative strategy was taken in which a duplex was

formed between U6(77-112) RNA and a complementary end-labelled DNA oligo, U6(77-92)rc (Fig. 5.5B). The melting temperature of the RNA/DNA duplex was estimated at approximately 54°C (default settings at http://bioweb.pasteur.fr/seqanal/interfaces/melting.html; Le Novere, 2001). Incubation in the presence of purified recombinant proteins at 37°C and subsequent native PAGE allowed detection of remaining duplex and DNA oligo freed by strand displacement. Reactions were carried out in the presence of a 20-fold excess of unlabeled DNA oligonucleotide to quench potential re-formation of the duplex with the displaced labelled 16-mer. Additionally, absence of significant strand displacement without proteins provided evidence that the duplex was thermodynamically stable under the assay conditions. A variety of substrates were created, mostly with similar melting temperatures, but different sequences in either the 3' single stranded part or in the duplex to allow to test for sequence specificity of strand displacement activity (see Appendix Fig. A1.5).

5.3.2 Proof of principle: Ded1p helicase unwinds the U6-like RNA/DNA duplex structure

To test the strand displacement assay, *S. cerevisiae* Ded1p helicase was used. Recombinant Ded1p has been shown to unwind an RNA/DNA duplex of unrelated sequence, but of similar design, including a 3' single stranded RNA tail and a duplex with a melting temperature of approximately 52°C (Cordin *et al.*, 2004). Here, 25 nM of Ded1p (kindly provided by Olivier Cordin) was able to unwind 250 nM of the U6-like duplex efficiently within minutes in the presence of 1 mM ATP (Fig. 5.5C). In contrast, only minimal strand displacement was observed in the absence of Ded1p or with Ded1p in the absence of ATP. This is comparable to what was found by Cordin *et al.* (2004) and shows that this substrate can be used to study strand displacement activity.

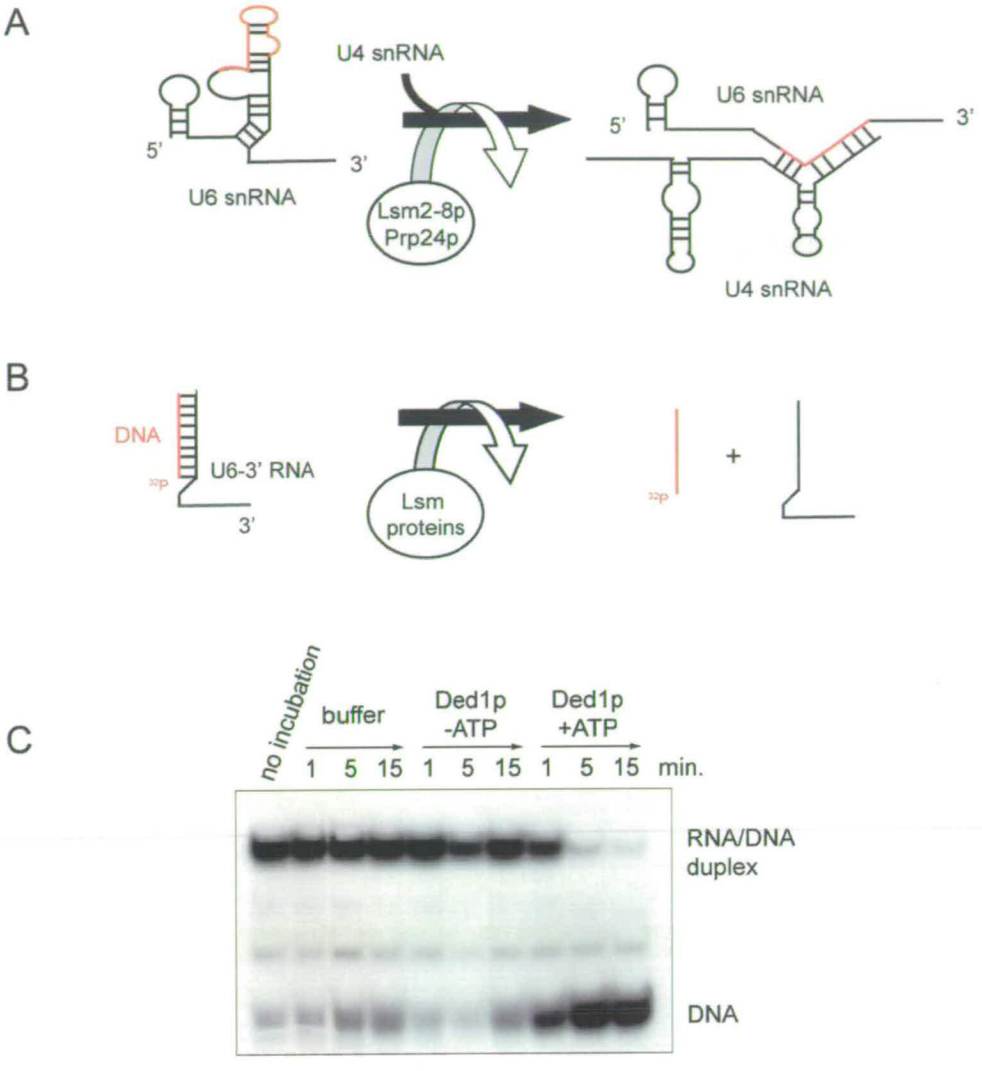


Fig. 5.5 Design of a strand displacement assay based on the U6 snRNA 3' stem-loop, and strand displacement activity of Ded1p helicase. **(A)** Schematic representation of the U6 snRNA secondary structure and U4/U6 structure; **B)** Schematic representation of the strand displacement assay. The substrate is an RNA/DNA duplex, resembling the 3' stem-loop structure, made with *in vitro* transcribed gel-purified U6-3' RNA and a complementary end-labelled DNA oligonucleotide; **(C)** Strand displacement over time with 25 nM Ded1p and 250 nM of substrate in the absence and presence of 1 mM ATP. Duplexes and oligonucleotides were separated by native PAGE.

5.4 Yeast Lsm complexes have strand displacement activity

Initial experiments with *in vitro* assembled yeast Lsm protein complexes displayed high levels of strand displacement activity. However, the discovery of co-purification of *E. coli* Hfq with Lsm8p in particular made interpretation of results problematic (see Appendix A1). Hfq-free Lsm complexes were then formed which were able to bind RNA (see 4.2.2), but were not able to promote U4/U6 annealing (see 4.3.1). These complexes failed to show strand displacement at rates which were seen in the earlier experiments. The substrate concentration was therefore decreased (from 250 nM) to 25 nM, which is approximately the same as the RNA concentrations used in the RNA bandshift assays (4.2.2). Both Lsm1-7p and Lsm2-8p now displayed strand displacement after 30-60 min (Fig. 5.6A), but a significant part of the substrate RNA was shown to be degraded (data not shown). It was unclear however, to what extent RNA degradation contributed to strand displacement and to what extent RNA was degraded after unwinding by the Lsm proteins. Reactions were therefore incubated for shorter periods of time (up to 20 min) and both Lsm1-7p and Lsm2-8p still displayed strand displacement activity (Fig. 5.6B), without significant RNA degradation (Fig. 5.6C). A higher level of strand displacement by Lsm2-8p than by Lsm1-7p corresponds to higher affinity RNA binding in the RNA bandshift assays (Fig. 4.1). It is uncertain however, whether this reflects a real difference between Lsm1-7p and Lsm2-8p or different levels of correctly formed complexes. Just like in the RNA bandshift assays, a large excess of protein over substrate is needed. This is likely to reflect the low level of correct heteroheptameric complexes rather than low affinity and/or strand displacement activity of these complexes. Addition of up to 1 mM ATP did not increase strand displacement activity (data not shown), suggesting that unwinding by the Lsm complexes does not require ATP.

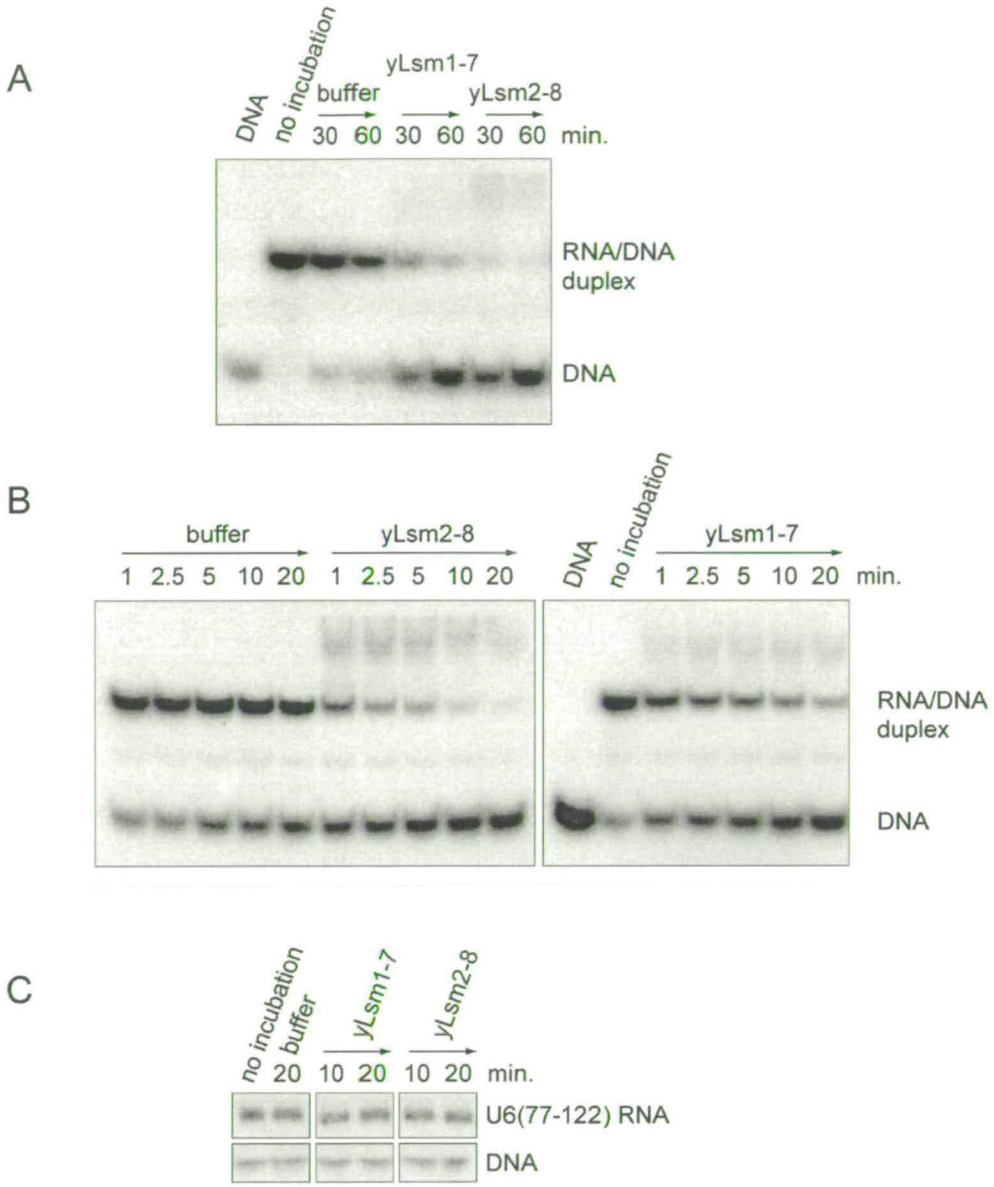


Fig. 5.6 Strand displacement by *in vitro* assembled yeast Lsm1-7 and Lsm2-8 complexes; **(A)** 25 nM of U6(77-112)/(77-92)rc duplex was incubated for 30 or 60 min with buffer, 1 μM of yLsm1-7p or yLsm2-8p; **(B)** Repeat of (A) but shorter incubations of up to 20 min; **(C)** Duplicate strand displacement reactions of (B) were separated by denaturing PAGE and analysed by Northern blotting with U6(77-92)rc probe showing absence of RNA degradation. The DNA half of the duplex substrate shows equal loading.

5.5 RNA-binding by recombinant Hfq is required for its strand displacement activity

Recombinant Hfq was shown to bind U4 and U6 snRNA (Fig. 4.2) and to promote annealing between these RNAs *in vitro* (Fig. 4.7 and 4.8). To investigate whether Hfq may promote U4/U6 annealing by destabilising the U6 3' stem-loop, recombinant Hfq was tested for strand displacement activity in the assay described above. As was seen for the hfq-free Lsm complexes, strand displacement activity for recombinant Hfq was not as robust as that detected for the Lsm-Hfq mixture (Appendix A1). A large excess of Hfq over duplex substrate was needed to detect strand displacement. With enough Hfq added however, most substrate was unwound within 60 minutes after what seems to be initial binding of the duplex by Hfq (Fig. 5.7A; compare to Fig. A1.4C). Strand displacement activity did not increase by addition of 1 mM ATP (data not shown), suggesting that Hfq does not require ATP to unwind the RNA/DNA duplex. As Hfq has been suggested to have ATPase activity (Sukhodolets and Garges, 2003) this finding is significant.

The Hfq C-terminal extension was shown to be non-essential for RNA binding and U4/U6 annealing activity. In contrast, RNA binding as well as annealing was undetectable for Hfq-Y55A, which carries a point mutation in a residue thought to be involved in RNA binding (Fig. 4.7A). Just like RNA-binding and U4/U6 annealing, strand displacement activity was similar for Hfq and HfqΔC, but almost non-existent for Hfq-Y55A (Fig. 5.7B. Northern analysis of U6(77-112) RNA after strand displacement (Fig. 5.7C) revealed that strand displacement by Hfq and HfqΔC was not due to RNA degradation, whereas the low level of strand displacement observed for the highest amount of Hfq-Y55A seems to be caused, at least in part, by RNA degradation, presumably by contaminating RNases.

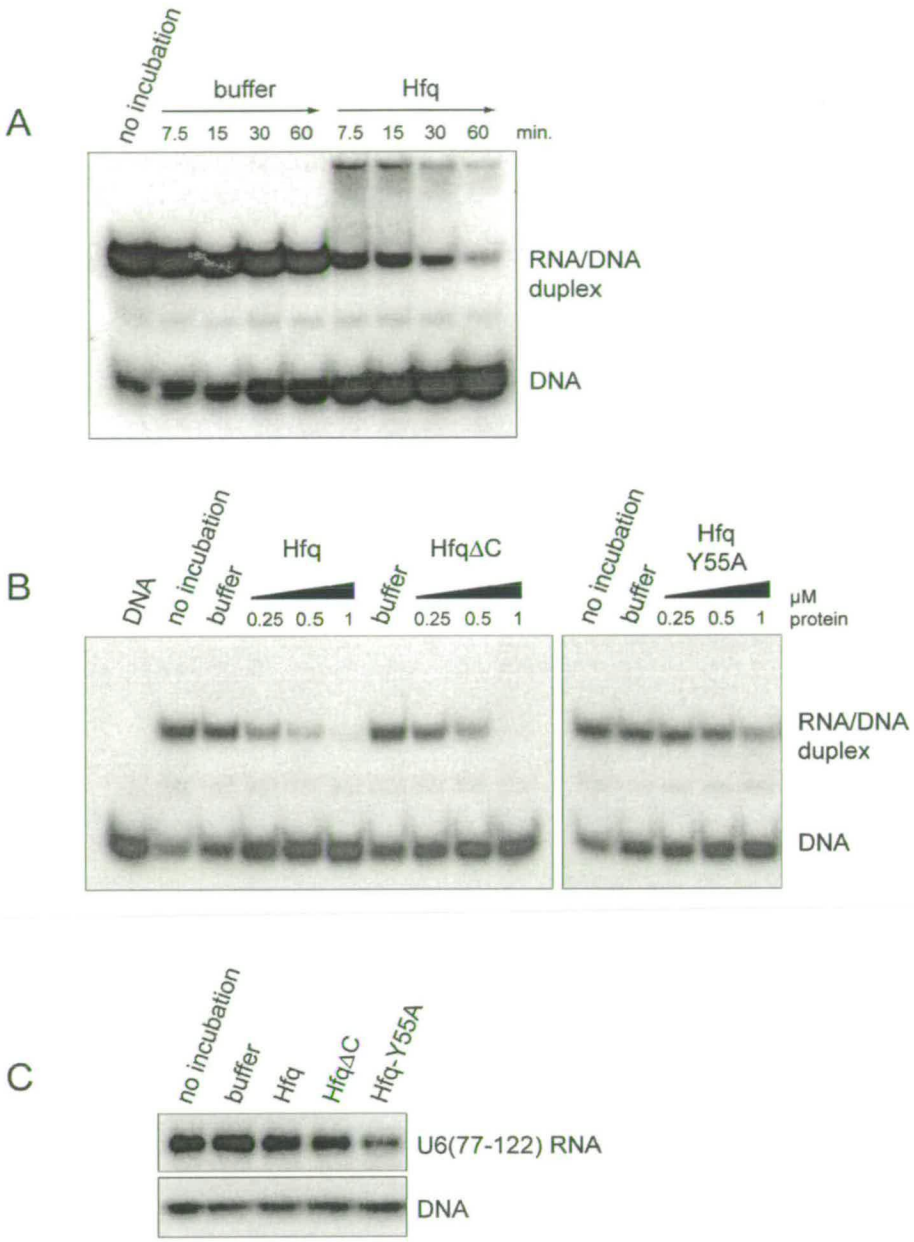


Fig. 5.7 Strand displacement by recombinant Hfq; **(A)** 25 nM of U6(77-112)/(77-92)rc duplex was incubated for up to 60 min with buffer or 1 μM of His-HA-Hfq (hexamer); **(B)** 25 nM of substrate was incubated with 1, 0.5 or 0.25 μM of His-HA-Hfq, His-Myc-HfqΔC or Hfq-Y55A-His for 60 min; **(C)** Duplicate strand displacement reactions of of the highest protein concentrations in (B) were separated by denaturing PAGE and analysed by Northern blotting with U6(77-92)rc probe showing absence of RNA degradation. The DNA half of the duplex substrate shows equal loading.

5.6 hLSm4 may be essential for strand displacement by the human LSm complexes

In vitro assembled human LSm1-7 and Lsm2-8 were tested in the above strand displacement assay and both complexes were shown to unwind the duplex substrate (Fig. 5.8A). In contrast to the effect that different assembly methods had on the U4/U6 annealing activities of the resulting complexes (4 M urea vs. 8M urea; see 4.3.3.1), only a minimal difference was observed in their strand displacement activities (Fig. 5.8A and B). This correlates with the absence of an effect on RNA binding by these complexes and suggests that RNA binding is sufficient for strand displacement activity. In contrast, there may be additional requirements for RNA-RNA annealing activity. The same may be true for the yeast Lsm complexes and these complexes may also display *in vitro* annealing activity when assembled under different conditions.

The same "incomplete" LSm complexes that were used as controls in the U4/U6 annealing assays (4.3.3.2) were used to determine the requirements for strand displacement activity. Similar to U4/U6 annealing activity displayed by the hexameric hLSm2-7 and pentameric hLSm23567 control complexes (Fig. 4.10) strand displacement activity was higher for hLSm2-7 than for hLSm1-7 and hLSm23567 only displayed minimal activity (Fig. 5.8C). To determine whether the presence of hLSm4 or the formation of a hexameric complex is essential for strand displacement activity a hexameric complex lacking hLSm4 (hLSm123567) and a pentameric complex with hLSm4 (hLSm4-8) were assembled and tested in the strand displacement assay. Minimal activity of the hLSm123567 and efficient strand displacement by hLSm4-8 suggest that hLSm4 is required for unwinding activity (Fig. 5.9A and B). Again, this is identical to what was found to be essential for annealing activity (Fig. 4.11A and B). To determine whether, like for U4/U6 annealing, hLSm4 is not only essential but also sufficient for strand displacement activity hLSm4 and hLSm48 were tested in the strand displacement assay (Fig.5.9C).

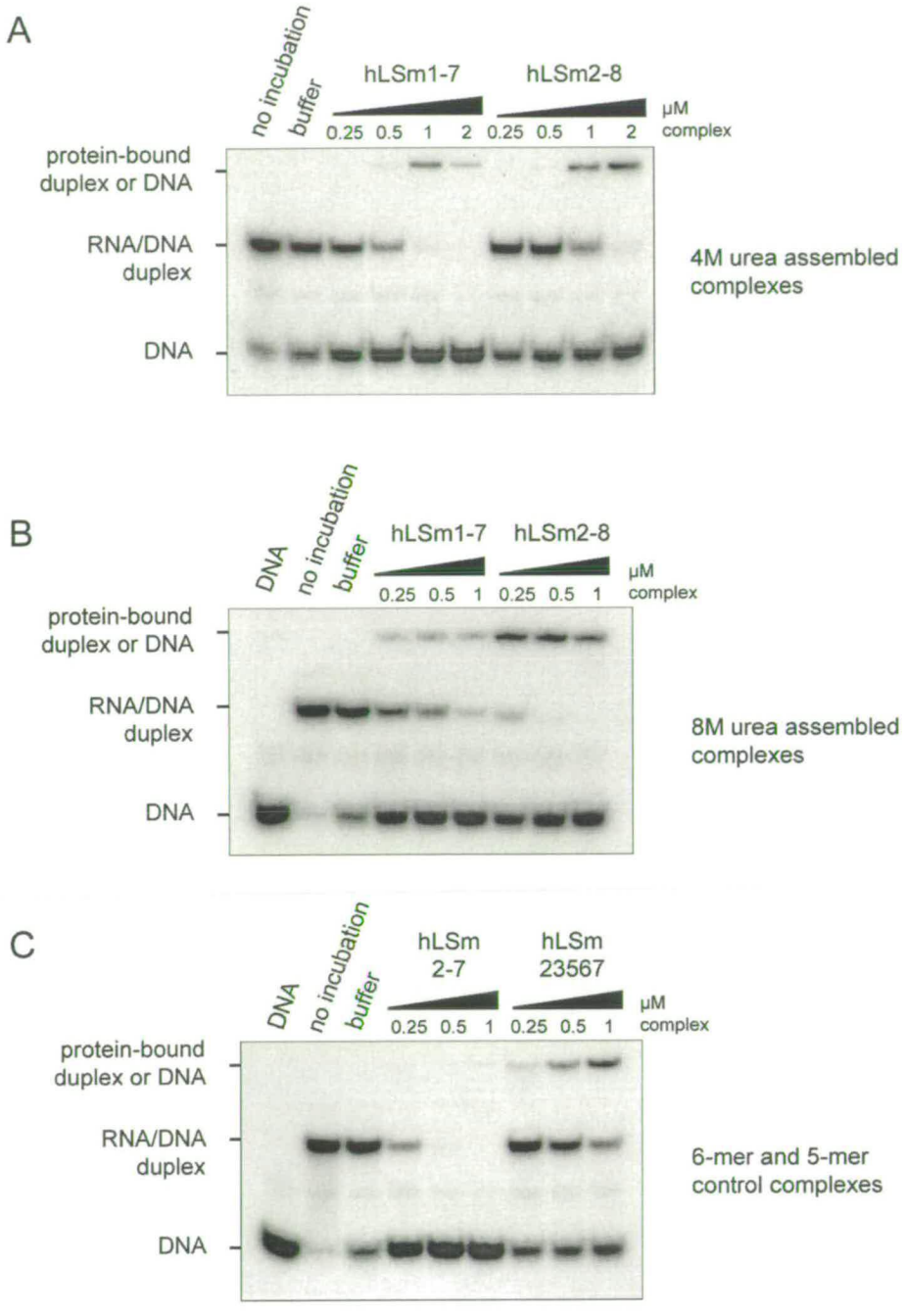


Fig. 5.8 Strand displacement activity of *in vitro* assembled human LSm complexes depends on the assembly method and may require hLSm4; **(A)** 25 nM of U6(77-112)/(77-92)rc duplex was incubated for 60 min with buffer, 2 μM of hLSm1-7 or 1 μM of hLSm2-8 (assembled using the 4 M urea method) and 2-fold dilutions thereof; **(B)** 25 nM of substrate was incubated for 60 min with buffer, 1 μM, 0.5 μM and 0.25 μM of hLSm1-7 or hLSm2-8 (assembled using the 8 M urea method); **(C)** 25 nM of substrate was incubated for 60 min with buffer, 1 μM, 0.5 μM and 0.25 μM of hLSm2-7 or hLSm23567.

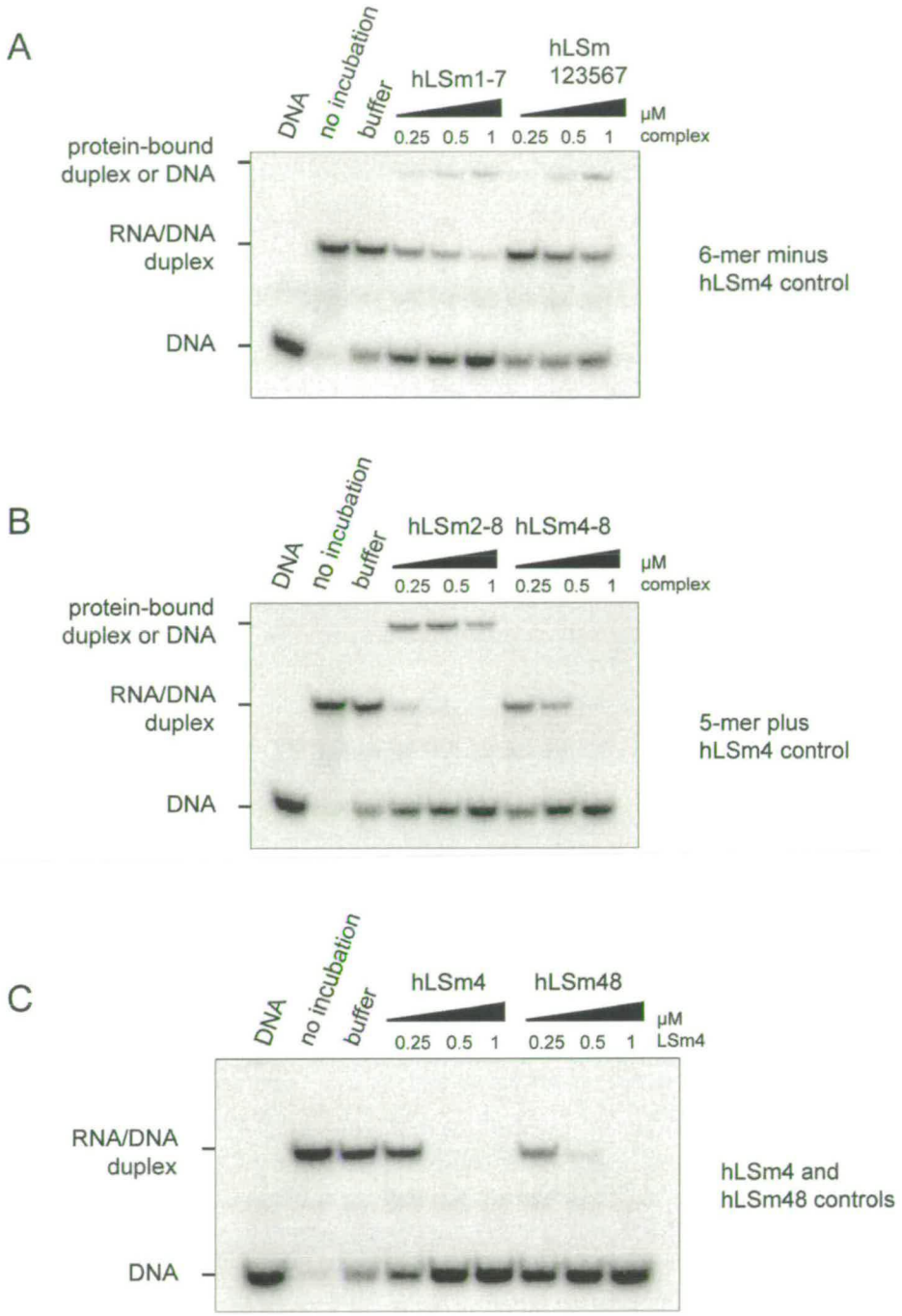


Fig. 5.9 hLSm4 is essential and sufficient for strand displacement activity of *in vitro* assembled human LSm complexes; **(A)** 25 nM of U6(77-112)/(77-92)rc duplex was incubated for 60 min with 1, 0.5 and 0.25 μM of hLSm1-7 or hLSm123567; **(B)** 25 nM of substrate was similarly incubated with 1, 0.5 and 0.25 μM of hLSm2-8 or hLSm4-8; **(C)** 25 nM of substrate was similarly incubated with hLSm4 or hLSm48 at 1, 0.5 and 0.25 μM of hLSm4 monomers.

Strand displacement activity for these proteins was as high as that of hLSm2-7 per mole of hLSm4, indicating that hLSm4 has annealing as well as strand displacement activity. The fact that the strand displacement activity of hLSm1-7 and hLSm2-8 was lower at the same molar concentration of hLSm4 suggests that hLSm1 as well as hLSm8 may somehow mask this hLSm4 activity. However, it can not be ruled out that the activity comes from a hLSm4-specific *E. coli* contaminant protein and that its activity is inhibited by hLSm1-7 and hLSm2-8 RNA binding.

5.7 Discussion

In this chapter I have shown that, although mutations that disrupt the U6 snRNA 3' stem-loop structure lead to more efficient U4/U6 annealing *in vitro*, the same mutations do not increase the level of U4/U6 duplex present *in vivo* for an *lsm6Δ* strain, nor do they rescue temperature-sensitive growth of this strain. Interestingly C69A had a more pronounced positive effect on U4/U6 annealing *in vitro* than did G63A. This may simply indicate a less stable stem-loop for U6-C69A than for U6-G63A. Both mutations disrupt a G-C basepair however, suggesting that the position of the disruption is important. Karaduman *et al.* (2006) noted the resemblance of the 3' stem-loop to a so-called kissing loop. It is likely that this loop and the complementary free 5' end of U4 snRNA interact, thus initiating formation of U4/U6 stem II by helix propagation. As C69 is close to the top of the stem-loop its mutation is likely to lower the energy barrier that needs to be overcome to initiate U4/U6 duplex formation. G63 is further down the stem and may therefore have less of an effect.

In vivo, binding by Prp24p and Lsm2-8p is likely to affect the stability of the U6 3' stem-loop. To examine the potential function of Lsm2-8p in destabilising this stem-loop and to investigate similar RNA chaperone functions of Lsm1-7p and Hfq, a strand displacement assay was designed using an RNA/DNA duplex resembling the 3' end of U6 snRNA. All Sm-like protein complexes that were tested, yeast and human Lsm1-7 and Lsm2-8, and *E. coli* Hfq, were able to displace the DNA oligonucleotide from the U6-3' RNA fragment. No ATP was required by any of

these Sm-like complexes, but results with an Hfq mutant and a variety of hLSm complexes indicate that RNA binding is essential for strand displacement activity. As an excess of protein over duplex substrate was needed for complete unwinding, it seems that it may be the process of RNA-binding itself that leads to strand displacement. I propose that binding of the uridine-rich sequence at the 3' end of U6 RNA along the inside of the doughnut-shaped Lsm complexes bends the RNA forcing open the bottom end of the duplex, effectively lowering the melting temperature. In addition, binding by extensions of specific Lsm proteins to the RNA duplex itself may contribute to duplex unwinding (Fig. 5.10). The results described in this Chapter and Chapter 4 suggest that Lsm4p may play an important role in the latter. Comparison of the naked U6 RNA structure to that bound by Prp24p and/or Lsm2-8p (Karaduman *et al.*, 2006) suggests that both destabilisation of the telestem and 3' stem-loop by Lsm2-8p binding occur *in vivo*, as presence of Lsm2-8p leads to reduced protection of nucleotides in the telestem and in the 3' stem-loop. As the structure of U6 snRNA with Lsm2-8p alone is not known, it cannot be ruled out however, that these effects are indirect effects of Lsm2-8p on Prp24p-U6 binding (see also Chapter 6). Preferred binding of the Lsm complexes to a U-rich sequence flanked by a stem-loop structure favours a model in which protein-RNA interactions extend beyond those between the single stranded RNA and the proposed nucleotide binding pockets along the inside of the ring.

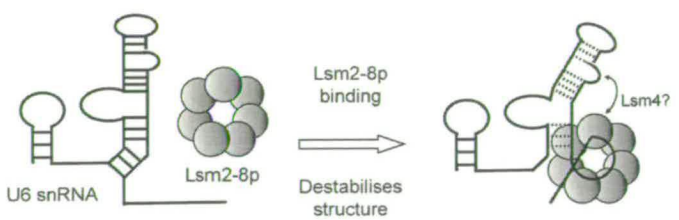


Fig. 5.10 A model for Lsm2-8p stem-loop destabilisation. The single stranded U-rich tail of U6 RNA is bound along the inside of the Lsm2-8p ring pulling apart the telestem. Additional interactions with other parts of the Lsm complex, possibly involving Lsm4p may further destabilise the 3' stem-loop structure.

Human LSm4 was shown to be essential and sufficient for annealing (Chapter 4) as well as strand displacement activity (this Chapter). This suggests that hLSm4 may be responsible for binding to and destabilisation of the stem-loop. Initial experiments with yeast Lsm4p suggest that this protein also has U4/U6 annealing activity. It would be interesting to investigate whether it is also able to perform strand displacement. However, no U4/U6 annealing activity was associated with the same yeast Lsm protein complexes that displayed RNA-binding and strand displacement activity. This suggests that destabilisation of the U6 3' stem-loop (at least as measured by the strand displacement assay) is not sufficient to promote U4/U6 annealing. As discussed in Chapter 4 the human LSm complexes lose their annealing activity when they are assembled using 8 M urea. This suggests that 8 M urea affects the hLSm4 functionality. However, an Lsm4p specific contaminant from *E. coli* can not be ruled out at this stage. Repeating these experiments with hLSm4 mutant proteins that are defective in complex formation and/or RNA-binding may shed some light on this. Additional experiments with yeast Lsm4p may also be useful.

In summary, I have shown strand displacement activity for *E. coli* Hfq as well as for the yeast and human Lsm complexes. The ability of Sm-like protein complexes to modulate RNA secondary structure may therefore be conserved. It allows these complexes to promote or disrupt RNA-RNA and RNA-protein interactions influencing a wide variety of RNA metabolic processes.

CHAPTER 6

Interactions between Prp24p and Lsm proteins

6.1 Introduction

Prp24p is an essential U6 snRNP protein that is important for U4/U6 di-snRNP formation. Recycling of U6 snRNA was shown to be impaired in splicing extracts immuno-depleted of Prp24p (Raghunathan *et al.*, 1998a) and recombinant Prp24p was shown to have U4/U6 annealing activity *in vitro*. This process does not require ATP, but is more efficient in the presence of Lsm2-8p (Ghetti *et al.*, 1995; Raghunathan *et al.*, 1998a; Rader *et al.*, 2002; this thesis). Yeast Prp24p has four RNA Recognition Motifs (RRMs; Shannon *et al.*, 1991; Rader *et al.*, 2002) and a highly conserved C-terminal decapeptide (Rader *et al.*, 2002; Bell *et al.*, 2000; Fig. 6.1). This C-terminal motif was termed SNNFL box (Kwan *et al.*, 2005), after its most conserved residues (SNDDFRKMFL) and was shown to be involved in a direct interaction with the Lsm2-8 complex, which was proposed to stabilise the Prp24p-U6 snRNA interaction (Rader *et al.*, 2002; Ryan *et al.*, 2002). A range of genetic interactions between mutations in *SNR14* (the U4 snRNA gene) and *SNR6* (the U6 snRNA gene) and *PRP24* that map to residues in these RRMs (Shannon *et al.*, 1991; Vidaver *et al.*, 1999) lead Rader and Guthrie (2002) to propose that RRM2 and RRM3 favour free U6 over U4/U6 RNA. At the same time, genetic interactions between mutations in the different RRMs lead them to propose that RRM1, RRM4 and the SNFFL box promote U4/U6 annealing. Lsm2-8p has been shown to be involved in U4/U6 di-snRNP formation, most likely as an RNA chaperone, both in yeast (Verdone *et al.*, 2004) and in humans (Achsel *et al.*, 1999; this work). Levels of U4/U6 RNA are reduced in *lsm6Δ* and *lsm7Δ* strains (Mayes *et al.*, 1999; Verdone *et al.*, 2004; this work). This may be caused, at least in part, by the reduced U6 snRNA levels in these strains. However, loss of efficient Lsm2-8p chaperone activity is likely to contribute. The exact roles Prp24p and the Lsm proteins play in the process of U4/U6 duplex formation are still unclear. In this chapter I examine interactions between these U6 snRNP polypeptides and propose a model for U4/U6 annealing.

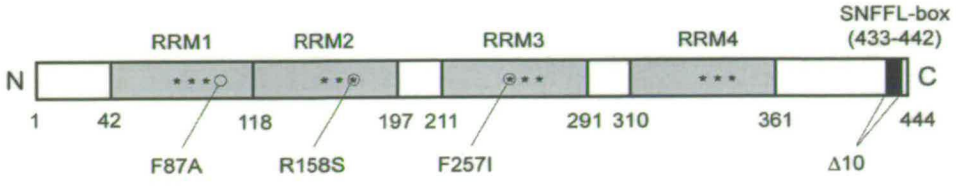


Fig. 6.1 Domains and mutations in yeast Prp24p. Prp24p contains four RNA recognition motifs (RRM1-4) and a highly conserved 10 aa motif near the C-terminus (SNFFL-box). Triple alanine mutations in each of the RRMs (RRM1-4sub; ***) affect residues that are proposed to be essential for RNA binding. R158S and F257I affect one of these residues in RRM2 and RRM3 respectively. F87A affects a highly conserved residue in RRM1. The prp24Δ10 mutation carries a deletion of the 10 residues covering the SNFFL-box. Amino acid coordinates are indicated at the bottom.

6.2 Investigating genetic interactions between *PRP24*, *LSM6* and *LSM7*

If the above proposed roles of the various Prp24p RRMs are correct, it may be expected that mutations in RRM2 and RRM3 could rescue slow growth phenotypes associated with *lsm6* and *lsm7* deletions, whereas mutations in RRM1 and RRM4 may cause synthetic lethality. To test this *LSM6* or *LSM7* was deleted in GLS618, a strain with a *prp24* deletion carrying a plasmid with *PRP24* under its own regulatory elements (pPR097, *CEN-URA3*). The resulting MRY15 (*lsm6Δ*) and MRY16 (*lsm7Δ*) strains were transformed with plasmids expressing *prp24* mutants (*CEN-HIS3*) and growth in the absence of wild-type Prp24p expression was tested for by streaking on 5-FOA plates (counter-selection against Ura+ cells; Fig. 6.2). The following *prp24* mutants were used: RRM1-4sub mutants with triple alanine mutations in putative RNA binding residues in each of the RRMs (Fig. 6.1), mutants with single point mutations in the RRM1-3 domains (F87A, R158S and F257I), a *prp24Δ10* mutant with a C-terminal 10 aa deletion of the residues involved in the interaction with Lsm2-8p, and various combinations of RRMsub mutations. *Prp24* mutations that were previously shown to be viable in the GLS618 (*LSM+*)

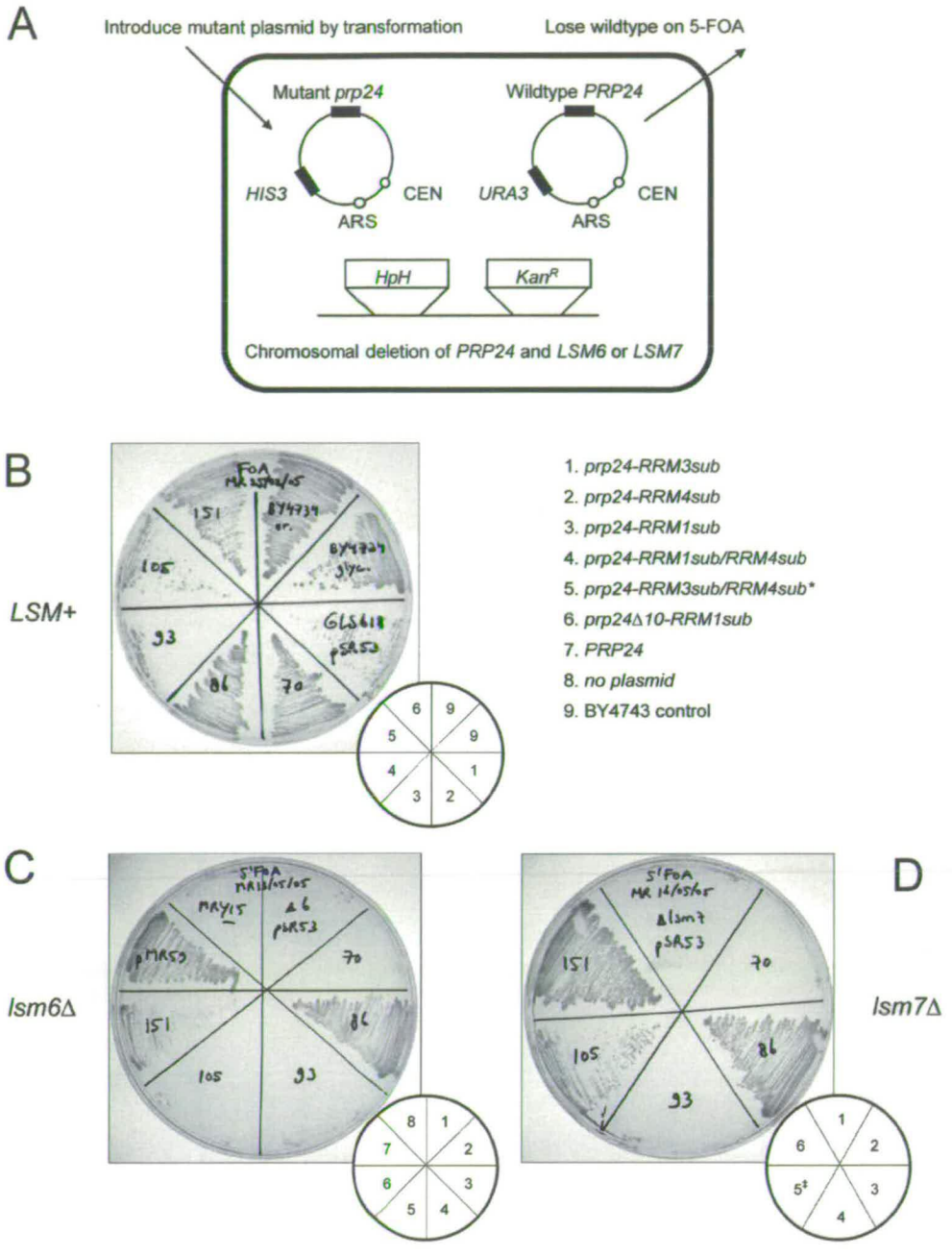


Fig. 6.2 Plasmid shuffle **(A)** in GLS618 (*LSM+*; **B**), MRY15 (*lsm6Δ*; **C**) and MRY16 (*lsm7Δ*; **D**). GLS618, MRY15 and MRY16 were transformed with the plasmids pSR53 (*prp24-RRM3sub*), pSR70 (*prp24-RRM4sub*), pSR86 (*prp24-RRM1sub*), pSR93 (*prp24-RRM1sub/RRM4sub*), pSR105 (*prp24-RRM3sub/RRM4sub**), pSR151 (*prp24Δ10-RRM1sub*) and pMR59 (*PRP24*). Counterselection against pMR58 (*PRP24-CEN-URA3*) on 5-FOA left cells dependent on the mutant *prp24* plasmids for survival. * Sequencing revealed the absence of the RRM4sub mutation in this plasmid. It was remade (pMR55) by site-directed mutagenesis of pSR53 and turned out to be inviable in the GLS618 background. ‡ Cells for *lsm7Δ* with pSR105 growing on 5-FOA were tested for growth on SD-Ura and grew; on repeat *lsm7Δ* pSR105 was inviable.

background, *prp24-RRM1sub*, *prp24-RRM3sub*, *prp24-RRM4sub*, *prp24-RRM1sub/RRM4sub*, *prp24-RRM3sub/RRM4sub*, *prp24Δ10*, *prp24Δ10/RRM1sub* (Rader *et al.*, 2002) all supported growth on 5-FOA (Fig. 6.2B). In contrast only three of these *prp24* mutants allowed growth in the MRY15 (*lsm6Δ*) and MRY16 (*lsm7Δ*) backgrounds: *prp24-RRM1sub*, *prp24Δ10* and the combined *prp24Δ10/RRM1sub* mutation (Fig. 6.2C and D; Fig. 6.3). In all cases growth was reduced for these *prp24* mutants compared to *PRP24*, with the exception of *prp24Δ10* in *lsm7Δ* (Fig. 6.3).

In addition to the triple alanine mutations in each of the RRM domains the effects of mutations in single residues in RRM2 and RRM3 were tested. These *prp24-R158S* and *prp24-F257I* mutations were shown to suppress cold-sensitive growth of a U6-A62G mutation, which supposedly causes hyperstabilisation of the U6 RNA 3' stem-loop structure, thus inhibiting U4/U6 formation (Fortner *et al.*, 1994; Vidaver *et al.*, 1999). The heat-sensitive phenotype caused by both *prp24* mutations as previously described (Vidaver *et al.*, 1999) was shown to be more severe for *prp24-R158S* than for *prp24-F257I* (Fig. 6.3). Like the *prp24-RRM2sub* and *RRM3sub* mutations these point mutations were also synthetic lethal with *lsm6Δ* and *lsm7Δ*. Kwan and Brow (2005) later showed that the *prp24-RRM1sub* mutations only affect residues that are poorly conserved between Prp24p homologues, suggesting that the residues mutated in RRM1sub are not essential for RNA-binding. A point mutation, F87A, in a highly conserved residue of RRM1 was shown to be heat-sensitive lethal (Kwan *et al.*, 2005). This mutation also turned out to be synthetic lethal with *lsm6Δ* and *lsm7Δ* (this work; data not shown).

All plasmids were sequenced to confirm the presence of the correct *prp24* mutations and absence of any others. Contrary to published data, plasmid pSR105 (RRM3/4sub) was shown to contain the *prp24-RRM3sub* mutation only and not the *prp24-RRM4sub* mutation. To test whether the double mutation *prp24-RRM3sub/RRM4sub* is really viable in the GLS618 (*LSM+*) background, pSR53 (RRM3sub) was used as a template for site directed mutagenesis creating pMR55 (RRM3/4sub). This double mutation was found to be lethal in *LSM+*, *lsm6Δ* and *lsm7Δ* backgrounds, unlike what was reported by Rader and Guthrie (2002).

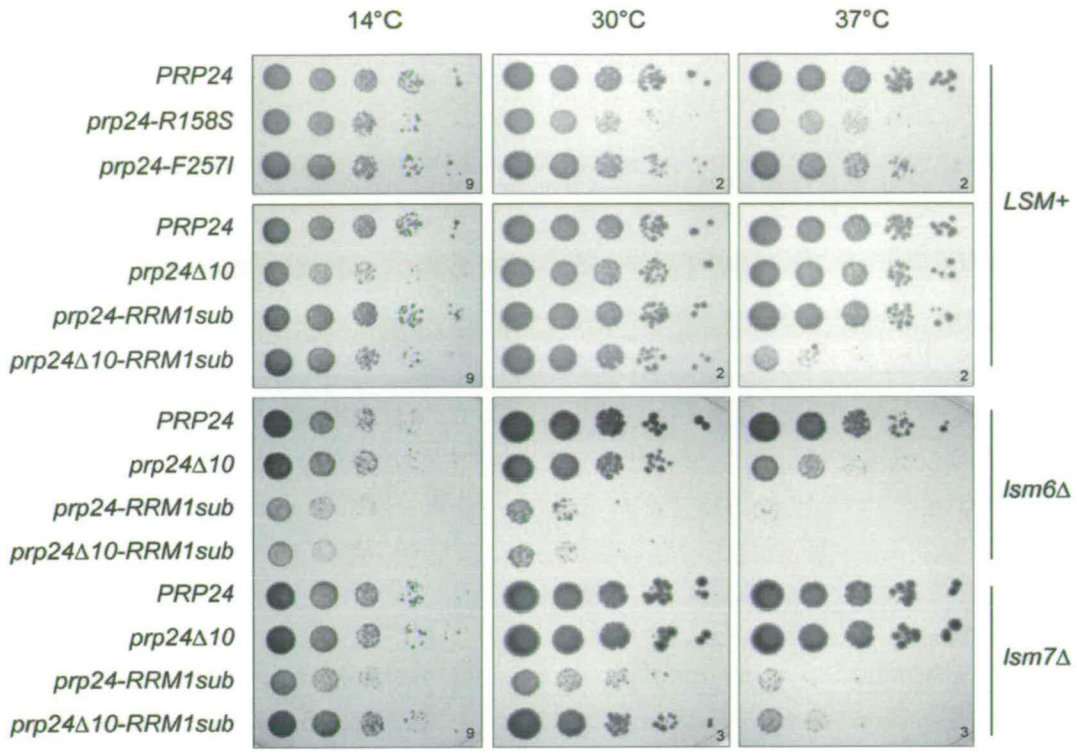


Fig. 6.3 Growth phenotypes of *prp24* mutants in *LSM+* (GLS618), *lsm6Δ* (MRY15) and *lsm7Δ* (MRY16). The following plasmids were shuffled into GLS618, MRY15 and MRY16 as in Fig. 6.1A: pMR59 (*PRP24*), pMR65 (*prp24-R158S*), pMR66 (*prp24-F257I*), pMR63 (*prp24Δ10*), pSR86 (*prp24-RRM1sub*) and pSR151 (*prp24Δ10-RRM1sub*). Resulting strains were grown overnight in YPDA and ten-fold serial dilutions of equal OD units of each culture were spotted on YPDA and grown at 14°C, 30°C and 37°C, for 2 days (2), for 3 days (3) or for 9 days (9).

In summary (Table 6.1), mutations in conserved residues of all four RRMs of Prp24p cause synthetic lethality with deletion of *LSM6* or *LSM7*. The only *prp24* mutations that allow growth in combination with *lsm6Δ* or *lsm7Δ* are *prp24-RRM1sub*, which does not affect conserved Prp24p residues, *prp24Δ10*, which affects the Prp24-Lsm2-8p interaction and the combined *prp24Δ10-RRM1sub* mutation. Interestingly, *prp24Δ10* confers cold-sensitivity on an *LSM+* strain, but has no effect on *lsm7Δ* at any temperature and leads to heat-sensitive instead of cold-sensitive growth in *lsm6Δ* (Fig. 6.3). In addition, *prp24Δ10* seems to alleviate part of the *prp24-RRM1sub* slow growth in the *lsm7Δ* background. *Prp24-RRM1sub* leads to slow growth of *lsm7Δ* at

all temperatures and is strongly heat-sensitive, whereas *prp24Δ10-RRM1sub* grows the same at 14 and 30°C as *PRP24* in *lsm7Δ* without significantly rescuing heat-sensitive growth.

Table 6.1 Overview of effects of *prp24* mutations on growth

	LSM+	*lsm6Δ*	*lsm7Δ*
PRP24	viable	Ts (weak)	Ts (weak)
RRM1sub	viable [a/c]	Cs / Ts	Cs / Ts
F87A	Ts [b]	†	†
RRM2sub	† [a/d]	n.t.	n.t.
R158S	Ts [d]	†	†
RRM3sub	Ts [a/c]	†	†
F257I	Ts [c]	†	†
RRM4sub	Ts [a]	†	†
RRM1/4sub	Ts [a]	†	†
RRM3/4sub	† [a*]	†	†
Δ10	Cs [a]	Ts	Ts
Δ10/RRM1sub	Cs / Ts [a**]	Cs / Ts	Cs / Ts [***]

† lethal; n.t. = not tested; Ts = heat sensitive; Cs = cold sensitive
[a] Rader and Guthrie (2002); this report
[b] Kwan and Brow (2005); this report
[c] Vidaver *et al.* (1999); this report
[*] Incorrect in Rader and Guthrie (2002); I've shown here that this combination is lethal
[**] As shown here and in Rader and Guthrie (2002), the RRM1sub mutation slightly alleviates Cs slow growth for the *prp24Δ10* mutation (this was not noted by R&G)
[***] In the *lsm7Δ* background, the *prp24Δ10* mutation alleviates slow growth caused by the RRM1sub mutation at all temperatures (see Fig. 6.3)

6.3 Investigating physical interactions between Prp24p and Lsm proteins *in vitro*

Interactions between Prp24p and Lsm2p, Lsm5p, Lsm6p, Lsm7p and Lsm8p were initially shown to exist by yeast two-hybrid analysis (Fromont-Racine *et al.*, 2000). These interactions were confirmed and extended to include Lsm4p by Rader and

Guthrie (2002). In addition Rader and Guthrie (2002) showed that deletion of the SNFFL box from the C-terminus of Prp24p abolished these interactions with the exception of the interaction with Lsm6p. Although Lsm3p showed an interaction with wild-type Prp24p that was not above that for the Lsm1p negative control as measured in a β-galactosidase assay, its interaction did significantly decrease in the absence of the SNFFL box. Interactions of Prp24p with Lsm5p, Lsm7p and Lsm8p were measured as the strongest in the same assay, suggesting that these may be direct interactions with Prp24p. An attempt was made by Page (2002) to investigate interactions of individual *in vitro* translated Lsm proteins with recombinant Prp24p. Page showed co-immunoprecipitation of *in vitro* translated Lsm3p, Lsm4p, Lsm5p, Lsm6p and Lsm8p (not Lsm1p, Lsm2p and Lsm7p) with His-Prp24p. However, similar repeat experiments by Silvia Galardi (this lab; unpublished data) did not find any significant interactions between Prp24p and Lsm polypeptides that were above a background of non-specific binding to beads. To attempt to elucidate these interactions I used affinity purified recombinant His-Lsm proteins in co-immunoprecipitation experiments with His-HA or His-Myc tagged Prp24p.

The level of non-specific pull-down of Lsm proteins was again high and was eventually partially resolved by addition of 0.1% Triton-X100 to the immunoprecipitation and wash buffers. Almost all Prp24p added was pulled down. In contrast, no Lsm6p or Lsm8p was pulled down (Fig. 6.4A, lanes 19-21, 25-27) and only small amounts of Lsm1p, Lsm2p, Lsm4p, Lsm5p and Lsm7p (Fig. 6.4A, lanes 4-9, 13-18, 22-24). However, background was still high for both Lsm1p and Lsm7p (lanes 6, 24). The only protein that is significantly pulled down with no background is Lsm3p (lanes 10-12). Lsm3p forms very stable homomultimeric complexes (Collins *et al.*, 2003; this work, Fig. 3.4C), which sometimes fail to resolve into monomers during SDS-PAGE. Insets in lane 10 indicate a probable Lsm3p 6- or 7-mer for the input sample. To test if Prp24p has a general affinity for multimeric Sm-like complexes, recombinant *E. coli* His-Myc-Hfq was used as a control (lanes 1-3). Hfq forms a similarly stable homohexameric complex (see inset), but does not interact with Prp24p under these conditions.

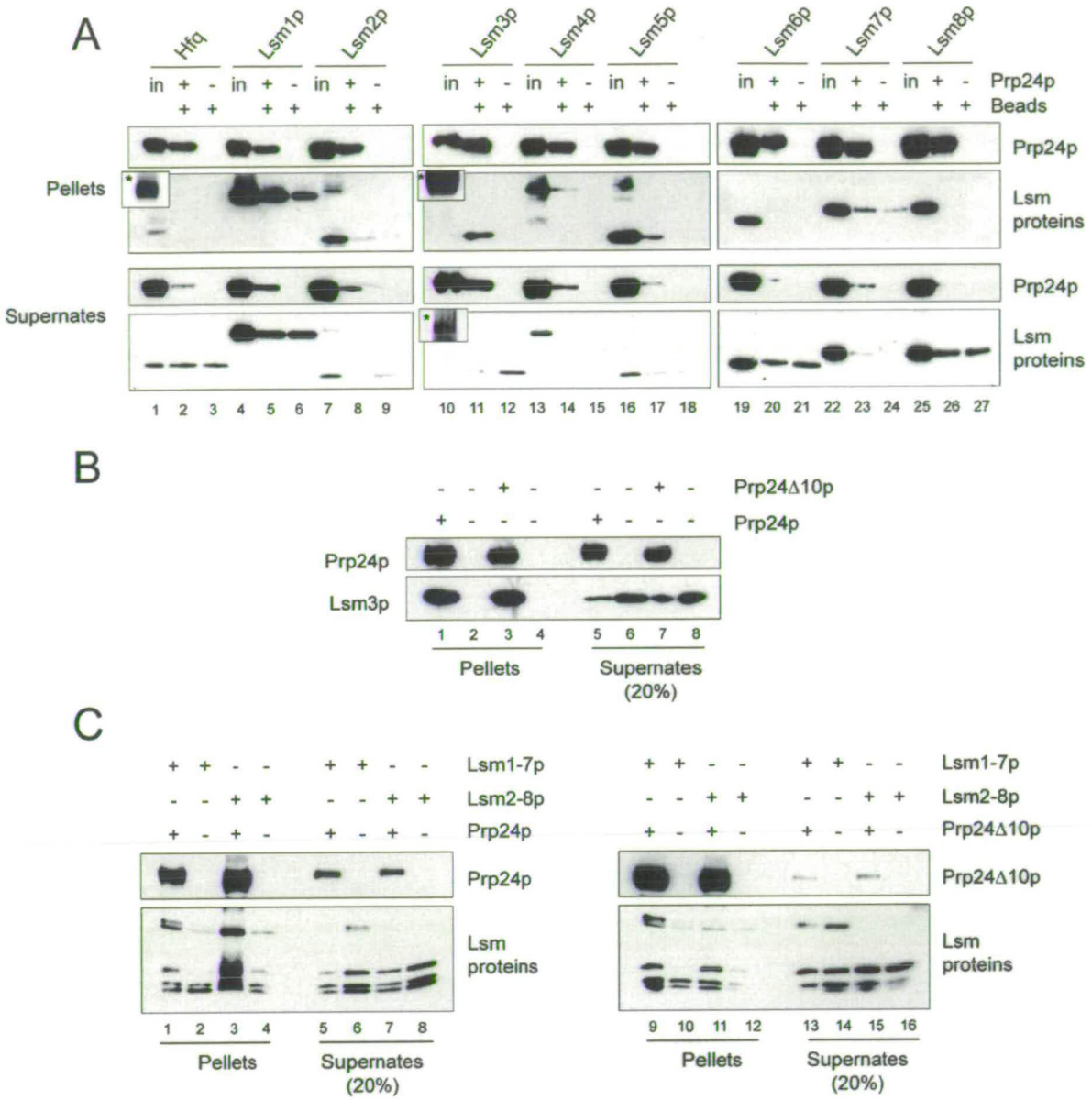


Fig. 6.4 Immunoprecipitation of recombinant Prp24p and Lsm proteins. **(A)** Recombinant His-HA-Prp24p was used in combination with affinity-purified recombinant His-Myc-Hfq and His-Myc-Lsm1p; His-Myc-Prp24p was used in combination with His-Lsm2p, -Lsm3p, -Lsm4p, -Lsm5p, -Lsm6p, -Lsm7p and His-HA-Lsm8p; **(B)** His-Myc-Prp24p or -Prp24Δ10p was used with His-Lsm3p; **(C)** His-HA-Prp24p or -Prp24Δ10p was used with Lsm1-7p (Lsm1 His-Myc tagged, all others His tagged) and His-Myc-Prp24p or -Prp24Δ10p was used with Lsm2-8p (Lsm8 His-HA tagged, all others His tagged). Approximately 0.5 μg of each recombinant protein was used in each reaction; HA-Prp24p was immunoprecipitated with α-HA agarose beads and Myc-Prp24p was immunoprecipitated with α-Myc agarose beads. Inputs (in), pull downs (pellets) and supernates were separated by SDS-PAGE and Western blots were probed with α-His-HRP antibody. * insets show bands on the same Western blots representing Hfq and Lsm3p multimers

Some of the Lsm proteins (e.g. Lsm4p, lanes 13-15) are lost from both pellets and supernates, suggesting that they stick to the beads during the initial incubation, but are subsequently lost during the washes. To test if the Prp24p-Lsm3p pull-down is SNFFL box specific, the pull-down was repeated with Prp24p and Prp24Δ10p (Fig. 6.4B). Lsm3p was specifically pulled down by both Prp24p and Prp24Δ10p, suggesting that the interaction does not depend on the SNFFL box. To test if Prp24p has specificity for Lsm2-8p over Lsm1-7p, *in vitro* assembled recombinant Lsm complexes were used in a pull-down with Prp24p (Fig. 6.4C, lanes 1-8). Both complexes showed significant background binding to the beads, but were pulled down more strongly in the presence of Prp24p. To test if this interaction is SNFFL box specific, the pull-down was repeated with Prp24Δ10p (lanes 9-16). The pull-down seemed slightly reduced for Lsm2-8p, but not for Lsm1-7p.

In summary, non-specific binding due to stickiness of Lsm proteins under immunoprecipitation conditions is a significant problem. Interactions found by yeast two-hybrid analysis were not confirmed by pull-down experiments. The strongest interaction found by co-immunoprecipitation between Prp24p and Lsm3p was the weakest by yeast two-hybrid and was independent of the SNFFL box. Although the results presented here are not conclusive, Prp24p may not differentiate between Lsm1-7p and Lsm2-8p, which would mean that the specificity of the interaction *in vivo* may depend entirely on their different cellular localisations.

6.4 Discussion

In this chapter I have shown that all *prp24* mutations, with the exception of *RRM1sub* and *prp24Δ10*, are synthetic lethal with deletions of *LSM6* and *LSM7*. This strongly suggests that both Prp24p and Lsm2-8p work together to anneal the U4 and U6 snRNAs. Genetic interactions between mutations in the Prp24p RRMs, mutations in U4 and U6 snRNAs, and their (suggested) effects on U4/U6 levels were previously used to propose distinct roles for the individual RRMs in either annealing or unwinding of the U4/U6 duplex (Fig. 6.5). The data presented here suggest that, although the individual RRMs may have different effects on the U6 structure and

thus have different effects on U4/U6 association, all four RRMs ultimately work together with Lsm2-8p to promote U4/U6 di-snRNP assembly.

The lack of a synthetic phenotype between *prp24-RRM3sub* and *prp24-RRM4sub* was the basis of the proposal that the functions of RRM3 and RRM4 may be antagonistic (Rader *et al.*, 2002). However, I have now shown that this observation was due to the absence of the double mutation and therefore incorrect. In fact, the double mutation is lethal suggesting that RRM3 and RRM4 instead may have partially redundant functions in U4/U6 annealing. This fits with the later proposal by Kwan and Brow (2005) that RRM3 and RRM4 may primarily promote U4/U6 association, possibly by binding to U4 RNA and/or U4 snRNP proteins. The Kwan & Brow model was based on *in vitro* RNA binding assays with truncated Prp24p mutants that showed that high affinity U6 RNA binding is due to RRM1 and not RRM4, and on the existence of synthetic lethal interaction of mutations in RRM1 and 2, but not RRM3 with U6 telestem mutations that promote U4/U6 association (Kwan *et al.*, 2005).

Rader and Guthrie (2002) proposed a partially redundant function for RRM1 and the SNFFL box in U4/U6 association based on the strong synthetic heat-sensitive phenotype between *prp24-RRM1sub* and *prp24Δ10*. The role of RRM1 in U4/U6 association was however disputed by Kwan and Brow (2005), who instead proposed a role for RRM1 in U4/U6 dissociation, based on synthetic lethal interactions between the *prp24-F87A* mutation (in RRM1) and mutations that disrupt the U6 telestem (and are thus proposed to promote U4/U6 annealing). Our finding of absence of a synthetic phenotype between *prp24-RRM1sub* and *prp24Δ10* in *lsm6Δ* and *lsm7Δ* backgrounds complicates matters even further, particularly since the double mutation even leads to better growth than the single *prp24-RRM1sub* mutation in the *lsm7Δ* background. The synthetic effect of *prp24-RRM1sub* alone with the *lsm* deletions found here (Fig. 6.3) supports a role for RRM1 in U4/U6 association. This suggests that both Prp24p-RRM1 and Lsm2-8p are needed to promote U4/U6 annealing.

A number of synthetic or suppressor effects were observed that are hard to explain, based on our current knowledge. None of these however, are in contradiction with the model proposed below. These effects include: partial suppression of slow growth of *prp24-RRM1sub lsm7Δ* by an additional *prp24Δ10* mutation; absence of cold-sensitive growth for *prp24Δ10* with *lsm6* or *lsm7Δ* compared to *prp24Δ10 LSM+*; and absence of a synthetic effect of *prp24Δ10* with *lsm7Δ* compared to a clear effect with *lsm6Δ*. Most of these involve deletion of the SNFFL-box, suggesting that this motif may have a function in U4/U6 association that is as yet unknown.

In addition to the interactions described above, there are many genetic interactions between mutations in *SNR14*, *SNR6* and *PRP24*. So far, these interactions have been used by various groups to try to assign specific functions in U4/U6 association and dissociation to the different motifs present in Prp24p (Fig. 6.5B). However, none of the models proposed so far manages to reconcile all known interactions and consequences of single mutations for U4/U6 formation. For example, the *prp24-R158S* and *prp24-F257I* mutation in RRM1 and RRM2 alleviate cold-sensitive growth of a U6-A62G mutation, a mutation which reduces U4/U6 levels. At the same time these *prp24* mutations themselves also lead to temperature-sensitive growth and a decrease in the level of U4/U6 (Vidaver *et al.*, 1999). The former is used as an argument to suggest that Prp24p-U6 interactions involving RRM2 and RRM3 stabilise an intramolecular structure in U6 RNA (the telestem) that antagonises U4/U6 basepairing. Reduced Prp24p-U6 binding in the U6-A62G mutant would therefore lead to a less stable telestem and increased U4/U6 association. No satisfactory explanation is given however, for reduced U4/U6 annealing in these *prp24* mutant strains with wild-type U6 RNA. Later, synthetic interactions between telestem destabilising mutations and mutations in RRM1 and RRM2, but absence of these interactions with mutations in RRM3 were used by Kwan and Brow (2005) to argue a role for RRM1 and 2 in telestem stabilisation and U4/U6 dissociation. However, Kwan and Brow (2005) did not explain the suppression by *prp24-F257I* (in RRM3) of the cold-sensitive growth caused by the 3' stem-loop stabilising U6-A62G mutation (Vidaver *et al.*, 1999), thus favouring a model in which RRM3 and 4 instead promote U4/U6 association.

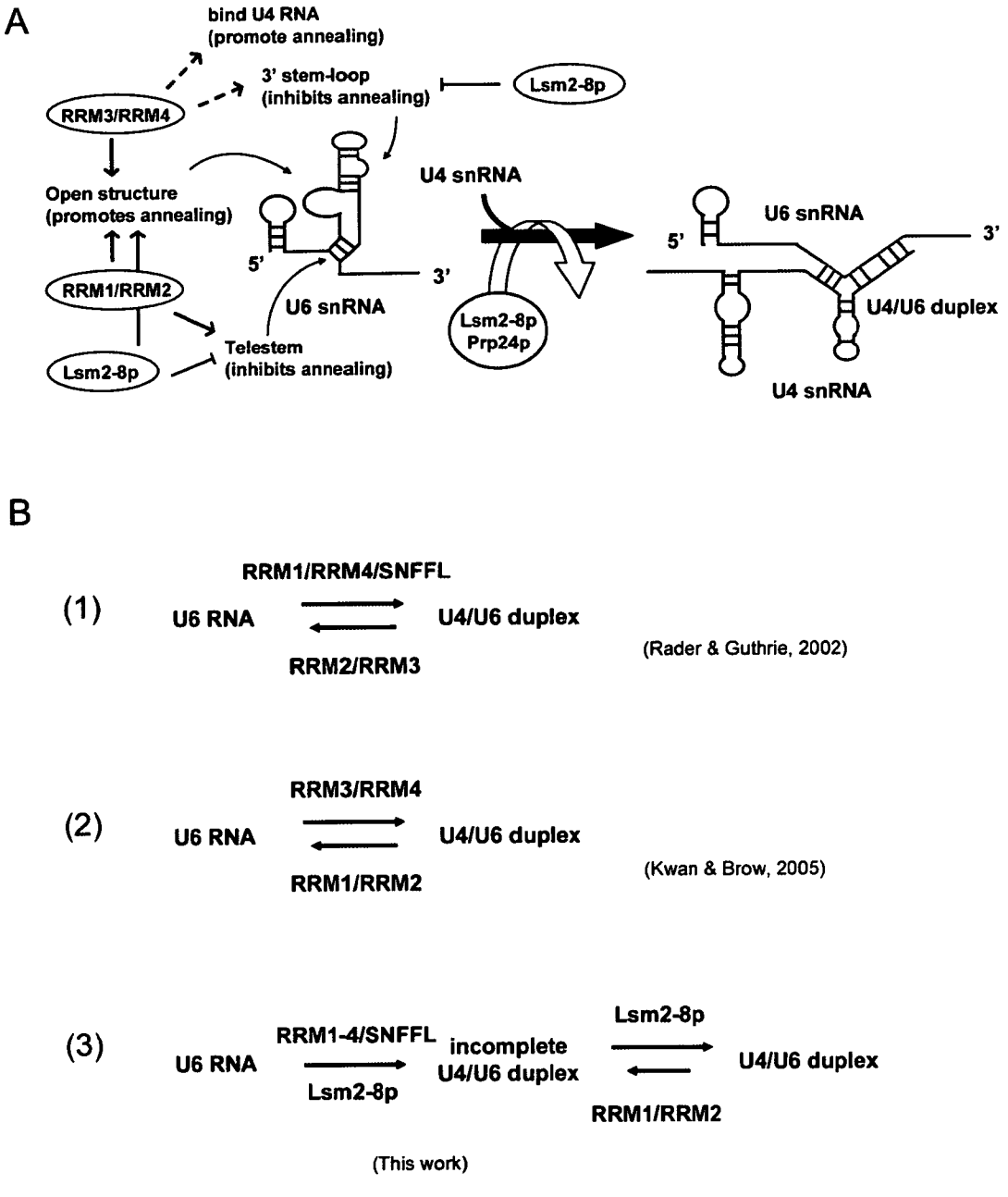


Fig. 6.5 A model for Prp24p and Lsm2-8p in U4/U6 annealing. **(A)** Schematic representation of effects Prp24p's RRM domains and Lsm2-8p have on the structure of U6 snRNA; **(B)** Proposed models about the role of the various Prp24p domains (and Lsm2-8p) in U4/U6 association and dissociation.

It may be possible to reconcile all known interactions in a model in which Prp24p has a role together with Lsm2-8p in U4/U6 di-snRNP formation (Fig. 6.5). Kwan and Brow (2005) built a strong case for a central role for RRM1 and 2 in binding of U6 RNA and a resulting stabilisation of the telestem. A more stable telestem opposes complete U4/U6 annealing. In addition SELEX (Ghetti *et al.*, 1995) and chemical probing experiments (Jandrositz *et al.*, 1995) show that Prp24p prefers to bind to a U4/U6 duplex in which stem II is not fully basepaired. Protection of the U6 RNA 3' stem-loop structure in a Prp24p-U6 RNA complex during chemical footprinting experiments suggests that Prp24p binds and possibly stabilises this structure. In contrast, this stem-loop region is less protected in the native U6 snRNP, i.e. in the presence of Lsm2-8p. In addition, the telestem is less protected in the native U6 snRNP. The overall U6 structure is more open in both the recombinant Prp24p-U6 complex and the native U6 snRNP in the regions covered by nt 29-33, 40-63 and 85-91 which are involved in extensive basepairing in the naked U6 RNA (Karaduman *et al.*, 2006). These results show that binding of Prp24p opens up the U6 RNA structure significantly, but that complete U4/U6 annealing may be inhibited by bound Prp24p, which was previously proposed by Ghetti *et al.* (1995) and Jandrositz and Guthrie (1995). Karaduman *et al.* (2006) has now shown that additional binding by Lsm2-8p opens up the U6 structure even further. Both direct effects of Lsm2-8p-U6 interactions and indirect effects of these interactions on Prp24p-U6 binding are likely to contribute. The Lsm2-8 complex could thus promote U4/U6 association both by opening up the U6 RNA structure and by destabilising Prp24p-U4/U6 binding. Kwan and Brow (2005) suggested that RRM3 and 4 may interact with U6 sequences downstream of the telestem in their *in vitro* binding experiments and may bind U4 RNA *in vivo*. The fact that mutations in RRM3 suppress the U6-A62G 3' stem-loop structure stabilising mutation may mean that RRM3 also binds and stabilises this stem-loop structure *in vivo*, but only if it is hyperstabilised by mutations in U6 RNA. Alternatively, reduced binding of Prp24p to U4/U6 due to the RRM3 mutation may aid annealing by increased Prp24p dissociation from the half-formed duplex. This could explain suppression of the U6-A62G cold-sensitive growth by mutations in RRM3 and the absence of synthetic interactions between RRM3 and telestem destabilising mutations.

In conclusion, to reconcile all existing data and explain the synthetic lethal interactions between *lsm* deletions and *prp24* mutations presented here, I would like to propose the following model (Fig. 6.5B). Proper and complete U4/U6 formation *in vivo* can only occur in the presence of both Prp24p and Lsm2-8p, and Prp24p leaves only when the di-snRNP is fully formed. This may be a quality control switch to ensure incorporation of only full length U6 RNA (to which Lsm2-8p binds) into a properly formed U4/U6 duplex. In this model all four Prp24p RRMs and the C-terminal SNFFL box cooperate with Lsm2-8p to promote formation of the U4/U6 di-snRNP, with Prp24p acting as a scaffold for proper U4/U6 alignment with the help of Lsm2-8p, while both destabilise inhibitory intramolecular structures in U6 RNA. A role for Prp24p in U4/U6 dissociation prior to spliceosome activation can not be ruled out. If it does function at this step RRM1 and 2 are likely to play a central role by stabilising the telestem (Fig. 6.5).

CHAPTER 7

A functional analysis of Lsm4p

7.1 Introduction

LSM4 of *S. cerevisiae*, also called *SDB23* or *USS1*, is an essential gene with a 564 bp ORF encoding a 187 amino acids long polypeptide (21.3 kDa). The N-terminal 92 amino acids include the Sm-domain, which is involved in protein-protein and protein-RNA interactions within the Lsm complexes. The C-terminal 95 amino acids of the protein are unstructured as predicted by 3D-PSSM (data not shown; Kelley *et al.*, 2000) and rich in asparagine (36%) and serine (17%), giving it a highly basic character (Fig. 7.1). The sequence of this C-terminal extension is poorly conserved and its function is as yet unknown. Many Lsm4p homologues from higher organisms have an abundance of arginine and glycine residues, often in the form of RG repeats, shown to be important for interactions with the SMN complex and believed to regulate snRNP assembly through symmetrical dimethylation of the arginine residues (Brahms *et al.*, 2001; Paushkin *et al.*, 2002). *S. cerevisiae* does not have a known SMN complex equivalent, providing a possible explanation for the absence of RG repeats in yeast Lsm4p. *SDB23* (suppressor of dumbbell phenotype 2, no. 3) was initially isolated in a high copy-number suppressor screen of a mutation in *DBF2*, which encodes a cell cycle protein kinase (Parkes *et al.*, 1992). The mechanistic basis of this suppressor activity is unknown. *SDB23* sequence homology to the Sm (snRNP core) proteins and subsequent demonstration of its requirement for U6 snRNA stability and pre-mRNA splicing established Uss1p (U-Six snRNP) as the first of seven U6 snRNP core proteins (Cooper *et al.*, 1995; Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999). *LSM4* was later also found in other screens: one looking for essential genes involved in cell morphology (Chun and Goebl, 1996), one looking for genes that promote *de novo* prion formation when present at high copy-number (Derkatch *et al.*, 2001), and one looking for synthetic lethality with *act1-120Ts* (Davierwala *et al.*, 2005). This poses a number of interesting questions: Does Lsm4p have functions in addition to those known for Lsm1-7p and Lsm2-8p in RNA processing? If so, does it perform these functions within the context of these Lsm complexes? What is the function of the unusual carboxy-terminal domain of yeast Lsm4p? In this chapter, I try to answer some of these questions by investigating phenotypes caused by overproduction or depletion of Lsm4p, and expression of Lsm4 C-terminal deletion mutants.

7.2 Investigating *LSM4* overexpression phenotypes

7.2.1 Overproduced Lsm4p and its C-terminus accumulate in foci

Spiller (2005; unpublished data) showed that both GFP-tagged full length Lsm4p (GFP-Lsm4; pMPSLsm4) and it C-terminus (GFP-Lsm4C; pMPSLsm4D2; Fig. 7.1) accumulate in foci when overexpressed from the *MET25* promoter. In contrast, the GFP-tagged N-terminal domain of Lsm4p (GFP-Lsm4ΔC2; pMPSLsm4D1; Fig. 7.1) failed to accumulate in foci. A repeat of these experiments confirmed these initial findings (Fig. 7.2). Extremely bright foci are present in a variable percentage of cells (10-60%) expressing GFP-Lsm4 and GFP-Lsm4C, whereas GFP-Lsm4ΔC2 localises throughout the cells, with occasional weak nuclear accumulation. This is the case both in the presence (BY4741; Fig. 7.2) and absence (MCY4-P_{GAL}-*LSM4* grown on glucose; Fig. 7.6B) of native Lsm4p. The size and number of foci increases after hypo-osmotic shock for both Lsm4p and Lsm4C, suggesting that at least some of them may be P-bodies. After osmotic shock GFP-Lsm4ΔC2 still localises throughout the cell, although very rarely (< 1% of cells) weak foci can be seen (data not shown).

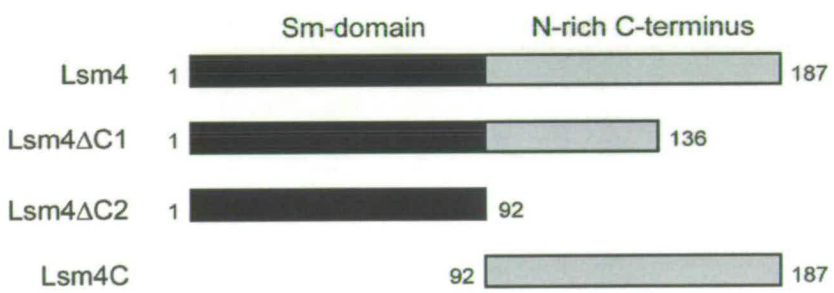


Fig. 7.1 Schematic overview of Lsm4p structure. The N-terminal domain of Lsm4p (aa 1-92) contains the Sm-domain (Sm1 and Sm2 motifs). The C-terminal extension (aa 92-187) is rich in Asn and Ser, and is predicted to be unstructured. Various Lsm4p fragments were used in this study. These includes a partial C-terminal deletion in pMC11 (covers aa 1-136, Uss1Δp; Cooper, 1995) which will be referred to as Lsm4ΔC1; a full C-terminal deletion mutant in pMC12 (covers aa 1-92, Uss1Δp*; Cooper, 1995) and pMPS4D1(covers aa 1-93, Lsm4D1; Spiller, 2005) which will be referred to as Lsm4ΔC2; and the C-terminus alone in pMPS4D2 (aa 92-187, Lsm4D2; Spiller, 2005) which will be referred to as Lsm4C.

It is unclear whether the Lsm4p foci are P-bodies, as suggested by Spiller (2005), as they are present throughout the log phase under normal growth conditions and they are often many times brighter (and bigger) than the P-bodies that are usually observed. I'd therefore like to name these foci UFO bodies, for Unidentified Lsm FOur bodies. In summary, these results indicate that the Lsm4 carboxy-terminal domain is essential as well as sufficient for aggregation in UFO bodies and possibly for localisation to P-bodies.

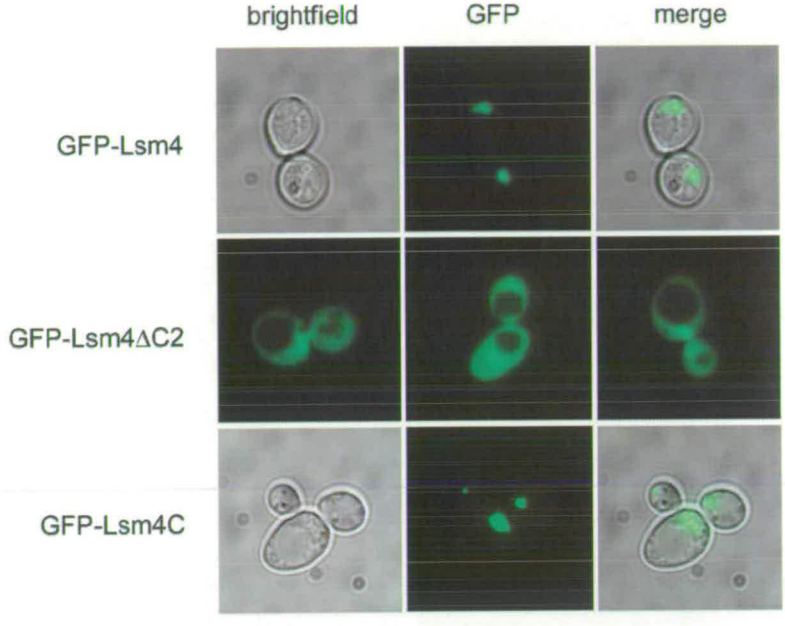


Fig. 7.2 Cellular localisation of overproduced full length Lsm4p and its N- and C-terminal domains tagged with fluorescent proteins. (**A**) GFP-Lsm4 (pMPSLsm4) and GFP-Lsm4C (pMPSlsm4D2), but not GFP-Lsm4ΔC2 (pMPSLsm4D1) accumulate in foci when overexpressed from the MET25 promoter. GFP-Lsm4D1 shows weak nuclear accumulation in some cells. Plasmids were transformed into BY4741 and cells were grown in SD-Ura -Met. Localisation was examined in live cells during log phase growth.

7.2.2 Toxic effects of *LSM4* overexpression

Lsm4p overproduction was shown to have a negative effect on cell growth and viability (Cooper *et al.*, 1995; Cooper, 1995). This effect is enhanced when combined with other mutations affecting U6 snRNP-related proteins. Cooper (1995; unpublished data) showed that overexpression of *LSM4* from the *GAL1* promoter

enhanced the temperature-sensitive phenotypes of *prp3-1*, *prp4-1* and *prp24-7* (U6 snRNP-related) splicing mutants, but not of a *prp2-1* splicing mutant (not U6 snRNP related). To extend this study, I examined the effects of *LSM4* overexpression on *lsm* mutant and deletion strains. *LSM4* overexpression (from pMR157-P_{GAL}-*LSM4*) had a negative effect on growth of all the strains that were tested. It aggravated slow growth of *lsm5Δ* (MRY27), *lsm6Δ* (MRY17) and *lsm7Δ* (MRY18) at permissive temperature and lead to slow growth of BMA38a wild-type cells at 36°C (Fig. 7.3). In addition, *LSM4* overexpression aggravated slow growth of *lsm8-1* (BP4), but only affected growth of *lsm1Δ* (AEMY25) at 14°C (data not shown). Overproduction of the C-terminal extension (Lsm4C from pMR158-P_{GAL}-*lsm4C*) had no noticeable effect on growth of any of these strains at any of the temperatures tested (Fig. 7.3; data not shown).

Thus it appears that *LSM4* overexpression is toxic to cells, in particular to those with defects in U6 snRNP-related proteins. In contrast, overproduction of the C-terminus of Lsm4p alone does not have a noticeable effect on growth. The latter indicates that aggregation of Lsm4p and Lsm4C in UFO bodies *per se* does not have a significant toxic effect on yeast, as has previously been seen for certain GFP aggregates in *C. elegans* (Link *et al.*, 2006).

7.2.3 Other Lsm proteins accumulate in UFO bodies upon Lsm4p overproduction

When *LSM4* is overexpressed from the *GAL1/10* promoter, in strain MCY4, GFP-Lsm1 (pGFP-N-Lsm1), -Lsm2 (pMPSLsm2), -Lsm6 (pGFP-N-Lsm6), -Lsm7 (pGFP-N-Lsm7) and even Lsm8-GFP (pGFP-C-Lsm8) localise to similar foci that are present during log phase growth (Fig. 7.4; data not shown). In contrast these GFP-tagged proteins show normal localisation (i.e. throughout the cell under normal growth conditions for Lsm2, -6 and -7p, to P-bodies under stress conditions for Lsm1p and to the nucleus for Lsm8p) when *LSM4* is expressed from its native promoter (Fig. 7.4A; data not shown). Strikingly, no nuclear concentration of GFP-Lsm8 is observed at all when MCY4 (P_{GAL}-*LSM4*) cells are grown in galactose-based medium, even when *LSM8* is deleted from the MCY4 genome (strain MRY65; Fig.

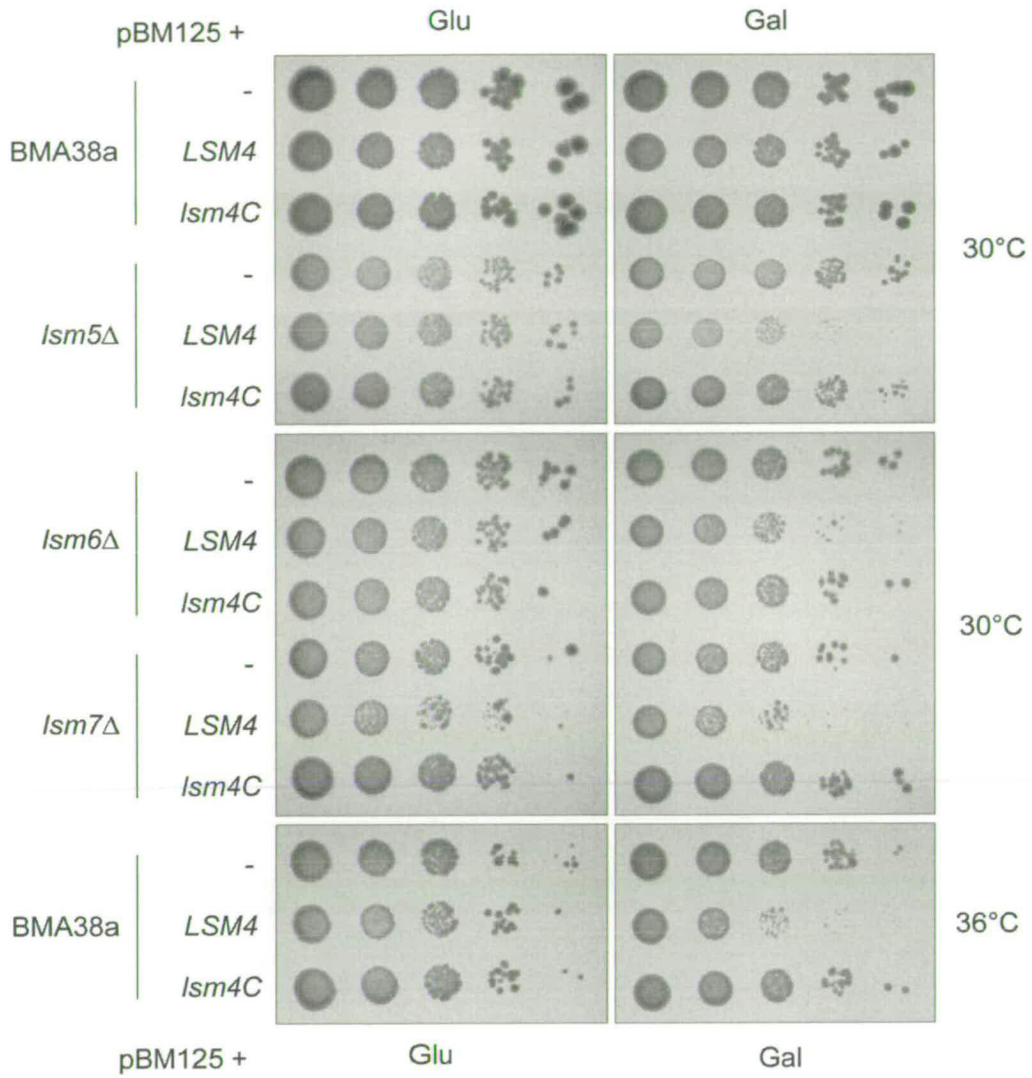


Fig. 7.3 Effects of Lsm4p or Lsm4C overproduction from the *GAL1* promoter on growth of *lsm5*Δ (MRY27), *lsm6*Δ (MRY17), *lsm7*Δ (MRY18) and wild-type (BMA38a) strains. Overexpression of *LSM4* (pMR157) aggravates slow growth of *lsm5*, *6* and *7* deletion strains at 30°C, and causes slow growth in a BMA38a wild-type strain at 36°C. Overexpression of *lsm4C* (pMR158) does not affect growth of any of these strains at any of the tested temperatures.

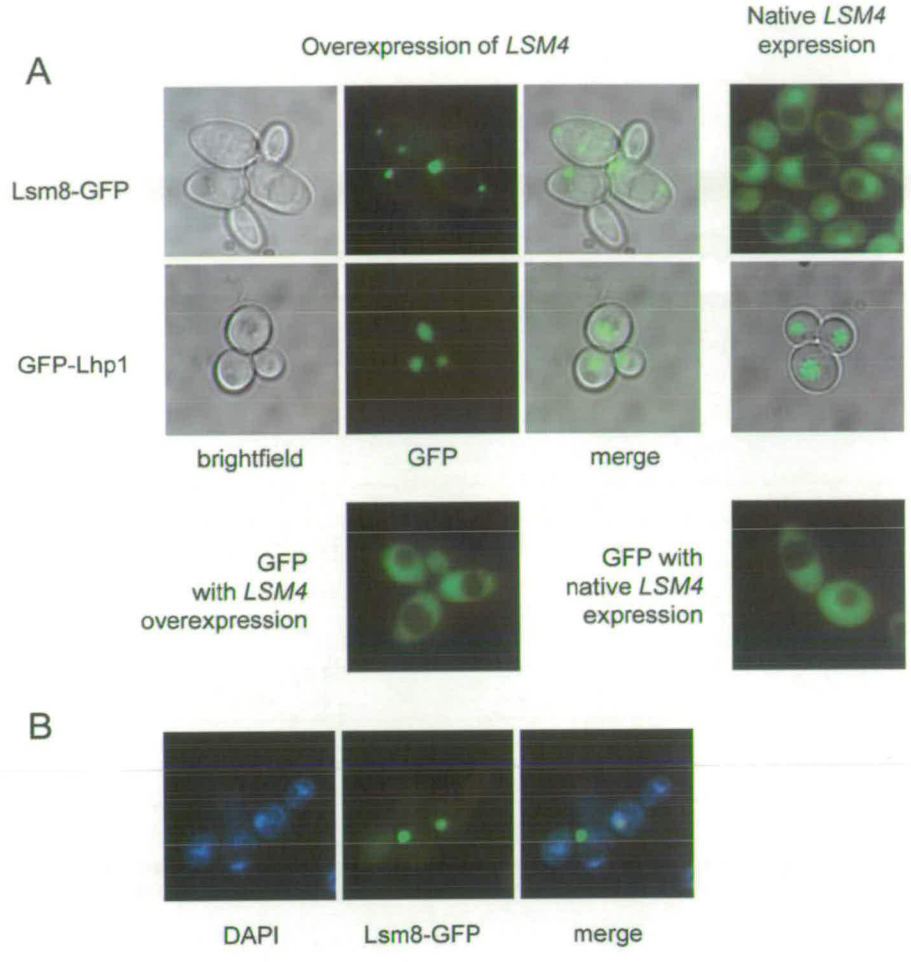


Fig. 7.4 Lsm8p aggregates when *LSM4* is overexpressed. **(A)** pMR83 (Lsm8-GFP), pMPS2 (GFP-Lhp1) or pGFP-N-FUS were transformed into MCY4 cells (P_{GAL}-*LSM4*) and grown in SDGal-Ura or in SD-Ura in the presence of pUSS1 (*LSM4* under its native promoter). Localisation of GFP-tagged proteins was investigated in live cells during log phase growth. **(B)** MRY65 (P_{GAL}-*LSM4*, *lsm8*Δ, pGFP-C-LSM8) was grown in SDGal-Ura. Log phase cells were fixed and DAPI stained; and Lsm8-GFP aggregates were shown to be cytoplasmic and in some cases to be closely associated with DAPI-stained nuclear DNA.

7.4B), so that cell viability depends on plasmid-based expression of Lsm8-GFP. As there is only a mild effect on cell viability (Fig. 7.3; data not shown), it suggests that only very low levels of Lsm8p are needed in the nucleus to support viability. This was previously shown for the *lsm8-1* mutant, which displays only a mild cold-sensitive phenotype, whereas the Lsm8 mutant protein level is strongly reduced (Pannone *et al.*, 2001). Accumulation into foci as a result of Lsm4p overproduction is specific to the Lsm proteins, as GFP alone localises throughout the cells as normal when expressed in MCY4 (*P_{GAL}-LSM4*) cells grown on galactose. Also, the exclusively nuclear Lhp1 protein tagged with GFP remains nuclear under these conditions and does not localise to foci (Fig. 7.4A). Co-localisation of Lsm8-GFP foci with DAPI-stained nuclear DNA showed that UFO bodies are mostly cytoplasmic, although regular close association with the DNA suggests possible nucleolar localisation (Fig. 7.4B). Co-localisation of overexpressed GFP-Lsm4 with tagged Nop1p and Pop1p (nucleolar proteins) confirmed localisation of some foci to the nucleolus (data not shown).

7.2.4 Lsm8p co-localises with Lsm4p, but not Lsm4C aggregates

It is likely that the negative effect of Lsm4p on growth is caused by aggregation of Lsm proteins in UFO bodies, leading to a decrease in soluble Lsm protein levels and decreased availability for essential processes. In contrast to the effect of overproduction of full length Lsm4p, overproduction of the C-terminus of Lsm4p alone does not have a noticeable effect on growth (Fig. 7.3). Indeed, Lsm8-RFP does not accumulate in foci when the Lsm4C is overproduced and still shows (weak) nuclear accumulation (Fig. 7.5A). Nor does Lsm8-RFP co-localise with GFP-Lsm4C aggregates that form upon overexpression from the *MET25* promoter, although Lsm8-RFP does fail to accumulate in the nucleus under these conditions (Fig. 7.5B). The significance of the latter is unclear, particularly as these experiments were performed in the presence of native Lsm8p, making it likely that Lsm8p is present in excess over the other Lsm proteins. In contrast, Lsm8-RFP co-localises with GFP-Lsm4 in UFO bodies, when GFP-Lsm4 is overexpressed from the *MET25* promoter (Fig. 7.5B).

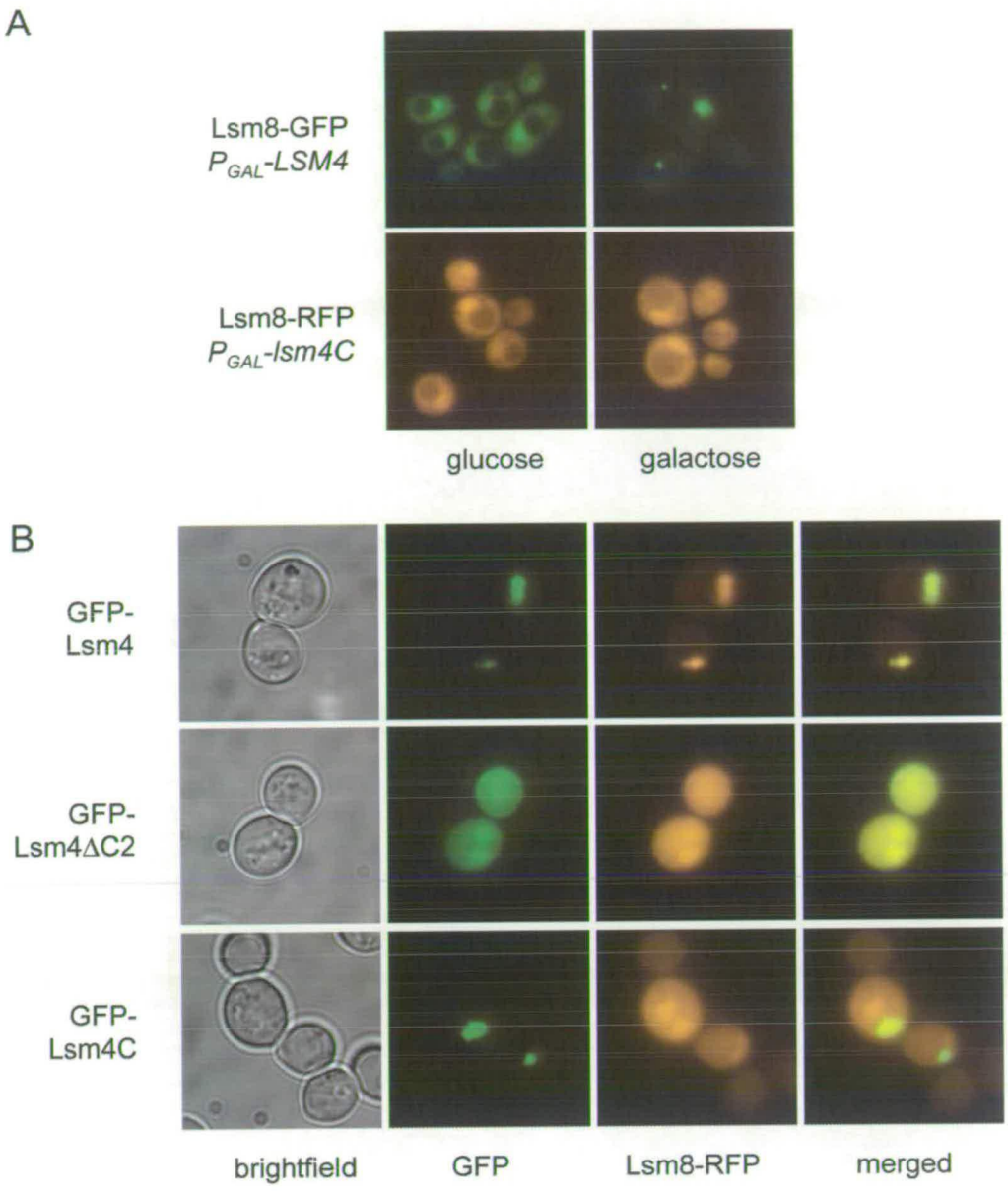


Fig. 7.5 Lsm8p localises in aggregates together with GFP-Lsm4, but not with GFP-Lsm4C. **(A)** pMR83 (Lsm8-GFP) and pMR163 (Lsm8-RFP) were transformed into MCY4 (P_{GAL}-*LSM4*) and BY4741 pMR158 (P_{GAL}-*lsm4C*) respectively. Cells were grown in SD-Ura-Met (MCY4 with pUSS1), SDGal -Ura-Met (MCY4), SD-Leu-Ura-Met or SDGal-Leu-Ura-Met (BY4741 pMR 158); **(B)** BY4741 was transformed with pMPSLsm4 (*LSM4*), -4D1 (*lsm4ΔC2*) and -4D2 (*lsm4C*) as well as pMR163 (Lsm8-RFP). Cells were grown in SD-Leu-Ura-Met. Localisation of fluorescently tagged proteins was examined in live cells during log phase growth. Note that all strains also express native *LSM8*.

7.2.5 Reduced U6 snRNA levels may be the cause of Lsm4 toxicity

Cooper *et al.*, (1995) showed that MCY4 (*P_{GAL}-LSM4*) cells accumulate *RP28* pre-mRNA when grown on galactose. When shifted to glucose splicing improved at early time points (5-10 h after shift) before deteriorating as a result of Lsm4p depletion (20 h after shift). This corresponded with an increase and subsequent decrease of U6 snRNA levels at the same time points compared to time point zero, i.e. just before shifting the cells from galactose to glucose (Cooper *et al.*, 1995). It is likely that the weak splicing defect displayed by MCY4 (*P_{GAL}-LSM4*) cells grown on galactose is caused by reduced levels of U6 snRNA due to Lsm protein aggregation with overproduced Lsm4p in UFO bodies.

In conclusion, it seems likely that overexpression of *LSM4* leads to aggregation of Lsm4p and other Lsm proteins into foci (UFO bodies). The resulting decrease in nuclear Lsm2-8p levels leads to reduced levels of U6 snRNA, which in turn affects splicing efficiency. This also explains why toxic effects of *LSM4* overexpression are more noticeable in strains that have other mutations affecting U6 snRNP proteins, as accumulative effects on U6 snRNA stability will enhance the mutant phenotypes. As the nuclear Lsm proteins play a role in many other RNA processing events, effects of their aggregation on these processes may contribute to Lsm4p toxicity as well. UFO bodies may not have biological significance and may simply be a result of prion-like aggregation of Lsm4p and associated proteins. Intriguingly, Derkatch *et al.* (2001) isolated *LSM4* in a high-copy screen for factors that are able to induce yeast prion *[PSI^+]* formation. All nine novel candidate prion genes isolated in this screen carry prion-like domains, i.e. Q-, N- or Q/N-rich domains and are themselves prone to aggregation (Derkatch *et al.*, 2004). This explains the requirement of the Lsm4p C-terminal asparagine-rich extension for aggregation of Lsm4p and aggregation of the C-terminal domain itself.

7.3 Investigating a role for Lsm4p in Lsm1-7p recruitment to P-bodies

The fact that GFP-Lsm4 as well as GFP-Lsm4C accumulate in foci when overexpressed, whereas GFP-Lsm4ΔC2 fails to aggregate even under stress conditions is suggestive of a role for the Lsm4p C-terminus in targeting Lsm1-7p to P-bodies.

7.3.1 UFO bodies are not P-bodies

As Lsm4p and Lsm4Cp aggregate in foci independent of growth conditions it is unlikely, but not impossible that UFO bodies are in fact P-bodies. To address this question Dcp2-RFP, a decapping enzyme subunit and P-body component, was expressed in the same cells that overproduced GFP-tagged Lsm4p or Lsm4ΔC2. Most foci that are present during log phase growth did not contain Dcp2-RFP, or only very small amounts (Fig. 7.6A). After hypo-osmotic shock however, the number and size of foci increased and Dcp2-RFP was seen to co-localise with most, but not all of these. This shows that Lsm4p localises to P-bodies under stress conditions, and that some UFO-bodies may even recruit decapping enzymes under these conditions. However, as not all foci contain Dcp2-RFP, in particular those present during log phase growth, there seems to be a non-P-body population of Lsm protein aggregates. Dcp2-RFP localised to foci after osmotic shock in the *lsm4ΔC2* mutant (Fig. 7.6B), showing that P-body localisation of Dcp2p does not depend on the Lsm4 C-terminus. This is in agreement with the observation by Sheth and Parker (2003) that localisation of Dhh1p to P-bodies was not noticeably affected in an *lsm1Δ* strain.

7.3.2 The Lsm4p carboxy-terminal domain is not essential for localisation to P-bodies

The above results show that deletion of the Lsm4p C-terminus prevents Lsm4ΔC2 from localising to P-bodies. As it is generally assumed that a complete Lsm1-7p complex is needed for it to localise to P-bodies (Ingelfinger *et al.*, 2002; Tharun *et al.*, 2005) the C-terminal deletion is likely to affect localisation of the entire Lsm1-7p complex. To study the effects of a C-terminal deletion of Lsm4p on

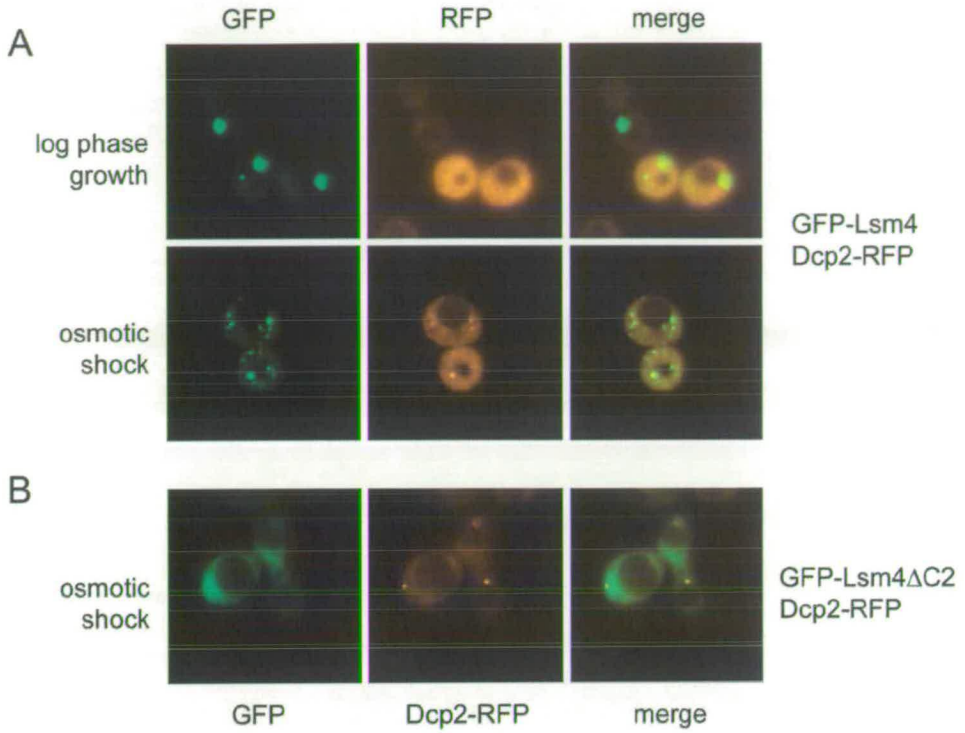


Fig. 7.6 Lsm4p forms UFO bodies during log phase growth when over produced, but also localise to P-bodies after osmotic shock. **(A)** Over-expressed GFP-Lsm4 forms strong aggregates during log phase growth that do not contain Dcp2-RFP, after osmotic shock the number of foci increases and many of these contain Dcp2-RFP; **(B)** GFP-Lsm4ΔC2 localises throughout the cell even after osmotic shock, whereas Dcp2-RFP localises to foci. BY4741 was transformed with pMPSLsm4 (GFP-Lsm4) and pRP1155 (Dcp2-RFP), MCY4 (P_{GAL}-*LSM4*) was transformed with pMPSLsm4D1 (GFP-Lsm4ΔC2) and pMR159 (Dcp2-RFP). Cells were grown in SD-Ura-Leu-Met and live cells were examined during log phase growth or 10-30 min after hypo-osmotic shock.

localisation of Lsm1p to P-bodies, MCY4 (P_{GAL}-*LSM4* strain) was transformed with pUSS1 (*LSM4*), pMC11 (*lsm4ΔC1*) or pMC12 (*lsm4ΔC2*) allowing growth on glucose with Lsm4 wild-type or mutant proteins expressed from the native *LSM4* promoter (Fig. 7.7A). These strains were in turn transformed with pGFP-N-Lsm1. Cells were grown in SD-Ura-Met and localisation of GFP-Lsm1 was examined during log phase growth and after hypo-osmotic shock. GFP-Lsm1 localised throughout the cytoplasm under normal growth conditions and concentrated in P-bodies when cells were subjected to hypo-osmotic stress (10 - 60 min incubation in

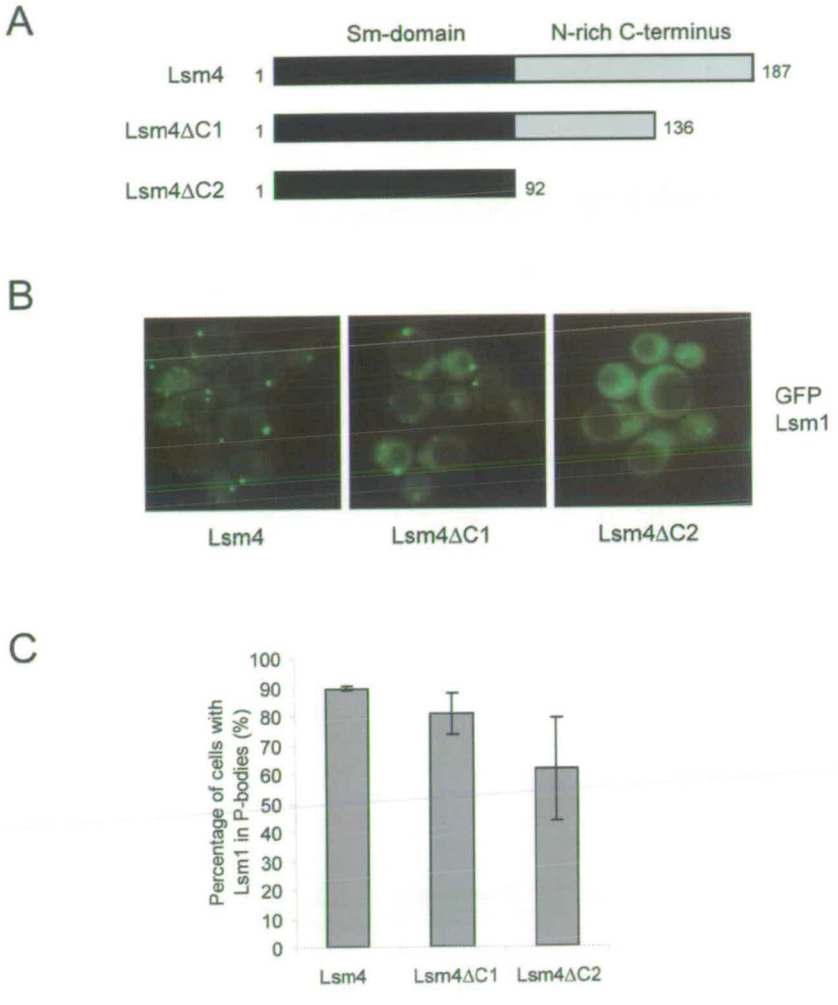


Fig. 7.7 The Lsm4 C-terminus is not essential for localisation of Lsm1p to P-bodies, but its absence leads to reduced localisation to P-bodies. **(A)** Schematic representation of Lsm4 deletion mutants; **(B)** GFP-Lsm1 localises to P-bodies after osmotic shock, but this localisation is reduced in the deletion mutants; **(C)** Quantification of the number of cells that show foci after osmotic shock. MCY4 (P_{GAL}-*LSM4*) was transformed with pUSS1 (*LSM4*), pMC11 (*lsm4ΔC1*) or pMC12 (*lsm4ΔC2*) and pGFP-N-Lsm1. Cells were grown in SD-Ura-Met and live cells were examined after hypo-osmotic shock. The percentage of cells in which foci were observed was calculated based on counts in three biological replicas (n = 200 per count).

water; Fig. 7.7B). Accumulation of Lsm1p in P-bodies in *lsm4ΔC2* shows that the Lsm4p carboxy-terminal domain is not essential as a P-body targeting signal for the Lsm1-7 complex. There did, however, seem to be a significant reduction in the intensity of the foci in the *lsm4ΔC2* strain compared to the *LSM4* strain. In addition the number of cells showing P-bodies seemed reduced. To quantify this, the number of cells showing P-bodies after hypo-osmotic shock was counted for all three strains. This was done in triplicate, counting 200 cells per strain per experiment. The majority of *LSM4* cells displayed foci (89 ± 1%), as did the *lsm4ΔC1* cells (81 ± 7%). In contrast, the number of *lsm4ΔC2* cells showing foci was significantly reduced (61 ± 18%; Fig. 7.7C).

7.3.3 Efficiency of Lsm protein recruitment to P-bodies is reduced in the absence of the Lsm4p C-terminus

The effect of the Lsm4p C-terminal deletion was more severe on P-body localisation of GFP-Lsm4ΔC2 (Fig. 7.2) than on GFP-Lsm1 (Fig. 7.7B). To rule out possible leaky *LSM4* expression from the genomic *GAL1/10* promoter an *lsm4* deletion strain with a *CEN-URA3-*P_{GAL}*-LSM4* plasmid (pMR157) was created (pMR68). This strain was transformed with pUSS1 (*LSM4*), pMC11 (*lsm4ΔC1*) or pMC12 (*lsm4ΔC2*; all three with the *LSM4* promoter). Counter-selection against pMR157 on 5-FOA medium left these centromeric plasmids as the only source of Lsm4 expression. The resulting strains (MRY71, 72 and 73) were transformed with pGFP-N-Lsm1 and GFP-Lsm1 localisation was investigated as under 7.3.2.

The effects on GFP-Lsm1 localisation were identical to those seen in MCY4 (P_{GAL}-*LSM4;* compare Fig. 7.7B and 7.8A). Upon closer investigation, there seemed to be a delay in the *lsm4ΔC* mutants in terms of GFP-Lsm1 localising to foci after hypo-osmotic shock. To quantify this effect, the number of cells that displayed P-bodies 5 min and 1 h after hypo-osmotic shock was counted in the *LSM4* and *lsm4ΔC2* strains. There was indeed a significant effect of the Lsm4p C-terminal deletion on the time after which GFP-Lsm1 could be seen to localise to P-bodies. Most *LSM4* cells displayed foci after 5 min (85%), with only a minor increase (to 90%) after 1 h. In

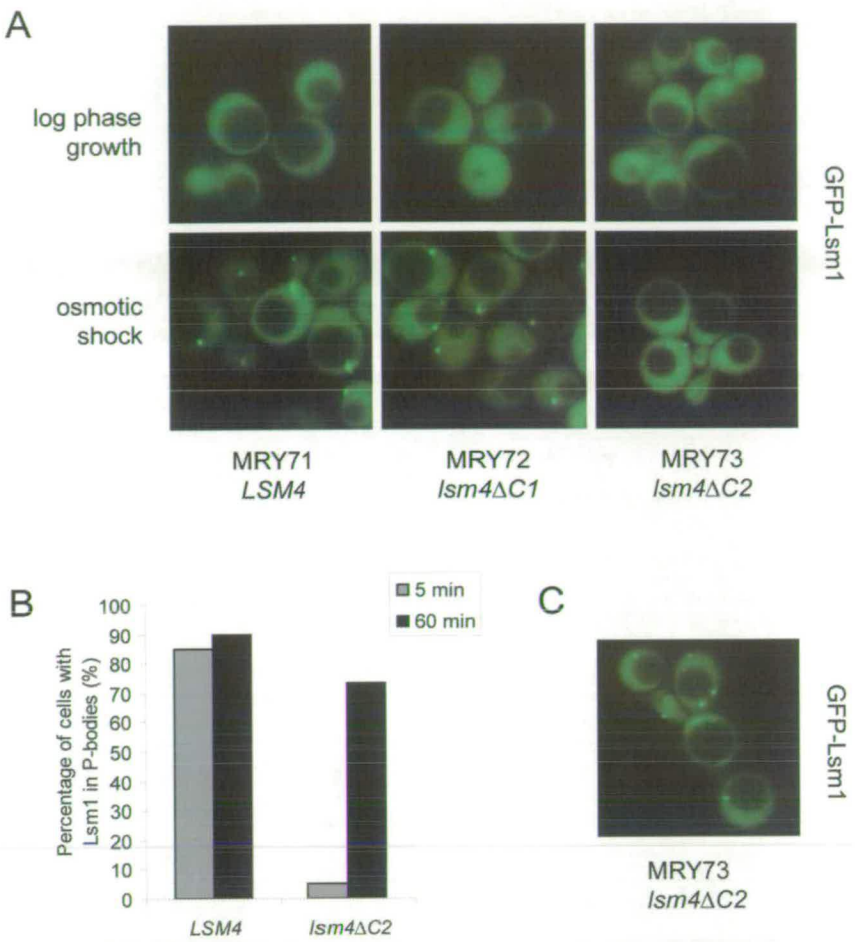


Fig. 7.8 GFP-Lsm1 is delayed in its localisation to P-bodies after osmotic shock in the absence of the Lsm4 C-terminal domain. **(A)** GFP-Lsm1 localises throughout the cell during log phase growth and localises to P-bodies 1h after osmotic shock, although at reduced levels for lsm4ΔC2; **(B)** The percentage of cells showing foci almost reaches its maximum 5 min after osmotic shock for *LSM4*, whereas less than 5% of *lsm4ΔC2* cells show foci. **(C)** Not all *lsm4ΔC2* cells show weaker foci than *LSM4*, but images in (A) are representative of the whole population. MRY71 (*LSM4*) , MRY72 (*lsm4ΔC1*) and MRY73 (*lsm4ΔC2*) were transformed with pGFP-N-Lsm1; cells were grown in SD-Ura-Met and localisation was examined during log phase growth or 5 min or 60 min after osmotic shock. Counts were performed once (n = 100).

contrast, only 4% of *lsm4ΔC2* cells displayed foci after 5 min, with an increase to 73% after 1 h.

To investigate whether the Lsm4p C-terminal deletion has the same effect on other Lsm1-7 complex members, similar experiments were carried out with GFP-Lsm2 and GFP-Lsm6. During log phase growth both localised throughout the cell, with some nuclear accumulation apparent for GFP-Lsm6 in both MRY71 (*LSM4*) and MRY73 (*lsm4ΔC2*; Fig. 7.9A). Under stress conditions (hypo-osmotic shock) both GFP-Lsm2 and GFP-Lsm6 accumulated in cytoplasmic foci, with a reduction in intensity for both in MRY73 (Fig. 7.9A). The GFP-Lsm2 foci were less pronounced than those for GFP-Lsm6, which together with the absence of nuclear accumulation for GFP-Lsm2 may indicate reduced complex formation for this tagged Lsm2p. In addition, there was a similar delay in the observed appearance of foci after hypo-osmotic shock (Fig. 7.9B). Taken together these data strongly suggest that the carboxy-terminal domain of Lsm4p is not essential for Lsm1-7p localisation to P-bodies. However, reduced and delayed localisation of Lsm1-7 complex members to P-bodies in an *lsm4ΔC2* strain suggest its importance for efficient accumulation of Lsm1-7p in P-bodies

7.3.4 The GFP-tag aggravates the *lsm4ΔC2* defect

The almost complete absence of GFP-Lsm4ΔC2 from P-bodies after osmotic shock seems to contradict the only slightly reduced P-body accumulation of GFP-Lsm1, -Lsm2 and Lsm6 in the *lsm4ΔC2* strain. Upon closer inspection, GFP-Lsm4ΔC2 was shown to localise weakly to foci after hypo-osmotic shock in a small fraction of cells (< 1%), where it co-localised with Dcp2-RFP (data not shown). In addition, cells grown into late stationary phase (OD600 ≥ 4) displayed accumulation of GFP-Lsm4ΔC2 in cytoplasmic foci together with Lsm1-RFP in more than 90% of cells (data not shown). GFP-Lsm4ΔC2 must therefore be able to incorporate into the Lsm1-7 complex and thus localises to P-bodies. Reduced accumulation in P-bodies is likely to reflect negative effects of the N-terminal GFP-tag, in addition to the C-terminal deletion.

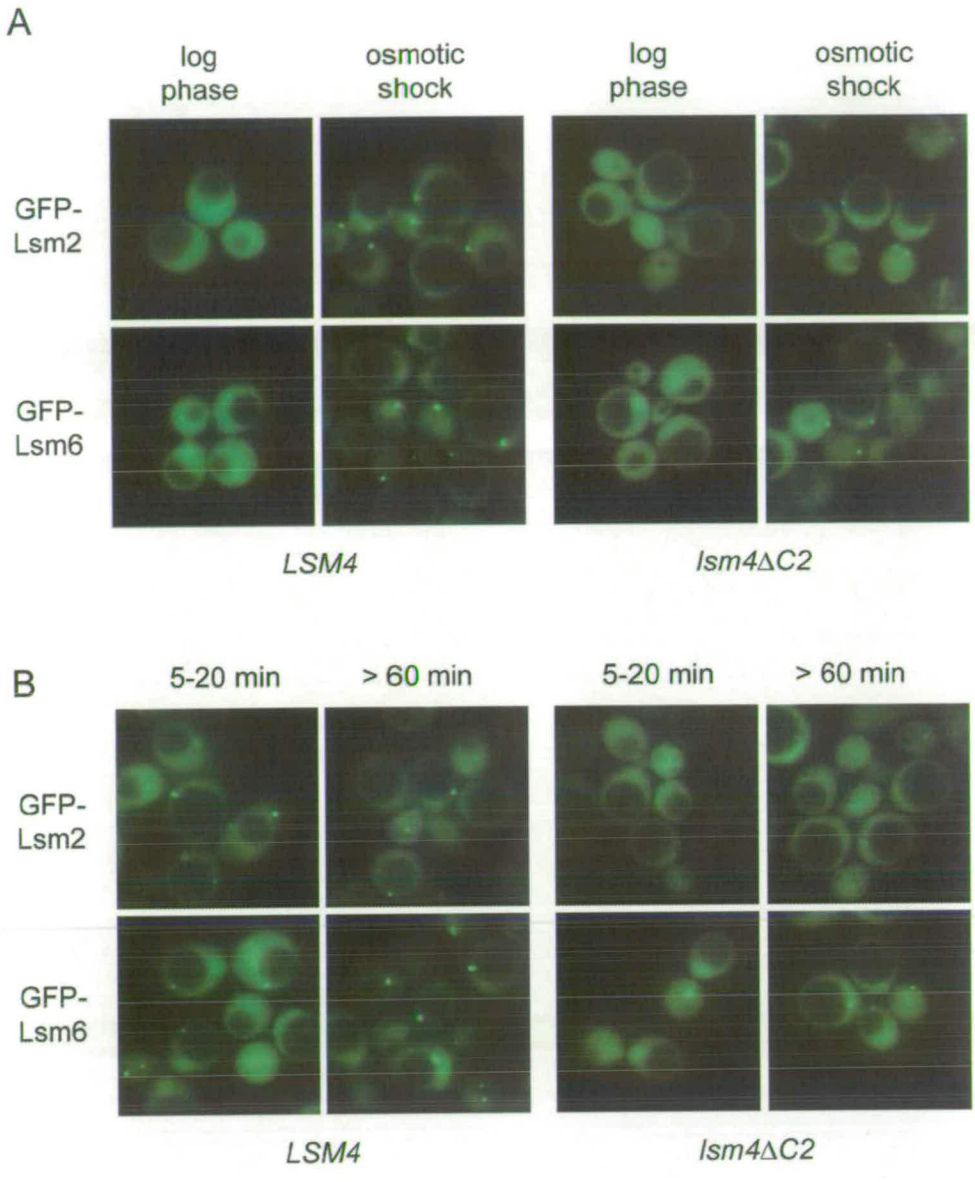


Fig. 7.9 GFP-Lsm2 and GFP-Lsm6 localise to P-bodies under stress conditions, but less efficiently in an *lsm4ΔC2* strain. **(A)** GFP-Lsm2 and GFP-Lsm6 localise throughout the cell with some nuclear accumulation for Lsm6p. Approximately 1 h after osmotic shock the majority of cells show strong foci in MRY71 (*LSM4*), but only weak foci in MRY73 (*lsm4ΔC*); **(B)** MRY71 starts showing P-body localisation of GFP-Lsm2 and -Lsm6 in the majority of cells after 5 - 20 min, whereas MRY73 only shows weak foci after 60 min or more. MRY71 and MRY73 strains were transformed with pGFP-N-Lsm2 and -Lsm6 plasmids. Cells were grown in SD-Ura-Met and live cells were examined during log phase growth or at different times after osmotic shock.

Growth was compared of *GFP-lsm4ΔC2* and *lsm4ΔC2* strains. Whereas the *lsm4ΔC2* strain was only weakly temperature sensitive, the *GFP-lsm4ΔC2* strain grew more slowly at all temperatures and was strongly temperature sensitive (Fig. 7.10). It therefore seems that the N-terminal GFP tag aggravates a growth defect caused by a C-terminal deletion in Lsm4p. It is likely that this is an indication of reduced complex formation, suggesting that the carboxy-terminal domain may also play a role in stabilising complex formation. In conclusion, the N-terminal GFP-tag in addition to the C-terminal deletion may prevent the mutant Lsm4 protein from localising efficiently to P-bodies, possibly by interfering with complex formation.

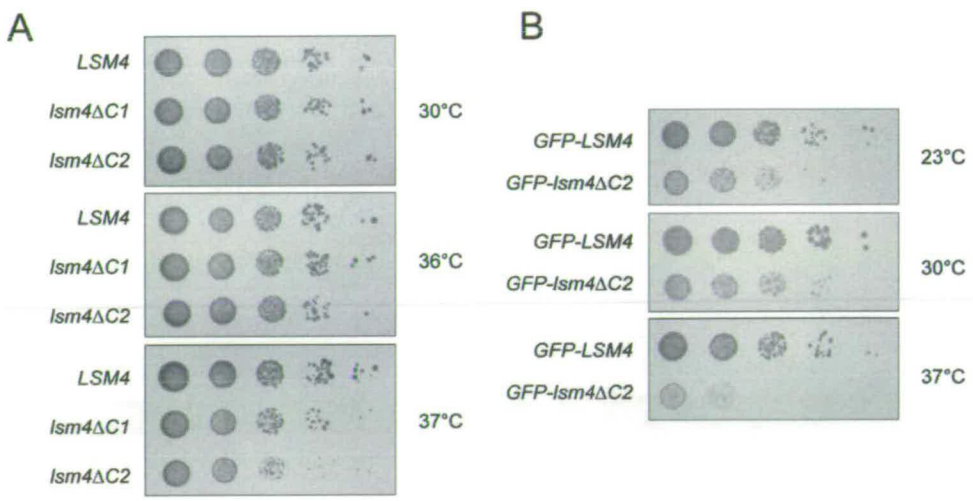


Fig. 7.10 Growth phenotypes of *lsm4* C-terminal deletion mutants. **(A)** *lsm4Δ* strains with plasmid-based *LSM4* expression from its native promoter, MRY71 (pUSS1-*LSM4*), MRY72 (pMC11-*lsm4ΔC1*) and MRY73 (pMC12-*lsm4ΔC2*) were plated in ten-fold serial dilution on YPDA and grown at 30°C, 36°C and 37°C. MRY73 shows slower growth at 37°C only; **(B)** P_{GAL}-*LSM4* strain MCY4 with GFP-tagged Lsm4p (pMPSLsm4) or Lsm4ΔC2 (pMPSLsm4D1) was plated on SD -Ura-Met and grown at 23°C, 30°C and 37°C. Cells expressing the GFP-tagged N-terminal domain of Lsm4p display a slow growth phenotype at all temperatures and are heat sensitive.

7.3.5 The Lsm4p C-terminus may contribute to efficient P-body formation

As the Lsm4 carboxy-terminal domain seems to be important for efficient recruitment of Lsm1-7p to P-bodies, it is possible that Lsm4p is also important for aggregation of other P-body components and ultimately efficient P-body formation. No significant effect was seen however on Dcp1p or Dcp2p localisation to P-bodies in an *lsm4ΔC2* strain (data not shown). A minor effect on the speed with which these decapping factors accumulate in foci after osmotic shock may exist however. This is subject to further investigations.

7.4 A potential link between Lsm4p and cytokinesis?

7.4.1 Depletion of Lsm4p leads to abnormal cell morphology

MCY4 (P_{GAL}-*LSM4*) cells show increasing numbers of abnormally shaped cells when Lsm4p is depleted by growing the cells on glucose (Fig.7.11D). Cells are often much enlarged and show (emerging) buds that do not separate, but elongate instead. MCY4 (P_{GAL}-*LSM4*) cells grown on galactose also show occasional cell shape abnormalities (Fig. 7.11B). In addition, nuclear morphology is often abnormal in cells overproducing Lsm4p (data not shown). Effects of *LSM4* overexpression on cell and nuclear shape may be explained by aggregation of Lsm4p, which effectively reduces

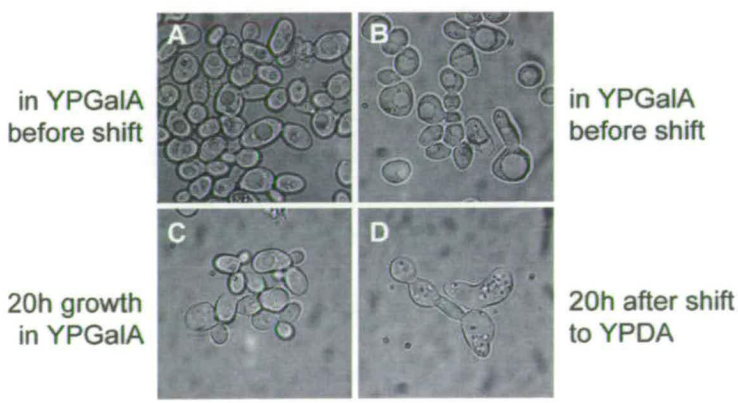


Fig. 7.11 MCY4 (P_{GAL}-*LSM4*) cells depleted of Lsm4p display abnormal cell shapes. MCY4 cells growing in YPGalA largely show normal cell morphologies **(A)** although a small percentage does show abnormal cell shapes **(B)**. After 20h of growth in YPDA more than half of the cells show abnormal cell shapes **(D)**, whereas the majority of those grown in YPGalA still look normal **(C)**.

its cellular concentration. A mechanism behind the effect of Lsm4p depletion on cell morphology is unclear, may be indirect and is no doubt complicated.

7.4.2 Increased Lsm4p production rescues a *dbf2-Ts* phenotype

LSM4 was originally isolated as *SDB23* in a high copy-number suppression screen of *dbf2-Ts* (Parkes *et al.*, 1992; Cooper *et al.*, 1995). Dbf2p is a kinase involved in exit from mitosis and subsequent cytokinesis (Hwa *et al.*, 2003; Corbett *et al.*, 2006). To show that the *dbf2* suppression activity lies within the *LSM4* ORF a 3.2 kb genomic *Bam*HI fragment was re-cloned into YEp24 and this high copy vector (pMR165) was transformed into *dbf2-1* (L119-7D) and *dbf2-2* (Amy1292) strains. The 2μ-*LSM4* plasmid was able to suppress the temperature-sensitive phenotype of the *dbf2* strains, although only at 35°C and not at 37°C as originally described (data not shown). The difference in the level of suppression between *dbf2-1* and *dbf2-2* described by Parkes and Johnston (1992) was not observed here either (data not shown).

Dbf2p kinase is activated by Cdc15p cell cycle kinase. As Cdc15p acts upstream of Dbf2p Lsm4p may similarly allow suppression of a *cdc15* mutation. However, the temperature-sensitive phenotype of a *cdc15-2* (Amy1277) strain could not be rescued by pMR168 (2μ-*LSM4*; data not shown). The 3.2 kb genomic fragment that rescues *dbf2-Ts* carries the *SWI4* promoter sequence in addition to the *LSM4* ORF. The *SWI4* gene encodes an essential transcription factor which controls gene expression at the G1/S transition of the cell cycle. *SWI4* transcription itself is cell cycle regulated, and this periodicity is crucial for cell cycle regulation (Foster *et al.*, 1993; MacKay *et al.*, 2001). The *SWI4* promoter contains a number of transcription factor binding sites that are involved in the cell cycle regulated expression of *SWI4* (Horak *et al.*, 2002). It would therefore be conceivable that the suppression of the *dbf2-Ts* phenotype is caused by the presence of additional copies of these promoter elements and not by increased levels of Lsm4p. To rule this out, the *LSM4* start ATG was mutated by site-directed mutagenesis of pMR165 to a TAG stop codon, effectively abolishing Lsm4 protein production. Significantly, the resulting plasmid (pMR168) was unable to rescue temperature-sensitive growth of *dbf2-1* and *dbf2-2* strains (data not shown). Increased Lsm4p production is therefore critical for *dbf2-Ts* suppression. It should be

noted that the level of Lsm4p overproduction from the 2μ vector is unlikely to have the same effect as overproduction from the MET25 or GAL1 promoter (i.e. aggregation). Overexpression of *LSM4* from these promoters was unable to rescue *dbf2* temperature-sensitive growth (data not shown). In contrast, the variable copy number of the 2μ vector may allow optimal levels of Lsm4p production needed for *dbf2* suppression. The mechanism for this suppression is unknown but may be linked to a role for Lsm4p in cytokinesis.

7.5 Discussion

In this chapter I have shown that Lsm4p forms aggregates when expressed at high levels. Aggregation into UFO bodies depends on the asparagine-rich prion-like C-terminal domain and overproduction of this domain alone leads to formation of similar aggregates. Lsm4p but not Lsm4C aggregates cause specific aggregation of other Lsm proteins. The resulting effect on U6 snRNA levels is the likely reason for toxicity of Lsm4p overproduction and the synthetic phenotypes with U6 snRNP mutants. I have shown that the ability of Lsm4p to aggregate is a pre-requisite for efficient localisation of Lsm1-7p to P-bodies, at least after hypo-osmotic shock. In addition, I have shown that depletion of Lsm4p leads to a cell morphology defect and, as interactions exist between *LSM4* and cell cycle factors, this may be linked to a role for Lsm4p in cytokinesis.

Although the Lsm4p C-terminal domain is poorly conserved between *S. cerevisiae* and higher eukaryotes, a function in recruitment to P-bodies may be conserved. The C-terminal extension of human LSm4 has a Q-rich stretch (Fig. 1.8), which may allow similar prion-like behaviour. Lsm4 proteins from other *Saccharomyces* species all have N-rich tails, although their exact positions are not conserved. Some of them even have extended N-rich stretches, with one (*S. bayanus*) showing a run of 24 asparagine residues interrupted by only one serine residue. Intriguingly, others show increased levels of glutamine residues, with *S. kluyveri* additionally showing a low level of RG repeats (SGD; data not shown). Lsm4p homologues from other species often, but not always, show N- and/or Q-rich C-terminal extensions (data not shown).

A function for Lsm4p in Lsm1-7p P-body localisation may therefore be conserved, at least in some organisms. Other Lsm proteins may have taken over this function by increasing the number of aggregation prone amino acid residues. This is a hypothesis that needs to be tested by comparing full sets of Lsm proteins from a range of species.

An intriguing question is how Lsm4p aggregation is prevented under normal growth conditions. Post-translational modifications, e.g. phosphorylation, of Lsm4p or other Lsm proteins, in particular its proposed neighbours Lsm1p, Lsm8p and Lsm7p may play a role. Such modifications allow the cell to respond to changes in conditions quickly and efficiently, and may allow it to regulate required levels and intracellular localisations of Lsm1-7p and Lsm2-8p. As Spiller (2005; unpublished results) showed that there is competition between these two complexes (e.g. there is a shift in localisation from the nucleus to the cytoplasm minutes after glucose starvation) this is a distinct possibility. The N-termini of Lsm1p and Lsm8p (see Chapter 9), and the C-terminus of Lsm4p itself in particular are likely targets for potential modifications.

Interestingly, it was shown that a C-terminal deletion mutant of the *Kluyveromyces lactis* Lsm4 protein (KlLsm4ΔC), which is able to support viability in the MCY4 (P_{GAL}-*LSM4*) background (Mazzoni and Falcone, 2001), led to decreased viability in the stationary phase (Mazzoni *et al.*, 2003a; Mazzoni *et al.*, 2003b). It was later shown that this was caused by increased levels of apoptosis (Mazzoni *et al.*, 2005) and that the same happened in deletion mutants of other proteins involved in mRNA turnover (*LSM1*, *DCP1*, *DCP2*), suggesting a direct connection between increased mRNA stability and apoptosis (Mazzoni *et al.*, 2003b). Mazzoni *et al.* (2003b) also suggested that the C-terminus of Lsm4p may have a specific role in mRNA decapping, as mRNA degradation, but not pre-mRNA splicing was affected in the *KlLsm4ΔC* mutant strain. My data now provide a possible explanation for the effect on mRNA decapping, i.e. reduced localisation of Lsm1-7p to P-bodies. Intriguingly, abnormal cell shapes, identical to those described here, were observed in stationary phase cells prior to apoptosis in the *KlLsm4ΔC* mutant strain.

Lsm4p may have a function in cytokinesis as several links between the two exist, not the least of which is the genetic interaction between *LSM4* and *dbf2*, which I have now conclusively shown to depend on *LSM4* expression. Dbf2p is a kinase involved in exit from mitosis and subsequent cytokinesis (Hwa *et al.*, 2003; Corbett *et al.*, 2006). It was shown that mutations in *DBF2* as well as in a number of other late cell cycle genes can cause abnormal cell morphologies, some of which are reminiscent of those observed for Lsm4p depleted cell (Chun *et al.*, 1996; Jiménez *et al.*, 1998 and references therein). It was suggested that this is due to an actin-based morphogenetic response to improper septin-ring formation, as some septin mutants show similarly elongated buds (Jiménez *et al.*, 1998). In addition, *LSM4* was one of three genes isolated in a screen looking for essential genes affecting cell morphology (Chun *et al.*, 1996) and I have now shown that depletion of Lsm4p also leads to abnormal cell morphology. The other two genes found in the Chun & Goebl study were *MOB1*, a component of the mitotic exit network (MEN) which associates with and is required for the activation and Cdc15p-dependent phosphorylation of the Dbf2p kinase (Luca *et al.*, 2001; D'Amours and Amon, 2004; Stoepel *et al.*, 2005) and *AME1*, an essential kinetochore protein associated with microtubules and spindle pole bodies (Pot *et al.*, 2005). It is possible that the Lsm4p depletion phenotype just like *dbf2* suppression by *LSM4* is caused by an effect on the mitotic exit network or on other factors involved in cell division or exit from mitosis.

Could it be that Lsm4p has other functions apart from those in mRNA decapping and splicing as part of the known Lsm complexes? Various yeast two-hybrid interactions found for *LSM4* suggest possible functional links (Fromont-Racine *et al.*, 2000; Uetz *et al.*, 2000). Most interestingly some of the genes that were found encode proteins involved in cell growth, proliferation and division. *CKB2* encodes a regulatory subunit of the CK2 protein kinase with roles in cell growth and proliferation (Glover, III, 1998). *DSE1* may participate in pathways regulating cell wall metabolism, and its deletion affects cell separation after division (Colman-Lerner *et al.*, 2001). *SLK19* encodes a kinetochore-associated protein which is a component of the FEAR (Fourteen Early Anaphase Release) regulatory network, which, like MEN, promotes Cdc14p release from the nucleolus during anaphase (Stegmeier *et al.*, 2002; Pfiz *et*

al., 2002). Cdc14p is a protein phosphatase required for mitotic exit. After release from the nucleolus it acts on key substrates effecting a decrease in CDK/B-cyclin (e.g. *CLB2*) activity necessary for cell cycle progression (Saunders, 2002). Other *LSM4* yeast two-hybrid interactors affect chromatin remodelling and maintenance. *SIN3* encodes the DNA binding subunit of the Sin3p-Rpd3p histone deacetylase complex involved in chromatin silencing at rDNA and telomeres (Silverstein and Ekwall, 2005). *STN1* encodes a telomere end-binding and capping protein and plays a key role in linking telomerase action with completion of lagging strand synthesis (Grossi *et al.*, 2004). An interaction between Lsm4p and any of these factors could underlie an as yet unknown function for the Lsm1-7 or Lsm2-8 complexes. Alternatively, Lsm4p may be part of another Sm-like complex with functions relating to cell division and cytokinesis.

It may be that the role for Lsm4p in Lsm1-7p recruitment to P-bodies is directly linked to cell division. A possible mechanism involves Lsm1-7p dependent degradation of cell cycle regulated mRNAs. This may be equivalent to the cell cycle regulated RNase MRP specific degradation of *CLB2* (cyclin) mRNA in TAM bodies (Gill *et al.*, 2006). Synthetic lethal interactions of *lsm1Δ*, *lsm6Δ*, *lsm7Δ* and *pat1Δ* with *nme1-P6*, an *MRP-Ts* mutant, suggest that there may even be a direct functional link between Lsm1-7p dependent RNA degradation and TAM bodies. Efficient *CLB2* mRNA degradation by the exosome and/or Xrn1p after endonucleolytic cleavage by RNase MRP may depend on Lsm1-7p. However, no abnormal cell shapes similar to those of Lsm4p depleted cells were observed for *lsm1Δ, lsm5Δ*, *lsm6Δ* or *lsm7Δ* cells, at least not under permissive growth conditions (data not shown). It would be interesting to see if depletion of one of the other essential Lsm proteins (*LSM2*, *LSM3* or *LSM8*) leads to abnormal cell morphology. It is possible that morphological defects caused by Lsm4p depletion are caused by inefficient degradation of cell cycle specific mRNAs through an Lsm1-7p dependent pathway. However, as a synthetic lethal phenotype exists between *LSM4* (Tet switch-off) and *act1-120Ts* (Davierwala *et al.*, 2005), a more direct effect of Lsm4p depletion on actin and cell morphology can not be ruled out.

In summary, Lsm4p plays a crucial role in efficient accumulation of Lsm1-7p in P-bodies and depletion of Lsm4p affects cell morphology. The two may or may not be linked. Regardless, effects on cell morphology may be linked to cytokinesis. Confirmation of a role for Lsm4p in cell division awaits concrete experimental evidence.

CHAPTER 8

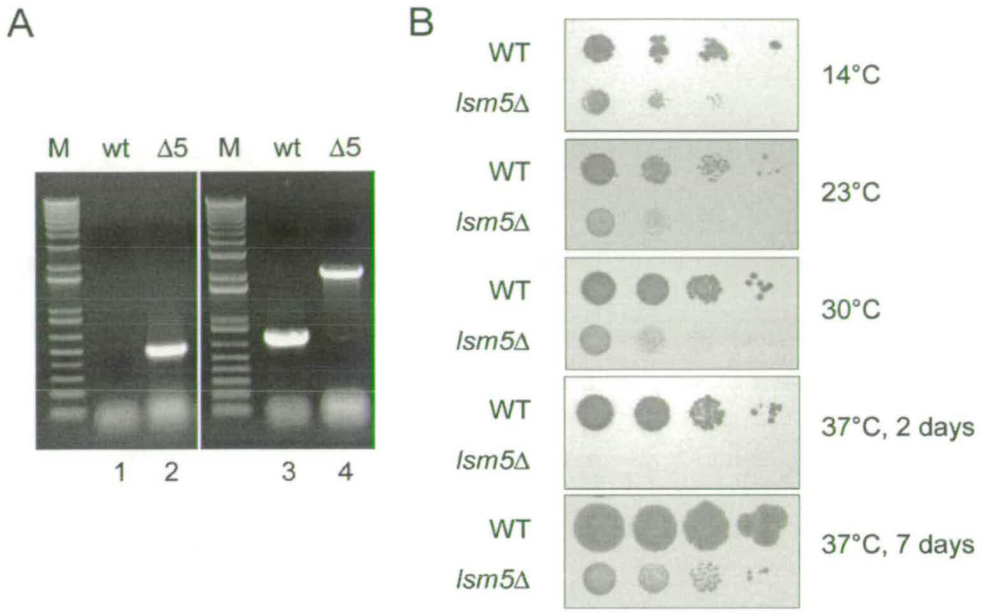
A functional analysis of Lsm5p

8.1 Introduction

LSM5 encodes a 93 aa (10.4 kDa) polypeptide and does not have significant N- or C-terminal extensions outside the Sm domain. Various reports show that *LSM1*, *LSM6* and *LSM7* are not essential for cell viability in *S. cerevisiae* (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999). However, *LSM5* was shown by Mayes *et al.* (1999) to be an essential gene, whereas others showed it to be dispensable for growth (Salgado-Garrido *et al.*, 1999; Pannone *et al.*, 2001). Also, in a genome-wide screen by Giaever *et al.* (2002) it was reported that *LSM5* is needed for cell survival and growth. Pannone *et al.* (2001) stated that, at 25°C, four spores from *LSM5/lsm5Δ* tetrads grew, whereas only two viable spores (*LSM5*) per tetrad were recovered after incubation at 30°C. Strain backgrounds, which were different for all the above studies (BMA64, BSY320, CY1 (YPH501) and BY4743 respectively), may also influence viability after *lsm5* deletion. In this chapter the effect of the *lsm5* deletion on two different strains is investigated, showing that *LSM5* is indeed a non-essential gene, but that growth is severely affected. In contrast to its relatively severe effect on growth rate compared to *lsm6* and *lsm7* deletions, the *lsm5* deletion shows a very similar effect on the levels of spliceosomal snRNAs, U6 and U4/U6 in particular.

8.2 *lsm5* deletion strains are viable but show severely reduced growth rates

Mayes *et al.* (1999) reported that no viable spores carrying an *lsm5Δ::TRP1* deletion in the BMA64 background were obtained after sporulation. However, when AEMY48 which carries a plasmid (pAEM75-*CEN-URA3-*P_{GAL}*-LSM5*) in addition to the genomic *lsm5* deletion was grown on 5-FOA-Glu at 23°C some very slow growing cells survived (data not shown). These *lsm5Δ* cells (MRY6) were able to grow on solid medium at 14°C and 23°C, grew very poorly at 30°C and died at 37°C (Fig. 8.2, MRY7c). When grown in liquid medium they display doubling times of approximately 6 h at 18°C, 8 h at 25°C and more than 16 h at 30°C (Fig. 8.1C), showing that these cells are extremely sick and heat sensitive. Effects of *lsm* deletions on growth of BY4741 and BY4742 strains tend to be less severe than those



C

Growth temperature	BMA64a *lsm5Δ* (MRY6)	BY4741 *lsm5Δ* (MRY27)
18°C	+ (6h)	+ (n.t.)
25°C	+ (8h)	+ (4h)
30°C	+/- (>16h)	+ (4h)
37°C	-	n.t.

Fig. 8.1 Creation of an *lsm5* deletion strain and its growth phenotype. **(A)** Colony PCR was performed on wild-type BY4741 and a Hygromycin resistant colony after transformation with an HpHMX6 deletion cassette targetting the *LSM5* ORF. Colony PCR with primers mro83 and rpTEF only results in a product for the deletion strain (lane 2); PCR with primers mro83and mro84 results in a small product (lane 3) covering the *LSM5* ORF in wild type and a large product (lane 4) covering the *HpHMX6* casette inserted at the *LSM5* locus in *lsm5*Δ. **(B)** Ten-fold serial dilutions of liquid BY4741 and MRY27 cultures were plated on YPDA and incubated at 14°C, 23°C, 30°C or 37°C. **(C)** MRY6 and MRY27 were grown in YPDA at 18°C, 25°C, 30°C or 37°C and their doubling times were estimated

seen for the same deletions in the BMA64 background. An *lsm5* deletion was made by direct insertion of an *HpHMX6* cassette replacing the entire *LSM5* ORF in BY4741 cells and subsequent selection on YPDA-Hyg plates at 23°C. Resulting colonies were tested by colony PCR for insertion of the *HpHMX6* cassette at the correct locus (Fig. 8.1A). The resulting MRY27 strain was shown to grow on solid YPDA medium at all temperatures (14°C, 23°C, 30°C and 37°C), but many times slower than BY4741 wild-type cells, particularly at higher temperature (Fig. 8.1B). When grown in liquid medium MRY27 cells displayed doubling times of approximately 4 h at both 25°C and 30°C, so considerably faster than MRY6 cells (Fig. 8.1C). Similar differences in growth phenotypes were observed between *lsm6* and *lsm7* deletion strains in the BMA38 background compared to the same deletions in the BY4741 background. In the BMA38 background these deletion strains are heat-sensitive lethal, whereas in the BY4741 background they are heat-sensitive slow growing (compare Fig. 5.2 and 6.3; data not shown).

8.3 The *Arabidopsis sad1/lsm5* mutation does not show a growth defect in yeast

A single point mutation in a conserved residue in the N-terminal domain of *Arabidopsis thaliana* Lsm5p, the E12K *sad1* (supersensitive to ABA and drought) mutation, was shown to make plants hypersensitive to abscisic acid (ABA) and osmotic stress (Xiong *et al.*, 2001). The plant hormone abscisic acid plays an important role in plant growth and development, including seed maturation, root growth and stress tolerance. The promoter of *RD29A*, a stress-responsive gene, contains elements that confer responsiveness to cold/osmotic stress and ABA. In a screen, *sad1* was found to superinduce expression of *RD29A* by ABA and osmotic stress, but not cold stress. It is not known what the mechanism is behind the effects the *sad1* mutation has on stress response and ABA signalling. One possibility suggested by Xiong *et al.* (2001) is that the *sad1* mutation affects the decay rate of mRNA for an early component in ABA signalling, possibly a regulatory component that modulates ABA biosynthesis. Interestingly, microarray analysis suggested that the *sad1* mutation does not affect general gene expression (Xiong *et al.*, 2001).

Although *S. cerevisiae* obviously does not have an ABA signalling pathway I decided to investigate whether a mutation of the equivalent glutamate residue in position 10 of Lsm5p to lysine leads to an abnormal phenotype in yeast. *LSM5* was cloned with its promoter and terminator sequences (327 bp upstream and 127 downstream of the ORF) into pRS313. The resulting pMR9 plasmid was used to introduce the E10K mutation by site directed mutagenesis in pMR10. Both plasmids, as well as pRS313 were transformed into AEMY48 (*lsm5Δ* P_{GAL}-*LSM5*) and selected for on SDGal-His medium. Subsequent counter-selection against the pAEM75 (*CEN-URA3-*P_{GAL}*-LSM5*) plasmid on 5-FOA-Glu plates resulted in strains MRY7 (*LSM5*), MRY8 (*lsm5-E10K*) and pMR7c (*lsm5Δ* pRS313). Fig. 8.2 shows that there was no difference in growth between the wild-type BMA64a and MRY7 strains at any of the tested temperatures, indicating that the plasmid-based copy of *LSM5* fully supplements the genomic deletion. No difference, however, was observed either between the *LSM5* and *lsm5-E10K* strains at any of the temperatures tested.

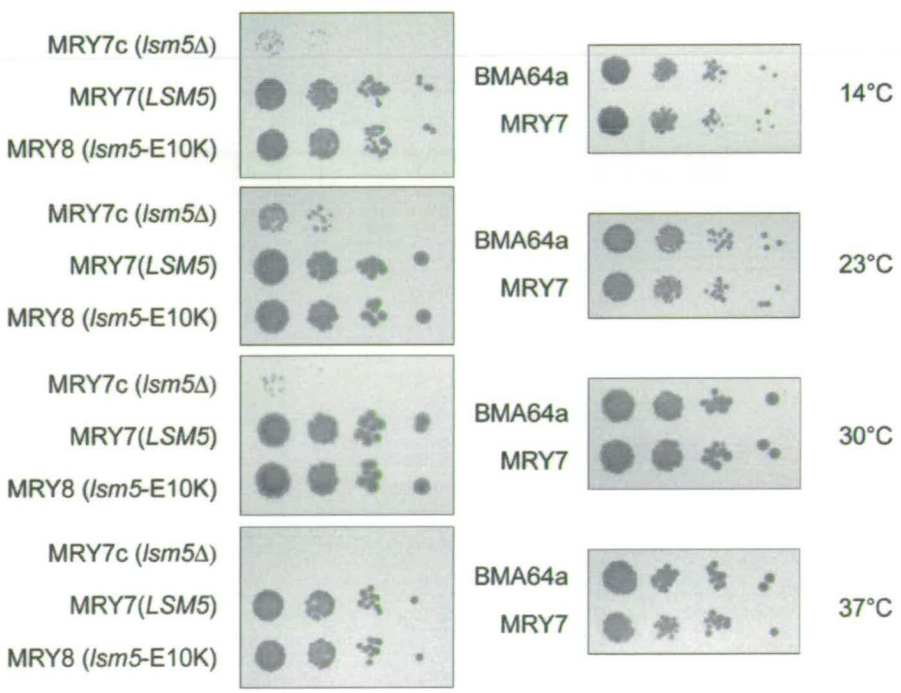


Fig. 8.2 *lsm5*Δ and *lsm5-E10K* growth phenotypes. *lsm5*Δ strain MRY6 carrying pRS313 (MRY7c), pMR9 (MRY7-*LSM5*) or pMR10 (MRY8-*lsm5-E10K*) as well as BMA64a (wild type) were grown in liquid medium, and 10-fold serial dilutions were plated on YPDA and grown at 14°C, 23°C, 30°C or 37°C.

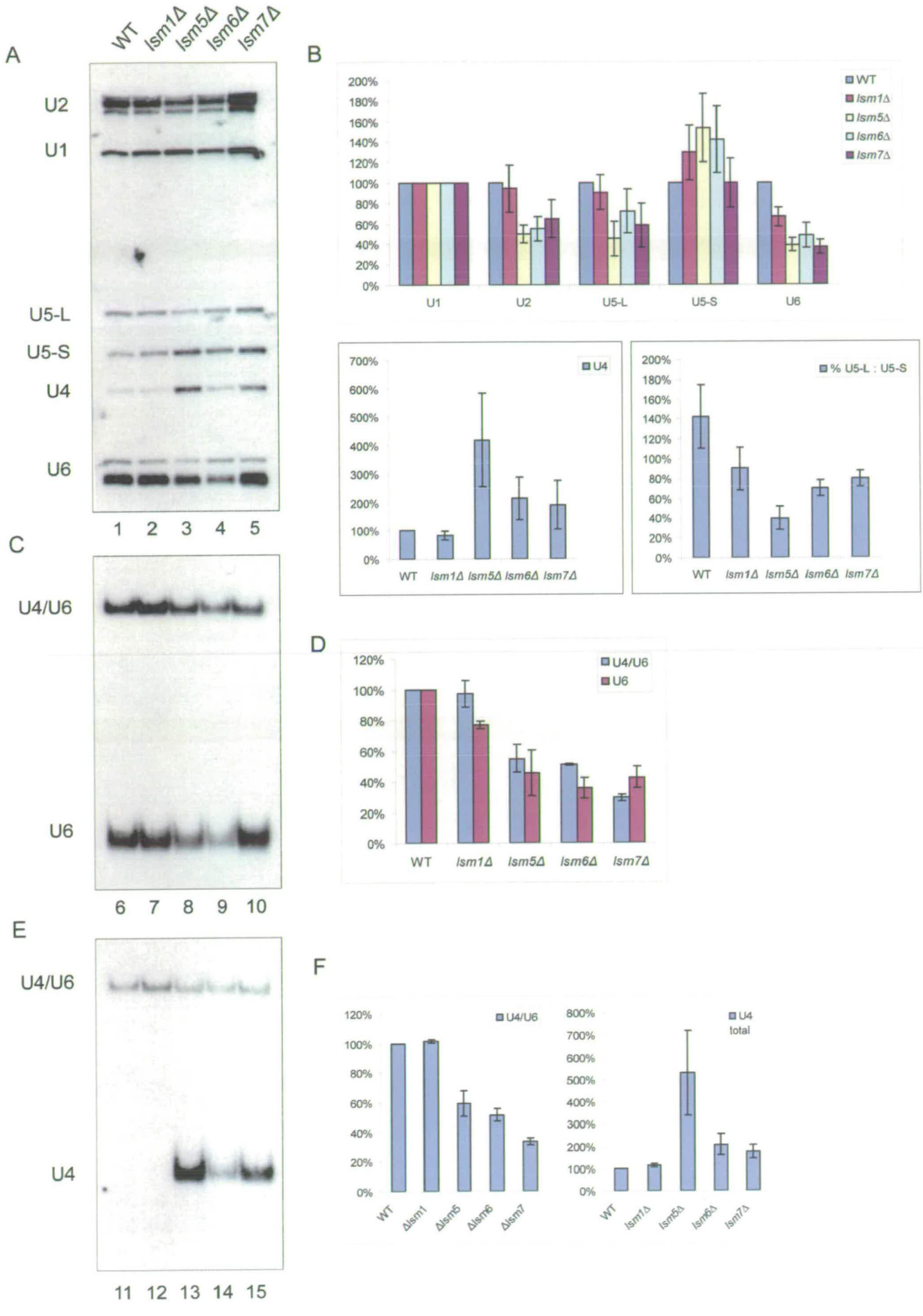
8.4 Effects of *lsm5Δ* on spliceosomal snRNA levels, a comparison with *lsm1Δ*, *lsm6Δ* and *lsm7Δ*

It is unclear why some of the Lsm proteins are essential and others are not, with the exception of Lsm1p which is assumed to function only as part of the non-essential Lsm1-7 complex. In contrast, it seems that the function of the Lsm2-8 complex is essential. It has been proposed that, of the wide variety of functions the nuclear Lsm proteins have in RNA processing, U6 snRNA stabilisation by the Lsm2-8 complex is the only one that is essential (Pannone *et al.*, 2001), as essential Lsm proteins become non-essential upon overexpression of U6 RNA (Mayes *et al.*, 1999; Pannone *et al.*, 2001). However, the Lsm2-8p subunits Lsm6p, Lsm7p and now Lsm5p have been shown to be non-essential. It is likely that the Lsm2-8 complex forms an alternative heterohexameric ring structure in the absence of one of these proteins, but an alternative heteroheptameric complex with one of the Lsm proteins occurring twice in one ring or with one of the homologous Sm proteins taking the place of the missing subunit cannot be ruled out. It is striking that all non-essential Lsm proteins are homologous to the SmE, F and G proteins that have been proposed to form a stable trimeric sub-complex (Raker *et al.*, 1996). The significance of this is unclear. It is clear however that deletion of non-essential as well as depletion of essential components of Lsm2-8p leads to destabilisation of U6 snRNA and reduction of the levels of U6 (Mayes *et al.*, 1999) and U4/U6 snRNAs (Mayes *et al.*, 1999; this thesis; Beggs lab unpublished observations). In addition, Verdone *et al.* (2004) suggested a specific role for Lsm6p in U4/U6 di-snRNP and Lsm7p in U4/U6.U5 tri-snRNP formation, as *lsm6Δ* extracts displayed reduced levels of annealing of exogenous U6 to U4 RNA, whereas *lsm7Δ* extracts displayed reduced U6 RNA incorporation into tri-snRNP particles. This suggests that each of the individual Lsm proteins may have a specific function in the context of the Lsm complexes. Analysis of the *lsm5Δ* strain may elucidate what this function is for the Lsm5 polypeptide.

8.4.1 *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains shows similar effects on the levels of spliceosomal snRNAs

The effect of the *lsm5* deletion on the levels of spliceosomal snRNAs was examined and compared to effects of other non-essential *LSM* gene deletions. Total RNA was extracted under non-denaturing conditions from wild-type BY4741 and isogenic *lsm1Δ* (Y01301), *lsm5Δ* (MRY27), *lsm6Δ* (MRY17) and *lsm7Δ* (MRY18) strains grown to OD_{600} 0.5 – 0.8 in YPDA at 30°C. The RNA was separated by denaturing and non-denaturing PAGE, Northern blotted and probed for the spliceosomal snRNAs (Fig. 8.3A, C and E). This was carried out for biological duplicates and quantification was performed on two separate Northern blots for each biological replica. The average values (shown in Fig. 8.3B, D and F) were corrected for equal loading by comparing the levels of U1 snRNA in each lane. The assumption that there is no effect of the *lsm* deletions on the levels of U1 RNA was shown to be correct by comparison to an scR1 loading control (Fig. 8.5). As expected the level of total U6 RNA (Fig. 8.3B) was reduced for *lsm5Δ* (39 ± 7% of WT), *lsm6Δ* (48 ± 12%) and *lsm7Δ* (37 ± 7%) compared to wild type. Surprisingly the level of U6 RNA was also reduced, but to a lesser extent, for *lsm1Δ* (66 ± 9%). A slight effect of depletion of some of the Lsm proteins on the levels of U5-L and U5-S had previously been reported by Mayes *et al.* (1999). Here a clear effect was seen of *lsm* deletions, leading to decreased levels of U5-L (91 ± 17% for *lsm1Δ*; 45 ± 17% for *lsm5Δ*; 72 ± 22% for *lsm6Δ*; 59 ± 22% for *lsm7Δ*) and increased levels of U5-S snRNA (129 ± 27% for *lsm1Δ*; 153 ± 33% for *lsm5Δ*; 142 ± 33% for *lsm6Δ*; 100± 24% for *lsm7Δ*).

Fig. 8.3 Deletions of *LSM5*, *LSM6* and *LSM7* lead to reduced levels of U2, U6 and U4/U6 snRNA, increased levels of U4 snRNA and a change in the ratio of U5-L over U5-S snRNA. RNA was extracted under non-denaturing conditions from BY4741 and isogenic *lsm1*Δ, *lsm5*Δ, *lsm6*Δ and *lsm7*Δ strains. The RNA was separated by denaturing and non-denaturing PAGE, Northern blotted and probed for the spliceosomal snRNAs. **(A)** Northern blot after denaturing PAGE and **(B)** quantification of all spliceosomal snRNAs. **(C)** Northern blot after non-denaturing PAGE probed for U6 RNA and **(D)** quantification of U4/U6 duplex and free U6 snRNA. **(E)** Northern blot after non-denaturing PAGE probed for U4 RNA and **(F)** quantification of U4/U6 duplex and total U4 snRNA. Levels were expressed as a percentage of WT and were corrected for equal loading by comparison to the amount of U1 snRNA present in each lane. Average values of biological duplicates (two Northerns each) are given.


A
WT
lsm1Δ
lsm5Δ
lsm6Δ
lsm7Δ
U2
U1
U5-L
U5-S
U4
U6
1 2 3 4 5
B
WT
lsm1Δ
lsm5Δ
lsm6Δ
lsm7Δ
U1
U2
U5-L
U5-S
U6
U4
% U5-L : U5-S
C
U4/U6
U6
6 7 8 9 10
D
U4/U6
U6
E
U4/U6
U4
11 12 13 14 15
F
U4/U6
U4 total
WT
Δlsm1
Δlsm5
Δlsm6
Δlsm7

Although there seemed to be considerable experimental variation, there was a significant effect on the U5-L over U5-S ratio, which changes from approximately 140% in wild-type to 90% in *lsm1Δ*, 40% in *lsm5Δ*, 70% in *lsm6Δ* and 80% *lsm7Δ*. Interestingly, deletions of *LSM5*, *LSM6* and *LSM7* also showed a clear effect on levels of U2 snRNA (50 ± 9% for *lsm5Δ*; 55 ± 12% for *lsm5Δ*; 65 ± 18% for *lsm7Δ*), whereas there was no significant effect for *lsm1Δ*. Finally, levels of U4 snRNA significantly increased for *lsm5Δ*, *lsm6Δ* and *lsm6Δ*, but not *lsm1Δ*. Levels of U4 RNA approximately doubled for *lsm6Δ* and *lsm7Δ* (214 ± 75% and 189 ± 86% of WT respectively), and on average quadrupled for *lsm5Δ* (419 ± 164%).

8.4.2 Levels of U6 and U4/U6 snRNA are decreased and U4 snRNA levels increased in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains

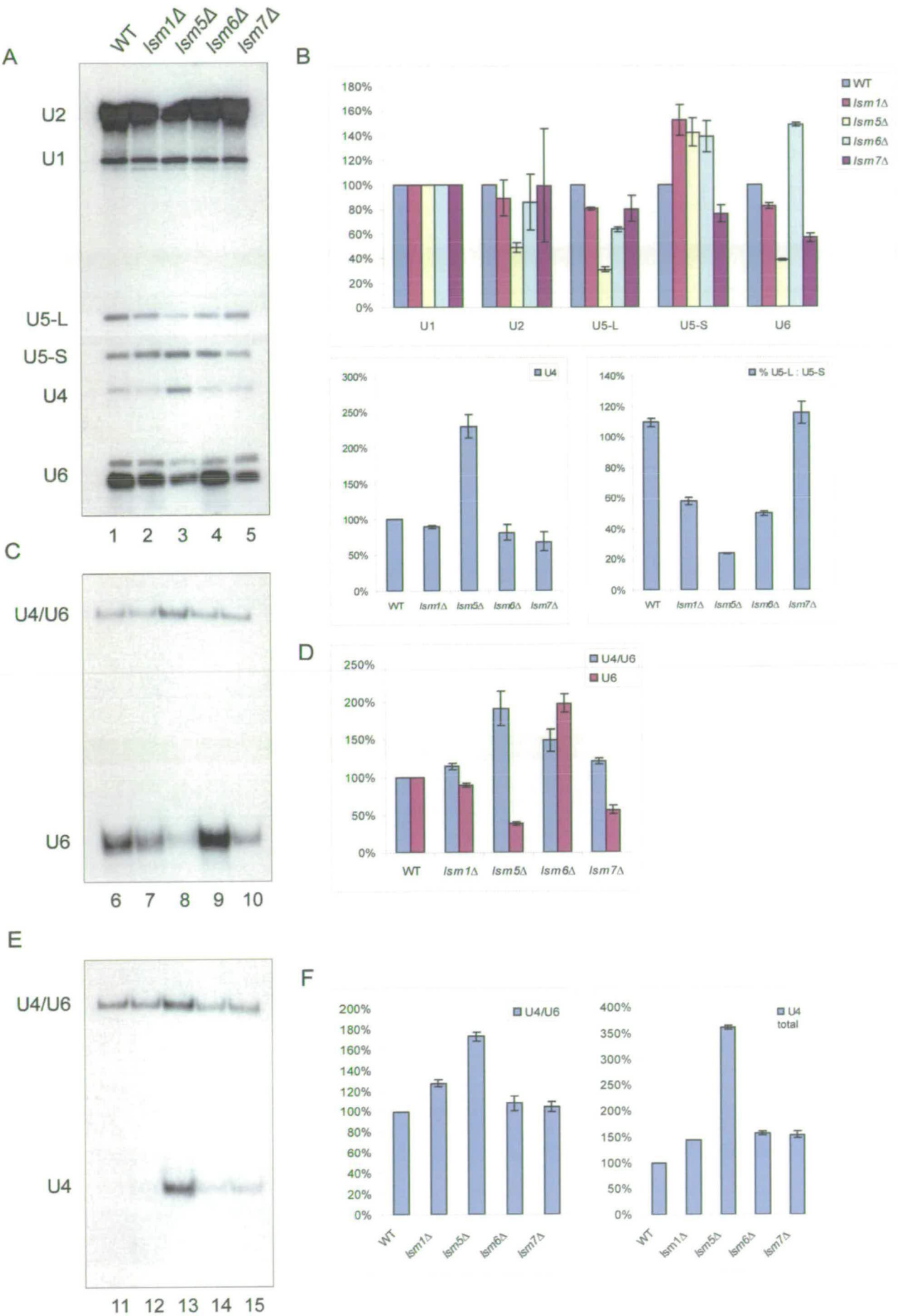
Quantification of levels of U4 and U6 snRNAs after non-denaturing PAGE (Fig. 8.3C-F) confirmed effects on the levels of U4 (529 ± 191% for *lsm5Δ*; 206 ± 48% for *lsm6Δ*; 175 ± 32% for *lsm7Δ*) and U6 snRNA (77 ± 2% for *lsm1Δ*; 46 ± 15% for *lsm5Δ*; 36 ± 7% for *lsm6Δ*; 43 ± 7% for *lsm7Δ*) found by denaturing PAGE (8.4.1). Levels of free U4 snRNA were high, in particular for *lsm5Δ*, while there was no detectable free U4 RNA in wild-type or *lsm1Δ* cells. In addition, levels of U4/U6 RNA were significantly reduced in *lsm5Δ* (55 ± 9%), *lsm6Δ* (51% ± 1%) and *lsm7Δ* (29% ± 2%), but not *lsm1Δ*.

8.4.3 Changed levels of spliceosomal snRNAs in *lsm* deletion strains are not fully corrected by increased U6 expression

As increased U6 expression is able to suppress growth defects of *lsm* deletion strains (Mayes *et al.*, 1999; Pannone *et al.*, 2001), we investigated whether U6 expression from pBS1191 (*CEN-URA3-SNR6*) in addition to expression from the *SNR6* genomic copy would have an effect on the changed levels of spliceosomal snRNAs in the same *lsm1Δ*, *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains (8.4.2). Total RNA was extracted under non-denaturing conditions from these strains carrying pBS1191 grown at 30°C in SD-Ura to OD_{600} 0.5 - 0.8; RNA was separated, Northern blotted, probed for spliceosomal snRNAs and bands were quantified. This was done for duplicate Northern blots of single biological samples. Graphs in Fig. 8.4 display average

values of quantifications of these Northern blots. Interestingly, the level of U6 snRNA, as estimated from Northern analysis after denaturing PAGE (Fig. 8.4A and B), did not increase for *lsm5Δ* (38 ± 1% of WT), but slightly increased for *lsm1Δ* (82 ± 3%) and *lsm7Δ* (56 ± 4%) and even increased to levels above those in wild-type cells for *lsm6Δ* (149 ± 2%). Levels of U4 snRNA were slightly decreased for *lsm6Δ* (81 ± 11%) and *lsm7Δ* (69 ± 14%) compared to wild type, in sharp contrast to the doubled levels of U4 in the absence of pBS1191 (Fig. 8.3). In *lsm5Δ*, however, the level of U4 RNA was still significantly above that of wild-type cells (230 ± 16%), but reduced compared to *lsm5Δ* in the absence of pBS1191 (Fig. 8.3). Similarly, the levels of U2 snRNA approached normal levels for *lsm6Δ* (89 ± 15%) and *lsm7Δ* (99 ± 46%), but were still down for *lsm5Δ* (49 ± 4% of WT). Finally, the change in the ratio of U5-L and U5-S snRNAs, due to respective decreased and increased levels in *lsm* deletion strains, was only corrected to a ratio similar to that of wild type (109 ± 3%) for *lsm7Δ* (115 ± 7%), but was still down for lsm5Δ (24 ± 0%), lsm6Δ (50 ± 2%) and lsm1Δ (58 ± 2%).

Fig. 8.4 Increased *SNR6* expression from pBS1191 only alleviates some effects deletions of *LSM5*, *LSM6* and *LSM7* have on the levels of U2, U4, U6 and U4/U6 snRNA, and the ratio of U5-L over U5-S snRNA. RNA was extracted under non-denaturing conditions from BY4741 and isogenic *lsm1*Δ, *lsm5*Δ, *lsm6*Δ and *lsm7*Δ strains carrying the pBS1191 (*CEN-URA3-SNR6*) plasmid and grown on SD-Ura. The RNA was separated by denaturing and non-denaturing PAGE, Northern blotted and probed for the spliceosomal snRNAs. **(A)** Northern blot after denaturing PAGE and **(B)** quantification of all spliceosomal snRNAs. **(C)** Northern blot after non-denaturing PAGE probed for U6 RNA and **(D)** quantification of U4/U6 duplex and free U6 snRNA. **(F)** Northern blot after non-denaturing PAGE probed for U4 RNA and **(F)** quantification of U4/U6 duplex and total U4 snRNA. Levels were expressed as a percentage of WT and were corrected for equal loading by comparison to the amount of U1 snRNA present in each lane. Average values for two Northerns of single biological samples are given.


A
WT
lsm1Δ
lsm5Δ
lsm6Δ
lsm7Δ
U2
U1
U5-L
U5-S
U4
U6
1 2 3 4 5
B
WT
lsm1Δ
lsm5Δ
lsm6Δ
lsm7Δ
U1
U2
U5-L
U5-S
U6
U4
% U5-L : U5-S
C
U4/U6
U6
6 7 8 9 10
D
U4/U6
U6
E
U4/U6
U4
11 12 13 14 15
F
U4/U6
U4 total

Quantification of the levels of U4, U6 and U4/U6 snRNAs after non-denaturing Northern analysis showed a similar trend (Fig. 8.4C-F). Levels of U4/U6 snRNA were back to normal or even higher than those for wild type for all *lsm* deletion strains carrying pBS1191 (114 ± 5% for *lsm1Δ*, 192 ± 23% for *lsm5Δ*, 149 ± 15% for *lsm6Δ* and 121 ± 4% for *lsm7Δ*). Free U6 snRNA approached the wild type level for *lsm1Δ* (89 ± 3%) and was significantly higher for *lsm6Δ* (198 ± 13%), but was still reduced in *lsm5Δ* (38 ± 2%) and *lsm7Δ* (57 ± 5%). There were still significantly increased amounts of free U4 snRNA present in *lsm5Δ*, *lsm6Δ and lsm7Δ* compared to wild type. The total level of U4 RNA as estimated from this non-denaturing Northern analysis, however, was similar between *lsm1Δ*, *lsm6Δ* and *lsm7Δ*, but was more than tripled in *lsm5Δ* (361 ± 4% compared to WT). Although there were some inconsistencies, the reasons for which are unclear, the general trend was the same between quantification of the denaturing and non-denaturing Northerns.

8.5 scR1 levels are increased in *lsm1Δ*

As mentioned above, loading was controlled for by comparing levels of U1 snRNA in the different lanes under the assumption that *lsm* deletions do not affect this RNA. As *lsm* deletions seem to affect levels of all other snRNAs, I wanted to test whether this assumption is correct by comparison to another RNA that is not known to be affected by the Lsm proteins. scR1 is the stable cytoplasmic non-coding RNA component of the signal recognition particle, which is involved in co-translational targeting of polypeptides with a signal peptide sequence to the endoplasmic reticulum. Lsm proteins are not known to be involved in its biogenesis or its degradation, although it was shown that depletion of Lsm3p and Lsm5p leads to reduced precipitation of scR1 by Lhp1p (Kufel *et al.*, 2002). The Northern blot after denaturing PAGE from Fig. 8.4A was probed for scR1 RNA and bands were quantified and levels were compared to those of U1 snRNA.

The results presented in Fig. 8.5 show that there was no significant effect of deletion of *LSM5*, *LSM6* or *LSM7* on the levels of scR1. This confirms that U1 snRNA was also not affected by these deletions, as the levels of U1 RNA were also comparable between wild type and the *lsm* deletion strains. Interestingly however, the level of scR1 RNA in *lsm1Δ* was almost three times (293 ± 32%) that of wild type.

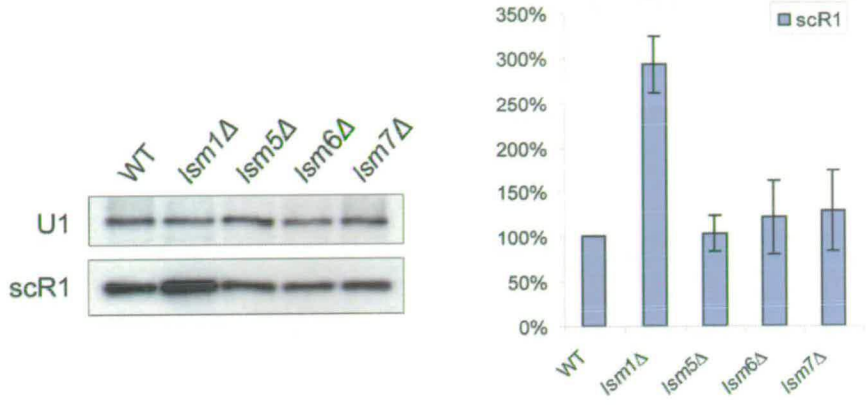


Fig.8.5 Levels of scR1 RNA are increased in an *lsm1Δ* strain, but not in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains. The denaturing Northern blot from Fig. 8.4 was probed for U1 and scR1 RNA. Bands were quantified and scR1 amounts were corrected for equal loading by assuming no effect of the *LSM* deletions on U1 snRNA. Approximately equal relative intensity for scR1 in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* compared to wild type shows that this assumption is correct. The amount of scR1 in *lsm1Δ*, however, has almost tripled. Average values for two Northerns of single biological samples are given.

8.6 Discussion

Roles for Lsm1p in mRNA decapping (Tharun *et al.*, 2000; Bouveret *et al.*, 2000) and for Lsm6p and Lsm7p in U4/U6 di-snRNP and U4/U6.U5 tri-snRNP formation respectively (Verdone *et al.*, 2004) were elucidated using the corresponding deletion strains. It seems likely that each of the subunits of the heteroheptameric Lsm complexes has a specific function, which it can probably only carry out in the context of this complex. The availability of a viable *lsm5* deletion strain will allow us to study functions associated with Lsm5p more easily. However, care should be taken in interpreting the results from analyses using this strain as together with the *LSM5* ORF, another ORF on the opposite strand of the chromosomal DNA is disrupted.

This ORF, *FTR1* (YER145C), nearly fully overlaps the *LSM5* ORF (282 bp) with the exception of the final 43 bp at the 3' end of *LSM5*, which presumably means the *LSM5* terminator sequence and *FTR1* promoter sequences also overlap. As the coding sequence of *FTR1* covers 1215 bp the *LSM5* promoter must lie within the *FTR1* open reading frame. Ftr1p is a high affinity iron permease involved in the transport of iron across the plasma membrane. It forms a complex with Fet3p and its expression is regulated by iron levels (Stearman *et al.*, 1996; Eide, 1998). *FTR1* is non-essential, but its deletion leads to reduced growth rates after several generations when grown on galactose, synthetic complete medium or a non-fermentable carbon source (Giaever *et al.*, 2002; Steinmetz *et al.*, 2002). Due to extensive overlap between these two genes, deletion of *LSM5* without affecting *FTR1* expression would require introduction of a non-sense mutation into the *LSM5* coding sequence that does not alter the Ftr1 protein sequence.

The *sad1* allele of *LSM5* shows a subtle defect in *Arabidopsis thaliana* that does not alter general gene expression, but seems to specifically affect positive feedback regulation of ABA biosynthesis genes by ABA and drought stress induction of ABA biosynthesis. It was suggested that the *sad1* mutation affects the decay rate of mRNA for an early component in ABA signalling. Negative regulators of ABA signalling, including protein phosphatase 2-C-like gene and *ABI1*, were indeed shown to be amongst those messages with reduced levels in the *sad1* mutant (Xiong *et al.*, 2001). Initial analysis of the equivalent E10K mutation in *LSM5* in yeast showed no growth defects. Since the *sad1* mutation in *Arabidopsis* displays only a subtle defect with regards to growth, development and stress response, this mutation may have a similar subtle effect in yeast. It is possible that this residue plays a role in targeting specific transcripts involved in stress response for degradation and that this function is conserved between *Arabidopsis* and yeast. It may therefore be interesting to investigate whether the *lsm5-E10K* strain has a growth defect under specific conditions, e.g. after osmotic shock, and if so what transcripts are affected under these conditions.

It is unclear whether the phenotypic effects that are seen for *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains are due to reduced efficiency of complex formation, reduced RNA-binding and/or reduced activity of alternative complexes, or loss of specific functions associated with the Lsm5, 6 and 7 proteins. Similar effects of all three deletion strains on U2, U5-L/U5-S, U4 and U6 snRNA levels suggest that this is due to one of the former. Effects on U2, U5 and U4 snRNA levels may be an indirect effect of the reduced levels of U6 snRNA. Alternatively, all three *lsm* deletions may have a direct effect on degradation and/or processing of these RNAs. Increased expression of U6 RNA rescues the U4/U6 to levels above wild type for all three mutants. The fact that U4/U6 levels are increased so readily may be due to the increased levels of free U4 snRNA, explaining why U4/U6 increases the most for *lsm5Δ*. Surprisingly however, effects of increased U6 snRNA expression on the other snRNAs vary for the three *lsm* mutants, suggesting that it is not a simple knock-on effect of reduced levels of U6 and U4/U6 RNA. The fact that effects are different for each of the *lsm* mutants may help us to elucidate mechanisms behind the effects on the spliceosomal snRNAs and may ultimately teach us more about functions of these individual Lsm proteins.

Increased levels of total U4 snRNA and reduced levels of U4/U6 suggest that there may be a direct link between U4 RNA degradation and U4 release from the spliceosome. As a result of inefficient U4/U6 di-snRNP formation U4 incorporation into the pre-spliceosome will be reduced, leading to an overall reduction in release of U4 RNA prior to activation. It is possible that the Lsm proteins thus indirectly affect U4 RNA degradation, as U6 overexpression (partially) alleviates increased U4 RNA levels in the *lsm* deletion mutants. Mutations in a number of other U6 snRNP-related splicing factors, Prp3p, Prp4p and Prp24 (Lygerou *et al.*, 1999), and a number of Prp19 complex components, including Prp19p and NTC25 (Lygerou *et al.*, 1999; Chen *et al.*, 2006) were shown to have reduced U6 and U4/U6 levels and to accumulate free U4 snRNA. However, none of these was shown to lead to increased levels of total U4 RNA[3]. It is therefore possible that the Lsm2-8 complex may have a more direct function in targeting U4 RNA for degradation after its release from the

[3] Care should be taken in interpreting levels of free U4 and U6 RNA, as Chen *et al.* (2006) showed that there is a marked difference between these levels in splicing extracts and total RNA extractions.

spliceosome. U4 snRNA was shown to be degraded by the exosome, as mature U4 snRNA accumulated in strains depleted of various exosome components (Allmang *et al.*, 1999). Lsm2-8p may therefore have a function in exosome-mediated U4 degradation, and as U4 snRNA levels were increased the most in the *lsm5Δ* strain Lsm5p may play a central role in this. Exosome mutations also led to accumulation of U5-S, but not U5-L, significantly increasing the U5-L/U5-S ratio (Allmang *et al.*, 1999). This suggests that the increase in this ratio in *lsm* mutant strains may similarly be caused by an effect of Lsm2-8p (and/or Lsm1-7p) on exosome-mediated degradation of U5-S.

Surprisingly, deletion of *LSM1* also affected the levels of spliceosomal snRNAs (except U1 and U4), although the effects were less pronounced than in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains. The reason for this remains unclear. Interestingly, however, Kufel *et al.* (2002) also showed similar effects on (nuclear) pre-tRNA processing in *lsm1*Δ, *lsm6*Δ and *lsm7*Δ strains at 37 °C, although these effects were much weaker than those observed for strains depleted of essential Lsm proteins. Spiller (2005; unpublished results) suggested that there may be competition between formation of Lsm1-7p and Lsm2-8p and a regulated exchange of shared subunits depending on growth conditions. If this is true, it may be expected that levels of Lsm2-8p would be higher in an *lsm1Δ* strain, which would have the opposite effect to that observed here. On the other hand, the cytoplasmic Lsm1-7 complex may act as a source of Lsm2-7 proteins for incorporation into the nuclear Lsm2-8p complex if the need for this arises. Deletion of *LSM1* may disturb this delicate balance reducing the flexibility of the cell to quickly respond to changes in conditions. Alternatively, the Lsm1-7 complex has a function in stabilising or processing these stable small RNAs, presumably in the cytoplasm. This suggests a cytoplasmic phase for U6 RNA and pre-tRNAs. U6 RNA has indeed been shown to shuttle between nuclei in a yeast heterokaryon assay (Olson and Siliciano, 2003) and has been shown to accumulate in the cytoplasm in the absence of Lsm2-8 proteins (Spiller *et al.*, 2007). Also, splicing of pre-tRNAs has been shown to take place in the cytoplasm (Yoshihisa *et al.*, 2003), showing that a role for Lsm1-7p in processing of these RNAs in the cytoplasm may

be possible. Alternatively, Lsm1-7p is present at low but functional levels in the nucleus.

Increased levels of scR1 RNA in *lsm1Δ*, but not in *lsm5Δ*, *lsm6Δ* and *lsm7Δ* seem to suggest a direct involvement of Lsm1p, and presumably Lsm1-7p, in scR1 degradation. Why would deletion of *LSM1*, but not deletion of any of the other three non-essential *LSM* genes affect scR1 degradation? In the absence of Lsm1p there is presumably no functional cytoplasmic Lsm complex, whereas in the absence of Lsm5, 6 or 7p there may be an Lsm1-7 complex lacking one of its subunits, which could be functionally competent in terms of scR1 degradation. It would be interesting to investigate whether scR1 degradation is carried out by the exosome or by Xrn1p and what role Lsm1-7p plays in this. Increased scR1 levels in an *rrp6* deletion strain (Allmang *et al.*, 1999) suggest that this exosome component may be involved in scR1 degradation. Lsm1-7p may therefore affect exosome-mediated degradation of this cytoplasmic non-coding RNA.

In summary, I have shown that *LSM5* is a non-essential gene. Its deletion has similar effects to those seen for *lsm6* and *lsm7* deletions: decreased U2, U6 and U4/U6 levels, a change in the U5-L/U5-S ratio and increased U4 snRNA levels. The strong increase in U4 snRNA levels for *lsm5Δ* suggest that Lsm5p may have a specific function in the exosome-mediated decay of U4 RNA.

CHAPTER 9

Dissection of Lsm1p and Lsm8p in sub-cellular localisation

9.1 Introduction

S. cerevisiae has at least two different Lsm complexes. The exclusively nuclear Lsm2-8 complex consists of the Lsm2 to Lsm8 proteins and forms the core of the spliceosomal U6 snRNP particle (Mayes *et al.*, 1999; Salgado-Garrido *et al.*, 1999; Spiller, 2005). A second complex is formed by the Lsm1 to Lsm7 proteins and is thought to be exclusively cytoplasmic. This Lsm1-7 complex has been shown to play a role in mRNA decapping by Dcp1p/Dcp2p and subsequent 5' to 3' degradation by Xrn1p exonuclease (Tharun *et al.*, 2000; Tharun *et al.*, 2001). It has been shown that this process occurs at specific cytoplasmic sites termed P- or Processing Bodies (Sheth *et al.*, 2003). Under conditions that warrant high levels of mRNA turnover such as osmotic shock or glucose starvation P-bodies increase in number and size (Teixeira *et al.*, 2005). Complex formation seems to be essential for Lsm1p and Lsm8p to localise to P-bodies and to nuclei respectively. This was shown for Lsm2-8p in yeast where Lsm8p loses its nuclear localisation when cells are depleted of Lsm2p or Lsm4p (Spiller, 2005; unpublished results). Human LSm4 was shown to lose its localisation to P-bodies when mutations in residues that are predicted to be involved in complex formation (G27D or L40N) were introduced (Ingelfinger *et al.*, 2002). Both complexes contain the Lsm2 to 7 proteins. Completion of the ring-shaped heteroheptamer by Lsm8p leads to nuclear localisation, whereas a ring completed by Lsm1p is (presumably) excluded from the nucleus. In contrast, the homologous Sm complex, which forms the core of the non-U6 spliceosomal snRNPs, accumulates in the nucleus only. SmB/D3, SmD1/D2 and SmF/E/G subcomplexes have been suggested to form as RNA-free particles and to assemble into the heptameric complex on the snRNAs (Raker *et al.*, 1996; Raker *et al.*, 1999). The basic residues in the C-terminal protuberances of the yeast SmB and D1 proteins have been shown to form separate nuclear localisation signals that are functionally redundant (Bordonné, 2000). The human SmB, D1 and D3 proteins were shown to contain similar signals important for nuclear localisation (Girard *et al.*, 2004). Nothing is known however about what leads to the nuclear accumulation of Lsm2-8p on the one hand and the nuclear exclusion of Lsm1-7p on the other. Since Lsm2 to 7p are shared between both complexes it would seem likely that specific localisation signals are present within the Lsm1 and Lsm8 proteins themselves. Tharun *et al.*

(2005) performed extensive mutational analysis of *LSM1* showing the importance of residues proposed to be involved in RNA binding and complex formation, and of the C-terminal domain for the functional competence of the Lsm1-7 complex. However, not much attention was given to what domains or residues are essential for specific localisation, apart from showing that (surprisingly) mutations in the putative RNA-binding residues did not significantly affect localisation to P-bodies. To investigate the requirement of different domains of the Lsm1 and Lsm8 proteins for their function and localisation in more depth, I created a number of mutant and hybrid proteins. These were all tested for their ability to support growth either of an *lsm1Δ* strain at non-permissive temperature or of an *lsm8Δ* strain. This showed that essential functions of Lsm8p are remarkably resilient to mutations. The cellular localisation of GFP-tagged versions of all hybrid and mutant proteins was investigated by fluorescence microscopy, uncovering a central role for the N-terminal domains of Lsm1p and Lsm8p for specific cellular localisation of the Lsm1-7 and Lsm2-8 complexes.

9.2 Creation and expression of Lsm1p and Lsm8p hybrids and deletion mutants

To decide where to make deletions and what domains of Lsm1p and Lsm8p to fuse together in hybrid polypeptides the two proteins were aligned with ClustalW (Fig. 9.1A). In addition, the 2D/3D structures of each of these proteins was predicted using the 3D-PSSM server (Fig. 9.1B and C; Kelley *et al.*, 2000b).

9.2.1 Structure prediction of Lsm1p and Lsm8p

The most highly conserved part between the Lsm1 and Lsm8 polypeptides is the part that covers the Sm1 and Sm2 motifs (Fig. 9.1A). These motifs form the Sm-fold, the hallmark of the Sm-like proteins, consisting of a five-stranded anti-parallel β-sheet which is involved in inter-subunit and protein-RNA contacts. The Sm1 motif is formed by β-strands β1 to β3 (aa 53–88 for Lsm1p and aa 12-45 for Lsm8p) and the Sm2 motif is formed by β4 and β5 (aa 94-113 for Lsm1p and aa 49-65 for Lsm8p; Fig. 9.1B and C). Crystal structures and cross-linking data have shown that RNA

binding residues in Sm(-like) proteins are located in loop 3 (between β2 and β3) and loop 5 (between β4 and β5). The consensus sequences for these so-called Knuckle motifs in eukaryotic Sm and archaeal Sm-like proteins are [His/Tyr]-Met-Asn for Knuckle I and Arg-Gly-Asp for Knuckle II (Khusial *et al.*, 2005). It is unknown how the Lsm proteins bind RNA, but it is presumed to occur in a similar fashion. The Lsm-ring subunits and the Sm rings have similar amino acid sequences in Knuckle II, however, most Lsm proteins diverge from the consensus sequence in the Knuckle I motif. Asparagine residue 3 is highly conserved but residues at positions 1 and 2 are often substituted by large hydrophobic residues. This has been suggested to be the cause of the broader RNA binding specificity of the Lsm rings (reviewed by Khusial *et al.*, 2005). This seems to hold true for yeast Lsm1p (Knuckle I: Tyr-Ala-Asn; Knuckle II: Arg-Gly-Glu) as well as yeast Lsm8p (Knuckle I: Asn-Thr-Asn; Knuckle II: Arg-Gly-Ser). Some of these putative RNA binding residues in Lsm1p were mutated and shown to affect both mRNA decay and mRNA 3' end protection, but not localisation to P-bodies (Tharun *et al.*, 2005).

Prediction of secondary structures outside the Sm motifs revealed an α-helical structure directly upstream of β1, another common feature of the Sm-fold (α1: aa 43-49 for Lsm1p and aa 3-8 for Lsm8p; Fig. 9.1B and C). In addition both proteins show potential α-helical structures in their C-terminal extensions (strong prediction for Lsm1p aa 135-159; weaker predictions for Lsm1p aa 124-129 and Lsm8p aa 93-104). Most of the N-terminal extension of Lsm1p is predicted to be unstructured. Based on the alignment and structure predictions (Fig. 9.1) fusions and deletions of N- and C-terminal domain were made outside the most highly conserved regions of the Sm motifs and care was taken to avoid structured regions. The N-terminal domain of Lsm1p was taken as aa 1-51 and that of Lsm8p as aa 1-10. The C-terminal domain of Lsm1p was taken as aa 122-172 and that of Lsm8p as aa 74-109, with the remaining residues in between the N- and C-terminal domains covering the Sm-domain.

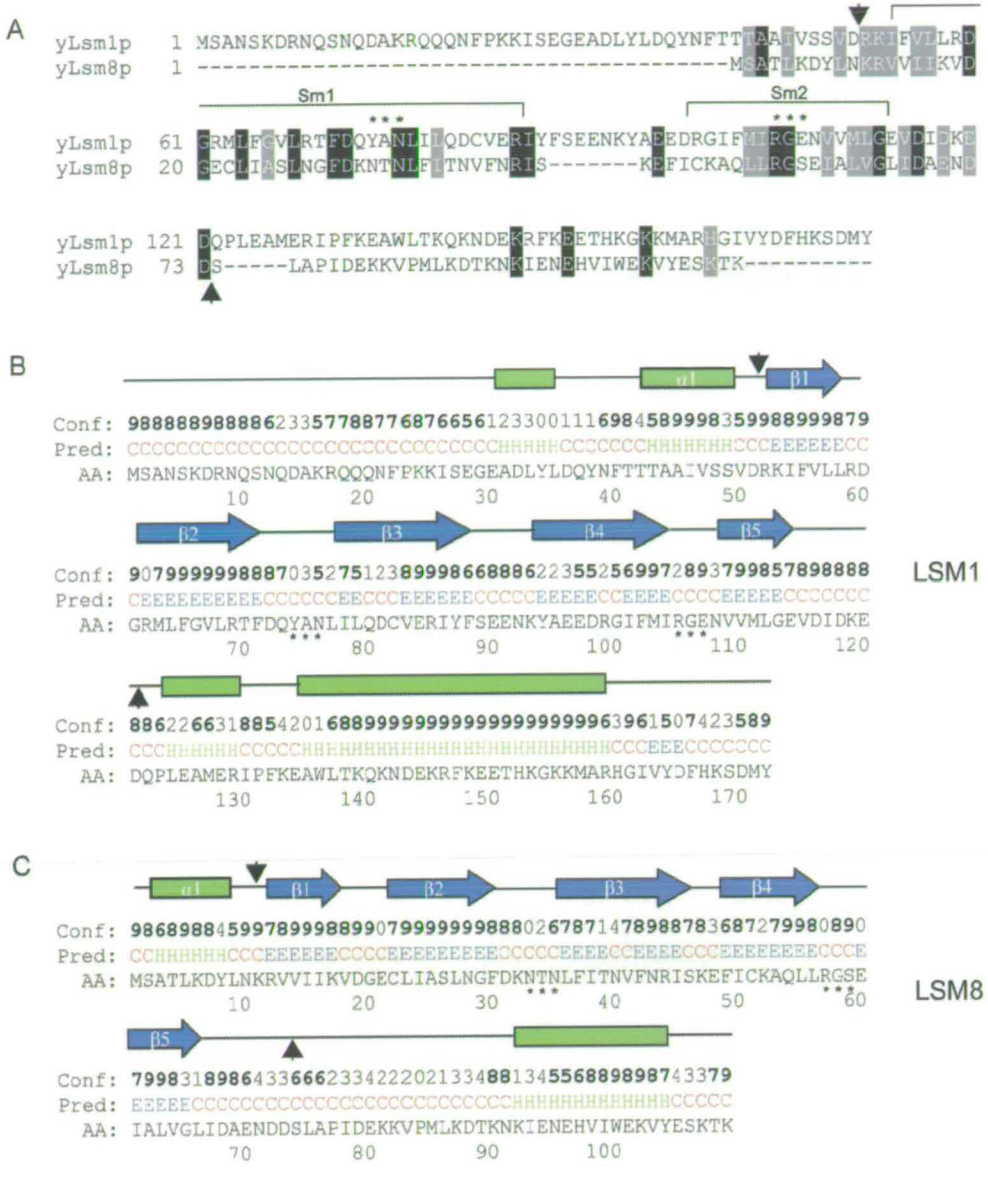


Fig. 9.1 Lsm1p and Lsm8p alignment (**A**; ClustalW; boxview) and 2D/3D structure predictions (**B,C**; 3D-PSSM). Arrowheads indicate the sites at which N- and C-terminal deletion mutants and hybrids of Lsm1 and 8 proteins were created. Asterisks indicate the amino acids forming the putative RNA-binding Knuckle domains. In **(B)** and **(C)** green boxes indicate regions that are predicted to form α-helices (H), blue arrows indicate regions that are predicted to form β-strands (E) and lines indicate regions that are predicted to form random coil (C). The corresponding numbers give the confidence score of these predictions, with 5-9 (in bold) indicating high confidence. α1 and β1-5 of the Sm-like fold are predicted correctly for both Lsm1p and Lsm8p.

9.2.2 Hybrids and mutant proteins are stably expressed

To test whether the hybrid and mutant proteins were stably expressed in yeast, I expressed GFP-tagged versions of these proteins and performed α-GFP Western analysis on total protein samples. Plasmids with N- or C-terminally tagged proteins expressed from the *MET25* promoter were transformed into MSP26. Cells were grown in SD-Ura-Met (or SD-Ura) to an OD_{600} of approximately 3 and harvested. Total protein was prepared for each, separated by SDS-PAGE and Western blotted. Anti-GFP antibody was used to detect the GFP-tagged proteins and anti-α-Tubulin antibody was used to detect α-Tubulin to test for equal loading (Fig. 9.2). All proteins, except LsmΔN8ΔC-GFP were expressed at similar levels in the absence of methionine, indicating that they are stably expressed. Expression was significantly lower, but not absent, in the presence of 1 mM of methionine as can be seen for GFP-Lsm118. This reflects incomplete repression of the MET25 promoter, which can be seen at the level of mRNA (Fig. 9.10) as well as protein (Fig. 9.2).

9.3 The N- and C-terminal domains of Lsm1p and Lsm8p do not suffice as localisation signals

To test whether the N- and C-terminal extensions of Lsm1p and Lsm8p contain any localisation signals, these domains were fused to the N- and C-termini of GFP respectively (in pGFP-N-FUS and pGFP-C-FUS). The resulting plasmids were transformed into MPS26, cells were grown in SD-Ura-Met and localisation was examined in live cells during log phase growth. To test if the Lsm1p domains would allow for GFP localisation to P-bodies under hypo-osmotic stress conditions, cells were incubated in water for 10-60 min before examining localisation in live cells. All GFP-fusions showed localisation identical to GFP alone, i.e. throughout the cell, excluding vacuoles (Fig. 9.3). This indicates that the terminal extensions of Lsm1p and Lsm8p by themselves do not suffice as localisation signals. This does not rule out that they may play a role in localisation, possibly as part of a signal sequence together with domains of other Lsm proteins. The human SmB protein is an example of this as, deletion of its C-terminal domain leads to reduced nuclear localisation,

even though it did not suffice as a nuclear localisation signal when fused to GFP (Girard *et al.*, 2004).

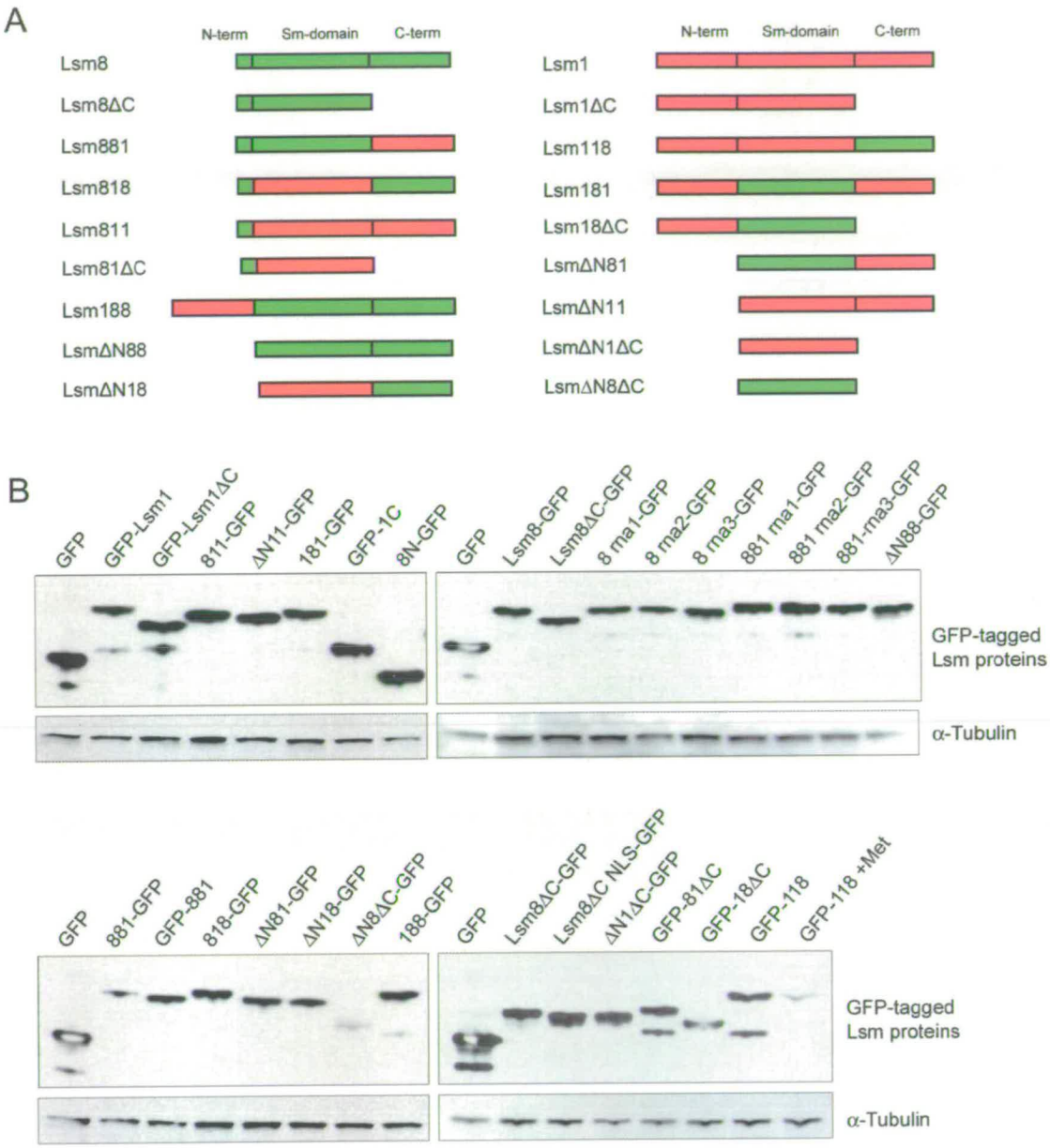


Fig. 9.2 Anti-GFP Western analysis of GFP-tagged Lsm1 and Lsm8 mutant and hybrid proteins. **(A)** Schematic overview of hybrids and deletion mutants of Lsm1p and Lsm8p. **(B)** MPS26 cells carrying the plasmids expressing the hybrid and mutant proteins GFP-tagged at the N- or C-terminus were grown into stationary phase in SD-Ura-Met and total protein samples were prepared of equal OD_{600} units. Proteins were separated by SDS-PAGE, Western blotted and probed with anti-GFP antibody. Equal loading was tested for by probing with anti-α-Tubulin antibody.

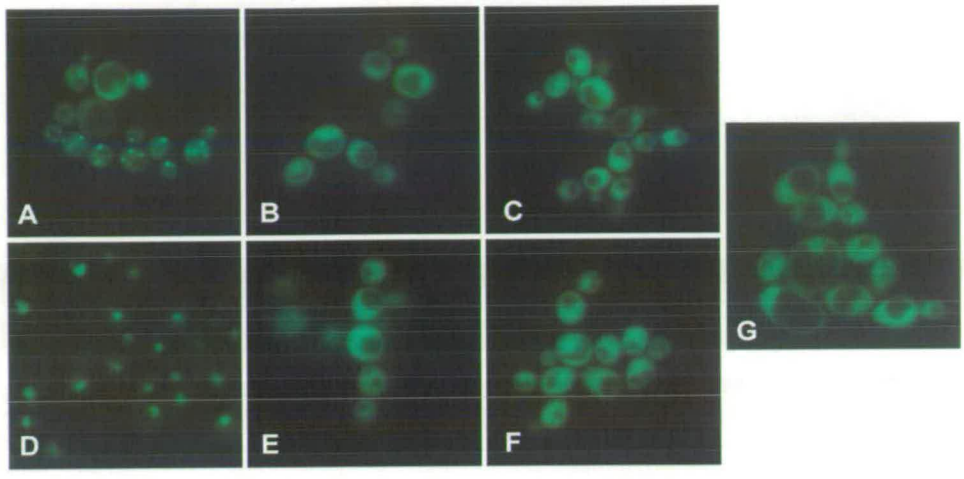


Fig. 9.3 Neither the N-terminal, nor the C-terminal extension of Lsm1p or Lsm8p is sufficient for localisation to P-bodies or nuclei respectively when fused to GFP. GFP-Lsm1 localises to P-bodies under stress conditions **(A)**, whereas Lsm1N-GFP **(B)** and GFP-Lsm1C **(C)** localise throughout the cell, like GFP **(G)** even under stress conditions. GFP-Lsm8 localises to the nucleus **(D)**, whereas Lsm8N-GFP **(E)** and GFP-Lsm8C **(F)** localise throughout the cell. MPS26 was transformed with pGFP-N-Lsm1, pMR144 (1N), pMR133 (1C), pGFP-N-Lsm8, pMR132 (8N), pMR156 (8C), or pGFP-N-FUS. Cells were grown in SD-Ura-Met and localisation was examined during log phase growth. Localisation under stress conditions was examined after hypo-osmotic shock in water for 10-60 min.

9.4 No single domain of Lsm8p is essential for nuclear localisation

Three different strategies were used to investigate the requirement of Lsm8p domains for its nuclear localisation. Firstly, mutations were made in the putative RNA binding residues to see if RNA binding is a prerequisite for nuclear localisation. Secondly, deletions of N- and/or C-terminal domains were made to see if, like for human SmBΔC (Girard *et al.*, 2004), localisation would be reduced. Thirdly, N-terminal, C-terminal or Sm domains of Lsm8p were replaced by those of Lsm1p to see if any particular domain from Lsm1p would lead to nuclear exclusion.

9.4.1 Reduced RNA binding may affect Lsm8p nuclear localisation

To test the effect of mutations in the putative RNA-binding Knuckle motifs, three different *rna* mutations were made in *LSM8*. Two of these were made in or near the Knuckle I motif: *lsm8 rna1* carried the mutations N28A and D31A, and *lsm8 rna2* carried the mutations T34A and N35A. A third, *lsm8 rna3*, was made in the Knuckle II motif and carried the mutations R57A, G58W and S59A. Based on analogous residues in Lsm1p (Fig. 9.1)(Tharun *et al.*, 2005) these would be expected to form the RNA-binding pocket (T34, N35, R57, S59) or be important for positioning of these residues (D31, G58). Mutations were introduced by site-directed mutagenesis using the Quik Change method in pGFP-N-Lsm8. Presence of the correct mutations and absence of any others in the *LSM8* coding sequence were checked by sequencing. The *lsm8* mutant genes could then be amplified by PCR and cloned into pGFP-C-FUS with or without the GFP fusion.

MPS26 was transformed with these plasmids expressing N- or C-terminally tagged mutant proteins from the *MET25* promoter. Cells were grown in SD-Ura or when fluorescence proved to be too weak in SD-Ura-Met, and localisation was examined in live cells during log phase growth. The *rna1* and *rna2* mutations showed to have no significant effect on nuclear accumulation of Lsm8p (Fig. 9.4A). In contrast, the *rna3* mutation in N-terminally tagged Lsm8p caused this protein to localise throughout the cell. However, when the same protein was tagged on the C-terminus it accumulated in the nucleus at levels comparable to those of wild-type GFP-Lsm8p (Fig. 9.4A). When, in addition to the *rna* mutations, the C-terminal domain of Lsm8p was replaced with that of Lsm1p all proteins failed to accumulate in the nucleus, irrespective of what side the GFP tag was on. Fluorescence of these hybrid proteins was extremely weak, potentially indicative of problems with correct folding, as expression levels were shown to be unaffected (Fig. 9.2). In contrast, replacement of the C-terminus of Lsm8p with that of Lsm1p in Lsm881p did not prevent accumulation of this hybrid protein in the nucleus, although nuclear levels were reduced compared to wild type (Fig. 9.4B).

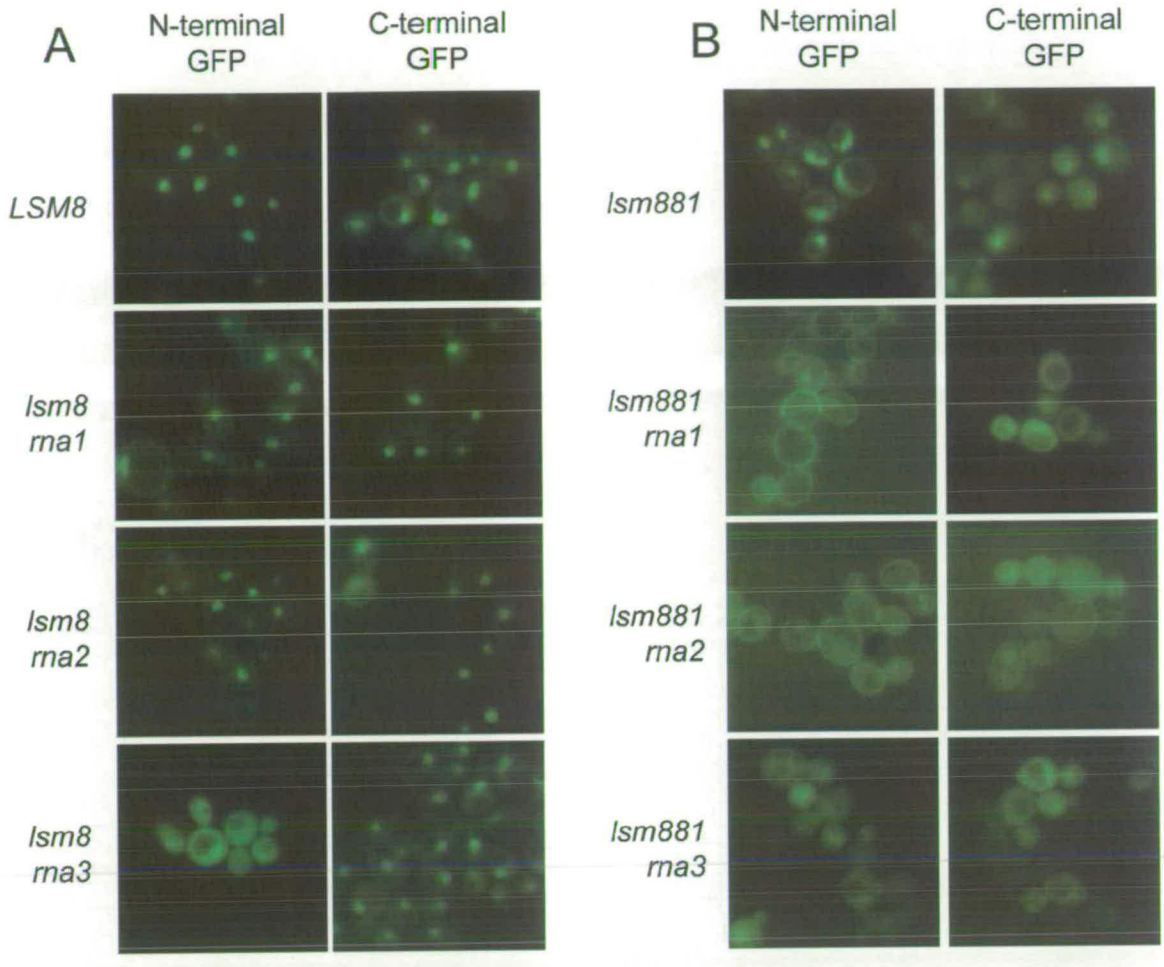


Fig. 9.4 Effects of mutations in putative RNA binding residues of Lsm8p on its nuclear localisation. MSP26 was transformed with plasmids **(A)** pGFP-N-Lsm8 and pMR83 (pGFP-C-Lsm8); pMR70, 71 and 72 (pGFP-N-Lsm8 rna1, 2 and 3); pMR92, 93 and 94 (pGFP-C-Lsm8 rna1, 2 and 3); **(B)** pMR70 and 84 (pGFP-N and -C-lsm881); pMR73, 74 and 75 (pGFP-N-Lsm881 rna1, 2 and 3); and pMR95, 96 and 97 (pGFP-C-Lsm881 rna1, 2 and 3). Cells were grown in SD-Ura (-Met) and localisation was examined in live cells during log phase growth. **(A)** Lsm8p, Lsm8-rna1p and -rna2p accumulate in the nucleus, but Lsm8-rna3p only accumulates in the nucleus when tagged on the C-terminus. **(B)** Mutant Lsm8 proteins that have their C-terminal domains replaced with that of Lsm1p in addition to the *rna* mutations fail to accumulate in the nucleus altogether and show extremely weak fluorescence, whereas Lsm811p shows reduced, but significant nuclear accumulation.

9.4.2 Any of the Lsm8p domains can be deleted or replaced without complete loss of nuclear accumulation

I have shown that the Lsm8p N- or C-terminal domains are not sufficient to allow GFP to accumulate in the nucleus (9.3). To test if these domains are essential for nuclear accumulation of Lsm8p, the termini were deleted or replaced with those of Lsm1p. Thus, the following mutant and hybrid proteins were created: Lsm8ΔCp, Lsm881p, LsmΔN88p and Lsm188p. To test if the central Sm-domain is essential for nuclear localisation (e.g. by giving Lsm2-8p U6 snRNA specific binding activity) this domain was replaced with that of Lsm1p in Lsm818p. In addition the Sm domain of Lsm1p was replaced with that of Lsm8p in Lsm181p. MPS26 was transformed with plasmids expressing these proteins, GFP tagged on the N- or C-terminus, from the *MET25* promoter. Localisation was examined in live cells during log phase growth.

The C-terminal domain of Lsm8p is not essential for nuclear accumulation as both Lsm8ΔCp and Lsm881p still accumulated in the nucleus (Fig. 9.5A). Both were significantly reduced however, particularly GFP-Lsm8ΔCp, suggesting that the Lsm8p carboxy-terminal domain does contribute to efficient nuclear localisation. The N-terminal domain is not essential for nuclear accumulation either, as both LsmΔN88p and Lsm188p still accumulated in the nucleus (Fig. 9.5B). Again, reduced nuclear localisation, particularly for Lsm188p, suggests that the N-terminal domain plays a significant role in localisation to the nucleus. Finally, failure of Lsm181p to accumulate in the nucleus and nuclear localisation of Lsm818p suggests that the Sm-domain of Lsm8p is neither essential nor sufficient for nuclear accumulation (Fig. 9.5C).

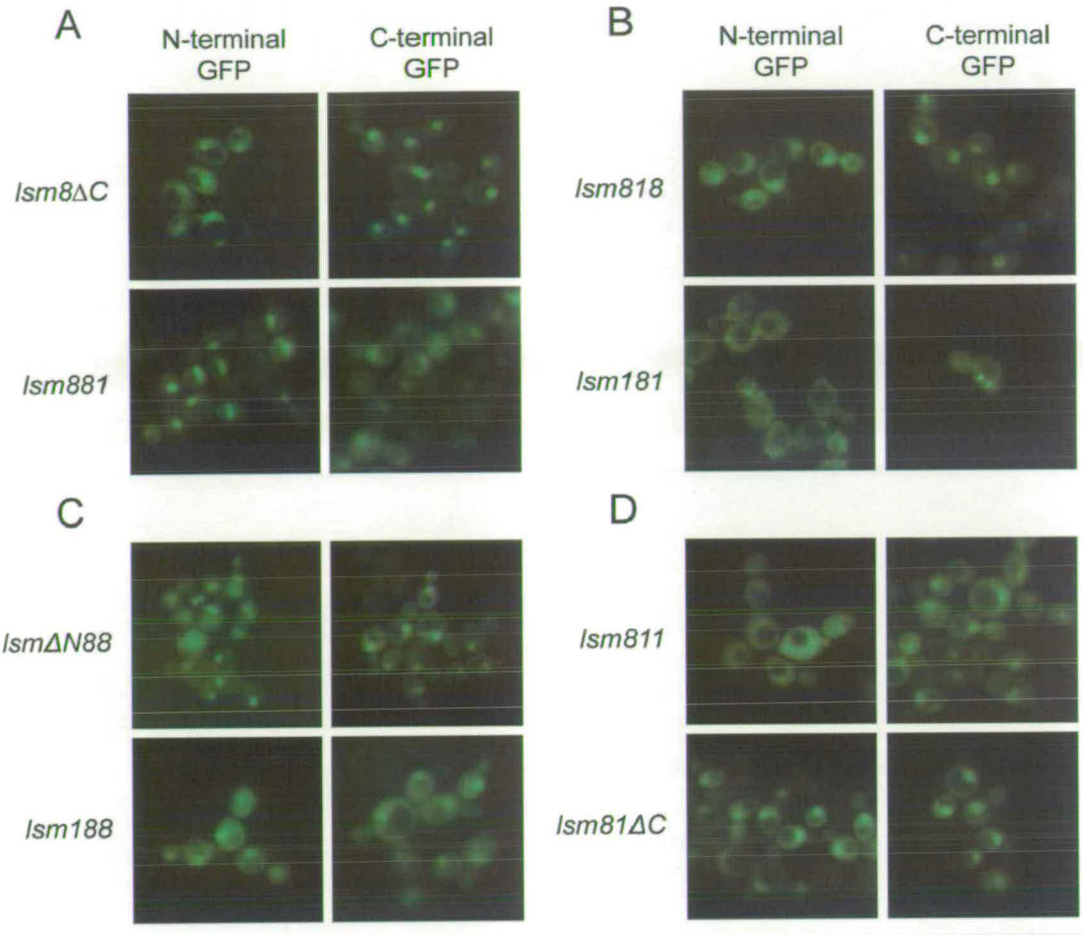


Fig. 9.5 None of the Lsm8p domains is essential for nuclear localisation, but the N-terminal domain plays a central role. MSP26 was transformed with plasmids **(A)** pMR70 and 80 (pGFP-N-Lsm881 and Lsm8ΔC), pMR84 and 104 (pGFP-C-Lsm881 and Lsm8ΔC); **(B)** pMR116 and 114 (pGFP-N-Lsm818 and 181), pMR125 and 123 (pGFP-C-Lsm818 and 181); **(C)** pMR126 and 140 (pGFP-N-Lsm188 and lsmΔN88), pMR117 and 141 (pGFP-C-Lsm188 and lsmΔN88); **(D)** pMR115 and 129 (pGFP-N-Lsm811 and Lsm81ΔC), pMR124 and pMR130 (pGFP-C-Lsm811 and Lsm81ΔC). Cells were grown in SD -Ura(-Met) and localisation was examined in live cells during log phase growth. **(A)** The C-terminal domain is not essential for nuclear localisation, but does contribute as N-terminally tagged Lsm8ΔC fails to accumulate strongly in the nucleus; **(B)** The central Sm-domain is not essential or sufficient for nuclear localisation but does contribute as N-terminally tagged Lsm818 shows significantly reduced nuclear localisation; **(C)** The N-terminal domain is not essential for nuclear localisation as LsmΔN88 and Lsm188 still accumulate in the nucleus, although at reduced levels; **(D)** The N-terminal domain of Lsm8p fused to the Sm-domain and C-terminal domain of Lsm1p allows these proteins to accumulate in the nucleus, showing that the N-terminal domain of Lsm8p as part of a functional Lsm protein is sufficient for nuclear accumulation.

9.4.3 The N-terminal ten amino acids of Lsm8p have a function in nuclear localisation

Generally, combining hybrids of Lsm8p and Lsm1p, or other mutations in Lsm8p with an N-terminal GFP tag seemed to affect nuclear accumulation more strongly than when the same mutations were combined with the C-terminal GFP tag (Fig. 9.4 and 9.5). This suggests that the N-terminus plays a role in nuclear accumulation, although I have already shown it is neither essential (Fig. 9.5B) nor sufficient on its own (Fig. 9.3E). To see if these ten residues are sufficient for nuclear accumulation in the context of the Lsm complex they were fused to the Sm domain and/or C-terminal domain of Lsm1p in Lsm81ΔCp and Lsm811p. MPS26 was transformed with plasmids expressing these hybrid proteins, GFP tagged on the N- or C-terminus. Localisation was examined in live cells during log phase growth. These hybrid Lsm proteins carrying only the N-terminal ten amino acids and no other part of Lsm8p accumulated in the nucleus, at least when tagged on the C-terminus (Fig. 9.5D). This indicates that these residues play a central role in nuclear localisation of Lsm8p.

9.5 Lsm1p, Lsm1p mutants and Lsm1/8p hybrids localise to P-bodies

It seems likely that Lsm1p has a nuclear exclusion signal as it is generally assumed that Lsm1p localises exclusively to the cytoplasm. Tharun *et al.* (2000) performed immuno-staining on cellular Lsm1p (HA-tagged on the genome) and showed it to be more prominent in the cytoplasm, whereas others have shown Lsm1p to be excluded from the nucleus altogether (Spiller, 2005). The reason for this apparent contradiction is unknown. Here, GFP-tagged Lsm1p was shown to localise throughout the cell, only excluding vacuoles, making it difficult to look for a nuclear exclusion signal. Apart from its presumed cytoplasmic localisation Lsm1p concentrates in P-bodies under stress conditions (Sheth *et al.*, 2003; Teixeira *et al.*, 2005). To investigate whether any of the Lsm1p domains plays a role in localisation to these foci a similar approach was taken as for Lsm8p (section 9.4). Deletions of N- and/or C-terminal domains were made as well as exchanging N-terminal, C-terminal or Sm domains for those of Lsm8p.

9.5.1 No single domain of Lsm1p is essential for localisation to P-bodies

Plasmids expressing N- or C-terminally GFP tagged mutant or hybrid proteins from the *MET25* promoter were transformed into AEMY25 (*lsm1Δ*). Cells were grown in SD-Ura-Met and localisation was examined in live cells during log phase growth or after hypo-osmotic shock (10-60 min incubation in water). The Lsm1p C-terminal domain seems not to be essential for localisation to P-bodies as Lsm1ΔCp and Lsm118p were shown to localise to P-bodies (Fig. 9.6A). The N-terminal domain and central Sm-domain seem not to be essential either, as LsmΔN11p and Lsm811p (Fig. 9.6B) as well as Lsm181p (Fig. 9.6C) were shown to accumulate in P-bodies under stress conditions. Localisation to P-bodies was however, significantly reduced compared to wild type, as only 5-20% of cells expressing these mutant or hybrid Lsm1 proteins showed foci, compared to 90% of cells expressing GFP-Lsm1. Interestingly, Lsm188p accumulated in the nucleus (Fig. 9.5B; Fig. 9.6D), but it accumulates in cytoplasmic foci under stress conditions. This suggests that the N-terminal domain of Lsm1p in combination with the Sm- and C-terminal domains of Lsm8p is sufficient to allow concentration in P-bodies.

9.5.2 Lsm1p, Lsm1p mutants and Lsm1/8p hybrids co-localise with Dcp2-RFP

To test whether the foci observed for Lsm1ΔCp and Lsm118p were indeed P-bodies, co-localisation with Dcp2p, another P-body component, was investigated. Dcp2-RFP was expressed from its own promoter from pRP1155 (*CEN-LEU2-DCP2-RFP*) and GFP-Lsm1, GFP-Lsm1ΔC, GFP-Lsm118 and GFP were expressed from the *MET25* promoter. Localisation was examined in live cells approximately 60 min after glucose starvation. For this purpose log phase cells grown in SD-Ura-Leu-Met were washed with and resuspended in glucose-free medium and incubated at room temperature. Dcp2-RFP was shown to localise to foci in all cells, and co-localised with GFP-Lsm1, Lsm1ΔC and Lsm118 foci (Fig. 9.7), showing that these mutant and hybrid proteins accumulate in P-bodies under stress conditions.

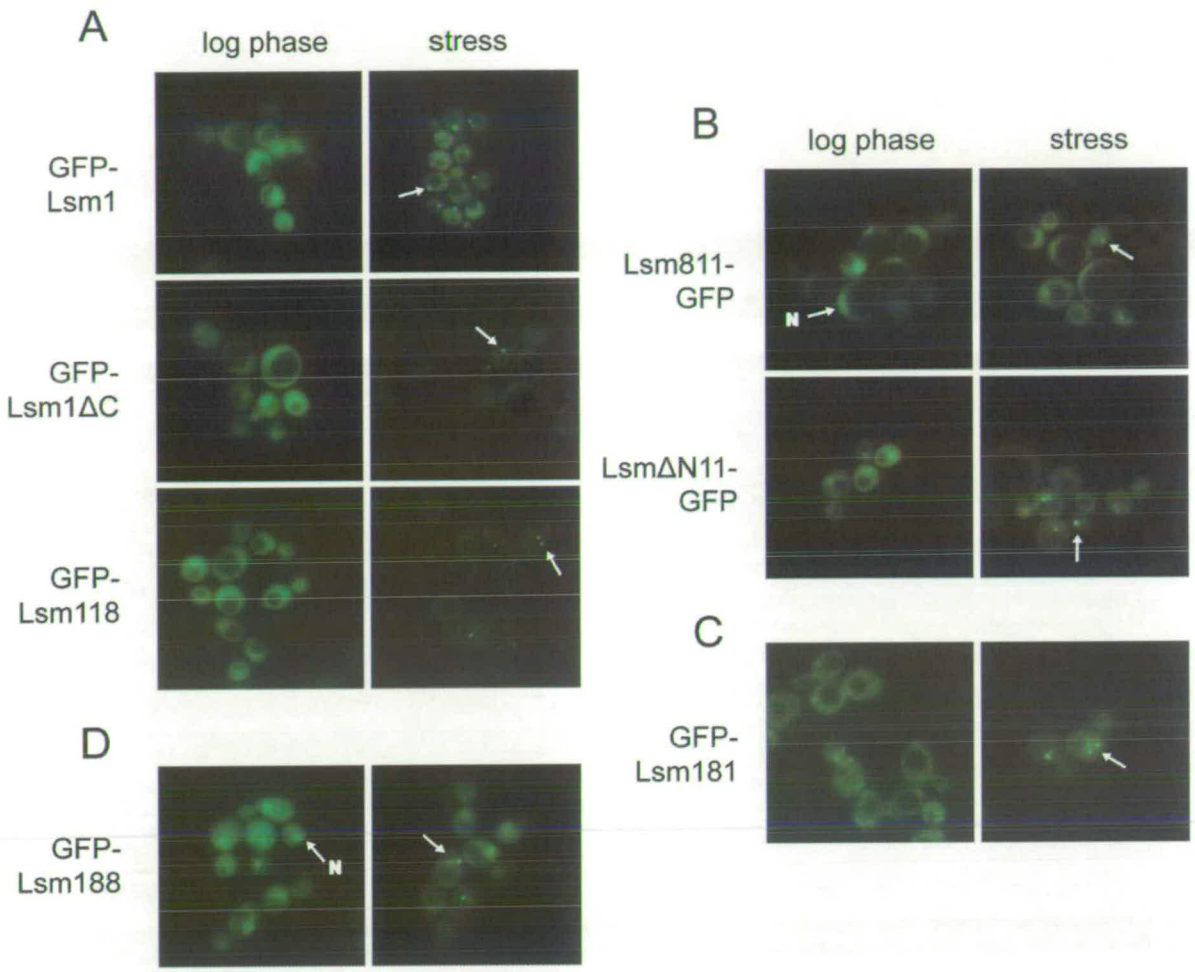


Fig. 9.6 None of the Lsm1p domains are essential for localisation to P-bodies under stress conditions AEMY25 (*lsm1*Δ) was transformed with plasmids **(A)** pGFP-N-Lsm1, pMR79 (pGFP-N-Lsm118) and pMR69 (pGFP-N-Lsm1ΔC); **(B)** pMR124 (pGFP-C-Lsm811) and pMR135 (pGFP-C-LsmΔN11); **(C)** pMR123 (pGFP-C-Lsm181); and **(D)** pMR126 (pGFP-N-Lsm188). Cells were grown in SD-Ura(-Met) and localisation was examined in live cells during log phase growth or after hypo-osmotic shock (stress). **(A)** The C-terminal domain is not essential for localisation to P-bodies as Lsm1ΔCp and Lsm118p still localise to P-bodies (indicated by arrows); **(B)** The N-terminal domain is not essential for P-body localisation as LsmΔN11p and Lsm811p still accumulate in P-bodies under stress conditions, although this is strongly reduced for Lsm811p; **(C)** The central Sm-domain is not essential for P-body localisation; **(D)** Although Lsm188p accumulates in the nucleus (N), the N-terminal domain of Lsm1p seems to be sufficient to allow localisation to P-bodies under stress conditions.

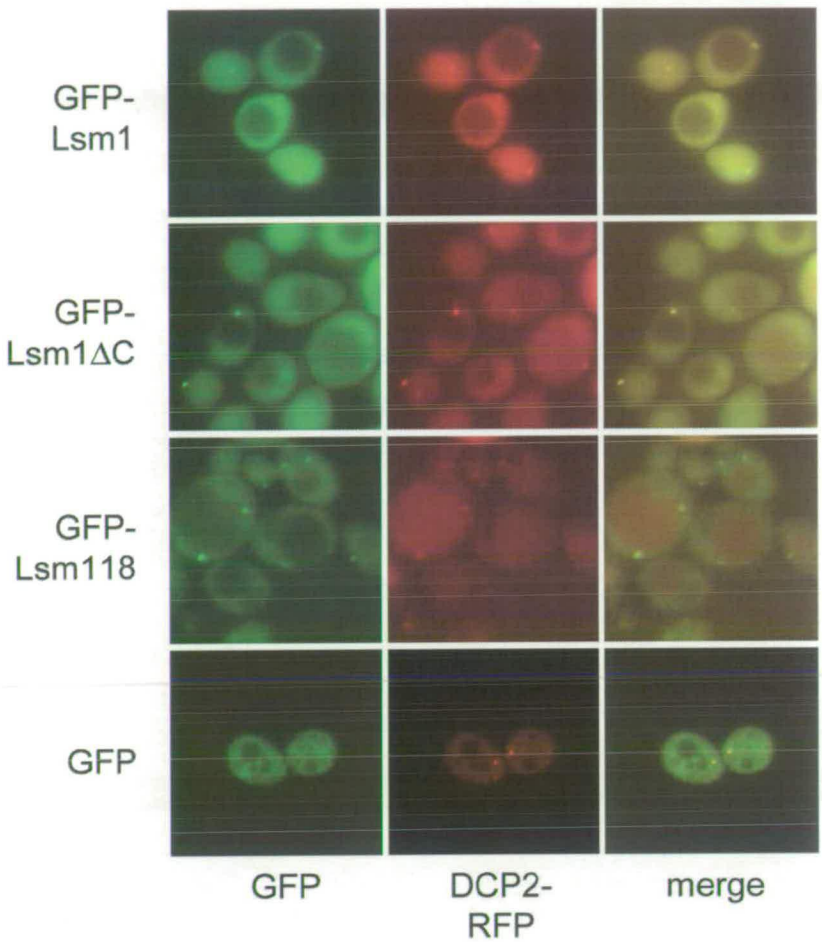


Fig. 9.7 Lsm1p, Lsm1ΔCp and Lsm118p co-localise with Dcp2p in cytoplasmic foci. AEMY25 (*lsm1Δ*) was transformed with pGFP-N-Lsm1, pMR69 (pGFP-N-Lsm1ΔC), pMR79 (pGFP-N-Lsm118) and pGFP-N-FUS together with pRP1155 (*DCP2-RFP*). Cells were grown in SD-Ura-Leu-Met and localisation was examined in live cells 60 min after glucose starvation.

9.6 The Lsm1p and Lsm8p N-terminal domains are required for distinct cellular localisation

It seems that the N-terminal domains of Lsm1p and Lsm8p play a central role in their localisation. To test what happens to localisation of mutant and hybrid proteins in their absence MPS26 was transformed with plasmids expressing N- or C-terminally GFP-tagged LsmΔN11, LsmΔN88, Lsm81ΔC, Lsm18ΔC, LsmΔN18, LsmΔN81, LsmΔN1ΔC and LsmΔN8ΔC mutant proteins from the *MET25* promoter. Cells were grown in SD-Ura-Met and localisation was examined in live cells during log phase growth and after hypo-osmotic shock. In the absence of the N-terminal domain both Lsm1p (LsmΔN11p) and Lsm8p (LsmΔN88p) accumulated in the nucleus as well as in foci (Fig. 9.8A). As Lsm81ΔCp accumulated in the nucleus and Lsm18ΔCp accumulated in P-bodies (Fig. 9.8B), it seems that in the absence of the C-terminal domain the N-terminal domain, not the Sm domain, determines the sub-cellular localisation. These data suggest that the N-terminal domains do indeed play a central role in determining the sub-cellular localisation of the Lsm1 and Lsm8 proteins, and presumably Lsm1-7 and Lsm2-8 complexes. The carboxy-terminal domains seem to have a function in localisation as well. LsmΔN18 mainly accumulated in the nucleus, whereas LsmΔN81 mainly accumulated in P-bodies (Fig. 9.8C). Although in the absence of the N-terminal domain these mutant and hybrid proteins do not have one distinct sub-cellular localisation, the carboxy-terminal domain, not the Sm domain, seems to determine the preferred site of accumulation. The Sm domain of Lsm1p alone accumulated both in nuclei and P-bodies, whereas the Sm domain of Lsm8p alone showed extremely weak fluorescence, but in some cells accumulated in foci (Fig. 9.8D). It therefore appears that in the absence of both N- and C-terminal extensions, the Sm domains of Lsm1p and Lsm8p do not have a strongly preferred localisation site. Their potential for P-body localisation (and nuclear localisation in case of LsmΔN1ΔC) does suggest that they incorporate into the Lsm complexes.

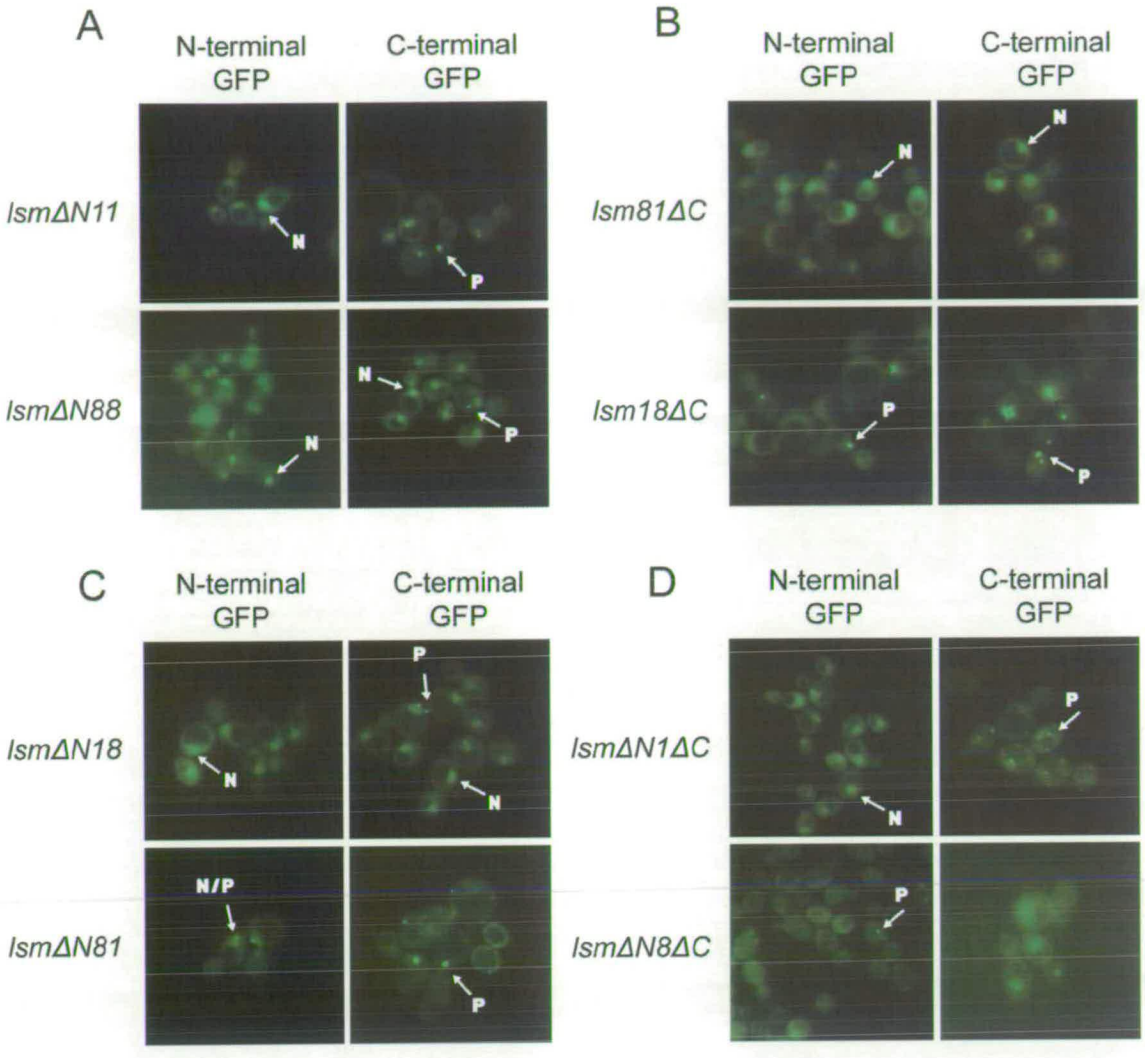


Fig. 9.8 The Lsm1p and Lsm8p N-terminal domains are required for distinct localisation. MPS26 was transformed with plasmids: **(A)** pMR134 and 140 (pGFP-N-LsmΔN11 and LsmΔN88), pMR135 and 141 (pGFP-C-LsmΔN11 and LsmΔN88); **(B)** pMR129 and 137 (pGFP-N-Lsm81ΔC and Lsm18ΔC), pMR130 and 138 (pGFP-C-Lsm81ΔC and Lsm18ΔC); **(C)** pMR143 and 147 (pGFP-N-LsmΔN18 and Lsm ΔN81), pMR145 and 148 (pGFP-C-LsmΔN18 and LsmΔN81); **(D)** pMR150 and 153 (pGFP-N-LsmΔN1ΔC and LsmΔN8ΔC), and pMR151 and 154 (pGFP-C-Lsm ΔN1ΔC and LsmΔN8ΔC). Cells were grown in SD-Ura-Met and localisation was examined in live cells during log phase growth and after hypo-osmotic shock. **(A)** In the absence of the N-terminal domain both Lsm1p and Lsm8p accumulate in the nucleus (N) as well as in foci (P) after osmotic shock; **(B)** Lsm81ΔC accumulates in the nucleus, whereas Lsm18ΔC accumulates in P-bodies; **(C)** LsmΔN18 and LsmΔN18 localise to both, but LsmΔN18 mainly accumulates in the nucleus, whereas LsmΔN81mainly accumulates in P-bodies; **(D)** The Sm domain of Lsm1p accumulates in both nuclei and P-bodies, whereas the Sm domain of Lsm8p shows extremely weak fluorescence, but may also accumulate in foci.

In summary, the N-terminal domains of Lsm1p and Lsm8p seem to most strongly confer preference for localisation to P-bodies or nuclei and seem to prevent localisation of Lsm8p to P-bodies and of Lsm1p to the nucleus. The C-terminal domains seem to play a secondary role as they determine the preferred site of accumulation in the absence of the N-termini. Finally, the central Sm domains do not seem to have a significant effect on localisation, although exchanging these domains between Lsm1p and Lsm8p does lead to reduced localisation to P-bodies and nuclei respectively.

9.7 A correlation between viability and correct localisation

All mutant and hybrid proteins, either GFP-tagged on the N- or C-terminus (pGFP-N-FUS or pGFP-C-FUS respectively) or without a tag (pMET25), were tested for their ability to support viability. Plasmids were transformed into AEMY25 (*lsm1Δ*) and MPS11 (*lsm8Δ* P_{GAL}-*HA-LSM8*) and tested for growth at a range of temperatures by streaking on SD-Ura (low level of expression) and SD-Ura-Met (high level of expression). Growth in MPS11 was tested at 18, 23, 30, 36 and 37°C, whereas growth in AEMY25 was tested at permissive (30°C) and non-permissive temperatures (36 and 37°C). The amount and rate of growth was compared to a positive control (pGFP-N-Lsm8 for MPS11; pGFP-N-Lsm1 for AEMY25) and a negative control (pGFP-N-FUS) and given a score: ++++ for growth like the positive control, – for background growth like the negative control and +++, ++ or + for anything in between.

There seemed to be a clear correlation between viability in *lsm8Δ* and accumulation in the nucleus (Fig. 9.9A; Table 9.1). All mutant and hybrid proteins that accumulated in the nucleus only were able to support viability to some extent. Even some of the proteins that accumulated in the nucleus as well as in P-bodies supported growth. There was a tendency for mutants and hybrids to support growth better at lower (18 and 23°C) than at higher temperatures (≥ 30°C). Mutant and hybrid Lsm8 proteins with a GFP-tag on the Lsm8p N-terminus generally showed less growth than the same proteins with a C-terminal tag or no tag, emphasising the importance of a

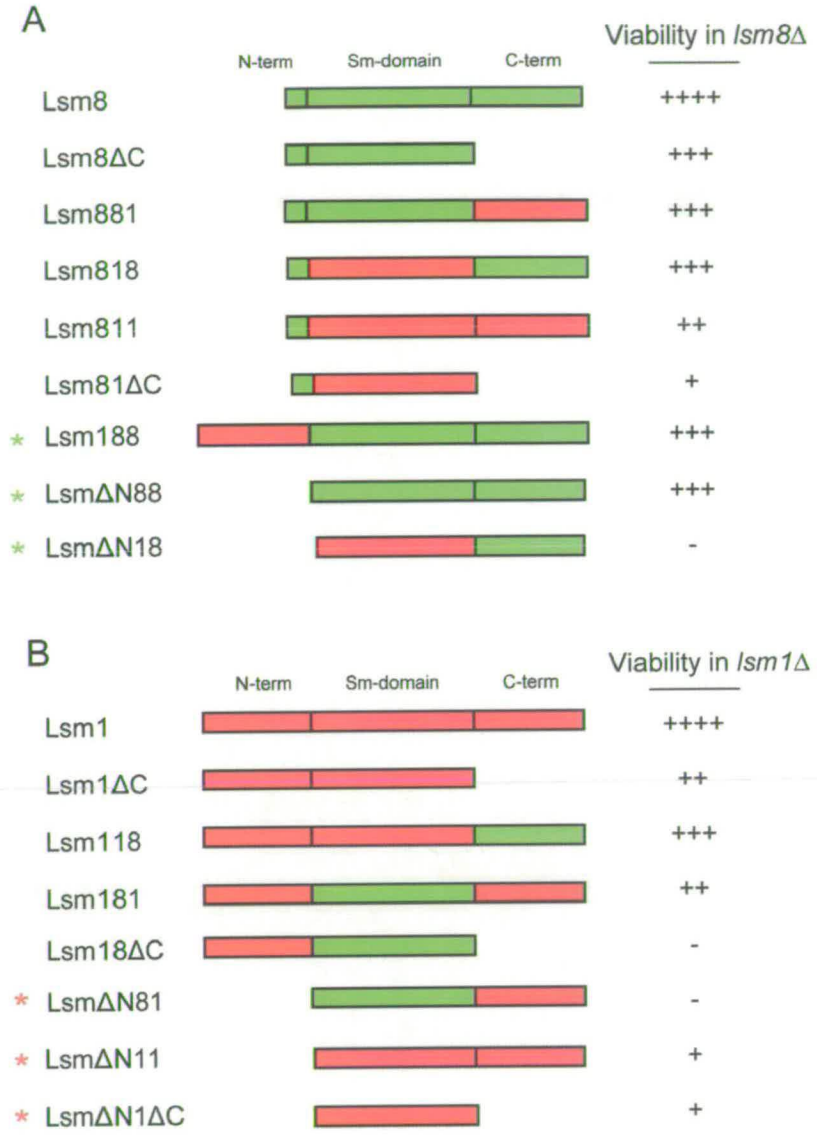


Fig. 9.9 Summary of localisation and viability of Lsm1 and Lsm8 hybrid and mutant proteins. **(A)** Mutant proteins that accumulate in the nucleus. Most proteins that accumulate in the nucleus are also able to support growth of an *lsm8Δ* strain; **(B)** Mutant proteins that accumulate in P-bodies. Most, but not all proteins that accumulate in P-bodies are also able to support growth of an *lsm1Δ* strain at non-permissive temperature. Lsm1p domains are indicated in red; Lsm8p domains are indicated in green. Viability was scored by comparison to the wild-type plasmid (++++) and the GFP only negative control (-). * Proteins that accumulate both in nuclei and P-bodies.

plasmid	Viable in *lsm8Δ* P_{GAL}-*HA-LSM8 LSM7-MYC* (MPS11) background at 18°C in SD -Ura	 -Ura-Met	at 23°C in SD -Ura	 -Ura-Met	at 30°C in SD -Ura	 -Ura-Met	at 36°C in SD -Ura	 -Ura-Met	at 37°C in SD -Ura	 -Ura-Met
pGFP-N-Lsm8	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++
pGFP-C-Lsm8	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++
pMET25-Lsm8	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++
pGFP-N-Lsm8Δ1	-	+	-	+	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm8Δ1	+++	++++	++	+++	-	++++	-	-	-	-
pMET25-Lsm8Δ1	+++	++++	++	++++	-	++++	-	-	-	-
pGFP-N-Lsm8 rna1	++++	++++	++++	++++	++++	++++	-	+++	-	+++
pGFP-C-Lsm8 rna1	++++	++++	++++	++++	++++	++++	++++	++++	+	+++
pMET25-Lsm8 rna1	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++
pGFP-N-Lsm8 rna2	++++	++++	++++	++++	++++	++++	-	+++	-	+++
pGFP-C-Lsm8 rna2	++++	++++	++++	++++	++++	++++	++++	++++	+	+++
pMET25-Lsm8 rna2	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++
pGFP-N-Lsm8 rna3	-	+	-	+	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm8 rna3	++	+++	++	+++	-	++	-	-	n.t.	n.t.
pMET25-Lsm8 rna3	++	+++	++	+++	-	++	-	-	n.t.	n.t.
pGFP-N-Lsm881 rna1	n.t.	n.t.	n.t.	n.t.	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm881 rna1	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pMET25-Lsm881 rna1	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pGFP-N-Lsm881 rna2	n.t.	n.t.	n.t.	n.t.	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm881 rna2	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pMET25-Lsm881 rna2	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pGFP-N-Lsm881 rna3	n.t.	n.t.	n.t.	n.t.	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm881 rna3	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pMET25-Lsm881 rna3	-	-	n.t.	-	-	-	-	-	n.t.	n.t.
pGFP-N-Lsm8ΔC	-	+	-	+	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm8ΔC	+++	++++	++	++++	-	+++	-	-	n.t.	n.t.
pMET25-Lsm8ΔC	+++	++++	++	++++	++	++++	-	-	n.t.	n.t.
pGFP-N-Lsm118	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm118	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pMET25-Lsm118	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pGFP-N-Lsm181	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm181	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pMET25-Lsm181	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.	n.t.	n.t.
pGFP-N-Lsm811	-	++	-	++	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm811	-	++	-	-	-	-	n.t.	n.t.	n.t.	n.t.
pMET25-Lsm811	-	+++	+	++	-	++	-	-	-	-
pGFP-N-Lsm818	-	+	-	+	-	-	n.t.	n.t.	n.t.	n.t.
pGFP-C-Lsm818	-	+++	-	+++	-	+++	-	-	-	-
pMET25-Lsm818	-	+++	+	+++	-	+++	-	+	-	-
pGFP-N-Lsm188	+	+++	+	+++	+	+++	-	+	-	-
pGFP-C-Lsm188	++	+++	++	+++	++	+++	-	+	-	-
pMET25-Lsm188	+	+++	+	+++	+	+++	-	+	-	+
pGFP-N-Lsm18ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-Lsm18ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-Lsm18ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-N-Lsm81ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-Lsm81ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-Lsm81ΔC	-	+	-	+	-	+	n.t.	-	n.t.	n.t.
pGFP-N-LsmΔN11	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-LsmΔN11	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-LsmΔN11	-	++	-	+	-	-	n.t.	-	n.t.	n.t.
pGFP-N-LsmΔN88	++	+++	++	+++	+	+++	-	+	n.t.	n.t.
pGFP-C-ΔN88	++	+++	++	+++	+	++	-	-	n.t.	n.t.
pMET25-LsmΔN88	++	+++	++	+++	++	+++	n.t.	++	n.t.	-
pGFP-N-LsmΔN1ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-LsmΔN1ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-LsmΔN1ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-N-LsmΔN8ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-LsmΔN8ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-LsmΔN8ΔC	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-N-LsmΔN18	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-LsmΔN18	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-LsmΔN18	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-N-LsmΔN81	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-C-LsmΔN81	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pMET25-LsmΔN81	n.t.	-	n.t.	-	n.t.	-	n.t.	-	n.t.	n.t.
pGFP-N-Lsm8ΔC NLS	-	-	-	++	-	+	-	-	n.t.	-
pGFP-C-Lsm8ΔC NLS	+	++++	++	++++	+	+++	-	++	n.t.	-
pMET25-Lsm8ΔC NLS	+	+++	+	++++	+	+++	-	++	n.t.	-
pMET25-lsm8 rna3 NLS	+	+++	++	+++	+	++	-	-	n.t.	-

Table 9.1 Viability of an *lsm8Δ* strain expressing mutant and hybrid Lsm8 proteins. Indicated plasmids were transformed into MPS11 and tested for growth on SD-Ura and SD-Ura-Met at 18, 23, 30, 36 and 37°C. Amount and rate of growth were compared to that for pGFP-N-Lsm8 and pGFP-N-FUS and given a score: ++++ for growth like the positive control, – for background growth and +++, ++ or + for anything in between. n.t. = not tested

plasmid	Viable in *lsm1*Δ (AEMY25) background at 36°C in SD -Ura	-Ura-Met	at 37°C in SD -Ura	-Ura-Met
pGFP-N-Lsm1	++++	++++	++++	++++
pGFP-C-Lsm1	++	+++	-	++
pMET25-Lsm1	++++	++++	+++	++++
pGFP-N-Lsm8$^{1-81}$	-	-	-	-
pGFP-C-Lsm8$^{1-81}$	-	-	-	-
pMET25-Lsm8$^{1-81}$	-	-	-	-
pGFP-N-Lsm1$^{1-8}$	+	+++	-	+
pGFP-C-Lsm1$^{1-8}$	-	+++	-	+
pMET25-Lsm1$^{1-8}$	+	+++	-	++
pGFP-N-Lsm1ΔC	-	++	-	+
pGFP-C-Lsm1ΔC	-	-	-	-
pMET25-Lsm1ΔC	n.t.	n.t.	n.t.	n.t.
pGFP-N-Lsm1$^{81-88}$	n.t.	-	-	-
pGFP-C-Lsm1$^{81-88}$	-	-	-	-
pMET25-Lsm1$^{81-88}$	++	++	-	+
pGFP-N-Lsm8$^{11-81}$	-	-	-	-
pGFP-C-Lsm8$^{11-81}$	++	++	-	+
pMET25-Lsm8$^{11-81}$	-	-	-	-
pGFP-N-Lsm8$^{18-88}$	-	-	-	-
pGFP-C-Lsm8$^{18-88}$	-	-	-	-
pMET25-Lsm8$^{18-88}$	-	-	-	-
pGFP-N-Lsm8$^{11-88}$	-	-	-	-
pGFP-C-Lsm8$^{11-88}$	-	-	-	-
pMET25-Lsm8$^{11-88}$	-	-	-	-
pGFP-N-Lsm8$^{11-81}$	-	-	-	-
pGFP-C-Lsm8$^{11-81}$	-	-	-	-
pMET25-Lsm8$^{11-81}$	+	+	-	-
pGFP-N-LsmΔN11	-	-	-	-
pGFP-C-LsmΔN11	-	-	-	-
pMET25-LsmΔN11	+	+	-	-
pGFP-N-LsmΔN88	-	-	-	-
pGFP-C-LsmΔN88	-	-	-	-
pMET25-LsmΔN88	-	-	-	-
pGFP-N-LsmΔN18	-	-	-	-
pGFP-C-LsmΔN18	-	-	-	-
pMET25-LsmΔN18	+	+	-	-
pGFP-N-LsmΔN81	-	-	-	-
pGFP-C-LsmΔN81	-	-	-	-
pMET25-LsmΔN81	-	-	-	-
pGFP-N-LsmΔN1ΔC	-	-	-	-
pGFP-C-LsmΔN1ΔC	-	-	-	-
pMET25-LsmΔN1ΔC	+	+	-	-
pGFP-N-LsmΔN8ΔC	-	-	-	-
pGFP-C-LsmΔN8ΔC	-	-	-	-
pMET25-LsmΔN8ΔC	-	-	-	-

Table 9.2 Viability of an *lsm1*Δ strain expressing mutant and hybrid Lsm1 proteins at non-permissive temperature. Indicated plasmids were transformed into AEMY25 and tested for growth on SD-Ura and SD-Ura-Met at 36 and 37°C. Amount and rate of growth were compared to pGFP-N-Lsm1 and pGFP-N-FUS and given a score: ++++ for growth like the positive control, – for background growth and +++, ++ or + for anything in between. n.t. = not tested

freely available Lsm8p N-terminal domain. Although the two correlate, reduced viability of Lsm8p mutants and hybrids may not necessarily be caused by reduced nuclear localisation in all cases. Instead they may both have the same cause, e.g. reduced complex formation. To test if increased nuclear localisation would allow increased levels of growth, the nuclear localisation signal (NLS) of simian virus 40 (SV40) large T antigen was added to Lsm8ΔCp and Lsm8-rna3p. However, neither showed increased nuclear accumulation or increased viability (Table 9.2; data not shown), although the fusions were stably expressed (Fig. 9.2). The reason for the failure of the SV40 NLS to localise these Lsm8 mutant proteins to the nucleus is unknown. It should be noted however, that the SV40 NLS could not efficiently replace the NLS-like motif of SmB either, whereas the same SV40 NLS fused to GFP was shown to allow for its nuclear accumulation (Bordonné, 2000). This may suggest that the SV40 NLS is not accessible due to association of the NLS-tagged proteins with other (L)Sm proteins.

The stringency for growth at non-permissive temperature for mutants and hybrids in the *lsm1Δ* background seemed to be higher. Only few mutants and hybrids showed some growth at 36°C, and even fewer at 37°C (Fig. 9.9B; Table 9.2). Although not all mutant and hybrid proteins that were able to accumulate in P-bodies supported growth at non-permissive temperatures, there was still a correlation between P-body localisation and viability, as all proteins that did support growth also accumulated in foci.

9.8 Levels of mutant and hybrid protein expression affect viability

As results in Table 9.1 and 9.2 show, the level of expression of mutant and hybrid proteins has a significant effect on their ability to support growth. Whereas expression of wild-type Lsm1p and Lsm8p in the presence of 1mM methionine (i.e. the MET25 promoter is repressed) still allowed growth at all temperatures, most mutants and hybrids showed a reduction in the amount of growth. Indeed, increased expression of *lsm118* from the *GAL1* promoter (pMR71) showed still more growth in

lsm1Δ at non-permissive temperature than when this hybrid protein was expressed from the *MET25* promoter (pMR89) in the absence of methionine (data not shown). Expression from the *MET25* promoter is significantly lower than from the *GAL1* promoter (Mumberg *et al.*, 1994; Niedenthal *et al.*, 1996). To estimate how much higher the level of expression is from the *MET25* promoter compared to natively expressed *LSM8*, Northern analysis was performed on total RNA from BMA38α (with pGFP-N-FUS) and MPS11 (*lsm8Δ* P_{GAL}-*LSM8*) transformed with pGFP-C-LSM8. Cells were grown in SD-Ura with or without 1 mM methionine to an OD_{600} of 0.5. Total RNA was extracted and used for Northern analysis. The Northern blot was probed with a single stranded DNA probe covering the full LSM8 coding

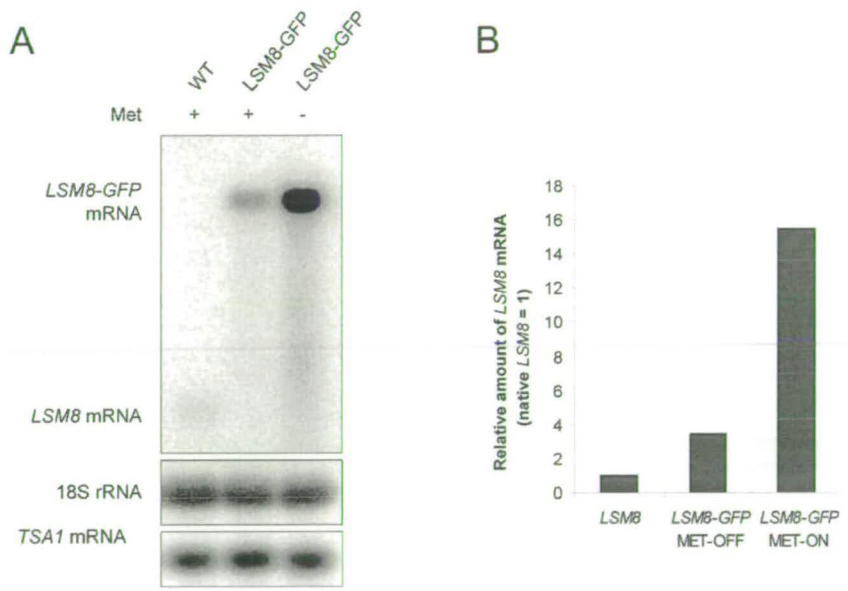


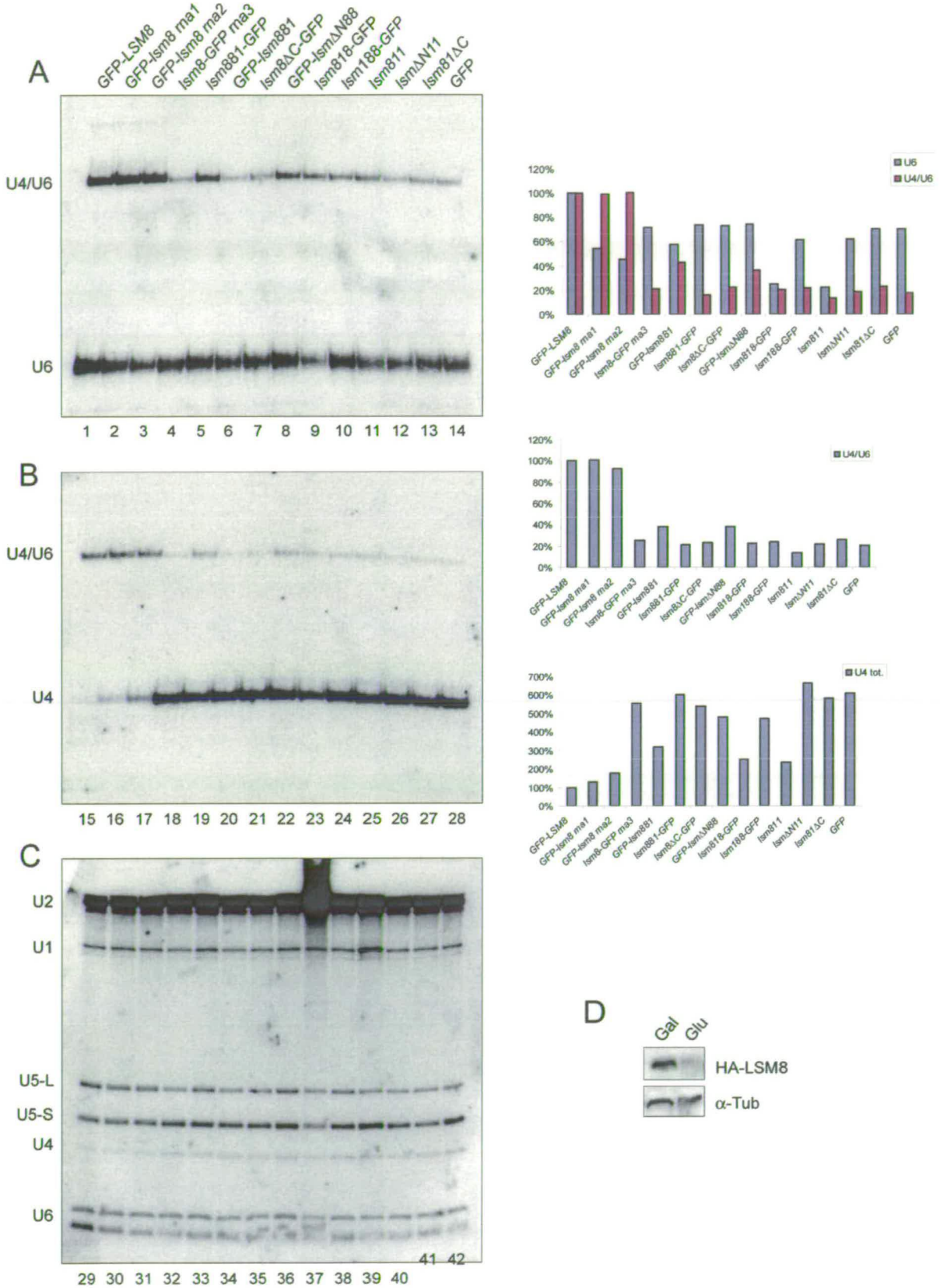
Fig. 9.10 *LSM8* probed Northern comparing the levels of native *LSM8* expression to those of *LSM8-GFP* from the pGFP-C-FUS *MET25* promoter under growth conditions in which the promoter is switched OFF (+Met) or ON (-Met). BMA38α (WT) was transformed with pGFP-C-FUS; MPS11 (P_{GAL}-*LSM8*) was transformed with pGFP-C-LSM8. Cells were grown in SD-Ura with or without 1 mM methionine to OD_{600} of 0.5. Total RNA was extracted and used for Northern analysis **(A)**. The Northern was probed with a single stranded DNA probe covering the full *LSM8* coding sequence. Loading was controlled for by probing the same Northern with oligonucleotide probes against 18S rRNA and *TSA1* mRNA. **(B)** Quantification of bands shows that under MET-OFF conditions the level of *LSM8* expression is approximately 3.5 times higher than expression from the native promoter. Under MET-ON conditions the expression is approximately 4.5 times higher than under MET-OFF conditions, making expression of *LSM8-GFP* under MET-ON conditions 16 times that of native *LSM8*.

sequence. Quantification of bands showed that in the presence of 1mM methionine the level of *LSM8-GFP* mRNA was approximately 3.5 times that of natively expressed *LSM8* mRNA. In the absence of methionine the expression was approximately 4.5 times higher than in the presence of methionine, making expression of *LSM8-GFP* from the activated *MET25* promoter approximately 16 times that of native *LSM8*. It is likely that many of the mutant and hybrid proteins would not support growth when expressed at normal levels. Higher levels of expression may be needed to make up for reduced half lives of mutant and hybrid proteins. Weak fluorescence for some of the mutant and hybrid proteins (particularly some of the deletion mutants) may be indicative of this. Alternatively, increased protein levels increase the chance of at least a fraction of the population to form functional complexes with the other Lsm proteins.

9.9 Effects of Lsm8 mutant and hybrid proteins on levels of spliceosomal snRNAs

To investigate whether Lsm8 mutant and hybrid proteins which are able to support viability also bind and stabilise U6 snRNA, total RNA was extracted from strains expressing these proteins and used for Northern analysis. MSP11, an *lsm8* deletion strain which depends on a *CEN-HIS3* plasmid expressing HA-Lsm8p from the *GAL1* promoter, was transformed with pGFP-N-FUS and pGFP-C-FUS based plasmids expressing the mutant and hybrid proteins from the *MET25* promoter. Cells were grown in SDGal-Ura at 30˚C into mid-log phase before shifting them to SD-Ura-Met and growing them for an additional 10h at 30ºC. To test if cells were depleted from HA-Lsm8p, total protein from cells before and 10 h after the shift to glucose was separated by SDS-PAGE. Western analysis showed an almost complete loss of HA-Lsm8p after 10 h (Fig. 9.11D). After 10 h equal numbers of cells (10 OD_{600} units) were harvested at OD_{600} 0.6-0.8 and total RNA was extracted under non-denaturing conditions. Total RNA was separated by native or denaturing PAGE. Northern analysis of U6 (Fig. 9.11A) and U4 snRNAs (Fig. 9.11B) after non-denaturing PAGE showed decreased levels of U4/U6 RNA for all mutants and hybrids except *lsm8 rna1* and *rna2*. Levels of free U6 snRNA were similarly decreased for all

Fig. 9.11 Levels of snRNAs in *lsm8* mutants and hybrids. **(A)** Northern probed for U6 RNA after non-denaturing PAGE of total RNA shows a decrease in the levels of U6 and U4/U6 RNA for most *lsm8* mutant and hybrids; **(B)** Same Northern probed for U4 RNA confirms the decrease in U4/U6 levels and shows an increase in the levels of total U4 RNA for most *lsm8* mutants and hybrids; **(C)** Northern probed for all five spliceosomal snRNAs after denaturing PAGE; **(D)** Western blot on total protein probed with α-HA antibody shows cells are almost entirely depleted of HA-Lsm8p after 10h of growth on glucose; similar levels of α-Tubulin confirm equal loading. MSP11 (*lsm8Δ* P_{GAL}*-HA-LSM8*) was transformed with *CEN-URA3* plasmids expressing the mutant and hybrid proteins from the *MET25* promoter. Cells were grown in SDGal-Ura into mid-log phase before shifting them to SD-Ura-Met and growing them for an additional 10h at 30°C. Equal numbers of cells (10 OD_{600} units) were harvested at OD_{600} 0.6-0.8 and total RNA was extracted under non-denaturing conditions. Equal amounts were separated by native (A/B) or denaturing (C) PAGE. Bands were quantified and loading was corrected for by comparing levels of U1 snRNA.


A
GFP-LSM8
GFP-lsm8 rna1
GFP-lsm8 rna2
lsm8-GFP rna3
lsm881-GFP
GFP-lsm881
lsm8ΔC-GFP
GFP-lsmΔN88
lsm818-GFP
lsm188-GFP
lsm811
lsmΔN11
lsm81ΔC
GFP
U4/U6
U6
1 2 3 4 5 6 7 8 9 10 11 12 13 14
B
U4/U6
U4
15 16 17 18 19 20 21 22 23 24 25 26 27 28
C
U2
U1
U5-L
U5-S
U4
U6
29 30 31 32 33 34 35 36 37 38 39 40 41 42
U6
U4/U6
120%
100%
80%
60%
40%
20%
0%
GFP-LSM8
GFP-lsm8 rna1
GFP-lsm8 rna2
lsm8-GFP rna3
GFP-lsm881
lsm881-GFP
lsm8ΔC-GFP
GFP-lsmΔN88
lsm818-GFP
lsm188-GFP
lsm811
lsmΔN11
lsm81ΔC
GFP
U4/U6
120%
100%
80%
60%
40%
20%
0%
GFP-LSM8
GFP-lsm8 rna1
GFP-lsm8 rna2
lsm8-GFP rna3
GFP-lsm881
lsm881-GFP
lsm8ΔC-GFP
GFP-lsmΔN88
lsm818-GFP
lsm188-GFP
lsm811
lsmΔN11
lsm81ΔC
GFP
U4 tot.
700%
600%
500%
400%
300%
200%
100%
0%
GFP-LSM8
GFP-lsm8 rna1
GFP-lsm8 rna2
lsm8-GFP rna3
GFP-lsm881
lsm881-GFP
lsm8ΔC-GFP
GFP-lsmΔN88
lsm818-GFP
lsm188-GFP
lsm811
lsmΔN11
lsm81ΔC
GFP
D
Gal
Glu
HA-LSM8
α-Tub

mutant and hybrids, including *lsm8 rna1* and *rna2*, but surprisingly, levels were only decreased by approximately 30-40% for most strains including the GFP only control. In contrast, levels of total U4 snRNA were significantly increased for all mutants and hybrids to levels two to six times that of wild type. The increase was the least for the *rna1* and *rna2* mutants (128% and 177% of wild type respectively) suggesting that the increase may be (partially) related to decreased levels of U4/U6 RNA (see also Chapter 8). Northern analysis of all snRNAs after denaturing PAGE generally confirmed the above findings (Fig. 9.11C; data not shown). Problems with proper separation of the RNA from lsm818-GFP (Fig. 9.11C) suggests that quantification data (Fig. 9.11A and B) for this hybrid may not be meaningful. In addition to reduced levels of U6 RNA and increased levels of U4 snRNA, a change in the levels of U5-L and U5-S snRNA was observed. Levels of U5-L RNA were reduced for all *lsm8* mutant and hybrid strains, whereas levels of U5-S RNA were increased compared to wild type, leading to a decrease in the U5-L over U5-S ratio for all mutants by 20-70% (compared to WT set at 100%). This was also observed for *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains (see Chapter 8), suggesting that this is a general phenomenon in *lsm* mutants, although the mechanism behind this is still unknown.

9.10 Discussion

In this chapter I have shown that the N-terminal domains of Lsm1p and Lsm8p play a central role in their localisation to P-bodies and nuclei respectively. They contribute to targeting these polypeptides to the correct sub-cellular sites as well as preventing accumulation of Lsm8p in P-bodies and of Lsm1p in the nucleus. The C-terminal domains play a secondary role as they determine the preferred site of accumulation in the absence of the N-termini. Finally, the central Sm domains have only a minor effect on localisation, as exchange of these domains between Lsm1p and Lsm8p leads to reduced localisation to P-bodies and nuclei respectively. In the absence of the N- and C-terminal domains however, the Sm domains do not show a preferred site of accumulation, and in hybrids of Sm domains of one protein with terminal extensions of the other it is the terminal domains that determine localisation specificity. RNA binding specificity of the Lsm1-7 and Lsm2-8 complexes may be a

contributing factor to their specific localisation. Limited effects of mutations in putative RNA binding residues of Lsm1p (Tharun *et al.*, 2005) and Lsm8p, and exchange of the Sm domains argue against this, unless the terminal extensions affect RNA binding specificity, e.g. by interacting with nearby stem-loop structures. Lsm2-8p has been shown to be instrumental in nuclear accumulation of the U6 snRNA (Spiller *et al.*, 2007), but my results, particularly those for *lsm8 rna3*, suggest that binding of Lsm2-8p to U6 snRNA may not be essential for nuclear accumulation of the Lsm proteins. However, residues outside the Sm-fold may contribute to RNA binding specificity. The RNA binding pockets formed by residues within the Sm-fold give the Lsm complexes specificity for a U-rich single stranded RNA sequence, but the presence of an adjacent stem-loop structure was proposed to stabilise the RNA-protein interaction (Vidal *et al.*, 1999; this thesis). It is the latter interaction to which other residues, for example those in the N-terminal domains, could contribute.

In terms of functionality Lsm8p seems to be more resistant to mutations than Lsm1p, as N- or C-terminal domains can be removed or replaced with those of Lsm1p and still support growth. Also, a mutant strain in which the central Sm domain is replaced with that of Lsm1p is still viable. Emphasising the functional flexibility of Lsm8p and the importance of the N-terminal ten amino acids of Lsm8p is the fact that Lsm811p, a hybrid with these ten residues fused to the Sm and C-terminal domains of Lsm1p, allows some growth of an *lsm8Δ* strain. Mutations in Lsm1p more readily affect viability, as deletion or replacement of any of the Lsm1p domains strongly affects viability at non-permissive temperature. A mutant strain expressing Lsm118p, in which the Lsm1p C-terminus is replaced with that of Lsm8p, shows most growth at *lsm1Δ* non-permissive temperatures. This is surprising as the C-terminal domain was suggested to be functionally important (Tharun *et al.*, 2005). Similarly affected growth for C-terminally tagged Lsm1p does indeed suggest that the C-terminal domain may have an essential function which is affected by addition of a tag. However, strongly reduced growth for Lsm1ΔCp suggests that part of this function may be to contribute to complex formation and/or protein stability which may be reduced in the absence of a C-terminal domain. There is a correlation between correct localisation of the mutant and hybrid proteins and their ability to support

growth. However, not all proteins that accumulate at the correct sites support viability. Obviously correct localisation is not sufficient if the mutant protein does not support essential functions of the corresponding complex. In addition, it has been suggested that temperature sensitivity of the *lsm1Δ* strain is related to Lsm1-7p function in protection of shortened 3' mRNA termini and not to Lsm1-7p function in decapping in P-bodies (He *et al.*, 2001; Tharun *et al.*, 2005). The difference in apparent resistance to effects of mutagenesis between Lsm1p and Lsm8p may be partially explained by the fact that Lsm8p is an essential protein and Lsm1p is not. Most of the mutant and hybrid Lsm8 proteins displayed heat-sensitive growth, but allowed growth at lower temperatures. Heat-sensitivity of some of the mutants and hybrids may be caused by reduced thermodynamic stability of these proteins or of the complex they are part of. This reduces the chance of the Lsm1p mutants and hybrids supporting growth at *lsm1Δ* non-permissive temperatures (36 or 37°C).

As stabilisation of U6 snRNA has been suggested to be the only essential function of the Lsm2-8 proteins (Pannone *et al.*, 2001), the Lsm8p mutants and hybrids that support viability presumably bind and stabilise U6 snRNA sufficiently. It was shown however, that there was only a weak correlation between levels of U6 and U4/U6 RNA and viability of the corresponding strain, suggesting that additional functions of Lsm8p may contribute to the growth phenotypes of the mutant strains. The U4/U6 levels were affected more strongly for most mutants than what was observed for *lsm5Δ*, *lsm6Δ* and *lsm7Δ* strains (see Chapter 8), whereas, surprisingly, free U6 levels were affected less for most. The reason for this is unclear, but may reflect early effects of Lsm8p depletion on U4/U6 levels with U6 RNA levels decreasing more slowly over time. In contrast, virtually unaffected U4/U6 levels for the *rna1* and *rna2* mutants suggest that Lsm2-8p function in U4/U6 di-snRNP formation is not affected for these mutants. Significantly reduced U6 RNA levels on the other hand, suggest that their ability to bind and stabilise free U6 RNA is affected. This seems slightly contradictory as RNA binding is thought to be essential for Lsm2-8p in promoting U4/U6 annealing. The reason for this is unclear, but may indicate that protein-RNA contacts outside the single-stranded U-rich tail are more important for

U4/U6 annealing whereas loss of the interaction with the U6 3' end is important in protecting U6 snRNA from degradation.

Effects of the Lsm8 mutant and hybrid proteins on U4 snRNA levels and U5-L/U5-S ratios suggest that the nuclear Lsm2-8p complex may be required for efficient exosome-mediated degradation of U4 snRNA as well as U5-S RNA (see also section 8.6).

Dissection of the Lsm1p and Lsm8p proteins has proved useful in determining domains important for localisation of these polypeptides. Further examination of the specific effects some of these mutants and hybrids may have on particular processes may teach us more about how the Lsm1-7 and Lsm2-8 complexes function.

CHAPTER 10

Final discussion

10.1 Sm-like proteins are RNA chaperones

Studies in the budding yeast *S. cerevisiae* have implicated the Lsm proteins in a wide range of RNA metabolic processes. The nuclear Lsm2 to 8 proteins, that form the Lsm2-8 complex and possibly alternative nuclear complexes, are required for efficient processing of pre-rRNAs, pre-tRNAs and pre-snoRNAs. In addition, Lsm2-8p is required for nuclear pre-mRNA decapping and subsequent 5' to 3' degradation. Arguably the most important function of Lsm2-8p is stabilisation of the spliceosomal U6 snRNA. The Lsm2 to 8 proteins form the core of the U6 snRNP; they are stably associated with the 3' end of U6 snRNA, protecting it from degradation, and are important for efficient association of the U4 and U6 snRNAs in the U4/U6 di-snRNP. Fernandez *et al.* (2004) argued that defects in the processing of other stable RNAs are indirect effects of reduced U6 RNA levels in the absence of a fully functional Lsm2-8 complex, as overexpression of *SNR6* suppressed some of them. This issue is unresolved.

Although depletion of Lsm proteins affects many different RNA processing events, the exact mechanism behind these effects still remains elusive. The eukaryotic Lsm proteins were suggested to function as RNA chaperones, modulating RNA-RNA and RNA-protein interactions. Later, the bacterial Hfq protein, which has limited sequence homology to the (L)Sm proteins, was shown to form a similar ring-like structure. This Sm-like protein functions as an RNA chaperone: e.g. it promotes interactions between small regulatory RNAs and their target mRNAs. This, together with the ability of the native human LSm2-8 complex to promote U4/U6 di-snRNP formation has led to the widely accepted view that all Sm-like proteins function as RNA chaperones. One definition of an RNA chaperone is a protein that assists in RNA folding by preventing misfolding or by resolving misfolded RNA structures (Herschlag, 1995). RNA chaperones have also been described as proteins that stimulate intermolecular annealing events between complementary nucleic acids (Tsuchihashi and Brown, 1994). As no direct evidence existed that eukaryotic Lsm proteins have either of these abilities, their RNA chaperone activity was still open for debate.

I have now shown that *H. sapiens* Lsm1-7p and Lsm2-8p, as well as *E. coli* Hfq are able to promote annealing between U4 and U6 RNAs *in vitro*, providing the first direct evidence that eukaryotic Lsm complexes are indeed RNA chaperones[4], just like their bacterial counterparts. Failure to show the same activity for the yeast Lsm complexes is attributed to technical difficulties in assembling functional complexes from purified recombinant proteins *in vitro*. Two models have been proposed for the way in which Hfq promotes annealing between two different RNA molecules. In the first, Hfq binds the two RNAs simultaneously, thus increasing their chances to interact. In the second, Hfq destabilises secondary structures in one of the RNAs, that would otherwise inhibit intermolecular basepairing interactions (reviewed by Valentin-Hansen *et al.*, 2004; see Fig. 5.1). Evidence exists for both models, suggesting that Hfq is able to employ either mechanism. To investigate the mode of action of the human and yeast Lsm complexes a strand displacement assay was developed. A stable secondary structure in the 3' half of U6 snRNA has been proposed to inhibit association with U4 snRNA (Fortner *et al.*, 1994; Brow *et al.*, 1995). The substrate in the strand displacement assay was therefore based on the 3' stem-loop, to test if Lsm2-8p is able to destabilise this structure. Hfq as well as human and yeast Lsm1-7p and Lsm2-8p displayed strand displacement activity, suggesting that each of these Sm-like complexes is able to destabilise RNA structure. I propose that this is a passive process in which binding of single stranded RNA along the inner pore of the ring-shaped complex bends the RNA, which in turn affects nearby RNA structures (Fig. 5.10). This model is supported by absence of strand displacement activity for an Hfq mutant that is impaired in RNA binding. However, strand displacement by the yeast Lsm complexes, but their failure to promote U4/U6 annealing suggests that RNA binding alone may not be sufficient. As human LSm4 was required and sufficient for both annealing and unwinding activities this protein may play a crucial role in the RNA chaperone function of the Lsm complexes. Recombinant yeast Lsm4p was also shown to promote U4/U6 annealing (it was not tested for strand displacement), suggesting that inactivation of this

[4] The definition of a molecular chaperone includes the condition that the chaperone does not form part of the final, folded structure. As Lsm2-8p forms part of the U4/U6 di-snRNP, it does not conform to the strictest definition of a chaperone. On the other hand, the U4/U6 duplex is stable once formed, even in the absence of the Lsm proteins.

polypeptide under denaturing conditions may be the reason for the absence of annealing activity for the *in vitro* assembled yeast Lsm complexes. Lsm4p may be involved in additional contacts with the U6 RNA, possibly the 3' stem-loop structure which I (and others) have shown stabilises the protein-RNA interaction. Interestingly, Vidal *et al.* (1999) showed that yeast Lsm4p crosslinks directly to U6 snRNA in yeast extracts and that it co-immunoprecipitates the U6 snRNA 3' half (nt 50-112) efficiently, but not U6(93-112). A central function for Lsm4p in the RNA chaperone activity of Lsm complexes awaits confirmation. Lsm4p mutants with point mutations in putative RNA-binding residues or residues involved in inter-subunit contacts may aid in this.

Lsm2-8p and Prp24p are both essential for efficient U4/U6 annealing *in vivo*. Lsm2-8p has been proposed to promote Prp24p-U6 RNA interactions (Rader *et al.*, 2002; Ryan *et al.*, 2002). Based on genetic interactions with mutations in U4 and U6 snRNA, that either promote or inhibit their annealing, each of the four Prp24p RRMs was previously proposed to function in either U4/U6 association or dissociation. I have now shown that mutations in conserved residues of each of the RRMs confer synthetic lethality with deletions of *LSM6* or *LSM7*. I therefore propose that all four Prp24p RRMs work together with Lsm2-8p to anneal U4 and U6 RNA (Fig. 6.5). Recent work by Karaduman *et al.* (2006) shows that binding of these proteins together leads to a U6 structure that is more favourable for annealing to U4 than when one or both of them are absent. The U4/U6 duplex in a Prp24p-U4/U6 complex formed *in vitro* was shown to be incomplete (Ghetti *et al.*, 1995; Jandrositz *et al.*, 1995). This suggests that Lsm2-8p, in addition to direct effects on the U6 structure, may promote complete U4/U6 annealing by destabilising the Prp24p-U4/U6 interaction.

Both Lsm1-7p and Lsm2-8p were able to promote strand displacement as well as U4/U6 annealing. This indicates that both complexes can function as RNA chaperones and that this activity may have only limited specificity. The latter is in line with the wide range of substrates these complexes have. Destabilisation of RNA

secondary structures in mRNA by Lsm1-7p may affect binding of proteins that are involved in decapping and decay.

10.2 Lsm protein promiscuity

The above shows that the Lsm complexes are promiscuous in their interactions with RNA, at least more so than the homologous Sm complexes. This was proposed to be due to different residues involved in formation of the nucleotide binding pockets (Khusial *et al.*, 2005) and emphasises the functional diversity of the Lsm proteins compared to the Sm proteins. Inter-subunit interactions of the Lsm proteins also seem to be more promiscuous than those of the Sm proteins. Formation of stable yeast Lsm3p homomultimers (Collins *et al.*, 2003; this thesis) and formation of various ring-like structures by different combinations of human LSm proteins (Zaric *et al.*, 2005) are but two examples. This underscores the considerable technical difficulties in assembling functional heteroheptameric complexes from recombinant subunits *in vitro*. Co-expression of seven Lsm proteins from a heptacistronic construct may overcome some of these problems, and may be helpful in allowing more detailed functional and structural examinations of the Lsm complexes.

Higher eukaryotes have the SMN complex which assists in proper snRNP assembly *in vivo*, thus preventing unwanted protein-protein and protein-RNA interactions. As *S. cerevisiae* is not known to have an equivalent RNP chaperone, this begs the question how budding yeast prevents undesirable interactions. The only proteins that have been implicated in snRNP assembly in yeast are Lhp1p (Xue *et al.*, 2000) and Brr1p (Noble *et al.*, 1996b). Their precise role in snRNP assembly is however, poorly understood and may not be related to that of the SMN complex in higher eukaryotes.

10.3 Potential new functions of Lsm proteins

Sm-like proteins occur in all three kingdoms of life. Most, but not all, bacterial genomes encode one Hfq or Hfq-like protein, whereas a few have been found to have two copies (reviewed by Valentin-Hansen *et al.*, 2004). Most archaeal genomes have

similarly been found to encode one or two Sm-like proteins that form homomultimeric complexes containing six or seven subunits (reviewed by Beggs, 2005). In contrast, upwards of 15 (L)Sm proteins have been found in eukaryotes. *S. cerevisiae* contains 15 known (L)Sm proteins and has at least 5 other proteins that, by sequence homology, are thought to form an Sm-like fold. Trypanosomes on the other hand have 17 (L)Sm proteins (Liu *et al.*, 2004; Tkacz *et al.*, 2006), with one Lsm ring apparently performing functions in both pre-mRNA splicing and mRNA decapping, and various Sm rings, with U snRNP-specific compositions. Finally, metazoa have 17 or more (L)Sm proteins, with complexes equivalent to those in budding yeast and at least one additional (L)Sm complex specific to the U7 snRNP The increasing number of (L)Sm proteins up the evolutionary ladder suggests multiple gene and possibly genome duplication events that allowed their functional diversification. The presumed advantage of having multiple (L)Sm proteins is the increase in the number of different functions that can be carried out. Intriguingly however, a wide (if not wider) variety of functions is similarly carried out by the single homohexameric Hfq protein. It seems nonetheless likely that each of the (L)Sm subunits has a specific function to play within the context of the complexes they are part of. One of the most striking examples is that of the Lsm1-7 and Lsm2-8 complexes, in which exchange of a single subunit leads to different localisations and functions. Interesting questions that remain include: Are there alternative complexes apart from the two that are known in yeast? If so, are they hexamers (e.g. Lsm2-7p), do they contain different combinations of Sm and Lsm proteins or do they even contain some of the other putative Lsm proteins?

With regards to the question of subunit specific contributions to Lsm complex functions our lab has made some progress over the past few years. Lsm6p and Lsm7p were shown to be required for efficient U4/U6 di-snRNP and U4/U6-U5 tri-snRNP assembly respectively (Verdone *et al.*, 2004). The underlying mechanism for this is unknown, but may involve protein-protein interactions of Lsm6p with di-snRNP and Lsm7p with tri-snRNP specific proteins. I have now uncovered specific functions for the Lsm4 polypeptide. As discussed above, it may be essential for the RNA chaperone function of the Lsm complexes. Additionally, I have shown that the

unusual C-terminal extension of Lsm4p shows prion-like behaviour and that this quality is important for fast and efficient accumulation of Lsm1-7p in P-bodies. Finally, a cell morphology defect observed in cells depleted of Lsm4p alludes to an unknown role for this protein in cytokinesis. As I have confirmed the non-essential nature of the Lsm5 protein, studying its specific functions is now more feasible. A dramatic increase in the level of U4 snRNA in the *lsm5Δ* strain suggests that Lsm5p may have a task in degradation of this small stable RNA. A similar effect of Lsm8p depletion suggests the involvement of the nuclear Lsm2-8 complex, and a smaller but consistent effect on U5-S indicates this snRNA may also be degraded in an Lsm2-8p dependent fashion. As these RNAs were shown to be degraded by the nuclear exosome, Lsm2-8p may act to stimulate and/or recruit exonuclease activity, possibly by affecting TRAMP activity. As the level of scR1 was significantly increased in an *lsm1Δ* strain, the Lsm1-7 complex may play a comparable role in promoting degradation of stable RNAs in the cytoplasm.

With this in mind, it is interesting to note that a strong yeast two-hybrid interaction was detected between Lsm2p and Air2p, which is a subunit of the TRAMP complex (Fromont-Racine *et al.*, 2000). This interaction could however, not be confirmed by co-immunoprecipitation (LaCava *et al.*, 2005), suggesting that a potential interaction may be transient. Defects in pre-tRNA, -rRNA and -snoRNA processing in *lsm* mutant strains are also consistent with a role for Lsm proteins in TRAMP/exosome mediated processing and/or degradation of these RNAs. Intriguingly, *E. coli* Hfq stimulates polyadenylation of mRNA by PAP1 (Hajnsdorf *et al.*, 2000; Le Derout *et al.*, 2003; Folichon *et al.*, 2005) and promotes degradation by PNPase, and can even be found in a complex with these two enzymes (Mohanty *et al.*, 2004). The polyadenylation and degradation of nuclear RNAs in eukaryotes by the TRAMP and exosome complexes may be an analogous mechanism preserved throughout evolution (Houseley *et al.*, 2006). Could Lsm2-8p have taken on the role of Hfq in promoting polyadenylation, in this case by the TRAMP complex?

Other parallels exist between Hfq and the Lsm proteins (reviewed by Wilusz and Wilusz, 2005), emphasising that we may learn more about the eukaryotic Lsm

proteins by looking at functions and mechanisms of their bacterial counterparts. One area of considerable interest is that of small regulatory RNAs. As Hfq is essential for the function of many of these sRNAs it is a real possibility that the eukaryotic Lsm complexes play a similar role in the function of miRNAs. As budding yeast does not have RNAi machinery, it may instead employ ncRNAs for post-transcriptional regulation in a manner equivalent to the Hfq-sRNA system in bacteria. This is subject to ongoing investigations in our lab.

10.4 Localisation of Lsm1-7p versus Lsm2-8p

It is not known what determines the different cellular localisations and functions of the Lsm1-7 and Lsm2-8 complexes. Involvement of the different domains of Lsm1p and Lsm8p in localisation of these proteins to the cytoplasm (to P-bodies under stress conditions) and nucleus respectively was investigated by creating (deletion) mutants and hybrids of various domains of these proteins. Their localisations suggest that the N-termini of both proteins play a central role in targeting them to their respective cellular locations. Surprisingly, exchange of the central Sm-domains, that contain the putative RNA-binding residues, had limited effects on localisation and functional competence of the resulting hybrid proteins. This suggests that protein-RNA interactions outside the Sm-domain may play a crucial role in the possible RNA-binding specificity of these complexes. The C-terminal extensions of both polypeptides were shown to be non-essential and to play a secondary role in determining localisation. In spite of the central role of the N-terminal extensions in localisation of the Lsm complexes, they do not function as localisation signals, as neither was sufficient for localisation of GFP to nuclei or P-bodies. This suggests that other Lsm polypeptides have a task in localisation. I have shown that at least one of these is Lsm4p. It is as yet unclear how functions of the N-termini of Lsm1p and Lsm8p are connected to a role for the Lsm4p C-terminus in localisation to P-bodies. One could imagine Lsm8p masking the aggregation-prone residues and exposing residues that could function as a nuclear localisation signal instead. The Lsm4p C-terminus was shown to be non-essential for Lsm8p nuclear localisation, but its deletion did lead to reduced accumulation in the nucleus (data not shown). Lsm1p on

the other hand may, together with Lsm4p, form a nuclear exclusion signal. The aggregation-prone residues may be similarly masked in Lsm1-7p, and, conceivably, a conformational change triggered by post-translational modifications under stress conditions may expose these, leading to fast accumulation in P-bodies. However, as the Lsm4p C-terminus is not essential for P-body localisation and Hfq was even seen to accumulate in foci together with Dcp2p (see appendix A3), the ability to bind RNA may be sufficient to allow P-body localisation of cytoplasmic Lsm complexes.

10.5 Final conclusions

The *in vitro* studies reveal that the RNA chaperone activities of Sm-like proteins appear to have been conserved from bacteria to eukaryotes. The *in vivo* results begin to distinguish functions specific to individual Lsm polypeptides and protein domains, however, undoubtedly others remain to be found. One of the difficulties is that complex formation is thought to be essential for all Lsm functions. The successful approach with Lsm1p and Lsm8p hybrids and deletion mutants suggests that a similar approach involving extensions of the other Lsm proteins (in particular the essential ones) may facilitate further functional analyses.

Appendices

A.1 Effects of Hfq co-purification on RNA binding, annealing and strand displacement by yeast Lsm proteins

I performed a large number of RNA bandshift, annealing and strand displacement experiments with purified recombinant yeast Lsm proteins without realising that some of them, Lsm8p in particular, also contained co-purified *E. coli* Hfq. Some discrepancies between results of these experiments lead me to investigate the additional protein band in the His-HA-Lsm8p preparation as discussed under 3.6.1. Results obtained for the Lsm-Hfq mixtures are difficult to interpret as it is, in most cases, impossible to distinguish between activities from Lsm, Hfq and possibly even mixed complexes. Later experiments, discussed in Chapters 4 and 5, allowed some distinction between these activities by comparing Lsm complexes purified from *hfq* deletion strains and recombinant Hfq. To put these later experiments in perspective results of earlier experiments with the Lsm-Hfq mixtures are presented and discussed here.

A.1.1 Effects of Hfq co-purification on RNA binding of yeast Lsm proteins

Recombinant His-tagged yeast Lsm proteins 1 to 8 were expressed in BL21(DE3) cells and purified by Ni-NTA affinity chromatography as described in Chapters 2 and 3. These purified proteins were used together with *in vitro* transcribed gel-purified U6 RNA in RNA bandshift assays to examine the ability of each of the Lsm polypeptides to interact with U6 snRNA. At protein concentrations in large excess over RNA Lsm5p, Lsm6p and Lsm8p seem to bind U6 RNA efficiently, with some others showing minor shifts (Fig. A1.1A). A comparison of the shift in mobility in the presence of these proteins to that observed for Prp24p (data not shown) suggested that Lsm multimers rather than Lsm monomers bind U6 RNA. Some, or maybe all, of the Lsm proteins may form homo-multimeric complexes some of which may bind RNA. Lsm3p was reported to form stable homo-multimeric rings (Collins *et al.*, 2003; this thesis, Fig. 3.4C) and additional bands after SDS-PAGE (and Coomassie staining or Western analysis) were at times observed for a number of the other Lsm proteins (data not shown) suggesting that some of these interactions are also

sufficiently stable to withstand heating in the presence of SDS. The observed absence of significant levels of U6 RNA binding by Lsm3p (Fig. A1.1A) is in agreement with the observation of Collins *et al.* (2003) that Lsm3p does not have high affinity for U-rich RNA. When lower, equimolar amounts of each of the recombinant Lsm polypeptides were used in a U6 RNA bandshift experiment only Lsm8p showed significant RNA binding (Fig. A1.1B). A mix of Lsm2 to 8 proteins showed a shift similar to that observed for a mix of Lsm5, 6 and 8p only (data not shown), indicating that an Lsm2-8p complex may not form under these conditions as discussed in Chapter 3. A similar shift was also observed with *in vitro* assembled Lsm2-8p (Fig. A1.2A), but not with Lsm1-7p (data not shown). The shift for Lsm2-8p occurred at a relatively low protein concentration, suggesting that a functional complex had formed. At higher protein concentrations multiple bands with decreased mobility were observed in RNA bandshift experiments (e.g. compare Lsm8p bandshift in Fig. A1.1A and B). This may be indicative of aggregation of complexes as discussed in 3.5.3, or may indicate multiple complexes interacting separately with one RNA molecule. Aggregation seems to be a common feature of Sm-like complexes as multimers of ring-shaped complexes were observed for yeast Lsm3p (Collins *et al.*, 2003), a number of archaeal Sm-like proteins (Mura *et al.*, 2003a; Mura *et al.*, 2003b) and *E. coli* Hfq (Arluison *et al.*, 2006).

Similar bandshifts by the *in vitro* assembled Lsm2-8 complex were observed for U4 RNA, U6rc RNA (reverse complement of U6 RNA) and U2Δ107 RNA (a shortened version of U2 snRNA, see Table 2.1.7.5; Fig. A1.1C-E), suggesting that RNA binding was not specific to U6 RNA. The RNA binding activities of the Lsm5p, Lsm6p and Lsm8p protein preparations were shown to be due to the presence of Hfq (Fig. 4.1A). It is unclear whether co-purified Hfq homohexamers were responsible for RNA binding or whether heteromultimeric complex containing both Hfq and Lsm proteins may have been present. Slight differences in U6-protein complex mobility between protein preparations suggest that the latter may have been the case.

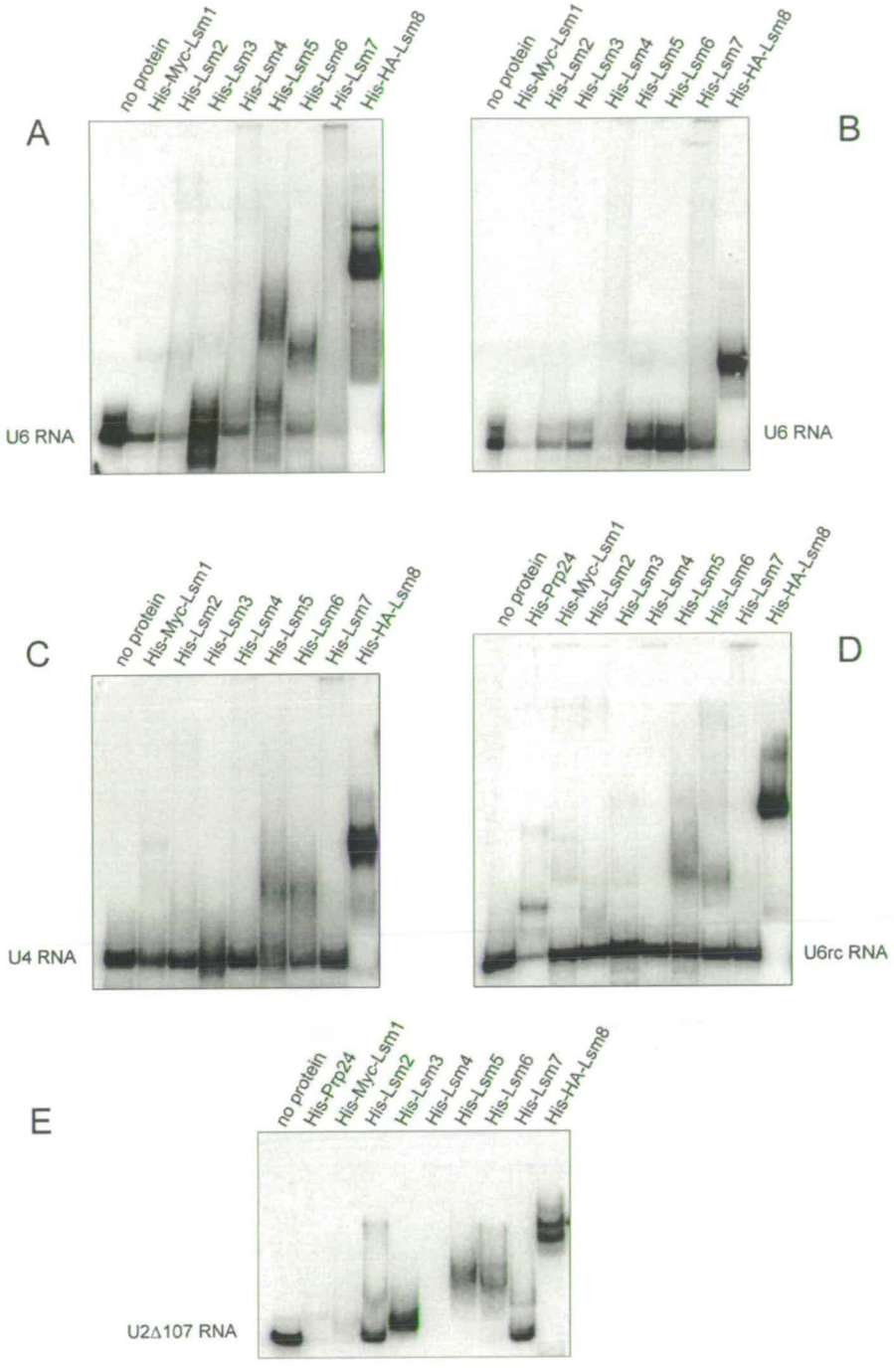


Fig. A1.1 RNA binding by recombinant yeast Lsm proteins purified by affinity chromatography from BL21(DE3) *hfq*+ strain. **(A)** Bandshift with 1.3 pmol (50 ng) of U6 RNA and excess amounts of purified recombinant Lsm proteins: 9 pmol of Lsm1, 160 pmol of Lsm2, 730 pmol of Lsm3, 26 pmol of Lsm4, 535 pmol of Lsm5, 540 pmol of Lsm6, 64 pmol of Lsm7 and 200 pmol of Lsm8; **(B)** Bandshift with 1.3 pmol of U6 snRNA and 25 pmol of each of the individual recombinant Lsm proteins; **(C)** Bandshift assay with 1.8 pmol of U4 RNA and individual Lsm proteins as under (A); **(D)** Bandshift assay with 1.3 pmol of U6 reverse complement RNA and individual Lsm proteins as under (A); **(E)** Bandshift assay with 1.3 pmol of U2Δ107 RNA and individual Lsm proteins as under (A).

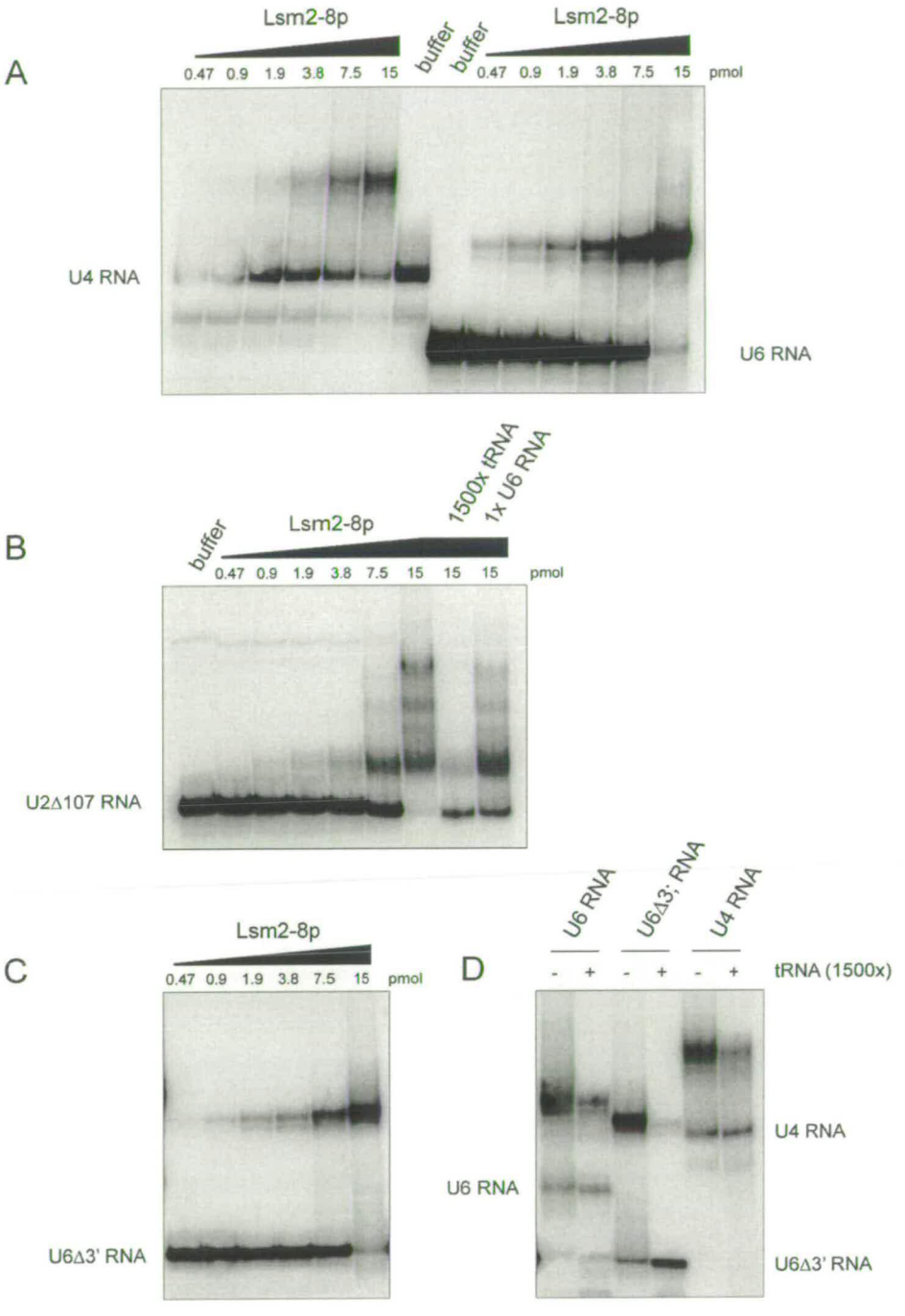


Fig. A1.2 RNA binding by *in vitro* assembled Lsm2-8p from purified recombinant components expressed in BL21(DE3) *hfq*+ strain. **(A)** Bandshift with 1.3 pmol of U4 or U6 RNA and 15 pmol of Lsm2-8p and 2-fold serial dilutions thereof; **(B)** Bandshift with 1.3 pmol of U2Δ107 RNA with same concentration range of Lsm2-8p and with excess tRNA or 1.3 pmol of U6 RNA; **(C)** Bandshift with 1.3 pmol of U6Δ3' RNA; **(D)** Bandshift assay with 1.3 pmol of U6, U6Δ3' or U4 RNA and 15 pmol of Lsm2-8p with and without excess competitor tRNA (75 µg).

In vitro assembled Lsm2-8p with natively purified His-HA-Lsm8 and 8 M urea purified His-Lsm2 to 7 (3.5.1) presumably contained approximately one seventh of the concentration of Hfq that was present in the His-HA-Lsm8 preparation. I assume that Hfq homohexamers and Lsm heteromultimers formed during renaturation, although formation of Hfq-Lsm heteromultimers cannot be ruled out. This mixture was used in RNA bandshift assays with U4, U6, U2Δ107 and U6Δ3' RNA (Fig. A1.2A-C) and protein was shown to bind all these RNAs. There seemed to be a slightly higher affinity for U6 RNA than for the other RNAs tested. Addition of 1500x excess competitor tRNA showed that there was a level of specificity for binding to U4 and U6 RNA, but slightly less for U6Δ3' RNA (Fig. A1.2D), suggesting that a U-rich sequence, in this case an Sm or Lsm binding site, may be favoured.

A.1.2 Effects of Hfq co-purification on U4/U6 annealing activity of yeast Lsm proteins

Recombinant yeast Lsm proteins purified from BL21(DE3) cells were used in a U4/U6 annealing assay (Fig. A1.3A). A small amount of U4/U6 duplex was formed in the absence of any protein. Lsm5p and Lsm8p (hfq+) preparations promoted formation of U4/U6 duplex most efficiently. Interestingly, the mobility of this duplex was slightly faster than that of the duplex formed in the absence of protein. This seems to be caused by Hfq and is discussed in more detail in Chapter 4. The fact that a small amount of the "faster" duplex was formed in the presence of Lsm1, Lsm3 and Lsm6p suggests that these proteins contained small amounts of co-purified Hfq as well. Lsm8p (hfq+) was also shown to promote annealing between U4 and U6Δ3' RNA (data not shown) indicating that the uridine-rich single stranded tail at the 3' end of U6 snRNA is not essential for annealing by this protein preparation. Neither a mixture of natively purified Lsm2 to 8p (hfq+) nor *in vitro* assembled Lsm2-8p (hfq+) displayed significantly higher annealing activities than the Lsm8p (hfq+) preparation alone, suggesting that all annealing activity is caused by co-purified Hfq. Absence of any annealing activity in Lsm8p purified from a *hfq* deletion strain and similar formation of a "fast" U4/U6 duplex with recombinant Hfq and Lsm8p from a *hfq*+ strain (Fig. A1.3B) confirms this. Interestingly, U4/U6 duplex formed by

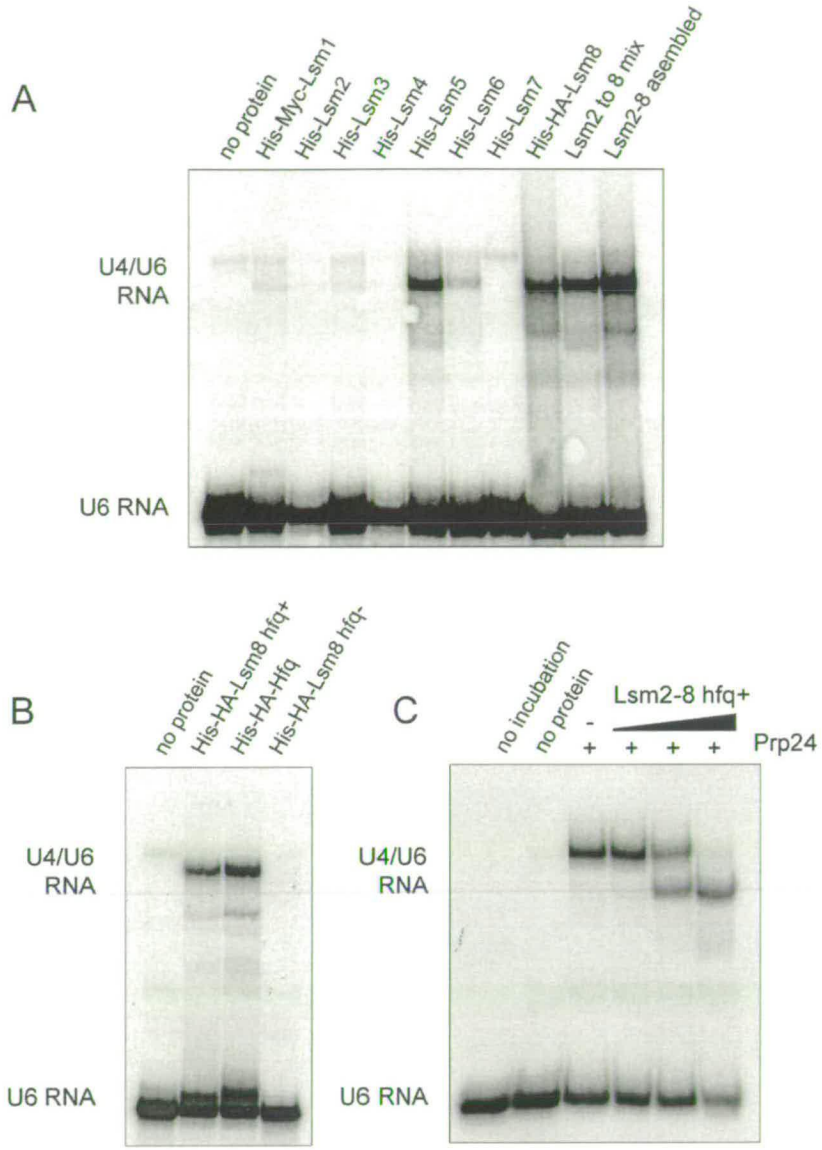


Fig. A1.3 U4/U6 annealing by individual yeast recombinant Lsm proteins and *in vitro* assembled Lsm2-8p from a BL21(DE3) *hfq*+ strain is due to co-purification with *E. coli* Hfq. **(A)** U4/U6 annealing with 0.2 pmol of U6 RNA, 0.3 pmol of U4 RNA and 17 pmol of Lsm1, 21 pmol of Lsm2, 97 pmol of Lsm3, 4.5 pmol of Lsm4, 71 pmol of Lsm5, 72 pmol of Lsm6, 8.5 pmol of Lsm7, 26 pmol of Lsm8, a mixture of Lsm2 to 8 (containing 1/7 of each of the above concentrations) and 3.5 pmol of *in vitro* assembled Lsm2-8; **(B)** U4/U6 annealing with 26 pmol Lsm8p purified from a *hfq*+ strain, 5 pmol of recombinant Hfq (monomers) and 49 pmol of Lsm8p purified from a Δ*hfq* strain; **(C)** U4/U6 annealing with 7.5 pmol of recombinant Prp24p and 2.25 pmol of Lsm2-8p (Hfq+) or 2-fold dilutions thereof.

recombinant Prp24p, which migrates like spontaneously formed duplex, was converted to the more mobile duplex when Hfq (co-purified with Lsm8p) was added (Fig. A1.3C). These data strongly suggests that Hfq affects the structure of the U4/U6 duplex (discussed in 4.3.2.2).

A.1.3 Effects of Hfq co-purification on strand displacement activity of yeast Lsm proteins

A potential mechanism for the annealing activity of Sm-like proteins is destabilisation of intramolecular structures that inhibit intermolecular basepairing interactions. Experiments investigating this mechanism are discussed in detail in Chapter 5. Initial strand displacement experiments with the Lsm-Hfq protein mixtures showed that either Hfq or Lsm complexes or both are able to "unwind" a RNA/DNA duplex resembling the U6 3' stem-loop structure. Increasing amounts of *in vitro* assembled Lsm2-8p (hfq+) lead to increasing amounts of strand displacement. Addition of 1 mM ATP to these reactions did not increase strand displacement activity (Fig. A1.4A). Surprisingly, Lsm8p (hfq+) only displayed very little activity although more Hfq was present in these reactions. The denaturing/renaturing steps of *in vitro* Lsm2-8p assembly therefore seem to lead to increased strand displacement activity. It is unclear whether this is due to formation of (more) Hfq homohexamers or formation of active Lsm2-8p complex or both. Strand displacement activity for Lsm2-8p (hfq+) was shown not to be caused by RNA degradation (Fig. A1.4B). Both *in vitro* assembled Lsm1-7p (hfq+) and Lsm2-8p (hfq+) were shown to "unwind" the duplex over time (Fig. A1.4C). Lsm2-8p seems to initially bind the substrate and subsequently displaces the DNA oligonucleotide. Strand displacement by Lsm1-7p (hfq+) and Lsm2-8p (hfq+) was shown not to be entirely specific for the U6-like substrate (Fig. A1.5). However, both complexes (or co-purified Hfq) did seem to have a preference for substrates with the U-rich single-stranded tail, particularly when the sequence in the stem-loop was replaced with a non-U6 sequence (Fig. A1.5A). Results of strand displacement assays to which competitor RNA (tRNA, U4, U6Δ3' or U6 RNA) was added indicate that binding and "unwinding" of the substrate depends at least in part on specific binding of the U-rich single stranded 3' tail and U6(-like) 3' stem-loop (Fig. A1.5C).

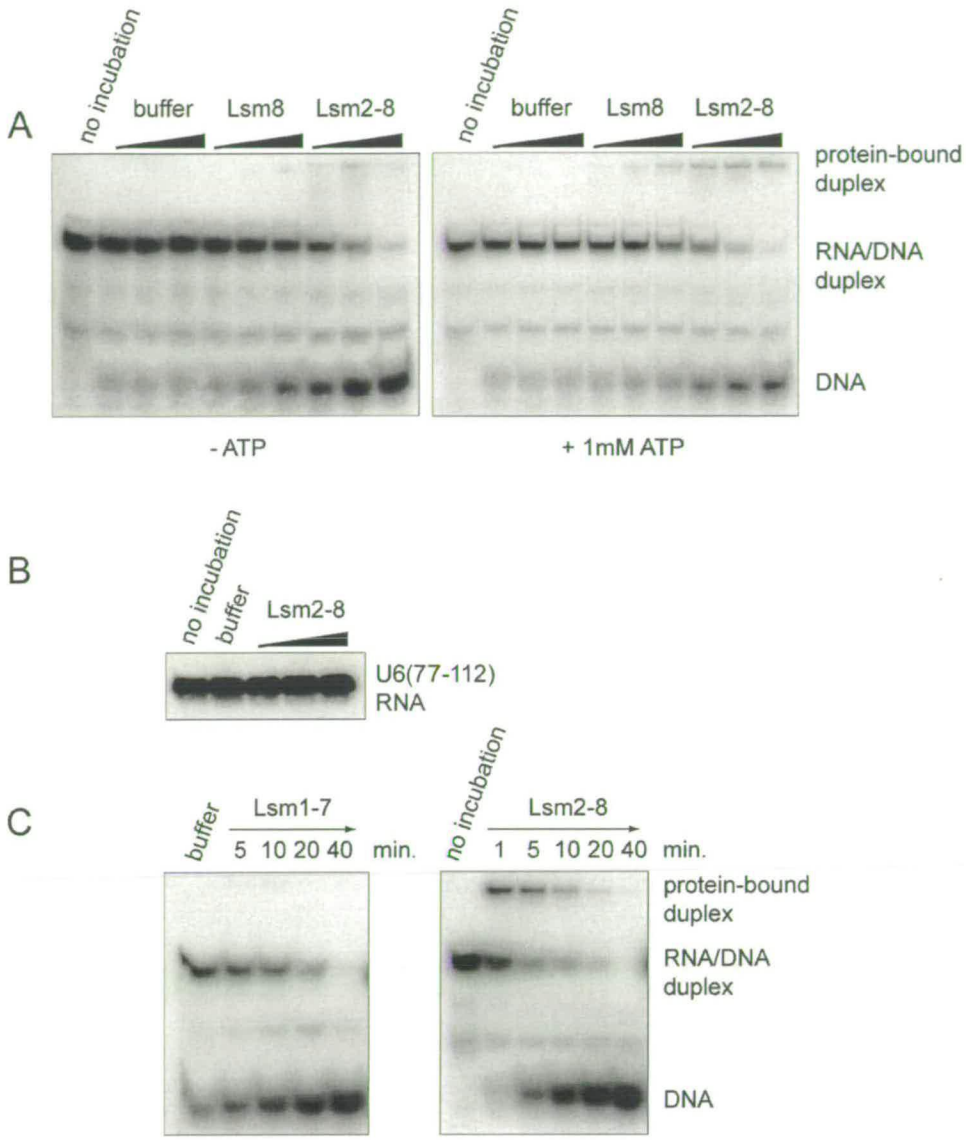


Fig. A1.4 Strand displacement activity of *in vitro* assembled Lsm1-7p and Lsm2-8p from a BL21(DE3) *hfq*+ strain may be due to co-purification with *E. coli* Hfq. **(A)** Strand displacement of 250 nM of U6(77-112)/(77-92)rc substrate with 7.2 μM of Lsm8 (monomers; Hfq+) and 2.4 μM of Lsm2-8 (Hfq+) and 2-fold dilutions thereof, after 10 min of incubation with and without 1 mM ATP; **(B)** Strand displacement by Lsm2-8 (Hfq+) is not due to RNA degradation. Duplicates of samples from (A) were separated by denaturing PAGE and U6(77-112) RNA was detected by Northern blot analysis; **(C)** Both *in vitro* assembled Lsm1-7p (Hfq+; 600 nM) and Lsm2-8p (Hfq+; 600 nM) display strand displacement of 250 nM of substrate over time.

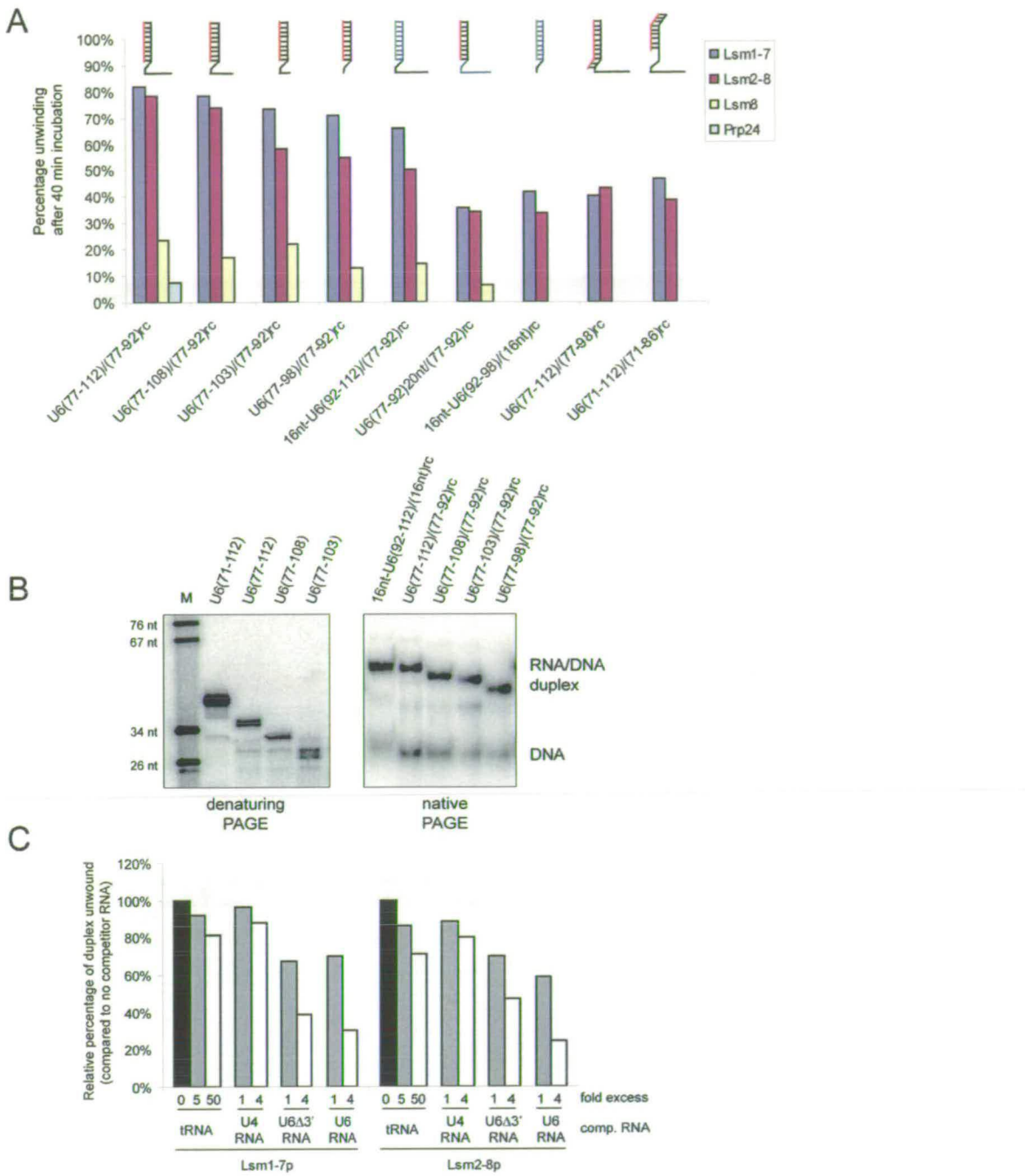


Fig. A1.5 Strand displacement activity of *in vitro* assembled Lsm1-7p and Lsm2-8p (Hfq+) shows a low level of specificity. **(A)** Quantification of strand displacement of various substrates at 250 nM by 600 nM of recombinant Lsm8p (hexamers, Hfq+), Lsm1-7p, Lsm2-8p or Prp24p after 60 min of incubation. Corrected for background unwinding in the absence of protein; **(B)** Various *in vitro* transcribed RNA fragments and RNA/DNA duplexes (separated by denaturing and native PAGE respectively) that were used as substrates in the strand displacement assays; **(C)** Relative strand displacement of U6(77-112)/(77-92)rc by Lsm1-7p and Lsm2-8p in the presence of various competitor RNAs. Unwinding in absence of competitor RNA was set to 100%.

Lsm2-8p complex that was assembled *in vitro* from recombinant Lsm2 to 8 proteins all purified under denaturing conditions displayed strand displacement, but no U4/U6 annealing activity (data not shown). This was one of the main reasons to investigate the additional band in the native Lsm8p purification, which turned out to be Hfq, as this band was absent in Lsm8p purified under denaturing conditions. This also suggested that Hfq was responsible for the annealing activity, but not for all strand displacement activity, which was later confirmed (Chapters 4 and 5).

A.2 Hfq interaction with Lsm8p

Hfq co-purifies with recombinant Lsm8p. What is the nature of this interaction? Do Hfq and Lsm8p form a heteromultimeric ring or other structure or do both proteins interact independently and simultaneously with the same RNA? Purified recombinant His-Myc-Hfq and His-HA-Lsm8p do not interact *in vitro* under immunoprecipitation conditions (Fig. A2.1). This may mean that stable homomultimeric complexes formed by purified Hfq and Lsm8p do not allow heteromultimeric complexes to form or that there is no RNA present to bridge the interaction. Purified recombinant His-HA-Lsm8p with co-purified Hfq shifts U6 RNA in a bandshift assay, whereas His-HA-Lsm8p from an *hfq* deletion strain does not (Fig. 4.1A). If this is due to a heteromultimeric Hfq-Lsm8p complex it would be expected that U6 RNA should co-immunoprecipitate with this complex. Anti-HA or α-His antibody was used to immunoprecipitate His-HA-Hfq or His-HA-Lsm8p purified from a wild-type or a *Δhfq* strain in the presence of *in vitro* transcribed U6 RNA (Fig. A2.2). Hfq immunoprecipitated with either antibody efficiently pulls down U6 RNA, whereas Lsm8p with co-purified Hfq only pulls down a small amount of U6 when immunoprecipitated with α-His antibody. This may mean that the tags are partially masked in a potential Lsm8p-Hfq heteromultimer or that the Hfq-U6 interaction is stronger than that of Lsm8p-Hfq-U6. In contrast, Lsm8p alone does not pull down any U6 RNA confirming the U6 bandshift data. In summary, Lsm8p and Hfq may form a heteromultimeric complex that binds U6 RNA.

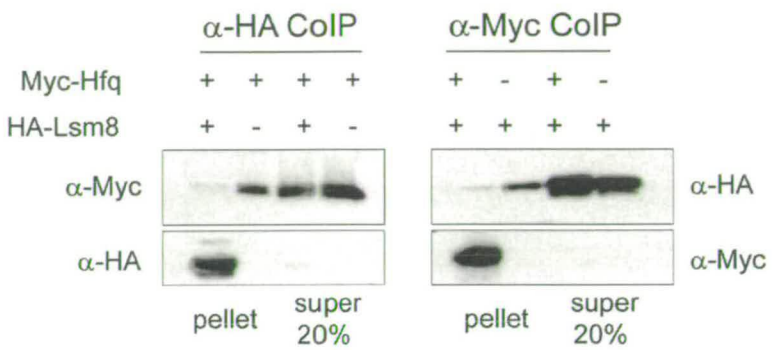


Fig. A2.1 Recombinant Hfq and Lsm8p do not interact *in vitro*. Approximately 0.5 μg of His-Myc-Hfq and His-HA-Lsm8p were mixed and immunoprecipitated with α-Myc or α-HA-conjugated agarose beads. Resulting pellets and supernates were separated by SDS-PAGE and Western blots were hybridised with α-Myc or α-HA antibodies. His-HA-Lsm8p and His-Myc-Hfq are efficiently and completely pulled down by the corresponding antibodies, but no Hfq or Lsm8p is co-immunoprecipitated. Pull downs are even reduced compared to the beads only negative controls.

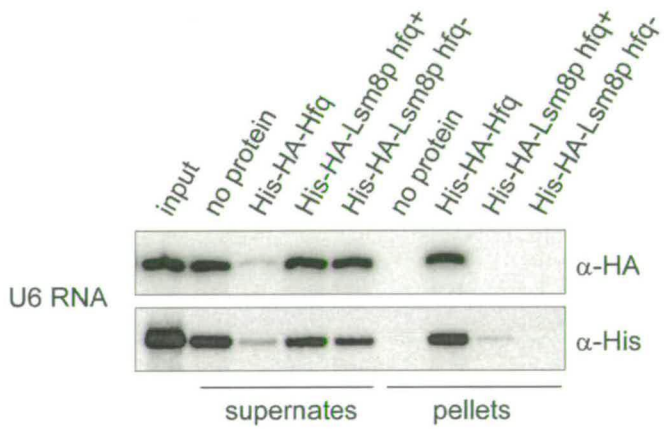


Fig. A2.2 His-HA-Lsm8p and native Hfq may form heteromultimeric complexes. Approximately 0.5 μg of purified recombinant protein (from *hfq+* or *hfq-* strains) was immuno-precipitated with α-HA or α-His antibodies in the presence of 7.5 ng of U6 RNA. RNA from pellets and supernates was separated by denaturing PAGE and visualised by hybridising the Northern blot with labelled U6-specific probe.

A.3 Effects of *hfq* expression in yeast

The interaction between yeast Lsm8p and Hfq (A2) as well as structural and functional similarities between Hfq and the Lsm complexes raises the question what effect *hfq* overexpression may have on yeast and what effect *LSM* overexpression may have on *E. coli*. In particular, can Hfq rescue *lsm* deletion phenotypes by carrying out some of functions of the Lsm complexes or does it interfere with these functions by interacting with Lsm8p and/or other Lsm proteins? GFP-Hfq (pMR128) was overexpressed in BY4741 cells and localisation was examined during log phase growth and after hypo-osmotic shock (Fig. A3.1). GFP-Hfq localised throughout the cytoplasm and was excluded from the nucleus (Fig. A3.1A). After hypo-osmotic shock GFP-Hfq aggregated strongly in cytoplasmic foci and co-localised to these foci with Dcp2-RFP, suggesting that Hfq localises to P-bodies under stress conditions. This presumably means Hfq interacts with poly(A) tails of mRNAs that aggregate in P-bodies and suggests that there may not be an essential P-body localisation signal in Lsm1-7p, unless Hfq is associated with the yeast Lsm proteins.

GFP-Hfq overexpression from the *MET25* promoter did not have any significant effects on growth of wild-type yeast or *lsm* deletion strains. The exception was *lsm7Δ* in which GFP-Hfq overproduction partially alleviated temperature-sensitive growth (data not shown). Overexpression of Hfq from the *MET25* promoter from a high copy 2μ plasmid (pMR127) however, was toxic in BY4741 (Fig. A3.2B) and BMA38a wild-type cells (data not shown). Surprisingly this effect was not as strong in *lsm1*, *6* or *7* deletion strains (Fig. A3.2A). Hfq expression had the least effect on *lsm7Δ* and even had a slightly positive effect under some conditions. In summary, Hfq seems to interfere with essential processes in wild-type cells, but may be able to partially functionally replace Lsm1-7p. Hfq may bind to shortened poly(A) tails and protect them from degradation.

Overexpression of individual Lsm proteins in *E. coli* did not have any discernible negative effects on its growth, nor was it able to suppress effects of an *hfq* deletion (data not shown).

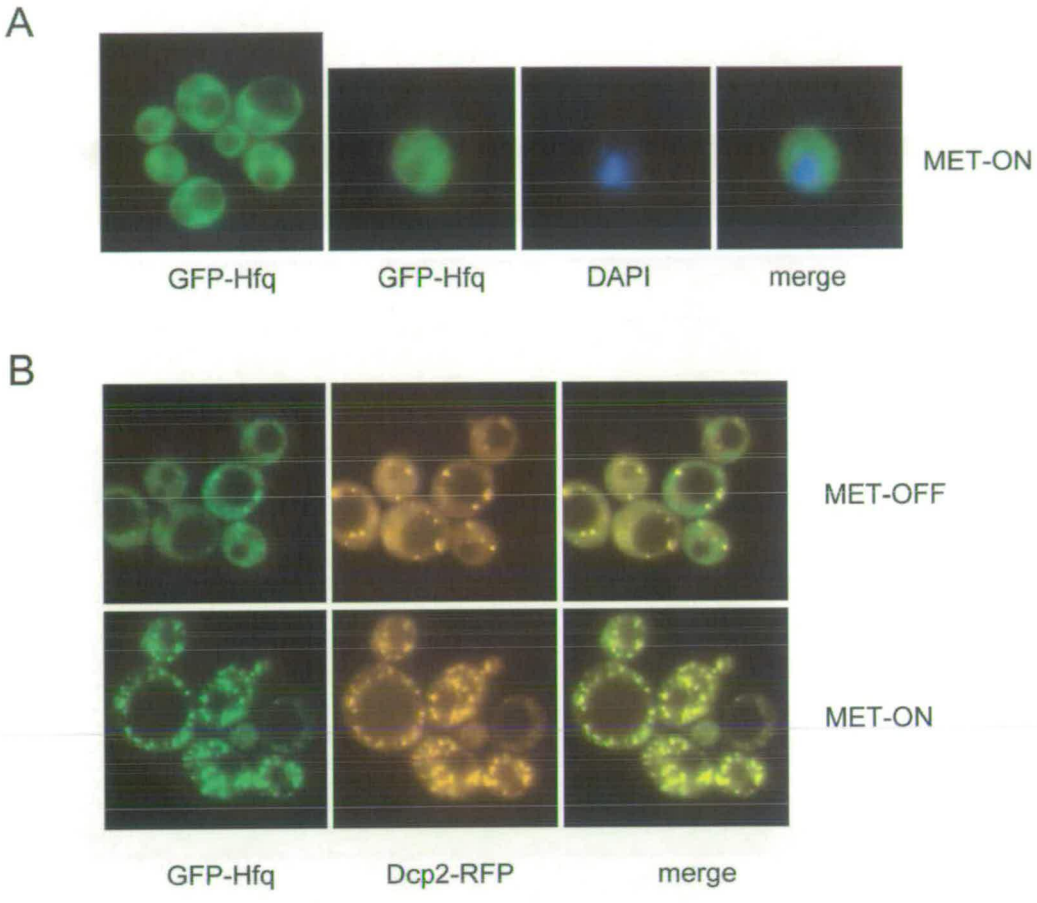


Fig. A3.1 GFP-Hfq localises to the cytoplasm in yeast and localises to P-bodies under stress conditions. **(A)** GFP-Hfq overexpressed from the *MET25* promoter (pMR128) localises throughout the cytoplasm and is excluded from the nucleus as can be seen in live cells and fixed cells in which nuclear DNA is DAPI-stained; **(B)** GFP-Hfq co-localises with Dcp2-RFP after hypo-osmotic shock. Under growth conditions in which P_{MET25} is repressed (MET-OFF) GFP-Hfq localises weakly to foci that contain Dcp2-RFP; under MET-ON conditions GFP-Hfq accumulates strongly in foci together with Dcp2-RFP. BY4741 was transformed with pMR128 (GFP-Hfq) and pRP1155 (Dcp2-RFP). Cells were grown in SD-Ura-(-Leu)(-Met) and localisation was examined in live (B) or fixed cells (A) during log phase growth or 10-30 min after hypo-osmotic shock.

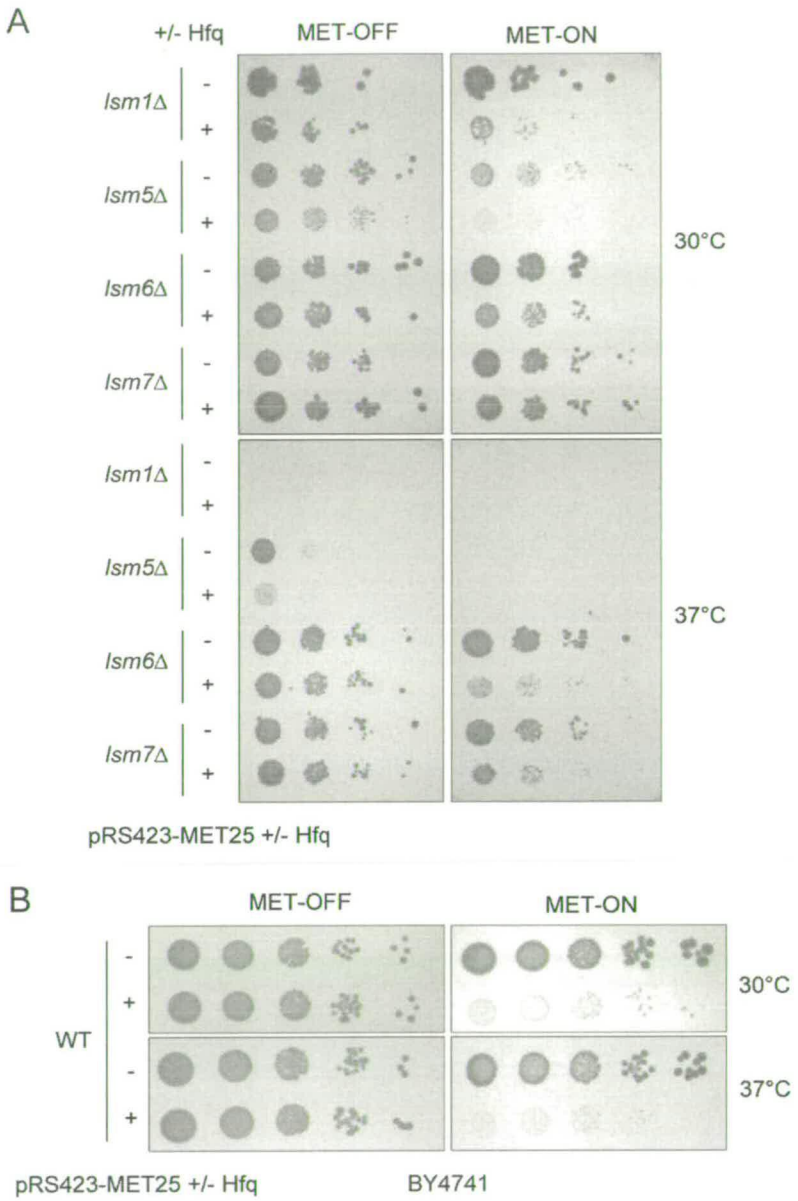


Fig. A3.2 Hfq overexpression has a detrimental effect on wild-type yeast growth, but less so on *lsm* deletion mutant strains. AEMY25 (*lsm1*Δ), MRY27 (*lsm5*Δ), MRY17 (*lsm6*Δ), MRY18 (*lsm7*Δ) and BY4741 (WT) were transformed with pRS4123-MET25 (2μ, *HIS3*) with (+) or without (-) Hfq. Cells were grown in SD -His and spotted in 10-fold serial dilutions on SD-His (MET-OFF) or SD-His-Met (MET-ON) and incubated at 30 or 37°C. **(A)** Hfq overexpression aggravates slow growth in *lsm1*Δ and *lsm6*Δ, is synthetic lethal with *lsm5*Δ, does not affect *lsm7*Δ at 30°C, but aggravates all temperature-sensitive phenotypes. A low level of Hfq expression (MET-OFF) even leads to slighlty better growth of *lsm7*Δ at 30°C; **(B)** Hfq overexpression has a strong negative effect on growth of wild-type yeast at 30 and 37°C.

References

Achsel,T., Brahms,H., Kastner,B., Bachi,A., Wilm,M., and Lührmann,R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.*, **18,** 5789-5802.

Achsel,T., Stark,H., and Lührmann,R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proc. Natl. Acad. Sci. U. S. A,* **98,** 3685-3689.

Albers, M. Identification and characterisation of Prp45p and Prp46p, two novel pre-mRNA splicing factors in *Saccharomyces cerevisiae.* 1999.
Ref Type: Thesis/Dissertation

Albrecht,M. and Lengauer,T. (2004). Novel Sm-like proteins with long C-terminal tails and associated methyltransferases. *FEBS Lett.,* **569,** 18-26.

Allmang,C., Kufel,J., Chanfreau,G., Mitchell,P., Petfalski,E., and Tollervey,D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.,* **18,** 5399-5410.

Altman,S. and Kirsebom,L. (1999). Ribonuclease P. *The RNA World,* . CSHL Press, Cold Spring Harbor, New York, pp. 351-380.

Altschul,S.F., Gish,W., Miller,W., Myers,E.W., and Lipman,D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.,* **215,** 403-410.

Anantharaman,V. and Aravind,L. (2004). Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability. *BMC. Genomics,* **5,** 45.

Andrei,M.A., Ingelfinger,D., Heintzmann,R., Achsel,T., Rivera-Pomar,R., and Lührmann,R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA,* **11,** 717-727.

Arluison,V., Hohng,S., Roy,R., Pellegrini,O., Regnier,P., and Ha,T. (2007). Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res.,* **35,** 999-1006.

Arluison,V., Mura,C., Guzman,M.R., Liquier,J., Pellegrini,O., Gingery,M., Régnier,P., and Marco,S. (2006). Three-dimensional structures of fibrillar Sm proteins: Hfq and other Sm-like proteins. *J. Mol. Biol.,* **356,** 86-96.

Askree,S.H., Yehuda,T., Smolikov,S., Gurevich,R., Hawk,J., Coker,C., Krauskopf,A., Kupiec,M., and McEachern,M.J. (2004). A genome-wide screen for Saccharomyces cerevisiae deletion mutants that affect telomere length. *Proc. Natl. Acad. Sci. U. S. A,* **101,** 8658-8663.

Audhya,A., Hyndman,F., McLeod,I.X., Maddox,A.S., Yates,J.R., III, Desai,A., and Oegema,K. (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in Caenorhabditis elegans. *J. Cell Biol.,* **171,** 267-279.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. (1994). *Current Protocols in Molecular Biology.* John Wiley and Sons, Inc..

Badis,G., Saveanu,C., Fromont-Racine,M., and Jacquier,A. (2004). Targeted mRNA degradation by deadenylation-independent decapping. *Mol. Cell,* **15,** 5-15.

Barbee,S.A., Estes,P.S., Cziko,A.M., Hillebrand,J., Luedeman,R.A., Coller,J.M., Johnson,N., Howlett,I.C., Geng,C., Ueda,R., Brand,A.H., Newbury,S.F., Wilhelm,J.E., Levine,R.B., Nakamura,A., Parker,R., and Ramaswami,M. (2006). Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. *Neuron,* **52,** 997-1009.

Bartel,D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell,* **116,** 281-297.

Beelman,C.A., Stevens,A., Caponigro,G., LaGrandeur,T.E., Hatfield,L., Fortner,D.M., and Parker,R. (1996). An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature,* **382,** 642-646.

Beggs,J.D. (2005). Lsm proteins and RNA processing. *Biochem. Soc. Trans.,* **33,** 433-438.

Bell,M., Schreiner,S., Damianov,A., Reddy,R., and Bindereif,A. (2002). p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J.,* **21,** 2724-2735.

Bentley,D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.,* **17,** 251-256.

Berget,S.M., Moore,C., and Sharp,P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U. S. A,* **74,** 3171-3175.

Bertani,G. (2004). Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.,* **186,** 595-600.

Bertrand,E. and Bordonné,R. (2004). Assembly and traffic of small nuclear RNPs. *Prog. Mol. Subcell. Biol.,* **35,** 79-97.

Bhattacharyya,S.N., Habermacher,R., Martine,U., Closs,E.I., and Filipowicz,W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell,* **125,** 1111-1124.

Blackburn,E.H. (1999). Telomerase. *The RNA World,* . CSHL Press, Cold Spring Harbor, New York, pp. 609-635.

Blom,N., Gammeltoft,S., and Brunak,S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.,* **294,** 1351-1362.

Blower,M.D., Nachury,M., Heald,R., and Weis,K. (2005). A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell,* **121,** 223-234.

Boeck,R., Lapeyre,B., Brown,C.E., and Sachs,A.B. (1998). Capped mRNA degradation intermediates accumulate in the yeast *spb8-2* mutant. *Mol. Cell Biol.,* **18,** 5062-5072.

Boelens,W.C., Palacios,I., and Mattaj,I.W. (1995). Nuclear retention of RNA as a mechanism for localization. *RNA,* **1,** 273-283.

Bonnerot,C., Boeck,R., and Lapeyre,B. (2000). The two proteins Pat1p (Mrt1p) and Spb8p interact in vivo, are required for mRNA decay, and are functionally linked to Pab1p. *Mol. Cell Biol.,* **20,** 5939-5946.

Boon, K. L. Analysis of yeast Prp8 protein and its role in U5 snRNP assembly. 2005.
Ref Type: Thesis/Dissertation

Bordonné,R. (2000). Functional characterization of nuclear localization signals in yeast Sm proteins. *Mol. Cell Biol.,* **20,** 7943-7954.

Bouveret,E., Rigaut,G., Shevchenko,A., Wilm,M., and Séraphin,B. (2000). A Sm-like protein complex that participates in mRNA degradation. *EMBO J.*, **19,** 1661-1671.

Brachmann,C.B., Davies,A., Cost,G.J., Caputo,E., Li,J., Hieter,P., and Boeke,J.D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast,* **14,** 115-132.

Brahms,H., Meheus,L., de,B., V, Fischer,U., and Lührmann,R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA,* **7,** 1531-1542.

Brengues,M., Teixeira,D., and Parker,R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science,* **310,** 486-489.

Brescia,C.C., Mikulecky,P.J., Feig,A.L., and Sledjeski,D.D. (2003). Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA,* **9,** 33-43.

Brow,D.A. (2002). Allosteric cascade of spliceosome activation. *Annu. Rev. Genet.,* **36,** 333-360.

Brow,D.A. and Guthrie,C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature,* **334,** 213-218.

Brow,D.A. and Vidaver,R.M. (1995). An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. *RNA,* **1,** 122-131.

Bruno,I. and Wilkinson,M.F. (2006). P-bodies react to stress and nonsense. *Cell,* **125,** 1036-1038.

Carrel,T.L., McWhorter,M.L., Workman,E., Zhang,H., Wolstencroft,E.C., Lorson,C., Bassell,G.J., Burghes,A.H., and Beattie,C.E. (2006). Survival motor neuron function in motor axons is independent of functions required for small nuclear ribonucleoprotein biogenesis. *J. Neurosci.,* **26,** 11014-11022.

Chan,S.P., Kao,D.I., Tsai,W.Y., and Cheng,S.C. (2003). The Prp19p-associated complex in spliceosome activation. *Science,* **302,** 279-282.

Chen,C.H., Kao,D.I., Chan,S.P., Kao,T.C., Lin,J.Y., and Cheng,S.C. (2006). Functional links between the Prp19-associated complex, U4/U6 biogenesis, and spliceosome recycling. *RNA,* **12,** 765-774.

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.

Chow,L.T., Gelinas,R.E., Broker,T.R., and Roberts,R.J. (1977a). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell,* **12,** 1-8.

Chow,L.T., Roberts,J.M., Lewis,J.B., and Broker,T.R. (1977b). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. *Cell,* **11,** 819-836.

Christiansen,J.K., Nielsen,J.S., Ebersbach,T., Valentin-Hansen,P., Sogaard-Andersen,L., and Kallipolitis,B.H. (2006). Identification of small Hfq-binding RNAs in Listeria monocytogenes. *RNA,* **12,** 1383-1396.

Chun,K.T. and Goebl,M.G. (1996). The identification of transposon-tagged mutations in essential genes that affect cell morphology in Saccharomyces cerevisiae. *Genetics,* **142,** 39-50.

Clamp,M., Cuff,J., Searle,S.M., and Barton,G.J. (2004). The Jalview Java alignment editor. *Bioinformatics.,* **20,** 426-427.

Cliften,P., Sudarsanam,P., Desikan,A., Fulton,L., Fulton,B., Majors,J., Waterston,R., Cohen,B.A., and Johnston,M. (2003). Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science,* **301,** 71-76.

Cokol,M., Nair,R., and Rost,B. (2000). Finding nuclear localization signals. *EMBO Rep.,* **1,** 411-415.

Coller,J. and Parker,R. (2005). General translational repression by activators of mRNA decapping. *Cell,* **122,** 875-886.

Coller,J.M., Tucker,M., Sheth,U., Valencia-Sanchez,M.A., and Parker,R. (2001). The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA,* **7,** 1717-1727.

Collins,B.M., Cubeddu,L., Naidoo,N., Harrop,S.J., Kornfeld,G.D., Dawes,I.W., Curmi,P.M., and Mabbutt,B.C. (2003). Homomeric ring assemblies of eukaryotic Sm proteins have affinity for both RNA and DNA. Crystal structure of an oligomeric complex of yeast SmF. *J. Biol. Chem.,* **278,** 17291-17298.

Collins,B.M., Harrop,S.J., Kornfeld,G.D., Dawes,I.W., Curmi,P.M., and Mabbutt,B.C. (2001). Crystal structure of a heptameric Sm-like protein complex from archaea: implications for the structure and evolution of snRNPs. *J. Mol. Biol.,* **309,** 915-923.

Collins,C.A. and Guthrie,C. (2000). The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.,* **7,** 850-854.

Colman-Lerner,A., Chin,T.E., and Brent,R. (2001). Yeast Cbk1 and Mob2 activate daughter-specific genetic programs to induce asymmetric cell fates. *Cell,* **107,** 739-750.

Conte,N., Charafe-Jauffret,E., Delaval,B., Adelaide,J., Ginestier,C., Geneix,J., Isnardon,D., Jacquemier,J., and Birnbaum,D. (2002). Carcinogenesis and translational controls: TACC1 is down-regulated in human cancers and associates with mRNA regulators. *Oncogene,* **21,** 5619-5630.

Conte,N., Delaval,B., Ginestier,C., Ferrand,A., Isnardon,D., Larroque,C., Prigent,C., Séraphin,B., Jacquemier,J., and Birnbaum,D. (2003). TACC1-chTOG-Aurora A protein complex in breast cancer. *Oncogene,* **22,** 8102-8116.

Cooper, M. Characterisation of a novel snRNP protein. 1995.
Ref Type: Thesis/Dissertation

Cooper,M., Johnston,L.H., and Beggs,J.D. (1995). Identification and characterization of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. *EMBO J.,* **14,** 2066-2075.

Corbett,M., Xiong,Y., Boyne,J.R., Wright,D.J., Munro,E., and Price,C. (2006). IQGAP and mitotic exit network (MEN) proteins are required for cytokinesis and re-polarization of the actin cytoskeleton in the budding yeast, Saccharomyces cerevisiae. *Eur. J. Cell Biol.,* **85,** 1201-1215.

Cordin,O., Tanner,N.K., Doere,M., Linder,P., and Banroques,J. (2004). The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J.,* **23,** 2478-2487.

Cougot,N., Babajko,S., and Séraphin,B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.,* **165,** 31-40.

D'Amours,D. and Amon,A. (2004). At the interface between signaling and executing anaphase--Cdc14 and the FEAR network. *Genes Dev.,* **18,** 2581-2595.

Davierwala,A.P., Haynes,J., Li,Z., Brost,R.L., Robinson,M.D., Yu,L., Mnaimneh,S., Ding,H., Zhu,H., Chen,Y., Cheng,X., Brown,G.W., Boone,C., Andrews,B.J., and Hughes,T.R. (2005). The synthetic genetic interaction spectrum of essential genes. *Nat. Genet.*, **37**, 1147-1152.

Decker,C.J. and Parker,R. (2006). CAR-1 and trailer hitch: driving mRNP granule function at the ER? *J. Cell Biol.*, **173**, 159-163.

Derkatch,I.L., Bradley,M.E., Hong,J.Y., and Liebman,S.W. (2001). Prions affect the appearance of other prions: the story of [PIN(+)]. *Cell*, **106**, 171-182.

Derkatch,I.L., Uptain,S.M., Outeiro,T.F., Krishnan,R., Lindquist,S.L., and Liebman,S.W. (2004). Effects of Q/N-rich, polyQ, and non-polyQ amyloids on the de novo formation of the [PSI+] prion in yeast and aggregation of Sup35 in vitro. *Proc. Natl. Acad. Sci. U. S. A*, **101**, 12934-12939.

Diéz,J., Ishikawa,M., Kaido,M., and Ahlquist,P. (2000). Identification and characterization of a host protein required for efficient template selection in viral RNA replication. *Proc. Natl. Acad. Sci. U. S. A*, **97**, 3913-3918.

Dunckley,T. and Parker,R. (1999). The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *EMBO J.*, **18**, 5411-5422.

Dunckley,T., Tucker,M., and Parker,R. (2001). Two related proteins, Edc1p and Edc2p, stimulate mRNA decapping in Saccharomyces cerevisiae. *Genetics*, **157**, 27-37.

Eide,D.J. (1998). The molecular biology of metal ion transport in Saccharomyces cerevisiae. *Annu. Rev. Nutr.*, **18**, 441-469.

Elble,R. (1992). A simple and efficient procedure for transformation of yeasts. *Biotechniques*, **13**, 18-20.

Evans,S.K. (2001). Telomeres. *Curr. Biol.*, **11**, R418.

Eystathioy,T., Jakymiw,A., Chan,E.K., Séraphin,B., Cougot,N., and Fritzler,M.J. (2003). The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA*, **9**, 1171-1173.

Eystathioy,T., Peebles,C.L., Hamel,J.C., Vaughn,J.H., and Chan,E.K. (2002). Autoantibody to hLSm4 and the heptameric LSm complex in anti-Sm sera. *Arthritis Rheum.*, **46**, 726-734.

Fabrizio,P., McPheeters,D.S., and Abelson,J. (1989). In vitro assembly of yeast U6 snRNP: a functional assay. *Genes Dev.*, **3**, 2137-2150.

Fernandez,C.F., Pannone,B.K., Chen,X., Fuchs,G., and Wolin,S.L. (2004). An Lsm2-Lsm7 complex in *Saccharomyces cerevisiae* associates with the small nucleolar RNA snR5. *Mol. Biol. Cell*, **15**, 2842-2852.

Filipowicz,W. and Pogacic,V. (2002). Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.*, **14**, 319-327.

Fischer,N. and Weis,K. (2002). The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. *EMBO J.*, **21**, 2788-2797.

Fleischer,T.C., Weaver,C.M., McAfee,K.J., Jennings,J.L., and Link,A.J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev.*, **20**, 1294-1307.

Folichon,M., Allemand,F., Régnier,P., and Hajnsdorf,E. (2005). Stimulation of poly(A) synthesis by Escherichia coli poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *FEBS J.*, **272,** 454-463.

Folichon,M., Arluison,V., Pellegrini,O., Huntzinger,E., Régnier,P., and Hajnsdorf,E. (2003). The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res.*, **31,** 7302-7310.

Fortner,D.M., Troy,R.G., and Brow,D.A. (1994). A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev.*, **8,** 221-233.

Foster,R., Mikesell,G.E., and Breeden,L. (1993). Multiple SWI6-dependent cis-acting elements control SWI4 transcription through the cell cycle. *Mol. Cell Biol.*, **13,** 3792-3801.

Franze de Fernandez,M.T., Eoyang,L., and August,J.T. (1968). Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature,* **219,** 588-590.

Fraser,M.M., Watson,P.M., Fraig,M.M., Kelley,J.R., Nelson,P.S., Boylan,A.M., Cole,D.J., and Watson,D.K. (2005). CaSm-mediated cellular transformation is associated with altered gene expression and messenger RNA stability. *Cancer Res.*, **65,** 6228-6236.

Friesen,W.J. and Dreyfuss,G. (2000). Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN). *J. Biol. Chem.*, **275,** 26370-26375.

Fromont-Racine,M., Mayes,A.E., Brunet-Simon,A., Rain,J.C., Colley,A., Dix,I., Decourty,L., Joly,N., Ricard,F., Beggs,J.D., and Legrain,P. (2000). Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins. *Yeast,* **17,** 95-110.

Fu,D. and Collins,K. (2006). Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev.*, **20,** 531-536.

Gao,M., Wilusz,C.J., Peltz,S.W., and Wilusz,J. (2001). A novel mRNA-decapping activity in HeLa cytoplasmic extracts is regulated by AU-rich elements. *EMBO J.*, **20,** 1134-1143.

Gavin,A.C., Aloy,P., Grandi,P., Krause,R., Boesche,M., Marzioch,M., Rau,C., Jensen,L.J., Bastuck,S., Dumpelfeld,B., Edelmann,A., Heurtier,M.A., Hoffman,V., Hoefert,C., Klein,K., Hudak,M., Michon,A.M., Schelder,M., Schirle,M., Remor,M., Rudi,T., Hooper,S., Bauer,A., Bouwmeester,T., Casari,G., Drewes,G., Neubauer,G., Rick,J.M., Kuster,B., Bork,P., Russell,R.B., and Superti-Furga,G. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature,* **440,** 631-636.

Geissmann,T.A. and Touati,D. (2004). Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J.*, **23,** 396-405.

Gerbi,S.A. and Lange,T.S. (2002). All small nuclear RNAs (snRNAs) of the [U4/U6.U5] Tri-snRNP localize to nucleoli; Identification of the nucleolar localization element of U6 snRNA. *Mol. Biol. Cell,* **13,** 3123-3137.

Ghetti,A., Company,M., and Abelson,J. (1995). Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA,* **1,** 132-145.

Giaever,G., Chu,A.M., Ni,L., Connelly,C., Riles,L., Veronneau,S., Dow,S., Lucau-Danila,A., Anderson,K., Andre,B., Arkin,A.P., Astromoff,A., El Bakkoury,M., Bangham,R., Benito,R., Brachat,S., Campanaro,S., Curtiss,M., Davis,K., Deutschbauer,A., Entian,K.D., Flaherty,P., Foury,F., Garfinkel,D.J., Gerstein,M., Gotte,D., Guldener,U., Hegemann,J.H., Hempel,S., Herman,Z., Jaramillo,D.F., Kelly,D.E., Kelly,S.L., Kotter,P., LaBonte,D., Lamb,D.C., Lan,N., Liang,H., Liao,H.,

Liu,L., Luo,C., Lussier,M., Mao,R., Menard,P., Ooi,S.L., Revuelta,J.L., Roberts,C.J., Rose,M., Ross-Macdonald,P., Scherens,B., Schimmack,G., Shafer,B., Shoemaker,D.D., Sookhai-Mahadeo,S., Storms,R.K., Strathern,J.N., Valle,G., Voet,M., Volckaert,G., Wang,C.Y., Ward,T.R., Wilhelmy,J., Winzeler,E.A., Yang,Y., Yen,G., Youngman,E., Yu,K., Bussey,H., Boeke,J.D., Snyder,M., Philippsen,P., Davis,R.W., and Johnston,M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature,* **418,** 387-391.

Gietz,D., St Jean,A., Woods,R.A., and Schiestl,R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.,* **20,** 1425.

Gietz,R.D. and Woods,R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.,* **350,** 87-96.

Gill,S.C. and von Hippel,P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.,* **182,** 319-326.

Gill,T., Aulds,J., and Schmitt,M.E. (2006). A specialized processing body that is temporally and asymmetrically regulated during the cell cycle in *Saccharomyces cerevisiae. J. Cell Biol.,* **173,** 35-45.

Girard,C., Mouaikel,J., Neel,H., Bertrand,E., and Bordonné,R. (2004). Nuclear localization properties of a conserved protuberance in the Sm core complex. *Exp. Cell Res.,* **299,** 199-208.

Glover,C.V., III (1998). On the physiological role of casein kinase II in Saccharomyces cerevisiae. *Prog. Nucleic Acid Res. Mol. Biol.,* **59,** 95-133.

Goldstein,A.L. and McCusker,J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae. Yeast,* **15,** 1541-1553.

Görnemann,J., Kotovic,K.M., Hujer,K., and Neugebauer,K.M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol. Cell,* **19,** 53-63.

Gottesman,S., Storz,G., Rosenow,C., Majdalani,N., Repoila,F., and Wassarman,K.M. (2001). Small RNA regulators of translation: mechanisms of action and approaches for identifying new small RNAs. *Cold Spring Harb. Symp. Quant. Biol.,* **66,** 353-362.

Gottschalk,A., Neubauer,G., Banroques,J., Mann,M., Lührmann,R., and Fabrizio,P. (1999). Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *EMBO J.,* **18,** 4535-4548.

Granneman,S. and Baserga,S.J. (2004). Ribosome biogenesis: of knobs and RNA processing. *Exp. Cell Res.,* **296,** 43-50.

Grossi,S., Puglisi,A., Dmitriev,P.V., Lopes,M., and Shore,D. (2004). Pol12, the B subunit of DNA polymerase alpha, functions in both telomere capping and length regulation. *Genes Dev.,* **18,** 992-1006.

Gumbs,A.A., Bassi,C., Moore,P.S., Falconi,M., Frigerio,I., Baron,A., Piemonti,L., Modlin,I., and Scarpa,A. (2002). Overexpression of the Sm-like proto-oncogene in primary and metastatic pancreatic endocrine tumors. *JOP.,* **3,** 109-115.

Hajnsdorf,E. and Régnier,P. (2000). Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc. Natl. Acad. Sci. U. S. A,* **97,** 1501-1505.

Halic,M. and Beckmann,R. (2005). The signal recognition particle and its interactions during protein targeting. *Curr. Opin. Struct. Biol.,* **15,** 116-125.

Hamm,J. and Mattaj,I.W. (1989). An abundant U6 snRNP found in germ cells and embryos of Xenopus laevis. *EMBO J.*, **8,** 4179-4187.

Hampsey,M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.*, **62,** 465-503.

Hatfield,L., Beelman,C.A., Stevens,A., and Parker,R. (1996). Mutations in trans-acting factors affecting mRNA decapping in Saccharomyces cerevisiae. *Mol. Cell Biol.*, **16,** 5830-5838.

He,W. and Parker,R. (2000). Functions of Lsm proteins in mRNA degradation and splicing. *Curr. Opin. Cell Biol.*, **12,** 346-350.

He,W. and Parker,R. (2001). The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3' termini from partial degradation. *Genetics,* **158,** 1445-1455.

Hentges,P., Van Driessche,B., Tafforeau,L., Vandenhaute,J., and Carr,A.M. (2005). Three novel antibiotic marker cassettes for gene disruption and marker switching in *Schizosaccharomyces pombe*. *Yeast,* **22,** 1013-1019.

Hermann,H., Fabrizio,P., Raker,V.A., Foulaki,K., Hornig,H., Brahms,H., and Lührmann,R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.*, **14,** 2076-2088.

Herschlag,D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.*, **270,** 20871-20874.

Horak,C.E., Luscombe,N.M., Qian,J., Bertone,P., Piccirrillo,S., Gerstein,M., and Snyder,M. (2002). Complex transcriptional circuitry at the G1/S transition in Saccharomyces cerevisiae. *Genes Dev.*, **16,** 3017-3033.

Houseley,J., LaCava,J., and Tollervey,D. (2006). RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.*, **7,** 529-539.

Hu,J., Xu,D., Schappert,K., Xu,Y., and Friesen,J.D. (1995). Mutational analysis of *Saccharomyces cerevisiae* U4 small nuclear RNA identifies functionally important domains. *Mol. Cell Biol.*, **15,** 1274-1285.

Hu,J., Xu,Y., Schappert,K., Harrington,T., Wang,A., Braga,R., Mogridge,J., and Friesen,J.D. (1994). Mutational analysis of the PRP4 protein of Saccharomyces cerevisiae suggests domain structure and snRNP interactions. *Nucleic Acids Res.*, **22,** 1724-1734.

Hwa,L.H., Yeong,F.M., and Surana,U. (2003). Inactivation of mitotic kinase triggers translocation of MEN components to mother-daughter neck in yeast. *Mol. Biol. Cell,* **14,** 4734-4743.

Ingelfinger,D., Arndt-Jovin,D.J., Lührmann,R., and Achsel,T. (2002). The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA,* **8,** 1489-1501.

Inoue,H., Nojima,H., and Okayama,H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene,* **96,** 23-28.

Jacquier,A. (1997). RNA-catalysed RNA processing. *Eukaryotic mRNA processing*, . Oxford University Press, Oxford, pp. 1-36.

Jandrositz,A. and Guthrie,C. (1995). Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J.*, **14,** 820-832.

Jarmolowski,A. and Mattaj,I.W. (1993). The determinants for Sm protein binding to Xenopus U1 and U5 snRNAs are complex and non-identical. *EMBO J.*, **12,** 223-232.

Jiménez,J., Cid,V.J., Cenamor,R., Yuste,M., Molero,G., Nombela,C., and Sánchez,M. (1998). Morphogenesis beyond cytokinetic arrest in Saccharomyces cerevisiae. *J. Cell Biol.*, **143,** 1617-1634.

Johnston,M. and Davis,R.W. (1984). Sequences that regulate the divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **4,** 1440-1448.

Jones,M.H. and Guthrie,C. (1990). Unexpected flexibility in an evolutionarily conserved protein-RNA interaction: genetic analysis of the Sm binding site. *EMBO J.*, **9,** 2555-2561.

Joyce,G.F. and Orgel,L.E. (1999). Prospect for the understanding of the origin of the RNA world. *The RNA World,* . CSHL Press, Cold Spring Harbor, New York, pp. 49-77.

Jurica,M.S. and Moore,M.J. (2002). Capturing splicing complexes to study structure and mechanism. *Methods,* **28,** 336-345.

Kajitani,M., Kato,A., Wada,A., Inokuchi,Y., and Ishihama,A. (1994). Regulation of the *Escherichia coli* hfq gene encoding the host factor for phage Q beta. *J. Bacteriol.*, **176,** 531-534.

Kambach,C., Walke,S., Young,R., Avis,J.M., de la,F.E., Raker,V.A., Lührmann,R., Li,J., and Nagai,K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell,* **96,** 375-387.

Karaduman,R., Fabrizio,P., Hartmuth,K., Urlaub,H., and Lührmann,R. (2006). RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J. Mol. Biol.*, **356,** 1248-1262.

Kastner,B., Bach,M., and Lührmann,R. (1991). Electron microscopy of U4/U6 snRNP reveals a Y-shaped U4 and U6 RNA containing domain protruding from the U4 core RNP. *J. Cell Biol.*, **112,** 1065-1072.

Kazantsev,A.V. and Pace,N.R. (2006). Bacterial RNase P: a new view of an ancient enzyme. *Nat. Rev. Microbiol.*, **4,** 729-740.

Kelley,J.R., Brown,J.M., Frasier,M.M., Baron,P.L., Schweinfest,C.W., Vournakis,J.N., Watson,D.K., and Cole,D.J. (2000a). The cancer-associated Sm-like oncogene: a novel target for the gene therapy of pancreatic cancer. *Surgery,* **128,** 353-360.

Kelley,J.R., Fraser,M.M., Hubbard,J.M., Watson,D.K., and Cole,D.J. (2003). CaSm antisense gene therapy: a novel approach for the treatment of pancreatic cancer. *Anticancer Res.*, **23,** 2007-2013.

Kelley,J.R., Fraser,M.M., Schweinfest,C.W., Vournakis,J.N., Watson,D.K., and Cole,D.J. (2001). CaSm/gemcitabine chemo-gene therapy leads to prolonged survival in a murine model of pancreatic cancer. *Surgery,* **130,** 280-288.

Kelley,L.A., MacCallum,R.M., and Sternberg,M.J. (2000b). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.*, **299,** 499-520.

Kellis,M., Patterson,N., Endrizzi,M., Birren,B., and Lander,E.S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature,* **423,** 241-254.

Khusial,P., Plaag,R., and Zieve,G.W. (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem. Sci.*, **30,** 522-528.

Kilic,T., Sanglier,S., Van Dorsselaer,A., and Suck,D. (2006). Oligomerization behavior of the archaeal Sm2-type protein from Archaeoglobus fulgidus. *Protein Sci.*, **15,** 2310-2317.

Kiss,T. (2002). Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell,* **109,** 145-148.

Klingauf,M., Stanek,D., and Neugebauer,K.M. (2006). Enhancement of u4/u6 small nuclear ribonucleoprotein particle association in cajal bodies predicted by mathematical modeling. *Mol. Biol. Cell,* **17,** 4972-4981.

Kozak,M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.,* **266,** 19867-19870.

Krogan,N.J., Cagney,G., Yu,H., Zhong,G., Guo,X., Ignatchenko,A., Li,J., Pu,S., Datta,N., Tikuisis,A.P., Punna,T., Peregrin-Alvarez,J.M., Shales,M., Zhang,X., Davey,M., Robinson,M.D., Paccanaro,A., Bray,J.E., Sheung,A., Beattie,B., Richards,D.P., Canadien,V., Lalev,A., Mena,F., Wong,P., Starostine,A., Canete,M.M., Vlasblom,J., Wu,S., Orsi,C., Collins,S.R., Chandran,S., Haw,R., Rilstone,J.J., Gandi,K., Thompson,N.J., Musso,G., St Onge,P., Ghanny,S., Lam,M.H., Butland,G., Altaf-Ul,A.M., Kanaya,S., Shilatifard,A., O'Shea,E., Weissman,J.S., Ingles,C.J., Hughes,T.R., Parkinson,J., Gerstein,M., Wodak,S.J., Emili,A., and Greenblatt,J.F. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature,* **440,** 637-643.

Kshirsagar,M. and Parker,R. (2004). Identification of Edc3p as an enhancer of mRNA decapping in Saccharomyces cerevisiae. *Genetics,* **166,** 729-739.

Kufel,J., Allmang,C., Petfalski,E., Beggs,J., and Tollervey,D. (2003a). Lsm Proteins are required for normal processing and stability of ribosomal RNAs. *J. Biol. Chem.,* **278,** 2147-2156.

Kufel,J., Allmang,C., Verdone,L., Beggs,J., and Tollervey,D. (2003b). A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic Acids Res.,* **31,** 6788-6797.

Kufel,J., Allmang,C., Verdone,L., Beggs,J.D., and Tollervey,D. (2002). Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol. Cell Biol.,* **22,** 5248-5256.

Kufel,J., Bousquet-Antonelli,C., Beggs,J.D., and Tollervey,D. (2004). Nuclear pre-mRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol. Cell Biol.,* **24,** 9646-9657.

Kwan,S.S. and Brow,D.A. (2005). The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA,* **11,** 808-820.

Lacadie,S.A. and Rosbash,M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Mol. Cell,* **19,** 65-75.

LaCava,J., Houseley,J., Saveanu,C., Petfalski,E., Thompson,E., Jacquier,A., and Tollervey,D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell,* **121,** 713-724.

LaGrandeur,T.E. and Parker,R. (1998). Isolation and characterization of Dcp1p, the yeast mRNA decapping enzyme. *EMBO J.,* **17,** 1487-1496.

Lange,T.S. and Gerbi,S.A. (2000). Transient nucleolar localization Of U6 small nuclear RNA in Xenopus Laevis oocytes. *Mol. Biol. Cell,* **11,** 2419-2428.

Le Derout,J., Folichon,M., Briani,F., Deho,G., Régnier,P., and Hajnsdorf,E. (2003). Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of *Escherichia coli* rpsO mRNAs. *Nucleic Acids Res.,* **31,** 4017-4023.

Le Novere,N. (2001). MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics.,* **17,** 1226-1227.

Lehmeier,T., Foulaki,K., and Lührmann,R. (1990). Evidence for three distinct D proteins, which react differentially with anti-Sm autoantibodies, in the cores of the major snRNPs U1, U2, U4/U6 and U5. *Nucleic Acids Res.*, **18,** 6475-6484.

Lehner,B. and Sanderson,C.M. (2004). A protein interaction framework for human mRNA degradation. *Genome Res.*, **14,** 1315-1323.

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. U. S. A,* **76,** 5495-5499.

Link,C.D., Fonte,V., Hiester,B., Yerg,J., Ferguson,J., Csontos,S., Silverman,M.A., and Stein,G.H. (2006). Conversion of green fluorescent protein into a toxic, aggregation-prone protein by C-terminal addition of a short peptide. *J. Biol. Chem.*, **281,** 1808-1816.

Liu,Q., Fischer,U., Wang,F., and Dreyfuss,G. (1997). The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell,* **90,** 1013-1021.

Liu,Q., Liang,X.H., Uliel,S., Belahcen,M., Unger,R., and Michaeli,S. (2004). Identification and functional characterization of lsm proteins in *Trypanosoma brucei*. *J. Biol. Chem.*, **279,** 18210-18219.

Longtine,M.S., McKenzie,A., III, Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P., and Pringle,J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast,* **14,** 953-961.

Luca,F.C., Mody,M., Kurischko,C., Roof,D.M., Giddings,T.H., and Winey,M. (2001). Saccharomyces cerevisiae Mob1p is required for cytokinesis and mitotic exit. *Mol. Cell Biol.*, **21,** 6972-6983.

Luukkonen,B.G. and Séraphin,B. (1998). Construction of an in vivo-regulated U6 snRNA transcription unit as a tool to study U6 function. *RNA,* **4,** 231-238.

Lygerou,Z., Christophides,G., and Séraphin,B. (1999). A novel genetic screen for snRNP assembly factors in yeast identifies a conserved protein, Sad1p, also required for pre-mRNA splicing. *Mol. Cell Biol.*, **19,** 2008-2020.

Lyng,H., Brovig,R.S., Svendsrud,D.H., Holm,R., Kaalhus,O., Knutstad,K., Oksefjell,H., Sundfor,K., Kristensen,G.B., and Stokke,T. (2006). Gene expressions and copy numbers associated with metastatic phenotypes of uterine cervical cancer. *BMC. Genomics,* **7,** 268.

Ma,Y., Dostie,J., Dreyfuss,G., and Van Duyne,G.D. (2005). The Gemin6-Gemin7 heterodimer from the survival of motor neurons complex has an Sm protein-like structure. *Structure.*, **13,** 883-892.

MacKay,V.L., Mai,B., Waters,L., and Breeden,L.L. (2001). Early cell cycle box-mediated transcription of CLN3 and SWI4 contributes to the proper timing of the G(1)-to-S transition in budding yeast. *Mol. Cell Biol.*, **21,** 4140-4148.

Madhani,H.D., Bordonné,R., and Guthrie,C. (1990). Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.*, **4,** 2264-2277.

Madocsai,C., Lim,S.R., Geib,T., Lam,B.J., and Hertel,K.J. (2005). Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs. *Mol. Ther.*, **12,** 1013-1022.

Majdalani,N., Vanderpool,C.K., and Gottesman,S. (2005). Bacterial small RNA regulators. *Crit Rev. Biochem. Mol. Biol.*, **40,** 93-113.

Mangus,D.A., Amrani,N., and Jacobson,A. (1998). Pbp1p, a factor interacting with Saccharomyces cerevisiae poly(A)-binding protein, regulates polyadenylation. *Mol. Cell Biol.*, **18,** 7383-7396.

Mangus,D.A., Smith,M.M., McSweeney,J.M., and Jacobson,A. (2004). Identification of factors regulating poly(A) tail synthesis and maturation. *Mol. Cell Biol.*, **24,** 4196-4206.

Maniatis,T. and Reed,R. (2002). An extensive network of coupling among gene expression machines. *Nature,* **416,** 499-506.

Mas,A., Alves-Rodrigues,I., Noueiry,A., Ahlquist,P., and Diéz,J. (2006). Host deadenylation-dependent mRNA decapping factors are required for a key step in brome mosaic virus RNA replication. *J. Virol.*, **80,** 246-251.

Matera,A.G. and Shpargel,K.B. (2006). Pumping RNA: nuclear bodybuilding along the RNP pipeline. *Curr. Opin. Cell Biol.*, **18,** 317-324.

Mathews,D.H., Sabina,J., Zuker,M., and Turner,D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288,** 911-940.

Mayes, A. E. Identification and characterisation of proteins interacting with Uss1p: a *Saccharomyces cerevisiae* Sm-like protein. 1998.
Ref Type: Thesis/Dissertation

Mayes,A.E., Verdone,L., Legrain,P., and Beggs,J.D. (1999). Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.*, **18,** 4321-4331.

Mazzoni,C. and Falcone,C. (2001). Isolation and study of KlLSM4, a *Kluyveromyces lactis* gene homologous to the essential gene LSM4 of *Saccharomyces cerevisiae. Yeast,* **18,** 1249-1256.

Mazzoni,C., Herker,E., Palermo,V., Jungwirth,H., Eisenberg,T., Madeo,F., and Falcone,C. (2005). Yeast caspase 1 links messenger RNA stability to apoptosis in yeast. *EMBO Rep.*, **6,** 1076-1081.

Mazzoni,C., Mancini,P., Madeo,F., Palermo,V., and Falcone,C. (2003a). A *Kluyveromyces lactis* mutant in the essential gene KlLSM4 shows phenotypic markers of apoptosis. *FEMS Yeast Res.*, **4,** 29-35.

Mazzoni,C., Mancini,P., Verdone,L., Madeo,F., Serafini,A., Herker,E., and Falcone,C. (2003b). A truncated form of KlLsm4p and the absence of factors involved in mRNA decapping trigger apoptosis in yeast. *Mol. Biol. Cell,* **14,** 721-729.

McConnell,T.S., Lokken,R.P., and Steitz,J.A. (2003). Assembly of the U1 snRNP involves interactions with the backbone of the terminal stem of U1 snRNA. *RNA,* **9,** 193-201.

McPheeters,D.S., Fabrizio,P., and Abelson,J. (1989). In vitro reconstitution of functional yeast U2 snRNPs. *Genes Dev.*, **3,** 2124-2136.

Medenbach,J., Schreiner,S., Liu,S., Lührmann,R., and Bindereif,A. (2004). Human U4/U6 snRNP recycling factor p110: mutational analysis reveals the function of the tetratricopeptide repeat domain in recycling. *Mol. Cell Biol.*, **24,** 7392-7401.

Merlin,C., McAteer,S., and Masters,M. (2002). Tools for characterization of *Escherichia coli* genes of unknown function. *J. Bacteriol.*, **184,** 4573-4581.

Mikulecky,P.J., Kaw,M.K., Brescia,C.C., Takach,J.C., Sledjeski,D.D., and Feig,A.L. (2004). *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat. Struct. Mol. Biol.*, **11,** 1206-1214.

Milligan,J.F. and Uhlenbeck,O.C. (1989). Synthesis of Small RNAs Using T7 RNA Polymerase. In Dahlberg,J.E. and Abelson,J.N. (Eds.), *Methods in Enzymology, Vol. 180 RNA Processing Part A, General Methods*, ., pp. 51-62.

Miroux,B. and Walker,J.E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.*, **260,** 289-298.

Mitchell,P. and Tollervey,D. (2001). mRNA turnover. *Curr. Opin. Cell Biol.*, **13,** 320-325.

Miura,F., Kawaguchi,N., Sese,J., Toyoda,A., Hattori,M., Morishita,S., and Ito,T. (2006). A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. *Proc. Natl. Acad. Sci. U. S. A.*

Mohanty,B.K., Maples,V.F., and Kushner,S.R. (2004). The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol. Microbiol.*, **54,** 905-920.

Moll,I., Leitsch,D., Steinhauser,T., and Bläsi,U. (2003). RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep.*, **4,** 284-289.

Møller,T., Franch,T., Hojrup,P., Keene,D.R., Bachinger,H.P., Brennan,R.G., and Valentin-Hansen,P. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell,* **9,** 23-30.

Mouaikel,J., Verheggen,C., Bertrand,E., Tazi,J., and Bordonné,R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell,* **9,** 891-901.

Mumberg,D., Muller,R., and Funk,M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.*, **22,** 5767-5768.

Mura,C., Cascio,D., Sawaya,M.R., and Eisenberg,D.S. (2001). The crystal structure of a heptameric archaeal Sm protein: Implications for the eukaryotic snRNP core. *Proc. Natl. Acad. Sci. U. S. A,* **98,** 5532-5537.

Mura,C., Kozhukhovsky,A., Gingery,M., Phillips,M., and Eisenberg,D. (2003a). The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs). *Protein Sci.*, **12,** 832-847.

Mura,C., Phillips,M., Kozhukhovsky,A., and Eisenberg,D. (2003b). Structure and assembly of an augmented Sm-like archaeal protein 14-mer. *Proc. Natl. Acad. Sci. U. S. A,* **100,** 4539-4544.

Nakanishi,K. and Nureki,O. (2005). Recent progress of structural biology of tRNA processing and modification. *Mol. Cells,* **19,** 157-166.

Niedenthal,R.K., Riles,L., Johnston,M., and Hegemann,J.H. (1996). Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast,* **12,** 773-786.

Nilsen,T.W. (1994). RNA-RNA interactions in the spliceosome: unraveling the ties that bind. *Cell,* **78,** 1-4.

Nilsen,T.W. (2002). The spliceosome: no assembly required? *Mol. Cell,* **9,** 8-9.

Noble,S.M. and Guthrie,C. (1996a). Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. *Genetics,* **143,** 67-80.

Noble,S.M. and Guthrie,C. (1996b). Transcriptional pulse-chase analysis reveals a role for a novel snRNP-associated protein in the manufacture of spliceosomal snRNPs. *EMBO J.*, **15,** 4368-4379.

Noueiry,A.O., Diéz,J., Falk,S.P., Chen,J., and Ahlquist,P. (2003). Yeast Lsm1p-7p/Pat1p deadenylation-dependent mRNA-decapping factors are required for brome mosaic virus genomic RNA translation. *Mol. Cell Biol.*, **23,** 4094-4106.

Olson,B.L. and Siliciano,P.G. (2003). A diverse set of nuclear RNAs transfer between nuclei of yeast heterokaryons. *Yeast,* **20,** 893-903.

Page, D. Analysis of the function of the Lsm8 protein complex in *Saccharomyces cerevisiae.* 2002. Ref Type: Thesis/Dissertation

Pan,X., Ye,P., Yuan,D.S., Wang,X., Bader,J.S., and Boeke,J.D. (2006). A DNA integrity network in the yeast Saccharomyces cerevisiae. *Cell,* **124,** 1069-1081.

Pannone,B.K., Kim,S.D., Noe,D.A., and Wolin,S.L. (2001). Multiple functional interactions between components of the Lsm2-Lsm8 complex, U6 snRNA, and the yeast La protein. *Genetics,* **158,** 187-196.

Pannone,B.K., Xue,D., and Wolin,S.L. (1998). A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.,* **17,** 7442-7453.

Parent,S.A., Fenimore,C.M., and Bostian,K.A. (1985). Vector systems for the expression, analysis and cloning of DNA sequences in *S. cerevisiae. Yeast,* **1,** 83-138.

Parker,R. and Sheth,U. (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell,* **25,** 635-646.

Parkes,V. and Johnston,L.H. (1992). SPO12 and SIT4 suppress mutations in DBF2, which encodes a cell cycle protein kinase that is periodically expressed. *Nucleic Acids Res.,* **20,** 5617-5623.

Paushkin,S., Gubitz,A.K., Massenet,S., and Dreyfuss,G. (2002). The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.,* **14,** 305-312.

Pfiz,S., Zimmermann,J., and Hilt,W. (2002). The yeast kinetochore protein Slk19 is required to prevent aberrant chromosome segregation in meiosis and mitosis. *Genes Cells,* **7,** 1033-1042.

Polevoda,B. and Sherman,F. (2001). NatC Nalpha-terminal acetyltransferase of yeast contains three subunits, Mak3p, Mak10p, and Mak31p. *J. Biol. Chem.,* **276,** 20154-20159.

Pot,I., Knockleby,J., Aneliunas,V., Nguyen,T., Ah-Kye,S., Liszt,G., Snyder,M., Hieter,P., and Vogel,J. (2005). Spindle checkpoint maintenance requires Ame1 and Okp1. *Cell Cycle,* **4,** 1448-1456.

Rader,S.D. and Guthrie,C. (2002). A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA,* **8,** 1378-1392.

Raghunathan,P.L. and Guthrie,C. (1998a). A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science,* **279,** 857-860.

Raghunathan,P.L. and Guthrie,C. (1998b). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.,* **8,** 847-855.

Raker,V.A., Hartmuth,K., Kastner,B., and Lührmann,R. (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol. Cell Biol.,* **19,** 6554-6565.

Raker,V.A., Plessel,G., and Lührmann,R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J.*, **15,** 2256-2269.

Reiter,N.J., Lee,D., Wang,Y.X., Tonelli,M., Bahrami,A., Cornilescu,C.C., and Butcher,S.E. (2006). Resonance assignments for the two N-terminal RNA recognition motifs (RRM) of the S. cerevisiae Pre-mRNA Processing Protein Prp24. *J. Biomol. NMR,* **36 Suppl 5,** 58.

Rosen,J., He,M., Umbricht,C., Alexander,H.R., Dackiw,A.P., Zeiger,M.A., and Libutti,S.K. (2005). A six-gene model for differentiating benign from malignant thyroid tumors on the basis of gene expression. *Surgery,* **138,** 1050-1056.

Ryan,D.E. and Abelson,J. (2002). The conserved central domain of yeast U6 snRNA: importance of U2-U6 helix Ia in spliceosome assembly. *RNA,* **8,** 997-1010.

Ryan,D.E., Stevens,S.W., and Abelson,J. (2002). The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA,* **8,** 1011-1033.

Salgado-Garrido,J., Bragado-Nilsson,E., Kandels-Lewis,S., and Séraphin,B. (1999). Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.*, **18,** 3451-3462.

Saunders,W.S. (2002). The FEAR factor. *Mol. Cell,* **9,** 207-209.

Sauter,C., Basquin,J., and Suck,D. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli. Nucleic Acids Res.*, **31,** 4091-4098.

Schaffert,N., Hossbach,M., Heintzmann,R., Achsel,T., and Lührmann,R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J.*, **23,** 3000-3009.

Schmitt,M.E., Brown,T.A., and Trumpower,B.L. (1990). A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae. Nucleic Acids Res.*, **18,** 3091-3092.

Schumacher,M.A., Pearson,R.F., Møller,T., Valentin-Hansen,P., and Brennan,R.G. (2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.*, **21,** 3546-3556.

Schümperli,D. and Pillai,R.S. (2004). The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol. Life Sci.*, **61,** 2560-2570.

Schuppli,D., Miranda,G., Tsui,H.C., Winkler,M.E., Sogo,J.M., and Weber,H. (1997). Altered 3'-terminal RNA structure in phage Qbeta adapted to host factor-less *Escherichia coli. Proc. Natl. Acad. Sci. U. S. A,* **94,** 10239-10242.

Schwartz,D.C. and Parker,R. (2000). mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Mol. Cell Biol.*, **20,** 7933-7942.

Schweinfest,C.W., Graber,M.W., Chapman,J.M., Papas,T.S., Baron,P.L., and Watson,D.K. (1997). CaSm: an Sm-like protein that contributes to the transformed state in cancer cells. *Cancer Res.*, **57,** 2961-2965.

Séraphin,B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J.*, **14,** 2089-2098.

Seto,A.G., Zaug,A.J., Sobel,S.G., Wolin,S.L., and Cech,T.R. (1999). *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature,* **401,** 177-180.

Shannon,K.W. and Guthrie,C. (1991). Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. *Genes Dev.*, **5,** 773-785.

Sheth,U. and Parker,R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science,* **300,** 805-808.

Sikorski,R.S. and Hieter,P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics,* **122,** 19-27.

Silverstein,R.A. and Ekwall,K. (2005). Sin3: a flexible regulator of global gene expression and genome stability. *Curr. Genet.*, **47,** 1-17.

Singh,R. and Reddy,R. (1989). Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. U. S. A,* **86,** 8280-8283.

Spiller, M. P. Localisation of the U6 snRNA and the Lsm proteins. 2005.
Ref Type: Thesis/Dissertation

Spiller, M. P., Boon, K. L., Reijns, M. A. M., and Beggs, J. D. The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. Nucleic Acids Res. 2007.
Ref Type: In Press

Staley,J.P. and Guthrie,C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell,* **92,** 315-326.

Stanek,D. and Neugebauer,K.M. (2004). Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer. *J. Cell Biol.*, **166,** 1015-1025.

Stanek,D., Rader,S.D., Klingauf,M., and Neugebauer,K.M. (2003). Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J. Cell Biol.*, **160,** 505-516.

Stark,H., Dube,P., Lührmann,R., and Kastner,B. (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature,* **409,** 539-542.

Stearman,R., Yuan,D.S., Yamaguchi-Iwai,Y., Klausner,R.D., and Dancis,A. (1996). A permease-oxidase complex involved in high-affinity iron uptake in yeast. *Science,* **271,** 1552-1557.

Stegmeier,F., Visintin,R., and Amon,A. (2002). Separase, polo kinase, the kinetochore protein Slk19, and Spo12 function in a network that controls Cdc14 localization during early anaphase. *Cell,* **108,** 207-220.

Steinmetz,L.M., Scharfe,C., Deutschbauer,A.M., Mokranjac,D., Herman,Z.S., Jones,T., Chu,A.M., Giaever,G., Prokisch,H., Oefner,P.J., and Davis,R.W. (2002). Systematic screen for human disease genes in yeast. *Nat. Genet.*, **31,** 400-404.

Stevens,S.W. and Abelson,J. (1999). Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc. Natl. Acad. Sci. U. S. A,* **96,** 7226-7231.

Stevens,S.W., Ryan,D.E., Ge,H.Y., Moore,R.E., Young,M.K., Lee,T.D., and Abelson,J. (2002). Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell,* **9,** 31-44.

Stoepel,J., Ottey,M.A., Kurischko,C., Hieter,P., and Luca,F.C. (2005). The mitotic exit network Mob1p-Dbf2p kinase complex localizes to the nucleus and regulates passenger protein localization. *Mol. Biol. Cell,* **16,** 5465-5479.

Storz,G., Opdyke,J.A., and Zhang,A. (2004). Controlling mRNA stability and translation with small, noncoding RNAs. *Curr. Opin. Microbiol.*, **7**, 140-144.

Streicher,K.L., Yang,Z.Q., Draghici,S., and Ethier,S.P. (2006). Transforming function of the LSM1 oncogene in human breast cancers with the 8p11-12 amplicon. *Oncogene.*

Sukhodolets,M.V. and Garges,S. (2003). Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry,* **42,** 8022-8034.

Takahashi,S., Suzuki,S., Inaguma,S., Cho,Y.M., Ikeda,Y., Hayashi,N., Inoue,T., Sugimura,Y., Nishiyama,N., Fujita,T., Ushijima,T., and Shirai,T. (2002). Down-regulation of Lsm1 is involved in human prostate cancer progression. *Br. J. Cancer,* **86,** 940-946.

Tan,E.M. and Kunkel,H.G. (1966). Characteristics of a soluble nuclear antigen precipitating with sera of patients with systemic lupus erythematosus. *J. Immunol.*, **96,** 464-471.

Tan,E.M. and Kunkel,H.G. (2006). Characteristics of a soluble nuclear antigen precipitating with sera of patients with systemic lupus erythematosus. J. Immunol. 1966. 96: 464-471. *J. Immunol.*, **176,** 1297-1304.

Tanaka,K.J., Ogawa,K., Takagi,M., Imamoto,N., Matsumoto,K., and Tsujimoto,M. (2006). RAP55, a cytoplasmic mRNP component, represses translation in Xenopus oocytes. *J. Biol. Chem.*, **281,** 40096-40106.

Tardiff,D.F. and Rosbash,M. (2006). Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. *RNA,* **12,** 968-979.

Teixeira,D., Sheth,U., Valencia-Sanchez,M.A., Brengues,M., and Parker,R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA,* **11,** 371-382.

Tharun,S., He,W., Mayes,A.E., Lennertz,P., Beggs,J.D., and Parker,R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. *Nature,* **404,** 515-518.

Tharun,S., Muhlrad,D., Chowdhury,A., and Parker,R. (2005). Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3' end protection. *Genetics,* **170,** 33-46.

Tharun,S. and Parker,R. (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Mol. Cell,* **8,** 1075-1083.

Thompson,J.D., Higgins,D.G., and Gibson,T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22,** 4673-4680.

Thore,S., Mayer,C., Sauter,C., Weeks,S., and Suck,D. (2003). Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya. *J. Biol. Chem.*, **278,** 1239-1247.

Tkacz,I.D., Lustig,Y., Stern,M.Z., Biton,M., Salmon-Divon,M., Das,A., Bellofatto,V., and Michaeli,S. (2006). Identification of novel snRNA-specific Sm proteins that bind selectively to U2 and U4 snRNAs in Trypanosoma brucei. *RNA.*

Tollervey,D. and Kiss,T. (1997). Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.*, **9,** 337-342.

Tomasevic,N. and Peculis,B.A. (2002). Xenopus LSm proteins bind U8 snoRNA via an internal evolutionarily conserved octamer sequence. *Mol. Cell Biol.*, **22,** 4101-4112.

Törö,I., Thore,S., Mayer,C., Basquin,J., Séraphin,B., and Suck,D. (2001). RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.*, **20,** 2293-2303.

Treich,I., Ho,L., and Carlson,M. (1998). Direct interaction between Rsc6 and Rsc8/Swh3,two proteins that are conserved in SWI/SNF-related complexes. *Nucleic Acids Res.*, **26,** 3739-3745.

Trotta,C.R. and Abelson,J. (1999). tRNA splicing: an RNA World add-on or an ancient reaction? *The RNA World,* . CSHL Press, Cold Spring Harbor, New York.

Tsuchihashi,Z. and Brown,P.O. (1994). DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.*, **68,** 5863-5870.

Tsui,H.C., Leung,H.C., and Winkler,M.E. (1994). Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli K-12*. *Mol. Microbiol.*, **13,** 35-49.

Tucker,M. and Parker,R. (2000). Mechanisms and control of mRNA decapping in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.*, **69,** 571-595.

Uetz,P., Giot,L., Cagney,G., Mansfield,T.A., Judson,R.S., Knight,J.R., Lockshon,D., Narayan,V., Srinivasan,M., Pochart,P., Qureshi-Emili,A., Li,Y., Godwin,B., Conover,D., Kalbfleisch,T., Vijayadamodar,G., Yang,M., Johnston,M., Fields,S., and Rothberg,J.M. (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature,* **403,** 623-627.

Urlaub,H., Raker,V.A., Kostka,S., and Lührmann,R. (2001). Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J.*, **20,** 187-196.

Valadkhan,S. (2005). snRNAs as the catalysts of pre-mRNA splicing. *Curr. Opin. Chem. Biol.*, **9,** 603-608.

Valadkhan,S. and Manley,J.L. (2001). Splicing-related catalysis by protein-free snRNAs. *Nature,* **413,** 701-707.

Valadkhan,S. and Manley,J.L. (2003). Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA,* **9,** 892-904.

Valentin-Hansen,P., Eriksen,M., and Udesen,C. (2004). The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol. Microbiol.*, **51,** 1525-1533.

van Dijk,E., Cougot,N., Meyer,S., Babajko,S., Wahle,E., and Séraphin,B. (2002). Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.*, **21,** 6915-6924.

van Nues,R.W. and Beggs,J.D. (2001). Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of *Saccharomyces cerevisiae*. *Genetics,* **157,** 1451-1467.

Vanacova,S., Wolf,J., Martin,G., Blank,D., Dettwiler,S., Friedlein,A., Langen,H., Keith,G., and Keller,W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS. Biol.*, **3,** e189.

Venema,J. and Tollervey,D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. *Annu. Rev. Genet.*, **33,** 261-311.

Verdone,L., Galardi,S., Page,D., and Beggs,J.D. (2004). Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr. Biol.*, **14,** 1487-1491.

Vidal, V. P. Detection and analysis of proteins interacting with the U6 snRNA in snRNPs and in spliceosomes in the yeast. 1998.
Ref Type: Thesis/Dissertation

Vidal,V.P., Verdone,L., Mayes,A.E., and Beggs,J.D. (1999). Characterization of U6 snRNA-protein interactions. *RNA,* **5,** 1470-1481.

Vidaver,R.M., Fortner,D.M., Loos-Austin,L.S., and Brow,D.A. (1999). Multiple functions of *Saccharomyces cerevisiae* splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics,* **153,** 1205-1218.

Villa,T., Pleiss,J.A., and Guthrie,C. (2002). Spliceosomal snRNAs: Mg(2+)-dependent chemistry at the catalytic core? *Cell,* **109,** 149-152.

Vincze,T., Posfai,J., and Roberts,R.J. (2003). NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res.,* **31,** 3688-3691.

Volland,C., Urban-Grimal,D., Geraud,G., and Haguenauer-Tsapis,R. (1994). Endocytosis and degradation of the yeast uracil permease under adverse conditions. *J. Biol. Chem.,* **269,** 9833-9841.

Wahle,M. and Keller,W. (1997). Capping, methylation, and 3'-end formation of pre-mRNA. *Eukaryotic mRNA processing,* . Oxford University Press, Oxford, pp. 280-309.

Wan,L., Battle,D.J., Yong,J., Gubitz,A.K., Kolb,S.J., Wang,J., and Dreyfuss,G. (2005). The survival of motor neurons protein determines the capacity for snRNP assembly: biochemical deficiency in spinal muscular atrophy. *Mol. Cell Biol.,* **25,** 5543-5551.

Wang,P., Palfi,Z., Preusser,C., Lucke,S., Lane,W.S., Kambach,C., and Bindereif,A. (2006). Sm core variation in spliceosomal small nuclear ribonucleoproteins from Trypanosoma brucei. *EMBO J.,* **25,** 4513-4523.

Watkins,N.J., Lemm,I., Ingelfinger,D., Schneider,C., Hossbach,M., Urlaub,H., and Lührmann,R. (2004). Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. *Mol. Cell,* **16,** 789-798.

Watson,N. (1988). A new revision of the sequence of plasmid pBR322. *Gene,* **70,** 399-403.

Wickner,R.B. and Leibowitz,M.J. (1976). Chromosomal genes essential for replication of a double-stranded RNA plasmid of Saccharomyces cerevisiae: the killer character of yeast. *J. Mol. Biol.,* **105,** 427-443.

Will,C.L. and Lührmann,R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.,* **13,** 290-301.

Wilusz,C.J. and Wilusz,J. (2005). Eukaryotic Lsm proteins: lessons from bacteria. *Nat. Struct. Mol. Biol.,* **12,** 1031-1036.

Wilusz,C.J., Wormington,M., and Peltz,S.W. (2001). The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell Biol.,* **2,** 237-246.

Worthington,M.T., Luo,R.Q., and Pelo,J. (2001). Copacabana method for spreading *E. coli* and yeast colonies. *Biotechniques,* **30,** 738-40, 742.

Wyers,F., Minet,M., Dufour,M.E., Vo,L.T., and Lacroute,F. (2000). Deletion of the PAT1 gene affects translation initiation and suppresses a PAB1 gene deletion in yeast. *Mol. Cell Biol.,* **20,** 3538-3549.

Wyers,F., Rougemaille,M., Badis,G., Rousselle,J.C., Dufour,M.E., Boulay,J., Regnault,B., Devaux,F., Namane,A., Seraphin,B., Libri,D., and Jacquier,A. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell,* **121,** 725-737.

Xiong,L., Gong,Z., Rock,C.D., Subramanian,S., Guo,Y., Xu,W., Galbraith,D., and Zhu,J.K. (2001). Modulation of abscisic acid signal transduction and biosynthesis by an Sm-like protein in Arabidopsis. *Dev. Cell,* **1,** 771-781.

Xue,D., Rubinson,D.A., Pannone,B.K., Yoo,C.J., and Wolin,S.L. (2000). U snRNP assembly in yeast involves the La protein. *EMBO J.,* **19,** 1650-1660.

Yan,Y., Rubinchik,S., Watson,P.M., Kelley,J.R., Fraser,M.M., Wood,A.L., Dong,J.Y., Gillanders,W.E., Boylan,A.M., Watson,D.K., and Cole,D.J. (2005). Establishing a murine pancreatic cancer CaSm model: up-regulation of CaSm is required for the transformed phenotype of murine pancreatic adenocarcinoma. *Mol. Ther.,* **11,** 363-372.

Yan,Y., Rubinchik,S., Wood,A.L., Gillanders,W.E., Dong,J.Y., Watson,D.K., and Cole,D.J. (2006). Bystander effect contributes to the antitumor efficacy of CaSm antisense gene therapy in a preclinical model of advanced pancreatic cancer. *Mol. Ther.,* **13,** 357-365.

Yong,J., Wan,L., and Dreyfuss,G. (2004). Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol.,* **14,** 226-232.

Yoo,C.J. and Wolin,S.L. (1997). The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell,* **89,** 393-402.

Yoon,H. and Donahue,T.F. (1992). Control of translation initiation in Saccharomyces cerevisiae. *Mol. Microbiol.,* **6,** 1413-1419.

Yoshihisa,T., Yunoki-Esaki,K., Ohshima,C., Tanaka,N., and Endo,T. (2003). Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol. Biol. Cell,* **14,** 3266-3279.

Zaric,B., Chami,M., Remigy,H., Engel,A., Ballmer-Hofer,K., Winkler,F.K., and Kambach,C. (2005). Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. *J. Biol. Chem.,* **280,** 16066-16075.

Zhang,A., Wassarman,K.M., Ortega,J., Steven,A.C., and Storz,G. (2002). The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol. Cell,* **9,** 11-22.

Zhang,D., Abovich,N., and Rosbash,M. (2001). A biochemical function for the Sm complex. *Mol. Cell,* **7,** 319-329.

Zhang,Z. and Dietrich,F.S. (2005). Mapping of transcription start sites in Saccharomyces cerevisiae using 5' SAGE. *Nucleic Acids Res.,* **33,** 2838-2851.

Zuker,M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.,* **31,** 3406-3415.