Structural characterisation of *yeast* Lsm protein complexes

Jens Moll

Master of Science (Biotechnology)

A thesis presented to Macquarie University in fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Chemistry & Biomolecular Sciences,

Macquarie University, Sydney NSW

Australia

August 2011

Content

Acknowledgements

A lot of hard work and effort went into the preparation of this thesis and many people guided and supported me along the way and deserve to be thanked. First of all, I thank my supervisor Dr. Bridget Mabbutt for giving me the opportunity to work on this exciting project and for her guidance and support throughout my thesis.

Thanks also goes to Dr. Robert Willows for advice on SPR experiments, Dr. Richard Hutton for ITC training, Dr. Grant Pearce who offered me beamtime for SAXS data collection, and Dr. Paul Curmi, Dr. Louise Brown and Dr. Paul Haynes for helpful discussions and inputs.

Special thanks to Dr. Stephen Harrop for sitting with me through countless hours of data collection at the Australian Synchrotron and for his advice on protein crystallisation experiments and crystallographic data processing.

All current and past members of the protein structure group deserve to be thanked, in particular Meghna for providing an excellent starting point for my own work, and for passing on her technical skills, Bhumika, Chandrika, Fran, Jan and Vani for their support, help, friendship, discussions and for keeping me company in the lab.

I am also grateful for the support I experienced from my family an friends and especially my fiancée Kerstin. Thanks for putting up with me and for making me laugh even in times of stress in the final stages.

Declaration

Where appropriate, work done in collaboration with other groups or individuals has been acknowledged. Outside these contributions, the material in this thesis is entirely my own work and to the best of my knowledge original. No part of this thesis has been submitted for a higher degree to any other university or institution. I consent to this thesis being made available for photocopy and loan.

Jens Moll

Department of Chemistry and Biomolecular Sciences

Macquarie University

Sydney NSW 2109

Australia

August 2010

Abstract

Lsm proteins are a family of RNA chaperones present in all kingdoms of life. Members of this protein family organise into ring-shaped quaternary structures and engage in the processing, sorting and regulation of a variety of RNA species. In archaea and bacteria, homomeric complexes of six or seven proteins are functional, whilst discrete heteromeric complexes of seven distinct Lsm proteins occur in eukaryotes. Eukaryotic Lsm assemblies modulate according to cellular localisation and RNA target, demonstrating that specific functionalities may exist for individual Lsm proteins.

In this study, I utilised Lsm polyproteins to pursue structural and functional studies of mixed Lsm rings, as well as to probe their quaternary dynamics. My work focused on two polyprotein forms, fusions of yeast Lsm[2+3] and Lsm[4+1]. Size exclusion chromatography in conjunction with static light scattering detects the formation of stable tetra- and octameric complexes, suggesting the formation of single and stacked tetrameric rings. Elevated populations of octamers are favoured at low ionic strength, indicating electrostatically-mediated packing of Lsm tetramers. A ring morphology for both tetrameric and octameric assemblies is confirmed by small angle X-ray scattering, estimating toroid dimensions to be 75 Å x 50 Å.

The simplified Lsm polyprotein complexes provide excellent probes of specific Lsm affinities for RNA sequences. Differential affinities of Lsm polyprotein towards U-rich G_5U_{10} and U_{10} oligonucleotides are detected by surface plasmon resonance and isothermal titration calorimetry. The highest affinity for these oligonucleotides are

observed for Lsm[2+3] (K_D= 34 ± 15 nM), possibly due to specific basic residues within the linker used to fuse Lsm2 and Lsm3 domains. Isolated Lsm polyprotein complexes were subjected to crystallographic studies, resulting in regular crystalline forms of Lsm[4+1] in three distinct morphologies. Subsequently, native datasets were collected and processed to 3 Å resolution. Extensive phasing attempts using molecular replacement have been made, but have not so far yielded a solution. This data will serve to solve the first crystal structure of a heteromeric Lsm protein complex at atomic resolution upon collection of a suitable heavy atom dataset, however, despite extensive screening, only weakly diffracting crystals were obtained from L-selenomethionine derivatives of the protein to date.

The results obtained from simplified Lsm complexes aid the understanding of natural Lsm assemblies *in vivo*. It is the dynamic reorganisation of Lsm complexes that likely contributes to the hurdle of obtaining quality diffracting crystals for Lsm complexes. My biophysical studies have, however, confirmed the likelihood of ring-shaped morphology of mixed Lsm rings *in vivo*, as well as differential affinities for RNA by their separate Lsm components.

List of abbreviations

List is based on the abbreviations accepted by JBC

Å	angstrom
A_{280}	absorbance at 280 nm
A_{260}	absorbance at 260 nm
*Af*Sm1	Sm protein from *Archeoglubus fulgidus*
ASU	asymmetric unit
Bicine	2-(bis(2-hydroxyethyl)amino)acetic acid
dn/dc	refractive index increment
D_{max}	maximal particle diameter
Hfq	bacterial Lsm protein
I3C	5-amino-2,4,6-triiodoisophthalic acid
k_a	association rate constant
K_A	equilibrium association constant
k_{av}	size exclusion distribution coefficient
k_d	dissociation rate constant
K_D	equilibrium dissociation constant
LB	Luria-Bertani
LLG	log likelihood gain
Lsm	Sm-like
Lsm[2+3]	Polyprotein consisting of fused Lsm2 and Lsm3
Lsm[4+1]	Truncated polyprotein consisting of fused Lsm4 and Lsm1
Lsm[4+1*ext*]	Polyprotein consisting of fused truncated Lsm4 and full length Lsm1

m_7G cap	5’ methyl guanosine cap
MALLS	multi angle laser light scattering
Mt Lsmα	*M. thermoautotrophicum* Lsmα
NSD	normalized spatial discrepancies
OB-fold	oligosaccharide/oligonucleotide binding fold
OD_{600}	optical density at 600 nm
*Pa*Sm1	Sm protein from *Pyrobaculum aerophilum*
P-body	processing body
PDB	Protein Data Bank
PEG	polyethyleneglycole
pka	acid dissociation constant
p(r)	electron distance distribution function
R_{eq}	surface Plasmon resonance signal at equilibrium
R_g	radius of gyration
RNase P	ribonuclease P
RNP	ribonucleo-protein
rRNA	ribosomal RNA
SAD	single wavelength anomalous dispersion
SAXS	small angle X-ray scattering
SEC	size exclusion chromatography
SeMet	L-selenomethionine
SIRAS	single isomorphous replacement plus anomalous scattering
SMN	Survival of Motor Neurons protein
snoRNA	small nucleolar RNA
snRNP	small nuclear ribonucleoprotein

sRNA	small RNA
Tacsimate	mixture of titrated organic acid salts
TCEP	tris(2-carboxyethyl)phosphine
TRAP	Trp RNA-binding attenuation protein
U	units
V_0	void volume of size exclusion column